Economics of Blockchain Storage

Yunshu Liu†, Zhixuan Fang†, Man Hon Cheung†, Wei Cai‡ and Jianwei Huang§

Abstract—**Miners in a blockchain system are suffering from the ever-increasing storage costs, which in general have not been properly compensated by the users' transaction fees. In the long run, this may lead to less participation of miners and jeopardize the blockchain security. In this paper, we study the economics of blockchain storage and identify the incentive issues related to this storage cost problem. More specifically, we model the interactions among users (who generate transactions) and miners in two stages, where the users set the transaction fees in Stage 1, and the miners select which transactions to include in Stage 2. Through characterizing the Nash equilibrium of the two-stage game, we find that the transaction fees indeed cannot cover the storage costs under the current practice in general, due to the negative externality and the unfair delay-based pricing. We also identify that a longer block interval can alleviate the concern by raising the transactions fees at the expense of larger delay.**

I. Introduction

The past decade has witnessed the fast growth of the Bitcoin [1], which is essentially a decentralized ledger proposed by Satoshi Nakamoto in 2008. As of 2019, Bitcoin has reached a market capitalization of over 150 billion USD.[1] As the enabling technology of Bitcoin, *blockchain* allows trustless but secure trading among different parties without a centralized authority [1]. Thus, blockchain has attracted tremendous attention from both the academia and industry. Various blockchain systems, such as Ethereum, Steemit, and IOTA, have been proposed for various applications.

The ever-increasing transaction volume imposes huge storage costs to *miners*.[2] Fig. 1 shows a typical blockchain operation process. When a user generates a new *transaction*,[3] he needs to propose a *transaction fee* as an incentive for miners to include the transaction into a future *block*.[4] Next, each miner will compete with others to produce a block and include (some of) the transactions. The miner who produces a new block can get the transactions' fees as commissions. Meanwhile, that miner also receives the *newly issued token* (also known as the *mining reward*) as bonus. Since the blockchain requires all miners to store the full transaction history locally, the booming of blockchain systems rapidly consumes huge amount of storage resource. From Feb. 2017 to Dec. 2018, the block size of Ethereum grows from 2 kilobytes to roughly 20 kilobytes. As a result, the storage size of an archive node grows nearly fourfold from 385 gigabytes to 1.8 terabytes.[5] Such a problem exists in many blockchain systems. The increase in block size causes the entire Bitcoin blockchain to reach 242.39 gigabytes.[6] Given the fact that there are many active miners in these popular blockchain systems, huge and continuous storage costs are incurred on miners.

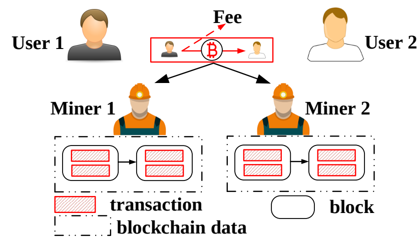

Fig. 1: Blockchain operation.

A key concern for such a rapid increase of storage size is that the storage costs may not be fully compensated by the current blockchain incentive mechanisms in the near future. Miners' storage costs are compensated by both the transaction fees and the mining reward. Currently, the mining reward accounts for the majority of miners' revenues (e.g. Bitcoin), which together with the transaction fees, can fully compensate storage costs. However, as the mining reward is designed to gradually decrease in many blockchain systems to give the cryptocurrency anti-inflationary properties (e.g. Bitcoin [1]), whether the transaction fees alone are sufficient to compensate the storage costs in the long run becomes a critical issue. In fact, the transaction fees cannot cover the ever-increasing storage costs currently [3]. For example, storing 1 kilobyte of data in Ethereum costs 0.656 USD on Apr 8th, 2019. Although

*This work is supported by the Shenzhen Institute of Artificial Intelligence and Robotics for Society, the General Research Fund CUHK 14219016 from Hong Kong UGC, the Presidential Fund from the Chinese University of Hong Kong, Shenzhen and Project 61902333 supported by National Natural Science Foundation of China.

†Y. Liu, Z. Fang, and M. H. Cheung are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: {ly018, zxfang, mhcheung}@ie.cuhk.edu.hk).

‡W. Cai is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China; Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China (e-mail: caiwei@cuhk.edu.cn).

§J. Huang is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China; Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China; the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: jianweihuang@cuhk.edu.cn).

[1] https://coinmarketcap.com/

[2] In blockchain, nodes participate in the consensus protocol and process the data are referred to as miners [2].

