Graph-Learning-Empowered Financial Fraud Detection: Progress and Future Directions

Enxia Li[1], Mengshi Chen[2], Sheng Xiang[1*], and Ling Chen[1]

[1]Australian Artificial Intelligence Institute, University of Technology Sydney, Sydney, Australia

[2]Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

[*]Address correspondence to: sheng.xiang@uts.edu.au

Abstract

Financial fraud detection is an important task in ensuring the integrity and security of financial systems. In recent years, it has been shown that graph learning, which utilizes the relational structure of data, can considerably enhance the detection of fraudulent financial activity by accurately modeling the complex patterns and relationships inherent in financial transactions. In this work, we provide a comprehensive survey on the emerging application of graph learning techniques in detecting and combating financial fraud that can serve as a guidepost for researchers and practitioners interested in leveraging the power of graph learning to create a safer, more secure financial environment. Specifically, we start by introducing the fundamental concepts of graph learning, outlining their unique advantages over traditional machine learning techniques in the context of financial fraud detection. It is worth mentioning that graph learning techniques enable end-to-end training from relational data input to fraud prediction, eliminating the need for additional feature engineering. We then delve into a systematic review of the recent advancements and methodologies in applying graph learning to various financial fraud scenarios, such as credit card fraud, insurance fraud, and money laundering. Furthermore, we provide unique insights regarding several critical challenges, such as data privacy, scalability, and the dynamic nature of financial networks, that are faced when implementing graph learning models in real-world financial ecosystems. We show that practical applications of graph learning still suffer from computational complexity and lack of interpretability, and we offer a forward-looking perspective on potential research directions and improvements that can boost the effectiveness of graph learning applied to financial fraud detection.

Introduction

The considerable losses resulting from financial fraud have consistently drawn the focus of academia, industry, and regulatory communities. As discussed, fighting financial fraud is beneficial for sustainable economic growth, which is one of the United Nations' sustainable development goals (SDGs) [1, 2]. For instance, online payment systems are critical in combating climate change, given that they can help to cut 400 million metric tons of annual carbon emissions by reducing the use of physical currency. However, fraudulent behaviors targeting online payments and credit card transactions have inflicted financial harm upon online payment users and hindered the promotion of online payment [3]. In the financial area, fraud behavior is widely modeled and discussed. For instance, Tergiman and Villeval experimentally studied the nature of lies and designed the 'Announcement Game' to evaluate how the introduction of reputation affects lies [4]. Also, Liu et al. studied the involvement of AI in credit scoring for loan lenders and found indications that AI techniques will benefit different aspects of lending, including fraud detection [5]. In addition, researchers have studied theoretical models in specific scenarios that are easily affected by fraud behaviors, such as mobile app ranking fraud [6–9]. Designing efficient and effective financial fraud detection methods can reduce the operating costs of service providers, protect the property of bank users, and help to establish a more sustainable finance system.

Due to the importance of financial fraud detection, a number of methods have been proposed in existing literature, falling into 2 broad categories: rule-based approaches and machine-learning-based approaches. Generally speaking, rule-based methods rely on a variety of rules generated by domain experts to identify suspicious behaviors. Though rule-based approaches are usually highly efficient and accurate in some scenarios, it is difficult for experts to identify proper rules when the context is complicated and dynamic. Moreover, fraudsters can learn the underlying logic of rules and then develop corresponding strategies to fool the system. In recent years, many machine-learning-based methods, such as logistic regression, decision tree, and support vector machine, have been developed to train models from historical data. To better capture the interrelation among the financial data (e.g., transactions) and identify potential fraud behaviors, graphs have been widely used to model financial activities. Moreover, in addition to the traditional machine learning models, deep learning techniques have been applied in various applications, such as anti–money laundering and credit card fraud detection, and achieved outstanding performance, thanks to the advance of graph machine learning.

Motivations

Traditional isolated fraud events have evolved into intricate networks of fraudulent activities. The sophistication of these activities, coupled with the massive and scattered nature of financial data, poses considerable challenges to existing fraud detection methods. In recent years, graph learning has emerged as a potential solution for addressing financial fraud effectively. By modeling entities as nodes and relationships as edges, graph learning techniques can capture complex patterns and relationships in the data, thus they are particularly suited for detecting organized fraud activities. Therefore, it is worthwhile to discuss the challenges in the use of graph learning in financial fraud detection, such as the sensitivity and complexity of graph data, as well as the interpretability and robustness of models based on graph learning, with some promising research directions in this important area.

Contributions

The main contributions of this work are summarized as follows.

1. A comprehensive analysis of applications of graph learning in the field of financial fraud detection is presented, highlighting their advantages over traditional approaches. We stress the novelty and adaptability of introducing graph learning in fraud detection and how it can improve traditional machine-learning-based and rule-based methods.

2. Financial fraud detection methods based on graph learning, including unsupervised, semisupervised, and supervised techniques, are categorized. The mechanisms and applications of each method are discussed in detail to show their advantages and availability in different scenarios.

3. Future research directions for financial fraud detection based on graph learning are discussed according to their desired features, including their dependability, scalability, and robustness, which are eagerly anticipated in modern financial fraud research trends.

Roadmap

In the next section, we provide the background of the problem of financial fraud detection and graph learning techniques, as well as the categorization of existing approaches. Then we introduce representative techniques in unsupervised approaches, semisupervised approaches, and supervised approaches. Finally, we highlight the remaining challenges and future research directions and share our conclusions.

Background

In this section, we first introduce the background of machine learning approaches for financial fraud detection, with a focus on graph-empowered approaches. We then introduce fundamental concepts of graph learning. Finally, we provide the classification of the introduced techniques.

Financial fraud detection via machine learning

Recent research in financial fraud detection can be broadly categorized into rule-based methodologies and machine-learning-based methodologies. For instance, Seeja and Zareapoor proposed a rule-based method for extracting association-based clues for detecting fraud [10]. Balagolla et al. proposed a blockchain-based approach to prevent fraudulent transactions [11]. Weinmann et al. utilized trace data to detect fraud by analyzing movements captured via a computer mouse [12]. Many machine-learning-based approaches have been proposed to help detect financial fraud. For instance, Fiore et al. studied feature extraction using neural networks and developed supervised classifiers to identify 362 fraudulent transactions [13]. Xiao et al. introduced a nonlinear optimization model to generate black-box attacks and evaluate the performance of different machine-learning–based credit card fraud detection models [14]. Xu et al. extracted additional categories of behavioral features from peer-to-peer lending transaction data, aiming to enhance various performance metrics in fraud detection [15]. Additionally, anomaly detection via machine learning techniques can also provide insights for fraud detection in the financial area [16–20]. In recent years, advanced machine learning techniques, such as graph representation learning, have shown superior performance in financial fraud detection [21, 22]. Therefore, existing machine-learning-based fraud detection solutions can be classified into graph-based and non-graph-based approaches. For instance, Fu et al. in 2016 studied the usage of automatic feature engineering in a convolutional neural network (CNN) [23]. Recently, financial data have been modeled by graphs, and graph neural network techniques have been deployed to detect financial fraud behaviors. For example, in 2020, Cheng et al. proposed a graph neural network model via spatiotemporal attention mechanism for the detection of fraudulent credit card transactions [24], and in 2021, Jing et al. proposed an effective learning method based on graph neural networks for credit card fraud detection with few data samples [25].

Graph empowered financial fraud detection

The realm of financial systems has witnessed profound changes in both structure and complexity over the past few decades. An important driver of this evolution is the intricate interconnectivity established between financial entities, transactions, and activities. Graph theory, a branch of mathematics dealing with networks of interconnected nodes and edges, offers a robust framework to represent and analyze this complex web of financial interdependencies. Research in graph theory and its application has been very active. For example, Zhou et al. proposed a novel "reduce-solve-combine" search strategy that integrates a problem reduction mechanism to effectively tackle the common challenge of identifying critical nodes [26]. The surge in financial malpractice has necessitated the exploration of advanced techniques, such as graph neural networks (GNNs), to identify and combat fraudulent activities within these interconnected financial systems.

The following are 2 representative financial applications that can greatly benefit from the modeling and analytical capabilities of graph structures:

- Networked-loan fraud detection: In a financial ecosystem, the manifestation of loans often entails a series of transactions and relationships between various entities, such as banks, borrowers, guarantors, companies, and intermediary organizations. By constructing a graph wherein entities are represented as nodes and transactions or relationships as edges, one can effectively capture the patterns and anomalies associated with potential fraudulent loan activities. For instance, a closed-loop of entities lending to each other without clear business justification could indicate a circular lending scheme, a common strategy to illicitly inflate assets or siphon off funds. As reported, companies participating in a real-world loan network are dynamic and complex [27]. Graph learning models, with their capacity to consider relational dependencies, offer a promising avenue for detecting such sophisticated networked-loan frauds that might remain elusive to conventional detection techniques.

