A Graph-Theoretical Analysis of the Lightning Network

Published Version:
Samoggia, S., Ferretti, S., Serena, L., Marzolla, M., D'Angelo, G. (2025). A Graph-Theoretical Analysis of the Lightning Network. New York : IEEE [10.1109/brains67003.2025.11302920].

Availability:
This version is available at: https://hdl.handle.net/11585/1039250 since: 2026-01-26

Published:
DOI: http://doi.org/10.1109/brains67003.2025.11302920

A Graph-Theoretical Analysis of the Lightning Network

1st Simone Samoggia
University of Bologna
Bologna, Italy
simone.samoggia2@studio.unibo.it

2nd Stefano Ferretti
University of Bologna
Bologna, Italy
s.ferretti@unibo.it

3rd Luca Serena
University of Bologna
Bologna, Italy
luca.serena2@unibo.it

4th Moreno Marzolla
University of Bologna
Bologna, Italy
moreno.marzolla@unibo.it

5th Gabriele D'Angelo
University of Bologna
Bologna, Italy
g.dangelo@unibo.it

***Abstract*—The Lightning Network is the most adopted layer-2 solution for Bitcoin, shifting many transactions off-chain through bidirectional channels to improve scalability and reduce costs. Despite its growing usage, few studies have been conducted to investigate its structure and how it evolved over time. In this work, we analyze the structural properties of the public Lightning Network using graph theory, collecting both real-time and historical data. We have then considered multiple filtering strategies to isolate different parts of the network, such as excluding single-connection nodes or central hubs. The results show that as the network has grown, its average connectivity has decreased, and its reliance on hubs has increased. This structure enables efficient routing, but also conceals some issues concerning the network resilience in case of hub failures.**

I. Introduction

Bitcoin, introduced in 2008, proved that a decentralized peer-to-peer network could reach consensus without any central authority [1]. Since the future success of cryptocurrencies was unimaginable at that time, Bitcoin's architecture was not designed to address issues such as scalability and transaction costs. In particular, Bitcoin's throughput is inherently limited by block size and interval [2]; as adoption grew, congestion caused higher fees and delays, making it unsuitable for everyday payments [3]. Despite the rise of several Distributed Ledger Technologys (DLTs), Bitcoin has retained its leading position in the crypto economy with over 60% of market capitalization [4]. To improve usability without changing the core protocol, layer-2 solutions were introduced, handling most transactions off-chain to reduce load while preserving Bitcoin's security. In particular, the Lightning Network (LN) enables bidirectional payment channels for routing transactions, providing fast and low-cost payments that significantly increase the throughput while preserving Bitcoin's security guarantees. While the LN solves some performance issues, it also introduces new questions. Unlike Bitcoin's base layer, which offers full transparency and a clear record of transactions, the Lightning Network does not provide an easy-to-access global view: nodes discover the network via direct links or gossip, while many channels remain private. However, the portion of the network that is retrieved corresponds to the relevant part for our analysis, since transaction routing takes place over publicly announced channels. In this study, graph theory is used to model the LN as a set of nodes and edges, capturing the relationships between participants through specific metrics. To conduct the analysis, we collected live public data from a Lightning node, combining such information with historical data, in order to analyze both the current network and its evolutions over the last years. Furthermore, we applied some filters, excluding the nodes with the most connections, with the aim of investigating the resilience of the network in case of hub unavailability. Our findings show that, as the network has grown, its average connectivity has decreased and its reliance on hubs has increased; failures or attacks targeting these hubs would result in significant fragmentation. Furthermore, the network still shows small-world characteristics, though these appear to have weakened over time. The results highlight the strengths and vulnerabilities of the current structure of the LN, and contribute to the broader discussion of what a decentralized, scalable payment network should look like.

