Method for Determining the Level of Centralization in BTC Lightning Nodes: A Centrality Analysis of the Lightning Network

Laura Atmanavičiūtė, Tomas Vanagas, Saulius Masteika

Vilnius University, Kaunas Faculty, Institute of Social Sciences
and Applied Informatics
Muitinės str. 8, LT-44280 Kaunas, Lithuania
laura.atmanaviciute@knf.vu.lt, tomas.vanagas@knf.vu.lt,
saulius.masteika@knf.vu.lt

Abstract. This study explores the Bitcoin Lightning Network (BLN), a Layer 2 solution for faster and cheaper transactions. Concerns about centralization have emerged due to the increasing concentration of power among specific nodes, named "hubs." Statistical measures like the Gini coefficient reveal a trend towards centralization, challenging the LN's decentralized nature. Consequently, further research is necessary to address this issue and ensure the integrity of the LN architecture. This paper aims to establish a method for determining the level of centralization within the BLN by applying centrality analysis techniques. Study revealed that over the six-year period Gini coefficient increased from 0.87 to 0.955, indicating significant inequality and apparent centralization of BLN nodes.

Keywords: Bitcoin, Lightning Network, Centralization, Data processing

1 Introduction

The Bitcoin Lightning Network (BLN) is a promising Layer 2 solution built on Bitcoin (BTC) that aims to make transactions faster and cheaper. When it was first introduced, LN fees were expected to be much lower than standard BTC transactions [1]. It was meant to be a way for users to send money to each other directly, without loading the transaction data to the whole BTC network [2]. Two users can agree to establish a direct channel by creating a multi-signed transaction on the blockchain [3]. When the channel is closed, only the final balance needs to be settled on-chain as a single transaction. Consequently, the system becomes capable of accommodating a significantly greater volume of payments [4].

DOI: https://doi.org/10.15388/LMITT.2024.1

However, a potential centralization issue has emerged. There is no robust answer as to whether the current distribution within the LN indicates a trend towards centralization, with a select few nodes maintaining a disproportionate share of the network's total channel capacity [5]. Can powerful, well-funded nodes, acting as hubs with extensive payment channels that process a large volume of transactions, gain undue influence? This dominance by a few hubs might lead to a more centralized system, contradicting the decentralized nature of BTC itself [6]. Based on this, the following hypothesis is proposed:

Hypothesis. An unequal distribution of channels among nodes within the BLN may be a contributing factor to a decrease in the network's overall decentralization.

To address the question of centralization, it's important to delve into the methods and ways, using specific coefficients, to measure the level of centralization. The challenge here is significant, as the data must first be extracted from the blockchain, categorized, and linked to obtain variables that can be appropriately integrated into methodology.

The aim of this study is to establish a method for determining the level of centralization within the BLN by applying centrality analysis techniques. Critical tasks towards achieving this goal include:

- Developing a comprehensive method for calculating centralization.
- Extracting, gathering, linking, and storing data from the BTC blockchain Layer 1 (L1) and Layer 2 (L2).
- Conducting experimental calculations and providing visual representation of the results.

This paper proposes a method for determining the centralization level of BLN nodes. It combines Gini coefficient to quantify the centralization and the Lorenz curve to visually present the results. The structure of the paper is as follows: The first part of the paper introduces the topic, outlining the research focus. The second part delves into the background of the method, analysing centrality aspects and coefficients relevant to assessing the BLN's centralization level. The third part consists of detailed explanation of data extraction and linking, including a proposed data retrieval and storage scheme. The fourth section presents the experimental setup and its results of the experiment. Finally, the last section presents the conclusions of the study.

