# Poster: Evolution of Ethereum: A Temporal Graph Perspective

Qianlan Bai[†], Chao Zhang[+], Yuedong Xu[+], Xiaowei Chen[*], Xin Wang[†]
[+] School of Information Science and Engineering, Fudan University
[†] School of Computer Science and Technology, Fudan University
[*]School of Computer Science and Engineering, The Chinese University of Hong Kong

***Abstract*—Ethereum is one of the most popular blockchain systems that supports more than half a million transactions every day. Whereas it remains mysterious what the transaction pattern is and how it evolves over time. In this paper, we study the evolutionary behavior of Ethereum transactions from a temporal graph point of view. It shows that there is no evidence that changes in average triplet closure duration is related to prices. We observe the macroscopic and microscopic burstiness of Ethereum transactions. We analyze the Gini indexes of the transaction graphs and the user wealth in which Ethereum is found to be very unfair since the very beginning, in a sense, "the rich is already very rich".**

## I. Introduction

As the second largest blockchain platform, the market value of Ethereum has reached 1573 millions dollars on Jan 1, 2020. Around half a million transactions are conducted on the Ethereum platform everyday. The boom of Ethereum has aroused great interests in the understanding of its social interactions. Analyzing transaction data provides a crucial way to know the development of Ethereum. A commonly used approach is to construct transaction graphs. In the existing literature, Ethereum are usually analyzed as static network, such as Chen *et.al* [3] characterized the activity of money transfer, contracts creation and contracts invocation using different graphs. The evolutionary behavior is largely overlooked.

In this paper, we study some basic questions: *How does Ethereum transaction pattern evolve over time, and how is it affected by Ether price?* To this goal, we first collect the transaction data and then construct user-to-user graph(UUG), which characterizes the trading relationship among externally owned accounts(EOAs). In order to understand the evolution of Ethereum transactions, we chop UUG into a sequence of temporal graphs with sliding and incremental time window.

Our graph analysis on UUG is carried out from three perspectives. Firstly, we measure the local graph structure that evolve over time. Secondly, the macroscopic burstiness that captures the aggregation extent of transactions made by nodes, and the microscopic burstiness that characterizes the inter-transaction time distribution is quantified. Third, the Gini indexes of degree, transaction and balance (wealth) distributions are computed to verify "the rich gets richer" effect.

Our major observations are briefly summarized as below:

- 1. The global clustering coefficient of UUG is very small. The correlation between Ether price and the local graph structure such as the proportion of closed triplets and the average closure time is not observed.
- 2. For most of nodes, their transactions are concentrated on a short duration of their active periods. The inter-transaction time intervals varies considerably, exhibiting burstiness.
- 3. The distribution of degree, transaction and wealth of nodes are always unfair since the genesis of Ethereum.

## II. Datasets and Basic Definitions

In this section, we describe our dataset. The construction of transaction graphs is introduced in detail.

### A. Overview of Datasets

We collect all the Ethereum transactions spanning from July 30th, 2015, the birth date of Ethereum, to February 9th, 2019. The total number transactions is around ***19 millions***, and the size of our dataset is around 34 GB. Each transaction records the following items: $\{$*Block ID, Sender, Receiver, Transaction Value*$\}$. We also collect the balance of each address at all time.

### B. Construction of Transaction Graphs

We focus on the fundamental properties of Ethereum transactions and the compound interactions among users. So we construct user-to-user graph (**UUG**), which captures the transaction patterns among users, that is, the most important aspect of Ethereum as a cryptocurrency. In UUG, each Externally Owned Account (EOA) is a node and two nodes form a *directed* edge if a transfer of Ether between them happens. The basic statistics of UUG is expressed in Table I.

