Topological Analysis of Bitcoin's Lightning Network

István András Seres[1], László Gulyás[1], Dániel A. Nagy[1], and Péter Burcsi[1]

[1]Department of Computer Algebra, Eötvös Loránd University

January 14, 2019

Abstract

Bitcoin's Lightning Network (LN) is a scalability solution for Bitcoin, which allows transactions to be issued with negligible fees and settle transactions instantly at scale. In order to use LN, funds need to be locked in payment channels on the Bitcoin blockchain (Layer-1), that one may use them in LN (Layer-2). LN is comprised of many payment channels forming a payment channel network. LN's promise is that a few payment channels already enable anyone to efficiently, securely and privately route payments across the whole network. In this paper we quantify the structural properties of LN and argue that LN's current topological properties could have been ameliorated in order to enable LN to reach its true potential.

Keywords: Bitcoin, Lightning Network, Network Topology, Payment Channel Network

1 Introduction

Recently the Bitcoin [5] network celebrated its 10th anniversary. During these years Bitcoin gained a huge popularity due to its publicly verifiable, decentralized, permissionless and censorship-resistant nature. This tremendous popularity and increasing interest in Bitcoin pushed its network's throughput to its limits. Without further advancements, the Bitcoin network can only settle 7 transactions per second (tps), while mainstream centralized payment providers such as Visa and Mastercard can process approximately 40,000 tps in peak hours. Moreover one might need to pay large transaction fees on the Bitcoin network, while also need to wait 6 new blocks to be published in order to be certain enough that the transaction is included in the blockchain.

To alleviate these scalability issues the Lightning Network (LN) is designed in 2016 [7], and launched in 2018, January. The main insight of LN is that transactions can be issued also off-blockchain in a trust-minimized manner achieving instant transaction confirmation times with negligible fees, whilst retaining the security of the underlying blockchain.

Bidirectional payment channels can be formed on-chain using a construction called Hashed Timelock Contracts (HTLC). Later several payments can take place in a payment channel. The main advantage of payment channels is that one can send and receive thousands of payments with essentially only 2 on-chain transations: the opening and closing channel transactions.

Using these payment channels as building blocks one might establish a

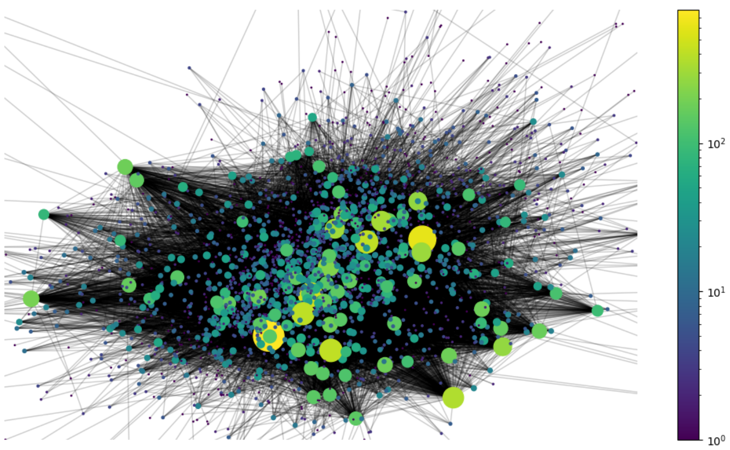


Figure 1: LN's topology. Nodes with higher degree are highlighted with lighter colors and larger circles.

payment channel network, where it is not necessary to have direct payment channels between nodes to transact with each other, but they could simply route their payments through other nodes' payment channels. Such a network can be built, because LN achieves payments to be made without any counterparty risk, however efficient and privacy-preserving payment routing remains a challenging algorithmic task [8].

Our contributions. We empirically measure and describe LN's topology and show how robust it is against both random failures and targeted attacks. These findings suggest that how LN's topology could be ameliorated in order to achieve its true potential.

2 Lightning Network's Topology

LN can be described as a weighted graph $G = (V, E)$, where V is the set of LN nodes and E is the set of bidirectional payment channels between these nodes. We took a snapshot[1] of LN's topology on the 10th birthday of Bitcoin, 2019 January 3rd. In the following we are going to analyze this dataset.

