UNIVERSITY OF PARIS 1 PANTHEON SORBONNE
MASTER 2 FINANCE TECHNOLOGY DATA
2022-2023

Network Analysis:
Analysis report on 2 days of Bitcoin transaction.

Group member:
Julien Dos Reis

- ***Introduction***

Bitcoin is a decentralized digital currency, there is no central authority behind it to issue or guarantee the money. Bitcoin, as other cryptocurrencies, operates on a peer-to-peer network, also called Blockchain, allowing transactions between users without the need for intermediaries. At the beginning of 2023, Bitcoin corresponds to around 40% of market shares according to coin360.com (1), followed by Ethereum.

These digital assets are also called cryptocurrencies because of the way they operate in the market. Cryptography is used to secure transactions and avoid coins stealing. These are algorithms, like SHA-256 for Bitcoin, used to create signatures for each transaction which is then added to the blockchain. This is a way to ensure that each transaction added to the blockchain is unique.

Additionally, encrypted public-keys are used to manage digital wallets. As transactions, each wallet is unique and refers to his public-key used to make transactions on the blockchain.

Due to his attributes, Bitcoin gained popularity over the years up to becoming a mean of payment for users. The number of Bitcoin transactions has significantly increased, tracking these transactions has become crucial to understand the currency’s use and value.

Since Bitcoin transactions are on the Blockchain, they’re publicly available. Each transaction is issued by a wallet, with a corresponding public-key, to another wallet also corresponding to a public-key. We will work here on a dataset corresponding to 2 days of Bitcoin transactions. This dataset is simplified to make it easier to understand as economic transactions. Each line corresponds to a payment from one entity to another.

- <u>Part I: Analysis of the dataset</u>

Please note that there are different units to value bitcoin in this dataset:

- In Satoshis: 1 BTC = 100 000 000 (it’s said in the GitHub repository that 1 BTC = 100 000 Satochis but I think this is an error since it doesn’t match with the value in USD indicated)
- In USD: since the data is from 07/08/2016 to 07/10/2016, at this time, according to Yahoo Finance data on BTC-USD was 1 BTC = 660 USD approximately

After creating a value in BTC column in a data frame we proceed to some descriptive statistics. First, we get rid of change transactions and only keep true payments to better analyse transactions between entities.

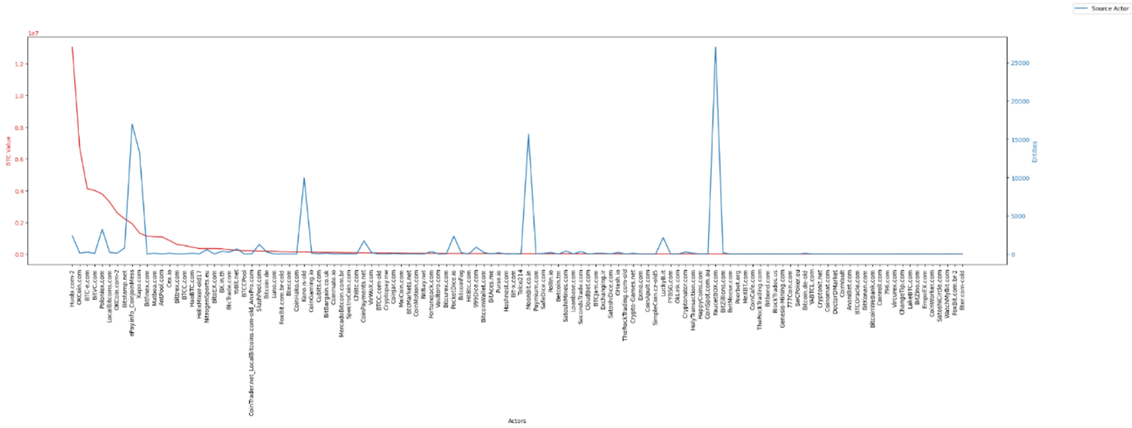
After that we’re left with 801 491 payments in our dataset and only 347 590 transactions by looking at unique hashes within the dataset.

We will now provide analysis on known entities in our dataset. To do so, we chose to keep the columns “V1_src_actor” since there is more unique values than for the V2 one, same for destination entities.

