[accepted manuscript before copyediting]
Tomoyo Sasao, Shin'ichi Konomi (2016). The Use of Historical Information to Support Civic Crowd- sourcing. Proceedings of the 4th International Conference on Distributed, Ambient, and Pervasive Interactions (DAPI 2016), Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016. pp.470-481. Lecture Notes in Computer Science 9749, Springer, Berlin/Heidelberg, 2016, ISBN: 978- 3-319-39861-7. http://dx.doi.org/10.1007/978-3-319-39862-4_43. The final publication is available at link.springer.com.

The Use of Historical Information to Support Civic Crowdsourcing

Tomoyo Sasao[1] and Shin'ichi Konomi[2]

[1]Department of Socio-Cultural Environmental Studies, The University of Tokyo
Kashiwa, Chiba, 277-8563, Japan
[2]Center for Spatial Information Science, The University of Tokyo
Kashiwa, Chiba, 277-8568, Japan
{sasaotomoyo, konomi}@csis.u-tokyo.ac.jp

Abstract. Context-aware notifications cannot be designed easily without knowing which context-aware notifications will be triggered and responded in time. In this paper, we discuss methods to improve the design of context-aware notifications. Using the data from our prior experiment, we identify main factors that influence citizens' responses to notifications and evaluate the predictability of quick responses using a simplified method. We then propose a model for designing civic crowdsourcing tasks based on historical information. We believe that creating well-designed notifications can decrease receivers' workloads and simultaneously expands the positive impacts of civic crowdsourcing on the quality of life in the city.

Keywords. Civic Crowdsourcing, Context-aware notification, Design

1 Introduction

This paper focuses on a smartphone-based locative media infrastructure for community-based crowdsourcing, which allows motivated citizens to generate tasks for citizens at large to help improve their local community and tackle social problems together. Recently, social ties of neighbors are becoming weaker and an increasing number of citizens are unconcerned spectators of their living environments. Raising citizens' awareness about local environments and triggering social actions are contemporary keys to achieve a sustainable community. We expect that smartphone-based context-aware notifications will play a key role in triggering awareness of social actions in proper place, time and situations, thereby providing solutions to pertinent issues in local communities.

A key problem of triggering citizens' actions by using context-aware notifications on smartphones is the difficulty to predict the timing to send notifications so as to maximize the chance that the receivers notice the notification and move into relevant action at the right time. Task designers may fail to create successful notifications even if they can define accurate task-relevant geo-fences easily. In our prior experiment, in which 19 citizens received community-related notifications during 28 days, only 30%

of the notifications were replied within 5 minutes. In other words, 70% of the notifications were replied more than 5 minutes after their delivery. This can be problematic since task designers often expect quick replies before the recipients move out of the corresponding geo-fences. This potential mismatch between the expectations of task designers and the actual behaviors of task workers can cause serious problems by degrading task quality, weakening task worker's motivation, and thereby undermining the sustainability of the crowdsourcing system.

In this paper, we present our methodology to use response logs of a mobile crowdsourcing system, which asks citizens small tasks through context aware notifications, so that task designers can grasp which areas are likely to generate quick *in-situ* reactions. In the case of our prior experiment, which used context-aware notifications for delivering crowdsourcing tasks in a local community, the system's response logs include the locations and timestamps of which task workers received and replied the notifications. We use them to find the areas with the high potential to collect *in-situ* reaction. Our work is in line with the previous research that focuses on historical information including the proposals of social navigation techniques [1][2], while little work has examined the potential of using historical information in the community-based computing. We also present our preliminary works to use historical information for giving advice to the designers of tasks and notifications.

