Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 7, No. 1, 2026

Graph Learning Against Adversaries: Bridging Structural Topology and Multimodal Semantics in Fraud Detection

Wenzhen Yang [1, 2], Xi Xiong [1, 2, *]

[1] School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China

[2] Advanced Cryptography and System Security Key Laboratory of Sichuan Province, Chengdu 610225, China

[*] **Corresponding author**: (Email: flyxiongxi@gmail.com)

Abstract: The digital economy's unprecedented expansion has been accompanied by increasingly sophisticated, networked fraudulent activities that exploit systemic vulnerabilities. Traditional tabular machine learning models evaluate entities in isolation, often failing to detect the collusive, multi-step strategies employed by modern adversaries. Consequently, Graph Neural Networks (GNNs) have emerged as the foundational architecture for advanced fraud detection, leveraging complex topological relationships to uncover illicit patterns. This paper systematically traces the evolutionary trajectory from general graph representation learning to Graph Anomaly Detection (GAD), and ultimately to the highly adversarial domain of fraud detection. We deconstruct the critical mathematical and structural challenges inherent in this domain, primarily extreme class imbalance and severe feature heterophily—a phenomenon where malicious entities actively deploy relational camouflage by linking to benign nodes, thereby neutralizing the smoothing mechanisms of standard homophily-assuming GNNs. We analyze the architectural adaptations developed to counteract these adversarial tactics. Furthermore, we critically re-evaluate recent advancements in the field, analyzing empirical evidence that questions the purported absolute superiority of Graph Transformers over meticulously optimized classic GNNs. Finally, we explore the emergent frontier of Large Language Model (LLM)-enhanced graph learning. We detail how sophisticated techniques, such as dual-granularity prompting and agentic semantic vectorization, are actively resolving the catastrophic information loss associated with early-stage multimodal feature encoding. By bridging discrete topological structures with continuous multimodal semantic reasoning, this paper charts the future trajectory of robust, scalable, and explainable adversarial graph learning.

Keywords: Graph Learning, Graph Neural Networks, Graph Anomaly Detection, Fraud Detection.

1. Introduction

The contemporary digital economy thrives on interconnectedness, with financial, e-commerce, and social networks generating massive relational data streams. This digital expansion, however, inevitably invites adversarial exploitation of systemic vulnerabilities. Traditional fraud detection methodologies, which historically relied on isolated, rule-based systems or tabular machine learning algorithms, are increasingly inadequate in this landscape. These legacy systems evaluate entities in a vacuum, extracting behavioral embeddings from isolated sequences of actions while failing to account for the sophisticated, collusive, and networked strategies employed by modern fraudsters.

In response to this paradigm shift, research has pivoted toward Graph Learning (GL) and GNN-based fraud detection. By modeling relational data as graphs, these techniques effectively capture complex topological dependencies, making them uniquely suited for detecting sophisticated, multi-step fraudulent activities.

This comprehensive overview aims to synthesize the evolutionary trajectory of graph learning, the theoretical underpinnings of graph anomaly detection, and the specialized application of these domains to fraud detection. The analysis delineates how advancements in general graph representation learning have directly catalyzed breakthroughs in GAD. Furthermore, it unpacks the critical migration of GAD methodologies into the highly adversarial, high-stakes domain of fraud detection. A central thesis of this report is that fraud detection is not merely a subset of GAD; rather, it is a uniquely adversarial environment characterized by severe class imbalance, extreme feature heterophily, and active relational camouflage. By linking GL, GAD, and Fraud Detection, this analysis highlights the profound architectural modifications required to transition from general-purpose topological learning to robust, adversarial anomaly identification.

1.1. Related Work

The proliferation of graph-structured data has spawned numerous survey papers and systematic reviews over the past decade. Early foundational works provided taxonomies of graph-based anomaly detection, primarily focusing on traditional, non-deep learning techniques such as network embedding, random walks, and spectral graph clustering [1]. Subsequent literature expanded into the deep learning era, providing extensive surveys on the application of GNNs for general node, edge, and graph-level classification tasks, documenting the shift from shallow embeddings to deep message-passing architectures [2].

However, a critical gap remains in the literature: existing surveys either focus on general GNN mechanics or broad GAD categorizations, rarely providing a unified analysis that links foundational graph learning to specialized fraud detection architectures [3]. By treating fraud detection as a mere sub-topic, previous reviews overlook the essential architectural adaptations needed to handle adversarial behavior. Furthermore, emerging trends like LLM-based

multimodal reasoning are often addressed in isolation. This review fills this void by comprehensively charting the lineage from basic GL principles to cutting-edge, LLM-enhanced frameworks.

