



Detecting Roles of Money Laundering in Bitcoin Mixing Transactions: A Goal Modeling and Mining Framework

Mingdong Liu^{1*}, Hu Chen² and Jiaqi Yan^{3*}

¹School of Economics and Management, Southeast University, Nanjing, China, ²School of Software Engineering, South China University of Technology, Guangzhou, China, ³School of Information Management, Nanjing University, Nanjing, China

Cryptocurrency has become a new venue for money laundering. Bitcoin mixing services deliberately obfuscate the relationship between senders and recipients, making it difficult to trace suspicious money flow. We believe that the key to demystifying the bitcoin mixing services is to discover agents' roles in the money laundering process. We propose a goaloriented approach to modeling, discovering, and analyzing different types of roles in the agent-based business process of the bitcoin mixing scenario using historical bitcoin transaction data. It adopts the agents' goal perspective to study the roles in the bitcoin money laundering process. Moreover, it provides a foundation to discover real-world agents' roles in bitcoin money laundering scenarios.

OPEN ACCESS

Edited by:

Xiao Fan Liu, City University of Hong Kong, SAR China

Reviewed by:

Daning Hu, Southern University of Science and Technology, China Hong-Liang Sun, Nanjing University of Finance and Economics, China

*Correspondence:

Mingdong Liu 230179735@seu.edu.cn Jiaqi Yan jiaqiyan@nju.edu.cn

Specialty section:

This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: 08 February 2021 Accepted: 07 May 2021 Published: 06 July 2021

Citation:

Liu M, Chen H and Yan J (2021) Detecting Roles of Money Laundering in Bitcoin Mixing Transactions: A Goal Modeling and Mining Framework. Front. Phys. 9:665399. doi: 10.3389/fphy.2021.665399 Keywords: goal modeling, money laundering, bitcoin mixing transactions, data analysis, agents' roles

INTRODUCTION

Financial crimes not only directly disturb the national financial order and affect social stability but also occur with other crimes to provide financial support for various types of organized crimes. Money laundering is a financial criminal activity, which mainly refers to the processing of illegal income by various means to cover up and conceal its source and nature. It not only damages the security of the financial system and the reputation of financial institutions but also destroys the normal economic order and social stability of the country. Since money laundering is such a harmful activity, anti-money laundering is, therefore, a worthwhile endeavor.

Money laundering is a complex activity involving many entities and relationships. With the development of the Internet, money launderers utilize advanced technology and multiple channels to cover up their criminal behaviors through numerous transactions. Cryptocurrency has become a new venue for money laundering. The simplest form of bitcoin money laundering is that the bitcoin transactions are made under pseudonyms. Criminals use pseudonymous bitcoin addresses to hide the illegal source of funds. However, as studies have revealed that the pseudonyms of bitcoin addresses can be broken by aggregating addresses into clusters with identified users [1], more and more third-party bitcoin mixing services emerged to provide additional anonymity [2]. It is reported [3] that at least 4,836 bitcoins stolen by hacking Binance were laundered through the crypto mixing service.

The emergence of bitcoin mixing services makes it difficult to trace suspicious money flow as they deliberately obfuscate the relationship between senders and recipients [4]. However, there are limited existing studies investigating the bitcoin mixing services. The difficulties lie in detecting different roles of bitcoin addresses as there are an enormous number of bitcoin addresses involved. One of the earliest studies [5] on mixing services revealed that they bundle a large number of small transactions

1



into a small number of large transactions to create all outgoing transactions, hiding the connections between input addresses and output addresses.

In this article, we propose that the key to demystifying bitcoin mixing services is to discover agents' roles in the money laundering process. As money laundering is usually committed by collusive money launders, multiple agents are involved in the process. Different agents have different roles that perform different tasks in the bitcoin mixing process to achieve the ultimate goal of money laundering. Identifying the agents' roles in the bitcoin mixing process will be helpful to understand the context of bitcoin money laundering. In this paper, we propose a goal-oriented approach to modeling, discovering, and analyzing different types of roles in the agent-based business process of the money laundering scenario using historical transaction data from bitcoin mixing services. To the best of our knowledge, this paper is the first to apply goal-oriented modeling to represent the agents in bitcoin mixing transactions. It provides a foundation to understand the role and task assignment at cryptocurrency transactions in money laundering scenarios.

