

Temporal Analysis of Transaction Ego Networks with Different Labels on Ethereum

Baoying Huang*, Jieli Liu[†], Jiajing Wu[§], Quanzhong Li*, Hao Lin[‡]

*School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China

[†]School of Software Engineering, Sun Yat-sen University, Zhuhai 519082, China

[‡]Merchants Union Consumer Finance Company Limited, Shenzhen 518000, China

[§]Corresponding Author: wujiajing@mail.sysu.edu.cn

Abstract—Due to the widespread use of smart contracts, Ethereum has become the second-largest blockchain platform after Bitcoin. Many different types of Ethereum accounts (ICO, Mining, Gambling, etc.) also have quite active trading activities on Ethereum. Studying the transaction records of these specific Ethereum accounts is very important for understanding their particular transaction characteristics, and further labeling the pseudonymous accounts. However, traditional methods are generally based on static and global transaction networks to conduct research, ignoring useful information about dynamic changes. Our work chooses six kinds of important account labels, and builds ego networks for each kind of Ethereum account. We focus on the interaction between the target node and neighbor nodes with temporal analysis. Experiments show that there is a significant difference between various types of accounts in terms of several network features, helping us better understand their transaction patterns. To the best of our knowledge, this is the first work to analyze the dynamic characteristics of Ethereum labeled accounts from the perspective of transaction ego networks.

I. INTRODUCTION

After the development of more than a decade, blockchain technology has become a buzzword and brought world-shaking innovations to human life. Generally speaking, blockchains are append-only distributed ledgers originated from Bitcoin [1]. Transactions on blockchain platforms are publicly verifiable based on open data architecture. Ethereum [2] is the largest public blockchain enabling Turing-complete smart contracts, which are computer programs that can automatically be executed in the Ethereum Virtual Machine (EVM). Since the outbreak of COVID-19, investors have found that Ethereum is suitable as a short-term safe-haven [3].

Therefore, increasing research efforts have been devoted to the analysis of the publicly accessible Ethereum transaction records. A considerable part of existing work modeled the transaction data into networks, which are a general language in describing interacting systems [4] [5] [6]. In the Ethereum transaction network, an Ethereum account is regarded as a node, and the edges between two nodes present the interactions between the two accounts with attributes such as transaction value, timestamp and so on.

Recent research [7] [8] [9] have analyzed the entire Ethereum transaction network from a static perspective for understanding the whole Ethereum ecosystem. They have built several static networks using transaction data, and analyzed network properties such as density and clustering coefficients,

comparing them with social networks and the Web. However, the daily new transaction volume of Ethereum is quite large, and static analysis leads to the loss of insights from a dynamic evolution perspective. Consequently, several recent studies [10] [11] [12] provided insights into the temporal and evolution properties of the entire Ethereum transaction network. But there is few studies investigating the microscopic structure of a kind of nodes in the Ethereum transaction network, which can give explanations for the different behaviors of accounts. Analyzing the Ethereum transaction network from a perspective of various types of nodes can provide us with a new understanding of their microscopic network structure, which also helps to label the pseudonymous accounts in Ethereum. In this paper, we are interested in studying the temporal changes of multiple transaction characteristics based on the ego network of each labeled account. The ego network consists of the central node “EGO” and the nodes directly connected to EGO (these are called “ALTER”), as well as the connections between the EGO and ALTER nodes, ALTER and ALTER nodes [13]. Ego network is widely used in social network analysis and helps to find many patterns [14] [15]. We mainly select six typical Ethereum account labels, including four normal labels (ICO, Mining, Gambling, Exchange) and two illegal labels (Ponzi and Phish), which are collected from Etherscan’s label word cloud¹. A normal account has no illegal activities such as deceiving users. While an illegal account is the opposite. Then, we collect the required transaction records from a blockchain data platform and build multiple ego networks. After analyzing how the transaction features change over time for each type of node, we discover the different features between various types of accounts and better understand their roles on Ethereum.

