Network and Conversation Analyses of Bitcoin

Brais Alvarez-Pereira\textsuperscript{1}, Matthew Ayres\textsuperscript{2}, Ana María Gómez Lopez\textsuperscript{3}, Shai Gorsky\textsuperscript{4}, Sean Hayes\textsuperscript{5}, Zhi Qiao\textsuperscript{6}, and Jessica Santana\textsuperscript{7}

\textsuperscript{1}European University Institute, \textsuperscript{2}Growth and Innovation Asia Pacific, \textsuperscript{3}Yale University, \textsuperscript{4}University of Utah, \textsuperscript{5}University of California at Riverside, \textsuperscript{6}National University of Singapore, \textsuperscript{7}Stanford University

Abstract. The Bitcoin marketplace provides a unique opportunity for information and social scientists to explore familiar patterns in new light. Trade manias, also often referred to controversially as economic bubbles, have been widely discussed in political-economy. In this paper, we identify moments of transition from sharp increases to sharp drops in the price of Bitcoin and apply network and conversation analyses around them. We isolate our analysis to the four largest peaks in the history of Bitcoin. Our findings illustrate how computationally intensive techniques may uncover signals of emergence of such phenomena in complex social systems.

Keywords: bitcoin, networks, sentiment, complexity

* Authors are listed alphabetically.
1 Introduction

Money is certainly one of the main institutions of human societies for several thousand years. The creation and distribution of money has been controlled by governments and banks for the past several centuries. In recent years, the expansion of mobile technologies, lower absolute computing costs, increase in remote data storage capacities, and the ubiquitous nature of PC systems have enabled novel, disruptive payment systems to challenge traditional banking and its controls.

Bitcoin, a notable and widely accepted such payment system, uses the distributed nature of the internet to conduct payments. Bitcoin uses no central controls – a stark contrast to traditional banking. This very fact raises a series of questions as to the nature of money, centralized vs. decentralized coordination, and systems stability and integrity. Between 2010 and 2014, Bitcoin continues to grow in transaction volume, value and number of exchanges.

1.1 Bitcoin Transactions

Setting up Bitcoin for personal use requires the installation of a free, open source Bitcoin wallet on a computer or mobile phone. This creates a unique personal address for transactions, and a public and private key for encryption. Each address accounts for a Bitcoin balance based on the number of Bitcoins purchased; they are held in the wallet. Bitcoins can be purchased from a wide range of Bitcoin Exchanges (such as itBit in $USD, Justcoin in €EUR). As all Bitcoin balances are public, using a new address for every transaction is considered a common practice to increase privacy and security.

Transactions require a third party accepting the currency and having a Bitcoin address. Payments are directed to the target address via the wallet and all transactions require detailed verification. The verification mechanism is the heart of the Bitcoin system: whereas banks and governments – centralized financial institutions – approve and back up every transaction in standard currencies, the approval algorithm in Bitcoin relies on a joint decentralized computational effort by its community members. Community members are encouraged to participate in the verification effort by competing for a reward in Bitcoins for every transaction approved. This verification process is known as Mining†. Miners track and compile transactions every 10 minutes into blocks of data and then convert them into a mathematical cryptographical riddle called a hash, reducing the ability to replicate data. Once a transaction has been verified it is appended to a public ledger called a Blockchain.

In the early days of the Bitcoin network, Mining was performed by individual computers. As the exchange rate of the Bitcoin currency grew, concerted efforts have been made to use increasing centralized computing power to gain more Bitcoins by competition – or by monopolizing the verification process and fabricating transactions.

Bitcoin acceptance is expanding both from the number of transactions and the number of exchanges. In 2013 alone, the total value of Bitcoins surpassed the total value of money assets from thirty-three national economies‡. As of July 2014, there were 13.1 million Bitcoins issued for $USD 7.8bn Market Cap with 31 currencies. The trading value of a Bitcoin was over $USD600.

