University of Nevada, Reno

COMPLEX NETWORK ANALYSIS OF CRYPTO CURRENCIES

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of Master of Science in
Computer Science and Engineering

by
Manoj Kumar Popuri
Dr. Mehmet Hadi Gunes / Thesis Advisor
December 2015

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

MANOJ KUMAR POPURI

Entitled

Complex Network Analysis Of Crypto Currencies

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

Mehmet H. Gunes, Ph.D., Advisor

Ming Li, Ph.D., Committee Member

Gokhan Pekcan, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

December, 2015

ABSTRACT

Analysis of the traditional currencies is not easy as the transactions are not centralized but rather take place over a large number of banks and commercial entities. Digital crypto currencies, however, require a public ledger to work. A crypto currency is a medium of exchange using cryptography to secure the transactions and to control the creation of new units. In this thesis, we analyze some of the popular crypto currencies. As the transaction data of crypto currencies are publicly available, we construct a network of transactions and extract the time and date of each payment for the analyzed crypto currencies. We investigate the structure of transaction network by measuring the network characteristics. In particular, we compare the evolution of Bitcoin and Litecoin currency systems, two of the currently most popular systems; analyze the wealth correlation with degree distribution for Bitcoin and litecoin; and investigate the transactions by the top 100 richest people in Bitcoin, Litecoin, Dash, Dogecoin, Peercoin, and Namecoin crypto currencies. Additionally, as the price of digital currencies are highly volatile, we perform a regression analysis on factors that affect the price of the Bitcoin currency in USD and derive a model with the factors that affects Bitcoin price.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Dr. Mehmet Hadi Gunes for his encouragement, support, and all the help throughout the thesis work. I would also like to thank my parents Chowdeswari and Raghavayya for their encouragement and sacrifices for my career and for believing in me.

Table of Contents

Abstract i
Acknowledgements ii
Table of Contents iii
List of Tables v
List of Figures vi

1 Introduction **1**

2 Background **6**
2.1 Bitcoin 6
2.1.1 Bitcoin Transactions 8
2.1.2 Bitcoin Market Value 10

3 Complex Network Analysis **12**
3.1 Tools 13
3.1.1 Stanford Network Analysis Platform (SNAP) 13
3.1.2 Neo4j 13
3.2 Bitcoin Network 14
3.2.1 Degree 14
3.2.2 Assortativity 21
3.2.3 Clustering 22
3.2.4 Anonymity 23
3.3 Litecoin Network 25
3.3.1 Degree 25
3.3.2 Assortativity 30
3.3.3 Clustering 31

4 Market Analysis **32**
4.1 Richest Bitcoin Addresses 32
4.2 Richest Litecoin Addresses 34
4.3 Richest Dash Addresses 35
4.4 Richest Dogecoin Addresses 36
4.5 Richest Peercoin Addresses 36
4.6 Richest Namecoin Addresses 37

5 Regression Analysis **39**
5.1 Regression Model with Market Capitalization Variable 40
5.2 Regression Model with All Predictor Variables - Full Model 43
5.3 Regression Model with Significant Predictor Variables - Reduced Model 47
5.3.1 Comparison of Full and Reduced Models 49
5.3.2 Model with Miners Revenue 49
5.3.3 SetpAIC Analysis 51
5.4 Validation of Assumptions 53
5.5 Variable Selection 54

6 Conclusion **56**

7 Future Work **58**

Bibliography **59**

List of Tables

3.1 In degree characteristics of yearly Bitcoin transactions 17
3.2 Out degree characteristics of yearly Bitcoin transactions 17
3.3 In degree characteristics of yearly Litecoin transactions 28
3.4 Out degree characteristics of yearly Litecoin transactions 28

List of Figures

1.1 Market capitalization of crypto currencies 2

2.1 Anatomy of a bitcoin transaction 8
2.2 Bitcoin transactions in detail 9
2.3 Bitcoin market price - Historical 10
2.4 Bitcoin market price - Last 6 months 11

3.1 PDF of in degree distribution of the Bitcoin transaction network for all transactions until Jan 2015. 15
3.2 CCDF of in degree distribution of the Bitcoin transaction network for all transactions until Jan 2015. 15
3.3 PDF of out degree distribution of the Bitcoin transaction network for all transactions until Jan 2015. 16
3.4 CCDF of out degree distribution of the Bitcoin transaction network for transactions until Jan 2015. 16
3.5 In degree distribution of the yearly Bitcoin transaction network. . . 18
3.6 In degree distribution of the Bitcoin transaction network - Combined. 19
3.7 Out degree distribution of the Bitcoin transaction network - Combined. 19
3.8 Out degree distribution of the Bitcoin transaction network. 20
3.9 In degree as a function of out degree $K_n^{in}(K_{out})$ for Bitcoin transactions. 21
3.10 Assortativity coefficients of the yearly Bitcoin transactions. 22
3.11 Clustering coefficient of yearly Bitcoin transactions. 23
3.12 Anonymity among MyWallet users 24
3.13 In degree distribution of the Litecoin transactions until Jan 2015. . . 26
3.14 CCDF of in degree distribution of the Litecoin transactions until Jan 2015. 26
3.15 Out degree distribution of the Litecoin transactions until Jan 2015. . 27
3.16 CCDF of out degree distribution of the Litecoin transactions until Jan 2015. 27
3.17 In degree distribution of the yearly Litecoin transactions. 29
3.18 Out degree distribution of the yearly Litecoin transactions 29
3.19 In degree as the function of out degree $K_n^{in}(K_{out})$ for Litecoin transactions. 30

3.20 Assortativity coefficients of the yearly Litecoin transactions. 31
3.21 Clustering of the network. 31

4.1 Percentage of Bitcoin wealth, the richest own during Dec 2014. . . . 33
4.2 Transaction pattern of the richest Bitcoin node. 33
4.3 In and out degree of the richest 100 Bitcoin nodes. 34
4.4 In and out degree of the richest 100 Litecoin nodes. 35
4.5 In and out degree of the richest 100 Dashcoins nodes. 35
4.6 In and out degree of the richest 100 Dogecoin nodes. 36
4.7 In and out degree of the richest 100 Peercoin nodes. 37
4.8 In and out degree of the richest 100 Namecoin nodes. 38

5.1 Regression analysis with market capitalization 41
5.2 Plot of market price and market capitalization 42
5.3 Regression analysis without market capitalization 43
5.4 Normal QQ plot for full model 44
5.5 Plot of the full model 46
5.6 Plot of reduced model 48
5.7 Plot of linear model with miners revenue 51
5.8 Normal QQ plot for reduced model 54

Chapter 1

Introduction

Currency is a medium of exchange, which aroused out of need to address the inefficiency of barter. Digital currency is a form of currency that is electronically created and stored [52]. The initial digital currencies were non crypto currencies, i.e., E-Gold in 1996 [19]. Crypto currencies are decentralized digital cash systems and there is no single overseeing authority [46]. In crypto currencies, encryption techniques are used to regulate the generation of units of currency and verify the transfer of funds [50].

The first public crypto currency is Bitcoin, proposed in 2008 by Satoshi Nakamoto, a pseudonym [41]. Even though the system went online in January 2009, Bitcoin had very few users and it didn't have any real world value for one year. Since the inception of the Bitcoin, over 48 million transactions took place. The market value of Bitcoins in circulation peaked at about 14 billion dollars on May 12, 2013, and as of Nov 23, 2015 is about 4.79 billion dollars. Figure 1.1 compares Bitcoin market capitalization with other popular currencies.

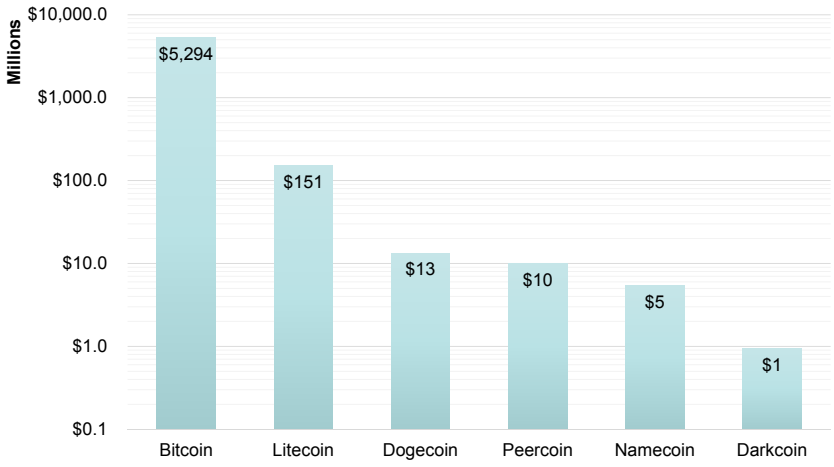

Figure 1.1: Market capitalization of crypto currencies

A problem in digital currencies to verify that the owner did not double-spend the coin. A common solution is to introduce a trusted central authority, or mint, that checks every transaction for double spending. After each transaction, the coin must be returned to the mint to issue a new coin, and only coins issued directly from the mint are trusted not to be double-spent. The problem with this solution is that the fate of the entire money system depends on the company running the mint, with every transaction having to go through them, just like a bank. We need a way for the payee to know that the previous owners did not sign any earlier transactions. For our purposes, the earliest transaction is the one that counts, so we don't care about later attempts to double-spend. The only way to confirm the absence of a transaction is to be aware of all transactions.

In the mint based model, the mint is aware of all transactions and decides which arrived first. To accomplish this without a trusted party, transactions must be publicly announced, and we need a system for participants to agree on a single history of the order in which they were received. The payee needs proof that at the time of each transaction, the majority of nodes agreed it was the first received. In case of the Bitcoin, Litecoin, Dogecoin, Dash, and Monero every payment is announced

on the network, and the payment is validated by checking consistency with the entire transaction history. To avoid fraud, it is necessary that the participants agree on a single valid transaction history. In case of the Nxt and Maidsafecoin, the distribution of the coins is decided by a central authority [33].

The Bitcoin system operates as an online peer-to-peer network, and anyone can join the system by installing a client application [41]. Instead of having a bank account maintained by a central authority, each user has a unique address that consists of a pair of public and private keys. Existing coins are associated to the public key of their owner, and outgoing payments have to be signed by the owner using the corresponding private key. After validation of transaction with the owner's public key, the successful transactions are formed into blocks at an approximate rate of 1 block per 10 minutes. To maintain privacy, a single user may use multiple addresses. Each participating node stores the complete list of previous transactions. New Bitcoins are awarded to the users based on their contribution in the generation of new blocks, a process known as mining. Another way to obtain coins is to purchase them from someone who already has coins using traditional currency.

The transactions of all the crypto currencies are available to anyone by installing the client and connecting to peer to peer network. Such detailed information is rarely available in financial systems, making the the crypto currency networks a valuable source of empirical data involving monetary transactions. Due to the anonymity of the crypto currencies and potentially unlimited number of pseudo identities a user could generate, however, it is hard to determine which observed phenomena are specific to the system and which results can be generalized.

An earlier study by Daniel et.al. analyzes the Bitcoin transaction network [24]. The main purpose of authors was to investigate the movement of money in detail and observe the dynamics of the network. In their analysis of Bitcoin data on May 7th 2013, they observe 17 million transactions among 13 million addresses where only a million of them had nonzero balance. According to their analysis there is a strong correlation between the balance and the indegree of individual nodes. They found that the Bitcoin network is gradually increasing since 2010 with some fluctuations, e.g., the boom in the exchange rate in 2011. According to their analysis both the in-degree and out-degree are highly heterogeneous with power law distributions. They also found that Bitcoin network is disassortative as except for only a brief period in the initial deployment where the number of nodes were few.

The study of networks has emerged in diverse disciplines as a means of analyzing complex relational data [45]. Network analysis has been applied to physical phenomena [56], biological systems [32], epidemics [6, 7], academic collaborations [5, 30, 31, 37], language [12], news media [48, 49], software development [13, 59], transportation [10], industry [40], online social networks [14, 15, 42], communications [25, 26], Internet [21, 23, 22, 28], synthetic topologies [2, 3, 8], visualization [51] and graph mining [27, 29].

In this thesis, we compare various crypto currencies as a network, by analyzing their complex relational data. We map the transaction network of Bitcoin and Litecoin digital currencies from their public transaction data and analyze the complex network of each digital currency. We download the complete list of transactions by installing the wallets of the digital currency, where the nodes represents the unique address with each user and the links are the transactions in between two

users. The number of inputs and outputs in each transaction vary as single user can create multiple addresses and can transfer to multiple address [16].

Then, we performed market analysis of richest users of the six popular crypto currencies, which includes Bitcoin, Litecoin, Dash, Dogecoin, Peercoin, and Namecoin [17]. At the time of this analysis, these six currencies occupies over 98 % of the total market capitalization of all 621 crypto currencies combined. In market analysis, we collect the incoming and outgoing transactions of the top 100 richest people in each crypto currency and formed a network of the top 100 richest address for each currency [55].

Finally, we collect the past six months data of the Bitcoin price in USD as shown in Figure 2.4 and all the factors that effects the price such as, miners revenue in USD, total transaction volume per day in USD, total outgoing transactions per day in USD, exchange trade value in USD, market capitalization in USD, transaction fees in USD, and cost per transaction in USD. We analyze this data to perform a multiple linear regression on the predictor variable to find the exact factors effecting Bitcoin price [58].

Chapter 2

Background

In this chapter, we present a summary of how Bitcoin network works. Other digital currencies have similar mechanisms.

2.1 Bitcoin

Bitcoin is an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Bitcoin transactions are computationally impractical to reverse and that would protect sellers from fraud. Bitcoin transactions are saved in blockchain, a single data file participants pass around to each other. The Blockchain is the fundamental data structure of the Bitcoin protocol, and it allows users to know who owns the currency. Anyone can perform transaction given they prove they own the Bitcoins and others can mathematically verify the transaction to ensure it's validity.

Bitcoin is run by 4 important rules:

1. *Governance* – an open source community of developers backed by the Bitcoin Foundation.
2. *Democratic* – if you don't like one of the changes, you are more than welcome to fork the chain and implement your own rules.
3. *Money Creation* – is given to the people, not to the central bankers.
4. *Deflationary by design* – money supply cannot be manipulated and is fixed at 21 million coins, each divisible up to 8 decimal.

The users get a public key and a private key by installing the client. While transferring the coin owner digitally signs the hash of previous transaction and the public key of the next owner and adds these to the end of the coin. The hash of the previous transaction is to verify the transaction and it can only decoded by the next owner as it was encrypted with the next owner's public key. The number of inputs and outputs in each transaction vary as single user can create multiple addresses and can transfer to multiple address.

The number of coins generated and distributed vary for each crypto currency, but the procedure for generation and distribution is same for all peer-to-peer crypto currencies. The coins are generated per block and the number of coins generated decreases with time while the difficulty in generation increases. The new coins are awarded to the users participating in the mining process, where users offer their computing power to solve the hash problem for generation of new blocks and adding verified transactions to the public leaguer.

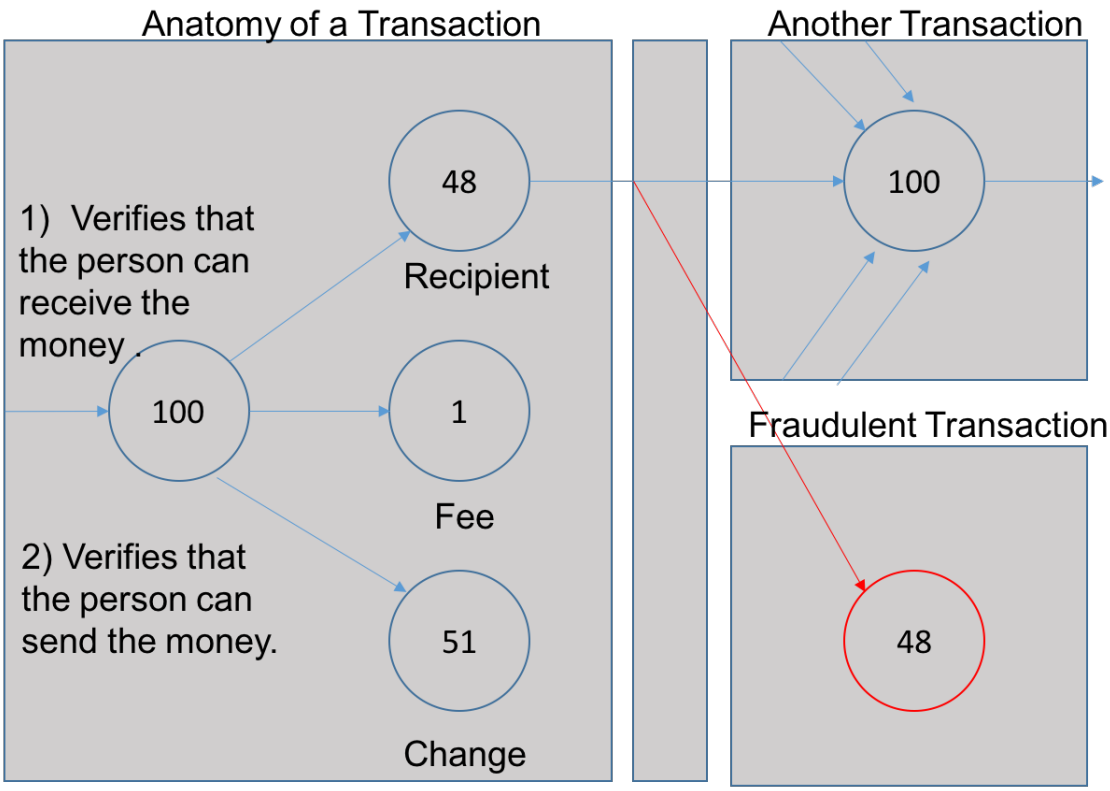

Figure 2.1: Anatomy of a bitcoin transaction

2.1.1 Bitcoin Transactions

The electronic coin is a chain of digital signatures. The input to a Bitcoin transaction contains the public key of the redeemer of the output transaction. The current owner should sign the transaction with the hash of the previous transaction. As indicated in Figure 2.1, output contains the actual amount that is being sent to the recipient, the change amount being sent back to the original sender, and the voluntary transaction fee attached to the output. The block chain prevents the double spend attack by giving other nodes the power to verify that transaction inputs were not already spent somewhere else.