The remainder of this paper is structured as follows. Section II introduces the process of collecting transaction data and building the ego network of various accounts. The analysis of network properties will be given in Section III. Finally, we will give a brief conclusion in Section IV.

II. DATASET AND EXPERIMENT SETUP

In this part, we firstly introduce the source and processing of data, and then explain the definition of ego network

¹<https://cn.etherscan.com/labelcloud>

constructed by using transaction data.

A. Data Collection

Etherscan is one of the most widely used Ethereum browsers, recording lots of account labels that are reliably provided by multiple users. A lot of previous research work on Ethereum also used the labeled accounts on Etherscan [16] [17]. We choose six types of the most common and important accounts on Ethereum as the targets. Especially, according to the crime report about cryptocurrency [18], ponzi schemes and phish scams have defrauded millions of dollars. Now we will introduce the definition of these accounts in detail.

- **ICO:** The accounts used to receive crowdfunding for the initial offering of the tokens
- **Mining:** The accounts of miners who participate in block mining
- **Gambling:** The accounts of the gambling platforms
- **Exchange:** The accounts of exchanges that allow users to sell or buy cryptocurrencies
- **Ponzi:** The accounts related to the manipulator of ponzi schemes
- **Phish:** The accounts related to the manipulator of phish scams

In addition, we collect all target labeled accounts as well as their complete *Ether*² transaction records. In order to construct the ego network, we also collect the transactions between their first-order neighbors. These transaction records are extracted from the latest Ethereum dataset on XBlock³, which is a blockchain data platform. Therefore, a total of 2584 accounts and nearly 38 million transaction records are collected.

B. Construction of Ego Networks

Let $G = (V, E)$ represents the ego network of an Ethereum account, where V represents the set of nodes, and E is the set of edges. V contains the central node “EGO” and its neighbor nodes “ALTER”. The transactions between EGO and ALTER nodes, ALTER and ALTER form the edge set E . Particularly, an edge can be represented as $e = \langle u, v, w, t \rangle$, indicating that at time t , account u transfers Ether of amount w to account v . Consequently, the ego network in our work is a weighted directed network.

Fig. 1 shows an example of an ego network. Moreover, we construct an ego network for each labeled account, classify the ego networks of the same label into one category, and take the average of various statistics as the final feature value of the label accounts. The details of each feature will be shown in the next section.

III. THE ANALYSIS OF NETWORK PROPERTIES

First, we focus on the local clustering coefficient of each labeled account to determine whether the nodes in the network are closely connected. Then we analyze several transaction

²Ether is the official cryptocurrency of Ethereum.

³<http://xblock.pro/xblock-eth.html>

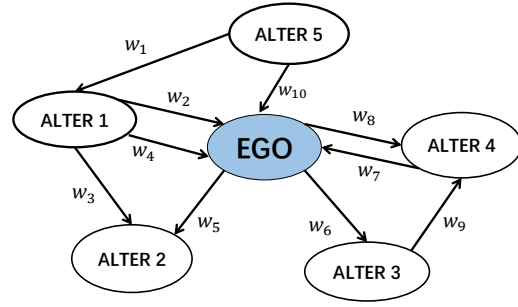


Fig. 1. An example Of ego network.

characteristics of labeled accounts. Finally, we obtain the unique transaction features of each labeled account, which also explain the role of the corresponding label in the Ethereum ecosystem.

A. Local Clustering Coefficient

The local clustering coefficient is a measure of the degree to which the neighbors of the central node in the graph tend to cluster together. Let $G_i = (V, E)$ represents the ego network whose ego is node i . Here the edge of node u pointing to node v is written as e_{uv} . Then N_i represents the neighbor nodes of node i and $k_i = |N_i|$. Hence the local clustering coefficient of node i is:

$$\tau_i = \frac{|\{e_{uv} : u, v \in N_i, e_{uv} \in E\}|}{k_i(k_i - 1)}. \quad (1)$$

Then we calculate the average local clustering coefficient of each type of ego network, and the results are shown in Table I. According to the table, the probability of transactions occurring between the neighbors of ICO accounts is about 18%, followed by Mining accounts and Exchange accounts. It tells us there is a higher probability that investors of ICO accounts will have transactions with each other. The coefficient of Gambling accounts is the smallest, which means the users participating in the gambling are generally not connected. Besides, the local clustering coefficient of Phish accounts is larger than that of Ponzi accounts, indicating that the neighbors of Phish accounts are also more likely to be connected. We speculate that there are fixed fraud partners in phishing activities.

