Blockchain Analysis Tool of a Cryptocurrency

Robert Werner
Institute for Software and Systems Engineering, Clausthal University of Technology
Arnold-Sommerfeld-Straße 1
38678 Clausthal-Zellerfeld, Germany
robert.werner@tu-clausthal.de

Sebastian Lawrenz
Institute for Software and Systems Engineering, Clausthal University of Technology
Arnold-Sommerfeld-Straße 1
38678 Clausthal-Zellerfeld, Germany
+49 5323 72-7176
sebastian.lawrenz@tu-clausthal.de

Andreas Rausch
Institute for Software and Systems Engineering, Clausthal University of Technology
Arnold-Sommerfeld-Straße 1
38678 Clausthal-Zellerfeld, Germany
+49 5323 72-8232
andreas.rausch@tu-clausthal.de

ABSTRACT

In recent years, cryptocurrencies have become more and more popular and the growing adoption has led to an increasing number of financial transactions being stored on the blockchain. Although cryptocurrencies have built a reputation as an anonymous means of payment, they are usually rather pseudonymous, transparent and everlasting logbooks about financial transactions, which are publicly available. Thus, analyzing a crypto address can reveal payment partners, money flows, behavior patterns and more. In this work, a program is presented, which provides an analysis of this kind and displays the results in a simple format. The technical properties of the blockchain that this analysis is based on are explained. This paper explains the possible impact of total transparency on the blockchain and our tool on our society.

CCS Concepts

•Information systems → World Wide Web → Web applications → Electronic commerce → Electronic data interchange • Security and privacy → Systems security → Distributed systems security •Computer systems organization → Architectures → Distributed architectures

Keywords

Blockchain; Cryptocurrency; Privacy; Trust; Transparency;

1. INTRODUCTION AND MOTIVATION

Since the introduction of the Bitcoin in the year 2009, the acceptance and establishment of cryptocurrencies have increased more and more. At its peak, one Bitcoin had a value of nearly 17.000 thousand euros in 2018. Moreover, many other cryptocurrencies, called altcoins, have been developed and established.

The acceptance of cryptocurrencies continues to rise and as they become a part of our everyday life. On one hand, cryptocurrencies are currently a very popular investment while on the other hand, they have already become an accepted form of payment in some domains. With the *estcoin* Estonia plans its own cryptocurrency [1]. Another popular and at the moment highly discussed project is Libra, a cryptocurrency developed and driven by Facebook. Libra's self-declared goal is nothing less than to become the world's most widely used cryptocurrency [2].

Many people see cryptocurrencies as a fast, simple and future-proof means of payment. On the other hand, however, cryptocurrencies are also popular in the darknet, because they are considered an anonymous means of payment. It is possible to buy credit card data, drugs and weapons in the darknet via bitcoin and the total volume of such Bitcoin transactions in 2018 was approximately 600 million dollars [3]. But, is a cryptocurrency really an anonymous means of payment?

The blockchain, as used by Bitcoin (BTC), for example, is an immutable ledger, which is stored on a large network of servers worldwide in a decentralized manner. On this ledger, all transactions are stored permanently, transparently and can be accessed by anyone. At least one sender and at least one receiver can be assigned to every transaction, which also makes it possible to assign all transactions to an address. This property offers several possibilities: For example, payments are verifiable and expenses can be transparently disclosed. At the same time, the payment partners and the possible dependence on them can be traced by the public. With the increasing adoption of transparent cryptocurrencies in the physical world, a person's transaction history could allow conclusions to be drawn about their social environment, their whereabouts and buying habits.

However, the traceability of a cryptocurrency bears consequences not only for the individual but also for the entire network and all of its users. This is because of the fact that the origin of the digital coins is always evident and there is a history tied to them. In the event that this history is investigated and indicates illegal activities, for example, these coins can be rejected by service providers and are considered tainted. As a result, two coins of the same denomination have a different value and are therefore not fungible i.e. interchangeable.

In any case, such an analysis takes time and requires a basic understanding of transactions on the blockchain. For that reason, it is reserved only for a small community, bearing the consequence that many users don't even know about this possibility.

