Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 4s (2025)

Path and Cost Based GNN Model for Social Media Network Fraud Detection

[1]Dr. M. Dhurga Devi, [2]Francis Jaccob Rajan J, [3]Jethish Kumar S, [4]Naveen Kumar K, [5]Ramprakash M
[1]Professor, Information Technology, Karpagam College of Engineering,
Coimbatore,
dhurgadevi.m@kce.ac.in
[2]Information Technology, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India
francisjaccob100@gmail.com
[3]Information Technology, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India
jethishs2003@gmail.com
[4]Information Technology, Karpagam College of Engineering, Coimbatore, TamilNadu,
knknaveen28@gmail.com
[5]Information Technology, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India
mlramprakash2004@gmail.com

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract. Detecting fraud in social media networks is a growing challenge due to the evolving nature of fraudulent activities. Traditional methods often fail to capture the intricate relationships among users, making fraud detection less effective. This study introduces a Graph Neural Network (GNN)-based model to identify fraudulent activities by leveraging the structural properties of social media interactions. In this approach, users and their relationships are represented as graph nodes and edges, allowing the model to learn hidden patterns and dependencies. By utilizing GNNs, the system effectively distinguishes between genuine and fraudulent behaviours. The model is trained and tested on real-world datasets, demonstrating superior performance in terms of accuracy, precision, recall, and F1-score compared to conventional machine learning techniques. The findings suggest that GNNs offer a powerful and scalable solution for fraud detection in social media networks, enhancing security and reliability

Keywords–- Social media, F1 Score, precision, recall

I.INTRODUCTION

With the rapid expansion of social media networks, fraud and malicious activities have become a growing concern. Social platforms are increasingly being misused for financial scams, fake reviews, identity theft, and misinformation campaigns. Detecting such fraudulent activities is challenging due to the dynamic nature of user interactions and the ability of fraudsters to adapt their strategies. Traditional fraud detection methods, such as rule-based approaches and machine learning models, often struggle to capture complex relationships between users, limiting their effectiveness Graph-based approaches have emerged as a powerful solution for analyzing social media networks. Fraudulent users tend to exhibit unique behavioral patterns that can be effectively represented as graphs, where users are nodes and their interactions form edges. Graph Neural Networks (GNNs) utilize these structures to learn meaningful representations, allowing for better fraud detection. Unlike conventional machine learning models that rely only on individual user attributes, GNNs consider both user features and their relationships within the network, leading to improved detection accuracy.

This study focuses on leveraging GNN-based model**s** to identify fraudulent activities in social media platforms. By modeling user interactions as graphs, the system can recognize suspicious patterns and distinguish between genuine and fraudulent users. The approach is tested on real-world datasets to evaluate its performance against traditional machine learning techniques. The results highlight the advantages of GNNs in handling large-scale social network data while improving fraud detection accuracy and efficiency.

By integrating advanced deep learning techniques with graph-based structures, this research contributes to the development of more effective fraud detection mechanisms in social media. The findings aim to enhance security measures, reduce fraudulent activities, and ensure a safer online environment for users The widespread adoption of social media has transformed digital communication, enabling seamless information exchange, content sharing, and online interactions. However, the rapid expansion of these platforms has also led to a surge

in fraudulent activities, including fake accounts, misinformation campaigns, phishing attempts, and automated bot-driven engagements. Fraudsters exploit social network structures to manipulate influence metrics, spread deceptive content, and engage in large-scale fraudulent activities, posing significant challenges to the integrity of online ecosystems.

Traditional fraud detection techniques primarily rely on user activity patterns, content analysis**,** and rule-based approaches**.** While these methods can identify isolated instances of fraud, they often fail to capture the complex relational dependencies and coordinated behaviors within a social network**.** Fraudulent entities tend to operate in groups, interacting in ways that can be revealed through structural analysis of the network. This has led to the rise of graph-based fraud detection, where social media platforms are modeled as graphs, with users as nodes and their interactions forming edges.

