IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 August 2025, accepted 22 September 2025, date of publication 24 September 2025, date of current version 1 October 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3614085

RESEARCH ARTICLE

Quantitative Analysis of Centralization in the Bitcoin Lightning Network Through Centrality Metrics

LAURA ATMANAVIČIŪTĖ, (Member, IEEE), TOMAS VANAGAS, AND SAULIUS MASTEIKA
Institute of Social Sciences and Applied Informatics, Kaunas Faculty, Vilnius University, 44280 Kaunas, Lithuania

Corresponding author: Laura Atmanavičiūtė (laura.atmanaviciute@knf.vu.lt)

ABSTRACT The Bitcoin Lightning Network, a Layer-2 more scalable solution for the Bitcoin blockchain, has emerged to address the scalability challenges faced by the Bitcoin network. However, the centralization of the Lightning Network has been a growing concern, as the concentration of highly active nodes within it could compromise the decentralized nature of the Bitcoin ecosystem. In this research paper, we conduct a quantitative analysis of centralization in the Lightning Network using various centrality metrics, such as Gini coefficient, Nakamoto coefficient, Herfindahl-Hirschman Index, Theil Index and Shannon entropy. The proposed methodology includes data collection, clustering nodes into entities and setting up the experimental environment. The quantitative analysis of centralization in the BLN reveals complex results, with the Gini coefficient for node capacity distribution increasing from 0.85 to 0.97 over eight yearly timestamps, indicating growing inequality. Meanwhile, the Nakamoto coefficient fluctuated, suggesting that while control over network resources is uneven, it may still be more decentralized than previously thought.

INDEX TERMS Bitcoin, lightning network, centralization, data processing, nodes.

I. INTRODUCTION

The Bitcoin Lightning Network (BLN) is a Layer 2 protocol built to address the scalability issues which are common to blockchain-based cryptocurrencies like Bitcoin (BTC). It enabled fast and inexpensive off-chain transactions through peer-to-peer channels [1], [2], [3]. As it is processing transactions off-chain, the Lightning Network (LN) is designed to address the challenges of scaling Bitcoin while adhering to the principles of its underlying peer-to-peer network. By enabling off-chain transactions, the LN aims to enhance network capacity while preserving the core principles of BTC [4], [5]. This architecture allows fast transactions that do not need to be settled or stored on the main blockchain.

The LN offers several key advantages, including rapid and low-cost transactions. By facilitating these efficient transfers, it contributes to improved blockchain scalability and enhanced user privacy [6]. The scalability of the LN is evident in its ability to handle many transactions simultaneously. As transactions are conducted off-chain, the LN can theoretically support a significantly higher number of transactions per second compared to the Bitcoin main chain, which is limited to approximately 7 transactions per second [7], [8]. This technology's potential to rival traditional payment systems is evident in its ability to enable instant, low-fee transactions globally and support micropayments within local economies [9]. Instant transactions allow a much higher transaction volume. Moreover, the LN addresses critical challenges associated with on-chain transactions, such as high fees and extended validation times, by utilizing off-chain payment channels [10], [11]. Only the opening and closing transactions of payment channels are recorded on the blockchain, minimizing on-chain activity. Since off-chain transactions also mean that transaction occur off the main BTC blockchain, they are not publicly recorded, and details are not visible on public ledger, enhancing the privacy [12]. Another way the LN improves privacy is through its use of onion routing. The LN employs a source-based onion routing

The associate editor coordinating the review of this manuscript and approving it for publication was Vivek Kumar Sehgal.

scheme like the one used in the Tor network, which obscures the identities of the sender and receiver by encrypting transaction data in layers [13].

While the LN offers advantages in scalability and transaction speed, it also has limitations and challenges. One of the primary limitations of the LN is the security vulnerabilities – while the LN enhances scalability by enhancing off-chain transactions, it also opened a potential for systematic attacks that could compromise multiple payment channels simultaneously [5], [14]. Liquidity issues are another limitation of the LN. Sufficient funds in payment channels are needed to successfully route transactions. The ability to route payments efficiently depends on finding paths with adequate liquidity, which can become problematic as channels may become unidirectional over time [15]. If the channel does not have funds, it cannot facilitate further transactions until it is topped up, leading to potential payment failures or delays.

One of the most crucial features Bitcoin provides is decentralization because it allows peer-to-peer transactions in a distributed and trustless system. The decentralized structure of the BTC network, maintained through proof-of-work, ensures a secure and transparent transaction ledger [16]. It provides independence from control and offers a globalized and decentralized payment network [17]. BTC achieves decentralized consensus through a network of miners rather than central authorities, as seen in the Nakamoto consensus [18]. Its decentralization is essential for the network's security, transparency, and independence from traditional banking systems.

Concerns have arisen regarding the potential for centralization within the Lightning Network (LN). The distribution of channel capacity among nodes raises questions about whether a small number of nodes are accumulating disproportionate control over the network [19]. The possibility of powerful, well-capitalized nodes operating as central hubs with extensive payment channels could lead to a centralized network structure, undermining Bitcoin's decentralized principles [20].

In this study, centralization refers to the degree to which a small number of nodes or entities accumulate disproportionate control over the Bitcoin Lightning Network. This control can be reflected in their number of connections (degree), the amount of liquidity they manage, or their role in facilitating transactions.

To analyze and determine the centralization of the BLN accurately, it is essential to identify appropriate metrics and methodology for analysis. However, the network's complexities, including challenges of clustering nodes into entities, further complicates the analysis. Selecting the most suitable centrality metrics and network variables is crucial for obtaining reliable results and demands meticulous attention.

The aim of this study is to conduct a quantitative analysis of centralization trends in the Bitcoin Lightning Network by employing different centrality metrics.

To achieve the aim of this study, the following objectives were raised:

1) Gather and process data from the BLN to create comprehensive data suitable for analysis.
2) Measure and analyze the changes in centrality metrics over time within the BLN.
3) Group the BLN nodes into distinct entities.
4) Using statistical methods, analyze how entities controlling specific nodes affect the overall centralization of the BLN.

Based on the emerging concerns, the following hypotheses are proposed:

- **H1**. The Bitcoin Lightning Network exhibits centralization around large entities, as clusters of nodes grouped by the entity hold a significant share of the network.
- **H2**. The centralization of the Bitcoin Lightning Network is increasing across time-series representations, following a non-linear trend detectable through centrality metrics.

While previous studies have typically relied on standard centralization metrics and consistently reported strong centralization, this paper integrates multiple centrality measures to provide a broader and more nuanced analysis of centralization trends within the BLN. Furthermore, by introducing a novel entity-level clustering approach, which groups multiple nodes under a single entity based on similarities in their alias naming patterns, this study uncovers centralized control structures that prior research has overlooked. Additionally, leveraging an extensive dataset spanning eight years (2018–2025), the analysis offers new and more accurate insights into the evolution and dynamics of network centralization.

The rest of the paper is structured as follows. Section II reviews the existing literature, emphasizing the importance of understanding centralization in the Bitcoin Lightning Network and identifying gaps in current knowledge and analyzing centrality metrics used for the research. Section III outlines the research methodology used in the analysis. Sections IV and V detail the research findings and comparative analysis of centrality metrics. Section VI presents the discussion, concluding with suggestions for future research directions. The key findings and implications of the study are provided in Section VII.

II. BACKGROUND AND RELATED WORK

The BLN facilitates fast and cost-effective off-chain Bitcoin transactions by establishing peer-to-peer channels between nodes that can be interconnected to create a routing path [21]. The Lightning Network facilitates frequent, rapid micro-transactions with minimal fees by using payment channels, which allow trustless exchanges between users [22]. These channels, developed from micropayment channel technology, extend the concept of one-way payment channels into two-way payment channels [23]. This setup allows for numerous payments while recording only two

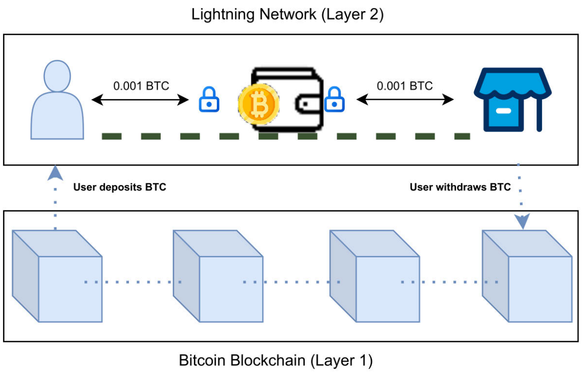

FIGURE 1. Lightning network.

transactions on the blockchain, reducing costs, and enhancing transaction speed. To establish a channel, both parties lock a certain amount of funds in a multisignature address on the blockchain. Transactions are then exchanged off-chain, with only the initial and final transactions being recorded on the blockchain. This method ensures that users can recover their funds if one party acts maliciously [2]. Fig. 1 illustrates the interaction between the BTC blockchain (Layer 1) and the Lightning Network (Layer 2).

Users deposit BTC into the blockchain to open a payment channel in the LN, enabling them to transact without recording each transaction on the blockchain. When users decide to finalize their transactions, they withdraw funds back to the BTC blockchain by closing the channel, which records the net result of their transaction.

A. EXISTING STUDIES ON BLN CENTRALIZATION

The centralization of the BLN recently has been an important topic within the cryptocurrency community. Existing research has shown that the BLN currently exhibits a high trend toward centralization, with a few powerful nodes acting as hubs [1]. These powerful nodes can be referred to as dominant nodes - individual nodes that either maintain a high number of channels or control a disproportionately large share of the total network capacity, thereby have significant influence over the network.

The authors of [4] analyzed the BLN over an 18-month period (from January 2018 to July 2019), focusing on degree distribution and Gini coefficient, and discovered that approximately 10% of nodes hold 80% of the bitcoins at stake in the BLN.

The paper [11] investigated degree and betweenness centrality patterns within the LN and discovered that from 2020 to 2022 the centrality has increased significantly within the LN and the Gini index has risen by more than 10% over that period.

The paper [19] found that the LN exhibits a centralized configuration, with the central nodes that play a crucial role in maintaining the network, the results of the Gini coefficient for node capacity distribution were 0.88 and coefficient for channel distribution was 0.75.

The paper [24] similarly observed the structure of the LN and noticed that there is a concentration of highly active nodes which can cause a potential centralization.

While existing research provides significant insights into the potential centralization of the BLN, it often overlooks important details necessary for a comprehensive understanding of BLN centralization. Firstly, many studies do not disclose their methods for data collection and processing from the LN, which is crucial for verifying and replicating findings. Secondly, there is a lack of deeper analysis of different centrality metrics, as most research focuses on the Gini coefficient. Furthermore, many analyses examine the centralization of individual nodes without attempting to group them into entities that control multiple nodes, which could reveal a higher degree of network centralization. Additionally, the dynamic nature of the LN is often neglected, with most research relying on static analysis at specific timestamps; representing these dynamic changes graphically would provide a clearer picture of network evolution. Lastly, existing studies often lack in providing recommendations on how the LN could be made more decentralized. Addressing these gaps would contribute significantly to a more thorough understanding of the BLN centralization and its implications.

To fully understand the structure and possible centralization of the BLN, it is essential to explore centrality concepts and metrics. Subsections II-B and II-C delve into the background of centrality aspects to gain insights into the distribution of nodes in the BLN, and centrality metrics, which could be employed to perform quantitative analysis.

B. TOPOLOGICAL CENTRALITY MEASURES AND THE ENTITY-BASED PERSPECTIVE

When assessing centralization within the BLN it is important to consider different centrality aspects. There are four main centrality aspects such as – betweenness centrality, degree centrality, weighted degree centrality, eigenvector centrality and closeness centrality.

One of the key aspects is betweenness centrality which highlights nodes that have significant influence by controlling the flow of information in the network [11]. It measures how often a node appears on the shortest path between other nodes. In the context of the LN, this highlights a node's importance in facilitating payment routes [25]. Nodes with high betweenness centrality are crucial for routing paths because they determine which paths for network traffic are the most efficient, unlike other centrality aspects such as degree centrality that focus on local connections [11]. This centrality measure is vital for understanding how information flows through the network and identifying nodes that act as intermediaries in the routing process.

$$C_i^{\{(\{betw\})\}} = \frac{1}{(N-1)(N-2)} \sum_{\substack{s,d=1 \\ s\neq d\neq i}}^{N} \frac{\sigma_{sd}(i)}{\sigma_{sd}} \tag{1}$$

As shown in Equation 1 N is used to represent the total number of nodes in the network, s and d is used to represent all pairs of nodes and i is the node which is passed through in the path between s and d.

Degree centrality is another centrality aspect when evaluating centralization in the LN. It quantifies the number of connections a node has, indicating its level of interaction within the network [26]. In the case of the LN, degree centrality quantifies the number of channels a node has. Nodes with high degree centrality are considered important hubs in the network, potentially exerting more influence due to their extensive connections [27]. Understanding the distribution of degree centrality across nodes can provide valuable insights into the network's structure and help identify potential central points of control or potential vulnerabilities.

$$C_i^{\{(\{deg\})\}} = k_i \tag{2}$$

As presented in Equation 2 k represents the total number of channels that node i has in the network.

