Measuring Decentralization in Emerging Public Blockchains

Yongpu Jia*, Changqiao Xu*, Zhonghui Wu*, Zichen Feng*, Yaxin Chen* and Shujie Yang*✉
*State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications, Beijing, China
Email: {jia980328, cqxu, zhwu, fengzichen1027, chenyx2139, sjyang}@bupt.edu.cn

2022 International Wireless Communications and Mobile Computing (IWCMC) | 978-1-6654-6749-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IWCMC55113.2022.9825341

***Abstract*—Bitcoin and Ethereum have always been the two major heavyweight infrastructures in the blockchain space. However, Low throughput and high cost hinder their further development. Recently, some emerging public blockchains have become popular. They all have efficient transaction confirmation mechanism and cheap interaction costs. However, the advantages are actually a sacrifice of decentralization. As we all know, decentralization is an essential feature of blockchain. Therefore, it requires the conceiving of up-to-date metrics of decentralization measurement. However, there is little research on the degree of decentralization of these emerging public blockchains in the past. This paper studies nine popular public chains such as Binance Smart Chain, Cardano, and Avalanche. Since these public chains mostly use the consensus mechanism of POS variants, the distribution of governance token balances on the chain can reflect the decentralization of the blockchain. Hence, We evaluate the distribution of the token balance of those public blockchains to indicate their decentralization degree. Two kinds of indicators are adopted and redesigned: information entropy and Gini coefficients. Among the nine public blockchains we selected, Cardano, Tron and Polkadot have a higher degree of decentralization, while Elrond and Binance Smart Chain have a lower degree of decentralization. We think our work will be helpful for future research on the degree of blockchain decentralization.**

***Index Terms*—blockchain, decentralization, Information entropy, Gini coefficient**

I. Introduction

In 2008, a mysterious man under the pseudonym Satoshi Nakamoto published the Bitcoin white paper [1]. It kicked off a technological revolution about trust and sharing. Since then, it has been possible to deliver value without any third-party institutions. Bitcoin is a great social experiment. Over the past decades, bitcoin has reached unprecedented heights both in value and usage. People extracted the critical technology in Bitcoin, calling it blockchain. The essence of the blockchain is a growing list of the immutable block that contain many transactions and the previous block's hash. Blockchain is also a distributed ledger that ensures that all nodes reach consensus on the public ledger in an untrusted and distributed environment through consensus algorithms.

Decentralization is an essential feature of blockchain. All transactions and events are recorded in the blockchain, owned and monitored by all participants, without central control. Blockchain's peer-to-peer system provides a verifiable ledger without centralized permissions and solves the single point of failure and the single point of trust [2]. Since there is no central server, blockchain nodes are spread across all corners of the world to maintain the blockchain together, and each node has same permissions to the ledger. In the blockchain, the unit of actors is an address rather than a specific person, enabling anonymity and privacy.

Driven by its promising features of immutability, transparency, and privacy features, blockchain has gained more and more focus from academia and industry. In the beginning, Bitcoin, as the most important application of blockchain, enables instant value transfer, improving our payment experience and pushing us from the Internet of Information to the Internet of Value. However, Bitcoin can only support a few simple scripts. After that, Ethereum [3] emerged, and Turing-complete smart contracts extended the blockchain application scenario to all aspects of the real world [4]. Smart contracts are computer programs triggered by transactions enforced based on consensus between untrusted parties. Through smart contracts, blockchain technology can be used in various scenarios in life such as finance, voting, supply chain, energy, and healthcare [5].

Bitcoin and Ethereum are still the two heavyweight blockchain infrastructures, but they also have limitations, for example, Bitcoin's low throughput and high confirmation latency. Although Ethereum has dramatically improved in the above two aspects, the cost of executing smart contracts is high. It usually costs tens of dollars to perform a simple transfer in the Ethereum and even hundreds of dollars when the blockchain network is congested. In order to solve the above problems, some newly-born public blockchains began to emerge, and they made significant breakthroughs in high performance and low cost through some novel designs, so they quickly became popular and had a large number of user groups—for example, Binance Smart Chain, Cardano, Avalanche.

Due to the well-known trilemma triangle of Blockchain, i.e., blockchain is challenging to balance decentralization, security, and scalability. Behind the excellent performance of the emerging public blockchain is a sacrifice of decentralization. This article analyzes the degree of decentralization of these public blockchains, aiming to provide new metrics for reference. The contributions of our research are as follows:

- We propose a method based on address balance allocation to assess the degree of decentralization in POS blockchain.

