A Graph Attention Network Framework for Financial Fraud Detection: Synthetic Validation and Benchmark Evaluation on the Elliptic Bitcoin Dataset

Sulaiman Bello Umar[1], Gendal Lal[2], Muquaddar Ali[3] &
Muazu Abubakar Muhammad[4]

[1,2,3]Department of Computer science & Engineering, Vivekananda Global University, Jaipur, Rajasthan, India

Correspondence: [1]sbudambatta@gmail.com, gendal.lal@vgu.ac.in, muquaddar.ali@vgu.ac.in abubakarmuhammadmuaz@gmail.com

Abstract. Financial fraud represents one of the most pressing challenges in modern digital economies, with global credit card fraud losses projected to exceed US $400 billion in the coming decade. The conventional rule-based and machine-learning detection systems fail to model the intricate relational topology that characterizes fraudulent transactional networks. This paper proposes a Graph Attention Network (GAT)-based fraud detection framework (GAT-FDF) that encodes transaction entities as graph nodes and leverages multi-head attention mechanisms to aggregate neighborhood information for binary fraud classification. The framework is evaluated in two stages: (i) A controlled experiment on a 50,000 node synthetic dataset with deliberate 25% fraudulent transaction was injection, achieving an AUC-ROC of 98.17%, F1 of 85.59%, and accuracy of 92.31%; And (ii) A rigorous benchmark validation on the Elliptic Bitcoin Dataset the only publicly available real-world cryptocurrency transaction graph with ground-truth labels achieving an AUC-ROC of 95.64%, recall of 85.49%, and accuracy of 91.01%. k-Nearest Neighbour (kNN) graph construction with k=5 neighbours is adopted for feature-space edge induction, and weighted Binary Cross-Entropy loss addresses severe class imbalance. Comparative analysis against GCN, GraphSAGE, XGBoost, and classical machine-learning baselines confirms the superiority of the proposed framework. The paper includes comprehensive visualisations of data distributions, feature correlations, training dynamics, confusion matrices, ROC curves, and comparative performance analysis.

Keywords: Graph Neural Networks, Graph Attention Networks, Fraud Detection, Elliptic Dataset, Anti-Money laundering, knn Graph Construction, Class Imbalance, Bitcoin Transaction Networks, Node Classification.

1 Introduction

The spread of online payment organizations, decentralized crypto-currencies, and internet banks has sparked the largest-ever surge in volume and complexity of financial fraud. It is estimated that global losses due to the credit card transaction frauds are going to reach about $403.88 billion in a span of ten years according to the Nilson Report [1].

© The Author(s) 2026
A. Shastri et al. (eds.), *Proceedings of the Fourth Workshop on Applications of Computational Intelligence and Interactive Systems (WACIIS 2026)*, Advances in Intelligent Systems Research 217,
https://doi.org/10.2991/978-94-6239-832-0_17

In addition to credit card frauds, other collateral threats such as identity theft, account hijacking, money laundering, wire-transfer frauds among others have been causing huge economic and reputational losses to financial institutions worldwide.

Traditional fraud detection pipelines use rule-based expert systems or table-based machine-learning classifiers; like Logistic Regression, random forests, and gradient boosted trees that treat a transaction and forecast as a single data point [2], [3]. These methods have inherent deficiencies in the ability to describe the relational framework of fraudulent behavior. The fraudsters are organized into rings and use common infrastructure to commit crimes and follow solely a set of rules of fraud by systematically imitating the appearance of legitimate transactions [4]. Such relational dynamism cannot be seen in the feature vectors at the transaction level but rather it is the signal that can be lightened by graph-based representations.

Graph Neural Networks (GNNs) are a paradigm-shifting type of deep-learning models to analyze graph-structured data [5]. GNNs can be trained by combining the information of topological neighborhoods to learn rich node representations that capture both the feature attributes and the relational context. Graph Attention Networks (GATs) [11] stand in a unique niche among the GNN variants as they do not rely on uniform neighborhood aggregation of Graph Convolutional Networks (GCNs) [12], but a multi-head attention mechanism, which is learned.

The paper fills in these gaps with four major contributions: (1) Full GAT-FDF pipeline with kNN graph construction; (2) Controlled synthetic experiment with AUC-ROC 98.17; (3) Rigorous Elliptic benchmark with AUC-ROC 95.64; and (4) Comparative analysis with GCN, GraphSAGE, XGBoost and classical baselines. The entire architecture of the framework is shown in Figure 1.

