Information Technology University

Analysis and Categorization of Smart Contracts on EVM-based Blockchain Networks (Polygon, Binance, Optimism)

Blockchain Project

Phase 2

MSCS22027 – Muhammad Qasim

BSCS21017 – Arham Irshad

Project Mentor: Fatima Shaheen &Abeeha Fatima

Nov 22, 2024

Table of Contents

Summary 3
Introduction 3
  Objective 3
Problem Statement 3
  Expected Outcomes 4
  Scope: 4
Literature Review 4
  Research Papers and Articles 4
Project Requirements 5
  Functional Requirements 5
  Non-Functional Requirements 5
  Resource Requirements 5
System Design and Architecture 6
  System Overview: 6
  Diagrams/Flowcharts: 6
    System Flow 6
    Categorization Workflow 6
    High-Level Diagram 7
  Module Breakdown: 7
Methodology and Detailed Design 8
  Requirements Analysis: 8
  Model/Algorithm Specification: 8
  Development Environment and Tools: 8
Future Work and Next Steps 8
  Phase 3 Roadmap: 8
  Remaining Milestones: 9
Conclusion 10
References 11

Summary

This project focuses on categorizing smart contracts across three prominent EVM-compatible blockchains: Binance Smart Chain, Polygon, or Optimism. Phase 2 of the research assumed that the metadata from approximately 10,000 smart contracts per blockchain was to be gathered, categorized, and later analyzed based on the foundational research that was conducted during phase one of the project. Some of the milestones realized were in categorizing contracts into tokens (tokens such as ERC-20, and ERC-721) DeFi apps, wallets, and games with distribution trend visualizations by each blockchain.

The data collected in phase two offers a clear understanding of the usage of contracts across the chosen networks and may help developers, investors, and researchers wanting to develop the blockchain technology to make optimal decisions depending on results obtained from phase 2.

Introduction

Due to the increasing popularity of EVM-compatible blockchains such as Polygon, Binance Smart Chain (BSC), and Optimism, there is a rapid uptake of DApps and smart contracts. This work explores the taxonomy and application of these smart contracts to identify their distribution by category of token, NFT, and DeFi.

Objective

To perform due diligence on ~10K contracts each from Polygon, Binance, and Optimism networks or broken down, for instance, into token smart contracts, NFTs, wallets, routers, and anything else. The purpose is to track different smart contract platforms and networks as well as provide a comprehensive breakdown of the contract type and quantity within those networks.

Problem Statement

The unprecedented growth of blockchain networks and dApps requires knowledge of the distribution of smart contracts between networks. Tokens or figures like NFTs, DeFi applications, and wallets affecting it represent the ecosystem's growth and its weaknesses. They analyze different categorization techniques and identify additional complex patterns in the gathered data sets at this phase.

Expected Outcomes

- **Categorization of Contract Types:** Establishment of the key categories studying tokens, NFTs, DeFi applications, wallets, and others.
- **Deployment Trends:** Comparison of the distribution of each type of contract in various networks.
- **Insights on Contract Usage:** An account of how each network is used, or suggestions for enhancements in either its operations or security measures.

Scope:

- The smart contract analysis for data extraction and metadata collection from BSC, Polygon, and Optimism.
- Division of contracts into certain categories (tokens, NFTs, DeFi, wallets, games, and more).
- The trend of the three blockchains and the distribution comparison of the two indices between each of the blockchains.
- Pilot study to establish the accuracy of the data as well as confirm the appropriateness of the categorization techniques.

Literature Review

Building on foundational research from Phase 1, additional insights were gained from studies such as:

- **Smart Contracts Taxonomy and Classification:** Increased knowledge about the nature of contract risks and improved approaches to contract categorization.
- **Blockchain Ecosystem Mapping:** Information on modern tendencies of contracts application and specific usage in the networks. Approaches and tools from these studies were used to design data collection techniques, data representation, and analysis techniques for this phase.

Research Papers and Articles

- Blockchain-based Smart Contracts: A Systematic Mapping Study
- Blockchain-based Smart Contracts: A Systematic Mapping Study of Academic Research.

- Vulnerabilities and Open Issues of Smart Contracts: A Systematic Mapping.

Project Requirements

Functional Requirements

- **Data Extraction:** Get smart contract details from the Polygonscan website | BscScan website and Optimism etherscan website respectively.
- **Contract Categorization:** Split the contracts into categories which are tokens (ERC-20), Non-Fungible tokens (ERC-721), games, decentralized finance, wallets, and others.
- **Reporting and Visualization:** Produce a report of the outcomes with graphical illustrations of the distribution of the types of contracts.

