Data Trading With Multiple Owners, Collectors, and Users: An Iterative Auction Mechanism

Xuanyu Cao, Yan Chen, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—In the big data era, it is vital to allocate the vast amount of data to heterogeneous users with different interests. To clinch this goal, various agents including data owners, collectors, and users should cooperate to trade data efficiently. However, the data agents (data owners, collectors, and users) are selfish and seek to maximize their own utilities instead of the overall system efficiency. As such, a sophisticated mechanism is imperative to guide the agents to distribute data efficiently. In this paper, the data trading problem of a data market with multiple data owners, collectors, and users is formulated and an iterative auction mechanism is proposed to coordinate the trading. The proposed mechanism guides the selfish data agents to trade data efficiently in terms of social welfare and avoids direct access of the agents' private information. We theoretically prove that the proposed mechanism can achieve the socially optimal operation point. Moreover, we demonstrate that the mechanism satisfies appealing economic properties such as individual rationality and weakly balanced budget. Then, we expand the mechanism to nonexclusive data trading, in which the same data can be dispensed to multiple collectors and users. Simulations as well as real data experiments validate the theoretical properties of the mechanism.**

***Index Terms*—Budget balance, data trading, iterative auction, individual rationality, optimization, social welfare.**

I. Introduction

In THE big data era, vast amount of data are generated and exploited by various agents. For example, numerous memes such as Twitter hashtags are produced in online social networks and millions of videos are uploaded to Youtube. Many software/APP developers may need certain online data (such as the click-through rate of some advertisements or mention count dynamics of some memes) to enhance the quality of their products. As another example, with the development of data procurement and storage capability, many organizations own databases of the statistics of their fields, e.g., hospitals may have data about the clinical performances of medicines. In order to conduct research, researchers need to access these data owned by organizations. In all these circumstances, we face the problem of allotting/trading data from the data owners (e.g., social networks/websites or organizations) to the data users (e.g., software companies or researchers). In fact, several data trading markets or companies have already emerged recently, such as the Data Marketplace, Big Data Exchange and Microsoft Azure Marketplace. However, these data markets are still at the incipient stage and lack appropriate regulations. Economically, the data agents are selfish and seek to maximize their own utilities instead of the overall system efficiency. As such, a sophisticated mechanism is imperative to guide the agents to distribute or trade data efficiently.

The problem of coordinating data trading in a data market falls into the general topic of resource trading/allocation in networks, for which abundant works have been done in the past decades. For communication networks, by using optimization and game theoretic techniques, researchers propose various algorithms to allocate power [1], [2] or channels [3], [4] to communication nodes or access points. For cognitive radio networks, spectrum resources are allotted among primary users and secondary users [5], [6]. For power networks or smart grids, power or voltage resources are distributed to devices and apparatuses in order to maintain high-performance and stable power systems [7]–[9]. The most relevant resource allocation/trading problem to this paper is the privacy trading problem [10]. In most privacy trading problems investigated in the current literature, a single data collector is aimed at collecting binary data from multiple data owners in order to estimate some statistics. From example, each data owner may have a binary answer (yes/no) to some problem and the data collector wants to estimate the proportion of data owners with the answer yes. The involved data are private and leakage of them to the data collector compromises the security of data owners. The loss from this compromising of privacy can be quantified by the differential privacy [11]. As such, data owners should be somehow compensated by the data collector. Additionally, data owners are selfish and may not report their true data to the data collector. Therefore, from the perspective of the data collector, a mechanism is needed to collect accurate data at a low cost from the data owners. To this end, Ghosh and Roth proposed an auction mechanism for a single data collector to collect data from multiple data owners [12]. Along this line, Fleischer and Lyu extended the auction mechanism to the scenario where individual data owner's valuation of the data privacy

Manuscript received September 12, 2016; revised November 10, 2016; accepted December 31, 2016. Date of publication February 13, 2017; date of current version May 17, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61672137. The work of Y. Chen was supported by the Thousand Youth Talents Program of China. The associate editor coordinating the review of this manuscript and approving it for publication was Zhenliang Zhang.

X. Cao and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: apogne@umd.edu; kjrliu@umd.edu).

Y. Chen is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610051, China (e-mail: eecyan@uestc.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSIPN.2017.2668144

was correlated with the data themselves [13]. Furthermore, Xu *et al.* proposed a contract-theoretic mechanism to collect general private data which are not necessarily binary [14].

However, there are two limitations of existing models of data trading in the aforementioned works [12]–[14]. First, in the existing models, there is only one single data collector. This is not the case in most real-world data market, where multiple data collectors (such as many companies or groups like Big Data Exchange) often coexist and compete with each other. Second, in most data markets, the data collectors usually do not exploit the data by themselves. Instead, they often sell the data to data users, who are not capable of collecting and storing massive datasets but need data to develop projects or conduct research. For example, many APP developers are small companies who cannot afford collecting necessary data to develop APPs and thus need to purchase data from professional data collecting companies. In other words, in data markets, besides data owners and collectors, there are data users who can make use of the data but are not able to collect data by themselves. In this paper, we take the above mentioned two limitations of existing works into consideration and investigate the data trading problem in a market with multiple data owners, collectors and users (in the following, we use the term *data agents* to refer to data owners, collectors and users).

Due to the existence of multiple collectors and users, the problem in this paper is significantly different from the data trading in [12]–[14]. Instead of maximizing the profit of a single collector as in previous works, we consider from a system designer's perspective and are aimed at maximizing the overall social welfare, which quantifies the operation efficiency of the data market. However, in practice, the data agents are usually selfish and seek to maximize their own utilities instead of the overall system performance. In order to coordinate the data trading among multiple selfish agents, we resort to the *iterative auction* mechanism, which is initially proposed in [15]. In iterative auction, the auctioneer announces the resource allocation and payment rules to the bidders. Then, the selfish bidders submit appropriate bids to the auctioneer with the goal of maximizing their own utilities. Based on the submitted bids, the auctioneer adjusts the resource allocation and payment rules and another round of auction starts. Through careful design of the mechanism, the iterative auction may converge to an operation point with satisfactory properties. The iterative auction has already been successfully applied to resource allocation in communication networks [16]–[19].

The contribution of this paper is epitomized in the following.

1) We present a data market model with multiple data owners, collectors and users who have heterogeneous utility functions. Considering from the perspective of the system designer, we formulate corresponding social welfare maximization problem.
2) An iterative auction mechanism is proposed to coordinate the data trading among the data agents. The mechanism avoids direct access to the data agents' utility functions, which are private information unknown to the system designer. The selfish nature of individual data agents is also respected in the mechanism.

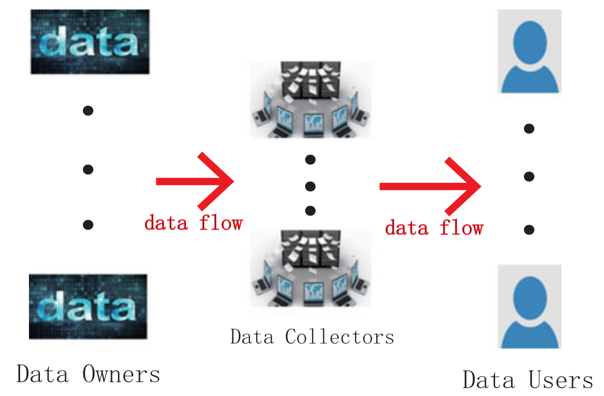


Fig. 1. A data market with multiple data owners, collectors and users.

3) We theoretically show that the proposed mechanism converges to the socially optimal operation point. We also analytically substantiate that the mechanism possesses appealing economic properties including individual rationality and weakly balanced budget.
4) We also extend the mechanism to the non-exclusive data trading scenario, where the same data can be used by multiple data users repeatedly.
5) Simulations as well as real data experiments are implemented to validate the theoretical results of the mechanism.

The roadmap of this paper is as follows. In Section II, our model of the data market is presented and the social welfare maximization problem is formulated. In Section III, we design an iterative auction mechanism to coordinate the data trading. The convergence analysis and economic properties of the proposed mechanism are presented in Section IV. Then, we extend the mechanism to the non-exclusive data trading scenario in Section V. In Section VI, simulation results and real data experiments are shown. Lastly, we conclude the paper in Section VII.

II. Model

In this section, we describe the model of a data market with multiple data owners, collectors and users in detail. Then, we formulate the associated social welfare maximization problem and motivate the iterative auction mechanism.

Consider a data market with M data owners, N data collectors and L data users as shown in Fig. 1. In real world, the data owners correspond to those sources or producers of the data such as websites with online user data or organizations with certain statistics. The data users can be any companies or individuals who either consume the data or exploit data to develop projects and to make profits. For example, a software company may need certain user record data to develop an APP. Often, in a data market, data users do not interact with the data owners directly due to the limited data collection, storage and processing capability of many data users. Instead, between data owners and users, there may exist data collectors who are able to collect, store and process massive datasets.

The collectors collect data from the owners through various methods such as web scraping for websites or direct inquiries to organizations with certain statistics. The specific data collection manner depends on the form of the data. After obtaining the (massive) data, the collectors store them and further process them to be more sanitary and user-friendly. Lastly, the collectors sell the data to users according to the different demands of users.

Different from prior works [12]–[14], we assume the existence of multiple data owners, collectors and users competing with each other, which is the case in reality as explained in Section I. This makes the problem more challenging because of the conflicting interests and selfishness of the data agents, which necessitates a framework different from the traditional auction theoretic approach in [12], [13] and contract theoretic approach in [14]. Next, we describe the data trading among the data agents and their utility functions in detail.

A. Data Owners

Suppose owner m (there are M data owners in total) entitles collector n to collect x_{mn} amount of data, which is the maximum amount of data that collector n can get from owner m. For instance, a website may give a data collector (e.g., a web scraper) access to a certain part of data in that website; an organization may allow a data collector to access certain records or statistics of the organization. Due to the exposure of its data, the owner m suffers a loss of $U_m(\mathbf{x}_m)$, where $\mathbf{x}_m = [x_{m1}, ..., x_{mn}]$. This loss may stem from compromise of privacy or leakage of lucrative information/technologies. For example, if a social network allows some of its users' data to be accessed by companies or researchers, its users' privacy will be compromised and the social network may lose popularity among online users.