- Group-based money-laundering detection: Money laundering, the act of disguising the origins of illegally obtained money, typically involves intricate transactional networks to obfuscate the trail of illicit funds. Group-based money laundering involves coordinated efforts by a consortium of individuals or entities to launder money in a manner that evades detection.

Representing financial transactions as a graph enables the identification of suspicious clusters or subgraphs where the patterns of transactions deviate from legitimate behavior. Graph learning models can be trained to recognize these anomalous patterns and to capture the relational dependencies between entities, thereby spotlighting potential laundering groups as validated [21].

In summary, as financial frauds grow in complexity and cunning, the development of methods that leverage the topological and relational properties of graphs, in conjunction with the power of graph learning models, is a promising frontier in the fight against these malicious activities. This survey aims to delve deeper into the methodologies and breakthroughs surrounding this interdisciplinary confluence of finance and graph neural networks.

Table 1: Categories, models, applications, and datasets of graph learning for financial fraud detection.

Category	Model	Application	Dataset(s)
Unsupervised	FlowScope [28]	Anti–money laundering	(see cited paper)
	AntiBenford [29]	Anti–money laundering	Ethereum
Semisupervised	GTAN [22]	Credit card fraud detection	Amazon, YelpChi, other (see cited paper)
	Federated metalearning [30]	Credit card fraud detection	(see cited paper)
	Semi-GNN [31]	Credit card fraud detection	AliPay
	InfDetect [32]	Insurance fraud detection	Ant Financial Services
	Scalable graph learning [33]	Anti–money laundering	(see cited paper)
	GraphSense [34]	Anti–money laundering	Bitcoin ransomware
Supervised	Tem-GNN [35]	Credit/loan risk assessment	AliPay
	HGAR [36]	Credit/loan risk assessment	Unspecified commercial bank in Asia
	DGANN [37]	Credit/loan risk assessment	Unspecified Korean payment service provider
	ST-GNN [38]	Loan default analysis	AliPay
	AMG-DP [39]	Loan default analysis	Ant Credit Pay
	Network learning [40]	Insurance fraud detection	Ant Finance
	PC-GNN [41]	Credit card fraud detection	Amazon, YelpChi
	MAHINDER [42]	Credit card fraud detection	AliPay
	HACUD [43]	Credit card fraud detection	Ant Credit Pay
	GAGNN [21]	Anti–money laundering	Amazon, YelpChi, UnionPay

Scope

This review focuses on the application of graph learning techniques in the domain of financial fraud detection. In order to structure our exploration and understand the evolving landscape of fraud detection methodologies based on graph learning, we adopt a taxonomy based on the nature of the learning paradigms employed, as shown in Table 1.

- Unsupervised graph learning techniques: Techniques in this category do not rely on labeled datasets for training. They aim to capture intrinsic patterns and anomalies within financial transaction graphs. By identifying deviations from normative behaviors, these models present avenues for uncovering novel, previously unidentified fraudulent strategies.

- Semisupervised graph learning techniques: Situated between the supervised and unsupervised paradigms, these techniques employ a blend of labeled and unlabeled data. Their primary advantage lies in their ability to leverage vast amounts of unlabeled data while utilizing a limited amount of labeled instances to guide the learning process. They often produce more generalizable models that can detect nuanced fraudulent patterns.

- Supervised graph learning techniques: This category encompasses methodologies where graph-based models are trained with labeled data comprising instances of both genuine and fraudulent financial activities. By leveraging the annotated datasets, these models aim to learn patterns and relationships indicative of fraud and to offer high precision in detecting malicious activities.

While the aforementioned taxonomy provides a clear and concise categorization, it is worth noting that the landscape of graph learning applications in fraud detection is vast and multifaceted. Several alternative taxonomies can be devised to characterize the diverse methodologies in this area. For instance:

- Based on graph construction: It is possible to categorize methods depending on how their graphs are constructed (e.g., static vs. dynamic, weighted vs. unweighted, or multilayered vs. single-layered).

- By fraud domain application: Techniques could also be grouped based on the specific financial fraud applications they target (e.g., credit card fraud, insider trading, or digital currency scams).

- Considering graph learning framework: Categorization could be based on the underlying graph learning framework (e.g., GraphSAGE, DeepWalk, or hybrid methods with tree-based machine learning methods).

As we venture into the depths of methodologies for financial fraud detection based on graph learning, it is paramount to acknowledge the diversity and richness of the techniques and their underlying principles. This survey, while emphasizing the supervised vs. semisupervised vs. unsupervised taxonomy, will touch upon the aforementioned categories, aiming to offer a comprehensive and holistic overview of the field.

Unsupervised Approaches

Overview

The unsupervised paradigm of graph learning techniques for financial fraud detection primarily focuses on the discovery of underlying patterns without the need for labeled data. By leveraging intrinsic properties and anomalies within financial transaction graphs, unsupervised graph learning methods offer the potential to unearth novel and sophisticated fraud strategies that might elude supervised techniques.

A crucial application benefiting from these unsupervised techniques is anti–money laundering, where the absence of predefined fraud patterns demands adaptable and evolving detection mechanisms. Although in this review, the unsupervised approaches are for anti–money laundering, they can be applied to fraud detection and risk assessment as well.

Anti–money laundering

In this section, 2 related works on anti–money laundering are discussed [28, 29]. Both of them introduce new algorithms to detect abnormal financial activities in graph data. They both introduce novel criteria into the task and are more efficient than all other existing methods.

FlowScope [28]

This paper addresses the detection of money laundering in bank transfer networks, highlighting the lack of flow tracking in current methods and inadequate theoretical guarantees. It introduces the FlowScope algorithm, which models the money transfer network as a graph $G = (V, E)$, where $V = X \cup W \cup Y$ represents the set of accounts. Here, W denotes the inner bank accounts, while X and Y correspond to outer accounts receiving net inflows and outflows, respectively. Each edge $e_{ij} \in E$ represents the aggregate monetary transfer from account v_i to v_j. The algorithm tracks transaction flows from source to destination, focusing on large deposits, internal transfers, and withdrawals, which are harder for fraudsters to hide. FlowScope also aims to predict intermediary layers in laundering schemes, using a maximum transaction step limit to mitigate risks. A novel density-based metric in FlowScope ensures efficient detection of dense money flows and caps the amount fraudsters can transfer. The algorithm proves more effective than existing benchmarks in various tests, and has a computational complexity nearly linear with respect to transaction volume.

- Anomalousness of ML for k-partite subgraph:

  Anomalous flow is characterized as flow from a node set A, through single or multiple layers of intermediate accounts M, to a different node set C. For an intermediary account v_i, let $f_i(S)$ and $q_i(S)$ denote the lower and upper bounds of the cumulative weighted out-degree and in-degree with respect to node group $S \subseteq V$. The anomalousness metric is defined as:

  $$g_k(S) = \frac{1}{|S|} \sum_{l=1}^{k-2} \sum_{v_i \in M_l} f_i(S) - \lambda(q_i(S) - f_i(S)), \tag{1}$$

  or equivalently,

  $$g_k(S) = \frac{1}{|S|} \sum_{l=1}^{k-2} \sum_{v_i \in M_l} (1+\lambda) f_i(S) - \lambda q_i(S), \quad k \geq 3, \tag{2}$$

  where λ (the imbalance cost rate) quantifies the adversity fraudsters encounter per unit of retention or shortfall, representing the cost of concealment.

- FlowScope algorithm:

  The FlowScope algorithm seeks the optimal subset S that maximizes $g(S)$. Initially, it constructs a priority hierarchy for nodes in S, where each node v_i is assigned a weight or priority, defined as:

$$w_i(S) = \begin{cases} f_i(S) - \frac{\lambda}{1+\lambda} q_i(S), & \text{if } v_i \in M_l \\ d_i(S), & \text{if } v_i \in A \cup C, \end{cases} \tag{3}$$

where $d_i(S)$ denotes a specified degree function for boundary accounts in $A \cup C$.

AntiBenford subgraph framework [29]

This paper proposes the AntiBenford subgraph framework, which aims to detect abnormal subgraphs in financial networks. Its goal is to identify a group of nodes within a large transaction or financial graph that considerably diverges from Benford's law, a principle that describes the expected distribution of the first digit of each number among data. The framework utilizes a dense subgraph discovery algorithm to identify these anomalous subgraphs, which are characterized by both their edge density and their deviation from Benford's law. With adequate experiments, the AntiBenford subgraph framework was proven to be able to detect anomalies in real-world networks that other anomaly detection strategies based on graphs may not notice.

