2013 IEEE International Conference on Distributed Computing in Sensor Systems

Quality of Contributed Service and Market Equilibrium for Participatory Sensing

Chen-Khong Tham and Tie Luo
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576
E-mail: {eletck, eleluo}@nus.edu.sg

***Abstract*—User-contributed or crowd-sourced information is becoming increasingly common. In this paper, we consider the specific case of participatory sensing whereby people contribute information captured by sensors, typically those on a smartphone, and share the information with others. We propose a new metric called Quality of Contributed Service (QCS) which characterizes the information quality and timeliness of a specific real-time sensed quantity achieved in a participatory manner. Participatory sensing has the problem that contributions are sporadic and infrequent. To overcome this, we formulate a market-based framework for participatory sensing with plausible models of the market participants comprising data contributors, service consumers and a service provider. We analyze the market equilibrium and obtain closed form expressions for the resulting QCS at market equilibrium. Next, we examine the effects of realistic behaviors of the market participants and the nature of the market equilibrium that emerges through extensive simulations. Our results show that, starting from purely random behavior, the market and its participants can converge to the market equilibrium with good QCS within a short period of time.**

I. Introduction

User-contributed or crowd-sourced information is becoming increasingly common. Together with the rise of social media, they are increasingly being relied on as alternative sources of information that supplement, or in some instances even replace, traditional information channels. One specific aspect of user-contributed or crowd-sourced information is participatory sensing whereby people contribute information captured by sensors, typically those on a smartphone, and share the information with other users or a service provider. The vast penetration of smartphones with a variety of built-in sensors such as GPS, accelerometer and camera amongst the population creates the potential of dense high-quality participatory sensing and makes it an appealing alternative to deployed sensors for large-scale data collection [1], [2].

There are several examples of smartphone applications that harness user-contributed data. Waze [3] is a community-based traffic and navigation application that enable drivers to share real-time traffic and road information in a particular area with other drivers, with the objectives of saving time and fuel costs for people on their daily commutes. Applications like Universal Studios Wait Times and Disneyland Wait Times collect user-contributed waiting time information for the various attractions of Disneyland and Universal Studios, respectively, supplementing the official waiting time information disseminated by the theme park operators.

Tie Luo is currently with the Institute for Infocomm Research (I^2R), Singapore. E-mail: luot@i2r.a-star.edu.sg

A Singapore-based smartphone application WeatherLah [4] receives crowd-sourced data in the form of a 'yes' or 'no' answer from each user about whether it is raining or not at a particular location. Although weather information from the Singapore National Environmental Agency (NEA) and other sources are available, they are usually based on satellite images taken at high altitude and may not reflect the actual fine-grained situation on the ground, which is where WeatherLah can be useful. Another application by the same developer, Mana Rapid Transit [5] invites iPhone users to submit a simple 'yes' or 'no' answer to the question "Is it crowded where you are right now?" to determine the level of crowdedness in the Mass Rapid Transit (MRT) subway stations and trains. This application has proven its worth during the two unfortunate major disruptions in the Singapore MRT system in December 2011 as the information provided led commuters to make alternative travel arrangements and avoid extreme over-crowding within the subway stations.

In [2], we presented *ContriSense:Bus*, a participatory sensing system comprising a client application on Android smartphones (Fig. 1) and a server or cloud back-end which performs spatio-temporal processing. Commuters contribute GPS traces while on public bus journeys which are processed to yield travel time measurements along segments delimited by two neighbouring bus-stops. Commuters can then query for the travel time of a specified bus journey comprising a number of segments. The system also informs the commuter making the query on the confidence level of the result for each segment.

Participatory sensing has the potential to achieve a greater sensing reach and coverage compared to the case of deployed sensors, especially when there are many data contributors. However, under normal circumstances, there are very few user contributions to the WeatherLah, Mana Rapid Transit and most other crowd-sourced or participatory applications. Thus, one serious weakness of participatory sensing is that user contributions are sporadic and infrequent, largely due to users' indifference and the cost to them in terms of mobile data charges, battery life and inconvenience. Even when and where there is data being contributed, the quality of the contributed data in terms of accuracy, resolution, frequency and timeliness may vary greatly as different contributors have different sensors, smartphone models and mobile data plans.

In [6], we explored several ways to incentivize participatory sensing and studied the fairness and social welfare characteristics of several algorithms to apportion the service quota of compelling services to a user based on the user's level of

978-0-7695-5041-1/13 $26.00 © 2013 IEEE
DOI 10.1109/DCOSS.2013.54

Fig. 1. *ContriSense:Bus* participatory sensing application for public transport on Android smartphones. Record and Query can happen simultaneously, with query results derived from earlier contributions.

contribution and demand for services.

In this paper, we tackle the challenge of attracting a regular stream of data contributions of reasonable quality through market-based mechanisms so that useful information can be extracted and passed on to service consumers who would pay for the information. First, in Section II, we motivate the need to consider information quality in participatory sensing. Next, in Section III, we formulate a market-based framework for participatory sensing with plausible models of the market participants comprising data contributors, service consumers and a service provider, and propose a new metric called Quality of Contributed Service (QCS) which characterizes the information quality and timeliness of a specific real-time sensed quantity achieved through participatory means. In Section IV, we analyze what happens at market equilibrium for the contributors and consumers, and design several algorithms to achieve the market equilibrium.

Next, in Section V, we examine the effects of realistic behaviors of the market participants and the nature of the market equilibrium that emerges through extensive simulations. Our results show that, starting from purely random behavior, the market and its participants can quickly converge to the market equilibrium with good QCS. Finally, we conclude in Section VI and discuss possible areas of future work.