First, we discuss the reasons why some notifications could induce quick responses. We analyze the response logs from our prior field study and find a primary factor that influences the acts of replying to notifications. Second, we test whether the variable, can be used for a statistical forecast of the areas that have high potential to generate quick responses. We validate the results by using the notification delivery and response logs of our prior one-month field study. Finally, we proposed a model of an interactive system that supports task designers to improve notifications. The system allows task designers to embed notifications in suitable locations and contexts. For example, if a new notification designed by a task designer is unlikely to induce sufficient *in-situ* responses, this model recommends to (a) change the content of the notification to make it easier to reply later in other places, (b) change the triggering time slot, and (c) change the triggering place to make the notification more likely to be responded quickly and *in situ*. We also discuss the effect of the proposed model for local community.

2 Prior Field Study

In our prior field study, we developed a smart phone app that receive context-aware notifications including small tasks about local community such as *"Take a photo of illegally parked cars"*, *"Do you see any garbage on the street? Choose yes or no"*, *"Please describe if you see any undesirable activities on this sidewalk. Input a short comment"* and so on and we also established a design environment for these notifications. Under the environment, citizens can design small community tasks by defining task formats, contents, and task triggers (see Fig. 1), and these notifications are made available on the smartphones of the residents. When the residents drop into

the geo-fences with the notifications, they are asked to do small tasks following the contents of the notifications.

We conducted an experiment in a small residential area (about 1km^2). The experiment has two parts; one is the notification design by a group of 4 citizens and the other is the notification use by 19 citizens in the local community. In the notification design part, safety-related notifications were generated, including the 21 notifications that are based on a colocated group work by the participants and the additional 21 notifications based on existing contents on the web. In the second part, we recruited the recipients of notification for one month from this target area. The recipients can respond to notifications if they want to perform the requested task. Nineteen citizens participated in this experiment. We collected month-long historical information including the locations and timestamps of all notifications and replies. We use these data to discuss the use of historical information in civic crowdsourcing.

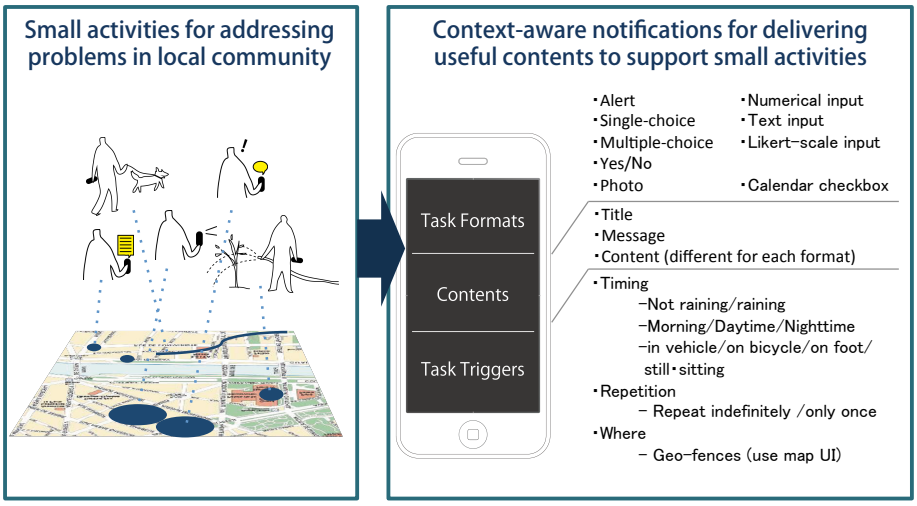


Fig. 1. Context-Aware Notification Design Environment (prototype in our prior work [3])

3 Related Works

3.1 Leveraging Historical Information

Worker's Mobility Information. Locations, transportation modes, traces, and so on are collected from crowdsourcing workers' mobile phones. Some researches use this personal historical information for estimating the most significant destinations for each person using clustering algorithms [4][5][6]. Moreover other researchers use those information for predicting and proposing more relevant to the user's next destination[7][8]. As another approach, some researchers use whole workers' historical information for building social mobility networks or throwing out coverage

where workers' works done. From these networks, they can simulate the movements of individuals as well as enclosing spatio-temporal stability of workers' behaviors[9][10]. In addition, Kazemi and Shahabi use location entropy to measure the total number of workers in that location as well as the relative proportion of their future visits to that location and they Least Location Entropy Priority (LLEP) Strategy as one of their task assignment protocol[11]. Reddy et al. presents framework of coverage based recruitment by processing workers' mobility historical data[12].

