Journal of Urban Technology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjut20

Managed Participatory Sensing with YouSense

Mattias Linnap & Andrew Rice
Published online: 09 Jun 2014.

To cite this article: Mattias Linnap & Andrew Rice (2014) Managed Participatory Sensing with YouSense, Journal of Urban Technology, 21:2, 9-26, DOI: 10.1080/10630732.2014.888216

To link to this article: http://dx.doi.org/10.1080/10630732.2014.888216

Journal of Urban Technology, 2014
Vol. 21, No. 2, 9–26, http://dx.doi.org/10.1080/10630732.2014.888216

Routledge
Taylor & Francis Group

Managed Participatory Sensing with YouSense

Mattias Linnap and Andrew Rice

ABSTRACT *Participatory sensing uses humans as intelligent sensors by asking them questions on mobile devices. It has been used to monitor effects of climate on plants, potholes in streets, and people's happiness in urban environments. These phenomena are difficult and expensive to measure with traditional electronic sensors. Previous participatory sensing projects have relied on the enthusiasm of the volunteers to notice and report events of interest. YouSense is a framework for centralized real-time management of the volunteers' efforts. It uses location tracking and model-based methods to target questions to participants only when they are in the optimal situation to contribute data.*

We have conducted two outdoor case studies to evaluate YouSense. Although it was effective at reducing participant effort wasted on unnecessary answers by 4× in short, focused studies, it was ineffective at targeting questions to valuable locations in long-term studies. Simulation results show that these outcomes are due to different participant behavior, and that centralized management is most effective at high numbers of questions.

KEYWORDS *participatory sensing; community sensing; volunteered geography; model-based sensing*

Introduction

The last decade has seen the rise of human computation, where volunteers collaborate to solve data creation and processing tasks (Quinn and Bederson, 2011). In volunteered geography, individuals create, edit, and analyze geographical datasets (Goodchild, 2007). A successful example is the OpenStreetMap project, which has created a road map of comparable quality to official sources in the United Kingdom (Haklay, 2010). Participatory sensing is an approach to volunteered geography that uses human senses via mobile phones for noticing events and answering questions about the physical world (Burke et al., 2006). It has been used for a variety of applications in biology, environmental and social science, and urban community campaigns. These phenomena are difficult and expensive to measure with traditional electronic sensors.

Previous participatory sensing campaigns have relied on the enthusiasm of the participants to notice events and report observations. We have built YouSense, a platform for real-time management of the participation process. It uses model-based sensing (Deshpande et al., 2004) to push questions to the participants when they are in the optimal situation to contribute new data, while avoiding

Correspondence Address: Mattias Linnap, Computer Laboratory, 15 JJ. Thomson Avenue, Cambridge CB3 0FD, United Kingdom. E-mail: mattias.linnap@cl.cam.ac.uk

annoying them with unnecessary requests. Targeting questions requires knowledge of participants' location and synchronization of their contributions to a central server. Our mobile client implementation uses sensor scheduling for improving energy-efficient location tracking.

We have evaluated YouSense in two outdoor case studies, one involving 24 participants for an hour, and one involving eight participants for nine weeks. In both cases, centralized management was beneficial, but with significantly different outcomes. In the short-term study, the questions were frequent enough to achieve a similar coverage area with all management methods, and management prevented effort wasted on unnecessary answers. In the long-term study, the benefit of management was in reminding participants to contribute answers, at the cost of increasing wasted effort. We show that these outcomes are due to different participant behavior in the studies and discuss the impact on the design of future participatory sensing studies.

The main contributions of this paper are:

- the design and architecture for a model-based management system for participatory sensing
- evaluation of centralized management in two outdoor case studies, showing 4× reduction in wasted participant effort in the short-term study, and 7× improvement in coverage area in the long-term study
- evaluation of participant behavior and willingness to answer questions in two studies, showing that centralized management may be ineffective for cyclist participants
- evaluation of the effects of participant behavior and study design parameters in simulation, showing that centralized management is most effective at high number of questions, and weekly synchronization to the server is sufficient
- evaluation of the energy efficiency of the YouSense platform on a nine-month location tracking study with 72 participants, showing that it uses 3.5–5.4 percent of battery per hour, and is practical on personal smartphones.

Related Work

Human computation and crowdsourcing are diverse fields of research, that involve solving problems with the help of members of the public. Public participants may be paid or volunteer. They help in areas that are not yet practical to attempt to automate such as object recognition from images, language understanding, translation, and various forms of data creation (Quinn and Bederson, 2011).

In volunteered geography (Goodchild, 2007), these members of the public create and edit geographical datasets. A successful example is the OpenStreetMap project, which provides an open-license road map of the world (Haklay and Weber, 2008). Volunteers create maps by vectorizing satellite images, cleaning up user-contributed GPS traces, and adding metadata about names, buildings, and geographical features. The quality of OpenStreetMap is comparable to the official Ordnance Survey road maps in the United Kingdom (Haklay, 2010). Volunteered geography has a long history: for example, the Christmas Bird Count has been running since 1900 to create a yearly census of bird species in the western hemisphere (Butcher et al., 1990).

Participatory sensing is an approach to volunteered geography that uses smartphones as the interface to human senses and decision-making (Burke et al., 2006). It is distinguished from other approaches to volunteered geography by the presence of the volunteer at the location of observations. For example, while roads can be mapped from satellite imagery or GPS traces, many phenomena of interest such as plant and animal life, or the condition of urban streets are not visible from existing electronic sensor datasets. Participatory sensing could be used to create new communication channels between governments and citizens in Smart Cities (Allwinkle and Cruickshank, 2011; Caragliu et al., 2011). The government can set up sensing campaigns to gather feedback about transport links, or citizens can set up a community-run platform to report problems to governments. One successful example of a community platform is FixMyStreet.com, which forwards location-tagged reports about graffiti and potholes to the appropriate local councils in the United Kingdom.