We assume that the data here are exclusive, i.e., the same data can only be assigned to one collector and one user. For example, software companies (data users) may need tailored data (e.g., click-through rate of specific web pages or advertisements in order to monitor the users' feedback) to develop their own softwares or APPs. These data are useful only to this user and are useless for others, i.e., these data are exclusive. In Section V, we extend the proposed mechanism to non-exclusive data trading scenario, where the same data can be used by multiple users.

We assume that owner m has C_m amount of data in total. In real world, when the data exposure or leakage is tiny, the data owner may hardly suffer any loss. However, if the data exposure is severe, e.g., larger than a certain threshold, the privacy loss will increase faster and faster with the amount of data exposure. In order to capture this second order property of loss function of data owners, we assume that the loss function U_m is a convex function.

B. Data Collectors

Suppose collector n (there are N data collectors in total) collects y_{mn} data from owner m. Clearly, y_{mn} is no larger than x_{mn}. When it is strictly smaller than x_{mn}, the collector n does not collect all the authorized data from owner m due to the loss from collection efforts. We assume that the collecting procedure incurs a loss of $V_n(\mathbf{y}_n)$ for collector n, where $\mathbf{y}_n = [y_{1n}, ..., y_{mn}]^T$. In real world, the collecting procedure can be data scraping from websites or direct inquiry to organizations etc., depending on the form and availability of the data. The collection and basic trimming/processing of the massive data need significant efforts of the collectors. In addition, the storage of the massive datasets also necessitate lots of apparatuses and devices. All of these contribute to the loss of the data collectors. Often, with the increase of the data to be collected, the difficulty (and hence efforts) of data collection increases faster and faster due to reasons such as the limitations on the internet connections and computers' processing speed (if the data amount is huge, collectors need to greatly enhance their internet connections or computer devices, which is costly). Therefore, we assume V_n is a convex function.

C. Data Users

Lastly, data user l (there are L data users in total) buys z_{nl} amount of data from collector n. The gain of user l is $W_l(\mathbf{z}_l)$, where $\mathbf{z}_l = [z_{1l}, ..., z_{Nl}]^T$. For instance, by exploiting the user feedback data such as click-through rate, a software/APP developer can enhance its product and makes more profits. As per conventions of the resource allocation literature, the gain function W_l is assumed to be a concave function.

D. Social Welfare Maximization

As the interests of the data agents conflict with each other (e.g., the data owners want to sell the data with high price while the data collector wants to gain the data at low cost) and the data agents are selfish, a system designer is needed to coordinate the agents' behaviors to maximize overall system efficiency or *social welfare*, which is defined as the difference between the total gain of users and total loss of owners and collectors. The corresponding social welfare maximization problem**SWM** can be formulated as follows.

$$\text{Maximize}_{\mathbf{X},\mathbf{Y},\mathbf{Z}} - \sum_{m=1}^{M} U_m(\mathbf{x}_m) - \sum_{n=1}^{N} V_n(\mathbf{y}_n) + \sum_{l=1}^{L} W_l(\mathbf{z}_l) \tag{1}$$

$$\text{s.t.} \sum_{n=1}^{N} x_{mn} \le C_m,\ \forall m, \tag{2}$$

$$\sum_{l=1}^{L} z_{nl} \le \sum_{m=1}^{M} y_{mn},\ \forall n, \tag{3}$$

$$y_{mn} \le x_{mn},\ \forall m, n. \tag{4}$$

The first constraint is the total data constraint at each data owner. The second constraint is the data constraint at each collector where the total amount of sold data is no larger than the amount of total collected data. The third constraint means that the data collected by a collector n from an owner m is no bigger than the data that owner m entitles collector n to collect.

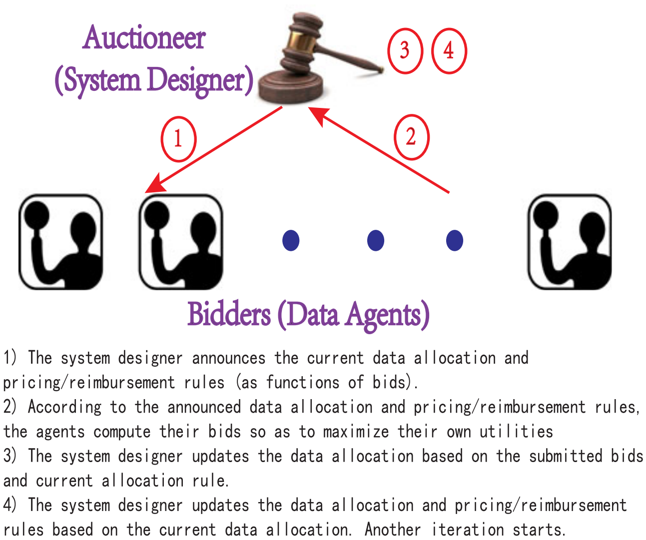


Fig. 2. An illustration of the proposed iterative auction mechanism, which iterates the four steps depicted in the figure.

SWM is a convex optimization problem and can be solved in a centralized manner by using state-of-the-art optimization toolbox such as CVX [20]. However, in real-world applications, we cannot directly solve the **SWM** to coordinate the data trading due to the following reasons.

1) First, data agents (data owners, collectors and users) are selfish and seek to maximize their own utilities instead of the social welfare. As a result, even if the system designer computes the socially optimal point by solving **SWM**, the optimal solution cannot be enforced given the selfishness of the data agents.
2) Second, the utility functions U, V, W are private information of the agents which is unknown to the system designer. Thereby, SWM cannot be solved at the system designer's side in a centralized fashion.

In order to elicit the private information of the agents and guide the selfish agents to cooperate to achieve social optimum, we resort to iterative auction mechanism [15]. The presumption of this mechanism is that the agents are price-takers, meaning that the each agent takes the announced prices as fixed and does not expect any impact of its action on the prices. This hypothesis holds when either (1) the agents have limited computational capability and thus limited rationality so that they do not consider the effects of their actions on pricing; or (2) the number of agents is large so that each agent has little influence on the prices.

III. Mechanism Design

In this section, we design an iterative auction mechanism for the data trading problem formulated in Section II. Our design goal is to guide the selfish agents to trade data at a socially optimal point while respecting each agent's private information, i.e., avoiding direct inquiry of the agents' utility functions. The proposed iterative auction mechanism is illustrated in Fig. 2. The system designer serves as the auctioneer and the data agents are the bidders. Analogous to many auction mechanisms in the literature [21], the agents submit *bids* to signal their valuations of the resources, or data in this context. The first step of the mechanism is that the system designer announces the data allocation and pricing/reimbursement rules to the agents. In the second step, based on these rules, each agent calculates and submits an appropriate bid in order to maximize her own utility in accordance with her selfishness. In the third step, the system designer computes the data allocation result according to the submitted bids and the data allocation rule. The aforementioned three steps are common in auction theory. The unique feature of iterative auction lies in the fourth step, in which the system designer adjusts the data allocation and pricing/reimbursement rules based on the data allocation results. Then, the system designer announces these new rules and another auction begins. This iterative process continues until the system designer observes convergence. In the following sections, we describe each step of the mechanism in more detail.

A. The System Designer's Problem

As explained in Section II, a difficulty for the system designer to solve the **SWM** is that the she is unaware of the loss and gain functions U, V, W, which are private information of the agents. Thus, the system designer has to replace these unknown functions with some known functions. In addition, denote the bid that owner m submits to the system designer by $\mathbf{s}_m = [s_{m1}, ..., s_{mn}] \succeq \mathbf{0}$, where $\succeq$ means componentwise inequality. Similarly, denote the bid of collector n by $\mathbf{t}_n = [t_{1n}, ..., t_{mn}]^T \succeq \mathbf{0}$ and the bid of user l by $\mathbf{r}_l = [r_{1l}, ..., r_{Nl}]^T \succeq \mathbf{0}$. The bids signal the agents' valuations of the data and should be incorporated into the loss and gain functions in the system designer's perspective. In the iterative auction mechanism, the system designer makes the following utility function replacements to avoid direct access of the private information of the agents:

$$U_m(\mathbf{x}_m) \leftarrow \sum_{n=1}^{N} \frac{s_{mn}}{2} x_{mn}^2, \tag{5}$$

$$V_n(\mathbf{y}_n) \leftarrow \sum_{m=1}^{M} \frac{t_{mn}}{2} y_{mn}^2, \tag{6}$$

$$W_l(\mathbf{z}_l) \leftarrow \sum_{n=1}^{N} r_{nl} \log z_{nl}. \tag{7}$$

Note that through these replacements, the convexity/concavity of the functions U, V, W are preserved. Then, the **SWM** is transformed into the following designer's allocation problem **DAP**.

$$\begin{aligned} \text{Maximize}_{\mathbf{X},\mathbf{Y},\mathbf{Z}} \ & \sum_{l=1}^{L}\sum_{n=1}^{N} r_{nl} \log z_{nl} - \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{s_{mn}}{2} x_{mn}^2 \\ & - \sum_{n=1}^{N}\sum_{m=1}^{M} \frac{t_{mn}}{2} y_{mn}^2 \\ \text{s.t.} \ & \text{the constraints } (2),\ (3) \text{ and } (4) \end{aligned}$$

Denote the dual variables associated with constraints (2), (3) and (4) by $\boldsymbol{\lambda} \in \mathbb{R}^M, \boldsymbol{\mu} \in \mathbb{R}^N, \boldsymbol{\eta} \in \mathbb{R}^{M\times N}$, respectively. The Lagrangian of **DAP** is:

$$L(\mathbf{X},\mathbf{Y},\mathbf{Z},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\eta}) = \sum_{m=1}^{M}\sum_{n=1}^{N}\frac{s_{mn}}{2}x_{mn}^2 + \sum_{n=1}^{N}\sum_{m=1}^{M}\frac{t_{mn}}{2}y_{mn}^2 - \sum_{l=1}^{L}\sum_{n=1}^{N} r_{nl}\log z_{nl} + \sum_{m=1}^{M}\lambda_m\left(\sum_{n=1}^{N}x_{mn} - C_m\right) + \sum_{n=1}^{N}\mu_n\left(\sum_{l=1}^{L}z_{nl} - \sum_{m=1}^{M}y_{mn}\right) + \sum_{m=1}^{M}\sum_{n=1}^{N}\eta_{mn}(y_{mn}-x_{mn}).$$

