Optimal Pricing Mechanism for Data Market in Blockchain-Enhanced Internet of Things

Kang Liu, *Member, IEEE*, Xiaoyu Qiu, Wuhui Chen, *Member, IEEE*, Xu Chen, *Member, IEEE*, and Zibin Zheng, *Senior Member, IEEE*

***Abstract*—With the rapid development of the Internet of Things (IoT) in the era of big data, the amount of collected data has increased dramatically. Data are one of the most important commodities in IoT. To maximize the utility of the collected data, it is crucial to design an open IoT data market that enables data owners and consumers to carry out data trading securely and efficiently. To address the challenge of security presented by an untrusted and nontransparent data market, we propose an edge/cloud-computing-assisted, blockchain-enhanced data market framework to support secure and efficient IoT data trading, with a particular focus on an optimal pricing mechanism. In this mechanism, an authorized market-agency works as a scheduler, determining the win-owner and its pricing strategy to the consumer. We formulate a two-stage Stackelberg game to solve the pricing and purchasing problem of the data consumer and the market-agency. In the first stage of the game, the market-agency gives the win-owner and its pricing strategy. In the second stage, the data consumer decides on its purchasing quantity of data. We consider competition between data owners and propose a competition-enhanced pricing scheme (CPS). We apply backward induction to analyze the subgame perfect equilibrium at each stage for both independent and CPSs. Lastly, we validate the existence and uniqueness of Stackelberg equilibrium, and the numerical results show the efficiency of the CPS.**

***Index Terms*—Blockchain, data market, Internet of Things (IoT), smart pricing mechanism, Stackelberg game.**

I. Introduction

RECENT years have witnessed the rapid development of the Internet of Things (IoT), along with the exploding data collected from various IoT node devices [1]. With the increasing requirements for new IoT network service, the massive amounts of IoT data have become a valuable commodity. Efficient utilization of the collected IoT data is a critical issue, which has created an emerging IoT business model called sensing as a service (SaaS) [2]. The basic idea of SaaS is that sensors are employed to collect data that could be vended to interested users or other devices [3]. To facilitate global data and anonymized SaaS transactions, the concept has been proposed of an IoT data market, which could automate the negotiation between data owners and consumers [4], [5].

Manuscript received February 18, 2019; revised April 15 2019, June 4, 2019, and July 13, 2019; accepted July 21, 2019. Date of publication July 26, 2019; date of current version December 11, 2019. This work was supported in part by the National Key Research and Development Plan under Grant 2018YFB1003803, in part by the National Natural Science Foundation of China under Grant 61802450 and Grant 61722214, in part by the Natural Science Foundation of Guangdong under Grant 2018A030313005, and in part by the Program for Guangdong Introducing Innovative and Entrepreneurial Teams under Grant 2017ZT07X355. *(Corresponding author: Wuhui Chen.)*

The authors are with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510275, China (e-mail: liuk73@mail.sysu.edu.cn; qiuxy23@mail2.sysu.edu.cn; chenwuh@mail.sysu.edu.cn; chenxu35@mail.sysu.edu.cn; zhzibin@mail.sysu.edu.cn).

Digital Object Identifier 10.1109/JIOT.2019.2931370

Data are considered a virtual commodity with the characteristics of variety, volume, and velocity [1], although traditional market models must be modified to enable an effective data market. Many third-party big data markets have been designed for exploiting the economic value of big data in the Internet, such as Airbnb and Uber [6], [7]. However, an Internet-oriented data market model cannot be applied to the data collected from IoT devices, because IoT data have unique characteristics, such as the requirements of privacy and real time, which demand a specific market model [3].

In this article, our aim is to design an automated negotiation and pricing mechanism for the IoT data market. To establish an effective IoT data market, the following challenges must be addressed.

1) Design a trustworthy market framework for IoT data trading that can meet the high-level requirements of security, privacy, and copyright protection. This is due to the fact that IoT data usually contain sensitive information and can be easily counterfeited or duplicated [3].
2) Develop an efficient negotiation and pricing mechanism for the IoT data market. The pricing mechanism for IoT data must of sophisticated design to ensure profit, fairness, and incentives for the participants. In addition, because IoT data are usually demanded in near real time with low delay tolerance [8], the proposed negotiation and pricing mechanism must be efficient to meet high-level real-time requests.

To address the first challenge, we apply blockchain technology in the IoT data market. Blockchain is a fully transparent, peer-to-peer (P2P), distributed ledger that allows transactions to occur directly between entities from different parties in a verifiable and permanent manner [9]–[11]. Previous studies have investigated the potential of using blockchain technology when establishing energy-trading systems for IoT, which motivated us to consider blockchain as a solution to the IoT data market [12], [13]. However, an unavoidable challenge when we employ blockchain in the IoT data market is that a secure and efficient consensus mechanism may not be well supported by the limited resource of IoT devices [3], [14]. To address this issue, we propose an edge/cloud-computing-assisted, blockchain-enhanced framework for the IoT data market. The shortcoming of resources can thus be compensated with the assistance of edge/cloud computing.

To address the second challenge, we propose an optimal pricing mechanism for the IoT data market. We introduce an authorized market-agency in the IoT data market, which is predefined as part of the smart contracts in the blockchain-based trading system. The market-agency enables data owners to set prices for their data dynamically to maximize their economic profits and determine the win-owner for the data consumer. The profit of both win-owner and consumer can be jointly maximized via a game-theory-based pricing algorithm. By introducing an incentive competition mechanism, the profit of the consumer can be further improved. The negotiation and pricing process can be automated and implemented efficiently to satisfy real-time requests.

The main contributions of this article are summarized as follows.

1) In this article, we exploit blockchain technology to establish a market system for IoT data. We propose an edge/cloud-computing-assisted, blockchain-enhanced framework to address the challenges of security, trust, and efficiency in the IoT data market.
2) We design an optimal pricing mechanism for the IoT data market. In particular, we adopt a Stackelberg game to jointly maximize the profits of win-owner and consumer. In this game, an authorized market-agency acts as the leader, determining the win-owner and giving the price for its data. The data consumer acts as the follower and determines its purchasing quantity of data from the win-owner.
3) Through backward induction, we first investigate the optimal quantity of purchased data in the second stage. We then investigate the pricing strategies of market-agency in the first stage. We consider competition between data owners and propose the competition-enhanced pricing scheme (CPS). We prove that the Stackelberg equilibrium is derived analytically for both independent pricing scheme (IPS) and CPS.
4) We conduct extensive numerical simulations to evaluate the performance of the proposed pricing schemes. Numerical results show that the proposed CPS is effective and efficient for IoT data trading.

The remainder of this article is organized as follows. Section II presents related research and Section III introduces the system components of the edge/cloud-computing-assisted blockchain-enhanced data market. We describe two trading models as well as the key trading operations. In Section IV, we propose an optimal pricing mechanism for the data market. We formulate a Stackelberg game for the pricing and purchasing problem and solve it using backward induction for both IPS and CPS. Section V presents the system evaluation and performance analysis, and Section VI concludes.

II. Related Work

With the advent of the era of big data, data products from various sources have become economic commodities. Although the study of the economics of data products is still in its initial stages, many researchers have addressed the problems of data valuation and trading mechanisms. Many third-party data trading markets have been designed. Cao *et al.* [7] formulated the trading problem of multiple data owners, collectors, and users and proposed an iterative auction mechanism to coordinate the trading among selfish agents in a socially optimal way without direct access to their private information. Niu *et al.* [15] proposed a secure mechanism that integrates truthfulness and privacy preservation in data markets. This secure mechanism can guarantee truthfulness and privacy by using homomorphic encryption and identity-based signature. However, the above centralized data market models may perform inefficiently when applied to IoT data markets, which usually have limited computing resources but massive transactions.

Recently, a number of studies have explored the potential of blockchain, IoT, and cloud (BIC)-based nonfinance solutions [16], [17]. Perboli *et al.* [18] proposed a standard digital strategy design methodology for blockchain projects and presented a use case of blockchain in logistics. Shuaib *et al.* [19] proposed a decentralized energy exchange system that allows for energy to be exchanged between producers and consumers using a form of smart electronic contracts based on blockchains. Xue *et al.* [20] presented a transaction management framework based on blockchain to facilitate P2P electronic trading in microgrid. Kang *et al.* [21] designed a localized P2P energy-trading system for electric vehicles called "P2P electricity trading system with consortium blockchain" (PETCON) by exploiting a consortium blockchain. These studies reveal that data sharing and data market are crucial for BIC-based nonfinance solutions. However, there is a lack of a smart pricing mechanism specially designed for the IoT data market.

Much effort has been made to develop secure, efficient, and low-complexity pricing models for the IoT data market. Zheng *et al.* [22] proposed a profit-driven data acquisition framework in the crowd-sensed data market by jointly considering the problems of profit maximization and payment minimization. Chen *et al.* [23] proposed an agent-based reinforcement learning system to mimic professional trading strategies and monitor unreasonable strategies made by traders. Wang *et al.* [24] proposed two pricing models for data trading in device-to-device communication networks, namely a Stackelberg game based one-buyer/multiple-sellers pricing model and an alternative ascending clock auction based one-seller/multiple-buyers pricing model. However, none of the above data pricing models considered the influence of competition between data owners on the pricing. Compared with other bilevel programming models, the Stackelberg game based pricing models enable a more complex market structure and more discriminatory pricing strategies [25]. Therefore, here, we develop a blockchain-enhanced IoT data market model and design a Stackelberg-game-based optimal pricing mechanism that introduces competition into the market. To our knowledge, this is the first study to adopt blockchain technology and competition-based smart pricing mechanisms in the IoT data market.

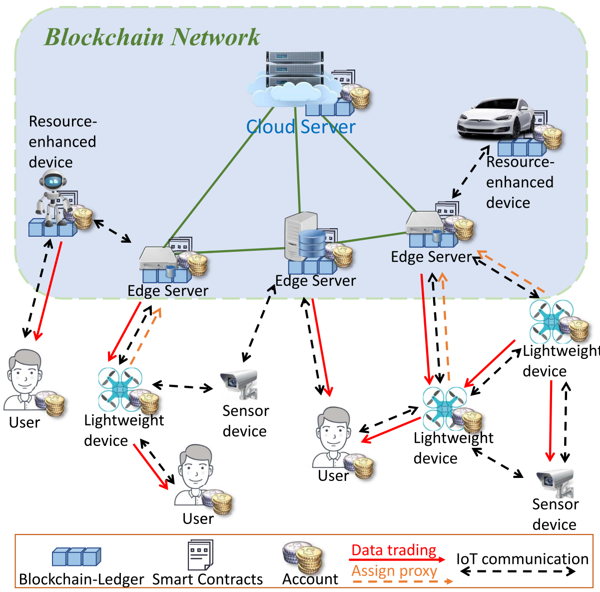


Fig. 1. System architecture of the blockchain-enhanced IoT data market.

