Behavioral Analysis of Bitcoin Users on Illegal Transactions

by

F. Z. Z. M. SAMSUDEEN
2014/IS/073

H. P. D. U. PERERA
2014/IS/058

Supervisor: Dr. M. G. N. A. S. Fernando

This dissertation is submitted to the University of Colombo School of Computing
In partial fulfillment of the requirements for the
Degree of Bachelor of Science Honours in Information Systems

University of Colombo School of Computing
35, Reid Avenue, Colombo 07,
Sri Lanka

January 07th 2019

Abstract

Bitcoin is a popular crypto currency that is used as a mode of investment and a medium for trading goods and services. The features such as anonymity, security and decentralization are the significant characteristics of Bitcoin. These features attract different types of users to adapt to this payment mechanism since 2009. In particular, it create several opportunities for criminals to involve in illegal and fraudulent activities. Thus, this research is based on the negative discussion around the Bitcoin.

The thesis pays particular attention on identifying user behavioral patterns and significant facts among illegal incidents that are of varied nature. The study examines each illegal incident of different nature so that it is able to identify common spending patterns among illegal users. The motivation for choosing this study lack of literature that covers illegal incidents of various nature and analysis for patterns and significant facts in patterns.

The analysis recognized spending pattern, popular exchanges used, and notable techniques used to cash out tainted Bitcoins. It shows that the illegal users utilize several techniques to cash out tainted Bitcoins. One way is to directly transfer to exchanges or Bitcoin washers or services which is an evaluation of a previous research. Another way is to transfer the Bitcoins in small amounts to several other their own addresses or to fresh wallets in subsequent transactions and thereafter transfer to exchanges. Bitcoins are transferred to fresh wallets in small amounts. They either transfer a constant amount of Bitcoins or in a certain proportion. These attempts are taken by illegal users with the intention of prevent being tracked by investigators. In addition, popular services such as Xapo.com, Helix Mixer and Poloniex.com were discovered as frequently used services.

Keywords—bitcoin, illegal, blockchain, analysis, behaviour, Transactions

Preface

The usage of Bitcoin has become wide due to the popular reasons such as speed, anonymity, security, convenience and decentralization. These features of Bitcoin lead the criminals to conduct their crimes easier than ever. This research study was an effort identifying behavioral patterns of users involving in illegal Bitcoin transactions. Our study concluded by identifying behavioral patterns among illegal Bitcoin users based on the past incidents. The study reveals a spending pattern among incidents and significant facts. In addition, some of our results also confirmed the findings of previous researches. The various approaches of transacting tainted coins were confirmed in our research. But, our novel findings revealed the different patterns of how transactions were done subsequently and how it differs based on the nature of illegal incident. We conclude the thesis successfully by identifying a novel behavioral pattern and thus ending with a note that the patterns also differ with the severity and value of Bitcoins transacted. In addition, in our study we confirmed previous study findings related to user spending patterns. The conclusion is solely our own work and has not been proposed in any other study related to the context of Bitcoin illegal use.

Acknowledgement

To begin with, it is a genuine pleasure to express our gratitude to our supervisor Dr. M. G. N. A. S. Fernando, Senior Lecturer, Coordinator (Computer Science Degree Programme), Department of Information Systems Engineering, University of Colombo School of Computing. The continuous guidance, scholarly advice and caring attention helped us solely and is mainly responsible for completing this study.

We owe a deep sense of thanks to Dr. T. A. Weerasinghe, Coordinator of the Final Year Project for her continuous guidance and motivation at every stage of the research. The timely suggestions and recommendations provided by the panel of examiners, lecturers inspired us to travel on right path throughout this research process. We would also like to extend utmost gratitude to Mr. Anupa Shyamlal for his kind help and continuous support on context of Bitcoin throughout the study.

We are extremely thankful to Network Operations Center, University of Colombo School of Computing for providing us with necessary technical environment and being supportive for the research.

We would also like to acknowledge Bitcointalk and Reddit forum community for extending their support by voluntarily providing information.

Last but not least, it is our privilege to thank our parents and family for the understanding and constant encouragement. We take this opportunity to thank each and every one who has been supportive in completing this research successfully.

Table of Content

List of Figures

List of Tables

Chapter 01

Introduction

1.1 Background to the Research

Many crypto currencies have come into usage in recent years for multiple purposes. Bitcoin developed by Satoshi Nakamoto came into usage from 2009 [1] and it is the most prominent crypto currency in terms of market capitalization of USD 67 billion as at 4th January 2019 [1].

Bitcoin can be used as an investment or as a medium of exchange [2]. But, Technology is referred to as a double edged sword which can be used both with positive and negative intentions. However, Bitcoin is discussed mostly on its negative aspect [3] since Bitcoin systems are being targeted by hackers and fraudsters [4] thus making it easy to compromise [3], [4].

The Bitcoin characteristics such as speed [5], anonymity, security, convenience [6] and decentralization have increased its user base. But, more importantly the decentralization feature has made the illegal users of Bitcoin more comfortable since there is no middle party to control their behavior [7].

It has been reported in a recent study [8] that one quarter (25%) of Bitcoin users and one half (50%) of Bitcoin transactions are involved in illegal activity. Therefore, Bitcoin is being discussed on its negative aspects due to its influence in Bitcoin network. Activities such as dark net marketplaces [9], [10], [11],[12] Ponzi scheme [13], [14], [15], ransomware [16], [17], [18], Bitcoin exploits [14], [7], Denial-Of-Service (DOS) attack [19], thefts [20], [7] and money laundering [13], [21], [22] have been discussed widely by previous researches and media. These negative aspects can be briefed as illegal activities [10].

[1] https://coinmarketcap.com/

1.2 Research Problem and Research Questions

The literature reveals in detail about illegal activities separately based on its nature or as a case study focusing on a single incident. According to careful investigation of the literature, it reveals that there is no evidence or evidences have not been documented properly by analyzing the user behavior of several illegal incidents as a whole study. Considering the increase in illegal users and its decentralization nature its timely important to focuses on providing a user behavioral analysis of illegal Bitcoin users involving on different illegal incidents.

This study aims to provide a broader picture on the patterns of illegal incidents that vary by nature as depicted in *Figure 1.1* as an example. The curves among incidents represent any possible pattern or significant facts among incidents. For example the curve between Incident Q and Y depicts that there is a common pattern in between two different type of illegal incidents (Theft and Ransomware) whereas the curve between Incident X and Y depicts that there is a common pattern in between two incidents under same nature of illegal activity (Ransomware).

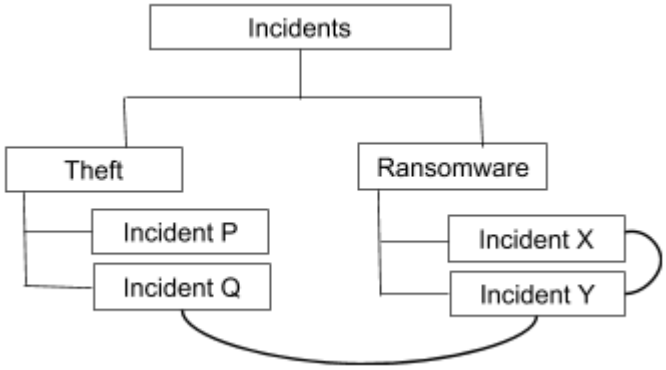


Figure 1.1: An example of analysis considering different illegal incidents

Considering this research problem, the generated research questions are as follows:

Main Question

What is the behavioral pattern or significant facts among Bitcoin users involving in different types of illegal incidents?

The following sub questions assists to find answers for the above main question

Sub Questions

1) What are the illegal incidents involving Bitcoin?
2) What are the common features that can be used to categorize these identified incidents into various categories?
3) What are the significant transactions among the identified illegal incidents?

In order to answer the main research question first it is required to identify the incidents that involved Bitcoin in an illegal manner. It would be easier for analyzing the identified incidents if they can be arranged in a categorized manner since human mind understands classified content better. Therefore finding the common features in order to recognize incident categories is another sub question that support to answer the main question. Thereafter, by identifying significant transactions that prevail in each of the illegal incidents; the patterns or significant facts can be discovered.

1.3 Goal

The ultimate goal of the research study is to provide a behavioral analysis of Bitcoin users involving in illegal incidents that are of varied natures. This is in order to provide timely information to the Bitcoin miners, Bitcoin users and relevant authorities with identified patterns or significant facts among past illegal incidents. To reach the goal, following are the objectives that have been set.

1.4 Objectives

1) Identify illegal incidents and categorize them according to their nature.
2) Identify significant facts in each transaction of the incident.
3) Recognize different patterns of user behavior on transactions among incidents and provide a comprehensive behavioral analysis of illegal Bitcoin users.

1.5 Significance of Research

This study is important to provide a more comprehensive idea on transacting patterns of Bitcoin illegal users. Hence it will mainly assist the Bitcoin miners, Bitcoin users, potential Bitcoin

users and related officials.

We cannot eliminate criminals from carrying out illegal activities. Neither we can stop them from spending the Bitcoin stored with them. But, creating awareness among the Bitcoin community and officials would lead to taking further actions to reduce the impact from it.

The Bitcoin miners validate the Bitcoin transactions. By providing them with timely information about the past incidents and the pattern of transacting, they are aware of the practices by illegal users. In future, if they experience any similar transaction pattern, they have the awareness not to mine those transactions but to broadcast it through the network as suspicious.

The study will be of use for the Bitcoin user community to a greater extent. Bitcoin operates in a decentralized environment. Thus, there is no centralized single party to create awareness or provide information about illegal users' behavior to the community. According to users comment in bitcointalk forum, revealing such information on unknown aspect of Bitcoin stays as an 'eye opener' and to be more vigilant when transacting.

In addition, the discovered patterns and significant facts would assist relevant officials to impose new rules and regulations on tracking suspicious illegal users and establishing controls on Bitcoin services such as exchanges. Thus if there is a way to identify the real illegal user of a Bitcoin address by mapping the address to the owner [23], it is possible for taking actions against them in the real world. Thus, this study paves a path for future researches that would assist to create a protocol in the Bitcoin network avoiding process of transactions that follow these patterns.

Thus, this research is the first attempt to consider several illegal incidents that are of different nature. Even though there is more detailed analysis carried out on individual incidents or of illegal activities that are similar in nature, this is the first attempt of research taking all into consideration. It is important to provide information, since regardless of the nature of incident, or severity of incident; they cause loss for the innocent party. Therefore, it is important to be aware of the patterns.

1.6 Methodology

A mixture of inductive and deductive research approach is practiced in this study. To validate the related findings in literature, the deductive approach is used. For the purpose of new findings, the inductive approach is practiced. The research methodology is in *Figure 1.2*. as shown in a high level diagram. The chain of transaction with respect to the illegal incidents is obtained by traversing via blockchain. In order to traverse transaction, the transaction id linking technique is used which is based on the N-ary data structure. The open source tool BlockSci is used for obtaining blockchain data by running scripts in Python programming language.

The traversed transaction data are visualized in the tool Gephi. The most significant data are summarized for further analysis. On the process of analysis, Bitcoin addresses and transaction chain are identified. Thereon, transaction data is traversed back in the blockchain to derive patterns and significant facts. As the last step, the patterns are evaluated using user feedback and based on literature.

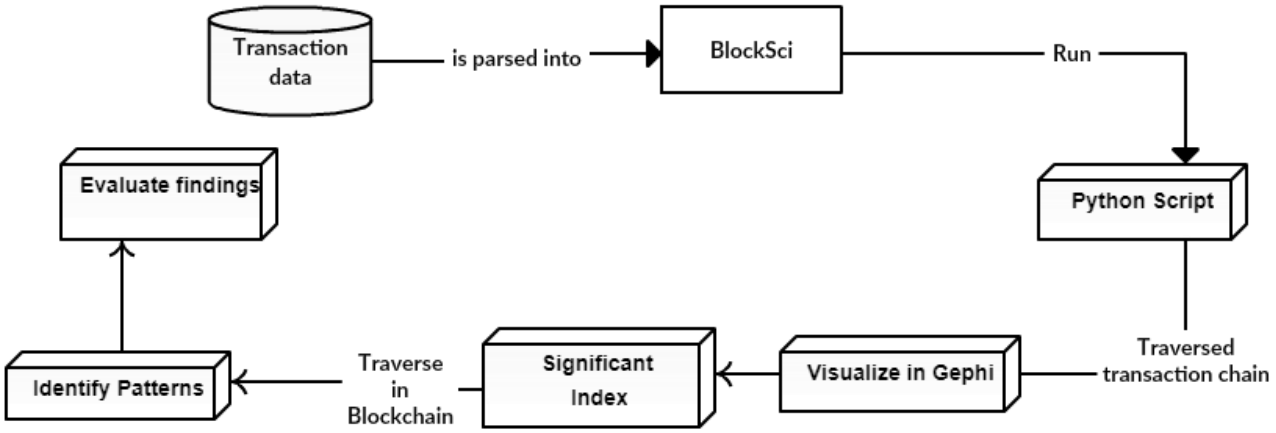


Figure 1.2: Research Methodology

1.7 Outline of the Thesis

The thesis is structured as follows. Chapter one gives an introduction to the research questions highlighting the significance of the research, goal, objectives along with the research approach. Chapter two gives background idea of Bitcoin and basic concepts. It also explores the existing analysis on individual illegal activities. The Chapter three presents the methodology carried out to

align with the objectives. It also includes an illustration of the data analysis. Chapter four represents the results of each illegal incident that were analyzed. The Chapter five presents the discussion of the results obtained along with a summary of new findings and validated finding. This chapter includes the evaluation of results as well. The last chapter six gives the conclusion of this study by highlighting the contribution and outlining the future work.

1.8 Delimitations of Scope

In this research, the illegal incidents considered were limited to the definition of 'Illegal Activity'. That is, the study did not consider every single illegal activity defined in accordance with general definition of law authorities

Specific country rules were not considered because legality of Bitcoin is different according to the country law. For example some countries consider Bitcoin as legal or illegal or restricted whereas some other countries are neutral on legality status of Bitcoin.

Chapter 02

Background

2.1Bitcoin

The Lehman Brothers was the fourth largest investment banker in the United States. It collapsed on October 2008. This incident created a lack of trustworthiness on all financial institutions among the investors. This made investors to require for a decentralized system with higher transparency in nature, low transaction fee and high returns for their investment. Within the next two months after the collapse, a paper was published on 'Cryptography Mailing list' on metzdowd.com titled as "Bitcoin: A Peer-to-Peer Electronic Cash System". This paper created hype among investors due to the positive impression it had and negative experience from fall of investment bank [7].

Bitcoin, developed by Satoshi Nakamoto came into usage from 2009 with the publishing of the white paper [1]. Even though multiple crypto currencies have come into usage in recent years for multiple purposes, Bitcoin is the most prominent in terms of market capitalization of approximately USD 67 billion as at 4th January 2019 [2].

The first Bitcoin to USD transaction was by Martti Malmi in 2009. He sold 5,050 Bitcoin for $5.02 to New Liberty Standard by using PayPal [3]. Whereas, the first transaction between two computers using were between Satoshi Nakamoto and Hal Finney[4]. However, the first real world transaction was by Laszlo Hanecz, who purchased two pizzas for 10,000 Bitcoins [24].

2.2 Blockchain

Bitcoin works based on a principle of a public ledger called Blockchain [8]. Blockchain is referred to as a decentralized and distributed database. It contains Blocks each of which consists of Bitcoin transactions [1] whose information are publicly visible [3] but it provides security using

[2] https://coinmarketcap.com/
[3] https://twitter.com/marttimalmi/status/423455561703624704?lang=en
[4] https://www.washingtonpost.com/news/the-switch/wp/2014/01/03/hal-finney-received-the-first-bitcoin-transaction-heres-how-he-describes-it/?noredirect=on&utm_term=.cad7daa06af2

blockchain technology [7]. Blockchain is also referred to as chain of blocks as each of the block is linked with the previous block [3]. The blockchain information is stored in a network of personal computers termed nodes. Whenever, a request for transaction appears, all nodes in the system are notified to access the information and process the transactions further [7].

2.2.1 Blockchain Data Structure

Blockchain has blocks which can be uniquely identified using block hash or the block height. A particular block contains metadata. These metadata are the parent hash, difficulty, timestamp, and nonce and Merkle root [25]. The Merkle Tree root is the generalized data structure that is used to summarize all the bitcoin transactions in the block. The Merkle Tree assists to prove the integrity of transactions in a block. The Merkle tree is build up using a "pair of nodes". These pair of nodes are repeatedly hashed till only one hash called the root or Merkle root [13] remains; as depicted in following *Figure 2.1*. This is used as the data structure for this study.

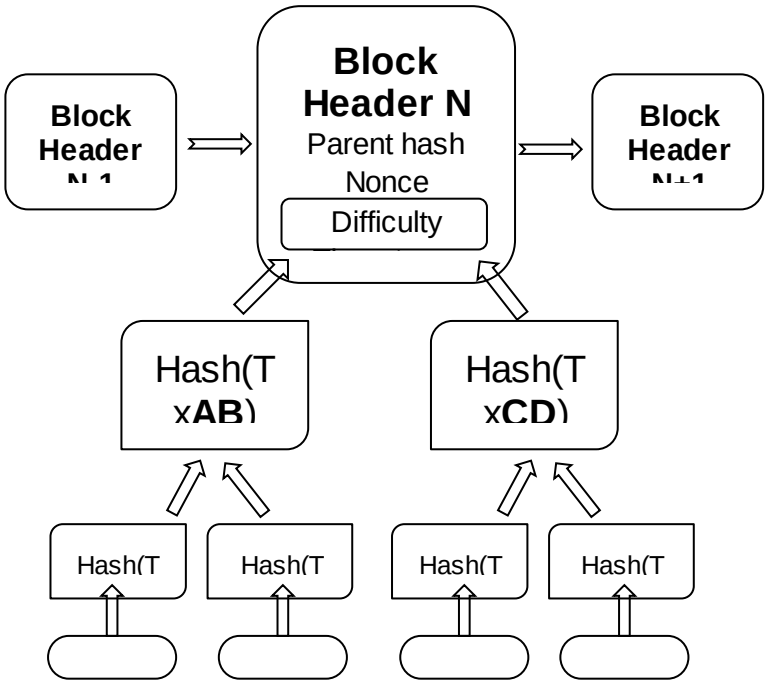


Figure 2.1: Blockchain Data Structure

2.3 Transactions

Bitcoin payments or transfers are carried out by generating transactions. All the transactions are recorded inside the blocks and the blocks are in form of merkle tree [26].

2.3.1 Input and Output of Transactions

A Bitcoin transaction has an input or a set of inputs pointing to an output or a set of outputs [27]. Both inputs and outputs are linked to Bitcoin addresses each showing an amount of Bitcoins transferred to a receiver address [25] where the output of one transaction is used as input for later transactions. The network does not allow spending coins that have been already spent. A user would spend an output by providing proof of ownership of the address. This proof is the private key respective to the address [27].

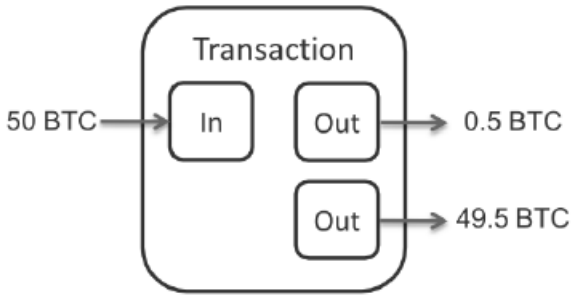


Figure 2.2: A transaction with single input and multiple outputs

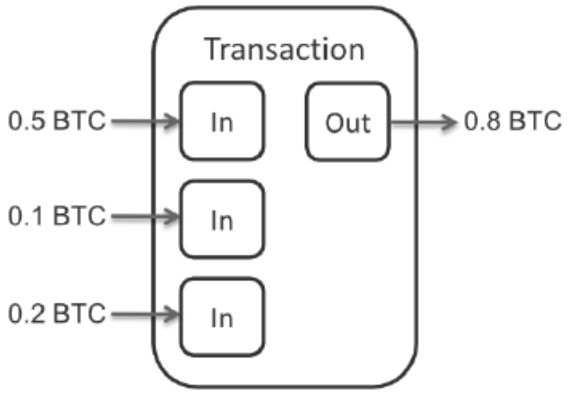


Figure 2.3: A transaction with multiple inputs and single output

The total values of the inputs must be distributed to the outputs as shown in the examples of *Figure 2.2* and *Figure 2.3*. There may be instances where the user has to send Bitcoins by combining a number of inputs of value equal or larger than the required payment in a transaction as in *Figure 2.3* which shows a transaction with multiple input Bitcoins for a single output. A transaction can also consist of multiple inputs and multiple outputs [28].

In blockchain, for the transaction to be allowed to proceed, the total value of the outputs should be less than the total value of the inputs. But, in case if the value of the inputs in user wallet is larger than the desired output to be sent, then the sender needs to create a transaction with two outputs namely the actual payment and 'change'. The actual payment is sent to the receiver address. The change is sent back to sender to himself a Bitcoin address controlled by sender [27]. For example in *Figure 2.4*, If Alice has to transfer 0.5 Bitcoins to Bob, but has an input with 50 Bitcoins, Alice will send 0.5 Bitcoins to Bob address and the remaining 49.5 Bitcoins will be sent

back to the Alice using "change address". In a transaction, when the output is still being unspent then it is referred to as Unspent Transaction Output (UTxO)[29].

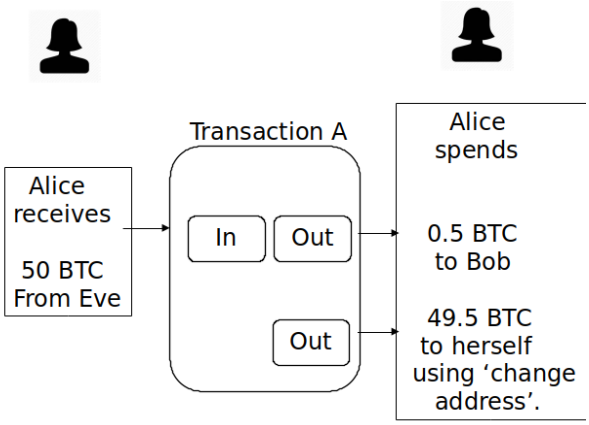


Figure 2.4: A transaction with change address

2.3.1 Bitcoin Mining

The transactions that occur are processed further only after solving a computation problem by the miners. This is done to include the transactions as confirmed into a block in blockchain. This process is called mining [30]. During this process, the transactions between Bitcoin users are registered, validated and maintained via the entire network which is called Bitcoin mining. This mining process enhances the data reliability. Since, the protocol makes sure that there is no possibility of fraud while conducting transactions [1], [3], [8]. As no any single node in the system can manipulate the data or try to change it since all the other nodes in network have access to accurate information [7]. Reliability is also ensured due to the usage of cryptographic proof for validation of transactions. During the process of validation, the nodes would ensure whether the Bitcoin belongs to the actual sender and it is still unspent. Once that information is validated, the transactions will be included in a block. Thereon, the block gets attached to a previous block. This chain of blocks refers to 'Blockchain' [1], [8].