Application Manipulation logs. Ratings for working quality, and worker's reaction when given the opportunity to task participation (Participation likelihood) can be collected from the workers' mobile phones. These data have potentials to be used in participation and performance based recruitment[12], and in creation of suitable incentives by person[13].

3.2 A Type of Mobile Crowdsourcing Tasks

Although researches based on ideas of using mobile devices and defining citizens as task workers for crowdsourcing are increasing recently, efficient approaches are different depending on what kind of data a task designer wants to collect or what kind of tasks a task designer wants to serve. In our research, crowdsourcing tasks require the workers to be physically located at that location and in registered contexts in order to perform the corresponding task. Additionally, the tasks don't have a worker list that is created from the worker's rating in advance, and they require people who pass close to the task location. While the number of research projects on mobile crowdsourcing in smart cities are increasing recently, few studies focus on the site-specific task, which need to be performed in that location. Kazemi and Shahabi focus on same type of mobile tasks with ours called *special tasks* such as taking a picture from a particular building[11] and presented some effective task assignment protocols including the one based on current worker's location and location entropy. Our focus in this paper is a method of enriching task qualities using historical information for such kind of spatial crowdsourcing.

3.3 Quick Responses

Research that focuses on quick responses mainly uses the following two approaches. First one is a smartphone based interaction approach, which focus on designing smartphone interface for quick input in performing crowdsourcing tasks. This includes the uses of unlock screen [14][15][16]. These interfaces make it easy to engage crowdsourcing workers in the place where the workers receive the task. They facilitate such engagement by decreasing the worker's workloads in that place. Second one is a physical interaction based approach, which focus on embedding the interaction in public spaces and making strict bonds that cannot be separated between a task receiving spot and a task performing spot. Gallacher el al. presented tangible questionnaire box and succeed to collect event attendees' relevant opinions during the

event [17]. However, there are few researches encouraging to quick responses for mobile crowdsourcing tasks by using historical information such as spots of notification received and responded.

4 Main Factors That Influence Citizens' Responses to Mobile Crowdsourcing Tasks

Context-aware notifications are triggered in many different situations and their contents may include all kinds of topics. Thus, it is difficult to forecast whether someone responds to a notification and whether the response is provided quickly. We next examine main factors that influence whether or not citizens respond to notifications and if their responses are provided quickly.

4.1 Notification Response

To search the main factors that influence the responses to notifications, we use binomial logistic regression with the dependent variable indicating whether or not the participant respond to a received notification. We prepared 51 candidate independent variables from various categories (i.e., notification receiver's characteristics such as demographic attributes, perceptions of the local community, familiarity with smart devices; characteristics of notifications such as their contents, context, and geo-fences; characteristics of the contexts at the time when receiving the notification such as impression of the received notification) and carefully selected relevant variables that may have significant effects on the result considering multicollinearity. Chi-squared test has been used to check the variables one by one and variance inflation factor (VIF) has been computed in incremental steps. The final model has the VIF that is under 5.

Finally, our comparison of two sets of notifications, one of which is designed by 4 local community group members and the other designed based on existing contents on the web, shows that the logistic regression models were statistically significant ($p<0.01$). The model of the former set of notifications explained 47.2%(Nagelkerke R^2) of the variance in a notification response event and correctly classified 86.9% of the cases. The model of the latter set of notifications explained 60.6%(Nagelkerke R^2) of the variance in a notification response event and correctly classified 84.3% of the cases.

The result we obtained from the analysis with the selected variables has revealed the main factors that influence whether or not notifications are responded for both sets of notifications. They are the variables concerned with the user's contexts when receiving the notifications. In particular, occurrences of other notification events before and after the received notification strongly impact the decision whether or not to respond to notifications.