The rest of this paper is organized as follows. *Related Work* reviews *related works*. *Goal Modeling and Mining in Money Laundering* formalizes the problem and presents the framework. *Case Study* provides a case study and presents algorithms for goal mining in the money laundering processes. This article concludes with contributions and future research plans in *Conclusion*.

RELATED WORK

Cryptocurrency Transaction Analysis

A cryptocurrency transaction is a basic unit describing cryptocurrency flow from input to output addresses. Every input address in a cryptocurrency transaction is a reference to an unspent transaction output (UTXO), which is an output address in a previous transaction that has not been referenced in other transactions. In the bitcoin system, addresses are the basic identities that hold virtual values, which can be generated offline to a public key using the bitcoin's customized hash function. **Figure 1** presents a basic example of UTXOs. It is composed of three transactions. In Transaction 1, address A is the input with 10 BTC and B, C, and D are the output addresses. All outputs in Transaction 1 are UTXOs before they are referenced by Transactions 2 and 3.

As the complete transaction history is publicly available, the transparency of cryptocurrency transactions enables statistical analysis and graphical visualization techniques. Some scholars produced an organized review of major works in cryptocurrency transaction analysis. For example, Chen et al. [6] reviewed the status, trends, and challenges in blockchain data analysis and summarized seven typical research issues of cryptocurrency transaction analysis into entity recognition, privacy identification, network risk parsing, network visualization and portrait, analysis of cryptocurrency market, etc. Liu et al. [7] surveyed knowledge discovery in cryptocurrency transactions and summarized the existing research that uses data mining techniques into three aspects, including transaction tracing and blockchain address linking, the analysis of collective user behaviors, and the study of individual user behaviors.

Both reviews have identified many studies on transaction tracing, showing that the mechanism of the pseudonymity of cryptocurrency addresses used in transactions can be broken by entity recognition (or blockchain address linking) and privacy identification techniques. For example, to identify money laundering in bitcoin transactions, Hu et al. [8] proposed four types of classifiers based on the graph features that appeared on the transaction graph, including immediate neighbors, deep walk embedding, node2vc embedding, and decision tree-based. Once addresses are identified, money flows can be immediately revealed, leading to no anonymity in the bitcoin system.

Because the original design of bitcoin transactions is easy to trace, several solutions have been proposed to improve its anonymity. One typical solution is a mixing service, which is widely used in underground markets like the Silk Road to facilitate money laundering. Mixing services aim to solve cryptocurrencies' traceability issues by merging irrelevant transactions with methods including swapping and conjoining. Only a few previous works have been carried out to demystify mixing services. For instance, in one of the earliest studies on mixing services [5], a simple graph analysis was carried out based on data collected from experiments of selected mixing services, and alternative anti-money laundering strategies were sketched to account for imperfect knowledge of true identities. Although an essential anti-money laundering strategy is not provided, Seo et al. [2] mentioned that money laundering conducted in the underground market can be detected using a bitcoin mixing service. These explorations revealed the importance of understanding bitcoin mixing services.

Data Mining in Money Laundering

Money laundering is a complex, dynamic, and distributed process, which is often linked to terrorism, drug and arms trafficking, and exploitation of human beings. Detecting money laundering is notoriously difficult, and one promising method is data mining [9]. Rohit and Patel [10] reviewed detection of suspicious transactions in anti-money laundering using a data mining framework and classified the literature into the rule-based approach, clustering-based approach, classification-based approach.



