NORTHWESTERN UNIVERSITY

Designing Flexible Coordination Systems to Advance Individual and Collective Goals in Physical Crowdsourcing

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Technology and Social Behavior

By

Yongsung Kim

EVANSTON, ILLINOIS

December 2020

ABSTRACT

Designing Flexible Coordination Systems to Advance Individual and Collective Goals in Physical Crowdsourcing

Yongsung Kim

Volunteer-based physical crowdsourcing systems connect individuals to make unique contributions to solve local and communal problems and enable new services. A key challenge in enabling such systems is attracting enough willing volunteers who can make useful contributions to achieve desired system goals. While most volunteer-based systems provide volunteers flexibility to attract more volunteers to make convenient contributions, it can be challenging to reach desired system goals with uncoordinated contributions. In contrast, other systems may direct volunteers to specific tasks to meet the desired system goals, but may fail to attract enough volunteers because they do not provide much-needed flexibility.

To overcome such challenges, this thesis introduces the idea of *flexible coordination* that combines the benefits of both approaches in providing flexibility and coordinating useful contributions. A flexible coordination system surfaces opportunities for volunteers to contribute that are within volunteers' routines and that are useful for achieving system goals. Unlike existing approaches that direct volunteers to go out of their routines or change their routines to meet the

desired system goals, a flexible coordination system allows volunteers to carry out their routines to maintain flexibility. In order to still collect useful contributions while maintaining flexibility, a flexible coordination system proactively suggests opportunities that are within volunteers' routines but that are also useful in advancing the desired system goals. Using the idea of flexible coordination, this thesis introduces *on-the-go crowdsourcing* systems that allow volunteers to just go about their days, focusing on their routines, and make convenient contributions that seamlessly fit into their routines but that are still useful for achieving desired system goals.

To enable the idea of flexible coordination and the design of on-the-go crowdsourcing systems, this thesis introduces three technical frameworks: (1) *Opportunistic Hit-or-Wait*, a decision-theoretic framework that surfaces opportunities for volunteers to make valuable, convenient, and coordinated contributions on-the-fly to improve the quality of service; (2) *4X*, a technical framework for multi-stage data collection processes that determine effective data collection strategies by reasoning about volunteers' dynamic changing state of interests and current knowledge about the world; and (3) *Opportunistic Supply Management*, a decision-theoretic framework that identifies and surfaces opportunities across the entire community in a way that can optimize the desired balance between the experience of volunteers and the goals of the system. Taken together, these frameworks demonstrate we can design volunteer-based systems that provide flexibility to volunteers and coordinate useful contributions to achieve globally effective outcomes by following volunteers' routines and surfacing opportunities at opportune moments when needs of volunteers align with that of a system.

Acknowledgements

I am grateful to many people for their support of my research. I would like to thank:

- my advisor Haoqi Zhang;
- my dissertation committee members Darren Gergle, Aaron Shaw, and Eric Horvitz;
- my mentors and collaborators Liz Gerber, Adam Fourney, Ece Kamar, Daniele Quercia, and Luca Aiello;
- my friends and colleagues in the BBQ Special Interest Group: Josh Hibschman, Gobi Dasu, Harrison Kwik, Kapil Garg, Leesha Maliakal, and Ryan Louie;
- members in the On-the-go Crowdsourcing Special Interest Group at Northwestern;
- the community and friends in the Delta lab and DTR at Northwestern;
- friends and colleagues in the Technology and Social Behavior program at Northwestern;
- Heavanston friend;
- and my parents Yoon Ho Kim and Kyung Jin Ko and my sister Myung Eun Kim.

Table of Contents

ABSTRACT 3

Acknowledgements 5

Table of Contents 6

List of Tables 9

List of Figures 10

Chapter 1. Introduction 17

1.1. Thesis and Contributions 22

1.2. Thesis Overview 29

1.3. Reader's Note 30

Chapter 2. Opportunistic Hit-or-Wait Framework 31

2.1. Introduction 31

2.2. Background 34

2.3. Coordinate On-the-go Crowds with Hit-or-Wait 37

2.4. Study 1: Simulation 45

2.5. Study 2: Field Deployment 52

2.6. Follow-up Interviews 61

2.7. Discussion and Future Work 64

Chapter 3. 4X Framework 68

3.1. Introduction 68

3.2. Background 72

3.3. 4X Framework 77

3.4. LES: Campus Data Collection with 4X 83

3.5. Study 1: Comparing 4X to Opportunistic and Directed Approaches 91

3.6. Study 2: 4X Yields Additional Contributions 101

3.7. Discussion and Future Work 110

Chapter 4. Opportunistic Supply Management Framework 117

4.1. Introduction 117

4.2. Related Work 121

4.3. Opportunistic Supply Management: A Framework for Community-Level Flexible Coordination Mechanisms 125

4.4. Simulation Study 134

4.5. Study 2: Field Deployment 143

4.6. Discussion 152

4.7. Conclusion 158

Chapter 5. Discussion 160

5.1. Flexible Coordination Revisited 160

5.2. A Mixed-Method Approach for Designing and Deploying Flexible Coordination Systems 172

5.3. Ethics 173

Chapter 6. Conclusion 178

References 180

List of Tables

3.1 Breakdown of contributions and notifications sent by condition in Study 1. 98

List of Figures

2.1 An illustrative example for demonstrating how Hit-or-Wait makes hit or wait decisions. The user is currently in the subregion A and is equally likely to go to the subregion B and the subregion C, or $P(B|A) = P(C|A) = 1/2$. We assume that the item is equally likely to be each subregion. The value of notifying the user in each subregion is the likelihood that the item is there following n searches, $P(item|n)$. Using the Bayes rule, we can compute $P(item|0) = 1/3$, $P(item|1) = 1/5$, $P(item|2) = 1/9$. 41

2.2 On-the-go crowdsourcing architecture. 43

2.3 Simulation results comparing the overall quality of search for node counting, Hit-or-Wait, and myopic optimal. The results show that Hit-or-Wait outperformed the node counting algorithm and approached the performance of the myopic optimal solution. 48

2.4 Percentage of missed opportunities for node counting, Hit-or-Wait, and myopic optimal solution. Hit-or-Wait algorithm missed 46.07% of opportunities (SD: 14.49) while node counting missed 78.68% of opportunities (SD: 10.53). 49

2.5 Effect of model accuracy on Hit-or-Wait performance in situations of uniform neighboring values and varied neighboring values. 51

2.6 Trouve, a prototype lost-and-found app for the study. (a) A user receives a notification indicating a lost item is nearby; (b) the user sees the details of the lost item; (c) after 30 seconds of clicking I am helping now, the user is prompted with a survey question about perceived disruptiveness of the 30-second search. 53

2.7 Responses from EMA and post-study survey about perceived disruption for 30-second searches on a 5-point Likert scale (1: not disruptive at all, 5: very disruptive). 57

2.8 Overall quality of search between Hit-or-Wait and myopic optimal solution. Hit-or-Wait was able to make near optimal decisions in actual use, and make efficient use of people's search efforts. 58

2.9 Comparison between the value gained if hit at the wait decision state vs. value gained from waiting and notifying at the later state. 59

2.10 Examples of missed opportunities after wait decisions: (a) taking shortcuts and did not go to adjacent states; (b) staying at the wait state (e.g., taking classes); (c) missing GPS location data. 60

2.11 Wait (top) and Hit (bottom) example visualizations. 62

3.1 The 4X framework scaffolds data in four stages. People first contribute opportunistically, marking new landmarks for tracking and responding when queried for information as they pass existing landmarks (eXplore). Collected data then draws people to target locations outside of their path, where they can be queried for additional contributions (eXpand). People may also be queried en route to the target location for contributions (eXploit). As data at a location fills (indicated by the size and the darkness of circles), users are directed to other regions and locations in need of contributions (eXterminate). 77

3.2 LES interfaces for contributing data and for receiving data of interest. Users respond to simple queries about locations they are in or are about to pass (a), and receives valuable information about nearby places of interest that they care about (b). All responses can be made through contextual notifications on the device's lock screen without needing to open the installed application. 84

3.3 Illustration of the notification policies for the Opportunistic, Directed, and 4X conditions in Study 1. In all conditions, users receive at-location notifications that ask them to contribute additional information when they are about to pass by a tracked task location. Users in the Directed and 4X conditions also receive at-distance notifications up to 300 meters away from the task location. The Directed condition sent these notifications regardless of whether preferred information is available (information is included when available), whereas the 4X condition only sent these notifications if preferred information is available. A star on a person indicates that preferred information of interest is available for that user at the time of notification. Finally, 4X also asks users to make contributions en route when they decide to go to the target location. 93

3.4 Percentage of users' contributions that built up data fidelity by scaffolding together multiple contributions beyond the first. 31.77% of all contributions made by users increased data fidelity beyond the first contribution (e.g., private tables available at a coffee shop), leading to higher-fidelity data from the scaffolded low-fidelity contributions (e.g., private tables near outlets available at a coffee shop). 97

3.5 Notification and interest preferences for Study 2. In this example, the participant wishes to be notified about private seating at coffee shops only when they are near windows (a), and reports a stronger interest in seating near windows than near outlets (b). 103

3.6 Total number of contributions in Study 2 by 4X stage. 4X yields 49.09% more data from directing users to locations of interest via eXpand and eXploit over data collected opportunistically via eXplore. 105

3.7 Users were more likely to go out of their way to locations when presented with information that is more valuable to them. 107

4.1 Each volunteer-based system has to manage the balance between system efficiency and the needs of volunteers, and each system may have their own desired balance. The opportunistic supply management framework allows the system to optimize the desired balance between system efficiency and the needs of volunteers by reasoning about different community-scale task notification policies that opportunistically decide who to send tasks across the community. As volunteers may or may not become available or accept tasks, supply management simulates over the possible unfoldings of adopting a task notification policy and chooses an optimal policy that can best achieve the desired goals. 118

4.2 An architectural diagram of the opportunistic supply management framework for modeling and optimizing the choice of task notification policies. The framework consists of (1) *models* of notification policies, volunteers (i.e. supply and pickup), and values; (2) the *simulator*, which uses the models to simulate and evaluate the effectiveness of policies; and (3) the *decision manager*, that determines the optimal policy. 128

4.3 An example workflow of the opportunistic supply management framework for a given policy. The supply management framework simulates over possible unfoldings of how many people may show up and who will accept or decline tasks. Given the possible outcomes and the intended goals, the supply management framework evaluates the value of enacting the policy and compares across policies to choose the optimal one. 131

4.4 Value functions for disruption and task completion. (a) presents the disruption value function that computes the disruption value as a function of the number of notifications being sent. We set the disruption value function so that the disruption value decreases drastically when more than 15-20% of users are notified. (b) presents the task completion value function that computes the task completion value as a function of the hour of the day at which the task is being completed. We set the task completion value function so that the task completion value decreases across four time-blocks at the same decreasing rate. 136

4.5 Supply management chooses the best policy that achieves the desired goals in balancing system efficiency and the needs of volunteers, while other fixed policies, the at-location policy (blue line) and the at-distance policy (orange line), would make trade-offs depending on the changing demand. Supply management (highlighted purple line) always chooses the policy with the highest expected value based on the changing demand. For example, supply management chooses the at-location policy when the number of task demand is less than 5, and chooses the at-distance policy when the number of task demand is more than 5. 139

4.6 Supply management makes decisions that can maximize different outcomes that different types of community may care about. Shown are dimensions of disruption as volunteer needs (blue bar) and task completion as system goals (orange bar) being prioritized for three different types of community, namely (1) volunteer-focused, (2) balanced, and (3) system-focused. While setting a value function that is system-goal focused leads to higher task completion value when compared to a value function that is balanced or volunteer-focused, the value function that is system-goal focused also leads to highest disruption (i.e. lowest value). 140

4.7 Supply management makes different decisions depending on varying demands and the goal orientation of the community, namely volunteer-focused, balanced, and system-goal focused. In this simulation, supply management chooses the at-location policy (blue) for the volunteer-focused community regardless of demands because it cares so much about disruption. In contrast, supply management chooses the at-location policy until demand is 2, and starts choosing the at-distance policy (orange) to get things done faster for the system-goal focused community. 141

4.8 For varying degree of noises in the model, the percent value gained via dynamic adjustment with regard to the value gained with non-adjustment with perfect model. 143

4.9 Apporte: a peer-to-peer delivery mobile application. (a) an example screenshot of a task notification that a potential volunteer receives when they pass by a pick-up location. (b) an example screenshot of an app when a potential volunteer clicks a notification. The potential volunteer can see the details of a request and can click “Accept” if they decide to help and click “Decline” otherwise. The volunteer can also text the requester if they need to coordinate with the requester. 144

4.10 Pick-up location (a) and Drop-off location (b). 148

4.11 Expected values of each policy computed with actual, realized supplies. Supply management chose the at-location policy during the deployment and the results show that the average value of using the at-location policy is higher than using the at-distance policy in practice. Error bars indicate standard error of the mean. 149

CHAPTER 1

Introduction

The growth of mobile devices in recent years has helped to bring about both commercial and volunteer-based physical crowdsourcing systems [**9, 126, 124**] that motivate large numbers of people to provide new and improved physical tasking services. In commercial systems, workers provide rides (Uber, Lyft), deliver groceries or meals (Instacart, Postmates, DoorDash), complete errands (TaskRabbit), and walk dogs (Wag). In volunteer-based systems, interested volunteers collect data at a large scale to help scientists conduct scientific inquiries (e.g. studying bird migration patterns) [**120**], citizen journalists gather information to help news organizations to cover smaller local events [**3, 130**], and citizens report infrastructural issues or city problems (e.g. broken street lights) to help local government agency track and fix the problems [**78**]. By attracting and coordinating crowds, these systems enable new services and solve problems that would not have been possible before.