TABLE I
THE LOCAL CLUSTERING COEFFICIENT OF LABELED EGOS

Label	ICO	Mining	Gambling	Exchange	Ponzi	Phish
τ	0.1812	0.1354	0.0242	0.1324	0.0379	0.1226

B. Ego's Life Cycle and Time Window

For an Ethereum account u with M transactions, we define its life cycle as $Life_u = t_M - t_1$, where t_M represents the time when the last transaction occurred and t_1 represents the time when the first transaction occurred. We count the life cycles of each kind of account and record its median and maximum values in Table II. It can be concluded that the life cycle of

TABLE II
THE LIFE CYCLE OF LABELED EGO (DAYS)

Label	ICO	Mining	Gambling	Exchange	Ponzi	Phish
Median	419.92	646.24	292.59	603.06	19.99	15.76
Mean	1700.18	2150.09	1527.35	2132.30	1642.22	1529.09

illegal accounts is much shorter than that of normal accounts, and their life cycle is less than 20 days. This is because the criminals usually transfer the illegal gains out of the accounts as soon as possible and no longer use the original accounts to enhance anonymity.

Afterwards, in order to observe the ego network of each type of label from a dynamic perspective, we refer to the previous work [12] and consider two types of time windows, the sliding window and the incremental window. Firstly, we take 1/5 of the life cycle of each type of labeled account as a fixed time span T , which is also the initial length of the two time windows. The sliding distance of the sliding window is T each time, while the size of the incremental window increases by T . Next, we use different time windows to extract transaction details of the egos, and obtain different characteristics in five phases.

C. The number of transactions between EGOs and ALTERs

In order to study the transaction activities of different types of accounts, we compare the evolution in the number of transactions between EGO and ALTER nodes during their life cycles in the incremental window mode. It can be seen from Fig. 2 that the number of transactions of illegal accounts is less than that of the normal accounts. In addition, the number of transactions on Exchange, Mining and Gambling accounts is relatively large. Because Exchange accounts are an intermediary on Ethereum, they generate a lot of transactions with users to buy and sell tokens. According to the description on Etherscan, we can know that Mining accounts are mining pools composed of many miners. So they usually return bonuses to miners multiple times. Additionally, Gambling accounts representing a kind of entertainment activity, also attract lots of users. Significantly, the curves of ICO accounts and Ponzi accounts are relatively similar, which shows that they have certain commonalities in the trading pattern. We presume that it is one of the reasons why it's difficult for users to distinguish whether the investment project is a scam or not.

For the purpose of further exploring the roles of the labeled accounts on Ethereum, the changes in the ratio of two kinds of transactions (in/out) in the incremental window mode are also recorded. An in-transaction is an interaction in which the EGO receives Ether, and an out-transaction is an interaction in which the EGO transfers Ether to another account.

The results in Table III show that only Mining, Exchange, and Ponzi accounts have a higher proportion of out-transactions than in-transactions. We suppose that the reasons for this phenomenon are as follows. Firstly, the rewards obtained from mining blocks will be distributed to users who