1.2. Contributions

This report provides a systematic, expert-level synthesis of the field by first mapping the theoretical continuum from general Graph Learning to specialized Fraud Detection, thereby highlighting the essential architectural adaptations at each evolutionary stage. It further standardizes the understanding of benchmark datasets by analyzing how inherent heterophily and extreme label imbalance dictate the design of specialized loss functions and sampling mechanisms. By formalizing the mathematical challenges of heterophily in fraud graphs, the review details how contemporary models transcend traditional homophily-based constraints, while simultaneously offering a nuanced critique of Graph Transformers—arguing that their reported superiority over classic GNNs is frequently a byproduct of suboptimal baseline tuning. Finally, the report explores the emergent frontier of LLM-enhanced paradigms, illustrating how semantic abstraction and multi-level enhancement resolve the deep-seated semantic-topological conflicts inherent in traditional graph vectorization.

2. Preliminaries

To establish a rigorous foundation for the subsequent analysis, it is necessary to define the core mathematical notations, structural definitions, and theoretical concepts governing graph-based anomaly and fraud detection.

Let a generic attributed graph be defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{X})$, where $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ represents the set of N nodes (representing entities such as users, financial accounts, or product reviews), and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges representing the structural interactions or relationships between these entities (such as monetary transfers, device sharing, or social connections). The topological structure of the graph is mathematically captured by the adjacency matrix $\boldsymbol{A} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$, where $\boldsymbol{A_{ij}} = 1$ if a direct edge exists between node v_i and v_j, and $\boldsymbol{A_{ij}} = 0$ otherwise. The attribute matrix is denoted as $\boldsymbol{X} \in \mathbb{R}^{\mathbb{N}\times\mathbb{d}}$, where each row $\boldsymbol{x_i} \in \mathbb{R}^{\mathbb{d}}$ corresponds to the d-dimensional raw feature vector of node v_i.

In the context of Graph Neural Networks (GNNs), the fundamental mechanism for learning representations is localized message passing, which is derived from the local Markov property on node features. For a standard spatial GNN layer l, the hidden representation $\boldsymbol{h_v^{(l)}}$ of a target node v is iteratively updated by aggregating messages from its local neighborhood $\mathcal{N}(v)$:

$$\boldsymbol{h_v^{(l)}} = \sigma\left(\boldsymbol{W^{(l)}} \cdot \mathrm{AGG}\left(\{\boldsymbol{h_u^{(l-1)}} : u \in \mathcal{N}(v) \cup \{v\}\}\right)\right) \quad (1)$$

Where $\boldsymbol{h_v^{(0)}} = \boldsymbol{x_v}$ represents the initial raw input features, represents a layer-specific learnable weight transformation matrix, $\boldsymbol{W^{(l)}}$ is a non-linear activation function (such as ReLU or LeakyReLU), and AGG is a permutation-invariant function (such as mean, sum, or max pooling) designed to compile the neighborhood context.

In Fraud Detection, the ultimate objective is to learn a discriminative mapping function $f: \mathcal{V} \to \{0,1\}$, or to assign a continuous anomaly score s_v to each node, where a higher score indicates a higher probability of the node being anomalous or fraudulent. This formulates fraud detection as a semi-supervised or fully-supervised node binary classification task. Crucially, this task is characterized by extreme label imbalance, as the set of anomalous nodes $\mathcal{V}_a$ is vastly outnumbered by the set of benign nodes $\mathcal{V}_b$ (i.e., $|\mathcal{V}_a| \ll |\mathcal{V}_b|$), rendering standard cross-entropy loss functions highly susceptible to majority-class overfitting [4].

A critical preliminary concept that dictates the success or failure of graph learning models in this domain is the Homophily Ratio. Homophily is a sociological principle stating that similar individuals tend to associate with one another. In graph representation learning, it quantifies the tendency of nodes to connect with nodes sharing the same ground-truth label [5]. The edge homophily ratio h for a graph is mathematically defined as the fraction of edges connecting nodes of the same class:

$$h = \frac{|\{(u,v) \in \mathcal{E} : y_u = y_v\}|}{|\mathcal{E}|} \quad (2)$$

Where y_u and y_u are the ground-truth labels of nodes u and v. Standard GNNs operate under the strong homophily assumption ($h \approx 1$), functioning essentially as spatial low-pass filters that smooth features across connected nodes. However, in fraud detection scenarios, graphs are inherently and strictly heterophilic ($h \ll 0.5$). Fraudsters actively deploy "camouflage" by connecting to benign users to mask their illicit activities, thereby generating a high volume of cross-class edges [6]. This mathematical reality fundamentally disrupts traditional message passing, necessitating the evolution of specialized graph anomaly detection frameworks.

