Information Propagation on Permissionless Blockchains

Oğuzhan Ersoy*, Zhijie Ren, Zekeriya Erkin, and Reginald L. Lagendijk

Cyber Security Group, Department of Intelligence Systems, Delft University of Technology

Abstract

Blockchain technology, as a decentralized and non-hierarchical platform, has the potential to replace centralized systems. Yet, there are several challenges inherent in the blockchain structure. One of the deficiencies of the existing blockchains is a convenient information propagation technique enhancing incentive-compatibility and bandwidth efficiency. The transition from a centralized system into distributed one brings along game theoretical concerns. Especially for the permissionless blockchains, information propagation should be incentive-compatible just like any other communication or computational costly operation. Another important issue is that information is relayed via gossip-like protocols causing excessive bandwidth usage. Each information is propagated at least twice: first to advertise its existence, second to announce that it is final and validated, i.e., added to the block.

In this work, we investigate two distinct aspects of the information propagation of the blockchains: incentive and routing mechanisms. For the former part, we analyze the necessary and sufficient conditions of the Sybil-proof incentive-compatible propagation methodology. We show the impossibility result of the Sybil-proofness in 1-connected network model. For the rest, we prove that the propagation decision is independent of the capabilities of the receiving side. Then, we formulate the generic fee sharing function which encourages rational participants to propagate information. Regarding the bandwidth efficiency, we study a special type of consensus protocols where the block owner (round leader) is validated before the block is created. We present a smart routing mechanism which the redundant communication cost from the size of the network to the scale of average shortest path length. Finally, we combine the incentive and routing mechanisms in a storage-efficient way.

Index terms— Blockchain, information propagation, incentive, Sybil-proof mechanisms, routing

*o.ersoy@tudelft.nl

1 Introduction

After acknowledgment of Bitcoin [1] as a digital cryptocurrency, blockchain has become a trending subject to not only the research community but also the industrial society because of the enormous application area. The blockchain can be defined as a decentralized immutable public ledger which is updated and secured in a distributed structure among the untrusted parties. This ledger consists of ordered blocks, which may be composed of transactions like in Bitcoin or smart contracts as in Ethereum [2].

There are variety of blockchain types trying to achieve decentralization and immutability in different settings. Main distinguishing characteristics are the access structure and the consensus protocol. Regarding the authorization, there are two types of blockchains: permissionless where anyone can join and contribute to the chain, and permissioned which requires an authorization to join the network. Consensus protocols are used to validate the ledger and to maintain consistency in the ledgers stored individually. In general, consensus on a chain can be achieved by either a consortium (e.g. Byzantine agreement [3]) among (subset of) all participants or the implicit validation of the legitimate blocks (e.g. Bitcoin). In any case, all information is advertised to the round leader(s), whose identity may not be available, after the validation it is included to the block and broadcast.

Authorization of the blockchain influences the other components of the blockchain as well. Since permissionless blockchains consist of anonymous or pseudonymous users, they require Sybil-proof consensus protocol and incentive-compatible participation [4]. First, in order to prevent Sybil attacks, participation or validation in a permissionless blockchain should require proof-of-something that cannot be fabricated like mining work [1], possession of stake [5] or storage of data [6]. Second, participation in each process of the system should be complied with the rational behavior. There should be an incentive for each non-free operation requiring communication or computational cost.

Rationalism. There is no surprise that transition from centralized system into distributed one brings along game theoretical concerns [7]. Indeed, rational behavior of the users have been observed in peer-to-peer networks [8, 9, 10]. Importance of the game theoretical analysis in blockchain technology is already demonstrated by the selfish mining attack [11]. In the long term, altruistic behavior is not sustainable for the permissionless blockchains, especially for cryptocurrencies [12, 13]. Existing blockchains, including Bitcoin and Ethereum, reward only the one who validate the transactions, not the ones who propagate them. Yet, the incentive for the propagation of these transactions is altruistically fulfilled [14]. Especially for the permissionless blockchains, information propagation should be incentive-compatible just like any other non-free operation.

Existing applications. The aforementioned problem seems to be underrated and mostly ignored since the most successful permissionless blockchain applications work without any direct incentive to propagate. If we continue with the Bitcoin example, there are two consecutive reasons for that: altruism,

and centralized mining. Bitcoin is evolved from altruistic participation into centralized mining pools. Early adopters were enthusiastic and contributed to information propagation regardless of the utility [14, 15]. At the moment, conversely, the environment turned into rational but centralized system where only a few number of pools control the network [13, 16, 17]. For now, a client needs to reach a node of the pool, instead of the all miners in that pool, then internal propagation is controlled by the pool. Consequently, the existing applications do not exhibit purely decentralized blockchain properties and do not conflict with the necessity of the incentive.

Incentive of the propagation is also relevant with the bounty on the transaction, namely transaction fee. Block rewards for the valid block generations do not encourage to validate the transactions with zero transaction fees [12]. Transaction fees encourage the participants to validate and add them to the block. Knowledge of a transaction is as precious as its transaction fee. The more it is known by the network, the less probability that a specific node has a chance to profit by its fee. For that reason, a rational participant has an incentive **not** to share the incoming transaction knowledge with the rest of the network. Contradictorily, the more a client pays for a transaction fee, the less rational nodes are willing to propagate, thereby the later it will be on the blockchain. The same problem holds for all permissionless blockchains supposed to work in a rational participation assumption.

Lack of incentive in information propagation in the peer-to-peer networks has been studied in the last decades [18, 19, 20, 21]. The proposed solutions are not applicable for the permissionless blockchains. In peer-to-peer solutions, participants are asked to provide a specific datum like the position of a peer or the answer of a query. In blockchains, it is requested to validate the transactions and place them into a valid block. In the latter case, participants can always compete for the reward without propagating, whereas in the former one, peers who have the datum do not need to propagate anymore and the others must propagate to have a chance. Recently, blockchain oriented propagation mechanisms have been proposed [22, 23]. Babaioff et al. [22] presented the lack of incentive to propagate transactions in the Bitcoin, and they provided a solution for the d-ary directed tree networks. In [23], the authors investigate the propagation of not only the transaction but also the block itself.

Routing. Another missing part regarding the propagation is the ineffective routing of the information from a provider to a validator. In a centralized system, propagation can be handled very efficiently by a predefined routing mechanism because location of the server is known and stable. For instance, in a network with millions of participants, direct routing would reduce the cost from order of millions to order of tens. This comes from the small world phenomenon [24], which states that the distance between any two nodes is extremely small regarding the network size. The dynamic structure of blockchain prevents from having stationary efficient information propagation routes.

It is reasonable and necessary to broadcast all validated information through the network since the public ledger is stored and validated by all contributing parties. In the existing blockchains, there is additional broadcasting of the

information before being validated. At first, every information is propagated throughout the network, then one party or a consortium validates it and after validation it is again propagated within the block. Even more redundancy is caused by the flooding of each information because a node will receive the same information from different neighboring nodes. This additional cost is already reduced by sending information hash to check whether the neighbor has it or not. If the size of the information is relative to hash, then the cost of a broadcasting would be significantly more than double of the network size.

For some protocols like Nakamoto consensus [1], the redundancy is inevitable because of the unpredictability of the block owner (round leader) who will create the new block. This is caused by the fact that the block owner is simultaneously validated with his proposed block. Nevertheless, there have been recent proposals where the legitimate block owner can be validated before the block is proposed [23, 25, 26, 27, 28], which we call *first-leader-then-block* consensus protocols. In a first-leader-then-block structure, it is possible to overcome redundant communication cost by routing information from a client to the round leader.