Participatory sensing systems may include other technology and electronic sensors. Sensors on the mobile phone, or external data sources such as weather reports or traffic cameras can be used to provide additional situational context to the reported observations. Participatory sensing is distinguished from people-centric sensing (Eisenman et al., 2006) by the participants being actively involved in the process and reporting their findings, opinions, or decisions, beyond carrying mobile sensors around.

Participatory Sensing Applications

Participatory sensing has been used for biology, environmental and social science, urban activism, and community-run campaigns (See Table 1). Despite the varied application domains, all these projects share a common architecture: a mobile client application for participants to report their observations, and a central server to collect, analyze, and distribute the results of the participatory sensing campaign to organizers. The custom mobile clients offer similar user interfaces and require similar electronic sensor features for tagging the observations with situational context.

Table 1 shows that the shared requirements of existing participatory sensing projects are:

- a user interface to enter observations with numerical, text, and multiple-selection tags
- situational context tagging with location, photos, and audio
- automated submission of the sampling reports to a central server for further analysis by the campaign creators

We believe that in addition to previous projects' features, a participatory sensing platform should include real-time participation management. Prior work has shown that participants get bored and stop contributing after a few weeks of initial enthusiasm (Reddy and Estrin, 2010). Without guidance on the most useful observation locations, the collected data may be clustered together both in time and space, whereas other areas remain unobserved. To manage the participatory sensing process, the platform should include:

- notifications that remind participants to report observations over long periods
- location-based targeting of the notifications to improve the spatial distribution of the observations.

Table 1: Application areas and data contribution methods of past participatory sensing projects.

Project	Application area	Input interfaces	Additional sensors	Further notes
Let's Do It World[a]	Environment	Multiple-choice	Camera	Rubbish mapping and collaborative clean-up campaigns with over 2.5 million volunteers from 19 countries.
GarbageWatch (Reddy and Estrin, 2010)	Environment	Multiple-choice	Camera	Mapping recyclable materials in outdoor general waste bins. The system includes participant pre- selection based on historical movement patterns and response rates, but no guidance to participants while the campaign is in progress.
BudBurst (Meymaris et al., 2008)	Biology	Multiple-choice	Camera	Mapping the stage of leafing, flowering and fruiting of plants.
Christmas Bird Count (Butcher et al., 1990)	Biology	–	–	The project does not currently use a mobile client application.
CycleStreets[b]	Urban Activism	Text	Camera	Mapping problems with cycle lanes and parking racks. The data is contributed to OpenStreetMap.
FixMyStreet[c]	Urban Activism	Text	Camera	Mapping graffiti, rubbish, broken pavement and street lights, and other urban problems. The data is forwarded to local governments.
Mappiness (MacKerron, 2012)	Social Science	Numerical, Multiple-choice	–	Investigation of happiness of people in various urban environments. The mobile client reminds participants to contribute data with a fixed number of randomly timed notifications per day.
NoiseTube (Maisonneuve et al., 2009)	Social Science	Multiple-choice	Microphone	Mapping noise levels and their subjective annoyance in urban environments.
Transafe (Salim, 2012)	Urban Safety	Text, Numerical	Camera, Microphone	Mapping perceived safety and emotions in urban environments.

[a]http://www.letsdoitworld.org/statistics
[b]http://www.cyclestreets.net
[c]http://www.fixmystreet.com

YouSense

YouSense is a real-time, model-based platform we have created to manage participatory sensing campaigns. Unlike prior work in participatory sensing, our aim is to increase the amount and quality of the contributed data with active guidance and questions to the participants while the campaigns are in progress. YouSense uses model-based sensing (Deshpande et al., 2004) to notify the participants of observation opportunities when they are in the optimal situation to contribute new data, while avoiding annoying them with unnecessary requests.

System Architecture

The high-level YouSense system architecture is composed of a mobile participatory sensing client, and a central server for campaign management (See Figure 1). An organizer interacts with the campaign management server. He starts a campaign and provides its metadata: geographical areas of interest, required electronic sensor context for reported observations, and the data entry interface presented to the participants. As the campaign proceeds, the researcher can export the data contributed by participants. The campaign management server collects all participants' locations, situational sensor context, and previous observations, and computes the optimal locations for new observation requests for each participant.

Model-Based Question Targeting

YouSense prompts participants with questions to collect observations of their surrounding environment and opinions. The management system has to trade off increasing the value of the dataset with more questions, and limiting the effort demanded from participants with fewer questions. YouSense uses model-based sensing to make decisions about when and whom to prompt with questions. Model-based sensing (Deshpande et al., 2004) has successfully reduced network use and energy requirements of electronic sensor networks of weather stations (Osborne et al., 2008), targeted electronic sensors in mobile phones for traffic-speed measurement (Krause et al., 2008) and been used for environmental sensing (Weinschrott et al., 2010).

This approach requires the survey organizers to define three mathematical models to describe their prior beliefs, expectations, and goals:

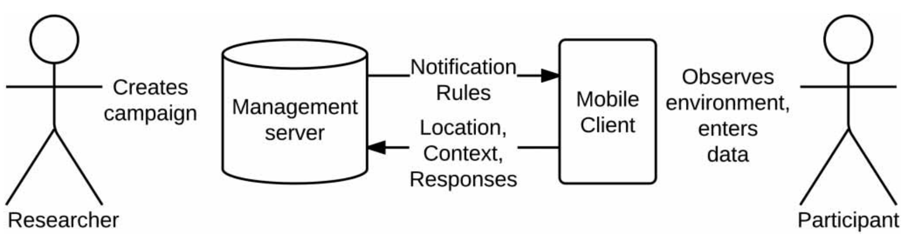

Figure 1: Overview of the YouSense system architecture. The management server collects electronic sensor readings from participants' mobile devices, and notifies them of observation opportunities.

