FraudLens: Graph Structural Learning for Bitcoin Illicit Activity Identification

Jack Nicholls, Dr. Aditya Kuppa, Assoc. Prof. Dr. Nhien-An Le-Khac

University College Dublin

School of Computer Science,

Ireland

- Crypto in the news:
  - June 2022: Binance enabled $2.35billion in laundering.
  - 2023: $500m in ransomware payments.
  - Tornado cash: $1billion laundered crypto.

- Increasing regulation on transparency and trading.

- Research focuses on GNN variations and enhancements rather than preprocessing and topology imbalance.

Binance Founder Pleads Guilty to Violating Money Laundering Rules

OFAC Sanctions Russian National Ekaterina Zhdanova for Using Cryptocurrency to Launder Money on Behalf of Russian Elites and Ransomware Groups

Crypto crime hits record $20 bln in 2022

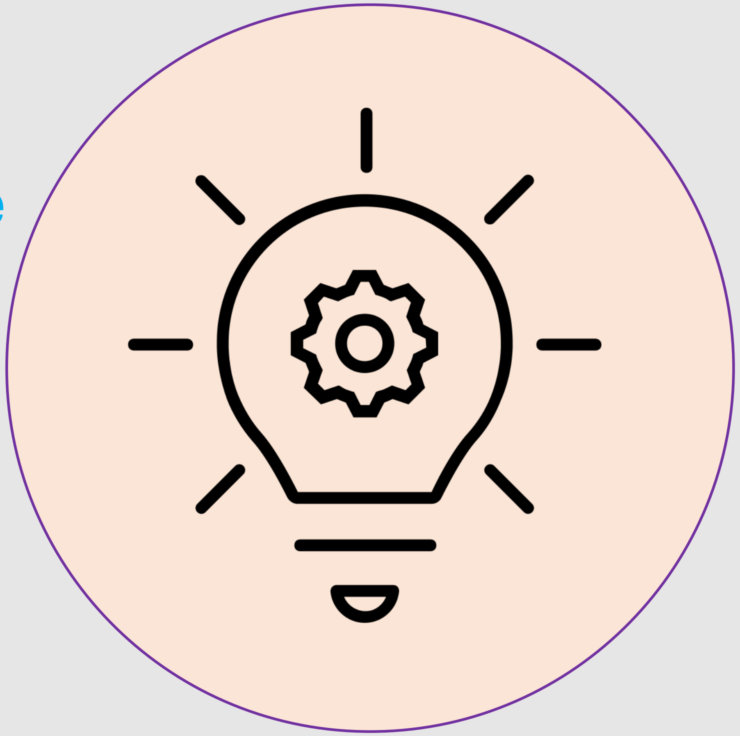
Contributions

- Identify label and topology imbalance issues impacting **GNN models** in identifying illicit activity in bitcoin.

- Propose two novel model-agnostic methods for graph structure learning that address the imbalances and discover fraudulent nodes in bitcoin transaction graphs.

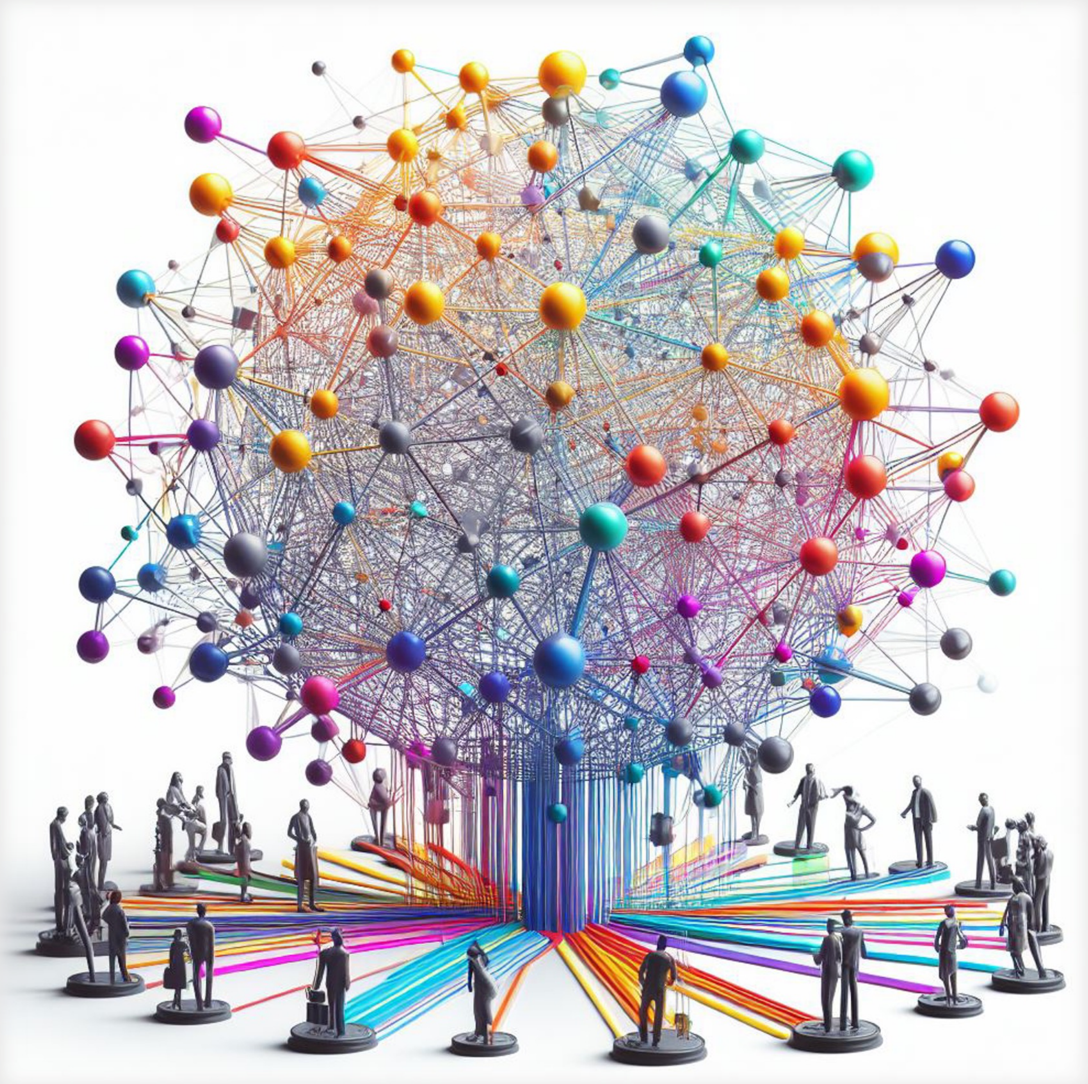
- Evaluate methods on a **highly imbalanced** and temporal **Elliptic Bitcoin dataset** to show performance improvement.

- **Compare** methods against **other imbalanced node classification** techniques on DBLP citation network to show **effectiveness**.



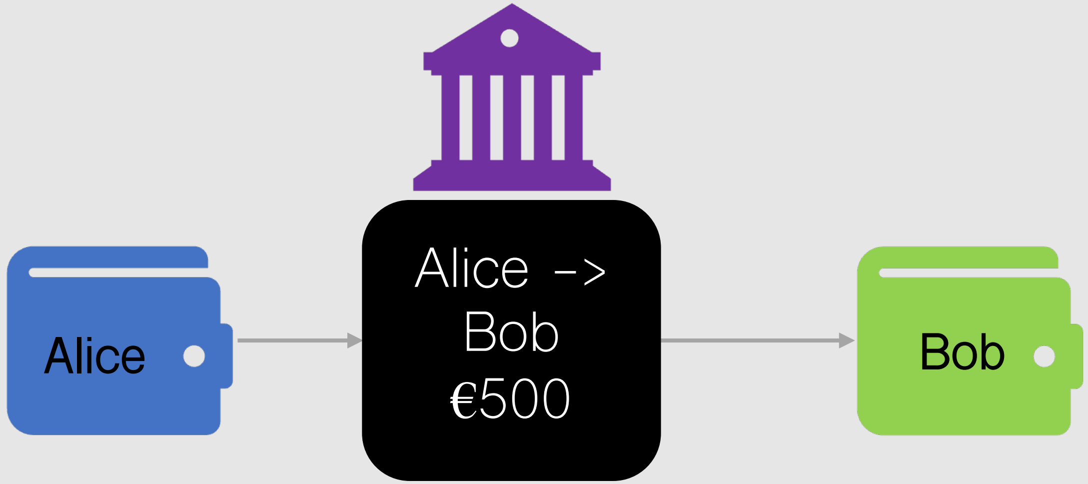
Cryptocurrency

- Computational method of transferring digital value between users.
- Does not require financial intermediary.
- Introduced blockchain technology.
- Two main models of development:
  - UTxO.
  - Account-based.
- Basis for digital currency.