1.1 Related Work

The lack of information propagation incentive in a peer-to-peer network has been known and studied in different settings [18, 19, 20, 21]. Kelinberg and Raghavan [20] proposed an incentive scheme for the answer to a query in a tree network model. Li et al. [21] focused on the peer discovery in a homogeneous network where each peer has the same probability to be a provider. Both systems try to locate a node (or a set of nodes) in the network, which has two main differences from the blockchain information propagation: participants do not compete against the ones who forwarded the message to them and participants cannot make up an answer for a query or location of a peer, i.e., either they have the right answer or not. Whereas in blockchain, every node is a leader candidate who can find the valid block which corresponds to the answer to a query.

In [18, 19], the authors analyzed the incentive problem for multi-level marketing which rewards referrals if their advertisement produces a purchase. In these marketing models, the reward is shared among all participants in the tree including the propagation path. Conversely, in our model, and [22, 23] which are focused on blockchain, it is shared between only the ones in the propagation path which is the direct path between the client and the leader. Besides all, an important difference between message propagation in blockchain and the other peer-to-peer systems is that a blockchain participant has an incentive not to propagate the message whereas others do not since they cannot generate the required information by themselves.

In [22], Babaioff et al. uncovered the incentive problem in the Bitcoin system where a rational miner has no incentive to propagate a transaction. They focused on a specific type of network, namely regular d-ary directed tree with a height H, and assumed that participants have the same processing power.

In this setting, they presented an incentive scheme and proved that it is also Sybil-proof. Abraham et al. recently proposed a consensus mechanism, Solidus, offering an incentive to propagate transactions and validated blocks (puzzles) [23]. In their proposal, the amount of processing fee passed to the next node is determined by the sender. Both works adopted signature chaining mechanisms to prevent any manipulation over the path and thereby secure shares of each contributors. Regarding the game theoretical analysis, former one [22] assumes tree structure which eliminates competition for the common neighbors, whereas the latter one [23] analyzes only the case of competitions between nodes for shared neighbors.

To the best of our knowledge, there is no existing work on a direct routing mechanism for dynamic blockchain networks. Nonetheless, Li et al. [21] presented a distributed routing scheme having the same structure as our solution. The main difference is that they focused on one-to-one routing which is dedicated to a single target, whereas we do one-to-all, which connects the complete network to the round leader. In addition, unlike [21], we insert alternative routes as a precaution and analyze the failure probability of these temporary routes caused from momently failing nodes.

1.2 Our Contributions

In this work, we investigate two information propagation related problems of blockchains: incentive and bandwidth efficiency. We present a generic incentive mechanism for information propagation, a routing mechanism compatible with first-leader-then-block consensus protocols and finally we combine them together in an efficient way.

Firstly, we formulate the ideal incentive function allocating the reward among the propagating participants for generic network model. In this manner, we obtain the following results:

- *Impossibility*: In 1-connected networks, it is not possible to design Sybil-proof fee sharing function which allocates the fee among all the contributors.
- *Equity*: Propagation decision of a node is independent from the neighbors' capacities. A rational node would propagate to either all of its neighbors or none of them.
- *Rationality:* Each rational node having probability less than C for being the round leader (among the ones having the transaction) will propagate the transaction given the following formula:

$$f^k_{[i]} = \begin{cases} F \cdot C(1-C)^{i-1} & \text{for } i < k, \\ F \cdot (1-C)^{k-1} & \text{for } i = k, \end{cases}$$

  where F is the total processing fee of the transaction, k is the length of the propagation path from a client to the round leader, and $f^k_{[i]}$ is the share of the i^{th} node in that path.

Secondly, we present a routing mechanism which reduces the communication cost of the information propagation from the size of the network to the scale of average shortest path length. It corresponds to a logarithmic scale of the network size for the random network models. This mechanism is compatible with all blockchains where the round leader can be verified before the block is announced. In addition, we analyze the failure probabilities of a transaction to reach the round leader in the presence of the node who may fail or censor.

Finally, we present an effective message propagation protocol which combines our incentive and routing mechanisms in a storage-efficient way.

The rest of the paper is organized as follows: Our blockchain model and notations are defined in Section 2. Section 3 formulates requirements of the incentive problem and computes the generic solution. Smart routing mechanism is presented in Section 4 and combined with incentive mechanism in Section 5. Finally, Section 6 concludes the paper.

2 Our Blockchain Model and Notations

In this section, we present general characteristics of our blockchain model and our notation. To keep it simple and consistent, in the rest of the paper, we will use the term transaction propagation, yet it can be replaced with any kind of information propagation.

Access structure. There are two types of blockchains: permissionless (public) and permissioned. In permissionless blockchain, anyone can join to the network and contribute without an authorization. Permissioned blockchain, conversely, has an authority deciding who can contribute to the ledger. Our work does not have a requirement in this manner, and we focus on the generic case: permissionless blockchain.

Participants. Since it is a permissionless blockchain, anyone can participate and contribute to the ledger directly. Moreover, there is no discrimination between participants, i.e., they all have abilities to prepare a transaction as a client and to propose a block as a round leader. For identification, each participant has a public and private key pair, and can be validated by his public key.

Network. It is a peer-to-peer network and each participant is matched with a node.

Consensus and leader election. Incentive mechanism defined in Section 3 works regardless of the consensus structure. Whereas, the routing mechanism requires special treatment, which we called *first-leader-then-block* consensus protocols. First-leader-then-block (*FLTB*) protocols can be defined as the consensus model where the round leader is validated before he proposes the block. Any leader election mechanism which is independent of the non-validated transactions can be converted into *FLTB* type. Examples of the *FLTB*-based blockchains are Bitcoin-NG [27] and several PoS-based ones [23, 25, 26, 28].

The rest of the definitions and notations are listed below:

- *Node*: A participant of the blockchain in the network. We may use the terms node and participant interchangeably.
- *Neighboring*: Direct connection in the network, adjacency in graph.
- *Client*: The source or the sender of a transaction. Client of a transaction T, denoted by c_T.
- *Round Leader*: The legitimate participant responsible to construct the block.
- *Intermediary Node*: A node on the transmission path between the round leader and a client.
- $\mathcal{L}^r$: The leader credential which validates the round leader for round r and can be verified by all nodes in the network. For example, it could be a special hash value in a Proof-of-work (PoW) protocol or the proof of possessing the chosen coin in a Proof-of-Stake (PoS) consensus. In general, regardless of the consensus mechanism, credentials are linked to the public key of the leader, and can be verified by a corresponding signature.
- $\pi(n_i)$: The probability of node n_i being the round leader, also referred as the capacity of node n_i. It corresponds to the mining power in PoW or the stake size in PoS protocols and is assumed to be greater than zero for every node in the network. $\pi(S)$ corresponds to the total probability of the all nodes in set S.
- $\mathcal{N}_K^T$: The set of nodes who know (have) the transaction T. $\mathcal{N}_K^{n,T}$ presents the set from the point of view of node n (including n itself).
- $\mathcal{N}_{NK}^T$: The set of nodes who do not know (have) transaction T yet. $\mathcal{N}_{NK}^{n,T}$ denotes the set from the point of view of node n and includes only the neighbors of n.

