Review of Privacy Enhancing Techniques in Bitcoin

Prerna Kalla
Indraprastha Institute of
Information Technology
Okhla Phase 3, Delhi, India
Email: prerna17042@iiitd.ac.in

***Abstract*—Bitcoin is a decentralized blockchain based cryptocurrency that has taken the world by storm. Since its introduction in 2009, it has grown tremendously in terms of popularity and market cap. The idea of having a decentralized public ledger while maintaining anonymity and security attracted the attention of developers and customers alike. Special nodes in the bitcoin network, called Miners are responsible for making the network secure by using a concept called Proof-of-Work. A certain degree of anonymity is also maintained as no personally identifiable information of a person like a name, address, aadhaar number etc is linked to the bitcoin wallet. In terms of bitcoin, a user is anonymous if different interactions of the user cannot be linked to each other or the user. Recent research shows that bitcoin is not that anonymous as it appears to be. The inherently public nature of blockchain technology makes it difficult to achieve privacy. The purpose of this paper is to review how varying degrees of user privacy is maintained in bitcoin cryptocurrency. This paper is divided into two main segments. The first segment explores privacy enhancing techniques adopted in bitcoin. The second segment critically analyzes these techniques.**

***Index Terms*—Anonymity, ToR, Bloom Filters, Altcoins, Zero-knowledge proofs, Mixing service, Zero-cash, Coinjoin, Mining, Blockchain**

I. Introduction

Bitcoin was introduced by Satoshi Nakamoto in 2009 as a decentralized cryptocurrency [1]. It is decentralized as no single organization owns bitcoin. All peers on bitcoin network are equal and have the same rights. A bitcoin user needs to install a bitcoin wallet which acts as an interface between bitcoin network and the user. The wallet generates a pair of the private and public key for the user. Public keys are used to send bitcoin to any other node and private address is used to redeem bitcoins. The bitcoin peer to peer network uses a public ledger called blockchain to store all its transactions [2]. All nodes in the network listen for valid transactions. A given set of valid transactions is collected into a block by the node and added into the blockchain through a process called mining [3]. Such nodes are called bitcoin miners. Miners are incentivized for adding a valid block in the blockchain. No transaction can be reversed once it is entered into the blockchain. Mining serves two purposes. First, it provides a technique to add new blocks to the blockchain. Second, it adds new coins in the network. Every miner is awarded some bitcoins for successfully adding a block. Currently, the reward is 12.5 BTC per block. This ensures that the bitcoin network is secure and stable. Mining requires solving a hard computation problem by giving proof-of-work.Proof of work means producing data with high level of difficulty which matches certain conditions. It is possible that more than 1 block is mined at a time or an attacker proposes a malicious block. Other nodes of the network will validate all the transactions of the block.

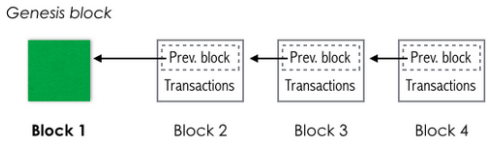

Fig. 1. Blocks in blockchain containing transactions [2]

If there is something wrong, the nodes will reject the block and not mine any new node on top of it. Nodes give their confirmation to valid block. A block with 6 confirmations is usually considered as valid. Bitcoin follows the policy of mining of the top of the longest blockchain. The longest chain will be the one which has got more confirmations and has the support of the network nodes [4]. To tamper with the blockchain, an attacker has to create a number of blocks than present in the blockchain and mine all of them on top of each other. Mining a single block is quite hard. It is not possible for an attacker to mine the entire blockchain alone. No individual has such computing power. Thus, mining keeps the bitcoin network secure and also allows the Bitcoin nodes to reach consensus. It is essential that all nodes in the network arrive at a consensus regarding the blockchain without trusting each other.

