Data Driven Analysis of Darknet Market Network Based on Bitcoin Transaction Data

By

D.W.H. Silva

15001301

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 Computer Science

University of Colombo School of Computing
35, Reid Avenue, Colombo 07,
Sri Lanka
July 2020

Declaration

I, D.W.H Silva and 2015/CS/130 hereby certify that this dissertation entitled "Data Driven Analysis of Darknet Market Network Based on Bitcoin Transaction Data" is entirely my own work and it has never been submitted nor is currently been submitted for any other degree.

Date	Student's Signature

I, Kasun de Zoysa, certify that I supervised this dissertation entitled "Data Driven Analysis of Darknet Market Network Based on Bitcoin Transaction Data" conducted by D.W.H Silva in partial fulfillment of the requirements for the degree of Bachelor of Science Honours in Computer Science.

Date	Supervisor's Signature

Abstract

In the past reason years, there are several research studies have been conducted to analyze the Darknet market network by using different approaches. However, from this research study, the novel approach has been presented to analyze the Darknet market network in order to investigate the dynamic behavior of the Darknet market network and the Darknet market users. The novel approach has been presented by introducing the analysis for the Darknet market network by using Bitcoin transaction data. The analysis has been presented in three phases by exploiting the graph models that have been constructed from the Bitcoin transaction data of the seven Darknet markets. Those seven Darknet markets are Abraxas Darknet market, Bluesky Darknet market, Cannabis Darknet market, Middle Earth Darknet market, Nucleus Darknet market, Pandora Darknet market, Sheep Darknet market.

In the first analysis phase, the overall Darknet markets behavior has been addressed by analyzing the transaction flow and the money flow of all Darknet markets. In this phase, the properties like inactiveness of the darknet markets have been discussed and derived conclusions based on overall Darknet markets behaviors.

In the second analysis phase, the analysis of the Darknet market has been addressed by exploiting the two graph models (Transaction graphs and User graphs) which constructed from Darknet market Bitcoin transaction data. The analysis has been conducted by measuring the connectivity in each Darknet market network and addressing the centrality measurements. In this phase, the conclusions based on scale-free networks and rich-get-richer properties have been addressed based on the graph models.

In the final phase, the main concern will be addressing the traceability analysis of the Darknet market users. In this phase, our approach has been succeeded to trace 1203 user entities inside the Darknet market networks. Additionally, the behavior of those user entities have been addressed according to the Darknet market transactions.

Preface

The existing approaches for analyzing the Darknet market network was unable to address the intended static and dynamic properties of the Darknet market networks due to several limitations. Therefore, our study has been introduced the novel approach for analyzing the static and dynamic properties of the Darknet market network by using the transactions that used Bitcoin Cryptocurrency.

There are seven Darknet markets that have been used to address the properties and behavior of the Darknet market network. At the end of this analysis, our work was able to introduce the novel approach to trace the Darknet market users and their behaviors using the surface web Bitcoin information. There are two types of data resource has been collected. The first resource is from www.walletexplorer.com which is used to collect the Darknet market transaction data and the second resource is from www.blockchain.com which was used to collect surface web Bitcoin information.

Acknowledgement

I would like to express my sincere gratitude to my research supervisor, Dr. Kasun de Zoysa, senior lecturer of the University of Colombo School of Computing and my research advisor Mr. Charith Elvitigala for providing the necessary guidance and supervisions throughout the research.

I would also like to extend my sincere gratitude to Dr. M. I. E Wickramasinghe, senior lecturer of the University of Colombo School of Computing and Dr. M.D.J.S Goonetillake, senior lecturer of the University of Colombo School of Computing for providing the necessary feedback on my research proposal and on my interim evaluation to improve this research work. I also take this opportunity to thank Dr. H.E.M.H.B Ekanayake, final year computer science project coordinator to provide the necessary assistance on the entire academic year.

Also, I would like to thank my family and my friends for their immense support and motivation to succeed in this research work. Finally, I would like to appreciate all of the people who are helping me to conclude my research work successfully.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

This chapter provides a detailed introduction to this research study. The background to this research study has discussed in section 1.1 and section 1.2 will be discussed the corresponding research problem and research questions. The relevant research aim and research objectives have discussed in section 1.3 and section 1.4 will be presented the justification for the identified research gap. The research methodology has addresses in section 1.5 and the outline of this dissertation has presented in section 1.6. Finally, the delimitations of the scope of this research have been addressed in section 1.7.

1.1 Background to the Research

As a result of the illegal merchandise on online Darknet marketplaces, the user of these illicit sites tends to trade illegal goods including drugs, hitman services, weapons, etc. In 2011, the darknet site called *Silk Road* has impacted to pioneer the illicit sites to the world via Dark web [1] and it has gained a huge amount of impact on the public society to attract on using Darknet markets to exchange illicit goods and services [1].

When considering the illegal merchandising, the vendors and the buyers in these markets always focus on protecting their anonymity within transactions. To provide the feasible anonymity to it's users, there are two main factors have been taken into consideration By the Darknet market itself. First, the transaction mechanism employed by the Darknet market and second, the Cryptocurrency that used in the Darknet market transactions. Concerning the transaction mechanism, there are three available transaction mechanisms that employed in the Darknet market have been identified. The detailed descriptions of those three approaches will be discussed in section 2.3. Aside from the transaction mechanisms, second main factor to be considered regarding the anonymity is the Cryptocurrency. Cryptocurrency has identified as one of the main anonymity providing factors inside the Darknet market [2]. In re-

cently, there are several popular Cryptocurrencies have been used inside the Darknet market transactions such as Bitcoin [3], Monero coin [4] and Etherum [5]. After several research studies, the Bitcoin has identified as the most trending and most widely used cryptocurrency among the Darknet markets [6,7]. However, the Bitcoin-related transactions have been considered as pseudonymous since the transactions occur via public key addresses of the involved parties without mentioning the real-world information of the involved parties [8, 9]. Further details about the usage of Bitcoin cryptocurrency on Bitcoin transactions have discussed in Section 2.4.

Along with those aspects of the Darknet market transactions and the Bitcoin cryptocurrency, this research study has been focused on address the data-driven analysis of the Darknet market network using the Bitcoin transaction data.

1.2 Research Problem and Research Questions

1.2.1 Research Problem

As mentioned in the section 1.1, after the huge publicity that gained from *SilkRoad* Darknet market, the user community and the income of the Darknet markets have been increased day by day with the help of nature in the Darkweb [10]. As a result, there are several analyses have been conducted to investigate this illicit merchandising and illicit users in the Darknet market networks. Those analyses have been discussed in the Chapter 2. However, there are several constraints and limitations have been identified in the existing approaches of analyzing the Darknet market network. Those constraints and limitations will be discussed in section 2.2. Therefore, the primary motivation for this research is to introduce proper analysis to mitigate those limitations and introduce a proper analysis of the Darknet market network in order to investigate properties and the dynamic behavior of the Darknet market network and investigate the Darknet market users who involved with the Darknet market tradings.

1.2.2 Research Questions

Considering the research problem, there are several research questions can be identified. Those research questions are as follows:

- What are the existing approaches for analyzing the Darknet market network?
- What are the existing analysis techniques based on the Bitcoin transaction data?
- How to construct the transaction graphs and user graphs using the Bitcoin transaction data of Darknet markets?

- How to analyze the Darknet market network by exploiting graph models (transaction graphs and user graphs)?
- How to trace the Darknet market users from using the Bitcoin data?

1.3 Research Aim and Objectives

1.3.1 Research Aim

There are two main research aims have been identified. First, presenting the analysis of the Darknet market network by addressing the several graph properties based on the constructed graph models. Second, presenting the traceability analysis of the Darknet market users by using the Bitcoin data.

1.3.2 Research Objectives

The objectives of this research study have defined to achieve the research aims by providing the solutions to the research questions that have discussed in section 1.2.2. The following objectives will be achieved throughout this research study,

- Identifying the existing approaches for analyzing the Darknet market network.
- Identifying the existing analysis techniques based on the Bitcoin transaction data.
- Constructing the graph models (transaction graphs and user graphs) using the Bitcoin transaction data of Darknet markets.
- Analyzing the Darknet market network by exploiting graph models (transaction graphs and user graphs) that have constructed using the Bitcoin transaction data from the Darknet markets.
- Identifying the Darknet market users by using the traceability analysis based on the Bitcoin data.

1.4 Justification for the research

When justifying the first research question, section 2.2 justified that analyzing the Darknet market network by using transaction data might be the proper approach for analyzing the Darknet market rather than the existing approaches such as analyzing using the drug-related listings [11–13], analyzing using the Darknet market discussion forums data [1,14], and analyzing using the product-related photos [15]. According to the justifying the second research question, section 2.5 has provided the justification by stating that graph-based analysis will be a proper approach of analyzing when using the Bitcoin transaction data rather than non-graph based approaches [16,17]. Further, section 2.5 has provided that a non-graph based approach will not be a proper approach when analyzing the dynamic and topological behaviors in the network by using the Bitcoin transaction data.

The third research question investigates the proper approach of constructing the transaction graphs and user graphs by using the Bitcoin transactions. Further, this research question helps to justify the choice of optimal and scalable address clustering tools and algorithms to generate the clusters of addresses. The details of constructing the graph models will be discussed in section 3.2.1. According to the justification of the fourth research question, section 3.2.2.1 and section 3.2.2.2 justified by providing the proper approach of analyzing the Darknet market network by exploiting the constructed graph models. In addition, this research question helps to justify the appropriateness of addressing the graph properties on the Darknet market network in order to investigate the behavior and the users of the network. Finally, when justifying the fifth research question, section 3.2.2.3 justified that the choice of the path to trace the Darknet market users will be possible with using the Bitcoin data.

1.5 Methodology

1.5.1 Research Approach

Since this research study relies on analyzing the existing data of the Bitcoin transactions in the Darknet markets, the quantitative research approach has been recognized as a suitable approach for this study. In addition, the quantitative approach will be provided feasible solutions to the research questions in order to obtain the results from evaluating the graph properties from the graph models and to obtain the results from traceability analysis over the Darknet market users.

1.5.2 Method of Data collecting

In this research study, the process of data collection was done using two main resources. First, Blockchain.com [18] and Second, WalletExplorer.com [19]. The purpose of collecting the data from [18] is to gather the real-world information of the Bitcoin users that are available on the surface web in order to use for the traceability analysis studies. In other hand, the purpose of collecting data from [19] is to gather the Bitcoin transaction data from the Darknet markets in order to use for graph models analysis. When considering the data collection from [19], the most active Darknet markets have been focused. The detailed description of the Data gathering and Dataset will be discussed in section 3.1.

1.5.3 Method of Analyzing

Before step into the analysis phase, the constructing of graph models has been done by modeling the two types of graph models namely, transaction graphs and using graphs. After the process of graph modeling, the three phases of analysis have been done using the quantitative approach. The first phase of analysis relies on the overall analysis of the all Darknet markets, second phase of analysis relies on addressing graph properties by using the constructed graph models and the third phase relies on traceability analysis over the Darknet market users. The detailed description of the analysis phase will be discussed in section 3.2.2.

1.6 Outline of the Dissertation

This dissertation has consisted of 6 chapters as mentioned below.

1. **Introduction**
   This chapter comprises the specific introduction to the research by describing several aspects of this research study.

2. **Literature Review**
   This chapter provides the details of the related works on existing analysis approaches of the Darknet markets, existing techniques for analyzing using Bitcoin transaction data and existing studies of analysis on Bitcoin graph models.

3. **Research Design**
   the chapter provides the details of the overall research design of this research study.

4. **Implementation**
   This chapter provides the overall implementation of this research study. This will includes the implementation of the web crawling, web scraping and analysis phases.

5. **Results and Evaluation**
   This chapter will provide the results and evaluation of this research study and it has been presented in three analysis phases.

6. **Conclusion**
   This chapter will discuss the overall derived conclusions from this research study. It consists the conclusions about research questions, conclusions about the research problem, limitations of this research study and future research works from this research

1.7 Delimitations of Scope

In this research study, there are three delimitations have been considered in the scope of the research. Those delimitations have mentioned below.

- **Implementing the clustering algorithm in order to construct the user graphs.**
  the research study does not consider implementing a clustering algorithm since the clustering technique provided from the BlockSci [20] tool has recognized as the more optimal and scalable approach.

- **Analysis of other complex networks such as Banks and Social networks.**
  analysis has only focused on the Darknet market network and it has not extended for other complex networks such as Banks and Social networks.

- **Analysis of using other cryptocurrencies such as Monerocoin [4] and Ethereum [5].**
  section 2.4 provides the justification to choose only the Bitcoin cryptocurrency among other cryptocurrencies that used in the Darknet market. Therefore, this research study has not focused to use the other cryptocurrency such as Monerocoin [4] and Ethereum [5].

1.8 Summary

This chapter has provided a precise introduction of the research using several subsections. Initially, the Background concepts of this research study have been discussed in section 1.1. Then in section 1.2, the research problem was defined, and

research questions have been introduced accordingly. Hereafter, the research aim and research objectives have been presented according to the research question. Then in section 1.4, the justification has been addressed in order to identify the research gaps. After the justification section, the research methodology has been introduced and in section 1.6 the outline of the dissertation has been presented. In the final section, the delimitation of the research scope has been discussed.

Chapter 2

Literature Review

In this chapter, a review of related work on data-driven analysis of Darknet markets using Bitcoin transactions has discussed. Section 2.1 provides a brief introduction to the Darknet market and its properties. Then, the existing analysis approaches to Darknet markets were discussed in section 2.2. Section 2.3 will be discussing the abvailable transaction mechanisms that used in the current Darknet market. Hereafter, the section 2.4 will be discussed the Cryptocurrency involvement in the Darknet markets and the importance of analyzing the Darknet market using the Bitcoin Cryptocurrency. Section 2.5 and section 2.6 will discuss the existing analysis approaches using the Bitcoin transaction data and the existing studies related to the graph-based analysis using the Bitcoin transactions respectively.

2.1 Introduction to Darknet Market

In terms of content accessibility, the web can be divided into three layers such as Surface web, Deep web and Dark web [21].The Surface web is the top layer that compromised web content that can be accessible (indexable) using a search engine [10]. On the other hand, the second layer is the Deep web which comprises web content that has not been indexed by search engines such as Google and Yahoo and constitutes of all information that resides in autonomous databases [10]. Therefore, the Deep web requires an additional layer of encryption to be visited. Besides the Surface web and Deep web, the bottom layer considers as the Dark web which intentionally conceals the web content. Furthermore, the Dark web requires special software such as "The Onion Router (Tor)" to be visited [16]. The Dark web has linked to Darknet markets (also known as Crypto markets) which are commercial websites that Dark web users have involved in trading legal as well as illegal goods such as weapons, drugs and hitman services. Figure 2.1 illustrates the abstract view between the Surface web, Deep Web and Dark Web.

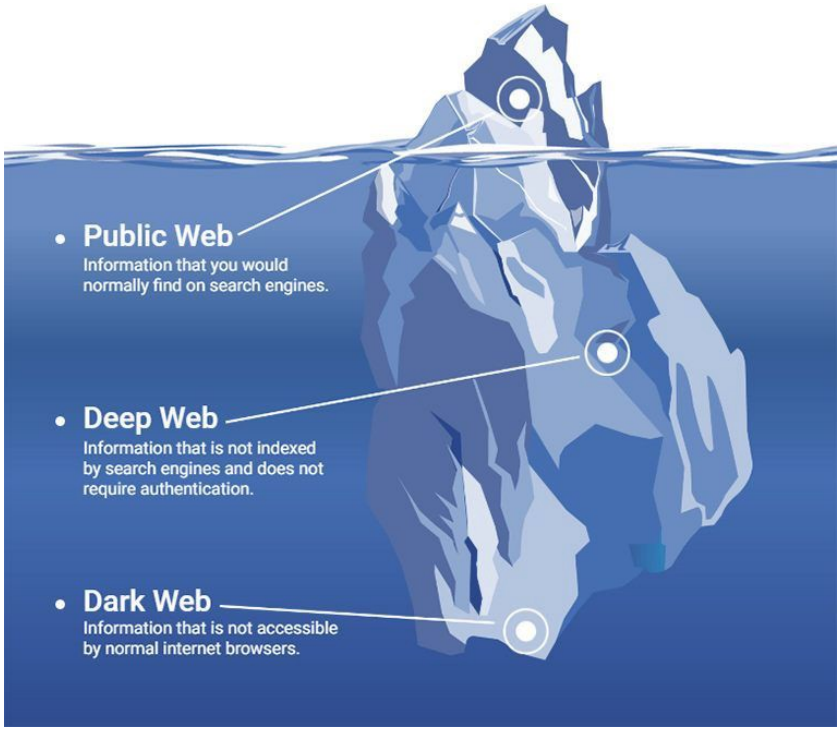


Figure 2.1: Abstract view between three layers of Webs using a picture of Iceberg

In 2011, the Darknet market called *Silk Road* has impacted to pioneer the illicit sites to the world via the Dark web. However, the *Silk Road* Darknet market has been shut down in late 2013 by the US Federal Bureau of Investigation (FBI) due to illicit drug trafficking, computer hacking and money laundering [1]. Shortly after the *Silk Road* shutdown, there was a huge amount of impact to increase the rapid activity on the other Darknet marketplaces. Therefore, there are more than 15 darknet marketplaces that are still active on the Dark web with highly operational on the illegal merchandising [22]. When concerning this illegal tradings, the vendors and the buyers of these markets always focus on protecting anonymity within their transactions on tradings over Darknet markets. From the Darknet market users perspective, the anonymity has been achieved by the two main network properties of the Darknet market [16, 23].First factor is the Tor hidden services which makes the IP addresses of both the client and the server unknown to each other and second factor is the cryptocurrency electronic payment system which is Bitcoin [3] that protects the true identity of the users (vendors and buyers) from the law enforcement agencies to trace users of the Darknet markets. Along with the anonymity and other properties of the Darknet markets, many researchers have been focused on addressing many aspects of the Darknet market and came up with several research studies based on many approaches for analyzing the Darknet Markets which will be further discussed in section 2.2.

2.2 Approaches for analyzing the Darknet Market Network

In the past recent years, there are several researchers have been focused to address different aspects of the Darknet market network by analyzing various properties such as communication between vendors and buyers, correlation and behaviors between the intra-market as well as inter-markets and most importantly, the anonymity of the market users. To address those properties, few approaches have been investigated by the researchers to analyze the Darknet market network.

One of the approaches used to analyze the Darknet market was based on the information of product-related listings[1] on the Darknet market websites such as Title of the product listings, shipping information, product prices, and product descriptions. Accordingly, several research studies have been conducted based on product listings. For instance, in 2016 one of the research studies was conducted to investigate illicit drug traffickings[2] from the Canadian perspective through the analysis of 3685 product-related listings on eight cryptomarkets [12]. Moreover, in 2017 there was a research study on geographical analysis over 11 countries on 92,980 product listings of cryptomarket called “Evolution” [13]. Similarly, in 2018 the researcher called ‘Ben R. Lane’ has introduced activities analysis called "EAST (Event Analysis of Systematic Teamwork)" and applied it to illicit trading on a Darknet market called “Dream Market” by using the 97,000 product listings [11]. However, some major limitations have been discussed by some researchers regarding the approach of analyzing the Darknet market network based on Darknet market product-related listings. One of the major limitations was that, the knowledge gathered through this kind of approach of analyzing was invaluable to design efficient policy for monitoring or repressive purposes against Darknet markets [12]. As a result of this limitation, this kind of approach cannot be used as the best approach for discussing the anonymity of the users (vendors and buyers) and their behaviors in the Darknet marketplaces.

Besides analyzing the Darknet market based on the product-related listings, some of the researchers have been introduced another approach for analyzing the Darknet markets which were based on the classification on product photos of the Darknet markets’ websites that vendors have posted [15]. For instance, there was a novel approach that has been introduced with deep neural networks to model photography styles and to link the multiple accounts of the same vendors in the Darknet market by using the product photos of 3 main darknet markets [10]. Even though this

[1] Product listings : lists of multiple products within a category, with each product represented by a photo, a price, and product name

[2] Drug trafficking : Illicit trade that involving sale of substances which are subject to drug prohibition laws

approach has demonstrated the feasibility of tracing markets' vendors' information, some limitations have been identified regarding the approach of analyzing based on the product photos of the Darknet markets. One of the main limitations was that, this photography style based analysis have drawbacks when there was intrinsic ambiguity arising from resale or photo plagiarizing [24].

To analyze the communication and behaviors between the vendors and the buyers, another alternative approach has been used to analyze the Darknet market network using the discussion forum data between vendors and buyers on the Darknet markets. For instance, one of the research studies has been proposed a classification method for analyzing the identities of the market users using the discussion forums data of the Darknet market [14]. Further, a researcher called *T. Reksna* has proposed a model to investigate the patterns in the Darknet market network based on discussion forums data of 26 darknet markets within the 4 years [1]. However, the limitations have been discussed according to the approach of analyzing the Darknet market network based on the discussion forum data. One of the main limitations is that this approach mainly focused on analyzing the internal behavior of the network rather than analyzing the external behavior between Darknet markets [25]. Therefore, there is a lack of analysis of the external correlation between Darknet markets. Further, one of the research studies has been identified that this kind of approach not sufficient for describing the network formation of Darknet markets [1].

Due to the aforementioned limitations in the approaches of analyzing the Darknet markets, this research study has been focused in to propose a novel approach to analyze the Darknet market network which is based on the transaction data of the Darknet market between vendors and buyers. The transaction data of the Darknet market has relied on several transaction mechanisms employed by the Darknet market itself. The behavior of those transaction mechanisms in the Darknet market will be discussed in section 2.3.

2.3 Transaction Mechanisms in the Darknet Market

The transactions of the Darknet market rely on different types of transaction mechanisms between vendors, buyers, and the Darknet market's platform administrators in order to maintain the success and durability of the Darknet markets [26]. The Darknet markets have employed different types of transaction mechanisms to provide an efficient and trusted environment to their users (vendors and buyers) to maintain higher reputation [2]. However, vendors and buyers always tend to prevent traditional ways of conducting illegal trades by getting advantages from the transaction mechanisms [16].

Among those transaction mechanisms, “Escrow Mechanism” is one of the widely used transaction mechanism which relies on the use of third-party escrow system. Escrow mechanism is a transaction arrangement that a trusted third party holds the funds to be transferred to the seller until the buyer receives the order, or the payment will be refunded to the buyer if the transaction is incorrect due to the failures on the vendor’s side [16, 26]. The Escrow transaction mechanism is depicted in Figure 2.2.

