

Thesis Dissertation

**STUDYING THE NETWORK STRUCTURE
OF THE BITCOIN TRANSACTION NETWORK**

Antonia Antoniou

UNIVERSITY OF CYPRUS



DEPARTMENT OF COMPUTER SCIENCE

May 2019

UNIVERSITY OF CYPRUS
DEPARTMENT OF COMPUTER SCIENCE

**Studying the Network Structure
of the Bitcoin Transaction Network**

Antonia Antoniou

Advisor:
Associate Professor Demetris Zeinalipour

Thesis submitted in partial fulfilment of the requirements for the award of
degree of Bachelor in Computer Science at University of Cyprus.

May 2019

Acknowledgments

I would like to express my sincerest thanks to my research supervisor, Dr. Demetris Zeinalipour for the excellent cooperation we had throughout this Individual Diploma Thesis. Without his assistance and guidance in every step of the process, I would not be able to complete this work.

Furthermore, I would like to express my thanks to my friends and family, who showed me their love and constant encouragement and helped me get through another chapter of my life.

Abstract

Bitcoin is a digital currency that is using the Blockchain Technology in order to allow people all over the world to transact with each other without involving any central bank. Bitcoin Technology and more specifically blockchain is becoming very popular these last few years. It has drawn the attention of many scientists, who are conducting many researches that concern its benefits and drawbacks, as well as what it has to offer as an alternative to some existing technologies.

This Thesis Dissertation is a re-visit of a prior study called “*Analyzing the Bitcoin Network: The First Four Years*”, which was aiming to get some insights on the evolution of the Bitcoin economy. The findings of that work were based on a period from 03.01.2009 until 10.04.2013 and they concerned the business and network perspective of the Bitcoin. This study’s datasets are based on a period from 03.01.2009 until 16.07.2014 and our goal is to observe the progress of the Bitcoin economy and network based on the comparative results regarding the transactions of the network. For the construction of our datasets we used the Bitcoin Core (also known as Bitcoin Client), two bash scripts that were using the API from the website “*Blockchain.com*” and a java program that was using the API from the website “*ip-api.com*”. For the management and the usage of those datasets a MariaDB database was used that was provided from the XAMPP tool.

Through this work and our findings, we concluded that although in a period of just fifteen months the Bitcoin Transaction Network has experienced a significant growth, most of the transactions had small Bitcoin values, which was the same finding as the prior study. Regarding the geographical characteristics of the network, we managed to extract some results based on the 69.32% of the total number of transactions, since the 30.68% of them were tagged with the IP address “127.0.0.1” (meaning they were initiated by the website “*Blockchain.com*”) or the IP address “0.0.0.0” (meaning they could not be linked to any IP address). Those results showed us that like the prior findings, the major markets of the network were the United States of America and Europe, more specifically Germany. Finally, we studied the Bitcoin Transaction Network and its relationship to Cyprus, which led us to the conclusion that Cyprus was participating in the network, but not as actively as we initially thought.

Contents

Chapter 1: Introduction	1
1.1 Motivation	1
1.2 Thesis Overview.....	2
1.3 Thesis Contribution.....	4
1.4 Thesis Outline	5
Chapter 2: Background.....	7
2.1 Peer-to-Peer Networks	7
2.2 Public-Key Cryptography	8
2.2.1 Encryption and Decryption.....	8
2.2.2 Digital Signing.....	9
2.3 Bitcoin	10
2.3.1 Bitcoin Technology.....	12
Chapter 3: Related Work.....	15
3.1 Blockchain Technologies and Applications.....	15
3.2 Bitcoin Data Available Online	16
3.3 Study: “Analyzing the Bitcoin Network: The First Four Years”	20
Chapter 4: Research Methodology.....	23
4.1 Architecture.....	23
4.1.1 Data Collection	23
4.1.2 Data Management and ER Diagram	27
4.2 Datasets	28
4.2.1 Bitcoin Transaction Data	28
4.2.2 Transaction’s IP Address	30
4.2.3 Transaction’s IP Geo-location	31
4.3 Research Questions Preview	31
Chapter 5: General Characteristics of the Bitcoin Transaction Network.....	33
5.1 Blockchain Size and Total Number of Blocks.....	34
5.2 Total Number of Transactions and Total Transacted Value in BTC	35
5.3 Bitcoin Growth.....	36
5.4 Bitcoin Market Price	38

5.5	Discussion	41
Chapter 6: Geographical Characteristics of the Bitcoin Transaction Network		42
6.1	Number of Distinct IP Addresses and their Geographical Representation	45
6.2	Number of Executed Transactions per Country	47
6.3	Discussion	49
Chapter 7: The Bitcoin Transaction Network with Regards to Cyprus.....		50
7.1	Cyprus Economy	50
7.2	Number of Distinct IP Addresses in Cyprus and their Geographical Representation	51
7.3	Number of Transactions Initiated from Cyprus	53
7.4	Discussion	54
Chapter 8: Conclusion.....		56
8.1	Conclusions	56
8.2	Limitations	57
8.3	Future Work	58
References.....		59
Appendix A.....		A-1

List of Figures

Figure 2.1	The design of a Peer-to-Peer Network, where nodes can directly	8
Figure 2.2	The process of encrypting and decrypting data	9
Figure 2.3	The process of digitally sign data using Public-Key Cryptography [2]	10
Figure 3.1	BlockCypher Website [28].....	17
Figure 3.2	SoChain Website [43]	17
Figure 3.3	CryptoID Website [30].....	18
Figure 3.4	BTC Website [29]	19
Figure 3.5	Blockchain Website [27].....	19
Figure 3.6	Final Data Model of M. Lischke and B. Fabian’s study [13]	20
Figure 3.7	Blockchain.info API Request Format for a Single Transaction [27]....	22
Figure 4.1	Data Structure of Bitcoin Blocks [25].....	24
Figure 4.2	ER Diagram.....	28
Figure 4.3	Transaction Data Fields	29
Figure 4.4	Input Data Fields	29
Figure 4.5	Output Data Fields	30
Figure 4.6	IP Address of the Transaction Initiator	30
Figure 4.7	IP Geo-location of the Transaction Initiator	31
Figure 5.1	Number of Total Transactions (2009-2014)	38
Figure 5.2	Bitcoin Market Price (USD) – Graph [35].....	40
Figure 6.1	IP nodes distributed on the world map [13].....	42
Figure 6.2	Development of particular Countries over Time [13].....	43
Figure 6.3	Transacted Bitcoin Value per Business Category per Country [13].....	44
Figure 6.4	IP Address Geo-location - Global Distribution.....	46
Figure 6.5	The 8 Countries with the biggest number of distinct IP addresses	47
Figure 6.6	Percentage of Distinct IP Addresses per Country	48
Figure 6.7	Percentage of Executed Transactions per Country	48
Figure 7.1	IP Address Geo-location - Cyprus Distribution.....	52
Figure 7.2	IP Addresses in Cyprus that participated in the Bitcoin Transaction Network.....	52

Figure 7.3 Number of Distinct IP Addresses for Countries Equivalent to Cyprus 53

Figure 7.4 Percentages for the Total Number of Executed Bitcoin Transactions for
Small Countries of the Network..... 54

List of Tables

Table 4.1	Data Structure of Bitcoin Blocks and Data Field Description [25]	27
Table 5.1	Bitcoin Network Statistics on Daily Basis [13]	34
Table 5.2	Number of Blocks, Size of Bitcoin Blockchain, Comparison	35
Table 5.3	Total Number of Transactions and Total Transacted Value in BTC	36
Table 5.4	Values used for the graph.....	37
Table 5.5	Some Bitcoin Market Prices (USD) - Price Table [35]	39
Table 6.1	Transactions tagged with different IPs (percentage of the total transactions contained in each dataset)	45

Chapter 1

Introduction

1.1	Motivation	1
1.2	Thesis Overview	2
1.3	Thesis Contribution	4
1.4	Thesis Outline	5

1.1 Motivation

Blockchain Technology is becoming very popular these last few years. It has drawn the attention of many scientists, who are conducting many and different researches that concern its possible uses, its benefits and drawbacks, as well as what it has to offer as an alternative to some existing technologies.

Upon the agreement between me and my supervisor that this Thesis would concern the “Blockchain Database Application”, my next step was to conduct a literature study. Through this study, I developed a better understanding of blockchain technology, its current uses and its possibilities. We discussed several possibilities of what we could do, and we concluded that we would like to analyse the blockchain technology from its network point of view. That led us decide to analyse the Bitcoin network, since it was operating for more than 9 years.

After our decision, we started looking for prior works and we found an article that was analysing the Bitcoin network from the first day of its operation until the April of 2013. With the knowledge that the Bitcoin network was a lot more popular after 2013 and there was not any study to analyse the network afterwards, we decided to re-visit this specific study and expand the analysed period in order to observe the evolution of the

Bitcoin network and come to some conclusions based on the comparison of the two studies.

1.2 Thesis Overview

Bitcoin is a digital currency also known as cryptocurrency, which is using the Peer-to-Peer Technology, the Public Key Cryptography and the Blockchain Technology in order to enable people all over the globe to transact with each other without the need of any central bank. It is becoming very popular these last few years, and to be more accurate the Blockchain Technology in general can be characterized as a hot topic. It has become the reason to initiate and conduct many and different researches that concern its possible uses, its benefits and drawbacks, as well as what it has to offer as an alternative to existing technologies. Many of those researches are focusing on the Bitcoin Network in order to analyse its security or maybe its involvement in illegal transactions, but also lots of those researches are aiming to develop computer software that will exploit the blockchain technology for purposes not related to cryptocurrencies and that it will be used all around the world in order to facilitate our lives.

One of the prior works that have been conducted about Bitcoin is called “*Analyzing the Bitcoin Network: The First Four Years*” and has been carried out by Matthias Lischke and Benjamin Fabian from the Institute of Information Systems, Humboldt University of Berlin. Their study was focused on the analysis of the Bitcoin Transaction Network by studying its economy and its network perspective from the 3rd of January 2009 until the 10th of April 2013.

Their results about Bitcoin Economy showed them that the biggest Bitcoin markets are located in the United States and Germany, while Bitcoin is mostly used within countries that are well developed and have good infrastructure. Moreover, they could extract some interesting results about the businesses distribution per country and they also managed to match the behaviour of some countries within the network with some events that happened during the specific period they were studying. Finally, their results about the network of the Bitcoin helped them to make several observations. Based on the degree distribution and power law over time they concluded that from 2010 the Bitcoin could be characterized as a scale-free network and based on the degree centrality on subsets of

their initial dataset they could identify the major hubs of the network. Finally, by studying the average clustering coefficient they observed the existence of the small world phenomenon in the Bitcoin Network as well as on a country level.

This Thesis Dissertation is based on that research and it is aiming in the analysis of the evolution of the Bitcoin Transaction Network. Our study is focusing to the Bitcoin Transaction Network from the 3rd of January 2009 until the 16th of July 2014, meaning that we expanded the initial dataset by fifteen months. In order to collect the information we needed and create our datasets, we used the Bitcoin Core also known as Bitcoin Client, the Application Programming Interface (API) from the website “*Blockchain.com*” to collect the Bitcoin Transaction Data and the IP address of the initiator of each transaction and the API from the website “*ip-api.com*” in order to convert the IP addresses to geo-location details. Once we collected all the data, we used a management tool called XAMPP, which provided us a MariaDB database.

Based on our datasets, we proceeded with the analysis of the network structure of the Bitcoin Transaction Network. In order to achieve that we set three research question, which are studied and answered within this Thesis. Our first research question is focusing on the characterization of the network, which is depending on some general metrics such as the size of the blockchain and the total number of blocks and executed transactions. Our findings concerning this question is that in a period of just fifteen months the Bitcoin Transaction Network had a significant growth regarding the executed transactions within the network, but regardless this growth we were able to observe that the total transacted Bitcoin value was low, which means that most of the transactions had small values, exactly like the prior findings.

Moving on to our second research question, we are focusing on the geographical distribution of all the transactions and distinct IPs of the network. Although the number of transactions carried out in the network is over 37 million, the number of distinct IP addresses we extracted is just 230,472. The datasets we used had transactions which we had to exclude from our study and that is because either they had been executed from the website “*Blockchain.com*” and their IP address was tagged as “127.0.0.1” or their IP address was not known and they were tagged as “0.0.0.0”. Those transactions were the 30.68% of our dataset, so the remaining 69.32% was used to extract our results. Our

analysis led us to the conclusion that like the prior findings the distinct IP addresses and generally lots of transactions were mainly originated from the United States of America and Europe and more specifically from Germany, while countries like China and Russia were also associated with a significant amount of transactions.

Finally, our third and last research question is focusing on the Bitcoin Transaction Network specifically associated with Cyprus. The main reason for our choice is because the geographical position of the island as a crossroad of three continents makes it a very popular financial and shipping hub. Our findings showed us that there were 82 distinct IP addresses that were originated from Cyprus and that 0.0068% of all the transactions within our datasets were initiated from Cyprus. Although Cyprus has participated in the network, this participation was not as big as we initially thought and by comparing Cyprus and Luxembourg, we can understand that Cyprus is not very interested in participating in the Bitcoin Transaction Network, at least for the period that we are analysing.

1.3 Thesis Contribution

This Thesis Dissertation will have a positive impact to any future work that may be done regarding the analysis of the Bitcoin Transaction Data. As the study that has been conducted before us was the motivation and the basis of this Thesis, so this Thesis could be the motivation for the next researcher that would like to observe the evolution of Bitcoin network in the following years. The contributions of this Thesis are the following:

- a. Datasets:** In order to complete this work, we had to create three datasets that would help us analyse the evolution of the Bitcoin Transaction Network. So, these datasets are available online for future use through GitHub under the repository "*dmsl/bitcoin*".
- b. Evolution of Bitcoin:** Based on the findings of this Thesis, the findings of the prior study and the comparison between them, we concluded that the Bitcoin Transaction Network have been significantly grown in a period of just fifteen months. So, this can lead us to the assumption that the growth of the network will be even more noticeable in the years that follow, and that will give us a better understanding of the network.

- c. Relationship between Bitcoin and Cyprus:** As a result of our analysis, we came to the conclusion that Cyprus cannot be characterized as an active member of the Bitcoin Transaction Network. However, our results could be used for comparison with future findings that would like to observe and analyse the relationship between them for a larger period of time.

1.4 Thesis Outline

This particular Thesis Dissertation begins with a small introduction in Chapter 1, which explains to the reader the motivation to carry out this work, its impact to future works and its general structure. In addition, it briefly explains the concept and the contents for each one of the chapters that is going to follow.

Moving on to Chapter 2, the “Background”, the reader will have the chance to get familiarized with some terms and technologies associated with Bitcoin, as well as the Bitcoin Technology itself. More specifically, in this section we are going to explain what the Peer-to-Peer Network and the Public-Key Cryptography are. In order to close the Chapter, we are going to describe what exactly Bitcoin is, while we are discussing some of its characteristics and giving definitions to some concepts closely associated with it.

Subsequently, in Chapter 3 we will review some prior researches and studies regarding the Blockchain Technologies and then we will continue mentioning a few pages available on the Internet that can provide blockchain data, either from Bitcoin either from other blockchain systems. Finally, we will conclude the Chapter by describing the concept of the prior study, which this Thesis is based on, and by explaining exactly which parts of that study we will re-visit.

In Chapter 4 the reader will be able to find all the relevant information regarding the datasets that have been used to carry out this study, as well as all the information that concern the tools, websites and the management technique that have been used to complete this work.

Furthermore, Chapter 5, Chapter 6 and Chapter 7 are consisted by an extended analysis of the Bitcoin Transaction Data from the 3rd of January 2009 until the 16th of July

2014. Chapter 5 is the “General Characteristics of the Bitcoin Transaction Network”, which refers to some general Bitcoin characteristics and presents the prior findings, the findings of the current study, a small comparison between the two results and a small discussion about our results and conclusions. Chapter 6 is the “Geographical Characteristics of the Bitcoin Transaction Network”, which refers to the geographical characteristics of Bitcoin and is consisted again from the prior findings, the findings of this study, a small comparison of the two results and a discussion about the results and our conclusions. Lastly, Chapter 7 is the “The Bitcoin Transaction Network with Regards to Cyprus”, which is discussing some geographical characteristics of Bitcoin with regards to our country.

Last but not least, in Chapter 8 the reader will be able to read a summary with all of the results, discussions and conclusions about the Bitcoin Transaction Network. Moreover, this final Chapter contains a small reference on the limitations and problems that we encountered during the conduct of this work and it closes with some suggestions about possible future work.

Chapter 2

Background

2.1	Peer-to-Peer Networks	7
2.2	Public-Key Cryptography	8
2.2.1	Encryption and Decryption	8
2.2.2	Digital Signing	9
2.3	Bitcoin	10
2.3.1	Bitcoin Technology	12

2.1 Peer-to-Peer Networks

Peer-to-Peer (P2P) Network [39][40] refers to a network among computers/nodes, where each node can directly communicate with other nodes inside the network, without involving any central server. In this type of networks, a node acts simultaneously as a server and as a client, meaning that it can send and receive files at the same time.

P2P Networks are available to anyone as long as they have an Internet connection and the necessary P2P software. Once someone become part of such a network, other network nodes can investigate his/her device for a specific file in order to transfer it to theirs, but typically only if the file is inside a folder that the owner of the device has chosen to share with the network.