II. Background

A. Graphs and Complex Systems

Graphs are data structures that are commonly used to represent networks. Specifically, a graph is a tuple $G = (V, E)$, where V is a set of N nodes, and $E \subseteq V \times V$ is a set of edges that connect pairs of nodes. Edges represent relationships of various types between nodes, and can be directed or undirected, depending on whether the relationship

M.M. is partially supported by the Istituto Nazionale di Alta Matematica "Francesco Severi" - Gruppo Nazionale per il Calcolo Scientifico (INdAM-GNCS) and by the ICSC National Research Centre for High Performance Computing, Big Data and Quantum Computing within the NextGenerationEU program - CUP: J33C22001170001. G.D'A. S.F. and L.S. are partially supported by the European Union - NextGenerationEU within the framework of NRRP Mission 4 - Component 2 - Investment 1.1 under the Italian Ministry of University and Research (MUR) programme "PRIN 2022" - grant number 2022N2NH42 SmartShires – CUP: H53D23003570006.

is one-way or mutual. In this paper we are concerned with undirected graphs, so the edges (u, v) and (v, u) are the same. A *path* $p = (v_0, \ldots, v_k)$ is a sequence of nodes where each consecutive pair is connected by an edge. An undirected graph is *connected* if a path exists between any two nodes; otherwise, it can be decomposed into multiple *connected components*, with the largest commonly referred to as the "main component".

Several metrics have been defined to analyze the structure of a graph; in this paper, we consider the following [5]:

- *Average Node Degree (AND)*, the average node degree.
- *Average Path Length (APL)*, the average number of edges in the shortest paths between all pairs of connected nodes.
- *Average Clustering Coefficient (ACC)*, the average fraction of a node's neighbors that are also directly connected to each other.
- *Betweenness centrality*, a measure of a node's importance, defined by how frequently it lies on the shortest paths between other pairs of nodes.
- *Network Centrality (NC)*, a measure of how evenly the degree centrality is distributed across the network. The metric is computed with the following formula [6]:

$$NC = \frac{\sum_{v \in V} (d_{\max} - \deg(v))}{(N-1)(N-2)} \tag{1}$$

where $d_{\max} = \max_{v \in V} \deg(v)$ is the maximum degree.

These metrics make it possible to classify graphs into different topologies, based on structural patterns and connectivity features. *Random graphs* are generated according to a probability distribution, such as the Erdős–Rényi model, where connections are generated by randomly selecting pairs of nodes. *Small world* graphs are characterized by a high clustering coefficient and a low APL, meaning that nodes tend to form tightly connected local groups, while still being able to reach nodes outside the cluster in a relatively "small" number of hops. A graph exhibits small-world properties if the small-world coefficient σ is greater than one, where:

$$\sigma = \left(\frac{ACC}{ACC_{\text{rand}}}\right) \Big/ \left(\frac{APL}{APL_{\text{rand}}}\right) \tag{2}$$

and ACC_{rand} and APL_{rand} are the clustering coefficient and average path length, respectively, of a random graph with the same number of nodes and edges.

Finally, in *scale-free* graphs, the degree distribution of nodes follows a power law, according to the following formula:

$$P(k) \sim k^{-\alpha}, \qquad k \geq x_{\min} \tag{3}$$

where $P(k)$ is the frequency of nodes with degree k, $x_{\min}$ is the minimum degree from which the power-law behavior holds, and α is the power law exponent, representing the form of the distribution.

B. Layer-2 Solutions

The LN, introduced in 2018, is a layer-2 protocol designed to address limitations in terms of scalability and transaction costs inherent in Bitcoin [7]. The idea is quite simple: two parties open a channel by committing a certain amount of bitcoins in an on-chain transaction. From that point on, they can send funds back and forth instantly and at negligible cost by updating the channel's balance, essentially signing off-chain commitments that reflect how the funds should be split if the channel were closed at that moment.

However, while the technical design is decentralized, the network structure that has emerged in practice does not respect that feature. Over time, certain nodes —especially those run by custodial services or large routing operators— have accumulated many channels and now sit in central positions within the network. These nodes handle a significant portion of the total routing activity, which raises questions about how decentralized the system really is when viewed as a graph. If one or more of these high-degree nodes went offline, the ability to route payments could be heavily impacted.

III. Results and Analysis

A. Network Setup

Data collection from the LN is challenging since, unlike Bitcoin's base layer, it is only partially observable [8]. We collected public topology snapshots using `lncli describegraph` on a Lightning node [9] anchored to a Bitcoin Core full node [10]. Although incomplete, this view is generally considered a reasonable approximation, since private channels cannot be used for routing between arbitrary nodes and only affect local connectivity between their endpoints [11]. On the other hand, for historical analysis we used the TimeMachine tool [12], which allowed us to reconstruct past topologies from archived gossip data.

