

Measurement and Analysis of the Bitcoin Networks: A View from Mining Pools

Authors: Canhui Wang[†], Xiaowen Chu^{*,†}, Yang, Qin[‡]

[†]Department of Computer Science, Hong Kong Baptist University

[‡]Department of Computer Science, Harbin Institute of Technology (Shenzhen)

*Email: chxw@comp.hkbu.edu.hk

July 25, 2020

Outline

1 Introduction

- Blockchain
- Bitcoin
- Mining Pool
- Related Work and Motivation

2 Our Work

- Data Collection
- Observation and Analysis

3 Conclusion

- Conclusion and Future Work

Section 1

1 Introduction

- Blockchain
- Bitcoin
- Mining Pool
- Related Work and Motivation

2 Our Work

- Data Collection
- Observation and Analysis

3 Conclusion

- Conclusion and Future Work

1.1. Blockchain

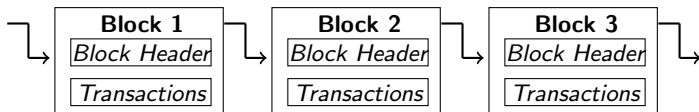


Figure: An Simplified Version of Blockchain

- **Blockchain is a Distributed Ledger Technology (DLT)** that was first proposed to solve the **Double-Spending Problem**¹ in 2008
- The entire blockchain is replicated and updated in full-node users. A **challenge**² is how to achieve a consistent state of blockchains of users, especially when there are large number of users?

¹Satoshi Nakamoto. *Bitcoin: A peer-to-peer electronic cash system*. Tech. rep. Manubot, 2019.

²Junfeng Xie et al. "A survey on the scalability of blockchain systems". In: *IEEE Network* 33.5 (2019), pp. 166–173.

1.1. Blockchain (cont'd)

- A not-bad **Solution**^{3,4} is the Proof-of-Work (PoW) protocol **because it works in a large-scale network** (>1000 nodes⁵)⁶, though it might not be as efficient as traditional Byzantine Fault Tolerant (BFT) solutions (<20 nodes⁵).

$$\text{SHA256} \left(\begin{array}{c} \text{Protocol version} \\ \text{Hash of previous block header} \\ \text{Merkle root of transactions} \\ \text{Timestamp} \\ \text{nbits} \\ \text{nonce} \leftarrow (\text{to be calculated}) \end{array} \right) \leq \text{Difficulty} \quad (1)$$

³Amos Fiat et al. "Energy equilibria in proof-of-work mining". In: *Proceedings of the 2019 ACM Conference on Economics and Computation*. 2019, pp. 489–502.

⁴PoW: Driven by an incentive mechanism, users compete to solve the PoW task. The user who first solve the PoW task (see inequality 1) can create a new block.

⁵Marko Vukolić. "The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication". In: *International workshop on open problems in network security*. Springer. 2015, pp. 112–125.

⁶It implies that an eventual consistency is guaranteed under the PoW assumption. ↻ 🔍 🔄

1.1. Blockchain (cont'd)

- **Permissionless Blockchains**

- ▶ Properties: **Anyone joins and/or leaves freely**
- ▶ Applications/Platforms: **Bitcoin (This work focuses on it)**, Ethereum

- **Permissioned Blockchains**

- ▶ Properties: **Participants must be authenticated and authorized**, such that they can access the network and complete certain operations (e.g., invoke, query transactions)
- ▶ Applications/Platforms: Hyperledger Fabric⁷, Quorum

⁷Canhui Wang and Xiaowen Chu. "Performance Characterization and Bottleneck Analysis of Hyperledger Fabric". In: *The Second IEEE International Workshop on Blockchain and Mobile Applications (BlockApp 2020) Co-located with the 40th IEEE International Conference on Distributed Computing Systems (ICDCS 2020)*. 2020, pp. 489–502.

1.2. Bitcoin

• Bitcoin⁸

- ▶ Bitcoin is a cryptocurrency **because it is built upon a public-key cryptography system**. Technical implementations mainly include public key, private key, and digital signature on a single node and the PoW protocol between different nodes.
- ▶ Bitcoin does not rely on third party **because only the distributed ledgers (i.e., blockchains) determine whether a transaction data is valid** or not. If yes, the transaction data will be permanently stored in the distributed ledgers. Thus, bitcoin can be transacted across different peers without trusting a third party.

⁸For example, Alice issues a transaction (along with her public key and digital signature) to Bob's public key. Upon the acceptance by blockchain, the transaction will be authenticatable, integrous and undeniable under the security assumptions in the PoW protocol and the public-key cryptography.

1.3. Mining Pool

- **Mining Pool** usually refers to an approach of solving the PoW task (see the 'nonce' field in inequality 1) by coordinating miners via a specific network protocol⁹.
 - ▶ With the popularity of Bitcoin, the "Difficulty" (see inequality 1) was getting increasingly difficult that ordinary miners cannot successfully solve it in ten minutes while mining rewards were very rewarding¹⁰ at that time. As a result, pooled mining becomes popular among miners.

⁹This approach requires miners to trust the pool operator. Miners request tasks from the operator and return the computing results through the network. Once news blocks are mined, the mining rewards are redistributed to miners via specific reward functions

¹⁰Blockchain.Info. In: <https://www.blockchain.com/explorer?view=btc>, Accessed on 23 July 2020.

1.4. Related Work and Motivation

- **Motivation 1:** Mining pools have become the dominant computing resources of the Bitcoin network. They are of great importance to both the security and performance of the Bitcoin network^{11,12}. Thus, it is important to study mining pools' computing resources.


¹¹Shuangke Wu et al. "Survive and Thrive: A Stochastic Game for DDoS Attacks in Bitcoin Mining Pools". In: *IEEE/ACM Transactions on Networking* 28.2 (2020), pp. 874–887.

¹²Some Real-World Cases of 51% Attacks: Different from small cryptocurrency systems such as Zcash and Bitcoin Gold that suffered from 51% attacks, the Bitcoin network has vast computing power such that 51% attacks requires a tremendous amount of computing power which is expensive.

1.4. Related Work and Motivation (cont'd)

- **Motivation 2:** The competition among mining pools is intense^{13,14}. Meanwhile, each mining pool's behaviors are dynamically adjusting for exploring rational strategies in an open Bitcoin network. Thus, though it is a challenging job, we aimed to learn something from mining pools' behaviors.

¹³Some Real-World Cases of Mining Pools' Birth and Death: During the observation period, many new mining pools have been created. Meanwhile, many old mining pools have disappeared.

¹⁴Rajani Singh, Ashutosh Dhar Dwivedi, and Gautam Srivastava. "Bitcoin Mining: A Game Theoretic Analysis." In: *IACR Cryptol. ePrint Arch.* 2018 (2018), p. 780. 

1.4. Related Work and Motivation (cont'd)

- **Motivation 3:** The impacts¹⁵ of mining pools (e.g., transaction fee, transaction delay) on Bitcoin end users. Mining pools' behaviors significantly affects the Bitcoin end users since it is the mining pools that process most of the users' transaction data.

¹⁵ Beltran Borja Fiz Pontiveros, Robert Norvill, and Radu State. "Monitoring the transaction selection policy of Bitcoin mining pools". In: *NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium*. IEEE. 2018, pp. 1–6.

1.4. Related Work and Motivation (cont'd)

- **Challenge 1:** The first challenge is **the inconsistent timestamps problem¹⁶ of transaction/block creators**, because nodes on the Bitcoin network could have different local system times.
- **Our Solution:** **We used one single machine's clock as the standard clock¹⁷.**

¹⁶Damiano Di Francesco Maesa, Andrea Marino, and Laura Ricci. "The bow tie structure of the bitcoin users graph". In: *Applied Network Science* 4.1 (2019), p. 56.

¹⁷For example, we used a local machine's clock as the standard system clock and calculate transaction delays for unconfirmed transactions. We used the remote BTC.com's clock as the standard system clock to calculate block intervals for confirmed blocks.

1.4. Related Work and Motivation (cont'd)

- **Challenge 2 (Practical):** The frequency of accessing local Mempool at Bitcoin Full Node when maintaining the timestamp list for unconfirmed transactions.
- **Our Solution:** In practice, to make it simple, we queried unconfirmed transactions every two seconds that is sufficient to satisfy the needs of maintaining the timestamp for unconfirmed transactions and not costly^{18,19}.

¹⁸A better solution could be implementing callback events, though things will be little complex then.

¹⁹Bellaj Badr, Richard Horrocks, and Xun Brian Wu. *Blockchain By Example: A developer's guide to creating decentralized applications using Bitcoin, Ethereum, and Hyperledger*. Packt Publishing Ltd, 2018.

Section 2

1 Introduction

- Blockchain
- Bitcoin
- Mining Pool
- Related Work and Motivation

2 Our Work

- Data Collection
- Observation and Analysis

3 Conclusion

- Conclusion and Future Work

2.1. Data Collection

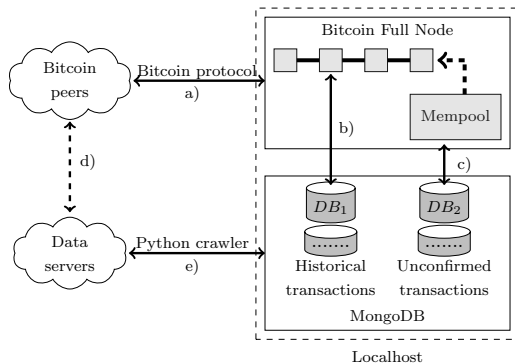


Figure: Our Measurement Architecture²⁰

²⁰Step a) ran Bitcoin Full node to **synchronize real-time Bitcoin traffics**. Step b) and c) set up two MongoDB databases to store **historical transaction data and unconfirmed transaction data**. Step d) and e) **obtained the block creator**.

2.1. Data Collection (cont'd)

Table: Block Processing Results²¹

Category	# of Blocks	Proportion	Time Span (mm/dd/yy)
Known pools	205,334	97.10%	02/25/2016 - 02/01/2020
Unknown pools	6,140	2.90%	02/25/2016 - 02/01/2020

Table: Top 4 Bitcoin Mining Pools^{22,23}

Mining Pool	Alias	# of Blocks	Time Span (mm/dd/yy)
AntPool*	N/A	33,576	02/25/2016 - 02/01/2020
BTC.com*	Block Trail	26,633	09/05/2016 - 02/01/2020
F2Pool*	Discus Fish	26,470	02/25/2016 - 02/01/2020
SlushPool*	Bitcoin.cz	16,287	02/25/2016 - 02/01/2020
ViaBTC	N/A	16,169	06/05/2016 - 02/01/2020

²¹ Around 97% blocks are created by mining pools.

²² Total number of processed blocks from Feb 25, 2016 to Feb 1, 2020 is 211,474.

²³ The symbol * indicates the top four major mining pools.

2.2. Observation and Analysis

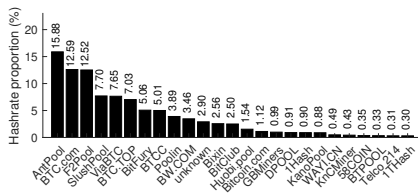


Figure: An Estimation of Blocks Created by Mining Pools from Feb 25, 2016 to Feb 01, 2020

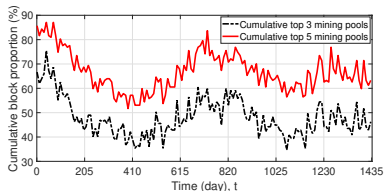



Figure: Cumulative Block Proportion of the Top Mining Pools from Feb 25, 2016 to Feb 01, 2020

2.2. Observation and Analysis (cont'd)

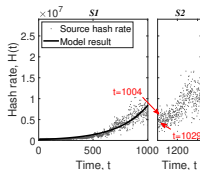
- **Observation 1:** During the observation span, over 94 percent of blocks are created by top 25 mining pools ²⁴.
- **Observation 2:** Bitcoin network relies heavily on a few top mining pools. ^{25,26}

²⁴It implies that over 94 percent of computing power and the generated bitcoins are controlled by these mining pools.

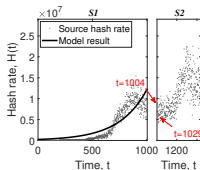
²⁵For example, more than 33 percent of Bitcoin blocks are created by only 3 mining pool entities every day.

²⁶Security Concerns: A trend of computing power centralization may raise security concerns, such as 51% attacks and selfish-mining attacks, to the Bitcoin network. 

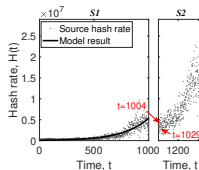
2.2. Observation and Analysis (cont'd)



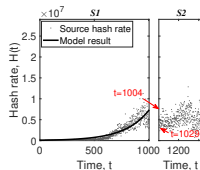
(a) AntPool



(b) BTC.com



(c) F2Pool



(d) SlushPool

Figure: Daily Hash Rate of the Top Mining Pools from Feb 25, 2016 to Feb 01, 2020

2.2. Observation and Analysis (cont'd)

- **Observation 3 (Summary):** In stage *S1*, hash rate of mining pool grows exponentially. Also, hash rate of Bitcoin network grows exponentially.

2.2. Observation and Analysis (cont'd)

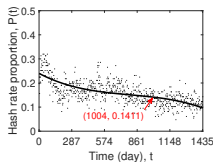
- **Observation 3 continued (Modeling and Validation):** Assume the growth rate λ of hash rate is a constant variable. Denote hash rate of a mining pool at time t be $H(t)$, we have

$$H(t) = e^{\lambda t} \cdot e^c \quad (2)$$

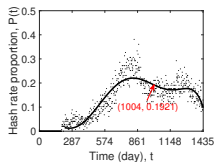
Table: Model Validation of $\ln(H(t)) = \ln(e^{\lambda t} \cdot e^c)$

Mining Entities	λ	c	R-square	Adj R-sq
AntPool	3.580×10^{-3}	12.3458	0.8746	0.8745
BTC.com	3.922×10^{-3}	12.3896	0.8047	0.8045
F2Pool	4.302×10^{-3}	11.1786	0.8083	0.8081
SlushPool	4.488×10^{-3}	11.3000	0.8389	0.8390
Bitcoin Network	4.048×10^{-3}	13.8813	0.9477	0.9477

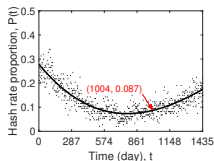
2.2. Observation and Analysis (cont'd)



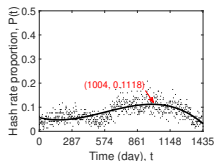
(a) AntPool



(b) BTC.com



(c) F2Pool



(d) SlushPool

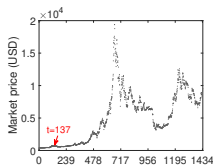
Figure: Daily Hash Rate Proportion of the Top Mining Pools from Feb 25, 2016 to Feb 01, 2020

2.2. Observation and Analysis (cont'd)

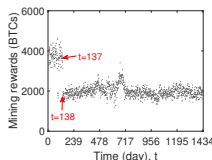
- **Observation 4 (Summary):** An exponential growth of hash rate does not mean an increasing market share (i.e., hash rate proportion) of a mining pool. On the converse, some mining pools (e.g., AntPool and F2Pool) decreased their market shares in spite of an exponential growth of hash rate. In literature, such a phenomenon is often called the Malthusian trap²⁷.

²⁷ Clement Tisdell and Serge Svizzero. "The ability in antiquity of some agrarian societies to avoid the Malthusian trap and develop". In: *Forum for Social Economics*. Vol. 49. 2. Taylor & Francis. 2020, pp. 202–227.

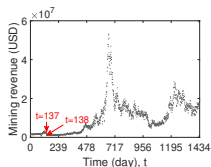
2.2. Observation and Analysis (cont'd)



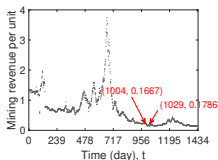
(a) Market Price



(b) Mining Rewards



(c) Mining Revenue



(d) Mining Revenue per Unit

Figure: Daily Mining Revenue and Mining Revenue per Unit from Feb 25, 2016 to Feb 01, 2020

2.2. Observation and Analysis (cont'd)

- **Observation 5 (Summary):** There is a threshold, where when the unit profit of hash rate is greater than the threshold, hash rate of Bitcoin network increases. When the unit profit of hash rate is lower than the threshold, hash rate of Bitcoin network decreases.

