Front. Comput. Sci., 2023, 17(6): 176215
https://doi.org/10.1007/s11704-023-2531-0

LETTER

BTC-Shadow: an analysis and visualization system for exposing implicit behaviors in Bitcoin transaction graphs

Ding BAO[1], Wei REN (✉)[1,2], Yuexin XIANG[1], Weimao LIU[3], Tianqing ZHU[1], Yi REN[4], Kim-Kwang Raymond CHOO[5]

1 School of Computer Science, China University of Geosciences, Wuhan 430074, China
2 Hubei Key Laboratory of Intelligent Geo-Information Processing, Wuhan 430074, China
3 NSFOCUS Technologies Group Co., Ltd., Beijing 100089, China
4 School of Computing Sciences, University of East Anglia, Norwich NR4 7TJ, UK
5 Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249-0631, USA

1 Introduction

A Bitcoin ledger comprises a sizable number of transaction records, which can be utilized to make it easier to track and analyze the traits and patterns of cryptocurrency-related transactions. To facilitate the visual analysis of Bitcoin, numerous tools with various aims have been developed. For example, MiningVis [1] and SuPoolVisor [2] are the analytics systems for Bitcoin mining pools, as well as [3−5] focus on the Bitcoin transaction graphs analysis.

However, due to our previous research requirements for Bitcoin transaction graphs, none of the available tools can provide exploring the features of connection related to the address and observe significant visual patterns. Specifically, using these tools is challenging to navigate to abnormal node clusters effortlessly from large node groups and then analyze the local interlink characteristics between transaction nodes with interactive analytics.

To handle these specific tasks, a study should plan appropriate visualization techniques for large-scale scope and adequate interactions for visual analysis. Therefore, we go over the types and characteristics of the Bitcoin transaction graph and propose a highly interactive Bitcoin analyzing system (hereafter referred to as BTC-Shadow). In contrast to previous studies as shown in Table 1, BTC-Shadow uses a novel node-cutting method and the fastest layout algorithm named ForceAtlas2 to generate graphs. As a result, BTC-Shadow enables more knowledgeable and precise visual inspection by supporting a higher level of human-computer interaction. A summary of our contributions in this paper is as follows:

- With a variety of customized interactive interfaces to display the attributes of transaction graphs, we create BTC-Shadow to disclose hidden characteristics of Bitcoin transactions.
- We construct a heterogeneous graph of Bitcoin transactions as well as an isomorphic graph that is not provided by similar tools or systems.
- We suggest a node-cutting method to reduce the number of nodes in graphs, which can make large-scale Bitcoin transaction graph analysis easier while preserving network structures possible.

Table 1 Comparison of visualization for Bitcoin transaction graph

Scheme	Graph interaction	Attributes display	Layout algorithm	Multiple workplace	Customization
BlockChainVis [3]	√	√	×	×	×
BitVis [4]	√	√	√	×	×
BiVA [5]	×	×	×	√	×
BTC-Shadow	√	√	√	√	√

Received August 13, 2022; accepted April 10, 2023

E-mail: weirencs@cug.edu.cn

2 Proposed scheme

2.1 Data process methods

We choose to crawl JSON format data by restful interface on the website btc.com for the crucial cause that it includes information on Bitcoin flows. Bitcoin flows are crucial for visual analysis because only through them can we trace unspent transaction output (UTXO).

2.2 Data visualization

2.2.1 Bitcoin transaction graph visualization

We depict transactions and the accounts involved in them as yellow and white nodes, and then we specify the Bitcoin flows between those nodes using colored lines. It is conceivable to see two different graph types: isomorphic, which just contains accounts, and heterogeneous, which also includes transactions. There are numerous characteristics to handle when these two graph types are applied to the Bitcoin graphs: 1) A transaction

has one or more than one input and output which include one or more addresses, called C-1; 2) The coinbase transaction is the reward of mining that is the first transaction for any blocks. The input of the coinbase transaction contains no address but has Bitcoins, called C-2.

