Quality of Contributed Service and Market Equilibrium for Participatory Sensing

Chen-Khong Tham and Tie Luo

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 a closed form expression 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.

Index Terms—Mobile computing, participatory sensing, network economics, incentive

1 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 (SP). 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.

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 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

- *C.-K. Tham is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576. E-mail: eletck@nus.edu.sg.*
- *T. Luo is with the Institute for Infocomm Research (I^2R), Singapore. E-mail: luot@i2r.a-star.edu.sg.*

Manuscript received 22 June 2013; revised 25 May 2014; accepted 3 June 2014. Date of publication 11 June 2014; date of current version 2 Mar. 2015. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TMC.2014.2330302

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 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, we survey recent literature on participatory sensing in Section 2 before we motivate the need to consider information quality (IQ) in participatory sensing in Section 3. Next, in Section 4, 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 5, we analyze what happens at market equilibrium (ME) for the contributors and consumers, and obtain a closed form expression for the resulting QCS at market equilibrium. We then design several algorithms to achieve the market equilibrium.

Next, in Section 6, 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. We also performed sensitivity analysis on several key parameters to investigate the stability of the market-based mechanism. Finally, we discuss some issues in Section 7 before concluding the paper in Section 8.

2 Related Work

The field of participatory sensing is multi-faceted and has received a lot of research attention in the past few years.

In order to improve the trustworthiness of participatory sensing, Dua et al. [7] used the trusted platform module (TPM) with modified phone hardware to ensure the integrity of sensor data and protect against malicious users who may tamper with sensor measurements. In this paper, we assume that users are not malicious and do not tamper with sensor measurements.

The issues of selection and recruitment of data contributors in order to obtain data from reputable contributors and achieve good coverage over a wide geographical area were presented in Reddy et al. [8]. In this paper, we consider a specific local region in which all contributors are welcome and an IQ metric is computed for each contributor.

Wang et al. [9] studied the issue of truth discovery from potentially noisy data contributed by multiple observers. Similar to our work in this paper, they considered the case of binary measurements and adopted a maximum likelihood with expectation maximization (EM) approach to derive the most probable ground truth, whereas we have used the simpler log-likelihood ratio metric. However, Wang et al. did not consider the larger system aspect of market equilibrium between data contributors and service consumers and the achievable aggregated information quality, which we have studied in detail in this paper.

Next, we shift our focus to network economics based schemes for participatory sensing which is closer to the topic of this paper. In order to motivate smartphone users to contribute to sensing tasks, Yang et al. [10] devised two auction-based incentive schemes: a platform-centric scheme and a user-centric scheme. Lee and Hoh [11] proposed a reverse auction mechanism that allows users to sell their sensing data to a service provider by bidding their desired selling prices. Recently, Luo et al. [12] proposed a mechanism that incentivizes participatory sensing based on all-pay auctions with a contribution-dependent prize.

There are important differences between these three approaches and the work presented in this paper. First, all of them only considered the interaction between a service provider or platform with data contributors and have not included service consumers who can pay for consumed services in their frameworks. This means that the service provider or platform has to finance the payouts to data contributors through some external means, e.g., advertising or other sources of revenue, whereas we have proposed a self-sustaining scheme with utility-maximizing data contributors and service consumers reaching a market equilibrium. Second, none of the aforementioned approaches have attempted to quantify the aggregated information quality achievable through participatory sensing.

The contributions of this paper are that it quantifies the aggregated information quality for participatory sensing in the form of the quality of contributed service metric, derives the market equilibrium achieved by data contributors and service consumers, and determines the level of QCS obtainable at that point.

3 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 1.

One of the main motivations of this paper is to characterize the information quality achievable through participatory sensing. In this paper, we consider the specific case of collecting time-sensitive binary event information. Although

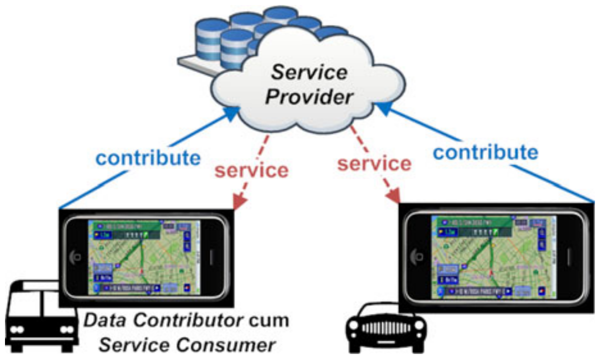


Fig. 1. System model for a participatory sensing application comprising contributors, a service provider and consumers.

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 *service provider* which aggregates the information contributed by many contributors and performs additional processing on the information received (see Fig. 1). *Consumers* invoke services which query the SP to seek the information that they desire which is derived from contributions and pay 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 shall be the focus of subsequent sections of this paper.

3.1 Information Quality of Contributions

Although there are a number of information utility measures [13], we focus on those related to the binary decision of whether a Phenomenon of Interest (PoI) is present or absent.

Following event detection theory [14], 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 [15]:

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

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:

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

A specific IQ metric used in decision fusion applications [16], [17] is the *log-likelihood ratio* S_i which characterizes the IQ in terms of the certainty of the presence or absence of the PoI at a sensor node i, defined as

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

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. Eq. (3) can be evaluated from Eqs. (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 Eq. (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 4.1.

3.2 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 $B = \{b_1, b_2, \ldots, b_n\}$ 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) [14], [18]:

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

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:

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

where S_i is defined in Eq. (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 [19], 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 Eq. (5) for the participatory sensing case where the FC is the service provider, we arrive at

$$Q_{\mathrm{SP}} = \sum_{i=1}^{\tilde{n}} q_i, \quad (6)$$

which is the aggregated cumulative IQ at the SP. Note that index i is used in place of k since contributions from many contributors, and each contributor may contribute more than once, 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

$$Q_{\mathrm{SP}} \geq \mathbb{B} \quad (7)$$

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

The summation structure in Eq. (6) will be augmented with time-decaying weights in Section 4.2 to form the quality of contributed service metric that is 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.

4 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 framework that will be developed in this section takes this into account.[1]

In the following 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.

4.1 Contributor

An arbitrary contributor $k \in \{1, \ldots, N_z\}$ makes contributions at a rate of λ_k per unit time, each with information quality q_k as defined in Section 3.1, where $\lambda_k \geq 0$.