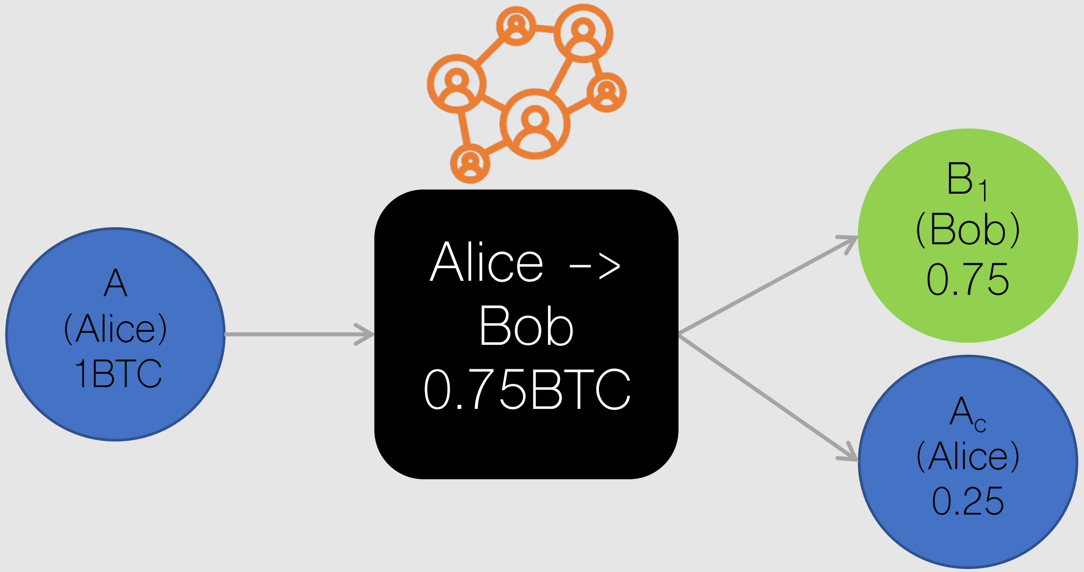


Bitcoin & UTxO

Traditional Account-Based Transactions



UTxO Bitcoin Transaction



- Decentralised
- Blockchain
- Digital Asset
- Programmable

UTxO: Unique method of transferring value without a financial intermediary.

An output represents Bitcoin that can be spent by a user who has the private key.

Bitcoin and Illicit Activity

Money Laundering

Dark Market Purchases

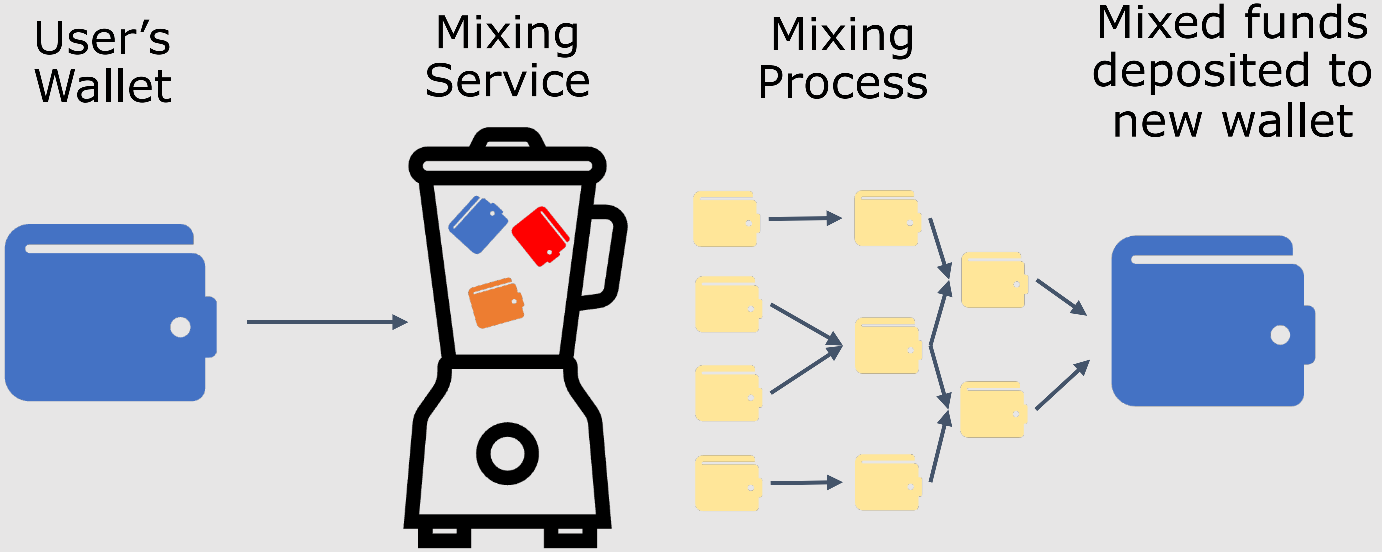
Terrorist Financing

Organized Crime Financing

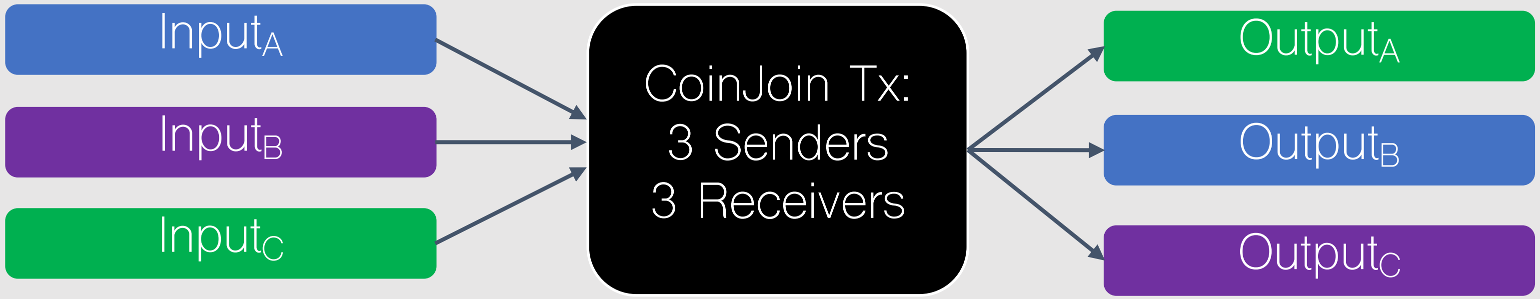
State bodies cybercrime

How to launder?

Mixing
Obfuscates origin of user's Bitcoin by blending them with many others.



CoinJoin
Multisignature transaction made available through privacy wallets and services.



