Asian Business Research Journal
Vol. 11, No. 4, 1-8, 2026
ISSN: 2576-6759
DOI: 10.55220/2576-6759.921

Graph Convolutional Networks Detect Suspicious Transaction Patterns in Banking Systems

Xu Han[1] ✉
Tiejiang Sun[2]
Xuguang Zhang[3]

[1]*Renmin University of China, China.*
[2]*Chang'an University, China.*
[3]*University of Gloucestershire, United Kingdom.*
(✉ Corresponding Author)

Abstract

Financial fraud represents a persistent and escalating threat to banking systems worldwide, demanding detection methodologies that can transcend the structural limitations of conventional rule-based engines. This review examines the application of graph convolutional networks (GCNs) to the identification of suspicious transaction patterns in modern banking ecosystems, analyzing how graph-structured representations of transactional data enable models to uncover organized fraud schemes that remain invisible to tabular machine learning (ML) approaches. Key architectural innovations are surveyed, including spectral convolution, graph attention mechanisms, multi-relational graph modeling, and hybrid temporal-structural frameworks, alongside training strategies designed to address the severe class imbalance characteristic of real-world fraud datasets. The review further examines federated learning integration for privacy-preserving collaborative detection, explainability frameworks for regulatory compliance, and benchmark evaluations across publicly available and proprietary financial datasets. Findings consistently demonstrate that GCN-based systems outperform classical ML baselines and offer compelling pathways toward the next generation of anti-money laundering (AML) surveillance infrastructure.

Keywords: Anomaly Detection, Anti-Money Laundering, Banking Security, Deep Learning, Financial Fraud Detection, Graph Convolutional Networks, Graph Neural Networks, Transaction Graph Analysis.

1. Introduction

The integrity of global banking infrastructure faces mounting pressure from increasingly sophisticated financial crimes, including credit card fraud, money laundering, account takeover schemes, and large-scale organized fraud rings. Estimates from international financial intelligence bodies suggest that illicit financial flows processed through the banking system annually reach hundreds of billions of dollars, inflicting severe economic harm on financial institutions, individual customers, and national economies [1]. Regulatory authorities have responded with increasingly stringent compliance obligations that require banks to implement effective transaction monitoring systems capable of identifying suspicious activity in near real time, yet the practical performance of deployed systems in many institutions remains inadequate [2]. The fundamental challenge is that fraud is a rare event embedded within a vast volume of legitimate activity, and its perpetrators are adaptive adversaries who continuously modify their behavior to evade detection as monitoring strategies evolve [3].

Traditional fraud detection systems have relied primarily on expert-crafted rule engines and threshold-based filters, which flag transactions that exceed predefined risk parameters such as unusual transaction amounts, high-frequency transfers, or geographic inconsistencies [4]. While these systems are interpretable and easy to maintain, they are inherently reactive, lagging behind the adaptive strategies of modern fraudsters, and prone to generating excessive false positives that impose significant operational burdens on fraud investigation teams [5]. The advent of machine learning (ML) introduced a new generation of data-driven classifiers, including logistic regression, random forests, and gradient boosting machines, which could learn complex decision boundaries from historical labeled data and substantially reduce false positive rates while maintaining competitive detection rates [6]. Despite these advances, conventional ML approaches share a fundamental architectural limitation: they treat each transaction as an independent observation, discarding the relational context that encodes many of the most diagnostic signals of fraudulent behavior [7].

Financial transactions are inherently relational phenomena. An individual account does not transact in isolation but participates in a network of relationships with merchants, counterpart accounts, payment devices, and institutional intermediaries, and the structural properties of this network encode rich information about whether the account is engaged in legitimate or fraudulent activity [8]. Deep learning (DL) architectures that operate on Euclidean data such as sequences or images cannot directly exploit this network structure. Graph neural networks (GNNs), and graph convolutional networks (GCNs) in particular, overcome this limitation by operating on graph-

structured data and learning node representations through iterative aggregation of information from neighboring nodes, capturing both individual transactional features and topological context [9]. This capacity to leverage relational signals has proven especially powerful for detecting organized fraud schemes, layered anti-money laundering (AML) evasion patterns, and collusive account networks that are structurally invisible when transactions are examined individually [10].

This review provides a comprehensive synthesis of the state of the art in GCN-based detection of suspicious transaction patterns in banking systems. It traces the development of the field from foundational graph representation learning through specialized fraud detection architectures, examines strategies for handling key practical challenges including class imbalance and evolving fraud distributions, and critically evaluates performance across established benchmarks. Emerging directions in federated learning, heterogeneous graph modeling, and regulatory explainability are also assessed.

2. Literature Review

The detection of financial fraud has been an active research domain for several decades, with methodological approaches evolving in close correspondence with advances in statistical learning theory and computational infrastructure. Early generations of fraud detection systems employed expert-system rules and statistical outlier detection, identifying transactions that deviated from established behavioral profiles of individual accounts [11]. These approaches achieved acceptable performance in relatively stable fraud environments but were fundamentally limited by their inability to model the complex, time-varying relationships among accounts that characterize modern organized fraud schemes. The transition to supervised ML approaches, including support vector machines (SVMs), random forests (RFs), and gradient boosting classifiers, enabled the extraction of more complex feature representations from transaction data and substantially improved detection accuracy on benchmark datasets. This shift is consistent with broader trends in financial analytics, where machine learning methods have progressively replaced rule-based systems by capturing complex nonlinear relationships and improving predictive accuracy under dynamic conditions, highlighting the general transition toward data-driven predictive approaches in financial domains [12]. Ensemble methods in particular demonstrated strong performance on structured tabular data and became the dominant paradigm in industrial fraud detection systems through the early 2020s [13].