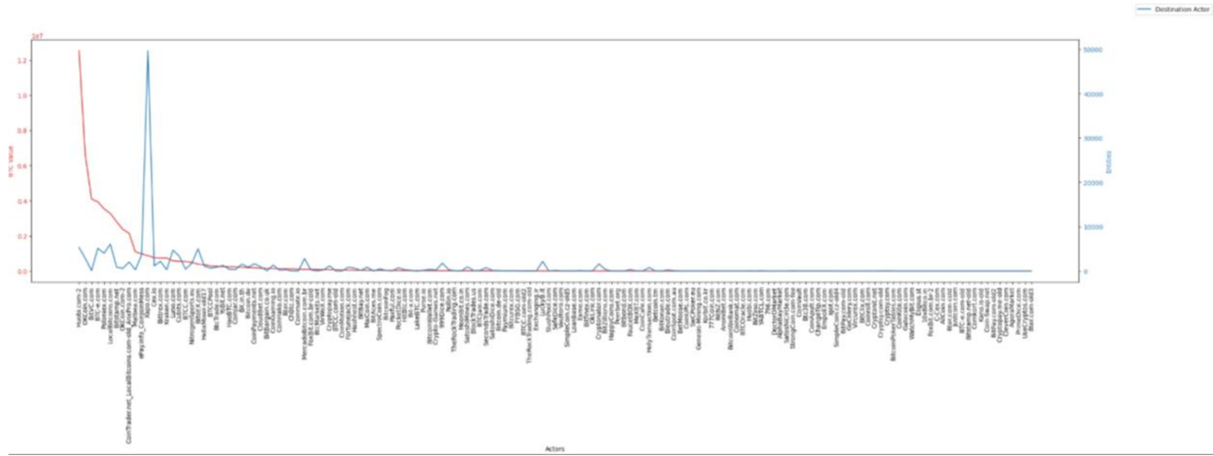
Source Actor	valueBTC	valueUSD	V1_src_actor
Huobi.com-2	19999.416219	1.304684e+07	2373
OKCoin.com	10204.016403	6.652176e+06	132
BTC-e.com	6322.698486	4.125722e+06	253
BitVC.com	6152.396558	4.012878e+06	96
Poloniex.com	5804.522373	3.789886e+06	3234
LocalBitcoins.com	5007.670470	3.276256e+06	218
OKCoin.com-2	3965.303604	2.595283e+06	125
Bitstamp.net	3401.235313	2.224462e+06	811
ePay.info_CoinJoinMess	2939.717425	1.918270e+06	16946
Xapo.com	2001.635154	1.315264e+06	13243
Bitfinex.com	1734.244777	1.127189e+06	1
Matbea.com	1670.227182	1.092923e+06	119
AntPool.com	1663.000000	1.082395e+06	2
Cex.io	1310.078262	8.540749e+05	123
Bittrex.com	927.541344	6.089071e+05	9

Destination Actor	valueBTC	valueUSD	V1_dst_actor
Huobi.com-2	19218.548653	1.253765e+07	5337
OKCoin.com	10131.666282	6.603969e+06	2861
BitVC.com	6285.946077	4.102348e+06	102
BTC-e.com	6044.149205	3.941943e+06	5110
Poloniex.com	5413.402797	3.539162e+06	4027
LocalBitcoins.com	5008.833768	3.277529e+06	6068
Bitstamp.net	4298.205609	2.808631e+06	878
OKCoin.com-2	3632.000391	2.373419e+06	599
CoinTrader.net_LocalBitcoins.com-old_AnxPro.com	3301.894700	2.149989e+06	2050
Matbea.com	1666.857571	1.090953e+06	286
ePay.info_CoinJoinMess	1496.964495	9.778128e+05	3669
Xapo.com	1335.717405	8.733877e+05	49678
Cex.io	1169.754878	7.627218e+05	1161
Bittrex.com	1135.561856	7.418130e+05	2201
Kraken.com	1131.388496	7.377953e+05	293

By comparing amounts and number of transactions between source and destination actors, we can see that some entities are in both tables. This can give us hints about which Entities are central in the Bitcoin industry. We will try to see further proofs by plotting results.

