Congestion Attacks in Payment Channel Networks

Ayelet Mizrahi
ayelem02@cs.huji.ac.il
The Hebrew University of Jerusalem

Aviv Zohar
avivz@cs.huji.ac.il
The Hebrew University of Jerusalem

ABSTRACT

Payment channel networks provide a fast and scalable solution to relay funds, acting as a second layer to slower and less scalable blockchain protocols. In this paper, we present an accessible, low-cost attack, in which the attacker paralyzes multiple payment network channels for several days. The attack is based on overloading channels with requests that are kept unresolved until their expiration time. Reaching the maximum allowed unresolved requests (HTLCs) locks the channel for new payments. The attack is in fact inherent to the way off-chain networks are constructed, since limits on the number of unresolved payments are derived from limits on the blockchain. We consider two main versions of the attack: one in which the attacker attempts to block as many high liquidity channels as possible, and one in which it tries to isolate individual nodes from the network. We evaluate the costs of both attacks on Bitcoin’s Lightning Network, and compare how changes in the network have affected the cost of attack. Specifically, we consider how recent changes to default parameters in each of the main Lightning implementations contribute to the attack. As we evaluate the attacks, we also look at statistics on parameters in the Lightning Network which are of independent interest and compare the various implementations of Lightning nodes. Finally, we suggest mitigation techniques that make the attack much harder to carry out.

1 INTRODUCTION

Payment channel networks such as The Lightning Network [26] and Raiden [23] are a second layer off-chain solution to the scalability problems of blockchains. Such payment channel networks that run on top of Bitcoin [21] and Ethereum [3] allow both a higher number of transactions and faster transaction resolution. These properties, that are in stark contrast to the blockchain’s slow transaction throughput and slow confirmation times make payment channel networks one of the leading approaches to increase the adoption of cryptocurrencies and may even allow low-fee micro-payments in these systems.

Payment channel networks require participants to lock funds into channels, which then allows them to perform payments using these locked funds. Payments can be relayed over several hops, traveling over multiple channels to their destination.

In this paper we evaluate an attack that locks funds in channels between honest participants that are potentially far away from the attacker, giving the attacker the ability to disrupt the transfer of payments throughout the network. The costs of running the attack are extremely low. We evaluate these costs in the Lightning Network where we show that using less than half a bitcoin, the attacker can indefinitely lock up channels holding the majority of the funds currently assigned to all channels.

Our attack is based on the inner workings of the main mechanism that makes payment channel networks possible: Hashed Time-Locked Contracts (HTLC). Essentially, as payments are set up to move along some path in the network, all channels along the path reserve some funds for the transfer that is about to take place. The number of simultaneously reserved and unresolved payments per path is limited. Our attack thus simply opens many small payment requests along extremely long paths and keeps them unresolved for as long as possible. In this way, all channels along the path are unable to relay other transfers.

The vulnerability can be attributed to three fundamental properties of off-chain payment networks.

1. Payments are executed in a trustless manner. Payments are executed using conditional payment contracts (in the form of transactions with HTLCs) that are exchanged between parties and are only sent to the blockchain if disputes arise. These contracts grow in size as more conditional payments are pending, and so the total number of pending payments is limited by transactions sizes that can be placed on the blockchain. Bitcoin’s Lightning Network is limited to at most 483 concurrent HTLCs [32], while Raiden is limited (due to gas costs) to at most 160 [22].

2. Expiration times are long. To allow nodes to recover their funds if a malicious partner closes a channel that is part of a pending payment, HTLC expiration times have been set to allow nodes sufficient time to appeal such closures.

In Bitcoin’s Lightning Network things are even more severe: due to lower expressiveness of its scripting language, HTLC expiration times accumulate over the length of the path, reaching up to 2016 blocks – which typically take the Bitcoin network two weeks to produce.

3. The privacy of payments. Payment Channel Networks utilize onion routing that does not allow intermediate nodes on the path to recognize where payments originate and where they are going, allowing the attacker to act with impunity.

The vulnerability. In order to paralyze channels, the attacker first adds a new node to the payment network. It then identifies a route suitable to attack, considering some restrictions on the path (maximum route length, locktime of intermediate nodes, remaining HTLC capacity) and maximizing the attack benefit (to lock channels with a large amount of funds). It opens channels with the source and target of the route, and requests many small payments through this path, exhausting the number of simultaneously open HTLCs (see Figure 1). Since the attacker is both the source and destination of this payment, it can choose to delay the final execution of the payment which would remove all pending HTLCs from the path. The path is then locked for long periods of time (up to several days). Just before expiration, the attacker sends an update_failure message to the previous node, which cancels the payment and reverts the state, avoiding a forced closure of the attacker’s channel.
This allows the attacker to re-run the attack once again and lock the same path for an additional period of time.

Figure 1: The adversary identifies routes suitable to attack

In this paper we evaluate the attack specifically on the Lightning Network, which is the most prominently used payment channel network. We evaluate two main attack scenarios: An attack on the entire network, which attempts to lock as many channels as possible and focuses on channels holding most of the funds in the network. This sort of disruption would severely hinder the connectivity of the entire Lightning Network and hurt its ability to relay payments. The main complexity in carrying out this attack is picking routes in a way that respects limits on the maximal delay incurred along the path, and still targets the channels with the highest connectivity and liquidity.

