Network-based indicators of Bitcoin bubbles

Bovet, Alexandre; Campajola, Carlo; Lazo, Jorge F.; Mottes, Francesco; Pozzana, Iacopo; Restocchi, Valerio; Saggese, Pietro; Vallarano, Nicoló; Squartini, Tiziano; Tessone, Claudio J.

Published in:
ArXiv pre-print

Publication date:
2018

Network-based indicators of Bitcoin bubbles

Alexandre Bovet[1,2], Carlo Campajola[3], Jorge F. Lazo[4], Francesco Mottes[5], Iacopo Pozzana[6], Valerio Restocchi[7], Pietro Saggese[8], Nicoló Vallarano[8], Tiziano Squartini[8*], and Claudio J. Tessone[9]

[1]naXys, University of Namur, Rempart de la Vierge 8, B-5000 Namur, Belgium
[2]ICTEAM, Université catholique de Louvain, Avenue George Lemaître 4, B-1348 Louvain-la-Neuve, Belgium
[3]Scuola Normale Superiore, I-56126 Pisa, Italy
[4]Lund University, Professorsgatan 1, Lund, Sweden
[5]Università di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy
[6]Birkbeck - University of London, Malet Street, WC1E 7HX London, UK
[7]ECS, University of Southampton, SO17 1BJ Southampton, UK
[8]IMT School for Advanced Studies Lucca, I-55100 Lucca, Italy
[9]URPP Social Networks, University of Zurich, CH-8050 Zürich, Switzerland
[*]tiziano.squartini@imtlucca.it

ABSTRACT

The functioning of the cryptocurrency Bitcoin relies on the open availability of the entire history of its transactions. This makes it a particularly interesting socio-economic system to analyse from the point of view of network science. Here we analyse the evolution of the network of Bitcoin transactions between users. We achieve this by using the complete transaction history from December 5th 2011 to December 23rd 2013. This period includes three bubbles experienced by the Bitcoin price. In particular, we focus on the global and local structural properties of the user network and their variation in relation to the different period of price surge and decline. By analysing the temporal variation of the heterogeneity of the connectivity patterns we gain insights on the different mechanisms that take place during bubbles, and find that hubs (i.e., the most connected nodes) had a fundamental role in triggering the burst of the second bubble. Finally, we examine the local topological structures of interactions between users, we discover that the relative frequency of triadic interactions experiences a strong change before, during and after a bubble, and suggest that the importance of the hubs grows during the bubble. These results provide further evidence that the behaviour of the hubs during bubbles significantly increases the systemic risk of the Bitcoin network, and discuss the implications on public policy interventions.

Introduction

Designed under the pseudonymous name of Satoshi Nakamoto, and introduced by a disruptive paper in 2008[1] while the world was challenged by the aftermaths of the financial crisis, Bitcoin is in essence a series of cryptographical protocols that solve the double-spending problem, i.e. prevent the same digital token from being spent more than once, in the absence of a third party that verifies and guarantees the validity of transactions. More in detail, Bitcoin consists of a decentralized peer-to-peer network, composed by users that transact bitcoins among them; once it is validated by a network of miners according to the consensus rules that are part of the protocol, these transactions are included in a public and distributed transactional database, the blockchain ledger[2–4]. Few years after the date of its release, this digital currency has showed to be able to attract an increasing number of users, both because of speculative reasons[5–7], and because of the trust of early adopters in the potentialities of this innovative technology[8,9]. In fact, the number of users and ergo the number of transactions within the bitcoin network has witnessed a remarkable burst which also has lead to an increment on its value in the market and consequently to some price bubbles and respective crashes[10–12]; on the other hand, the novelties introduced by the bitcoin protocol have allowed a numerous number of innovative analyses and make the bitcoin network a particularly interesting case of study[13]. Indeed, a remarkable feature of the transaction verification mechanism on which bitcoin system relies on is that the transaction history since the creation of the currency is openly accessible. The availability of the complete transaction history allows to investigate the structural properties of the network of bitcoin users and to examine their relations with its different growing phases. The structure and dynamics of the network of bitcoin users has only recently started to be investigated. Looking at the network of transactions between addresses, Kondor et al.[14] have shown that the in-degree distribution of nodes, i.e. the number of incoming transactions of nodes, relates with nodes wealth distribution. Parino et al.[15] have investigated the network of the international bitcoin flow to identify socio-economic factors driving its adoption by country.

Bubble	Start	End
1	2012-05-25	2012-08-18
2	2013-01-03	2013-04-11
3	2013-10-07	2013-11-23

Table 1. Time intervals of the three Bitcoin bubbles occurring between May 2012 and January 2014[11].

Here, we reconstruct the network of transaction between users by merging addresses apparently owned by different users making it closer to reality than the raw network of addresses[16,17]. We study the evolution of the network global quantities, such as the variation coefficient of the degree distributions, the sizes of the largest strongly and weakly connected components and we also investigate the evolution of the local structure of the networks by examining so-called *network motifs*[18]. Network motif, defined as statistically recurrent subgraphs, were shown to implement simple functionalities that contribute to the complex behaviour of the system as a whole[18]. This modular organization at the local scale was also shown to be common to a wide range of real networks[19]. In particular, the abundance of certain triadic motifs has been identified as a early-warning signal for topological collapse of inter-banking networks[20].

