Towards Deanonymization of Mixing Services in Bitcoin

1st Chábeli Castaño Arango*
Department of Computer Science
Florida International University
ccast353@fiu.edu

2nd Roberto Luna-Garcia
Department of Computer Science
Florida International University
rluna017@fiu.edu

3th Steven Cutchin
Department of Computer Science
Boise State University
stevencutchin@boisestate.edu

4th Gaby G. Dagher
Department of Computer Science
Boise State University
gabydagher@boisestate.edu

***Abstract*—Bitcoin transactions are pseudonymous, which means that even when addresses or addresses can be connected one to another, it is really hard to connect them with outside entities. On top of that there have been a proliferation of bitcoins mixing sites in recent years. These sites operate mostly on the dark web and their main mission is launder bitcoins by making vast amounts of complex transactions with them and to make their association with a single owner even harder. Our mission is to make that distinction easier. In this paper we plan to introduce two novel heuristics. Our OI heuristic is designed to parse blockchain data in a way where we only receive information we deem is of interest. We introduce HOLO which aims to taint bitcoin addresses in reference to a fixed output. We also visualize the blockchain and our heuristic in an easily digestible manner.**

***Index Terms*—Bitcoin; Taint Analysis; Visualization; Mixing Services;**

I. Introduction

Bitcoin remains one of the most prominent crypto-tokens on the market today, with about half of the market share in the crypto-space. The decentralized and pseudonymous nature of the blockchain has attracted many illicit services, such as tumbling, with more than half of all Darknet users tumbling their Bitcoin [1]. Such tumblers, or mixing services, make it hard for individuals to know if they are receiving funds from illicit actors. This leads to distrust of the system and can harm the Bitcoin economy as a whole.

A major problem with detecting Bitcoin laundering is the inherent privacy of the system. Even though the public ledger provides some clues about ownership, it's difficult to determine who is a legitimate wallet and who is an illicit wallet. One challenge is following the flow of Bitcoin and determining which outputs are direct beneficiaries of a dirty input. In order to combat this, there are many proposed heuristics that aim to, in some way, classify dirty wallets.

*These authors contributed equally.

Taint analysis has become an emerging interest in Bitcoin and attempts to assign a certain amount of Taint to a transaction or address. There are many heuristics that attempt to accurately Taint chains of transactions such as Poison, Haircut, and FIFO being the most common [8] [10] while LIFO and TIHO are proposed in [10].

In this paper, we propose a heuristic we will be referring to as HOLO (Highest Out Lowest Out). In contrast to previous methods, our heuristic taints inputs based on outputs. We assign the largest amount of taint to both the highest and lowest inputs. This way we can cover the largest transaction, which is believed to be important, and also not be fooled by peeling chains. While our heuristic is better equipped to deal with peeling chains, it is still vulnerable to being fooled by dusting.

The contributions of this paper are as follows:

1) We propose two novel heuristics OI (Output to Input) for trimming addresses from the Bitcoin address graph while traversing the blockchain backwards, and HOLO (Highest Out Lowest Out) for backwards address taint analysis.
2) We design a method to visualize transactions known to belong to darknet users and illicit services to identify meaningful patterns.
3) We implemented our approach and our experimental evaluation results show that our approach is scalable and the total tainted addresses are reduced.

II. Related Work

A number of papers explore the illicit nature of Bitcoin and how illicit actors use Bitcoin for illegal purposes such as Foley et al. [1], Möser et al. [2], Balthasar and Hernandez-Casto [3], Crawford and Guan [4], Simin et al. [5], and Spagnuolo et al. [6]. Foley et al. [1], and Spagnuolo et al. [6] both focus on general crimes committed on the dark web using Bitcoin,

such as buying drugs and sex services or hiring killers on demand. [1] estimates 46% of Bitcoin Transactions involve illicit activities and 26% of all users are involved in those activities. These activities usually happen on the so called dark net (which is accessed using Tor browser) and about 64% of dark net users tumble their bitcoin. [2], [3], [4], and [5], are all papers that examine Bitcoin Mixers at a closer level. [4] and [5] compare and contrast tens of mixers and report their findings. [4] concludes a good number of mixers are scams. We have found that these papers lack simple visualizations to detect patterns in each mixing service and we attempt to bridge that gap.