While degree centrality evaluates the connections of nodes, weighted degree centrality also considers their importance by incorporating the capacity, or liquidity, into calculations [28]. This is particularly useful in networks where not all connections are equal. Weighted degree centrality can be particularly useful for analyzing centralization trends in the LN. The nodes in the network are connected through channels, and each of them has a capacity, a locked amount of BTC which can be interpreted as weights. By applying weighted degree centrality, it is possible to identify nodes that not only have a high number of connections but also those that facilitate a significant volume of transaction [11]. The weighted degree centrality formula is below (Eq. 3):

$$C_i^{(w)} = \sum_j w_{ij} \tag{3}$$

where w_{ij} is the weight of the channel between node i and node j. The weight of the channel is the capacity between two nodes – the amount of BTC that is locked up in a payment channel. It represents the total amount of BTC that can flow through the channel. The weighted degree centrality of a node i would be the sum of the capacities of all channels connected to that node.

Eigenvector centrality considers both the direct and indirect connections of a node, giving more importance to connections with nodes that are already well-connected within the network [26]. Nodes with high eigenvector centrality are well-connected to other influential nodes, enhancing their overall importance in the network [29]. Eigenvector and closeness centrality both consider the distance between nodes when identifying the shortest paths between them [30]. Distance between nodes refers to the number of nodes required to traverse from one node to another, which helps in determining how quickly the transaction can go through the network. This centrality aspect is valuable for identifying nodes that may have indirect but significant influence due to their connections with other central nodes.

$$\lambda C_i^{\{(\{eig\})\}} = \sum_{j=1}^{N} a_{ji} C_j^{\{(\{eig\})\}} \tag{4}$$

In Eq. 4, a_{ji} is an indicator to know if there is a channel between nodes j and i. If nodes have a channel together, then $a_{ji} = 1$, if there is no channel, $a_{ji} = 0$. Meanwhile, $C_j^{\{(\{eig\})\}}$ is the importance of node j, which is assigned to each node according on how many important nodes it is connected to.

Closeness centrality is another centrality aspect that assesses how quickly a node can interact with other nodes in the network [4]. Nodes with high closeness centrality can reach other nodes more efficiently, which can enhance the speed of information dissemination or transaction routing [31]. Analyzing closeness centrality can reveal nodes that play a critical role in maintaining network connectivity and ensuring efficient communication across the Lightning Network.

$$C_i^{\{(\{clos\})\}} = \frac{1}{\sum_{j=1}^{N} d_{ij}} \tag{5}$$

N is the total number of nodes in the network, d_{ij} is the shortest path distance between node i and node j. As shown in eq. 5, nodes with higher closeness centrality can more efficiently route transactions in the network. For the LN, this means how many nodes it takes to route a payment. The fewer the nodes, the closer they are.

Analyzing the centralization of Bitcoin Lightning nodes alone is not precise, as node owners are often the same entities. These entities may appear as separate units in the network topology, but they belong to a single verifier.

A method that allows grouping different, yet related, nodes is necessary for accurate evaluation. To address this, this paper introduces a novel aspect of centrality analysis by grouping individual nodes into entities. This method identifies and aggregates nodes under common control and this way enables the assessment of centrality at the entity level. This approach has the potential to reveal that entities controlling multiple nodes might have significant influence over the network's topology and transaction flow. By analyzing these entities, this method aims to provide a more accurate and comprehensive view of network centralization. The entity-based centrality analysis could reveal potential risks associated with a few entities having disproportionate control over the LN, offering crucial insights that were previously overlooked in traditional node-based analyses.

To conclude subsection II-B, this research is focusing primarily on the analysis of degree centrality and weighted degree centrality, which provide critical insights into the structure and potential centralization within the BLN. These metrics are particularly useful because they not only highlight how well-connected certain nodes are but also consider the capacity or liquidity of channels, offering a more nuanced understanding of influence within the network. These two aspects are not only the focal points of the research but are also directly incorporated into the calculation of the centrality metrics, providing a detailed measure of centralization within the network.

Betweenness, eigenvector, and closeness centrality are not directly used in the calculations, because they are all based on routing and do not consider channel capacities, which are crucial in this research. They may serve as points for discussion and in another research, offering additional perspectives on node influence and information flow. By focusing on degree and weighted degree centralities - which consider both direct connections and the capacity of nodes, this approach enables a more comprehensive understanding of how centralization impacts network efficiency and resilience.

In subsection II-C, the analysis of degree and weighted degree centrality will be explored further, as these metrics are specifically chosen for their ability to capture both the number of connections and the capacity of channels. This allows for a more nuanced understanding of the network's structure and centralization trends.

C. QUANTITATIVE METRICS FOR DEGREE AND WEIGHTED DEGREE CENTRALIZATION ANALYSIS

In this paper, several centrality metrics are analyzed to provide a thorough view of the potential centralization of BLN, focusing specifically on degree and weighted degree centralities. The metrics considered include the Gini coefficient, Lorenz curve, Nakamoto coefficient, Herfindahl–Hirschman Index (HHI), Theil index, Shannon entropy, Atkinson index, k-core decomposition, PageRank, and Katz centrality. These coefficients were initially chosen based on their capability to measure various forms of inequality, although not all metrics were selected for detailed analysis.

The Gini coefficient, originally introduced in 1921 to measure income and wealth distribution [32], has since found applications in various fields. In physics, for example, it has been used as a morphological measurement of strongly lensed galaxies in the image plane [33]. In agricultural economics, the Gini coefficient has been employed to assess the distribution of direct payments among agricultural farms, aiming for a more balanced allocation of resources [34]. In medicine, the Gini coefficient has been applied to evaluate the equity in the distribution of mental health resources in China, helping to assess how accessible healthcare is across different regions [35].

In recent years, the Gini coefficient has also been increasingly applied to blockchain networks, where it is used to analyze wealth distribution and decentralization. One such example is its use in blockchain network studies to measure the decentralization of wealth across users [36]. This demonstrates the versatility of the Gini coefficient in measuring inequality across various fields.

In the case of the BLN, the Gini coefficient is a commonly used metric that measures the inequality of channel distribution within the Lightning Network (LN). A higher Gini coefficient indicates greater centralization and is recognized as a strong indicator of overall network centralization, especially when combined with other measures [4], [37]. The coefficient is calculated based on the Lorenz curve, which illustrates the cumulative distribution of the variable. The Gini coefficient ranges from 0 to 1, where 0 represents perfect equality with everyone having an equal share, and 1 represents total inequality where one entity possesses everything [4], [18]. A lower Gini coefficient suggests a more balanced (decentralized) network, while a higher coefficient indicates a more uneven (centralized) distribution of the resource being analyzed. It can be measured using the following formula (Eq. 6):

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2N^2\bar{x}}, \tag{6}$$

N is used to represent a total number of nodes, x_i and x_j represent capacity of nodes and x is an average capacity across all nodes. As shown in Eq. 6, a higher G indicates greater inequality in capacity distribution across the network.

The Lorenz curve is a valuable function, interpreted as a graphical representation and used to measure inequality within a population. It illustrates the proportion of the distribution held by a certain percentage. The Lorenz curve is also related to the Gini coefficient as it quantifies inequality by comparing the area between the perfect equality line (45-degree angle) and the observed Lorenz curve. It is computed by subtracting the Lorenz curve area from the perfect equality line area, and then dividing by the total area under the equality line [11]. For this paper, the Lorenz curve will be used to provide a visual dynamic representation of channel distribution among nodes based on their channel capacity.

The Lorenz curve has been applied in various fields beyond the LN. For example, [38] proposed a method for estimating the Lorenz curve to analyze size distributions and inequality across different scientific disciplines, highlighting its relevance in social sciences and environmental studies. In the field of quantum mechanics, [39] introduced the concept of quantum relative Lorenz curves, applying it to noncommutative geometry and quantum entanglement, which highlights the adaptability of the Lorenz curve to complex systems. [40] used the Lorenz curve in social choice theory to assess aversion to inverse downside inequality, emphasizing its importance in understanding societal preferences regarding inequality. These examples demonstrate the broad applicability of the Lorenz curve in both sociological and technological research.

The Nakamoto coefficient is a combined method of the Gini coefficient and Lorenz curve. It is a recent metric employed to analyze the level of decentralization within cryptocurrency networks. Study [41] shown that tokens such as Chainlink (LINK) and Polygon (MATIC) exhibit high levels of centralization, where a small number of wallets control a significant share of the wealth. The Nakamoto coefficient helped to identify this centralization. The Nakamoto coefficient quantifies a blockchain's decentralization by identifying a minimum number of entities required to control more than half (51%) of mining power, transaction volume or total wealth within the network [42]. Initially it emerged as a metric to assess the susceptibility of the Bitcoin network to a majority attack [41]. Overall, this coefficient answers the question how many nodes must be compromised to undermine the network's integrity?

Higher value of the minimum Nakamoto coefficient means that the system is more decentralized. It is defined as (see Eq. 7):

$$K = \min \left\{ n \in \mathbb{N} : \sum_{i=1}^{n} x_i > \frac{1}{2} \sum_{i=1}^{N} x_i \right\}, \tag{7}$$

where n is the number of nodes being considered and N is the total number of nodes in the network. x_i is the capacity of the node, or the number of channels the node has. The condition for this formula is that the sum of values controlled by the top n nodes must be greater than half of the total value in the network. The Nakamoto coefficient, K, gives the smallest number of entities that can control more than half of the network.

The Herfindahl-Hirschman Index (HHI) measures the market concentrations and can be widely used to assess market effectiveness. In the healthcare sector, the HHI revealed significant centralization among healthcare providers, which impacts competition and patient costs [43]. In banking, higher HHI values indicated increased concentration and provided a nuanced perspective on market dynamics [44]. Similarly, in the sports funding context, the HHI was used to demonstrate the concentration of funding among Olympic sports, indicating a centralized approach to allocation [45].

The HHI starts near zero in highly competitive markets, and it can climb up to ten 10,000 if a single node dominates the network. Overall, value below 1500 indicates an unconcentrated market with many players, value between 1500 and 2500 signifies a moderately concentrated marked with some dominant companies, or nodes, and if the results are above 2500 it suggests a highly concentrated market held by a few significant nodes [46]. In the BLN analysis, a lower HHI value would suggest a more distributed set of nodes, supporting the goal of decentralization, while higher values might indicate emerging centralization tendencies, where a few nodes begin to dominate.

It can be measured by using the equation 8 below [47].

$$HHI = \sum_{i=1}^{n} 10000 * (H_i)^2, \quad H_i = \frac{h_i}{C}, \tag{8}$$

where h_i is the capacity held by node i, and C is the total network's capacity by that timestamp.

The Theil index, also known as the Theil coefficient, is a measure used to assess inequality. It can also be used to assess the centralization of networks, evaluating the distribution of network capacity or number of channels among nodes [48]. It ranges between 0 to inequality, where 0 signifies a perfect equality in the network, while higher values show an increasing inequality between nodes.

The Theil index had been applied in various studies to assess centralization and inequality. For example, it was used to measure regional differences in urban energy consumption, highlighting significant disparities and centralization of energy use [49]. Similarly, it was used to analyze civil vehicle ownership distribution, revealing higher ownership in urban areas compared to rural regions, indicating centralization [50]. In regional poverty analysis, the Theil index showed economic disparities in the Xiamen-Zhangzhou-Quanzhou city cluster, with centralization of development in certain regions [51].

In the context of the BLN, the Theil index can provide valuable information into the centralization within the network. A lower Theil index value would indicate a more balanced distribution of node capacity or number of channels, while higher values would point to centralization.

The Theil index is measured using the eq. 9 below:

$$T = \frac{1}{N} \sum_{i=1}^{n} \frac{x_i}{x_j} \ln \left(\frac{x_i}{x_j} \right), \tag{9}$$

where N is the total number of nodes, x_i is a capacity of a single node, and x_j is a capacity across all nodes.

However, when comparing Theil index results across different time periods, it is crucial to use normalized values rather than raw values. The normalization process adjusts the differences in the scale of data and ensures that results are comparable. By normalizing the index values to a range between 0 and 1, it is easier to interpret and compare.

To be able to calculate normalized Theil values, first it is important to measure maximum Theil index, which occurs when all the weight is concentrated in one node (Eq. 10).

$$T_{max} = ln(N) \tag{10}$$

Next step in normalizing Theil index values is to divide the original Theil index by the maximum value (Eq. 11).

$$T_{norm} = \frac{\frac{1}{N} \sum_{i=1}^{n} \frac{x_i}{x_j} \ln \left(\frac{x_i}{x_j} \right)}{\ln (N)} = \frac{T}{T_{max}} \tag{11}$$

Shannon entropy is a fundamental concept in information theory that quantifies uncertainty or randomness in a set of outcomes [52]. Shannon entropy has also been employed in various fields. For instance, in the social sciences, Shannon entropy has been used to assess community vulnerability to natural disasters, illustrating its utility in understanding the resilience and stability of communities under stress [53]. In technological context, Shannon entropy is pivotal in information processing tasks, such as data compression and communication channel capacity. It quantifies the amount of information that can be reliably transmitted over a channel, thereby informing the design and optimization of communication systems [54].

