Steered Crowdsensing: Incentive Design towards Quality-Oriented Place-Centric Crowdsensing

Ryoma Kawajiri
The University of Tokyo
Tokyo, Japan
kawajiri@ics.t.u-tokyo.ac.jp

Masamichi Shimosaka
The University of Tokyo
Tokyo, Japan
simosaka@ics.t.u-tokyo.ac.jp

Hisashi Kahima
Kyoto University
Kyoto, Japan
kashima@i.kyoto-u.ac.jp

ABSTRACT
Crowdsensing technologies are rapidly evolving and are expected to be utilized on commercial applications such as location-based services. Crowdsensing collects sensory data from daily activities of users without burdening users, and the data size is expected to grow into a population scale. However, quality of service is difficult to ensure for commercial use. *Incentive design* in crowdsensing with monetary rewards or gamifications is, therefore, attracting attention for motivating participants to collect data to increase data quantity. In contrast, we propose *Steered Crowdsensing*, which controls the incentives of users by using the game elements on location-based services for directly improving the quality of service rather than data size. For a feasibility study of steered crowdsensing, we deployed a crowdsensing system focusing on application scenarios of building processes on wireless indoor localization systems. In the results, steered crowdsensing realized deployments faster than non-steered crowdsensing while having half as many data.

Author Keywords
Mobile Crowdsourcing; Incentive Mechanism; Gamification

ACM Classification Keywords
H.4.m Information Systems Applications: Miscellaneous

INTRODUCTION
The proliferation of smart phones makes it increasingly feasible to build context-aware systems that gather large-scale sensory data. One of the important applications of context-aware systems is location-based services such as enhanced local search and recommendation, games, and social network services. Numerous research prototypes of context-aware systems have been demonstrated, enabling, for example, meter-level localization based on wireless technologies [11, 12, 13, 21], prediction of location categories that participants visit (e.g., workplaces, homes, or cafes) [15], and understanding of building-scale behavior analysis [25]. On the other hand, the quality of collected data is becoming a bottle-neck for promoting the application of the context-aware systems to commercial services.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
UbiComp '14, September 13 - 17, 2014, Seattle, WA, USA
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2968-2/14/09...$15.00.
http://dx.doi.org/10.1145/2632048.2636064

One approach for solving the data quality problems is to utilize advanced machine learning methods such as *semi-supervised learning*, and *transfer learning* to reduce the quantity of necessary data. The main idea of these learning methods is to extract knowledge from unlabeled data or other field data that are more abundant or lower cost. For example, Lane et al. [23] show activity recognition such as "Walking" and "Exercising" can be improved by transferring knowledge from other users who have similar with demographic information or sensory data. Even with these methods, the sensor data need to cover spatial, temporal, and demographic areas.

In recent years, alternative mainstream approach is *Opportunistic Crowdsensing*, which aims to increase the quantity of data collected by sensing users daily activities without burdening users. The most important related work for opportunistic crowdsensing system is CrowdSense@Place developed by Chon et al. [14]: one of the largest examples of mobile crowdsensing. They recruited 85 participants to collect data for two months in Seoul, South Korea. According to their reports, the crowdsensing can provide relatively high coverage levels in popular places even with a small number of contributors. Almost places in the real world are, however, unpopular and uncovered by directly increasing quantity of data. Therefore, some researchers proposed to utilize *Gamification*, which has been mainly used just to improve user experience or user engagement. The gamifications in most proposals aim to improve quality of data indirectly by directly increasing quantity. In contrast, Chon et al. [14] also reported that it is difficult for only growth of a population to achieve high place-temporal coverage of data collection, and is necessary to recruit spatially distributed participants. Just increasing quantity is not able to be scaled up due to monetary cost and time consumption. That is, as shown Figure 1, it is important to control the incentives by rewards such as game points or money for directly improving the quality of service rather than the size of data (e.g., localization and categorization accuracy).

To control users incentives for quality of services, *incentive design* has started to attract attention in crowdsensing field in recent years [30, 28, 27, 31, 22, 20]. In this context, gamifications are also used to guide users to uncovered location [30, 28, 27]. Due to noise of sensor data or sensing skills of users, increasing the coverage does not always directly im-

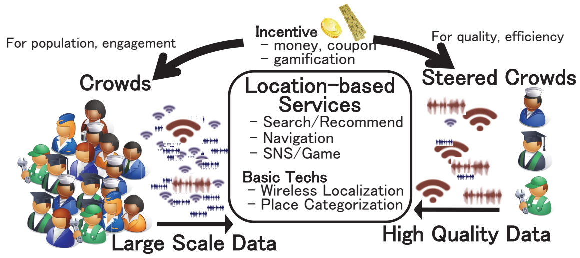


Figure 1. Comparison between existing crowdsensing and steered crowdsensing.

prove the quality of the services. In contrast, Koutsuopoulos [22] formalized the incentive mechanism as an auction model in crowdsensing by directly considering the quality. The auction model arisen from *crowdsourcing* context is appropriate for complex collection task. In contrast, it is inappropriate for Opportunistic Crowdsensing, because the auction process is unnatural and troublesome for the users.

In this paper, we propose a framework called *Steered Crowdsensing*, which directly increases the quality of service even in natural crowdsensing settings that keep the process simple to collect data. The simple process is as follows: the system determines the rewards for data collection such as *game points* and *coupon points*, then users decide whether participate to collect. We formalize the calculation of *points* paid for users by introducing a *quality indicator* of service in online machine learning setting. In this paper, by illustrating application scenarios, we confirm the feasibility of introducing steered crowdsensing to commercial location-based services in the real world and the merits for both the users and services of improving the quality of crowdsensing. In the formalization, by considering the probabilistic model of users and a trade-off of the explicit rewards of the services, we show possibilities to reduce both the users workload and the system payment to users.

For the first step of quantitative evaluation of steered crowdsensing, we deployed a crowdsensing system focusing on building processes of wireless indoor localization. We recruited 18 people for five weeks, and conducted the experiment on six floors of a university facility building. To assess the incentives of gamification and the direct controls of the quality separately, we gradually introduced them to the system week by week. Based on our deployment experience and analysis of the collected dataset, we report the following findings: 1) gamification with monetary rewards in our field trials strongly motivated some participants to collect data, 2) the strength of incentives is controllable by adjusting the game elements, and 3) quality of localization can be raised thanks to the user incentives controlled by our framework.

The contributions of this paper are as follows: 1) Proposal and formalization of a *steered crowdsensing* framework for effective data collection, 2) the quantitative evaluation of incentive design such as gamification on crowdsensing in the real world, and 3) illustration of effectiveness of controls for quality of service rather than quantity of data. We believe the analysis and findings we present provide valuable insights useful for builders of crowdsensing systems.

RELATED WORK

In this section, we describe place-centric crowdsensing and researches into steered crowdsensing. First, we explain place-centric crowdsensing and show the necessity of effective data collection scheme. Then, we illustrate incentive design ideas considered effective for improving the crowdsensing system.