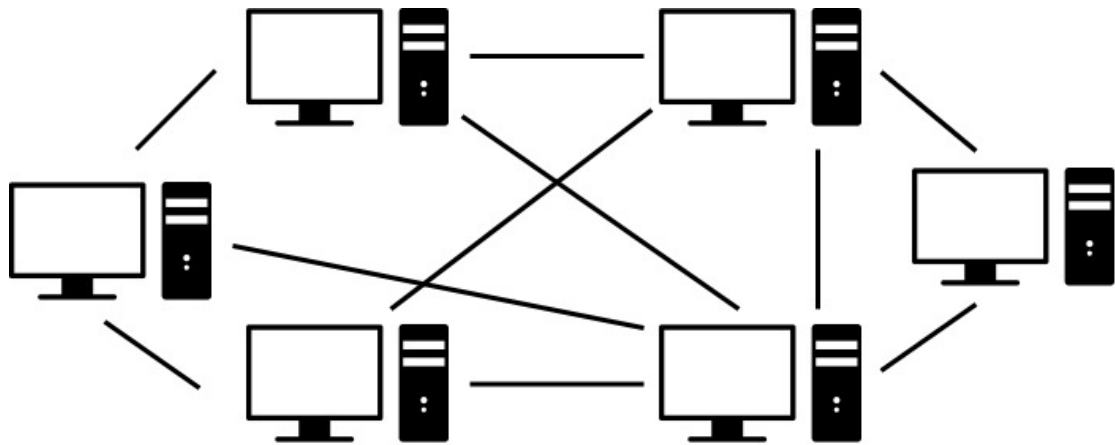


Figure 2.1 The design of a Peer-to-Peer Network, where nodes can directly communicate with each other without involving a central server [40]

2.2 Public-Key Cryptography

Public-Key Cryptography [2][41][42] is an encryption technique that belongs to the category of Asymmetric Ciphers and is also known as public-key encryption. The goal of public-key encryption is to ensure confidentiality, authenticity and non-repudiation regarding private communications between two parties over a network.

Public-Key Cryptography uses pairs of keys, which are generated with the use of expensive mathematical calculations and their size should be between 1024 and 3072 bits, in order to be considered computationally secure. A pair of keys is consisted by a public key, which is publicly known to everybody within the network and by a private key, which is kept a secret and the only person who has it, is the one who have generated it.

2.2.1 Encryption and Decryption

Due to the expensive mathematical operations that the asymmetric algorithms use, public-key cryptography is not the ideal choice for encrypting large amount of data. However, it is perfectly suitable for the encryption of small messages, for example the encryption of the key for a symmetric block cipher like AES [3], and for digital signatures.

Let's imagine that there are two persons, Actor A and Actor B, that want to exchange encrypted messages, using their personal pairs of keys. Actor A must encrypt his/her message using the public key that belongs to Actor B, before sending the message to Actor B through the network. Once Actor B receives the message, he/she will decrypt it using his/her private key and he/she will be able to read the original message from Actor A.

Basically, Figure 2.2 represents the process of exchanging messages, which were encrypted using Public-Key Cryptography. Generally, anyone that knows the public key of an entity/person, he/she can encrypt a message with the specific public key and only the owner of the corresponding private key will be able to decrypt and read the original message.

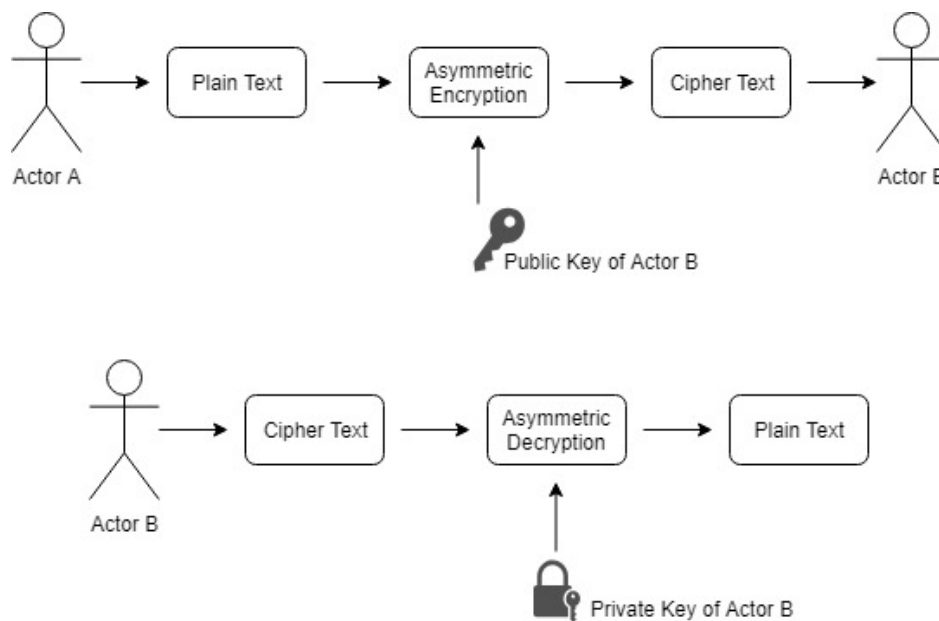


Figure 2.2 The process of encrypting and decrypting data using Public-Key Cryptography [2]

2.2.2 Digital Signing

The reverse process of what described above, also works. However, its usage it is not recommended for standard encryption, because if someone encrypts sensitive data

with his/her private key and send the data through the network, anyone that knows his/her public key will be able to decrypt and “stole” the data.

Nevertheless, encrypting messages with the private key is also useful, not because someone will prevent others from reading those messages, but because it is a way to verify that the messages arrived to their destination without any alteration and they were originated from the right person, the person that had access to the specific private key.

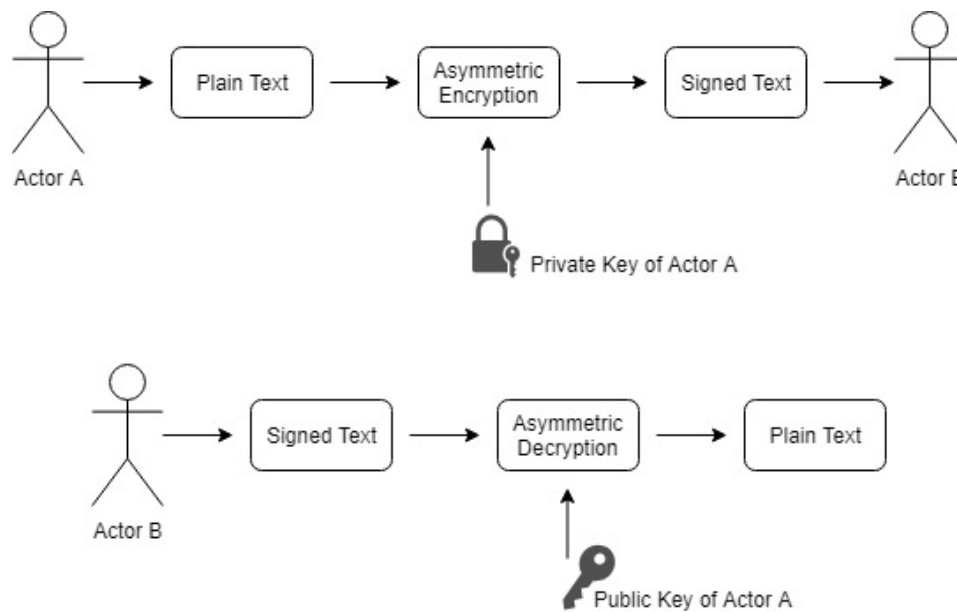


Figure 2.3 The process of digitally sign data using Public-Key Cryptography [2]

2.3 Bitcoin

Bitcoin [24][25][26] is a cryptocurrency (digital/electronic currency) that uses peer-to-peer technology and it does not depend on any authority or bank (decentralized).

Bitcoin was invented by Satoshi Nakamoto, whose identity remains a mystery until today, as an open-source software in 2009 and it is a software that anyone can download and participate to its network. *Bitcoin (BTC)* can be used either for products/services payment or someone could exchange BTC for other currencies such as USD or EUR.

Bitcoin transactions are verified by network nodes and then they are stored in the blockchain. The Bitcoin blockchain can be described as a public ledger book that stores inside all the transactions that have ever been made in the network. The blockchain, as the word describe by itself, is a chain of blocks interconnected and every block contains a hash value (based on all the transactions contained in the block) of its previous block. Furthermore, Bitcoin is designed in such way, that once a transaction is made and verified it cannot be altered easily, as if you want to change a block then you will have to modify all the following blocks. As the number of blocks inserted in the blockchain is increasing the difficulty of changing a block is increasing as well.

Every transaction in the network has to be validated by someone in the network. That someone is called miner. There are many miners in the Bitcoin network and their goal is to keep the blockchain consistent and updated by verifying transactions. All the Bitcoin nodes have their own copy of the blockchain, which is updated every 10 minutes, when a new block that contains numerous transactions inside it, is added to the blockchain by a miner and it is broadcasted towards the rest of the Bitcoin community.

Bitcoin is not like other currencies that exist all over the world. Since Bitcoin is not controlled by any central bank, bitcoins are created by the nodes participating in the network and they are given as a reward to miners. In order to be able to control the number of bitcoins created, there is a generation algorithm used by the network, which defines the rules that you must follow for bitcoin creation, as well as their creation rate. The mining reward started at 50 bitcoins and every 210,000 mined blocks that reward is decreased by half. Eventually the creation of new bitcoins will stop approximately by 2140, when the total number of bitcoins will reach 21 million, and then the only reward for miners will be the amount of BTC that will be collected as transaction fees.

Another important characteristic of the Bitcoin network is that it is specifically designed in order to provide its users a kind of anonymity. As mentioned earlier, anyone that wants to install the Bitcoin client and create his/her digital wallet can participate in the network and initiate or receive a transaction. A transaction can be made between any node in the network, but there are no personal details of the nodes stored anywhere. The only information about the nodes that is stored are the public keys of the source and the destination addresses. In this way, no one can really know the source and the destination

addresses except the 2 nodes involved in the transaction. Last but not least, each public key that exists in the network corresponds to a specific private key, that is stored in a digital wallet. In case that someone lose his/her private key, there is no way to recover it and all the bitcoins that are saved in his/her electronic wallet will be lost forever.

2.3.1 Bitcoin Technology

In the following few pages we will introduce and explain some very popular concepts that are associated with Bitcoin. We will start from the Units of Bitcoin, then we will continue with the description of the concepts of Decentralization, Blockchain, Transaction, Mining and Wallet.

All around the world people are using the term *Bitcoin* (฿) the same way they use the term *Euro* (€), but there are also two alternative units, that can be used to address smaller amounts. The first alternative unit is called *satoshi* (*sat*), which corresponds to 0.00000001 bitcoins, and the second alternative unit is called *millibitcoin* (*mBTC*), which corresponds to 100,000 satoshis or 0.001 bitcoins.

Moving on to our next concept we find the word Decentralization, which is very important when we are talking about Bitcoin. When someone says that the Bitcoin network is decentralized, the first thing that comes in mind, is that Bitcoin does not have a central server or that it does not depend on any central authority or bank. Beyond that, the decentralized structure of Bitcoin network also means that there is no central storage for the blockchain, and anyone can download and store it on his/her computer. Furthermore, there is no need of any approval before anybody can become part of the network, become a miner, create a Bitcoin address or participate in any transaction within the network. Finally, the creation of bitcoins itself is decentralized, meaning that the supply of bitcoins is limited and it is not fully available yet, due to the fact that bitcoins are created every time as a reward to the fastest miner and until a new block is broadcasted through the network, is impossible for anyone to know who is the miner who created it.

As mention earlier, all the Bitcoin data is stored in the Blockchain, which can be described as a public ledger report that contains all the transactions executed in the network. Its implementation is basically a linked list of blocks, and since a linked list is

very similar to a chain hence the word blockchain. Every block that exists or will be created and inserted in the blockchain contains one or more network transactions and the hash value of the previous block. This feature of the blockchain allows the users to maintain and validate the blockchain. To achieve that, a user needs to begin from a certain block and be able to travel back to the first Bitcoin block, which is also called the genesis block.

All the transactions of the network are consisted of one or more inputs and outputs. When a user wants to transfer bitcoins to somebody in the network, he/she must explicitly determine all the Bitcoin addresses that will receive the bitcoins and the exact number of bitcoins that will be sent to each address. Each address and the corresponding amount are defined in an output, while the bitcoins that will be transferred are defined in an input. The capability of a transaction to contain multiple inputs and outputs allows the user to send bitcoins to multiple Bitcoin addresses, including his/her own if not all the bitcoins from the inputs is needed. Last but not least, each output corresponds only to one input and that is how the Bitcoin network can detect if somebody is trying to double spend bitcoins.

Our next concept is Mining. The nodes of the network, that are trying to confirm the validity of transactions and then group them into a block are called miners. The main goal of mining is to maintain the consistency and the completion of the blockchain, as well as to make sure that nobody is altering the blockchain, but the mining process require lots of processing power. Once a miner creates a new block, the block must to be broadcasted to the rest of the network in order to be verified and accepted. The only way for a block to be verified and accepted by the network, is to contain a *proof-of-work (PoW)*.

The PoW system that is used in the Bitcoin, basically requires from the miner to detect a number, which is called nonce. The nonce found by the miner is considered correct only if you hash the contents of the particular block with it, and the result of that operation is smaller than a target number, which is called difficulty target. Every time that 2,016 blocks are generated, the difficulty target is auto adjusted based on the recent performance of the network, because Bitcoin aims to maintain the average time between the creation of two blocks at 10 minutes. As more and more blocks are mined by the day,

the difficulty target constantly increases and as a result mining requires huge amounts of processing power.

Finally, a Wallet can be described as a way for the user to store his/her private keys, which they provide access to the user to manage his/her digital Bitcoin assets. There are many different categories of wallets, such as wallets owned by Bitcoin users themselves, online wallets provided by a third-party and hardware wallet.

Wallets owned by Bitcoin users can be separated to full clients and lightweight clients. In the case of a full client, a full copy of the Bitcoin blockchain is downloaded to the user's device and the verification of the transactions is made explicitly by the client. It is the most secure and reliable way in order to use the Bitcoin network, but due to the size of the client and the complexity that is required to verify the whole blockchain, it is not suitable for some devices. On the other hand, a lightweight client means that the client depends on other full clients, in order to send and receive transactions, which does not require a full copy of the blockchain stored in the user's device and makes it suitable for devices that are using low power and low bandwidth, such as smartphones. The drawback of having a lightweight client, is that absolute trust is required from the user towards the full client, as it may report wrong values to the user and in addition, the client does not verify the transactions, which implies that the user must trust the miners for the validation of the entire blockchain.

Wallets provided by third parties, are providing similar functionalities as the wallets owned by the user, but their use is easier. In this case, the private keys of the user, is stored in the servers of the wallet provider and requires the user to completely trust the third party, because if a provider is malicious or the servers' security measures are not powerful enough, this may lead to the stealing of the user's bitcoins.

Hardware wallets are another type of wallets, which store the user's private keys offline, in order to avoid their potential theft, while help the user transact in the network easier.

Chapter 3

Related Work

3.1	Blockchain Technologies and Applications	15
3.2	Bitcoin Data Available Online	16
3.3	Study: “ <i>Analyzing the Bitcoin Network: The First Four Years</i> ” .	20

Blockchain Technology is a rather new and popular research field and is drawing a lot of attention from the research community and it has been studied extensively in the last few years. These researches are focusing in many aspects related to blockchain, as well as to lots of applications of the blockchain technology.

In the following section we will shortly review some related work, which concerns the application of blockchain technology in every-day aspects of life, such as education and healthcare, as well as some related work concerning *Internet of Things (IoTs)*. Subsequently, we will discuss about blockchain technology and its use in cryptocurrencies and we will close this chapter with some Bitcoin related work.

3.1 Blockchain Technologies and Applications

As mentioned earlier, blockchain technology is extensively studied by the research community [1][10][14][15], which is hoping that blockchain could be used in order to improve some aspects of our lives [21], but also tries to study its security aspect and its involvement in illegal operations [7].

For example, a recent study is focusing in the use of the blockchain technology in education [17]. In this study, researches propose a solution that will connect all learning data from various Learning Management Systems, Learning Record Stores, institutions and organizations. Their objective is to overcome the limitations that exist with the

current system, concerning the slow onboarding process when students are moving between learning environments, which is caused by the difficulty to access the students' learning history.

On the other hand, another recent study is focusing in the use of blockchain technology in order to improve the healthcare [4]. The objective of researchers in this study, is to ensure the privacy and security of patients' data with the use of a private blockchain, while the advantages on the medical research field and the patients' treatment will significantly increase, due to the sharing of information among hospitals, clinics, patients, providers and insurance companies.

Moreover, there are many studies regarding the use of blockchain technology in the field of IoTs [6][8][9][18][19] and Smart Cities [11][16][20]. Researchers think that the blockchain as a data management tool is the key to solve a fair amount of the challenges that are attached to the IoTs. For example, they believe that blockchain is the solution to bridge the current different administrative domains, in order to create end-to-end trust and communication among IoTs devices of different domains without the need of a central authority. Furthermore, it is believed that many problems regarding the scalability and the security of IoTs will be eliminated, because there will be no single point of failure.

3.2 Bitcoin Data Available Online

As far as Bitcoin is concerned, it is a topic well debated and lots of people have studied it, whether they are researchers or not. There are many websites on the internet, which are providing information, and some of them are providing statistics and charts as well, about Bitcoin blocks, transactions and Bitcoin addresses. Let's discuss about some of these websites and what they can provide to the user.

- **BlockCypher:** It provides Bitcoin data, such as block information, transaction information and Bitcoin address information. In addition, it can offer to the user data about other blockchains such as Litecoin and Dogecoin. It provides to the user a free of charge API, that can be used with Ruby, Python, Java, PHP, Go, Node.js and command line. However, their free of charge API is limited to 3 requests per second

and 200 requests per hour. If the user would want a higher limit, then a payment is required.