The contributor incurs some cost arising from sensing and contributing, e.g., telecommunication charges and battery consumption. This cost is dependent on the desired IQ as it is usually more costly in terms of sensing and computation time and energy (e.g., more samples, more complex signal processing), and even communication (e.g., verification with other information sources), to produce a sensor reading that has high information quality (e.g., accurate and complete, with low uncertainty). In this paper, we approximate the IQ-dependent cost incurred by contributor k for each contribution with the expression $c_k q_k$, where c_k is some constant.[2]

In return, a contribution i will receive remuneration r_i from the service provider, 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 $t^z_{1,2,3,\ldots}$ and consumption events (originating from several consumers) enter the platform at time instances $t^s_{1,2,3,\ldots}$. 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 $t^s \in W^+_{t^z}$.

1. Our framework subsumes time-*insensitive* cases too, as will be shown later in Section 5.3.

2. Quality-dependent cost is adopted by several other works such as [12] and [20]. 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 4.2.

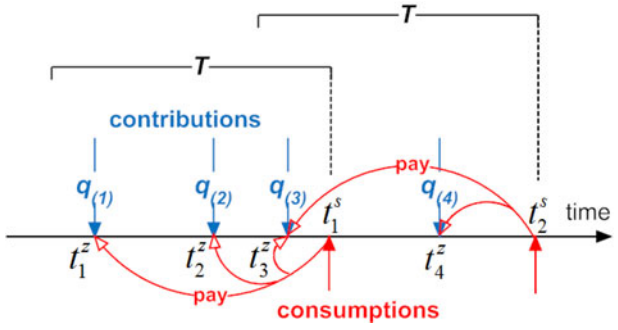


Fig. 2. Interactions between contributions and consumptions according to temporal sequence. T indicates the length of each consumable window.

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 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 a risk of loss to the SP when the revenue from selling the services does not cover the cost of buying them. This remuneration scheme carries no risk of loss as it uses a *balance equation* among consumers and contributors.
- It implies that remuneration is *postponed*: when a contributor makes a contribution, he will only receive remuneration for it after an interval of up to T time units. 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. 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 stability to the system and some assurance of IQ to the SP and consumers, but also motivates us to use the concept of "review period" (RP) in the analysis of the market equilibrium that will be presented in Section 5.4, and the mechanisms to achieve it that will be presented in Section 5.5.

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.

4.2 Quality of Contributed Service

In this section, we develop a new metric called quality of contributed service 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 measure of each contribution and the cumulative IQ at the service provider presented in Section 3.

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 made 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

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

In Eq. (10), $W^-_{t^s}$ and $q_{(i)}$ are the consumable window and IQ 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

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

where $\Delta t^z_i \triangleq t^s - t^z_i, t^z_i \in W^-_{t^s}$.

The definition of w_i in Eq. (11) is similar to the discount factor in dynamic programming and the Bellman equation, and the exponential weighted moving average (EWMA) [21]. This can be seen by ignoring the normalizing term and noticing that $e^{-\Delta t^z_i} = (e^{-a})^{\frac{\Delta t^z_i}{a}}$ for such a that $0 < e^{-a} < 1$ and that $\frac{\Delta t^z_i}{a}$ equals the number of epochs between the two points in time.

The QCS defined in Eq. (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 Eqs. (5) and (6).

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

4.3 Consumer

An arbitrary consumer $k \in \{1, \ldots, N_s\}$[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 and the consumer is likely to monitor it continuously at some rate μ_k. Furthermore, 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. The 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 [25] 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 a common concept in economics. Property 3 captures the effect of 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 abovementioned properties are:

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

and

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

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

5. Telecommunication charges incurred by the consumer is subsumed in the price p. We do not consider dynamic pricing in this paper and leave that for future work. In practice, dynamic pricing encounters several difficulties [22], [23]. Flat pricing, in addition to being simpler, may be more appealing to consumers.

6. This simpler case was considered in [24] in a different setting.

Under the same assumption of rationality as in the case of contributors, the objective of a consumer is to maximize his payoff or 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 $\mu_k \geq 0$.

5 Market Equilibrium Analysis

User-contributed sensing and services are relatively recent developments. 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 the 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 are the specific achievable QCS and the contribution and consumption levels at the market equilibrium?

In this section, we study these issues and perform a market equilibrium analysis using the models for contributor and consumer, and the QCS definition, presented in Section 4 above, before deriving specific algorithms to achieve the ME.

5.1 Contributor

The first parameter of interest is contributor k's remuneration received per unit time, R_k, as introduced in Eq. (9). To derive this, recall the scenario in Fig. 2 and imagine that it happens continuously, constituting a large sample space. Hence, on average, the system will receive a total payment of $p\sum_{k=1}^{N_s} \mu_k T$ in each period of T, and distribute $(1-\eta)$ portion of it to the contributors following the remuneration allocation scheme stipulated in Eq. (8). Thus, a contributor k will receive in period T the average remuneration of

$$(1-\eta)p\left(\sum_{k=1}^{N_s} \mu_k T\right) \frac{\lambda_k T \cdot q_k}{\lambda_{-k} T \odot q_{-k} + \lambda_k T \cdot q_k},$$

where λ_{-k} and q_{-k} are the contribution rates and information qualities associated with all contributors other than k, i.e., $\lambda_{-k} \triangleq \{\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_{N_z}\}$ and $q_{-k} \triangleq \{q_1, \ldots, q_{k-1}, q_{k+1}, \ldots, q_{N_z}\}$, and $\odot$ denotes the inner product.

For notational convenience, let $U \triangleq \sum_{k=1}^{N_s} \mu_k$ be the aggregate consumption rate, $R_{all} \triangleq (1-\eta)pU$ be the total remuneration that all contributors receive per unit time, and $\Sigma_{-k} \triangleq \lambda_{-k} \odot q_{-k}$. Thus, we obtain R_k as

$$R_k = \frac{\lambda_k q_k}{\Sigma_{-k} + \lambda_k q_k} R_{all.}$$

Letting $z_k \triangleq \lambda_k q_k$, which we shall call the *contribution level* of contributor k, we transform contributor k's objective, from Eq. (9), into

$$\text{maximize } \pi_k^z = \frac{z_k}{\Sigma_{-k} + z_k} R_{all} - c_k z_k,$$

where $z_k \in [0, \infty)$ is the decision variable. A contributor can always achieve a certain z_k by adjusting rate λ_k and/or quality q_k according to his own preference (e.g., sometimes q_k may be hard to control). This gives the user some flexibility.