- a spatial world model that captures the current knowledge about the phenomena of interest. In participatory sensing, it is based on answers contributed by the users, but it can also incorporate other external datasets.
- a spatial goal model that describes the value of additional contributions from any location. For example, it can be set to collect K responses from each location, L responses from each participant, or M responses that agree.
- a cost model that describes the maximum effort demanded from each participant. For example, it could be set to allow questions at regular intervals or a maximum number of questions per day. In studies with monetary incentives, the cost model can be based on the available total budget.

The YouSense platform uses the world and value models to decide if the contribution from a participant's current location would be valuable to the organizers, and the cost model to decide if the participant can be currently interrupted with a question (See Figure 2).

In the two evaluation studies presented below, we used a grid-based world model where all answers from the same grid cell are considered to be from the same location. The goal model was set to collect exactly one answer from each cell, and the cost model to allow one question prompt per two minutes or one hour (See Figure 3).

Other studies should use models that best match their organizers' expectations about the phenomena of interest and the value of contributions from different locations and participants. However, we believe that most participatory sensing studies can be run with a small set of pre-designed models, requiring only the choice of configuration parameters (for example, the grid size and question frequency) to the organizers.

Management Approaches

The YouSense platform uses centralized management: question targeting decisions are made based on all participants' locations and contributions. To evaluate the benefits of this approach, we have implemented four approaches to management:

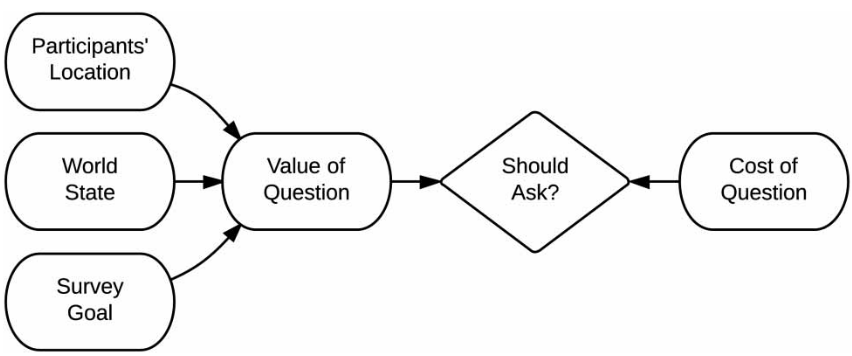

Figure 2: The YouSense model-based notification targeting architecture. The system decides whether to notify participants with a question based on the current expected value of the question, and the cost of annoying a participant.

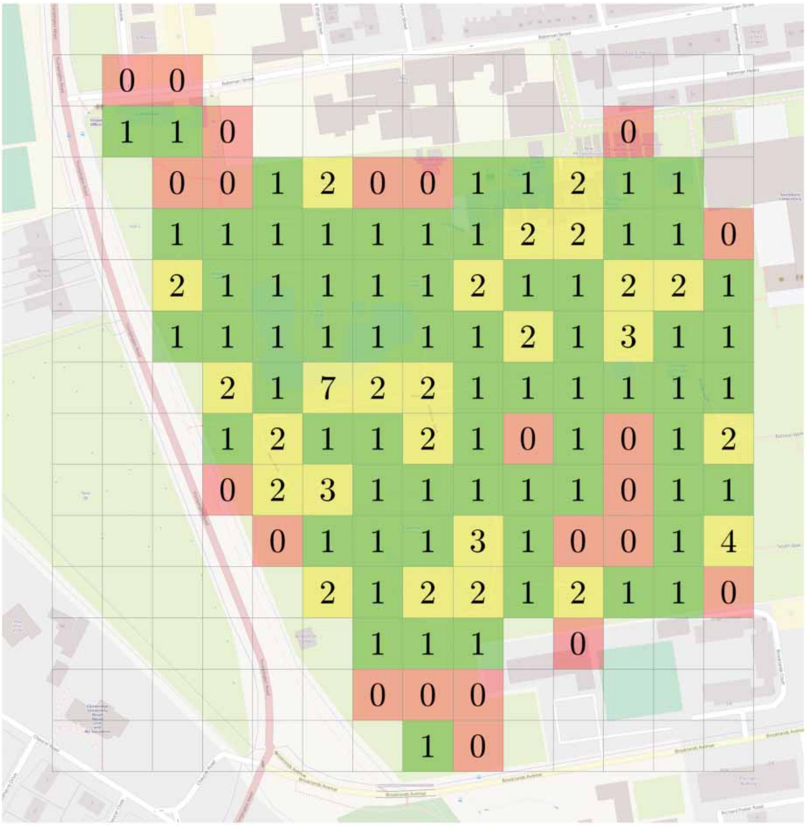

Figure 3: The world and value models at the end of the Botanical Gardens case study. The world model is represented as a spatial grid. The survey goal was set to collect one answer from each grid cell. Green cells are sampled precisely according to the goal model, yellow cells are over-sampled, and red cells are under-sampled. Participants did not visit the empty cells, therefore these cannot be sampled with any management approach.

- manual, where participants can contribute observations when they notice events of interest, but are not prompted with questions. This has been used by most past projects.
- timed management, where participants are prompted with questions at regular intervals regardless of their location. This has been used in a small number of past projects, for example Mappiness (MacKerron, 2012).
- individual management, where participants are prompted with questions based on their own locations and previous answers. This approach can be useful for studies in areas without network connectivity.
- centralized management, where questions are based on all participants' locations and previous answers. The world and goal models are synchronized between participants.

