

SOK: Cryptocurrency Networking Context, State-of-the-Art, Challenges

Maya Dotan¹ Yvonne-Anne Pignolet² Stefan Schmid³ Saar Tochner¹ Aviv Zohar¹

¹ Hebrew University, Israel ² DFINITY, Switzerland

³ Faculty of Computer Science, University of Vienna, Austria

ABSTRACT

Cryptocurrencies such as Bitcoin are realized using distributed systems and hence critically rely on the performance and security of the interconnecting network. The requirements on these networks and their usage, however can differ significantly from traditional communication networks, with implications on all layers of the protocol stack. This paper is motivated by these differences, and in particular by the observation that many fundamental design aspects of these networks are not well-understood today. In order to support the networking community to contribute to this emerging application domain, we present a structured overview of the field, from topology and neighbor discovery to block and transaction propagation. In particular, we provide the context, highlighting differences and commonalities with traditional networks, review the state-of-the-art, and identify open research challenges. Our paper can hence also be seen as a call-to-arms to improve the foundation on top of which cryptocurrencies are built.

1 INTRODUCTION

Cryptocurrencies provide means to exchange digital assets relying on strong cryptography and, in contrast to centralized digital currencies and central banking systems, offer *decentralized* control. Cryptocurrencies are thus typically realized as distributed systems. Accordingly, much existing research in the blockchain field focused on the underlying cryptographic primitives and on improved distributed blockchain protocols, e.g., consensus.

The network required to connect the distributed system, however, has received relatively little attention. However, there is increasing evidence that the network can become the bottleneck and root-cause for some of the most pressing challenges blockchains face today. For example, the propagation of transactions and blocks (or other control messages for the execution of consensus algorithms), require unicast and multicast communication services. Blockchain miners which write transactions to the blockchain, are connected through dedicated miner P2P networks [10], in addition to the public blockchain P2P network. Studies show that the cryptocurrency network layer is critical for scalability [17, 47], security [32] and privacy [31] of a blockchain, and that an efficient network layer enables higher transaction throughput and stronger resilience against malicious actors [32]. Networking issues are also not limited to overlays on the network layer. Besides node discovery and data routing, the provided network functionality for example includes the encoding and

transmission of data, and error correction, as well as measurements of the network performance.

Interestingly, cryptocurrency networks and their usage, can differ significantly from traditional communication networks. For example, cryptocurrency networks may come with different requirements (e.g., related to anonymity), may need to serve different traffic mixes (e.g., more frequent broadcast of transactions and states of the blockchain), may need different routing mechanisms (e.g., source routing), or may be more dynamic (e.g., channels and fees in PCNs). Obviously, the fact that no protocol participant can be trusted (not only users issuing transactions but also miners and thus information providers may act maliciously), often requires a different design, and the incentives of all participants must be considered.

This paper is a call-to-arms to the networking community to identify the unique requirements of cryptocurrency networks and address the open issues.

Our Contributions. This paper aims to provide a fast introduction to cryptocurrency network issues with a focus on open research challenges. To this end, we provide a structured overview of cryptocurrency networks, introducing the different aspects and their context, highlighting differences and commonalities with traditional networks, and reviewing the state-of-the-art. At the end of each section, we identify and discuss research questions.

This paper hence specifically targets junior and senior researchers with different backgrounds (e.g., in networking, algorithms, or game theory) who would like to get an overview of the state-of-the-art and start working in this area. It can also serve experts and decision-makers in the networking industry as well as interested laymen.

Related Work. This paper surveys cryptocurrency network issues with a focus on research challenges. The most closely related papers in this area are surveys and systematizations of knowledge efforts on cryptocurrencies and blockchains in general [12]. Gudgeon et al. [36] provide a comprehensive survey of offchain networks, and Neudecker et al. [55] survey the network layer of permissionless blockchains, with a focus on attacks and the design space (but less on research questions, e.g., revolving around incentives or mining). Katkuri presents a survey of data transfer and storage techniques in prevalent cryptocurrencies and suggests improvements [45]. The focus is on aspects related to the broadcast networks underlying Ethereum, Nano and IOTA. Gervais et al. [32] give a thorough security analysis of proof-of-work blockchain systems; their focus is on the consensus layer (i.e., block generation) whereas the network layer is abstracted. Troncoso et al. [66]

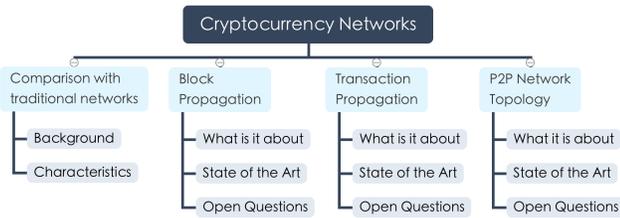


Figure 1: Paper Organization

show a broader perspective covering numerous systems apart from Bitcoin and Tor but also abstract from the network layer. A recent paper by Delgado-Seguara et al. [19] explores the characteristics of the peer-to-peer network established by Bitcoin, but abstracts from the design space of the network layer. Delgado et al. [19] provide an in-depth study of the Bitcoin P2P network.

Organization. The remainder of this paper is organized as follows (see Figure 1). Section 2 provides an overview of some basic aspects of cryptocurrency networks. These include incentives, topology, communication pattern and security. Next, we discuss P2P core protocol aspects, namely block propagation, transaction propagation and P2P network topologies in Sections 3–5. In each of these sections, we give some background information, present the state of the art followed by open research questions that the writers of this paper find interesting. Finally, in Section 6 we conclude this SOK.

2 CRYPTOCURRENCY NETWORKS IN PERSPECTIVE

2.1 Background

Before delving into the details, we provide some background and introduce preliminaries.

Cryptocurrencies permit mutually distrusting parties to engage in financial operations securely. They guarantee that a transaction issued by Alice to send money to Bob reaches its destination at most once and only if Alice’s balance is sufficiently high. Analogously, operations with multiple senders and receivers are typically supported. The guarantees hold despite Byzantine behavior of a fraction of the participants (nodes) maintaining the cryptocurrency service. To this end, cryptocurrencies rely on distributed ledger technology, which serves as a transaction database, containing the global history with all transactions. To build this global history, typically consensus-based blockchain solutions are employed. A blockchain consists of a replicated linked list of immutable blocks, each block comprising batches of transactions. This list is maintained by a large number of nodes to tolerate malicious behavior of a small group of nodes and still reach agreement on the blocks and their content with a consensus protocol. An honest node will only propose and agree on blocks that contain valid transactions: i.e., the transaction is properly signed by the current owners of the funds, it has not been executed already, and the senders’ balance is high enough.¹ Using such a blockchain, virtual currency can

¹To facilitate validity checks, many cryptocurrencies require outgoing transactions to link to a previous incoming transaction. Thus, an

be transferred from senders to receivers in a fully distributed manner, cutting out any middle man or trusted third party. This feature has gained enormous visibility and is envisioned to transform the financial sector and potentially bring disruptive innovation to many other sectors that traditionally rely on trusted third parties.

Bitcoin and other cryptocurrencies run on top of a **P2P network**. Over this network, nodes send and receive blocks and transactions, which are the basic data structures of cryptocurrencies. In permissionless protocols, such as Bitcoin and Ethereum, any machine can join the network and become a node of the P2P network. A node bootstraps its operation with a *discovery protocol* to establish connections with other nodes in the system.

In most cryptocurrencies there are two roles a node can assume: peer or miner. Peers can create and send transactions. Peers verify the correctness of received transactions and blocks and relay them if valid. Miners do anything a peer does, but they also generate blocks. Transactions and blocks are typically propagated in the network using a flooding or gossip protocol. E.g., when a node either creates or receives a transaction or block in Bitcoin, it announce that item and may request it from peers if receiving an announcement of an item it does not have yet. A node does not forward invalid items. A node keeps valid items in memory (the mempool) and answer requests for them. This way, each node in the network will eventually learn about every new item. The underlying wire protocol prescribes the data encoding and how to use which transport protocols. E.g., Bitcoin clients establish a TCP connection and perform a protocol-level three-way handshake informing each other of the height of the blockchain as they know it and the software version they use [42]. To support encryption and authentication, Ethereum defines the TCP-based DEVp2p protocol [25]. After a handshake, all exchanged messages are encrypted and authenticated via key material generated during the handshake.

In a later of this paper, we provide a detailed discussion of block propagation and transaction propagation mechanisms and measurements. Furthermore, incentives to support reliable information forwarding and the different topologies of the virtual currency networks are presented.