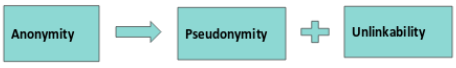

Fig. 2. Anonymity in bitcoin

Another major aspect which attracted people towards bitcoin was its claim of anonymity [1]. Bitcoin is pseudonymous. Pseudonymous means hiding behind an alias which in this case is the public key of the user. Some level of privacy is maintained as no personally identifiable information of a

person is linked to the persons bitcoin wallet. But if you know a bitcoin address, you can trace all transactions made by that address as the ledger is public. Hence it has become a common practice to use a different address for each transaction. Yet, there are other ways through which user privacy can be compromised. It is important to add unlinkability to bitcoin to achieve the goal of full autonomy. A way to keep yourself completely anonymous is to download the entire blockchain (around 156 GB in size as of now) and generate a new address for each transaction [5]. The problem with this approach is that most bitcoin users install wallets on their smartphones which do not have sufficient memory to store the chain. Bitcoin users are generally divided into two categories - full nodes and lightweight nodes. Full nodes store the entire blockchain and validate all the transactions in the block. Lightweight clients store only block headers. They are interested in only those transactions which concern them. Lightweight nodes can request full nodes to send them transactions pertaining to them [6]. But this is a major privacy breach as lightweight nodes have to reveal their addresses to full weight nodes. A technique called bloom filter is used to obtain required transactions without revealing the actual address [7]. This is done by sending a pattern of the required address instead of the actual address. However, this does not provide complete anonymity as a third party monitoring network traffic can still figure out public addresses. Further developments led to the introduction of new techniques like ToR encryption and Peer-to-peer encryption and authentication. Section 3 discusses threats to privacy in bitcoin. Section 4 explores privacy enhancing techniques in detail and Section 5 critically analyzes them.

II. Review of Literature

The literature review is broad enough to cover relevant work done in this field. All cited sources are pertinent to the topic of study. All of the cited sources have been published in the last 5 years, hence it covers recent research done in this field. There is no evidence of any bias.

III. Threats to Anonymity and Privacy in Bitcoin

Bitcoin addresses are public key hashes rather than real identities [5]. No personally identifiable information of a person is linked to it. This gives a certain level of anonymity to the bitcoin user. The goal of anonymity for bitcoin is that it should not be possible to link any user with his/her addresses or transactions. Bitcoin satisfies this goal partially, hence it is pseudonymous. There exist several tools to link addresses of a person [8]. By knowing the public address of a person, it is possible to track all the transactions done by the person to date. Figure 3 shows how online websites like https://blockchain.info/address/ allows anyone to trace other peoples transaction history [9].

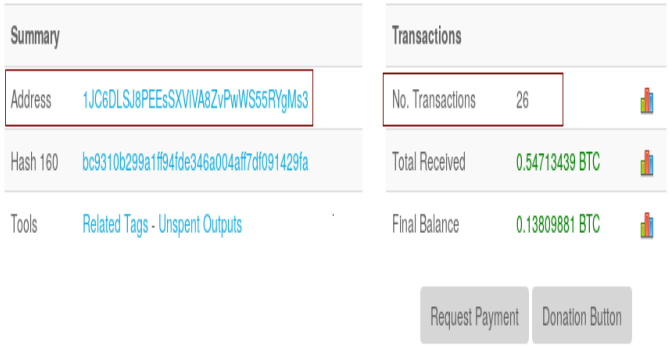

Fig. 3. Tracing transaction history through online tools for a bitcoin address [9]

In this section, I will discuss categories of threats to privacy and anonymity for a bitcoin user.

A. Peer to Peer Deanonymization

A trivial solution to keep your transactions unlinkable to each other is by generating a new address for each and every transaction [1]. However, this technique fails as there are several ways in which multiple addresses of a person can link together. Suppose, Alice has 5 BTC with one public address and 4 BTC with another public address. She wants to send 8 BTC to Bob. So she needs to merge her money from both accounts to pay 8 BTC to Bob. 1 BTC will be returned as the change to her third bitcoin address. Since bob gets payment from 2 separate addresses and he sees the change address, he can link all the three addresses to Alice. If Bob follows her transaction history on the blockchain by tracing her public addresses, he can discover her other addresses and Alices privacy is breached. Similarly, other linking techniques exist to deanonymize user. Such techniques can often produce false positives and lose accuracy with time. However, they can work well if planned carefully by the adversary.

B. Network Layer Deanonymization

Lightweight nodes request transactions of their interest from full nodes. They do so by sending a list of their addresses, transaction etc. This is a major breach of privacy as lightweight nodes reveal their public addresses to full nodes. A solution for this is to use bloom filters, which sends the pattern of the address instead of the actual address [10]. Bloom filters give false positives and do not provide complete anonymity. Another issue is that while transmitting transactions, their IPs addresses are also attached to it. Hence network level anonymization is needed. If not maintained, it poses a threat to user anonymity and privacy.

