Analysis of Tainted Transactions in the Bitcoin Blockchain Transaction Network

María Óskarsdóttir(✉), Jacky Mallett, Arnór Logi Arnarson, and Alexander Snær Stefánsson

Reykjavík University, Reykjavík, Iceland
{mariaoskars,jacky}@ru.is

Abstract. Blockchain technology, with its decentralised peer-to-peer network and cryptographic protocols, has led to a proliferation of cryptocurrencies, with Bitcoin at the forefront. The blockchain publicly records all Bitcoin transactions which can be used to build a dynamic and complex network to give a representation of the transactions in the underlying monetary system. Despite the cryptographic guarantees there exist inconsistencies and suspicious behavior in the chain. We reported on two such anomalies related to block mining in previous work. In this paper, we build a network using bitcoin transactions and apply techniques from network science to analyse its complex structure. We focus our analysis on sub-networks induced by the two sets of anomalies, and investigate how inequality in terms of wealth and anomaly fraction evolves from the blockchain's origin. Thereby we present a novel way of using network science to detect and investigate cryptographic anomalies.

Keywords: Bitcoin · Transaction network · Cryptography · Blockchain

1 Introduction

The blockchain is a publicly available ledger that stores all transactions made using bitcoin, the first cryptocurrency. The blockchain technology, proposed by Nakamoto in 2008, is based on an open peer-to-peer network to authenticate transactions using cryptographic technologies and implement a decentralized distributed digital ledger. Its introduction has led to a proliferation of cryptocurrencies in recent years [16]. The public bitcoin blockledger is now –12 years later– the most prominent and impactful version. To date, it records over half a billion bitcoin transactions which it stores in 620,000 blocks on the blockchain. In total, 18 million bitcoins are currently stored in over 46 million digital wallets, accompanied by details of the transactions they have been used in. The impact of this novel technology and the accompanying financial system is already considerable and it has attracted researchers from various disciplines, including cryptography, economics and network science.

By construction, the bitcoin blockledger lends itself extremely well to network analysis since all transactions using the ledger are publicly recorded, with

R. M. Benito et al. (Eds.): COMPLEX NETWORKS 2020, SCI 944, pp. 571–581, 2021.
https://doi.org/10.1007/978-3-030-65351-4_46

information about both the originator and the recipient. The dynamic nature of blockchain, the vast amount of transactions, intricate patterns, richness of node and edge features, exogenous effects (such as of markets and the economy) all contribute to the complexity of the network and its analysis. The bitcoin transaction network has been studied before to some extent, including investigation of the acquisition and spending behaviour of bitcoin owners [19]. The network shows evidence of the Pareto principle during the first four years, in that linear or sub-linear preferential attachment drive the network's growth and wealth distribution [9]. More recently, there has been a data driven analysis of price fluctuations, user behaviour, and wealth accumulation in the bitcoin transaction network, including an investigation of the richest wallets [17]. Finally, an analysis of the transaction network for the first nine years after its creation identified a causal relationships between the movements of bitcoin prices and changes of the transaction network topology [4]. As the bitcoin infrastructure has evolved, a number of measures have been introduced to address the inherent scaling limitations of a peer-to-peer network, a recent review of research on the bitcoin transaction network, identified three types of these networks, namely the Bitcoin Address Network, the Bitcoin User Network and the Bitcoin Lightning Network. In addition, the authors conclude that distribution of bitcoin is very uneven and the network is becoming increasingly more sparse [21].

Another stream of research is focused on anomalies and suspicious behaviour in the bitcoin blockledger using data science and machine learning. In an attempt to find anomalous transactions, [18] extracted features from the transaction network, from the origin until 2014, and applied k-means clustering to find outliers. Similar approaches have been proposed by other researchers [14,15]. Some studies investigate certain types of suspicious behaviours. Firstly, to identify ponzi schemes, transactions and wallets related to known schemes were extracted and compared to regular transactions and wallets in a supervised learning setting [3]. Secondly, researchers have looked into money laundering specifically, using network methods, in particular network representation learning and supervised machine learning models [8]. Recently, Elliptic[1] introduced a public data set which contains several sub-networks for the blockchain transaction network, with rich node features and labels for licit and illicit transactions. Researchers have trained several supervised learning methods to detect illicit transactions and compared their performance [22]. Others have also worked with the Elliptic dataset [1,11,20], for example using active learning to address the high class imbalance in the data set [11].