Non-Functional Requirements

- **Performance:** Data extraction and data processing for the management of contracts number more than 10,000 per the given network.
- **Usability:** Gather numerous contracts and organize them systematically, thus there are markers of clear categorizations of contracts for easy analysis.
- **Security:** API data protection, the fact that the API data especially when in transit or storage needs to be protected from access by unauthorized persons, for example, API keys.

Resource Requirements

- **APIs:** Ability to leverage API from Polygonscan, BscScan, and Optimism Etherscan.
- **Development Tools:** For data processing and analysis, Python is majorly used with the help of libraries like the request, to interact with APIs that provide the data, pandas, used for data handling and for getting the visual outputs like bar or line charts we will be using matplotlib or seaborn.
- **Hardware:** Computation capability to accommodate the amount of data storage and processing required to support at least thirty thousand smart contracts.

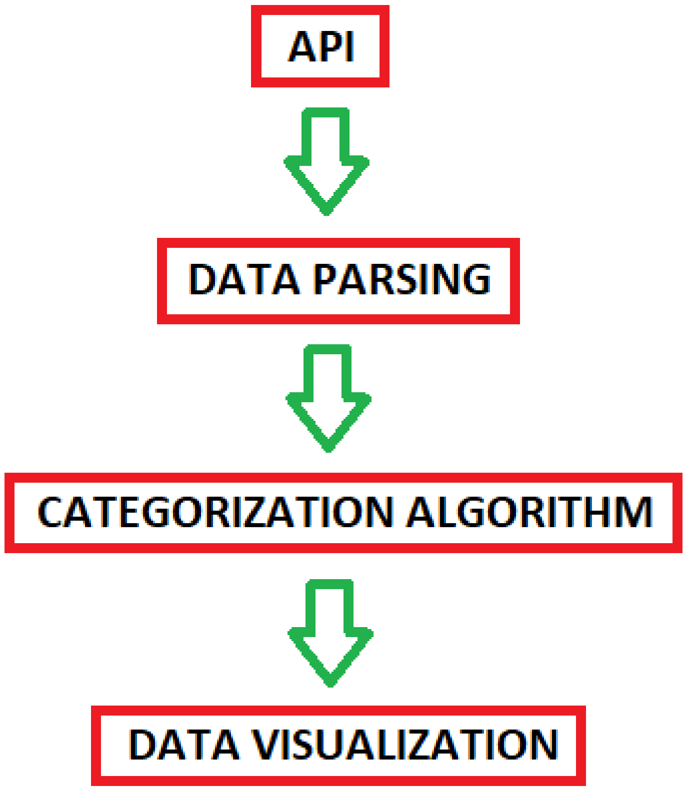
System Design and Architecture

System Overview:

Using API (BscScan, Polygonscan, Optimism Etherscan), the system downloads smart contract metadata, and then it sorts the contracts into groups based on outlined classes.

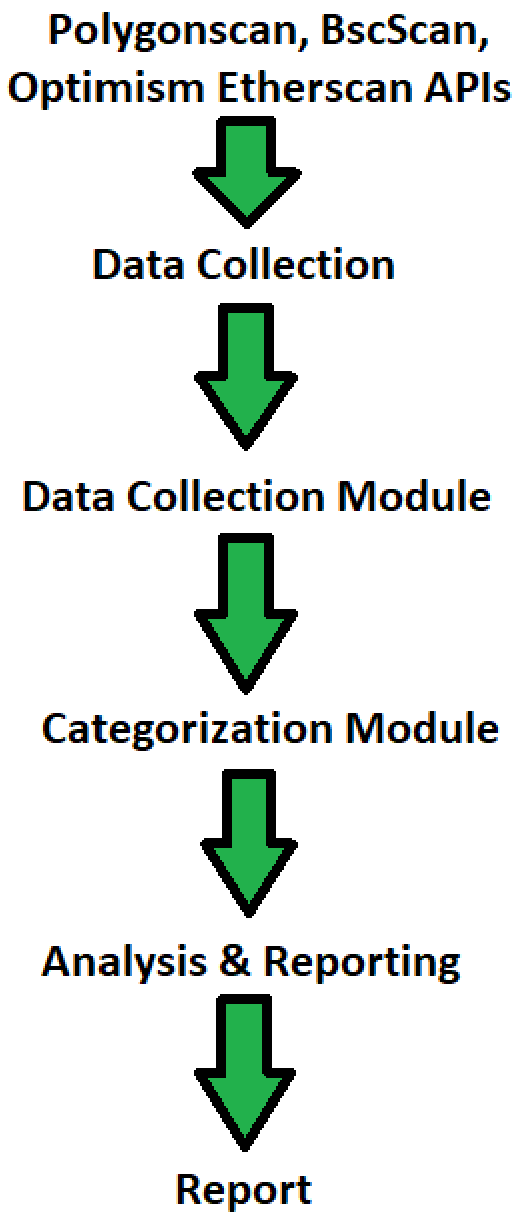
Diagrams/Flowcharts:

System Flow



Categorization Workflow

High-Level Diagram



Module Breakdown:

- **Data Collection:** It uses API calls to pull metadata.
- **Processing & Categorization:** Categorizes the contract features (parameters such as the type of standard the contract follows such as ERC-20 or ERC-721).

- **Visualization:** Tends to produce graphs and charts for use in analysis and reporting of trends.

Methodology and Detailed Design

Requirements Analysis:

- **Software:** Python, Pandas, Matplotlib, Seaborn, API Libraries (Requests) None.
- **Hardware:** Systems that can handle datasets with records exceeding thirty thousand in count.

Model/Algorithm Specification:

- **Categorization Algorithm:** Seven metadata fields including ABI and contract type are parsed using a custom script to classify the contract.
- **Data Validation:** Checksum matching prevents error by comparing results and duplicate filtering removes any data that is duplicated.

Development Environment and Tools:

- Python 3.x
- **Libraries:** Imports, Libraries Used: Requests, Pandas, Matplotlib and Seaborn

Repository Link:

https://github.com/arham017/Analysis-and-Categorization-of-Smart-Contracts-EVM-based-Blockchain-Networks-

Future Work and Next Steps

Phase 3 Roadmap:

- An analysis of the category of contracts as per their performance.
- Emphasis is placed on identifying loopholes and flaws, which are often embodied in contracts.
- Machine learning for automated classification which enhances the logistics of the process.

Remaining Milestones:

- Detailed analysis of contract type of all three networks.
- Setting up a well-grounded reporting system for stakeholders.

Conclusion

The overall goal of the project has been accomplished as it was designed to classify smart contracts of Binance, Polygon, and Optimism separately using the metadata obtained from blockchain analytics. By connecting APIs and creating categorization algorithms, contracts were successfully sorted into relevant categories like tokens and NFTs or used in DeFi. It was a revelation on its own and was able to make some sense of different trends and uses of smart contracts of these blockchains.

However, there were obstacles including API rate limits and missing metadata and we applied remedies for the problems for example using cache methods and backup classification on artifacts. These adaptations guaranteed that the system was strong and could be expanded as and when needed. Initial testing of our prototype showed a high degree of accuracy with regards to contract classification, which suggests that this paradigm can be extended for other purposes, related to blockchain analysis.

As future work, the project provides groundwork for pursuing research directions, including: expanding categorization to more blockchains, improving the algorithm for better accuracy, and developing a live monitoring framework of smart contract activity over time. This work will benefit developers, researchers, and another user of blockchain technology and can help advance understanding of the decentralized ecosystem.

References

Alharby, M., & Van Moorsel, A. (2017). Blockchain-based Smart Contracts: A Systematic Mapping Study. ArXiv. https://doi.org/10.5121/csit.2017.71011

de Sousa Matsumura, G., dos Santos, L.B.R., Conceição, A.F.d., Vijaykumar, N.L. (2021). Vulnerabilities and Open Issues of Smart Contracts: A Systematic Mapping. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2021. ICCSA 2021. Lecture Notes in Computer Science(), vol 12952. Springer, Cham. https://doi.org/10.1007/978-3-030-86973-1_1

M. Alharby, A. Aldweesh and A. v. Moorsel, "Blockchain-based Smart Contracts: A Systematic Mapping Study of Academic Research (2018)," 2018 International Conference on Cloud Computing, Big Data and Blockchain (ICCBB), Fuzhou, China, 2018, pp. 1-6, doi: 10.1109/ICCBB.2018.8756390. keywords: {Smart contracts;Blockchain;Bitcoin;Systematics;Databases;Peer-to-peer computing;Blockchain;Smart contract;Systematic mapping}