# Understanding the Decentralization of DPoS: Perspectives From Data-Driven Analysis on EOSIO

**Preprint** · January 2022

CITATIONS
0

READS
61

**5 authors**, including:

Jieli Liu
Sun Yat-Sen University
**15** PUBLICATIONS **130** CITATIONS
SEE PROFILE

Weilin Zheng
Sun Yat-Sen University
**10** PUBLICATIONS **168** CITATIONS
SEE PROFILE

Jiajing Wu
Sun Yat-Sen University
**96** PUBLICATIONS **1,684** CITATIONS
SEE PROFILE

Zibin Zheng
Sun Yat-Sen University
**405** PUBLICATIONS **19,467** CITATIONS
SEE PROFILE

# Understanding the Decentralization of DPoS: Perspectives From Data-Driven Analysis on EOSIO

Jieli Liu, Weilin Zheng, Dingyuan Lu, Jiajing Wu, *Senior Member, IEEE,* and Zibin Zheng, *Senior Member, IEEE*

***Abstract*—Recently, many Delegated Proof-of-Stake (DPoS)-based blockchains have been widely used in decentralized applications, such as EOSIO, Tron, and Binance Smart Chain. Compared with traditional PoW-based blockchain systems, these systems achieve a higher transaction throughput and are well adapted to large-scale scenes in daily applications. Decentralization is a key element in blockchain networks. However, little is known about the evolution of decentralization in DPoS-based blockchain networks. In this paper, we conduct a systematic analysis on the decentralization of DPoS with data from up to 135,000,000 blocks in EOSIO, the first successful DPoS-based blockchain system. We characterize the decentralization evolution of the two phases in DPoS, namely block producer election and block production. Moreover, we study the voters with similar voting behaviors and propose methods to discover abnormal mutual voting behaviors in EOSIO. The analytical results show that our methods can effectively capture the decentralization evolution and abnormal voting phenomena in the system, which also have reference significance for other DPoS-based blockchains.**

***Index Terms*—Blockchain, Delegated Proof-of-Stake, decentralization, EOSIO, network analysis**

## I. Introduction

NOWADAYS, blockchain has aroused great attention among people and has been widely applied in many fields such as finance, healthcare, and Internet of Things (IoTs). Technically, blockchain [1] is an append-only distributed ledger database combined with hybrid techniques like peer-to-peer networks, cryptography, and consensus mechanisms. Different from traditional centralized systems, blockchain provides users a decentralized environment that can avoid the single point failure of a system. Thus decentralization is a core element in blockchain networks.

The most famous blockchain system is Bitcoin [2] and Ethereum [3], which is based on the Proof-of-Work (PoW) consensus protocol [4]. However, with the wide application of blockchain, both Bitcoin and Ethereum suffer from the low scalability problem and can not meet the growing application requirement. The Delegated Proof-of-Stake (DPoS) consensus protocol is an efficient and flexible blockchain consensus mechanism famous for its high scalability in block production. Unlike traditional PoW-based consensus protocols, DPoS concentrates the block production process in the hands of a small set of block producers (also called super nodes) elected by the entire network, and thus improves the transaction throughput. The original DPoS was proposed in 2014 [5], and it has given rise to many variant versions adopted in a series of successful blockchains like EOSIO [6], Tron [7] and Binance Smart Chain [8]. The framework of all these DPoS consensus protocols can be summarized into two phases:

1) Block producer election: This phase dynamically elects a limited set of block producers to produce blocks. The voting weights are different across the stakes of voters. A fixed number of top candidates receiving the highest voting weight can become block producers.
2) Block production: The elected block producers begin a round of block production orderly. Each elected block producer has a fixed time slot to produce blocks. If a block is not produced at the scheduled time, this block will be skipped and will not affect the subsequent block production. Finally, the block production rewards are distributed to these elected block producers and a new round of block producer election and block production will be started.

Out of the importance of blockchain security, recently many efforts have been devoted to analyzing the decentralization in different blockchain systems. Intuitively, if a blockchain system is not decentralized enough, it can be easily controlled by a small minority of parties. In this case, the blockchain system is fragile and cannot ensure the facticity of the ledger records when facing threats like 51% attack [9] and selfish mining [10]. Existing research [11], [12] studied the decentralization of Bitcoin and Ethereum in terms of mining power, bandwidth, etc. For DPoS-based blockchains, some studies [13], [14] have theoretically analyzed the decentralization from the protocols and pointed out several clear vulnerabilities. Some other studies [15], [16] measured the decentralization of blockchain systems with the entropy metric. This metric is useful in comparing the decentralization among different protocols according to the block production records of blockchain systems. However, there is no study quantifying the decentralization evolution in DPoS-based blockchains. And an in-depth study on the behaviors of voters and block producers can help us timely capture the abnormal phenomena such as voting manipulation in a DPoS-based blockchain system. Recent reports [13], [17] have pointed out that it is possible to achieve voting collusion in DPoS-based blockchains. Yet there is still a lack of study to capture the abnormal voting behaviors in DPoS-based blockchain systems.

Manuscript received January xx, 2022. The research is supported by xxx). (*Corresponding Author: Jiajing Wu*)

J. Liu and Z. Zheng are with the School of Software Engineering, Sun Yat-sen University, Zhuhai 519082, China. W. Zheng, D. Lu, and J. Wu are with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China. (Email: liujli7@mail2.sysu.edu.cn, wujiajing@mail.sysu.edu.cn)

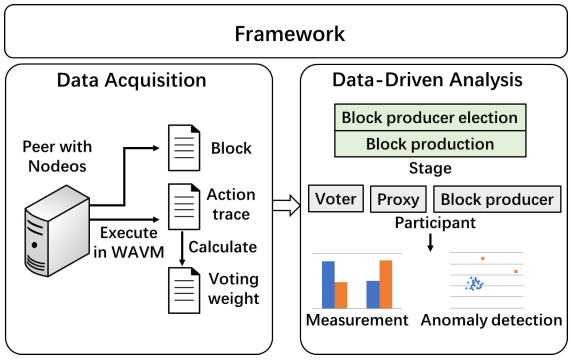



Fig. 1. The analysis framework.

To fill this gap, we conduct a data-driven analysis with the blockchain data from EOSIO. EOSIO [6] is a typical and the first successful DPoS-based blockchain system. Up to now, the number of transactions in EOSIO has exceeded 5.0 billion according to the statistic of *eosflare.io* [18]. Our analysis framework is shown in Fig. 1. Firstly, we collect the block data as well as the action trace data, and calculate the voting weight data according to the action trace data. Then in the analysis, we conduct a decentralization evolution measurement by characterizing all voters, voting proxies, and block producers participating in the DPoS consensus process. Following the analysis, we design network-based methods to uncover the abnormal voting behaviors in the system. In particular, we investigate the abnormal voting gangs with similar voting behaviors and mutual voting behaviors. The analytical results show that our study can reveal the decentralization evolution and effectively capture some abnormal voting activities in a DPoS-based blockchain.

In summary, we make the following contributions in this paper:

- To the best of our knowledge, our work is the first data-driven study on the decentralization evolution of DPoS. Based on the data from a typical DPoS-based system named EOSIO, we quantify its decentralization and capture the abnormal voting behaviors in the system.
- We report some voting gang anomalies with our methods, including gangs with similar voting behaviors and mutual voting behaviors. Moreover, we also analyze the phenomena behind the detected results.
- We find that the EOSIO proxies account for an increasingly large proportion of the total voting power, and the set of block producers changes less and less. Besides, the overall decentralization in EOSIO is also affected by these factors. Our findings and discussion can provide a reference for other DPoS-based blockchains.

The remaining sections of this paper are organized as follows. Section II and Section III first provide the background and detail the data collection. Section IV and Section V next introduce the decentralization measurement study and analysis on the abnormal voting phenomena. Section VI discusses some implications from our findings for DPoS-based blockchains and Section VII provides the related work. Finally, we conclude the paper and discuss the future work in Section VIII.

## II. Background

In this section, we provide the background required to understand the operation mechanism of EOSIO and the DPoS consensus in EOSIO.

### A. Background of EOSIO

EOSIO is a blockchain system built for providing an operating environment for a large scale of commercial DApps, being regarded as the foundation of blockchain 3.0 [13]. Different from the most famous blockchain platforms like Bitcoin and Ethereum, EOSIO adopts DPoS as its consensus mechanism, which allows it to achieve a far higher transaction throughput. Recent years have seen a growth in the popularity of EOSIO in DApp transactions since its high performance and a waiver of transaction fees.

