2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) | 978-1-6654-1540-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICBAIE52039.2021.9389894

2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE 2021)

Anonymity Analysis of Bitcoin, Zcash and Ethereum

Yitao Zhou[1,a*], Judong Wu[2,b], Shengxin Zhang[3,c]
[1]New York University Shanghai, Shanghai, China. 200122
[2]Guangdong University of Technology, Guangzhou City, China. 510520
[3]St Joseph High School, Connecticut, United States. 06611
[a]yitao.zhou@nyu.edu
[b]plgawu@gmail.com
[c]szhang2022@students.sjcadets.org

Abstract: **As an innovative type of decentralized model, blockchain is a growing list of blocks linked by cryptography. Blockchain incorporates anonymity protocol, distributed data storage, consensus algorithm, and smart contract. The anonymity protocols in blockchain are significant in that they could protect users from leaking their personal information. In this paper, we will conduct a detailed review and comparison of anonymity protocols used in three famous cryptocurrencies, namely Bitcoin, Zcash, and Ethereum.**

Keywords: Blockchain, Anonymity, Network Analysis

I. INTRODUCTION

Concepts of blockchains may vary, but intuitively blockchain is a special type of database that is shared between nodes in a peer-to-peer network. Symbolic applications of blockchain include Bitcoin, Zcash, and Ethereum. The name blockchain comes from blocks that hold information and these blocks are chained together onto the previous block, forming a chain of data. Blockchain applications use different techniques to achieve the same goals such as decentralization, trustless, tamper-proof, transparency, and security. However, Bitcoin, Zcash, and Ethereum achieved different anonymity goals through diverse design.

Anonymity in blockchain means no central authority is tracing each user's personal information. This is entirely different from conventional transactions with the centralized authority tracking all user information and transactions. Each user can interact with the blockchain network with a randomly generated address [1]. And many addresses can be generated by the same user to avoid identity exposure. This is identity anonymity. Meanwhile, some cryptocurrencies have taken transaction anonymity into consideration, where transactions are secure, typical confidential cryptocurrencies with transaction anonymity include Zcash and Monero. Different protocols are suggested by Bitcoin, Zcash, and Ethereum with the intention to preserve the anonymity of blockchain transactions, or the identity of users, or both.

In this paper, we are going to compare and explore the similarities and differences in the anonymity protocols for the three cryptocurrencies. In Section II, we will first introduce the fundamental backgrounds about the three cryptocurrencies. In Section III, we will respectively discuss what anonymity protocols are enforced. Then in Section IV, we will compare the similarities and differences in respect of the three cryptocurrencies. The final conclusion will be summarized in section IV.

II. BACKGROUND

A. Bitcoin

On October 30, 2008, Satoshi Nakamoto published a paper titled "Bitcoin: A Peer-to-Peer Electronic Cash System" on the cryptography forum, introducing the concept and principles of Bitcoin and the blockchain technology. In his design, Bitcoin is a peer-to-peer electronic cash system. Anyone can join the Bitcoin system at any time and transfer Bitcoin to each other through anonymous addresses without going through a financial institution. With the putting forward of Bitcoin, the payment system based on trust is broken, being replaced by a payment system based on cryptographic proof [2].

B. Zcash

With the influence of Bitcoin on the market, people begin to value cryptocurrencies. Zcash was first introduced in 2013 with the name Zerocoin. As one of the developing cryptocurrency blockchain, Zcash has received support from a large number of users since its introduction in 2016. Similar to some other altcoins, Zcash is also created based on the code of bitcoin. Thus, Zcash and Bitcoin shared many similarities in data and forms. However, Zcash also has a specialty, which is different from other cheap coin imitations. The founders of Zcash show concentration on the anonymity of their users, and they develop a unique anonymity system based on zk-SNARK, a zero knowledge proof.

C. Ethereum

Ethereum is one of the most successful blockchain projects which marks the advent of the blockchain 2.0. Ethereum has a decentralized replicated virtual machine, called Ethereum Virtual Machine (EVM), which can execute Turing-complete scripts and run decentralized applications. Everyone on this network act as a public ledger [3] who will keep a copy of the state of this EVM through running a node. Additionally, any participant can perform arbitrary computation through broadcasting a transaction into the network. Whenever such a transaction is broadcasted, other participants on the network verify, validate, and execute the computation. This causes a state change in the EVM, which is committed and propagated throughout the entire network. All transactions as well as EVM's present state are stored in the blockchain and synchronized by other nodes via peer-to-peer network.