TABLE I: Sizes of UUG.

| type of nodes | #of nodes | #of edges | #of transactions |
|---|---|---|---|
| EOA | 41722479 | 89194399 | 192300454 |

Temporal graphs are constructed to explore the evolution of Ethereum system.Two types of time windows are considered, the sliding window and the incremental window. We observe that more than 70% of nodes have a lifetime (the interval between the first and last transaction time in our dataset) below 180 days, and thus setting the sliding window to be 180 days is suitable to evaluate the graph dynamics. The sliding window is shifted with a granularity of 45 days (i.e. $1/4$ of this time window) that is useful to compare the graphs at different stages. The incremental window expands from 180 days to 1260 days with the same granularity. Although the

This work was supported by the Shanghai-Hong Kong Collaborative Project under Grant 18510760900. Yuedong Xu is the corresponding author.
Annex to ISBN 978-3-903176-28-7© 2020 IFIP

time window is a hyper parameter, it is carefully selected and its choice does not influence the main conclusions.

## III. Motifs of UUG

Transaction patterns usually are the basic building blocks of the entire graph that reveal the microscopic behaviors on how the network is formed. In general, the subgraphs containing a few nodes and edges induced from the original graph, also named *motifs* [1], are of particular interests to the network science research community.

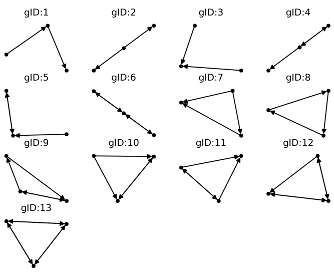


Fig. 1: 3-nodes motifs.

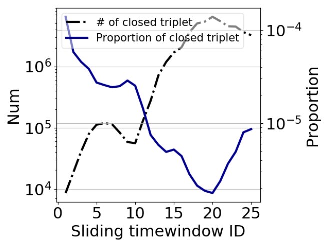


Fig. 2: Number and proportion of closed triplets in UUG.

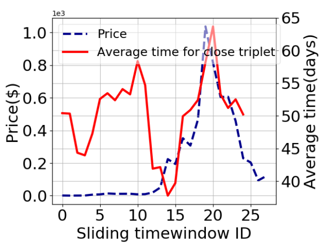


Fig. 3: Average time for each closed triplet.

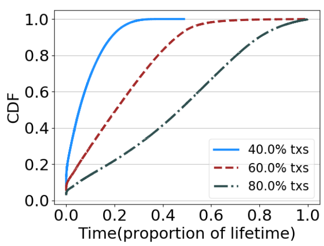


Fig. 4: Busy period ratio distribution.

We hereby investigate 3-node motifs of UUG. As a directed graph, there are thirteen motifs(Figure 1). We classify seven motifs as *closed triplet* that has transactions between each pair of nodes, and name *open triplet* for the remaining six motifs as a unity. We illustrate the closed triplets in Figure 2. In the very beginning, the number of closed triplets is around $10^4$, and it grows to around $4 \times 10^6$ in the last sliding window. Although this number decreases considerably between the seventh window and the tenth window, the overall trend is the expansion of hundreds of times. This manifests more and more EOAs are conducting transactions with others. The proportion of closed triplets among all the 3-node motifs is almost strictly decreasing from the beginning to the twentieth sliding window, which descends to nearly the order of millionth. The low ratio of closed triplet directly leads to a very low global clustering coefficient, up to $10^{-4}$. This means the transaction pattern of UUG is dominated by the mode which a vast majority of nodes merely interact with a small amount of nodes.

Another interesting question is how long time it will take to form a closed triplet. Figure 3 shows the average needed time for an open triplet to be closed. The average closure time varies at different sliding windows, i.e. ranging between 37 days and 64 days, while their differences are not prominent, and most of the closure time oscillates between 50 days and 55 days. We further plot the dynamics of Ether price in Figure 3. The correlation between the Ether price and the closure time is relatively weak. Especially, when the Ether price increases fifty to one hundred times, the closure time remains the same or experiences a relatively small increase.

**Observation 1:** The global cluster coefficient of Ethereum is very small. Although the number of closed triplet has increased, it is still negligible compared to the number of open triplet. The average closure time fluctuates at different time and there is no evidence that it affected by the price of Ether.

## IV. Burstiness of Transactions in UUG

We witnessed much progress in the study of Ethereum transactions, little is known about the dynamics characteristics. Burstiness is common temporal measure of the dynamics of complex systems. Here, we analyze the burstiness of Ethereum from both macroscopic and microscopic perspectives.