3 Incentive Mechanism

In a permissionless setting, incentive-compatibility and rational behavior have been already observed [4] and studied in information propagation manner [22, 23]. Conventional incentive instrument, namely transaction fee, almost always refers to the reward of the round leader. Here, we describe *processing fee* which consists of the reward to propagate and to validate transactions. Thereby, rational participants are encouraged to not only validate transactions but also propagate them. How to determine the fee is out of the scope of this paper but we assume that each processing fee is predefined by either the client or a known function. We focus on how to automatically allocate the fee among all the contributors of the process.

Fee sharing function. The fee sharing function allocates the processing fee of a transaction among the propagating nodes and the round leader. Suppose

that k nodes are involved in the processing of a transaction with processing fee F, where $k-1$ of the nodes are in the direct path between the client and the round leader. $f^k_{[i]}$ denotes the share of i^{th} node in the propagation path, $f^k_{[k]}$ is the share of the round leader which corresponds to the conventional transaction fee, and $\sum_{i=1}^{k} f^k_{[i]} = F$.

In the rest of the section, we formulate the necessities of the fee sharing function used to share an arbitrary transaction T with a fee F among propagating participants and the round leader. An ideal incentive function should satisfy the following properties:

1. *Sybil-proofness*: An intermediary node as well as the round leader should not benefit from introducing Sybil nodes to the network.

2. *Game theoretically soundness*: A transaction should not be kept among a subset of the network. There should be adequate incentive for rational nodes willing to propagate, thence it will eventually reach to the whole network.

By formulating these conditions, we achieve the following theorem (where C is a constant which can be chosen according to the network connectivity):

Theorem 1. *In a 2- or more connected blockchain network, each rational node $n \in \mathcal{N}_K^T$ with $\pi(n) < C \cdot \pi(\mathcal{N}_K^{n,T})$ propagates message T without introducing Sybil nodes, if the processing fee F is shared by the following method:*

$$f^k_{[i]} = \begin{cases} F \cdot C(1-C)^{i-1} & \text{for } 1 \leq i < k, \\ F \cdot (1-C)^{k-1} & \text{for } i = k. \end{cases}$$

Proof of the theorem is divided into the following sections. The requirements are formulated in Sections 3.1 and 3.2, and the fee sharing function satisfying them is computed in Section 3.3.

3.1 Sybil-Proofness

Here, we use the same definition of Sybil nodes in [22]: fake identities sharing the same neighbors with the original node that do not increase connectivity of the network. Because of the Sybil-proof consensus algorithm, Sybil nodes do not increase the capacity of their owner, i.e., the probability of being the round leader.

We investigate the problem in two different settings: 1-connected networks and the rest. k-connected network means that removal of any $k-1$ nodes does not disconnect the network. In 1-connected networks, there exist a bridge which is the only connection between two distinct subnetworks. Though 1-connected network model seems to be unrealistic topology for permissionless blockchains, it is important to see the intuition behind the non-competition effect.

1-connected networks. In 1-connected networks, there are critical nodes which have special positions in the propagation paths between some node pairs.

A critical node for a node pair appears in all possible paths between these two nodes. The following lemma shows that non-competing advantage of critical nodes makes it impossible to have a Sybil-proof incentive mechanism for 1-connected networks.

Lemma 2 (Impossibility Lemma)**.** *In order to deviate nodes from introducing Sybil nodes in 1-connected networks, processing fee should be shared between the first propagating node and the round leader.*

Proof. Assume that, because of 1-connectedness of the network, a node n may have a critical position for a transaction T, meaning that it is certain he will be included in the propagation path of that transaction.

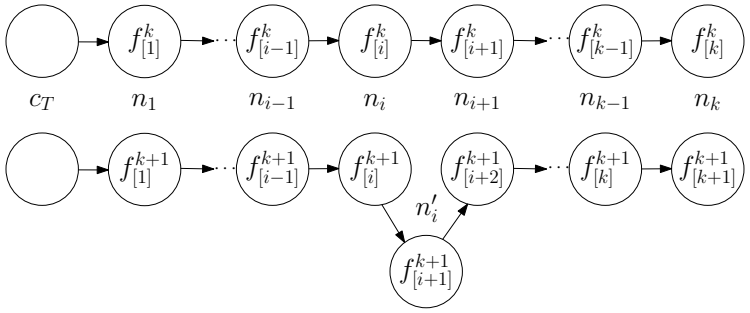

Figure 1: The fee sharing before and after a Sybil node $n_{i'}$ added by the node n_i

Now, we investigate the share of a node n_i with and without a Sybil node. As given in Figure 1, n_i is the i^{th} node in the propagation path and his corresponding fee shares are $f^k_{[i]}$ and $f^{k+1}_{[i]} + f^{k+1}_{[i+1]}$. In order to demotivate n_i, $f^k_{[i]}$ should be greater than or equal to $f^{k+1}_{[i]} + f^{k+1}_{[i+1]}$. Since the position of node would change for different transactions and rounds, the condition should hold for all positions:

$$
\begin{aligned}
\forall\, i \in \{1, \ldots, k\}, \qquad & f^k_{[i]} \geq f^{k+1}_{[i]} + f^{k+1}_{[i+1]} \\
\text{(summing for all } i\text{'s)} \implies \qquad & \sum_{i=1}^{k} f^k_{[i]} \geq \sum_{i=1}^{k} f^{k+1}_{[i]} + \sum_{i=1}^{k} f^{k+1}_{[i+1]} \\
\text{(Definition of } \mathcal{F}) \implies \qquad & F \geq F - f^{k+1}_{[k+1]} + F - f^{k+1}_{[1]} \\
\implies \qquad & f^{k+1}_{[k+1]} + f^{k+1}_{[1]} \geq F \\
\text{(Definition of } \mathcal{F}) \implies \qquad & f^{k+1}_{[k+1]} + f^{k+1}_{[1]} = F\,.
\end{aligned}
$$

Therefore, there will not be any incentive for the rest to propagate T which contradicts with rational behavior. □

Eclipse and partitioning. Note that this monopolized behavior is similar to the eclipse and partitioning attacks where the adversary separates the network into two distinct group and controls all the connections between them [29, 30]. Indeed, Lemma 2 can be generalized to the case where the adversary is able to control all the outgoing connections of a client. In that case, there is no way to deviate the adversary from creating Sybil nodes for that specific transaction. We assume that client nodes are able to defend against the eclipse attacks using the countermeasures defined in [29].

2- or more connected network. In a 2-connected network, there are multiple paths between any two nodes, including the client and the round leader. Therefore, we can immediately focus on the multiple paths case where there are competing paths for the same transaction and the round leader includes one of them to the block. As there are multiple paths/options, nodes can be demotivated from introducing Sybil nodes by following conditions:

- *Intermediary nodes*: If share of the round leader decreases as the propagation path length increases, then he will choose the shortest path for each transaction. In that case, introducing Sybil nodes will decrease his chance to be included in the block. Therefore, providing larger gain to the leader for choosing the shortest path is sufficient and can be formulated as $f^{k}_{[k]} > f^{k+1}_{[k+1]}$.
- *Round leader*: In some cases, round leader is determined before the block is created or even several rounds earlier [25, 26, 27]. Since the round leader is guaranteed to be in the propagation path, it is needed to be taken into account separately. In the case of s Sybil nodes, his share will change from $f^{k}_{[k]}$ to $\sum_{i=0}^{s} f^{k+s}_{[k+i]}$ for some k. For Sybil-proofness against the round leader, $f^{k}_{[k]} \geq \sum_{i=0}^{s} f^{k+s}_{[k+i]}$ is required.

