Available online at www.sciencedirect.com
ScienceDirect
ICT Express 7 (2021) 278–282
www.elsevier.com/locate/icte

DQ: Two approaches to measure the degree of decentralization of blockchain

Jaeseung Lee[a], Byungheon Lee[b], Jaeyoung Jung[c], Hojun Shim[d], Hwangnam Kim[a,*]

[a] School of Electrical Engineering, Korea University, Seoul 02841, Republic of Korea
[b] Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Republic of Korea
[c] Department of Agricultural Economics and Rural Development, Seoul National University, Seoul 08826, Republic of Korea
[d] Department of New Business Development, Hanwha Life, Seoul 07345, Republic of Korea

Received 23 March 2021; received in revised form 17 June 2021; accepted 2 August 2021
Available online 16 August 2021

Abstract

Blockchain technology innovatively removes the need for intermediaries and establishes a trust-less trust system by peer-to-peer networks and distributed ledger technology. It is still difficult to clarify the specific degree of decentralization of blockchain even though the decentralized characteristic of the system is derived from this system structure. Namely, the criteria for determining whether decentralized are not well established and do not reflect the real world well. So we propose two approaches to measure the degree of decentralization of blockchain system that currently exists: Censorship resistance and geographical diversity. This paper outlines the meaning of each decentralization quotient and explains how those indices can be applied in respect of the user protection.

Keywords: Blockchain; Decentralization quotient; Distributed network; Privacy

1. Introduction

What is Decentralization? Decentralization is one of the words most frequently used in the field of blockchain, and even it is often viewed as a kind of blockchain itself. However it is also the word most difficult to define [1].

In 2008, Satoshi Nakamoto, a pseudonymous developer, proposed a cryptocurrency called Bitcoin, which uses block concept and chaining each block to solve the trust problem [2], a challenge in a distributed network with decentralized ledgers. After Bitcoin was introduced, people realized that what Bitcoin uses is the blockchain technology, and considered that the blockchain technology can remove the need for intermediaries and facilitate digital trust by recording the essential information in an open network [3–6].

Decentralization is the most important feature of the blockchain system. This word contains complex concepts such as economy, technology, and philosophy. There are therefore various ways to define it [7]. The most popular way is to define the system by contrasting it to the centralization. The centralized system is the system that has only one central point that has the power to decide the system's status [8]. From any so, the decentralized system can be said to be the system that has no point to decide the system's status at one's own will, which means that peer-to-peer consensus is required. There is no need to trust anybody, but users can trust the system, which is called a trust-less trust system. For the decentralization system, users do not need to trust anyone but trust the system. In addition, it means that the measurement of decentralization is related to the measurement of how reliable the system itself is.

Also, decentralization can be interpreted differently in terms of the utility. We acknowledge that a decentralized system can offer user privacy and autonomous control of the system, which are essential features of the consumer protection [9]. For this reason, a decentralized system can be said to be an excellent system for customer protection.

However, one question is left. How can we judge if this platform or network is decentralized and reliable? For those who use the network, it is more important to determine whether this network is decentralized or not than the abstract concept of decentralization.

* Corresponding author.
E-mail addresses: tkxkd0159@korea.ac.kr (J. Lee), topkids123@snu.ac.kr (B. Lee), gadget0316@snu.ac.kr (J. Jung), joy.shim@hanwha.com (H. Shim), hnkim@korea.ac.kr (H. Kim).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2021.08.008

Table 1
The symbols in equations.

Symbol	Description	Equation
DQ_{cr}	Decentralization Index in terms of censorship resistance	
DQ_{geo}	Decentralization Index in terms of geographical diversity	
n	Total number of nodes participating in the network	(1), (2)
H_i	The hash power of ith node	(1)
$DQ_{cr,pow}$	DQ_{cr} in a PoW system	(1)
S_i	The share of the ith nod	(2)
$DQ_{cr,pos}$	DQ_{cr} in a PoS system.	(2)
N_t	Total number of countries	(3)
N	The number of countries where nodes in the blockchain are located	(3)
n_i	The number of nodes belonging to ith country	(3)
$DQ_{geo,target}$	DQ_{geo} for the selected platform	(3), (4)
$DQ_{geo,excl}$	DQ_{geo} when all nodes are located in one country	(4)
$DQ_{cr,equal}$	DQ_{geo} when all nodes are equally distributed across the entire countries	(4)

In past, consensus protocols and fault tolerance are used for assessing the degree of decentralization [10]. However, there is a limit because it is only the prior-assessment. Even though systems are designed with good protocols, they can become centralized and this would make the network hard to guarantee consumer protection, depending on the direction of use. To solve this problem, we need a quantitative approach to establish post-evaluation criteria.