4.2 Quick Response

To explore main factors that influence citizens' quick responses, we classified all notification events in our prior experiment (n=4460) into the following 4 classes according to response patterns:

(1) notifications that are responded immediately (i.e., within 5 minutes),
(2) notifications that are responded more than 5 minutes after their receipt, and immediately after the receipt of other notifications,
(3) notifications responded more than 5 minutes after their receipt, and *not* immediately after the receipt of other notifications, and
(4) notifications that are not responded.

Figure 2 shows spatial distributions of each pattern of notifications. We can see that there are a number of overlapping areas among these spatial distributions, and that we can't clearly distinguish Pattern 1 from the other patterns only based on these spatial distributions.

Next, we examine the events during the period between 5 minutes before and after the receipt of notifications, such as the mean numbers of the other notifications received in this period and the mean numbers of the other notifications replied in this period (see Table1). The result shows that Pattern 1 has much higher average numbers of other notifications replied in this period than other patterns. We thus use this variable for forecasting the places where citizens respond to notifications quickly.

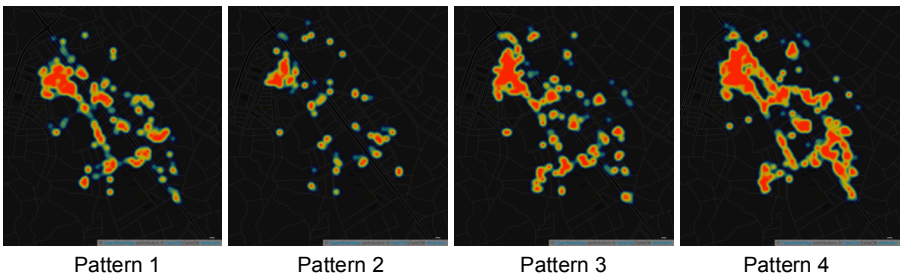


Fig. 2. Heat Maps of Emerging Notifications

Table 1. Classification of notifications based on the patterns of their responses

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
	Responded immediately (≦ 5 min.)	Responded later (> 5 min.) (immediately after other notifications)	Responded later (> 5 min.) (*not* immediately after other notifications)	No responses
Number of notifications (%)	660 (14.8%)	372 (8.3%)	1,129 (25.3%)	2,299 (51.5%)
Mean response time (h:m:s)	0:01:29	1:04:32	3:10:13	N/A
Mean number of other notifications received during the ±5 min period	4.02	4.75	3.34	4.15
Mean number of other notifications replied during the ±5 min period	4.88	0.20	0.24	0.55

5 Evaluating the Predictability of Quick Responses

In this section, we evaluate the predictability of the places where citizens respond to notifications quickly using the response logs. If such prediction is confirmed to be feasible, designers of civic crowdsourcing tasks can create efficient notifications based on predictions.

According to the result of Section 3, the notifications that participants responded quickly (i.e., Pattern 1) have much higher numbers of other notifications responded in the ±5 minutes period. We infer from the result that quick responses might be forecast from the spatial distribution of the numbers of responses. Then, as a preliminary analysis, we verify the predictability of quick responses based on the spatial distribution of the number of responses using the log file from our prior experiment.

5.1 Method

The log file includes the data from the 4-week period. We use the data from the first 3 weeks to build a forecasting model, and the data from the last 1 week to evaluate the model.

First, we have constructed a spatial distribution of response events using the data from the first 3 weeks. The data are aggregated based on grids with different sizes (see Table 2).