In the context of the BLN, Shannon entropy can be used to measure the network's centralization by analyzing the distribution of node capacity or number of channels. A higher Shannon entropy indicates a more decentralized network, as node capacity or the number of channels is more evenly distributed among nodes, while lower entropy suggests centralization, with a few nodes holding a significant part of the network [1], [24].

Shannon entropy is measured using the Eq. 12 below [55].

$$SE(x) = -\sum_{y=1}^{n} x_i \log(x_i) \tag{12}$$

where x_i is the capacity or the number of channels a single node has.

However, it is also important to calculate normalized Shannon entropy to be able to interpret and compare the results across different time periods and datasets. The normalized entropy values range between 0 and 1, which can help to assess how it changed throughout time.

To calculate normalized Shannon entropy value, first the maximum value of the Shannon entropy must be measured. The equation 13 for the maximum entropy is below, where n is the total number of nodes.

$$SE_{max} = log(n). \tag{13}$$

Normalized Shannon entropy can be measured using the following Eq. 14.

$$SE_{norm}(x) = \frac{-\sum_{y=1}^{n} x_i \log(x_i)}{\log(n)} = \frac{SE(x)}{SE_{max}}, \tag{14}$$

where $SE(x)$ is Shannon entropy value and SE_{max} is the maximum entropy value.

This research has employed several well-known coefficients to measure various aspects of centralization and inequality, including Gini coefficient, Lorenz curve, Nakamoto coefficient Herfindahl-Hirschman Index (HHI), Theil index, and the Shannon entropy. Each of these metrics was chosen for their ability to highlight different facets of centralization within the BLN. However, there are other metrics commonly used to evaluate centralization, which were considered but decided not to be included in the analysis for specific reasons. These include metrics such as the Atkinson index, K-core decomposition, PageRank, and Katz centrality.

The Atkinson index, which is widely used in economic studies to measure income inequality [56], was one such measure considered. While the Atkinson index is highly effective in evaluating inequality by adjusting the sensitivity to different parts of the income distribution, it is less applicable in a network context where individual preferences about inequality aversion are not as directly relevant [57]. Therefore, the Atkinson index was not included due to its less straightforward interpretation for network topologies and centrality.

K-core decomposition was also reviewed as a potential metric. This measure identifies cores of highly interconnected nodes, which can help in understanding the robustness and resilience of networks [58]. However, it is more suitable for analyzing clusters rather than providing an overall measure of centralization [59], leading to its exclusion from this study.

PageRank and Katz centrality were similarly evaluated. PageRank, often used in ranking nodes based on their connections [60], and Katz centrality, which measures the influence of a node within a network by considering direct and indirect connections [61]. However, these metrics are more focused on ranking nodes rather than quantifying the overall distribution of resources or influence, which is the primary focus of this centralization analysis.

In conclusion, the selected metrics provide a comprehensive assessment of BLN centralization, capturing key aspects of inequality. This selection forms a solid foundation for research methodology.

III. RESEARCH METHODOLOGY

Understanding the centralization within the BLN is crucial for assessing its reliability, scalability, and potential vulnerabilities. Evaluating reliability involves assessing how decentralization enhances the network's resilience, as a more distributed structure reduces the risk of failures and ensures that the network can continue to function smoothly even if some nodes go offline. Centralization can affect scalability by creating bottlenecks if most transactions are controlled by a small number of nodes. Potential vulnerabilities arise from centralization making the network more prone to attacks or service interruptions if key nodes are compromised [5], [14].

While the LN is known as a promising second layer solution for the BTC, a systematic analysis of its network topology and the identification of the most influential entities and nodes have been unexplored. This paper aims to advance the understanding of the LN topology by developing a methodology that incorporates a set of centrality metrics, including Gini index, Lorenz curve, Nakamoto coefficient, HHI index, Theil index, and Shannon entropy, to quantify and visualize network centralization. These specific centrality and inequality metrics were selected due to their effectiveness in analyzing inequalities and centralization in various network contexts. Prior literature has validated the use of these metrics in cryptocurrency and blockchain network analyses due to their sensitivity in capturing inequalities and concentration risks [36], [62]. By combining these metrics, the study ensures a comprehensive assessment of centralization within the BLN.

A significant challenge in assessing the centrality of the LN is also inseparable to the complexities of the network itself. The success of clustering nodes into entities remains an open question, especially given the presence of multiple entities using different nodes, which could impact the reliability of results, as does the precise number of active channels within the network. Furthermore, mobile channels of the LN (e.g., Phoenix, Éclair, private channels) introduce uncertainty into the data. And lastly, selecting the most appropriate centrality metrics and network variables to accurately measure centralization within the LN is a complex task which requires careful consideration.

The entity-level clustering method employed is based on alias similarity. Such name-based resolution techniques have been successfully used in prior research to identify entities operating multiple nodes in blockchain and financial networks, where systematic alias naming conventions help uncover centralized structures [63], [64]. The established effectiveness of this simpler alias-based clustering approach lends credibility to its applicability and accuracy in the context of the BLN.

To address the challenges this paper adopts a quantitative research approach using various centrality metrics and proposes a systematic methodology for centrality analysis in the BLN. By employing a set of centrality metrics and network properties, this paper aims to provide a thorough and comparable framework for assessing a network's centralization.

This study will be conducted through several stages. First, in subsection III-A, data will be collected on nodes, channels and transactions within the BLN using the LND Database Reader stack, which integrates LND (Lightning Network Daemon), a Golang program, and a MySQL database. Next, in subsection III-B, for Hypothesis 1, a clustering analysis will be conducted to group nodes into entities. Following this, in subsection III-C, for Hypothesis 2, the collected data will be analyzed using centrality metrics such as the Gini index, Lorenz curve, Nakamoto coefficient, HHI index, Theil index, and Shannon entropy to evaluate the distribution of influence and connectivity across the network. These metrics will be calculated separately for both individual nodes and clustered entities. In the subsection III-D, experimental setup employed to implement this methodology is described. The results will then be visualized to enhance the interpretability and dynamics of the findings. Finally, a comparative analysis of the metrics will be performed to identify how each one reflects network centralization and to explore potential discrepancies in the results.

A. DATA COLLECTION AND PREPROCESSING

The real-time data collection for this research involves using the LND Database Reader software stack, which incorporates three key components: Lightning Network Daemon (LND) [65], a Golang-based program, and a MySQL database (Fig. 2). LND was selected as BLN node for the data collection because of its feature to connect to BTC layer 1 via Neutrino protocol [66]. Neutrino is a lightweight BTC client that enables BLN data collection without the need to run a BTC full node.

Neutrino also known as a BTC light client, allows LND to interact with the BTC network by only downloading a subset of the blockchain data. This approach significantly reduces the resource requirements for running BLN data collection instance, making it more accessible to run even several instances if needed.

LND Database Reader stack is designed for reproducibility using Docker containerization, with each component deployed as a separate container orchestrated through Docker Compose configuration. This containerized approach ensures consistent deployment across different computational environments and facilitates research validation by independent researchers.

LND stores its data in a BoltDB database [67], which is file-based and optimized for fast access within the Go programming language (Golang). However, accessing the database in real-time poses a challenge, as LND locks the BoltDB file, preventing simultaneous access by external programs. To overcome this, the LND Database Reader begins by initializing its configuration parameters, including database paths and connection strings. It then creates a temporary copy of the BoltDB file, allowing the system to read the data without interrupting LND's operations.

Once the database copy is created, the Golang program extracts relevant data such as 'channel announcements' and 'node announcements,' which are key messages broadcast across the BLN. The program iterates through each announcement in the LND database, processing them sequentially. Channel announcements contain critical details about new payment channels, including their unique identifiers (ID) and the nodes involved. Similarly, node announcements inform the network about new nodes joining the BLN, containing the IDs of the nodes and other information such as the node's operator or its public key.

The extracted data is then transferred to a MySQL database for long-term storage. MySQL was selected for its robust data management capabilities, including its ability to prevent duplicate entries during data insertion. This ensures the data collected from the BLN remains consistent and up to date, enabling the tracking of real-time network activity.

However, this method captures only real-time data. For this research, historical data is also necessary to understand how the network evolved over time. To address this, additional data is gathered from external sources – the Lightning Network Research [68] and the BTC blockchain.

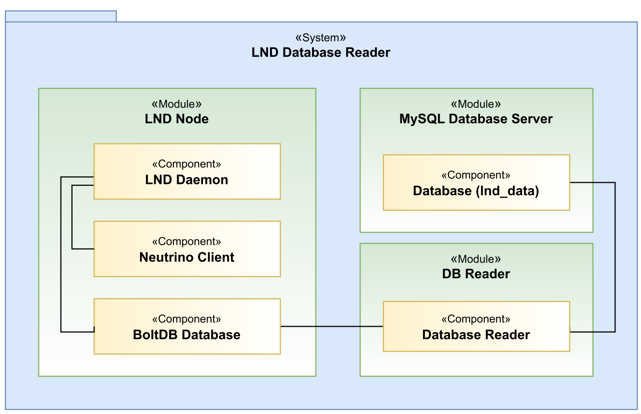

FIGURE 2. LND database reader system structure.

The LN Research repository collects BLN data, including the same 'channel announcements' and 'node announcements' collected in real-time, but dating back to previous years. The messages were imported into the same MySQL database used for real-time data, allowing for direct comparison between historical and current data (Fig. 3).

This approach of integrating on-chain and off-chain data has been successfully used in previous studies to analyze temporal network dynamics and assess changes in structural properties of cryptocurrency systems. For example, [69] and [70] demonstrated the effectiveness of combining Bitcoin blockchain data with Lightning Network topology for uncovering network evolution and centralization trends. Their work validates the methodological soundness of using unified datasets built from both BLN and BTC sources, as implemented in this research.

To collect data on when BLN channels were opened or closed, the research utilizes the BTC blockchain, as this information is not available directly within the BLN itself. A full BTC node, in conjunction with an Electrum node, is used to index and retrieve transaction data efficiently. To deploy BTC full node and Electrum node was selected MyNodeBTC operating system [71] for its convenience of deployment and management (Fig. 4). The Electrum node allows for quick access to specific blockchain transactions, including those that locked BTC into payment channels or closed them. Each transaction associated with a channel is identified by its 'ShortChannelID,' which includes the block height, transaction index, and output index.

A critical aspect of this research is the ability to link data collected from the BLN with BTC blockchain data, creating a unified dataset. The process of connecting these datasets is centered around the 'ShortChannelID,' which acts as a unique key across the different databases.

In the BLN dataset, the 'ShortChannelID' is used to uniquely identify each payment channel. This ID is recorded when a channel is announced via the LN's gossip messages and is stored in the MySQL database. In parallel, the same 'ShortChannelID' is recorded on the BTC blockchain when BTC is locked into a payment channel. By cross-referencing the 'ShortChannelID' in the BLN and blockchain datasets, the research can establish a direct link between the announcements of channel creation in the LN and the underlying transactions recorded on the Bitcoin blockchain.

In addition to linking channels, node data is aggregated by combining individual nodes under common aliases. The MySQL database stores both 'NodeID' and 'Alias', a user-friendly name associated with each node. This information helps to track the activities of entities that control multiple nodes, providing a higher-level view of network centralization. By analyzing historical data on nodes and aggregating them under shared aliases, the research can observe the growth of specific entities over time and their influence on the network.

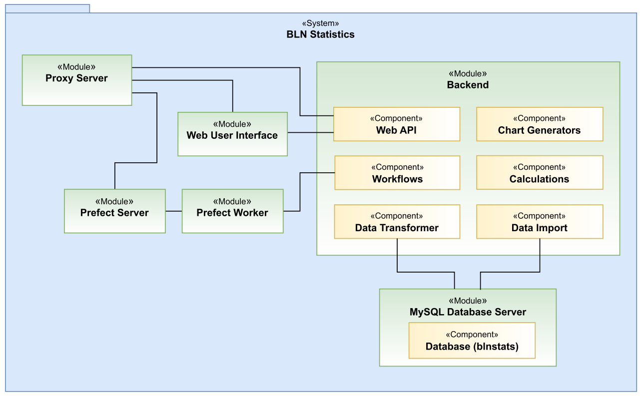

FIGURE 3. BLN statistics system structure.

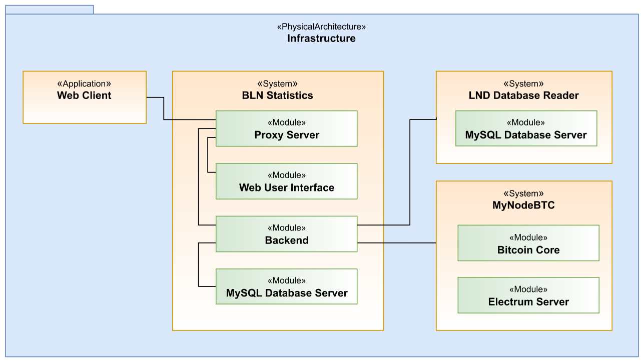

FIGURE 4. High level overview of BLN Statistics system with data collection.