Generally, data mining in money laundering consists of the rule-based classification and machine learning approaches. In a rule-based approach, ontologies or other forms of rules are adopted to classify suspicious transactions. For instance, Rajput et al. [11] proposed an ontology-based expert system for suspicious transaction detection. The ontology consists of domain knowledge and a set of Semantic Web rules, and the native reasoning support in the ontology was used to deduce new knowledge from the predefined rules about suspicious transactions. A Secure Intelligent Framework for Anti-Money Laundering was presented to make use of an intelligent formalism by using ontologies and rule-based planning [12]. Bayesian approaches were adopted to assign a risk score to money laundering-related behavior [13]. It was designed based on rules suggested by the State Bank of Pakistan in its 2008 regulations to declare a transaction as suspicious.

Machine learning algorithms were also applied to group or classify the data, so as to predict suspicious money laundering transactions. Chen et al. [14] provided a comprehensive survey of machine learning algorithms and methods applied to detect suspicious transactions, including typologies, link analysis, behavioral modeling, risk scoring, anomaly detection, and geographic capability. A support vector machine-based classification system was developed to handle large amounts of data and take the place of traditional predefined-rule suspicious transaction data-filtering systems [15]. However, the limitation of the machine learning approach lies in its data dependence, with sometimes limited adaptability and scalability. As stated in [6], the model requirement of historical data makes it difficult to identify illicit operations performed by newcomers.

GOAL MODELING AND MINING IN MONEY LAUNDERING

In this section, we propose to model agents' roles in the money laundering process with goal-oriented modeling techniques. We adopt the goal modeling and mining approach [16] and assume that different agents have different roles, which perform different tasks in the process to achieve the ultimate goal of money laundering. As shown in **Figure 2**, the goal modeling and mining framework consists of three phases. In the data collection phase, we will collect not only cryptocurrency transaction data but also the domain data about the address entity information. The model discovery phase consists of the address miner, role miner, and process miner. The address miner will collect all related address information. The role miner will detect the goals of the agent who owns the addresses. The process miner will present the money laundering process.

In the blockchain analysis phase, we first analyze the goal of money laundering, which can be decomposed into three



sub-objectives, **namely**, **placement**, **layering**, **and integration**. Placement aims to cut off the connections between illegal funds and upstream crimes of money laundering. It can be realized by the task of introducing illegal funds into the financial system. This task can be decomposed into sub-tasks like depositing money or remitting cash. Placement always involves breaking down the original large amount of funds into a lot of small ones. So, many soldiers will be employed to execute the tasks to introduce multiple illegal funds into the financial system without suspicion.

Layering is transferring the funds among different accounts or institutions so that the initial source of funds will be difficult to track. Layering can be realized by the task of obscuring the sources of the money. This can be decomposed into sub-tasks like transferring money or purchasing insurance from different institutions. Layering means that funds have entered the financial system for circulation. Soldiers may continue to do a lot of basic tasks, while communicators' main role is to transfer the funds.

Integration means to integrate the funds into the legitimate economy. Integration can be realized by the task of legalizing illegal funds. This task can be decomposed into sub-tasks like transferring money or purchasing insurance from different institutions, transferring overseas, withdrawing cash, or investing. Finally, the legalized funds will be possessed by the money laundering organizers. Communicators and organizers will work together to complete the task.

In the money laundering process, three roles are involved, as follows:

1) Organizer: Organizers are the core of the organization. As described in Figure 3, the organizer's goal is to organize money laundering, which can be decomposed to allocate resources and organizing paths. The goal of allocate resources can be decomposed into allocate funds and allocate people, which can be realized by the tasks of assign funds to agents. The goal of organizing paths can be realized by the task of assign funds to agents. The goal of organizing paths can be decomposed into communicate with customers and plan links.

Organizers organize the money laundering process collaborating with communicators and soldiers. Particularly, as shown in **Figure 4**, the organizer depends on communicators to communicate with customers while depending on soldiers to assign funds to agents. Communicators execute tasks from the organizer and employ soldiers to introduce illegal assets into the financial system.