Place-centric Crowdsensing

Place-centric crowdsensing has been widely studied [11, 14, 30, 28, 27, 26]. For example, Azizyan et al. [11] proposed SurroundSense, in which smartphone sensors are used to build a localization system on the basis of sensor fingerprints. Tuite et al. [30] introduced an online real world game called PhotoCity to motivate its participants to take photos at targeted locations to create 3D building models. In these applications, the systems need to gather data by considering noises and variances analyzed from collected data as well as quantity and coverage of data. In present systems, the incentive designs are elaborated and manually controlled for guiding to locations lacking data. In this paper, we formalize the incentive design by introducing *quality indicator* and realize systematical crowdsensing.

Incentive Designs making Crowdsensing Practical

Recently incentive designs succeeded in other fields are exploited to make the crowdsensing practical. We illustrate two directions: *quantity-oriented* design and *quality-oriented* design. These technologies discussed in both directions are able to be applied to direct control of data quality in steered crowdsensing.

Quantity-Oriented Incentive Design

There are many ideas to increase the quantity of data of crowdsensing, such as paying monetary rewards, embedding the collection process in the location-based services. *Gamification* is one of most popular ideas for increasing the quantity of data [17]. For example, gamification is applied to many social networks services, wearable devices for health services, and so on [4, 3]. Many useful heuristic design guides have been developed thanks to gamification for a long time in the video game field [32]. From gamification, research into quantitative evaluation for gamification has emerged. For example, Anderson et al. [10] build user behavior models for receiving badges on a Q&A website. Some researchers have applied gamification to crowdsensing [30, 28]. However, the game elements indirectly increase the quality by directly increasing quantity and coverage. We present formalization for directly increasing quality of data.

Quality-Oriented Incentive Design

Quality-oriented incentive design has been studied mainly in crowdsourcing field. In the crowdsourcing field, there are various commercial internet marketplaces for and researches into improving quality of data and reducing cost. In the crowdsensing field, a few commercial services [5, 2] and analytic researches [31, 22, 20] have just begun to emerge. Especially, Koutsuopoulos [22] formalized the incentive design on

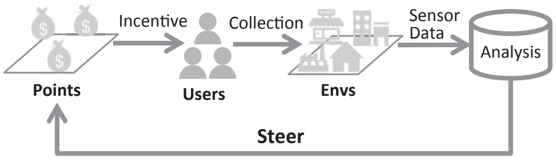


Figure 2. The flow of steering users.

crowdsensing as an auction model by considering costs paid to users and quality of service. The auction model includes a type of price negotiation process to reduce the cost and ensure the quality of data. This is appropriate for the system that one unit of collection task needs higher cost and to reduce the cost is necessary (e.g., writing a review of a restaurant, using special devices for measuring air pollution). However, in many crowdsensing settings, one unit of collection task is very low cost (e.g., choosing a category of users' current location, sensing with a common smart phone on background processes). In these settings, the negotiation process is troublesome for users and unnatural. Therefore, crowdsensing on auction model prolong the time until collection and might miss up-to-date data (e.g., real-time weather forecasting, urban parking space management). To remove the troublesome and collect data without burdening users, we omit the negotiation process and propose the calculation of the price directly in our formalization.

Reddy et al. [27] also formalized a recruitment process for crowdsensing to select participants for improving the quality of data and quantitatively evaluated the system by omitting negotiation processes. Instead of the negotiation processes, the system considers availability predictions of participants for the recruitment and reputations as data collections are introduced for incentivization. Therefore, the system relies on past coverage and participant behavior. Moreover, the reputation model cannot be used to directly control incentives for data collection or improve quality of data. In contrast, we directly formalize the way to incentivize participants for high-quality data, and the formalization does not necessarily require the participants' past information.

STEERING CROWDS IN LOCATION-BASED SERVICES

The process of steered crowdsensing is illustrated by Figure 2. Users on location-based services are given incentives by *points* of each location for data collection (e.g. game points, coupon points). Then, the users decide whether and where to collect data on the basis of the points of each location. The important part on steered crowdsensing is the feedback to the points on the basis of data analysis to directly improve the quality of services. In this section, we show application scenarios of steered crowdsensing and confirm that displaying *points* on each location to the users is useful to control their incentives to collect for data. Even if only monetary rewards and gamification are utilized, incentives are able to be largely raised. However, it is necessary to ensure the quality of service is improving. Then, for showing how to control the points, we formalize the scenarios by introducing a *quality indicator* for the location-based services. Finally, we specify the calculation of the points by introducing a simple user behavior model and improvements in quality.

Application Scenarios

In this section, by illustrating application scenarios, we confirm the feasibility of introducing steered crowdsensing to commercial location-based services in the real world and the merits for both the users and services by the improving the quality of crowdsensing.

For checking the feasibility of introduction, we need to consider the three following elements: 1) the systems must be able to display points to users to motivate them to move to a specific location, 2) the location information must be obtainable by check-in or authentication, and 3) the users must collect accompanying information of the location, for example, scanning Wi-Fi, taking pictures, inputting a category or writing a review of a store. There are mainly three types of location-based services: (a) game/social network services, (b) local search and recommendation, and (c) crowdsourcing the market for sensing.

Games and Social Network Services

We can find many commercial services: Ingress [6] and Colopl [1] for games, and Foursquare [4] and Yelp [8] for social network services. Users of the services are commonly given a kind of points by visiting specified locations. In many games, as a proof of being at the right location, localization embedded in smart phone or photography of monuments is requested. On most social networks services, the locations are reported by users themselves, since the accuracy of the localization technologies is at the building level ($\sim$ 100 meters). The improvements of data quality should increase the accuracy of *automatic check-in* [19].

Search and Recommend

There are many commercial services for local search and recommendations [7, 8].These services can motivate users by giving higher ranks or point cards to visit stores that need more data. Additional offers can be shown to users, for example, reviewing the stores or photographing meals. These data are able to improve search and recommendation.

Crowdsourcing Market for Sensing

There are a few commercial services that explicitly give monetary incentives to participants to recruit [5, 2]. Unlike games or coupons, frequent changes of monetary rewards might provoke user's antipathy. Recently, incentive design researches for these situations has emerged [31, 22, 20]. Improving of crowdsensing might reduce not only monetary rewards of the system but also man-hours of the users. Our formalization can be also used for crowdsourcing market and might be able to accelerate the speed of data collection.

