Hybrid Graph Representation Learning for Money Laundering Detection

Author: Marius Frija (mfrija@tudelft.nl)
Responsible Professor: Dr. Kubilay Atasu
Supervisor: Halil Çağrı Bilgi

TU Delft

1. Introduction

- **Money laundering** is the process of making illegally obtained funds appear legitimate by moving them through a **series of transactions** and ultimately reintroducing them into the legal financial system. Traditional anti-money laundering systems suffer from high false positive rates and lack large-scale analytical capabilities.
- Financial transacation data can be modelled as **graphs**:
  - **Edges**: financial transactions
  - **Nodes/vertices**: financial entities

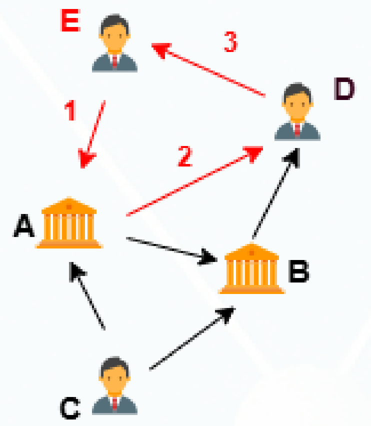

Figure 1: Money laundering cyclic pattern

- **Graph Neural Networks** (GNNs) have recently shown strong potential in modeling transaction networks and detecting suspicious patterns through the local message passing mechanism [1, 2].

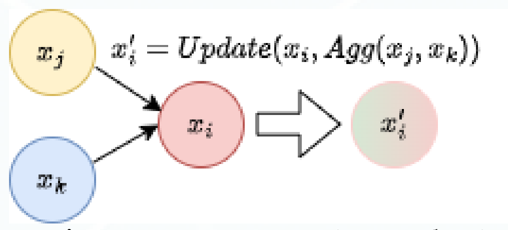

Figure 2: Message Passing Mechanism

- The emergence of **Graph Transformers** (GTs) [4, 6] introduces attention-based mechanisms that overcome the limitations of GNNs by modeling more intricate long-range dependencies between the components of graphs. However, unifying the GNN and GT paradigms into a hybrid model for money laundering detection has been little explored.

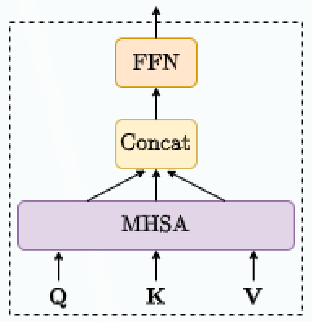

Figure 3: Architecture of a Transformer block

- This research proposes GraphFuse, a **hybrid architecture** that unifies GNNs with a GT to generate rich edge-level representations for detecting money-laundering activities in financial transaction graphs. Furthermore, we investigate three configurations of the GT module for achieving a trade-off between performance and computational complexity.

2. Research Questions

(1) How does the **late-fusion GNN-GT architecture** compare to standalone GNN and Graph Transformer models in terms of money laundering detection?

(2) Under which global attention mechanism **linear**, **sparse** or **full** does the hybrid model achieve the **best trade-off** between **performance** and computational **efficiency**?

References

[1] Egressy, B., von Niederhäusern, L., Blanuša, J., Altman, E., Wattenhofer, R., and Atasu, K. (2024). Provably Powerful Graph Neural Networks for Directed Multigraphs. Proceedings of the AAAI Conference on Artificial Intelligence.
[2] Bilgi, H. Ç., Chen, L. Y., and Atasu, K. (2024). Multigraph Message Passing with Bi-Directional Multi-Edge Aggregations (MEGA-GNN). arXiv preprint arXiv:2412.00241.
[3] Corso, G., Cavalleri, L., Beaini, D., Liò, P., & Veličković, P. (2020). Principal Neighbourhood Aggregation for Graph Nets. Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 13260–13271.
[4] Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., Shen, Y., & Liu, T.-Y. (2021). Do Transformers Really Perform Badly for Graph Representation?. NeurIPS, 34, 28877–28888.
[5] Altman, E,. Blanuša, J., von Niederhäusern, L., Egressy, B., Anghel, A., and Atasu, K. (2023). Realistic Synthetic Financial Transactions for Anti-Money Laundering Models. NeurIPS 2023 Datasets and Benchmarks Track. arXiv:2306.16424.s
[6] Lin, J., Guo, X,. Zhu, Y., Mitchell, S., Altman, E., and Shun, J. (2024). Fraudgt: A simple, effective, and efficient graph transformer for financial fraud detection. In Proceedings of the 5th ACM International Conference on AI in Finance (ICAIF). ACM, 2024.

3. Proposed Model

Local Message Passing Module

- Captures local patterns among nearby accounts and their transactions leveraging message passing.
- GNN Backbones: PNA [3] or MEGA-PNA [2] with edge features.

Global Attention Module

- Attention computation on edges captures more intricate illicit patterns in the extended sub-graph.
- Three GA mechanisms: Linear, Sparse and Full.