The Bitcoin Miners those who are doing the mining process will be rewarded with Bitcoins. This is how new Bitcoins are added to the Bitcoin system. Initially, bitcoin miners were rewarded with 50 Bitcoins for mining. Thereafter 25 Bitcoins were offered. Currently, 12.5 are being rewarded [7]. In addition, when initiating a transaction, the buyer and seller can also pay a

‘transaction fee’ to the miner as a bonus for verifying the transactions. However, these fees are optional, but 97% of the transactions in 2014 include a transaction fee. This fee will also depend on the standard client software, but normally it is 0.0001 bitcoin (usually less than 0.1% of total transaction value) [3], [16].

A heavy amount of energy is consumed during mining process. Mining process is the main requirement to keep Bitcoin usage going. In the white paper, Satoshi Nakamoto proposed Bitcoin as a peer-to-peer network using proof-of-work to record the history of entire transactions publicly without relying on trust and prevent double spending. So it is computationally not practical for an fraudster to change the transactions [1] and spend same Bitcoin on more than one transaction [9]. So this protocol would not waste the energy as the honest nodes are in control of majority of CPU power. Therefore it argues the energy consumption as a way of making the Bitcoin transactions sustainable since it avoids the transactions being compromised by fraudsters.

2.4 Wallets

When a user wants to sell or buy Bitcoin, they needs to have a wallet which stores the private keys. Bitcoin wallets are stored in several ways such as desktop, mobile, webs, and hardware wallets. If a user uses an exchange, it creates wallet for every user that are stored in their system [30].

2.4.1 Bitcoin Addresses

Bitcoin addresses are used while performing Bitcoin transactions [26]. Each of the Bitcoin address is mapped to a unique public and private key pair which helps to transfer the ownership of Bitcoins among different Bitcoin addresses or users. Generally a user has hundreds of different Bitcoin addresses which are usually stored in their digital bitcoin wallet [6], [25]. Each address is 160 bits long. Therefore, even though the addresses are generated randomly, overlap or collision in addresses does not occur. But it is not necessary to always generate unique addresses. The addresses that are used only once are termed as disposable addresses. Bitcoin addresses can be reused as well [27]. But, reusing Bitcoin address is traceable because the flow of Bitcoin can be traced from one known or unknown address to another [18] which leads to privacy leaks. Therefore, Bitcoin community and previous researches has encouraged using a different Bitcoin address for every

transaction [27].

2.5 Features of Bitcoin

The usage of Bitcoin became significant due to the popular reasons such as speed [5], anonymity, security, convenience [6] and decentralized system with less transaction cost. The reduced cost is because there is no middle party involved to control the Bitcoin in comparison to traditional payment methods [1]. But in [7], the author critique that while Bitcoin exist as a decentralized system, it should have a formal structure, rules and a proper line of communication for better management. But Bitcoin lack these requirements. It lack legal explanation in between the Bitcoin user community and Bitcoin exchanges and no coordination among the exchanges as well [7].

Even with bitcoin lacking of formal structure, rules and a proper line of communication it has been made legal in countries such as Australia, Brazil, Canada, France, Germany, Italy, Japan, South Africa, Turkey, UK, US, Chile, The Netherlands, Singapore, Spain etc. According to Cryptocurrency G20(Summit on Financial Markets and the World Economy) results[5]. These countries have adopted to use Bitcoin and encouraged their people. In China Bitcoin has been banned including all crypto currency trading [5]. Bitcoin is banned or its usage is discouraged in Bangladesh, Bolivia, China, Ecuador, Iceland, India, Russia, Sweden, Thailand and Vietnam [10]. Some of the countries as Mexico, Russia, South Arabia and South Korea have titled Bitcoin as restricted in regulation of crypto market. However, some countries such as Argentina, India, Indonesia, Jamaica, Rwanda and Senegal are waiting for the world's approach to make an informed decision [5].

Regardless of several criticisms and concerns on the legality of Bitcoin, currently not only online businesses, but also traditional retailers are also beginning to accept Bitcoins as a payment method [6]. By June 2011, Wikileaks started accepting bitcoin as donations and now many companies like Expedia, Microsoft are accepting Bitcoin [31].

[5] https://g20.org/en/summit/

Even though Bitcoin is accepted as a payment method by several parties, the value of Bitcoin is highly volatile. In 2013, the market price of a Bitcoin fluctuated between $13 and $1200 USD [6] whereas according to CoinDesk bitcoin Price Index in 2017, the price of Bitcoin fluctuated between $1,000 to nearly $20,000 [31]. Regardless of it volatility, it is shown in [2] that users who use Bitcoin for purpose of exchange do not react to past returns or volatility, but the investors only concern the volatility in Bitcoin returns.

The volatility of Bitcoin is not completely a disadvantage. In [6], the author mentions that while Bitcoin is volatile; it provides financial stability when a fiat national currency is unstable. For an example, in 2012 the Iranian Rial was undergoing a hyperinflation and a scarcity of USD in Iran. At that instance, the value of Bitcoins was considered to be more stable than Iranian Rial. The funds in bitcoin were easily transferred to Iran over the internet [6].

Bitcoin provides numerous benefits for its users. In addition, on a positive note the author in [11] mentions that since Bitcoin works through a peer-to-peer network, they are not subject to bank or third party regulations thus making them less influenced by economic or political issues that affect other currencies. However, Bitcoin also has its own risks for users such as fluctuations in exchange rate, Bitcoin exchange hacks [16], payment market such as banking sector, financial markets, securities exchange companies, payment processors, and trust and escrow companies disruption [30].

2.6 Illegal Activities

Technology is referred to a double edged sword. Bitcoin is an example of using this sword in both positive and negative perspectives. Bitcoin can be used as an investment or as a medium of exchange. In [2], the author concludes that Bitcoins are used as more as a speculative asset than a currency. However, on any purpose it contains both different legal and illegal intentions for both users and service providers [9]. The illegal activities related to Bitcoin cover a wide range of crimes such as murders for hire, funding terrorism, drug or weapon or organ trafficking, Ponzi schemes, forgeries, unlawful gambling, money laundering, illegal mining, computer hacking, spreading ransomware and outright theft [8], [10], [33]. Due to this unlawful involvement of bitcoin the public trust in Bitcoin is getting declined [20].

At least 25% of Bitcoin users and around 44% of Bitcoin transactions are associated mainly with illegal activities. It is discovered that around 24 million Bitcoin users use Bitcoin primarily for illegal purposes [9]. Whereas as per a recent study in [8] reports that one quarter of Bitcoin users and one half of Bitcoin transactions are involved in illegal activity. In addition, it also highlights that approximately $72 billion on yearly basis involves Bitcoin.

As per recent study [8],a highlightable fact is that illegal users tend to transact more in smaller amounts repeatedly with a certain party to avoid getting noticed. In addition, it is noted that the illegal users are holding less Bitcoin due to Bitcoin seizure incidents by FBI [8]. This is supported by the finding in [9], where the authors highlight that the users who are spending Bitcoin on illegal goods had about 25%-45% more Bitcoin (with the 95% confidence interval) than those who doesn't spend Bitcoin on illegal goods.

Following is a literature review on some major illegal activities that involves Bitcoin.

2.6.1 Scam

Scams based on Bitcoin can be classified into mainly four groups such as high yield investment programs or ponzi scheme, mining investment scams, scam wallet and scam exchanges according to a classification identified by studying 192 scam incidents [14].

2.6.1.1 Scam wallet

This category includes fraudulent services that pretend to be as real Bitcoin wallets. This provides several features of an online wallets while the operator clear some proportion or all of the currency from customer wallet. For example services such as Onion Wallet, Easy Coin, and Bitcoinwallet.in transfers to innocent customers wallet were literally delivered to scammer's address [14].

2.6.1.2 Scam exchanges

This is a type of insider scam where, fraudulent exchange operators steal user funds by themselves. The legal action on exchanges can also lead to seizure of Bitcoins, thus causing a loss of

funds for the users. As, sometimes it will not be clear as to which is the main reason for the breach [34].

The exchanges of the scammers have at most existed for about three months. But, they are also known to pull out very least amount of Bitcoins from innocent party. However, exchanges such as CoinOpend and btcQuick were in operation for less than a month, but they pulled out about hundreds of thousands of dollars from its customers [14].

2.6.1.3 Ponzi scheme

High-yield investment programs also referred to as ponzi scheme is an online version of a financial scam in which investors are promised extremely high rates of return on their investments. The returns to existing investors are made from the reserve deposited by new investors [35]. Generally, in theory, Ponzi schemes collapse when withdrawal requests exceed the cash reserves of the service. But, in practice the operators find it more profitable to scam once the deposits are made. For example 'Leancy and Cryptory', the scheme stopped paying as soon as the funds stopped flowing in. These operators failed to maintain appearance of credibility by complying withdrawal requests after new deposits stopped, but they chose not to and ended up scamming users [14].

Though there are no broad thought of legal restrictions towards scamming in global level, some Bitcoin exchanges in countries with greater emphasis on anti-money laundering (AML) are regulated by financial regulators. Thus, exchanges may be pressured into shutting down [15].

However, the investors are generally aware of the fraudulent nature of their investment; they are nonetheless being masterfully deceived by con artists. In addition, it is reasonable that some innocent users believe in the credibility of the investment made [35].

2.6.2 Theft

Over one-third (⅓) of money in the Bitcoin system was lost [7] due to Bitcoin being vulnerable to software hacks and network based attacks [3]. These attacks are commonly termed as Cyber Attacks referring to any action that violates the security of the exchange system. Cyber-

attacks on Bitcoin wallets can be due to security flaws in system, mistakes of Bitcoin users such as negligence or ignorance and a Denial of Service (DoS) attack [7].

Both security flaws in exchange system and mistakes of Bitcoin users would lead to a hack or theft of Bitcoins [7]. Thus theft is very popular among cyber criminals [20] whereas a DoS (denial-of-service) attack refers to an explicit attempt by a malicious party to deny legitimate access of a service for users [19] and disrupt the operation of the Bitcoin exchange and lower the monetary value of the Bitcoin, even though attackers cannot steal Bitcoins. Sometimes DoS attack is performed by attackers before purchasing Bitcoins by reducing its value making it volatile, whereas some threaten the owners of the exchanges to pay them money to bring an end to the attack [7]. In addition, the fact that an exchange is surviving in a competitive market with a considerable volume of transactions also increase the possibility of experiencing a breach. As hackers find it worthwhile to launch an attack on that particular exchange [30]. The Bitfloor exchange was exposed to a security breach when unauthorized individuals obtained access to the backup private keys that controls cash flow accounts of the exchange [34].

Another significant fact in [3], [15] is that the transactions are not reversible in blockchain which is a disadvantage because it is impossible to correct errors occurred due to a theft. But acts as an advantage to the dishonest party thus allowing funds being stolen or taking without the permission of Bitcoin owners [15].

However, in few years' time Bitcoin has the likelihood of being recognized as a common medium of exchange. This change would make theft and scam to be less disturbing for the Bitcoin community. As, the need for trading fiat currency against Bitcoin reduces which in turn creates a reduction in demand for Bitcoin exchanges. Thus reducing, the number of investors and traders in exchanges [30].

2.6.3 Darknet

There are mainly four methods of accessing websites namely surface web, deep web, dark web and darknet. Surface web could be accessed by any standard web browsers available. Deep web can be used to access websites or content using search engine if the link is available [36]. The

websites that are with the data intentionally hidden through any standard web browsers is termed as dark web where there are lots of illicit activities occur. Darknet refers to a network that is encrypted and existing on internet which can be accessed only by using some special browsers [11]. As in [36] the research results prove 57% of content in darknet is illegal, whereas 47% of all bitcoin transactions involve illegal trading on darknet [8].

According to [36] terrorists use darknet as a tool for fundraising, money transfers and illegal purchases of explosives and weapons mainly using Bitcoin for payment. One example for a popularly known darknet marketplace was 'Silk Road' which was operating via hidden network called TOR [9], [37]. It is estimated that this service earned US $1.2 billion in Bitcoins [36] and generated millions of USDs as a commission. However, as per [11], Ross Ulbricht the main operator was traced down and seized by FBI in October 2013. In addition after the closure of Silk Road, Silk Road 2.0 emerged in year 2014 [11]. Thereafter, other marketplaces such as Evolution and Agora evolved quickly [12].

There are numerous marketplaces operating through TOR browser and I2P which is operated as a hidden service in deep web [10]. There are mainly four factors that would facilitate darknet businesses for both seller and buyers; namely identity and flexibility, anonymity, ease of associating in cyberspace, and lack of deterrence [11]. However, as in [7], [38] authors highlight that mainly anonymity of bitcoin transactions give criminals as an enabler tool to operate without getting noticed by legal authorities.

As observed in numerous real world incidents that prove some criminals use only Bitcoin to conduct illegal activities whereas the study in [10] indicates that the same idea will be applicable even for cash transactions conducted using fiat currencies. Thus highlighting less necessity to implement additional rules and regulations especially for Bitcoin.

2.6.3.1 Marketplace scam

While darknet activities are illegal, some of the darknet marketplace operators chose to carry out an exit scam. Generally, all the payments made by customers are held in 'escrow' by the marketplace. Vendor prices on darknet markets are often quoted inclusive of a marketplace fee. The escrow system is to reduce any quarreling in between buyers and sellers and reduce scams. This will

remain with and managed by market place until buyer notifies that the goods or service have been received. This escrow system pave path to exit scams, where the operators disappear along with Bitcoins held in escrow. Some such exit scams are Sheep Marketplace in 2013, Pirate Market in 2014, Evolution in 2015 and Nucleus in 2016 [8].

2.6.4 Ransomware

The advanced encryption algorithms are used to protect valuable information in the businesses [39]. But on another perspective, ransomware are generated using advanced encryption algorithms in order to negatively affect an individual or business. Ransomware are similar to other computer virus such as Trojan horse, worms and spyware [39], [40] and it is defined as the emergence of cyber hack jacking threat in new form in the cyberspace. Ransomware also referred to as crypto virus [39]. A main feature of ransomware is asking for a ransom after files has been encrypted [1], [18].

Generally ransomware can be divided into two basic types such as crypto ransomware and locker ransomware. Crypto ransomware encrypts files and data whereas locker ransomware locks the computer or the device by avoiding the victims from using it [16].

Ransomware has become a significant problem [16] due to its rapid growth in global level [41]. The ransomware attacks were growing fast and significantly since the period of Bitcoin came into usage [16],[42]. Whereas in [42] the researcher mentions one of the main reason for the growth of ransomware is due to the increasing ease of use of Bitcoin systems for payment purposes. As of now it has been an efficient way of earning money directly [17], [42]. Users are preferring to pay the ransom since they are desperate to get their data back [40].

However, according to A. Kharraz et al. about 2 of the 15 malware family were involved in Bitcoin payment transactions from period 2006 to 2014 whereas a considerable 88.22% of ransomware samples used prepaid online payment systems such as Moneypak, Paysafecard, and Ukash cards as a payment method [41]. However, according to [17] the limited geographic availability of these payment methods led the path to conduct payments by Bitcoin which is an international payment method. Another fact is that the attackers switched to using Bitcoin is to avoid reversing transactions or track them [17] and to protect their identity [41].

Few examples would be CryptoLocker [17], CTB Locker which was the serious malware during 2016 [39]. In CryptoLocker, they required the affected party to pay the criminal the ransom with bitcoin by providing their bitcoin address and within the time duration as they defined. Generally, criminals store the decryption key in their server in order to decrypt the files after the victim pays the ransom. It has been identified 771 ransoms, for a total of 1226 Bitcoins. However, some bitcoin addresses receive Bitcoins for one payment as well as they are reused for several payments [1], [17]. In addition, according to [43], there is an existence of connections between CryptoLocker to bitcoin services namely Bitcoin Fog and BTC-e, and to the Sheep Marketplace scam happened in 2013. Later, CryptoLocker 2.0 a modified version was released. It was written in a different language than CryptoLocker and was believed to be released by different attackers from previous version [16]. Meanwhile, the CTB Locker was the serious malware ever during 2016 [39].

The victims of the ransomware are both individuals and business [16], [44] but with the growth of ransomware, the victims being targeted are also changing. The companies are being targeted since ransom values get higher [16], [42]. One such significant example would be Hollywood Presbyterian Medical Center case in US which was about a crypto ransomware. They ended up paying 40 Bitcoin ($17,000) to recover from attackers in Turkey [42]. In addition, it is also revealed in [16] that ransomware attackers are focusing the countries that have weak or reduced law enforcement.

In order to spread ransomware, mainly social engineering techniques are carried out. In addition it is also spread by downloads, visiting unnecessary websites or clicking on a malicious link or ad. Moreover the removable drives such as USBs or portable desktop or laptop drives are also a possibility for attacks [39]. Thereby, security practices on Bitcoin significantly depend on how well users protect themselves from common cyber threats such as malware, social engineering, negligence and data corruption [3]. Especially carelessness is the main reason for losses [45].

A highlightable finding in [18] indicates that some ransomware attackers directly sent the ransom payments received to known parties such as exchange services and gambling. In addition, it also revealed that some ransomware attackers has specifically transacted multiple times with the same party [18].

2.7 Tainted Coins

Tainted coins are Bitcoins which has involved in some sort of crime [46]. If a bitcoin address is tainted, it is visible across the network. This is due to the digital signature mechanism in Bitcoin. The publicly available transaction history can be used to examine how a tainted Bitcoin behave in the network [47]. When a Bitcoin user receive Bitcoins from a sender, the Bitcoin user can check whether the receiving Bitcoins has involved in fraudulent activity in past. Thereafter, determine whether to continue the transaction with accepting Bitcoin or not. [46] This exception has led to deflation of the monetary value of the some Bitcoins since honest users are reluctant to receive tainted Bitcoin [46] due to the probability of refusal by the next user [4], [47]. An example would be Mt. Gox, a Bitcoin exchange based in Japan locked Bitcoin holder's account with tainted coins after an incident of theft where 43,000 Bitcoins were robbed from another Bitcoin trading platform Bitcoinica [47].

The untainted coins are the ones that are freshly mined [4]. According to [4], the more 'tainted' the chain of transactions is, the stronger the link in between the Bitcoin addresses is. For example, if my wallet is stolen, whenever the robber tries to bank the money at an exchange, the criminals can be arrested [48].

However, by performing CoinJoin transactions, it enables the users to move their Bitcoin from tainted inputs to untainted output addresses. In addition to coinJoin, if users exchange their output by an anonymous broadcasting method, it is not possible to link inputs to outputs even by dishonest users in the mixing strategy [49].

According to [29], the author highlights theoretically, it is possible to trace tainted coins and transaction history since the transaction point to previous transaction in the blockchain. Therefore, in order to analyze these tainted coins there are three algorithms namely poison, haircut and FIFO. Poison is where all the outputs are completely tainted by all inputs. Haircut is when the output is tainted by the percentage of input value tainted. The third method FIFO states the withdrawals from an account are considered to be realized against the deposits first made to it [29].

2.7.1 Strategies

Illegal users practice some strategies such as using CoinJoin [18], [27], Mixing Services or tumblers [18], [50]. This is performed in order to increase the complexity of tracing them via transaction chain by identifying Bitcoin addresses accordingly. In addition, as per [51] users also proceed with strategies such as long chains and fork merge structures in the transaction graph.

However, practicing these strategies does not ensure un traceability. A previous study in [51] discovers that all the isolated large transactions in the system had a close relationship to a single large transaction belonging to November 2010. This was discovered even with the long chain and fork merge structures used. In addition, if the groups are tallying with tags from external sources such as forums or websites such as blockchain.info, walletexplorer.org, it is clearly possible to trace or to de anonymize a larger section of the Bitcoin transaction network. Even if any strategies are used but the address is tagged [18].

2.7.1 CoinJoin

CoinJoin is a solution for the privacy problem of pseudonymity. This does not require changing the protocol of Bitcoin system. In this, several inputs of Bitcoin from different users combined into a single transaction to spend it to several other users [27]. The Figure 2.5 below summarizes the process of CoinJoin [29].

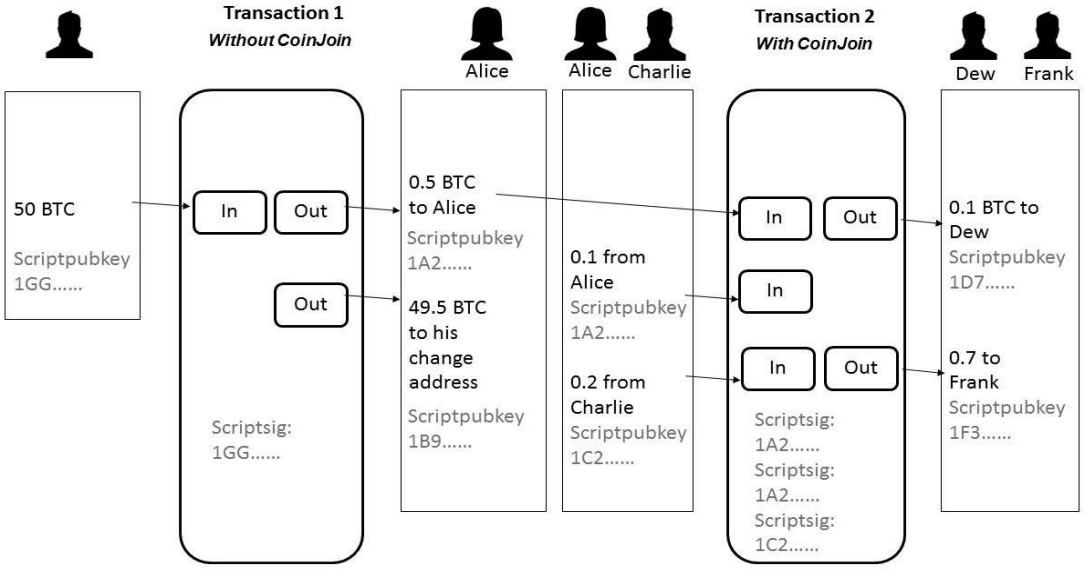


Figure 2.5: CoinJoin Process Summary

According to a recent study [49], the coin mixing is the most prevalent approach among Bitcoin community. It is encouraged to maintain anonymity. However it is of less usage in the present time due to its expensive nature.

2.7.2 Bitcoin Exchanges

A Bitcoin Exchange is an online platform which converts between Bitcoins and freely convertible currency [34]. Some of the exchanges behaves similar to a bank where they offer fixed interest on the customer savings. There are popular exchanges such as are Bitstamp, Kraken, Coinbase etc. [7]. The exchange stores all of customer details including bank information. When a customer wants to sell Bitcoins via an exchange, they have to first transfer those Bitcoins to their wallet in the exchange. The exchange create a wallet for every customer in their system and one can sell or buy Bitcoins thereon [30]. The exchanges charge customers for the services they provide. The services one can expect from an exchange are buying, selling and maintaining the bitcoin wallet of the user. A transaction fee is charged from users by exchanges. The transaction fee provided to miner is for validation in blockchain. But, the transaction fees paid to exchanges are for the services it provides. The fees vary from exchange to exchange. Most of the Chinese bitcoin exchanges charge fixed transaction fee of 0.2% [7].

Exchanges ease the process for a bitcoin user however; it has a major risk of getting hacked by attackers [16]. In [7] author concludes on recommending exchanges to clearly disclose all the details of the cyber-attacks on them to their customers. Thus leading to better transparency in the way they operate. But Bitcoin exchange rate remains unchanged from disclosed hacks and scams. This indicates regardless of the risk bitcoin exchanges carries, it manages to self-sustain its exchange rate [15].

Some examples of the major attacks were Mt. Gox attack losing 450 million dollars, attack on Bitfinex exchange leading to reduction in value of bitcoin by 23% and DDoS attack on Bitfinex and Bitcoin-e Exchanges [7], Bitfloor loss of 24,000 Bitcoins in an attack [30].