[3] The transaction serves different purposes in different systems e.g. Bitcoin transaction is a transfer of cryptocurrency between different addresses.

[4] A block is a container of transactions and it contains the cryptographic hash of the previous block, a time-stamp, and the data [2].

[5] https://wiki.parity.io/FAQ#what-are-the-parity-ethereum-disk-space-needs-and-overall-hardware-requirements

[6] https://www.statista.com/statistics/647523/worldwide-bitcoin-blockchain-size/

this is not cheap, it still cannot cover the storage costs, given that every miner needs to store the data *permanently*. As a result of the systematic reduction of mining reward over time, the increasing storage costs and insufficient transaction fees may discourage miners from staying in the system, and may lead to a high entrance barrier for incoming miners. For example, the number of full nodes storing Bitcoin blockchain declines by 11% in 2018.[7] Such a problem is catastrophic to the blockchain in the long-run: With fewer miners, the blockchain will be more vulnerable to a single point of failure [2], and the whole system will be less decentralized, becoming easier for the malicious miners to launch attacks [4]. To maintain a healthy decentralized ecosystem, it is critically important to systematically investigate the issue of insufficient transaction fees.

Despite of the heated discussion on the above issue in the technical community,[8] to the best of our knowledge, there has not been a *theoretical* study on this topic. This motivates us to take the first step in this paper to understand the economics of blockchain storage. More specifically, we model the storage-aware transaction fee determination as a two-stage game among users and miners. In Stage 1 (the fee setting stage), users choose the transactions fees to minimize their delays, considering the fees set by other users. In Stage 2 (the mining stage), miners maximize their own payoffs by selecting the proper subset of transactions to include from the pool. The key results and contributions are summarized as follows:

- *Economics of blockchain storage:* To the best of our knowledge, this is one of the first theoretical studies on the economics of blockchain storage. We analyze the reasons behind the insufficient fees for covering storage costs in blockchain, and we identify the related incentive issues in the mining process.
- *Closed-form characterization of system equilibrium:* We propose a model to characterize how users set transaction fees by considering the storage costs in blockchain, which applies to many current systems such as Bitcoin, Ethereum, and Litecoin. Generalizing previous works, our model allows users to choose transaction fees from a continuous interval considering the transactions' delays and storage costs. We are able to derive the Nash equilibrium of the model in closed-form.
- *Insights on redesigning blockchain storage mechanism:* We find that in general the transactions fees cannot cover the storage costs for two reasons: the negative externality and the unfair delay-based pricing. We also identify that a longer block interval can increase the transaction fees, but will lead to a larger delay.

II. Literature Review

Previous works on the users' transaction fees mainly focus on the *transaction delay*.[9] In [5], Huberman *et al.* built a game theoretic model about blockchain and found that users propose transaction fees to shorten the transaction delay. In [6], Easley *et al.* modeled the users choose between a fixed level transaction fee and zero fee. The analysis showed the ratio of users who pay fees increases with the transaction delays. In [7], Li *et al.* applied queuing theory to model users' transaction delays. At a given fee level, they derived how miners' expected profits change with the transaction delays. In [8], Peter R. Rizun proved that users will still pay a fixed level transaction fees even without block size limit.

Overall, these previous works have not considered the economics aspect of storage costs. Besides, previous models are based on a restricted assumption of the fixed level transaction fee, while in practice users set fees from a continuous interval according to the transaction delays. In this paper, we consider the situation where users set transaction fees from a continuous interval considering storage costs in blockchain.

III. System Model

We introduce the blockchain system model in this section. We first describe the storage-aware transaction fee determination in Section III-A, and model it as a two-stage Stackelberg game in Section III-B. Sections III-C and III-D introduce the detailed models in Stage 1 and 2, respectively.

A. Storage-aware Transaction Fee Determination

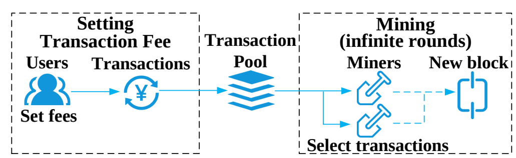

Fig. 2: Two-stage storage-aware transaction fee determination.