Each owner transfers the coin to the next by digitally signing a hash of the

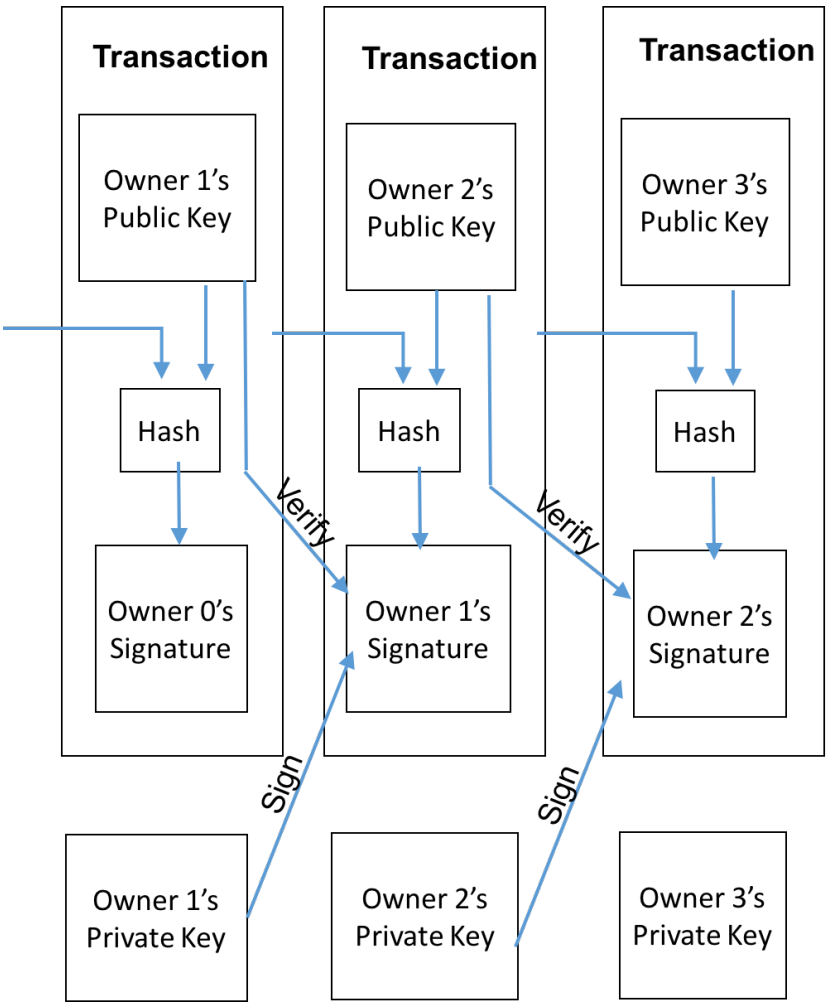

Figure 2.2: Bitcoin transactions in detail

previous transaction and the public key of the next owner and adding these to the end of the coin as shown in Figure 2.2. A payee can verify the signatures to verify the chain of ownership.

The new transactions are announced by the users on the network, and these transactions are formed into blocks at time varying for different crypto currencies. These blocks from the block-chain, where each block references the previous block. Hence, double spending, i.e., spending the money more than once, require the re-computation of previous blocks. Such double spending attack can be possible only

if the attacker has greater than 50 percentage of the hash rate of the whole, which is practically infeasible with large scale deployment of these networks. The block generation time for Bitcoin is approximately one block per 10 minutes, the block generation for Litecoin is approximately one block per 2.5 minutes, and the block generation time for all other currencies is around 4-6 minutes per one block; with a varying number of transactions per block for each crypto currency.

2.1.2 Bitcoin Market Value

The price of Bitcoin is highly volatile [17]. Starting from 2009 we can divide the price into three groups, in 2010 Bitcoin gained real dollar value and at that time 1 Bitcoin was around 0.01 dollars. By mid 2011 Bitcoin gained public attraction due to various reasons and the market value started increasing drastically [11]. This public inclination in price continued until 2013, at that time 1 Bitcoin was about 1200 USD and then from 2013 till now the market price is unstable. As of Dec 1, 2015, 1 Bitcoin is around 350 dollars, the three phases can be observed in Figure 2.3.

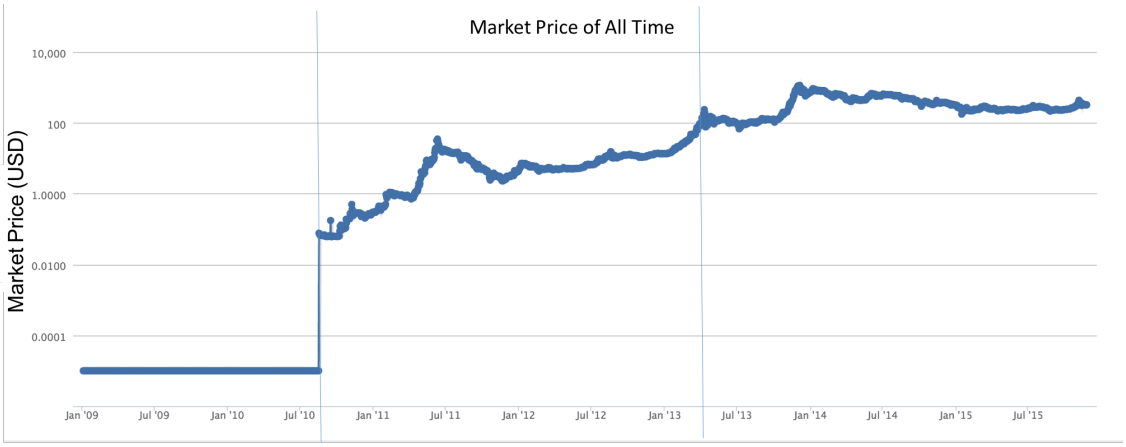

Figure 2.3: Bitcoin market price - Historical

Figure 2.4: Bitcoin market price - Last 6 months

Chapter 3

Complex Network Analysis

In the context of network theory, a complex network is a graph with non-trivial topological features that do not occur in simple networks such as lattices or random graphs but often occur in graphs modelling real systems [39, 45].

The crypto currencies are based on a peer to peer network connected through the Internet. The transactions are validated based on the proof of work system where each node stores the list of all previous transactions. The transactions of the crypto currencies are available to anyone by installing the client and connecting to the peer to peer network. Such detailed information is rarely available in financial systems, making the crypto currency networks a valuable source of empirical analysis of monetary transactions. However, as these networks are anonymous and each user can create unlimited number of addresses, which appears as separate nodes while analyzing the network, it is hard to determine which observed phenomena are specific to the crypto currency system, and which results are general for commercial transactions [20].

3.1 Tools

As the data that we are analyzing is massive, we need the tools which can effectively analyze the massive data up to millions of nodes. So we are using the following tools:

3.1.1 Stanford Network Analysis Platform (SNAP)

Stanford Network Analysis Platform (SNAP) is a general purpose network analysis and graph mining library [57]. It is written in C++ and easily scales to massive networks with hundreds of millions of nodes, and billions of edges. It efficiently manipulates large graphs, calculates structural properties, generates regular and random graphs, and supports attributes on nodes and edges.

3.1.2 Neo4j

Neo4j is a robust transactional property graph database [44]. Due to its graph data model, Neo4j is highly agile and fast. Neo4j scales up and out, supporting tens of billions of nodes and relationships, and hundreds of thousands of transactions per second. Distributed across multiple machines, Neo4j uses a Graph query language Cypher whose syntax provides a familiar way to match patterns of nodes and relationships in the graph [43].

3.2 Bitcoin Network

We downloaded the Bitcoin data-set from casejobs web database interface of the *Do the rich get Richer* project by Daniel Kondor [], and decoded the data collected from the wallet. We crawled `Blockchain.info` and `Bitinfocharts.com` for daily transaction data and the richest node data.

In our network, the nodes are the addresses assigned to Bitcoin users and the edges are the transaction between two nodes. The network we are analysing is comprising of N = 49,390,594 nodes, total incoming transactions Ein = 151,933,127, and the outgoing transactions Eout = 151,857,042 edges. We also divide the network in different points of time to study the growth of the network over the years.

Bitcoin network is a growing network where the number of unique addresses created every year increases exponentially. The major increase in the number of unique addresses occurred after the first boom in 2011 and the second one when the Bitcoin market value crossed 1000 USD.

3.2.1 Degree

A network can be an exceedingly complex structure, as the connections among the nodes can exhibit nontrivial patterns [4]. To study a network, we need to develop simplified measures that reflects the network characteristics in an understandable way. The degree distribution captures the underlying structure of a network by summarizing the degree characteristics of the nodes [47];

$$P(K) = \binom{n-1}{K} P^K (1-P)^{n-1-K}$$

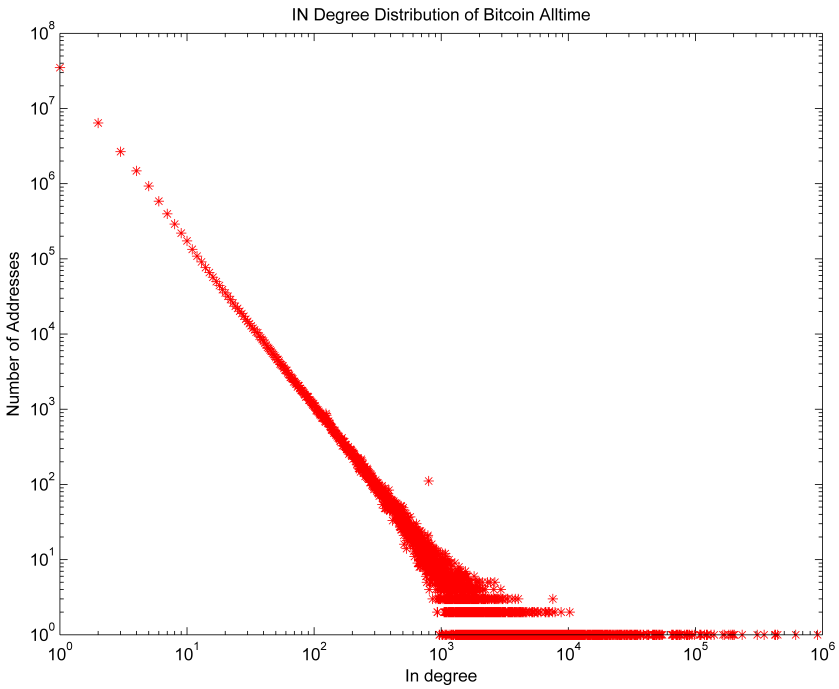

Figure 3.1: PDF of in degree distribution of the Bitcoin transaction network for all transactions until Jan 2015.

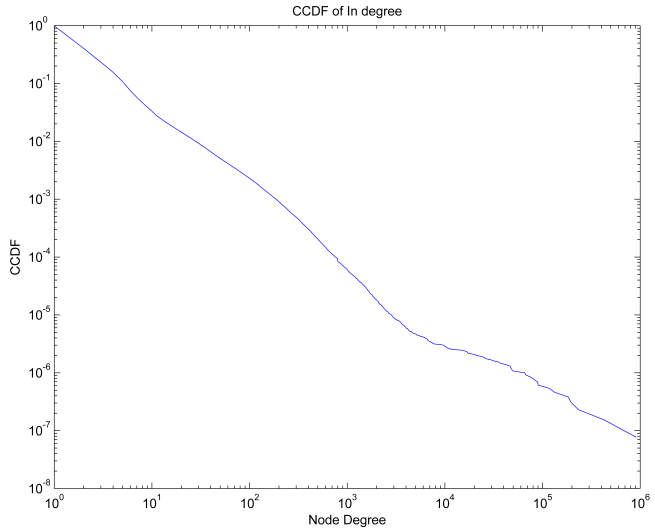

Figure 3.2: CCDF of in degree distribution of the Bitcoin transaction network for all transactions until Jan 2015.

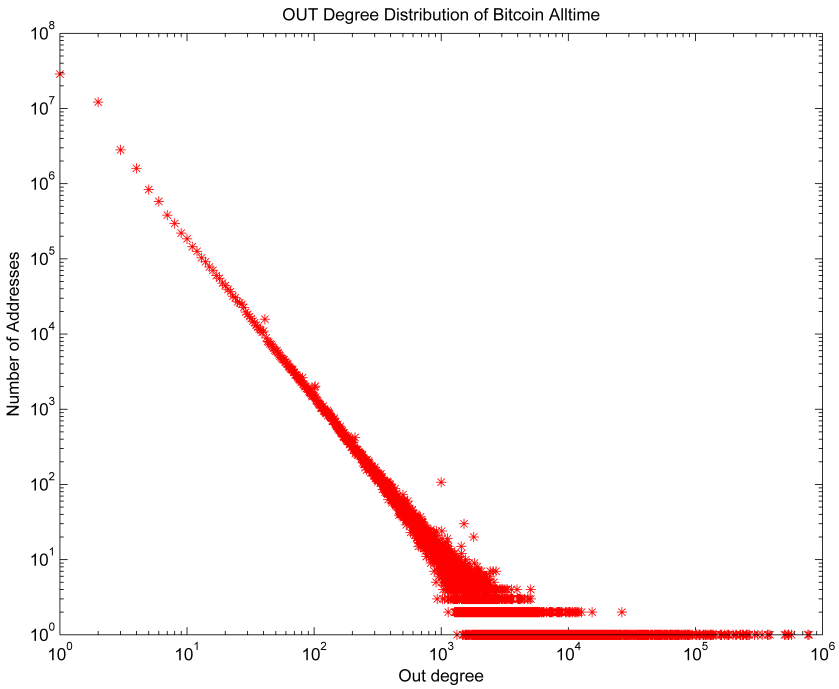

Figure 3.3: PDF of out degree distribution of the Bitcoin transaction network for all transactions until Jan 2015.

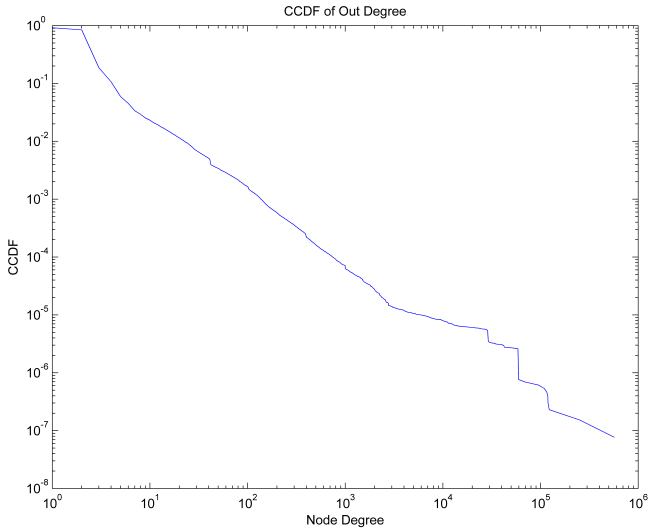

Figure 3.4: CCDF of out degree distribution of the Bitcoin transaction network for transactions until Jan 2015.