Results

Evolution of global network measures

To be able to detect patterns occurring at different time scales, we construct two time series of networks, using a integration time of one day and one week respectively. We refer to Wheatley et al. 2018.[11] to identify the start and end date of the various bubbles and focus on the three first bubbles, reported in Table 1. Our observation period starts 5 months before the onset of the first bubble, namely on December 5th 2011, and finishes on December 23rd 2013, one month after the burst of the last bubble taken into account.

We explored the evolution of several network properties that can be relevant to highlight events that might be related to irrational exuberance of agents in the market or to crisis triggering phenomena such as liquidity imbalances. The following measures are defined:

- Number of nodes (wallets) in the network Bubbles are characterised by a rising number of active users, especially at the end of the critical period.
- Size of largest strongly and weakly Connected Components.
- Ratio between the (total, in/out) degree of the most connected and second most connected node, dubbed as “Degree Gap Ratio”.
- Ratio between the in and out degree of the most connected node, dubbed as “Hub in/out Degree Ratio”.

As a first step we analyse the time series of the number of active users alongside the price evolution (Fig 1a), aggregated at the weekly time scale, which shows the strong correlation between the two, especially during bubble periods. This is easily explained by the herding feedback mechanism[10], where an initial price hike is followed by an increased popularity of the asset that lures more users into the market. The increase in activity of typically buying users (since they are getting into the market) further boosts the growth of price which eventually grows at unsustainable rate and collapses, either because of growth slowdown that triggers speculators to sell or after some negative news release that leads to panic sales.

Next we try to quantify how central are the most connected nodes in the circulation of money and whether their centrality in the network of transactions is related or not to the price surge and fall. To do so, we compute the size of the largest Strongly Connected Component (SCC) and Weakly Connected Component (WCC), namely the largest set of vertices that have a path running between them in both directions, for the SCC, and in at least one direction for the WCC. We make this analysis on the full network and on the networks where the most connected (SCCm1, WCCm1) and the two most connected nodes (SCCm2, WCCm2) are removed from the network with all of their edges. This would highlight the role of the largest and second largest hubs as intermediaries in the market. In particular the WCC is relevant to show if there are some fractions of the economy that never interact with others or that do so only through one of the two hubs, while the SCC can be an indicator of how efficiently a bitcoin can pass from one end to the other of the network and how this happens thanks to hubs. We plot these indicators in Fig 1b, 1c. We see how the vast majority of users are always connected to the network since the largest WCC includes almost all nodes. However, the importance of the first hub changes throughout the time period we consider, decreasing in time and in particular during the first bubble. We see that the first bubble eliminates a second important hub (WCCm1), and this could be