The storage-aware transaction fee determination is a two-stage decision process as illustrated in Fig. 2:

1) Setting the transaction fees: Users generate transactions that will enter the transaction pool. For each newly generated transaction, the corresponding user will set a transaction fee, paid to the miner who records the transaction in the blockchain.
2) Mining: The mining stage comprises of infinite *rounds of mining*. One round of mining corresponds to the period from reaching consensus on one block to reaching consensus on the immediately next block. In each round, miners select transactions from the pool. Once a miner finds the next block, *all* miners need to store the new block to reach consensus. The miner who finds the block will get the transaction fees from his selected transactions and mining reward. The transactions in the new block are recorded into the blockchain (and deleted from the pool).

Next, we introduce the model of the two-stage game.

B. Two-stage Stackelberg Game

We model the decision process as a two-stage Stackelberg game, as illustrated in Fig. 3. There are 2 users and M miners.

[7]https://bitcoinist.com/bitcoin-nodes-10k-reachable/

[8]https://ethereum-magicians.org/search?q=state%20rent

[9]Transaction delay is the time gap between a transaction is generated and it is recorded in blockchain.

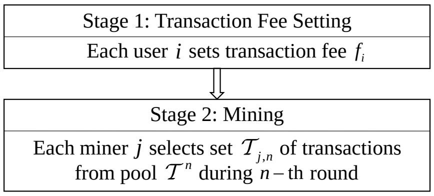

Fig. 3: Two-stage Stackelberg game.

- In Stage 1, each user i sets a transaction fee f_i that applies to all his transactions. In each round of mining, each user i will generate one transaction with probability p_i. We label the transactions from user i based on generation time as $tx_{i,k}$ $(k = 1, 2, \ldots)$, denoted as a sequence $(tx_{i,k})_{k=1}^{\infty}$.
- Stage 2 comprises infinite rounds of mining. Without loss of generality, we examine the n-th round of mining ($n = 1, 2, \ldots$). The newly generated transactions enter transaction pool $\mathcal{T}^n$, which is the set of all unrecorded transactions till n-th round. Each miner j selects set $\mathcal{T}_{j,n} \subseteq \mathcal{T}^n$ and includes $\mathcal{T}_{j,n}$ in his block to be mined. Then, a miner finds the new block. We assume the zero communication delay such that miners reach consensus instantly. All the miner store the new block and the model enters $(n+1)$-th round of mining.

Currently, the mining reward accounts for majority of miners' revenue, which can support a steady number of miners. Hence we assume that the number of miners is fixed in our model.

Example: Fig. 4 shows an example of Stage 2, the n-th round of mining, for the case of $M = 2$. We use $\emptyset$ to

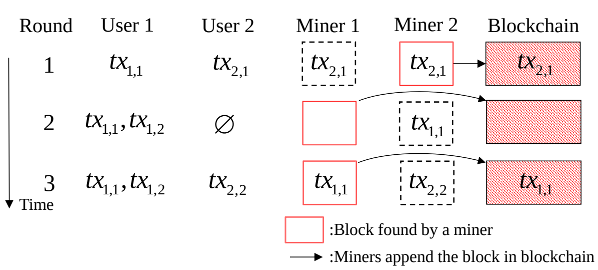

Fig. 4: Stage 2 operation.

represent no unrecorded transactions, red rectangle to represent the block found by a miner, and arrow to represent that a block is appended to all miners' local storage of blockchain.

- Round 1: Users 1 and 2 generate transactions $tx_{1,1}$ and $tx_{2,1}$ respectively. Both miner 1 and 2 choose to include $tx_{2,1}$. Miner 2 finds the next block, thus $tx_{2,1}$ is recorded into the blockchain and $tx_{1,1}$ is left to the next round.
- Round 2: User 1 generates transaction $tx_{1,2}$. Miner 1 includes no transaction and miner 2 includes $tx_{1,1}$. The block found by miner 1 records no transaction, and $tx_{1,1}$ and $tx_{1,2}$ are left to the next round.
- Round 3: User 2 generates $tx_{2,2}$. Miner 1 and 2 include $tx_{1,1}$ and $tx_{2,2}$, respectively. The block found by miner 1 records $tx_{1,1}$, while $tx_{1,2}$ and $tx_{2,2}$ are left to the next round.

Next we introduce the assumptions of our model.

1) Assumptions: To facilitate the later analysis, we make following assumptions about mining process:

(a) One block can contain at most 1 transaction regardless of the transaction size.
(b) The total transaction generation probability per round is smaller than 1. That is $p_1 + p_2 < 1$.
(c) Each round of mining (i.e. *block interval*) has a fixed time duration L.