Estimated p_{addr}	Deanonymization rate with 3-tuples	
	Actual	Predicted
0.64	41%	43%
0.86	59.9%	65.6%

Fig. 4. Deanonymization rate of addresses [8]

IV. PRIVACY ENHANCING TECHNIQUES IN BITCOIN

This section describes the common techniques used in bitcoin to enhance user privacy.

A. Bloom Filters

Bloom filter is a probabilistic search filter which provides a way to provide a pattern without specifying it directly [7]. This is done by using hash functions. The required patterns are passed through hash function whose output is used in a bit array. These filters allow a simple payment verification (SPV) node to specify a pattern corresponding to its addresses and transactions which are hashed in the filter. This filter is sent to any full node which stores the entire blockchain. The full node will filter out transactions of interest according to the specified patterns and return it to the SPV. This hides the identity of the SPV as no address or transaction detail is sent directly and provides a certain level of anonymity. However, the more specific the pattern is, the lesser is the privacy. There is a tradeoff between privacy and precision.

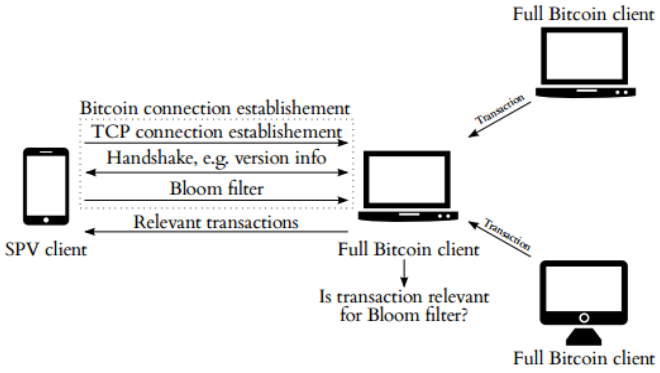

Fig. 5. How bloom filters work [7]

B. Peer to Peer Mixing

A proposal to increase anonymity in bitcoin is peer to peer mixing [1]. A usual bitcoin transaction contains some input values which are debited from signature providers and some output values which are credited to the corresponding recipient addresses. The idea of peer to peer mixing is to group a set of bitcoin users who want to do some transaction and mix them so that it is not known which transaction output corresponds to which input. Anonymity can be increased by increasing the number of participants in the mixing set. Once all peers in the mix are found, the transaction is constructed and send around to all peers to collect signatures. All peers will check if the output desired by them is present in the transaction or not. If yes, the peer can sign the transaction can forward it to other peers [11]. If no, then the peer can refuse to sign the transaction and a new peer-to-peer will have to be created. This protects peers from the theft of money. Once the transaction is signed by all the peers, it will be broadcasted on the network where it will be hopefully mined. One important property to note is that all individual transactions of all the peers in a mix should have the same bitcoin transfer value. This technique can help in increasing privacy significantly [12].

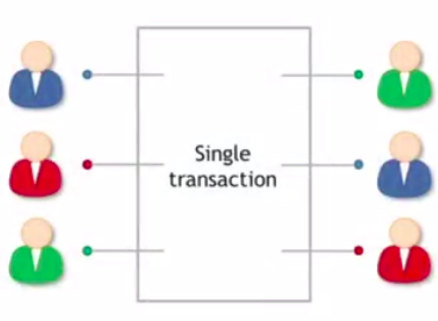

Fig. 6. Peer to peer mixing [13]

C. Third party mixing services

To protect anonymity, an intermediary is used which mixes all transactions of same values for different users to unlink their identity from their actual address [1]. These values are decided on the basis of standard chunk sizes predefined by the mixing service. Certain practices are followed to maximize anonymity like using a series of mixes and automating client-side software [13]. This technique allows anonymity from external users as well as internal users participating in the mix. To stay in business, third-party mixes need to earn the networks trust. Cryptographic warranties are provided to the user to assure them of their privacy and security [14].

