The Impact of Realistic Laundering Subgraph Perturbations on Graph Neural Network Based Anti-Money Laundering Systems

Author:
Tom Joshua Clark
t.j.clark@tudelft.nl

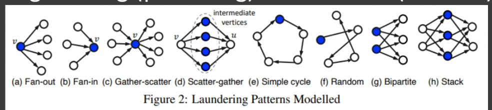
Supervisors:
Dr. Zeki Erkin, z.erkin@tudelft.nl
Dr. Kubilay Atasu kubilay.atasu@tudelft.nl

Introduction

- 2–5% of the global GDP is laundered annually [1].
- Traditional KYC checks and rule-based monitoring struggle to detect complex laundering schemes.
- Future AML investigations will likely have insights into comprehensive transaction graphs across institutions.
- As AML investigations become more effective, money laundering tactics are likely to evolve, leading to the emergence of more complex laundering patterns.

Background

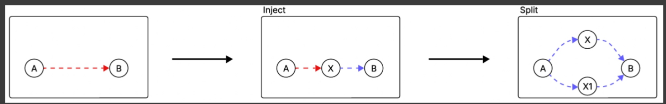
- Common Money Laundering Patterns include fan-in/out, scatter-gather/gather-scatter, cycles, bipartite flows, and stacks [2].
- Graph Neural Networks are a powerful tool for analyzing relational data and uncovering hidden relationships.
- Synthetic Data generates labeled transaction data for analyzing money laundering patterns, benchmarking AML techniques, and training ML models.
- Adversarial attacks exploit model sensitivity by making changes to graph data to degrade detection performance during training (poisoning) or inference (evasion) [3].



Laundering Patterns from [2]

Methodology and Experimental Setup

- We cluster and perturb only the laundering subgraphs in the transaction graph under the assumption that these are the nodes and edges an adversary has control over.
- We introduce a graph perturbation framework to modify laundering subgraphs using three parameterized actions: Inject Intermediary Nodes, Merge Nodes, and Split Nodes, which simulate real-world realistic laundering tactics.
- We evaluate how four distinct perturbation presets (that increase and decrease complexity) used to generate 78 perturbed variants affect the F1-score of pre-trained state-of-the-art MEGA-GNN models [4].



Scatter-Gather Pattern simulated with Preset 4

Conclusions and Future Work

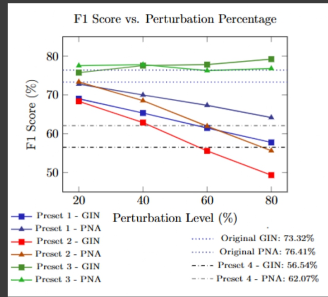
- We introduce a novel perturbation framework that can degrade GNN performance while maintaining plausible behavior by preserving the shape of patterns, highlighting the importance of evaluating model adaptability under various structural changes.
- Future work could expand the framework to simulate more varied adversary behavior and implement rule-based perturbations.
- Future work could involve designing more experiments and presets to investigate how different actions and types of perturbations affect the resilience and generalization capabilities of various GNN architectures.

Results and Discussion

- Structural perturbations can impact and degrade the performance of GNN models while maintaining the operational plausibility and realism of laundering patterns.
- The PNA model performed considerably better than the GIN model on both the original dataset and the perturbed datasets.
- Our framework can generate data for "what-if" scenarios by perturbing synthetic datasets, which allows us to investigate how laundering strategies may evolve.
- Increasing structural complexity decreases performance

Transaction Dataset	GIN	PNA
HI-Small (Original)	73.32% ± 0.06	76.41% ± 0.09
Preset 4 (Inject+Split)	56.54% ± 0.75	62.07% ± 0.52

F1 Scores (%) for GIN and PNA on original dataset and dataset perturbed by Preset 4



F1 Scores (%) across perturbation levels

References

[1]United Nations Office on Drugs and Crime, "Money laundering." [Online]. Available: https://www.unodc.org/unodc/en/money-laundering/overview.html
[2] E. Altman, J. Blanuˇsa, L. von Niederh¨ausern, B. Egressy, A. Anghel, and K. Atasu, "Realistic synthetic financial transactions for anti-money laundering models," 2024. [Online]. Available: https://arxiv.org/abs/2306.16424
[3] J. Ma, S. Ding, and Q. Mei, "Towards more practical adversarial attacks on graph neural networks," 2021. [Online]. Available: https://arxiv.org/abs/2006.05057
[4] H. C, a˘grı Bilgi, L. Y. Chen, and K. Atasu, "Multigraph message passing with bi-directional multi-edge aggregations," 2024. [Online]. Available: https://arxiv.org/abs/2412.00241