Studying the Effects of Task Notification Policies on Participation and Outcomes in On-the-go Crowdsourcing

Yongsung Kim, Emily Harburg, Shana Azria, Aaron Shaw,
Elizabeth Gerber, Darren Gergle, Haoqi Zhang
Northwestern University, Evanston, IL USA
{yk, eharburg, shanaazria2016}@u.northwestern.edu, {aaronshaw, egerber, dgergle, hq}@northwestern.edu

Abstract

Recent years have seen the growth of physical crowdsourcing systems (e.g., Uber; TaskRabbit) that motivate large numbers of people to provide new and improved physical tasking and delivery services on-demand. In these systems, opportunistically relying on people to make convenient contributions may lead to incomplete solutions, while directing people to do inconvenient tasks requires high incentives. To increase people's willingness to participate and reduce the need to incentivize participation, we study *on-the-go crowdsourcing* as an alternative approach that suggests tasks along people's existing routes that are conveniently on their way. We explore as a first step in this paper the design of *task notification policies* that decide when, where, and to whom to suggest tasks. Situating our work in the context of practical problems such as package delivery and lost-and-found searches, we conducted controlled experiments that show how small changes in task notification policy can influence individual participation and actions in significant ways that in turn affect system outcomes. We discuss the implications of our findings on the design of future on-the-go crowdsourcing technologies.

Introduction

Recent years have seen the growth of physical crowdsourcing systems that motivate large numbers of people to provide new and improved physical tasking and delivery services on-demand. With a few button clicks, people can connect to workers who provide rides (Uber, Lyft), deliver groceries or meals (Instacart, Postmates, DoorDash), complete errands (TaskRabbit), and walk dogs (Wag). Enabled by network-connected mobile devices, mobile apps help connect people to tasks and lower barriers to participation, and market algorithms effectively coordinate the provision of physical resources on-demand. Together, these system components facilitate transactions and interactions to scale services to a broader market and worker population in ways that are transforming entire service sectors.

While physical crowdsourcing systems may change who can participate and what tasks can be done, applications are still limited by the fact that contributions are gathered via one of two approaches: *opportunistic* or *directed*. Opportunistic approaches rely on contributions from workers in situations where the workers themselves decide when and where to contribute. While convenient for workers, opportunistic tasking behaviors can lead to incomplete tasks or slow task completion (Han et al. 2015; Thebault-Spieker, Terveen, and Hecht 2015). As a result, commercial physical crowdsourcing systems mainly use a directed approach to assign tasks to workers that best address system needs. However, a directed approach often requires contributions that are largely outside workers' existing routines and as a result strong incentives are required (Thebault-Spieker, Terveen, and Hecht 2015). For example, a directed approach may lead to the assignment of tasks that require significant travel and incentives in order to increase workers' willingness to complete the task (Teodoro et al. 2014).

We study *on-the-go crowdsourcing* as a hybrid approach that seeks to bring about the best elements of both opportunistic and directed approaches. On-the-go crowdsourcing systems use people's existing routes to complete physical tasks that are conveniently on their route, but do so in ways that also address system needs.

We focus in this paper on understanding the challenges in designing effective *task notification policies* that determine when, where, and to whom to suggest tasks. Unlike directed approaches that assume workers are insensitive to when and where task assignments are sent so long as the tasks are generally nearby, we hypothesize that people contributing on-the-go may be quite sensitive to exactly when, where, and how task suggestions are presented. Task notification policies can thus affect individual participation and actions that in turn affect global tasking outcomes and can influence the success of an on-the-go crowdsourcing system.

To better understand such challenges, we design task notification policies to study 1) how slight changes in notification radius can affect individual participation and user recruitment, and 2) how slight changes in the timing of notifications can affect individual task actions and system outcomes. We conducted controlled experiments and the results show that small changes in task notification policies can affect individual participation and actions in significant ways that ultimately affect system outcomes. Specifically, we found that small changes in the notification radius had a drastic effect on user recruitment and user participation in package delivery settings. In addition, we found that small changes in the timing of notifications had a significant impact on individual actions, which in turn affected global search coverage in lost-and-found settings.

This paper makes the following contributions:

- We design and experimentally assess different task notification policies in order to understand how small policy changes can significantly affect participation and actions in ways that influence system outcomes.
- We identify and report on situational factors that affect on-the-go participation that are not prominent in existing physical crowdsourcing systems, such as *providing little or no travel detour*, *limiting the need to walk back*, and *assessing the likelihood of a user's hands availability*.
- We discuss the implications of our findings for the design of on-the-go crowdsourcing systems, and highlight the need for new design and algorithmic work to advance this promising new model for physical crowdsourcing.

Related work

Existing physical crowdsourcing systems rely on one of two approaches for soliciting contributions: directed and opportunistic. The directed approach routes tasks to workers who are either nearby or who fit the task request criteria. Many existing on-demand services use the directed approach to ensure task completion and task quality. However, the directed approach requires high incentive for inconvenient tasks, such as tasks that are at inconvenient locations (Teodoro et al. 2014) or that require large travel distances (Thebault-Spieker, Terveen, and Hecht 2015). Long travel distances not only deter workers from participating (Teodoro et al. 2014; Musthag and Ganesan 2013), but also prevent requesters who cannot afford the associated higher cost for the service. The opportunistic approach relies on workers to contribute as they wish, e.g., by browsing and selecting tasks from a list. Some community-driven systems, such as *time-banking* systems which support members of a community exchanging tasking services, use the opportunistic approach because it provides workers freedom and convenience to choose when to complete tasks and what tasks to complete. Since this relies solely on worker's task selection behavior, it often leads to task incompletion (Han et al. 2015) and can leave requesters whose tasks have not been completed unsatisfied with the service. On-the-go crowdsourcing seeks to bring about the best elements of both opportunistic and directed approaches to permit access to a large number of people who can contribute through their existing routines in ways that better achieve system goals.

