Tracing Bitcoins Across Privacy Enhancing Technologies

Master-Thesis von Andreas Rothenhäuser aus Miltenberg
Tag der Einreichung:

1. Gutachten: Prof. Dr. Stefan Katzenbeisser
2. Gutachten: Spyros Boukoros, Nikolaos Alexopoulos

TECHNISCHE
UNIVERSITÄT
DARMSTADT

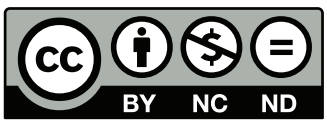
Computer Science
Security Engineering Group

Tracing Bitcoins Across Privacy Enhancing Technologies

Vorgelegte Master-Thesis von Andreas Rothenhäuser aus Miltenberg

1. Gutachten: Prof. Dr. Stefan Katzenbeisser
2. Gutachten: Spyros Boukoros, Nikolaos Alexopoulos

Tag der Einreichung:



Erklärung zur Master-Thesis

Hiermit versichere ich, die vorliegende Master-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 13.11 .2018

(Andreas Rothenhäuser)

Abstract

Test

Contents

List of Figures

List of Tables

1 Introduction

The importance of looking at money flows in criminal prosecution was already known back in 80 BC, when Marcus Tullius Cicero phrased the question "Cui bono?" - "who benefits?" in his speech "Pro Roscio Amerino" [1]. Since then this idea has evolved to the wide field of financial investigation. Today following the money trail to identify suspects, find backers or explore the hierarchy of a criminal organization is considered a basic task by law enforcement authorities around the world. The ongoing Trump-Russia investigation of Robert Mueller is just one example of how the money trail can convict criminals of all kinds [1, 2].

In 1994 the Internal Revenue Service (IRS) published a whole book dedicated to the art of financial investigation. The book was written as a guide for investigators, summarizing techniques that can be applied to almost any criminal case. Not only does it showcase how certain financial information can be interpreted and leveraged to build a successful case against a suspect, but it also lists potential sources of such information. Financial institutions usually enforce Know Your Customer (KYC) policies in order to meet the Anti-Money Laundering (AML)/Combating the Financing of Terrorism (CFT) regulations of their home country or the countries they do business in. That means they collect personal information, like names, addresses or solvency about every customer and business partner they engage with. Furthermore they keep track on the financial transactions they do on behalf of their customers in bank statements. This is why the authors of the book note that financial institutions "are probably the single most important source of information available to the financial investigator" [3].

The recent rise of the so called cryptocurrencies like Bitcoin poses a major problem to law enforcement. In 2012 the Federal Bureau of Investigation (FBI) assesses, that "law enforcement faces difficulties detecting suspicious activity, identifying users, and obtaining transaction records" [4]. Financial institutions like the European Central Bank (ECB) follow along by stating problems with the enforcement of AML/CFT regulations [5]. Unlike traditional financial systems the Bitcoin protocol is designed to work in a decentralized and anonymous manner without any kind of financial institution to manage it. With the absence of a centralized institution collecting all the valuable information, investigators can not just walk in somewhere with a search warrant and retrieve it as needed. Thus many of the traditional techniques of financial investigation as described in the aforementioned book can not be applied in the world of Bitcoin & Co. Law enforcement agencies are confronted with a whole new challenge that demands for new investigative techniques.

In this paper I propose a scheme that enables law enforcement agencies to trace money flows in the Bitcoin realm. Opposing many of the previous works on blockchain analysis my goal is not to deanonymize *most of the Bitcoin users* or *the average Bitcoin user*, but full disclosure of the actions of just a few individuals within a predefined section of the blockchain. The problem will be solved in two stages. Stage one maps Bitcoin addresses to their corresponding users, utilizing feature vectors known from stylometry. For my feature set I borrow ideas from previous work in the area of Bitcoin tracing and deanonymization and supplement them with new features. The resulting feature vectors are input to a Hierarchical Agglomerative Clustering (HAC) algorithm that clusters Bitcoin addresses according to their respective owners. Once individual

