

Analysis of Bitcoin Exchange Using Relationship of Transactions and Addresses

Seongho Hong*, Heeyoul Kim**

*Computer Science and Engineering, Kyonggi University, Korea

201511873@kyonggi.ac.kr, heeyoul.kim@kgu.ac.kr

Abstract— This bitcoin system is a system made to support transactions between users without the help of any financial institution. The Bitcoin Exchange is a service that enables exchanges of bitcoins mined by others with cash. To use this service, the miner should send bitcoins from his bitcoin wallet to the deposit address provided by the Exchange. The bitcoins delivered to the deposit address are automatically moved to one of the Exchange's bitcoin addresses by the Exchange system. Since what is known to the user is only the deposit address for the user, the user can only guess that the Exchange has taken the bitcoins from his deposit address when he/she sees the transfer of his/her bitcoins to another address. We will present a method to find out a few bitcoin addresses through preliminary work and find out the associated Exchange addresses using the characteristics of the deposit addresses and system addresses of the Exchange through transactions with the relevant addresses on the blockchain.

Keywords— Blockchain, Bitcoin, Exchange, Transaction, Analysis

I. INTRODUCTION

The bitcoin[1] system is a blockchain system made for the first time for electronic settlement of trust-based models. Through the consensus process called Proof of Work, the bitcoin system always makes common values and creates blocks in which new values can be recorded. The process of creating the blocks is called mining. The miner obtains a certain quantity of bitcoins as compensation for the mining. The quantity of bitcoin obtained when creating blocks decreases over time so that no compensation can be obtained by creating blocks from about 2040. The bitcoin system is designed so that the total number of created bitcoins becomes 21 million, and quantities beyond it cannot be created.

Over time, as the value of bitcoins has been recognized by some people so that bitcoins became the target of transactions[2], the possibility for other coins to be recognized for their values has arisen. People began to exchange bitcoins, which have begun to be recognized for their values has arisen. People began to exchange bitcoins, which have begun to be recognized for their value, with other coins, which were thought to be recognized for their values, and some people exchanged bitcoins with cash[3]. In 2010, the first Exchange, that is, a bitcoin market, was established to transact different coins[4].

To use the Exchange, the user must transfer their bitcoins from his/her existing bitcoin client wallet to the deposit address assigned by the Exchange and the bitcoins are delivered to the Exchange after several confirmations. Although this transfer process per se is simple, the process through which the bitcoins are handled thereafter inside the Exchange is not simple. The Exchange moves the bitcoins from the deposit addresses of many users to internal system addresses according to its internal policy and redistributes the bitcoins among system addresses for efficient management.

In general, the Exchange does not disclose the way in which bitcoins are handled internally and its financial status. Although the quantity of bitcoins held by the Exchange by finding out all addresses owned by the Exchange, the Exchange does not disclose its addresses to the outside. This paper presents a method to distinguish the addresses owned by the Exchange from transactions based on heuristic behaviors of the Exchange and expand new addresses using the relationships of the addresses to find out the size and the number of addresses of the Exchange.

II. RELATED WORKS

Because of anonymity, which is one of the core values of bitcoins, BitIodine[5] uses data outside the bitcoin blockchain such as the data crawled from the web to categorize bitcoin addresses. External data are those data that are generated voluntarily or involuntarily by the user and based on this information; addresses can be classified after clustering. As a result of the classification, the user becomes to obtain addresses classified by tags such as “miner”, “one-time”, “malware”, and so on. Although BitIodine is meaningful in that it is as an automated classification and analysis tool, according to experimenters, it is a tool that cannot be easily used in general environments. The work was done in a cloud environment with 26 instances and 68.4GB of memory and the memory is not optional. It is not easy to actually use BitIodine because its requirements exceed the requirement of general specifications.

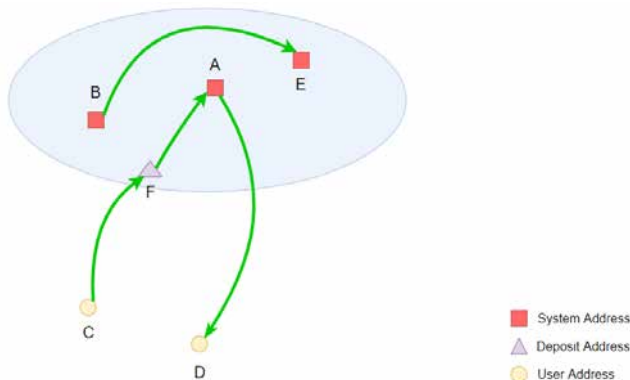
Methods to infer the relationships of bitcoin addresses through several simple constraints even without using any tool such as BitIodine were proposed. If multiple input addresses exist in one transaction, it can be inferred that they were owned by one user[6] under the assumption that they do not use multi-sig. When there are two addresses at the output of a

transaction and one of the addresses is an address found for the first time on the blockchain, if an unspent bitcoin delivered to the address is delivered again to the user at the input, it can be inferred that the two addresses were derived from one address.