In spite of the blockchain's structural and operational properties that are designed to safeguard it, i.e. the decentralized peer-to-peer network, cryptographic protocols, validation of transactions, openness etc., inconsistencies and suspicious behaviour have been observed and reported. These have been connected with colluding miners [6], enhanced performance mining [5,7],

[1] Elliptic is a cryptocurrency intelligence company focused on safeguarding cryptocurrency ecosystems from criminal activity.

the so-called Patoshi pattern which appears in the first 30,000 blocks [13] and selfish mining, where miners publish the blocks they mine selectively [10].

In this paper we use network science to analyse the complex network of bitcoin transactions with respect to two particular anomalies which we have identified in blocks mined in the early years of the blockchain [12][2]. Given the magnitude of these anomalies –the blocks in question represent well over 3 million bitcoin– we investigate whether they may have led to false conclusions about some aspects of bitcoin transactions. We construct sub-networks of transactions that originate with the anomalous, or *tainted*, blocks and compare the structural properties of the sub-networks with the full network as well as sub-networks that arise from non-tainted blocks. Furthermore, motivated by the analysis of wealth distribution presented by Kondor et al. (2014) [9] and irregularities observed there, we compare the evolution of Gini coefficients of node features in the various sub-networks.

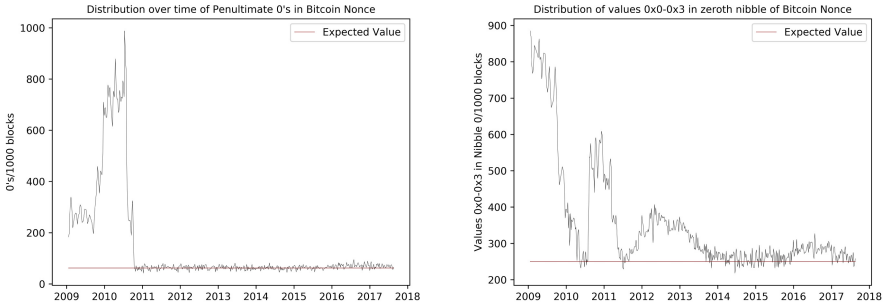
In the next section we discuss the two anomalies on which the analyses in this paper are based. Then we describe our methodology and present the results. The paper concludes with a summary of our findings and directions for future work.

2 Background

The origin story of bitcoin is that the technology originated with a posting by a Satoshi Nakamato to the cryptography mailing list in 2008, followed by a slow expansion in 2009-10 as early adopters installed mining software and began creating bitcoins. Although there has been some question as to whether a single individual could have developed and tested this system, simply due to the range of expertise required, this story has been broadly accepted by researchers.

At the end of 2019 we performed a simple frequency analysis of the hexadecimal values (nibbles) by position, in the bitcoin blockchain [12]. This revealed two distinct anomalous patterns, both in the nonce which is a key part of the proof of work performed by all miners to obtain bitcoins. One anomaly occurs in the first hexadecimal position (nibble) of the block's nonce field as shown in Fig. 1b where in a disproportionate number of blocks this has a value in the range 0–3, and the other is in the penultimate position of the nonce where an abnormal number of 0's occur in the first 18 months of mining, Fig. 1a. We refer to these as the P anomaly and the the Z anomaly, respectively. Both patterns seem to be associated either with the originators of bitcoin or very early adopters. The Extended Patoshi anomaly in the first nibble of the nonce is a notable feature of the first months of mining, part of which has already been attributed by Sergio Lerner to mining by Nakamato. The second, “”penultimate zero”, pattern can also be seen almost from the start of mining, and is either part of Nakamoto's mining, or that of a very early adopter. After accounting for the expected number of blocks that would contain these values, (6.25% in the penultimate zero case,

[2] The paper is currently under review, but will be shared upon request.



