mathematics

Article

FinGraphFL: Financial Graph-Based Federated Learning for Enhanced Credit Card Fraud Detection

Zhenyu Xia and Suvash C. Saha *

School of Mechanical and Mechatronic Engineering, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia; 13758750098@163.com
* Correspondence: suvash.saha@uts.edu.au; Tel.: +61-2-9514-3183

Abstract: In the field of credit card fraud detection, traditional methods often struggle due to their reliance on complex manual feature engineering or their inability to adapt to rapidly changing fraud patterns. This paper introduces an innovative approach called FinGraphFL, which merges graph-based learning with the principles of federated learning and improves security through differential privacy. FinGraphFL utilizes Graph Attention Networks to analyze dynamic relationships between daily credit card transaction records, enhancing its ability to detect fraudulent activities. With the addition of differential privacy, the model allows multiple financial institutions to collaborate to refine the detection model without sharing sensitive data, thus improving adaptability and accuracy. The results are tested in two public datasets that show that FinGraphFL achieves accuracy rates of 0.9780 and 0.9839, significantly outperforming traditional methods. Building on these results, FinGraphFL sets the stage for future advances in real-time learning and global financial collaboration, ensuring simultaneous progress in security and privacy protections.

Keywords: privacy preserving; federated learning; local federalization; differential privacy

MSC: 68T07

check for updates

Academic Editor: Raymond Lee

Received: 4 March 2025
Revised: 7 April 2025
Accepted: 23 April 2025
Published: 24 April 2025

Citation: Xia, Z.; Saha, S.C. FinGraphFL: Financial Graph-Based Federated Learning for Enhanced Credit Card Fraud Detection. *Mathematics* **2025**, *13*, 1396. https://doi.org/10.3390/math13091396

1. Introduction

In the contemporary financial environment, credit card fraud poses a significant challenge for both consumers and financial institutions. This illicit activity encompasses a range of fraudulent actions, from the use of unauthorized cards to complex, cyber-enabled schemes. Recent reports from the Federal Trade Commission have indicated a marked increase in credit card fraud incidents, with annual losses ranging in billions of dollars [1]. Furthermore, the globalization of financial markets adds another layer of complexity. In today's interconnected world, cross-border transactions are common, providing fraudsters with additional opportunities and presenting challenges in tracking and prosecuting such activities [2]. Anti-fraud transactions at this international level require coordinated multi-party cooperation and the use of more sophisticated detection systems and prevention methods. Given these challenges, there is an urgent need for innovative and effective strategies to combat credit card fraud. The banking industry is constantly searching for advanced technological solutions to stay ahead of fraudsters, which has led to an increasing interest in leveraging artificial intelligence and machine learning to revolutionize fraud detection systems [3].

In the domain of credit card fraud detection, traditional machine learning methods such as logistic regression, random forests, and k-nearest neighbor (k-NN) algorithms were once widely employed, valued for their simplicity and interpretability [4]. However, these methods often rely on laborious manual feature engineering, making them unable to adapt to today's ever-changing fraud patterns and the endless emergence of new fraud methods. With technological advancements, deep learning models incorporating cutting-edge techniques like Transformers, Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and Unified Message Passing models (UniMP) have begun to gain attention [5]. These models significantly enhance the detection of subtle fraud activities through their capabilities in automatic feature extraction and complex pattern recognition.

Although deep learning has advanced significantly in theory, its practical implementation still faces considerable obstacles. A key issue is the reliance of deep learning techniques on extensive labeled datasets, which poses a particular challenge for small and mid-sized financial firms with limited access to sufficient data (Challenge 1). These institutions often find it difficult to gather diverse, high-quality samples that cover the full spectrum of fraudulent activities, leading to suboptimal model performance and poor generalization in real-world applications. Moreover, the constantly evolving tactics of credit card fraud can quickly make existing detection systems outdated. For resource-constrained organizations unable to perform regular model updates, their datasets may fail to reflect the latest fraudulent patterns, diminishing the reliability of their fraud detection mechanisms (Challenge 2). Consequently, there is a pressing need for adaptive learning frameworks capable of responding to emerging threats. These systems should not only detect diverse fraud strategies but also enable seamless updates without compromising data privacy.

To address the two primary challenges that deep learning faces in credit card fraud detection, we propose a graph-based federated learning framework, named FinGraphFL, that integrates the latest advancements in Graph Neural Networks (GNNs) and differential privacy. This approach is designed to enhance fraud detection capabilities across institutions while safeguarding user privacy. (1) Addressing Challenge 1: Our strategy employs a federated learning architecture that capitalizes on the benefits of collaborative model training. Each institution independently trains customized models using locally sourced data within this framework. This practice significantly reduces reliance on large, labeled datasets and diminishes the computational demands typically associated with deep learning. Furthermore, it bolsters data privacy by keeping sensitive information localized, minimizing the need for centralized data storage. (2) Addressing Challenge 2: In our approach to credit card fraud detection, we incorporate GNNs as local models within the federated learning setup. This configuration enables institutions to update their models collaboratively without the need to exchange sensitive data. GNNs are adept at identifying intricate patterns within transaction networks, enhancing the model's ability to detect fraud while maintaining data privacy through the federated structure. We construct transaction graphs for each institution's data and deploy Graph Attention Networks (GATs) managed by the federated server. To further enhance privacy, we introduce a novel technique that injects Laplacian noise into the gradients after local model training, utilizing graph embeddings from pooled GAT convolutions. This technique ensures differential privacy and permits reversible perturbations. The denoised gradients are then aggregated through a similarity-based attention mechanism tailored to each institution, optimizing the update process and significantly enhancing the system's ability to handle diverse and evolving fraud patterns.

This research aims to bridge the gap between privacy and efficiency in credit card fraud detection systems. By introducing a novel integration of personalized federated learning and graph models, we seek to establish a framework that enhances detection capabilities while strictly adhering to privacy and regulatory standards, catering to the diverse and evolving landscape of financial fraud. Our contributions are manifold and significant:

- Decentralized Training for Enhanced Data Security: FinGraphFL leverages decentralized model training within a federated framework, enabling each financial institution to process and analyze data locally. This setup reduces the dependency on shared, massive datasets, thereby minimizing data exposure and enhancing security against breaches.
- Enhanced Detection with Differential Privacy through Graph Attention Mechanisms: We incorporate graph attention networks (GATs) within our federated learning model to analyze intricate connections within transaction data. This method provides a targeted approach to fraud detection, allowing the model to focus on key relationships and anomalies. Our use of differential privacy techniques further ensures that the privacy of sensitive data is maintained while allowing for accurate and secure model enhancements.
- Practical Applicability: Through experimentation and analysis, FinGraphFL demonstrates effectiveness and practicality in real-world financial security scenarios. This bridges the gap between theoretical research and industry practice, offering banks and financial institutions a viable, efficient tool for credit card fraud detection.

The structure of this paper is outlined as follows. Section 2 surveys prior studies on credit card fraud detection, graph-based machine learning, and federated learning, establishing the background for our work. Section 3 introduces our framework, FinGraphFL, elaborating on its theoretical basis and technical design. Section 4 evaluates the framework through comprehensive experiments and analyzes the findings. Lastly, Section 5 summarizes the contributions and suggests avenues for further research.

2. Related Work

The financial sector faces growing challenges in detecting credit card fraud due to increasingly sophisticated fraudulent activities. While traditional methods struggle to analyze intricate transactional patterns, modern machine learning techniques—particularly deep learning, graph neural networks (GNNs), and federated learning—have demonstrated superior capabilities in addressing these challenges. Beyond enhancing detection accuracy, these approaches enable decentralized data analysis while preserving user privacy, making them well suited for financial applications. This section examines the role of these advanced methodologies in transforming fraud detection systems.

2.1. Traditional Machine Learning and Deep Learning Methods

In the field of credit card fraud detection, traditional machine learning methods have long been the norm. Techniques such as decision trees, support vector machines (SVMs) [6,7], logistic regression [8,9], and ensemble methods like random forests [10,11] and gradient boosting machines [12,13] are extensively employed to identify potential fraudulent activities. These algorithms utilize historical transaction data to detect fraud patterns, offering the advantages of relatively low computational complexity and good interpretability. However, they may encounter limitations when dealing with complex, high-dimensional, non-linear data, which could impact their effectiveness in certain situations.