This paper is organized as follows. Section 2 describes our proposed design of Decentralization Quotient (DQ) in two sides and Section 3 shows analysis results using this designed quotient. Section 4 concludes the paper

2. Design of decentralization quotient

As mentioned in the introduction, there can be a wide variety of views that measure the degree of decentralization. Therefore, our scope has been narrowed and refined to the point that the system should not produce a result against the user's will if a non-malicious user uses the blockchain system. Hence, we designed DQs and then divided them by censorship resistance and geographic diversity from a consumer perspective to define the final values.

2.1. Censorship resistance

In this subsection, we define the degree of decentralization associated with censorship resistance. One of the serious problems with existing centralized networks is that our data can be arbitrarily censored by third parties like central institutions. A recent notable example is Facebook's self-censorship and filtering of information related to COVID-19 [11]. In this case, the individual loses the right to judge the correctness of the information and make his (or her) own choices.

To address these issues, we discuss the decentralized index in terms of censorship resistance and apply it to each platform to assess which one is a more decentralized network and allow users to determine their entry to the network based on this metric. Assuming extreme decentralization, it is a unanimous opinion of the individual's overall opinion on the network loading of that data. In addition, extreme centralization is the exclusive censorship of an institution. Therefore, the decentralized quotient can be derived by applying block generation probability, which is the process by which network participants load data into the blockchain, to the Gini coefficient to adapt to each consensus algorithm.

First of all, if we look at the Proof of Work (PoW) [2], the block creation privilege is taken by the node that solved the hash puzzle first. Therefore, assuming that all nodes in the network have the same hash power, we can eventually assume that everyone's opinions in the network are probabilistic equally reflected and converge decentralized. However, since this is an algorithm that prioritizes liveness over safety, it requires a lot of computing power to solve hash puzzles to compensate for it. Therefore, it is essential to provide incentives for continuous network operation. As a result, this led to the emergence of mining pools with concentrated hash power. The higher the hash power, the higher the probability of obtaining block generation authority, which can claim the possibility that blockchain networks will be censored by individuals or organizations with higher hash power. Therefore, we can stack the hash rate for each node in ascending cumulative fashion to yield the index. The higher the index, the closer the network is to the decentralized and it ranges from 0 to 1. This can be expressed as an equation as follows,

$$DQ_{cr,pow} = 1 - 2 \times \left(\frac{1}{2} - \frac{1}{n} \sum_{i=1}^{n} H_i \right) \tag{1}$$

where all variables are shown in Table 1.

In other case, the Proof of Stake (PoS) [12] family, there are various variations, but the basis is that the probability of block creation increases proportionally to the stake. Therefore, if all nodes in the network have the same stake, they converge to full decentralization. However, under the capitalist system, a completely equal distribution situation is impossible. Therefore, we can evaluate how close the network is to the decentralized by listing each node's coin holdings in an ascending cumulative manner to produce the quotient. This can be expressed by the following equation:

$$DQ_{cr,pos} = 1 - 2 \times \left(\frac{1}{2} - \frac{1}{n} \sum_{i=1}^{n} S_i \right) \tag{2}$$

where all variables are shown in Table 1. The Delegated Proof of Stake (DPoS) [13] created to increase Transactions Per Second (TPS) is similar to parliamentary democracy. In

Table 2
Censorship-resistant DQ of Bitcoin, Ethereum.

Types of DQ	Bitcoin	Ethereum
DQ_{cr}	0.0110	0.0108

other words, this algorithm was not considered when we designed our decentralized index because it was intended to weaken decentralization and increase scalability in the blockchain trilemma. The Practical Byzantine Fault Tolerant (PBFT) [14] family was not considered for similar reasons.

2.2. Geographical diversity

In this subsection, we define the degree of decentralization in relation to geographical diversity. Geographical diversity is as an indicator of how nodes are distributed in various regions. It is very important to maintain decentralization because if most nodes of the network are located in the same country, the availability of that network is not guaranteed and can be dependent on the country's policies. Thus, we can calculate the relative index accordingly by setting it as a reference value if we assume that the number of nodes present in each country is the same.