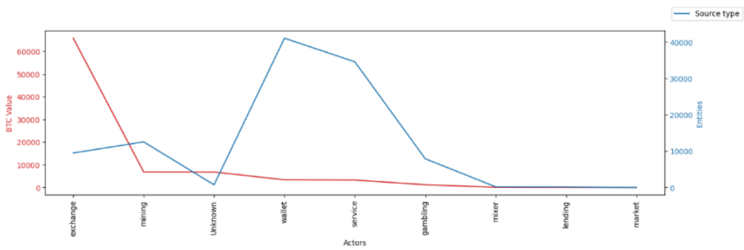


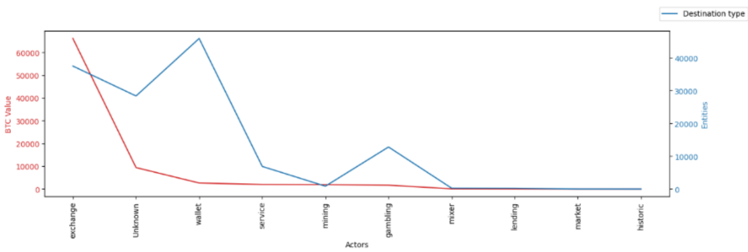
This first graph is showing the entities amount of bitcoin sent and number of transactions done. We can that see entities sending huge amounts are not the ones doing the higher number of transactions. As we can see, Faucet.com is doing huge amounts of micro-transactions, this kind of behaviour must be considered when talking about centrality of entities. We will see after is these patterns can be explained.



This is slightly different for Entities receiving the transactions since that, except for Xapo.com and Huobi.com-2, number of transactions seems to be much more related to the amount of Bitcoin received.

We will now focus on the same kind of analysis but with the type of actors found in the dataset. To be noted that there is a huge amount of unknown type of actors while knowing entities doing the transactions. To reduce the impact of these unknown type of actors, I've proceeded to a quick investigation by looking at some entities doing numerous transactions to manually fill in the type, as for example with Huobi.com or ePay. You can see all entities affected by this changes in exchanges and service lists in the code.





Both graphs above summarize Bitcoin transactions in value and in count for each type of actor. We can see that exchanges are the core of huge amounts of Bitcoin transactions while wallets retrieve huge number of micro-transactions. By looking at these graphs, we can ask ourselves why there is such differences between type of actors?

Let's look at the descriptive statistics of this dataset for value in Bitcoin transactions.

```
count    801491.000000
mean          2.983354
std          52.073332
min           0.000000
25%           0.000800
50%           0.012104
75%           0.113504
max       10000.000000
```

We can see that up to 75% of transactions are below 0.1135 BTC (=around $75 USD).

We can easily say that most of these small transactions are provided by wallet type of actors while exchanges transactions should be mainly in the last quartile of this table.

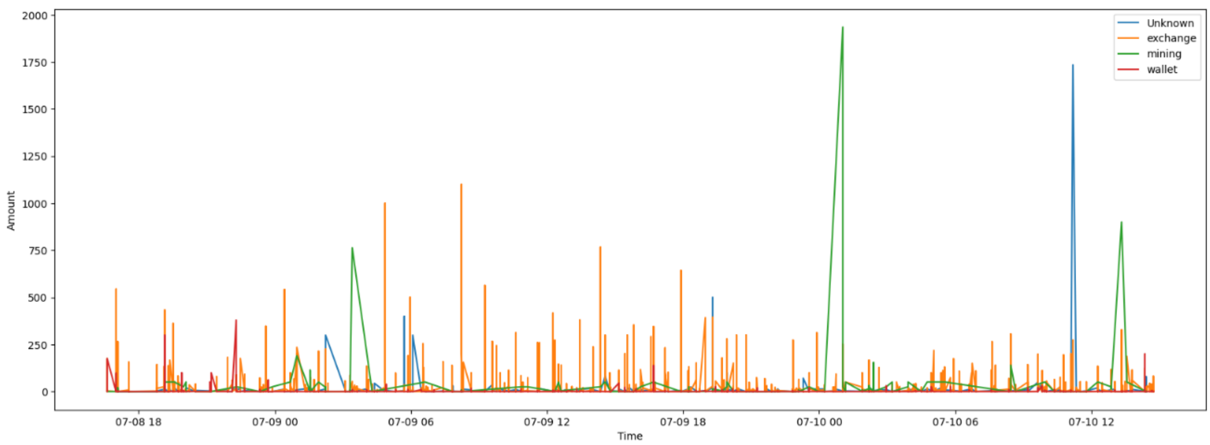
This differences in transactions patterns between type of entities perfectly reflects the diverse range of uses cases about Bitcoin transactions.

Since exchanges are platforms ensuring trades of different currencies between buyer and sellers, they must provide liquidity to the market to facilitate buying and selling. It is also a way for users to enter for the first time in the market with the possibility to trade cryptocurrencies with traditional currencies such as USD. So the amount of Bitcoin has to be superior to other type of entities.