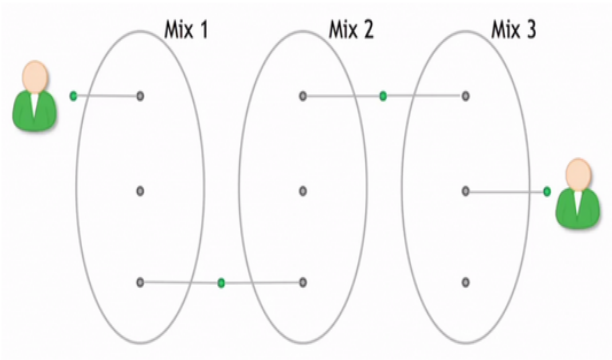

Fig. 7. Third party mixing services [12]

D. Altcoins with built-in privacy

Instead of having a third party to mix transactions, it can be directly incorporated into the protocol [1]. It eliminates the need to trust any third party and provides a cryptographic guarantee of mixing. Altcoin like zerocoin uses zero-knowledge proofs to make a statement, without revealing any other information. Every user generates a public serial number and a private secret number. The hash of this pair is posted on the blockchain as a cryptographic commitment. An arbitrary zerocoin in the blockchain of the required value can be chosen as input to the new transaction [15]. This provides full anonymity. No one can figure out which serial

number belongs to which zerocoin user. Similarly, zero-cash was developed as a more efficient implementation of zerocoin with better cryptographic proofs to make it more secure [16].

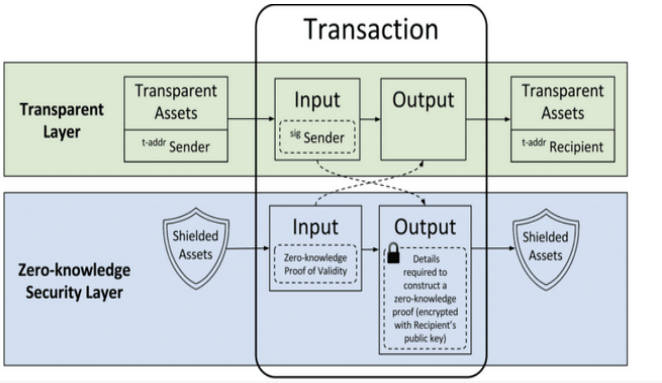

Fig. 8. How zero-knowledge proofs work [15]

E. ToR

In the original implementation of bitcoin, all nodes communicate openly throughout the network. All information was broadcasted in the form of plaintext. This vulnerability can be exploited to perform network deanonymization. A simple solution to this problem is to transmit encrypted data over the network. This is where ToR comes in. It stands for The Onion Router [17]. It is an open source software used to communicate anonymously over the network Bitcoin users are encouraged to use Tor along with bitcoin core software to increase user anonymity. Original bitcoin core software was later modified to offer support for ToR. This is mainly done so that privacy of the SPVs is not compromised. There is one entry node, one exit node, and at least one relay node in between [10]. Any data to be communicated is encrypted at the sender side and is sent to the receiver through random paths. ToR provides anonymity service that bitcoin currently lacks.

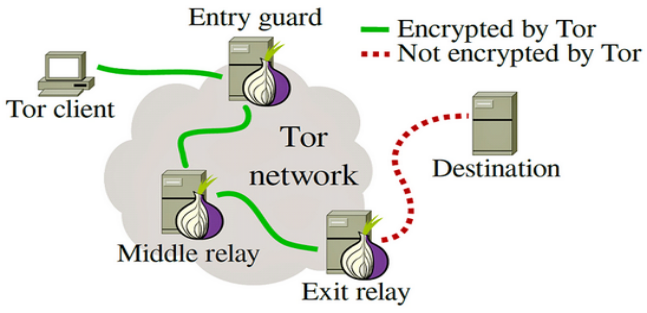

Fig. 9. How ToR works [17]

V. ANALYSIS OF PRIVACY ENHANCING TECHNIQUES

This section describes the pros and cons of the above-discussed privacy techniques.

A. Bloom Filters

Bloom filters are easy to use for an SPV client for whom anonymity is not of utmost importance. They give a probabilistic answer as they are based on patterns which can also be inaccurate. According to [7], using bloom filters leads to leakage of users private information. A user who uses many addresses (¿20) faces the risk of revealing some or all of its addresses. A malicious attacker who is observing network for some time can leak users address. Moreover, the precision of bloom filters decreases when a number of patterns are added to it. It is also known to give false positives. Additional verification has to be performed by the SPV to eliminate false positives. Figure 10 shows the number of false positives increases as the number of wallet addresses increases. Figure 11 shows that the number of true positives decreases as the number of wallet addresses increases. Bloom filters work as a temporary solution but they will fail in the long run to provide full anonymity to the user.