For a generic formulation without referring to a particular user, we rewrite the above by removing the subscript k and denoting Σ_{-k} by Σ_o which represents other contributors' aggregate contribution level, and arrive at

$$\text{maximize } \pi^z = \frac{z}{\Sigma_o + z} R_{all} - cz. \tag{15}$$

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

$$z^* = \sqrt{\frac{\Sigma_o R_{all}}{c}} - \Sigma_o \tag{16}$$

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

Proof. The first and second order derivatives of π^z with respect to z are:

$$\frac{\partial \pi^z}{\partial z} = \frac{\Sigma_o}{(\Sigma_o + z)^2} R_{all} - c,$$
$$\frac{\partial^2 \pi^z}{\partial z^2} = -\frac{2\Sigma_o}{(\Sigma_o + z)^3} R_{all}.$$

Since $\partial^2 \pi^z / \partial z^2 < 0$,[7] π^z is strictly concave in z. Hence, the optimal z that maximizes π^z is determined by letting the marginal utility $\partial \pi^z / \partial z$ equal to 0, which leads to Eq. (16).

The condition $c < R_{all}/\Sigma_o$ ensures that $z^* > 0$. Otherwise and intuitively, an overly high cost will make the payoff negative, and the optimal action for a contributor to take is to not contribute at all. □

In deriving Eq. (16), R_{all} is treated as exogenously given instead of being a function of z. The validity of this is ensured by our slotted approach using RPs together with a mechanism to achieve the ME, which will be introduced later in Sections 5.4 and 5.5.

For a particular contributor, as Σ_o is the *other* contributors' aggregate contribution level, Theorem 1 actually tells him his *best response* or *rational action set*, in game-theoretic terms. Hence, we can derive the following game-theoretic result.

Theorem 2. *In a homogeneous setting where all contributors bear the same unit cost c, there exists a unique Nash equilibrium in which every contributor adopts the same optimal contribution level z^*, given by*

$$z^* = \frac{N_z - 1}{N_z^2} \frac{R_{all}}{c}, \tag{17}$$

and each contributor's payoff is maximized as

$$\pi^* = \frac{R_{all}}{N_z^2}. \tag{18}$$

Proof. We model the competition between the contributors as a Cournot game [26]. This competition arises as a result of remuneration sharing as in Eq. (8). Each player k tries to maximize his payoff as in (15), by choosing the best strategy z. It can then be shown that when each player bears the same cost c, the optimal strategy z^* will also be the same. As this is an intuitive result, we omit the rigorous proof for brevity.

Thus, $\Sigma_o = (N_z - 1)z^*$, and (17) follows from (16). The maximum payoff π^* is then obtained from (15). □

7. We ignore the case of $\Sigma_o = 0$ or $R_{all} = 0$. These are trivial cases, but are tricky to deal with mathematically. In fact, as will become clear later, these cases do not occur at market equilibrium.

5.2 Consumer

Similarly, for a generic formulation without referring to a particular user, we remove the subscript k from the consumer's objective function and arrive at

$$\text{maximize } \pi^s = \psi(\mu)\beta Q - \mu p. \tag{19}$$

Theorem 3. *The optimal consumption rate for maximizing a consumer's utility is determined by*

$$\psi'(\mu^*) = \frac{p}{\beta Q}, \tag{20}$$

and the necessary and sufficient condition for it to have a unique positive solution and for $\pi_s^* > 0$ is

$$\psi'(0^+) > \frac{p}{\beta Q}. \tag{21}$$

Otherwise, $\mu^* = 0$.

Proof. It follows from the concavity of $\psi(\mu)$ that π^s is also strictly concave in μ. Furthermore, since $\pi^s|_{\mu=0} = 0$ and $\pi^s|_{\mu\to\infty} < 0$ (due to Property 4), $\pi_s^* > 0$ if and only if $\partial \pi^s/\partial \mu|_{\mu\to 0^+} > 0$, which leads to (21) since $\partial \pi^s / \partial \mu = \psi'(\mu)\beta Q - p$. As π^s is strictly concave, it is also the necessary and sufficient condition for the optimal μ that maximizes π^s to be uniquely determined by the first-order condition $\partial \pi^s / \partial \mu = 0$, which is equivalent to (20). □

Similar to R_{all}, Q has been treated as exogenously given rather than being a function of μ. The validity if this is again ensured by considering a slotted approach involving RPs, which will be explained later.

Taking for $\psi(\cdot)$ the two exemplifying non-linear functions, Eqs. (12) and (13), we expand Eq. (19) into

$$\pi^s = \beta Q \log(1 + \mu) - \mu p, \tag{22}$$

$$\pi^s = \beta Q (1 - e^{-\mu}) - \mu p, \tag{23}$$

respectively, where we let $a = 1$ without affecting the principle in the results.

Corollary 1. *The optimal consumption rate for maximizing a consumer's utility (22) and (23) are given by*

$$\mu^* = \frac{\beta Q}{p} - 1 \tag{24}$$

and

$$\mu^* = \log \frac{\beta Q}{p}, \tag{25}$$

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

In a homogeneous setting where all consumers share the same β and function $\psi(\cdot)$, they will all adopt the same μ^*

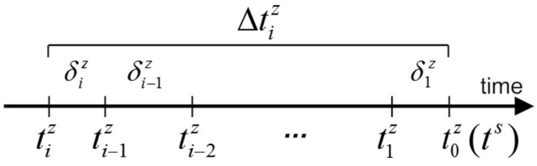


Fig. 3. Decomposing Δt_i^z into δ_j's: time sequence of $t_j^z, j = 0, 1, \ldots, i$.

and hence the aggregate consumption rate $U^* = N_s\mu^*$. Recalling the definition of R_{all}, we have

$$R_{all}^* = (1-\eta)pN_s\mu^*, \tag{26}$$

where μ^* is determined as shown above.

5.3 QCS and Its Approximation

Recall that $Q \triangleq \mathbb{E}_{t^s}[Q(t^s)]$ and, as per Eqs. (10) and (11),

$$Q(t^s) = \sum_{i=1}^{\tilde{n}} q_{(i)}w_i = \frac{\sum_{i=1}^{\tilde{n}} q_{(i)}e^{-\Delta t_i^z} - e^{-T}\sum_{i=1}^{\tilde{n}} q_{(i)}}{1-e^{-T}}.$$

Statistically, $q_{(i)}$, Δt_i^z and $\tilde{n}$ are independent of each other, and hence, from the earlier definition of $Q \triangleq \mathbb{E}_{t^s}[Q(t^s)]$

$$Q = \frac{\mathbb{E}[q_{(i)}]}{1-e^{-T}}\left(\mathbb{E}\left[\sum_{i=1}^{\tilde{n}} e^{-\Delta t_i^z}\right] - \mathbb{E}[\tilde{n}]e^{-T}\right). \tag{27}$$