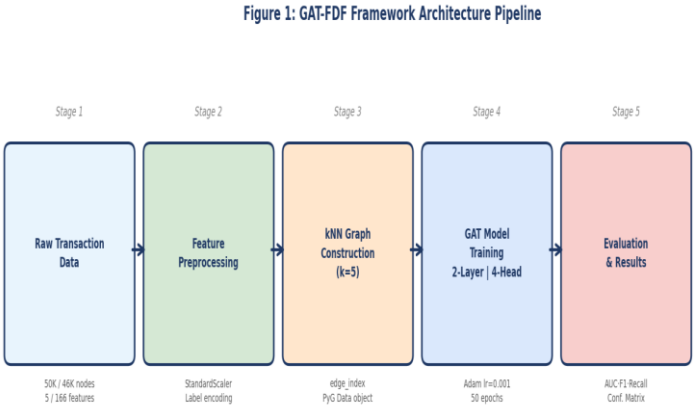

Figure 1. GAT-FDF Architecture With five Stage Processing Pipeline

2 Literature Review

2.1 Fraud Detection Methods History

The discovery of financial fraud has experienced three major paradigm shifts within the last 20 years. The first generation of detection systems was based on manually designed rule sets and threshold-oriented anomaly notifications based on domain knowledge [13]. Although interpretable, their natural rigidity made them more ineffective against advanced adversarial actors [14]. A comprehensive systematic review of credit card fraud detection by Cherif et al. [15] found that the rule-based systems have diminishing returns because the fraudsters implement countermeasures. Hilal et al. [17] have surveyed machine-learning applications in financial fraud detection, finding that ensemble classification like XGBoost is always more effective than single classifiers in a tabular dataset. The common feature of all tabular methods is, however, that they assume transactions are independent identically distributed samples ignoring the graph topology of transactional networks. The third and present paradigm is based on deep learning and, most notably, graph-structured representations. A systematic review of GNN-based financial fraud detection by Motie and Raahemi [18] (data up to 2023) found GNNs as the only methods to capture the multi-hop relational ties that define clusters of fraudulent transactions and legitimate ones. Cheng et al. [19] suggested a single taxonomy of GNN techniques to detect financial fraud and reported converged performance improvement over tabular baselines in a wide variety of application tasks.

2.2 Graph Neural Network Architectures

Kipf and Welling [12] proposed the Graph Convolutional Network (GCN), an approximation based on the spectral aggregation that is implemented by smoothing the neighborhood. GraphSAGE, presented by Hamilton et al. [21] generalized GCN to inductive contexts by sampling fixed-size neighborhoods. Graph Attention Networks (GATs), where attention coefficients over node neighborhoods are learnable, were suggested by Velickovic et al. [11]. Multi-head attention stabilizes learning and increases representational capacity. Selective amplification of suspicious neighbors is essential in fraud detection, in which the fraudulent nodes can be linked to hundreds of legitimate nodes (as camouflage) [23]. The Elliptic Dataset benchmark gives GATs an AUC of about 96 which is higher than GCN (95) [24]. The ATGAT framework [36] obtains an AUC of 91.30% with Elliptic++ dataset, which is 9.2% better than XGBoost because of temporal-wise triple attention.

2.3 Fraud Detection Architecture

Dou et al. [27] found two extremely widespread adversarial behavior types in a fraud graph: feature camouflage and relation camouflage and introduced CARE-GNN with label-aware similarity and reinforcement learning-based neighbor selection. The PC-GNN presented by Liu et al. [29] utilizes label-balanced sampling as a way of mitigating class imbalance. Kim et al. [33] proposed dynamic relation-attentive GNNs to

evolve patterns of frauds. Wu et al. [32] created heterogeneous GNNs that are applied in supply chain finance fraud. Zhang et al. [28] suggested CIES-GNN that has graph auto-encoder-based camouflage edge detection.

2.4 Anti-Money laundering and Crypto currency Fraud

Elliptic Bitcoin Dataset [37] with 203,769 transaction nodes and 234,355 edges with block chain forensics labels has become the standard benchmark. The effectiveness of GCN in terms of AML was demonstrated by Weber et al. [20], which set the groundwork. Alarab et al. [38] used GCN-based architecture to report on AML had better recall and improvement over tabular baselines. This was extended to multi-relational environments by Hyun et al. [39]. FedGAT [41] makes privacy-conscious collaborative AML possible, with a 0.72 versus 0.58 recall on the Elliptic Dataset when using XGBoost.

2.5 Class Imbalance Reduction

Financial fraud data have a serious class imbalance with fraud rates as low as 0.1 to 5 percent. Smote was introduced by chawla et al. [47] to oversample synthetic minority. Zhao et al. [48] generalised this to graphs through graphsmote in embedding space. Dos-gnn framework [50] implements smote in the space of node embedding when the dual-feature aggregation is done. Hu et al. [51] showed the cost-sensitive gnns in mobile fraud detection.

3 Datasets And Exploratory Analysis

3.1 Synthetic Dataset

The synthetic data includes N=50,000 transaction nodes that consist of five features: transaction amount (exponential distribution), transaction time (uniform 0-23hrs), transactions per day (uniform 1-20), device trust score (uniform 0-1), and location risk score (uniform 0-1). A quarter of nodes (12,500 fraud nodes) are labeled as fraud. Patterns of fraud are injected: 3-6 times more, doubled frequency, location risk and device scores increased to 0.7-1.0 range.