To the best of our knowledge, this paper is the first to quantify the Quality of Contributed Service (QCS) for participatory sensing, derive the market equilibrium and determine the level of obtainable QCS at that point.

II. Participatory Sensing and Information Quality

Participatory sensing can be employed to gather sensor information in: (1) continuous-valued form, such as temperature and other environmental parameters, travel durations in ContriSense [2] and wait time durations in queues, e.g. at theme parks such as Disneyland and Universal Studios, or (2) binary form, such as the presence of absence of an event in event detection applications, or the 'yes' or 'no' responses in the WeatherLah and Mana Rapid Transit applications, as described in Section I.

One of the main motivations of this paper is to characterize the information quality (IQ) achievable through participatory sensing. In this paper, we consider the specific case of collecting time-sensitive binary event information. Although this is more restrictive than the continuous-valued case, we start with this in order to rigorously study the expected equilibrium conditions and the IQ that can be achieved. We leave the study of the continuous-valued form of participatory sensing for future work.

Similar to [6], we consider a participatory sensing scenario in which *contributors* contribute raw sensor data and/or processed sensor information over a wireless or cellular connection to a server or *service provider* (SP) which aggregates the information contributed by many contributors and performs additional processing on the information received. *Consumers* query the SP to seek the information they desire, paying a token sum for this information. This ecosystem comprising contributors, consumers and the SP will only be sustainable if each party derives some utility from this arrangement. This issue is the focus of subsequent sections of this paper.

A. Information Quality of Contributions

Although there are a number of information utility measures [7], we focus on those related to the binary decision of whether a Phenomenon of Interest (PoI) is present or absent.

Following event detection theory [8], we are concerned with the detection accuracy of the system whose IQ is reflected in the degree of confidence that an event of interest has occurred. In this section, we develop the relationship between the IQ of an individual contribution by an individual contributor to the target IQ of the system in terms of the target probabilities of detection and false alarm, P_d and P_f, respectively.

We let hypothesis H_1 denote the presence of a PoI; H_0 denotes the corresponding absence of the PoI. The probabilities $P(H_1) = p$ and $P(H_0) = 1 - p$, where $0 < p < 1$, are assumed to be known *a priori* and can be based on historical information. Each contributor *independently* senses and collects data about the environment periodically. When conditioned upon the hypothesis H_i, $i \in \{0, 1\}$, observations are assumed to be independently and identically distributed (i.i.d.) by each contributor as well as across contributors.

The independent signal y_k observed by a contributor k is:

$$y_k = \begin{cases} w_k & \text{if } H_0 \text{ (PoI is absent);} \\ f(r_k) + w_k & \text{if } H_1 \text{ (PoI is present),} \end{cases}$$

where $w_k \sim \mathcal{N}(0, \sigma_w^2)$ is the noise seen by contributor k that follows a normal distribution with zero mean and standard deviation σ_w; r_k is the distance between contributor k and the PoI; and f is a function that monotonically decreases with increasing r_k.

For each sampled signal y_k, contributor k makes a per-sample binary decision $b_k \in \{0, 1\}$ such that:

$$b_k = \begin{cases} 0 & \text{if } y_k < \mathbb{T}_k; \\ 1 & \text{otherwise,} \end{cases}$$

where $\mathbb{T}_k$ is the per-sample threshold of contributor k.

The per-sample probability of false alarm p_0^k by contributor k is independent of its location, and given by [9]:

$$p_0^k = P(b_k = 1|H_0) = Q(\frac{\mathbb{T}_k}{\sigma_w}) \tag{1}$$

where $Q(x)$ is the Gaussian Q-function of a standard normal distribution. The corresponding per-sample probability of detection p_1^k (where $p_1^k > p_0^k$ from the characteristics of the Q-function) at contributor k is dependent on the distance r_k between contributor k and the PoI, and given by:

$$p_1^k = P(b_k = 1|H_1) = Q(\frac{\mathbb{T}_k - f(r_k)}{\sigma_w}) \tag{2}$$

A specific IQ metric used in decision fusion applications [10], [11] is the *log-likelihood ratio* S_i which characterizes the information quality (IQ) in terms of the certainty of the presence or absence of the PoI at a sensor node i, defined as

$$S_i \triangleq \log \frac{P(b_i|H_1)}{P(b_i|H_0)} = \log \Lambda(b_i) \tag{3}$$

where $H_{1,0}$ corresponds to the case that the PoI is actually present or absent, and $b_i = \{1, 0\}$ corresponds to node i's decision on whether the PoI is present or absent, respectively. Eqn. (3) can be evaluated from Eqn.s (1) and (2).

In our case of participatory sensing, the contributor k contributes a decision b_k and provides an IQ measure q_k which reflects his certainty on the presence or absense of the event. We use the quantity S_i given by Eqn. (3) above to be the IQ measure q_k of the contribution from contributor k, which can be evaluated either by the contributor himself or the SP. This quantity will be used in the system model for a contributor that will be developed in Section III-A.

B. Cumulative Information Quality at Service Provider

In decision fusion applications, the role of the fusion center (FC) is to detect the presence of the PoI by making a global binary decision $\hat{H} = \{H_0, H_1\}$ based on the decisions that it has received from a set of n sensor nodes. Let $B = \{b_1, b_2, ..., b_n\}$ be the set of per-sample binary decisions that the FC receives from each sensor node in a time epoch. The optimal decision fusion rule for the FC using aggregated data from all the sensor nodes is the Likelihood Ratio Test (LRT) [8] [12]:

$$\Lambda(B) = \frac{P(b_1, b_2, ..., b_n|H_1)}{P(b_1, b_2, ..., b_n|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1-p}{p} \tag{4}$$

The FC makes the decision that the PoI is present ($\hat{H} = H_1$) if $\Lambda(B) \geq \frac{1-p}{p}$, and decides that the PoI is absent otherwise.