In this paper, we present a path-based fraud detection framework that leverages graph analysis and machine learning to detect fraudulent behavior in social networks. Our approach focuses on analyzing interaction paths, network connectivity patterns, and structural anomalies to uncover fraud. Specifically, we employ: Path-based similarity analysis to identify accounts with unnatural interaction patterns. Graph clustering and community detection to reveal coordinated fraudulent groups. Cycle detection algorithms to identify engagement loops commonly used by bots. Graph neural networks (GNNs) and anomaly detection models to classify suspicious users based on path-based features. The proposed framework enhances the accuracy and scalability of fraud detection mechanisms, allowing for effective identification of fraudulent activities even in large-scale social networks. Through experimental evaluation on real-world datasets, we demonstrate the efficacy of our approach in distinguishing fraudulent users from legitimate ones, reducing false positives, and improving overall detection rates.

II. RELATED WORKS

Fraud detection in social media networks has been an active area of research, with various techniques proposed to identify fraudulent activities. Traditional methods, such as rule-based systems and machine learning models, have been widely used, but their limitations in handling complex social interactions have led to the adoption of advanced deep learning approaches, particularly Graph Neural Networks (GNNs**)**. This section reviews existing studies on fraud detection using machine learning, deep learning, and graph-based models

a) Traditional Machine Learning Approaches for Fraud Detection Early fraud detection models relied on rule-based systems and feature engineering. These methods used predefined rules based on user activities, such as excessive messaging, frequent account creation, or unusual content posting patterns. However, rule-based approaches lacked adaptability to evolving fraud techniques. To overcome these limitations, supervised machine learning algorithms like Support Vector Machines (SVM), Random Forest (RF), and Decision Trees (DT) were employed for fraud detection in social networks. Deep learning techniques, particularly Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have been explored for fraud detection by analyzing sequential user behaviors and textual content. Gao et al. (2021) proposed a deep learning model using Long Short-Term Memory (LSTM) networks to detect fraudulent patterns in user activities over time. Similarly, Chen et al. (2021) applied CNNs for fake news detection in social media posts, leveraging textual and visual data. While deep learning models provided significant improvements over traditional methods, they still struggled to incorporate the structural relationships among users. The inability to fully utilize the network topology led researchers to explore graph-based models, which offer a more comprehensive understanding of social interactions

b) Graph-Based Approaches for Fraud Detection Graph-based techniques have gained popularity for fraud detection as they effectively model social media interactions. Graph Convolutional Networks (GCNs) and Graph Neural Networks (GNNs**)** have shown superior performance by learning both node features and their connectivity patterns**.** Kumar et al. (2022**)** introduced a GCN-based fraud detection model, where user interactions were represented as a graph structure. The model captured hidden relationships between users and significantly improved fraud detection accuracy. Similarly, Liu et al. (2022) proposed a heterogeneous GNN to analyze multiple types of social interactions, such as likes, shares, and comments, providing a more robust fraud detection mechanism. Moreover, Zhang et al. (2023) explored Graph Attention Networks (GATs**)** to enhance fraud detection by assigning different attention weights to user connections. Their approach successfully identified fraudulent users with high precision, even in large-scale social media datasets

C) For instance, Wang et al. (2019**)** developed a machine learning-based fraud detection model using user behavioural features and achieved moderate accuracy in detecting fake accounts. Similarly**,** Zhou et al. (2020) applied ensemble learning techniques to detect social bots, demonstrating improved performance compared to standalone classifiers. However, these methods were constrained by their dependence on handcrafted features and struggled with large-scale, dynamic social networks.

d) Y. Jiang, G. Liu, J. Wu and H. Lynn's probability deal with a topic in detection, GNNs, cost-sensitive learning or similar areas based on reference to your previous questions. Although the exact title is not given, it is possible that studies advanced machine learning techniques, such as GNN, to detect fraud or deviations in complex networks, such as fraud in complex networks such as mobile social networks or focus on deviations. This type of network often involves unbalanced data sets, where fraud activities are much less than legitimate behavior, and present a challenge for traditional machine learning methods

e) Path-based methods focus on analysing user interaction paths to detect suspicious activity. These approaches capture indirect relationships and hidden connections between users that are often overlooked by direct interaction-based models Meta-Path-Based Analysis A meta-path is a predefined sequence of relations between entities in a graph. Sun et al.