The second attack we consider is one that targets single nodes and paralyzes all channels that connect them to the network.

As far as we are aware, while exhaustion of the HTLCs of a channel is known to paralyze the channel, the attack that we describe has never been evaluated for its effects on the network, or on individual nodes. In particular, there are no available estimates of the cost to attackers from executing either version of the attacks we propose.\(^1\)

The remainder of this paper is structured as follows: Section 2 gives more background related to the Lightning Network. Section 3 explores the current network’s basic characteristics and statistics, specifically focusing on parameters relevant for the attack. In Sections 4-5 we present the two attacks and evaluate their effectiveness and costs. Section 6 presents our proof of concept experiments - paralyzing channels we construct in test-networks in order to validate the attack. In Section 7 we explore suggestions and mitigation techniques to reduce the vulnerability of the network. Section 8 discusses related work. We conclude in Section 9.

\(^1\)We were able to find public record describing the basic idea of the attack, which was raised as a git issue in the BOLT [7] and mentioned in some correspondence in the Lightning-dev list [30]. We note that no full evaluation of the consequences of the attack were raised. Due to the public nature of these posts, we did not perform a disclosure of the vulnerability to the devs.

2 BACKGROUND ON THE LIGHTNING NETWORK

We introduce some of the basic properties of the Lightning Network. The Lightning Network is made up of a collection of channels, each established between a pair of nodes. Channels are provided with funds (either by both nodes or by one) that are locked into a Bitcoin output that is redeemable only by both participants. Channels can then be used to transfer funds between the two participants. This is done by exchanging signed transaction messages between them re-allocating the funds in the channel. As of the end of 2019, the Lightning Network has more than 10k nodes and 35k channels and holds a total capacity of around 860 BTC.

The Lightning Network then allows payments to occur over longer paths, essentially shifting payments from the originating node to its successor in the path, and from that successor onward. The network utilizes source-routing, and its topology is gossiped between nodes in the network to allow payments to be sent. We use this information on the topology to evaluate the different attacks.

Hashed Time Locked Contracts (HTLCs). The main obstacle that the protocol overcomes in order to allow for longer paths is ensuring that payments are done in a trustless manner – nodes that follow the protocol cannot lose their funds even if others on the path misbehave. The way this is achieved is using conditional payments called “Hashed Time-Locked Contracts” (HTLCs).

HTLCs promise an intermediate node on the channel that it can receive funds if it submits a cryptographic proof (pre-image of a hash) within a given timeframe (specified as a specific chain height). Each transaction that occurs in the Lightning Network is first set up by adding an additional HTLC output to every channel on its path. Once these are set up, the payment is executed by propagating the pre-image from the payment’s recipient back along the path towards the sender. Once the pre-image arrives at some intermediate node, it can essentially guarantee that it can receive the funds (if it posts the transaction with the pre-image to the blockchain). The conditional payment is then removed from the channel and is replaced by a non-conditional reallocation of the funds.

The main problem with the approach above is that if several payments are being set up, the number of HTLCs on a channel grows. This implies that the transaction that will eventually be posted to the blockchain will be large – setting a natural limit on the number of HTLCs that can be simultaneously open on a channel.

HTLC Timeouts. Usually, channels are set up quickly and do not wait long for the pre-image to propagate. An update_failure message may sometime be returned instead of the pre-image if one of the intermediate nodes cannot or will not relay the payment. However, malicious nodes may withhold the pre-image and not propagate it back (or alternatively not complete the channel set up with HTLCs).

In such cases, HTLCs are designed to expire. This is done using a CheckLockTimeVerify (CLTV) instruction, which essentially does not allow the HTLC to be redeemed after a certain block height. In order to ensure that intermediate nodes do not lose funds, outgoing HTLCs must expire before incoming HTLCs do. Each node specifies a parameter cltv_expiry_delta which specifies the difference in timeouts it is willing to tolerate. The timeout of payments is...
We begin our exploration of the current state of the Lightning Network by listing the default values for various parameters in the main implementations of the Lightning protocol. These are of interest since, as we show later below, most nodes use the defaults, and thus these heavily influence the state of the Lightning Network and its vulnerability to our attack.

3 LIGHTNING NETWORK ANALYSIS

3.1 Default Parameter Values

The BOLT (Basis of Lightning Technology) [32] specifications detail the protocol of Lightning Networks. These were originally drafted in late 2016 to allow several implementations to work together. In our work, we focus on the main three implementations: LND [14], C-Lightning [4], and Eclair [6].

