World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

Adaptive financial fraud detection using graph neural networks and reinforcement learning

Pius Businge [1,*], Ivan Asiimwe Agaba [1], Jude Innocent Atuhaire [1], Faith Isabella Nayebale [2], Joram Gumption Ariho [1], Brian Mugalu [1], Denis Musinguzi [3], Curthbert Jeremiah Malingu [1] and Collin Arnold Kabwama [1]

[1] *Department of Computer Science, Maharishi International University, Fairfield, Iowa, USA.*
[2] *Department of Business Administration, Maharishi International University, Fairfield, Iowa, USA.*
[3] *Department of Electrical and Computer Engineering, Makerere University, Kampala, Uganda.*

World Journal of Advanced Research and Reviews, 2025, 28(02), 842-847

Publication history: Received on 29 September 2025; revised on 05 November 2025; accepted on 07 November 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.28.2.3761

Abstract

Fraud detection in financial systems is critical to maintain security and trust in these systems. Traditional fraud detection methods often struggle with the dynamic fraud patterns, and this requires method that can adapt in real-time. Traditional systems require large volumes of labeled data which is difficult to obtain given the private nature of financial systems. In this work, we introduce a framework that utilizes reinforcement learning to enhance fraud detection capabilities. We treat fraud detection as a sequential decision-making process. Reinforcement learning agents can learn and refine optimal strategies to identify fraudulent activities as they adapt to new transaction data. The method combines transaction history, and user behavior to enhance detection efficiency. Through evaluation on benchmark datasets, we show that this approach significantly improves detection rates and reduces the incidence of false positives which can hurt business profits. It makes the systems robust while facilitating real-time processing. Our results suggest that reinforcement learning agents, when combined with well-designed reward mechanisms, can outperform traditional models in detecting fraudulent activities.

Keywords: Financial Fraud Detection; Reinforcement Learning; Graph Neural Networks (GNNS); Deep Q-Network (DQN)

1. Introduction

Financial fraud has become a critical challenge in modern financial systems, with global losses estimated to exceed $6 trillion annually**.** Beyond the direct monetary impact, fraud undermines public trust in financial institutions and threatens the stability of the global economy. The rapid digitalization of financial services has expanded both the volume and complexity of financial transactions, creating new opportunities for exploitation. Contemporary fraud schemes are increasingly sophisticated, often spanning multiple transactions, accounts, and even institutions hence rendering detection methods based on isolated events largely ineffective. Moreover, fraud patterns evolve rapidly, frequently outpacing the update cycles of traditional detection systems.

Regulatory constraints further compound these challenges. Strict data protection policies limit the sharing of financial data across institutions, reducing the availability of comprehensive datasets for fraud detection. Additionally, fraudulent transactions represent only a minute fraction of total activity, leading to highly imbalanced datasets that challenge the performance of standard machine learning models.

* Corresponding author: Pius Businge

Traditional approaches, such as rule-based systems and static machine learning models, struggle to address these issues effectively. Rule-based systems, though interpretable, require constant manual updates to remain effective against new fraud patterns. Static machine learning models are more flexible but often treat transactions as independent events, ignoring the intricate relationships among entities. While deep learning techniques have recently improved fraud detection accuracy, they too face limitations: most models fail to capture temporal dependencies and inter-entity relationships, struggle with concept drift**,** rely on centralized data (raising privacy and compliance concerns), and use fixed decision thresholds that may become suboptimal over time.

To overcome these limitations, we propose a hybrid framework that integrates Graph Neural Networks (GNNs) and Reinforcement Learning (RL) for adaptive financial fraud detection. GNNs are well-suited for modeling financial ecosystems, as entities (e.g., accounts, users, and institutions) and transactions can naturally be represented as nodes and edges in a graph. This structure allows the model to capture both spatial and relational dependencies in financial data. Reinforcement learning complements this by enabling dynamic adaptation, learning optimal detection strategies through continuous feedback as fraud patterns evolve.