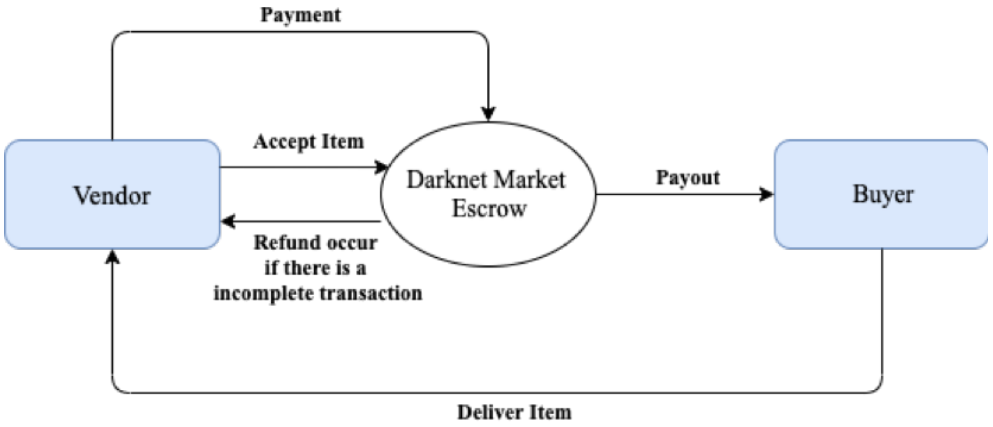


Figure 2.2: Escrow Transaction Mechanism in the Darknet Market

However, there were some major drawbacks that have been introduced by the escrow mechanism since it relies on a third-party trusted system. Because the Darknet market administrators can simply steal the money that held temporary in the third-party escrow system which his/her under control and leave the Darknet market [26]. Therefore, to mitigate this drawback in escrow mechanism, an alternative mechanism has introduced which is called a “Multi-signature” mechanism.

In a multi-signature mechanism, multiple parties (vendors, buyers, and the Darknet market platform) which involved in the trade have to agree to release the funds to the third party escrow system to hold until the shipping is complete [2, 26]. Therefore, two out of three multi-signatures (cryptographic signatures of vendors, buyers and Darknet market administrators) should need to sign to release the funds to the escrow system [26]. However, this mechanism has provided a difficult experience for the buyers and the darknet market itself due to the difficulties of setting up the procedure in the correct manner [26]. As a result, some of the Darknet markets have followed a novel mechanism called the “Finalized Early” mechanism which is a direct and fast transaction processing approach without the use of third party escrow mechanisms [26].

Along with those transaction mechanisms, darknet markets tend to provide the expected efficiency to their users (vendors and buyers) for gaining a better reputation for the Darknet market itself [2]. In addition to the Darknet market transaction mechanism, there is another major existing factor called “Cryptocurrency” which is

a part of the payment system in the Darknet market and the main factor to secure the anonymity (secure the real identity from the outside observers) of the Darknet market users. The Cryptocurrency involvement in the Darknet market transaction will be discussed in section 2.4.

2.4 Cryptocurrency involvement in the Darknet market transaction and Bitcoin cryptocurrency

The cryptocurrency is a part of the Darknet market transaction mechanism which is a medium of payment between the vendors and the buyers. The major usage of the cryptocurrency is that ensuring the expected anonymity of the Darknet market users [2]. There are several cryptocurrencies that have been used in the Darknet market transactions as the currency of payment such as Bitcoin [3], Monerocoin [4] and Etheruem [5]. As a result of the analysis between those major cryptocurrencies, a few studies have stated that Bitcoin is the most popular and the most used cryptocurrency on the Darknet market in the day-to-day transactions [6, 7]. Therefore, this reason leads this research study to focus on analyzing the Darknet market network using the transaction data based on the Bitcoin cryptocurrency.

Bitcoin is a decentralized digital cryptocurrency that relies on a cryptography algorithm and a peer to peer network [3]. Unlike a traditional banking system, a transaction using Bitcoin cryptocurrency occurs without the help of central authority and under the pseudonym[3] (public key address) of the involved parties in the transactions [3]. The transactions using the Bitcoin cryptocurrency can be categorized into two part as single-input transactions and multi-input transactions. In the single-input transactions, the transactions occurred using the single public key address (pseudonym) as an input and while in multi-input transactions, the transactions occur using a multiple public-key addresses as an input [8]. However, there is a major risk that has occurred in both types of transactions relevant to the user identity since Bitcoin transactions considered pseudonymous. Considering the single-input transactions,there is a main observation that, if there are any public-key addresses (pseudonym) in the particular transaction ever link to his/her real identity then every transaction that belongs to the particular user can be traced [27]. In contrast, according to the multi-input transactions, multiple public key addresses (as inputs to the transactions) can be linked to the single entity called wallets [9]. As a result, if there is a possibility to link the particular user's identity to any of the addresses in the input of the transaction, then none of the addresses in the input of transaction will be remained as anonymous since the wallet belongs to the single user/entity [27].

[3]https://www.buybitcoinworldwide.com/anonymity/

Therefore, this research study has focused on the traceability analysis of Darknet market users and addressing the Darknet market network behavior that involved the Bitcoin transactions by using those pseudonymous observations as aforementioned. However, to obtain this analysis over the Darknet market network over Bitcoin transactions, there is a need for a proper approach of analysis that needs to be conducted using the Bitcoin transaction data. The approaches of analysis by using the Bitcoin transaction data will be discussed in section 2.5.

2.5 Approaches of analysis by using the Bitcoin transaction data

Based on several studies [16, 17, 28–30] there are two main approaches have been identified in the literature regarding the analysis using the Bitcoin transaction data.First, Non-graph based analysis and Second, Graph-based analysis.

In the non-graph based analysis, the analysis has been based on the statistical data on the Bitcoin network such as the number of Bitcoin transactions and the number of long-chain transactions [16, 17]. Therefore this approach has derived conclusions by evaluating the data on the statistical operations [16, 17]. Since the Bitcoin transaction network is always evolving with the time , the statistical analysis will not be a proper approach to be used to analyze for investigate the topological behaviors of the network [17]. Therefore, the graph-based analysis has recognized as a better approach to analysis the network structure to investigate behaviors of the network [17]. Therefore, this approach has been chosen in this research study for analyzing the behaviors of the Darknet market network using the Bitcoin transaction data. In the past few years, there are some research studies that have been conducted regarding the graph-based analysis using the Bitcoin transaction data. Those studies will be discussed in section 2.6.

2.6 Studies regarding the Graph-based analysis with using the Bitcoin transaction data

In recent years, several research studies [17, 28–30] have been conducted according to the graph-based analysis using the Bitcoin transaction data to address many aspects. Those analyses have been conducted through deriving the two graph structures namely, the transaction graphs and the user graphs. The transaction graph is the directed multigraph that can directly model using the Bitcoin network [28]. Further, the vertices of the transaction graph represent the public key addresses in the

transaction and edges of the transaction graph represented a particular transaction from a source address to target address [29]. Consequently, there are several recent analyses that have been introduced [28,30–32] which is relevant to constructing the Bitcoin transaction graphs to find the interesting properties of the Bitcoin economy. Furthermore, there were research studies which are regarding the analysis of the Bitcoin transaction graph by measuring network characteristics (e.g.: degree distribution and degree correlation) [23,28,32]. Moreover, in [33], the Bitcoin transaction graph has extracted to evaluate the privacy provision in the Bitcoin blockchain network to address the security vulnerabilities in real-world transactions.

Aside from the transaction graph-based research studies, there are several studies [28, 30] that have been conducted through evaluating the user graph of Bitcoin users. The user graph is a directed multigraph where the vertices represent the clusters of multiple public-key addresses that belong to the same user/entity and edges represent that the particular transaction between one cluster to another cluster; if there exists a transaction from one address in a particular cluster to another address in another particular cluster [30]. To derive the user graphs, there is a need for a clustering approach to preparing the vertices (collection of public key addresses that belongs to the same user/entity). For this concern, "Meiklejohn" has been introduced two heuristics to construct clusters of Bitcoin public addresses that belong to the same users [28]. Those two heuristics has mentioned below,

- **Heuristic one:** If two (or more) addresses are inputs to the same transaction, they are controlled by the same user; i.e., for any transaction t, all public keys $\in$ inputs(t) are controlled by the same user.

- **Heuristic one:** The one-time change address[4] (the bitcoin addresses that used to send the remaining change money back to the sender in bitcoin transaction) [34] is controlled by the same user as the input addresses; i.e., for any transaction t, the controller of inputs(t) also controls the one-time change address public key $\in$ outputs(t) (if such an address exists).

Using the above heuristics, there are few studies have been conducted by analyzing the topological behaviors by constructing Bitcoin user graphs [28, 30]. Using this graph-based analysis approach, this research study has been focused into data driven analysis of Darknet market network based on the Bitcoin transaction data.

[4]https://support.bitpay.com/hc/en-us/articles/115003063823-What-is-a-bitcoin-change-address-

2.7 Summary

This chapter focused on providing an extensive review of the existing analysis of the Darknet market network and analysis based on Bitcoin transaction data. Initially, it discussed the Darknet market and existing approaches of analysis such as analysis based on product-related listings, analysis based on product photographs and analysis based Darknet markets discussion forums. Next it has done the comparison between those approaches and transaction data-based analysis. Then discussed the cryptocurrency usage in the Darknet market transaction and the importance of using Bitcoin cryptocurrency in the analysis. Finally, the approaches of analyzing using the Bitcoin transaction data were discussed and compared the effectiveness of the graph-based analysis using several research studies.

Chapter 3

Design

3.1 Dataset

There are two datasets have used in this research study. First, dataset used in this research study was the publicly available resource named “Blockchain.com” [18] which consists the real-world Bitcoin information of the Bitcoin users that were available in the surface web such as tagged Bitcoin addresses [35] (the Bitcoin address that label with the short name and external link),the URLs for the Bitcoin Talk profiles [36],etc. In [18], the information was available in four categories. Namely, submitted links (all tagged Bitcoin addresses that are submitted to the [18]) , signed messages (all tagged Bitcoin addresses that are submitted to the [18] by signing a particular message using the private key of the submitter) , “BitcoinTalk” profile [37] (The list of Bitcoin addresses that belongs to “BitcoinTalk” forum users) and “Bitcoin - OTC” profiles [36] (The list of Bitcoin addresses that belongs to the Bitcoin Over-The-Counter (OTC) marketplace). The Table 3.1 depicted the number of Bitcoin addresses consists in the dataset of “Blockchain.com” that collected in four categories.

Table 3.1: The Number of Bitcoin addresses that crawled according to several categories from Blockchain.com

Category	Number of Bitcoin Addresses crawled
Submitted Links	3,650
Signed Messages	26,700
Bitcoin Talk Profiles	2,550
Bitcoin OTC Profiles	4,700

Second dataset used in this research study was the publicly available resource named “walletexplorer.com” [19] which contains the Bitcoin transactions of the Darknet markets. The Bitcoin transactions from the seven Darknet markets have been col-

lected and information such as transaction ids, timestamp of the transactions, transaction fees, input addresses of the transactions and output addresses of the transactions, the relevant tags[1] for Bitcoin addresses(Market Tags and User Tags) have included in the dataset. The Table 3.2 depicted the number of Bitcoin transactions have collected in each Darknet market from the “walletexplorer.com”.

Table 3.2: Number of Bitcoin Transactions crawled from the WalletExplorer.com

Darknet Market	Number of Transaction crawled
Abaraxas Market	119,065
Bluesky Market	55,106
Cannabis Market	2,829
Middle Earth Market	34,149
Nuclues Market	146,381
Pandora Market	55,757
Sheep Market	53,639

3.2 Research Design

The research design encompasses four main steps: Web crawling process, Data extracting process, Graph modeling process, and analysis. 3.1 depicted the detailed diagram of the proposed research design.

[1]Tags : Short name for the given Bitcoin addresses. In [19], there is a two types of tags such as Market tags i.e:"SheepMarketPlace","CannabisMarketPlace" and User Tags i.e: "046576b9af","00000014ea". These tags can be use to identify the given Bitcoin address either User/service or market itself.

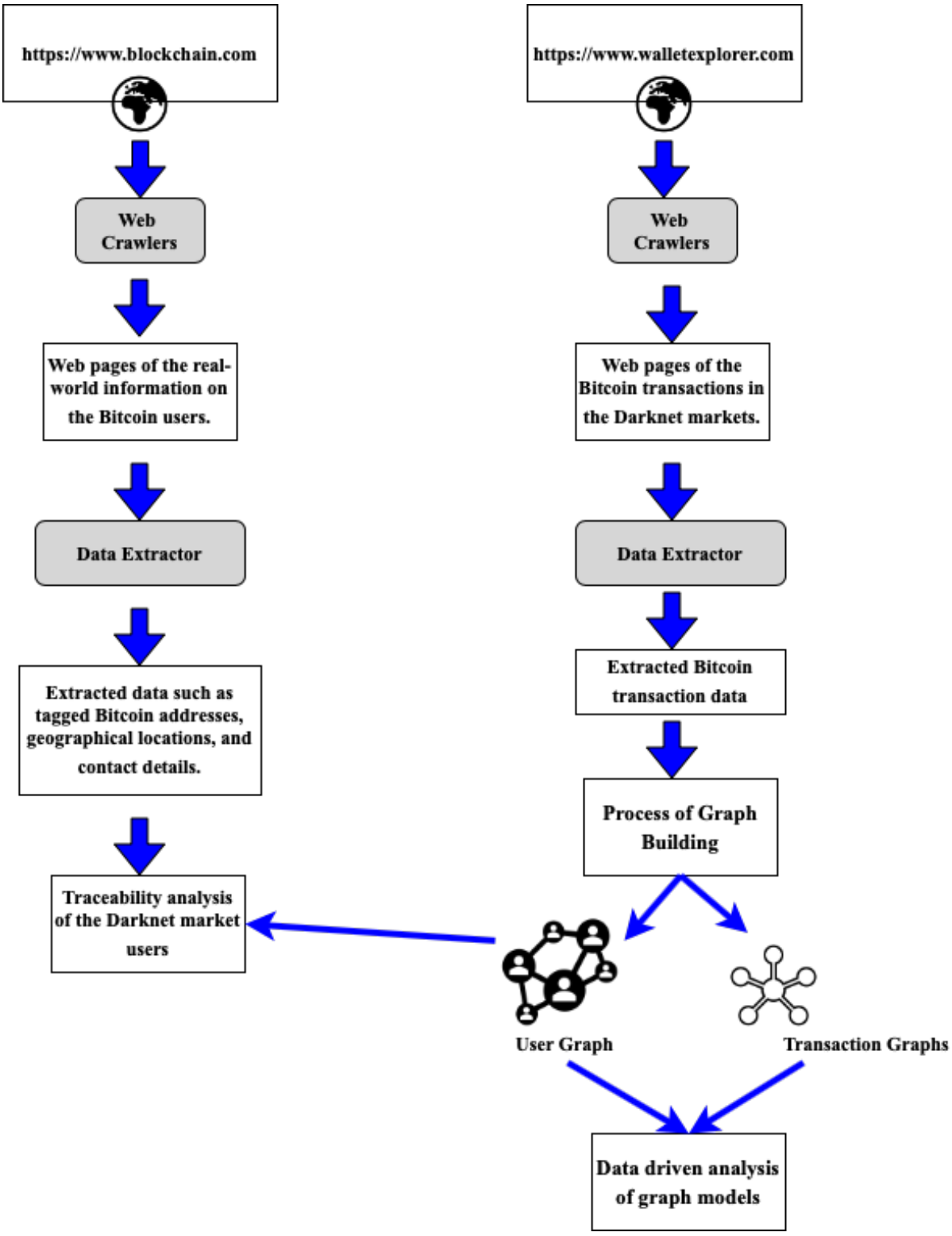


Figure 3.1: Overall Research Design

Initially, the web pages have been crawled from sources blockchain.com [18] and walletexplorer.com [19] by using the implemented web crawlers. As mentioned in Section 3.1, the crawled web pages have contained the real-world Bitcoin information of the Bitcoin users and Bitcoin transactions of the seven Darknet markets. In

the second step, the crawled web pages have been used to extract the Bitcoin data in the data extracting process. In the data extracting process, the real-world Bitcoin information of the Bitcoin users such as Bitcoin addresses along with the tags, URLs for Bitcoin Talk forum profiles and URLs for Bitcoin OTC profiles has been extracted from the web pages of [18]. Similarly, the Bitcoin transaction data has been extracted from the web pages of [20]. In the third step, there were two types of graphs have been created as Transaction graphs and User graphs using the extracted Bitcoin transaction data from [19]. The graph modeling process will be discussed in Section 3.2.1. In the final step, the analysis has been done using the extracted real-world Bitcoin information of the Bitcoin users from the data extracting process and the created graphs models from the graph building process. The process of analysis has been discussed in section 3.2.2.

3.2.1 Graph modeling process

In the graph modeling process, there are two types of graphs have been created. First, the **Transaction graphs** (directed multigraph that can directly model using the Bitcoin network) and second, **User graphs** (directed multigraph where the vertices represent the clusters of multiple public-key addresses that belong to the same user/entity and edges represent that the particular transaction between one cluster to another cluster) with using the extracted Bitcoin transactions from the [19]. In the Transaction graphs modeling, the edges have been represented as the particular transaction between the input and output Bitcoin addresses and the vertices have been represented the input and output Bitcoin addresses in the particular transaction. Considering the Transaction graph modeling, there are two scenarios that have been taken into consideration.

- **Single-input transactions:** The transactions that occur using the single input Bitcoin address. The creation of vertices and edges in the single-input transaction has shown in Figure 3.2.
  - Consider the example of Bitcoin transaction as follows. Note that the transaction's format in the example is
    {<timestamp of the transaction >, <input addresses of transaction >,<(output addresses of transactions, amount)>}

∗ **Transaction 1:** {t1; a1 ; (a2,10),(a3,30)}

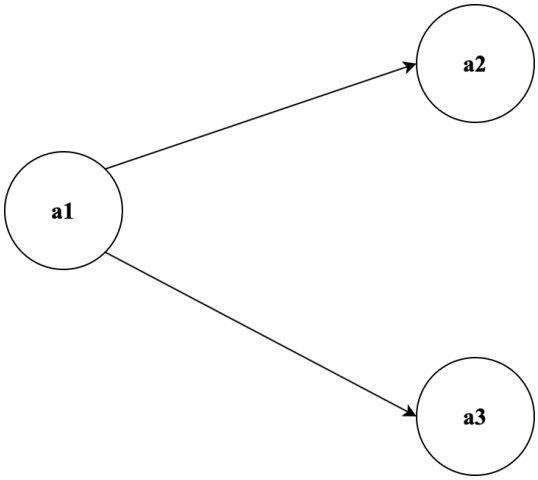


Figure 3.2: Example transaction graph for the single-input transaction

- **Multi-input transactions:** The transactions that occur using the multiple input Bitcoin addresses. The creation of vertices and edges in the multi-input transaction has shown in 3.3.

  – Consider the example of Bitcoin transaction as follows. Note that the transaction's format in the example is
  {<timestamp of the transaction >, <input addresses of transaction >,<(output addresses of transactions, amount)>}

    ∗ **Transaction 1:** {t1; a1 ; (a2,10),(a3,30)}

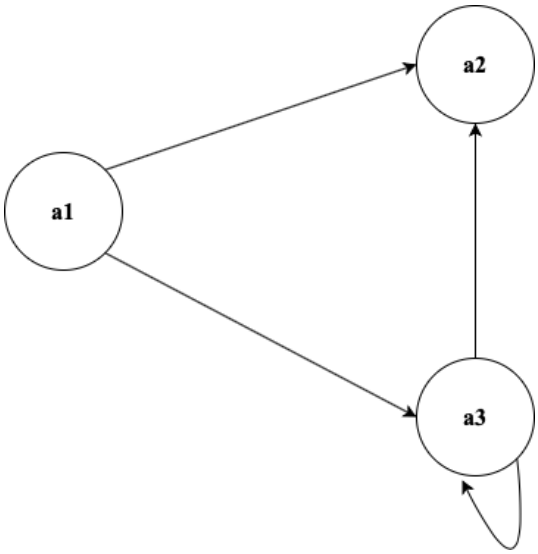


Figure 3.3: Example Transaction graph for the multi-input transactions

Besides the transaction graph modeling, the second graph model that has created in the graph modeling process is User graphs. In order to create user graphs, there is a clustering approach has been conducted as a preprocessing step. Therefore, the Bitcoin addresses in the transactions have used to obtain the clusters of Bitcoin addresses in order to derive the vertices of the user graphs. The process of clustering has been done using the BlockSci Kalodner2017 tool which is the well-known Bitcoin analysis tool available with many features such as address clustering and address tagging (the process of labeling the Bitcoin addresses with the proper names and external links). The clustering algorithm in the BlockSci tool has relied on the two heuristics that have discussed in section 2.6. According to those well-known heuristics, *a single cluster can be considered as a single user since the cryptographic nature of the Bitcoin cryptocurrency* [28]. Therefore, the particular cluster of Bitcoin addresses has been considered as the single vertex in the user graph and the single edge of the user graph will be represented as the transactions between two clusters if there exists a transaction between the Bitcoin addresses inside those particular clusters. The example of deriving the user graph from the Bitcoin transaction will be depicted in Figure 3.4.

- Consider the example of three Bitcoin transactions as follows. Note that the transaction's format in the example is

  {<timestamp of the transaction >, <input addresses of transaction >,<(output addresses of transactions, amount)>}

  * **Transaction 1:** {t1; a3,a5 ; (a3,10)}
  * **Transaction 2:** {t2; a5 ; (a2,20),(a1,10)}
  * **Transaction 3:** {t2; a2 ; (a1,40),(a4,10)}

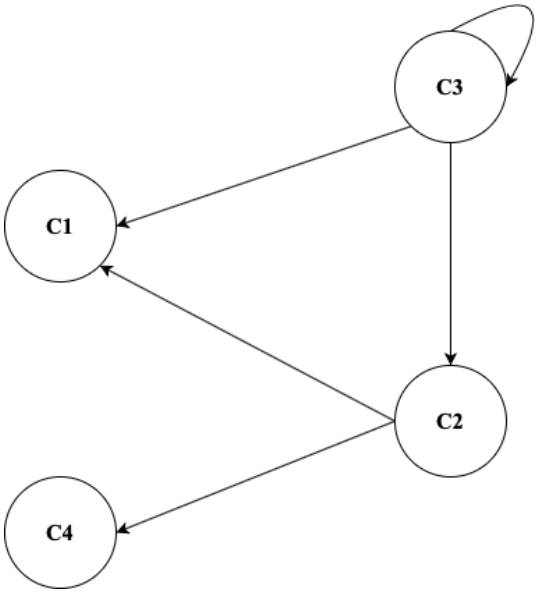


Figure 3.4: Example user graph for the given Transaction data

– Note that cluster C1 contains the address a1, Cluster C2 contains the address a2 , Cluster C3 contains the address a3 and a5, and Cluster C4 contains the address a4.