After data is collected using this approach, combining the data from both BLN and BTC blockchain, the data is prepared to be clustered into nodes and to conduct a quantitative analysis of centralization trends in the BLN by employing different centrality coefficients.

Managing and orchestrating long-lived processes, Prefect server was selected. Prefect is a workflow management system that enables task scheduling, execution and monitoring of complex data pipelines [72]. In this research Prefect facilitated tasks such as data import, processing, calculation and chart generation. By using Prefect, these processes were organized into workflows with defined dependencies to ensure that each task was executed in correct sequence. This greatly reduces complexity for bootstrapping the system and managing in the future. Prefect's monitoring capabilities allow real-time insights into the status of each workflow, enabling quick detection and resolution of any issues during execution. Automated workflow management ensures reproducibility and facilitates validation of results across different research environments.

The following subsection III-B presents how collected and processed data was used to cluster nodes to entities.

B. NODE CLUSTERING AND ENTITY IDENTIFICATION

The methodology for Hypothesis 1 is based on the manual clustering of nodes based on alias data gathered from historical BLN records. In this context, entity clustering refers to the process of grouping individual nodes that appear to be operated by the same organization or individual, based on similarities in their alias names. This method is based on the patterns and conventions in node aliases to identify entities under common ownership or control, also known as Alias-based Entity Resolution [73]. Although the clustering approach is relatively simple, it is methodologically sound and supported by prior literature. [63] demonstrated the practical effectiveness of simple alias-based entity resolution techniques in relational datasets, highlighting their utility even without complex computational methods. Similarly, [64] validated the successful use of alias-based clustering in cryptocurrency networks, where entity identification was effectively achieved through simple alias naming patterns, the exact approach taken in this study.

This analysis leveraged on the tendency of node operators to include the names of companies or entities in their full aliases, which often reflect systematic naming schemes used for branding of the companies. By examining these aliases, patterns emerged that suggested certain nodes were operated by the same entity or group. The largest entities by network capacity – such as Bitfinex, LNBIG, River, ACINQ, Kraken, Bitrefill, NiceHash, etc. – often use their company names directly in their node aliases, making them easily identifiable.

Initially, a comprehensive list of node aliases was collected from historical data and prepared for the analysis. The data was standardized by normalizing text cases and removing special characters to ensure consistency. This preparation was crucial for accurately detecting naming patterns without interference from formatting discrepancies or variations caused by inconsistent data.

The analysis focused on identifying common prefixes, suffixes, and full names of companies or entities within the aliases. For instance, nodes with aliases like "bfx-lnd0", " bfx-lnd1" and "bfx-lnd2" shared the prefix "bfx-ln" indicating operation by the "bitfinex.com" entity. Similarly, aliases such as "Bitrefill.com/gift-cards", "Bitrefill" and "Bitrefill Routing" were grouped under the "bitrefill.com" entity. Identifying these prefixes allowed for the grouping of nodes that share the same company or entity name in their aliases.

Further examination revealed that some operators employed thematic naming conventions or included unique identifiers (IDs) within their aliases. For example, LNBIG uses aliases like "LNBIG.com [lnd-20]", "LNBIG.com [lnd-21]" and "LNBIG.com [lnd-22]" incorporating both the entity name and unique node identifiers. This systematic naming not only aids internal tracking for the operators but also facilitates external analysis in identifying clusters of nodes under common control.

Unique sequences of letters and numbers within aliases also offered insights into potential associations between nodes. Some operators include identifiers like serial numbers or internal codes (e.g. "node204.fmt.mempool.space", "node201.va1.mempool.space"), which, when matched across multiple aliases, indicate a common source. This approach leverages the operators' systematic naming for ease of management while aiding the clustering process.

Finally, the list of all entities and their capacities is generated to analyze how the liquidity is distributed among the network and if Hypothesis 1 can be confirmed.

While this alias-based clustering method is practical and effective, it is not without limitations. Node aliases are unverified and manually assigned by operators, which may lead to false positives, by grouping unrelated nodes, or false negatives, by missing common ownership.

The following subsection III-C describes subsequent steps in data processing, providing a more detailed methodology for performing qualitative analysis of centralization in the BLN.

C. FURTHER PROCESSING STEPS

Following data collection and entity aggregation, the "BLN Stats" system, introduced by the authors, implements a comprehensive analysis pipeline that transforms channel and node announcement data into quantitative metrics suitable for centralization analysis.

Before node metrics transformation can begin, the system performs blockchain block and associated transaction synchronization. This stage ensures temporal consistency between BLN data and corresponding Bitcoin blockchain states. The blockchain synchronization process creates temporal reference points that align BLN events with their underlying on-chain transactions. This ensures that subsequent node metrics calculations are based on clean, validated datasets with accurate alignment between off-chain Lightning Network topology and on-chain blockchain transactions.

Next stage involves transforming BLN announcements and blockchain transaction data into structured node metrics. This transformation process aggregates channel information at the node level, computing fundamental network statistics - node degree (number of connections), weighted degree (total channel capacity). The system processes historical snapshots at regular monthly intervals, to capture network evolution over time. The temporal granularity enables tracking of individual node growth patterns and network participation changes over time.

The quantitative centralization analysis employs multiple statistical coefficients, each providing different perspectives on network concentration. The system implements standardized calculations for Gini coefficient, Herfindahl-Hirschman Index (HHI), Theil Index, Shannon Entropy, and Nakamoto coefficient. Each coefficient calculation processes the distribution of network resources (capacity or connections) across participants, generating values that indicate the degree of centralization. The system computes coefficients for both capacity-based and connectivity-based distributions, providing comprehensive coverage of different centralization aspects. Temporal coefficient series enable trend analysis and identification of centralization patterns over network evolution.

The final processing stage is output generation which produces both quantitative datasets and visual representations suitable for research publication and further analysis. Automated chart generation creates temporal evolution plots, distribution analysis visualizations, and comparative charts across different network metrics.

The complete research methodology described in this study has been implemented as an open-source software available through a public GitHub repositories at VUKNF-Fintech-Research-Group. The containerized architecture enables researchers to reproduce the entire analysis pipeline using Docker, eliminating environment-specific dependencies and ensuring consistent deployment across different computational platforms. The repositories include installation and setup instructions and default workflows that enable independent verification of research results. Researchers can either utilize the full system initialization workflow to process complete historical datasets or execute individual processing components for focused analysis. The modular design supports customization while maintaining methodological consistency. This implementation facilitates collaborative research efforts and enables extension of the centralization analysis methodology to other cryptocurrency lightning networks.

The experimental setup employed to implement this methodology is described in the following subsection III-D. It outlines the practical procedures, tools, and environment used to conduct the analysis, detailing the specific conditions under which the centralization metrics for the BLN were measured and evaluated.

D. EXPERIMENTAL SETUP

The methodology for Hypothesis 2 employs a quantitative analysis approach trough static analysis. The BLN is a dynamic network with nodes and channels constantly being added or removed. Static analysis is used to freeze the network and capture a snapshot of the network's structure at a specific point in time, making it possible to compare centrality measures. The study examines network snapshots at eight distinct points in time, beginning in March 2018 – at the same time when the Lightning Labs' lnd was released [74], the first LN implementation – and concluding in March 2025, using the most recent data available at the time of the experiment.

While March data points are used as representative timestamps for annual comparison, the analysis is based on full monthly data spanning from March 2018 to March 2025, enabling a continuous and accurate time-series representation of centrality trends.

TABLE 1. Lightning Network nodes and channels at specific timestamps.

Abbr.	Timestamp	Date	Number of nodes	Number of channels
T1	1519855474	March, 2018	450	828
T2	1551391683	March, 2019	4626	34232
T3	1583014153	March, 2020	6683	37111
T4	1614550557	March, 2021	9629	39979
T5	1646088233	March, 2022	21524	89739
T6	1677621623	March, 2023	19666	86532
T7	1709244541	March, 2024	18541	61675
T8	1740780152	March, 2025	16866	54533

This approach enables the tracking of centralization trends and the assessment of how key metrics such as the Gini coefficient, Lorenz curve, Nakamoto coefficient, HHI, Theil Index, and Shannon entropy evolve over time. The research is not only concerned with the static centralization at each timestamp but also seeks to understand the dynamic shifts in these metrics across different periods, providing a comprehensive view of centralization.

Table 1 shows the dataset used in this study. Over the period from March 2018 to March 2025, there is an apparent growth in both the number of nodes and channels. Starting with 450 nodes and 828 channels in 2018 (T1), the network expanded dramatically to reach 21,524 nodes and 89,739 channels by 2022 (T5). There was a steady decrease since 2022 (T5) to 16,866 nodes and 54,533 channels, the latest trend reflects that the BLN is scaling down.

In this study, the analysis is divided into two primary scopes of results. First, **channel capacity**, which is focused on degree centrality and measures how many channels a node has within the network at different timestamps. This offers insights into the distribution across the LN and helps identify the most connected nodes. Second, **node capacity**, which is assessed through weighted degree centrality and measures the total capacity in each node. This weighted analysis is more complex and highlights the nodes with the most influence in the network. By analyzing both channel capacity and node capacity, the study provides a comprehensive assessment of network centralization, combining the distribution of connections with the weight of each node. This way this study offers a better understanding of the network's structure.

IV. CENTRALIZATION ANALYSIS OF THE LIGHTNING NETWORK

This section presents the distribution of entities in the BLN and calculated centralization coefficients, focusing on key metrics introduced in the methodology. Comparisons across timestamps allow for tracking changes in the network's centralization over time. Additionally, the findings are assessed in relation to previous studies that utilized similar datasets and calculation methods.

A. DISTRIBUTION OF ENTITIES

Visual representation of bubble graph was utilized to illustrate the distribution of nodes within the network on the exact

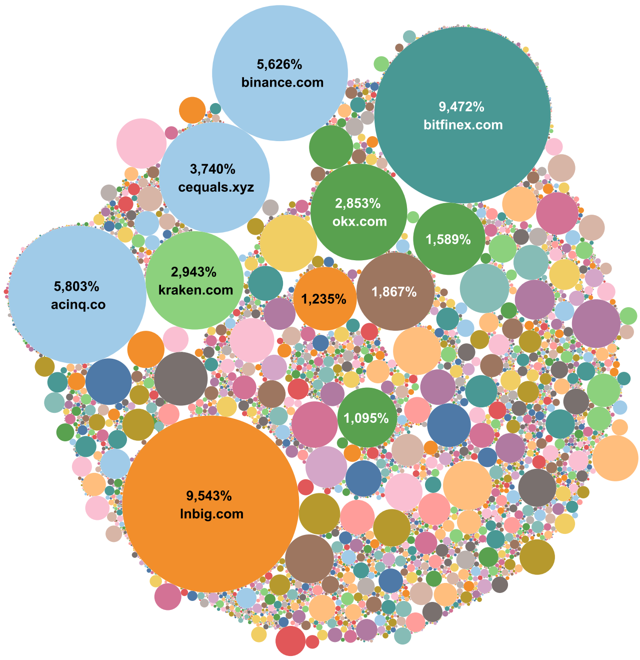

FIGURE 5. Distribution of entities in Lightning Network.

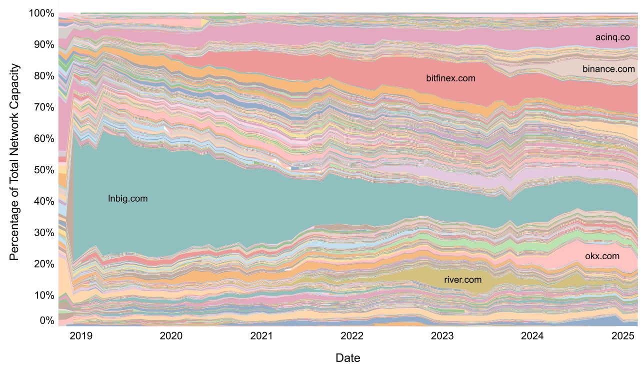

FIGURE 6. Total network capacity distribution among entities in Lightning Network.

date of 1st March 2025 (see Fig. 5) – this date is chosen for visualization, because it is the most recent and accurate data of BLN nodes. This graph showcases the capacities of the entities, providing a picture of BLN's topology. By representing entities like Bitfinex, LNBIG, Binance, ACINQ, Kraken and Cequals with proportionally sized bubbles, the graph highlighted how these entities, through their multiple high-capacity nodes, play pivotal roles in the network's structure. A total number of 15,530 entities that own a total of 4,839 BTC, is used in this graph and 10% of entities control 96.5% of the network.

Additionally, an area chart was created to represent the presence and growth of these entities throughout the entire existence of the LN, from 2018 to 2025 (see Fig. 6).

It provides insights into how the sizes of entities within the network have evolved over time. Initially, LNBIG was the largest entity in the BLN and dominated a significant portion of the network's capacity. As the overall network capacity increased, LNBIG's percentage of the total capacity decreased, and other entities such as Bitfinex, ACINQ, and River emerged. This evolution indicates that the network has transitioned from being dominated by a single major player to a more diversified structure with multiple significant entities. Despite the rise of these larger entities, there remains a multitude of smaller nodes and entities that do not have the same level of influence on the network's overall dynamics. However, the emergence of several major players also raises concerns about potential centralization, as these entities could collectively have substantial control over the network.