Prior theoretical work highlights the possibility of coordinating physical crowd work to enable new applications by effectively leveraging people's existing routines (Sadilek, Krumm, and Horvitz 2013; Chen et al. 2014). Sadilek et al. introduced Crowdphysics and conducted empirical study with geo-tagged tweets to demonstrate that it may be possible to coordinate long-distance package delivery by leveraging people's existing routines with minimal diversion and short wait time (Sadilek, Krumm, and Horvitz 2013). Chen et al. introduced an approach for optimizing the sequence of physical tasks to assign to workers so as to minimize travel detours (Chen et al. 2014). While similarly focused on the idea of promoting convenient contributions, the methods in these prior works broadly assume that tasks will be accepted whenever assigned, which makes them a poor match for on-the-go crowdsourcing in which helpers may or may not perform tasks presented to them. We find through our studies on task notification policies that how we suggest tasks can significantly affect system outcomes, and present later in the paper the need for further technical work to support on-the-go crowdsourcing systems in practice.

Existing physical crowdsourcing applications that leverage people's existing routines focus primarily on data collection and rely on the opportunistic approach. One approach is to attach sensors onto people and objects that travel regular routes to passively collect sensing data, e.g., on bikes (Eisenman et al. 2007), street sweepers (Aoki et al. 2008), and public transportation (Aberer et al. 2010). This approach does not require active user involvement but can only collect data using machine sensors. Another approach is to opportunistically collect data from active user. For example, Tiramisu crowdsources bus information from bus riders to provide real-time arrival time (Zimmerman et al. 2011), and Twitch uses smartphone unlocks to collect coarse-grained census data (Vaish et al. 2014). This approach leverages human sensors, but are limited to the situations and locations in which people opportunistically contribute data. In contrast, our on-the-go approach actively suggests tasks to potential contributors using task notification policies that must balance the goals of promoting convenient and valuable contributions to advance individual and system goals.

When delivering tasks to potential helpers, task notification policies need not only consider interruptible moments in which people are more likely to accept notifications (Ho and Intille 2005; Fischer, Greenhalgh, and Benford 2011) but also situational factors that can significantly affect one's interest in contributing to a task. In this direction, our studies contribute a set of small but significant factors, or *channel factors* (Ross and Nisbett 2011), that affect participation in on-the-go crowdsourcing systems. Understanding these factors can help us to more accurately identify situations in which someone on-the-go may be able to help and be used to extend models that consider only distance-based measures for the cost of diversion (Horvitz and Krumm 2012).

Task Notification Policies for On-the-go Crowdsourcing

We envision on-the-go crowdsourcing systems that suggest tasks to potential helpers who happen to pass by task locations. These systems implement *task notification policies* that decide when, where, and to whom to suggest tasks in a way that balances individual convenience with the system's needs for timely, accurate, and complete solutions. To make decisions, a task notification policy may consider input factors such as a potential helper's location, their distance to tasks, their likely future routes, the urgency of tasks, the likelihood that others will become available in the near future, and so on. Based on this input, a policy specifies the conditions under which to notify a potential helper about a task opportunity. For example, a policy may specify a task radius

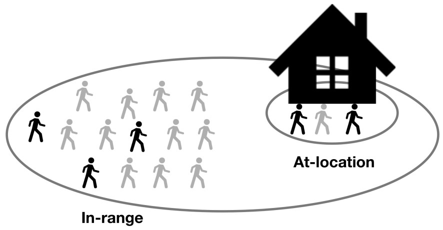

Figure 1: We study two task notification techniques, *in-range* and *at-location*, that differ only in the radius within which to notify potential helpers of task needs. In this illustration, black and grey colors respectively denote people's decision to help or not help when notified.

within which to notify people about tasks, which in turn affects the size and composition of the pool of helpers who receive notifications. A policy may also specify the precise timing of notifications, affecting when and where the helper sees the notification and in turn may influence their decisions on whether to contribute and which tasks they perform.

We describe below two core challenges in designing task notification policies with illustrative scenarios drawn from practical problems such as package delivery and lost-and-found searches. We introduce and compare a number of notification techniques to help to illustrate these challenges.

Balancing individual disruption and the quality-of-service

Scaling a physical tasking service with an on-the-go crowd requires balancing desired qualities of service with the disruption to individual routines in order to promote efficiency and ensure long-term community viability. Given a set of task requests and incomplete knowledge of who might be willing and able to help, a system must decide whom to suggest tasks to from among the pool of potential helpers who pass by a task location. Setting an aggressive notification policy that notifies many potential helpers can quickly identify those who agree to help but may disrupt numerous others who cannot help. While this may help satisfy demand right now, it risks notification blindness over time that will eventually reduce the pool of potential helpers. Setting a conservative notification policy that suggests tasks only to those who are deemed most likely to help may risk recruiting too few helpers, which can lead to task delays or incomplete tasks. Over time, this approach risks overburdening the best helpers and can lead to temporal burnout that causes the best helpers today to eventually exit the system for good.

Illustrative scenario: Community package delivery To illustrate the challenges of managing the supply of helpers and deciding whom to notify of task opportunities, consider a package delivery scenario where the goal is to leverage people's existing travel for picking up packages from a mail center and delivering them to others who could not pick up packages on their own. The system allows people to submit requests for pickup and then notifies potential helpers passing by to perform the pickup. As potential helpers approach the mail center, the system must decide which sets of helpers to notify, for example, based on how likely they are able to help and on whether deliveries are time critical.

In such a scenario, identifying an effective task notification policy is non-trivial. Using an aggressive task notification policy allows us to reach people who are able and willing to help, such as someone who was already going to the package center or someone who wasn't going to the center but is willing to help after a notification. However, at the same time, we are also sending notifications to many people who are not able and willing to help right now (e.g. they don't currently have physical ability to carry the packages or they are too far from the package center). If significant disruption continues, it may lead this group of people to eventually ignore notifications, and result in their leaving the pool of likely helpers in the future.