A group of papers look into Taint Analysis, such as Gifari et al. [7], Hercog and Povše [8], Tironsakku et al [9], and Tironsakku et al, [10].

Tironsakku et al, [9] uses backward address taint analysis, meaning they taint any address that sends bitcoin to a tainted address. We will also be using backward address taint analysis in our implementation, however in order to determine taint Tironsakku et al use what they call the Baseline method. The Baseline method considers all outputs of every transaction on the blockchain within a given time frame, and considers them to be tainted, they then use input sharing and output sharing clustering to analyze their chains. Meanwhile, our OI heuristic starts with one address and follows all transactions, for a given time frame, recursively, where the input of our current transaction is an output. It does not consider outputs as groups. This is due to the fact that there is no limitation to being an output in a transaction, as opposed to an input which requires a signature and can be grouped with the common input heuristic. We also consider taint to be a value instead of absolute such as in poison and in [9]

Tironsakku et al, [10] study five different kinds of taint analysis. Poison, which taints all outputs of a tainted transaction. Haircut, which taints each output in proportion to its value compared to the total output. FIFO, the first input correlates to the first output and so on. LIFO, which is the opposite of FIFO. They also propose a new heuristic, TIHO (Taint in, Highest out), which prioritises the distribution of the tainted inputs to the highest value output. We believe our heuristic does a better job of taking into account Peeling Chains discussed in [3] and [4], which are overlooked by Haircut, FIFO, and TIHO.

There are other papers that focus on visualizing the blockchain, such as McGinn et al. [11], Kinkeldey et al. [12], Shrestha and Vassileva [13], Di Battista et al. [14], and Ahmed et al. [15]. The majority of these papers focus on general Bitcoin visualization, such as visualizing blocks, transactions, or actors on the blockchain. Ahmed et al. [15] specifically uses FIFO tracking in a system they called the Taintchain to propagate through transactions from known thefts or scams to implement taint analysis. While most similar to ours, we want to have a visualization with a focus on Taint Analysis of mixing services, as opposed to thefts or other crimes.

Other related works such as Maurer et al. [16] delve into programs that are built to oppose Taint Analysis such as CoinJoin.

III. Preliminaries

Bitcoin is a peer to peer pseudonymous blockchain, which functions as electronic cash. It was proposed by Satoshi Nakamoto in [17]. Bitcoin uses proof of work in order to validate transactions. Instead of a central authority, special miner nodes are tasked with validating each transaction and adding them to data structures known as blocks. Miners are incentivized by coinbase transactions, where they are paid Bitcoin if they are the first to validate a block. Once validated, the block propagates through the system as all the nodes eventually accept it into the blockchain.

A transaction in the bitcoin network consists of inputs and outputs. These inputs and outputs in turn consist of addresses. Many of the transactions in the blockchain are many to many, but there can be any number of outputs and any number of inputs excluding zero.

The Bitcoin network is also said to be immutable. The system is immutable so long as no entity controls a majority of the nodes in the system. Due to the sheer size of the blockchain, and the computationally expensive proof of work, this is thought to be very unlikely.

Tumblers/Mixers in Bitcoin serve to ”clean” any Bitcoins that are from a known criminal account. Many mixers exist mainly on the Tor network and all use different methods in an attempt to break the correlation between Bitcoin submitted to the mixer and Bitcoin that exits the mixer [4] [5]. Mixers will create new addresses for users to send bitcoin to, and will even allow them to specify a certain delay between the initial mixing and when the user receives their funds.

Mixers come and go as they are constantly playing cat and mouse with law enforcement. Many of the most notorious mixing services are no longer running, and many of them suffer from ”Evil Twin” mixers that copy their namesake in order to steal bitcoin from unsuspecting users [4]. There are two main types of mixers, centralized and decentralized. Centralized mixers are controlled by a central entity or person, while decentralized mixers attempt to use algorithms in order to further obfuscate their methods.