Thus, the Karush-Kuhn-Tucker (KKT) conditions of **DAP** can be written as follows.

$$\text{Primal Feasibility}: \sum_{n=1}^{N} x_{mn} \le C_m, \forall m, \tag{8}$$

$$\sum_{l=1}^{L} z_{nl} \le \sum_{m=1}^{M} y_{mn},\ \forall n, \tag{9}$$

$$y_{mn} \le x_{mn},\ \forall m,n, \tag{10}$$

$$\text{Dual Feasibility}: \boldsymbol{\lambda} \succeq \mathbf{0},\ \boldsymbol{\mu} \succeq \mathbf{0},\ \boldsymbol{\eta} \succeq \mathbf{0}, \tag{11}$$

$$\text{Complementary Slackness}: \tag{12}$$

$$\lambda_m\left(\sum_{n=1}^{N}x_{mn} - C_m\right) = 0, \forall m, \tag{13}$$

$$\mu_n\left(\sum_{l=1}^{L}z_{nl} - \sum_{m=1}^{M}y_{mn}\right) = 0, \forall n, \tag{14}$$

$$\eta_{mn}(y_{mn} - x_{mn}) = 0, \tag{15}$$

$$Stationarity: s_{mn}x_{mn} + \lambda_m - \eta_{mn} = 0, \forall m,n, \tag{16}$$

$$t_{mn}y_{mn} - \mu_n + \eta_{mn} = 0, \forall m,n, \tag{17}$$

$$-\frac{r_{nl}}{z_{nl}} + \mu_n = 0, \forall n,l. \tag{18}$$

From equations (16), (17) and (18), we obtain the *data allocation rule*:

$$x_{mn} = \frac{\eta_{mn}-\lambda_m}{s_{mn}}, y_{mn} = \frac{\mu_n - \eta_{mn}}{t_{mn}}, z_{nl} = \frac{r_{nl}}{\mu_n},\ \forall m,n,l. \tag{19}$$

The data allocation rule prescribes how the data are allocated given the submitted bids $\mathbf{S} = [s_{mn}]_{M\times N}, \mathbf{T} = [t_{mn}]_{M\times N}, \mathbf{R} = [r_{nl}]_{N\times L}$. The allocation rule is parameterized by the Lagrangian multipliers $\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\eta}$. Given a set of $\{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\eta}\}$, an allocation rule is defined according to Eq. (19), i.e., a relationship between the data allocation and the bids is specified. As stated in the first step of the mechanism in Fig. 2, besides data allocation rule, the system designer also needs to specify the data pricing/reimbursement rule, i.e., the price and reimbursement of data as functions of the bids of the agents. In other words, for owner m, given its bid $\mathbf{s}_m$, the system designer needs to reimburse $f_m(\mathbf{s}_m)$ amount of money to compensate her loss due to privacy compromise. Similarly, the system designer will reimburse $g_n(\mathbf{t}_n)$ amount of money to collector n given her bid $\mathbf{t}_n$. Furthermore, the system designer will charge user l $h_l(\mathbf{r}_l)$ amount of money given her bid $\mathbf{r}_l$. As a mechanism designer, we need to appropriately design the pricing/reimbursement functions f_m, g_n, h_l so that the data allocation will gradually converge to the socially optimal point, i.e., the optimal point of **SWM**. In the following sections, we specify how to design these pricing/reimbursement functions in detail.

B. Owners' Problems

For owner m, if she bids $\mathbf{s}_m$, she will get an reimbursement of $f_m(\mathbf{s}_m)$ as well as a loss of $U_m(\frac{\eta_{m1}-\lambda_m}{s_{m1}}, ..., \frac{\eta_{mn}-\lambda_m}{s_{mn}})$, according to the data allocation rule in Eq. (19). Hence, the utility maximization problem of owner m can be written as:

$$\text{Maximize}_{\mathbf{s}_m\succeq\mathbf{0}}\ f_m(\mathbf{s}_m) - U_m\left(\frac{\eta_{m1}-\lambda_m}{s_{m1}}, ..., \frac{\eta_{mn}-\lambda_m}{s_{mn}}\right). \tag{20}$$

The first order optimality condition of owner m's problem is:

$$\frac{\partial f_m(\mathbf{s}_m)}{\partial s_{mn}} + \frac{\partial U_m}{\partial x_{mn}}\frac{\eta_{mn}-\lambda_m}{s_{mn}^2} = 0, \forall n. \tag{21}$$

In order to design a suitable f_m such that the data allocation will converge to the socially optimal point, we need to compare Eq. (21) with the optimality condition of **SWM**. To this end, we write the Lagrangian of **SWM** as follows:

$$\tilde{L}(\mathbf{X},\mathbf{Y},\mathbf{Z},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\eta}) = \sum_{m=1}^{M}U_m(\mathbf{x}_m) + \sum_{n=1}^{N}V_n(\mathbf{y}_n) - \sum_{l=1}^{L}W_l(\mathbf{z}_l) + \sum_{m=1}^{M}\lambda_m\left(\sum_{n=1}^{N}x_{mn} - C_m\right) + \sum_{n=1}^{N}\mu_n \times\left(\sum_{l=1}^{L}z_{nl} - \sum_{m=1}^{M}y_{mn}\right) + \sum_{m=1}^{M}\sum_{n=1}^{N}\eta_{mn}(y_{mn}-x_{mn}).$$

The constraints of **SWM** and **DAP** are the same and the only difference is the objective function. Thus, in the KKT conditions of **SWM**, the primal feasibility, dual feasibility and complementary slackness conditions are the same as those of **DAP**, i.e., equations (8)–(15), while stationarity condition of **SWM** is:

$$\frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}} + \lambda_m - \eta_{mn} = 0,\ \forall m,n, \tag{22}$$

$$\frac{V_n(\mathbf{y}_n)}{\partial y_{mn}} - \mu_n + \eta_{mn} = 0,\ \forall m,n, \tag{23}$$

$$-\frac{W_l(\mathbf{z}_l)}{z_{nl}} + \mu_n = 0,\ \forall n,l. \tag{24}$$

Combining equations (21) and (22), we derive:

$$\frac{\partial f_m(\mathbf{s}_m)}{\partial s_{mn}} = \frac{\lambda_m - \eta_{mn}}{s_{mn}^2}\frac{\partial U_m}{\partial x_{mn}} = -\frac{(\lambda_m - \eta_{mn})^2}{s_{mn}^2}. \quad (25)$$

Therefore, we set the reimbursement rule of owner m to be $f_m(\mathbf{s}_m) = \sum_{n=1}^{N} \frac{(\lambda_m - \eta_{mn})^2}{s_{mn}}$.

C. Collectors' Problems

For collector n, if she bids $\mathbf{t}_n$, she will get a reimbursement of $g_n(\mathbf{t}_n)$ and a loss of $V_n(\frac{\mu_n - \eta_{1n}}{t_{1n}}, ..., \frac{\mu_n - \eta_{mn}}{t_{mn}})$. Thereby, the utility maximization problem of collector n is:

$$\text{Maximize}_{\mathbf{t}_n \succeq \mathbf{0}}\ g_n(\mathbf{t}_n) - V_n\left(\frac{\mu_n - \eta_{1n}}{t_{1n}}, ..., \frac{\mu_n - \eta_{mn}}{t_{mn}}\right). \quad (26)$$

The optimality condition of collector n's problem is:

$$\frac{\partial g_n(\mathbf{t}_n)}{\partial t_{mn}} + \frac{\partial V_n}{\partial y_{mn}}\frac{\mu_n - \eta_{mn}}{t_{mn}^2} = 0,\ \forall m. \quad (27)$$

Combining equations (23) and (27) yields:

$$\frac{\partial g_n(\mathbf{t}_n)}{\partial t_{mn}} = \frac{\eta_{mn} - mu_n}{t_{mn}^2}\frac{\partial V_n}{\partial y_{mn}} = -\frac{(\mu_n - \eta_{mn})^2}{t_{mn}^2}. \quad (28)$$

So, the reimbursement function of collector n should be $g_n(\mathbf{t}_n) = \sum_{m=1}^{M} \frac{(\mu_n - \eta_{mn})^2}{t_{mn}}$.

D. Users' Problems

For user l, if she bids $\mathbf{r}_l$, she will be charged $h_l(\mathbf{r}_l)$ and has a gain of $W_l(\frac{r_{1l}}{\mu_1}, ..., \frac{r_{Nl}}{\mu_N})$. Thus, the utility maximization problem of user l is:

$$\text{Maximize}_{\mathbf{r}_l \succeq \mathbf{0}}\ -h_l(\mathbf{r}_l) + W_l\left(\frac{r_{1l}}{\mu_1}, ..., \frac{r_{Nl}}{\mu_N}\right). \quad (29)$$

The optimality condition of user l's problem is:

$$-\frac{\partial h_l(\mathbf{r}_l)}{\partial r_{nl}} + \frac{\partial W_l}{\partial z_{nl}}\frac{1}{\mu_n} = 0, \forall n. \quad (30)$$

Combining equations (24) and (30) yields:

$$\frac{\partial h(\mathbf{r}_l)}{\partial r_{nl}} = \frac{1}{\mu_n}\frac{\partial W_l}{\partial z_{nl}} = \frac{1}{\mu_n}\cdot \mu_n = 1. \quad (31)$$

Thus, we design the price function of user l to be $h_l(\mathbf{r}_l) = \sum_{n=1}^{N} r_{nl}$.

E. Summary of Algorithm

The owners' problem (20), the collectors' problem (26) and the users' problem (29) together specify how the bids are chosen in the second stage of the mechanism in Fig. 2. Then, in the third stage, the system designer computes the new data allocation result based on these submitted bids and the data allocation rule in Eq. (19). In the fourth stage, we update the dual variables $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\eta}$ (or equivalently, update the data allocation rule and data pricing/reimbursement rule) by invoking the subgradient method:

Algorithm 1: The Proposed Iterative Auction Mechanism.