- Anomaly Score:

  For each vertex $v \in V$ in a financial network $G = (V, E)$, an anomaly score $s(v)$ is defined using a χ^2 score compared to Benford's distribution:

  $$s(v) = \Sigma_{d=1}^{9} \frac{(X_d^u - \mathbb{E}(X_d^u))^2}{\mathbb{E}(X_d^u)}, \tag{4}$$

  where X_d^u is the number of transactions, or edges in the graph, related to u whose first digit is d.

- AntiBenford Subgraph Detection:

  The AntiBenford subgraph detection algorithm is based on the above anomaly score. First, the score of each vertex is computed. Then edge $e(u, v)$ is weighted using function $f(u, v) = \sqrt{s(u) \cdot s(v)}$. Then, most of the densest subgraph problem is solved by minimizing the objective $max_{S \subset V} \frac{\Sigma f(u,v)}{|S|}$, which represents a subgraph with large average weight for each edge. Another objective $max_{S \subset V} \frac{e(S) + \Sigma s(v)}{|S|}$ is utilized to avoid including vertices that have low scores themselves but have higher degree towards an anomalous set of nodes.

Semisupervised Approaches

Overview

Semisupervised graph learning approaches in the domain of financial fraud detection leverage the strength of both labeled and unlabeled data to uncover and predict fraudulent activities. This hybrid approach provides a robust framework for scenarios where labeled data is limited, yet unlabeled transactions are abundant. By combining knowledge of known instances of fraud with the intrinsic patterns present in the broader dataset, semisupervised graph learning techniques have demonstrated remarkable accuracy and adaptability. Three prominent applications that have substantially benefited from semisupervised graph learning techniques are:

- Credit card fraud detection: In this domain, the rapid influx of transaction data often results in a scenario where only a fraction of the data is labeled, typically the confirmed fraud cases. Semisupervised methods can bridge this gap, effectively using the limited labeled data to guide the model's learning from the vast swaths of unlabeled transactions, identifying potentially suspicious activities that might have otherwise been overlooked.
- Insurance fraud detection: Insurance fraud encompasses a wide range of illicit activities, from exaggerated claims to false information on policy applications. As in other domains, labeled instances of insurance fraud are scarce compared to the volume of overall claims and applications. Semisupervised approaches adeptly analyze patterns within this largely unlabeled data, helping to identify fraudulent claims or applications by learning from the smaller number of known fraud cases. This approach is particularly effective in adapting to new methods of fraud that constantly emerge in the insurance sector.
- Anti–money laundering (AML): Money laundering schemes often manifest as intricate patterns that subtly deviate from normative behaviors. While there might be known instances or patterns of money laundering, the ever-evolving tactics used by perpetrators demand a more flexible detection mechanism. Here, semisupervised graph learning techniques prove invaluable, given that they can extrapolate from known laundering patterns to detect newer, subtler schemes in the broader, mostly unlabeled dataset.

The versatility and adaptability of semisupervised graph learning techniques, which cater to diverse applications and constantly evolving threats, make them a cornerstone in the modern arsenal against financial fraud.

Credit card fraud detection

Here we introduce Semi-GNN [31], federated metalearning [30], and the gated temporal attention network (GTAN) [22]. Semi-GNN uses graph-attention networks for financial fraud detection. GTAN achieved the best performance in both semisupervised and supervised fraud detection tasks, according to the experimental results reported in [22].

Semi-GNN [31]

Current financial fraud detection methods such as neural networks [44] and SVM [45] mainly use machine learning to identify fraud patterns, focusing on users' statistical features from profiles, behaviors, and transactions. However, these methods often overlook user interactions, which are crucial in financial settings. Traditional machine learning approaches, which use models such as logistic regression or neural networks, fail to consider these interactions and the rich data they offer, such as social connections and merchant transactions. Additionally, most fraud detection data is unlabeled, a challenge not fully addressed by existing graph-based methods, which also lack interpretability. Semi-GNN addresses these issues by leveraging both labeled and unlabeled data for fraud detection, bridging supervised and unsupervised learning. It confronts the challenge of integrating sparse labeled data with abundant unlabeled data to enhance detection performance using novel techniques.

- Low-level view-specific user embedding:

$$h_k^u = \sigma\left(\sum_{i \in N_v^u} a_{ui}^{(k)} W^{(k)} x_i\right), \tag{5}$$

  where N_v^u is the set of neighbors of user u in the v-th view-specific graph, $a_{ui}^{(k)}$ is the attention coefficient between user u and neighbor i in the k-th layer, $W^{(k)}$ is the weight matrix in the k-th layer, and x_i is the input feature of neighbor i. The function σ is a nonlinear activation function.

- High-level view-specific user embedding:

$$h_L^u = \sigma\left(\sum_{k=1}^{L} \phi_k h_k^u\right), \tag{6}$$

  where L is the number of layers, h_k^u is the low-level embedding of user u in the k-th layer, and ϕ_k is the layer-specific weight. The function σ is a nonlinear activation function.

Federated metalearning [30]

The publicly available datasets for fraudulent credit card detection are imbalanced and limited. Approximately 2% of all credit card transactions are associated with fraudulent activities. Owing to concerns related to data security and privacy, individual banks are typically unwilling to share their transaction datasets with one another. Moreover, traditional techniques [23, 46] neglect the exploration of the relationships among training samples. They focus on the comparison with samples from different classes.

To tackle these problems, the paper proposes a novel metalearning-based model for the task of detecting fraudulent credit card activities. This model leverages a federated learning strategy. This approach enables various banks to collectively train a shared model while maintaining the privacy of their imbalanced training data, which is stored within their individual private databases. The proposed approach demonstrates notably improved performance in comparison to state-of-the-art methods when evaluated on 4 publicly available datasets. The paper also introduces a novel metalearning-based classifier to effectively learn discriminative feature extraction on unseen class samples. This strategy delivers remarkably superior performance beyond other metalearning-based methods.

- Feature extraction model: The paper uses a “K-tuplet” (K-tuple) network utilizing ResNet-34 as the feature extraction tool. The output of this model is considered to be network features.

The feature extraction model, denoted as f_{network}, produces feature maps to represent the feature extraction function.

- Relation model: The relation model, denoted as J_{relation}, depicts the similarity between the support set $D_{c_{\text{support}}}$ and the query set $D_{c_{\text{query}}}$. The relation model takes the network features, applies a sigmoid function, and outputs similarity scores.

- Federated metalearning framework: The paper applies federated learning to facilitate the collaborative sharing of datasets among different banks, enabling the construction of an efficient fraud detection model without compromising the privacy of each bank's customers. The objective loss function is defined as:

$$\min l(x_i; y_i; w), \quad \text{where} \quad l(x_i; y_i; w) = \frac{1}{n_c} \sum_{i \in D_i} L_{c_{\text{relation}}}(x_{i_c}; y_{i_c}; w). \tag{7}$$

  Here, $n_{\text{whole}} = \sum_{c=1}^{C} n_c$ represents all the data samples involved in the whole process. $L_{c_{\text{relation}}}$ is the loss function for the relation model.

- K-tuple network: An enhanced metric learning approach, referred to as the deep K-tuple network, has been developed. This network extends the capabilities of the triplet network by enabling simultaneous comparisons with K negative samples within each minibatch.

- Federated learning: The paper uses a horizontal federated learning framework, which is used in scenarios where datasets have similar feature space but differ in samples. The server initializes all parameters of the fraud detection model. At each communication round t, a random fraction seed of banks is selected. These banks establish direct communication with the server, retrieve the latest global model parameters, and compute the average gradient of the loss function f_c using their respective private datasets at the current fraud detection model parameters, employing a fixed learning rate η.

GTAN [22]

In this work, the authors note that in real-world applications, only a small portion of transactions are labeled as fraudulent or nonfraudulent. This lack of labeled data makes it difficult to train effective machine-learning models for fraud detection.

To address this challenge, the authors propose a semisupervised learning approach that leverages both labeled and unlabeled data. Their method, called attribute-driven graph representation learning, uses the attributes of transactions to construct a graph and then applies graph representation learning to capture the complex relationships between transactions. By employing this approach, the model can learn from both labeled and unlabeled data, thereby enhancing its capacity to identify fraudulent transactions.