Concerning wallet type of actor, they also play a big role. This is a way to store and manage cryptocurrencies. This role is similar to traditional banks. That's why there is a lot of transactions happening with these platforms but for less amounts than other such as exchanges.

Since the dataset is providing a timestamp, informing at which minutes a transaction happened, we will now try to see if we can highlight intra-day patterns focusing on this type of actors that we have.

Distribution of the amount of transaction for the 3 most important source actor, without considering Unknown type. We can't see clear patterns of transactions in these 2 days. But as we can see, exchanges actors are much more active, compared to the others, where there is constantly transactions happening.



- <u>Part 2: Network Analysis</u>

We will now focus on a network analysis of these transactions. This will help us to better understand structure and dynamics of this complex system. We will explore relationships between different actors in the network and try to identify patterns but also anomalies. This allows us to identify clusters of frequently involved actors.

To study these concerns, we will work on the 100 most important transactions of the dataset.

```
Number of nodes: 32
Number of edges: 57
Average degree: 3.5625
```

This gives us a total of 32 nodes with 57 edges. A node corresponds to an entity. Edges are links between two nodes, for our case it's corresponding to a transaction from an entity to another.

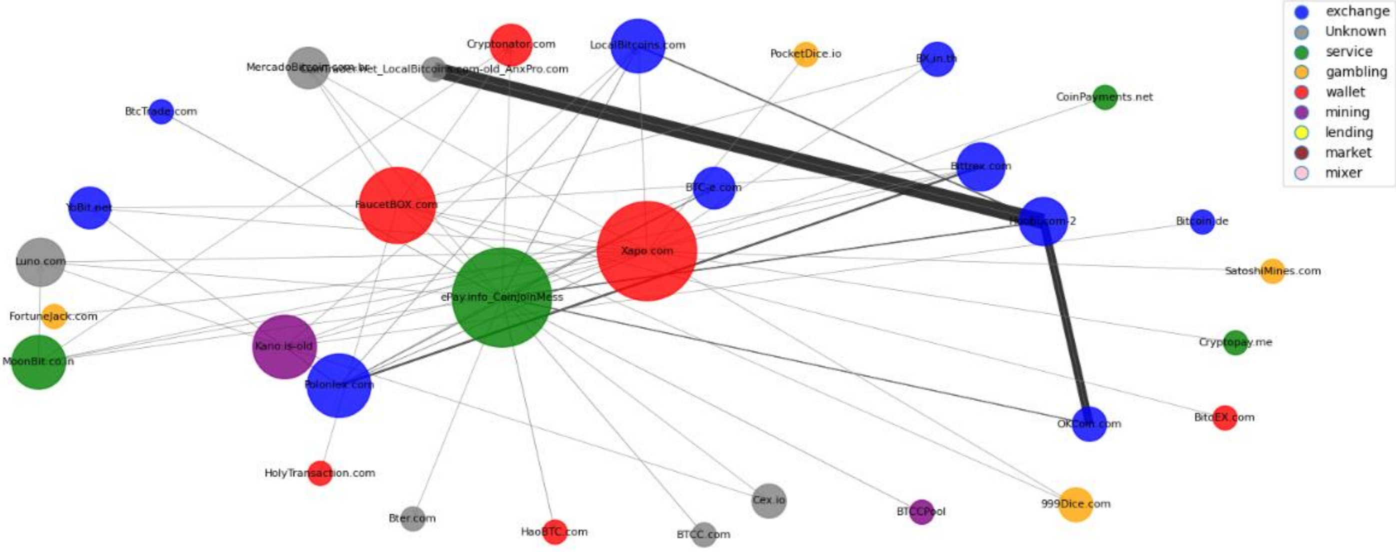
The average degree is 3.6, which shows the average number of edges connected to each node in the graph.

Following this, we will build our network graph but first we need to compute some parameters, such as degree centrality. There is different ways to compute centrality for network analysis such as betweenness centrality or closeness centrality, but degree centrality seems to be more accurate for our topic since it is used to measure the popularity of a node. The more edges will be linked to a node, the higher the degree centrality score will be. So we compute this parameter to assign each entities her centrality score. This will allow us to modify the size of a node proportionally to his centrality score.

