Proceedings of the 9th International Conference on Applied Informatics
Eger, Hungary, January 29–February 1, 2014. Vol. 1. pp. 201–211
doi: 10.14794/ICAI.9.2014.1.201

Quantitative analysis of Bitcoin exchange rate and transactional network properties*

Imre Szücs, Attila Kiss

Eötvös Loránd University, Department of Information Systems, Budapest, Hungary
icsusz@gmail.com, kiss@inf.elte.hu

Abstract

The role of Bitcoin – open source virtual peer-to-peer money – in finance has become more important with the increasing acceptance by service providers. Nevertheless several financial institutes and governments explain their revulsion against Bitcoin, due to the unknown financial risks behind it which could have an impact on the global financial world. In this paper we examine the relationship between BTC/USD exchange rate and the network properties of the underlying transactional graph. The main goal of our research is to get a deeper understanding on the behavior of Bitcoin and ground further researches on exploring the financial risk. To characterize the transactional graph network analysis techniques, while to examine the relationship data mining and time series analysis techniques were used.

Keywords: Bitcoin, Network analysis, Graph, Data mining, Time series

MSC: MSC: AMS classification numbers: 62-07, 91D30, 91G80

1. Introduction

In recent years a new kind of payment system came into alive, which split the financial experts' opinion, the so called Bitcoin [1]. Bitcoin is basically a peer to peer payment system, which is also called cryptocurrency because both the creation and the transfer mechanism are controlled by cryptographic algorithms [2] [3].

Its development was mentioned firstly in 2008, and released in 2009 by Satoshi Nakamoto [4]. The identity of the author is unknown, and it is also unknown whether one person or a group stands behind the development.

*This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013).
This work was completed with the support of the Hungarian and Vietnamese TET (grant agreement no. TET 10-1-2011-0645).

From financial point of view one of the riskiest properties of Bitcoin is its extremely high volatility [5]. In 2011 the BTC/USD exchange rate moved from 0.3 to 32, than back to 2.0. In 2013 the Bitcoin system was attacked by several authorities, like FBI shut down the Silk Road on-line black market, China restricted BTC usage and the European Banking Authority issued a warning on the risks of using Bitcoin. Nevertheless Bitcoin spread as a payment system and more and more service provider accept the payment with it. It is still a question whether the digital money can be just a new evolutionary state in the money history.

The increasing popularity and the high financial risk together require analyzing Bitcoin deeper and try to understand its relation to the traditional financial world. In this paper we analyze the Bitcoin transactional network and specify that about 200 days are needed for the network to be formed. The relationship between the transactional network's graph properties and the BTC/USD exchange rate is also explored and the most important graph properties are determined.

2. Background

2.1. Data

Bitcoin transactional data can be downloaded and analyzed [6]. For the analysis we have used the Bitcoin transactions from the 1^{st} of January in 2009 to 1^{st} of April in 2013.

2.2. Bitcoin database and graph

The Bitcoin Network Data Tools provide the Bitcoin data in the following relational data structure:

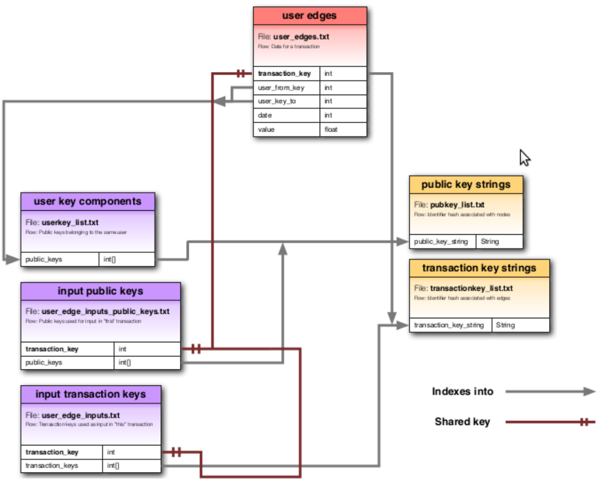

Figure 1: Bitcoin database schema [6]

For graph analysis purposes we used a directed, weighted, time-dependent graph representation of the user transactions where the vertices represents the Bitcoin users, the edges represents the transactions, the edge weights are the values of the transactions and all edge has a datetime property:

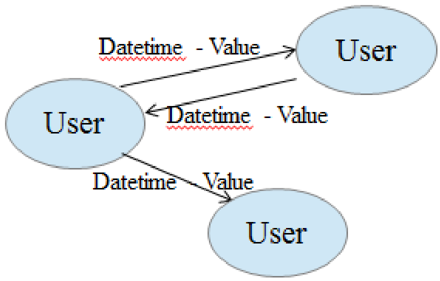

Figure 2: Bitcoin transactions' graph representation

The transactional graph contained 6.339.769 vertices and 37.450.461 directed edges.