Gephi is an open source software made to explore, manipulate and explore networks. It can handle networks with over 20,000 nodes and allows personalized node design. Gephi also supports many clustering algorithms that can be applied to graphs, such as Force Atlas 2 [18].

For visualizing our graphs we will be using the Event Graph layout. This layout maps our nodes to a value on the horizontal axis. The event graph requires multiple values to be filled in such as Scale of Order, Vertical force, and Gravity. The scale of order is by default 10 and can be changed to make the graph tighter or sparser on the x axis. The vertical scale determines how strongly unconnected nodes are repelled from each other. The Gravity variable sets the main attracting force that can prevent islands from drifting away from the rest of the nodes.

IV. Solution

Our solution will parse data from the Bitcoin blockchain using the Blockstream.info API, visualize the fore-mentioned

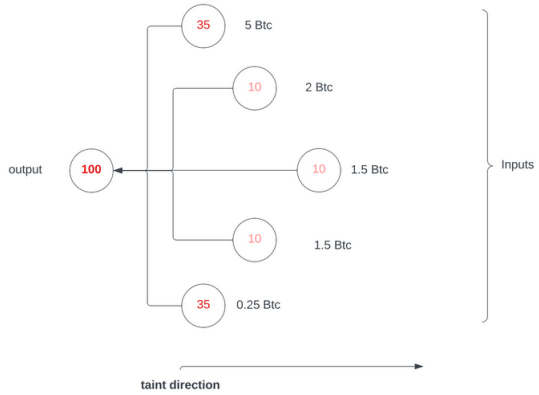

Fig. 1. HOLO, which is a combination of TIHO and LIFO

data with a visualization tool, and apply heuristics for following blockchain transactions of interest and taint analysis of a selected chain of transactions.

We choose an address that recently has sent to a dark web address and ask the Blockstream.info API for any transactions where this address is an output. We then add all inputs of that transaction into an array, where we then check if each transaction where the input is an output and if it has occurred in the past month. Using our OI heuristic we remove certain data from our date set. We then write this data into a Node file and an Edge file both of which are in CSV.

In order to visualize the data we import the CSV files into Gephi. Gephi uses special keywords in each CSV file to determine the nodes and the edges, it then connects those edges and nodes as a network graph. We can achieve a desired visualization using the built in algorithms in Gephi.

OI is our traversal heuristic in which we try to maintain a direct relationship between our seed transaction and the rest of our chain. HOLO is our taint analysis heuristic and attempts to equally taint the addresses with the highest and lowest input values, and then equally distributing remaining taint to the addresses in between.

A. Heuristics

Our OI heuristic is as follows: In cases where a transaction has outputs other than our target address, we will not consider those outputs. We assume that they are not directly related to our seed transaction. This allows our taint analysis to function similarly to taint analysis done from input to output. It eliminates the challenge of bi-directional taint analysis.

HOLO is a mixture of TIHO and LIFO suggested in [10]. We first order the list of Inputs by amount of BTC sent and will taint the largest transaction and the lowest transaction by 70 percent. We chose 70 to avoid any transactions in between the largest and lowest having more taint.

B. Getting Blockchain Data

In order to get the Blockchain data we created a python script that make calls to the Blockstream.info API. Using cryptocurrency tracking sites such as walletexplorer.com, we