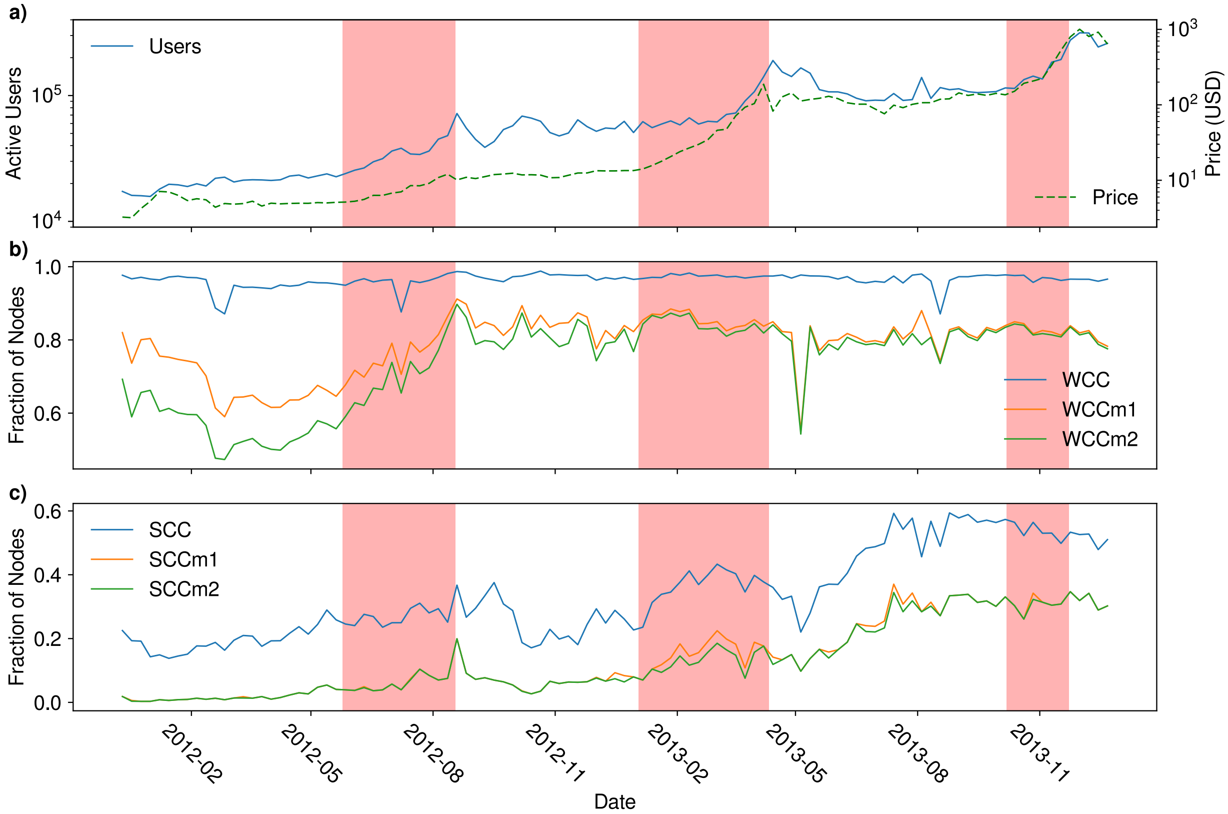

Figure 1. Bitcoin network and price evolution. The measures are performed on networks reconstructed each week. The three bubbles are indicated by a shaded area (from onset until burst day). a) Number of active users, i.e. number of nodes in the networks, as a function of time (blue) and Bitcoin price in USD (green). b) Size of the largest weakly connected component (WCC) of the full network divided by the total number of nodes (blue), largest weakly connected component after the removal of the highest degree node (WCCm1, orange) and the removal of the two nodes with highest degree (WCCm2, green). c) Similar measure than in b) but using the largest strongly connected component instead of the largest weakly connected component.

related to the unmasking of the Ponzi scheme behind Bitcoins Savings & Trust[11] who might have been that player. After that event, the centrality of the first hub stabilises in the WCC and no other node has similar importance in that almost 20% of nodes are connected to the market only through it.

Similar conclusions on the centrality of the first hub come from the gap between the SCC and SCCm1 which is again close to 20% of the total nodes in the network. The circulation of money in the network rises as time goes on, and after the first bubble the first hub, although very central, is not any more the only passage point. This is a signal that smaller nodes start trading among themselves without relying only on the intermediation of the large exchange (likely Mt. Gox) and the velocity of money through the network is not completely hindered if the first hub goes in distress thanks to these secondary channels. It is interesting to notice that this doesn't happen at the expense of the importance of the first hub, which is indeed increasing its centrality, and neither through the formation of a secondary hub, since the second largest node is irrelevant to the size of the SCC and WCC.

However, we do not find consistent signals fore-running bubble onsets or bursts in these global measures. This could be related to the developing nature of the market, which is in constant evolution and for which these global features undergo such big changes that even comparing between time frames is extremely hard. We thus turn to more microscopic measures, and expose the results in the next section.

Heterogeneity measures

In this section we examine several measures of heterogeneity for three distributions, namely the in-degree, out-degree and total degree. Specifically, we use a daily aggregation windows for network and cover two years of transactions, from January 2012 to December 31st 2013. For each type of network and distribution, we compute the variation coefficient, i.e., $V = \frac{\sigma}{\bar{x}}$, where σ and