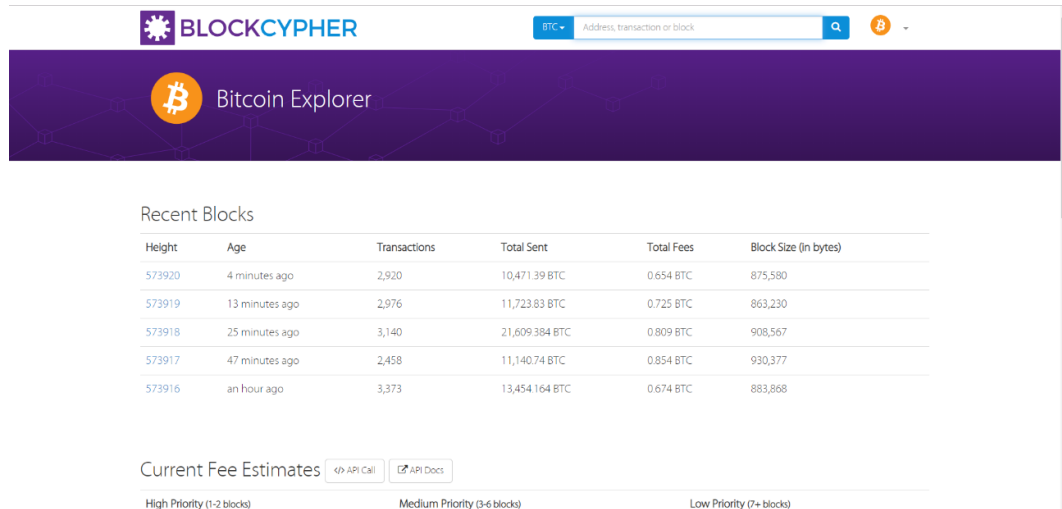


Figure 3.1 BlockCypher Website [28]

- **SoChain:** It provides data about Bitcoin, Dogecoin, Litecoin and other blockchains in real time, such as block information, transaction information and Bitcoin address information. It provides to the user a free of charge API, that allows 5 requests per second and if the user would want a higher limit, then he/she would have to contact the provider.

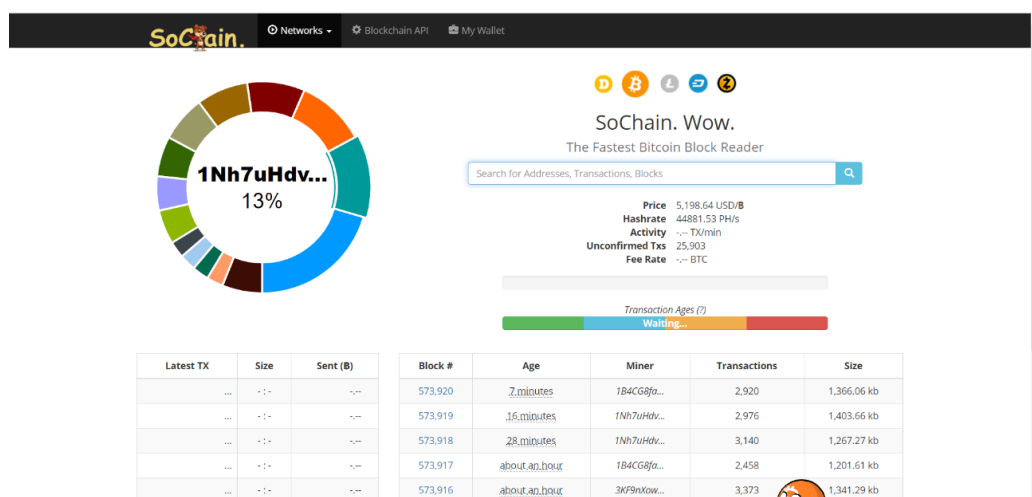


Figure 3.2 SoChain Website [43]

- **CryptoID:** It provides data about Bitcoin and many other blockchains, such as block information, transaction information and Bitcoin address information. It provides to the user a free of charge API, that allows 1 request per 10 seconds. An API key could be requested for free by the user, in order to eliminate some of the limitations, if it is necessary.

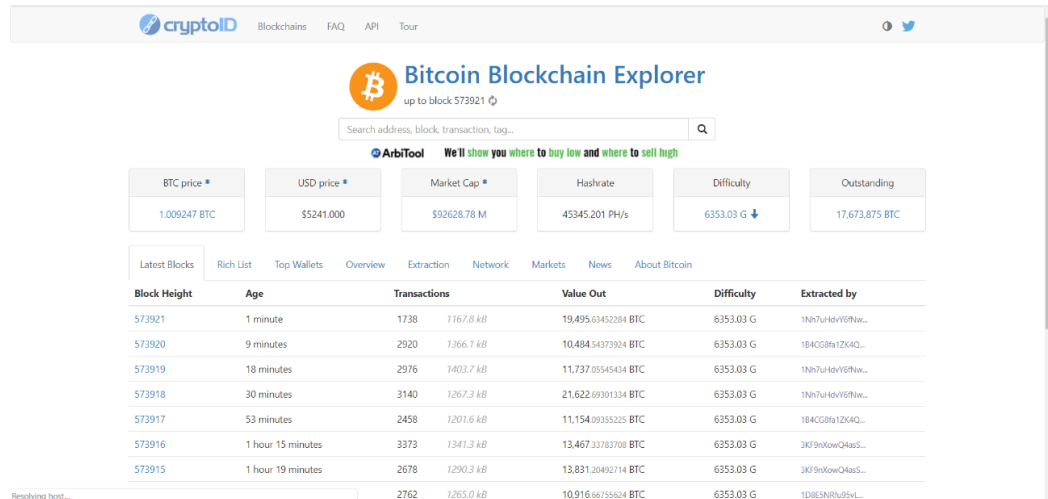
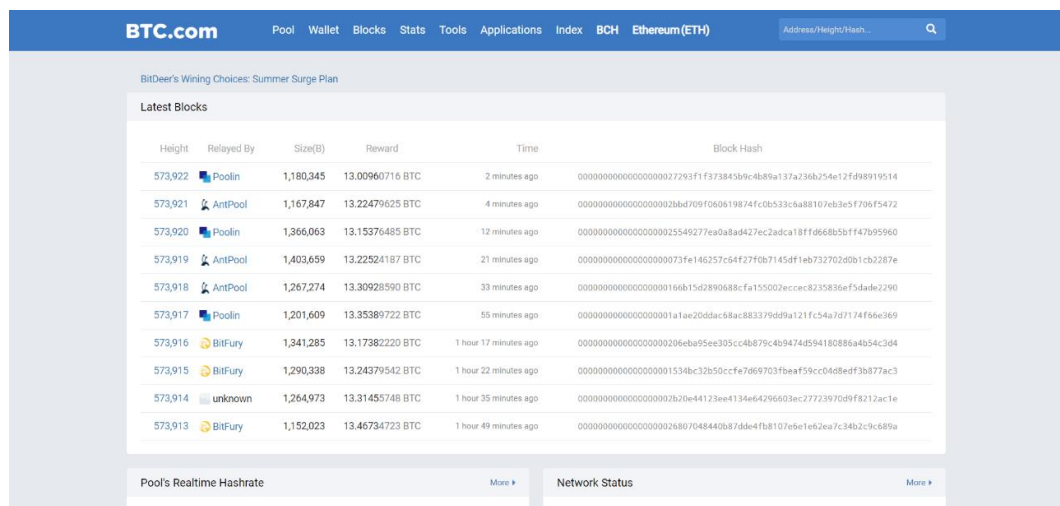
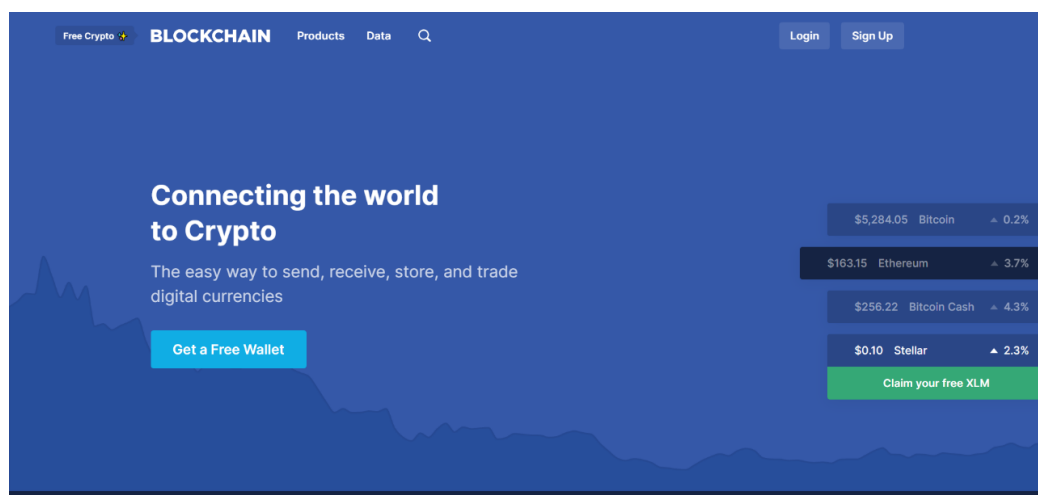


Figure 3.3 Bitcoin Blockchain Explorer Website [30]

- **BTC:** It provides data about Bitcoin, Bitcoin Cash and Ethereum blockchains, such as block information, transaction information, Bitcoin address information and information for some Bitcoin mining pools. It provides to the user a free of charge API, that allows maximum 50 batch API requests.



- **Blockchain:** It provides services and products such as online wallet, strategies for investors, lockbox (hardware wallet) and an API. In addition, it provides data and charts about Bitcoin, Bitcoin Cash and Ethereum blockchains, such as block information, transaction information, Bitcoin address information, market price and information for some Bitcoin mining pools. It provides to the user a free of charge API, that however have some limitation to the number of requests per second. An API key could be requested by the user, in order to eliminate some of the limitations, and the cost that the user has to pay is relative to the number of requests required by the user.



3.3 Study: “Analyzing the Bitcoin Network: The First Four Years”

Our study is a revisit of the study “Analyzing the Bitcoin Network: The First Four Years” by Matthias Lischke and Benjamin Fabian [13] from the Institute of Information Systems, Humboldt University of Berlin. Their study was aiming to examine the economy and the transaction network of Bitcoin, in the first four years of its operation. In their article they explained that the reason behind choosing the specific period was to reduce the amount of data that they needed to analyse, as well as to offer the opportunity to someone else after them to analyse a different period and compare the results of the two studies.

They started their study by explaining what the Bitcoin Technology is and how it works. Subsequently, they discussed some works that were related to their subject and of course they explained the methods that they used to complete their work. By telling that, we mean that they explained how they collected and managed their data, what tools they used, as well as the different network metrics that they used to analyse the Bitcoin network.

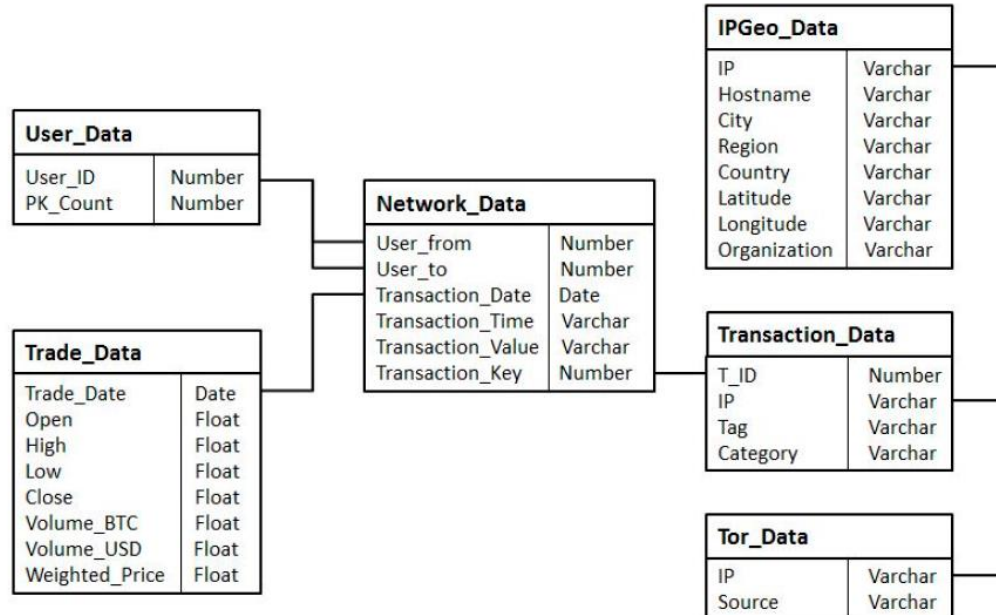


Figure 3.6 Final Data Model of M. Lischke and B. Fabian’s study [13]

Before they started to present the results of their study, they illustrated the structure and the contents of their datasets and then they presented their analysis, which was separated in two parts. The first part of their analysis was business-related, and the second part concerned the network structure and characteristics of Bitcoin. Finally, they concluded their study by pointing out some limitations their research had and by suggesting what could be analysed in future works.

Our study although is a revisit of M. Lischke and B. Fabian's work, does not cover both parts of their analysis, but it is just focusing on the business-related part, in order to compare the results only on that aspect.

The business-related part of the prior research contains three different areas of interest. The first one is some characteristics and statistics of Bitcoin network, such as the number of executed transactions, the size of the blockchain or changes of the Bitcoin market price. In order to extract those results, they used only the transaction data that were collected from the publicly available Bitcoin client.

The second area that was investigated is the Bitcoin's business statistics. For that they needed the transaction data, as well as the business tag of a transaction. In order to obtain the business tag of each transaction they used a JSON API provided by the website "Blockchain.info". Due to the limit rate of the website they needed more than 40 days to collect all the data that was necessary for this analysis and even then they did not have a business tag for all of the transactions, because the business tag was only a voluntary field that could be completed by the initiator of a transaction and not all of them completed that particular field.

The third and last area of their business-related analysis is the geographic characteristics of the Bitcoin. In order to analyse this part, they needed the transaction data like before, the IP address that was associated with each transaction and the geo-location information that was extracted from the IP addresses. In order to collect all the data, they used the website "*Blockchain.info*", which with some techniques introduced by Kaminsky [12] provided them with high probability the IP address of the initiator of each transaction, if it was known and also the website "*ipinfo.io*" [37], which could provide them all the relevant geo-location information of an IP.

Single Transaction

- [https://blockchain.info/rawtx/\\$tx_hash](https://blockchain.info/rawtx/$tx_hash)
- You can also request the transaction to return in binary form (Hex encoded) using ?format=hex

```
{
  "hash": "b6f6991d03df0e2e04da99fcd6bc418aac66049e2cd74b80f14ac86db1e3f0da",
  "ver": 1,
  "vin_sz": 1,
  "vout_sz": 2,
  "lock_time": "Unavailable",
  "size": 258,
  "relayed_by": "64.179.201.80",
  "block_height": 12200,
  "tx_index": "12563028",
  "inputs": [
    {
      "prev_out": {
        "hash": "a3e2bcc9a5f776112497a32b05f4b9e5b2405ed9",
        "value": "100000000",
        "tx_index": "12554260",
        "n": "2"
      },
      "script": "76a914641ad5051edd97029a003fe9efb29359fcee409d88ac"
    }
  ],
  "out": [
    {
      "value": "98000000",
      "hash": "29d6a3540acf0a0950bef2bdc75cd51c24390fd",
      "script": "76a914641ad5051edd97029a003fe9efb29359fcee409d88ac"
    },
    {
      "value": "2000000",
      "hash": "17b5038a413f5c5ee288caa64cfab35a0c01914e",
      "script": "76a914641ad5051edd97029a003fe9efb29359fcee409d88ac"
    }
  ]
}
```

Figure 3.7 Blockchain.info API Request Format for a Single Transaction [27]

Chapter 4

Research Methodology

4.1	Architecture	23
4.1.1	Data Collection	23
4.1.2	Data Management and ER Diagram	27
4.2	Datasets	28
4.2.1	Bitcoin Transaction Data	28
4.2.2	Transaction's IP Address	30
4.2.3	Transaction's IP Geo-location	31
4.3	Research Questions Preview	31

4.1 Architecture

Once we decided which kind of data we needed to collect, we needed to choose a way to download them and a tool that would help us manage the enormous amount of data that we would collect.

4.1.1 Data Collection

In order to gather the Bitcoin Transaction Data, we needed access to the entire Bitcoin blockchain, so we downloaded the Bitcoin Core [23]. Once somebody install the Bitcoin Core in his/her computer, some binary files are being downloaded locally, which some of them, named like "*blk*.dat*", are containing the Bitcoin blockchain. Figure 4.6 represents the block structure of every bitcoin block.