Trading Bitcoins via an exchange gets recorded in the trade history of that specific Bitcoin Exchange. Therefore, transactions recorded in the blockchain are recognized as those were sent from exchange because blockchain records only the real transactions which are deposits and withdrawals of bitcoin. Thus, the real owner cannot be accurately identified as the transactions details of the owner are present in the exchange's database [30]. This overcame one of the main risks mentioned by Satoshi Nakamoto to a certain extent. He mentioned that revealing the owner of a key is a risk since it could reveal other transactions that were carried out by the same owner [52].

2.7.3 Bitcoin Mixtures

Bitcoin Mixing Services or tumblers are where a specialized middle party breaks the link between sender and receivers addresses. This is done by mixing transactions such that the connection of input and output transaction is not identifiable from a third party view [18].

Mixing services are frequently used to conduct illegal activities such as the trading of illegal goods. For example, Silk Road was known for trading variety of drugs, harmful code, hacking technologies, bank card details and theft accounts. In addition, Ransomware attackers have enforced victims to pay them in bitcoin by a certain deadline to restore encrypted files. Such cases also include the use of a mixing service to avoid tracking by investigative agencies. However, according to study [50], a generic de mixing algorithm has been designed to find input output relationships among the mixed transactions.

Some examples of mixing services known to be linked with ransomware attackers are BitcoinFog.info and Helix Mixer. Helix Mixer is one of the most widely used mixing services. The mixing services also play a significant role in money laundering and illegal activities using bitcoin [18].

Chapter 03

Methodology

3.1 Outline of the Problem

Bitcoin is a medium that several individuals or groups adapt to it in illegal intentions. As per the research context, previous study addresses by providing a wider analysis of illegal activities. This research study was focused on analyzing the behavior of the illegal users as to highlight how they transact in the network bitcoin. Thus, intending to differentiate honest users from dishonest users. This study also assist legal authorities, officials, bitcoin miners and other interested parties to take decisions and implement controls via rules and regulations.

3.2 Research Approach

A broad analysis of the behavior of bitcoin illegal users paved the path to study in this context. The beginning of the study focused on studying about specific incidents and then moved towards formulating general patterns on user behaviour. Therefore, theoretically an inductive method was used in the research. Detailed information on illegal incidents was first gathered and preliminary patterns from separate specific incidents were obtained. Finally, generalized conclusions were produced based on the individual patterns and facts derived from the analyzed cases. Therefore it can be highlighted that an inductive method was applied as this study developed a theory of how the illegal bitcoin users are transacting in the network.

However, some of the findings of previous researches also get confirmed in this study. Thus, a deductive approach was practiced during the evaluation of the already defined patterns by prior research studies. In brief, this study was a combination of both inductive and deductive research approaches.

3.3 Research Design

The research design was the guideline for the study to address the research questions. Research design (Figure 3.1) shows the important stages regarding the approach to answer the research problem.

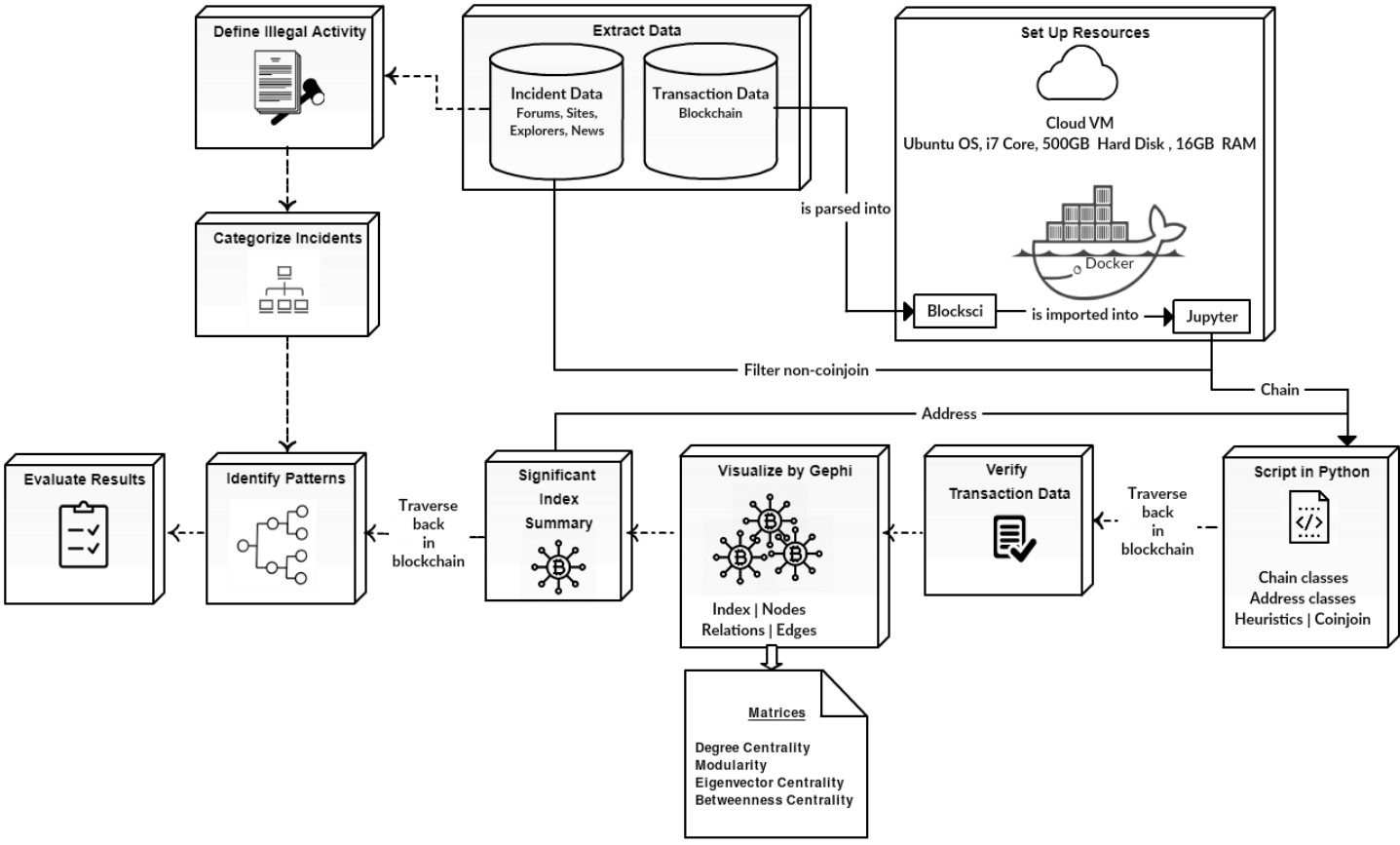


Figure 3.1: Research design

There are mainly two data types in data extraction phase. The incident data was extracted from publicly available data sources. Whereas transaction data was from the Bitcoin Blockchain. The extraction of the incident data led to the definition of illegal activity so which helped to categorize the incidents.

On parallel to that, the resources required for analysis were set up. During the process of resource setting, blockchain data was parsed into installed tool BlockSci. Thereafter, BlockSci was imported into the notebook, Jupyter to run python scripts according to the parameters. The scripts

included filtering of non coinjoin incidents transactions using heuristic parameter, transaction data data using chain classes in BlockSci was formulated. Thereafter, through traversing back in the blockchain the initial data derived from blockchain was verified. In order to visualize the data, Gephi was used along with various metrics. The data preparation was in the form of indexes as nodes and the relationships in between indexes as the edges. The most significant indexes were then summarized for further interpretation of data. The interpretation of data was carried out using Address classes in BlockSci which consisted of addresses and address types. On the process of interpretation, traversing data back in the blockchain was the method used in deriving patterns and significant facts. Finally some of the patterns obtained were evaluated based on previous findings. In addition, new findings were evaluated using user feedback obtained via a survey.

3.4 Resource Set Up

Ubuntu OS was used for installing Blocksci. The unique features like consistency, security are provided by Ubuntu 18.04 LTS Server Version. Since the Blockchain was approximately of size 157.3GB as at March 2018, it required greater amount of CPU power for running multiple applications parallely. Therefore i7 was chosen as it did have overall high end performance. Thereby, a Cloud Virtual Machine (VM) which allowed the research team to interact remotely was used along with the specifications of Ubuntu 18.04 LTS Server Version as OS, i7 as core, 500GB as Hard Disk and 16GB as RAM. Both Docker, to run as the Operating System level virtualization and Jupyter notebook, to enable integration among research team was used [12].

3.4.1 BlockSci

The Bitcoin Blockchain is described as public chain of time stamped transactions that involve pseudo anonymous identity wallet addresses. As the adaptation on Bitcoin gets higher and higher there are tools developed for analyzing the transactions of Bitcoin in Blockchain [17].

In order to analyse this growing transactions Kalodner et al released a blockchain analysis tool called BlockSci in 2017 [53]. Few main properties of BlockSci are [53],

1. Supports Bitcoin, Litecoin, Namecoin and Zcash.
2. It consists of library of useful analytic and visualization tools.
3. The transactions are transferred on peer to peer network and revealed via a common interface enabling different types of analysis.

BlockSci was used at Princeton for research and educational purposes [53]. In [54], BlockSci library was used to analyze the bitcoin blockchain from 2009 till August in 2017 to examine how the bitcoin usage grew over time.

In addition to BlockSci, cryptocurrency analytics tools such as Chainalysis, Elliptic and GraphSense are also in practise [18].

In this study, BlockSci was used for gaining transaction data from the blockchain. Jupyter Notebook enabled creating, sharing exploring and interacting with the scripts real time [12]. Python programming language was used in writing scripts as BlockSci supports python. Both the BlockSci and Jupyter notebook were hosted on Cloud. The Docker was used in order to enable a single platform that integrated all research team members [12].

During the installation of resources, as the initial stage the blockchain was parsed into Blocksci. Then BlockSci was imported into Jupyter notebook, thereafter "chain class" was created to pull data from Blockchain. This class was constructed by parsing a string that represented the file path to the BlockSci data files generated by blocksci_parser. As shown in *Script 1*.

```
import blocksci
chain = blocksci.Blockchain("/root/blocksci-data")
```

Script 3.1: Importing Blocksci and Creating Class "Chain"

3.5 Data Extraction

There are two types of data required under data extraction section. One of them is data on illegal incidents involving Bitcoins. Since blockchain is decentralized, there is no one central entity to request for such data or extract from a central repository but data is available publicly. The incident data was obtained from the online open data sources such as Bitcointalk, reddit forums, websites, blockchain explorer, official news sites and data from previous researches. The transaction data was the second type of data required. The transaction data was available on blockchain.

3.5.1 Illegal Incident Data

The attributes such as name of the incident, approximate date incident occurred, value involved in both USD and in terms of coins (bitcoin), transaction ids, bitcoin addresses, nature of incident were collected. For the period from 2012 to 2018, data was collected on 33 illegal incidents. However, due to time limitation and resource constraints 10 of the illegal incidents were used for analysis. The data for period from 2009 to 2014 was available on bitcointalk.org. After the data collection, the data was processed to make it to a homogeneous format. All of the collected data were cross checked for more accuracy with multiple sources. The following are the sources used in extracting incident data for this study.

3.5.1.1 bitcointalk.org

Bitcoin Talk is an online forum available for bitcoin community. The forum is organized into sub forums along with specific topics. Several previous researches have used data from this source for their researches [14], [34].

3.5.1.2 reddit.com/r/Bitcoin/

Reddit is a discussion website used widely for open discussion which allows users to post questions and answers to communities, known as subreddits. Several previous researches has used data solely from this source for their researches [1], [2], [14].

3.5.1.3 bitcoinwhoswho.com

Bitcoinwhoswho is a site which provides all available information about bitcoin addresses. It provides information on wallet name, IP address, last transactions with flagged addresses, keyword search, tags, transaction history and transaction alerts of that specific bitcoin address. The illegal incident details that are obtained are posted as blog post along with the source that was obtained from.

3.5.1.4 Walletexplorer.org

WalletExplorer is a bitcoin blockchain explorer. It provides bitcoin blockchain public ledger of transactions, bitcoin addresses, and bitcoin blocks. It has been used in previous researches as one of main data sources to identify the exchanges and miners. [4], [8].

3.5.1.5 Other

In addition, data was also extracted from sites such as coindesk.com, bleepingcomputer.com. CoinDesk is a popular news site providing updates relevant to bitcoin and digital currencies. Bleepingcomputer is an information security and technology news publication site. These have been used as sources in previous researches as in [3].

The transaction ids of extracted incident data were checked for the coinjoin using the *Script 3.2*. It used basic structural features to recognize whether the transaction id given **(**d656…..c92e [6]) is a coinjoin transaction or not. A Coinjoin transaction does not allow to specifically identifying by matching the sender and receiver. Therefore filtering the transactions which are not with coinjoin was important for proceeding with the analysis.

```
tx = chain.tx_with_hash("d656e3bdde46e9f54107195b13afa757fd9eeb4b1a7f8dc6e210e71224f7c92e")
blocksci.heuristics.is_coinjoin(tx)
```

Script 3.2: Checking Transactions for Coinjoin

3.5.2 Transaction Data

blockchain.info

The data related to all bitcoin transactions was obtained from blockchain from 2009 to March 2018 up to the block height of 514463.

[6] d656e3bdde46e9f54107195b13afa757fd9eeb4b1a7f8dc6e210e71224f7c92e

3.6 Definition of Illegal Activity

According to the knowledge obtained from studying the incident data, a definition for illegal activity was formulated which differs from the general definition for illegal activities.

"Any activity that involves bitcoin which brings a financial disadvantage to one or more parties with or without their knowledge while the opposite party gain benefits financially from its outcome with their knowledge"

3.7 Incident Categorization

The identified illegal incidents were categorized into different sub categories based on the nature of incidents. The nature was identified by using its similar characteristics and setting up boundaries. The boundary lines and categorization was based on understanding the basic categorization of illegal incidents in bitcointalk forum and with reading further in detail about the incidents. The purpose of categorization was to make the process of analysis easier since when there are categories to analyze it is more comprehensive for the mind of the reader to understand clearly.

Even though there were overlaps of the incidents classification, the incidents were categorized under five subcategories in a way that it minimized the severity of overlap.

3.7.1 Boundary Lines of Subcategories

1. **Hack** - The incidents where wallets owned to a particular exchange or a platform were hacked by outsiders thus leading to the collapse of the exchange or platform. Hack has brought financial disadvantage to bitcoin holders since they lost their Bitcoins and exchanges or platform also had to refund users in some cases where they were also affected financially. Moreover exchange would be impacted with negative reputation. In addition, the hackers has obtained the benefit financially.

2. **Ransomware** - The incidents which spread a malware that locked or encrypt databases, files, PCs or any electronic copy that made the owner inaccessible to the resources. Then victims

were requested to pay ransoms in bitcoin to provide the accessibility or promise to send decryption key. This has brought financial benefit to attackers whereas it has delivered financial loss to the victim. Some of the attackers never unlock or provide decryption key to the victim in some incidents.

3. **Known Theft** - The incidents where bitcoin holder knowingly sent bitcoin to criminal because of the fear due to either threatening or blackmail. The bitcoin holder has suffered from the financial loss. The receiving party has got financial benefit.

4. **Scams** - There were incidents where the exchange or the platform has stolen the users' wallet and disappeared by closing down their exchange. In some scenarios, exchange was closed down by issuing a notice by falsely claiming that they were hacked. Sometimes the exchange would make their website unavailable either by issuing a notice or without issuing a notice. This has brought financial loss to bitcoin holders while the exchange has gone away with the money by deceiving.

5. **Fake Agencies** - This was where scammers pretended to be an already existing popular exchanges or government organisation. Their intention was to steal bitcoin by communicating with customers pretending to be honest. This has brought financial loss to victim party while the others be benefited.

After incidents were entered into these subcategories, it was noted that there were similar features which can be used to merge incidents into some main categories. It enabled better analysis in terms of comparison of the incidents.

3.7.2 Definitions of Main Categories

The Main Categories were created based on how attacker has committed the financial loss to the victim party. That is either attacker directly dealt with bitcoin user or as a third party. In addition to this aspect, it was also considered whether the bitcoin user lost his Bitcoins with his knowledge or not.

1) **Hack**- Incidents where the attacker as the middleman came in between the bitcoin holder and the exchange or the platform to cause financial harm to either or both of them without their knowledge [*Figure 3.2*].

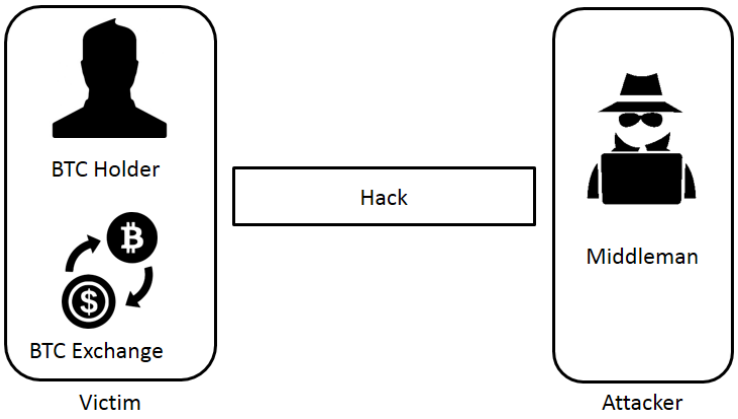


Figure 3.2: Main category 'Hack'

2) **Personal Losses** - Included subcategories of 'Ransomware' and 'Known Theft' where the effect of the bitcoin loss was solely born by the individual or a group of bitcoin users. The bitcoin holders were harmed by the attacker by communicating or dealing directly with them with their knowledge [*Figure 3.3*].

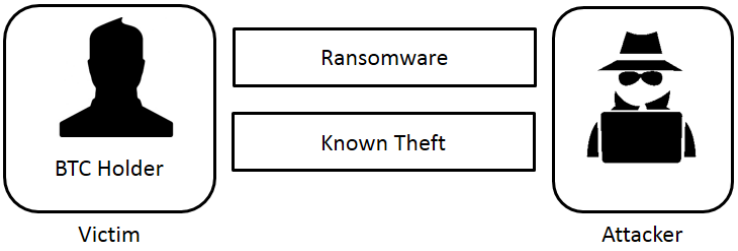


Figure 3.3: Main category 'Personal Loss'

3) **Scams** - Included subcategories of 'Scam' and 'Fake Agencies' where frauds were done by a scammer or exchange or platform itself pretending to be a real service provider. This would led to the financial loss to the bitcoin holders of such exchange without their knowledge [*Figure 3.4*].

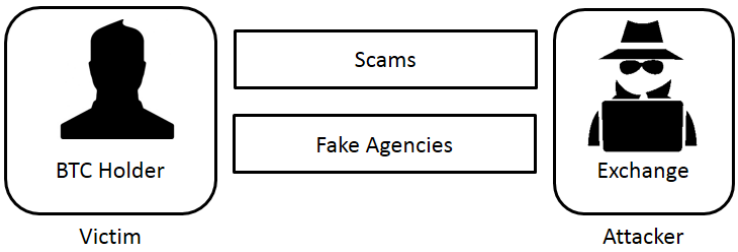


Figure 3.4: Main category 'Scams'

The *Table 3.1* briefs the categorization

Main Categories	**Hack**	**Personal Losses**	**Scams**
Subcategories	Hack	Ransomware	Scams
		Known Theft	Fake Agencies

Table 3.1: Summary of main and sub categorization of illegal incidents

3.8 Main Scripts

The N-ary tree represents the interconnection of transactions with other transactions (*Figure 3.5*). The tree structure created the way by trying out several scripts that suggested many constraints in order to get the required results for this study. The tree structure and the arrangement of addresses are linked together can be described using a random instance in the tree. In (*Figure 3.5)* T2 169...Tig address sends 5 Bitcoins to 1DV...NQs address holding the transaction id 118du…..2a8u7. In T3 1DV...NQs address sends 4.52809038 Bitcoins to 1Nc...ytD address holding the transaction id 7d5uc…..5c2e1 and 1.28 Bitcoins to 1xt...vsn address holding the transaction id 1frtu…..ud2e1. Likewise every output address can spend its bitcoins or else it can keep bitcoins in the address itself without spending as labeled 'unspent'.

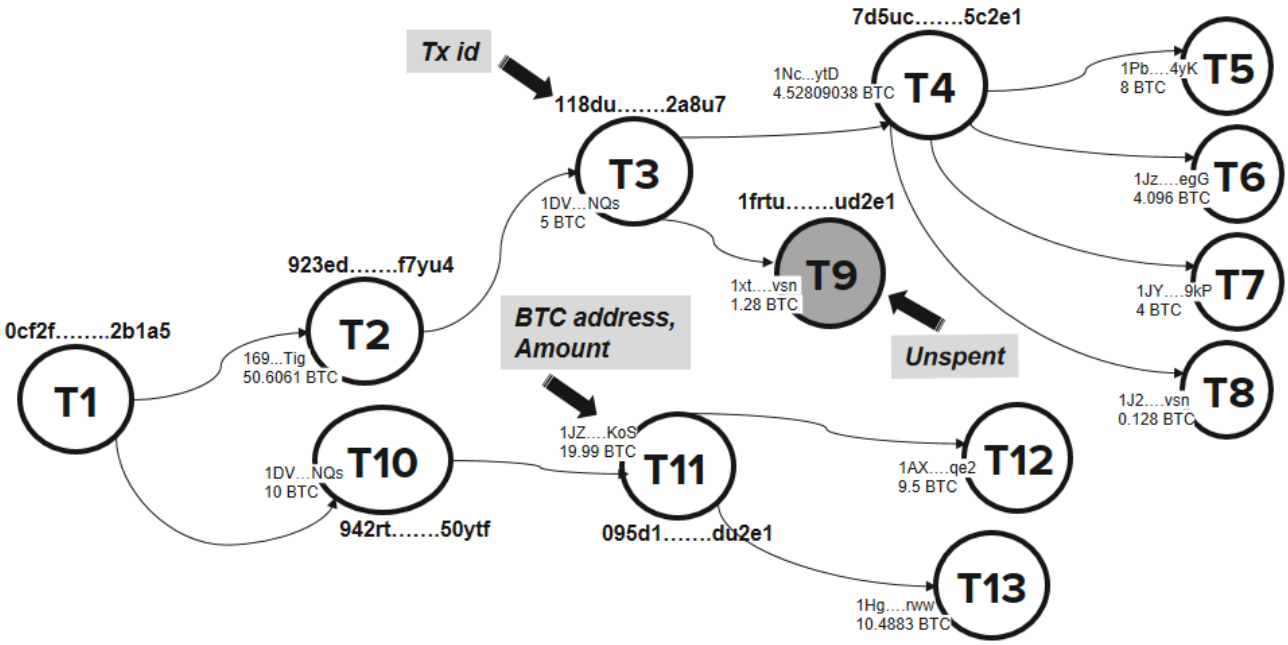


Figure 3.5: N-ary tree chart for bitcoin transactions

After studying the structure of the n-ary tree chart, it was clear to use recursion programming scripts to get the transaction data with respect to the relationship of the chained transactions. As recursion is direct and appropriate when the structure has a similar iterative format.