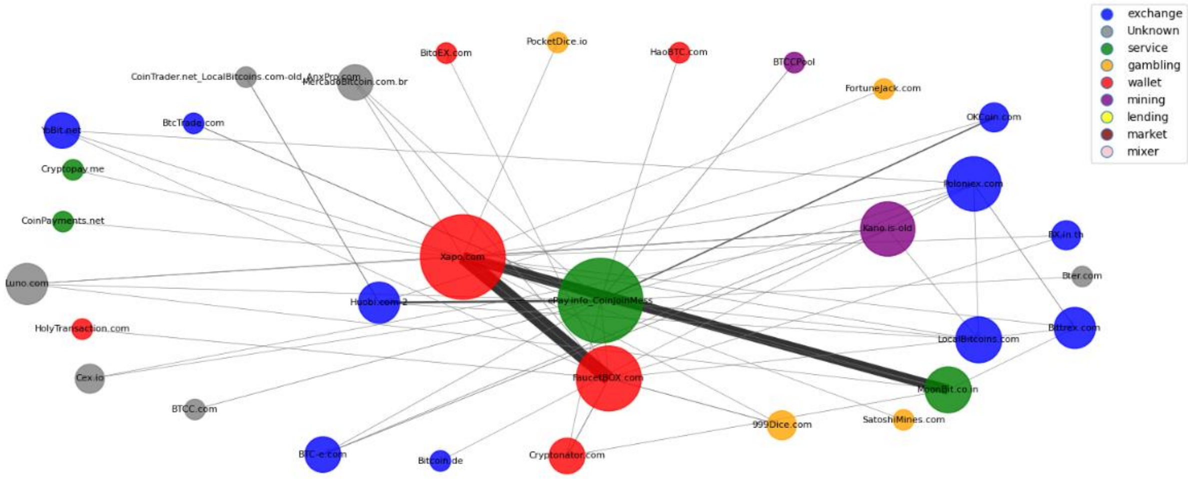
We then assign a colour to each type of entities within our dataset. By doing so, we will assign each node his corresponding type, this will make the graph much more understandable.

Finally, we assign a size parameter to edges, two type of sizes: one regarding the amount of transactions, the second regarding the number of transactions between entities.

When plotting these graphs, we will use the spring layout as layout for positions of nodes. Their positions on the graphs are based on the connectiveness between each other. Nodes connected by strong edges are closer together while nodes connected by weak edges are farther apart.



In this first graph, edges size is based on amount of transactions between the two nodes, in BTC. We can see a huge amounts between Huobi.com-2 and coinTrader.net even though there isn't a huge number of transactions. Moreover, we also see that the biggest nodes are wallets and service type of actors.



This time, by setting the edge size according to the number of transactions, we can see different links between entities. The ones highlighted here are the connections between Xapo.com with FaucetBOX.com and MoonBit.co.in. To be noted that edges linked to Huobi.com-2 are almost the same size has the others.

These discoveries made us think ask ourselves what the explanation for these strong links between specific entities are. We haven't managed to find additional information about it, FaucetBOX.com has closed, but having this kind of strong edges between specific entities in only 2 days of Bitcoin transactions seems weird in the first place.

- ***Conclusion:***

In conclusion, the analysis of Bitcoin transactions plays a crucial role in understanding the dynamics and structure of the cryptocurrency market. Researchers, developers, and other stakeholders can gain valuable insights by studying the behaviour of individual actors in the network, as well as the overall health and security of the network.

As a result of this project, we might leave with more questions than at the beginning but that's why having a proper understanding of the stakes is very important to ensure stability on the market.

This technology is designed in a way that it makes it difficult to understand and study. I think that there is a lot more to do when analysing these kinds of networks' graphs. Moreover, additional data could have helped with that. Having more days of transactions could have helped understanding patterns, but we would have to deal with huge amounts of data. Also, more information about the actors, but we understand that this is difficult considering the way cryptocurrencies transactions are build, that's why it is challenging. While building this project, I was also asking myself if having more recent data would change some patterns we've seen. Indeed, since 2016, the cryptocurrency market has changed and developed a lot.

(1)

Source:

- https://corporatefinanceinstitute.com/resources/cryptocurrency/cryptocurrency-exchanges/
- http://cazabetremy.fr/Teaching/bitcoinClass/BitcoinNetwork.html
- Exploring the Bitcoin Network, by Annika Baumann, Benjamin Fabian and Matthias Lischke, Institute of Information Systems
- Dissecting bitcoin blockchain: Empirical analysis of bitcoin network (2009–2020), by Pranav Nerurkar et al., Science Direct

Contributions: Working alone