3. Graph Anomaly Detection

3.1. Evolution of Graph Learning Paradigms

Graph learning represents a profound paradigm shift in machine learning, transitioning from the processing of independent and identically distributed (i.i.d.) Euclidean data to the modeling of complex, non-Euclidean topological dependencies. Before the widespread adoption of deep graph learning, algorithms relied heavily on manual heuristic feature engineering and shallow network embedding techniques such as DeepWalk [7] and Node2Vec [8]. These early methodologies, while revolutionary in their ability to map graph topologies into latent vector spaces via random walks, were inherently transductive. They struggled to scale to massive industry datasets and failed to seamlessly integrate rich, multidimensional node attributes with complex structural properties, limiting their effectiveness in dynamic anomaly detection scenarios. The introduction of Graph Neural Networks—most notably Graph Convolutional Networks (GCN) [9], Graph Attention Networks (GAT) [10], and GraphSAGE [11]—transformed the landscape by bridging deep learning with graph theory. GNNs leverage localized message passing to map nodes into low-dimensional continuous vector spaces while preserving both structural roles and feature similarities simultaneously.

In the context of Graph Anomaly Detection (GAD), these foundational graph learning architectures are both a breakthrough and a bottleneck. They are highly adept at encoding structural dependencies, which is paramount because an anomaly is often a contextual state rather than an absolute one. An entity's behavior (e.g., transaction volume) may appear perfectly normal in isolation but highly

anomalous when evaluated against the behavioral baseline of its local network community. However, the fundamental mechanism of standard GNNs—neighborhood feature smoothing—directly conflicts with the very nature of anomalies. Since anomalies deviate significantly from their reference objects by definition, iteratively averaging their features with benign neighbors dilutes their unique signatures. This phenomenon homogenizes the representations of normal and abnormal nodes, rendering anomalies invisible in the final embedding space. Consequently, the field of graph learning had to evolve specialized branches and objective functions to address GAD effectively.

3.2. Node-Level Detection Tasks

Graph Anomaly Detection operates across various structural granularities—including nodes, edges, subgraphs, and entire graphs. However, node-level detection remains the most prominent, complex, and thoroughly researched area, given its direct applicability to identifying malicious actors, spam accounts, and fraudulent entities in real-world systems. The evolution of GAD research has systematically categorized node-level anomalies into several distinct typologies based on the origin and nature of their irregularity:

Attribute Anomalies: These are nodes whose raw feature values (X_v) are statistical outliers compared to the global feature distribution of the entire graph, regardless of their structural position or connectivity.

Structural Anomalies: These anomalies manifest purely through topological irregularities. They often appear as densely connected subgraphs (fraud rings or cliques) or as bridge nodes inappropriately connecting disparate communities, deviating wildly from the standard structural patterns of the network.

4. GNN-based Fraud Detection

4.1. Benchmark Datasets

The rigorous evaluation and advancement of GNN-based fraud detectors rely heavily on standardized benchmark datasets. These datasets encapsulate organic, real-world anomalies across diverse domains, capturing complex relation types, severe feature heterophily, and varied scale distributions. Enriching the variety and scale of these datasets is fundamental, as synthetic anomalies often fail to replicate the sophisticated evasion tactics of real-world fraudsters. The statistical properties of the most prominent benchmark datasets utilized in state-of-the-art evaluations are listed in Table 1.