Let us define $\overline{q} \triangleq \mathbb{E}[q_{(i)}]$ and assume that the *aggregate* stream of contributions is a Poisson point process with rate $\Lambda = \sum_{k=1}^{N_z}\lambda_k$ and is independent of the consumption time t^s. It then follows that $\mathbb{E}[\tilde{n}] = \Lambda T$. To tackle the term $\mathbb{E}[\sum_{i=1}^{\tilde{n}} e^{-\Delta t_i^z}]$, decompose $\Delta t_i^z = \sum_{j=1}^{i}\delta_j$ where $\delta_j = t_{j-1}^z - t_j^z, j = 1, \ldots, i$ and $t_0^z \triangleq t^s$ (see Fig. 3). Since δ_j are exponentially i.i.d. in $[0, T]$,

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{\tilde{n}} e^{-\Delta t_i^z}\right] &= \mathbb{E}\left[\sum_{i=1}^{\tilde{n}}\prod_{j=1}^{i} e^{-\delta_j}\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{\tilde{n}}\prod_{j=1}^{i}\mathbb{E}[e^{-\delta_j} \mid \delta_j < T]\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{\tilde{n}}\prod_{j=1}^{i}\int_0^T e^{-\delta}\frac{\Lambda e^{-\Lambda\delta}}{1-e^{-\Lambda T}}d\delta\right] \\ &= \mathbb{E}\left\{\sum_{i=1}^{\tilde{n}}\left[\frac{\Lambda[1-e^{-(\Lambda+1)T}]}{(\Lambda+1)(1-e^{-\Lambda T})}\right]^i\right\}.\end{aligned}$$

Let us define $f(x) \triangleq \frac{1-e^{-xT}}{x}$ and $g(x) \triangleq \frac{f(x+1)}{f(x)}$. Then, the above transforms to

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{\tilde{n}} g^i(\Lambda)\right] &= \mathbb{E}\left[\frac{g(\Lambda)}{1-g(\Lambda)}(1-g^{\tilde{n}}(\Lambda))\right] \\ &= \frac{g(\Lambda)}{1-g(\Lambda)}\left(1-\sum_{k=0}^{\infty} g^k(\Lambda)\frac{(\Lambda T)^k}{k!}e^{-\Lambda T}\right) \\ &= \frac{g(\Lambda)}{1-g(\Lambda)}\left(1-e^{-[1-g(\Lambda)]\Lambda T}\right).\end{aligned}$$

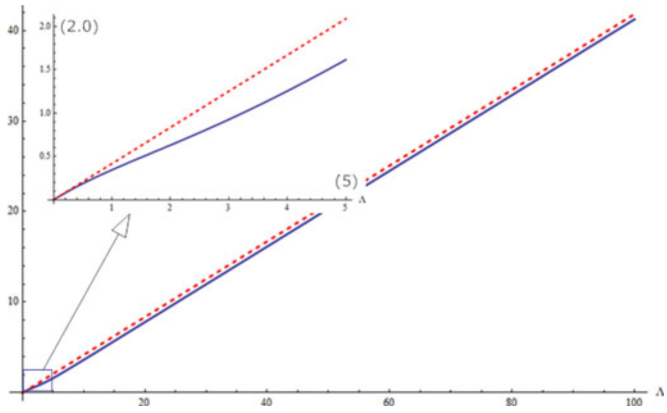


Fig. 4. Verifying the QCS approximation. The blue solid curve represents the target function $h(\Lambda)$. The red dashed line represents the approximating function $\kappa\Lambda$. T=1 hour. Therefore, $\kappa|_{T=1} \approx 0.418$.

Substituting this into Eq. (27) obtains

$$\begin{aligned}Q &= \frac{\overline{q}}{1-e^{-T}}\left[\frac{g(\Lambda)}{1-g(\Lambda)}(1-e^{-[1-g(\Lambda)]\Lambda T}) - \Lambda T e^{-T}\right] \\ &\triangleq \overline{q}\cdot h(\Lambda),\end{aligned} \tag{28}$$

where the definition of $h(\Lambda)$ can be readily seen.

Remark. This analysis covers *time-insensitive* services as well, which is readily obtained by setting the data lifetime T to be infinite:

$$\lim_{T\to\infty} Q = \overline{q}\frac{g(\Lambda)}{1-g(\Lambda)} \equiv \overline{q}\frac{f(\Lambda+1)}{f(\Lambda)-f(\Lambda+1)} = \overline{q}\Lambda. \tag{29}$$

Approximation. Eq. (28) is a complex expression and we adopt an approximation for the sake of tractability. According to [27], for a *given* $\tilde{n}$ (e.g., taking its mean value ΛT), the event times Δt_i^z are distributed like the *order statistics* of $\tilde{n}$ independent random variables that are *uniformly* distributed in $[0, T]$. Therefore,

$$\mathbb{E}\left[\sum_{i=1}^{\tilde{n}} e^{-\Delta t_i^z}\right] = \Lambda T\int_0^T \frac{e^{-x}}{T}dx = \Lambda(1-e^{-T}).$$

Substituting this into Eqn. (27) obtains

$$Q = \overline{q}\Lambda\frac{1-(T+1)e^{-T}}{1-e^{-T}}. \tag{30}$$

Defining $\kappa \triangleq \frac{1-(T+1)e^{-T}}{1-e^{-T}}$, which is a constant, we arrive at $Q = \kappa\overline{q}\Lambda$.

We compare $h(\Lambda)$ in Eq. (28) against its approximated version $\kappa\Lambda$ in Fig. 4 for $\Lambda \in [0, 100]$ where the portion of $\Lambda \in [0, 5]$ is magnified for a clearer view. We see that, when Λ is sufficiently large, e.g., $\Lambda > 20$, the target function can be approximated with error less than 6 percent. As Λ is an *aggregated* rate and considering each unit time as an hour, this is a reasonably good approximation.

5.4 Market Equilibrium

As aforementioned, we have treated R_{all} and Q as exogenously given when deriving the optimal z for contributors and optimal μ for consumers, respectively. The validity of this is ensured by our approach of slotting the system into review periods, which we now describe.

The system unfolds over time as consecutive RPs. The length of an RP is application dependent and can be, for example, a small multiple of T. Let us index the RPs by $m = 1, 2, \ldots$ and denote the variables we have used with RP indices explicitly. For particular examples, $z^*(m)$, which is the optimal z to use in RP m, is determined at the beginning of RP m, when the value of $R_{all}(m-1)$ becomes available. Therefore, $R_{all}(m-1)$ can be treated as exogenously given when determining $z^*(m)$. Hence, a more precise understanding of Eqs. (16) and (17) is obtained by including these RP indices. In addition, because $R_{all}(m-1)$ is determined by $\mu(m-1)$, so $Q(m)$, which is determined by $z(m)$, is also determined by $\mu(m-1)$. Since $\mu(m-1)$ does not affect $\mu^*(m)$, so $Q(m)$ can be treated as exogenously given when determining $\mu^*(m)$. Similarly, Eq. (20) is precisely understood by including the RP indices.