This dissertation focuses on a core challenge that volunteer-based systems must address: attracting *enough volunteers* who can make *useful contributions* to achieve desired system goals. Without recruiting enough volunteers, the volunteer-based systems can't provide artifacts or services to attract people to use the systems. Prior research shows that many volunteer-based systems fail due to the lack of participation [**8**]. Even when the systems have enough participation, the systems can't achieve their desired goals if contributions are not useful in meeting the system needs [**140**]. For example, tasks that are appealing to volunteers may not be the ones that the system needs most help, and the system often suffers from poor quality of services as the result of the mismatch between individual needs and the system needs.

Most volunteer-based physical crowdsourcing systems seek to attract volunteers by providing flexibility to volunteers in deciding when and which task they contribute to [**53, 46, 9**]. This flexibility allows volunteers to meet their own needs, for example contributing to tasks that suit their schedules and routines, that are convenient for them, or that are of interests to them. But the flexibility provided to volunteers and the opportunistic nature of their contributions collected can make it hard to meet the desired system goals. For example, a volunteer-based lost-and-found service may ask community members to look for a lost item anywhere they'd like along their route, but this ease and convenience makes it difficult for the service to ensure good search coverage.

In contrast, other systems may direct volunteers to specific tasks to meet the desired system goals [**98**]. However, such a coordinated approach may fail to attract enough people because it does not provide much-needed flexibility to volunteers as the systems coordinate contributions under the assumption that volunteers are committed to participation. For example, prior work in physical crowdsourcing focuses on optimizing effective task assignments to minimize deviations from volunteer routines while maximizing system efficiency, which largely assumes that individuals will accept tasks when asked [**98, 23, 68**]. But these prior techniques break down in real-world settings where any given volunteer may or may not go near task locations, and may or may not accept the tasks.

To address the core challenge in attracting enough volunteers who can make useful contributions to achieve desired system goals, this thesis introduces the idea of *flexible coordination* that combines the benefits of both approaches to provide flexibility and coordinate useful contributions. A flexible coordination system surfaces opportunities for volunteers to contribute that are within volunteers' routines and that are useful for achieving system goals. Unlike existing approaches that direct volunteers to go out of their routines or change their routines to meet the desired system goals, a flexible coordination system allows volunteers to carry out their routines and activities as they wish to maintain flexibility. In order to still collect useful contributions while maintaining flexibility, a flexible coordination system proactively suggests opportunities that are within volunteers' routines but that are also useful in advancing the desired system goals. By doing this, a flexible coordination system provides the necessary flexibility to volunteers while still achieving the desired system goals with well-coordinated opportunistic contributions.

Using the idea of flexible coordination, this thesis introduces *on-the-go crowdsourcing* that allows people to just go about their days, focusing on their routines, and make convenient contributions to physical crowdsourcing tasks that seamlessly fit into their routines but that are still useful for achieving the desired system goals. For example, a community-based lost-and-found service may only ask community members to look for lost items along their existing route, but still try to ensure good search coverage by controlling when and where to ask for their help. A community-based package delivery may only ask community members to deliver packages along their existing routes, but still try to meet both the needs of volunteers in not being over-disrupted or over-burdened and the needs of systems in delivering items in a timely manner.

Existing physical crowdsourcing systems are limited in that they either passively wait for volunteers to complete tasks opportunistically—which can lead to many missed opportunities that lead to poor quality of services (e.g. poor search coverage in lost-and-found setting)—or directly assign tasks that best address system needs (e.g. searching for a lost item in a region where no one searched)—which require strong incentives to volunteers who may have to go out of their way to complete tasks. In contrast, on-the-go crowdsourcing systems send task notifications in situations where volunteers are likely able to contribute along their existing routines to tasks that are valued by the system. For example, an on-the-go lost-and-found system may follow along a user's route, predict their future trajectories, and proactively send a task notification only in a region where their contribution is most valued by the system in ensuring good search coverage to find a lost item.

To realize the idea of flexible coordination and the design of on-the-go crowdsourcing systems, we need to address two core challenges. First, existing solutions for task optimization and task assignment optimize over a fixed set of opportunities, but we do not know a priori which set of opportunities may become available. Existing approaches can optimize over a fixed set of opportunities only because they largely assume that systems can prescribe or pre-determine what each individual must do by directing or changing people's routines. In contrast, we have to recognize that there is no fixed set of opportunities because opportunities may dynamically arise depending on how a person's routine is carried out, and how that in turn changes people's future trajectories, interests and availability in helping.

Second, the quality of opportunities is relative to an individual's and other's routines, and how our knowledge of the world may change; therefore, evaluating the quality of an opportunity in isolation is ineffective. For example, opportunities are good relative to other opportunities that may arise within an individual's routine. In the lost-and-found example, a system may only look at where a user is right now and evaluate the value of their search in the current region. However, evaluating the quality of an opportunity in isolation is ineffective because it does not consider a user's entire routine, and other opportunities that may arise within the routine (e.g. people's future trajectories and possible opportunities arise in future locations). This can prevent us from recognizing other opportunities that may arise and that those potentially available opportunities may be better than the current opportunity.

Likewise, opportunities are good relative to other opportunities that may arise *across multiple people's routines*; therefore, evaluating the quality of an opportunity within a single person's routine may be ineffective. The success of achieving desired collective goals is dependent on many people's engagement; for example, delivering all packages in a timely manner to community members requires a group of people's participation. Therefore, evaluating the quality of an opportunity within a single person's routine (e.g. evaluating when and whether the user is willing and able to deliver a package) is ineffective because the quality of an opportunity to achieve desired collective goals should be evaluated based on the uncertain engagement and availability of all potential volunteers. For example, in a community-based package delivery service, a system may need to understand, across the population of potential volunteers, how many people may become available will pass by the package center, who will be willing to help when being asked, and whether or not all items will be delivered in a timely manner.

To address these challenges, our technical approach builds an understanding of how people's routines may unfold, reasons about opportunities that may become available within their routines, and sets conditions under which to surface an opportunity to contribute by comparing across possible opportunities that may arise within or across people's routines. Instead of optimizing over a fixed set of opportunities, which we do not know a priori, our technical approach optimizes over immediate and possible situations people might be in without ever prescribing or pre-determining what each person must do. Instead of evaluating the quality of opportunities in isolation, our technical approach evaluates the quality of opportunities by taking into consideration the uncertainty in people's routines, reasoning across possible opportunities that may arise within or across people's routines, and considering the expected value of opportunities in achieving individual needs and collective goals.

In doing this, our technical approach takes inspiration from Horvitz's work on flexible computation [**55, 61**] to reason about computational strategies under scarce resources that can achieve optimal outcomes under uncertainty. As flexible computation treats computation power as a scarce resource, by analogy, our technical approach treats volunteer attention and contributions as scarce resources, which differ from computation power in that individuals may have varying availability and willingness to help depending on their situations, may or may not decide to help when they are asked to help, may want to be minimally disrupted, and have dual needs in quality of services as a requester and user experience as a volunteer. To preserve the opportunistic nature in participation and meet the desired goals, our technical approach reasons about uncertainty in availability and participation and desired goals, and surfaces task needs at opportune moments when volunteers can conveniently contribute to meet the goals. Instead of assuming future situations are fixed, our technical approach generates strategies that are custom-tailored to varying situations where potentially available resources are uncertain, and that can still provide optimal solutions to reach the desired outcomes with available resources.

1.1. Thesis and Contributions

My thesis statement is:

> By building an understanding of how people's routines may unfold, reasoning about opportunities that may become available within their routines, and setting

> conditions under which to surface an opportunity to contribute within people's routines, we can design *flexible coordination* systems that provide flexibility to volunteers and coordinate useful contributions to achieve globally effective outcomes.

The core contribution of this thesis is the idea of *flexible coordination* and three technical frameworks that address the general challenges in enabling the idea of flexible coordination and the design of *on-the-go crowdsourcing* systems. The sections below highlight the contribution of each technical framework.

1.1.1. Opportunistic Hit-or-Wait Framework

Flexible coordination allows us to advance the design of physical crowdsourcing systems by enabling volunteers to contribute within their existing routines in a way whereby their contributions are convenient and also useful for achieving desired system goals. To do this, systems need to proactively send task notifications in situations where volunteers are likely able to contribute along their existing routines to tasks that are valued by the system. However, as people go about their days, we do not know a priori which opportunities may arise within their routines because of the uncertainty in their future trajectories. Therefore, we cannot use existing task assignment or optimization techniques that optimize over a fixed set of opportunities by prescribing or pre-determining what each individual must do that might be outside of their routines or that might require them to change their routines. Instead, we need to build up an understanding of a user's routine, reason about which opportunities may arise within a user's routine, and evaluate the quality of opportunities that may arise when surfacing an opportunity to a user.

To determine good opportunities when people can make useful contributions towards the system goals while following one's changing mobility patterns, we need to address a core technical challenge in managing trade-offs between taking a current opportunity and waiting for a better opportunity in the future. While the current opportunities are known at any given moment, the uncertainty in a user's future trajectories makes future opportunities uncertain. As a result, we may wait for the best match of a user to a task, only to find that the user never comes near such tasks in their realized routes. We may also be overly eager to capture current opportunities, sending a user the first task they come across, only to realize that adopting this strategy may require us to give up on better matches that become available later.

In order to identify good opportunities within a user's routine, our core idea is to evaluate the quality of an opportunity not in isolation but by understanding how a user's routine may unfold, what opportunities may thus arise, and how waiting for those opportunities may compare to taking a current opportunity. To realize this idea, we introduce *opportunistic Hit-or-Wait*, a general decision-theoretic mechanism that intelligently controls decisions over when to notify a person of a task among many tasks that they can contribute to along their existing routes, in ways that reason both about system needs across tasks and about a user's changing patterns of mobility. To stay within a user's routine while still eliciting useful contributions, Hit-or-Wait follows a user's locations, predicts their future trajectories, and models system decisions about whether to send a task that is near a volunteer now or to wait for a better opportunity in the future. Rather than surface an opportunity based on the quality of the current opportunity without considering what situations the person may later come across (e.g. solely based on how many people have looked for a lost item in the current region), Hit-or-Wait continually reasons about the tradeoffs between current and future situations the user may come across (e.g. where else might the person come across and how well searched those other areas may be), and make decisions about whether to surfacing a current opportunity or waiting for a future opportunity. This allows us to incorporate what we know about how the person's routine may unfold and use that information to make on-going tradeoffs on when and which opportunity to present. In other words, we avoid being over eager and avoid missing opportunities by explicitly reasoning about how a current opportunity compares to possible future opportunities.

1.1.2. 4X Framework

While taking into consideration the unfolding of future trajectories allows us to better understand which opportunities may arise within a user's routine, there are other factors, such as our changing knowledge of the world, that can also affect which opportunities may arise within a user's routine and how we should evaluate the quality of opportunities.

Consider mobile and physical crowdsourcing systems that engage volunteers to report data about the dynamically changing state of the world to help understand it and to enable new services (e.g. eBird [**120**], See Click Fix [**1**]). In such systems, volunteers' willingness to go to places and contribute data not only depends on their current locations but also depends on which information is available to them. For example, people may be willing to contribute data opportunistically if they are already at a location and can conveniently contribute data. However, people's willingness to go to a target location that is outside of their routine to contribute additional data may depend on what information is currently available and whether the information is of their interests. For example, notifying people interested in events such as free food or guest lectures may inspire some people to go; while there, they may be willing to contribute additional information about what food is available and what the lecture is about.

However, existing approaches are limited in how they strike the balance between the needs of data contributors (e.g. data collection tasks to be not disruptive and personally relevant) and the data collection goals (e.g. data coverage). On one hand, allowing people to actively contribute data along their existing routines when it is convenient for them to do so can better meet the needs of data contributors, but the opportunistic nature of the contributions makes it difficult to meet specific data needs (e.g. high data coverage) since data is only collected opportunistically along one's existing routine. On the other hand, directing people to fulfill a specific task that is not necessarily in their immediate vicinity can meet specific system goals such as increasing data coverage wherever it is needed, but it requires people to deviate from their existing routines and may require high incentives, such as monetary rewards that might be cost-prohibitive, to offset the disruption.

To overcome such shortcomings, we introduce a new hybrid approach that collects data opportunistically and uses the collected data to selectively notify people based on our understanding of the world and their interests. A core idea is to progressively build up our understanding of the world, and use that understanding to notify more people about opportunities that are still within their interests and convenience. With this approach we can simultaneously scaffold people's interests while building up our knowledge of the world further. Unlike existing approaches that direct people out of their routine to meet specific data collection goals, which can cause disruption to people, our approach offsets the "cost" of deviation with the "value" of the personally-relevant information. Unlike existing approaches that only elicit opportunistic contributions, which may fail to meet desired system goals, our approach directs people out of their way with personally-relevant information and captures opportunities en-route or at target locations. This allows systems to become more directed or remain opportunistic depending on our refined understanding of the world.