978-1-6654-1540-8/21/$31.00 ©2021 IEEE

III. ANONYMITY PROTOCOLS

A. Bitcoin Anonymity

In the white paper of Bitcoin, Santoshi Nakamoto has already taken identity anonymity into consideration [2]. He stated that although the inherent property of Bitcoin, which is announcing all transactions, prevents transaction anonymity from happening, by guaranteeing that public keys of both parties are anonymous, identity anonymity can be achieved. With such a method, the outside observer can only see that some party is sending an amount of Bitcoin to some other party, but cannot link the transaction to anyone. Meanwhile, Nakamoto also suggested that new pseudonyms (addresses) should be changed for each transaction to further prevent information leakage.

A user in the Bitcoin blockchain can generate several pseudonymous addresses, however, this does not guarantee perfect identity anonymity [4]. Since every transaction is recorded by the public ledger with addresses of both sender and receiver, sometimes the flow of coins is traceable through analyzing trading patterns, thus addresses could be possibly linked to the real identity. One example raised by Bonneau et al. is that a user of Bitcoin blockchain may need to assemble the coins from multiple addresses she owns, during which process information that all these addresses are owned by her is revealed because of gathering transactions [5]. Another example is that Meiklejohn et al. successfully clustered addresses belonging to the same user through developing a new clustering heuristic based on change addresses [6].

Meanwhile, network data sent through nodes are also vulnerable to the leakage of real identity. It is possible that nodes leak IP addresses when broadcasting transactions even when network data walks over anonymous networks such as Tor. Biryukov et al. point out a DoS attack to disconnect Tor exit nodes from the Bitcoin network will leak full control of information flow to the attacker [7].

The most famous fix to the leakage of users' identities is mixing service. It literally means a random exchange of one user's coins with another user's coins through centralized services or peer-to-peer, so that the ownership of coins gets indistinguishable for attackers. The coin mixing solves the problem of users' addresses being linked, but the related centralized services are still risky in leaking users' private information [5]. Therefore, altcoins with integrated unlinkability are proposed. For example, Zcash transactions use a special type of zero knowledge proofs called zk-SNARK, which reveal no information about the transaction amount or recipients, making transactions untraceable. Specific details of Zcash anonymity will be discussed in Section III-B.

B. Zcash Anonymity

Zerocoin, the predecessor of Zcash first put forward the idea of using zero knowledge proof for anonymity, in which the coins are destroyed and recreated for the protection of information. The idea of using zero knowledge proof for anonymity was adopted and developed in different cryptocurrencies, including Dcash, Montero and Zcash. Zcash, created in 2016 by Matthew Green, could fix the anonymity deficiency of Bitcoin. Green developed a privacy system that is made up with two significant protocols: shield pool and zk-SNARK.

In zk-SNARK, a prover sends a statement to the verifier, which turns the transaction validity function into a mathematical representation. The verifier breaks down the logical steps into the smallest possible operations, creating an arithmetic circuit. Figure 1, is the illustration of an sample arithmetic circuit. The verifier checks whether the results of inputs equals to the answer in gates in order to prove this statement as valid [13]. As the first cryptocurrency to employ zk-SNARK, Zcash uses this algorithm to verify the statement provided by the prover and examine the validity of the statement without leaking information regarding the transaction.

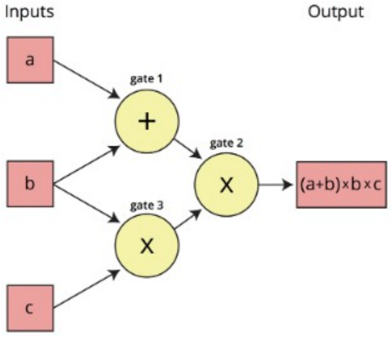


Fig. 1. How zk-SNARK works in Zcash

By using zk-SNARK, transaction information is private regardless of whether the user is using transparent pool or shield pool. Only the participants that are involved in the transactions can receive detailed information. Meanwhile, With zk- SNARK, the recipient might not even know the address of the person that he is making a deal with. This is known as the privacy sending except the shield sending on Zcash. Besides, a transaction can also apply both shield address and zk-SNARK at the same time, as indicated in Zcash's official website, a message can be fully encrypted on the blockchain through the encryption of shield pool, but the user can further increase the message security by applying zk-SNARK. Hence, these two major characteristics guarantee and ensure the strong privacy system of Zcash.

