A Hybrid Pricing Mechanism for Data Sharing in P2P-based Mobile Crowdsensing

Xiao Zeng, Lin Gao, Changkun Jiang, Tong Wang, Juan Liu, Baitao Zou

***Abstract*—Mobile crowdsensing (MCS) is becoming more and more popular with the increasing demand for various sensory data in many wireless applications. In the traditional server-client MCS system, a central server is often required to handle massive sensory data (e.g., collecting data from users who sense and dispatching data to users who request), hence it may incur severe congestion and high operational cost. In this work, we introduce a *peer-to-peer (P2P)* based MCS system, where the sensory data is stored in user devices locally and shared among users in an P2P manner. Hence, it can effectively alleviate the burden on the server, by leveraging the communication, computation, and cache resources of massive user devices. We focus on the economic incentive issue arising in the sharing of data among users in such a system, that is, how to incentivize users to share their sensed data with others. To achieve this, we propose a *data market*, together with a hybrid pricing mechanism, for users to *sell* their sensed data to others. We first study how would users choose the best way of obtaining desired data (i.e., sensing by themselves or purchasing from others). Then we analyze the user behavior dynamics as well as the data market evolution, by using the evolutionary game theory. We further characterize the users' equilibrium behaviors as well as the market equilibrium, and analyze the stability of the obtained equilibrium.**

I. Introduction

A. Background and Motivations

With the increasing demand for various sensory data in our daily life, a novel sensing scheme called *Mobile CrowdSensing (MCS)* [1] has become more and more popular in recent years. The key idea of MCS is to employ a large amount of user devices (e.g., smartphones) with various built-in sensors to collect the desired data, instead of deploying dedicated sensor networks. By crowdsourcing the capabilities of massive user devices, this novel sensing scheme can achieve a higher and more flexible sensing coverage with a lower deploying cost. Thus, it has attracted a wide range of applications in environment and community monitoring [2]–[6].

A basic MCS framework often consists of two parties: (i) a set of participating users (clients) for sensing the desired data by using their carried devices and (ii) a central platform (server) for providing the necessary control and organization for the system. In the traditional *server-client* MCS architecture [7]–[13], the server is also responsible for the necessary processing of data, such as collecting data from users who sense, storing and manufacturing data (e.g., data aggregation and mining), and dispatching the associated data to users who request. However, in a large system with massive data demand,

This work is supported by the National Natural Science Foundation of China (Grant No. 61771162).

Authors are with the Department of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen, China. Lin Gao is the *Corresponding Author*. Email: gaol@hit.edu.cn

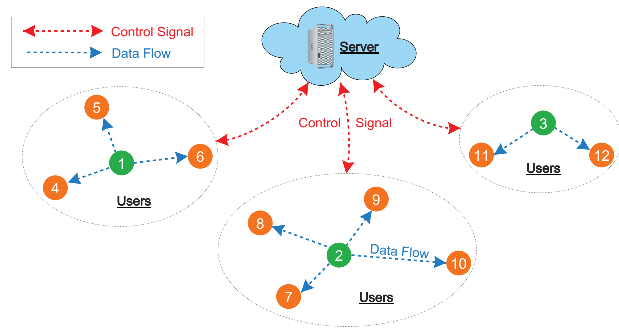

Fig. 1: Illustration of a P2P-based MCS System. Green users 1–3 share the sensory data with orange users 4–12.

such a server-client system may incur severe congestion and high operational cost on the server.

To this end, some researchers have started to develop the *peer-to-peer* (P2P) based MCS systems, e.g., MPSDataStore [14], SmartP2P [15], and LL-Net [16]. In a P2P-based MCS system, the server is no longer responsible for the transferring, saving, and processing of sensory data. Instead, the sensory data is processed and saved in user devices locally and shared among different users (whenever needed) in a P2P and on-demand manner. Moreover, the data sharing between users can be achieved via WiFi and Bluetooth (when they are locally connected) or Internet (when they are not locally connected). Clearly, in such a P2P-based MCS system, the server only needs to provide the necessary control and organization for the system. For example, the server needs to keep track of each user's IP address, data occupancy information (i.e., what data he has), and data request information (i.e., what data he wants), with which it can help a user find and connect to another user who has his desired data.

Figure 1 illustrates an example of a P2P-based MCS system, where green users $\{1–3\}$ share the sensory data with orange users $\{4–12\}$. That is, user 1 shares his sensory data with users $\{4–6\}$, user 2 shares his sensory data with users $\{7–10\}$, and user 3 shares his sensory data with users $\{11, 12\}$. The server only exchanges the necessary control signal with users (to help users find and connect to other users who have their desired data), while the data flows occur within users.

It is easy to see that *incentive* becomes one of the critical issues in such a P2P-based MCS system. The reason is that users need to consume some resources (e.g., bandwidth and energy) in order to share their data with others. Thus, without proper incentives, they may not be willing to share data with others. In [17], [18], Jiang *et al.* proposed a data sharing market to address the incentive issue in the data sharing among users. The key idea is to allow users to *sell* their sensed data to other users, instead of sharing data freely with others. However, they considered a simplified scenario where a user can share (sell) data with an *unlimited* number of users in a given time period. In practice, however, a user can only share data with a limited set of users due to, for example, the physical-layer link capacity constraint. In this work, we will study such a P2P-based MCS system with the *limited* data sharing.

Note that with the limited data sharing constraint, the market will become much more complicated due to the following reason. With the unlimited data sharing (as in [17], [18]), a user is guaranteed to obtain certain data (if he requests) as long as there exists one user owning the data. With the limited data sharing, however, he may not be able to obtain the data, as the users owning the data may not be able to share the data with all requesting users (e.g., when only few users own the data and many users request the data). This actually introduces certain uncertainty (regarding the supply of data) in the data market, and hence greatly complicates the problem.