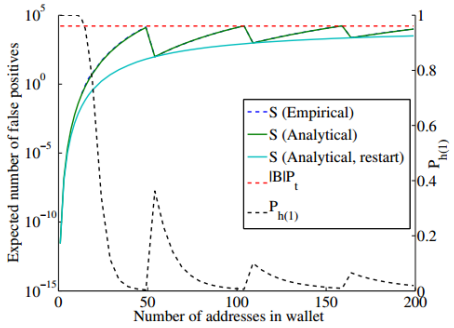

Fig. 10. False positives using bloom filters [7]

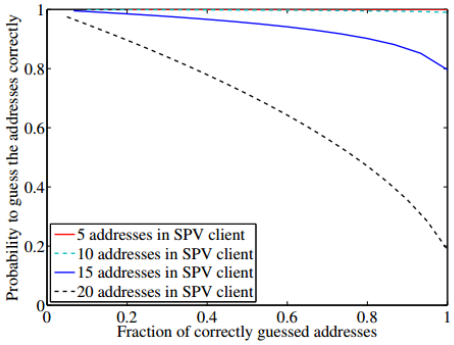

Fig. 11. True positives using bloom filters [7]

B. Peer to Peer Mixing

Peer to peer mixing allows peers the right to refuse to sign the transaction to reduce money laundering. This also makes it vulnerable to Denial Of Service attacks [1]. Another issue is that all peers in the mixing set can see individual inputs and outputs within the transaction. Hence, such mixing only provides external anonymity. There is no anonymity within the set. This defeats the purpose of mixing as if a peer within

the mix is malicious, he can leak users private information for whatever reasons. The additional restriction imposed is that the transfer value of bitcoin has to be the same in the mix. A valid node may have to wait for a while to get a suitable mix [12].

C. Third Party Mixing Services

The third party mixes provide a good degree of anonymity if they are not malicious. According to [1], clients need to trust the third party mixing services. The key idea behind bitcoin was its decentralized nature. To ask users to trust a third party defeats the original idea behind the development of bitcoin. To stay in business, the mixing service needs to build up a reputation and gain users trust which may take some time. The biggest issue is that though the users in the mix are anonymous to each other, they are not anonymous to the mixing service [13]. Hence the anonymity is partial. If the mixing service is malicious, it may also steal bitcoins. Currently, there is no reputed third party mixing service.

D. Altcoins with built-in privacy

This technique is the only one which provides full anonymity without any risk as it incorporates privacy enhancing techniques within the cryptocurrency [16]. The disadvantage is that altcoins which provide full anonymity are not as popular as bitcoin [18]. They may catch up in the future [1]. Altcoins use zero-knowledge proofs for achieving built-in-privacy [15]. Figure 12 shows the comparison of Zcash and Bitcoin.

	ZCASH	BITCOIN
PRIVACY	zk-SNARKs	n/a
BLOCK TIME	2.5 mins	10 mins
BLOCK SIZE	2MB	1MB
MINING ALGORITHM	Equishash	SHA-256
DIFFICULTY ADJUSTMENT	Every block	Every 2016 blocks

Fig. 12. Comparison between bitcoin and Zcash [19]

Figure 13 represents the comparison between various existing altcoins on the basis of the following five categories- Internal Unlinkability, Theft Resistance, DoS Resistance, Bitcoin-compatible and number of transactions [1]. A fully shaded circle means the property is fully present whereas a half shaded circle implies that the particular property is partially present. The absence of circle represents the absence of the property. An interesting observation is that blindcoin is bitcoin compatible and satisfies most of the properties. Bitcoin developer community can use properties of blindcoin to improve its existing code.