Unlike Ethereum, the identity of an account in EOSIO is a human-defined string with up to 12 characters in length. Each account is created by an existing account via the *newaccount* interface and can deploy a contract on itself via the *setcode* interface of the system account *eosio*. The main currency circulating in EOSIO is EOS tokens, and resources of smart contracts such as CPU and RAM can be obtained via EOS token mortgage. A transaction in EOSIO contains multiple actions, and each action corresponds to an invocation of a contract. According to the initiator in contract invocation, these actions can be divided into calling actions and inline actions, for which are called by users and triggered by contracts respectively [13]. Besides, EOSIO provides a flexible role-based permission management allowing users to delegate their permissions to achieve a high-level control on other users.

### B. DPoS and DPoS in EOSIO

The DPoS consensus has become increasingly popular and widely adopted in several successful blockchain systems such as EOSIO, Tron and Binance Smart Chain (BSC). Instead of solving the PoW puzzles, DPoS decides its block producers according to the votes of the entire stakeholders, thus achieving high scalability in block production. The DPoS consensus consists of the block producer election phase and the block production phase, and the details of these in EOSIO are provided as follows:

*1) Phase 1: Block Producer Election:* This phase elects 21 block producers for block production. To become a block producer candidate, anyone in EOSIO who possesses enough hardware resources and the full ledger data can register via the *regproducer* interface of *eosio*. And the top 21 candidates with the highest voting weight can become block producers and obtain block production rewards.

Anyone in EOSIO can vote for the block producer candidates in two ways. The first way is voting directly through setting the *producers* parameter of the *voteproducer* interface as the list of at most 30 selected candidates. The second way is voting through a proxy by setting the *proxy* parameter of the *voteproducer* interface as the chosen proxy, and then the proxy can vote for block producer candidates on behalf of all proxied users. A user can register to become a voting proxy

by setting the *isproxy* parameter of the *regproxy* interface as 1, otherwise, a user can cancel the registration by setting the *isproxy* parameter as 0.

The voting weight of a voter is related to the voting time and the staked tokens for CPU and bandwidth, which can be calculated as:

$$weight = 10000 * stake * 2^{index}, \quad (1)$$

where $stake$ is equal to the amount of delegated bandwidth and computation, and $index$ can grow incrementally every week if voters update their voting, calculated as: $index = \frac{1}{52} * \lfloor \frac{t_{vote}-t_{init}}{7*t_{day}} \rfloor$. In this formula, $t_{vote}$ represents the Unix timestamp when a user performs the *voteproducer* operation, $t_{init}$ represents the Unix timestamp of Jan. 1, 2000, and $t_{day}$ denotes the number of seconds per day. As we can see, with the same amount of stakes, the voting weight of recent votings is higher than that of earlier votings. This design can encourage voters to keep their votes updated. Besides, once a user updates the number of stakes via *delegatebw* (increase mortgage) or *undelegatebw* (reduce mortgage), the voting weight will be automatically updated. For a voting proxy, its voting weight will be the sum of voting weight owned by all proxied voters. If a voter votes for multiple candidates, each chosen candidate can receive an equal voting weight according to the stakes and voting time of the voter.

*2) Phase 2: Block Production:* In this phase, the elected 21 block producers from the block producer election begin a round of block production on behalf of the stakeholders in EOSIO. They validate the transactions, construct the valid transactions into blocks and then produce blocks orderly. Each elected block producer has a fixed 3-second to produce 6 blocks. Therefore, a round of block production lasts 63 seconds. If a block is not produced at the scheduled time because of network delay or other reasons, this block will be skipped and will not affect the subsequent block production. The new block confirmation process is also conducted among the 21 block producers. Once a new block is confirmed by more than 2/3 of the block producers (i.e. at least $21 \times 2/3 + 1 = 15$ block producers) through signed messages, this block will be appended to the blockchain. With DPoS, EOSIO can generate a new block every 0.5 seconds averagely with a throughput of up to 8,000 TPS [19], which achieves higher performance than many other blockchain systems.

## III. Data Collection

We conduct our experiments on up to 135,000,000 blocks in EOSIO, which cover the transaction data from Jun. 8, 2018 to Aug. 5, 2020. To measure the decentralization of EOSIO, we collect three types of blockchain data, including block header, transaction actions, and the voting weight data by running an EOSIO full node and replaying all transactions.

- **Block header:** The block header in each block provides a summary for the entire block, which includes information such as the block producer, the block timestamp, the transaction Merkle root, etc. We obtain the block header data by starting a core service daemon of EOSIO named *Nodeos* [20] to synchronize data on the mainnet.
- **Transaction actions:** A transaction in EOSIO contains a set of actions, and each action is an invocation of a smart contract. However, only the calling actions are recorded in the blockchain, which are operations performed by users. While the inline actions, which are triggered by calling actions, are not recorded in the on-chain data and can be obtained only by replaying all transactions. To obtain the full transaction action data, we make use of the action trace data, which are the detailed run-time data of smart contract invocation generated by the Web Assembly Virtual Machine (WAVM). With *history_file_plugin* [21], we collect all action traces in JSON format. Then we extract the related actions such as *delegatebw*, *undelegatebw*, *regproxy*, *voteproducer*, etc from the action traces.
- **Voting weight data:** The voting weight received by each block producer candidate is in constant change since it is affected by the change of voters and their stakes. Thus we calculate the voting weight data of each candidate by traversing the transaction actions and record the data in periodically.

## IV. Decentralization Measurement Study

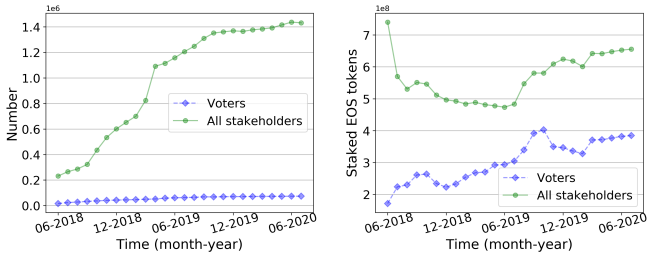
In this section, we present a decentralization measurement study in EOSIO considering the processes of block producer election and block production. We try to characterize all voters, proxies, and block producers participating in the DPOS consensus process. Based on the analysis results, we reveal the evolution of decentralization and discuss some findings in the EOSIO network.

### A. Block Producer Election in EOSIO

*1) Overview of Block Producer Election:* Based on the statistics of the collected dataset, we find out that there are totally 2,009,168 accounts. Among these accounts, 1,739,839 accounts have staked EOS tokens as stakeholders while only 84,668 accounts have participated in the block producer election as voters, occupying 4.87% of all stakeholders. The result indicates that only a small fraction of all stakeholders have taken part in the voting process.

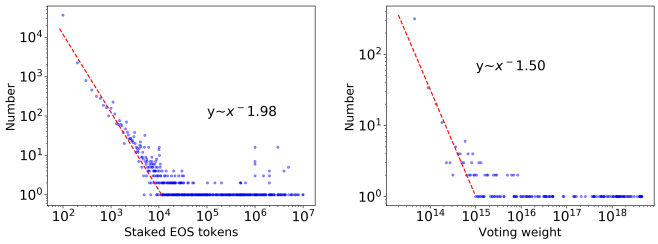
We then calculate the number evolution of accounts for voters and all stakeholders, as shown in Fig. 2(a). We found that the number of stakeholders and voters increase over time, but the number growth rate of voters is far short of the number growth rate of all stakeholders. This infers that many stakeholders do not care about the block production election, or even do not understand the voting mechanism in EOSIO. The power to decide who will become block producers thereby is held by a small number of stakeholders.

Besides, from the amount evolution of staked tokens held by voters and all accounts shown in Fig. 2(b), we observe that the staked tokens held by voters account for half or less of the total network for a long time. Moreover, with the passage of time, the proportion of staked tokens held by voters increases. The waiver of trading fees in EOSIO provides a user-friendly environment, and thus DApp users will have less incentive to mortgage their EOS tokens. However, for application developers, voters, and block producer candidates,




(a) Number evolution (b) Staked token evolution

Fig. 2. The number evolution and staked token evolution of voters and all stakeholders.