Since observations across sensor nodes are i.i.d., the *cumulative log-likelihood ratio* S_{FC} at the FC is:

$$S_{\mathrm{FC}} = \log \Lambda(B) = \log \prod_{i=1}^{n} \Lambda(b_i) = \sum_{i=1}^{n} S_i \tag{5}$$

where S_i is defined in Eqn. (3) above. The summation property of the log-likelihood ratio is particularly useful and will be exploited later.

The level of S_{FC} achieved reflects the degree of confidence in the global binary decision and can be regarded as the cumulative information quality at the FC. Following Wald [13], the hypothesis H_1 that the PoI is present, i.e. an event of interest has occurred, is highly confident when the cumulative log-likelihood ratio satisfies

$$S_{\mathrm{FC}} \geq \mathbb{B}$$

where $\mathbb{B} = \log(\frac{P_d}{P_f})$, and P_d and P_f are the target detection and false alarm probabilities, respectively.

The i.i.d. requirement is satisfied in the participatory sensing case since each contributor makes his own observation and decision. We assume that observations and decisions by the same contributor at two different points in time are also independent. Rewriting Eqn. (5) for the participatory sensing case where the FC is the Service Provider (SP), we arrive at

$$Q_{\mathrm{SP}} = \sum_{i=1}^{\tilde{n}} q_i \tag{6}$$

for the aggregated cumulative IQ at the SP. Note that index i is used in place of k since contributions from many contributors are aggregated at the SP, and $\tilde{n}$ is the number of such contributions in a valid time interval that will be defined in the next section.

A similar test

$$Q_{\mathrm{SP}} \geq \mathbb{B} \tag{7}$$

can be performed to determine whether there is high confidence in the global binary decision at the SP.

The summation structure in Eqn. (6) will be augmented with time-decaying weights in Section III-B to form the Quality of Contributed Service (QCS) metric proposed in this paper. QCS can be viewed as the *cumulative time-decaying* or *timeliness-weighted log-likelihood ratio* of the global decision at the SP on the presence of the PoI. This is a natural extension since the confidence level in each contribution decreases over time due to the fact that the status of the PoI is more likely to change as a longer time elapses.

III. System Model

Most participatory sensing applications are *time sensitive* in nature, due to their objective of sourcing for up-to-date information. This means that the value or usefulness of user-contributed data decays with time and may even become worthless after a certain period of time. The Quality of Contributed Service (QCS) framework that will be developed in this section takes this into account.[1]

In the following sub-sections, we will develop the system model for a contributor in the participatory sensing system before presenting the definition of QCS, followed by the model for a consumer. Note that a user can be both a contributor and a consumer although we treat them as separate roles here.

A. Contributor

An arbitrary contributor $k \in \{1, ..., N_z\}$ makes contributions at a rate of λ_k per unit time, each with information quality (IQ) q_k as defined in Section II-A, where $\lambda_k \geq 0$.

The contributor incurs some cost arising from sensing and contributing, either in terms of telecommunication charges or battery consumption. Let us denote the IQ-dependent cost by

[1] Our framework subsumes time-*insensitive* cases too, as shown in the Appendix [14].

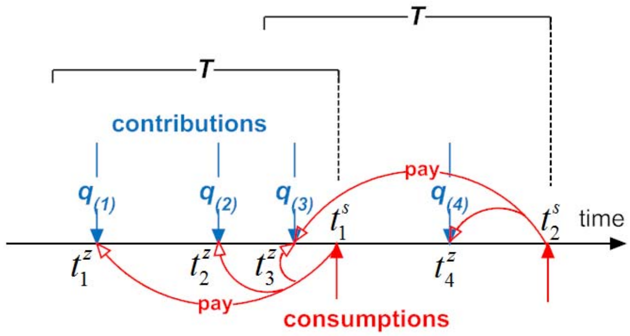


Fig. 2. Interactions between contributions and consumptions according to temporal sequence. T indicates the length of each consumable window.

c_k. Each contribution will thus incur a cost of $c_k q_k$ to the contributor.[2]

In return, a contribution i will receive remuneration r_i from the service provider (SP), where r_i depends on both *demand* for information by consumers and *supply* of contributions by other contributors. The SP operates a platform that does not just connect one consumer to one contributor, but connects an indefinite number of consumers to an indefinite number of contributors. Fig. 2 shows contribution events (originating from several contributors) enter the platform at time instances $t^z_{1,2,3,\dots}$ and consumption events (originating from several consumers) enter the platform at time instances $t^s_{1,2,3,\dots}$. We consider the case where timely data are valuable whereas outdated data are worthless. As such, we bring in the notion of *lifetime* of a contribution, denoted by T as seen in the figure.[3] Accordingly, we define two sliding time windows for consumers and contributors, respectively: (i) $W^-_{t^s} \triangleq [t^s - T, t^s]$ is the *consumable window* of the consumption that enters the platform at t^s - only contributions with $t^z \in W^-_{t^s}$ are relevant to this consumption; (ii) $W^+_{t^z} \triangleq [t^z, t^z + T]$ is the *valid window* of the contribution that enters at t^z - this contribution is only valid to consumptions whose $t^s \in W^+_{t^z}$.

