

Ledgit: A Service to Diagnose Illicit Addresses on Blockchain using Multi-modal Unsupervised Learning

Xiaoying Zhi², Yash Satsangi¹, Sean Moran¹, Shaltiel Eloul¹

CTO, JP Morgan Chase¹ ETH Zurich²

Introduction

- **Distributed ledger** technology benefits society by enabling an ecosystem of decentralized finance but its pseudo-anonymous nature enables new routes for financing **illicit** activities.
- Current solution for identifying digital wallets that are involved in illicit activities rely on **investigations** run by government agencies that are expensive and difficult to scale leading to outdated list of addresses.
- **Publicly available data** such as transaction data and community reports is real-time and contains useful information that can help identify if a digital wallet is associated with illicit activities or not.
- We present **Ledgit**: an online service that lets an user query the risk score of a **Bitcoin wallet address** based on publicly available data.

Date	Abuse Type	Description
Feb 4, 2022	ransomware	I lost over \$ 57,000 however I was able to recover and reclaim all my lost money easily and free of charge through an online recovery platform link : https://moneybacklimited.com ... If you are a victim of online scam and you want to recover all your lost money free of charge you are required to urgently visit the link https://moneybacklimited.com to track, trace, reclaim and recover all your lost money free of charge as soon as possible., The recovery process is easy and fast. visit the link https://moneybacklimited.com right away to recover and reclaim all your lost money free of charge.
Feb 4, 2022	other	This scammer is pretending to trade for investors but just takes the money, then blocks you

Figure 1: Anonymous public reports from **BitcoinAbuse.com**

Ledgit

- **Ledgit**: An automated real-time service: that lets an user query the risk level of a Bitcoin wallet address involved in illicit activities based on public reports. Ledgit tackles the challenge of:
- Unsupervised spam detection: online platforms are filled with reports that can be **spam** or advertisement.
- Multi-modality: mining features and patterns from dynamic **transaction graphs** and combining them with textual and meta-data from public reports.