Each of the implementations uses slightly different default values for parameters of interest. These are depicted in Table 1, along with ranges or values specified in the BOLT.

| Parameter                  | LND  | C-Lightning | Eclair | BOLT   |
|----------------------------|------|-------------|--------|--------|
| cltv_expiry_delta          | 40   | 14          | 14     | -      |
| min_final_cltv_expiry      | 40   | 10          | 9      | 9      |
| locktime_max               | 2016 | 2016        | 2016   | <5 \cdot 10^6 |
| htlc_minimum_msat          | 1000 | 1000        | 1      | -      |
| max_concurrent_htlcs       | 483  | 30          | 30     | ≤ 483  |
| fee_base_msat              | 1000 | 1000        | 1000   | -      |
| fee_proportional_millionths| 1    | 10          | 100    | -      |

Table 1: Default Parameters

Recent changes to the defaults have in fact made our attack easier to carry out: LND changed their cltv_expiry_delta default from 144 to 40 blocks (on 12 Mar 2019) [27], which allows chaining more nodes in each path without reaching the locktime_max limit. Nodes running an old version may still hold the 144 default that was used prior to that.

Additionally, a locktime_max of 2016 was agreed upon by Lightning developers, in the 2018 Adelaide meeting to set the BOLT 1.1 specs [5]. This is an increase of previous values used in some implementations. Again, this allows for longer routes and longer expiration delays that make the attack more damaging and easier to carry out.

3.2 Network Statistics

We introduce some statistics on the parameters announced by nodes in channels on the Lightning Network. In order to perform the calculations, we took snapshots of the Lightning Network mainnet. The information was obtained using the describegraph command of LND. Our results correspond to a network snapshot taken on Jan 1st, 2020. We include other analysis with snapshots taken at different times for comparison.

In Figure 3 we present the most common values of four of the parameters announced by nodes. It is clear that very few values are used. The remaining values appeared less than 1.1% each (which we grouped together as "other").

Figure 3A presents the distribution of htlc_minimum_msat, the minimum amount in millisatoshi (msat) that the node will be willing to transfer. The results indicate that 99.9% configure values ≤ 1000 msat, which is less than 0.0001 USD. While this allows nodes to relay micropayments of this size, it will also allow us to send extremely small payments and will make the attack cheaper to carry out.

In Figure 3B we obtain the distribution of fee_base_msat, the constant fee (in msat) the node will charge per transfer. The results indicate that 97.8% of the network define fee_base_msat ≤ 1000 msat.

We give the defaults used in mainnet. Testnet behavior differs slightly.

We ignore disabled channels and channels with nodes that do not reveal their policies.
The snapshot from 9 Mar 2019 was taken from an external source [28, 29]. The main change that can be observed is the decreased use in the value 144, and the increase in the use of the value 40. We attribute this to the fact that LND changed their default cltv_expiry_delta from 144 to 40 on Mar 2019 [27]. As LND nodes update to the latest defaults which correspond to the different implementations mentioned previously. In Figure 3D we see that the defaults constitute approximately half of the network’s capacity in multiple channels.

Finally, we examine the distribution of cltv_expiry_delta. This is the minimum difference in HTLC timeouts the forwarding node will accept. We recall from table 1 that 144, 40 and 14 are the most common values. The previous statistics imply that the attack’s cost is extremely low. We note, that when traversing a channel in two different directions, the values of cltv_expiry_delta that are used may be different, as these correspond to the values set by nodes. We add an analysis of the cltv_expiry_delta distribution by nodes (rather than by channel), which we do by looking at the most common value of cltv_expiry_delta used by each node. In this case we see that 54.3% of the nodes use the value 40, 33.5% use 144 and 8.2% use the value 14. The remaining values (such as 30, 9, 4) are each associated with less than 1.5% of the nodes.

The value 30, which is common when looking at the distribution of channels is not often used by nodes. In fact, this value is used primarily by a single entity that controls some 25 nodes: LNBIG [13]. These nodes are extremely central to the network, holding approximately half of the network’s capacity in multiple channels.

How do values change over time? In our attack, the route length we can compose is often limited by the values of cltv_expiry_delta. We are therefore also interested in how these values change over time. Figure 4 shows the changes in cltv_expiry_delta during a 10 month-period. We show only the most common values.

The main change that can be observed is the decreased use in the value 144, and the increase in the use of the value 40. We attribute this to the fact that LND changed their default cltv_expiry_delta from 144 to 40 on Mar 2019 [27]. As LND nodes update to the latest

4The snapshot from 9 Mar 2019 was taken from an external source [28, 29]
version, and as old channels are closed and new ones are opened, the statistics gradually change. We can also deduce from the magnitude of the change, that LND constitutes a considerable percentage of the network. We will examine this further below.

3.3 Tagging Nodes by Implementation

Our attack is based on overloading channels to their maximum HTLC capacity, which is determined by the minimum \( \text{max}_{\text{concurrent_htlcs}} \) of the peers in the channel. The \( \text{max}_{\text{concurrent_htlcs}} \) is a configured parameter set by the node’s owner. Default values differ between implementations as described in Table 2.

| implementation | \( \text{max}_{\text{concurrent_htlcs}} \) |
|----------------|------------------|
| LND            | 483              |
| C-Lightning    | 30               |
| Eclair         | 30               |

Table 2: Default \( \text{max}_{\text{concurrent_htlcs}} \) in the mainnet configuration for different implementations

The information of which client a node runs and what value of \( \text{max}_{\text{concurrent_htlcs}} \) it uses is not accessible publicly. Hence, we use the data nodes do publish via channel update messages in order to infer which implementation they run and deduce the \( \text{max}_{\text{concurrent_htlcs}} \) defaults. Here, we rely strongly on the assumption that most users do not change default values too much. This assumption is supported by the statistics we show in Section 3.2.