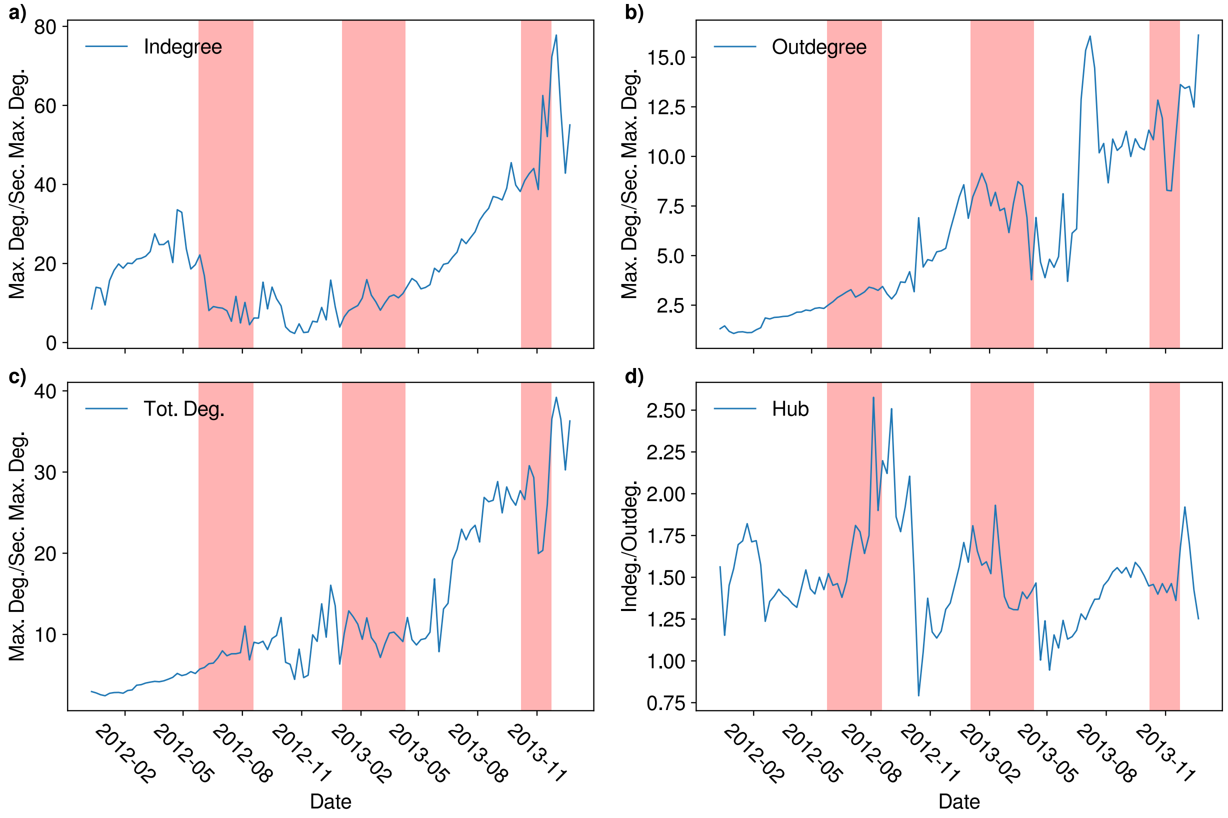

Figure 2. Degree gap ratios of the two largest hubs. The three bubbles are indicated by shaded area (from onset until burst day). a) Ratio of the in-degree of the two nodes with highest in-degree. b) Ratio of the out-degree of the two nodes with highest out-degree. c) Ratio of the total degree of the two nodes with highest total degree. d) Ratio of in-degree over out-degree for the largest hub.

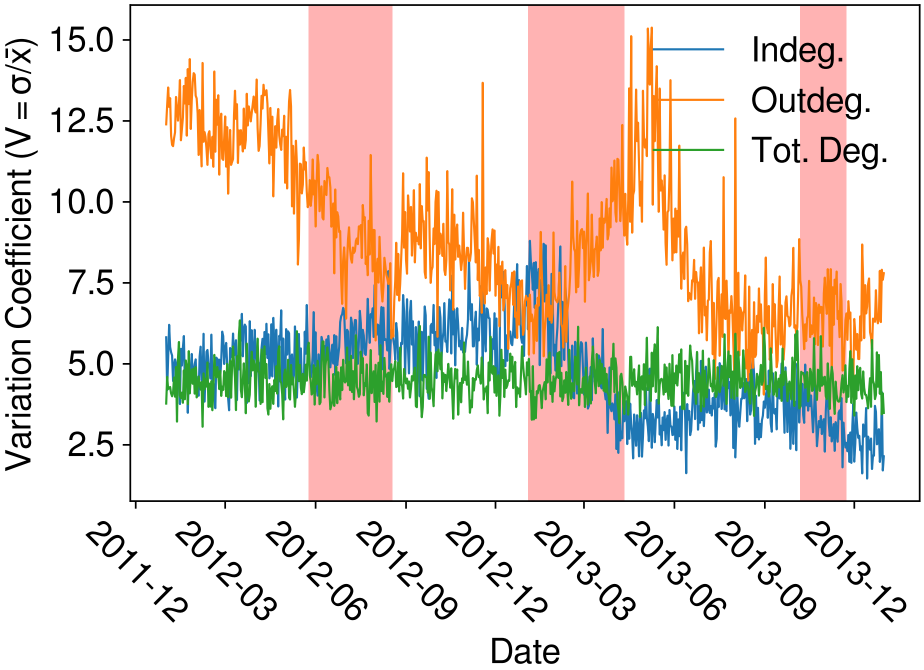

Figure 3. Heterogeneity evolution of the degree distributions. Variation coefficient of the in-degree (blue), out-degree (orange) and total degree (green) distributions of daily networks. The bubbles are indicated by the pink shaded areas (from onset until burst day).

$\bar{x}$ are the standard deviation and the average of the empirical distribution, respectively. As all the distributions display heavy tails, this measure of heterogeneity may depends on the size of the networks. To account for the varying size of the networks at every point in time, we build the distributions only considering $N_{sample} = 1784$ nodes for each daily network, where N_{sample} is the number of nodes the smallest network possesses, and measure the average value of V over 100 random sample for each distribution. Our results, displayed in Figure 3, show that during the second bubble, which occurred between January 3rd 2013 and April 11th 2013, the in-degree distribution becomes far more homogeneous than before the bubble, whereas the out-degree distribution exhibit a surge of heterogeneity.

Fig. 2 shows the relative degrees of the largest and second largest hub, as well as the out-degree-in-degree ratio of the largest hub. Specifically, the latter measure suggests that the largest hub lowers the number of people it buys from with respect to the number of people it sells to, which is expected during a bubble, since more and more low-degree nodes enter the market following the price surge (see Fig. 1). Interestingly, figures 2 (a) and 2 (b) show that the the in-degree (out-degree) of the largest hub grows (decreases) with respect to that of the second largest hub, suggesting that the second largest hub follows a similar dynamics to that of the largest hub, but to a greater extent. Indeed, these results suggest that, during the second bubble, the second largest hub increases the number of customers it sells to, whereas it lowers the number of customers it buys from. It is worth noting that this is purely a structural change, since the largest hub keeps a null trade balance throughout this period.