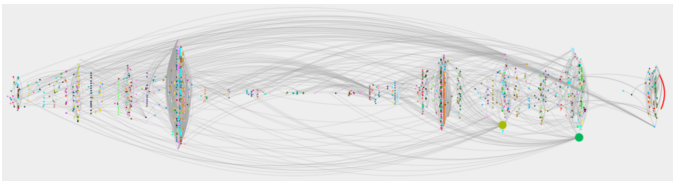

Fig. 2. All transactions associated with an address; Complex

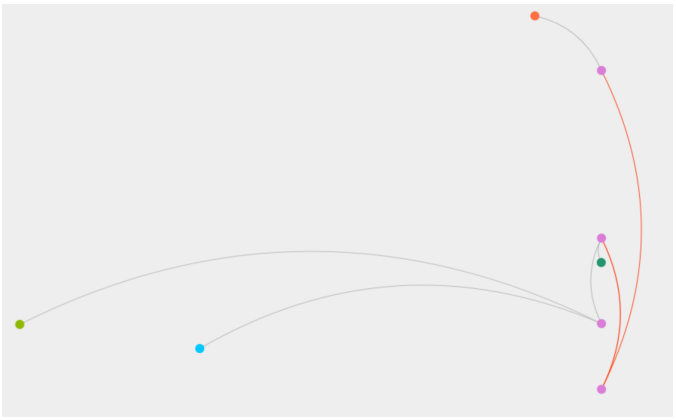

Fig. 3. All transactions associated with an address; Simple

picked a series of known compromised addresses and transactions where these addresses have been involved.

Our first function takes one of those addresses and its corresponding transaction -we will call them our seed address and seed transaction respectively- as parameters. Then, it verifies if the address is an output for the transaction. If this is true, it will return the inputs for that same transaction, which will allows us to start tracing the bitcoins backwards. Having those inputs, we will then check were they are outputs and so on in a recursive manner.

In order to establish some kind of time frame assume that a block is created approximately every ten minutes. The number of blocks created within a month is estimated to be be 4320. Having this number and the block height for the transaction seed, we can establish an upper and lower bound for our data collection.

C. Visualization

Our CSV files are separated into nodes and edges. For our nodes we define 6 columns, Name and Id, which are required by Gephi to identify our nodes,Transaction List, Taint Value, Start Block, and End Block. Transaction list is a list of every transaction we have found this particular node to be a part of. Start block and end block are taken from an ordered list of all the blocks an address has appeared in. We narrow it down to two values in order to create a time interval column in Gephi, which allows us to view our graph at multiple stages throughout the month. Taint value is the taint value our heuristic has assigned to that specific node.

For our edges file we have the required columns Source, Target, Taint Category and Weight. The source column indi-

cates which node is the 'source' of an outgoing edge, while the target column specifies the target. For the weight column we have chosen to use our taint values to more easily visualize the tainted relationship between nodes. The taint category column is used to determine which transaction is our seed transaction, and it can be easily identified in our visualization.

In order for our graph to work with the Event Graph layout we have to create a new column in our nodes file. We will call this column X index. This is necessary because it is not possible to map directly from the block heights, they are too large and exceed the Gephi canvas. This makes them gather at the edge of the screen in a straight line. We calculate the values in the X index by subtracting the last block an address appeared in from itself minus 2,000. This means our first index will be -2,000 and will not exceed 2,320. This allows Gephi to draw our graph within its canvas. we chose to use the last block an address appeared in to better visualize the addresses near our seed transaction. Edges are colored by category, meaning that our seed transaction is easily distinguishable with a red line, while the rest of the transactions are gray. Figures 2 and 3 shows a Graph made in Gephi after implementing our parsing and tainting heuristic.

Due to the variability in our data, we separated our graphs into two groups, simple and complex. A complex graph is one that provided us with a large amount of data that would allow us to visualize a larger network. A simple graph is one with less data, and a less complex graph. Figure 2 is an example of a complex graph, and figure3 is a simple graph.

The complex graphs allow us to make more meaningful observations in regards to mixing of Bitcoins. Given their nature, we suspect that an address that produces a complex graph is more likely to have been involved in mixing within that month. In Figure 2, based on the address 1EPHNmtC-nWZU4A7Vm68TDHmKrHoGaQ9GuA, we see that some nodes are larger than others, this is because we chose to resize nodes based on out-degree. We paid special attention to the addresses because they may signal the ending of a mixing cycle, where a mixer has to return money to users, which would produce nodes with high out-degrees. We also take not of the addresses that are heavily connected by have a small size. This means they have a high in degree. As mentioned in [3], some mixers use centralized addresses once they commence their mixing. This could be an indication of one of those addresses.