### A. Macroscopic Burstiness

We follow the next step to quantify the macroscopic burstiness. First, the lifetime of a node is defined as the time interval between the first and final transaction made by him. We next find the minimum time needed by each node to conduct a certain percentage of transactions. This time interval quantifies the bursty period for each node. Based on this, we define the least required time to complete a certain percentage of transactions divided by the lifetime of this node as *busy period ratio*. Figure 4 shows the results in three scenarios. One can observe that 80% of nodes complete 40% of the transactions with only 13.81% of their lifetime; 80% of nodes completed 60% of transactions with 38.02% of their lifetime, and 80% of nodes use 70.46% of their lifetime to complete 80% of transactions. Our measurement shows the clear bursty evidence of almost all the meaningful nodes as a whole.

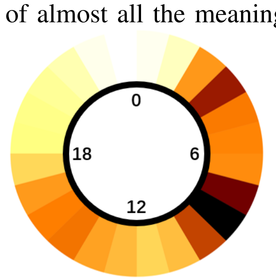


Fig. 5: Burstiness of transactions in UUG during one day.

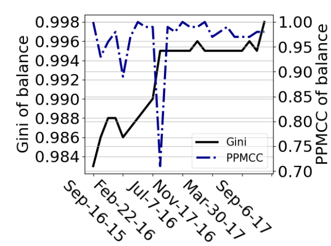


Fig. 6: Matthew effect of balance.

The burstiness of global transactions is measured by the average number of transactions in each hour on a daily basis (GMT is adopted). Figure 5 illustrates the intensity of bursty transactions made by EOAs. The most active period is 7*am* ∼10*am*, and the second one is 3*am*∼4*am*. The most active period corresponds to 15 *pm*∼18*pm* at Beijing Time (GMT+8), the afternoon working hours in Asian countries. Due to the escalating interests toward blockchains in these countries, we conjecture that this is the very reason accounting for the global burstiness in terms of the number of transactions in each day.

### B. Microscopic Burstiness

The microscopic burstiness refers to the significantly enhanced activity level over short periods of time followed by long periods of inactivity for a single node. Goh and Barabasi proposed to characterize the (microscopic) bursty of social events using orthogonal measures [4]. They quantify two distinct mechanisms causing burstiness: the inter-event time distribution and the *memory*. We adopt the same method to examine the burstiness and memory of Ethereum transactions.

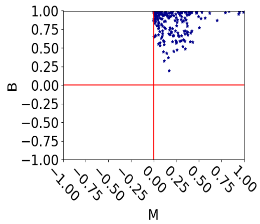


Fig. 7: MB of normal nodes.

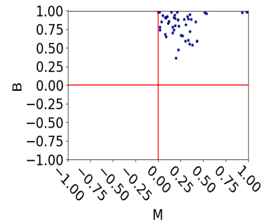


Fig. 8: MB of exchanges.

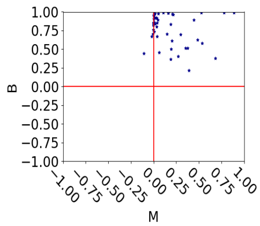


Fig. 9: MB of mining pools.

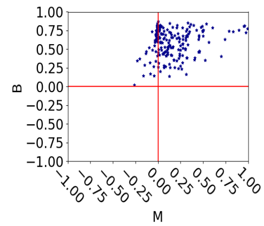


Fig. 10: MB of phish hack.

*B* is used to measure the inter-event time distribution and it is defined as Eq.1.

$$B \equiv \frac{\sigma - <\tau>}{\sigma + <\tau>} \tag{1}$$

$<\tau>$ and $\sigma$ represent the mean and standard deviation of inter-event times. Here *B* takes the value of -1 for regular time series and 0 for random.

*M* is memory coefficient to measure two-point correlations between consecutive inter-event times as the following:

$$M \equiv \frac{1}{n-2} \sum_{i=1}^{n-2} \frac{((\tau_i) - \langle\tau\rangle_1)(\tau_{i+1} - \langle\tau\rangle_2)}{(\sigma_1)(\sigma_2)} \tag{2}$$

The closer M is to 1 (-1), the greater the probability that the time interval after the long interval is long (short).