On the other hand, using a conservative task notification policy reaches people who are already going to or at a package center. This may lead to higher percentage of pick-ups since it's likely to reach people who are able and willing to help, and would alleviate issues of notifying people who can't or don't want to help right now. However, it also misses opportunities to notify likely helpers who would have helped if we notified them, and as a result we may overly rely on people who often go to package center and cause them to burn out eventually.

In-range and at-location notification techniques To better understand the challenges in setting task notification policies to balance the goals of reaching willing helpers and disrupting others, we introduce and compare *in-range* and *at-location* notification techniques to study how notification radius might affect user recruitment and individual participation (see Figure 1). In-range notifications suggest tasks to potential helpers who pass by regions that are within a certain range of the task location. At-location notifications monitor a user's distance to the task location and only notifies them of a task when they are at the location. As the notification radius shrinks, we expect to reach fewer potential helpers but see that those receiving notifications complete a higher percentage of tasks suggested to them. In our first experiment, we use these two notification techniques in the context of package delivery to study users' willingness to pick up packages and overall user recruitment with regard to small distance changes.

Promoting contributions where they are most needed

Another core challenge in setting task notification policies is deciding when to suggest tasks. We assume that on-the-go crowdsourcing should minimally disrupt people's routines, and not require the attention of potential helpers until a task request is made. The timing of task requests may affect where helpers complete tasks, which tasks are completed, and which tasks are skipped. Contributions may be valued more or less depending on task's importance and time-criticality, as well as how many people can complete the task and are likely to come across it. A task notification policy risks either (1) sending a notification too early, which

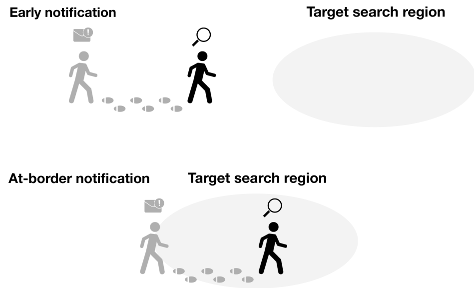

Figure 2: We study two task notification techniques, *early notification* and *at-border notification*, that differ only in the timing of when notifications are sent to helpers to alert them of a task need.

can lead to helpers completing a low-valued task when their routes would have led them to a nearby higher-valued task; or (2) passing on an opportunity to send a task or sending a task too late, which may lead to lost opportunities.

Illustrative scenario: Lost and found searches To illustrate the challenges of determining when to suggest tasks to potential helpers, we consider a lost & found scenario where the goal is to effectively coordinate the search for lost items with people who can help look for lost items along their existing routes. A requester who has lost an item marks the region(s) where it may have been lost, and the system notifies potential helpers who pass by the marked region. As a helper walks through a region, the system must decide when to notify the helper with the goal of steering their attention toward areas that most require their search efforts.

A general challenge in timing notifications is to avoid wasting search efforts while at the same time avoiding missed search opportunities. For example, if we notify helpers too early before they reach the target search region, the helpers may search for lost items in an unintended region and thus their searches will be wasted. If we notify people right at the beginning of the search region, by the time they respond to the notification and start looking for the item, they might have already passed some parts of the target search region, thereby leaving some areas unsearched. In either case, individual search efforts affect overall search coverage, and the timing of notifications thus affect search coverage globally and the overall efficiency of the system.

Pre-tracking and early notification techniques Many existing systems rely on *geofencing* for location-based notifications that trigger notifications anywhere around 100 meters from a centroid of a geo-fence (Rodriguez Garzon and Deva 2014). We posit that geofencing is inadequate for setting task notification policies because it is limited in its precision and ability to steer users to start searching at precise locations. To overcome this, we introduce *pretracking* as a general method for precisely notifying helpers in a more timely manner. Pre-tracking triggers fine-grained GPS monitoring once a potential helper approaches a region of interest. Upon entering the region, it continuously makes decisions based on precise tracking of a helper's location and uses this information to time when to deliver tasks.

Using a pre-tracking technique, we design two notification techniques, *early notification* and *at-border notification*, to study how the timing of notifications affects individual actions and global system outcomes. The early notification technique sends notifications ahead of time, by taking into account factors such as the helpers' interaction time with notifications and their walking rate, so that they might start a task at the beginning of a target region by the time they decide to help. The at-border notification technique sends notifications right at the border of the target region where the system wants the helper to start performing the task. In our second experiment, in the context of lost & found, we study how slight changes in the timing of notifications can affect where helpers' searches are made in ways that affect the value of their individual searches and overall search outcomes.

Controlled Experiments

As a first step to explore the challenges mentioned in the previous section, we designed two controlled experiments to study the challenges in 1) balancing individual participation and user recruitment, and 2) promoting convenient contributions that are most valued. We will also discuss on-the-go situational factors we discovered in the experiments.

Experiment 1: The effect of notification radius on user recruitment and participation

Methods We conducted a two-week long, within-subjects experiment for a package delivery scenario in order to understand 1) how sensitive user recruitment and participation are to notification radius changes, and 2) situational factors influence helping behaviors. We measured the number of notifications sent for user recruitment and the task pickup rate for user participation. We used a within-subject design to limit the effect that large differences in participant's routines could have on our findings; some participants may frequent task location areas more than others, which could unduly influence the number of notifications being sent.

For the experiment, we developed a prototype of Libero, a mobile application that collects package delivery requests and routes them to potential helpers. In our setting, since requesters receive package notifications from the package center through their email, they can either forward the email to the system or take a screenshot of the email if they want to request that their package be picked up. First, a helper receives package pick-up request notifications (Figure 3a) as they pass by the package center. Next, if the helper decides to pick up the package, she is redirected to "Others' Requests" where she sees a list of the packages with package details (Figure 3b). To pick up a package, the helper selects the package she wants to pickup and, after confirming the pickup, she is provided with the image of the requester's package notification email. As this occurs, the requester whose package was just picked up, receives an email and an in-app notification updating her of the status of her request, revealing the helper's name as well as the date and time of pickup.