III. Blockchain-Enhanced Data Market Framework for IoT

In this section, we provide a blockchain-enhanced data market framework for IoT that can address the challenges of security and efficiency in IoT data trading. Fig. 1 depicts the general architecture and entities of the blockchain-enhanced IoT data market.

A. System Architecture Overview

The IoT data market in Fig. 1 is represented as an ecosystem including IoT devices, edge/cloud servers, and blockchain network. In our framework, we divide the IoT devices into two categories based on their resources, that is, resource-enhanced devices and lightweight devices. Resource-enhanced devices are IoT devices with sufficient resources which can act as mining nodes and create blocks for the blockchain system, including smart robots, vehicular devices, and other smart IoT devices. Lightweight devices are the devices with restrict resources which cannot work as blockchain peers, including sensors, unmanned aerial vehicles (UAVs), personal wearable devices, etc. Sensor devices, such as cameras, radars, and meters, are responsible for collecting data. Most of the IoT devices and users have private accounts, enabling them to trade data with other peers. Edge servers provide computational resources to facilitate the blockchain consensus and smart data trading & pricing processes. In this framework, the blockchain network is integrated with the edge/cloud servers, forming a novel blockchain-edge/cloud ecosystem that can support a series of smart contracts for secure and efficient data trading interactions.

In this IoT data market, data transactions are performed among different parties that usually have conflicting interests. Thus, the lack of trust always exists in the market. To address the challenge of trust, we introduce a blockchain to record all transactions in the market. We adopt the proof-of-work (PoW) consensus protocol to provide high security and trust for data trading [21], [26], [27]. For efficiency and protection of private data, we apply a specifically designed permissioned consortium blockchain in the IoT data market system [28]. However, the limited computing and storage capacities of IoT devices cannot efficiently support the PoW consensus protocol in the blockchain-based system. To address this challenge, edge/cloud computing is introduced to the framework. In our framework, IoT devices can offload blockchain-related tasks, such as mining and auditing tasks, and trading-related tasks to the edge/cloud servers [26], [27], [29]. Edge servers are deployed close to devices, providing edge computing with low latency and cost [30], [31]. The cloud server acts as an authority and provides wide-area monitoring and controlling for all systems, thus further improving system security [32], [33]. Moreover, to ensure the efficiency of blockchain consensus, edge servers can also offload their workloads to the cloud when they have insufficient computing resources for their local tasks. The blockchain network cooperates with edge/cloud computing in a symbiotic manner. Edge/cloud servers provide computation resources for blockchain tasks, while the blockchain network supports secure device-to-edge and device-to-cloud data interactions.

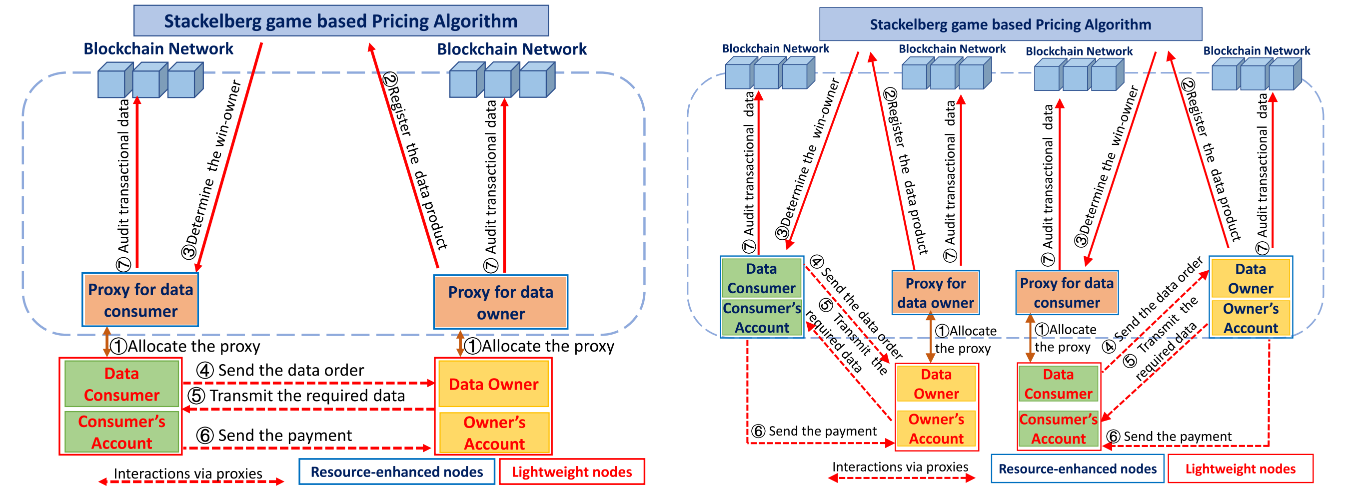
In the framework, trading related protocols are predefined in the smart contracts which are deployed in all blockchain peers. Smart contracts support efficient and secure data interactions by utilizing blockchain and edge/cloud computing. In a blockchain-based system, each smart contract acts as a finite state machine which executes laid down protocol when an action has been activated based on an instance. The trading models and optimal pricing mechanism of our framework are all included in the smart contracts. When local resources of IoT devices are inadequate, the corresponding smart contracts will be triggered, and help coordinate the cooperation and task offloading among IoT devices and edge/cloud servers.

B. Blockchain-Enhanced IoT Data Trading Models

The proposed blockchain-enhanced IoT data market framework provides two trading models, namely P2P model and proxy-based model.

In the P2P model, IoT devices involved in the trading act as blockchain peers that conduct consensus process and store the increased blockchain data. The P2P model is proposed for IoT devices with sufficient resources (resource-enhanced devices in Fig. 1), while the P2P model offers IoT devices optimal guarantees for their transactions.

The proxy-based model is designed for devices that do not have sufficient resources to work as blockchain peers (lightweight devices in Fig. 1). In the proxy-based model, lightweight devices can entrust blockchain peers as their proxies, performing all trading-related processes for them, including auditing tasks and smart pricing tasks. The transactional record is stored and audited in the blocks of the proxies. A proxy node can be a neighbor resource-enhanced device or edge server. Generally, the lightweight devices with fewer neighbor proxy nodes have higher priority for selecting their proxies. When a lightweight device has multiple neighbor



(a) (b)

Fig. 2. Key operations of proxy-based trading model. Key operations of data trading between (a) two lightweight devices and (b) resource-enhanced devices and lightweight devices.

proxy nodes, priority is given to the proxy node with more available resources.

The P2P and proxy-based models generally coexist in our IoT data trading market. All transactional data can be audited into a unified blockchain. IoT devices set trading models based on their economic benefits based on their economic benefits, capacities, and security requirements. A device can also change its trading model if its capacity changes (e.g., switching from the P2P model to the proxy-based model when its resources are occupied by other tasks). With these two trading models, all IoT devices can trade their data efficiently and securely.

C. Operations of IoT Data Trading

We now discuss the general trading operations in the proposed market framework. Fig. 2 illustrates the key operations of the proxy-based model. In the scenario of Fig. 2(a), the trading occurs between two lightweight devices. Both data owner and data consumer should entrust their proxies to perform the blockchain consensus process and smart pricing mechanism. In the scenario of Fig. 2(b), transactions occur between resource-enhanced devices and lightweight devices. In this case, only the lightweight devices entrust their proxies to perform the blockchain consensus process and smart pricing mechanism. Resource-enhanced devices perform the blockchain consensus process locally. For P2P model, all devices are resource-enhanced devices, which perform the blockchain consensus process locally. Operations of the P2P model are similar with those of proxy-based model, except that proxies are removed in Fig. 2. Then, we will describe the operations of data trading in detail.

1) System Initialization: The blockchain-based trading system needs initialization with cryptography to prevent malicious attacks such as Sybil attacks. In the system, each node (IoT device or edge server) becomes a legitimate entity by registration with a trusted authority that is deployed on the cloud server. Smart contrasts deployed in the system ensure the authority of IoT nodes and the security of transactions. When joining the system, each node obtains a unique true identity, its public and private keys, and its digital signature by using the elliptic curve digital signature algorithm (ECDSA) [17], [34], [35]. The authority also allocates a wallet address for each node as its account of data coins. After registration, the authority generates a mapping list for each node and stores the list in the cloud. Then, the system allocates blockchain peers to edge/cloud servers and resource-enhanced devices according to the related smart contracts.

2) Determining Data Owner and Price: After initialization, data owners in the market register their data in the cloud, which can be accessed by all blockchain peers. Authorized market-agencies are predefined by the smart contracts for homotypic data owners. The market-agency can provide the data consumer and optimal data owner (win-owner) a suitable price for the data. The win-owner can be any IoT node that provides the data products. When a lightweight node is selected as the win-owner, it entrusts its proxy to perform all its trading and auditing processes. Determining the win-owner and its price is one of the most challenging tasks in our market framework. We design a game theory based optimal pricing mechanism to determine the win-owner and its price. The optimal pricing mechanism is laid down in the smart contracts, and executed once any consumer requests some data product from the system. Section IV explains this process in detail.

3) Transaction Generation: After confirming the win-owner and its price, the market-agency sends the win-owner an order message for the desired data with the consumer's ECDSA signature. The win-owner verifies the order as well as the consumer's identity. After that, the win-owner sends the required data to the consumer directly or via the cloud. The consumer checks and confirms whether the data is valid. If the data is confirmed valid, the consumer sends the corresponding data coins to the win-owner's public wallet address.

After the payment is finished, a transaction is generated and broadcasted to all blockchain peers in the network. The transaction data includes the description of data product, encrypted

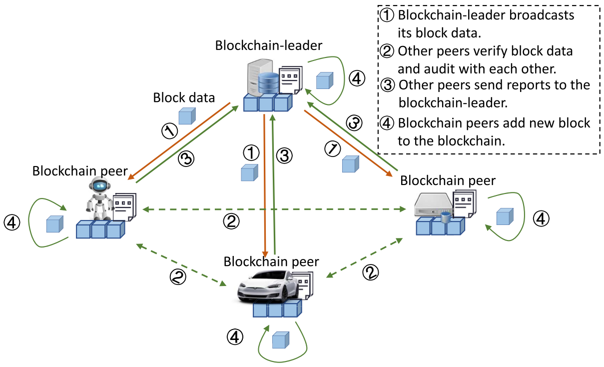


Fig. 3. Blockchain consensus process for data trading.

signatures and account information of trading peers, trading time, and payment information. Each blockchain peer gathers and verifies all generated transactions for a certain period, and then generates a block by structuring the transactions with its signature.