2 Background of the method

To objectively assess centralization within the LN, it's important to explore different aspects of centrality. There are five main aspects that can be considered when assessing the level of centralization of the BLN – degree, weighted degree, betweenness, eigenvector and closeness centrality. **Degree centrality** evaluates the number of channels a node has with other nodes – it identifies highly connected nodes but not the significance of those connections [2]. This limitation can be addressed by considering **weighted degree centrality**, which incorporates channel capacity into calculations [7]. **Eigenvector centrality** is variant of degree centrality and measures a node's influence in the network based. In simpler terms, degree centrality counts nodes and eigenvector centrality measures the influence of a node [8]. Meanwhile, **closeness** and **betweenness centrality** consider the distance between nodes when trying to find the shortest connection between them [7]. **Closeness centrality** is used for calculating how close a node is to all other nodes in the network [2]. It helps understand the efficiency of network, but it might not directly address the concern of the centralization if all nodes have similar closeness. Another approach is **betweenness centrality**, which measures how frequently a node is on the shortest path connecting other nodes – in the context of the LN, it indicates a node's significance for routing payment [9, 10]. For this paper, specifically weighted degree centrality aspect is employed, because it counts not only the number of channels a node has, but also considers the capacity of each connection.

After choosing to assess BLN centralization through weighted degree centrality aspect, it is important to delve into coefficients which can be employed to quantify this centrality measure. One of the most common coefficients is Gini coefficient, which measures the inequality of channel distribution within the LN. Higher Gini coefficient suggests a greater centralization – it is known as a strong indicator of overall network centralization, especially if utilized with other measures [2, 11]. Another well-known coefficient is Herfindahl–Hirschman index (HHI), which can also be used as a method when determining network's centralization. HHI is a traditional metric used to assess market's concentration and is often used to measure market efficiency [12].

Gini coefficient is more insightful than the HHI as it directly measures inequality in capacity distribution and shows how influential few nodes in the network might be. Meanwhile, HHI is less sensitive to imbalances, which is one of the main issues of network centralization. Additionally, the Gini coefficient acts like a score between 0 and 1, where 0 signifies everyone having an equal share of the resource and 1 represents a scenario where one individual has everything [2, 9], which makes this measure easy to interpret, while HHI doesn't have a straightforward interpretation. A lower Gini coefficient points towards a network with a more balanced distribution (decentralized), while a higher value suggests an uneven spread (centralized) of the resource being analysed. It can be measured using the following formula:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2N^2 \bar{x}}$$

N is used to represent a total number of nodes, x_i and x_j represent capacity of nodes and $\bar{x}$ is an average capacity across all nodes.

The Gini coefficient represents the difference between the line representing perfect equality and the actual distribution depicted by the Lorenz curve. It's calculated by subtracting the area below the Lorenz curve from the area below the line of perfect equality, and then dividing this result by the total area under the line of perfect equality [10]. The Lorenz curve visually illustrates how channels are distributed among nodes based on the size of their channel capacity. It compares this distribution to a perfectly equal scenario represented by a line at a 45-degree angle, known as the line of equality. The area between this line and the Lorenz curve is utilized to calculate the Gini index [13]. The analysis of the Gini coefficient and Lorenz curve for channel capacity distribution is leveraged to determine the level of centralization of BTC lightning nodes.

Existing research [2, 6, 9, 10, 11] utilizes the Gini coefficient revealing a growing trend of uneven distribution of channel capacity within the LN. While these studies provide valuable insights, there still are some limitations – such as detailed descriptions of data processing, which limits replication of study. This paper addresses these gaps by proposing a data collection and linking scheme, along with different timestamps than compared to related research.

3 Data processing for proposed method

To research the centralization level of the BLN, data is gathered from 3 primary sources – LN Research [14], Bitcoin Core [15] and Electrum Node [16]. LN Research investigates the LN data, while Bitcoin Core validates transactions and confirms blocks. Electrum Nodes act as intermediaries –

they don't store entire blockchain and are relevant for this paper to collect spending transaction information. Data retrieval and storage is presented in Figure 1.