Figure 3.1 and Figure 3.2 present the in degree distribution of the Bitcoin transactions. Similarly, Figure 3.3 and Figure 3.4 present the out degree distribution of the Bitcoin transactions. We find that the degree distributions follow a power law, which makes Bitcoin network a scale free network, for both in degree and out degree distributions. The power laws of the degree distributions are $\alpha_{in}) \sim -2.21$ and $\alpha_{out} \sim -2.10$.

To understand the evolution of the network, we calculated the degree distributions of the network yearly. Table 3.1 and Table 3.2 present the characteristics of the yearly Bitcoin transactions. We observe that there is a huge variation in degree distribution from 2009 to 2011, but thereafter the distribution is more stable while the number of the nodes is considerably increasing.

Table 3.1: In degree characteristics of yearly Bitcoin transactions

Year	α	*Nodes*	*Edges*	*MaxDegree*	*AvgDegree*
2009	1.94	32,699	98,611	,1,257	3.01
2010	2.00	122,167	374,712	1,826	3.07
2011	2.13	1,610,899	5,198,488	118,016	3.23
2012	2.14	3,780,767	14,570,562	913,847	3.85
2013	2.19	5,082,351	16,338,332	16,969	3.21
2014	2.21	38,761,711	115,352,422	636,092	2.98

Table 3.2: Out degree characteristics of yearly Bitcoin transactions

Year	α	*Nodes*	*Edges*	*MaxDegree*	*AvgDegree*
2009	1.78	32,699	95,499	1,528	2.92
2010	1.83	122,167	478,271	5,8829	3.19
2011	1.88	1,610,899	5,461,888	59,297	3.39
2012	1.90	3,780,767	14,130,630	570,898	3.73
2013	1.92	508,235	16,442,626	69,919	3.23
2014	2.10	38,761,711	116,241,889	1,765,959	3.01

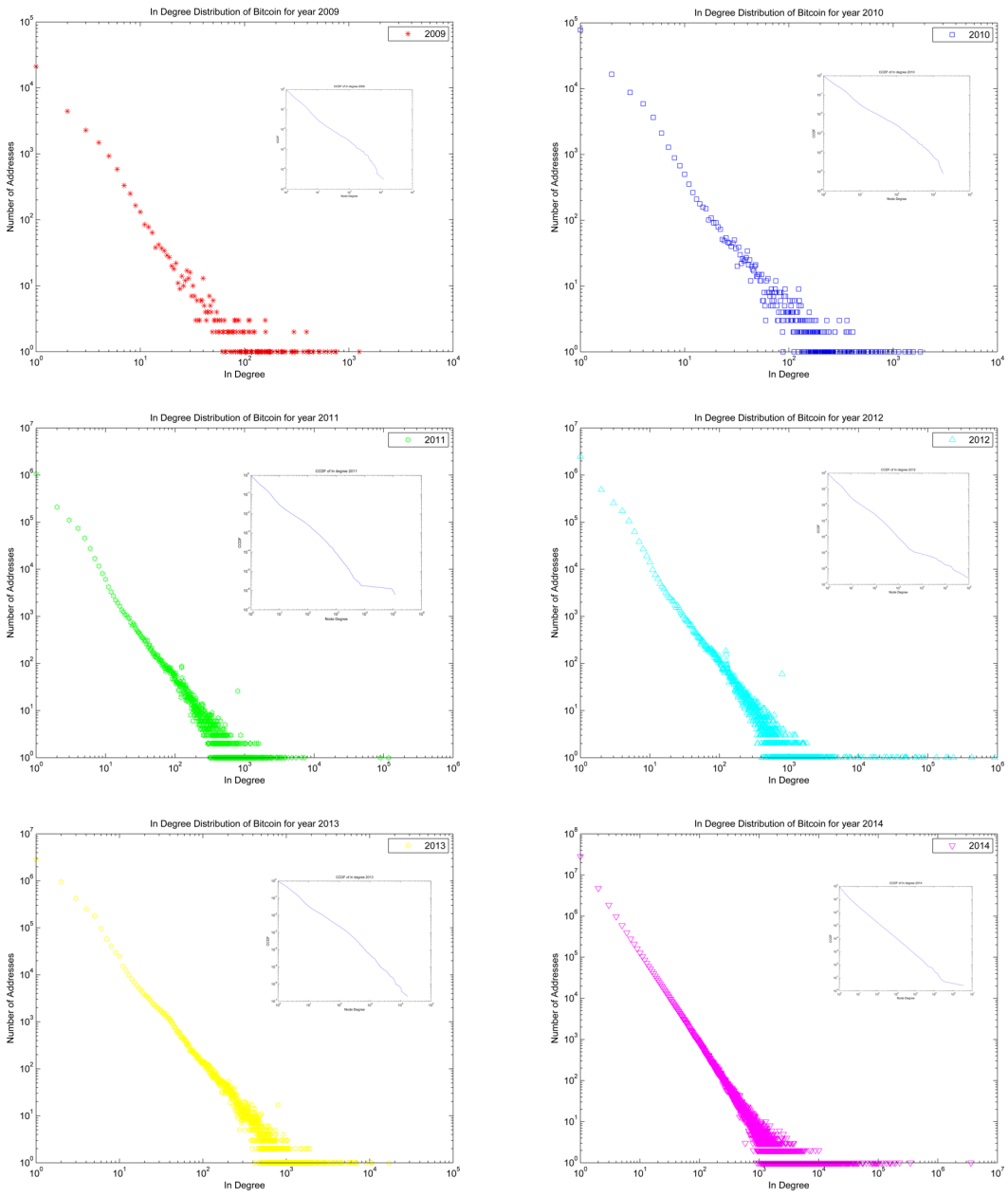

Figure 3.5: In degree distribution of the yearly Bitcoin transaction network.

Figure 3.5 and Figure 3.6, present in degree distributions of transactions per each year.

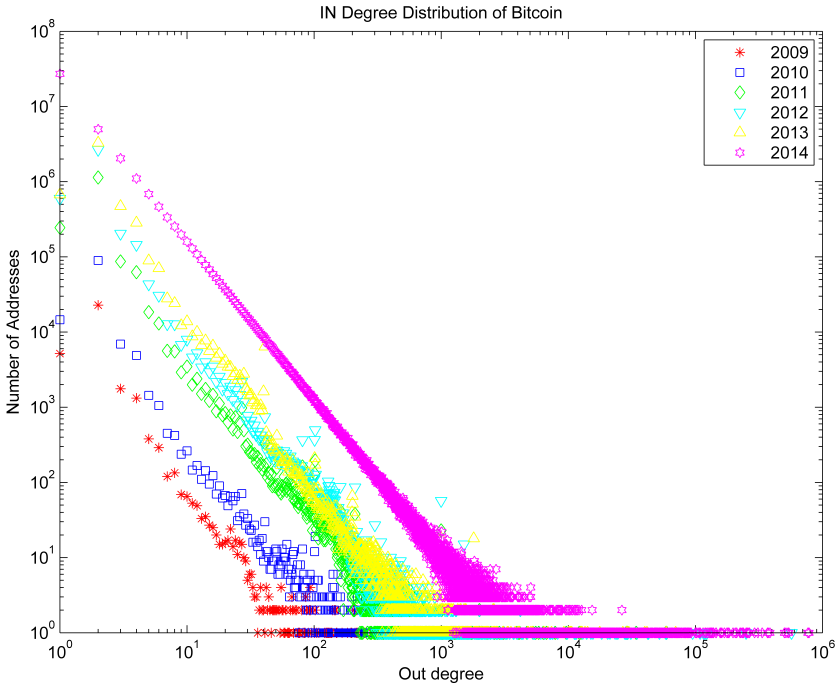

Figure 3.6: In degree distribution of the Bitcoin transaction network - Combined.

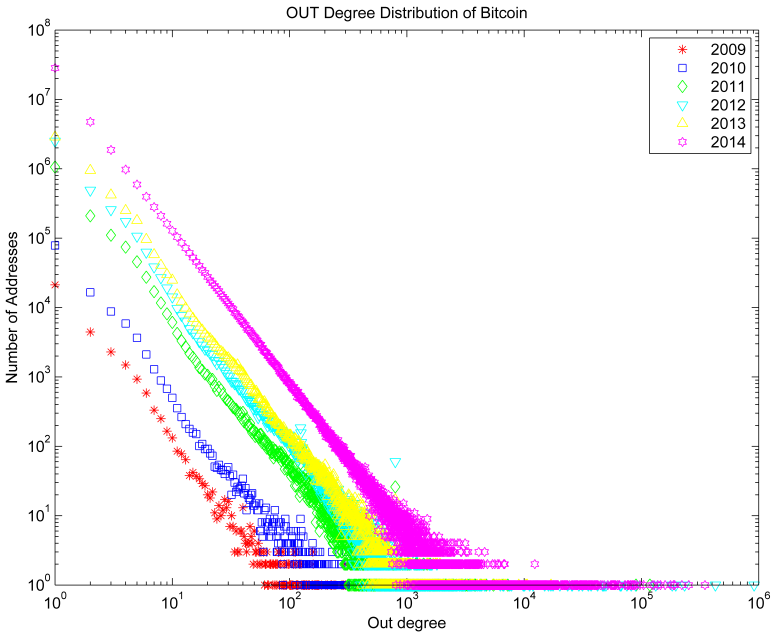

Figure 3.7: Out degree distribution of the Bitcoin transaction network - Combined.

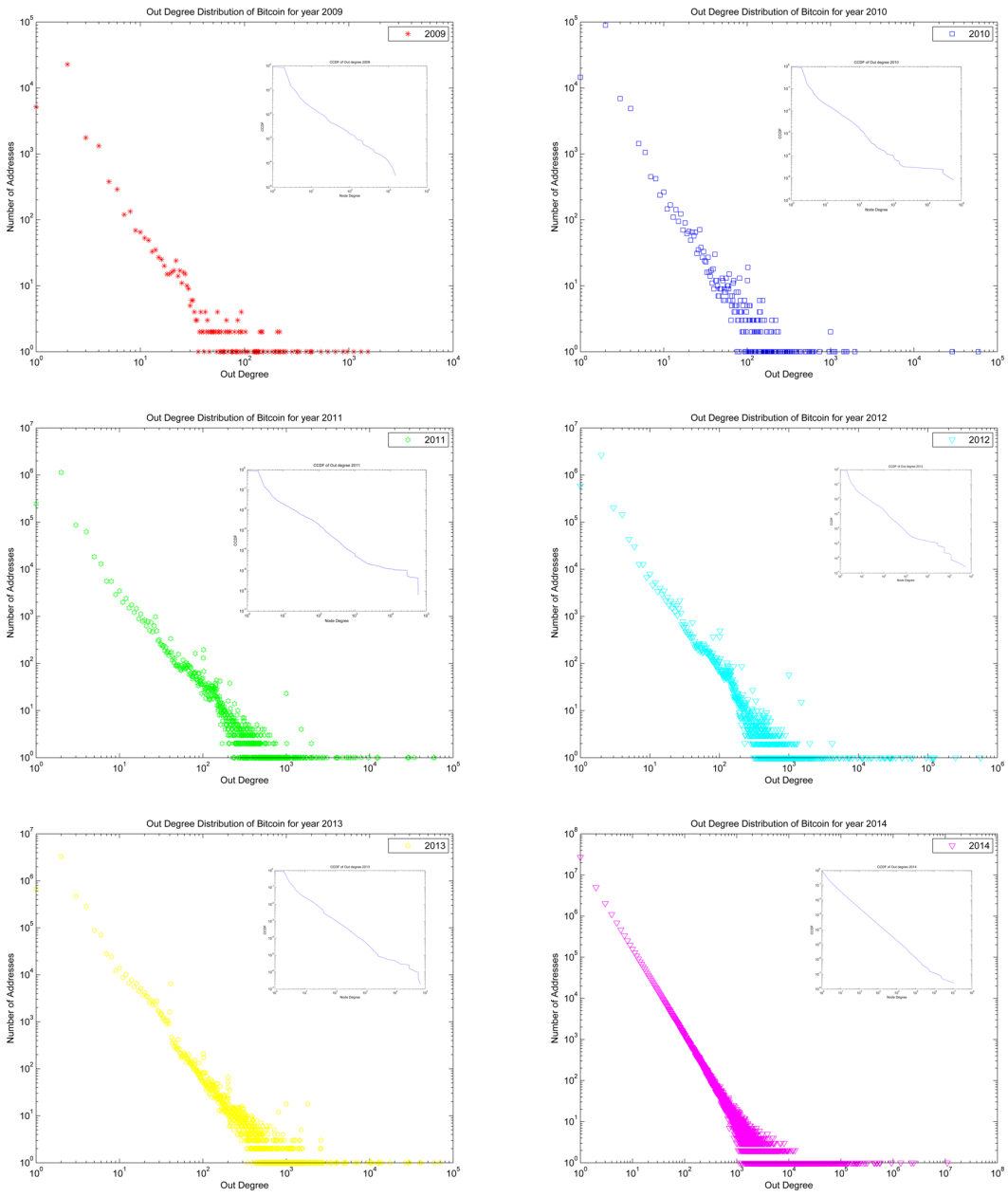

Figure 3.8: Out degree distribution of the Bitcoin transaction network.

Figure 3.7 and Figure 3.8, present out degree distributions of transactions per each year.

3.2.2 Assortativity

We computed the nearest neighbour degree function $K_n^{in}(K_{out})$, which measures the in degree K_{in} of the nodes with respect to out degree K_{out}. Figure 3.9 presents the degree correlations for the Bitcoin network. In the graph, we observe that here is a disassortative behaviour between the In and out Degrees of the nodes. That is, the nodes with high out degree tend to connect to the node with low in degree.

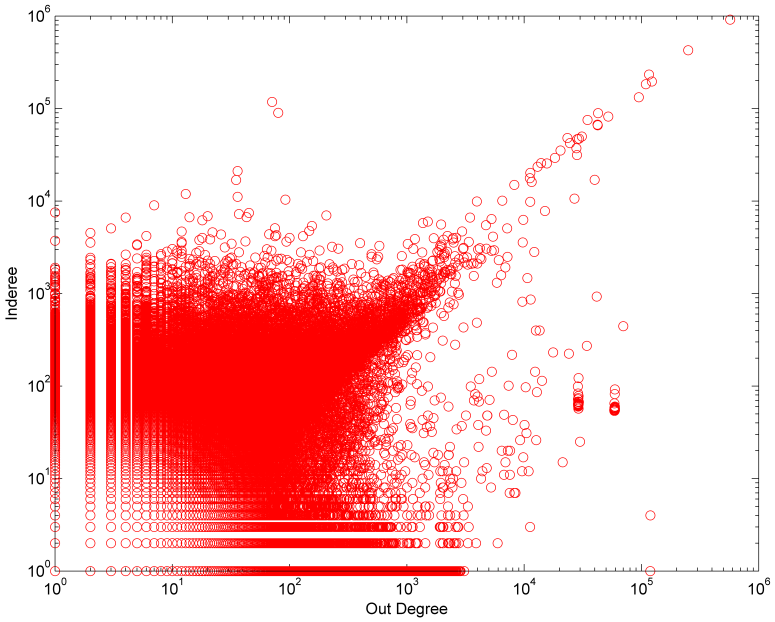

Figure 3.9: In degree as a function of out degree $K_n^{in}(K_{out})$ for Bitcoin transactions.

The assortativity coefficient is the Pearson correlation coefficient of degree between pairs of linked nodes. The assortativity coefficient is calculated as:

$$r = \Sigma_j k \frac{jk(e_j k - q_j q_k)}{\frac{2}{q}}$$

Positive values of r indicate a preference to link between nodes of similar degree, while negative values indicate relationships between nodes of different de-

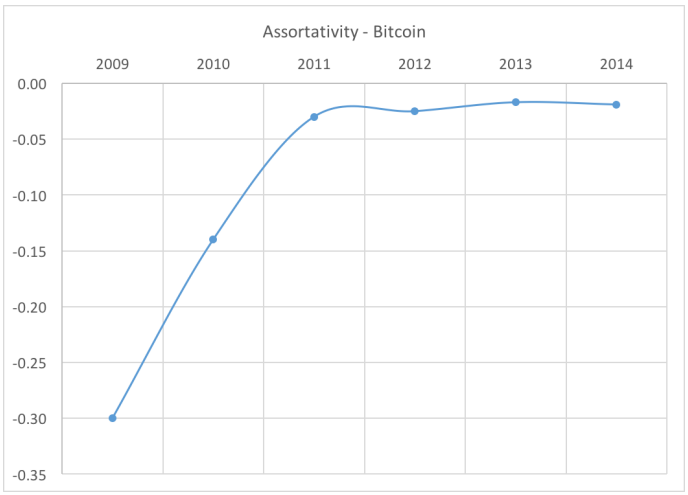

Figure 3.10: Assortativity coefficients of the yearly Bitcoin transactions.

gree. The value of r lies between [1, 1]. When r = 1, the network is said to have perfect assortative mixing patterns, i.e., cliques among same degree nodes, while at r = 1 the network is completely disassortative, i.e., star graphs. When r = 0 the network is non-assortative. Figure 3.10 presents the yearly assortativity coefficients of the Bitcoin transactions.