With the advancement of big data technologies and increased computational power, deep learning methods have demonstrated significant performance improvements in the detection of credit card fraud. In particular, deep neural networks, including convolutional neural networks (CNNs) [14,15] and recurrent neural networks (RNNs) [16,17], are noted for their robust feature extraction capabilities. Long Short-Term Memory Networks (LSTMs) [18] are widely used for analyzing transaction data due to their ability to process time series data. Deep learning approaches are capable of autonomously identifying complex patterns, thus improving detection performance. However, these methods may require more data and computational resources, and their "black-box" nature often results in lower interpretability.

2.2. Graph Neural Networks

GNNs represent a significant breakthrough in deep learning, particularly for their unique ability to process and extract insights from graph-structured data. In financial fraud detection scenarios, where transactions inherently create interconnected networks, GNNs offer distinct advantages in capturing complex relational patterns that traditional methods often miss. Various GNN architectures have proven effective for this task: graph convolutional networks (GCNs) excel at aggregating neighborhood information, graph attention networks (GATs) can learn dynamic importance weights between connected transactions, and Graph Isomorphism Networks (GINs) provide enhanced discriminative power for fraud pattern recognition.

GCNs adapt convolutional operations to graph-structured data, enabling systematic feature extraction through neighborhood aggregation. By propagating and transforming node features across adjacent connections, GCNs effectively identify local anomaly patterns characteristic of fraudulent transactions [19,20]. Building upon this framework, GATs introduce learnable attention coefficients that dynamically quantify the relevance of neighboring nodes. This adaptive weighting mechanism proves particularly valuable in financial fraud scenarios, where the significance of transaction relationships varies substantially [21,22]. GINs concentrate on capturing the structural information of graphs by considering the isomorphism between different graph structures. This method is advantageous for distinguishing between genuine and fraudulent transaction patterns that may appear similar but have subtle structural differences [23,24].

Despite their potent capabilities, GNNs face challenges when applied to large-scale graph data, such as transaction networks. The complexity and size of these graphs require significant computational resources and efficient storage solutions. Moreover, the complexity of GNN models makes their tuning a complicated task, necessitating a deep understanding of the underlying principles of graph theory and the specific characteristics of the credit card fraud detection problem at hand.

2.3. Federated Learning

Federated learning has emerged as a paradigm-shifting framework for decentralized model development, enabling collaborative training across distributed entities while maintaining data localization [25]. This approach addresses critical challenges in financial fraud detection, where privacy regulations (e.g., GDPR, CCPA) and competitive concerns traditionally prevent cross-institutional data sharing. The framework operates through coordinated parameter aggregation rather than raw data exchange, allowing participating banks to retain sensitive transaction records on-premises while still benefiting from collective learning. Such architecture achieves dual objectives: (1) improving detection accuracy through diversified training samples across institutions, and (2) guaranteeing client-level privacy through secure aggregation protocols.

Recent innovations in federated learning have addressed critical limitations including communication overhead and inter-client data distribution disparities. The Federated Averaging (FedAvg) framework [26,27] establishes a foundational approach, where decentralized model training alternates with synchronized parameter averaging, effectively reconciling localized adaptation with global generalization. Subsequent developments have introduced sophisticated optimization methodologies [28] that significantly reduce bandwidth requirements while improving convergence rates through adaptive gradient techniques. From a privacy perspective, cryptographic aggregation schemes [29,30] provide mathematical guarantees for protecting participant contributions during model synchronization, strengthening the fundamental privacy-preserving properties of federated architectures.

Despite these advancements, federated learning still faces issues like model drift, where the model's performance may degrade over time due to the non-IID (independent and identically distributed) nature of data across different nodes. Strategies such as model personalization [31,32] are being explored to mitigate this, with the aim of more effectively tailoring the global model to local data distributions and dynamically adapting to changing data patterns. In this study, we introduce a personalized approach to credit card fraud detection, leveraging a federated learning framework enhanced with an attention mechanism for adaptive weight adjustments. We have also incorporated graph models to enhance the model's performance by capturing complex transaction patterns and correlations, which traditional methods might miss. This approach not only improves credit card fraud detection but also aids in understanding customer behavior patterns while maintaining privacy.

3. Methodology

In this section, we will introduce our proposed method FinGraphFL in detail, including its theoretical basis and corresponding implementation details.

3.1. Overview of FinGraphFL

In this section, we introduce FinGraphFL, a federated learning framework custom-designed for detecting credit card transaction fraud within financial institutions (as shown in Figure 1). It specifically addresses the unique challenges faced by small to medium-sized financial entities. Our approach integrates Graph Attention Networks (GATs) and differential privacy techniques to ensure the privacy of these institutions while boosting the framework's fraud detection capabilities. We utilize transaction similarity graphs along with GATs to effectively capture and analyze complex transactional relationships. Moreover, to accurately represent the distribution of customer data and prevent data leakage, we devise an innovative differential privacy method using graph embeddings, significantly enhancing the model's ability to safeguard user data.

3.2. Integrating GAT for Enhanced Fraud Detection in Federated Learning

Unlike traditional graph convolutional networks (GCN), GAT employs a unique attention mechanism that dynamically assesses the importance of nodes within a transaction graph, thus improving fraud detection by focusing on significant transactional patterns. Implementing GAT within a federated learning setup ensures that each participating entity can train on their dataset, reap the benefits of shared learning, and contribute to a robust, collective fraud detection model. Furthermore, federated learning avoids the need for clients to directly share their local datasets, ensuring compliance with data sharing restrictions and regulations while fully protecting the privacy of bank customers. This approach represents a significant advancement in merging cutting-edge AI techniques with the essential needs for privacy and collaboration in the financial sector.

GAT utilizes the self-attention mechanism to determine the importance of neighboring nodes relative to a current node in a graph, thereby enhancing the ability to effectively capture and utilize complex relational data. In GAT, each node updates its embedding by aggregating the embeddings of neighboring nodes, which are weighted by attention coefficients. These coefficients are determined using a shared attention mechanism, enabling the model to concentrate more on relevant neighbors. Mathematically, this process can be described as:

$$\alpha_{ij} = \text{softmax}\Big(\text{LeakyReLU}\Big(\mathbf{a}^T[\mathbf{W}h_i \,||\, \mathbf{W}h_j]\Big)\Big), \tag{1}$$

where α_{ij} indicates the attention coefficients between nodes i and j, h_i and h_j are the embedding vectors of nodes i and j, $\mathbf{W}$ is a learnable weight matrix, $\mathbf{a}$ is a learnable parameter of the attention mechanism, and $[\cdot||\cdot]$ denotes concatenation. This mechanism prioritizes nodes that significantly contribute to the graph's overall learning objectives in the feature representation, ensuring that their influence is adequately reflected in the network learning process.

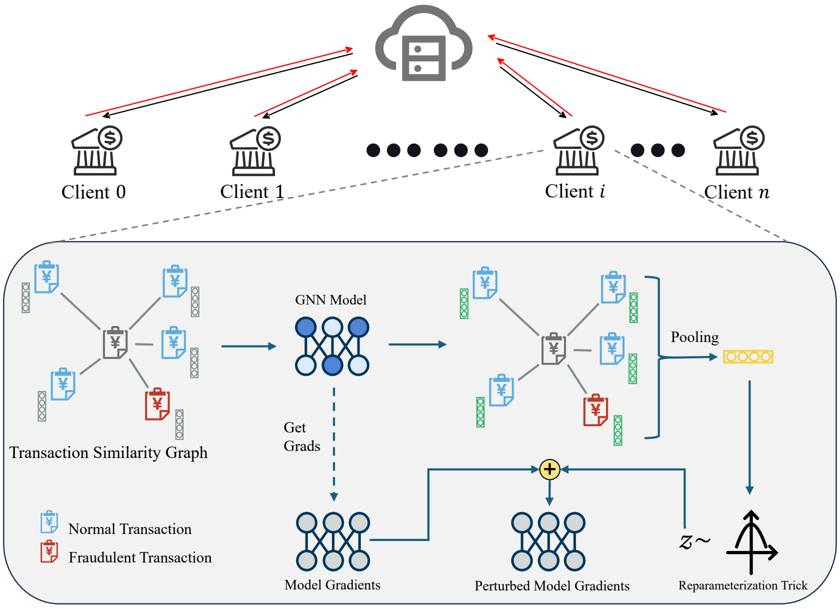

Figure 1. FinGraphFL Framework: This figure illustrates the FinGraphFL framework for federated learning-based credit card fraud detection. The nodes in the transaction similarity graph represent transactions, classified as normal (blue) or fraudulent (red). Each client independently trains a graph neural network (GNN) model on local data and computes gradients, which are then perturbed for privacy using differential privacy techniques before being sent to a central server. The server aggregates these perturbed gradients based on the similarity of data distributions among clients, as determined through pooling techniques. This process enhances model performance while maintaining privacy and security across the network.