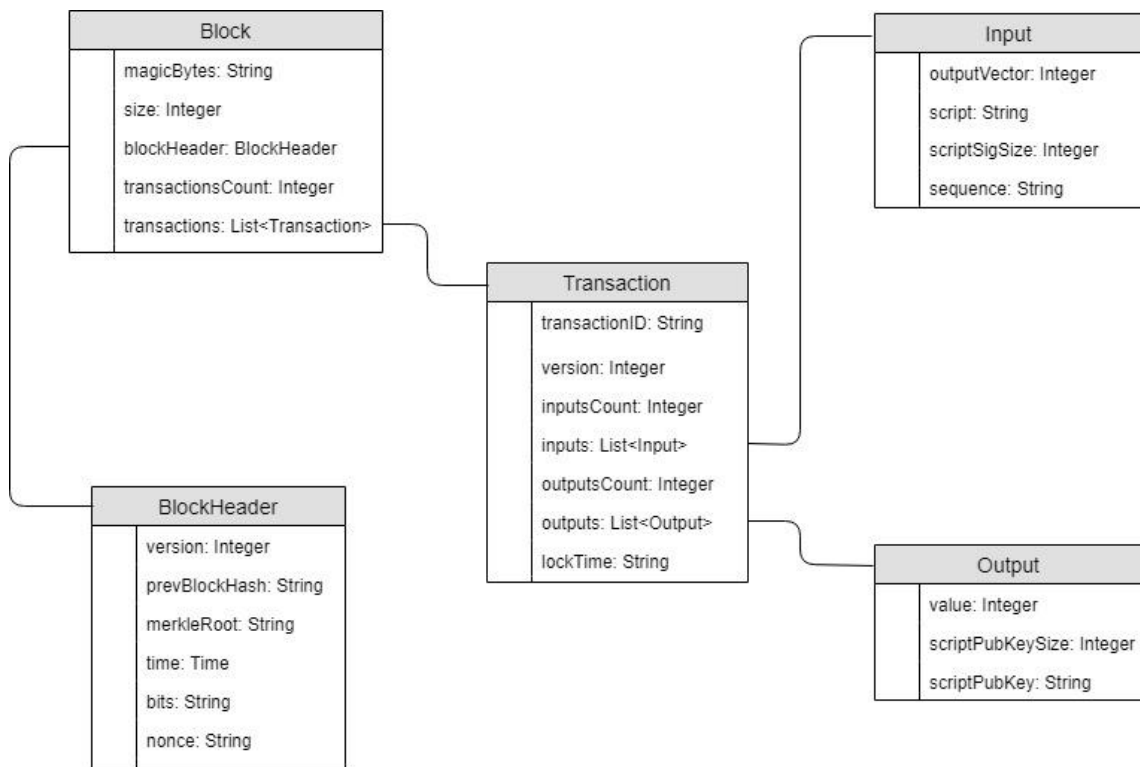


Figure 4.1 Data Structure of Bitcoin Blocks [25]

Each binary file contains several blocks and all the information is stored in hexadecimal format. However, if someone wanted to separate the blocks from a file, it would not be very difficult, due to the fact that each block begins with a standard sequence of 4 bytes, “f9beb4d9” (magic bytes). Table 4.1 explains exactly the meaning of each field.

Block		
Magic Bytes	4 bytes	“f9beb4d9”, a way to identify a new block in a blk*.dat file.
Size	4 bytes	The number of bytes following until the end of the block. (Stored in Little-endian)
Block Header	80 bytes	The header of the block.

Transactions Count	VarInt (1/3/5/9 bytes)	The number of transactions included in the specific block. If the first byte is smaller than the byte “fd” then that byte is the transaction counter else if the first byte is “fd” then the next 2 bytes represents the transaction counter. If the first byte is “fe” then the next 4 bytes is the transaction counter or else if the first byte is “ff” then the next 8 bytes are the transaction counter. (If the counter is bigger than one byte then the bytes that contain the counter is stored in Little-endian)
Transactions	Depends on Transactions Count	The data of the transactions contained in the block.
Block Header		
Version	4 bytes	The current version of the block. (Stored in Little-endian)
Previous Block Hash	32 bytes	A 256-bit hash value that connects the current block with the previous one.
Merkle Root	32 bytes	A 256-bit hash value of the current block, that is based on all transactions in the block.
Time	4 bytes	Timestamp in seconds. (Stored in Little-endian)
Bits	4 bytes	A smaller version of the target. (Stored in Little-endian)
Nonce	4 bytes	A number, that every time a miner tries to get a block hash below the target it increases. (Stored in Little-endian)
Transaction		
Transaction ID	32 bytes	A double hashed 256-bit value of the transaction. (Stored in Little-endian)
Version	4 bytes	The version of transaction data that is used. (Stored in Little-endian)

Inputs Count	VarInt (1/3/5/9 bytes)	The number of inputs included in the specific transaction.
Inputs	Depends on Inputs Count	The inputs that are going to be used in the transactions.
Outputs Count	VarInt (1/3/5/9 bytes)	The number of outputs included in the specific transaction.
Outputs	Depends on Outputs Count	The outputs that are going to be used in the transactions.
Locktime	4 bytes	The block height or a timestamp, when the transaction will be final. (Stored in Little-endian)
Input		
outputVector	4 bytes	The number of one of the outputs from a previous transaction. (Stored in Little-endian)
ScriptSig Size	VarInt (1/3/5/9 bytes)	The size of the following script in bytes.
ScriptSig	Depends on ScriptSig Size	The script that unlocks the previous output in order to be able to spend it. (Stored in Little-endian)
Sequence	4 bytes	Usually has the value “0xffffffff”.
Output		
Value	8 bytes	The number of Bitcoins that will be transferred. (Stored in Little-endian)
ScriptPubKey Size	VarInt (1/3/5/9 bytes)	The size of the following script in bytes.

ScriptPubKey	Depends on ScriptPubKey Size	The script that locks the output in order to be able to spend it later. (Stored in Little-endian)
--------------	------------------------------------	---

Table 4.1 Data Structure of Bitcoin Blocks and Data Field Description [25]

Since the data from the Bitcoin Core was downloaded in binary files, I tried to use the same Python2.7 data scrapper tool [5] as the prior study, but since the version of Bitcoin Core had changed the tool could not work. In order to overcome this particular problem, we decided to use the Blockchain.info website, which gave us the Bitcoin Transaction Data and the IP address of the initiator of the transaction in human readable form.

The data we needed to download was retrieved by the website using bash scripts that were created and they were using cURL commands towards the website. The output of those bash scripts were some files in the format of *Comma Separated Values (CSV)*. Unfortunately, due to the data limit that was set by the website administrator, the collection of the data took us approximately three months to be completed.

Finally, based on the IP addresses that were downloaded, we wanted to obtain the Geo-location Information of each transaction, so we used the “*IP Geolocation API*” [36], which provided us the Country Name, the Region Name, the City Name and the Latitude and Longitude of a given IP address.

4.1.2 Data Management and ER Diagram

Once we gathered all the data that was essential to our study, we wanted a tool to manage them. Our choice was to use the XAMPP environment [44], which is a free Apache distribution [22] and it provided us a MariaDB database (server version 10.1.36) [38]. Figure 4.7 shows the ER Diagram, that is used to structure our data within the database.

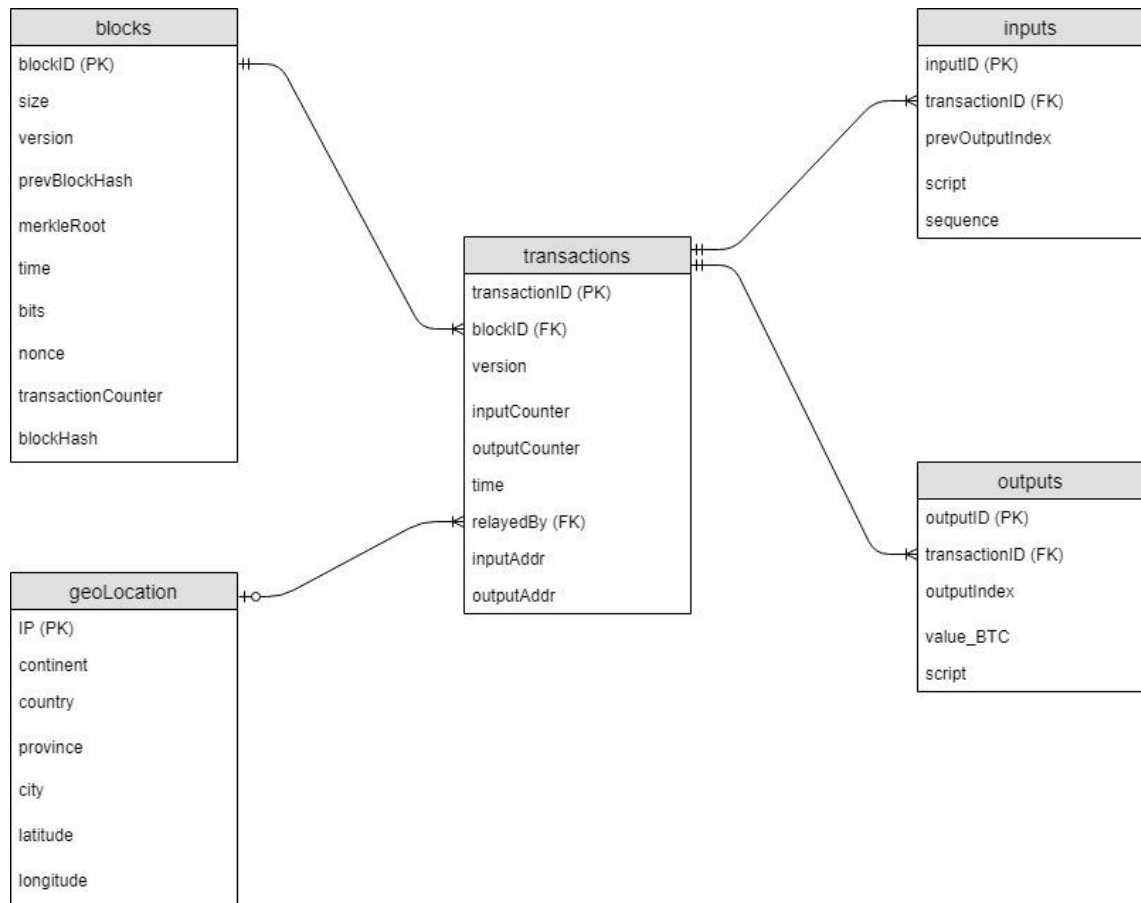


Figure 4.2 The ER Diagram that is used to structure and manage the datasets that are used in this Thesis Dissertation

4.2 Datasets

In order to be able to conduct this study, some data related to Bitcoin had to be collected. Based on all the data we collected, we separated them in datasets, which helped us analyse the Bitcoin transaction data and make some observations about the Bitcoin network.

4.2.1 Bitcoin Transaction Data

Firstly, we needed to collect the Bitcoin transaction data, which was the key in order to fulfil the entire analysis. Our study is focusing on the Bitcoin transaction data from the 3rd of January 2009 until the 16th of July 2014, which means we expanded the dataset for

fifteen months more than the prior study. The reason that forced us to expand our dataset only by fifteen months and not six years (until the 3rd of January 2019) was that the data that were needed to be collected for such a long period, were too many and the available time we had was not enough.

Basically, the Bitcoin transaction dataset contains all the transactions that were carried out in the Bitcoin network. For each one of the transactions, we collected the bitcoin addresses of the sender and the receiver, the transacted Bitcoin value and many other information that come along with the transaction as shown in Figures 4.1, 4.2 and 4.3. This dataset was used in order to analyse and discuss the General Characteristics of the Bitcoin Transaction Network, which follow in Chapter 5.

#	Name	Type
1	transactionID 	varchar(64)
2	blockID	int(11)
3	version	int(11)
4	inputCounter	int(11)
5	outputCounter	int(11)
6	time	datetime
7	relayedBy	varchar(50)
8	inputAddr	varchar(5000)
9	outputAddr	varchar(5000)

Figure 4.3 Transaction Data Fields

#	Name	Type
1	inputID 	int(11)
2	transactionID	varchar(64)
3	prevOutputIndex	int(11)
4	script	varchar(200)
5	sequence	bigint(11)

Figure 4.4 Input Data Fields

#	Name	Type
1	outputID 	int(11)
2	transactionID	varchar(64)
3	outputIndex	int(11)
4	value_BTC	double
5	script	varchar(200)

Figure 4.5 Output Data Fields

4.2.2 Transaction's IP Address

Moving on to our next dataset we have the Transactions' IP Address Dataset. The reason that we need this dataset is because later in Chapter 5, is used to analyse and discuss the section of "Geographical Characteristics of the Bitcoin Transaction Network". In order to be able to do it, we needed the IP address of the initiator of each transaction, so we used the same website that the prior study used, which could provide it to us.

#	Name	Type
1	transactionID 	varchar(64)
2	blockID	int(11)
3	version	int(11)
4	inputCounter	int(11)
5	outputCounter	int(11)
6	time	datetime
7	relayedBy	varchar(50)
8	inputAddr	varchar(5000)
9	outputAddr	varchar(5000)

Figure 4.6 IP Address of the Transaction Initiator

4.2.3 Transaction's IP Geo-location

Moving on to our final dataset we have the Transactions' IP Geo-location Dataset. The reason that we need this dataset is because in Chapter 7, is used to analyse and discuss the section of “Geographical Characteristics of the Bitcoin Transaction Network” and the section of “The Bitcoin Transaction Network with Regards to Cyprus”. In order to be able to do it, we needed the Geo-Location information of the initiator of each transaction, which could be calculated based on the IP address of the initiator of each transaction that we collected in the previous dataset. This particular dataset is consisted of the IP address, the Country Name, the Region Name, the City Name and the Latitude and Longitude of the initiator of each transaction.

#	Name	Type
1	IP 	varchar(50)
2	Country	varchar(50)
3	Region	varchar(100)
4	City	varchar(100)
5	Latitude	varchar(10)
6	Longitude	varchar(10)

Figure 4.7 IP Geo-location of the Transaction Initiator

4.3 Research Questions Preview

In the following Chapters we will analyse the Bitcoin Transaction Network and based on our results and the comparison with the prior findings we will be able to get some inside knowledge about its network structure and characteristics, as well as its growth rate and its popularity among different countries.

More specifically, we will set and answer three research questions regarding the Bitcoin Transaction Network. The first research question is going to characterize the

network from a general perspective, meaning that we will describe the network and its growth just by comparing the size of its blockchain, the total number of blocks and transaction within the blockchain, the total transacted bitcoin value and its market price. Moving on to our second research question, we will study the popularity of Bitcoin network among countries from all over the world, such as the United States of America, Russia and Europe. Based on our analysis in that part we will get some insight about the countries with the largest usage of Bitcoin and by having also access to the prior findings we will be able to answer the question of whether United States and Europe are still the most active Bitcoin markets. Last but not least, our third and final research question will be focusing on the Cypriot activity within the Bitcoin Transaction Network and whether the Bitcoin Transaction Network has any positive, negative or neutral impact on Cypriot Economy or generally on Cyprus.

Chapter 5

General Characteristics of the Bitcoin Transaction Network

5.1	Blockchain Size and Total Number of Blocks	34
5.2	Total Number of Transactions and Total Transacted Value in BTC	35
5.3	Bitcoin Growth	36
5.4	Bitcoin Market Price	38
5.5	Discussion	41

In order to start the analysis about Bitcoin, we decided to examine some general characteristics of the network. In the following section we are going to analyse the growth of the Bitcoin blockchain and more specifically we are going to compare the results of Matthias Lischke and Benjamin Fabian’s study [13] with the results of this study.

M. Lischke and B. Fabian’s analysis was focused on the Bitcoin network’s activity from the 3rd of January 2009 until the 10th of April 2013. The basis of their analysis was the Bitcoin transaction dataset, which was publicly available and extracted from a full node with Bitcoin client (version 0.5.3.1) [23]. The dataset consisted of 230,686 blocks and its size was 1.51 GB. Since the data from the Bitcoin node was downloaded in binary files, they used a Python2.7 data scrapper tool [5], which gave them the information they needed in human readable form.

Through their study, they observed some general statistics that described the network. By the 10th of April 2013, approximately 6.3 million user entities were involved in over 15.8 million transactions in the network. They chose to observe the changes of the transacted value from day to day as shown in Table 5.1. In this way they came to the conclusion that most of the transactions had small values and that according to the

statistics they got, there was a low activity in the network compared to the distribution of users and transactions.

	Median	Mean	Sd	Skew	Min	Max	Correl (ExRate)
Transaction Value (BTC)	173,457	910,053	2,231,647	7	50	29,958,714	0.199
Number of Active Users	1637	4049	5243	2	1	36,120	0.730
Number of Transactions	3678	24,084	38,303	2	1	189,284	0.680

Table 5.1 Bitcoin Network Statistics on Daily Basis [13]

In the same way to M. Lischke and B. Fabian, we chose to analyse some general characteristics of the Bitcoin network. This analysis is focusing on the Bitcoin network's activity from the 3rd of January 2009 until the 16th of July 2014 at 18:24 and on the progress that has been made in the Bitcoin network.

5.1 Blockchain Size and Total Number of Blocks

In order to extract some statistics and be able to compare our results with the prior findings, we used the Bitcoin Transaction Dataset that we discussed earlier. The dataset is containing 311,042 blocks and its size is approximately 20.5 GB.

The most important and easy observation that anyone could make, is that the Bitcoin blockchain is sufficiently larger. In their prior study, M. Lischke and B. Fabian used a dataset that concerned the first four years of Bitcoin's operation, whereas in this study we use a dataset that concerns just fifteen more months. However, as we can see in Table 5.2, the size of the blockchain is much larger.

	03/01/2009 - 10/04/2013	03/01/2009 - 16/07/2014	11/04/2013 - 16/07/2014	Comparison
Total Number of Blocks	230,686	311,042	82,563	Approx. 1.35 times more
Blockchain Size	1.51 GB	20.5 GB	18.99 GB	Approx. 13.57 times bigger

Table 5.2 Number of Blocks, Size of Bitcoin Blockchain, Comparison

To be more accurate, by the 16th of July 2014 in the current dataset there were approximately 1.35 times more blocks and in addition the size of the dataset is about 13.57 times bigger than the prior study.