2. Bitcoin Price

Bitcoin price or exchange rate§ is established in the same manner as a stock market, through an auction mechanism. Buyers enter bids and sellers enter offers simultaneously. When a bid and an offer overlap, Bitcoins are exchanged. The last price at which a bid and offer have agreed becomes the exchange rate. According to its founders, Bitcoin was established as alternative money (or currency) (Nakamoto 2008), a possibility that resonated heavily among some of the early adopters of Bitcoin technology (Grinberg 2011).

In the history of political-economy there were and still are vigorous debates on the nature of money: as an instrument of credit, as means of exchange, or anything in between. Moreover, the traditional view of money as

---

† For a thorough analysis on the economics of Bitcoin mining and the mechanisms of community participation, refer to Kroll et al. 2013.
‡ This comparison to foreign money stocks was obtained from the CIA World Factbook, as cited in Castronova 2014: 239, note 1.
§ Almost anything that can be purchased in Bitcoin is priced in a national currency (generally $USD) and then converted into Bitcoins at the current exchange rate, which implies that its exchange rate is in fact the same as its price.
means of exchange argues that money has to store value. Bitcoins can and have been exchanged for goods and services, and thus can easily be considered as means of exchange. Proponents of the idea that money has to store value may argue that the fact that transactions in Bitcoin are not backed by a well-established institution may detach Bitcoin from the notion of fundamental value. Hence, they would add, Bitcoin cannot be considered money in the traditional sense. From either perspective, the only “value” that can be ascribed to Bitcoin is its market price, as determined by its current exchange rate. In other words, bitcoins are worth whatever people deem them to be worth: what determines Bitcoin’s price is what people are willing to pay for it. Husler et al. (2013), for example, start from the idea that other currencies store value, arguing that the fact that Bitcoin’s price is completely determined by people’s beliefs is one of the reasons behind its particularly high volatility. However, if the idea of money as a store of value is abandoned, Bitcoin’s history can be treated not as an isolated study-case of “valueless” quasi-money, but as a data-rich example of trade in financial assets in general. The current version of the paper will not pursue the theoretical implications of either of these directions: we do not attempt to define any value that Bitcoin may store, and instead focus our attention on the preliminary empirical evidence that we have collected and analyzed.

The term trade mania refers to periods of trade for a financial asset in which the asset’s price reaches unprecedented heights and then collapses. The term economic bubble expands the notion of trade mania and includes the idea that an asset’s price is considerably higher than the value it actually stores for a substantial period of time. As we do not try to define and measure any intrinsic value that Bitcoin may store, our analysis then refers to the underpinnings of the recorded trade manias in Bitcoin’s history. Our paper seeks to understand how Bitcoin’s price is affected by people’s beliefs, trust, and expectations, as reflected through network and conversation analyses.

Nominated either as trade manias or as economic bubbles, extreme price fluctuations in commodity markets have been well observed and documented from the 17th century onward. Probably the earliest known example is the tulip mania in Holland, which lasted from 1634 to February 1637. Other famous instances are the South Sea Company bubble in 1720, the Mississippi bubble between 1719 and 1721, and certainly the 1929 crash, where both stock and real estate markets collapsed. More recent instances include Japan’s real estate bubble in the 1980s, the dot-com bubble from 1995 to 2000, or the bubbles in warrants issued by Chinese companies between 2005 and 2008 studied by Yu and Xiong (2011), as well as bubbles in other emerging markets as studied by Reinhart and Rogoff (2009). These crashes in trade involve very different assets from Bitcoin–material goods (such as tulips or houses), companies’ stock, or warrants. However, all of these examples share a pattern which defines them as variations of the same phenomenon: the period of price increase displays very high growth rates that are fed by expectations of future price or rising returns, and the subsequent period of price decay is always significantly shorter than the period of the build-up. Bitcoin’s price (exchange rate) has experienced several periods of fast increase followed by a sharp decline, which coincide with the aforementioned pattern.