2.2. Observation and Analysis (cont'd)

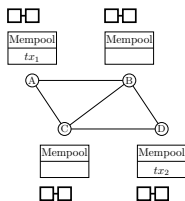
- **Observation 5 continued (Modeling and Validation):** Assume mining pools are rational to maximize their mining revenues and the unit cost per hash rate is a constant denoted by ξ . Mining revenue of a given day t is denoted by $R(t)$. Hence, the net profit of a mining pool of day t is as follows.

$$Net(t) = R(t) - \xi \cdot H(t) \quad (3)$$

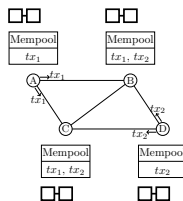
In literature, such a model is often called the Cournot Competition model²⁸. Following this model, the unit cost per hash rate can be proved to convergent to a stable state. Our empirical results show a stable state: the unit cost per hash rate $\xi = 0.1667$ USD per GHash/s.

²⁸Kostas Bimpikis, Shayan Ehsani, and Rahmi Ilkılıç. “Cournot competition in networked markets”. In: *Management Science* 65.6 (2019), pp. 2467–2481. 

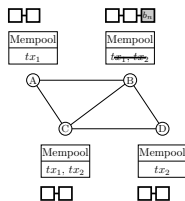
2.2. Observation and Analysis (cont'd)



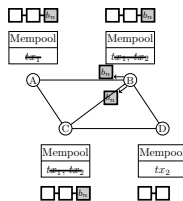
(a) Create Tx



(b) Broadcast Tx



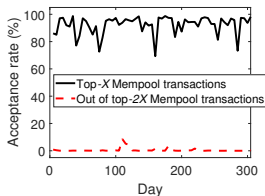
(c) Create Blk



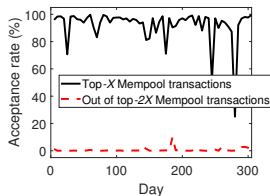
(d) Broadcast Blk

Figure: Broadcast Unconfirmed Transactions across the Bitcoin network (a,b,c,d)

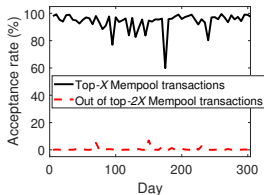
2.2. Observation and Analysis (cont'd)



(a) AntPool



(b) F2Pool



(c) Create Block

Figure: Acceptance rate of local top- X transactions with the highest *feerate* from March 6, 2018 to January 3, 2019

2.2. Observation and Analysis (cont'd)

- **Observation 6:** The factor 'feerate' significantly impacts transaction collection strategy of mining pools. Other factors, such as waiting time, transactions size, have no significant results.

2.2. Observation and Analysis (cont'd)

- **Observation 6 continued (Modeling and Validation):** Assume mining pools are rational to maximize their mining revenues. Given a subset of n unconfirmed transactions, each transaction i has a transaction fee p_i and a transaction size w_i . The maximum block size is c . Let $x_i = 1$ means that mining pool choose to collect transaction i ; otherwise, $x_i = 0$. Thus, we have,

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n p_i x_i \\ & \text{subject to} && \sum_{i=1}^n w_i x_i \leq c \\ & && x_i \in \{0, 1\}, \quad i \in \{1, 2, 3, \dots, n\} \end{aligned} \tag{4}$$

In literature, such a model is often called the Knapsack model²⁹.

²⁹ Mohamed Baza et al. "Blockchain-based charging coordination mechanism for smart grid energy storage units". In: *2019 IEEE International Conference on Blockchain (Blockchain)*. IEEE. 2019, pp. 504–509.

Section 3

1 Introduction

- Blockchain
- Bitcoin
- Mining Pool
- Related Work and Motivation

2 Our Work

- Data Collection
- Observation and Analysis

3 Conclusion

- Conclusion and Future Work

3.1. Conclusion and Future Work

- ① Mining pool's unit cost per hash rate will be stable according to the Cournot Competition model and our observations (our empirical result: around $\xi=0.1667$ USD per GHash/s). If this holds, then an interesting question is how does the rule affects the security of Bitcoin in the near future?
- ② Mining pools are stuck in a Malthusian trap where an exponential growth of hash rate does not mean an increasing market shares. If this holds, then an interesting question is how do mining pools escape the trap in the future?
- ③ If mining pools are rational to maximize their mining revenues, then a practical (may not optimal) solution is to adopt 'feerate-first' principle. If this holds, then an interesting question is how does affects end users, e.g., fee competition among Bitcoin end users?

Thank You!