Blockchains are typically managed by a peer-to-peer network of nodes which collectively create and validate new blocks. Interestingly, however, cryptocurrency networks and their applications and usage, can differ significantly from traditional communication networks. These differences influence the required performance, security, and incentives, and touch all layers of the network stack. Indeed, while early solutions relied on either a centralized issuer [62, 72] or creating inter-user credit [30], which required users to trust the original issuer; decentralized systems more critically depend on the *network* to connect their constituent parts. One main challenge is the fact that cryptocurrencies are a relatively recent concept and many aspects are not well explored and documented, even

attempt to double spend consists in getting the same transaction into multiple blocks that the recipients consider valid by mistake.

compared to other components of blockchains and cryptocurrencies. With the exception of Bitcoin and Ethereum, many cryptocurrencies lack substantial documentation about their operational details, other than information scattered in the source code repositories. At the same time, the dependability of cryptocurrencies is becoming increasingly important, and vulnerabilities and inefficiencies are a major concern given the corresponding direct financial implications.

2.2 Characteristics

Before diving into the details of the state-of-the-art and research challenges, we give an overview of some of the distinguished characteristics of cryptocurrency networks.

A first important aspect concerns **incentives**. Since there is no central party or consortium paying for the resources (bandwidth, CPU, storage, ..) necessary to maintain the distributed ledger, the participating nodes must be remunerated through the protocol directly. Similarly to traditional peer-to-peer networks, nodes maintaining distributed ledgers require incentives to motivate nodes to propagate information (transactions, blocks, control information) between them. In addition, cryptocurrencies also need an incentive system to motivate nodes to verify blocks and the transactions included in them and discard invalid ones. In Bitcoin, nodes generating blocks are called miners. As a remuneration of their work the creator of block obtains a block reward and a fee for each transaction in the block. Thus, miners have an incentive to keep the knowledge of any transaction that offers a high fee to themselves instead of forwarding them, as any other node that becomes aware of the transaction will compete to include it in a block first and claim the associated fee. Additional incentive questions involve the use of lightweight nodes (know as SPV wallets) that rely on messages from a full-node for their operation. SPV wallets do not hold a complete copy of the blockchain and so must rely on other nodes to track payments sent to them.

Another interesting characteristic is formed by the prevalent **communication pattern**. Many cryptocurrency networks are characterized by frequent broadcast operations, e.g., related to the communication of transactions and states of the blockchain. Furthermore, systems such as Bitcoin do not follow a complex multihop routing scheme but employ a simple flooding-based strategy where all peers in the network replicate the information that has been flowing through the system so far, i.e., keep a complete copy of the blockchain. Hence, there is no need to forward queries to other peers, as all information should to be available at a neighbor.

The **network topology** in cryptocurrencies can be fairly different from traditional networks. For example, while Bitcoin and Ethereum rely on flat random graph topologies, Cardano [15] uses different roles influencing how nodes connect to each other and to users (see Section 5).

In terms of **security** and dependability, cryptocurrencies critically depend on a correct functioning of the consensus layer, and the knowledge of the set of information consensus is to be agreed on (e.g., blocks and transactions). Flooding or

gossip protocols are used for the propagation of the required information to all peers of the network. While unstructured P2P networks have been used for decades (e.g., Gnutella) and were extensively analyzed, the considered adversarial models do not match well the threats to blockchain systems. For example, anonymity providing networks (i.e., onion routing networks [33]) have different requirements regarding information propagation than blockchain based systems. Commonly considered requirements in anonymity providing networks are high performance, low bandwidth cost, resistance to traffic analysis, and resistance to denial of service (DoS). For example, distributed denial of service attacks can be used to gain advantages in mining, voting, and other business and protocol-related activities [3, 43, 68]. To prevent malicious nodes from flooding the network with invalid blocks, nodes use a store-and-forward propagation model, where each node downloads the full block and verifies it prior to propagating it to its peers. This model allows nodes to identify any node which propagates invalid blocks as malicious, and limits the effect of such attacks to the nodes which are directly attacked.

There are also implications on **performance**. While individual nodes may support high transaction rates, the distributed propagation can introduce novel kinds of bottlenecks. Indeed, one of the main issues in blockchain systems is their scalability. Increasing the number of transactions processed by the system naturally requires more resources such as bandwidth and storage, and especially the consensus protocols underlying cryptocurrencies can slow down execution, or even effect security. For example, Nakamoto’s consensus protocol which relies on the longest chain for Bitcoin is known to be susceptible to attacks by weaker attackers as transaction throughputs increase. The main underlying cause of this decline in security is the fact that blocks containing more transactions propagate more slowly through the network, which causes forks to form in the blockchain [17]. As a result, a great deal of work has been devoted to improving block propagation times (see in Section 3). In order for Bitcoin to function as a decentralized system, it must allow nodes to receive blocks at a higher rate than the rate at which blocks are being produced. Indeed, if blocks are produced at a higher rate than a node is capable of receiving them, then said node cannot keep track of balances stored in the blockchain, cannot determine whether or not transactions and blocks are valid, and is in effect excluded from the Bitcoin network. The block propagation time to reach the majority of the network does not depend solely on a receiving node’s bandwidth, but also on the network topology, the bandwidth of all nodes, and the manner in which blocks propagate.

3 BLOCK PROPAGATION

3.1 What is it about?

Blocks in cryptocurrency protocols are used to establish common state. They form the input the consensus protocol strives to reach agreement on. Blocks order transactions, thus the state of the network can be constructed by following the ordering of transactions included in blocks in the consensus

chain. Transactions once included in a block deep in the consensus chain are considered confirmed. Therefore, block propagation is an issue of utmost importance to the consensus process. How fast miners learn about new blocks, and how quickly blocks can be created and validated are crucial for the efficiency of a cryptocurrency.

A block consists of a header and a set of transactions. These transactions can be relayed by the sender together with the block, but this wastes bandwidth if they are already stored in the mempool of the receiver.

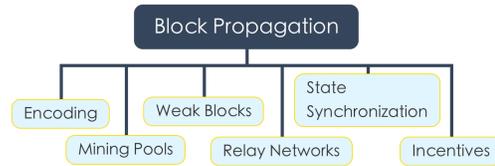
Blocks are typically re-propagated to all connected peers as soon as basic validity of the announcement has been established (e.g. after the proof-of-work check). In Bitcoin, propagation uses the `NewBlock` and `NewBlockHashes` messages. The `NewBlock` message includes the full block and is sent to a small fraction of connected peers (usually the square root of the total number of peers). All other peers are sent a `NewBlockHashes` message containing just the hash of the new block. Those peers may request the full block later if they fail to receive it from anyone within reasonable time.

Blocks can be relayed with a compressed encoding. Efficient propagation of blocks is critical to achieving consensus, reducing storage bloat, overcoming network firewall bottlenecks, and allowing scaling to a large number of transactions per second. Delayed blocks can lead to forks [17]: based on measurements of the rate of information propagation in the network, the propagation delay in the network can be the primary cause for blockchain forks. blockchain forks should be avoided as they are symptomatic for inconsistencies among the replicas in the network. As a mitigation strategy, the authors propose pipelining the block’s delivery, i.e., starting to transmit blocks before they have been fully validated.

One of the reasons for such delays is churn. Imtiaz et al. [40] report that almost all (97%) Bitcoin nodes are connected intermittently only, which results in significant numbers of unsuccessful exchanges, roughly twice the figure for continuously connected nodes. In particular, they demonstrate experimentally that this churn leads to a 135% average increase in block propagation time (i.e., 336.57 ms vs 142.62 ms), and can lead to as high as an 800-fold increase in the worst case measured.

Also permissioned blockchain networks based on Byzantine Fault Tolerance (BFT) consensus algorithms are highly affected by the propagation time. Nguyen et al. [56] demonstrate how a network delay leads to a 30 times larger offset in the consensus layer in Hyperledger.

Often, miners collaborate in mining pools to share the risks and rewards of finding blocks. To this end, a dedicated server is connected to a node that acts as a gateway to a cryptocurrency network. This node gathers newly transmitted transactions and newly built blocks to construct a new block template. The template header is then sent via a mining pool server to the miners which attempt to complete it to a valid block. In the simplest approach for Bitcoin, the miners try different values for the nonce field in the header. If the resulting hash has enough leading zeros for the current difficulty level, i.e., when the block is completed, it is sent back to the server, which then uses the gateway node to publish



the newly formed block to the network and distributes the reward among the miners of the pool corresponding to their contributions.

In 2017, Bitcoin derived at least 95% of its mining power from 10 mining pools; in the Ethereum network, 6 pools are responsible for 80% of the mining power [48].

3.2 State of the Art

A measurement campaign [53] showed that 2016-2018 it takes between 2 and 20 seconds until 90% of the nodes announce the reception of a new block, with a tendency for shorter propagation latency in the more recent past. This is a vast improvement considering that Donet et al. [20] reported that after 84 seconds new blocks have been reached 50% of the nodes and less than 1% of the blocks is known by 90% of the nodes in the same time.