5.2 Total Number of Transactions and Total Transacted Value in BTC

After the observation about the size of the blockchain and the number of the blocks contained in the blockchain, we wanted to study which differences we could identify between the two studies that concerned the total number of transactions and the total number of transacted Bitcoin value.

As we can see in Table 5.3, the total number of transactions carried out in the Bitcoin network during the following fifteen months, is larger than the number of transactions executed in the first four years and more specifically, the total number of executed transactions in the current dataset is approximately 2.33 times larger than the total number of transactions of the initial dataset. The current dataset that is used for the analysis contains over 37 million transactions that they were carried out in the Bitcoin network by people all around the globe, with a total transacted Bitcoin value of about 86 million. Furthermore, we can observe that although the total number of transactions is considerably bigger, the total number of transacted Bitcoin value is only 1.32 times larger.

Following the above observation, it is reasonable to conclude that most of the transactions carried out in the Bitcoin network continued to have small values.

	03/01/2009 - 10/04/2013	03/01/2009 - 16/07/2014	11/04/2013 - 16/07/2014	Comparison
Number of Transactions	15,898,625	37,029,456	21,130,831	Approx. 2.33 times more
Total output value (BTC)	65,864,246	86,950,661	21,086,415	Approx. 1.32 times more

Table 5.3 Total Number of Transactions and Total Transacted Value in BTC

5.3 Bitcoin Growth

Following the above observations, we wanted to know exactly at which rate the Bitcoin network has grown. In order to achieve that, a graph was created. The values that were used to create this graph is the total number of transactions until a specific date as shown in Table 5.4. For example, until the 3rd of January 2009 we had 1 transaction and until the 12th of October 2009, we had 25,030 transactions (the 1 transaction from 2009 is included in the number 25,030).

	Number of Transactions per Year (2009-2014)
03/01/2009	1
12/10/2009	25,030
12/04/2010	51,348
13/10/2010	128,151

13/04/2011	407,557
12/10/2011	1,677,142
13/04/2012	2,820,743
30/07/2012	5,456,202
12/10/2012	7,917,958
04/01/2013	10,725,281
12/04/2013	16,022,580
13/10/2013	25,330,588
03/01/2014	30,358,785
16/07/2014	37,029,456

Table 5.4 Values used for the graph

Figure 5.1 plots the rate at which the total number of transactions increases over the years. By observing the graph, we understand that the total number of transactions in the Bitcoin network was increasing exponentially during the following fifteen months and based on the data that we have collected so far, there is no indication that the growth of the Bitcoin network was going to slow down.

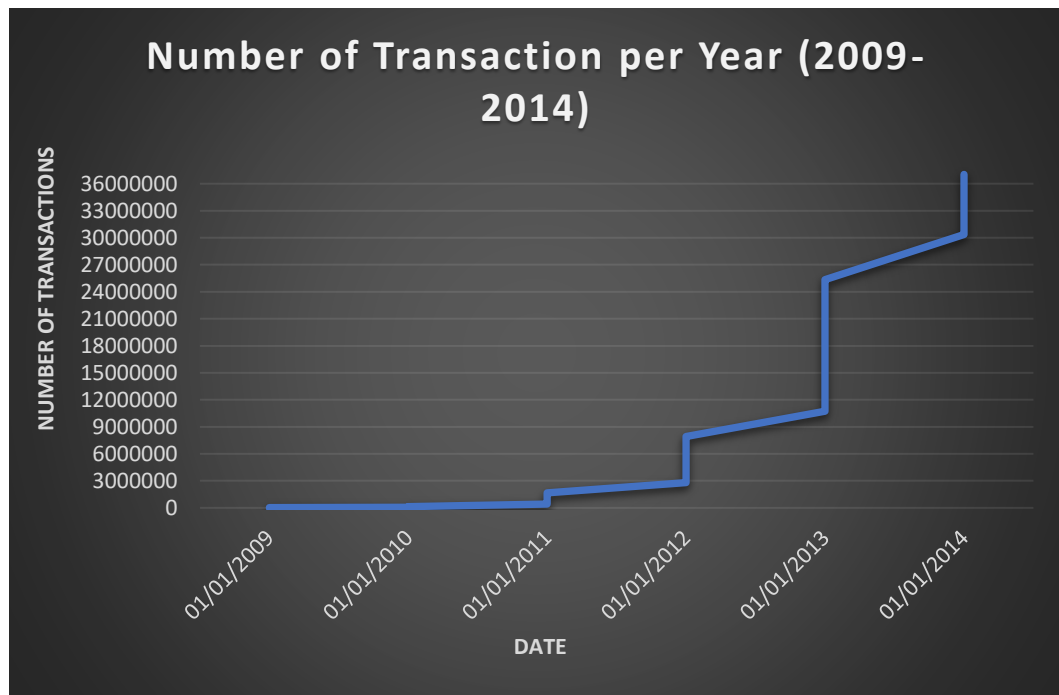


Figure 5.1 Number of Total Transactions (2009-2014)

5.4 Bitcoin Market Price

Another characteristic of the Bitcoin network that we can observe, regardless of this study's datasets, is the Bitcoin price from the very beginning. Based on Table 5.5 and Figure 5.2, we can see that at the beginning, from January 2009 until the end of 2010, the Bitcoin was basically worthless. Then, at the beginning of 2011, Bitcoin value started to increase, when in February 2011, 1 BTC was approximately equal to 1 *United States Dollar (USD)*.

Date	1 BTC to USD	Date	1 BTC to USD
03/01/2009	0	02/12/2015	360.98
03/01/2010	0	03/06/2016	568

03/01/2011	0.299998	01/01/2017	997.729875
10/02/2011	1.1	28/03/2017	1,046.127625
02/06/2011	10.57	31/08/2017	4,748.255
02/01/2012	5.4999	30/10/2017	6,105.87422
02/06/2012	5.279	01/12/2017	10,883.912
02/01/2013	13.4	17/12/2017	19,783.06
03/06/2013	120.00002	02/01/2018	15,005.8567
28/11/2013	1,009	01/02/2018	9,083.2583
03/01/2014	806.21	02/05/2018	9,221.426
11/04/2014	344.22166	13/06/2018	6,315.7
25/11/2014	381.03	04/12/2018	3,961.493
03/06/2015	226.29	03/01/2019	3,865.7975

Table 5.5 Some Bitcoin Market Prices (USD) - Price Table [35]

By the end of November 2013, we see the first time that the Bitcoin value exceeded the amount of 1,000 USD. Then, in April 2014 we observe the lowest Bitcoin value since November 2013, when Bitcoin value dropped to approximately 340 USD, after that Bitcoin value continue dropping and then started increasing again. By March 2017, Bitcoin value exceeded the amount of 1000 USD and continued its increase until the

middle of December 2017, when Bitcoin value reached its highest value at 19,783.06 USD.

Finally, by February 2018 Bitcoin value dropped 50% and 1 BTC was about 10,000 USD. The Bitcoin value continued to drop and by October 2018, 1 BTC was worth around 6,300 USD. These days, due to the fact that Bitcoin value is unstable, 1 BTC is about 7100 USD with a lot of changes from day to day.

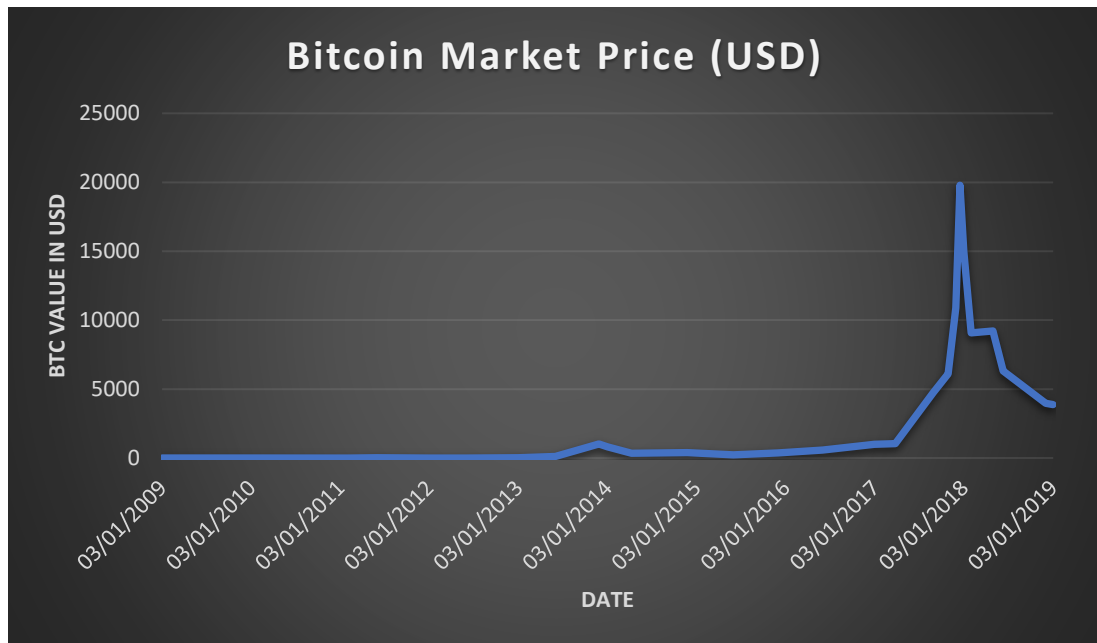


Figure 5.2 Bitcoin Market Price (USD) – Graph [35]

The observation that we can make, from the above graph is that the Bitcoin value is not very stable. In the last 2 years, Bitcoin value has changed rapidly from 1,000 USD to the impressive amount of approximately 20,000 USD and again back approximately to the much lower amount of 4,000 USD. Although, for the last 4 months the Bitcoin value seemed to gain some stability and then again to gain more value, nobody can be certain that the price would remain stable.

5.5 Discussion

In conclusion, based in the above measurements we can certainly say that the Bitcoin blockchain has significantly grown. In a period of just fifteen months the data that are stored in the blockchain has increased about 13.57 times, the transactions that were carried out in the network are more than doubled and the transacted Bitcoin value has also increased approximately by 1.32 times.

Those results show us, that although there are a lot more transactions carried out in the network, there are not many bitcoins transacted each time (neither in the prior study transactions had large transacted values). The above observation is completely normal and what we expected, since the Bitcoin market price was higher, and someone could send and receive larger amounts of USD just by sending or receiving the same or fewer number of bitcoins.

Furthermore, we can clearly observe the exponential growth of the Bitcoin Transaction Network in accordance to the transactions executed within the network, since in a period of fifteen months the network has become more than twice as big in comparison with the period of the first four years. Finally, we can safely come to the conclusion that the more popular the Bitcoin Network becomes, the more value its bitcoins are getting, although due to the fact that Bitcoin is just an electronic cryptocurrency and it does not hold any market price stability people are very careful and sceptical about its use.

Chapter 6

Geographical Characteristics of the Bitcoin Transaction Network

6.1	Number of Distinct IP Addresses and their Geographical Representation	45
6.2	Number of Executed Transactions per Country	47
6.3	Discussion	49

Moving on to our next research question, our goal is to study and analyse the geographical characteristics of the Bitcoin Transaction Network, while we will gain some important information about the geographical distribution of the Bitcoin economy.

In the prior study the two researchers manage to find 40,329 distinct geo-location details and based on the 72.4% of all the transactions within their dataset they extracted some statistics about the Bitcoin economy. Furthermore, they concluded that a percentage of approximately 10.7% were carried out by the website “*Blockchain.info*” and a percentage of around 16.6% could not be related to any *Internet Protocol (IP)* address.

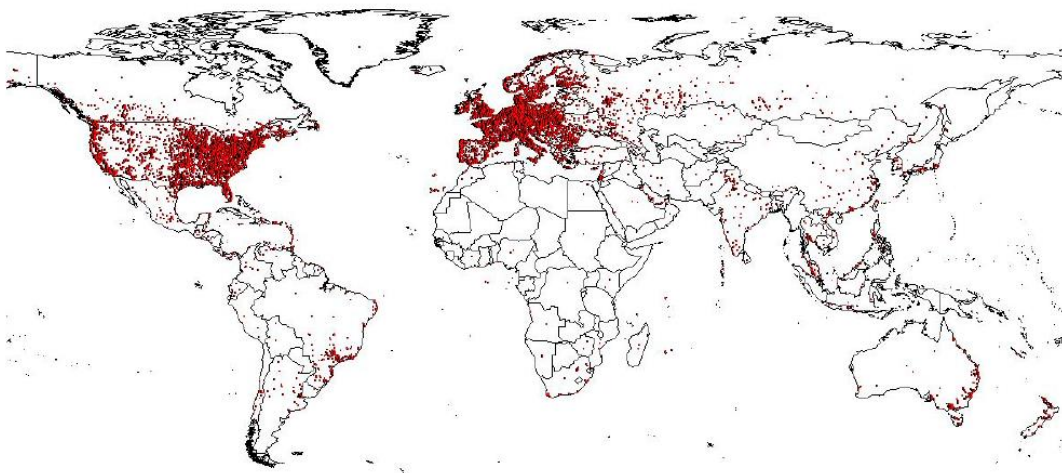


Figure 6.1 IP nodes distributed on the world map [13]

Subsequently, their findings showed them that the United States and German markets were the two most active ones, while a positive relationship between the usage of Bitcoin and well-developed countries emerged, since the most IP nodes were located in the United States and Europe as can be seen in Figure 6.1. Furthermore, based on some particular countries, such as the United States, Germany, Russia, China and some others they manage to depict the development of those countries over a period of time in accordance to the number of IP nodes.

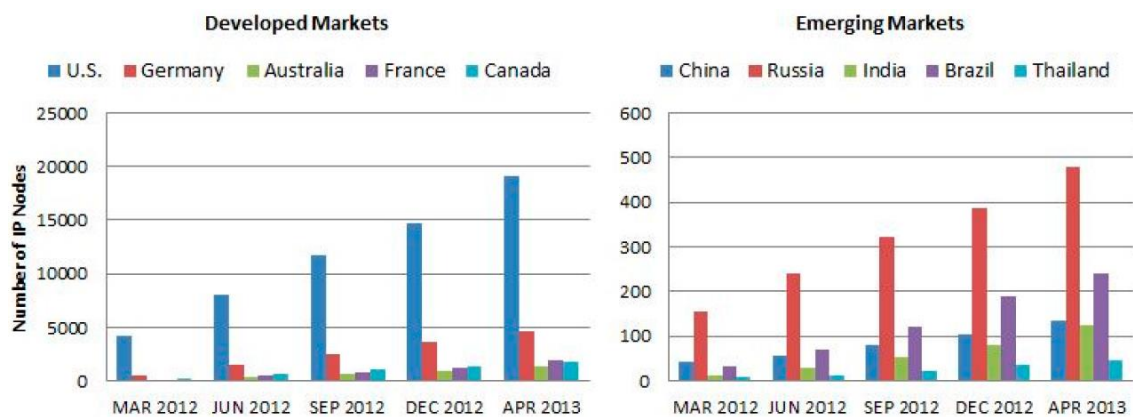


Figure 6.2 Development of particular Countries over Time [13]

Finally, they were able to extract some statistics, which were invoking the transacted Bitcoin value of each country and the business categories they had found. Those statistics showed that the European Countries such as Germany and France were focusing on the mining aspect of the Bitcoin Transaction Network, whereas other countries such as the United States and China were focusing on the gambling aspect of the Bitcoin Transaction Network.

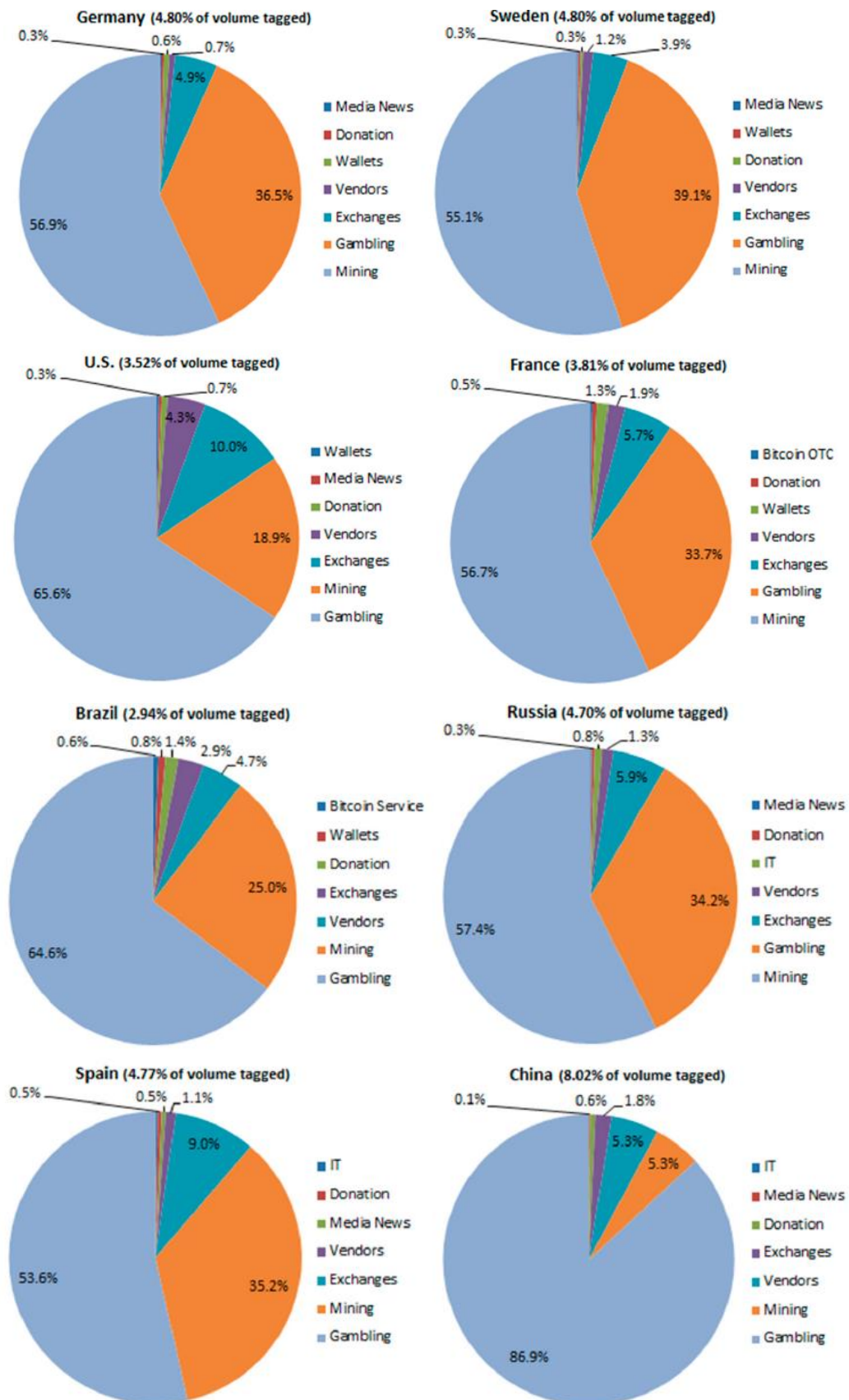


Figure 6.3 Transacted Bitcoin Value per Business Category per Country [13]

In this section we are going to see some statistics referring to the number of distinct IP addresses and their distribution over the globe, as well as some statistics referring to the total number of executed transactions per country. Unfortunately, we will not be able to extract statistics about the Business categories for each country, since the website “*Blockchain.com*” does not provide the info of Business Tag anymore. Finally, we will discuss our results with regards to the prior findings and we will come to some conclusions based in the comparison.