The model is designed to handle the challenge of limited labeled data in credit card fraud detection by leveraging labeled and unlabeled data in a semisupervised learning approach. The technical contributions can be summarized as follows:

- Attribute embedding and feature learning: The authors introduce a novel approach for transforming the numerical attributes of each record into a tensor format, symbolized as $X_{num} \in \mathbb{R}^{N\times d}$. Here, N is the total count of transactions, and d represents the dimensions of the features. They adopt distinct attribute embedding layers to formulate categorical attributes into a separate matrix $X_{cat} \in \mathbb{R}^{N\times ds}$. These embeddings are then merged to form a comprehensive categorical representation of each transaction, computed by add-pooling as $x^{(u)}cat = \sum ix^{(u)}cat, i$, where i encompasses various categories such as card, transaction, and merchant. The notation $x^{(u)}cat \in \mathbb{R}^{1\times d}$ signifies the aggregated category embedding vector for each transaction record indexed by u.

- Gated temporal attention networks: The authors utilize a sequence of transaction embeddings, denoted as $X = \{x_{t0}, x_{t1}, ...x_{tn}\}$, to capture the temporal dynamics of each transaction. This is achieved by integrating both categorical and numerical attributes as inputs to the gated temporal attention network (GTAN). The formulation is given by $x_{ti} = x^{(ti)}_{num} + x^{(ti)}_{cat}$. Initially, at the first temporal graph attention layer, $H_0 = X$ is set as the base input embedding matrix. They apply a multihead attention mechanism to evaluate the importance of each neighboring transaction and accordingly update the embeddings, enhancing the model's ability to discern relevant temporal patterns.

- Risk embedding and propagation: The authors suggest embedding the partially observable risk attributes (or labels) into a unified feature space to align them with other node features. This process involves creating risk embedding vectors for nodes with labels and assigning zero vectors for those without. The combined node features, now inclusive of risk embeddings, are represented as $x_{ti} = x^{(ti)}_{num} + x^{(ti)}_{cat} + \hat{y}^{(t1)}W_r$, with W_r denoting the tunable parameters for risk embedding. This method allows for a more nuanced incorporation of risk factors into the overall feature set, potentially enhancing the model's predictive accuracy.

- Fraud risk prediction techniques: Upon aggregating the transaction embeddings, the authors utilize a sophisticated 2-layer multi-layer perceptron (MLP) to estimate fraud risk. This process is articulated as $\hat{y} = \sigma(\mathrm{PReLU}(HW_0 + b_0)W_1 + b_1)$, where $\hat{y} \in \mathbb{R}^{N\times 1}$ encapsulates the predicted risk levels for all transactions. The parameters W and b are adjustable within the MLP, allowing the model to fine-tune its predictions based on the learned embeddings. This advanced prediction methodology is expected to yield more accurate and reliable fraud risk assessments.

Insurance fraud detection

InfDetect [32]

The InfDetect paper introduces a large-scale, graph-based fraud detection system for e-commerce insurance that leverages advanced graph learning techniques. This approach employs both supervised and unsupervised graph learning algorithms, including DeepWalk (unsupervised), graph neural networks (supervised and semisupervised node classification), and DistRep (supervised edge classification), to analyze complex fraud patterns in e-commerce transactions.

Various graph types, such as device-sharing, transaction, and friendship graphs, are analyzed. The system preprocesses these graphs to optimize classification performance by removing isolated accounts. Additionally, the data processing phase involves feature collection and processing, employing techniques such as data scaling, categorical feature encoding, and denoising autoencoder (DAE) for unsupervised feature transformation.

The system's primary objective is to identify fraudsters in the claim stage by classifying accounts or orders as fraudulent. The methodology is particularly effective in detecting fraudster gangs using graph learning techniques. The paper includes both qualitative and quantitative evaluations, demonstrating the system's enhanced fraud detection capabilities over traditional rule-based methods. This novel approach contributes substantially to the field of fraud detection in e-commerce, showcasing the power of graph learning in uncovering and analyzing fraudulent activities.

Anti–money laundering

This subsection delves into two sophisticated anti–money laundering (AML) methods: Scalable graph learning and GraphSense. Scalable graph learning employs graph convolutional neural networks to analyze complex financial networks, tackling challenges in large, noisy datasets. It emphasizes graph compression and neural network scalability. GraphSense, on the other hand, analyzes Bitcoin transactions related to ransomware using graph-based methodologies, focusing on transaction and cluster analysis in the Bitcoin blockchain. Both methods highlight innovative approaches to combating financial crimes through advanced data analysis techniques.

Scalable graph learning [33]

As to AML, there is a "needle-in-a-haystack" problem of entity classification and hidden pattern discovery in extensive, constantly changing, high-dimensional, time-series transaction datasets with high noise-to-signal ratios, combinatorial complexity, and nonlinearity. Datasets frequently exhibit fragmentation, inaccuracy, incompleteness, and inconsistency, both within individual organizations and across them. Automating the synthesis of information from diverse data streams proves to be a formidable challenge, often necessitating the involvement of human analysts who may have limited resources at their disposal.

Moreover, The challenge of AML involves dealing with vast, high-dimensional graph data that map billions of relationships (edges) among millions of entities (nodes). In the realm of AML transaction monitoring, a node entity can represent either an individual account or a collection of linked accounts whose connections are already established or inferred through clustering. The known attributes encompass explicitly defined data, constituting information that has been specifically collected through standard "know your customer" procedures or multimodal data extracted from public or partner information sources. This category also includes observable transactions and any accompanying flags or suspicious activity reports that have been filed. These challenges highlight the complexity and scale of the problem, as well as the societal impact of failing to effectively address it.

The technical contribution of this paper is centered around the use of scalable graph convolutional neural networks for forensic analysis of financial data in AML. The paper discusses some key concepts and results:

- Graph compression with Ligra+: The researchers applied a graph compression tool called Ligra+ and conducted experiments on both a simulated AML graph and several deep learning benchmark graphs. Ligra+ serves as a compression system that aims to reduce space usage while maintaining competitive or improved performance compared to running algorithms on uncompressed graphs using multicore machines. The researchers managed to obtain a compression ratio of up to 2 ×, although compressing larger graphs presents greater challenges. This development creates the potential for memory-efficient training and inference in deep neural networks by harnessing compression and reordering techniques.

- Scalable graph convolutional neural networks: The authors offer an initial exploration of scalable graph convolutional neural networks for the forensic analysis of vast, dense, and ever-changing financial data. They present preliminary experimental findings based on a sizable synthetic graph generated by their proprietary data simulator, known as AMLSim, featuring 1 million nodes and 9 million edges.

- High efficiency: The paper explores prospects for achieving high efficiency in both computation and memory usage, and it also provides findings from a straightforward experiment involving graph compression. The outcomes of the study substantiate the initial hypothesis that employing graph deep learning for AML holds considerable potential in combating illicit financial practices.

GraphSense [34]

The GraphSense paper introduces a novel methodology for analyzing Bitcoin transactions related to ransomware using advanced graph learning techniques. This approach extends beyond traditional clustering heuristics and is applied to 35 different ransomware families. Key to this methodology are 2 different types of network representations over the entire Bitcoin blockchain: the address graph and the cluster graph. The address graph represents each Bitcoin address as a vertex and each transaction as a directed edge, facilitating the computation of summary statistics, such as transaction numbers and value flows. The cluster graph partitions addresses into maximal subsets or clusters, likely controlled by the same real-world actor, using the multiple-input clustering heuristic. This method contributes substantially to dataset expansion, identifying numerous Bitcoin addresses associated with ransomware attacks and tracing their outgoing relationships to key addresses. Furthermore, it assesses the minimum direct financial impact of each ransomware family by analyzing monetary flows to these key addresses. Implemented on the open-source GraphSense platform, this approach enables comprehensive transaction extraction and analysis, offering new insights into the ransomware economy and enhancing the understanding of its financial implications through graph analysis.

Supervised Approaches

Overview

Supervised graph learning techniques for financial fraud detection employ labeled data, which comprises instances of both genuine and fraudulent financial activities, to train models. These labeled datasets serve as foundational knowledge, allowing graph learning models to learn intricate patterns and relationships indicative of fraud and other financial risks. The strength of supervised graph learning techniques lies in their ability to achieve high precision in detecting and categorizing financial activities based on past labeled examples.

The following applications represent notable areas where supervised graph learning techniques have been extensively applied.

- Credit/loan risk assessment: Supervised graph learning techniques are used to evaluate the potential risk associated with lending money to individuals and institutions. By analyzing past borrowing behavior and related financial activities, these models can predict the likelihood of timely loan repayment, aiding financial institutions in making informed lending decisions.
- Loan default analysis: In this domain, supervised models analyze transactional data and financial histories to predict the probability of a borrower defaulting on a loan. Such insights can guide lenders in devising risk mitigation strategies and shaping their lending policies.
- Insurance fraud detection: Insurance claims come with diverse data points, from personal histories to event details. Supervised graph learning techniques parse this data to identify patterns that might indicate fraudulent insurance claims, ensuring that insurers can validate claims with higher accuracy.
- Credit card fraud detection: Given the vast amounts of daily credit card transactions, detecting fraudulent activities requires models that can rapidly and accurately identify suspicious patterns. Supervised graph learning techniques, trained on historical instances of fraud, excel in pinpointing such activities amidst a sea of legitimate transactions.
- Anti–money laundering: While semisupervised methods are crucial for AML due to evolving money laundering tactics, supervised graph learning techniques play a vital role when there is a substantial amount of labeled data on known laundering techniques. These models can detect and flag transactions that align with known illicit money movement strategies, acting as a first line of defense.