We modeled the decentralized index in terms of geographical diversity by applying a variance and rescaling output. Also, the auxiliary coefficient was designed to give weights accordingly as nodes are distributed in various countries. This index can be expressed:

$$DQ_{geo,target} = \left(2 - \frac{log_{N+1}N_t - log_{N_t+1}N_t}{log_2 N_t - log_{N_t+1}N_t}\right) \times \sqrt{\frac{\sum_{i=1}^{N_t}(n_i - \mu)^2}{N_t}}, \tag{3}$$

$$DQ_{geo} = \frac{(DQ_{geo,excl} - DQ_{geo,target}) - DQ_{geo,equal}}{DQ_{geo,excl} - DQ_{geo,equal}}. \tag{4}$$

where μ is the total number of nodes divided by total number of countries, and other variables are shown in Table 1. DQ_{geo} is defined by normalizing $DQ_{geo,target}$ between 0 and 1. The higher the DQ_{geo}, the more distributed it is.

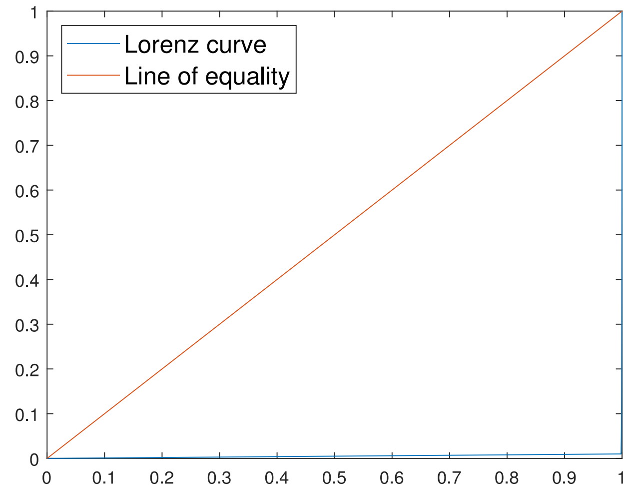
3. Analysis

Based on the quotients we designed earlier, we can measure the degree of decentralization of several blockchain platforms. We applied it to Bitcoin and Ethereum, which are the most platforms.

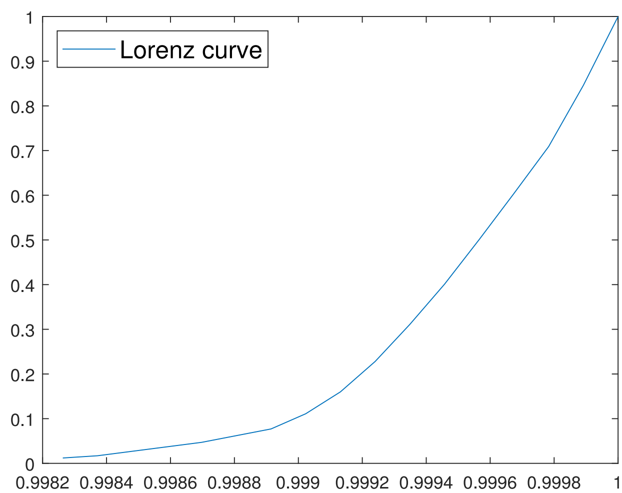
3.1. Censorship resistance

Bitcoin and Ethereum's compensation system is a winner-take-all structure. Thus, this led to a situation in which a hash power of 99 percent was concentrated in 17 mining pools in Bitcoin and a hash power concentrated in 16 mining pools in Ethereum [15,16].

In Figs. 1 and 2, we represented the cumulative share of nodes from the lowest to the highest hash rate of the two platforms. In addition, the calculated DQ values based on this can also be seen in Table 2. We can see that the degree of decentralization of blockchain was measured very low due to the concentration of hash power



(a) Graphical representation for the whole nodes



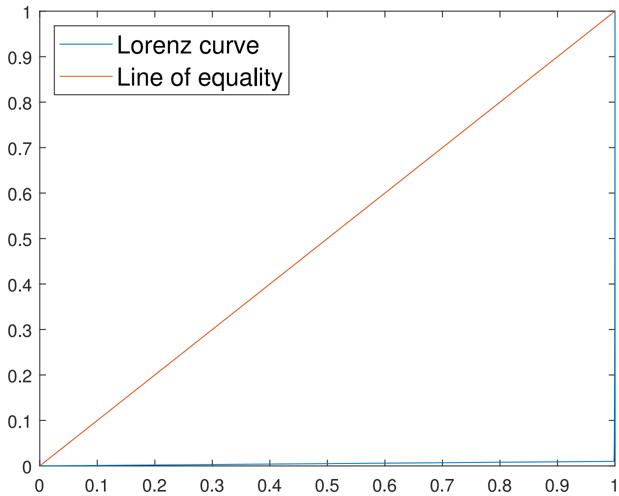
(b) Magnified graph for the top nodes

Fig. 1. Cumulative share of BTC node from the lowest to the highest hash rate.