- Our measurement use two indicators, information entropy and Gini coefficient. The analysis shows that when the degree of decentralization of the two chains is largely different, the two indicators can get the same result for comparing the two chains. However, when the two degrees of decentralization are similar, the two indicators may give different results. In general, both indicators can measure the degree of decentralization of the blockchain. To a certain extent, the two indicators are consistent with the comparison of decentralization between public chains.
- Most of the past research is on Bitcoin and Ethereum and less on other public blockchains. We evaluated the degree of decentralization of the nine recently emerging public blockchains. The results show that Cardano, Tron, and Polkadot have a higher degree of decentralization among the nine public chains we selected. In comparison, Elrond and Binance Smart Chain have a lower degree of decentralization.

The rest of this article is organized as follows: Section II discusses the related work; section III introduces the measurement metrics of the degree of decentralization; in the section IV, we evaluate and analyze the nine public chains; and finally, section V presents the conclusions.

II. Related Work

Decentralization is an essential feature of blockchain, and scholars have done much research in this area [6]–[10]. In terms of indicators of the degree of decentralization of blockchain. Wu, et al. in [11] proposes a method for assessing the degree of blockchain decentralization based on information entropy. Using this method, it is concluded that Bitcoin is more decentralized than Ethereum. Balaji S Srinivasan, et al. in [12] propose an indicator called the minimum Nakamoto coefficient to compare the degree of decentralization of blockchains. The authors define a blockchain system as multiple subsystems, then calculate the difficulty of manipulating the subsystem by controlling as many entities as possible, and finally use the minimum of each result as a decentralized index of the entire system. The higher the numerical value of the system, the higher the degree of decentralization. Li, et al. in [13] Starting with Bitcoin and Steem compares the degree of decentralization of the DPOS and POW blockchains. they find that compared with Steem, Bitcoin tends to be more decentralized among top miners but less decentralized in general. Multiple metrics were used in [14] by lin, et al. to measure the degree of blockchain decentralization, and a sliding window approach was taken to obtain more fine-grained information. Their results demonstrate that the use of sliding windows could reveal additional cross-interval information overlooked by the fixed window based measurements, thus enhancing the effectiveness of measuring decentralization in terms of continuous trends and abnormal situations. A systematic approach is used in [15] by Gochhayat, et al. to measure the overall decentralization of the blockchain from three levels: the governance layer, the network layer, and the storage layer. They noticed that, with time, both Bitcoin and Ethereum networks tend to behave like centralized systems where a few nodes govern the whole network. However, most research has focused on Bitcoin and Ethereum. With the rapid development of the blockchain field, some blockchains with novel protocols have performed exceptionally well, but less research has been done on them. These new public blockchains are more likely to adopt the POS consensus mechanism rather than POW as in Bitcoin and Ethereum. As far as we know, our paper is the first research work to assess the degree of decentralization of these new public blockchains from the perspective of address balance distribution.

III. Measurement Metrics

Most high-performance public blockchains now adopt the POS consensus mechanism or its variants. So the number of governance tokens can directly affect the weight of producing blocks in accounting. The distribution of governance tokens can reflect the degree of decentralization of the public blockchains. Uneven distribution of governance tokens, i.e., a small number of addresses hold most of the governance tokens, and those addresses can grasp most of the block accounting rights, indicating that the degree of decentralization of the public blockchain is low. Conversely, if the distribution of governance tokens is relatively even, it indicates that the chain is highly decentralized. This paper uses two indicators, information entropy and Gini coefficient, to quantify each public blockchain's token distribution and evaluate the degree of decentralization of each blockchain. Next, we introduce the two measurements separately.

A. Information Entropy

Entropy is a physical concept of thermodynamics, a measure of the degree of system chaos or disorder. The larger the entropy, the more chaotic the system, the less information it carries. The smaller the entropy, the more orderly the system is and the more information it carries. Information entropy borrows from the concept of entropy in thermodynamics to describe the amount of event information on average. So mathematically, information entropy is an expectation of the amount of information contained in an event. Depending on the properties of entropy, we can use it to measure the degree of randomness or disorder of an event. Extending to the blockchain scenario, we can use information entropy to assess the randomness and chaos of the distribution of tokens at each address of the public blockchain. The calculation method of information entropy is as follows:

$$p_i = \frac{x_i}{\sum_{i=1}^{n} x_i} \tag{1}$$

$$E = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{2}$$

Where x_i denotes the number of governance tokens held by the address ranked i and n denotes the number of addresses. In this work, $n = 1000$, which indicates that information entropy or Gini coefficient is computed based on the balance of top-1000 addresses. The greater the entropy value, the greater the randomness of the token distribution, the higher the degree of disorder, and the higher the degree of decentralization of the blockchain.