The goal of this paper is to review and discuss the anonymity of cryptocurrencies. We first introduce the Blockchain technology and explain how transactions in blockchain work. In the next subsection, we introduce block explorers, a tool to investigate the history of a coin and discuss the limits of this

ICBCT'20, March 12–14, 2020, Hilo, HI, USA

ACM ISBN 978-1-4503-7767-6/20/03.
https://doi.org/10.1145/3390566.3391671

approach. In section 3 we present our Blockchain analysis tool (available online at [4]), which goes beyond the “classic” block explorers. Finally, we show the evaluation of our tool and discuss the results in the conclusion.

2. BACKGROUND

2.1 Blockchain and Transactions

The blockchain consists of a series of chronologically sorted blocks [5], with each of them consisting of at least one transaction. Transactions on the Bitcoin blockchain always consist of at least one input and one output (generation transactions excluded) [6], [7]. The combined input value is equal to the output value. Each transaction has a unique transaction ID and each input is a UTXO [1] under the control of the sending party. This prevents coins from being spent multiple times (double spending) which would cause uncontrolled inflation for the cryptocurrency. It is only in generation transactions [2] that new coins may be created [8], which will be distributed to the miners or stakers as the block reward. The block reward serves as an incentive for consensus creation as well as coin distribution [9].

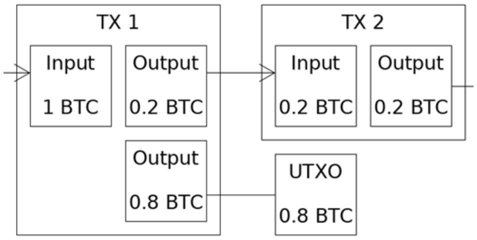


Figure 1. Transaction inputs and outputs

Figure 1 shows, using two Bitcoin transactions in which each input consists of a preceding output of equal value. If an output has not yet been reused as an input, it is called UTXO and can still be used as input.

To increase privacy, most wallets by default generate a new receiving address for each transaction requested. This prevents all transactions of an individual to occur on a single address, therefore making it more challenging to trace. The additional addresses of the same person can be located in the following scenario:

If the value of the desired outputs is greater than a single UTXO (which is being used as an input), several UTXOs are combined with each other until the value of the desired output is reached or exceeded. Therefore, if the first input comes from a known address, it can be assumed that all further inputs, even if they come from another address, are controlled by the same wallet or owner.

Figure 2 shows a transaction consisting of two inputs from two different addresses. It can be assumed that the addresses 'bc1qA...' and 'bc1qB...' belong to the same person.

If the value of all desired outputs is less than the value of the UTXOs (Inputs), the difference is sent back to one of many change addresses of the sender. These new addresses can later be determined when used in combination with an already known address, as explained above.

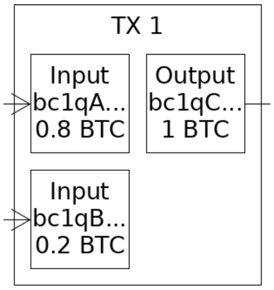


Figure 2. Input combinations

In TX 1 Figure 3, 0.8 BTC is sent from address A to address B. However, since only one UTXO with a value of 1 BTC is available, the remaining amount is sent to a change address A'. Since in TX 2 the addresses A and A' are combined as input, it is obvious that both addresses are controlled by the same person.

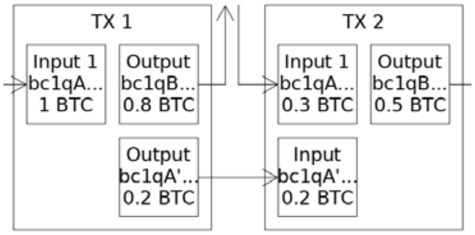
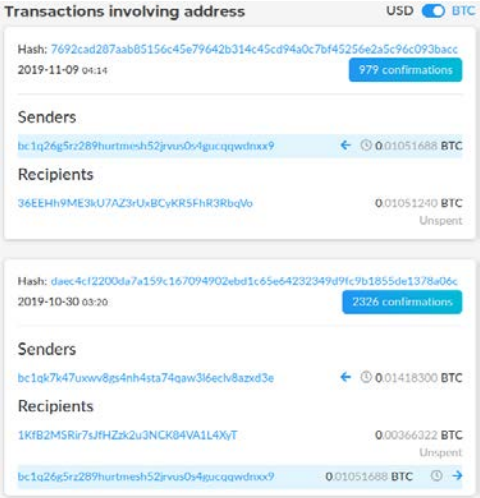