The application of supervised graph learning techniques spans a broad spectrum of financial domains, offering tools that are both precise and responsive to the intricacies of each application.

Credit/loan risk assessment

In this subsection, we delve into 3 cutting-edge methodologies. First, Tem-GNN [35] is notably adept at extracting credit risk insights from dynamic graphs. The high-order graph attention representation (HGAR) [36] approach employs a higher-order graph attention mechanism to discern the potential of loan defaults. Lastly, DGANN [37] introduces an end-to-end dynamic graph neural network for predicting risks associated with networked loans.

Tem-GNN [35]

The authors highlight 3 challenges in credit risk prediction from temporal graphs: time interval irregularity, integration of structural and temporal information, and static and short-term temporal factors. In response, the Tem-GNN model integrates a pathway for static factors, short-term graph encoders, and a time-series model. Key features include:

- Static feature learning model: Using an MLP, the model extracts stable user features and projects them into a high-level space.
- Short-term graph encoder: This encoder focuses on recent neighbor information through a graph convolution module.
- Temporal Model: The model incorporates an interval-decayed attention mechanism defined as:

$$\alpha'(t) = \frac{\exp(s'(h_t))}{\sum_{i\in\{0,1,\ldots,T\}} \exp(s'(h_{ti}))}, \tag{8}$$

  where $s'(h_t)$ is the output of a single-layer feed-forward neural network with a tanh activation function. To handle interval irregularity:

$$g(\delta_t) = \frac{1}{\log(e + \delta_t)}, \tag{9}$$

  where δ_t is the time interval between the event t and now.
- User embedding and risk score calculation: The method utilizes temporal attention to define the user embedding and an MLP for risk score prediction.

HGAR [36]

The authors of this work note that companies are allowed to guarantee each other to enhance loan security, but when more and more companies are involved, they form complex guarantee networks. These complex structures can be a double-edged sword for the economy, especially during economic downturns, when many companies may fail to repay loans, causing a domino effect of defaults.

To address this challenge, the paper proposes a high-order graph attention representation method (HGAR) for the default risk assessment of networked guarantee loans. The main contributions of HGAR can be summarized as follows:

- Dual roles of nodes: Nodes in the guarantee network have dual roles, guarantor and guarantee.
- Graph attentional layer: This layer captures node importance in a network.
- High-order adjacent approximation: This approximation preserves dual node roles and high-order adjacency.

- Loan default prediction layer: This layer uses both first-order and high-order attentional representation for default prediction.

DGANN [37]

Loans from commercial banks form complex directed-network structures, thus an adaptive strategy to efficiently identify and address any systematic crises is necessary. Therefore, the authors propose an end-to-end dynamic graph-based attention neural network (DGANN) to predict risk guarantee relationships among lenders by learning on the interconnected loans in a network. Their approach can be split into 3 parts:

- Graph convolution network (GCN) with structure attention: Uses temporal guarantee networks for high-order graph representation.
- Graph recurrent network (GRN) with temporal attention: Extracts edge attributes and learns temporal patterns. The model updates the hidden state as:

$$r_t = \sigma(W_r * [h_{t-1}, e_t]), \tag{10}$$

$$z_t = \sigma(W_z * [h_{t-1}, e_t]), \tag{11}$$

$$\hat{e}_h = \tanh(W_e * [r_t \odot h_{t-1}, e_t]), and \tag{12}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{e}_h, \tag{13}$$

  where r_t and z_t denote the reset and update gates of the tth object respectively, $\hat{e}_h$ represents the representation of the candidate hidden layer, and W are the weights dynamically updated during the model training phase.

- Prediction layer: Estimates risk probability with a global view of guarantee networks.

Loan default analysis

In this subsection, we delve into 2 methods for loan default prediction. Formally, the loan default prediction task falls under detection of node anomalies and edge anomalies in the loan network. The ST-GNN [38] targets small and medium-sized enterprises (SMEs), addressing data deficiency issues often seen with online financial institutions. By adeptly extracting genuine supply chain relationships from the SME graph, this approach enhances loan default prediction accuracy. Conversely, AMG-DP [39] emphasizes the intricate properties intrinsic to loan default prediction: communicability, complementation, and induction. By leveraging relation-specific layers and attention mechanisms, AMG-DP ensures precise predictions, even for new users with sparse data.

ST-GNN [38]

This research addresses data scarcity in financial risk analysis for SMEs, particularly in online financial institutions such as Ant Financial and WeBank, where acquiring credit-related SME data is challenging.

Key concerns include extracting authentic supply chain relationships from the SME graph and using these relationships in financial risk analysis. The solution involves 2 components:

- Spatial-temporal aware graph neural network (ST-GNN): Designed to capture both the local graph structure and temporal dynamics of the SME graph, ST-GNN is trained in a semisupervised manner for link prediction, and focuses on identifying genuine supply chain links. It assigns confidence scores to each edge, refining the SME graph into a supply-chain graph by removing low-confidence edges.

- Loan default prediction module: Utilizing the refined supply-chain graph, this module aggregates data from neighboring SMEs into a target SME, enabling more accurate financial risk assessments. The method incorporates ST-GNN for supervised node classification to predict future loan repayment failures of SMEs.

In summary, ST-GNN efficiently extracts supply chain relationships and predicts loan defaults for SMEs, offering a novel approach in the realm of financial risk analysis for SMEs.

AMG-DP [39]

There are 3 key intrinsic properties of the problem of loan default prediction:

- Communicability: The patterns of loan default behaviors are discerned through the examination of the local network structure.

- Complementation: Describing the profile of a financial defaulter is intricate and challenging to capture comprehensively with just one source of information, particularly for new users who have limited prior behavioral data.

- Induction: In practical financial scenarios, new users emerge on a daily basis, necessitating models to retain the capacity to make predictions for this category of users who possess only limited profile information.

Previous solutions have been ineffective in harnessing multiplex relationships in financial settings, thus have neglected these crucial inherent features of loan default detection. The authors propose a novel attributed multiplex graph-based loan default prediction approach (AMG-DP) to address these issues. The authors also acknowledge that the factors leading a user to fail in meeting required repayments are intricate and that expressing such multiplex information is challenging with feature-based approaches. The AMG-DP model is designed with relation-specific receptive layers that utilize an adaptive breadth function to integrate vital information obtained from the local structure in each facet of the AMG. Additionally, the model incorporates multiple propagation layers to investigate higher-order connectivity information. It leverages a relation-specific attention mechanism to highlight pertinent information throughout the end-to-end training process.

The adaptive fusion function is implemented with a relation-specific attention mechanism, which is formulated as follows:

$$\alpha_{u,r} = \frac{\exp(\gamma_r \tanh(W_\gamma h_u + b_\gamma))}{\sum_{r' \in R} \exp(\gamma_{r'} \tanh(W_\gamma h_u + b_\gamma))}, \tag{14}$$

where $\{W_\gamma, b_\gamma, \gamma_r\}$ form the parameter set of the adaptive fusion function. The function $\tanh(\cdot)$ is used as the activation function.

The relation-specific receptive layer is formulated as follows:

$$\alpha_{u,i} = \frac{\exp(v_r^T \tanh(W_r^{(l)}[h_u^{(l-1)}||h_i^{(l-1)}||e_{u,i}^{(l-1)}]))}{\sum_{(u,j) \in E_r} \exp(v_r^T \tanh(W_r^{(l)}[h_u^{(l-1)}||h_j^{(l-1)}||e_{u,j}^{(l-1)}]))}. \tag{15}$$

The final user representation is calculated as follows:

$$z_u = \sum_{r \in R} \alpha_{u,r} \cdot h_{u,r}. \tag{16}$$

The model has been evaluated on a large-scale real-world dataset and has shown effectiveness compared with state-of-the-art models. The model also has the inductive ability to predict whether new users are likely to default on loans by aggregating their neighbors with multiplex relations.

Insurance fraud detection

Network learning [40]

This paper focuses on the detection and prevention of fraudulent insurance claims, particularly those made by organized groups of fraudsters. This issue is important in the context of Alibaba's return-freight insurance, which faces numerous potentially fraudulent claims daily. Deliberate misuse of the insurance policy can result in substantial financial losses. The paper outlines 3 key challenges in fraud detection:

- Concept drift: This term describes the situation in which new forms of fraud emerge over time and become progressively less predictable. This primarily happens because fraud detection systems use changing features that do not stay constant.