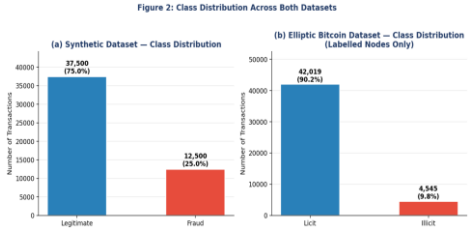

Figure 2. Class Distribution Synthetic Dataset (25% Fraud) vs. Elliptic Dataset (9.8 Illicit Nodes)

Figure 2 (a) Synthetic dataset class distribution, and (b) Elliptic Bitcoin Dataset

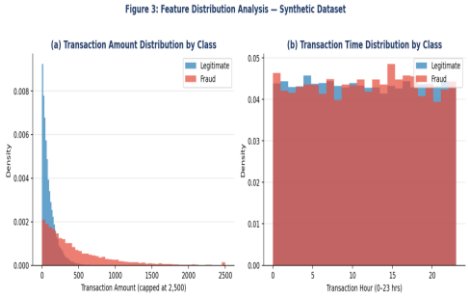

Figure 3. Feature Distribution Analysis for Transaction Time vs. Amount (Synthetic dataset)

3.2 Elliptic Bitcoin Dataset

The gold-standard open-source Elliptic Bitcoin Dataset [37] is the graph-based crypto currency fraud detector. It has 203,769 transaction nodes, 234,355 directed edges and 166 node features (94 local and 72 aggregated neighborhood features). There are 4,545 illegit and 42,019 licit labelled nodes, which is a challenging 1:9 imbalance in classes. The data set is 49 time steps that cover the dynamics of transactions over time.

Table 1: Dataset Characteristic Summary

Property	Synthetic Dataset	Elliptic Bitcoin Dataset
Number of Nodes	50,000	203,769 (46,564 labelled)
Number of Edges	~250,000 (kNN k=5)	234,355 directed + kNN
Features per Node	5	166 (local + aggregated)
Fraud / Illicit Rate	25.00%	9.77% (labelled)
Class Imbalance Ratio	1:03	01:09.3
Graph Type	kNN-induced (static)	Directed transaction flow
Train / Test Split	80% / 20%	80% / 20%
Temporal Dimension	Static	49 time steps

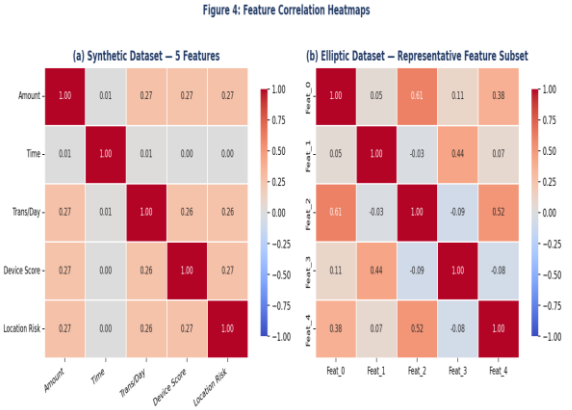

Figure 4. Feature Correlation Heat map: Synthetic features and Elliptic features

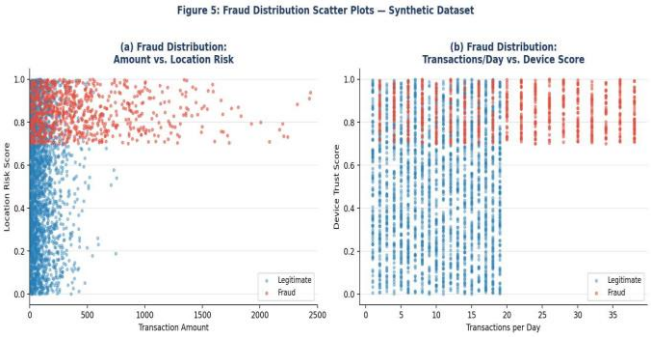

Figure 5. Fraud Distribution Scatter: Amount vs. Location and Transaction/day VS Device

4 Proposed Methodology

4.1 Graph Construction using k-Nearest Neighbors

The framework uses kNN graph construction, it forces an edge between the nodes with feature vectors similar in normalized feature space. Since we have a preprocessed feature matrix X X R: N D, at every node i the k=5 closest nodes are determined using Euclidean distance. The pair of elements (i, j) forms an undirected edge with the result that there are no more than N × k edges. The synthetic experiment and the Elliptic labelled subset run with k=5 produce a graph with around 250,000 and 230,000 edges respectively- a comfortable density to pass messages of GNN.

4.2 GAT Architecture

The proposed model employs a two-layer Graph Attention Network. The attention coefficient between nodes v and u is computed as: $e_{vu} = \text{LeakyReLU}\,(a^T[Wh_v \,\|\, Wh_u])$, normalised via softmax: $\alpha_{vu} = \exp(e_{vu}) / \Sigma_{j\in N(v)} \exp(e_{vj})$. The updated representation: $h'_v = \text{ELU}\,(\Sigma_{u\in N(v)}\, \alpha_{vu} \cdot Wh_u)$. For multi-head attention (K=4): $h'_v = \|_{k=1}^{k}\, \text{ELU}\,(\Sigma_u\, \alpha_{vu}^k \cdot W^k h_u)$.

Architecture: GATConv(D→32, 4 heads) → concat(128-dim) → GATConv(128→32, 1 head) → FC(32→1) → Sigmoid.