4) Consensus Process: In a blockchain-based system, consensus is the process to reach an agreement of the whole blockchain peers onto a trust about their transactions. In our system, the PoW based consensus process can be conducted with the assistance of edge/cloud. Edge servers and proxies work as miners in the consensus process. The consensus process of our system is illustrated as Fig. 3. First, each blockchain peer competes with others to faster give a valid PoW. The blockchain peer that first provides a valid PoW is selected as the blockchain-leader, and obtains the authority of updating the blockchain ledger. Then the blockchain-leader broadcasts its block data with its signature to other blockchain peers. For mutual supervision and verification, all other blockchain peers audit the blockchain-leader's block data and broadcast their audit results to each other with their signatures. Each blockchain peer compares its own audit result with others, and sends a report to the blockchain-leader. If all blockchain peers agree on the block data, they reach consensus. Finally, the blockchain-leader sends a reply to all other blockchain peers for permit updating the distributed ledger. The blockchain-leader's block will be added into the blockchain, and the transactions in the new block are considered as confirmed transactions.

IV. Optimal Pricing Mechanism for IoT Data Market

In the blockchain-enhanced IoT data market, the consumer obtains the required data by paying the win-owner. Determining the win-owner and giving the optimal price to maximize the profits of all trading parties constitute the IoT data pricing problem. In this section, we model the pricing problem as a Stackelberg game, in which the market-agency acts as the leader, determining the win-owner and the price of its data. The consumer acts as the follower, determining the quantity of purchased data.

A. Problem Formulation

The interactions of data owners and data consumer are illustrated in Fig. 4. In the simplest scenario, the market consists

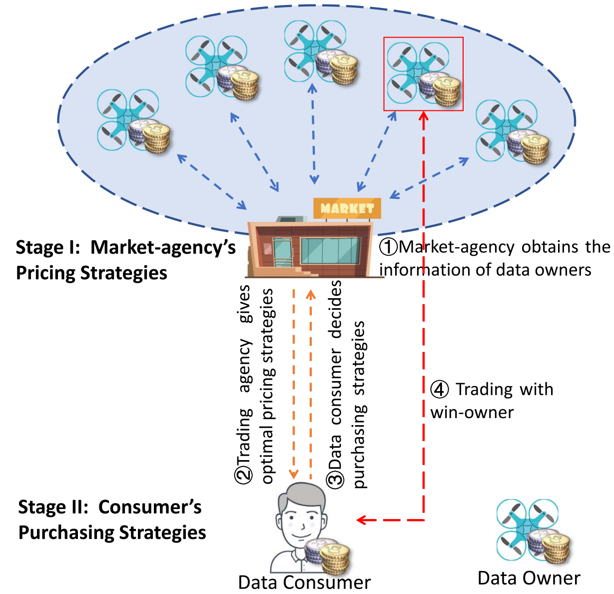


Fig. 4. Stackelberg game-based optimal pricing mechanism.

of some data owners, a data consumer, and a market-agency. Data owners are IoT devices that are responsible for gathering, storing, and transferring data to the consumer. The market-agency is predefined in the smart contracts and works as a scheduler when a consumer asks for some data. We consider that there is a group of N eligible data owners, whose set is denoted by $\mathcal{N} = \{1, \dots, N\}$, which can provide the required raw data. The market-agency can obtain the list of data owners and the trading-related information for both consumer and owners. In our trading model, for a given consumer (i.e., data buyer), each data owner $i \in \mathcal{N}$ sets the optimal price for its data, and the consumer determines the purchasing strategy, i.e., the quantity of purchased data. We denote the unit price and quantity of data purchased from data owner i as p_i and x_i, respectively. Similar, the optimal price and optimal quantity of data purchased from i are denoted as p_i^* and x_i^*, respectively.

The market-agency selects the win-owner based on the service quality of data owners and the purchasing willingness of the data consumer. We denote the maximum quantity of data that the data owner i can provide as $x_i^{\max}$ and denote the minimum quantity of data required by the data consumer as $x_{\min}$. Because of replicability, the data of an owner can be sold to multiple consumers; thus, the reserves of data are theoretically infinite. However, the collection and transmission of data introduce energy costs to the data owners. Here, we consider that each data owner $i \in \mathcal{N}$ has predefined acquisition cost and transmission cost for its data, which are denoted as a_i and t_i, respectively. Each owner $i \in \mathcal{N}$ must set the price of its data higher than its cost. We denote the cost of unit data set by data owner i as $c_i = a_i + t_i$. In contrast, the data consumer has an acceptable maximum price, which is denoted as $p_{\max}$. Let $\mathbf{x} \triangleq (x_1, \dots, x_N)$ and $\mathbf{x}^* \triangleq (x_1^*, \dots, x_N^*)$ represent the quantity profile of purchased data, and the optimal quantity profile of purchased data, respectively. Similarly, $\mathbf{p} \triangleq (p_1, \dots, p_N)$ and $\mathbf{p}^* \triangleq (p_1^*, \dots, p_N^*)$ represent the price profile and the optimal

TABLE I
NOTATIONS

Symbol	Definition
$\mathcal{N}$	The set of data owners
N	The number of data owners
B	The purchasing willingness of data consumer
x_i	The quantity of data purchased from data owner i
x_i^*	The optimal quantity of data purchased from data owner i
x_i^{max}	The maximum quantity of data provided by data owner i
x_{min}	The minimum quantity of data required by data consumer
p_i	The unit price set by data owner i
p_i^*	The optimal price set by data owner i
a_i	The acquisition cost of unit data set by data owner i
t_i	The transmission cost of unit data set by data owner i
c_i	The cost of unit data set by data owner i
p_{max}	The maximum accepted price set by data consumer
q_i	The transmission quality of data owner i
w_i	The purchasing willingness of consumer from data owner i
y_i	The auxiliary pricing strategy of data owner i
$\mathbf{x}$	The quantity profile of purchased data
$\mathbf{x}^*$	The optimal quantity profile of purchased data
$\mathbf{p}$	The price profile of all data owners
$\mathbf{p}^*$	The optimal price profile of all data owners
$\mathbf{p}_{-i}$	The price profile of other owners except i
$\mathbf{y}$	The auxiliary pricing profile of all data owners
$\mathbf{y}^*$	The optimal auxiliary pricing profile of all data owners
$\mathbf{y}_{-i}$	The auxiliary pricing profile of other owners except i

price profile of data owners, respectively. The major notations used in this article are listed in Table I.

In our model, the consumer purchases data from the win-owner via the market-agency. The transaction is actually performed between the consumer and the market-agency, which can be modeled as a two-stage Stackelberg game [36], [37]. In this game, the market-agency acts as the leader and sets the price profile $\mathbf{p}$ for data owners in stage I. The data consumer acts as the follower and decides the quantity profile of the purchased data $\mathbf{x}$ in stage II. Fig. 4 illustrates this Stackelberg game and the optimal pricing mechanism.

Although there would be various kinds of value in terms of data products, we consider the analysis service is the primary utility of IoT data products in our data market. The classification accuracy and satisfaction rate of the data analysis service are mainly determined by the quantity of data. Therefore, we consider the quality of the model, which is defined in [8], as the quality of data from owner i

$$\rho(x_i) = \alpha_1 + \alpha_2 \ln(x_i + 1) \tag{1}$$

where α_1 and α_2 are curve fitting parameters of (1) modeled to the real-world experiments.

Besides the quantity of data, the satisfaction of the consumer is also affected by the quality of the transmission, such as data arrival rate, link quality, and communication speed. We denote the transmission quality of data owner i as q_i. The transmission power of i is related to the transmission cost. We define the signal-to-noise ratio of i as $\text{SNR}_i = \eta t_i$, where η is a predefined parameter. Based on the Shannon–Hartley theorem, the transmission quality of i is

$$q_i = B \log_2 (1 + \eta t_i) \tag{2}$$

where B is a predefined factor related to the purchasing willingness of the data consumer.

Therefore, the satisfaction function of the consumer is

$$B_{\text{sat}}(x_i, q_i) = q_i \rho(x_i) = q_i[\alpha_1 + \alpha_2 \ln(x_i + 1)]. \tag{3}$$

Accordingly, the utility of the consumer is given as its satisfaction minus the payment

$$BU_i(x_i, p_i) = q_i[\alpha_1 + \alpha_2 \ln(x_i + 1)] - x_i p_i. \tag{4}$$

The consumer decides x_i, i.e., the quality of data purchased from data owner i, to maximize its utility, forming the consumer's subgame for i which can be written as follows.

Problem 1 (Consumer's Subgame for i):

$$\begin{aligned} &\underset{x_i}{\text{maximize}} \ \ BU_i(x_i, p_i) \\ &\text{subject to} \ \ x_i \in \left[x_{\min}, x_i^{\max}\right], \forall i \in \mathcal{N}. \end{aligned} \tag{5}$$

In our optimal pricing mechanism based on whether to consider the competition between data owners, the market-agency could apply two pricing schemes: 1) IPS and 2) CPS.

B. Independent Pricing Scheme

We first consider the IPS. In this scheme, data owners (i.e., data sellers) set their pricing strategies independently without considering other owners' pricing strategies. The two-stage Stackelberg game can be divided into a series of subgames between the consumer and each seller. The utility of data owner i is given as its economic revenue minus the cost, which can be written as

$$SU_i(x_i, p_i) = x_i(p_i - c_i). \tag{6}$$

The market-agency chooses the win-owner to maximize its profit. Thus, the utility of the market-agency is

$$SU(\mathbf{x}, \mathbf{p}) = \max_{i \in \mathcal{N}} SU_i(x_i, p_i). \tag{7}$$

Therefore, the agency decides the pricing profile $\mathbf{p}$ to maximize its utility, forming the agency's subgame in IPS which can be written as follows.

Problem 2 (Agency's Subgame in IPS):

$$\begin{aligned} &\underset{\mathbf{p}}{\text{maximize}} \ \ SU(\mathbf{x}, \mathbf{p}) \\ &\text{subject to} \ \ p_i \in \left[c_i, p_{\max}\right], \forall i \in \mathcal{N} \end{aligned} \tag{8}$$

Problems 1 and 2 together form the IPS Stackelberg game. The objective of this game is to find the Stackelberg equilibrium point in which the utility of the leader is maximized as the follower adopts its best response.

Definition 1: The point $(\mathbf{x}^*, \mathbf{p}^*)$ is the Stackelberg equilibrium of the IPS if the following conditions $SU(\mathbf{x}^*, \mathbf{p}^*) \geq SU(\mathbf{x}^*, \mathbf{p})$, and $BU_i(x_i^*, p_i^*) \geq BU_i(x_i, p_i^*), \forall i \in \mathcal{N}$, are satisfied.