Formalization towards Quality-Oriented Crowdsensing

In this section, we formalize a mechanism displaying *points* (coupon) that motivate users to improve the quality of data. For this formalization, we quantify the quality of data in online machine learning settings [18]. Note that the formalization proposed in this section is applicable to both problems of setting place categorization and wireless indoor localization. For simplicity, we focus on wireless localization here. In this paper, locations are given as labels $l \in \mathcal{L} = \{l_1, l_2, \dots\}$. Points displayed to users are denoted by $\boldsymbol{a}_t \in \mathbb{R}^{|\mathcal{L}|}$, where

$|\mathcal{L}|$ is size of all location labels $\mathcal{L}$. Wireless fingerprints are denoted by x. Here we consider a setting of crowdsensing where sensory data can be continuously collected and analyzed. Sensory data arrive at regular time intervals, i.e., additional data obtained at t-th session are represented by $D_t = \{x_j^{(t)}, l_j^{(t)}\}_{j=1:N_t}$. The additional data D_t are gathered by crowds while the points $\boldsymbol{a}_t$ are displayed with probability $p(D_t|\boldsymbol{a}_t)$. We define a *quality indicator* for quantifying quality of location-based services given by data until t-th session, and it is represented by $Q(\cup_{\tau=1:t} D_\tau)$. We also consider improvements in the quality to be differences in the quality caused by additional data D_t. That is, the improvements are defined by $S_t(D_t) \triangleq Q(\cup_{\tau=1:t} D_\tau) - Q(\cup_{\tau=1:t+1} D_\tau)$. Optimized points displayed to users $\tilde{\boldsymbol{a}}_t$ are obtained by maximizing the expectation of $S_t(D_t)$ over the distribution of $p(D_t|\boldsymbol{a}_t)$ as follows:

$$\begin{aligned}\tilde{\boldsymbol{a}}_t &= \arg\max_{\boldsymbol{a}_t} E_{D_t}[S_t(D_t)|\boldsymbol{a}_t] \\ &= \arg\max_{\boldsymbol{a}_t} \int p(D_t|\boldsymbol{a}_t) S_t(D_t) \mathrm{d}D_t. \end{aligned} \tag{1}$$

When effectively developing crowdsensing for indoor localization systems, the quality of the services $Q(\cup_{\tau=1:t} D_\tau)$ should be defined as the expected loss of the localization. It is generally difficult to calculate the improvements directly from all the additional data D_t [29]. We assume that the all the improvements $S_t(D_t)$ can be approximated by factorizing with the improvements of the one sample $s_t(x, l)$. When $N_t^{(l)}$ is defined as the number of samples of l including D_t, the approximation can be represented by

$$S_t(D_t) \triangleq \sum_{(x,l)\in D_t} s_t(x,l) - \lambda \sum_{l\in\mathcal{L}} \rho(N_t^{(l)}), \tag{2}$$

where $\rho(N_t^{(l)})$ is a regularization function preventing concentrate data to one label since the effects of multiple samples on the same class becomes small, and λ is a trade-off parameter of the regularization effect. The designs of the improvements $s_t(x, l)$ will be explained at the end of this section. The optimized points are rewritten as

$$\begin{aligned}\tilde{\boldsymbol{a}}_t = \arg\max_{\boldsymbol{a}_t} & E_{D_t}[\sum_{(x,l)\in D_t} s_t(x,l)|\boldsymbol{a}_t] \\ & - \lambda E_{D_t}[\sum_{l\in\mathcal{L}} \rho(N_t^{(l)})|\boldsymbol{a}_t]. \end{aligned} \tag{3}$$

For simplicity, we assume that the probabilities of each location are also independently calculated. Naturally, sensory data x are independently generated for each measurement. When $D_t^{(l)}$ represents the data of location l included in data D_t, the probabilities of whole data generation become the following stochastic point process:

$$\begin{aligned}p(D_t|\boldsymbol{a}_t) &= \prod_{l\in\mathcal{L}} p(D_t^{(l)}|\boldsymbol{a}_t) \\ &= \prod_{l\in\mathcal{L}} \left(p(N_t^{(l)}|\boldsymbol{a}_t) \prod_{(x,l)\in D_t^{(l)}} p(x|l) \right). \end{aligned} \tag{4}$$

Here, we assume this point process can be factorized into the number of samples on all locations and sensor data. The probability $p(N_t^{(l)}|\boldsymbol{a}_t)$ is representing a user behavior model: how often users visit the location l. Note that, from the independencies for locations and measurements, $\mathrm{d}D_t$ can be divided into $\sum_{l\in\mathcal{L}} \mathrm{d}D_t^{(l)}$, and $\mathrm{d}D_t^{(l)}$ can be also divided into $\mathrm{d}x$ and $\sum N_t^{(l)}$. Then, by substituting the above, the first term of Equation 3 becomes

$$\begin{aligned} & E_{D_t}[\sum_{(x,l)\in D_t} s_t(x,l)|\boldsymbol{a}_t] \\ &= \sum_{l\in\mathcal{L}} \int \left(p(D_t^{(l)}|\boldsymbol{a}_t) \sum_{(x,l)\in D_t^{(l)}} s_t(x,l) \right) \mathrm{d}D_t^{(l)} \\ &= \sum_{l\in\mathcal{L}} \left(\sum_{N_t^{(l)}=0}^{\infty} p(N_t^{(l)}|\boldsymbol{a}_t) N_t^{(l)} \int p(x|l) s_t(x,l) \mathrm{d}x \right) \\ &= \sum_{l\in\mathcal{L}} \left(E_{N_t^{(l)}}[N_t^{(l)}|\boldsymbol{a}_t]\ E_x[s_t(x,l)|l] \right). \end{aligned} \tag{5}$$

$E_{N_t^{(l)}}[N_t^{(l)}|\boldsymbol{a}_t]$ is the expectation of $N_t^{(l)}$, i.e., the frequency of visits to locations l. $E_x[s_t(x, l)|l]$ is an expectation of the improvements by a sample measured on location l and only depends on location l. For simplicity, we here rewrite it as follows:

$$s_t^{(l)} \triangleq E_x[s_t(x,l)|l]. \tag{6}$$

Similarly, the second term becomes

$$E_{D_t}[\sum_{l\in\mathcal{L}} \rho(N_t^{(l)})|\boldsymbol{a}_t] = \sum_{l\in\mathcal{L}} E_{N_t^{(l)}}[\rho(N_t^{(l)})|\boldsymbol{a}_t]. \tag{7}$$

$E_{N_t^{(l)}}[\rho(N_t^{(l)})|\boldsymbol{a}_t]$ also depends on the user behavior models. Then, the optimized points $\boldsymbol{a}_t$ can be represented by

$$\begin{aligned} & \tilde{\boldsymbol{a}}_t = \arg\max_{\boldsymbol{a}_t} \\ & \sum_{l\in\mathcal{L}} \left(s_t^{(l)} E_{N_t^{(l)}}[N_t^{(l)}|\boldsymbol{a}_t] - \lambda E_{N_t^{(l)}}[\rho(N_t^{(l)})|\boldsymbol{a}_t] \right). \end{aligned} \tag{8}$$

Since this equation shows the optimized points $\boldsymbol{a}_t$ depend on the user behavior model, we consider the model in the next section.