Implementation

We have built a prototype implementation of the YouSense platform. The mobile client is available for Android smartphones, and the campaign management server can be installed on any commodity server with Python and PostgreSQL.

Mobile Client Interface

The mobile client user interface is designed for long-term campaigns. The most common interaction pattern is via the targeted notifications that ask the user to report a participatory sensing observation (See Figure 4[c]), followed by data entry (See Figure 4[b]). Enthusiastic participants have the option of browsing all running campaigns and contributing data manually (See Figure 4[a]). This interaction pattern is common in most existing participatory sensing projects and does not require persistent background location tracking.

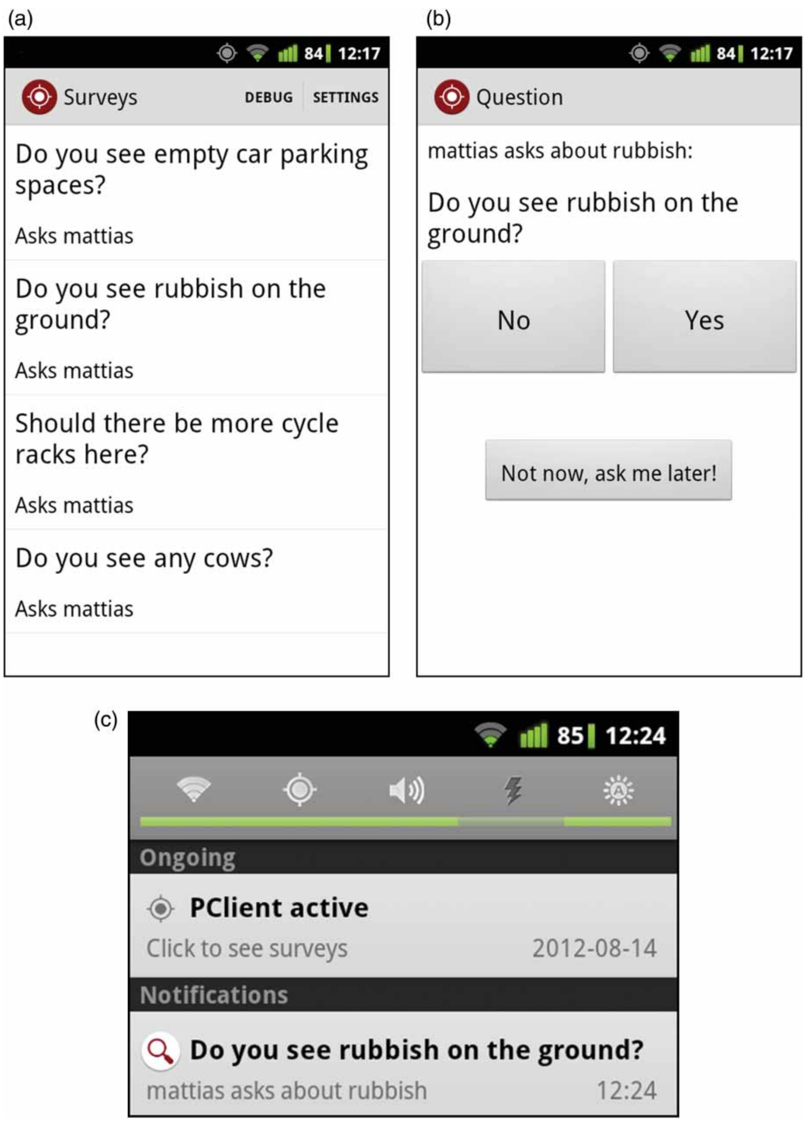

Figure 4: Android mobile client user interface. (a) List of active participatory sensing campaigns. The user can contribute observations without being prompted with a notification. (b) Data entry interface for campaigns with a Boolean answer. The user has the option to provide annoyance metadata to improve notification targeting and (c) Notifications requesting the user to report a participatory sensing observation.

Feedback from preliminary testing indicated that participants are concerned about battery life and privacy while running YouSense. We addressed these concerns with user-configurable tracking behavior: the users have an option to pause YouSense background location tracking manually, and to pause it automatically if the battery level falls below 30 percent.

Energy-Efficient Question Targeting for Participatory Sensing

Targeting questions according to the spatial world and goal models requires knowledge of the participants' current location. Modern mobile phone operating systems provide positioning using cell towers, WiFi access points, and GPS. YouSense uses GPS, as cell tower positioning is too inaccurate, and WiFi positioning is unavailable in areas of interest for many participatory sensing projects, such as bike paths and parks. However, tracking the location at all times with GPS and sending it to the server in real-time is infeasible, as it would deplete the battery life of a typical smartphone in 4–6 hours.

However, tracking participants' locations at all times with GPS is unnecessary. People spend most of their time indoors, and move outdoors infrequently. Based on our experimental data, mobile phones detect movement with accelerometers in 10.3 percent of the time, and GPS is available 5.7 percent of the time. By using lower-power accelerometers and WiFi signals for movement detection, YouSense is able to reduce high-power GPS usage to periods where the user is likely to be moving outdoors.

Centralized management requires regular synchronization of world and goal models between the participants' mobile phones and the management server. Mobile data network use is energy-intensive and may be expensive for participants depending on their data plan. Time-sensitive events, such as location updates and observation reports, are uploaded once per hour. Time-insensitive data, such as debug logs and additional sensor data for system evaluation are uploaded only when the phone is connected to a WiFi network. Simulation results below show that this was unnecessarily frequent—even weekly synchronization is sufficient for long-term studies.