V. Experiments

A. Implementation and Setup

Most taint heuristics are compared and contrasted to existing heuristics in order to determine effectiveness. However, the main types of taint analysis such as Poison, Haircut, and FIFO are designed to function from input to output while taking into account what amount of bitcoin is transferred between accounts. We have determined that it would not be an accurate comparison to apply those heuristics in a way they were not designed to be implemented. This is why we have decided to focus on the scalability of our system. we will also be examining how many transactions our OI heuristic has parsed for every thousand blocks.

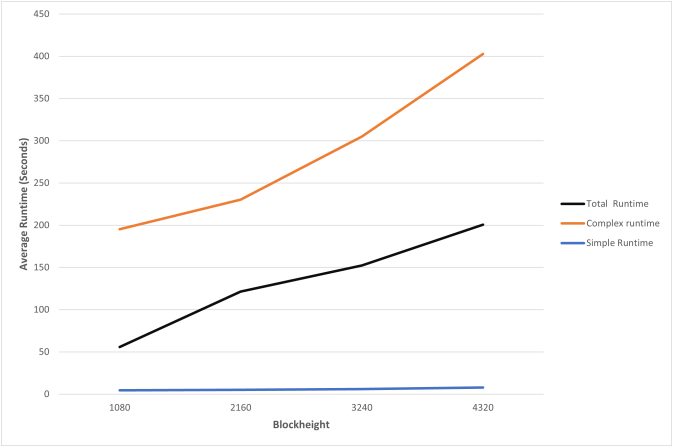

Fig. 4. Average Runtime Across all Addresses

For our experiments we will evaluate 26 addresses that we determined to have sent to a malicious dark web address such as Silk Road and other illicit marketplaces. These addresses were taken from WalletExplorer.com

B. Scalability

We performed a run-time experiment in order to measure how long it takes our python program to process different amount of transactions and write them to our CSV files. Since we assumed that 4320 blocks represents approximately a month, we decided to further divide that into weeks, being 1080 blocks per week. Then, we calculate the time in seconds that it took to process one, two, three and four weeks separately. In most cases we obtain linear results, meaning that each increment yields a number larger than the previous experiment, but there were some exceptions. We think that there are many factors that can affect this experiment. One of them is the fact that we are working with data that can be modify in real time, thus any change to the blockchain might also affect our results. Another one is that there might be certain variation in the API response times. We ran a total of 26 addresses for each four week interval. We then took an average of each week and plotted them on figure 3. We also separately plotted the average run times for our simple and complex addresses. They show that our complex data requires a larger amount of time to process as the addresses processed increases. We expect our lines to not be exactly straight or curved due to the variability in our data.

C. Weekly Transaction

In order to determine if any period of time is more active than any others, we decided to plot each address on a graph, as shown in figure 4. Our green line represents 17 of our simple addresses that had little to no activity and were limited to week 4. They appear as one line but are in reality all following the same path. Among our complex addresses we find that the transaction activity usually peak around week 2 or 3, which