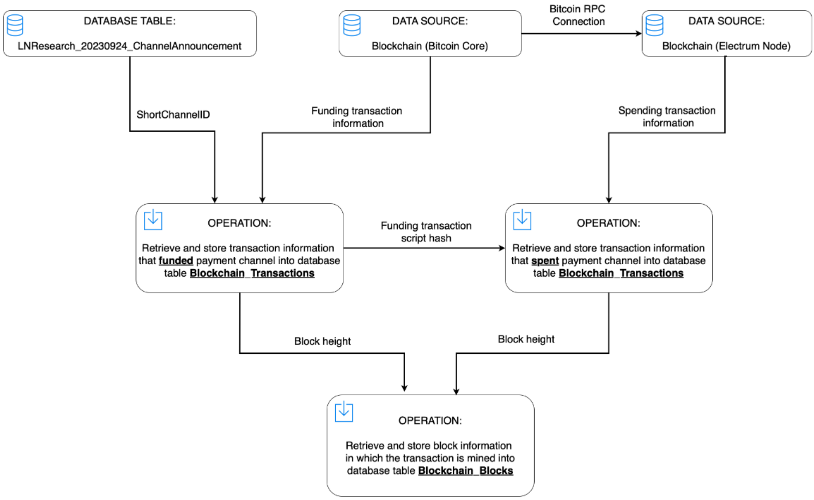

Figure 1. Data retrieval and storage

First, L2 data from LN Research is gathered. This research provides data of LN transactions by exchanging information through gossip protocol. For this study, relevant are 'Channel Announcement' messages. They provide information about the creation of a new payment channel in the LN – such as unique identifier (ID) of the channel ('ShortChannelID'), and nodes, which participates in this channel IDs.

While LN Research provides all relevant information about LN transactions, it lacks data about the L1 in which these transactions take part. To collect relevant data from L1, operating system 'MyNodeBTC' is utilized which has Bitcoin Core and Electrum Server installed. It helps to synchronize a Bitcoin's full node and an Electrum node for the transaction indexing. The BTC blockchain keeps a transaction dataset, which includes all transactions that have transpired on the BTC network. This database records all critical data of each transaction, including **timestamps** – the precise date and time at which BTC was locked within a transaction, **transaction amounts** – the quantity of BTC that was locked, and **channel status** – an indicator which indicated whether the transaction was utilized for opening a LN channel (spent) or is still unused (unspent). Transactions which were identified as spent, were further investigated by assigning the specific block height within the blockchain, where the spending transaction has occurred.

Bitcoin Core software was configured to index all transactions in the blockchain by enabling 'txindex' flag in the configuration, while Electrum Server has various indexes to support Electrum Light Wallets, including transaction script hash index. Using Bitcoin Core and Electrum Server node it is possible to efficiently retrieve BLN payment channel funding and spending transaction data from the blockchain without fully scanning entire blockchain, as Electrum Server can leverage its prebuilt indexes.

3.1. Data linking

The data is linked by connecting data collected by LN Research to the relevant blockchain transactions which opened the channels. Blockchain data retrieval process starts by iterating through every 'Channel Announcement' message in the LN research dataset and retrieving transaction which opened BLN payment channel from Bitcoin Core node. The link is facilitated by the 'ShortChannelID', which consists of the block height, the transaction index within the block, and the transaction output index.

Bitcoin Core's node is requested to return transaction based on block height, transaction index in the block and transaction output index in the transaction ('ShortChannelID') and inserting returned data to the database table 'Blockchain_Transactions'. After inserting the channel funding transaction details, it is required to find when transaction output in question was spent. For this part of the process, Electrum Server Node can return this data by using RPC's 'blockchain.scripthash.get_history' function.

At the end of the blockchain data retrieval process there should be same number of records in both 'LNResearch_20230924_ChannelAnnouncement' and 'Blockchain_Transactions' database tables. This verifies that data from both sources has been successfully linked.

After the transaction details have been retrieved, it is necessary to retrieve data about the block in which the transaction has occurred. Information about the blockchain block contains a timestamp which shows when the transactions in question have been mined. Data about the blockchain blocks are stored in a different database table 'Blockchain_Blocks'.

The full dataset of both L1 and L2 allows researchers to have a full picture of the BLN by joining database tables together and filtering the dataset to any moment of time of BLN existence and applying calculations.