B. Gini Coefficient

Gini coefficient is a standard indicator of inequality in the economic context, with a maximum value of 1 representing extreme inequality and a minimum of 0 representing an absolute average. The Lorenz curve starts with the poorest individuals and describes the relationship between the cumulative percentage of total income and the cumulative number of people. The Gini coefficient measures the area relationship between the Lorenz curve and the supposed absolute equality line. It can be thought of as the ratio of the area that lies between the line of equality and the Lorenz curve over the total area under the line of equality. When measuring the degree of blockchain decentralization, the Gini coefficient can describe the distribution of blockchain governance tokens. Its calculation formula is as follows:

$$A = \frac{n \sum_{i=1}^{n} x_i}{2} \tag{3}$$

$$B = \sum_{i=1}^{n} i x_i \tag{4}$$

$$G = \frac{A - B}{A} \tag{5}$$

The smaller the Gini coefficient, the more evenly distributed the governance tokens, and the higher the degree of blockchain decentralization.

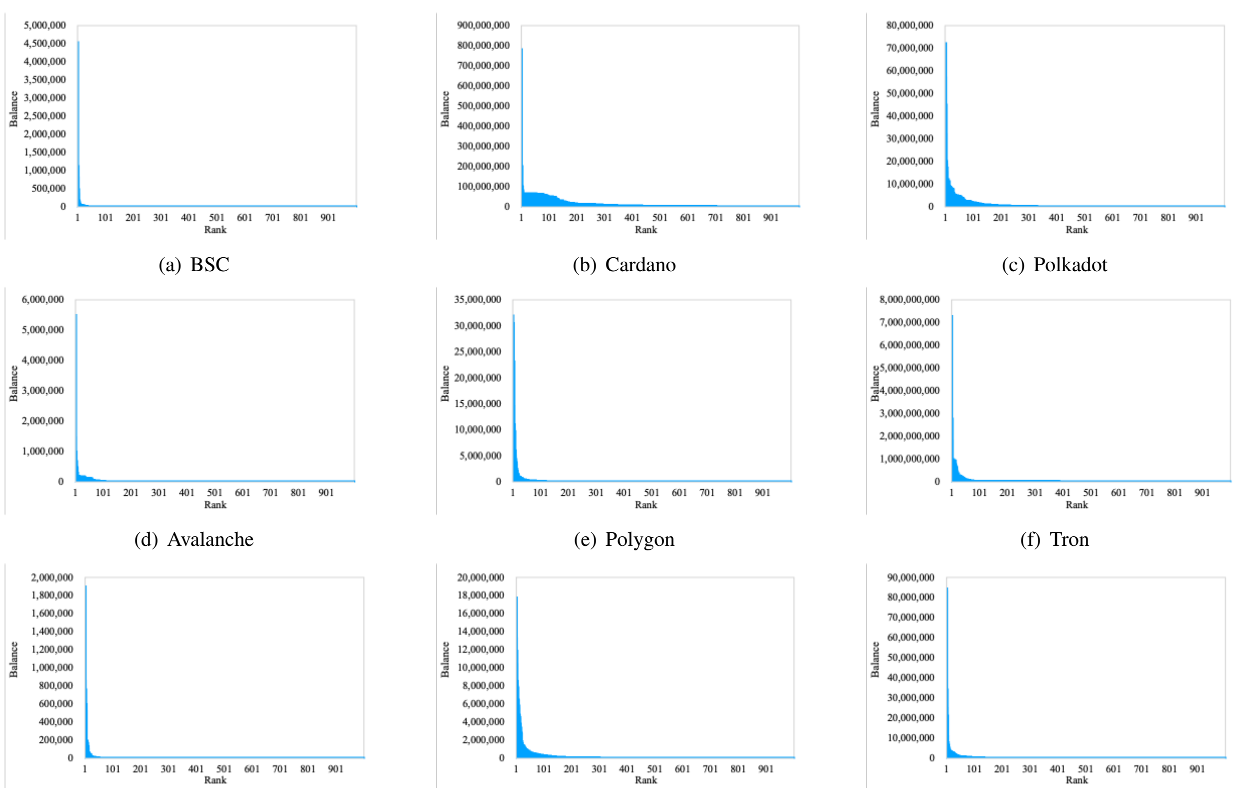
TABLE I
DATA SOURCES

Chain	Token	Blockchain Explorer
Binance Smart Chain	BNB	bscscan.com
Cardano	ADA	adastat.net
Polkadot	DOT	polkascan.io
Avalanche	AVAX	snowtrace.io
Polygon	MATIC	polygonscan.com
TRON	TRX	tronscan.org
Elrond	EGLD	explorer.elrond.com
Cosmos	ATOM	atomscan.com
Fantom	FTM	ftmscan.com

IV. Experimentation and Analysis

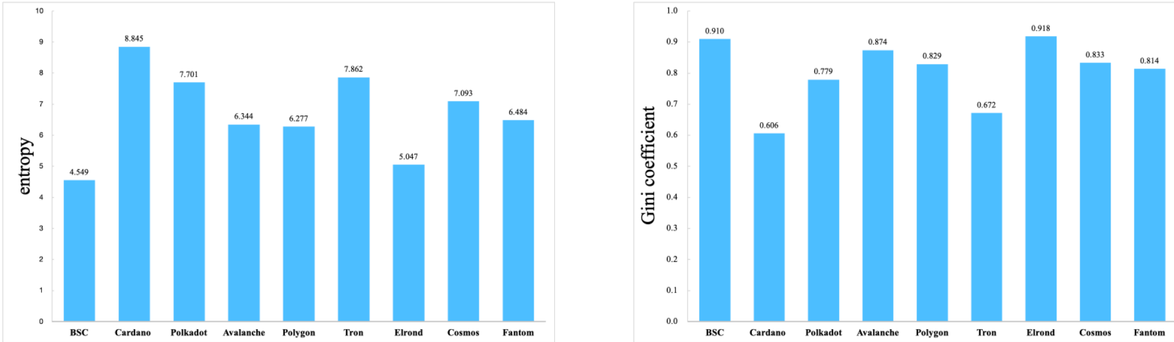
A. Data Collection

On December 16, 2021, we selected the top 1000 addresses with balances from the official blockchain browsers of several blockchains. The domain name of each blockchain browser is shown in Table I. In the selection process, we try to remove the contract address as much as possible because now the ecological technologies such as Defi and cross-chain on the blockchain are prevalent, and many contracts with a large



(a) BSC (b) Cardano (c) Polkadot
(d) Avalanche (e) Polygon (f) Tron
(g) Elrond (h) Cosmos (i) Fantom

Fig. 1. Address Balance Distribution



(a) Infomation entropy

(b) Gini coefficient

Fig. 2. Quantitative Analysis of the Degree of Decentralization

TABLE II
RANK OF DECENTRALIZATION

Decentralization	**High**			**medium**				**low**	
Entropy	Cardano	Tron	Polkadot	Cosmos	Fantom	Avalanche	Polygon	Elrond	BSC