However, the encryption of the shield pool is not unbreakable. Floyd has identified patterns in certain kinds of Zcash transactions that weaken their anonymity [12]. Most of the users in Zcash simply encrypt their address with a shield pool, but they operate this mechanism in a simple, direct and easily identifiable way, such as retrieving the same amount of Zcash coins as they just deposited in a short time. Thus, it is easy to discover the identity of the dealers.

C. Ethereum Anonymity

Ethereum achieves anonymity through its account-based model, meaning all account states are stored on nodes locally instead of being transferred with blocks. Nodes reach a consensus by comparing StateRoot which is the Merkle root of the global state. For personal usage, users are not required to upload their identity to sign up for an account. In terms of anonymity, it is hard to track assets and identify real people if users keep aware of the vulnerabilities inherent in an account-based model, but we cannot expect everyone using Ethereum to have the same awareness of protecting their identity. It is likely that some people would prefer to share their account addresses while socializing with others. Then all their previous records in the Ethereum network would be leaked to the public because their transactions would be traceable. This is the shortcoming of the account-based model in guaranteeing anonymity.

With the development of applications, Ethereum is losing its anonymity, Béres et al. [8] shows a new method to deanonymize Ethereum users by applying several quasi-identifiers stemming from address reuse (time-of-day, activity, transaction fee and transaction graph). "Careless usage easily reveals links between deposits and withdraws and also impacts the anonymity of other users, since if a deposit can be linked to a withdrawal, it will no longer belong to the anonymity set" [8]. Especially, it is becoming easier to link real people through Ethereum Name Service (ENS) since ENS supplies human- readable names. For example, some well-known ENS like *consensys. eth, metamask. eth*, etc [9] have high relevancy with their owners who are public figures in Ethereum community. Additionally, apart from ENS, there are other risks that you would never know before your identity gets revealed, like online trackers and cookies [10]. Even when you use a mixer like Tornado Cash to transact, which makes tracing funds more difficult, Chain analysis can still track the transactions [11].

Fortunately, "Ethereum Layer2 scaling strategy has basically succeeded", as the co-founder of Ethereum, Vitalik Buterin said on twitter. As the extension of Layer1 (Ethereum mainnet), Layer2 network can carry out some transactions with high frequency and low value which make no sense to mainnet. The main function of Layer2 is to reduce Layer1 network load and improve network throughput. Meanwhile, some solutions adopted secure cryptography to optimize the anonymity problem of account-based models. Especially zk-Rollup, the zero knowledge proof solution for Ethereum Layer2, has grabbed more attention and might be possible to solve all anonymity concerns with a single anonymity protocol.

The ZK-Rollup solution consists of two types of users: transactors and relayers. Transactors create transfers off-chain and broadcast the transfers to network. Relayers conduct three operations. First, they collect and roll up transfers into a single transaction to create a rollup block. Then they generate a zk-SNARK proof which prove validity of block state transitions (from previous merkle root to new one). Finally, the state of merkle root, transactions and zk-SNARK proof will be submitted to rollup smart contract. After smart contract finished implementing verification of proof, the state of merkle root which record all transactions will be written to Ethereum network as CALLDATA.

Transactors in Ethereum layer2 can hide their address because relayers bundle transfers and index addresses off-chain, without relying on Ethereum layer1 storage. Furthermore, even the relayers are malicious, it is hard for them to commit an invalid or manipulated state because of the zk-SNARK proof. The worst action that relayers can do is only refusing to serve transactors.

zk-Rollup not only improves the privacy of transactions, but also scalability. Because zk-Rollup has used zk-SNARK proof, it is able to reduce computation by only verifying the proof of state instead of verifying all detailed transaction data. Meanwhile, the new structure of transfer data in zk- Rollup include an indexed "from" and "to" address, a value to transact, the transaction fee and nonce as illustrated in Table I. This new structure can speed up transaction since it only contains 15 bytes instead of Ethereum layer1's current 68 bytes. Thus, the number of transactions in a fix-sized block can be enlarged.