[1] `http://www.thelatinlibrary.com/cicero/sex.rosc.shtml`

users are successfully identified the financial links between them can be trivially looked up on the blockchain in stage two by traversing the transaction graph. Unlike other state of the art approaches of Bitcoin address clustering this scheme provides robust results even in the presence of privacy enhancing technologies like Bitcoin mixing services.
The paper will be organized as follows: In chapter 2 I will discuss previous work, that has been done in the field of blockchain analysis and the assesment of anonymity properties in the bitcoin ecosystem. In chapter 3 I will briefly introduce some background aspects like the workings of blockchains and the Bitcoin protocol, special demands in the context of criminal investigations and some basic techniques of athorship attribution. In chapter 4 I will develop the precise problem statement of this thesis and explain the methodology of my approach. In chapter 5 I will describe the experiments I set up to assess the performance of my approach and provide the evaluation results.

2 Related Work

A lot of work has already been done regarding the anonymity properties of Bitcoins. Most of the papers that have been written on this topic can be roughly divided into three different categories based on their perspective and approach. The first category engages in the so-called blockchain analysis. As the name implies, blockchain analysis tries to leverage the information contained in the Bitcoin blockchain itself. Opposing to that, the second category applies techniques of network analysis. These approaches draw information from the analysis of the Bitcoin Peer-to-Peer (P2P) network traffic. The last category of papers deals with a variety of privacy enhancing technologies that have been proposed to improve the Bitcoin protocol over the years.

2.1 Blockchain Analysis

A very prominent approach based on blockchain analysis was already mentioned by the inventor of Bitcoin [6] and later used by Reid and Harrigan in [7]. It utilizes the fact that Bitcoins, similar to real coins, need to be combined with other coins, if the amount that needs to be payed exceeds the value of any single coin the payee possesses. This combination of Bitcoins is done by creating so called multi-input transactions which have two or more coins as an input and at least one output. [7] as well as others ([8, 9]) apply Bitcoin address clustering based on this observation in order to reduce the anonymity set of the Bitcoin system. The authors conclude, that multi-input transactions testify a common owner of all inputs of the transaction and thus map the corresponding Bitcoin addresses to the same superordinate entity. If an attacker now manages to identify one single Bitcoin address, the other addresses belonging to the same entity are compromised as well.
Another approach building on top of the aforementioned multi-input heuristic is to identify so called change addresses. This heuristic again relies on an inherent property of the Bitcoin protocol. Since unspent Bitcoins need to be spent in total or not at all, a common use case is to add a special output to the transaction that receives the change (the money you get back, when you did not pay the exact amount). If the change address of a transaction is known, it can be mapped to the same entity as the inputs, reducing the anonymity set of the whole system even further. The change address heuristic was first explored in [9]. Lateron it was refined in [8] to account for changes in the usage of Bitcoins.
Although both heuristics might have yielded good results in the early days of Bitcoins they are not suited very well to the problem I want to solve. The multi-input heuristic is based on the assumption, that those transactions can only be issued by one single entity, whereas the change-address heursitic assumes, that each transaction necessarily has one change address. Back then these assumptions might have been reasonable, but the usage of Bitcoins has gone a long way since. Big mining pools regularly issue transactions with a lot of outputs, containing no change to reward their miners, thus breaking the assumption of the change-address heuristic. In addition there are services that implement CoinJoin. The idea is basically to merge transactions of different users into a single, shared one in order to improve privacy (for more information see 2]. The resulting transactions have inputs that belong to different users, thus breaking the assumption of the multi-input heuristic. Both approaches exploit a common practice and still hold true in many cases today, but the number of cases to prove them wrong is growing quickly.

