Bitcoin Network Measurements for Simulation Validation and Parameterisation

Muntadher Fadhil; Gareth Owen; Mo Adda
University of Portsmouth, Buckingham Building, Portsmouth, United Kingdom
{Muntadher.sallal; Gareth.owen; Mo.Adda}@port.ac.uk

***Abstract*— Bitcoin is gaining increasing popularity nowadays, even though the crypto-currencies field has plenty of digital currencies that have emerged before the adoption of Bitcoin idea. Bitcoin is a decentralized digital currency which relies on set of miners to maintain a distributed public ledger and peer-to-peer network to broadcast transactions. In this paper, we analyse how transaction validation is achieved by the transaction propagation round trip and how transaction dissemination throughout the network can lead to inconsistencies in the view of the current transactions ledger by different nodes. We then measure the transaction propagation delay in the real Bitcoin network and how it is affected by the number of nodes and network topology. This measurement enables a precise validation of any simulation model of the Bitcoin network. Large-scale measurements of the real Bitcoin network are performed in this paper. This will provide an opportunity to parameterise any model of the Bitcoin system accurately.**

Keywords—Bitcoin ; Propagation delay ;Simulation Validation

I. INTRODUCTION

Bitcoin is a decentralized peer-to-peer electronic currency that allows online payments between two parties without any form of central authority [1]. The system was proposed in 2008 by Satoshi Nakamoto and deployed as a payment system in January 2009 [2], [3]. Bitcoin relies on a cryptographic protocol that operates on top of the Bitcoin peer-to-peer network. The user's identity in Bitcoin is represented by a public key as opposed to their name or other identifiable information [4]. Furthermore, Bitcoin is also the name of the currency that this network enables where, one Bitcoin (BTC) has an equivalent value in British pounds (GBP).

Bitcoin is considered as a reliable currency which allows global transactions to be processed as fast as local ones. In addition, it offers a public history of all transactions that have ever been processed. It also introduces such new payment strategies, such as micropayment, contract, and escrow transactions.

Bitcoin follows a distributed trust mechanism which relies on distributed validation and tracking of transactions. Based on this mechanism, a Bitcoin transaction has to be broadcasted to all nodes within the network to reach a consensus about which transactions are valid. The consensus is recorded in a publicly distributed ledger which is shared by the entire network.

As Transactions are validated against the public ledger, inconsistency in the replicas of ledger is unavoidable. This introduces uncertainty about the validity of a given transaction which may lead to an attacker being able to spend a Bitcoin twice.

In this work, we present measurements of the transaction propagation delay as well as measurements of the real Bitcoin network. These measurements are important to validate and parameterise any simulation model of Bitcoin network. We further analyse transaction validation in the Bitcoin network and how the consistency of the public ledger is affected by transaction propagation.

The paper is organised as follows: Section II focuses on giving an overview of the Bitcoin system and briefly describing the Bitcoin networking aspects. In Section III, we discuss in details the information propagation in the Bitcoin network and analyse the double spending attack which is caused by the transaction propagation delay. In addition, related work in measuring and analysing Bitcoin information propagation and in modelling approaches to avoid double spending attacks will be outlined. In Section IV, measurements of the transaction propagation delay as well as measurements of the real Bitcoin network parameters will be presented. In Section V, we conclude the paper and discuss the future work.

II. BACKGROUND

In this section we provide a general overview of the Bitcoin system. We focus on the Bitcoin protocol by discussing the basic operation of the Bitcoin network and how the globally consistent state is achieved. We then give a brief description of the relevant aspects of Bitcoin which are block chain and the network structure.

A. The Bitcoin protocol

The Bitcoin protocol is built on the basis of creation and distribution of public record of all the Bitcoins in the system. This record considers each entry as a transaction by which the transfer of virtual currency is accomplished. Each transaction consists of inputs and outputs. A transaction's output which indicates the new owner of the transferred Bitcoins, will be referenced as inputs in future transactions to create new

output/outputs [5]. Transactions are formed as a directed graph which helps when giving constants about transaction record.