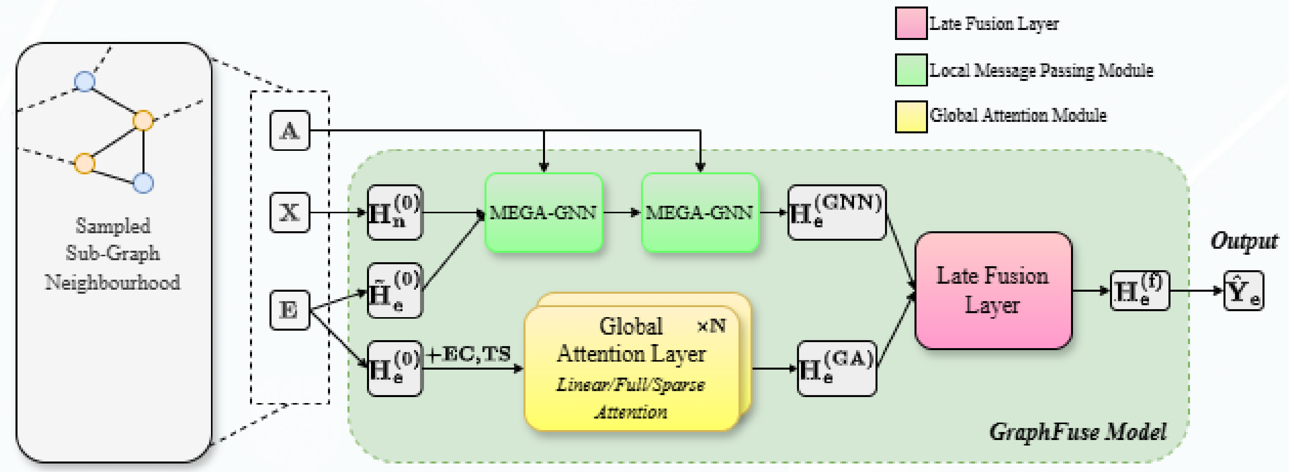

Figure 4: Architecture of the GraphFuse model.

Late Fusion Layer

- Fuses together the local and global representations, generating meaningful edge-level embeddings for the money laundering detection task.
- Simple yet effective MLP(Concat(.)) layer.

Edge Encodings

- Edge Centrality - captures edge importance based on the centrality of its source and destination node.
- Transaction Signature - injects relevant information about the financial activity of associated accounts.

4. Experimental Setup & Results

Classification Results

	AML Small-HI	AML Small-LI
GNN Baseline		
PNA [2]	61.20 ± 2.24	16.10 ± 2.38
MEGA-GNN Baseline		
MEGA-PNA [7]	73.10 ± 1.46	44.87 ± 1.62
+ Ego IDs	73.74 ± 1.55	45.37 ± 1.45
GT Baseline		
Multi-FraudGT [14]	76.13 ± 0.95	47.01 ± 2.22
GraphFuse-PNA		
Linear Attention	66.12 ± 2.29	30.39 ± 2.24
Sparse Attention	64.24 ± 2.06	28.83 ± 1.94
Full Attention	64.29 ± 3.07	26.30 ± 1.08
GraphFuse-MEGA-PNA		
Linear Attention (w/o EC,TS)	75.81 ± 1.14	46.66 ± 0.37
+ EC	75.86 ± 1.56	46.05 ± 1.06
+ TS	75.72 ± 0.85	47.13 ± 0.19
+ EC + TS	**76.89 ± 0.88**	**47.56 ± 0.36**
Sparse Attention	76.29 ± 0.99	47.47 ± 0.17
Full Attention	76.40 ± 0.28	46.93 ± 0.60

Table 1: Classification performance (F1 score (%) ± std). Standard deviations are calculated over 5 runs with different random seeds.We highlight the best and second-best results..

Training Efficiency and Fraud Detection Performance Analysis

	Size (# params)	Epoch time (s)	Avg. Throughput (trans/s)	Test F1 (%)
GraphFuse-MEGA-PNA				
Linear Attention	$212.6 \cdot 10^3$	626.7 ± 4.0	$\approx 53 \cdot 10^3$	76.89 ± 0.88
Sparse Attention	$162.5 \cdot 10^3$	1441.9 ± 23.9	$\approx 21 \cdot 10^3$	76.29 ± 0.99
Full Attention	$112.2 \cdot 10^3$	1648.1 ± 2.9	$\approx 15 \cdot 10^3$	76.40 ± 0.28

Table 2: Training and estimative inference efficiency comparison of the three GraphFuse model variants with different Global Attention Module configurations.

Observations

- The effectiveness of the hybrid model is supported by the improvement upon PNA and MEGA-PNA across both datasets and for all attention configurations.
- A notable gain of 0.76% over the state-of-the-art Multi-FraudGT model demonstrates the efficacy of the late-fusion GNN-GT architecture.
- The Linear Attention model variant achieves a favorable performance-efficiency trade-off.

Experimental Setup

- Datasets: IBM AML datasets [5] of synthetic & realistic financial transactions between individuals, companies and banks (Small-HI and Small-LI).
- Baselines: PNA [3], MEGA-PNA [2], Multi-FraudGT [6].
- Evaluation Metric: Minority (illicit) class F1 score.

5. Conclusions and Future Work

This research has demonstrated the effectiveness of unifying GNNs with GTs for money laundering detection tasks. Our proposed hybrid model closely surpassess state-of-the-art performance by 0.76 p.p. on large synthetic datasets. Furthermore, we introduced three attention-wise configurations that offer a trade-off between scalabilty and fraud detection performance.

Future Work

- Considering structure-aware positional encodings for increasing the expressivity of the Global Attention Module.
- Incorporating the Reverse Message Passing [1] mechanism and investigating more complex Late Fusion mechanisms.
- Assessing the effectiveness of the model on real-world transactional data.