B. Solution and Contributions

In this work, we study a general P2P-based MCS system with the *limited* data sharing, where the sensory data is shared among different users in a P2P manner (without passing through the server as shown in Figure 1), and each user can share his sensory data with a limited number of users. We focus on the *economic incentive* issue arising in the data sharing among users in such a system, that is, how to incentivize users to share their sensed data with others.

Inspired by [17], [18], we introduce a *data market* to address the above incentive issue, where users choosing to sense data can *sell* the sensed data to other users, instead of sharing data freely with others. Note that users who purchase data from others cannot resell the purchased data to others, for the purpose of copyright protection. We further propose a *hybrid pricing* scheme for the data market, where a user needs to pay a fixed wholesale price plus a varying price proportional to his achieved revenue when purchasing data from others. Clearly, such a hybrid pricing generalizes both the revenue sharing scheme and the wholesale pricing scheme.

In such a data market, users can obtain the desired data in two different ways: *sensing* by themselves and *purchasing* from others. For example, a user with a high sensing cost may choose to purchase data from others, while a user with a low sensing cost may choose to sense data by himself. Note that a user may choose to sense the data that he is not at all interested in, as he can potentially sell the data in the data market to obtain some profit. We are interested in the following key problems arising in such a market:

- *What is the best decision for each user, regarding the way of obtaining his desired data (i.e., sensing by himself or purchasing from others)?*
- *How will the user best decision (behavior) change over time and how will the whole market evolve?*
- *What is the equilibrium of the user behavior evolution and the market evolution?*

We study the above problems systematically by using the evolutionary game theory [19].[1] Specifically, we first analyze the user best decision (behavior), taking the limited data sharing constraint into consideration. Then we study the user behavior dynamics and the market evolution using evolutionary game, based on which we characterize the user equilibrium behaviors as well as the market equilibrium, and analyze the stability of the derived equilibrium.

We further analyze how the equilibrium changes with the hybrid pricing parameters, i.e., the fixed wholesale price and the varying price factor. We show that the social welfare under equilibrium first increases and then decreases with the fixed wholesale price and the varying price factor. The reason is that a smaller price will attract less users to sense data (and sell data) and drive more users to purchase data. Thus, with very small prices, the users who sense data cannot share (sell) their sensed data with all requesting users, hence causing certain social welfare loss. Note that this is quite different from the results in [17], [18] (with the unlimited data sharing), where the social welfare under equilibrium always decreases with both parameters. In summary, the key contributions of this work are as follows.

- *More Practical Model:* We consider a P2P-based MCS system with limited data sharing, where each user can only share his sensed data with a limited number of users in a given period. This model is more practical than the existing model in the literature.
- *Market Design and Game-theoretic Analysis:* We introduce a data sharing market, together with a hybrid data pricing scheme, and analyze the user behavior as well as the market equilibrium systematically, by using the evolutionary game theory. Such an equilibrium analysis can help us to understand how the market evolves and where it is likely to evolve to.
- *Observations and Insights:* Through theoretic analysis and numerical results, we obtain some new observations and insights that are different from those in the existing work with the unlimited data sharing. The most important one is that the social welfare under equilibrium first increases and then decreases with the fixed wholesale price and the varying price factor. This implies that a smaller price is no longer better from the social perspective. On the contrary, a relatively higher price is desired in order to attract a sufficient number of users to sense data.

The rest of the paper is organized as follows. In Section II, we present the system model. In Section III, we derive the socially optimal solution. In Section IV, we analyze the system from the game-theoretic perspective. We provide simulation results in Section V and finally conclude in Section VI.

II. System Model

A. Network Model

We consider a P2P-based MCS system with a set $\mathcal{N} = \{1, 2, \dots, N\}$ of participating users (clients), who can share

[1]Evolutionary game theory has been widely used in wireless networks for analyzing various dynamic scenarios, such as TV white space information market [20], [21] and WiFi community network [22], [23].

the sensed data with each others in a P2P manner (without passing through the server) as shown in Figure 1. We consider multiple types of location-based data, each associated with a particular location. Let $\mathcal{I} = \{1, 2, \dots, I\}$ denote the set of all locations (or data types). Each user can sense one or multiple locations, depending on his mobility, device type, and energy budget. Each location $i \in \mathcal{I}$ is associated with a weight ω_i, denoting the importance of the location. For example, a hotspot location often has a larger weight than a non-hotspot location. Moreover, different users may have different personal preferences for the same data (location), which is captured by a user-dependent data value. Let $v_{n,i}$ denotes the user-dependent value of data i for user n. Then, the *utility* of data i for user n is defined as the product of the data weight and the user-dependent data value, i.e., $\omega_i \cdot v_{n,i}$.

Each user can obtain his desired data through either the sensing of himself or the sharing of other users. The latter case may happen when the user is not able to sense the data (e.g., due to the mobility or device capability constraint), or when the user's sensing cost is very large (e.g., due to the energy budget constraint). The data sharing between users can be based on WiFi, Bluetooth, and Internet connections. To facilitate the data sharing among users, the server needs to keep track of each user's IP address, data occupancy information (e.g., what data he has), and data request information (e.g., what data he wants). Note that the server does not need to store and process the sensory data, that is, all the sensory data will be stored and processed in the user devices locally.