These structural changes are consistent with the changes in the heterogeneity of the in-degree and out-degree distributions, and suggest that there are two hubs that centralise the market by selling bitcoins to most of the traders that enter the market during the bubble, resulting in a significant increase of the systemic risk. Indeed, if only a few hubs account for most of the transactions in the network, if at any point in time one of them fails, the whole network may crash. This is exactly what happened on April 10th 2013, when Mt Gox, the major Bitcoin exchange, broke under the high trading volume, triggering the burst of the bubble.

Triadic motifs analysis

Triadic motifs, i.e. all the possible directed patterns connecting three vertices, are the natural generalizations of directed clustering coefficients and the starting point for the understanding of a complex network self-organization in communities. Thirteen, non-isomorphic, triadic directed patterns (reported in Fig. 6) can be identified and classified. Given a real, binary, directed matrix $\mathbf{A}$, the motifs occurrences N_m can be written as reported in table 2.

Motif m	N_m
1	$\sum_{i\neq j\neq k}(1-a_{ij})a_{ji}a_{jk}(1-a_{kj})(1-a_{ik})(1-a_{ki})$
2	$\sum_{i\neq j\neq k}a_{ij}(1-a_{ji})a_{jk}(1-a_{kj})(1-a_{ik})(1-a_{ki})$
3	$\sum_{i\neq j\neq k}a_{ij}a_{ji}a_{jk}(1-a_{kj})(1-a_{ik})(1-a_{ki})$
4	$\sum_{i\neq j\neq k}(1-a_{ij})(1-a_{ji})a_{jk}(1-a_{kj})a_{ik}(1-a_{ki})$
5	$\sum_{i\neq j\neq k}(1-a_{ij})a_{ji}a_{jk}(1-a_{kj})a_{ik}(1-a_{ki})$
6	$\sum_{i\neq j\neq k}a_{ij}a_{ji}a_{jk}(1-a_{kj})a_{ik}(1-a_{ki})$
7	$\sum_{i\neq j\neq k}a_{ij}a_{ji}(1-a_{jk})a_{kj}(1-a_{ik})(1-a_{ki})$
8	$\sum_{i\neq j\neq k}a_{ij}a_{ji}a_{jk}a_{kj}(1-a_{ik})(1-a_{ki})$
9	$\sum_{i\neq j\neq k}(1-a_{ij})a_{ji}(1-a_{jk})a_{kj}a_{ik}(1-a_{ki})$
10	$\sum_{i\neq j\neq k}(1-a_{ij})a_{ji}a_{jk}a_{kj}a_{ik}(1-a_{ki})$
11	$\sum_{i\neq j\neq k}a_{ij}(1-a_{ji})a_{jk}a_{kj}a_{ik}(1-a_{ki})$
12	$\sum_{i\neq j\neq k}a_{ij}a_{ji}a_{jk}a_{kj}a_{ik}(1-a_{ki})$
13	$\sum_{i\neq j\neq k}a_{ij}a_{ji}a_{jk}a_{kj}a_{ik}a_{ki}$

Table 2. Classification and definitions of the triadic motifs abundances.