This alias-based clustering methodology illuminated aspects of centralization within the BLN by revealing how certain operators might exert influence due to managing multiple high-capacity nodes. Entities such as Bitfinex, LNBIG, River, ACINQ, and Kraken emerged as significant players in the network. By operating numerous nodes that are easily identifiable through their aliases, these entities impact the overall topology and centralization metrics of the network.

By grouping nodes according to the names of companies or entities in their full aliases, the study provided a foundational understanding of potential centralization within the BLN. This initial clustering step was crucial for more in-depth analysis using centralization metrics. These findings support Hypothesis 1, demonstrating that centralization appears around large entities, because the clusters of nodes grouped by the entity control a significant share of the network.

B. CENTRALITY METRIC RESULTS

This subsection presents the results obtained using each of the selected centrality metrics. The findings illustrate how centralization in the BLN has evolved over time and highlight the differences captured by each measure.

1) GINI COEFFICIENT

The Gini coefficient values for degree centrality (based on channel distribution) show a clear trend of increasing centralization over time – it rises from 59.6% at T1 to 75.9% at T6 (see Fig. 7). This indicates a growing inequality in the distribution of channels among nodes, which means that over time more channels are concentrated in a smaller number of nodes over time. The most significant growth is noticed between the first and second year of the LN – Gini coefficient increased from 59.8% to 74.6%. This increase is then followed by more gradual change in the next timestamps. Although, it's important to mention that the latest data shows that Gini coefficient slightly decreased from 76.4% in 2022, to 74,1% in 2025, which indicates that centralization slightly decreased in the last 3 years.

In contrast, the Gini coefficients for node capacity are significantly higher, starting at 84.9% at T1 and rising to 97.0% by T8, as shown in Fig. 8. This indicates that in terms of node capacity, the amount of liquidity each node controls, centralization is even more noticeable than in the channel distribution. The sharpest increase occurs between T2 (89.4%) and T3 (93.3%), reflecting growing concentration of liquidity in fewer nodes.

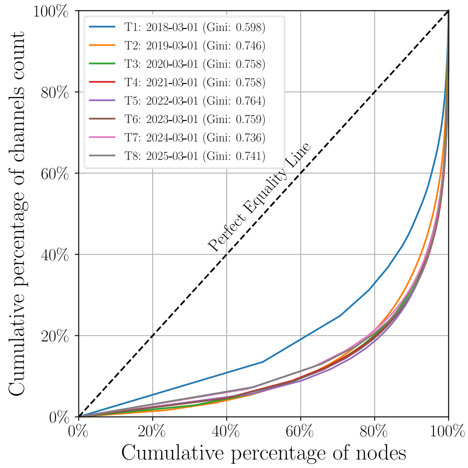

FIGURE 7. **Lorenz curves of BLN nodes on degree centrality.**

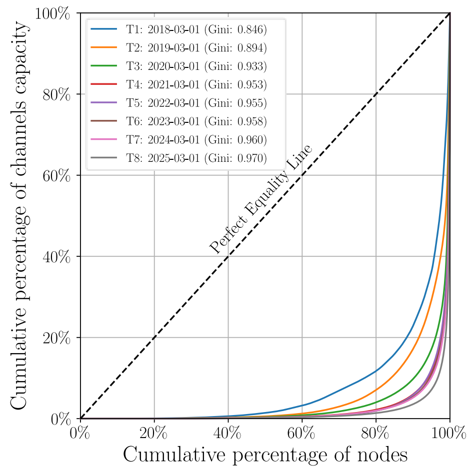

FIGURE 8. **Lorenz curves of BLN nodes on weighted degree centrality.**

These findings align with the trends identified in previous research on the LN. Study [20] reported a similar increase in centralization, with the average Gini coefficient of 76% for the number of channels in nodes and average of 95% for node capacity in 2022. The results in this study are slightly higher, with the coefficient of 76.4% for the number of channels and coefficient of 95.5% for the node capacity in 2022.

2) NAKAMOTO COEFFICIENT

For channel distribution, the Nakamoto coefficient for nodes starts at 40 in T1 (see Fig. 10), and it indicates that at the network's initial stages, only 40 nodes were needed to control over half of the channels, which means that control was concentrated in the hands of few nodes. As the network matures, this figure increases, reaching a peak of 685 in T5 before dropping steadily to 510 until T8. The sharp rise between T3 (220) and T5 (685) suggests that channel distribution has become more decentralized, with more nodes needed to hold significant control over the network. However, this does not indicate a sustained trend, as the Nakamoto coefficient begins to decline after T5. While this rise initially matched with a period of growth in node participation (see Fig. 9), the later decrease in node count suggests that overall distribution has become less balanced in subsequent periods. This can be observed when calculating the Nakamoto coefficient based on the node capacity.

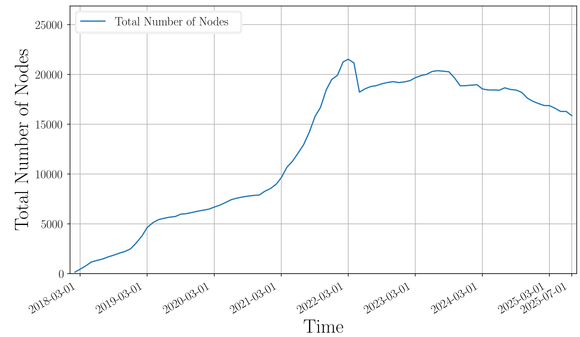

FIGURE 9. **Total number of BLN nodes in the Lightning Network.**

For entities (see Fig. 10 and Fig. 11), Nakamoto coefficient is progressing similarly to the coefficients of nodes, but the overall values are lower. This indicates that clusters of nodes belonging to single entities hold a higher concentration of control compared to individual nodes acting independently. When nodes are grouped into entities, fewer of these entities are needed to control a substantial portion of the network's channels or capacity. This suggests that while the network may appear more decentralized when viewed at the node level, centralization remains evident at the entity level, where a smaller number of clusters still hold significant control over the network's resources. This observation supports Hypothesis 1, indicating that despite an increase in nodes, the influence of key entities continues to play a dominant role in the network's structure.

The node capacity results show a different picture, revealing a higher concentration of liquidity among fewer nodes. The Nakamoto coefficient begins at a much lower value – 13 at T1, as shown in Fig. 11. Over time, this number increases to 81 at T5, then drops to 27 until T8. These lower coefficients indicate that node capacity is more centralized, and fewer nodes are controlling a significant part of the network 's capacity. This concentration shows that there are risks to network stability if these key nodes fail or act maliciously. Nakamoto coefficient for entities shows the same trend as before – the progress is similar to nodes, but Nakamoto coefficient itself is lower. Research [75] indicated that the Nakamoto coefficient, alongside the Gini, revealed a significant concentration in BTC and Ethereum, where a small number of participants control over 51% of the wealth and this reveals that there is a security risk for these networks.

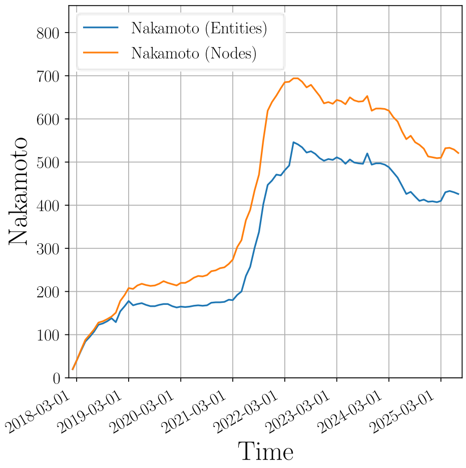

FIGURE 10. Nakamoto coefficient of BLN nodes on degree centrality.

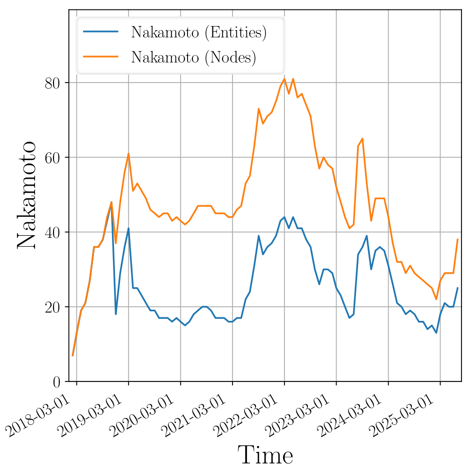

FIGURE 11. Nakamoto coefficient of BLN nodes on weighted degree centrality.

These results and differences between channel number and node capacity reveal that the channel infrastructure is becoming more distributed among nodes, but node capacity remains centralized.

3) HERFINDAHL-HIRSCHMAN INDEX

The Herfindahl-Hirschman index (HHI) for degree centrality starts at a high value of 153.3 at T1 (see Fig. 12), indicating a strong concentration of channels among a few nodes. But this value decreases significantly to 26.9 at T2 and remains stable at around 26 for the next two years, before declining further to 16.2 at T5 and increasing back to 26.9 up until T8. These results show that the network is becoming more decentralized in channel distribution, and it becomes more balanced over time with fewer nodes dominating the network. The HHI for entities shows a distinct pattern, with an extreme rise between T2 and T3, followed by a gradual decrease since and a slight increase again in T8. The peak in T3 may indicate the emergence of a large entity that temporarily raised the HHI; however, over time, the network became more balanced.

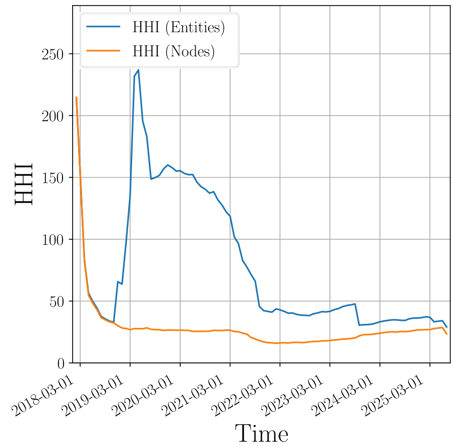

FIGURE 12. HHI of BLN nodes on degree centrality.

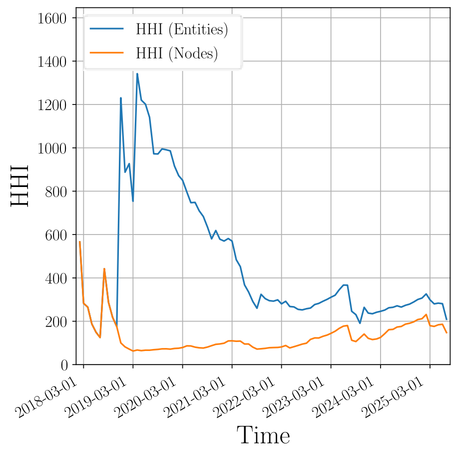

FIGURE 13. HHI of BLN nodes on weighted degree centrality.

The weighted degree centrality HHI values are higher, starting at 282.8 and decreasing to 62.7 in the following year, but then fluctuating across later timestamps as shown in Fig. 13.

The HHI value at T8 is 179.6. These higher values indicate that there is a concentration of liquidity in smaller groups of nodes. Fluctuation indicated that while some decentralization occurs, capacity control remains centralized in a limited number of dominant nodes. The HHI for entities also rises sharply at T2 and T3, then gradually decreases, confirming that while large influential entities initially emerged in the network, it became more balanced over time.

This pattern mirrors early observations in BTC, where [47] noticed that the sharp rise in HHI for BTC at the beginning of it was due to the small number of participants at that time. In the early phases, the Top 100 addresses controlled most BTC and the clustering of addresses within this group led to increased HHI values.

4) THEIL INDEX

For degree centrality, the Theil index starts at 0.146 at T1 and steadily increases to 0.183 at T4 (see Fig. 14), indicating a rising inequality in the number of channels per node as a higher value of Theil index signifies greater inequality in the distribution. Some nodes might be establishing significantly more channels than others. From T4 to T8 it slightly decreases, suggesting a minor reduction in inequality, which means that new nodes with fewer channels could be balancing the network. For entities, the normalized Theil index is higher than nodes – rising to over 0.22 between T2 and T4 - and its overall trend mirrors the HHI observed for nodes. This indicates that the network is less equal when assessed through clusters of nodes, highlighting a greater concentration of control among entity clusters.

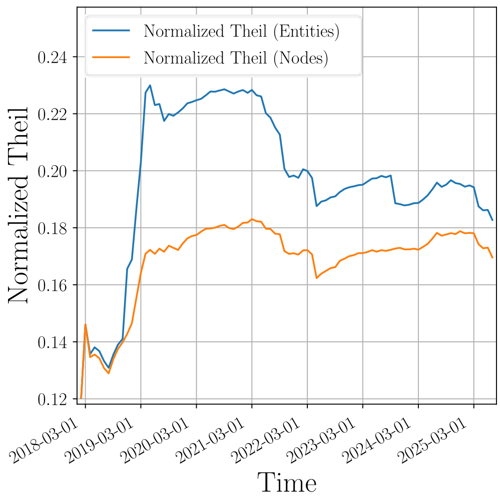

FIGURE 14. Theil index of BLN nodes on degree centrality.