B. User Model

As mentioned previously, to obtain the data in a particular location, a user can choose to (i) act as a *sensor* and sense the data directly or (ii) act as a *requester* and request the data from others. Note that a user can also choose to not obtain the data, for example, when he is not interested in the data or when the cost of obtaining data is larger than the utility of the data for him. In this case, the user will act as an *alien* and neither sense nor request the data. More specifically,

- *Sensor:* A sensor acquires data via sensing directly, and hence will incur some sensing cost (e.g., energy consumption). Meanwhile, a sensor can share the sensed data with other users. As reward, he can ask for certain monetary payment from the users who get data from him.
- *Requester:* A requester acquires data from the sharing of other users, and hence needs to bear the data transfer cost (e.g., data uploading and downloading cost). Moreover, he may also need to share some benefit with the user who shares the data with him as reward.
- *Alien:* An alien neither senses the data, nor requests the data from others. This often occurs when the user is not interested in the data.

Due to the physical-layer link capacity constraint as well as other possible constraints, a sensor can only share data with a limited number of other users in a given time period. Moreover, for the purpose of copyright protection, requesters cannot share the obtained data to others. Let K denote the number of users that a sensor can share data with. When $K \to \infty$, our model degenerates to the traditional model with unlimited data sharing (e.g., those in [17], [18]). In this sense, our model generalizes the existing model in the literature.

We consider a general sensing cost model, where the same user may have different sensing costs for different data, and different users may have different sensing costs for the same data. Let $c_{n,i}$ denote the sensing cost of user n for data i. To simplify the later analysis while not affecting the meaningful insight, we assume that all users have the same average uploading cost s_u and average downloading cost s_d for any data. Thus, the average data transfer cost between a sensor and a requester is $s = s_u + s_d$. This cost will be fully beared by the requester.

C. Data Market

To provide necessary incentive for sensors to share their data with requesters, we introduce a *data market* as in [17], [18], where sensors *sell* data to requesters, instead of sharing data freely. Obviously, a proper data pricing scheme is the core of the data market. We propose a *hybrid pricing* scheme, which includes both the revenue sharing scheme [24] and the wholesale pricing scheme [25] as special cases. With the hybrid pricing, a requester needs to pay a fixed wholesale price plus a varying price proportional to his achieved revenue when purchasing data from others. Formally,

Definition 1 (Hybrid Pricing Mechanism)**.** *Suppose that a requester can obtain a total benefit b from the data shared by a sensor. Then, the payment of the requester is:*

$$\gamma(b) = (1 - \mu) \cdot b + p, \tag{1}$$

where $\mu \in [0, 1]$ is the revenue sharing factor and $p \geq 0$ is the fixed wholesale price.

It is easy to see that the hybrid pricing scheme degenerates to the pure revenue sharing scheme when $p = 0$ and to the pure wholesale pricing scheme when $\mu = 1$.

Moreover, with the limited data sharing, a sensor can only share data with a limited number of requesters, and hence some requesters may not be able to acquire the desired data. Let $\alpha \in [0, 1]$ denote the probability that a requester acquires the data successfully, called the *serving probability* of requester. Note that with the unlimited data sharing, α is always equal to 1. With the limited data sharing, however, α can be smaller than 1, as sensors may not be able to serve all requesters.

Definition 2 (Serving Probability)**.** *Suppose that one sensor can serve at most K requesters. Let N_S and N_R denote the number of sensors and requesters, respectively. Then, the serving probability of requester is:*

$$\alpha = \min \left\{ \frac{N_S \cdot K}{N_R}, \ \ 1 \right\}. \tag{2}$$

Obviously, when the number of sensors N_S is very small and the number of requesters N_R is very large, it is likely that some requesters cannot be served.

D. User Behavior and Payoff

Now we model the user behavior and define the user payoff. Without loss of generality, we consider an arbitrary user n and an arbitrary data i in the later analysis. For writing convenience, we omit the subscripts n and i whenever there is no confusion caused. Hence, we can write the location weight ω_i as ω, and write the sensing cost $c_{n,i}$ and user-dependent data value $v_{n,i}$ as c and v, respectively.

Note that users can be fully characterized by c and v, as all other parameters are identical for all users. Hence, we will use (c, v) to characterize the user type. Different users have different c and v, which are independent and identically distributed (iid). For simplicity, we assume that both v and c follow independent and uniform distributions over $[0, 1]$ across all users, and hence their joint probability density function (pdf) is $f_{cv} = 1$ if $v, c \in [0, 1]$, and otherwise $f_{cv} = 0$.

A user can choose different ways to obtain data or choose not to obtain data. We denote the user strategy by

$$\chi \in \{\widetilde{S},\ \widetilde{R},\ \widetilde{A}\}$$

where $\chi = \widetilde{S}$ denotes that the user chooses to be a sensor, and $\chi = \widetilde{R}$ denotes that the user chooses to be a requester, and $\chi = \widetilde{A}$ denotes that the user chooses to be an alien. Let $U(\chi)$ denote the net benefit (i.e., payoff) of user when choosing a particular strategy χ. Note that $U(\chi)$ also depends on the user type (c, v), and we omit (c, v) for presentation convenience. The objective of the user is to choose the proper strategy χ that maximizes his payoff. Next, we define the user payoff $U(\chi)$ under different user strategies χ.

Definition 3 (Sensor Payoff). *A sensor acquires the data via sensing, and hence can achieve a certain direct benefit*

$$b_S = \omega \cdot v - c.$$

Moreover, the sensor can also achieve a certain sharing income by selling the data to requesters. Let ρ denote the ***average sharing income*** *that a sensor can achieve by selling his data to requesters. Then, the user's payoff when choosing to be a sensor is*

$$U(\widetilde{S}) = b_S + \rho, \tag{3}$$

where ρ will be derived in Section IV.