With these concepts, we are now ready to introduce the demand and supply based remuneration scheme. A consumer will have to pay a price of p for each instance of consumption j. This amount, less a *commission rate* of η deducted by the SP, will be shared by all the contributions made in j's consumable window $W^-_{t^s_j}$. Conversely, an instance of contribution i will receive payment from all instances of consumption happening during i's valid window $W^+_{t^z_i}$. The remuneration r_i is calculated as

$$r_i \triangleq (1-\eta)p \sum_{t^s_j \in W^+_{t^z_i}} \frac{q_{(i)}}{\sum_{t^z_l \in W^-_{t^s_j}} q_{(l)}} \tag{8}$$

where the subscript of q with parentheses indicates that it pertains to the IQ of an instance of contribution in order to differentiate it from q_k which refers to the IQ associated with contributions from a *contributor* k.

This remuneration scheme (8) has two important features:

- It is *risk free* for the SP, in the sense that the SP does not act as an reseller who *buys* from one market and then *sells* to another market, which presents the risk of loss to the SP when the revenue of selling does not cover the cost of buying. The remuneration scheme (8) carries no risk of loss as it uses a *balance equation* among consumers and contributors.
- It implies that remuneration is *postponed* - a contributor can only receive remuneration at interval T after making a contribution. This is analogous to the real life situation where an employee only receives his salary a certain period (e.g. a month) later. In many participatory sensing applications such as traffic monitoring, the interval T is fairly short such as one hour, which should be acceptable to contributors. In fact, such a postponed scheme has the advantage that contributors will be *forward-looking* and tend to maintain their contribution levels and only make adjustments after some time when they review the payoff. This not only provides some stability in the system and IQ assurance to the SP and consumers, but also motivates us to use the concept of "review period" (RP) in the design of a mechanism to achieve market equilibrium that will be presented in Section IV-A.

Denote by R_k the total remuneration received by contributor k per unit time, and denote by π^z_k his *payoff* per unit time. Under the common assumption that users are rational, we assume that a contributor k's objective is to maximize his own payoff, i.e.

$$\text{maximize } \pi^z_k = R_k - c_k \lambda_k q_k \tag{9}$$

where the decision variables are λ_k and q_k, and we will analyze R_k later. This optimization will be conducted for each time slot, which is the RP just mentioned. Therefore, λ_k and q_k may vary from RP to RP.

B. Quality of Contributed Service (QCS)

In this sub-section, we develop a new metric called Quality of Contributed Service (QCS) which characterizes the information quality and timeliness of a real-time sensed quantity achieved in a participatory manner. The QCS metric extends the information quality (IQ) measure of each contribution and the cumulative IQ at the service provider (SP) presented in Section II.

QCS can be defined with respect to an individual consumption, which reflects a particular one-time *consumer experience* of using the service, or with respect to the whole system, reflecting the expected consumer experience. These two perspectives can be mde concrete: (1) a single consumption that happens at t^s will experience an *instantaneous* QCS of

$$Q(t^s) \triangleq \sum_{t^z_i \in W^-_{t^s}} q_{(i)} w_i \equiv \sum_{i=1}^{\tilde{n}} q_{(i)} w_i \tag{10}$$

and, (2) the system QCS is given by

$$Q \triangleq \mathbb{E}_{t^s}[Q(t^s)]$$

[2]Quality-dependent cost is also adopted by many other works such as [15]. In the case of constant contribution cost regardless of quality, it can be easily shown that, in any Nash Equilibrium, each contributor will contribute at the maximum quality so as to maximize his payoff. We will not consider this case in this paper.

[3]In addition, a *time-decaying effect* is associated with each contribution and will be formulated in Section III-B.

In Eqn. (10), $W_{t^s}^-$ and $q_{(i)}$ are the consumable window and information quality of a contribution, respectively, $\tilde{n}$ is the number of contributions (treated as a random variable) in $W_{t^s}^-$, w_i is the normalized *time-decaying factor*, defined as

$$w_i \triangleq \frac{e^{-\Delta t_i^z} - e^{-T}}{1 - e^{-T}}, \tag{11}$$

where $\Delta t_i^z \triangleq t^s - t_i^z, t_i^z \in W_{t^s}^-$.

The definition of w_i in Eqn. (11) is similar to the discount factor in dynamic programming and the Bellman equation, and the exponential weighted moving average (EWMA) [16]. This can be seen by ignoring the normalizing term and noticing that $e^{-\Delta t_i^z} = (e^{-a})^{\frac{\Delta t_i^z}{a}}$ for such a that $0 < e^{-a} < 1$ and that $\frac{\Delta t_i^z}{a}$ equals the number of epochs between the two points in time.

The QCS defined in Eqn. (10) is a *cumulative time-decaying* or *timeliness-weighted quality of contribution*: the more contributions, or the higher the quality of the contributions, or the more up-to-date the contributions are in the consumable window, the higher will be the QCS value. We have exploited the summation property of the cumulative log-likelihood ratio shown in Eqns (5) and (6).

Note that the time-decaying factor w_i does not affect the remuneration r_i as shown in in Eqn. (8). This prevents the remuneration from diminishing too rapidly in order to ensure that contributors are motivated to contribute.

C. Consumer

An arbitrary consumer $k \in \{1, ..., N_s\}$[4] consumes the service (e.g. query for a phenomenon of interest or PoI) at a rate of μ_k per unit time, for which he pays a price of p for each consumption. We assume $\mu_k \geq 0$ and $p > 0$. Similar to the above, μ_k may vary from RP to RP while being unchanged within each RP. The price p is fixed in each RP.[5] In addition, the service time of each consumption is assumed to be negligible.