Section 3.2.2 will be discussed the work of analysis doing by exploiting those constructed graph models.

3.2.2 The Analysis Phase

The analyzing has been done in three phases. First, analyzing the overall Darknet markets. Second, the analyzing of transaction graphs and user graphs of each darknet markets by addressing the graph properties and third, the traceability analysis of Darknet market users by using the information of the Bitcoin users in the surface web. Section 3.2.2.1 will be discussed in the first phase of the analysis, section 3.2.2.2 will be discussed in the second phase of analysis and 3.2.2.3 will be discussed in the second phase of analysis.

3.2.2.1 Analysis Phase I: Overall market analysis

In this analysis phase, the overall analysis of all Darknet markets will be addressed. The overall analysis includes several factors to be addressed. First factor that consider is the transaction flow of the overall Darknet markets. The transaction flow of the overall Darknet markets will be analyzed to find the hyper operation phases and their significant properties in the overall transaction flow of the Darknet markets. Second factor that consider is the money flow of the overall Darknet markets. Apart from the overall transaction flow, the analysis of the money flow in the overall Darknet market will address the incoming and outgoing payments in the overall Darknet markets and the comparison between them (incoming payments and outgoing payments). The results and the evaluation of the analysis phase I will be discussed in Section 5.1

3.2.2.2 Analysis Phase II: Analyze the graph models by addressing graph properties

There are several graph properties that have been addressed by applying to both transaction graphs and user graphs in each Darknet market.Those graph properties have been addressed in two different categories. Namely,Graph connectivity analysis and centrality analysis. Densification measures, Degree Distribution measures and Clustering Coefficient measures will be addressed under the connectivity analysis and centrality measures will be addressed under centrality analysis. The graph properties that consider for the analysis phase II have described below.

- **Densification :** The densification property has been used to observe the evaluation of graphs and to obtain whether the Darknet market network will be dense over time [38]. In this research study, there are three measurements have been obtained according to the densification analysis of Transaction graphs and User graphs. The First measurment is the observing evaluation of the number of nodes and edges over time. Second measurment is observing the evaluation of the average in-degree and out-degree over time (sum of the in or out degrees of the nodes divided by the number of nodes in the particular timestamp) [39]. Third is observing the evaluation of the percentage of nodes in the giant strongly connected component over time [39].

- **Degree Distribution :**The degree distribution property has been used to observe the connectivity of Bitcoin addresses and clusters of Bitcoin addresses over time [40]. In this research study, there are three types of degree distributions have been measured from the transaction graphs and user graphs. First, In-degree distribution. Second, Out-degree distribution and Third, the Degree distribution of the undirected graph that converted from the multi directed graph.

- **Centrality Measures :**The Centrality measure has been used to obtain the most central and most active vertices in the Darknet market network [41]. In this study, there are four types of centrality measures have been obtained from the Transaction graphs and User graphs.

  - **Normalized Degree Centrality** which is a fraction of nodes that particular node connected to [41]. Normalized Degree centrality ($C_D(v)$) of a node v in a simple graph that has N number of nodes can be retrieved from the equation 3.1:

$$C_D(v) = \frac{degree(v)}{(N-1)} \tag{3.1}$$

  - **Normalized in-degree centrality** which is a fraction of nodes that particular node's incoming edges are connected to [39]. Normalized in-degree centrality ($C_I(v)$) of a node v in a simple graph that has N number of nodes can be retrieved from the equation 3.2:

$$C_I(v) = \frac{In\_degree(v)}{(N-1)} \tag{3.2}$$

- **Normalized out-degree centrality** which is a fraction of nodes that particular node's outgoing edges are connected to [39]. Normalized Out-degree centrality ($C_O(v)$) of a node v in a simple graph that has N number of nodes can be obtained from the equation 3.3:

$$C_O(v) = \frac{Out\_degree(v)}{(N-1)} \tag{3.3}$$

- **PageRank centrality** is a centrality measure that normalizing the number and quality of links to a particular node and evaluate the importance of that particular node in the network [42]. The PageRank of node v (PR(v)) can be obtained as follows:

  * Assume that nodes T1 ... Tn have links to the node v. Then the PageRank of node v can be obtained from the equation 3.4:

$$PR(v) = (1-d) + d(\frac{PR(T_1)}{C(T_1)} + \cdot + \frac{PR(T_n)}{C(T_n)}) \tag{3.4}$$

  * Note that the d is the damping factor which can be set between 0 and 1(usually 0.85) and C(T1) can be defined as the number of outgoing links from the node T1.

- **Clustering Coefficient:** The Clustering Coefficient has been used to observe the degree of nodes in the Darknet market which tends to cluster together []. In this research study, the clustering coefficient of the nodes has been measured from the simple undirected graph which was converted from the original multi directed graph. The clustering coefficient (C(u)) of the node u can be obtained from the equation 3.5:

$$C(u) = \frac{2T(u)}{degree(u)(degree(u)-1)} \tag{3.5}$$

  - Note that T(u) is the number of triplets (three nodes that are connected by the two edges or three edges) through the node u and deg(u) is the degree of u.

The details of the results and conclusions have made from the analysis of transaction graphs and user graphs by using the aforementioned graph properties will be discussed in Chapter 5.

3.2.2.3 Analysis Phase III: Traceability Analysis between Darknet market users and Surface web Bitcoin users

The traceability analysis has done in three steps. In the First step, the Bitcoin addresses of the surface web users that have collected from the Blockchain.com [18] have used. In this step, the clusters of Bitcoin addresses have obtained from those surface web Bitcoin addresses. In order to obtain those clusters, the BlockSci [20] tool has been used which is discussed in section 3.2.1. In the second step of the traceability analysis has been done by comparing the vertices(known to be clusters of Bitcoin addresses) of the user graphs in each Darknet market and the clusters obtained from the surface web Bitcoin addresses in the first step. Therefore, the investigation has been done to check whether there is one or more Bitcoin addresses of the Darknet markets have been clustered inside the clusters of surface web users(obtained from the first step). In another word, there should need to have the same clusters in the list of vertices of the user graphs and list of clusters in the surface web users. Since those trace clusters can be recognized as the users(or single entity) inside the Darknet market networks, those behaviors of traced clusters can be considered as the behaviors of the users in the Darknet market network (refer section 3.2.1). Therefore, in the final step, the investigation and analysis have been done for obtaining the behaviors and information about those traced users inside the Darknet markets.

3.3 Summary

This chapter has provided a detailed description of the research design. In section 3.1, the detailed description of the Dataset has been provided. After that chapter comprised a detailed description of the research design in section 3.2. In that section research design has presented in four phases namely web crawling process, Data extracting process, Graph modeling process which was based on building transaction graphs and user graphs, and analysis phase which is based on addressing graph properties and traceability analysis of the Darknet market users.

Chapter 4

Implementation

This chapter consists of the overall implementation details of the analysis process. Beginning with the introduction to the software tools used in this implementation(refer section 4.1) this chapter shows the overall implementation details in five sections. In section 4.2.1 discussed the implementation details of the web crawling and data extracting process. In section 4.2.2 discussed the implementation details of the graph model building process. Then section 4.2.3, section 4.2.4 and section 4.2.5 will show the overall description of implementation details of the analysis phase I, analysis phase II and analysis phase II respectively.

4.1 Software Tools

All the implementation has been done by using the Python 3.5 version. Python [43] has provided numerous libraries which convenient in the implementation of the web crawling process, data extracting process and graph analysis process. Due to that reason, Python has been chosen as a programming language to implement the intended solution.

In the web crawling process, the *Urllib2* [44] was the main library used in the implementation. *Urllib2* library has provided the solution to fetching the URLs and open the web pages that intended to crawled. In the data extracting process, the *Lxml* [45] was the main library used in the implementation which has provided a convenient way for handling the HTML files and extracting the data from the crawled web pages. In the graph building and Darknet market network analysis process, there were two main libraries have been used namely, *NetworkX* [46] and *BlockSci* [20]. The NetworkX library used in several scenarios. First, it has used for the building process of transaction graphs and user graphs using Bitcoin transaction data and then *NetworkX* has used for the graph analysis process by addressing the several graph properties which were provided by the *NetworkX* library itself such as degree distri-

bution, centrality analysis, and densification analysis. Specifically, the *NetworkX* has been provided as an efficient graph analysis library for the complex network with a high number of vertices and edges. Apart from the *NetworkX* library, the second main library which was used in the graph building process is the *BlockSci*. The *BlockSci* is the open-source software platform for Blockchain analysis and it has consisted of several functionalities such as address tagging, address clustering, and Blockchain transaction analysis. The implementation of the vertices in the user graphs of the Darknet market network has been done by using the address clustering feature of the *BlockSci* library. According to this user graph building scenario, Addressing clustering functionality has provided the clusters of addresses for the given Bitcoin addresses of Darknet market transaction and those clusters have used as the vertices for the user graphs. Besides the process of building user graphs, *BlockSci* library has used for the traceability analysis of the Darknet market users as well. In the process of traceability analysis, the *BlockSci* module has been used to get the clusters of addresses of the given Bitcoin addresses which were available on the surface web (refer section 3.2.2.3).

Apart from the main implementations, there was another software tool has been used for graph visualization called *Gephi* [47] which were recognized as a high-performance open-source graph visualization tool for complex network.

4.2 Implementation Details

The overall implementation of this research can be categorized into five main categorized regarding the research design that has discussed in chapter 3.

1. Web crawling and data extracting process.
2. Graph Models Building from Darknet market transaction data.
3. Analysis Phase I: Overall market analysis.
4. Analysis Phase II: Analyze the graph models by addressing graph properties.
5. Analysis Phase III: Traceability Analyzing between Dark web market users and Surface web Bitcoin users.

4.2.1 Web crawling and web scraping

As mentioned in Section 3.1, the web crawling and the data extracting process has been done from two different sources namely, Blockchain.info [18] and Walletexplorer.com [19]. Therefore separate implementations have been done to crawled and extracted the data from each data resource. In the implementation of web crawlers, all the HTML files are fetched according to the given URL using *FETCH()* function. As mentioned in Algorithm 1, the URL will be updated sequentially until exceeding the *limit* URL parameter and each HTML file will be saved in the local file system in each iteration using the *SAVE()* function.

Algorithm 1: Algorithm for Web Crawlerling Process

```
WebCRawler (URL, limit)
    inputs : The Website URL; limit value that indicate the limit of the web
             pages that need to crawled
    output: None
    for i=0 to offset do
        Page = FETCH(URL+i);
        SAVE(Page);
    end
```

Apart from the web crawling process, the data extracting process has been used to extract the relevant data such as Darknet market transaction data (from [19]) and Bitcoin user data (from [18]).Then after the data extracting process, it will return the JSON files which contains the extracted data. As mentioned in Algorithm 2, data will be extracted from the EXTRACT_DATA() function using the crawled HTML files. Then each extracted records will be saved as a JSON file to a local file system. Specifically, there are separate functionalities that have been implemented for adding data to the JSON files according to both sources. For an instance, if the records from source [19] then the data that added to the JSON file are “Input address”, “output address”, “input value”, “output value” and “timestamp”.

Algorithm 2: Algorithm for Data Extracting Process

WebScraper *File_Path*

inputs : Directory Path that inclused HTML files

output: JSON Files

$HTML\_Files = GET\_LIST(File_path)$;

foreach $f \in HTML\_Files$ **do**

$Records = EXTRACT\_TRANSACTIONS(f)$;

foreach $r \in Records$ **do**

$JSON = ADD(r["..."])$;

return $JSON$;

4.2.2 Graph Building from Darknet market transaction data

Two types of graphs have been implemented based on Darknet markets transaction data. First graph model is the Transaction graph.The transaction graphs have been implemented by using the *NetworkX* [46] library. The edges and the vertices of the transaction graph have been implemented by considering the scenarios that have mentioned in section 3.2.1. However, there was a separate Python script has been implemented to prepare the vertices and edges for the implementation of the transaction graph by using the extracted data from the data extracting process. This python script has created separate *JSON* files for both edges and vertices for each Darknet market. The implementation of this Python script has mentioned in Code 4.1 and Code 4.2. The code 4.1 shows the implementation of obtaining the vertices in the transaction graph. The vertices will have some seperate attributes. Therefore vertices will be in the combination of either input transaction address and its label or output transaction address and its label. Code 4.2 shows the implementation of the edges in the transaction graph and the edges have included the data of transactions such as input address, output address, input value, the output value, and timestamp.Consider that input value,output value and timestamp will be setting as edge attributes.

```
import json

vertices = []
def create_vertices(path,market_name):
    chdir(path)
    with open('extracted_tx_data'+market_name+.json') as
    json_file:
        data = json.load(json_file)
        for key,value in data.items():
              input_label = value['tx_sender']
              tx_inputs = value["tx_inputs"]
              tx_outputs = value["tx_outputs"]

              for input_addr in tx_inputs:
                  if input_addr['input_addr'] not in addresses:
                      label_in = {'label':input_label} vertices.
    append(list([input_addr['input_addr'],label_in]))

              for output_addr in tx_outputs:

                  if output_addr['output_addr'] not in addresses
    :

                      if output_addr['output_tag'] == '(change
    address)':
                          label_out = {'label':input_label}
                      else:
                          label_out = {'label':output_addr['
    output_tag']}
                      vertices.append(list([output_addr['
    output_addr'],label_out]))

         with open('vertices_'+market_name+.json', 'w') as
    file_handler:
              file_handler.write(json.dumps(vertices))
```

Listing 4.1: Python Code for Implementation of the vertices in one transaction graph

```
import json

edges = []
def create_edges(path,market_name):
    chdir(path)
    with open('extracted_tx_data_of_'+market_name+.json') as
   json_file:
        data = json.load(json_file)
        for key,value in data.items():
            input_label = value['tx_sender']
            tx_fee = value['tx_fee']
            tx_timestamp = str(value["tx_date"])+' '+str(value
   ["tx_time"])
            tx_inputs = value["tx_inputs"]
            tx_outputs = value["tx_outputs"]

            for input_addr in tx_inputs:
                input_val = input_addr["input_val"]
                for output_addr in tx_outputs:
                    output_val = output_addr["output_val"]
                    val = {'input_val':input_val,'output_val':
   output_val}
                    ts = {'timestamp':tx_timestamp}
                    edges.append(list([input_addr['input_addr'
   ],output_addr['output_addr'],ts,val]))
            with open('edges_'+market_name.json', 'w') as
   file_handler:
                file_handler.write(json.dumps(edges))
```

Listing 4.2: Python Code for Implementation of the edges in one transaction graph

After preparing the vertices and edges for the transaction graphs, Code 4.3 shows the implementation procedure of the transaction graph using another Python script. According to Code 4.3, the edges and the vertices will be loaded from the JSON files which were created from the previously mentioned Python script (Code 4.1 and Code 4.2) for implementing edges and vertices for each Darknet market. Accordingly, the transaction graphs have been created using the NetworkX library function called NetworkX.MultiDiGraph().

```
import networkx as nx
import json

market_vertices = json.load(open(".../vertices_example_market.
    json"))
market_edges = json.load(open(".../edges_example_market.json")
    )
//function for creating graphs
def create_multigraph(vertices_list,edges_list,market_name):
    tx_MultiDigraph = nx.MultiDiGraph()
    tx_MultiDigraph.add_nodes_from(vertices_list)
    tx_MultiDigraph.add_edges_from(edges_list)
    print(market_name+ " market Multi Directed graph created")
    return tx_MultiDigraph
```

Listing 4.3: Python Code for Implementation of transaction graph using JSON Files previously created as Vertices and Edges

The second type of graph is the User graph which has implemented using the *NetworkX* [46] and *BlockSci* [20] libraries. As mentioned in section 3.2.1, the vertices of these graphs represent the set of Bitcoin addresses that can cluster together using the *BlockSci* library. However, a separate Python script has been implemented to obtain the JSON files of edges and vertices of each user graph prior to the graph building process. Code 4.4 shows the implementation of vertices in the user graph and Code 4.5 shows the implementation of the edges in the user graph. In code 4.4, the address list of the transaction has been converted to the particular cluster IDs using the *BlockSci* library in two steps. First, the transaction addresses converted to the *BlockSci* compatible string format by using the CHAIN.ADDRESSES_FROM_STRING() function and second, get the particular cluster IDs for a given transaction address by using the CLUSTER_MANAGER.CLUSTER_WITH_ADDRESSES() function.Finally, the vertices will be saved as the JSON format. In code 4.5 shows that the input addresses and output addresses will be converted to the particular cluster IDs the same as mentioned in the code 4.4. Then those transactions with cluster IDs saved as a JSON file for user graph building process.

```
1
2  import json
3  import pandas as pd
4  import numpy as np
5  from tqdm import tqdm_notebook as tqdm
6  from collections import defaultdict
7  import networkx as nx
8  from forex_python.bitcoin import BtcConverter
9  import blocksci
10
11
12 chain = blocksci.Blockchain("/root/blocksci-data/")
13 cm = blocksci.cluster.ClusterManager("/root/blocksci-data/
       clusters-h1/", chain)
14
15 market_vertices = json.load(open(".../vertices_example_market.
       json"))
16 market_edges = json.load(open(".../edges_example_market.json")
       )
17
18 //create the cluster IDs for the given addresses
19 def get_node_list(addr_list,market_name):
20     cluster_id_list = []
21     for addr in addr_list:
22         address = chain.address_from_string(str(addr[0]))
23         cluster = cm.cluster_with_address(address)
24         cluster_id_list.append(cluster.index)
25     with open('vertices_for_usr_graph'+market_name.json', 'w')
       as file_handler:
26                 file_handler.write(json.dumps(cluster_id_list)
       )
```

Listing 4.4: Python Code for Implementation of vertices in one user graph

```
1
2  import json
3  import pandas as pd
4  import numpy as np
5  import networkx as nx
6  import blocksci
7
8  chain = blocksci.Blockchain("/root/blocksci-data/")
9  cm = blocksci.cluster.ClusterManager("/root/blocksci-data/
       clusters-h1/", chain)
10 //create the cluster IDs for the given addresses
11 def get_edges_for_user_graph(edges_list,market_name):
12     edges = []
13     for edge in edges_list:
14
15         input_address = chain.address_from_string(str(edge['
       input_addr']))
16         output_address = chain.address_from_string(str(edge['
       output_addr']))
17         input_cluster = cm.cluster_with_address(input_address)
18         output_cluster = cm.cluster_with_address(
       output_address)
19
20         //additional attributes like timestamp will be added
       as a Edge attributes
21         row ={
22             'input_addr':edge['input_addr'],
23             'input_cluster':input_cluster.index,
24             'input_val':edge['input_val'],
25             'output_addr':edge['output_addr'],
26             'output_cluster':output_cluster.index,
27             'output_val':edge['output_val'],
28             'timestamp':edge['timestamp']
29
30         }
31         edges.append(list([row['input_cluster'],row['
       output_cluster'],{'output_val':row['output_val'],'timestamp
       ':row['timestamp']}]))
32
33     with open('edges_for_usr_graph'+market_name.json', 'w') as
       file_handler:
34                 file_handler.write(json.dumps(edges))
```

Listing 4.5: Python Code for Implementation of edges in one user graph

After preparing the vertices and edges for the transaction graphs, the user graphs has been build using the same procedure as mentioned in the code 4.3.

4.2.3 Analysis Phase I: Overall markets analysis

In the analysis phase I, the implementation has been done by considering two factors. First, the transaction flow in the overall Darknet markets and money flow in the overall Darknet markets. The implementation for analyzing the transaction flow of the overall Darknet markets has done by implementing the python script for calculating the number of transactions against the timestamps. For calculating the transactions against timestamps from the Darknet markets, the data extracted from the web scrapping process in each Darknet market has been used. Code 4.6 shows that the transactions count in the each timestamp will be counted in particular Darknet market. Code 4.7 shows that after calculating those transactions in each Darknet market, then those each transaction counts that calculated in each Darknet market will be joined according to the timestamps.

```
from os import chdir
import json

tx_count_with_date= {}
def counting_tx_in_one_market(path,market_name):
    chdir(path)
    with open('extracted_tx_+market_name+.json') as json_file:
        data = json.load(json_file)
        for key,value in data.items():
            date = value['tx_date']
            if date in s:
                tx_count_with_date[date] += 1
            else:
                tx_count_with_date[date] = 1
        final_txs_count = []
        for key,value in tx_count_with_date.items():
            if key not in final_txs_count:
                txs ={"date": key,"number of txs":value}

                final_txs_count.append(txs)
        with open('final_transaction_count+market_name+.json',
    'w') as file_handler:
            file_handler.write(json.dumps(final_txs_count))
```

Listing 4.6: Python Code for counting the transactions against timestamp in one Market

The implementation for analyzing the money flow of the overall Darknet markets has done by calculating the incoming money and the outgoing money of each Darknet market. This calculation was done by using the constructed transaction graphs (refer section 4.2) of each Darknet market. Firstly, identify the market's nodes from the attributes of vertices by using the address tags (refer section 4.2). Then identify the incoming edges and outgoing edges for particular market's nodes in each Darknet market transaction graphs. Then total incoming payment and total outgoing payment from the markets will be calculated against the timestamp by using the edge attributes called "input_val" and "timestamp". Code 4.8 shows the aforementioned functionality for calculating the incoming money flow of the Darknet market.The implementation for the outgoing money flow same as the Code 4.8.

```
from os import chdir
from os import listdir
from os.path import isfile, join
import json

def counting_tx_in_one_market(path,market_name):
    chdir(path)
    json_files = [f for f in listdir(path) if isfile(join(path
   , f)) and f.endswith('.json')]
    records = []
    tx_count_with_date= {}
    for file_name in json_files:
        with open(file_name) as json_file:
            data = json.load(json_file)
            for key,value in data.items():
                date = key
                if date in s:
                    tx_count_with_date[date] += value["txs"]
                else:
                    tx_count_with_date[date] = value["txs"]
    all_txs_count = []
    for key,value in tx_count_with_date.items():
        if key not in all_txs_count:
            txs={"date":key,"txs":value}
            all_txs_count.append(txs)

    with open('final_all_transaction_count.json', 'w') as
   file_handler:
        file_handler.write(json.dumps(all_txs_count))
```

Listing 4.7: Python Code for counting the transactions against timestamp in all Market

```
1
2  def get_incoming_edges(G,tag):
3
4      receieve_tx_list = []
5      for u,v,data in G.edges(data=True):
6          if G.nodes[v]['label'] == tag:
7              receieve_tx_list.append(list([u,v,{'label':G.nodes
       [v]['label'],'output_val':data['output_val'],'timestamp':
       data['timestamp']}]))
8      return receieve_list
9
10 def get_incoming_money_flow(receive_list):
11     recieve_flow ={}
12     final_recieve_flow = []
13     for tx in receive_list:
14         date = tx[2]['timestamp']
15         if date in recieve_flow:
16             recieve_flow[date] += float(tx[2]['output_val'])
17         if date not in recieve_flow:
18             recieve_flow[date] = float(tx[2]['output_val'])
19
20     for key,value in recieve_flow.items():
21         if key not in final_recieve_flow:
22             tx={
23                 "date":key,
24                 "money":value
25             }
26
27             final_recieve_flow.append(tx)
28     return final_recieve_flow
```

Listing 4.8: Python Code for calculate the incoming money values against timestamp in one Darknet market

4.2.4 Analysis Phase II: Analyze the graph models by addressing graph properties

In the analysis Phase II, all the analysis has been done by using the transaction graphs and user graphs of each Darknet markets in order to address graph properties. The graph properties include the centrality analysis like normalized in-degree centrality, normalized out-degree centrality, and Page-rank centrality and connectivity analysis properties like densification, degree distribution, and clustering coefficients. Specifically, the connectivity analysis has been done by using the equal range timestamps

according to each Darknet market transaction's duration. For instance, if the transaction duration for the "Cannabis" Darknet market is from 2014/04/01 06:36:53 to 2014/08/25 11:01:28, then this duration has been divided into equal twenty timestamp instances. Code 4.9 will describe the implementation of this division of timestamps. The various connectivity analysis functions have been calculated according to this each timestamp in each darknet market. The functions such as calculating the number of edges and vertices over time, average in-degree over time and percentage of nodes in the maximum strongly connected component .

```
from datetime import datetime

#get the date ranges and divided into equal space timestamp
def get_equally_timestamp(start, end, intv):
    start = datetime.strptime(start,"%Y-%m-%d %H:%M:%S")
    end = datetime.strptime(end,"%Y-%m-%d %H:%M:%S")
    diff = (end  - start ) / intv

    for i in range(intv):
        yield (start + diff * i).strftime("%Y-%m-%d %H:%M:%S")
    yield end.strftime("%Y-%m-%d %H:%M:%S")
```

Listing 4.9: Python Code for dividing the equal timestamp instances for particular transaction flow

Apart from the connectivity analysis, the centrality analysis has addressed using the transaction and user graphs in each Darknet market. The centrality measures like normalized in-degree centrality, normalized out-degree centrality, and Page-rank centrality have been addresses in each darknet market's transaction graphs and user graphs by using the NetworkX functions such as NETWORKX.PAGERANK(), NETWORKX.GRAPH.OUT_DEGREE() and NETWORKX.GRAPH.IN_DEGREE().