Definition 4 (Requester Payoff). *A requester acquires the data via the sharing of other users and needs to bear the total data transfer cost. Hence, if acquiring the data successfully, he can achieve a certain direct benefit*

$$b_R = \omega \cdot v - s.$$

Besides, the requester needs to pay a hybrid price $\gamma(b_R)$ defined in Definition 1 to the sensor who shares data with him. Thus, the requester's payoff when acquiring the data successfully is $b_R - \gamma(b_R)$. Notice that a requester can only be served with a serving probability α given in Definition 2. Hence, the user's average payoff when choosing to be a requester is

$$U(\widetilde{R}) = \alpha \cdot \big(b_R - \gamma(b_R)\big). \tag{4}$$

Finally, an alien neither achieves benefit, nor incurs cost. Hence, the user's payoff when choosing to be an alien is

$$U(\widetilde{A}) = 0. \tag{5}$$

E. Problem Formulation

Each user can choose a role from the strategy set $\{\widetilde{S}, \widetilde{R}, \widetilde{A}\}$ according to his type (c, v) as well as the other users' choices. Thus, we formulate the system as a non-cooperative game.

- *Players:* the set of all users, i.e., $\mathcal{N} = \{1, 2, \dots, N\}$;
- *Strategies:* each user can choose a strategy χ from the strategy set $\{\widetilde{S}, \widetilde{R}, \widetilde{A}\}$;
- *Utilities:* each user's payoff when choosing a particular strategy χ is $U(\chi)$ defined in (3)-(5).

For analytical convenience, we assume a large network with an infinite number of users (i.e., $N \to \infty$) similar as in [17], [18]. Thus, the impact of a single user's action on the whole market can be ignored. That is, each user's best choice depends on the choices of a mass of users, rather than that of a particular user. This assumption is mainly used for facilitating the theoretic analysis and obtaining the closed-form result. Our analysis, however, can be applied to any finite number of users.

III. Social Optimum

Before analyzing the game equilibrium, we first provide the socially optimal solution as a benchmark, in which all users cooperate with each other to maximize their total payoff (called social utility or social welfare).

For notational convenience, we denote χ_{cv} as the role of a user with type (c, v). Then, the total number of sensors (i.e., those choosing $\chi_{cv} = \widetilde{S}$) is:

$$N_S = \int_c \int_v I_{cv}^S \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c,$$

where $I_{cv}^S = 1$ if $\chi_{cv} = \widetilde{S}$, and 0 otherwise. Similarly, the total number of requesters (i.e., those choosing $\chi_{cv} = \widetilde{R}$) is:

$$N_R = \int_c \int_v I_{cv}^R \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c,$$

where $I_{cv}^R = 1$ if $\chi_{cv} = \widetilde{R}$, and 0 otherwise. Thus, we have the following serving probability for requesters:

$$\alpha = \min\left\{ \frac{\int_c \int_v I_{cv}^S \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c}{\int_c \int_v I_{cv}^R \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c} \cdot K,\quad 1 \right\}. \tag{6}$$

Then, the social utility can be formally defined as follows.

Definition 5 (Social Utility). *Suppose the role of a user with type (c, v) is χ_{cv}. Then, the social utility generated by all users can be defined as follows:*

$$\begin{aligned} W = &\int_c \int_v (\omega \cdot v - c) \cdot I_{cv}^S \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c \\ &+ \alpha \cdot \int_c \int_v (\omega \cdot v - s) \cdot I_{cv}^R \cdot f_{cv}\ \mathrm{d}v\mathrm{d}c. \end{aligned} \tag{7}$$

In (7), the first term is the total achieved benefit of all sensors and the second term is the total achieved benefit of all requesters. Note that the social utility consists of the direct benefits of sensors and requesters only, as the total payment of requesters and the total income of sensors cancel out.

Based on the above, we can formulate the social utility maximization problem as follows.

$$\begin{aligned} \max_{\{\chi_{cv}\}} \quad & W \\ s.t. \quad & \chi_{cv} \in \{\widetilde{S}, \widetilde{R}, \widetilde{A}\}, \quad \forall c, v \in [0,1]. \end{aligned} \tag{8}$$

It is challenging to solve the above problem directly, mainly due to the non-linearity introduced by α. In the following, we will transform this optimization problem into an equivalent linear programming.

We first notice that when $\alpha < 1$ (i.e., when some requesters cannot acquire data), we can always improve the social utility by removing some requesters with lower direct benefits (hence increasing the serving probability of those requesters with higher direct benefits). We summarize this observation in the following lemma.

Lemma 1. *Under the optimal solution of* (8)*, the serving probability must be* $\alpha^* = 1$*, or equivalently,*

$$K \cdot \int_c \int_v I_{cv}^S \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c \geq \int_c \int_v I_{cv}^R \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c$$

That is, all requesters can acquire data successfully.

Under the above condition, we can rewrite the social utility into the following equivalent form:

$$\begin{aligned} \widetilde{W} = & \int_c \int_v (\omega \cdot v - c) \cdot I_{cv}^S \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c \\ & + \int_c \int_v (\omega \cdot v - s) \cdot I_{cv}^R \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c. \end{aligned} \tag{9}$$

Therefore, we can transform the above problem into the following equivalent optimization problem:

$$\begin{aligned} \max_{\{\chi_{cv}\}} \quad & \widetilde{W} \\ s.t. \quad & \chi_{cv} \in \{\widetilde{S}, \widetilde{R}, \widetilde{A}\}, \quad \forall c, v \in [0,1], \\ & K \cdot \int_c \int_v I_{cv}^S \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c \geq \int_c \int_v I_{cv}^R \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c. \end{aligned} \tag{10}$$

It is easy to see that (10) is a linear programming, and hence can be solved by many classic methods. Due to space limit, we will skip the detailed solving process here. For convenience, we denote the optimal solution of (8) or (10) by χ_{cv}^*, and denote the corresponding maximum social utility by W^*.