Therefore recursion function in Python was focused to gain the results on the circulation of the illegal bitcoin input addresses and output addresses initially using the following *Pseudocode 3.1*

Input: Transaction id related to illegal incident; txid
Output: Related address details to a given txid; input address, output address, unspent output address

```
01. extract details from chain tx
02. define results (tx)
03. if any tx has outputs:
04.     for each output:
05.         If any output is spent
06.             write input address, output address
```

```
07.                 do results recursively for every output
08.             else:
09.                 return unspent output address
10. else:
11.     return unspent output address
```

Pseudocode 3.1: Related Illegal Input and Output Addresses for a Given Transaction Id

The script for *Pseudocode 3.1* is as follows *Script 3.3*

```
txid = input('Enter transaction id related to an illegal incident: ')
tx = chain.tx_with_hash(txid)

file=open('./address.txt','w')

def results(tx):
if any(tx.outputs.is_spent):
        for in, out in zip(tx.ins, tx.outs):
                if out.is.spent:
                        data = str(in.address) + str(out.address)
                        file.write(data)
                        results(out.spending_tx)
                else:
                        return tx.outputs.unspent.address
else:
        return tx.outputs.unspent.address

f.close()
```

Script 3.3: Related Illegal Input and Output Addresses for a Given Transaction Id

The results from the script were computationally complex due to the time and memory consumption. In addition, high processing power was required for Gephi Software for visualization. As, some transactions had ScriptHashAddress along with wrapped_addresses with different types of

addresses requirements. Thus, high processing was needed. One such example for the results with different types of addresses is represented in *Table 3.2*.

	Output Address
Result from Script	ScriptHashAddress(344....1Md [7], wrapped_address = MultisigAddress(2 of 4 multisig with addresses 13d….vVrc[8], 1Da….SbB [9], 1Gv….WbU [10], 18V….BAJ [11]))
Result to be Feed into Gephi	344….1Md [7] (only recipient address)

Table 3.2: Different address types in results

When a result was with a ScriptHashAddress, only the first address i.e the recipient address was required for the analysis and the rest of the address are required in order to spend bitcoins as an additional security for the transaction. For example, Multisignature addresses require another user or users to sign a transaction before it can be broadcast for mining in Blockchain. So those Multisignature addresses were not considered with regards to this study. So as denoted in the *Table 3.2,* though the actual result for output address is with 344...1Md [7], 13d….Vrc [8] ,1Da…..SbB [9], 1Gv….WbU [10], 18V….BAJ [11] addresses, the required address for further consideration is just 344….1Md [7] address. Therefore it was needed to adjust the script to be more efficient providing only the relevant information needed for the analysis.

Therefore, scripts were created to get related illegal transaction ids instead of bitcoin addresses was used. Though it solved the problem of computational power to a certain extent, transaction ids resulted out 256-bit hash which had longer number of digits leading to slow server results.

[7] 344pUP56enuGjbPdyubYEqoxB6VaFmD1Md
[8] 13dCNU7T38Ca3zp4mMBSmP6FGyBzq6vVrc
[9] 1DayuQZkBCt4MYYA5Hr8awXvmJDXLndSbB
[10] 1GvFkgaLV69PtrTcMC9XznqcXZxRHWvWbU
[11] 18VibwUc5CNG8TFZNRMSY6LMadED3qGBAJ

Thereafter, a script considering the circulation of the input and output index was focused on. Since, the index is of lesser number (lesser than 9 digits in the dataset). Index represents the sequence order of a given transaction id. Thus, obtaining the indexes as the results was quicker than transaction ids. Thereon further details like respective bitcoin addresses related to that index was obtained if that index represented worth of seeing more insights. The pseudocode (*Pseudocode 3.2*) that was used in order to automate obtaining related illegal input and output indexes for a given transaction id.

Input: Transaction id related to illegal incident; txid
Output: Related address details to a given txid; input address, output address, unspent output address

```
01. extract details from chain tx
02. define results (tx)
03. if any tx has outputs:
04.     for each output:
05.         If any output is spent
06.             write input index, output index
07.             do results recursively for every output
08.         else:
09.             return unspent output index
10. else:
11.     return unspent output index
```

Pseudocode 3.2: Related illegal input and output indexes for a given transaction id

The script for Pseudocode 3.2 is as follows Script 3.4

```
txid = input('Enter transaction id related to an illegal incident: ')
tx = chain.tx_with_hash(txid)

file=open('./address.txt','w')
```

```
def results(tx):
if any(tx.outputs.is_spent):
        for in, out in zip(tx.ins, tx.outs):
                if out.is.spent:
                        data = str(in.index) + str(out.index)
                        file.write(data)
                        results(out.spending_tx)
                else:
                        return tx.outputs.unspent.address
else:
        return tx.outputs.unspent.index

f.close()
```

Script 3.4: Related illegal input and output indexes for a given transaction id

In the *Script 3.4*, first selection construct was used in identifying the outputs if there were outputs available for given transaction id. Then recursion was applied for repeated function calls for each output if there were outputs available for considered output i.e. the output taken into consideration out of all outputs in one term. Recursion gets terminated when the transaction has no more transaction relationship further as denoted in T9 as Unspent [*Figure 3.5*].

Every illegal transaction id available in the data set related to a particular incident was provided as the input to the script. For each transaction id approximately 1.3 million records were obtained in format of indexes and related indexes pairs. The Poison heuristics was used in the study even though FIFO is methodical [29] since the open source tool BlockSci does not possess FIFO implementation.

3.9 Illustration of the Analysis

The following illustrates the technique of how the data related to an instance of an illegal incident was analysed. For the illustration, a hack incident on NiceHash was considered.It happened on December 6th 2017. It had five transaction ids reported. The transaction ids were,

d656e3bdde46e9f54107195b13afa757fd9eeb4b1a7f8dc6e210e71224f7c92e,
60f4666603e8baa87a80821c93924e5c7430a3d2eb84fe65c23140ccf6a8a640,
4b417dcfc06b7ccdb61006d3e7025baa192b53d36cf8e9476edede4bc7f3dac4,
26fe37f72fd8017044719fe2bd8fdd6f85f9eb8e05fa6e1ee873672999d1a0d1,
09cb2313be757d29bba8f99cbdbec2045e5e3705ce10d9e2cb15e559e5119edd.

These above ids were made as an input to the script and the data was appended upto 6.5 million records (1.3 million records per transaction id). This is a sample of output data (index and related index pair) obtained for the incident NiceHash (*Figure 3.6*).

Source	Target
278632091	283732308
283732308	283736830
283736830	283761304
283761304	283768900
283768900	283808512
283808512	283839519
283839519	284110284
284110284	287510062
287510062	298593742
298593742	298596286
298596286	298727558
298727558	298817394
298817394	299166814
299166814	299181059
299181059	300262409
300262409	300561817
300561817	302380903
302380903	302420122
302420122	303545122
302420122	302435156
302435156	302440587
302440587	304222210

Figure 3.6: Output data sample (Index, Related Index) for incident NiceHash

3.9.1 Transaction Data Verification

The output transaction data from the script was verified with the blockchain data by sketching up a N-ary tree manually for the chain. The sketch for the data verification was drawn by traversing back the blockchain as depicted in following Figure 3.7.

A sample of the output shown in Figure 3.6 demonstrates in its first record, index 278632091 has transacted with the transaction of index 283732308. Then in the second record, index 283732308 has then transacted with transaction of index 283736830 and so on.

The index 278632091 assigned for the transaction id 60f4…..a640 [12] has the output address 1En….B4rq [13] (*Figure 3.7.1)* which appeared in the index 283732308 assigned to the transaction id ae8b….180dc [14] (*Figure. 3.7.2*). This is the relationship of the two indexes 278632091 and 283732308 that the script resulted out.

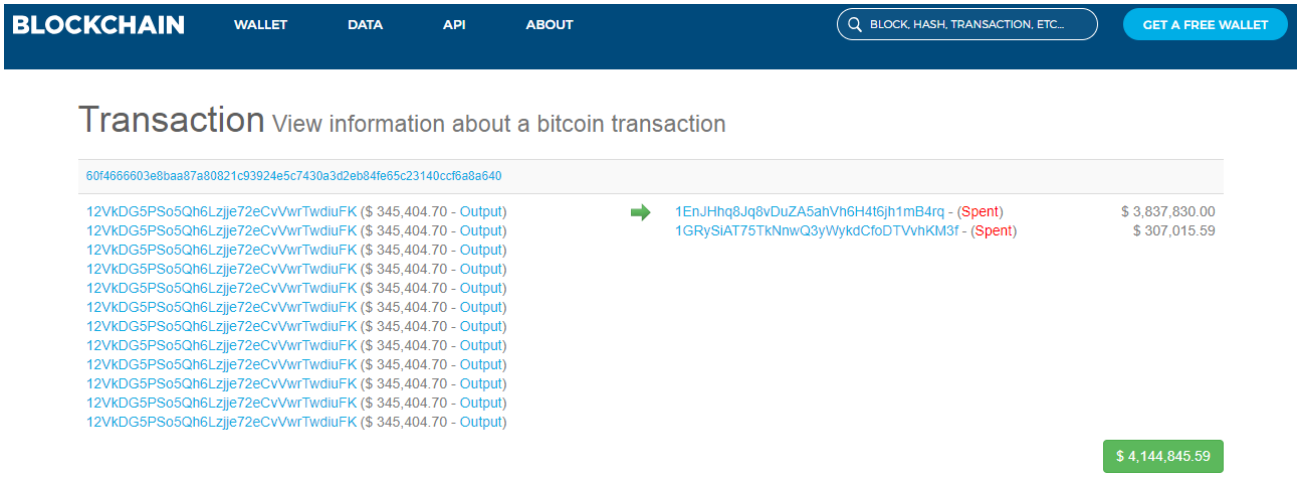


Figure 3.7.1: Transaction data verification

Further, index 283732308 has the output address 1F9….b4K [15] (*Figure 3.7.2*) which appeared in the index 283736830 assigned for the transaction id 9386…..1dae3 [16] (Figure 3.7.3). This was continuously verified for a transaction. Therefore, it was concluded the obtained traversed transaction data via script represents the blockchain transaction data.

[12] 60f4666603e8baa87a80821c93924e5c7430a3d2eb84fe65c23140ccf6a8a640
[13] 1EnJHhq8Jq8vDuZA5ahVh6H4t6jh1mB4rq
[14] ae8b41c1db662b4d8063d11729cb6b438035d8f6094e113490fb2a6b337180dc
[15] 1F9E4dnwYxVidpNfBzjiJssE8ew5kVkb4K
[16] 938654b8386a3cfe8420297b65b36a221153eea15705aa3227d070fe4671dae3

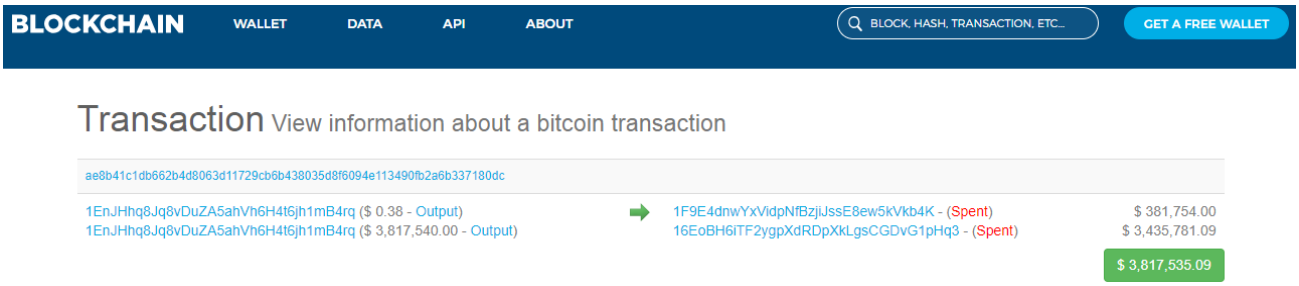


Figure 3.7.2: Transaction data verification

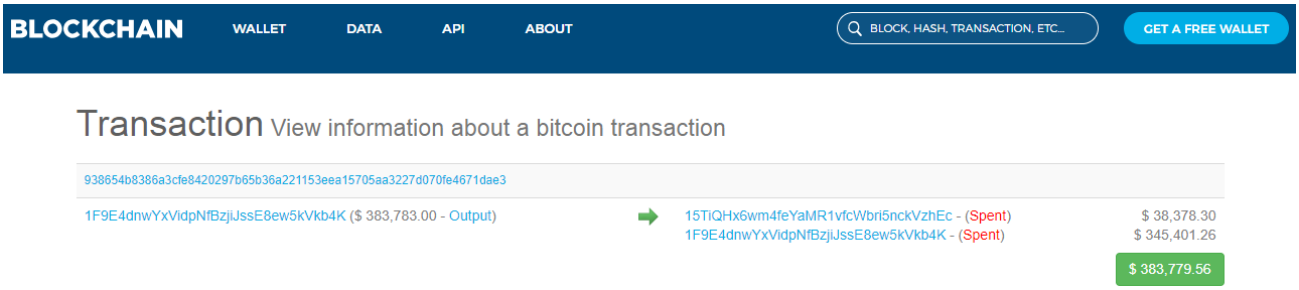


Figure 3.7.3: Transaction data verification

3.9.2 Data Visualization

The traversed transaction data were fed into Gephi. Gephi is graph visualization and exploration software which is available publicly[17]. It has been designed to associate with data in representing it, the structures, and shapes and in different colors to reveal hidden patterns in graphs and networks. This also highlights the outliers. Gephi has been used previously in researches to represent the transaction data therefore, due to its significant features it was used for visualization for this study. Data was fed into gephi, filtered for noisy data and explored the data in the bitcoin network as the initial step towards the analysis of results.

First the data was imported in CSV (Comma Separated Values file) format into gephi. Gephi considered indexes as the nodes, the relationships between indexes as the edges and "Directed" as the graph type since the edges have a direction associated with them. The direction indicates from and to addresses the bitcoins were transferred. The *Figure 3.8* shows the visual that is obtained once data is fed into gephi.

[17] https://gephi.org/

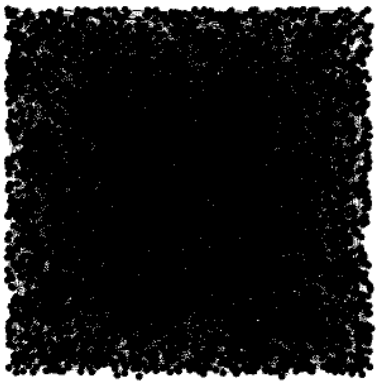


Figure 3.8 : Graph just after feeding initial data into Gephi

Thereafter the ForceAtlas 2 layout was chosen as it is a continuous graph layout algorithm for user friendly visualization as recommended in [55]. Under ForceAtlas 2, parameters such as scale, gravity were adjusted and adjusted in order to prevent overlaps until the visualization got smooth. In addition, the Fruchterman Reingold layout was used to improve the look of the graph [56] The visual obtained after applying layout is as shown in *Figure 3.9*

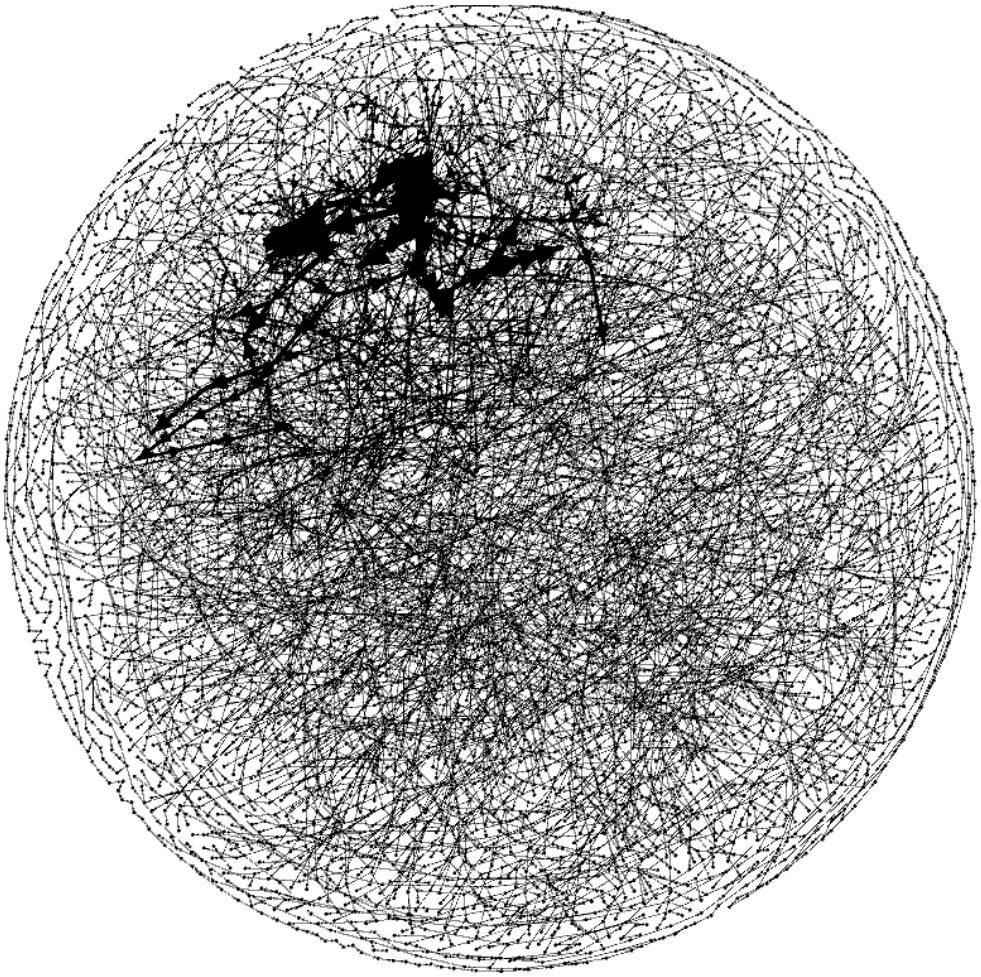


Figure 3.9 : Graph after running layouts

Thereafter, metrics were computed in statistics such as Degree Centrality, Modularity, Eigenvector Centrality, Betweenness Centrality in order to obtain further insights of the transactions and their relationships [56], [57]. During the process the noisy data were removed through filters. Since, it kept only the significant nodes and the active participants in the network to remain. This was done with the intention of identifying the nodes which should be paid further attention to.

In the graphs, the Degree represented the number of connections held by each illegal transaction index denoting the number of transactions each index had with other indexes. Thus, according to poison heuristics; the related party is also considered as illegal. Therefore, the Degree denoted the connected illegal transactions.The size of the node denoted the strength of the node indicating how well the specific transaction has influenced by several numbers of transactions. The metrics Average Degree is the average number of related and adjacent indexes that are linked to a

specific index. The following *Figure 3.10* is the representation of Degree Centrality for the incident NiceHash. By using the zooming features the indexes to be analysed were obtained. The significant index was obtained 305655991, 305667670, 305701925, 305817614, 305714185, 305881864 which held higher degree. Thus, it was taken taken into consideration for further analysis.

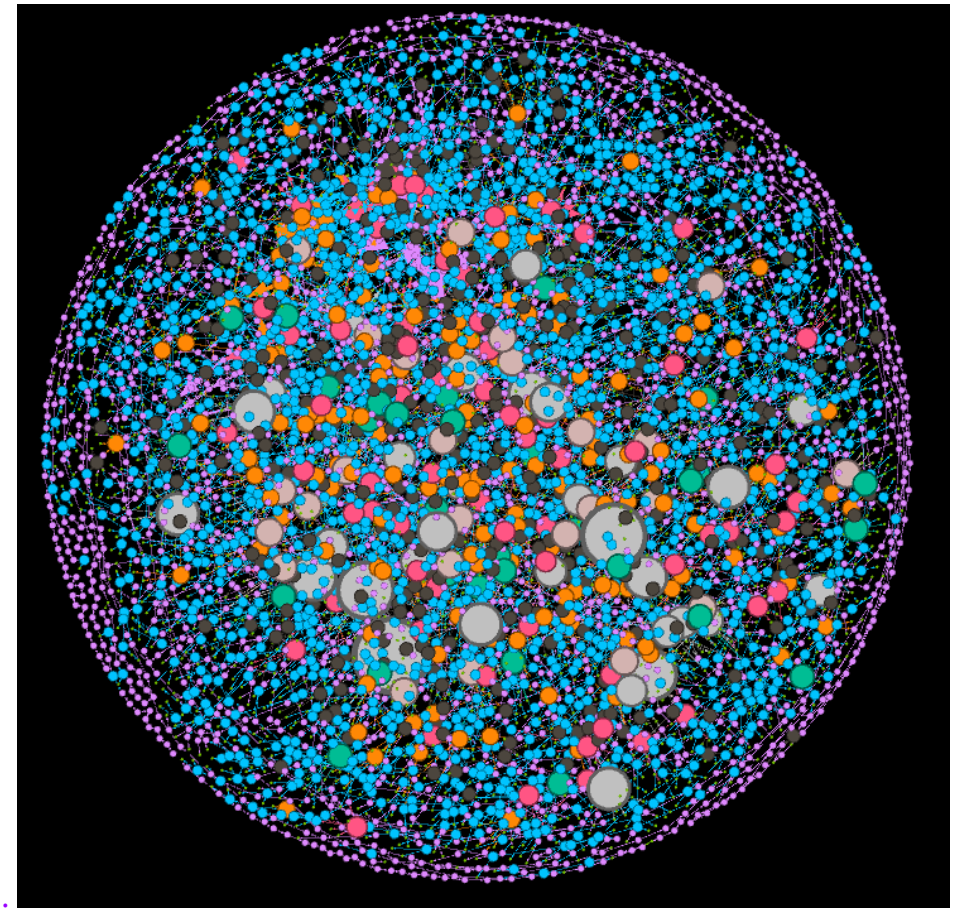


Figure 3.10: Degree Centrality for Incident NiceHash

Modularity in the study measured how well the network was disintegrated into modular communities of illegal bitcoin users. For the incident Nicehash contained 6 communities with resolution of 5.0 which were represented in different colors in *Figure 3.11*. This measure was just to get an overall idea how the indexes related to this incident were spread in the network.