4 Experimental setup and results

To capture the changes of the LN, six data snapshots were selected for the experiment, taken on June 1st of each year, starting from 2018 – the year when LN was presented – and ending with 2023. This approach not only allows tracking changes and identifying trends in the LN structure over time, but it also addresses the challenge of the LN's rapidly evolving data. The experiment analyses the distribution of channel capacity across the BLN nodes. The experiment utilizes a dataset containing 495,755 channel announcement records. A node is considered existing if it has at least one channel open during that timeframe.

The results reveal a concerning trend towards centralization. Figure 2 presents Lorenz curves for the BLN nodes on weighted degree centrality aspect captured at six specific timestamps. It was created by retrieving data from the intermediate database table at specific moments of time. After this, all the nodes were sorted in ascending order based on the BTC amount and then cumulative percentages of the whole network were calculated in 1% granularity to calculate Lorenz curve. Figure 2 shows that over the time Lorenz curve is progressively moving further away from the perfect equality line across the six timestamps, which means that inequality between BLN nodes is increasing. This trend also aligns with the results of calculated Gini coefficient, which increased from 0.87 in 2018 to 0.955 in 2023, with an average of 0.926. This indicates a great inequality, suggesting a concentration of channel capacity among specific groups of nodes.

Furthermore, Figure 3 visually represents the results of Gini coefficients of weighted degree centrality for the BLN nodes and reinforces the observation. In the two years of LN, inequality for BLN nodes increased significantly. Starting with 0.87 Gini coefficient at the start of LN and reaching 0.936 in 2020, there is a huge decrease in network's decentralization.

The experimental results proves that the BLN is exhibiting tendencies towards the centralization, especially in channel capacity distribution. The research confirms the initial hypothesis – the Gini coefficient rose significantly over time and indicated unequal distribution of channel capacity among nodes. This distribution, which also was visualized by Lorenz curves deviating further from perfect equality, aligns with the hypothesis that uneven channel distribution is a contributing factor to a decrease in the network's overall decentralization.

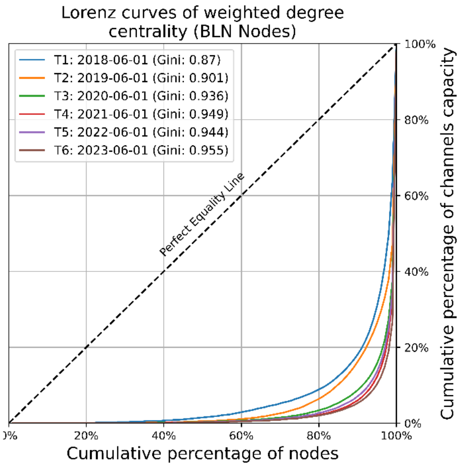

Figure 2. Lorenz curves of weighted degree centrality for Bitcoin Lightning Network nodes

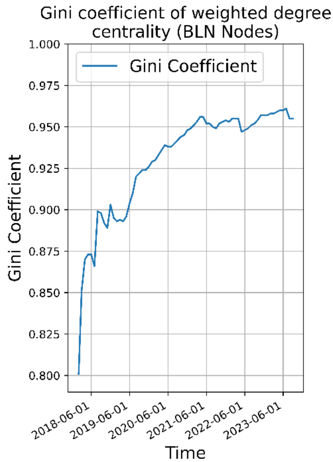

Figure 3. Gini coefficient of weighted degree centrality for Bitcoin Lightning Network nodes

5 Conclusions

This study established a method for assessing the level of centralization in BTC lightning nodes using the Gini coefficient and Lorenz curve. Building upon prior research that identified a concerning trend of uneven channel capacity distribution, this paper expands that research with a detailed data collection and linking scheme, along with a different range of timestamps. Gini coefficient was proposed as a reliable method because it measures inequality and shows how influential nodes in the network can be. Along with the Gini coefficient, the Lorenz curve depicted channel capacity distribution. This combined approach enabled comprehensive analysis and trend identification.