- 2) Communicator: Communicators are at the middle level of the organization. The communicator's goal is to transmit information in the money laundering activity. The goal of transmitting information can be realized by the task of communicate with customers, which can be decomposed to inputs and outputs for customers.
- 3) Soldier: The soldier's goal is to become agents with chips. In Figure 5, soldiers are employed by the organizer to deal with some basic tasks. They are not related to the core organization, while they are the key to facilitate the flow of illegal funds into the financial system. The goal of agents with chips can be realized by the task of split funds and assign chips.

CASE STUDY

In this section, we present a case study to demonstrate the proposed framework. We investigated a popular mixing service. In this bitcoin mixing service, the smallest unit of deposit is 0.001 BTC. As shown in **Figure 5**, the deposit is divided into different chips, which are $2^k \times 0.001$ BTC. Users can send these chips to one or more different withdrawal addresses, as shown in **Figure 5**. In our experiment, we sent all five chips of 0.031 BTC to one withdrawal address.

We tracked withdrawal transactions through the btc.com website. In **Figure 6**, the box represents the address and the oval box represents the transaction (the text is the transaction time). The value on the line between the address and the transaction represents







the bitcoin value that the address entered or output in this transaction. The address in green is our export address. We noticed that there are five exit addresses in the money laundering network, which are consistent with the chip division. These five addresses seem to correspond to five address pools of different amounts. We selected three other addresses in the address pool





with a bitcoin amount of 0.002 and found that these transactions were also combined to produce output from different address pools.

Deposit transactions are shown in **Figure** 7. In the figure, purple denotes the source address of the experiment and yellow-green is the entry address provided by the bitcoin mixing service. It can be seen that the money laundering entry address in this experiment is directly used to generate an address pool of 0.008 BTC.

As shown in **Figure 8**, we find that there are three categories of addresses: *entry addresses (communicator), exit addresses (communicator),* and *kernel addresses (soldier).* When a laundering request is issued, it will generate an *entry address* to receive the bitcoin from a user. After a while, some entry addresses and *kernel addresses* are combined as the inputs of one mixing transaction to generate some *exit addresses.* An important feature of the exit addresses is that their amount is $2^k \times 0.001$ BTC. When it decides to send $Y \times 0.001$ BTC to one laundering output address, some *exit addresses* are selected according to *Y*. For example, if $Y = 0.031 = (16 + 8 + 4 + 2 + 1) \times 0.001$ BTC, then five addresses holding 0.016, 0.008, 0.004, 0.002, and 0.001 BTC are selected from its pools. These *exit addresses* are treated as inputs for a withdrawal transaction to send to the output address specified by the user.

Definitions for Identifying the Mixing Transactions

The transaction T with m input addresses (a_1^I, \ldots, a_m^I) and n output addresses (a_1^O, \ldots, a_n^O) is described as follows:

$$(a_1^I, c_1^I), \dots, (a_m^I, c_m^I) \to (a_1^O, c_1^O), \dots, (a_n^I, c_n^O),$$
 (1)

where c_i^I and c_j^O $(1 \le i \le m, 1 \le j \le n)$ are the amount of input address a_i^I and output address a_j^O , respectively. Obviously, $\sum_{i=1}^m c_i^I \ge \sum_{j=1}^n c_i^O$. And the fee of T is $\sum_{i=1}^m c_i^I - \sum_{j=1}^n c_i^O$.

For a transaction *T*, the set of input addresses is defined as $A^{I}(T) = \{a_{1}^{I}, \ldots, a_{m}^{I}\}$ and the set of output addresses is defined as $A^{O}(T) = \{a_{1}^{O}, \ldots, a_{n}^{O}\}.$

Actually, we find two types of mixing transactions to generate the exit addresses, defined as follows:

[Definition 1] A transaction *T* with *n* outputs is type I transaction, if $c_i^O = 2^k \times 0.001$ BTC $(1 \le j \le n)$ and n > 1.

[Definition 2] A transaction T with n outputs is type II transaction, if $c_j^{O} = 2^k \times 0.001$ BTC ($1 \le j \le n-1$), $c_n^{O} \ne 2^l \times 0.001$ BTC, and n>2.