Since the latter condition includes the former one ($f^{k+1}_{[k]} > 0$), Sybil proofness condition can be formulated as:

$$\forall\, k \geq 1, \forall\, s \geq 1 \quad f^{k}_{[k]} \geq \sum_{i=0}^{s} f^{k+s}_{[k+i]}\,. \tag{1}$$

3.2 Game Theoretically Soundness

The decision of the propagation of a transaction can be analyzed as a simultaneous move game where each party takes action without knowing strategies of the others. All players (nodes in our case) are assumed to be rational and they decide their actions deducing that the others will also act rationally. Some nodes may cooperate with each other. We assume that colluding neighboring nodes already share everything with each other and take actions as one. In other words, they act like a single combined node in the network which can be seen as Sybil nodes.

Here, we investigate the sharing (propagating) decision by comparing the change in the expected rewards for a transaction T. At the beginning, each

transaction is shared with some nodes, at least with the neighbors of the client. We will find the required condition to propagate through the whole network. We first investigate the propagation decision by comparing the change in the expected rewards immediately after the action. Then, we extend our analysis with a permanence condition which guarantees that the ones who propagate will not suffer from any future actions.

We show that sharing decision of a node is independent from the probability of his neighboring nodes being the round leader. Instead, it depends on his own probability against the rest who knows the transaction.

Lemma 3 (Equity Lemma)**.** *Propagation decision of a node is independent from the neighbors' capacities. A rational node would propagate to either all of its neighbors or none of them.*

Proof. Let a transaction T (with processing fee F) is known by a node n, and its distance to the c_T is k. The expected reward of node n can be defined as a function $R(\cdot)$ whose input corresponds to the capacities of the nodes who received T from n, then

$$R(X) = \frac{f_{[k]}^{k} \cdot \pi(n) + f_{[k]}^{k+1} \cdot X}{\pi(\mathcal{N}_K^{n,T}) + X}.$$

We show that $R(\cdot)$ is a monotone function. In order to show that a function is monotone, it is enough to show that the sign of its derivative does not change in the domain range. For our case, it can be seen that the sign is independent of the input:

$$\begin{aligned} R'(X) &= \frac{f_{[k]}^{k+1}\left(\pi(\mathcal{N}_K^{n,T}) + X\right) - \left(f_{[k]}^{k}\pi(n) + f_{[k]}^{k+1}X\right)}{\left(\pi(\mathcal{N}_K^{n,T}) + X\right)^2} \\ &= \frac{f_{[k]}^{k+1}\pi(\mathcal{N}_K^{n,T}) - f_{[k]}^{k}\pi(n)}{\left(\pi(\mathcal{N}_K^{n,T}) + X\right)^2}. \end{aligned}$$

Since $R(\cdot)$ is a monotone function, then it achieves the maximum value at one of the boundary values. In our case, the boundary values are $X = 0$ where no neighbors received the transaction and $X = \pi\left(\mathcal{N}_{NK}^{n,T}\right)$ where all neighbors received it. Here, we omit the fact that $\pi(\cdot)$ is also a monotone function. Thereby, we can say that a rational node would maximize the profit by propagating to either all of its neighbors or none of them. □

Lemma 3 states that a rational node decide to propagate to either all of his neighbors or none of them. This result simplifies to evaluate interfering multiple node decisions.

Lemma 4 (Propagation Lemma)**.** *Let a node $n \in \mathcal{N}_K^T$, $\mathcal{N}_{NK}^{n,T} \neq \emptyset$ where the distance between n and c_T is k. All neighbors of n will be aware of T if*

$$\frac{f_{[k]}^{k+1}}{f_{[k]}^{k}} > \frac{\pi(n)}{\pi(\mathcal{N}_K^{n,T})}.$$

Proof. Assume that some of the neighbors of n are not aware of T, i.e., $\mathcal{N}_{NK}^{n,T} \neq \emptyset$. From Lemma 3, we know that n did not propagate the transaction to any of his neighbors. Therefore, at the moment, the only way that he profits from T is being the round leader with a reward $f_{[k]}^{k}$.

Table 1: Expected reward of n from T regarding possible decisions of n and the rest of $\mathcal{N}_K^{n,T}$.

		$\mathcal{N}_K^{n,T}$ (excluding n)	
	Decision	not Propagate	(some) Propagate
n	not Propagate	$\frac{f_{[k]}^{k}\cdot\pi(n)}{\pi(\mathcal{N}_K^{n,T})}$	$\frac{f_{[k]}^{k}\cdot\pi(n)}{\pi(\mathcal{N}_K^{n,T})+\pi(CN)+\pi(NCN_2)}$
	Propagate	$\frac{f_{[k]}^{k}\cdot\pi(n)+f_{[k]}^{k+1}\cdot\pi(\mathcal{N}_{NK}^{n,T})}{\pi(\mathcal{N}_K^{n,T})+\pi(\mathcal{N}_{NK}^{n,T})}$	$\frac{f_{[k]}^{k}\cdot\pi(n)+f_{[k]}^{k+1}\cdot\pi(NCN_1)+\alpha f_{[k]}^{k+1}\cdot\pi(CN)}{\pi(\mathcal{N}_K^{n,T})+\pi(\mathcal{N}_{NK}^{n,T})+\pi(NCN_2)}$

Table 1 presents expected reward of n with respect to each possible action of n and $\mathcal{N}_K^{n,T}$. Propagation decision of $\mathcal{N}_K^{n,T}$ may not include of all its members, thereby all possible decisions are taken into account. Here, CN corresponds to the common neighbors of n and $\mathcal{N}_K^{n,T}$, NCN_1 distinct neighbors of n and NCN_2 distinct neighbors of $\mathcal{N}_K^{n,T}$ (who decide to propagate), i.e., $CN \bigcup NCN_1 = \mathcal{N}_{NK}^{n,T}$. Since CN is received the transaction from both n and the rest of the $\mathcal{N}_K^{n,T}$, α represents the percentage of the ones in CN decided to continue with the one including n.

If all participants of $\mathcal{N}_K^{n,T}$ decide not to propagate with their neighbors, then n will benefit from propagating T in the case of

$$\frac{f_{[k]}^{k}\cdot\pi(n)+f_{[k]}^{k+1}\cdot\pi(\mathcal{N}_{NK}^{n,T})}{\pi(\mathcal{N}_K^{n,T})+\pi(\mathcal{N}_{NK}^{n,T})} > \frac{f_{[k]}^{k}\cdot\pi(n)}{\pi(\mathcal{N}_K^{n,T})} \iff \frac{f_{[k]}^{k+1}}{f_{[k]}^{k}} > \frac{\pi(n)}{\pi(\mathcal{N}_K^{n,T})}.$$

If (some participants in) $\mathcal{N}_K^{n,T}$ decide to propagate T, then n will benefit from propagating T in the case of

$$\frac{f_{[k]}^{k}\cdot\pi(n)+f_{[k]}^{k+1}\cdot\pi(NCN_1)+\alpha f_{[k]}^{k+1}\cdot\pi(CN)}{\pi(\mathcal{N}_K^{n,T})+\pi(\mathcal{N}_{NK}^{n,T})+\pi(NCN_2)} > \frac{f_{[k]}^{k}\cdot\pi(n)}{\pi(\mathcal{N}_K^{n,T})+\pi(CN)+\pi(NCN_2)}$$
$$\Longleftarrow \frac{f_{[k]}^{k+1}}{f_{[k]}^{k}} > \frac{\pi(n)}{\pi(\mathcal{N}_K^{n,T})+\pi(CN)+\pi(NCN_2)} \quad \text{and} \quad NCN_1 \neq \emptyset.$$