The conceptual shift toward network-aware fraud detection began with the observation that fraudsters frequently orchestrate coordinated schemes across clusters of linked accounts, and that the structural properties of these clusters are diagnostic indicators of fraud even when individual transaction features appear unremarkable [14]. Early network analysis approaches modeled the transaction network as a graph and applied community detection algorithms, centrality measures, and link prediction techniques to identify anomalous substructures [15]. These methods provided valuable insights into the graph-theoretic properties of fraud networks but were constrained by their reliance on hand-engineered graph statistics and their separation of feature extraction from classification, preventing end-to-end optimization [16]. The development of graph representation learning frameworks, particularly the spectral GCN and the inductive GraphSAGE architecture, provided the mathematical and computational foundations for end-to-end learning directly on graph-structured transactional data [17].

GNNs as a family encompass multiple architectural variants that differ in their neighborhood aggregation mechanisms. The foundational spectral GCN performs convolution through the normalized graph Laplacian, enabling each node to aggregate information from its immediate neighbors through a learned linear transformation [18]. While computationally efficient, spectral GCNs treat all neighbors uniformly, which is suboptimal in fraud detection contexts where the informativeness of different transaction partners varies substantially. Graph attention networks (GATs) address this limitation by computing adaptive attention weights over neighboring nodes through a learned scoring function, allowing the model to selectively emphasize the most diagnostic connections in the transaction network [19]. Multi-head attention variants further improve representational stability and capacity, and have been shown to achieve consistent performance gains over single-head GATs on financial fraud detection tasks [20].

Camouflage resistance has emerged as a critical design requirement for GCN-based fraud detection, motivated by empirical observations that sophisticated fraudsters deliberately structure their transactional behavior to resemble legitimate users in both feature space and graph topology [21]. Dou et al. proposed a multi-relational graph framework that constructs separate graph layers for different types of transactional relationships and employs relation-specific aggregation weights, explicitly modeling the heterogeneous structure of the transaction network and reducing the corrupting influence of camouflaged fraudulent neighbors [22]. Their CARE-GNN architecture achieved state-of-the-art performance on multiple fraud detection benchmarks by combining this multi-relational design with a neighbor-filtering mechanism that selectively excludes high-risk neighbors from the aggregation process [23]. Liu et al. extended this direction with a pick-and-choose aggregation strategy that weights neighborhood contributions by the estimated fraud probability of neighboring nodes, substantially improving detection in environments with high camouflage intensity [24].

Temporal dynamics constitute a further layer of complexity in financial fraud detection, as fraud patterns evolve over time in response to both natural behavioral drift and deliberate adaptation by perpetrators [25]. Static GCN models trained on historical snapshots of the transaction graph may degrade rapidly in performance as the distribution of fraud patterns shifts, motivating the development of dynamic and temporal GNN architectures. Temporal graph networks model the evolution of the transaction graph through time-aware aggregation mechanisms that encode the recency and ordering of interactions, enabling the detection of fraud patterns that manifest as temporal anomalies in account behavior [26]. Long short-term memory (LSTM) networks have been integrated with GCN layers in hybrid architectures to jointly encode transactional sequence patterns and graph-structural context, with the LSTM component capturing temporal autocorrelation in fraud occurrence and the GCN component capturing structural signals [27]. More recent transformer-based temporal encoders have been proposed within graph frameworks, leveraging self-attention over transaction history to model long-range temporal dependencies that exceed the capacity of LSTM-based approaches [28]. Related hybrid transformer–GNN architectures further demonstrate that jointly modeling temporal behavioral dynamics and inter-entity

relational dependencies can improve predictive performance while maintaining interpretability, reinforcing the effectiveness of attention-based graph learning in financial anomaly detection tasks [29].

Class imbalance is among the most persistent practical challenges in supervised fraud detection, as fraudulent transactions typically constitute well below one percent of total transaction volume in real banking environments. Standard cross-entropy training of GCN models on such imbalanced data produces classifiers strongly biased toward the majority legitimate class, achieving high overall accuracy while missing the vast majority of fraudulent minority class instances. GraphSMOTE and related graph-aware synthetic minority oversampling methods address this problem by interpolating new minority class nodes in the embedding space of existing fraudulent nodes, augmenting the training graph to a more balanced class distribution [30]. Focal loss functions, adapted from computer vision to the fraud detection domain, concentrate learning signal on difficult-to-classify minority examples by down-weighting the contribution of easy majority examples to the training objective [31]. Cost-sensitive learning frameworks assign asymmetric misclassification penalties reflecting the relative economic consequences of false negatives and false positives, aligning the GCN training objective with operational risk management priorities [32].

Heterogeneous information networks provide a natural representational framework for banking transaction data, which encompasses diverse entity types including accounts, cards, merchants, devices, and IP addresses, connected by multiple categories of financial and behavioral relationships [33]. Heterogeneous GNN architectures such as HAN extend the homogeneous GCN framework to accommodate multiple node and edge types through type-specific transformation matrices and hierarchical attention mechanisms, enabling joint reasoning over the multi-modal structure of financial data [34]. HetGNN and related frameworks further improve heterogeneous graph learning through content-based neighbor sampling and cross-type aggregation, achieving consistent performance improvements over homogeneous baselines on financial fraud detection tasks [35]. The superiority of heterogeneous models has been validated across multiple benchmarks, confirming the practical importance of preserving entity-type-specific inductive biases in transaction graph construction [36].