TABLE I NEW STRUCTURE OF A SINGLE TRANSFER IN ZK-ROLLUP

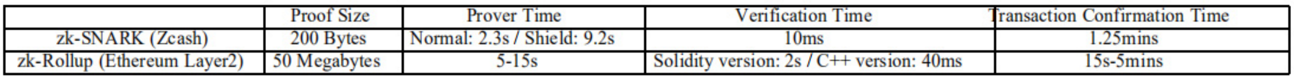
From	To	Value	Fee	Nonce
3 bytes	3 bytes	6 bytes	1 byte	2 bytes

IV. COMPARISON OF ANONYMITY PROTOCOLS

For the purpose of achieving anonymity in a complex blockchain system, different techniques are used. For Bitcoin and Ethereum Layer1, the most famous two cryptocurrency projects in blockchain, only identity anonymity is considered. To achieve identity anonymity, both Bitcoin and Ethereum have proposed mixing services. Ethereum has Tornado Cash, and Bitcoin also has some famous projects in mixing service including Mixcoin [14] and CoinJoin [15]. It is analyzed by Foxley that" Tornado Cash is more readily compared to existing coin mixers on Bitcoin because of its retail focus" [11]. However, the mixing service is still vulnerable to attack, since ordinary people's casual operations on Tornado Cash or Coinjoin may leak information about their wallets. Thus, with the development of blockchain technology, people view transaction anonymity as the solution of flaws in identity anonymity, therefore, Ethereum layer2 and Zcash have taken transaction anonymity into consideration.

As stated in section III-B and section III-C, Zcash applies zero knowledge proof called zk-SNARK to achieve transaction anonymity, and Ethereum Layer2 has build upon zk-SNARK, roll it up with the state of merkle root to form a smart contract, calling it zk-Rollup. Through the application of zk-SNARK and zk-Rollup, it becomes possible for Zcash and Ethereum Layer2 to efficiently perform secret arbitrary transactions that are verifiable by anyone.

TABLE II COMPARISON OF ZK-SNARK (ZCASH) AND ZK-ROLLUP (ETHEREUM LAYER2), DATA FROM [16], [17]



	Proof Size	Prover Time	Verification Time	Transaction Confirmation Time
zk-SNARK (Zcash)	200 Bytes	Normal: 2.3s / Shield: 9.2s	10ms	1.25mins
zk-Rollup (Ethereum Layer2)	50 Megabytes	5-15s	Solidity version: 2s / C++ version: 40ms	15s-5mins

The similarity between zk-Rollup and zk-SNARK is using separate merkle trees to store transactions and generate a zero knowledge proof to verify the new state of merkle tree. The difference of zk-Rollup and zk-SNARK is that zk-SNARK only support anonymity, while zk-Rollup have taken scalibility into consideration. Therefore, zk-Rollup is more practical to be enforced. Zhang et al. shows Zcash have around 95% linkable transactions, meaning low transaction anonymity [18]. And in Ethereum Layer2, transaction anonymity will be possibly enabled by default through zkSync. Furthermore, Aztec has deployed a new scripting language called Nori to support transaction anonymity [19]. Although zk-Rollup is still in experiment phase, but it is worth to look forward.

We also conducted a comparison of proof size, prover time, verification time and transaction confirmation time for zk-SNARK and zk-Rollup, which are summarized in Table II. The reason that might cause the difference could be the different ways of utilizing this algorithm. Since zk-SNARK works directly as an algorithm in Zcash, it have a less computation intensity which causes the speed to be faster. Besides, in comparing transaction cost and size of transaction, zk-SNARK has won the competition in both of them, but zk-Rollup does have a faster transaction speed than zk-SNARK.

V. CONCLUSION

In this paper, we first make a brief introduction of three famous cryptocurrencies, namely Bitcoin, Ethereum and Zcash. Bitcoin takes care of identity anonymity, Ethereum has just proposed Layer2 to guarantee both identity and transaction anonymity, and Zcash has proposed zk-SNARK zero knowledge proof to guarantee transaction anonymity at birth.