4.2.5 Analysis Phase III: Traceability Analyzing between Dark web market users and Surface web Bitcoin users

In the traceability analysis, the implementation has been done for tracing the Darknet market users from the Surface web Bitcoin users. This analysis has been done using the vertices list of the user graphs and the newly clustered list of Bitcoin users from the surface web. The implementation process for clustering the surface web Bitcoin users has been done by using the same clustering process which has done for creating vertices in user graphs of Darknet markets and it's done by using the extracted data from source [18]. Code 4.10 shows the implementation of traceability analysis between Darknet market users and surface web Bitcoin users. In Code 4.10, the

clusters of Darknet market users and clusters of surface web Bitcoin users have been compared using the CLUSTER_COMPARE() function and return all trace cluster IDs of the surface web Bitcoin users. Then using those tracing clusters, all the information has been obtained relevant to the particular cluster IDs such as Bitcoin tags and URLs for the Bitcoin talk and Bitcoin OTC profiles which mentioned in section 3.1.

```
1
2  def cluster_compare(user_graph_cluster_list,
       surface_web_cluster_list):
3
4       for m_name,market_data in user_graph_cluster_list.items():
5           for b_name,surfaceWeb_info_data in
       surface_web_cluster_list.items():
6               matching_clusters =[]
7               m =[]
8               for market_cluster,value in market_data:
9                   if any(market_cluster in s for s in
       surfaceWeb_info_data):
10                      //check_in_list : comparing function for
       given clusters with labels
11                      if(check_in_list(matching_clusters,
       market_cluster,value['label'])):
12
13                      matching_clusters.append({'cluster':
       market_cluster,'label':value['label']})
14                          m.append(market_cluster)
15              c = unique(m)
16              print(c)
```

Listing 4.10: Python Code for comparing the clusters of User graphs and Cluster of surface web Bitcoin users

4.3 Summary

In this chapter, the overall implementation of this research study has addressed. The overall implementation has described in two sections. The first section discussed the software tools used in the implementation process and in the second section, the overall implementation process has described in five sub-sections. In section 4.2.1 has discussed the implementation process of data collection. In the section 4.2.2 has discussed the implementation process of graph modeling. In section 4.2.3, section 4.2,4 and 4.2.5 has discussed the implementation process done for the each analysis phases.

Chapter 5

Results and Evaluation

This chapter will be given the results and evaluation of this research study. The results and evaluation will be presented in three analysis phase. Section 5.1 will be discussed the results and evaluation of the analysis phase I which regarding the overall market analysis. Section 5.2 will be discussed the results and evaluation of the analysis phase II and the analysis of the transaction graph models and user graph models using different graph properties will be addressed in that section. In Section 5.3, the traceability analysis of the Darknet market will be addressed. Additionally, the behavior of the traced nodes in the user graph models will be addressed under the traceability analysis work.

5.1 Analysis Phase I: Overall market analysis

In the analysis Phase I, the overall analysis has done considering the transactions of all seven Darknet markets. As mentioned in section 3.2.2, this analysis was done by addressing two factors. First, the overall transaction flow of all Darknet markets and second, overall money flow of all Darknet markets. Accordingly, section 5.1.1 will be present the results and evaluation of the analysis in the overall transaction flow of all Darknet markets and section 5.1.2 will be present the results and evaluation of the analysis in the overall money flow of all Darknet markets.

5.1.1 Results and evaluation of the analysis in the overall transaction flow

```
Hyperoperation phase:
First Hyperoperation phase:
        second Peak = 2612 txes
        At = 2013-11-19
        Starts = 2013-10-06
        Ends = 2013-12-01
Second Hyperoperation phase:
        second Peak = 4357 txes
        At = 2015-03-03
        Starts = 2015-01-01
        Ends = 2015-03-19
Third Hyperoperation phase:
        Peak = 5431 txes
        At = 2015-10-28
        Starts = 2015-08-23
        Ends = 2015-11-06

<matplotlib.lines.Line2D at 0xa0a90f290>
```

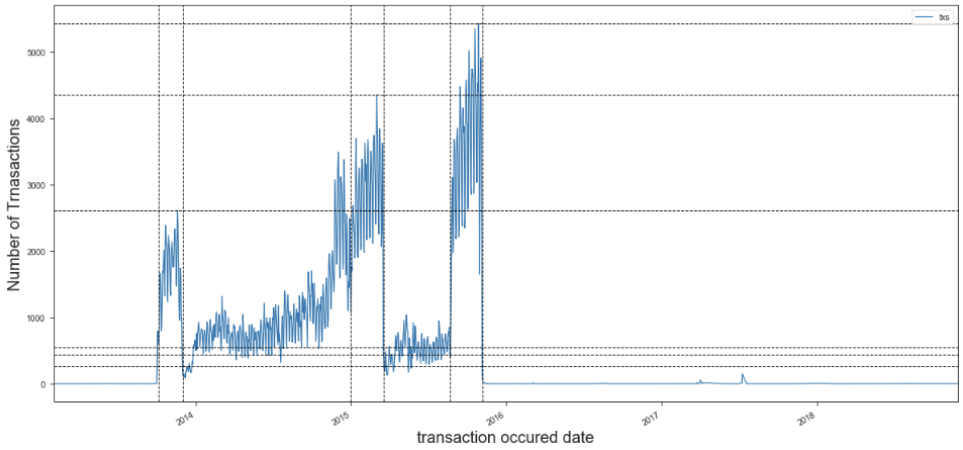


Figure 5.1: Overall Transaction Flow of all seven Darknet Markets

Figure 5.1 shows that most of the transactions have occurred during the period of October 2013 to November 2015. This indicates that all majority transactions are occurring in the 2 years duration.However, the timestamp for all duration was nearly Th 5 years(2013 to 2018).Therefore, this behavior of all these transactions will show that some of the markets have remained inactive after the particular timestamp.The inactiveness of the market will occur due to many reasons such as due to the exit scams(confidence trick where an established business stops shipping orders while continue to receive payments for new orders), due to the raids, due to the market shutdowns by the law enforcement agencies and due to the voluntarily shutdowns [16].

Apart from the inactiveness of the transactions, Figure 5.1 shows that another observation regarding the hyper operation phases in the transactions of all Darknet markets. According to the results shown in Figure 5.1, There are three main hyper operation periods that can be identified. The first hyper operation period can be identified during the period of 16^{th} of October 2013 to 01^{st} of December 2013 and it consists of a peak of 2612 transactions. The second hyper operation period can

be identified during the period of 01st of January 2015 to 19th of March 2015 and it consists of the second most peak with 4357 transactions. The last hyper operation period occurred in the duration of 23rd of August 2015 to 06th of November 2015 and is consists of the highest peak with the 5431 transactions. Therefore, the results of these hyper operational phases indicate that the most active transaction period is the third hyper operational period among all of the seven Darknet markets.

5.1.2 Results and evaluation of the analysis in the overall Money flow

As mentioned in section 3.2.2 The overall money flow of all the seven markets can be categorized into two factors. First, the incoming money flow of the markets and second, the outgoing money flow of the markets. Figure 5.2 shows that the incoming money flow of all seven Darknet markets against the timestamps and Figure 5.3 shows that the outgoing money flow of all seven Darknet markets against the timestamps.

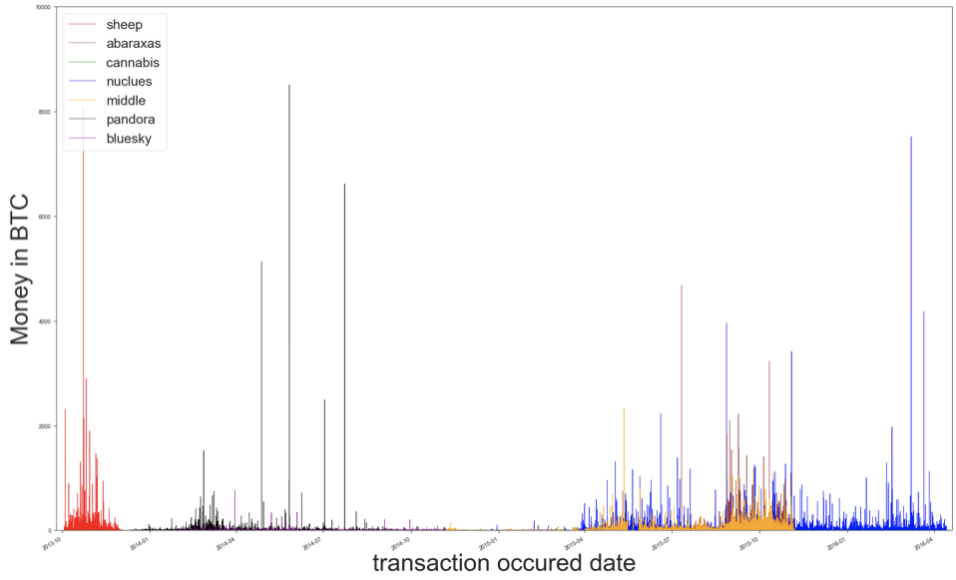


Figure 5.2: Overall Incoming Money Flow of all seven Darknet Markets

According to Figure 5.2, it shows that most of the Darknet markets are tends to have a low-income status during their transactions period. However, there are few spikes(highest incomings) that can be identified mainly in *Pandora*, *Middle Earth*, *Sheep* and *Nucleus* Darknet markets. The Darknet market like *Pandora* and *Bluesky* keeps a low level of incoming rate and it indicates that The Darknet market users are not usually having a good interaction with those Darknet markets when their tradings. The incoming flow of the Nucleus Darknet market shows that the tradings

of the Nucleus market remain a more static manner during its transaction period and it can be categorized as the most wealthy market among seven Darknet markets.

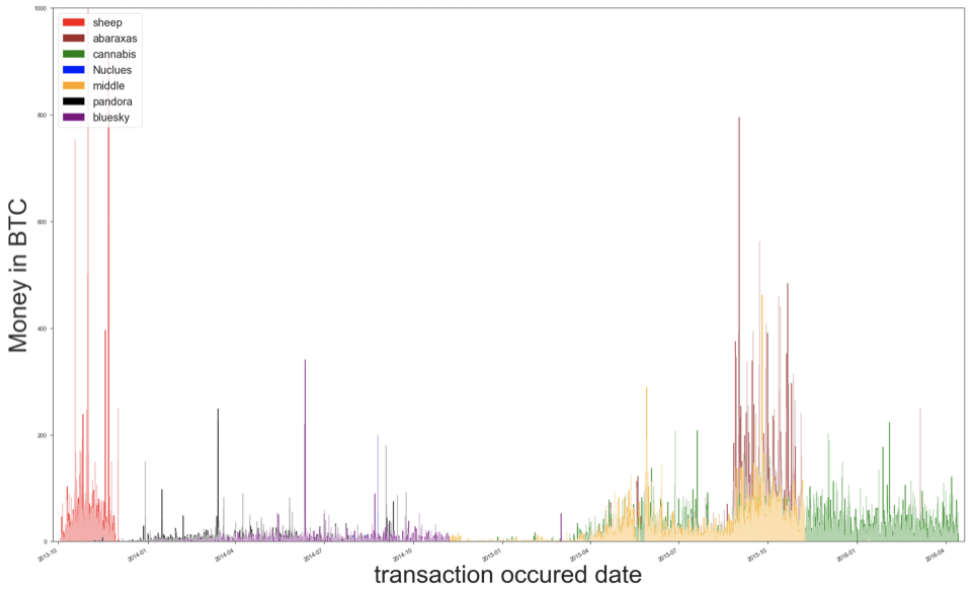


Figure 5.3: Overall Outgoing Money Flow of all seven Darknet Markets

According to Figure 5.3, it shows that outgoing money flow has a similar behavior as the incoming money flow of the Darknet markets. Additionally, it shows that the Darknet markets that have the most incoming money flow have the most outgoing money flow as well. It indicates that Darknet users tend to do their tradings with some popular Darknet markets like *Nucleus*, *Middle Earth* market rather than small unpopular Darknet markets.

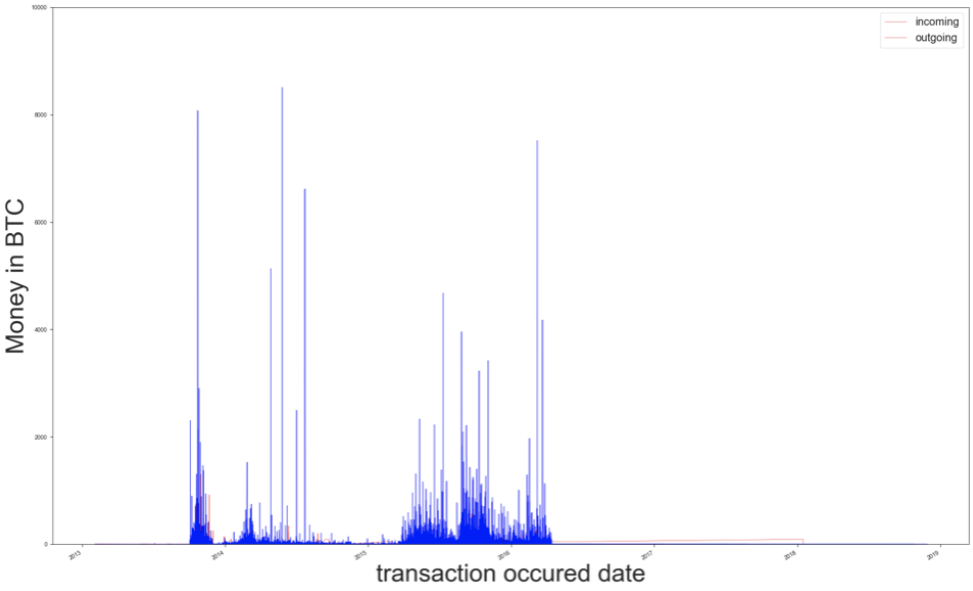


Figure 5.4: Overall comparison between Incoming and Outgoing money flow of all seven Darknet Markets

When comparing the outgoing and incoming money rate of the overall money flow, Figure 5.4 shows that the overall income rate is higher than the overall outcome rate. Therefore we can identify that all seven Darknet markets have success tradings when considering the overall money flow among all seven Darknet markets.

5.2 Analysis Phase II: Analyze the graph models by addressing graph properties

In this phase, the analysis has been done by addressing several graph properties for constructed two graph models namely, Transaction graph models and User graph models. As mentioned in section 3.2.2.2, graph properties have been addressed in three different categories for each constructed graph model(transaction graph model and user graph model). The three different sections are,

- Graph connectivity analysis
  - Densification analysis over time
  - Degree distribution analysis
  - Clustering coefficient analysis
- Centrality analysis
  - Normalized degree centrality analysis
  - Normalized In-degree centrality analysis
  - Normalized Out-degree centrality analysis
  - Page-rank centrality analysis

5.2.1 Graph connectivity analysis

The analysis of the connectivity of graph models has addressed by using three ways. First, the densification analysis of both transaction graphs and user graphs in each Darknet market network. Second, the degree distributions of the transaction graphs and User graphs. Third, the clustering coefficients of both graph models. Since the interest of this section relies on the connectivity of the network, all the measurements for analyzing the graph connectivity has obtained by the graph models after converting from default Multi Directed Graphs (refer section 4.2.2) to simple graphs (Graphs that have no self-loops and parallel edges).

5.2.1.1 Densification Analysis over time

The densification analysis of the graph models has been measured by using the three measurements. First, the number of vertices and edges of the graph over time. Second, average out-degree of the graph over time. Third, the Percentage of vertices in the maximum strongly connected component in the graph over time. Specifically, these three measurements have been measured in twenty equal timestamp instances in each Darknet market. These timestamp instances are relevant to each Darknet market transaction duration.

The results of the densification analysis have depicted in the bellow figures. Note that there are only a few significant results have been mentioned and the rest of the results have mentioned in Appendix A. Additionally, The results and evaluation of each measurement will be described according to the two graph models namely Transaction graphs and User graphs.

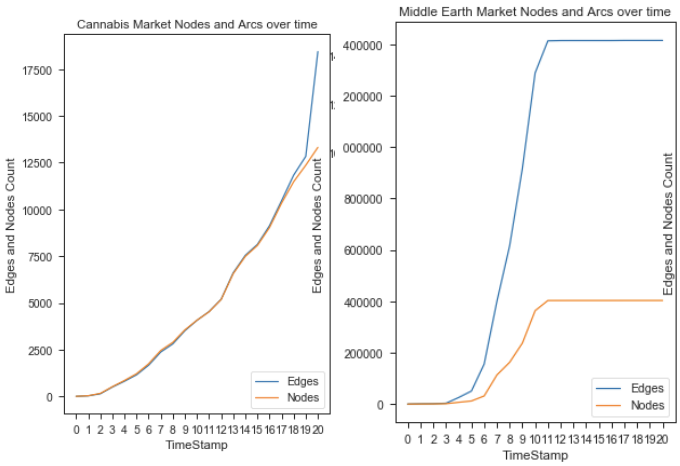
- **Results and evaluation for the number of vertices and edges in the graph over time**

Transaction Graphs

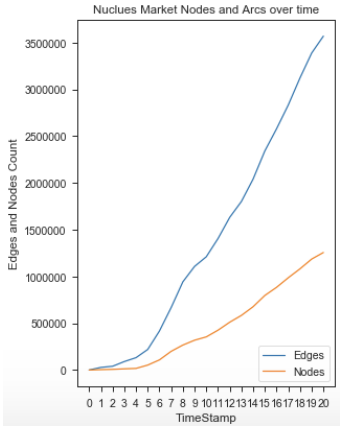
According to Figure 5.5 and Figure A.1 (refer to Appendix A), the overall transaction graphs show that the number of edges and the number of vertices (Node) increases when time increased. Additionally, those plots have shown that the increase in the number of edges and vertices is slightly more than linear. As an exception, Figure 5.5b highlights that after the eleventh timestamp, the number of edges and Nodes was constant over time. It highlights that transactions have stopped and that particular Darknet market has become inactive after the particular timestamp. According to Figure A.7, it highlights that This inactive behavior has occurred in the *Middle Earth*, *Bluesky*, *Abraxas*, *Sheep* and *Pandora* Darknet markets after particular timestamp instances. Therefore, this observation has been verified the results that discussed in the section 5.1.1 regarding the inactiveness of the Darknet market.

User Graphs

Figure 5.6 and Figure A.2 (refer to Appendix A) show that the number of vertices and the number of edges has increased when time increased in overall user graphs. Additionally, those plots have highlighted that this measurement gets the same behavior in both Transaction graphs and User graphs in each particular Darknet market. Therefore, it stated that the property of increasing the number of edges and the number of vertices over time inherited from the Transaction graph to the User graphs. Specifically, the inactiveness property of the market network also inherited to the User graphs which have mentioned in the results of Transaction graphs.