To realize this idea, we introduce *4X*, a technical framework for multi-stage data collection processes that determine effective data collection strategies by reasoning about dynamically changing state of the world, people's locations, and their willingness to deviate from their routine based on the current knowledge of the world. To do this, 4X models people's interests in information about the world; understands how the current state of the world matches their interests—which in turn affects which opportunities may become available; makes decisions about which data collection opportunity to surface in a way that does not over-extend their interests but that is still useful for gathering more data. For example, 4X first collects low-effort, low-fidelity opportunistic contributions when no data is available and when a user is passing by a location where they can conveniently contribute. 4X then draws other users to places outside of their routines where the data and their interests align, and while they are en route, 4X elicits further contributions at a place where it needs more contributions to ensure high data coverage. Instead of using a single data collection strategy regardless of situations on the ground, 4X reasons about how and when to enact certain data collection strategies based on the changing state of our knowledge of the world, people's interests and locations in a way that can simultaneously achieve both needs of data collectors and system goals.

1.1.3. Opportunistic Supply Management Framework

While Hit-or-Wait and 4X evaluate the quality of opportunities within a single individual's routine, opportunities can be good relative to other opportunities that may arise across multiple people's routines. For example, in a community-based peer-to-peer delivery service, delivering all packages in a timely manner to community members requires a group of people's participation; therefore, rather than evaluating the quality of an opportunity with a single person's routine (e.g. evaluating when and whether the user is willing and able to deliver a package), we need to evaluate the quality of an opportunity based on the uncertain engagement and availability of all potential volunteers. This may require the system to reason about how many users and when to engage them with opportunities to deliver items in a way that is still within people's routines and that considers changing availabilities and willingness of people.

A core challenge in finding good opportunities across the entire community is to reason about the uncertainty in engagement across the community. Existing technical solutions largely assume that people are available and will participate when coordinating contributions across the community to meet desired system goals. This takes away the flexibility for people to conveniently contribute through their routines that we are trying to preserve with flexible coordination. However, by maintaining the flexibility, we cannot be sure who will be available to contribute and whether they will actually contribute, thereby making it hard to know if a policy—a set of conditions that determine when and whom to notify of tasks—would be effective without knowing how it might, across a community of volunteers, lead to good outcomes that are aligned with the goals of the community. For example, an effective system must manage the tradeoffs imposed by being too aggressive in recruitment—which can be overly disruptive and result in a low task pickup rate—and being too restrictive in recruitment— which can involve too few volunteers, overburden the ones that are involved, and leave a disproportionately large number of task demands unfulfilled.

To overcome this challenge, we introduce *Opportunistic Supply Management*, a general decision-theoretic framework for modeling and optimizing the choice of task notification policies that find opportunities across the community to meet the needs of volunteers and system efficiency. Supply management follows community members' routines and integrates models that describe how task notification policies affect the available supply of volunteers and their likelihood to accept tasks, and how that in turn affects system efficiency and the needs of volunteers. Using these models, Supply Management simulates the possible outcomes that may result from adopting a task notification policy and chooses an optimal policy for a given situation (or set of situations) that best achieves intended system goals and desired volunteer experiences in expectation. With this approach, Supply Management can reason about how the world will unfold, take into consideration people's availability and willingness to help, and devise custom-tailored strategies that adapt to changing situations without ever imposing on what each individual must do. Unlike existing task assignment solutions that only consider how each individual can best contribute to the system, supply management considers how to leverage volunteer efforts across the community to best meet system goals in ways that still ensure good volunteer experiences by not overburdening or disrupting potential volunteers.

1.2. Thesis Overview

- Chapter 2 introduces *Opportunistic Hit-or-Wait*, a decision-theoretic framework that surfaces opportunities for volunteers to make valuable, convenient, and coordinated contributions on-the-fly to improve the quality of service.
- Chapter 3 introduces *4X*, a technical framework for a multi-stage data collection processes that determine effective data collection strategies by reasoning about people's dynamic changing state of interests and current knowledge about the world.

- Chapter 4 *Opportunistic Supply Management*, a decision-theoretic framework that identifies and surfaces opportunities across the entire community in a way that can optimize the desired balance between the experience of volunteers and the goals of the system.
- Chapter 5 discusses implications and principles for flexible coordination, as well as the generalizability and limitations of flexible coordination.
- Chapter 6 reviews the contributions of the thesis and proposes a vision for the future of flexible coordination.

1.3. Reader's Note

Throughout this dissertation, we will reference existing approaches that provide flexibility for people to contribute as they wish as *opportunistic* approaches, and approaches that coordinate contributions without providing flexibility as *directed* approaches.

CHAPTER 2

Opportunistic Hit-or-Wait Framework

We consider how the idea of flexible coordination may enable on-the-go crowdsourcing systems that allow volunteers to make convenient contributions within their routines but that are still useful in achieving desired system goals. In this chapter, we discuss the core challenge in evaluating the quality of opportunities that may arise within a person's routine to surface good opportunities to the user.

2.1. Introduction

The growth of mobile devices in recent years has helped to bring about mobile [**49, 34, 4**] and physical [**9**] crowdsourcing systems that help connect people to solve local, communal problems. In these mobile and physical crowdsourcing systems, people make small contributions toward a larger collective problem, such as tracking animal species or air quality for citizen science projects or providing rides or delivering packages in commercial applications. In these systems, opportunistically relying on people to do convenient parts of the problem often leads to incomplete solutions [**51, 126**]. For example, volunteer-based time banking systems may complete only a fraction of the tasks requested, even days after the requests [**51**]. Yet, directing people to do inconvenient tasks decreases their willingness to complete them and therefore requires higher incentives. For example, tasks can require significant travel that strongly decreases people's willingness to complete tasks [**126**].

To overcome such shortcomings, we use the idea of flexible coordination to enable *on-the-go crowdsourcing* as an alternative model for enabling people to make convenient contributions that are within their existing routines but are nevertheless effective in achieving desired system goals [**114, 32, 77**]. For example, a community-based lost-and-found service may only ask community members to look for lost items along their existing route, but still try to ensure good search coverage by controlling when and where to ask for a user's help. By following people's routines and notifying people of tasks when they are likely able to help, we can increase people's willingness to participate and reduce the need to incentivize people. Unlike existing commercial services where workers are mostly available on-demand, on-the-go crowdsourcing systems attempt to follow people's changing state in their routines and surface opportunities as they become available and when they can best contribute towards the desired system goals. Adopting the idea of flexible coordination, On-the-go crowdsourcing thus attempts to make effective use of every potential volunteer toward a collective goal, while using only people's existing mobility and notifying them of tasks on the way that they can best help with. As such, it aims to achieve much of the benefits of explicit coordination but without requiring volunteers to go out of their way, to actively seek out tasks, or to reason about which task they should contribute to.

To determine good opportunities when people can make useful contributions towards the system goals while following one's changing mobility patterns, we need to address a core challenge of evaluating the quality of opportunities in comparison to other opportunities that may arise within a user's routine. In practice, among the many tasks a user may encounter during their routine, deciding which one to notify them about directly affects which tasks are completed and what outcomes are prioritized. Given the uncertainty in participation and potential volunteer's future trajectories, it is not at all obvious when to engage a volunteer with which opportunity to best leverage their efforts. One may pre-determine which opportunity to surface to a user based on a set of criteria. For example, in a lost-and-found setting, one may determine which task to surface to a user that is within their routine based on how well a region is being searched. However, by pre-determining when to surface an opportunity, we may wait for the best match of a volunteer to a task (e.g., where the region is less-searched and thus the user's search effort is highly valued), only to find that the person never comes near such tasks in their realized routes. We can also be overly opportunistic, sending a person the first task they come across, only to realize that adopting this strategy while aiming to avoid over-disruption may require us to give up on better opportunities that become available later. This illustrates a general challenge for flexible coordination whereby effective coordination relies on not only evaluating opportunities in isolation, but careful consideration of the trade-offs between presenting current opportunities and waiting for possible future opportunities whose availability is dependent on how a person's routine may unfold, and not under the control of the system.

To resolve this challenge, we introduce *Opportunistic Hit-or-Wait*, a general decision-theoretic framework that intelligently controls decisions over when to notify a person of a task among many tasks that they can contribute to along their existing routes, in ways that reason both about system needs across tasks and about a volunteer's changing patterns of mobility. To stay within people's routines while still eliciting useful contributions, Hit-or-Wait follows a user's current locations and predicts their future trajectories, and reasons about which opportunities that may unfold within a user's routine when making decisions about whether to send a task that is near a user now or to wait for a better opportunity in the future. Rather than surface an opportunity based on the quality of the current opportunity without considering what situations a user may later come across (e.g. solely based on how many people have looked for a lost item in the current location), Hit-or-Wait continually reasons about the tradeoffs between current and future situations a user may come across (e.g. where else might the person come across and how well searched those other areas may be), and make decisions about whether to surface the current opportunity or wait for a future opportunity.

We evaluate Hit-or-Wait both in simulations and in a field deployment in the context of community-based lost-and-found, where we use Hit-or-Wait to indirectly coordinate people to look for a lost item on their way by recruiting individuals to collectively search for the item across small subregions (i.e. tasks) they may encounter. In simulations, we found that Hit-or-Wait significantly outperforms our baseline task notification strategy and approaches the performance of the myopic optimal solution which has full knowledge about future trajectories. In a field study with 25 participants, we found that Hit-or-Wait coordinated small, opportunistic contributions to achieve globally effective solutions by minimizing disruptions and maximizing the value of individual contributions. In other words, Hit-or-Wait was able to follow and stay within a user's routine while still eliciting contributions that are useful in achieving globally effective outcomes in the real-world. Interviews with field study participants further suggest that highlighting an individual's contribution to the global goal may help people value their contributions more.

2.2. Background

A general challenge facing all crowdsourcing systems is the dual need to recruit contributors and to make effective use of contributions to best address task needs. In online crowdsourcing, system designers can reason about such needs separately; once a person is recruited to an effort, they land on a website where the most valued, compatible task (i.e. user can contribute) can be delivered. This separation of concerns allows one to design mechanisms for motivating and recruiting users (e.g., [**11, 13, 132, 133**]) independently of mechanisms for coordinating contributions (e.g., [**12, 102, 139, 25**]), regardless of whether tasks are assigned to workers by a system (as is typical) or dynamically determined by workers (as in [**139, 25**]).

In contrast, in mobile and physical crowdsourcing systems the tasks that a person can readily contribute to depend largely on the person's physical location relative to the location of tasks [**63**]. In other words, the tasks that are most in need of completion or that best match a worker's abilities cannot be readily presented unless the person can be motivated to arrive at the task's location. As a consequence, efforts to motivate and coordinate physical crowds cannot consider these two problems in isolation, and instead require the design of system-level mechanisms like Hit-or-Wait that are capable of reasoning jointly about the needs of the system and the changing availability of contributors.

The different models of mobile and physical crowdsourcing lead to particular challenges and tradeoffs for recruiting and coordinating workers. In the *opportunistic*, or *pull-based approach*, it is up to the workers to choose which tasks to contribute to. Even though a system can display, upon request, nearby tasks a worker can best contribute to, this approach can lead to many missed opportunities as it is only effective if workers actively look for tasks as they move about so as to happen upon high-valued tasks nearby [**51**].

In the *directed*, or *push-based approach*, workers are assigned tasks that best address system needs given worker locations and characteristics with the assumption that users' future routes can be determined by the system (e.g., in commercial services like Uber and PostMates) or the routes are known a priori [**65, 134, 21**]. This admits the use of standard optimization techniques to maximize the efficiency of the system through effective task assignments [**18**], but requires strong incentives to recruit on-demand workers who may have to go out of their way to complete tasks [**126, 124**].

In our *on-the-go approach*, workers are sent proactive task notifications in *situations where they are likely able to contribute to tasks that are valued by the system*. Given many tasks with differing values and priorities that need to be completed, effectively connecting users to tasks requires mechanisms to manage a delicate balance between recruiting users in situations where they are able to help and making efficient use of their efforts. While existing task assignment mechanisms are effective for directed approaches [**23, 68, 69**], they are not effective in the on-the-go setting where the user determines their own future routes and thus may never reach the tasks they are assigned. Even if these mechanisms took into account the uncertainty in future routes [**18**], pre-assigning tasks to users is still ineffective as it unnecessarily pre-determines who should do what tasks, which in the on-the-go setting will depend on the tasks that users actually encounter in their routes. Instead of assigning tasks, Hit-or-Wait offers a more flexible approach that reasons about whether to surface a task need at the current location or to wait to surface a different task need at a later time. To do this, Hit-or-Wait uses decision-theory over predictive models of people's routes and models of system needs to determine, on-the-fly, when to engage users for opportunistic contributions that are convenient to them, valued by the system, and that ultimately lead to globally effective solutions.

While prior work in *opportunistic planning* [**59, 67, 60**] had considered the problem of choosing which task to present given uncertainty over a user's route, this choice was static and assumed that a system had to make a decision at a fixed moment in time [**60**]. When unsure of which tasks a user may encounter, such a system may resort to asking a user directly for information about their route, which reduces uncertainty but adds extra effort on the part of the user. In order to reason flexibly about changing conditions without user intervention, Hit-or-Wait moves away from optimizing among a set of tasks towards optimizing over immediate situations and possible future situations. This allows Hit-or-Wait to coordinate contributions dynamically, by controlling when and whether to engage a volunteer as they move from place to place.