(a) Staked token distribution of voters (b) Received voting weight distribution of block producer candidates

Fig. 3. The staked token distribution of voters (for proxies the tokens of proxied accounts are accumulated) and the received voting weight distribution of block producer candidates, calculated at the end of the 135,000,000 blocks.

they have a strong motivation to mortgage their EOS tokens. Especially, block producer candidates always canvass other stakeholders for electoral support to ensure their position, and thus the amount of staked tokens of voters exhibits an increasing trend.

*2) Distribution Statistics:* We investigate the staked token distribution among all voters and the received voting weight distribution among all block producer candidates. The distributions of the mapping values with a fitting line $y \sim x^{-\alpha}$ are shown in Fig. 3(a) and Fig. 3(b). Both of the two distributions are in line with the power-law distribution, which indicates that a small number of voters hold a large amount of staked tokens, and a few block producer candidates have received a large voting weight. We then calculate the fraction of stakes held by the richest voters, and find out that the top 5% of the voters possess more than 95% of the stakes. As for block producer candidates, the top 10% of the candidates have received more than 95% of the voting weight in the whole network. Hence, the voting resource is unevenly distributed in both voters and block producer candidates. As one of the security risks, most of the voting powers are concentrated on only a few rich voters so that the voting results basically depend on a small number of voters. On the other hand, the extremely uneven distribution of received voting weight among block producer candidates can lead to the rich-get-richer phenomenon, also named as the *Matthew effect* [22]. In the long run, it is easy to form block production monopoly among the most popular block producers.

*3) Statistics of Voting Proxies:* Although anyone who stakes EOS tokens can participate in the governance of EOSIO with their voting power, block producer candidates should be carefully chosen by considering some factors such as resources, community contribution, and location of block producer candidates. However, not every stakeholder has enough incentive to investigate different candidates and stays abreast of every new proposal in EOSIO, especially those who only hold a tiny amount of EOS tokens.

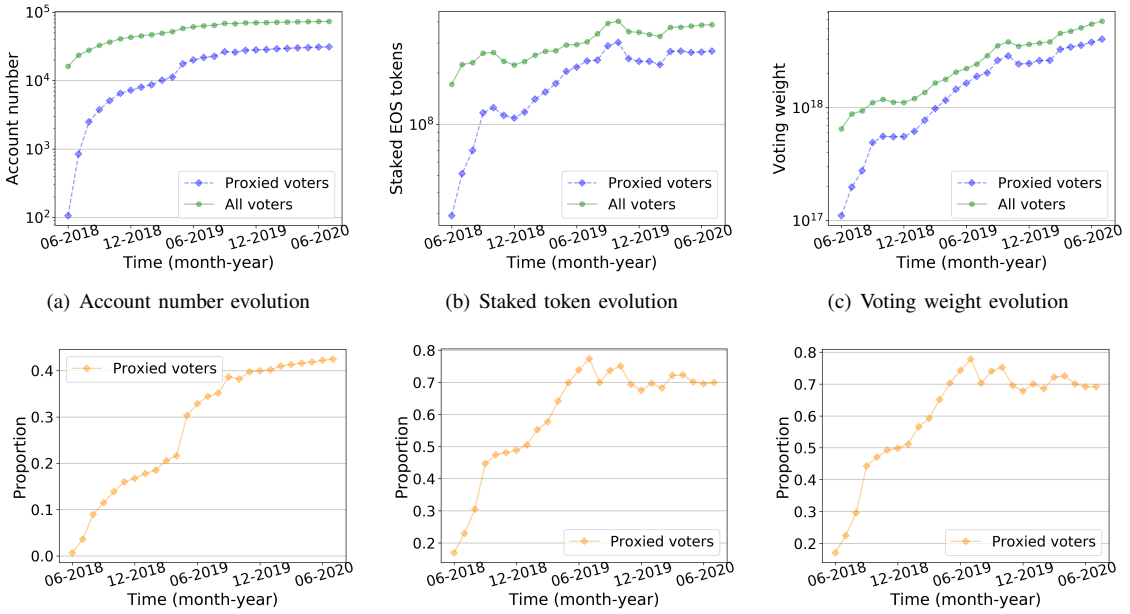
Voting proxies are accounts that perform voting on behalf of other accounts. They vote for block producer candidates with voting power collected from proxied accounts, and communicate to the proxied accounts which block producer candidate they choose. Each user can delegate its voting power to a voting proxy, and can also withdraw its voting power from the proxy at any time. In general, voting proxies can remove the high barrier of participating in governance for users and fully arouse the initiative of individual users, thereby increasing user participation in the voting process.

According to our statistics, 1,792 accounts have registered to become voting proxies, and 924 of them have been delegated as proxies via the *voteproducer* interface. It is worth noting that 40,561 of the 84,668 voters have voted through voting proxies, which means more than 47% of the voters execute their voting power via voting proxies. Fig. 4(a)-4(c) show the account number evolution, staked token evolution, and voting weight evolution of all voters and proxied voters. By cooperating with the corresponding value share evolution of the proxied voters in all voters in Fig. 4(d)-4(f), the increasing use of proxies can be observed. Moreover, after May 2019, nearly 70% of the voting weights are delegated to voting proxies. Under this trend, the votes of voting proxies can exercise considerable influence over the election results, since most of the voting powers are centralized in the voting proxies. A valuable research issue is whether there exist abuses of rights among the voting proxies, and we will discuss this issue in Section V.

**Finding 1:** The number of stakeholders increases rapidly with the boom of the EOSIO DApp ecology, but only a small number of them participated in the voting. The voting power is concentrated among a small number of voters.
**Finding 2:** The received voting weights are unevenly distributed among the block producers, making it possible to form a block production monopoly in the future.
**Finding 3:** Proxies account for an increasingly large proportion of the total voting power. Though the use of proxies can arouse more stakeholders to participate in the consensus government, it also centralizes most of the voting power in a small part of the accounts.

## B. Block Production in EOSIO

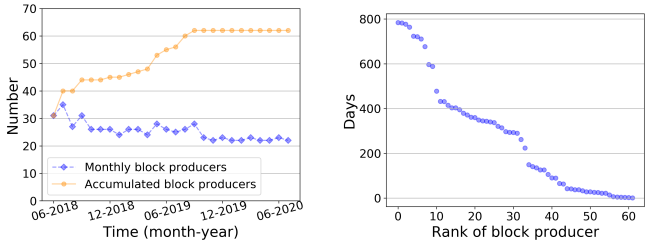
*1) Overview of Block Production:* When focusing on block production, we found that there are 615 accounts registering to become block producer candidates. Among these candidates, 602 candidates have received voting weights from the voters




(a) Account number evolution (b) Staked token evolution (c) Voting weight evolution

(d) Proportion evolution in account number (e) Proportion evolution in staked tokens (f) Proportion evolution in voting weights

Fig. 4. Statistics about the proxied voters. (a)-(c) show the account number evolution, staked token evolution, and voting weight evolution of all voters and proxied voters. (d)-(f) display the value share evolution of the proxied voters in all voters for account number, staked tokens, and voting weights.




(a) Number evolution of block producers

(b) Distribution of block producers' accumulative days participating in block production, calculated at the end of the 135,000,000 blocks

Fig. 5. Statistics about block producers, including (a) the number evolution and (b) distribution of days participating in block production.

and 62 candidates have become block producers. Fig. 5(a) shows the number evolution of block producers, and both the monthly number and accumulated number are given. Although the 21 elected block producers can be updated in each round according to their real-time voting weight, only around 25 different accounts are elected as block producers each month. Moreover, the chosen block producers are relatively fixed after September 2019.

Fig. 5(b) shows the distribution of days that block producers participating in block production, from which we can observe that about 10 block producers participating in block production for more than 600 days during the 135,000,000 blocks which cover about 790 days. These statistics illustrate that despite the large number of candidates, the set of block producers has small variations. Many accounts participate in block production for a long term. Once these accounts are in collusion, it is easy to achieve double-spend attacks since EOSIO can only tolerate no more than 1/3 of malicious block producers. In theory, the block producers acting in malicious manners can be voted out. Yet in practice, block producers can earn considerable rewards in block production so that most of them possess a large number of EOS tokens and can become one of the largest stakeholders, making it even harder to vote the malicious block producers out.