Table 2. Grids used for constructing spatial distribution

	Grid-4	Grid-16	Grid-64	Grid-256	Grid-1024	Grid-4096
Num of cells	2×2	4×4	8×8	16×16	32×32	64×64
Length of cell	1295m	324m	162m	81m	41m	20m

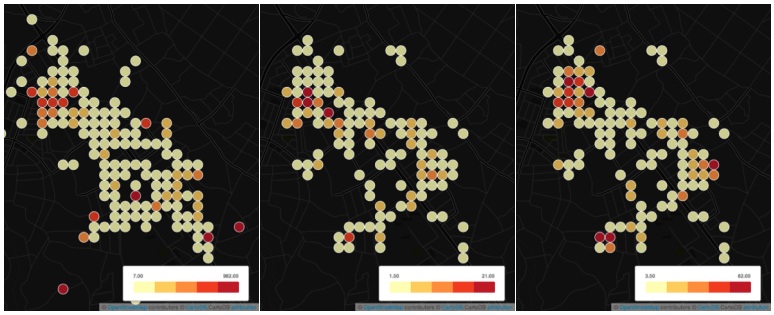
Second, we divide the cells into two groups based on the median of the number of response events. The group that has values larger than the median (G^+) is regarded as areas where quick responses can be obtained frequently. We expected that the areas with the other group (G^-) would not induce quick responses.

Third, we have constructed a spatial distribution of triggered notifications that were responded quickly (i.e., Pattern 1) using the data from the last week, and aggregated the events for each cell. We compare the numbers of quickly responded notifications between the two groups using independent t-test.

5.2 Result

We found that, in Grid-256, Grid-1024 and Grid-4096, cells in group G^+ had statistically significantly larger numbers of notification events that got a quick response (Pattern1) in the last week: Grid-256[$t(24.833)=3.209$, $p<0.01$], Grid-1024[$t(27.515)=2.875$, $p<0.01$], Grid-4096[$t(28.847)=2.105$, $p<0.05$]. Based on this result, it is possible to forecast the places in which many notifications are quickly responded by using historical response logs for Grid-256, Grid-1024, and Grid-4096.

We also found that in Grid-256, Grid-1024 and Grid-4096, cells in group G^+ had statistically significantly larger numbers of ignored notification events (Pattern 4) in the last week: Grid-256[$t(24.468)=2.765$, $p<0.05$], Grid-1024[$t(28.820)=2.723$, $p<0.05$], Grid-4096[$t(29.576)=2.494$, $p<0.05$]. This result shows that, although we can forecast the places in which many notifications are quickly responded, the same places may have a large number of ignored notifications. It is difficult to divide these two phenomena by only using a spatial distribution of response events (see Fig. 3). To cope with this issue, we could consider other conditions such as the ones based on time.



number of response (model data) | response pattern1 (test data) | response pattern 4 (test data)

Fig. 3. Heat map based on the count each Grid

6 Designing Civic Crowdsourcing Tasks Based on Historical Information

The result of Section 4 shows that the places where notification recipients tend to respond quickly can be predicted based on historical information. Historical information will increase gradually, thereby improving prediction accuracy. As the accuracy improves, such prediction can be used to provide suggestions for designing effective notifications that induce quick responses. We propose a model for supporting civic crowdsourcing task designers to design effective context-aware notifications. The model is based on notifications that have already been embedded in a local community as well as their usage logs. This model can be used to provide advice to improve notifications based on the places predicted to have high or low chances of acquiring quick responses. For example, if a notification designer sets a notification in a place with a high chance of acquiring quick responses, this model can let the designer know that the place is appropriate for a site-specific task such as *"take a picture of this park."* Meanwhile, if a notification designer wants to collect quick responses but sets the notification in a place with low chance of acquiring quick responses, the model will recommend to change the content of the notification so that responses can be useful even when they are collected at a later point in time based on the memories of respondents, or tweak the timing and the location of the notification to collect responses as quickly as possible.