Table 1. Statistics of datasets

Dataset	Nodes	Edges	Fraud (%)
YelpChi	45,954	3,846,979	14.5%
Amazon	11,944	4,398,392	9.5%
Elliptic	203,769	234,355	9.8%
T-Finance	39,357	21,222,543	4.6%
DGraph-Fin	3,700,550	4,300,999	1.3%
T-Social	5,781,065	73,105,508	3.0%

These datasets expose the critical, defining challenges of fraud detection that separate it from standard GAD. For instance, the YelpChi [6] and Amazon [6] datasets are characterized by highly dense, multi-relational connections (e.g., users sharing the same IP address, users reviewing the same product within the same month). This high density exponentially increases multi-hop neighborhood sizes, leading to computational bottlenecks and severe over-smoothing. Conversely, massive industrial financial datasets like DGraph-Fin [12], T-Finance [13] and T-Social [13] introduce severe class imbalance—often featuring anomaly ratios well below 5%. This extreme imbalance causes traditional cross-entropy loss functions in standard GNNs to degenerate, forcing the model to over-optimize for the majority benign class while almost entirely ignoring the critical fraudulent minority [4].

4.2. Challenges of Graph Heterophily

The most pervasive and mathematically disruptive technical obstacle in graph-based fraud detection is heterophily. As defined in the preliminaries, homophily assumes that connected nodes share similar characteristics and ground-truth labels. The entire foundational architecture of standard GNNs (such as GCN and GraphSAGE) relies explicitly on this assumption; their aggregation functions act as spatial low-pass filters that homogenize representations to capture localized similarities [5].

However, empirical analysis reveals that fraud graphs inherently exhibit a mixed homophily-heterophily pattern [14]. Fraudsters are economically motivated to avoid connecting with other fraudsters, as dense clusters of malicious activity are easily detected by simple rule-based density algorithms. Instead, they strategically execute "camouflage attacks" by linking themselves to high-reputation, normal nodes—such as a fraudulent account conducting micro-transactions with legitimate merchants or a spam bot following verified celebrity accounts. This deliberate relational engineering artificially inflates their trustworthiness and avoids triggering security tripwires, resulting in highly heterophilic edges.

Consequently, when a standard GNN aggregates information across these camouflaged edges, the distinct, anomalous feature signals of the fraudster are heavily diluted by the overwhelming volume of benign neighbor features. Through iterative message passing layers, the fraudster's latent representation becomes virtually indistinguishable from that of a normal user, completely neutralizing the model's discriminative power.

4.3. Graph Learning Methodologies in Fraud Detection

To resolve the profound inadequacies of standard GNNs in the face of heterophily, relational camouflage, and extreme label imbalance, researchers have engineered a suite of specialized, highly discriminative graph learning architectures explicitly tailored for the fraud detection domain. To address the inherent complexities of fraud data, DGA-GNN [15] introduces decision tree binning and recursive feedback grouping to resolve feature non-additivity and neighbor distinguishability. By discretizing continuous attributes—such as account age—and segregating aggregation paths for suspected fraudsters and benign users, this approach prevents detrimental cross-class smoothing and preserves feature discriminability. Complementing these structural adaptations, frameworks like PC-GNN [16] and CARE-GNN [6] optimize performance in imbalanced, multi-relational environments through adaptive sampling and reinforcement learning; while PC-GNN employs subgraph sampling to mitigate minority-class gradient domination,

CARE-GNN utilizes RL agents to dynamically filter camouflaged connections, thereby maximizing the signal-to-noise ratio before GNN aggregation.

5. Advanced Insights and Emerging Trends

As digital financial networks continue to scale into the billions of nodes and fraudsters adopt increasingly complex, long-range strategies, the reliance on purely structural, k-hop spatial message-passing has revealed computational bottlenecks and semantic limitations. This realization has driven the field toward more advanced, highly parameterized architectures—namely the adaptation of Graph Transformers (GTs) and the emergent integration of Large Language Model (LLM)-enhanced graph networks.

5.1. Graph Transformers

Graph Transformers represent a major paradigm shift aimed at overcoming the intrinsic mathematical limitations of traditional GNNs, specifically the phenomena of "over-squashing" and limited structural expressivity. Over-squashing occurs when a GNN attempts to compress exponentially growing multi-hop neighborhood information into a fixed-size vector at the target node, leading to severe information loss. By leveraging global self-attention mechanisms originally developed for natural language processing, GTs can theoretically capture long-range, higher-order dependencies across the entire graph without being strictly bound by the localized topological constraints of spatial GNNs [17], and enhancing performance on heterophilic graphs [18].

5.1.1. Applications in Fraud Detection

In the realm of financial fraud detection, GTs have demonstrated substantial promise. Modern fraud schemes are rarely isolated incidents; they involve complex, multi-step transaction trajectories—such as money laundering fan-out patterns, where an originator sends illicit funds through layers of intermediate accounts to multiple beneficiaries. Tracking these long-range topological paths is critical, yet notoriously difficult for standard GNNs limited to 2 or 3 layers of aggregation.