The above methods were used to analyse the correlation between the trend of occurrence of Ransomware and the increase/decrease of bitcoin usage and arrest the Dread Pirate Roberts(DPR), and the methods were also proposed as methods to cluster bitcoin networks and partially reveal the anonymity of the targets.

III. HEURISTIC METHOD FOR ANALYSIS

In this study, transactions are analysed using the characteristics of the Exchange and diverse heuristics, and addresses related to the Exchange are found to find out those addresses of the Exchange that could not be easily found. Figure 1 shows the Exchange address group on the bitcoin network. There are two types of addresses in the Exchange: Deposit address D user's use of the Exchange use and system address S for Exchange operation. All addresses except these addresses are expressed as external address U. In Figure 1, system address can send bitcoins to other system address(from B to E) and user address(from A to D). deposit address, it can only receive bitcoin from user address(from C to F), and always send its rest of bitcoin to system address(from F to A). User address can receive bitcoin from system address when he withdraw it from exchange(from A to D).



[Figure 1. Bitcoin Network and Exchange Group]

D : a set of deposit addresses of the exchange

S : a set of system addresses of the exchange

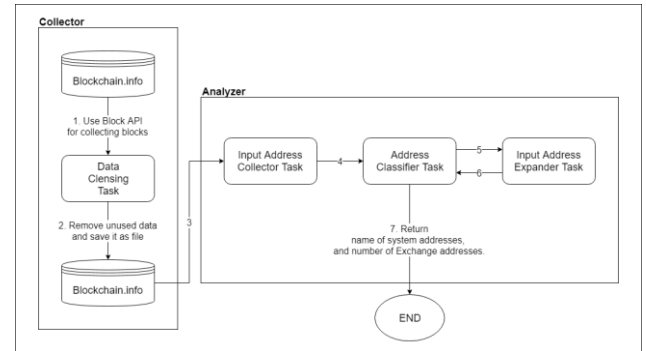
E : a set of addresses of the exchange ($E = D \cup S$)

I_T : a set of input addresses in transaction T

O_T : a set of output addresses in transaction T

For analysis, first, actual transactions are conducted through the Exchange to obtain initial system address S_a , and based on this, Exchange addresses can be searched through the following three steps. The individual steps have the roles

of expanding the Exchange addresses, classifying the addresses, and searching for new expandable addresses, respectively. The structure of the entire analysis system and the flow of each step are shown in Figure 2.



[Figure 2. Process of Bitcoin Analyzer]

A. Step 1

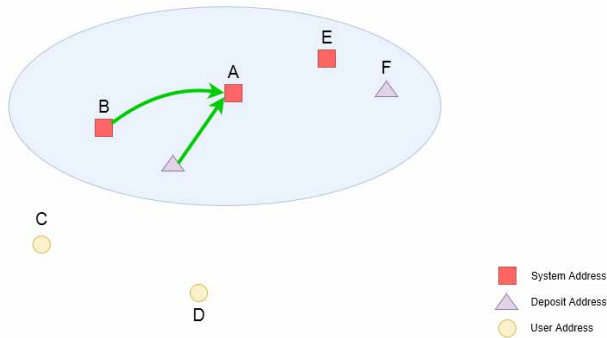
Although the Exchange creates deposit addresses by user and assigns them to users, the deposit addresses are managed by the Exchange and the Exchange has the private keys of the addresses. Therefore, heuristics regarding known multiple-input addresses can be used. That is, if there are previously known Exchange addresses among the input addresses of one transaction, it can be inferred that all addresses in the input of the transaction are Exchange addresses. Therefore, new Exchange addresses can be found from initial address S_a .

$$\exists (x \in I_T) \wedge (x \in E) \rightarrow I_T \subset E$$

B. Step 2

In step 2, the Exchange addresses obtained through Step 1 are classified to see if they belong to system address S . If Exchange addresses exist among the input addresses in a transaction and the Exchange addresses also exist among the output addresses, it can be judged that the Exchange addresses in the output belong to system address S . This is due to two premises.

