Bitcoin Research Overview

Kevin Smith

April 25, 2018

1 Background

1.1 Bitcoin

Bitcoin is a decentralized cryptocurrency first proposed in 2008. As with traditional currency, users can send and receive fractional units of Bitcoin to one another.[1] But there are several important distinctions between Bitcoin and traditional currency. Primarily, Bitcoin is decentralized—there is no central bank, no analogy to the Federal Reserve, no regulatory body, no entity single entity which mints new coins. This decentralization has several consequences:

- Rather than central authorities (banks) managing and validating transactions, **Bitcoin uses a peer-to-peer network** to track the flow of currency between wallets. If Alice wants to send Bitcoin to Bob, then Alice broadcasts this transaction to her peers in the network, who propagate this information to their peers, and so on, until the network reaches a consensus that the transaction has occurred. (Note that this peer network is *not* the network that we study in this project.)
- **The entire history of Bitcoin transactions is publicly available**. Anyone can download the record via online blockchain explorers[2] or through the official Bitcoin client.[3]
- **Bitcoin users are anonymous**. It is considered bad practice to re-use a Bitcoin address, and a typical user's Bitcoin assets are split over many unaffiliated addresses.[4] While some techniques exist to associate several addresses with a single user [1] [2], it is impractical for our purposes to link Bitcoin addresses and individual users.

Bitcoin is a tantalizing object of study—a huge quantity of high-quality data is freely available, but because de-anonymization is impractical, traditional approaches to studying economic networks do not apply.

The network that we propose to study is the *transaction network*, defined in [3]. Each vertex in the transaction network is a Bitcoin transaction, and weights on the edge $t_1 \to t_2$ are the amount of Bitcoin flowing from output addresses in t_1 to input addresses in t_2. In other words, this edge weight is the amount of Bitcoin involved in the transaction t_2 which was "financed" by transaction t_1. Using a Blockchain explorer, it is possible to download the transaction graph from any given day (or any other time interval).

[1] The smallest Bitcoin unit (analogous to the "cent" in USD) is the Satoshi, with a value of 10^{-8} Bitcoin.

[2] Example blockchain explorer: `https://blockchain.info/`

[3] Bitcoin Core: `https://bitcoin.org/en/download`

[4] Imagine that to buy a cup of coffee, you pay Starbucks from one of your 20 bank accounts, open a brand new account, and drain the remainder of your old account into the new one. This is how a typical Bitcoin transaction proceeds.

1.2 Network Motifs

Network motifs, introduced in [4] in 2002, provide a way to quantify the local structure of a complex network. An n-motif is a connected, n-node digraph, typically with $n \leq 5$. One way to quantify local graph structure is to calculate the *motif distribution*; that is, the fraction of n-node subgraphs which are isomorphic to each possible n-motif. Graphs with similar motif distributions have similar structure at the local level.

Daily transaction graphs typically have $\sim 10^6$ nodes, making it computationally infeasible to calculate their motif distributions exactly. Fortunately, a sampling algorithm exists [5], making it possible to uniformly sample a fixed fraction of n-node subgraphs. Therefore it is possible estimate the motif distributions for each daily transaction graph.

2 Research Questions

In general terms, we want to study how local-level structure in the Bitcoin transaction graph correlates with macro-level properties of the currency. We propose using motif distributions to quantify local graph structure. Macro-level properties of interest include the direction and magnitude of daily Bitcoin price fluctuations, as well as the number of daily Bitcoin transactions. Specifically, we aim to test the following hypotheses:

- Hypothesis 1: Motif distributions in daily transaction graphs will differ between (i) days with a large price increase, (ii) days with a moderate price increase, (iii) days with a small price change, (iv) days with a moderate price decrease, and (v) days with a large price decrease.
- Hypothesis 2: Motif distributions in daily transaction graphs will differ between (i) days with a large transaction volume, (ii) days with a moderate transaction volume, and (iii) days with a small transaction volume.

If we can confirm either hypothesis to a reasonable level of significance, we ask an additional question:

- Can we train a machine learning algorithm to predict daily price movement or daily transaction volume from daily motif distribution?

References

[1] D. Ron and A. Shamir, "Quantitative analysis of the full bitcoin transaction graph," in *International Conference on Financial Cryptography and Data Security*, pp. 6–24, Springer, 2013.

[2] A. Biryukov, D. Khovratovich, and I. Pustogarov, "Deanonymisation of clients in bitcoin p2p network," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 15–29, ACM, 2014.

[3] F. Reid and M. Harrigan, "An analysis of anonymity in the bitcoin system," in *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pp. 1318–1326, IEEE, 2011.

[4] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, "Network motifs in the transcriptional regulation network of escherichia coli," *Nature genetics*, vol. 31, no. 1, p. 64, 2002.

[5] S. Wernicke, "A faster algorithm for detecting network motifs," in *International Workshop on Algorithms in Bioinformatics*, pp. 165–177, Springer, 2005.