(a) Z anomaly (b) P anomaly

Fig. 1. Anomalous patterns discovered by frequency analysis of the hexadecimal values by position in the bitcoin blockchain.

and 25% in the Patoshi anomaly in the first nibble), we estimate that approximately one third of all coins mined at the first difficulty level are obtained from blocks mined with these features. Across the entire ten years of both patterns, well over 3 million bitcoins appear to have been obtained from blocks with these distinguishing features. The magnitude of these two patterns clearly warrants further investigation into any associated patterns in the transactions associated with the coins mined in these blocks. Previous research into early transactions in the bitcoin network has thrown up evidence of suspicious clusters, notably Shamir and Dorit's work [19] which discovered a large number of coins being progressively consolidated into a small number of apparently connected wallets, however generally research in this area has not had a clear marker in the blocks themselves on which to attach suspicion.

3 Methodology

3.1 Bitcoin Transaction Network

To carry out our analysis, we extract the entire bitcoin blockchain from origin to November 2019. Using these blocks, we create a database of transactions, with information about the *from* transaction and one or more *to* transactions which correspond to the movement of bitcoin between wallets. Wallets that received the miner's reward coins (otherwise known as coinbase transactions) from blocks with the two patterns are marked as tainted, and as these coins are transferred to other wallets, the percentage taint for each pattern is calculated and updated for the receiving wallet. This allows us to accrue information on the from and to nodes (wallet addresses) of the transaction, as well as the amount that was transferred, the transactions' tainted P ratio and tainted Z ratio and the timestamp of each transaction. In this way we obtain an edgelist of timestamped transactions from which we create a directed network.

3.2 Generation of Sub-networks

Having identified two types of anomalous transactions in the coinbase, namely the Z and the P anomaly, we continue to investigate their prominence in and effect on the bitcoin transaction network. To do this, starting from the full network, we extract sub-networks of transactions that have an origin with a specific set of coinbase transactions. We consider five sets of coinbase(cb) transactions as listed below.

$^{T}\mathcal{Z}$ = {cb|The Z anomaly is in the nonce of the cb block}
$^{T}\mathcal{P}$ = {cb|The P anomaly is in the nonce of the cb block}
$^{T}\mathcal{Z} \cap {}^{T}\mathcal{P}$ = {cb|The Z and the P anomalies are in the nonce of the cb block}
$\neg {}^{T}\mathcal{Z}$ = {cb|The Z anomaly is not in the nonce of the cb block}
$\neg {}^{T}\mathcal{P}$ = {cb|The P anomaly is not in the nonce of the cb block}

We create a sub-network for each set using snowball sampling. In the snowball sampling, we start off with a sub-network of source nodes that consists of the coinbase transactions in the respective set. Any transaction that is linked to one of these source nodes in the full network is added to the sub-network. Subsequently, any transaction in the full network that is linked to one of the most recently added transaction in the sub-network, is also added to the sub-network. This process is repeated until no more transactions can be added. Since the full network is timestamped and directional, the process will terminate.

As a result, we obtain, in addition to the full network –which we refer to as *All*– five sets of sub-networks, each one originating with the sub-sets listed above. We refer to these as *Tainted Z*, *Tainted P*, *Tainted P&Z*, *Not Z* and *Not P*, respectively. These sub-networks and the full network are created for each month starting in January 2010 until May 2012.

Due to the size of the entire dataset it is not feasible to build the sub-networks with the snowball sampling technique using all the nodes in each set. Therefore we randomly sample 1000 nodes from each set before doing the snowball sampling. This is repeated ten times for each sub-network in each month that we analyse. The values shown in the plots below are the mean value for each measure in these ten samples.

3.3 Network Measures

In order to compare the characteristics of the sub-networks to those of the full network, we consider several network measures.

First we measure basic properties of the networks. The first three basic measures are the number of nodes, density and diameter [2]. Number of nodes is simply the total number of nodes in the respective sub-network. The second measure is the network's density, or the number of edges divided by the maximum possible number of edges. It gives an intuition of how well connected the network is. Finally, diameter measures the length of the longest shortest path in the network. For any given pair of nodes, there is a path between them that is shorter than any other path between them. The diameter is the longest of such

paths in the network and represents the size of the network. Since computing the shortest path between all pairs of nodes in a network can get quite time consuming as networks grow in size, we randomly sample 1000 pairs of nodes from each network and use those pairs to estimate the networks' diameters.