Then, we respectively discuss and compare the technologies employed by different cryptocurrencies to achieve anonymity. Since Bitcoin and Ethereum layer1have public ledgers and account-based models, they are nearly impossible to achieve transaction anonymity. Meanwhile, identity anonymity is also not a guarantee because of non-standard usage. Luckily, Bitcoin and Ethereum layer1 applied mixing services to enhance identity anonymity, although still requiring standard usage. Meanwhile, Ethereum has proposed Layer2 which takes both identity anonymity and transaction anonymity into consideration by developing zk-SNARK zero knowledge proof into zk-Rollup. Since their similarity in adopting zk-SNARK for the sake of transaction anonymity, we also conducted a comparison between the proof size, prover time, verification time and transaction confirmation time for zk-SNARK and zk-Rollup. We found that zk-Rollup in Ethereum has a longer proof time and verification time compared with zk-SNARK in Zcash, but zk-Rollup does have a faster transaction speed than zk-SNARK. Although Zcash's anonymity system and zero knowledge proof are relatively faster and stronger than Ethereum's currently, there are greater potentials for Ethereum to improve since it has higher scalibility.

REFERENCES

[1] Z. Zheng, S. Xie, H. N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: a survey," International Journal of Web and Grid Services, vol. 14, no. 4, p. 352, 2018.

[2] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System,"Manubot, 2019.

[3] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," Ethereum Project Yellow Paper, 2014.

[4] R. Zhang, R. Xue, and L. Liu, "Security and Privacy on Blockchain,"ACM Computing Surveys, 2019.

[5] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll and E. W. Felten, "SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies," 2015 IEEE Symposium on Security and Privacy, San Jose, CA, 2015, pp. 104-121, doi: 10.1109/SP.2015.14.

[6] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. Mccoy, G. M. Voelker, and S. Savage, "A fistful of bitcoins," Proceedings of the 2013 conference on Internet measurement conference - IMC '13, 2013.

[7] A.Biryukov and I. Pustogarov, "Bitcoin over Tor isn't a Good Idea," 2015 IEEE Symposium on Security and Privacy, San Jose, CA, 2015, pp. 122-134, doi: 10.1109/SP.2015.15.

[8] F. Be´res, I. A. Seres, A. A. Benczu´r, and M. Quintyne-Collins, "Blockchain is Watching You: Profiling and Deanonymizing Ethereum Users," arXiv.org, 13-Oct-2020.

[9] T. Copeland, "We tracked 133,000 Ethereum names and exposed their secrets," Decrypt, 21-Feb-2020. [Online]. Available: https://decrypt.co/19423/we-tracked-133000-ethereum-names-and-exposed-their-secrets.

[10] S. Goldfeder, H. Kalodner, D. Reisman, and A. Narayanan, "When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies," arXiv.org, 16-Aug-2017. [Online]. Available: https://arxiv.org/abs/1708.04748.

[11] W. Foxley, "Developers of Ethereum Privacy Tool Tornado Cash Smash Their Keys," CoinDesk, 19-May-2020. [Online]. Available: https://www.coindesk.com/developers-of-ethereum-privacy-tool-tornado-cash-smash-their-keys.

[12] D. Floyd, "Zcash Privacy Weakened by Certain Behaviors, Researchers Say," CoinDesk, 09-May-2018. [Online]. Available: https://www.coindesk.com/zcash-privacy-weakened-by-certain-behaviors-researchers-say..

[13] K. Peters, "What is zk-SNARK?," Investopedia, 01-Dec-2020. [Online]. Available: http://www.investopedia.com/terms/z/zksnark.asp.

[14] J. Bonneau, A. Narayanan, A. Miller, J. Clark, J. A. Kroll, and E. W. Felten, "Mixcoin: Anonymity for Bitcoin with Accountable Mixes," SpringerLink, 03-Mar-2014.

[15] M. Gregory, "CoinJoin: Bitcoin privacy for the real world." Post on Bitcoin forum. 2013.

[16] Network, "Practical ZK-SNARKs for Ethereum," Medium, 26-Aug- 2020. [Online]. Available: https://medium.com/coinmonks/practical-zk- snarks-for-ethereum-140cbddcb55d.

[17] Matter-Labs, "matter-labs/awesome-zero-knowledge-proofs," GitHub. [Online].Available: https://github.com/matter-labs/awesome-zero-knowledge-proofs.

[18] Z. Zhang, W. Li, H. Liu and J. Liu, "A Refined Analysis of Zcash Anonymity," in IEEE Access, vol. 8, pp. 31845-31853, 2020, doi: 10.1109/ACCESS.2020.2973291.

[19] J. Andrews, "Aztec: zkRollup Layer 2 + Privacy", Medium, 2020. [Online]. Available: https://medium.com/aztec-protocol/aztec-zkrollup-layer-2-privacy-1978e90ee3b6.