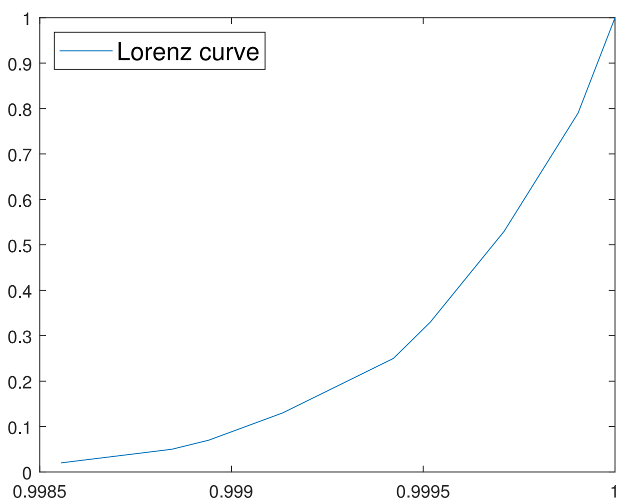
3.2. Geographical diversity

To compute DQ_{geo}, we set N_t to 195 according to official UN standards [17]. Table 3 shows the calculated DQ_{geo} values based on [18,19]. Bitcoin is a bit more decentralized than Ethereum in this perspective.

When looking at decentralization in terms of geographical diversity, we found that it measures higher than in terms of censorship resistance. Actually, there are 83 countries for Bitcoin and 108 countries for Ethereum that include one or more nodes. Fig. 3 is a picture visualized with a choropleth map so



(a) Graphical representation for the whole nodes



(b) Magnified graph for the top nodes

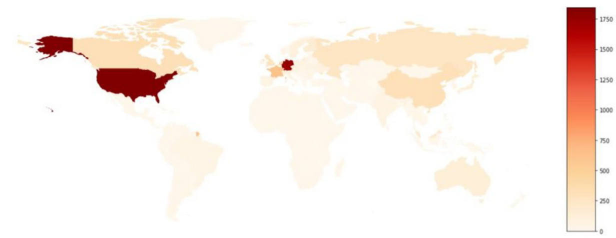
Fig. 2. Cumulative share of ETH node from the lowest to the highest hash rate.

that the degree of decentralization can be easily grasped. There are some countries where nodes are concentrated, but we can see that the other nodes are quite overall distributed around the world.

4. Discussion

We proposed two kinds of decentralization indices and applied them to popularized blockchain platforms. Our most innovative contribution is that we have attempted a quantitative approach to the concept of decentralization, an abstract yet qualitative domain. For example, Joe Lubin's DTPS [20] is a similar attempt to define decentralization, but it is just simply multiplied by the qualitative constant they set themselves.

The strength of our proposed method is that it is numerically easy to calculate based on data that is not hard to obtain.



(a) Distribution of Bitcoin full Node

(b) Distribution of Ethereum full Node

Fig. 3. Geographic distribution of full nodes on a blockchain platform.

Table 3
Geographical DQ of Bitcoin, Ethereum.

Types of DQ	Bitcoin	Ethereum
$DQ_{geo,target}$	377.7091	550.9375
$DQ_{geo,excl}$	561.7068	788.1325
$DQ_{geo,equal}$	0	
DQ_{geo}	0.3276	0.3010

Since PoW and PoS are consensus methods currently borrowed from most blockchain platforms, our proposed method allows users to evaluate the degree of decentralization and can use the method to select a platform.

The limitation of our proposed approach is that this is all postmortem analysis. Although blockchain network designers can use DQ to post-analyze how decentralized that network is, it is difficult to use as a preliminary analysis when starting the design. Furthermore, in geographical methods, the nationality of the country in which the node operates may not match the nationality of the owner of this node, which makes it ambiguous which region belong. However, we think we can reach an agreement about this point within the community as the blockchain ecosystem matures in the future.