1: Initialize $\mathbf{X}^{(0)}, \mathbf{Y}^{(0)}, \mathbf{Z}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\eta}^{(0)}$ to be non-negative. Set the time index τ to be 0.
2: Repeat the following until convergence:
3: The system designer announces $\boldsymbol{\lambda}^{(\tau)}, \boldsymbol{\mu}^{(\tau)}, \boldsymbol{\eta}^{(\tau)}$.
4: $\tau \leftarrow \tau + 1$.
5: Each owner m solves its problem (20) to get $\mathbf{s}_m^{(\tau)}$.
6: Each collector n solves its problem (26) to get $\mathbf{t}_n^{(\tau)}$.
7: Each user l solves its problem (29) to get $\mathbf{r}_l^{(\tau)}$.
8: The system designer computes the new $\mathbf{X}^{(\tau)}, \mathbf{Y}^{(\tau)}, \mathbf{Z}^{(\tau)}$ according to the current allocation rule (19) and the submitted bids $\mathbf{S}^{(\tau)}$, $\mathbf{T}^{(\tau)}$ and $\mathbf{R}^{(\tau)}$.
9: The system designer updates the dual variables:

$$\lambda_m^{(\tau)} = \left(\lambda_m^{(\tau-1)} + \alpha\left(\sum_{n=1}^{N} x_{mn}^{(\tau)} - C_m\right)\right)^+, \forall m \quad (35)$$

$$\mu_n^{(\tau)} = \left(\mu_n^{(\tau-1)} + \alpha\left(\sum_{l=1}^{L} z_{nl}^{(\tau)} - \sum_{m=1}^{M} y_{mn}^{(\tau)}\right)\right)^+, \forall n \quad (36)$$

$$\eta_{mn}^{(\tau)} = \left(\eta_{mn}^{(\tau-1)} + \alpha\left(y_{mn}^{(\tau)} - x_{mn}^{(\tau)}\right)\right)^+, \forall m, n. \quad (37)$$

$$\lambda_m \leftarrow \left(\lambda_m + \alpha\left(\sum_{n=1}^{N} x_{mn} - C_m\right)\right)^+,\ \forall m \quad (32)$$

$$\mu_n \leftarrow \left(\mu_n + \alpha\left(\sum_{l=1}^{L} z_{nl} - \sum_{m=1}^{M} y_{mn}\right)\right)^+,\ \forall n \quad (33)$$

$$\eta_{mn} \leftarrow (\mu_{mn} + \alpha(y_{mn} - x_{mn}))^+, \forall m, n, \quad (34)$$

where $\alpha > 0$ is the step length and $x^+ = \max\{x, 0\}$. The proposed iterative auction mechanism is summarized in Algorithm 1. We remark that Algorithm 1 is a distributed algorithm: each data agent solves its own utility maximization problem in a parallel manner and the interactions between the agents. Algorithm 1 clearly resolves the two difficulties for directly solving **SWM** in Section II-D: (i) each agent maximizes her own utility in accordance with her selfishness; (ii) the system designer does not direct access the private information of the agents, i.e., the loss/gain functions U, V, W. Instead the system designer gradually and implicitly elicits this information through iterative auctions.

IV. Convergence and Economic Properties of the Mechanism

In this section, we theoretically show that the proposed iterative auction mechanism for data trading can indeed converge to the socially optimal operating point, i.e., the optimal point of **SWM**. Moreover, we prove that the mechanism has two appealing economic properties, i.e., individual rationality and weakly balanced budget, which makes the mechanism economically viable.

A. Convergence Analysis

When designing the mechanism in Section III, we make a connection between the data allocation rule, the optimality condition of each agent's utility maximization problem and the KKT conditions of **SWM**. Intuitively, the mechanism should guide the data allocation towards the solution of **SWM**. In this section, we rigorously demonstrate this convergence result. To make the analysis tractable, we assume that the step size α in the update of dual variables (35)–(36) and (37) is very small, which is a reasonable assumption in the literature of subgradient method in optimization theory [22] and LMS algorithm in adaptive signal processing [23]. Thus, we can approximate Algorithm 1 with a continuous-time version by taking the time slot to be α. From Eq. (35), we know that λ_m is always non-negative. If $\lambda_m^{(\tau-1)} > 0$, since α is very small, the quantity inside the parenthesis of Eq. (35) is still positive. Thus, $\lambda_m^{(\tau)} = \lambda_m^{(\tau-1)} + \alpha(\sum_{n=1}^N x_{mn}^{(\tau)} - C_m)$. Noting that the time slot length is α, a small positive number, we have $\frac{d\lambda_m}{d\tau} = \sum_{n=1}^N x_{mn} - C_m$. If $\lambda_m^{(\tau-1)} = 0$, we can similarly derive that $\frac{d\lambda_m}{d\tau} = (\sum_{n=1}^N x_{mn} - C_m)^+$. Define the notation (for $x, y \in \mathbb{R}$ and $y \geq 0$):

$$(x)_y^+ = \begin{cases} x, \text{ if } y > 0, \\ x^+, \text{ if } y = 0. \end{cases} \tag{38}$$

Then, we have:

$$\frac{d\lambda_m}{d\tau} = \left(\sum_{n=1}^N x_{mn} - C_m\right)^+_{\lambda_m}. \tag{39}$$

Similarly, we have:

$$\frac{d\mu_n}{d\tau} = \left(\sum_{l=1}^L z_{nl} - \sum_{m=1}^M y_{mn}\right)^+_{\mu_n}, \tag{40}$$

$$\frac{d\eta_{mn}}{d\tau} = (y_{mn} - x_{mn})^+_{\eta_{mn}}. \tag{41}$$

Now, we are ready to state the convergence result.

Theorem 1: Suppose the step size α in Algorithm 1 is small enough. Then, the data allocation $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ of Algorithm 1 converges to the optimal point of **SWM**. Moreover, the dual variables $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\eta})$ of Algorithm 1 converge to the dual optimal point of **SWM**.

Proof: Denote the dual optimal point of **SWM** by $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\eta}^*)$. Define the Lyapunov function:

$$H(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\eta}) = \frac{1}{2}\sum_{m=1}^M (\lambda_m - \lambda_m^*)^2 + \frac{1}{2}\sum_{n=1}^N (\mu_n - \mu_n^*)^2 \tag{42}$$

$$+\frac{1}{2}\sum_{m=1}^M \sum_{n=1}^N (\eta_{mn} - \eta_{mn}^*)^2. \tag{43}$$

Taking derivative of Z with respect to the (continuous) time τ yields:

$$\frac{dH}{d\tau} = \sum_{m=1}^M (\lambda_m - \lambda_m^*)\frac{d\lambda_m}{d\tau} + \sum_{n=1}^N (\mu_n - \mu_n^*)\frac{d\mu_n}{d\tau} + \sum_{m=1}^M\sum_{n=1}^N (\eta_{mn} - \eta_{mn}^*)\frac{d\eta_{mn}}{d\tau} \tag{44}$$

$$= \sum_{m=1}^M (\lambda_m - \lambda_m^*)\left(\sum_{n=1}^N x_{mn} - C_m\right)^+_{\lambda_m} + \sum_{n=1}^N (\mu_n - \mu_n^*)\left(\sum_{l=1}^L z_{nl} - \sum_{m=1}^M y_{mn}\right)^+_{\mu_n} + \sum_{m=1}^M\sum_{n=1}^N (\eta_{mn} - \eta_{mn}^*)(y_{mn} - x_{mn})^+_{\eta_{mn}} \tag{45}$$

$$\leq \sum_{m=1}^M (\lambda_m - \lambda_m^*)\left(\sum_{n=1}^N x_{mn} - C_m\right) + \sum_{n=1}^N (\mu_n - \mu_n^*)\left(\sum_{l=1}^L z_{nl} - \sum_{m=1}^M y_{mn}\right) + \sum_{m=1}^M\sum_{n=1}^N (\eta_{mn} - \eta_{mn}^*)(y_{mn} - x_{mn}), \tag{46}$$

where we use equations (39)–(40) and (41) to get Eq. (45). The reason of inequality (46) is as follows. If $\lambda_m = 0$, then $(\sum_{n=1}^N x_{mn} - C_m)^+_{\lambda_m} = (\sum_{n=1}^N x_{mn} - C_m)^+ \geq \sum_{n=1}^N x_{mn} - C_m$. Since $\lambda_m - \lambda_m^* = -\lambda_m^* \leq 0$, we have $(\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)^+_{\lambda_m} \leq (\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)$. If If $\lambda_m > 0$, we evidently have $(\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)^+_{\lambda_m} = (\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)$. In all, we always have $(\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)^+_{\lambda_m} \leq (\lambda_m - \lambda_m^*)(\sum_{n=1}^N x_{mn} - C_m)$ and similar inequalities hold for the other two terms in (45), leading to inequality (46). In Step 5, the optimal point of the problem (20) should satisfy the optimality condition (21). Noting the form of the reimbursement function f we design in Section III-B, we have:

$$-\frac{(\lambda_m - \eta_{mn})^2}{s_{mn}^2} + \frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}}\frac{\eta_{mn} - \lambda_m}{s_{mn}^2} = 0, \tag{47}$$

which leads to:

$$\lambda_m = \eta_{mn} - \frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}}. \tag{48}$$

Similarly, from the optimality condition (27), we get

$$\mu_n = \eta_{mn} + \frac{\partial V_n(\mathbf{y}_n)}{\partial y_{mn}}. \tag{49}$$

And from the optimality condition (30), we obtain:

$$\mu_n = \frac{\partial W_l(\mathbf{z}_l)}{\partial z_{nl}}. \tag{50}$$

Denote the optimal point of **SWM** by $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$. Since SWM is a convex optimization problem, KKT condition is necessary and sufficient for optimality. Hence, the primal optimal point $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$ together with dual optimal point $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\eta}^*)$ should satisfy the stationarity condition (22), (23) and (24), which can be further rewritten as:

$$\lambda_m^* = \eta_{mn}^* - \frac{\partial U_m(\mathbf{x}_m^*)}{\partial x_{mn}}, \tag{51}$$

$$\mu_n^* = \eta_{mn}^* + \frac{\partial V_n(\mathbf{y}_n^*)}{\partial y_{mn}}, \tag{52}$$

$$\mu_n^* = \frac{\partial W_l(\mathbf{z}_l^*)}{\partial z_{nl}}. \tag{53}$$

Hence, according to equations (48) and (51), we have:

$$\begin{aligned}&\sum_{m=1}^{M}(\lambda_m - \lambda_m^*)\left(\sum_{n=1}^{N} x_{mn} - \sum_{n=1}^{N} x_{mn}^*\right)\\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\eta_{mn} - \frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}} - \eta_{mn}^* + \frac{\partial U_m(\mathbf{x}_m^*)}{\partial x_{mn}}\right)\\ &\times (x_{mn} - x_{mn}^*),\end{aligned} \tag{54}$$

which can be further rewritten as:

$$\begin{aligned}&\sum_{m=1}^{M}(\lambda_m - \lambda_m^*)\left(\sum_{n=1}^{N} x_{mn} - \sum_{n=1}^{N} x_{mn}^*\right)\\ &+ \sum_{m=1}^{M}\sum_{n=1}^{N}(\eta_{mn} - \eta_{mn}^*)(x_{mn}^* - x_{mn})\\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\frac{\partial U_m(\mathbf{x}_m^*)}{\partial x_{mn}} - \frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}}\right)(x_{mn} - x_{mn}^*).\end{aligned} \tag{55}$$

