Visual Analytics Project

Lorella Landi 1912333
Andrea Gaetano Ferrara 1887505
Giovanni Ruocco 1891327

13/09/2021

Bitcoin Transactions Visualization

A thesis presented for the course of Visual Analytics
MSc in Engineering in Computer Science

SAPIENZA
UNIVERSITÀ DI ROMA

Contents

1 Abstract

Purpose: The main purpose is to present a data visualization system to illustrate the data regarding the global phenomenon of Bitcoin giving an interesting insight of its behavior during its first steps in life. More precisely we will show its characteristics connections in the year 2010, year of its creation. We will illustrate below the main functionality and design of the system and its possible usage.

Analysis: The analysis is based on different values related to bitcoin transactions: while in the firsts graphs the analysis is based on the number of transactions, inputs and outputs, for the following graphs are also analyzed the value of inputs and outputs, in terms of bitcoins exchanged.

2 Introduction

2.1 Bitcoin

Bitcoin (₿) is a decentralized digital currency, without a central bank or single administrator, that can be sent from user to user on the peer-to-peer bitcoin network without the need for intermediaries. Transactions are verified by network nodes through cryptography and recorded in a public distributed ledger called a blockchain. The cryptocurrency was invented in 2008 by an unknown person or group of people using the name Satoshi Nakamoto.

2.2 The Project

The project is based on all the transactions that took place in 2010. The reason that prompted us to analyze that specific year is related to the aim of showing who are interested in the history of Bitcoin all the specific values of the transactions in the year in which the Bitcoin acquired monetary value, showing: the number of transactions, inputs and outputs, the first 3 daily buyers and sellers, the trend in the number of Bitcoin transactions and the PCA.

2.3 The Databases

2.3.1 Bitcoin Blockchain Historical Data

The Bitcoin database we analyzed was extracted from the Kaggle database "Bitcoin Blockchain Historical Data". From this we decided to study the transactions and informations to them connected:

```
HASH, BLOCK_TIMESTAMP, INPUT_COUNT, OUTPUT_COUNT,INPUT_VALUE,
OUTPUT_VALUE, FEE, INPUTS[INDEX, ADDRESSES, VALUE],
OUTPUTS [INDEX, ADDRESSES, VALUE]
```

General Information
Format : JSON
Dimension : 58 MB
Year : 2010
AS rule: #tuples * #transactions + (#INPUTtuples * $\sum input\_count$) + ($\#OUTPUTtuples * \sum output\_count$) = 2.105.790

2.3.2 Bitcoin Market price

The Bitcoin Market price we analyzed was extracted from **blockchain.info**. From this we decided to study the 2010's bitcoin values:

```
TIMESTAMP, PRICE
```

General Information
Format : CSV
Dimension : 2 KB
Year : 2010

3 Precomputation and Analytics

3.1 Precomputation

We initially spent some time analyzing the data at our disposal, making preliminary analyzes. As a result, the entire dataset present on kaggle had some non-essential values for our analysis. This made the job not easy, in fact the first thing we had to do was adapt the data to the visualization and to the precomputation, exctrating kaggle data in a json file and using python algorithms to synthesize the database for our purpose. Moreover inside this json file there was a nested json for each Input and Output, so our work was to inspect several times the internal values for how many inputs/outputs there were. The number of transactions per day and this graphical trend were also calculated using python scripts.

3.2 Analytics

Clearly behind all the visualization we show in the next section there is an important analytical work. To show the ranking of every day of the first 3 buyers and sellers, we dynamically calculate how much each single user spends/receives every day, to draw up the ranking chart. It was not possible to know in advance every single combination, since filters can be used on the chart and the user can select the transactions on which he wants to rank; indeed for the Transaction Graph with a simple toggle the user can filter out only the transactions that have a degree lower than the daily average, where in our case degree corresponds to the number of operations (inputs and outputs).

4 Visual Environment

The entire system was designed to allow the user to analyze all the transactions that took place in 2010 and the top 3 buyers and sellers day by day, graphs showing the trend of bitcoin over time by transaction number and value and the PCA.

