

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
II SEMESTER 21-22
DSE CL ZG628T DISSERTATION

Dissertation Outline

BITS ID No. 2019HC04178

Name of Student: Anjali Sunder Naik

Email ID: sundernaik.anjali@in.bosch.com

Mobile: +91-8147757114

Name of Supervisor: Dr. Vishnu Prasad V J

Designation of Supervisor: Project Manager

Qualification and Experience: PHD Physics, Senior Technical Architect

E- mail ID of Supervisor: vishnu.prasad@in.bosch.com

Mobile Number of Supervisor: +91-7598033881

Topic of Dissertation: Analysis, detection & mitigation of felonious wallet accounts over the Ethereum blockchain network using machine learning techniques

Name of First Examiner: _____

Designation of First Examiner: _____

Qualification and Experience: _____

E- mail ID of First Examiner: _____

Name of Second Examiner: _____

Designation of Second Examiner: _____

Qualification and Experience: _____

E- mail ID of Second Examiner: _____

(Signature of Student)

(Signature of Supervisor)

Date: 26/11/2021

Date: -----

Problem statement (what is the problem being addressed)

As of 2021, a survey from Coin Market Cap indicates that there are nearly over 6,000 digital coins in the market, a severe increase from just a handful since 2013. However, a large portion of these cryptocurrencies might not be that significant. The total market cap of all the crypto assets, including stable coins and tokens has shown a significant rise from year 2020 and has hit 2.4 trillion. Cryptocurrencies has vast potential of revolutionizing and transforming compliance-free peer-to-peer transactions. However, an end user must overcome certain challenges related to privacy, security, and control. As the transactions are recorded in a publicly distributed ledger known as blockchain, hackers have a large attack surface to gain access to critical and sensitive data. In the rapidly growing crypto currency space, the technological advent of cryptocurrencies and their respective benefits has been veiled with several illicit financing activities operating over the network such as ransomware, terrorist financing, hacking, data manipulation during transaction process, phishing, fraud, money laundering, bribery etc. Chainalysis, a firm that tracks every crypto currency transaction and serves as an advisor to an array of government authorities has published a report that shows that the amount of cryptocurrency spent on dark net markets rose 60% to reach a new high of \$1.15billion from July 2020 to June 2021.

Business process flow, if any

NA

Objective of the project

In this work, the primary focus is on the Ethereum network, which has seen over 1373 million transactions since its inception. Propelled with the rise in use of machine learning techniques in the research dimensions of financial domain, this is an attempt to explore the possibility to use various machine learning algorithms to analyze and detect the illicit accounts using the transaction history. Many criminals don't typically transfer funds directly to and from their linked addresses when transacting with regulated exchanges. A vast majority of bad actors will move their funds at least one time. CipherTrace analysts found that a typical cryptocurrency exchange's dark market exposure will typically double at two hops out (transactions once removed from the exchange). Various machine learning algorithms are evaluated on publicly available accounts flagged by the Ethereum community for their illegal activity coupled with valid accounts.

Uniqueness of the project

- This project will provide a compact comparison of various machine learning techniques to identify fraudulent activities in the Ethereum network.

Benefit to the organization

The analysis would benefit the organization in below ways

- There would be more trust and worldwide adoption of the distributed ledger technology and possible regulation can be achieved in its usage.
- The illicit wallet accounts can be tracked and validated which would internally enable the law enforcement agencies take appropriate actions on the fraudulent activities
- The analysis data would serve as a benchmark for future analysis.

Scope of work

- Collection of blacklisted Ethereum wallet accounts data and its relevant historical transactions.
- Cleansing of data and identification of relevant attributes
- Data preprocessing & raw feature extraction
- Identification of Machine Learning algorithms for the obtained data set
- Data Analysis using various algorithms and evaluation of key metrics
- Hyperparameter tuning to identify ideal model architecture
- Visualizations of results through a user interface

Resources needed for the project, including people, hardware, software, etc.

Programming Language	Python
Hardware+ Cloud Platform	Google Collab+ GPU

Potential challenges & risks in doing the project

- Collection of meaningful data and the amount of data collected
- Scaling data

Background of previous work done in the chosen area

- Work on various blockchain platforms to understand the internals of distributed ledger technology

Solution architecture, if any

NA

Detailed Plan of Work (as follows)

#	Task	Expected date of completion	Names of Deliverables
1	Research and data collection from various sources	3 weeks	Research & Data
2	Data cleansing & identification of data attributes	2 weeks	Feature extraction
3	Identification & Analysis of various Machine Learning algorithms for the obtained data set	6 weeks	Models
4	Hyperparameter tuning to identify the best model architecture	3 weeks	Optimal Model
5	Visualization of results	2 weeks	Metrics UI

Supervisor's Rating of the Technical Quality of this Dissertation Outline

EXCELLENT / GOOD / FAIR/ POOR (Please specify): _____

Supervisor's suggestions and remarks about the outline (if applicable).

Date_____

(Signature of Supervisor)

Name of the supervisor:

Email Id of Supervisor

Mob # of supervisor: