Available online at www.sciencedirect.com
ScienceDirect
ICT Express 10 (2024) 380–386
www.elsevier.com/locate/icte

Overview and empirical analysis of wealth decentralization in blockchain networks

Mindaugas Juodis, Ernestas Filatovas, Remigijus Paulavičius*

Institute of Data Science and Digital Technologies Vilnius University, Akademijos str. 4, Vilnius, LT-08663, Lithuania

Received 31 March 2023; received in revised form 30 December 2023; accepted 5 February 2024
Available online 7 February 2024

Abstract

The decentralization paradigm has made blockchain one of the most disruptive technologies today. When evaluating the level of decentralization, the key metric for most public blockchain networks is the degree of decentralization of the resources responsible for determining who generates the blocks. In turn, it facilitates a greater understanding of both security and scalability on a blockchain. This work provides an overview of the current state-of-the-art on wealth decentralization, which has not yet received the attention it deserves. We collect data, calculate various wealth decentralization metrics, and compare our results with research on the same methodology. As the amount of data for various blockchains increases rapidly, it is helpful to have techniques to aggregate data for statistical analysis. We introduce and provide conservative estimates of decentralized group metrics based on the reduced data and compare them with full-data measurements. Our research considers both the Layer 1 blockchains of Bitcoin and Ethereum, along with Layer 2 blockchains such as Arbitrum, Optimism, and Polygon.

Keywords: Blockchain; Decentralization; Wealth distribution; Bitcoin; Ethereum; Arbitrum; Optimism; Polygon

1. Introduction

Blockchain is recognized as one of the most disruptive technologies [1] introduced in Bitcoin [2], which is based on the decentralization paradigm. Public blockchains such as Bitcoin and Ethereum are ledgers of transactions operated by nodes on blockchain networks in a decentralized manner. Due to its decentralized nature and features such as immutability, traceability, transparency, and security, blockchain technology drives the rise of decentralized platforms and applications in various domains [3], such as finance [4], healthcare [5], social media [6], etc. Blockchain decentralization is most commonly viewed as a reduced level of centralized control. Therefore, it could be understood as a process of the permanent transfer of control from centralized ownership to a distributed network.

Although the designs of various blockchain networks often differ significantly, fortunately, they share enough traits to evaluate and compare their decentralization. Such traits include the same consensus protocols, similar governance mechanisms, comparable network structure, information on the holder's address balances for wealth distribution, etc., and their decentralization could be evaluated from these perspectives [7,8]. However, as a rule, public blockchains are based on cryptocurrencies and have monetary assets (coins) distributed between blockchain participants. Cryptocurrencies play an important role in the usage and development of blockchain ecosystems; therefore, the evaluation of wealth decentralization is very relevant.

Considering market capitalization, user base, and popularity, there are two dominating blockchain networks: Bitcoin (BTC) and Ethereum (ETH). Bitcoin is based on Proof-of-Work (PoW) consensus protocol, while Ethereum originally used PoW too, but on 15 September 2022, it shifted to Proof-of-Stake (PoS) protocol. These two types of consensus protocols (PoW and PoS) dominate most blockchain networks today [9–11]. Those networks are comparable regarding wealth since wealth decentralization involves the distribution of coins across blockchain participants, which is unrelated to the network consensus protocol. Additionally, aligning with prevailing trends in the blockchain industry, we explored three

---

* Corresponding author.
E-mail addresses: mindaugas.juodis@mif.vu.lt (M. Juodis), ernestas.filatovas@mif.vu.lt (E. Filatovas), remigijus.paulavicius@mif.vu.lt (R. Paulavičius).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2024.02.002

prominent Layer 2 protocols, Arbitrum (ARB), Optimism (OP), and Polygon (MATIC), employing the same statistical analysis. Moreover, since the Ethereum case involves a transition from two different consensuses, we analyze and measure the decentralization of this network in both states (PoW and PoS), considering two different timeframes.

We also focus on providing all popular metrics in one place for the same data set. Note that an important quantification metric such as the Herfindahl–Hirschman index (HHI) is seldom used in blockchain decentralization research area. This study shows that considering this metric enriches the decentralization quantification process. Additionally, our work introduces the methodology of simplified statistics using group-based constructions and compares these measurement results with full data measurements. Also, we show that simplified data collection and construction could give considerably conservative estimates compared to the entire sample statistics.

To sum up, the main contributions of this work are the following:

- It reviews approaches to measuring wealth decentralization for blockchain networks.
- It proposes a group-based metric based on the Herfindahl–Hirschman index to assess the measurement's conservatism and estimate the most relevant data points.
- It collects data, employs metrics, and provides measurements for prominent blockchain networks, including Bitcoin and Ethereum, along with widely adopted Ethereum Layer 2 solutions: Arbitrum, Optimism, and Polygon.

The remainder of the paper is organized as follows. Section 2 reviews related literature on various aspects of wealth decentralization. Section 3 provides definitions and an overview of the decentralization metrics used throughout the article. Section 4 gives quantification and calculation results with a comparison of it provided in previous research works. Finally, Section 5 concludes the work.

2. Related works on wealth decentralization

Wealth decentralization is commonly defined as uniformity in wealth distribution and is about the distribution across ownership of addresses and concentration. Thus it captures the decentralization of monetary assets in tokens and native cryptocurrencies distributed across blockchain users. However, there is a big difference between addresses and individual holders. For example, many Top Bitcoin addresses belong to exchanges, companies, exchange-traded funds (ETFs), or parties representing many respective holders. Ideally, the initial distribution of a cryptocurrency should be as fair as possible to get off to a good start.

To remind Bitcoin's evolution, note that there was no "premine" (the practice of privately mining coins, usually by the developers) nor "instamine", where many cryptocurrencies/tokens are brought into existence at once. Meanwhile, Ethereum had a public presale of about 60M ethers. The public sale comprised 80% of the initial supply, while the other 20% was allocated to the Ethereum Foundation and early Ethereum contributors.

The authors of [12] discovered that Bitcoin and Ethereum have very high levels of wealth concentration. While in [13], it was observed that Bitcoin is 12% more decentralized than Ethereum in mining power and 9% more decentralized in wealth. [14] apply the Gini index (see for more information on it) to quantify wealth and consensus decentralization. An article [15] found that the statement "2% of addresses control 95% of the supply". We repeated equivalent calculations on 10.11.2022 using BitInfoCharts[1] data and found that 2.24% of BTC addresses (not individual holders) control 94.82% of the circulating supply, which is quite close to this statement. In [16], it was observed that PoS-type blockchains are susceptible to the Matthew effect.

Table 1 summarizes works on wealth decentralization, emphasizing the metrics used – Gini index, Nakamoto coefficient, Shanon entropy, and Herfindahl–Hirschman index (HHI) (see Section 3) – and the blockchains analyzed. The list is short, and most of the works consider only one or two metrics. The most popular metric is the Gini index, while the second is the Nakamoto coefficient. Other metrics are rarely found in the literature. Note that some of the works mentioned in Table 1 use specific granularity, which is actually a specific time horizon. Also, another approach is when the method is applicable to all time horizons, i.e., daily, weekly, monthly, then we say that we have all granularities. The work of [17] considered more blockchains, but the authors limited their research only to the Top 10, 30, and 50 addresses, which is a considerably small number of data points. In [18], authors use various blockchains, but their method involves a heuristic approach.

Overall, the literature on wealth decentralization is growing but limited, and all of this motivates one to pay more attention to the analysis of decentralization through the lens of wealth. Furthermore, we note that, with respect to Layer 2 protocols, there is no literature on wealth decentralization; therefore, this work provides pioneering results in this direction. Further, we investigate the measurement of wealth decentralization by applying all four metrics and also introduce a new block-based metric.

3. Decentralization metrics

This section briefly reviews the main statistical metrics commonly used to quantify blockchain decentralization and introduces a new group-based Herfindahl–Hirschman index.

Herfindahl–Hirschman Index (HHI) is a standard measure of concentration and could be applied to measure the decentralization of blockchains. Let a participant (holder) share be denoted by a variable x_i $(1 \leq i \leq N)$, which corresponds to the percentage of assets (coins or tokens) held on the ith address so that x_i are ordered and normalized to 1. The Herfindahl–Hirschman index is calculated by squaring

---

[1] https://bitinfocharts.com/.

Table 1
Literature summary on wealth decentralization.

Source	Gini index	Shanon entropy	HHI	Nakamoto coefficient	Used blockchains	Contribution and limitations
[19]	–	–	+	–	ETH, ETH tokens	Daily granularity
[17]	+	+	–	+	BTC, ETH, BCH, LTC, Dash, Doge, BCH, ETC	Their approach uses only the Top 10, 30, and 50 addresses.
[7]	–	+	–	–	BTC, ETH, tokens	They use a decentralization index with construction based on Shanon Entropy.
[18]	+	–	–	+	BTC, ETH, BCH, LTC, Doge, Monacoin	Multiple input heuristic method, weekly granularity
[14]	+	–	–	–	BTC	All granularities
[13]	–	+	–	–	BTC, ETH	Weekly granularity
[20]	+	–	–	+	BTC, ETH	All granularities
This study	+	+	+	+	BTC, ETH, ARB, OP, MATIC	All granularities

each participant's share x_i on a blockchain and summing the resulting numbers

$$HHI = \sum_{i=1}^{N} x_i^2, \text{ where } \sum_{i=1}^{N} x_i = 1 \text{ and } x_1 \geq \cdots \geq x_N. \quad (1)$$

An application of HHI is to evaluate the concentration of holders on a particular blockchain. For example, measuring what share of all coins is taken by single addresses or groups of addresses. In this work, we introduce a group-based HHI and use this modified estimate to provide information on the size of the blocks (U) and the structure of the sample needed for the blockchain to withstand the challenges of decentralization.

The *group-based Herfindahl–Hirschman index* is defined as follows:

$$HHI_{\text{group}} = \sum_{i=1}^{n} U_i^2, \quad U_i = \sum_{j=m_i}^{m_{i+1}} x_j \text{ and } \sum_{i=1}^{n} m_i = N. \quad (2)$$

Here, we employ a square sum of groups of input variables rather than a straightforward square sum. This formulation improves the statistical dispersion, providing a more accurate metric estimation. For more theoretical insights, we refer to [21].

We utilize the notation $HHI_{N:y}$, where we partition our sample of N shares into y groups. When the number of groups (N/y) is not an integer, we use the floor function and select the highest integer less than or equal to N/y. We start with a group of size 2 and increase the size of the groups in each iteration by multiplying by 2. Note that we fixed our sample size to 10,000 so that throughout the article, we could achieve comparability with state-of-the-art in all metrics. Different sample sizes and different groups of samples are subject to investigation for future research.

Theil L, Shannon entropy is a centralization metric commonly used in physics, especially in quantum information theory (Von Neumann entropy), thermodynamics, and statistical physics. Shannon entropy (S) is defined as follows: $S = \sum_{i=1}^{N} x_i \log x_i$. In the general case, the values of Shannon entropy are unbounded. However, when the number of available states (N) is fixed, the Shannon entropy takes values from an interval $[0, \log(N)]$. Maximal centralization corresponds to a Shannon entropy value equal to zero, and decentralization grows with the growth of the entropy value.

The Gini index is an inequality measure widely used in economics and social statistics and is defined as follows: $G = \frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left| x_i - x_j \right|$. The Gini index, also called the coefficient, is the most popular in economics but can be applied in various other fields. Thus, it is usually used for measuring the decentralization of blockchains. The Gini index takes values from 0 (complete decentralization of wealth) to 1 (absolute centralization). The Lorenz curve (see Section 4.4) graphically represents the percentage of wealth accumulated by various parts of the participants ordered by the size of their wealth. The line of equality is at a 45° angle and represents the perfectly equal distribution of wealth. The area between the line of equality and the Lorenz curve can be used to determine the Gini index.

Nakamoto coefficient is a relatively new centralization metric in the analysis of cryptocurrency-based blockchains, defined as follows: $K = \min\left\{ n \in N : \sum_{i=1}^{n} x_i > \frac{1}{2} \sum_{i=1}^{N} x_i \right\}$. and could be seen as the minimum number of participants in a given subsystem required to reach 51% of the total capacity. [20] pointed out that the Gini index, while indicative, is not a good measure of the vulnerability of a decentralized system to compromise. The Nakamoto coefficient is much more susceptible to rapid changes — a few large transactions from large wallet addresses can readily skew it. The Gini index is substantially more stable and considers more information from multiple addresses.

We note that evaluating a blockchain using various metrics is not straightforward. Some challenges are that users can have multiple addresses, custodial wallets represent multiple users, etc. Also, concerning peeling chains and dust wallets, hundreds of millions of wallet addresses must be excluded to reliably interpret the blockchain data for these purposes. Most often, researchers introduce a requirement of a minimum balance, the so-called threshold, or consider only addresses with Top balances.

4. Estimation and comparison of wealth decentralization metrics

This section investigates and provides calculations on basic decentralization metrics introduced in the previous section.

We also compare our results with those given by the authors, who used the same sample size and calculation technique at different times.

4.1. Considered data

For the evaluation, we consider data consisting of 10,000 data points, which are addresses of Top balances with the largest amounts. This number of data points captures the entirety of blockchain addresses, achieving a well-balanced compromise between the size of the data set and the emission of blockchain coins (around 56% for BTC, more than 83% for ETH and even higher for Layer 2 cases). In the blockchain, users may control multiple addresses, making it practically impossible to determine all such addresses and their owners' relations. However, since this is not expected to significantly affect the results, we assume that a user controls a single address.

For Bitcoin, we considered balance data as of 30.06.2023 collected from https://explorer.btc.com/btc/top-address. In the Ethereum case, we had considered https://etherscan.io/accounts and distinguished two points in time: 01.09.2022, while Ethereum was still using a PoW-type consensus algorithm, and 30.06.2023 when Ethereum had already switched to PoS and the community had several months to adapt to this change. With respect to Layer 2, we use the Top 10,000 data extracted on 30.06.2023 from the following sources: https://arbiscan.io, https://optimistic.etherscan.io, https://polygonscan.com.

4.2. Evaluation of Herfindahl–Hirschman index

Herfindahl–Hirschman index (HHI) gives only a little information when used in its classical form (square sum of all data points). Moreover, it has not been found in the literature to be applied to the assessment of blockchain decentralization.

First, the original HHI calculated for all 10,000 samples gives us a value of 19. Note that a blockchain can be considered decentralized if the HHI is less than 2500 since the value above 2500 is an indication of the concentration risk. The obtained value of 19 is impressively low and shows a very good decentralization factor. Nevertheless, we must note that calculations are performed assuming that addresses do not collide (each address is treated individually).

Table 2 shows results for $N = 10{,}000$ and $y = 512, 256, 128, 64, 32, 16, 8, 4, 2$. Note that in our case, x_i is the Top ith address balance divided by the sample sum.

We see that even under very conservative case scenarios and correlation assumptions, we still have very good HHI estimates for ETH and BTC. Results in Table 2 show that for the Bitcoin case, for the Herfindahl Hirschman index to exceed the 2500 threshold, it is groups of 512 required, while in the ETH case, it is already for groups of size 90, we have an index crossing concentration threshold. Note that this analysis is performed for addresses ranked and ordered by size. The findings indicate a significantly higher concentration within the Layer 2 protocols. This distinction from the BTC and ETH blockchains can be attributed to the relative novelty of these protocols, which results in fewer circulating nodes. Another contributing factor to this difference is the token distribution mechanism, which inherently initiates the decentralization process with a degree of concentration.

Table 2
Representation of group-based HHI (values rounded to whole numbers) for groups of rank-ordered 10,000 sample.

Metric	BTC	ETH-PoW	ETH-PoS	ARB	OP	MATIC
HHI_{10000}	19	298	693	2760	1636	1665
$HHI_{10000:2}$	37	442	886	5345	3262	2544
$HHI_{10000:4}$	71	628	1115	5809	4347	3579
$HHI_{10000:8}$	133	882	1416	6533	5596	4673
$HHI_{10000:16}$	238	1116	1662	6999	6265	5677
$HHI_{10000:32}$	403	1452	2026	7362	7357	6331
$HHI_{10000:64}$	655	1953	2550	7663	8518	7034
$HHI_{10000:128}$	1032	2732	3334	7977	9207	8073
$HHI_{10000:256}$	1683	3829	4364	8292	9599	9001
$HHI_{10000:512}$	2602	5074	5426	8623	9757	9385

When data samples were randomly shuffled and the same grouping was used (see Table 3) ETH showed unexpectedly good results. ETH has very close values for higher HHI splits, that is, 1024 and higher. Also, ETH shows a very moderate pace with respect to the smallest groups to the highest. The explanation is that the two largest addresses form the HHI at the higher stage. Still, we have a very good decentralization situation for the other 96 million addresses, which is very positive for ETH and, in some sense, even better than BTC.

Regarding Layer 2, as shown in Table 3, we observe that in the case of ARB, whether we start from the first block or consider the full sample, we surpass the concentration threshold. This phenomenon can be attributed to the dominance of the Top 1 address within the sample, exerting a significant influence on the entire distribution. In the cases of OP and MATIC, the situation is somewhat less extreme. However, the first two addresses collectively account for more than 85% of the sample, contributing substantially to the concentration estimates.

Note that the random shuffle sample (RSS) differs from the rank-ordered sample (ROS). In practice, observable collusion typically does not occur among rank-ordered addresses, but may involve addresses of various sizes. When checking the HHI for RSS, our aim is not to estimate collusion effects, but to identify group sizes that exceed the 2500 concentration threshold. For BTC, in both cases, we have good results and groups such as 512 (ROS) or 2048 (RSS), which is very good for decentralization. This implies that a practical concentration is not achievable, as only addresses in 10 or fewer groups could potentially form a concentration on the blockchain. For Ethereum, we have that (ROS) estimates are somewhat high but still in line with positive decentralization trends. Conversely, the results from RSS are very positive, indicating a strong inclination toward high decentralization when looking from the all-address structure.

4.3. Evaluation of Shanon entropy

In [13] authors employing Shanon entropy demonstrate decentralization for Bitcoin and Ethereum on Top 10,000

Table 3
Representation of group-based HHI (values rounded to whole numbers) for groups of randomized 10,000 sample.

Metric	BTC	ETH-PoW	ETH-PoS	ARB	OP	MATIC
HHI_{10000}	19	298	693	2760	1636	1665
$HHI_{10000:2}$	20	299	694	2760	1637	1666
$HHI_{10000:4}$	22	300	695	2761	1872	1669
$HHI_{10000:8}$	26	303	698	2762	1873	1670
$HHI_{10000:16}$	34	309	704	2776	1874	1671
$HHI_{10000:32}$	49	325	716	2781	1876	1674
$HHI_{10000:64}$	82	362	743	2788	1917	1685
$HHI_{10000:128}$	147	415	787	2808	1938	1815
$HHI_{10000:256}$	272	530	907	2891	1987	1860
$HHI_{10000:512}$	513	756	1182	3167	2076	1991
$HHI_{10000:1024}$	1014	1221	1590	3470	2258	2247
$HHI_{10000:2048}$	2019	2324	2487	6776	2962	3637
$HHI_{10000:4096}$	3766	4102	3393	7491	3892	5352

address balances as of 10.11.2018. The values $S_{\text{BTC}}^{10.11.2018} \approx 11.33$, and $S_{\text{ETH}}^{10.11.2018} \approx 10.38$ were obtained. As decentralization increases with the growth of the entropy value, we see that the wealth of Bitcoin was approximately 9.2% more decentralized than Ethereum. However, both values are much closer to the maximum than to the minimum, so we can conclude from the Shanon entropy that the decentralization of both BTC and ETH is quite high.

In our study, we recalculate the Shanon entropy for the dataset representing the situation more than four years later, that is, on 30.06.2023, also using 10,000 samples: $S_{\text{BTC}}^{30.06.2023} \approx 11.55$, and $S_{\text{ETH}}^{30.06.2023} \approx 8.61$. We see that after more than four years, the gap between Bitcoin and Ethereum is rising, and the decentralization is now 22% better in Bitcoin's favor. Shannon entropy values for ARB, OP, and MATIC (see Table 4) do not exhibit substantial differences but are notably lower compared to those of BTC and ETH.

4.4. Evaluation of Gini index

Using our dataset of 10,000 data points, we calculate the Gini index for Bitcoin (see Fig. 1), resulting in a value of 0.6538. This value closely aligns with the result of 0.65 reported in [20], although there are differences between calculations. In particular, they employ a threshold considering accounts with ≥185 BTC per address, while in our case the threshold is ≥161 BTC. This lower threshold includes more addresses with smaller balances, potentially resulting in slightly higher Gini.

In the same vein, the calculated Gini index for the Top 10,000 Ethereum addresses is 0.8298 (PoW)/0.8457 (PoS), and it is around 30% worse than for the BTC case.

To sum up, Bitcoin and Ethereum's wealth decentralization can be on par with (or even better than) many countries (compared with countries' wealth Gini index, not the income Gini index). For example, according to the [22] global wealth report, many major countries have a Gini wealth index higher than 0.7: Brazil (0.89), China (0.704), the United Kingdom (0.717), and the United States (0.85).

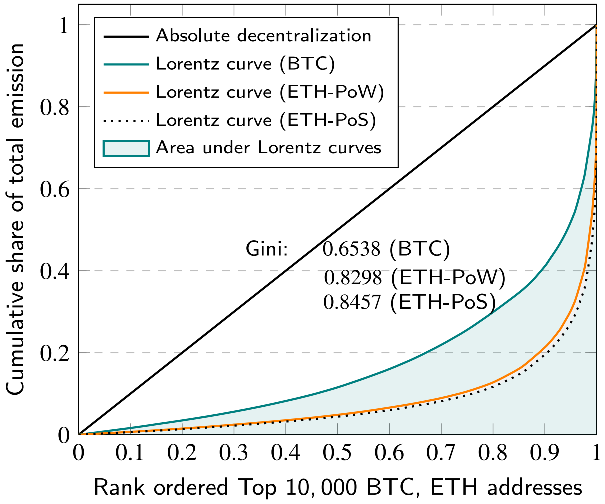

Fig. 1. Lorentz curves and Gini indices for Layer 1 blockchains: BTC, ETH-PoW & ETH-PoS.

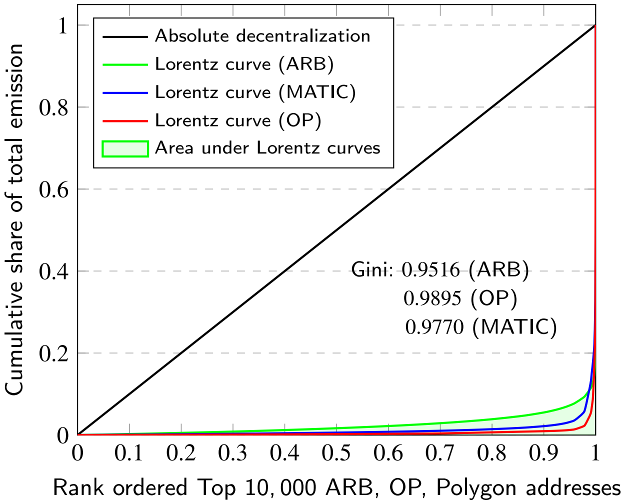

Fig. 2. Lorentz curves and Gini indices for Layer 2 blockchains: ARB, OP & MATIC.

Concerning Layer 2, it is apparent that the indexes are significantly higher compared to Layer 1 blockchains, Bitcoin and Ethereum (see Figs. 1 and 2), or numerous countries. This can be attributed to its relatively recent development and the concentration within the first five top addresses.

