Effectivenes simulation of a rebalancing algorithm for the Lightning Network under partial participation

Proposal

May 14, 2020

Organization	FHNW, School of Business, Basel
Study program	Business Information Technology
Author	Tobias Koller
Supervisor	Prof. Dr. Kaspar Riesen
Project sponsor	Prof. Dr. Thomas Hanne
Expert	René Pickhardt

1 Introduction

1.1 Lightning technology

The Lightning Network is a network that utilizes Bitcoin as its underlying system. It can, therefore, be described as a "second layer" protocol building upon the Bitcoin "base layer". Bitcoin is a decentralized peer-to-peer money system with no central entities. The system was designed with security and robustness being the main objectives, sacrificing other properties such as transaction throughput (speed).

The Bitcoin system consists of nodes each maintaining a ledger of historic transactions. All new transactions must be distributed to all nodes and validated by them. Transactions are therefore public information and must be stored by all nodes. To allow many people to run a node, therefore promoting a decentralized network, the hardware requirements must be as low as possible. This is why there is a limitation of new transactions that can be recorded in the network causing this low transaction throughput.

The Lightning technology offers a solution to this issue of scaling by adding a second network on top. In this network participants open payment channels with each other. Transactions within these channels are only visible to the two partners but stay invisible to the rest of the network. While opening and closing a channel each requires one transaction in the base layer (Bitcoin) unlimited transactions with almost no throughput restriction can be facilitated within a channel during its lifetime.

It is important to note the difference between a Lightning node and a Bitcoin node. While they can run on the same system, they operate in two different networks. A Bitcoin node works well individually but a Lightning node needs to have access to a Bitcoin node.

For a node to pay another node that it has no direct channel open with, he can simply route the transaction via other nodes and their channels. Since the network graph is public, the path can be chosen by the initiator of a transaction. This is called source routing.

A payment channel is always opened between two nodes. One of the participants acts as the initiator and provides funds for the channel in the form of bitcoin. This leads to the total capacity being allocated to his or her balance within the channel. As soon as he starts to make payments towards the other node, his or her balance decreases, and the partner's balance increases (total capacity remains constant). Transactions can only be executed if the amount is smaller or equal than the channel's capacity and if the sending node has enough local balance. Channels and their capacities are announced to the network but the distribution of balances remains private to the channel partners.

1.2 Path finding problem

Nodes trying to find a path in the Lightning Network work with limited information. While they know what channels are available and what their capacities are, they do not know about the balances and therefore whether the nodes can forward their payment or not. Hence, it is likely that a payment attempt fails because a node had insufficient balance. The paying node needs to find another route and retry the payment until it succeeds. If the payment fails repeatedly it can cause delays that are bad for the user experience.

1.3 Starting point

René Pickhardt's and Mariusz Nowostawski's publication "Imbalance measure and proactive channel rebalancing algorithm for the Lightning Network" [1] serves as a base to formulate the

question for this thesis. In their work, they present a solution for the pathfinding problem in a privacy-aware payment channel network. The proposed solution includes a rebalancing protocol which the nodes of the network should follow to achieve a higher balancedness (for itself but also the entire network). It consists of instructions to proactively rebalance their channels within their friend of a friend's network, redistributing the relative funds owned in a channel but leaving total funds owned unchanged.

Rebalancing is an activity where one node engages in a circular payment that pays itself. This is only possible when the node has at least two channels with different peers. The payment gets routed **out** through one channel and is **received back** over another. On the way, it can use one or more hops to find back to the sender node. This procedure enables a node to change the balances of the individual channels while the total node balance stays the same. In practice, there would be a fee collected by the intermediate nodes whose channels are used. In the proposed rebalancing protocol nodes would forego the fee and only participate in the rebalancing attempt if their balancedness improves as well.

1.4   Problem statement

These payment channel networks are decentralized by nature and no protocol change can be forced upon the node operators. Therefore, the question arises on how effective this protocol change will be assuming only partial participation of nodes. What are the effects of different levels of participation on the imbalance measure [1] of the network during repeated rebalancing cycles? What is the effect of different levels of participation on the network's ability to route payments between random nodes?

2   Objectives

This section outlines the different objectives that are part of the bachelor thesis. There are mandatory objectives that must be achieved and optional objectives that can be targeted if enough time is available. Although some part of the simulation was already done by René Pickhardt and Mariusz Nowostawski, all the simulation code will be written from scratch.

2.1   Rebuild the Lightning Network

To be able to run any simulation a replication of the Lightning Network must be built. This includes the extraction of node and channel information from the actual network.

One channel in the Lightning Network should be modeled as two directed edges between the same vertices, the network model resulting in a directed graph. This is because edge properties such as "balances" and "fees charged for routing" are different in each direction.

One has to be aware that not all channels in the network are public. The protocol foresees the option to open private channels that are not announced to the network. This is intended for Lightning participants that do not want to partake in any routing activity but only want to send and receive. They are always start- or endpoints of a transaction and often represented by mobile wallets. Since they are not publicly known they are not part of the reconstructed network.