After computing the attention weights for all neighboring nodes, GAT utilizes these weights to update the embedding of the current node. The embedding update is performed as follows:

$$h_i^l = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}^{l-1}\mathbf{W}h_j^{l-1}\right), \tag{2}$$

where h_i^l indicates the updated embedding for node i at l-th layer, σ denotes a non-linear activation function (e.g., ReLU), $\mathcal{N}(i)$ represents the neighbors set of node i, and α_{ij}^{l-1} is the attention coefficient between node i and node j at $l-1$-th layer.

GAT's attention mechanism dynamically adjusts the weights of relationships between transaction records, effectively highlighting critical features and interactions within the data. This adaptability significantly improves the performance of the model by prioritizing transactions that are rich in information over those with lesser relevance, thus increasing the accuracy of fraud detection. Moreover, GAT's ability to process inputs of various sizes, such as transactions involving varying numbers of participants or different stages of the transaction process, along with its inherent flexibility, makes it particularly effective for handling the diverse and distributed datasets commonly found in federated learning environments.

3.3. Enhancing Privacy in Federated Learning Through Differential Privacy and Graph Embeddings

In this work, we propose a novel privacy enhancement method that is different from existing approaches. To enhance data privacy within federated learning environments, we leverage graph embeddings generated by GAT. These embeddings encapsulate both the structural properties and transactional characteristics of each client's local data, thereby serving as a compact and expressive representation of dataset semantics. Our proposed framework exploits these graph embeddings to estimate inter-client dataset similarity, which is then used to guide the generation of personalized Laplace noise. This noise is subsequently applied to perturb and recover each client's local gradient updates.

The framework implements an adaptive privacy preservation strategy that dynamically calibrates noise injection based on inter-client data distribution similarities. This personalized differential privacy mechanism enables more precise gradient estimation among clients with comparable data patterns while maintaining rigorous privacy guarantees for dissimilar participants. The approach optimizes the privacy–utility trade-off by employing a novel sensitivity metric derived from graph embedding characteristics:

$$\mathcal{S} = (\mathrm{std}(H - h_G) + (\max(h_G) - \min(h_G)))/2, \tag{3}$$

$$\mathcal{N}_{std} = (\mathcal{S}/\epsilon), \tag{4}$$

where H denotes the node-level embeddings produced by GAT, and h_G represents the global graph embedding obtained via pooling operations. The sensitivity metric $\mathcal{S}$ incorporates two key components: (1) the standard deviation term $\mathcal{N}_{std}$ measures the dispersion of node embeddings relative to the global representation, capturing client-specific data patterns; and (2) the range term $\max(h_G) - \min(h_G)$ reflects the overall variation in graph-level features across the federated system. The noise standard deviation $\mathcal{N}_{std}$ is then derived by scaling $\mathcal{S}$ according to the privacy budget ϵ, which governs the strength of privacy protection—lower ϵ values enforce stricter guarantees. This adaptive mechanism ensures the injected noise sufficiently preserves privacy while minimizing unnecessary degradation of model utility.

After each client completes local model training, we apply the proposed personalized differential privacy mechanism on each client to perturb the original gradient obtained from local training. This process ensures that the aggregated data used for model updates does not compromise the privacy of individual data points.

$$\mathcal{G}_{perturb} = \mathcal{G} + \mathcal{L}(0, \mathcal{N}_{std}), \tag{5}$$

where $\mathcal{G}_{perturb}$ represents the perturbed gradient, $\mathcal{G}$ is the original gradient, and $\mathcal{L}(0, \mathcal{N}_{std})$ is the Laplacian noise added to the gradient. This noise, with a mean of zero and a standard deviation of $\mathcal{N}_{std}$, is strategically employed to ensure privacy.

Following the privacy-preserving transformation, the federated server computes pairwise client similarities using their respective graph embeddings. These similarity metrics, combined with the noise-perturbed gradients, are then distributed to all participating clients. Each client subsequently employs a denoising procedure to reconstruct neighboring clients' gradients using the following estimator:

$$\mathcal{G}_{rec} = \mathcal{G}_{perturb} - \mathcal{L}(0, \mathcal{N}_{std}), \tag{6}$$

where $\mathcal{G}_{rec}$ denotes the reconstructed gradient vector after noise reduction, $\mathcal{G}_{perturb}$ represents the noise-corrupted gradient received from the server, and $\mathcal{L}(0, \mathcal{N}_{std})$ indicates the Laplace-distributed random noise sampled from the same distribution used in the initial perturbation phase. This compensation mechanism allows clients to approximate the original gradients while maintaining differential privacy guarantees.

Finally, the model aggregation process employs an adaptive weighting scheme based on two factors: inter-client graph embedding similarity and gradient perturbation characteristics. We quantify embedding similarity using normalized dot products, assigning greater influence to clients with more compatible data distributions. The final aggregated gradient is computed as follows:

$$\mathcal{G}_{agg} = \sum_{i=1}^{N} w_i \cdot (\mathcal{G}_{rec})_i, \tag{7}$$

where the weighting coefficient w_i for the i-th client is dynamically determined by both its similarity to other participants and the reliability of its gradient reconstruction. This dual-factor approach ensures optimal balance between collaborative learning efficiency and individual data privacy protection.

4. Experiment

In this section, we conduct comprehensive experiments to verify the performance of our proposed FinGraphFL, including comparative experiments to verify its fraud transaction detection performance and ablation studies to test the contribution of each of its components to the final results.

4.1. Experimental Settings

We first introduce our experimental settings in this work in detail.

4.1.1. Benchmark Models

In this experiment, we evaluate and compare several credit card fraud detection benchmark models commonly used in the financial industry. These benchmark models reflect the methods widely adopted within the industry, and by comparing their performance with that of FinGraphFL, we seek to explore the effectiveness of the FinGraphFL model in the realm of credit card fraud detection. The specific benchmark models include the following:

1. Logistic Regression: A classic statistical method used for binary classification problems, which determines whether a transaction is fraudulent by estimating probabilities, suitable for linearly separable datasets.
2. KNNs (K-Nearest Neighbors): An instance-based learning method that determines the category of a test data point by finding its K nearest neighbors and basing the decision on the neighbors' categories; it is suitable for datasets where data points exhibit clear similarities.
3. Histogram-Based Gradient Boosting Classifier (HGBC): An ensemble learning method based on decision trees that improves performance by constructing multiple decision trees and integrating their predictions, particularly suitable for large-scale data.
4. Support Vector Machine (SVM): SVM distinguishes between different categories by finding the optimal separating hyperplane in the dataset, performing well in high-dimensional spaces, especially suitable for datasets where the number of features exceeds the number of samples.
5. Random Forest Classifier: An ensemble learning technique that constructs multiple decision trees and uses their average or majority voting for the final prediction, aimed at reducing overfitting and enhancing accuracy.

6. AdaBoost Classifier: An adaptive boosting technique that combines multiple weak classifiers to form a strong classifier, with each weak classifier given higher weight on the data where the previous classifier made errors.
7. Multi-Layer Perceptron Classifier (MLP): A basic feedforward neural network with at least one hidden layer, capable of learning non-linear patterns in data, suitable for complex classification problems.

We also compared our proposed FinGraphFL with existing SOTA federated learning methods to conduct an in-depth study of the performance differences of our proposed federated learning method that combines GAT and differential privacy mechanisms. The specific federated learning methods compared are as follows:

8. FedProx: FedProx is a federated learning method designed to handle heterogeneity in device hardware and data distribution by modifying the traditional federated averaging algorithm to include a proximal term, which helps stabilize training across unevenly distributed and partial datasets.
9. Personalised FL: Personalized Federated Learning (Personalised FL) is a method that tailors the federated learning process to individual users or devices by allowing local models to deviate from the global model, which helps optimize model performance based on user-specific data characteristics.
10. FedAMP: FedAMP is a federated learning method that enhances model convergence and performance by applying an adaptive mixing parameter to aggregate updates more effectively across clients with non-IID data distributions.
11. FedFomo: FedFomo is a federated learning method that aims to improve model performance by prioritizing the aggregation of client updates based on a measure of regret, comparing potential model updates with actual ones to optimize learning outcomes.
12. APFL: APFL (Adaptive Personalized Federated Learning) is a federated learning strategy that improves personalization by adapting models to individual clients using a combination of local and global updates, thereby enhancing overall performance with personalized tuning.
13. PFedMe: PFedMe is a method in personalized federated learning that optimizes a personalized model for each client by employing a Moreau envelope-based regularization technique, ensuring better convergence properties and personalization effectiveness across heterogeneous data distributions.
14. APPLE: APPLE (Agnostic Personalized Private Learning) is a federated learning method designed to enhance user privacy and model personalization by integrating differential privacy and learning personalized models that can be adapted to different data distributions of individual clients.
15. ATT: The federated learning method based on the attention mechanism implements a personalized federated learning update process by calculating the similarity of gradients between customers.