A consumer is associated with a *QCS valuation factor*, β_k, which represents how "generous" or "stringent" a consumer valuates the QCS, denoted by Q. In other words, $\beta_k Q$ is the "satisfaction level" or "psychological price" a consumer rates the service to be at, e.g. a low β_k indicates a "hard-to-satisfy" consumer. Thus, a consumer gains a *utility* of $\beta_k Q - p$.

However, this view only treats each consumption in *isolation*, whereas consumptions tend to occur successively in practice since the PoI constantly changes. In this situation, a consumer's utility would not evolve in an *additive* manner[6] as $\mu_k \beta_k Q - \mu_k p$, but rather, non-linearly as $\psi_k(\mu_k)\beta_k Q - \mu_k p$, where $\psi_k(\mu_k)$ is a non-linear function. This function $\psi_k(\cdot)$ satisfies:

1) $\psi_k(0) = 0$;
2) monotonically increasing and concave in μ_k;
3) $\psi_k(\mu_k) \sim \mu_k$ when $\mu_k \to 0^+$, where $\sim$ is a Bachmann-Landau notation [20] meaning "asymptotically equal";
4) $\psi_k(\mu_k) = o(\mu_k)$ when $\mu_k \to \infty$, where $o(\cdot)$ is also a Bachmann-Landau notation meaning "asymptotically dominated by".

Property 2 captures the effect of decreasing marginal utility as consumption increases, which is common in economics. Property 3 emulates the scenario that when the consumption rate is extremely low, consecutive consumptions can be treated as isolated. Property 4 is similar to Property 2.

Two examples satisfying the above properties are:

$$\psi_k(x) = \frac{1}{a}\log(1 + ax), a > 0 \tag{12}$$

and

$$\psi_k(x) = 1 - \frac{1}{a}e^{-ax}, a > 0. \tag{13}$$

Under the same assumption of rationality as in the case of contributors, the objective of a consumer is to maximize his utility received per unit time, or formally

$$\text{maximize } \pi_k^s = \psi_k(\mu_k)\beta_k Q - \mu_k p \tag{14}$$

where the decision variable is $\mu_k \geq 0$.

IV. Market Equilibrium

User-contributed sensing and services are still at an early stage of development. In this new paradigm, since users making contributions are not obligated to do so, but are altruism- or incentive-driven (for example, this paper considers monetary incentive), there are two pertinent questions of interest:

- Does a market equilibrium exist? In other words, will the system stabilize at a certain QCS level?
- If the answer is "yes", what specifically are the achievable QCS and the contribution and consumption levels at the market equilibrium?

We studied these issues and performed a market equilibrium (ME) analysis using the models for contributor and consumer, and the QCS definition, presented in Section III above. Due to the space constraint of this paper, only the key theoretical results of this analysis are presented below. The full derivation and results of the analysis can be found in the Appendix to this paper [14].

Let Σ_o represent other contributors' aggregate contribution level, $U \triangleq \sum_{k=1}^{N_s} \mu_k$ be the aggregate consumption rate and $R_{all} \triangleq (1-\eta)pU$ be the total remuneration that all contributors receive per unit time.

Theorem 1. *The optimal contribution level for maximizing a contributor's payoff is given by*

$$z^* = \sqrt{\frac{\Sigma_o R_{all}}{c}} - \Sigma_o$$

provided that $c < R_{all}/\Sigma_o$, *or otherwise* $z^* = 0$.

Corollary 1. *The optimal consumption rate for maximizing a consumer's utility of* (13) *is given by*

$$\mu^* = \log\frac{\beta Q}{p}$$

respectively, provided that $p < \beta Q$, *or otherwise* $\mu^* = 0$.

[4] We use k as a generic index. Consumer k should not be deemed to be the same as contributor k.

[5] We do not consider dynamic pricing in this paper and leave that for future work. In practice, dynamic pricing encounters several difficulties [17], [18]. Flat pricing, in addition to being simpler, is appealing to consumers.

[6] This simpler case was considered in [19] in a different setting.

Theorem 5. *Under the consumer model of* (13)*, the QCS at market equilibrium is*

$$Q_{me} = -C_2 \cdot \Omega(-\frac{p}{\beta C_2})$$

where $C_2 = \kappa p(1-\eta)(1-\frac{1}{N_z})N_s/c$ *and* $\Omega(\cdot)$ *is the Lambert W-function.*

The Lambert W-function, also called the omega function, is the inverse function of $f(W) = We^W$. For instance, $\Omega(e) = 1$, $\Omega(-1/e) = -1$, and $\Omega(1) = 0.56714$ (the "omega constant").

A. Algorithms to Achieve Market Equilibrium

The theoretical results above guide us towards the design of a mechanism for the system to achieve the ME, which we present here as three Algorithms 1, 2 and 3.

At the end of each RP, the SP will announce R_{all} and N_z in the elapsed RP, for each contributor to decide on his contribution level z in the next RP. On the other hand, consumers do not need to rely on the SP to disseminate information on Q because they can experience the (changing) QCS instantaneously and thus, can adjust their consumption rates μ promptly. As a result, they do not even need to follow the RPs.

Throughout this paper, we do not treat N_z merely as the number of (registered) contributors who may or may not contribute, but the *effective* number of contributors who are actually contributing. This reflects the real situation where participatory sensing usually has a large population to serve as *potential* contributors but the pool of "active contributors" is usually much smaller. Finally, note that the pool of active contributors does not always have to contain the same set of users to achieve the ME, as newly joined contributors can also be guided by the mechanism described above.