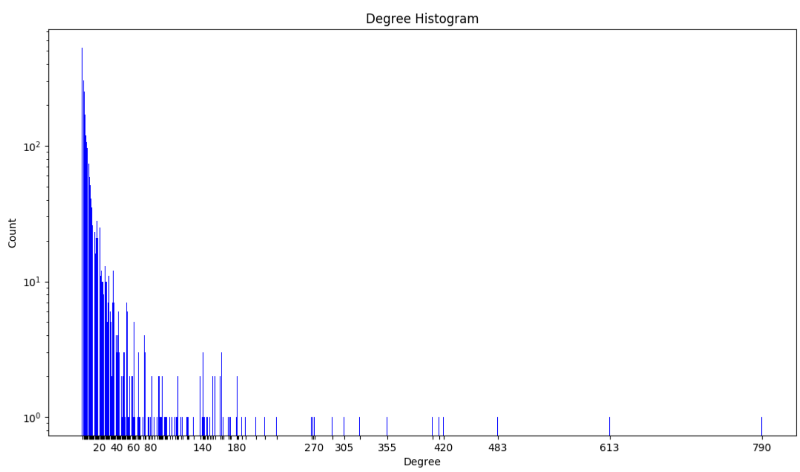
Number of nodes	2344
Number of payment channels	16617
Average degree	7.0891
Connected components	2
Density	0.00605
Total BTC held in LN	543.61855₿
s-metric	0.6878
Maximal independent set	1564
Bridges	530
Diameter	6
Radius	3
Mean shortest path	2.80623
Transitivity	0.1046
Average clustering coefficient	0.304
Degree assortativity	-0.2690

Figure 2: LN at a glance: basic properties of the LN graph.

LN gradually increased adoption and attraction throughout 2018, which resulted in 3 independent client implementations (c-lightning, eclair and lnd) and 2344 nodes joining LN as of 2019, January 3rd. The density of a graph is defined as $D = \frac{2|E|}{|V||V-1|}$ which is the ratio of present and potential edges. As it is shown in Figure 2. LN is quite a parse graph. This is further justified by the fact that LN has 530 bridges, edges which deletion increases the number of connected components. Although LN is consisted of 2 components, the second component has only 3 nodes. The low transitivity, fraction of present and possible triangles in the graph, highlights the sparseness of LN as well.

Negative degree assortativity of the graph indicates that on average low degree nodes tend to connect to high degree nodes rather than low degree nodes [6]. Such a dissortative property hints a hub and spoke network structure, which is also reflected in the degree distribution, see Figure 3.

Figure 3: LN's degree distribution



Average shortest path length is 2.80623, which signals that payments can easily be routed with a few hops throughout the whole network. Although this is far from being a straightforward task, since one also needs to take into consideration the capacity of individual payment channels along a candidate path.

When a new node joins LN, it needs to select which other nodes it is trying to connect to. In the *lnd* LN implementation one of the key goal of a node is to optimize its centrality by connecting to central nodes. This phenomena sets up a preferential attachment pattern. Other LN implementations rely on their users to create channels manually, which also most probably end up users connecting to high-degree nodes. Betweenness centrality of a node v is given by the expression $g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$,

[1]https://graph.lndexplorer.com

where σ_{st} is the total number of shortest paths between node s and t, whilst $\sigma_{st}(v)$ is the number of those paths, that pass through v. Closeness centrality of a node v is defined as $CC(u) = \frac{N}{\sum_{u \neq v} d(u,v)}$, where N is the number of nodes in the graph and $d(u,v)$ is the distance between node u and v. Closeness centrality measures how close a node is to all other nodes.