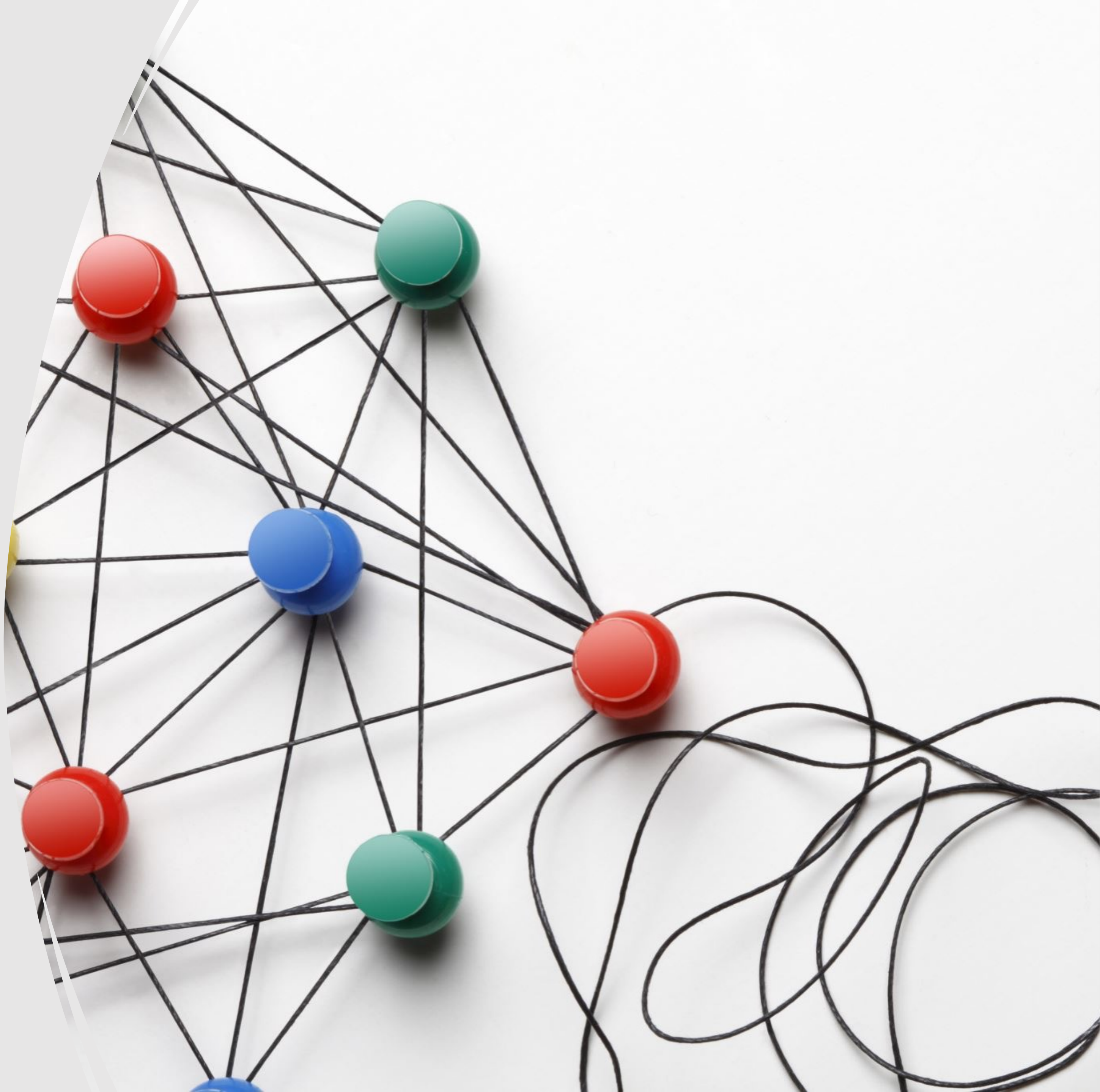
Heuristics

Denonymise the Bitcoin network

Group inputs into clusters

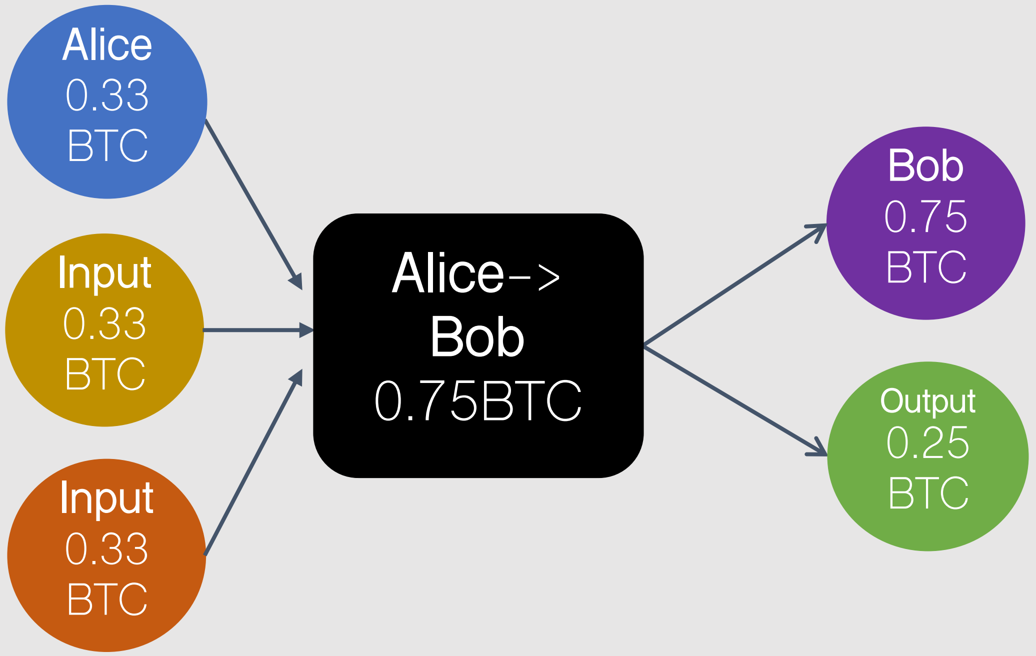
Heavy assumptions

Broken



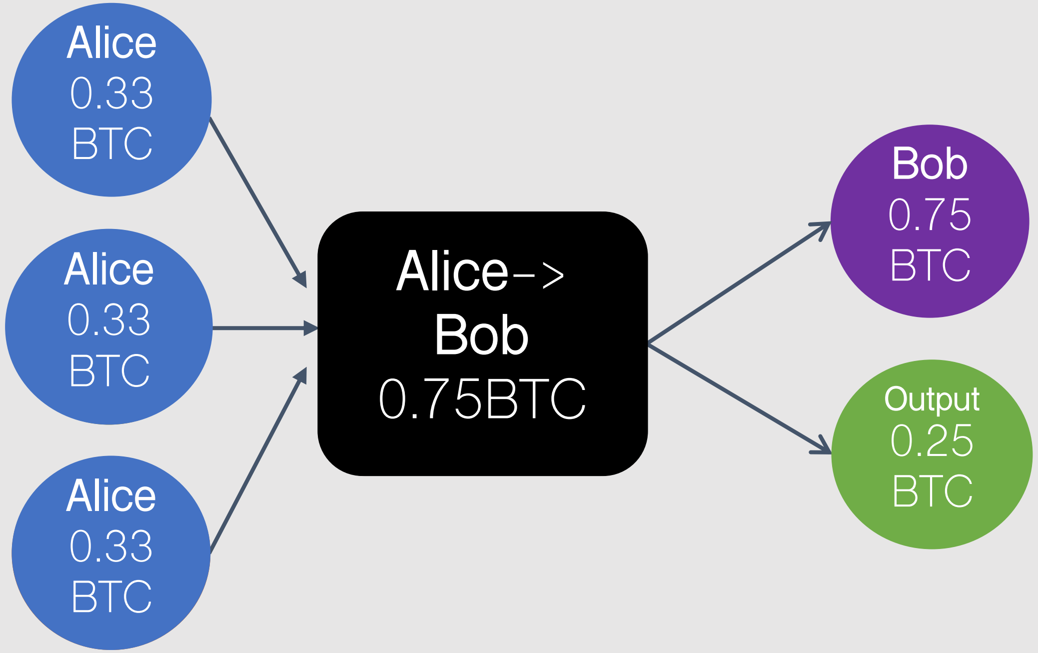
Heuristics

Multi-Input/Co-Spend
Clusters the inputs in a transaction and links them to a controlling entity

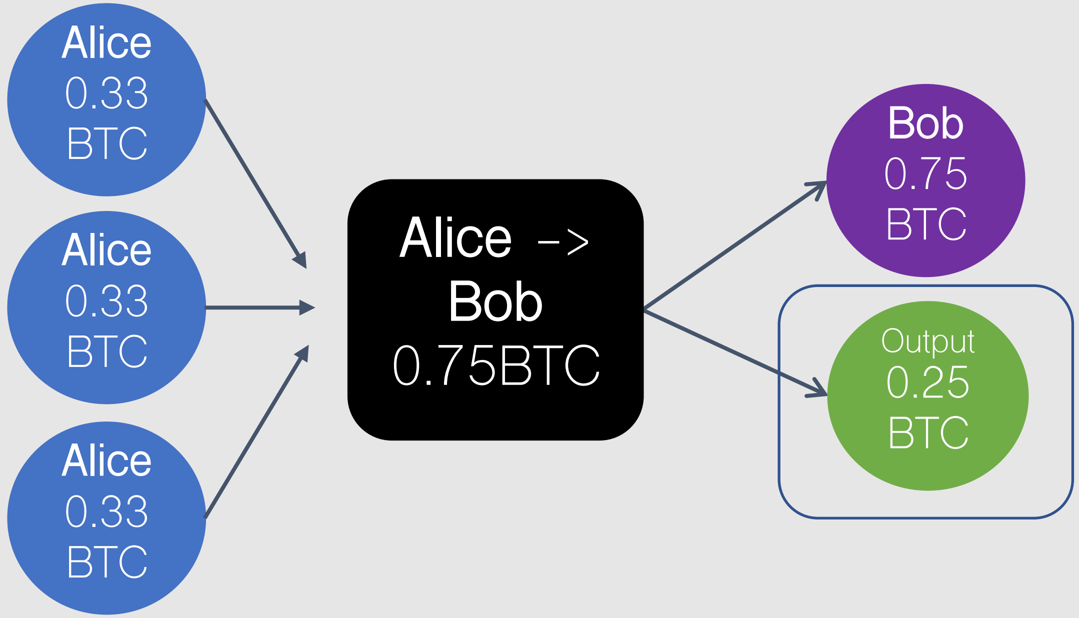


Heuristics

Multi-Input/Co-Spend
Clusters the inputs in a transaction and links them to a controlling entity