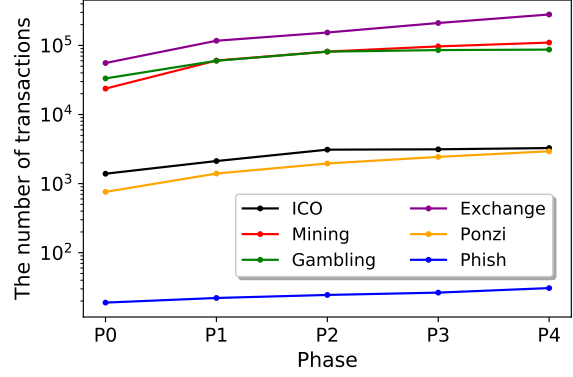


Fig. 2. Changes of the number of transactions at each phase

TABLE III
CHANGES IN THE PROPORTION OF THE NUMBER OF IN/OUT TRANSACTIONS

	Phase	P0	P1	P2	P3	P4
ICO	in	92.36%	90.83%	62.84%	62.92%	61.24%
	out	7.64%	9.17%	37.16%	37.08%	38.76%
Mining	in	0.19%	0.14%	0.17%	0.17%	0.16%
	out	99.81%	99.86%	99.83%	99.83%	99.84%
Gambling	in	55.73%	60.13%	61.45%	61.31%	61.27%
	out	44.27%	39.87%	38.55%	38.69%	38.73%
Exchange	in	43.83%	44.18%	41.34%	33.25%	26.67%
	out	56.17%	55.82%	58.66%	66.75%	73.33%
Ponzi	in	29.04%	28.28%	28.38%	28.98%	30.58%
	out	70.96%	71.72%	71.62%	71.02%	69.42%
Phish	in	61.15%	61.82%	61.46%	61%	57.78%
	out	38.85%	38.18%	38.54%	39%	42.22%

invest in miners, which is a long-term stable mechanism. Secondly, when an Exchange account is used as the initiator of transactions, it means that users are buying Ether from this platform. Generally speaking, investors will continue to purchase Ether multiple times and then sell them all at the highest price point. And thus Exchange accounts have more in-transactions than out-transactions. Thirdly, a Ponzi account intends to act like a normal project at first to attract more people. Then, a small part of the investment of latecomers will also be used as bonuses to return to the initial users several times.

Moreover, the in-transaction ratio of ICO accounts is relatively large at the initial phase of token issuance, and the later transaction volume gradually decreases. The other two kinds of accounts with a higher proportion of in-transactions are Gambling and Phish. A gambling activity is a game where the dealer makes a profit, so the number of users who win money is less than the total number of participants. Phish accounts are fraudulent, so they deceive multiple users to obtain Ether, and then transfer out all the illegal gains at the final phase.

D. The amount of transactions between EGO and ALTER

Besides the number of transactions, the transaction amount is also an extremely important feature. Therefore, we observe the transfer amount of Ethereum accounts at each phase in the