- Label uncertainty: In the past, the rule-based fraud detection system assigned a risk tag to each account, but it is uncertain if the accounts labeled as 'no observable risk' are genuinely risk-free or not.

- Excessive human effort: Traditional insurance fraud detection settings require considerable human effort for labeling and evaluation tasks. The authors aim to focus on automated risk control that requires negligible human effort, except for a periodical evaluation conducted by insurance professionals.

Importantly, to uncover organized fraudsters and capture complex relations among colluding fraudsters, the paper constructs 3 types of graphs: the device-sharing graph, the transaction graph, and the friendship graph.

- Device-sharing graph: This graph reveals the relations between accounts that share a device. It consists of 2 types of vertices: device vertices (representing user machine IDs, UMIDs) and account vertices. Edges exist between device vertices and account vertices, indicating the log-in activities in the history. The device-sharing graph captures the device-sharing relationships among fraudsters and regular customers.

- Transaction graph: This graph represents the fund exchange relations between accounts. It consists of account vertices and edges that indicate the existence of established transactions between accounts. The transaction graph provides information about financial interactions between fraudsters and regular customers.

- Friendship graph: This graph is built based on the friendship relationships at Alipay, a product of Ant Financial that has social networking features. It captures the social connections between accounts. The friendship graph can reveal social networks and connections among fraudsters and regular customers.

The paper compares and analyzes these graphs to determine which is most suitable for fraud detection. Specifically, it focuses on the device-sharing graph, which exhibits contrasting patterns between colluding fraudsters and regular customers.

Credit card fraud detection

Researchers have presented techniques to enhance detection rates for credit card fraud detection. Among these are the pick-and-choose GNN (PC-GNN) [41] approach, which seeks to address the class imbalance through a supervised GNN-based approach; the MAHINDER [42] model, a unique tool that capitalizes on multi-view attributed heterogeneous information networks to detect defaulters; and HACUD [43], a dedicated method that utilizes a hierarchical attention mechanism for cash-out user detection in an attributed heterogeneous information network.

PC-GNN [41]

The primary challenge tackled by the PC-GNN paper is the issue of imbalanced classes in graph-based fraud detection. This problem arises due to the considerable disparity between the number of instances of the majority class (nonfraudulent entities) and the minority class (fraudulent entities). This imbalance often leads to poor performance of detection algorithms, particularly for the minority class (i.e., the fraudsters), which is more important.

The paper illustrates 3 key challenges in creating graph neural networks for fraud detection that grapple with class imbalance:

- Redundant link information: Fraudsters often employ deceptive tactics, such as camouflaging their activities by generating misleading information, to make identifying them more challenging. For example, spammers may utilize legitimate accounts to post spammy reviews, creating numerous connections between the spam reviews and genuine users, effectively concealing their true intentions within the pool of legitimate users.

- Lack of necessary link information: Perpetrators of fraud usually avoid trading with each other to escape being caught together. For instance, in financial situations, individuals engaging in fraud would steer clear of conducting transactions with each other to evade joint detection. Therefore, there might not exist a connection between the 2 nodes, which can negatively impact the performance of methods based on GNNs.

- Dilution of minority class features: This difficulty arises during the process of gathering messages in GNNs, and can result in the dilution of features associated with the minority class. In imbalanced settings, it is common for most of the neighbors connected to a central node to belong to the majority class. Consequently, the characteristics of fraudulent neighbors can be easily neglected, and predictions can be heavily influenced by the benign majority.

To address these challenges, the paper proposes a graph-learning-based imbalanced learning approach for graph-based fraud detection. A label-balanced sampler is designed to select nodes and edges for training. The probability assigned for each node is inversely proportional to its label frequency. This ensures that nodes from the minority class are more likely to be chosen. To address application challenges, the paper suggests using a neighborhood sampler to select neighbors based on a distance function with adjustable parameters. To focus on the fraud-labeled nodes, unnecessary links can be removed by selecting neighbors that are distant from the target based on distance measurements. Meanwhile, essential links that contribute to fraud prediction can be established by selecting nodes similar to the fraud class and considering them as neighbors. Specifically, the 3 steps of the PC-GNN framework can be described as follows:

- Pick: The central nodes are chosen using a balanced sampler to create a well-balanced subgraph for training in minibatches. This step helps to ensure that the minority class is sufficiently represented in the training process.

- Choose: With a distance function that can be adjusted using parameters, the neighborhood of the minority class is sampled more than required, and that of the majority class is sampled less. This step helps to ensure that the model is exposed to a balanced representation of the classes during training.

- Aggregate: Messages from chosen neighbors and various relationships are combined to obtain the ultimate representations of the target. This step helps to ensure that the model captures the relevant information from the graph structure for the prediction task.

MAHINDER [42]

The problem of detecting default users on online payment service platforms has at least one of these 2 challenges: (a) It is challenging to capture the inherent aspects of default users, and there is an urgent need for a more accurate approach to modeling user profiles. (b) Criminals engaging in illicit activities might intentionally create intricate behaviors, such as moving money between numerous users and attempting to lengthen the money transfer path to evade regulations. This adversarial nature of financial default makes it challenging to detect.

It is important to mention that the current task involves identifying users at risk of default using multiview data from diverse information networks. This particular approach has not been explored previously.

Therefore, the authors proposed a novel approach called MAHINDER (Multiview Attributed Heterogeneous Information Network based financial DEfault useR detection) to address these challenges. MAHINDER uses multiple user relationships to construct a multiview attributed heterogeneous information network (MAHIN) for better user profile modeling. It also takes into account the local structures of assigned metapaths to extract detailed behavioral patterns at a finer level. The main contribution of MAHINDER can be summarized as follows:

- Problem Statement: The issue of identifying financial defaulters is structured as a binary classification problem within a multiview attributed heterogeneous information network (MAHIN). Given a MAHIN $G = \{V, E, X_V, X_E\}$ consisting of m specific views, the purpose is to detect defaulters from the target user set $U \subseteq V$. Each user $u \in U$ is assigned a label $y_u \in \{0, 1\}$ to indicate whether they are a defaulter or not.

- The MAHINDER Model: The MAHINDER model is proposed to address the challenges of detecting default users. The model uses multiple user relationships to construct an MAHIN for better user profile modeling. The approach also takes into account the nearby patterns of assigned metapaths to uncover detailed behavioral trends. The model uses metapaths from different perspectives to comprehensively model user profiles. It also includes a path encoder based on metapaths to grasp the local structural patterns present in both nodes and links. Attention mechanisms are used at various levels, including nodes, links, and metapaths, to automatically determine the importance of different elements. The MAHINDER model includes attention mechanisms to meticulously model the selected paths. There is an additional attention layer applied to different metapaths to filter out irrelevant perspectives.

HACUD [43]

The authors address the challenge of detecting cash-out fraud in credit payment services, an important fraud problem in financial services. Cash-out fraud involves users seeking cash gains through illegal or insincere means. The authors state that conventional detection methods rely mainly on the statistical features of users and often overlook the interaction relations between them. They propose modeling the problem as a classification task in an attributed heterogeneous information network (AHIN) and introduce the hierarchical attention-based cash-out user detection (HACUD) model, which uses metapath-based neighbors and a hierarchical attention mechanism.

The HACUD model is detailed as follows:

- Feature transformation: The original user features are transformed into latent representations. The transformation function is:

$$f' = \sigma(Wf + b), \tag{17}$$

  where W is the weight matrix, b is the bias vector, σ is the activation function (ReLU is used), and f is the original feature vector.

- Neighbor feature fusion: The latent representations of a user and their neighbors, derived from each metapath, are combined. The fusion function $g(\cdot, \cdot)$ can be concatenation, addition, or element-wise product.

- Hierarchical attention mechanism: Different users may have different preferences over features based on metapaths and attribute information. The attention weight of metapath ρ for user u is:

$$\beta_{u,\rho} = \frac{\exp(z_\rho^T f''_{uC})}{\sum \exp(z_\rho^T f''_{uC})}, \tag{18}$$

  where z_ρ is the importance of metapath ρ and f''_{uC} is the representation of neighbors for user u based on metapath ρ.

- Model learning: The model aims to minimize the loss function:

$$L(\Theta) = \sum (y_u \log(p_u) + (1 - y_u) \log(1 - p_u)) + \lambda ||\Theta||_2^2, \tag{19}$$

  where y_u and p_u are the ground truth and predicted cash-out probability of user u, respectively, Θ is the model's parameter set, and λ is the regularizer parameter. The model is trained using SGD or one of its variants.