Figure 3. Change addresses

2.2 Blockexplorer

In order to make the blockchain readable for humans, there is a multitude of Block explorers. These block explorers can enlist all blocks and transactions and offer the user the possibility to verify transactions or check the status of the network.


Transactions involving address USD BTC

Hash: 7692cad287aab85156c45e79642b314c45cd94a0c7bf45256e2a5c96c093bacc
2019-11-09 04:14 979 confirmations

Senders
bc1q26g5rz289hurtmesh52jrvus0s4gucqqwdnxx9 ← 0.01051688 BTC

Recipients
36EEHh9ME3kU7AZ3rUxBCyKR5FhR3RbqVo 0.01051240 BTC
Unspent

Hash: daec4cf2200da7a159c167094902ebd1c65e64232349d9fc9b1855de1378a06c
2019-10-30 03:20 2326 confirmations

Senders
bc1qk7k47uxwv8gs4nh4sta74qaw3l6eclv8azxd3e ← 0.01418300 BTC

Recipients
1KfB2MSRir7sJfHZzk2u3NCK84VA1L4XyT 0.00366322 BTC
Unspent
bc1q26g5rz289hurtmesh52jrvus0s4gucqqwdnxx9 0.01051688 BTC →

Figure 4. Block explorer transactions list [10]

Most block explorers also offer the possibility to view the capital and any transactions of an address, which allows for manual analysis. However, this is very time-consuming and requires extensive technical knowledge about the respective cryptocurrency. It is not immediately clear which other addresses

[1] UTXO (Unspent Transaction Output): an output from a preceding transaction that has not yet been used as an input in another transaction.

[2] the first transaction in a block without an input from a previous transaction

the owner controls, and which addresses and to what extent they are transacting with.

In the extract from the block explorer shown above (see Figure 4), two transactions are listed in connection with the light blue shaded address. This makes it possible to trace individual transactions and the assets of this address. The transactions are presented in detail and link to all occurring addresses, previous outputs, and subsequent inputs. This makes it possible to track cash flows manually.

3. IMPLEMENTATION

The most important decision for the analysis tool was the choice of the Cryptocurrency to be analyzed. Criteria for this were mindshare, analyzable properties and the simplicity of working with it.

Although Bitcoin is by far the most widespread cryptocurrency, it is difficult to work with its blockchain size which is greater than 210 GB. In order to make the analysis tool largely compatible with Bitcoin, the decision was made to choose the currency based on Bitcoin called PIVX. The size of PIVX's blockchain is approximately 17 GB3 and is therefore relatively easy to process compared to that of Bitcoin. In addition to transparent transactions, as known from Bitcoin, PIVX also offers the possibility to create anonymous transactions and, by participating in the Proof of Stake consensus algorithm or operating a master node, to generate income directly from the blockchain.

PIVX offers a variety of analyzable data and combines transparency and anonymity on the blockchain while maintaining basic compatibility with Bitcoin.

In order to analyze the transaction history, the blockchain needs to be in a data format that allows drawing links between multiple transaction inputs, as well as finding the past and future usage of a specific UTXO. Since the PIVX Node stores the whole Blockchain we have direct and fast access to the ledger. However, the reason for the node storing the whole Blockchain is solely to verify all transactions in the network. That makes it easy to get transaction details immediately via JSON-RPC[4], which lists all TXOs that are used as an input and all outputs with the corresponding value and target address, but makes it difficult to get the address of an input, future uses of output or a list of transactions authorized by an address.

In order to find those transactions, we need to assign the inputs of a transaction to an address and index all existing addresses, which can be achieved by parsing the blockchain into a relational database.