We perform classification using defaults from Table 1 in order to infer the implementation of each node. Where nodes deviated from defaults, we use a score that is weighed according to the parameter that was changed in order to infer the most likely original implementation. The weights we have use are:

- cltvexpirydelta: 0.75
- htlc_minimum_msat: 0.2
- fee_proportional_millionths: 0.05

For each node in the network we check the compatibility of these parameters in the policies of its channels with the implementation defaults (using the weights) and decide the label.\(^5\)

The results of the implementation inference at different times are presented in Figure 5. We see that our analysis resulted in tagging approximately 90% of the nodes as LND. When we then reduce the network graph to nodes labeled LND alone, we find that we remain with 89% of the network’s original capacity (in BTC). Hence, LND nodes are both the most common nodes, and also the ones that hold most of the liquidity in the network.

For our purposes, this implies that channels between LND nodes are most likely to have \( \text{max}_{\text{concurrent_htlcs}} \) set to 483. A small minority of channels involves either an Eclair node or a C-Lightning node, which would imply that their \( \text{max}_{\text{concurrent_htlcs}} \) is set to 30 (at least in one direction).

In the next two sections we explain two different modes of attack that we evaluated. In the first attack, we attempt to lock up channels with as much liquidity as possible throughout the network. Our assumption is that by tying up a large fraction of the liquidity of the network we will severely restrict its operations. We evaluate the costs of such attacks and their efficiency. In our second attack, we target a specific node, and attempt to lock all of its channels, effectively isolating it from the payment network. Here the degree of a node will be key to the cost of attack.

4 ATTACKING THE ENTIRE NETWORK

In this section we consider a malicious node that wishes to disrupt the entire network’s operation. Our attacker uses a greedy algorithm in order to pick routes and paralyze as much liquidity as possible. For each route, the attacker will initiate \( \text{max}_{\text{concurrent_htlcs}} \) payments, and withhold the response, turning all channels along the path unavailable for new requests. Just before expiration, the attacker will announce a failure to complete the payment. This step is repeated for multiple disjoint routes making the network less and less connected.

The main challenge faced by the attacker is to use routes composed of channels with similar \( \text{max}_{\text{concurrent_htlcs}} \) so that we

\(^5\)The results were robust to the use of different classification methods that we tried.
Algorithm 1 outputs $R_G$ — a partition of $G$’s channels into disjoint routes that can be paralyzed for at least $\tau_{min}$ blocks. We apply Algorithm 1 separately to subgraphs with similar $max_{concurrent_htlcs}$. We unify the outputs $R = \bigcup_G R_G$ and sort the set by routes capacities in decreasing order as we want to attack routes with higher capacity first. Note that routes produced by the algorithm are circular (from the attacker to itself) and require the attacker to open two channels to begin and end each route.

For many channels in the network, the value set for $cltv_{expiry} \_delta$ is different depending on the direction we traverse the channel (this is because nodes may have set different values for this parameter). Our greedy approach excelled at picking channels that keep the network’s capacity that the attacker succeeds in locking as a function of the resources it invests (the number of channels it is required to open). We find for example, that the attacker can lock 20% of the network’s capacity using only 90 channels, and can lock 90% using 1044 channels.

**4.1 Evaluation**

We run the attack with $\tau_{min} = 432$, i.e. locking channels for at least 3 days. We present the attack results, stopping after 1500 attacked routes. Each route requires the attacker to open exactly two channels.

Using our inferred node implementations from Section 3.3, we partition the network into two sub-graphs:

1. The network graph reduced to LND nodes. Which has $max_{concurrent_htlcs}$ defaults that are 483.
2. The complementary graph, that consists of all channels with at least one Eclair or C-Lightning node. These use a default $max_{concurrent_htlcs}$ of 30.

We visualize the results in Figure 6, presenting the fraction of the network’s capacity that the attacker succeeds in locking as a function of the resources it invests (the number of channels it is required to open). We find for example, that the attacker can lock 20% of the network’s capacity using only 90 channels, and can lock 90% using 1044 channels.

![Figure 6: Fraction of attacked network capacity](image-url)
In Figure 7 we run the attack changing the number of days that channels remain locked for. The results indicate that the number of attacker channels required to lock paths for different periods (from 1 to 6 days) differs only slightly. This can be explained by the relation between the large locktime_max (2016 blocks) value, the small cltv_expiry_deltas and the 20-hop route length constraint. In other words, most of the liquidity of the network can be attacked using routes that consist of small cltv_expiry_deltas, allowing the attacker to high timeouts and withhold the payments for a long period.

Figure 7: Fraction of attacked network capacity for different lock periods

We show more details on the results in Figure 8. The figure shows that the attacker succeeds in attacking long routes (exploiting maximum route length), and that most of the routes are locked for more than the 3 days that were set as the minimal lock time.