4.1.2. Training Device and Parameter Configuration

In this study, we utilized PyTorch version 2.2.0 as the experimental platform, operating under Python 3.12. Our simulation experiments were conducted on a computer equipped with an Intel i5 processors (Intel Corporation, Santa Clara, CA, USA) at 3.7 GHz, 64.0 GB of installed RAM, and an NVIDIA RTX 4070 GPU (NVIDIA Corporation, Santa Clara, CA, USA) with 12.0 GB of RAM.

Table 1 presents the optimized hyperparameter configuration that yields peak performance for FinGraphFL. Through extensive empirical validation, we established the following training protocol: The federated learning framework operates for 300 communication rounds, with each client performing 10 local training epochs per round. We employ the Adam optimizer with a base learning rate of 0.005, augmented by L2 regularization (weight decay = 1×10^{-4}) to prevent model overfitting. For the objective function, binary cross-entropy is adopted due to its effectiveness in fraud detection tasks, providing robust gradient signals for model updates.

Table 1. Parameter configuration.

Parameters	**Parameter Values**
communication rounds	300
local epochs	10
dropout	0.3
batch size	128
learning-rate	0.005
optimizer	Adam
weight decay	1×10^{-6}
loss function	BinaryCrossEntropyLoss

4.2. Dataset and Financial Institution Client Configuration

1. **2018 4th 'HaoDai Cup' China Risk Management Control and Capability Challenge Dataset (2018CN):** This dataset contains transaction records of credit card holders from September 2013. It covers transactions over two days, totaling 284,302 transactions, of which 483 were marked as fraudulent. The dataset is highly imbalanced, with fraud transactions accounting for only 0.172% of the total. For privacy reasons, all features except "Time" and "Amount" have been transformed into numerical results through Principal Component Analysis (PCA), with original features and background information undisclosed. V1 to V28 are the principal components obtained from PCA. The "Time" feature records the seconds elapsed from the first transaction in the dataset, and the "Amount" feature represents the transaction amount, useful for cost-sensitive learning. The response variable "Class" indicates whether each transaction is fraudulent, with 1 for fraud and 0 for non-fraud.
2. **European Cardholders' Credit Card Transaction Records in 2023 (2023EU):** This dataset compiles credit card transactions of European cardholders in 2023, containing over 550,000 anonymized records to ensure the security of cardholder identities. Each record consists of a unique transaction identifier (id) and 28 anonymous features (V1-V28), which may include information about the transaction's time, location, etc. The dataset also includes the transaction amount ("Amount") and a binary label ("Class") indicating whether each transaction is fraudulent (1 for fraud, 0 for non-fraud). Notably, to ensure accuracy for research and model development, this dataset has been specially processed to balance the number of fraudulent and non-fraudulent transactions. Such a balanced dataset helps algorithms learn and identify fraud patterns more effectively without bias from overrepresentation of any category.

The class distribution comparison of 2018CN and 2023EU datasets is shown in Figure 2. The dataset for our study is methodically segmented into distinct subsets, each tailored to correspond with a predetermined number of financial institution clients. We allocate each subset exclusively to a specific financial institution, ensuring there is no data overlap among clients. This precise segmentation is crucial for minimizing the risk of information leakage and provides each financial institution with a secure and isolated environment conducive to effective data processing and model training. To protect and maintain data integrity, we employ advanced randomization techniques during the dataset subdivision process. This approach is vital to ensure that each subset allocated to a financial institution is not only unique, but also represents a true microcosm of the larger dataset. This method enhances the security measures in place by preventing predictable data handling patterns, thereby safeguarding sensitive information. In addition, we put a strong emphasis on maintaining a uniform distribution of data across all subsets to ensure consistency in model performance across the federated learning network. By ensuring that each subset accurately mirrors the overall characteristics of the dataset, we maintain the integrity and diversity of the data. This strategic distribution is critical for enabling the models developed by various financial institutions to generalize effectively, which in turn leads to more robust and reliable fraud detection outcomes. This comprehensive approach ensures that our federated learning framework can deliver high-performance results while adhering to stringent security and privacy standards.

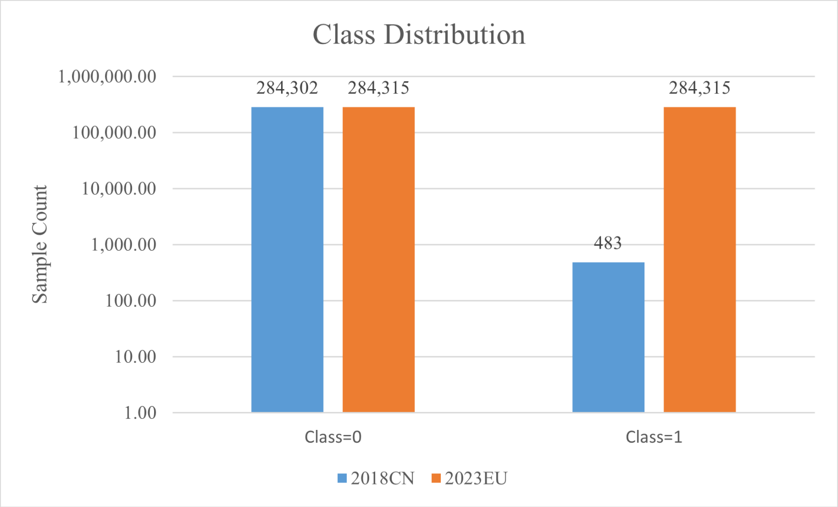

Figure 2. Class distribution comparison in 2018CN and 2023EU datasets.

4.3. Comparison with SOTA Benchmark Models

In the traditional model category (as shown in Table 2), there are notably high accuracy scores, with MLP at 0.9991, KNN at 0.9993, and SVM at 0.9990. These figures reflect the models' performance when trained on extensive datasets amalgamated from multiple banking institutions. It is crucial to underscore, however, that achieving such a comprehensive level of data pooling is not just challenging but typically impossible in actual practice. The practical deployment of these models is severely restricted by stringent privacy laws and regulatory frameworks that prohibit the sharing of sensitive financial data between different entities.

Shifting the focus to federated learning models, FinGraphFL stands out with its notable strengths. It achieves an accuracy of 0.9780 in the 2018CN dataset and maintains impressive performance with an accuracy of 0.9839 in the 2023EU dataset. These figures underscore FinGraphFL's capability to adeptly handle diverse data distributions, a critical advantage in the evolving arena of financial fraud detection.

In the realm of federated learning models, FedProx and PFedMe demonstrated very high accuracies in the 2018CN dataset, both achieving 0.9994, but their performance in the 2023EU dataset dropped significantly to 0.5008 and 0.5006, respectively. This indicates that while these models excel at handling older datasets, they struggle with newer, potentially more complex datasets. Furthermore, FinGraphFL showed slightly lower performance in the 2018CN dataset with an accuracy of 0.9780, but improved significantly in the 2023EU dataset with an accuracy of 0.9839. This reflects FinGraphFL's robust adaptability to continuously changing and varied data distributions, particularly vital in the rapidly evolving field of financial fraud detection. The adoption of differential privacy techniques by FinGraphFL also provides it with an advantage in addressing significant data privacy issues, a challenge that traditional models often fail to overcome.

Table 2. Comparison with state-of-the-art benchmark models: This table presents accuracy comparisons for various models, including FinGraphFL, against SOTA benchmark models on 2018CN and 2023EU. The settings for the FinGraphFL model include a differential privacy mechanism with epsilon = 0.005, graph attention networks (GATs) with a degree of 8 in 2018CN and 5 in 2023EU, and 2 attention heads.