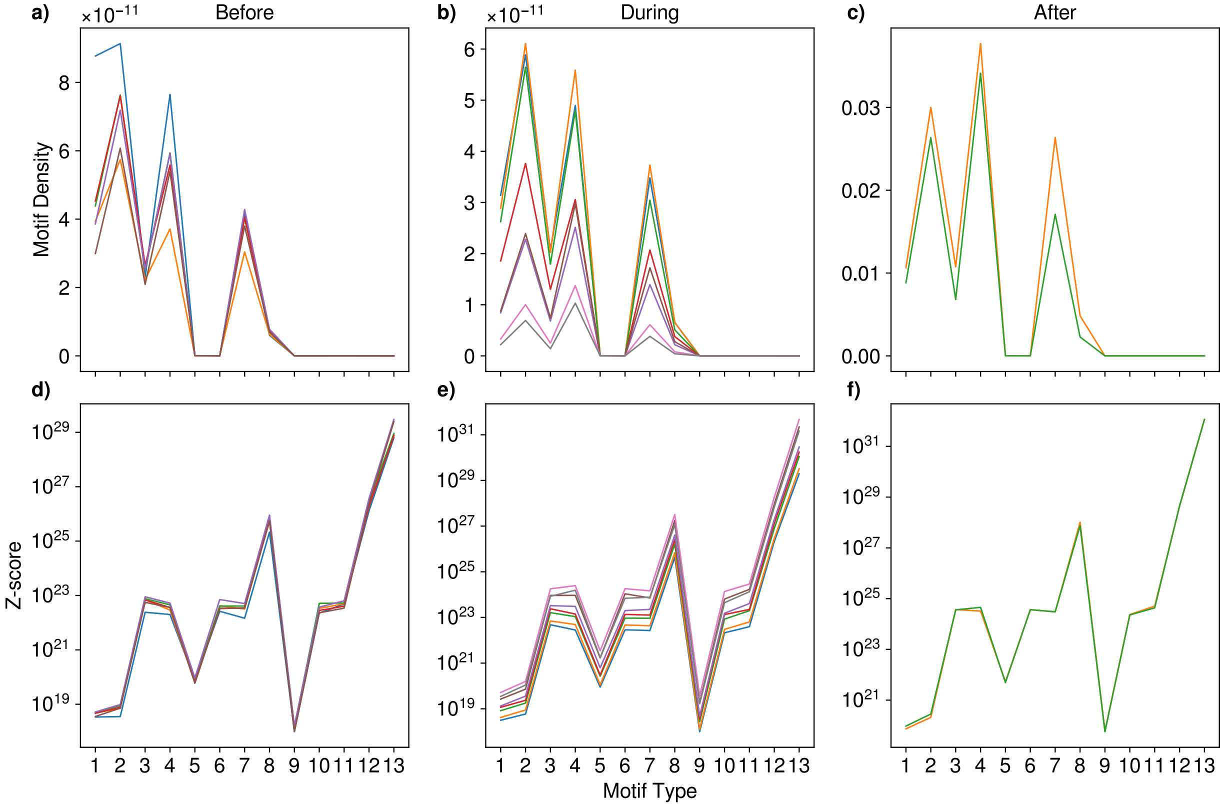

Figure 4. **Frequency evolution of the 13 triadic motifs**.

The three upper panels of Figure 4 show the density of the 13 triadic motifs during three periods that have been chosen to monitor the system *before*, *during* and *after* the bubble happening in the period of time from December 5th 2011 until September 18th 2012, taken on weekly aggregated networks. The three lower panels of Figure 4 show the z-scores of the same motifs.

Upon visually inspecting the three upper panels, it is apparent that the bubble is indeed characterised by motifs profiles that differ from both the previous and the following period. It is apparent that the same motifs (i.e. 1, 3, 6) are overrepresented during the three periods: upon inspecting these three kinds of motifs, we see that they are constituted by a basic unit of two non-reciprocated dyads. We suspect these to characterise the topological structure of the hubs, reflecting a huge selling activity in all periods. This is particularly evident, however, when considering the pre-bubble period from December 5th 2011 until May 25th 2012: this evidence leads us to assume that monitoring the hubs connectivity may be useful to detect upcoming critical activity.

These structural changes are consistent with the changes in the heterogeneity of the in-degree and out-degree distributions, and suggest that there are two hubs that centralise the market by selling bitcoins to most of the traders that enter the market during the bubble, resulting in a significant increase of the systemic risk. Indeed, if only a few hubs account for most of the transactions in the network, if at any point in time one of them fails, the whole network may crash.

Conclusions

In this paper we analyse the impact of structural properties of the Bitcoin transaction network on the generation and crash of bubbles in the exchange with respect to fiat currencies. Specifically, we examine network features such as heterogeneity of the degree distributions and frequency of connectivity patterns (i.e., motifs). We find significant changes in these properties during the period of price bubbles. A more detailed analysis unveils that, during the first bubble, the frequency of motifs indicating the relationship hubs have with new, low-degree users changes significantly; this suggests that hubs take an even more important role in becoming liquidity providers. These results are confirmed in the second bubble: There, by analysing the heterogeneity of the in-degree, out-degree, and total degree distributions, we find that there is a significant widening (narrowing) of the out-degree (in-degree) distributions, whereas the total degree does not change its distribution significantly. By performing additional analyses on the two largest hubs, we find that these structural changes - similar to what is observed during the first bubble - is likely to be caused by the *centralising* role hubs take on as liquidity providers.

Although we find that measures can explain well some price bubbles but not others, these results highlight that tracking properties of hubs in the transaction network is key for understanding the underlying mechanisms of a bubble. Moreover, at least in the first three Bitcoin bubbles, the behaviour of hubs significantly increased the systemic risk of the Bitcoin economy, eventually leading to systemic failure and sudden price crashes.

These results also suggest that Bitcoin bubbles are difficult to forecast, but can be prevented, or at least alleviated, by introducing policies that aim at reducing the importance of large hubs in the network. In future work, we plan to extend our analysis by introducing new structural measures and by covering all the bubbles that happened to date.

Methods

User network reconstruction

An element of the Bitcoin protocol is that it attempts to preserve anonymity of users in a way that is better defined as *pseudonimity*: transactions take place without the need of a third party and users cannot be directly linked to real users or to an identity[3]. A transaction, thus, does not identify the payer or the payee in any way. However, by exploiting the properties of the protocol, like the fact that the transaction history is publicly available, it is possible to trace and cluster addresses that are owned by the same user, collapsing in that way a network of addresses into a network of users. The principle that drove our approach is to minimise as much as possible the number of false positives, that is, the addresses that are linked together as if they were owned by the same user but they are not. The approach is based on two heuristics introduced by Meiklejohn *et al.*[17], that we describe here.