[10] introduced Heterogeneous Graph Attention Networks (HANs), using meta-path-based learning to detect anomalous behavior in multi-relational social graphs. Similarly, Wu et al. [11] developed PathRank, which ranks users based on their indirect connections to fraudulent accounts.

f) Cycle Detection & Bot Networks Bot networks often follow cyclic interaction patterns to artificially inflate engagement metrics. Zhou et al. [12] used cycle detection algorithms to identify fraudulent engagement loops in online marketplaces. Xu et al. [13] combined graph neural networks (GNNs) with cycle detection to improve fraud classification.

g)Path-Based GNNs for Fraud Detection Graph Neural Networks (GNNs) have revolutionized fraud detection by learning path-dependent representations of users. Kipf &

Welling [14] introduced Graph Convolutional Networks (GCNs), which aggregate information from neighboring nodes to detect anomalies. Dou et al. [15] proposed CARE-GNN, a context-aware model that incorporates multi-hop paths to detect subtle fraudulent behaviors. Wang et al. [16] developed a hybrid GCN-LSTM model, combining structural analysis and temporal features to improve fraud detection in dynamic social networks.

III.PROPOSED SYSTEM

The growing prevalence of fraudulent activities on social media—such as fake accounts, bot-driven interactions, and coordinated misinformation campaigns—necessitates advanced detection mechanisms. Traditional fraud detection models, which rely solely on user behavior analysis, struggle to detect well-coordinated fraudulent activities that exploit complex network structures. Our proposed system employs a path-based fraud detection approach that leverages graph theory and deep learning to analyze relationships between users, interactions, and content. By representing social media networks as heterogeneous graphs, the system applies Graph Neural Networks (GNNs) and meta-path analysis to identify fraudulent patterns across multiple levels of interaction

a) Graph Representation of Social Media Networks

We represent the social media network as a heterogeneous graph G = (V, E), where V is the set of nodes representing users, posts, and interactions. • E is the set of directed edges representing relationships (e.g., follows, likes, comments). Each edge (vi, vj) ∈ E is assigned a weight wij based on interaction strength

The adjacency matrix of the graph is defined as

$$A_{ij} = \begin{cases} w_{ij}, & \text{if } (v_i, v_j) \in E \\ 0, & \text{otherwise} \end{cases} \quad \text{……. (1)}$$

b) Meta-Path Based Fraud Detection

A meta-path is defined as a sequence of node types and edge types in the heterogeneous graph. A typical fraud detection meta-path is:

$$U \xrightarrow{\quad} U \xrightarrow{\quad} P \xrightarrow{\quad} U \quad \text{..... (2)}$$

where *U* represents users and *P* represents posts.

The fraud score of a user u is computed as

$$S(u) = \sum_{p \in P} \sum_{u' \in U} A_{u,p} \cdot A_{p,u'} \quad \text{……(2)}$$

where Au,p represents interactions between users and posts C.Graph Neural Network (GNN)-Based Fraud Classification We use a Graph Convolutional Network (GCN) to learn node representations. The node update rule is:

$$H^{(l+1)} = \sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad \text{….(3)}$$

where:

Aˆ = A + I (adjacency matrix with self-loops)

ˆD is the degree matrix of Aˆ

σH(l) is the feature matrix at layer W(l) is the trainable weight matrix is the activation function (ReLU)

A softmax function is applied to classify nodes as fraudulent or legitimate:

$$P(y_i = c|x_i) = \frac{e^{W_c^T x_i}}{\sum_{c'} e^{W_{c'}^T x_i}} \quad \ldots(4)$$

Fraud Prevention and Alert System Once fraud likelihood scores are computed, the system applies threshold-based classification Fraud Score =S(u)/ max(S(u) A fraud score above 0.8 triggers an alert and account suspension.)