6.1 Number of Distinct IP Addresses and their Geographical Representation

Based on the IP of the initiator of each transaction, we managed to collect 230,472 distinct geo-location details, which belonged to 175 distinct countries. A percentage of 17.23% of all the transactions is tagged with the IP address “127.0.0.1”, which means that those transactions have been carried out by the website “*Blockchain.com*” and there is also a percentage of 13.45% of all transactions, which is tagged with the IP address “0.0.0.0”, which means that those transactions could not be linked to any IP address.

The remaining 69.32% is used in order to extract the different graphs and statistics, such as the total number of executed transactions per country that we are going to discuss in a while.

	03/01/2009 - 10/04/2013	03/01/2009 - 16/07/2014
Transactions tagged as “127.0.0.1”	10.7%	17.23%
Transactions tagged as “0.0.0.0”	16.6%	13.45%
Transactions used for analysis	72.4%	69.32%

Table 6.1 Transactions tagged with different IPs (percentage of the total transactions contained in each dataset)

With the help of “*Google Fusion Tables*” [34] we managed to make a geographical representation of the global distribution of the distinct IP Addresses we collected. As we can see in Figure 6.4, the most IP addresses are associated with the United States and Europe, while some countries like China and Russia are also associated with a significant amount of IP addresses.

More specifically, United States has the biggest number of distinct IP addresses, which is 65,939 IPs and almost 3 times bigger than the second biggest number that belongs to Germany and it is 22,389 IPs. Based on Figure 6.5, we can see that the top eight countries regarding the number of distinct IP addresses, are basically the United States or countries within Europe, while countries like China, Russia, Australia and Canada are also associated with a lot of IP addresses, but those IP address are less than those linked to United States and Europe.

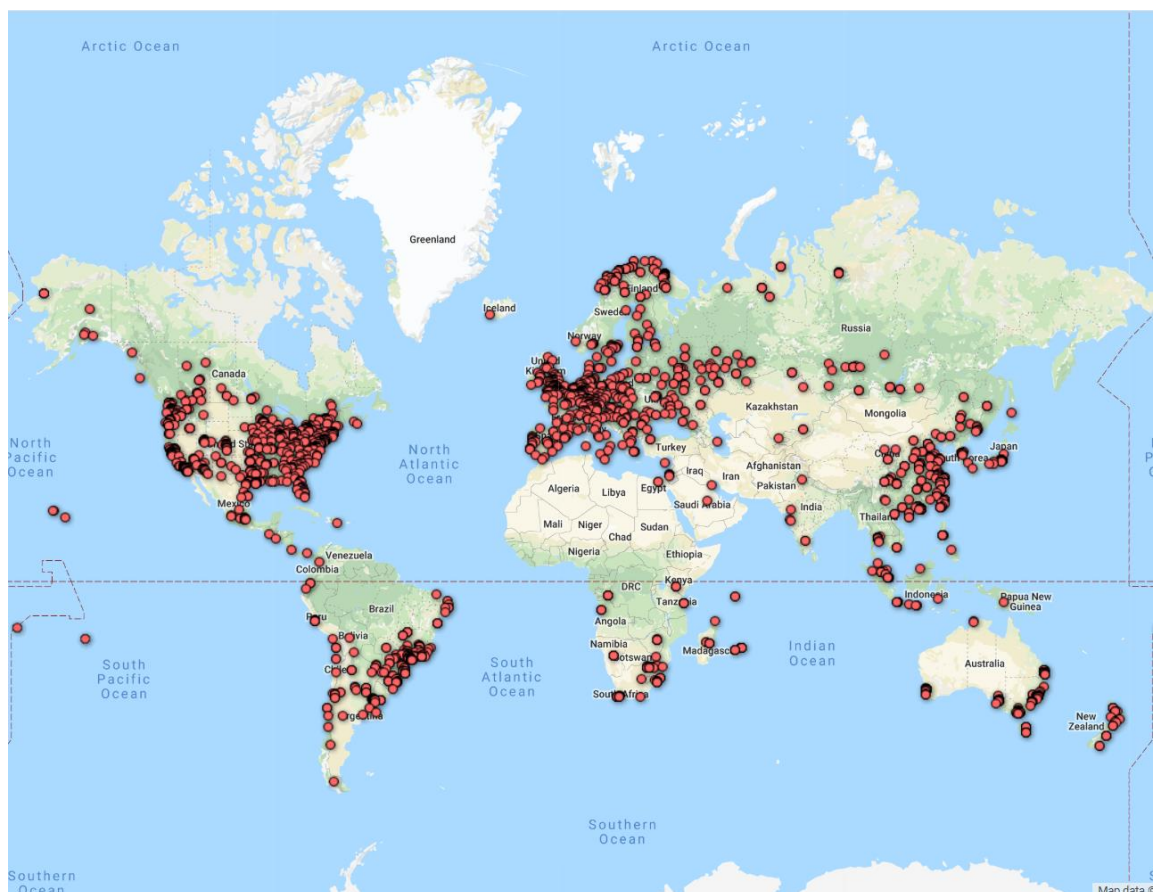


Figure 6.4 IP Address Geo-location - Global Distribution

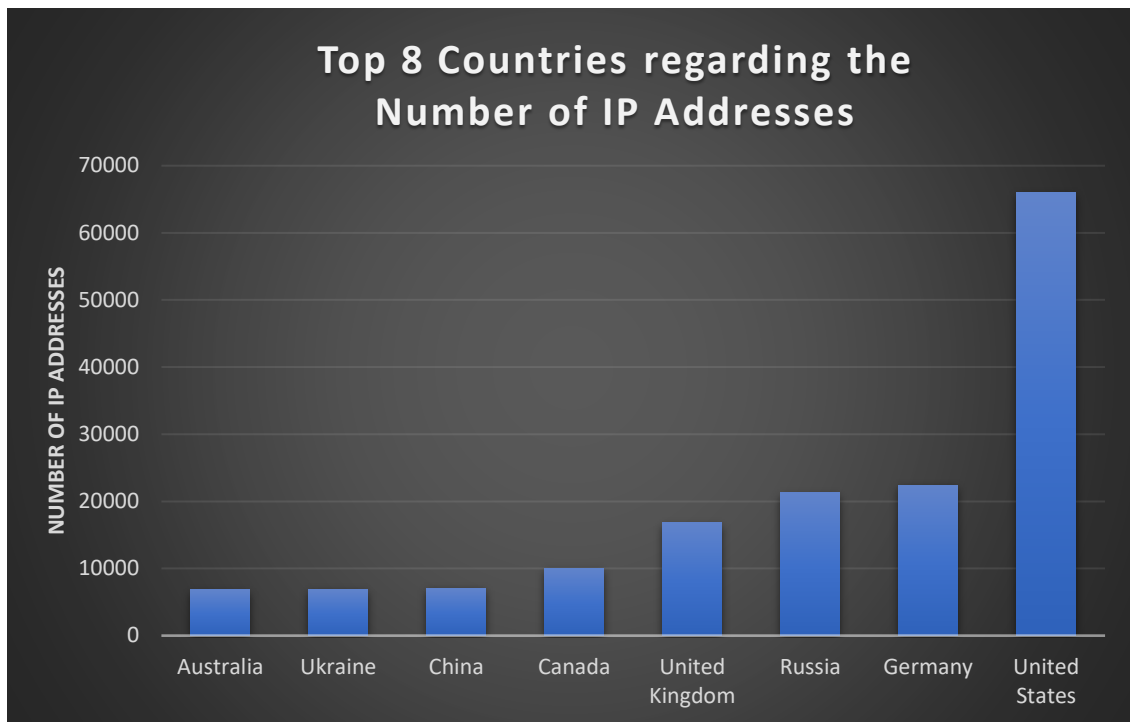


Figure 6.5 The 8 Countries with the biggest number of distinct IP addresses

6.2 Number of Executed Transactions per Country

Since we got the global distribution of the distinct IP addresses per country, we proceeded with some statistics regarding the total transactions carried out within the network for the eight countries with the biggest number of distinct IP addresses. In Figure 6.6 and Figure 6.7 we managed to depict the percentage of the total transaction carried out per country alongside the percentage of their distinct IP addresses. Although we can clearly see that over the 68% of all the Distinct IP address originates from those eight countries, we can also observe that only about 52% of all the executed transactions were initiated from those countries. Considering that the IP addresses are associated with 175 distinct countries, the above two percentages are telling us that these eight countries are the major markets of Bitcoin Transaction Network and that over the half amount of the executed transactions are initiated by just those few countries.

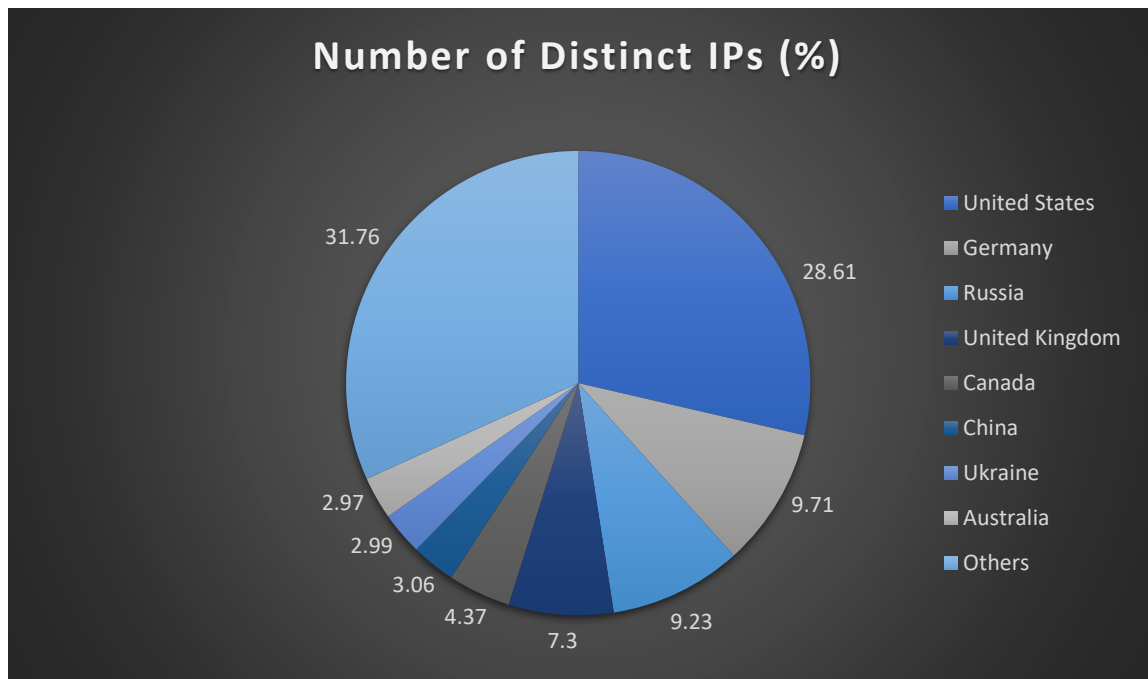


Figure 6.6 Percentage of Distinct IP Addresses per Country

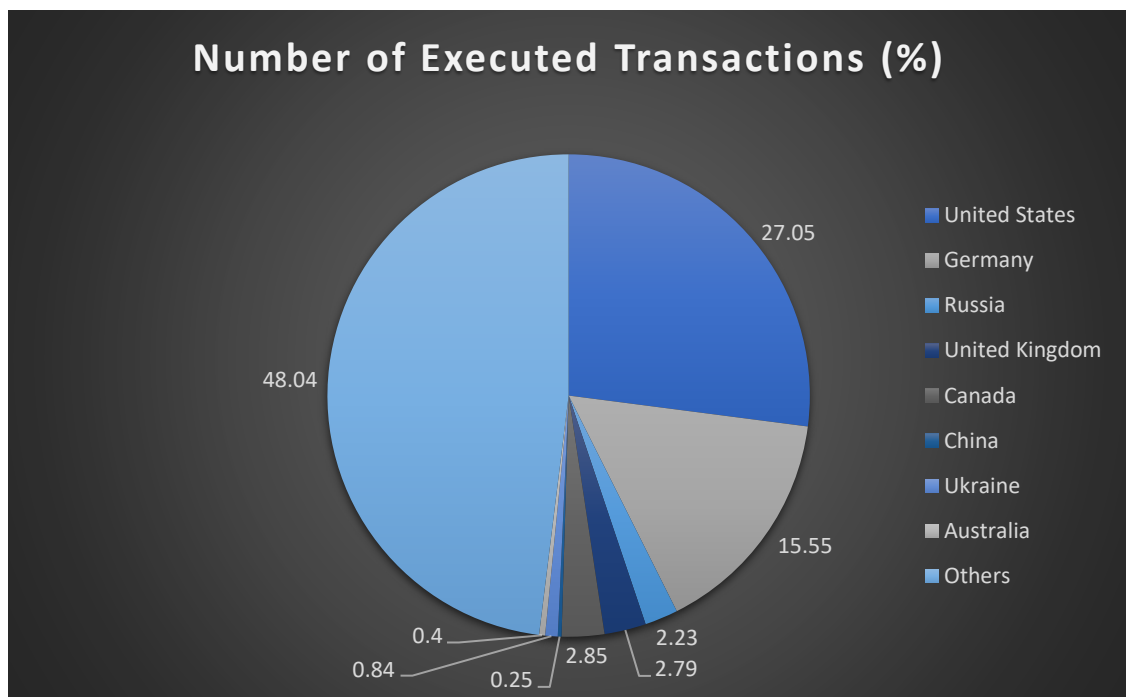


Figure 6.7 Percentage of Executed Transactions per Country

6.3 Discussion

Summarizing everything we have seen in this particular chapter, led us to the conclusion that the Bitcoin Transaction Network had two prime markets, the first one located in the United States of America followed by the second one in Germany. Furthermore, we observed that countries like Canada, Russia and United Kingdom are also actively participating in the network with a significant amount of executed transactions, while other countries like China, Ukraine and Australia were categorized within the eight countries with the most distinct IP address, but they had a less active involvement in the network.

Chapter 7

The Bitcoin Transaction Network with Regards to Cyprus

7.1	Cyprus Economy	50
7.2	Number of Distinct IP Addresses in Cyprus and their Geographical Representation	51
7.3	Number of Transactions Initiated from Cyprus	53
7.4	Discussion	54

7.1 Cyprus Economy

Cyprus Economy [31][32][33] was characterized by the World Bank as an economy with high incomes in 2001. On the 1st of May 2004, the Republic of Cyprus become part of the European Union and by the 1st of January 2008 it joined Eurozone, meaning that it replaced the Cypriot Pound with Euro as its official currency.

During the 2012 and 2013 Cyprus experienced the Cypriot Financial Crisis, which had dominated its economy, but in the last few years the country has started overcoming the challenges of that Crisis. Although Cyprus is a rather small country, it managed to impress the European Union and exceed their expectations, by exiting the financial assistance programme that was in by 2016, which was earlier than expected.

During the following year an economic growth was noted and now the priority of Cyprus is to maintain its economy, while it improves its financial status and attracts new investors that will invest in the island and boost its economy.

Last but not least, Cyprus can be characterized as a financial hub as it is very famous as an entry point used from countries outside European Union in order to invest within Europe or from countries in the West to invest into Russia and Eastern Europe. Moreover,

the geographical position of the island (crossroad of three continents and its proximity to the Suez Canal) gave it the opportunity to become a shipping hub for a lot of shipping companies and enterprises, which are using Cyprus as their management centre.

Based on all the above facts, we decided to investigate whether Bitcoin Network had any influence on Cyprus. We also wanted to investigate whether Cyprus participated in the Bitcoin Network and if it did, how big its participation was.