Compressed block encoding. Many proposals to minimize the bandwidth consumption for block propagation exist. One such proposal for Bitcoin addresses the inefficiency of broadcasting blocks with all the transactions included. By the time a new block is created, it is very likely that most peers have these same transactions stored in their mempool. As such, relaying new blocks causes inbound bandwidth spikes for receiving nodes and potentially large outbound bandwidth spikes for nodes that receive blocks before their peers, since they will flood the network with the new, raw block data.

Xtreme Thinblocks (XThin) [67], deployed in Bitcoin Unlimited (BU) clients uses Bloom filter encoding the transaction IDs in nodes’ mempool, thus only missing transactions must be exchanged. In an alternative, the Compact Blocks protocol [1], deployed in the Bitcoin Core, Bitcoin ABC, and Bitcoin Unlimited clients, the block’s transaction IDs are announced shortened to 6-bytes. If the receiver has missing transactions, it requests them separately. Thus, if the receiver is missing many transactions, Compact Blocks have an extra roundtrip time compared to Xthin, which may cost more if enough transactions are missing. Graphene [58] combines the use of a Bloom Filter with Invertible Bloom Lookup Tables (IBLTs) [34] to shrink the size of the sender’s Bloom filter by increasing its false positive rate, and correcting any false positives at the receiver with an IBLT. In comparison to XThin, Graphene uses significantly lower bandwidth both when the receiver is and is not missing transactions. However, Graphene may use an additional roundtrip time to repair missing transactions. Velocity [16] is an approach that exploits the fact that typically several peers already have (parts of the) data in a block. To this end, it applies Fountain codes, which provide a mechanism by which information can be

encoded such that the resulting segments can be probabilistically re-assembled into the original data when the number of the received segments exceeds a threshold.

Stratum Mining Protocol. Stratum [41, 59] is the de-facto standard mining communication protocol used by blockchain-based cryptocurrency systems. It enables miners to reliably and efficiently fetch jobs from mining pool servers.

Stratum was initially a proposal for an open source client-server overlay protocol to support lightweight clients. The Stratum mining protocol extended this proposal to a networking protocol for pooled mining services on the Bitcoin network and many other blockchain protocols. The protocol establishes client-server connections using plain TCP sockets between mining clients and a pool operator or server to distribute new work defined through a blockchain’s proof of work protocol in human readable format.

Recabarren et al. in [60] exploit Stratum’s lack of encryption to develop passive and active attacks on Bitcoin’s mining protocol, with important implications on the privacy, security and even safety of mining equipment owners. Active attacks can hijack shares submitted by miners, and their associated payouts, by modifying TCP packet surreptitiously without causing disconnections and session resets. To mitigate such attacks, the authors proposes Bedrock, a Stratum extension that protects the privacy and security of mining participants with mining cookies. Each miner shares a secret with the pool and includes in its puzzle computations, preventing attackers from hijacking the solutions.

Weak blocks. In order to speed up block propagation even further, one approach is to let miners broadcast blocks they are working on before they have finished the corresponding proof of work. More precisely, so called *weak* or *near* blocks whose proof of work is insufficient for the target difficulty, can be disseminated early. As a consequence, when the block is fully mined the corresponding payload has been received and validated by most nodes already and only the headers needs to be broadcast and processed [2].

Traditionally, the weak blocks are discarded in Bitcoin, wasting their proof of work entirely, contrary to the mission of securing the chain with any and all sufficient proofs of work. Weak blocks by definition have shorter interarrival times and can be used by miners to both receive strong confirmation signals for weak transactions (transactions in weak blocks) as well as anticipate forks sooner. Many updates to Nakamoto Consensus have been proposed that utilize similar ideas, yet no protocol change to utilize weak blocks has made its way into the Bitcoin Core source code. Some, such as BitcoinNG [26], exploit weak blocks to store and propagate transactions. The *key* blocks serve to elect a new leader, granting that miner the right to extend the chain with *weak* blocks. The protocol splits rewards between miners of previous leader elections to incentivize against malicious behavior such as selfish mining or hidden block extension attacks. Another similar proposal termed *Flux* [74] augments the existing Bitcoin protocol with weak blocks such that chains of weak blocks or sub-chains contribute to a chain’s proof of work. Using the heaviest chain rule as its consensus rule, it can

provide faster transaction confirmation times by ensuring that *key* blocks that link to sub-chains contain transactions included in the sub-chain’s *weak* or *sub*-blocks. This can work in practice without the buy-in from all miners as well. One can imagine that if a certain number of miners opts in for broadcasting *weak* blocks, *key* blocks (which would also serve as legacy Bitcoin blocks) could ensure that the dominant chain contains some sub-chains of *weak* blocks.

Relay Networks. In parallel to the public P2P protocol, separate relay networks have been designed to increase network efficiency for miners. The first such system for Bitcoin, called Bitcoin relay network [52], achieved this by disseminating blocks without full block verification and retransmitting known transactions. It consisted of a few nodes (supported by donations) scattered around the globe, all of which peer with each other. Another approach, Falcon [6], relies on cut-through routing for faster block propagation in addition to minimal validation and a hand-optimized topology. More recently, FIBRE [10] has been initiated to provide a similar service by combining cut-through routing with compact blocks and forward error correction over UDP (the normal Bitcoin protocol uses TCP) for registered users. Both Falcon and FIBRE can greatly reduce block propagation times and block orphan rates in the Bitcoin network, as shown in [57]. However, it is important to note that neither was designed, nor is suitable, to scale Bitcoin. Bitcoin cannot rely on these relay networks to achieve higher throughput, since the use of a relay network to scale, places the control over which transactions are included in the blockchain, and which miners may participate, in the hands of its operator. For example, the relay network operator may choose (or be coerced) to propagate blocks only from one group of miners, and reject all others, or to propagate blocks only to one group of miners, and not to others. Worse still, it might reject all blocks which contain transactions involving a specific address, effectively banning its owner from using it.

State Synchronization. For new nodes to be able to contribute to the P2P network quickly, Ethereum provides a state synchronization protocol. The first message sent by two Ethereum peers after the handshake describes their status containing information on the node’s protocol version, network ID (multiple Ethereum networks exist), the hash of the genesis block, the best known block hash and the currently used difficulty. Only connections to nodes operating on the same network ID and genesis block are maintained. Based on their best block hashes the nodes will then synchronize their available information.

When a node joins the Ethereum network, it obtains a local copy of the full blockchain by first requesting block headers, which include block meta information such as parent block hash, miner address. After it has compiled a list of missing block hashes, the node then sends requests to retrieve full block contents and verify the validity of the blockchain.

In Ethereum two validation mechanisms can be distinguished: 1) block header validation and 2) blockchain state validation. Block header validation, ensures that a block’s parent block hash, block number, timestamp, difficulty, gas

limit, and valid proof-of-work hash are correct. In contrast, blockchain state validation consists of sequentially executing all transactions and thus requires significantly more computation and time. In order to reduce the time for new nodes to synchronize and validate the entire blockchain, the *fast sync mode* has been introduced. Instead of running blockchain state validation on all blocks since genesis (as necessary in Bitcoin), header validation is run until a pivot point block close to the most recent head of the blockchain is chosen (using `GET_RECEIPTS` messages for meta information including gas consumption, transaction logs, and status code). At the pivot point, a fast sync node utilizes `GET_NODE_DATA` messages to download a global state database at that block. From the pivot point onward, the node performs full blockchain validation.

Incentives for Block Propagation. For cryptocurrencies to function properly, blocks need to be propagated promptly upon their creation to all other users in the network. This is crucial both to the liveness and to the security of these protocols. To this end, there is a need to examine that miners are incentivised to follow block propagation rules, and not to vary from them. Selfish mining attacks [27, 63] aim to increase the relative fraction of blocks mined by an attacker through timing the release of blocks created by an attacker. The strategies differ based on how long the attacker waits before publishing blocks from a secret chain. These attacks show that there are cases in which miners can profit by not propagating blocks as soon as they are created.