Similarly, from equations (50) and (53), we obtain:

$$\begin{aligned}&\sum_{n=1}^{N}(\mu_n - \mu_n^*)\left(\sum_{l=1}^{L} z_{nl} - \sum_{l=1}^{L} z_{nl}^*\right)\\ &= \sum_{n=1}^{N}\sum_{l=1}^{L}\left(\frac{\partial W_l(\mathbf{z}_l)}{\partial z_{nl}} - \frac{\partial W_l(\mathbf{z}_l^*)}{\partial z_{nl}}\right)(z_{nl} - z_{nl}^*).\end{aligned} \tag{56}$$

And combining equations (49) and (52) yields:

$$\begin{aligned}&\sum_{n=1}^{N}(\mu_n - \mu_n^*)\left(\sum_{m=1}^{M} y_{mn}^* - \sum_{m=1}^{M} y_{mn}\right)\\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\eta_{mn} + \frac{\partial V_n(\mathbf{y}_n)}{\partial y_{mn}} - \eta_{mn}^* - \frac{\partial V_n(\mathbf{y}_n^*)}{\partial y_{mn}}\right)\\ &\times (y_{mn}^* - y_{mn}),\end{aligned} \tag{57}$$

which can be rewritten as:

$$\begin{aligned}&\sum_{n=1}^{N}(\mu_n - \mu_n^*)\left(\sum_{m=1}^{M} y_{mn}^* - \sum_{m=1}^{M} y_{mn}\right)\\ &+ \sum_{m=1}^{M}\sum_{n=1}^{N}(\eta_{mn} - \eta_{mn}^*)(y_{mn} - y_{mn}^*)\\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\frac{\partial V_n(\mathbf{y}_n)}{\partial y_{mn}} - \frac{\partial V_n(\mathbf{y}_n^*)}{\partial y_{mn}}\right)(y_{mn}^* - y_{mn}).\end{aligned} \tag{58}$$

Moreover, since the primal optimal point $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$ together with dual optimal point $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\eta}^*)$ should satisfy the KKT conditions of **SWM**, including conditions (8)–(15) (this part of KKT conditions coincides with that of **DAP**), from the complimentary slackness conditions, we have:

$$\lambda_m^*\left(\sum_{n=1}^{N} x_{mn}^* - C_m\right) = 0, \tag{59}$$

$$\mu_n^*\left(\sum_{l=1}^{L} z_{nl}^* - \sum_{m=1}^{M} y_{mn}^*\right) = 0, \tag{60}$$

$$\eta_{mn}^*(y_{mn}^* - x_{mn}^*) = 0. \tag{61}$$

Further notice that $\lambda_m, \mu_n, \eta_{mn} \geq 0$ and $\sum_{n=1}^{N} x_{mn}^* \leq C_m, \sum_{l=1}^{L} z_{nl}^* \leq \sum_{m=1}^{M} y_{mn}^*$. Thus, we get:

$$(\lambda_m - \lambda_m^*)\left(\sum_{n=1}^{N} x_{mn}^* - C_m\right) \leq 0, \tag{62}$$

$$(\mu_n - \mu_n^*)\left(\sum_{l=1}^{L} z_{nl}^* - \sum_{m=1}^{M} y_{mn}^*\right) \leq 0, \tag{63}$$

$$(\eta_{mn} - \eta_{mn}^*)(y_{mn}^* - x_{mn}^*) \leq 0. \tag{64}$$

Adding the six equations and inequalities (55), (56), (58), (62), (63) and (64) gives:

$$\begin{aligned}&\sum_{m=1}^{M}(\lambda_m - \lambda_m^*)\left(\sum_{n=1}^{N} x_{mn} - C_m\right)\\ &\quad + \sum_{n=1}^{N}(\mu_n - \mu_n^*)\left(\sum_{l=1}^{L} z_{nl} - \sum_{m=1}^{M} y_{mn}\right)\\ &\quad + \sum_{m=1}^{M}\sum_{n=1}^{N}(\eta_{mn} - \eta_{mn}^*)(y_{mn} - x_{mn}),\\ &\leq \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\frac{\partial U_m(\mathbf{x}_m^*)}{\partial x_{mn}} - \frac{\partial U_m(\mathbf{x}_m)}{\partial x_{mn}}\right)(x_{mn} - x_{mn}^*)\\ &\quad + \sum_{n=1}^{N}\sum_{l=1}^{L}\left(\frac{\partial W_l(\mathbf{z}_l)}{\partial z_{nl}} - \frac{\partial W_l(\mathbf{z}_l^*)}{\partial z_{nl}}\right)(z_{nl} - z_{nl}^*)\\ &\quad + \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\frac{\partial V_n(\mathbf{y}_n)}{\partial y_{mn}} - \frac{\partial V_n(\mathbf{y}_n^*)}{\partial y_{mn}}\right)(y_{mn}^* - y_{mn})\\ &\leq 0\end{aligned} \tag{65}$$

The last inequality of (65) is due to the convexity/concavity of the functions U, V, W. Specifically, since U_m, V_n are convex functions and W_l is concave function, we have:

$$(\nabla U_m(\mathbf{x}_m^*) - \nabla U_m(\mathbf{x}_m))^T (\mathbf{x}_m^* - \mathbf{x}_m) \geq 0, \ \forall m, \quad (66)$$

$$(\nabla W_l(\mathbf{z}_l) - \nabla W_l(\mathbf{z}_l^*))^T (\mathbf{z}_l - \mathbf{z}_l^*) \leq 0, \ \forall l, \quad (67)$$

$$(\nabla V_n(\mathbf{y}_n) - \nabla V_n(\mathbf{y}_n^*))^T (\mathbf{y}_n - \mathbf{y}_n^*) \geq 0, \forall n. \quad (68)$$

Adding inequalities (66)–(67) and (68) together over all m, n, l yields the last inequality of (65). Combining the inequalities (46) and (65), we obtain $\frac{dH}{d\tau} \leq 0$. Thus, according to LaSalle's invariance principle [24], $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\eta})$ converges to $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\eta}^*)$. Comparing equations (48), (49) and (50) with equations (51)–(52) and (53), we conclude that $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ converges to $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$. ∎

B. Economic Properties

Implementation of the proposed iterative auction mechanism in real-world data trading market necessitates brilliant economic properties of the mechanism. In this section, we show that the proposed mechanism has appealing economic properties. First, the proposed mechanism is clearly *efficient* since it converges to the socially optimal point. Second, the proposed mechanism possesses the *incentive compatibility* property because in each auction iteration, each agent is maximizing her own utility selfishly. To ensure that each agent complies to the mechanism voluntarily, the mechanism needs to guarantee that every agent has non-negative utility, i.e., the mechanism should be *individually rational*. This is shown in the following proposition.

Proposition 1: Assume that $U_m(\mathbf{0}) = 0, V_n(\mathbf{0}) = 0, W_l(\mathbf{0}) = 0, \forall m, n, l$. Then, when Algorithm 1 converges, every data agent has non-negative utility, i.e., the proposed mechanism is individually rational.

Proof: As shown in Theorem 1, when Algorithm 1 converges, $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ becomes $(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*)$ and $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\eta})$ becomes $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\eta}^*)$. Thus, according to the allocation rule (19), the bids $(\mathbf{S}, \mathbf{T}, \mathbf{R})$ become $(\mathbf{S}^*, \mathbf{T}^*, \mathbf{R}^*)$ defined as follows:

$$s_{mn}^* = \frac{\eta_{mn}^* - \lambda_m^*}{x_{mn}^*}, \quad (69)$$

$$t_{mn}^* = \frac{\mu_n^* - \eta_{mn}^*}{y_{mn}^*}, \quad (70)$$

$$r_{nl}^* = z_{nl}^* \mu_n^*. \quad (71)$$

Since U_m is convex, we have:

$$0 = U_m(\mathbf{0}) \geq U_m(\mathbf{x}_m^*) + \nabla U_m(\mathbf{x}_m^*)^T (\mathbf{0} - \mathbf{x}_m^*), \quad (72)$$

which can be rewritten as:

$$\sum_{n=1}^{N} \frac{\partial U_m(\mathbf{x}_m^*)}{\partial x_{mn}} x_{mn}^* - U_m(\mathbf{x}_m^*) \geq 0. \quad (73)$$

By Eq. (51), we further derive:

$$\sum_{n=1}^{N} (\eta_{mn}^* - \lambda_m^*) x_{mn}^* - U_m(\mathbf{x}_m^*) \geq 0, \quad (74)$$

which by Eq. (69) can be written as:

$$\sum_{n=1}^{N} \frac{(\lambda_m^* - \eta_{mn}^*)^2}{s_{mn}^*} - U_m(\mathbf{x}_m^*) \geq 0. \quad (75)$$

Note that the left hand side is exactly the utility of owner m when Algorithm 1 converges. So, owner m has non-negative utility. Similarly, from the convexity of V_n, we have:

$$V_n(\mathbf{y}_n^*) \leq \sum_{m=1}^{M} y_{mn}^* \frac{\partial V_n(\mathbf{y}_n^*)}{\partial y_{mn}}, \quad (76)$$

which by equations (52) and (70) can be rewritten as:

$$\sum_{m=1}^{M} \frac{(\mu_n^* - \eta_{mn}^*)}{t_{mn}^*} - V_n(\mathbf{y}_n^*) \geq 0. \quad (77)$$

Notice that the left hand side is just the utility of collector n when Algorithm 1 converges. We thus assert that each collector has non-negative utility. From the concavity of W_l, we obtain:

$$W_l(\mathbf{z}_l^*) \geq \sum_{n=1}^{N} z_{nl}^* \frac{\partial W_l(\mathbf{z}_l^*)}{\partial z_{nl}}, \quad (78)$$

which by equations (53) and (71) is be written as:

$$-\sum_{n=1}^{N} r_{nl}^* + W_l(\mathbf{z}_l^*) \geq 0. \quad (79)$$

Hence, each user has non-negative utility. Overall, we conclude that the mechanism is individually rational. ∎

We can further show that the system designer has *weakly balanced budget*, i.e., the income (through the data reimbursement/pricing) of the system designer in the mechanism is non-negative when Algorithm 1 converges. In other words, the system designer does not need to inject any money into the data market in order to implement the mechanism.

