WILEY | Hindawi

Research Article

A Near-Optimal Truthful Online Auction for Efficient Crowdsourced Data Trading with Dynamic Data Owners and Dynamic Data Requests

Zhenni Feng [ID],[1] Chen Zhang,[1] Junchang Chen,[1] and Tong Liu[2]

[1]*Donghua University, Shanghai, China*
[2]*Shanghai University, Shanghai, China*

Correspondence should be addressed to Zhenni Feng; fzn@dhu.edu.cn

Data is an extremely important asset in a modern scientific and commercial society. The life force behind powerful artificial intelligence (AI) or machine learning (ML) algorithms is data, especially lots of data, which makes *data trading* significantly essential to unlocking the power of AI or ML. Data owners who offer crowdsourced data and data consumers who request data blocks negotiate with each other to make an agreement on data assignment and trading prices via a data trading platform; consequently, both sides gain profit from the process of data trading. A great many existing studies have investigated various kinds of data sharing or trading as well as protecting data privacy or constructing a decentralized data trading platform due to mistrust issues. However, existing studies neglect an important characteristic, i.e., dynamics of both data owners and data requests in trading crowdsourced data collected by IoT devices. To this end, we first construct an auction-based model to formulate the data trading process and then propose a near-optimal online data trading algorithm that not only resolves the problem of matching dynamic data owners and randomly generated data requests but also determines the data trading price of each data block. The proposed algorithm achieves several good properties, such as a constant competitive ratio for near-optimal social efficiency, incentive compatibility, and individual rationality of participants, via rigorous theoretical analysis and extensive simulations. We further design a decentralized data trading platform in order to construct a practical data trading process incorporating the proposed data trading algorithm.

1. Introduction

Data is an extremely import asset in modern scientific and commercial society. As predicted by IDC, there will be 55.7 billion connected devices worldwide by 2025, 75% of which would be connected to an IoT platform [1]. Furthermore, data generated by these IoT devices are estimated to be 73.1 ZB. Then, most of the data are collected from video surveillance and industrial IoT applications. Almost all companies are aggressively turning to AI or ML technology to gain competitive advantages from valuable data. Undoubtedly, data is the life force behind these AI or ML algorithms, especially a vast amount of data. Consequently, *data trading*, as a convenient and promising method of sharing data, is significantly essential to unlocking the power of AI or ML.

Data trading is different from the general concept of data sharing. In mobile crowdsensing systems, workers usually share their sensing data and get reward in return [2–4]. Application users must disclose their data records, e.g., web browsing and online shopping orders, to application service providers in an unconscious or forced manner in order to get accessible to their services. The concept of sharing data between application users and application service providers is expanded to data trading which allows data owners and data consumers to proactively decide whether to participant in and to further specify what kinds of data they want or how much they expect to get in return for disclose their data.

Many recent studies [5–9] have paid attention to propose data trading approaches. Various kinds of datasets,

such as raw data samples, range counts, and aggregate statistic results, are traded between data owners and data consumers. To negotiate the data trading process between them, data trading platforms or data brokers are usually introduced to support transmission of data trading messages or traded data. To protect data privacy, some recent studies employ encryption algorithms and then to disclose encrypted data to data consumers; other studies introduce privacy-preserving schemes such as differential privacy or its variations to control the level of data privacy. Some other existing studies [10–15] consider participants or the data platform are untrustworthy; they propose decentralized data trading platforms based on the blockchain technology. Almost all existing studies, however, ignore an important observation of dynamics for both data owners and data requests in IoT data trading, where crowdsourced data collected by diverse smart IoT devices are traded. A data owner is not always available to provide data because a smart device is sometimes occupied by its owner or cannot provide data service due to resource constraints, e.g., out of battery and intermittent connection. Meanwhile, every data requests are generated according to a data consumer's demand, which is unknown in advance. Therefore, we regard both data owners and data requests have dynamic properties.

In the paper, we model dynamic properties of both data owners and data requests in the scenario of IoT data trading. For example, when a data request which requests the real-time noise level located in a block is submitted, a smartphone can serve as a data collector and a data owner if the owner of the smartphone passed by the block then. Consequently, we claim that smart devices (i.e., data owners) are intermittent and only available to trade their data during a specific time period, which is called *active time*, because of limited resource, mobility, or occupied by device owners. Furthermore, data requests are randomly generated by data consumers and then submitted to a data trading platform which runs on an edge server or a cloud server.

There are several main technical challenges to solve the crowdsourced data trading problem. *Firstly*, data owners are allowed to dynamically join in or leave the data trading process and data requests are randomly generated according to the demand of applications. Such uncertain and unpredictable data requests make the data trading process rather complicated. *Secondly*, it is very difficult for the data trading platform to make an efficient matching between data owners and data requests, because both valuation of data blocks and active time are private information of data owners. *Finally*, rational and strategic data owners are not willing to offer their valuable data or report their private information truthfully except with suitable compensation.

To this end, we propose a truthful online auction-based data trading algorithm containing two key components, which resolves two subproblems of how to match dynamic data owners with randomly generated data requests and how to determine trading prices of data blocks. On the one hand, an online auction model is constructed to formulate dynamic data trading process, and an efficient online matching algorithm based on a greedy scheme is further proposed to achieve near-optimal system efficiency with a constant

competitive ratio of 1/2. On the other hand, the trading price of each data block is computed according to a *critical value*, which is the highest bidding price that a data owner would win a bid. Both rigorous theoretical analysis and extensive simulations demonstrate desirable properties of our proposed online data trading algorithm, e.g., *individual rationality*, *incentive compatibility*, and *near-optimality* on system efficiency. Finally, we design a decentralized data trading platform based on the blockchain technology which incorporates the proposed truthful online auction-based data trading algorithm as a key component. Based on the decentralized data trading platform, we further design a whole process of data trading in order to avoiding misbehavior of participants, such as refusing to offer payment.

Major technical contributions in this paper are summarized as follows:

(i) It is the first work, to the best of our knowledge, which takes account of dynamic behaviors of both data owners and data requests in the problem formulation of the data trading process

(ii) We propose a truthful online auction-based data trading algorithm which not only determines the matching rule with incomplete information but also computes proper data trading prices between data owners and data consumers

(iii) We have demonstrated that the proposed algorithm achieves several good properties via rigorous theoretical analysis and extensive simulations

(iv) We design a decentralized data trading platform in order to construct a practical data trading process incorporating the proposed data trading algorithm

The rest of the paper is organized as follows. System model and problem formulation of the crowdsourced data trading problem are presented in Section 2. Then, the proposed online data trading algorithm is discussed in Section 3 along with rigorous theoretical analysis. In Section 4, we further design a decentralized data trading platform based on the blockchain technology. Section 5 provides extensive simulations and numerical results to demonstrate desirable properties of the proposed algorithm. We review related work in Section 6 and finally conclude the paper in Section 7.

2. System Model and Problem Formulation

We first introduce participants in the data trading process and describe the data trading model between data owners and data consumers; the mathematical formulation of data trading is then provided.

2.1. System Model. In a data trading market for sharing IoT data, there are mainly two kinds of participants, *end devices* and *edge servers or cloud servers*. End devices who collect data are *data owners*; edge servers or cloud servers who buy data from data owners are *data consumers*.