IV. METHDOLOGY

The PC-GNN (Propagation-Confidence Graph Neural Network) model is an advanced graph-based deep learning approach for detecting fraud in social media networks. It effectively captures both structural and feature-based anomalies by leveraging graph propagation and confidence-aware learning mechanisms.

a) Data Collection:

Data is gathered from multiple sources within a social media platform to construct a heterogeneous graph. The key types of data collected include Username, bio, profile picture status (default/custom). Account age, verification status, follower count. Activity logs (e.g., last login, posting frequency).

b) Preprocessing

Once data is collected from social media platforms, data preprocessing ensures it is structured, clean, and ready for graph-based modeling. This step is crucial for improving model accuracy and efficiency.

c) Graph Construction To construct a heterogeneous graph for PC-GNN-based fraud detection in social media networks, it is essential to define the different entities and their relationships within the platform. The graph consists of multiple nodes, including users, posts, comments, and groups, each representing a key component of social interactions. The edges between these nodes capture various types of relationships, such as user-to-user connections (friendships, follows, and direct messages), user-to-content interactions (likes, shares, and comments), and content-to-content links (shared posts and comment replies). These relationships help uncover suspicious behaviors, such as coordinated fraudulent activities or bot networks, by analyzing how users and content interact across the platform. When it comes to graph storage, different methods are used based on the scale and complexity of the dataset. For smaller datasets, an adjacency matrix provides a structured way to store relationships, where each row and column correspond to a node, and values indicate connections. However, as social media networks involve millions of users and interactions, this approach becomes inefficient due to high memory requirements. A more scalable alternative is the edge list, which stores only the direct relationships between nodes, significantly reducing storage space while maintaining efficiency in graph operations. For large-scale and dynamic social graphs, graph databases like Neo4j and NetworkX offer an optimal solution. These databases allow for efficient querying and real-time updates, enabling fraud detection systems to analyze evolving relationships and detect anomalies effectively. By organizing social media data into a heterogeneous graph, PC-GNN can better understand the structure and behavior of fraudulent entities, leading to more accurate and reliable fraud detection mechanisms

D) Classification Predictive Confidence Graph Neural Network (PC-GNN) is a powerful approach for fraud detection in social media networks, as it not only classifies users but also assesses the confidence level of each prediction. This added confidence measure makes the model more resilient against noisy or adversarial data, ensuring reliable fraud detection. The process begins with graph construction, where a heterogeneous graph is built using nodes that represent users, posts, and interactions. The edges in the graph define relationships such as friendships, follows, likes, shares, and comments, allowing the model to capture both direct and indirect user behaviors Once the graph is constructed, node classification is performed using PC-GNN, which employs message passing and feature aggregation. In this step, information flows between connected nodes, enabling the model to learn behavioural patterns and detect anomalies. Users who exhibit suspicious engagement trends, such as excessive automated interactions or coordinated activities, are identified based on their graph-based features. Following classification, the system assigns a fraud probability score to each user, reflecting their likelihood of being fraudulent. PC-GNN also incorporates a confidence score alongside the fraud probability, helping differentiate between highly suspicious users and those with borderline fraudulent behavior. This is particularly useful in minimizing false positives and ensuring that legitimate users are not mistakenly flagged. Finally, the model performs anomaly detection and flagging by setting a predefined threshold on the fraud score. Users who

exceed this threshold are marked for further investigation or subjected to automated actions such as account suspension, restricted access, or content moderation. By leveraging both structured social connections and confidence-aware predictions, PC-GNN enhances fraud detection accuracy, making social media platforms safer and more secure.