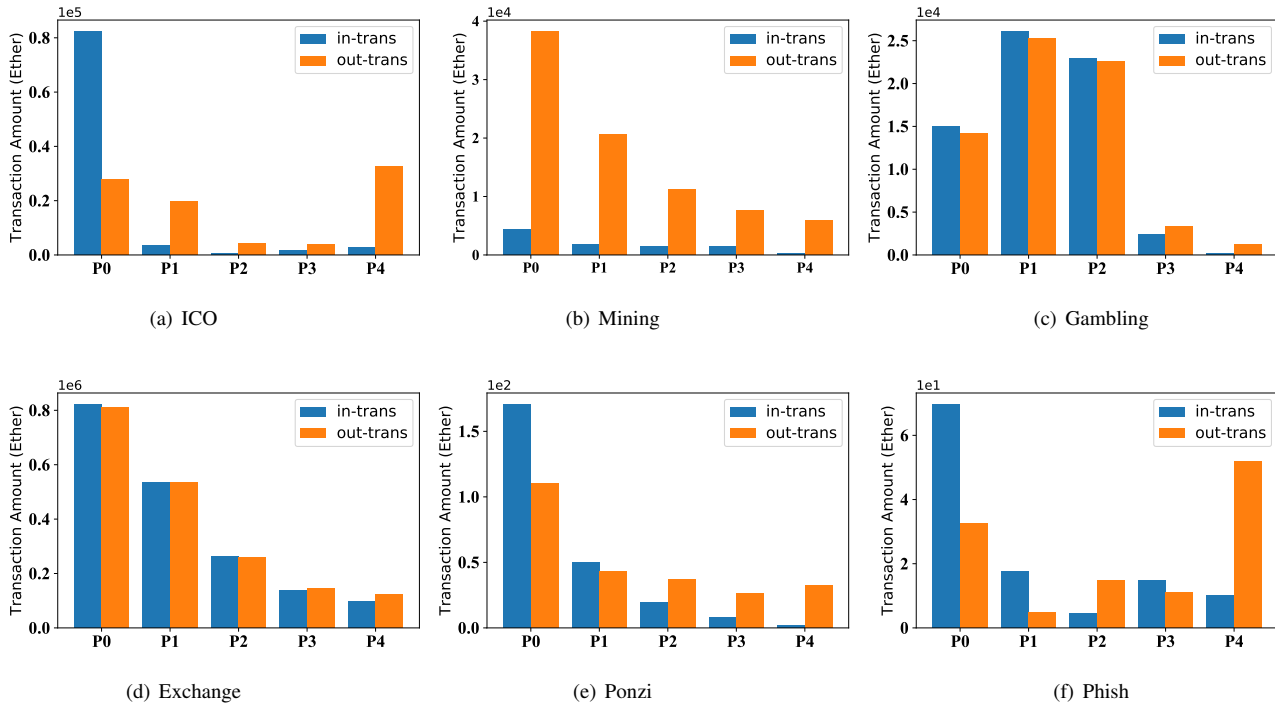


Fig. 3. Changes of the transaction amount at each phase

sliding window mode, and analyze the difference and essence. It can be seen from Fig. 3 that ICO accounts initially receive a large amount of investment and then gradually transfer it to the token-related accounts. Interestingly, the total amount transferred out from Mining accounts is always greater than the total amount transferred in, because most of the balance in the accounts comes from mining rewards. These rewards are not be recorded in the transactions, but they can be queried through the Ethereum browser. In addition, when no player participates in gambling games, Gambling accounts will transfer the balance out, indicating that Gambling accounts are essentially a profitable entity. In general, only the transfer-in and transfer-out amount at all phases of Exchange accounts are basically the same, which also reflects the role of Exchange accounts as an intermediary for the sale of tokens. Finally, we find that both types of illegal accounts have reaped huge profits, which means it's urgent to identify and combat fraudulent accounts on Ethereum. It is worth noting that Phish accounts usually transfer out a large amount of Ether in the final phase. While Ponzi accounts transfer Ether out in batches, and some Ether will be returned to investors.

IV. CONCLUSIONS AND FUTURE WORK

With the rapid development of Ethereum, a large number of accounts on Ethereum have involved in various business roles. For exploring the transaction behavior of these accounts and obtaining a more comprehensive description of their roles in the Ethereum ecosystem, we used transaction records to construct multiple ego networks with transaction amount and

time information for each type of account. By analyzing the dynamic changes of the interaction characteristics of EGO (labeled accounts) and ALTER (first-order neighbor accounts), we found that the trading of Mining accounts and Exchange accounts is more stable. In addition, the two illegal accounts, Ponzi and Phish, also have obvious abnormalities in the ratio of the amount of in/out transactions. The dynamic changes of transaction characteristics also enabled us to label unknown accounts. Most of all, we understood the evolutionary phases of scams, which can provide some help in identifying fraudulent activities on Ethereum.

In the future, we will continue to expand our research by investigating more network properties in the dynamic transaction network, such as transaction objects and network motifs. We expect the corresponding results to tell us whether there is a specific transaction rule and help us discover hidden communities in the network. Furthermore, we will consider more account labels, and more dynamic changes in complex network characteristics.

ACKNOWLEDGMENT

The work described in this paper is supported by the National Key R&D Program of China (2020YFB1006005), the National Natural Science Foundation of China (61973325) and the Natural Science Foundations of Guangdong Province (2021A1515011661)

REFERENCES

- [1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," <https://bitcoin.org/bitcoin.pdf>, 2008.