We divide time into time slots of equal size. Auctions between data owners and data consumers are executed
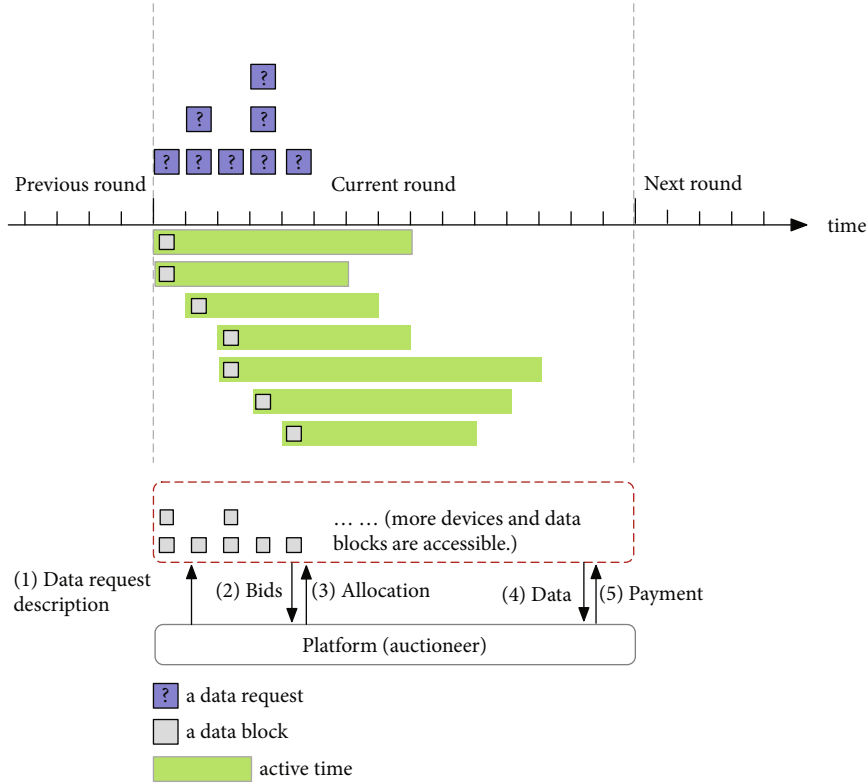
FIGURE 1: The auction process repeated round by round. Data blocks offered by dynamic end devices and dynamic data requests arrives at the first five time slots are depicted in the picture.

round by round. Data owners are short-sighted so that they expect to get as much profit as possible in the current round. Without loss of generality, we only consider the auction process in a single round.

Data requests are submitted randomly and dynamically. We assume that a practical data requirement can be decomposed into several smaller data requests, each of which can be satisfied by a single data block generated by a single data owner. Let τ_j denote the number of data requests arriving at time slot j. A data request submitted at the time slot j is denoted by $r_{j,k}, k \leq \tau_j$. The set of all data requests is denoted by $R = \{r_{j,k} | j = 1, 2, \cdots, T, k \leq \tau_j\}$, where T is the total number of time slots in each round. As we introduced in Section 1, there will be plenty of connected devices such as smartphones and in-vehicle sensing devices in future. Crowdsensing has become extremely popular recently, and the assetization of personal data is a future trend. So, we assume that there are a sufficient number of data owners that every data request is matched to a data block at its arriving time slot for simplicity.

Data owners are dynamic because their data are only accessible during their *active time*. The active time of a data owner i is a time period described by $[s_i, e_i)$, where s_i and e_i are the start time slot and end time slot (not included) of her active time. Each data owner can sell at most one data block or sell her data block at most once during her active time because of limited resources. A data owner i has a valuation v_i for her data block which indicates that she would not

trade her data block with a trading price lower than v_i. Similarly, the reported active time is possibly different from her real active time.

2.2. Data Trading Model Based on the Auction Mechanism. The interaction between data owners and data consumers is modeled by an auction mechanism, as shown in Figure 1. In a data trading market, data owners share their data blocks with others, and then they are compensated according to trading prices; data consumers get data blocks and pay to data owners. There exists a third-party trading platform to manage the auction process. The auction process is described as follows:

(1) The platform sends data requests to data owners

(2) Each data owner generates a bid B_i which reports the active time and valuation for her data block and then sends her bid to the platform

(3) The platform matches bids to data requests and determines the trading time slot and trading price (payment) for each selected bid. Then, platform return matching results to data owners and data consumers

(4) Each data owner whose bid is selected uploads her data block at a specific time slot

(5) The data consumer pays for the data owner according to the negotiated trading price

The platform must determine a *matching rule* and a *trading price rule*. We use $B = \{B_i | i = 1, 2, \cdots, N\}$ to denote the set of all bids submitted by all data owners, where N is the total number of data owners in a round. The matching rule **X** is a matrix of indicator variables; each element, $x_{i,j} \in \{1, 0\}, 1 \le i \le N, 1 \le j \le T$, represents whether bid B_i is selected at time slot j or not. Actually, we should denote the matching rule by $x_{i,j}(B)$ since each $x_{i,j}$ is determined by B. For simplicity, we use $x_{i,j}$ instead of $x_{i,j}(B)$ in the rest of the paper. According to the trading price rule, payment to each bid B_i is denoted by $p_i(B) \in \mathbf{R}$.

In each round, a data owner i submits at most one bid, $B_i = (\tilde{s}_i, \tilde{e}_i, b_i), 1 \le \tilde{s}_i < \tilde{e}_i \le T + 1, b_i \ge 0$, where $[\tilde{s}_i, \tilde{e}_i)$ is the reported active time, and b_i is the bid price. The bid price may be different from real valuation v_i, becasue data owners are usually *selfish*. Similarly, the reported active time period may be different from her real active time.

Next, we discuss utilities of data owners and explain why data owners are selfish.

Definition 1 (utility of a data owner). The utility of a data owner is the difference between the trading price and her valuation if the bid of the data owner is selected. Otherwise, her utility is zero. The utility is computed as follows:

$$u_i^o(B) = \sum_{j=s_i}^{e_i-1} x_{i,j}(p_i(B) - v_i), \qquad (1)$$

where $\sum_{j=s_i}^{e_i-1} x_{i,j} \le 1$ holds because a data owner only trade her data block at most once during her active time.

A data owner is selfish so that she probably selects strategies solely to maximize her utility. The data owner probably misreports start time or end time of her active time as well as to charge a higher price than her valuation. Data owners cannot report earlier arrivals or delayed departure, because it is easy to detect their absence and thus they would be punished. Therefore, there are three of strategic behaviors for selfish data owners, i.e., delayed arrival, earlier departure, and misreporting valuation.

Definition 2 (utility generated by a data request). The utility of a data consumer who publishes a data request is the difference between the amount of profit that the data consumer get from data and the trading price. If the k-th data request $r_{j,k}$ at time slot j is matched to the bid B_i of a data owner i, then the utility generated by the data request $r_{j,k}$ is computed as follows:

$$u_{j,k}^r(B) = x_{i,j}\left(\varphi_j - p_i(B)\right), \qquad (2)$$

where φ_j is the amount of profit that the data requester get from the traded data block.

Consequently, utilities of all data requests at time slot j are

$$\sum_{k=1}^{\tau_j} u_{j,k}^r(B) = \sum_{i=1}^{N} x_{i,j}\left(\varphi_j - p_i(B)\right), \qquad (3)$$

where $\sum_{i=1}^{N} x_{i,j} = \tau_j$ holds because a data request can be satisfied by a single data block, and there are τ_j data owners selected at time slot j.