IV. Game Equilibrium

In this section, we will study the game equilibrium systematically. With a little abuse of notation, we denote χ_{cv} as the strategy choice of a user with type (c, v) and $U_{cv}(\chi_{cv})$ as the payoff of a user with type (c, v) choosing strategy χ_{cv}. Then, the game equilibrium can be formally defined as follows.

Definition 6 (Nash Equilibrium)**.** *A strategy profile* $\{\chi_{cv}^\dagger, \forall c, v \in [0,1]\}$ *is an Nash equilibrium, if and only if*

$$U_{cv}(\chi_{cv}^\dagger) \geq U_{cv}(\chi_{cv}), \quad \forall \chi_{cv} \in \{\widetilde{S}, \widetilde{R}, \widetilde{A}\}$$

for all users with all types.

That is, under an Nash equilibrium, none of the users can improve his payoff by changing his strategy choice.

We first notice that given a strategy profile $\{\chi_{cv}, \forall c, v \in [0,1]\}$, the whole market will be divided into three parts according to user choices: the sensor set $\mathcal{S}^S$, the requester set $\mathcal{S}^R$, and the alien set $\mathcal{S}^A$. That is,

$$\begin{aligned} \mathcal{M}^S &\triangleq \{(c,v) | \chi_{cv} = \widetilde{S}\}, \\ \mathcal{M}^R &\triangleq \{(c,v) | \chi_{cv} = \widetilde{R}\}, \\ \mathcal{M}^A &\triangleq \{(c,v) | \chi_{cv} = \widetilde{A}\}, \end{aligned}$$

For convenience, we refer to them as the mark partitions of sensors, requesters, and aliens, respectively.

In the following, we will first derive the average sharing income ρ of sensors, and then analyze the user best response as well as the game equilibrium.

A. Average Sharing Income: ρ

We now derive the average sharing income ρ of sensors. Notice that when a requester with type (c, v) acquires the data successfully, he can achieve a direct benefit: $b_R = \omega \cdot v - c$. Thus, according to the hybrid pricing mechanism in Definition 1, his payment to the sensor (who shares data with him) is:

$$\gamma(b_R) = (1 - \mu) \cdot (\omega \cdot v - c) + p.$$

For notational convenience, we will write $\gamma(b_R)$ as $\gamma_{cv}(b_R)$ or simply γ_{cv}, as it depends on the user type.

We further note that a requester can acquire data successfully with a probability α, i.e., the serving probability defined in (6). Thus, the total payment of all requesters (denoted by Ω) can be computed in the following way:

$$\Omega = N \iint_{\mathcal{M}^R} \alpha \cdot \gamma_{cv} \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c. \tag{11}$$

The above payment will be shared by all sensors equally. The total number of sensors is:

$$N_S = N \iint_{\mathcal{M}^S} f_{cv} \, \mathrm{d}v\mathrm{d}c. \tag{12}$$

Thus, the average sharing income of each sensor is:

$$\rho = \frac{\Omega}{N_S} = \alpha \cdot \frac{\iint_{\mathcal{M}^R} \gamma_{cv} \cdot f_{cv} \, \mathrm{d}v\mathrm{d}c}{\iint_{\mathcal{M}^S} f_{cv} \, \mathrm{d}v\mathrm{d}c}, \tag{13}$$

where α is the serving probability defined in (6).

B. User Best Response

We now analyze the user best response under a given market partition $\{\mathcal{M}^S, \mathcal{M}^R, \mathcal{M}^A\}$.

Specifically, given a market partition $\{\mathcal{M}^S, \mathcal{M}^R, \mathcal{M}^A\}$, we can derive the average sharing income ρ of each sensor according to (13) and the serving probability α for each requester according to (6). Then, we can compute the user payoffs as sensor and requester explicitly according to (3) and (4). Therefore, the best response of each user can be characterized as follows:

1) A user with type (c, v) will choose to be a sensor, i.e., $\chi_{cv} = \widetilde{S}$, if and only if his payoff as a sensor is higher than that as a requestor or an alien, i.e.,

$$U_{cv}(\widetilde{S}) \geq U_{cv}(\widetilde{R}) \quad \text{and} \quad U_{cv}(\widetilde{S}) \geq U_{cv}(\widetilde{A}) = 0$$

Substituting (3)-(5), we have:

$$v \geq \max\left\{\frac{c - \rho - \alpha \cdot \mu \cdot s - \alpha \cdot p}{\omega \cdot (1 - \alpha \cdot \mu)},\ \frac{c - \rho}{\omega}\right\}.$$

That is, a user with type (c, v) will choose to be a sensor, if and only if the above condition holds.

2) A user with type (c, v) will choose to be a requester, i.e., $\chi_{cv} = \widetilde{R}$, if and only if his payoff as a requestor is higher than that as a sensor or an alien, i.e.,

$$U_{cv}(\widetilde{R}) \geq U_{cv}(\widetilde{S}) \quad \text{and} \quad U_{cv}(\widetilde{R}) \geq U_{cv}(\widetilde{A}) = 0$$

Substituting (3)-(5), we have:

$$\frac{\mu \cdot s + p}{\mu \cdot \omega} \leq v \leq \frac{c - \rho - \alpha \cdot \mu \cdot s - \alpha \cdot p}{\omega \cdot (1 - \alpha \cdot \mu)}.$$

That is, a user with type (c, v) will choose to be a requester, if and only if the above condition holds.