We find that the in-out degree correlation coefficient is negative, except for only a brief period in the initial phase. After mid-2010, the degree correlation coefficient stays between r ≈ -0.012 and r ≈ -0.015, reaching a value of r ≈ -.016 by 2014, suggesting that the network is disassortative. In general, for large scale-free networks, assortativity vanishes as the network size increases [36] and similar behavior is observed in the Bitcoin network.

3.2.3 Clustering

We also measured the average clustering coefficient, which measures local density of edges.

$$C = \frac{1}{n}\Sigma_{i=1}^{n}\frac{\Delta_i}{d_i(d_i-1)/2}$$

Where d_i is the degree of node i, $\Sigma_{i=1}^{n}$ runs over all the nodes, and Δ_i is the number of triangle containing node i. To calculate Δ_i, we ignored the directionality of the edges. The average clustering coefficient C is the average of local clustering coefficients of all the nodes n.

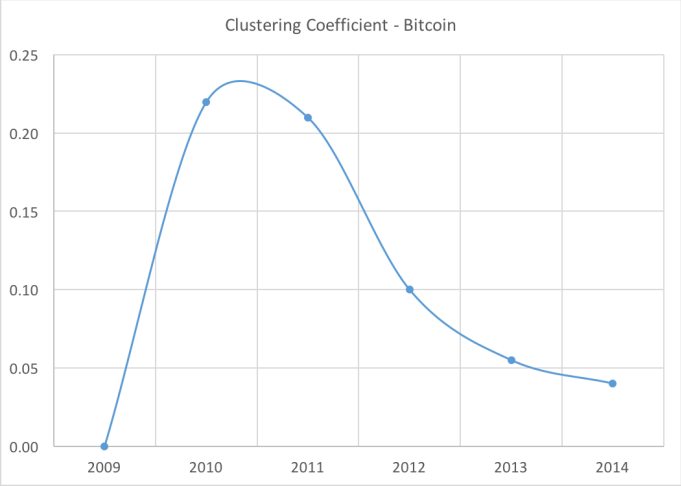

Figure 3.11: Clustering coefficient of yearly Bitcoin transactions.

Figure 3.11 presents yearly clustering coefficients of the Bitcoin transactions. We observed that in the initial phase C is high, fluctuating around 0.15. This is because the initial transactions may be the placed by few initial user transferring money between their own accounts to test the network. After the initial phase, the clustering coefficient reduces from 0.007 in 2012 to around 0.052 in 2014, which is much higher than a random network of similar size.

3.2.4 Anonymity

Even though Bitcoin data is anonymous, an active attacker can observe the IP address of a transaction request and match it to an actual user. Hence, some users

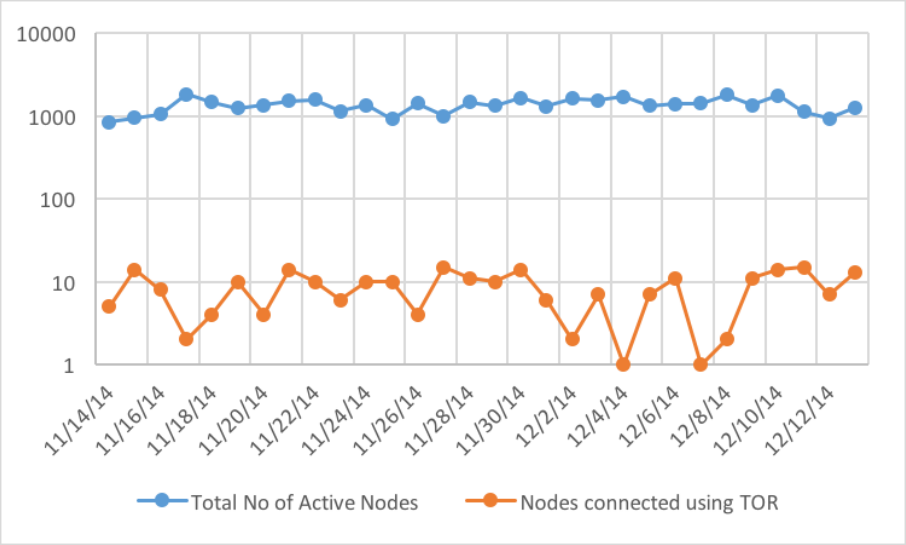

Figure 3.12: Anonymity among MyWallet users

might be interested in hiding their IP address even when communicating with the network.

Anonymizer technologies allow one to hide a user's IP address and are widely used [35]. Tor is currently the most popular anonymizer network with millions of users [1]. We wanted to analyze the percentage of users using Bitcoin anonymously [38]. We tracked the IP addresses connected to the Bitcoin network and compared those IP addresses with the exit nodes of Tor anonymizer network.

We crawled `blockchain.info` for the online nodes and compared the list with the Tor network exit nodes every hour. We compared IP addresses for 30 days to find the percentage of users connecting to the Bitcoin network using Tor anonymizers. We observed that among 800 to 2000 connects to mywallet at a given time only up to 20 nodes are using anonymizers as shown in Figure 3.12.

3.3 Litecoin Network

Litecoin is refereed to as the silver form of Bitcoin where the protocol is designed so that custom hardware cannot be used for mining. Even through Litecoin market value is 1 % of Bitcoin, the Litecoin network has a total N= 6,990,919 unique addresses and E = 6,486,325 edges.

3.3.1 Degree

We calculated the in degree and out degree distributions of the network in Figure 3.13 - Figure 3.14 and Figure 3.15 - Figure 3.16. Unlike Bitcoin network, the Litecoin network growth is continuous. The degree distributions show a power law pattern with an exponent of $\alpha_{in} \sim -2.14$ for in degree distribution and $\alpha_{out} \sim -2.01$ for out degree distribution.

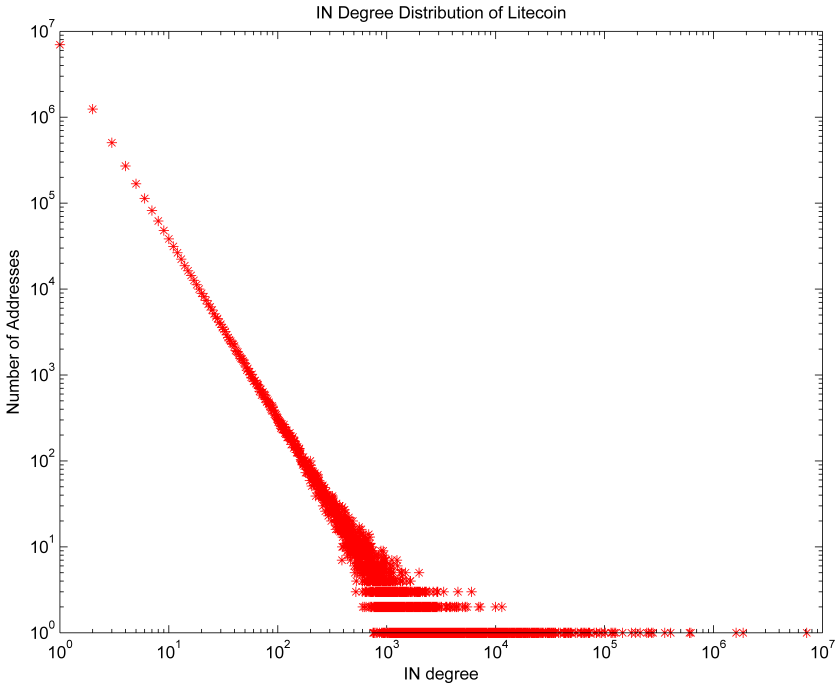

Figure 3.13: In degree distribution of the Litecoin transactions until Jan 2015.

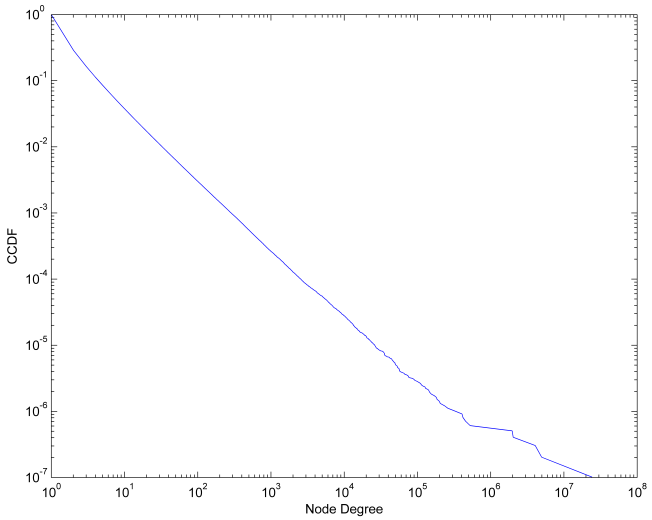

Figure 3.14: CCDF of in degree distribution of the Litecoin transactions until Jan 2015.

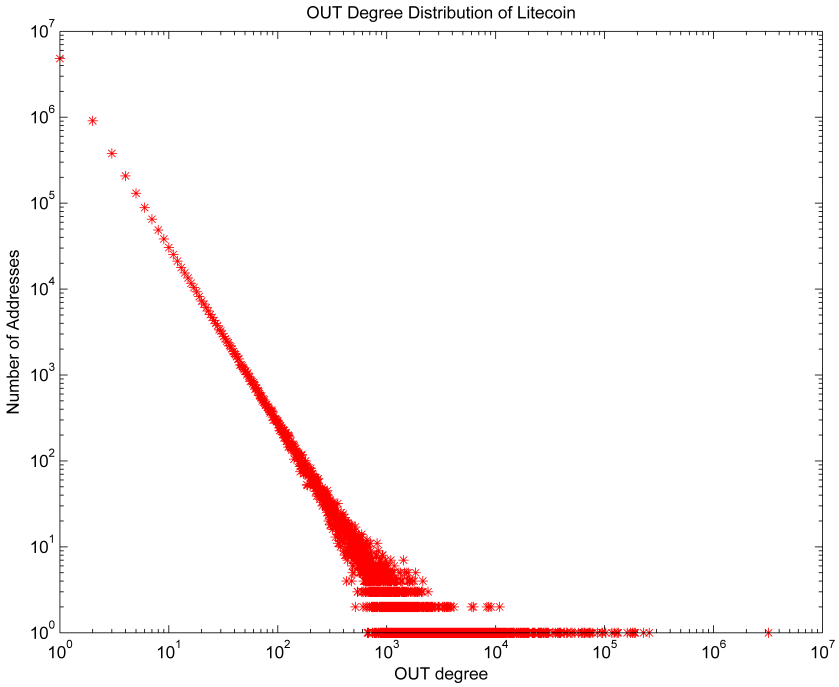

Figure 3.15: Out degree distribution of the Litecoin transactions until Jan 2015.

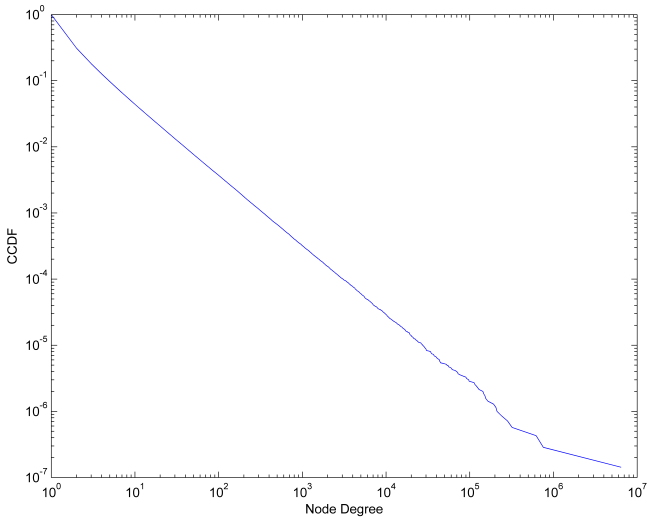

Figure 3.16: CCDF of out degree distribution of the Litecoin transactions until Jan 2015.

Table 3.3: In degree characteristics of yearly Litecoin transactions

Year	α	*Nodes*	*Edges*	*MaxDegree*	*AvgDegree*
2011	1.8	22400	754734	170892	33.69
2012	1.9	545576	10391318	1124344	19.04
2013	2.0	2546672	25208855	2765143	9.89
2014	2.2	6735643	19850699	360129	2.94

Table 3.4: Out degree characteristics of yearly Litecoin transactions

Year	α	*Nodes*	*Edges*	*MaxDegree*	*AvgDegree*
2011	2.2	22400	63163	4037	2.81
2012	2.1	545576	2484673	395841	4.55
2013	2.0	2546672	17876786	734660	7.01
2014	2.1	6735643	32031470	1373967	4.75

Table 3.3 and Table 3.4 present yearly Litecoin network characteristics. The in degree and out degree are stable for the entire lifetime of the Litecoin network as the network is continuously growing since the introduction of the currency.

Figures 3.17 and Figures 3.18 present in degree and out degree distributions of yearly Litecoin transactions.

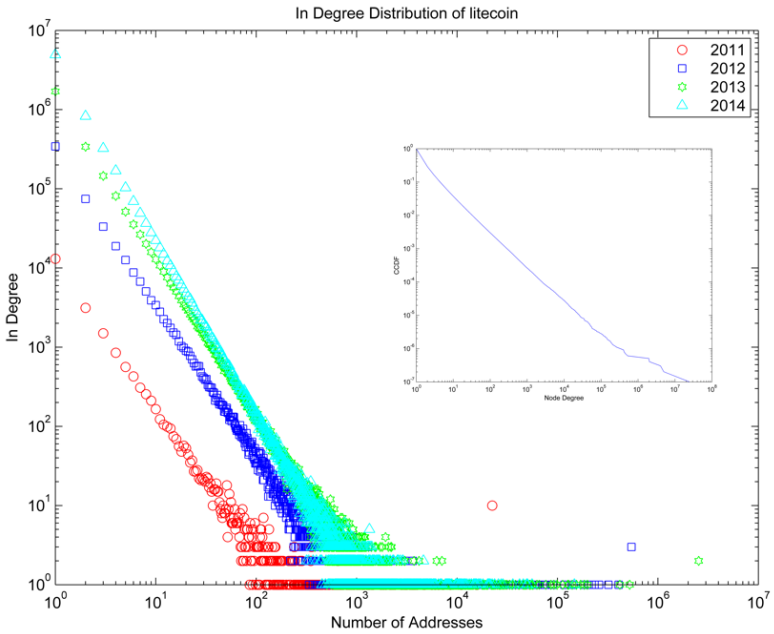

Figure 3.17: In degree distribution of the yearly Litecoin transactions.

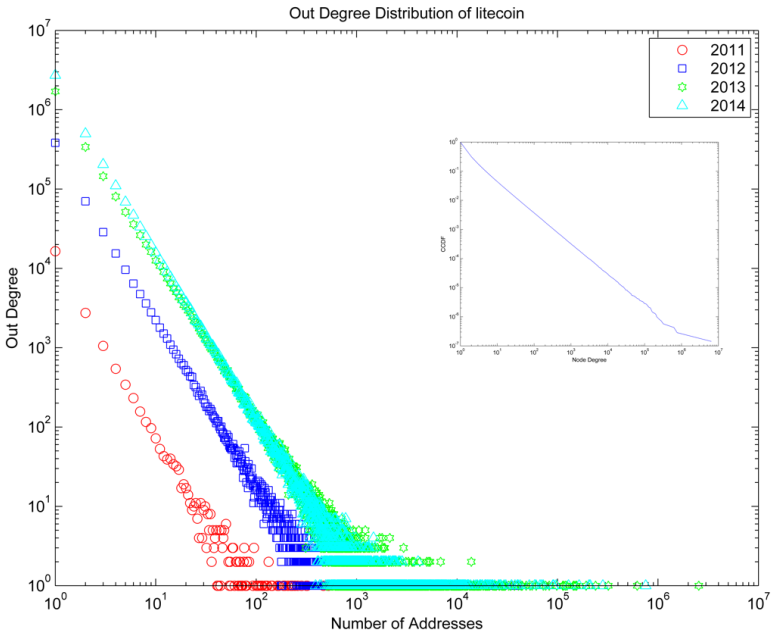

Figure 3.18: Out degree distribution of the yearly Litecoin transactions

3.3.2 Assortativity

We compute the degree correlation function, i.e., the in degree K_{in} of the nodes with out degree K_{out}, for the network in Figure 3.19. We find that the in-out degree correlation is dissortaative as the nodes with high degree have low in degree. This distribution is clearly different from Figure 3.9 for Bitcoin where the very high degree nodes connected to other very high degree nodes.