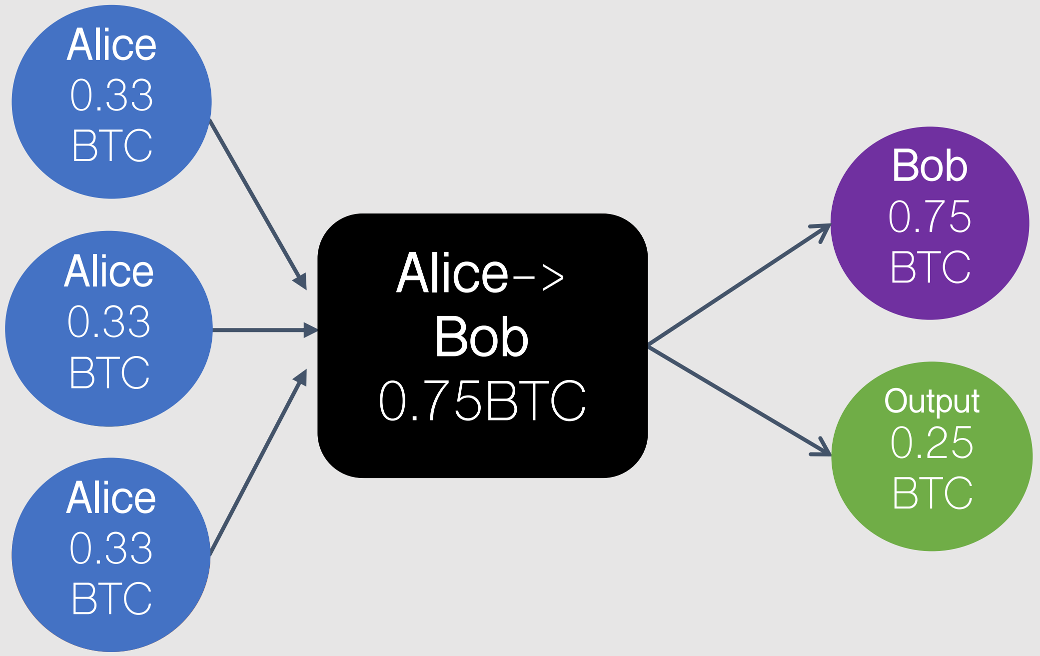


Change Address
Classifies one of the outputs as change in a standard transaction

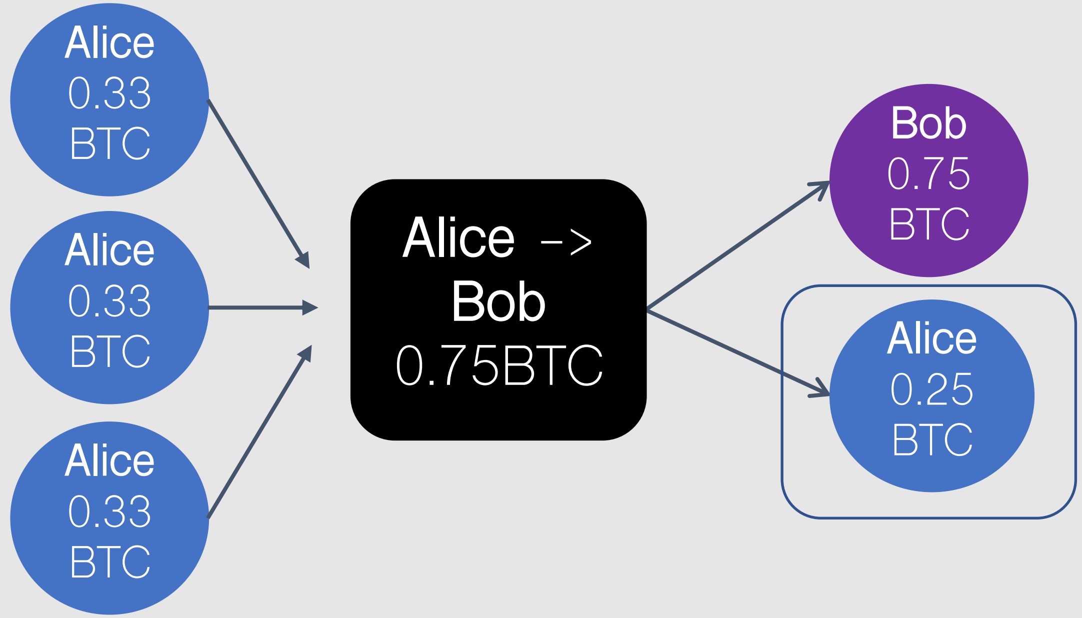


Heuristics

Multi-Input/Co-Spend
Clusters the inputs in a transaction and links them to a controlling entity



Change Address
Classifies one of the outputs as change in a standard transaction



Smallest amount must be change in transaction.

Deep Learning in Illicit Activity Identification

Heuristics have high avg. error rate (**63.46%** for co-spend, **92.66%** for change address)[1].

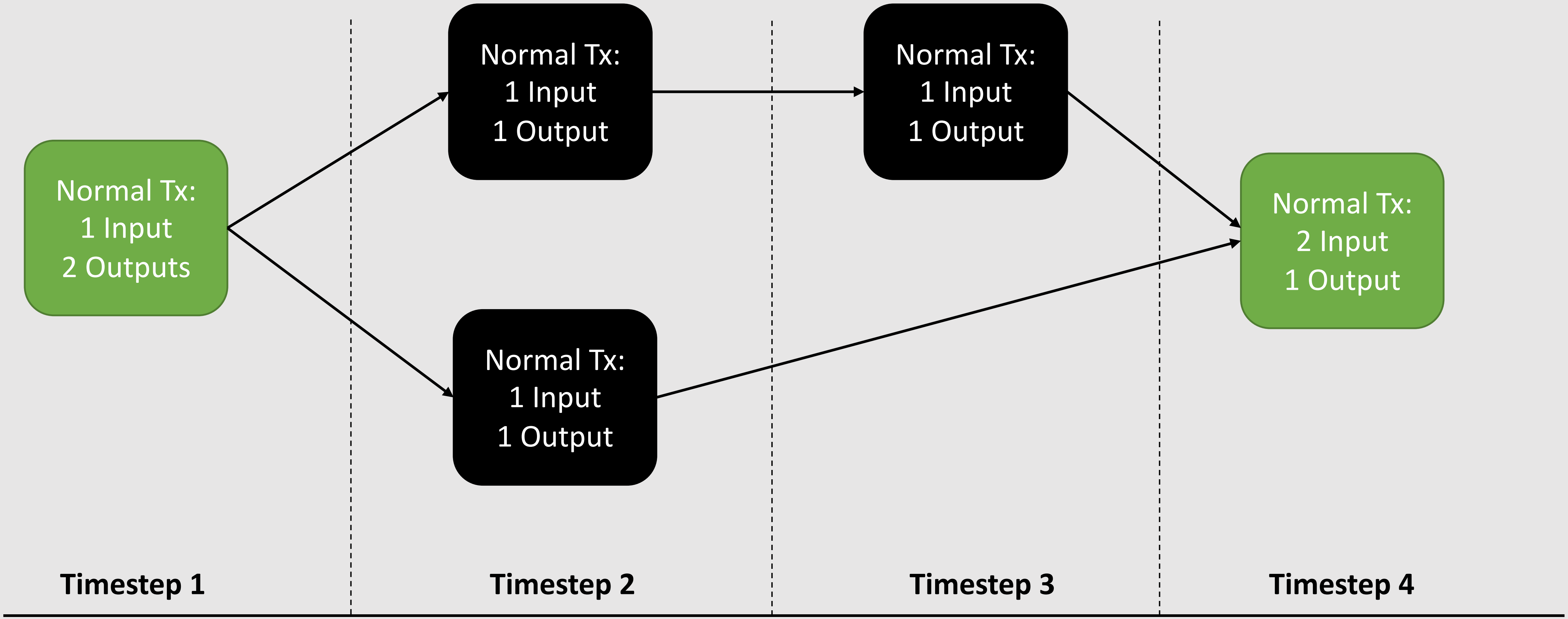
Complementing heuristics with ML[2,3].