Experimental Results

We evaluated the YouSense platform on two outdoor case studies. The Botanical Gardens case study tested three methods of management in a controlled, one-hour study with 24 participants (Linnap and Rice, 2014), and showed that centralized management can decrease wasted effort by $4\times$, but increase coverage area only by 8.6 percent. In this paper, we compare these results to a nine-week study with eight participants, where management was effective at increasing the coverage area, but not in decreasing the wasted effort. We explain the different outcomes based on participant behavior during the studies.

We show that the YouSense location tracker doubles energy use of the mobile phones in the background state (screen off), but remains efficient enough for use on personal phones on a eight-month study with 72 participants. It was impractical to conduct a controlled outdoor experiment for every combination of management models and configuration parameters. The sensitivity of the results to chosen parameters and participant behavior was evaluated in simulation, using the location traces collected during the outdoor studies. The simulation results show that the main factor influencing the effectiveness of centralized management is participant behavior, and that additional questions have a diminishing return on the coverage area.

Effectiveness of Centralized Management

Evaluation Studies. The effectiveness of centralized management was evaluated in the Botanical Gardens and Cambridge case studies. In the Botanical Gardens case study, 24 participants walked the gardens for 43 to 73 minutes, and were prompted with the question "Do you see any bees?" The participants were split into three groups for testing three management approaches:

- timed management, where participants are prompted with a question every two minutes
- individual management, where participants are prompted when entering a location with no previous answers from themselves
- centralized management, where participants are prompted when entering a location with no previous answers from any participant in the group.

In the Cambridge case study, eight participants ran YouSense on their personal mobile phones for nine weeks while living in Cambridge. They were prompted with a number of different questions motivated by previous participatory sensing projects, including "Should there be more cycle racks here?" "Do you see any empty parking spaces?" and "Do you see any rubbish on the ground?" The study tested centralized management and manual, unprompted contributions in a long-term casual participation scenario.

Participant Selection. The 24 participants of the Botanical Gardens study were recruited by e-mail. Seventy-nine percent of them were members of the Computer Laboratory, and 75 percent of them were male. Fifty-eight percent of the participants were provided with an Android smartphone suitable for the study, the rest used their personal phones. The participants were rewarded with free food for taking part, but were offered no performance-based incentives.

The participants were split into three equal groups of eight for each of the three management approaches. They were unaware of their group, and how or when the system prompts them with questions. As the behavior of the individuals influences the results, in the ideal case, members from each management approach would walk in groups of three, and respond to questions with identical delays. As a practical approximation, they were asked to walk naturally in groups of no more than three people, and participants who were expected to walk together were assigned to different management approaches.

The eight participants of the long-term Cambridge study were members of the Computer Laboratory. The participants used their personal Android smartphones, with YouSense running in the background. All participants commuted in the city by bicycle.

Management Models. As described in the previous section, the YouSense framework uses a number of models to target questions to locations considered valuable to the researchers. In both evaluation studies, the world model was represented as a grid over space. Answers from the same grid cell were considered to be from the same location. The goal model was set to one answer from each grid cell. Cells with no answers represent lack of coverage, and multiple answers from a cell are considered wasted effort. The cost model was set to unlimited number of questions, but with a minimum time gap between question prompts. The quantitative configuration parameters are described in Table 2. Other studies may use different world model representations, different value models, and different cost models based on the study organizers' prior beliefs and expectations.

Coverage Area and Wasted Effort. We evaluated the success of participatory sensing studies based on two quantitative metrics: coverage area and wasted effort.

- coverage area: the fraction of cells with answers, out of visited cells in the study area. Cells never visited by any participant are excluded, as they cannot be sampled with any participatory sensing method.
- wasted effort: the fraction of answers which were unnecessary to satisfy the goal model. This includes answers contributed from outside the study area, and for the chosen evaluation models, second and subsequent answers from a cell.

Table 3 summarizes the results of the two studies. As the Cambridge case study involved a number of different questions, the table presents the results of the one with the most answers: "Do you see any rubbish on the ground?"

In the Botanical Gardens study, all participant groups were prompted with questions. The centralized management approach achieves 8.6 percent larger coverage area than time-based questions, and reduces participant effort wasted on answers that are not valued according to the organizers' goal model by four times. We believe the modest increase in the coverage area is due to frequent questions and a small study area: as participants moved within a small closed garden and were asked questions up to every two minutes, all methods received enough answers to cover most of the area. The main benefit of individual and centralized management is in preventing question prompts when the answer would not be valuable: the number of *useful* answers is similar for all three methods.

In the Cambridge case study, the same group of participants were prompted with questions using the centralized management approach, but also had the opportunity to contribute data manually. The managed approach achieves a seven times larger coverage area than manual contributions, but a larger proportion of them are wasted effort. As the total number of manual unprompted

Table 2: Participant set and configuration parameters for the two studies.

	Botanical Gardens	Cambridge
Participants	24	8
Main travel method	Walking	Cycling
Study length	1 hour	9 weeks
Study area	400 × 400 meters	City of Cambridge
Grid size	28 × 28 meters	210 × 220 meters
Coverage goal	1 answer from cell	1 answer from cell
Minimum question gap	2 minutes	1 hour
Synchronization interval	1 minute	1 hour

Table 3: Coverage area and wasted effort metrics for the two case studies. Coverage area is the fraction of visited goal cells with answers. Wasted effort are answers that are not useful according to the goal model, or all but the first answer from a cell in these studies. Note that the over-sampled area is a subset of the covered area.

	Botanical Gardens			Cambridge	
	Timed	Individual	Centralized	Centralized	Manual
Visited cells	104	97	114	179	179
Covered (1+ answers)	75 (72%)	74 (76%)	92 (81%)	66 (37%)	9 (5.0%)
Missed (0 answers)	29 (28%)	23 (24%)	22 (19%)	113 (63%)	170 (95%)
Over-sampled (2+ answers)	59 (57%)	51 (53%)	25 (22%)	17 (9.5%)	3 (1.2%)
Total answers	224	188	129	199	19
Useful answers	75 (33%)	74 (39%)	92 (71%)	66 (33%)	9 (47%)
Wasted effort	149 (67%)	114 (61%)	37 (29%)	133 (67%)	10 (53%)

contributions is low, we believe the main benefit of management in this case was reminding participants to contribute during the nine-week study.