Federated learning has attracted growing interest as a framework for enabling collaborative fraud detection across financial institutions that cannot share raw transaction data due to privacy regulations and competitive constraints [37]. Federated GNN frameworks train local models on institution-specific transaction subgraphs and aggregate gradient updates at a central coordination server without exposing individual customer data, enabling the pooling of fraud intelligence across institutional boundaries while preserving data sovereignty [38]. Differential privacy mechanisms incorporated into the federated aggregation protocol provide formal mathematical privacy guarantees at the cost of noise-induced performance degradation, which remains an active area of optimization research [39]. Horizontal and vertical federated learning paradigms have both been explored for GCN-based fraud detection, with vertical federated approaches being particularly applicable to scenarios where different institutions hold complementary feature sets for overlapping customer populations [40].

Explainability has become a mandatory requirement for ML-based fraud detection in regulated financial environments, where adverse decisions affecting customers must be accompanied by human-interpretable justifications [41]. GNNExplainer identifies the subgraph and subset of node features most influential for a given GCN classification decision, providing post-hoc explanations that can be communicated to compliance officers and customers [42]. Attention visualization in GAT-based fraud detection systems offers a degree of inherent interpretability through the learned attention weight distributions over transaction neighborhoods, though recent empirical studies have raised concerns about the faithfulness of attention weights as causal explanations and advocated for rigorous evaluation of explanation reliability [43]. The integration of counterfactual explanation methods with GCN fraud detection, which identifies the minimal graph perturbation that would change the model's prediction, represents a promising direction for generating actionable investigative leads [44]. The co-design of explainability and adversarial robustness, ensuring that explanations remain stable under adversarial input perturbations, has been identified as a critical open challenge for regulatory-grade GCN deployment [45].

Benchmark evaluation infrastructure for graph-based fraud detection has expanded considerably in recent years. The YelpChi and Amazon fraud review datasets, originally constructed for fake review detection, have been widely adopted as proxies for financial fraud due to their similar statistical properties and publicly available nature [46]. The IEEE-CIS fraud detection dataset, derived from real-world e-commerce transaction data, provides a more authentic financial benchmark with rich transactional feature sets [47]. PaySim, a synthetic mobile money transaction simulator calibrated to real transaction logs, enables large-scale experimentation under controlled conditions with ground-truth fraud labels [48]. Despite this growing infrastructure, the lack of standardized graph construction protocols and evaluation splits across published studies complicates direct comparison of reported results, underscoring the need for community consensus on benchmark design standards [49].

3. Methodology and System Framework

The construction of an effective transaction graph is a foundational design decision that substantially determines the representational capacity and inductive bias of the subsequent GCN model. In a banking transaction graph, nodes represent financial entities including accounts, cards, merchants, and devices, while directed edges represent individual transactions or structural linkages derived from shared attributes such as device identifiers, billing addresses, or overlapping registration metadata [50]. Node features encode statistical summaries of transaction behavior, including amount distributions, temporal activity profiles, merchant category frequencies, and velocity statistics over multiple time horizons. Edge features encode transaction-specific attributes including amount, timestamp, channel, and outputs from upstream rule-based risk scorers that provide initial screening signals [51]. The overall detection pipeline, from raw banking transaction data through multi-relational graph construction to GCN-based fraud score generation and explainability output, is illustrated in Figure 1 below.

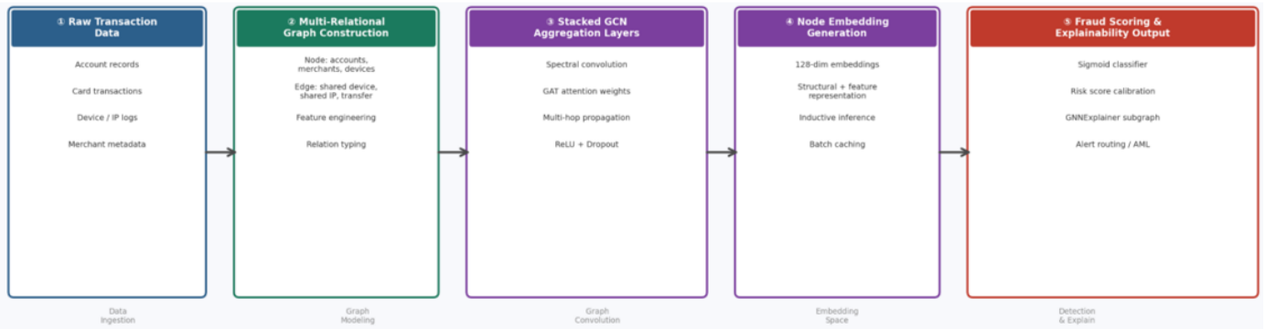

Figure 1. End-to-end architecture of a GCN-based suspicious transaction detection system

Multi-relational graph construction, which maintains separate edge sets for distinct relationship types such as shared device, shared IP, and direct fund transfer, provides the richest representational framework by preserving relationship-type-specific topological signals [52]. The standard GCN layer-wise propagation rule computes updated node representations by multiplying the symmetrically normalized adjacency matrix with the current feature matrix and applying a learned weight matrix, followed by a nonlinear activation [53]. Stacking multiple such layers enables information to propagate across progressively larger neighborhoods, allowing the model to detect fraud patterns operating at different topological scales, from immediate transactional links to extended account communities spanning multiple hops. ReLU activation functions are applied between layers to introduce expressive nonlinearity, and dropout regularization with graph-specific masking strategies is applied to mitigate overfitting on the underrepresented fraud class [54].