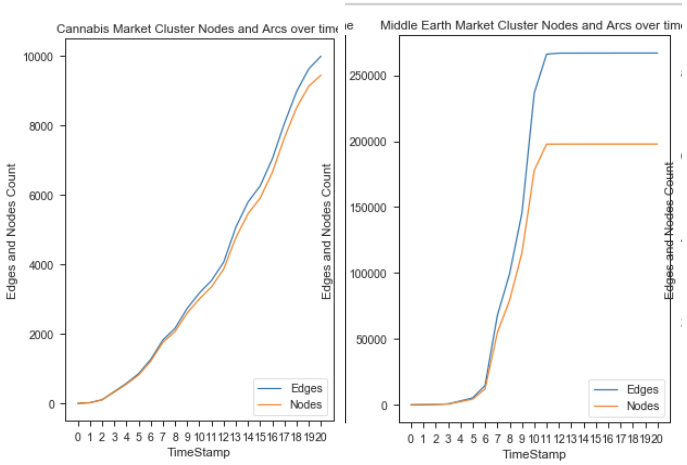


(a) *Cannabis* Market (b) *Middle Earth* Market

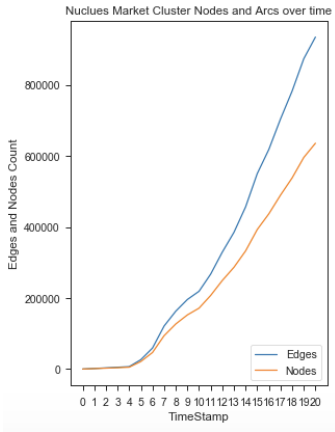


(c) *Nuclues* Market

Figure 5.5: Number of vertices and edges growth against time in the Transaction graph

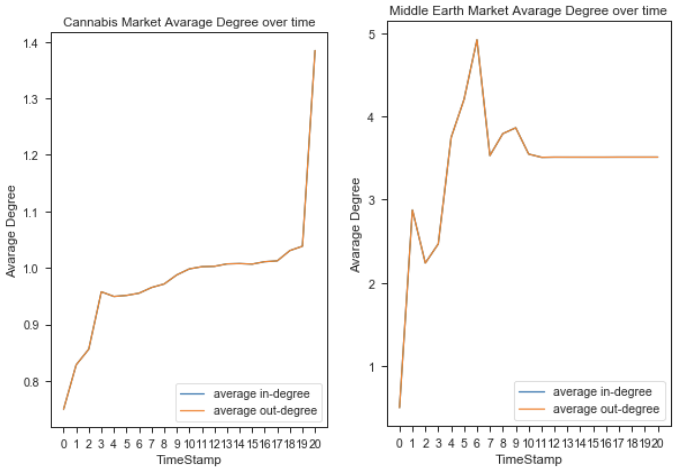


(a) *Cannabis* Market (b) *Middle Earth* Market



(c) *Nuclues* Market

Figure 5.6: Number of vertices and edges growth against time in the User graph



(a) *Cannabis* Market (b) *Middle Earth* Market

Figure 5.7: Average Out-degree against time in Transaction graphs

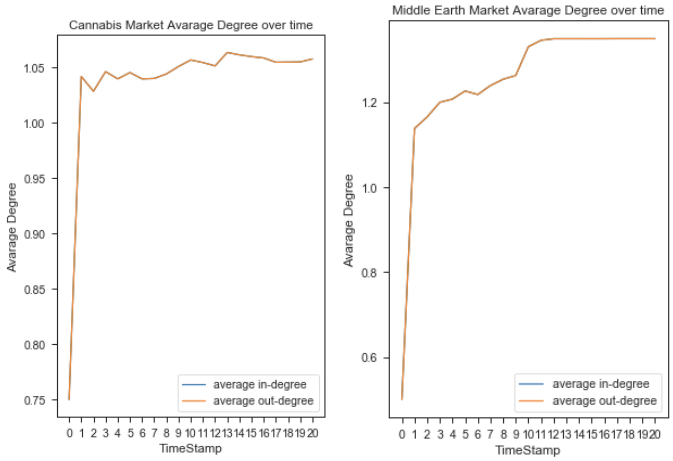
- **Results and evaluation for the average out-degree over time**

Transaction Graphs

In Figure 5.7a shows that the average out-degree(or in-degree) of the transaction graph is increasing over time. This highlights that the number of edges increased much faster than the number of vertices that represent the densification power law [48] of the particular network. Therefore, this behavior has occurred in both the *Cannabis* Darknet market network and the *Sheep* Darknet market (refer Figure A.3). However, in Figure 5.7b shows that the increase of the out-degree only limited to the particular timestamps and the out-degree will be constant after the eleventh timestamp. This behavior represents that the market network has not been dense over time due to the aforementioned inactiveness of the market network. This behavior has occurred in *Middle Earth*, *Bluesky*, *Abraxas*, *Sheep* and *Pandora* Darknet markets which interpreted the inactiveness of the market network as mentioned before.

User Graphs

Unlike in the transaction graphs, Figure 5.8 shows that the user graphs have shown that out-degree(or in-degree) of the graph is increasing over time in all the Darknet market networks (refer Figure A.10). This means that the results of the increasing out-degree (or in-degree) in the user graphs of each Darknet market has shown that the number of edges is increasing much faster than the number of vertices. Therefore, this behavior stated that each user graph of the seven Darknet markets followed the densification power law [48]. However, user graphs of the *Middle Earth*, *Bluesky*, *Abraxas*, *Sheep* and *Pandora* Darknet markets have interpreted the inactiveness property after the particular timestamp as aforementioned.



(a) *Cannabis* Market (b) *Middle Earth* Market

Figure 5.8: Average Out-degree against time in User graphs

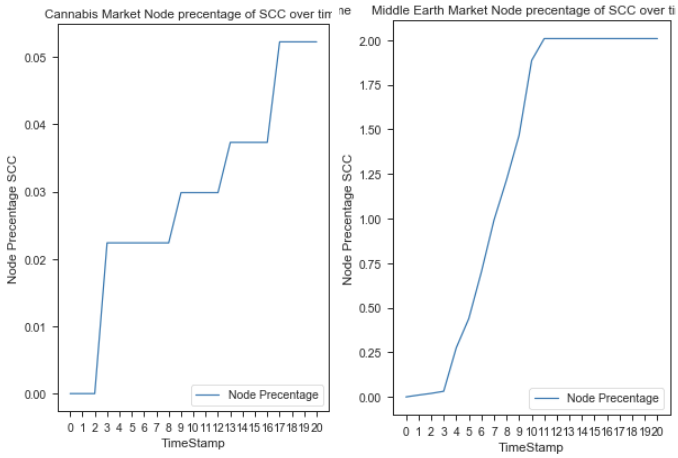
- **Results and evaluation of the Percentage of vertices in the maximum strongly connected component in the graph over time**

Transaction Graphs

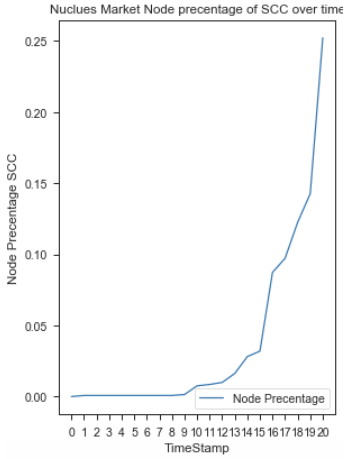
Figure 5.9 shows that the percentage of the vertices in the maximum strongly connected component is slightly increasing over time in each transaction graph of the Darknet market. Therefore, it shows that even though the number of vertices increases fast (refer Figure A.1) over time, the number of vertices in the maximum strongly connected component is increased much faster over time by providing much more robust to the particular Darknet market network. Specifically, Figure 5.9c shows that the time spend to increase the percentage of the nodes in the maximum strongly connected component is much higher than the other Darknet market network which stated that the network is not very active (Transaction of the Nucleus market was not growing very fastly) and network tends to have a slow growth to be robust. Additionally, the results of this measurement also interpret the inactiveness of the few darknet markets as mentioned in section 5.1.1.

User Graphs

Figure 5.10 and Figure A.6 (Refer Appendix A) show that similar behavior as mentioned in the transaction graphs which means that the percentage of the vertices in the maximum strongly connected component increases over time. Additionally, the inactiveness property of the market has inherited to the User graphs from the transaction graphs and that inactiveness has occurred in the same timestamp instances which relevant to the particular Darknet markets as mentioned in the section 5.1.1.

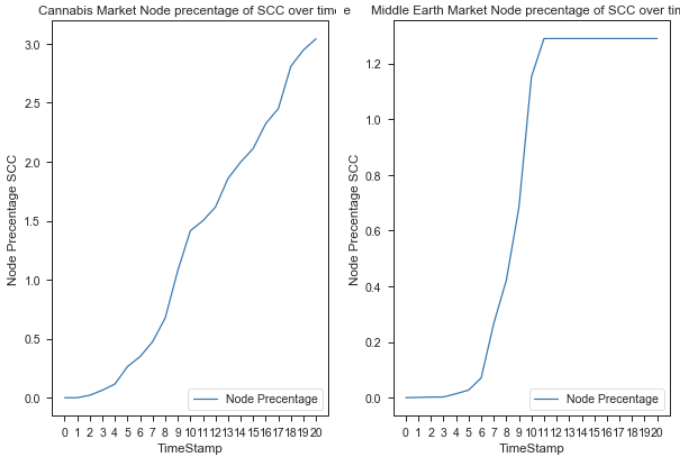


(a) *Cannabis* Market (b) *Middle Earth* Market



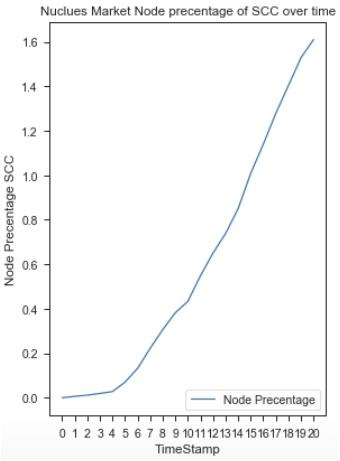
(c) *Nuclues* Market

Figure 5.9: Percentage of vertices in the maximum strongly connected component in the Transaction graph over time



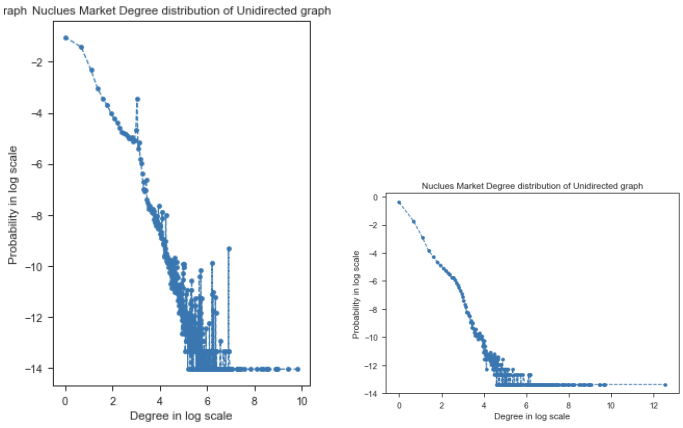
(a) *Cannabis* Market

(b) *Middle Earth* Market



(c) *Nuclues* Market

Figure 5.10: Percentage of vertices in the maximum strongly connected component in the User graph over time



(a) Transaction Graph (b) User Graph

Figure 5.11: Degree Distribution in the Transaction graph and User graph of the Nuclues Market

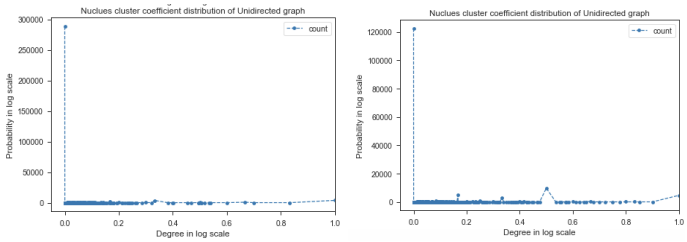
5.2.1.2 Degree Distribution Analysis

The degree distribution of the graph models has addressed in three measurements. Namely, the degree distribution of the undirected graph models, in-degree distribution of the directed graph models and out-degree distribution of the directed graph models (refer section 3.2.2.2). Note that all the degree distributions have plotted in the log-log scale in order to identify where the degree distribution gets normal distribution or power-law distribution [49].

The results and the evaluation of the degree distribution have presented in both Transaction graph models and user graph models. However, there are only a few significant results have been mentioned and the rest of the results have mentioned in Appendix A.

- **Results and evaluation of the degree distribution**

Figure 5.9 and Appendix A shows that all the degree distributions including the in-degree and out-degree distribution of the directed transaction graphs and degree distribution of the undirected transaction graphs in each Darknet market show that they follow the power-law distribution [49]. Therefore, those distributions indicate that few vertices having higher edges and many vertices having few edges [49]. Since each Darknet market network follows the power-law distribution, these Darknet market networks consider as the scale-free networks [49]. Therefore, these results stated that the evolution of each Darknet market network over time assume to be follows the preferential attachment or rich-get-richer principle [49] which means that vertices of the network that already have a high number of edges more likely to establish new edges to them compare to the vertices with a low number of edges when the evaluation of the network occurs over time. Therefore, it indicates that the Bit-



(a) Transaction Graph (b) User Graph

Figure 5.12: Degree Distribution in the Transaction graph and User graph of the Nuclues Market

coin addresses that have involved in a high number of transactions are more likely to involves in many transactions in the future as well. Apart from the power-law distributions, there is another observation can be obtained from the degree distribution which has shown in Figure A.7 and Appendix A. It can be noticed that there are some outliers (spikes) that can be observed in the convergence area of the degree distribution. Those spikes have shown to be related to the particular topological pattern which more likely due to the unexpected user behaviors rather than the natural economy in the Darknet market network [30]. As shown in Figure A.8 and Appendix A, we can observe that all the degree distributions in the user graphs have a similar pattern which that has previously observed from the transaction graphs. Therefore those results of the user graphs show that the patterns and principles that have derived from the degree distribution of transaction graphs are valid for the user graphs as well.

5.2.1.3 Clustering Coefficient Analysis

In the clustering coefficient analysis, the local clustering coefficient of a vertex has been considered. Therefore, the local clustering coefficient has been measured in each Darknet market with respect to both graph models. The results and the evaluation of the clustering coefficient measurement have presented in both graph models (Transaction graph and User graph). However, there are only a few significant results have been mentioned and the rest of the results have mentioned in Appendix A.

- **Results and evaluation for the Clustering Coefficient analysis**

Figure 5.12a and Appendix A shows the local clustering coefficient of the degree distribution results according to the transaction graphs of Darknet markets. Those results show that most of the vertices have zero local clustering coefficients and only a few have higher values of the clustering coefficient. According to equation 3.5(refer section 3.2.2.2), the value of zero in the local clustering coefficient indicates that

the vertices with a degree less than two. Since there is a majority number of vertices get zero value, it states that the value of the average local clustering coefficient (global clustering coefficient) will be close to the value zero. Therefore it interprets that these kinds of scale-free networks tends (refer section 5.2.1.2) to have a low global clustering coefficient and this finding also verified in [50]. Further, these results interpret that there are only a few numbers of transaction hubs (addresses that involve in a higher number of transactions) in the Darknet market networks. Apart from the local clustering coefficient measures from the transaction graphs, Figure 5.12b and Figure A.14(refer Appendix A) show that the local clustering coefficient measurements from the User graphs show the same results as the transaction graphs. Therefore those derived conclusions from the results of transaction graphs are also valid for the results from User graphs.

5.2.2 Centrality analysis

The centrality analysis of the Darknet market networks has been addressed by using the four centrality measures (refer section 3.2.2.2) namely, normalized degree centrality, normalized in-degree centrality normalized out-degree centrality and Pagerank centrality. The results and evaluation of these three measurements have presented using both constructed graph models (Transaction graphs and User Graphs.). However, there are only a few significant results have been mentioned.

- **Results and evaluation for the centrality analysis from the Transaction graphs**

Table 5.1: Highest Centrality Measurement values Obtained from the each Darknet market Transaction graph

	Degree		In-Degree		Out-Degree		Page Rank	
Market Name	Label	Value	Label	Value	Label	Value	Label	Value
Cannabis	f5ed660882	0.169155	f5ed660882	0.169155	6f717b2883	0.043297	f5ed660882	0.055061
Middle Earth	123b335398	0.035630	00e52be96b	0.003737	123b335398	0.035611	123b335398	0.020923
Bluesky	c9dffbdf04	0.060969	LocalBitcoins.com-old	0.008339	c9dffbdf04	0.060963	c9dffbdf04	0.024817
Nuclues	MoonBit.co.in	0.014787	Paymium.com	0.001375	MoonBit.co.in	18565	MoonBit.co.in	0.011186
Sheep	15323043d4	0.024109	15323043d4	0.024109	ad0b75d50e	0.021623	ad0b75d50e	0.005085
Pandora	c9dffbdf04	0.228177	CoinJoinMess	0.010914	c9dffbdf04	0.228175	c9dffbdf04	0.032897
Abaraxas	MoonBit.co.in	0.025958	0232c8ec20	0.002898	MoonBit.co.in	0.014786	MoonBit.co.in	0.019090

Table 5.1 shows that there are two types of labels that can be obtained from the centrality measures. As mentioned in section 3.1, the labels that can be either user/service labels labels (eg: $f5ed660882$, $123b335398$,MoonBit.Co.In) or market labels (eg: CannabisRoadMarket,MiddleEarthMarketplace). According to Table 5.1, most of the central nodes in the transaction graphs of each Darknet market was

labeled as a User label. For an instance, in the Cannabis Darknet market, the node $f5ed660882$ is considered as the most central node in the market network when considering the local connectivity, i.e the degree centrality and considering the whole network, i.e the Page-rank centrality. This behavior can be obtained in the *Middle Earth*, *Bluesky*, *Abraxas*, *Sheep* and *Pandora* Darknet market as well. However, in *Nucleus* and *Abraxas* Darknet markets shows that their most central node was labeled as a Service label and it is MoonBit.Co.In (One of the popular Bitcoin faucet[1] service). According to the degree centrality measures, it is known to measure the nodes that have involved in most transactions (either receive or send payments). Specifically, In out-degree centrality, it is known to measure that nodes that have sent most payments in the network and in in-degree centrality, it is known to measure that nodes that have received most of the payments. Therefore, we can observe that the majority of nodes that previously considered as most central in the network are more likely to send than receiving payments. As an exception in the Cannabis market, the most central node, i.e $f5ed660882$ is more likely to receive than sending payments. Further, when considering the degree centrality measures and Page-rank centrality measures, Table 5.1 highlights that there is a big correlation between two measures. Additionally, Table 5.1 shows that the Page-rank centrality value of most central nodes likely has low values when in the large size of Networks.

- **Results and evaluation for the centrality analysis from the User graphs**

Table 5.2: Highest Centrality Measurement values Obtained from the each Darknet market User graph

	Degree		In-Degree		Out-Degree		Page Rank	
Market Name	Label	Value	Label	Value	Label	Value	Label	Value
Cannabis	CannabisRoadMarket	0.946751	CannabisRoadMarket	0.441733	CannabisRoadMarket	0.505018	CannabisRoadMarket	0.223361
Middle Earth	MiddleEarthMarketplace	0.312905	MiddleEarthMarketplace	0.261359	MiddleEarthMarketplace	0.051546	MiddleEarthMarketplace	0.090787
Bluesky	BlueSkyMarketplace	0.401400	BlueSkyMarketplace	0.277215	c9dffbdf04	0.161148	BlueSkyMarketplace	0.100709
Nuclues	NucleusMarket	0.450751	NucleusMarket	0.286366	NucleusMarket	0.164385	NucleusMarket	0.109862
Sheep	SheepMarketplace	0.385191	SheepMarketplace	0.327027	SheepMarketplace	0.058165	SheepMarketplace	0.124455
Pandora	c9dffbdf04	0.419557	PandoraOpenMarket	0.162153	c9dffbdf04	0.419553	PandoraOpenMarket	0.066245
Abaraxas	AbraxasMarket	0.452724	AbraxasMarket	0.370422	AbraxasMarket	0.082302	AbraxasMarket	0.120151

Like in transaction graphs, Table 5.2 shows that the vertices with the highest centrality values in each centrality measure in each Darknet market along with it's labeled. Unlike in the transaction graphs, the user graphs have shown that the most central node(when considering the degree centrality and Page-rand centrality) in each Darknet market can be labeled as itself. For instance, in the Cannabis Road Darknet market, the most central node is the node with label *"CannabiRoadMarket"*. This behavior can be observed in six out of seven Darknet market User graphs

[1]https://thecoinshark.net/en/bitcoin-faucets-what-is-it-and-how-much-you-can-earn/

namely,*"Cannabis"*, *Bluesky*, *Middle Earth*, "Nucleus", *Sheep*, *Abraxas* Darknet market user graphs. Further, when considering the most central node in those six Darknet market, it shows that those central nodes are equally central when considering receive and send payments. Because those central nodes have the highest value in each Darknet market when considering the in-degree and out-degree centrality measure (refer Table 5.2). However, in *"Pandora"* Darknet market it shows that the node with label *"PandoraOpenMarket"*(self-market label) is more likely to receive payments than send payments and Pandora user graph shows exceptional behavior among other Darknet markets' User graphs.

5.3 Analysis Phase III: Traceability Analysis between Dark web market users and Surface web Bitcoin users.

In the Analysis Phase III, the main concern is to trace the Darknet market users by using the extracted Bitcoin information from the surface web(refer section 3.2.2.3). As mentioned in section 3.1, this Bitcoin information has been collected from [19] and there are 37600 records have used for this analysis. However, this number of records has been reduced to 9881, after the clustering process that has done from these records (refer section 3.2.2.3). These 9881 clusters have used to trace whether the Bitcoin addresses used in Darknet markets are inside those clusters. Therefore, those 9881 clusters have been compared with the clusters of addresses from the Darknet markets(the vertices of the User graphs). As a result of that process, there are 1203 clusters have been identified and those 1203 clusters have consisted of at least one Bitcoin address which is used in the Darknet markets transaction. In another word, those 1203 clusters can be recognized as traced vertices in the user graphs which constructed from the Darknet markets transaction data. Since the vertices of the Darknet market user graphs are recognized as the single entity or the user in the Darknet market (refer section 3.2.2.3), those 1203 clusters can be identified as the single users or single entities in the Darknet market.

Table 5.3: Top 10 nodes that have most transactions in all Darknet markets

Traced Cluster IDs	Number of Transactions	Tag
2560366	52706	LocalBitcoins.com-old
262134758	43986	MoonBit.co.in
261	38679	CoinJoinMess
1381240	13964	Bitstamp.net-old
2250106	12752	BitBargain.co.uk
137821	11571	[00000014ea]
14809960	10736	[046576b9af]
18168070	5983	BTC-e.com
421008	5164	VirWoX.com
7024680	4354	CoinBox.me

The resultant 1203 clusters were used to analyze their behaviors in the Darknet market networks and there are several findings have been observed. First, the activeness inside the Darknet market network of those 1203 entities(or clusters) has been analyzed. Table 5.3 shows that top 10 clusters that have most transactions inside all seven Darknet markets networks. According to Table 5.3, cluster ID 2560366 has the most transactions and it can be recognized as the most active node among the traced entities from the surface web. Further, when considering the "tags" among those top 10 entities, it shows that the most active nodes are some Bitcoin services, i.e: "LocalBitcoins.com-old", "Bitbargain.co.uk", "MoonBit.co.in". Among those Bitcoin services, "LocalBitcoins.com" well known Bitcoin exchange service and "MoonBit.co.in" is a popular Bitcoin faucet. Therefore, it can be stated that there is a high probability that has, those Bitcoin services are interacting within both the surface web and Darkweb in a highly active manner. However there some user nodes such as 00000014ea and 046576b9af also have the average transaction rate among the Darknet markets. In Figure 6.2 shows how those transactions were evaluated against the timestamp and significantly, It shows that cluster ID 261 shows some static behavior from its transactions instead of being inactive over time.