User Behavior Modeling

In this section, by considering the specific model of the effects of displaying points to user behavior $E_{N_t^{(l)}}[N_t^{(l)}|\boldsymbol{a}_t]$, we show how to calculate points to users to improve the quality. We assume the points motive users to collect data. We simply consider the frequencies of collection by users are proportional to the points of the location.

$$E_{N_t^{(l)}}[N_t^{(l)}|\boldsymbol{a}_t] = c a_t^{(l)}, \tag{9}$$

where c is a constant number representing the effects of the points. To avoid concentrating the points in the same location, for a regularization function ρ, we apply two norms commonly used as $\rho(N_t^{(l)}) = (N_t^{(l)})^2$. By instituting the above, Equation (8) becomes the quadratic programming of $\boldsymbol{a}$; like below

$$\tilde{\boldsymbol{a}}_t = \arg\max_{\boldsymbol{a}_t} \sum_{l \in \mathcal{L}} a_t^{(l)} s_t^{(l)} - c\lambda (a_t^{(l)})^2. \tag{10}$$

By solving the above and resetting the constant variable, the steered elements become

$$\tilde{a}_t^{(l)} = \frac{1}{2c\lambda} s_t^{(l)}. \tag{11}$$

That is, the points are simply proportional to the expected improvements of location l and inversely proportional to the effects c of the point to the user. The resultant point displayed to the users is intuitive. However, the point can be controlled by installing accurate user behavior models. Specifically, the probabilities of the user behaviors become more complex. For example, the effects of the points c can vary in accordance with locations, time, and users. Costs can be reduced by selecting the most effective time or users even in the same location. Because, we did not have data for estimating the parameters, we do not consider these differences in the paper. We will instead show the analysis of the experiment below for the next steps of deployments.

Formalization with Location-based Services

Thus far, we have considered only the quality of the crowdsensing system. For collaborating with location-based services, the rewards, promotion, and the cost of the services should also be considered. When $R_t(D_t)$ denotes a reward or cost function given additional data D_t such as purchases of products or payment of points, the reward function considering the quality of crowdsensing becomes

$$\tilde{R}_t(D_t) = R_t(D_t) + \mu S_t(D_t), \tag{12}$$

where μ is the rate to change the quality to the rewards. In general, the points have constraints such as minimum $a_{\min}$ and maximum $a_{\max}$ ($a_{\min} \leq \boldsymbol{a}_t \leq a_{\max}$). Then, the optimized steered element can be calculated by

$$\begin{aligned} \tilde{\boldsymbol{a}}_t &= \arg\max_{\boldsymbol{a}_t} E_{D_t}[\tilde{R}_t(D_t) | \boldsymbol{a}_t] \\ &\text{s.t.} \quad a_{\min} \leq \boldsymbol{a}_t \leq a_{\max}. \end{aligned} \tag{13}$$

This can integrate the short-term profits of services and the long-term profits due to the quality of crowdsensing. For example, when the system pays the costs to users such as coupons or monetary rewards, the cost function becomes

$$R_t(D_t) = -\sum_{l \in \mathcal{L}} a_t^{(l)} N_t^{(l)}. \tag{14}$$

By substituting the above, Equation (13) also shows quadratic programming of $\boldsymbol{a}$, and the optimized steered elements also becomes

$$\tilde{a}_t^{(l)} = \max(a_{\min}, \min(a_{\max}, c' s_t^{(l)})), \tag{15}$$

where c' is a rewritten coefficient. This form is almost the same as Equation (15) except for the limitation of the minimum $a_{\min}$ and the maximum $a_{\max}$.

Improvements in Quality

As described above, the improvements $s(x, l)$ in quality by one sample (x, l) become important for calculating the points. The designs of the improvements $s_t(x, l)$ from one sample (x, l) can be done in various ways depending on the domain of crowdsensing. Here, referencing ideas in Active Learning [29, 24] fields, we introduce examples in some crowdsensing problem settings. Cohn et al. [16] show that the variance of estimation errors is the most effective way to support the expected error reduction. For example, on wireless indoor localization, one sample (x, l) consists of location l and Wi-Fi fingerprint x. The crowdsensing system estimates the probability $p(\tilde{l}|x)$ that fingerprint x is measured on location $\tilde{l}$. When $\delta_{\text{loc}}(l, \tilde{l})$ denotes a error function that is commonly a distance between l and $\tilde{l}$, the improvements are defined as

$$s(x, l) = \text{Var}_{\tilde{l}}[\delta_{\text{loc}}(l, \tilde{l}) | x]. \tag{16}$$

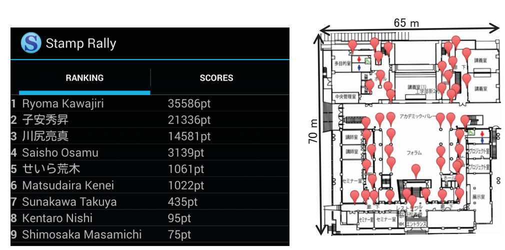
In place categorization, a category c labeled by a user is also given with sensor data x. Thus, one sample can be written as (x, c, l). The crowdsensing systems estimate probability $p(\tilde{c}|l, x)$ that location l with sensor data x is categorized to $\tilde{c}$. When $\delta(c, \tilde{c})$ is an error function (which commonly becomes 1 if $\tilde{c}$ is a misclassified label), the improvements can be denoted by

$$s(c, x, l) = \text{Var}_{\tilde{c}}[\delta_{\text{cat}}(c, \tilde{c}) | l, x]. \tag{17}$$

Many crowdsensing systems directly estimate locations or categories, not probability. In these cases, Abe et al. [9] proposed *a bagging method* that builds multiple estimators from multiple datasets resampled randomly and calculates pseudo probability outputs from the estimations.

EXPERIMENTAL SETTINGS

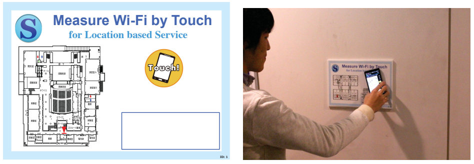
We built a crowdsensing system on a university facility by focusing on deployment of wireless indoor localization as one of the fundamental technologies of mobile sensing. The system is designed to act like a gamified location-based service in retail facilities. For the first step of quantitative evaluation, we compare steered crowdsensing with crowdsensing without data analysis feedback and analyze the relationship of user behavior with contexts such as time, place, and users.Game elements for controlling the incentives of users were amounts of points at each measurement location. Conditions of game elements were changed each week. There are mainly three types of crowdsensing: Naïve CS, Gamified CS and Steered



(a) Game: scoring and ranking.

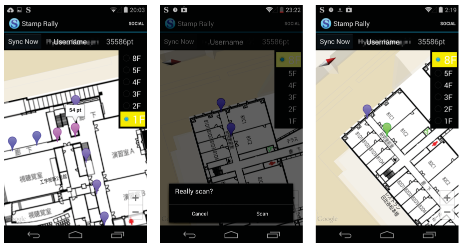
(b) Poster locations on the 2nd floor.

Figure 3. System overview.



(a) Appearance example. (b) Tapping with smartphone.

Figure 4. Smartposters with NFC.



(a) Show points. (b) Confirmation. (c) Collected.

Figure 5. Screenshots of Smartphone App.

CS. In Naïve CS, participants were provided a mobile app that only had measurements functions implemented. On Gamified CS, the app shows the score depending on the amount of points of a measurement location and the ranking of the score (Figure 3(a)). On Steered CS, the points are controlled to reflect contribution toward data quality as shown by Equation (15). The experiment had 18 participants and was held on six floors of a building for five weeks. The size of one floor is about 60 meters $\times$ 70 meters (Figure 3(b)). In the following, we detail the design of the crowdsensing system and the conditions of data collection in this study.

