Defending Blockchain from Malicious Contracts Using Anomaly Detection

DUY HUYNH · SON HOANG DAU · XIAODONG LI

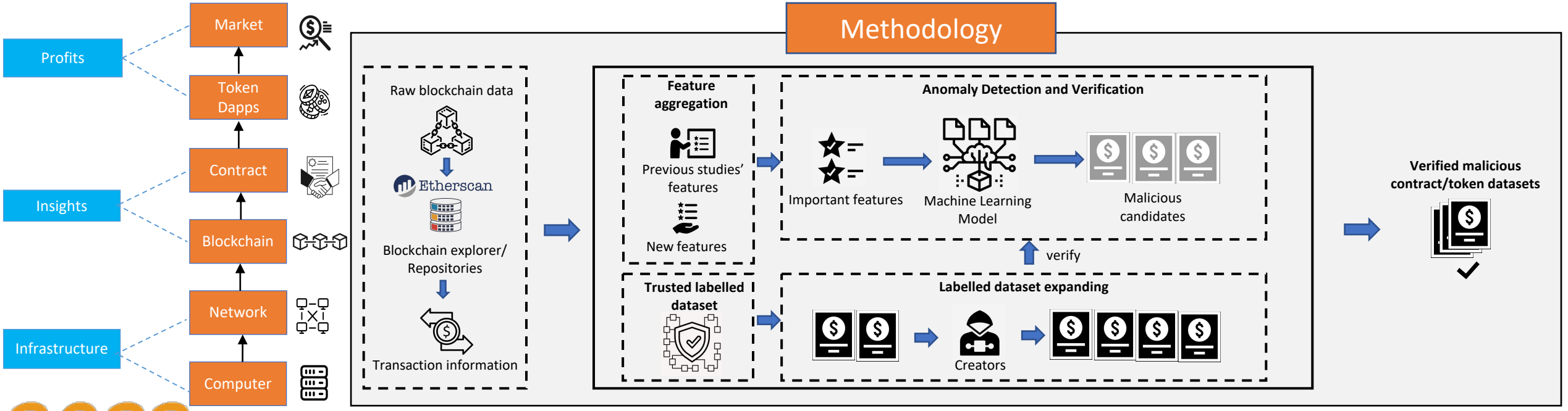
Introduction

PROBLEM

- Blockchain became a “paradise” of latent financial scams along to its rocket development
  - E.g: Ponzi, Phishing, Honeypot, Wallet theft, Rug-pull, etc
- Crypto investors lost millions of dollars by scam tokens and malicious contracts

MOTIVATION

- Previous studies focus on solving problems at infrastructure and operation levels while financial scams widespreads mainly at profit level
- Many studies aim at tackling each kind mallicious contract using classification, but those requires labeled datasets
- Anomaly detection is widely used in financial domains and does not require labelled datasets which are unavailable in new emerging blockchains



2022 Australian Computer Science Week

Defending Blockchain from Malicious Contracts Using Anomaly Detection

DUY HUYNH · SON HOANG DAU · XIAODONG LI

What is done so far

Data collection

- We first tried on Ethereum blockchain data
- 8 million blocks data was downloaded from etherscan:
  - 480 million normal transactions
  - 690 million internal transactions
- Labelled data set was downloaded from Xblock.pro repository
  - 3795 contracts with more than 200 Ponzi contracts
- We grouped transaction information by contract address and saved them as CSV format to facilitate further steps

Features aggregation

- Aggregated new features base on different aspects
  - Eg: Transaction, Time series, Creator, etc.
- Data nomarlization
- Doing statistic to show features' characteristic

Future works

- Choosing different anomaly detection techniques to solve the problem
- Finding the most appropriate model and the most important features list
- Verifying the out-come dataset and characterizing mallicious contracts

References

- Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system. Cryptography Mailing list at https://metzdowd.com, 03 (2009)
- Dinh, T. T. A., Liu, R., Zhang, M., Chen, G., Ooi, B. C., & Wang, J. (2018). Untangling blockchain: A data processing view of blockchain systems. IEEE Transactions on Knowledge and Data Engineering, 30(7), 1366–1385. ISSN 1041-4347.
- Criminals have made off with over $10 billion in 'DeFi' scams and thefts this year. https://www.cnbc.com/2021/11/19/over-10-billion-lost-to-defi-scams-and-thefts-in-2021.html, 2021
- A.S. Sabau, Survey of clustering based financial fraud detection research, Inform. Econ. 16 (1) (2012) 110–122.
- Ahmed, Mohiuddin, Abdun Naser Mahmood, and Md Rafiqul Islam. "A survey of anomaly detection techniques in financial domain." Future Generation Computer Systems 55 (2016): 278-288.