NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

Network Analysis and Traceability

On Ethereum Blockchain

Student: Su Voon Hou

Supervisor: Dr Sourav Sen Gupta

PROJECT OBJECTIVES

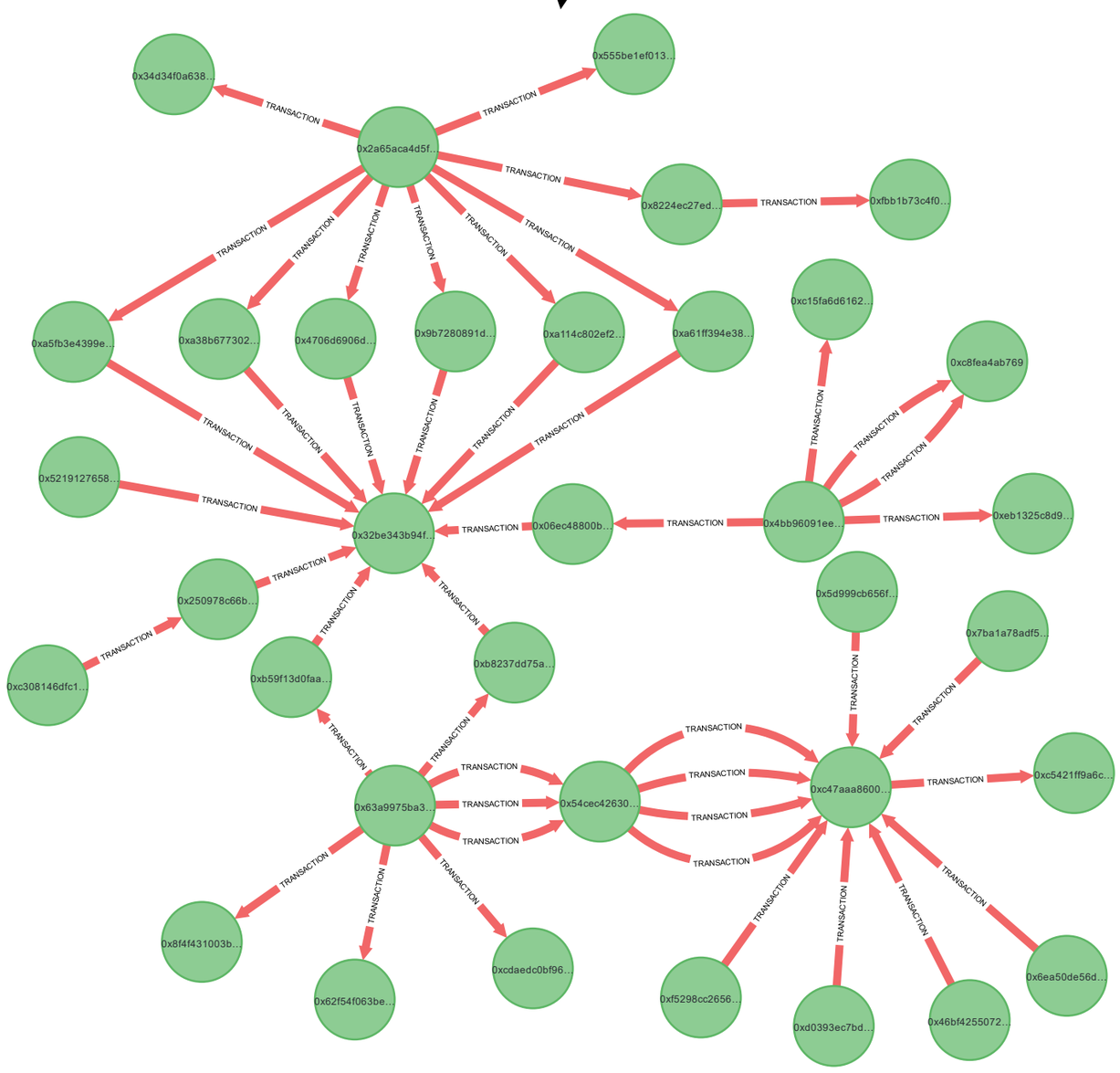
The goal of the project is to implement an end-to-end analytics toolbox, named **EtherNet** to perform data mining, network and traceability analysis on Ethereum by automating the transformation of tabular blockchain data into its graph representation. **EtherNet** aims to:

1. Provide a scalable and efficient storage system for Ethereum blockchain data
2. Provide a consistent access layer for ETL workflows
3. Enable discovery and documentation of existing graphs

EXISTING WORKFLOW ISSUES

1. Lack of automated ETL workflows
2. Difficult to extract required data
3. No means of managing downloaded data

from_address	to_address	other_data	block_number
0xd3b1fad2...	0x1625a9f7...	...	0
0x4bc3c203...	0xfe611a3d...		1
0x40af81b3...	0x5716678d...		2
0x9786a242...	0xa25a8dcb...		3



ETHERNET'S AUTOMATED END-TO-END WORKFLOW

1. User uses the Python SDK to submit an ETL request
2. Missing data in local Hive database is identified
3. Missing data is downloaded from Google BigQuery
4. Missing data is added to local Hive database
5. Required data to construct network/graph is exported to CSV files
6. Network/graph is constructed in Neo4j

Data Compression Ratios

- 5.5 token_transfers
- 11.8 traces
- 3.7 transactions