Mobile App

To eliminate the effects of skill variances, actions such as checking-in on the location-based service are realized by simply tapping a *smartposter*. When tapping on the poster with a mobile phone installed with the app, the embedded location information is read as soon as it starts to scan Wi-Fi networks. NFC tags are used to embed location information. Figure 4 shows an example (the size is A4 = 297 mm $\times$ 210 mm) of one of our posters. Figure 5 shows basic functions of the app. Poster locations and points were shown on a floor map (Figure 5(a)). The range of the points was set from 0 to 100, i.e., $a_{\min} = 0$ and $a_{\max} = 100$. When a poster was touched, the app asked the user for confirmation to save the sensor data (Figure 5(b)). The color of locations sensed that day became green (Figure 5(c)). The scores and ranking are calculated every 15 minutes, i.e., the session intervals are set to 15 minutes. We updated the functions of the app by using a service for privately distributing mobile apps.

Data Collection

Location

The location was a 12-floor university facility building that has entrances on the 1st and 2nd floors. We used six floors: 1st to 5th and 8th. There are mainly lecture rooms on the 1st to the 4th floors, a library on the 5th floor, and laboratories on the 8th floor. A total of 190 posters were placed at intervals of about five meters on the walls of the hallways since this corresponds to distance between shops in retail facilities (Figure 3(b)). There are many campus access points.

Participants

We recruited 18 undergraduate students studying in the facility. Nine participants mainly use the 1st to the 3rd floors, and the rest use the 4th floor. Participants were informed that the aim of the experiment was to evaluate crowdsensing for wireless indoor localization and were paid at least JPY 5,000 ($\approx$ USD 50) for all five weeks. The gamification was started in the middle of the experiment, and participants were told the rules of the game and payment amounts at the same time. Thirteen participants used Nexus 7 tablet computers (2012), and the other five used smart phones including one Nexus 5, two SHARP SH series devices, and two Sony SOL series devices. Eleven participants were lent Nexus 7 tablets (2012) from our laboratory, and the rest used their own devices.

Collection Rules

Similar to real world commercial applications, the number of measurement times were limited to three times in every session (= 15 minutes) and once for the same poster per day. Each participant was requested to gather at least 25 measurements every week. The intentions of this rule were to simulate passive incentives not affected by gamification and to obtain a sufficient amount of data in non-gamified weeks. We limited the collection during the night and weekends for safety reasons; participants were allowed to collect data between 8:00 and 20:00 from Monday to Friday. We sent reminder mails to participants each Thursday.

Flow of Experiment

The experimental period was from January 10th to February 14th, 2014. The conditions of game elements were changed each week, and the scores were reset at the beginning of each week. The 1st week was Naïve CS; the functions of the mobile app were provided only for measurement. The 2nd week was Gamified CS. After the end of the 1st week, we announced the rules of the game and payment amount and updated the mobile app. To accelerate the incentives of the participants, we designed additional payments: the top ranked participant of each week was paid JPY 10,000 ($\approx$ USD 100), the 2nd JPY 7,000 ($\approx$ USD 70), the 3rd JPY 4,000 ($\approx$ USD 40), the 4th JPY 2,000 ($\approx$ USD 20), and the 5th to 7th JPY 1,000 ($\approx$ USD 10). The points of the posters were fixed to $a_t = 10$ in this week. From the 3rd week, the points of the posters varied for steering participants. We announced it to the users at the beginning of the week [1]. The points are calculated on the basis of the analysis of collected data in this week. We give details in the next sections.In the 4th week, to confirm whether differences between the 2nd and 3rd weeks were caused by novelty effect or not, we set back to Gamified CS to evaluate the effects of point variance without data analysis, but the points randomly varied.Thus, the points were set

[1]On the 4th day of the 3rd week, we received some complaints from participants that almost all points were below 10. We therefore limited minimum points of posters to 10 on the 5th day of this week.

Week	Name	Gamification	Points $\boldsymbol{a}_t$ calculation	Notes
1st	Naïve CS	-	-	-
2nd	Gamified CS	score & rank	const.	-
3rd	Steered CS	score & rank	steer (Eq. (15))	-
4th	Gamified CS	score & rank	rand. (1st day), const. (2-5th day)	-
5th	Mixed CS	score	const. (1-3rd day), steer (4-5th day)	The 2nd day was a national holiday. Some students returned home.

Table 1. The flow of the experiment.

uniformly random from $a_{\min} = 5$ to $a_{\max} = 50$ on the 1st day and were fixed $a_t^{(l)} = 10$ on the other days. In the 5th week, to evaluate the effects of the ranking, we eliminated the ranking and the payments were in proportion to their scores. The calculation of points was constant on from the 1st to 3rd days while steering was applied on the 4th and 5th days. From this week, many students had started to return home. Since the 2nd day of this week was a national holiday, the assignments were reduced to 20 times for this week. Table 1 summarizes the flow of the experiment.

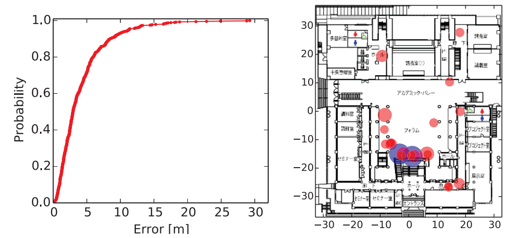
Points Calculation for Steered CS

In steered CS, the points are calculated on the basis of Equation (15) in every session. In this simple equation, the points are simply proportional to the expectation $s_t^{(l)} \triangleq \int s_t(x,l)p(x|l)\mathrm{d}x$ of improvements in quality on each location except for the limitation of the minimum $a_{\min} = 10$ and the maximum $a_{\max} = 100$ in this experiment. The expectations $s_t^{(l)}$ of improvements in quality can be calculated by a sampling method. The generation models of RSSIs on each location are built by Gaussian distributions from collected data. By setting prior distribution of the mean of Gaussian distribution, the points of locations that have no data become relatively high. Therefore, the coverage of data tends to increase rapidly in the first stage of data collection. After the coverage is accomplished, the points reflect the variances of localizations and keep increasing the quality of localizations.

Localization Performance on Controlled Experiment

In this section, we report a preliminary experiment to confirm the basic performances of wireless indoor localization in our system. Many wireless indoor localization methods have been developed. Since our steered crowdsensing algorithm can be thought of as meta-algorithm of online training for wireless localization methods, any localization method can be applied to our system. We used a method proposed by Kawajiri et al. [21]. The method perform well thanks to exploiting a loss metric for the learning step in proportion to distance error. However, their classification method can only predict poster locations, not locations among posters. Therefore, we extend the classification method to a regression method that can predict middles of posters by averaging some classes of candidates [13].

We collected test data and training data for evaluating localization accuracy. Considering data variance according to time, we collected the test data for two days at one-week intervals. Considering data variance according to devices, we



(a) Cumulative distance error. (b) Locations with large errors (> 10 meters).