Another case in which miners can profit by deviating from vanilla block propagation is **SPV mining**. In a similar vein, the “SPV (simplified payment validation) mining” concept can decrease the block propagation latency, by avoiding the full verification of blocks and instead relaying them partly unchecked. Originally, the approach has been developed to expedite mining: In order to build on top of the previous block and extend the chain, miners need the hash of the previous block. However, they do not need the full block with all the transaction data in order to start mining. It is in fact sufficient to only have the hash of the previous block header in order to mine a valid block. The incentive for SPV mining is a rush to mine blocks as fast as possible to increase profits. Waiting to download the full block and validate all of the transaction results in idle time which can accumulate to lost profits. Therefore miners may be tempted to find the next block before they have even had time to download and verify the previous block. Like this, miners avoid putting any transactions in the block (apart from the coinbase transaction that rewards the miner), since they cannot know which transactions were in the previous block. Including transactions could result in double spending (which would deem the block invalid). SPV mining is one of the reasons that empty blocks appear on the blockchain [70]. Moreover, SPV mining increases the probability of an invalid block being used to extend the chain and mine the next block linking to an invalid block (since the transactions are not validated by the following block, or even multiple blocks). This in turn results in the network being less reliable for payments as double spends

are more likely. In fact SPV mining has already caused a split in the network in the past: In 2015 there was a change implemented in the Bitcoin protocol (regarding enforcing BIP66 strict DER signatures) that was supposed to go into action after 95% of the network updated their software. The way in which this was implemented is the following: Once 950 of the last 1,000 blocks were version 3 (v3) blocks, all upgraded miners would reject version 2 (v2) blocks. On 4 July 2015, shortly after the threshold was reached, a small miner (part of the non-upgraded 5%) mined an invalid block. Unfortunately, it turned out that roughly half the network hash rate was mining without fully validating blocks, and built new blocks on top of that invalid block, causing a split.

SPV mining is also an issue in the Ethereum, and in 2018 it was discovered that F2Pool (one of the largest mining pools at the time) was engaged in SPV mining (creating a disproportionate amount of empty blocks).²

Another similar strand of attack is Spy mining among competing mining pools. Spy mining occurs when attackers join the pools of others to obtain hints about new blocks appearing on the network. When a spy detects such information via the changed headers sent to it in the Stratum protocol, it can notify its other pool to avoid wasted work. Thanks to SPV, the miners can start mining a new block without seeing the old blocks content.

3.3 Open Questions

Latency and throughput of crypto currencies depend on the efficiency of block propagation. Therefore, mechanisms and incentives to spread newly minted blocks as quickly as possible while minimizing bandwidth waste are crucial. At the same time, the system design must not forego safety and the protocols in place must ensure that blocks contain valid transactions despite the hunt for speed.

OPEN QUESTION 1. *How to accelerate block propagation?* *Several approaches to optimize the exchange of information in blocks between two nodes as well the dissemination in the network have been presented. Yet, both dimensions provide opportunities for empirical studies of the status quo and subsequently to devise new approaches to overcome this barrier to faster transaction rates.*

OPEN QUESTION 2. *How to incentivise mining blocks with many valid transactions?* *SPV mining still poses risks, thus designing a system preventing this behavior, e.g., with economic incentives is still open.*

OPEN QUESTION 3. *How to broadcast blocks within a pool?* *Pools which are very well connected internally and therefore can disseminate newly minted blocks fast, have an advantage in mining over their competitors. Within a pool one could consider a more permissioned model with a more structured overlay topology for speed, balancing the possible velocity gains with the risk of new attacks.*

²<https://medium.com/@ASvanevik/why-all-these-empty-Ethereum-blocks-666acbbf002>

OPEN QUESTION 4. **Efficient network design for mining pools:** *How to get headers to miners quickly, how to disseminate the accomplished work in whole network?*

OPEN QUESTION 5. **Should pools be “allowed”?** *If not can they be prevented? What tools can be used? Can pools be monitored? There are also questions related to incentives (e.g., to force miners to have copies of the blockchain to make pools less attractive).*

4 TRANSACTION PROPAGATION

4.1 What it is about?

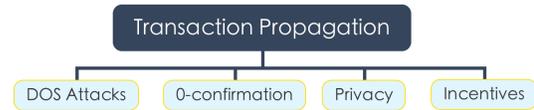
One of the main services provided to users in cryptocurrencies is the propagation of their transactions. Users’ transactions must reach miners in order to be included in blocks, and similarly miners are interested in obtaining transactions of users in order to be able to collect their associated fees. Hence, delays in transaction propagation result in possible delays for transaction confirmation for users, and may cause losses of funds for miners.

Bitcoin’s method for propagating transactions is flooding-based, using the same underlying P2P network used to relay blocks. Thus, if a user sends a transaction to a node, it is potentially sent to all other peers of that node, and then propagated onwards. Transaction relaying has been measured to be slower than block propagation [20]. While 50% of blocks were broadcast to 25% of the nodes in less than 22 seconds, 17 minutes are needed to relay 50% of the transactions to the 25% of the nodes in the sample. Similarly, Neudecker [53] reports that since 2017 transactions propagate to 50% of all nodes within around 5 seconds and to 90% of all nodes within around 15 seconds, both with a trend to increase over the horizon of the measurement campaign.

Kim et al.[46] observed that transactions dominate the network traffic, with clients using different strategies for transaction dissemination. Different protocols have been suggested, such as Geth which relays transactions to all peers, and Parity which forwards them only to a subset of size square root of number of peers in the network.

4.2 State of the Art

Amplified DDoS. One main concern with respect to transaction propagation is that attackers will try to send many transactions to nodes. If each such transaction is later propagated to the entire network unconditionally, then attackers would be able to amplify any large scale DoS attack—sending a single message causes many more messages to be generated throughout the network. Thus, an attempt is made to charge the attacker for sending messages. Unlike with blocks, that are rate-limited by the very fact that they require a proof-of-work in order to be valid, transactions bear no such limits. Their main cost for senders are reflected in the fees paid by the sender. These fees are not guaranteed. In fact, if a transaction is propagated to the entire network and later is not entered to the blockchain, its fees are not collected. Thus, in Bitcoin, miners only convey transactions that enter



their mempool – these in turn are transactions that pay a sufficient fee that is likely to lead to their inclusion in a block.

Similar concerns apply when users wish to increase the fees of their transactions (possibly after finding out the the fee is insufficient to enter a block quickly enough), or to users who wish to double-spend previous un-confirmed transactions. If users are allowed to replace transactions with very modest fee increases, and the new transactions are propagated everywhere, again an attack is possible: attackers will simply send one transaction and then replace it with newer transactions with insignificant fee increases, thereby flooding the network once again. Thus, miners typically set a minimal fee increase to replace a transaction in the mempool and have it re-sent.

Similarly, if a user double spends a transaction and redirects its funds elsewhere, miners will not typically relay the double spend. This serves both as a countermeasure for transaction flooding and as a tool to support 0-confirmation recipient policies (which we discuss below).

Information Eclipsing and 0-Confirmation policies. Some users may not wish to wait until transactions are included in blocks. They then can adopt a somewhat risky policy of accepting 0-confirmation transactions, i.e., they can consider funds received if a transaction is propagated through the network (with fees that they estimate are sufficient to enter a block). Since these transactions are not yet included in a block, they are vulnerable to double spends. Most notably, double spends by miners that need only a single block that contains a conflicting transaction before the double spent transaction is included in a block itself (known as Finney attacks).

Due to the policy of miners mentioned above, double-spend transactions are not entered into the mempool of nodes and are not further propagated to others. This implies that some nodes are essentially “eclipsed” by their neighbors and may never find out about the existence of a double spend, simply because these neighbors first received one transaction (and forwarded that one to the node), but did not relay a conflicting transaction that was later received, leaving the node unaware of its existence. Information eclipsing hinders nodes that would like to accept 0-confirmation policies.

Still, such attacks require mining power which can present some barrier to the attack. The question of how easy it is to attack 0-confirmation transactions without holding mining power naturally arises. Karame et al. [44] analyzed such attacks. They discuss countermeasures such as waiting a while before accepting the payment (and checking to see if conflicting transactions are propagated) and judge that these are not always sufficient to prevent attacks. They propose modifying protocol rules so that nodes forward double spending transactions instead of dropping them to avoid information eclipsing. They however do not analyze the effect of denial of service attacks that may be aggravated as a result.

Privacy. An active attacker may wish to identify the node from which transactions originated. By actively connecting to several nodes, it is possible that a curious observer will observe the transaction origin, or will be able to deduce it.

Biryukov et al. [8] present a deanonymization method for a significant fraction of Bitcoin users that correlates their pseudonyms with public IP addresses. The method explicitly targets peers behind NAT or firewalls, and can differentiate between nodes with the same public IP. They show ways to counteract the fact that nodes may use TOR to hide their IP and essentially utilize an anti DoS countermeasure in Bitcoin to cut off access to TOR. The main deanonymization technique is to identify nodes via the set of nodes it connects to (its entry nodes), transactions are mapped to a set of entry nodes, which uses a fingerprint to associate together transactions with similar sets, which suggests they belong to the same clients. The authors primarily propose to change the set of entry nodes often to avoid such correlations.