Our proposed framework introduces a novel graph-based architecture that jointly models temporal**,** spatial**,** and semantic patterns in transaction data. We further employ a Deep Q-Network (DQN) to dynamically adjust both feature weights and fraud classification thresholds, enhancing the model's adaptability to changing fraud behaviors. To address data imbalance, we integrate specialized training techniques that improve the model's sensitivity to rare fraudulent cases. Through extensive experiments, we demonstrate that our method significantly outperforms state-of-the-art baselines in terms of accuracy**,** false-positive reduction**,** and resilience to concept drift**.**

The remainder of this paper is organized as follows: Section 2 reviews related work in fraud detection, GNNs, and RL. Section 3 presents relevant methodology. Section 4 details the proposed approach**.** Section 5 discusses the experimental setup and results. Finally, Section 6 concludes the paper and outlines future directions.

2. Related work

2.1. Financial Fraud Detection

Traditional fraud detection was based on rule-based systems and statistical methods. Bhattacharyya et al [8] compared several data mining approaches and highlighted class imbalance as a major challenge to such systems. With advancements in machine learning, fraud detection improved because of better feature extraction techniques for instance transaction aggregation [11]. Pozollo et al [1] tackled concept drift using sliding windows. Deep learning approaches further enhanced detection capabilities with Roy et al [2] demonstrating the superiority of neural networks in capturing non-linear patterns and Wang et al [6] combined autoencoders with random forests for imbalanced data. However, key challenges remained unaddressed, for instance handling the interconnected nature of transactions and evolving fraud patterns.

2.2. Graph Based Approaches in Fraud Detection

Graphs are a natural representation of financial systems and graph-based methods have shown promise in financial fraud detection through their ability to model complex transaction relationships [10]. Akoglu et al [3] surveyed graph-based anomaly detection methods highlighting techniques like community detection and subgraph mining for identifying fraudulent patterns. Liu et al [4] introduced an isolation-based method for graph structured data. The introduction of graphical neural networks brought significant advances. Kipf et al [5] introduced graph convolution networks (GCNs) and Hamilton et al [9] proposed GraphSAGE for dynamic network embedding. More recent works have further enhanced these approaches, with Dou et al developing attention based GNNs for detecting camouflaged fraudsters, and Liu et al. integrating Multiview financial relationships in a heterogeneous graph framework.

2.3. Reinforcement Learning in Fraud Detection

Reinforcement learning has shown great potential in various financial applications, including trading and risk management. Sutton and Barto [7] highlighted the ability of reinforcement learning models to learn optimal policies in dynamic environments which is key for financial fraud detection systems. Despite this, the application of reinforcement learning to fraud detection has been limited.

Our work aims to fill the gaps of previous approaches by incorporating graphical neural networks and reinforcement learning.

2.4. Graph Representation of Financial Transactions

We model financial transaction networks as a heterogeneous graph $G = (V, E, X)$where:

- $V = v_1, v_2, \dots, v_n$ is the set of nodes representing entities such as users and merchants.
- $E = e_1, e_2, \dots., e_m$ is the set of edges representing transactions between entities.
- $X = x_1, x_2, \dots, x_n$ is the set of node feature vectors, where $x_i = R^d$ is the feature vector of node v_i

Each edge $e_k = (v_i, b_j, t_k, f_k)$ represents a transaction from entity vi to entity $v_j$at time $t_k$with feature vector f_k. The graph G is dynamic, evolving over time as new transactions occur.

2.5. Graph Neural Networks

Graph neural networks are deep learning models designed to operate on graph structured data. They function by using a neighborhood aggregation scheme to iteratively update a node's representation by aggregating information from its neighbors, enabling it to learn from the relationships within the data. Each node in the graph has a node embedding that represents the features of the node. These embeddings can be used for multiple tasks. Fraud detection can be treated as an edge classification task where the goal is to predict whether a transaction is fraudulent or not. A general form of the node update in a GNN layer can be expressed as:

$$h_i^{l+1} = \sigma\left(W^{(l)}.\,\text{AGGREGATE}\left(h_j^{(l)}: j \in N(i)\right) + b^{(l)}\right)$$

where h_i^l is the feature vector of node I at layer l, N(i) is the set of neighbors of node I, $W^{(l)}$

and $b^{(l)}$ are learnable parameters and σ is a nonlinear activation function. The AGGREGATE function can take various forms, such as mean, sum, or more sophisticated pooling operations.