---

[1]Defined as the inequality of the distribution of a nodes channel balance coefficients

2.1.1 Optional: Develop heuristic for intial channel balances

As the distribution of a channel's capacity is private between the two nodes there is no way to model the balances exactly. In the previous experiment, the total capacity was allocated randomly to one of the two nodes. Since the channel was once funded by one of the nodes this represents an acceptable approximation. However, it ignores the fact that many transactions might have been routed since the channel was first opened.

A probing experiment on testnet [2] showed that most channels are highly unbalanced, meaning the capacity is merely on one side of the channel. An optional objective is the development of a heuristic to model the balance distribution in a way that takes into account that channels were used since they were opened.

2.2 Reproduce simulation results from previous study

Before I can answer further questions regarding the partial participation, the findings from [1] must be reproduced with the constructed network. To do this the measurements for balancedness and improvement in routing capabilities should be transferred. Later, other measures can be introduced to compare the different levels of participation.

2.3 Simulate different levels of participation

With objective 1 & 2 completed, we can now simulate partial participation in the new rebalancing protocol. The participation should be simulated in 10% intervals or smaller. At first, the participants should be determined randomly. A heuristic to chose participating nodes can be optionally introduced later.

2.3.1 Optional: Define different balance measurement

In the underlying research [1] the measure for balance is represented by the similarity of all the channels of a node. Meaning, if all channels of a node have the same relative distribution of funds they are all equal, leading to a Gini coefficient of 0. This represents a perfect balancedness. The proposed algorithm targets a small Gini coefficient.

Alternatively, another measure for balancedness could be defined and measured in parallel.

2.3.2 Optional: Define different success measurements

The overall goal of proactive rebalancing is to increase the network's ability to route payments. The previous work [1] measures this with the success rate of random payments between all pairs of nodes along the cheapest path (based on fees) and the median possible payment size along the same paths.

While those measures should be recorded in the different scenarios as well to provide a basis of comparison, other measures could be defined. One would be the median number of attempts to successfully route a payment, starting with the cheapest path.

2.3.3 Optional: Non-random selection of participants

In the beginning, the nodes participating in the protocol change will be selected randomly. However, this is not likely to resemble the real world as different types of node owners share characteristics that make them more or less likely to adopt the change. Nodes can be categorized by size, either total balance or number of channels, or by their centrality in the graph (closeness). These results can give insights about what participants are crucial to adopt the protocol change.

2.4 Analyze and visualize results

This section will define which parameters will be simulated and what measures are being recorded. Further, it shows how this data can be analyzed and what diagrams could help to visualize the results. All plotted graphs are based on random data.

2.4.1 Levels of participation

The share of nodes participating in the proposed rebalancing is the most important parameter of the experiment and it represents the independent variable. In steps of 10% all levels (from 10% to 100%) should be simulated to get a clear picture.

Likely, the different levels of participation lead to different **average** Gini coefficients at the end of the rebalancing operation. This correlation can be shown by a line chart. When different node selection heuristics are used (see section 2.3.3) the different heuristics can each be represented by an individual line.

To show the distribution of the Gini coefficients in more detail a box plot with the same axes can be used.
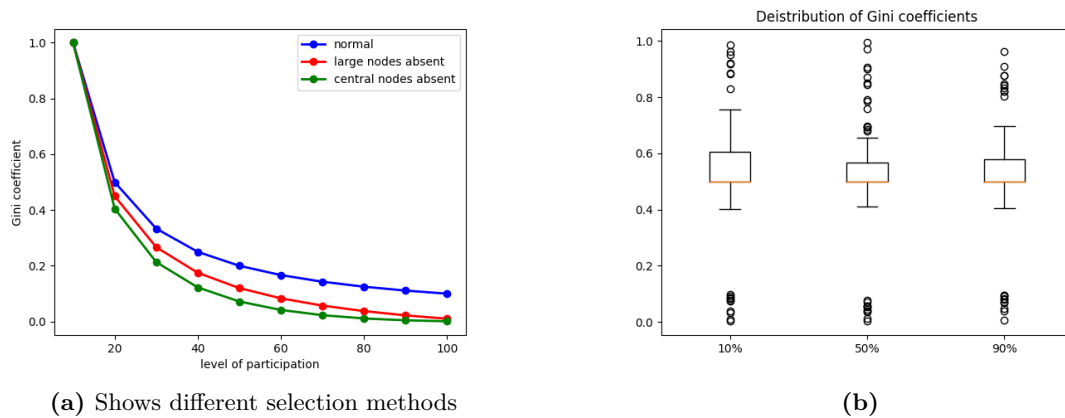

(a) Shows different selection methods

(b)

Figure 2.1: Ways to visualize resulting Gini coefficients.