The primary dataset consists of a graph snapshot taken on February 6, 2025, containing 16 688 nodes and 50 565 channels. In addition, we used TimeMachine to generate historical network topologies from 2019 through 2024, allowing us to perform a longitudinal analysis of several key metrics to assess how the network structure has changed over time. Four types of filters have been applied to the network: *unfiltered*, including all publicly visible nodes and channels; *min_2*, excluding nodes with degree 1; *top10* and *top100*, which are extensions of the *min_2* filter that additionally remove the 10 and 100 nodes with higher degree, respectively.

B. Analysis on 2025 Snapshot

The first question we have investigated is whether the LN exhibits "small-world" properties, which is a common pattern in many real-world networks, and is typically taken as a sign of efficient information flow with localized structure [13]. To this aim, we first computed the metrics for a random graph of a similar size, using the formulas provided in [14]. The APL and ACC of a random graph with n nodes and m edges are:

$$\text{APL}_{\text{rand}} \approx \frac{\ln n}{\ln(2m/n)}, \qquad \text{CC}_{\text{rand}} \approx \frac{2m}{n^2}. \tag{4}$$

Then, applying Eq. (2) we obtain $\sigma \approx 479$, meaning that the graph exhibits strong small-world behavior.

To determine whether the network exhibits scale-free behavior, we plotted the node degree distribution on a log-log scale. Although the power-law fit identifies $x_{\min}$ and $\alpha = 2.367$, suggesting that scale-free behavior applies to nodes with high degree, the degree distribution at these values appears noisy, as seen in Figure 1. This is likely due to the small number of hubs present in the network, making statistical fluctuations more visible in the upper tail of the distribution.

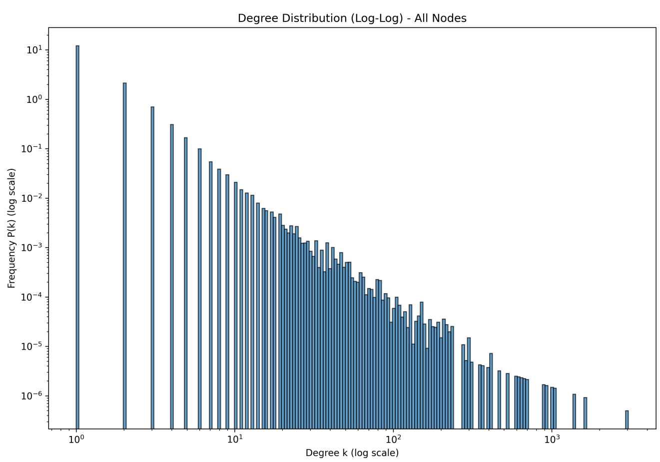

Fig. 1: Degree Distribution of the 2025 snapshot.

The analysis of the betweenness centrality, shown in Figure 2, highlights once again the influence of the few hubs, which turn out to be key intermediaries in routing across the network. The hubs are public routing nodes that are typically run by companies that provide services for the LN.

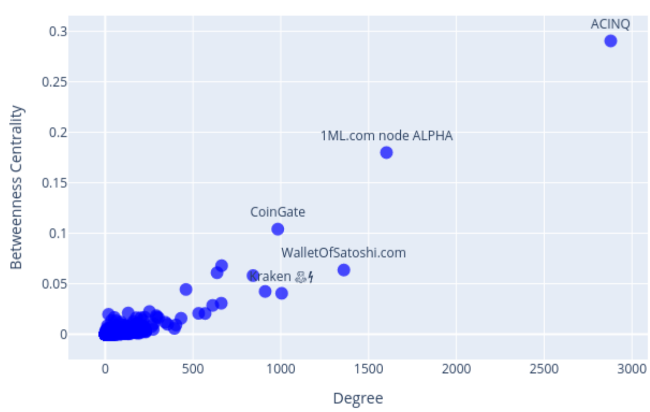

Fig. 2: Relationship between node degree and betweenness centrality on the 2025 snapshot.

C. Analysis over Time

In this section, we show the main metrics of the Lightning Network over the 2019–2025 period. As shown in Figure 3, the network exhibits a rapid growth between 2019 and 2022, especially in the number of channels. After that, the network seems to stabilize, even though the drop in channels from 2022 to 2025 is still significant. This change is reflected in the AND as well, which falls consistently over time, suggesting that while the number of nodes remains high, the overall density of connections is decreasing.

