Haruspicy of Algorand Network via Topological Data Analysis

Double Master Maths & Info MAFI 2023

1 Introduction, Motivation and Related Works

Topological Data Analysis (TDA) is a mathematical technique that is used to analyze complex data sets and identify patterns and structures in the data [1, 11, 2, 3]. It uses tools from topology [6], which is a branch of mathematics that studies properties of shapes and spaces that are preserved under continuous deformations.

TDA can be used to analyze data from a wide range of fields, including biology, physics, engineering, sports and economics. It can be used to identify clusters or groups of data points, detect boundaries or shapes in the data, and extract relevant features or variables that capture the underlying structure of the data. TDA has proven to be a powerful tool in machine learning, data science, and data visualization.

Graphs and networks structures have been successfully analyzed by means of TDA as witnessed by the following papers [10, 8, 4, 7, 1].

2 Algorand Blockchain

Algorand (`http://algorand.com`) also has a transaction graph that represents the flow of transactions between Algorand addresses. The Algorand transaction graph is a directed graph, where each node represents an Algorand address and each directed edge represents a transaction between two addresses.

Like the Ethereum transaction graph, the Algorand transaction graph is also a complex and dynamic network, with new transactions being added to the graph continuously as new blocks are added to the Blockchain. The graph can be analyzed using various techniques, including Topological Data Analysis (TDA), to identify patterns and structures in the data. This is exactly the purpose of the current project.

By analyzing the Algorand transaction graph using TDA, it is possible to identify clusters or groups of addresses that are likely controlled by the same entity, as well as detect anomalies or outliers in the transaction patterns that may indicate fraudulent or malicious activity. This information can be used to improve the security and stability of the Algorand network and to develop more effective tools and techniques for analyzing and monitoring Blockchain networks.

The Algorand Blockchain is a public blockchain, which means that anyone can access the transaction data and analyze the transaction graph. There are several ways to access the transaction data for Algorand, depending on your needs and expertise.

One way to access the transaction data is to use a blockchain explorer that supports Algorand. A blockchain explorer is a web application that allows users to view transaction data and other information about a blockchain. Some examples of blockchain explorers that support Algorand are AlgoExplorer (`https://algoexplorer.io/`) and PureStake Explorer (`https://goalseeker.purestake.io/algorand/mainnet/`).

Another way to access the transaction data is to use a blockchain API that provides programmatic access to the blockchain data. Some examples of blockchain APIs that support Algorand are AlgoExplorer API, AlgoDex API, and PureStake API.

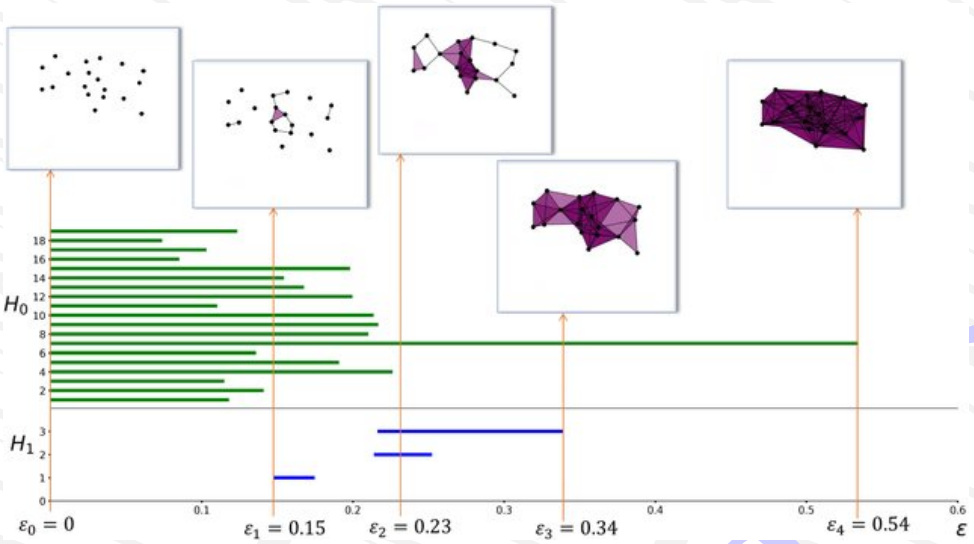

Figure 1: From data clouds to barcodes (example from [12]).

If you are interested in conducting more advanced analysis of the transaction graph using Topological Data Analysis (TDA), you may need to use specialized software tools or develop your own algorithms. There are several open-source TDA software packages available, such as GUDHI (`https://gudhi.inria.fr/`) and TDAstats (`https://cran.r-project.org/web/packages/TDAstats/index.html`), that you can use to analyze the transaction graph data.