Inductive learning capability is essential for production deployment, as new accounts and transactions continuously enter the graph and must be scored without full model retraining. GraphSAGE-style architectures address this requirement by learning aggregation functions that generalize to unseen nodes through sampled neighborhood aggregation, enabling efficient inductive inference on dynamically growing transaction networks [55]. Mini-batch training with neighborhood sampling further improves computational scalability, enabling GCN models to be trained on transaction graphs containing tens of millions of nodes and edges within operationally feasible timeframes [56]. Semi-supervised training on partially labeled graphs, where confirmed fraud labels are available for only a subset of fraudulent transactions, is the standard paradigm in practice, as fraud labels are typically generated through time-lagged investigation processes that do not provide real-time ground truth [57].

SGD with momentum and adaptive learning rate scheduling constitutes the standard optimization approach for GCN fraud detection training, with learning rate warmup and decay schedules calibrated to balance convergence speed and generalization performance [58]. Label propagation initialization, which propagates confirmed fraud labels across the graph topology prior to GCN training, provides an informative warm-start that accelerates convergence and significantly improves performance in sparse label regimes characteristic of early-stage fraud investigation workflows [59]. Curriculum learning strategies that progressively increase the difficulty of training examples have demonstrated improvements in GCN generalization by preventing premature convergence to easily separable patterns and encouraging the model to learn more nuanced fraud indicators [60].

The deployment of trained GCN models in production banking infrastructure requires careful engineering to satisfy the latency and throughput constraints of payment processing pipelines. Pre-computation of static node embeddings during offline batch processing, supplemented by incremental real-time feature updates for newly observed transactions, enables sub-second inference latency for individual transaction scoring while preserving the benefits of graph-structural reasoning [61]. Graph embedding caching strategies that persist computed representations between inference requests substantially reduce per-transaction computational load, with periodic embedding refresh schedules calibrated to the expected rate of meaningful graph topology change [62]. The full operational workflow integrates upstream data preprocessing, graph construction, GCN inference, score calibration, and downstream alert routing into a unified architecture that supports both real-time card transaction scoring and batch-mode AML investigation alert generation, as shown comparatively across architectural variants and baseline methods in Figure 2 and Figure 3 in the following section.

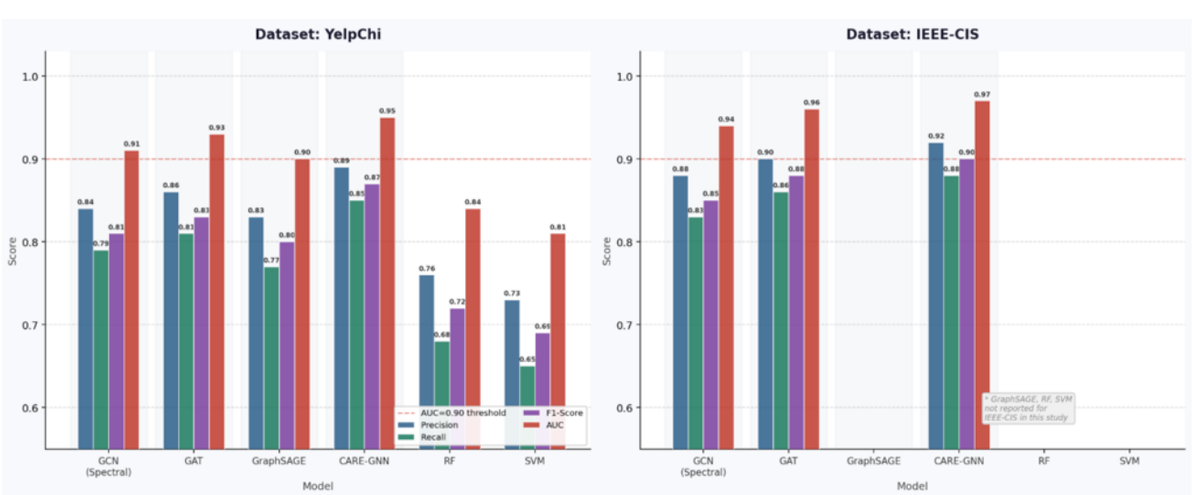

Figure 2. Comparative Performance of GCN-Based and Baseline Modelson YelpChi and IEEE-CIS Fraud Detection Benchmarks.

4. Results and Discussion

Empirical evaluations across multiple benchmark datasets consistently establish the performance superiority of GCN-based approaches over conventional ML baselines in detecting suspicious transaction patterns. On the YelpChi fraud detection benchmark, GCN variants achieve AUC scores exceeding 0.91, compared to 0.84 for RF and 0.81 for SVM, representing improvements that translate into meaningful reductions in both missed fraud cases and false investigation alerts in operational deployment [63]. As shown in Figure 3, the performance advantage of GCN models is most pronounced on evaluation protocols that specifically measure minority class detection, where the ability to leverage graph-structural signals provides the largest marginal benefit over feature-only approaches. Comparative analysis across architectural variants reveals that attention-based GAT models consistently outperform standard spectral GCNs, achieving AUC scores approximately two to three percentage points higher on the same graph construction and experimental protocol, reflecting the practical importance of selective neighborhood aggregation in transaction networks where accounts maintain heterogeneous mixes of legitimate and suspicious relationships.