Figure 6. Performance of localization.

used two types: a smartphone (Nexus 4) and a tablet computer (Nexus 7, 2012). Test data were collected in a total of 719 samples by inputting locations at intervals of about five meters by tapping a display showing floor maps. Training data were collected by using the smartposters. To tap once each poster by one worker needs more than one hour labor. We collected training data in a total 3678 samples, containing at least 11 samples per poster for few weeks. Figure 6 shows the results. Figure 6(a) shows cumulative distance error. The median of the errors is 2.9 meters, and the 95 percentile of the errors is 11.2 meters. Figure 6(b) shows locations estimated with errors of more than 10 meters on the 2nd floor. Blue circles show locations estimated on different floors.

ANALYSIS AND DISCUSSION

On the basis of our deployment experiments and analysis of the collected dataset, we report three findings: 1) gamification with monetary rewards in the experimental settings can reinforce the incentive of some participants for collection, 2) user behaviors are controllable by adjusting the game elements thanks to the measurement frequency $E[N_t^{(l)}|a_t^{(l)}]$ in relation to points, and 3) quality of localization can be raised thanks to the user incentive controlled by our framework. Finally, we discuss the limitations of the experiment.

Incentive Reinforcement

Comparing data collected for each week, we consider the incentivization of participants by gamification and confirm that skills variances and novelty effects are vanishingly small. Figure 7 shows the amount of data according to weeks: 629 samples in the 1st week, 1946 in the 2nd, 1066 in the 3rd, 2085 in the 4th, and 733 in the 5th. Figure 8 shows the amount of data according to days, and Figure 9 shows the amount of data according to participants for each week. The vertical black lines show their assignments. The participants are sorted by the number of samples. The horizontal black lines divide participants into those who received additional payment and those that did not from the 2nd to 4th weeks. Figure 10 shows the transition of points and the poster coverage for each week. The poster coverage is defined by $(\#\ \text{measured posters})/(\#\ \text{all posters})$. The blue lines show the median of the points at that time, and blue regions show the points range from the 30th to 70th percentile. In the 1st week on Naïve CS, most participants had done only their basic assignments. In the 2nd week on Gamified

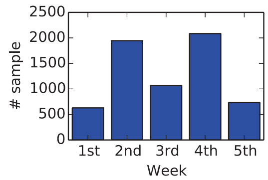


Figure 7. Amount of data according to weeks.

CS, there is a marked increase on the 1st week. The differences in the weeks strongly depend on each person (Figure 9). However, some people remarkably increased data and seemed to be strongly encouraged by gamification and monetary rewards. In the 3rd week on Steered CS, a lot of data were collected too. The points became very high early time in the first day and reduced later in the week. After a few measurements, all points of the posters quickly became low. By the end of the 1st day, most points of posters were around 10 points. However, as shown in below analyses, these low scores means the quality of localization is sufficient. The amount of data in the 3rd week was, however, less than that in the 2nd week. In the 4th week on Gamified CS with random points, the amount of data increased again and was more than that in the 2nd week. This result shows that the novelty effect is negligible and the differences are mainly caused by the ways to control the points. In the 5th week, some people returned to their homes. Therefore, the amount of the data is reduced.

After the experiments ended, we presented participants with a questionnaire about what motivated or discourage them to collect data. Fifteen participants responded. Questions and answers were written in Japanese. Many participants commonly mentioned that the competition motivated them in the early days of each week, but the large point differences between them and the top ranked participants discourage them.

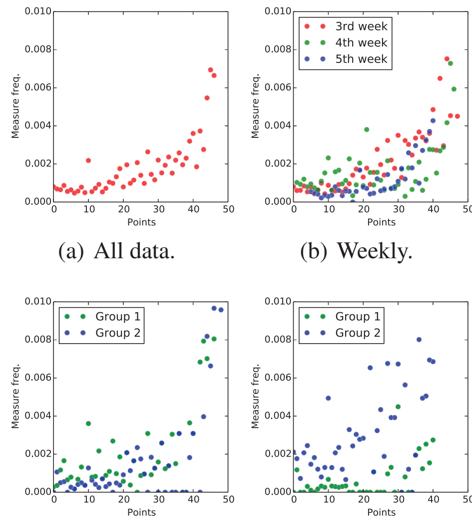
"I tried to become a top ranker at first, but when I realized it was impossible to go up in the rankings, I could not get motivated at all."

Especially in the 3rd week on Steered CS, the large differences in the points compared with the early days had a discouraging effect.

"The variance of points is not bad, but the big difference in points depending on days really discouraged me."

Incentive Control and User Behavior Model

We first confirm that the incentives of users for collecting data are controllable by changing the points. Then, for more efficient data collection, we illustrate that user behaviors are different among floors, people, and weeks. Figure 11 shows the relationship between points $a_t^{(l)}$ and frequency of measurements $E[N_t^{(l)}|a_t^{(l)}]$. The horizontal axes show points displayed to participants, and the vertical axes show the frequency of measurements. Figure 11(a) shows the relationship in all data. To calculate the points for Steered CS in this experiment, we assume the user behavior models in which the frequencies of collection by users are proportional to the points of the location. In fact, the frequency tends to increase



(a) All data. (b) Weekly.

(c) 2nd floor. (d) 4th floor.

Figure 11. Points $a_t^{(l)}$ and frequency $E[N_t^{(l)}|a_t^{(l)}]$ according to contexts (weeks, floors, users).

as a function of points. This shows that the incentive for collecting can be controlled by the game elements. In the below section, we explore spatial, temporal, and demographic variances for further analyses.

Weekly Variances

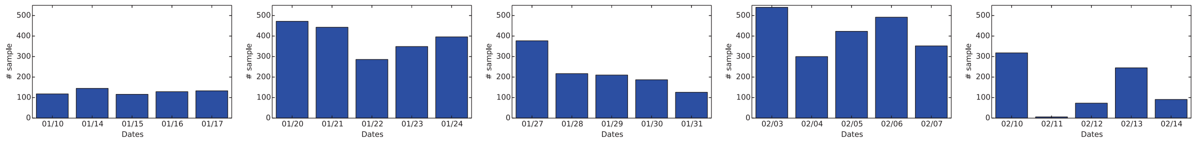
Figure 11(b) compares the incentives among the 3-5th weeks. The differences in each week can be seen. In the 4th week, the frequency of collection in low points (around 10 to 20) is comparatively higher than in the other weeks. Some participants reported that they had time for collection because they did not have lectures this week. On the other hand, in the 5th week, more than 30 points are required to incentivize participants strongly. In the application scenario of a retail facility, for example, holding an opening day or avoiding unpopular weeks are more effective for collecting data.

