Logbook

Napatchol Thaipanich (nt375)

2/7/25

- Desired Outcome: ML model to predict whether a token is a scam or not and categorise it (can't tell exactly which one is a scam; it has to be a percentage)
- Requires raw data and actual transactions with a mix with scam and non-scam: Currently, no dataset; need to trace back to transactions from the scam token summary and transaction address
- Rug-Pull-Dataset can be used to categorise
- Consider integrating **blockchain transaction analysis** and **social media sentiment analysis**.
- TODO:
  - Tell all the students I left to ask the professor to support
    - Clear the storage
    - Based on this list: https://www.comparitech.com/crypto/cryptocurrency-scams/
    - Get transactions for each month and map them with the lists, including the above and others that the professor provides.

2/14/25

- Explore APIs to retrieve **transaction data, contract information, and Twitter data**, as demonstrated in the attached YAML file and image. I also came across a **Worldwide Crypto & NFT Rug Pulls and Scams Tracker** website, https://www.comparitech.com/crypto/cryptocurrency-scams/
- ![[Screenshot 2568-02-14 at 01.04.08.png]]
- TODO:
  - academic literature and walkthroughs of bitcoin/blockchain forensics processes - to help identify the different types of blockchains - and scoping the statement
  - determine the indicators of soft/hard scams

4/6/25

- Understanding the requirement of critical review

4/7/25

- Read the article, trying to identify the scope of the project

4/8/25

- discuss with professor
  - questions that I have to define in the critical review (literature review)
    1. Purpose
       - What do you want to evaluate, the tool, or analyze the entire block entered to define the dataset?
    2. dataset
       - Amount of Transaction involved: The number of blocks I have to query

        - Amount of platform involved: such as Ethereum, Serena, Bitcoin, or all
    3. period of the data collection

4/9/25

- Read the article, trying to identify the scope of the project

4/10/25

- Scope the topic by writing a critical review

4/14/25 - 4/15/25

- Get feedback from the supervisor, and discuss which paper I should use in a critical review
- These are three papers that the supervisor suggested:
    - Taxonomy of Anti-Fraud Measures within Token Economy: Insights from Rug Pull Schemes
    - Analytical Tools for Blockchain: Review, Taxonomy, and Open Challenges
    - Token Spammers, Rug Pulls, and Sniper Bots: An Analysis of the Ecosystem of Tokens in Ethereum and in the Binance Smart Chain (BNB)
- Summarize those three papers again in detail

4/16/25 - 4/18/25

- Rewrite the critical review and send it to the professor to discuss further.

4/26/25

- Having a meeting with the professor
    - Feedback on critical review
        - Summary:
            - Statistics in the paragraph should be cited [1]
            - In the conclusion, more papers may be added that support this paper's summary.
        - Critique:
            - Add a positive point.
            - It should balance the positive and the improvement points.
            - Suggesting the paper on the on-chain point: find it.
        - Synthesis
            - Idea: Verify the transaction or only the data analysis
            - It can be further applied to code analysis, including smart contracts (involving code security), and maybe to statistics that support testable claims or to finding examples.
        - Professor suggesting: (Does it relate to my ideal project or not)
            - https://dl.acm.org/doi/abs/10.1145/3623376?casa_token=chlkZrNLzloAAAAA%3A2oUVW7unK12AeVgL7WSRMUVhcc7razn5wB1avzcvZQojQD-lkOxLVSNnEgKJ5tMZ8PsL7SDVsLZv
            - https://ieeexplore.ieee.org/document/8823898/
- Alongside the critical review:
    - should find the details of

        - The way to gather the dataset
            - Cost of gathering the dataset or
            - What is the limit and time to consume if it's free, and how much is it offered?
- TODO:
    - Read the articles that the professor suggested
    - Groom the synthesis
    - Revise the critical review following the professor's suggestion
    - Send it to the professor by Wednesday

4/28/25 - 4/30/25

- revise critical review
    - add the number of citation
    - research more about on-chain prevention mechanisms

5/2/25

- revise from feedback
    - summary:
        - Clearly state the value of unifying language in academic research (science communication benefits).
        - Highlight that the taxonomy aids in comparing methods and their benefits/drawbacks across studies.
        - Identify whether the authors suggest overlapping toolsets or hybrid approaches to address research gaps.
    - Critique:
        - Assess if the authors' claimed contributions align with their motivation and evidence.
        - Replace vague praise like "inspirational starting point" with measurable and specific benefits (e.g., rigor, novelty).
        - Explain why the paper is compelling—what makes it useful: structure, research gaps, or unique classification?
        - Justify your concern with small sample size (10 papers) by:
        - Doing background reading on other taxonomies/literature reviews for comparison.
        - Discuss the inclusion of preprints (e.g., arXiv):
        - What are the pros/cons of using non-peer-reviewed sources?
        - What validation steps would be needed to use them responsibly?
        - Briefly reflect on the value of peer review:
        - Does peer review guarantee quality?
        - What are its limitations?
        - Why might peer-reviewed work still need scrutiny?
        - Expand critique on lack of on-chain examples:
        - Does the scope of the paper exceed its evidence?
        - Would narrowing the scope or presenting on-chain as a theoretical goal be better?
    - Synthesis:
        - Elaborate more fully on the limitations and implications:
        - What are the difficulties in developing cross-layer (off-chain, transactional) detection tools?