An example of this application is FraudGT [19], a specialized graph transformer explicitly engineered for large-scale financial transaction graphs. A critical insight driving FraudGT is that unlike standard social networks where node profiles contain the most information, financial graphs store the vast majority of crucial forensic data (transaction amount, currency type, micro-second timestamps) within the edge features. FraudGT addresses this by deploying edge-based message passing gates that actively regulate the flow of information, selectively passing only the most critical transactional features to neighboring nodes while discarding noise. Concurrently, it employs an edge attribute-based attention bias, allowing the global self-attention mechanism to heavily weight metadata indicative of fraud.

5.1.2. Reassessing Performance: Graph Transformers vs. Optimized GNNs

Though advancements in graph representation learning have seen a surge in the adoption of GTs, emerging empirical evidence suggests that the efficacy of GTs may be significantly overstated.

In a comprehensive hyperparameter study, Luo et al. [20] demonstrate that classic message-passing GNNs with modest but systematic tuning can match or outperform state-of-the-art GTs on the vast majority of node-classification benchmarks, calling into question claims of GT universal superiority. At the same time, Zhou et al. [21] show that much of the practical advantage attributed to GTs depends on their multi-head self-attention (MHA) component, which can introduce global noise by ignoring local topology and be computationally prohibitive on large graphs. Replacing MHA with carefully designed propagation/transformation (P/T) combinations plus an effective FFN often recovers or improves node-level accuracy. Furthermore, recent comprehensive benchmarking reveals that even simple shallow methods, such as tree ensembles (e.g., XGBoost, Random Forest) integrated with basic neighborhood aggregation, consistently outperform state-of-the-art GTs in supervised graph anomaly detection [22]. These shallow, tree-based methods naturally handle the miscellaneous, heterogeneous tabular features common in fraud datasets by executing robust feature-space partitioning to construct complex, disjoint decision boundaries.

Taken together, these findings support a cautious stance: Graph Transformers are a powerful tool, but their superiority is not universal. Carefully optimized spatial aggregation, robust feature-space partitioning (or simple ensemble methods), and attention to dataset-specific heterogeneity often yield equal or better results with lower complexity.

5.2. Large Language Model Enhanced Fraud Detection

While GNNs and GTs excel at topological pattern recognition and numerical feature processing, they fundamentally operate in a modality-isolated regime. In real-world fraud detection—such as identifying fake e-commerce reviews, malicious social media bots, or phishing inducements—nodes are heavily attributed with unstructured, complex semantic text (e.g., user bios, detailed review text, transaction memos) [23]. Traditional graph pipelines process this multimodal data by passing it through static, early-stage textual encoders (e.g., Word2Vec, initial BERT embeddings) to generate fixed, low-dimensional numerical vectors before feeding them into the GNN. This early vectorization causes catastrophic information loss, destroying the nuanced semantic cues, sentiment variations, and contextual reasoning that distinguish a legitimate user from a camouflaged fraudster.

To bridge the fundamental gap between discrete topological structures and continuous multimodal semantics, the field is rapidly adopting Graph-Enhanced Large Language Models (LLMs). These cutting-edge frameworks utilize LLMs not merely as auxiliary feature extractors, but as core reasoning engines capable of directly resolving semantic-topological conflicts. MLED [24] pioneers the integration of external semantic knowledge into GNN pipelines through its type-level and relation-level enhancers, which utilize frozen LLMs to widen the contextual gap between fraudsters and benign entities while dynamically weighting edge importance through multimodal fusion. Complementing this, DGP [25] framework addresses the challenges of information overload and signal dilution by feeding LLMs structured prompts that preserve the fine-grained text of target nodes while applying bi-level semantic abstraction to neighbors. By further utilizing Markov diffusion kernels to prune irrelevant metapaths and statistical aggregation to condense numerical features, DGP resolves the inherent conflicts between

massive graph scale and finite LLM context windows, ensuring high-fidelity reasoning in complex fraud networks. FLAG [23] utilizes semantic similarity neighbor sampling to dynamically prune camouflaged structural neighbors that lack semantic alignment with the target node. In the industrial domain, TransactionGPT [26], a specialized 3D-Transformer foundation model pre-trained on billions of real-world financial records, represents the frontier of holistically understanding consumer behavior and semantic transaction trajectories without relying solely on traditional network topologies. These models collectively signify a decisive transition in the field from purely structural anomaly detection to holistic, multimodal semantic reasoning.

6. Conclusions

6.1. Summary

The application of Graph Neural Networks to Fraud Detection represents a critical, highly specialized evolution from general Graph Learning and standard Graph Anomaly Detection. While early GL models successfully captured complex structural dependencies in benign datasets, their inherent reliance on homophily and spatial feature smoothing rendered them highly vulnerable to the adversarial, evasive nature of fraud. Fraudsters actively deploy abnormality camouflage and engineer dense heterophilic connections, systematically breaking standard message-passing paradigms.

In response, the field has developed highly sophisticated discriminative GNNs. Furthermore, while Graph Transformers theoretically offer superior long-range dependency tracking via global attention, rigorous benchmarking reveals that their supremacy is frequently overstated. When classic GNNs or simple tree ensembles with neighborhood aggregation are subjected to rigorous hyperparameter tuning, they routinely match or exceed the performance of GTs at a fraction of the computational cost. Ultimately, the most profound recent leap in the field is the integration of Large Language Models. By fusing deep semantic abstraction with topological context through frameworks like MLED and DGP, modern systems can effectively counter neighborhood camouflage and solve the catastrophic information loss associated with early-stage vectorization, setting a new standard for adversarial graph learning.

6.2. Future Directions

Looking forward, the trajectory of GNN-based fraud detection is defined by several highly promising research avenues that seek to address current limitations in scalability, generalization, and interpretability:

Foundation Models and Zero-Shot Generalization: Current fraud detectors are largely domain-specific, requiring extensive retraining and struggling with out-of-distribution (OOD) generalization when deployed on new platforms. Future research must prioritize Cross-Scenario GAD Foundation Models capable of zero-shot detection. By leveraging advanced transfer learning and invariant feature representations, models could be pre-trained on massive, multi-domain graphs and deployed successfully on unseen, label-scarce target datasets without catastrophic degradation.

Explainable and Causal GAD: The inherent "black-box" nature of deep graph learning severely limits its adoption in highly regulated financial sectors, where auditors require explicit justification for flagged accounts. Future models must transition from identifying simple correlations to establishing causation. Frameworks like CaT-GNN [27] point toward a future where models not only output probabilistic anomaly scores but provide mathematically rigorous, human-readable causal graphs detailing exactly why a specific transaction trajectory was deemed fraudulent.

Agentic LLM Frameworks: Expanding upon prompt-based models like DGP, future systems will likely deploy multi-agent LLM frameworks dedicated specifically to the early-stage vectorization phase. Rather than autonomously navigating the entire graph topology, specialized AI agents could be deployed to collaboratively process complex, multimodal raw node attributes before any message passing occurs. By interacting and synthesizing this multi-modal evidence, these agents would generate highly expressive, aligned semantic vectors or structured prompts prior to graph aggregation. This agentic preprocessing fundamentally resolves the information loss inherent in traditional vectorization, providing downstream discriminative models with a significantly richer, context-aware foundation for structural fraud detection.

Acknowledgment

We would like to thank the developers and maintainers of essential graph anomaly detection libraries and benchmarking frameworks, such as PyGOD, DGFraud, and the GADBench project. The availability of these open-source resources, along with foundational frameworks like the Deep Graph Library (DGL) and PyTorch Geometric, has been instrumental in standardizing evaluations, enabling fair comparisons, and accelerating advanced research in graph-based fraud detection.

References

[1] L. Akoglu, H. Tong, and D. Koutra, “Graph based anomaly detection and description: A survey,” Data Mining and Knowledge Discovery, vol. 29, no. 3, pp. 626–688, May 2015.

[2] H. Qiao, H. Tong, B. An, I. King, C. C. Aggarwal, and G. Pang, “Deep graph anomaly detection: A survey and new perspectives,” IEEE Transactions on Knowledge and Data Engineering, vol. 37, no. 9, pp. 5106–5126, Sept. 2025.

[3] D. Cheng, Y. Zou, S. Xiang, et al., “Graph neural networks for financial fraud detection: A review,” Frontiers of Computer Science, vol. 19, no. 9, Art. no. 199609, 2025.

[4] F. Xu, N. Wang, H. Wu, et al., “Revisiting graph-based fraud detection in sight of heterophily and spectrum,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, no. 8, 2024, pp. 9214–9222.