Intuitively, the above, particularly the case of $z^*(m)$ with respect to $R_{all}(m-1)$, says that the contributors assume that the remuneration in the coming slot will remain the same as in the previous slot. This is referred to as the naive or *static expectation* [28], which is also adopted in other works such as [20], [24], [29], [30].

Now, we present the theoretical derivation and results for the market equilibrium.

Theorem 4. *Under the consumer model of (22), the QCS at market equilibrium is*

$$Q_{me} = \frac{pC_1}{\beta(C_1 - 1)}, \qquad (31)$$

where $C_1 = \kappa\beta(1-\eta)(1-\frac{1}{N_z})N_s/c$. *The QCS converges to* Q_{me} *on the condition that* $C_1 > 1$ *and*

$$\frac{pC_1}{\beta(C_1 - 1)} = Q(1). \qquad (32)$$

Proof. Let us rewrite Eqs. (17), (26) and (30) with RP indices explicitly, as

$$R_{all}(m) = (1-\eta)pN_s\left[\frac{\beta Q(m)}{p} - 1\right], \qquad (33)$$

$$z^*(m+1) = \frac{N_z - 1}{N_z^2}\frac{R_{all}(m)}{c}, \qquad (34)$$

$$Q(m+1) = \kappa N_z z^*(m+1), \qquad (35)$$

respectively, where Eq. (33) is based on Eq. (24) and we take $\overline{q}\Lambda = N_z z^*$. Therefore,

$$\begin{aligned} Q(m+1) &= \frac{\kappa(N_z-1)}{cN_z}(1-\eta)pN_s\left[\frac{\beta Q(m)}{p} - 1\right] \\ &= C_1 Q(m) - \frac{pC_1}{\beta}, \end{aligned} \qquad (36)$$

which is a recursive equation. If there exists an ME, i.e., the series $Q_{m=1,2,\ldots}$ converges to a certain quantity which we denote by Q_{me}, then we can obtain Q_{me} by letting $m \to \infty$, as

$$Q_{me} = C_1 Q_{me} - \frac{pC_1}{\beta},$$

whose solution is Eq. (31).

Now, we prove the existence of ME, i.e. convergence condition Eq. (32). Let us recursively expand Eq. (36) as

$$Q(m+1) = C_1^m Q(1) - \frac{p}{\beta}\sum_{i=1}^{m} C_1^i.$$

First, $C_1 > 1$ must hold, for otherwise $Q(m)$ will be negative after a certain m. To see this, suppose $C_1 < 1$ and we will have $Q_\infty = -\frac{p}{\beta}\frac{C_1}{1-C_1} < 0$; similarly, suppose $C_1 = 1$ and we will have $Q_\infty = Q(1) - \frac{p}{\beta}\cdot\infty < 0$.

Furthermore,

$$\begin{aligned} Q(m+1) &= C_1^m Q(1) - \frac{pC_1}{\beta}\frac{1-C_1^m}{1-C_1} \\ &= \frac{pC_1}{\beta(C_1-1)} + C_1^m\left[Q(1) - \frac{pC_1}{\beta(C_1-1)}\right]. \end{aligned}$$

Since $C_1 > 1$, $Q(1) - \frac{pC_1}{\beta(C_1-1)} = 0$ must hold so that $Q(m)$ converges to $\frac{pC_1}{\beta(C_1-1)}$ when $m \to \infty$. Eqn. (32) is thus proven. □

However, note that the convergence condition Eq. (32) is not stable since a small perturbation of $Q(1)$ can prevent the QCS from converging. We now derive the following theorem to get a stable ME.

Theorem 5. *Under the consumer model of (23), the QCS at market equilibrium is*

$$Q_{me} = -C_2 \cdot \Omega\left(-\frac{p}{\beta C_2}\right), \qquad (37)$$

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, discovered by Lambert [31] and Euler [32], and 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").

Proof. Similarly, rewrite Eq. (26) as

$$R_{all}(m) = (1-\eta)pN_s \log\frac{\beta Q(m)}{p}$$

based on Eq. (25). The above, together with Eqs. (34) and (35), lead us to

$$Q(m+1) = \kappa\frac{N_z - 1}{cN_z}(1-\eta)pN_s \log\frac{\beta Q(m)}{p}.$$

Assuming that the ME exists, we let $m \to \infty$ and obtain

$$Q_{me} = C_2 \log\frac{\beta Q_{me}}{p}, \qquad (38)$$

whose solution is a closed form expression given by Eq. (37). □

5.5 Algorithms to Achieve Market Equilibrium

The theoretical results above guide us towards the design of a mechanism for the system comprising contributors, consumers and a service provider to achieve the market equilibrium, which we present here as three Algorithms 1, 2 and 3, respectively.

Algorithm 1 Algorithm for Contributor

for $m = 1 \rightarrow \infty$ **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 Eq. (17):
    $z \leftarrow \frac{N_z(m-1)-1}{N_z(m-1)^2}\frac{R_{all}(m-1)}{c}$
  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 Eq. (25):
  $\mu \leftarrow \log \frac{\beta Q}{p}$
  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 Eq. (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

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 QCS instantaneously and thus, adjust their consumption rates μ promptly. As a result, they do not even need to follow the RPs.

Throughout this paper, we do not regard N_z as the number of (registered) contributors who may or may not contribute, but as the *effective* number of contributors who are actually contributing. This reflects the real situation where participatory sensing usually has a large population of *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 an on-board car computer or smartphone can be configured to contribute and/or consume at exponentially distributed intervals, as determined by the algorithms presented in this section.

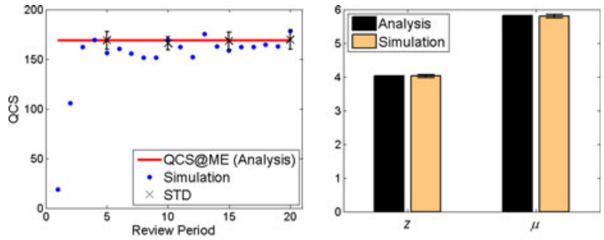
6 Performance Evaluation

In this section, we conduct discrete-event driven simulations to verify our theoretical analysis of the market equilibrium 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 the ME.