Apart from the transaction rate and the activeness of the traced entities, the 1203 clusters have been used to analyze the total income of those clusters inside all seven Darknet markets. Table 5.4 shows that the top 10 clusters that have the most total income from the transactions among Darknet markets. It indicates that the entities like 261 and 2560366 have become the frontrunners in the total income. Notably, those frontrunners in the total income also become the frontrunners when considering the most active nodes inside the Darknet markets. However, significantly cluster

ID 18168070 has some more total income despite being inactive up to some context.

Table 5.4: Top 10 nodes that have most income in all Darknet markets

Traced Cluster IDs	Money(BTC)	Tag
261	22798.154766	CoinJoinMess
2560366	20121.908885	LocalBitcoins.com-old
18168070	16807.989293	BTC-e.com
1381240	16312.749904	Bitstamp.net-old
2250106	8607.526904	BitBargain.co.uk
137821	5722.159051	[00000014ea]
421008	5535.616270	VirWoX.com
1310890	3938.946484	[000f5614ea]
263048303	3756.819557	[0545f5614f5]
1479481	2554.440044	[789040f576fa]

5.4 Summary

In this chapter, the results and evaluation of the research study have presented in three sections. In section 5.1, the results and evaluation of the overall market analysis have been presented. In section 5.2 the results and evaluation of the analysis phase II have presented. Further, the results and the evaluation of the graph properties have been addressed. In the final section, the results of the traceability analysis have been presented and it shows that the possibility of tracing the Darknet market users using the surface web Bitcoin information.

Chapter 6

Conclusions

This chapter presents the overall conclusions about the research questions and research problems. Then, the chapter will examine the limitation of this research study and finally, the chapter will conclude by giving the future implications from this research study.

6.1 Conclusions about research questions

Research Question 1: What are the existing approaches for analyzing the Darknet market network?

There were several approaches have identified through the literature review for analyzing the Darknet market network. The main approaches that identified were analyzing the Darknet market through product-related listings, analyzing through the classification of product photos of the Darknet markets, analyzing through the discussion forum data between vendors and the buyers on the Darknet market network and analyzing through the Bitcoin transaction data. The limitations and advantages have been identified and described in Chapter 2.

Research Question 2: What are the existing analysis techniques based on the Bitcoin transaction data?

The existing analysis techniques based on the Bitcoin transaction data have addressed in the Literature review(refer Chapter 2) and those identified techniques have categorized into two main techniques such as graph-based analysis and non-graph-based techniques. The limitations and advantages of these techniques have identifies and described in Chapter 2.

Research Question 3: How to construct the transaction graphs and user graphs using the Bitcoin transaction data of Darknet markets?

As mentioned in section 3.2.1, there two types of graph models have constructed within this research study. First, the transaction graph models and second, the user graph models. In transaction graph models, the vertices represent the Bitcoin public key addresses of the Bitcoin transactions and the edges represent the particular transaction between the two or more Bitcoin transaction addresses (refer section 3.2.1). In user graph models, the vertices have been constructed by using a clustering technique in order to represent clusters of addresses and the edges have been constructed by considering the transaction occurring between the addresses inside those clusters (refer section 3.2.1). The detailed description of the construction process of the Transaction graphs and User graphs has been mentioned in Chapter 3.

Research Question 4: How to analyze the Darknet market network by exploiting graph models (transaction graphs and user graphs)?

As mentioned in Chapter 5, the analysis of the Darknet market has done in three analysis phases. The first and second analysis phases have obtained the analysis of the properties and behaviors of the Darknet market network. Specifically, in the second analysis phase, the topological patterns and behaviors of the Darknet market network have addressed by exploiting the constructed graph models (Transaction graphs and User graphs). Further, there were seven Darknet markets have used and Both graph models have constructed according to each Darknet market transaction data. The conclusions derived from the first and second analysis phases have mentioned below.

- Conclusions from the Analysis Phase I
  - Results with analysis of overall transaction flow among the Darknet market shows that all the Darknet markets that used in this analysis have maintained their transactions mostly in a short period of time.
  - Results with analysis of overall money flow among the Darknet markets shows that Darknet market users are used to doing their tradings with popular Darknet markets rather than small unpopular Darknet markets.
- Conclusions from the Analysis Phase II
  - The results of connectivity analysis from the graph models have obtained that some of the Darknet markets have shown inactiveness behavior in the transactions after the particular timestamp. The reasons for the inactiveness of the darknet market network have described in section 5.1.1.

- The results of the densification analysis from both transactions and user graph models stated that all seven Darknet markets' networks have keen to follow the densification power law [48] (The number of edges grows faster than the number of vertices) (refer section 5.2.1.1). Further, it stated that over the time all the networks are likely to be more dense until those Darknet markets start to behave inactive.
- The results of the degree distribution from both graph models stated that degree distribution of these Darknet markets networks have followed the power-law distribution [49] and the majority out of the seven Darknet markets can be recognized as scale-free networks. Therefore, we came to the conclusion that the majority of the Darknet market networks have high probability to obey the rich get richer property when the network grows over time (refer section 5.2.1.2).
- The results of the clustering coefficient stated that every Darknet market have few numbers of transaction hubs (the nodes that involve for most of the transactions in the market network) and a high number of nodes that have less than the degree two (refer section 5.2.1.3).This stated that the most of users in the Darknet markets more used to interact with the popular nodes and most of the users were act as a isolated entities in the market with doing few transactions.
- The results of the centrality analysis stated that most of the central and active nodes in the network are the Darknet market itself and other nodes that involves the transactions have low centrality values than the market network it self (refer section 5.2.2). Therefore this finding verified the conclusions that derived from the section 5.2.1.3.

Research Question 5: How to trace the Darknet market users from using the Bitcoin data?

As mentioned in section 5.3, it shows that the approach this research study follows for tracing the Darknet market entities (or users) has succeeded and the results of the traceability analysis in the third phase of analysis show that there are 1203 entities(Darknet market users) have identified in the Darknet market network. Further, the results of this analysis show that most of the traced entities are well-known Bitcoin services,i.e: "LocalBitcoins.com-old" and"MoonBit.co.in", and there is a high probability that those Bitcoin services are interact within both the Darkweb and surface web highly active manner. Additionally, the results of the traceability analysis have observed that those traced Bitcoin services are keen to have more income with a high number of transactions inside the Darknet markets.

6.2 Conclusions about research problem

Section 1.2.1, discussed that there is a need for proper analysis over Darknet market networks in order to investigate the properties and dynamic behaviors of the Darknet market network and trace the users of the Darknet markets with the Bitcoin users in the surface web. Section 3.2 explained the proposed research design for the proper analysis over the Darknet market networks and chapter 5 has explained the derived conclusions from the results of analysis by using the three phases of analysis. According to those results of the analysis, it concluded that this research has been provide a precise approach to investigate the properties and dynamic behavior inside the Darknet markets network and provide a proper approach to investigate the Darknet market users and their behavior inside the Darknet markets.

6.3 Limitations

The main limitation of this research study is that this analysis work only limited to the seven Darknet markets. Therefore, it can be investigated more topological patterns and behaviors around the Darknet market if there is more number of Darknet market networks available for analyzing. Apart from the limitation in the number of Darknet markets, there is another limitation can be identified in the process of traceability analysis. Since the process of traceability analysis relies on the crawled data set from the Blockchain.info [18] , it has a limited number of few Bitcoin transactions. Therefore, the results of the traceability analysis have limited to this few numbers of Bitcoin transaction data from the Blockchain.info [18].

6.4 Implication for further research

This research study can be extended to several areas. As mentioned in section 6.3, this analysis work can be extended with more number of Darknet market networks and analyze more properties in the Darknet market network. Apart from that, this research study also can be extended using a developed novel clustering approach for constructing the user graphs. Since this research study addressed the properties and dynamic behaviors of the Darknet market network, another research study can be proposed for generalizing these network findings with other large networks such as social networks, food networks, etc. Since the process of traceability analysis of the Darknet market users based on the clustering approach, there can be another further research implication that can be done to improve this process of traceability analysis using another approach like tracing users by identifying their IP addresses.

6.5 Summary

This chapter has discussed the overall conclusions for this research study. It begins by presenting the conclusions about the research questions and research problems. Then the chapter has provided the limitation of this research study and then finally concluded by presenting the future implications.

References

[1] Toms Reksna. *Complex Network Analysis of Darknet Black Market Forum Structure*. PhD thesis, 2017.

[2] Romain Espinosa. Scamming and the reputation of drug dealers on Darknet Markets. *International Journal of Industrial Organization*, 67(July), 2019.

[3] Satoshi Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash Syste. 2009.

[4] Malte Möser, Kyle Soska, Ethan Heilman, Kevin Lee, Henry Heffan, Shashvat Srivastava, Kyle Hogan, Jason Hennessey, Andrew Miller, Arvind Narayanan, and Nicolas Christin. An Empirical Analysis of Traceability in the Monero Blockchain. *Proceedings on Privacy Enhancing Technologies*, 2018(3):143–163, 2018.

[5] Sergei Tikhomirov. Ethereum: State of knowledge and research perspectives. In *Lecture Notes in Computer Science*, volume 10723 LNCS, pages 206–221. Springer, Cham, 2018.

[6] Seunghyeon Lee, Changhoon Yoon, Heedo Kang, Yeonkeun Kim, Yongdae Kim, Dongsu Han, Sooel Son, and Seungwon Shin. Cybercriminal Minds: An investigative study of cryptocurrency abuses in the Dark Web. *The Network and Distributed System Security Symposium (NDSS)*, (February), 2019.

[7] Daniel Moore and Thomas Rid. Cryptopolitik and the darknet. *Survival*, 58(1):7–38, 2016.

[8] Jordi Herrera-joancomart. Research and Challenges on Bitcoin Anonymity.pdf. In *9th InternationalWorkshoponData Privacy Management*, volume LNCS 8872, pages 1–14, 2014.

[9] Sevil Guler. *Secure, Shared Bitcoin Wallet*. PhD thesis, UNIVERSITY OF TARTU, 2015.

[10] Kristin Finklea. Dark Web. Technical report, Washington D.C., 2017.

[11] Ben R. Lane, David Lacey, Neville A. Stanton, Anita Matthews, and Paul M.L. Salmon. The dark side of the net: Event analysis of systemic teamwork (EAST) applied to illicit trading on a darknet market. *Proceedings of the Human Factors and Ergonomics Society*, 1:282–286, 2018.

[12] J. Broséus, D. Rhumorbarbe, C. Mireault, V. Ouellette, F. Crispino, and D. Décary-Hétu. Studying illicit drug trafficking on Darknet markets: Structure and organisation from a Canadian perspective. *Forensic Science International*, 264(October 2017):7–14, 2016.

[13] Julian Broséus, Damien Rhumorbarbe, Marie Morelato, Ludovic Staehli, and Quentin Rossy. A geographical analysis of trafficking on a popular darknet market. *Forensic Science International*, 277(October):88–102, 2017.

[14] Martijn Spitters, Femke Klaver, Gijs Koot, and Mark Van Staalduinen. Authorship Analysis on Dark Marketplace Forums. *Proceedings - 2015 European Intelligence and Security Informatics Conference, EISIC 2015*, 1:1–8, 2016.

[15] Xiangwen Wang, Peng Peng, Chun Wang, and Gang Wang. You Are Your Photographs. In *Asia Conference on Computer and Communications Security*, pages 431–442, 2018.

[16] Christian Janze. Are cryptocurrencies criminals best friends? examining the coevolution of bitcoin and darknet markets. *AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation*, 2017-Augus(October), 2017.

[17] Annika Baumann, Benjamin Fabian, and Matthias Lischke. Exploring the Bitcoin network. *WEBIST 2014 - Proceedings of the 10th International Conference on Web Information Systems and Technologies*, 1(April 2014):369–374, 2014.

[18] Blockchain.com - The Most Trusted Crypto Company. `https://www.blockchain.com/`. [Online; accessed 05-January-2020].

[19] WalletExplorer.com: smart Bitcoin block explorer. `https://www.walletexplorer.com/`. [Online; accessed 05-January-2020].

[20] Harry Kalodner, Steven Goldfeder, Alishah Chator, Malte Möser, and Arvind Narayanan. BlockSci: Design and applications of a blockchain analysis platform. *ArXiv*, 1709.02489, 2017.

[21] S Suneetha. Unveiling Deep Web , a High-Quality , Quantitative Information Resource. *International Journal of Latest Trends in Engineering and Technology*, 9(2):167–174, 2017.

[22] Www.darknetstats.com. DARK NET MARKETS COMPARISON CHART. https://www.darknetstats.com/dark-net-markets-comparison-chart/.

[23] Kyle Soska and Nicolas Christin. Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. *Proceedings of the 24th USENIX Security Symposium*, (August):33–48, 2015.

[24] Yiming Zhang, Yujie Fan, Liang Zhao, Wei Song, Shifu Hou, Chuan Shi, Yanfang Ye, Xin Li, Jiabin Wang, and Qi Xiong. Your style your identity: Leveraging writing and photography styles for drug trafficker identification in darknet markets over attributed heterogeneous information network. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pages 3448–3454, 2019.

[25] Luo. *An exploratory investigation into the darknet marketplace discussion forum Agora*. PhD thesis, 2017.

[26] Andrea Evangelista, L Allodi, and M Cremonini. *Darknet markets competitive strategies in the underground of illicit goods Darknet Markets : Competitive Strategies in the Underground of Illicit Goods Student :*. PhD thesis, Eindhoven University of Technology, 2018.

[27] Joseph Bonneau, Arvind Narayanan, Andrew Miller, Jeremy Clark, Joshua A. Kroll, and Edward W. Felten. Mixcoin: Anonymity for bitcoin with accountable mixes. *International Conference on Financial Cryptography and Data Security*, 8437:486–504, 2014.

[28] Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M. Voelker, and Stefan Savage. A fistful of Bitcoins: Characterizing payments among men with no names. *Communications of the ACM*, 59(4):86–93, 2016.

[29] Michael Fleder, Michael S. Kester, and Sudeep Pillai. Bitcoin Transaction Graph Analysis. *ArXiv*, abs/1502.0(February), 2015.

[30] Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. Uncovering the bitcoin blockchain: An analysis of the full users graph. In *Proceedings - 3rd IEEE International Conference on Data Science and Advanced Analytics, DSAA 2016*, number February 2019, pages 537–546, 2016.

[31] Micha Ober, Stefan Katzenbeisser, and Kay Hamacher. Structure and anonymity of the bitcoin transaction graph. *Future Internet*, 5(2):237–250, 2013.

[32] Dániel Kondor, Márton Pósfai, István Csabai, and Gábor Vattay. Do the rich get richer? An empirical analysis of the Bitcoin transaction network. *PLoS ONE*, 9(2), 2014.

[33] Elli Androulaki, Ghassan O. Karame, Marc Roeschlin, Tobias Scherer, and Srdjan Capkun. Evaluating user privacy in Bitcoin. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7859 LNCS, pages 34–51. Springer, Berlin, Heidelberg, 2013.

[34] Dmitry Ermilov, Maxim Panov, and Yury Yanovich. Automatic bitcoin address clustering. In *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, volume 2017-Decem, pages 461–466, 2017.

[35] Yazan Boshmaf, Husam Al Jawaheri, and Mashael Al Sabah. BlockTag: Design and applications of a tagging system for blockchain analysis. *IFIP Advances in Information and Communication Technology*, 562:299–313, 2019.

[36] #bitcoin-otc. `https://bitcoin-otc.com/`. [Online; accessed 06-January-2020].

[37] Bitcoin Forum - Index. `https://bitcointalk.org/`. [Online; accessed 19-January-2020].

[38] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data*, 1(1), 2007.

[39] Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. Data-driven analysis of Bitcoin properties: exploiting the users graph. *International Journal of Data Science and Analytics*, 6(1):63–80, 2018.

[40] Tab Snijders. Accounting for degree distributions in empirical analysis of network dynamics. *Dynamic social network modeling and ...*, (December):1–16, 2003.

[41] Junlong Zhang and Yu Luo. Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network. *Advances in Intelligent Sytems Research*, 132(Msam):300–303, 2017.

[42] Guy Pujolle and Otto Spaniol. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer networks and ISDN systems*, 24(2):107–108, 2003.

[43] Welcome to python.org. `https://www.python.org/`.

[44] urllib2 - extensible library for opening urls¶. https://docs.python.org/2/library/urllib2.html. [Online; accessed 14-January-2020].

[45] Xml and html with python. https://lxml.de/. [Online; accessed 23-January-2020].

[46] Aric Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using networkx. 2008.

[47] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. Gephi: An open source software for exploring and manipulating networks, 2009.

[48] Jure Leskovec, Jon M. Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *KDD '05*, 2005.

[49] Andrew T. Stephen and Olivier Toubia. Explaining the power-law degree distribution in a social commerce network. 2009.

[50] Liudmila Ostroumova and Egor Samosvat. Global clustering coefficient in scale-free networks. In *WAW*, 2014.

Appendices

Appendix A

Diagrams

A.1 Evaluation of Vertices and Edges over time

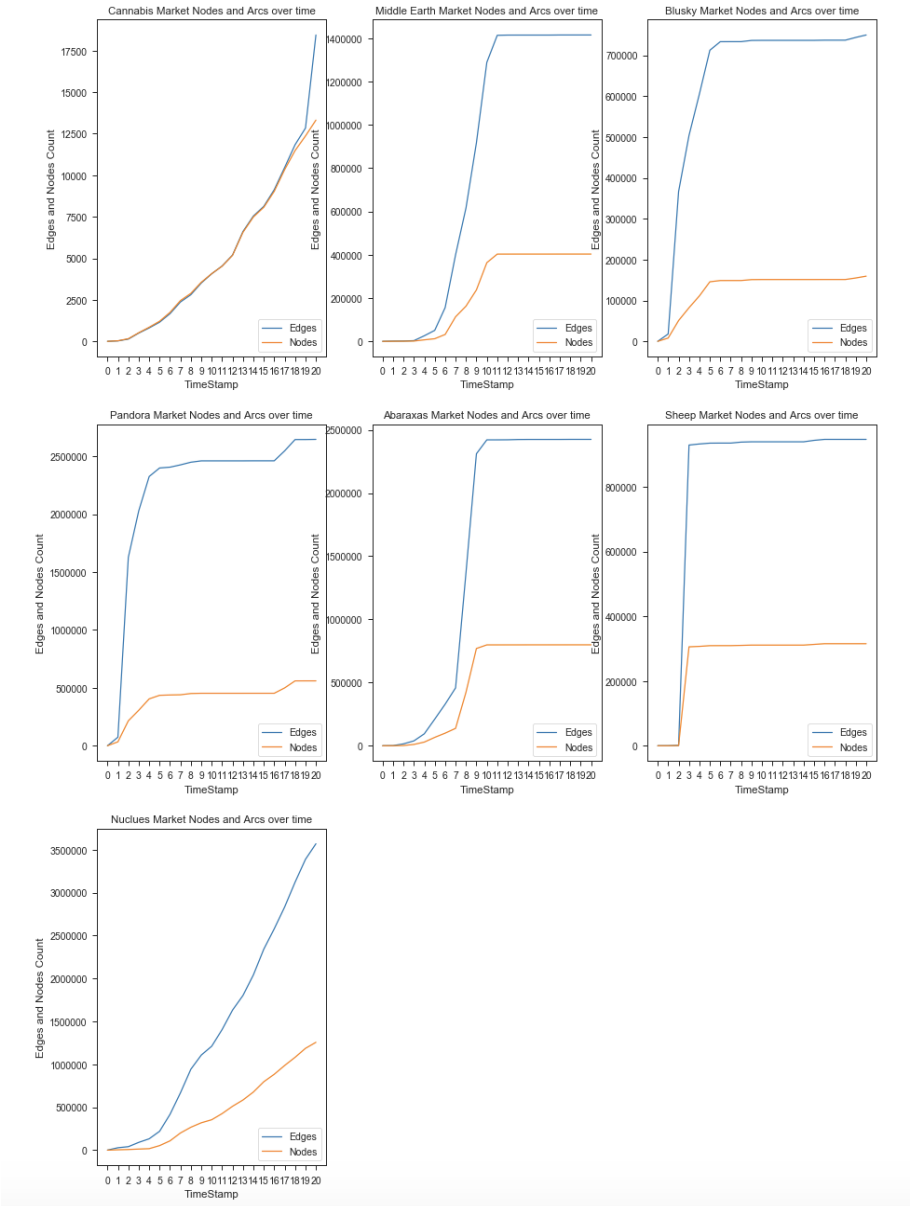


Figure A.1: Evaluation of Vertices and Edges of the Transaction Graph Over Time in each Darknet market