In practical settings, it may be too onerous for users to manually follow the steps in the algorithms presented here. An application running on, e.g. an on-board car computer, can be configured to contribute and/or consume at exponentially distributed intervals determined by the mechanism described in this section.

Algorithm 1 Algorithm for Contributor

for $m = 1 \to \infty$ **do**
  if $m = 1$ **then**
    Randomly choose a contribution level z
  else
    Receive $R_{all}(m-1)$ and $N_z(m-1)$ from the SP
    Determine z^* according to $z^* \leftarrow \frac{N_z(m-1)-1}{N_z(m-1)^2}\frac{R_{all}(m-1)}{c}$
  end if
  Choose λ and q such that $\lambda q = z^*$
  Contribute at the chosen level (i.e. at exponentially distributed intervals of mean $1/\lambda$ and quality q) till the end of the RP
end for

Algorithm 2 Algorithm for Consumer

Randomly choose the initial consumption time t_s
loop
  Consume service at t_s and pay price p to the SP
  Experience QCS and obtain a satisfaction level of $\beta Q(t_s)$
  Determine μ according to $\mu \leftarrow \log \frac{\beta Q}{p}$
  Consume at the chosen rate (i.e. at exponentially distributed intervals of mean $1/\mu$)
end loop

Algorithm 3 Algorithm for SP

Set a countdown timer $tm \leftarrow ||\text{RP}||$ (duration of RP) associated with callback function `endOfRP`
loop
  Wait for an incoming event
  if event=contribution **then**
    Evaluate and record the contribution with timestamp
  else if event=consumption **then**
    Serve the consumer, i.e. provide aggregated information with QCS Q, and receive payment p
    Remunerate contributors in the consumable window according to Eqn. (8)
  end if
end loop

CALLBACK `endOfRP`:

if tm fires **then**
  Calculate and announce R_{all} and N_z
  Reset $tm \leftarrow ||\text{RP}||$
end if

V. Performance Evaluation

In this section, we first conduct discrete-event driven simulations to verify our theoretical analysis of market equilibrium (ME) by examining key parameters such as QCS and contribution and consumption levels, as well as to evaluate the speed of convergence and the parameters of the market-based mechanism for achieving ME.

A. Market-Based Mechanism

Four cases are considered in evaluating Algorithms 1, 2 and 3:

1) *Homogeneous Users with Optimal Adjustments*: All users are homogeneous. Each contributor is able to adjust his rate and quality of contribution to achieve the ME based on information provided by the SP, following Algorithm 1. Each consumer is also able to adjust his rate of consumption based on his perceived QCS, following Algorithm 2. This case can be treated as the system operating under ideal conditions.
2) *Homogeneous Users with Sub-optimal Adjustments*: This is similar to Case 1 above, with the difference that contributors and consumers are unwilling or unable to adjust their behaviors precisely to the optimal settings,

due to various real-life factors such as indifference or lack of knowledge.

3) *Heterogeneous Users with Optimal Adjustments*: Due to different human usage patterns, smartphone models and mobile data plans in use, the unit cost c of making a contribution is different for different contributors. Each contributor is still able to adjust his rate and quality of contributions based on information provided by the SP. For consumers, we take into account the different psychological factors of different people by considering different user-specific β. Each consumer is still able to adjust his rate of consumption based on his perceived QCS.

4) *Heterogeneous Users with Sub-Optimal Adjustments*: This is a combination of Cases 2 and 3 above.

In the simulations, QCS is computed as the actual experienced QCS which is determined using Eqn. (10) with actual arrival patterns of contributors and consumers according to two random Poisson point processes. This gives the actual perceived QCS result under realistic operating conditions. Similarly, the R_{all} value in Algorithm 1 is computed as the actual remuneration that the SP paid to all the contributors averaged over time.

The simulation setup is as follows. Contributors and consumers enter the system as two Poisson point processes with mean $\Lambda = \sum_{k=1}^{N_z} \lambda_k$ and $U = \sum_{k=1}^{N_s} \mu_k$, respectively. The time unit is hour. T=1, $||\text{RP}||$=24, p=1, η=0.3, c=1, β=2, N_z=N_s=100. As explained in Section IV-A, the population size can be arbitrarily large, but the number of active contributors is usually much smaller, and assumed here to be fairly stable. With these settings, the theoretical result of Theorem 5 above gives the theoretical Q_{me} to be 168.612.[7]

As aforementioned, since $z = \lambda q$, a contributor can either adjust λ or q or both to achieve a certain z. In the simulations, we let $\lambda \hookleftarrow \mathcal{U}(0, 2)$ (where "$\hookleftarrow$" means "draws from" and $U(a, b)$ means uniformly distributed between a and b), and each contributor adjusts q as per $q = z/\lambda$ where z is specified in Algorithm 1. In the initial RP, $q \hookleftarrow \mathcal{U}(0, 1)$. In the event that Q drops to as low as $\beta Q \leq p$ for a consumer, he will choose $\mu \hookleftarrow \mathcal{U}(0, 2)$. In fact, this did not happen in the simulations, meaning that $\beta Q > p$ was always satisfied.