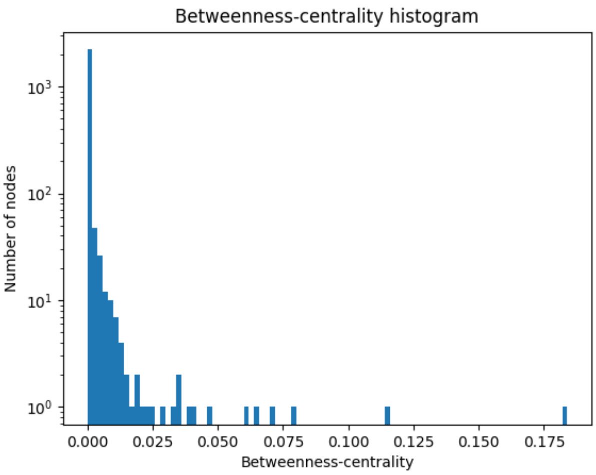


Figure 4: LN's betweenness centrality

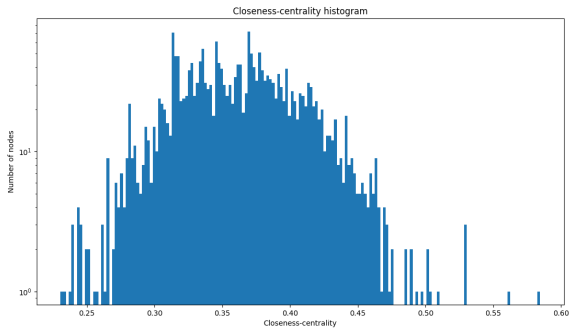


Figure 5: LN's closeness centrality

Small-world architectures, like LN, exhibit high clustering with short path lenghts. The appropriate graph theoretic tool to asses clustering is the clustering coefficient [9]. Local clustering coefficient measures how well a node's neighbors are connected to each other, namely how close they are to being a clique. If a node u has $deg(u)$ neighbors, then between these $deg(u)$ neighbors could be at maximum $\frac{1}{2}deg(u)(deg(u)-1)$ edges. If $N(u)$ denotes the set of u's neighbors, then the local clustering coefficient is defined as $C(u) = \frac{2|(v,w):v,w \in N(u) \wedge (v,w) \in E|}{deg(u)(deg(u)-1)}$.

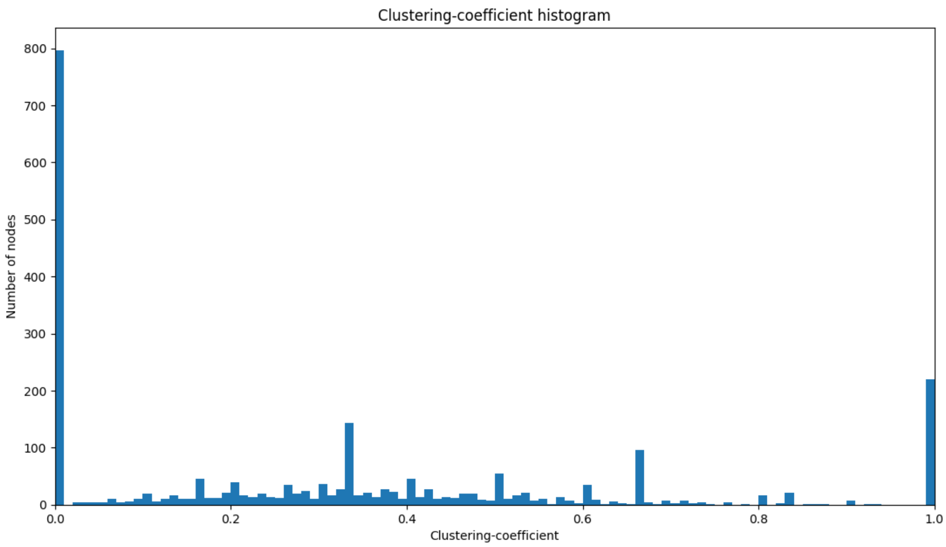


Figure 6: Local clustering coefficient of LN

LN's local clustering coefficient distribution suggestively captures that LN is essentially comprised of a small central clique and a loosely connected periphery.