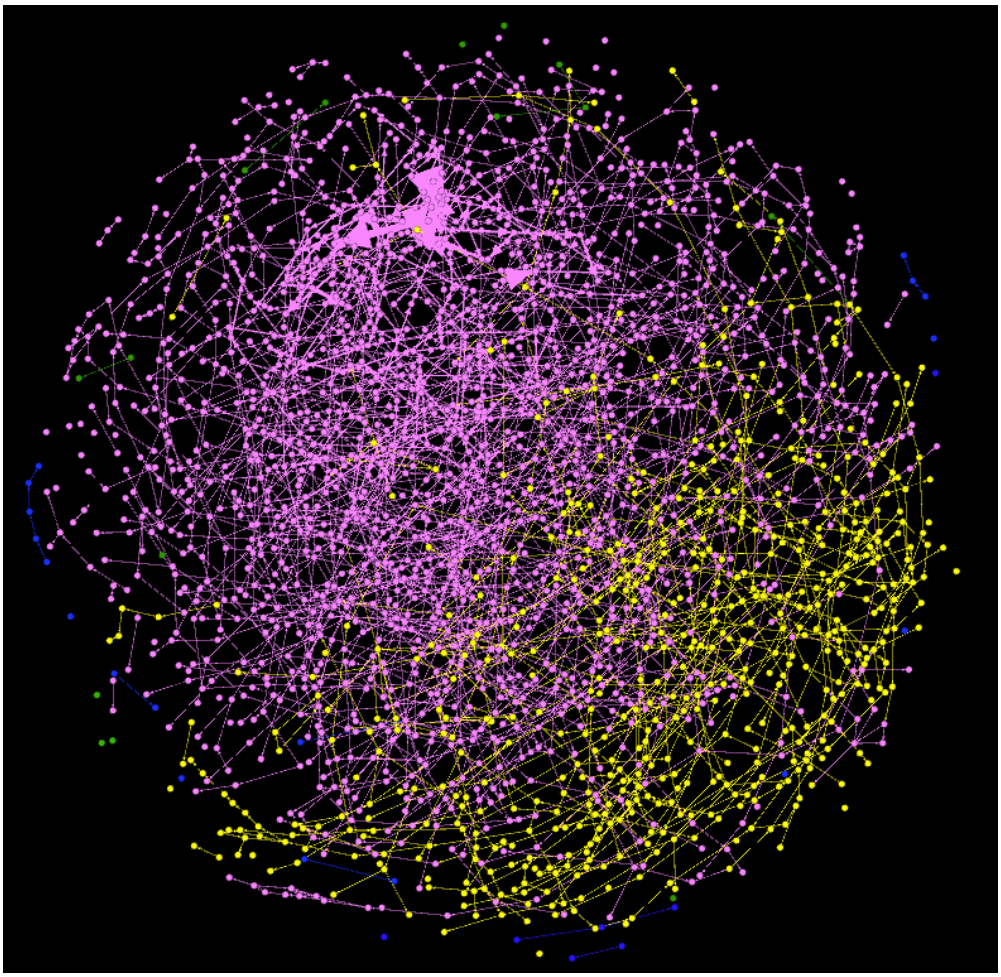


Figure 3.11: Modularity for Incident NiceHash

Eigenvector Centrality measured the importance of an index in a network based on connections of rest of other indexes. For instance, a high eigenvector score meant that an index was connected to many indexes who themselves had high connection. The following *Figure 3.12* is an example for the representation of Eigenvector Centrality for the incident NiceHash. The significant indexes were 305813297, 305813540, 305781355, and 305724948.

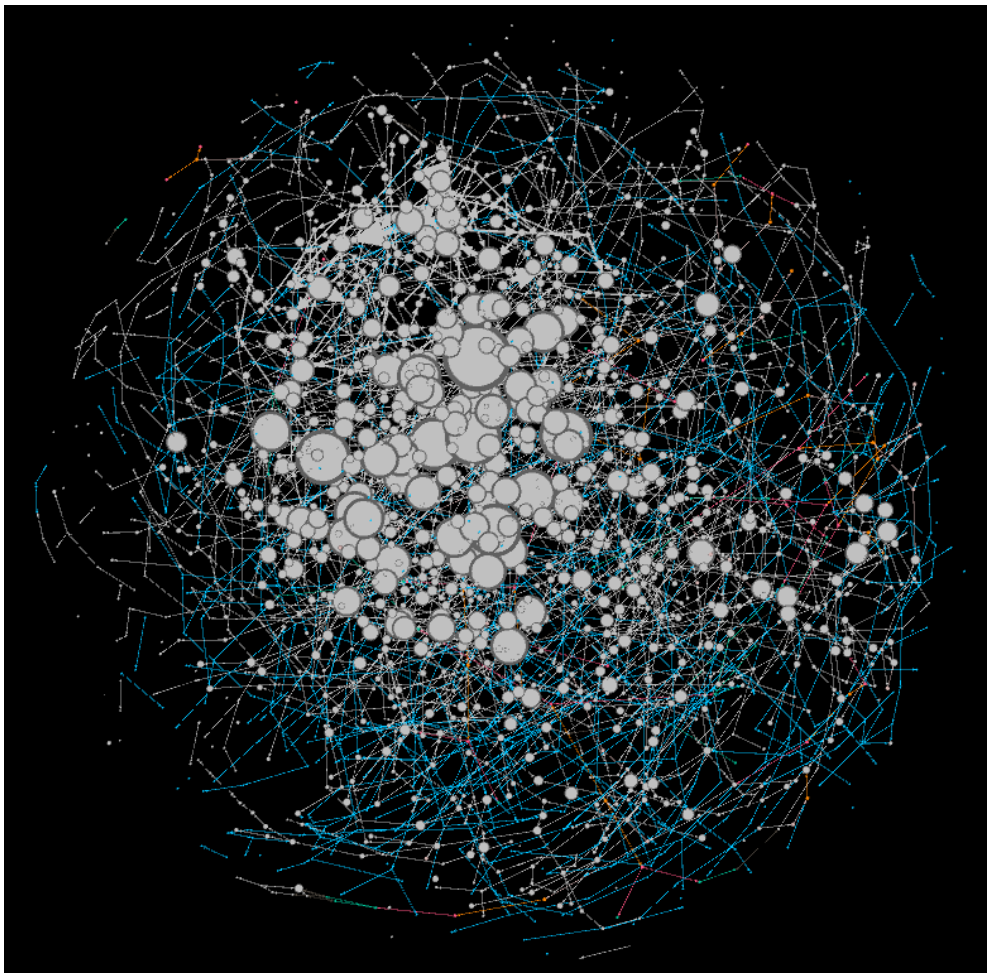


Figure 3.12: Eigenvector Centrality for Incident NiceHash

Betweenness Centrality measured to what extent an index was required to go to another index in this study. If an index is with a high betweenness, it often appears on shortest paths between indexes in the network. That is Betweenness centrality decides how the connectivity of the indexes in the graph is. The following *Figure 3.13* is an example for the representation of Betweenness Centrality for the incident NiceHash having 298174792, 305592308, 305603846, 305616674 were indexes of high betweenness.

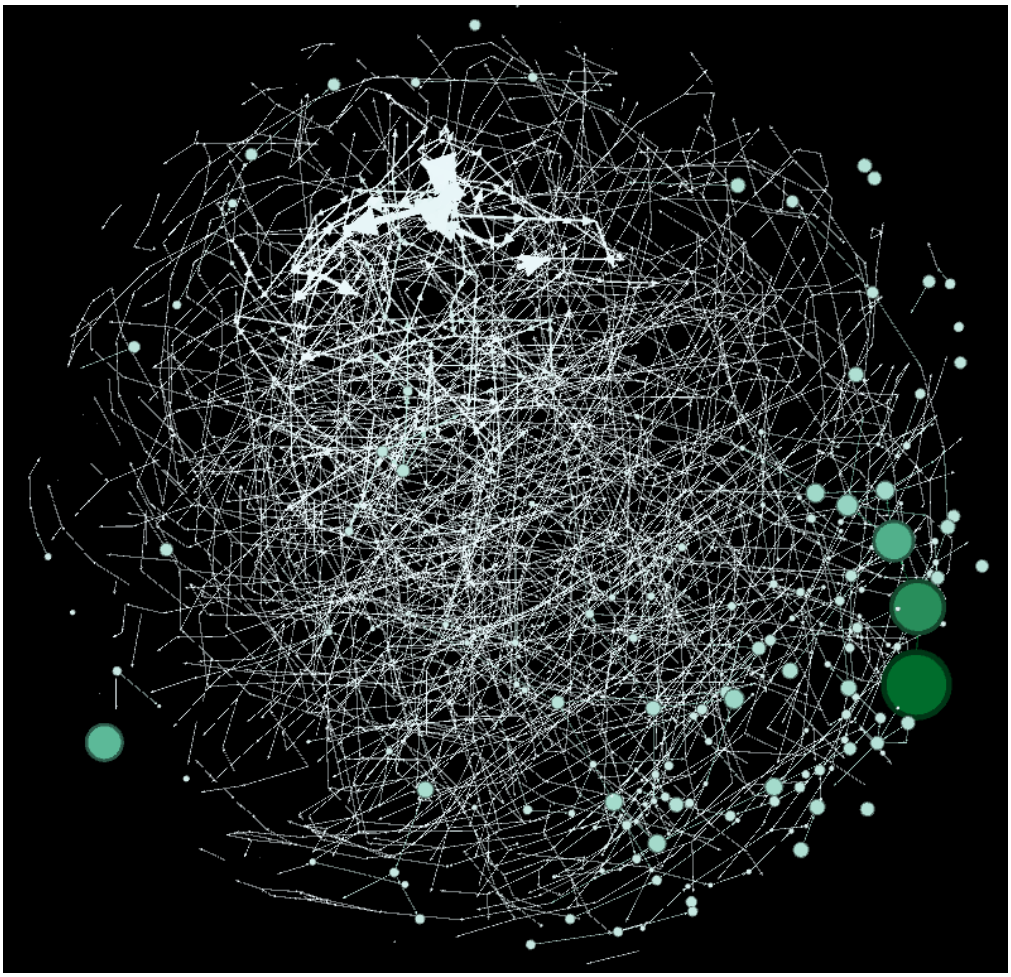


Figure 3.13: Betweenness Centrality for Incident NiceHash

3.9.3 Significant Index Summary

Metrics helped to identify the important index in the network. After computing the metrics such as Degree Centrality, Modularity, Eigenvector Centrality and Betweenness Centrality using Gephi. It detected the indexes that were of more importance according to the behavior and appearance of the network. For instance, in Incident NiceHash the following indexes were summarized as significant indexes in different aspects which were further considered in analysis process.

305655991, 305667670, 305701925, 305817614, 305714185, 305881864 indexes were with higher number of linked indexes. Whereas, 305813297, 305813540, 305781355, 305724948 indexes were connected to several indexes who themselves are connected to many other indexes. The

indexes such as 298174792, 305592308, 305603846, 305616674 305755805, 305657443, 305759195, 305798672, 305806775 maintained connectivity in the network by preventing separation within the graph.

Once the indexes were obtained, the bitcoin addresses related with significant indexes and its related indexes were obtained using the following *Script 3.5*. For instance, the high degree index was 305655991; then the script is used to obtain its addresses involved in that transaction.

```
tx = chain.tx_with_index(305655991)
for transaction in tx.outs:
    print(transaction.address)
```

Script 3.5: Related illegal bitcoin addresses for a given index

Thereafter the transaction id was used to analyse the spending pattern to identify significant facts or patterns by traversing through blockchain. On parallel to that the identified addresses were analysed with walletexplorer, blockchain tags, bitcointalk forum, reddit to identify the tagged probable owner of the address. The Results obtained as a result of this analysis in similar manner for other incidents are briefed in Results Chapter in details.

Chapter 04

Results & Analysis

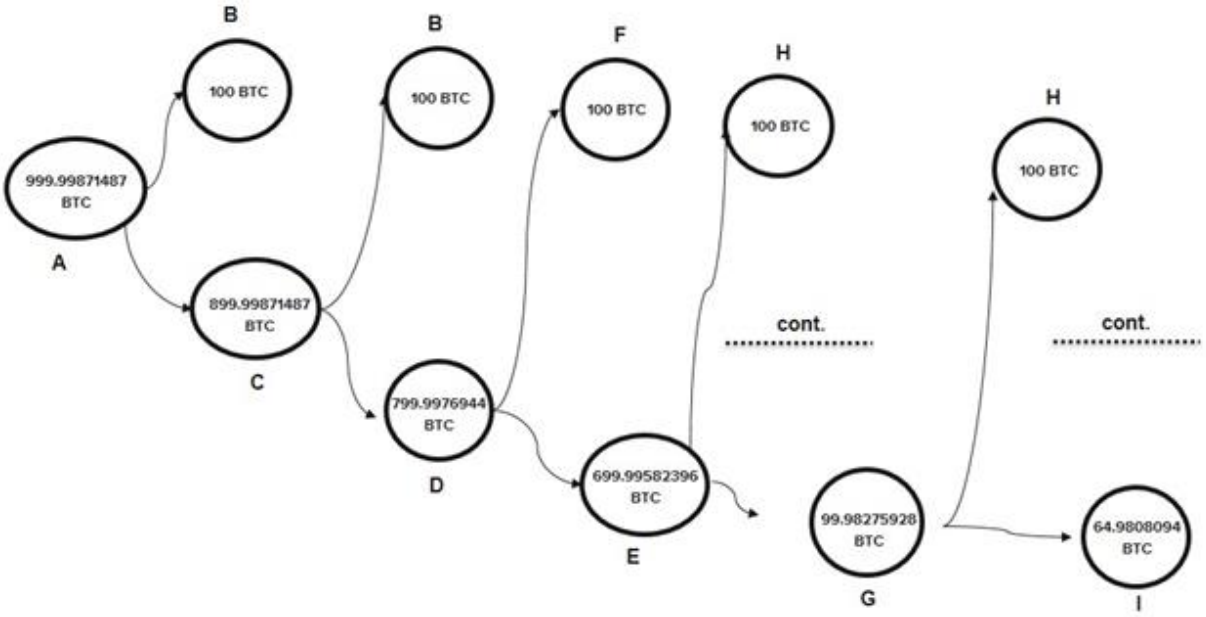
The following chapter summarizes the significant results for each illegal incident that was analyzed.

4.1 NiceHash

NiceHash, a cryptocurrency mining marketplace was formed in 2014 in Slovenia, a country in Central Europe. This served as a marketplace for miners to rent out their hash rate (i.e. a measure of miner's performance) to others. During the early December 2017, NiceHash has been hacked due to a security breach, causing a loss of 4,736.42 bitcoins [18] [58].

The analysis of results shows that the illegal party has been sending Bitcoins to several other different wallets and addresses in subsequent transactions before sending out them to an exchange or a service (*Figure 4.1*). The Bitcoins have been distributed subsequently in certain proportions. For instance, 100 of Bitcoins were sent constantly whereas the rest to another wallet which appeared to be freshly created wallets. This maybe simply for the purpose of transacting in small amounts. It can also be noticed that the fresh wallets are created and they have been mainly used with sole intention of 2 types of transactions which is to receive the amount and to send that amount. Thereafter, no transactions have been recorded in the fresh wallets.

[18] https://bitcointalk.org/index.php?topic=2535366.0;all



A - 1EnJHhq8Jq8vDuZA5ahVh6H4t6jh1mB4rq	[059337b6fb]	Wallet
B - 1F9E4dnwYxVidpNfBzjiJssE8ew5kVkb4K	[82f2ccfc22]	Wallet
C - 16EoBH6iTF2ygpXdRDpXkLgsCGDvG1pHq3	[e8f9216a56]	Wallet
D - 1CD6sbQgU9tS8ge1eertfxz4UD3RiUuZdr	[d06a4cf1bf]	Wallet
E - 1HnNdTHftJZLznYjgU9vLw7hTeVWE3151t	[059337b6fb]	Wallet
F - 19arQXiR63XQdbvzL3NfzXGwPXjG6jEb7h	[cde314a885]	Wallet
G - 197F9RdmF3WcBETjKo2XKj8jzDLESniSP3	[461cb29c5f]	Wallet
H - 1GqmHHy2uF3cZ6gHshuFSFZJeo9binA1AR	[72c787e808]	Wallet
I - 1GMH31jsHyhwUiCAZbLEY84ZKEYykvcSFC	[77baa2bbe6]	Wallet

Figure 4.1: Transactions to freshly created wallets

In addition, similar patterns can be observed in Tx id 26f…..0d1 [19] where proportion of bitcoins divided are decreased gradually by sending a constant of 100 BTCs to addresses 1BN…..HGp [20], 1Gq…..1AR [21] and the rest to fresh wallets in subsequent transactions.

Another similar type of pattern that can be observed is that, bitcoins are distributed among two addresses or to more addresses in a certain proportions such as starting from ratio in percentage of 89 % : 11% approximately. This ratio gradually decreases in proportion of 85%: 15%, 75%: 25%, 67%: 33% and so on. This continues till the wallet reaches to approximately 4.94 Bitcoins. And

[19] 26fe37f72fd8017044719fe2bd8fdd6f85f9eb8e05fa6e1ee873672999d1a0d1
[20] 1BNkb37S4F78SuH79HTnewEJhsyXLevHGp
[21] 1GqmHHy2uF3cZ6gHshuFSFZJeo9binA1AR

thereafter it gets transferred to an address 1Kg…...bfr [22] in a fresh wallet in every different hacked transaction. As for example Tx id : 4b417…….3dac4 [23].

4.2 VenusLocker

VenusLocker is a ransomware type virus which was spreaded via an infectious email letters [59]. The results of this in transaction traversal shows that there are few popular exchanges that is closely related with transactions. They are Poloniex.com, Xapo.com, Luno.com and untagged wallets such as [000245c2b8], [000194c053]. In addition, another notable result would be that hacker had immediately made some transactions with Helix Mixer as in *Figure 4.2*.

Tx id - d38c3fc4bebd70bd8e511487a1aad60e5324445974bb7e3ea723a6ae90484dbb

Figure 4.2: Hacker's immediate transaction with Helix Mixer

Thereon, it is also revealed that illegal party has been sending bitcoins to several other different wallets and addresses in subsequent transactions before sending out them to an exchange or a service. The Bitcoins have been distributed subsequently in certain proportions as shown below in *Figure 4.3*.

[22] 1KgLhtSQG9QtUgHiVmawVqDTX6JyZCCbfr
[23] 4b417dcfc06b7ccdb61006d3e7025baa192b53d36cf8e9476edede4bc7f3dac4

Tx id : b0506a47ac75e9e34e0ca1a5a77de1483a8eb094b6e55fc9381dafe86ca828d9

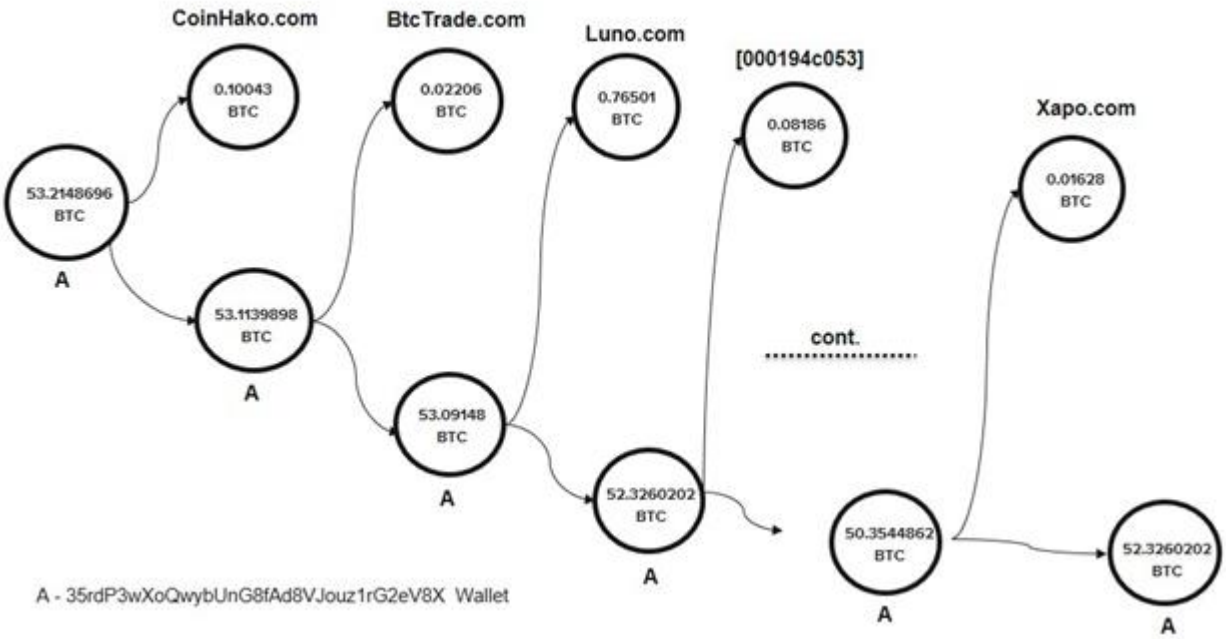
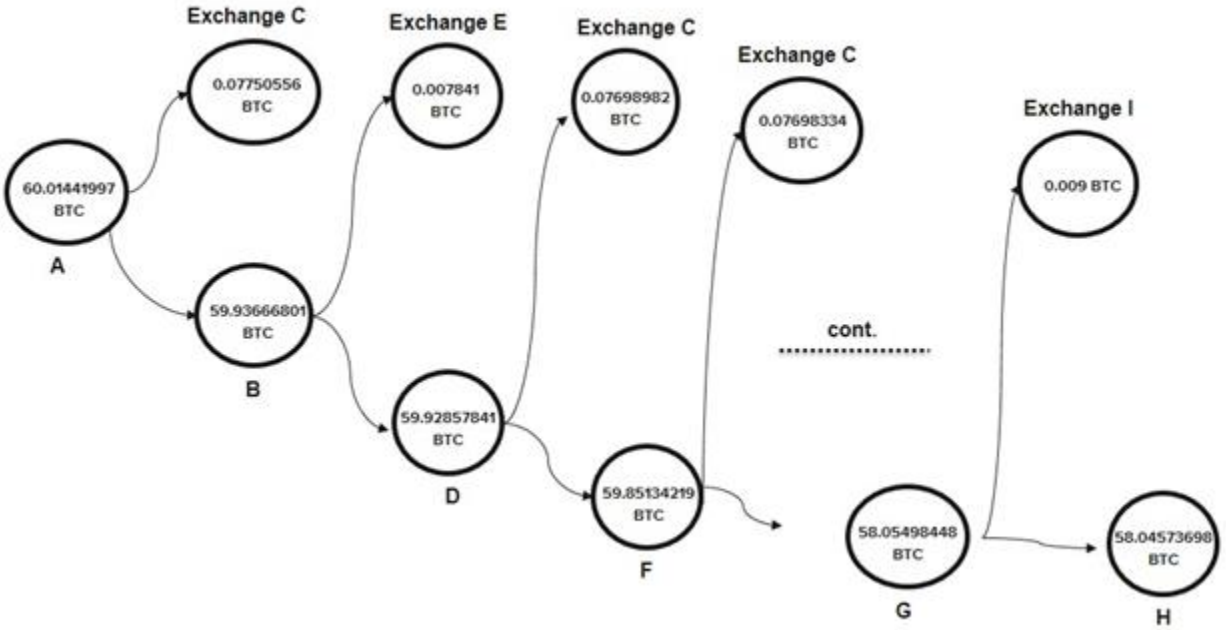


Figure 4.3: Bitcoin amount transacted to exchanges from same wallet

In addition, the following *Figure 4.4* shows how bitcoins were being sent to freshly create new wallets.