Proposition 2: When Algorithm 1 converges, the income of the system designer through data reimbursement/pricing in the mechanism is non-negative. In other words, the mechanism has weakly balanced budget.

Proof: The income of the system designer through data reimbursement/pricing is:

$$\sum_{l=1}^{L} h_l(\mathbf{r}_l^*) - \sum_{m=1}^{M} f_m(\mathbf{s_m}^*) - \sum_{n=1}^{N} g_n(\mathbf{t}_n^*) \quad (80)$$

$$= \sum_{l=1}^{L} \sum_{n=1}^{N} r_{nl}^* - \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{(\lambda_m^* - \eta_{mn}^*)}{s_{mn}^*} - \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{(\mu_n^* - \eta_{mn}^*)^2}{t_{mn}^*} \quad (81)$$

$$= \sum_{l=1}^{L}\sum_{n=1}^{N} z_{nl}^{*}\mu_n^{*} - \sum_{m=1}^{M}\sum_{n=1}^{N} x_{mn}^{*}(\eta_{mn}^{*} - \lambda_m^{*}) - \sum_{n=1}^{N}\sum_{m=1}^{M} y_{mn}^{*}(\mu_n^{*} - \eta_{mn}^{*}) \tag{82}$$

$$= \sum_{m=1}^{M}\sum_{n=1}^{N} \eta_{mn}^{*}(y_{mn}^{*} - x_{mn}^{*}) + \sum_{n=1}^{N}\mu_n^{*}\left(\sum_{l=1}^{L} z_{nl}^{*} - \sum_{m=1}^{M} y_{mn}^{*}\right) + \sum_{m=1}^{M}\sum_{n=1}^{N} x_{mn}^{*}\lambda_m^{*} \tag{83}$$

$$\geq 0 \tag{84}$$

where Eq. (82) comes from equations (69)–(70) and (71). The reason of the last step is: $\eta_{mn}^{*}(y_{mn}^{*} - x_{mn}^{*}) = 0, \mu_n^{*}(\sum_{l=1}^{L} z_{nl}^{*} - \sum_{m=1}^{M} y_{mn}^{*}) = 0$ due to complimentary slackness (60) and (61) and $x_{mn}^{*} \geq 0, \lambda_m^{*} \geq 0$. ■

V. Extension to Non-Exclusive Data Trading

In previous sections, we assume that the data are exclusive, i.e., the same data can be dispensed to only one user and one collector. However, in many real-world data markets, the data can be *non-exclusive*, i.e., the same data can be allotted to multiple collectors and users. For example, many software/APP developers (data users) may want to access the same online data of some social network (data owner); or many researchers (data users) may want to use the same data from an organization (data owner) to conduct research. In this section, we formulate the data trading problem with non-exclusive data and extend the proposed mechanism in Section III to this scenario.

Since the same data can be distributed to multiple collectors, different collectors' data can overlap each other. To avoid purchasing the same data from different collectors, we assume that each user buys data from only one single collector. Equivalently, from the collectors' perspective, each collector n serves a set of users $\mathcal{L}_n$ and users in $\mathcal{L}_n$ only purchase data from collector n. For example, in real world, a data collection company may occupy the most of the share of the local market in some region and becomes the monopoly in the local region. Basically all data users in this region will purchase data only from this data collector. Note that the sets $\mathcal{L}_n, n = 1, ..., N$ are mutually exclusive and $\bigcup_{n=1}^{N} \mathcal{L}_n = \{1, ..., L\}$. Each user l purchases from its designated collector z_{ml} amount owner m's data. Other notations are the same as the exclusive data trading model in Section II. The social welfare maximization problem for non-exclusive data trading can be formulated as follows.

$$\text{Maximize}_{\mathbf{X},\mathbf{Y},\mathbf{Z}} - \sum_{m=1}^{M} U_m(\mathbf{x}_m) - \sum_{n=1}^{N} V_n(\mathbf{y}_n) + \sum_{l=1}^{L} W_l(\mathbf{z}_l) \tag{85}$$

$$\text{s.t. } y_{mn} \leq x_{mn},\ \forall m, n, \tag{86}$$

$$x_{mn} \leq C_m,\ \forall m, n, \tag{87}$$

$$z_{ml} \leq y_{mn},\ \forall m, n, l \in \mathcal{L}_n. \tag{88}$$

The first constraint means that the data collected by collectors should be no more than the data authorized by the owners. The second constraint is the data constraint at each owner. Instead of total data constraint in the exclusive data trading scenario, the data constraint becomes individual data constraint in the non-exclusive data trading scenario. The third constraint indicates that the data purchased by users are no greater than the data collected by collectors. Similar to the exclusive data trading scenario, it is inviable to directly solve this social welfare maximization problem and enforce the solution for the data agents. Hence, we go through similar procedures as in Section III to obtain an iterative auction mechanism which can achieve the social optimum while respecting agents' private information (their loss/gain functions) and selfishness. The mechanism is summarized in Algorithm 2 and the design details are omitted. In Algorithm 2, we denote the Lagrangian multipliers corresponding to constraints (86), (87) and (88) by $\boldsymbol{\mu} \in \mathbb{R}^{M\times N}, \boldsymbol{\lambda} \in \mathbb{R}^{M\times N}$ and $\boldsymbol{\eta} \in \mathbb{R}^{M\times L}$, respectively.

VI. Simulations and Real Data Experiments

In this section, we present simulations as well as real data experiments to validate the theoretical results for the proposed iterative auction mechanism. We consider both exclusive data trading and non-exclusive data trading.

A. Simulations

Consider a data market with $M = 2$ data owners, $N = 2$ data collectors and $L = 4$ data users. The total data amount of owners 1 and 2 are set to be 2 and 4, respectively. The owners' convex loss functions are defined as follows:

$$U_m(\mathbf{x}_m) = a_m\left(\sum_{n=1}^{2} e^{x_{mn}} - 2\right),\ m = 1, 2, \tag{98}$$

where $a_1 = 0.1, a_2 = 0.3$. The collectors' convex loss functions are defined as:

$$V_n(\mathbf{y}_n) = b_n \sum_{m=1}^{2} y_{mn}^2,\ n = 1, 2, \tag{99}$$

where $b_1 = 0.5, b_2 = 1$. The users' concave gain functions are:

$$W_l(\mathbf{z}_l) = c_l \sum_{n=1}^{2} \log(1 + z_{nl}),\ l = 1, 2, 3, 4, \tag{100}$$

where $c_1 = \frac{3}{2}, c_2 = \frac{7}{6}, c_3 = \frac{5}{6}, c_4 = \frac{1}{2}$.

We first consider the exclusive data trading scenario. We simulate the proposed iterative auction mechanism in Algorithm 1. In Fig. 3, we validate the convergence behavior of the mechanism. The relative error used in Fig. 3 is $\max\{\frac{||X-X^*||_F}{||X^*||_F}, \frac{||Y-Y^*||_F}{||Y^*||_F}, \frac{||Z-Z^*||_F}{||Z^*||_F}\}$, where $||\cdot||_F$ means the Frobenius norm. As guaranteed by Theorem 1, the mechanism converges to the socially optimal point, i.e., the mechanism is efficient. We further investigate the economic properties of the mechanism through simulations in Fig. 4. We report the utilities of the owner 1, collector 1 and user 1 as the algorithm gradually converges. As asserted in Proposition 1, the mechanism is individually rational: the three data agents in Fig. 4

Algorithm 2: The Iterative Auction Mechanism for Non-exclusive Data Trading.

1: Initialize $\mathbf{X}^{(0)}, \mathbf{Y}^{(0)}, \mathbf{Z}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\eta}^{(0)}$ to be non-negative. Set the time index τ to be 0.

2: Repeat the following until convergence:

3: The system designer announces $\boldsymbol{\lambda}^{(\tau)}, \boldsymbol{\mu}^{(\tau)}, \boldsymbol{\eta}^{(\tau)}$.

4: $\tau \leftarrow \tau + 1$.