	2018CN Accuracy	**2023EU Accuracy**
Logistic Regression	0.9988	0.9648
KNN	0.9993	0.9899
HGBC	0.9989	0.9810
SVM	0.9990	0.9615
Random Forest	0.9988	0.9317
AdaBoost Classifier	0.9989	0.9568
MLP	0.9991	0.9977
FedProx	0.9994	0.5008
Personalised FL	0.9979	0.9392
FedAMP	0.9980	0.5004
FedFomo	0.9926	0.9265
APFL	0.9989	0.9031
PFedMe	0.9994	0.5006
APPLE	0.9988	0.9490
ATT	0.9886	0.9888
FinGraphFL	0.9780	0.9839

In conclusion, while FinGraphFL might slightly lag behind some models in terms of raw accuracy, it compensates for this with its robust approach to data privacy. This makes FinGraphFL particularly well suited for financial institutions seeking to manage the intricacies of fraud detection under strict privacy regulations, thereby underscoring its potential as a valuable tool for both current and future scenarios in financial fraud prevention.

Although the aforementioned models exhibit commendable accuracy, their performance in handling imbalanced datasets is less than optimal when assessed using the ROC-AUC curve. This is evident from their results on the highly imbalanced 2018CN dataset, as shown in Figure 3. The ROC-AUC curve is an essential metric for evaluating a model's performance on imbalanced datasets because it assesses both the sensitivity and the false positive rate. In contrast, FinGraphFL not only maintains high accuracy, but also excels with its ROC-AUC score. It achieves an area under the ROC curve of 0.9670, demonstrating a robust ability to effectively manage imbalanced datasets. The reality of credit card fraud detection involves inherently imbalanced datasets where fraudulent transactions are much less frequent than legitimate ones. The ability of FinGraphFL to maintain high accuracy and an impressive ROC-AUC score under such conditions is particularly significant, highlighting its effectiveness in real-world scenarios of credit card fraud detection. This highlights its potential and importance as a practical tool in the field, adept at navigating the complexities of imbalanced data.

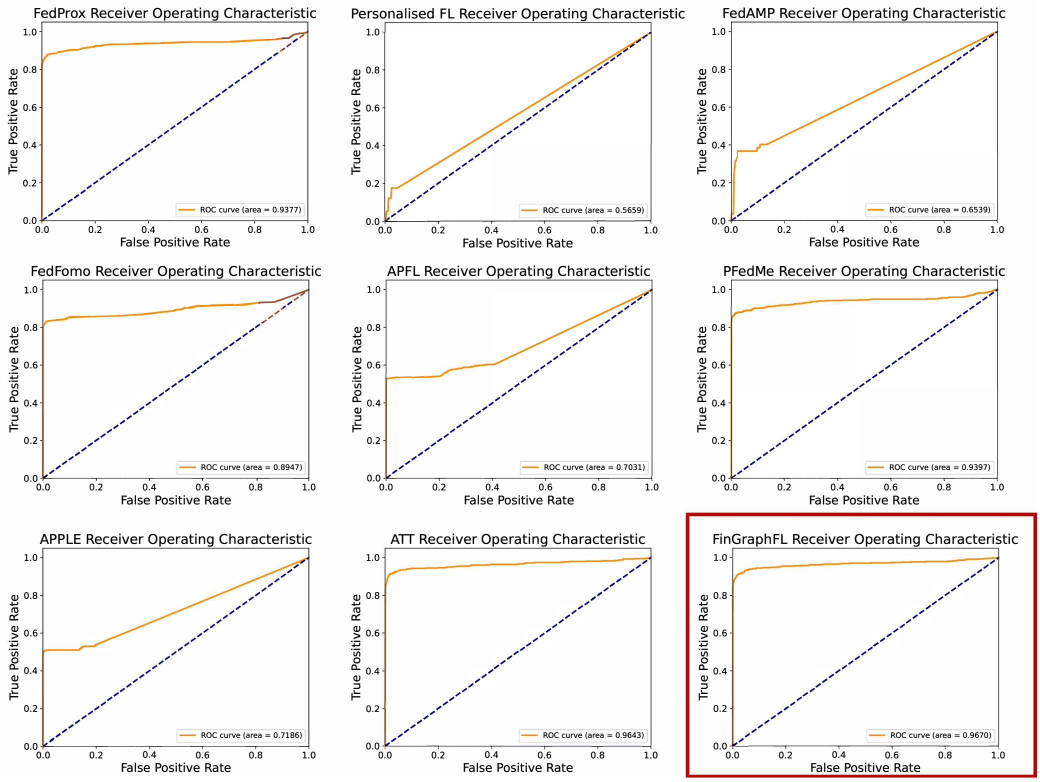

Figure 3. Receiver Operating Characteristic (ROC) curves for various credit card fraud detection models on the imbalanced 2018CN dataset.

4.4. Ablation Study

To further investigate the performance and characteristics of FinGraphFL, targeted ablation experiments were designed and conducted. The ablation study conducted in this research is based on the 2018CN dataset. This dataset was chosen due to its extreme class imbalance, which closely resembles real-world scenarios in credit card fraud detection. Class imbalance refers to a situation where one class (in this case, fraudulent transactions) is significantly underrepresented compared to the other class (non-fraudulent transactions). By using a dataset with such extreme class imbalance, the study aims to evaluate the robustness and effectiveness of the proposed method under conditions that more closely mirror the challenges faced in real-world fraud detection scenarios. This choice enhances the relevance and applicability of the findings to practical applications in the field of credit card fraud detection.

4.4.1. Ablation Study on Different Epsilon

In evaluating the differential privacy mechanism within the FinGraphFL model, various epsilon values were tested to assess their impact on model performance. The results, as shown in Figure 4, detail the accuracy of the test set for different levels of epsilon.

Contrary to conventional privacy–accuracy trade-off assumptions, our experiments reveal a non-monotonic relationship between privacy budget (*eps*) and model performance. At $eps = 0.005$, the model achieves near-optimal accuracy (0.978), suggesting that the introduced noise may serve as an implicit regularizer that prevents overfitting, which is particularly beneficial for complex financial datasets. This phenomenon indicates enhanced generalization capability, where the model learns robust patterns without memorizing training-specific artifacts.

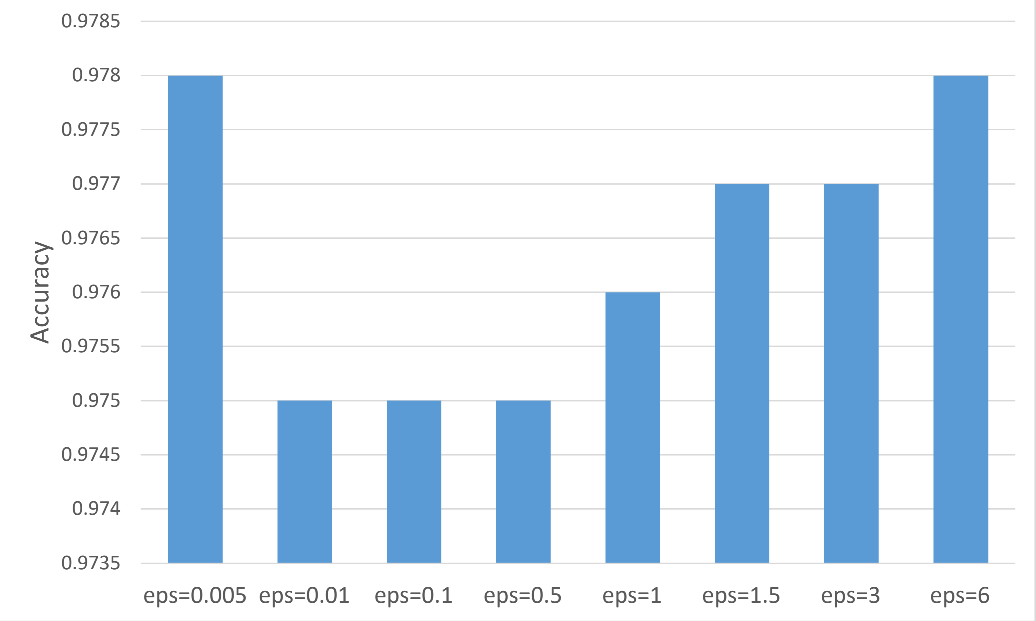

Figure 4. Test set accuracy of different epsilon.

Remarkably, the framework maintains consistently high accuracy (0.975) across moderate privacy budgets (*eps* = 0.01–0.5), demonstrating the resilience of our adaptive feature extraction mechanism. The model's architecture appears to effectively distill discriminative patterns even under significant noise injection, a critical advantage for privacy-sensitive financial applications.

As privacy constraints relax (*eps* = 1–6), we observe the expected gradual accuracy improvement, plateauing at 0.978 for *eps* = 6. This progression confirms our framework's capacity to leverage richer feature representations when permitted by privacy requirements. The results underscore a fundamental design principle: optimal privacy-preserving learning requires careful calibration between data utility and protection strength, rather than simply maximizing either dimension.