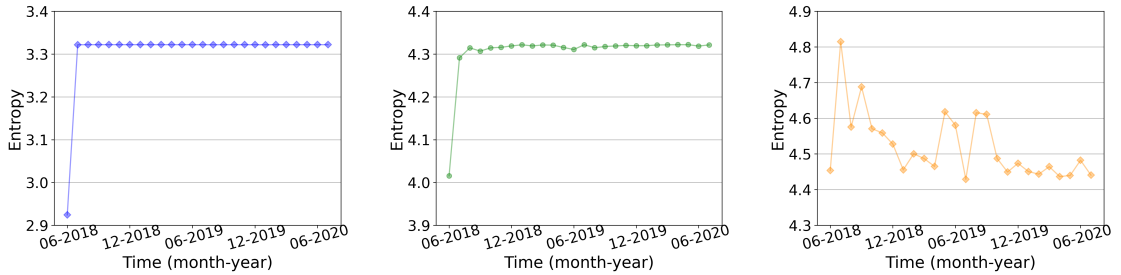
*2) Quantitative Analysis with Information Entropy:* Previous studies have introduced using the information entropy metric to quantify the degree of decentralization in block production [23], [16]. Here we compare the decentralization degree of different months with information entropy. The information entropy metric for a month $i$, namely $H(X_i)$, can be calculated as follows:

$$p_j = \frac{x_j}{\sum_{j=1}^{n} x_j}, \text{where } x_j \in X_i, \tag{2}$$

$$H(X_i) = -\sum_{j=1}^{n} p_j log_2 p_j, \tag{3}$$

where $X_i$ stores the number of blocks produced by each producer in month $i$, $x_j$ denotes the number of blocks produced by block producer $j$, and $n$ denotes the measuring range of top-$n$ block producers. The greater the value of $H(X_i)$ is, the greater the decentralization degree is in the month, since the block production is more random and disorderly.

We display the quantitative results of information entropy in Fig. 6 by setting $n = 10, 20$, and the number of block producers in each month. From Fig. 6(a) and Fig. 6(b), we can observe that the information entropy of the first month is much smaller than that of the latter months for $n = 10$ and $n = 20$, which indicates that block production in the early days of EOSIO was more centralized among the top stakeholders. While as shown in Fig. 6(c), block production in




(a) Entropy with $n = 10$ (b) Entropy with $n = 20$ (c) Entropy with $n$ = the number of all block producers

Fig. 6. Evolution of the information entropy.

months Oct. 2018-Apr. 2019, Jul. 2019, Oct. 2019-Jul. 2020 were also relatively centralized in general. Combining with Fig. 5(a), we can infer that this phenomenon may be due to the less replacement of block producers in these months.

**Finding 4:** Despite the relatively large amount of candidates, the set of block producers has small variations in general.
**Finding 5:** The block production was more centralized among the top stakeholders in the early days of EOSIO, and then it was more centralized in general for the later months due to the small variations of the block producer set.

## V. Exploring the Voting Gang Anomalies

In the previous section, the voting election and block production in EOSIO are characterized statistically. The obtained observations have exposed some problems that may give rise to power centralization such as large stakeholders working together, proxies abusing their power, and block producers colluding. In this section, we identify and analyze the voting gang anomalies during the block producer election in EOSIO. Firstly, we develop a voter clustering method to detect those who potentially work together and share their voting targets. After that, we discuss the mutual voting behaviors in EOSIO and propose an anomaly detection method to identify suspicious voting gangs.

### A. Voter Clustering Analysis

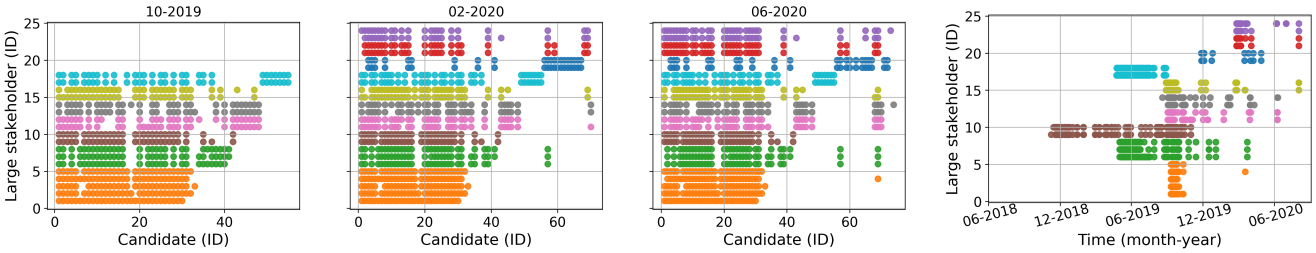
According to the statistics given in Section IV-B, there are totally 615 accounts registering as block producer candidates. Since each voter can vote for up to 30 block producer candidates and the votes can be updated in each round, the probability that two voters often vote for two groups of similar candidates is relatively low. Therefore, voters who often support similar block producer candidates are likely motivated by common interests. A great concern in EOSIO is that some large stakeholders may form alliances and communicate with each other before voting to solidify the positions of specific block producer candidates [13], causing voting manipulation and unfair competition. Hence, we propose a clustering method based on the voting similarity of voters and analyze the clustering results.

**Algorithm 1** Voter Clustering Based on Similar Voting Behaviors

**Input:** The set of voters $V$, voting records of voters $R$, selected threshold $\theta$.
**Output:** A set of voter clusters with similar voting behaviors $C$.

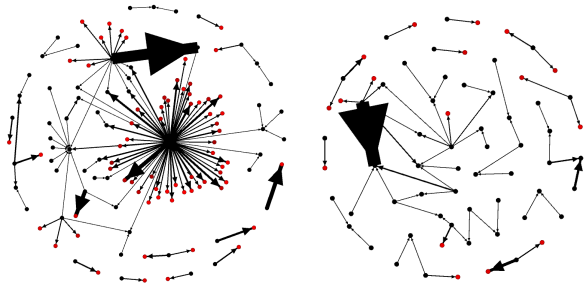
$visited = \emptyset$ //visited marker
$C = \emptyset$ //cluster set
**for** voter $v_i$ in $V - visited$ **do**
  add $v_i$ to $visited$
  $C_i = \{v_i\}$
  $N_\theta = \{v \in V, v! = v_i | \text{Similarity}(R(v), R(v_i)) \geq \theta\}$
  **for** voter $v_j$ in $N_\theta$ **do**
    **if** $v_j$ not in $visited$ **then**
      add $v_j$ to $visited$
      add $v_j$ to $C_i$
      **for** voter $v_k \in \{v \in V, v! = v_j | Similarity(R(v), R(v_j)) \geq \theta\}$ **do**
        add $v_k$ to $N_\theta$
  **if** len($C_i$) != 1 **then**
    add $C_i$ to $C$

*1) Method Design:* To group the voters which have similar voting behaviors into a cluster, our method considers the similarity of the dynamic voting records of each voter. The approach contains two stages. Firstly, we collect the voting status of each voter every sample time, and represent the candidates voted by the voter into a set with a maximum length of 30. In the experiment, we record the voting status at the end of each month and thus we obtain 26 sets as the sampled voting records during Jun. 8, 2018 to Aug. 5, 2020 for each voter, which can be used in measuring the similarity of voters. Secondly, we propose a similarity-based voter clustering method which can automatically output the clusters with similar voting behaviors. Algorithm 1 shows the pseudocode of the method, where the input contains the set of voters $V$, the sampled voting records $R$, and a similarity threshold $\theta$. The set $visited$ in the algorithm is used to mark the visited voters, and we treat each unvisited voter as a center. For each center, we select out voters whose voting similarity




(a) Three voting snapshots (b) Voting time distribution

Fig. 7. Three voting snapshots and the voting time distribution of the top 10 detected clusters with the most staked tokens among the large stakeholders, where votes in the same cluster are assigned the same color.



(a) Large stakeholders (b) Proxies

Fig. 8. Relationships between account creators and the detected clusters, where clusters owning one account creator are colored red.

is greater than or equal to $\theta$, add them to $visited$ and the cluster that the center belongs to. And then, if a center $v_i$ has members behaving similarly in its cluster $C_i$, voters who are similar to these members are also added into $C_i$. Finally, the cluster $C_i$ is added to the output cluster set $C$. Technically speaking, the proposed method detects voter clusters based on the similarity of the dynamic voting records, and the similarity measurement can be Jaccard's coefficient, Pearson correlation coefficient, etc.