Based on Kondor et al. (2014), we focus on the Gini coefficient, clustering coefficient and the degree correlation to quantify the inequality in the network and sub-networks [9]. Firstly, we use the Gini coefficient to characterize the heterogeneity of the distribution of in-degree, out-degree, transaction amount, tainted Z ratio and tainted P ratio. Generally, the Gini coefficient is defined as

$$G = \frac{2\sum_{i=1}^{n} i x_i}{n\sum_{i=1}^{n} x_i} - \frac{n+1}{n} \tag{1}$$

where $\{x_i\}$ is a monotonically non-decreasing ordered sample of size n. Thus, $G = 0$ indicates perfect equality, or every node being equal in terms of the value being considered, whereas $G = 1$ indicates complete inequality. As in [9] we measure this for the distribution of in-degree, out-degree and transaction amount, but in addition we compute the distribution of tainted P and tainted Z ratios.

Secondly, we look at the assortativity or degree correlation of the network [2]. We compute it using the Pearson correlation coefficient of the out- and in-degrees of connected node pairs

$$r = \frac{\sum_e (j_e^{out} - j^{\bar{out}})(k_e^{in} - k^{\bar{in}})}{L\sigma_{out}\sigma_{in}} \tag{2}$$

where for the edge e that links node v_{from} to v_{to}, j_e^{out} is the out-degree of node v_{from} and k_e^{in} is the in-degree of node v_{to},

$$k^{\bar{in}} = \sum_e k_e^{in}/L \quad \text{and} \quad \sigma_{in}^2 = \sum_e (k_e^{in} - k^{\bar{in}})^2/L. \tag{3}$$

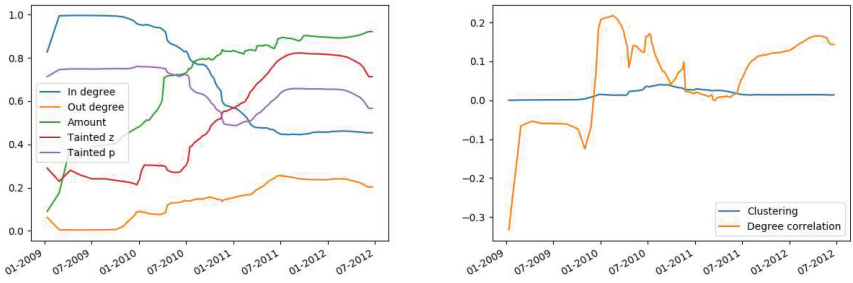
σ_{out} and $k^{\bar{out}}$ are computed in a similar way. Degree correlation measures the nodes' tendency to be linked to nodes with a similar degree. In an assortative network (where $r > 0$) high degree nodes are linked to other high degree nodes and low degree nodes are linked to other low degree nodes. In disassortative networks ($r < 0$), in contrast, high degree nodes have a tendency to connect to low degree nodes, creating a hub and spoke structure.

Finally, we measure the networks' clustering coefficient, that is, the density of triangles in the networks, given by

$$C = \frac{1}{N}\sum_v \frac{2\Delta_v}{d_v(d_v - 1)} \tag{4}$$

where Δ_v is the number of triangles with node v and d_v is the degree of node v. The sum runs over all nodes in the network [2]. To compute C we must ignore the directionality of the network. The clustering coefficient measures how connected then nodes are in their closest neighborhoods.

4 Results



(a) Gini coefficient.

(b) Clustering coefficient and degree correlation.

Fig. 2. Evolution of the network's characteristics.

Figure 2 shows the evolution of some of the network's characteristics as presented by Kondor et al. (2014) [9], namely the Gini-coefficient of in-degree, out-degree and amount in Fig. 2a and the degree correlation and clustering coefficient in Fig. 2b. Since we are looking at transactions only, and not wallets, these graphs are slightly different from the ones presented in [9], although the trends are very similar, except for the clustering coefficient. However, given this close similarity, we continue to work with the network of transactions only. In addition, we have added the Gini coefficient for tainted P ratio and tainted Z ratio in the plot in Fig. 2a. We can see that both start off relatively low, but increase sharply in mid 2010, with the tainted Z inequality increasing much more than the tainted P inequality.