In considering a dynamic sequence of decisions over whether to hit or wait, our approach bears resemblance to the use of decision-theoretic methods in online crowdsourcing that optimally control what tasks to allocate and when to stop allocating tasks [**28**]. Whereas efficiency is the primary reason for using decision-theory in earlier work, in our setting, the use of decision theory is further motivated by its ability to empower a seamless and lightweight form of interaction that requires no attention of potential volunteers until a task request is made. Following arguments made by Kim et al. [**77**], we hypothesize that increasing the ability for people to conveniently contribute to local, communal problems may help to engage and sustain contributions over time, providing important benefits beyond any efficiency gain in a single scenario. As previous work has shown that highlighting the uniqueness and benefits of user contributions can elicit more contributions [**8, 108**], we study how volunteers perceive the value of their contributions and explore ways to better communicate how Hit-or-Wait decisions make effective use of volunteers' efforts.

2.3. Coordinate On-the-go Crowds with Hit-or-Wait

In this section, we briefly review the core challenges of coordinating on-the-go contributions within a user's routine; introduce opportunistic *Hit-or-Wait*, an individual-level flexible coordination mechanism for coordinating on-the-go contributions in a way that achieves desired global outcomes; and describe our technical architecture that supports integrating Hit-or-Wait into on-the-go crowdsourcing systems.

As a reminder, there are several core challenges when attempting to coordinate on-the-go contributions: First, there is greater uncertainty around worker participation because on-the-go crowds consist of mobile community members and not dedicated workers. Second, task notifications need to be sensitive to the opportunistic nature of participation and cannot be overly burdensome or disruptive to potential volunteers. Third, the system needs to be able to predict future routes based on current movement patterns, and make decisions while reasoning about the uncertainty of the predicted routes. Finally, the overall uncertainty that surrounds participation and future routes makes successful pre-defined task assignment implausible, and requires solutions to make decisions about when to engage potential volunteers in an online manner.

To address these challenges, we present opportunistic Hit-or-Wait as a general decision-theoretic mechanism for coordinating on-the-go contributions. Hit-or-Wait aims to dynamically coordinate contributions in a way that achieves effective global outcomes by considering both current and future situations that may unfold within a user's routine, and to notify potential volunteers of tasks that they can conveniently contribute to and that most need their help.

2.3.1. Opportunistic Hit-or-Wait

We consider an on-the-go crowdsourcing setting with a set of tasks $\mathbf{T} = \{T_1, T_2, T_3, \ldots\}$ that are distributed across a physical space. Tasks may be of varying values that denote their priority, importance, or fit for a volunteer; task values are assumed known or can be estimated by the system. For any potential volunteer who may be able to contribute, we consider the problem of deciding, on-the-fly, which task to notify the volunteer of among the possibly many tasks the volunteer passes by. To make these decisions, the system can make use of an available *movement model*, which predicts a potential volunteer's future trajectories given historical data and the volunteer's current contexts. In order to not overly burden potential volunteers, we assume that each potential volunteer may be notified of at most one task within a given time horizon. The goal is to notify volunteers of tasks that they reach that are most valued, but with the caveat that given uncertainty in future routes it is possible to notify too early and miss a higher valued task that is reached later, or to pass on a valued task now, when in fact there are no better future opportunities on the horizon (or none at all).

To approach this problem, we model a sequence of *Hit-or-Wait* decisions with a Markov Decision Process (MDP) over a finite time horizon. A MDP consists of a set of states $s \in S$, available actions a in each state s, a transition function $P(s'|s)$ representing the likelihood of reaching state s' from state s, and a reward function $R(s,a)$ that defines the value of taking action a at state s. In Hit-or-Wait, states in the MDP represent possible situations the volunteer may reach. Each state s encodes the location of the volunteer, the task that is at that location (if any), and additionally, other contextual information about the volunteer's particular situational context (e.g., just left work). A volunteer transitions from state to state probabilistically, based on the movement model which provides the transition function $P(s'|s)$. Upon reaching a state that contains a task, the system has two possible actions: *hit* or *wait*. Hitting in state s with a task T notifies the volunteer of the task, and results in a reward that denotes the expected value of the volunteer completing the task. Waiting results in no reward and triggers a transition to the next state, while hitting triggers a transition to a terminal state to model only notifying a volunteer of at most one task.

In order to determine whether to notify the volunteer of a task in a given situation, we compute using the MDP an optimal policy π such that $\pi^t(s)$ denotes the decision to hit or to wait when the volunteer is at state s at time t. Computing this policy compares the expected value of hitting now with the expected value from making a decision later if we wait. Formally, we can represent the value of the optimal policy as:

$$V^t(s) = max(R(s,hit), \sum_{s'} P(s'|s)V^{t-1}(s'))$$

Which states that the expected value of the best decision $V^t(s)$ is the maximum of the expected value of hitting now and the expected value of the best future decisions. Using this recurrence relation, we can solve for the optimal Hit-or-Wait decisions using dynamic programming.

Example Scenario: Lost and Found. To better illustrate how *Hit-or-Wait* can be used in an on-the-go crowdsourcing setting, we describe the algorithm in the context of a community-based lost-and-found scenario. Given a person who lost an item somewhere in a large region, the goal is to coordinate volunteers' existing on-the-go mobility to effectively search for the item. volunteers contribute to small tasks that each request a search in a smaller subregion where the item may have been lost. The system must decide for each potential volunteer, whether to notify them to search in a subregion they are in, or to wait for another opportunity. The goal is to maximize the value of notifying by notifying a user in a less-searched region and wait if they are in a well-searched region.

To model and solve this problem, we can construct a Hit-or-Wait MDP for the lost-and-found scenario as follows: states represent subregions in the large region where the person might have lost an item that contain the search tasks, and the reward models the likelihood that the item is in each subregion. For instance, we may model the reward for searching in a subregion as the

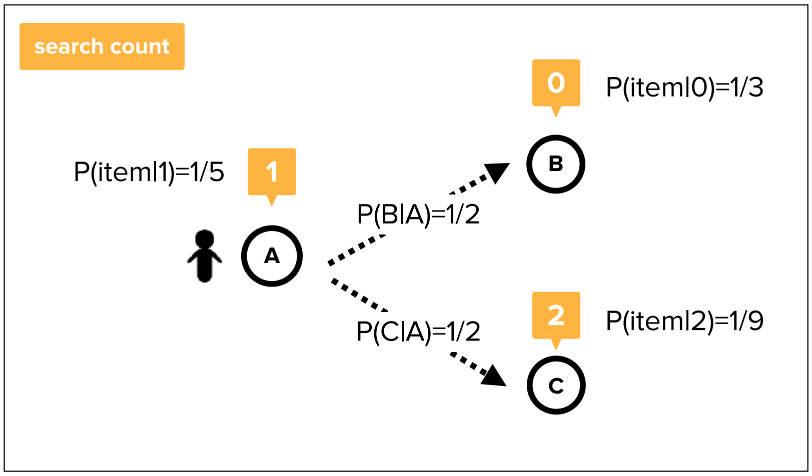


Figure 2.1. An illustrative example for demonstrating how Hit-or-Wait makes hit or wait decisions. The user is currently in the subregion A and is equally likely to go to the subregion B and the subregion C, or $P(B|A) = P(C|A) = 1/2$. We assume that the item is equally likely to be each subregion. The value of notifying the user in each subregion is the likelihood that the item is there following n searches, $P(item|n)$. Using the Bayes rule, we can compute $P(item|0) = 1/3$, $P(item|1) = 1/5$, $P(item|2) = 1/9$.

likelihood that the item is there following n (unsuccessful) searches, or $P(item|n)$. Assuming the likelihood of finding the item is conditionally independent given the item is in the subregion, we can compute $P(item|n)$ using Bayes' rule: $P(item|n) = \frac{P(item)P(n|item)}{P(item)P(n|item)+P(\overline{item})P(n|\overline{item})} = \frac{P(item)P(1|item)^n}{P(item)P(1|item)^n+P(\overline{item})}$, where $P(item)$ denotes the prior probability that the item is in the subregion, and $P(n|item)$ denotes the likelihood of n unsuccessful searches given that the item is there. Given that an item is more likely to be there after fewer searches, as a potential volunteer walks around the neighborhood where the item may have been lost, Hit-or-Wait will tend to notify them to search in a less-searched subregion than in other subregions they might encounter. Unlike an approach that may pre-determine when to notify a potential volunteer of a task solely based on the reward—which may miss opportunities if a user does not come across the region in their actual realized route, Hit-or-Wait constantly reasons about trade-offs between the value of notifying a user of current task and the value of waiting and making decisions in the future.

We refer to Figure 2.1 as a simple example to illustrate how Hit-or-Wait makes effective hit or wait decisions. In this example, the user is currently in the subregion A, and based on the user's historical route data, the user has the equal likelihood of reaching the subregion B and the subregion C, $P(B|A) = P(C|A) = 1/2$. We assume that the item is equally likely to be in one of the three subregions, $P(item) = 1/3$. We also assume that the likelihood of finding the item after the first search given that the item is in the subregion as $P(1|item) = 1/2$. The value of notifying the user in each subregion is the likelihood that the item is there following n searches, $P(item|n)$. Using the Bayes rule, we can compute $P(item|0) = 1/3$, $P(item|1) = 1/5$, $P(item|2) = 1/9$.

In order to determine whether to notify the user of a task in the subregion A or wait for better opportunities in the subregion B or C, we compute $V^t(s) = max(R(s,hit), \sum_{s'} P(s'|s)V^{t-1}(s'))$. We can compute the first term R(A, hit), the value of notifying the user in the subregion A, which is $P(item|1) = 1/5$. Then, we can compute the second term $P(B|A) * V(B) + P(C|A) * V(C)$, which is the expected value of best future decisions in the subregion B and C. Because we already know $P(B|A) = P(C|A) = 1/2$, we only need to compute V(B) and V(C), which is the value of the best decision in the subregion B and in the subregion C, respectively. In the subregion B and C, the best decision is to hit because waiting will lead to missed opportunities as the user will leave the search region entirely. Given that the best decision is to hit in the subregion B and C, the value of notifying the user in the subregion B is $R(B,hit) = P(item|0) = 1/3$, and the value of notifying the user in subregion C is $R(C,hit) = P(item|2) = 1/9$. Therefore, the expected value of best future decisions in the subregion B and C is $1/2 * 1/3 + 1/2 * 1/9 = 2/9$. Because

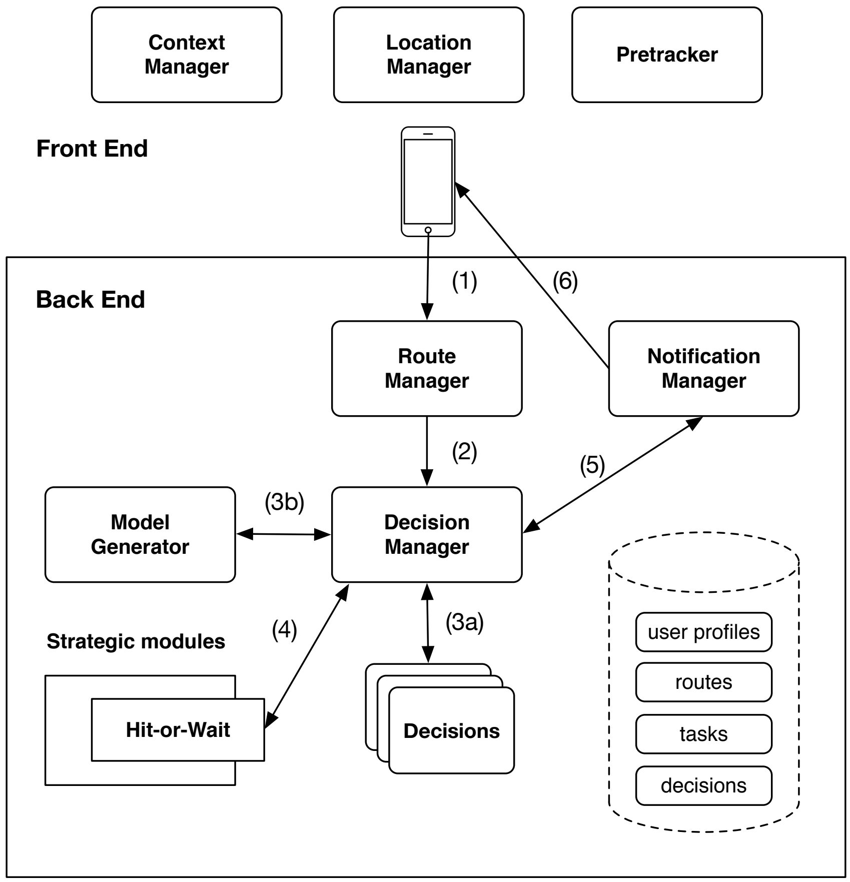


Figure 2.2. On-the-go crowdsourcing architecture.

the expected value of making decisions later (i.e. 2/9) is greater than the value of hitting in the subregion A right now (i.e. 1/5), we make a wait decision in the subregion A.

2.3.2. On-the-go Crowdsourcing Architecture

Building on-the-go crowdsourcing applications powered by Hit-or-Wait requires an architecture that can track location data, sense user's context, and make decisions of when and which tasks to notify based on user location and context. The architecture is described below and shown in Figure 2.2.