Graph Neural Networks show promise in classification and deanonymisation tasks[4,5].

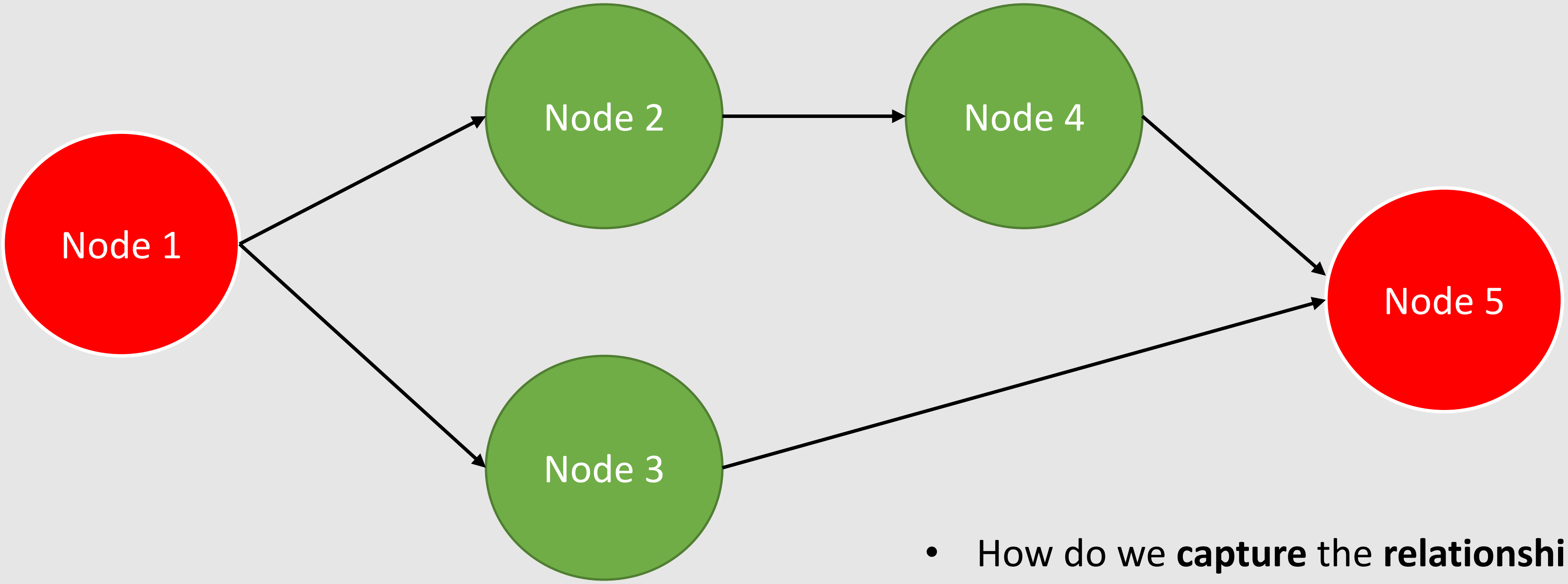
Bitcoin is naturally a graph.

Bitcoin Graph – Transaction Level

Classic Edges – Transaction flow



Illicit/Licit Labels



- How do we **capture** the **relationship** between **illicit** nodes?
- Can we **restructure** the graph based on underlying properties and similarity between nodes?
- Does this **improve** model's performance?

Bitcoin Graph Topology

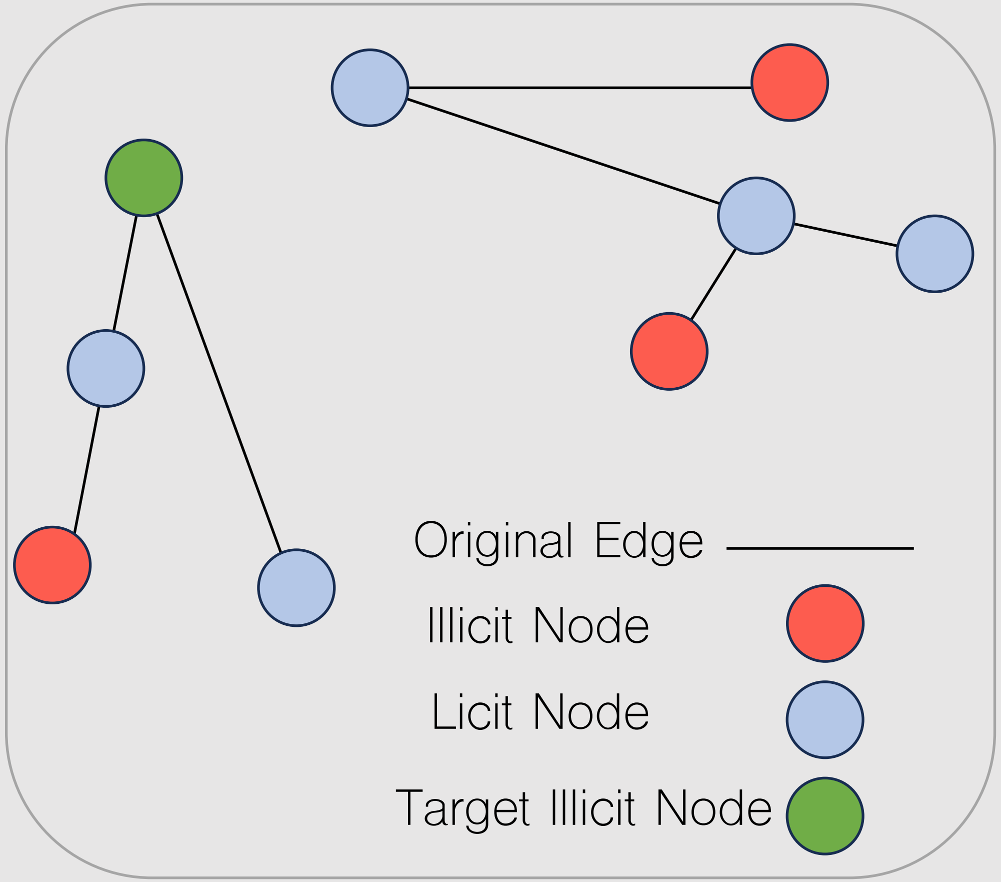
- Topology imbalance in Bitcoin is a major issue in illicit activity detection.

- Three key aspects of graph class-imbalance are unique against classical class-imbalanced tasks in ML.
    1. Graph data is unique and non-Euclidean. Traditional methods may struggle to handle **complex connectivity patterns** in graph data.
    2. Mishandling the graph relationships through under and oversampling can disrupt the **rich relational information**.
    3. Specialized techniques are needed to preserve and leverage the information.

Edges based on Affinity (EA)

Edges created based on node connectivity through Personalised Page Rank (**PPR**). Edge is created if connectivity score reaches parameter threshold.