Each transaction input should have a digital signature that unlocks the previous transactions' output. This signature is created only by the user who possesses an appropriate private key. This ensures that Bitcoins can only be spent by their owners. In addition, the sum of the values of the inputs should be equal to or greater than the sum of all outputs.

B. Block Chain

Block chain is simply the ledger of all transactions, grouped into blocks. Every block is linked with previous blocks by including the unique hash of the previous block in its header. The first block in the block chain is known as the genesis block and it has no references to previous blocks. A branch is a path in the block chain which starts from a leaf block to the genesis block [6]. Block chain technology is deemed as the most important invention in the field of cryptography and security of decentralised networks, because it allows an immutable record of all transactions to be created, such that it is resistant to modification from the most resourceful attackers. Block chain is publicly visible and allows nodes within the network to agree to be confident about the transfer of money between users [7],[8]. Any valid transactions disseminated in the Bitcoin network are collected in a block by miners. After that, this block requires a degree of computational effort before it will be accepted by other nodes as valid. A group of nodes known as miners provide this effort when they solve the computational problem, and for which they are rewarded with a small number of Bitcoins. The solution to the problem is easy to verify but difficult to calculate and as such the solution can be considered a proof of work (POW) [6]. Blocks are chained together, and thus, modifying a block becomes exponentially harder with the passage of time, as all subsequent blocks must also modified[9].

C. Bitcoin peer-to-peer network

In the Bitcoin network, as shown in Fig.1, each peer connects randomly with other peers over a TCP channel [10]. Each node maintains a list of IPs of peers that the node established connections with. For the purpose of making denial of service impractical, just the valid transactions and blocks are propagated, whereas invalid transactions and blocks are discarded. Furthermore, Bitcoin network achieves a reputation protocol by which each node maintains a penalty score for every connection. Once a node receives a corrupted message from a particular connection of its connections, it increases the penalty score of the connection and bans the misbehaving IP when the score reaches the value of 100. Bitcoin network nodes are classified into two groups. Servers which can accept incoming connections and those which can't (clients), because

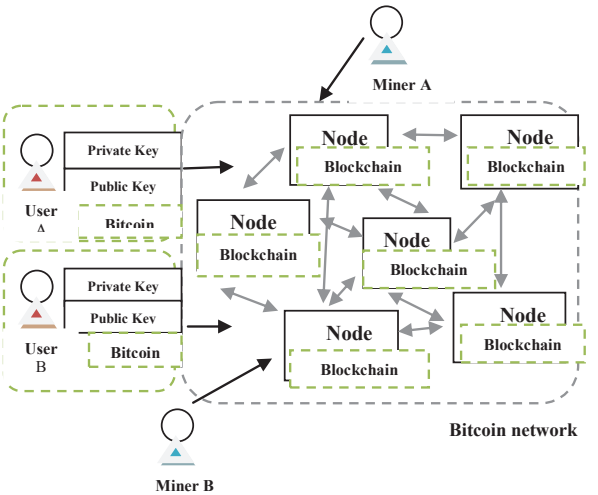

Figure 1: The structure of the Bitcoin network

they are behind NAT or firewall. Peers in the Bitcoin network maintains up to 8 outgoing connections and accept up to 117 incoming connections. Bitcoin peer stays connected to the 8 outgoing connections until it is restarted, whereas connections will be replaced if any of the outgoing connections drop [10].

III. INFORMATION PROPAGATION AND RELATED WORKS

There are two types of information that are propagated in the Bitcoin network: Transactions and Blocks. Transactions are responsible for transferring values, whereas blocks are used to ensure a chronological ordering of transactions across all nodes in the Bitcoin network and also form part of the ledger [11]. To broadcast a transaction, a user simply connects to a number of peers within the network and sends it to them. Each peer maintains history of forwarded transactions for each connection and if it has not seen that transaction before then it will rebroadcast it to all of its peers [12]. In the following, we discuss how the Bitcoin transaction propagation affects the synchronization of the public ledger and the role inconsistency of the public ledger plays in making double spending attack achievable. Finally, we close this section by highlighting related work on measuring, analysing, and speeding up Bitcoins transaction propagation.