2.2 Network Analysis

The approaches based on network analysis are a bit more invasive, since they require an attacker to participate in the Bitcoin P2P network other than in the originally intended way. Opposing the first category, the target is not to reduce the anonymity set, but to link Bitcoin addresses directly to IP addresses. In [10] the authors describe an apporoach based on detection of unusual relay patterns in the Bitcoin P2P network. They propose two patterns that, whenever observed, reveal the true originator of a Bitcoin transaction to everybody listening. Firstly, if one peer is the only one to relay a certain transaction to the network, it must be the true originator and its IP address can be linked to the corresponding Bitcoin addresses. Secondly, if one peer is the only one to relay a single transaction multiple times the same assertion holds. In order to be able to observe such relay patterns across the network the attacker needs to have a connection to quasi every participating peer.
An even more sophisticated approach was proposed by Biryukov [11]. Again the attacker needs to connect to as much peers as possible in order to succeed. Conforming the protocol a newly connected peer will advertize its own IP address to the eight random peers it directly connects to. These in turn will relay the new IP address to their own peers and so on. An completely connected attacker can now log the first eight peers he receives a new IP address from, having a good chance of retreiving the eight random peers of the newly connected peer. Using this eight peers as a fingerprint for following transactions received over the network can unmask the true originator of the transaction. Again a link between the transaction and an distinct IP address is established.
Both approaches can immediately deanonymize the originator of any transaction advertized to the network, but they have a serious drawback, which makes them inapplicable to my scenario. Since the links between transactions and the originators IP address are uncovered at the moment the transactions are issued, this technique is of no use to an investigator that has to solve a crime committed in the past.

2.3 Privacy Enhancing Technologies

The last category of papers does not deal with the Bitcoin protocol itself, but with privacy enhancing technologies that can be found in the Bitcoin ecosystem. Since they are meant to improve the anonymity of Bitcoin users, any work on assessing or attacking these also adds to the discussion on anonymity in the Bitcoin system. There are two important kinds of privacy enhancing technologies I want to discuss. The first is Bitcoin mixing and the second is the so called CoinJoin protocol.
The idea of Bitcoin mixing is based on the work on anonymous electronic messaging of Chaum [12]. A user *A* can send a certain amount of Bitcoin to the mixing service and the service will return the same amount to a new Bitcoin address of the user's choice. When the service is returning the money it will not use the coins formerly sent by *A*, but coins of another user it engaged with previously. Since *A* receives randomly choosen coins of the service provider's funds, only the mixing service itself can link *A*'s inputs and outputs afterward. The service provider will usually demand a small fee for this and guarantuee to delete any paper trail on the transactions. The analysis of real world implementations of this scheme shows results that are twofold. On the one hand, it is indeed hard to correlate inputs and outputs of mixed coins, when Bitcoin

mixing is properly implemented [13]. On the other hand many of the implementations out there do not provide proper unlinkability [13, 14].
The second privacy enhancing technology, namely CoinJoin, was first proposed by Gregory Maxwell in 2013 [15]. It has found adoption by multiple projects lateron[1]. Yanovich et al. have proven the problem of linking of inputs to outputs in this scheme to be NP-hard [16].
Both examples show, that it is possible to achieve transaction unlinkability within the Bitcoin protocol. This makes it even harder for a financial investigator to track flows of Bitcoin. Opposing all the previously mentioned schemes my approach is designed to cope with such privacy enhancing technologies without utilizing implementation specific weaknesses.