Our project was thought mainly to give some insight and interesting information on the year most significative for the development of the Bitcoin phenomenon. We aimed to do this through several visualization briefly explained in what follows in the next sections.
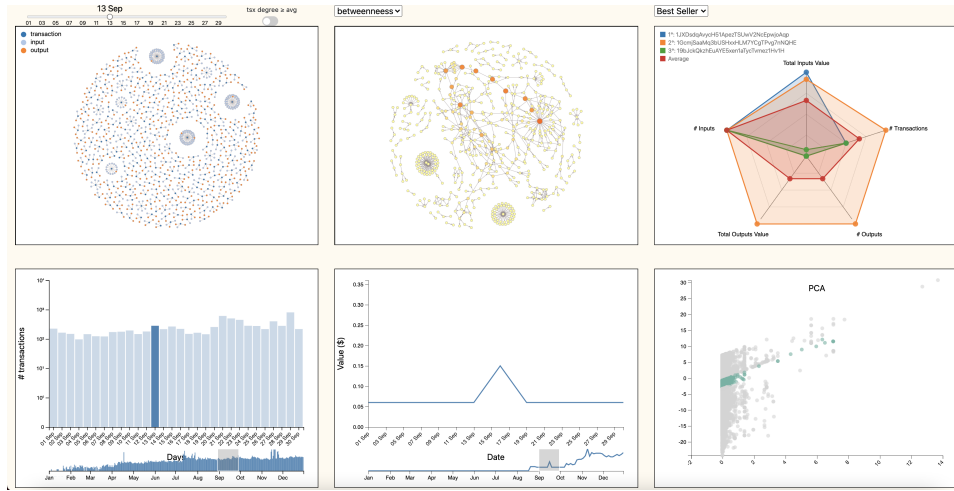
4.1 Transaction Graph

The Transaction Graph shows all the transactions that occurred in a given day. Each transaction is connected to its inputs and outputs which correspond to people in the network. The user can brush to select one or more transaction to show its information throughout the project. Moreover we implemented some filters to help select only the more significant transactions:

- Select one transaction by click

- Select more transactions by brush

- Filter by transaction degree

4.2 Network Graph

The Network Graph shows how users are connected: two users are connected to each other when one is input of a transaction where another user is output. With a simple toggle the color of a node changes depending on betweenness

centrality, starting from yellow to darker colors up to red. Moreover we implemented some filters to help select only the more significant transactions:

- Select one user by click

- Select more users by brush

4.3 Radar Chart

The Radar Chart displays the differences in values in the top 3 addresses selected through the brush or simply in a given day. The ranking parameter can be switched from best sellers to best buyers.Also a line for the average is implemented to better show the difference of the top 3 from the rest of the selection.

4.4 Transaction Bar Chart

The Transaction Bar Chart shows the number of transactions of each day. It can be brushed in any way but it is focused on single months in order to better show the connection with the other graphs.
By clicking on a bar the user can select the day to analyze.

4.5 Prices Line Chart

The Prices Line Chart shows the close value of Bitcoin for each day of 2010. We supposed that to an increase in transactions in a day an increase in value could be observed as a consequence, or vice versa. In order to bring the user to a conclusion we added a percentage increase tooltip for the days.

4.6 PCA

The PCA shows all the transactions used in the whole project, takes the different parameters such as 'input_value', 'output_value', 'input_count', 'output_count', 'fee' and then reduces them to a 2-dimension graph that can be displayed.