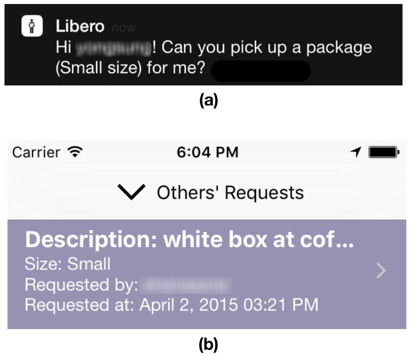

Figure 3: (a) Libero sends notifications to people who pass by a package center; (b) potential helpers can see the details of the package.

We used two notification conditions: *at-location notifications* that notify users when they enter a task location, and *in-range notifications* that notify users when they pass by a region that is within 100 meters of the task location. To implement at-location notification, we placed a bluetooth low energy beacon with the broadcasting signal power of 4dB (which translates into 1-3 meters in distance) and 1200ms as the interval. To implement in-range notification, we used a geo-fence with 100 meters as the radius.

We recruited 16 people who had iPhone 5 or above via flyers and local university mailing lists. The average participant age was 25 (SD: 3) with 9 male and 7 female participants. They were randomly assigned to one notification condition and asked to switch to the other condition after one week. The order of conditions for each person was counterbalanced. One participant stopped participating during the second week of the study due to personal reason, so we excluded that participant's data from the quantitative analysis.

In both conditions, when the participants accepted a request notification they were asked to deliver packages from a specified pick-up location and take them to a specified drop-off location. We chose a coffee shop that is close to the nearest train station from a school building where many classes take place as a pick-up location, and set up a drop-off location in a building that is next to the school building. We chose these two locations as pick-up and drop-off locations because the participants frequently visit places that are nearby and as a result they would not have to deviate much from their travel routine.

An author served as a requester during the experiment. Since we did not want the package size to affect willingness to help, we only requested packages that were small enough to be carried by one hand.

The participants were also asked to rate statements related to the perceived cost of disruption at two times during the study: once after the first week and another time after the second week. We also interviewed participants who delivered packages at least once in order to better understand what were the factors affecting their willingness to pick up. The study lasted for two weeks and the participants received a $25 gift card as compensation.

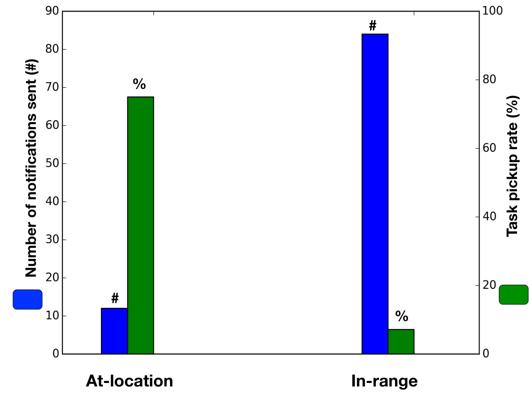

Figure 4: Comparison between at-location and in-range condition in terms of number of notifications sent and task pickup rate.

Results Our results showed that even the small distance changes that occurred across our two notification conditions had a drastic and inverse effect on both user recruitment and user participation (Figure 4). There were 7 times more opportunities presented in the in-range condition than in the at-location condition, with 84 notifications sent to the participants in the in-range condition, while only 12 notifications were sent in the at-location condition. On average, the participants in the in-range condition received 5.6 notifications (SD: 4.37), while the participants in the at-location received 0.8 notifications (SD: 0.8). A Wilcoxon Signed-Ranks test shows that there is a significant difference between the two conditions ($Z=3.49$, $p<0.001$).

We saw a significant effect in the reverse direction for task pickup rate, which was almost 10 times lower in the in-range notification condition than the at-location condition with an average of 7.13% (SD: 14.79) task pickup rate in the in-range condition and the average of 75% (SD: 28.87) in the at-location condition. Wilcoxon Signed-Ranks test results show that there is a significant difference between two condition ($Z=3.01$, $p=0.003$).

Although there is a significant difference in task pickup rate and number of notifications sent, we found that participants' perceived cost of disruption was low in both conditions. The participants' response on a 5-point likert scale asking about the perceived cost of disruption had a mean response of 1.8 (SD: 0.77) for the at-location condition and 2.33 (SD: 1.11) for the in-range condition (1 indicates not disruptive at all while 5 indicates very disruptive), and there was no significant difference between the conditions ($Z=1.29$, $p=0.2$).

While it is limited to our experiment, our results show the potential for leveraging on-the-go crowds to deliver packages in which people are willing and able to pick up tasks when they are suggested at the right time. During the study, there were 13 packages delivered by 5 different participants. Three participants delivered packages in both conditions while 2 participants only delivered in one condition (one in at-location, one in in-range). On average, the participants received 0.79 notifications (SD:0.56) per day.

Discussion Our exploration with in-range and at-location notifications showed that small changes in distance for task notification policy resulted in significant opposing effects on individual participation and user recruitment. It resulted in high task pickup rate and low numbers in user recruitment for the at-location condition, while resulting in low task pickup rate and large numbers in user recruitment for the in-range condition.

The results highlight the trade-offs of setting task notification policies to achieve desired outcomes; they do not to show that either notifying at-location or in-range is better than the other. Instead, our results show that small changes in task notification policies can affect individual participation in significant ways. It should also be noted that the significance of such a shift may vary depending on who the users are. For example, some people may be less sensitive to task distance, while others would be more sensitive to it.