The x-axis can also be replaced by the number of rebalancing operations while the lines represent the different %'s of participation. The number of data series might be reduced to avoid noise. This chart could help understand how fast a balanced network can be achieved and in what period of rebalancing the most progress is made.
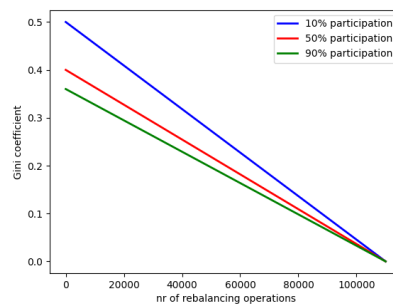

Figure 2.2: Development during rebalancing operations.

2.4.2 Ability to route

The figures to measure the ability to route include "success rate", "median possible payment" and possibly "the median number of retries". Those should be displayed in relation to the independent variable, the participation level. Again, multiple lines can be used to represent the different node selection heuristics.
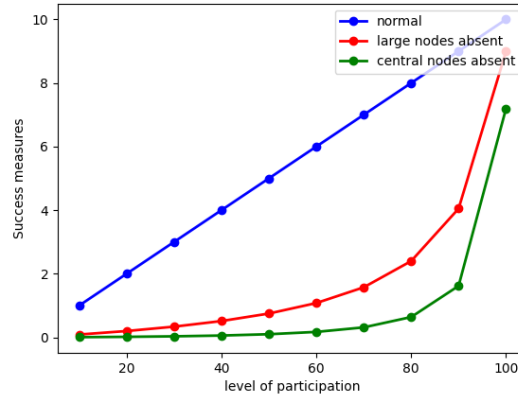

Figure 2.3: Display different success measures

To find out what percentage of cheapest paths between all nodes can route a certain amount of satoshis a cumulative distribution chart could be used. The x-axis displays the number of satoshis to route and the y-axis displays the ratio of all cheapest paths that can accommodate this. Different lines represent different percentages of participation.
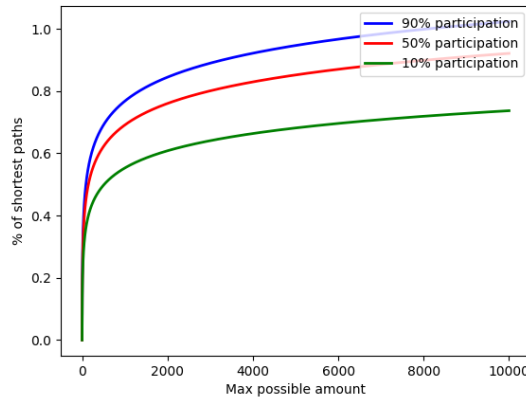

Figure 2.4: Max payable satoshis on cheapest paths

3 Method

To model the network, public information from the Lightning Network is used. From a Lightning node, all the channel and node information can be extracted.

For all further manipulations and calculations, the programming language Python will be used. This includes writing code that facilitates:

- The selection of nodes, participating in the protocol change.

- Implement the proposed algorithm [1, p. 3].
- Performing rebalancing in the network.
- Storing different network states for different scenarios.
- Calculate different performance measures.
- Aggregate data.
- Plot graphs to visualize the results.

4 Assumptions

All models of the Lightning Network and simulations within the realm of this thesis are based on the **public** Lightning Network as this is retrievable by a Lightning node. Private nodes and channels are entirely excluded.

The simulation only takes into account how the network changes when the nodes do rebalancing. Economic transactions that participants are performing in the real network are not simulated. They would constantly reallocate balances and undo some of the work the rebalancing algorithm is performing.

In the simulation, the feasible rebalancing paths can be easily calculated as we have perfect information about all the channel balances in the network. In reality, rebalancing attempts would fail often since balances are not sufficient or nodes do not participate since they do not improve their balancedness.

5 Project schedule

The following table gives an overview of the thesis's schedule. At the end of each week indicated the mentioned objective is planned to be achieved. The final submission deadline is Friday, August 7 (midday). However, to account for unforeseeable events the completion is scheduled two weeks earlier.

Week	Objective / Task	Description
18	Submit proposal [draft]	First version of the proposal will be submitted to René Pickhardt and Kaspar Riesen.
19	Proposal approved [final]	Final version will be submitted to the project sponsor for approval.
20	Rebuild the Lightning Network (2.1)	Retrieve node and channel information and decide for an appropriate Python package to deal with networks.
22	Reproduce simulation results from previous study (2.2)	Confirm similar results fount in previous study [1] with own code and dataset.
25	Simulate different levels of participation (2.3)	Write software to run the simulation. Record the described measures.
27	Analyze and visualize results (2.4)	Aggregate the data in a way that makes them easier to present and read. Write software to plot the data in a meaningful way. Derive conclusions.
29	Submit thesis for review	Document all the steps, simulation results, and findings [draft].

Week	Objective / Task	Description
30	Finalize thesis	Final version ready for submission.

Table 5.1: Time schedule for objectives

References

[1] R. Pickhardt and M. Nowostawski, *Imbalance measure and proactive channel rebalancing algorithm for the Lightning Network.* 2019.

[2] S. Tikhomirov, R. Pickhardt, A. Biryukov, and M. Nowostawski, *Probing Channel Balances in the Lightning Network.* 2020.