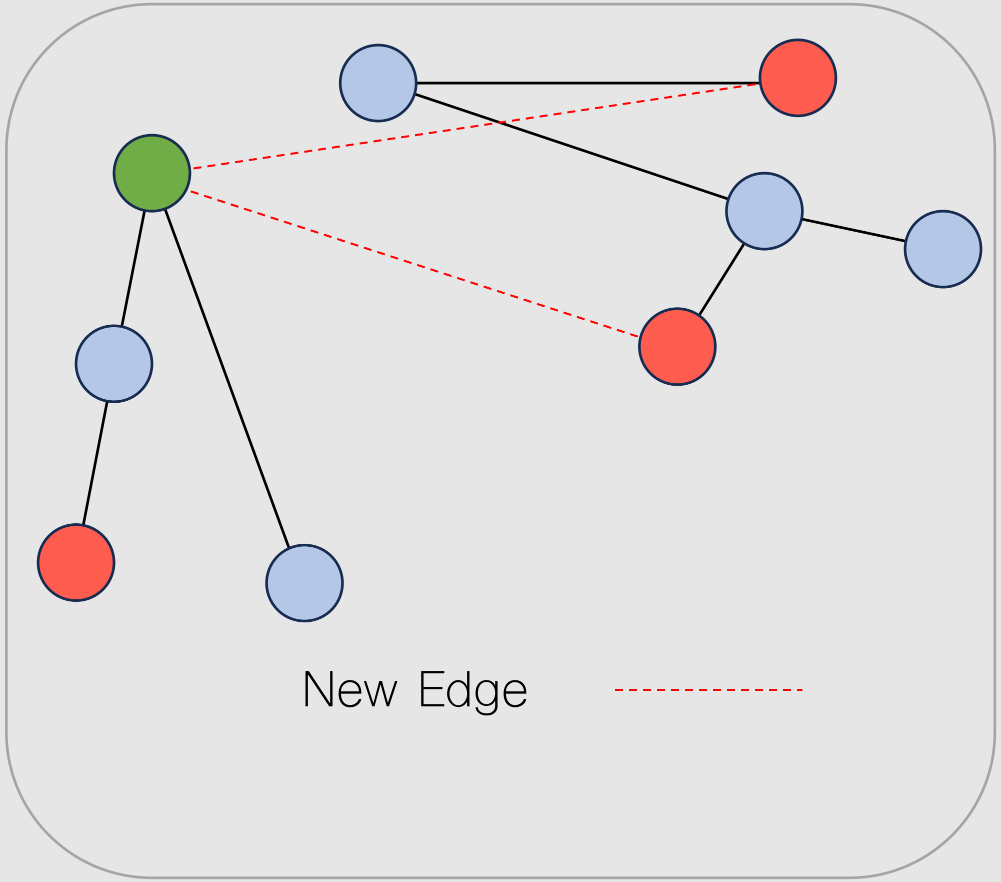
Original Graph



Measure connectivity influence of illicit nodes using PPR.

Establish new edges if connectivity over threshold.

Restructured Graph



Edges based on Affinity (EA)

To restructure a graph using EA:

- Using temporal graph, G, and create subgraph, G_L, with labelled illicit nodes (V_L).
- Pick random nodes, u_i and V_L, from G and G_L respectively.
- Apply function beta (PPR) to measure connectivity influence between V_L and u_i.
- Select all nodes, u_i, with the highest affinity to V_L and select all the edges between them to create new adjacency matrix A^*.

Algorithm 1 Edges based on Affinity (EA) Method

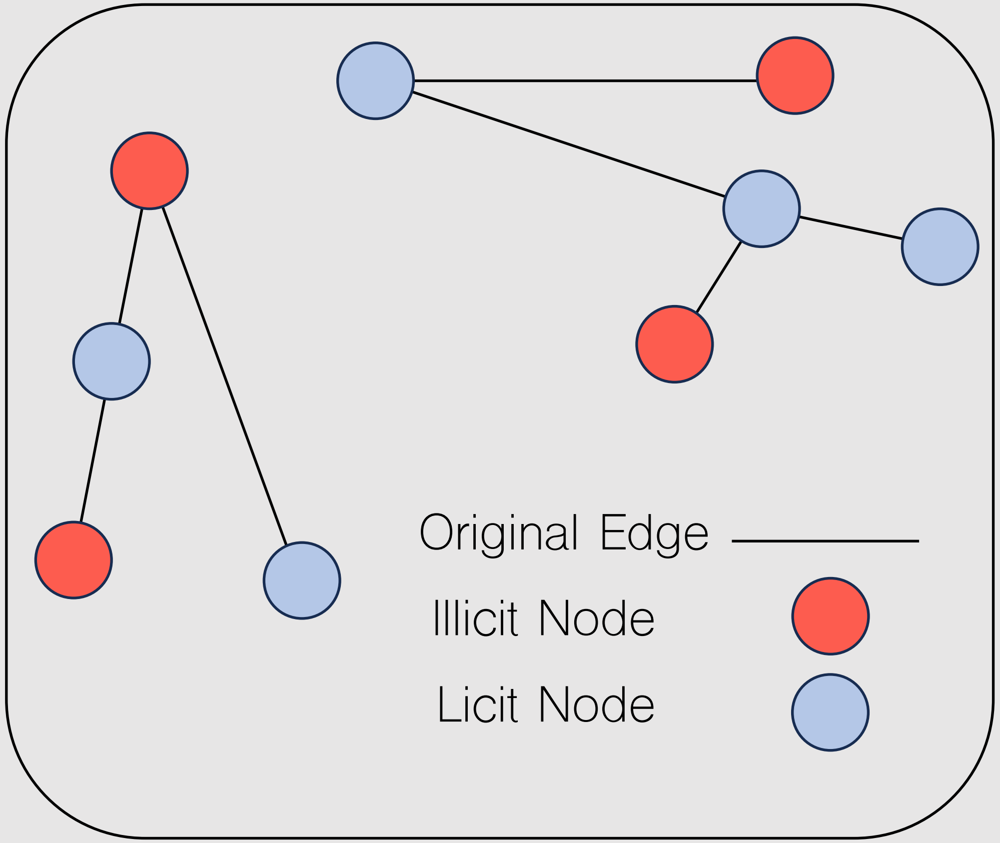
Require: Original graph G per temporal step, G'_l graph containing only labeled illicit nodes at training time and target ratio $p \in (0, 1)$

1: $n, n' \leftarrow$ Pick random nodes from G and G'_l, respectively
2: **for** $(G_t, n_t) \in \{(G, n), (G', n')\}$ **do**
3: $\quad \mathbf{s}_t \leftarrow$ Calculate connectivity scores of nodes $\beta(G_t, n_t)$
4: $\quad S_t \leftarrow$ Select k nodes having the largest scores in $\mathbf{s}_t$
5: $\quad S \leftarrow S_t$ if $G_t = G$ otherwise $S' \leftarrow S_t$
6: **end for**

Edges based on Node Features (ENF)

Edges created based on node feature similarity. MLP calculates embeddings and sigmoid function used to find probabilistic cut-off.

Original Graph

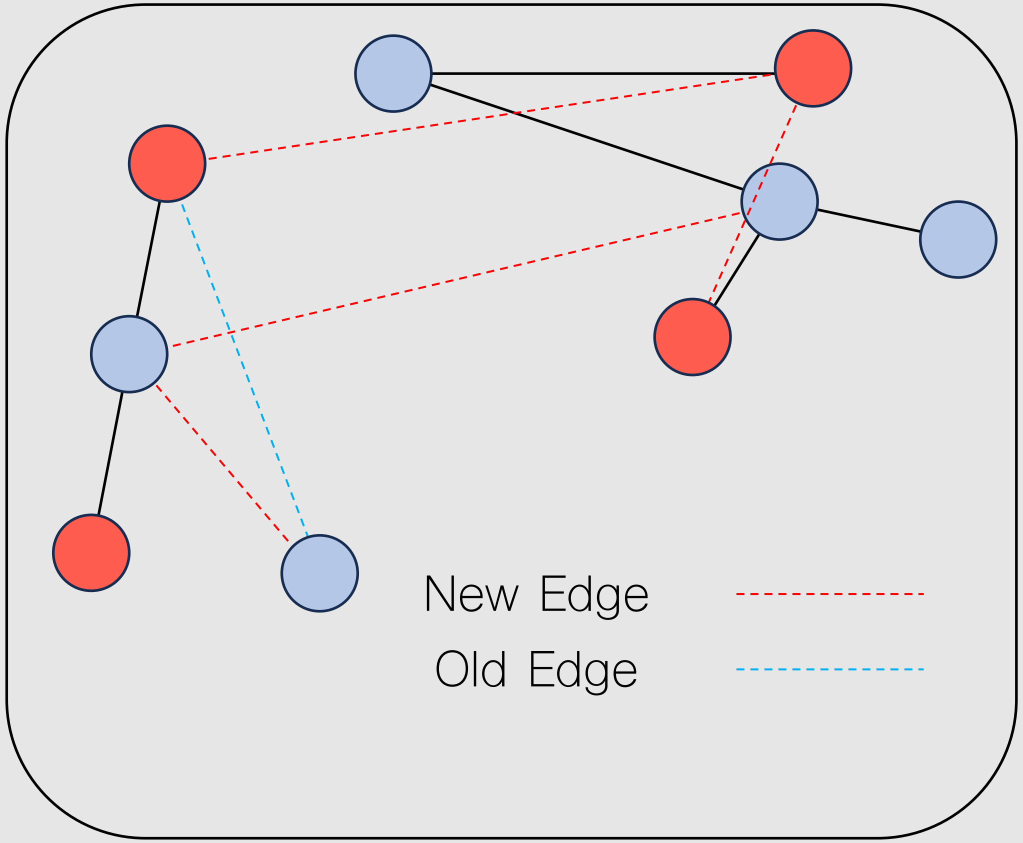


Node features similarity is compared using **MLP**.