2.6. Reinforcement learning

Reinforcement learning is a framework for learning to make sequential decisions in an uncertain environment to maximize a cumulative reward in the absence of guidance from a human user. In reinforcement learning, learning is through trial and error. Reinforcement learning consists of the relationship between an agent, environment and a goal. This relationship is formulated in terms of a Markov decision process (MDP). The agent learns about a problem by interacting with its environment. The environment provides information on its current state. The agent then uses that information to determine which actions to take to maximize its long-term reward. If a particular action obtains a high reward signal from the surrounding environment, the agent is encouraged to take that action again when in a similar state in the future. The fraud detection problem is an offline learning problem.

The RL problem is typically formulated as a Markov Decision Process (MDP), defined by a tuple (S, A, P, R, γ) where:

- S is the set of states
- A is the state of actions
  - $P: S \times A \times \rightarrow [0,1]$ is the transition probability function
  - $R: S \times A \rightarrow R$ is the reward function

$\gamma \in [0,1]$ is the discount factor

The goal of RL is to learn a policy π: $S \rightarrow A$ that maximizes the expected cumulative discounted reward:

$$J(\pi) = E_{\tau \sim \pi}[\sum_{t=0}^{\infty} \gamma^t \; R(s_t, a_t)]$$

Where $\tau = (s_0, a_o, s_0, a_1, \dots.)$ is a trajectory sampled according to policy π.

2.7. Q-Learning and Deep Q-Networks

Q-learning is a value based off-policy reinforcement learning algorithm used to fins the optimal action-selection policy for a given finite Markov Decision Process (MDP). The aim of Q-learning is to learn a Q-value function Q(s, a) which represents the expected future reward when taking an action in a state and the following the optimal policy thereafter.

The Q-value function estimates the quality of a state-action pair. The Q-value function is updates using the Q-learning update rule defined as follows.

$$Q(s\_t, a\_t) \leftarrow Q(s\_t, a\_t) + \alpha[r_t + \gamma max'_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

Where α is the learning rate

Deep Q-Networks extend Q-learning by using a deep neural network to approximate the Q function. The networks are trained to minimize the loss:

$$L(\theta) = E_{(s,a,r,s')\sim D}\left[r + \gamma max'_a Q(s', a', \theta^-) - Q(s, a; \theta)\right)^2]$$

Where D is the replay buffer of past experiences and θ^- are the parameters of a target network that is periodically updated for stability.

3. Methodology

3.1. Experimental setup

We evaluate our model on two real-world finance-related fraud detection datasets:

3.2. Paysim Mobile Money Dataset

This is a synthetic dataset with simulation of mobile money transactions based on a sample if real transactions extracted from one month of financial logs from a mobile money service implemented in an African country. The dataset contains 6,362,620 transactions, among which 8,213 are fraudulent. Each transaction record includes 11 features such as transaction type (CASH IN, CASH OUT, DEBIT, PAYMENT and TRANSFER), amount, sender and recipient information, and initial and final account balances. The data spans 744 hours of simulations, with each step representing an hour.

3.3. IEEE-CIS Fraud Detection Dataset

This dataset is obtained from a real-world e-commerce fraud detection challenge, composed of 590,000 online transactions. Each transaction is described by two types of features: transaction features including product, card information, addresses, emails, and various merchant features M1-M9 and identity features including Device Type, Device Info, and identity indicators id_12-id_38. The transactions are chronologically ordered, with transaction DT representing the time delta from a reference point. The dataset is split into training, and testing sets, maintaining the temporal nature of fraud patterns. This dataset offers rich contextual information through both transaction and identity features.

For all datasets, we perform the following preprocessing steps: we impute the missing values using mean values for numerical features and mode for categorical features, we scale the features using min-max normalization for numerical features and we encode categorical variables using one-hot encoding. We extracted temporal features including hour of the day, day of the week, and time difference between consecutive transactions. For the IEEE-CIS dataset, we merged the transaction and identity information using the Transaction ID as the key thus handling cases where identity information is missing.

3.4. Baselines

We compared the performance of our approach with the following methods:

3.4.1. XGBoost

This is an optimized gradient boosted library designed to be highly efficient, flexible and portable. XGBoost is popular fraud detection method due to high performance and ability to handle imbalanced datasets.

3.4.2. Graph Convolution Network

This is a variant of GNNs that can take advantage of the structural information in graphs. GCNs have shown promise in fraud detection by capturing the relationships between entities in financial transaction networks.