[5] J. Zhu, Y. Yan, L. Zhao, et al., “Beyond homophily in graph neural networks: Current limitations and effective designs,” in Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 7793–7804.

[6] Y. Dou, et al., “Enhancing graph neural network-based fraud detectors against camouflaged fraudsters,” in Proc. 29th ACM Int. Conf. Information and Knowledge Management (CIKM), 2020.

[7] B. Perozzi, R. Al-Rfou, and S. Skiena, “DeepWalk: Online learning of social representations,” in Proc. 20th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), 2014, pp. 701–710.

[8] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” in Proc. 22nd ACM SIGKDD Int. Conf.

Knowledge Discovery and Data Mining (KDD), 2016, pp. 855–864.

[9] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv preprint arXiv:1609.02907, 2016.

[10] P. Veličković, et al., "Graph attention networks," arXiv preprint arXiv:1710.10903, 2017.

[11] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[12] X. Huang, Y. Yang, Y. Wang, C. Wang, Z. Zhang, J. Xu, L. Chen, and M. Vazirgiannis, "DGraph: A large-scale financial dataset for graph anomaly detection," in Proc. Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track, 2022.

[13] J. Tang, J. Li, Z. Gao, and J. Li, "Rethinking graph neural networks for anomaly detection," arXiv preprint arXiv:2205.15508, 2022.

[14] Y. Liu, J. Cheng, J. Li, et al., "A pre-training and adaptive fine-tuning framework for graph anomaly detection," arXiv preprint arXiv:2504.14250, 2025.

[15] M. Duan, T. Zheng, Y. Gao, et al., "DGA-GNN: Dynamic grouping aggregation GNN for fraud detection," in Proc. AAAI Conf. Artificial Intelligence, vol. 38, no. 10, 2024, pp. 11820–11828.

[16] Y. Liu, et al., "Pick and choose: A GNN-based imbalanced learning approach for fraud detection," in Proc. Web Conf. (WWW), 2021.

[17] Y. Shi, Z. Huang, S. Feng, H. Zhong, W. Wang, and Y. Sun, "Masked label prediction: Unified message passing model for semi-supervised classification," in Proc. Int. Joint Conf. Artificial Intelligence (IJCAI), 2021.

[18] L. Müller, M. Galkin, C. Morris, et al., "Attending to graph transformers," Transactions on Machine Learning Research, 2024.

[19] J. Lin, X. Guo, Y. Zhu, et al., "FraudGT: A simple, effective, and efficient graph transformer for financial fraud detection," in Proc. 5th ACM Int. Conf. AI in Finance (ICAIF), 2024, pp. 292–300.

[20] Y. Luo, L. Shi, and X.-M. Wu, "Classic GNNs are strong baselines: Reassessing GNNs for node classification," Advances in Neural Information Processing Systems, vol. 37, pp. 97650–97669, 2024.

[21] J. Zhou, X. Chen, C. Xie, S. Yu, Q. Xuan, and X. Yang, "Rethinking graph transformer architecture design for node classification," arXiv preprint, Oct. 2024.

[22] J. Tang, F. Hua, Z. Gao, P. Zhao, and J. Li, "GADBench: Revisiting and benchmarking supervised graph anomaly detection," in Proc. NeurIPS Track on Datasets and Benchmarks, 2023.

[23] C. Yang, H. Liu, D. Wang, et al., "FLAG: Fraud detection with LLM-enhanced graph neural network," in Proc. 31st ACM SIGKDD Conf. Knowledge Discovery and Data Mining (KDD), vol. 2, 2025, pp. 5150–5160.

[24] T. Huang, Y. Wang, Q. Li, et al., "Can LLMs find fraudsters? Multi-level LLM enhanced graph fraud detection," in Proc. 33rd ACM Int. Conf. Multimedia (MM), 2025, pp. 1530–1538.

[25] Y. Li, J. Hu, B. Hooi, et al., "DGP: A dual-granularity prompting framework for fraud detection with graph-enhanced LLMs," arXiv preprint arXiv:2507.21653, 2025.

[26] Y. Dou, Z. Jiang, T. Zhang, et al., "TransactionGPT," arXiv preprint arXiv:2511.08939, 2025.

[27] Y. Duan, G. Zhang, S. Wang, et al., "CAT-GNN: Enhancing credit card fraud detection via causal temporal graph neural networks," arXiv preprint arXiv:2402.14708, 2024.