3) A user with type (c, v) will choose to be an alien, i.e., $\chi_{cv} = \widetilde{A}$, if and only if his payoff as an alien (i.e., 0) is higher than that as a sensor or a requester, i.e.,

$$U_{cv}(\widetilde{R}) \leq U_{cv}(\widetilde{A}) = 0 \quad \text{and} \quad U_{cv}(\widetilde{R}) \leq U_{cv}(\widetilde{A}) = 0$$

Substituting (3)-(5), we have:

$$v \leq \min\left\{\frac{c - \rho}{\omega},\ \frac{\mu \cdot s + p}{\mu \cdot \omega}\right\}.$$

That is, a user with type (c, v) will choose to be an alien, if and only if the above condition holds.

After obtaining the best response of every user, we actually get a new market partition, called the newly derived market partition and denoted by $\{\widetilde{\mathcal{M}}^S, \widetilde{\mathcal{M}}^R, \widetilde{\mathcal{M}}^A\}$.

Similarly, under the newly derived market partition $\{\widetilde{\mathcal{M}}^S, \widetilde{\mathcal{M}}^R, \widetilde{\mathcal{M}}^A\}$, we can derive a new average sharing income (denoted by $\widetilde{\rho}$) of each sensor according to (13) and the serving probability (denoted by $\widetilde{\alpha}$) for each requester according to (6). Then, we can compute the user payoffs as sensor and requester again according to (3) and (4), and derive the user best response accordingly. This is the key idea of user behavior dynamics and market evolution.

C. Game Equilibrium Analysis

We now analyze the equilibrium of game, i.e., the equilibrium state of the above user behavior dynamics or market evolution.

According to Definition 6, a strategy profile is a game equilibrium, if and only if none of the users has the incentive to change his strategy. This implies that under a game equilibrium, the market partition $\{\mathcal{M}^S, \mathcal{M}^R, \mathcal{M}^A\}$ will no longer change, and hence the average sharing income ρ of sensor and the serving probability α of requester will no longer change accordingly. Thus, we have the following necessary and sufficient condition for the equilibrium.

Lemma 2. *A strategy profile $\{\chi^{\dagger}_{cv}, \forall c, v \in [0, 1]\}$ is an Nash equilibrium, if and only if*

(a) $\widetilde{\mathcal{M}}^S = \mathcal{M}^S,\ \widetilde{\mathcal{M}}^R = \mathcal{M}^R,\ \widetilde{\mathcal{M}}^A = \mathcal{M}^A$;

(b) $\widetilde{\rho} = \rho,\ \widetilde{\alpha} = \alpha$.

Note that the two conditions in Lemma 2 are equivalent. On one hand, if $\widetilde{\mathcal{M}}^S = \mathcal{M}^S$, $\widetilde{\mathcal{M}}^R = \mathcal{M}^R$, $\widetilde{\mathcal{M}}^A = \mathcal{M}^A$, i.e., the newly derived market partition is same as the original one, then the new average sharing income $\widetilde{\rho}$ must be same as the original one ρ as they are computed by (**??**) in the same way, and the new serving probability $\widetilde{\alpha}$ must be same as the original one α as they are computed by (6) in the same way. On the other hand, if $\widetilde{\rho} = \rho$ and $\widetilde{\alpha} = \alpha$, then the payoffs as requester and sensor for each user will not change, and hence all users will keep the same choice, which implies that the market partition will not change.

D. Dynamic Algorithm

In the previous subsection, we have characterized the conditions of game equilibrium. Now we propose a best response based dynamic algorithm to compute the above game equilibrium.

The key idea of the dynamic algorithm is as follows: every user repeatedly updates his strategy in a myopic manner, i.e., each strategy update is based only on the current market partition. To describe such a dynamic process, we will introduce a virtual time-slotted system, similar as in [17], [18]. Each user hypothetically updates his strategy in every time slot based on the current market partition (formed by the users' strategy choices in the previous time slot). Such a strategy update process repeats until none of the users change strategy. Obviously, the dynamic algorithm stops when it reaches an equilibrium state.

Let $t = 1, 2, ...$ denote the time slots (each with a sufficiently small time length) in the virtual time system. For notational convenience, we denote $\chi_{cv}(t)$ as the strategy of a user with type (c, v) at time slot t, $\{\mathcal{M}^S(t), \mathcal{M}^R(t), \mathcal{M}^A(t)\}$ as the market partition at time slot t, and $\rho(t)$ and $\alpha(t)$ as the average sharing benefit of sensor and serving probability of requester at time slot t. The detailed algorithm is given Algorithm 1.

V. Simulations

In this section, we will provide simulation results regarding the average sharing income, social utility, and the proportions of users with different types, etc. We will show how these outcomes change with different system configurations and parameters. To get clear insights, while not affecting the generality, we assume $\omega = 1$ and $s = 0.1$ in our simulations. We will compare our game equilibrium outcome with the social optimum outcome.

A. Evolution of Market

From the Figure 2, we assume that in an ideal case of $K = 5, \eta = 0.98, p = 0.01$, the P2P-based MCS system can reach the best result of the social optimal model, because the ultimate distribution of the user type of the two models are almost the same; and figure 3 shows the serving probability can reach 1