Figure 5 presents the convergence dynamics across different eps values, revealing two key insights: (1) all configurations achieve stable convergence within the training budget, and (2) stronger privacy protection (lower epsilon) only marginally extends the required training epochs while maintaining final performance. These observations validate the practical viability of our approach across varying privacy requirements.

4.4.2. Ablation Study on Different Number of Clients

Figure 6 illustrates the impact of varying numbers of clients on the test set accuracy of the FinGraphFL model. At four clients, the model achieves an accuracy of approximately 0.9620. As the number of clients increases to eight, accuracy improves to around 0.9718, indicating that the involvement of more clients allows the model to leverage a richer data resource, thus enhancing learning outcomes.

Our experiments reveal a non-linear relationship between client participation and model performance. The optimal configuration emerges with 10 clients, achieving peak accuracy of 0.9780. This sweet spot appears to balance sufficient data diversity against the computational complexity of coordination, maximizing the collective learning potential. However, expanding to 12 clients results in a significant performance drop (accuracy = 0.9455), likely due to crossing a critical threshold in per-client data volume that impedes effective local model training. Interestingly, while further increasing to 16 clients yields partial recovery (accuracy = 0.9647), it fails to match the 10-client benchmark. This non-linear pattern suggests the existence of the following: (1) a minimum data requirement per client for meaningful local learning, and (2) diminishing returns from excessive federation that outweigh the benefits of added diversity.

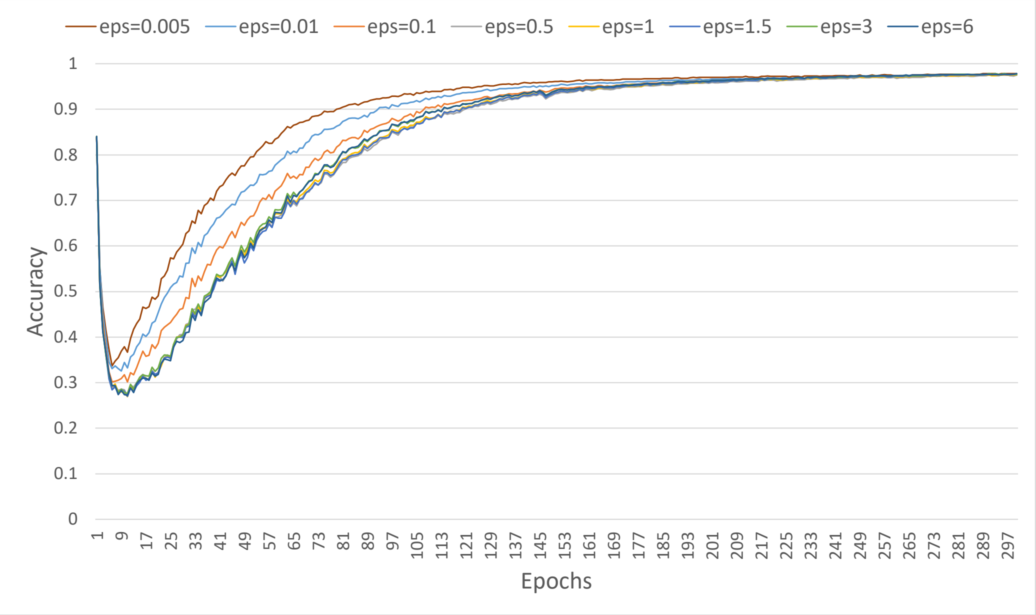

Figure 5. Accuracy trajectories across different epsilon values.

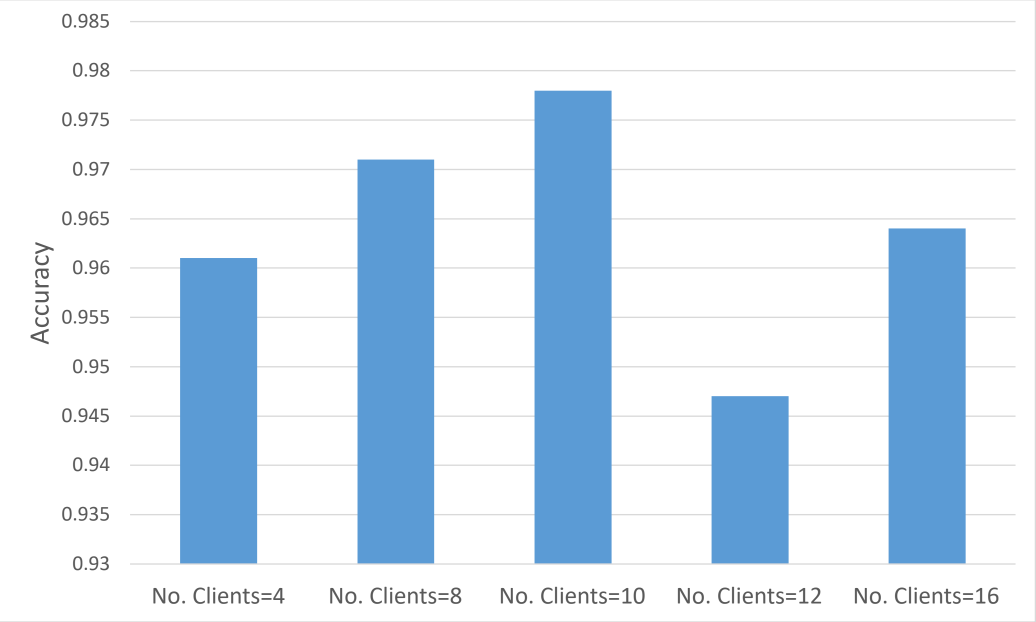

Figure 6. Test set accuracy of different number of clients.

Figure 7 demonstrates several key convergence properties: First, all configurations begin with comparable initial accuracy, confirming consistent initialization. The four-client and eight-client models show rapid convergence, benefiting from richer local datasets. Notably, the 10-client configuration combines the fastest convergence with highest final accuracy, representing the ideal operating point. While 12-client and 16-client setups require more epochs to stabilize, they ultimately achieve comparable performance levels, demonstrating the framework's resilience to federation scale.

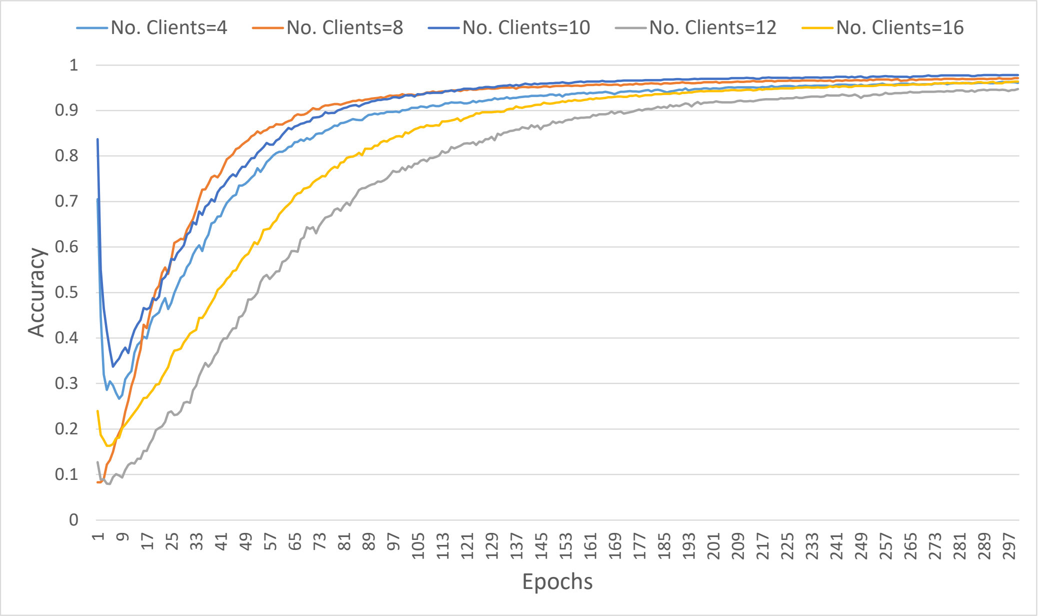

Figure 7. Accuracy trajectories across different number of clients.