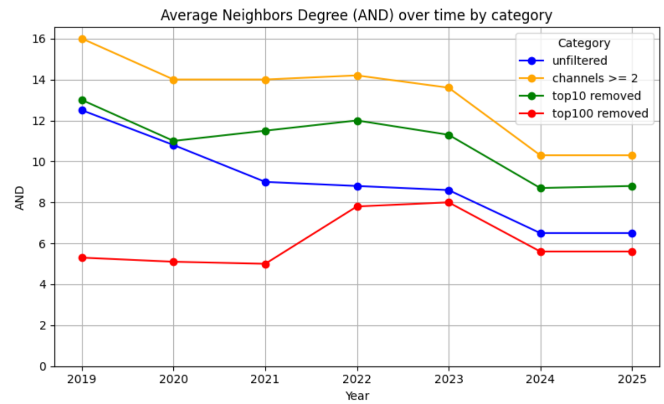

Fig. 3: Average Node Degree over time.

As shown in Figure 4, the ACC also declines, especially after 2021, indicating fewer local clusters or a shift towards a more hub-based structure. Centralization values stay within a narrow range, without a clear trend in either direction.

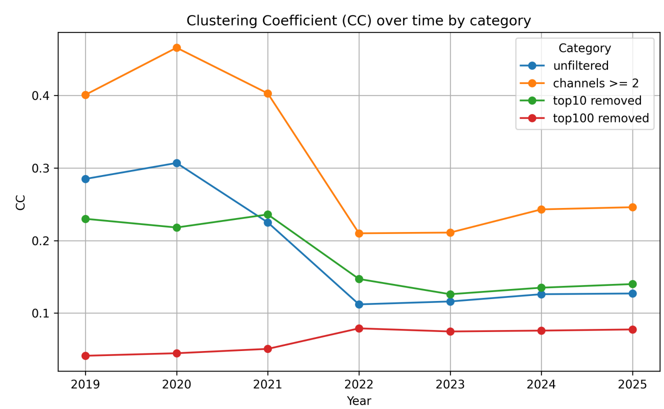

Fig. 4: Average Clustering Coefficient over time.

One point that stands out is the rise in the number of connected components, over 500 by 2025 as highlighted in Figure 5. That might reflect the presence of low-activity or poorly connected nodes entering the network. Taken together, the data hints at a network that is growing at the edges while becoming more reliant on a smaller core for overall connectivity.

By applying the *min-2* filter, the average degree remains higher since we remove all nodes with a single connection, which tend to lower the average. In contrast, the ACC decreases significantly, with values roughly halved compared to the unfiltered configuration. Given that the APL decreases only marginally, this suggests a weakening of the small-world property. The structure appears more cohesive and less

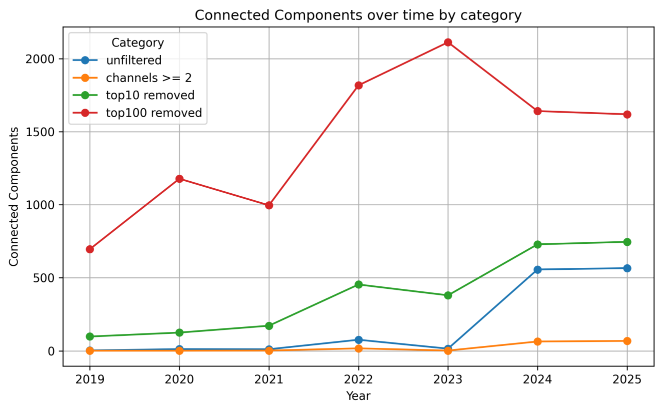

Fig. 5: Number of connected components over time.

fragmented compared to the *unfiltered* setting, though the number of connected components still rises in the final years.

The *top10* filter shows that removing the 10 highest-degree nodes (less than 1% of the network) has measurable effects: the AND drops by 20%, clustering halves, and network centrality falls by over 50% (Fig. 6), confirming the heavy skew toward a few hubs. The network also fragments, with connected components increasing tenfold and path lengths slightly rising. These changes suggest that a small number of nodes are critical for connectivity; their removal causes disproportionate degradation, leaving many participants unable to make payments. From a security perspective, this highlights the risk of targeted attacks on hubs, which could severely impair LN connectivity and efficiency [15]. While certain users could still communicate using an alternative path, others might be unable to make payments.