Data from both the BTC blockchain L1 and L2 was extracted, gathered, linked, and stored successfully by connecting LN Research data and queries to Bitcoin Core and Electrum Server nodes. This dataset ensured a complete picture of the BLN for further calculations assessing the centralization of the network.

The experimental calculations, using the Gini coefficient and Lorenz curves for the six timestamps, confirmed the initial hypothesis. The research revealed an increase in the Gini coefficient – from 0.87 in 2018 to 0.955 in 2023, signifying a growing inequality in channel capacity distribution among nodes. This trend is also highlighted by Lorenz curves, which is progressively moving further away from perfect equality. The results of the experiments suggest a concerning shift towards centralization within the BLN – especially regarding channel capacity distribution among nodes.

References

[1] Poon J, Dryja T. (2016). The Bitcoin Lightning Network: Scalable Off-Chain. DRAFT Version 0.5.9.2. 2016. URL: https://lightning.network/lightning-network-paper.pdf.

[2] Lin, J.-H., Primicerio, K., Squartini, T., Decker, C., & Tessone, C. J. (2020). Lightning network: a second path towards centralisation of the Bitcoin economy. New Journal of Physics, 22(8), 83022. doi:10.1088/1367-2630/aba062

[3] Martinazzi, S., Flori, A. (2020). The evolving topology of the Lightning Network: Centralization, efficiency, robustness, synchronization, and anonymity. PloS One, 15(1), e0225966–e0225966. doi:10.1371/journal.pone.0225966

[4] Divakaruni, A., Zimmerman, P. (2022). The Lightning Network: Turning Bitcoin into money. Finance Research Letters 52.

[5] Carotti, A., Sguanci, C., & Sidiropoulos, A. (2023). Rational Economic Behaviours in the Bitcoin Lightning Network. doi:10.48550/arxiv.2312.16496

[6] Masteika, S., Rebždys, E., Driaunys, K., Šapkauskienė, A., Mačerinskienė, A., & Krampas, E. (2023). Bitcoin double-spending risk and countermeasures at physical retail locations. International Journal of Information Management, 102727. doi:10.1016/j.ijinfomgt.2023.102727

[7] Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks, 32(3), 245–251. doi:10.1016/j.socnet.2010.03.006

[8] Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. The Journal of Mathematical Sociology, 2(1), 113–120. doi:10.1080/0022250X.1972.9989806

[9] Mahdizadeh, M. S., Bahrak, B., & Sayad Haghighi, M. (2023). Decentralizing the lightning network: a score-based recommendation strategy for the autopilot system. Applied Network Science, 8(1), 73–33. doi:10.1007/s41109-023-00602-2

[10] Zabka, P., Foerster, K.-T., Decker, C., & Schmid, S. (2022). Short Paper: A Centrality Analysis of the Lightning Network. In Financial Cryptography and Data Security (pp. 374–385). Cham: Springer International Publishing. doi:10.1007/978-3-031-18283-9_18

[11] Lin, J.-H., Marchese, E., Tessone, C. J., & Squartini, T. (2022). The weighted Bitcoin Lightning Network. Chaos, Solitons and Fractals, 164, 112620. https://doi.org/10.1016/j.chaos.2022.112620

[12] Cheng, L., Zhu, F., Liu, H., & Miao, C. (2021). On Decentralization of Bitcoin: An Asset Perspective. doi:10.48550/arxiv.2105.07646

[13] Juodis, M., Filatovas, E., & Paulavičius, R. (2024). Overview and empirical analysis of wealth decentralization in blockchain networks. ICT Express, 00(00), 1–7. doi:10.1016/j.icte.2024.02.002

[14] Decker, C. (2023). Lightning Network Gossip. URL: https://github.com/lnresearch/topology

[15] Bitcoin Core Developers. (2024). Bitcoin Core. URL: https://bitcoincore.org/

[16] T. Voegtlin. (2011). Electrum Bitcoin Wallet. URL: https:// electrum.org/