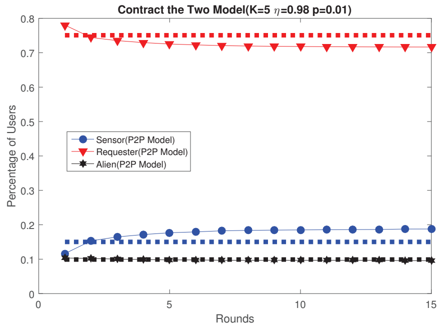

Fig. 2: Distribution of The User Type Between 2 Models

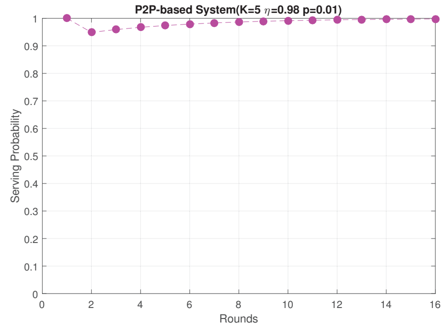

Fig. 3: The Serving Probability

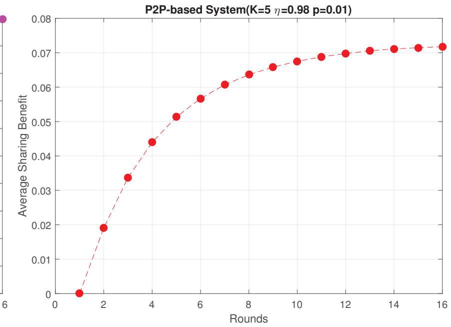

Fig. 4: The Average Sharing Income

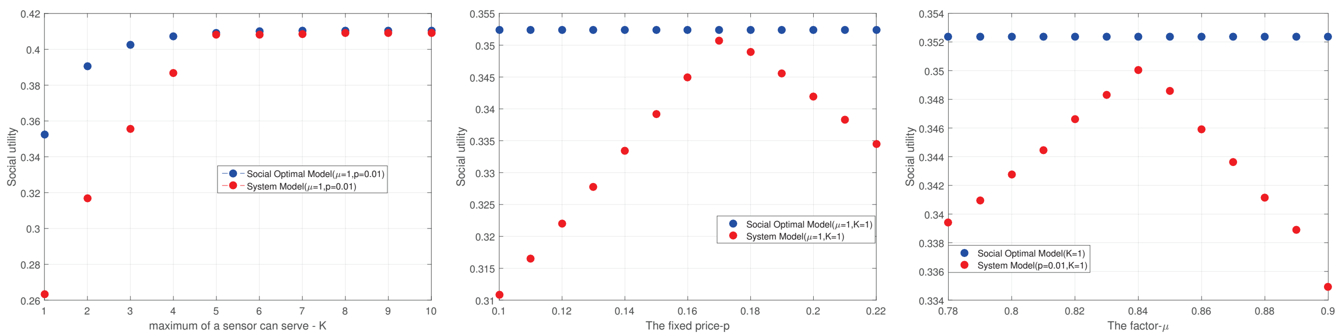

Fig. 5: Social Utility with the Change of K, p, and μ

Algorithm 1: Best Response Based Dynamic Algorithm

Initialization: $\chi_{cv}(0),\ \forall c, v \in [0, 1]$;

while *at each time slot* $t = 1, 2, ...$ **do**

Get the current market partition according to the user strategy update in the previous time slot:

$$\{\mathcal{M}^S(t), \mathcal{M}^R(t), \mathcal{M}^A(t)\} \leftarrow x_{cv}(t-1)$$

Compute the average sharing benefit of sensor and the serving probability of requester:

$$\rho(t), \alpha(t) \leftarrow \{\mathcal{M}^S(t), \mathcal{M}^R(t), \mathcal{M}^A(t)\}$$

Derive the best strategy of each user:

$$x_{cv}(t), \forall c, v \leftarrow \rho(t), \alpha(t)$$

if $x_{cv}(t) = x_{cv}(t-1), \forall c, v$ **then**

Break; /* *reach equilibrium* */

after iteration for about 16 times. And we can see from 4, the average sharing benefit is also increasing as the increases of iteration times and tends to convergent. Under this particular case, the requesters can acquire the data with a low price, and the new model (P2P-based MCS model) can perform pretty well.

B. Impact of K

The factor K can influence the results of the P2P-based MCS system, as is shown in the first picture of Figure 5, the blue points and red points are the best social utility under the social optimal model and P2P-based MCS model, respectively. They both under the case of $\mu = 1, p = 0.01$, we can see that K will more and more seriously effect the difference of social utility of the two models, especially K is small than 5, hence we should adjust the price scheme to narrow the gap.

We further see the next two pictures of Figure 5, we choose $K = 1$ to simulate, and we can see when $\mu = 0.84, p = 0.1$ or $\mu = 1, p = 0.17$ (though p is still small, it has increased 10 times), the gap can be narrowed, this means when a task need more sensors, the requesters have to pay a little more money to get the sensory data.

C. Impact of Pricing Policy

From above analysis, we know that while K is small, the requesters need to pay more for the data to reach the maximum social utility. What we want to study is whether the new pricing scheme which aims at the best social utility will do harm to the benefit of requesters or sensors.

Simulation results as is shown in Figure 6 and Figure 7 have told us the answer. Actually, both the *net* benefit of the requesters and average sharing income of sensors are increased, due to the increase of the serving probability which is shown in Figure 8, more requesters can get the data from the sensors (requester who doesn't obtain the sensory data will not pay any money for it), and this make the social utility increase.

VI. Conclusion

In this work, we introduced a peer-to-peer (P2P) based MCS system, where the sensory data is stored in user devices locally and shared among users in an P2P manner. We focused on the economic incentive issue arising in the sharing of data among users in such a system, that is, how to incentivize users to share