2.3. Graph analysis environment

For general data manipulation purposes we have used the R environment (R version 3.0.3) [7], while for graph analysis purposes the igraph R package (version 0.7.1) was used [8].

3. Experiments

The first question had to be answered was "How long does it take the graph to be formed?" In one hand it is interesting from graph characteristic point of view, while on the other hand it is important from technical and performance point of view to determine how large graph should be handled and analyzed which characterize Bitcoin transactions well.

Using 100 and 200 days observation windows, the relationship between the properties of Bitcoin transactional graph and BTC/USD exchange rate was examined, and the most important factors were determined. As a measure of the predictive power, the correlation coefficients were calculated between the time delayed graph properties and the BTC/USD exchange rate.

3.1. Formation time

To determine the formation time several graph properties were analyzed in time. We have used 100 points in time with different observation windows from 10 days to 360 days with 10 days steps, and illustrated the average of the given graph property as a function of the observation window size.

Number of vertices and Number of Edges: Both Number of vertices and edges in the graph shows that in case the observation window is above 200 days, the graph became stable.

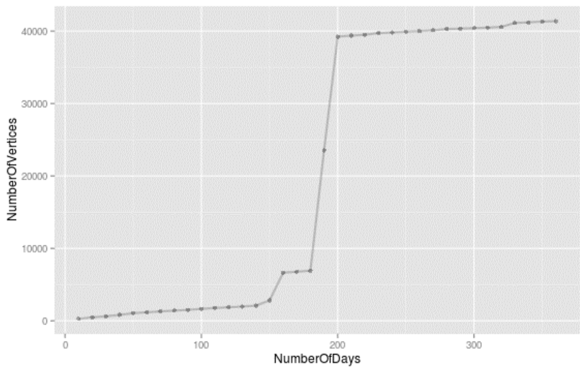

Figure 3: Bitcoin transactional graph: Number of vertices as a function of observation window in days

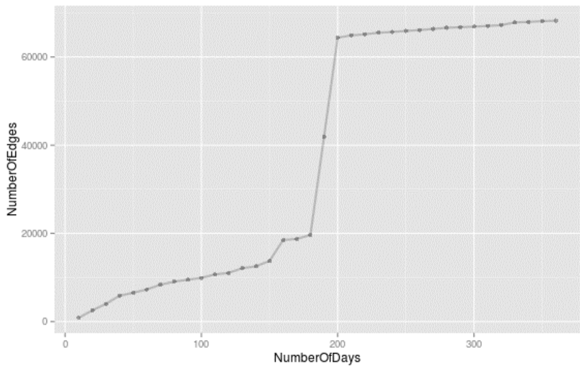

Figure 4: Bitcoin transactional graph: Number of edges as a function of observation window in days

Transitivity (clustering coefficient): measures the probability that the adjacent vertices of a vertex are connected. The transitivity is calculated as the ratio of the number of closed triplets and the number of connected triplets of vertices.

Diameter: The diameter of a graph is the maximum eccentricity of any vertex in the graph, or alternatively it is the greatest distance between any pair of vertices:

$$d = \max_{v \in V} \epsilon(v)$$

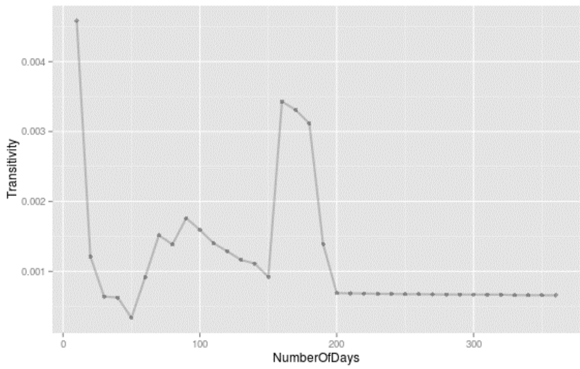

Figure 5: Bitcoin transactional graph: Transitivity as a function of observation window in days

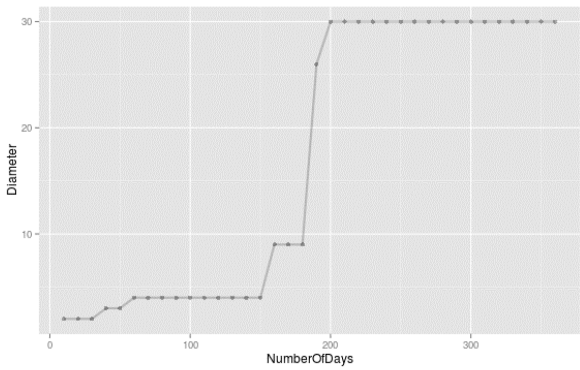

Figure 6: Bitcoin transactional graph: Diameter as a function of observation window in days

where $\epsilon(v)$ is the eccentricity of a vertex v.