Figure 8: Histogram of route lengths (including attacker’s edges) and route lock times

Figure 9 explores how the attack would work on the Lightning Network at different times. We use snapshots taken over several months. The results generally show that the attack gets easier as time passes. This can be explained by the changes made to default parameters – increasing locktime_max to 2016 in all implementations, and decreasing cltv_expiry_delta from 144 to 40 in LND. Both changes make it easier to construct long routes with high timeouts.

Figure 9: Fraction of attacked network capacity in different snapshots

Finally, we present an estimate of the attack costs. There are two types of costs associated with the attack:

1. The cost of opening channels. The attacker pays the fee required to place channel funding transactions on the blockchain. We estimated the cost of opening a channel to be 1 USD, which is typically higher than transaction fees that we observed over the last 3 months [2].

2. The cost of provisioning channels with liquidity. Attackers must lock enough liquidity in their channels to later be able to request enough payments over routes they seek to paralyze. Locked funds are not spent, and will return to the attacker once it completes the attack. The calculation of the payment amounts takes into account the minimal amounts nodes are willing to relay, and accounts for all proportional and absolute fees these payments incur over the entire route (using data from the actual network).

Figure 10 displays our evaluation of the costs. It clearly separates the two types of costs mentioned above (non-refundable blockchain fees and locked liquidity). Our results show that the attacker can paralyze 650 BTC of liquidity in the Lightning Network for 3 days using less than 0.25 BTC. We lock at-least 2000 satoshis in each channel (we chose 2000 since we observed existing channels with this capacity), but it is possible to also lock less [25].

5 ATTACKING HUBS - ATTACK ON A SINGLE NODE

In this section we consider an attack aimed at disconnecting a single node from the network for an extended period of time. Here, the
adversary connects to the victim node and paralyzes its adjacent channels one by one using the following steps:

1. The adversary connects to the victim with a new channel.
2. It then initiates a payment to itself via a route that begins with its connection to the victim, and then traverses a single target channel back and forth multiple times, then the path returns to the attacker. It is important to note, that such paths that traverse channels back and forth are possible (see section 6).
3. The attacker makes multiple payment requests over this path until the target channel reaches max_concurrent_htlcs. Note that in this case, the attacker’s own channel is usually not maxed out, and can be used to attack another target channel.

Figure 11 depicts the attack. We note that the attack is still possible to carry out if the victim does not accept direct connections (but at a somewhat lower efficiency). In this case, we would connect to neighbors of the victim.

Once the target channel is paralyzed, we move to the next one and apply the same method. We will need to open a new channel between the adversary and the target node every time that the former reaches its max_concurrent_htlcs. Yet, at each payment we withhold only two HTLCs on the adversary’s channel while it is possible to reach up to 18 HTLCs in the target channel at the same time. In other words, in order to attack all of the victim’s channels, the adversary needs to open a small number of channels relative to the victim’s degree.

5.1 Evaluation

We evaluate the attack on prominent nodes in the network. Table 3 summarizes our results. The names of nodes were taken from our snapshot data directly. The last entry in the table relates to an attack on LNBIG that isolates all 25 nodes from the rest of the network, without paralyzing links between the nodes themselves. Paths were set so that all links are paralyzed for at least 3 days in each iteration.

| Alias           | % of Network | Node’s Degree | Attacker Channels |
|-----------------|--------------|---------------|-------------------|
| ACINQ           | 6%           | 581           | 86                |
| BlueWallet      | 4.8%         | 303           | 36                |
| LNBIG [Ind-01]  | 3.6%         | 427           | 48                |
| Bitrefill       | 3.3%         | 145           | 20                |
| LIGHTNING       | 3%           | 181           | 25                |
| LNBIG (25 nodes)| 47.3%        | 4994          | 566               |

Table 3: Attack on Selected Nodes

We evaluated the cost of attack on all nodes in the network using a snapshot from Jan 1st 2020, isolating each node for 3 days. Figure 12 presents a histogram of the degree of nodes, and shows the relation between the degree and the number of channels attackers needed to perform the attack on each node. Each node is represented by a point in the graph. The number of channels is not directly determined by the degree, because different nodes have set up different values of cltv_expiry_delta.

We see that most nodes have a very low degree and are extremely easy to isolate. Even nodes with high degree, require far fewer channels than the degree to attack.

Figure 12: Degree Analysis

In an additional evaluation, we estimate the cost of isolating nodes running one of the major implementations, assuming default values are used by it and its neighbors. As before, we find the
number of channels the attacker needs to open in order to isolate a node for 3 days for different degrees. We present the results in Figure 13. Notice, that implementations differ due to their different default values. We recall from Section 3.3 that ~ 90% of nodes run LND. Figure 13 shows that these are the easiest to attack.

Eclair is hardest to attack due to its higher default $\text{cltv expiry}_\text{delta}$ value. LND’s value of $\text{cltv expiry}_\text{delta}$ is higher than C-Lightning (40 vs 14), but it is still easier to attack due to its different locktime_max value (CLTV values were low enough so that they did not form a constraint—the number of hops was the main restriction on path length).