Neudecker and Hartenstein evaluate a form of deanonymization in [54]. They compare clustering approaches using transaction data (like signatures for different public keys that appear together in transactions) with clustering based on network data. Finding a correlation between the two likely means that both approaches in fact approximate the desired outcome. They show that for the majority of users no correlation between network information and the clustering performed on blockchain data could be found. A small number of participants do exhibit correlations that might make them susceptible to network based deanonymization attacks.

The Dandelion protocol [11] is a suggestion to provide better network-layer anonymity for transacting users. The main construction in Dandelion is based on having a forwarding phase for transactions before they are widely distributed, trying to mask the origin of transactions. Nodes agree on some full ordering of the network (i.e. some Hamiltonian cycle) which is changed every few minutes in order to avoid the adversary learning it. Transactions are first broadcast along this path for a random (small) number of steps. After this phase, broadcast is done through diffusion as it is done today in Bitcoin. Dandelion++ [29] extends Dandelion to defend against adversaries that are allowed to disobey the protocol. The core improvement is moving from forwarding through a long line graph to a referral 4-regular graph in the initial phase (before diffusion).

Incentives. The incentive of a node to participate in the dissemination of transactions is unclear. As an example, consider the case of a large transaction with high fee. In Bitcoin, the miner’s incentive is to not propagate the transaction to other miners, in order to reduce competition (so it can include the transaction in one of its own blocks and claim the fee).

Although this field was studied thoroughly in the context of P2P networks (examples in [22], [21]), one of the first paper examining this for cryptocurrencies was [5]. In this paper, the authors offer to augment the protocol with a scheme that rewards information propagation, while balancing it with the incentive to decrease competition. They show that their

scheme is sybil-proof (robust against creating clones) and has low overhead (a total reward that is not too high).

An improvements is proposed in [24], where additionally to awarding the propagating nodes, the authors propose “smart routing”. In this mechanism, nodes directly route the transactions to a round leader, which is known in advance (first-leader-then-block (FLTB) consensus protocols). This mechanism increases the bandwidth efficiency by reducing the propagation of redundant transactions.

4.3 Open Questions

If the miners do not have access to a large number of transactions to put into blocks, the throughput and latency of the cryptocurrency will be suboptimal. In addition to the technical limits of spreading transactions widely, incentives play a very important role here and must be designed carefully.

OPEN QUESTION 6. *How to avoid transaction floods?* *An attacker may try to flood the network with meaningless transactions and thus cause the nodes to waste resources on them. How can one quickly identify bad transactions and discard them? How to trade off the verification cost, punishment mechanisms and velocity?*

OPEN QUESTION 7. *How to balance DOS prevention and 0-confirmation requirement?* *On the one hand, nodes strive to avoid DoS attacks by not allowing the propagation of double-spend transactions. On the other hand, users want low-latency cryptocurrency systems and thus favor 0-confirmation policies. Mechanisms to meet these two conflicting goals would allow for a better user experience.*

OPEN QUESTION 8. *How to model and analyse cryptocurrencies?* *The costs and benefits of propagating the different protocol messages deserves a more thorough analysis under realistic utility assumptions. This will uncover further weaknesses of current mechanisms and inform the development of superior approaches.*

OPEN QUESTION 9. *How to implement a suitable reward system?* *Currently prevalent mechanisms incentivise nodes to keep transactions with high fees to themselves, instead of propagating them widely. To alleviate this shortcoming, one must find better methods to align the incentives of nodes and cryptocurrency users. In addition, strategies that allow a superior method to be adopted quickly must be developed.*

5 TOPOLOGY OF THE P2P NETWORK

5.1 What is it about?

The Bitcoin P2P network topology is formed by each peer connecting to 8 nodes (outbound connections) and accepting up to 125 in-coming connections. Outbound destinations are randomly selected among known identities. In other cryptocurrencies, nodes are assigned roles that influence the topology. For example, Cardano distinguishes between mutually exclusive core, relay, and edge node roles [15]. The core nodes

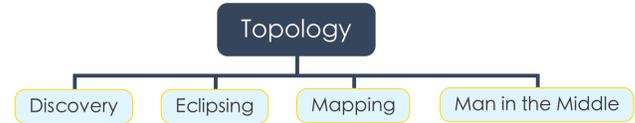
create blocks and run the consensus protocol, in other words, the core nodes maintain the blockchain. Relay nodes protect the core nodes, serving as intermediaries between the public internet where the edge nodes reside and the core node. If relay nodes are attacked, this may lead to a service interruption, but the integrity of the core nodes (and thus the Cardano blockchain) is not compromised. Relay nodes are fully under the control of the federated committee of initial Cardano stakeholders. Edge nodes create the payload of the blocks. They can be run by anyone on their computer to create currency transactions. They cannot directly communicate with core nodes, only with relay nodes and with other edge nodes. Also Ripple distinguishes between different roles for nodes, partitioning them in to superpeers and leafs[61]. A nodes in the leaf role does not route messages and only connects to super peers. In the superpeer role, a peer accepts incoming connections from other leaves and superpeers up to the configured slot limit. It also routes messages.

Before being able to send and receive protocol messages a node has to find other nodes to connect to join the network. This *discovery* process typically relies on static information sources and/or distributed hash table (DHT) approaches, discussed in the first part of this section. Knowledge of the network topology can give parties an advantage in the dispersal of information (blocks, transactions) which can lead to security risks. Because of this, there has been extensive research and development of tools and techniques aimed at exploring and mapping the Bitcoin P2P topology. The same holds true for many other crypto-currencies. We present here some of these works and their main contributions.

5.2 State of the Art

Discovery. To join a cryptocurrency P2P network, a new node must find other nodes to connect with. In Bitcoin, a node first tries to connect to nodes it knows from participating previously. If no connections can be established this way, or if the node connects for the very first time, it queries a list of well known DNS seeds. As a last resource, it will try to connect to hardcoded seed nodes. The DNS seeds are maintained by Bitcoin community members: some of them provide dynamic DNS seed servers which automatically get IP addresses of active nodes by scanning the network; others provide static DNS seeds that are updated manually and are more likely to provide IP addresses for inactive nodes [9]. With the first messages exchanged between new peers, they inform each other of a random subset of locally known addresses with a timestamp of at most 3 hours ago. With this mechanism a list of addresses is maintained at each node. Each node will also accept incoming connections (up to 125 by default). The address of a new node is propagated through the network, so all peers can learn about it eventually.

Ethereum and Cardano use mechanisms inspired by Kademlia [49], a Distributed Hashtable (DHT) approach that has already been widely used for file sharing. Kademlia assigns key-value pairs to sets of peers based on the distance between the key ID and the node IDs and a routing table structure is



maintained that allows to find responsible nodes recursively in a logarithmic number of steps.

While Kademlia and its derivatives have been used for many years, there is no formal proof for its performance and robustness. More theory-oriented approaches [13, 37, 73] on the other hand, have not established themselves as alternatives. [37] shows how to maintain clusters of size $O(\log n)$, each containing more than two thirds of honest nodes with high probability. Even when the system size can vary polynomially with respect to its initial size, the communication cost induced by each node arrival or departure is $O(\text{polylog } n)$. The approach guarantees robustness to a Byzantine adversary controlling a fraction $1/3$ of the nodes (could be $1/2$ with the application of cryptography). The proofs guarantee polylogarithmic maintenance and sampling overhead and rely on assumptions of a synchronous network. Alternatives, e.g., [73], based on the BLS threshold signature scheme, demonstrate how a DHT with quorums of logarithmic size, the time and message complexity is polylogarithmic if up to a third of nodes per quorum is malicious. To this end, any request needs quorum approval before getting answered or continued to avoid SPAM and DOS attacks and prevent wrong responses. To remain robust despite churn, routing tables are maintained according to the Cuckoo Rule. Compared to other approaches, this method requires the creation and verification of many signatures, which is demanding.

Instead of a DHT-based approach, Brahms [13] proposes churn and Byzantine resistant sampling. The paper presents an attack-resilient gossip-based membership protocol and shows how to extract independent uniformly random node samples from the stream of node ids gossiped by the first. It draws its power from an assumption of limited bandwidth availability: Byzantine nodes cannot send messages unlimitedly, if one node sends more often than expected, it is ignored. In Brahms unsynchronized gossip rounds, nodes send addresses they know to some other nodes (to reinforce the knowledge for underrepresented nodes), and request known addresses from other nodes (to spread existing knowledge). If more than the expected number of address are received in a round, Brahms does not update its view in that round to prevent malicious influence. Furthermore, locally known history also influences next views to avoid poisoning. Together with a sampling algorithm this ensures that nodes have an approximately uniform sample of the nodes in the system, even as long as every joining correct node knows some correct other node. Brahms offers a tradeoff between communication and storage and can thus be adapted to different needs. Such a sampling approach can also be run on top of other solutions. **Mapping the network.** In [7], Ben Mariem et al. characterize and analyze Bitcoins' P2P network topology and main properties with a purely network measurements-based approach. They present a BTC crawler in order to discover and track

all active nodes of the BTC P2P network. They also discuss a passive approach to reconstruct the topology of a blockchain P2P network, which can unveil miners. Another approach is presented in [18], uncovering the topology of the Bitcoin network by processing orphaned and conflicting transactions. More precisely, by sending different double-spending transactions one can determine if these nodes are connected. This is possible because of how a node’s mempool maintains orphaned transactions and processes conflicting transactions. Their techniques allow mapping the topology of the Bitcoin network in roughly 8 hours. The technique enables accurate topology reconstruction because of how nodes deal with both new transactions and orphaned transactions. In [50], Miller et al. determine the Bitcoin topology by utilizing the update method of the timestamp field in the `addrMan` of nodes in the network to learn about the topology. With this approach, they show that the deployed Bitcoin topology does not resemble a random graph.