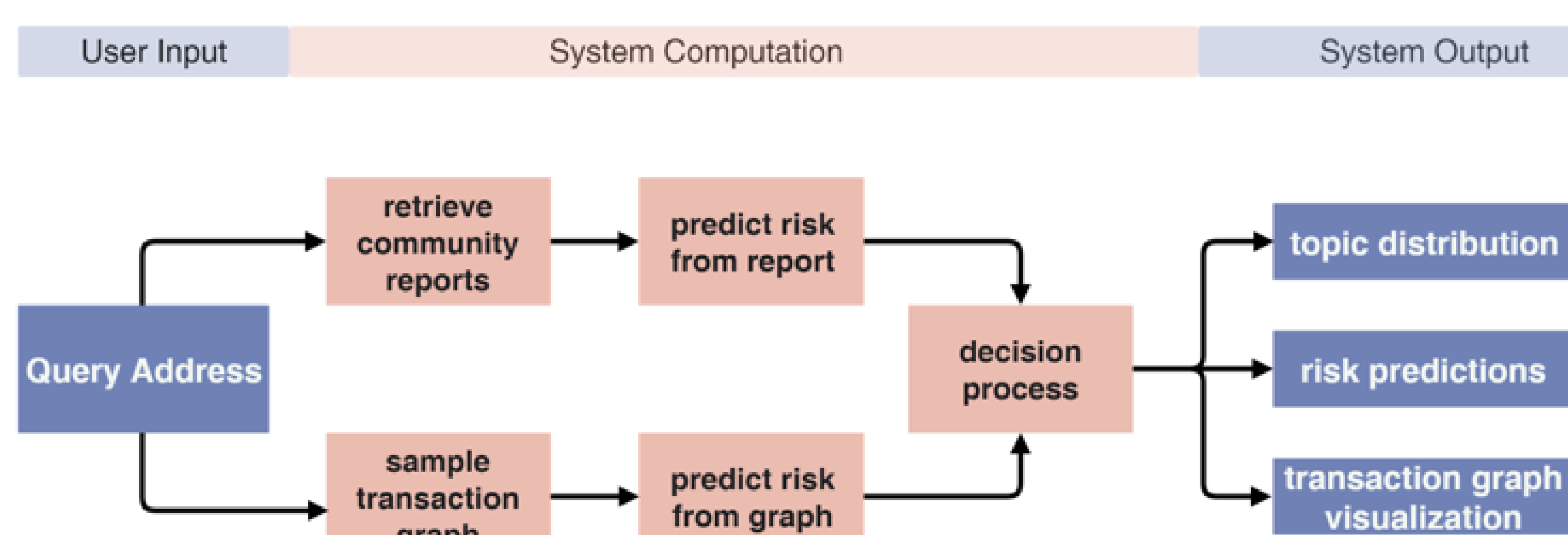


Figure 2: System overview

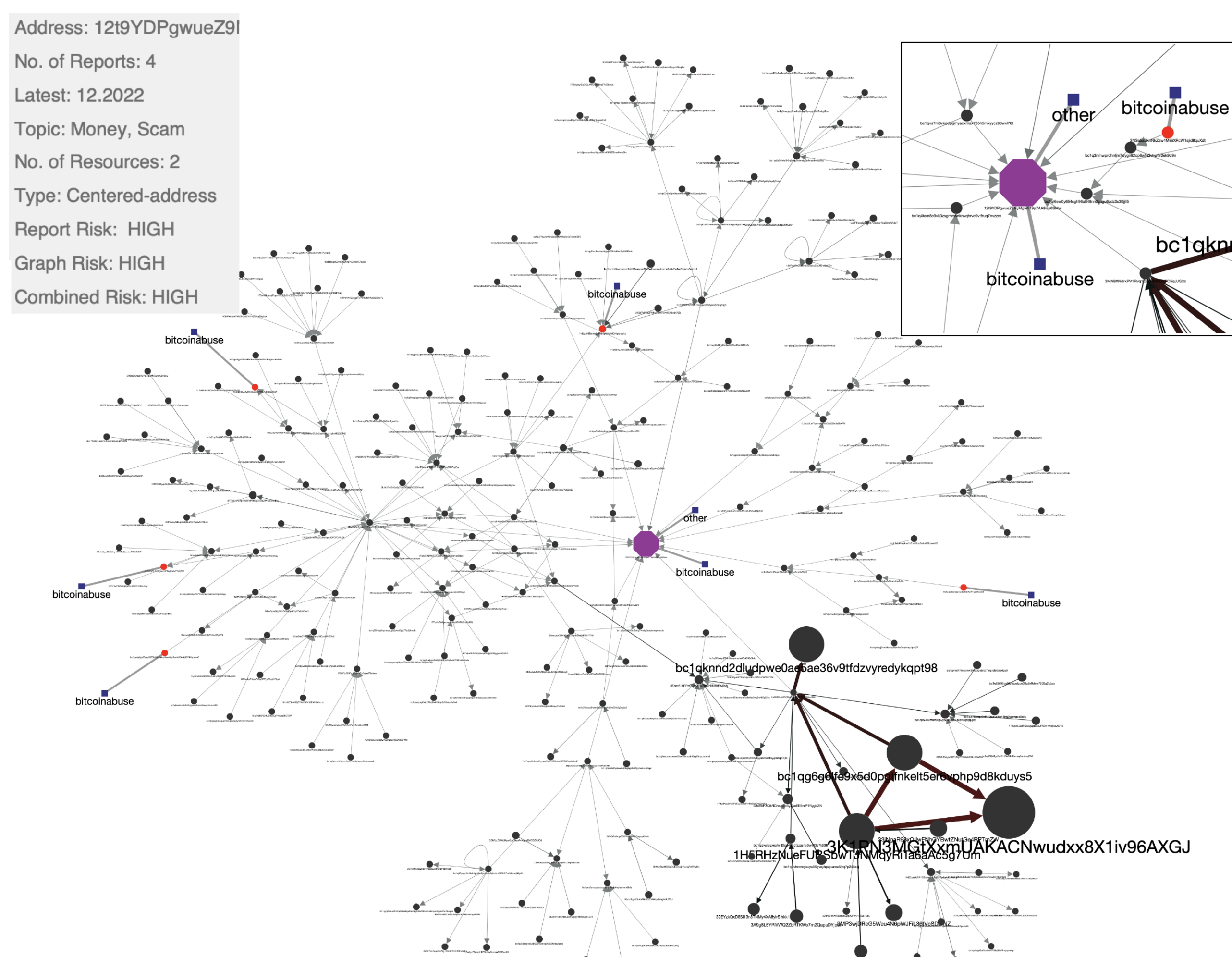


Figure 3: Snapshot of user interface

Algorithm

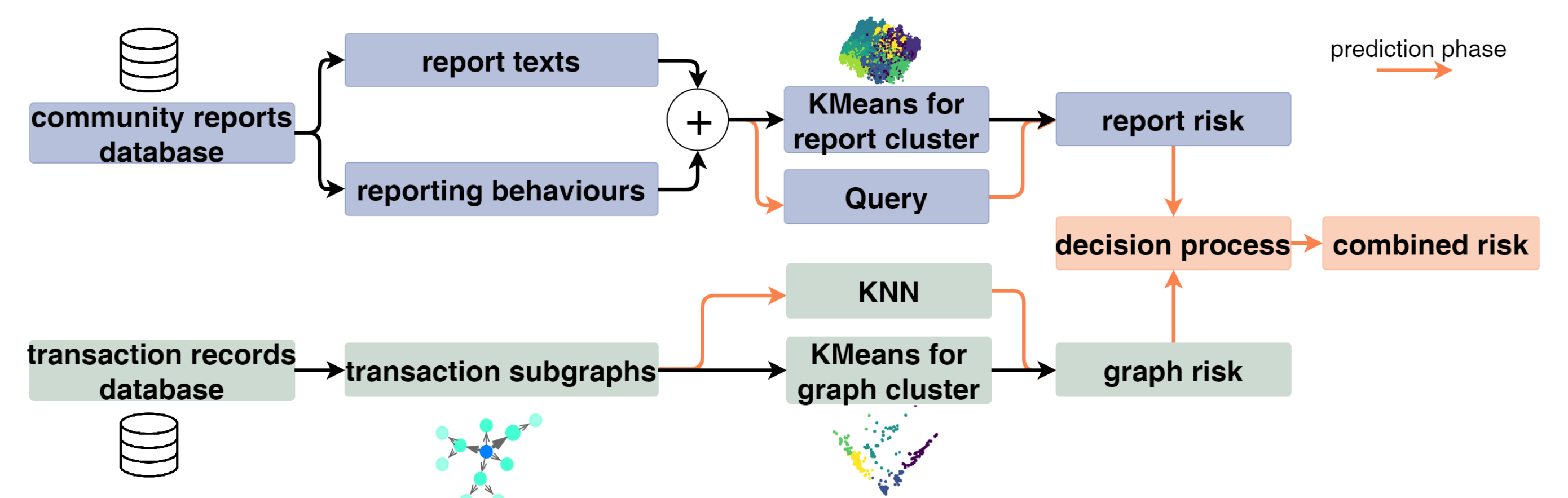


Figure 4: Data processing and training pipeline

- BitcoinAbuse reports are broken into text and metadata. The text is encoded using universal sentence encoder to capture **textual content**.
- The metadata helps in **cross victim/address** complaints.
- From transaction graphs, information such as **volume of transactions** and neighborhood behaviour is detected.
- Based on each modality, addresses are **clustered** into risky and non-risky clusters.
- The final risk score is obtained by combining risk scores from BitcoinAbuse reports and transaction graph.