5 Case Studies

5.1 1st Case

Our first case study is the one in which a Bitcoin lover wants to ascertain whether an increase or decrease in the monetary value of Bitcoin corresponds to a change in the number of transactions. To help our user reach a conclusion on the matter, we have introduced a tooltip that shows, when the cursor is moved over the function that analyzes the value of bitcoin over time, which are the changes in price and transactions compared to the last recorded day.
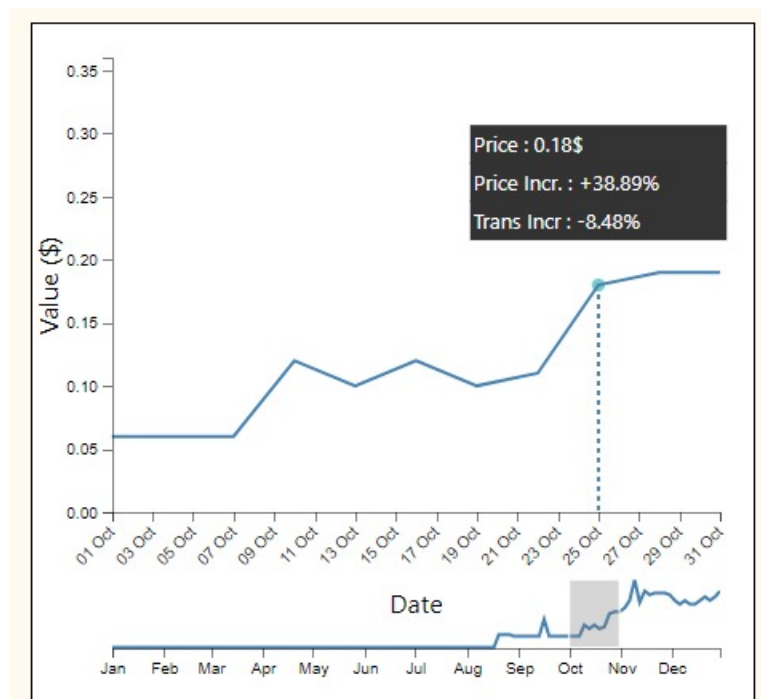
5.2   2nd Case

We all know that one of the reasons why Bitcoin has had this important growth is its decentrality. It is reasonable to assume that in the first months of bitcoin's life, when fewer users had access to the system than today, the risk of a prevalent user on the market is higher. Thanks to the Transaction Graph and the Network Graph the user can observe this phenomenon.
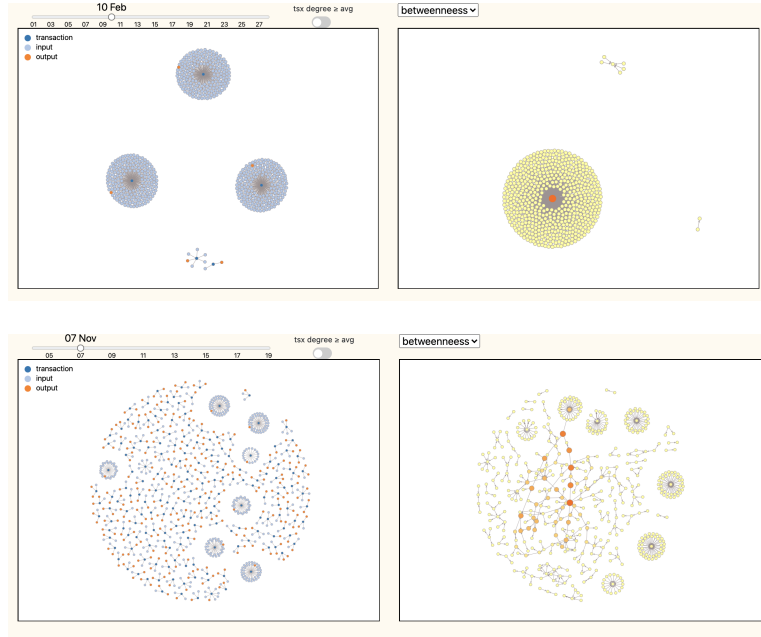
In the previous images we compared two days of 2010: February 10 and November 7. As we can see in the first image we have the impression of witnessing a phenomenon of a centralized system, in fact all the inputs and outputs converge in a single user. This is not the case in the second image, which clearly shows a fully distributed system.

9

6 Related Works

In the literature there are several different articles and visualization around Bitcoin, especially considering the mediatic attention it acquired in the last few years. However, most of them focus their attention on the latest transaction in stead of studying the past, like the following ones:

- **DailyBlockChain**

- **Bitcoin Interaqt**

- **Bitforce5**

7 Conclusions

In conclusion, this project can be both a starting point for its further development and improvement and a useful system that can be adapted for other types of databases. This work can certainly be deepened to analyze other features. One of these may be the possibility of having, for the first three graphs, a monthly view and not just a daily one. Furthermore, it could be interesting to provide the tools to compare the characteristics studied over the years.