3.1 Parsing

In order to fill the relational database with data, a script iterates over each block and every transaction contained in it and adds them to the database.

To speed up the later real-time analysis, some evaluations can already be made during the parsing process. For example, an input can be assigned an address and a value by viewing the previous output. Similarly, an output can be assigned the Block Reward, if the parent transaction does not have any inputs.

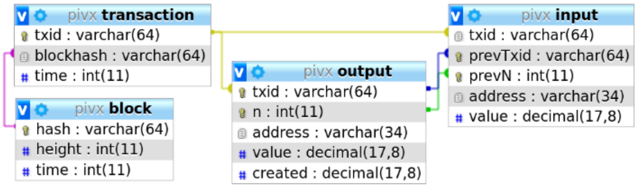


Figure 5. Relational database model

In the database model (see Figure 5), each input and output were assigned the corresponding address and value. In addition, the field “created'” was assigned to each output, in which, the Block Reward can be stored.

Although the initial parsing process is a lengthy procedure, it can be performed incrementally because the database is being stored permanently and the blocks on the blockchain are in chronological order. This makes it possible to update the database to the latest state of the blockchain within seconds.

3.2 Analysis

Thanks to the flexible relational database, it is now possible to capture all incoming and outgoing transactions to a specific address with relatively little effort.

With the help of these transactions, all payment partners can now be located and sorted according to the frequency of interaction or payment amount.

In addition, it is now possible to derive other addresses (change addresses and other receiving addresses) of the same user from a transaction.

This is done by taking all addresses from all inputs from transactions that contain a known address as an input.

4. EVALUATION

The analyzed data is now presented visually and can be interpreted.

From the diagram (see Figure 6) above you can see that a single destination address receives about a third of all transactions sent by the address being analyzed.

This diagram in Figure 7 shows the exact same address being analyzed and is clearly showing, that the same destination address is actually receiving more than half of the total value sent.

Furthermore, the many “non-standard” transactions (see Figure8) show that this person participated in the consensus and received Staking Reward.

The analysis of a different address (see Figure 9) gives us the information that the user has carried out most outgoing transactions via the zerocoin protocol, i.e. privately. It is clear the user was interested in privacy.

The diagram (see Figure 10) shows an even distribution over many source addresses. This indicates an operator of a master node, as the address is paid regularly by many different consensus participants.

[3] as of March 2019

[4] JavaScript Object Notation Remote Procedure Call: The file format used to communicate with the node. Return values are in JSON.