Average path length is defined as the average number of steps along the shortest paths for all possible pairs of network vertices.

$$l_G = \frac{1}{n(n+1)} \sum_{i \neq j} d(v_i, v_j),$$

where n is the number of vertices in graph G, $d(v_i, v_j)$ is the shortest distance between vertices v_i and v_j.

Weakly connected components of a given directed graph means it has an undirected path from each vertex to every other vertex and vice versa inside the component. While in case of *strongly connected components* there is a directed path from each

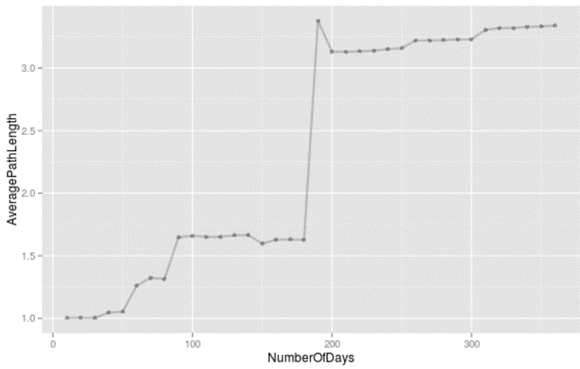

Figure 7: Bitcoin transactional graph: Average path length as a function of observation window in days

vertex to every other vertex and vice versa. *Average weakly and / or strongly connected cluster size* means the average of the number of vertices of the weakly / strongly connected components. The *maximal weakly connected cluster size* means the number of vertices of the largest weakly / strongly connected cluster in the graph.

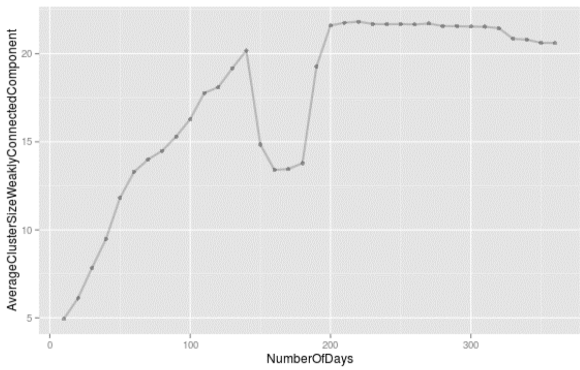

Figure 8: Bitcoin transactional graph: Average cluster size of weakly connected components as a function of observation window in days

Closeness centrality of a given vertex measures how many steps is required to access every other vertex from the given vertex, which can be calculated as 1/sum(d(v,i), i!=v), where d(v,i) means the length of the shortest path from vertex v to vertex i.

Degree centrality of a vertex is the number of its adjacent edges.

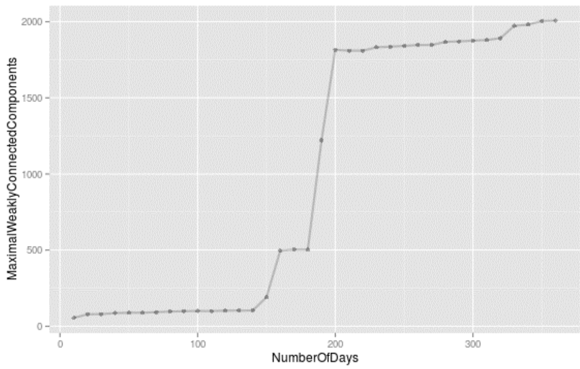

Figure 9: Bitcoin transactional graph: Cluster size of the largest weakly connected component as a function of observation window in days

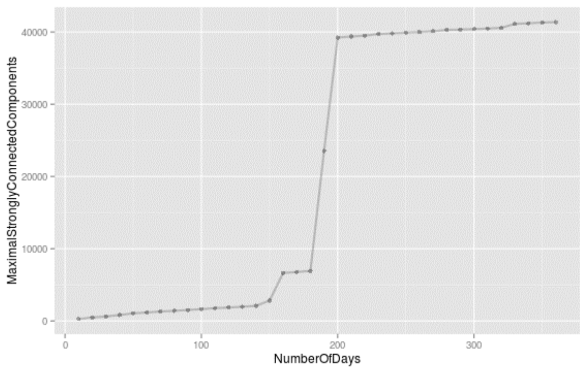

Figure 10: Bitcoin transactional graph: Cluster size of the largest strongly connected component as a function of observation window in days

Eigenvector centrality scores correspond to the values of the first eigenvector of the graph adjacency matrix.