Definition 3 (social efficiency). The social efficiency is defined as the sum of utilities of all participants. It is computed as follows:

$$
\begin{aligned}
\Delta &= \sum_{i=1}^{N} u_i^o(B) + \sum_{j=1}^{T} \sum_{k=1}^{\tau_j} u_{j,k}^r(B) \\
&= \sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j}(p_i(B) - v_i) + \sum_{j=1}^{T} \sum_{i=1}^{N} x_{i,j}\left(\varphi_j - p_i(B)\right) \\
&\overset{(a)}{=} -\sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j}v_i + \sum_{j=1}^{T} \tau_j\varphi_j \\
&\quad + \left(\sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j}p_i(B) - \sum_{j=1}^{T} \sum_{i=1}^{N} x_{i,j}p_i(B)\right) \\
&\overset{(b)}{=} C - \sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j}v_i,
\end{aligned}
\qquad (4)
$$

where (a) uses $\sum_{i=1}^{N} x_{i,j} = \tau_j$, and (b) follows from exchanging two summations and replacing $\sum_{j=1}^{T} \tau_j\varphi_j$ with a constant parameter C whose value is not related to $x_{i,j}$.

2.3. Problem Formulation. In the paper, we aim to design an auction mechanism for the data trading market so that the maximum social efficiency is achieved as well as following properties, i.e., *individual rationality*, *incentive compatibility*, and *computation efficiency*.

Definition 4 (individual rationality). An auction mechanism satisfies the property of individual rationality if and only if every data owner has a nonnegative utility, i.e., $u_i^o \ge 0, i = 1, 2 \cdots N$.

Definition 5 (incentive compatibility). An auction mechanism is incentive-compatible (which is also called truthful) if and only if, for each data owner i, she cannot increase her utility by misreporting her private information, i.e.,

$$u_i^o(B_i \cup B_{-i}) \ge u_i^o(\tilde{B}_i \cup B_{-i}), \qquad (5)$$

where $B_i = (s_i, e_i, v_i)$ and $\tilde{B}_i = (\tilde{s}_i, \tilde{e}_i, b_i)$ are not the same, which means any of $s_i \le \tilde{s}_i$, $e_i \ge \tilde{e}_i$, and $v_i \ne b_i$ holds; B_{-i} denotes the set of all bids except B_i.

```
Input: The set of bids B.
Output: The matching rule X = {x_{i,j}|i = 1, 2,···,N, j = 1, 2,···,T}.
1 A ⟵ ∅, j ⟵ 1, X ⟵ 0;//A is the set of all active bids which can provide accessible data block at current time slot.
2 while j ≤ T do
3      Remove expired bids (bids with e_i = j) from A;
4      Add newly active bids (bids with s_i = j) to A;
       /*Greedy select the first τ_j bids with lowest bid price at each time slot j.          */
5      for k ⟵ 0 to τ_j do
            //Loop τ_j times.
6           Choose a bid B_{(i')} with the lowest bid price and match it to the k-th data request at current time slot, i.e., x_{(i'),j} ⟵ 1;
7           A ⟵ A − B_{(i')};//Remove B_{(i')} from A.
8      end for
9      j ⟵ j + 1
10 end while
11 return X
```

ALGORITHM 1: Online matching algorithm.

In the following, we offer the mathematical formulation of our problem. We need to determine the matching rule of $X = \{x_{i,j}|i = 1, 2,···,N, j = 1, 2,···,T\}$ by solving the optimization problem defined in (6) as well as the trading price rule $\{p_i|i = 1, 2,···,N\}$ satisfying definitions of (4) and (5).

$$\arg \max_{x_{i,j}} \Delta = \arg \min_{x_{i,j}} \sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j} v_i,$$

$$s.t. \sum_{j=s_i}^{e_i-1} x_{i,j} \leq 1, \sum_{j \in \{t|t<s_i\} \cup \{t|t\geq e_i\}} x_{i,j} = 0, \forall i,$$

$$\sum_{i=1}^{N} x_{i,j} = \tau_j, \forall j, \quad (6)$$

$$x_{i,j} \in \{0, 1\}, \forall i, \forall j.$$

From the objective of the problem, we can see that maximizing social efficiency is equivalent to minimizing sum of valuation from selected data owners.

Our problem cannot be solved by classic optimization algorithms solving linear programming for several reasons. *Firstly*, valuation v_i for each data owner is not accessible. *Secondly*, both data owners (with their data blocks) and data requests join the data trading market dynamically, which means that $x_{i,j}$ should be determined online without future information. *Thirdly*, the solution of Eq. (6) fails to provide any information about trading prices.

3. Near-Optimal Online Data Trading Algorithm

In the section, we propose an online data trading algorithm to determine both the matching rule and the trading price rule. Our online data trading algorithm contains two components, which solves the matching subproblem in Section 3.1 and trading price subproblem in Section 3.2. For simplicity, we assume that every data owner submits a bid which is

exactly the same as her private information, and then we prove that data owners would honestly report her private information under the given trading price rule in Section 3.3.

3.1. Online Matching Algorithm Based on a Greedy Scheme. We propose an online matching algorithm to match data blocks of data owners to data requests using a greedy strategy. The basic idea of this algorithm is to greedily select one with the lowest bid price from current available data blocks to satisfy the newly submitted data request. The selection is executed at the beginning of every time slot. As shown in Algorithm 1, the algorithm maintains a set of active bids which have not been matched to any data request; the set is updated at the beginning of each time slot, i.e., appending or removing bids into the active set according to active time of bids. At each time slot j, the first τ_j bids with lowest bid prices that are selected.

3.2. Computing Trading Prices Based on Critical Data Owners. Unfortunately, a VCG-based payment scheme [16] is inapplicable to the online auction mechanism, because the online matching algorithm is not optimal. In the paper, we propose a trading price determination scheme based on *critical bid* which guarantees that each data owner reports private information truthfully. The basic idea is to set the trading price of a selected bid B_i as the bid price of the first bid that makes B_i fails. The first bid that makes B_i fails is the critical bid of B_i. Actually, if $B_i = (s_i, e_i, b_i)$ is selected at time slot $\bar{j}$ according to the matching rule, the critical bid $c(B_i)$ of B_i is the bid with the highest bid price which are selected during the time period of $[\bar{j}, e_i - 1]$ other than B_i.

Main steps of computing trading price for a selected bid are shown in Algorithm 2. Firstly, remove $B_i = (s_i, e_i, b_i)$ from the set of all bids B. Secondly, employ the matching rule proposed in Algorithm 1 to find all bids that are selected earlier than $\bar{j}$ and remove all of them from the active set of bids. Finally, find the bid (i.e., the critical bid) with the highest bid price from all bids that are selected during the time period $[\bar{j}, e_i - 1]$ and return the bid price of the critical bid

```
Input: A selected bid B_i = (s_i, e_i, b_i), time slot j̄ that B_i is selected (i.e., x_{i,j̄} = 1), the set of all bids B.
Output: The trading price p_i of the data block that is associated to B_i.
1   A ⟵ ∅, j ⟵ 1, p_i ⟵ b_i
2   B ⟵ B − B_i; //Remove B_i from the set of all bids.
3   while j < e_i do
4        Remove expired bids (bids with e_i = j) from A;
5        Add newly active bids (bids with s_i = j) to A;
6        for k ⟵ 0 to τ_j do
7             Choose a bid B_{i'} = (s_{i'}, e_{i'}, b_{i'}) from A with the lowest bid price
8             A ⟵ A − B_{i'};
             /*Find the highest bid price from all bids that are selected during [j̄, e_i − 1]        */
9             if j ≥ j̄ and b_{i'} > p_i then
10                      p_i ⟵ b_{i'};
11              end if
12        end for
13        j ⟵ j + 1;
14  end while
15  return p_i;
```

ALGORITHM 2: Trading price determination algorithm (for a selected bid).

as the trading price. Similarly, we can repeat the procedure in Algorithm 2 for each selected bid. Besides, if a bid is not selected, then the trading price of the data block that is associated to the bid is zero.