3.5. Evaluation Metrics

We use the following metrics for evaluation:

3.5.1. AUC-ROC

Area Under the Receiver Operating Characteristic curve. This metric provides an aggregate measure of performance across all possible classification thresholds.

3.5.2. F1-score

The harmonica mean of precision and recall. It provides a single score that balances precision and recall.

3.5.3. AUC-PR

Area Under the Precision and Recall curve. This metric focuses on the minority class when dealing with an imbalanced dataset.

3.6. Implementation Details

We implement our approach using PyTorch and Pytorch Geometric. We constructed a transaction graph where each transaction is a node, and edges are created based on temporal proximity and feature similarity. We connected transactions that occur within an hour of each other and have a cosine similarity of their feature vectors above a threshold of 0.9.

Our model is composed of 3 graph convolution layers with 128 hidden units each with a ReLU activation function and apply batch normalization after each convolution later. For the RL agent, we use a DQN with two hidden layers of 256 units each. We use the AdamW optimizer with a learning rate of 0.001 and a batch size of 128.We trained the model for 50 epochs. All experiments were conducted on a single A40 GPU with 48GB of VRAM.

4. Results

Table 1 Results of the evaluation of the baseline methods and our GCN-RL approach

Method	Paysim			IEEE-CIS		
	AUC-ROC	AUC-PR	F1-Score	AUC-ROC	AUC-PR	F1-Score
XGBoost	0.948	0.412	0.756	0.952	0.425	0.768
GCN	0.978	0.578	0.879	0.980	0.582	0.882
GCN-RL	0.997	0.675	0.925	0.993	0.654	0.930

Table 1 shows the results of all the models. Our proposed model outperforms all the other models across all datasets and evaluation metrics. XGBoost achieves a reasonable AUC-ROC score of 0.948 but a relatively low AUC-PR, indicating its limitations in handling highly imbalanced fraud detection tasks. The graph convolution network had an AUC-PR score of 0.578 and an F1-score of 0.878.

The performance of the model can be attributed to the combination of graph neural network and reinforcement learning which enables adaptive feature learning and decision making, while the temporal-spatial semantic graph convolution captures financial patterns. The results show that our model effectively balances the trade-off between detection accuracy and false positive rate, making it suitable for practical deployment in financial fraud detection systems.

5. Conclusion and future work

This work presents the GCN-RL, a framework combining GNNs with RL for financial fraud detection. We show that by combining these 2 machine learning approaches, we obtain model that is robust to the imbalance in fraud detection dataset. In addition, it can adapt to changing patterns of fraudsters in real time. Despite the performance of our method, we believe it can be improved further by incorporating federated learning to preserve the privacy of data.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," Expert Syst. Appl., vol. 41, no. 10, pp. 4915–4928, 2014.

[2] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, and P. Beling, "Deep learning detecting fraud in credit card transactions," in Proc. Syst. Inf. Eng. Des. Symp., 2018, pp. 129–134.

[3] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey," Data Mining Knowl. Discov., vol. 29, no. 3, pp. 626–688, 2015.

[4] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," ACM Trans. Knowl. Discov. from Data, vol. 6, no. 1, pp. 1–39, 2012.

[5] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in Proc. Int. Conf. Learn. Representations, 2017.

[6] D. Wang et al., "Deep fraud detector: A deep learning framework for financial fraud detection," in Proc. IEEE Int. Conf. Data Mining, 2019, pp. 1361–1366.

[7] R. S. Sutton and A. G. Barto," in Reinforcement Learning: An Introduction. Cambridge, MA, USA: MIT Press, 2018.

[8] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," Decis. Support Syst., vol. 50, no. 3, pp. 602–613, 2011.

[9] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," IEEE Data Eng. Bull., vol. 40, no. 3, pp. 52–74, 2017.

[10] Collin Arnold Kabwama et al., "Graph attention networks for credit card fraud detection: A relational learning approach," World Journal of Advanced Research and Reviews, vol. 26, no. 3, pp. 2580–2585, Jun. 2025, doi: 10.30574/wjarr.2025.26.3.2400.

[11] Curthbert Jeremiah Malingu et al., "Application of LLMS to Fraud Detection," World Journal of Advanced Research and Reviews, vol. 26, no. 2, pp. 178–183, May 2025, doi: 10.30574/wjarr.2025.26.2.1586.