A. Transaction propagation

Bitcoin uses a gossip-like protocol to broadcast information throughout the network [7]. Therefore, transactions are not forwarded directly in order to avoid sending a transaction to a node that already received it from other nodes. Instead, a node announces to its neighbour nodes about the transaction availability once the transaction has been verified. As shown in Fig.2, transactions are disseminated through the network using a protocol, which includes propagating two types of messages, an INV message and a GETDATA message.

When a node receives a transaction from one of its neighbours, it sends an INV message containing the hash of the transaction to all of its peers. When a node receives an INV, it checks whether the hash of a transaction has been seen before. If it has not been seen before, the node will request the transaction by sending a GETDATA message. In terms of receiving a GETDATA message, a node responds by sending the transaction's data.

An INV message is not propagated to all of the connected peers at the same time, instead, every 100ms it is sent to a random selected peer of all connected peers. Therefore, the required time for forwarding the INV message relies on the number of connected nodes [13]. Due to the above broadcasting scenario, a delay in transaction propagation happens, and this delay combines between time which takes to validate the transaction and propagate it. Essentially, propagation delay pertains to many issues in the Bitcoin due to the inconsistency of the public ledger which comes up with the opportunity for an attacker to abuse the network consensus. Specifically, inconsistency in the public ledger will induce the dishonest nodes to disturb the confirmation operation of a valid transaction by broadcasting a conflicting transaction with the same amount of coin during the period of confirmation, in which the valid transaction waits to be added to the block chain. This type of attack is called double spending attack in which the attacker attempts to spend the same transaction output more than once.

Double spending attacks happen when an attacker creates two transactions (T_A and T_M) with the same input (same source of Bitcoin) and different outputs (different recipients, suppose we have two transactions, T_A will go to the majority of peers and T_M will go to the vendor). We can consider the double spending attack as successful when T_A is confirmed before T_M. This means the majority of peers accept T_A while the vendor accepts just T_M. This will lead to the acceptability of T_A by subsequent blocks as an original transaction and the vender can not redeem the T_M because it is considered as an invalid transaction because it is trying to spend money which has already been spent [14].

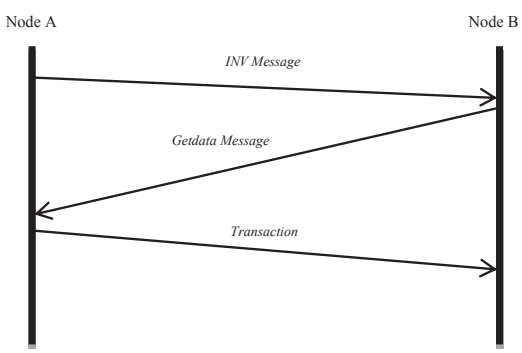

Figure 2: Transaction propagation protocol between Nodes A and B

The second scenario of double spending is when an attacker secretly mines a branch which includes, the transaction that returns the payment to himself, while disseminating the merchant's transaction [15]. The attacker will not broadcast this branch until the merchant's transaction gets confirmed. In this scenario, the merchant is going to be confident about the transaction and then he will consider delivering the product. Furthermore, the attacker has to be sure that the secret branch is longer than the public branch, so if necessary, continue extending the secret branch. Finally, an attacker broadcasts the secret branch when he confirms that the secret branch is longer than the public one. Typically, this is computationally expensive, requiring the attacker to control 50% of the computing available in the network.

B. Related works

Most previous analytical studies of the Bitcoin network have presented measurements of the network that are linked to the information propagation delay. Recent [11] research has shown that the number of nodes in the Bitcoin network and the structure of overlay network have a great impact on transaction propagation time. Their results showed that the transaction propagation time is improved by reducing the number of nodes from 6000 to 2000. Furthermore, their results demonstrated that the overlay topology of the Bitcoin network, which is not geographically localised, offers inefficient transaction propagation time. Transaction propagation delay in real Bitcoin network has been measured in [13],[16] by developing a Bitcoin client that tracks how transactions are disseminated through the network by listening for INV messages.However, previous propagation delay measurements do not represent the real propagation delay as it does not indicate the exact time by which peers announce transactions.