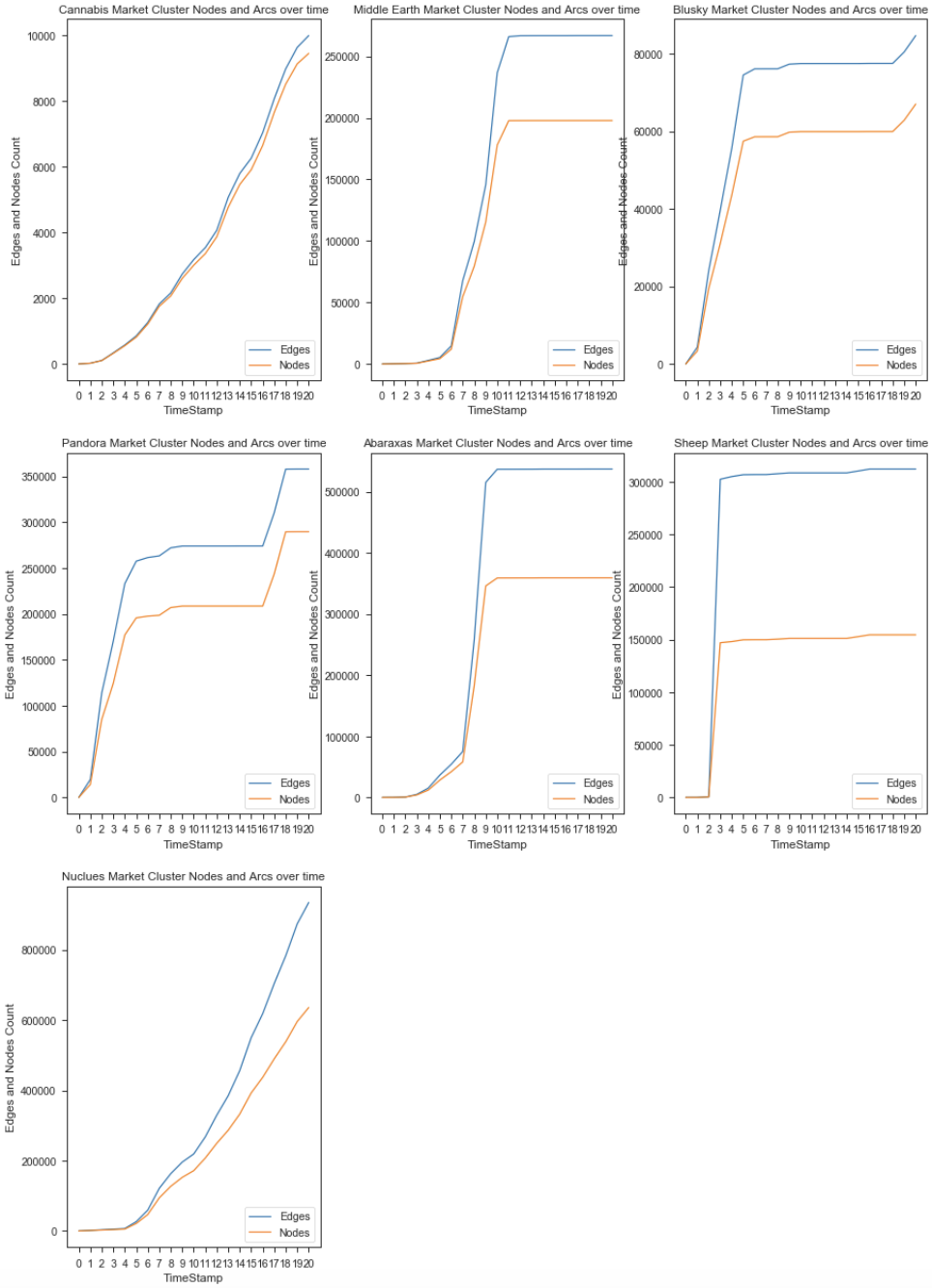


Figure A.2: Evaluation of Vertices and Edges of the User Graph Over Time in each Darknet market

A.2 Evaluation of Vertices and Edges Average Outdegreeover time

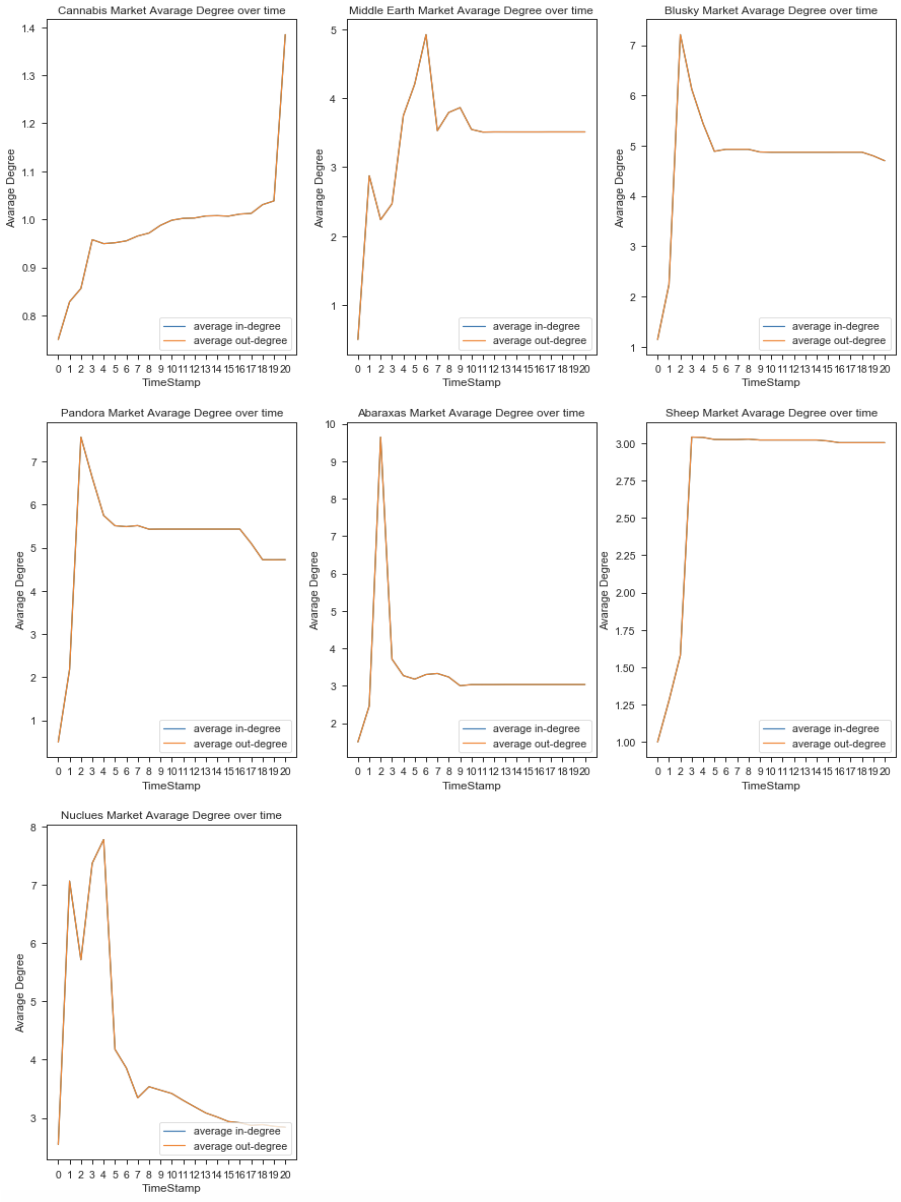


Figure A.3: Evalution of Average Out Degree of the Transaction Graph Over Time in each Darknet market

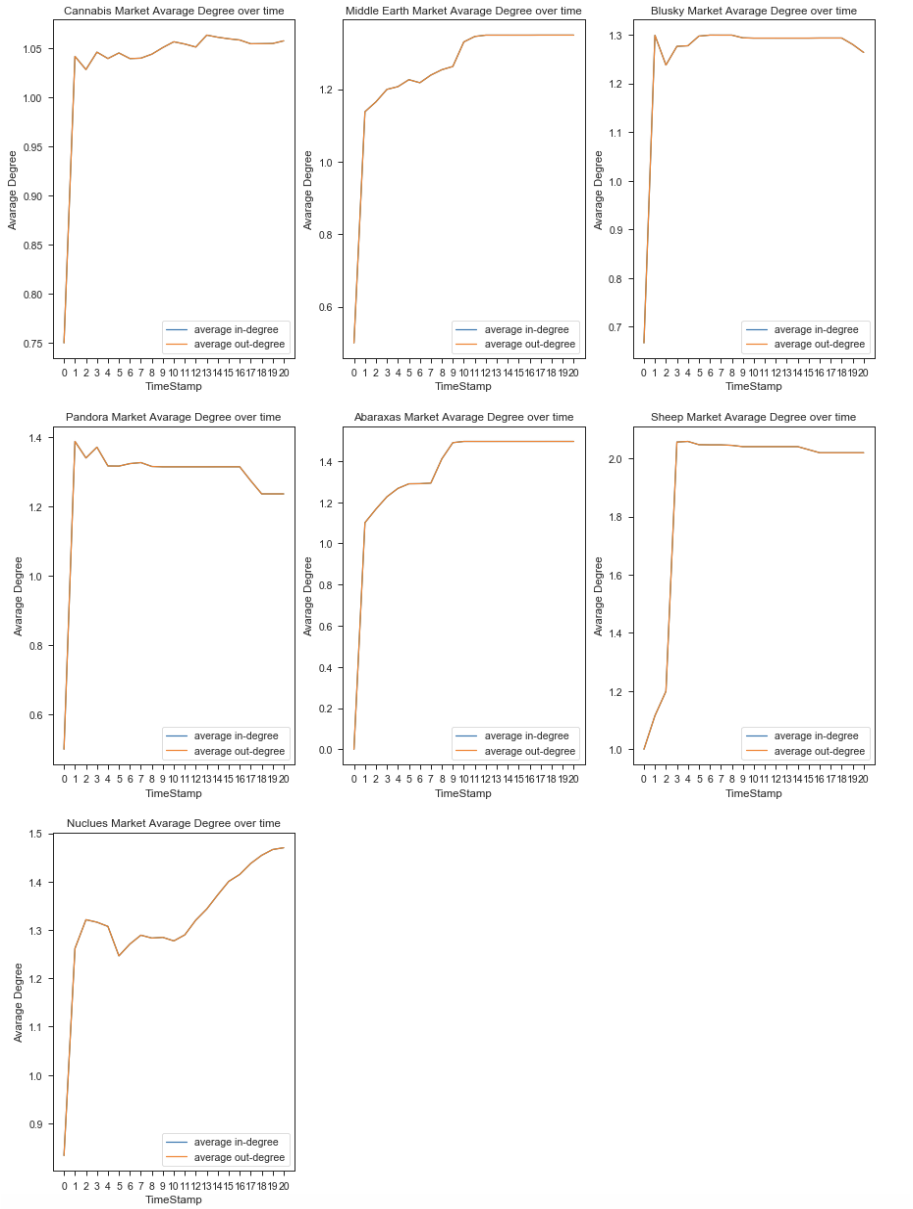


Figure A.4: Evalution of Average Out Degree of the User Graph Over Time in each Darknet market

A.3 Percentage of vertices in the maximum strongly connected component in the over time

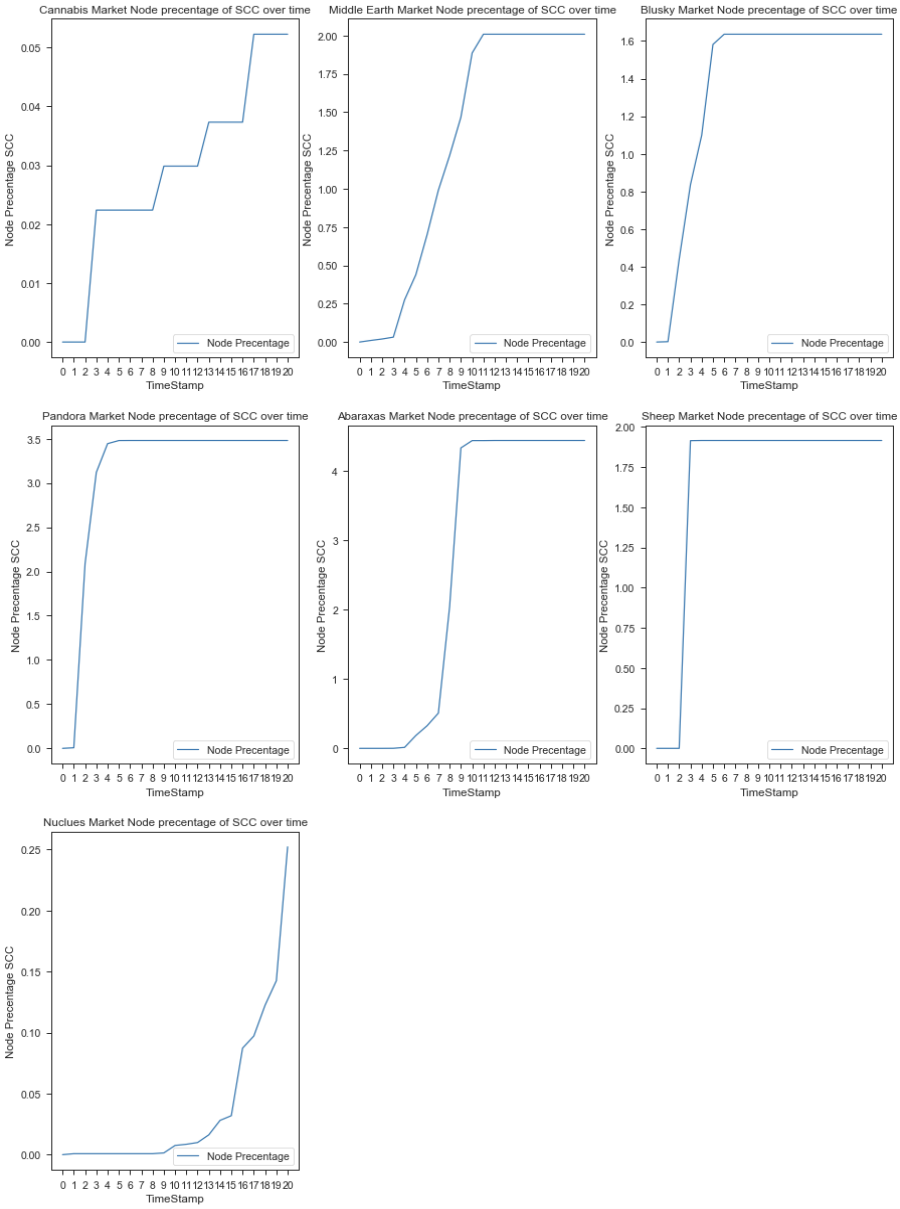


Figure A.5: Percentage of vertices in the maximum strongly connected component in the Transaction graph over time in each Darknet market

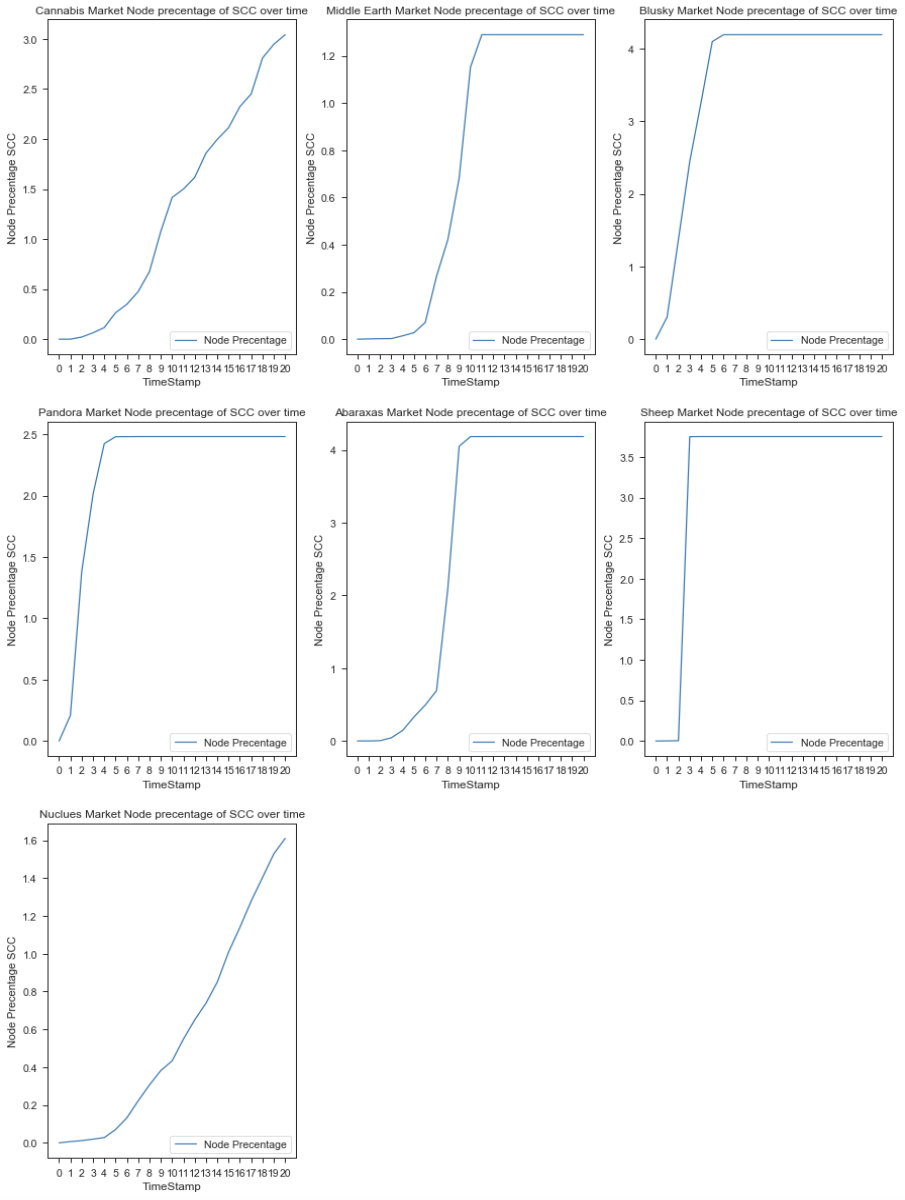


Figure A.6: Percentage of vertices in the maximum strongly connected component in the User graph over time in each Darknet market