The following experiment that we conducted demonstrates our ability to block payments over a channel between victims Alice and Bob.

**Experiment 2 (Blocking the Victim’s Channel).** We repeat the setup of Experiment 1: $\text{Attacker1} \leftrightarrow \text{Alice} \leftrightarrow \text{Bob} \leftrightarrow \text{Attacker2}$. We create 483 different payments from Attacker1 to Attacker2. Again, Attacker2 does not instantly respond with the secret.

We now try to establish one more additional payment from Alice to Bob. The payment fails. Just before HTLC expiration, Attacker2 responds with update_fail_htlc messages for all payments and now, an additional payment does succeed.

The above experiment also blocks the payment from Alice to Bob if some of the 483 payments are in the reverse direction (from Attacker2 to Attacker1), demonstrating that the max Concurrent HTLCs limit applies to HTLCs in either direction. We additionally tried the experiment above in paths that contained loops (even including back and forth traversals of a single channel). These are all allowed by nodes. We used this in a proof of concept experiment to attack a single hub.

**Experiment 3 (Back and Forth Attack on a Single Hub).** We set up the network $\text{Attacker1} \leftrightarrow \text{Hub} \leftrightarrow \text{Node} \leftrightarrow \text{Attacker2}$. Attacker1 initiates payments to itself in the following route, and does not respond with the secrets:

1. $\text{Attacker1} \rightarrow \text{Hub}$.
2. 9 times back and forth on Hub $\leftrightarrow$ Node.
3. $\text{Node} \rightarrow \text{Attacker1}$.

After 26 such payments, Hub $\leftrightarrow$ Node holds 26 · 18 = 468 unresolved payments. Sending an additional payment will fail (27 · 18 = 486 > 483). Hence, Attacker1 sends 2 additional payments that are meant to fill up the remaining HTLC quota:

- Similarly to the previous paths, only going back and forth 7 times on the channel connecting Hub $\leftrightarrow$ Node
- Finally, a single payment to Attacker2: $\text{Attacker1} \rightarrow \text{Hub} \rightarrow \text{Node} \rightarrow \text{Attacker2}$

We now try to establish a payment from Hub to Node, which fails, confirming that we did paralyze this channel. A payment from Attacker1 to Hub succeeds. In fact, Attacker1 succeeds sending 428 (483 - 7 - 1) more payments to Hub while Hub $\leftrightarrow$ Node is blocked. These “free” 428 payments may be used to attack other channels connected to this Hub.

In the next experiment we tried paths with varying max Concurrent HTLCs values, and verified that the minimal value constrains such paths. We further checked that only the edge with the minimal value is fully locked.

**Experiment 4 (Varying max Concurrent HTLCs).** We set up the network $\text{Attacker1} \leftrightarrow \text{Alice} \leftrightarrow \text{Bob} \leftrightarrow \text{Attacker2}$. Attacker1 does not respond with the HTLC secrets. Attacker1, Alice, Carol and Attacker2 have max Concurrent HTLCs configured as 483, while Bob configured max Concurrent HTLCs to be 30. We create 30 different payments from Attacker1 to Attacker2. An additional payment from Attacker1 to Attacker2 fails. An additional payment including Bob in the route fails. An additional payment from Attacker1 to Alice, or from Attacker2 to Carol succeeds. 453 additional payments from Carol to Attacker2 are accepted and wait for Attacker2 to respond.

\[\text{https://github.com/ayeletmz/Lightning-Network-Congestion-Attacks}\]
In the additional experiments, we also verified that paths are indeed limited to 20 hops and to 2016 block lock-time in total (aggregated over the entire path).

7 SOLUTIONS

In this section we discuss several proposed adjustments to payment channel network protocols that may help mitigate the attack. Specifically, we discuss some ideas that were raised in the Lightning-dev mailing list [7, 30], as well as our own suggestions. We discuss weaknesses and strengths of each such suggestion.

**Enforcing fast HTLC resolution.** This is our most drastic suggestion (and perhaps the most controversial one): While HTLC expiration times allow nodes to remain secure and provide sufficient time to publish transactions to the network, we propose the addition of another time-out mechanism. Specifically, if HTLC secrets are not propagated fast enough from one’s neighbor the channel with this neighbor should be closed.

Each node should announce to its successor in the path its own deadline for resolving the HTLC. The node would then be able to communicate an earlier deadline for HTLC resolution to its next hop. If the timeout arrives, and the HTLC was not fulfilled or canceled, the node will wait for the HTLC to naturally expire, but will close the channel with its neighbor.

To avoid having all channels along the path closed due to a failure to complete the HTLC in time, and specifically to avoid closing channels between compliant nodes, the last node in the path will provide proof of the channel closure to its predecessors (this can be done using a zero-knowledge proof for example).

We stress that this proposed mechanism does not replace the HTLC timeouts that still ensure the safety with regards to the current payment. Our mechanism is a way to disconnect misbehaving peers from the network in order to prevent them from repeating the attack many times at no cost.

We note that it is risky to add behavior that automatically closes channels, and so this proposal warrants further evaluation. We leave this to future work.