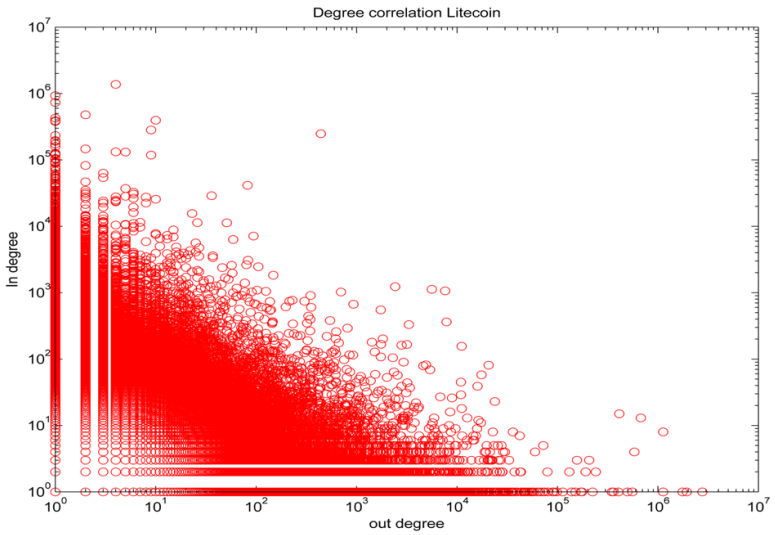

Figure 3.19: In degree as the function of out degree $K_n^{in}(K_{out})$ for Litecoin transactions.

Figure 3.20 presents assortativity coefficient of yearly Litecoin transactions. We find that the In-Out degree correlation coefficient is non-assortative except for the first two year. Assortativity of Litecoin is tending toward 0 indicating a non-assortative behavior.

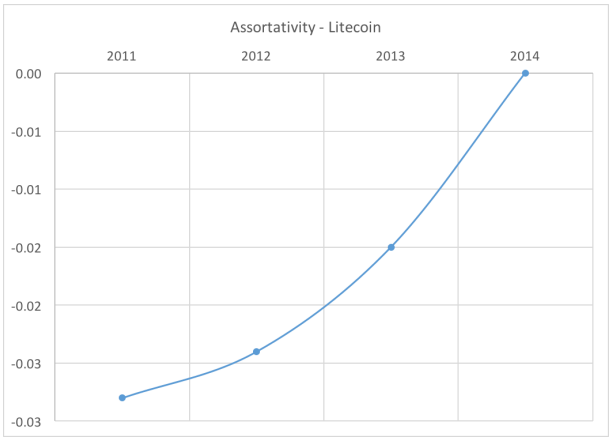

Figure 3.20: Assortativity coefficients of the yearly Litecoin transactions.

3.3.3 Clustering

We also measured the average clustering coefficient, the C value in Figure 3.21 We observed that, in the initial phase C is high. After the initial phase the clustering coefficient reduces from 0.33 in 2012 to around 0.05 in 2013, and has become 0.032 in 2014.

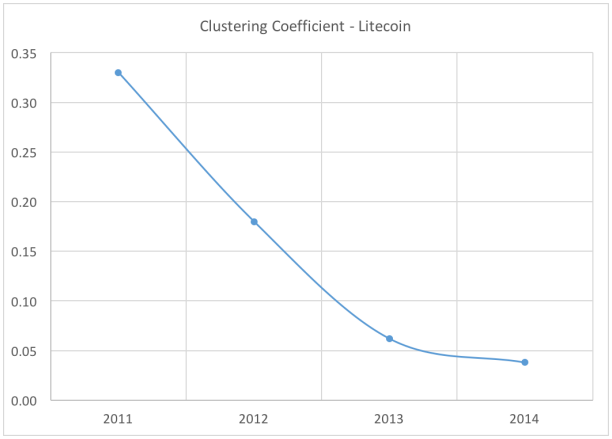

Figure 3.21: Clustering of the network.

Chapter 4

Market Analysis

In this chapter, we analyze the richest users of popular crypto currencies. The top 100 richest people in various crypto currencies might reveal the secrets about how a rich person becomes more rich, how he accumulates the money, and so on. Hence, we pick the top 100 richest people in each crypto currency and we analyze their degrees compared to their wealth [11]. We back traced all the coins in the wallets of the top 100 nodes. We analyzed the data for specific patterns and the relations among the top 100 nodes.

4.1 Richest Bitcoin Addresses

We collected the data of the top 100 richest addresses in the Bitcoin network and analysed for unique patterns in their behaviour. The total Bitcoins in circulation are 14,917,575 BTC with a market value of 377.93 USD and a market capitalization of 5,632,274,268 USD as of Dec 1, 2015. The top 100 richest nodes in Bitcoin hold

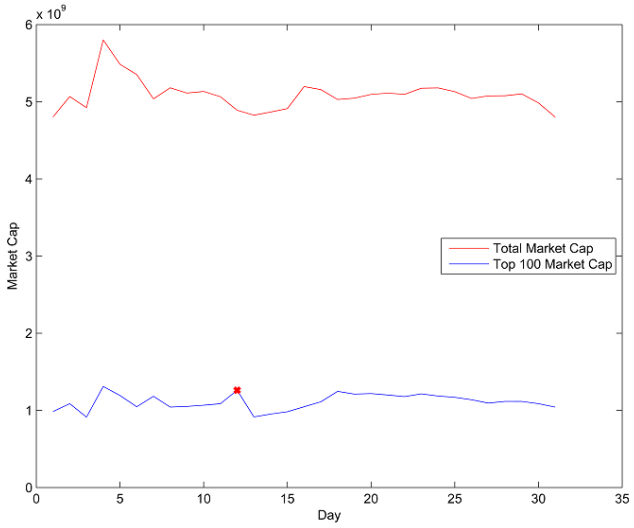

Figure 4.1: Percentage of Bitcoin wealth, the richest own during Dec 2014.

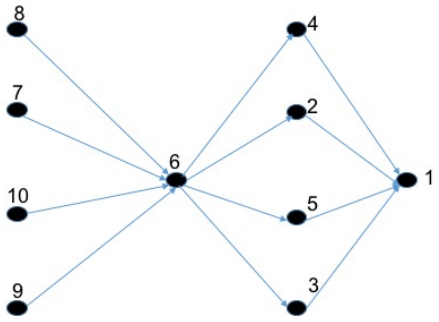

Figure 4.2: Transaction pattern of the richest Bitcoin node.

19.88 % of wealth as shown in Figure 4.1.

We noticed couple of interesting behaviours among richest Bitcoin users. For instance, the richest node transfers his/her bitcoins to four new addresses and then on the same day transfers all coins back into a single new address, which becomes the new richest address as shown in Figure 4.2.

Figure 4.3 shows the in and out degree of the top 100 nodes. We observe that the incoming transactions to the richest people are through mining nodes, which specifies that most of the richest nodes are miners. We also observe that the ap-

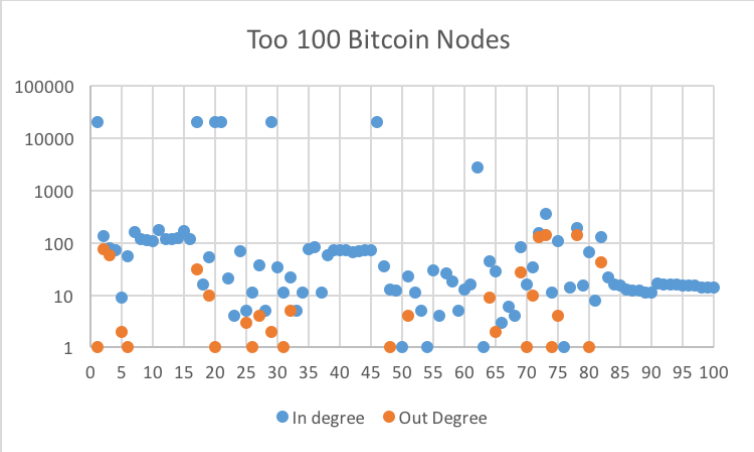

Figure 4.3: In and out degree of the richest 100 Bitcoin nodes.

proximately 73 % of the richest people have 0 out degree, which means that they just accumulate money without spending it.

4.2 Richest Litecoin Addresses

The total Litecoin in circulation are 43,455,110 LTC with a market value of 0.00959 USD and a market capitalization of 157,184,844 USD as of Dec 1, 2015. The 48.89 % of the total market capitalization of the Litecoin is hold by the richest 100 people. We observed that the behaviour of the top 100 addresses in the Litecoin network are similar to the Bitcoin's richest people. We find that among the 100 richest nodes 82 % of the nodes have 0 out degree as shown in Figure 4.4. We observe an interesting pattern among the richest Litecoin users where more than two thirds of the 100 richest nodes simply transfer their Litecoins into a new account while paying a small transaction fee.

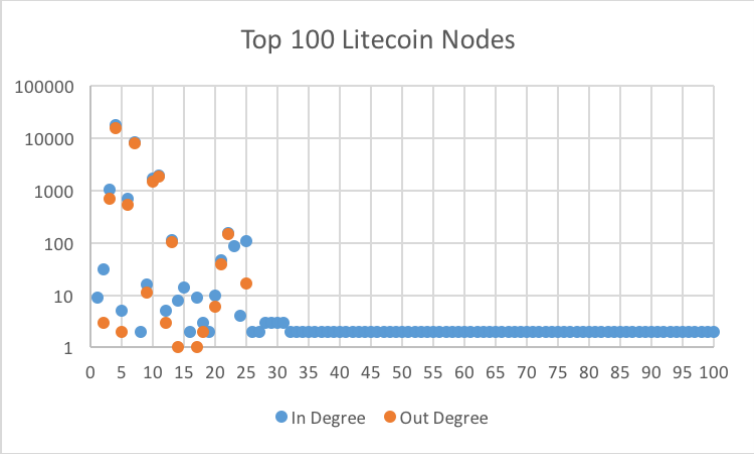

Figure 4.4: In and out degree of the richest 100 Litecoin nodes.

4.3 Richest Dash Addresses

The total Dash in circulation are 6,035,717 DASH with a market value of Dashcoins is about 0.00574 USD and a market capitalization of 13,056,033 USD as of Dec 1, 2015. Note that Darkcoin was renamed as Dash. The top 100 richest nodes in Dash hold 34.82 % of wealth. About 80 % of the richest nodes have a 0 out degree as shown in Figure 4.5. We also observed that the top 100 nodes in Dash keep

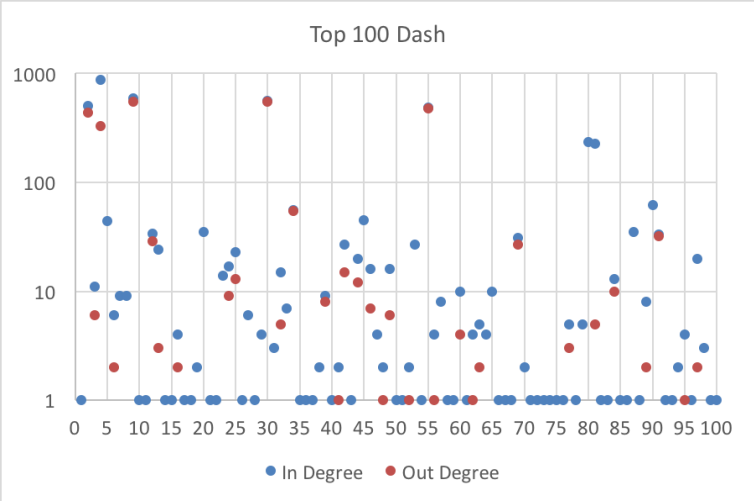

Figure 4.5: In and out degree of the richest 100 Dashcoins nodes.

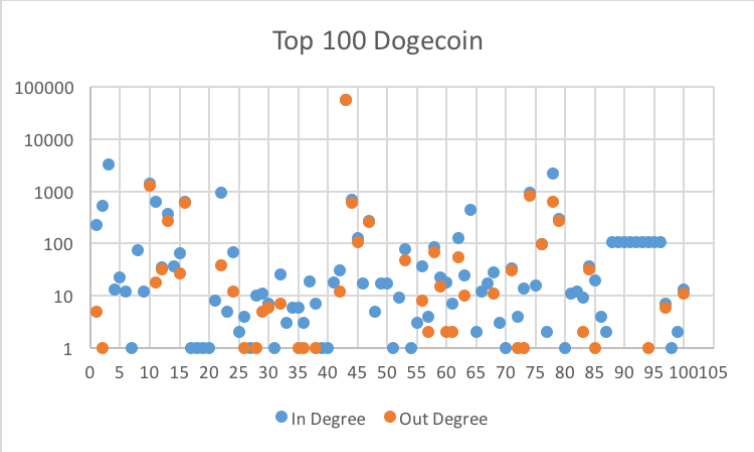

Figure 4.6: In and out degree of the richest 100 Dogecoin nodes.

changing.

4.4 Richest Dogecoin Addresses

The total Dogecoins in circulation are 102,091,461,013 DOGE with a market value of 0.00000034 USD and a market capitalization of 13,019,691 USD as of Dec 1, 2015. The top 100 richest nodes in Dogecoin network hold 45.20 % of wealth. We find that unlike other coins, Dogecoin richest nodes are there since the introduction of the network while 54 % of the nodes have 0 out degree as shown in Figure 4.6.

4.5 Richest Peercoin Addresses

The total peercoins in circulation are 22,814,995 PPC with a market value is 0.00122 USD and a market capitalization of 10,495,248 USD as of Dec 1, 2015. The top 100 richest nodes in Peercoin hold 58.71 % of wealth. We observed that only less than 19 % of nodes have a 0 out degree, indicating almost all rich nodes are active in

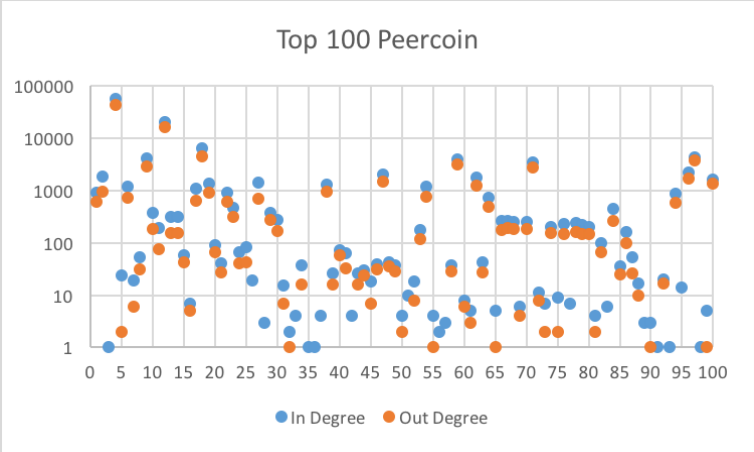

Figure 4.7: In and out degree of the richest 100 Peercoin nodes.

transactions rather than accumulating wealth Figure 4.7.

4.6 Richest Namecoin Addresses

The total Namecoins in circulation are 13,052,300 NMC with a market value of 0.00127 USD and a market capitalization of 6,255,649 USD as of Dec 1, 2015. The top 100 richest nodes in Namecoin hold 75.13 % of wealth. We found that the top 100 nodes also existed since the introduction of the network while 75 % of the nodes have 0 out degree as shown in Figure 4.8.

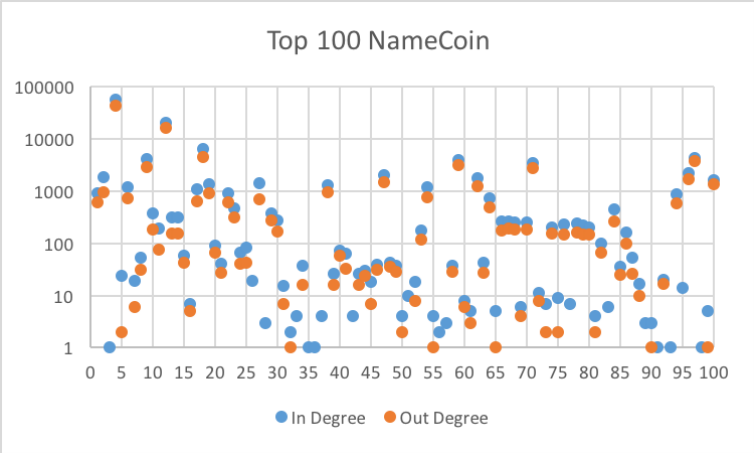

Figure 4.8: In and out degree of the richest 100 Namecoin nodes.