The two studies had considerable differences in scale, in participant travel methods, and in outcomes. In the next section, we analyze participant behavior in the two studies, and in the section following that, we show that this is a significant factor in the differing outcomes.

Participant Behavior

The YouSense centralized management targets *questions* to locations where *answers* would be valuable according to the goal model, but it cannot compel participants to actually answer from there. If participants move between being prompted with a question and contributing an answer, then management will be less effective at covering all goal cells exactly once.

Table 4 and Figure 5 show that the distribution of time taken to answer and the distance moved in the two evaluation case studies were significantly different. In the Botanical Gardens, the median time taken to answer was 14.6 seconds, and in Cambridge, 14.2 minutes. These result in median distance moved between YouSense prompting a participant with a question and receiving an answer of five meters in the gardens (well within the 28-meter grid cell size), but 579 meters in Cambridge (more than double the 210-meter grid cell size). As a result, centralized

Table 4: Time taken to answer by the participants in the two studies, and the resulting movement distance between the system prompting the participant with a question and receiving an answer. Note the difference in units.

	Botanical Gardens	Cambridge
Mean question to answer time	20.1 seconds	76.3 minutes
Median	14.6 seconds	14.2 minutes
95th percentile	45.1 seconds	325.3 minutes
Mean question to answer distance	9.0 meters	1091 meters
Median	5.0 meters	579 meters
95th percentile	26.6 meters	3330 meters

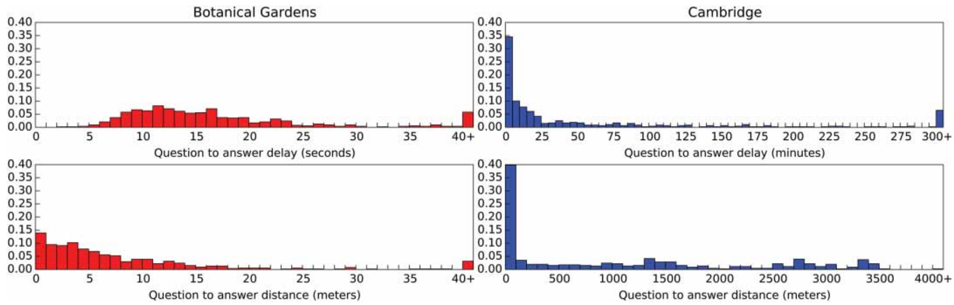

Figure 5: Histogram of time taken to answer by the participants in the two studies, and the resulting movement distance between the system prompting the participant with a question and receiving an answer. Note that Cambridge time delays are in minutes. The final bin contains all values larger than the upper bound of the histogram.

management is unable to target contribution of answers to grid cells where new data are useful.

Based on post-study interviews with the participants, these results are due to different enthusiasm and travel methods. In the Botanical Gardens, the participants were on foot, and remained motivated throughout the one-hour study. Some participants in the timed management group held the phone in their hand for the entire study, as they received questions every two minutes. In the Cambridge case study, the participants tended to visit new locations while cycling, and were reluctant to take the phone out of their pocket to answer quickly. As a result, they often delayed answering until reaching their destination, and the answers were clustered around home and work.

Energy Efficiency of the System

The energy-efficiency of the YouSense framework was evaluated on a separate eight-month location-tracking study of 72 participants. It did not involve participatory sensing, and collected only electronic sensor data from mobile phones. However, the location tracking and server synchronization components were operated identically to the nine-week Cambridge study.

The 72 participants of the location tracking study were recruited from Tartu, Estonia, using snowball sampling. Sixty-five percent of them were female, and 35 percent male, all with at least a Bachelor's degree. They were provided with Android smartphones for the duration of the study, which they used as their primary personal mobile device.

Energy use of YouSense was measured by logging the charger connection events and battery level changes on the phone at 1 percent granularity. We believe that energy use is influenced by four main factors: design of the location tracker, the hardware model of the mobile device, movement patterns of the participant, and use of other apps by the participant. To reduce the influence of other apps, we present energy use while the screen is off—and most other apps should not be active. To reduce the influence of personal movement patterns, the results are averaged over all participants using the same device. Twelve participants were excluded from the location tracker energy efficiency analysis, as they were using uncommon device models.

Table 5: Mean and standard deviation battery level drop per hour with charger disconnected. All time and the DeviceAnalyzer baseline count both screen on and off time, with charger disconnected.

		Screen Off		All time		
Phone Model	Number of Participants	YS Off	YS On	YS Off	YS On	DeviceAnalyzer Baseline
Galaxy S4	4	3.3 ± 0.0	4.9 ± 0.3	4.5 ± 0.0	5.4 ± 0.3	4.1 ± 1.5
LG E610	44	1.3 ± 0.5	3.3 ± 0.9	1.5 ± 0.7	3.5 ± 1.0	6.0 ± 3.2
LG P700	7	6.2 ± 5.2	4.7 ± 1.2	5.7 ± 4.6	5.1 ± 1.3	5.4 ± 3.6
LG P880	5	1.4 ± 0.7	4.1 ± 1.0	2.3 ± 1.0	4.8 ± 1.2	7.7 ± 3.6

Table 5 shows the mean ± standard deviation battery level drop per hour with the charger disconnected. The mean and standard deviation are computed over the distinct participants' mean values, with participants with under one hour of data for the particular activity state excluded. For reference, the rightmost column includes mean energy use observed over all time (with charger disconnected) on the same device models in the DeviceAnalyzer dataset of 12500 Android devices (Wagner et al., 2013).