A sigmoid function is used to decide whether an edge is created or not.

Low similarity scores are considered **noisy** and removed.

Restructured Graph



Edges based on Node Features (ENF)

- For each temporal graph, G, calculate embeddings, Z, for each node, u, against random node, v.

$$Z(u) = \theta\big(X\,(u)\big)$$

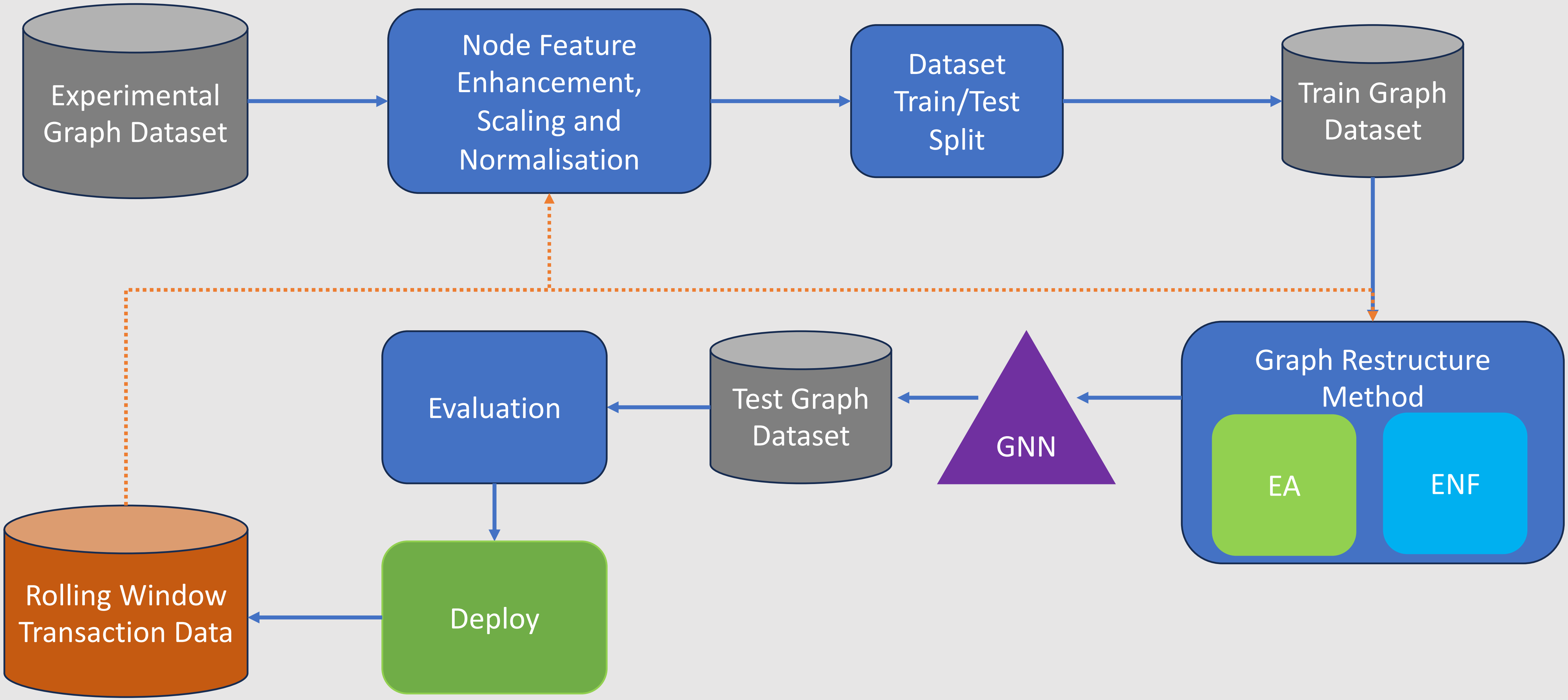
$$\pi_{u,v} = sigmoid\,(Z(u)Z(v)^T)$$

Where $\boldsymbol{\theta}$ is a two-layer perception network, $\boldsymbol{\pi_{u,v}}$ denotes the strength of similarity between node u and v.

- Create probability of forming edges using learning attention weights $\boldsymbol{\pi_{u,v}}$ in a parameterized matrix $\mathbf{P}_{uv} = \{\pi_{u,v}\}$

$$p_{uv} = \frac{\exp(\pi^{uv})}{\sum_u \exp(\pi^{uv})}$$

Pipeline



Experimental Graph Dataset
Node Feature Enhancement, Scaling and Normalisation
Dataset Train/Test Split
Train Graph Dataset
Graph Restructure Method
EA
ENF
GNN
Test Graph Dataset
Evaluation
Deploy
Rolling Window Transaction Data

Graph Neural Networks (GNNs)

- Relationship between data points (edges)
- Requires nodes, edges, and node features.
- Different architectures focus on different aggregation methods.
- Creates embeddings representative of nodes and their neighbourhood.

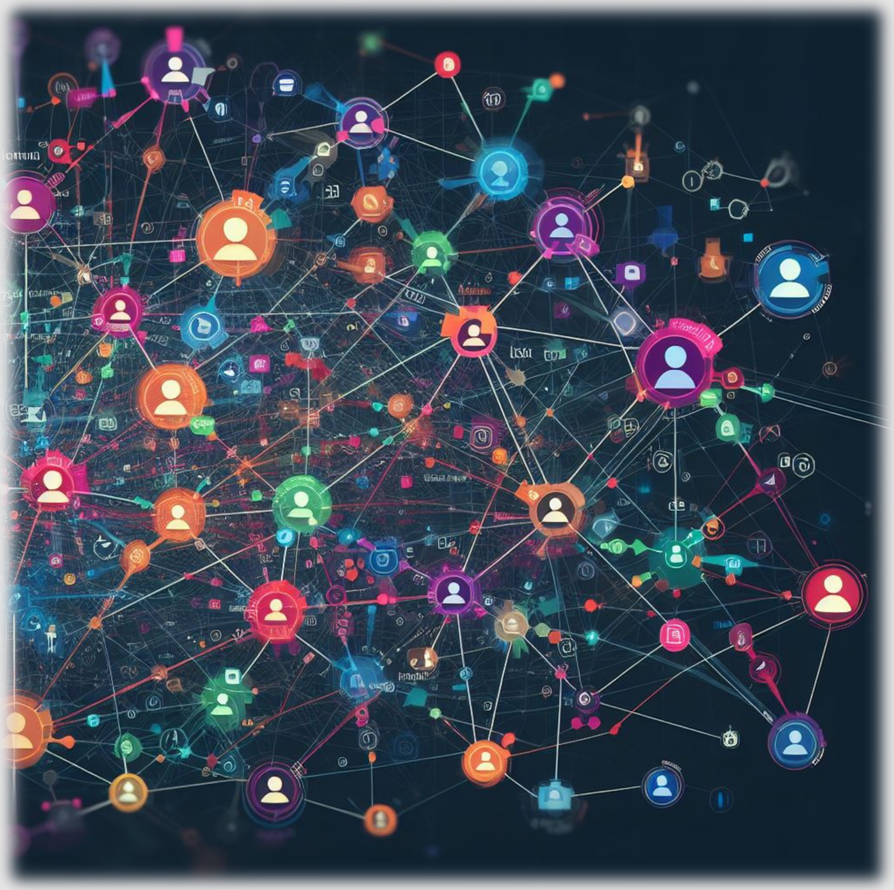
GNN
Graph Convolutional Network (GCN)[6]
Graph Attention Transformer (GAT)[7]
GraphSAGE[8]
Generalised PageRank (GPRGNN)[9]
Explore-to-Extrapolate Risk Minimization (EERM)[10]

Experiment

H: Can we improve GNN performance of an imbalanced node classification task using proposed EA and ENF methods?