Chapter 5

Regression Analysis

Bitcoin price is highly volatile and it depends on various factors similar to other currencies [11]. In this chapter, we want to determine the factors affecting the price of the Bitcoin and develop a mechanism to predict market value of Bitcoins. To achieve this, we performed regression analysis, a statistical process for estimating the relationships among the variables [34]. In regression analysis, we call the variable for which we want to find a relation the *dependent variable* and the variables from which the relationship is derived from as *predictor variables*.

We performed regression on the last 6 months of Bitcoin values shown in 2.4. The dependent variable for analysis is the Bitcoin market price while we considered predictor variables are miners revenue in USD , total transaction volume per day in USD, total outgoing transactions per day in USD, exchange trade value in USD, market capitalization in USD, transaction fees in USD, and cost per transaction in USD [54, 18]. All the predictor variables are in USD.

We perform the regression based on the results from Analysis of variance (ANOVA)

tables. ANOVA is a collection of statistical models used to analyze the differences among group means and their associated procedures such as ”variation” among and between groups [53]. In ANOVA setting, the observed variance in a particular variable is partitioned into component attributable to different sources of variation. In its simplest form, ANOVA provides a statistical test of whether or not the means of several groups are equal, and therefore generalizes the t-test to more than two groups. As doing multiple two-sample t-tests would result in an increased chance of committing a statistical type I error, ANOVAs are useful for comparing (i.e., testing) three or more means (i.e., groups or variables) for statistical significance.

We finally, perform the stepAIC analysis to test our predictions. StepAIC selects the model based on Akaike Information Criteria (AIC), a measure of the relative quality of statistical models for a given set of data [9]. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection.

5.1 Regression Model with Market Capitalization Variable

The matrix plot with all of the predictor variables is shown in Figure 5.1. We observe that market capitalization is directly related to the market price of Bitcoin. ANOVA indicate that *Market Price(Y) = Market Capitalization (X)* where the market price is only dependent on the market capitalization. The r-squared value is 0.9981 and p-value is less than 2.2e-16. These values indicate that the market capitalization is direct dependent on the market price.

We also performed ANOVA analysis of the model with only market capitalization as shown in Figure 5.2. The model indicates an almost perfect match.

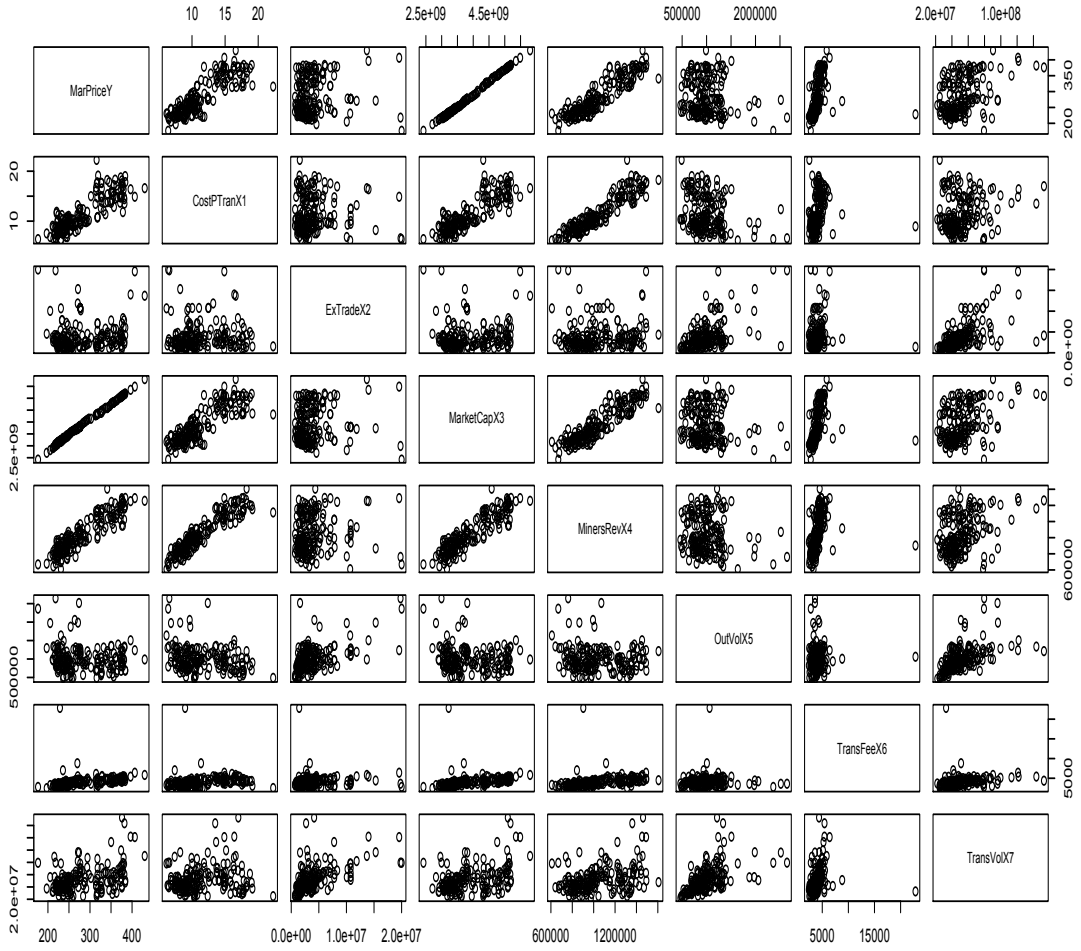

Figure 5.1: Regression analysis with market capitalization

```
anova(lm10)
Analysis of Variance Table

Response: MarPriceY
            Df Sum Sq Mean Sq F value    Pr(>F)
MarcapX7     1 587135  587135   92935 < 2.2e-16
Residuals  179   1131       6
```

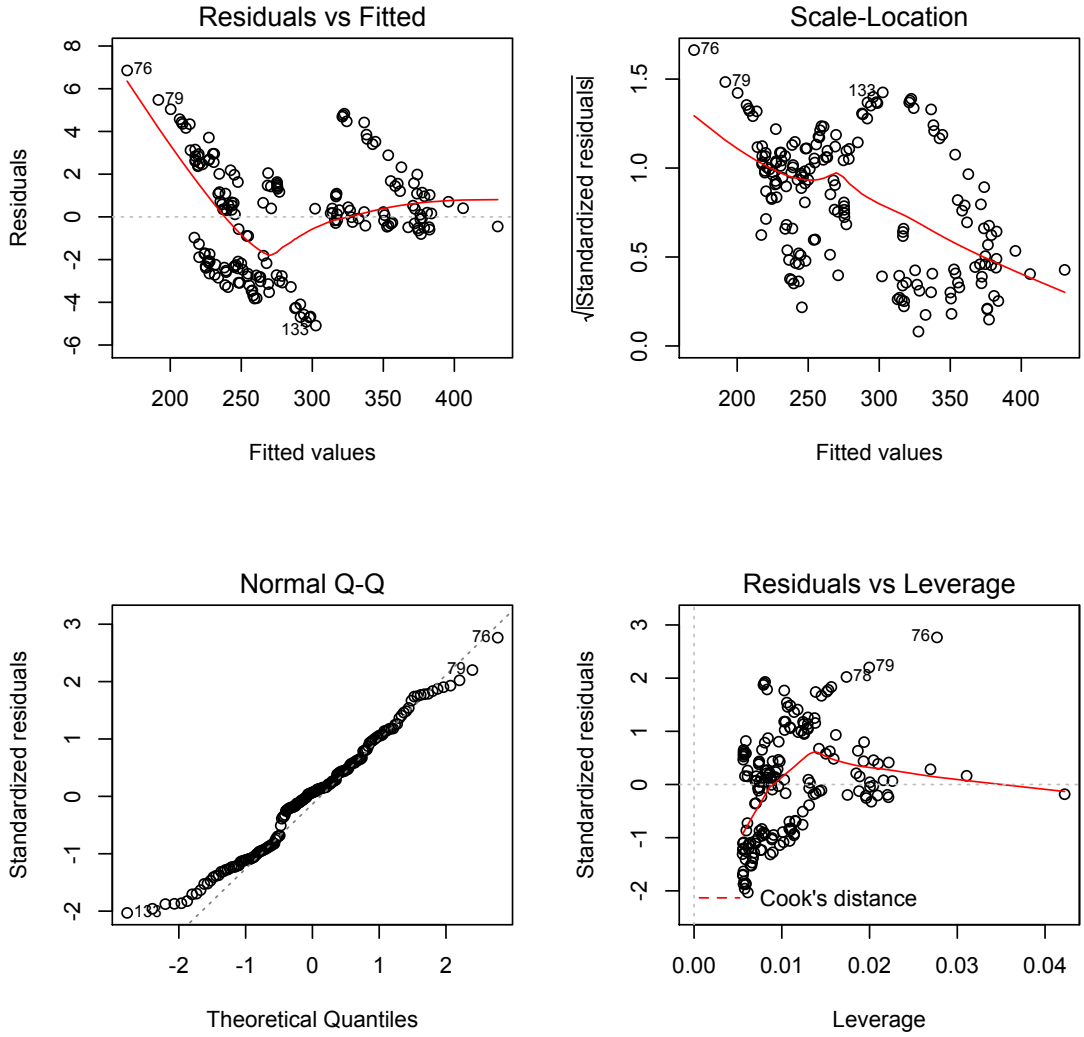

Figure 5.2: Plot of market price and market capitalization

However, as the *Market capitalization = Market price * Total number of Bitcoins*, we can not use the variable which itself is derived from the market price. Hence, we removed the market capitalization and performed the regression analysis as shown in Figure 5.3.

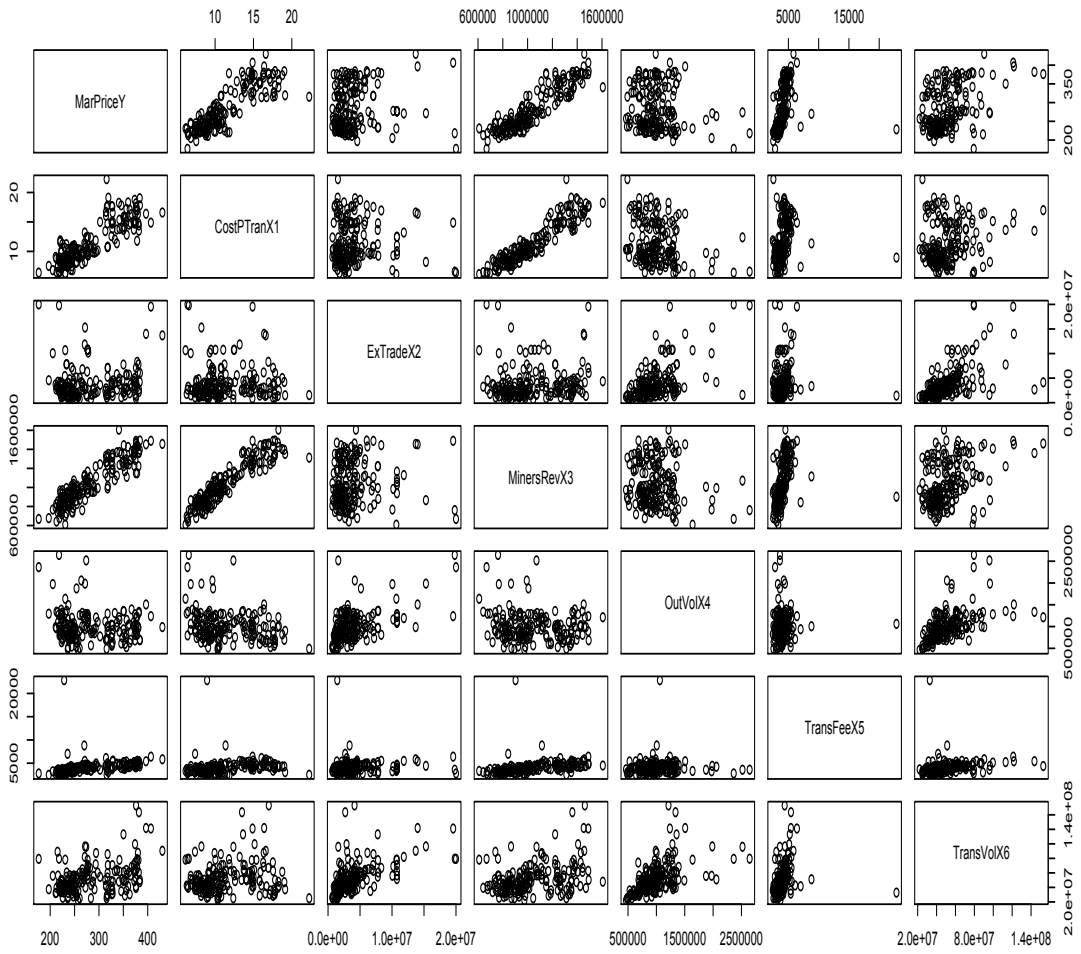

Figure 5.3: Regression analysis without market capitalization

5.2 Regression Model with All Predictor Variables - Full Model

We initially run the model with all predictor variables and used the ANOVA test to calculate the threshold of significance for each variable. The registered values below the significance threshold of 0.05 indicate that the independent predictor variables explain the variation of the dependent variable.

```
> anova(lm1)
Analysis of Variance Table

Response: MarPriceY
              Df Sum Sq Mean Sq  F value    Pr(>F)
CostPTranX1    1 428681  428681 865.6386 < 2.2e-16 ***
ExTradeX2      1   8135    8135  16.4261 7.615e-05 *
MinersRevX3    1  52674   52674 106.3651 < 2.2e-16 ***
OutVolX4       1    151     151   0.3057    0.5810
TransFeeX5     1    755     755   1.5247    0.2186
TransVolX6     1  11702   11702  23.6297 2.596e-06 ***
Residuals    174  86168     495
```

Figure 5.4 presents Normal QQ plot for the full model. The multiple R-squared value is about 0.8535 and the p-value is less than 2.2e-16. We observe that the market price is dependent on all of the predictor variables.

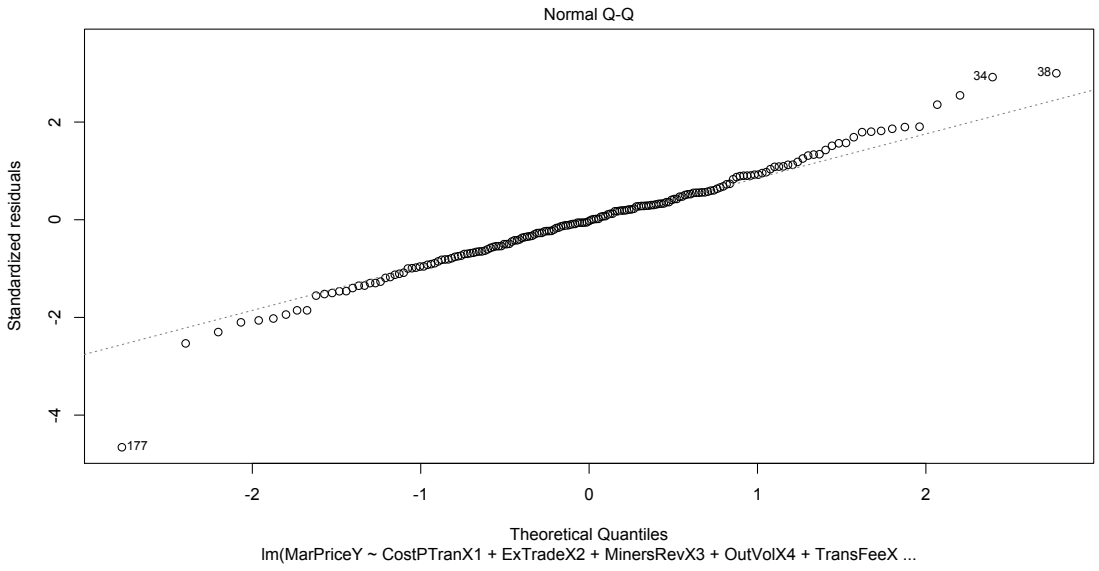

Figure 5.4: Normal QQ plot for full model

```
summary(lm1)


Call:
lm(formula = MarPriceY ~ CostPTranX1
+ ExTradeX2 + MinersRevX3 +
OutVolX4 + TransFeeX5 + TransVolX6)

Residuals:
     Min       1Q   Median       3Q      Max
-49.890  -14.355   -0.571   12.216   65.978

Coefficients:
              Estimate Std. Error
t value Pr(>|t|)
(Intercept)  6.847e+01   1.108e+01    6.179  4.45e-09 ***
CostPTranX1  4.603e+00   1.499e+00    3.070   0.00248 **
ExTradeX2    1.877e-09   6.327e-07    0.003   0.99764
MinersRevX3  1.339e-04   2.513e-05    5.328  3.05e-07 ***
OutVolX4    -1.654e-05   6.749e-06   -2.450   0.01526 *
TransFeeX5   1.529e-03   1.076e-03    1.422   0.15690
TransVolX6   6.285e-07   1.293e-07    4.861  2.60e-06 ***
---
Signif. codes:  0    ***    0.001
   **    0.01    *    0.05    .    0.1 1

Residual standard error: 22.25 on
174 degrees of freedom
Multiple R-squared:  0.8535,  Adjusted R-squared:  0.8485
F-statistic:   169 on 6 and 174 DF,  p-value: < 2.2e-16
```

The summary of the full model with all the predictor variables and the plot of

the full model is shown in Figure 5.5. We from different linear models by adding and removing predictor variables. The multiple R-squared value of the model with all the predictor variables is 0.8535 so the reduced model should be one with R-squared value greater that or equal to 0.8535.