Proposal	Class	Internal Unlinkability	Theft Resistance	DoS Resistance	Bitcoin-compatible	# Transactions
CoinJoin	P2P	●			●	1
Shuffle Net	P2P	●			●	1
Fair Exchange	P2P	●			●	4
CoinShuffle	P2P	●	●	◐	●	1
Mixcoin	distr.	◐	◐	●	●	2
Blindcoin	distr.	●	◐	●	●	4
CryptoNote	altcoin	●	●	●		0
Zerocoin	altcoin	●	●	●		2
Zerocash	altcoin	●	●	●		0

Fig. 13. Comparative evaluation of altcoins [1]

E. ToR

ToR is a general anonymous network service. It is easy to integrate ToR with bitcoin as bitcoin core was modified to provide in-built support for ToR. ToR aims to hide who is talking to whom on the network. Yet, there exist several ways through which ToR users can be de-anonymized. By linking bitcoin to ToR, we link bitcoin with all the attacks and vulnerabilities associated with ToR. Additionally, using ToR with bitcoin makes bitcoin vulnerable to man in the middle attacks [10]. It also makes bitcoin software slow. The figure shown below shows the clients state after the man in the middle attack [17]. By performing Man in the middle attack, not only is the identity of the person is revealed, but the person can also be robbed of bitcoins. Some researchers argue that ToR increases existing problems rather than solving them.

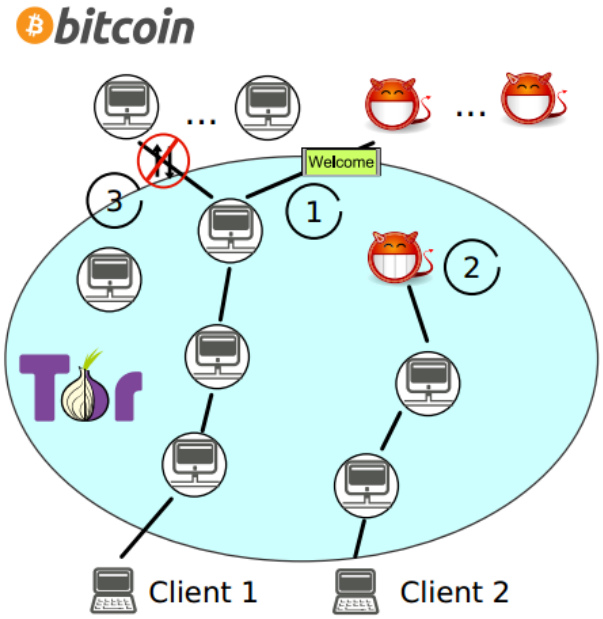

Fig. 14. Man in the middle attack using ToR [10]

VI. Conclusion

This paper highlights the need for implementing new techniques that would improve user anonymity in cryptocurrencies like bitcoin. Many users believe bitcoin to be anonymous. Researchers have proved that this is far from reality. Without privacy, bitcoin is worse than centralized banking. There exist several techniques which aim to enhance the pseudonymity provided by bitcoin. These techniques succeed partially in doing so. None of them provides total anonymity that can be adopted by bitcoin. Techniques like using bloom filters, ToR, peer to peer mixing, and decentralized mixing services are flawed. Table 1 shows all the mentioned techniques with their issues in brief. They would not work in the long run. Using ToR with bitcoin makes bitcoin vulnerable to furthermore attacks. The best approach to achieve full anonymity without involving any third party is to change the bitcoin protocol to have an inbuilt privacy feature like Zerocash. This would truly make bitcoin anonymous. Many ongoing experiments on achieving complete privacy are occurring on various altcoins. These groundbreaking changes should be incorporated into bitcoin. However, a major challenge lies in convincing the bitcoin developer community and bitcoin users to support this move [20].

TABLE I
Privacy Techniques Assessment

Privacy Technique	Issues
Bloom Filters	False positives,low precision,leaks information
P2P Mixing	No internal anonymity, DoS Attack
Third Party Mixing	Trust third part,Risk of theft
AltCoins	Not as popular as bitcoin
ToR	Man in the middle attacks

VII. Overall Assessment

Cryptocurrencies are the future of banking. It is of utmost importance to remove any vulnerability in them. Bitcoin is the most popular cryptocurrency. There is a huge debate over bitcoin privacy within the bitcoin developer community. Researchers have created several techniques to fully anonymize bitcoin but none have proven to work perfectly. After reviewing many research papers on this topic, I conclude that there is no single existing technique in use currently that can be said to be the best for increasing privacy in bitcoin. I feel that a major change needs to be brought into the bitcoin core software to achieve privacy without creating further risks for bitcoin owners. Experiments done on altcoins show that privacy within the protocol can be achieved efficiently. Altcoins like zerocoin and zero-cash have successfully achieved complete privacy, though they are not as popular as bitcoin. It would be interesting to analyze how inbuilt privacy can impact the way in which bitcoin operates and is an open area for research.