7.2 Number of Distinct IP Addresses in Cyprus and their Geographical Representation

In the previous Chapter we saw and discussed the geographical distribution of the distinct IP addresses that exist within our dataset in a global scale. Now, let's see the geographical distribution of distinct IP addresses especially in Cyprus. In Cyprus we could find 82 distinct IP addresses associated with the Bitcoin Transaction Network, which is about the 0.036% of all the distinct IP addresses of the network, and that addresses are graphically represented in Figure 7.1. The geographical coordinates for each IP address was extracted based on the City that it belonged, so it is logical that these coordinates are not 100% accurate. So, for the IP addresses that belong to the same City the coordinates are the same and that is why Figure 7.1 does not depict 82 different locations, but only 14.

An observation that anyone could make, is that there are IP addresses associated with the Bitcoin Transaction Network in the five out of six districts of Cyprus. Nicosia seems to be the most active district for the time that we are studying and if you consider that in the years that follow the University of Cyprus and the University of Nicosia are actively researching the blockchain technologies, with the second one also accepting Bitcoin payments, we can safely assume that Cyprus activity in the network will be greater. To be more accurate, there is one distinct IP address in Famagusta, three IP addresses in Paphos, seven IP addresses in Larnaca, thirteen IP addresses in Limassol and fifty-eight IP addresses in Nicosia.



Figure 7.1 IP Address Geo-location - Cyprus Distribution

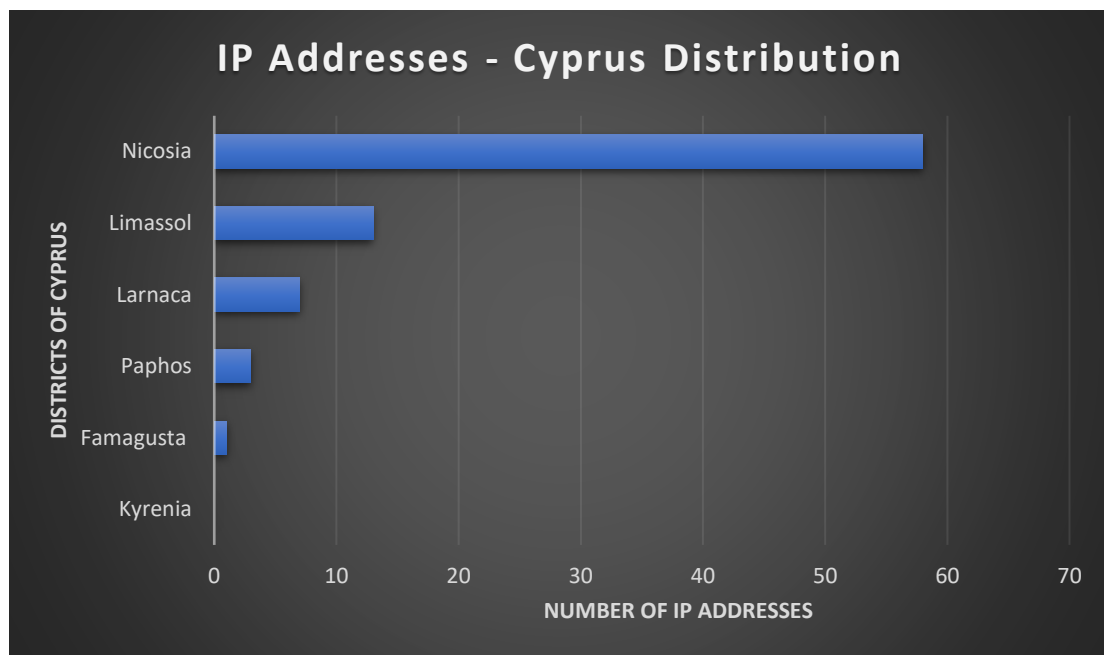


Figure 7.2 IP Addresses in Cyprus that participated in the Bitcoin Transaction Network

7.3 Number of Transactions Initiated from Cyprus

Similarly to the previous Chapter, where we extracted some statistics about the total number of transactions carried out within the network by each country, now we will focus our attention on the total number of transactions initiated from Cyprus and the comparison of the results with the results of other countries, which have equivalent infrastructure and size. The countries that are used for this analysis are Cyprus, Montenegro, Malta, Luxembourg and Iceland, which are all members of the Small States of Europe.

As shown in Figure 7.3, the distinct IP addresses for these small countries are varying between 59 and 165, while Cyprus has 82 distinct IP addresses. This particular observation cannot lead us to any conclusions, but along with Figure 7.4 we can get some useful insights. Firstly, we can observe that Cyprus has the fewest distinct IP addresses after Montenegro, while Luxembourg has the most IP addresses followed by Iceland. According to the second graph we can clearly see that the three countries with the fewest IP addresses also have fewer total number of transactions (between 0.001% and 0.008%), while Luxembourg and Iceland have a lot more (between 0.13% and 0.17%).

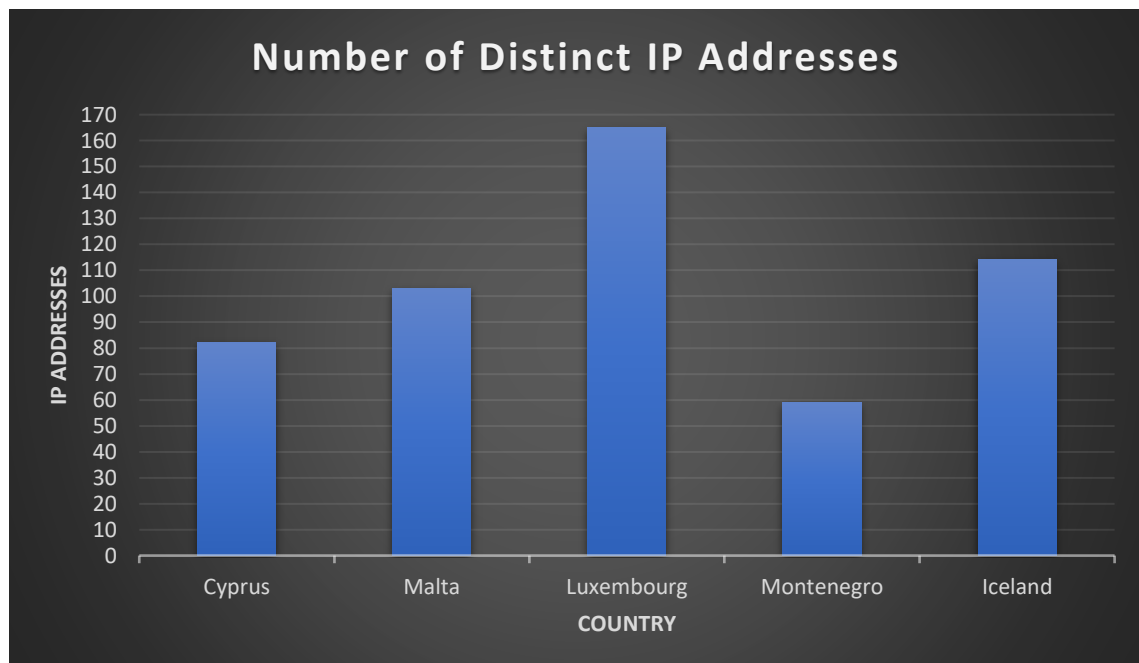


Figure 7.3 Number of Distinct IP Addresses for Countries Equivalent to Cyprus

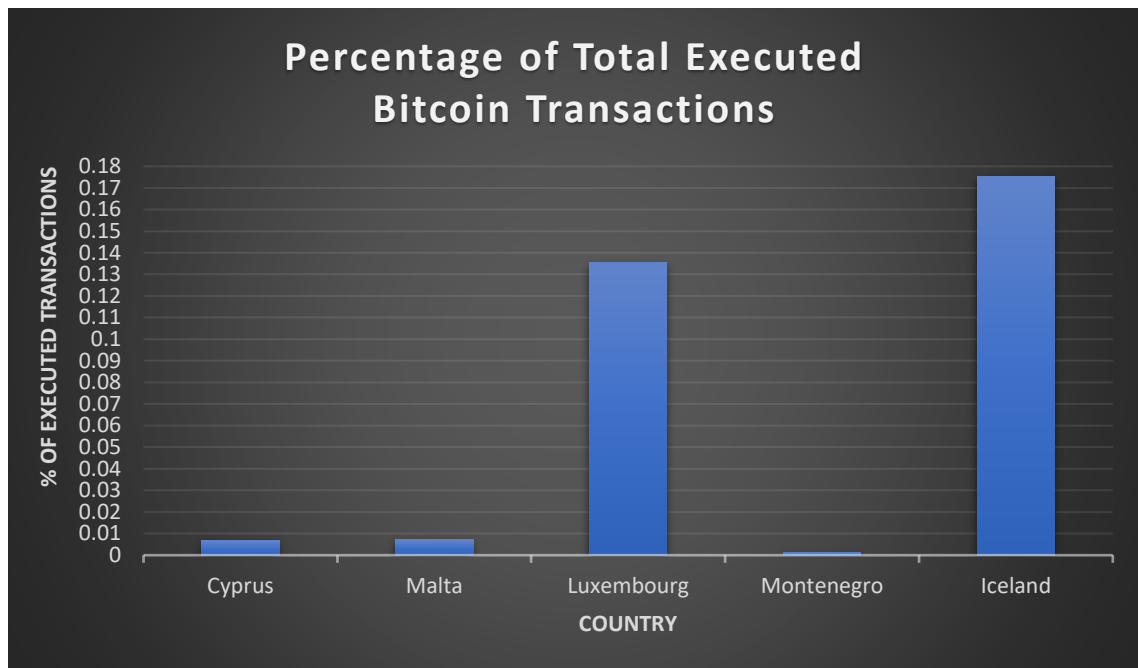


Figure 7.4 Percentages for the Total Number of Executed Bitcoin Transactions for Small Countries of the Network

7.4 Discussion

The analysis of this final Chapter was aiming to answer the question whether Cyprus is an active market of the Bitcoin Transaction Network, not in the same way as the United States or Germany as we saw before, but in a way that corresponds to the size, population and infrastructure of the country. In order to achieve that we compared the statistics that we extracted for Cyprus with the same statistics for countries that belong to the Small States of Europe and are similar to Cyprus.

Our results were based on a group of five countries (Cyprus, Malta, Montenegro, Luxembourg and Iceland) and they showed us that Cyprus is not a very active market of Bitcoin Transaction Network after all. Although Cyprus has more than 80 IP addresses associated with the network, those addresses are corresponding only to about 0.0068% of the total number of transactions that were carried out within the network, while Luxembourg has been associated with about 0.13% of all the transactions and it has also been linked with the double amount of IP addresses that are associated with Cyprus.

In conclusion, we can certainly say that Cyprus is not as active as we expected within the Bitcoin Transaction Network, at least based on the period that we are analysing. However, someone could expand the datasets of our study (maybe until 2018 or 2019) and investigate whether there was a progress of the relationship between them in the years that follow, based on the fact that as we mentioned earlier in the years that follow Cypriot Universities have started researching the blockchain technology and one of them has also started accepting payments through Bitcoin.

Chapter 8

Conclusion

8.1	Conclusions	56
8.2	Limitations	57
8.3	Future Work	58

8.1 Conclusions

Summarizing our work, we can finally give an answer to the three research questions that we have set at the beginning of this Thesis Dissertation. First of all, the Bitcoin blockchain has started to grow exponentially after the first four years. Through our comparative results between our study and the prior one, we observed that the current dataset of the Bitcoin Transaction Network has become more than 13 times larger just by studying fifteen more months, while the Bitcoin transacted value has increased about 1.3 times. Although the network has a lot more transactions than before, the transacted bitcoins are not too many, like they were not many in the prior findings as well. Our last observation about the general characteristics of the network, is that although bitcoins have gained more value over time and more people are interested in using the network, the market price of Bitcoin is unstable.

Moving on to the results of our second research question, we concluded that United States of America and Germany can be still characterized as the two major Bitcoin markets followed by the countries of Russia, United Kingdom and Canada, which are pretty active within the Bitcoin Transaction Network as well. Furthermore, the countries of China, Ukraine and Australia can neither be characterized as major Bitcoin markets nor as minor, due to the fact that although these particular countries do not have as many executed transactions compared to the countries we just mentioned, they are listed in the

top eight countries associated with the most IP addresses that are participating in the Bitcoin Transaction Network.

Finally, our last research question was focused on the participation of Cyprus within the Bitcoin Transaction Network and whether Cyprus was interested in the network. For this part of our analysis we chose to compare Cyprus with other countries like Malta and Luxembourg, which are all countries that belong to the Small States of Europe. Based on our results, we were led to the conclusion that although Cyprus has participated in the network, this participation was not as big as we initially thought. Lastly, by comparing Cyprus and Luxembourg we understood that Cyprus is not very interested in participating in the Bitcoin Transaction Network, at least until the mid of July 2014, since the transactions initiated from Cyprus were only about 0.0068% of the transactions executed all around the world, whereas the transactions initiated from Luxembourg were about 0.13%.

8.2 Limitations

During this study, unfortunately some limitations and problems occurred, which we had to overcome in some way. More specifically, our first idea for this study was to analyse the Bitcoin Transaction Data from the 3rd of January 2009 until the 3rd of January 2019, so we downloaded the Bitcoin raw data that are publicly available through the Bitcoin Client.

When we tried to use the Python data scrapper that was used in the prior research, we faced our first problem, because the specific tool could not work due to the change of the Bitcoin Client version. Subsequently, we tried to find another tool to convert the raw data to human readable form in order to manage them through our database, but we could not find the right tool and that led us decide to use the “Blockchain.com” website to download the transaction data along with the IPs of the transactions’ initiators.

Another limitation we had, was that the specific website had a data limit, which was set by the website administrator and that was preventing us from downloading big amounts of data from the same IP address for a specific period of time. In order to overcome this limitation, we used the resources from our University’s Unix Labs so as to

request the data we needed from the website from many different IP addresses in parallel. In this way we increased the data we could gathered per day, but again it was not enough. Although we managed to increase the amount of data we could gather per day, the transactions carried out in the Bitcoin network during these ten years were over 370 million. Since we could download approximately 2.5 million transactions per day, we would need approximately 5 months to gather all the data.

Our last solution to overcome the limited access we had to the website and the limited time we had in our disposal, was to narrow down the period we wanted to analyse from ten years to five and a half, which meant that the original dataset from the prior study would be expanded only by fifteen months.

8.3 Future Work

Some future works that could be conducted, is the analysis of the Bitcoin Transaction Data for a much larger period of time than the one that we are using and even observe the relationship between the network and Cyprus in order to investigate whether there are any changes in the years that follow. Furthermore, with enough computational power someone could not only expand the analysed period of the Bitcoin Transaction Data, but he/she could also re-visit the second part of Matthias Lischke and Benjamin Fabian's study, which was focused on some network metrics and on the investigation of the presence of the small world phenomenon on the Bitcoin, but unfortunately it was out of scope for this Thesis Dissertation.

References

- [1] Néstor Álvarez-Díaz, Jordi Herrera-Joancomartí, and Pino Caballero-Gil. 2017. Smart contracts based on blockchain for logistics management. In Proceedings of the 1st International Conference on Internet of Things and Machine Learning (IML '17). ACM, New York, NY, USA, Article 73, 8 pages.
DOI: <https://doi.org/10.1145/3109761.3158384>
- [2] Elias Athanasopoulos. 2018. Asymmetric Encryption and RSA. Lecture Notes pp. 18-23.
- [3] Elias Athanasopoulos. 2018. Advanced Encryption Standard (AES). Lecture Notes.
- [4] Md Zakirul Alam Bhuiyan, Aliuz Zaman, Tian Wang, Guojun Wang, Hai Tao, and Mohammad Mehedi Hassan. 2018. Blockchain and Big Data to Transform the Healthcare. In Proceedings of the International Conference on Data Processing and Applications (ICDPA 2018). ACM, New York, NY, USA, 62-68.
DOI: <https://doi.org/10.1145/3224207.3224220>
- [5] I. Brugere. 2013. Bitcoin-Transaction-Network-Extraction.
- [6] Francesco Buccafurri, Gianluca Lax, Serena Nicolazzo, and Antonino Nocera. 2017. Overcoming Limits of Blockchain for IoT Applications. In Proceedings of the 12th International Conference on Availability, Reliability and Security (ARES '17). ACM, New York, NY, USA, Article 26, 6 pages.
DOI: <https://doi.org/10.1145/3098954.3098983>
- [7] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. 2018. Detecting Ponzi Schemes on Ethereum: Towards Healthier Blockchain Technology. In Proceedings of the 2018 World Wide Web Conference (WWW '18). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1409-1418.
DOI: <https://doi.org/10.1145/3178876.3186046>
- [8] Roberto Di Pietro, Xavier Salleras, Matteo Signorini, and Erez Waisbard. 2018. A blockchain-based Trust System for the Internet of Things. In Proceedings of

the 23rd ACM on Symposium on Access Control Models and Technologies (SACMAT '18). ACM, New York, NY, USA, 77-83.