These findings highlight three important characteristics of FinGraphFL: (1) robustness to federation size variations, (2) graceful degradation rather than catastrophic failure when crossing data volume thresholds, and (3) consistent eventual convergence regardless of client count. The results suggest that while client number selection impacts training efficiency, it does not fundamentally limit the framework's effectiveness—an important property for real-world deployments where participant numbers may fluctuate.

4.4.3. Ablation Study on Different Number of Attention Heads

Table 3 shows the performance of two local models, GAT (Graph Attention Network) and UniMP (Unified Message Passing), under different numbers of attention heads, using accuracy and ROC-AUC as evaluation metrics.

Table 3. Ablation study for different number of attention heads.

	No. Heads = 1		No. Heads = 2		No. Heads = 4		No. Heads = 8		No. Heads = 12		No. Heads = 16	
	Accuracy	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**
GAT	0.9756	0.9714	0.9780	0.9670	0.9641	0.9597	0.9714	0.9604	0.9615	0.9707	0.9152	0.9715
UniMP	0.9789	0.9657	0.9748	0.9588	0.9357	0.9744	0.9387	0.9740	0.9323	0.9699	0.8349	0.9753

GAT generally exhibits optimal performance with a single attention head, achieving an accuracy of 0.9756 and an ROC-AUC of 0.9714. However, with an increasing number of heads, both metrics tend to decline for GAT, particularly when the number of heads reaches 16, where accuracy drastically decreases to 0.9152, and ROC-AUC slightly decreases to 0.9715. Given the extreme imbalance of the 2018CN dataset, this significantly impacts model performance, especially under multi-head attention configurations. The dataset's imbalance often causes models to overfit the majority class and underperform in recognizing minority classes, severely affecting overall model performance. Increasing the number of attention heads in networks like GAT and UniMP helps to capture more complex features, but it can also exacerbate model biases due to dataset imbalances, leading to insufficient focus on key features of minority classes. Similarly, the UniMP model also shows strong performance with a single head, achieving an accuracy of 0.9789 and an ROC-AUC of 0.9657. However, as more heads are added, its performance gradually worsens, especially at 16 heads, where accuracy significantly drops to 0.8349.

For effective credit card fraud detection, an optimal model must demonstrate both strong overall classification performance and particular sensitivity to minority-class fraudulent transactions. The GAT architecture is especially well adapted for this application due to its localized attention mechanism, which excels at identifying intricate transactional relationships and subtle indicators of fraudulent activity. Experimental results indicate that GAT delivers superior performance when implemented with limited attention heads (typically one or few), achieving an optimal balance between capturing local transaction patterns and maintaining model simplicity—thereby avoiding the potential overfitting issues associated with multiple attention heads.

4.4.4. Ablation Study for Different Mean Node Degree of Transaction Similarity Graph

Table 4 reveals the performance of the local GAT model within the FinGraphFL framework for different mean node degrees in the context of credit card fraud detection. The data clearly show marked differences in model performance across various node degree configurations, offering significant insight for designing effective fraud detection systems.

Table 4. Ablation study for different mean node degree of transaction similarity graph.

Degree = 2		Degree = 4		Degree = 8		Degree = 10	
Accuracy	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**	**Accuracy**	**ROC-AUC**
0.9423	0.9425	0.9667	0.9369	0.9780	0.9670	0.9671	0.9350

With a node degree of two, the GAT model achieves an accuracy of 0.9423 and an ROC-AUC of 0.9425, indicating that while the model performs adequately at lower connectivity, its ability to differentiate between fraudulent and legitimate transactions is limited. As the node degree increases to four, the accuracy significantly improves to 0.9667, although the ROC-AUC slightly decreases to 0.9369. This suggests that while the model has improved overall recognition accuracy, its sensitivity to the minority class (i.e., fraudulent transactions) may have diminished. Increasing the degree of the node further to eight, the GAT model reaches its maximum performance with an accuracy of 0.9780 and a ROC-AUC of 0.9670. This indicates that at a moderate degree of nodes, the model balances overall accuracy with the ability to recognize fraudulent transactions effectively, which is crucial for the detection of credit card fraud. However, increasing the degree of the node to 10 maintains a high accuracy level of 0.9671 but results in a reduced ROC-AUC of 0.9350, suggesting that excessive connectivity may lead to overfitting or other performance degradation issues in complex data environments.

From these observations, it can be concluded that a moderate node degree (such as degree 8) may be optimal for credit card fraud detection tasks, as it not only provides high accuracy but also maintains a robust ROC-AUC value, effectively identifying fraudulent transactions. This configuration helps the model capture sufficient complexity in transaction patterns while avoiding the problems of information redundancy and noise that can arise with excessively high node degrees. Therefore, in practical applications, selecting an appropriate node degree is a key factor in enhancing the performance and practicality of graph-based credit card fraud detection models.

5. Conclusions and Future Work

The FinGraphFL framework represents a significant advancement in credit card fraud detection by seamlessly integrating federated learning with graph attention networks. This design enables the framework to effectively model complex transactional patterns across multiple financial institutions while maintaining strict data privacy guarantees. By overcoming key limitations of conventional fraud detection systems, Fin-

GraphFL enhances both adaptability and detection accuracy through advanced graph-based representation learning.

A central innovation of FinGraphFL lies in its personalized differential privacy mechanism, which not only safeguards sensitive user and institutional data but also empowers small and medium-sized financial institutions to improve the effectiveness of their fraud monitoring systems. By leveraging inter-client similarity, this mechanism enables clients with semantically similar datasets to better reconstruct each other's gradient updates, thereby achieving stronger model utility while preserving privacy. Such a design is particularly critical in addressing the dynamic, imbalanced, and adversarial nature of modern financial fraud. Extensive experiments on public datasets demonstrate that FinGraphFL consistently outperforms conventional baselines, validating its potential as a practical and privacy-preserving solution for collaborative fraud detection in decentralized environments.

Looking ahead, the continued development of FinGraphFL offers several promising directions for both theoretical enhancement and real-world deployment. First, future work may explore extending the framework to support real-time fraud detection, enabling institutions to detect and respond to malicious behavior with minimal latency. Second, incorporating more expressive neural architectures—potentially beyond graph attention networks—could enhance the framework's ability to capture complex and subtle fraud patterns that vary across institutions and evolve over time.

Third, an important area of improvement lies in refining the similarity computation used in the personalized privacy mechanism. Although the current dot product-based approach is computationally efficient, it may not fully capture the nuanced relationships across heterogeneous datasets. Designing more expressive or learnable similarity functions may help further reduce the privacy-utility trade-off, especially in highly diverse or non-IID settings.

Furthermore, we acknowledge the practical challenges associated with deploying FinGraphFL in real-world financial environments. These include substantial variations in data distributions, discrepancies in feature schemas across institutions, evolving fraud behaviors, and diverse IT infrastructure capabilities. Additionally, financial institutions may be subject to different regulatory standards and compliance requirements, making unified system design more complex. In practice, FinGraphFL can be deployed with secure aggregation protocols and privacy-preserving communication channels, and configured to comply with the most stringent applicable regulations across participants. However, designing a federated framework that dynamically adapts to heterogeneous infrastructure and legal contexts remains a critical direction for future research. Addressing these challenges—through system-level optimization, compliance-aware aggregation mechanisms, and robust deployment protocols—will be essential for scaling FinGraphFL into production-grade, cross-institutional fraud detection systems.

Finally, we emphasize the importance of considering the broader ethical and societal implications of deploying fraud detection models like FinGraphFL. While our work is conducted entirely on public and de-identified datasets, real-world deployment may introduce challenges such as algorithmic bias, false positives in transaction blocking, or disproportionate impact on underrepresented or vulnerable user groups. In particular, although our similarity-based gradient aggregation is designed to enhance personalization and utility, there is a potential risk that it could reinforce existing data imbalances if minority patterns are underrepresented in client datasets. Such unintended effects may result in lower fraud detection accuracy for certain population segments or demographic groups.

To address these risks, future work should incorporate fairness-aware modeling and evaluation practices. These may include bias auditing during federated training, performance disaggregation across client subgroups, and the use of fairness constraints or regularizers in local model objectives. Additionally, stakeholder engagement and regulatory oversight will be essential to ensure that deployed systems operate transparently and responsibly in sensitive financial contexts.

In addition to fairness, we also recognize the importance of enhancing the explainability of FinGraphFL's predictions. While the use of graph attention networks (GATs) inherently provides some level of interpretability through learned attention weights, more systematic approaches—such as visualizing attention distributions or integrating explanation modules like GNNExplainer—could be employed to better support transparency and accountability. These tools may help financial institutions understand the rationale behind specific fraud predictions, facilitating model auditing and improving stakeholder trust. We consider the integration of explainability mechanisms an important direction for future research and deployment.