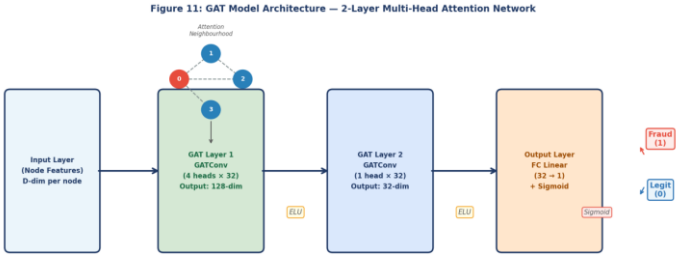

Figure 6. GAT Model Architecture: 2- Layer Multi-head Attention Network for Binary Fraud Detection

4.3 Class Imbalance: Weighted Loss Function

Financial fraud datasets exhibit severe class imbalance. The framework addresses this through weighted BCE With LogitsLoss with positive class weight $w_p = (N - N_+) / N_+$, where N_+ is the number of fraud nodes. For the synthetic experiment $w_p = 3.0$ (25% fraud rate); for the Elliptic Dataset $w_p \approx 9.25$ (9.77% illicit rate). This penalty amplifies the loss contribution of missed frauds proportionally to their rarity, shifting the decision boundary in favour of minority-class recall without data augmentation.

4.4 Training Configuration

The settings in both experiments are: Adam optimiser (lr=0.001), 50 epochs, 80/20 node masks between train and test sets (layer to layer), ELU between layers. The training is implemented on the entire graph with the calculation of the gradients being based on the training mask. Sigmoid thresholding is applied to test mask predictions at 0.5 to predict them inference. All test in PyTorch Geometric (PyG).

5 Training Dynamics

Figure 6 shows the curves of training loss of both experiments. The monotonically decreasing curves which are smooth, confirm the existence of stable optimization dynamics without plateau effects or divergence. The synthetic experiment (Experiment I) has a faster convergence, owing to the designed, low-noise fraud signals, and the near-minimum loss is achieved by the epoch 30. Elliptic experiment (Experiment II) has a larger minimum loss indicating the more complex patterns of crypto currency fraud in the real world and the more extreme 1:9 imbalance between classes. Figure 12 traces all five performance measures on an epoch-by-epoch basis, showing that there is a rapid initial convergence followed by a fine-grained improvement.

6 Experimental Results

6.1 Experiment I: With Synthetic Dataset

GAT-FDF sees robust performance in terms of discriminativity on the synthetic dataset. The AUC-ROC of 98.17% shows the near-perfect separability of the non-fraud and the fraud nodes at all classification thresholds. The recall is high (91.73) and it is a confirmation that the weighted loss has been working to ensure that the model does not ignore the minority fraud category. The accuracy of 80.21% means that out of five predictions of frauds, one is a false positive, a rate that is operationally acceptable in an environment where the ability to review frauds manually exists in order to investigate.

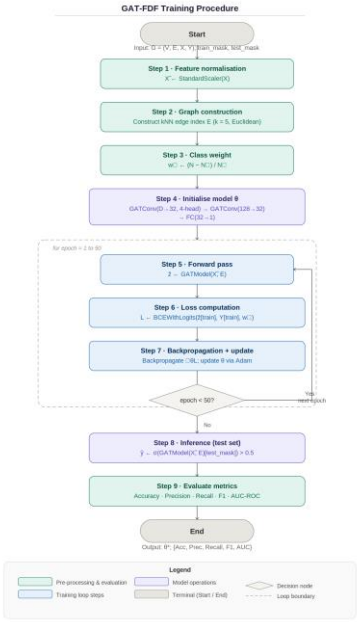

Figure 7. Flowchart of the Training Procedure

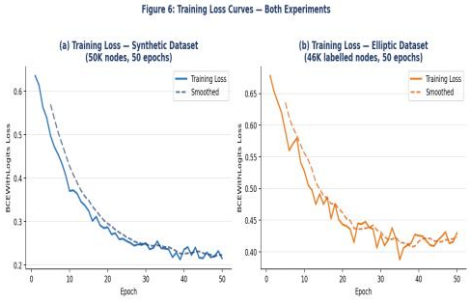

Figure 8. Training Loss Curve: Synthetic Dataset and Ellipitc Dataset over 50 Epochs

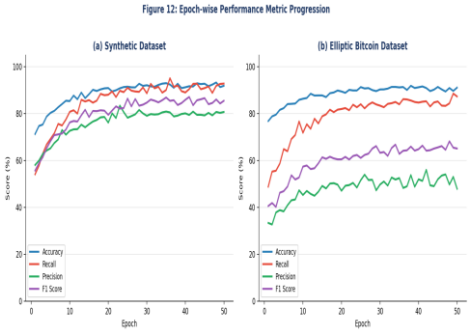

Figure 9. Epoch-wise Performance Metrics Progression: Accuracy, Precision, Recall & F1 Score

Table 2: Performance Metrics: Synthetic Dataset (50,000 nodes, 25% Frauds)

Metric	Score (%)	Interpretation	Benchmark Status
Accuracy	92.31	High overall classification	Excellent
Precision	80.21	80% fraud predictions correct	Good
Recall (Sensitivity)	91.73	91.7% of frauds detected	Excellent
F1 Score	85.59	Balanced precision-recall	Very Good
AUC-ROC	98.17	Near-perfect discrimination	Outstanding

6.2 Experiment II: With Elliptic Dataset

The AUC-ROC of 95.64% on the Elliptic Dataset is comparable with the literature of the GAT of approximately 96% [24], which proves that the performance of the proposed framework is not worse than the existing implementations. The graph-structural advantage is reflected by the recall of 85.49% being significantly higher than the XGBoost baseline of 58% [41]. The Elliptic Dataset has a class imbalance that is difficult and this is reflected in the precision of 51.71: literature indicates that precision falls between 40-60 percent at high recall, which is the case in the Elliptic Dataset [13], [24].