Figure 3 shows an example that illustrates the usage of the model for the support of designing notifications. The goal here is to increase the number of quickly responded notifications. First, a notification designer creates a context-aware notification without worrying about how quickly it might be responded. The designer specifies (1) notification contents (e.g., task requests and questions), (2) areas that trigger notification (e.g., geo-fences), and (3) other detailed trigger contexts (e.g., time span, weather, activity patterns such as walking speed or activity categories such as walking, biking, driving, etc.). Our model uses these pieces of information input from

the notification designer, and provides pertinent advice to improve the quality of the notification. First, this model judges whether or not a designed trigger area is likely produce quick responses. If the notification has been set in a "quick response area," our model shows a message "your notification is set in a context that can produce quick responses" on the screen, and the designer can revise the notification contents based on the message. Meanwhile, if the notification has not been set in a "quick response area," our model can create the following three kinds of advice and shows the most suitable one on the screen; (1) change the time slot of the notification, (2) change the area of the notification, (3) modify the content of the notification by taking into account that no quick responses may be obtained.

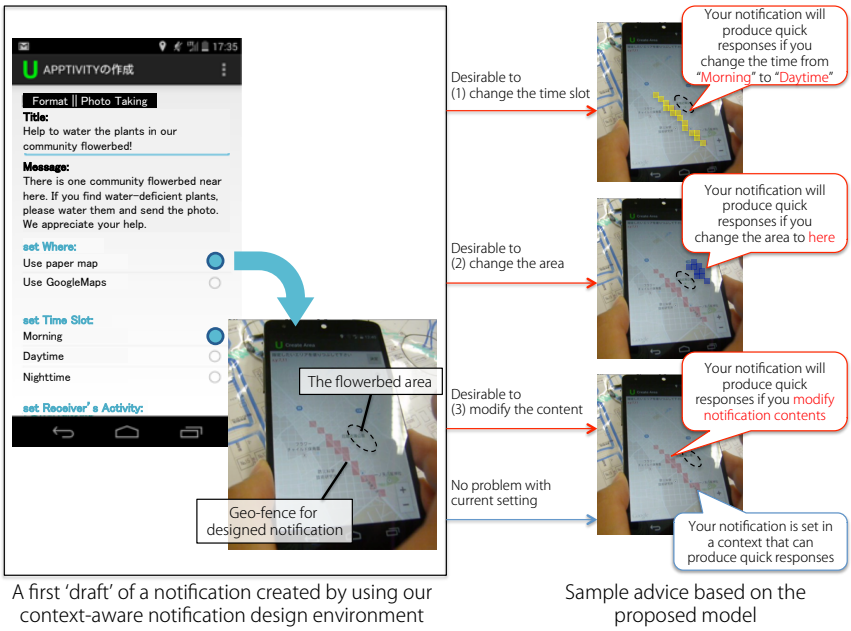


Fig. 4. Supporting Model of Designing Notifications

6.1 Limitation

We discuss the limitations of our model as follows:

Different historical data. The model is based on the predictability of quick responses using historical information. In this study, we only use the logs from our prior experiment, which was conducted during a one-month period involving only 19 participants. We also focused on crowdsourcing for safety in a specific local community. Thus our data may be biased for these specific conditions, and we cannot make strong general conclusions based on these data. However, it is possible to extract different predictor variables which are suitable in different conditions by using

the processes described this study. We can do so by using different historical information collected in different conditions. In future, we will clarify the minimum amount of data required for reasonable prediction and build a sophisticated forecast algorithm using time stamp data and so on.

Consideration of "no response." This study supposed that "quick response places" are the areas in which many people respond quickly after receiving notifications. We then forecast spatial distribution of the amount of quick responses. The reason why we focus on forecasting the amount of quick responses rather than the rate of quick responses is that our current purpose is to allow notification designers to collect many quick responses. While this can be highly useful for notification designers, it may not reduce the number of ignored notifications. In future, we will also consider forecasting the rate of quick responses.

7 Conclusion and Future Works

This paper focuses on the difficulty of controlling the timing of which context-aware notifications are responded. To solve this problem, we discussed methods to improve notification design environments, rather than the environments of notification recipients. Nevertheless, we believe that creating well-designed notifications can decrease receivers' workloads and simultaneously expands the positive impacts of civic crowdsourcing on the quality of life in the city. Using the data from our prior experiment, we found main factors that influence citizens' responses to mobile crowdsourcing tasks and evaluated the predictability of quick responses using a simplified method. We then proposed a model for designing civic crowdsourcing tasks based on historical information.