Variances of People and Locations

Figures 11(c) and 11(d) compare the relationships between points and the frequency with the two groups of users on each floor. Group 1, shown by green circles, mainly uses the 1st to 3rd floors, and group 2, shown by blue circles, mainly uses the 4th floor for lectures. Group 2 is very active on the 4th floors but quite inactive on the 2nd floors. In contrast, group 1 is active on from the 2nd floors but quite inactive on the 4th floor. These results show that the reactions of points vary in accordance with locations and users. In fact, some participants mentioned the following in the survey:

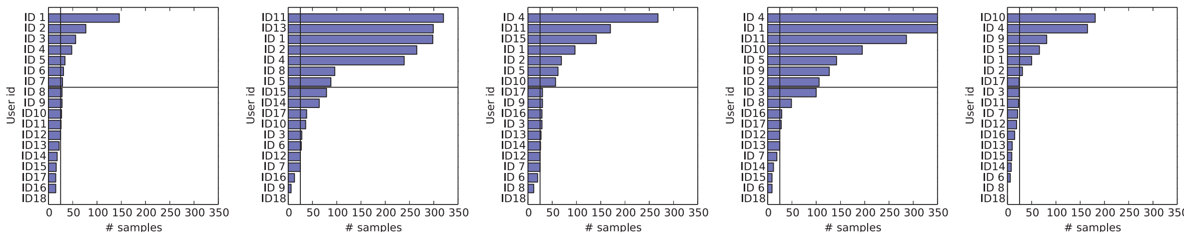
"I just completed the assignments every week, so I always collected at similar places (e.g. nearby lecture rooms)."

For effective collection by using this analysis, for example, group 2 can be mainly targeted on the 4th floor with low points. If there is a deadline for deployment of the localization system, the points of the 1st and the 2nd floors should be set comparatively high.



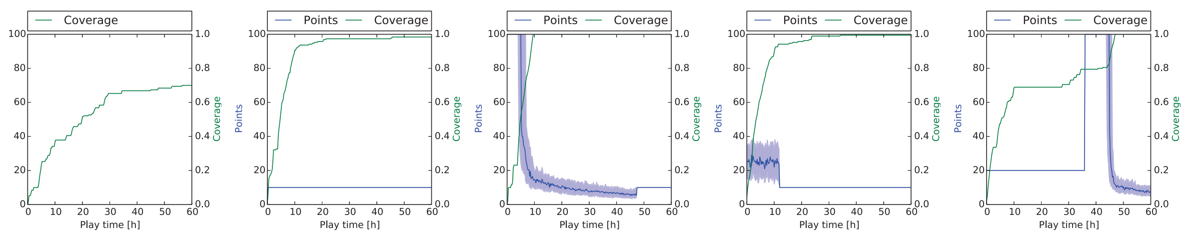
(a) 1st week (Naïve CS). (b) 2nd week (Gamified CS). (c) 3rd week (Steered CS). (d) 4th week (Gamified CS). (e) 5th week (Mixed CS).

Figure 8. Amount of data according to dates.



(a) 1st week (Naïve CS). (b) 2nd week (Gamified CS). (c) 3rd week (Steered CS). (d) 4th week (Gamified CS). (e) 5th week (Mixed CS).

Figure 9. Amount of data according to participants: The vertical axes are sorted by the number of samples.



(a) 1st week (Naïve CS). (b) 2nd week (Gamified CS). (c) 3rd week (Steered CS). (d) 4th week (Gamified CS). (e) 5th week (Mixed CS).

Figure 10. Transitions of points and coverage of data.

Quality Improvements by Steering

This section explains that steering can realize a high quality of data, especially the accuracy of the indoor localization.

Coverage

Figure 12 shows locations of 0 samples when the first 200 samples were given on 2nd floor. There is no big difference between Naïve CS in the 1st week (Figure 12(a)) and Gamified CS in the 2nd week (Figure 12(b)). Gamified CS with random points in the 4th week (Figure 12(d)) and on 1st day of the 5th week (Figure 12(e)) slightly distributed the locations of 0 samples. However, the number of the locations of 0 samples was not affected. On the other hand, Steered CS on the 3rd (Figure 12(c))f and 4th day of the 5th week (Figure 12(f)) reduced the locations of 0 samples. Specifically on the 4th day of the 5th week, the 2nd floor has no locations of the 0 samples. These results show that controlling the quality improvements can reduce location bias of data.

(a) 1st week (Naïve CS). (b) 2nd week (Gamified CS). (c) 3rd week (Steered CS).

(d) 4th week (Gamified CS). (e) First 3 days of 5th week (Gamified CS). (f) Last 2 days of 5th week (Steered CS).

Figure 12. Uncovered locations with first 200 samples on the 2nd floor.

Localization accuracy

Figure 13 shows the relationship between the number of samples and the error of estimations in each week. The horizontal axis shows the number of samples, and the vertical axis shows the error of estimation. The filled markers show 95 percentile errors, and the unfilled markers show median errors. The horizontal dashed line and the dotted line respectively show the 95 percentile error and median error on the preliminary controlled experiment. Naïve CS in the 1st week and Gamified CS with constant points in the 2nd week

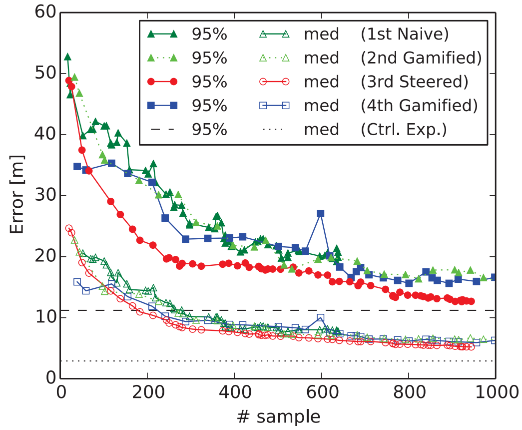


Figure 13. The accuracy of localization and the number of samples.

	median	95%tile
1st & 2nd	0.05	0.061
1st & 4th	0.15	0.31
2nd & 4th	0.93	0.69

(a) Not Steered CSs.

	median	95%tile
1st	1.2×10^{-4}	1.6×10^{-4}
2nd	2.1×10^{-5}	2.1×10^{-6}
4th	5.4×10^{-6}	1.5×10^{-6}

(b) Steered CS and others.

Table 2. The p-values obtained from paired t-test applied to the performances of each week from 100 to 950 samples with intervals of 100 samples.

draw similar curves. Gamified CS with random points in the 4th week shows comparatively small errors until around 100 samples and draws similar curves from around 100 samples. According to a paired t-test, the differences between them are not statistically significant at the 0.01 level (Table 2(a)). In contrast, Steered CS (the 3rd week) draws curves almost below the other curves. The 95 percentile errors of Steered CS shows especially large improvements: less than 20 meters with 250 samples. Other CSs require more than 500 samples for the same performance, i.e., Steered CS needs only half the amount of data. The differences between steered CS and others are statistically significant at the 0.01 level according to a paired t-test (Table 2(b)). In addition, we compare the estimators of each week with 500 samples. The estimator of Gamified CS in the 2nd week shows 2.5% (18/719 samples) with errors of more than 25 meters, whereas the estimator of Steered CS in the 3rd week shows only 1.9% (7/719 samples). The above shows that simple gamification is not sufficient and requires and the control of game elements for high quality of data. On the other hand, even with less than 1000 samples, the curves of Steered CS are close to the lines of the controlled experiment using the dataset of 3678 samples collected by hard work as shown preliminary experiment. Collecting data on the basis of quality may possibly lead to data being collected more efficiently than in the naively controlled experiment.

Discussion and Limitation

In this section, we discuss the possibilities to collect data more effectively and the limitations of this experiments.