Transaction Id: 318ac7f93bc86c1ecae1fbacfd1942d293a539c625f35c6adeeab61d8675eafa



A -	1PR9kXXCzoMNLXqVJHZBMZcfYF4B7qjNG6	[dc07337c06]	Wallet
B -	19aex5hLQBmCcXmTcypSxkAXKqDawAF9n	[059c52ddf5]	Wallet
C -	323AvZBQDX2CzfamjAy5W8VBW8CkNK5Wqg	[0000e74af0]	Exchange
D -	1LTFMxcz79VsnLnN6Ti9UGkwK5XcC24nC6	[6637a8bdd1]	Wallet
E -	1LtmG3BLWeTm9V8Hm9tXVvaZU1oyoBCPYs	[006572b4e7]	Exchange
F -	1L1o5smk9zdkVkwQanXzbfGVaoune9XuKQ	[a2877fec1e]	Wallet
G -	18dQJ8XgZpNLetrjWiZbkpMiHMp5QADNAV	[d145697b3d]	Wallet
H -	1KYkXekBju6vUMFY3fM1URhUyFTSkZfVFa	[b93a134937]	Wallet
I -	1Ndbdq5uTeiT8Mn42W3yPZw6n6QPmgFqJo	[04af64e01c]	Exchange

Figure 4.4: Bitcoin amount transacted to exchanges from several wallets

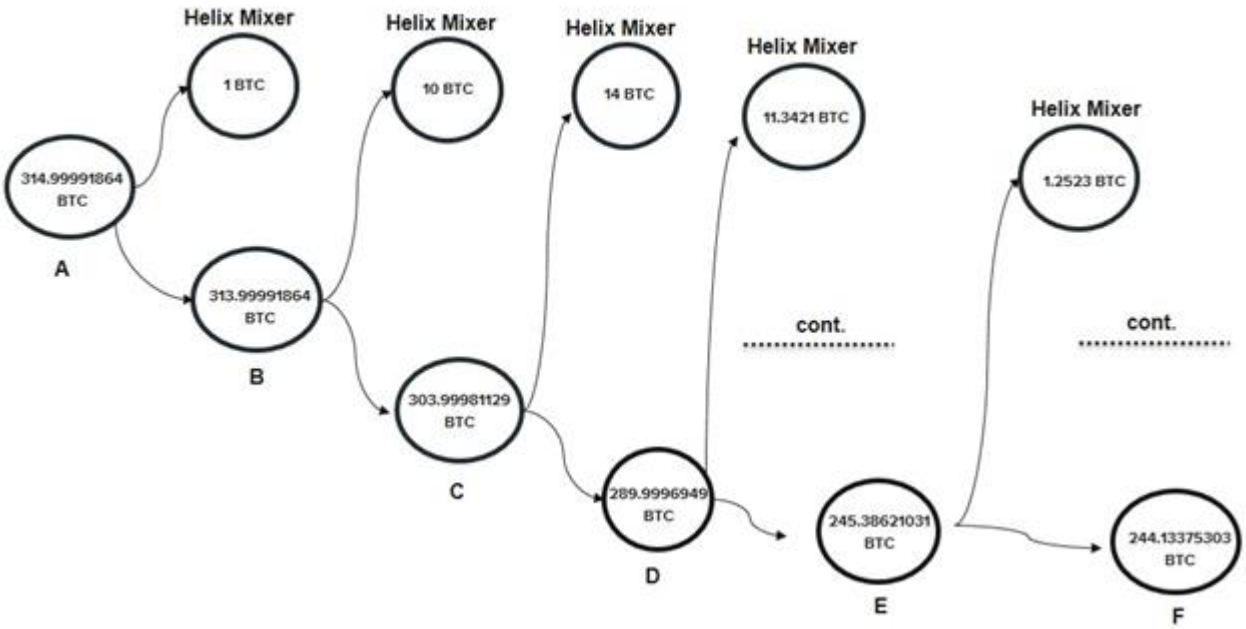
4.3 Shapeshift.io

Shapeshift.io is a Switzerland based cryptocurrency exchange service that offers trading cryptocurrencies through its website and its API globally. On 7th April, 2016, it faced a security breach which compromised on the server infrastructure of platform. On these continuous incidents, ShapeShift lost $230,000. The hack was proved to be an inside job by an employee who sold key security data to an outsider [24] [60].

The analysis reveals a significant fact that there are transactions tagged to 'Helix Mixer'. Helix Mixer is a bitcoin mixer accessed via darknet administered by Grams. A pattern can be observed in the transaction chain as shown in the *Figure 4.5* which was drawn considering the starting transaction id 1855…..2cff [25]. A significant amount is transacted to freshly created wallets whereas the rest is sent to Helix Mixer as depicted in *Figure 4.5*. Thereafter, the fresh wallet does not indicate any record of activities.

[24] https://news.bitcoin.com/looting-fox-sabotage-shapeshift/, https://bitcointalk.org/index.php?topic=1449715.0

[25] 1855667af03f4cda150f2cbf5d12948c6f1883256a0570f36e66223a0ebc2cff



A -	1LchKFYxkugq3EPMoJJp5cvUyTyPMu1qBR	[129a49215e]	Wallet
B -	1PZzrgrxGYWaLhwpoMu4E7T2A1wdd7jczB	[7a79f023ce]	Wallet
C -	1DcuXaMdnF2r8S2JdfYzTXTSgcid7ZxJTW	[3c5d1e7703]	Wallet
D -	12zSugV9sqTFp5arHEuooZCXB8B9PUQWpd	[7e62cfa5f4]	Wallet
E -	16iCMmpvFbhEZFeo3aTVAivJUAiU33ormw	[714cbed68c]	Wallet
F -	1D8d2LYFAMEjEsMQ3kmzUAUYN1kS57idc2	[7d9e8cc838]	Wallet

Figure 4.5: Bitcoin amount transacted to Helix Mixer and several fresh wallets

Another notable results in the transactions traversed indicate that mostly Polenix.com is used as the popular exchange. One of the significant results is that next transaction that comes out from exchange has been made with the address 344….1Md [26] that has been tagged as 'Richest Bitcoin Address'.

In further analysis, it can be observed that an address 1DU….Uru [27] which belongs to Bittrex exchange has higher degree of sending bitcoins continuously to its change addresses before transferring to a new wallet. In addition, this address has been listed as the sixth richest on the Tether crypto list[28].

[26] 344pUP56enuGjbPdyubYEqoxB6VaFmD1Md
[27] 1DUb2YYbQA1jjaNYzVXLZ7ZioEhLXtbUru
[28] https://www.reddit.com/r/btc/comments/7ih0hd/guess_who_controls_over_half_a_billion_tethers/

4.4 WannaCry

WannaCry Ransomware is a type of malicious software that blocks user access to files or systems, holding files or entire devices under their control using encryption techniques, till the victim pays the demanded ransom in return for the decryption key. WannaCry Ransomware Attack happened in May 2017. According to statistics of this attack 300,000 computers including entities such as hospitals, companies, universities and government organization across 150 countries had a loss of hundreds of millions to billions of dollars. It has been proved the initiative of this ransomware was from North Korea or related agencies [29] [59].

The result of this transaction shows that there are few popular exchanges that is related in transactions traversal. They are Poloniex.com, Bittrex.com, HaoBTC.com, BTC-e.com, Xapo.com, CoinGaming.io and untagged wallet 00000ff1c2. The same wallet 00000ff1c2 has been highly used repeatedly to cash out in Alphabay incident too which can be carefully interpreted to be a bitcoin service.

Another notable result would be that the following addresses in their respective wallets have involved in conjoin transactions to mix their coins. This was carried out to make it harder for outside parties to determine which party or parties was making a particular transaction.

1MNm91r5tCCYrrVDQZZXFjX6iLmHodfGQJ - [798f69dcbc]
1HMP1MdAk6Vg4PLVJsjQKukLRRwpvgkzYX - [0d1adfe792]
135EKnnTBBjC9uUcTEAkdZKvstVfA4mf2c - [1272d9cce7]
1CFqVxksgY8p98SpSHpj1o9ZC8xq5THtdq - [04e232922c]
3QkdFTCphYvS4EhBmgaSN996QWiCZTEV5n - [ce1abf0e1f]

4.5 Alphabay

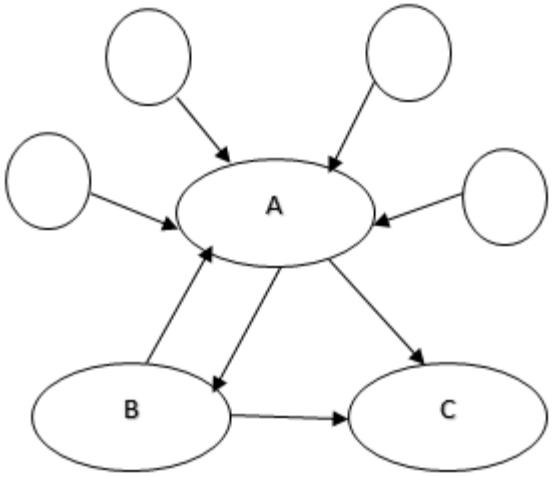
AlphaBay Market, operated in Thailand was an online darknet market which was launched in December 2014 and operated via Tor network. It operated under an escrow system which paved the way for the scam. AlphaBay went offline due to a scam with 1,479 bitcoins transferred from a

[29] https://www.theguardian.com/technology/2017/may/12/nhs-ransomware-cyber-attack-what-is-wanacrypt0r-20

bitcoin wallet which were identified to be used by those behind the darknet site to other bitcoin wallets. During that period, there are numerous orders pending in its escrow system. It was shut down by 13th July 2017[30].

The results show that there are more number of ScriptHashAddresses which are wrapped with MultisigAddresses in transaction traversal. As such in the very first transaction, it has send to bitstamp.net, xapo.com and to other untagged addresses which are possibly bitcoin services based on activities of its wallet. The significant fact is that ScriptHashAddresses also have been used to transact in the chain providing evidence that they also have some sort of relationship other than with main recipients. Thus, it is also important to notice AlphaBay Market allowed using Multisig for payments during its active time.

Moreover, three connected wallets repeatedly transacted with each other as shown in below *Figure 4.6*. Based on transaction activities of the wallets, all three appears to be wallets of a bitcoin service. Finally all bitcoins have been transferred to B leaving 0 balances at the other wallets.



A - 13KMJQ9uqAkXok7CM22tN7jFjvd7dzkRPW, 1Kr6QSydW9bFQG1mXiPNNu6WpJGmUa9i1g [00000ff1c2] Wallet
B - 3D2oetdNuZUqQHPJmcMDDHYoqkyNVsFk9r [4d0aa5dbef] Wallet
C - 1FTgECWAnepRwHcouefWpANHgywQ35DNcp [04bb2fc371] Wallet

Figure 4.6: Wallets connected to each other

30 https://www.reddit.com/r/AlphaBay/comments/6lbu32/alphabay_down_shit_vendor_review_as_well_buyer/, https://news.bitcoin.com/major-darknet-marketplace-alphabay-goes-down-exit-scam-speculations-arise/

4.6 Cryptorlocker

CryptoLocker Ransomware started spreading since September 2013 that encrypted files and demanded ransom. This was sent via a spammed message that contained a malicious attachment. This created almost USD 519,991 direct financial impact. CryptoLocker opened the gates to many other ransomware variants. [18] Later on, the botnet of CryptoLocker was brought down by Operation Tovar, a white-hat campaign[31].

In the study of the incident CryptorLocker, it could be discovered that an address 121….PM4[32] belonging to wallet [2fcbc8653a] had obtained Bitcoins from Agora and Evolution darknet market and paid in bitcoins to Agora market and Black bank market. Thereby, [2fcbc8653a] is possibly a wallet belonging to a darknet market supplier.

4.7 MyBTGWallet

Bitcoin Gold (BTG) is one of the forks of bitcoin which was released on 24th October 2017 [33]. MyBTGwallet.com is an online wallet creator that only stores data on the browser. This website cheated investors out of $3.3 million in November 2017 by promising to allow them to claim their Bitcoin Gold. This scam was done by the owners of MyBTGWallet who was a member of Bitcoin Gold community and also promoted the project on Bitcoin Gold. The owners falsely lead their users into a trap by telling that the Bitcoin Gold could be claimed in an easy and cheap manner, which allowed them to be trapped by a malicious scheme. The users had been required to submit their private keys in order to create their Bitcoin Gold wallets. Analysis of the website code showed the website stored these private keys when they were submitted. Then the website owner had access to all of the funds that were in wallets and the operators of the website diverted money into a different bitcoin addresses. The owners of this mybtgwallet project made changes in the GitHub code as well [34].

[31] https://www.bleepingcomputer.com/virus-removal/torrentlocker-crypt0l0cker-ransomware-information
[32] 121dBo5epQEDJZVpZDuBYBwV5Y2xeXTPM4
[33] https://99bitcoins.com/the-bitcoin-gold-hard-fork-explained-coming-october-25th/
[34] https://news.bitcoin.com/bitcoin-gold-wallet-stole-private-keys-scooped-3-3-million/, https://bitcointalk.org/index.php?topic=2412182.0, https://www.reddit.com/r/CryptoCurrencies/comments/7db42c/httpsmybtgwalletcom_seems_to_be_scam/

The results of this transaction traversal showed that there are few popular exchanges that are closely related with the transactions. They are Bittrex and Bitflyer.jp.

In addition, the results highlight that there has been continuous transactions from the exchange 'hitbtc.com' to another address 3Jj….4FC [35] belonging to wallet [457b8ced80] as shown in *Figure 4.7*.

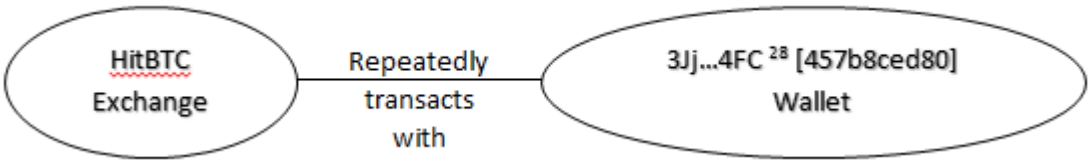


Figure 4.7: Repetitive transactions from HiBTC exchange to another bitcoin address

With further analysis, it can be observed that this wallet is active to date, but the specific address has been tagged as involved in many illegal incidents by bitcointalk forum users. The common incident on the tagged address is that when some users copied their own bitcoin addresses from hitbtc.com, the address is getting changed or replaced to this above illegal address 3Jj….4FC [28] automatically which is a malware.

4.8 Blackmail

Several people received different versions of emails claiming that the recipient's computer has been used to create a video of adult websites that the recipient visiting and threaten that it will be sent to recipients' contacts if they do not paid $200-$400 in BTC within 20-24 hours. It appeared in many languages and has been consistent format[36].

As per analysis, it can be noted that majority of the transactions have been performed significantly through exchanges such as Poloniex.com, Matbea.com, Cubits.com. Some individuals identified as involved especially in blackmail incidents are John Doe of bitcoin address 16F….BhY [37] and Natalie Wallis of bitcoin address 1Fo….MBG [38].

[35] 3JjPf13Rd8g6WAyvg8yiPnrsdjJt1NP4FC
[36] https://bitcoinwhoswho.com/blog/2017/10/09/blackmail-scam-run-on-russian-wallet-matbea/#more-540
[37] 16FqwvZdty3H3B5hjHK268YvDoDFaUxBhY
[38] 1FoR6263co7SA9BAA3qrrH4sT7qSFnMMBG

Among the blackmail incidents, it is significant that the Bitcoins are immediately cashed out via exchanges since the money received on blackmailing is not relatively a notable large bitcoin value.

4.9 Fake Agency Support

This includes a Coinbase support phone scam where a phone number '1-888-455-1155' which is not a real Coinbase support number were shown up in a lot of web search results. When users search in Google typing "coinbase phone support" they obtained a phone number from Google search results that leads them to this scam in which an operator tells them to send money in bitcoin[39].

As per analysis, it can be highlighted that the transactions have been performed significantly through exchanges such as Cex.io, Luno.com, Bittrex and has cashed out from exchanges immediately.

4.10 Gatecoin

Gatecoin is an exchange established in Hong Kong, mainly facilitating services for bitcoin and Ethereum tokens. It was founded by was founded in July, 2013 by an investment banker, Aurelien Menant. Gatecoin was hacked in between 9th and 12th of May 2016. The hackers accessed the hot wallets of both bitcoins and ethereum stealing 250 bitcoins and 185,000 ethers[40]. This is known to be representing 15% of Gatecoin crypto asset deposits [61].

A major portion of the immediately sent inputs are still unspent. Due to that reason, a clear insight on the tainted bitcoin circulation cannot be obtained. However, the minor amount of Bitcoins that were spent indicates that mostly Poloniex.com, OKCoin.com, Bter.com, Xapo.com are very commonly used exchanges and those Bitcoins have been cashed out immediately.

[39] https://bitcoinwhoswho.com/blog/2017/12/17/fake-coinbase-support-phone-number-1-888-455-1155/, https://www.reddit.com/r/Bitcoin/comments/77hxl0/my_bitcoin_at_coinbase_got_hacked/
[40] https://news.bitcoin.com/gatecoin-official-statement-hot-wallet-breach-losses-estimated-2m-usd/

The following *Table 4.1* has been summarized to show the incidents and their respective categories that has been used to obtain the findings.

Hack	***Personal Losses***	***Scam***
NiceHash	WannaCry	BTGwallet
Gatecoin	CryptoLocker	AlphaBay
	Venuslocker	
ShapeShift	Blackmailing	Fake Agency Support

Table 4.1. Summary of analyzed incidents under their respective categories

Chapter 05

Discussion and Evaluation

5.1 Discussion

The findings from this study suggest that there are patterns and significant facts among illegal incidents of different nature.

A significant pattern that can be seen in several incidents is that illegally transacted Bitcoins are being sent to different wallets which are newly created wallets. Generally, the total bitcoin value is sent to two new addresses as shown below *Figure 5.1* in a variety of methods. As per recent study of [8], illegal users tend to transact more in smaller amounts repeatedly with a certain party to avoid getting noticed. In addition, it is noted that illegal users are holding less bitcoin due to bitcoin seizure incidents by FBI [8].

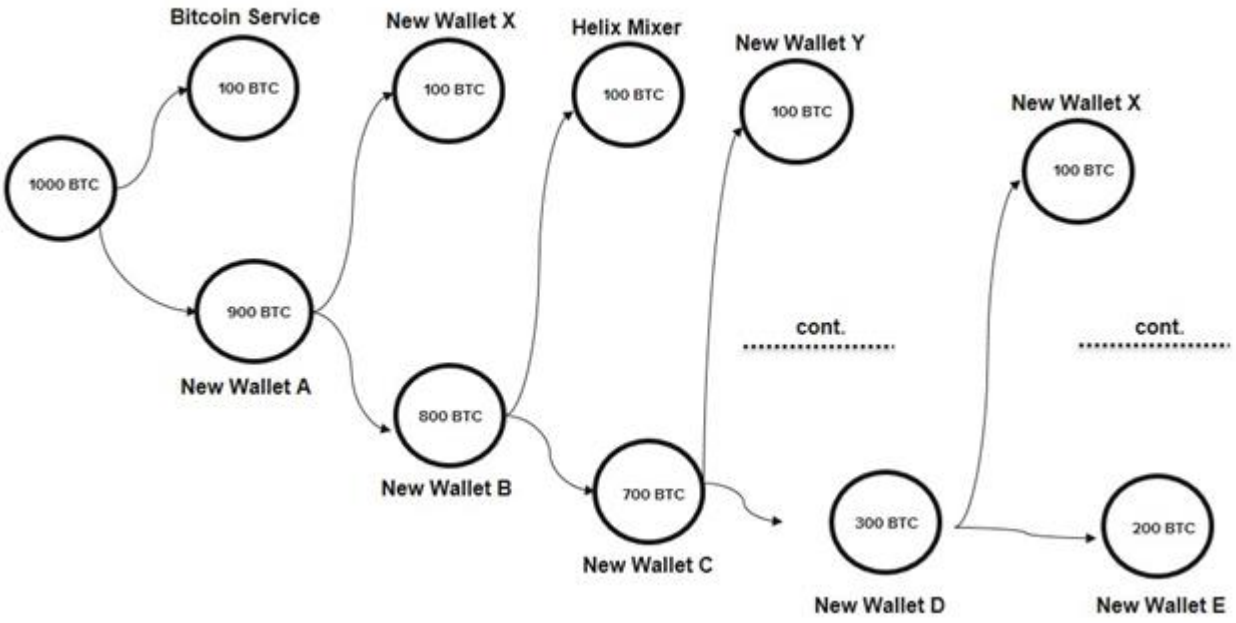


Figure 5.1: Common pattern that prevails among illegal users in their transactions subsequent to the incident

1) The small value is sent to either

a) A bitcoin service which includes an exchange, a bitcoin washer, or a mixing service provider or a retailer. Eg : To Bitcoin Service or Helix Mixer as shown in *Figure 5.1*
b) A newly created wallet. Eg : To New Wallet X or Y as shown in *Figure 5.1*
c) A newly created wallet which has been used in a previous very recent transaction. Eg : To New Wallet X which is further away from initial transactions and uses New Wallet X after several transactions as shown in *Figure 5.1*

2) Bitcoin value is generally transacted to two addresses on most cases in their transaction chain. The total Bitcoin value is divided into two portions. One contains a relatively larger portion and another less amount. Two main technique of dividing the portions are notably highlighted as,

a. The smaller portion having constant bitcoins and in subsequent transactions, the smaller portion maintains the constant amount as depicted in *Figure 5.2*

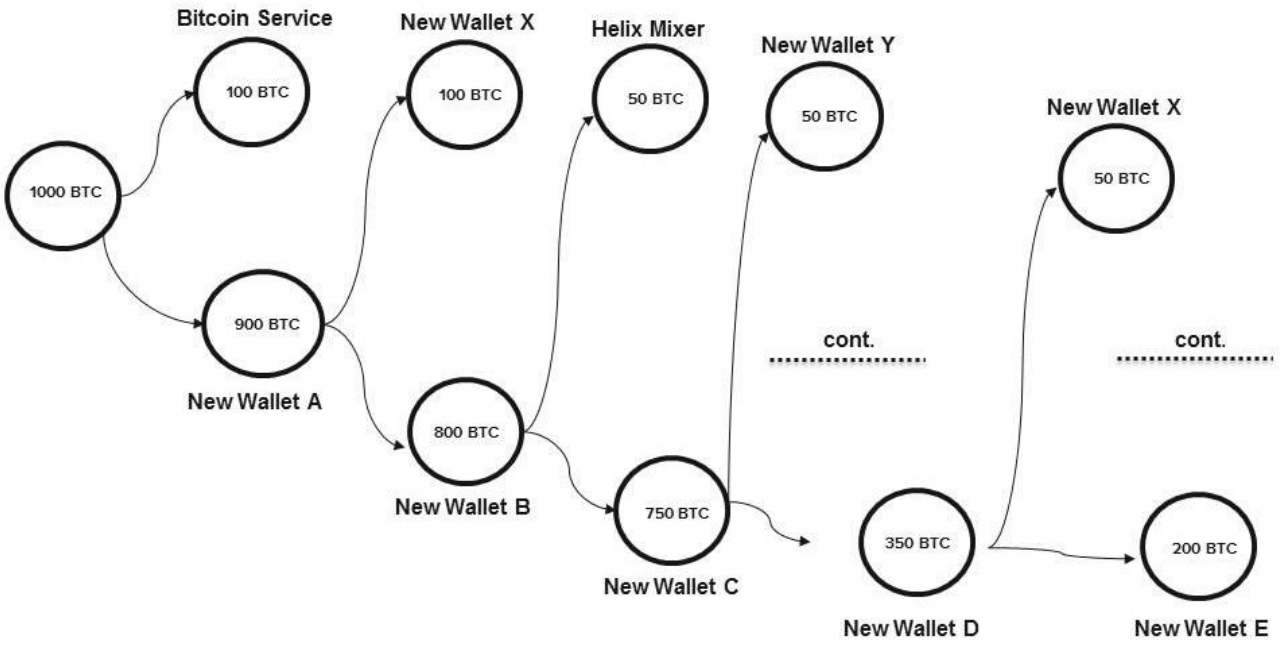


Figure 5.2: Transacting a constant amount of bitcoins in subsequent transactions

a) The two portions are divided into a certain percentage and it continues in subsequent transactions.

As shown in *Figure 5.3*, the bitcoins are spread or distributed in a certain proportion which reduces in subsequent transactions.