2.3.2.1. Architecture. The *Location Manager* and *Context Manager* collect user's location data and other contextual information and communicate them to the back-end.

The *Pretracker* helps deliver precise, fine-grained notifications by managing device's location accuracy depending on user's current location. For example, if a user is far from a task region, it decreases the location accuracy and increases it once the user is nearby a task region. Since it dynamically manages location accuracy, it can both save device battery and deliver fine-grained notifications at or near a task location.

The *Route Manager* processes incoming user location data in the back-end. It maps latitude and longitude pairs to states, stores new trips, or updates the existing ones in the database.

The *Model Generator* computes models that are required for strategic modules. For example, it produces the movement model for state transitions using previous route histories.

The *Decision Manager* takes as input user profiles (including routes, contexts), models, and tasks, and generates as output decisions based on a strategic module (e.g. Hit-or-Wait).

The *Notification Manager* delivers notifications to the the front-end based on the decisions made by the Decision Manager when a user meets the notification criteria, which includes but is not limited to conditions over the user location and the frequency of notifications (e.g., to model disruption and to avoid over-notifying users of tasks).

2.3.2.2. Flows. Figure 2.2 demonstrates how the various components interact with each other. The Route Manager receives raw GPS coordinates from the user device (1), preprocesses and stores the data, sends them to the Decision Manager (2). The Decision Manager first checks whether or not there exists decisions computed for the given location (3a), if there exists decisions, it sends the decisions to the Notification Manager (5). Otherwise, the Decision Manager requests models from Model Generator (3b). Together with the user location, the generated models, user profile, and task needs, the Decision Manager chooses a strategic module to compute decisions (4), and finally sends the decisions to the Notification Manager. The Notification Manager considers notification criteria such as the distance to a task location, user profiles, and delivers a task notification if all the criteria are met (6).

2.4. Study 1: Simulation

We conducted a simulation study in a community-based lost-and-found setting to understand (a) the performance of Hit-or-Wait mechanism for indirectly coordinating contributions towards global goals, and (b) the effect of movement model accuracy on the performance of Hit-or-Wait.

2.4.1. Dataset and Modeling

To train a movement model and simulate the routes of on-the-go volunteers, we scraped running routes from publicly available RunKeeper data in Chicago and its northern suburban area. The dataset contains 5,983 running routes from 2,419 users. It contains a total of 590,860 latitude and longitude pairs for an average of 98.76 points per user.

We model each subregion where an item may have been lost by representing individual road segments as states. We gather road segment data from OpenStreetMap, which treats each segment as a connection between street intersections, represented as sequence of latitude and longitude pairs that construct the segment. We preprocessed our data following the steps from [**81**] but adopted the following heuristic for converting GPS traces into a sequence of adjacent road segments. For each latitude and longitude pair in a runner's GPS trace, we sought a road segment within 40 meters in the OpenStreetMap dataset. If we couldn't find the nearest road segments, we marked the road as Unnamed Road. We eliminated repeated road segments to finish constructing the sequence of segments.

We used the processed data to generate a population-based movement model where the transition probability from one road segment to the next is trained using the frequencies observed in the data. We consider a first-order Markov model where predictions of next locations are conditioned only on current locations. We used population-based model instead of individual-based model because there were not enough individual route histories to train an accurate individual-based model. For routes on which no training data exists, we used a simple model trained across our dataset that assigns a probability distribution over going straight, turning left or right.

2.4.2. Simulation I: The Efficiency of Hit-or-Wait

2.4.2.1. Study Procedure. We compare the performance of Hit-or-Wait with other flexible coordination solutions: a simple *node counting* algorithm and with a *myopically optimal* solution given full knowledge of people's routes. The node counting algorithm notifies a person of a task in a subregion if and only if the search count in that subregion is the lowest among all subregions. This algorithm follows a user's routine and makes efficient use of presented opportunities if they are most valuable to the system, but the lack of knowledge of a user's entire routine and other opportunities that may unfold within the routine makes this algorithm prone to miss opportunities (e.g. waiting for opportunities may fail to recruit volunteers who do not approach areas with low search count). The myopically optimal solution is omniscient of a volunteer's routes and notifies a volunteer in the subregion they come across that has the lowest search count. This algorithm is an ideal solution for flexible coordination and serves as an upper bound on the performance of

Hit-or-Wait given its perfect knowledge of a user's routine to surface best opportunities within the routine.[1]

To set up lost-and-found scenarios, we chose a road segment in our dataset with the highest foot traffic and included 41 nearby road segments to form the area for our study. Within this area, we randomly selected 10 road segments to represent the search subregions where the item may have been lost. We set the reward for searching in each subregion to the likelihood that the item is there after n people have (unsuccessfully) searched in that subregion (i.e., $P(item|n)$). For each trial of our simulation, we randomly sampled 100 routes from 428 running routes from 269 unique runners.

2.4.2.2. Measures and Analysis. We measure the performance of our algorithm against other algorithms by considering the *overall search quality* and *the number of missed opportunities*. Overall search quality provides a measure of how likely a search effort (i.e., the number of searches in each subregion) is to result in finding a lost item. For simplicity, we assume that the item is equally likely to be in each state,[2] and that searches are independent conditional on the item being in the search region. We set the likelihood of finding the item after the first search given that the item is in the subregion as 0.67. We let $V(s,n)$ denote the likelihood of having found an item after n searches when the item is in state s, and compute the quality of search as: $QoS = \sum_s V(s,n)/|S|$.

[1]To make the baseline comparison informative and compelling, we refrained from (a) comparing to approaches that notify users of tasks at non-nearby locations, which differs from our setting; and (b) comparing to directed approaches that pre-assign users tasks that may never be on their actual routes, as their performance would be similar to or worse than our chosen baseline.

[2]For the simplicity of the measure we treat the likelihood that an item is in a state as a constant, when in practice search counts contribute information about where the lost item is.

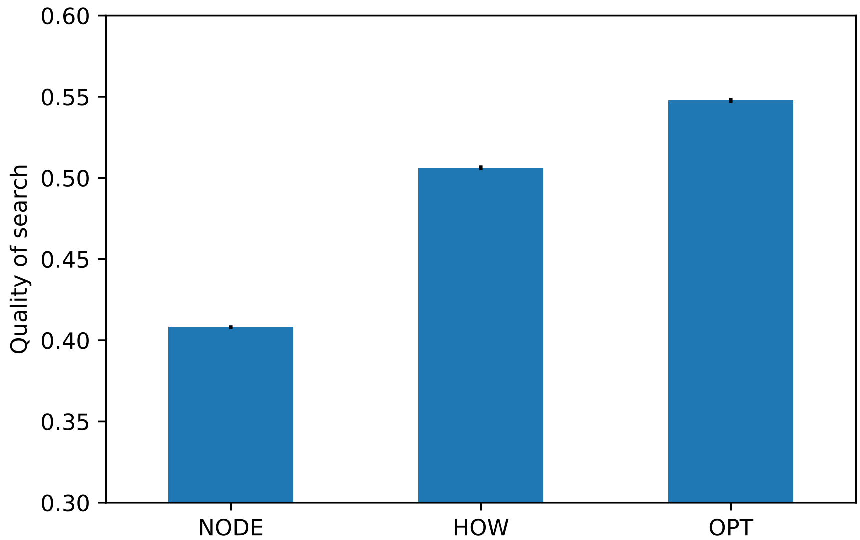


Figure 2.3. Simulation results comparing the overall quality of search for node counting, Hit-or-Wait, and myopic optimal. The results show that Hit-or-Wait outperformed the node counting algorithm and approached the performance of the myopic optimal solution.

For the number of missed opportunities, we measure occurrences where a person, given their actual route, could have been notified to search along their route but were not notified (regardless of the value of contribution).

2.4.2.3. Results of Simulation I. Figure 2.3 shows that Hit-or-Wait outperformed the node counting algorithm and approached the performance of the myopic optimal solution. It shows that Hit-or-Wait achieved 92.43% of the value of the myopic optimal solution, whereas node counting only achieved 74.5% of the value of the myopic optimal. Compared to node counting, Hit-or-Wait makes use of more of potential volunteers' efforts by drastically lowering the percentage of missed opportunities compared to node counting algorithm; see Figure 2.4. On average, Hit-or-Wait algorithm missed 46.07% of opportunities (SD: 14.49) while node counting

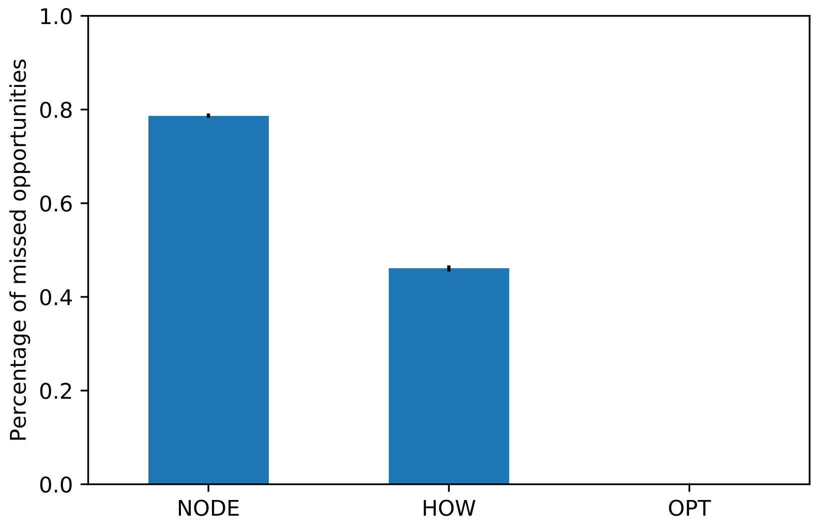


Figure 2.4. Percentage of missed opportunities for node counting, Hit-or-Wait, and myopic optimal solution. Hit-or-Wait algorithm missed 46.07% of opportunities (SD: 14.49) while node counting missed 78.68% of opportunities (SD: 10.53).

missed 78.68% of opportunities (SD: 10.53) by waiting for people to enter the least searched regions. While node counting notified users of the highest valued tasks exclusively, many users never reached such tasks; this led to a high percentage of missed opportunities that ultimately resulted in a lower overall quality of search than Hit-or-Wait. The myopic optimal solution has full knowledge of people's future routes, and thus misses no opportunities. This suggests that, by using an understanding of how a user's routine may unfold, Hit-or-Wait considers other opportunities that may arise within the routine to make more informed decisions about when to engage a user with which opportunities to achieve globally effective outcomes.

2.4.3. Simulation II: The Effect of Model Accuracy on Hit-or-Wait

2.4.3.1. Study Procedure. To understand the effect of movement model accuracy on the performance of Hit-or-Wait, we compare the performance of Hit-or-Wait using more and less accurate movement models across two types of situations: *uniform neighboring values* and *varied neighboring values*. In situations of uniform neighboring values, tasks in neighboring states are uniform in value; in such situations, we hypothesize that movement model accuracy has less impact on the performance of Hit-or-Wait because the value of future decisions is largely invariant. In situations of varied neighboring values, neighboring tasks differ in value; incorrect predictions of future routes are thus more likely to affect the quality of Hit-or-Wait decisions.

To set up an illustrative scenario, we chose a search region that consisted of three road segments that are in the area we had chosen in Simulation I for which the movement model is strongly discriminative. This allows us to observe different decisions when using Hit-or-Wait with our trained model and a less accurate model that transitions to neighboring states uniformly at random. We considered all 44 running routes that passed by this region, and considered each route as an instance of a potential search. We set the current road segment with value 0.6, and set the mean value of the neighboring road segments to 0.5. For uniform neighboring values, this should result in hit decisions at the current road segment regardless of the movement model. For varied neighboring values, we uniformly sampled a value in a range of 0.8 to 1 and set it as the value of the neighboring road segment more likely to be reached (and 1 minus that value for the other neighbor to preserve the mean of 0.5). This should allow a more accurate movement model to make wait decisions when it has strong predictions of reaching more valued states, whereas a less accurate movement model may still decide to hit.

Figure 2.5. Effect of model accuracy on Hit-or-Wait performance in situations of uniform neighboring values and varied neighboring values.

2.4.3.2. Measures and Analysis. To study how movement model accuracy affects the performance of Hit-or-Wait with the trained model and a uniformly at random model, we measure the percentage of value captured with respect to the myopic optimal solution in situations of uniform neighboring values and varied neighboring values. We chose this measure instead of *overall search quality* because we are only looking at specific moments where we vary the task values, which means that there are no accumulated searches across the regions to compute overall search quality.

2.4.3.3. Results of Simulation II. Figure 2.5 shows the performance of Hit-or-Wait algorithm with a uniformly at random movement model and our trained model in the situations of uniform neighboring values and varied neighboring values. As we hypothesized, Hit-or-Wait using the uniformly at random movement model still achieves good performance in situations of uniform neighboring values, but not in the case of varied neighboring values. In the varied neighboring values case, Hit-or-Wait with our trained model captured 95.08% of the values of what myopic optimal was able to achieve, while Hit-or-Wait with the uniformly at random movement model only captured 67.66% of the values of myopic optimal. In this particular example, Hit-or-Wait made the same decisions as OPT in the case of uniformly neighboring values, as (a) incorrect predictions did not lead to any missed opportunities (e.g., the user ends up in a region without a task); and (b) the value gained for hitting in subsequent states is identical regardless of next states.