        - Why might the authors not have pursued implementation?
        - Is there a current state-of-the-art in on-chain detection, or is the gap real and limiting?
        - Discuss the challenge of predicting rug pulls in chaotic token markets:
        - Compare forensic (post hoc) vs predictive methods (real-time detection).
        - Emphasize the need for contextual and code-based evidence for better prediction.
    - Final Touch:
        - Add citations or examples (e.g., tools like UNCX) where helpful.
        - Maintain academic tone and clarity throughout.

5/26/25

- think about the project plan

5/27/25

- write and email to supervisor asking about early deliverable submission
- Tasks:
    - Complete your initial survey of rug pull scams and the means by which they were detected
    - Identify where quantitative techniques can be used to predict rug pull scams
    - Identify whether code analysis can help protect against rug pulls by indicating the existence of backdoors or policy loopholes

6/2/25

- 2 Workshops
    - Social, ethical, and professional practice issues and considerations
    - Data gathering and analysis (including basic statistics)
- Do the initial survey of rug pull scams and how they were detected
- Identify where quantitative techniques can be used to predict rug pull scams.

6/3/25

- 1 Workshop
    - Dissertation writing

6/6/25

- Ask for the early deliverables.
    - Early deliverables are about research, demonstration or prototype.
    - Question:
        - What do you want to do? Pick one of the research topics
        - What are the existing solutions?
        - Any technique to use in pre-processing
        - How much data is required
- Todo:
    - A draft and an idea to ask the above questions
    - Within 2 weeks or early (if it is possible to answer all those questions)

6/9/25 - 6/13/25

- Exploring research to find the answer to the above questions by
  - collecting data
  - pre-processing
  - training basically

6/13/25

- Draft the proposal and timeline
- Mail to report to my professor
- Question from my supervisor
  - The goal is to extend Gigahorse with rug-pull-specific code features and identify potential rug-pull-related functions.
  - Do I understand your goals, or did you have something else in mind?
  - What data does the CRPWarner dataset provide that is useful to your project?
  - What will you use EtherScan for? Identifying new projects or directly sourcing code?
  - How do you plan to link these elements?
  - Do you plan to implement a toolchain that periodically scans for new projects, attempts to find code, and then performs feature extraction and evaluation to identify rug pull features?

6/15/25

- Draft Email to answer above questions and send to my supervisor

6/16/25

- write a plan including
  - timeline
  - tasks in detail
- Task: exploring more feature extraction technique, and more

6/17/25 - 7/1/25

- do the experiment exploring

  - feature extraction techniques like

    - Opcode-Based Features
    - Gigahorse Features (Static Code Analysis)
    - Bytecode-Level Features
    - Source Code Features (if verified source exists)

  - simple model like

    - Traditional Machine Learning Models

Model	Notes
Random Forest	Good baseline; handles tabular features and non-linear interactions

Model	Notes
XGBoost / LightGBM	Efficient and high-performance on sparse/opcode frequency vectors
Logistic Regression	Works well with binary/multiclass labels for simpler baselines
SVM (Linear/RBF)	Effective with high-dimensional opcode vectors (requires scaling)
KNN	Simple but less effective on large datasets with high dimensionality

- Deep Learning Models (MLP & CNN)

Model Type	Input Type	Suitable For	Source
CodeBERT	Solidity Code	Syntax & semantic analysis	CodeBERT (2020)
GraphCodeBERT	Code + Data Flow	Logic flow, function detection	GraphCodeBERT (2021)
GNN (e.g., GCN)	Gigahorse output	Control/Data flow-based detection	IEEE Access 2021
CNN	Bytecode	Pattern recognition	arXiv 2105.02917
LSTM/GRU	Opcode sequence	Sequence anomaly detection	NDSS 2018 Workshop
ByteTransformer	Bytecode tokens	Bytecode-level classification	arXiv 2305.14767

- Sequence Models (RNN, LSTM, GRU)

- In sum up,

Source	Example Feature Types	Best for Models
Bytecode	Opcode freq, entropy, n-grams	XGBoost, MLP, CNN
Gigahorse	Storage access, JUMPs, selfdestruct	RF, GNN, MLP
Source Code	Modifiers, regex flags, token logic	CodeBERT, Regex + RF
Graph	CFG size, backdoors, loop detection	GCN, GAT
Custom Rules	Mint/Leak/Limit flag heuristics	Any classifier

6/17/25

- Having a meeting with my supervisor
  - Showing my experiments during the past week, and told him about my concern, and