We view the combination of privacy protection, fairness guarantees, and interpretability as essential for the responsible adoption of graph-based federated learning systems. We encourage continued interdisciplinary collaboration to ensure that FinGraphFL and similar frameworks contribute to equitable, transparent, and trustworthy financial AI solutions.

Author Contributions: Conceptualization, Z.X.; Methodology, Z.X.; Software, Z.X.; Validation, Z.X.; Formal analysis, Z.X.; Investigation, Z.X.; Data curation, Z.X.; Writing—original draft, Z.X.; Writing—review & editing, S.C.S.; Visualization, Z.X.; Supervision, S.C.S.; Project administration, S.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors gratefully acknowledge the computational resources provided by the University of Technology Sydney. Furthermore, AI-assisted technology (ChatGPT 4) was employed for language editing and grammatical refinement.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Commision, F.T. *Consumer Sentinel Network Data Book 2021*; Federal Trade Commission: Washington, DC, USA, 2022.
2. Gundur, R.; Levi, M.; Topalli, V.; Ouellet, M.; Stolyarova, M.; Chang, L.Y.C.; Mejía, D.D. Evaluating Criminal Transactional Methods in Cyberspace as Understood in an International Context. 2021. Available online : https://www.crimrxiv.com/pub/48bmtkg0/release/3 (accessed on 28 June 2024).
3. Mahalakshmi, V.; Kulkarni, N.; Kumar, K.P.; Kumar, K.S.; Sree, D.N.; Durga, S. The role of implementing artificial intelligence and machine learning technologies in the financial services industry for creating competitive intelligence. *Mater. Today Proc.* **2022**, *56*, 2252–2255. [CrossRef]
4. Awoyemi, J.O.; Adetunmbi, A.O.; Oluwadare, S.A. Credit card fraud detection using machine learning techniques: A comparative analysis. In Proceedings of the 2017 International Conference on Computing Networking and Informatics (ICCNI), Lagos, Nigeria, 29–31 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–9.
5. Najadat, H.; Altiti, O.; Aqouleh, A.A.; Younes, M. Credit card fraud detection based on machine and deep learning. In Proceedings of the 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 7–9 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 204–208.
6. Rtayli, N.; Enneya, N. Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *J. Inf. Secur. Appl.* **2020**, *55*, 102596. [CrossRef]
7. Kumar, S.; Gunjan, V.K.; Ansari, M.D.; Pathak, R. Credit card fraud detection using support vector machine. In Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2021, Pune, India, 3–5 May 2023; Springer: Singapore, 2022; pp. 27–37.
8. Hussein, A.S.; Khairy, R.S.; Najeeb, S.M.M.; Alrikabi, H.T.S. Credit Card Fraud Detection Using Fuzzy Rough Nearest Neighbor and Sequential Minimal Optimization with Logistic Regression. *Int. J. Interact. Mob. Technol.* **2021**, *15*, 24–42. [CrossRef]

9. Alenzi, H.Z.; Aljehane, N.O. Fraud detection in credit cards using logistic regression. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 540–551. [CrossRef]
10. Xuan, S.; Liu, G.; Li, Z.; Zheng, L.; Wang, S.; Jiang, C. Random forest for credit card fraud detection. In Proceedings of the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 27–29 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
11. Dileep, M.; Navaneeth, A.; Abhishek, M. A novel approach for credit card fraud detection using decision tree and random forest algorithms. In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6 February 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1025–1028.
12. Taha, A.A.; Malebary, S.J. An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE Access* **2020**, *8*, 25579–25587. [CrossRef]
13. Mishra, A.; Ghorpade, C. Credit card fraud detection on the skewed data using various classification and ensemble techniques. In Proceedings of the 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 24–25 February 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.
14. Fu, K.; Cheng, D.; Tu, Y.; Zhang, L. Credit card fraud detection using convolutional neural networks. In *Neural Information: 23rd International Conference, ICONIP 2016, Kyoto, Japan, October 16–21, 2016, Proceedings, Part III*; Springer: Cham, Switzerland, 2016; pp. 483–490.
15. Sadgali, I.; Sael, N.; Benabbou, F. Fraud detection in credit card transaction using neural networks. In Proceedings of the 4th International Conference on Smart City Applications, Casablanca, Morocco, 2–4 October 2019; pp. 1–4.
16. Benchaji, I.; Douzi, S.; El Ouahidi, B. Credit card fraud detection model based on LSTM recurrent neural networks. *J. Adv. Inf. Technol.* **2021**, *12*, 113–118. [CrossRef]
17. Roy, A.; Sun, J.; Mahoney, R.; Alonzi, L.; Adams, S.; Beling, P. Deep learning detecting fraud in credit card transactions. In Proceedings of the 2018 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 27 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 129–134.
18. Wiese, B.; Omlin, C. Credit card transactions, fraud detection, and machine learning: Modelling time with LSTM recurrent neural networks. In *Innovations in Neural Information Paradigms and Applications*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 231–268.
19. Liu, G.; Tang, J.; Tian, Y.; Wang, J. Graph neural network for credit card fraud detection. In Proceedings of the 2021 International Conference on Cyber-Physical Social Intelligence (ICCSI), Beijing, China, 18–20 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6.
20. Jing, R.; Tian, H.; Zhou, G.; Zhang, X.; Zheng, X.; Zeng, D.D. A GNN-based Few-shot learning model on the Credit Card Fraud detection. In Proceedings of the 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), Beijing, China, 15 July–15 August 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 320–323.
21. Shi, F.; Zhao, C. Enhancing financial fraud detection with hierarchical graph attention networks: A study on integrating local and extensive structural information. *Financ. Res. Lett.* **2023**, *58*, 104458. [CrossRef]
22. Liu, C.; Sun, L.; Ao, X.; Feng, J.; He, Q.; Yang, H. Intention-aware heterogeneous graph attention networks for fraud transactions detection. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Virtual Event, 14–18 August 2021; pp. 3280–3288.
23. Gu, K. Deep Learning Techniques in Financial Fraud Detection. In Proceedings of the 7th International Conference on Cyber Security and Information Engineering, Brisbane, Australia, 23–25 September 2022; pp. 282–286.
24. Wang, Y.; Zhang, J.; Guo, S.; Yin, H.; Li, C.; Chen, H. Decoupling representation learning and classification for gnn-based anomaly detection. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, 11–15 July 2021; pp. 1239–1248.
25. Yang, Q. Toward responsible ai: An overview of federated learning for user-centered privacy-preserving computing. *ACM Trans. Interact. Intell. Syst. (TIIS)* **2021**, *11*, 32. [CrossRef]
26. Yang, W.; Zhang, Y.; Ye, K.; Li, L.; Xu, C.Z. Ffd: A federated learning based method for credit card fraud detection. In *Big Data—BigData 2019, Proceedings of the 8th International Congress, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, 25–30 June 2019*; Proceedings; Springer: Cham, Switzerland, 2019; pp. 18–32.
27. Abdul Salam, M.; Fouad, K.M.; Elbably, D.L.; Elsayed, S.M. Federated learning model for credit card fraud detection with data balancing techniques. *Neural Comput. Appl.* **2024**, *36*, 6231–6256. [CrossRef]
28. Zheng, W.; Yan, L.; Gou, C.; Wang, F.Y. Federated meta-learning for fraudulent credit card detection. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, Online, 7–15 January 2021; pp. 4654–4660.
29. Byrd, D.; Polychroniadou, A. Differentially private secure multi-party computation for federated learning in financial applications. In Proceedings of the First ACM International Conference on AI in Finance, New York, NY, USA, 15–16 October 2020; pp. 1–9.

30. Kanamori, S.; Abe, T.; Ito, T.; Emura, K.; Wang, L.; Yamamoto, S.; Le, T.P.; Abe, K.; Kim, S.; Nojima, R.; et al. Privacy-preserving federated learning for detecting fraudulent financial transactions in japanese banks. *J. Inf. Process.* **2022**, *30*, 789–795. [CrossRef]
31. Long, G.; Tan, Y.; Jiang, J.; Zhang, C. Federated learning for open banking. In *Federated Learning: Privacy and Incentive*; Springer: Cham, Switzerland, 2020; pp. 240–254.
32. Ge, J.; Xu, G.; Lu, J.; Xu, C.; Sheng, Q.Z.; Zheng, X. FedAGA: A federated learning framework for enhanced inter-client relationship learning. *Knowl.-Based Syst.* **2024**, *286*, 111399. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.