3 Project and Research Proposal

The group of MAFIs students would need to conduct research that makes (significant) contributions to the field of Topological Data Analysis as applied to Blockchain Technology. This would require identifying a novel research question or problem that has not yet been addressed, and developing an innovative approach to analyzing the transaction graph data.

Some potential directions for research and works in this area could include (but is not limited to):

- Developing new TDA algorithms or methods that are specifically designed for analyzing blockchain transaction graphs.
- Investigating the relationship between the topological properties of the blockchain transaction graph and various network characteristics, such as degree distribution, clustering coefficient, or centrality measures.
- Developing new metrics or measures to evaluate the performance of TDA algorithms in the context of blockchain transaction graph analysis.
- Applying TDA to analyze the transaction graph data of specific blockchain networks, such as Ethereum (see [5]) or Algorand, to identify patterns or structures that are unique to those networks.
- Exploring the potential applications of TDA in blockchain, such as identifying fraud or money laundering activities, predicting network congestion or transaction fees, or improving the scalability and efficiency of the blockchain network.
- Developing TDA-based approaches for clustering and classifying blockchain addresses based on attributes such as transaction history, balance, or social network.
- Evaluating the robustness and vulnerability of blockchain systems to targeted attacks using TDA-based models and simulation tools.

- Investigating the use of TDA in the analysis of cross-chain transactions and interactions between different blockchain networks.
- Developing TDA-based methods for predicting the behavior and actions of individual blockchain users, such as predicting which addresses are likely to become involved in fraudulent activity.
- Applying TDA to the analysis of smart contracts and decentralized applications (DApps) built on blockchain networks, in order to identify patterns and structures in the interactions between contract code and user behavior.

In addition to conducting novel and innovative research, it is also important to present your results in a clear and concise manner, and to provide rigorous experimental evaluation and validation of your methods. This will increase the likelihood that your research will be accepted by the scientific communities. Note that there are existing works on TDA and blockchains networks including [9, 5]. In particular, the authors of [5]

1. offer a novel perspective to risk analysis of crypto-assets by "dissecting" hidden linkages between *token price* and *local geometry* of the Ethereum transaction graph,
2. propose the *Betti pivot* as a new illustrative measure on the Ethereum transaction graph and
3. report first results for crypto-token price anomaly detection.

Note that this project is very similar to a twin project but on Tezos Blockchain.

References

[1] G. Carlsson. Topology and Data. *Bull. of the AMS*, 46(2):255 – 308, 2009.

[2] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 454 – 463, 2000.

[3] R. Ghrist. Barcodes, the persistent topology of data. *Bull. of the AMS*, 45(1):61 – 75, 2008.

[4] Risi Kondor, John C Miller, and Eric B Foxall. Spectral networks and topological data analysis. *Journal of Machine Learning Research*, 17(1):771–798, 2016.

[5] Yitao Li, Umar Islambekov, Cuneyt Gurcan Akcora, Ekaterina Smirnova, Yulia R. Gel, and Murat Kantarcioglu. Dissecting Ethereum Blockchain Analytics: What We Learn from Topology and Geometry of the Ethereum Graph? In *SDM*, pages 523–531. SIAM, 2020.

[6] J. P. May. *A Concise Course in Algebraic Topology*. The University of Chicago Press, 1st edition, 1999.

[7] Facundo Memoli and Sayan Mukherjee. Exploring topological order: Graph entropy measures and persistent homology. *IEEE Transactions on Information Theory*, 57(10):6772–6787, 2011.

[8] Sayan Mukherjee and Josh Reed. Topological data analysis of biological networks. In *2009 IEEE International Workshop on Genomic Signal Processing and Statistics*, pages 27–30. IEEE, 2009.

[9] Dorcas Ofori-Boateng, Ignacio Segovia-Dominguez, Cuneyt Gurcan Akcora, Murat Kantarcioglu, and Yulia R. Gel. Topological anomaly detection in dynamic multilayer blockchain networks. In *ECML/PKDD (1)*, volume 12975 of *Lecture Notes in Computer Science*, pages 788–804. Springer, 2021.

[10] Mason A Porter, Jesse H Bettencourt, and MEJ Newman. Persistent homology of complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(09):P09005, 2006.

[11] A. Zomorodian. Topological data analysis. *Adv. in Appl. and Comp. Topology*, 70:1 – 39, 2007.

[12] İsmail Güzel and Atabey Kaygun. A new non-archimedean metric on persistent homology. *Computational Statistics*, 37:1963 – 1983, 2022.