A.4 Degree Distribution

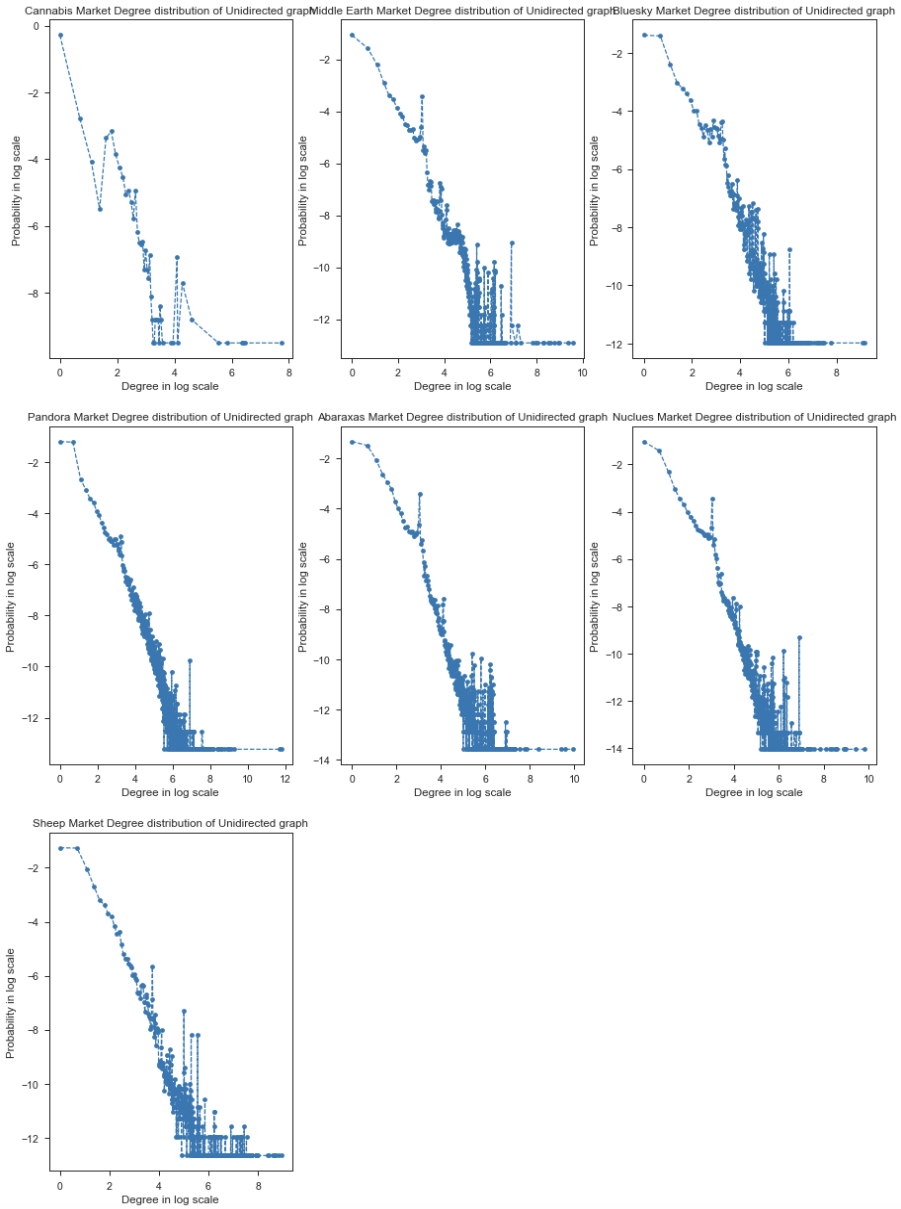


Figure A.7: Degree Distribution in Transaction Graph for each Darknet market

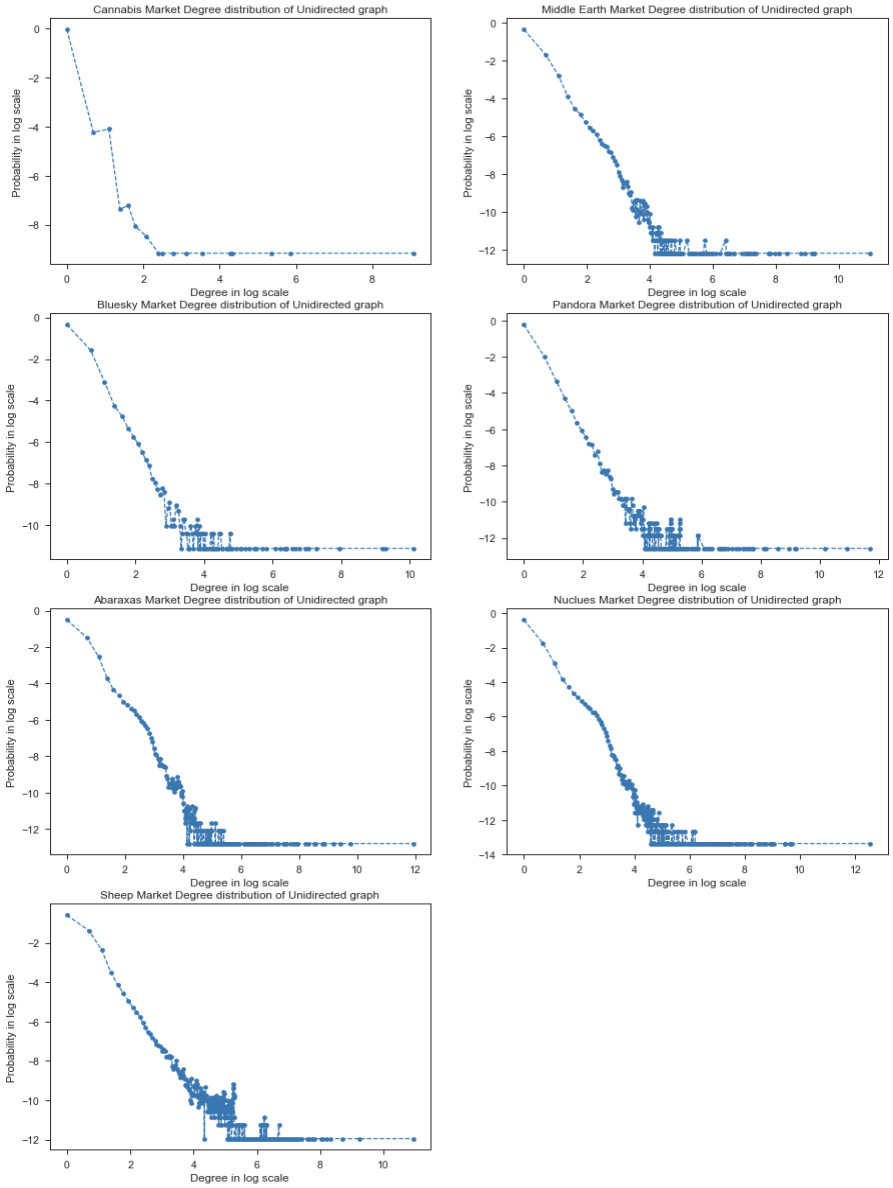


Figure A.8: Degree Distribution in User Graph for each Darknet market

A.5 InDegree Distribution

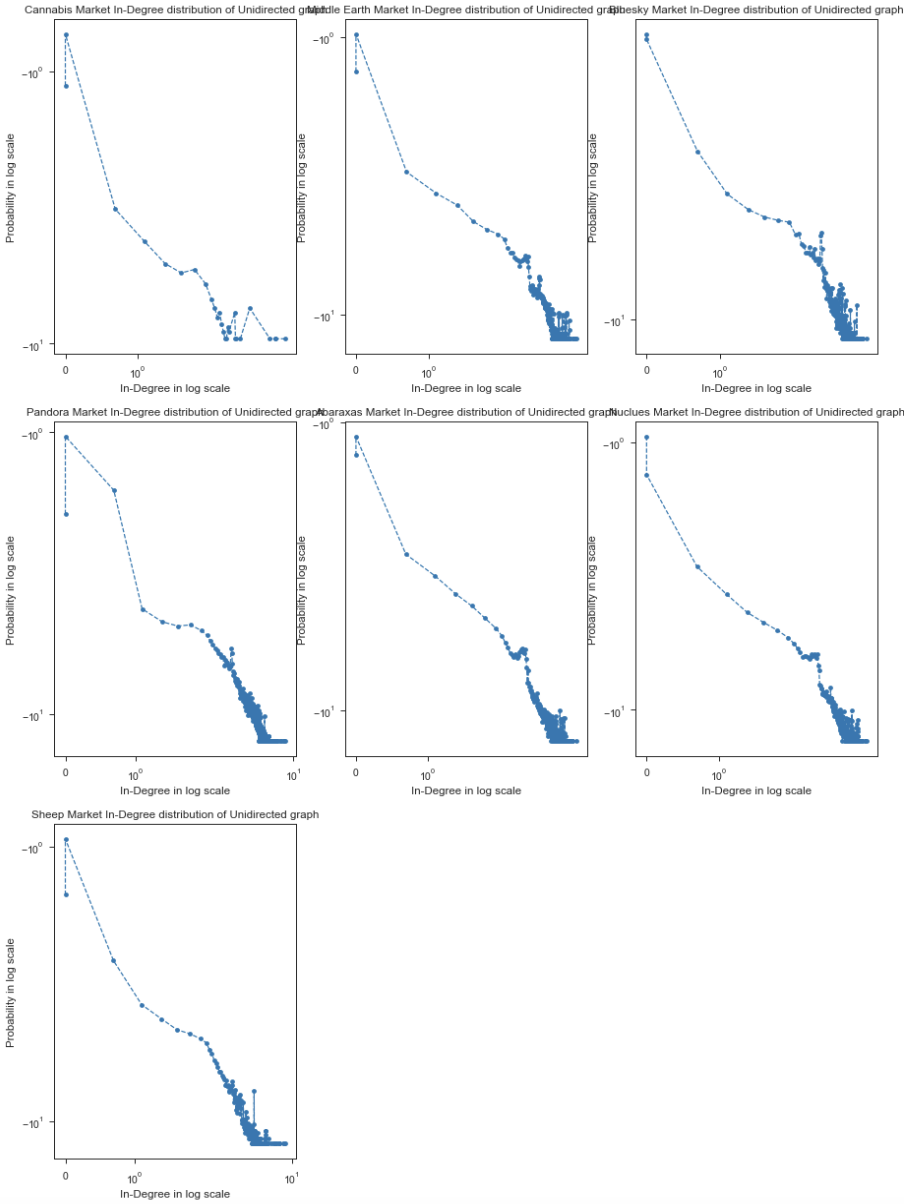


Figure A.9: InDegree Distribution in Transaction Graph for each Darknet market

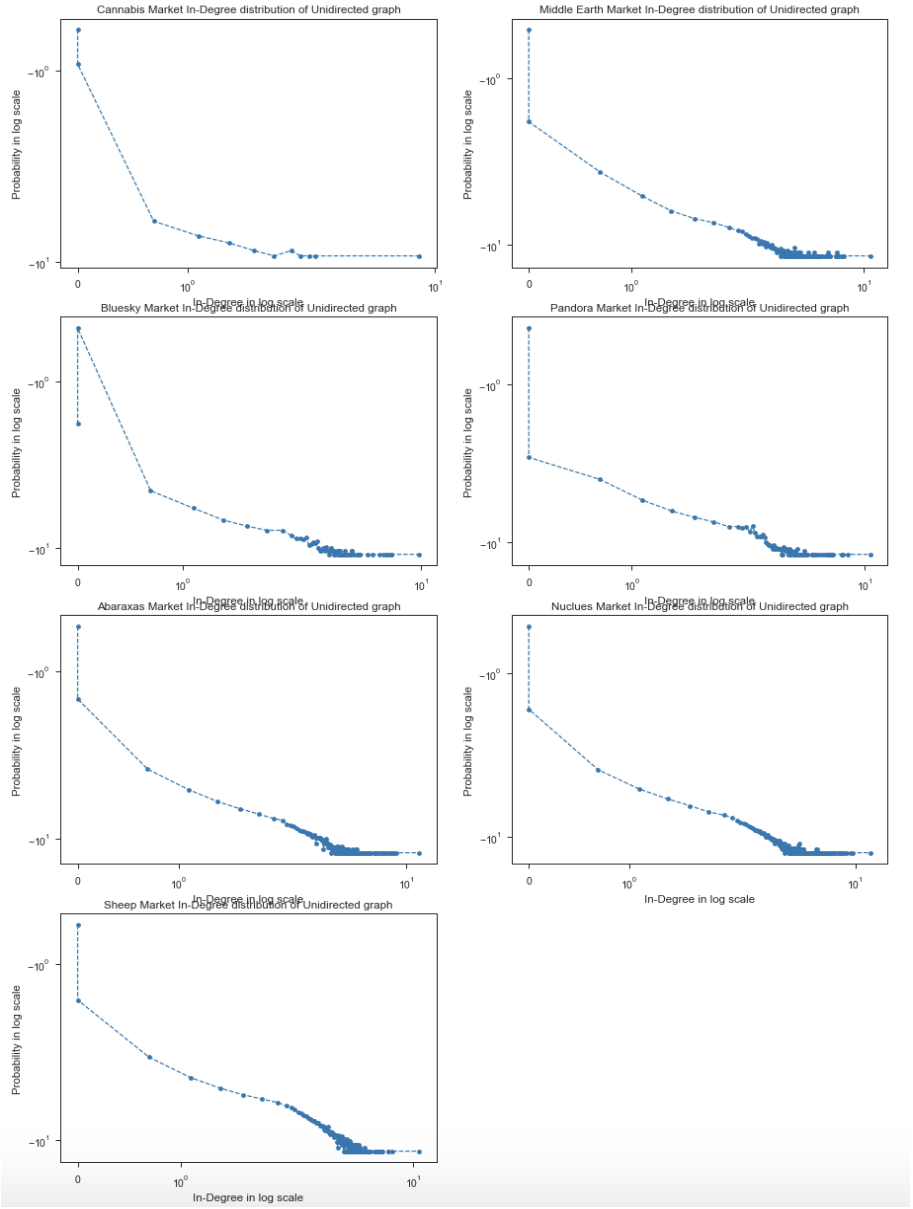


Figure A.10: InDegree Distribution in User Graph for each Darknet market

A.6 OutDegree Distribution

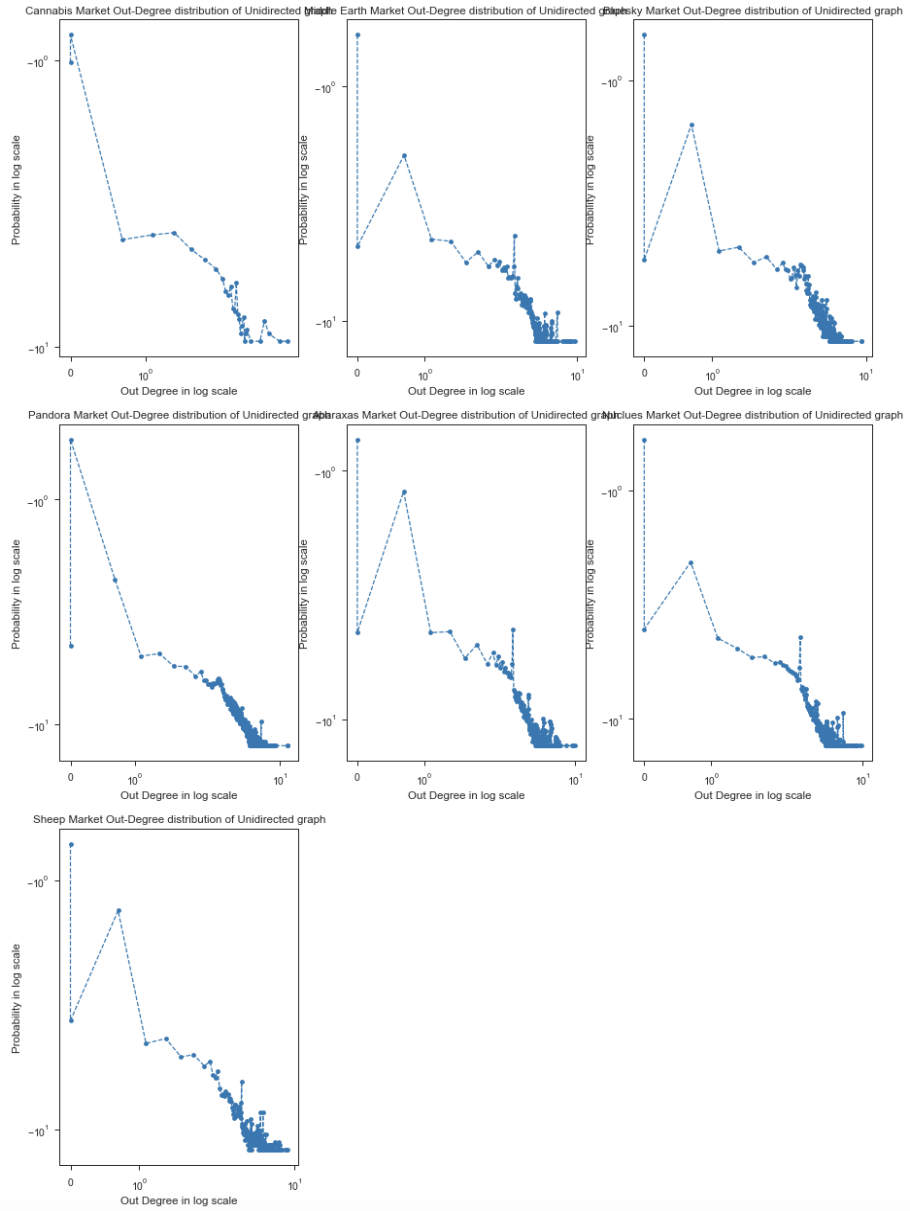


Figure A.11: OutDegree Distribution in Transaction Graph for each Darknet market

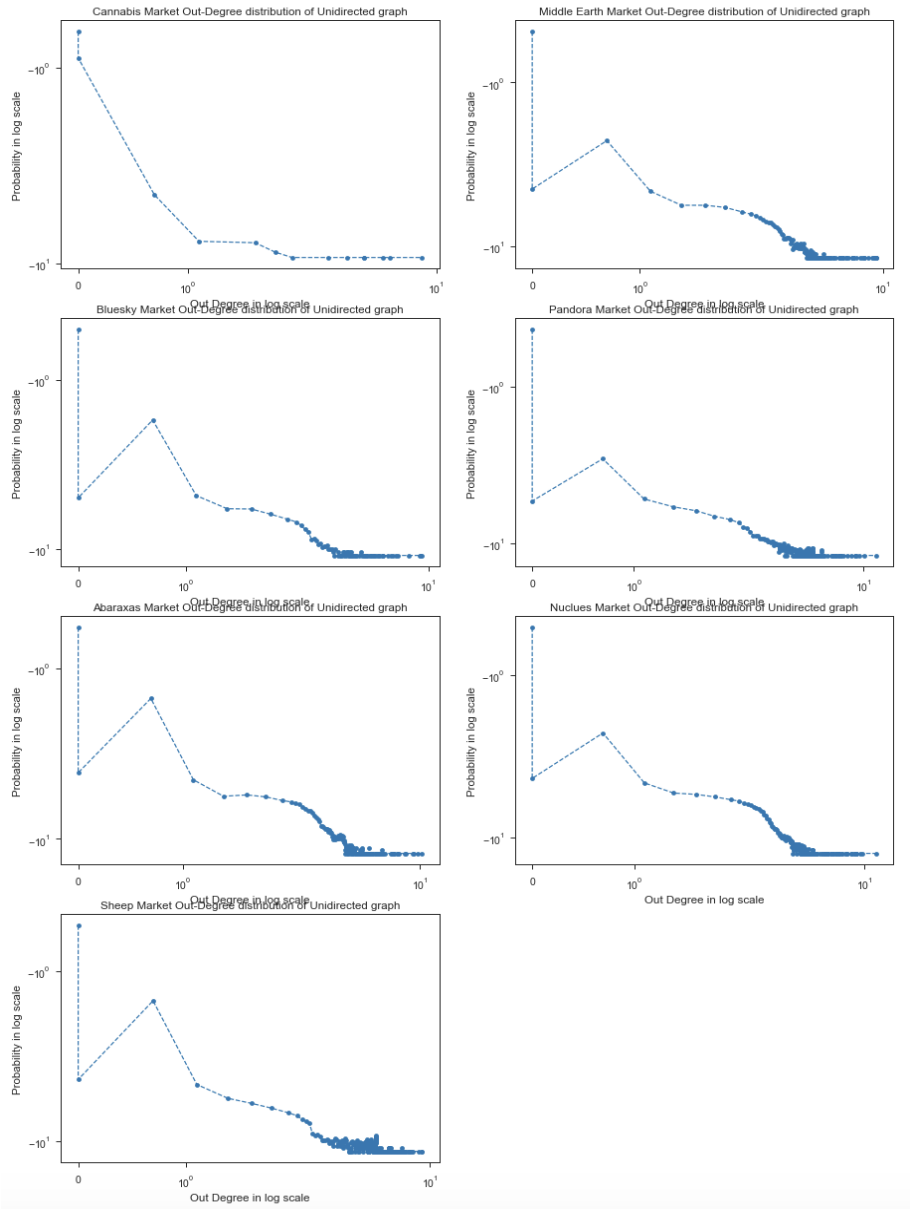


Figure A.12: OutDegree Distribution in User Graph for each Darknet market

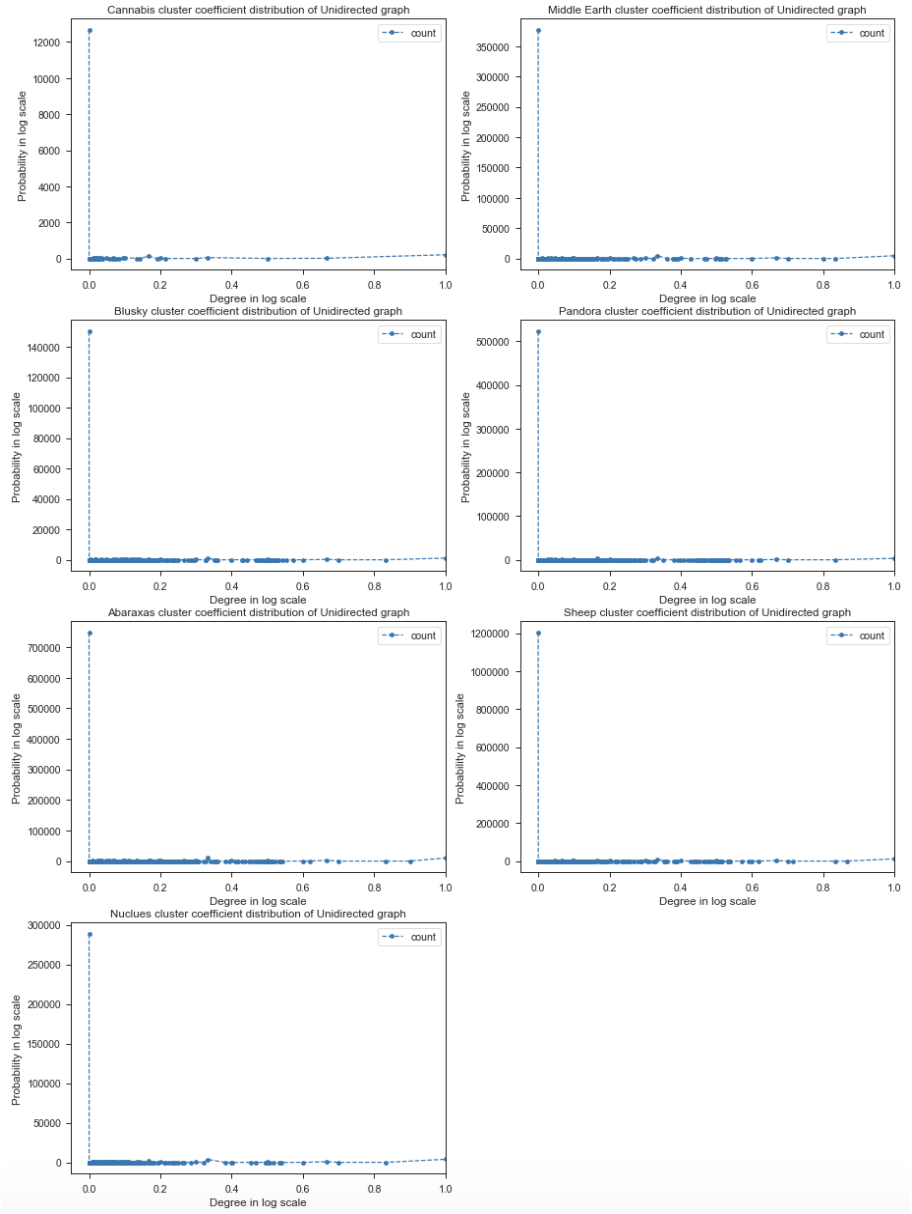


Figure A.13: Clustering Coefficient in Transaction Graph for each Darknet market

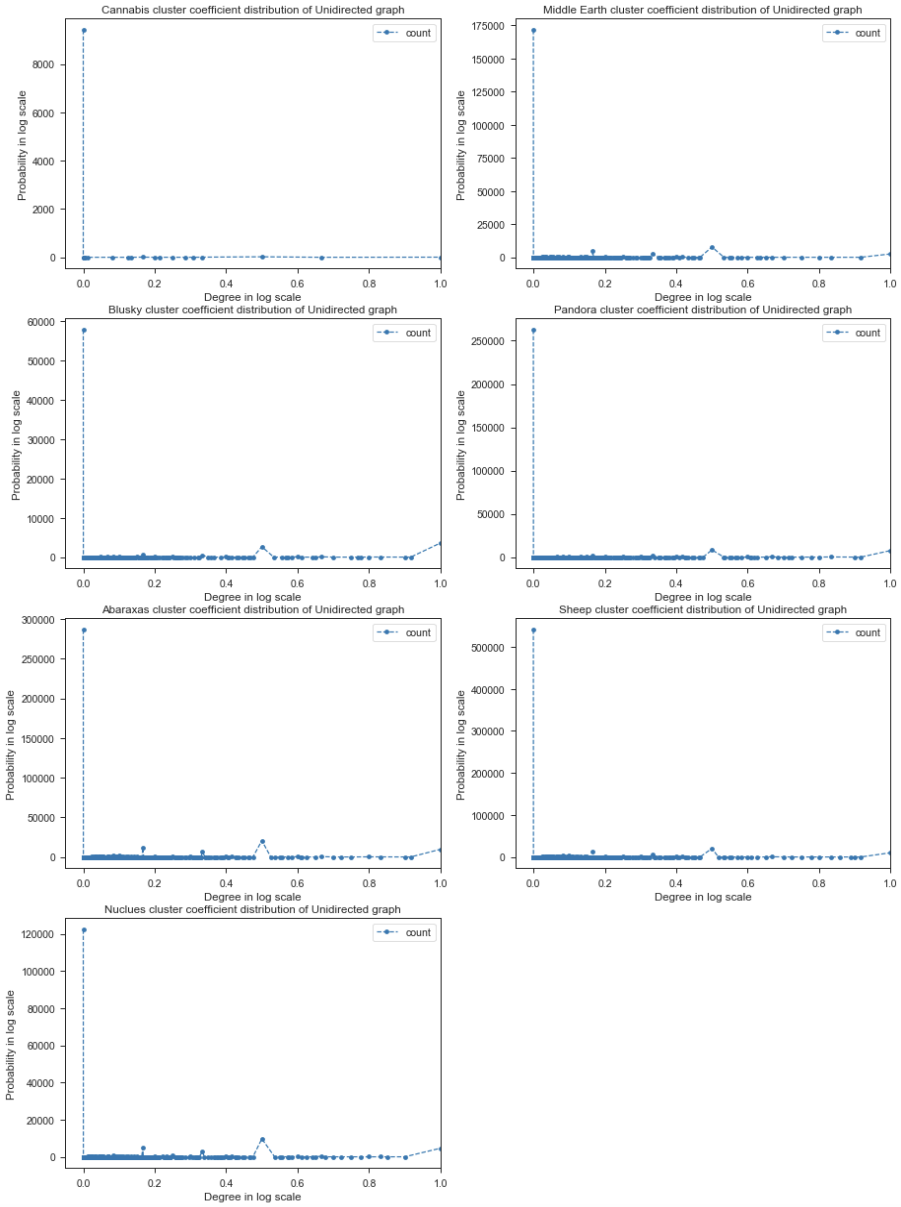


Figure A.14: Clustering Coefficient in User Graph for each Darknet market