YouSense doubles energy use of the phone in background tracking mode while the screen is off. However, the resulting battery level drop including other daily phone use is the range of 2.5 to 6.4 percent per hour, therefore YouSense can be used continuously on primary personal phones if they are charged overnight. Completely controlled energy measurements are not possible on participants' personal phones due to other activity, but could be achieved with a controlled movement pattern, mobile phones with all other background applications disabled, and the PowerBenchmark hardware kit (Rice and Hay, 2010).

Effects of Participant Behavior in Simulation

It is infeasible to conduct a controlled outdoor experiment for every combination of management models and configuration parameters. Instead, we investigated the effects of participant behavior and the chosen parameters in simulation. The simulator uses the same management algorithms as the outdoor experiments. In addition to the survey configuration parameters, it needs two additional input datasets for the human factors in participatory sensing: a movement trace of the participants and a model for time taken to answer in different situations.

For the movement traces, we used the largest available dataset: YouSense GPS logs of 72 participants of the eight-month Tartu location-tracking study. To simulate different participant behavior, we used three models:

- immediate, where participants respond without delay
- cyclist, where participants respond as soon as their movement speed drops below 4 m/s for one minute
- delayed, where participants respond with delays picked randomly from the same distribution as observed in the Cambridge case study.

Figures 6 and 7 show the coverage area and wasted effort arising from varying three parameters: the participant answer delays, the maximum number

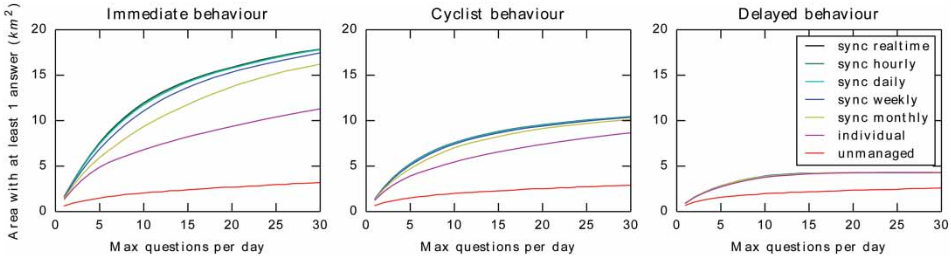

Figure 6: Coverage area resulting from three behavior models, maximum number of questions and management synchronization interval in simulation. Management is most effective with frequent questions and quick responses.

of questions per day, and the frequency of world model synchronization in the centralized management.

The results show that centralized management is most effective for diligent participants and a high number of questions per day. With the answer delays observed in the Cambridge case study, management has limited benefits for increasing area coverage or reducing wasted effort. However, the delays cannot be entirely attributed to cycling: participants must have also postponed answers at other times.

The simulation also evaluates the choice of synchronizing the world model once per hour in the Cambridge study. The results show that even weekly synchronizations would have been sufficient. To minimize the impact on battery life, we recommend future studies to upload answers and download new world models when the mobile phone is plugged into a charger.

Discussion and Future Work

We have conducted two outdoor case studies of YouSense: one to evaluate three management approaches on 24 highly enthusiastic participants for one hour, and one to evaluate centralized management against manually contributed answers with eight casual participants over nine weeks. In both cases, centralized management was the most effective approach, but with significantly different results. In the short-term study, questions were frequent enough to achieve similar coverage with all management methods. Centralized management was

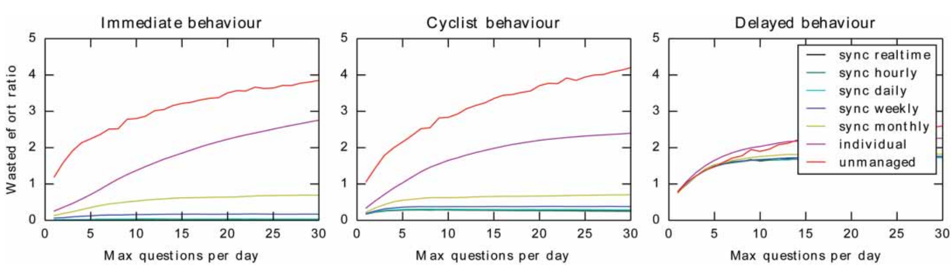

Figure 7: Wasted effort ratio (wasted effort to useful effort) resulting from three behavior models, maximum number of questions, and management synchronization interval in simulation. Management is most effective with frequent questions and quick responses.

effective at reducing wasted effort on answers not valued by the goal model by 4 × by inhibiting questions when the participants were in an already-sampled location.

In the long-term study, participants were reluctant to answer while cycling, and moved a median of 579 meters between being prompted with a question and responding. As a result, centralized management was ineffective at targeting answers to locations valued by the goal model. However, management increased the coverage area by reminding participants to contribute at all.

Analysis of different management approaches in simulation confirms that participant answer delays are an important factor in deciding whether to use centralized management in a participatory sensing study. Further studies of participant behavior are needed to determine in which conditions and travel modes quick responses can be expected.

We believe that the effectiveness of managed participatory sensing depends on a number of additional factors, including the overlaps between participants' movement, the chosen world and value models for the study, and perhaps further human factors. Further work is required to evaluate their effects, and compile guidelines for organizers of studies. If centralized management is used, the synchronization of world models between participants can be as infrequent as once per week. We recommend synchronizing when the phone is charging to minimize the impact on battery life.

Conclusion

YouSense is a platform for model-based management of participatory sensing campaigns. Unlike past projects, it actively notifies participants with questions when they are in the optimal location to contribute new data. The results from two evaluation case studies and simulations show that centralized management is most effective with frequent questions and quickly-responding participants.