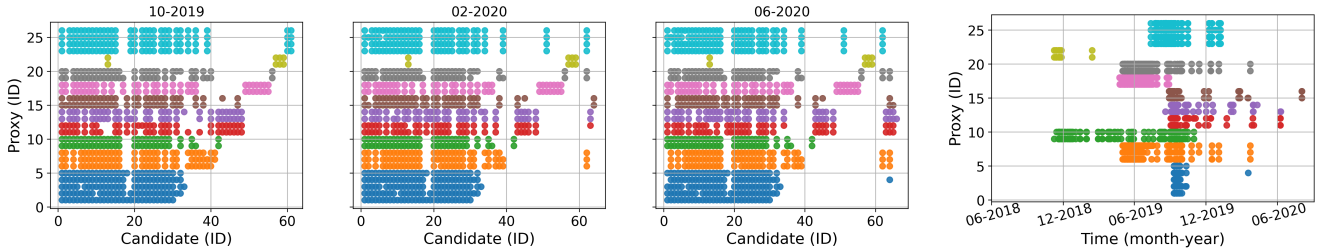
*2) Detecting in Large Stakeholders:* As mentioned in Section IV-A, the staked token distribution among all voters is extremely uneven. About 5% of the voters possess more than 95% of the stakes in the statistics. Hence the votes of large stakeholders make a great effect on the governance of EOSIO. To investigate the large stakeholders with similar voting behaviors, we apply Algorithm 1 on the top 5% richest voters (the proxied stakes are accumulated for proxies when ranking) in the sampled voting records, consisting of 2,723 voters. With Jaccard's coefficient as the similarity measure and $\theta = 0.9$, we detect 88 clusters including 293 voters.

Fig. 7(a) displays the votes of the large stakeholders in the top 10 detected clusters with the most staked tokens, where the y-axis depicts the index of the detected large stakeholders and the x-axis depicts the index of the candidates. If a voter $y$ votes for block producer candidates $x_1, x_2$ and $x_3$, then ($x_1$, $y$), ($x_2$, $y$) and ($x_3$, $y$) will be colored in the corresponding color of its cluster. From this figure, we can observe that voters in the same detected cluster have similar voting behaviors in both time and space, which can be illustrated by the very similar block producer candidates they support in these different snapshots. Fig. 7(b) shows the distribution of time when the large stakeholders in the top 10 detected clusters with the most staked tokens operate the *voteproducer* action, where different clusters are assigned different colors. To our surprise, though the voters in the same cluster are grouped according to the sampled voting records, the voting actions of these voters occur at very similar timestamp sequences.

We further analyze from the account creators of the detected voters, and visualize the relationships between account creators and the 88 clusters in Fig. 8(a). In each edge of the figure, a target node represents a detected cluster of large stakeholders, a source node represents an account creator that creates one or more stakeholders in a cluster, and the edge thickness is related to the number of accounts created. From the figure, we can observe that most accounts in the same cluster are created by the same account, implying that these voters are likely to be controlled by the same entity and therefore they have similar voting behaviors. Such kind of gathering phenomena in voting among large stakeholders can easily cause block production monopoly—the allies of these large stakeholders can obtain a solid position in the election. Moreover, they will earn more rewards (i.e, EOS tokens) from block production to consolidate their advantage in the election.

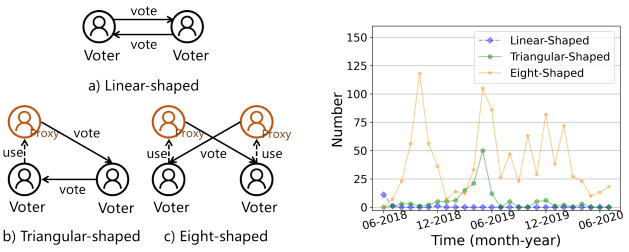
*3) Detecting in Voting Proxies:* Voting proxies are accounts that can execute the voting power of proxied accounts. Recently, we have seen the rising use of proxies in voting. Usually, users delegate their voting rights to a proxy out of trust in the proxy's voting decision. However, we can not exclude the presence of "selfish" proxies that only vote for their alliance members but do not consider the governance of the community. In this way, if users delegate their voting power to these proxies, they will actually increase the competitiveness of a particular alliance and the degree of centralization. Therefore, there is an urgent need to understand the voting behaviors of voting proxies. Particularly, proxies having similar voting behaviors deserve our attention. And these proxies are suspected of being manipulated by the same entity and using different identities to attract the authorization of other stakeholders.

To this end, we apply Algorithm 1 on all proxies in our dataset with Jaccard's coefficient as the similarity measure and $\theta = 0.9$. We totally detect 35 clusters including 162 proxies. Fig. 9(a) and Fig. 9(b) visualize three voting snapshots and the voting time distribution of proxies in the top 10 detected




(a) Three voting snapshots (b) Voting time distribution

Fig. 9. Three voting snapshots and the voting time distribution of the top 10 detected clusters with the most staked tokens among the proxies, where votes in the same cluster are assigned the same color.




(a) Mutual voting patterns (b) Number evolution

Fig. 10. Three mutual voting patterns and their number evolution.

clusters with the most staked tokens, where different clusters are assigned different colors. Similar to the observations drawn from Fig. 7(a) and Fig. 7(b), the proxies in the same detected clusters vote for almost the same group of block producer candidates across time, and the time they operate the *voteproducer* action shows a convergent trait.

Fig. 8(b) displays the account creation relationships between account creators and the 35 detected clusters, and a target node denotes a cluster of voting proxies and a source node denotes an account creator that creates one or more proxies in a cluster. As we can see, among 17 of the 35 clusters, proxies in the same cluster are created by one account. Besides, we observe that many proxies in the same cluster have common features in their account name. For example, proxies *'hashfineos44'*, *'hashfineos14'*, *'hashfineos33'*, *'hashfineos55'* and *'hashfineos13'* have the common prefix 'hashfineos-' in their name, proxies *'reallyreally'*, *'windowwindow'*, *'citycitycity'*, *'familyfamily'*, etc. have reduplicated words in their name. We also observe that the account *'octgenerator'* creates 53 proxies detected in the same cluster, and all these proxies only vote for *'oraclegogogo'*. Though it is impossible to verify that the detected accounts in the same cluster belong to an entity due to the pseudonymous nature, our method helps reveal the abnormal voters who potentially work together.

### B. Mutual Voting Anomaly Detection

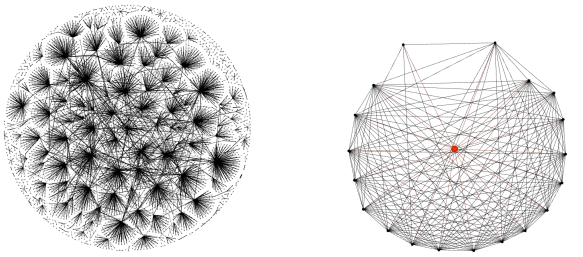
In this part, we aim to investigate the mutual voting anomalies, which are suspected of trading votes and sharing the rewards of block production with gang members. We first propose and discuss three mutual voting patterns, and then we develop an algorithm to explore the potential mutual voting anomalies.

*1) Mutual Voting Pattern Analysis:* As shown in Fig. 10(a), we consider three types of mutual voting patterns in the analysis: a) *linear-shaped* pattern, in which two voters vote for each other directly to enhance their roles; b) *triangular-shaped* pattern, where a voter $a$ use a proxy to vote for another voter $b$ and receive the vote back from the voter $b$; and c) *eight-shaped* pattern, in which two proxies and two proxied voters are involved and the two proxied voters vote for each other via proxies. Among these three kinds of mutual voting patterns, the linear-shaped pattern and the triangular-shaped pattern are relatively abnormal voting behaviors since there exist votes from non-proxy entities.

The number evolution of these patterns in different months is displayed in Fig. 10(b), and the occurring time of the mutual voting actions in each pattern are constrained within seven days. We observe that the most straightforward mutual voting pattern, namely the linear-shaped pattern, occurred mainly between some famous super nodes such as eosnationftw, eoslaomaocom, and argentinaeos in June 2018. For the triangular-shaped pattern, its occurrence number reached a peak in May 2019. By analyzing the voting records, we find that the occurrence of this peak is related to the obvious mutual voting behaviors between games.eos, eos.fish, starteosiobp, and dilidilifans. Firstly, the votes between starteosiobp and eos.fish, games.eos and eos.fish, dilidilifans and eos.fish are in triangular-shaped patterns with proxy start13.io, proxyfordili, and starteos.io respectively. Secondly, dilidilifans and games.eos, starteosiobp and dilidilifans, games.eos and starteosiobp are in eight-shaped patterns via their proxy, which provides further proof for the close relationship among these voters.