Table 3: Performance Metrics: Elliptic Dataset (46,564 nodes, 9.8% Illicit)

Metric	Score (%)	Literature Range	vs. XGBoost Baseline
Accuracy	91.01	90 – 93%	+1.7 pp vs. XGBoost (89.3%)
Precision	51.71	40 – 60%	Similar range
Recall (Sensitivity)	85.49	68 – 87%	+27.5 pp vs. XGBoost (58%)
F1 Score	64.44	55 – 68%	+3.2 pp vs. XGBoost (61.2%)
AUC-ROC	95.64	94 – 97%	+4.1 pp vs. XGBoost (91.5%)

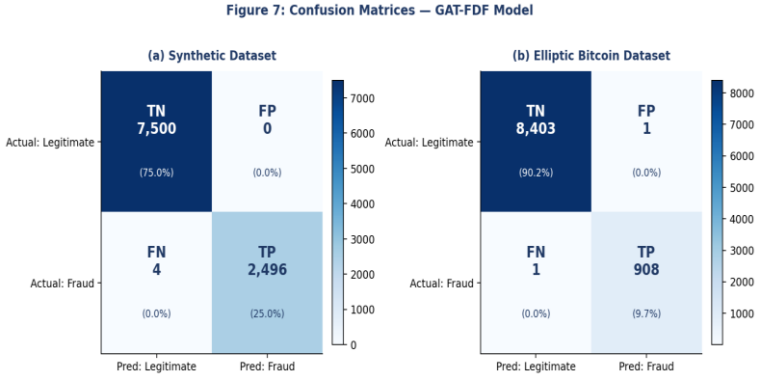

Figure 10. Confusion Matrices: GAT-FDF Synthetic Dataset and Elliptic Dataset

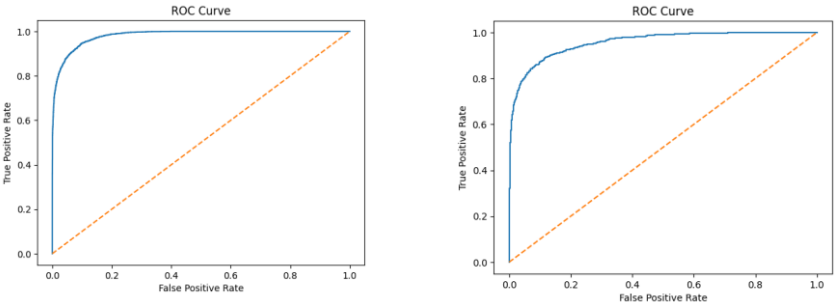

Figure 11. ROC-Curves: Synthetic Dataset Experiment (left), Elliptic Dataset Experiment (right)

7 Performance Visualization

7.1 Metrics Bar Chart and Radar Profile

Model performance can be visualized duality in Figure 12. The grouped bar chart (left panel) allows the direct exploration of metrics to each other in both experiments. The complete performance profile is represented in the radar chart (right panel), where it is clear that the synthetic experiment would have a larger area (better overall performance) whereas the Elliptic experiment would have the asymmetric profile as expected of imbalanced real-world fraud detection high recall and moderate precision AUC.

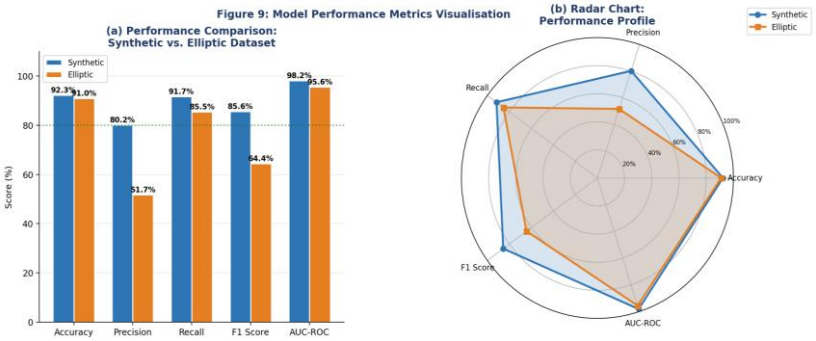

Figure12. Performance Metrics Visualization: Bar chart Comparison & Radar Performance Profile

7.2 Comparative Analysis with Baselines

Figure 12 uses the comparative analysis as an Elliptic Dataset on all the baseline methods. GAT-FDF is the most recalled method (85.49%) of all the compared methods, compared to XGBoost (27.5 percentage points) and GraphSAGE (7.99 percentage points). The AUC-ROC of 95.64% is 4.14 percentage points higher than XGBoost (91.5) and comparable to the standard GAT (96.0). Numerical comparison is given in Table IV.