Our experiment could achieve the best effects for only gamification. Especially on Steered CS, there were largely different scores among participants early on. This reduced the total number of samples. This is not a problem for the quick deployment of the system in the short-term. However, in the case of long-term deployments (e.g., detecting the variance of categories and Wi-Fi access points), users' engagements is required. One solution for this is that participants in the bottom of ranking should be treated favorably. However, too much balancing might make the top participants feel unmotivated. For this kind of problem, dynamic game difficulty balancing is proposed in gamification [32]. Additionally, for the simplicity of the computation, we used the same models for all users, times, and locations in this experiment. However, as shown above, user behavior depends on various contexts. Therefore, embedding user behavior models could lead to the efficiency improvements of data collection.

The experiment described in this paper was limited in spatial scale (one building) and to a single application scenario (wireless localization). In larger scale crowdsensing, the mobility cost cannot be negligible for incentivizing users to collect. There are other applications for which the demands of the data collection vary more dynamically (e.g. parking availability reporting, real-time weather forecasting). To evaluate in those application scenarios, we will require precise forecasting technologies for user behaviors (e.g. mobility patterns) and objectives for sensing (e.g. variances of weather). Our formalization can be, however, compatible with other user behavior models and the quality definitions.

CONCLUSION

In this paper, we proposed a new crowdsensing framework with incentive design, called *steered crowdsensing*. To increase the quality rather than quantity of data directly, the framework controls incentives by introducing gamifications to location-based services, while continuous data analyses consider the data qualities. We also proposed the application scenarios of steered crowdsensing and the formalization of steered crowdsensing with online learning settings for raising the quality of services. For the first step of quantitatively evaluating steered crowdsensing, we deployed a crowdsensing system focusing on wireless indoor localization. We recruited 18 people for five weeks on six floors of a university facility building. We varied scoring methods and used gamified and non-gamified conditions to compare them. Based on our deployment experiment and analyses, we found the following findings: 1) gamification with monetary rewards in our field trials strongly incentivized some participants to collect data, 2) strength of incentives is controllable, and 3) quality of localization can be raised thanks to the user incentive controlled by our framework. In the results of our experiment, steered crowdsensing realized deployments of wireless indoor localizations faster than non-steered crowdsensing while having only half the amount of data.

Our future work includes a long-term demonstration experiment in real location-based services and retail facilities. In addition, as mentioned on our analysis, embedding user behavior models may possibly further improve the efficiency of data collection. For application to commercial services, constraints such as budgets or deadlines should be considered for deploying crowdsensing systems.

REFERENCES

1. Colopl. `http://colopl.co.jp`.
2. Field agent. `http://www.fieldagent.net`.
3. Fitbit. `http://www.fitbit.com`.
4. Foursquare. `https://foursquare.com`.
5. Gigwalk. `http://gigwalk.com`.
6. Ingress. `http://www.ingress.com`.
7. Tabelog. `http://tabelog.com`.
8. Yelp. `http://www.yelp.com`.
9. Abe, N., and Mamitsuka, H. Query learning strategies using boosting and bagging. In *Proc. ICML 1998*, Morgan Kaufmann Publishers Inc. (1998), 1–9.
10. Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J. Steering user behavior with badges. In *Proc. WWW 2013*, International World Wide Web Conferences Steering Committee (2013), 95–106.
11. Azizyan, M., Constandache, I., and Roy Choudhury, R. Surroundsense: mobile phone localization via ambience fingerprinting. In *Proc. MobiCom 2009*, ACM (2009), 261–272.
12. Bahl, P., and Padmanabhan, V. N. RADAR: An in-building Rf-based user location and tracking system. In *Proc. INFOCOM 2000*, vol. 2, IEEE (2000), 775–784.
13. Bahl, P., Padmanabhan, V. N., and Balachandran, A. A software system for locating mobile users: Design, evaluation, and lessons. *online document, Microsoft Research* (2000).
14. Chon, Y., Lane, N. D., Kim, Y., Zhao, F., and Cha, H. Understanding the coverage and scalability of place-centric crowdsensing. In *Proc. UbiComp 2013*, ACM (2013), 3–12.
15. Chon, Y., Lane, N. D., Li, F., Cha, H., and Zhao, F. Automatically characterizing places with opportunistic crowdsensing using smartphones. In *Proc. UbiComp 2012*, ACM (2012), 481–490.
16. Cohn, D. A., Ghahramani, Z., and Jordan, M. I. Active learning with statistical models. *CoRR cs.AI/9603104* (1996).
17. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. Gamification. using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, ACM (2011), 2425–2428.
18. Fu, K. S. *Sequential methods in pattern recognition and machine learning*, vol. 14. Academic press New York, 1968.
19. He, W., Liu, X., and Ren, M. Location cheating: A security challenge to location-based social network services. In *Proc. ICDCS 2011*, IEEE (2011), 740–749.
20. Jaimes, L. G., Vergara-Laurens, I., and Labrador, M. A. A location-based incentive mechanism for participatory sensing systems with budget constraints. In *Proc. PerCom 2012*, IEEE (2012), 103–108.
21. Kawajiri, R., Shimosaka, M., Fukui, R., and Sato, T. Frustratingly simplified deployment in WLAN localization by learning from route annotation. In *Proc. ACML 2012*, JMLR (2012).
22. Koutsopoulos, I. Optimal incentive-driven design of participatory sensing systems. In *Proc. INFOCOM 2013*, IEEE (2013), 1402–1410.
23. Lane, N. D., Xu, Y., Lu, H., Hu, S., Choudhury, T., Campbell, A. T., and Zhao, F. Enabling large-scale human activity inference on smartphones using community similarity networks (csn). In *Proc. UbiComp 2011*, ACM (2011), 355–364.
24. Lomasky, R., Brodley, C. E., Aernecke, M., Walt, D., and Friedl, M. Active class selection. In *Proc. ECML 2007*. Springer, 2007, 640–647.
25. Meschtscherjakov, A., Reitberger, W., Lankes, M., and Tscheligi, M. Enhanced shopping: a dynamic map in a retail store. In *Proc. UbiComp 2008*, ACM (2008), 336–339.
26. Ouyang, R. W., Srivastava, A., Prabahar, P., Roy Choudhury, R., Addicott, M., and McClernon, F. J. If you see something, swipe towards it: Crowdsourced event localization using smartphones. In *Proc. Ubicomp 2013*, ACM (2013), 23–32.
27. Reddy, S., Estrin, D., and Srivastava, M. Recruitment framework for participatory sensing data collections. In *Proc. Pervasive 2010*. Springer, 2010, 138–155.
28. Rula, J., and Bustamante, F. E. Crowd (soft) control: moving beyond the opportunistic. In *Proc. HotMobile 2012*, ACM (2012), 3.
29. Settles, B. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.
30. Tuite, K., Snavely, N., Hsiao, D.-y., Tabing, N., and Popovic, Z. Photocity: training experts at large-scale image acquisition through a competitive game. In *Proc. CHI 2011*, ACM (2011), 1383–1392.
31. Yang, D., Xue, G., Fang, X., and Tang, J. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In *Proc. MobiCom 2012*, ACM (2012), 173–184.
32. Zichermann, G., and Cunningham, C. *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media, Inc., 2011.