Results

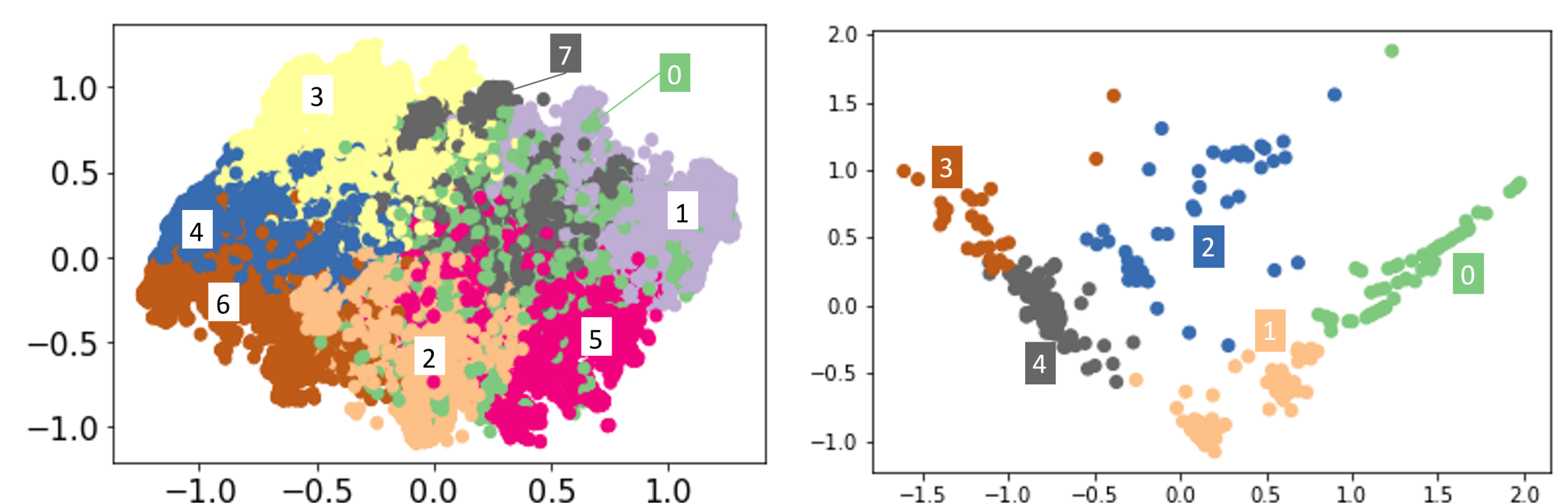


Figure 5: (left) Report clusters; (right) Transaction graph clusters

cluster	avg. txn value	avg. txn std	addresses
0	104.65	880.90	20.11%
1	6.51	37.48	24.93%
2	37.57	386.94	12.18%
3	11.58	77.51	8.22%

Table 1: Cluster profiles for transaction graphs

cluster #	reports	duration (hour)	addresses
0	1.42	140.78	5.80%
1	3.02	210.42	11.78%
2	1.96	98.92	18.24%
3	1.83	218.51	9.36%
4	1.95	118.97	15.95%
5	9.05	544.57	14.85%
6	2.42	87.50	9.31%
7	2.32	157.72	14.71%

Table 2: Cluster profiles for reports cluster

	report risk	graph risk	combined risk
accuracy	0.7467	0.8400	0.9067
precision	0.9444	0.9756	0.9400
recall	0.6667	0.7843	0.9215
F1-score	0.7816	0.8695	0.9306

Table 3: Evaluation of three clustering methods assuming 'illicit' as 'positive' class

References

- [1] Blockchain.com, 2008.
- [2] Bitcoinabuse, 2022.
- [3] Shaltiel Eloul, Sean J Moran, and Jacob Mendel. Improving streaming cryptocurrency transaction classification via biased sampling and graph feedback. In *Annual Computer Security Applications Conference*, pages 761–772, 2021.
- [4] Arjun Mukherjee, Abhinav Kumar, Bing Liu, Junhui Wang, Meichun Hsu, Malu Castellanos, and Riddhiman Ghosh. Spotting opinion spammers using behavioral footprints. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 632–640, 2013.
- [5] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Technical report*, Mambot, 2008.