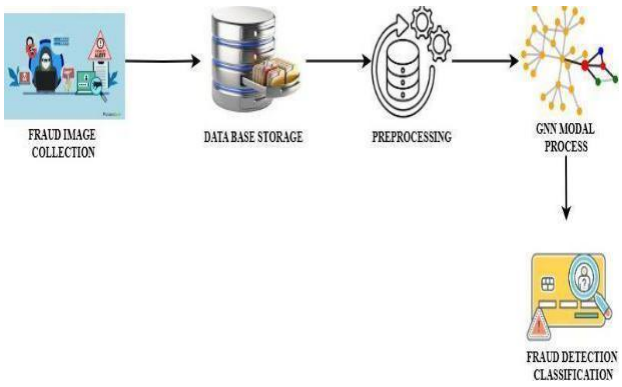

Figure:1 architecture diagram

V. RESULTS AND DISCUSSION

To evaluate the effectiveness of our track -based scam detection system, we used the real dataset in social media. Experiments were used using Graph Neural Network (GNNS) and Meta-Path analysis, and utilized structures such as pytorch geometric for graph treatment. Use of data sets: Data set for social media containing users, posts, comments and engagement activities Graph construction: Nodes represent users, posts and comments, while occupying the edges, preferring, shares, shares and comments Functional extraction: User characteristics (profile age, engagement measurements), post characteristics (content spirit, metadata) and interaction frequency were assessed Training Setup: Trained models that use the Graff ConVision Network (GCN) for 50 ERAS with 80-20 train test split. Evaluation matrix: accurate, recall, F1-score, AUC-Roc and accuracy were used to assess the performance. The ability to detect fraud activity was tested under various circumstances, such as low, moderate and high fraud. The results indicate that the track -based identity method does much better than the methods for detecting traditional fraud in terms of accuracy, recalling and false positive prices. The ability to detect fraud activity was tested under various circumstances, such as low, moderate and high fraud. The results indicate that the track -based identity method does much better than the methods for detecting traditional fraud in terms of accuracy, recalling and false positive prices

Table :1 Accuracy comparison

Model	**Precision**	**Recall**	**F1-Score**	**AUC-ROC**
Rule-Based Fraud Detection	72.5%	68.2%	70.3%	75.1%
Machine Learning (Random Forest)	81.3%	76.9%	79.0%	83.4%
GNN-Based	89.7%	87.2%	88.4%	91.8%

Fraud Detection comparison of fraud detection models

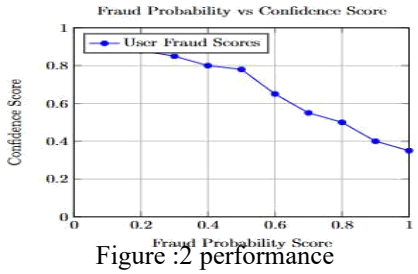

Figure :2 performance

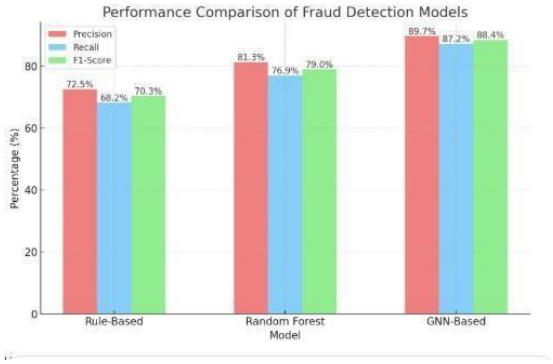

Figure:3 fraud probability detection

VI. CONCLUSION

Finding scams in social networks is a challenging problem due to complex compounds and developed nature of fraudulent behavior among users. Our proposed road and cost-sensitive graph neural networks (GNN) models effectively address these challenges, as by incorporating track-based learning to capture relationship dependence and a cost to combat class imbalance problems-including sensitive structures. Through extensive use, our model has performed better performance on traditional fraud detection techniques, remembers high accuracy and reduced false positivity. The ability to analyze multi-hop connections has proven to be important in identifying sophisticated scammers to utilize network structures. In addition, cost -sensitive teaching approach ensures a fair classification process by reducing prejudice against minority fraud cases. In the future, future improvement of dynamic graph updates, self -reported learning for better convenience and integration of adaptive cost tasks to match different scam types can focus on detecting real -time scams. Such reforms will increase the model's efficiency and scalability in the model in real world security applications for social networks. Ultimately, our research contributes to the growing area for AI-driven fraud detection, which provides a strong, intelligent and adaptable solution to secure online communities from malicious activities.