Note that $NCN_1 = \emptyset$ means that all the neighbors of n are also neighbors of $\mathcal{N}_K^{n,T}$ who decide to propagate. Therefore, in any case, if $\frac{f_{[k]}^{k+1}}{f_{[k]}^{k}} > \frac{\pi(n)}{\pi(\mathcal{N}_K^{n,T})}$ is satisfied, then all neighbors of n will be aware of the transaction. □

Corollary 5. *Let $f_{[k]}^{k+1} \geq C \cdot f_{[k]}^{k}$ for some constant $C \in (0,1)$. $\mathcal{N}_K^T$ will continue to expand until there is no more node $n \in \mathcal{N}_K^T$ having neighbors in $\mathcal{N}_{NK}^T$ and satisfying $\pi(n) < C \cdot \pi(\mathcal{N}_K^{n,T})$.*

Remark I. Here, it is possible to define different C_k values for each distance k, i.e., $f_{[k]}^{k+1} \geq C_k \cdot f_{[k]}^{k}$. One might argue that, as the distance increases, it could be possible to find nodes satisfying $\frac{\pi(n)}{\pi(\mathcal{N}_K^{n,T})} < C_k$ for smaller C_k values. However, as seen in Section 5, this is not always the case. In addition, the intermediate node may not know the exact distance, thus using the same C value would make the decision simpler.

Remark II. Note that the propagation decision is based on $\mathcal{N}_K^{n,T}$ instead of $\mathcal{N}_K^T$ since the latter one may not be available. This could lead to better consequences for propagation because nodes may predict $\mathcal{N}_K^T$ greater than its actual size and decide accordingly. Nonetheless, a carefully chosen C value will lead the nodes to share it with an overwhelming probability.

Remark III. Being the round leader should be more appealing than being an intermediary node, thus the round leader would try to fulfill the round block capacity to maximize his profit. The system may not work in full capacity if the nodes gain the same reward from propagating instead of validating (as the round leader) transactions. In Corollary 5, the propagation condition is given as $f_{[k]}^{k+1} \geq C \cdot f_{[k]}^{k}$. We fix the condition in favor of the round leader:

$$\forall\, k, \quad f_{[k]}^{k+1} = C \cdot f_{[k]}^{k}\,. \tag{2}$$

Permanence condition. In the simultaneous move analysis, we investigated one step at a time, i.e., what will happen immediately after the decision of propagation. However, all possible future actions should be taken into account. For example, the sender of a transaction should consider possibility of the further propagation done by the receiver. From Lemma 3, capacities of the neighboring nodes does not have any influence on the sharing decision. Unless the processing fee share decreases, which is caused by some possible future actions like increased path length, the same lemma will be satisfied. If the share of a propagating node is non-decreasing with respect to the path length, then the ones who propagate will not suffer from any future actions. This can be formulated as

$$\forall\, i < k, \quad f_{[i]}^{k} \geq f_{[i]}^{k+1}\,. \tag{3}$$

3.3 Fee Sharing Function

With the equations obtained from the required conditions, we can uniquely determine the fee sharing function and conclude Theorem 1. First, using per-

manence condition (3), Sybil-proofness condition (1), can be reduced to $f^k_{[k]} \geq f^{k+1}_{[k+1]} + f^{k+1}_{[k]}$:

$$\begin{aligned}
\forall\, k \geq 1, \quad f^k_{[k]} &\geq f^{k+1}_{[k+1]} + f^{k+1}_{[k]} \geq f^{k+2}_{[k+2]} + f^{k+2}_{[k+1]} + f^{k+1}_{[k]} \\
&\geq f^{k+3}_{[k+3]} + f^{k+3}_{[k+2]} + f^{k+2}_{[k+1]} + f^{k+1}_{[k]} \geq \cdots \\
\forall\, s \geq 1, \qquad &\geq f^{k+s}_{[k+s]} + \sum_{i=0}^{s-1} f^{k+i+1}_{[k+i]} \geq f^{k+s}_{[k+s]} + \sum_{i=0}^{s-1} f^{k+s}_{[k+i]} \\
&= \sum_{i=0}^{s} f^{k+s}_{[k+i]}\,.
\end{aligned}$$

Therefore, we can update the Sybil-proofness condition as:

$$\forall\, k \geq 1, \quad f^k_{[k]} \geq f^{k+1}_{[k+1]} + f^{k+1}_{[k]}\,. \tag{4}$$

Then, we can obtain the following equations:

$$\begin{aligned}
\text{Using (4)} \qquad & \sum_{i=1}^{k} f^i_{[i]} \geq \sum_{i=1}^{k} f^{i+1}_{[i+1]} + \sum_{i=1}^{k} f^{i+1}_{[i]} \\
\implies \quad & F = f^1_{[1]} \geq f^{k+1}_{[k+1]} + \sum_{i=1}^{k} f^{i+1}_{[i]} \\
\text{Using (3)} \quad \implies \quad & F \geq f^{k+1}_{[k+1]} + \sum_{i=1}^{k} f^{k+1}_{[i]} = F \\
\implies \quad & f^k_{[i]} = f^{k+1}_{[i]} \text{ and } f^k_{[k]} = f^{k+1}_{[k+1]} + f^{k+1}_{[k]}.
\end{aligned} \tag{5}$$

After all, we can finalize the fee sharing function which corresponds to Theorem 1. Using (2) and (5), the share of the round leader can be computed:

$$f^k_{[k]} = f^{k-1}_{[k-1]}(1-C) = \cdots = F \cdot (1-C)^{k-1}. \tag{6}$$

Using (5) and (6), the share of an intermediary node can be computed:

$$\forall\, i < k, \quad f^k_{[i]} = f^{i+1}_{[i]} = F \cdot C(1-C)^{i-1}\,.$$

3.4 Discussion

Integration. Implementation of the incentive mechanism should take into account the security and efficiency concerns. The propagation path should be immutable in a way that an adversary cannot add or subtract any node neither in the propagation process nor during the block generation. At the same time, storage efficiency is also essential since these path logs are needed to be stored in the ledger by every node. Both existing incentive-compatible blockchain solutions adopted a signature chaining mechanism where each propagated message

includes public key of the receiver and signature of the sender [22, 23]. This protocol prevents any manipulation over the path and thereby secures the shares of each contributors. It requires additional storage which is the signatures of the contributors. Although signature chaining solution requires the knowledge of the public key of the receiver and stores signatures of each sender, it is generic and can be applied to any blockchain. In Section 5, we present a novel and storage-efficient solution which is feasible for *FLTB* blockchains. It is embedded into routing mechanism and does not require the knowledge of the public keys of the neighboring nodes.

Determining C parameter. C value plays an important role to make sure that there will be incentive to propagate a transaction for some nodes until it reaches to the whole network. On the one hand, as the choice for the C value increases, it will be easier to satisfy the propagation condition since there will be more chance to find nodes satisfying $\pi(n) < C \cdot \pi(\mathcal{N}_K^T)$. On the other hand, the higher C value, the lower fee remains for the rest of the propagation path. It significantly reduces the fee of the round leader, thereby the incentive. For these reasons, it is required to choose a moderate C value, e.g., a reasonable choice would be $C = \frac{2}{N_{con}}$ where N_{con} denotes default number of connections of a node. For example, in Bitcoin network where $N_{con} = 8$, nodes will propagate unless they assume that their mining power is greater than 25% of the ones having the transaction. Even at the very beginning, at least N_{con} nodes have the information, $C = \frac{2}{N_{con}}$ setting would provide overwhelming probability to have nodes willing to propagate according to Corollary 5.