2.1 Scale-freeness of LN

LN might exhibit scale-free properties as the s-metric suggests. S-metric was first introduced by Lun Li et al. in [3] and defined as $s(G) = \sum_{(u,v)\in E} deg(u)deg(v)$. The closer to 1 s-metric of G is, the more scale-free the network. Diameter and radius of LN suggests that LN is a small world. Somewhat scale-freeness is also exhibited in the degree distribution of LN. Majority of nodes have very few payment channels, although there are a few hubs who have significantly more connections as it can be seen in Figure 3.

The scale-freeness of LN is further justified also by applying the method introduced in [2]. The maximum-likelihood fitting asserted that the best fit for the empirical degree distribution is a power law function with exponent $\gamma = -2.1387$. The goodness-of-fit of the scale-free model was ascertained by the Kolmogorov-Smirnov statistic. We found that the p-value of the scale-free hypothesis is $p = 0.8172$, which is accurate within 0.01. Therefore the scale-free hypothesis is plausible for the empirical degree distribution of LN.

3 Robustness of LN

It is a major question in network science how robust a given network is. LN, just like Bitcoin, is a permissionless network, where nodes can join and leave arbitrarily at any point in time. Nodes can also create new payment channels or close them any time. Furthermore as new payments are made, capacities of payment channels are changing steadily. Despite the dynamic nature of LN, its topology's characteristics remain constant after all. In this section we investigate how resilient is LN, whether it can effectively withhold random node failures or deliberate attacks.

Measuring robustness means that one gradually removes nodes and/or edges from the graph, until the giant component is broken into several isolated components. The fraction of nodes need to be removed from a network to break it into multiple isolated connected components is called the percolation threshold and denoted as f_c. In real networks percolation threshold can be well estimated by the point where the giant component first drops below 1% of its original size [1].

3.1 Random Failures

Random failures are a realistic attack vector for LN. If nodes happen to be off-line due to bad connections or other reasons, they can not participate in routing payments anymore. Such a failure can be modeled as if a node and its edges are removed from the graph.

For scale-free networks with degree distribution $P_k = k^{-\gamma}$, where $2 < \gamma < 3$ the percolation threshold can be calculated by applying the Molloy-Reed criteria, ie. $f_c = 1 - \frac{1}{\frac{\gamma-2}{3-\gamma}k_{min}^{\gamma-2}k_{max}^{3-\gamma}-1}$, where k_{min} and k_{max} denote the lowest and highest degree nodes respectively. This formula yields $f_c = 0.9797$ for LN in case of random failures. This value is indeed close to the percolation threshold measured by network simulation as shown in Figure 10, that is, LN provides an evidence of topological stability under random failures.

Network	f_c
Internet	0.92
WWW	0.88
US Power Grid	0.61
Mobil Phone Call	0.78
Science collaboration	0.92
E. Coli Metabolism	0.96
Yeast Protein Interactions	0.88
LN	0.96

Figure 7: Random failures in networks. Values of critical thresholds for other real networks are taken from [1].

3.2 Targeted attacks

Targeted attacks on LN nodes are also a major concern as the short history of LN has already shown it. On 2018 March 21st [2], 20% of nodes were down due to a Distributed Denial of Service (DDoS) attack against LN nodes. Denial of Service (DoS) attacks are also quite probable by flooding HTLCs. These attack vectors are extremely harmful, especially if they are coordinated well. One might expect that not only state-sponsored attackers will have the resources to attack a small network like LN. In the first attack scenario we removed 30 highest-degree nodes one by one starting with the most well-connected one and gradually withdraw the subsequent high-degree nodes. We recorded the number of connected components. As it is shown in Figure 8. even just removing the highest-degree node[3] fragments the LN graph into 37 connected components! Altogether the removal of the 30 largest hubs incurs LN to collapse into 424 components, although most of these are isolated vertices. This symptom can be explained by the experienced dissortativity, namely hubs tend to be at the periphery.