Assumption (a) characterizes the reality of block size limit, which has been adopted in [6] [7]. Assumption (b) assumes that blockchain has enough space to record all the transactions. This is because the maximum outflow rate of transaction pool is 1 transaction per round (as every block contains at most 1 transaction) and inflow rate per round is $p_1 + p_2 < 1$. In reality, a stable system like Bitcoin is able to record all the transactions and the block size can be adjusted to accommodate the increasing transaction volume. Assumption (c) serves as an approximation of the block generation in real-world systems (e.g. average block interval is 10 minutes in Bitcoin [1]). It allows us to rigorously analyze the system and generates enough insights.

C. Stage 1: Transaction Fee Setting

In Stage 1, each user i sets the *transaction fee* $f_i \geqslant 0$ (for all his transactions) to maximize his payoff, and we formulate the users' fee interactions as a non-cooperative game.

We assume that all transactions of user i consume s_i bytes of storage ($i = 1, 2$). This simplifies the analysis but still captures the reality that different users may generate different sizes of transactions.

1) Transaction delay: Each user faces a *tradeoff* between paying a high fee and suffering a high transaction delay. Since miners prefer higher-fee transactions, user i will experience a lower delay by setting a higher fee than the other user. However, if the other user sets a very high fee, user i would be better off by setting a lower fee and baring a higher delay.

Here we define the delay of transaction $tx_{i,k}$:

- Generation time: For transaction $tx_{i,k}$ generated at n-th round mining, its generation time is $t_{i,k}^g = n$. For example, in Fig. 4, $tx_{1,1}$'s generation time is $t_{1,1}^g = 1$.
- Recorded time: We define:

$$t_{i,k}^r = \begin{cases} l, & \text{if } tx_{i,k} \text{ is recorded at the } l\text{-th round,} \\ \infty, & \text{if } tx_{i,k} \text{ never gets recorded in blockchain.} \end{cases}$$

- Delay: $d_{i,k} = t_{i,k}^r - t_{i,k}^g$, which is the recorded time minus the generation time. For example, in Fig. 4, $tx_{1,1}$'s recorded time and generation time are $t_{1,1}^r = 3$ and $t_{1,1}^g = 1$, respectively. Thus, delay is $d_{1,1} = 2$.

2) Users' payoff functions: We first characterize user i's payoff from one transaction, and then we define his payoff function in terms of the long-term average.

- User i's payoff from a single transaction $tx_{i,k}$:

$$v_{i,k} = \begin{cases} R_i - f_i - \gamma d_{i,k}, & \text{if } tx_{i,k} \text{ is recorded in blockchain,} \\ -\gamma d_{i,k}, & \text{otherwise,} \end{cases} \quad (1)$$

where R_i represents user i's level of satisfaction of having one transaction recorded in blockchain and γ is the delay cost factor. The transaction delay $d_{i,k}$ corresponds to the cost for the user, and γ reflects the patience level of users. A higher γ means that the users are less patient.

- User i's long-term average expected payoff function is:

$$\mathbb{E}(u_i) = \lim_{T\to\infty} \frac{\sum_{k=1}^{N_i(T)} \mathbb{E}(v_{i,k})}{T}, \tag{2}$$

where $\mathbb{E}(v_{i,k})$ is user i's expected payoff from $tx_{i,k}$ and $N_i(T)$ is the number of user i's generated transactions from round 1 to T.

3) Game formulation: We formulate users' transaction fee decision process as a non-cooperative game, where users set fees simultaneously to maximize their own payoffs.

Definition 1 (Stage 1: Fee setting game): A fee setting game is a tuple $\Omega = (\mathcal{I}, \mathcal{F}, \boldsymbol{U})$ defined by:

- Players: The set of users $\mathcal{I} = \{1, 2\}$.
- Strategies: Each user $i \in \mathcal{I}$ chooses a transaction fee f_i from the strategy space $\mathcal{F}_i = [0, \infty)$. The feasible set of all strategy profiles is $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$.
- Payoffs: The vector $\boldsymbol{U} = (\mathbb{E}(u_1), \mathbb{E}(u_2))$ contains two users' expected payoff functions defined in Equation (2).

D. Stage 2: n-th Round of Mining

The mining stage consists of infinite rounds of mining. Without loss of generality, we consider the n-th round of mining in Stage 2, where each miner selects which transaction to include in order to maximize his payoff. We formulate the interactions among the miners as a non-cooperative game.