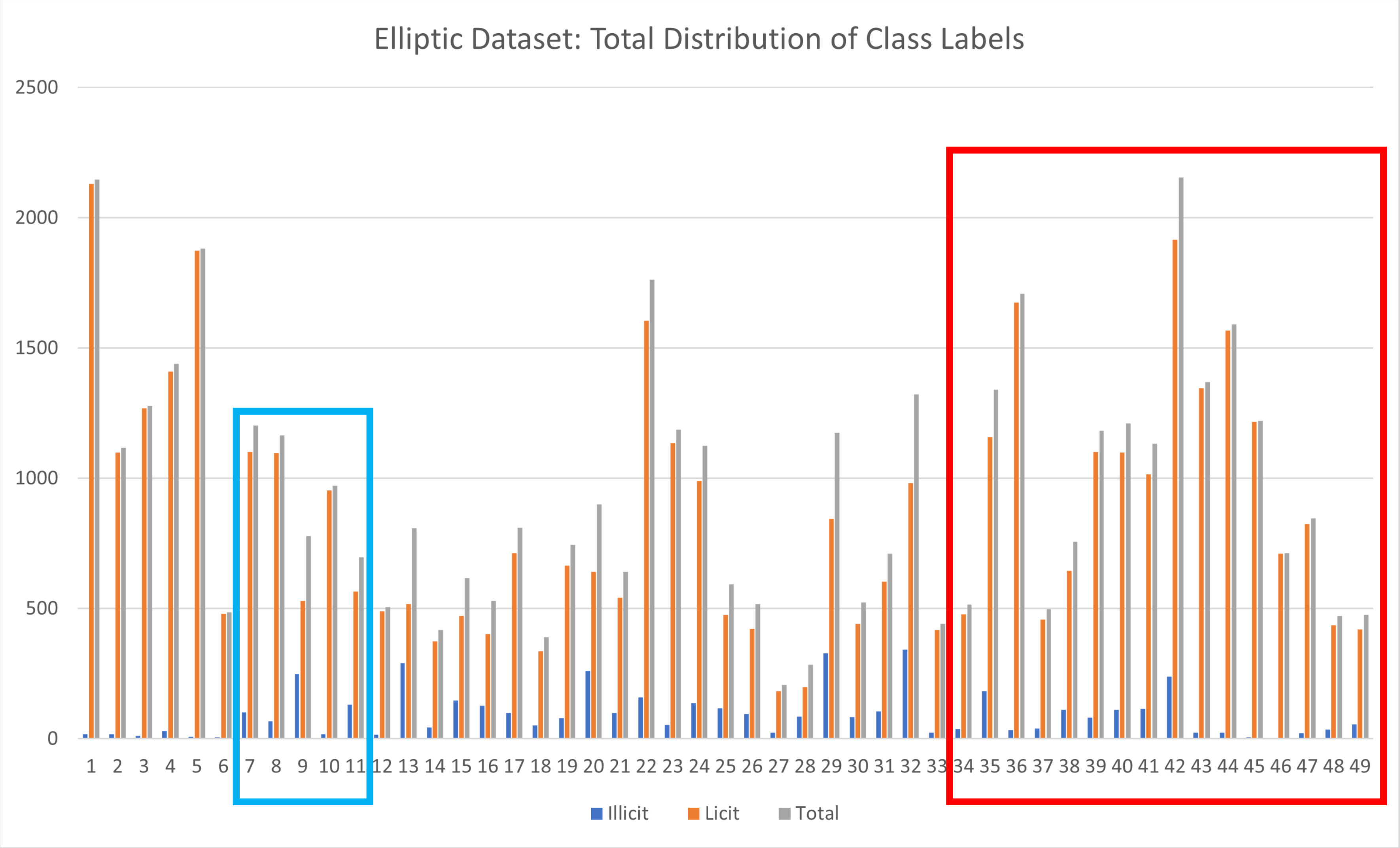
1. **Elliptic Bitcoin temporal graph dataset**:
   - Train 5 GNNs using new structured graphs from EA and ENF.
   - Compare against **baseline random forest**.

2. **Compare node imbalance techniques** against proposed EA and ENF methods on **DBLP citation network**.
   - Demonstrate model agnostic and multi-domain applicability of methods.



ELLIPTIC Bitcoin Dataset

- Largest labelled dataset for cryptocurrency illicit activity.
  - 203,769 nodes
  - 49 time-steps
  - 166 features
  - 21% labelled licit
  - 2% labelled illicit
- Labelled through heuristics-based reasoning.
- Popularly researched.
- We train on the first **7-11 timesteps**
- We test on time steps **34-49 timesteps**.
- Elliptic++ (2023)



Results – Elliptic Dataset

F1-score is metric of evaluation.

G^{EA} = Graph created from EA

G^{ENF} = Graph created from EA

G = Original Graph

Even in highly imbalanced temporal steps 43-49, GNNs identify illicit transactions.

ENF shown to be the most impactful method of graph restructuring.

Graph	GNN-Arch/Model	Timestep 34-38	Timestep 39-42	Timestep 43-46	Timestep 47-49
None	RF	85.49	78.67	0.00	0.00
G^{EA}	GCN	51.33	52.33	49.16	44.07
G^{ENF}		47.59	48.38	49.62	46.61
G		48.86	47.79	48.69	**46.65**
G^{EA}	GAT	56.75	53.76	48.16	46.432
G^{ENF}		68.58	58.19	47.32	52.97
G		50.39	50.26	**49.01**	46.80
G^{EA}	Graph-SAGE	61.54	59.25	54.15	46.55
G^{ENF}		65.86	62.73	49.58	46.97
G		56.06	50.23	49.39	46.41
G^{EA}	GPR-GNN	67.92	63.44	48.22	44.02
G^{ENF}		72.73	61.37	49.73	46.69
G		67.34	**67.25**	47.91	44.34
G^{EA}	EERM	76.05	78.09	62.65	49.91
G^{ENF}		76.35	78.34	63.92	50.45
G		73.05	75.33	59.45	49.42

Results – DBLP Citation Network

ENF and EA method tested with GAT.

ENF consistently **outperforms** against other node imbalance classification techniques.

Method	DBLP	Timestep 34-38	Timestep 39-42	Timestep 43-46	Timestep 47-49
ReNode	52.70	54.02	50.38	48.95	50.38
RECT	51.40	54.18	51.67	47.83	46.25
DR-GCN	54.30	52.04	50.36	48.91	45.67
ENF with GAT	56.80	68.58	58.19	49.01	52.97
EA with GAT	53.70	56.75	53.76	48.16	46.43

Discussion

- GNN models can identify illicit transactions well in each timestep segment even with heavy class imbalance.
- Edge Affinity (EA) and Edge Node Features (ENF) consistently outperform original graph.
- EA and ENF are model and domain agnostic.
- Preprocessing of MLOps Pipeline.
- Potential for identifying mixing and CoinJoin operations.
- Wider applicability in financial cybercrime activity detection

Future Work & Limitations

- Improving performance and testing on more datasets.

- Rich node features required to gauge similarity.

- Integrating LLM to interpret transactions and create narratives for investigation.

Thank You

Jack Nicholls

Email: jack.nicholls@ucdconnect.ie

Dr. Aditya Kuppa

Email: aditya.kuppa@ucd.ie

Assoc Prof. An Le

Email: an.lekhac@ucd.ie

LinkedIn: https://www.linkedin.com/in/jack-nicholls93/

Twitter/X: https://twitter.com/JackPNicholls

Sources

1. Gaihre, A., Pandey, S., & Liu, H. (2019). Deanonymizing Cryptocurrency with Graph Learning: The Promises and Challenges. *2019 IEEE Conference on Communications and Network Security, CNS 2019*, 2019–2021. https://doi.org/10.1109/CNS.2019.8802640
2. Möser, M., & Narayanan, A. (2021). *Resurrecting Address Clustering in Bitcoin*. http://arxiv.org/abs/2107.05749
3. Kappos, G., Yousaf, H., Stütz, R., Rollet, S., Haslhofer, B., & Meiklejohn, S. (2022). *How to Peel a Million: Validating and Expanding Bitcoin Clusters*. http://arxiv.org/abs/2205.13882
4. Weber, M., Weidele, D. K. I., Domeniconi, G., Bellei, C., Leiserson, C. E., Chen, J., & Robinson, T. (2019). Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. *ArXiv*, *10*.
5. Gaihre, A., Pandey, S., & Liu, H. (2019). Deanonymizing Cryptocurrency with Graph Learning: The Promises and Challenges. 2019 IEEE Conference on Communications and Network Security, CNS 2019, 2019–2021. https://doi.org/10.1109/CNS.2019.8802640
6. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings, 1–14.
7. Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive Representation Learning on Large Graphs. http://arxiv.org/abs/1706.02216
8. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2017). Graph Attention Networks. http://arxiv.org/abs/1710.10903
9. Chien, E., Peng, J., Li, P., & Milenkovic, O. (2020). Adaptive Universal Generalized PageRank Graph Neural Network. http://arxiv.org/abs/2006.07988
10. Qitian Wu, Hengrui Zhang, Junchi Yan, and David Wipf. 2022. Handling Distribution Shifts on Graphs: An Invariance Perspective. arXiv preprint arXiv:2202.02466