As a future work, we will develop methods to improve the environments of notification recipients to decrease their cognitive workloads by controlling the amount and frequency of notifications. They are to be integrated with the model proposed in this paper.

References

1. Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T.: Edit wear and read wear. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '92), Penny Bauersfeld, John Bennett, and Gene Lynch (Eds.) pp. 3-9. ACM, New York (1992)
2. Dieberger, A., Dourish, P., Hook, K., Resnick, P. and Wexelblat, A.: Social navigation: techniques for building more usable systems. *interactions, 7, 6,* pp. 36-45. ACM Press (2000)
3. Sasao, T.: Support Environment for Co-designing Micro Tasks in Suburban Communities. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '15), pp. 231-234. ACM, New York (2015)
4. Ashbrook, D., Starner, T.: Using GPS to learn significant locations and predict movement across users. In: *Personal and Ubiquitous Computing*, pp. 275–286 (2003)

5. Kim, M., Kotz, D., Kim, S.: Extracting a mobility model from real user traces. In: *Proceedings of Infocom*, pp. 1–13. IEEE, Los Alamitos (2006)
6. Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L.: Discovering personal gazetteers: an interactive clustering approach. In: *Proceedings of GIS*, pp. 266–273. ACM, New York (2004)
7. Bhattacharya, A., Das, S.: LeZi-update: an information-theoretic approach to track mobile users in PCS networks. In: Proceedings of Mobicom, pp. 1–12. ACM, New York (1999)
8. Krumm, J., Horvitz, E.: Predestination: Inferring destinations from partial trajectories. In: *Proceedings of Ubicomp,* pp. 243–260. ACM, New York (2006)
9. Hsu, W., Dutta, D., Helmy, A.: CSI: A Paradigm for Behavior-oriented Delivery Services in Mobile Human Networks. *ACM Transactions on Networking* (2008)
10. Eagle, N., Pentland, A.: Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing* 10(4), pp. 255–268 (2006)
11. Kazemi, L. and Shahabi, C. GeoCrowd : Enabling Query Answering with Spatial Crowdsourcing. Proc. *Int'l Conf. Advances in Geographic Information Systems*, pp. 189–198. (2012)
12. Reddy, S., Estrin, D., Srivastava, M.: Recruitment framework for participatory sensing data collections. In *Proceedings of the 8th international conference on Pervasive Computing* (Pervasive'10), Berlin, Heidelberg, pp. 138-155. (2010)
13. Feng, Z., Zhu, Y., Zhang, Q., Zhu, H., Yu, J., Cao, J., & Ni, L. M.: Towards truthful mechanisms for mobile crowdsourcing with dynamic smartphones. In *Distributed Computing Systems (ICDCS), 2014 IEEE 34th International Conference,* pp. 11-20. IEEE (2014)
14. Vaish, R., Wyngarden, K., Chen, J., Cheung, B., Bernstein, M.S.: Twitch crowdsourcing: crowd contributions in short bursts of time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). pp. 3645-3654. ACM, New York (2014)
15. Truong, K.N., Shihipar, T., Wigdor, D.: Slide to X: Unlocking the potential of smartphone unlocking. In *the Proceedings of CHI 2014: The ACM Conference on Human Factors in Computing Systems,* Toronto, Ontario, pp. 3635-3644. (2014)
16. Banovic, N., Brant, C., Mankoff, J., Dey, A.: ProactiveTasks: the short of mobile device use sessions. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (MobileHCI '14). pp. 243-252. ACM, New York (2014)
17. Gallacher, S., Golsteijn, C., Wall, L., Koeman, L., Andberg, S., Capra, L., Rogers, Y.: Getting quizzical about physical: observing experiences with a tangible questionnaire. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). pp. 263-273. ACM, New York (2015)