Traditional economics has largely theorized these observed phenomena as bubbles (benchmark against some intrinsic value). The first models of classic “rational bubbles” are based on fully rational and perfectly informed agents, and impose a very rigid structure on the problem which imply - for obvious reasons - that these would not be particularly insightful to understand the behavior of Bitcoin’s price. The new generation of rational models introduces incentives, market frictions and non-standard preferences. These innovations bring into play interesting phenomena such as investment herding, limited liability or perverse incentives of key market players, which help to sustain and propagate bubbles. Still, the assumption of agents’ perfect rationality seems too extreme if we want to explain the volatile behavior of Bitcoin prices. More recent and richer behavioral models depart from the assumption of perfect rationality. These models introduce variations such as optimistic and pessimistic agents, feedback trading – when a group of traders base their expectations on past price movements, biased self-attribution, or representativeness heuristic and conservatism bias. Many of the ideas from these streams of the literature are based on phenomena well documented in psychology and have

** For a review of bubbles in history see Kindleberger (2000). For a review of asset price bubbles in general see Scherbina (2013).

††† These models study if rational bubbles can exist or not under certain conditions, such as the existence or absence of common knowledge of other agents’ information. See for example Allen et al. (1993) or Conlon (2004).

‡‡‡ See for example DeMarzo et al. (2008), Allen and Gorton (1993), or Scherbina (2008).


*** See for example Shiller (2002)


been tested in empirical studies. Because these underscore the crucial role played by information, several of these may be suited for the study of Bitcoin bubbles.

In a recent paper, close in spirit to our research, Garcia et al (2014) analyze the role of social interactions in the creation of Bitcoin price bubbles. The authors use vector auto-regression and four socio-economic indicators about Bitcoin: price on on-line exchanges, volume of word-of-mouth communication in social media, volume of information search and user base growth to identify two positive feedback loops that lead to price bubbles in the absence of exogenous information shocks. Their study also identifies a relationship between external events - through spikes in information search - and drastic price declines.

3. Network Analysis

The recent rise of the Internet and digital social domains has led to a surge of interest and research in formal network science. Indeed, Bitcoin seems very susceptible to formal network analysis, as each individual transaction since its inception has been recorded – and all are publicly available (alongside their time and volume). Examples of recent longitudinal Bitcoin network analysis using this public data include evaluations of shared Bitcoin wallet authorities (Meiklejohn et al. 2013) and the Bitcoin transaction graph (Ron and Shamir 2012, Ober et al. 2013). Similarly to the approach taken by Kondor, Pósfai, Csabai and Vattay (2014), we have defined the vertices of the network model to be individual Bitcoin wallets, and the directed edges to be transactions in Bitcoin (pointing from dispenser to receiver of Bitcoins). Kondor, Pósfai, Csabai and Vattay have focused on two long-term periods of trade in Bitcoin which they have termed the “initial phase” and the “trading phase”. The initial phase, from the inception of Bitcoin until the fall of 2010, is associated in their work with low activity. In the following trading phase, Bitcoin has started to function as a “real currency”. The authors use various network measures to characterize the global features of the Bitcoin network along the abovementioned periods.

Our approach was different. We have focused on the periods of the most notable peaks in Bitcoin’s exchange rate to US dollars. The periods we have identified around these peaks are June 2011, August 2012, April 2013 and November-December 2013. As of today, we have completed the preliminary analysis of the first three peaks.

![Fig. 1. Bitcoin to $USD exchange rate and notable peaks](image)

Around each peak we chose a period consisting of 20 blocks from the blockchain, spanning around approximately 5 days. In addition, as the peaks may signify “regime shifts” in the patterns of Bitcoin trading, we have selected as control periods three periods of the same length of some 20 blocks that were before the significant increase in the exchange rate has started and in which the volatility of the exchange rate was minimal. For each block we have calculated several network measures.
These data have been analyzed in three levels:

1. We have plotted each of the measures as a function of time and searched for points of rapid change in them.

2. We have divided each peak period into four sub-periods: price increase until the absolute peak; first exchange rate drop and temporary short lived recovery of the exchange rate; main exchange rate drop; and beginning of renewed sustained increase in the exchange rate. Notably, all the three peaks we analyzed demonstrated the same abovementioned pattern. We then searched to see if and which of the network measures significantly change between the sub-periods. Figure 2 exemplifies this division into sub-periods.