6.1 Market-Based Mechanism

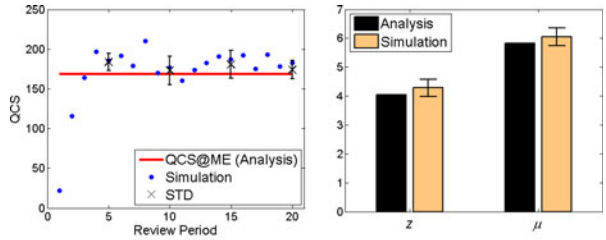
Four cases are considered in evaluating Algorithms 1, 2 and 3:

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.



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

Fig. 5. Case 1: Homogeneous users with optimal adjustments.



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

Fig. 6. Case 2: Homogeneous users with sub-optimal adjustments.

4) *Heterogeneous users with sub-optimal adjustments*. This is a combination of Cases 2 and 3 above.

In the simulations, QCS is not computed using the analytical expressions, i.e., Eq. (28) or (30), but computed as the actual experienced QCS which is determined using Eq. (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 not computed using the analytical expression of Eq. (26), but 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. $T = 1, ||RP|| = 4, p = 1, \eta = 0.3, c = 1, \beta = 2, N_z = N_s = 100$. As explained in Section 5, 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.[8]

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 percent 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 ± 50 percent from the homogeneous case, i.e., each $c_k \hookleftarrow \mathcal{U}(0.5, 1.5)$ and each $\beta_k \hookleftarrow \mathcal{U}(1,3)$.

6.1.1 Results

The results of Case 1 are shown in Fig. 5. We can see from the convergence trajectory in Fig. 5a that if users make the optimal adjustments, the system converges to the ME in only 4 RPs and the converged QCS matches well with the theoretical value of Q_{me}. The error bars show the mean and standard deviation of the QCS values (indicated as 'STD') at review periods 5, 10, 15 and 20 over 10 simulation runs.

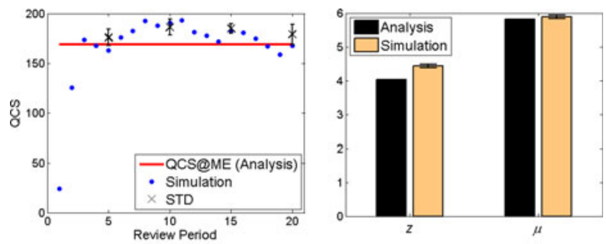
The significant value of the achieved QCS shows that the SP is able to achieve a good cumulative IQ that exceeds the IQ threshold, i.e., it is able to make the global decision with high confidence, as discussed in Section 3.2 and expressed by Eqs. (6) and (7), taking into account the fact that the QCS value is a timeliness-weighted sum of the IQ value of each contribution.

Fig. 5b compares the optimal contribution level z^* and consumption rate μ^*, which are the analytical values at the ME, 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 RPs, and the results for $m > 20$ are similar to $5 \leq m \leq 20$) for all users over 10 simulation runs. We can see that $\bar{z}$ and $\bar{\mu}$ are almost identical to the theoretical values, which validates our analysis. In this case, the error bars for the simulation results, which indicate the standard deviation of $\bar{z}$ and $\bar{\mu}$, were determined from 10×15 data points (10 simulation runs and RPs 6-20, in which the system is deemed to have converged).

In Case 2 where users make sub-optimal adjustments, Fig. 6a shows that slight fluctuations in QCS occur, but are nevertheless still centred around the theoretical Q_{me}. In Fig. 6b, 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. 7a shows that, interestingly, user heterogeneity raises the Q_{me} of the homogeneous case by 5 to 18 percent. This is attributed to the increased contribution level $\bar{z}$ as can be seen in Fig. 7b.

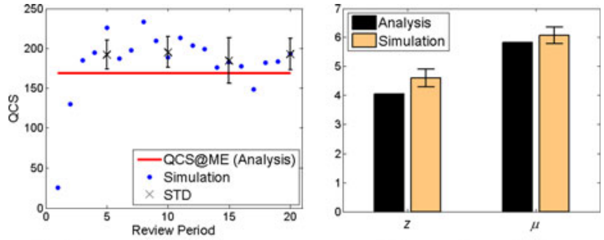
Finally, Fig. 8 shows the results for Case 4, which is the most comprehensive and realistic experiment. We can see that the results demonstrate a combination of the results



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

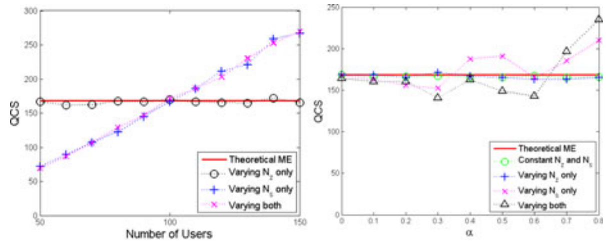
Fig. 7. Case 3: Heterogeneous users with optimal adjustments.

8. The Lambert W-function is a multi-valued function. The other solution of 0.508861 is not meaningful and should be ignored.



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

Fig. 8. Case 4: Heterogeneous users with sub-optimal adjustments.



(a) Varying number of users. (b) Varying α.

Fig. 9. Sensitivity of QCS with respect to varying parameters.

from Cases 2 and 3: the convergence trajectory fluctuates like in Case 2, and the QCS is higher than Q_{me} like in Case 3. This observation applies similarly to Fig. 8b as well. The key message is that, even under fairly high heterogeneity (maximal ± 50 percent deviation) and sub-optimal adjustments (standard deviation of 50 percent 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. The greater variability results in broader error bars for QCS, $\overline{z}$ and $\overline{\mu}$ compared to the earlier cases.

6.2 Sensitivity Analysis

We have thus far assumed that N_z and N_s are constant in our analysis and simulations, based on static expectation [28] as discussed earlier in Section 5.4. However, these are expected values and the actual values can vary in practice. This section investigates how unknown or wrong estimates of N_z and N_s will affect the performance.