**Reducing route length.** We suggest lowering the maximum allowed route length (currently 20 hops). The network graph is a small world network [29] – it is highly connected, and a smaller number of hops should still suffice. We point out that shortest paths between nodes in the network have an average of less than 3 hops and that the network diameter is \(\sim 6\) [29, 31], which are significantly lower than the 20 allowed hops. In Figure 14 we show the fraction of successfully attacked capacity (with respect to the attack described in Section 4), assuming that different max route lengths are allowed. The figure shows that attackers need many more channels to attack if they are forced to use shorter route lengths.

**Setting number of max concurrent payments based on trust level.** Currently, each node configures \(\text{max\_concurrent\_htlcs}\) to bound the maximum transfers it is willing to hold concurrently. Most of the nodes use the default value configured by the implementation they run, and in all cases this value may not exceed the number 483 which is derived from the blockchain’s limitations. We suggest changing the way nodes configure this parameter, adjusting the value according to the what level of trust they have in particular peers. Setting a high \(\text{max\_concurrent\_htlcs}\) for some peer effectively allows it to route many concurrent payments through your node and to do more damage if it is malicious. Therefore, newly created channels with unknown and untrusted nodes should default to a low \(\text{max\_concurrent\_htlcs}\).

**Loop Avoidance.** As our experiments show, it is possible to construct paths that visit the same node several times, including traversals of the same channel back and forth. It is relatively simple for nodes to disallow such paths. Since HTLCs that belong to the same path use the same hash, they can be easily recognized and rejected. This will make some of our attacks harder to carry out, and will not hurt the usability of the payment network.

**Non-refundable fees for HTLC setup.** For the sake of completeness, we mention an idea that was discussed (and dismissed) in [7]. Pre-paying a non-refundable fee for route establishment was suggested as a way to mitigate the attack. The idea was to raise costs for the attacker, forcing it to pay fees even if payments are not eventually routed along the path. There are several reasons to avoid this proposal. The first relates to the effectiveness—it does not prevent the attack. As we have seen in the evaluation of our attack cost (Figure 10), even if the attacker pays the fees, the cost will remain low. Secondly, forcing fee payment in cases of failure will cause other problems: one can intentionally fail to complete HTLC setup causing senders to lose fees. Strategic attackers may collect fees from setup and at the same time fail to relay the payment. Lastly, non-refundable fees will also alter the way honest nodes route – they will be less willing to attempt to route via paths that are likely to fail.

8 RELATED WORK

The structural properties of the Lightning Network and its topology have been studied in [31]. The work goes on to study the robustness
of the network to targeted attacks as well as to random node failures, applying techniques that are often used to assess the resilience of scale free networks such as the internet. A similar study of the network’s robustness appears in [12].

A DDoS attack on the Lightning Network occurred in March of 2018. Many nodes were flooded with traffic and Around 200 Lightning nodes were taken offline [35]. Several studies explore more sophisticated attacks on the Lightning Network. Some focus on privacy issues, and others on isolating nodes or disrupting the network in other ways. In [29], an attack that disrupts the liquidity balance of channels is explored. The attacker initiates payments that move all the liquidity to one side, effectively blocking payments in that direction (payments in the other direction are still possible). Using this technique, they exploit node isolation (all liquidity is pushed away from the victim on all links but the attacker’s links). Our attack differs from the attack in [29], as they require direct connections to the victim node, as well as locked liquidity in high amounts (up to the liquidity the victim has), and also require the payment of fees for large transactions.

A similar attack uses payment griefing but avoids paying the fees [24]. In this variant of the attack, the attacker still sends a payment in one direction that unbalances the channel in order to isolate a node. This time it withholds the HTLC pre-image in order to lock the amount, and never really executes the payment. Unlike our attack, this attack still requires large amounts of locked funds, but does indeed avoid paying most of the fees (channel establishment is still needed).

[34] presents a denial-of-service attack based on route hijacking within the Lightning Network. Exploring how routes are created by the most common Lightning implementations, they find that attacking few nodes can result with a denial-of-service to a large fraction of the network. They show how connecting with few channels to the network offering low fees draws most of the routes which yields a potent attack. Our work does not rely on the routing strategy of nodes to attack the network.

[17] explores a clever attack that allows the “short-cutting” of long payment channels on which the attacker appears twice. This allows the attacker to steal the fees of intermediate nodes. The payment is carried out, but nodes in between the two attacker nodes are tricked into thinking the payment was not executed.

The privacy of payments in the Lightning Network is known to be relatively weak. Discovering the current liquidity balance of a channel can be accomplished using techniques from [11]. The main idea is to try different payment sizes over a channel. Learning the response yields lower and upper bounds on the available liquidity.

Due to the above-mentioned privacy reasons and other practical concerns, channels along the network do not reveal their balance, causing a significant decrease in transaction success rates. [33] explores this tradeoff between privacy and utility in PCNs, considering adding noise to channels as well (which adds privacy, but lowers efficiency).