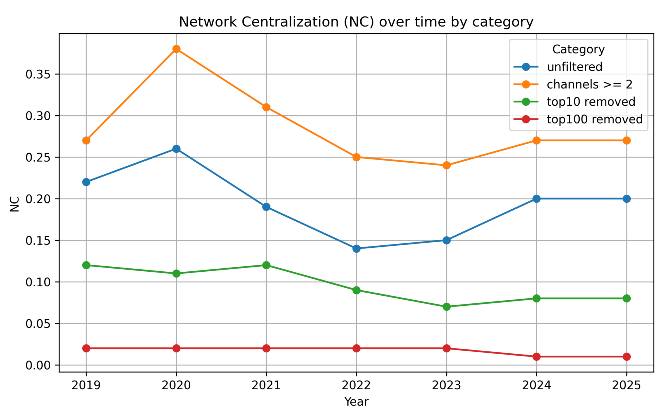

Fig. 6: Network centrality over time.

Applying the *top100* filter amplifies these effects: average degree drops sharply, clustering falls below 0.08, and the network fragments into thousands of components (over 2000 in 2023), making communication largely impossible. Despite the large remaining node count, the structure becomes highly fragmented. As expected, centralization values are low (mostly at 0.01 or 0.02), simply because the nodes that contributed most to centralization have been removed. This does not imply a decentralized structure, but rather lack of coordination and cohesion, as evidenced by the high number of connected components.

IV. Conclusions

In this paper we analyzed the structure of the Lightning Network using graph theory, focusing on how its topological features affect decentralization. By combining data collected with a live node with historical snapshots, we were able to investigate how the LN has evolved since its deployment. Our study confirms that the network exhibits both small-world and scale-free properties, characterized by a few highly connected hubs and many nodes with only a small number of links. This structure supports efficient routing but also raises concerns about resilience in the event of hub failures. In fact, the removal of the highest-connected nodes significantly increases the number of connected components, impairing communication for a large portion of the network. Even when communication remains possible, a higher number of hops is required to complete it, highlighting a marked degradation in overall efficiency.

References

[1] S. K. Panda, A. Sathya, and S. Das, "Bitcoin: Beginning of the cryptocurrency era," in *Recent advances in blockchain technology: Real-world applications*. Springer, 2023, pp. 25–58.

[2] C. Decker and R. Wattenhofer, "Information propagation in the bitcoin network," in *IEEE P2P 2013 Proceedings*. IEEE, 2013, pp. 1–10.

[3] K. Croman, C. Decker, I. Eyal, A. E. Gencer, A. Juels, A. Kosba, A. Miller, P. Saxena, E. Shi, E. Gün Sirer *et al.*, "On scaling decentralized blockchains: (a position paper)," in *International conference on financial cryptography and data security*. Springer, 2016, pp. 106–125.

[4] A. Jain, "Studying the analysis of cryptocurrency and its market dominance," *Available at SSRN 4751851*, 2024.

[5] M. Magnani and M. Marzolla, *Path-Based and Whole-Network Measures*. New York, NY: Springer New York, 2017, pp. 1–16.

[6] V. Modrak, S. M. Radu, and J. Modrak, "Metrics in organizational centralization and decentralization," *Polish Journal of Management Studies*, vol. 10, 2014.

[7] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016.

[8] I. A. Seres, L. Gulyás, D. A. Nagy, and P. Burcsi, "Topological analysis of bitcoin's lightning network," in *Mathematical Research for Blockchain Economy: 1st International Conference MARBLE 2019, Santorini, Greece*. Springer, 2020, pp. 1–12.

[9] L. Labs, "lnd - lightning network daemon," https://github.com/lightningnetwork/lnd, 2024.

[10] B. C. Developers, "Bitcoin core," https://github.com/bitcoin/bitcoin, 2024.

[11] Voltage. (2023) What are the differences between public and private channels? [Online]. Available: https://www.voltage.cloud/blog/what-are-the-differences-between-public-and-private-channels

[12] P. Zabka, K.-T. Förster, C. Decker, and S. Schmid, "A centrality analysis of the lightning network," *Telecommunications Policy*, 2024.

[13] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, June 1998.

[14] M. D. Humphries and K. Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence," *PloS one*, vol. 3, no. 4, p. e0002051, 2008.

[15] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the internet under intentional attack," *Physical review letters*, 2001.