5: Each owner m solves the following problem to get $\mathbf{r}_m^{(\tau)}$:

$$\text{Maximize}_{\mathbf{r}_m \succeq \mathbf{0}} \sum_{n=1}^{N} \frac{\left(\lambda_{mn}^{(\tau)} - \mu_{mn}^{(\tau)}\right)^2}{r_{mn}} - U_m\left(\frac{\mu_{m1}^{(\tau)} - \lambda_{m1}^{(\tau)}}{r_{m1}}, \ldots, \frac{\mu_{mn}^{(\tau)} - \lambda_{mn}^{(\tau)}}{r_{mn}}\right). \tag{89}$$

6: Each collector n solves the following problem to get $\mathbf{s}_n^{(\tau)}$:

$$\text{Maximize}_{\mathbf{s}_n \succeq \mathbf{0}} \sum_{m=1}^{M} \frac{\left(\sum_{l\in\mathcal{L}_n} \eta_{ml}^{(\tau)} - \mu_{mn}\right)^2}{s_{mn}} - V_n\left(\frac{\sum_{l\in\mathcal{L}_n} \eta_{1l}^{(\tau)} - \mu_{1n}^{(\tau)}}{s_{1n}}, \ldots, \frac{\sum_{l\in\mathcal{L}_n} \eta_{Ml}^{(\tau)} - \mu_{mn}^{(\tau)}}{s_{mn}}\right). \tag{90}$$

7: Each user l solves the following problem to get $\mathbf{t}_l^{(\tau)}$:

$$\text{Maximize}_{\mathbf{t}_l \succeq \mathbf{0}} - \sum_{m=1}^{M} t_{ml} + W_l\left(\frac{t_{1l}}{\eta_{1l}^{(\tau)}}, \ldots, \frac{t_{Ml}}{\eta_{Ml}^{(\tau)}}\right). \tag{91}$$

8: The system designer computes the new $\mathbf{X}^{(\tau)}, \mathbf{Y}^{(\tau)}, \mathbf{Z}^{(\tau)}$ according to:

$$x_{mn}^{(\tau)} = \frac{\mu_{mn}^{(\tau)} - \lambda_{mn}^{(\tau)}}{r_{mn}^{(\tau)}}, \tag{92}$$

$$y_{mn}^{(\tau)} = \frac{\sum_{l\in\mathcal{L}_n} \eta_{ml}^{(\tau)} - \mu_{mn}^{(\tau)}}{s_{mn}^{(\tau)}}, \tag{93}$$

$$z_{ml}^{(\tau)} = \frac{t_{ml}^{(\tau)}}{\eta_{ml}^{(\tau)}}. \tag{94}$$

9: The system designer updates the dual variables:

$$\lambda_{mn}^{(\tau)} = \left(\lambda_{mn}^{(\tau-1)} + \alpha\left(x_{mn}^{(\tau)} - C_m\right)\right)^+, \forall m, n, \tag{95}$$

$$\mu_{mn}^{(\tau)} = \left(\mu_{mn}^{(\tau-1)} + \alpha\left(y_{mn}^{(\tau)} - x_{mn}^{(\tau)}\right)\right)^+, \forall m, n, \tag{96}$$

$$\eta_{ml}^{(\tau)} = \left(\eta_{ml}^{(\tau-1)} + \alpha\left(z_{ml}^{(\tau)} - y_{mn}^{(\tau)}\right)\right)^+, \forall m, n, l \in \mathcal{L}_n. \tag{97}$$

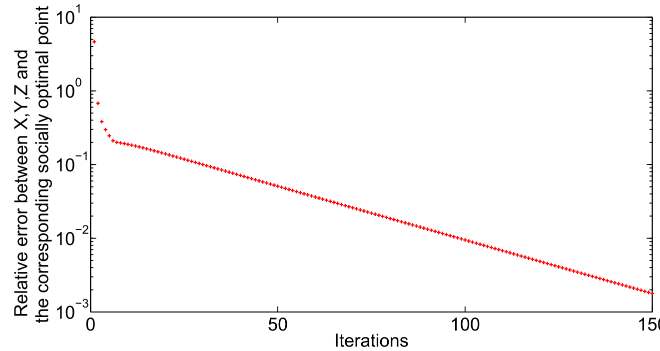


Fig. 3. Convergence of the iterative auction mechanism to the socially optimal point, i.e., the optimal point of **SWM**.

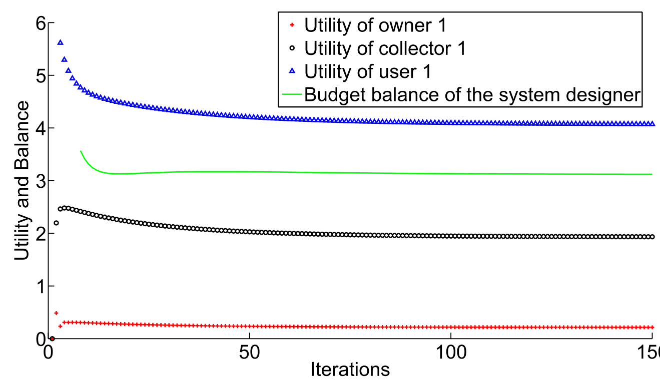


Fig. 4. The utilities of owner 1, collector 1 and user 1 and the budget balance (income) of the system designer.

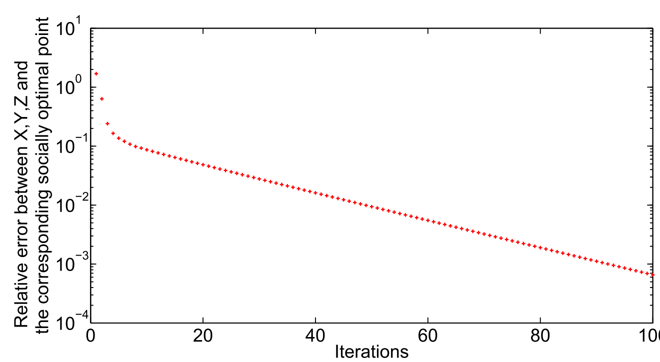
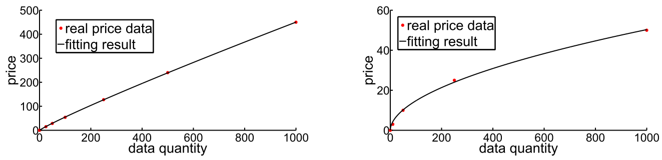


Fig. 5. Convergence of the iterative auction mechanism to the socially optimal point: non-exclusive data trading.

have non-negative utilities when the algorithm converges. Furthermore, we show the budget balance (income) of the system designer and find that as assured by Proposition 2, the budget balance is non-negative when the algorithm converges. Next, we turn to the non-exclusive data trading scenario. We set $\mathcal{L}_1 = \{1, 2\}, \mathcal{L}_2 = \{3, 4\}$. Other simulation setup remains unchanged and we simulate the iterative auction mechanism in Algorithm 2. As exhibited in Fig. 5, the mechanism still converges to the socially optimal point.



(a) (b)

Fig. 6. Fitting the real-world data price. (a) Wealth score data price: $y = 0.821x^{0.9131}$. (b) Text analytics data price: $y = 1.267x^{0.5329}$.

B. Real Data Experiments

In this section, we use real data to get the loss/gain functions of the data agents and investigate the performance of the proposed mechanism on them. We still consider a data market with $M = 2$ owners, $N = 2$ collectors and $L = 4$ users. We first use real data prices to estimate the users' gain functions. To this end, we fit the prices of the two datasets, namely the wealth score dataset and the text analytics dataset, in the Microsoft Azure Marketplace [25] (a data trading platform) with the function $y = ax^b$. The fitting results are shown in Fig. 6, which are very accurate. The sum of these two price functions can be regarded as the mean user gain function. To introduce heterogeneity into users' gain functions, we multiple a coefficient onto this mean user gain to get individual users' gains as follows:

$$W_l(\mathbf{z}_l) = c'_l \sum_{n=1}^{2} \alpha_n z_{nl}^{\beta_n}, \; l = 1, 2, 3, 4 \tag{101}$$

where $\alpha_1 = 0.821, \alpha_2 = 1.267, \beta_1 = 0.9131, \beta_2 = 0.5329, c'_1 = 1/2, c'_2 = 5/6, c'_3 = 7/6, c'_4 = 3/2$.

Next, we estimate the owners' loss functions. In [14], a relationship between the information loss and the privacy breach level in anonymization is obtained from real data [26]. Specifically, the privacy leakage is quantified by the k-anonymity, which means that the probability that an individual item being re-identified by an attacker is no higher than $1/k$. Thus, $1/k$ can be regarded as the loss of the data owner. (total data amount − IL) can be regarded as the effective amount of data obtained by a collector, where $\mathtt{IL}$ means the information loss. The relationship between k and $\mathtt{IL}$ is estimated to be $\mathrm{IL} = -0.4804k^{-0.2789} + 0.7883$, which can be rewritten as $1/k = (2.0816(0.7883 - \mathrm{IL}))^{3.5855}$. We set 0.7883 to be the total amount of data and thus $y = (2.0816x)^{3.5855}$ can be regarded as the average owners' loss function. By varying the coefficients, we introduce heterogeneity to the loss function and finally set:

$$U_m(\mathbf{x}_m) = a'_m \sum_{n=1}^{2} (\theta_n x_{mn})^{3.5855}, \; m = 1, 2, \tag{102}$$

where $\theta_1 = 1.5816, \theta_2 = 2.5816, a'_1 = 5, a'_2 = 15$. As for the collectors' loss functions V_n, it is hard to find corresponding real data and we directly use quadratic functions in simulation setups for them. Other experiment setups are the same as those of simulations.

With the loss/gain functions estimated from real data, we test the performance of the proposed iterative auction mechanism. We first consider the exclusive data trading. The total data amounts of owner 1 and owner 2 are 0.25 and 0.5, respectively. As shown in Fig. 7, the mechanism still converges to the socially optimal point. In Fig. 8, we further observe that the individual rationality and weakly balanced budget still hold as the utilities of owner 1, collector 1 and user 1 as well as the budget balance of the system designer are all non-negative. Then, we change to the non-exclusive data trading and alter the total data amounts of owner 1 and owner 2 to be 0.2 and 0.4, respectively. We remark that the mechanism still converges to the socially optimal point, as illustrated in Fig. 9.

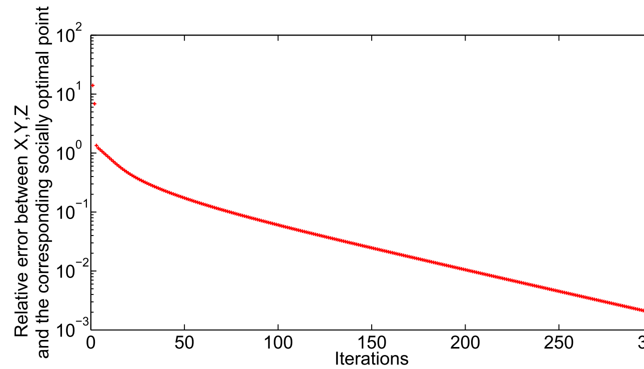


Fig. 7. Convergence of the iterative auction mechanism to the socially optimal point in real data experiment.