Client ($0-$capacity) nodes. The main goal of the propagation incentive mechanism is to make sure that the transactions are received by the nodes who are capable of validating transactions as well as creating blocks. For that reason, we mainly focused on the nodes having capacity greater than zero, i.e., $\pi(\cdot) > 0$. Nevertheless, a client node can be seen as a potential capacity node because of the possible propagation of the client. Regarding Lemma 3 and permanence condition (3), a rational node, who decided to propagate, would benefit from propagating to the client nodes as well. At the same time, a client node will always benefit from propagating any transaction since otherwise it will not have any chance to gain a fee.

Decentralization effect. In the conventional permissionless blockchains, all rewards including block reward and transaction fees are given to the block owner. In other words, nodes have only one incentive to participate to the network: being round leader. The less chance individual participants have to be the round leader, the more they are motivated to join into centralized forms (e.g. mining pools) [31, 16]. Conversely, processing fee is shared with all propagators nodes. In addition, since many transactions are included in a single block, aiming processing fees of (some) transactions has significantly more chance than being the round leader. Thereby, it is reasonable to conclude that incentive mechanism would have positive impact on the decentralization of the permissionless blockchains.

4 Routing Mechanism

As a non-hierarchical peer-to-peer network, the blockchain ledger is validated by all participants individually. This requires to broadcast every data and blocks over the network since every node needs to keep record of the chain to validate new blocks. In existing permissionless blockchains, every transaction is broadcast throughout the network by the client, then the new block including (some of) these is constructed and broadcast by the round leader. Hence, each transaction is broadcast at least twice. Even more messages are send to check the neighbors' awareness on the transaction.

In Nakamoto-like consensus protocols, the round leader is validated simultaneously with his proposed block where the redundant propagation of the client is inevitable. In *FLTB* protocols, on the other hand, it is possible to validate the round leader before the block is proposed. It enables to determine a direct route between each client and the round leader. Our routing mechanism in Algorithm 1 finds the shortest paths between clients and the round leader for each round. Instead of sending each transaction to all nodes in the network, it is relayed over the shortest path between the client and the leader. The distance between (almost) any two nodes in a connected graph is dramatically smaller than the size of the network [24]. This is equivalent to cost reduction from $O(N)$ to $O(\ln N)$ in a random network of size N [32, 33].

Our protocol can be divided into two parts: *Recognition Phase* where the routes are determined and *Transaction Phase* where the transactions are propagated (see Figure 2). First, in the recognition phase, the round leader is recognized throughout the network and his credential is propagated with a standard gossip protocol. Each node n_i learns his closest node towards the round leader, *gradient node* (gn_i), who is the first node forwarding the credential. In the transaction phase, each client forwards his transaction to (some of) his neighbors. Then, each node, receiving a transaction for the first time, directly transmits to his gradient node. Here, the reason for clients to broadcast more than one neighbor is that one path could yield a single point of failure. It could be caused by the nodes who fail or maliciously censor some of the transactions. As presented in the experimental results, forwarding transaction to a few of the neighbors (precisely N_{con}) is sufficient. Note that, the routing mechanism works under asynchronous network assumptions since a client does not have to wait for all nodes but N_{con} of his neighbors. Similarly, for an intermediary node, waiting for the first credential message is enough to propagate received transactions.

Locational privacy. There have been several papers investigating anonymity in the permissionless blockchain networks, especially for the Bitcoin network [34, 35, 36]. It is found out that matching public keys and IP addresses can be done by eavesdropping. In this manner, *FLTB*-based blockchains may exposed to DoS (denial-of-service) attacks against to the round leader. We want to stress that our routing mechanism does not leak any more locational information about the position of the leader other than the *FLTB* protocols do. It just takes advantage of the announcement of the leader which is done exactly in the same manner with the basic *FLTB* protocols. Therefore, our routing

Algorithm 1 The Routing Mechanism

Recognition Phase
Leader provides his credential $\mathcal{L}^r$ to his neighbors.
for Node n_1 to n_N **do**
  if First time receiving $\mathcal{L}^r$ **then**
    Store ID of the sender (gradient) node n_j, i.e., $gn_i \leftarrow n_j$
    Propagate $\mathcal{L}^r$ to neighbors.
  end if
end for

Transaction Phase
Client provides transaction T to his neighbors.
for Each node n_i receiving T **do**
  if First time receiving T **then**
    Send it to the gn_i
  end if
end for

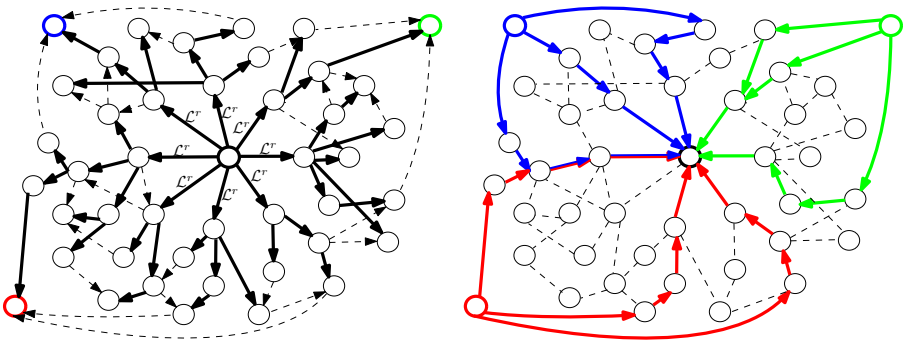

Figure 2: The Routing Mechanism. The left one illustrates the Recognition Phase and connections to the gradient nodes are shown with bold solid lines. On the right, three clients and their transaction paths are presented.

mechanism does not cause any additional vulnerabilities for DoS-like attacks against the round leader. Yet, it is possible to improve the locational privacy via anonymity phase where the message is first forwarded in a line of nodes, then diffused from there [37]. The extra cost of anonymity would be a few nodes on the line which is still proportional to the logarithmic size of the network.

4.1 Analysis

In order to simulate permissionless blockchain networks, Barabási-Albert (BA) [32] and Erdős-Rényi [33] graph models have been used (specifically for the Bitcoin) [38, 39]. In this manner, we combine both models for our simulation.

We use Barabási-Albert (BA) model [32] for our blockchain graph which simulates peer discovery in a peer-to-peer network. It starts with a well-connected small graph and each new node is connected to some of the previous nodes with a probability proportional to their degrees. We start with 50 nodes in Erdős-Rényi model [33] with edge probability of 1/2, meaning that on average each node has 25 connections. Then, each new node is added by connecting with N_{con} nodes in the network. For each (N, N_{con}) pair analyzed, we generated 1000 different graphs with 100 transactions using Python graph library [40].

Bandwidth gain. In [41], the average shortest path length between any two nodes, i.e., the average path length, of a BA graph is shown to be in the order of $\frac{\ln N}{\ln \ln N}$. Hence, our routing protocol reduces the communication cost of a message transaction from $O(N)$ to $O(N_{con} \cdot \frac{\ln N}{\ln \ln N})$. The communication gain is up to 99% for scaled networks (see Table 2), which can be verified by counting the average number of nodes visited per transaction. Here, we assume that the first arriving credential is coming from the node which is closest to the leader with respect to the number of nodes in between. In other words, the delay between any two nodes is computed by the node-distance.