- [2] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," <https://ethereum.github.io/yellowpaper/paper.pdf>, 2014.
- [3] E. Mnif, A. Jarboui, and K. Mouakhar, "How the cryptocurrency market has performed during COVID 19? A multifractal analysis," *Finance Research Letters*, vol. 36, p. 101647, 2020.
- [4] F. Victor and B. K. Lüders, "Measuring Ethereum-based ERC20 token networks," in *Proceedings of the International Conference on Financial Cryptography and Data Security*. Frigate Bay, St. Kitts and Nevis: Springer, 2019, pp. 113–129.
- [5] J. Wu, J. Liu, W. Chen, H. Huang, Z. Zheng, and Y. Zhang, "Detecting mixing services via mining Bitcoin transaction network with hybrid motifs," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, to be published, doi: 10.1109/TSMC.2021.3049278.
- [6] Y. Huang, J. Wu, W. Ren, C. K. Tse, and Z. Zheng, "Sequential restorations of complex networks after cascading failures," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 400–411, 2021.
- [7] D. Lin, J. Wu, Q. Yuan, and Z. Zheng, "Modeling and understanding Ethereum transaction records via a complex network approach," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 11, pp. 2737–2741, 2020.
- [8] J. Wu, J. Liu, Y. Zhao, and Z. Zheng, "Analysis of cryptocurrency transactions from a network perspective: An overview," *Journal of Network and Computer Applications*, vol. 190, p. 103139, 2021.
- [9] J. Wu, Q. Yuan, D. Lin, W. You, W. Chen, C. Chen, and Z. Zheng, "Who are the phishers? Phishing scam detection on Ethereum via network embedding," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 2, pp. 1156–1166, 2022.
- [10] J. Geng, Y. Li, L. Fang, and P. Chen, "VW-DBG: A dynamically evolving Bitcoin transaction network model," *IEEE Transactions on Network Science and Engineering*, 2021, to be published, doi:10.1109/TNSE.2021.3116243.
- [11] D. Lin, J. Chen, J. Wu, and Z. Zheng, "Evolution of Ethereum transaction relationships: Toward understanding global driving factors from microscopic patterns," *IEEE Transactions on Computational Social Systems*, 2021, to be published, doi:10.1109/TCSS.2021.3093384.
- [12] L. Zhao, S. Sen Gupta, A. Khan, and R. Luo, "Temporal analysis of the entire Ethereum blockchain network," in *Proceedings of the Web Conference 2021*. Ljubljana, Slovenia: ACM, 2021, pp. 2258–2269.
- [13] F. Odella, "Using ego-network in surveys: Methodological and empirical research issues," in *Proceedings of the International Conference on Network Science*. Bloomington, Indiana: Springer, 2006.
- [14] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins, "Microscopic evolution of social networks," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Las Vegas, Nevada, USA: ACM, 2008, pp. 462–470.
- [15] V. Arnaboldi, M. Conti, A. Passarella, and F. Pezzoni, "Analysis of ego network structure in online social networks," in *Proceedings of the 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*. Amsterdam, Netherlands: IEEE, 2012, pp. 31–40.
- [16] Q. Yuan, B. Huang, J. Zhang, J. Wu, H. Zhang, and X. Zhang, "Detecting phishing scams on Ethereum based on transaction records," in *Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. Seville, Spain: IEEE, 2020, pp. 1–5.
- [17] W. Chen, X. Guo, Z. Chen, Z. Zheng, and Y. Lu, "Phishing scam detection on Ethereum: Towards financial security for Blockchain ecosystem," in *Proceedings of the International Joint Conference on Artificial Intelligence*. Yokohama, Japan: IJCAI, 2020, pp. 4506–4512.
- [18] Chainalysis, "The Chainalysis 2021 crypto crime report," <https://go.chainalysis.com/2021-Crypto-Crime-Report.html>, 2021.