There are $M \geqslant 2$ miners. We use β_j to denote the probability of miner j finding the next block within each round, so β_j represents miner j's mining power.

1) Miners' decisions: Miner j selects a set $\mathcal{T}_{j,n}$ of transaction from the pool $\mathcal{T}^n$ based on transactions' fees and storage costs. Each block contains either one transaction or not. Thus $\mathcal{T}_{j,n} \subseteq \mathcal{T}^n$ and the cardinality of $\mathcal{T}_{j,n}$ satisfies $|\mathcal{T}_{j,n}| \in \{0, 1\}$.

Since for any user i, transaction tx_{i,k_1} and tx_{i,k_2} ($k_1 \neq k_2$) yield the same fees and storage costs, miner j's strategy falls into one of following three cases:

1) Not selecting any transaction.
2) Selecting the earliest transaction from user 1 in the pool.
3) Selecting the earliest transaction from user 2 in the pool.

To simplify the description, we use $\mathcal{T}_{j,n} = \emptyset, \{1\}$, and $\{2\}$ to represent the above three cases, respectively.

2) Miners' payoff functions: Miner j's payoff consists of cost and revenue.

- Cost: Miners have both mining cost and storage cost.
  - Mining cost: We assume that the mining cost is proportional to mining power and mining time [9] [10]. Let us denote C^m as the mining cost per unit mining power per unit time, hence miner j's mining cost in the n-th round is $\beta_j L C^m$, where L is the duration of each round.
  - Storage cost: Let C^s denote the storage costs of permanently storing 1 byte. We assume that all miners have the same storage costs per byte. If miner j finds the next block and includes $tx_{i,k}$ in the block, he will incur the storage cost $s_i C^s$ for storing that transaction. In the blockchain consensus protocol, if another miner l successfully finds the next block, miner j still needs to store that block and bare the storage costs for the transaction which miner l selects [11]. Thus the storage costs for miner j is:

$$\begin{cases} \sum\limits_{i\in\mathcal{T}_{j,n}} s_i C^s, & \text{if } j \text{ finds next block (w.p. } \beta_j\text{)}, \\ \sum\limits_{i\in\mathcal{T}_{l,n}} s_i C^s, & \text{if } l \text{ finds next block (w.p. } \beta_l,\ l \neq j\text{)}, \end{cases} \tag{3}$$

where *w.p.* represents "with probability".

- Revenue: Only the miner who successfully finds the next block gets both the *mining reward* R^m and the transaction fee from his selection. Thus the revenue for miner j is:

$$\begin{cases} R^m + \sum\limits_{i\in\mathcal{T}_{j,n}} f_i, & \text{w.p. } \beta_j, \\ 0, & \text{w.p. } 1-\beta_j. \end{cases} \tag{4}$$

Combining the revenue and the cost, miner j's payoff in the n-th round of mining is as follows:

$$w_{j,n}(\mathcal{T}_{j,n}, \mathcal{T}_{-j,n}) = \beta_j \sum_{i\in\mathcal{T}_{j,n}} (f_i - s_i C^s) - \sum_{l\neq j} \beta_l \sum_{i\in\mathcal{T}_{l,n}} s_i C^s + \beta_j (R^m - LC^m). \tag{5}$$

3) Game formulation: We formulate the n-th round of mining as a non-cooperative game, where miners select the transactions simultaneously to maximize their own payoffs.

Definition 2 (Stage 2: Transaction selection game): The transaction selection game in Stage 2 is a tuple $\Gamma = (\mathcal{M}, \mathcal{X}, \boldsymbol{W})$ defined by:

- Players: The set of miners $\mathcal{M} = \{1, 2, \dots M\}$.
- Strategies: Miner j's strategy space is $\{\emptyset, \{1\}, \{2\}\}$. The feasible set of all strategy profiles is $\mathcal{X} = \{\emptyset, \{1\}, \{2\}\}^M$, where exponent M represents Cartesian product.
- Payoffs: The vector $\boldsymbol{W} = (w_{1,n}, w_{2,n}, \dots, w_{M,n})$ contains the payoff functions of M miners defined in Equation (5).

IV. Nash Equilibrium Analysis

In this section, we analyze Nash equilibrium of the two-stage game by backward induction. We derive Stage 2's equilibrium in Section IV-A and Stage 1's equilibrium in Section IV-B, respectively.