Flooding. In Table 2, we count only one redundant communication for each information. Even more redundancy is caused by the flooding of each information because the same information is received from different neighboring nodes. In other words, the total redundancy is not N, but on average $N_{con} \cdot N$. This additional redundancy is already reduced by the sending the hash of the information to check whether the neighbor has it or not. However, if the size of the information is relative to hash, then the cost of a broadcasting would be significantly more than double of the network size. To conclude, since our mechanism does not suffer from the flooding effect, the actual communication gain would be much higher than the result in Table 2.

Table 2: Experimental results for APL: Average Path Length, ANVN: Average Number of Visited Nodes per transaction and Gain: Communication Gain.

$G(N, N_{con})$	G(100,8)	G(500,8)	G(1000,8)	G(2000,8)	G(2000,16)	G(5000,8)	G(5000,16)
APL	1.86	2.45	2.64	2.81	2.50	3.03	2.72
ANVN	11	14	15	16	29	18	32
Gain	89%	97%	98%	99%	98%	99%	99%

Failing transmissions. Since each transaction is propagated among a small set of nodes, we need to take into account the possibility of propagation failure which can be caused by the nodes who fail or censor the transaction. The failure probability of a transaction can be approximated by $\left(1-(1-h)^{\frac{\ln N}{\ln \ln N}-1}\right)^{N_{con}}$ where h denotes the probability of a node in the network who fails or censors the transaction. Table 3 shows that the percentage of the failing transactions is nearly negligible and can be reduced even more by increasing the number of paths.

Table 3: Experimental results for the percentage of a transaction that did not reach to the round leader in the presence of nodes failing or censoring the transaction with probability h.

$G(N, N_{con})$	G(100,8)	G(500,8)	G(1000,8)	G(2000,8)	G(2000,16)	G(5000,8)	G(5000,16)
$h = 0.2$	0.04%	0.30%	0.60%	1.09%	0.004%	1.57%	0.01%
$h = 0.3$	0.30%	1.63%	2.76%	3.62%	0.07%	5.49%	0.12%

5 Effective Message Propagation

In this section, we show how to deploy both of the incentive and routing mechanisms for any blockchain having a first-leader-then-block consensus protocol. At first glance, they seem to conflict with each other because the incentive mechanism is used to encourage propagation while the routing mechanism helps to reduce redundant propagation. We combine them in a way that rational nodes are encouraged to propagate only the transactions which are coming from the predefined paths of the routing mechanism. As demonstrated in Algorithm 2, we use the same infrastructure with the routing mechanism, and we include proofs of the intermediary nodes such that their contributions cannot be denied. Each transaction path is defined and secured by a path identifier which includes the public keys of the propagating nodes. Blocks consist of transactions as well as their path identifiers used to claim processing fee shares.

In the recognition phase, each intermediary node conveys the leader credential and the path identifier. Incoming and outgoing path identifiers of a node n are denoted by IN_n and OUT_n, which are used to validate and secure the propagation path. The round leader ℓ produces the initial identifier, $OUT_\ell = H(\mathcal{L}^r, PK_\ell)$, and propagates to his neighbors. Each node n updates the identifier coming from the gradient node by $OUT_n = H(IN_n, PK_n)$. This operation is done just for the gradient node (first one sending $\mathcal{L}^r$), then updated identifier and the credential are forwarded to the neighbors. Nodes may ignore the subsequent identifiers except a client who stores the first N_{con} ones for the transaction phase.

After the routing paths are determined, each client delivers the signed transaction and the incoming identifier to his N_{con} neighbors. The first receiving

nodes, check the signature, then add their public keys to the transaction and forward it to their gradient nodes. From that point, each intermediary node in the path first checks the validity of the path via the public keys included and his own identifier, then forwards the transaction including his public key to the gradient node.

Once transactions are received by the round leader, he includes the valid ones into the block. The block consists of the credential, hash of the previous block and valid transactions with their paths. Then, the block is propagated throughout the network.

Algorithm 2 Message Propagation Protocol

Recognition Phase
Leader l propagates $\mathcal{L}^r$
for Each node n_i **do**
  if First time receiving $\mathcal{L}^r$ and $IN_{n'}$ **then**
    if $\mathcal{L}^r$ is valid **then**
      Assign $IN_{n_i} \leftarrow IN_{n'}$ and gradient node as n'
      Compute $OUT_{n_i} = H(IN_{n_i}, PK_{n_i})$
      Propagate $\mathcal{L}^r$ and OUT_{n_i} to neighbors.
    end if
  end if
end for

Transaction Phase
Client c_T provides $Signed(T, IN_{c_T})$ (and $\mathcal{PK} = \emptyset$) to the first N_{con} gradient nodes.
for Each node n_i receiving $Signed(T, IN_{c_T})$ and $\mathcal{PK}$ **do**
  if First time receiving T **then**
    if Signature path holds **then**
      Update $\mathcal{PK} \leftarrow \mathcal{PK} \bigcup \{PK_{n_i}\}$
      Send $Signed(T, IN_{c_T})$ and $\mathcal{PK}$ to the gradient node.
    end if
  end if
end for

Incentive for block propagation. As a consequence of the incentive and routing mechanisms, intermediary nodes also have incentives to propagate the block since they share processing fees. Even more, the ones who are closer to the leader would have higher motivation since they probably gain from more transactions.

Storage efficiency. Any propagation incentive mechanism requires additional data storage than the data itself to keep track of the propagation path. Previous works having incentive [23, 22] utilize signature chains where each node signs the transaction and the public key of receiver. Therefore, additional to the transaction, the signature package of each propagating node is included. On the

other hand, our solution with the path identification benefits from the recognition phase of the routing protocol, and its additional storage requirement is only the public keys of propagating nodes and a signature of the client. Since the ability to claim propagation reward and the validation of the path need to be available, our propagation mechanism demands minimal storage components.

Privacy of the intermediary nodes. Signature chains and the proposed path identifier yield direct connection between nodes network ID and their public keys. Unlike signature chains, our solution consists of two phases and the propagating nodes validate it by checking whether their input is preserved or not. This enables us to tackle the privacy issue by replacing plain public keys with commitments. Instead of directly including a public key, each node can obscure it in a simple commitment with a random number ($CT_i = H(PK_i, R_i)$). All verifications can be handled with the commitments, while claiming propagation reward requires to reveal it. The commitment version uses the same network structure without compromising the identities of the nodes except clients and the round leader. The location of the round leader and clients will be known to their neighbors. They may need to update their key pairs or replace their connections for the next rounds.

6 Conclusion

In this work, we investigated two information propagation related problems of blockchains: incentive and bandwidth efficiency. We presented an incentive mechanism encouraging nodes to propagate messages, and a routing mechanism reducing the redundant communication cost.

We analyzed the necessary and sufficient conditions providing incentive to propagate messages as well as to deviate participants from introducing Sybil nodes. We studied different types of network topologies and we showed the impossibility result of the Sybil-proofness for 1-connected model. We formulated the incentive-compatible propagation mechanism, and proved that it obeys the rational behavior. We presented a new aspect of the consensus algorithms, namely first-leader-then-block protocols. We proposed a smart routing mechanism for these protocols, which reduces the redundant information propagation from the size of the network to the scale of average shortest path length. Finally, we combined both mechanisms in a compatible and efficient way.

Future work and open questions. In Section 3.4, we mentioned the parameter choice and possible outcomes of the incentive mechanism. Detailed effect of incentive model and parameter choice are left as a future work. Another open question is the effect of the incentive mechanism on the topology of the network. Nodes would benefit from increasing their connection to contribute more transaction propagations, i.e., it would increase the connectivity of the network. Using that result, a rigorous analysis on the choice of the C parameter can be done. Finally, there are open problems regarding the economics of the processing fee: analyzing the accuracy of the de facto formulas in the existing cryptocurrencies with respect to the cost of the propagation and validation and

investigating the possible impacts of the sharing the fee like decentralization effect.