First, we examine analytically the partial derivatives of Q_{me} with respect to N_z and N_s, respectively. By differentiating Eq. (38), we obtain

$$\frac{\partial Q_{me}}{\partial N_z} = \frac{\frac{C_3}{N_z^2}\log\frac{\beta Q_{me}}{p}}{1-\frac{C_3}{Q_{me}}(1-\frac{1}{N_z})},$$

$$\frac{\partial^2 Q_{me}}{\partial N_z^2} = \frac{\left[\frac{2C_3}{Q_{me}N_z^2} - \frac{C_3}{Q_{me}^2}\left(1-\frac{1}{N_z}\right)\right]Q'_z - \frac{2C_3}{N_z^3}\log\frac{\beta Q_{me}}{p}}{1-\frac{C_3}{Q_{me}}(1-\frac{1}{N_z})},$$

where $C_3 = \kappa p(1-\eta)N_s/c$ and $Q'_z \triangleq \partial Q_{me}/\partial N_z$,

$$\frac{\partial Q_{me}}{\partial N_s} = \frac{C_4 \log\frac{\beta Q_{me}}{p}}{1-\frac{C_4 N_s}{Q_{me}}},$$

$$\frac{\partial^2 Q_{me}}{\partial N_s^2} = \frac{2C_4 Q'_s/Q_{me} - C_4 N_s (Q'_s/Q_{me})^2}{1-\frac{C_4 N_s}{Q_{me}}},$$

where $C_4 = \kappa p(1-\eta)(1-\frac{1}{N_z})/c$ and $Q'_s \triangleq \partial Q_{me}/\partial N_s$. Note that the above four expressions are not fully reduced because the R.H.S. contains Q_{me} and its derivatives Q'_z and Q'_s. However, this suffices for evaluating:

$$\left.\frac{\partial Q_{me}}{\partial N_z}\right|_{N_z=100} = 0.02, \quad \left.\frac{\partial Q_{me}}{\partial N_s}\right|_{N_s=100} = 2.03,$$

$$\left.\frac{\partial^2 Q_{me}}{\partial N_z^2}\right|_{N_z=100} = -0.0004, \quad \left.\frac{\partial^2 Q_{me}}{\partial N_s^2}\right|_{N_s=100} = 0.0033,$$

with the other parameters kept at the same values as in previous simulation set-ups.

Therefore, by Taylor series expansion, we have

$$Q_{me} = 168.612 + 0.02(N_z - 100) - 0.0002(N_z - 100)^2 + \dots, \tag{39}$$

$$Q_{me} = 168.612 + 2.03(N_s - 100) + 0.0016(N_s - 100)^2 + \dots, \tag{40}$$

treating N_z and N_s as variables, respectively, and evaluating at their expected value of 100. The above shows that the theoretical Q_{me} is not affected much by the changes in N_z, but is more sensitive to those in N_s.

To verify this, we carry out simulations by varying N_z and N_s, respectively, from 50 to 150, and measure the corresponding QCS at market equilibrium by averaging the QCS over 10 simulation runs at $m = 10$. The reason for choosing $m = 10$ is because we observed that the convergence trajectories stabilize at $m = 4$ to 6.

We obtain three sets of results from varying N_z only, N_s only, and both N_z and N_s, respectively, and present them in Fig. 9a. We can see that the QCS is robust to changes in N_z, while it varies approximately linearly with N_s at a slope of about $\frac{270-168.6}{150-100} = 2.028$. These observations mirror well the expressions of Eqs. (39) and (40).

In order to examine how the QCS reacts to the changes of N_z and N_s in practice where user dropout and enrolment happen in a random manner, we carry out another set of simulations by varying both N_z and N_s according to the Gaussian distribution, i.e., as $N_z \hookleftarrow \mathcal{N}_{tr}(\overline{N}_z, \alpha\overline{N}_z, 0.1\overline{N}_z, 5\overline{N}_z)$ and $N_s \hookleftarrow \mathcal{N}_{tr}(\overline{N}_s, \alpha\overline{N}_s, 0.1\overline{N}_s, 5\overline{N}_s)$. Here, $\mathcal{N}_{tr}(\overline{m}, \sigma, a, b)$ denotes the *truncated normal distribution* with mean $\overline{m}$ and standard deviation σ and is bounded within $[a, b]$, and $\overline{N}_z = \overline{N}_s = 100$. This setup means that in 95 percent of the time, N_z is within $[(1-2\alpha)\overline{N}_z, (1+2\alpha)\overline{N}_z]$, and in the remaining 5 percent of the time, $\overline{N}_z$ can be as low as 10 or as high as 500. N_s is varied in a similar manner. This setup is able to cover very large variations that may arise in practice.

The results are shown in Fig. 9b. In the cases of constant N_z and N_s and the case of varying N_z only, the QCS is stable and remains close to the theoretical QCS at ME, which is consistent with the observations in Section 6.1 and the sensitivity analysis above. In the other two cases that involve N_s, we find that although certain deviations do appear, as predicted by the sensitivity analysis, the QCS does not deviate significantly from the theoretical ME, particularly when $\alpha < 0.7$, which safely covers fairly large variations in N_s. The stability of the QCS is due to the Gaussian distribution where the effect on the QCS of a higher-than-expected number of users at some moments is offset to a large extent by the effect of a lower-than-expected number of users at other moments.

7 Discussion

In the analysis in Sections 3.1 and 3.2, we have made the assumption that observations and contributions are independently and identically distributed for each contributor at different points in time as well as across different contributors. In practice, a contributor's observation and contribution may be correlated with his earlier observation and contribution, or those of another contributor.

The i.i.d. assumption is commonly made in the sensor networks literature for analytical tractability. In our case, taking correlation or dependencies into account would make the analysis significantly more difficult as it is hard to model how each user's observations are correlated with those of other users, or his own at different points in time. Recent results [33] indicate that the optimal sensor density to achieve the mean-squared error IQ metric in wireless sensor networks decreases as spatial correlation increases. Further study is required to see how these results can be extended to the participatory sensing case. Having said this, the i.i.d. assumption is likely to be justifiable in situations where the physical phenomenon, e.g., traffic or weather condition, changes over a short period of time, or can be different when sampled at slightly different locations, such as when the contributor is moving.

Furthermore, as stated in Section 2 of this paper, we assume that contributors are not malicious and do not misbehave by intentionally providing low quality data or attempt to gain unfair payoffs. Dealing with misbehaving users in participatory sensing is an important topic which would require a separate study altogether.

8 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. We have taken a market-based approach whereby each data contributor is motivated by obtaining a share of consumer payment from the service provider, according to his contribution rate and quality, and timeliness of the data; on the other hand, consumers choose the service consumption rate based on how well the QCS meets their satisfaction levels. Both contributors and consumers are not altruistic, but rational, behaving in the manner that maximizes their respective payoffs or utilities.

Our findings indicate that participatory sensing can be used for fairly reliable sensing purposes when certain incentives, e.g., monetary, and a market framework are 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 analyze the case when user contributions are in the form of continuous-valued measurements, instead of the case of binary decisions that has been considered in this paper.

Acknowledgements

Most of this work was done when Tie Luo was at the Department of Electrical and Computer Engineering, National University of Singapore, under the EDASACEP project which was part of the DVCaaS TSRP research programme funded by SERC, A*STAR Singapore.