[1] https://en.bitcoin.it/wiki/User:Gmaxwell/state_of_coinjoin

3 Background

3.1 The Blockchain

In 2008 Satoshi Nakamoto introduced the now famous payment scheme 'Bitcoin' and with it the notion of a blockchain. The Bitcoin blockchain is a permanent and inmutable history of all Bitcoin transactions that have ever been issued.
Basically a block is a collection of transaction statements like "Alice sends 10 Bitcoin to Bob". These statements are collected by so-called miners. The miners verify the statements and integrate them into a new block. By adding a cryptographic hash of the previous block, the blocks are chained toghether. Miners have to compete in calculating the next valid block with other miners as only the first block will be added to the blockchain and awarded with newly generated Bitcoins. Calculating a valid block entails more than just gathering enough transaction statements and adding the hash of the previous block. In order to discourage malicious entities from adding nonsensical blocks to the blockchain a miner has to prove, that he spent considerable work into the creation of the new block. This is done by adding a nonce to each block and changing its value until the hash value of the current block contains a certain amount of leading zeroes. Only when a new block is made of valid transactions and contains the correct nonce to satisfys the proof-of-work condition, it will be accepted by the community and considered the next block of the blockchain. Miners will start over trying to add the next block on top of it.
The structure of the blockchain guarantees for some important properties like inmutability, publicly verifyable transactions and prevention of double spending.
Inmutability is ensured by the combination of two facts. First, changing the contents of a previously accepted on block entails redoing the work of finding the correct nonce. Secondly, since the hash of this block has been integrated into the following one, the entire blockchain has to be recomputed from the point of the manipulation on. Honest miners will always advance on the longest chain they find and thus the cheater will have to outpace all the other miners alone, if he wishes the community to accept his manipulation. So as long as the majority of computational power is on the side of honest players, the chances of manipulating the blockchain are very small.
In addition, transactions are publicly verifyable, because the blockchain itself is open to the public and contains every transaction the community ever agreed on. Everyone has the option to look up and verify any transaction.
Double spending is prevented by requiring every transaction statement to explicitly name the past transactions that source the amount to be spent in the newly advertized transaction. For example, if Alice wants to send 10 Bitcoins to Bob, Alice needs to name previous transactions, that in turn sent the amount of 10 Bitcoins to her. This way she can prove, that she indeed owns the 10 Bitcoins she wants to send. It does not matter which and how many transactions Alice names, as long as she has not called upon the same resources before and they total to at least the amount of 10 Bitcoins. Again all this constraints can be validated by anyone, by looking up al related transactions in the blockchain.

3.2 The Bitcoin Protocol

The Bitcoin protocol controls how the blockchain is distributed, transaction statements and new blocks are advertized and how users can interact with each other. At its core the Bitcoin protocol is just a P2P protocol, where everyone can talk to everyone else. Although the original Bitcoin client as proposed by Satoshi Nakamoto is still widely used, multiple implementations from various institutions exist nowadays.
There are no central servers everyone needs to connect to. Instead a newly connected client just chooses a random set of peers to connect to. If the client has never connected to the network before, it will first request the complete blockchain from its peers and download and verify it block-wise. This can take some time, since the blockchain is almost at 200GB in size at the time of this writing.
Afterward the client can issue own transactions, by simply sending the transaction statement to its peers via the network. On receipt the peers will verify the transaction and resend it to their respective peers until everyone connected to the network knows of the new transaction. By simply drawing upon a tiny bit of more resources in a transaction statement, than is actually paid to someone, one can leave the rest as a fee to reward the miner, who adds the transaction to a block.
Miners are alway listening for such new transactions and collect them for integration in a new block. Usually a larger few results in faster integration of the transaction into the blockchain. Miners, in turn, will advertize new blocks to the network in the same way. If the block satisfies all requirements, each client will add it to its local copy of the blockchain upon receipt.

3.3 Stylometry

Stylometry builds on pattern recognition to identify the author of a work. It has successfully been applied to various fields, like writing, music, painting and computer code.
Unique patterns in style allow an analyst to differenciate between single authors. Depicting the style by numerical or logical features enables computers to aid in analysing it. A typical stylometric analysis would include the phases of pre-processing, feature crafting, feature extraction and prediction.
In the pre-processing phase the data analysed needs to be transformed to a computer-readable shape. Depending on the application field this can entail multiple steps, where in the end the data needs to be digitalized and in a well defined format.
The phase of feature crafting is tightly coupled with the data itself. The goal of this phase is to identify reasonable and discriminative features to describe the style of a single piece of data. For example, if you want to identify the author of a business letter the valediction might not be as discriminating a feature as the spelling mistakes, because there are only very few valedictions commonly used, but many mistakes one can make. Adding to the quality of spelling mistakes as example feature is the fact, that not knowing how to spell out certain words is a very personal issue. On the other hand looking at spelling mistakes, if you want to find the painter of some picture might not be a good idea as well, because chances are good there is no text to find mistakes in at all. The total of all features identified in this phase is called the feature set.
Once the analyst has decided on a feature set, the phase of feature extraction can take place. In this phase every piece of data is looked at. For each feature of the feature set a (usually)