3.3. Theoretical Analysis. In the subsection, we prove that the proposed online auction mechanism which contains two components of an online matching algorithm (Algorithm 1) and a trading price determination algorithm (Algorithm 2) satisfies several good properties aforementioned.

To prove the auction mechanism is incentive-compatible, it is equivalent to prove that it satisfies following two conditions: (i) the matching rule in Algorithm 1 is *monotonic*, and (ii) the trading price of the data block associated to each bid is equal to the *critical value*.

Definition 6 (monotonicity). The matching rule is monotonic if and only a data owner whose bid $B_i = (s_i, e_i, b_i)$ is selected would also win if she reports a more attractive bid $\tilde{B}_i = (\tilde{s}_i, \tilde{e}_i, \tilde{b}_i)$ with a lower bid price or a longer active time period, i.e., $\tilde{s}_i \leq s_i, \tilde{e}_i \geq e_i, \tilde{b}_i \leq b_i$.

Definition 7 (critical value). For a data owner whose bid $B_i = (s_i, e_i, b_i)$ is selected, the critical value of the bid is the highest bid price $b_{i'}$ that the data owner submits a bid $B_{i'} = (s_i, e_i, b_{i'} \geq b_i)$, and the new bid B_i' is still selected.

Theorem 8 (incentive compatibility). *The proposed auction mechanism is incentive-compatible, because the matching rule is monotonic, and the trading price is set as the critical value.*

Proof. First of all, we show that the matching rule is monotonic. Suppose a bid $B_i = (s_i, e_i, b_i)$ is selected at time slot j according to the matching rule. We replace the bid B_i with another bid $\tilde{B}_i = (\tilde{s}_i, \tilde{e}_i, \tilde{b}_i)$, where $\tilde{s}_i \leq s_i, \tilde{e}_i \geq e_i, \tilde{b}_i \leq b_i$. Obvi-

ously, $\tilde{B}_i$ would be selected at time slot j or earlier. Therefore, the matching rule is proved to be monotonic.

Then, we check whether the trading price computed by Algorithm 2 is exactly the critical value. Suppose a data owner whose bid $B_i = (s_i, e_i, b_i)$ is selected at time slot j and the trading price of this bid computed by Algorithm 2 is $\dot{p}_i$. Therefore, there must be another bid B^0 whose bid price is $\dot{p}_i$ and is selected during the time period of $[j, e_i − 1]$. If the data owner submits another bid $\hat{B}_i = (s_i, e_i, \dot{p}_i − \xi)$ instead of B_i, where $\xi > 0$, then $\hat{B}_i$ would be selected during its active time and makes B^0 fails. On the contrary, if the data owner submits another bid $\bar{B}_i = (s_i, e_i, \dot{p}_i + \zeta)$ instead of B_i, where $\zeta > 0$, then $\bar{B}_i$ would not be selected since its bid price is higher than all selected bids during its active time. So, we have verified that the trading price computed by Algorithm 2 is the critical value. We therefore conclude that the proposed auction mechanism is incentive-compatible.   □

Theorem 9 (individual rationality). *The proposed online auction mechanism is individually rational.*

Proof. For a data owner whose bid fails, her utility is zero. For a data owner whose bid is selected at any time slot, we can compute the trading price of her data block according to Algorithm 2. Suppose a bid $B_i = (s_i, e_i, b_i)$ is selected at time slot j, there must be another bid $B_{i'} = (s_i, d_{i'}, b_{i'})$ chosen at time slot j, and p_i is updated to $b_{i'}$ (line 10 in Algorithm 2). According to the update rule of trading prices, the final trading price would be $p_i \geq b_{i'}$. We can see that $b_{i'} \geq b_i$; otherwise, $B_{i'}$ would be selected at time slot j instead of B_i according to the matching rule in Algorithm 1. Since we have demonstrated that the auction mechanism is incentive-compatible, i.e., every data owner would report their valuation truthfully, and we can get that $p_i \geq b_{i'} \geq b_i = v_i$. Therefore, the utility of data owners are always nonnegative.   □

Theorem 10 (competitive ratio). *The online matching algorithm achieves a competitive ratio of 1/2, i.e., $\Delta_{online}/\Delta^* \geq 1/2$, where Δ_{online} and Δ^* denote the resulting social efficiency of the online matching algorithm and the optimal solution of Eq. (6), respectively.*

Proof. The competitive ratio is computed by introducing a parameter a whose value is 0 initially. For a bid $B_i = (s_i, e_i, v_i)$ that is selected both in the online matching rule and the optimal solution, increase a by $\varphi_j - v_i$ (suppose B_i is selected at time slot j). For a bid B_i that is selected in the optimal solution at time slot j to a data request but not in the online matching rule, suppose this data request is matched to another bid $B_{i'}$ with a bid price of $v_{i'}$ in the online case, i.e., $v_{i'} \leq v_i$, increase a by $\varphi_j - v_{i'}$. Then, we can get $a \geq \Delta^*$.

For each bid that is matched by the online matching rule, its matching value $\varphi_j - v_i$ is added to a at most twice, i.e., $a \leq 2\Delta_{online}$. Therefore, we have $2\Delta_{online} \geq a \geq \Delta^*$. □

4. Decentralized Data Trading Platform

In this section, we design a decentralized data trading platform based on the blockchain system.

4.1. Overview of System Architecture. In this chapter, we construct a decentralized data trading platform based on the blockchain technology to in order to avoid some distrust issues caused by data owners and data consumers, such as refusing to pay or cheating. As shown in Figure 2, the proposed data trading platform contains three layers: an application layer, a blockchain layer, and a storage layer. The application layer is the most important part which need much more efforts to design and implement while the other two layers are realized based on existing development tools mentioned in Section 4.3:

(i) Application layer: the application layer, which is functioned as a client of the proposed decentralized data trading platform, mainly includes three components of *user management*, *online auction*, and *data trading process*. Each components are realized by one or more smart contracts built upon the underlying blockchain system. The user management component is responsible for management of all participants, such as maintaining user information of registered data owners and data consumers, removing invalid data owners which are not active for a long time, and management of deposit which are provided by participants in order to reduce malicious behavior. Either a data owners or a data consumer should register an account to participate in the data trading process. The online auction component is responsible for realizing the online data trading algorithm proposed in Section 3 which then informs both data owners and data consumers of matching results, trading prices, and maximum social welfare. The data trading process components are
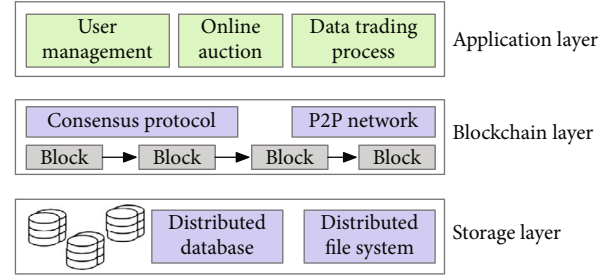

FIGURE 2: System architecture of the proposed decentralized data trading platform.

designed to organize the whole process of online data trading, including registration of participants, online data trading process, data upload and data transmission, and payment transfer. A lot of crucial information of data trading process, e.g., bid information, matching results, and trading prices, is coded into transition blocks and saved by the blockchain system