2.5. Study 2: Field Deployment

Following our simulation study, we conducted a 10-day long field deployment of Hit-or-Wait in the lost-and-found domain to understand (a) the performance of Hit-or-Wait in comparison to a myopic optimal solution—our upper bound—in the real world, and (b) users' perceived disruptions. In addition to the simulations, this study allows us to explore the balance between hitting and waiting, the consequences of wait decisions, and when and why wait decisions may fail.

2.5.1. Trouve: Lost-and-Found Application

We developed a prototype, *Trouve*, a lost-and-found mobile application where users can request searches for lost items and it notifies people who pass by possible lost item regions to request that they look for the items. A user who lost an item can post a request by providing a lost item description and a possible region where they might have lost the item. When a potential volunteer passes by a subregion in the potential search region, they receive a notification (Figure 2.6a) asking if they can help look for the lost item there and then. Once they click the notification, the

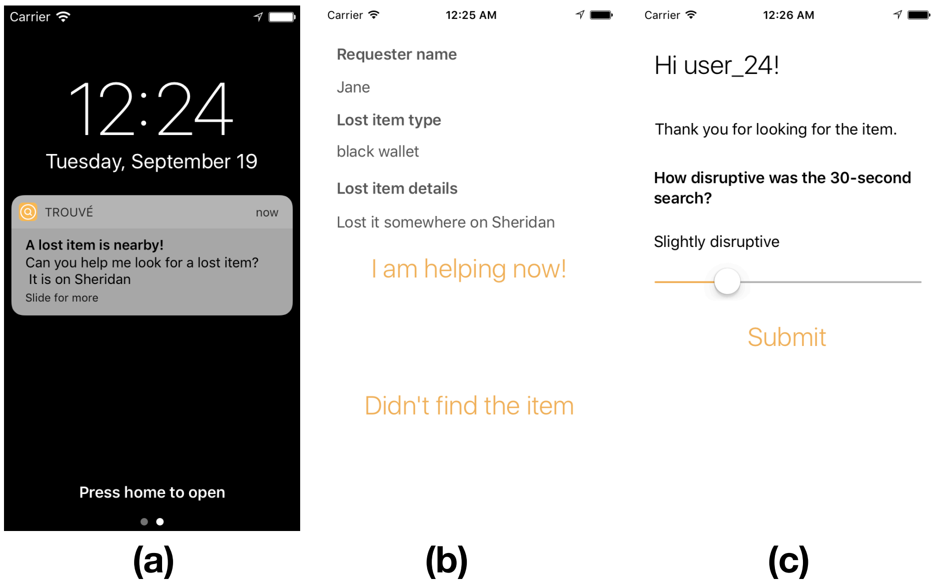


Figure 2.6. Trouve, a prototype lost-and-found app for the study. (a) A user receives a notification indicating a lost item is nearby; (b) the user sees the details of the lost item; (c) after 30 seconds of clicking I am helping now, the user is prompted with a survey question about perceived disruptiveness of the 30-second search.

relevant information is shown to the user (Figure 2.6b). If the user decides to help, they can click "I am helping now!" to indicate their search attempt. After 30 seconds, a survey question about the perceived disruption of the 30-second search is shown to the user (Figure 2.6c).

2.5.2. Study Procedure

We recruited 25 people who had an iPhone 5S or above with iOS 10+ via flyers and local university mailing lists. 13 participants were male and 12 were female; the average age was 21.7 (SD: 2.79). Participants consented to enrollment and then received the study instructions that asked participants to look for lost items for approximately 30 seconds when notified while traveling along their existing routes, and noted that searching was not mandatory for their participation. One of the authors acted as the requester, using Trouve to post lost items for participants to find.

We generated a movement model using the same procedure as in our simulation study, but with the training data coming from 51 routes from 11 recruited participants who used our location tracking app for a week prior to the study. Throughout the study we collected additional 1,490 routes and continually updated our model each time new route data was collected.

We chose two search regions near the university campus: one near the south side and another near the north side of the campus. Each search region included 5 road segments as its subregions; around 70-100 meters for each road segment. We sought to have a mixture of both high traffic and low traffic pedestrian streets within each search region, so we interviewed students who frequently traverse the regions and used the pre-study location data to help guide our choice of roads.

For each lost item request for a search region, one road segment was randomly selected as the lost item location. Based on common requests on the university's lost & found group, a lost item region was described as "somewhere on *street name #1* and *street name # 2*." Requested lost items included a wallet, a coin purse, and trinkets. The interval between task notifications (per user) was 4 hours to help minimize the overall disruption caused to the participants. Following each unsuccessful search in a subregion, we updated the search count and reduced the reward for subsequent searches in that subregion.[3] The search requests expired either when someone found the item or if no one found the item after 3 days.

Since the primary focus of the study was coordinating searches across road segments within a search region and not handing off found items or delivering them to a lost-and-found center, we asked the participants to simply take a picture of the found item and send it via SMS or email to the researchers. This way, we were able to verify whether or not the participants actually found the item. In the discussion, we will discuss more complex scenarios of Hit-or-Wait in which the volunteers have to travel with and then hand off the found items. The participants received a $25 gift card as compensation.

2.5.3. Measures and Analysis

To understand the performance of Hit-or-Wait in the real-world, we considered two new measures: *the perceived cost of disruption* and *the value of waiting*. To measure the perceived cost of disruption, we used both ecological momentary assessments (EMA) and post-study survey, where the participants were asked to rate their perceived disruption on a 5-point Likert scale (ranging from “1: not disruptive at all,” to “5: very disruptive”). The EMA was delivered to the participant via their smartphone (see Figure 2.6c) 30 seconds after they clicked “I am helping now!” We used both EMA and post-study survey to complement each other’s strengths and weaknesses. On one hand, while EMA allows us to collect user responses while their memory is fresh, the

[3] Due to an error in Hit-or-Wait implementation, we encoded the value of searching after n searches as $(1-P(1|item))^{n+1}$ in Study 2. While this value is decreasing in the number of searches as we would want, a more accurate estimate of the value of search should be based on $P(item|n)$ as shown earlier. Compared to using $P(item|n)$ as the reward function, our implementation overvalues states with more (unsuccessful) searches. This can lead Hit-or-Wait to make more hit decisions in well-searched regions when in actuality waiting for a less searched region would have been more valuable. As a result, the performance of Hit-or-Wait in our deployment may have been lower than if we had implemented the more accurate reward function. The error did not otherwise affect our measures, analyses, or findings.

responses were collected only when the participants decided to help, and therefore we miss responses when they found the tasks disruptive and did not help. This measure more effectively captures reflection on the disruption of the 30-second search task itself. On the other hand, the post-study survey allows us to capture participants' reflection on the amount of disruption they experienced throughout the study both including times when they decided to help and when they declined. One downside of this measure is that the participants' memory may not be as accurate after 10 days of study and their reporting may exhibit recall or recency bias.

In addition to *overall search quality* measure from Study 1, we added a measure that captures the value of waiting. We considered the wait decisions Hit-or-Wait made and compared the value of the eventual outcome (e.g., based on whether and where a person eventually searched) to the value if we just sent them the task then and there. To measure the real-world performance, we made this comparison by using the actual number of searches performed in the subregion thus far to compute the likelihood that the item is still in that subregion (i.e., $P(item|n)$), instead of using the expected value of waiting as computed by Hit-or-Wait. As the system might make multiple wait decisions until it makes a hit decision, we considered only the first wait decisions for comparison.

2.5.4. Study Results

2.5.4.1. Searching with Low Disruption. Over the course of 10 days, the participants received 248 notifications and conducted 60 searches along their routes (24.19% acceptance rate). We found that the 30-second on-the-go searches were not disruptive to the users when they decided to help. Figure 2.7 shows that both the EMA (N=60 from 23 out of 25 participants) and post-study survey responses (N=24) about the perceived cost of disruption were low; the average rating

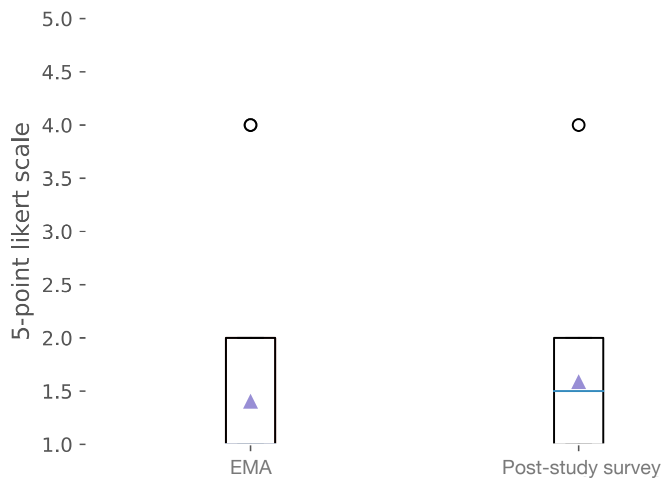


Figure 2.7. Responses from EMA and post-study survey about perceived disruption for 30-second searches on a 5-point Likert scale (1: not disruptive at all, 5: very disruptive).

from EMA is 1.39 (SD: 0.58) and the average rating from post-study survey is 1.58 (SD: 0.72), values that fall between "not disruptive at all" and "slightly disruptive."[4] Our interview findings also mirror the survey responses, as P2 said: *"So everyday I walk passed [street name] and [street name], and usually I am on my phone when I am walking and I see the alert. I usually just keep walking and look around my path and look for the item...It's not bad at all and really easy."*

Among the searches, 4 different participants found 4 items out of the 9 search requests that were made. While finding items was not the primary focus of our study, this finding demonstrates how effective coordination can make use of smaller contributions from many users to find lost items in large regions.

2.5.4.2. Maximizing User Contributions. The results show that Hit-or-Wait was effective in maximizing the user contributions by notifying the tasks where they were most needed. For 57

[4]One of the participants did not fill out the post-study survey and never responded to the emails.

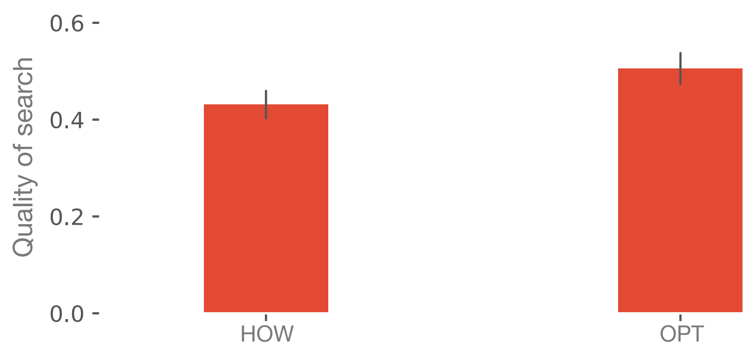


Figure 2.8. Overall quality of search between Hit-or-Wait and myopic optimal solution. Hit-or-Wait was able to make near optimal decisions in actual use, and make efficient use of people's search efforts.

searches that took place in the study, the average quality of search was 0.43 for Hit-or-Wait (SD: 0.21) and 0.51 for myopic optimal (SD: 0.24), indicating that Hit-or-Wait captured 84.31% of the value of what myopic optimal is able to achieve (Figure 2.8).[5] Closer analysis shows that Hit-or-Wait made hit decisions in subregions with a higher search count than myopic optimal only 9.68% of the time (24 out of 248), and in 77.42% of the times (192 out of 248) it made decisions identical to the myopic optimal. These results, together with the self-reported perceived disruption, suggest that even without a full knowledge of people's entire routine, individual-level flexible coordination mechanisms such as Hit-or-Wait can make near optimal decisions to elicit useful contributions from people while still staying within their routines.

We found that Hit-or-Wait also made effective wait decisions. Figure 2.9 illustrates the value of waiting and shows that deciding to wait led to future decisions with a 67.6% increase in value compared to immediately notifying users. A paired t-test shows that there is a significant

[5]We excluded 3 searches from this analysis since they were missing the GPS location data needed to compute the measure.