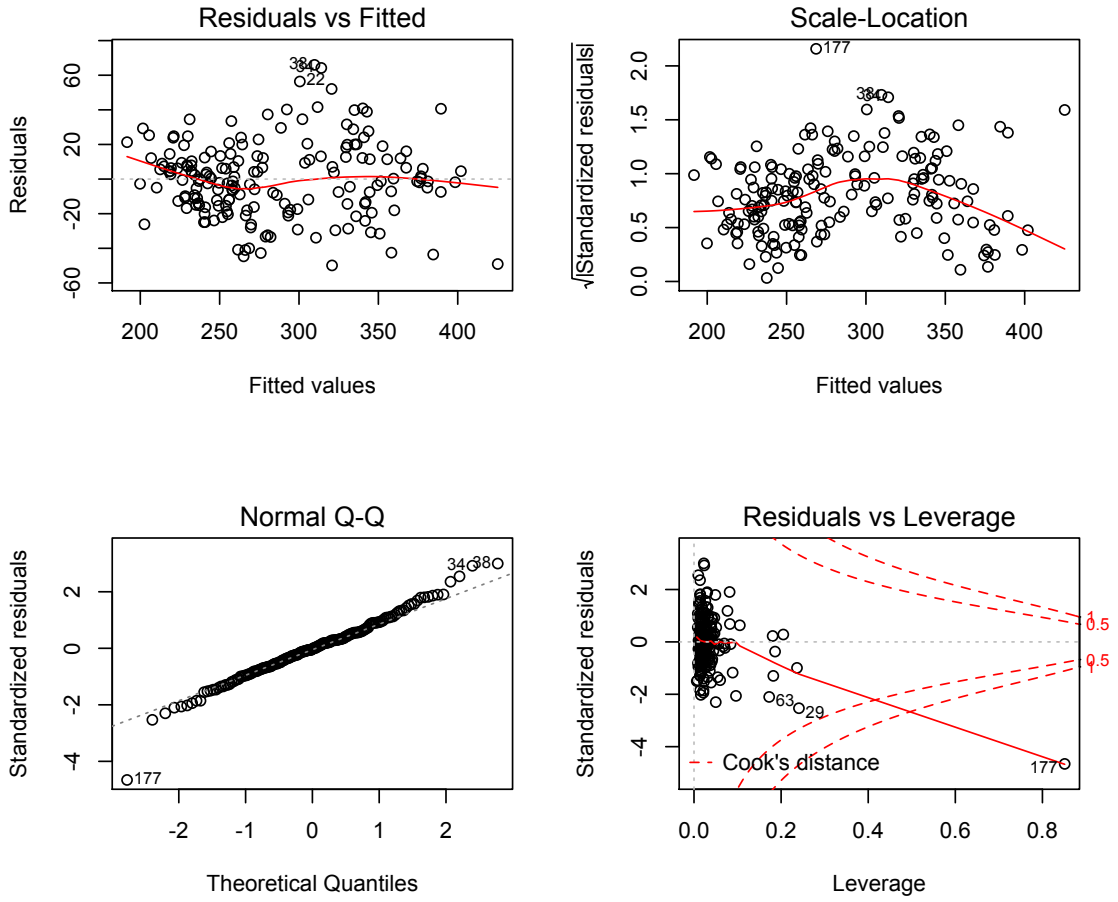

Figure 5.5: Plot of the full model

5.3 Regression Model with Significant Predictor Variables - Reduced Model

To find the best predictor variables, we tried different possible models using the six predictor variables and compared every model to the full model. If the reduced and full model yield the same the R-Squared value, we remove the dropped variable from the reduced model. After analyzing all of the linear models, we only drop the *exchange trade* variable and keep the rest of the predictor variables.

```
> summary(lm8)

Call:
lm(formula = MarPriceY ~ CostPTranX1 +
MinersRevX3 + OutVolX4 +
    TransFeeX5 + TransVolX6)

Residuals:
   Min      1Q  Median      3Q     Max
-49.88  -14.36   -0.57   12.22   65.97

Coefficients:
              Estimate Std. Error
t value Pr(>|t|)
(Intercept)  6.847e+01  1.101e+01    6.218 3.58e-09 ***
CostPTranX1  4.604e+00  1.477e+00    3.117  0.00214 **
MinersRevX3  1.339e-04  2.478e-05    5.402 2.12e-07 ***
OutVolX4    -1.653e-05  6.380e-06   -2.591  0.01037 *
TransFeeX5   1.529e-03  1.072e-03    1.426  0.15559
TransVolX6   6.287e-07  1.202e-07    5.229 4.82e-07 ***
---
```

```
Signif. codes:  0     ***     0.001
   **     0.01     *     0.05     .     0.1          1

Residual standard error: 22.19 on
175 degrees of freedom
Multiple R-squared:  0.8535,
Adjusted R-squared:  0.8493
F-statistic: 203.9 on 5 and 175 DF,  p-value: < 2.2e-16
```

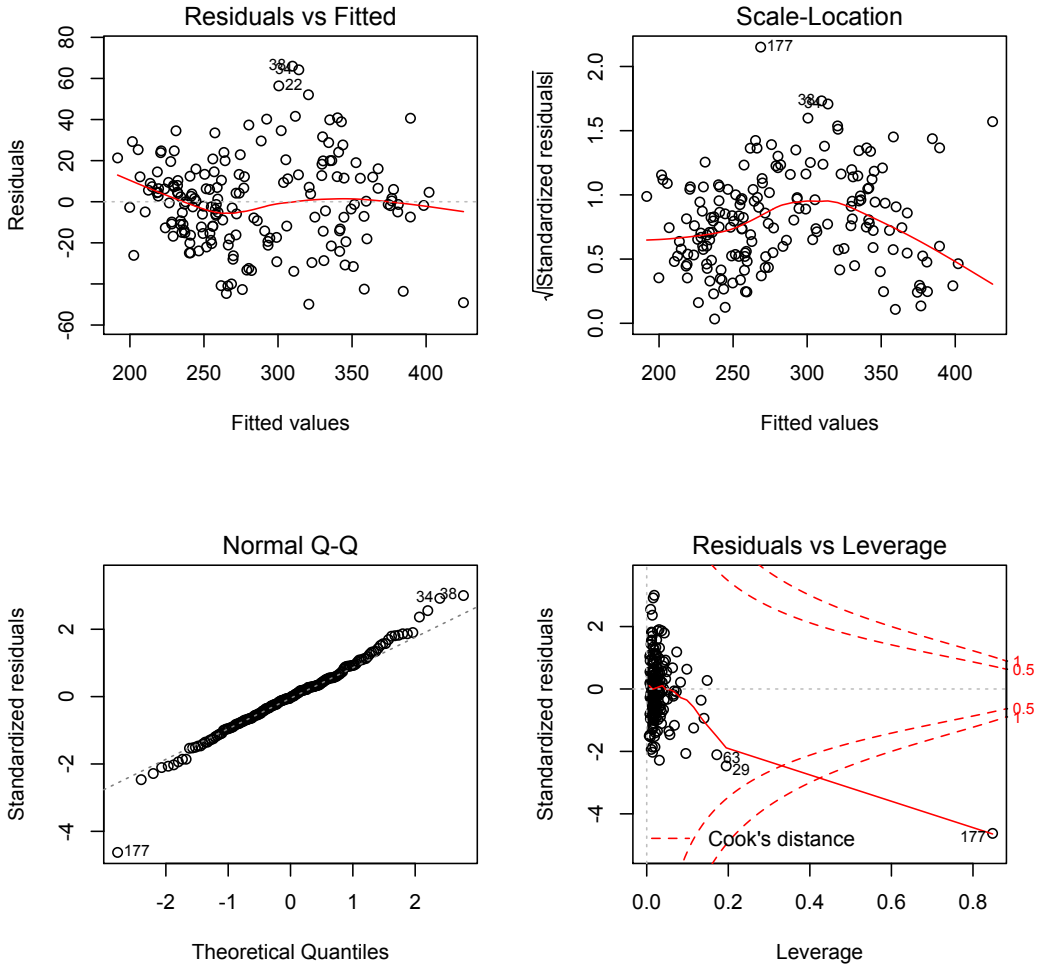

Figure 5.6: Plot of reduced model

The summary of the reduced model and the plot of the Reduced model in Figure 5.6. Here we observe that the model without exchange trade variable has R-squared value equal to the full model, which indicates that exchange trade has no effect on the market price of Bitcoin.

5.3.1 Comparison of Full and Reduced Models

We compared, the full model and reduced model in the following. We observe that both models yield the same R-squared value of 0.86168.

```
> anova(lm1,lm8)
Analysis of Variance Table

Model 1: MarPriceY ~ CostPTranX1 +
ExTradeX2 + MinersRevX3 + OutVolX4
+TransFeeX5 + TransVolX6
Model 2: MarPriceY ~ CostPTranX1 +
MinersRevX3 + OutVolX4 + TransFeeX5 +
TransVolX6
  Res.Df   RSS Df Sum of Sq  F Pr(>F)
1    174 86168
2    175 86168 -1 -0.004357  0 0.9976
```

5.3.2 Model with Miners Revenue

This section includes the summary of the model with only *miners revenue* as the predictor variable. Even through the model only contains miners revenue, the R-squared value is almost equal to the R-squared value of the full model (0.8284 vs

0.8535 for the full model). Hence, we can select the model with only *miners revenue* variable as our model, as it is the second best. From this analysis, we observe that miners revenue plays a significant role in prediction of the Bitcoin market price.

```
> summary(lm12)

Call:
lm(formula = MarPriceY ~ MinersRevX3)

Residuals:
     Min       1Q   Median       3Q      Max
-70.791  -17.581   -1.014   15.385   70.800

Coefficients:
             Estimate Std. Error
t value Pr(>|t|)
(Intercept) 4.566e+01  8.288e+00   5.509 1.24e-07 ***
MinersRevX3 2.278e-04  7.750e-06  29.397  < 2e-16 ***
---
Signif. codes:  0    ***    0.001
   **    0.01    *    0.05    .    0.1         1

Residual standard error: 23.75 on
179 degrees of freedom
Multiple R-squared:  \textbf{0.8284},
Adjusted R-squared:  0.8274
F-statistic: 864.2 on 1 and 179 DF,
p-value: < 2.2e-16
```

The plot of the *miners revenue* based model is shown in Figure 5.7.

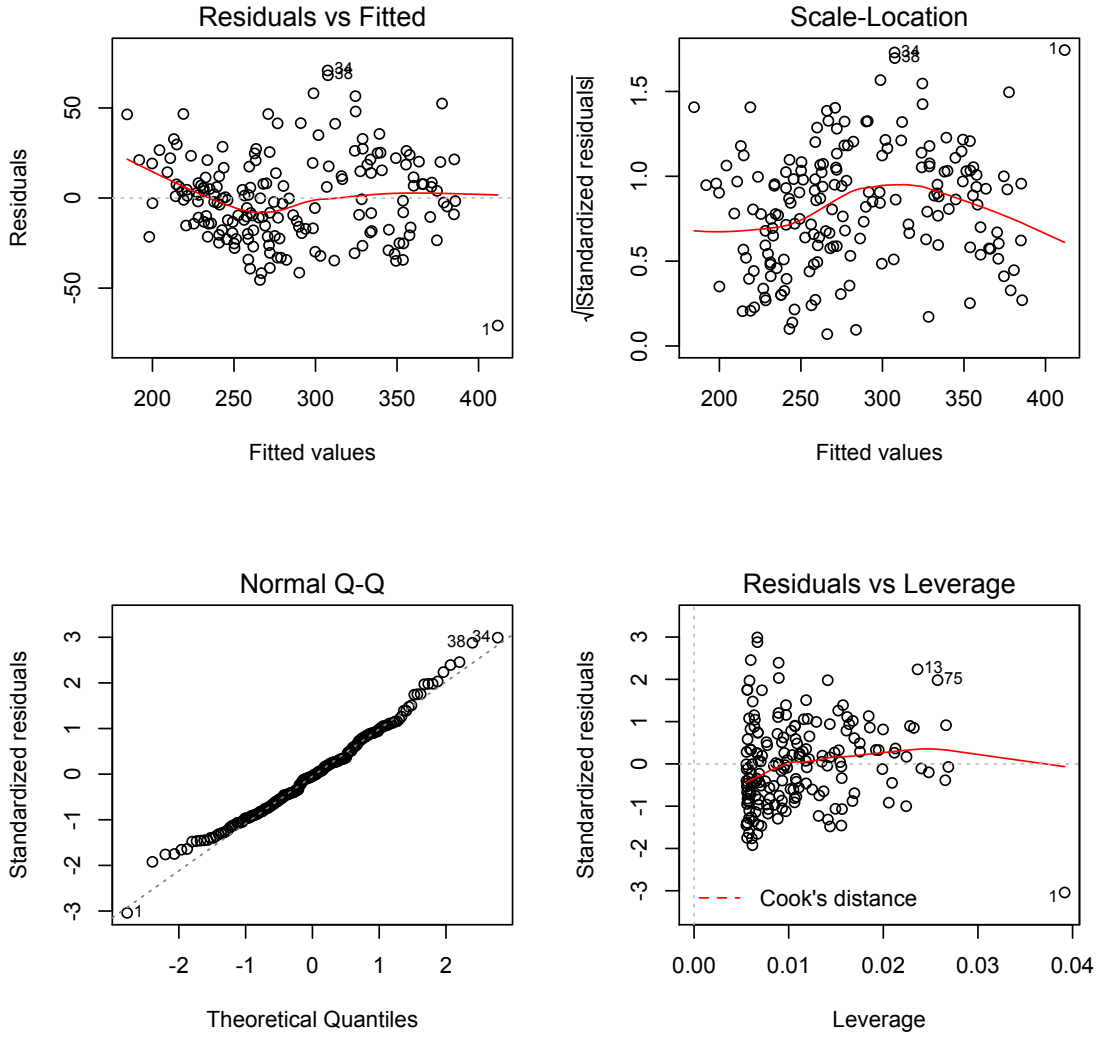

Figure 5.7: Plot of linear model with miners revenue

5.3.3 SetpAIC Analysis

StepAIC analysis [9] performs regression analysis on the full model by adding and removing variables in both directions, and selects the model with least AIC value as the reduced model. We use stepAIC analysis to validate our models.

```
> step=stepAIC(lm1,direction="both")
Start:  AIC=1129.97
MarPriceY ~ CostPTranX1 + ExTradeX2 +
MinersRevX3 + OutVolX4 +
TransFeeX5 + TransVolX6

               Df Sum of Sq     RSS     AIC
- ExTradeX2     1         0.0   86168  1128.0
<none>                       86168  1130.0
- TransFeeX5    1      1000.9   87169  1130.1
- OutVolX4      1      2973.3   89141  1134.1
- CostPTranX1   1      4668.8   90837  1137.5
- TransVolX6    1     11701.9   97870  1151.0
- MinersRevX3   1     14056.7  100225  1155.3

Step:  AIC=1127.97
MarPriceY ~ CostPTranX1 + MinersRevX3
+ OutVolX4 + TransFeeX5 + TransVolX6

               Df Sum of Sq     RSS     AIC
<none>                       86168  1128.0
- TransFeeX5    1      1001.5   87170  1128.1
+ ExTradeX2     1         0.0   86168  1130.0
- OutVolX4      1      3305.8   89474  1132.8
- CostPTranX1   1      4783.8   90952  1135.8
- TransVolX6    1     13464.1   99632  1152.2
- MinersRevX3   1     14370.3  100538  1153.9
> step$anova
Stepwise Model Path
Analysis of Deviance Table
```

```
Initial Model:
MarPriceY ~ CostPTranX1 + ExTradeX2
+ MinersRevX3 + OutVolX4 +
TransFeeX5 + TransVolX6


Final Model:
MarPriceY ~ CostPTranX1 + MinersRevX3
+ OutVolX4 + TransFeeX5 + TransVolX6



         Step Df     Deviance Resid. Df
Resid. Dev      AIC
1                                  174    86168.11 1129.966
2 - ExTradeX2  1 0.004356982       175    86168.11 1127.966
```

The Step AIC analysis found (MarPriceY CostPTranX1 + MinersRevX3 + OutVolX4 + TransFeeX5 + TransVolX6) as the final model.