- Defined the tasks of the early deliverables
    - Project plan: Plan for July and August. Link actions to your project aims. This should include research, design, implementation/experimentation, and validation phases as appropriate.
    - Project documentation: can be the summary
    - Any software or corpus items: Python notebook, experiment note

7/2/25 - 7/4/25

- Do a project plan for July and August, including: (***Link actions to your project aims***)
    - research phases
    - design phases
    - implementation/experimentation phases
    - validation phases
- Do a Project documentation
    - Lit Review
    - Problem analysis and/or early design documents
    - a clear project idea
    - requirements needed to finish the project,
- code part, techniques level

7/7/25 - 7/17/25

- Do the experiment on groundtruth of CRPWarner about feature extractions and models

7/18/25 - 7/21/25

- Do the experiment on groundtruth + large sample of CRPWarner about models which pick up from the previous week
- Do pseudo-label on CRPWarner Large

7/21/25

- summarised the professor's feedback
    - literature review
        - content:
            - should provide information on the history and methods of cryptocurrency-related rug pull scams, including the difficulties in proving that a scam has occurred – What evidence is required to identify a rug pull scam? -- the attributes of rug pull scams and the key indicators
            - section: the limitations of existing tools. -- RPHunter, CRPWarner and other tools
                - RPHunter: e.g. only able to operate over transaction logs and graphs. -- restricting them to a forensic/retrospective role.
            - the issue that my tool aims to solve and how it impacts
            - any additional information about supervised and semi supervised learning
        - how to write:
            - summarises the key findings
            - identifies the project scope

  - implementation:
    - add more validation/test strategy (before the end of the month) -- according to a set of goals
- TODO:
  - the intentions behind your tool
  - identify the limitations of existing tools. -- RPHunter, CRPWarner and other tools
  - think about any technical challenges or questions you might want to discuss
  - We should discuss what **the main research question** for your project is in our next meeting
    - e.g.
      - Is there a higher level of evidence/detail that you believe would benefit any forensic/prosecution process aimed at rug pull scammers?
      - What challenges do existing methods have in providing evidence, and how do you think these could be reduced?
  - discuss what the specific functional objectives of your tool are.
    - do you have any target false report rates or accuracy metrics that you consider a minimum acceptable limit?
    - test strategy will allow you to identify where tests are intended to indicate a simple pass or fail, or whether they are there to characterise your tool
      - providing statistics or qualitative data which allows you to identify what it is capable of, and how it compares to existing approaches
        - optimise
        - increase the level of detail
        - speed up the analytical process
        - Which attributes are most desirable for a tool that would either be used in a forensic context?
  - limitation of contract analysis
    - ??? in case the source code is not public should be switch to transaction or not
    - Are you able to identify any estimates regarding the proportion of public vs private contract code?
    - What does this mean for attempts to determine if rug pulls are occurring – can this only be done over public contracts?

7/22/25

- The competitor analysis should depict what they can do to work, what cases do not work, and what is my tool to solve that problem. e.g. CRPWarner is the tool to use when analysing source code before it announces idea:
- get the bytecode from the chain and reverse engineer to get source code <-- The plan has changed from last week. TF-IDF from source code is the best technique to extract features, but the percentage of open source is low, around 1 - 5%. Covert bytecode to source code might be the key.
- The professor's already agreed on the idea to combine with the transaction behaviour analysis.
- should observe the result, and the ML should notice the difference between a rug pull event and an organic failure TODO:
- statistics from the final tool
- draft ui

7/22/25 - 8/5/25

- getting source from chain using etherscan v2
- do the pipeline for feature extraction
- Experiment with Torch graph classification on my laptop, but sometimes segfaults happen, so i decided to extract features from the graph and push them to ML, like xgboost
- Try to use my fusion to do pseudo-labelling
- I do evaluate on a small test, 0.2 of 69 items. We will conduct a more thorough valuation of the added pseudo-labelled data in each round.
- new I am not satisfied with this result much

8/5/25

- having meeting with my professor

think about/ todo:

- indication of my tool
    - potential role
    - benefit of my tools
    - fields that the model used to concern to classify label
    - low positive fault, ideally
    - set up document

8/5/25 - 8/15/25

- Do the ui demo
- Set up the installation README markdown sheet
- Adjust the threshold and return the threshold of each label as the professor required
- Run self learning and expanding labels CRPWarn + RPHunter

8/15/25

submit the corpus in the demo ui and backend by using gitlab and the config on the moodle

8/19/25

- Fix installation sheet
- Updata dataset V5
- Set the structure of dissertation
- Clean up the logbook
- feedback on dissertation outline:
    - literature review
        - after analysing the research gap: Mention back to the research question
        - Describe about optuna and tools
    - methodology
        - try to fill by diagrams e.g. classes
    - writing order: literature review --> methodology --> implementation --> introduction --> conclusion

8/20/25

- Complete literature review chapter

8/20/25 - 9/1/25

- Complete whole dissertation including 5 chapters and abstract

8/26/25 - 8/31/25

- record and edit live demo