A. Miners' Transaction Selection Game in Stage 2

We start by analyzing Stage 2. Here, we use $-j$ to represent all the miners other than miner j.

Definition 3 (Nash Equilibrium): A strategy profile $\mathcal{T}_n^* = (\mathcal{T}_{j,n}^*, \mathcal{T}_{-j,n}^*)$ constitutes a *Nash equilibrium* (NE) if

$$w_{j,n}(\mathcal{T}_{j,n}^*, \mathcal{T}_{-j,n}^*) \geqslant w_{j,n}(\mathcal{T}_{j,n}, \mathcal{T}_{-j,n}^*), \forall \mathcal{T}_{j,n}, \forall j \in \mathcal{M}. \tag{6}$$

For the simplicity of presentation, we define $h_i \triangleq f_i - s_i C^s$ as the net fee of transaction $tx_{i,k}$. Given the users' decisions in Stage 1, we define the highest-net-fee transaction in the pool as

$$tx_{i^*,k} = \arg \max_{tx_{j,l} \in \mathcal{T}^n} h_j.$$

We summarize the NE in the n-th round of mining of Stage 2 in the following theorem.

Theorem 1: The Nash equilibrium of Stage 2 is:

- Case I: The pool contains no transactions, then each miner j's strategy is $\mathcal{T}^*_{j,n} = \emptyset$.
- Case II, III and IV: The pool contains transactions.
  - If $h_{i^*} \geqslant 0$, each miner j's strategy is $\mathcal{T}^*_{j,n} = \{i^*\}$.
  - If $h_{i^*} < 0$, each miner j's strategy is $\mathcal{T}^*_{j,n} = \emptyset$.

We leave all the proof of lemmas, propositions, and theorems in online appendix [12] due to space limits.

The intuition of Theorem 1 is two-fold.

1) Transaction priority: The net fee h_i determines transaction $tx_{i,k}$'s priority ($tx_{i,k} \in \mathcal{T}^n$). The higher net fee h_i means miners will always record transactions from user i in the blockchain first (higher priority than $tx_{-i,k} \in \mathcal{T}^n$).
2) Threshold fee: The threshold fee for miners to select transactions $tx_{i,k}$ is $s_i C^s$, which is insufficient to cover M miners' total storage costs $M s_i C^s$.

Negative externality: The key reason behind the second intuition is the *negative externality*. That is, the miner only considers individual payoff maximization and neglects the storage costs he imposes on other miners, which leads to insufficient-fee transactions being recorded in blockchain.

B. Users' Fee-setting Game in Stage 1

Based on the NE in Stage 2, we analyze Stage 1.

Without loss of generality, we assume $p_1 \leqslant p_2$, meaning that user 1 generates fewer transactions than user 2. We define mixed strategy NE to facilitate the analysis.

Definition 4 (Mixed Strategy Nash Equilibrium): Let $\mathscr{B}^2$ be the space of all probability measures on $D = [0,\infty) \times [0,\infty)$. A two-dimension probability measure $\boldsymbol{\omega}^* \in \mathscr{B}^2$ constitutes a *mixed strategy Nash equilibrium* if the following inequality holds for all $i \in \mathcal{I}$ and $\forall \boldsymbol{\omega}_i \in \mathscr{B}$,

$$\begin{aligned} &\int_D \mathbb{E}(u_i(f_i, f_{-i})) d\big(\boldsymbol{\omega}^*_i(f_i) \times \boldsymbol{\omega}^*_{-i}(f_{-i})\big) \\ &\geqslant \int_D \mathbb{E}(u_i(f_i, f_{-i})) d\big(\boldsymbol{\omega}_i(f_i) \times \boldsymbol{\omega}^*_{-i}(f_{-i})\big). \end{aligned} \tag{7}$$

We denote f_i^{NE} as user i's fee at the NE of Stage 1 and we summarize the NE of Stage 1 in the following theorem.

Theorem 2: The unique mixed strategy Nash equilibrium in Stage 1 is:

- User 1 chooses the transaction fee f_1^{NE} uniformly in the interval of $[s_1 C^s, s_1 C^s + \gamma d]$.
- User 2 chooses the transaction fee according to

$$f_2^{\mathrm{NE}} \begin{cases} = s_2 C^s, & \text{w.p. } 1 - \frac{p_1}{p_2}, \\ \in U(s_2 C^s, s_2 C^s + \frac{\gamma p_1}{1 - p_1 - p_2}), & \text{w.p. } \frac{p_1}{p_2}, \end{cases}$$

where U stands for uniform distribution.