We use backward induction to analyze the IPS Stackelberg game. This game can be divided into a series of subgames between the data consumer and each data owner. We denote this noncooperative game as $\mathcal{G} = \{\mathcal{N}, \{p_i\}_{i \in \mathcal{N}}, \{SU_i(x_i, p_i)\}_{i \in \mathcal{N}}\}$. Data owners must set their prices via the market-agency. The subgames $\mathcal{G}$ are performed between the consumer and the market-agency for each owner $i \in \mathcal{N}$ independently. Therefore, we can solve the pricing and purchasing problem by deriving the Stackelberg equilibrium for each subgame.

1) Consumer's Purchasing Strategy in Stage II: Given the unit price of a data owner $i \in \mathcal{N}$, i.e., p_i, the consumer maximizes its utility by determining its optimal purchasing strategy x_i^*.

We derive the first- and second-order derivatives of the consumer's utility in (4) with respect to x_i, which can be written as follows:

$$\frac{\partial BU_i}{\partial x_i} = \frac{q_i \alpha_2}{1 + x_i} - p_i \tag{9}$$

and

$$\frac{\partial^2 BU_i}{\partial x_i^2} = -\frac{q_i \alpha_2}{(1 + x_i)^2} < 0. \tag{10}$$

These derivatives show that $BU_i(x_i, p_i)$ is a strictly concave function. We obtain the optimal response function of the data consumer by solving $[(\partial BU_i)/(\partial x_i)] = 0$, as follows:

$$x_i^* = \frac{q_i \alpha_2}{p_i} - 1. \tag{11}$$

2) Market-Agency's Pricing Strategies in Stage I: Based on the optimal purchasing strategy decided by the consumer for each data owner in Stage II, the data owners give their pricing strategies to maximize their utilities. Substituting (11) into (6) allows the utility of data owner i to be rewritten as

$$\begin{aligned} SU_i\left(x_i^*, p_i\right) &= x_i^*(p_i - c_i) \\ &= (q_i \alpha_2 + c_i) - \left(p_i + \frac{q_i \alpha_2 c_i}{p_i}\right). \end{aligned} \tag{12}$$

We derive the first- and second-order derivatives of $SU_i(x_i^*, p_i)$ with respect to p_i, which can be written as

$$\frac{\partial SU_i}{\partial p_i} = \frac{q_i \alpha_2 c_i}{p_i^2} - 1 \tag{13}$$

and

$$\frac{\partial^2 SU_i}{\partial p_i^2} = -2\frac{q_i \alpha_2 c_i}{p_i^3} < 0. \tag{14}$$

Therefore, $SU_i(x_i^*, p_i)$ is strictly concave with respect to p_i. Based on $[(\partial SU_i)/(\partial p_i)] = 0$, we obtain the optimal pricing strategy of i, which is given as

$$p_i^* = \sqrt{q_i \alpha_2 c_i}. \tag{15}$$

We also obtain the optimal utility of i by substituting (15) into $SU_i(x_i^*, p_i)$

$$SU_i\left(x_i^*, p_i^*\right) = \left(\sqrt{q_i \alpha_2} - \sqrt{c_i}\right)^2. \tag{16}$$

Therefore, based on (16), the data owner with maximal utility is selected as the win-owner. The utility of the market-agency under independent pricing can be obtained as

$$SU(\mathbf{x}, \mathbf{p}) = \max_{i \in \mathcal{N}} \left(\sqrt{q_i \alpha_2} - \sqrt{c_i}\right)^2. \tag{17}$$

C. Competition-Enhanced Pricing Scheme

We next consider the CPS. In this pricing scheme, the competition between data owners is considered in the leader's subgame. Each data owner sets its price to maximize its profit and to compete with other peers for the trading deal. We define the purchasing willingness of the data consumer with data owner i as $w_i = (q_i/p_i)$. There is a strong purchasing willingness if the transmission quality q_i is high or if the data owner sets a low price p_i for its data. We define the competitiveness of i as

$$\chi_i(p_i) = \frac{w_i}{\frac{1}{N}\sum_{j \in \mathcal{N}} w_j} = \frac{N q_i}{\sum_{j \in \mathcal{N}} \frac{q_j p_i}{p_j}}. \tag{18}$$

The market-agency sets its strategy, i.e., determines the win-owner and corresponding data price, by jointly considering the purchasing willingness of the consumer and the profit of the win-owner. Therefore, the utility of the market-agency for the data owner i is given as the product of competitiveness and the profit of i

$$\Pi_i\left(p_i, \mathbf{p}_{-i}, x_i\right) = \chi_i(p_i) x_i (p_i - c_i) \tag{19}$$

where $\mathbf{p}_{-i}$ is the price profile of other owners except i.

The market-agency selects the data owner with maximal utility for it as the win-owner. Therefore, the utility of the market-agency under competition-enhanced pricing is given as

$$\Pi(\mathbf{x}, \mathbf{p}) = \max_{i \in \mathcal{N}} \Pi_i\left(p_i, \mathbf{p}_{-i}.x_i\right). \tag{20}$$

Thus, in the CPS the market-agency decides the pricing profile $\mathbf{p}$ to maximize its utility defined in (20), forming the market-agency's subgame in CPS which can be written as follows.

Problem 3 (Market-Agency's Subgame in CPS):

$$\begin{aligned} &\underset{\mathbf{p}}{\text{maximize}} \quad \Pi(\mathbf{x}, \mathbf{p}) \\ &\text{subject to} \quad p_i \in \left[c_i, p_{\max}\right], \forall i \in \mathcal{N}. \end{aligned} \tag{21}$$

Problems 1 and 3 together form the CPS Stackelberg game. The objective of this game is to find the Stackelberg equilibrium in which the market-agency's utility is maximized as the follower adopts its best response.

Definition 2: The point $(\mathbf{x}^*, \mathbf{p}^*)$ is the CPS Stackelberg equilibrium if the following conditions, $\Pi(\mathbf{x}^*, \mathbf{p}^*) \geq \Pi(\mathbf{x}^*, \mathbf{p})$, and $BU_i(x_i^*, p_i^*) \geq BU_i(x_i, p_i^*)$, $\forall i \in \mathcal{N}$, are satisfied.

We use backward induction to analyze the CPS Stackelberg game. This game can also be divided into a series of CPS subgames between the data consumer and each data owner. In stage I of the CPS Stackelberg game, data owners compete with each other to maximize their own profit and competitiveness, which forms the owners' noncooperative game $\mathcal{G} = \{\mathcal{N}, \{p_i\}_{i \in \mathcal{N}}, \{\Pi_i(p_i, \mathbf{p}_{-i}, x_i)\}_{i \in \mathcal{N}}\}$. We can solve this pricing and purchasing problem by deriving the equilibrium point of each CPS subgame. Note that the consumer's subgame in CPS is the same as that in IPS. Thus, the optimal response function of the consumer in CPS can also be obtained by (11). Therefore, we only need to investigate the market-agency's pricing strategies in stage I.

Based on the optimal purchasing strategy decided by the data consumer for each data owner in stage II, the data owners give their pricing strategies to maximize their utilities. The pricing mechanism of the market-agency considers the utility of each owner and the purchasing willingness of the consumer for the data. Therefore, given the profile of the purchasing strategies $\mathbf{x}$ of the consumer, the owners compete with each other by setting their individual pricing strategies, forming the noncooperative game $\mathcal{G}$. Each owner $i \in \mathcal{N}$ gives its price p_i to the market-agency to maximize the utility of the market-agency for i, i.e., $\Pi_i(p_i, \mathbf{p}_{-i}, x_i^*)$. There should be a Nash equilibrium in the noncooperative game $\mathcal{G}$, and we next investigate the Nash equilibrium of $\mathcal{G}$.

Theorem 1: A unique Nash equilibrium exists in $\mathcal{G} = \{\mathcal{N}, \{p_i\}_{i\in\mathcal{N}}, \{\Pi_i(p_i, \mathbf{p}_{-i}, x_i^*)\}_{i\in\mathcal{N}}\}$.

Proof: We first prove the existence of the Nash equilibrium point of $\mathcal{G}$. For convenience, we define auxiliary pricing strategies as $\{y_i = (1/p_i)\}_{i\in\mathcal{N}}$. Let $\mathbf{y} \triangleq ((1/p_1), \ldots, (1/p_N))$ and $\mathbf{y}^* \triangleq ((1/p_1^*), \ldots, (1/p_N^*))$ represent the auxiliary pricing profile and the optimal auxiliary pricing profile of all data owners, respectively. Let $\mathbf{y}_{-i}$ represent the auxiliary pricing profile of other owners except i. Given the consumer's optimal purchasing strategy for owner i, i.e., x_i^*, the utility of the market-agency for i in (19) can be rewritten as

$$\begin{aligned}\Pi_i\big(y_i, \mathbf{y}_{-i}, x_i^*\big) &= \frac{N q_i y_i}{\sum_{j\in\mathcal{N}} q_j y_j}(q_i\alpha_2 y_i - 1)\left(\frac{1}{y_i} - ci\right)\\ &= -\frac{N}{\sum_{j\in\mathcal{N}} \frac{q_j}{q_i} y_j}\Big[q_i\alpha_2 c_i y_i^2 - (q_i\alpha_2 + c_i)y_i + 1\Big].\end{aligned} \tag{22}$$

Because $y_i = (1/p_i)$ can be regarded as a continuous and monotone function of p_i, the data owners' noncooperative game $\mathcal{G}$ is equivalent to the noncooperative game $\mathcal{G}' = \{\mathcal{N}, \{y_i\}_{i\in\mathcal{N}}, \{\Pi_i(y_i, \mathbf{y}_{-i}, x_i^*)\}_{i\in\mathcal{N}}\}$. Therefore, Theorem 1 can be proved if we prove that a unique Nash equilibrium exists in $\mathcal{G}'$.