Input-based Heuristic: The first and safest one exploits the fact that, in principle, if two addresses are input of the same transaction, then they are controlled by the same user. This property is also transitive, which means that if a transaction includes in the input the addresses A and B, and a second transaction includes the addresses B and C, then it is safe to assume that A,B,C addresses belong to the same user.

One-time change addresses: the second one detects addresses appearing in the output of a transaction and that can be attributed to the owner of the inputs.

In our work we adopt the first heuristic, but we modify the second one, as in the working paper by Tessone[16], with the aim of using the second heuristic only in the case in which it is safe to assume that the one-time change address belongs to the owner of the inputs.

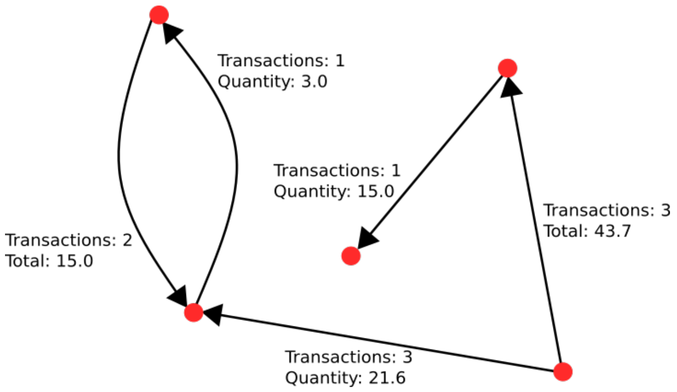

Figure 5. An illustrative example of the construction of the network. Nodes, corresponding to users, are connected by directed edges, corresponding to transactions. Edges are annotated with the total amount of transactions occurred and the total quantity of bitcoins transferred during the observation window.

Others techniques use different approaches[21–24], where the one-time change addresses are defined using similar heuristics. Another approach, proposed by Cazabet et al.[25], is based on a combination of the input-based heuristic and of the Louvain community detection algorithm to detect addresses that are likely to belong to the same user. The reason why we did not use these methods is that we privileged the heuristics preserving the analysis from errors in detecting addresses not belonging to the same users as if they were controlled by a single entity, i.e. avoiding false positives.

Note that these techniques do not allow to reproduce the perfect network of users, since real users can use different wallets that not necessarily are linked together by a transaction; thus, the reader should not consider the network obtained as a perfect representation of the real network of users, but as an approximation that clusters addresses minimizing the presence of false positives.

Once the addresses are grouped by wallet, we build the network in the following way: two nodes (wallets) i and j are connected via a directed edge (i,j) if at least one transaction from i to j occurs during the considered time integration window. Edges are annotated with the number of transactions occurred and the total quantity of bitcoins transferred. The structure of the network is illustrated with an example in Figure 5.

Null models

In order to verify the statistical significance of our results we have compared them with a properly-defined null model. Inspired by the empirical regularities of the degree distribution we have employed the Directed Random Grapg Model (DRGM). It is an Exponential Random Graph Model (ERGM) defined within the constrained Shannon entropy-maximization framework. Briefly speaking, one solves the following problem

$$\max_{\mathbf{P}}\left\{S[P]-\sum_i\theta_i\left[\sum_{\mathbf{A}}P(\mathbf{A})X(\mathbf{A})-\bar{X}_i\right]\right\} \tag{1}$$

where the vector of constraints reads $\vec{C}(\mathbf{A})=\{\vec{k}^{out},k^{in}\}$ and $C_0=\langle C_0\rangle=1$ sums up the normalization condition of the searched probability distribution. The solution to the problem above reads

$$P(A)=\frac{e^{-H(\mathbf{A},\vec{\theta})}}{Z(\vec{\theta})} \tag{2}$$

with $H(\mathbf{A},\vec{\theta})=\vec{\theta}\cdot\vec{C}$ summing up the proper topological constraints. In the DRGM case, our Hamiltonian reads with $H(\mathbf{A},\theta)\equiv\theta L$ a position leading to the probability function

$$P(\mathbf{A})=p^L(1-p)^{N(N-1)-L} \tag{3}$$

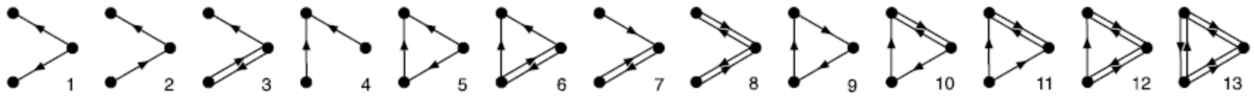

Figure 6. The 13 possible triadic motifs involving three connected vertices[20].

with $p\frac{e^{-\theta}}{1+e^{-\theta}} \equiv \frac{x}{1+x}$. The comparison between observed and expected properties on the ensemble has been carried out by employing the z-score index, defined as

$$z_X = \frac{N_X - \bar{X}}{\sigma_X}; \tag{4}$$

in our case, X represents the abundance of the triadic motifs shown in fig. 6. The *z-score* is a standardised variable measuring the difference between the observed and the expected value in units of standard deviation. If X is normally distributed under the null model, then values within $z = \pm 1$, $z = \pm 2$, $z = \pm 3$ would (approximately) occur with a 68%, 95%, 99% probability respectively. If the observed value of X corresponds to a large positive (negative) value of z_X then the quantity X is over(under)-represented in the data, and not explained by the null model.