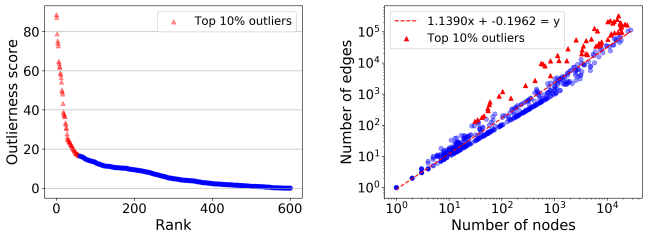
As for the eight-shaped pattern, it occurs more frequently than the linear-shaped pattern and triangular-shaped pattern because voting proxies are active. We then investigate the occurrence number peak of the eight-shaped pattern. The maximum peak reached in October 2018, mainly caused by the frequently mutual voting of acroeos12345 and eoseouldotio with a common proxy votetochange. This mutual voting relationship is natural since acroeos12345 and eoseouldotio are both the organizers of the proxy votetochange[1]. For the second peak reached in May 2019, we observe that there exists a large connected component formed by the mutual voting relationships of proxied voters. And among the 17 proxied voters in

[1]https://www.alohaeos.com/vote/proxy/votetochange



(a) The voting network (b) An example of near-clique anomaly

Fig. 11. Visualization of the voting network (a) and near-clique anomaly (b).




(a) Outlierness score distribution (b) The Egonet Density Power Law

Fig. 12. Result of the near-clique anomaly detection. (a) shows the distribution of outlierness scores, and (b) shows the visualization of the Egonet Density Power Law, where the red dotted line is the fitting line of power law. The top 10% outliers in near-clique pattern with the largest outlierness score are assigned red and marked with triangles.

this connected component, the average clustering coefficient is 0.431 and 26 triangles are constructed by their relationships, indicating that these voters tend to gather together.

*2) Mutual Voting Gang Detection:* To investigate the mutual voting gang anomalies, we build a voting network based on the voting actions, where each node denotes an account and each edge denotes the voting relationship between a pair of accounts. Then, we visualize the voting network by randomly selecting 5,000 edges. As shown in Fig. 11(a), the voting network contains many hub nodes whose one-hop neighborhood (also named egonet) is in a near-star pattern and the one-hop neighbors are almost not connected. While in voting collusion, the neighbors of a node tend to have more connections to maintain their role. Therefore, this kind of voting anomaly can be described by the near-clique pattern (Fig. 11(b)) that the one-hop neighbors of a node are very well connected. Based on this observation, we first detect the near-clique anomalies among the block producer candidates, and then discover the mutual voting gangs within the voting network of the anomalies. The whole analysis process contains three steps:

**Step 1: Near-clique anomaly detection.** Here the OddBall algorithm [24] is applied to identify the suspicious block producer candidates which have a relatively closed connected neighborhood. To spot near-clique anomalies, the OddBall algorithm makes use of two features extracted from the egonet of each node, namely $N_i$ (the number of neighbors of node $i$) and $E_i$ (the number of edges in the egonet of node $i$). Based on the observation that $N_i$ and $E_i$ for nodes in a network follow a *Egonet Density Power Law* $E_i \propto N_i^{\alpha}$ $(1 \leqslant \alpha \leqslant 2)$ [24], the outliers whose $E_i$ deviates from and is much greater than the expected value $CN_i^{\alpha}$ can be identified as near-clique anomalies. And the outlierness score of a node can be calculated as follow according to its distance to the fitting line:

$$Score(i) = \frac{max(E_i, CN_i^{\alpha})}{min(E_i, CN_i^{\alpha})} * log(|E_i - CN_i^{\alpha}| + 1), \quad (4)$$

where $\frac{max(E_i, CN_i^{\alpha})}{min(E_i, CN_i^{\alpha})}$ is the penalty coefficient measures the times that $E_i$ deviates from $CN_i^{\alpha}$ for a node $i$, and $log(|E_i - CN_i^{\alpha}| + 1)$ is a logarithmic distance measure.

**Step 2: Network reconstruction.** In order to represent the voting intensity between a pair of nodes, we assign a weight to each edge in the network. And the weight $w_{ij}$ of the edge between nodes $i$ and $j$ is decided by the voting intensity from $i$ to $j$ (denoted by $I_{ij}$) and the voting intensity from $j$ to $i$ (denoted by $I_{ji}$), i.e., $w_{ij} = I_{ij} + I_{ji}$. Here the voting intensity is calculated according to the voting frequency, duration, and voting power as Equation (5).

$$I_{ij} = \frac{1}{3} * (\frac{F_{ij}}{\sum_{u \in N(i)} F_{iu}} + \frac{T_{ij}}{\sum_{u \in N(j)} T_{uj}} + \frac{P_{ij}}{\sum_{u \in N(j)} P_{uj}}), \quad (5)$$

where $N(\cdot)$ represents the neighbors of a node, $F_{ij}$ represents the voting frequency from $i$ to $j$, $T_{ij}$ is the cumulative voting time duration and $P_{ij}$ is the average voting weight from $i$ to $j$. The item about the voting frequency measures the willingness of $i$ voting to $j$, and the items about the voting time duration and average voting weight measure $i$'s voting share of $j$. Then, we reconstruct the voting network by only keeping the relationships between the block producer candidates in the egonet of the detected accounts with near-clique neighborhood, paving the way for investigating the mutual voting gangs.

**Step 3: Community detection.** We apply the Louvain algorithm [25] on the weighted network to detect suspicious voting gangs, and the community detection results are obtained according to the voting intensity calculated by Equation (5) and modularity optimization. After obtaining the clustering results, we weed out the accounts with only one edge in the reconstructed network since these accounts have less possibility to participate in voting collusion in their cluster. And the outputs are clusters with more than one node.

*Results:* The distribution of the calculated outlinerness score is shown in Fig. 12(a), and here we choose the top 10% outliers with the highest outlierness score as the detected anomalies (totally 60 accounts), which are assigned red and marked with triangle in the figure. Fig. 12(b) shows the visualization of the *Egonet Density Power Law*. We can observe that these detected anomalies deviate from and are above the fitting line of the *Egonet Density Power Law*. The reconstructed network based on the detected anomalies contains 334 nodes and 1,604 undirected edges. And finally, we obtain 11 abnormal voting gangs by community detection.

In particular, we find out that a detected cluster, which includes 41 accounts, has 14 accounts coinciding with more than half of the accounts allegedly involved in mutual voting reported in [17]. The voting relationships in this cluster are

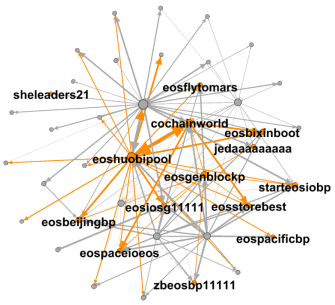



Fig. 13. Visualization of the voting relationships in the detected cluster of *'eoshuobipool'*, where the accounts reported in [17] are labeled and assigned the color orange.

shown in Fig. 13, where the detected labeled accounts and their relationships are assigned the color orange. As we can see in this figure, though there are some accounts having not been reported in [17], they have strong voting relationships with the cluster members. Moreover, 24 of the 26 reported accounts have been detected in the output clusters. Besides, we observed that there is a huge detected cluster consisting of 105 accounts. Among these accounts, 102 are named in a regular pattern with 'hayd' or 'g44t' as the first four letters and 'ge' as the last two letters, such as *'haydiojxgege'*, *'haydiojygage'*, *'g44tomrygige'* and *'g44tomzugqge'*. And 2 of the last 3 accounts are created by two accounts in this naming rule. These observations illustrate that this three-step method can help us notice some abnormal mutual voting gangs.

## VI. Discussion

Previous sections have presented a decentralization evolution analysis on EOSIO, a typical DPoS-based blockchain system, and reported some abnormal voting phenomena in the system. In this section, we discuss some implications from our findings for decentralization enhancement in DPoS-based blockchain systems.

**First**, voters in EOSIO only account for a small fraction of stakeholders, and their voting power is distributed unevenly, which means a few voters possessing a large number of stakes are sufficient to centralize the network. Moreover, since each voter can vote for up to 30 block producer candidates at the same time with no discount on the voting weight that each candidate can receive, the cost of voting manipulation for malicious account gangs is relatively low. Therefore, except to encourage the stakeholders to be active in the governance of EOSIO, the community can also look for a more reasonable value for the number of votes allowed per account. Existing work [26] has studied how to find the optimal number of votes allowed per account with governance game. Besides, another possible solution can be the introduction of an attenuation coefficient when a user votes for multiple candidates. In this way, a voter can hold different ratings for the multiple chosen candidates and give corresponding voting weights based on the ratings.