The authors of [35] suggest to exploit two mechanisms of transaction accumulation to map the Bitcoin network: (i) a node accumulates input transactions before propagating them, and then propagates them in the original order; (ii) nodes do not propagate contradicting transactions, so by propagating a contradiction to two nodes, they can tell which node is closer to the target (the one which got it first).

Similarly, other networks have been analyzed, including Ethereum and Monero. [46] unveils that in Ethereum more than 40% (12%) of the nodes use US (Chinese) IP addresses. The most used ASs used by the nodes can be assigned to cloud hosting providers, dominating residential or commercial operators. In [14] the authors developed tools to collect Monero’s P2P network information, including its network size, distribution (geo-wised), and connectivity. They set up 4 Monero nodes across the world with their tools. They found many interesting properties about this network: 87% of the nodes have a degree smaller than 8 (approx. 17% of the overall number of edges), and they mapped the complete outgoing connections of 99% nodes in the network. They succeeded to connect to 85% of the nodes that they discovered. They argue that this is a vulnerability since it enables network level attacks (eclipse, BGP hijacking, DOS).

There has also been some work that aims to fix some of the mapping issues raised. One approach is to secure blockchain network communication using SCION [69]: the design and use of a cryptocurrency network such as Bitcoin and Ethereum. The goal of using SCION is to facilitate client node communication with a more direct, point-to-point infrastructure to prevent attacks and numerous cases of lost cryptocurrency from network hijacks and BGP route poisoning/spoofing.

Eclipsing / Splitting the network. Splitting the Bitcoin network can have severe consequences, as the shorter chain produced will not survive and as a result, many transactions are rolled back (and potentially double-spent), the revenue of miners from these blocks is lost. Partitions affect the ability of participants to operate on transactions. This may cause exchanges to stop receiving and sending the cryptocurrency, and merchants to be unable to get paid. There are several

ways to isolate nodes in the network: for example, by disrupting the routing of traffic between them, or by causing nodes to connect only to attackers (Eclipsing). We list here a few well-known eclipse attacks, and some techniques suggested in the literature to detect and avoid them.

In Bitcoin, nodes choose their peers from a list of stored IP addresses. This list is limited in size and IPs must be evicted if fresh ones are placed inside. IPs are placed in the list pseudo-randomly in a way that is based on the IP itself and the IP of the advertising node. In [39], the authors explore ways of isolating nodes in Bitcoin by affecting the way that nodes choose their peers. The main idea of the attack is to announce many IPs to the node that are either controlled by the attacker or that have no node behind them. The node eventually evicts all IPs of honest nodes from the list, and will only connect to the attacker. The idea is to cause collisions in the placement of similar IPs and of IPs advertised by the same node, thus the attack above needs some minimal number of IP addresses controlled by the attacker to succeed. This number is not high in practice. Also Ethereum is the target of eclipse attack constructions. In [71], Ethereum’s peer-to-peer network is partitioned without monopolizing the connections of the victim, which is possible due to the block propagation design of Ethereum. In this attack, the attacker can potentially keep the victim from receiving a block almost indefinitely. This attack could be used as an infrastructure for a double spend attack. The authors also present an exploit that can force a node to accept a longer chain with lower total difficulty than the main chain (also using the block sync mechanism). In this attack a node that newly connects to the network and receives a chain that is longer than the valid chain but has a lower total difficulty because the adversary advertised a higher total difficulty than honest nodes. The attacker is therefore disconnected from the network. The authors highlight a bug in Ethereum’s difficulty calculation as well. This can be used in an attack that prevents the victim from synchronizing with the valid chain. The paper also outlines countermeasures.

Two attacks on Bitcoin exploiting the networking stack are presented in [4]. First, BGP hijacking that is used to disconnect parts of the network. The network is shown to be poorly distributed, so that relatively few prefixes need to be hijacked in order to partition miners from each other. Once a partition is fixed, natural churn allows nodes to connect across the partition and blocks are once again propagated. A second attack proposed in the paper utilizes the fact that Bitcoin traffic at the time was un-encrypted. Intervention in the content of announcements of new blocks and transactions as well as requests for the data of recently announced blocks was shown to severely delay block propagation. As a result nodes are left uninformed of the latest blocks in the chain for longer periods, which causes miners to waste time mining blocks that will be discarded and will yield no reward, and users to be unaware of funds they may already have received.

A game theoretic approach is used in [64] to manage the list of known peers. Consequently, attackers need to corrupt a large number of nodes to eclipse a node successfully. Similarly

to the Bitcoin protocol, the paper assumes that acquiring IPs from the same prefix is cheaper than acquiring the same number of IPs from multiple prefixes, and utilizes this fact in the peer selection mechanism to increase the attack cost.

In [51] the authors investigate new “Stubborn Mining” attacks which combine eclipse attack with selfish mining [28] attacks. In this work, the authors consider the same model against users who are also eclipsed in the network and show the effect to which eclipsed users help a stubborn mining attacker. Overall, eclipse attacks empower adversarial agents with a larger strategy space to continue running attacks, and when paired with stubborn mining strategies, enable an attacker to improve their relation fraction of block rewards beyond traditional selfish mining strategies.

In [65], Tran et al. present the EREBUS attack, that partitions the Bitcoin network without any routing manipulations, which makes the attack undetectable (even against bug fixes specifically addressing partitioning attacks). Adversaries who may control large transit ISPs, are able to mount the attack. The adversary utilizes a large number of network addresses reliably over an extended period of time. A fix attempt to this attack is suggested here. This work enables Bitcoin core to prefer to connect to peers which are on different source ASNs to try to reduce the probability of any single host/path/hijack is relied on by a peer. The authors focus on the process of building the AS map, including simple filtering suggestions such as treating prefixes only reachable via a common upstream as if they were hosted directly on that upstream (by pulling routing information from diverse sources).

SABRE [3] presents a secure and scalable Bitcoin relay network resilient to routing attacks, designed to run alongside the existing peer-to-peer network and is easily deployable. The authors use properties of BGP to predict where would be a good place to host relay nodes – locations that are inherently protected against routing attacks and on paths that are economically-preferred by the majority of Bitcoin clients. In addition, they provide resiliency through soft/hardware co-design through the use of caching, and offloading most operations to hardware (programmable network devices). This enables SABRE relay nodes to sustain load even when originating by DDoS attackers.

The effectiveness of mining pools can also be hampered by Distributed Denial of Service Attacks (DDoS) in order to disrupt their operations. As a consequence a competing mining pool is slowed down giving an advantage to other pools. This in turn may encourage individual miners to leave unreliable mining pools and join the attacker’s pool as a result. After currency exchanges, Bitcoin mining pools are the most frequent victim of DDOS attacks [68]. Of 49 mining pools, 12 experienced (repeated) DDoS attacks. Based on a game theoretic model where pools can select between investing funds into additional mining equipment or DDOS attacks [43], larger mining pools have a slightly greater incentive to attack than smaller mining pools.

Man in the middle attacks. In [23], the authors study the impact of Man-In-The-Middle Attacks on Ethereum. The paper looks closely at the feasibility of MITM and double

spending attacks on simulated network corresponding to the real Ethereum topology with real network components. They show the impact of such attacks, also gathering public information about the network, and mimicing the structure of its biggest 10 mining pools connected through 5 BGP routers, and performing BGP hijacking and ARP spoofing. The authors find that the attack is almost infeasible in the public context (because of its structure), but in the case of route hijacking (e.g. if Ethereum is deployed over a WAN in a consortium environment, and an adversary that has control on the border gate) could double-spend through either BGP hijacking or ARP spoofing with a success rate up-to 80%.