As on-the-go crowdsourcing systems designers, we have to consider the trade-offs in possibly overburdening a small subset of people who are more likely to help with disrupting larger crowds who might not be able to help. In either case, there is a potential risk of losing helpers in the long run due to burnout or disruptions. Since the cost of disruption was considered low at the two ends of the distance spectrum where we were able to find potential helpers, there might exist a "goldilocks" zone where we can get enough potential helpers and low enough disruptions to mitigate overburdening or disrupting participants at both the individual and aggregate levels.

Limitations Due to the small sample size, we used a within-subject experimental design that may lead to fatigue and learning bias. Regarding the potential fatigue issue in a within-in subject design, we saw that two people who helped in the first week did not help in the second week. There was one dropout in each of the counterbalanced conditions (i.e., one in the in-range, the other in the at-location). For the other three participants, two people picked up more packages and one picked up same amount of packages compared to first week. Based on these empirical results, it appears unlikely that fatigue was an issue that influenced our results.

We also used the task pick-up rate as a measure for individual participation and the number of notifications sent for user recruitment, which served as an intermediary measure for system outcomes. In the future, we may conduct a large-scale, between subject, longitudinal study to explore actual number of pickups and timeliness of delivery to better measure individual participation and overall system outcomes, and we can further assess user attrition rate to gain insight into possible burnout that is caused by too much disruption.

Unlike other physical tasking services which use monetary incentives, we did not use payment per task as an incentive mechanism since it would be a confounding factor for understanding how people's participation varies as its convenience (i.e. distance to task locations) varies.

Another potential limitation of this experiment is the simulated package center, a coffee shop in a region where our participants frequent, which might have influenced people's willingness to help. Some people reported feeling awkward about picking up packages and not purchasing a coffee. In the future, we plan to conduct a field deployment with an actual college package center in order to eliminate such a deterrent to participation.

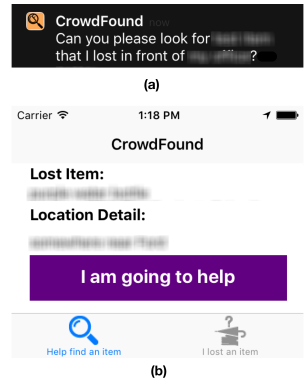

Figure 5: (a) CrowdFound sends notification to people who pass by possible lost item location; (b) the details of the lost item are also provided to the helpers.

Experiment 2: The effect of notification timing on most valued contributions

Methods We conducted a two-week long, within-subjects experiment in order to understand 1) how the particular timing of a notification can affect an individual's actions and the global system outcomes, and 2) to help reveal what the situational factors that affect an individual's search behavior.

To facilitate the experiment, we developed a prototype of CrowdFound, a mobile application where users can post lost item search requests that then notifies people who pass by possible lost item locations. A user who lost an item can post a request by providing a lost item type, a detailed description of the item, an approximate location where she believes she lost the item on a map, additional location details, and a picture of the item if she has one. When a potential helper is in the vicinity of the tagged location of a lost item, she receives a notification (Figure 5a) asking if she can help look for the missing item. Once she clicks the notification, she is shown the relevant description of the item (Figure 5b). If the user decides to help, she clicks "I am going to help" and a 30 second countdown is shown on the screen. After 30 seconds, a thank-you message is shown, along with the option to either select: "I found the item" or "I couldn't find it". If the user clicks "I found the item", an e-mail is sent to the requester informing her that her lost item has been found and connecting her to the helper so that she may retrieve her item. For this experiment, we set the default search time as 30 seconds and the diameter of the search region as 40 meters.

We used two notification conditions: in the *early notification* condition we notified the users 20 meters in advance of the target search region, and for the *at-border* condition we notified users immediately upon entering to the region. We chose 20 meters based on a calculation that added the mean

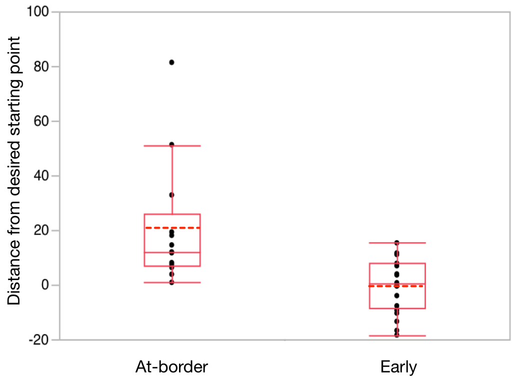

Figure 6: Distance from where they were supposed to start searching and where the participants actually searched in at-border and early notification. Means are shown as dotted lines and medians as solid lines. The box whiskers indicate range including outliers.

interaction time (7 seconds) for reading and opening notifications (numbers derived from a pilot study) to the mean time it takes for people to read the descriptions in the lost item description view (6 seconds, as derived from a usability testing). We then multiplied this sum by the average walking rate (1.4m/s) to set the early notification distance.

We recruited 15 people (8 male and 7 female) who had iPhone 5 mobile devices or above from undergraduate class and convenient samples. The participants' age ranged from 19 to 34 (mean: 23.33, SD: 3.64). They were randomly assigned to one condition and then switched to the second condition after one week. The order of conditions for each person was counterbalanced.

One of the authors served as a requester during the experiment, and the participants were asked to find lost items in both conditions. Each participant received 8 requests from among 5 different types of items (item types were based on actual posts on Northwestern University's lost&found Facebook group). The items ranged in size, but we tried to avoid bulky items which might be too obvious to spot. Descriptions of the item locations used similar prepositional phrases (e.g. somewhere, near, in front of) to those found in common descriptions posted on the lost&found group.

After the experiment, the participants were asked to rate statements on a 5-point likert scale about the perceived cost of disruption. Since we believe that there will be no difference in terms of the number of notifications sent between the two conditions and its effect on the perceived cost of disruption, we only sent out one post-study survey at the conclusion of the experiment. We also performed post-interviews with participants who searched for lost items and asked them about the times they helped or didn't help, how they responded to the notifications, what were their search strategies, and what features might have helped them to search.