An interesting insight of Theorem 2 is that the user who generates more transactions pays lower net fee per transaction. The reason is that as user 2 generates more transactions than user 1, user 1's transactions suffer a higher delay under a low priority than user 2's. Thus user 1 has more incentive to set a higher net fee than user 2 to achieve a higher priority and hence reduces his delay cost.

Unfair delay-based pricing: The key reason behind this insight is the *delay-based pricing*, where the user who suffers a higher delay will pay a higher net fee.

Delay-based pricing is *unfair* because the user who consumes more storage resource not necessarily pays more. It also implies that we can increase the expected fee by increase the delays. We will further elaborate it in the next section.

To sum up, transaction fees cannot cover storage costs because of *negative externality* and *unfair delay-based pricing*.

Next we derive the condition of the transaction fees cannot cover storage costs.

Lemma 1: As the number of round T approaches infinity, the sum of all transaction fees cannot cover the sum of all miners' costs for storing transactions if and only if

$$\frac{\gamma \min(p_1^2, p_2^2)}{1 - p_1 - p_2} < C^s (M-1)(p_1 s_1 + p_2 s_2).$$

We combine Lemma 1 with the blockchain practice. In a popular blockchain system, there are many miners (M is large). To provide good user experience, the block size will be adjusted to avoid large delay (which means p_i is small). Thus, the condition in Lemma 1 can be satisfied and the transaction fees are unlikely to cover storage costs in practice.

V. Impact of Longer Block Interval

In this section, We show that *a longer block interval* can increase fees to compensate the storage costs, meanwhile it will increase the delay.

A. Longer Block Interval Increases Fees

The length of the block interval is related to the delay experienced by the users. The longer block interval leads to the longer delay. Thus, users are more willing to increase the transaction fees under the delay-based pricing.

With this idea, we can adjust the block interval time[10] to let the total revenues cover the total costs, which facilitates the sustainable mining. Specifically, as mining reward decreases, we can increase the block interval L, such that the mining reward and expected fee together exactly cover all miners' storage costs and mining costs. The corresponding condition is:

$$R^m + \mathbb{E}(f^{\mathrm{NE}}) - C_{\mathrm{total}} - L C^m = 0, \tag{8}$$

where $\mathbb{E}(f^{\mathrm{NE}})$ denotes the expected fee at NE and C_{total} denotes the total storage cost of all miners. The block interval also affects transaction generation probability per round p_i. For user i, transaction generation rate per minute C_i is considered

[10]Tuning the block interval time is practical. E.g., the mining difficulty of Bitcoin is adjusted roughly every two weeks to keep the block interval constant [1]. Thus, increasing mining difficulty will increase the block interval.

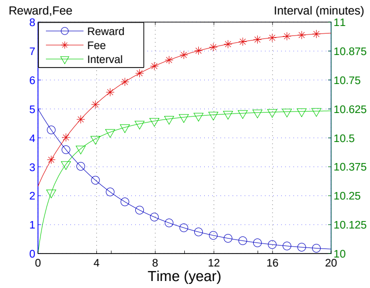

Fig. 5: As mining reward decreases, the corresponding fee and block interval required to balance the revenues and the costs.

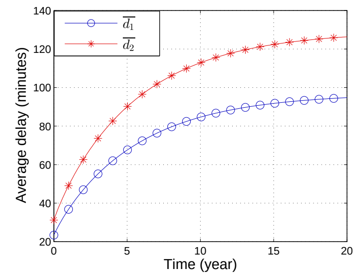

Fig. 6: The corresponding average delays of transactions as block interval changes with time.

as a constant and not affected by the block interval. Thus, the transaction generation probability per round satisfies

$$p_i = LC_i. \tag{9}$$

Based on delay-based pricing, the expected fee $\mathbb{E}(f^{\mathrm{NE}})$ will increase as block interval L increases. Thus, we can choose the proper block interval L to let the total revenues cover the total costs according to Equation (8) and (9). The mathematical details are shown in online appendix [12].