5.3 Open Questions

OPEN QUESTION 10. **Topology information hiding:** *How can the overlay topology be efficient yet make it hard for attackers to learn it and mount eclipse and hijacking attacks? Answers to this question provide discovery mechanisms that strike a balance between containing truthful information and DOS resistance, e.g., using overlay rotation and sharding mechanisms that minimize the information necessary to participate in a crypto currency network.*

OPEN QUESTION 11. **Giant honest component:** *Most cryptocurrencies employ a flooding-based strategy to broadcast on a topology constructed with a (pseudo) random process. Non-honest nodes may choose to drop messages or forward outdated and wrong information. For a broadcast to succeed, thus enough honest nodes must be connected to each other via at least one path, not containing bad nodes. To this end, graphs must provide large connected components that consist of only honest nodes.*

OPEN QUESTION 12. **Link failure models:** *Traditional failure models consider the number of faulty nodes as the main parameter when analyzing P2P networks. For a more nuanced analysis, link failures must be taken into account as well. Link failures can be modelled as random processes or in a worst-case fashion that is bounded in some way (e.g., a strongly connected union of available links when considering a time interval [38]). More granular models and analyses tailored to the cryptocurrency conditions and constraints are necessary to better understand current cryptocurrency networks and to build the basis for future designs.*

6 CONCLUSION

This paper presented an overview of the cryptocurrency networking aspects, with a focus on open research questions, and towards this goal, covering also background and state-of-the-art. We believe that the networking aspects have not yet received the attention they deserve, and hope that our paper can contribute toward more research in this space.

Table 1 provides a summary of these questions, structured around the methodologies needed to address them, from algorithms, over network protocols, game theory, security, to cryptography.

Open Question	Methodologies
How to accelerate block propagation?	PROT, ALG, GAME, SEC, CRYPTO
How to incentivise mining blocks with many valid transactions?	GAME
How to broadcast blocks within a pool?	ALG, PROT
Efficient network design for mining pools	PROT, ALG
Should pools be allowed?	GAME
How to avoid transaction floods?	SEC, GAME
How to balance DOS prevention and 0-confirmation requirements?	PROT, GAME, ALGP
How to model and analyze crypto currency?	PROT, GAME
How to implement a suitable reward system?	GAME
Topology information hiding	SEC, CRYPT, ALG
Giant honest component	ALG
Link failure models	PROT, ALG

Table 1: Open network questions and methodologies needed to address them: ALGORITHMS, network PROTOCOLS, GAME theory, SECURITY, CRYPTOGRAPHY.

Acknowledgments. Research in part supported by the Vienna Science and Technology Fund (WWTF) project WHATIF, grant ICT19-045, 2020-2024.

REFERENCES

- [1] [n.d.]. Bitcoin Improvement Proposals Compact Block Relay. <https://github.com/bitcoin/bips/blob/master/bip-0152.mediawiki>. Accessed: 2019-09-23.
- [2] Gavin Andresen. 2015. *Weak Block Thoughts*. <https://github.com/ethereum/devp2p/blob/master/caps/eth.md>
- [3] Maria Apostolaki, Gian Marti, Jan Müller, and Laurent Vanbever. 2018. SABRE: Protecting Bitcoin against Routing Attacks. *arXiv preprint arXiv:1808.06254* (2018).
- [4] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. 2017. Hijacking bitcoin: Routing attacks on cryptocurrencies. In *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, 375–392.
- [5] Moshe Babaioff, Shahar Dobzinski, Sigal Oren, and Aviv Zohar. 2012. On bitcoin and red balloons. In *Proceedings of the 13th ACM conference on electronic commerce*. 56–73.
- [6] Soumaya Basu, Ittay Eyal, and Emin Gün Sirer. 2016. *Falcon: Relay Network for Bitcoin Blocks*. <https://www.falcon-net.org/>
- [7] Sami Ben Mariem, Pedro Casas, and Benoît Donnet. 2018. Vivisectioning Blockchain P2P Networks: Unveiling the Bitcoin IP Network. In *ACM CoNEXT Student Workshop*.
- [8] Alex Biryukov, Dmitry Khovratovich, and Ivan Pustogarov. 2014. Deanonymisation of clients in Bitcoin P2P network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 15–29.
- [9] Bitcoin. 2019. *P2P Guide*. <https://bitcoin.org/en/p2p-network-guide>
- [10] BlueMatt. 2016. FIBRE. <http://bitcoinfibre.org/public-network.html>
- [11] Shaileshh Bojja Venkatakrisnan, Giulia Fanti, and Pramod Viswanath. 2017. Dandelion: Redesigning the bitcoin network for anonymity. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 1, 1 (2017), 22.
- [12] Joseph Bonneau, Andrew Miller, Jeremy Clark, Arvind Narayanan, Joshua A Kroll, and Edward W Felten. 2015. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. In *Proc. IEEE Symposium on Security and Privacy (SP)*. IEEE, 104–121.
- [13] Edward Bortnikov, Maxim Gurevich, Idit Keidar, Gabriel Kliot, and Alexander Shraer. 2009. Brahms: Byzantine resilient random membership sampling. *Computer Networks* 53, 13 (2009), 2340–2359.
- [14] Tong Cao, Jiangshan Yu, Jérémie Decouchant, Xiapu Luo, and Paulo Verissimo. 2019. Exploring the Monero Peer-to-Peer Network. *IACR Cryptology ePrint Archive* 2019 (2019), 411.
- [15] Cardano. 2019. *P2P Topology*. <https://cardanodocs.com/cardano/topology/>
- [16] Nakul Chawla, Hans Walter Behrens, Darren Tapp, Dragan Boscovic, and K Selçuk Candan. 2019. Velocity: Scalability improvements in block propagation through rateless erasure coding. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 447–454.
- [17] Christian Decker and Roger Wattenhofer. 2013. Information propagation in the bitcoin network. In *Proc. IEEE P2P Conference*. IEEE, 1–10.
- [18] Sergi Delgado-Segura, Surya Bakshi, Cristina Pérez-Solà, James Litton, Andrew Pachulski, Andrew Miller, and Bobby Bhattacharjee. 2018. TxProbe: Discovering Bitcoin’s Network Topology Using Orphan Transactions. *arXiv preprint arXiv:1812.00942* (2018).
- [19] Sergi Delgado-Segura, Cristina Pérez-Solà, Jordi Herrera-Joancomartí, Guillermo Navarro-Arribas, and Joan Borrell. 2018. Cryptocurrency networks: A new p2p paradigm. *Mobile Information Systems* 2018 (2018).
- [20] Joan Antoni Donet, Cristina Pérez-Solà, and Jordi Herrera-Joancomartí. 2014. The bitcoin P2P network. In *International Conference on Financial Cryptography and Data Security*. Springer, 87–102.
- [21] Fabio A Drucker and Lisa K Fleischer. 2012. Simpler sybil-proof mechanisms for multi-level marketing. In *Proceedings of the 13th ACM conference on Electronic commerce*. 441–458.
- [22] Kirill Dyagilev, Shie Mannor, and Elad Yom-Tov. 2010. Generative models for rapid information propagation. In *Proceedings of the First Workshop on Social Media Analytics*. 35–43.
- [23] Parinya Ekparinya, Vincent Gramoli, and Guillaume Jourjon. 2018. Impact of man-in-the-middle attacks on ethereum. In *2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS)*. IEEE, 11–20.
- [24] Oğuzhan Ersoy, Zhijie Ren, Zekeriya Erkin, and Reginald L Lagendijk. 2018. Transaction propagation on permissionless blockchains: incentive and routing mechanisms. In *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*. IEEE, 20–30.
- [25] Ethereum. 2020. *Ethereum Wire Protocol*. <https://github.com/ethereum/devp2p/blob/master/caps/eth.md>
- [26] Ittay Eyal, Adem Efe Gencer, Emin Gün Sirer, and Robbert Van Renesse. 2016. Bitcoin-ng: A scalable blockchain protocol. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. 45–59.
- [27] Ittay Eyal and Emin Gün Sirer. 2014. Majority is not enough: Bitcoin mining is vulnerable. In *International conference on financial cryptography and data security*.
- [28] Ittay Eyal and Emin Gün Sirer. 2018. Majority is not enough: Bitcoin mining is vulnerable. *Commun. ACM* 61, 7 (2018), 95–102.
- [29] Giulia Fanti, Shaileshh Bojja Venkatakrisnan, Surya Bakshi, Bradley Denby, Shruti Bhargava, Andrew Miller, and Pramod Viswanath. 2018. Dandelion++: Lightweight cryptocurrency networking with formal anonymity guarantees. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 2, 2 (2018), 29.
- [30] Ryan Fugger. 2004. Money as IOUs in social trust networks & a proposal for a decentralized currency network protocol. *Hypertext document*. Available electronically at <http://ripple.sourceforge.net> 106 (2004).
- [31] Arthur Gervais, Srdjan Capkun, Ghassan O Karame, and Damian Gruber. 2014. On the privacy provisions of bloom filters in lightweight bitcoin clients. In *Proceedings of the 30th Annual Computer Security Applications Conference*. 326–335.
- [32] Arthur Gervais, Ghassan O Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, and Srdjan Capkun. 2016. On the security and performance of proof of work blockchains. In *Proc. of the 2016 ACM SIGSAC conference on computer and communications security*.
- [33] David Goldschlag, Michael Reed, and Paul Syverson. 1999. *Onion routing for anonymous and private internet connections*. Technical Report. NAVAL RESEARCH LAB WASHINGTON DC CENTER FOR HIGH ASSURANCE COMPUTING SYSTEMS . . .
- [34] Michael T Goodrich and Michael Mitzenmacher. 2011. Invertible bloom lookup tables. In *Allerton*. IEEE.