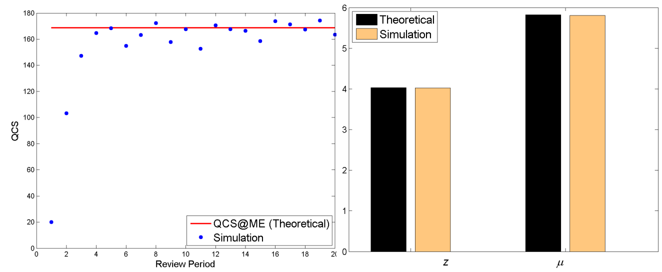
In Cases 2 and 4, contributors and consumers deviate from the optimal settings following a normal distribution with standard deviation 50% of the optimal settings, i.e. $z_k \hookleftarrow \mathcal{N}(z^*, 0.5z^*)$ and $\mu_k \hookleftarrow \mathcal{N}(\mu^*, 0.5\mu^*)$. The z_k's and μ_k's are independently generated.

In Cases 3 and 4, the heterogeneity is characterized by a random deviation of maximal $\pm 50\%$ from the homogeneous case, i.e. each $c_k \hookleftarrow \mathcal{U}(0.5, 1.5)$ and each $\beta_k \hookleftarrow \mathcal{U}(1, 3)$.

1) Results: The results of Case 1 are shown in Fig. 3. We see from the convergence trajectory that if users make the optimal adjustments, the system converges to the ME in only 4 or 5 RPs and the converged QCS matches well with the theoretical value of Q_{me}.

This shows that the SP is able to achieve a good cumulative IQ that exceeds the IQ threshold, i.e. it is able to make

[7]The Lambert W-function is a special multi-valued function and the other solution of 0.508861 should not be taken.

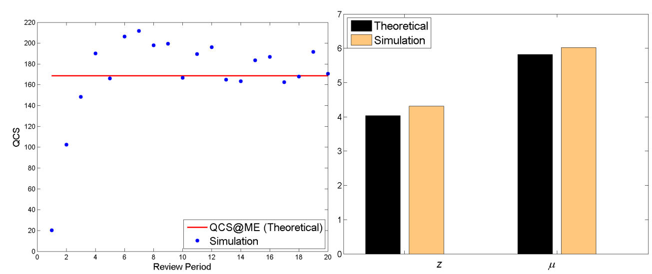


(a) Convergence trajectory. (b) $\bar{z}$ and $\bar{\mu}$ at ME.

Fig. 3. Case 1: Homogeneous users with optimal adjustments.

the global decision with high confidence, as discussed in Section II-B and expressed by Eqn.s (6) and (7), taking into account the fact that QCS is a timeliness-weighted sum of the IQ value of each contribution.

Fig. 3b compares the optimal contribution level z^* and consumption rate μ^*, which are the analytical ME values, with $\bar{z}$ and $\bar{\mu}$, which are the simulation results. $\bar{z}$ and $\bar{\mu}$ are calculated as the average of z and μ in RPs of $m = 6$ till 20 (since ME is observed to be achieved after 4 or 5 RPs, and the results for $m > 20$ are similar to $m \leq 20$) for all users over 10 simulation runs. In Fig. 3b, we see that $\bar{z}$ and $\bar{\mu}$ are almost identical to the theoretical values, which validates our analysis.



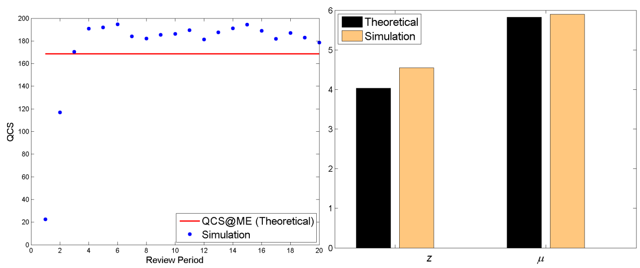
(a) Convergence trajectory. (b) $\bar{z}$ and $\bar{\mu}$ at ME.

Fig. 4. Case 2: Homogeneous users with sub-optimal adjustments.

In Case 2 where users make sub-optimal adjustments, Fig. 4a shows that slight fluctuations in QCS occur, but are nevertheless still centred around the theoretical Q_{me}. In Fig. 4b, we see that $\bar{z}$ and $\bar{\mu}$ in this case differ slightly from their optimal values: although z_k and μ_k are generated from the normal distribution with means z^* and μ^*, respectively, the simulation average is not equal to the optimal settings. This is because of a non-linear effect: the impact of z_k when it is above the optimal setting is larger than its impact when it is below the optimal setting; similarly for μ.

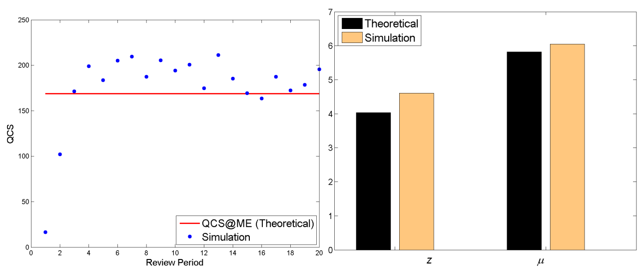
For Case 3, the result in Fig. 5a shows that, interestingly, the user heterogeneity raises the Q_{me} of the homogeneous case by $5 - 18\%$. This is attributed to the increased contribution level $\bar{z}$ as seen in Fig. 5b.

Finally, let us look at Fig. 6 for Case 4 which is the most comprehensive and realistic experiment. We can see that the results demonstrate a combination of the results from Cases 2



(a) Convergence trajectory. (b) $\bar{z}$ and $\bar{\mu}$ at ME.

Fig. 5. Case 3: Heterogeneous users with optimal adjustments.