Participants may make mistakes in their reports. In Cambridge, queue lengths of car traffic at red lights are available via the SCOOT system embedded at intersections. We plan to evaluate the quality and timeliness of participatory sensing by comparing the human-contributed reports to the ground truth from electronic sensors, and by cross-participant validation in campaigns where no ground truth is available.

Although the current YouSense prototype is energy-efficient enough to be used on personal smartphones if the battery is charged overnight, there are further possibilities for energy use optimization. Trajectory tracking instead of location tracking can be used to detect visits to areas of interest with less energy use (Kjærgaard et al., 2011), and vibration circuits in future mobile phones may lead to more efficient motion detection (Guha et al., 2010).

Acknowledgments

The authors thank Ripduman Sohan and Daniel Wagner for their help in reviewing drafts of the paper.

Notes on Contributors

Mattias Linnap is a PhD student in the Computer Laboratory at the University of Cambridge. His research focuses on energy-efficient location tracking on smartphones, and using location-based models to improve Participatory Sensing.

Andrew Rice is a lecturer in the Computer Laboratory at the University of Cambridge. His research focus is in the use of smart phones as a computing platform, and currently considers issues of understanding modern phone use and associated energy consumption.

Bibliography

S. Allwinkle and P. Cruickshank, "Creating Smart-er Cities: An Overview," *Journal of Urban Technology* 18: 2 (2011) 1–16.

J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. Srivastava, "Participatory Sensing," paper presented at ACM World Sensor Web Workshop (Boulder, 2006).

G.S. Butcher, M.R. Fuller, L.S. McAllister, and P.H. Geissler, "An Evaluation of the Christmas Bird Count for Monitoring Population Trends of Selected Species," *Wildlife Society Bulletin* 18: 2 (1990) 129–134.

A. Caragliu, C. Del Bo, and P. Nijkamp, "Smart Cities in Europe," *Journal of Urban Technology* 18: 2 (2011) 65–82.

A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong, "Model-driven data acquisition in sensor networks," paper presented at 30th International Conference on Very Large Data Bases (Toronto, 2004).

S. Eisenman, N. Lane, E. Miluzzo, R. Peterson, G. Ahn, and A. Campbell, "Metrosense project: People-centric sensing at scale," paper presented at ACM World Sensor Web Workshop (Boulder, 2006).

M. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal* 69: 4 (2007) 211–221.

S. Guha, K. Plarre, D. Lissner, S. Mitra, B. Krishna, P. Dutta, and S. Kumar, "AutoWitness: Locating and Tracking Stolen Property While Tolerating GPS and Radio Outages," paper presented at 8th ACM Conference on Embedded Networked Sensor Systems (Zurich, 2010).

M. Haklay, "How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets," *Environment and Planning B: Planning and Design* 37: 4 (2010) 682–703.

M. Haklay and P. Weber, "OpenStreetMap: User-Generated Street Maps," *IEEE Pervasive Computing* 7: 4 (2008) 12–18.

M.B. Kjærgaard, S. Bhattacharya, H. Blunck, and P. Nurmi, "Energy-efficient Trajectory Tracking for Mobile Devices," paper presented at 9th ACM International Conference on Mobile Systems, Applications and Services (Washington, 2011).

A. Krause, E. Horvitz, A. Kansal, and F. Zhao, "Toward Community Sensing," paper presented at International Conference on Information Processing in Sensor Networks (St. Louis, 2008).

M. Linnap and A. Rice, "The Effectiveness of Centralized Management for Reducing Wasted Effort in Participatory Sensing," paper presented at 1st International Workshop on Crowdsensing Methods, Techniques and Applications (Budapest, 2014).

G. MacKerron, "Happiness economics from 35000 feet," *Journal of Economic Surveys* 26: 4 (2012) 705–735.

N. Maisonneuve, M. Stevens, M.E. Niessen, and L. Steels, "NoiseTube: Measuring and Mapping Noise Pollution with Mobile Phones," paper presented at 4th International Symposium on Information Technologies in Environmental Engineering (Thessaloniki, 2009).

K. Meymaris, S. Henderson, P. Alaback, and K. Havens, "Project BudBurst: Citizen Science for All Seasons," paper presented at American Geophysical Union Fall Meeting (San Francisco, 2007).

M.A. Osborne, S.J. Roberts, A. Rogers, S.D. Ramchurn, and N.R. Jennings, "Towards Real-Time Information Processing of Sensor Network Data Using Computationally Efficient Multi-output Gaussian Processes," paper presented at International Conference on Information Processing in Sensor Networks (St. Louis, 2008).

A.J. Quinn and B.B. Bederson, "Human Computation," paper presented at ACM CHI Conference on Human Factors in Computing Systems (Vancouver, 2011).

S. Reddy and D. Estrin, "Recruitment Framework for Participatory Sensing Data Collections," paper presented at 8th International Conference on Pervasive Computing (Helsinki, 2010).

A. Rice and S. Hay, "Decomposing Power Measurements for Mobile Devices," paper presented at 8th International Conference on Pervasive Computing and Communications (Mannheim, 2010).

F. Salim, "Probing Streets and the Built Environment with Ambient and Community Sensing," *Journal of Urban Technology* 19: 2 (2012) 47–67.

D.T. Wagner, A. Rice, and B. Alastair, "Device Analyzer: Understanding Smartphone Usage," paper presented at 10th International Conference on Mobile and Ubiquitous Systems (Tokyo, 2013).

H. Weinschrott, F. Dürr, and K. Rothermel, "Stream Shaper: Coordination Algorithms for Participatory Mobile Urban Sensing," paper presented at 7th International Conference on Mobile Adhoc and Sensor Systems (San Francisco, 2010).