References

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

[2] Gavin Wood. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum Project Yellow Paper*, 151, 2014.

[3] Leslie Lamport, Robert Shostak, and Marshall Pease. The Byzantine generals problem. *ACM Trans. Program. Lang. Syst.*, 4(3):382–401, July 1982.

[4] Gareth W Peters and Efstathios Panayi. Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money. In *Banking Beyond Banks and Money*, pages 239–278. Springer, 2016.

[5] Sunny King and Scott Nadal. Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper, August*, 19, 2012.

[6] Stefan Dziembowski, Sebastian Faust, Vladimir Kolmogorov, and Krzysztof Pietrzak. Proofs of space. In *Annual Cryptology Conference*, pages 585–605. Springer, 2015.

[7] Dov Monderer and Moshe Tennenholtz. Distributed games: From mechanisms to protocols. In *In Proc. 16th National Conference on Artificial Intelligence (AAAI-99*. Citeseer, 1999.

[8] Eytan Adar and Bernardo A Huberman. Free riding on gnutella. *First monday*, 5(10), 2000.

[9] Bernardo A Huberman and Rajan M Lukose. Social dilemmas and internet congestion. *Science*, 277(5325):535–537, 1997.

[10] Jeffrey Shneidman and David C. Parkes. *Rationality and Self-Interest in Peer to Peer Networks*, pages 139–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[11] Ittay Eyal and Emin Gün Sirer. Majority is not enough: Bitcoin mining is vulnerable. In *International conference on financial cryptography and data security*, pages 436–454. Springer, 2014.

[12] Kerem Kaskaloglu. Near zero bitcoin transaction fees cannot last forever. In *The International Conference on Digital Security and Forensics (DigitalSec2014)*, pages 91–99. The Society of Digital Information and Wireless Communication, 2014.

[13] Joshua A Kroll, Ian C Davey, and Edward W Felten. The economics of bitcoin mining, or bitcoin in the presence of adversaries. In *Proceedings of WEIS*, volume 2013, 2013.

[14] Emin Gün Sirer. Bitcoin runs on altruism, http://hackingdistributed.com/2015/12/22/bitcoin-runs-on-altruism/, 2015.

[15] Thomas J Housel, Wolfgang Baer, and Johnathan Mun. 2 a new theory of value. *Intellectual Capital in Organizations: Non-Financial Reports and Accounts*, 1:16, 2014.

[16] Yoad Lewenberg, Yoram Bachrach, Yonatan Sompolinsky, Aviv Zohar, and Jeffrey S. Rosenschein. Bitcoin mining pools: A cooperative game theoretic analysis. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 919–927, Richland, SC, 2015. International Foundation for Autonomous Agents and Multiagent Systems.

[17] Andrew Miller, Ahmed Kosba, Jonathan Katz, and Elaine Shi. Nonoutsourceable scratch-off puzzles to discourage bitcoin mining coalitions. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, pages 680–691, New York, NY, USA, 2015. ACM.

[18] Fabio A Drucker and Lisa K Fleischer. Simpler sybil-proof mechanisms for multi-level marketing. In *Proceedings of the 13th ACM conference on Electronic commerce*, pages 441–458. ACM, 2012.

[19] Yuval Emek, Ron Karidi, Moshe Tennenholtz, and Aviv Zohar. Mechanisms for multi-level marketing. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 209–218. ACM, 2011.

[20] Jon Kleinberg and Prabhakar Raghavan. Query incentive networks. In *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pages 132–141. IEEE, 2005.

[21] Cuihong Li, Bin Yu, and Katia Sycara. An incentive mechanism for message relaying in unstructured peer-to-peer systems. *Electronic Commerce Research and Applications*, 8(6):315–326, 2009.

[22] Moshe Babaioff, Shahar Dobzinski, Sigal Oren, and Aviv Zohar. On bitcoin and red balloons. In *Proceedings of the 13th ACM conference on electronic commerce*, pages 56–73. ACM, 2012.

[23] Ittai Abraham, Dahlia Malkhi, Kartik Nayak, Ling Ren, and Alexander Spiegelman. Solidus: An incentive-compatible cryptocurrency based on permissionless Byzantine consensus. *arXiv preprint arXiv:1612.02916*, 2016.

[24] Jeffrey Travers and Stanley Milgram. The small world problem. *Phychology Today*, 1:61–67, 1967.

[25] Iddo Bentov, Ariel Gabizon, and Alex Mizrahi. Cryptocurrencies without proof of work. In *International Conference on Financial Cryptography and Data Security*, pages 142–157. Springer, 2016.

[26] Iddo Bentov, Charles Lee, Alex Mizrahi, and Meni Rosenfeld. Proof of activity: Extending bitcoin's proof of work via proof of stake [extended abstract]. *ACM SIGMETRICS Performance Evaluation Review*, 42(3):34–37, 2014.

[27] Ittay Eyal, Adem Efe Gencer, Emin Gun Sirer, and Robbert Van Renesse. Bitcoin-ng: A scalable blockchain protocol. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 45–59. USENIX Association.

[28] Aggelos Kiayias, Alexander Russell, Bernardo David, and Roman Oliynykov. Ouroboros: A provably secure proof-of-stake blockchain protocol. Cryptology ePrint Archive, Report 2016/889, 2016.

[29] Ethan Heilman, Alison Kendler, Aviv Zohar, and Sharon Goldberg. Eclipse attacks on bitcoin's peer-to-peer network. In *USENIX Security Symposium*, pages 129–144, 2015.

[30] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. Hijacking bitcoin: Routing attacks on cryptocurrencies. In *IEEE Symposium on Security and Privacy*, pages 375–392. IEEE, 2017.

[31] Arthur Gervais, Ghassan Karame, Srdjan Capkun, and Vedran Capkun. Is bitcoin a decentralized currency? *IEEE Security & Privacy*, 12(3):54–60, 2014.

[32] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

[33] Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[34] Alex Biryukov, Dmitry Khovratovich, and Ivan Pustogarov. Deanonymisation of clients in bitcoin p2p network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 15–29. ACM, 2014.

[35] Giulia Fanti and Pramod Viswanath. Deanonymization in the bitcoin p2p network. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 1364–1373. Curran Associates, Inc., 2017.

[36] Philip Koshy, Diana Koshy, and Patrick McDaniel. *An Analysis of Anonymity in Bitcoin Using P2P Network Traffic*, pages 469–485. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[37] Shaileshh Bojja Venkatakrishnan, Giulia Fanti, and Pramod Viswanath. Dandelion: Redesigning the bitcoin network for anonymity. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 1(1):22, 2017.

[38] Till Neudecker, Philipp Andelfinger, and Hannes Hartenstein. Timing analysis for inferring the topology of the bitcoin peer-to-peer network. In *Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016 Intl IEEE Conferences*, pages 358–367. IEEE, 2016.

[39] Martí Berini Sarrias. Bitcoin network simulator data explotation. Master's thesis, Universitat Oberta de Catalunya, 2015.

[40] Tams Nepusz. IGraph: High performance graph data structures and algorithms, http://igraph.org/python/, 2006–.

[41] Agata Fronczak, Piotr Fronczak, and Janusz A Hołyst. Average path length in random networks. *Physical Review E*, 70(5):056110, 2004.