Graph-level centrality score can be calculated from vertex-level centrality measure, with the formula C(G)=sum(max(c(w),w)-c(v),v), where c(v) is the centrality of vertex v. The graph-level centrality score can be normalized by dividing by the maximum theoretical score for a graph with the same number of vertices, using the same parameters.

Each graph property examination implies that the Bitcoin transaction graph's formation time is about 200 days. The main characteristic of the graph remain the same after 200 days, while before 200 days the properties varies considerably.

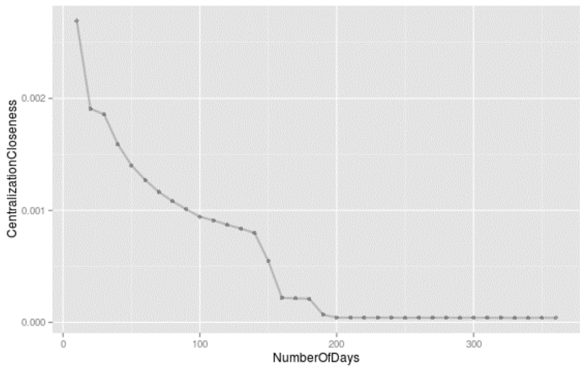

Figure 11: Bitcoin transactional graph: Closeness centrality as a function of observation window in days

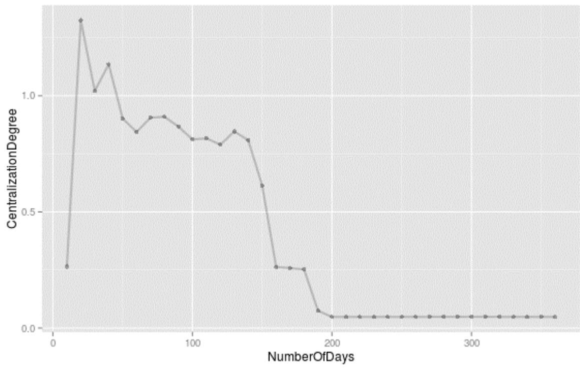

Figure 12: Bitcoin transactional graph: Degree centrality as a function of observation window in days

3.2. Relationship between graph properties and BTC/USD exchange rate

The relationship between the transactional graph properties and the BTC/USD exchange rate was measured by correlation, using graph properties calculated with 100 and 200 days observation window. The reason of the 200 days is the formation time of the transactional graph, while it was also interesting to use a shorter time window as well for being able to examine shorter effects.

Based on correlation it was found, that the most important factors are:

- Number of edges
- Diameter
- Average path length

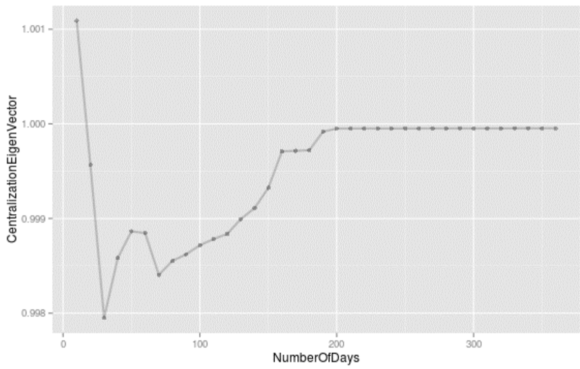

Figure 13: Bitcoin transactional graph: Eigenvector centrality as a function of observation window in days

Correlation coefficient	**Observation window**	
Graph property	**100 days**	**200 days**
Number of vertices	0,29	0,15
Number of edges	0,78	0,73
Transitivity	-0,05	0,13
Maximum Size of Strongly Connected Components	0,26	0,13
Diameter	0,7	0,69
Number of weakly connected components	0,17	-0,14
Maximum size of Weakly Connected Components	0,31	0,20
Average size of Weakly Connected Components	0,36	0,34
Average path length	0,83	0,81
Degree centralization	0,00	0,22
Eigenvector centralization	0,07	-0,14

Table 1: Correlation between Graph properties and BTC/USD exchange rate using 100 and 200 days observation windows

3.3. Predictive power

We have examined the correlation coefficient as a function of the time delay between the given graph property and the BTC/USD exchange rate, to simulate a kind of prediction power of the graph property. For the most important factors, the following result we got:

Figure 14 shows that the correlation decreasing nearly linearly with time delay, and the initial 0.75 correlation coefficient move to about 0.65 with a few days delay.