Figure 3 shows the evolution of the networks' diameter, number of nodes and density. Note the log scale on the y-axis. We can see that the sub-networks are both smaller and denser than the full transaction network, which is to be expected, since they are samples of the full network. The sub-networks are smaller because their origin can only be traced to particular subsets of coinbase transactions, and yet as time goes by they mix in with all the other transactions, and hence the measures presented in Fig. 3 converge. The diameter is more fuzzy in the beginning, but eventually, all networks show a similar tendency in this regard.

Figure 4 shows the evolution of the Gini coefficient for in-degree, out-degree, transaction amount, tainted Z and tainted P, in addition to the degree correlation and clustering coefficient for each of the five sub-networks on a monthly basis. In each plot, the red line denotes the whole network, and we can see how the values for each sub-network all converge towards each other and are

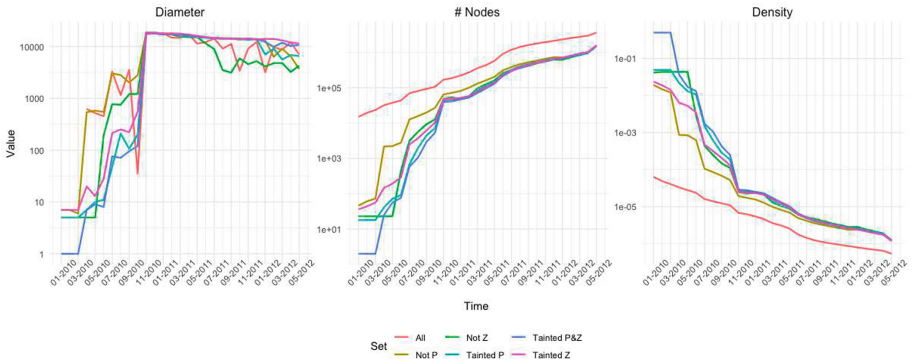


Fig. 3. Evolution of diameter, number of nodes and density in the network of all transactions and in the five sub-networks.

slowly nearing the red line. Moreover, we see that in the beginning, the in-degree tends to be more equally distributed in the sub-networks than in the whole network, whereas for out-degree there is an opposite behavior, the distribution of out-degree is less equal in the sub-networks. We also see that in the tainted P and tainted $P\&Z$ networks, the inequality in the amount distribution increases in early 2010 and remains very high. In terms of the Gini coefficient for tainted Z ratio, the inequality in the tainted P is very high early on, and we see the opposite effect in terms of the Gini coefficient of tainted P, here the tainted Z sub-network scores very high, at least until November 2010. Both sub-networks of not tainted transactions have a high clustering coefficient in the beginning, whereas all converge to the same low value towards the end of the period. The Not P sub-network behaves differently from the other ones. In terms of out-degree, tainted Z and tainted P it dips in April 2010 and jumps at the same time in terms of in-degree and clustering coefficient. Its amount inequality remains high throughout the whole period. For degree correlation, all sub-networks show a similar trend, except for the tainted $P\&Z$ sub-network which takes a downwards turn in September 2010 and stays negative for a couple of months. This particular observation clearly demonstrates an irregularity that needs to be studied further.

The evolution of the various Gini coefficients in the full network in comparison to the sub-networks can tell us a great deal about how the tainted coinbase transactions have blended in with the other transactions, thus hiding in plain sight. It also informs us of points in time where the transaction network ought to be investigated more in-depth. In terms of in-degree, the Gini coefficient is much lower in the sub-networks than in the full network, which indicates a more homogeneous in-degree distribution. The opposite holds for the out-degree, there is more inequality in the out-degree in the tainted networks. This could indicate that owners of tainted bitcoin were behaving differently when trading them, while mixing them with untainted coins. In terms of amount inequality, it is the highest in the tainted sub-networks. It is interesting to see such a high