The withdrawal transactions are based on the *exit addresses*, described as follows:

[Definition 3] A transaction *T* with *m* inputs and one output is a *withdrawal transaction*, if

1. $c_i^{I} \neq c_j^{I}$, $i \neq j$, $1 \le i, j \le m$; 2. $c_i^{I} = 2^k \times 0.001$ BTC, $1 \le i \le m$ or m-1.

We design **Algorithm 1** to find the mixing transactions. There are three stages in the algorithm. The first stage (step 1) is to find all type I or II transactions T_1 . In stage 2 (steps 2–10), we try to find the withdrawal transactions T_2 based on T_1 . In the last stage, the transactions in T_1 are selected as the mixing ones according to T_2 .



Algorithm 1

Input: T, bitcoin transaction set from t_0 to t_1 Output: T_C , detected mixing transactions 1. Finding transaction set T_1 from T, where $t \in T_1$ is type I or type II transaction. 2. $T_2 = \emptyset$ 3. for each $T \in T - T_1$ if T is withdrawal transaction then 4. if $c_i^J = 2^k \times 0.001$ BTC and $a_i \in A^O(T_i)$ and $a_i \in A^I(T)$, $T_1 \in T_1$, $1 \le i \le m$ then 5. 6. $T_2 = T_2 \cup \{T\}$ 7. endif 8. endif 9. endfor 10. T_C=Ø 11. for each $T \in T_1$ if all outputs of T are the inputs of some transactions in T_2 12 13. $T_C = T_C \cup \{T\}$ 14. endif 15. endfor

Kernel addresses appear in different mixing transactions as the inputs. We use **Algorithm 2** to find them.

Algorithm 2

```
Input: T_C, detected mixing transactions
Ouput: A_k, kernel addresses
1. A_k = \emptyset
2. for each T \in T_C
3. for each T_s \in T_C, and T_s \neq T
4. A_k = A_k \cup (A^l(T) \cap A^l(T_s))
```

In the 2020 BTC transactions, we find 4,689 type I transactions and 3,124 type II transactions. With **Algorithm 1**, we determine that 2,687 are mixing transactions with 47,433 associated withdrawal

transactions. The number of mixing transactions in each month of 2020 is shown in Figure 9.

With **Algorithm 2**, we find 2,451 kernel addresses, of which 2,143 (87%) are in one wallet [0005190b7a] according to walletexplorer. com. This proves that these addresses we discovered belong to an organization that has not been revealed before and can prove to be controlled by the mixing service provider.

We tracked money laundering transactions and discovered the transaction structure. From historical bitcoin transactions, we found a large number of mixing transactions with significant characteristics and kernel addresses related to the bitcoin mixing service. We can estimate the scale of money laundering based on such role analysis.

CONCLUSION

In this paper, we propose that the key to demystifying bitcoin mixing services is to discover agents' roles in the money laundering process and present a goal-oriented modeling framework to model different roles in the money laundering process. The framework consists of data collection, model discovery, and blockchain analysis. With this framework, the three roles of the organizer, soldier, and communicator are analyzed in the money laundering process of placement, layering, and integration.

We then apply the proposed framework to investigate a popular bitcoin mixing service. Specifically, we identify two types of mixing transactions to generate the exit addresses. We propose two algorithms to analyze the roles of the soldier and communicator in the money laundering process. With the identified roles, we can estimate the scale of money laundering in the bitcoin mixing service.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