![Fig. 2. Bitcoin to $USD exchange rate, 13-20 August 2012, and division into sub-periods](image)

3. We have compared the network measures around each peak and the respective control period to check for statistically significant changes in them.

In addition to the quantitative analysis, we have created visualizations of the network structure for each peak. Figures 3, 4 and 5 show these.
In order to describe and look into the dynamics of the Bitcoin transaction network, we will study it as a complex system and evaluate the overall performance throughout all the peak periods. As is well known in the ecosystem,
MacArthur (1955) applied Shannon’s information measure to the flows in an ecosystem network as an important tool to define the system diversity and stability which is:

\[ H = -k \sum_{i,j} \left( \frac{T_{ij}}{T_i} \right) \log \left( \frac{T_{ij}}{T_i} \right) \]

Where \( H \) is the diversity of flows in the network, \( k \) is a scalar constant, and \( T_i \) signifies the sum of \( T_{ij} \) over all combinations of \( i \) and \( j \).

Later in Ulanowicz (2004) extended this definition and decomposed \( H \) into two parts:

\[ H = AMI + H_c \]

Where

\[ AMI = k \sum_{i,j} \left( \frac{T_{ij}}{T_i} \right) \log \left( \frac{T_{ij}T_i}{T_iT_{ij}} \right) \]

And

\[ H_c = -k \sum_{i,j} \left( \frac{T_{ij}}{T_i} \right) \log \left( \frac{T_{ij}^2}{T_iT_{ij}} \right) \]

AMI is called the average mutual information inherent in the flow structure, and \( H_c \) is the residual diversity/freedom. In other words, the overall complexity of the flow structure, as measured by MacArthur’s index, can be resolved into a component that gauges how orderly and coherently the flows are connected and a residual that measures the disorder and freedom that remains.

Similarly, in the Bitcoin network the flow is the transaction between each node and the above three quantity can well tell us the complexity, order and freedom in the whole network, from which we can find out how the diversity or freedom is changing through the whole network.

Apart from this information, we also measured the following basic factors in the Bitcoin network to combine together forming a complete map of the structures changing over different peak periods: connectance (edges/nodes²), path Length (maximum and mean), clustering (number of clusters and the size of the largest cluster) and so on. The key results of the three peak periods are shown below:

<table>
<thead>
<tr>
<th>( \Delta \text{USD/Hr} )</th>
<th>Event 1 (R2 = .87)</th>
<th>Event 2 (R2 = .87)</th>
<th>Event 3 (R2 = .35)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diameter</strong></td>
<td>-0.81</td>
<td>AMI/Hc</td>
<td>-2.19</td>
</tr>
<tr>
<td><strong>T..</strong></td>
<td>-1.54</td>
<td>Avg. Path</td>
<td>-0.7</td>
</tr>
<tr>
<td><strong>Hc</strong></td>
<td>2.32</td>
<td>Tc</td>
<td>18.7</td>
</tr>
<tr>
<td><strong>Gen</strong></td>
<td>1.378</td>
<td>T..</td>
<td>-16.25</td>
</tr>
<tr>
<td><strong>AMI/Hc</strong></td>
<td>0.92</td>
<td>AMI</td>
<td>0.49</td>
</tr>
</tbody>
</table>
4. Conversation Analysis

In addition to the network analysis of the Bitcoin community, we considered the relationship between the Bitcoin price and conversations within and about Bitcoin. Conversations are both drivers and results of important social phenomena like economic bubbles. Previously, conversations were difficult to analyze rigorously, requiring sophisticated equipment and physical observation of participants. As social interaction expands to online spaces, however, entire conversations across large communities are captured as they occur.