Anti–money laundering

GAGNN [21]

The main challenge addressed in this paper is the detection of money laundering activities in large-scale transaction networks. Traditional methods for detecting money laundering often fail to capture the complex and evolving patterns of these activities. Moreover, they often suffer from high false-positive rates and are unable to effectively leverage the rich relational information in transaction networks.

To address these challenges, the authors propose the group-aware graph neural network (GAGNN) as part of their novel deep graph learning framework, which is designed to capture group-level money laundering patterns. The framework is capable of learning the complex and evolving patterns of money laundering activities, reducing false positives, and effectively leveraging the rich relational information in transaction networks.

The technical contributions of this paper can be summarized as follows:

- Group-aware graph neural network: The authors propose a novel graph neural network model, GAGNN, which is designed to capture the group structure in financial transactions. The model is defined by the following equation:

$$h_v^{(l+1)} = \sigma\left(W^{(l)} \cdot \text{CONCAT}\left(h_v^{(l)}, \frac{1}{|\mathcal{N}(v)|} \sum_{u \in \mathcal{N}(v)} h_u^{(l)}, \frac{1}{|\mathcal{G}(v)|} \sum_{g \in \mathcal{G}(v)} h_g^{(l)}\right)\right), \tag{20}$$

  where $h_v^{(l+1)}$ is the hidden state of node v at layer $l + 1$, σ is the activation function, $W^{(l)}$

is the weight matrix at layer l, CONCAT is the concatenation operation, $\mathcal{N}(v)$ is the set of neighbors of node v, $\mathcal{G}(v)$ is the set of groups that node v belongs to, $h_u^{(l)}$ is the hidden state of node u at layer l, and $h_g^{(l)}$ is the hidden state of group g at layer l.

- Group embedding: The authors introduce a method to learn the embedding of groups. The group embedding is computed as the average of the embeddings of the nodes in the group. The group embedding is updated during the training process. The group embedding is defined by the following equation:

$$h_g^{(l+1)} = \frac{1}{|\mathcal{V}(g)|} \sum_{v \in \mathcal{V}(g)} h_v^{(l+1)}, \tag{21}$$

  where $h_g^{(l+1)}$ is the hidden state of group g at layer $l+1$, $\mathcal{V}(g)$ is the set of nodes in group g, and $h_v^{(l+1)}$ is the hidden state of node v at layer $l+1$.

- Group-aware loss function: The authors propose a group-aware loss function to train the GAGNN model. The loss function encourages the model to assign similar labels to the nodes in the same group. The group-aware loss function is defined by the following equation:

$$\mathcal{L} = \mathcal{L}_{\text{node}} + \lambda \mathcal{L}_{\text{group}}, \tag{22}$$

  where $\mathcal{L}_{\text{node}}$ is the node classification loss, $\mathcal{L}_{\text{group}}$ is the group consistency loss, and λ is a hyperparameter to balance the 2 terms.

Discussion and Future Work

Rise of financial fraud gangs

Traditionally, financial fraud was often committed by individuals or small groups acting in isolation. However, with the advancement of technology and the increasing sophistication of fraudsters, we are now seeing the emergence of large-scale, organized financial fraud gangs. These gangs often have complex structures and employ advanced techniques, making their activities harder to detect and prevent. Although similar issue has been noticed and discussed, there is still a need for further research on detection algorithms [47].

To tackle this challenge, future research could focus on developing more sophisticated detection algorithms that are capable of identifying complex patterns and relationships indicative of organized fraud. This development could involve the use of advanced machine learning techniques, including deep learning and graph neural networks, which are capable of modeling complex structures and relationships. Additionally, cooperation between financial institutions and regulatory bodies could be enhanced to share information and intelligence about suspected fraud gangs.

Sensitivity and complexity of financial data

Graph data, representing entities (e.g., individuals, accounts) as nodes and relationships (e.g., transactions, connections) as edges, is increasingly being used in financial fraud detection. However, this type of data is often sensitive (due to privacy concerns) and complex (due to the large number of nodes and edges and the potential for complex relationships). This situation has been hindering traditional fraud detection, according to Sangers et al. [48]. Privacy and security are highly expected in graph data when detecting financial fraud.

Future work could focus on developing methods for effectively and securely handling sensitive graph data. This could involve the use of privacy-preserving techniques, such as differential privacy or federated learning, which allow for the analysis of sensitive data while minimizing the risk of privacy breaches. Additionally, new methods for handling the complexity of graph data, such as scalable graph processing algorithms or graph simplification techniques, can be developed.

Interpretability and robustness of GNN-based models

Graph learning strategies such as the use of GNN-based models have shown great promise in financial fraud detection due to their ability to model complex relationships in graph data. However, these models often lack interpretability, meaning it can be difficult to understand why they have made a particular prediction. Additionally, they may not be robust to changes in the data or to adversarial attacks. Much effort has been made in related research by using attention mechanisms and designing feature extraction methods [49, 50]. At the same time, financial fraud detection calls for interpretability due to the complexity of the data.

To address the interpretability challenge, future work can focus on developing methods for explaining the predictions of graph learning models. This can involve the use of techniques such as local interpretable model-agnostic explanations (LIME) or Shapley additive explanations (SHAP), which provide insights into the contribution of each feature to a model's prediction. To improve the robustness of graph learning models, we can investigate methods for making these models more resistant to changes in the data or to adversarial attacks. These goals can be achieved by involving techniques such as adversarial training or robust optimization. Moreover, real-world graph data are usually dynamic and time-evolving. Dynamic GNNs [51, 52] can help model temporal dependencies in financial networks.

Conclusion

In this review, we have provided a comprehensive overview of the current landscape of financial fraud detection, with a particular focus on the use of graph learning techniques. We have traced the evolution of financial fraud from isolated events to complex, organized activities and discussed how this has necessitated a shift in detection methods. Then, we delved into the use of graph learning methods for financial fraud detection. We discussed how these methods, by modeling entities as nodes and relationships as edges, are able to capture the complex patterns and relationships inherent in financial fraud activities. We highlighted the advantages of graph learning methods over traditional methods, including their ability to handle heterogeneous data and their capacity for learning high-level features. However, we also acknowledged that the use of graph learning methods in financial fraud detection is not without its challenges. In the final part of our survey, we discussed these challenges in detail, including the sensitivity and complexity of graph data and the issues of interpretability and robustness in models based on graph learning. We highlighted the need for methods that can handle sensitive graph data securely and efficiently, as well as techniques that can improve the interpretability and robustness of graph learning methods.

In conclusion, while graph learning methods hold great promise for financial fraud detection, there are still many open research problems. Future research can focus on addressing the challenges we have identified, as well as exploring new methods and techniques for leveraging the power of graph learning methods. We hope that this survey will serve as a valuable resource for researchers and practitioners in this field and inspire further work in the fight against financial fraud.

Acknowledgments

Author Contributions

E.L. led the literature collection and drafted the manuscript. S.X. and M.C. assisted in the literature analysis and contributed to the synthesis of findings. L.C. guided the overall structure of the manuscript and performed critical revisions. All authors read and approved the final manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

1. AlFalahi L and Nobanee H. Conceptual Building of Sustainable Economic Growth and Corporate Bankruptcy. Available at SSRN 3472409 2019.
2. Máté D, Sadaf R, Oláh J, Popp J, and Szűcs E. The effects of accountability, governance capital, and legal origin on reported frauds. Technological and Economic Development of Economy 2019;25:1213–31.
3. Bhattacharyya S, Jha S, Tharakunnel KK, and Westland JC. Data mining for credit card fraud: A comparative study. Decis. Support Syst. 2011;50:602–13.
4. Tergiman C and Villeval MC. The way people lie in markets: Detectable vs. deniable lies. Management Science 2023;69:3340–57.
5. Liu Y, Li X, and Zheng Z. Smart Natural Disaster Relief: Assisting Victims with Artificial Intelligence in Lending. Information Systems Research 2023.
6. Jin C, Yang L, and Hosanagar K. To brush or not to brush: Product rankings, consumer search, and fake orders. Information Systems Research 2023;34:532–52.
7. Delkhosh F, Gopal RD, Patterson RA, and Yaraghi N. Impact of Bot Involvement in an Incentivized Blockchain-Based Online Social Media Platform. Journal of Management Information Systems 2023;40:778–806.

8. Zhu H, Xiong H, Ge Y, and Chen E. Discovery of ranking fraud for mobile apps. IEEE Transactions on knowledge and data engineering 2014;27:74–87.

9. Papanastasiou Y, Yang SA, and Zhang AH. Improving dispute resolution in two-sided platforms: The case of review blackmail. Management Science 2023.