- [35] Matthias Grundmann, Till Neudecker, and Hannes Hartenstein. 2018. Exploiting transaction accumulation and double spends for topology inference in bitcoin. In *International Conference on Financial Cryptography and Data Security*. Springer, 113–126.
- [36] 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.
- [37] Rachid Guerraoui, Florian Huc, and Anne-Marie Kermarrec. 2013. Highly dynamic distributed computing with byzantine failures. In *Proceedings of the 2013 ACM symposium on Principles of distributed computing*. 176–183.
- [38] Bernhard Haeupler. 2016. Analyzing network coding (gossip) made easy. *Journal of the ACM (JACM)* 63, 3 (2016), 1–22.
- [39] Ethan Heilman, Alison Kendler, Aviv Zohar, and Sharon Goldberg. 2015. Eclipse attacks on bitcoin’s peer-to-peer network. In *24th {USENIX} Security Symposium*.
- [40] Muhammad Anas Imtiaz, David Starobinski, Ari Trachtenberg, and Nabeel Younis. 2019. Churn in the Bitcoin Network: Characterization and Impact. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 431–439.
- [41] Bitcoin Inc. 2015. *Stratum Mining Protocol*. https://en.bitcoin.it/wiki/Stratum_mining_protocol
- [42] Bitcoin Inc. 2018. *Network*. <https://en.bitcoin.it/wiki/Network>
- [43] Benjamin Johnson, Aron Laszka, Jens Grossklags, Marie Vasek, and Tyler Moore. 2014. Game-theoretic analysis of DDoS attacks against Bitcoin mining pools. In *International Conference on Financial Cryptography and Data Security*. Springer, 72–86.
- [44] Ghassan O. Karame, Elli Androulaki, and Srđjan Capkun. 2012. Double-spending Fast Payments in Bitcoin. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*.
- [45] Sunny Katkuri. 2018. A survey of data transfer and storage techniques in prevalent cryptocurrencies and suggested improvements. *arXiv preprint arXiv:1808.03380* (2018).
- [46] Seoung Kyun Kim, Zane Ma, Siddharth Murali, Joshua Mason, Andrew Miller, and Michael Bailey. 2018. Measuring Ethereum network peers. In *Proceedings of the Internet Measurement Conference 2018*. 91–104.
- [47] Uri Klarman, Soumya Basu, Aleksandar Kuzmanovic, and Emin Gün Sirer. 2018. bloxroute: A scalable trustless blockchain distribution network whitepaper. *IEEE Internet of Things Journal* (2018).
- [48] Loi Luu, Yaron Velner, Jason Teutsch, and Prateek Saxena. 2017. Smartpool: Practical decentralized pooled mining. In *26th {USENIX} Security Symposium*.
- [49] Petar Maymounkov and David Mazieres. 2002. Kademlia: A peer-to-peer information system based on the xor metric. In *International Workshop on Peer-to-Peer Systems*. Springer, 53–65.
- [50] Andrew Miller, James Litton, Andrew Pachulski, Neal Gupta, Dave Levin, Neil Spring, and Bobby Bhattacharjee. 2015. Discovering bitcoin’s public topology and influential nodes. (2015).
- [51] Kartik Nayak, Srijan Kumar, Andrew Miller, and Elaine Shi. 2016. Stubborn mining: Generalizing selfish mining and combining with an eclipse attack. In *2016 IEEE European Symposium on Security and Privacy (EuroSecP)*. IEEE, 305–320.
- [52] Bitcoin Relay Network. 2017. *high-speed block-relay system for miners*. <http://www.bitcoinrelaynetwork.org/>
- [53] Till Neudecker. [n.d.]. Characterization of the Bitcoin Peer-to-Peer Network (2015-2018). ([n.d.]).
- [54] Till Neudecker and Hannes Hartenstein. 2017. Could network information facilitate address clustering in Bitcoin?. In *International conference on financial cryptography and data security*. Springer, 155–169.
- [55] Till Neudecker and Hannes Hartenstein. 2018. Network layer aspects of permissionless blockchains. *IEEE Communications Surveys & Tutorials* 21, 1 (2018), 838–857.
- [56] Thanh Son Lam Nguyen, Guillaume Jourjon, Maria Potop-Butucaru, and Kim Thai. 2019. Impact of network delays on Hyperledger Fabric. *arXiv preprint arXiv:1903.08856* (2019).
- [57] Kai Otsuki, Yusuke Aoki, Ryohei Banno, and Kazuyuki Shudo. 2019. Effects of a Simple Relay Network on the Bitcoin Network. In *Proceedings of the Asian Internet Engineering Conference*. 41–46.
- [58] A Pinar Ozisik, Gavin Andresen, Brian N Levine, Darren Tapp, George Bissias, and Sunny Katkuri. 2019. Graphene: efficient interactive set reconciliation applied to blockchain propagation. In *Proceedings of the ACM Special Interest Group on Data Communication*. 303–317.
- [59] Slush Pool. 2019. *Stratum Mining Protocol*. <https://slushpool.com/help/stratum-protocol/>
- [60] Ruben Recabarren and Bogdan Carbutar. 2017. Hardening stratum, the bitcoin pool mining protocol. *Proceedings on Privacy Enhancing Technologies* 2017, 3 (2017), 57–74.
- [61] Ripple. 2019. *Overlay*. <https://github.com/ripple/rippled/tree/develop/src/ripple/overlay>
- [62] Ronald L Rivest. 1997. Electronic lottery tickets as micropayments. In *International Conference on Financial Cryptography*. Springer, 307–314.
- [63] Ayelet Sapirshstein, Yonatan Sompolinsky, and Aviv Zohar. 2016. Optimal selfish mining strategies in bitcoin. In *International Conference on Financial Cryptography and Data Security*.
- [64] Saar Tochner and Aviv Zohar. 2018. How to pick your friends—a game theoretic approach to p2p overlay construction. *arXiv preprint arXiv:1810.05447* (2018).
- [65] Muoi Tran, Inho Choi, Gi Jun Moon, Anh V Vu, and Min Suk Kang. 2020. A Stealthier Partitioning Attack against Bitcoin Peer-to-Peer Network. (2020).
- [66] Carmela Troncoso, Marios Isaakidis, George Danezis, and Harry Halpin. 2017. Systematizing decentralization and privacy: Lessons from 15 years of research and deployments. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 404–426.
- [67] Peter Tschipper. 2016. BUIP010: Xtreme Thinblocks. In *Bitcoin Forum (1 January 2016)*. <https://bitco.in/forum/threads/buip010-passed-xtreme-thinblocks>, Vol. 774.
- [68] Marie Vasek, Micah Thornton, and Tyler Moore. 2014. Empirical analysis of denial-of-service attacks in the Bitcoin ecosystem. In *International conference on financial cryptography and data security*. Springer, 57–71.
- [69] Aleksandar Vorkapic. 2018. Secure Blockchain Network Communication using SCION.
- [70] Canhui Wang, Xiaowen Chu, and Qin Yang. 2019. Measurement and analysis of the bitcoin networks: A view from mining pools. *arXiv preprint arXiv:1902.07549* (2019).
- [71] Karl Wüst and Arthur Gervais. 2016. *Ethereum eclipse attacks*. Technical Report. ETH Zurich.
- [72] Beverly Yang and Hector Garcia-Molina. 2003. PPay: micropayments for peer-to-peer systems. In *Proceedings of the 10th ACM conference on Computer and communications security*. 300–310.
- [73] Maxwell Young, Aniket Kate, Ian Goldberg, and Martin Karsten. 2013. Towards practical communication in Byzantine-resistant DHTs. *IEEE/ACM Transactions on Networking (ToN)* 21, 1 (2013), 190–203.
- [74] Alexei Zamyatin, Nicholas Stifter, Philipp Schindler, Edgar R Weippl, and William J Knottenbelt. 2018. Flux: Revisiting Near Blocks for Proof-of-Work Blockchains. *IACR Cryptology ePrint Archive* 2018 (2018), 415.