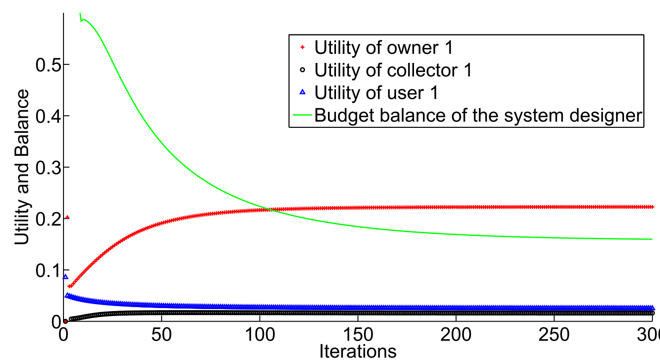


Fig. 8. The utilities of owner 1, collector 1 and user 1 and the budget balance (income) of the system designer in real data experiment.

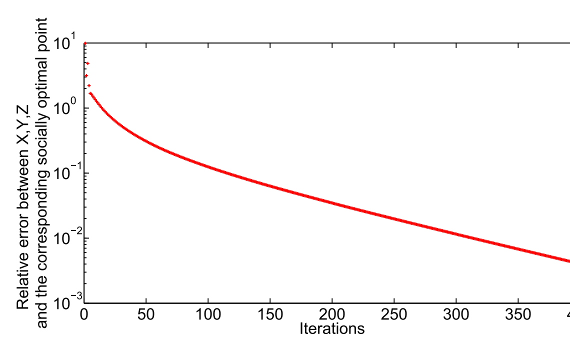


Fig. 9. Convergence of the iterative auction mechanism to the socially optimal point in real data experiment: non-exclusive data trading.

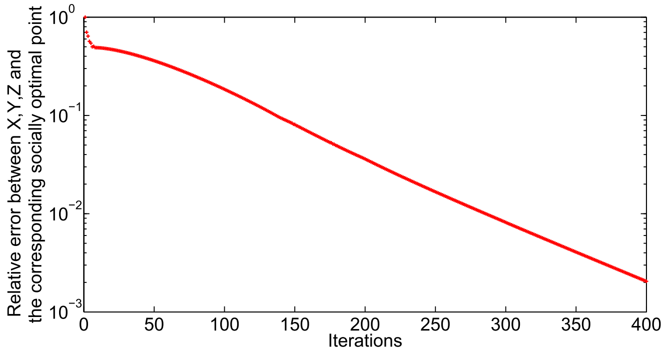


Fig. 10. Convergence of the iterative auction mechanism to the socially optimal point in the comparison experiment.

Lastly, we endeavor to compare the proposed iterative auction mechanism with the contract-theoretic approach in [14]. The model of [14] consists of multiple data owners and one single data collector without the notion of data users. To accommodate to this, we consider $M = 4$ owners, $N = 1$ collector and $L = 1$ user in our model. As per setups of real data experiments, we set the loss function of owners to be:

$$U_m(x_m) = a''_m (2.0816 x_m)^{3.5855}, m = 1, 2, 3, 4, \qquad (103)$$

where $a''_1 = 5, a''_2 = \frac{25}{3}, a''_3 = \frac{35}{3}, a''_4 = 15$. The total data amount of each owner is 0.08. Moreover, we set the gain function of the single user to be:

$$W_1(z_1) = 0.82105 z_1^{0.5329}. \qquad (104)$$

The loss function of the single data collector still takes the quadratic form previously used, i.e., $V_1(\mathbf{y}_1) = \frac{1}{2}\mathbf{y_1}^T\mathbf{y_1}$. Since the model in [14] only considers linear owner loss, we use $\tilde{U}_m(x_m) = 2.0816 a''_m x_m$ for [14]. Besides, the model in [14] sets the collector's gain to be a square root function. Hence, we use $\tilde{W}_1(z_1) = 0.82105 z_1^{0.5}$ for [14]. Note that the collector in [14] plays the role of end user and we translates that into the user in our model. In the model of [14], we need to specify a required total amount of data, i.e., $q_{\text{req}} = \sum_{m=1}^{M} x_m$, which we set to be 0.16, i.e., the half of the sum of total data amounts of all the owners. We first simulate the proposed iterative auction mechanism, which still converges to the socially optimal point, as illustrated in Fig. 10. The socially optimal point is $X = Y = [0.08, 0.08, 0.074, 0.074]^T, Z = 0.3015$ and the optimal social welfare (which is obtained by the proposed mechanism) is 0.373. Then, we simulate the contract-theoretic approach of [14], which gives the data allocation $X = Y = [0.080.0800]^T, Z = 0.16$ and a social welfare of 0.2812. Thus, we observe that the proposed mechanism can achieve a higher social welfare than [14].

According to the experiments and simulations, a practical issue of the proposed iterative auction mechanism is that it may need hundreds of iterations to converge. This requires the bidders (agents) to bid for hundreds of times. A common solution to this issue is to equip each bidder with some bidding software, which can automatically bid for the agent according to some preset bidding rule such as the one specified in the proposed iterative auction mechanism. With the help of such bidding softwares, the bidding processes can be very fast and accomplish hundreds of iterations quickly, making the proposed mechanism practical. In fact, fast iterative bidding with the assist of bidding softwares is already used in practice such as the eBay auction.

VII. Conclusion

In this paper, we study the data trading problem with multiple data owners, collectors and users. We present an iterative auction mechanism to guide the selfish agents to behave in a socially optimal way without direct access of their private information. We theoretically prove the convergence as well as economic properties (individual rationality and weakly balanced budget) of the mechanism. Simulations and real data experiments are carried out to confirm the theoretical properties of the proposed mechanism.

References

[1] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[2] J. Huang, V. G. Subramanian, R. Agrawal, and R. A. Berry, "Downlink scheduling and resource allocation for ofdm systems," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 1, pp. 288–296, Jan. 2009.

[3] Z. Han, Z. Ji, and K. J. R. Liu, "Fair multiuser channel allocation for ofdma networks using nash bargaining solutions and coalitions," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1366–1376, Aug. 2005.

[4] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 77–97, Sep. 2009.

[5] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.

[6] Z. Ji and K. J. R. Liu, "Cognitive radios for dynamic spectrum access-dynamic spectrum sharing: A game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, May 2007.

[7] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[8] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 940–951, May 2013.

[9] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2011, pp. 1–8.

[10] A. Roth, "Buying private data at auction: The sensitive surveyor's problem," *ACM SIGecom Exchanges*, vol. 11, no. 1, pp. 1–8, 2012.

[11] C. Dwork, "Differential privacy: A survey of results," in *Proc. Int. Conf. Theory Appl. Models Comput.*, 2008, pp. 1–19.

[12] A. Ghosh and A. Roth, "Selling privacy at auction," *Games Econ. Behav.*, vol. 91, pp. 334–346, 2015.

[13] L. K. Fleischer and Y.-H. Lyu, "Approximately optimal auctions for selling privacy when costs are correlated with data," in *Proc. 13th ACM Conf. Electron. Commerce*, 2012, pp. 568–585.

[14] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu, "Privacy or utility in data collection? A contract theoretic approach," *IEEE J. Sel. Top. Signal Process.*, vol. 9, no. 7, pp. 1256–1269, Oct. 2015.

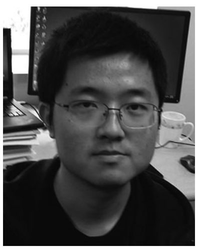
[15] F. P. Kelly, A. K. Maulloo, and D. K. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[16] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "A double-auction mechanism for mobile data-offloading markets," *IEEE/ACM Trans. Netw.*, vol. 23, no. 5, pp. 1634–1647, Oct. 2015.

[17] G. Iosifidis and I. Koutsopoulos, "Double auction mechanisms for resource allocation in autonomous networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 1, pp. 95–102, Jan. 2010.

[18] Y. Liu, M. Tao, and J. Huang, "An auction approach to distributed power allocation for multiuser cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 237–247, Jan. 2013.

[19] Q. Cao, Y. Jing, and H. V. Zhao, "Iterative double-auction-based power allocation in multiuser cooperative networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4298–4303, Sep. 2015.
[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[21] V. Krishna, *Auction Theory*. New York, NY, USA: Academic, 2009.
[22] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," in *Lecture Notes of EE392o*, Stanford University, 2004.
[23] S. S. Haykin, *Adaptive Filter Theory*. Noida, India: Pearson Educ. India, 2008.
[24] H. K. Khalil and J. Grizzle, *Nonlinear Systems*, vol. 3. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.
[25] (2016). [Online]. Available: https://datamarket.azure.com/browse/data
[26] A. Asuncion and D. Newman, "Uci Machine Learning Repository," 2007.



Xuanyu Cao received the Bachelor's degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013. He is currently working toward the Ph.D. degree from the University of Maryland, College Park, MD, USA. His generic research interest includes network science, including game-theoretic analysis of network users' behaviors, optimization and mechanism design for network resource allocation, fundamental performance limits of stochastic networks and signal processing over networks. He received the first prizes in Chinese National Mathematics Contest in 2007 and 2008, and also the Jimmy Lin scholarship from the Department of Electrical and Computer Engineering, University of Maryland.

Yan Chen (SM'14) received the Bachelor's degree from the University of Science and Technology of China, Chengdu, China, in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, in 2007, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2011. He was with the Origin Wireless Inc. as a Founding Principal Technologist. Since September 2015, he has been a full Professor with the University of Electronic Science and Technology of China. His research interests include multimedia, signal processing, game theory, and wireless communications.

He received multiple honors and awards, including the Best Student Paper Award at the IEEE ICASSP in 2016, the best paper award at the IEEE GLOBECOM in 2013, the Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, the University of Maryland in 2011, and the Chinese Government Award for outstanding students abroad in 2010.

K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of the University of Maryland, College Park, MD, USA, in 2007, where he is a Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on smart radios for smart life.

He received the 2016 IEEE Leon K. Kirchmayer Technical Field Award on graduate teaching and mentoring, the IEEE Signal Processing Society 2014 Society Award, and the IEEE Signal Processing Society 2009 Technical Achievement Award. Recognized by Thomson Reuters as a Highly Cited Researcher, he is a Fellow of AAAS. He is a member of IEEE Board of Director. He was the President of IEEE Signal Processing Society, where he has served as the Vice President Publications and Board of Governor. He has also served as the Editor-in-Chief of IEEE Signal Processing Magazine. He also received teaching and research recognitions from the University of Maryland including the university-level Invention of the Year Award; and the college-level Poole and Kent Senior Faculty Teaching Award, the Outstanding Faculty Research Award, and the Outstanding Faculty Service Award, all from A. James Clark School of Engineering.