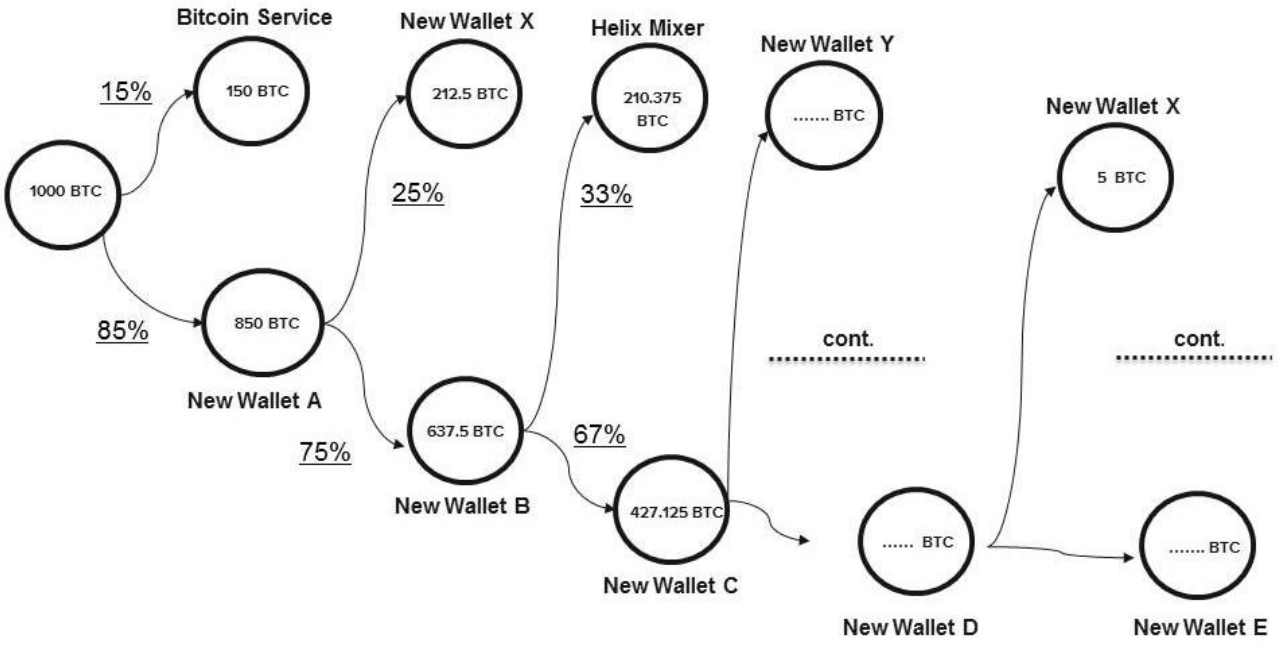


Figure 5.3: Transacting a bitcoins in a certain proportions

This way of transacting continuously in constant or a certain proportion can also be interpreted as that the illegal users have automated the process of transacting in this manner.

The pattern of continuously transacting to two addresses by creating new wallets and minimizing the entire value is to reduce being noticed by transacting in large amounts. Even though there are transactions which involves spending to more than 2 outputs, this specific pattern is very commonly observed.

In addition, much of wallets were newly created to transact. It is observable two types of transactions in new wallets. A transaction to receive the amount and to send it back to other wallets. The balance remains zero (0) and no activities can be seen thereafter.

Generally users have hundreds of different bitcoin addresses which are stored in their digital bitcoin wallet [6], [25]. But, reusing bitcoin address is traceable because the flow of bitcoin can be traced from one known or unknown address to another [18] leading to privacy leaks. As per this study finding, some criminals transact to new bitcoin addresses in to their new wallet.

When criminal are cashing out via a bitcoin service, there are few patterns identified as

1) Directly send the illegally obtained Bitcoins to exchanges or mixers without using any intermediate addresses. Illegal incidents such as Blackmailing, Fake coinbase, Alphabay scam has used directly an exchange to cash out. In addition, As such the illegal incidents of Shapeshift and VenusLocker have used their tainted Bitcoins to purify with the Helix Mixer.

2) Transfer the Bitcoins to fresh wallets created newly or to new addresses. The transactions are done subsequently by continuously reducing the total bitcoin value in single wallet or address. Thereafter, transfer to exchanges or a bitcoin washer or a service in small amounts. It can be interpreted that creating several new addresses/wallets to transfer is an attempt to protect them from being noticed of transacting in larger values. This would, in turn, make the follow-ups difficult.

3) Transfer the Bitcoins to fresh wallets created newly or to new addresses. The transactions are done subsequently by continuously reducing the total bitcoin value in single wallet/address. Thereafter, transfer to untagged addresses which can possibly be an exchange or a bitcoin washer or a service. It could be an exchange or a bitcoin washer or a service due to the wallet activity. In addition, using untagged addresses which are publicly less known exchange services or may be an attempt to manipulate.

In addition, All of the incidents, at one point transfer to a bitcoin service even after transferring to several new wallets and addresses. There are some popular services such as Poloniex.com, Bittrex.com, Xapo.com, and an untagged address 00000ff1c2. In a previous research [18], it is highlighted that certain exchanges such as Xapo.com,BTC-e.com, LocalBitcoin.com, Kraken.com.

In addition, it is observed criminals are using an untagged or less known bitcoin service. This can also be interpreted as an attempt to increase complexity in investigations.

When a criminal transacts through an exchange or any bitcoin service. The tainted Bitcoins circulate within the exchange and when it comes out, it exits from a different user, so it’s hardly

distinguishable to trace the related party as there won't be records in Blockchain. The records will be in the bitcoin service's database only. This can be interpreted as criminals transact through multiple exchanges which are registered under different countries and they comply with different law enforcement. Then, different rules and regulations on exchanges may or may not pave the path to reveal the identity of the criminal. So the results highlight, this common pattern followed so far.

However, it's not completely untraceable, since the operator of the bitcoin exchange/ service will have sufficient proof required to reveal the outbound transaction back to the origin of input tainted coins, as the circulation within the service gets recorded in their database. So, still, the criminals are under the vulnerability of revealing the truth of the crime. In addition, as per blockchain explorer[41], it is clear that large services like exchanges, wallet services and mixing services have the capability to track and identify a large subset of user activity. It is evident according to the findings of [51] study discovered that all the isolated large transactions in the system had a close relationship to a single large transaction belonging to November 2010, even though the users tried to hide this reality with several strategies such as long chains and fork-merge structures in the transaction graph.

5.1.2 Discussion Summary

Validated Findings	**New Findings**
Ransomware criminals cash out via a bitcoin services, gambling and mixing services to avoid being tracked by investigators [18]. In addition, some ransomware attackers particularly transacted several times with the same party [18].	1) Directly send the illegally obtained Bitcoins to exchanges or mixers without using any intermediate address. Illegal incidents such as Blackmailing (personal loss), Fake coinbase (scam), Alphabay (scam) has used directly an exchange to cash out.
As per recent study [8],illegal users tend to transact more in smaller amounts repeatedly	2) Transfer the Bitcoins to new wallets or to new addresses. The transactions are done subsequently by

[41] https://www.blockchain.com/explorer

with a certain party to avoid getting noticed. In addition, it is noted that illegal users are holding less bitcoin due to bitcoin seizure incidents by FBI [8].	continuously reducing the total bitcoin value. Thereafter, transfer to exchanges or to a bitcoin washer or to a service in small amounts. It can be interpreted that creating several new wallets is an attempt to protect themselves from being noticed of transacting in larger values. This would, in turn make the following ups difficult. 3) Transfer the Bitcoins to fresh wallets or to new addresses. The transactions are done subsequently by continuously reducing the total bitcoin value in single wallet or address. Thereafter, transfer to untagged addresses which can possibly be an exchange or a bitcoin washer or a service. It could be an exchange or a bitcoin washer or a service due to the wallet activity. In addition, using untagged addresses which are publicly less known exchange services or may be an attempt to manipulate.
	Bitcoin value is generally transacted to two types of addresses. They are to different exchanges or to new wallets multiple times. The total Bitcoin value is divided into two portions. One contains a relatively larger portion and other with less amount. Two main technique of dividing the portions are notably highlighted as, 1. The smaller portion having constant bitcoins and in subsequent transactions, the smaller portion maintains the constant amount.

	OR 2. The two portions are divided into a certain percentage and it continues in subsequent transactions.
Popular exchanges or services used to cash out are Xapo.com, Helix Mixer [18].	Popular exchanges such as Poloniex.com, Bittrex, CEX.IO, and untagged wallet address 00000ff1c2. In addition, using an untagged or less known bitcoin service can also be interpreted as an attempt to reduce further investigations.

Table 5.1: Interpretation Summary

5.2 Evaluation

As in [62], it has been highlighted that many of the proposed solutions in blockchain related researches are lack of solid evaluation on their effectiveness. However, recent researches in [23], [63], [64] has used deep neural network and unsupervised feature learning approach to evaluate the results obtained.

As in recent research study [23], on bitcoin address linking; authors used deep neural network for the purpose of testing the efficiency of the method used. In addition, as in [63] and [64], for the purpose of frauds detection in bitcoin network, an feature learning approach of K-means has been used. Whereas in [65], a feature learning approach together with performance measures and validation techniques used for evaluation.

But, accordingly the approach of machine learning cannot be implemented as only few results can be tested using some features. Therefore, in this study the evaluation technique of obtaining feedback from real bitcoin users was used.

User feedback

This approach is to obtain feedback from real bitcoin users about their spending patterns of Bitcoins by issuing an online survey. Due to the lack of reliability and response from illegal users with regards to their bitcoin usage practices, the study focused to obtain feedback about legal users' behavioral patterns. The survey is available via this link[42] and attached as appendix. This inquired questions to validate every findings of this study because the illegal user responses may not be reliable.

The Survey targeted on validating the findings from this study. The questions were formulated in an indirect way to obtain direct answers to validate the patterns and significant facts. This was posted to bitcoin forum Under two different topics in 'Bitcoin Discussion' and ' off-topic ' section using the user name " Rosecuppy123 ". In addition, this was also posted in Reddit forum in r/Bitcoin using the user name "ray123amee". There were about 27 complete responses for the survey. The graph of the responses are attached as appendix and below is a summary of the responses.

The majority of the respondents for the survey is frequent bitcoin users who has used bitcoins more than 50 times. There is a variety of the number of bitcoin addresses they used starting from one address to above 50 addresses. Majority of them are using bitcoin exchanges. The popular exchanges are LocalBitcoins, Bittrex, Binance. Additionally Cryptopia, Coins Pro, Jaxx, MorphToken, Bitcoin ATM and Coinb.in were used whereas few said that they use almost all exchanges and some preferred not to mention the exchange.

It was evident that majority of the number of wallets they maintained was in range of 2 to 5. The reason behind using more wallets than single wallet was mostly because of the security aspect and increase in privacy. Some users pointed out that storing in one wallet would not secure and so that scattering across multiple locations would make transactions secure. Some other users have mentioned that there is a high risk in loss keeping only one wallet and so that many wallets are used by them as backups. Another popular reason was they prefer using different wallets for different use

[42] https://goo.gl/forms/7mesAY61CyYlkThx2

cases. For instance, hardware wallet for secure savings and mobile, web wallets for everyday use in small amounts.

The user feedback also highlight that most of the users send bitcoins to their own addresses belonging to existing wallets before sending someone else. However, they do not create new wallets for every single transaction before cashing out via an exchange.

The purposes of using bitcoins were inquired at the end of the survey to identify whether the respondent is a legal or illegal user. The popular intentions of the users were investment and goods and service trade. But few users have responded for darknet marketplace and other purposes which may be more towards to illegal purposes. However, since we cannot rely on the response for the purpose of using bitcoins ; this study did not use response for this to evaluate the overall feedback.

In Conclusion, it can be summarized that majority of the frequent users use around 2 - 5 wallets for the storage for security purposes. The users are also of practise of sending bitcoins to their own addresses belonging to existing wallets before sending someone else. However, they do not create new wallets for every single transaction before cashing out via an exchange unlike an illegal users. So, we can confirm illegal users create several new wallets with new addresses to spread coins.

In addition, popular exchanges are LocalBitcoins, Bittrex, Binance, Cryptopia, Coins Pro, Jaxx, MorphToken, Bitcoin ATM, Coinb.in and so on. Whereas illegal users are more into exchanges such as Xapo.com, Helix Mixer[18], Bittrex, CEX.IO, Poloniex.com and uses exchanges that are untagged.

However, based on this evaluation approach of user feedback, results cannot be validated. Since this approach will not work and are not suitable for researches of this nature. Due to number of reasons. One of the reasons is that an illegal person can pretend to be a legal individual and can provide feedback which will be misleading. Another reason would be that most of the respondents in the bitcointalk forum and reddit were not willing to provide information as they were of more suspicious of the google survey link provided. In addition, users were asking for further information

about the project and were more suspicious on providing information. Few of the responses are attached in *Figure 5.4*. This can be validated with the experience during the research process.

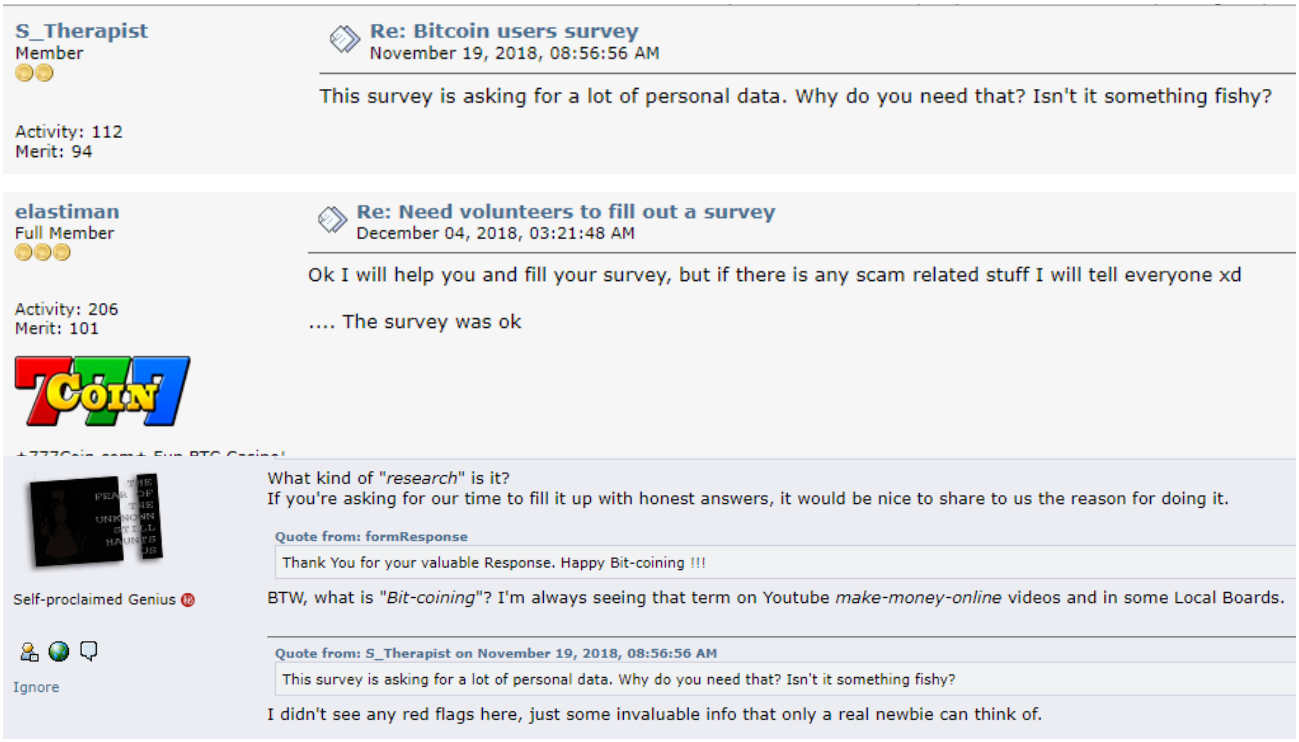


Figure 5.4 : Screenshots of bitcoin user comments with regards user feedback survey

In addition, as for an example when asked for mobile money user transaction data from an private entity through the police department of Sri Lanka in order to analyses user behavior in relation to bitcoin user behavior. It was not possible to obtain the data even with help of officials. Therefore, user feedback evaluation is not a suitable approach for user behavioral analysis of bitcoin illegal incidents.

Chapter 06

Conclusion and Future Work

6.1 Conclusion

Bitcoin is a crypto currency that is being used by millions of people for both legal and illegal intentions. The decentralized and anonymization feature of Bitcoin has drastically increased the rate of Bitcoin being misused, particularly its involvement in illegal activities. This triggered to conduct a comprehensive analysis of several illegal incidents of different nature. The main objective is to provide a more comprehensive and timely information about illegal user behaviors. This aims for the bitcoin community, authorities, lawmakers and officials to intervene and reduce future illegal incidents. In order to analyze the Bitcoin Blockchain with respect to the illegal incidents, it was required to automate the data extraction process. Thereafter, the transaction data were processed with regards to incidents and results were visualized using Gephi to perform the analyses of Bitcoin transactions. Thereon, the significant transactions were obtained and analyzed further.

The study revealed that, illegal users repeatedly transact with bitcoin exchanges or mixtures or to newly created wallets. The illegal users transact more in smaller chunks of bitcoins repeatedly. In addition, they also transact in smaller chunk by sending constant value of bitcoins or in certain proportion of bitcoins that gradually reduces in subsequent transactions. The study also confirms the on the popularly used exchanges and mixtures such as Xapo.com, Helix Mixer, Bittrex, Cex.io. In addition, it revealed the exchange services Poloniex.com and untagged services being well used. Accordingly, we can conclude there are similar patterns among similar natured and dissimilar natured incidents.

6.2 Contribution

The study has contributed to research community and bitcoin community in several means.

The major contribution of this study is the notable patterns on how the bitcoins have been spread. The significant facts on how the bitcoins that were used in illegal transactions has circulated and the means of how it was cashed out. The findings specific to this research are, that illegal users

use new wallets to spread bitcoins in subsequent transactions, the transactions are done continuously by reducing the total bitcoin value in single wallet or address using a constant bitcoin amount or by a certain proportion. In addition, this study also contributed to identify popular exchanges that are used such as Poloniex.com.

In addition, we had provided validations for the findings made by previous researchers. It was validated that the illegal users tend to transact more in smaller amounts repeatedly to prevent getting noticed. They often use exchanges or mixing services to prevent tracking by investigators. In addition, the behavior of ransomware criminals were also confirmed whereas, some of popular exchanges or services such as Xapo.com, Helix Mixer that are usually used by criminals were confirmed by our findings.

The next contribution is the coverage of ten illegal incidents that involved bitcoin which are of dissimilar nature. There has no evidence of previous researches that covered illegal incidents that are of different nature. Even though the previous researchers carried out for separate crimes involving Bitcoin, this research is the first wider analysis.

Another major contribution to the research community is the transaction data of illegal incidents. It has been attached in the appendix section for the use of research community. In addition, the soft copy can be accessed via this google sheet[43]. This data contribution from this study is significant as the incidents data involving Bitcoins are scattered in bitcointalk forum, online forums and official websites. Based on the experience, there is no single repository to obtain illegal incident data after the year of 2014. Thus, this dataset will act as a single repository to get incident data involving bitcoin. The data gathered from 2012 to 2018 on 33 illegal incidents including its respective 331 transaction ids will be of much use for future researchers. In addition, the details such as incident name, nature of the incident, transaction ids, bitcoin addresses, and sources of data are included in the dataset. On another perspective, the dataset will be of much use to the bitcoin community users. Since, they are interested to know about the bitcoin world as observed in the forum (*Figure 5.5*) and are shown the risky side of using bitcoin as well.

[43] https://docs.google.com/spreadsheets/d/1fOUIA9J4-lJKhgXqh2_zH6__t1BBRjaPMExH3GeFi6w/edit?ts=5bd593ec#gid=1694620899

Figure 6.1: Screenshots of bitcoin user comments with regards to data availability of incidents

6.3 Ethical Considerations

This study obtained data from publicly available sources such as Blockchain.info, online forums, official websites etc. In addition, data obtained via survey from online communities has an ethical consideration which is preserved and users were given consent of data confidentiality. Another ethical consideration would be including certain screenshots of comments from bitcoin user community for the purpose of thesis. Since these forums are open it is not unethical in extracting their feedback for information purposes. Therefore, there were no issues arisen on ethical perspectives.

6.4 Delimitations

This research study was limited to selected number of incidents out of total collected data, for which its transaction ids were available from a reliable source. So only ten incidents were considered covering five subcategories under three main categories of illegal activities.

The number of records for each transaction id that was considered was limited to approximately 1.3 million records since the time consumption in gaining results and the memory requirement in visualizing the results was high.

6.5 Future Work

One direct future work would be to consider a wider range of many other illegal incidents that has happened using bitcoin. It would also be a value addition to consider ScriptHashAddresses. Since, the wrapped addresses may represent some kind of a relationship to the illegal entities since those addresses are required for some sort of a consent (i.e. based on the requirement. As for example. Multisignature addresses require another user or users to sign a transaction before it can be broadcast onto the Blockchain) in order to spend the tainted Bitcoins. So the analysis can be broadened if those addresses also taken into the further investigation.

Another future work would be to use Big data analytics. One of the limitations of this study is transaction data restricted to 1.3 million records for each incident due to limitations on data processing, time consumption in getting results and the memory issues in visualizing the results. Therefore, a big data analytics will help to solve these barriers.

In addition, another significant future work would be for the purpose of evaluation. A model of unsupervised learning in machine learning can be used by training the transactions data and testing it with a new set of incident transaction data. So the association rule in machine learning would assist to discover the relations in transactional data.

Moreover, another highlightable recommended future work would be to map the behavior of Bitcoin illegal users to the other crypto currencies or with fiat currency illegal users. This would then enable to identify similarities among different currency illegal users.

Bibliography

1. S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” 2008. [Online]. Available: https://bitcoin.org/en/bitcoin-paper. [Accessed: May 07, 2018].

2. D. G. Baur, K. Hong, and A.D. Lee, “Bitcoin: Medium of exchange or speculative assets?” *Journal of International Financial Markets, Institutions and Money*, 2017.

3. R. Raeesi, “The Silk Road, Bitcoins and the Global Prohibition Regime on the International Trade in Illicit Drugs: Can this Storm Be Weathered?” *Glendon Journal of International Studies*, vol. 8, no. 1, 2018.

4. R. Böhme, N. Christin, and B. Edelman, "Bitcoin: Economics, Technology, and Governance," *Journal of Economic Perspectives*, pp. 213-238, 2015.

5. L. Trautman, “Virtual Currencies Bitcoin & What Now After Liberty Reserve, Silk Road, and Mt. Gox. Richmond,” *Journal of Law and Technology, pp.* 1-108, 2014.

6. M. Ly, “COINING BITCOIN’S “LEGAL-BITS”: EXAMINING THE REGULATORY FRAMEWORK FOR BITCOIN AND VIRTUAL CURRENCIES,” *Harvard Journal of Law & Technology*, vol. 27, no. 2, 2014.

7. V. Marella, "Bitcoin: A Social Movement under Attack," *Selected Papers of the IRIS - European Journal of Philosophy and Public Debate*, 2017. [Online]. Available: http://aisel.aisnet.org/iris2017/1/ [Accessed: Aug. 17, 2018].

8. S. M. Foley, J. R. M. Karlsen, and T. J. M. Putniii, “Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed Through Cryptocurrencies?” *SSRN Electronic Journal*, 2018.