**Second**, voting proxies are popularly used recently, making the votes more centralized. Some voting proxies even attract new users by claiming that delegating the voting power to proxies can bring new users high profit with zero risk. However, these voting proxies are likely to become large staking pools, which not only violates the original design intention of voting proxies but also raises some security concerns similar to those brought by mining pools in Bitcoin [12], such as 51% attacks. With the proposed clustering method, we have identified some suspicious voting proxy gangs in which the members have similar voting behaviors in both time and space. The proxies in these detected clusters probably have made use of the collected voting power to vote for their alliance members. We suggest that users can utilize our method to investigate the voting behaviors of voting proxies before choosing a voting proxy.

**Third**, from the collected data, we observe that the block producer set has small variations, and this fact makes the system rather vulnerable to centralized control and malicious attacks. For example, Xu et al. [13] proposed a type of voting attack executed by block producers—the block producers can collude to blacklist the voters with large stakes which will threaten the authority of the block producers themselves. Since there is no protocol to prevent this type of vulnerability, it is vital to select reliable nodes in block producer election. As introduced earlier, our method can detect the block producers participating in mutual voting and thus can provide a reference for analyzing block producers in DPoS-based blockchain.

**Last but not least**, our analysis methods can be generalized to other DPoS-based blockchain systems like Tron, Binance Smart Chain and OKExChain [27] since these systems employ similar settings in their voting mechanism.

## VII. Related Work

We present the related work in this section, which mainly includes the decentralization analyses of DPoS-based blockchain systems and existing analyses in EOSIO.

### A. Decentralization Analysis of DPoS-based Blockchains

Decentralization is an important characteristic that can tell blockchain systems from traditional centralized systems. At present, there have been a few studies on analyzing the decentralization of blockchain systems. Most of the existing researches pointed out that Bitcoin and Ethereum have not achieved true decentralization in terms of mining power, network resources such as bandwidth, etc [28], [11], [12]. Since PoW-based blockchains become more and more centralized, Chen et al. [29] proposed tools to timely quantify the decentralization in terms of mining power in PoW-based blockchain systems. While for DPoS-based blockchains, the block production process does not require every node in the network to participate in block production directly. Instead, users can vote for a specified number of block production candidates, and the elected block producers generate blocks in turns. Thus in DPoS-based blockchains, the mining power is a weak influence factor of decentralization.

For DPoS-based blockchain systems, Jeongy [26] explored the relationship between decentralization and the number of votes allowed per account, and found that the "one vote per account" rule for centralization mitigation may be not necessary.

Rebello et al. [14] analyzed several quorum-based consensus protocols including the EOSIO protocol, and pointed out some clear vulnerabilities of these protocols. Kwon et al. [15] measured the degree of decentralization in several blockchain systems based on PoW, PoS, and DPoS with the entropy metric and explain why it is difficult to design a fully decentralized system. Li and Palanisamy [16] conducted a comparison of decentralization between Bitcoin and Steem [30], which are based on PoW and DPoS respectively. They found that compared with Steem, Bitcoin tends to be less decentralized in general and more decentralized among top stakeholders via some distributions and the entropy metric. However, there is a lack of in-depth analyses studying the decentralization evolution and uncovering the abnormal voting phenomena in DPoS-based systems.

### B. EOSIO Analysis

Nowadays, blockchain has attracted intensive interests of researchers. Existing studies cover several aspects including blockchain architecture and performance [1], [31], privacy and security [32], [33], smart contracts and transactions [34], [35], etc. As a successful and typical DPoS-based blockchain system, EOSIO has received attention of scholars these years.

In terms of architecture and security, Xu et al. [13] provided a comprehensive analysis on EOSIO's architecture, performance, and economy. Lee et al. [36] introduced four attacks in EOSIO, whose root cause lies in the characteristics of EOSIO. Quan et al. [37] proposed a static analysis tool to detect the fake-transfer vulnerabilities of EOSIO's smart contracts. He et al. [38] proposed a tool that can detect four popular vulnerabilities in EOSIO smart contracts at the WebAssembly code level. Lee et al. [39] proposed a kind of attack that can disturb the fairness of incentive policy by manipulating the block production schedule in EOSIO, and this attack can bring additional rewards or losses to block producers. In terms of blockchain transaction analysis, Huang et al. [19] investigated EOSIO and the associated DApps via measurement study. With a focus on security issues, they discovered some real-world attacks and developed detection techniques for fraudulent activities in EOSIO. Zheng et al. performed [21] a statistical analysis on 7 well-processed datasets in EOSIO, and outlined some possible research directions based on the proposed datasets. Zhao et al. [40] abstracted the records of four major activities in EOSIO as networks, and obtained some interesting observations via network metric analysis. However, none of them quantify the decentralization characteristic and investigate the abnormal voting phenomena based on the real transaction data in EOSIO.

## VIII. Conclusion and Future Work

This paper performed the first decentralization evolution study on DPoS-based blockchains. Based on the blockchain data from EOSIO, a typical DPoS-based blockchain system, we characterized the activities of voters, voting proxies, and block producers participating in the DPOS consensus process and obtained many insightful findings. Moreover, we presented methods to discover the abnormal voting gangs with similar voting behaviors and mutual voting behaviors. Some suspected voting manipulation phenomena have been revealed in our analysis. Besides, we provided some implications for enhancing the decentralization of EOSIO, which can also provide a reference for other DPoS-based blockchain systems.

For future work, we will expand our study in two directions. First, there may exist other complex abnormal voting patterns needing to be explored, we plan to investigate more abnormal voting phenomena in the system by combining with the token trading data and the resource trading data in EOSIO. Second, we want to extend this analysis to other DPoS-based blockchain systems, and build a decentralization monitoring platform to timely quantify the degree of decentralization and capture the abnormal phenomena in different DPoS-based blockchain systems.

## References


[1] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *Proc. 2017 IEEE Int. Congr. Big Data*. Honolulu, HI,USA: IEEE, 2017, pp. 557–564.

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

[3] G. Wood *et al.*, "Ethereum: A secure decentralised generalised transaction ledger," http://gavwood.com/Paper.pdf, 2014.

[4] H. Finney, "RPoW: Reusable Proofs of Work," https://cryptome.org/rpow.htm, 2004.

[5] D. Larimer, "Delegated proof-of-stake (dpos)," https://github.com/bitshares-foundation/bitshares.foundation/blob/master/papers/BitSharesBlockchain.pdf, 2014.

[6] Block.one, "EOS.IO technical white paper v2," https://github.com/EOSIO/Documentation/blob/master/TechnicalWhitePaper.md, 2018.

[7] T. Foundation, "Tron: Advanced decentralized blockchain platform," https://tron.network/static/doc/white_paper_v_2_0.pdf, 2018.

[8] Binance, "Binance Smart Chain," https://github.com/binance-chain/bsc, 2020.

[9] C. Ye, G. Li, H. Cai, Y. Gu, and A. Fukuda, "Analysis of security in blockchain: Case study in 51%-attack detecting," in *Proc. Int. Conf. Dependable Syst. Their Appl.* Dalian, China: IEEE, 2018, pp. 15–24.

[10] I. Eyal and E. G. Sirer, "Majority is not enough: Bitcoin mining is vulnerable," in *Proc. Int. Conf. Financ. Cryptogr. Data Secur.* Barbados: Springer, 2014, pp. 436–454.

[11] A. E. Gencer, S. Basu, I. Eyal, R. van Renesse, and E. G. Sirer, "Decentralization in Bitcoin and Ethereum networks," in *Proc. Int. Conf. Financ. Cryptogr. Data Secur.* Curaçao: Springer, 2018, pp. 439–457.

[12] C. Wang, X. Chu, and Y. Qin, "Measurement and analysis of the bitcoin networks: A view from mining pools," in *Proc. 6th Int. Conf. Big Data Comput. Commun. BigCom 2020*. IEEE, 2020, pp. 180–188.

[13] B. Xu, D. Luthra, Z. Cole, and N. Blakely, "EOS: An architectural, performance, and economic analysis," https://whiteblock.io/wp-content/uploads/2019/07/eos-test-report.pdf, 2018.

[14] G. A. F. Rebello, G. F. Camilo, L. C. Guimaraes, L. A. C. de Souza, and O. C. M. Duarte, "Security and performance analysis of quorum-based blockchain consensus protocols," https://www.gta.ufrj.br/ftp/gta/TechReports/RCG20c.pdf, 2020.