(a) Convergence trajectory. (b) $\bar{z}$ and $\bar{\mu}$ at ME.

Fig. 6. Case 4: Heterogeneous users with sub-optimal adjustments.

and 3: the convergence trajectory is fluctuating like in Case 2, and the QCS is higher than Q_{me} like in Case 3. This observation applies similarly to Fig. 6b as well. The key message is that, even under fairly high heterogeneity (maximal $\pm 50\%$ deviation) and sub-optimal adjustments (standard deviation of 50% about the optimum), the market can still converge to an equilibrium close to the theoretical ME and the user activity level is also well predicted by the analysis.

B. Summary of Findings

- The market-based mechanism *can* lead the users to the market equilibrium (ME) that is in agreement with the ME analyzed theoretically.
- This mechanism is *effective*, with a fast convergence speed and accommodates well user heterogeneity and sub-optimal behavior adjustments.
- This mechanism is *robust* to the randomness of user behavior: the market consistently converges to the same equilibrium regardless of how users start off in the initial period when the mechanism is not yet in effect.

VI. Conclusion

Participatory sensing has so far been regarded as a "best effort" or "opportunistic" form of sensing that is inferior to deployed sensors. This paper has quantified the quality of service of participatory sensing systems whose service relies solely on user contributions by proposing the concept of Quality of Contributed Service (QCS). We take a market-based approach whereby each contributor contributing data is motivated by obtaining a share of consumer payment from the service provider, according to his contribution rate and quality; on the other hand, consumers choose the service consumption rate based on how well the QCS meets his satisfaction level. Both contributors and consumers are not altruistic but are rational, behaving (contributing or consuming) in the manner that maximizes their respective payoffs or utilities. We derive the optimal contribution level and consumption rate and prove the existence of the market equilibrium, at which QCS is shown to be significant and non-trivial. Our findings indicate that participatory sensing can be used for reliable sensing purposes when certain incentives, e.g. monetary as in this paper, and a market framework are properly set up. In future work, we plan to study the effects of dynamic pricing on the achievable information quality in participatory sensing, as well as study the achievable information quality for user contributions in the form of continuous-valued data rather than that of binary events considered in this paper.

References

[1] UCLA Urban Sensing Research Area, http://research.cens.ucla.edu/urbansensing/.

[2] J. K.-S. Lau, C.-K. Tham, and T. Luo, "Participatory cyber physical system in public transport application," in *Proceedings of CCSA, IEEE/ACM UCC*, 2011.

[3] "Waze - outsmarting traffic, together," 2012. [Online]. Available: http://www.waze.com/

[4] Buuuk, "WeatherLah iPhone application," 2012. [Online]. Available: http://itunes.apple.com/us/app/weatherlah/id411646329?mt=8

[5] ——, "Mana Rapid Transit iPhone application," 2012. [Online]. Available: http://itunes.apple.com/sg/app/mana-rapid-transit/id501961346?mt=8

[6] T. Luo and C.-K. Tham, "Fairness and social welfare in incentivizing participatory sensing," in *IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (IEEE SECON 2012)*, Jun 2012.

[7] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *Int'l Jnl of High Perf. Comp. Appl.*, vol. 16, no. 3, 2002, (Initially appeared as Xerox PARC Tech. Rep. P2001-10113).

[8] R. Viswanathan and P. K. Varshney, "Distributed Detection With Multiple Sensors: Part I - Fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, 1997.

[9] T. Q. S. Quek, D. Dardari, and M. Z. Win, "Energy Efficiency of Dense Wireless Sensor Networks: To Cooperate or Not to Cooperate," *IEEE JSAC*, vol. 25, no. 2, 2007.

[10] H.-X. Tan, M.-C. Chan, W. Xiao, P.-Y. Kong, and C.-K. Tham, "Information quality aware routing in event-driven sensor networks," in *IEEE INFOCOM*, 2010.

[11] W.-L. Yeow, C.-K. Tham, and W.-C. Wong, "Optimizing application performance through learning and cooperation in a wireless sensor network," in *IEEE MILCOM*, 2007.

[12] S. C. A. Thomopoulos, R. Viswanathan, and D. K. Bougoulias, "Optimal Distributed Decision Fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 25, no. 5, 1989.

[13] A. Wald, "Sequential Tests of Statistical Hypotheses," *Annals of Mathematical Statistics*, vol. 16, no. 2, 1945.

[14] C. Tham and T. Luo, "Appendix to 'Quality of Contributed Service and Market Equilibrium for Participatory Sensing'." [Online]. Available: http://cnl-ece.nus.edu.sg/cktham/ThamLuoDCOSS13Appendix.pdf

[15] S. Ren, J. Park, and M. van der Schaar, "Profit maximization on user-generated content platforms," in *49^{th} Allerton Conference on Communication, Control, and Computing*, 2011.

[16] C. C. Holt, "Forecasting trends and seasonal by exponentially weighted averages," *Office of Naval Research Memorandum*, vol. 52, 1957.

[17] J. Park and M. van der Schaar, "Incentive provision using intervention," in *IEEE INFOCOM*, 2011.

[18] P. Nabipay, A. Odlyzko, and Z.-L. Zhang, "Flat versus metered rates, bundling, and bandwidth hogs," in *NetEcon*, 2011.

[19] S. Ren, J. Park, and M. van der Schaar, "User subscription dynamics and revenue maximization in communications markets," in *IEEE INFOCOM*, 2011.

[20] N. G. de Bruijn, *Asymptotic Methods in Analysis*. Amsterdam: North-Holland, 1958.