In our analysis, we use distance from the actual search area to the intended search point as a measure of individual behavior. We use coverage of the search area as a way

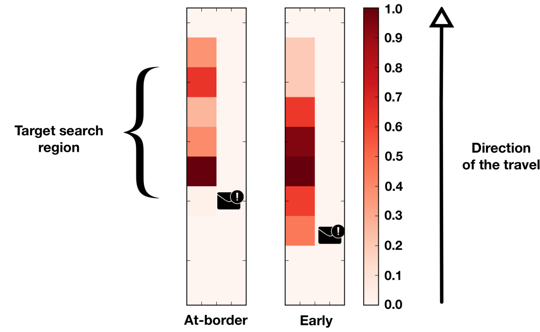

Figure 7: A heatmap shows the relative search distribution in each condition when participants were heading north in 30 seconds time window. In at-border condition people missed some of the areas at the beginning of the region, while in early notification some of the searches were wasted outside of the intended search region.

to measure system outcomes. Wasted efforts mean that the searches happened outside of the target search region, and missed opportunities means that helpers were in the target search region but did not cover some areas.

Results Our results show that small changes in notification timing have a significant effect on where individual searches occurred and they also affect the aggregate area covered by all participants.

Figure 6 shows the distance from where the participants started searching to the border of the target search region where they were supposed to search. The distance was closer to the desired search border in the early notification condition than in the at-border notification condition.[1] Wilcoxon test shows that there is a significant difference (Z=3.4, $p<.001$) between the at-border (mean: 20.55 meters, SD: 22.70) and the early notification condition (mean: -0.42 meters, SD: 10.18). This indicates that, on average, by the time the helpers searched in the at-border condition they had already passed almost half of the desired search region, while those in the early notification condition started 0.42 meters before the border of the desired search region.

Locations where the individual helper started searching also influenced the areas they covered within the region. In the at-border notification condition the participants missed some of the desired search areas located near the beginning of the target area, while in the early notification condition some people started searching too early and outside of the desired target region. Figure 7 shows the relative search distribution surrounding the target search region in each condition when participants were heading on a northerly trajectory. We could not clearly see the difference in the relative coverage distribution between the conditions when the participants were heading south, since there were only five searches in the at-border condition in total and we had to discard two searches due to GPS inaccuracy.

[1] We excluded 4 searches in which app open time exceeded 50 seconds from when they received notifications.

An individual's on-the-go interaction with the mobile devices also influenced their participation. Most of the people would open the app via notifications and read the details of lost items as they were walking. It took 21.62 seconds (SD: 15.12) on average from the time the notification was received to the time the participants started helping, and they walked an average of 22.41 (SD:18.75) meters. However, a few participants would stop and read the notifications because they were afraid that they might miss the region they were supposed to search and didn't want to circle back.

Although limited to our experiment, our results show that it was possible to find lost items with relatively little effort and little disruption on the part of the on-the-go helpers. At least six items (out of eight) were found by six different participants. It took 15.38 hours (SD: 17.22) on average from the time the item was posted until it was found. It also took 10.65 hours (SD: 11.31) on average to find an item since the time the first notification was sent to a helper. The helpers searched 38 times out of the 68 notifications sent to them, resulting in a 55.88% task acceptance rate. Also, the mean rating for the perceived cost of disruption on 5-point likert scale was 2.4 (SD: 0.91), indicating that the notifications were not disruptive (1 indicates not disruptive at all while 5 indicates very disruptive).

Discussion Our experiment with early and at-border notifications shows that small changes in the timing of a notification can have significant effects on values of individual actions and overall system outcomes. While early notifications permitted people to start their search for lost items without missing areas at the beginning, some of their searches happened prior to the intended search region. On the other hand, in the at-border condition, some people missed areas at the beginning of the region and as a result those areas were never covered.

Limitations Although in our experiment we only focused on individual searches within a single small region, in a real world setting the size of a lost item region may be large and it will require systems to break down a large search region into smaller regions to solicit help. In such cases, at a system-level, we need to reason about helper's future routes, value of asking help in a certain region, and coordinate searches across the regions in order to minimize the wasted efforts or missed opportunities to accomplish desired system outcomes. That being said, in the future, we could conduct a large-scale study with multiple lost item locations and a larger number of participants.

On-the-go situational factors

We now discuss some of the situational factors that influenced a helper's willingness to contribute. We identified on-the-go situational factors such as *little or no travel detour, having to walk back, and the availability of the worker's hands* that are not prominent in existing physical crowdsourcing systems.

Existing regular route, having to walk back, missing the right moment In lost and found searches, most of the people tended to search for items along their current routes and they mainly covered areas that correspond to those routes, and they were unwilling to walk back in the opposite direction of their destination. As one participant said: *"It just seemed more natural to look, or not natural, convenient for me to look in the direction I was already walking especially. Because it's usually when I am walking I have a destination so it doesn't make much sense to me to turn around and look in the other way I just came. It makes more sense to make that sweep when I am already walking that way."*

We found similar results in our package delivery settings. Participants mentioned that sometimes they missed the notifications or checked the notifications too late and didn't want to walk back and pick up packages. One participant said: *"Sometimes I couldn't feel the notification, and sometimes I checked the notification after passed the [task location], then I was like 'oh well...' and I kept walking."*

Factors like having to walk back served as a channel factor for participants, in which situations they were dissuaded to participate.

Hands Availability One of the main reasons that the helpers didn't pick up packages was that they couldn't carry the package. One participant mentioned that he had to bring both his lunchbox and a coffee so he couldn't deliver the package. Another participant said: *"I did not have hands to carry [so I didn't pick up]. Also one time, I had to put a package in my backpack [to help]..."* This finding showed that for physical tasks such as delivery it is important for systems to understand whether or not the potential helpers have enough hands to do the task.