The same number of accounts are selected from normal accounts, exchanges, mining pools and fraud accounts respectively, and then we map the transaction of these accounts on the (*M*,*B*-space). Figure 8 and Figure 9 show the transactions of exchanges and mining pools are all busty, we conjecture this may be caused by the burstiness of transactions during on day. Compared with Figure 7, Figure 10 shows the memory of phishing accounts are distributed on a wider range.

**Observation 5:** From perspective of macroscopic burstiness, the transactions of most of nodes are concentrated on a small part of the lifetime of the nodes and the transactions during one day is distributed on some hours. From perspective of microscopic burstiness, the arrival time distribution of transactions of nodes is very highly devise.

## V. The Rich Gets Richer in UUG

The Matthew effect can be observed in many aspects of social and economic system [2]. It is sometimes known as the adage "the rich get richer". We measure it to know whether Ethereum will evolve into an extremely unhealthy economic system. In Ethereum network, we measure the Matthew effect with Gini coefficient, a commonly used measure for the income gap. The Gini coefficient is in the range $[0, 1]$, when it is above $0.5$, it means that income is too unbalanced.

From Figure 11 and Figure 12, we discover the Gini coefficient of degree and transaction number is above 0.5. It means the degree and the transaction number distribution is unbalance in all time periods. But there is no sign on the direction that the Gini coefficient is moving toward both in degree distribution and transaction number distribution. Hence, we can conclude that in terms of degree distribution and transaction number distribution, "Always unfair but not the rich gets richer." Meanwhile, we hope that Gini coefficient of nodes' balances can reveal the (un)fairness of wealth distribution in Ethereum. As the first step, we assume that a node is "dead" if his balance is 0, and he does not trade with any others afterwards. Then Figure 6 shows the Gini coefficient of balance increase from 0.982 to 0.998, it means Ethereum is "extremely unfair in terms of the balance since its birth".

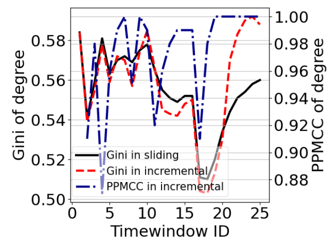


Fig. 11: Matthew effect of degree.

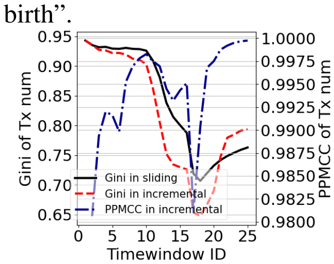


Fig. 12: Matthew effect of Tx num.

A subsequent question is "will the rich still be rich in the near future?". In Gini coefficient, there is no differentiation on the node identity. We then introduce the Pearson product-moment correlation coefficient (PPMCC) to explore this question. In statistics, PPMCC is a measure of linear correlation between two variables, and when it is above 0.6, there is a strong correlation among variable. We evaluate PPMCC of the degree, transaction number and balance of nodes in consecutive windows. Figure 11, Figure 12 and Figure 6 show that they are all above 0.6. This implies that PPMCC is very large between consecutive windows, which means that the nodes with rich degrees, rich transaction number and large health are still rich.

**Observation 6:** The distribution of degree and transaction number are always unfair but not the rich get richer. The distribution of wealth is extremely unfair since very beginning.

## VI. Conclusion

We conduct an evolution of Ethereum from the perspective of temporal graph analysis. By analyzing it through various metrics, we obtain many new observations and insights, which help people have a understanding of the evolution of Ethereum.

## References


[1] Alon U. "Network motifs: theory and experimental approaches". Nature Reviews Genetics, 2007.

[2] Arney, Chris. "Networks, Crowds, and Markets: Reasoning About a Highly Connected World". Mathematics and Computer Education, 2013.

[3] Chen, Ting, et al. "Understanding ethereum via graph analysis". INFOCOM 2018-IEEE Conference on Computer Communications. IEEE 2018.

[4] K.-I. Goh and A.-L. Barabási. "Burstiness and memory in complex systems". EPL (Europhys. Lett.), 2008.