The strategy space of the auxiliary pricing strategy for owner i, i.e., y_i, is defined to be $[(1/p_{\max}), (1/c_i)]$, which is a nonempty, compact subset of the Euclidean space. From (22), $\Pi_i(y_i, \mathbf{y}_{-i}, x_i^*)$ is apparently continuous in $[(1/p_{\max}), (1/c_i)]$. We take the first- and second-order derivatives of $\Pi_i(y_i, \mathbf{y}_{-i}, x_i^*)$ with respect to y_i to prove its concavity, which can be written as

$$\frac{\partial \Pi_i}{\partial y_i} = -\frac{q_i\alpha_2 c_i y_i^2 + (2q_i\alpha_2 c_i y_i - q_i\alpha_2 - c_i)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}\right) - 1}{\left(\sum_{j\in\mathcal{N}}\frac{q_j y_j}{q_i}\right)^2} \tag{23}$$

and

$$\frac{\partial^2 \Pi_i}{\partial y_i^2} = -2\frac{q_i\alpha_2 c_i\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}\right)^2 + (q_i\alpha_2 + c_i)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}\right) + 1}{\left(\sum_{j\in\mathcal{N}}\frac{q_j y_j}{q_i}\right)^3}. \tag{24}$$

We easily derive that $[(\partial^2\Pi_i)/(\partial y_i^2)] < 0$. Therefore, we can prove that $\Pi_i(y_i, \mathbf{y}_{-i}, x_i^*)$ is strictly concave with respect to y_i and that the Nash equilibrium exists in game $\mathcal{G}'$.

Then, we prove the uniqueness of Nash equilibrium in $\mathcal{G}'$. Based on $[(\partial\Pi_i)/(\partial y_i)] = 0$, we have:

$$q_i\alpha_2 c_i y_i^2 + (2q_i\alpha_i c_i y_i - q_i\alpha_2 - c_i)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}\right) - 1 = 0. \tag{25}$$

Therefore, the optimal auxiliary pricing strategy of i is:

$$y_i^* = \sqrt{\left(\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{1}{c_i}\right)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{1}{q_i\alpha_2}\right)} - \sum_{j\neq i}\frac{q_j y_j}{q_i}. \tag{26}$$

Consider y_i is subject to $[(1/p_{\max}), (1/c_i)]$, we obtain the optimal auxiliary pricing function of i as:

$$y_i^* = \mathcal{F}_i(\mathbf{y}) = \begin{cases} \frac{1}{c_i} & z_i > \frac{1}{c_i} \\ z_i & \frac{1}{p_{\max}} < z_i \le \frac{1}{c_i} \\ \frac{1}{p_{\max}} & z_i \le \frac{1}{p_{\max}} \end{cases} \tag{27}$$

where

$$z_i = \sqrt{\left(\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{1}{c_i}\right)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{1}{q_i\alpha_2}\right)} - \sum_{j\neq i}\frac{q_j y_j}{q_i}$$

is the optimal auxiliary pricing function of i.

Let $\mathbf{y}^*$ denote the Nash equilibrium of $\mathcal{G}'$. This Nash equilibrium satisfies $\mathbf{y}^* = \mathcal{F}(\mathbf{y})$, where $\mathcal{F}(\mathbf{y}) = (\mathcal{F}_1(\mathbf{y}), \ldots, \mathcal{F}_N(\mathbf{y}))$. $\mathcal{F}_i(\mathbf{y})$ is the optimal auxiliary pricing function of i, as shown in (27). The uniqueness of the Nash equilibrium in $\mathcal{G}'$ can be proven by showing that the optimal auxiliary pricing function of i is a standard function [38].

Definition 3: A function $\mathcal{F}(\mathbf{y})$ is a standard function when the following properties are guaranteed:

1) Positivity: $\mathcal{F}(\mathbf{y}) > 0$;

2) Monotonicity: if $\mathbf{y} \le \mathbf{y}'$, *then* $\mathcal{F}(\mathbf{y}) \le \mathcal{F}(\mathbf{y}')$;

3) Scalability: for all $\lambda > 1$, $\lambda\mathcal{F}(\mathbf{y}) > \mathcal{F}(\lambda\mathbf{y})$.

This section proves that $\mathcal{F}_i(\mathbf{y})$ satisfies the three properties of a standard function.

First, for the positivity, we have $\mathcal{F}_i(\mathbf{y}) > 0$ by (26), thus proving the positivity of $\mathcal{F}_i(\mathbf{x})$.

Then, we prove the monotonicity of $\mathcal{F}_i(\mathbf{y})$ with respect to $\mathbf{y}$. For $\mathbf{y}' \ge \mathbf{y}$, we have $\sum_{j\neq i}[(q_j y_j')/q_i] \ge \sum_{j\neq i}[(q_j y_j)/q_i]$. Therefore, we can prove the monotonicity of $\mathcal{F}_i(\mathbf{y})$ with respect to $\mathbf{y}$ by proving the monotonicity of $\mathcal{F}_i(\mathbf{y})$ with respect to $\sum_{j\neq i}[(q_j y_j)/q_i]$ instead.

By differentiating (26) with respect to $\sum_{j\neq i}[(q_j y_j)/q_i]$, we obtain the following:

$$\frac{\partial\mathcal{F}_i(\mathbf{y})}{\partial\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}\right)} = \frac{\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{q_i\alpha_2 + c_i}{2q_i\alpha_2 c_i}}{\sqrt{\left(\sum_{j\neq i}\frac{q_j y_j}{q_i} + \frac{q_i\alpha_2 + c_i}{2q_i\alpha_2 c_i}\right)^2 - \left(\frac{q_i\alpha_2 - c_i}{2q_i\alpha_2 c_i}\right)^2}} - 1 > 0. \tag{28}$$

Based on (28), we have $\mathcal{F}_i(\mathbf{y}') - \mathcal{F}_i(\mathbf{y}) \ge 0$. Therefore, the optimal auxiliary pricing function $\mathcal{F}_i(\mathbf{y})$ is always monotonically increasing with $\mathbf{y}$.

Lastly, for scalability, we must prove that $\lambda\mathcal{F}_i(\mathbf{y})-\mathcal{F}_i(\lambda\mathbf{y}) > 0$, for all $\lambda > 1$. The steps of proving the positivity of $\lambda\mathcal{F}_i(\mathbf{y})-\mathcal{F}_i(\lambda\mathbf{y})$ are shown in (29), at the bottom of this page.

Thus far, we have proven that $\mathcal{F}_i(\mathbf{y})$ satisfies the three properties of a standard function. Therefore, a unique Nash equilibrium exists in $\mathcal{G}'$. The proof is completed. ■

The optimal pricing function of data owner i in $\mathcal{G}$ can be obtained by substituting $y_i = (1/p_i)$ into (27), which we can rewrite as

$$p_i^* = \mathcal{F}'_i(\mathbf{p}) = \begin{cases} p_{\max} & z'_i > p_{\max} \\ z'_i & c_i < z'_i \le p_{\max} \\ c_i & z'_i \le c_i \end{cases} \tag{30}$$

where

$$z'_i = \frac{1}{\left(\sqrt{\left(\sum_{j\neq i}\frac{q_j}{q_i p_j}+\frac{1}{c_i}\right)\left(\sum_{j\neq i}\frac{q_j}{q_i p_j}+\frac{1}{q_i\alpha_2}\right)}-\sum_{j\neq i}\frac{q_j}{q_i p_j}\right)}.$$

The utility maximization of the market-agency defined in (20) is a convex optimization problem. To obtain the maximum utility of the market-agency, we apply the low-complexity gradient-based searching algorithm to achieve the approximate Nash equilibrium point $\mathbf{y}^*$ of $\mathcal{G}'$. Furthermore, we can obtain the optimal price profile of all data owners, i.e., $\mathbf{p}^* \triangleq ((1/y_i^*), \ldots, (1/y_i^*))$. The market-agency chooses the optimal data owner as the win-owner for maximizing its utility.

In particular, we adopt Algorithm 1 to obtain the unique Nash equilibrium and solve the optimal pricing problem. Using this gradient-based algorithm, the leader's optimal pricing strategy can be obtained. The follower can also obtain its optimal purchasing strategy.

Algorithm 1 Iterative Gradient Algorithm to Find the Stackelberg Equilibrium of the CPS

Initialization:
Set the initial input $\mathbf{p}^{(0)} = \{p_i^{(0)}\}_{i\in\mathcal{N}}$ and $\mathbf{x}^{(0)} = \{x_i^{(0)}\}_{i\in\mathcal{N}}$, where $p_i^{(0)} \in [a_i + t_i, p^{max}]$ and $x_i^{(0)} \in [x_{min}, x_i^{max}]$, $\mathbf{y}^{(0)} = \{\frac{1}{p_i^{(0)}}\}_{i\in\mathcal{N}}$, $1 \leftarrow t$, $1 \leftarrow \tau$, a precision threshold $\varepsilon \ll 1$;
while $(\tau > \varepsilon)$ **do**
  for all data owners $i \in \mathcal{N}$ **do**
    The consumer decides the quantity of its purchasing data $x_i^{(t)}$ from data owner i based on (11);
$$x_i^{(t)} = \frac{q_i\alpha_2}{p_i^{(t-1)}} - 1 = q_i\alpha_2 y_i^{(t-1)} - 1.$$
  end for
  for all data owners $i \in \mathcal{N}$ **do**
    The market-agency updates its auxiliary pricing strategy profile for data owners using a gradient-assisted search algorithm:
$$y_i^{(t)} = y_i^{(t-1)} + \mu\frac{\partial \Pi_i\left(y_i^{(t-1)}, \mathbf{y}_{-i}^{(t-1)}, x_i^{(t)}\right)}{\partial y_i},$$
    where μ is the step size of the price update. $y_i^{(t)}$ is subject to $[\frac{1}{p_{max}}, \frac{1}{a_i+t_i}]$.
  end for
  $\tau \leftarrow \frac{\sum_{i\in\mathcal{N}}\|y_i^{(t)}-y_i^{(t-1)}\|}{\sum_{i\in\mathcal{N}}\|y_i^{(t-1)}\|}$,
  $t \leftarrow t+1$,
end while
Obtain the pricing profit $\mathbf{p}^{(t)} = \{\frac{1}{y_i^{(t)}}\}_{i\in\mathcal{N}}$.
Search the optimal utility of the market-agency:
$\Pi(\mathbf{x}, \mathbf{p}^{(t)}) = \Pi_l(p_l^{(t)}, \mathbf{p}_{-l}^{(t)}, x_l^{(t)})$, where $l = \arg\max_i \Pi_i(p_i^{(t)}, \mathbf{p}_{-i}^{(t)}, x_i^{(t)})$.
Output: l, $p_l^{(t)}$, $x_l^{(t)}$.

V. System Evaluation and Performance Analysis

In this section, we first evaluate the security, trustworthiness, efficiency, scalability, profitability, and stability of the proposed blockchain-enhanced IoT data market framework. Then, we conduct extensive numerical simulations to analyze the performance of the proposed optimal pricing mechanisms, including both IPS and CPS. We compare the performance of CPS with that of IPS and focus on investigating the advantages of CPS. In particular, the benefits brought by introducing competition between data owners are examined.