In terms of weighted degree centrality, or node capacity, the Theil index values are higher, starting at 0.289 at T1 and increasing to 0.435 at T8 (see Fig. 15). This indicates that capacity is becoming more concentrated among fewer nodes and liquidity distribution is becoming centralized in a smaller number of nodes. The inequality in the distribution of node capacity is growing. For entities, the normalized Theil index is also higher, confirming similar findings as observed with degree centrality.

To sum up both degree and weighted degree centrality results, the Theil index for node capacity is consistently higher than for the number of channels, suggesting that capacity is more unevenly distributed than connectivity. The number of channels is becoming more balanced and at the same time, node capacity is becoming more concentrated among fewer nodes.

5) SHANNON ENTROPY

Normalized Shannon entropy for degree centrality values begins at 0.854 at T1 and fluctuates during the first half of the year (see Fig. 16). After this, there is a drastic decrease from T1 to T4, where entropy decreases from 0.854 to 0.817, indicating a slight increase in inequality in the distribution of the number of channels per node. From T4 to T8 entropy rises slightly to 0.822, suggesting a

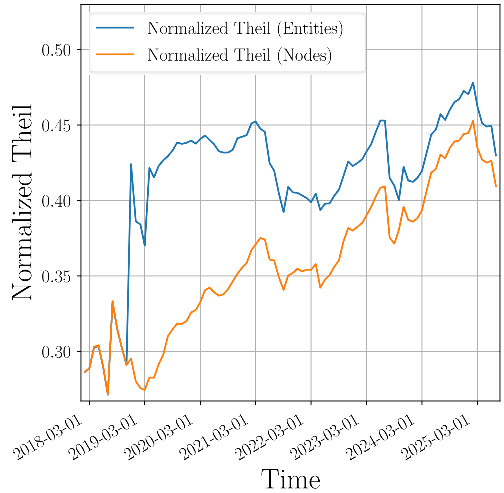

FIGURE 15. Theil index of BLN nodes on weighted degree centrality.

minor move toward more even distribution. The values of Shannon entropy remain consistently high, suggesting that the channel distribution is balanced throughout the different timestamps. For entities, the decrease between T1 and T2 is more gradual and balanced, indicating that, the distribution of channels among entities is less equal than among individual nodes.

In contrast, Normalized Shannon entropy for weighted degree centrality starts at a lower value of 0.711 at T1 and experiences similar fluctuations during the first year, rising to 0.726 at T2, then decreasing to 0.629 at T4 (see Fig. 17). After this decrease the entropy increases again and finally drops to 0.565 at T8. This lower overall value suggests that capacity distribution among nodes is more concentrated compared to channel distribution. The decline over time implies that, although there might be some decentralization occurring, it is not as pronounced as the changes seen in channel distribution. For entities, at T2, normalized Shannon entropy increases, while for nodes, there is a sharp decline. Afterward, normalized Shannon entropy for entities decreases gradually, indicating that capacity distribution among entities remains more balanced than among individual nodes.

In study [36] Shannon entropy was used to evaluate decentralization in various blockchain networks, with results highlighting that in 2018 the values for BTC and ETH were 11.33 and 10.38, respectively. By 2023, these results have shifted to 11.55 for BTC and 8.61 for ETH. In this current study, weighted degree centrality entropy for BLN at T6 (8.69) was lower than both BTC and ETH. This suggests that while channel distribution is becoming more decentralized, node capacity remains concentrated in fewer nodes, making BLN more centralized in terms of resource control compared to both BTC and ETH.

While the above metrics provide isolated insights, the following section synthesizes these results through summary metrics to better understand overall trends.

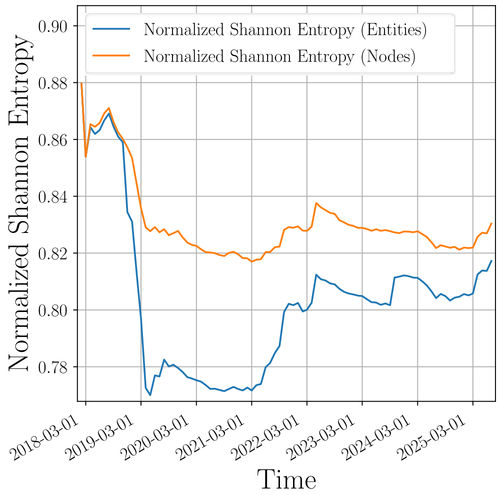

FIGURE 16. Shannon entropy of BLN nodes on degree centrality.

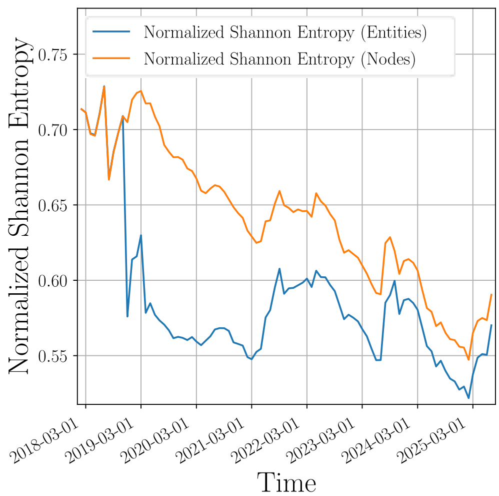

FIGURE 17. Shannon entropy of BLN nodes on weighted degree centrality.

V. COMPREHENSIVE SUMMARY OF CENTRALIZATION METRICS

The analysis of centralization in the BLN across various centrality metrics reveals a complex picture of the network 's evolution. While some metrics indicate that the network is becoming more decentralized, others suggest the opposite, highlighting the complexity of centralization within the system.

Table 2 provides a comprehensive summary of centralization metrics based on degree centrality for individual nodes across eight timestamps.

As shown in Table 2, the Gini coefficient for degree centrality increases from 0.598 at T1 to 0.764 by T5 and then slightly decreases at T8 (0.741), indicating a general rise in inequality in the distribution of channels among nodes over time. This suggests that some nodes are accumulating more channels compared to others, leading to a less equal network in terms of connectivity.

TABLE 2. Centrality metrics for number of channels (degree centrality) for individual nodes.

Centrality metric	T1	T2	T3	T4	T5	T6	T7	T8
Gini coefficient	0.598	0.746	0.758	0.758	0.764	0.759	0.736	0.741
Nakamoto coefficient	40	208	220	274	685	644	619	510
HHI	153.3	26.9	26.5	26.4	16.2	18.1	24.1	26.9
Normalized Theil index	0.146	0.164	0.177	0.183	0.172	0.171	0.172	0.178
Normalized Shannon entropy	0.854	0.836	0.823	0.817	0.828	0.829	0.828	0.822

However, when viewing centralization through entities, the control of the network held by key players becomes evident (see Table 3). For instance, the clustering of nodes by entities supports Hypothesis 1, pointing to a continued concentration of control despite minor fluctuations in Gini values as larger entities that control a significant portion of the network's capacity dominate.

TABLE 3. Centrality metrics for number of channels (degree centrality) for entities.

Centrality metric	T1	T2	T3	T4	T5	T6	T7	T8
Gini coefficient	0.598	0.756	0.770	0.774	0.774	0.769	0.748	0.753
Nakamoto coefficient	40	178	165	180	481	511	488	410
HHI	153.3	134.8	155.4	118.7	42.9	41.7	33.3	36.9
Normalized Theil index	0.146	0.203	0.225	0.229	0.200	0.195	0.189	0.194
Normalized Shannon entropy	0.854	0.797	0.775	0.772	0.800	0.805	0.811	0.806

The Nakamoto coefficient significantly increases at both levels, from 40 at T1 to 510 at T8 for nodes, and from 40 to 410 for entities. This increase implies that a larger number of nodes and entities are required to control 50% of the total channels over time, indicating decentralization in terms of control distribution. However, the consistently lower Nakamoto coefficient at the entity level suggests that fewer entities are needed to hold majority control compared to individual nodes. This discrepancy highlights that while the network appears more decentralized at the node level, entities consolidate control, leading to centralization at a higher level.

The HHI decreases dramatically at the node level, from 153.3 at T1 to 26.9 at T8, indicating reduced concentration among the top nodes and a more competitive environment. At the entity level, the HHI decreases initially but remains significantly higher than the node-level HHI throughout, ending at approximately 36.9 at T8. This suggests that concentration remains higher among entities, reflecting their influence in maintaining control despite apparent decentralization among individual nodes.

Both the Normalized Theil Index and Normalized Shannon Entropy exhibit minor fluctuations at both levels. The Theil

TABLE 4. Centrality metrics for node capacity (weighted degree centrality) for individual nodes.

Centrality metric	T1	T2	T3	T4	T5	T6	T7	T8
Gini coefficient	0.846	0.894	0.936	0.953	0.955	0.958	0.960	0.970
Nakamoto coefficient	13	61	43	44	81	52	44	27
HHI	282.8	62.7	78.9	109.7	81.5	144.3	125.4	179.6
Normalized Theil index	0.289	0.274	0.333	0.371	0.354	0.390	0.394	0.435
Normalized Shannon entropy	0.711	0.726	0.668	0.629	0.646	0.610	0.606	0.565

Index increases slightly more at the entity level, indicating greater inequality among entities compared to nodes. The Shannon Entropy decreases initially and then increases slightly, with lower values at the entity level, suggesting a less even distribution of channels among entities. These trends reinforce the notion that entities contribute to a more unequal and concentrated network structure.

Table 4 presents the centralization metrics based on weighted degree centrality (node capacity).

For node capacity, the Gini coefficient increases from 0.846 at T1 to 0.970 at T8, highlighting a significant rise in inequality in the distribution of capacities among nodes. This means that a small number of nodes are holding a larger proportion of the network's total capacity over time. Entity-based clustering results in Table 5 highlights this centralization effect even further, as entities increase their influence on network capacity.

The Nakamoto coefficient at the node level increases from 13 at T1 to 81 at T5, suggesting initial decentralization, but then decreases to 27 at T8, indicating a shift toward centralization. At the entity level, the Nakamoto coefficient remains lower throughout, decreasing from 44 at T5 to 18 at T8. This suggests that fewer entities are needed to control the majority of the network's capacity compared to individual nodes, highlighting the centralizing effect of entities and their capacity to consolidate control.

The HHI at the node level decreases initially but fluctuates, ending higher at T8 than at T2, indicating inconsistencies in concentration among top nodes. At the entity level, the HHI is significantly higher, peaking at T3, reflecting a high concentration of capacity among top entities. The elevated HHI values at the entity level demonstrate that entities contribute to a more concentrated distribution of capacity, reinforcing centralization trends.

Both the Normalized Theil Index and Normalized Shannon Entropy indicate growing inequality and uneven distribution of capacity at both levels. The Theil Index increases more sharply at the entity level, reaching 0.462 at T8 compared to 0.435 at the node level. The Shannon Entropy decreases over time, with lower values at the entity level, suggesting a less even distribution of capacity among entities. These metrics confirm that entities exacerbate inequality within the network.

For degree centrality, while the Gini coefficient indicates a trend toward higher inequality overall, with minor recent declines, the Nakamoto coefficient and HHI suggest increasing decentralization and decreasing concentration. These differences occurs because of the different aspects the metrics capture – the Gini coefficient measures inequality across the entire distribution, while the Nakamoto coefficient focuses on the minimum number of nodes controlling a majority and the HHI measures network‘s concentration. The HHI might decrease, suggesting less concentration among top nodes, even when the Gini coefficient indicates rising inequality. The influence of dominant nodes might reduce, while overall inequality still grows due to disparities among smaller nodes.

In weighted degree centrality, the increasing inequality shown by the Gini coefficient and Theil index aligns with the decreasing Shannon entropy, all pointing toward growing inequality. However, the fluctuations in the Nakamoto coefficient and HHI indicate inconsistencies in the trend towards centralization, highlighting the complexity of the network.

TABLE 5. Centrality metrics for node capacity (weighted degree centrality) for entities.

Centrality metric	T1	T2	T3	T4	T5	T6	T7	T8
Gini coefficient	0.846	0.901	0.939	0.956	0.951	0.959	0.959	0.971
Nakamoto coefficient	13	41	16	16	44	25	31	18
HHI	282.7	753.6	850.3	569.8	280.3	310.8	245.9	299.1
Normalized Theil index	0.289	0.370	0.441	0.452	0.399	0.432	0.420	0.462
Normalized Shannon entropy	0.711	0.630	0.559	0.548	0.601	0.568	0.580	0.538

As noted in study [75], a high Gini coefficient can coexist with a high Nakamoto coefficient. This means that while resources are unevenly distributed, control over the majority can still be spread across many nodes, creating an illusion of decentralization.

The observed trends support Hypothesis 1, which suggests that the BLN exhibits centralization around large entities, as clusters of nodes grouped by the entity hold a significant share of the network. The aggregation of nodes into entities reveals that entities consolidate control, leading to increased centralization.

Additionally, the results align with Hypothesis 2, confirming that centralization follows a non-linear pattern over time rather than consistent increase, with key nodes maintaining disproportionate influence. The consistent rise in inequality metrics at both levels indicates that resources are becoming more unevenly distributed, with entities playing a significant role in this process.