DOI: <https://doi.org/10.1145/3205977.3205993>

- [9] Lingjun Fan, J. Ramon Gil-Garcia, Derek Werthmuller, G Brian Burke, and Xuehai Hong. 2018. Investigating blockchain as a data management tool for IoT devices in smart city initiatives. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age (dg.o '18), Anneke Zuiderwijk and Charles C. Hinnant (Eds.). ACM, New York, NY, USA, Article 100, 2 pages. DOI: <https://doi.org/10.1145/3209281.3209391>
- [10] Adishesu Hari and T. V. Lakshman. 2016. The Internet Blockchain: A Distributed, Tamper-Resistant Transaction Framework for the Internet. In Proceedings of the 15th ACM Workshop on Hot Topics in Networks (HotNets '16). ACM, New York, NY, USA, 204-210.
DOI: <https://doi.org/10.1145/3005745.3005771>
- [11] Simona Ibba, Andrea Pinna, Matteo Seu, and Filippo Eros Pani. 2017. CitySense: blockchain-oriented smart cities. In Proceedings of the XP2017 Scientific Workshops (XP '17). ACM, New York, NY, USA, Article 12, 5 pages. DOI: <https://doi.org/10.1145/3120459.3120472>
- [12] D. Kaminsky. 2011. Black ops of TCP/IP. Black Hat USA. pp. 44. Available: <https://de.slideshare.net/dakami/black-ops-of-tcpip-2011-black-hat-usa-2011>.
- [13] Matthias Lischke, Benjamin Fabian. 2016. Analyzing the bitcoin network: The first four years. Future Internet.
- [14] C. Mohan. 2017. Tutorial: blockchains and databases. Proc. VLDB Endow. 10, 12 (August 2017), 2000-2001. DOI: <https://doi.org/10.14778/3137765.3137830>
- [15] Arvind Narayanan. 2018. Blockchains: Past, Present, and Future. In Proceedings of the 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (SIGMOD/PODS '18). ACM, New York, NY, USA, 193-193. DOI: <https://doi.org/10.1145/3196959.3197545>
- [16] C. Năsulea, S. Mic. 2018. Using Blockchain as a Platform for Smart Cities. Journal of E-Technology Volume. vol. 9.
- [17] Patrick Ocheja, Brendan Flanagan, and Hiroaki Ogata. 2018. Connecting decentralized learning records: a blockchain based learning analytics platform.

- In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18). ACM, New York, NY, USA, 265-269.
DOI: <https://doi.org/10.1145/3170358.3170365>
- [18] Mayra Samaniego and Ralph Deters. 2016. Using Blockchain to push Software-Defined IoT Components onto Edge Hosts. In Proceedings of the International Conference on Big Data and Advanced Wireless Technologies (BDAW '16), Djallel Eddine Boubiche, Hani Hamdan, and Ahcène Bounceur (Eds.). ACM, New York, NY, USA, Article 58, 9 pages.
DOI: <https://doi.org/10.1145/3010089.3016027>
 - [19] Davor Svetinovic. 2017. Blockchain Engineering for the Internet of Things: Systems Security Perspective. In Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security (IoTPTS '17). ACM, New York, NY, USA, 1-1. DOI: <https://doi.org/10.1145/3055245.3055256>
 - [20] Melanie Swan. 2018. Blockchain Enlightenment and Smart City Cryptopolis. In Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems (CryBlock'18). ACM, New York, NY, USA, 48-53. DOI: <https://doi.org/10.1145/3211933.3211942>
 - [21] Hoang Tam Vo, Lenin Mehedy, Mukesh Mohania, and Ermyas Abebe. 2017. Blockchain-based Data Management and Analytics for Micro-insurance Applications. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17). ACM, New York, NY, USA, 2539-2542. DOI: <https://doi.org/10.1145/3132847.3133172>
 - [22] (2019). Apache Software Foundation [Online]. Available: <https://www.apache.org/>.
 - [23] (2019). Bitcoin Core [Online]. Available: <https://bitcoin.org/en/bitcoin-core/>.
 - [24] (2019). Bitcoin Official Website [Online]. Available: <https://bitcoin.org/en/>.
 - [25] (2018). Bitcoin Wiki [Online]. Available: https://en.bitcoin.it/wiki/Main_Page
 - [26] (2019). Bitcoin - Wikipedia [Online]. Available: <https://en.wikipedia.org/wiki/Bitcoin>.
 - [27] Blockchain.info [Online]. Available: <https://www.blockchain.com/>.

- [28] BlockCypher [Online]. Available: <https://live.blockcypher.com/btc/>.
- [29] BTC [Online]. Available: <https://btc.com/>.
- [30] CryptoID [Online]. Available: <https://chainz.cryptoid.info/>.
- [31] (2019). Cyprus - Wikipedia [Online].
Available: <https://en.wikipedia.org/wiki/Cyprus>.
- [32] (2018). Economy of Cyprus - Cyprus Profile [Online].
Available: <https://www.cyprusprofile.com/en/economy/>.
- [33] (2019). Economy of Cyprus - Wikipedia [Online].
Available: https://en.wikipedia.org/wiki/Economy_of_Cyprus.
- [34] Google Fusion Tables [Online]. Available: <https://fusiontables.google.com/>.
- [35] (2019). History of bitcoin - Wikipedia [Online]. Available:
https://en.wikipedia.org/wiki/History_of_bitcoin
- [36] (2019). IP Geolocation API [Online]. Available: <http://ip-api.com/>
- [37] (2019). IP Data for All Your Business Needs [Online]. Available:
<https://ipinfo.io/>
- [38] (2019). MariaDB.org [Online]. Available: <https://mariadb.org/>.
- [39] (2019). P2P (Peer to Peer) Definition - TechTerms [Online]. Available:
<https://techterms.com/definition/p2p>.
- [40] (2019). Peer-to-peer - Wikipedia [Online]. Available:
<https://en.wikipedia.org/wiki/Peer-to-peer>.
- [41] (2019). Public-key cryptography - Wikipedia [Online]. Available:
https://en.wikipedia.org/wiki/Public-key_cryptography.
- [42] Public key cryptography - IBM knowledge Center [Online]. Available:
https://www.ibm.com/support/knowledgecenter/en/SSB23S_1.1.0.13/gtps7/s7pkey.html.
- [43] SoChain [Online]. Available: <https://chain.so/>.
- [44] (2019). XAMPP Installers and Downloads for Apache Friends [Online].
Available: <https://www.apachefriends.org/index.html>.

Appendix A

1) Bash Script that is taking as an argument a bitcoin block hash, and then it downloads from the website “Blockchain.com” the data of the declared block, it calculates the next block hash and moves on to that block. The same procedure is executed N times, where N can be changed within the code by the user. Its output is a CSV file with the blocks’ details that is called “blockEntries.csv” and a text file that contains the transactions of those blocks.

```
1#!/bin/bash
2
3if [ $# -eq 0 ]; then
4    echo "Please provide a Bitcoin block hash"
5    exit
6elif [ $# -gt 1 ]; then
7    echo "You only have to provide a Bitcoin block hash"
8    exit
9fi
10
11BLOCKHASH=$1
12i=0
13
14while [ $i -lt 15000 ]; do
15    REQUEST="curl https://blockchain.info/rawblock/$BLOCKHASH"
16    $REQUEST > block.txt
17
18    BLOCKID=$(cat block.txt | tr "," "\n" | grep "\"height\"" | tr ":" "\n"
19              | grep -v "height")
20
21    if [ -z "$BLOCKID" ]; then
22        echo "Something went wrong. There are no details for block $BLOCKHASH"
23        echo "Next time start from block $BLOCKHASH"
24        exit
25    fi
26
27    SIZEBYTES=$(cat block.txt | tr "," "\n" | grep "\"size\"" | tr ":" "\n"
28               | grep -v "size" | head -1)
29    TEMP1=$(bc <<< "scale=3; $SIZEBYTES / 1000")
30    if [[ $TEMP1 == \.* ]]; then
31        SIZEKB="0$TEMP1"
32    else
33        SIZEKB="$TEMP1"
34    fi
35    BLOCKSIZE="$SIZEKB kB"
36    BLOCKVERSION=$(cat block.txt | tr "," "\n" | grep "\"ver\"" | tr ":" "\n"
37                  | grep -v "ver" | head -1)
38    PREVBLOCKHASH=$(cat block.txt | tr "," "\n" | grep "\"prev_block\""
39                   | tr ":" "\n" | grep -v "prev" | tr -d "\"")
40    MERKLEROOT=$(cat block.txt | tr "," "\n" | grep "\"mrkl_root\""
41                | tr ":" "\n" | grep -v "root" | tr -d "\"")
42    BLOCKTIME=$(cat block.txt | tr "," "\n" | grep "\"time\"" | tr ":" "\n"
43               | grep -v "time" | tr -d "\"" | head -1)
44    BLOCKDATETIME=$(TZ=UK date -d @"$BLOCKTIME" +%Y-%m-%d %H:%M:%S)
45    BITS=$(cat block.txt | tr "," "\n" | grep "\"bits\"" | tr ":" "\n"
46           | grep -v "bits")
47    TEMP=$(cat block.txt | tr "," "\n" | grep "\"nonce\"" | tr ":" "\n"
48          | grep -v "nonce")
```

```

49 if [[ $TEMP -lt "0" ]]; then
50     NONCE=$((4294967296+$TEMP))
51 else
52     NONCE="$TEMP"
53 fi
54 TXCOUNTER=$(cat block.txt | tr "," "\n" | grep "\"n_tx\"" | tr ":" "\n"
55             | grep -v "tx")
56
57 echo "$BLOCKID,$BLOCKSIZE,$BLOCKVERSION,$PREVBLOCKHASH,$MERKLEROOT,
58       $BLOCKDATETIME,$BITS,$NONCE,$TXCOUNTER,$BLOCKHASH" >> blockEntries.csv
59
60 cat block.txt | tr "," "\n" | grep "\"hash\"" | tail -n +2 | tr ":" "\n"
61 || grep -v "hash" | grep -v "tx" | tr -d "\"" >> transactions.txt
62
63 BLOCKHASH=$(cat block.txt | tr "," "\n" | grep "\"next_block\""
64             | tr ":" "\n" | grep -v "next" | tr -d "\"" | tr -d "[\"]")
65 i=$((i+1))
66 done
67 echo "Next time start from block $BLOCKHASH"

```

2) Bash Script that is taking as an argument a text file that contains a transaction hash in each row, and then it downloads from the website “Blockchain.com” the data of those transactions. Its output is three CSV files, the first one is called “txEntries.csv” and it contains the transactions’ details, the second one is called “inputEntries.csv” and it contains the details of the transactions’ inputs and the last one is called “outputEntries.csv” and it contains the details of the transactions’ outputs.

```

1 #!/bin/bash
2
3 if [ $# -eq 0 ]; then
4     echo "Please provide a file with a transaction hash in each line"
5     exit
6 elif [ $# -gt 1 ]; then
7     echo "You only have to provide a file with a transaction hash in each line"
8     exit
9 fi
10
11 file=$1
12 lines=$(cat $file)
13
14 for line in $lines; do
15     request="curl https://blockchain.info/rawtx/$line"
16     $request > tx.txt
17
18     txID="$line"
19     blockID=$(cat tx.txt | tr "," "\n" | grep "\"block_height\"" | tr ":" "\n"
20             | grep -v "block_height")
21
22     if [ -z "$blockID" ]; then
23         echo "Something went wrong. There are no details for transaction $txID"
24         continue
25     fi
26
27     version=$(cat tx.txt | tr "," "\n" | grep "\"ver\"" | tr ":" "\n"
28             | grep -v "ver")
29     inputCounter=$(cat tx.txt | tr "," "\n" | grep "\"vin_sz\"" | tr ":" "\n"
30                 | grep -v "vin_sz")
31     outputCounter=$(cat tx.txt | tr "," "\n" | grep "\"vout_sz\"" | tr ":" "\n"
32                    | grep -v "vout_sz")

```

```

33 time=$(cat tx.txt | tr "," "\n" | grep "\"time\"" | tr ":" "\n"
34 | grep -v "time")
35 txTime=$(TZ=UK date -d @"$time" +%Y-%m-%d %H:%M:%S')
36 relayed=$(cat tx.txt | tr "," "\n" | grep "\"relayed_by\"" | tr ":" "\n"
37 | grep -v "relayed_by" | tr -d "\"")
38 inputAddr=$(cat tx.txt | tr -d "\n" | sed -e 's/"out"/\n&/g'
39 | grep -v "\"out\"" | tr "," "\n" | grep "\"addr\"" | tr ":" "\n"
40 | grep -v "addr" | head -"$inputCounter" | tr -d "\""
41 | tr "\n" "/")
42 outputAddr=$(cat tx.txt | tr -d "\n" | sed -e 's/"out"/\n&/g'
43 | grep "\"out\"" | tr "," "\n" | grep "\"addr\"" | tr ":" "\n"
44 | grep -v "addr" | tail -"$outputCounter" | tr -d "\""
45 | tr "\n" "/")
46
47 echo "$txID,$blockID,$version,$inputCounter,$outputCounter,$txTime,$relayed,
48 $inputAddr,$outputAddr" >> txEntries.csv
49
50 i=0
51 while [ $i -lt "$inputCounter" ]; do
52     j=$((i+1))
53     prevOutIndex=$(cat tx.txt | tr -d "\n" | sed -e 's/"out"/\n&/g'
54 | grep -v "\"out\"" | tr -d "\n" | sed -e 's/"value"/\n&/g'
55 | grep "\"value\"" | sed -e 's/"n":/\n&/g'
56 | sed -e 's/,/"script"/\n&/g' | grep "\"n\"" | grep -v "}]")
57 | tr -d "\"n:" | head -"$j" | tail -1)
58     inputScript=$(cat tx.txt | tr -d "\n" | sed -e 's/"script"/\n&/g'
59 | grep "script" | grep "spent" | head -"$inputCounter"
60 | sed -e 's/{/"sequence"/\n&/g' -e 's/]/.\n"weight"/\n&/g'
61 | grep "script" | tr ":" "\n" | grep -v "script" | tr -d "\"}"
62 | head -"$j" | tail -1)
63     inputSequence=$(cat tx.txt | tr -d "\n" | sed -e 's/"sequence"/\n&/g'
64 | sed -e 's/"witness"/\n&/g' | grep "sequence" | tr ":" "\n"
65 | grep -v "sequence" | tr -d "," | head -"$j" | tail -1)
66
67     echo "$txID,$prevOutIndex,$inputScript,$inputSequence">>inputEntries.csv
68     i=$((i+1))
69 done
70
71 i=0
72 while [ $i -lt "$outputCounter" ]; do
73     j=$((i+1))
74     outputIndex=$(cat tx.txt | tr -d "\n" | sed -e 's/"out"/\n&/g'
75 | grep "\"out\"" | sed -e 's/"value"/\n&/g' | grep "\"value\""
76 | sed -e 's/"n":/\n&/g' | sed -e 's/,/"script"/\n&/g'
77 | grep "\"n\"" | grep -v "]" | tr ":" "\n" | grep -v "n"
78 | head -"$j" | tail -1)
79     valueBTC=$(cat tx.txt | tr -d "\n" | sed -e 's/"out"/\n&/g'
80 | grep "\"out\"" | sed -e 's/"value"/\n&/g' | grep "\"value\""
81 | sed -e 's/"n":/\n&/g' | grep "value" | tr ":" "\n"
82 | tr -d "," | grep -v "value" | head -"$j" | tail -1)
83     value=$(bc <<< "scale=10; $valueBTC / 100000000")

```

3) Java code that is taking as an argument a text file that contains an IP address in each row, and then it downloads from the website “IP Geolocation API” the geo-location information associated with each IP. Its output is a CSV file that is called “IPs.csv” and it contains that geo-location information.

```
1 import java.io.BufferedReader;
2 import java.io.File;
3 import java.io.FileReader;
4 import java.io.FileWriter;
5 import java.io.InputStreamReader;
6 import java.io.PrintWriter;
7 import java.io.Writer;
8 import java.net.URL;
9 import java.net.URLConnection;
10
11 public class IPcurlRequest {
12
13     public static void main(String [] args){
14
15         if(args.length != 1){
16             System.err.println("Please provide a file with an IP in each row");
17             System.exit(1);
18         }
19
20         try {
21             File input = new File(args[0]);
22             BufferedReader in1 = new BufferedReader(new FileReader(input));
23             String ip;
24             int count = 0;
25             while((ip = in1.readLine()) != null) {
26                 count++;
27             }
28             String [] ipTable = new String[count];
29             in1.close();
30
31             in1=new BufferedReader(new FileReader(input));
32             int k = 0;
33             while((ip = in1.readLine()) != null) {
34                 ipTable[k] = ip;
35                 k++;
36             }
37
38             String fileName = "IPs.csv";
39             File IPs = new File(fileName);
40             if(IPs.exists() == false)
41                 IPs.createNewFile();
42             else {
43                 IPs.delete();
44                 IPs.createNewFile();
45             }
46             Writer fileWriter = new FileWriter(fileName, true);
47             PrintWriter printWriter = new PrintWriter(fileWriter);
48             String country = null, region = null, city = null;
49             String latitude = null, longitude = null;
50
```

```

51     for(int i = 0; i < ipTable.length; i++){
52         String fields = "?fields=country,regionName,city,lat,lon";
53         String req = "http://ip-api.com/line/";
54         URL url = new URL(req + ipTable[i] + fields);
55         URLConnection con = url.openConnection();
56         BufferedReader in = new BufferedReader(new
57             InputStreamReader(con.getInputStream()));
58
59         String inputLine;
60         int j = 0;
61         while ((inputLine = in.readLine()) != null) {
62             switch(j){
63                 case 0:
64                     country = inputLine;
65                     break;
66                 case 1:
67                     region = inputLine;
68                     break;
69                 case 2:
70                     city = inputLine;
71                     break;
72                 case 3:
73                     latitude = inputLine;
74                     break;
75                 case 4:
76                     longitude = inputLine;
77                     j = -1;
78                     break;
79             }
80             j++;
81         }
82         in.close();
83
84         printWriter.println(ipTable[i] + "," + country + "," + region +
85             "," + city + "," + latitude + "," + longitude);
86         if(i%2==0){
87             Thread.sleep(800);
88         }
89     }
90     printWriter.close();
91 }catch(Exception e){
92     e.printStackTrace();
93 }
94 }
95 }

```