References

[1] UCLA Urban sensing research area. (2009) [Online]. Available: 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 *Proc. CCSA, IEEE/ACM Int. Conf. Utility Cloud Comput.*, 2011.
[3] Waze - outsmarting traffic, together. (2014) [Online]. Available: http://www.waze.com/
[4] Buuuk, WeatherLah iPhone application. (2013) [Online]. Available: http://itunes.apple.com/us/app/weatherlah/id411646329?mt=8
[5] Buuuk, 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 *Proc. IEEE Conf. Sens., Mesh Ad Hoc Commun. Netw.*, Jun. 2012, pp. 425–433.
[7] A. Dua, N. Bulusu, W. Feng, and W. Hu, "Towards trustworthy participatory sensing," in *Proc. 4th USENIX Workshop Hot Topics Security*, 2009, p. 8.
[8] S. Reddy, D. Estrin, and M. Srivastava,, "Recruitment framework for participatory sensing data collections," in *Proc. 8th Int. Conf. Pervasive Comput.*, 2010, vol. LNCS 6030, pp. 138–155.
[9] D. Wang, T. Abdelzaher, and L. Kaplan, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proc. 11th ACM Int. Conf. Inf. Process. Sens. Netw.*, 2012, pp. 233–244.
[10] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. 18th ACM Annu. Int. Conf. Mobile Comput. Netw., MobiCom*, 2012, pp. 173–184.
[11] J.-S. Lee and B. Hoh, "Sell your experiences: A market mechanism based incentive for participatory sensing," in *Proc. Int. Conf. Pervasive Comput. and Commun.*, 2010, pp. 60–68.
[12] T. Luo, H.-P. Tan, and L. Xia, "Profit-maximizing incentive for participatory sensing," in *Proc. IEEE INFOCOM*, 2014, pp. 127–135.
[13] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *Int. J. High Perf. Comp. Appl.*, vol. 16, no. 3, pp. 293–313, 2002.
[14] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I - Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
[15] T. Q. S. Quek, D. Dardari, and M. Z. Win, "Energy efficiency of dense wireless sensor networks: To cooperate or not to cooperate," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 459–470, Feb. 2007.
[16] 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 *Proc. IEEE INFOCOM*, 2010, pp. 1–9.
[17] W.-L. Yeow, C.-K. Tham, and W.-C. Wong, "Optimizing application performance through learning and cooperation in a wireless sensor network," in *Proc. IEEE Military Commun. Conf.*, 2007, pp. 1–7.
[18] S. C. A. Thomopoulos, R. Viswanathan, and D. K. Bougoulias, "Optimal distributed decision fusion," *IEEE Trans. Aerospace Electron. Syst.*, vol. 25, no. 5, pp. 761–765, Sep. 1989.
[19] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Stat.*, vol. 16, no. 2, pp. 117–186, 1945.
[20] S. Ren, J. Park, and M. van der Schaar, "Profit maximization on user-generated content platforms," in *Proc. 49th Allerton Conf. Commun., Control, Comput.*, 2011, pp. 135–142.
[21] C. C. Holt, "Forecasting trends and seasonal by exponentially weighted averages," *Office Naval Res. Memorandum*, vol. 52, 1957.
[22] J. Park and M. van der Schaar, "Incentive provision using intervention," in *Proc. IEEE INFOCOM*, 2011, pp. 571–575.

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

[24] S. Ren, J. Park, and M. van der Schaar, "User subscription dynamics and revenue maximization in communications markets," in *Proc. IEEE INFOCOM*, 2011, pp. 2696–2704.

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

[26] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.

[27] O. J. Boxma and U. Yechiali, Poisson processes, ordinary and compound. (2007) [Online]. Available: http://www.math.tau.ac.il/~uriy/Papers/encyc57.pdf

[28] G. W. Evans and S. Honkapohja, *Learning and Expectations in Macroeconomics*. Princeton, NJ, USA: Princeton Univ. Press, 2001.

[29] M. H. Manshaei, J. Freudiger, M. Felegyhazi, P. Marbach, and J.-P. Hubaux, "On wireless social community networks," in *Proc. IEEE INFOCOM*, 2008.

[30] A. Zemlianov and G. de Veciana, "Cooperation and decision making in wireless multi-provider setting," in *Proc. IEEE INFOCOM*, 2005, pp. 387–397.

[31] J. Lambert, "Observationes variae in mathes in puram," *Acta Helvitica, Physico-Mathematico-Anatomicobotanico-Medica*, no. 3, pp. 128–168, 1758.

[32] L. Euler, "De serie lambertina plurimisque eius insignibus proprietatibus," *Acta Acad. Scient. Petropol.*, vol. 2, pp. 29–51, 1783.

[33] J. Wu and N. Sun, "Optimum sensor density in distortion-tolerant wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2056–2064, Jun. 2012.

Chen-Khong Tham received the MA and PhD degrees in electrical and information sciences engineering from the University of Cambridge, United Kingdom, and was an Edward Clarence Dyason Universitas21 fellow at the University of Melbourne, Australia. He is an associate professor at the Department of Electrical and Computer Engineering (ECE) of the National University of Singapore (NUS). His current research focuses on sensor network infrastructures and real-time sensor data analytics involving cyber-physical systems, wireless sensor networks, cloud computing, and participatory sensing. From 2007 to 2010, he was on secondment at A*STAR Institute for Infocomm Research (I2R) Singapore and served as a principal scientist and the department head of the Networking Protocols Department, and a programme manager of the Personalization, Optimization & Intelligence through Services (POISe) (Services) Programme. From 2006 to 2009, he was the programme manager of a multi-institution research programme on UWB-enabled Sentient Computing (UWB-SC) funded by A*STAR Singapore. He is an associate editor of the *IEEE Internet of Things Journal (IoT-J)* and is in the Editorial Board of the *International Journal of Network Management*. He was the general chair of IEEE SECON 2014, IEEE AINA 2011 and IEEE APSCC 2009 conferences. He and his coauthors received the Best Paper award at IEEE ICUWB 2008 and ACM MSWiM 2007.

Tie Luo received the PhD degree in electrical and computer engineering from the National University of Singapore (NUS). He is a scientist at the Institute for Infocomm Research (I2R), A*STAR, Singapore. He was a postdoctoral research fellow of the Communications & Networks Lab at NUS from 2011 to March 2012 and a postdoctoral research fellow of the Computer Networks & Distributed Systems Lab at NUS from late 2008 to 2010. His research interests include crowdsourcing and participatory sensing, network economics, pervasive computing, and mobile social networks.