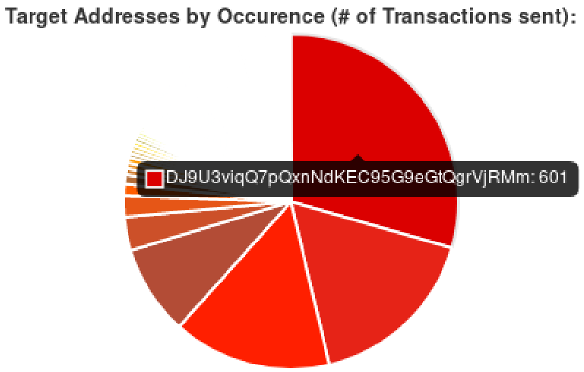


Figure 6. Target address with the highest frequency

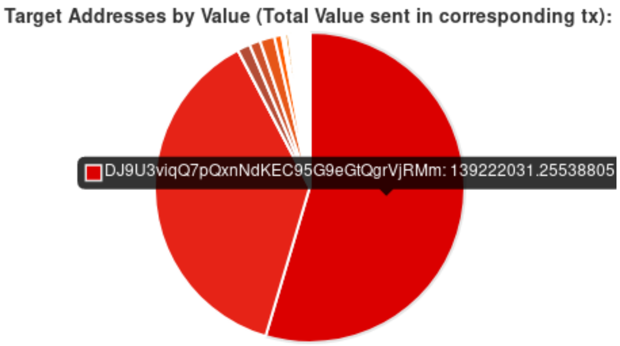


Figure 7. Target address with the highest value

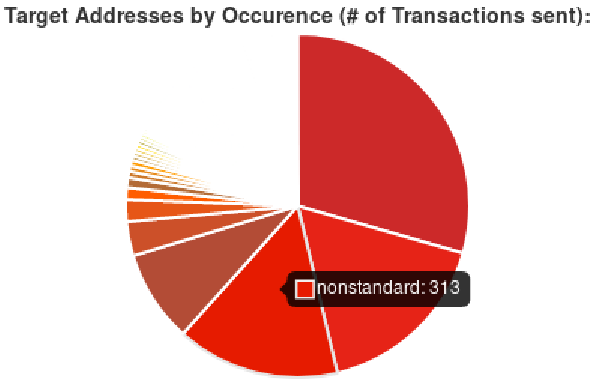


Figure 8. Staking analysis

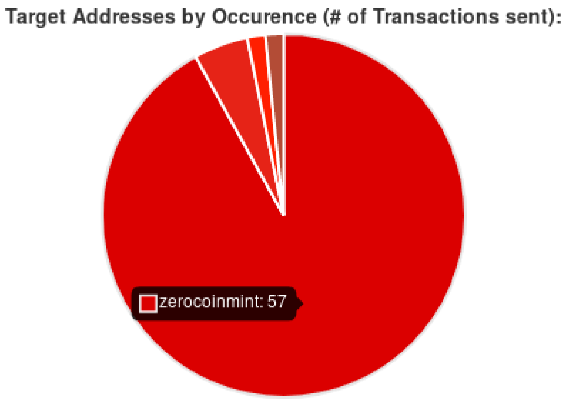


Figure 9. Zerocoin analysis

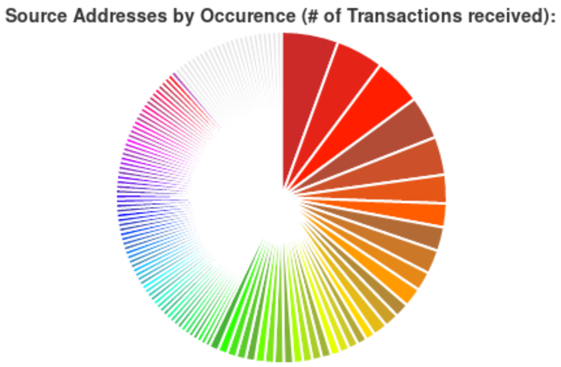


Figure 10. Masternode analysis

5. DISCUSSION

Carrying out a meaningful analysis is made difficult particularly by transactions containing simultaneous multiple inputs and multiple outputs. Even though it is possible to determine the frequency at which two addresses were in contact, it is not possible to determine transferred values between an address from the input pool and an address from the output pool. In such cases, it would make sense to analyze the owner of the address (i.e., all input addresses).

The use of change addresses also complicates finding the destination address with a 100% guarantee, as long as the change address is not being used in combination with an address that is already known. Consequently, an analysis of a person usually becomes more effective after a longer period of time.

6. CONCLUSION

The blockchain analysis tool successfully shows how much information a public address on a transparent blockchain can contain and how these can be automatically evaluated. In particular, the connections to other addresses are of great interest and could in the future lead to many applications. For example, the purchasing behavior of consumers can be traced precisely and compared to their social environment. This can be done without the explicit consent or awareness of the target group due to the publicly observable blockchain. The analysis of cash flows or partnerships of competing companies could also become more relevant in the future.

The presented tool creates awareness for the transparency on the blockchain and can serve as a decision aid when choosing the technology for the desired application purpose. For example, governments, banks, businesses, and citizens should be able to consciously choose for or against a government-backed, possibly transparent, currency, such as the planned Estcoin [11].

At the same time, this paper also makes it clear that the analysis results do not always have to be complete. This may be due to a lack of data, a person using multiple wallets, or a person putting a special focus on privacy. Even in such cases, the blockchain can be analyzed, albeit with considerably more effort. With 'timing analysis' and 'amount analysis' not only linked inputs and outputs are considered, but also the temporal sequence and values of supposedly independent transactions [12] Automating this reliably is not very realistic because individual assumptions and interpretations would often have to be made. However, the blockchain does not provide a time frame, so that a manual analysis can be executed over a longer period of time.