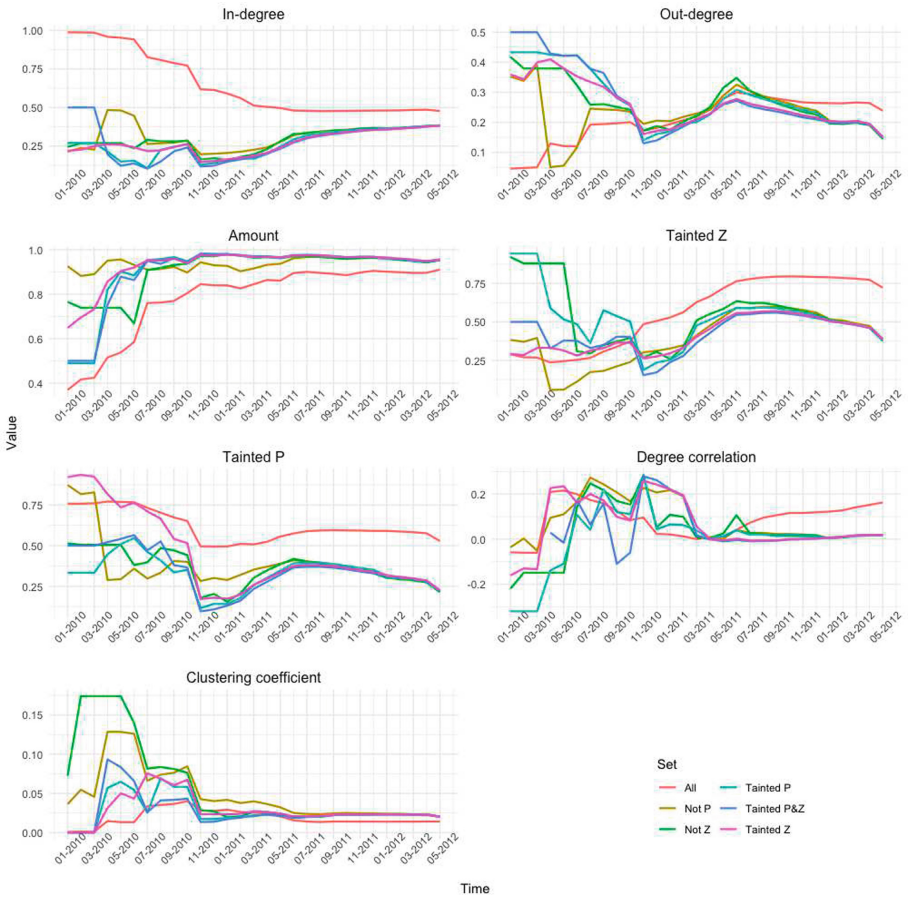


Fig. 4. Evolution of Gini coefficients of in-degree, out-degree, transaction amount, tainted Z ratio and tainted P ratio, as well as degree correlation and clustering coefficient for the whole transaction network and five types of sub-networks.

tainted P inequality in the tainted Z network and a high tainted Z inequality in the tainted P network in the first year. Finally, the networks' assortativity raises many questions, because of the varied patterns in the sub-networks. Furthermore, the fact that the Tainted $P\&Z$ network becomes disassortative for two months is highly irregular. All of these observations require further investigation, for example by looking at the degree distribution of the sub-networks, and a closer inspection of the structure of transactions at various moments.

5 Conclusion

In this paper we used network science to detect and investigate cryptographic anomalies. Based on two types of anomalies, we constructed sub-networks of

bitcoin transactions and compared their structural properties. We saw that the distribution of several node properties, such as in-degree, transaction amount and tainted ratio is different in the sub-networks when compared to the full network. This is apparent in the networks until late 2010, when the properties start to converge to what is observed in the full network. In particular, degree correlation of the sub-network with both anomalies shows a great deviation from the rest at the same time as both these anomalies were prominent in block mining. This paper has an additional contribution. The size of the blockchain and its transactions places a prohibitively high computational complexity on analysing its network behaviour, the technique used here of sampling when creating the sub-networks has allowed us to adequately estimate the networks' properties as Figs. 3 and 4 show. Using this as a basis for similar methods to compress computation time for block chain transaction analysis is worth exploring.

Further work is needed to get a full grasp on what exactly is happening in the networks we examined. Our analysis is based on monthly updates of the network, whereas weekly or daily updates might give a better sense of when and how the anomalies are having an effect on transaction patterns. Moreover, we are looking at a network of transaction only, and not including the wallets. Having wallets as nodes would change the network structure and may well provide other insights. Finally, we have only analysed transactions until mid 2012. In our continued work, our plan is to consider the entire blockchain.