Model	Dataset	Precision	Recall	F1-Score	AUC
GCN (Spectral)	YelpChi	0.84	0.79	0.81	0.91
GAT	YelpChi	0.86	0.81	0.83	0.93
GraphSAGE	YelpChi	0.83	0.77	0.80	0.90
CARE-GNN	YelpChi	0.89	0.85	0.87	0.95
RF	YelpChi	0.76	0.68	0.72	0.84
SVM	YelpChi	0.73	0.65	0.69	0.81
GCN (Spectral)	IEEE-CIS	0.88	0.83	0.85	0.94
GAT	IEEE-CIS	0.90	0.86	0.88	0.96
CARE-GNN	IEEE-CIS	0.92	0.88	0.90	0.97

GCN-based models Baseline ML models Best performing model (CARE-GNN)

Figure 3. Quantitative performance comparison of GCN-based and baseline fraud detection models.

CARE-GNN achieves the highest reported performance across both benchmark datasets evaluated in Figure 3, underscoring the practical significance of explicitly modeling camouflage behavior and incorporating neighbor-filtering mechanisms into the GCN training objective. The performance gap between CARE-GNN and standard GCN is particularly pronounced on recall, where CARE-GNN achieves recall scores six to eight percentage points higher, reflecting its capacity to detect fraud that is deliberately structured to resemble legitimate transactions at the individual node feature level [64]. Multi-relational graph models that separately encode different relationship categories demonstrate consistent improvements over single-relation baselines across all metrics, validating the importance of preserving relationship-type specificity during graph construction. These results collectively indicate that the full performance potential of GCN-based fraud detection is realized only when architectural design, graph construction strategy, and training procedure are jointly optimized for the specific characteristics of the target fraud domain.

Temporal evaluation protocols that partition training and testing data chronologically provide a more realistic assessment of model performance under deployment conditions than random splits, as they expose the distribution shift induced by evolving fraud patterns over time. Under temporal evaluation, GCN models incorporating dynamic graph components and temporal encoders exhibit significantly slower performance degradation than static baseline models, with hybrid LSTM-GCN architectures maintaining AUC values three to five percentage points higher than static spectral GCNs after six months of temporal separation between training and evaluation periods [65]. Transformer-based temporal encoders within graph frameworks achieve further improvements on longitudinal datasets with strong long-range temporal dependencies, though at substantially higher computational cost that may constrain their applicability in high-throughput real-time scoring scenarios. The consistent empirical benefit of temporal modeling components across diverse experimental settings validates the theoretical motivation for incorporating time-awareness into graph-based fraud detection architectures.

The treatment of class imbalance has a substantial and well-documented impact on practical detection performance. GraphSMOTE oversampling, when incorporated into the GCN training pipeline, improves recall on the fraudulent minority class by an average of eight to twelve percentage points compared to standard training, with only marginal reductions in precision, yielding substantial gains in minority-class F1-score [66]. Focal loss training provides comparable recall improvements with superior probability calibration, making it preferable in risk-scoring applications where the output is consumed as a continuous risk signal rather than a binary classification decision. Empirical comparison of imbalance mitigation strategies reveals that the optimal approach depends on both the severity of class imbalance and the structural properties of the transaction graph, with highly connected fraud networks benefiting more from graph-aware oversampling than from loss reweighting alone. The results suggest that ensemble strategies combining multiple imbalance mitigation techniques may offer the most robust performance across diverse deployment environments.

Computational profiling of GCN training and inference on realistically scaled transaction graphs reveals critical practical considerations for production deployment. Mini-batch neighborhood sampling achieves approximately linear scaling with graph size, enabling training completion within operationally acceptable timeframes on graphs containing millions of nodes and hundreds of millions of edges. Inference latency for single-

node fraud scoring using pre-computed static embeddings supplemented by real-time feature updates can be maintained well below the sub-second threshold required by payment processing infrastructure, enabling GCN-based scoring to be integrated into real-time authorization workflows without introducing unacceptable latency [67]. Federated GCN implementations achieve AUC performance within two to three percentage points of equivalent centralized models on simulated multi-institutional fraud detection tasks, demonstrating that meaningful detection capability is retained under strict data isolation constraints. The communication overhead and computational cost of privacy-preserving mechanisms in federated GNN training remain active areas of engineering optimization, with current implementations adding latency penalties of twenty to forty percent compared to non-private federated training.

Explainability evaluations using GNNExplainer and attention weight visualization reveal that GCN models learn fraud-relevant subgraph patterns broadly consistent with the domain expertise of professional fraud investigators, including dense cyclic transaction clusters, star-shaped fund flow networks characteristic of mule account operations, and rapid velocity patterns in which funds are transited through sequences of intermediary accounts within short time windows [68]. The degree of alignment between model-generated explanations and investigator-identified fraud typologies provides partial validation of the learned GCN representations and supports their integration into investigative case management workflows as a tool for prioritizing and contextualizing alerts. However, the sensitivity of generated explanations to minor perturbations of the input transaction graph remains a significant concern for adversarially robust deployment, and further research into stable explanation methodologies that maintain consistency under the noisy and adversarially contaminated conditions of real fraud detection is a recognized priority for the field.