VI. DISCUSSION

This section interprets the results compared to the existing literature, outlines the study's contributions, discusses practical implications, and identifies future research directions.

This study investigated the centralization in the BLN using a range of centrality metrics, incorporating an entity-based analysis that highlights the influence of clusters of nodes. The findings provide a detailed view of the network's centralization dynamics, revealing both consistent and conflicting trends across different metrics, with entity clustering showing more concentrated control compared to node-based metrics alone. These results are contextualized through comparison with existing research on network centrality, highlighting implications for understanding control and influence in decentralized networks. Additionally, the study addresses methodological limitations and discusses how these insights may inform future approaches to assessing centralization in decentralized financial networks.

A. REFLECTION ON LITERATURE

The BLN has emerged as a second-layer solution designed to address Bitcoin's scalability challenges by enabling rapid and low-cost transactions through an off-chain framework. This architecture allows for the establishment of peer-to-peer channels where transactions occur off-chain, thus preserving the decentralized nature of Bitcoin [1], [2]. Despite these advantages, discussions on the degree of centralization within the BLN raise critical questions about the future of decentralization within this network. Findings from studies suggest a concerning trend, where a small subset of nodes may hold a disproportionate share of the network's total capacity, potentially compromising the decentralization goal. These observations relate directly to Hypothesis 1, which states that larger entities controlling specific nodes might increasingly dominate the network, potentially impacting decentralization. Current study's findings on node centralization dynamics in the BLN align closely with trends identified in existing literature. Like prior research indicating that roughly 10% of nodes hold 80% of the BLN's capacity [4], our analysis reveals that 10% of nodes control 96.5% of the network capacity, suggesting that more recent data indicate even more significant centralization.

The investigation into centralization metrics within the BLN reveals various aspects of node centrality that could influence control and power dynamics within the network. Degree centrality, for instance, identifies high-activity hubs that may exert significant influence through numerous connections [26], [27]. Weighted degree centrality, on the other hand, measures the liquidity or capacity of these connections, providing a nuanced understanding of node significance beyond simple connection counts [28]. These two metrics reveal different aspects of centralization, and the results of the study showed that centralization based on weighted degree centrality is more apparent than on degree centrality.

Several statistical tools provide a framework to measure and interpret inequality within the network. The Gini coefficient, a commonly used metric for inequality, suggests that higher values indicate greater centralization; studies have reported Gini coefficients as high as 0.88 for node capacity and 0.75 for channel distribution within the BLN, emphasizing a centralized configuration [4], [37]. Current study average results confirm with the previous research – 0.92 for node capacity and 0.73 for channel distribution. Relatedly, the Lorenz curve graphically represents this distribution of network resources, highlighting the disparity among nodes [11]. The Nakamoto coefficient, integrating insights from both the Gini coefficient and Lorenz curve, specifies the minimum number of entities required to control the network, thus quantifying decentralization risk [41], [42]. Additionally, the Theil index and Shannon entropy provide further insights: the Theil index evaluates inequality in node capacity and transaction volume distribution, while Shannon entropy reflects the uniformity of node capacity distribution, where higher entropy values suggest a more decentralized structure [1], [24], [48], [53].

Findings of the study also support Hypothesis 2, which suggests that centralization within the BLN may be gradually increasing over time. Metrics such as the Gini coefficient for node capacity and weighted degree centrality indicate a progressive trend toward more significant inequality, while fluctuations in the Nakamoto coefficient highlight the nuanced dynamics of this centralization process. These metrics collectively offer a comprehensive view of centralization trends within the BLN, and their implications suggest both challenges and considerations for the network's development and sustainability.

B. CONTRIBUTION ON THE RESEARCH

This research contributes to the understanding of centralization dynamics in the BLN by analyzing multiple inequality and centralization. In line with Hypothesis 1, which highlights that larger entities controlling specific nodes may dominate the network, findings indicate that increasing capacity inequality could indeed pose centralization risks as a small number of nodes accumulate a larger proportion of the network's total capacity. Additionally, consistent with Hypothesis 2 - that centralization in the BLN is gradually increasing over time - the study observes growing trends in capacity inequality, underscoring potential centralization over time despite some metrics showing signs of decentralization.

Unlike previous research that largely presented centralization as uniformly strong and increasing, this study highlights a more nuanced insight. The introduction of entity-level clustering further differentiates this study by analyzing previously unrecognized centralized structures at the entity level, offering insights that extend beyond conventional node-level analysis. These nuanced findings show the importance of adopting multiple metrics and perspectives to accurately assess centralization dynamics, contributing meaningful advancements to both academic research and practical strategies for maintaining the decentralized integrity of the BLN.

By differentiating between degree centrality and weighted degree centrality, the study reveals that while inequality is growing in both aspects, the impact on decentralization is complex, with increasing inequality in connectivity accompanied by enhanced decentralization (as control spreads across more nodes), whereas increasing capacity inequality poses potential centralization risks due to capacity concentration among fewer nodes. This analysis highlights the importance of using multiple metrics to capture different dimensions of centralization, uncovers nuanced insights into the network's structural changes, contributing valuable insights that inform network development, policy-making, and future research to ensure the BLN remains secure, efficient, and aligned with its decentralized principles.

C. IMPLICATIONS FOR PRACTICE

The practical implications of this research emphasize the need for proactive measures to address growing inequality and potential centralization within the BLN. Consistent with Hypothesis 1, which suggests that the centralization of the BLN may increase as larger entities controlling specific nodes dominate the network, it is essential for network participants to regularly monitor inequality and centralization trends using metrics like the Gini coefficient and Theil index to detect early signs of centralization [1], [4]. This monitoring can reveal shifts in power dynamics, allowing for timely interventions.

Moreover, in line with Hypothesis 2, that centralization in the BLN is gradually increasing over time, encouraging an equal distribution of capacity is crucial. This can be achieved by incentivizing smaller nodes, supporting new users, and promoting diverse connectivity to help maintain decentralization [76]. Additionally, enhancing network security by mitigating single points of failure and implementing redundancy measures is vital in reducing risks associated with high-capacity nodes that may increase centralization [11].

Adjusting network policies and protocols to balance efficiency with decentralization, educating the community about the importance of decentralization, and fostering collaborative efforts among stakeholders are also critical steps. By implementing these practices, the BLN could effectively address the challenges posed by centralization and ensure security, efficiency, and alignment with the foundational principles of blockchain technology.

D. LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

This study identifies several limitations in assessing centralization within the BLN. A key limitation is the use of static data rather than dynamic data, which may not fully capture the network's evolvement over time. Although March of each year was selected as a reference point for presenting results in figures, the full analysis relies on monthly network snapshots from March 2018 to March 2025, minimizing sampling bias. The inherent complexity of the network further complicates efforts to accurately cluster nodes into entities and determine the exact number of active channels.

Moreover, the alias-based clustering method used in this study assumes that nodes create their aliases based on their entity name, which may not always be accurate. This manual approach could introduce errors and biases. Future research could benefit from exploring alternative clustering methods, to more accurately group entities.

While a comparative analysis of centrality metrics has been conducted, this study focused only on degree centrality and weighted degree centrality. Other important measures like betweenness, closeness, and eigenvector centrality were not quantified, because they are based on routing and do not consider channel capacities. Future research should include these additional centrality aspects to provide a more complete understanding of the network's structure.

VII. CONCLUSION

This research concludes that the centralization dynamics within the BLN are complex and multifaceted. While some centrality metrics indicate increasing centralization, others suggest a potential for decentralization, revealing nuanced results of network centralization. The study highlights the importance of employing diverse metrics to accurately assess centralization and emphasizes the need for entity-based analysis to uncover the influence of grouped nodes. These findings provide critical insights for stakeholders seeking to enhance the resilience and decentralization of the BLN, guiding future strategies and policy development.

DATA AND CODE AVAILABILITY

Data and code used in this study are openly accessible. The scripts used for data collection are available at VUKNF-Fintech-Research-Group/lnd-dbreader. The repository containing the code for calculating the centrality metrics and creating the graphs is available at VUKNF-Fintech-Research-Group/blnstats. Additionally, a database of all the results is accessible at blnstats.knf.vu.lt.

REFERENCES

[1] S. Martinazzi and A. Flori, "The evolving topology of the lightning network: Centralization, efficiency, robustness, synchronization, and anonymity," *PLoS ONE*, vol. 15, no. 1, Jan. 2020, Art. no. e0225966.

[2] S. Bartolucci, F. Caccioli, and P. Vivo, "A percolation model for the emergence of the Bitcoin lightning network," *Sci. Rep.*, vol. 10, no. 1, p. 4488, 2020.

[3] A. Howell, T. Saber, and M. Bendechache, "Measuring node decentralisation in blockchain peer to peer networks," *Blockchain, Res. Appl.*, vol. 4, no. 1, Mar. 2023, Art. no. 100109.

[4] J.-H. Lin, K. Primicerio, T. Squartini, C. Decker, and C. J. Tessone, "Lightning network: A second path towards centralisation of the Bitcoin economy," *New J. Phys.*, vol. 22, no. 8, Aug. 2020, Art. no. 083022.

[5] S. Tikhomirov, P. Moreno-Sanchez, and M. Maffei, "A quantitative analysis of security, anonymity and scalability for the lightning network," in *Proc. IEEE Eur. Symp. Secur. Privacy Workshops (EuroS&PW)*, Sep. 2020, pp. 387–396.