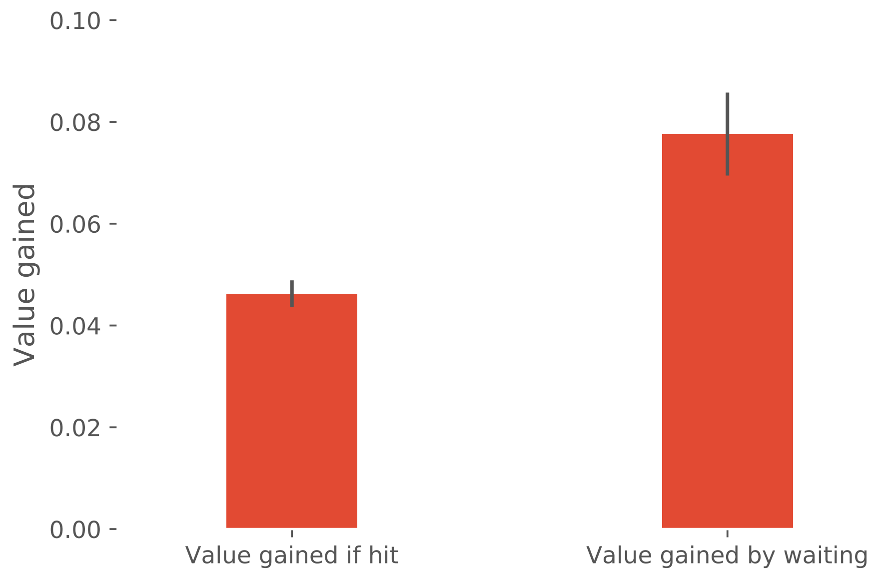
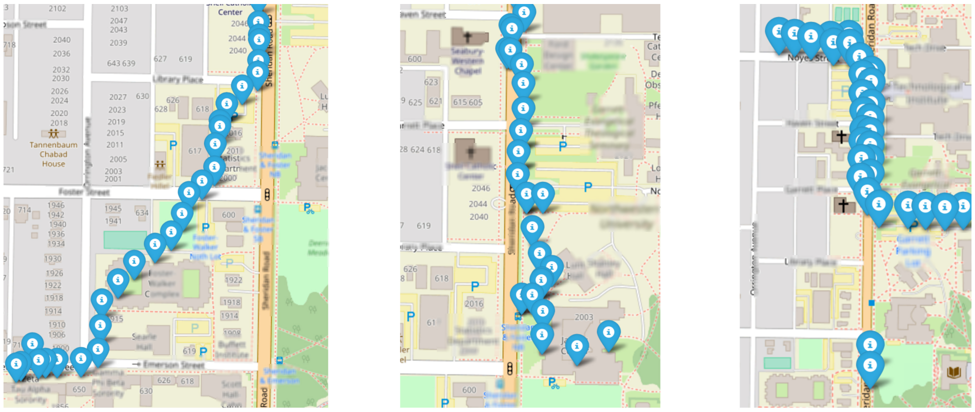


Figure 2.9. Comparison between the value gained if hit at the wait decision state vs. value gained from waiting and notifying at the later state.

difference in the value of waiting versus not waiting (t = 3.98, df = 120, $p < 0.0001$). In other words, by understanding a user's routine and opportunities that may unfold within the routine, flexible coordination mechanisms such as Hit-or-Wait do not have to resort to whichever opportunities that arise within a user's routine, but instead it can compare across the value of the present and possible future opportunities to make better decisions about when to engage people with an opportunity.

2.5.4.3. When and Why Missed Opportunities Happen. Waiting for a better opportunity poses a risk of completely missing the opportunity to notify. Our results show that the variance for value gained from a wait decision is quite large (M: 0.0776, SD: 0.0893), and it is mainly due to the fact that there is zero value gained when missed opportunities occur. Our results show that 45.16% (56 out of 124) of the wait decisions resulted in missed opportunities.



(a) **(b)** **(c)**

Figure 2.10. Examples of missed opportunities after wait decisions: (a) taking shortcuts and did not go to adjacent states; (b) staying at the wait state (e.g., taking classes); (c) missing GPS location data.

From 52 instances, outside of situations where uncertainty in future routes naturally led to users going outside of the task region, we identified three other reasons that resulted in missed opportunities (Figure 2.10). First, contrary to our model, people sometimes took unexpected routes or did not move to adjacent states (Figure 2.10a). For example, some people took shortcuts or trespassed in ways that were unexpected by our model and this led to missed opportunities where the system may have notified them in other subregions. In the future we could have models with more fine-grained state spaces, where the state is more granular than a road segment.

Second, people sometimes stopped moving and stayed at the location where the system made a wait decision (Figure 2.10b). These instances occurred when the participants were on their way to classes or home, and they only passed one of the subregions since their trip was cut short. Our system did not have the notion of terminal state, but in the future it could predict whether or not the current state will be the terminal state so that we can prevent such missed opportunities.

Third, inaccuracy or inconsistency in location tracking also caused some missed opportunities (Figure 2.10c). There are many reasons such technical failures can happen (e.g., turning off Wi-Fi and thus lowering the location tracking accuracy; switching between LTE and Wi-Fi while walking around the campus; turning on low-power or airplane mode). For the rest of 4 instances it did not notify due to technical failure.

2.6. Follow-up Interviews

In the follow-up interviews after the field deployment, we sought to understand how volunteers perceive the value of their contributions toward the larger goal of finding the item in a large search region. We also explored ways to represent and visualize the value of contributions and use it as a tool to better communicate seemingly opaque Hit-or-Wait decisions.

2.6.1. Interview Setup

We invited participants for an optional 30-min interview after the field deployment and interviewed 7 participants who helped at least once during the study. Each interview lasted around 30 minutes. We chose 4 different scenarios to highlight a high-level idea of how *Hit-or-Wait* works: 1) A Hit decision is made because a user is at a road with no searches; 2) a user is at a road with some searches but the user is likely to go to another road with no searches, so it makes a Wait decision at the current road and makes a Hit decision if the user reaches the subregion with no searches; 3) a user is at a road with a few searches and the user is likely to go to a road with a fewer searches, so it makes a Wait decision at the current road, and a Hit decision if the user reaches the road with the fewer searches; 4) a user is at a road with a few searches and the user is likely to go to a road with more searches, so it makes a Hit decision at the current road.

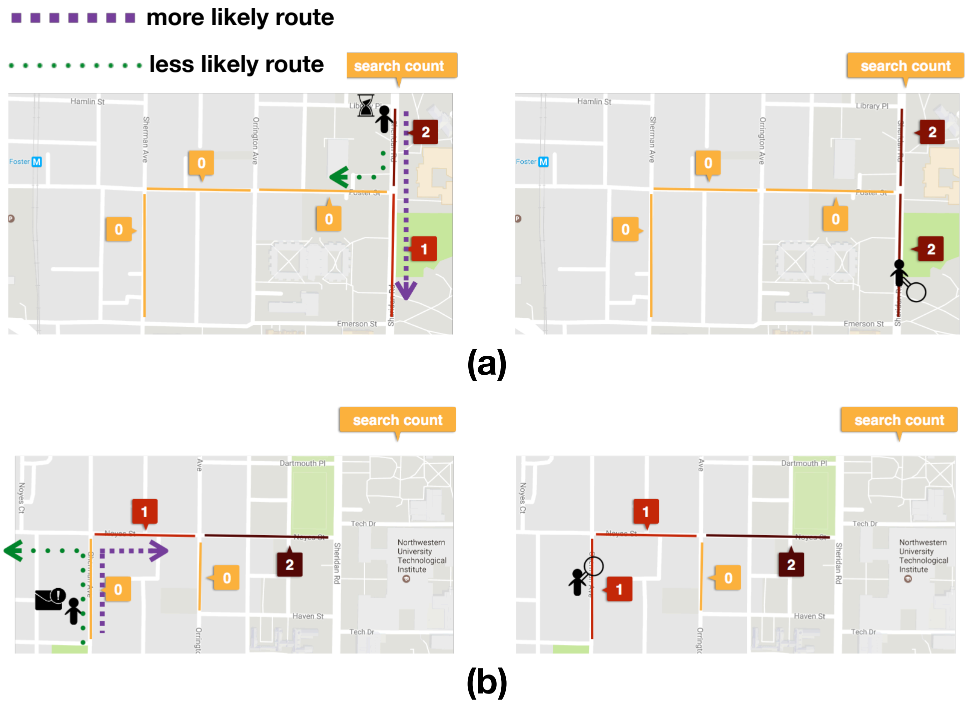


Figure 2.11. Wait (top) and Hit (bottom) example visualizations.

During the interviews, we first asked users to recall their searches and tell us about the perceived value of them. We chose different contribution scenarios from a user's actual contributions (when the user's searches did not cover all four scenarios, we showed other users' searches instead) and showed the visualizations for those searches (some examples of the visualizations are shown in Figure 2.11). We then asked the participants to walk us through how they thought the system worked based on the visualizations. After showing the participants all of the visualizations, we again asked them about the perceived value of their searches, and we elicited their suggestions about how the system could more clearly communicate its goals and highlight the value of their contributions. The interview participants received a $5 gift card as compensation.

2.6.2. Interview Findings

Some participants perceived the value of their contributions solely on the basis of whether they found the item, and as a result they did not regard their contributions as valuable if they were unable to find the item. For example, P6 described how she thought that her contribution was not valuable: *"Well, clearly wasn't that valuable, because I never found anything."*

Some participants also assumed that the system did not take into consideration other people's searches, and perceived their contributions either as redundant or too miniscule to be valuable. One participant (P6) explained how she thought that the system was notifying everyone who passed by and as a result many people would have searched in a same region: *"there was nothing to stop someone else from doing the exact same search that I did even if I already searched that area, right?"* On the contrary, another participant assumed that she was the only one searching in a large search region and did not feel her search was ever going to be useful (P2).

However, when the participants understood the high-level idea of the mechanism of Hit-or-Wait—predicting likely routes and considering other people's searches—either through the visualizations or verbal description from the interviewer, they stated that their contributions were more valuable. P7 said: *"Oh, definitely valuable because it carefully calculates who has already [searched], so I don't feel like I am just another person who's like useless."* Another participant P2 said: *"I guess it is a lot more valuable. Because I guess I've never thought the computer was taking in how other people are doing it."*

The participants also mentioned that highlighting an individual's contribution as part of the global goal may help them value their contributions more. As P7 stated: *"Maybe also having information like if someone does find the item, then I would know I was just being helpful...So I was helping part of that even if I wasn't the exact person to find it."* The participant also said that emphasizing the uniqueness of her contributions could have helped her feel the contributions to be more valuable: *"It's nice to know that I am the first person to search like there...If I saw this while I was searching, that would've made sense and I may have felt like it's valuable."*

To summarize, our interview results show that it's important to communicate the global goal of the system and highlight the parts of the goal that the users are contributing to so as to help them to be cognizant of the value of their contributions.

2.7. Discussion and Future Work

In this chapter, we introduced Hit-or-Wait, a general decision-theoretic mechanism for flexible coordination that coordinates opportunistic contributions within a user's routine to achieve effective global outcomes. We demonstrated the effectiveness of Hit-or-Wait through simulations and a field deployment, which highlighted Hit-or-Wait's ability to follow and stay within a user's routine and elicit useful contributions by deciding on-the-fly whether to notify or wait for better opportunities. In the rest of the section, we discuss the applicability of Hit-or-Wait to more complex scenarios and other domains; tractability and scalability; limitations of myopic Hit-or-Wait; and lastly, the general need for system-level coordination in on-the-go crowdsourcing systems.

2.7.1. Complex Scenarios and Applicability to Other Domains

As the goal of this chapter is to explore ways to implicitly coordinate user contributions towards global outcomes, we excluded subsequent scenarios where volunteers have to hand off found items or factors that could affect the willingness and convenience besides user's current location in the field deployment. To make applications like Trouve a full-fledged system with Hit-or-Wait, we could take into consideration the cost of diversion [**60, 114**] from a user's existing route or predicted destination [**82, 142**] to a hand-off location when computing the value of notifying the user. We could also include parameters that capture busyness, schedules, existence of companions, and other situational factors that are known to affect task acceptance rate [**63, 77**].

While we studied Hit-or-Wait in the context of a community-based lost-and-found, the general mechanism can be applied to other domains such as community sensing or for other community-based peer-to-peer services. For instance, in community sensing, Hit-or-Wait can be used to support the global goal of ensuring data coverage and fidelity, even when using low-effort contributions [**131, 128**]. Depending on where and how much data has been collected at different locations, Hit-or-Wait can decide when to ask for additional pieces of information, for example by making the decision to wait should the user be likely to reach other locations where data coverage is low. In community-based peer-to-peer services such as timebanking, Hit-or-Wait can be used to achieve the community goal of effectively providing help for each other by accounting for different skills, abilities, and preferences [**32, 66**]. Hit-or-Wait can encode the value of contributions based on required skills, priority, as well as volunteer preferences. For example, depending on a task's urgency, Hit-or-Wait can effectively coordinate opportunistic contributions to prioritize high-valued, urgent tasks that a user may encounter on their route.

2.7.2. Tractability and Scalability

Our MDP model for Hit-or-Wait scales well for reasonable state spaces; there are no immediate tractability concerns at the community- or neighborhood-scale that on-the-go crowdsourcing systems are intended to be deployed in. In cases where the state space becomes large, either in larger-scale systems or by including other contextual information, we can scale further by employing standard techniques such as using coarser or factored state representations [**48**].

2.7.3. Limitations of Myopic Hit-or-Wait

One of the limitations of the current Hit-or-Wait implementation is that it makes decisions on an individual basis without regard to the possible future routes and decisions of other volunteers who may arrive. By taking into consideration others' future routes and decisions, *futuristic Hit-or-Wait* can potentially coordinate contributions more effectively, especially in cases where the contributions are contingent on differentiating factors among volunteers. For instance, if only certain people have access to locations (e.g., returning a book to university library), then by predicting who will come across which locations we can more effectively coordinate these scarce resources where they are most needed. Realizing the benefits of futuristic Hit-or-Wait will require overcoming the computational challenges imposed by (a) reasoning about the potential routes and decisions of future volunteers; and (b) considering the interdependencies of how current decisions can affect future decisions. Resolving these challenges to provide globally optimal solutions through opportunistic coordination will require applying and advancing existing decision-theoretic methods.