(ii) Blockchain layer: we choose a consortium blockchain system as a bridge between the application layer and the storage layer for the purpose of simplicity and efficiency. Each node saves the complete chain of transaction blocks composed of all transactions in the blockchain system. The Byzantine fault-tolerant distributed consensus protocol is used internally, which can provide stable and reliable services even if there are a small number of malicious nodes

(iii) Storage layer: we introduce a distributed database to store a large amount of traded data and to support for the process of efficient data transmission. That is, traded data is not necessary to upload and save in the blockchain system. Instead, traded data blocks are uploaded to the distributed database, and only hash values and locations of traded data blocks are sent to the data trading platform or data consumers. The distributed database used our proposed decentralized data trading platform is the InterPlanetary File System (IPFS), which manages files through a distributed file management system

4.2. The Whole Process of Online Data Trading. We expand the online data trading process in Section 2.2 to derive a whole process of online data trading upon a decentralized data trading platform as shown in Figure 3. For brevity, we omit two phases of deposit management and user registration. Either a data owner or a data consumer pays an amount of deposit before participating in the data trading process in order to avoid mistrust behaviors such as refusing to offer payment, refusing to collect or upload data. We list all phases of the whole process of online data trading as follows:

(1) Each data consumer submits a data request at a randomly chosen time slot

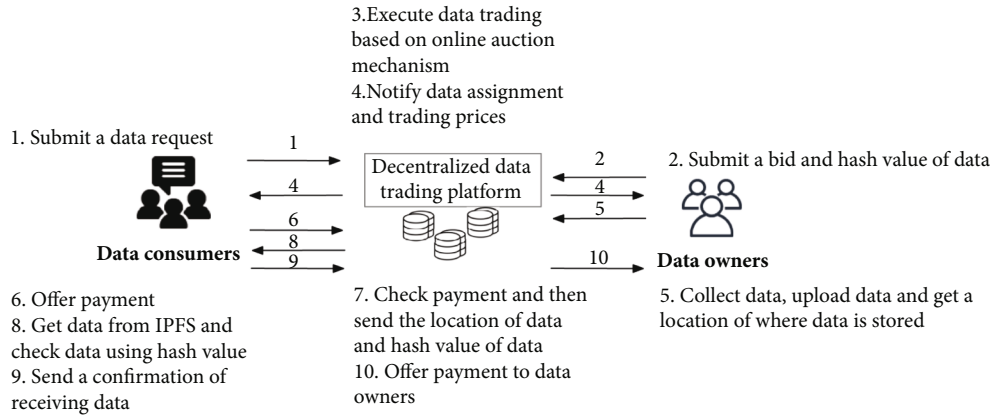(2) Each data owner submits a bid including information of bidding price and active time

FIGURE 3: The whole process of online data trading upon the decentralized data trading platform.

(3) The proposed near-optimal online data trading algorithm is executed

(4) Data matching results and trading prices are sent to all participants

(5) Each data owner whose bid is chosen uploads data to the distributed file system, which would return a location which describes where the data block is stored along with hash value of data

(6) A data consumer offers payment to the data trading platform according to trading prices derived by the online data trading algorithm

(7) The data trading platform checks payment and then sends the location of data along with hash value to the data consumer

(8) The data consumer obtains the traded data block via the data location and check data using hash value

(9) The data consumer sends a conformation of receiving data

(10) Data owners finally receive payment

For instance, there is a data request generated that requires road temperature and humidity data sometime. At the same time, three data owners with the required data submitted bids and their bid prices are 5, 7, and 10. According to our proposed algorithm, the platform matches the request with bid 1 and determines the trading price (payment) as 7. After the data owner uploads the data and the platform checks payment, the data storage location and hash value are sent to the data consumer. The data owner will receive payment after the data consumer check data.

4.3. Development Tools. We list several development tools used to construct the decentralized data trading platform. We employ the first two development tools to construct the decentralized data trading platform and further employ the latter two to develop a web-based user interface for participants.

(i) FISCO BCOS: FISCO BCOS provides a series of visual middleware tools which greatly simplify the process of building a chain of transaction blocks. It is very useful because it supports Java, Python, and other SDKs and also provides desirable properties in terms of security and scalability

(ii) WeBase: WeBase is an open source middleware platform developed by WeBank which provides a simple method of directly linking the blockchain system and upper applications. The WeBase middleware platform encapsulates complicated implementation details of the underlying blockchain system, greatly reducing efforts of developing decentralized applications and improving development efficiency of developers

(iii) Spring-Boot: developers can use the Spring Boot framework to avoid complex XML annotation assignment in Spring. The spring MVC three-tier architecture pattern can realize loose coupling between different functional modules, which greatly improves scalability and development efficiency of the system

(iv) Vue: Vue is one of the mainstream front-end development frameworks recently. In the development process, we only need to pay attention to the view layer, which is equipped with a complete third-party class library, which is loosely coupled with the back-end development framework

5. Numerical Illustration

In the section, we perform extensive simulation and report simulation results to show performance of the proposed online auction mechanism for data trading with dynamic data owners.

5.1. *Methodology and Simulation Settings.* We compare our proposed online auction mechanism with an optimal data trading algorithm with an optimal matching rule based on complete information and an incentive-compatible trading price rule. Suppose that active time of data owners are known and submitted in advance, we employ the Hungarian algorithm to find the optimal matching and then introduce the VCG payment scheme to compute trading prices. That is to say, since we have complete information about all bids and data requests in the auction and make decisions off-line, we can solve the optimization problem in Eq. (6) using the Hungarian algorithm. The VCG scheme can be utilized to determine trading prices, stimulating data owners to report their private information truthfully.

In simulation experiments, we suppose that both data requests and data owners are generated randomly. Generally, the Poisson distribution can also be used to formulate the number of events in a specified interval. For example, the number of calls received during any minute has a Poisson probability distribution with a specific mean. Therefore, both arrivals of data owners and data requests are generated from Poisson distributions, i.e., the number of data owners newly joining the data trading market at each time slot follows a Poisson distribution with a parameter $\lambda_o = 10$; the number of data requests submitted at each time slot follows a Poisson distribution with a parameter $\lambda_t = 3$. The active time length, i.e., the number of slots of active time, follows a uniform distribution. The maximum value of active time length is 20, and the minimum value is 5. The valuation every data owner or data requester is randomly chosen from a uniform distribution, ranging from 1 to 10. The number of time slots in a round is 50. The same parameters are used for the baselines. Settings of all parameters used in simulation experiments are listed in Table 1.

In following pictures, the online auction mechanism and the optimal data trading algorithm are denoted by "Online" and "Opt" in the legends, respectively. Furthermore, these approaches under different valuations of data owners' are evaluated, denoted by "Online, $v= [1, 5]$" and "Opt, $v= [1, 5]$" for example.

These two approaches are evaluated with extensive simulations based on three metrics of *social efficiency*, *competitive ratio*, and *running time*. We conduct several groups of experiments and report comparison results of these two approaches. Each point in these figures is average result over 100 runs.

5.2. *Numerical Results*

5.2.1. *Evaluation on Social Efficiency.* According to Eq. (4), C is a constant which is irrelevant with matching and trading prices of bids, and we instead evaluate performance of sum of valuation of all data owners whose have been matching, i.e., $\sum_{i=1}^{N} \sum_{j=s_i}^{e_i-1} x_{i,j} v_i$, in Figure 4. As shown in Figure 4(a), there is a decrease in sum of valuation for all matched data owners when there are more available data owners with a larger arriving rate λ_o. It is obvious that sum of valuation is on the increase along with the number of time slots in a round, since there are more data requests are satisfied and

TABLE 1: Summary of default settings.