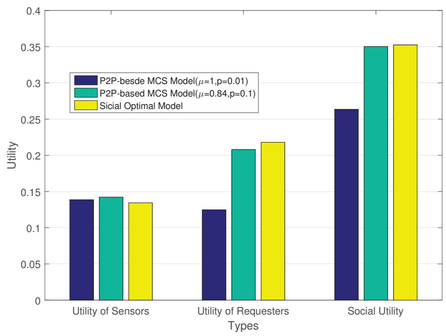

Fig. 6: Users' Net Benefits Under Different Price Schemes

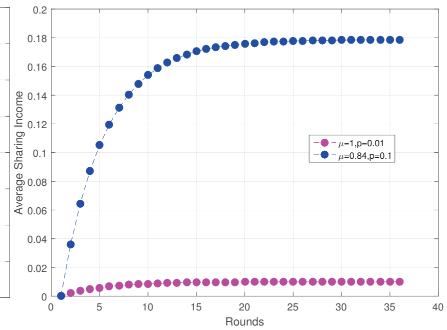

Fig. 7: The Change of ρ Under Different Prices

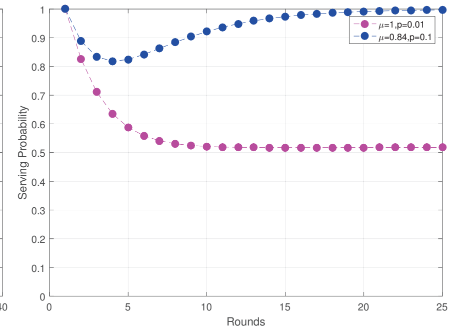

Fig. 8: The Change of α Under Different Prices

their sensed data with others. To achieve this, we proposed a *data market*, together with a hybrid pricing mechanism, for users to *sell* their sensed data to others. We analyzed the user behavior dynamics as well as the data market evolution, by using the evolutionary game theory, and characterized the users' equilibrium behaviors as well as the market equilibrium. There are several interesting and important directions for the extension of this work. First, it is important to study the user behavior dynamics with bounded rationality. Second, it is important to study the MCS system with multiple coupled data.

REFERENCES

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile Crowdsensing: Current State and Future Challenges", *IEEE Communications Magazine*, 2011.
[2] B. Hull, V. Bychkovsky, Y. Zhang, *et al.*, "CarTel: A Distributed Mobile Sensor Computing System," *Proc. ACM SenSys*,2006.
[3] John R. Vacca, "Handbook of Sensor Networking: Advanced Technologies and Applications" *CRC Press*, 2015.
[4] M. R. Ra, B. Liu, T. F. La Porta, *et al.*, "Medusa: A programming framework for crowd-sensing applications," *Proc. ACM Mobisys*, 2012.
[5] NoiseTube, http://www.noisetube.net/.
[6] WeatherLah, http://www.weatherlah.com/.
[7] C. Jiang, L. Gao, L. Duan, and J. Huang, "Data-Centric Mobile Crowdsensing," *IEEE Transactions on Mobile Computing*, 2017.
[8] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand, "Motivating Smartphone Collaboration in Data Acquisition and Distributed Computing," *IEEE Transactions on Mobile Computing*, 2014.
[9] T. Luo and C.-K. Tham, "Fairness and Social Welfare in Incentivizing Participatory Sensing," *Proc. IEEE SECON*, 2012.
[10] X. Zhang, Z. Yang, Z. Zhou, *et al.*, "Free Market of Crowdsourcing: Incentive Mechanism Design for Mobile Sensing," *IEEE Transactions on Parallel and Distributed Systems*, 2014.
[11] C. Jiang, L. Gao, L. Duan, and J. Huang, "Exploiting Data Reuse in Mobile Crowdsensing," *Proc. IEEE GLOBECOM*, 2016.
[12] Y. Singer and M. Mittal, "Pricing Mechanisms for Crowdsourcing Markets," *Proc. ACM WWW*, 2013.
[13] L. Gao, F. Hou, and J. Huang, "Providing Long-term Participation Incentive in Participatory Sensing," *Proc. IEEE INFOCOM*, Apr,2015.
[14] J. Niwa, K. Okada, T. Okuda, and S. Yamaguchi, "MPSDataStore: A Sensor Data Repository System for Mobile Participatory Sensing," *ACM SIGCOMM workshop on Mobile Cloud Computing*, 2013.
[15] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with Smartphones," *IEEE Internet Computing*, 2012.
[16] Y. Kaneko, K. Harumoto, S. Fukumura, S. Shimojo, and S. Nishio, "A Location-based Peer-to-Peer Network for Contextaware Services in a Ubiquitous Environment," *IEEE Symposium on Applications and the Internet Workshops*, 2005.
[17] C. Jiang, L. Gao, L. Duan, and J. Huang, "Economics of Peer-to-Peer Mobile Crowdsensing," *Proc. IEEE GLOBECOM*, 2015.
[18] C. Jiang, L. Gao, L. Duan, and J. Huang, "Scalable Mobile Crowdsensing via Peer-to-Peer Data Sharing," *IEEE Transactions on Mobile Computing*, 2017.
[19] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, 1998.
[20] Y. Luo, L. Gao, and J. Huang, "An Integrated Spectrum and Information Market for Green Cognitive Communications," *IEEE Journal on Selected Areas in Communications*, 2016.
[21] Y. Luo, L. Gao, and J. Huang, "MINE GOLD to Deliver Green Cognitive Communications," *IEEE Journal on Selected Areas in Communications*, 2015.
[22] Q. Ma, L. Gao, Y.-F. Liu, and J. Huang, "A Game-Theoretic Analysis of User Behaviors in Crowdsourced Wireless Community Networks," *Proc. IEEE WiOpt*, 2015.
[23] Q. Ma, L. Gao, Y.-F. Liu, and J. Huang, "Economic Analysis of Crowdsourced Wireless Community Networks," *IEEE Transactions on Mobile Computing*, 2016.
[24] L. He and J. Walrand, "Pricing and Revenue Sharing Strategies for Internet Service Provider," *IEEE Journal on Selected Areas in Communications*, 2006.
[25] C. Gizelis and D. Vergados,"A Survey of Pricing Schemes in Wireless Networks," *IEEE Commu. Surveys and Tutorials*, 2012.