10. Seeja K and Zareapoor M. FraudMiner: A novel credit card fraud detection model based on frequent itemset mining. The Scientific World Journal 2014;2014.

11. Balagolla E, Fernando W, Rathnayake R, Wijesekera M, Senarathne A, and Abeywardhana K. Credit card fraud prevention using blockchain. In: *2021 6th international conference for Convergence in Technology (I2CT)*. IEEE. 2021:1–8.

12. Weinmann M, Valacich J, Schneider C, Jenkins JL, and Hibbeln MT. The path of the righteous: Using trace data to understand fraud decisions in real time. MIS Quarterly 2021.

13. Fiore U, De Santis A, Perla F, Zanetti P, and Palmieri F. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences 2017.

14. Xiao J, Tian Y, Jia Y, Jiang X, Yu L, and Wang S. Black-box attack-based security evaluation framework for credit card fraud detection models. INFORMS Journal on Computing 2023.

15. Xu JJ, Chen D, Chau M, Li L, and Zheng H. PEER-TO-PEER LOAN FRAUD DETECTION: CONSTRUCTING FEATURES FROM TRANSACTION DATA. MIS quarterly 2022;46.

16. Zhou F, Wang G, Zhang K, Liu S, and Zhong T. Semi-Supervised Anomaly Detection via Neural Process. IEEE Transactions on Knowledge and Data Engineering 2023.

17. Yang Y, Zhang K, and Fan Y. Analyzing firm reports for volatility prediction: A knowledge-driven text-embedding approach. INFORMS Journal on Computing 2022;34:522–40.

18. Tang J, Li J, Gao Z, and Li J. Rethinking graph neural networks for anomaly detection. In: *International Conference on Machine Learning*. PMLR. 2022:21076–89.

19. Xiao K, Liu Q, Liu C, and Xiong H. Price shock detection with an influence-based model of social attention. ACM Transactions on Management Information Systems (TMIS) 2017;9:1–21.

20. Sun J, Xiao K, Liu C, Zhou W, and Xiong H. Exploiting intra-day patterns for market shock prediction: A machine learning approach. Expert Systems with Applications 2019;127:272–81.

21. Cheng D, Ye Y, Xiang S, Ma Z, Zhang Y, and Jiang C. Anti-Money Laundering by Group-Aware Deep Graph Learning. IEEE Transactions on Knowledge and Data Engineering 2023:1–13.

22. Xiang S, Zhu M, Cheng D, et al. Semi-supervised Credit Card Fraud Detection via Attribute-driven Graph Representation. In: *AAAI*. 2023.

23. Fu K, Cheng D, Tu Y, and Zhang L. Credit card fraud detection using convolutional neural networks. In: *International Conference on Neural Information Processing*. Springer. 2016:483–90.

24. Cheng D, Wang X, Zhang Y, and Zhang L. Graph Neural Network for Fraud Detection via Spatial-temporal Attention. IEEE TKDE 2020:1–1.

25. Jing R, Tian H, Zhou G, Zhang X, Zheng X, and Zeng DD. A GNN-based Few-shot learning model on the Credit Card Fraud detection. In: *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*. IEEE. 2021:320–3.

26. Zhou Y, Li J, Hao JK, and Glover F. Detecting critical nodes in sparse graphs via "reduce-solve-combine" memetic search. INFORMS Journal on Computing 2023.

27. Cheng D, Niu Z, and Zhang Y. Contagious Chain Risk Rating for Networked-guarantee Loans. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining 2020.

28. Li X, Liu S, Li Z, et al. FlowScope: Spotting Money Laundering Based on Graphs. In: *AAAI Conference on Artificial Intelligence*. 2020.

29. Chen T and Tsourakakis C. Antibenford subgraphs: Unsupervised anomaly detection in financial networks. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2022:2762–70.

30. Zheng W, Yan L, Gou C, and Wang FY. Federated Meta-Learning for Fraudulent Credit Card Detection. In: *International Joint Conference on Artificial Intelligence*. 2020.

31. Wang D, Qi Y, Lin J, et al. A Semi-Supervised Graph Attentive Network for Financial Fraud Detection. 2019 IEEE International Conference on Data Mining (ICDM) 2019:598–607.

32. Chen C, Liang C, Lin J, et al. InfDetect: a Large Scale Graph-based Fraud Detection System for E-Commerce Insurance. 2019 IEEE International Conference on Big Data (Big Data) 2019:1765–73.

33. Weber M, Chen J, Suzumura T, et al. Scalable Graph Learning for Anti-Money Laundering: A First Look. In: *NeurIPS*. Worksop on Challenges and Opportunities for AI in Financial Services. 2018.

34. Paquet-Clouston M, Haslhofer B, and Dupont B. Ransomware Payments in the Bitcoin Ecosystem. arXiv preprint 2018. https://doi.org/10.48550/arXiv.1804.04080.

35. Wang D, Zhang Z, Zhou J, et al. Temporal-Aware Graph Neural Network for Credit Risk Prediction. In: *SDM*. 2021.

36. Cheng D, Tu Y, Ma Z, Niu Z, and Zhang L. Risk assessment for networked-guarantee loans using high-order graph attention representation. In: *IJCAI*. AAAI Press. 2019:5822–8.

37. Cheng D, Wang X, Zhang Y, and Zhang L. Risk Guarantee Prediction in Networked-Loans. In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. IJCAI'20. Yokohama, Yokohama, Japan, 2021.

38. Yang S, Zhang Z, Zhou J, et al. Financial Risk Analysis for SMEs with Graph-based Supply Chain Mining. In: *International Joint Conference on Artificial Intelligence*. 2020.

39. Hu B, Zhang Z, Zhou J, et al. Loan Default Analysis with Multiplex Graph Learning. Proceedings of the 29th ACM International Conference on Information & Knowledge Management 2020.

40. Liang C, Liu Z, Liu B, et al. Uncovering Insurance Fraud Conspiracy with Network Learning. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval 2019.

41. Liu Y, Ao X, Qin Z, et al. Pick and Choose: A GNN-based Imbalanced Learning Approach for Fraud Detection. Proceedings of the Web Conference 2021 2021.

42. Zhong Q, Liu Y, Ao X, et al. Financial defaulter detection on online credit payment via multi-view attributed heterogeneous information network. In: *Proceedings of The Web Conference 2020*. 2020:785–95.

43. Hu B, Zhang Z, Shi C, Zhou J, Li X, and Qi Y. Cash-Out User Detection Based on Attributed Heterogeneous Information Network with a Hierarchical Attention Mechanism. Proceedings of the AAAI Conference on Artificial Intelligence 2019;33:946–53.

44. Kirkos E, Spathis C, and Manolopoulos Y. Data Mining techniques for the detection of fraudulent financial statements. Expert Syst. Appl. 2007;32:995–1003.

45. Ravisankar P, Ravi V, Rao GR, and Bose I. Detection of financial statement fraud and feature selection using data mining techniques. Decis. Support Syst. 2011;50:491–500.

46. Ki Y and Yoon JW. PD-FDS: Purchase Density based Online Credit Card Fraud Detection System. In: *Proceedings of the KDD 2017: Workshop on Anomaly Detection in Finance*. Ed. by Anandakrishnan A, Kumar S, Statnikov A, Faruquie T, and Xu D. Vol. 71. Proceedings of Machine Learning Research. PMLR, 2018:76–84. URL: https://proceedings.mlr.press/v71/ki18a.html.

47. Wang J, Guo Y, Wen X, Wang Z, Li Z, and Tang M. Improving graph-based label propagation algorithm with group partition for fraud detection. Applied Intelligence 2020;50:3291–300.

48. Sangers A, Heesch M van, Attema T, et al. Secure multiparty PageRank algorithm for collaborative fraud detection. In: *Financial Cryptography and Data Security: 23rd International Conference, FC 2019, Frigate Bay, St. Kitts and Nevis, February 18–22, 2019, Revised Selected Papers 23*. Springer. 2019:605–23.

49. Khalid F, Javed A, Ilyas H, Irtaza A, et al. DFGNN: An interpretable and generalized graph neural network for deepfakes detection. Expert Systems with Applications 2023;222:119843.

50. Wang D, Lin J, Cui P, et al. A semi-supervised graph attentive network for financial fraud detection. In: *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE. 2019:598–607.

51. Xu Y, Zhang W, Zhang Y, Orlowska M, and Lin X. TimeSGN: Scalable and Effective Temporal Graph Neural Network. In: *2024 IEEE 40th International Conference on Data Engineering (ICDE)*. IEEE. 2024:3297–310.

52. Xu Y, Zhang W, Xu X, Li B, and Zhang Y. Scalable and effective temporal graph representation learning with hyperbolic geometry. IEEE Transactions on Neural Networks and Learning Systems 2024.