5.4 Validation of Assumptions

We use = 0.05 level of significance as threshold. Hence, if the p-value less than 0.05, we reject the null hypothesis and retain it otherwise.

By analyzing the residual versus fitted of the full model and reduced models, we can say that there is no observed pattern in the residuals. Normal QQ plot show the normality of the full model in Figure 5.4 and the reduced model in Figure 5.8.

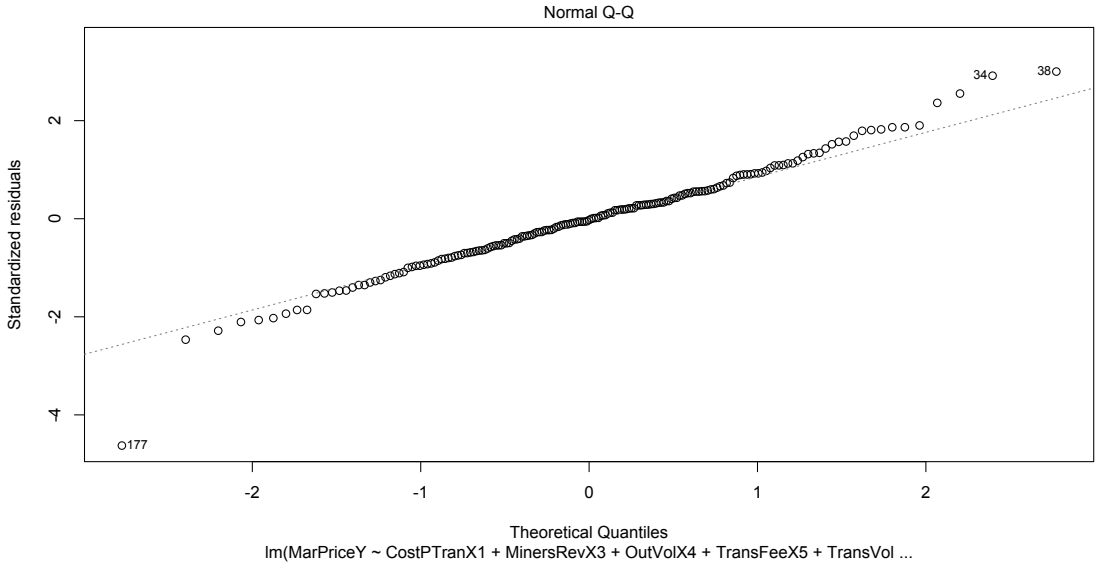

Figure 5.8: Normal QQ plot for reduced model

5.5 Variable Selection

We select the model without exchange trade as the reduced model based on two criterion

1. Multiple R-squared value of reduced model (i.e., model without exchange trade variable) is same as the multiple R-Squared value of full model.
2. Analysis of deviance table predicted that the model without trade value can predict the market price better compared to the model with exchange trade value.

We also noticed that miners revenue itself is a significant indicator of the market value. The multiple R-squared value of the miners revenue model with is 0.8284 with a p-value less than 2.2e-16.

According to our analysis and stepAIC result, we conclude that the highly volatile Bitcoin market price can best be predicted by the cost per transaction, transaction fee, total outgoing transactions, total number of transactions, and miners revenue.

Chapter 6

Conclusion

We have performed a detailed analysis of the popular digital currencies, namely, Bitcoin, Litecoin, Dash, Dogecoin, Peercoin, and Namecoin. We also analyzed their top 100 richest nodes. After becoming popular after mid-2011, Bitcoin is characterized by a dissasortative degree correlation and power law in- and out-degree distributions. Similarly, Litecoin network has disassortative degree correlation and power law in- and out-degree distributions after inception.

We found that majority of nodes in the Bitcoin network are gaining money from mining. The characteristics of richest nodes in Bitcoin, Litecoin and Dash networks are similar. While richest 100 nodes in Bitcoin, Litecoin and Dash keep changing, richest nodes in Dogecoin, Namecoin, and Peercoin networks are often the same nodes. We also found that majority of the richest nodes among all networks except peercoin are interested in accumulating money.

We found that market price of Bitcoin depends on the internal variables of Bitcoin such as miners revenue, transaction fee, transaction volume, daily output vol-

ume, and cost per transaction rather that on the foreign exchange trade value.

Chapter 7

Future Work

Although the results presented here have demonstrated the analysis of Bitcoin, Litecoin networks in detail, we can perform more detailed analysis of the data to get more insight. However, such analysis require considerable computing power due to large scale of the currency transactions.

Right now due to the limited amount of data available for the Dogecoin, Namecoin, Dash, and Peercoin networks, their transaction networks are not analyzed in detail. In future, we would like to analyze the remaining currencies in detail and compare the results with the Bitcoin and Litecoin.

There is a possibility to analyze the flow of currency, i.e., from the origination of money to the current balances. This can be achieved by viewing the network in detail and tracing the transactions to individual nodes.

Bibliography

[1] An overview of anonymity technology usage. *Computer Communications*, pages –, 2013.

[2] Mehmet Burak Akgun and Mehmet Hadi Gunes. Bipartite internet topology at the subnet-level. In *IEEE International Workshop on Network Science (NSW 2013)*, West Point, NT, April 2013. IEEE, IEEE.

[3] Mehmet Burak Akgun and Mehmet Hadi Gunes. Impact of multi-access links on the internet topology modeling. In *IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2013)*, San Francisco, CA, August 2013. IEEE, IEEE.

[4] Joan Antoni, Donet Donet, and P Cristina. The Bitcoin P2P network. *1st Workshop on Bitcoin Research*, pages 1–15, 2014.

[5] Engin Arslan, Mehmet Hadi Gunes, and Murat Yuksel. Analysis of academic ties: A case study of mathematics genealogy. In *2011 IEEE GLOBECOM Workshops*, pages 125–129, 2011.

[6] Adrienne E. Breland, Karen A. Schlauch, Mehmet Hadi Gunes, and Frederick C. Harris, Jr. Fast graph approaches to measure influenza transmission across geographically distributed host types. In *BCB '10: Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology*, page 594–601, New York, NY, USA, 2010. ACM, ACM.

[7] Adrienne.E. Breland, Mehmet Hadi Gunes, Karen A. Schlauch, and Frederick C. Harris, Jr. Mixing patterns in a global influenza a virus network using whole genome comparisons. In *2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–8, 2010.

[8] Mehmet Burak and Mehmet Hadi Gunes. Link-level network topology gen-

eration. In *2011 31st International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 140–145, 2011.

[9] Kenneth P. Burnham, David R. Anderson, and Kathryn P. Huyvaert. AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65(1):23–35, 2011.

[10] Dorothy P. Cheung and Mehmet Hadi Gunes. A complex network analysis of the united states air transportation. In *ASONAM '12: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*, page 699–701, Washington, DC, USA, 2012. IEEE Computer Society, IEEE Computer Society.

[11] Pavel Ciaian, Miroslava Rajcaniova, and Artis Kancs. Economics of BitCoin Price Formation. *EERI Research paper series*, No. 8:2–22, 2014.

[12] Grace Crosley and Mehmet Hadi Gunes. *Using Complex Network Representation to Identify Important Structural Components of Chinese Characters*, page 319–328. Springer International Publishing, 2014.

[13] Andrew Dittrich, Mehmet Hadi Gunes, and Sergiu Dascalu. *Network Analysis of Software Repositories: Identifying Subject Matter Experts*, volume 424 of *Studies in Computational Intelligence*, pages 187–198. Springer Berlin Heidelberg, 2013.

[14] Esra Erdin, Eric Klukovich, and Mehmet Hadi Gunes. An analysis of friend circles of facebook users. In *The 9th IEEE Workshop on Network Measurements (WNM)*, Clearwater Beach, FL, Oct 2015.

[15] Esra Erdin, Eric Klukovich, Mehmet Hadi Gunes, and Gurhan Gunduz. Posn: A personal online social network. In *30th International Information Security and Privacy Conference*, 2015.

[16] and M. Harrigan F. Reid. An analysis of anonymity in bitcoin system. *arXiv*, 1107.4524(2):215–226, 2011.

[17] C Fink and T Johann. Bitcoin Markets. *Available at SSRN 2408396*, (2013):1–19, 2014.

[18] J Fletcher. Multiple Linear Regression. *Measurement*, 338(jan28 3):b167–b167, 2009.

[19] E Gold. Www.e-gold.com. *Egold websit*, (1).

[20] Andy Greenberg. Crypto Currency. *Forbes*, 187(8):40–42, 2011.

[21] Mehmet Hadi Gunes. Complex network discovery: router-level internet topology mapping. Phd, University of Texas at Dallas, Richardson, TX, USA, 2008. ¡p¿Adviser-Sarac, Kamil¡/p¿.

[22] Mehmet Hadi Gunes and Kamil Sarac. Resolving anonymous routers in internet topology measurement studies. In *The 27th IEEE Conference on Computer Communications (INFOCOM 2008)*, pages 1076–1084, 2008.

[23] Mehmet Hadi Gunes and Kamil Sarac. Resolving ip aliases in building traceroute-based internet maps. *IEEE/ACM Transactions on Networking*, 17(6):1738–1751, 2009.

[24] C. Istvan K. Daniel, P. Marton and V. Gabor. Do the rich get richer? an empirical analysis of the bitcoin transaction network. *arXiv*, V3:1308.3892, 2014.

[25] Hasan Tarik Karaoglu, Mehmet Burak Akgun, Mehmet Hadi Gunes, and Murat Yuksel. Multi path considerations for anonymized routing: Challenges and opportunities. In *2012 5th International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–5, 2012.

[26] Hasan Tarik Karaoglu, Murat Yuksel, and Mehmet Hadi Gunes. On the scalability of path exploration using opportunistic path-vector routing. In *2011 IEEE International Conference on Communications (ICC)*, pages 1–5, 2011.

[27] Hakan Kardes and Mehmet Hadi Gunes. Structural graph indexing for mining complex networks. In *ICDCSW '10: IEEE 30th International Conference on Distributed Computing Systems Workshops*, page 99–104, Washington, DC, USA, 2010. IEEE Computer Society, IEEE Computer Society.

[28] Hakan Kardes, Mehmet Hadi Gunes, and Talha Oz. Cheleby: A subnet-level internet topology mapping system. In *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS)*, pages 1–10, 2012.

[29] Hakan Kardes, Mehmet Hadi Gunes, and Kamil Sarac. Graph based induction of unresponsive routers in internet topologies. *Computer Networks*, 2015.

[30] Hakan Kardes, Abdullah Sevincer, Mehmet Hadi Gunes, and Murat Yuksel. Six degrees of separation among us researchers. In *2012 IEEE/ACM*

International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pages 654–659, 2012.

[31] Hakan Kardes, Abdullah Sevincer, Mehmet Hadi Gunes, and Murat Yuksel. *Complex Network Analysis of Research Funding: A Case Study of NSF Grants*. Number XII in Lecture Notes in Social Networks. Springer Vienna, 2014.

[32] Kakajan Komurov, Mehmet Hadi Gunes, and Michael A. White. Fine-scale dissection of functional protein network organization by statistical network analysis. *PLoS ONE*, 4(6):e6017, 06 2009.

[33] Dániel Kondor, Márton Pósfai, István Csabai, and Gábor Vattay. Do the Rich Get Richer? An Empirical Analysis of the Bitcoin Transaction Network. *PLoS ONE*, 9(2):e86197, 2014.

[34] Steff Lewis. Regression analysis. *Practical neurology*, 7(4):259–264, 2007.

[35] Bingdong Li, Esra Erdin, Mehmet Hadi Gunes, George Bebis, and Todd Shipley. *An Analysis of Anonymizer Technology Usage*, volume 6613 of *Lecture Notes in Computer Science*, pages 108–121. Springer Berlin Heidelberg, 2011.

[36] Jörg Menche, Angelo Valleriani, and Reinhard Lipowsky. Asymptotic properties of degree-correlated scale-free networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 81(4):046103, 2010.

[37] S. Mercan, Ufuk Ozkanli, and Mehmet Hadi Gunes. Analyzing funding network in turkey. In *2nd International Symposium on Computing in Informatics and Mathematics (ISCIM 2013)*, Tirana, Albania, 09/2013 2013.

[38] Ian Miers, Christina Garman, Matthew Green, and Aviel D. Rubin. Zerocoin: Anonymous distributed e-cash from bitcoin. In *Proceedings - IEEE Symposium on Security and Privacy*, pages 397–411, 2013.

[39] Melanie Mitchell. Complex systems: Network thinking. *Artificial Intelligence*, 170(18):1194–1212, 2006.

[40] Tony Morelli and Mehmet Hadi Gunes. Video game industry as a social network. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 1183–1188, 2012.

[41] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *White Paper*, 1(2):1–9, 2008.

[42] Jeffrey Naruchitparames, Mehmet Hadi Gunes, and Sushil J. Louis. Friend recommendations in social networks using genetic algorithms and network topology. In *2011 IEEE Congress on Evolutionary Computation (CEC)*, pages 2207–2214, 2011.

[43] Neo4j. neo4j: World's Leading Graph Database, 2012.

[44] Peter Neubauer. InfoQ: Graph Databases, NOSQL and Neo4j. *InfoQ*, 2010.

[45] Mark Newman. *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA, 2010.

[46] Micha Ober, Stefan Katzenbeisser, and Kay Hamacher. Structure and Anonymity of the Bitcoin Transaction Graph. *Future Internet*, 5(2):237–250, 2013.

[47] K Panagiotou and a Steger. On the Degree Distribution of Random Planar Graphs. *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '11)*, pages 1198–1210, 2010.

[48] Daniel Ramos, Mehmet Hadi Gunes, Donica Mensing, and DavidM. Ryfe. *Mapping Emerging News Networks: A Case Study of the San Francisco Bay Area*, volume 424 of *Studies in Computational Intelligence*, pages 237–244. Springer Berlin Heidelberg, 2013.

[49] David Ryfe, Donica Mensing, Hayreddin Ceker, and Mehmet Hadi Gunes. Popularity is not the same thing as influence: A study of the bay area news system. *#ISOJ Journal*, 2(2), 2012.

[50] T. SABLIK. New private currencies like bitcoin offer potential and puzzles. *Publications,Richmondfed*, 7(6):676–678, 2013.

[51] David S. Shelley and Mehmet Hadi Gunes. Gerbilsphere: Inner sphere network visualization. *Computer Networks*, 56(3):1016 – 1028, 2012. ¡ce:title¿(1) Complex Dynamic Networks (2) {P2P} Network Measurement¡/ce:title¿.

[52] M. Shoaib, M. Ilyas, and M. Sikandar Hayat Khiyal. Official digital currency. *8th International Conference on Digital Information Management, ICDIM 2013*, pages 346–352, 2013.

[53] Lars Sthle and Svante Wold. Analysis of variance (ANOVA). *Chemometrics and Intelligent Laboratory Systems*, 6(4):259–272, 1989.

[54] Xiaogang Su, Xin Yan, and Chih-Ling Tsai. Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(3):275–294, 2012.

[55] Peter Šurda. *Economics of Bitcoin : is Bitcoin an alternative to at currencies and gold ?* PhD thesis, 2012.

[56] Guoxun Tian and Mehmet Hadi Gunes. *Complex Network Analysis of Ozone Transport*, page 87–96. Springer International Publishing, 2014.

[57] STAANFORD University. SNAP. *Forbes*, (8):40–42, 2011.

[58] David Yermack. Bitcoin Economics. *Technology Review*, 117(2):12, 2014.

[59] Christopher Zachor and Mehmet Hadi Gunes. *Software Collaboration Networks*, volume 424 of *Studies in Computational Intelligence*, pages 257–264. Springer Berlin Heidelberg, 2013.