The probability of double spending attack in fast payments, have been measured in the previous research through analytical models, based on measurements in real Bitcoin network [12]. In terms of avoiding double spending attack, [17] introduced some counter-measures to avoid double spending attacks and proposed a prototype system, which is applied in vending machines. The main idea of this system is to set a server that will observe the transaction. When transaction propagation reaches over 40 nodes, the server will give a signal, which means that the transaction has been confirmed. Unfortunately, this solution is limited because the attacker's transaction could still be propagated to the majority of nodes.

In [18] a new protocol has been proposed which tackles the problem of inconsistency in the public ledger by reducing the information propagation time. This solution claims that the information propagation could be pipelined instead of waiting to receive the transaction. In other words, any node can immediately forward an invitation message (INV Message) that includes a list of hashes of available transactions, rather than waiting for receiving transactions. Another change has been proposed in the same theory. This change increases the geographical connectivity in Bitcoin network in order to offer

faster information propagation. However, this theory reduces the propagation delay with a very low rate because the transaction still needs to visit the all nodes in the Bitcoin network. Additionally, the transaction verification time still remains inefficient due to the size of the public ledger.

A model for faster transaction propagation has been presented in [19] by considering some modifications in the transaction dissemination protocol. The core idea of this model is that when nodes receive a transaction, they check whether this transaction has been seen before in their pool. In case the transaction has not been seen before, they add the transaction to their pool and forward it to the other nodes. Otherwise, they directly forward the transaction to other neighbours without adding it to their pool. This scenario allows the fake transaction to be received by the node that issues the original transaction.

IV. BITCOIN PARAMETERS MEASUREMENT :

Large scale parameters of real Bitcoin network are difficult to predict, therefore, some certainty is required before any simulation model of Bitcoin system would be implemented. To this purpose, we provide accurate measurements of different parameters of the real Bitcoin network. In the following subsections, the measurement of transaction propagation as well as large scale measurements of real Bitcoin network will be presented. Propagation delay is very important for the validation of simulation measurements as many aspects of the Bitcoin network such as network topology, clients' behaviour, and processing delay affect it. By offering these measurments, validation of any Bitcoin simulation model would be possible through comparing the propagation delay measurements that will be collected from the Bitcoin simulator to the same measurements that have been collected in this experiment. Also the real Bitcoin network measurements are essential to parametrize any model of Bitcoin.

A. TRANSACTION PROPAGATION MEASUREMENT

In this section, we investigate how fast a transaction propagates in the Bitcoin network and how this is impacted by the number of nodes. A transaction propagation delay was measured in the prior research by setup a Bitcoin client which keeps listening for INV messages. However, we present a novel methodology by which the transaction propagation delay is accurately measured as the measurements are indicated when peers receive transactions. Experiment methodology and results are presented in the following subsections.

1) Experiment description and methodology

To measure the propagation delay, the Bitcoin protocol was implemented and used to establish connections to many points in the network, in order to measure the time that a transaction takes to reach each point. Specifically, we first implemented a measuring node, which behaves exactly like a normal node with the following functionalities. The measuring node connects to 14 reachable peers in the network. Furthermore, it

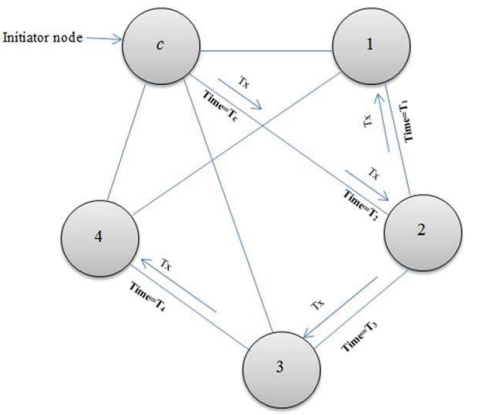

Figure 3: Illustration of propagation experimental setup

is able to create a valid transaction and send it to one peer of its connections, and then it tracks the transaction in order to record the time by which each peer of its connections announces the transaction. To measure how fast a transaction is exchanged between the connected nodes, we calculated the time by which the transaction is propagated by our measuring node and reached each node of our measuring nodes connections. Specifically, suppose a client c has connections (1,2 ,3,…., n), c propagates a transaction at time T_c , and it is received by its connected nodes at different times ($T_1, T_2, T_{3,\ldots,} T_n$) as illustrated in Fig.3. The time differences between the first transaction propagation and subsequent receptions of the transaction by connected nodes were calculated ($\Delta t_{c,1}$,…., $\Delta t_{c,n}$) according to Eq.1:

$$\Delta t_{c,n} = T_n - T_c \qquad (1)$$

In order to get accurate measurements, the timing information was collected by running the experiment 1000 times as errors such as loss of connection and data corruption, are expected to happen in case of dealing with real network. At each run, the measuring node is randomly connected to 14 nodes. The timing information contains the hash of the transaction, the announcing nodes IP, transaction confirmation time and a local time stamp, which represents propagation time once the transaction was received.

2) Results

The propagation measurements from this experiment are shown in Fig.4. The number of connected nodes represents the sequence of the random nodes that the measuring node connects with at each run. Fig 4 indicates that during the first 13 seconds, transaction has been propagated faster and 6 nodes received it with low variance of delays. It should be noted that the transaction propagation delays is dramatically increased over nodes (9,10,14) which means that the transaction has been received by these nodes with significantly larger variances of delays. Obviously, these results reveal that

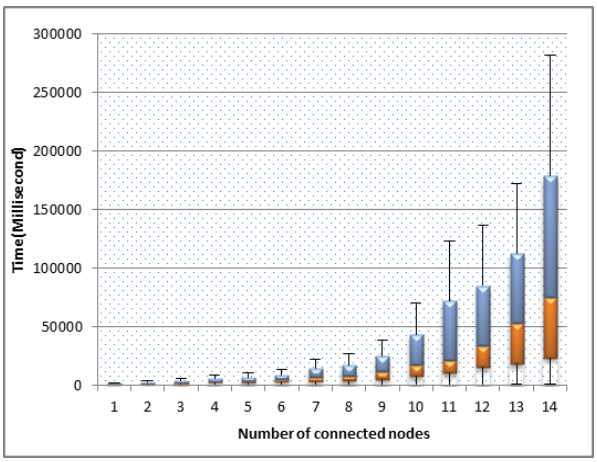

Figure 4: Transaction propagation time in real Bitcoin network

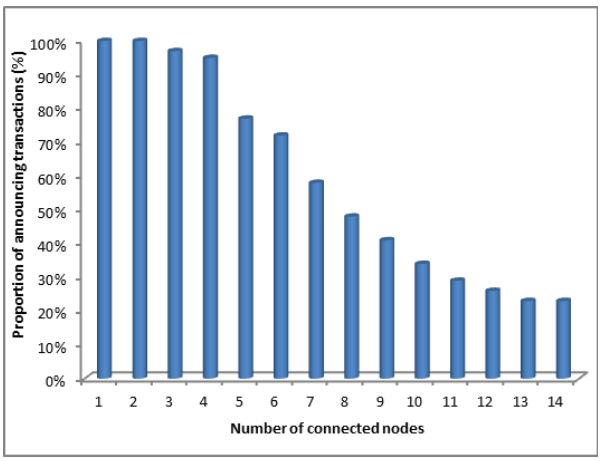

Figure 5: Proportion of each node announced the transaction

the propagation delay negatively corresponds with the number of nodes, as the total duration of subsequent announcements of the transaction by the remaining nodes increases with larger numbers of connected nodes. This happened due to each node being connected to large segments of the network, while the connected nodes were not geographically localized. On the other hand, transaction verification at each node affects trickling transaction to the remaining nodes. However, there are possible ways that can improve information propagation in the Bitcoin network which in turn would reduce the probability of double spending attacks. Reducing the noncompulsory hops in that the transaction passes through in conjunction with increasing the locality of connectivity are considered as a possible scenario that would achieve significantly faster information propagation. This can be achieved by applying a clustering theory by which the Bitcoin network nodes are fully partitioned into clusters depend on its geographical location. To evaluate any clustering theory based on improving information propagation, major changes are required to the Bitcoin protocol which would have to be accepted by the Bitcoin community. Therefore, Bitcoin model which behaves as close as real Bitcoin network is required. Both clustering theory and Bitcoin model are considered as our future work.