References

1. Alarab, I., Prakoonwit, S., Nacer, M.I.: Comparative analysis using supervised learning methods for anti-money laundering in bitcoin. In: Proceedings of the 2020 5th International Conference on Machine Learning Technologies, pp. 11–17 (2020)
2. Barabási, A., et al.: Network Science. Cambridge University Press, Cambridge (2016)
3. Bartoletti, M., Pes, B., Serusi, S.: Data mining for detecting bitcoin ponzi schemes. In: 2018 Crypto Valley Conference on Blockchain Technology (CVCBT), pp. 75–84. IEEE (2018)
4. Bovet, A., Campajola, C., Mottes, F., Restocchi, V., Vallarano, N., Squartini, T., Tessone, C.J.: The evolving liaisons between the transaction networks of bitcoin and its price dynamics (2019). arXiv preprint: arXiv:1907.03577
5. Courtois, N.T., Grajek, M., Naik, R.: The unreasonable fundamental incertitudes behind bitcoin mining (2013). arXiv preprint: arXiv:1310.7935
6. Dev, J.A.: Bitcoin mining acceleration and performance quantification. In: 2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–6. IEEE (2014)
7. Eyal, I., Sirer, E.G.: Majority is not enough: bitcoin mining is vulnerable. In: Christin, N., Safavi-Naini, R. (eds.) Financial Cryptography and Data Security. FC 2014. Lecture Notes in Computer Science, vol. 8437, pp. 436–454. Springer, Berlin, Heidelberg (2014). https://doi.org/10.1007/978-3-662-45472-5_28
8. Hu, Y., Seneviratne, S., Thilakarathna, K., Fukuda, K., Seneviratne, A.: Characterizing and detecting money laundering activities on the bitcoin network (2019). arXiv preprint: arXiv:1912.12060

9. Kondor, D., Pósfai, M., Csabai, I., Vattay, G.: Do the rich get richer? an empirical analysis of the bitcoin transaction network. PloS one **9**(2) (2014)
10. Li, S.-N., Yang, Z., Tessone, C.J.: Mining blocks in a row: a statistical study of fairness in bitcoin mining. In: 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 1–4. IEEE (2020)
11. Lorenz, J., Silva, M.I., Aparício, D., Ascensão, J.T., Bizarro, P.: Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity (2020). arXiv preprint: arXiv:2005.14635
12. Mallett, J.: A report on cryptographic anomalies in the bitcoin blockchain (2020)
13. McGinn, Dan, McIlwraith, Doug, Guo, Yike: Towards open data blockchain analytics: a bitcoin perspective. R. Soc. Open Sci. **5**(8), 180298 (2018)
14. Monamo, P., Marivate, V., Twala, B.: Unsupervised learning for robust bitcoin fraud detection. In: 2016 Information Security for South Africa (ISSA), pp. 129–134. IEEE (2016)
15. Monamo, P.M., Marivate, V., Twala, B.: A multifaceted approach to bitcoin fraud detection: global and local outliers. In: 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 188–194. IEEE (2016)
16. Nakamoto, S., Bitcoin, A.: A peer-to-peer electronic cash system. Bitcoin (2008). https://bitcoin.org/bitcoin.pdf
17. Pavithran, D., Al-Karaki, J.N., Thomas, R., Shibu, C., Gawanmeh, A.: Data-driven analysis of price change, user behavior and wealth accumulation in bitcoin transactions. In: 2019 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1–6. IEEE (2019)
18. Pham, T., Lee, S.: Anomaly detection in the bitcoin system-a network perspective (2016). arXiv preprint: arXiv:1611.03942
19. Ron, D., Shamir, A.: Quantitative analysis of the full bitcoin transaction graph. In: Sadeghi, A.R. (ed.) Financial Cryptography and Data Security FC 2013. Lecture Notes in Computer Science, vol. 7859. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39884-1_2
20. Turner, A.B., McCombie, S., Uhlmann, A.J.: Discerning payment patterns in bitcoin from ransomware attacks. Journal of Money Laundering Control (2020)
21. Vallarano, N., Tessone, C., Squartini, T.: Bitcoin transaction networks: an overview of recent results (2020). arXiv preprint: arXiv:2005.00114
22. Weber, M., Domeniconi, G., Chen, J., Weidele, D.K.I., Bellei, C., Robinson, T., Leiserson, C.E.: Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics (2019). arXiv preprint: arXiv:1908.02591