A. Evaluation of Blockchain-Enhanced IoT Data Market

1) Security: The blockchain-enhanced IoT data market framework guarantees trading security and privacy protection through standard cryptographic primitives. The decentralization and digitally signed transactions ensure that no adversary

$$\begin{aligned}\lambda\mathcal{F}_i(\mathbf{y}) - \mathcal{F}_i(\lambda\mathbf{y}) &= \lambda\sqrt{\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{1}{c_i}\right)\left(\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{1}{q_i\alpha_2}\right)} - \lambda\sum_{j\neq i}\frac{q_j y_j}{q_i} - \sqrt{\left(\sum_{j\neq i}\frac{\lambda q_j y_j}{q_i}+\frac{1}{c_i}\right)\left(\sum_{j\neq i}\frac{\lambda q_j y_j}{q_i}+\frac{1}{q_i\alpha_2}\right)} \\ &\quad + \sum_{j\neq i}\frac{\lambda q_j y_j}{q_i} \\ &= \sqrt{\left(\lambda\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{\lambda}{c_i}\right)\left(\lambda\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{\lambda}{q_i\alpha_2}\right)} - \sqrt{\left(\lambda\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{1}{c_i}\right)\left(\lambda\sum_{j\neq i}\frac{q_j y_j}{q_i}+\frac{1}{q_i\alpha_2}\right)} > 0, \forall\lambda > 1\end{aligned} \tag{29}$$

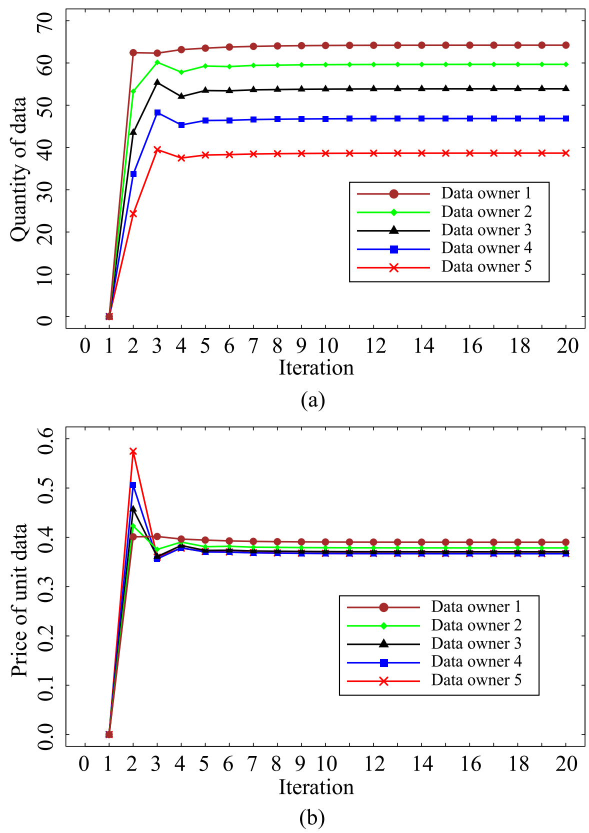


Fig. 5. Convergence of the (a) quantity of data and (b) price of data.

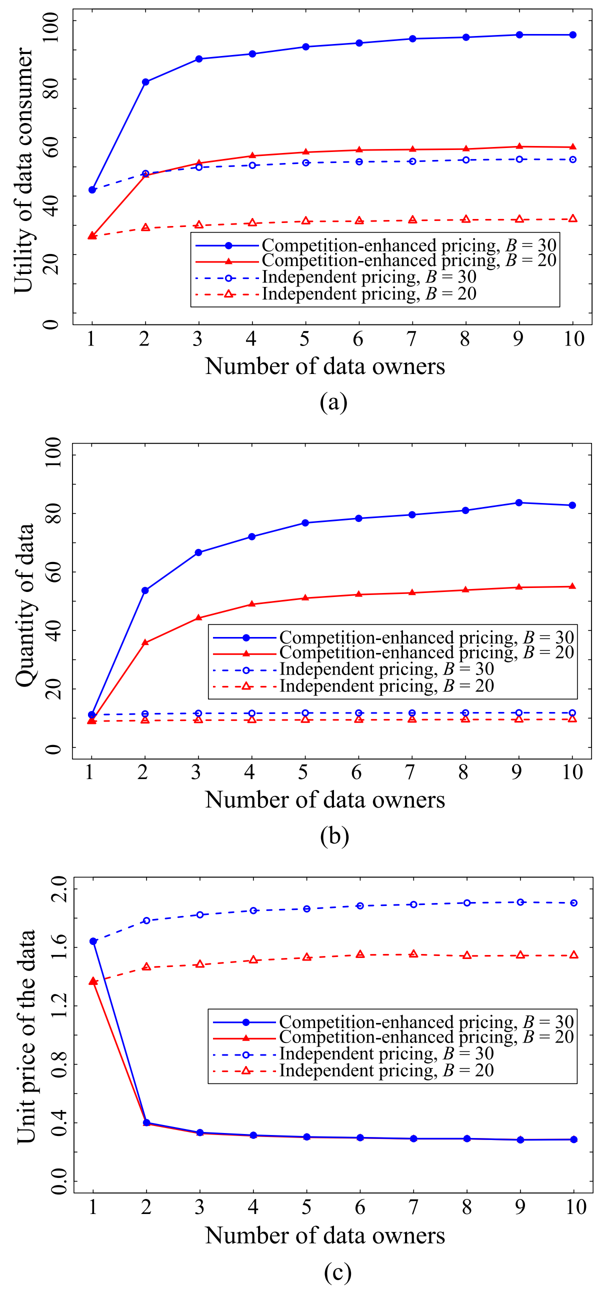


Fig. 6. Impacts brought by competition between data owners on the (a) utility of the consumer, (b) quantity of purchased data, and (c) price of data.

can pose as a user or corrupt the network [39]. A public history of transactions is recorded in the blocks and shared by all blockchain peers, ensuring the validity of transactions and preventing invalid trading such as double-spending.

2) Trustworthiness: In the blockchain-enhanced market model, decentralized consensus and public ledger of the blockchain provide a trustworthy environment for interactions of IoT devices. The authorized market-agency which is predefined in smart contract helps schedule the trading and sets pricing for data owners. Therefore, there is actually no private data exposed to third parties during the trading.

3) Efficiency: With the cooperation of edge/cloud computing and blockchain, data trading between IoT devices can be performed with low latency and energy consumption. The game-theory-based smart pricing mechanism enables data pricing in a low-complexity manner. The negotiation and pricing process can be implemented automatically and efficiently to satisfy real-time requests.

4) Scalability: In our blockchain-enhanced IoT data market framework, we exploit edge servers to improve the scalability. Moreover, the proposed data market framework enables a proxy-based trading model in which proxy peers are assigned to devices/servers to conduct the blockchain-related tasks. Such a consortium-blockchain trading model further enhances the scalability of the system.

5) Profitability: The intermediary market agent acts as the leader in the Stackelberg game based optimal pricing mechanism. Therefore, it can decide the win-owner and price strategy while considering the profit of both win-owner and consumer. In particular, with the competition between data owners, the consumer is encouraged to purchase data at a competitive price.

6) Stability: The market stability is also guaranteed in the proposed data market. First, the blockchain-based framework ensures the trust and credit of market participants. Transaction fraud and default are recorded in the blockchain and inform all peers in the market. Second, CPS enables consumer to buy data at a low price, thus enhancing economic vitality and preventing a currency crisis.

B. Performance Analysis of the Optimal Pricing Mechanism

We study the performance of the proposed smart data pricing mechanism by numerical simulations. To illustrate the

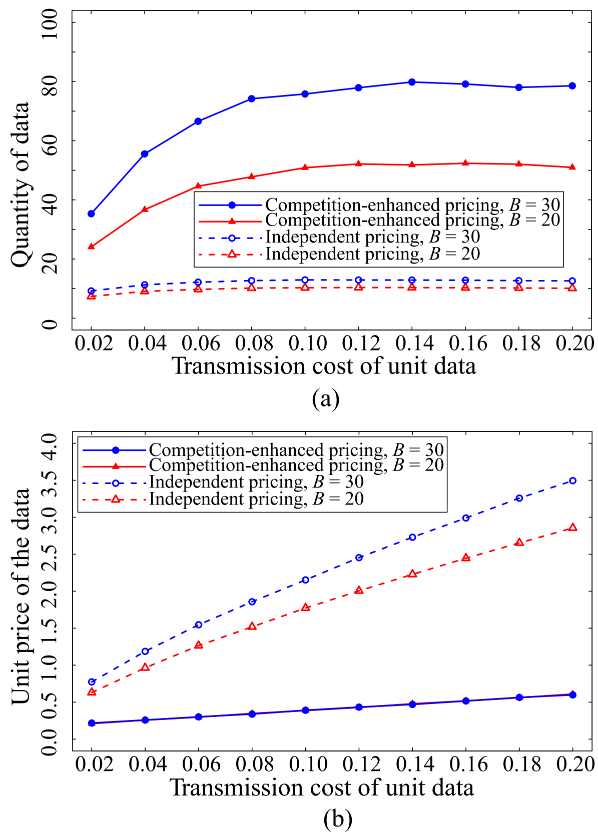


Fig. 7. Impacts brought by transmission cost on the (a) quantity of data and (b) price of data.

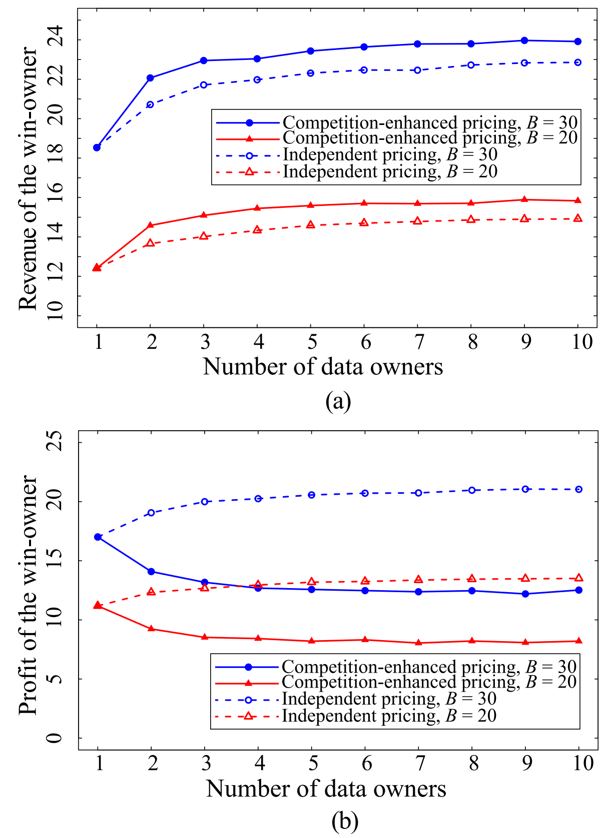


Fig. 8. Impacts brought by competition between data owners on the (a) revenue of the win-owner and (b) profit of the win-owner.

impacts of different parameters on performance, we consider a group of N data owners selling data when a consumer requests some data.