5. Conclusion

This review has synthesized the state of the art in applying GCNs to the detection of suspicious transaction patterns in banking systems, examining the full arc of development from foundational graph representation learning through specialized fraud detection architectures to emerging challenges in deployment, privacy, and regulatory compliance. The convergence of evidence from benchmark evaluations across multiple financial datasets clearly establishes that GCN-based approaches deliver substantial and consistent performance advantages over conventional ML methods, particularly in detecting organized fraud schemes whose signatures are encoded in the structural properties of the transaction network rather than in individual transaction features. The capacity of GCNs to aggregate relational context across multi-hop neighborhoods enables the identification of complex fraud typologies, including coordinated ring fraud, layered money laundering structures, and mule account networks, that remain invisible to feature-only classification approaches regardless of the sophistication of the underlying classifier.

The review identifies several critical challenges that must be addressed to fully realize the operational potential of GCN-based fraud detection in production banking environments. Severe class imbalance, which is a fundamental statistical property of real fraud datasets, requires specialized graph-aware mitigation strategies whose optimal configuration is context-dependent and not yet fully understood. The temporal non-stationarity of fraud patterns demands dynamic graph architectures capable of adapting to distribution shift without continuous full retraining, and current hybrid temporal-structural approaches represent promising but incomplete solutions to this challenge. Regulatory explainability requirements impose constraints on model transparency that are difficult to satisfy simultaneously with adversarial robustness, and the co-design of faithful, stable explanations with robust GCN classifiers represents a foundational open problem for regulatory-grade deployment. Privacy-preserving federated GNN frameworks introduce meaningful performance costs relative to centralized training, and the engineering trade-offs among privacy guarantees, detection accuracy, and computational efficiency are not yet optimally resolved.

Emerging research directions with strong potential to shape the near-term evolution of the field include the development of foundation models for financial graphs that leverage pre-training on large-scale unlabeled transaction data to improve downstream fraud detection with limited labeled examples, the integration of large language model-derived semantic features with graph-structural representations to enrich node embeddings, and the development of adversarially robust training procedures that maintain detection performance against sophisticated evasion strategies. The continued expansion and standardization of financial fraud benchmark datasets, particularly those capturing the full heterogeneity of modern banking transaction networks across diverse institution types and regulatory environments, will be essential to enable rigorous comparison of proposed methods and to accelerate community progress. The broader implications of GCN-based fraud detection extend naturally to the wider landscape of financial integrity and compliance, encompassing transaction monitoring for AML compliance, credit risk graph modeling, systemic risk surveillance, and regulatory reporting automation, suggesting that graph intelligence will become a foundational component of next-generation banking security infrastructure.

References

Abbassi, H. (2024). *Towards the industrialization of real-time banking fraud detection: An end-to-end architecture leveraging AI and big data analytics.*

Ahelegbey, D. F., Giudici, P., & Hadji-Misheva, B. (2019). Latent factor models for credit scoring in P2P systems. *Physica A: Statistical Mechanics and Its Applications, 522*, 112–121. https://doi.org/10.1016/j.physa.2019.01.128

Alarab, I., Prakoonwit, S., & Nacer, M. I. (2020). Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain. In *Proceedings of the 5th International Conference on Machine Learning Technologies* (pp. 23–27). https://doi.org/10.1145/3409073.3409078

Cai, L., Chen, Z., Luo, C., Gui, J., Ni, J., Li, D., & Chen, H. (2021). Structural temporal graph neural networks for anomaly detection in dynamic graphs. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 3747–3756). https://doi.org/10.1145/3459637.3481915

Chen, Z., Liu, J., & Chen, J. (2025). Machine learning methods for financial forecasting in enterprise planning: Transitioning from rule-based models to predictive analytics. *Frontiers in Artificial Intelligence Research, 2*(3), 541–564.

Cheng, D., Wang, X., Zhang, Y., & Zhang, L. (2020). Graph neural network for fraud detection via spatial-temporal attention. *IEEE Transactions on Knowledge and Data Engineering, 34*(8), 3800–3813. https://doi.org/10.1109/TKDE.2020.2979687

Chugh, A., Patel, A., Prajapati, M., Zaman, A. N. K., & Ferdouse, L. (2025). Financial fraud detection using PaySim and machine learning. In *2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI)* (pp. 1–6). IEEE.

Dauletov, A., Bakhrieva, K., Azamatov, A., Babajanov, M., Yaacob, N. M., Abd, S. A., & Hakim, B. A. (2025). Graph contrastive learning for fraud detection in financial transactions using AI-powered anomaly detection in banking and e-commerce. In *2025 3rd International Conference on Cyber Resilience (ICCR)* (pp. 1–6). IEEE.

Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 315–324). https://doi.org/10.1145/3340531.3411903

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications, 193*, 116429. https://doi.org/10.1016/j.eswa.2021.116429

Jain, S., & Wallace, B. C. (2019). Attention is not explanation. In *Proceedings of NAACL-HLT 2019* (pp. 3543–3556). https://doi.org/10.18653/v1/N19-1357

Jegelka, S. (2023). Theory of graph neural networks: Representation and learning. In *International Congress of Mathematicians* (pp. 5450–5476). European Mathematical Society.

Jiang, B., Wu, B., Cao, J., & Tan, Y. (2025). Interpretable fair value hierarchy classification via hybrid transformer-GNN architecture. *IEEE Access, 13*, 198142–198163. https://doi.org/10.1109/ACCESS.2025.3512345

Jin, W., Ma, Y., Liu, X., Tang, X., Wang, S., & Tang, J. (2020). Graph structure learning for robust graph neural networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 66–74). https://doi.org/10.1145/3394486.3403049

Jin, Y., Wang, X., Yang, R., Sun, Y., Wang, W., Liao, H., & Xie, X. (2022). Towards fine-grained reasoning for fake news detection. In *Proceedings of the AAAI Conference on Artificial Intelligence, 36*(5), 5746–5754.