Parameter	Default value
Rate of data owners, λ_o	10
Rate of data requests, λ_t	3
Range of active time length	[5, 20]
Range of data owners' valuation	[1, 10]
Number of time slots in a round, T	50

more data owners are chosen. The performance of Opt outperforms Online in terms of sum of valuation.

To stimulate data owners to honestly report their valuation, the trading price of a data block is usually no lower than the valuation that data owner claims. We further introduce a metric of *overpayment ratio* to show that data requests should pay extra money to ensure social efficiency. The overpayment ratio is the amount of extra expenditure (i.e., the difference between the trading price and the valuation of a data block) to the valuation. The performance of overpayment ratio is shown in Figure 5. Compared to Opt, the proposed approach of Online must pay higher prices to encourage data owners' cooperation since less information is known in the setting of Online. We can also get that performance of Opt is stabler than Online with different arriving rates of data owners or different numbers of time slots in a round.

5.2.2. *Evaluation on Competitive Ratio.* We plot empirical CDF of competitive ratio of the proposed online auction with different parameters in Figure 6. In each parameter setting, simulations are repeated 1,000 times. The result on competitive ratio of each run is regarded as a sample; all these samples are utilized to derive the empirical CDF. When the valuation range of data owners varies, we can see that the competitive ratio is always above the bound of 0.5. To simplify the simulation, we set the valuation of each data request to be the upper bound of the valuation of data owners and then compute social efficiency of Online.

To compare the proposed online auction mechanism with optimal algorithm in detail, we further compare ratio of sum of valuation between different methods with different parameter settings; results are shown in Tables 2 and 3. We can see that the sum of valuation of Online is a little higher than that of Opt. Results of ratio of sum of valuation range from 1.03 to 1.06. Ratio of sum of valuation remains stable, even if number of time slots increases or arriving rate of data owners increases. With smaller range of valuations, i.e., $v = [1, 5]$, the ratio is smaller since data owners participate in a more competitive auction, and it is easier to induce their truthfulness.

5.2.3. *Evaluation on Individual Rationality.* We plot an empirical CDF of individual rationality of the proposed online auction with different parameters in Figure 7. In each parameter setting, simulations are repeated 10 times, and 5,900 bids are selected. The utility corresponding to every bid is regarded as a sample. When the valuation range of data owners varies, we can see that the individual rationality
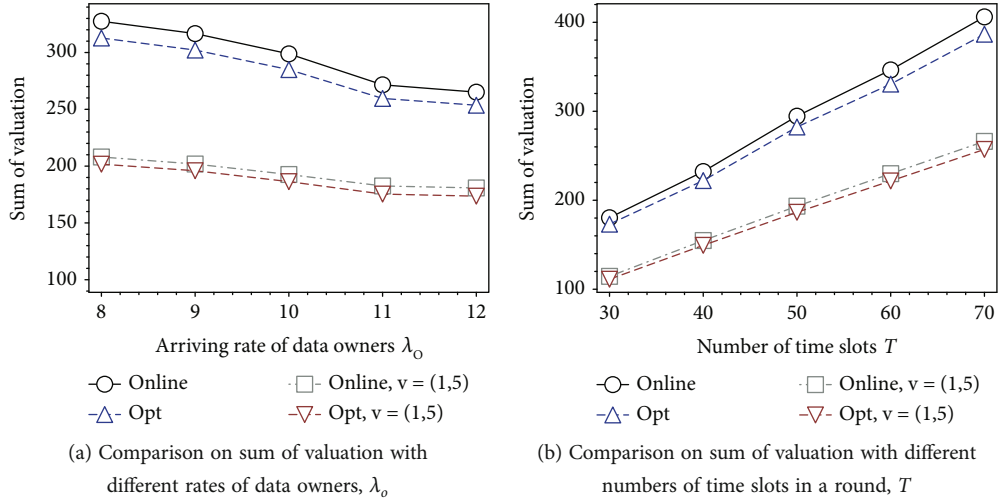
(a) Comparison on sum of valuation with
different rates of data owners, λ_o

(b) Comparison on sum of valuation with different
numbers of time slots in a round, T

FIGURE 4: Performance of sum of valuation.

(a) Comparison on overpayment ratio with
different rates of data owners, λ_o

(b) Comparison on overpayment ratio with different
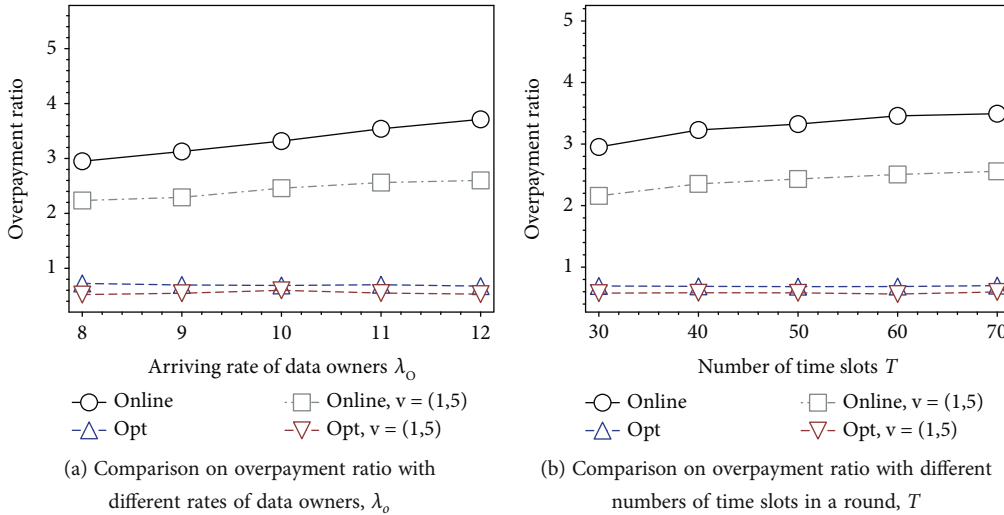numbers of time slots in a round, T

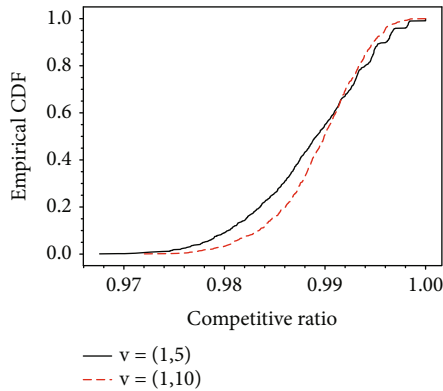FIGURE 5: Performance of overpayment ratio.

FIGURE 6: Empirical CDF of competitive ratio.

is always above the bound of 0. The reason why the shape in the figure is not a curve is that valuation is always taken as an integer in our simulation,

5.2.4. Evaluation on Truthfulness. In order to prove the truthfulness of proposed method, we randomly select two bids and record their utilities when they misreport their bid prices.

For the first bid with valuation 2 (bid 1), we vary his bid price from 1 to 9 and record his utilities. Similarly, for the second bid with valuation 6 (bid 2), we conduct the same experiment.

The results are shown in Figure 8, and we can see that for any one of these two bids, claiming a bid price which is different from his valuation does not improve his utility. Therefore, we verify that the property of truthfulness is satisfied and conclude that data owners have no incentive to misreport their valuation.