The default parameter values of the numerical experiments are set as follows: $\alpha_1 = 0.5$, $\alpha_2 = 1$, $p_{\max} = 5$, $x_{\min} = 5$, $N = 5$, and $\{x_i^{\max} = 100\}_{i\in\mathcal{N}}$. We assume that $\{a_i\}_{i\in\mathcal{N}}$ and $\{t_i\}_{i\in\mathcal{N}}$ follow uniform distribution among [0.05, 0.1] and [0.04, 0.08], respectively. We repeat all numerical simulations 500 times and use the median values as results. The parameter B indicates the purchasing willingness of the consumer for the data, so we also investigate the impacts of B on performance in the analyses. We analyze and compare the performance of game theory-based smart pricing mechanisms, i.e., CPS and IPS. We focus on investigating the benefits brought by competition between data owners in CPS.

1) Efficiency of CPS: The solution of the pricing problem in CPS is the equilibrium point of the noncooperative Stackelberg game. We obtain an approximate equilibrium point by using the iterative gradient algorithm (Algorithm 1). Fig. 5 shows the convergence of price and quantity of data, i.e., p_i and x_i, with iterations by Algorithm 1. In the scenario of Fig. 5, there are five data owners with different cost c_i. We can find that the price and quantity of data from all the five data owners converge quickly with Algorithm 1. The equilibrium point can be obtained within ten iterations. These results show the efficiency of the proposed algorithm. The results also show that different data owners can achieve similar optimal prices, indicating that competition between data owners enables stable data prices in the market.

2) Investigation of the Utility of the Data Consumer: We next investigate the utility of the data consumer under CPS and IPS. We focus on the impacts brought by competition between data owners and the transmission cost. The number of data owners, i.e., N, indicates the degree of competition in CPS. We compare the performance of CPS with that of IPS in which there is no competition between data owners. Note that these two pricing schemes are the same when $N = 1$.

Fig. 6(a) shows the general utility of the data consumer, which is defined in (4), with N. We see that the consumer's utility increases with the aggravation of market competition. The results show that under CPS, the consumer can obtain more utility from the win-owner. The results also show that the data consumer obtains more utility when the purchasing willingness B is larger. Fig. 6(b) and (c) shows the impacts of N on the quantity and price of data and indicate that the consumer can purchase more data at a lower price under CPS. Under IPS, the quantity of data is nearly unchanged with N. But under CPS, the quantity of data increases with N and the price of data decreases with N. When $N > 5$, the market saturation makes the price tend to be stable. This result shows that under CPS, competition between owners produces more benefits to the consumer, allowing consumers to purchase data at a lower price. Furthermore, the results also show that the data consumer purchases significantly more data when B increases

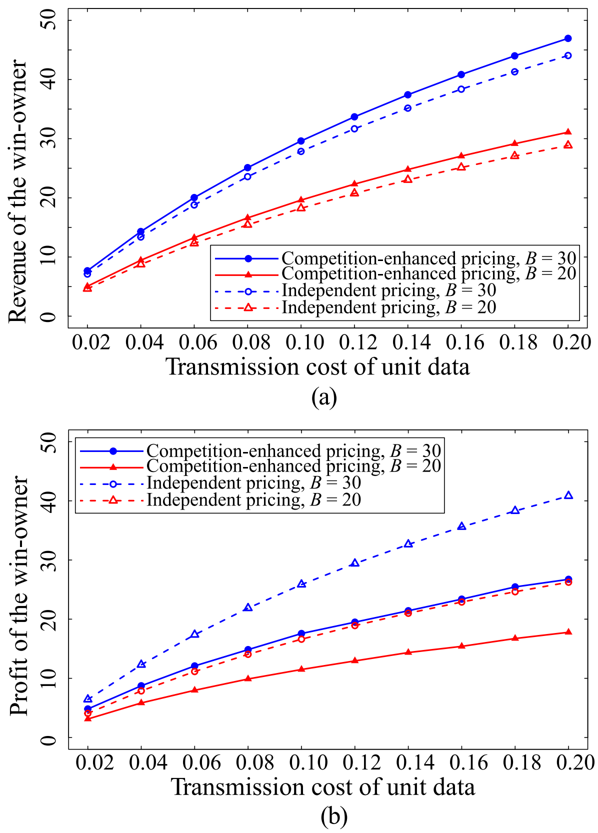


Fig. 9. Impacts brought by transmission cost on the (a) revenue of the win-owner and (b) profit of the win-owner.

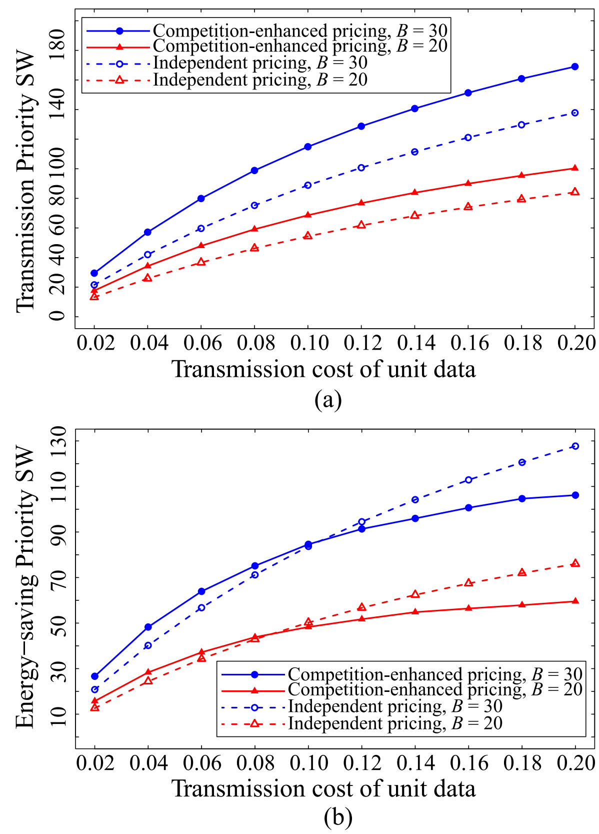


Fig. 10. Impacts brought by transmission cost on (a) transmission priority SW and (b) energy-saving priority SW.

under CPS than under IPS, indicating that the consumer's purchasing willingness can be better satisfied under CPS.

We next investigate the impacts brought by the transmission cost of data owners. Transmission cost influences the expenditure of the win-owner and the quality of its service. Here, we let all data owners set a uniform transmission cost t. Fig. 7 shows the impacts of t on the price and quantity of the data and indicates that the consumer can purchase more data at a lower price under CPS than under IPS. This difference arises because market competition hinders the increase in the price of data under CPS when t is increased. In addition, under CPS, the quantity of trading data increases as t is increased and tends to be stable because the increased transmission quality encourages the consumer to purchase more data. When the transmission quality is sufficiently high, the increased t reduces the consumer's purchasing willingness. Under IPS, the quantity of trading data is almost unchanged as t increases. The results also show that the price of data is not affected by B under CPS, demonstrating that data owners under CPS set their prices mainly based on cost and competition.

3) Investigation of the Profitability of the Win-Owner: We next investigate the profitability of the win-owner under CPS and IPS. We use economic revenue and profit to measure the profitability of the win-owner. The revenue of the win-owner is defined as the payment from the consumer, and the profit of the win-owner is calculated by (6).

Fig. 8 shows the impact of N on the revenue and profit of the win-owner and indicates that the revenue of the win-owner increases with N under both IPS and CPS, and the win-owner obtains more revenue under CPS. Fig. 8 also shows that the win-owner obtains less economic profit under CPS than under IPS. This difference arises because competition makes all owners reduce their profits under CPS. The win-owner can sell more data, thus obtaining more revenue even though the price is lower. Furthermore, we note that the win-owner can obtain more revenue and profit when B increases, indicating that the data consumer is inclined to obtain more data when its purchasing willingness is high.

Fig. 9 shows the impact of t on the revenue and profit of the win-owner and indicates that the profit of the win-owner is increased with t, meaning that providing a data service with the maximal quality is the optimal transmission strategy of owners. The results also show that the win-owner can obtain higher profit under CPS than under IPS even if the price is lower. This is due to the fact that under CPS the consumer is willing to purchase more data at a lower price, making the win-owner gain a higher profit.

4) Investigation of Social Welfare: Social welfare (SW) is the global evaluation of the data market, influencing the system performance and stability. Here, we define the SW function of the consumer and the win-owner as follows:

$$\mathrm{SW} = q[\alpha_1 + \alpha_2 \ln(x+1)] - \gamma x t \tag{31}$$

in which q, x, and t are the transmission quality, the quantity of data, and the transmission cost of the win-owner, respectively. γ is a parameter indicating the impact of energy consumption on SW. Here, we consider two types of SW, energy-saving priority SW and transmission priority SW, in which we set $\gamma = 4$ and $\gamma = 8$ in (31), respectively. Fig. 10 shows the impacts of t on these two types of SW, and both are seen to increase as t increases. This relationship arises because the increased t improves the transmission quality and counteracts the influence brought by increased energy consumption, thus improving SW. The results also show that transmission priority SW under CPS is higher than under IPS. Energy-saving priority SW under CPS is higher than under IPS when $t > 0.1$. This demonstrates that when the transmission cost is low, the system can save more energy under CPS and thus perform in a more stable manner.

VI. Conclusion

In this article, we have presented a blockchain-enhanced data market framework and an optimal pricing mechanism for IoT. To address the challenges of security and efficiency, we have proposed an edge/cloud-computing-assisted, blockchain-enhanced IoT data market framework. In particular, we designed an optimal pricing mechanism to support efficient trading in the IoT data market by using a game theory based model. We formulated a pricing and purchasing problem between the data consumer and the market-agency that sets prices for data owners. We adopted a two-stage Stackelberg game to jointly maximize the profits of the data consumer and the market-agency, and proposed a CPS that considers the competition between data owners. Through backward induction, we investigated the optimal purchasing strategy of the consumer in the second stage and the pricing strategies of the market-agency in the first stage. We prove that the Stackelberg equilibrium is derived analytically. We conducted numerical simulations to evaluate the performance of the proposed pricing scheme, which show the effectiveness and efficiency of the CPS compared with the IPS. With competition-enhanced pricing, the consumer can purchase more data at a lower price and the win-owner can also efficiently gain more economic profit.