Address	Total Addresses Visited	Addresses Tainted	Percentage
1ALRsFWtAoetAaht3yQ7pwhwmNSiunxMxR	223	46	20.63
15uyvmNQtLPyzeNcBCvuvgH4f7MUN6XFKF	93576	553	0.59
1BA1FjTAm5Xb5c8XMGqJ377SSmDgSRgZWV	631	63	9.98
1KFsFbmuwwpoViLG3TVQLhZwND6Z7cZSsr	1118	225	20.13
1CQiErmibvpEGvCdQQT9h5cEhtbEVWZNCt	166	35	21.08
1PhyakEn4RpNZVDi6E9tPJKFwx9kHUm3YN	5	2	40.00
1MFEDYqa5Ezqtd8QR58RcgV6HGNU2ejFqx	15771	2	0.01
1573tdS4xad6faPMoQS3iTKwBmKgsfQVuP	251	2	0.80
1LdLiYhEXPuP4Mk9GvCKwkveMfYxz8DyAd	251	2	0.80
1CvnBQ9zkPYwG77oZoSkQFYCpKxhFquNji	5	2	40.00
17ZCySiULt6SV2DF6Ke8RJsdD6ct1MFQzJ	3	2	66.67
37NsSsEKqruFSZ9vWCFGNJjv2BooRwFqma	996	154	15.46
1EPHNmtCnWZU4A7Vm68TDHmKrHoGaQ9GuA	31246	1595	5.10
3EQPT2yUi46GSnXYCkN4RyMf6rMGeYEfEK	58	10	17.24
3EKj2EEjLcA8TAknrbZZ1AGLsMF3Ggtm1A	8	3	37.50
1LssZGrJSszcx7DFRDTRknPvbf8JZhPgGk	50	11	22.00
1FcA7HQXdTP8ZF4YLE7oNBX8c8B5mEngwg	533	108	20.26
3NquBJyPbhZFoaeEUoKcMKRFx8aFk9eEdQ	305	5	1.64
39f2Sm9FmovDvRXkWSTSNuWwb7JjQmUmjz	8	4	50.00
1PhMtSbi39Qhcu3pwqv9s5pt4M6R84iZyo	3	2	66.67
14geEFfPXT9K5Vao9DcoFeikPDJyc7QZUM	24	3	12.50
19MXDBEyGL1zPp6K6QR4yDHA7aDSdJU7UW	3	2	66.67
1LgGUMbfLBVFX8RTWCFdsuDaZm6bcgWSGN	19012	1905	10.02
1PzD1ima5fXfwFg3PJAJpffjqRHLkajPLj	38	9	23.68
3JNyjJTu5uw4YwKU5utDrCKFAURn7zHJV7	54781	501	0.91
34jqRto9zZo8qy73w6aMKu48qL8eabyZKb	121	9	7.44
1Z1yC8fKbG2RKmarVaVcTe6MQW8czeagv	55	7	12.73

TABLE I
EVALUATION OF OI HEURISTIC

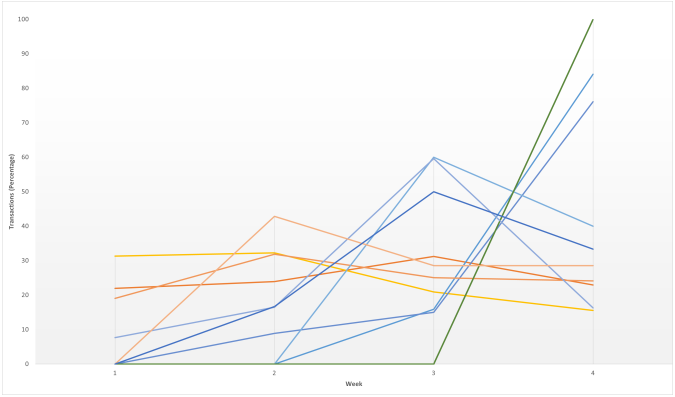

Fig. 5. Weekly transactions of each address as a percentage of total transactions for the month

lines up well with mixing services that offer delays for a week or longer. This leads us to believe that the time period with the most likelihood of mixing are weeks 2 and 3.

D. Parsing Transactions

In order to calculate how effective our OI heuristic is, we ran an experiment that compares the total number of visited addresses to the total number of tainted addresses. We found that on average, our OI heuristic tainted, on average, approximately 22% of the addresses as shown in table 1. This percent of transactions are the ones we consider to have the highest likelihood of being part of a network with a direct correlation to our illicit address. This allows our taint heuristic to not tag every single address, avoiding some of the issues present in other heuristics such as poison.

VI. CONCLUSION AND FUTURE WORK

In this paper, we design and implement heuristics for backwards address taint analysis and data trimming. The implementation of our taint analysis and time based visualization, showed that we are able to identify certain peaks of activity with the highest probability of mixing for each tainted address. We are also able to clearly visualize highly connected nodes of interest based on their in or out degree, which is a possible signal of a centralized mixing address or the end of a mixing cycle. Our data trimming heuristic reduces the number of nodes we apply any amount of taint to.

Most of the current taint analysis happening with the Bitcoin blockchain is from Input to Output, as more and more research is done into backwards taint analysis, we hope to see more heuristics develop that allow this method to become a mainstay. Future research could focus on a tainting heuristic that can take into account both dusting and peeling chains concurrently, which remains a challenge due to the nature of each method. Additionally, future work could involve refining the heuristics that trim addresses in order to be more certain we do not lose relevant information.