REFERENCES

- Reid F, and Harrigan M. An Analysis of Anonymity in the Bitcoin System. 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (2011). p. 1318–26. doi:10.1109/PASSAT/SocialCom.2011.79 An Analysis of Anonymity in the Bitcoin System.
- Seo J, Park M, Oh H, and Lee K. "Money Laundering in the Bitcoin Network: Perspective of Mixing Services," in 2018 International Conference on Information and Communication Technology Convergence (ICTC) (2018) 1403–1405. Available at: https://www.researchgate.net/publication/ 329491846_Money_Laundering_in_the_Bitcoin_Network_Perspective_ of_Mixing_Services. doi:10.1109/ictc.2018.8539548
- Maria O. Chipmixer Used To Launder BTC from Recent Binance Hack. *Cryptocurrencynews* (2019). Available at: https://cryptocurrencynews.com/ chipmixer-binance-hack-btc-laundering/ (Accessed May 30, 2021).
- 4. Wu L, Hu Y, Zhou Y, Wang H, Luo X, Wang Z, et al. *Towards Understanding* and Demystifying Bitcoin Mixing Services (2020).
- Möser M, Böhme R, and Breuker D. An Inquiry into Money Laundering Tools in the Bitcoin Ecosystem. 2013 APWG ECrime Researchers Summit (2013) 1–14. doi:10.1109/eCRS.2013.6805780
- Chen Z, Van Khoa LD, Teoh EN, Nazir A, Karuppiah EK, and Lam KS. Machine Learning Techniques for Anti-money Laundering (AML) Solutions in Suspicious Transaction Detection: A Review. *Knowl Inf Syst* (2018) 57: 245–285. doi:10.1007/s10115-017-1144-z
- Liu XF, Jiang X-J, Liu S-H, and Tse CK. Knowledge Discovery in Cryptocurrency Transactions: A Survey. *IEEE Access* (2021) 9:37229–37254. doi:10.1109/ACCESS.2021.3062652
- Hu Y, Seneviratne S, Thilakarathna K, Fukuda K, and Seneviratne A. Characterizing and Detecting Money Laundering Activities on the Bitcoin Network (2019). ArXiv:1912.12060 [Cs]. Available at: http://arxiv.org/abs/ 1912.12060 (Accessed May 30, 2021).
- 9. Watkins RC, Reynolds KM, Demara R, Georgiopoulos M, Gonzalez A, and Eaglin R. Tracking Dirty Proceeds: Exploring Data Mining Technologies as

AUTHOR CONTRIBUTIONS

ML: Software, Investigation, Case Implementation, Validation, Methodology, Conceptualization. HC: Validation, Software, Data Curation, Case Implementation, Formal analysis, JY: Conceptualization, Editing, Visualization, Review, Supervision.

Tools to Investigate Money Laundering. *Police Pract Res* (2003) 4(2):163–78. doi:10.1080/15614260308020

- Rohit KD, and Patel DB. Review on Detection of Suspicious Transaction in Antimoney Laundering Using Data Mining Framework. *Int J Innovative Res Sci Technol* (2015) 1(8):129–133.
- Rajput Q, Khan NS, Larik A, and Haider S. Ontology Based Expert-System for Suspicious Transactions Detection. *Comp Inf Sci* (2014) 7(1):103–114. doi:10. 5539/cis.v7n1p103
- Sobh TS. An Intelligent and Secure Framework for Anti-money Laundering. J Appl Security Res (2020) 15(4):517–46. doi:10.1080/19361610.2020. 1812994
- Khan NS, Larik AS, Rajput Q, and Haider S. A BAYESIAN APPROACH FOR SUSPICIOUS FINANCIAL ACTIVITY REPORTING. Int J Comput Appl (2015) 35(4):181–7.
- Chen W, Zheng Z, Fronzetti Colladon A, and Remondi E. Blockchain Data Analysis: A Review of Status, Trends and Challenges Using Social Network Analysis to Prevent Money Laundering. J Comp Res Development Expert Syst Appl (2017) 5567(9):185349–7058. doi:10.1016/j.eswa.2016.09.029
- Tang J, and Yin J. Developing an Intelligent Data Discriminating System of Anti-money Laundering Based on SVM. Machine Learning And Cybernetics Proceedings of 2005 International Conference on (2005).
- Yan J, Hu D, Liao S, and Wang H. Mining Agents' Goals in Agent-Oriented Business Processes. ACM Trans Manag Inf Syst (2015) 5(4):20–120.22. doi:10. 1145/2629448

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu, Chen and Yan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.