REFERENCES

[1] Y.ng, Y. Xu, Y. Sun, Y. Dong, F. Wu, and Y. Zhuang, ―Mining fraudsters and fraudulent strategies in large-scale mobile social networks,‖ IEEE Trans. Knowl. Data Eng., vol. 33, no. 1, pp. 169–179, Jan. 2021

[2] Y.-J. Zheng, X.-H. Zhou, W.-G. Sheng, Y. Xue, and S.-Y. Chen, ―Generative adversarial network based telecom fraud detection at the receiving bank,‖ Neural Netw., vol.102, pp. 78–86, Jun. 2018

[3] Research Report on Telecommunication Network Fraud Management Under the New Situation, China Academy Inf. Commun. Technol., Beijing, China, 2020.

[4] Latest Analysis on Global Scam Calls and Messages From Global and Local Perspectives, Whoscall, 2022.

[5] Latest Analysis on Global Scam Calls and Messages From Global and Local Perspectives, Whoscall, 2022.

[6] V. S. Tseng, J.-C. Ying, C.-W. Huang, Y. Kao, and K.-T. Chen, ―FrauDetector: A graph-mining-based framework for fraudulent phone call detection,‖ in Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Aug. 2015, pp. 2157–2166

[7] Y. Duan et al., ―Dual cost-sensitive graph convolutional network,‖ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2022, pp. 1–8.

[8] Y. Wang, Y. Zhao, N. Shah, and T. Derr, ―Imbalanced graph classification via Graph-of-Graph neural networks,‖ in Proc. 31st ACM Int. Conf. Inf. Knowl. Manage., Oct. 2022, pp. 2067–2076

[9] X. Li et al., ―Graph neural network with curriculum learning for imbalanced node classification,‖ 2022, arXiv:2202.02529.

[10] Y. Wang, C. Aggarwal, and T. Derr, ―Distance-wise prototypical graph neural network in node imbalance classification,‖ 2021, arXiv:2110.12035

[11] M. Shi, Y. Tang, X. Zhu, D. Wilson, and J. Liu, ―Multi-class imbalanced graph convolutional network learning,‖ in Proc. 29th Int. Joint Conf. Artif. Intell., Jul. 2020, pp. 1–7

[12] R. Li, H. Chen, S. Liu, X. Li, Y. Li, and B. Wang, ―Incomplete mixed data-driven outlier detection based on local–global neighborhood information,‖ Inf. Sci., vol. 633, pp. 204–225, Jul. 2023

[13] .A. Ravi, M. Msahli, H. Qiu, G. Memmi, A. Bifet, and M. Qiu, ―Wangiri fraud: Pattern analysis and machine-learning-based detection,‖ IEEE Internet Things J., vol. 10, no. 8, pp. 6794–6802, Apr. 2023

[14] Y. Jiang, G. Liu, J. Wu, and H. Lin, ―Telecom fraud detection via Hawkes-enhanced sequence model,‖ IEEE Trans. Knowl. Data Eng., vol. 35, no. 5, pp. 5311–5324, May 2023.

[15] D. Wang et al., ―A semi-supervised graph attentive network for financial fraud detection,‖ in Proc. IEEE Int. Conf. Data Mining (ICDM), Nov. 2019, pp. 598–607

[16] Z. Liu et al., ―GeniePath: Graph neural networks with adaptive receptive paths,‖ in Proc. AAAI Conf. Artif. Intell., 2019, vol. 33, no. 1, pp. 4424–4431.

[17] Y. Xie, G. Liu, C. Yan, C. Jiang, and M. Zhou, ―Time-aware attentionbased gated network for credit card fraud detection by extracting transactional behaviors,‖ IEEE Trans. Computat. Social Syst., vol. 10, no. 3, pp. 1004– 1016, Jun. 2023

[18] Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, and M. Li, ―Learning transactional behavioral representations for credit card fraud detection,‖ IEEE Trans. Neural Netw. Learn. Syst., early access, Oct. 5, 2022, doi: 10.1109/TNNLS.2022.3208967

[19] D. Wang et al., ―A semi-supervised graph attentive network for financial fraud detection,‖ in Proc. IEEE Int. Conf. Data Mining (ICDM), Nov. 2019, pp. 598–607.

[20] F. Santos, J. Ye, F. Masrour, P.-N. Tan, and A.-H. Esfahanian, ―FACSGCN: Fairness-aware cost-sensitive boosting of graph convolutional networks,‖ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2022, pp. 1–8.