[15] Y. Kwon, J. Liu, M. Kim, D. Song, and Y. Kim, "Impossibility of full decentralization in permissionless blockchains," in *Proc. 1st ACM Conf. Adv. Financ. Technol.* Zurich, Switzerland: Association for Computing Machinery, 2019, pp. 110–123.

[16] C. Li and B. Palanisamy, "Comparison of decentralization in DPoS and PoW blockchains," in *Proc. Int. Conf. Blockchain*. Virtual Conference: Springer International Publishing, 2020, pp. 18–32.

[17] A. BERMAN, "EOS developer acknowledges claims of 'collusion' and 'mutual voting' between nodes," https://cointelegraph.com/news/eos-developer-acknowledges-claims-of-collusion-and-mutual-voting-between-nodes, 2018.

[18] eosflare.io, "eosflare.io," https://eosflare.io/, accessed Jan 31, 2021.

[19] Y. Huang, H. Wang, L. Wu, G. Tyson, X. Luo, R. Zhang, X. Liu, G. Huang, and X. Jiang, "Understanding (mis)behavior on the EOSIO blockchain," in *Proc. ACM Meas. Anal. Comput. Syst.* Boston, Massachusetts, USA: Association for Computing Machinery, 2020.

[20] EOSIO, "EOSIO core nodeos," https://eos.io/for-developers/build/nodeos/, accessed Jan 31, 2021.
[21] W. Zheng, Z. Zheng, H. Dai, X. Chen, and P. Zheng, "Xblock-EOS: Extracting and exploring blockchain data from EOSIO," *Inf. Process. Manag.*, vol. 58, no. 3, p. 102477, 2021.
[22] M. Perc, "The matthew effect in empirical data," *J. R. Soc. Interface*, vol. 11, no. 98, p. 20140378, 2014.
[23] K. Wu, B. Peng, H. Xie, and Z. Huang, "An information entropy method to quantify the degrees of decentralization for blockchain systems," in *Proc. 2019 IEEE Int. Conf. Electron. Inf. Emerg. Commun.* Beijing, China: IEEE, 2019, pp. 1–6.
[24] L. Akoglu, M. McGlohon, and C. Faloutsos, "Oddball: Spotting anomalies in weighted graphs," in *Proc. Pacific-Asia Conf. Knowl. Discov. Data Min.* Hyderabad, India: Springer, 2010, pp. 410–421.
[25] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech. Theory Exp.*, vol. 2008, no. 10, p. P10008, 2008.
[26] S. E. Jeong, "Centralized decentralization: Does voting matter? Simple economics of the DPoS blockchain governance," https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3575654, 2020.
[27] OKEx, "Okexchain documents," https://okexchain-docs.readthedocs.io/en/latest/index.html, 2021.
[28] A. Beikverdi and J. Song, "Trend of centralization in Bitcoin's distributed network," in *Proc. 2015 IEEE/ACIS 16th Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distributed Comput.* Takamatsu, Japan: IEEE, 2015, pp. 1–6.
[29] R. Chen, I.-P. Tu, K.-E. Chuang, Q.-X. Lin, S.-W. Liao, and W. Liao, "Endex: Degree of mining power decentralization for proof-of-work based blockchain systems," *IEEE Network*, vol. 34, no. 6, pp. 266–271, 2020.
[30] C. Li and B. Palanisamy, "Incentivized blockchain-based social media platforms: A case study of Steemit," in *Proc. 10th ACM Conf. Web Sci.* Boston, MA, USA: Association for Computing Machinery, 2019, p. 145–154.
[31] H. Huang, J. Lin, B. Zheng, Z. Zheng, and J. Bian, "When blockchain meets distributed file systems: An overview, challenges, and open issues," *IEEE Access*, vol. 8, pp. 50 574–50 586, 2020.
[32] M. C. K. Khalilov and A. Levi, "A survey on anonymity and privacy in Bitcoin-like digital cash systems," *IEEE Commun. Surv. & Tutorials*, vol. 20, no. 3, pp. 2543–2585, 2018.
[33] 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 Trans. Syst. Man, Cybern. Syst.*, 2021, to be published, doi: 10.1109/TSMC.2021.3049278.
[34] T. Durieux, J. a. F. Ferreira, R. Abreu, and P. Cruz, "Empirical review of automated analysis tools on 47,587 ethereum smart contracts," in *Proc. ACM/IEEE 42nd Int. Conf. Softw. Eng.* Seoul, South Korea: Association for Computing Machinery, 2020, p. 530–541.
[35] J. Wu, J. Liu, Y. Zhao, and Z. Zheng, "Analysis of cryptocurrency transactions from a network perspective: An overview," *J. Netw. Comput. Appl.*, vol. 190, p. 103139, 2021.
[36] S. Lee, D. Kim, D. Kim, S. Son, and Y. Kim, "Who spent my EOS? On the (in)security of resource management of EOS.IO," in *Proc. 13th USENIX Work. Offensive Technol.* Santa Clara, CA: USENIX Association, 2019.
[37] L. Quan, L. Wu, and H. Wang, "EVulHunter: Detecting fake transfer vulnerabilities for EOSIO's smart contracts at webassembly-level," *arXiv preprint arXiv:1906.10362*, 2019.
[38] N. He, R. Zhang, H. Wang, L. Wu, X. Luo, Y. Guo, T. Yu, and X. Jiang, "EOSAFE: security analysis of EOSIO smart contracts," in *Proc. 30th USENIX Secur. Symp.* Virtual Conference: USENIX Association, 2021, pp. 1271–1288.
[39] D. Lee and D. H. Lee, "Push and pull: Manipulating a production schedule and maximizing rewards on the EOSIO blockchain," in *Proc. Third ACM Work. Blockchains, Cryptocurrencies Contract.* Auckland, New Zealand: Association for Computing Machinery, 2019, pp. 11–21.
[40] Y. Zhao, J. Liu, Q. Han, W. Zheng, and J. Wu, "Exploring EOSIO via graph characterization," in *Proc. Int. Conf. Blockchain Trust. Syst.* Dali, China: Springer, 2020, pp. 475–488.

**Jieli Liu** received her B.Eng. in Software Engineering from Sun Yat-sen University, Guangzhou, China, in 2019. She is currently studying toward a Ph.D. degree in the School of Software Engineering, Sun Yat-sen University. Her current research interests include blockchain, network science, data mining, and machine learning with graphs.

**Weilin Zheng** is currently pursuing the M.Eng. degree with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China. His research interests include performance monitoring and optimization, blockchain computing power utilization, blockchain data analysis, and blockchain-based decentralized applications.

**Dingyuan Lu** received the B.Eng. in computer science and technology from Shanxi University, Taiyuan, China, in 2018. He is currently pursuing an M.Eng. degree in the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China. His current research interests include applications of network science, blockchain, and machine learning with graphs.

**Jiajing Wu** (S'11–M'14–SM'19) received the B.Eng. degree in communication engineering from Beijing Jiaotong University, Beijing, China, in 2010, and the Ph.D. degree from Hong Kong Polytechnic University, Hong Kong, in 2014. She was awarded the Hong Kong Ph.D. Fellowship Scheme during her Ph.D. study in Hong Kong (2010–2014).

In 2015, she joined Sun Yat-sen University, Guangzhou, China, where she is currently an Associate Professor. Her research focus includes blockchain, graph mining, network science. She serves as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS.

**Zibin Zheng** received a Ph.D. degree from the Chinese University of Hong Kong in 2011.

He is currently a Professor of Data and Computer Science with Sun Yat-sen University, China. He serves as the Chair of the Software Engineering Department, Pearl River Young Scholars, and the Founding Chair of the Services Society Young Scientists Forum (SSYSF). In the past five years, he published over 120 international journal and conference papers, including three ESI highly-cited papers, 40 ACM/IEEE TRANSACTIONS papers. According to Google Scholar, his papers have more than 6300 citations, with an H-index of 41. His research interests include blockchain, services computing, software engineering, and financial big data. He was the recipient of several awards, including the outstanding Thesis Award of CUHK, in 2012, the ACMSIGSOFT Distinguished Paper Award at ICSE2010, the Best Student Paper Award at ICWS2010, and IBM Ph.D. Fellowship Award. He served as CollaborateCom'16 General co-Chair, ICIOT'18 PC co-Chair, and IoV'14 PC co-Chair.