4.5. Evaluation of Nakamoto coefficient

For our data collected, the calculation of the Nakamoto coefficient for BTC is equal to 5401 and is a considerable increase compared to [20], where it was 456. For Ethereum PoW, the Nakamoto coefficient is 285, while in [20] – 72. It is a significant improvement in both cases since the higher values of the Nakamoto coefficient indicate better decentralization. Note that both Ethereum and Bitcoin Nakamoto calculations are straightforward, since we can sum up top balances inside our sample and get the result.

Table 4
Summary of obtained wealth decentralization metrics values.

Metric	BTC	ETH-PoW	ETH-PoS	ARB	OP	MATIC
HHI	18.97	297.6	693.8	2760	1636	1665
Shanon ent.	11.55	9.41	8.61	4.02	4.37	5.04
GINI index	0.6538	0.8298	0.8457	0.9516	0.9895	0.9770
Nakamoto c.	5401	285	275	2	2	3

As expected, Arbitrum, Optimism, and Polygon show quite low Nakamoto coefficients (see Table 4) due to their high concentration of Top 3 addresses, which by their sum value takes a dominant part of the overall emission.

4.6. Summary

Table 4 summarizes the values obtained for all four metrics applied to the BTC and ETH, along with Layer 2 protocols ARB, OP, and MATIC, each analyzed with a full dataset of 10,000 samples. Unlike the Gini index (which ranges from 0 to 1), the Nakamoto coefficient defines a single threshold, often considered the number of nodes needed to compromise a blockchain system. Due to this difference, it is possible, though unlikely, to obtain a small Gini index value (high wealth distribution) with a small Nakamoto index (a small number has 51% control) or a high Shanon entropy or high HHI (both showing high concentration). Similarly, it is also possible to get a high Gini index value with a high Nakamoto coefficient, inferring an uneven wealth distribution across the whole population, but a more fair distribution in participants who control 51% of wealth.

5. Conclusions and future research

This work conducted an overview of the current state-of-the-art of various measurements for decentralization in blockchains focusing on wealth decentralization. Using four statistical metrics (Herfindahl–Hirschman index, Shannon entropy, Gini index, and Nakamoto coefficient), we quantified blockchain decentralization based on wealth. Additionally, a novel group-based Herfindahl–Hirschman index was introduced, employing an adjusted sample technique, facilitating a more detailed analysis of the decentralization structure within the blockchain sample. We found that for wealth decentralization, both the Bitcoin and Ethereum blockchains are comparable in all the most popular decentralization metrics, but, in general, Bitcoin shows better results. Moreover, we found that Ethereum's shift to the PoS consensus protocol had little impact on wealth decentralization. However, Layer 2 blockchains, namely Arbitrum, Optimism, and Polygon, exhibit significantly higher centralization across all investigated metrics. Finally, the data collected and all calculations performed are freely available, ensuring complete replicability, reusability, and further development.

In our upcoming research, we plan to explore the implementation of group-based construction with various metrics and analyze time-series data within the wealth decentralization cycle. Moreover, the novel metric can also be applied to Monte Carlo simulations. Finally, we are interested in surveys and analyses about transaction decentralization.

Funding statement

The research work of Ernestas Filatovas was funded by the Research Council of Lithuania (LMTLT), agreement No. S-MIP-21-53.

Data availability

Data underlying this article can be accessed at https://github.com/blockchain-group/blockchain-decentralization-measurement and used under the Creative Commons Attribution license.

CRediT authorship contribution statement

Mindaugas Juodis: Data curation, Conceptualization, Writing – original draft, Methodology. **Ernestas Filatovas:** Data curation, Conceptualization, Writing, Review, Editing, Methodology. **Remigijus Paulavičius:** Supervision, Conceptualization, Writing, Review, Editing, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ernestas Filatovas reports financial support was provided by Research Council of Lithuania (LMTLT).