As we have seen, in the Lightning Network each intermediate node along a payment route adds a delay (cltv_expiry_delta) to the total cltv expiry of the path. This leads to high lock-times which contributes to our attack. Sprites, a new type of payment channels, are introduced in [20]. They reduce the time a path can be locked due to a pending payment. Sprites are implemented in the Raiden network on the Ethereum blockchain [23], where a “Preimage Manager” smart contract resolves disputes in parallel. This sort of technique is not applicable to the Bitcoin blockchain, due to its more limited scripting language.

In [8, 10, 15] several protocols improving upon privacy issues in off-chain payment channels are suggested. While [8, 10] support only single-hop payments, the proposal in [15] fits linked payment channels (but provides weaker privacy guarantees).

Other advances in payment channel networks appear in [1, 18], which discuss delegating dispute handling to third party services (called watchtowers) that continuously monitors the blockchain on behalf of others checking if channels closures that should be “appealed” were sent to the blockchain. The watchtower then appeals on behalf of the users and redeems funds. This allows users to be less available and still remain safe.

[19] presents a technical overview of Bitcoin’s payment channel networks. In [16] they construct multi-hop locks for secure and privacy preserving PCNs. Additional work on off-chain protocols can be found in a wider survey [9].

9 CONCLUSIONS AND FUTURE WORK

In this paper we discussed a fundamental vulnerability that arises in payment channel networks as part of the construction of trust-less multi-hop payments. We presented two types of attacks: the first aims to lock as many high liquidity channels as possible for an extended period, and the second isolates hubs from the rest of the network. We evaluated these attacks over the Lightning Network. We examined the network’s properties and different parameters set by the three main implementations of the Lightning Network. We showed how recent changes in default parameters agreed upon by Lightning Devs have made the attack easier to carry out. Our results show that it is possible to disrupt the Lightning Network by locking most of its liquidity using less than half a bitcoin.

Further work must be conducted in order to mitigate this type of attack. We have suggested several solutions to reduce the success rate of these attacks, but such mitigation is generally harder due to the nature of the attack: it relies on several fundamental properties of payment channel networks, and the blockchain.

Another way in which this work can be extended, is to evaluate the effect of max_concurrent_htlcs on naturally occurring network congestion. The limit on concurrent HTLCS sets a limit on the rate of payments the network can handle. This will require measuring typical response times of nodes to an HTLC secret, and to failure messages.

ACKNOWLEDGMENTS

This research was supported by the Israel Science Foundation (grant 1504/17) and by a grant from the HUJI Cyber Security Research Center in conjunction with the Israel National Cyber Bureau.

REFERENCES

[1] Georgia Avrakioti, Felix Laufenberg, Jakub Sliwinski, Yuyi Wang, and Roger Wattenhofer. 2018. Towards secure and efficient payment channels. arXiv preprint arXiv:1811.12740 (2018).
[2] BitInfoCharts. 2020. Bitcoin Avg. Transaction Fee historical chart. https://bitinfocharts.com/comparison/bitcoin-transactionfees.html#3m
Patrick McCorry, Surya Bakshi, Iddo Bentov, Sarah Meiklejohn, and Andrew Giulio Malavolta, Pedro Moreno-Sanchez, Clara Schneidewind, Aniket Kate, Ethan Heilman, Leen Alshenibr, Foteini Baldimtsi, Alessandra Scafuro, and Elias Rohrer, Julian Malliaris, and Florian Tschorsch. 2019. Discharged Payment Elias Rohrer. 2018-2019. Lightning Network Snapshots. https://gitlab.tu-berlin. Olaoluwa Osuntokun (Roasbeef). 2019. Git Commit: “lnd: https://github.com/gitlab.tu-berlin.de/rohrer/discharged-pc-data/tree/master/snapshots diff-356ddb2e7efca712327c3b2d94d3afd3.

Joseph Poon and Thaddeus Dryja. 2016. The bitcoin lightning network: Scalable lower default CLTV delta from 144 to 40”. https://github.com/lightningnetwork/lightning-rfc

Matthew Green and Ian Miers. 2017. Bolt: Anonymous payment channels for decentralized currencies. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, Dallas, TX, USA, 473–489.

Cezary Dziemian. 2018. Summary of the Second Lightning Development Summit. https://lists.linuxfoundation.org/pipermail/lightning-dev/2018-December/001595.html. [Lightning-dev].

EmelyanenkoK. 2017. lightning-rfc issue #182: Payment channel congestion via spam-attack. https://github.com/lightningnetwork/lightning-rfc/issues/182.

Lewis Gudgeon, Pedro Moreno-Sanchez, Stefanie Roos, Patrick McCorry, and Arthur Gervais. 2019. SoK: Off The Chain Transactions. IACR Cryptology ePrint Archive 2019 (2019), 360.

Ethereum: A next-generation smart contract and decentralized application platform. URL https://github.com/ethereum/wiki/wiki/White-Paper (2014).

C-Lightning. 2020. A Scala implementation of the Lightning Network. https://github.com/Eclair.

Vitalik Buterin et al. 2014. Ethereum: A next-generation smart contract and decentralized application platform. URL https://github.com/ethereum/wiki/wiki/White-Paper (2014).

Patrick McCorry, Siamak F Shahandasti, and Feng Hao. 2016. To-