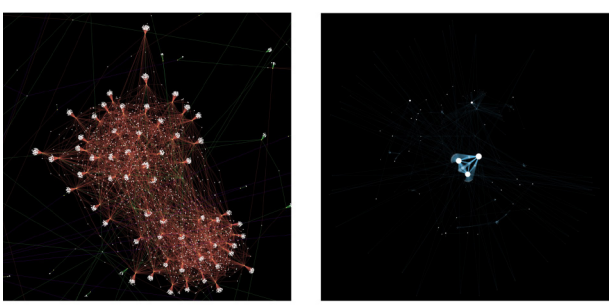
For C-1, when Bitcoins from a single output is held by two or more accounts, we believe that the first account that appears in the output is the owner of those Bitcoins. In the isomorphic graph, each input is connected to each production in a single transaction, whereas in the heterogeneous network, each input and output is connected to its transaction. We consider each empty input address of a coinbase transaction to be a special address hash for C-2.

2.2.2 Node-cutting strategy

It would be challenging for users to spot abnormal features in the case of observing a lot of transactions. We describe a node-cutting strategy for the Bitcoin transaction graph to reduce the impact of so many transactions.

Node-cutting strategy: Given an account node that has only one transaction and is in the output, if the transaction has only one input and output, the account node and its connection edge will be cut and the diameter of the input node will be enlarged.

Only 8.6% of Bitcoin addresses are used more than once, according to statistical results of Kalodner et al. [6], yet this accounts for 51% of all input occurrences, which is the inspiration behind the suggested node-cutting strategy. We can assume that the main part of the Bitcoin transaction network is made up of just 8.6% of Bitcoin addresses acting as bridges between two nodes. However, single-used Bitcoin addresses are like unnecessary leaves and will be removed by the node-cutting approach.



(a) (b)

Fig. 1 Typical graphs of gambling addresses. (a) Heterogeneous; (b) isomorphic

After discussing the visualization scheme, we finally use ForceAtlas2 to create the layout of nodes.

3 User cases

3.1 Use case 1: pattern observation

Finding complex relationships among nodes in BTC-Shadow is straightforward because the layout algorithm places the nodes reasonably. For instance, in the largest connected network as depicted in Fig. 1 by BTC-Shadow, the gambling pattern, in which people typically view Bitcoins as bets, is apparent. However, due to the poor visualization strategies used by BlockChainVis, BiVA, and BitVis, these Characteristic patterns are not visible there.

3.2 Use case 2: subgraph analysis

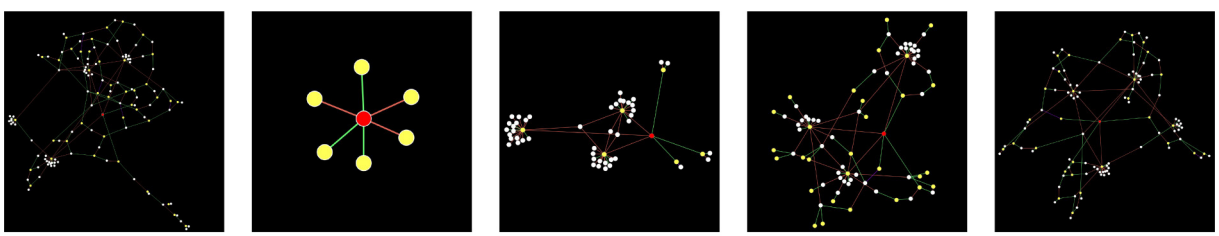
A subgraph analysis of the complete graph becomes a key approach to analyzing Bitcoin addresses because the subgraph minimizes the loss of the graph while focusing on one address. Two methods are not provided in other tools [3,5], to generate the subgraph in BTC-Shadow. The most practical one allows users to create an n-hop graph from any address. Another strong and competent method of BTC-Shadow is that it allows users to construct a customized filter to produce a subgraph that is more diverse and finely tuned compared to other tools [4]. Fig. 2 displays one instance. The address participating tumbler (i.e., mixer) is represented by the red nodes in the graphs. With a reduced n value, we can see that the entire graph is more comprehensible.