Ultimately it is up to the people to decide whether or not they want a transparent, pseudonymous blockchain, like Libra [13], to store some of their most sensitive data publicly and on a permanent record.

7. ACKNOWLEDGMENTS

This paper evolved from the research project “Recycling 4.0” (digitalization as the key to the Advanced Circular Economy using the example of innovative vehicle systems) which is funded by the European Regional Development Fund (EFRE | ZW 6-85017297) and managed by the Project Management Agency NBank.

8. REFERENCES

[1] Browne ,R. “Estonia won’t issue national cryptocurrency estcoin, never planned to,” 2018. [Online available] https://www.cnbc.com/2018/06/04/estonia-wont-issue-national-cryptocurrency-estcoin-never-planned-to.html. [Accessed: 11-Nov-2019].

[2] Laaff, E. K. M. “Kryptowährung Libra: Einmal mit Facebook zahlen, bitte! | ZEIT ONLINE,” 2019. [Online available] https://www.zeit.de/digital/internet/2019-06/kryptowaehrung-libra-facebook-bitcoin-blockchain. [Accessed: 11-Nov-2019]

[3] “Bitcoin boomt als Zahlungsmittel im Darknet,” Wirtschaftswoche, 2019. [Online available] https://www.wiwo.de/finanzen/geldanlage/kryptowaehrung-bitcoin-boomt-als-zahlungsmittel-im-darknet/23884938.html. [Accessed: 11-Nov-2019].

[4] “GitHub - rw501/Blockchain-Analysis-Tool-of-a-Cryptocurrency: includes a blockchain parser to mysql.” [Online available] https://github.com/rw501/Blockchain-Analysis-Tool-of-a-Cryptocurrency. [Accessed: 15-Nov-2019].

[5] Zheng, Z., Xie, S., Dai, H., Chen X. and Wang, H. “An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends,” *Proc. - 2017 IEEE 6th Int. Congr. Big Data, BigData Congr. 2017*, pp. 557–564, 2017.

[6] Nakamoto, S. “Bitcoin: A Peer-to-Peer Electronic Cash System.” Manubot, 2019 [Online available] https://git.dhimmel.com/bitcoin-whitepaper/ [Accessed: 13-Jan-2020]

[7] Antonopoulos, A. M. Mastering Bitcoin: Unlocking Digital Crypto-Currencies, 1st ed. O’Reilly Media, Inc., 2014.

[8] “Coinbase - Bitcoin Wiki.” [Online available] https://en.bitcoin.it/wiki/Coinbase. [Accessed: 15-Nov-2019].

[9] Draupnir, M. “What is the Bitcoin Mining Block Reward?,” 2016. [Online available] https://www.bitcoinmining.com/what-is-the-bitcoin-block-reward/. [Accessed: 15-Nov-2019].

[10] “Bitcoin / Address / bc1q26g5rz289hurtmesh52jrvus0s4gucqqwdnxx9 — Blockchair.” [Online available] https://blockchair.com/bitcoin/address/bc1q26g5rz289hurtmesh52jrvus0s4gucqqwdnxx9. [Accessed: 15-Nov-2019].

[11] “E-Residency 2.0 Whitepaper.” [Online available]: https://s3.eu-central-1.amazonaws.com/ereswhitepaper/e-Residency+2.0+white+paper+English.pdf [Accessed: 15-Nov-2019

[12] Van Wirdum, A. “Is Bitcoin Anonymous? A Complete Beginner’s Guide | Bitcoin Magazine,” 2015. [Online available] https://bitcoinmagazine.com/articles/is-bitcoin-anonymous-a-complete-beginner-s-guide-1447875283. [Accessed: 15-Nov-2019].

[13] “Libra’s mission is to enable a simple global currency and financial infrastructure that empowers billions of people. An Introduction to Libra.” [Online available]: https://libra.org/en-US/wp-content/uploads/sites/23/2019/06/LibraWhitePaper_en_US.pdf [Accessed: 15-Nov-2019]