5. Conclusion

In this paper, we proposed DQ as two quantitative approaches. Since each individual tries to maximize his or her own interests, we have identified a dilemma in which decentralized systems built for decentralized purposes converge towards centralization. In other words, this dilemma seems difficult to solve until now unless each participant in the network tries to maintain a decentralized system with a high level of consciousness. Even in this difficult situation, various blockchain platforms to be developed can apply DQ during service operation and receive feedback from these values to build a more decentralized structure. These will serve as good milestones to preserve the royalty of the blockchain. In addition, consumers can identify which platform is more decentralized and choose optionally through our proposed DQ. In conclusion, it will eliminate the less decentralized blockchain platform and eventually form a virtuous cycle of the entire blockchain ecosystem.

CRediT authorship contribution statement

Jaeseung Lee: Writing – original draft, Conceptualization, Methodology, Validation, Software, Data curation, Visualization. **Byungheon Lee:** Conceptualization, Investigation. **Jaeyoung Jung:** Data curation, Visualization. **Hojun Shim:** Investigation. **Hwangnam Kim:** Funding acquisition, Supervision, Validation, Resources, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the MSIT (Ministry of Science and ICT), Republic of Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-01835) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation), Republic of Korea, and the Human Resources Program in Energy Technology of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20204010600220).

References

[1] V. Buterin, The meaning of decentralization, 2017, URL https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274.
[2] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, Tech. Rep., 2008.
[3] V. Karthika, S. Jaganathan, A quick synopsis of blockchain technology, Int. J. Blockchains Cryptocurrencies 1 (1) (2019) 54–66.
[4] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, H. Wang, Blockchain challenges and opportunities: A survey, Int. J. Web Grid Serv. 14 (4) (2018) 352–375.
[5] S. Park, S. Oh, H. Kim, Performance analysis of DAG-based cryptocurrency, in: 2019 IEEE International Conference on Communications Workshops, ICC Workshops, 2019, pp. 1–6, http://dx.doi.org/10.1109/ICCW.2019.8756973.
[6] S. Park, H. Kim, DAG-based distributed ledger for low-latency smart grid network, Energies 12 (18) (2019) 3570.
[7] S.P. Gochhayat, S. Shetty, R. Mukkamala, P. Foytik, G.A. Kamhoua, L. Njilla, Measuring decentrality in blockchain based systems, IEEE Access 8 (2020) 178372–178390.
[8] R. Rezazadeh, The concept of centralization and decentralization: An analysis and evaluation, Int. Rev. Adm. Sci. 27 (4) (1961) 425–430.
[9] C. Troncoso, M. Isaakidis, G. Danezis, H. Halpin, Systematizing decentralization and privacy: Lessons from 15 years of research and deployments, Proc. Priv. Enhanc. Technol. 2017 (4) (2017) 404–426.
[10] A.H. Lone, R.N. Mir, Consensus protocols as a model of trust in blockchains, Int. J. Blockchains Cryptocurrencies 1 (1) (2019) 7–21.
[11] J. Koetsier, Facebook deleting coronavirus posts, leading to charges of censorship, 2020, URL https://www.forbes.com/sites/johnkoetsier/2020/03/17/facebook-deleting-coronavirus-posts-leading-to-charges-of-censorship/?sh=4c57a21e5962.
[12] P. Vasin, BlackCoin's proof-of-stake protocol v2, 2014, URL https://blackcoin.org/blackcoin-pos-protocol-v2-whitepaper.pdf.
[13] D. Larimer, Delegated proof of stake (DPOS), 2018, URL https://how.bitshares.works/en/master/technology/dpos.html#id2.
[14] M. Castro, B. Liskov, Practical Byzantine fault tolerance and proactive recovery, ACM Trans. Comput. Syst. (TOCS) 20 (4) (2002) 398–461.
[15] Bitcoin hashrate distribution, 2021, URL https://blockchair.com/bitcoin/charts/hashrate-distribution. (Accessed 21 March 2021).
[16] Ethereum hashrate distribution, 2021, URL https://blockchair.com/ethereum/charts/hashrate-distribution. (Accessed 21 March 2021).
[17] United Nations Statistics Division, Methodology–standard country or area codes for statistical use (M49), 2019.
[18] Bitcoin nodes distribution, 2021, URL https://bitnodes.io/. (Accessed 21 March 2021).
[19] Ethereum mainnet statistics, 2021, URL https://www.ethernodes.org/. (Accessed 21 March 2021).
[20] How we get to a decentralized web, 2021, URL https://consensys.net/blog/events-and-conferences/joe-lubins-full-speech-from-devcon-5-how-we-get-to-a-decentralized-world-wide-web/. (Accessed 13 June 2021).