Detecting Bot Activity in the Ethereum Blockchain Network

Morit Zwang[1], Shahar Somin[1,2], Alex 'Sandy' Pentland[2], and Yaniv Altshuler[1,2]

[1]Endor Ltd. [2]MIT Media Lab, Cambridge, MA, USA
morit@endor.com, {shaharso, pentland, yanival}@media.mit.edu

The Ethereum blockchain network is a decentralized platform enabling smart contract execution and transactions of Ether (ETH) [1], its designated cryptocurrency. Ethereum is the second most popular cryptocurrency with a market cap of more than 100 billion USD, with hundreds of thousands of transactions executed daily by hundreds of thousands of unique wallets. Tens of thousands of those wallets are newly generated each day. The Ethereum platform enables anyone to freely open multiple new wallets [2] free of charge (resulting in a large number of wallets that are controlled by the same entities). This attribute makes the Ethereum network a breeding space for activity by software robots (bots). The existence of bots is widespread in different digital technologies and there are various approaches to detect their activity such as rule-based, clustering, machine learning and more [3,4]. In this work, we demonstrate how bot detection can be implemented using a Network Theory.

Being a platform used for human interactions, the Ethereum network can be described and modeled by a Network Theory approach. The degree distribution of such networks, for example, often displays a power law distribution [5]. This phenomenon can also be observed in the Ethereum network when constructing a graph network that represents Ethereum transactions between wallets—where each wallet is a vertex and a transaction between two wallets is an edge.

Previous research has demonstrated that time differences between consecutive events in many human activities display a power law distribution as well. The time difference between consecutive transactions in this work refers to the number of minutes between every transaction and its prior transaction. The time difference was calculated for the transactions of each wallet separately, and we created a histogram from the time difference of all the transactions of all wallets in the Ethereum network. The histogram of the time difference distribution of an arbitrary one-week sample shows that, indeed, the time difference between consecutive transactions demonstrates a power law distribution. (Fig. 1)

It can be observed that the distribution of time differences between the consecutive transactions of Ethereum wallets does not perfectly fit the power law model and is characterized by multiple spikes. Each spike represents a collection of highly correlated wallets which deviate from the expected power law distribution rather than resembling spontaneous human activity. Anomalies from the power law model in human interaction networks might represent the occurrence of potentially interesting events [6]. In this case, we assume that transactions which are anomalous to the power law model represent non-human behavior executed by bots. This assumption is based on the nature of the anomalies (spikes occurring at a very specific time difference) and on the observation of other properties common to the anomalous transactions, such as having the same transaction value.

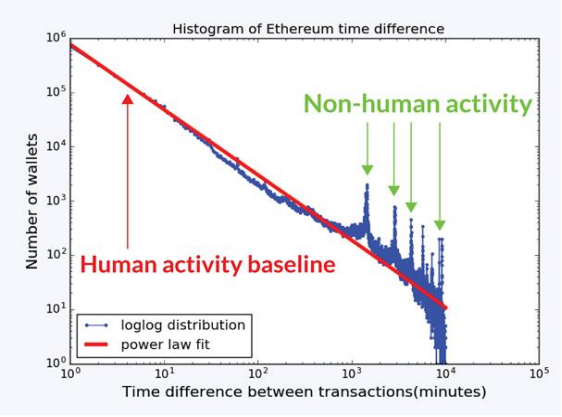

Fig. 1: An arbitrary one-week sample

References

[1] V. Buterin et al., “A next-generation smart contract and decentralized application platform,” white paper, 2014.

[2] Dannen, C. (2017). Introducing Ethereum and Solidity. Berkeley: Apress.

[3] Gu, G., Perdisci, R., Zhang, J., & Lee, W. (2008). Botminer: Clustering analysis of network traffic for protocol-and structure-independent botnet detection.

[4] Wang, A. H. (2010, June). Detecting spam bots in online social networking sites: a machine learning approach. In IFIP Annual Conference on Data and Applications Security and Privacy (pp. 335-342). Springer, Berlin, Heidelberg.

[5] Callaway, D. S., Newman, M. E., Strogatz, S. H., & Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. Physical review letters, 85(25), 5468.

[6] Akoglu, L., McGlohon, M., & Faloutsos, C. (2010, June). Oddball: Spotting anomalies in weighted graphs. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 410-421). Springer, Berlin, Heidelberg.