[6] G. Kappos, H. Yousaf, A. Piotrowska, S. Kanjalkar, S. Delgado-Segura, A. Miller, and S. Meiklejohn, "An empirical analysis of privacy in the lightning network," Tech. Rep., 2021.
[7] D. Khan, L. T. Jung, and M. A. Hashmani, "Systematic literature review of challenges in blockchain scalability," *Appl. Sci.*, vol. 11, no. 20, p. 9372, Oct. 2021.
[8] G. Tripathi, M. A. Ahad, and G. Casalino, "A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges," *Decis. Anal. J.*, vol. 9, Dec. 2023, Art. no. 100344.
[9] A. Kurt, S. Mercana, E. Erdin, and K. Akkaya, "Enabling micro-payments on IoT devices using Bitcoin lightning network," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2021, pp. 1–3.
[10] A. Kurt, S. Mercan, E. Erdin, and K. Akkaya, "3-of-3 multisignature approach for enabling lightning network micro-payments on IoT devices," Tech. Rep., 2021.
[11] P. Zabka, K.-T. Förster, C. Decker, and S. Schmid, "A centrality analysis of the lightning network," *Telecommun. Policy*, vol. 48, no. 2, Mar. 2024, Art. no. 102696.
[12] F. Beres, I. A. Seres, and A. A. Benczur, "A cryptoeconomic traffic analysis of Bitcoin's lightning network," Tech. Rep., 2020.
[13] R. Pickhardt and M. Nowostawski, "Imbalance measure and proactive channel rebalancing algorithm for the lightning network," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2020, pp. 1–5.
[14] J. Harris and A. Zohar, "Flood & loot: A systemic attack on the lightning network," in *Proc. 2nd ACM Conf. Adv. Financial Technol.*, Oct. 2020, pp. 202–213.
[15] V. Sivaraman, S. B. Venkatakrishnan, M. Alizadeh, G. Fanti, and P. Viswanath, "Routing cryptocurrency with the spider network," in *Proc. 17th ACM Workshop Hot Topics Netw.*, Nov. 2018, pp. 29–35.
[16] G. Danezis and S. Meiklejohn, "Centrally banked cryptocurrencies," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 2016.
[17] M. S. Ismail, S. I. Hussain, and M. S. M. Noorani, "Detecting early warning signals of major financial crashes in Bitcoin using persistent homology," *IEEE Access*, vol. 8, pp. 202042–202057, 2020.
[18] D. Grossi, "Social choice around the block: On the computational social choice of blockchain," Tech. Rep., 2022.
[19] A. Carotti, C. Sguanci, and A. Sidiropoulos, "Rational economic behaviours in the Bitcoin lightning network," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2024, pp. 594–601.
[20] S. Masteika, E. Rebždys, K. Driaunys, A. Šapkauskienė, A. Mačerinskienė, and E. Krampas, "Bitcoin double-spending risk and countermeasures at physical retail locations," *Int. J. Inf. Manage.*, vol. 79, Dec. 2024, Art. no. 102727.
[21] F. Waugh and R. Holz, "An empirical study of availability and reliability properties of the Bitcoin lightning network," 2020, *arXiv:2006.14358*.
[22] Z. Wang, R. Zhang, Y. Sun, H. Ding, and Q. Lv, "Can lightning network's autopilot function use BA model as the underlying network?" *Frontiers Phys.*, vol. 9, Jan. 2022, Art. no. 794160.
[23] J. Poon and T. Dryja, "The Bitcoin lightning network: Scalable off-chain instant payments," Tech. Rep., 2016.
[24] A. Divakaruni and P. Zimmerman, "The lightning network: Turning Bitcoin into money. Federal reserve bank of Cleveland," Tech. Rep., 2022.
[25] M. S. Mahdizadeh, B. Bahrak, and M. S. Haghighi, "Decentralizing the lightning network: A score-based recommendation strategy for the autopilot system," *Appl. Netw. Sci.*, vol. 8, no. 1, pp. 33–73, Oct. 2023.
[26] N. Meghanathan, "Clique size and centrality metrics for analysis of real-world network graphs," in *Encyclopedia of Information Science and Technology*, 4th ed., Hershey, PA, USA: IGI Global, 2018, pp. 6507–6521.
[27] S. Gómez, "Centrality in networks: Finding the most important nodes," in *Business and Consumer Analytics: New Ideas*. Cham, Switzerland: Springer, 2019, pp. 401–433.
[28] T. Opsahl, F. Agneessens, and J. Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths," *Social Netw.*, vol. 32, no. 3, pp. 245–251, Jul. 2010.
[29] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *J. Math. Sociol.*, vol. 2, no. 1, pp. 113–120, Jan. 1972.
[30] L. Atmanavičiūtė, T. Vanagas, and S. Masteika, "Method for determining the level of centralization in BTC lightning nodes: A centrality analysis of the lightning network," *Vilnius Univ. Open Ser.*, pp. 6–14, May 2024.
[31] M. Lischke and B. Fabian, "Analyzing the Bitcoin network: The first four years," *Future Internet*, vol. 8, no. 1, p. 7, Mar. 2016.
[32] C. Gini, "Measurement of inequality of incomes," *Econ. J.*, vol. 31, no. 121, pp. 124–126, Mar. 1921.
[33] M. K. Florian, N. Li, and M. D. Gladders, "The Gini coefficient as a morphological measurement of strongly lensed galaxies in the image plane," *Astrophys. J.*, vol. 832, no. 2, p. 168, Dec. 2016.
[34] R. Grochowska, A. Pawłowska, and A. Skarżyńska, "Searching for more balanced distribution of direct payments among agricultural farms in the CAP post-2020," *Agricult. Econ. (Zemědělská ekonomika)*, vol. 67, no. 5, pp. 181–188, May 2021.
[35] X. Fan, W. Zhang, Y. Guo, J. Cai, and B. Xie, "Equity assessment of the distribution of mental health beds in China: Based on longitudinal data from 2011 to 2020," *BMC Health Services Res.*, vol. 22, no. 1, Nov. 2022.
[36] M. Juodis, E. Filatovas, and R. Paulavičius, "Overview and empirical analysis of wealth decentralization in blockchain networks," *ICT Exp.*, vol. 10, no. 2, pp. 380–386, Apr. 2024.
[37] J.-H. Lin, E. Marchese, C. J. Tessone, and T. Squartini, "The weighted Bitcoin lightning network," *Chaos, Solitons Fractals*, vol. 164, Nov. 2022, Art. no. 112620.
[38] T. Sitthiyot and K. Holasut, "A simple method for estimating the Lorenz curve," *Humanities Social Sci. Commun.*, vol. 8, no. 1, Nov. 2021.
[39] F. Buscemi and G. Gour, "Quantum relative Lorenz curves," *Phys. Rev. A, Gen. Phys.*, vol. 95, no. 1, Jan. 2017.
[40] W. H. Chiu, "Intersecting Lorenz curves and aversion to inverse downside inequality," *Social Choice Welfare*, vol. 56, no. 3, pp. 487–508, Apr. 2021.
[41] B. Kusmierz and R. Overko, "How centralized is decentralized? Comparison of wealth distribution in coins and tokens," in *Proc. IEEE Int. Conf. Omni-Layer Intell. Syst. (COINS)*, Aug. 2022, pp. 1–6.
[42] L. Zhang, X. Ma, and Y. Liu, "Sok: Blockchain decentralization," Tech. Rep., 2023.
[43] J. Rój, "Competition measurement of hospitals in Poland: The Herfindahl–Hirschman index approach," *Ekonomika*, vol. 95, no. 1, pp. 166–181, Apr. 2016.
[44] S. Herwald, S. Voigt, and A. Uhde, "The impact of market concentration and market power on banking stability–evidence from Europe," *J. Risk Finance*, vol. 25, no. 3, pp. 510–536, Apr. 2024.
[45] A. C. Weber, V. De Bosscher, and H. Kempf, "Positioning in Olympic winter sports: Analysing national prioritisation of funding and success in eight nations," *Eur. Sport Manage. Quart.*, vol. 18, no. 1, pp. 8–24, Jan. 2018.
[46] I. Pavic, F. Galetic, and D. Piplica, "Similarities and differences between the CR and HHI as an indicator of market concentration and market power," *Brit. J. Econ., Manage. Trade*, vol. 13, no. 1, pp. 1–8, Jan. 2016.
[47] L. Cheng, F. Zhu, H. Liu, and C. Miao, "On decentralization of Bitcoin: An asset perspective," 2021, *arXiv:2105.07646*.
[48] H. Theil and A. Finezza, "A note on the measurement of racial integration of schools by means of informational concepts," *J. Mathematical Sociol.*, vol. 1, no. 1, pp. 187–194, 1971.
[49] L. Zhang, Z. Yang, J. Liang, and Y. Cai, "Spatial variation and distribution of urban energy consumptions from cities in China," *Energies*, vol. 4, no. 1, pp. 26–38, Dec. 2010.
[50] K. Yang, Y. Shi, Y. Luo, D. Xia, and X. Zhou, "The spatiotemporal characteristics of Chinese civil vehicles' possession in the context of rapid economic development from 1996 to 2015," *Sustainability*, vol. 10, no. 9, p. 2999, Aug. 2018.
[51] W. Pan, H. Fu, and P. Zheng, "Regional poverty and inequality in the Xiamen-Zhangzhou-Quanzhou city cluster in China based on NPP/VIIRS night-time light imagery," *Sustainability*, vol. 12, no. 6, p. 2547, Mar. 2020.
[52] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
[53] D. K. Yoon, J. E. Kang, and Y. J. Kim, "Measuring community vulnerability to natural disasters in Korea using the entropy method," *Int. J. Mass Emergencies Disasters*, vol. 35, no. 2, pp. 8–29, Aug. 2017.
[54] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, "Chain rules for smooth Min- and max-entropies," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2603–2612, May 2013.
[55] S. Lahmiri and S. Bekiros, "Chaos, randomness and multi-fractality in Bitcoin market," *Chaos, Solitons Fractals*, vol. 106, pp. 28–34, Jan. 2018.
[56] I. García and J. A. Molina, "The effects of region on the welfare and monetary income of Spanish families," *Urban Stud.*, vol. 38, no. 13, pp. 2415–2424, Dec. 2001.
[57] J. Creedy, "The Atkinson inequality measure and inequality aversion," Victoria Univ. Wellington, Tech. Rep., 2019.

[58] C. Giatsidis, F. D. Malliaros, N. Tziortziotis, C. Dhanjal, E. Kiagias, D. M. Thilikos, and M. Vazirgiannis, "A k-core decomposition framework for graph clustering," Tech. Rep., 2016.
[59] I. Malvestio, A. Cardillo, and N. Masuda, "Interplay between k-core and community structure in complex networks," *Sci. Rep.*, vol. 10, no. 1, Sep. 2020.
[60] J. He, "The Google PageRank algorithm," in *Proc. 2nd Int. Academic Conf. Blockchain Inf. Technol. Smart Finance (ICBIS)*, 2023, pp. 417–422.
[61] M. Lin, W. Li, L. J. Song, C. Nguyen, X. Wang, and S. Lu, "Sake," *ACM Trans. Knowl. Discovery From Data*, vol. 15, no. 4, pp. 1–21, 2021.
[62] Q. Lin, C. Li, X. Zhao, and X. Chen, "Measuring decentralization in Bitcoin and Ethereum using multiple metrics and granularities," in *Proc. IEEE 37th Int. Conf. Data Eng. Workshops (ICDEW)*, Apr. 2021, pp. 80–87.
[63] I. Bhattacharya and L. Getoor, "Collective entity resolution in relational data," *ACM Trans. Knowl. Discovery From Data*, vol. 1, no. 1, p. 5, Mar. 2007.
[64] P. Moreno-Sanchez, M. B. Zafar, and A. Kate, "Listening to whispers of ripple: Linking wallets and deanonymizing transactions in the ripple network," *Proc. Privacy Enhancing Technol.*, vol. 2016, no. 4, pp. 436–453, Oct. 2016.
[65] Lightning Labs. *Lnd–channel Database*. GitHub. Accessed: Nov. 4, 2024.
[66] *Neutrino: Privacy-Preserving Bitcoin Light Client*, GitHub, Lightning Labs, Palo Alto, CA, USA.
[67] B. Johnson. *Boltdb: An Embedded Key/Value Database for Go*. Accessed: Dec. 20, 2024. [Online]. Available: https://github.com/boltdb/bolt
[68] C. Decker. (2023). *Lightning Network Gossip*. Accessed: Jan. 12, 2025.
[69] I. A. Seres, L. Gulyás, D. A. Nagy, and P. Burcsi, "Topological analysis of Bitcoin's lightning network," in *Proc. 1st Int. Conf. MARBLE Math. Res. Blockchain Economy*, Santorini, Greece. Cham, Switzerland: Springer, 2019, pp. 1–12.
[70] D. Thakkar, S. Sabale, and A. Waghmare, "Exploring the efficiency of off-chain vs. on-chain transactions in blockchain network," *Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol.*, vol. 10, no. 3, pp. 40–44, May 2024.
[71] *Introduction to MyNode*, MyNode Documentation, MyNode, Huntsville, AL, USA.
[72] *Getting Started With Prefect 3.0*, Prefect Documentation, Prefect Technologies, Washington, DC, USA.
[73] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 4, pp. 133–145, Oct. 2002.
[74] Lightning Labs. (2018). *Lightning Labs Blog—The Official Blog of Lightning Labs*. Accessed: Jan. 12, 2025.
[75] A. R. Sai, J. Buckley, and A. L. Gear, "Characterizing wealth inequality in cryptocurrencies," *Frontiers Blockchain*, vol. 4, Dec. 2021.
[76] A. Biryukov, G. Naumenko, and S. Tikhomirov, "Analysis and probing of parallel channels in the lightning network," in *Financial Cryptography and Data Security*. Cham, Switzerland: Springer, 2022, pp. 337–357.

LAURA ATMANAVIČIŪTĖ (Member, IEEE) received the master's degree in financial technology from Vilnius University, where she is currently pursuing the Ph.D. degree in informatics engineering, where she focuses on zero-knowledge proofs and the adaptation of blockchain technology. She was a Senior Laboratory Assistant and a Researcher with the FinTech, Blockchain, and Cryptocurrency Research Group, contributing to studies on network centralization metrics and cryptocurrency ecosystems. She was also with the FinTech study Program Committee, Contributing to Program Development in Financial Technology. Prior to her current academic role, she gained extensive experience in accounting and process automation, implementing robotic process automation. Her research interests include decentralized technologies, advanced cryptographic methods, and the intersection of finance and technology.

TOMAS VANAGAS is currently pursuing the bachelor's degree in information systems and cyber security with Vilnius University Kaunas Faculty, focusing on blockchain and cryptocurrency technologies. He has over six years of experience in system administration and decentralized systems, overseeing critical IT infrastructures, containerized application clusters, and monitoring solutions in his role as a Systems Administrator. He also serves as a Laboratory Assistant, designing lab tasks on cryptocurrencies, blockchain mechanics, and second-layer solutions, such as the Lightning Network. Prior to this, he gained a robust technical foundation as an Electrician-Technician and an Electrician-Automatician, working with electrical control systems, automation, and PLC programming. His research interests include the intersection of cybersecurity and decentralized technologies.

SAULIUS MASTEIKA received the master's degree in financial technologies. He is currently a Full Time Professor and the Head of Fintech, Blockchain, and Crypto Currency Research Group with the Institute of Social Sciences and Applied Informatics, Kaunas Faculty, Vilnius University. He has authored over 40 scientific papers, including 25 publications in Clarivate Analytics Web of Science journals with impact factor and WoS conference proceedings. He has been an Associated Partner and a Principal Investigator in several international research projects, including the Marie Curie Actions EU Project "Training in Modern Quantitative Methods and High-Performance Computing for Finance" (FP7-PEOPLE), "Blockchain Network Online Education for Interdisciplinary European Competence Transfer" (LMT, No. LT01-KA203-047044), "Creation of Oscillations-Based Paradigm for Simulation of Social Systems Behavior" (ŠMM, No. K-01-137), and "Implementation of a Blockchain-Based Payment System for Encouraging the Use of Tourism Services after the COVID-19 Pandemic" (LMT, No. K-718-05-0006). He leads the Research and Development Group developing digital identity solutions using blockchain technologies. His research interests include computational finance, blockchain technologies, information systems, algorithmic and high-frequency trading, and hedging strategies in crypto markets. He has served as an Expert for the Agency for Science, Innovation, and Technology (MITA) and the Lithuanian Business Support Agency (LVPA) in blockchain, fintech, and artificial intelligence. He is the Founder and the Chair of the Scientific Committee for Lithuania's first master's degree, specializing in research on information systems, which has been running for eight years. He represents the university in national and international institutions shaping science and innovation policy, significantly advancing fintech and blockchain research.