5.2.5. Evaluation on Running Time. To show computation efficiency of proposed online auction mechanism, running time of both the matching algorithm and the trading price determination algorithms are recorded in Tables 4 and 5,

TABLE 2: Ratio of sum of valuation when the number of time slots in a round changes.

T	30	40	50	60	70
Ratio (Online)	1.04253	1.04609	1.04401	1.04854	1.05032
Ratio (Online, $v = [1, 5]$)	1.03155	1.03763	1.03783	1.03822	1.03606

TABLE 3: Ratio of sum of valuation when arriving rate of data owners changes.

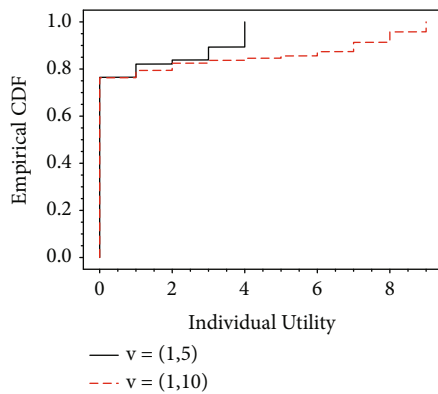| λ_o	8	9	10	11	12
Ratio (Online)	1.04813	1.04819	1.04862	1.04670	1.04565
Ratio (Online, $v = [1, 5]$)	1.03109	1.03092	1.03488	1.04189	1.04236

FIGURE 7: Empirical CDF of utilities.

respectively. Obviously, we can find that the offline solution, i.e., Opt, takes much longer time than the online solution for both. When the number of time slots is larger, there would be more data consumers and more data requests to be matched at current round, certainly causing longer running time. In addition, with a smaller range of valuations, shorter time is required for both Online and Opt, which is because a data request is more likely to be handled in shorter delay, and the matching task at each time slot is simpler.

6. Related Work

We review related work from the following two aspects and point out that existing solutions cannot be applied to solve our problem.

6.1. Decentralized Data Trading Based on the Blockchain Technology. A great many research papers have paid attention to design a data trading market or platform based on blockchain technology [10–14] because of the absence of a trustworthy and centralized data trading platform, single point of failure, and DDoS. Dai et al. claim that both data brokers and buyers are dishonest, and none of them is accessible for raw data; data processing and analysis algorithms encoded in smart contracts are deployed in a secure data trading platform supported by the hardware of Intel's Software Guard Extensions- (SGX-) based secure execution environment [10]. Another trusted data trading platform

employing both the blockchain technology and trusted execution environment (TEE) is implemented by Su et al., where the trusted trading platform contains a special kind of nodes each of which supports TEE and serves as a trust exchange for exchanging data or payment between data sellers and data buyers [11]. Ha et al. introduce a decentralized private data trading marketplace called "Digital Me" based on the blockchain technology [12], where data sellers and data buyers trade personal data directly without trustworthy servers. The AI agent included in "Digital Me" serves as a trading assistant to recommend trading prices based on a user's personal data and data transaction history data. He et al. propose a distributed and trusted data trading platform based on blockchain technology to detect misbehavior of participants and a dataset similarity comparison scheme based on MinHash for detecting illegal resale efficiently is then employed [13]. Zheng et al. deploy smart contracts to solve the problem of data matching and reward distribution in a distributed data trading platform and then introduce proxy reencryption to guarantee secure data transmission, where trading data are encrypted, and only valid data requesters are allowed to decrypt trading data [14]. Nguyen et al. design a distributed ledger-based IoT data trading system along with three typical data trading protocols for city-level environmental monitoring using NB-IoT connections and further analyze the cost of data trading in terms of end-to-end transmission latency and energy consumption [17]. To achieve a good trade-off between the privacy and data utility, Sabounchi and Wei exploit the blockchain techniques and contract theory to design a blockchain-based peer-to-peer data trading mechanism [18]. The trustless environment of Internet of Electric Vehicles, including fuel vehicles and EVs, encounters trading disputes and conflicting interests among trading parties. To address it, Sadiq et al. exploit consortium blockchain to maintain transparency and trust in trading activities. Smart contracts are used to tackle trading disputes and illegal actions [19]. Although many valuable personal data are generated by individuals, only centralized service providers get profit from the data. Yoon et al. propose a blockchain-based personal data trading system using DID (Decentralized Identifiers) and VC (Verifiable Credentials), and the proposed system allows users to collect personal data in their own data storage provided by the system [20].
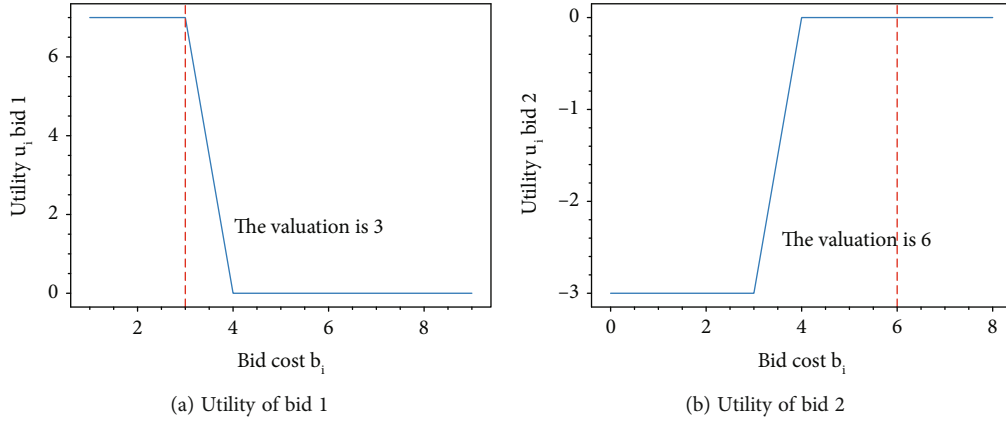
(a) Utility of bid 1

(b) Utility of bid 2

FIGURE 8: Verification of truthfulness.

TABLE 4: Running time of matching when the number of time slots in a round changes.

T	30	40	50	60
Running time (Online)	0.00172	0.00172	0.00250	0.00266
Running time (Online, $v = [1, 5]$)	0.00078	0.00219	0.00141	0.00391
Running time (Opt)	0.00344	0.00469	0.00594	0.00906
Running time (Opt, $v = [1, 5]$)	0.00250	0.00422	0.00500	0.00672

TABLE 5: Running time of trading price determination when the number of time slots in a round changes.

T	30	40	50	60
Running time (online)	0.11077	0.17767	0.27264	0.38186
Running time (online, $v = [1, 5]$)	0.10905	0.18358	0.27734	0.39044
Running time (opt)	0.28500	0.50982	0.83949	1.41682
Running time (opt, $v = [1, 5]$)	0.27357	0.56167	0.92322	1.40695

6.2. Trading Data with Different Levels of Privacy. Existing studies [5–8] investigate how to trade private data where the authors propose pricing functions for personal data or other private data to compensate data owners' different levels of privacy loss. Private data with different privacy loss is generally traded at different prices. Furthermore, data in different formats e.g., data samples and range counts, are returned to data consumers. Higher prices should be paid to a higher level of privacy loss caused by traded data undoubtedly. A few desirable properties, e.g., arbitrage-freeness, budget feasibility, and performance accuracy based on traded data, are considered when designing pricing functions.