numerical value is extracted and added to the feature vector. The feature vector is subsequently used to describe the style of the corresponding piece of data.
In the last phase a prediction regarding the author of each piece of data needs to be made. The outcome of this phase can slightly differ based on the concrete problem statement. For authorship verification a simple 'yes' (the hypothesis of *A* being the author is accepted) or 'no' (the hypothesis of *A* being the author is rejected) can be enough. For authorship attribution one of many authors needs to be assigned to every piece of data. Both problems require a reference work for each author to compare the computed feature vectors to. In a more open scenario it might even be enough to just cluster the single pieces of data by the distance to each other to discover the number of different authors.

3.3.1 A Word on Interpretability

It is increasingly popular to combine stylometry with machine learning techniques to automatically find features. These schemes heavily lighten the workload of analysts, because no or only minimal human intervention is required in the feature crafting phase. This goes at the expense of interpretability, since sometimes not even experts can explain why certain features have been chosen and why they add to the discrimination of authors [17]. Depending on the application this may or may not be an acceptable property.

4 Problem Statement + Methodology

5 Experiments + Evaluation

Bibliography

[1] T. McCarthy, "Trump and 'collusion': what we know so far about mueller's russia investigation," 2018.

[2] J. Rubin, "Mueller follows the money trail," 2018.

[3] D. Vogel *et al.*, *Financial Investigations: A Financial Approach to Detecting & Resolving Crimes*. DIANE Publishing, 1994.

[4] FBI, "Bitcoin virtual currency: Intelligence unique features present distinct challenges for deterring illicit activity," 2012.

[5] ECB, "Virtual currency schemes – a further analysis," 2015.

[6] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[7] F. Reid and M. Harrigan, "An analysis of anonymity in the bitcoin system," in *Security and privacy in social networks*, pp. 197–223, Springer, 2013.

[8] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage, "A fistful of bitcoins: characterizing payments among men with no names," in *Proceedings of the 2013 conference on Internet measurement conference*, pp. 127–140, ACM, 2013.

[9] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin," in *International Conference on Financial Cryptography and Data Security*, pp. 34–51, Springer, 2013.

[10] P. Koshy, D. Koshy, and P. McDaniel, "An analysis of anonymity in bitcoin using p2p network traffic," in *International Conference on Financial Cryptography and Data Security*, pp. 469–485, Springer, 2014.

[11] A. Biryukov, D. Khovratovich, and I. Pustogarov, "Deanonymisation of clients in bitcoin p2p network," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 15–29, ACM, 2014.

[12] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, 1981.

[13] M. Moser, R. Bohme, and D. Breuker, "An inquiry into money laundering tools in the bitcoin ecosystem," in *eCrime Researchers Summit (eCRS), 2013*, pp. 1–14, IEEE, 2013.

[14] T. de Balthasar and J. Hernandez-Castro, "An analysis of bitcoin laundry services," in *Nordic Conference on Secure IT Systems*, pp. 297–312, Springer, 2017.

[15] G. Maxwell, "Coinjoin: Bitcoin privacy for the real world," 2013.

[16] Y. Yanovich, P. Mischenko, and A. Ostrovskiy, "Shared send untangling in bitcoin," *The Bitfury Group white paper*, 2016.

[17] W. Knight, "The dark secret at the heart of ai," 2017.

6 Glossary

AML Anti-Money Laundering

CFT Combating the Financing of Terrorism

ECB European Central Bank

HAC Hierarchical Agglomerative Clustering

IRS Internal Revenue Service

KYC Know Your Customer

FBI Federal Bureau of Investigation

P2P Peer-to-Peer

A Some Appendix