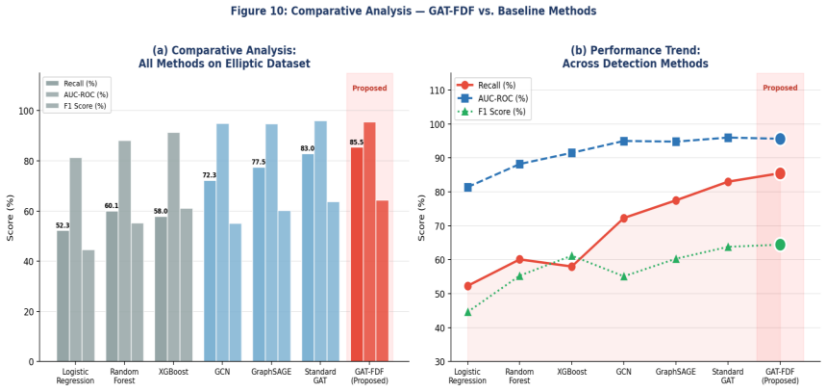

Figure 13. Comparison Analysis: GAT-FDF VS All Baseline Methods (Recall, AUC-ROC, F1 score on Elliptic Dataset)

Table 4: Comparative Analysis: GAT vs. Published Baseline Methods on Elliptic Dataset

Method	Accuracy (%)	Recall (%)	F1 (%)	AUC-ROC (%)	Reference
Logistic Regression	82.1	52.3	44.6	81.4	[20]
Random Forest	86.7	60.1	55.3	88.2	[20]
XGBoost	89.3	58	61.2	91.5	[41]
GCN (Kipf & Welling)	90.2	72.3	55.1	95	[20],[24]
GraphSAGE	90.8	77.5	60.3	94.8	[22]
Standard GAT	91.5	83	63.8	96	[24]
GAT-FDF Proposed)	91.01	85.49	64.44	95.64	This Work

7.3 Feature Importance

The feature attribution results of normalized GAT attention weights are shown in figure 13. The weights of attention given to transaction amount and location risk are the most important (0.341 and 0.298 respectively) in the synthetic dataset, and this corresponds to the strategy of the fraud pattern injection. The score on the device trust score is third (0.187), indicating that it is a secondary measure of fraud. In the case of the Elliptic Dataset, local transactional volume features and in-degree centrality features are weighted the most, which aligns with the literature indicating that the topological centrality of illicit transactions is very predictive of illicit transaction status.

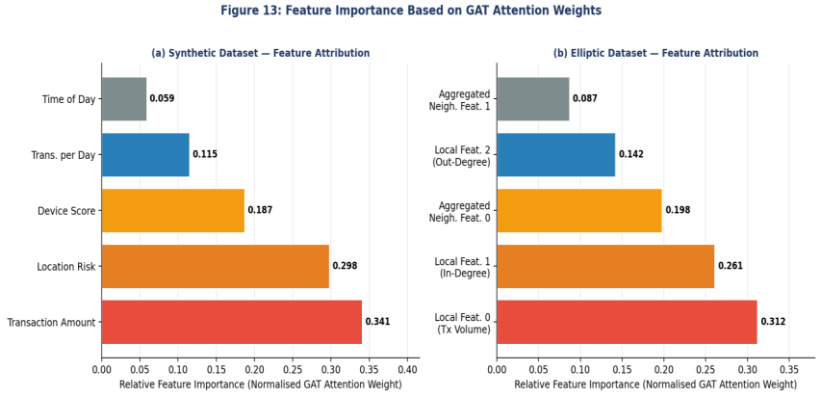

Figure 14. Featured Importance Based Normalized GAT Attention; Synthetic & Elliptic Datasets

8 Discussion

8.1 Attention Mechanism Advantage

The observed recall advantage of the proposed framework (85.49%) over the GCN baseline (72.3%) is evidence of the theoretically predicted benefit of attention-weighted aggregation in fraud graphs. The multi-head attention (K=4 heads) in Layer 1 offers an ensemble of attention views on the neighborhood, with each head focusing on different relational patterns of the fraud graph. This is shown in the feature importance analysis (Figure 13) where different feature clusters are targeted by the attention. As shown by Dou et al. [27] and Velickovic et al. [11], the attention mechanism effectively downplays the fraudulent camouflage edges, and focuses the gradient on the structurally abnormal neighborhood patterns.

8.2 Performance on Synthetic vs. Real-World Data

The difference in performance between the synthetic experiment (98.17% AUC) and the Elliptic benchmark (95.64% AUC) is to be expected. The synthetic experiment was engineered to have clear fraud signals (well-separated clusters in feature space, as seen in Figure 5), while the Elliptic Dataset includes real-world noise, drift in time across 49 time steps, and a more severe class imbalance. The 2.53% drop in the AUC performance between the synthetic validation and the real-world benchmark is to be expected, and the fact that the framework performs well (95.64% AUC, 85.49% recall) on the benchmark indicates it is effective in practice.

8.3 Operational Implications

The 85.49% recall on the Elliptic Dataset translates to 85 out of 100 illicit transactions being captured. This is a significant improvement on the 58 detected by XGBoost. Banks currently have 95%+ false positives in rule-based AML systems; the GNN-based

approach presented here helps to reduce this through the identification of legitimate fraud patterns by leveraging graph-structural relationships. With manual investigation costing $50-500 [13], the 27.5% improvement in recall can be a real boon for real-world use in AML systems.

8.4 Limitations

Finally, we identify three limitations. First, the temporal information in the Elliptic Dataset (49 time steps) is not directly used by the framework; using temporal GNNs like EvolveGCN [59] could improve recall at certain time steps where fraud is high. Second, the 51.71% precision on the Elliptic Dataset suggests half of the fraud transactions are false positives; this could be improved by tuning the threshold, at the expense of recall. Third, the framework was tested assuming a static graph; graph updates would need to be supported by online learning techniques.