REFERENCES

[1] Foley, Sean and Karlsen, Jonathan R and Putniņš, Tālis J. *"Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies?"* The Review of Financial Studies, 2019.

[2] Möser, Malte and Böhme, Rainer and Breuker, Dominic. *"An inquiry into money laundering tools in the Bitcoin ecosystem"* 2013 APWG eCrime Researchers Summit, 2013.

[3] Thibault de Balthasar and Julio Hernandez-Castro. *"An Analysis of Bitcoin Laundry Services"* NordSec, 2017.

[4] Crawford, Jesse and Guan, Yong. *"Knowing your Bitcoin Customer: Money Laundering in the Bitcoin Economy"* 2020 13th International Conference on Systematic Approaches to Digital Forensic Engineering (SADFE), 2020.

[5] Simin Ghesmati and Walid Fdhila and Edgar Weippl. *SoK: How private is Bitcoin? Classification and Evaluation of Bitcoin Mixing Techniques* Cryptology ePrint Archive, Paper 2021/629, 2021.

[6] Spagnuolo, Michele and Maggi, Federico and Zanero, Stefano. *"BitIodine: Extracting Intelligence from the Bitcoin Network"* Financial Cryptography and Data Security, 2014.

[7] Gifari, Abidzar and Anggorojati, Bayu and Yazid, Setiadi. *"On preventing bitcoin transaction from money laundering in Indonesia: Analysis and recommendation on regulations"* 2017 International Workshop on Big Data and Information Security (IWBIS), 2017.

[8] Hercog, Uroš and Povše, Andraž *Taint analysis of the Bitcoin network* arXiv, 2019

[9] Tironsakkul, Tin and Maarek, Manuel and Eross, Andrea and Just, Mike. *"Tracking Mixed Bitcoins"* Data Privacy Management, Cryptocurrencies and Blockchain Technology, 2020

[10] Tironsakkul, Tin and Maarek, Manuel and Eross, Andrea and Just, Mike *"Probing the Mystery of Cryptocurrency Theft: An Investigation into Methods for Taint Analysis"* arXiv, 2019

[11] D. McGinn and David Birch and David Akroyd and Miguel Molina-Solana and Yike Guo and William John Knottenbelt. *"Visualizing Dynamic Bitcoin Transaction Patterns"* Big Data, 2016.

[12] Christoph Kinkeldey and Jean-Daniel Fekete and Tanja Blascheck and Petra Isenberg *"Visualizing and Analyzing Entity Activity on the Bitcoin Network"* arXiv, 2019.

[13] Shrestha, Ajay Kumar and Vassileva, Julita. *"Bitcoin Blockchain Transactions Visualization"* 2018 International Conference on Cloud Computing, Big Data and Blockchain (ICCBB), 2018.

[14] Giuseppe Di Battista and Valentino Di Donato and Maurizio Patrignani and Maurizio Pizzonia and Vincenzo Roselli and Roberto Tamassia. *"Bitconeview: visualization of flows in the bitcoin transaction graph"* 2015 IEEE Symposium on Visualization for Cyber Security (VizSec), 2015.

[15] Ahmed, Mansoor and Shumailov, Ilia and Anderson, Ross. *"Tendrils of Crime: Visualizing the Diffusion of Stolen Bitcoins"* arXiv, 2019

[16] Felix Konstantin Maurer and Till Neudecker and Martin Florian. *"Anonymous CoinJoin Transactions with Arbitrary Values"* 2017 IEEE Trustcom/BigDataSE/ICESS, 2017.

[17] Nakamoto, Satoshi. *"Bitcoin: A peer-to-peer electronic cash system"* http://www.bitcoin.org/bitcoin.pdf, 2009.

[18] Jacomy, Mathieu and Bastian, Mathieu and Heymann, Sebastien *"Gephi: An Open Source Software for Exploring and Manipulating Networks"* Third International ICWSM Conference, 2009.