Surprisingly, we noticed that not all of the connected nodes received the transaction except rare cases in which all the 14 connected nodes announced the transaction. Fig.6 shows proportions of announcing transactions for each node. Each proportion was calculated over 1000 runs. Nodes 1, 2, 3 and 4 are almost announced transactions within proportions between 90-100. The proportion dramatically declined at node 5 and continued to go down to reach 23 at node 14. This pointed to the issue caused by network partitions in which the network is divided into two or more partitions due to network outages or link failure, so that no information flow between partitions is possible. Network partitions are more likely to happen within the network topology, which is not geographically localized. However, network partitions can be done by an attacker to impair main Bitcoin functions. We leave a further analysis of this issue as future work.

B. Bitcoin network measurements

In this section, we present the measurements of two Bitcoin network parameters which are number of the reachable nodes and link latencies between peers. These parameters are considered as the most influential parameters in the Bitcoin network due to their direct impact on the information propagation delay in the network. Therefore, these parameters are important to parameterise any model of Bitcoin accurately. For this purpose, Bitcoin client was implemented and used to crawl the entire Bitcoin network through establishing connections to all reachable peers in the Bitcoin network. Every five minutes the snapshot of IP addresses of all reachable peers was published by the developed crawler. We discovered that the crawler learned 313676 IP addresses but was only able to connect to 5378 peers. This indicates that the Bitcoin network size is presently around 5400 nodes.

Fig 6 shows the distribution of latencies in the real Bitcoin network. The crawler was connected to around 5000 network peers and observing a total of 20,000 ping/pong messages. It should be noted that the measured distribution only represents the latency between our crawler and other peers in the network.

As these measurements have a direct impact on the information propagation time, it is necessitated to perform these measurements when any model of Bitcoin is built. Though, attaching the measured distribution to the model would give an accurate estimate of the time delay that is taken by a transaction to reach different peers in the network.

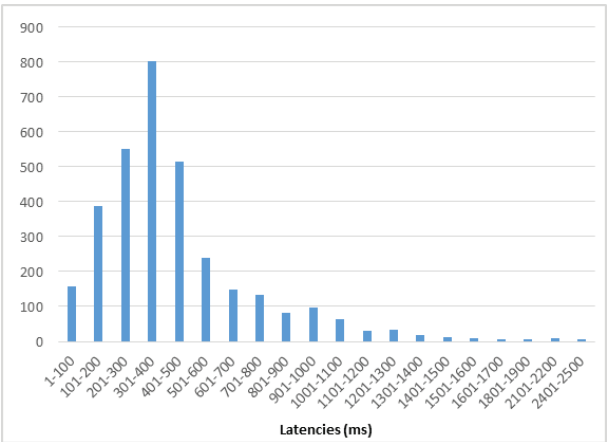

Figure 6: Latencies distribution between the measurement node and other peers

V. Conclusion

A brief background of Bitcoin system and block chain technology was presented in this paper. In addition, we analysed the information propagation in the real Bitcoin network. We have also discussed how propagation delay could affect the security by offering an opportunity to double spend the same coins, thereby abusing the consistency of the public ledger. Furthermore, previous studies to analyse and measure the information propagation delay were explained briefly.

In order to offer an opportunity to validate and parameterise any model of Bitcoin network, different kinds of measurements have been presented in this paper. We implemented a novel methodology to measure the transaction propagation delay in real Bitcoin network. Our measurements show that the transaction propagation time is significantly affected by the number of the connected nodes and the network topology which is not geographically localised. In addition, partitions in the connection graph are actively detected. Finally, the size of the Bitcoin network and distribution of latencies between nodes are accurately measured in this paper.

Future work:

The future work will be to examine clustering as a mechanism to improve the propagation delay. Our proposed approach claims that the fully Bitcoin network nodes could be partitioned into clusters depend on geographical location. Each node will be included in a cluster that correlates to its geographical location. We claim that giving rise to the locality of connectivity can affect the information dissemination by reducing the noncompulsory hops that the transaction passes through, so that, on other hand, this could minimize the propagation delay.