3.3 Use case 3: Bitcoin flow trace

The studies [3,5] employ arrow lines to depict the flow of bitcoin between neighbors. While Bitcoins from multiple transactions are pooled into one account, this representation is unable to determine where mixed Bitcoins originated. BTC-Shadow can search where output is used as the input in another transaction and where input is coming from to track Bitcoin flows. If a search result exists, the corresponding lines will be colored in a distinctive color of blue, as shown in Fig. 3.

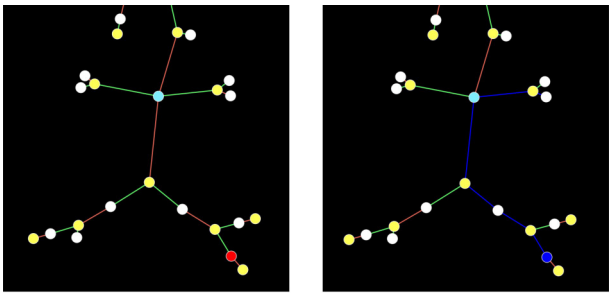
4 Conclusion

We present BTC-Shadow, a highly interactive analysis tool for Bitcoin data that allows for user customization. Users can interact with BTC-Shadow further on GitHub, where the tutorial and its executable version are available. We believe that this tool will be beneficial to a range of stakeholder groups, including law enforcement.



(a) (b) (c) (d) (e)

Fig. 2 n-hop graphs with different n value. (a) Entire graph; (b) $n = 1$; (c) $n = 2$; (d) $n = 3$; (e) $n = 4$



(a) (b)

Fig. 3 Bitcoin trace via UTXO. (a) Cyan node mixes two Bitcoin flows; (b) the Bitcoin flow of red node

Acknowledgements The research was financially supported by the CCF-NSFOCUS Kun-Peng Scientific Research Fund (No. CCF-NSFOCUS2021008), the Provincial Key Research and Development Program of Hubei (No. 2020BAB105), the National Natural Science Foundation of China (Grant No. 61972366), the Knowledge Innovation Program of Wuhan - Basic Research (No. 2022010801010197), the Foundation of Hubei Key Laboratory of Intelligent Geo-Information Processing (No. KLIGIP-2021B06), and the Opening Project of Nanchang Innovation Institute, Peking University (No. NCII2022A02). The work of K.-K. R. Choo was supported only by the Cloud Technology Endowed Professorship.

References

1. Tovanich N, Soulie N, Heulot N, Isenberg P. MiningVis: visual analytics of the Bitcoin mining economy. IEEE Transactions on Visualization and Computer Graphics, 2022, 28(1): 868–878
2. Xia J Z, Zhang Y H, Ye H, Wang Y, Jiang G, Zhao Y, Xie C, Kui X Y, Liao S H, Wang W P. SuPoolVisor: a visual analytics system for mining pool surveillance. Frontiers of Information Technology & Electronic Engineering, 2020, 21(4): 507–523
3. Bistarelli S, Santini F. Go with the -bitcoin- flow, with visual analytics. In: Proceedings of the 12th International Conference on Availability, Reliability and Security. 2017, 38
4. Sun Y, Xiong H, Yiu S M, Lam K Y. BitVis: an interactive visualization system for bitcoin accounts analysis. In: Proceedings of 2019 Crypto Valley Conference on Blockchain Technology (CVCBT). 2019, 21−25
5. Oggier F, Phetsouvanh S, Datta A. BiVA: Bitcoin network visualization & analysis. In: Proceedings of 2018 IEEE International Conference on Data Mining Workshops (ICDMW). 2018, 1469−1474
6. Kalodner H, Möser M, Lee K, Goldfeder S, Plattner M, Chator A, Narayanan A. BlockSci: design and applications of a blockchain analysis platform. In: Proceedings of the 29th USENIX Conference on Security Symposium. 2020, 153