Gao et al. propose a pricing rule based on an auction-based model where both task description and bid prices in bids are possible to disclose sensitive information of data owners; they employ differential privacy schemes at both stages of data collection and trading price determination [21]. Another research paper [22] using the auction-based model introduces geoindistinguishability to quantify privacy loss of geographical locations and then pays for their sensing cost as well as privacy breach. Zhang et al. point out that dis-

closure of raw social media data of users probably cause privacy leakage, because anonymous user IDs can be linked to real users, and they propose a novel mechanism based on a notion of ε-text indistinguishability to guarantee different user privacy as well as to achieve high data utility [23]. Wang et al. study the value of data privacy in a game-theoretic model of trading private data, and they propose that the value of ε units of privacy is measured by the minimum payment of all nonnegative payment mechanisms, under which an individual's best response at a Nash equilibrium is to report the data with a privacy level of ε [24]. Cai et al. study the trading of multiple correlated queries on private web browsing history data and propose TERBE which is a novel trading framework for correlated queries based on private web browsing histories [25].

All existing studies, however, neglect the dynamic behavior of data owners as well as randomly generated data requests of data consumers in IoT data trading. It is reasonable that end devices only trade their data intermittently due to limited resources or mobility. In this paper, we further design a decentralized data trading platform which extends our previous work [26].

7. Conclusion

In the paper, we have investigated the data trading problem with dynamic data owners, aiming to share IoT data and to take full advantages of big data. Existing studies have neglected an important observation that a data owner is not always available to trade her data blocks except in her active time period. To this end, we have proposed a truthful and efficient online data trading algorithm which not only resolves the problem of matching dynamic data owners and randomly generated data requests with near-optimal social efficiency but also determines the trading price of each data block which ensures the incentive-compatibility and individual rationality of participants.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] "IoT growth demands rethink of long-term storage strategies, says IDC," https://www.idc.com/getdoc.jsp?containerId=prAP46737220.

[2] T. Liu, W. Wu, Y. Zhu, and W. Tong, "Accuracy-guaranteed event detection via collaborative mobile crowdsensing with unreliable users," in *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 729–744, Springer, 2019.

[3] C. Yang, P. Li, T. Zhang et al., "Mobile data sharing with multiple user collaboration in mobile crowdsensing (short paper)," in *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 356–365, Shanghai, China, 2019.

[4] T. Liu, D. Li, C. Cao, H. Gao, C. Li, and Z. Feng, "Joint location-value privacy protection for spatiotemporal data collection via mobile crowdsensing," in *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 87–103, Virtual Event, 2021.

[5] C. Niu, Z. Zheng, F. Wu, S. Tang, X. Gao, and G. Chen, "Unlocking the value of privacy: trading aggregate statistics over private correlated data," in *ACM International Conference on Knowledge Discovery & Data Mining (KDD)*, pp. 2031–2040, London, UK, 2018.

[6] S. Zheng, Y. Cao, and M. Yoshikawa, "Money cannot buy everything: trading mobile data with controllable privacy loss," in *IEEE International Conference Mobile Data Manage (MDM)*, pp. 29–38, Versailles, France, 2020.

[7] C. Li, D. Y. Li, G. Miklau, and D. Suciu, "A theory of pricing private data," *Communications of the ACM*, vol. 60, no. 12, pp. 79–86, 2017.

[8] Z. Cai and Z. He, "Trading private range counting over big IoT data," in *IEEE International Conference on Distributed Computing System (ICDCS)*, pp. 144–153, Dallas, TX, USA, 2019.

[9] Z. Feng, J. Chen, and Y. Zhu, "Uncovering value of correlated data: trading data based on iterative combinatorial auction," in *IEEE Internationale Conference on Mobile Ad-Hoc and Smart System(MASS)*, pp. 260–268, Denver, CO, USA, 2021.

[10] W. Dai, C. Dai, K.-K. R. Choo, C. Cui, D. Zou, and H. Jin, "SDTE: a secure blockchain-based data trading ecosystem," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 725–737, 2019.

[11] G. Su, W. Yang, Z. Luo, Y. Zhang, Z. Bai, and Y. Zhu, "BDTF: A blockchain-based data trading framework with trusted execution environment," in *2020 16th International Conference on Mobility, Sensing and Networking (MSN)*, pp. 92–97, Tokyo, Japan, 2020.

[12] M. Ha, S. Kwon, Y. J. Lee, Y. Shim, and J. Kim, "Where WTS meets WTB: a blockchain-based marketplace for Digital Me to trade users' private data," *Pervasive and Mobile Computing*, vol. 59, article 101078, 2019.

[13] Y. He, H. Zhu, C. Wang, K. Xiao, Y. Zhou, and Y. Xin, "An accountable data trading platform based on blockchain," in *IEEE International Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, Paris, France, 2019.

[14] S. Zheng, L. Pan, D. Hu, M. Li, and Y. Fan, "A blockchainbased trading platform for big data," in *IEEE International Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 991–996, Toronto, ON, Canada, 2020.

[15] Z. Feng and J. Chen, "Blockchain based mobile crowd sensing for reliable data sharing in IoT systems," in *2021 IFIP Networking Conference (IFIP Networking)*, pp. 1–3, Espoo and Helsinki, Finland, 2021.

[16] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds., *Algorithmic Game Theory*, Cambridge University Press, 2011.

[17] L. D. Nguyen, I. Leyva-Mayorga, A. N. Lewis, and P. Popovski, "Modeling and analysis of data trading on blockchain-based market in IoT networks," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6487–6497, 2021.

[18] M. Sabounchi and J. Wei, "Blockchain- enabled peer-to-peer data trading mechanism," in *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 1410–1416, Halifax, NS, Canada, 2018.

[19] A. Sadiq, M. U. Javed, R. Khalid, A. Almogren, M. Shafiq, and N. Javaid, "Blockchain based data and energy trading in internet of electric vehicles," *IEEE Access*, vol. 9, pp. 7000–7020, 2021.

[20] D. Yoon, S. Moon, K. Park, and S. Noh, "Blockchainbased personal data trading system using decentralized identifiers and verifiable credentials," in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 150–154, Jeju Island, Republic of Korea, 2021.

[21] G. Gao, M. Xiao, J. Wu, S. Zhang, L. Huang, and G. Xiao, "DPDT: a differentially private crowd-sensed data trading mechanism," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 751–762, 2020.

[22] W. Jin, M. Xiao, M. Li, and L. Guo, "If you do not care about it, sell it: trading location privacy in mobile crowd sensing," in *IEEE International Conference on Computer Communications (INFOCOM)*, pp. 1045–1053, Paris, France, 2019.

[23] J. Zhang, J. Sun, R. Zhang, Y. Zhang, and X. Hu, "Privacy-preserving social media data outsourcing," in *IEEE International Conference on Computer Communications (INFOCOM)*, pp. 1106–1114, Honolulu, HI, USA, 2018.

[24] W. Wang, L. Ying, and J. Zhang, "The value of privacy: Strategic data subjects, incentive mechanisms and fundamental limits," in *ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pp. 249–260, Antibes Juan-Les-Pins, France, 2016.

[25] H. Cai, F. Ye, Y. Yang, Y. Zhu, and J. Li, "Towards correlated queries on trading of private web browsing history," in *IEEE International Conference on Computer Communications (INFOCOM)*, pp. 99–108, Toronto, ON, Canada, 2020.

[26] Z. Feng, J. Chen, and T. Liu, "An online truthful auction for IoT data trading with dynamic data owners," in *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 554–571, Springer International Publishing, Virtual Event, 2021.