Conversation analysis complements network analysis in identifying pivotal factors in the emergence of economic bubbles. While network analysis nicely evaluates structural transitions, conversation analysis addresses social psychological and cultural dynamics. We included conversation analysis in our approach to gain insight into changes in the network structure.

We recognized that the Bitcoin community faces two audiences, an internal community and an external environment that spread information and set regulations. Thus, we first divided our analysis by external and internal audience.

National media reflect the state of the reported object, but also influence that object through the report. Using LexisNexis, we collected all newspaper articles published within ten-day windows of our price peaks of interest (June 2011, August 2012, April 2013, and November-December 2013) that mention the keyword “bitcoin.” These articles are summarized in Table 1.

<table>
<thead>
<tr>
<th>Peak</th>
<th>Article Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 2011</td>
<td>9</td>
</tr>
<tr>
<td>Aug 2012</td>
<td>3</td>
</tr>
<tr>
<td>Apr 2013</td>
<td>105</td>
</tr>
<tr>
<td>Nov/Dec 2013</td>
<td>262</td>
</tr>
</tbody>
</table>

Table 1. Newspaper Article Counts by Peak Windows, Source: LexisNexis

Just like the Roman fora of antiquity, online forums play central roles in the collection and diffusion of information. The policy, norms, and markets of a community are both reflected and molded in the forum. We identified Bitcointalk.org as the dominant forum within the Bitcoin community. With about 100,000 users, this forum publishes a monthly online magazine and operates through a social media type of network. Interestingly, Bitcointalk.org experienced the highest volume of online participants during one of our peaks (April 10, 2013). This forum is summarized in Table 2.

<table>
<thead>
<tr>
<th>Total Members</th>
<th>342,055</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Posts</td>
<td>7,917,248</td>
</tr>
<tr>
<td>Ave. Posts Per Day</td>
<td>4,650</td>
</tr>
<tr>
<td>Total Topics</td>
<td>326,737</td>
</tr>
<tr>
<td>Ave. Topics Per Day</td>
<td>214</td>
</tr>
<tr>
<td>Total Boards</td>
<td>176</td>
</tr>
<tr>
<td>Ave. Users Online Per Day</td>
<td>353</td>
</tr>
<tr>
<td>Male : Female Ratio</td>
<td>4.5:1</td>
</tr>
</tbody>
</table>

Table 2. Bitcointalk.org Summary Statistics (as of July 21, 2014)

To narrow down our analysis of conversations internal to the Bitcoin community, we identified three hypotheses.

4.1 Self-Fulfilling Prophecy

Merton (1948) describes a self-fulfilling prophecy as the result of a feedback loop wherein the prophecy generates social action that produces the prophesized outcome. Merton uses a bank run to explain how “rumours” of insolvency affect the emergent outcome (Merton 1968). In the same way, we expect discussion of a bubble to precede an actual bubble.
4.2 Reputation Effect

Despite Bitcoin’s efforts to operate without trust mechanisms, reputation remains critical to any online marketplace. A negative reputation decreases an exchange participant’s leverage in the market. Moreover, reputation is a key differentiator when the market displays low dimensionality (Ghose et al. 2009). We thus hypothesized that an increase in enquiries or reports on scams and other negative reputation displays might predicate a sell-off of bitcoin and subsequent drop in the bitcoin price.

4.3 Technical Vulnerability

Our final hypothesis was that increased instances of technical troubleshooting would, similarly to reputation, cause a sell-off of bitcoin. The broader assumption was that as vulnerabilities are exposed in a system like bitcoin, lower-risk individuals would exit the market and the lower demand would cause the price to drop.