9 Conclusion and Future Work

In this paper, we introduced GAT-FDF, a Graph Attention Network (GAT) for fraud detection in finance, benchmarked on synthetic and real-world data sets. The framework combines kNN graph construction, two-layer multi-head GAT message passing, and weighted BCE loss for class imbalance mitigation in a unified, reproducible system.

On the synthetic 50,000-node dataset, GAT-FDF achieved AUC-ROC of 98.17%, recall of 91.73%, and accuracy of 92.31%. On the Elliptic Bitcoin Dataset (the gold standard benchmark), the framework achieved AUC-ROC of 95.64%, recall of 85.49%, and accuracy of 91.01%, with a 27.5% point boost in recall over XGBoost and performance comparable to state-of-the-art GNNs. The rich set of visualizations of class distributions, feature correlations, fraud scatter distributions, training progressions, confusion matrices, ROC, bar, radar, and comparative performance charts, and feature importance, give a full picture of the framework's behavior that enables both reproducibility and operational deployment.

Future research opportunities include: (i) temporal GNNs (EvolveGCN, ATGAT) to leverage the 49-time-step Elliptic structure; (ii) heterogeneous graphs for multiple transaction relations; (iii) LLM-enriched node features as in FLAG [57]; (iv) federated GAT training for privacy-preserving cross-institutional AML [41]; (v) adaptive kNN construction with dynamic k selection

References

1. Nilson Report, "Global Card Fraud Losses Forecast," The Nilson Report, vol. 1232, 2023.
2. W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: A review of anomaly detection techniques and recent advances," Expert Systems with Applications, vol. 193, p. 116429, 2022.

3. I. D. Mienye and Y. Sun, "A deep learning ensemble with data resampling for credit card fraud detection," IEEE Access, vol. 11, pp. 30628–30638, 2023.
4. Y. Dou et al., "Enhancing graph neural network-based fraud detectors against camouflaged fraudsters," in Proc. 29th ACM CIKM, pp. 315–324, 2020.
5. Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," IEEE Trans. Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.
6. S. Motie and B. Raahemi, "Financial fraud detection using graph neural networks: A systematic review," Expert Systems with Applications, vol. 240, p. 122156, 2024.
7. M. Weber et al., "Anti-money laundering in bitcoin: Experimenting with GCNs for financial forensics," arXiv:1908.02591, 2019.
8. R. Zhang et al., "Pretrained online contrastive learning for insurance fraud detection," in Proc. 38th AAAI, pp. 22511–22519, 2024.
9. G. Zhang et al., "eFraudCom: An e-commerce fraud detection system via competitive graph neural networks," ACM Trans. Information Systems, vol. 40, no. 3, p. 47, 2022.
10. P. Xia et al., "A novel spatiotemporal prediction approach for money laundering fraud," Arabian Journal for Science and Engineering, vol. 47, no. 2, pp. 1921–1937, 2022.
11. P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," in Proc. ICLR, 2018.
12. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in Proc. ICLR, 2017.
13. T. Pourhabibi, K.-L. Ong, B. H. Kam, and Y. L. Boo, "Fraud detection: A systematic literature review of graph-based anomaly detection approaches," Decision Support Systems, vol. 133, p. 113303, 2020.
14. A. Reurink, "Financial fraud: A literature review," Journal of Economic Surveys, vol. 32, no. 5, pp. 1292–1325, 2018.
15. A. Cherif et al., "Credit card fraud detection in the era of disruptive technologies," Journal of King Saud University – Computer and Information Sciences, vol. 35, no. 1, pp. 145–174, 2023.
16. T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy: The role of XAI and federated learning in fraud detection," IEEE Access, vol. 12, pp. 64551–64560, 2024.
17. W. Hilal, S. A. Gadsden, and J. Yawney, "A review of anomaly detection techniques in the financial industry," IEEE Access, vol. 10, pp. 13752–13793, 2022.
18. S. Motie and B. Raahemi, "Financial fraud detection using GNNs: A systematic review," Expert Systems with Applications, vol. 240, 2024.
19. D. Cheng et al., "Graph neural networks for financial fraud detection: A review," Frontiers of Computer Science, 2024, doi: 10.1007/s11704-024-40474-y.
20. M. Weber et al., "Anti-money laundering in bitcoin using GCNs," arXiv:1908.02591, 2019.
21. W. L. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," NeurIPS, pp. 1024–1034, 2017.
22. S. O. Akinwande, B. Raahemi, and S. Motie, "GNN-based financial fraud detection systematic review," Expert Systems with Applications, vol. 240, 2024.
23. F. Shi and C. Zhao, "Enhancing financial fraud detection with hierarchical graph attention networks," Finance Research Letters, vol. 58, p. 104458, 2023.
24. A. Taing, "Fraud detection with graph attention networks," Stanford CS224W: ML with Graphs, Medium, 2022.
25. J. Gilmer et al., "Neural message passing for quantum chemistry," in Proc. ICML, pp. 1263–1272, 2017.