We use the numerical results to demonstrate this approach. In Fig. 5, we illustrate the evolution of the mining reward, transaction fee, and block interval with time. The initial value of mining reward is $R^m = 5$ and is halved every four years, following the pattern in Bitcoin [1]. Based on the mining reward level at a given time, we determine the transaction fee and block interval L such that the total revenue exactly cover the total costs, which is the solution of Equation (8) and (9). Other parameters are $M = 11$, $C^m = 1$, $C^s = 0.5$, $s_1 = 1$, $s_2 = 1.5$, $C_1 = 0.037$ and $C_2 = 0.055$.

In Fig. 5, as the mining reward decreases, we can observe that the longer block interval will cause users to pay more fees. This can help compensate the maintenance expense under decreasing mining reward. On the other hand, it actually significantly increases the transactions' average delay, as we will discuss in the next section.

B. Longer Block Interval Increases Delays

We define the average delay of transaction sequence $(tx_{i,k})_{k=1}^{\infty}$ as

$$\overline{d_i} = \lim_{T\to\infty} \frac{\sum_{k=1}^{N_i(T)} \mathbb{E}\left(d_{i,k}\right) L}{N_i(T)}. \tag{10}$$

where $N_i(T)$ is the number of user i's generated transactions from round 1 to T, $\mathbb{E}(v_{i,k})$ is user i's expected payoff from $tx_{i,k}$ and L is the block interval.

As the block interval increases with time in Fig. 5, we illustrate the corresponding average delays in Fig. 6. We can observe that both users' delays significantly increase as block interval in Fig. 5 increases with time. The reason is that the increase of block interval L leads to the increase of transaction delay. The delay increase force users to pay more fees, which replaces the shrinking mining reward and cover the storage costs.

To sum up, a longer block interval can alleviate the storage cost issue in blockchain, but it causes a poor user experience, because users pay more fees but suffer larger delays.

VI. Conclusion

In this paper, we studied the economics of blockchain storage and identified some potential issues in the current incentive mechanisms. Specifically, we modeled the storage-aware transaction fee determination as a two-stage game and derived the Nash equilibrium in closed-form. We revealed that insufficient fees for storage costs are mainly due to the negative externality and the unfair delay-based pricing. We also found that a longer block interval means sufficient fees to cover storage costs, but causes a larger delay.

In the future work, we will consider a more general case where miners have heterogeneous storage costs and the block generation time is a random variable. We will also design an incentive mechanism to address the insufficient fees for storage cost problem in blockchain.

References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," https://bitcoin.org/bitcoin.pdf, 2008.

[2] W. Cai, Z. Wang, J. B. Ernst, Z. Hong, C. Feng, and V. C. Leung, "Decentralized applications: The blockchain-empowered software system," *IEEE Access*, vol. 6, pp. 53 019–53 033, 2018.

[3] V. Buterin, "Blockchain resource pricing," https://ethresear.ch/t/a-simple-and-principled-way-to-compute-rent-fees/1455, 2018.

[4] "Bitcoinwiki. weaknesses," https://en.bitcoin.it/wiki/Weaknesses.

[5] G. Huberman, J. Leshno, and C. C. Moallemi, "Monopoly without a monopolist: An economic analysis of the bitcoin payment system," working paper, Columbia Business School, 2017.

[6] D. Easley, M. O'Hara, and S. Basu, "From mining to markets: The evolution of bitcoin transaction fees," *Journal of Financial Economics*, 2019.

[7] J. Li, Y. Yuan, S. Wang, and F.-Y. Wang, "Transaction queuing game in bitcoin blockchain," in *IEEE Intelligent Vehicles Symposium*, June 2018.

[8] P. R. Rizun, "A transaction fee market exists without a block size limit," *Block Size Limit Debate Working Paper*, 2015.

[9] L. W. Cong, Z. He, and J. Li, "Decentralized mining in centralized pools," *National Bureau of Economic Research*, no. 25592, Feb 2019.

[10] M. Carlsten, H. Kalodner, S. M. Weinberg, and A. Narayanan, "On the instability of bitcoin without the block reward," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Oct 2016, pp. 154–167.

[11] W. Wang, D. T. Hoang, Z. Xiong, D. Niyato, P. Wang, P. Hu, and Y. Wen, "A survey on consensus mechanisms and mining management in blockchain networks," https://arxiv.org/abs/1805.02707v1, 2018.

[12] "Online appendix," https://www.dropbox.com/s/3wb53t2pd0a4szw/Liuyunshu_OnlineAppendix.pdf?dl=0.