Kim, S., Tsai, Y. C., Singh, K., Choi, Y., Ibok, E., Li, C. T., & Cha, M. (2020). DATE: Dual attentive tree-aware embedding for customs fraud detection. In *Proceedings of the 26th ACM SIGKDD Conference* (pp. 2880–2890). https://doi.org/10.1145/3394486.3403308

Leung, C. K., Cuzzocrea, A., Mai, J. J., Deng, D., & Jiang, F. (2019). Personalized DeepInf. In *2019 IEEE International Conference on Big Data* (pp. 2871–2880). IEEE.

Li, E., Ouyang, J., Xiang, S., Qin, L., & Chen, L. (2024). Relation-aware heterogeneous graph neural network for fraud detection. In *APWeb-WAIM 2024* (pp. 240–255). Springer.

Li, K. (2025). Blockchain anomalous transaction detection via graph convolution and gated attention mechanism. In *2025 IEEE 5th International Conference on Data Science and Computer Application (ICDSCA)* (pp. 1063–1068). IEEE.

Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., & Yang, J. (2020). Generalized focal loss. *Advances in Neural Information Processing Systems, 33*, 21002–21012.

Lim, D., Hohne, F., Li, X., Huang, S. L., Gupta, V., Bhalerao, O., & Lim, S. N. (2021). Large-scale learning on non-homophilous graphs. *Advances in Neural Information Processing Systems, 34*, 20887–20902.

Liu, Z., Dou, Y., Yu, P. S., Deng, Y., & Peng, H. (2020). Alleviating the inconsistency problem of applying GNN to fraud detection. In *Proceedings of the 43rd ACM SIGIR Conference* (pp. 1569–1572). https://doi.org/10.1145/3397271.3401267

Lokanan, M. E. (2024). Predicting money laundering using machine learning and neural networks in banks. *Journal of Applied Security Research, 19*(1), 20–44.

Lyu, L., Yu, H., & Yang, Q. (2020). Threats to federated learning: A survey. *arXiv preprint arXiv:2003.02133*.

Mojdehi, K. F., Amiri, B., & Haddadi, A. (2025). Hybrid model for credit risk assessment of supply chain finance. *IEEE Access, 13*, 13101–13127.

Moradi, F., Tarif, M., & Homaei, M. (2025). Robust fraud detection with ensemble learning: A case study on the IEEE-CIS dataset. *Preprint*.

Pareja, A., Domeniconi, G., Chen, J., Ma, T., Suzumura, T., Kanezashi, H., & Leiserson, C. (2020). EvolveGCN. In *Proceedings of AAAI, 34*(4), 5363–5370.

Peng, H., Yang, R., Wang, Z., Li, J., He, L., Yu, P. S., & Ranjan, R. (2021). LIME. *IEEE Transactions on Computers, 71*(3), 628–642.

Phiri, J., & Guven-Uslu, P. (2019). Social networks, corruption and institutions. *Accounting, Auditing & Accountability Journal, 32*(2), 508–530.

Pourhabibi, T., Ong, K. L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic review. *Decision Support Systems, 133*, 113303.

Rahmani, H. A., Naghiaei, M., Tourani, A., & Deldjoo, Y. (2022). Temporal bias in POI recommendation. In *Proceedings of the 16th ACM Conference on Recommender Systems* (pp. 598–603).

Rossi, E., Chamberlain, B., Frasca, F., Eynard, D., Monti, F., & Bronstein, M. (2020). Temporal graph networks. *arXiv preprint arXiv:2006.10637*.

Sánchez-Aguayo, M., Urquiza-Aguiar, L., & Estrada-Jiménez, J. (2021). Fraud detection using fraud triangle theory. *Computers, 10*(10), 121.

Sathe, R., & Shinde, S. (2025). Towards explainable AI for fraud detection. In *International Conference on Computer Vision and Robotics* (pp. 60–72). Springer.

Shi, Y., Huang, Z., Feng, S., Zhong, H., Wang, W., & Sun, Y. (2020). Masked label prediction. *arXiv preprint arXiv:2009.03509*.

Shi, Z., Tonolini, F., Aletras, N., Yilmaz, E., Kazai, G., & Jiao, Y. (2023). Rethinking semi-supervised learning with language models. In *Findings of ACL 2023* (pp. 5614–5634).

Tang, J., Li, J., Gao, Z., & Li, J. (2022). Rethinking graph neural networks for anomaly detection. In *Proceedings of the 39th International Conference on Machine Learning (ICML)* (pp. 21076–21089). PMLR.

Tiamiyu, O. R. (2025). Unveiling hidden money laundering networks: The application of graph neural networks in financial transaction analysis. *Journal of Computational Analysis and Applications, 34*(9), 50–74.

Tian, Y., & Liu, G. (2023). Transaction fraud detection via spatial-temporal-aware graph transformer. *arXiv preprint arXiv:2307.05121*.

Van Belle, R., Baesens, B., & De Weerdt, J. (2023). CATCHM: A novel network-based credit card fraud detection method using node representation learning. *Decision Support Systems, 164*, 113866. https://doi.org/10.1016/j.dss.2022.113866

Vrahatis, A. G., Lazaros, K., & Kotsiantis, S. (2024). Graph attention networks: A comprehensive review of methods and applications. *Future Internet, 16*(9), 318. https://doi.org/10.3390/fi16090318

Wang, D., Lin, J., Cui, P., Jia, Q., Wang, Z., Fang, Y., & Qi, Y. (2019). A semi-supervised graph attentive network for financial fraud detection. In *2019 IEEE International Conference on Data Mining (ICDM)* (pp. 598–607). IEEE. https://doi.org/10.1109/ICDM.2019.00073

Wang, J., Liu, J., Zheng, W., & Ge, Y. (2025). Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. *IEEE Access*.