VI. References

[1] Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system. Retrieved from http://www.bitcoin.org/bitcoin.pdf.

[2] Ron, D., & Shamir, A. (2013). Quantitative analysis of the full bitcoin transaction graph. In Financia Cryptography and Data Security (pp. 6–24). Springer.

[3] Moore, T., & Christin, N. (2013). Beware the middleman: Empirical analysis of Bitcoin-exchange risk. In Financial Cryptography and Data Security (Vol. 7859, pp. 25–33). Springer.

[4] Biryukov, A., Khovratovich, D., & Pustogarov, I. (2014). Deanonymisation of clients in Bitcoin P2P network. arXiv Preprint arXiv:1405.7418.

[5] Heilman, E., Kendler, A., Zohar, A., & Goldberg, S. (2015). Eclipse Attacks on Bitcoin's Peer-to-Peer Network. In: 24th USENIX Security Symposium(USENIX Security 15), Washington, D.C., USENIX Association.

[6] Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T., & Capkun, S. (2013). Evaluating user privacy in bitcoin. In Financial Cryptography and Data Security (pp. 34–51). Springer.

[7] Ober, M., Katzenbeisser, S., & Hamacher, K. (2013). Structure and Anonymity of the Bitcoin Transaction Graph. Future Internet, 5(2), 237–250. doi:10.3390/fi5020237.

[8] Koshy, P., Koshy, D., & Mcdaniel, P. (2014). An analysis of anonymity in bitcoin using p2p network traffic. In Proceedings of FinancialCryptography and Data Security (FC'14). Springer.

[9] Barber, S., Boyen, X., Shi, E., & Uzun, E. (2012). Bitter to better—how to make bitcoin a better currency. In Financial cryptography and data security (pp. 399–414). Springer.

[10] Biryukov, A., Khovratovich, D., & Pustogarov, I. (2014). Deanonymisation of clients in Bitcoin P2P network. arXiv Preprint arXiv:1405.7418.

[11] Miller, A., & Jansen, R. Shadow-Bitcoin: Scalable simulation via direct execution of multithreaded applications. IACR Cryptology ePrint Archive 2015 (2015), 469.

[12] Karame, G. O., Androulaki, E., & Capkun, S. (2012). Two Bitcoins at the Price of One? Double- Spending Attacks on Fast Payments in Bitcoin. IACR Cryptology ePrint Archive, 2012, 248.

[13] Neudecker, T., Andelfinger, P., & Hartenstein, H. (2015). A simulation model for analysis of attacks on the Bitcoin peer-to-peer network. In Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on (pp. 1327-1332). IEEE.

[14] Karame, G. O., Androulaki, E., & Capkun, S. (2012). Two Bitcoins at the Price of One? Double- Spending Attacks on Fast Payments in Bitcoin. IACR Cryptology ePrint Archive, 2012, 248.

[15] Rosenfeld, M. (2014). Analysis of Hashrate-Based Double Spending, 13. Cryptography and Security. Retrieved from http://arxiv.org/abs/1402.2009.

[16] Decker, C., & Wattenhofer, R. (2013, September). Information propagation in the bitcoin network. In Peer-to-Peer Computing (P2P), 2013 IEEE Thirteenth International Conference on (pp. 1-10). IEEE.

[17] Bamert, T., Decker, C., Elsen, L., Wattenhofer, R., & Welten, S. (2013). Have a snack, pay with Bitcoins. IEEE P2P 2013 Proceedings, 1–5. doi:10.1109/P2P.2013.6688717.

[18] Stathakopoulou, C.(2015).A faster Bitcoin network. Tech. rep., ETH, Zurich,. SemesterThesis, supervised by Decker.C and Wattenhofer.R.

[19] Karame, G. O., Androulaki, E., Roeschlin, M., Gervais, A., & Èapkun, S. (2015). Misbehavior in Bitcoin: A Study of Double-Spending and Accountability. ACM Transactions on Information and System Security (TISSEC), 18(1), 2.