- 1) **Non-Benefit:** External users will not deposit money at a system address where no profit will be accrued when the money has been deposited.
- 2) **Preventing User's Doubt:** Exchange users do not have a private key for the deposit address, but they can check the details of their depositing wallets through a service such as blockchain.info. The Exchange will ensure that no transaction detail that can confuse users would occur.



[Figure 3. Transaction Network on Exchange]

Based on the above premises, in the network, there should be only transactions in the forms of from deposit address to system address and from system address to system address. Step 2 can be expressed as follows.

$$(\exists(x \in I_T) \wedge (x \in E)) \wedge (\exists(y \in O_T) \wedge (y \in E)) \rightarrow y \in S$$

C. Step 3

Step 3 is a process of finding new addresses based on the classified set information. Applying Fig. 3, if a transaction has Exchange system address S as an output, it can be judged that all addresses in the input are Exchange addresses. Based on the foregoing, among input addresses, new ones that have not been classified into Exchange addresses already can be added to Exchange addresses. However, separately from the premises, since the administrator might have transmitted bitcoins for the first system address of the Exchange for setting initial bitcoins for Exchange operation, in Step 3, addresses are explored after an interval of at least about 144 blocks(= one day) after the relevant system addresses was found.

$$\exists(x \in O_T) \wedge (x \in S) \rightarrow I_T \subset E$$

IV. HEURISTIC METHOD FOR ANALYSIS

An experiment to apply the proposed analysis method to the Gopax Exchange[7] was conducted. The block data for analysis were obtained by using the block API provided by blockchain.com[8], and only the block height, hash, and transaction information were extracted and processed into the form of a file. The new addresses were stored and classified using the processed bitcoin blockchain data and a data analyser in which steps 1 ~ 3 were implemented, and the results obtained in each step were stored in the form of logs. The results obtained from previous cycles and those obtained from the current cycle are compared and if no change is found, it is judged that all addresses have been found. Both the data collector and analyser were implemented with Python, and in the blockchain, transactions between the 470,000th block when the operation of Gopax began and 540,000th block were analysed. The system address

'367f4YWz1VCFaqBqwbTrzwi2b1h2U3w1AF' was used as the initial address for the search of Exchange addresses. The results obtained through the analyser are as follows.

TABLE 1. ANALYZING RESULT

Step	# of System Address	# of Exchange Address
1-1	1	1867
1-2	1	2236
1-3	1	2657
1-4	1	2746
1-5	1	2755
1-6	1	2757
2	3	2757
3	3	8044
2	32	8044

In this table, first step were repeated continuously. These same steps are represented by slash and numbers to distinguish it from other steps. Through each step, we found the associated addresses at one initial bitcoin address according to the criteria. We could find diverse addresses by connecting the individual methods of the steps. As a result, we were able to figure out 8044 bitcoin addresses possessed by the Exchange had and could see that 32 out of them were internal system addresses used by the Exchange.

V. CONCLUSIONS

Through the abovementioned method, the quantity of addresses owned by the Exchange and the roles of some of those addresses could be figured out. For more accurate analysis, the classification between deposit addresses and system address should be made clearer, and those system addresses that only send money to external addresses could be found from among the system addresses. The current relationship analysis cannot classify between system addresses dedicated to money transfer and money transfers between systems. If the foregoing can be grasped, the structures and roles can be more concrete and the next goal will be presenting measures to solve related problems.

ACKNOWLEDGMENT

This research was supported by Ministry of Science ICT Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2018R1C1B6002903).

REFERENCES

- [1] Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2010.
- [2] Morten Brandvold, Price Discovery on Bitcoin Exchanges, 2015.
- [3] Stephanie Lo, Bitcoin as Money, 2014.
- [4] Matthias Lischke, Analyzing the Bitcoin Network, The First Four Years, 2016.
- [5] Michele Spagnuolo, Bitlodge: Extracting Intelligence from the Bitcoin Network, 2014.
- [6] Fergal Reid, An Analysis of Anonymity in the Bitcoin System, 2012.
- [7] (2018) Gopax website. [Online]. Available: <http://www.gopax.co.kr/>
- [8] (2018) Blockchain Info website. [Online]. Available: <http://www.blockchain.com/explorer/>



Seongho Hong is in the undergraduate course of Computer Science and Engineering Department, Kyonggi University. His major research interests include blockchain.



Heeyoul Kim received the B.E degree in Computer Science from KAIST, Korea, in 2000, the M.S. degree in Computer Science from KAIST in 2002, and the Ph.D. degree in computer science from KAIST in 2007. From 2007 to 2008, with the Samsung Electronics as a senior engineer. Since 2009 he has been a faculty member of Department of Computer Science at Kyonggi University. His major research interests include cryptography, security, and blockchain.