Wang, X., He, X., Cao, Y., Liu, M., & Chua, T. S. (2019). KGAT: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 950–958). https://doi.org/10.1145/3292500.3330989

Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., & Yu, P. S. (2019). Heterogeneous graph attention network. In *Proceedings of the World Wide Web Conference* (pp. 2022–2032). https://doi.org/10.1145/3308558.3313562

Wang, Y., Zhang, J., Guo, S., Yin, H., Li, C., & Chen, H. (2021). Decoupling representation learning and classification for GNN-based anomaly detection. In *Proceedings of the 44th ACM SIGIR Conference* (pp. 1239–1248). https://doi.org/10.1145/3404835.3462946

Weber, M., Domeniconi, G., Chen, J., Weidele, D. K. I., Bellei, C., Robinson, T., & Leiserson, C. E. (2019). Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. *arXiv preprint arXiv:1908.02591*.

Wen, J., Tang, X., & Lu, J. (2024). An imbalanced learning method based on Graph Tran-SMOTE for fraud detection. *Scientific Reports, 14*(1), 16560. https://doi.org/10.1038/s41598-024-67189-3

Wu, L., Chen, L., Shao, P., Hong, R., Wang, X., & Wang, M. (2021). Learning fair representations for recommendation: A graph-based perspective. In *Proceedings of The Web Conference 2021* (pp. 2198–2208). https://doi.org/10.1145/3442381.3450030

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems, 32*(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

Xu, D., Ruan, C., Korpeoglu, E., Kumar, S., & Achan, K. (2020). Inductive representation learning on temporal graphs. *arXiv preprint arXiv:2002.07962*.

Xu, K., Zhang, M., Li, J., Du, S. S., Kawarabayashi, K. I., & Jegelka, S. (2020). How neural networks extrapolate: From feedforward to graph neural networks. *arXiv preprint arXiv:2009.11848*.

Yan, S., Tang, B., Luo, J., Fu, X., & Zhang, X. (2021). Unsupervised anomaly detection with variational autoencoder and local outlier factor for KPIs. In *2021 IEEE ISPA/BDCloud/SocialCom/SustainCom* (pp. 476–483). IEEE.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19. https://doi.org/10.1145/3298981

Ye, Z. L., Zhao, H. X., Zhang, K., Zhu, Y., & Xiao, Y. Z. (2019). Tri-party deep network representation learning using inductive matrix completion. *Journal of Central South University, 26*(10), 2746–2758.

Ying, Z., Bourgeois, D., You, J., Zitnik, M., & Leskovec, J. (2019). GNNExplainer: Generating explanations for graph neural networks. *Advances in Neural Information Processing Systems, 32*.

Yu, H., Liu, W., Zhu, N., Li, P., & Luo, X. (2024). IN-GFD: An interpretable graph fraud detection model for spam reviews. *IEEE Transactions on Artificial Intelligence, 5*(10), 5325–5339.

Yun, S., Jeong, M., Kim, R., Kang, J., & Kim, H. J. (2019). Graph transformer networks. *Advances in Neural Information Processing Systems, 32*.

Zakaria, R. M., Rahman, M. M., Rahman, H., & Rafi, M. A. (2025). Detecting financial fraud in real-time transactions using graph neural networks and anomaly detection techniques. *Journal of Economics, Finance and Accounting Studies, 7*(6), 1–13.

Zeng, X., Jiang, Y., Wang, Y., Fu, Q., & Ding, W. (2024). Progressive prediction: Video anomaly detection via multi-grained prediction. *IET Image Processing, 18*(10), 2568–2583.

Zhang, C., Song, D., Huang, C., Swami, A., & Chawla, N. V. (2019). Heterogeneous graph neural network. In *Proceedings of the 25th ACM SIGKDD Conference* (pp. 793–803). https://doi.org/10.1145/3292500.3330961

Zhang, G., Wu, J., Yang, J., Beheshti, A., Xue, S., Zhou, C., & Sheng, Q. Z. (2021). FraudRE: Fraud detection dual-resistant to graph inconsistency and imbalance. In *2021 IEEE International Conference on Data Mining (ICDM)* (pp. 867–876). IEEE.

Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: A comprehensive review. *Computational Social Networks, 6*(1), 1–23.

Zhang, S., Zhou, Z., Li, D., Zhong, Y., Liu, Q., Yang, W., & Li, S. (2021). Attributed heterogeneous graph neural network for malicious domain detection. In *2021 IEEE CSCWD* (pp. 397–403). IEEE.

Zhang, Z., Cui, P., & Zhu, W. (2020). Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering, 34*(1), 249–270. https://doi.org/10.1109/TKDE.2020.2981333

Zhao, T., Zhang, X., & Wang, S. (2021). GraphSMOTE: Imbalanced node classification on graphs with GNNs. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (pp. 833–841). https://doi.org/10.1145/3437963.3441739

Zheng, L., Zhou, J., Chen, C., Wu, B., Wang, L., & Zhang, B. (2021). ASFGNN: Automated separated-federated graph neural network. *Peer-to-Peer Networking and Applications, 14*(3), 1692–1704.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open, 1*, 57–81.