Figure 15 shows that the correlation reduction is stronger than in the case of number of edges, and in about 20 days it reaches about the 50% of its initial value.

Figure 16 shows that the correlation decreasing nearly linearly with time delay.

In case correlation is used as a measure of predictive power of graph properties

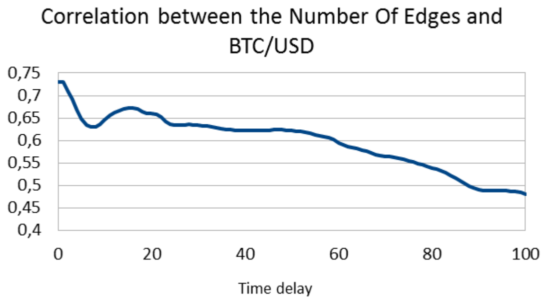

Figure 14: Correlation coefficient as a function of the time delay between the Number of edges and the BTC/USD exchange rate

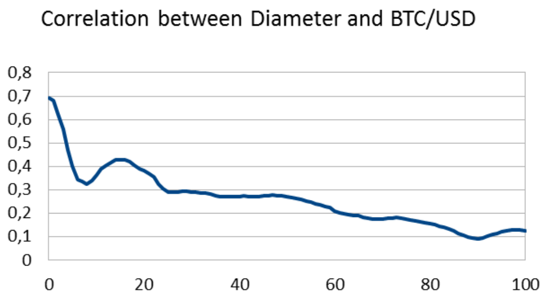

Figure 15: Correlation coefficient as a function of the time delay between diameter and the BTC/USD exchange rate

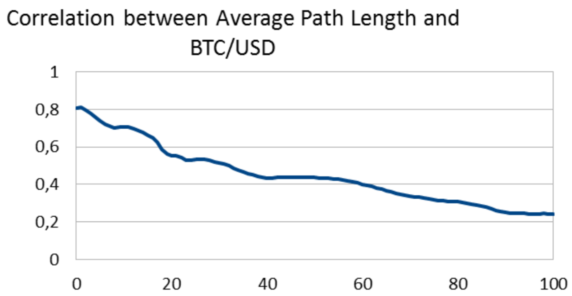

Figure 16: Correlation coefficient as a function of the time delay between average path length and the BTC/USD exchange rate

to BTC/USD exchange rate, it can be said that the predictive power of graph properties decreasing quickly in time.

4. Summary and future plans

In this paper we have analyzed the Bitcoin transaction graph with network analysis tools with the aim to get a better understanding of this much criticized payment system. We have shown that the transactional graph has a 200 days formation time, by calculating several graph properties. The relation to the traditional financial world was examined by measuring the relationship between the Bitcoin graph and BTC/USD exchange rate. It was found that the number of edges, diameter and the average path length in the graph has the stronger relationship to BTC/USD exchange rate. It was shown that the predictive power of the graph properties related to BTC/USD exchange rate decreasing quickly in time.

This paper was our first step in analyzing Bitcoin payment system. More detailed analysis of the transaction graph is required to get a clear overview on its behavior. Our future plan is to examine the existence and behavior of time-dependent clusters. It is also a plan to use community properties when analyzing the relationship of Bitcoin to the traditional financial world. And the final aim is to support inferring to financial risks related to Bitcoin with network analysis tools.

References

[1] BITCOIN, Wikipedia, `https://en.bitcoin.it/wiki/Introduction`

[2] REID, F. HARRIGAN, M., , An Analysis of Anonymity in the Bitcoin System, Security and Privacy in Social Networks (2013), pp 197-223, ISBN: 978-1-4614-4139-7

[3] BARBER, S., BOYEN, X., SHI, E., UZUN, E., Bitter to Better — How to Make Bitcoin a Better Currency, Financial Cryptography and Data Security (2012), pp 399-414, ISBN: 978-3-642-32946-3

[4] NAKAMOTO, S., Bitcoin: A Peer-to-Peer Electronic Cash System (2009), `https://www.bitcoin.org/bitcoin.pdf`

[5] BITCOIN Charts, `http://bitcoincharts.com/`

[6] BRUGERE, I., Bitcoin Network Data Tools, Laboratory for Computational Population Biology, University of Illinois at Chicago

[7] R – `http://www.r-project.org`

[8] CSARDI, G., NEPUSZ, T., The igraph software package for complex network research, InterJournal, Complex Systems 1695. (2006), `http://igraph.org`