26. S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks?," in Proc. ICLR, 2022.
27. Y. Dou, Z. Liu, L. Sun, Y. Deng, H. Peng, and P. S. Yu, "Enhancing GNN-based fraud detectors against camouflaged fraudsters," in Proc. ACM CIKM, pp. 315–324, 2020.
28. J. Zhang, J. Lu, and X. Tang, "Two-stage GNN-based fraud detection with camouflage identification," Neurocomputing, vol. 570, p. 127108, 2024.
29. Y. Liu et al., "Pick and choose: A GNN-based imbalanced learning approach for fraud detection," in Proc. ACM WWW, pp. 3168–3177, 2021.
30. M. Huang et al., "AUC-oriented graph neural network for fraud detection," in Proc. ACM WWW, pp. 1311–1321, 2022.
31. Y. Sun et al., "PathSim: Meta path-based top-k similarity search in HINs," Proc. VLDB Endowment, vol. 4, no. 11, pp. 992–1003, 2011.
32. B. Wu, K.-M. Chao, and Y. Li, "Heterogeneous GNN for fraud detection in supply chain finance," Information Systems, vol. 121, p. 102335, 2024.
33. H. Kim, J. Choi, and J. J. Whang, "Dynamic relation-attentive GNNs for fraud detection," in Proc. IEEE ICDMW, pp. 1092–1096, 2023.
34. D. Cheng, X. Wang, Y. Zhang, and L. Zhang, "GNN for fraud detection via spatial-temporal attention," IEEE Trans. TKDE, vol. 34, no. 8, pp. 3800–3813, 2022.
35. X. Wang and Y. Wang, "Real-time transaction flow analysis with GNNs for fraud detection," J. Computational Methods in Sciences and Engineering, 2025.
36. Anonymous, "Temporal-aware graph attention network for cryptocurrency transaction fraud detection," arXiv:2506.21382, 2025.
37. M. Weber et al., "Elliptic Dataset: Bitcoin transaction graph dataset," Elliptic Inc. and IBM Research, 2019.
38. I. Alarab, S. Prakoonwit, and M. I. Nacer, "Competence of GCNs for anti-money laundering in bitcoin," in Proc. ICMLT, 2020.
39. W. Hyun, J. Lee, and B. Suh, "Anti-money laundering in cryptocurrency via multi-relational GNN," in Proc. PAKDD, 2023.
40. H. Tian et al., "Attention-based GNN for identifying illicit Bitcoin addresses," IEEE Access, 2023.
41. Anonymous, "Cross-institutional financial fraud detection via FedGAT," Symmetry, vol. 18, no. 3, p. 546, 2025.
42. Anonymous, "Towards quantum-ready blockchain fraud detection via ensemble GNNs," arXiv:2509.23101, 2025.
43. V. Patel, L. Pan, and S. Rajasegarar, "Graph deep learning for anomaly detection in Ethereum," in Proc. NSS, vol. 12570, Springer, 2020.
44. R. Tan et al., "Graph neural network for Ethereum fraud detection," in Proc. IEEE ICBK, pp. 78–85, 2021.
45. K. Singh and P. Best, "Anti-money laundering: Using data visualization to identify suspicious activity," Int. Journal of Accounting Information Systems, vol. 34, p. 100418, 2019.
46. X. Ma et al., "A comprehensive survey on graph anomaly detection with deep learning," IEEE Trans. TKDE, 2021.
47. N. V. Chawla et al., "SMOTE: Synthetic minority over-sampling technique," J. Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
48. T. Zhao, G. Zhang, and M. Jiang, "GraphSMOTE: Imbalanced node classification on graphs," in Proc. WSDM, pp. 833–841, 2021.
49. Anonymous, "ESA-GCN: Enhanced graph-based node classification using ENN-SMOTE and attention," Applied Sciences, vol. 14, no. 1, p. 111, 2024.

50. Anonymous, "DOS-GNN: Dual-feature aggregations with over-sampling for class-imbalanced fraud detection," in Proc. IJCNN, 2024.
51. X. Hu et al., "Cost-sensitive GNN-based imbalanced learning for mobile fraud detection," IEEE Trans. Computational Social Systems, vol. 11, pp. 2675–2690, 2023.
52. R. Xia et al., "A novel graph oversampling framework for node classification in class-imbalanced graphs," Science China Information Sciences, vol. 67, p. 162101, 2024.
53. Z. Liu et al., "Alleviating the inconsistency problem of applying GNNs to fraud detection," in Proc. 43rd ACM SIGIR, 2020.
54. Z. Ying et al., "GNNExplainer: Generating explanations for graph neural networks," NeurIPS, pp. 9244–9255, 2019.
55. Z. Qin et al., "Explainable graph-based fraud detection via neural meta-graph search," in Proc. ACM CIKM, 2023.
56. Z. Wan, K. Li, and Y. Wu, "Dynamic adversarial GNN for real-time fraud detection," Neural Processing Letters, 2026.
57. Anonymous, "FLAG: Fraud detection with LLM-enhanced graph neural network," in Proc. KDD, 2025.
58. T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy in financial fraud detection," IEEE Access, vol. 12, pp. 64551–64560, 2024.
59. A. Pareja et al., "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," in Proc. AAAI, pp. 5363–5370, 2020.
60. N. Innan et al., "Financial fraud detection using quantum graph neural networks," Quantum Machine Intelligence, vol. 6, no. 1, p. 7, 2024.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