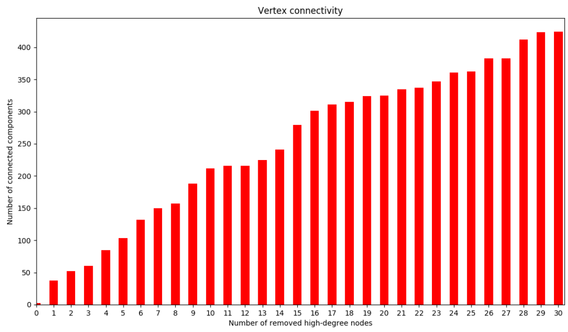


Figure 8: LN's vertex connectivity, when all the 30 largest hubs are removed one by one

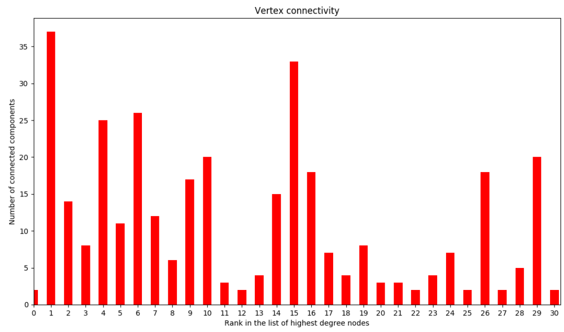


Figure 9: LN's vertex connectivity if only one high-degree node is removed from the graph.

We reasserted the targeted attack scenario, but for the second time we only removed one of the 30 largest hubs and recorded the number of connected components. As it can be seen in Figure 9 most of the hubs, 25, would leave behind several disconnected components.

Such network fragmentations are unwanted in case of LN, because they would make payment routing substantially more challenging (one needs to split the payment over several routes) or even impossible (there would be no routes at all).

Furthermore we estimated the percolation threshold by simulating two attacking strategies. In the first scenario we removed high degree nodes one by one (high degree removal attack, HDR) and in the second we removed nodes with the highest betweenness centrality (high betweenness removal attack, HBR). Note that in both cases after each node removal we recalculated which node has the highest degree or betweenness centrality in order to have a more powerful attack. We found out that $f_c = 0.1627$ for removing high degree nodes, while for removing high betweenness centrality nodes $f_c = 0.1409$, therefore choosing to remove high betweenness centrality nodes is a better strategy as it can also be seen in Figure 12.

Figure 10: Real networks under targeted attacks. Values of critical thresholds for other real networks are taken from [1] and [4].

Network	f_c
Internet	0.16
WWW	0.12
Facebook	0.28
Euroroad	0.59
US Power Grid	0.20
Mobil Phone Call	0.20
Science collaboration	0.27
E. Coli Metabolism	0.49
Yeast Protein Interactions	0.16
LN	0.14

Node outage not only affects robustness and connectivity properties, but also affects average shortest path lengths. Although the outage of random nodes does not significantly increase the average shortest path lengths in LN, targeted attacks against hubs increase distances between remaining nodes. The spillage of high-degree nodes not only decrease the amount of available liquidity but also rapidly

[2] https://www.trustnodes.com/2018/03/21/lightning-network-ddos-sends-20-nodes
[3] http://rompert.com/