Personal situation, time availability and temporal preference Participants' schedules and mood also influenced their willingness to pick up items. Sometimes they were in a hurry or they just didn't feel like doing it. One participant stated: *"It also depends on my schedule and mood. If I have a meeting with my advisor and get nervous, I wouldn't pick up even if I am there. If there is something important or you are nervous, you don't care about other things."*

In lost and found searches, the common reason for not being able to help was the lack of time. Participants were not willing to help or even pull out their phone to check the notification when they were rushing to classes or meetings. One participant said: *"Most of the time if I am rushing for class [and] I know I am already late, so I wouldn't check my phone."*

Although it differs individually, people have different time periods when they are not willing to help for various reasons. For instance, one participant stated that he never searched after class during lunch time, because he was hungry and didn't feel like helping. Another participant was not willing to search in the morning because there were many people heading to classes and work, and he didn't want to stop and block the crowds in the busy morning: *"I am always rushing to work in the morning...There are a lot of people walking, [so] it will be weird if I just stop and look around."*

Being with other people We found that whether or not participants were alone or with other people influenced their willingness to help. For some people, being with others motivated them to participate together: *"This was the time when I was looking for the gloves...I just looked around like the same route, but my friend who was more curious and went to the grass area...He looked around a little bit here and there, and he gave up...It wasn't about helping, it was more like a game that I should find something. It was very amusing to me, and it was a challenge, and I really wanted find this."* In contrast, some participants felt it was odd or socially unacceptable to search for items when they were with other people. One participant said: *"If I am walking with other people then I wouldn't, this is kind of weird."*

Weather conditions Current weather condition influenced not only participants' willingness to search but also their search time and search regions. Many participants didn't even want to check their phone because it was too cold on some of the days during the experiment period. Even if they were willing to help despite the weather condition, people's search behavior changed: *"Oh, one time it rained and this part [lawn area] was really muddy, and I didn't wanna go look in there. So I just kept looking on the sidewalk while I walked, but I didn't like go search for a minute."*

Discussion We found that situational factors such as time availability, the timing of the task, weather conditions, and a convenient physical location were all important situational factors, and this result is in line with previous findings in mobile crowdsourcing services (Teodoro et al. 2014). We also found that there are other on-the-go specific factors such as having to walk back or re-trace one's prior route, taking a detour that is in the opposite direction of the destination, and having one's hands available, are important situational factors for on-the-go helpers. Extending the work from Horvitz and Krumm (Horvitz and Krumm 2012), in which distance was taken into account for calculating the cost of diversion in opportunistic routing, in on-the-go settings we might account for individual differences and the situational factors identified in this study when calculating the cost of diversion and incorporate it into task suggestion decisions.

Design Implications

Our research seeks to broadly realize the vision of leveraging people's planned travel to create previously unavailable opportunities for physical crowdsourcing over large geographic areas (Sadilek, Krumm, and Horvitz 2013). In this paper we studied the challenges in designing task notification policies for on-the-go crowdsourcing systems. From two controlled experiments in package delivery and lost-and-found settings, we found that small changes in notification radius and timing can have a significant effect on individual participation and actions that in turn affect global outcomes. In this section, we discuss the implications of our findings on the design of future on-the-go crowdsourcing technologies and applications.

Technical Advancements

Focusing on a general population of people on-the-go necessarily implies that helpers may or may not perform tasks presented to them at any given moment. Ensuring that system goals are achieved thus requires managing the constantly changing levels of availability, attention, and interests of people on the move. This motivates the need to plan and execute actions *opportunistically* based on available mobility resources (Horvitz, Koch, and Subramani 2007; Kamar, Horvitz, and Meek 2008), and to make adjustments over the course of problem solving as partial solutions are made available (Malone and Crowston 1994). Future work can fill a void in the existing literature on task planning and routing (Hinds 2002; Shahaf and Horvitz 2010; Wooldridge 2001; Grosz and Kraus 1996) by providing flexible frameworks that will enable on-the-go crowds to achieve a wide range of tasks and objectives.

Since the precise means for supporting system goals with an on-the-go crowd is often unknown a priori, coordinating complex physical behaviors will rely upon dynamic monitoring and flexible utilization of resources. First, we need models for monitoring and predicting the supply of potential helpers built not only on our understanding of people's mobility routines but also their willingness to take on a task in a given situation, taking into consideration the costs of diversion and disruption as well as any channel factors that affect participation. Second, we need decision-theoretic frameworks that use such models to optimize and adjust task notification policies to best support desired individual, community, and system outcomes.

To balance individual disruption and the quality of service, we need a *supply management framework* for on-the-go crowdsourcing that can simultaneously reason over people, tasks, and time to opportunistically govern who to send tasks to based on needs and the situation on the ground. Based on the urgency of tasks and the availability of helpers, such a framework can help us reason about which tasks to suggest to whom so as to maintain a healthy pool of helpers by avoiding unnecessarily disrupting or overburdening helpers, while also being sensitive to system goals to ensure the timely completion of tasks.

To promote contributions where they are most needed, we need a *hit-or-wait framework* that can reason over tasking opportunities a person can encounter on their route and devise policies that decide whether to *hit* (i.e. send a task request) or to *wait* (i.e. to hold off on a request) for a better opportunity. Based on the importance of tasks and people's likelihood of reaching them and helping, such a framework can help us make informed decisions about when to suggest tasks to potential helpers en-route so as to best leverage individual efforts in ways that advance system goals.

Design Opportunities

On-the-go crowdsourcing can open up new ways for people to help others in their communities and neighborhoods. By surfacing task needs in situations where potential helpers are likely able and ready to help (e.g., in package pickup), on-the-go crowdsourcing can facilitate more connections between helpers and requesters to bring significant benefits through minimal efforts. By coordinating opportunistic contributions from willing helpers each contributing a small part to completing a larger task (e.g., in lost-and-found), on-the-go crowdsourcing can lower the barrier to participation and connect requesters to a larger population of helpers. Tapping into a large crowd of potential helpers and facilitating convenient help should make existing community support systems more effective and efficient. By scaling the efforts of an engaged community, on-the-go crowdsourcing can also spur new community-based services designed to help people in need (e.g., the elderly, people with disabilities, new moms, etc) that may otherwise be difficult to support.