Gini coefficient	Cardano	Tron	Polkadot	Fantom	Polygon	Cosmos	Avalanche	BSC	Elrond

number of tokens may be related to staking or cross-chain bridges, in which the tokens belong to many people rather than one. The distribution of address balances in the top 1000 of each public chain is shown in Figure 1.

B. Quantitative Analysis

Using the top 1000 address balances of each public blockchain as the dataset, and using two measurement methods of information entropy and Gini coefficient respectively for quantitative analysis, we obtained the results shown in Figure 2. The information entropy and Gini coefficients of each blockchain are: BSC{4.549, 0.910}; Cardano{8.845, 0.606}; Polkadot{7.701, 0.779}; Avalanche{6.344, 0.874}; Polygon{6.277, 0.829}; Tron{7.862, 0.672}; Elrond{5.047, 0.918}; Cosmos{7.093, 0.833}; Fantom{6.484, 0.814}. In order of entropy, the nine chains are ranked as Cardano, Tron, Polkadot, Cosmos, Fantom, Avalanche, Polygon, Elrond, and BSC. This is also the ranking of the degree of decentralization reflected by information entropy from high to low. In order of Gini coefficients from small to large, the ranking of these nine public chains is Cardano, Tron, Polkadot, fantom, polygon, Cosmos, avalanche, BSC, Elrond, which is the ranking of the degree of decentralization reflected by the Gini coefficient from high to low. We observed that the rankings obtained by the two indicators were slightly different. When the information entropy of the two chains was similar, the Gini coefficient was also similar, and vice versa. When the two chains are similar in degree of decentralization, the two indicators may get different results, but the indicator values of the two chains are similar. If the nine public chains selected in this paper are divided into three categories according to the degree of decentralization, the results obtained by the two indicators are roughly same. The results are as follows: in these nine public chains, Cardano, Tron, and Polkadot have a high degree of decentralization, avalanche, polygon, Cosmos, fantom have a medium degree of decentralization, and BSC and Elrond have a low degree of decentralization. This means that both indicators can measure the degree of decentralization of the blockchain. To a certain extent, the two indicators are consistent in comparing the degree of decentralization between public blockchains.