r

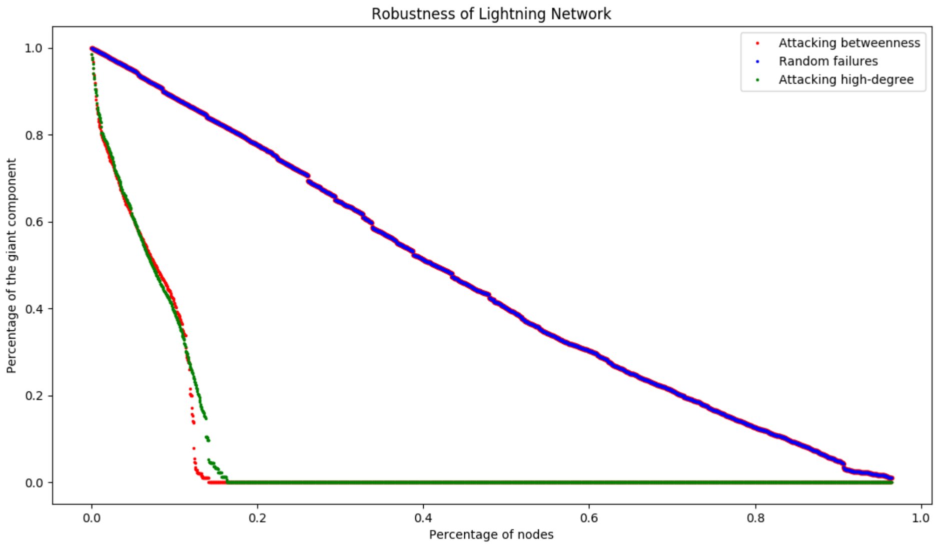
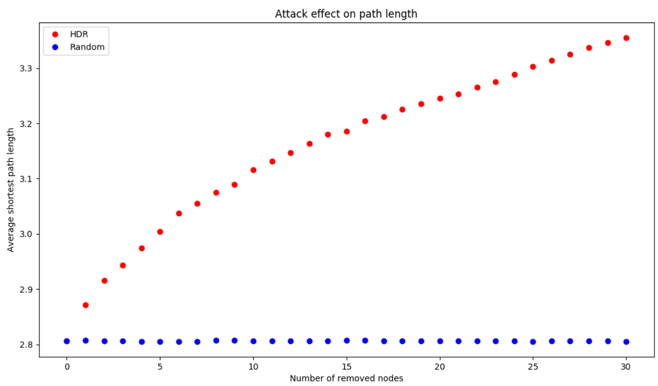


Figure 12: Percolation thresholds for various attack scenarios: $f_c^{HDR} = 0.1627$, $f_c^{HBR} = 0.1409$, $f_c^{RND} = 0.9645$

increase the necessary hops along payment routes as Figure 11 suggests. This could cause increased ratio of failed payments due to larger payment routes.

3.3 Improving LN's resilience against random failures and attacks

Figure 11: High degree removal (HDR) attack effects average shortest path lengths



Designing networks which are robust to random failures and targeted attacks appear to be a conflicting desire [1]. For instance a star graph, the simplest hub and spoke network, is resilient to random failures. The removal of any set of spokes does not hurt the connectedness of the main component. However it can not withstand a targeted attack against its central node, since it would leave behind isolated spokes.

Nonetheless, we could still enhance the network's attack tolerance by connecting its peripheral nodes [1] and mandating newcomers to connect to not only hubs as current implementations do but also to at least a few random nodes. This would largely increase robustness of LN even against targeted attacks.

4 Conclusion

In summary, a better understanding of the network topology is essential to improve the robustness of complex systems, like LN. Network resilience depends on topology. LN is well approximated by the scale-free model and also its attack tolerance properties are similar to that of scale-free networks; namely, LN is robust against random failures, however it is quite vulnerable against targeted attacks. Shiny figures of LN's topology, like Figure 1, convey only false sense of security and robustness, therefore to provide robust Layer 2 solutions for blockchains, the community needs to aim at building resilient network topologies.

5 Acknowledgements

We are grateful for the insightful comments and discussions to Chris Buckland, Olaoluwa Osuntokun and Alex Bosworth.

References

[1] Albert-László Barabási et al. *Network science.* Cambridge university press, 2016.

[2] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[3] Lun Li, David Alderson, John C Doyle, and Walter Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics*, 2(4):431–523, 2005.

[4] Yuan Lin, Wei Chen, and Zhongzhi Zhang. Assessing percolation threshold based on high-order non-backtracking matrices. In *Proceedings of the 26th International Conference on World Wide Web*, pages 223–232. International World Wide Web Conferences Steering Committee, 2017.

[5] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[6] Mark EJ Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.

[7] Joseph Poon and Thaddeus Dryja. The bitcoin lightning network: Scalable off-chain instant payments. *See https://lightning. network/lightning-network-paper. pdf*, 2016.

[8] Stefanie Roos, Pedro Moreno-Sanchez, Aniket Kate, and Ian Goldberg. Settling payments fast and private: Efficient decentralized routing for path-based transactions. *arXiv preprint arXiv:1709.05748*, 2017.

[9] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440, 1998.