To evaluate these hypotheses, we collected post content and metadata from three forum boards of interest: Economics, Speculation, Scam Accusations, Reputation, Technical Support, and Development & Technical Discussion. These forums are summarized in Table 3. For our analysis, we grouped these by our hypotheses into economic, reputation, and technical content.

<table>
<thead>
<tr>
<th>Forum</th>
<th>Total # Topics</th>
<th>Total # Posts</th>
<th>Date Created</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economics</td>
<td>12,688</td>
<td>536,640</td>
<td>12/12/2009</td>
</tr>
<tr>
<td>Speculation</td>
<td>8,727</td>
<td>439,853</td>
<td>07/13/2011</td>
</tr>
<tr>
<td>Scam Accusations</td>
<td>2,357</td>
<td>49,029</td>
<td>06/27/2012</td>
</tr>
<tr>
<td>Reputation</td>
<td>1,191</td>
<td>6,973</td>
<td>04/19/2013</td>
</tr>
<tr>
<td>Technical Support</td>
<td>4,827</td>
<td>35,235</td>
<td>02/03/2010</td>
</tr>
<tr>
<td>Development &amp; Technical Discussion</td>
<td>8,116</td>
<td>105,770</td>
<td>12/30/2009</td>
</tr>
</tbody>
</table>

Table 3. Select Forum Summary (as of August 8, 2014)

As an initial analysis, we compared the volume of posts in given forums with the change in price. An example of this comparison, taken from the April 2013 and August 2013 peak windows, is shown in Figure 6.
Fig. 6. Comparison of Bitcoin price (top dark line in each figure) against Economics, Reputation, and Technical forum posts between April 6-14, 2013 (top figure) and August 23-31, 2013 (bottom figure).

4.4 Sentiment Analysis

To determine if the forum content indicated a negative discussion, which we hypothesized would lead to a drop in the bitcoin price, we performed sentiment analysis on the content. The content was divided into daily documents containing the posts or articles for that day. We prepared the text for analysis by tokenizing, filtering tokens by length, stemming, and filtering stopwords. Tokenizing the document created a “bag of words” by splitting the document into a series of “tokens” or character sets. Filtering tokens by length restricted the retained content to those tokens of at least 3 characters but no more than 999 characters. Stemming, using the Porter algorithm, reduces the remaining words to their minimal stem. This permits tokens such as “worked” and “work” to be evaluated as equivalent. Finally, filtering stopwords removes words common to the English language such as “a,” “and,” and “the.”

We used a linear Support Vector Machine (SVM) to create a model against which our article and forum content could be analyzed for sentiment. To train the SVM, we used SentiWordNet, a dictionary of words scored across positivity, negativity, and objectivity (Esuli and Sebastiani, Baccianella et al). Based on a word’s positivity and negativity scores, we split the SentiWordNet dictionary into positive and negative word lists. These lists served as the training documents for our SVM.

Once the SVM was created and the model had been trained, we applied the model to our article and forum content. An example of our results is shown in Figure 7.
4.5 Word Frequency Analysis

To evaluate our first hypothesis, that discussion of a bubble would generate a bubble, we analyzed the occurrence of the keyword “bubble” within the economic forums. An example of this analysis is shown in figure 8.

4.6 GLS Time Series Regression

We did find a potential correlation between the sentiment of econ discussion on the forum and the price. A drop in price appeared to occur in parallel to a drop in discussion sentiment. Evaluating a potential correlation across time-based trends requires a time series analysis. We chose to employ a Generalized Least Squares (GLS) model with a first-order autoregressive error term process to avoid assumptions of independent observations. We are in the process of expanding our sentiment dataset beyond the peak windows before we can reliably evaluate the results of the GLS time series analysis.

5. Conclusion

The Bitcoin marketplace provides a unique opportunity for information and social scientists to explore traditional phenomena such as economic bubbles in new light. Through network and conversation analysis, we have undertaken an innovative approach to evaluating the emergence of these phenomena. It is our hope that this research facilitates greater analysis of this complex system.
References