Acknowledgment

I would like to thank Dr. Anubha Gupta for guiding me and providing me an opportunity to write a review paper.

References

[1] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "Sok: Research perspectives and challenges for bitcoin and cryptocurrencies," in *2015 IEEE Symposium on, Security and Privacy (SP)*. IEEE, 2015, pp. 104–121.

[2] A. Barrera, "A guide to bitcoin (part i): A look under the hood," 2014. [Online]. Available: http://tech.eu/features/808/bitcoin-part-one/

[3] I. Eyal, "The miner's dilemma," in *2015 IEEE Symposium on, Security and Privacy (SP)*. IEEE, 2015, pp. 89–103.

[4] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, "On the security and performance of proof of work blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 3–16.

[5] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin," in *International Conference on Financial Cryptography and Data Security*. Springer, 2013, pp. 34–51.

[6] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *2016 IEEE Symposium on, Security and Privacy (SP)*. IEEE, 2016, pp. 839–858.

[7] A. Gervais, S. Capkun, G. O. Karame, and D. Gruber, "On the privacy provisions of bloom filters in lightweight bitcoin clients," in *Proceedings of the 30th Annual Computer Security Applications Conference*. ACM, 2014, pp. 326–335.

[8] A. Biryukov, D. Khovratovich, and I. Pustogarov, "Deanonymisation of clients in bitcoin p2p network," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2014, pp. 15–29.

[9] B. Info, "Bitcoin address," 2018. [Online]. Available: https://blockchain.info/address/1JC6DLSJ8PEEsSXViVA8ZvPwWS55RYgMs3

[10] A. Biryukov and I. Pustogarov, "Bitcoin over tor isn't a good idea," in *2015 IEEE Symposium on, Security and Privacy (SP)*. IEEE, 2015, pp. 122–134.

[11] E. K. Kogias, P. Jovanovic, N. Gailly, I. Khoffi, L. Gasser, and B. Ford, "Enhancing bitcoin security and performance with strong consistency via collective signing," in *25th USENIX Security Symposium (USENIX Security 16)*, 2016, pp. 279–296.

[12] Learningspot, "Bitcoin mixing to improve anonymity." [Online]. Available: http://learningspot.altervista.org/bitcoin-mixing-to-improve-anonymity/

[13] L. Spot, "Bitcoin decentralized mixing." [Online]. Available: http://learningspot.altervista.org/bitcoin-decentralized-mixing/

[14] M. Andrychowicz, S. Dziembowski, D. Malinowski, and L. Mazurek, "Secure multiparty computations on bitcoin," in *2014 IEEE Symposium on,Security and Privacy (SP)*. IEEE, 2014, pp. 443–458.

[15] A. Managl, "What is monero? an in-depth guide," 2017. [Online]. Available: https://coincentral.com/what-is-monero/

[16] E. B. Sasson, A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, and M. Virza, "Zerocash: Decentralized anonymous payments from bitcoin," in *2014 IEEE Symposium on,Security and Privacy (SP)*. IEEE, 2014, pp. 459–474.

[17] A. Tiwari, "Everything about tor: What is tor? how tor works?" 2017. [Online]. Available: https://fossbytes.com/everything-tor-tor-tor-works/

[18] A. Miller, A. Juels, E. Shi, B. Parno, and J. Katz, "Permacoin: Repurposing bitcoin work for data preservation," in *2014 IEEE Symposium on, Security and Privacy (SP)*. IEEE, 2014, pp. 475–490.

[19] 99finesse, "Fundamental report: Bitcoin private and the value of privacy on the blockchain," 2018. [Online]. Available: https://medium.com/@99finesse/fundamental-report-bitcoin-private-and-the-value-of-privacy-on-the-blockchain-e552fb7952be

[20] E. Anceaume, T. Lajoie-Mazenc, R. Ludinard, and B. Sericola, "Safety analysis of bitcoin improvement proposals," in *2016 IEEE 15th International Symposium on, Network Computing and Applications (NCA)*. IEEE, 2016, pp. 318–325.