A simpler analysis discounting the increasing volume of the network is obtained by considering the index $\hat{N}_m = \frac{N_m}{N}$, i.e. by dividing the abundance of a given motif m by the total number of nodes in a particular snapshot.

Motifs detection

An exact counting of the network motifs present in the reconstructed networks of transactions was performed, on a reduced set of 17 time points in the period of the first bitcoin bubble. The 13 three-node network motifs analysed are the same ones as those described by Squartini et al.[20] and are represented in Fig. 6. Null models are built for each time point with the procedure described above. The expected number of network motifs $\bar{X}$ and their standard deviation σ_X are henceforth obtained, allowing us to calculate the *z-score* (equation 4).

Acknowledgements

All authors are grateful to Alberto Antonioni, Eugenio Valdano and IMT Lucca for the organisation of the workshop Complexity 72H during which this research was performed. C.J.T. acknowledges financial support of the University of Zurich through the University Research Priority Programme on Social Networks. V.R. was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorised to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

1. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system (2008).

2. Halaburda, H. & Sarvary, M. Beyond bitcoin. *The Econ. Digit. Currencies* (2016).

3. Antonopoulos, A. M. *Mastering Bitcoin: Programming the Open Blockchain* (O'Reilly Media, Inc., 2017).

4. Glaser, F. Pervasive decentralisation of digital infrastructures: a framework for blockchain enabled system and use case analysis. *Hawaii Int. Conf. on Syst. Sci. 2017 (HICSS-50)* (2017).

5. Gandal, N., Hamrick, J., Moore, T. & Oberman, T. Price manipulation in the bitcoin ecosystem. *J. Monet. Econ.* (2018).

6. Chu, J., Nadarajah, S. & Chan, S. Statistical analysis of the exchange rate of bitcoin. *PLoS ONE* **10**, e0133678 (2015).

7. Hayes, A. S. Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics Informatics* **34**, 1308–1321 (2017).

8. Böhme, R., Christin, N., Edelman, B. & Moore, T. Bitcoin: Economics, technology, and governance. *J. Econ. Perspectives* **29**, 213–38 (2015).

9. Gervais, A., Karame, G., Capkun, S. & Capkun, V. Is bitcoin a decentralized currency? *IEEE security & privacy* **12**, 54–60 (2014).

10. Garcia, D., Tessone, C., Mavrodiev, P. & Perony, N. The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *J. R. Soc. Interface* **11:20140623.** (2014).

11. Wheatley, S., Sornette, D., Huber, T., Reppen, M. & Gantner, R. N. Are bitcoin bubbles predictable? combining a generalized metcalfe's law and the lppls model. *Swiss Finance Inst. Res. Pap. No. 18-22.* (2018).

12. Gerlach, J.-C., Demos, G. & Sornette, D. Dissection of bitcoin's multiscale bubble history from january 2012 to february 2018. *arXiv preprint arXiv:1804.06261* (2018).

13. Foley, S., Karlsen, J. & Putniņš, T. J. Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? (2018).

14. Kondor, D., Csabai, I. & Vattay, G. Do the rich get richer? an empirical analysis of the bitcoin transaction network. *PLoS ONE* **9(2):**, e86197.doi:10.1371/journal.pone.0086197 (2014).

15. Parino, F., Gauvin, L. & Beiro, M. G. Analysis of the bitcoin blockchain: Socio-economic factors behind the adoption. *arXiv preprint arXiv:1804.07657* (2018).

16. Tessone, C. & D., G. Bitcoins: the transition towards centralisation of a decentralised economy. *In preparation* (2018).

17. Meiklejohn, S. *et al.* A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference*, 127–140 (ACM, 2013).

18. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).

19. Milo, R. *et al.* Superfamilies of evolved and designed networks. *Science* **303**, 1538–1542 (2004).

20. Squartini, T., van Lelyveld, I. & Garlaschelli, D. Early-warning signals of topological collapse in interbank networks. *Sci. Rep.* **3,3357**, DOI:10.1038 (2013).

21. Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T. & Capkun, S. Evaluating user privacy in bitcoin. In *International Conference on Financial Cryptography and Data Security*, 34–51 (Springer, 2013).

22. Tasca, P., Hayes, A. & Liu, S. The evolution of the bitcoin economy: extracting and analyzing the network of payment relationships. *The J. Risk Finance* **19**, 94–126 (2018).

23. Harrigan, M. & Fretter, C. The unreasonable effectiveness of address clustering. In *Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016 Intl IEEE Conferences*, 368–373 (IEEE, 2016).

24. Ron, D. & Shamir, A. Quantitative analysis of the full bitcoin transaction graph. In *International Conference on Financial Cryptography and Data Security*, 6–24 (Springer, 2013).

25. Remy, C., Rym, B. & Matthieu, L. Tracking bitcoin users activity using community detection on a network of weak signals. In *International Workshop on Complex Networks and their Applications*, 166–177 (Springer, 2017).