References

[1] R. Paulavičius, S. Grigaitis, A. Igumenov, E. Filatovas, A decade of blockchain: Review of the current status, challenges, and future directions, Informatica 30 (4) (2019) 729–748.

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

[3] F.A. Sunny, P. Hajek, M. Munk, M.Z. Abedin, M.S. Satu, M.I.A. Efat, M.J. Islam, A systematic review of blockchain applications, IEEE Access 10 (2022) 59155–59177.

[4] P. Treleaven, R. Gendal Brown, D. Yang, Blockchain technology in finance, Computer 50 (9) (2017) 14–17.

[5] E. Balistri, F. Casellato, C. Giannelli, C. Stefanelli, BlockHealth: Blockchain-based secure and peer-to-peer health information sharing with data protection and right to be forgotten, ICT Express 7 (3) (2021) 308–315.

[6] B. Guidi, A. Michienzi, L. Ricci, Assessment of wealth distribution in blockchain online social media, IEEE Trans. Comput. Soc. Syst. (2022) 1–12.

[7] L. Zhang, X. Ma, Y. Liu, SoK: Blockchain decentralization, 2022, arXiv preprint arXiv:2205.04256.

[8] J. Lee, B. Lee, J. Jung, H. Shim, H. Kim, DQ: Two approaches to measure the degree of decentralization of blockchain, ICT Express 7 (3) (2021) 278–282.

[9] E. Filatovas, M. Marcozzi, L. Mostarda, R. Paulavičius, A MCDM-based framework for blockchain consensus protocol selection, Expert Syst. Appl. 204 (2022) 117609.

[10] R. Paulavičius, S. Grigaitis, E. Filatovas, A systematic review and empirical analysis of blockchain simulators, IEEE Access 9 (2021) 38010–38028.

[11] S. Zhang, J.-H. Lee, Analysis of the main consensus protocols of blockchain, ICT Express 6 (2) (2020) 93–97.

[12] Q. Lin, C. Li, X. Zhao, X. Chen, Measuring decentralization in Bitcoin and Ethereum using multiple metrics and granularities, in: 2021 IEEE 37th International Conference on Data Engineering Workshops, ICDEW, IEEE, 2021, pp. 80–87.

[13] K. Wu, B. Peng, H. Xie, Z. Huang, An information entropy method to quantify the degrees of decentralization for blockchain systems, in: 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication, ICEIEC, IEEE, 2019, pp. 1–6.

[14] M. Gupta, P. Gupta, Gini coefficient based wealth distribution in the Bitcoin network: A case study, in: Computing, Analytics and Networks, Springer Singapore, 2018, pp. 192–202.

[15] M. Weymans, Bitcoin wealth is becoming more evenly distributed over time, 2022, URL: https://ecommerceinstitut.de/bitcoin-wealth/.

[16] T. Bol, M. de Vaan, A. van de Rijt, The Matthew effect in science funding, Proc. Natl. Acad. Sci. 115 (19) (2018) 4887–4890.

[17] B. Kusmierz, R. Overko, How centralized is decentralized? Comparison of wealth distribution in coins and tokens, in: 2022 IEEE International Conference on Omni-Layer Intelligent Systems, COINS, IEEE, 2022, pp. 1–6.

[18] C. Campajola, R. Cristodaro, F.M. De Collibus, T. Yan, N. Vallarano, C.J. Tessone, The evolution of centralisation on cryptocurrency platforms, 2022, arXiv preprint arXiv:2206.05081.

[19] L.W. Cong, K. Tang, Y. Wang, X. Zhao, Inclusion and Democratization through Web3 and DeFi? Initial Evidence from the Ethereum Ecosystem, Technical Report, National Bureau of Economic Research, 2023.

[20] B.S. Srinivasan, L. Lee, Quantifying decentralization, 2017, news.earn.com.

[21] M. Juodis, A. Račkauskas, A central limit theorem for self-normalized sums of a linear process, Statist. Probab. Lett. 77 (15) (2007) 1535–1541.

[22] Credit Suisse, Bitcoin Wealth Is Becoming More Evenly Distributed Over Time, Technical Report, 2021, URL: https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html.