References

[1] F. Liang, W. Yu, D. An, Q. Yang, X. Fu, and W. Zhao, "A survey on big data market: Pricing, trading and protection," *IEEE Access*, vol. 6, pp. 15132–15154, 2018.

[2] D. Niyato, M. A. Alsheikh, P. Wang, D. I. Kim, and Z. Han, "Market model and optimal pricing scheme of big data and Internet of Things (IoT)," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 1–6.

[3] K. Mišura and M. Žagar, "Data marketplace for Internet of Things," in *Proc. Int. Conf. Smart Syst. Technol. (SST)*, Osijek, Croatia, 2016, pp. 255–260.

[4] C. Perera, R. Ranjan, and L. Wang, "End-to-end privacy for open big data markets," *IEEE Cloud Comput.*, vol. 2, no. 4, pp. 44–53, Jul./Aug. 2015.

[5] L. G. Pitta and M. Endler, "Market design for IoT data and services the emergent 21st century commodities," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Natal, Brazil, 2018, pp. 410–415.

[6] P. V. Desai, "A survey on big data applications and challenges," in *Proc. 2nd Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, Coimbatore, India, 2018, pp. 737–740.

[7] X. Cao, Y. Chen, and K. J. R. Liu, "Data trading with multiple owners, collectors, and users: An iterative auction mechanism," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 2, pp. 268–281, Jun. 2017.

[8] Y. Jiao, P. Wang, S. Feng, and D. Niyato, "Profit maximization mechanism and data management for data analytics services," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2001–2014, Jun. 2018.

[9] M. Liu, F. R. Yu, Y. Teng, V. C. M. Leung, and M. Song, "Joint computation offloading and content caching for wireless blockchain networks," in *Proc. IEEE INFOCOM Workshops*, Honolulu, HI, USA, 2018, pp. 517–522.

[10] K. Suankaewmanee, D. T. Hoang, D. Niyato, S. Sawadsitang, P. Wang, and Z. Han, "Performance analysis and application of mobile blockchain," in *Proc. Int. Conf. Comput. Netw. Commun. (ICNC)*, Maui, HI, USA, Mar. 2018, pp. 642–646.

[11] O. Novo, "Blockchain meets IoT: An architecture for scalable access management in IoT," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1184–1195, Apr. 2018.

[12] R. Li, T. Song, B. Mei, H. Li, X. Cheng, and L. Sun, "Blockchain for large-scale Internet of Things data storage and protection," *IEEE Trans. Services Comput.*, to be published. doi: 10.1109/TSC.2018.2853167.

[13] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1495–1505, Apr. 2019.

[14] P. K. Sharma, M.-Y. Chen, and J. H. Park, "A software defined fog node based distributed blockchain cloud architecture for IoT," *IEEE Access*, vol. 6, pp. 115–124, 2017.

[15] C. Niu, Z. Zheng, F. Wu, X. Gao, and G. Chen, "Achieving data truthfulness and privacy preservation in data markets," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 1, pp. 105–119, Jan. 2019.

[16] V. Gatteschi, F. Lamberti, C. Demartini, C. Pranteda, and V. Santamaria, "To blockchain or not to blockchain: That is the question," *IT Prof.*, vol. 20, no. 2, pp. 62–74, Mar./Apr. 2018.

[17] P. Urien, "Blockchain IoT (BIoT): A new direction for solving Internet of Things security and trust issues," in *Proc. 3rd Cloudification Internet Things (CIoT)*, Paris, France, 2018, pp. 1–4.

[18] G. Perboli, S. Musso, and M. Rosano, "Blockchain in logistics and supply chain: A lean approach for designing real-world use cases," *IEEE Access*, vol. 6, pp. 62018–62028, 2018.

[19] K. Shuaib, J. A. Abdella, F. Sallabi, and M. Abdel-Hafez, "Using blockchains to secure distributed energy exchange," in *Proc. 5th Int. Conf. Control Decis. Inf. Technol. (CoDIT)*, Thessaloniki, Greece, 2018, pp. 622–627.

[20] L. Xue, Y. Teng, Z. Zhang, J. Li, K. Wang, and Q. Huang, "Blockchain technology for electricity market in microgrid," in *Proc. 2nd Int. Conf. Power Renew. Energy (ICPRE)*, Chengdu, China, 2017, pp. 704–708.

[21] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain, "Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3154–3164, Dec. 2017.

[22] Z. Zheng, Y. Peng, F. Wu, S. Tang, and G. Chen, "Trading data in the crowd: Profit-driven data acquisition for mobile crowdsensing," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 2, pp. 486–501, Feb. 2017.

[23] C.-T. Chen, A.-P. Chen, and S.-H. Huang, "Cloning strategies from trading records using agent-based reinforcement learning algorithm," in *Proc. IEEE Int. Conf. Agents (ICA)*, Singapore, 2018, pp. 34–37.

[24] J. Wang, C. Jiang, Z. Bie, T. Q. S. Quek, and Y. Ren, "Mobile data transactions in device-to-device communication networks: Pricing and auction," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 300–303, Jun. 2016.

[25] D. Niyato, D. T. Hoang, N. C. Luong, P. Wang, D. I. Kim, and Z. Han, "Smart data pricing models for the Internet of Things: A bundling strategy approach," *IEEE Netw.*, vol. 30, no. 2, pp. 18–25, Mar./Apr. 2016.

[26] F. Xu, F. Yang, C. Zhao, and C. Fang, "Edge computing and caching based blockchain IoT network," in *Proc. 1st IEEE Int. Conf. Hot Inf. Centric Netw. (HotICN)*, Shenzhen, China, 2018, pp. 238–239.

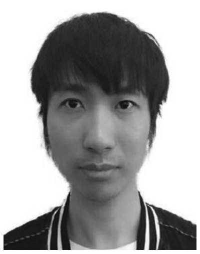
[27] Y. Jiao, P. Wang, D. Niyato, and K. Suankaewmanee, "Auction mechanisms in cloud/fog computing resource allocation for public blockchain networks," *IEEE Trans. Parallel Distrib. Syst.*, to be published. doi: 10.1109/TPDS.2019.2900238.

[28] H. G. Do and W. K. Ng, "Blockchain-based system for secure data storage with private keyword search," in *Proc. IEEE World Congr. Services (SERVICES)*, Honolulu, HI, USA, 2017, pp. 90–93.

[29] W. Chen *et al.*, "Cooperative and distributed computation offloading for blockchain-empowered industrial Internet of Things," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2019.2918296.

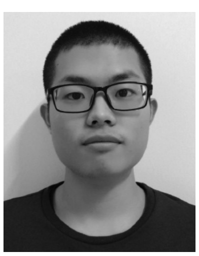
[30] Z. Li, Z. Yang, S. Xie, W. Chen, and K. Liu, "Credit-based payments for fast computing resource trading in edge-assisted Internet of Things," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2019.2908861.

[31] N. Seydoux, K. Drira, N. Hernandez, and T. Monteil, "A distributed scalable approach for rule processing: Computing in the fog for the SWoT," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. (WI)*, Santiago, Chile, 2018, pp. 112–119.
[32] W. Chen, Y. Yaguchi, K. Naruse, Y. Watanobe, and K. Nakamura, "QoS-aware robotic streaming workflow allocation in cloud robotics systems," *IEEE Trans. Services Comput.*, to be published. doi: 10.1109/TSC.2018.2803826.
[33] Z. Li, Z. Yang, and S. Xie, "Computing resource trading for edge-cloud-assisted Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3661–3669, Jun. 2019. doi: 10.1109/TII.2019.2897364.
[34] S. Lamba and M. Sharma, "An efficient elliptic curve digital signature algorithm (ECDSA)," in *Proc. Int. Conf. Mach. Intell. Res. Adv.*, Katra, India, 2013, pp. 179–183.
[35] M. Y. Jung and J. W. Jang, "Data management and searching system and method to provide increased security for IoT platform," in *Proc. Int. Conf. Inf. Commun. Technol. Convergence (ICTC)*, Jeju, South Korea, 2017, pp. 873–878.
[36] G. El Rahi, S. R. Etesami, W. Saad, N. B. Mandayam, and H. V. Poor, "Managing price uncertainty in prosumer-centric energy trading: A prospect-theoretic Stackelberg game approach," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 702–713, Jan. 2019.
[37] S. Kim, "A novel Bitcoin mining scheme based on the multi-leader multi-follower Stackelberg game model," *IEEE Access*, vol. 6, pp. 48902–48912, 2018.
[38] Z. Han, D. Niyato, W. Saad, T. Baar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[39] G. Zyskind, O. Nathan, and A. S. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," in *Proc. IEEE Security Privacy Workshops*, San Jose, CA, USA, 2015, pp. 180–184.



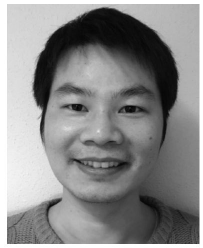
Kang Liu (M'19) received the B.Sc. degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2008, the M.Sc. degree from the College of Physics and Technology, Central China Normal University, Wuhan, in 2013, and the Ph.D. degree from the Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma, Japan.

He is currently a Post-Doctoral Fellow with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His current research interests include game theory, Internet of Things, and blockchain technology.



Xiaoyu Qiu is currently pursuing the B.S. degree with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China.

His current research interests include edge/cloud computing, cloud robotics, computation offloading, and artificial intelligence in edge/cloud computing.



Wuhui Chen (M'15) received the bachelor's degree from the Software College, Northeast University, Shenyang, China, in 2008, and the master's degree from the School of Computer Science and Engineering, University of Aizu, Aizuwakamatsu, Japan, in 2011.

He is currently an Associate Professor with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His current research interests include signal processing, service computing, social computing, big data mining, and visualization.

Xu Chen (M'13) received the Ph.D. degree in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2012.

He is a Full Professor with Sun Yat-sen University, Guangzhou, China, where he is the Vice Director of the National and Local Joint Engineering Laboratory of Digital Home Interactive Applications. His current research interests include mobile computing, cloud computing, game theory, and resource allocation.

Zibin Zheng (SM'16) received the Ph.D. degree from the Chinese University of Hong Kong, Hong Kong, in 2012.

He is a Professor with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His current research interests include service computing and cloud computing.

Prof. Zheng was a recipient of the Outstanding Ph.D. Thesis Award of the Chinese University of Hong Kong in 2012, the ACM SIGSOFT Distinguished Paper Award at International Conference on Software Engineering 2010, the Best Student Paper Award at International Conference on Web Services (ICWS) 2010, and the IBM Ph.D. Fellowship Award in 2010. He has served as a PC Member of the IEEE Cloud, ICWS, International Conference on Service Computing, International Conference on Service-Oriented Computing, and International Conference on Service-Oriented System Engineering.