V. CONCLUSION AND FUTURE WORK

This paper studies the degree of decentralization of nine emerging public blockchains(BSC, Cardano Polkadot, Avalanche, Polygon, Cosmos, Tron, Elrond, and Fantom). Our research starts with the distribution of address balances in these POS public blockchains, using two indicators of information entropy and the Gini coefficient to quantify the degree of decentralization of the blockchain. In general, Cardano, Tron and Polkadot have a higher degree of decentralization in these nine public chains, while Elrond and BSC have a lower degree of decentralization. We believe that the approach and work in this paper could facilitate future research on the degree of blockchain decentralization. However, the actual degree of decentralization of blockchain may be more complex. In terms of the address balance-based approach adopted in this article, the degree of decentralization of the blockchain may be overestimated because one person can spread the coins across many addresses; it may also be underestimated because the wallet of the exchange is also in it, at least nominally, the coins in this address belong to thousands of users. In addition, some addresses may have lost tokens long ago (due to reasons such as the loss of private keys), which can also lead to underestimating the degree of decentralization. Future work could do further research in this direction.

References

[1] Nakamoto, Satoshi. ”Bitcoin: A peer-to-peer electronic cash system.” Decentralized Business Review (2008): 21260.

[2] Zyskind, Guy, and Oz Nathan. ”Decentralizing privacy: Using blockchain to protect personal data.” 2015 IEEE Security and Privacy Workshops. IEEE, 2015.

[3] Wood, Gavin. ”Ethereum: A secure decentralised generalised transaction ledger.” Ethereum project yellow paper 151.2014 (2014): 1-32.

[4] Buterin, Vitalik. ”A next-generation smart contract and decentralized application platform.” white paper 3.37 (2014).

[5] Abou Jaoude, Joe, and Raafat George Saade. ”Blockchain applications–usage in different domains.” IEEE Access 7 (2019): 45360-45381.

[6] Gencer, Adem Efe, et al. ”Decentralization in bitcoin and ethereum networks.” International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, 2018.

[7] Beikverdi, Alireza, and JooSeok Song. ”Trend of centralization in Bitcoin’s distributed network.” 2015 IEEE/ACIS 16th international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD). IEEE, 2015.

[8] Tschorsch, Florian, and Björn Scheuermann. ”Bitcoin and beyond: A technical survey on decentralized digital currencies.” IEEE Communications Surveys & Tutorials 18.3 (2016): 2084-2123.

[9] Kwon, Yujin, et al. ”Impossibility of full decentralization in permissionless blockchains.” Proceedings of the 1st ACM Conference on Advances in Financial Technologies. 2019.

[10] Gervais, Arthur, et al. ”Is bitcoin a decentralized currency?.” IEEE security & privacy 12.3 (2014): 54-60.

[11] Wu, Keke, et al. ”An information entropy method to quantify the degrees of decentralization for blockchain systems.” 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC). IEEE, 2019.

[12] Balaji S Srinivasan et al., ”Quantifying decentralization”, 2017, [online] Available: https://news.earn.com/quantifying-decentalization-e39db233c28e.

[13] Li, Chao, and Balaji Palanisamy. ”Comparison of decentralization in dpos and pow blockchains.” International Conference on Blockchain. Springer, Cham, 2020.

[14] Lin, Qinwei, et al. ”Measuring decentralization in bitcoin and ethereum using multiple metrics and granularities.” 2021 IEEE 37th International Conference on Data Engineering Workshops (ICDEW). IEEE, 2021.

[15] Gochhayat, Sarada Prasad, et al. ”Measuring Decentrality in Blockchain Based Systems.” IEEE Access 8 (2020): 178372-178390.