9. J. Bohr, and M. Bashir, “Who Uses Bitcoin? An Exploration of the Bitcoin Community,” in *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, Toronto, Canada, 2014.

10. C. Janze, “Are Cryptocurrencies Criminals Best Friends? Examining the Co-Evolution of Bitcoin and Darknet Markets,” in *Proceedings of the Americas Conference on Information Systems (AMCIS)*, America, 2017. [Online]. Available:

https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1041&context=amcis2017 [Accessed: Aug. 10, 2018].

11. S. Kethineni, Y. Cao, and C. Dodge, "Use of Bitcoin in Darknet Markets: Examining Facilitative Factors on Bitcoin-Related Crimes," *American Journal of Criminal Justice*, vol. 43, no. 2, pp. 141-157, 2017.

12. A. Baravalle, M. S. Lopez, and S. W. Lee, "Mining the Dark Web - Drugs and fake ids," in *2016 IEEE 16th International Conference on Data Mining Workshops*, Barcelona, 2016.

13. M. Vasek, M. Thornton, and T. Moore, "Empirical Analysis of Denial-of-Service Attacks in the Bitcoin Ecosystem," in *18th Conference on Financial Cryptography and Data Security*, Barbados, 2014, pp. 57-71.

14. M. Vasek, M. Thornton, and T. Moore, "There's No Free Lunch, Even Using Bitcoin: Tracking the Popularity and Profits of Virtual Currency Scam," in *18th Conference on Financial Cryptography and Data Security*, Barbados, 2014, pp. 44-61.

15. T. Moore, and N. Christin, "Beware the Middleman: Empirical Analysis of Bitcoin-Exchange Risk," in *18th Conference on Financial Cryptography and Data Security*, Barbados, 2014, pp.25-33.

16. R. Richardson, and M. North, "Ransomware: Evolution, Mitigation and Prevention," in *Scholars' Press*, 2018. [Online]. Available: http://scholarspress.us/journals/IMR/pdf/IMR-1-2017.%20pdf/IMR-v13n1art2.pdf [Accessed: Aug. 12, 2018].

17. D. Huang, M. Aliapoulios, V. Li, L. Invernizzi, E. Bursztein, K. McRoberts, J. Levin, K. Levchenko, A. Snoeren, and D. McCoy, "Tracking Ransomware End-to-end," in *2018 IEEE Symposium on Security and Privacy (SP)*, San Francisco, 2018.

18. M. Paquet-Clouston, B. Haslhofer, and B. Dupont, "Ransomware Payments in the Bitcoin Ecosystem," 2018. [Online]. Available: https://arxiv.org/abs/1804.04080 [Accessed: Aug. 15, 2018].

19. M. Karami, and D. McCoy, "Understanding the Emerging Threat of DDoS-as-a-Service," 2018. [Online]. Available: https://www.usenix.org/conference/leet13/workshop-program/presentation/karami [Accessed: Aug. 12, 2018].

20. S. Barber, X. Boyen, and E. Shi, "Bitter to Better—How to Make Bitcoin a Better Currency," in *16th Conference on Financial Cryptography and Data Security*, Bonaire, 2012, pp. 399–414.

21. A. Gifari, B. Anggorojati, S. Yazid, "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.

22. D. Bryans, "Bitcoin and Money Laundering: Mining for an Effective Solution," *Digital Repository @ Maurer Law*, 2018. [Online]. Available: https://www.repository.law.indiana.edu/ilj/vol89/iss1/13/ [Accessed: Aug. 12, 2018].

23. W. Shao, H. Li, M. Chen, C. Jia, C. Liu, and Z. Wang, "Identifying Bitcoin Users Using Deep Neural Network," in: J. Vaidya, and J. Li (eds), Algorithms and Architectures for Parallel Processing, *ICA3PP 2018*. Lecture Notes in Computer Science, Springer, Cham, 2018, vol. 11337, pp. 178-192.

24. C. Rose, "The Evolution of Digital Currencies: Bitcoin, A Cryptocurrency Causing A Monetary Revolution," *International Business & Economics Research Journal (IBER)*, vol. 14, no. 4, p.617, 2015.

25. G. Karame, E. Androulaki, and S. Capkun, "Double-spending fast payments in bitcoin," in *2012 ACM conference on Computer and communications security - CCS '12*, Raleigh, NC, USA, 2012.

26. E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating User Privacy in Bitcoin," in: A. R. Sadeghi (eds), *Financial Cryptography and Data Security*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2013, vol. 7859, pp. 34-51.

27. J. Barcelo, "User Privacy in the Public Bitcoin Blockchain," 2018. [Online]. Available: https://pdfs.semanticscholar.org/549e/7f042fe0aa979d95348f0e04939b2b451f18.pdf [Accessed: Aug. 15, 2018].

28. J. V. Monaco, "Identifying Bitcoin users by transaction behavior," in *SPIE DSS*, 2015. doi: 10.1117/12.2177039.

29. R. Anderson, I. Shumailov, M. Ahmed, and A. Rietmann, "Bitcoin Redux," *17th Annual Workshop on the Economics of Information*, 2018.
30. N. Bhaskar, and D. Chuen, *Bitcoin Exchanges. Handbook of Digital Currency*, 2015, pp. 559-573.
31. CoinDesk. (2017, Dec. 29). *From $900 to $20,000: Bitcoin's Historic 2017 Price Run Revisited* [Online]. Available: https://www.coindesk.com/900-20000-bitcoins-historic-2017-price-run-revisited/ [Accessed: Aug. 18, 2018].
32. CoinDesk. (2018, Jan. 20). *What can you buy with bitcoins?* [Online]. Available: https://www.coindesk.com/information/what-can-you-buy-with-bitcoins/ [Accessed: May 08, 2018].
33. G. Hurlburt, and H. Bojanova, "Bitcoin: Benefit or Curse?" *IEEE Computer Society Press*, 2014.
34. T. Moore, N. Christin, and J. Szurdi, "Revisiting the Risks of Bitcoin Currency Exchange Closure," 2018. [Online]. Available: https://www.semanticscholar.org/paper/Revisiting-the-Risks-of-Bitcoin-Currency-Exchange-Moore-Christin/a9d7fac962ba0cc707406b34e45c2bea32875e48 [Accessed: Aug. 13, 2018].
35. T. Moore, J. Han, and R. Clayton, "The Postmodern Ponzi Scheme: Empirical Analysis of High-Yield Investment Programs," in A. D. Keromytis (eds), *Financial Cryptography and Data Security*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2012, vol. 7397, pp.41-56.
36. G. Weimann, "Going Dark: Terrorism on the Dark Web. Studies in Conflict & Terrorism," vol. 39, no. 3, pp. 195-206, 2016. doi: 10.1080/1057610X.2015.1119546.
37. N. Christin, "Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace," in *Proceedings of the IW3C2 WWW 2013 Conference*, Rio de Janeiro, Brazil, 2013, pp. 213–224.
38. D. Ron, and A. Shamir, "How Did Dread Pirate Roberts Acquire and Protect his Bitcoin Wealth?," *IACR Cryptology ePrint Archive*, 2013, [Online]. Available: https://eprint.iacr.org/2013/782.pdf [Accessed: Jun. 21, 2018]. doi: 10.1007/978-3-662-44774-1_1,
39. R. Upadhyaya, and A. Jain, "Cyber ethics and cyber crime: A deep dwelved study into legality, ransomware, underground web and bitcoin wallet," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, Noida, 2016.

40. P. B. Pathak, "A Dangerous Trend of Cybercrime: Ransomware Growing Challenge," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 5, no. 2, 2016. [Online]. Available: http://ijarcet.org/wp-content/uploads/IJARCET-VOL-5-ISSUE-2-371-373.pdf [Accessed: Aug. 12 2018].
41. A. Kharaz, W. Robertson, D. Balzarotti, L. Bilge, and E. Kirda, "Cutting the Gordian Knot: A Look Under the Hood of Ransomware Attacks," *Detection of Intrusions and Malware, and Vulnerability Assessment*, vol. 9148, no. 3, pp. 3-24, 2001. doi: 10.1007/978-3-319-20550-2_1.
42. C. Everett, "Ransomware: to pay or not to pay?" *Computer Fraud & Security*, vol. 2016, no. 4, pp.8-12, 2016.
43. K. Liao, Z. Zhao, A. Doupe, and G. Ahn, "Behind closed doors: measurement and analysis of CryptoLocker ransoms in Bitcoin," in *2016 APWG Symposium on Electronic Crime Research (eCrime)*, Toronto, ON, Canada, 2016. doi: 10.1109/ECRIME.2016.7487938.
44. G. Grant, and R. Hogan, "Bitcoin: Risks and Controls," *Journal of Corporate Accounting & Finance*, vol. 26, no. 5, pp.29-35, 2015.
45. K. Krombholz, A. Judmayer, M. Gusenbauer, and E. Weippl, "The Other Side of the Coin: User Experiences with Bitcoin Security and Privacy," in J. Grossklags, and B. Preneel (eds), *Financial Cryptography and Data Security*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2017, vol. 9603, pp.555-580. [Online]. Available: https://www.sba-research.org/wp-content/uploads/publications/TheOtherSideOfTheCoin_FC16preConf.pdf [Accessed: Jun. 14, 2018]
46. C. Fromknecht, "One-Time, Zero-Sum Ring Signature," 2015. [Online]. Available: https://scalingbitcoin.org/papers/one-time-zero-sum-ring-signature-conner-fromknecht-2015.pdf [Accessed: Aug. 15, 2018]
47. A. Gervais, G. Karame, S. Capkun, and V. Capkun, "Is Bitcoin a Decentralized Currency?" *IEEE Security & Privacy Press, vol.* 12, no. 3, pp. 6-7, 2014. [Online]. Available: https://eprint.iacr.org/2013/829.pdf [Accessed: Aug. 13, 2018].
48. R. Anderson, I. Shumailov, M. Ahmed, "Making Bitcoin Legal," [Online]. Available: https://www.cl.cam.ac.uk/~rja14/Papers/making-bitcoin-legal.pdf [Accessed: Aug. 12, 2018].
49. T. Ruffing, and P. Moreno-Sanchez, "ValueShuffle: Mixing Confidential Transactions for Comprehensive Transaction Privacy in Bitcoin," in M. Brenner et al. (eds), *Financial*

Cryptography and Data Security. Lecture Notes in Computer Science, Springer, Cham, 2017, vol. 10323, pp. 133–154.

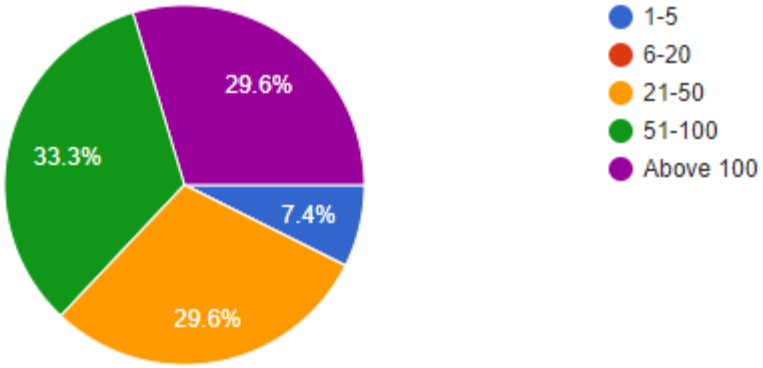
50. Y. Hong, H. Kwon, S. Lee, and J. Hur, "Poster: De-mixing Bitcoin Mixing Services," [Online]. Available: https://pdfs.semanticscholar.org/5425/fb2c0a039bc16e5a4fe31a1b493094631462.pdf [Accessed: Aug. 22, 2018].
51. D. Ron, and A. Shamir, "Quantitative Analysis of the Full Bitcoin Transaction Graph," in A. R. Sadeghi (eds), *Financial Cryptography and Data Security*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2013, vol. 7859, pp. 6-24.
52. F. Reid, and M. Harrigan, "An Analysis of Anonymity in the Bitcoin System," *Security and Privacy in Social Networks*, pp. 197-223, 2012.
53. H. Kalodner, S. Goldfeder, A. Chator, M. Möser, and A. Narayanan, "BlockSci: Design and applications of a blockchain analysis platform," 2017. [Online]. Available: https://arxiv.org/abs/1709.02489 [Accessed: Aug. 12, 2018].
54. G. S. Stanford, and T. H. Stanford, "Cointopia: Blockchain Analysis using Online Forums," 2017. [Online] Available: https://web.stanford.edu/class/cs224w/projects/cs224w-87-final.pdf [Accessed Aug. 12, 2018].
55. D. McGinn, D. Birch, D. Akroyd, M. Molina-Solana, Y. Guo, and W. Knottenbelt, "Visualizing Dynamic Bitcoin Transaction Patterns", *Big Data*, vol. 4, no. 2, 2016.
56. M. Rodrigues, J. Gama, and C. A. Ferreira, "Identifying Relationships in Transactional Data" in J. Pavón, N. D. Duque-Méndez, and R. Fuentes-Fernández (eds), *Advances in Artificial Intelligence – IBERAMIA 2012*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2012, vol. 7637, pp. 81-90.
57. E. Rojas, D. Gorton, and S. Axelsson, "Using the RetSim simulator for fraud detection research," *International Journal of Simulation and Process Modelling*, vol. 10, no. 2, 2015.
58. CoinDesk. (2017, Dec. 6). *Cryptocurrency Mining Market NiceHash Hacked* [Online]. Available: https://www.coindesk.com/62-million-gone-cryptocurrency-mining-market-nicehash-hacked [Accessed: Oct. 18, 2018].
59. M. Conti, A. Gangwal, and S. Ruj, "On the Economic Significance of Ransomware Campaigns: A Bitcoin Transactions Perspective," *Computers & Security*, vol. 79, pp. 162-189, 2018. doi: 10.1016/j.cose.2018.08.008.

60. CoinDesk. (2016, Apr. 18). *ShapeShift Lost $230k in String of Thefts, Report Finds* [Online]. Available: https://www.coindesk.com/digital-currency-exchange-shapeshift-says-lost-230k-3-separate-hacks/ [Accessed: Aug. 22, 2018).
61. CoinDesk. (2016, May 16). *Gatecoin Claims $2 Million in Bitcoins and Ethers Lost in Security Breach* [Online]. Available: https://www.coindesk.com/gatecoin-2-million-bitcoin-ether-security-breach [Accessed: Oct. 18, 2018].
62. J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where Is Current Research on Blockchain Technology?—A Systematic Review," *PLoS ONE*, vol. 11, no. 10, 2016. doi: 10.1371/journal.pone.0163477
63. V. R. Patil, A. P. Nikam, J. S. Pawar, and M. S. Pardhi, "Bitcoin Fraud Detection using Data Mining Approach," *Journal of Information Technology and Sciences*, vol.4, no. 2, 2018.
64. D. Zambre, and A. Shah, "Analysis of Bitcoin Network Dataset for Fraud," 2013. [Online]. Available: http://snap.stanford.edu/class/cs224w-2013/projects2013/cs224w-030-final.pdf [Accessed: Oct. 21, 2018].
65. M. Bartoletti, B. Pes, and S. Serusi, "Data mining for detecting Bitcoin Ponzi schemes," 2018. [Online]. Available: https://arxiv.org/pdf/1803.00646 [Accessed: Oct. 21, 2018].

Appendix A: Survey

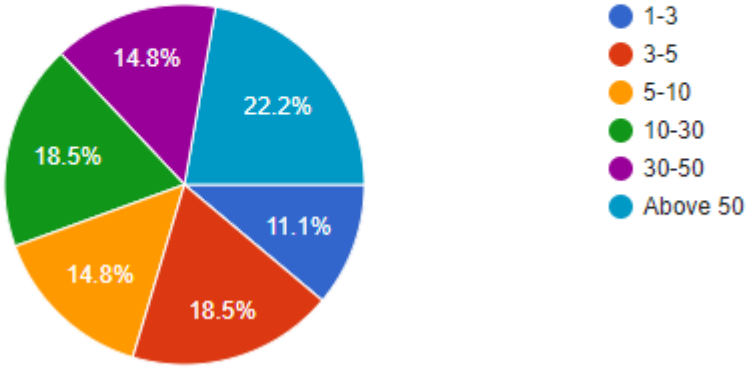
How many times you have used Bitcoin?

27 responses



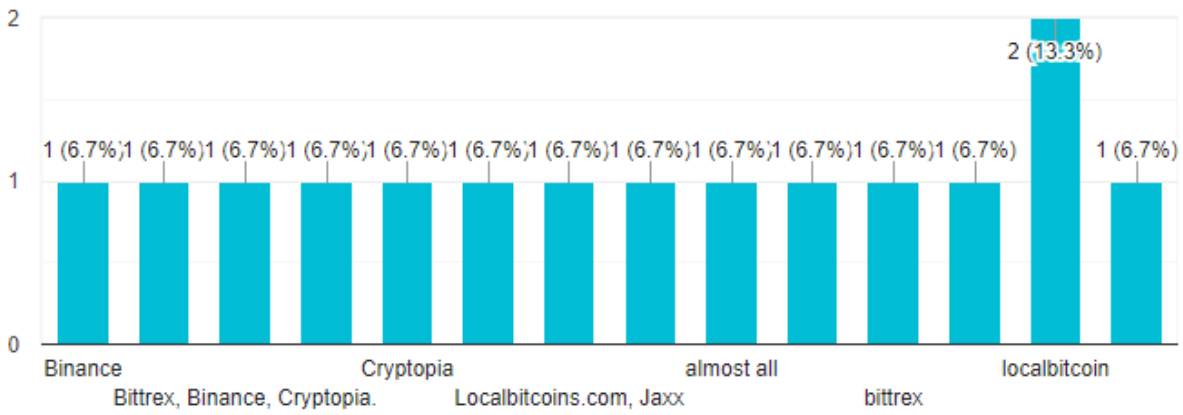
How many bitcoin addresses do you have?

27 responses



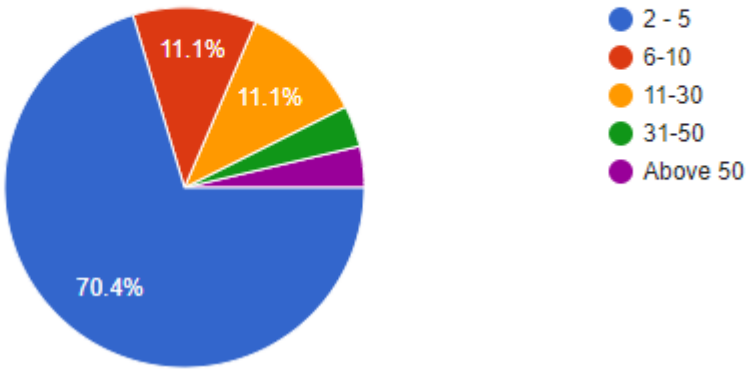
What exchanges you prefer?

15 responses



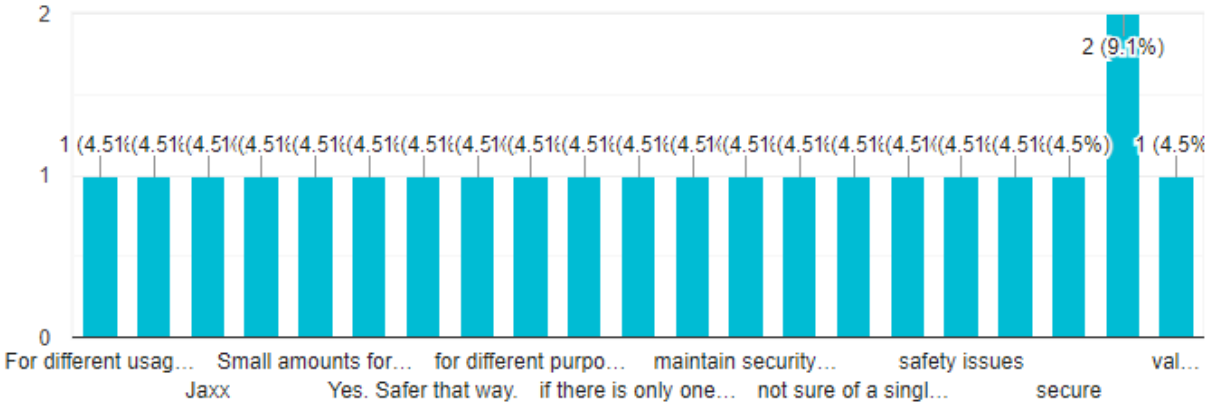
How many wallets do you have?

27 responses



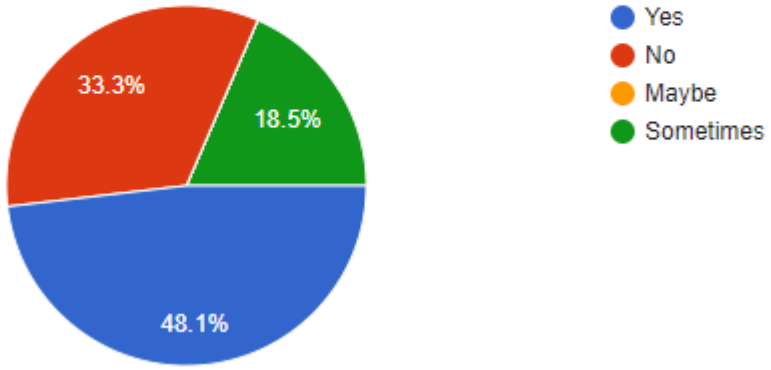
Why do you prefer using many wallets?

22 responses



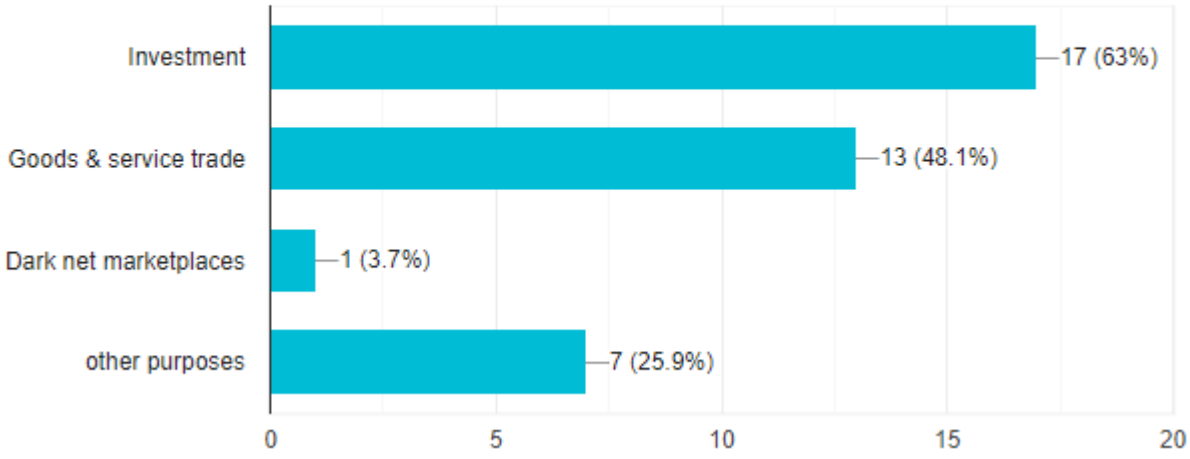
Do you send bitcoins to your own addresses before sending bitcoin to another person

27 responses



Why do you use bitcoin?

27 responses



Appendix B : Data