On-the-go crowdsourcing can also open up new ways of providing commercial physical tasking services. First, on-the-go crowdsourcing can enable more flexible ways of working that fit into workers' existing routines and mobility. For example, this can allow a driver to pickup and drop off a passenger on their way to picking up a child from school. Second, on-the-go crowdsourcing systems can reduce travel costs to make affordable services that would otherwise have been prohibitively expensive to scale. This can enable new services that, by effectively coordinating convenient contributions from on-the-go workers, can provide a quality of service that matches or exceeds existing models but at just a fraction of the cost.

In order to realize these and other design opportunities, future work should seek to better understand the motivations required for people to contribute on-the-go and design solutions that address practical issues such as establishing trust. For example, there is a need to better understand how physical tasks can conveniently fit into people's routines in ways that enrich their lives. There is also a need for designs that assure requesters that tasks will be completed well and on-time, given a service that is supported with an on-the-go crowd instead of workers who respond on-demand. We look forward to designing new on-the-go crowdsourcing systems that build on people's motivations and that address such practical challenges in future work.

Acknowledgement

We thank members of the Design, Technology, and Research program and the Delta Lab for their valuable feedback and helpful discussions. This work was funded by National Science Foundation grant #1618096, a Microsoft Research FUSE Labs Award, and a Segal Design Cluster fellowship.

References

Aberer, K.; Sathe, S.; Chakraborty, D.; Martinoli, A.; Barrenetxea, G.; Faltings, B.; and Thiele, L. 2010. OpenSense: Open Community Driven Sensing of Environment. In *SIGSPATIAL International Workshop on GeoStreaming*.

Aoki, P. M.; Honicky, R. J.; Mainwaring, A.; Myers, C.; Paulos, E.; Subramanian, S.; and Woodruff, A. 2008. Common Sense: Mobile Environmental Sensing Platforms to Support Community. In *Ubicomp*.

Chen, C.; Cheng, S.-F.; Gunawan, A.; Misra, A.; Dasgupta, K.; and Chander, D. 2014. TRACCS: A Framework for Trajectory-Aware Coordinated Urban Crowd-Sourcing. In *HCOMP*.

Eisenman, S. B.; Miluzzo, E.; Lane, N. D.; Peterson, R. A.; Ahn, G.-S.; and Campbell, A. T. 2007. The BikeNet Mobile Sensing System for Cyclist Experience Mapping. In *SenSys*.

Fischer, J. E.; Greenhalgh, C.; and Benford, S. 2011. Investigating Episodes of Mobile Phone Activity As Indicators of Opportune Moments to Deliver Notifications. In *MobileHCI*.

Grosz, B. J., and Kraus, S. 1996. Collaborative plans for complex group action. *Artificial Intelligence* 86(2):269–357.

Han, K.; Shih, P. C.; Bellotti, V.; and Carroll, J. M. 2015. It's Time There Was an App for That Too:: A Usability Study of Mobile Timebanking. *IJMHCI* 7(2):1–22.

Hinds, P. 2002. *Distributed work*. MIT Press.

Ho, J., and Intille, S. S. 2005. Using Context-aware Computing to Reduce the Perceived Burden of Interruptions from Mobile Devices. In *CHI*.

Horvitz, E., and Krumm, J. 2012. Some help on the way: Opportunistic routing under uncertainty. In *Ubicomp*.

Horvitz, E.; Koch, P.; and Subramani, M. 2007. Mobile opportunistic planning: methods and models. In *User Modeling*.

Kamar, E.; Horvitz, E.; and Meek, C. 2008. Mobile opportunistic commerce: mechanisms, architecture, and application. In *AAMAS*.

Malone, T. W., and Crowston, K. 1994. The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)* 26(1):87–119.

Musthag, M., and Ganesan, D. 2013. Labor Dynamics in a Mobile Micro-task Market. In *CHI*.

Rodriguez Garzon, S., and Deva, B. 2014. Geofencing 2.0: Taking Location-based Notifications to the Next Level. In *Ubicomp*.

Ross, L., and Nisbett, R. E. 2011. *The person and the situation: Perspectives of social psychology*. Pinter & Martin Publishers.

Sadilek, A.; Krumm, J.; and Horvitz, E. 2013. Crowdphysics: Planned and Opportunistic Crowdsourcing for Physical Tasks. In *ICWSM*.

Shahaf, D., and Horvitz, E. 2010. Generalized task markets for human and machine computation. In *AAAI*.

Teodoro, R.; Ozturk, P.; Naaman, M.; Mason, W.; and Lindqvist, J. 2014. The Motivations and Experiences of the On-demand Mobile Workforce. In *CSCW*.

Thebault-Spieker, J.; Terveen, L. G.; and Hecht, B. 2015. Avoiding the South Side and the Suburbs: The Geography of Mobile Crowdsourcing Markets. In *CSCW*.

Vaish, R.; Wyngarden, K.; Chen, J.; Cheung, B.; and Bernstein, M. S. 2014. Twitch Crowdsourcing: Crowd Contributions in Short Bursts of Time. In *CHI*.

Wooldridge, M. J. 2001. *Multi-agent systems: an introduction*. Wiley Chichester.

Zimmerman, J.; Tomasic, A.; Garrod, C.; Yoo, D.; Hiruncharoenvate, C.; Aziz, R.; Thiruvengadam, N. R.; Huang, Y.; and Steinfeld, A. 2011. Field trial of tiramisu: crowdsourcing bus arrival times to spur co-design. In *CHI*.