2.7.4. Towards Community-level Coordination

Hit-or-Wait's ability to minimize disruptions and eliminate coordination costs [**95**] increases the ability for people to conveniently and effectively contribute to local, communal problems. In this way, Hit-or-Wait can potentially help encourage and sustain more contributions over time. We hope that such new ways of contributing to local, communal problems can provide social benefits, and create new ways of interacting with, supporting, and becoming a part of a community.

Future work on developing community-level mechanisms for flexible coordination may look beyond making effective use of individual contributions to considering how to engage the community of volunteers as a whole. For example, a supply management framework may be able to balance the demands from requesters with disruption to potential volunteers, considering both system goals such as the quality of service but also the needs of volunteers such as not overburdening volunteers, so as to maintain a healthy pool of future volunteers [**77**]; see Chapter 4. In a different vein, as some of our participants from the interviews indicated that they may be willing to deviate from their routes, there may also be opportunities to coordinate mixed models of contributing that engage both on-the-go volunteers and more dedicated volunteers to collectively respond to local, communal needs.

CHAPTER 3

4X Framework

While Hit-or-Wait builds up an understanding of a user's future trajectories to better reason about opportunities that may arise within a user's routine and evaluate the quality of opportunities, there are other factors, such as our knowledge of the world, that can also affect which opportunities may arise within a user's routine and ways in which we should evaluate the quality of opportunities. In this chapter, we discuss the core challenge in taking into consideration our knowledge of the world and people's changing interests when evaluating the quality of opportunities within a user's routine in the context of participatory sensing platforms.

3.1. Introduction

Participatory sensing has developed into an effective method for actively engaging large numbers of people to report data about the dynamically changing physical world to help us understand it and to enable new services [**127, 86, 104**]. For example, birding hobbyists record their observations to help scientists track migration patterns [**120**]. Citizens in the U.S. make 3-1-1 calls to help city planners understand where city resources are needed [**78, 1**]. Users of mobile services such as Google Maps and Foursquare actively contribute data about places to help others make plans around accessibility, dietary needs, and family needs.

Despite successful applications, meeting both the needs of users who contribute data and the system's goals for data collection remains a critical challenge for participatory sensing. On one hand, addressing the needs of data contributors–such as the desire for physical data collection tasks to be minimally disruptive [**131**], personally relevant [**120, 5**], and generally of value to them [**78, 107**]–is necessary for engaging enough contributors to actively make contributions. On the other hand, achieving data collection goals–such as obtaining *high-fidelity data* with detailed information about objects or events of interest at fine enough temporal and spatial scales—is necessary for understanding certain phenomena of interest and ensuring the usefulness of services that depend on such data [**30, 5**].

All participatory sensing systems must achieve some kind of balance between the needs of users and the system's data collection goals if they are to remain viable, but existing approaches are limited in how they strike this balance. *Opportunistic approaches*, where people are asked to actively contribute data along their existing routines when it is convenient for them to do so [**143, 120**], can better meet the needs of data contributors, but the opportunistic nature of the contributions makes it difficult to meet specific data needs. Contributors may not frequent certain locations, which makes it difficult to ensure high data coverage across locations and to keep dynamic data fresh. Even for locations that people frequently visit or pass by, the effort to contribute high-fidelity data can still be prohibitively high (e.g., filling out a full survey while walking around) [**78, 1**]. Low-effort opportunistic approaches can attract more contributors (e.g., [**128, 131**]), but using a low-effort approach has typically resulted in settling for low-fidelity data.

Directed approaches, where people are asked to fulfill a specific task that is not necessarily in their immediate vicinity, can be used to target specific system goals such as increasing data fidelity and coverage wherever it is needed. However, with a directed approach, meeting the needs of data contributors is difficult as it requires them to deviate from their existing routines [**114, 76**]. This disruption is offset either with monetary incentives [**68, 3, 130**], which may be infeasible or cost-prohibitive in many domains, or through dedicated volunteers [**129, 27**], who may be difficult to recruit at scale in many domains.

To overcome such shortcomings, we introduce a new hybrid approach that collects data opportunistically and uses the collected data to selectively notify people based on our understanding of the world and their interests so that we can reach out to more people to collect more data about the world. This approach uses the idea of flexible coordination to progressively build up our understanding of the world in such a way that it notifies people only when they can conveniently contribute within their routines, and directs people to places only when it contributes directly to their goals or interests. Unlike existing approaches that direct people out of their routine to meet specific data collection goals, which can cause disruption to people, our new hybrid approach offsets the “cost” of deviation with the “value” of the personally-relevant information. Unlike existing approaches that only elicit opportunistic contributions, which may fail to meet desired system goals, our approach provides personally-relevant information to meet people’s goals while finding opportunities for data collection en-route or at-location. This allows systems to become more directed or remain opportunistic depending on our current, refined, richer understanding of the world.

To realize this idea, we introduce *4X*, a framework for multi-stage data collection processes that determine effective data collection opportunities by reasoning about changing states of the world, people’s locations, and their willingness in deviating off of their routine based on our knowledge of the world. To do this, 4X models people’s interests in information about the world; understands how the current state of the world matches their interests—which in turn affects which opportunities may become available; makes decisions about which data collection opportunity to surface in a way that does not overextend their interests but that is still useful for gathering more data to have a richer understanding of the world. For example, 4X may first collect low-effort, low-fidelity opportunistic contributions when no data is available and when a user is passing by a location where they can conveniently contribute. 4X then may draw other users to places outside of their routines where the data and their interests align, and while they are en route, 4X may elicit further contributions at a place where it needs more contributions to ensure high data coverage. Instead of using a single data collection strategy regardless of situations on the ground, 4X reasons about how and when to enact certain data collection strategies based on the changing state of the world and a user's current location and interests in a way that simultaneously achieves both needs of data collectors and system goals.

To demonstrate and evaluate the effectiveness of the 4X framework, we implemented it in *LES* (Low-Effort Sensing), a low-effort sensing application on iPhone and Apple Watch that we designed to collect dynamically changing information about places and events around college campuses, such as coffee shops, libraries, and free food events. Through two user studies of LES (N = 95, N = 18), we demonstrate the advantages of 4X over opportunistic and directed approaches (Study 1), and the extent to which collected data can be used to promote additional data contributions from interested users who go out of their way (Study 2). Results from Study 1 show that 4X used collected data to create 34% more data collection opportunities without increasing reported disruption over an opportunistic approach; results further show that 4X is significantly less disruptive than a directed approach that notifies users of tasks at a distance regardless of whether there is data of interest to the user. Results from Study 2 show that 4X yielded 49% more data by directing users to locations of interest where they made additional contributions en route to and at target locations. These study results demonstrate the effectiveness of dynamic data collection processes that use multiple data collection strategies to better achieve desired outcomes as opportunities arise based on people's changing state of interest and situations on the ground.

3.2. Background

We are interested in advancing participatory sensing approaches that engage people to *actively contribute data about physical locations that they are in or are willing to go to* [**127, 85, 86, 17, 42, 70, 104**]. This is in contrast to prior work on machine-sensor based participatory sensing approaches that passively collect data from mobile device sensors or other custom sensors attached to people or objects (e.g., [**35, 6, 2, 41, 5, 134, 16**]).[1] While these prior systems can sense some attributes about the physical world, machine sensors are limited in what they can measure without active human participation.

More recent work involves humans-in-the-loop by using machine sensors to collect data and using remote crowd workers to analyze the data to derive useful information. This approach allows for sensing a wider range of phenomena, but is still limited by what machine sensors can observe. For example, Project Sidewalk asks online crowd workers to label Google Street View images with respect to accessibility issues present, like missing sidewalk ramps [**115**]. Though contributors to this project can help to interpret the collected data, this approach cannot easily capture dynamically changing information since the collected data is often stale by the time of analysis (e.g., as happens with seating availability at a coffee shop). Recent work on Zensors [**87**] is able to provide dynamic information about a location by sending a live camera feed to crowd workers and machine vision algorithms in real-time. However, this approach is

[1] This passive machine sensing approach is sometimes referred to as "opportunistic sensing" in the literature [**17, 85, 86, 42, 104**]. However, this should not to be confused with our use of the term, "opportunistic data collection approach," that describes an active, participatory sensing approach in which people actively contribute data along their existing routines when it is convenient for them.

still limited by machine vision algorithms that can only identify a narrow range of phenomena, and crowd workers who have limited access to context and who cannot move across the physical space where the phenomenon is occurring.

Given the limitations inherent in using machine sensors, we focus in this chapter on better ways to actively engage people in participatory sensing. In what follows, we highlight the limitations of existing active participatory sensing approaches that restrict their data collection process to a single data collection mechanism, and argue for the need for hybrid, multi-stage approaches such as 4X to overcome the limitations of existing opportunistic and directed approaches.

3.2.1. Existing Approaches to Participatory Sensing: Opportunistic and Directed

To engage people in active data collection, *opportunistic data collection* approaches support users who provide data when it is convenient for them to do so, as part of their existing routines or through activities they are interested in. Example initiatives engage citizens and hobbyists to report infrastructural issues such as potholes, graffiti and broken street lights in the community [**78, 1**], to track the presence of various bird species [**120**], and to answer questions about locations people are at, such as the amount of time to get through airport security [**101**]. While these examples do not require users to travel out of their routine, engaging volunteers to provide useful data still requires designing for the interests and goals of data contributors [**143**]. System designers must balance the need for collecting high-fidelity data that is valid—which may require more stringent data collection protocols and thus more effort and interest on the part of contributors—with the competing goal of recruiting many participants to advance coverage [**30**]. While successful initiatives such as eBird [**120**] have garnered large-scale use and resulted in data collected from many contributors, recruiting enough willing contributors to projects is generally still a challenge [**131**].

To promote convenient contributions from a wider base of casual users, *low-effort opportunistic data collection* approaches introduce lightweight interaction techniques that seek to minimize the effort required to contribute. One approach presents tasks during smartphone unlocks to, for example, collect coarse-grained census data [**131**] and answers to microtasks [**128**]. Other approaches infer environmental data through immersive interactions embedded into a user's habit-building practice (e.g., while going for a run) [**93**], or ask users to complete small tasks (e.g., look for a lost item; pick up and deliver a package) along their route while they are on-the-go [**77, 76**]. While these approaches may be useful for recruiting more contributors and increasing data throughput, they are generally limited to collecting low-fidelity data when high-fidelity data might be desired. Moreover, these approaches cannot direct users to help fill gaps in data coverage since data is only collected opportunistically along one's existing routine.

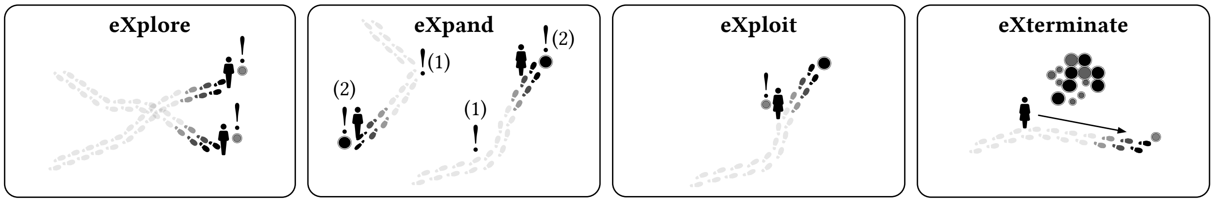
In contrast to opportunistic approaches, *directed data collection* approaches actively direct users out of their routine towards tasks and areas where data is needed to target specific gaps in coverage or fidelity. Due to the deviation required from a user's routine, directed approaches require high incentives which some prior works address by providing monetary rewards in return for a data contribution [**68, 3, 130**]. While paying participants allows for rich forms of data collection, the costs may be prohibitive for scaling many services, particularly those that require active monitoring or tracking dynamically changing information (e.g., campus events; the state of city infrastructure).

In the absence of monetary rewards, other directed approaches use gamification to align data goals with game mechanics; they direct users to collect useful data and travel as a byproduct of gameplay. For instance, PhotoCity recruited users to take photos of target areas from specific angles in order to make 3D models of buildings [**129**]. A contemporary example is Niantic's Ingress Prime, which pits players against rival factions to create an immersive game environment in which dedicated gamers travel out of their existing routes to take in-game actions that produce crowdsourced data, such as walking routes, as a byproduct. While effective for some use cases, designing appealing game mechanics that address a wide range of desired data collection goals remains challenging.

3.2.2. A Conceptual Framework for A Hybrid Approach to Participatory Sensing

Prior systems either opportunistically or directly ask users to complete tasks based on their convenience [**131, 128**], domain interests [**120**], or system's data collection needs [**3**], and in doing so, unnecessarily restrict their data collection process to a single data collection mechanism. Our work on 4X, in contrast, leverages the benefits of both opportunistic and directed approaches, allowing for the collection process to adapt based on the state of available data and user interest. This enables the collection of high-fidelity and high-coverage data while being minimally disruptive to users and providing them data of interest. Specifically, our approach considers ways to strategically gather initial pieces of dynamically changing, location-specific data that is used to attract users out of their way to locations of interest and inspire further data contributions, while other works focus on only using pre-existing data to recommend locations [**29**]. Unlike existing opportunistic data collection approaches that stop at using collected data as feedback to users when they contribute data (e.g., to see how their responses align with the rest of the community [**131**]; to know which nearby birds they have tracked so far [**120**]; to see summaries of collected ESM data to increase compliance [**62**]), 4X actively uses data collected by some users to provide relevant information of interest to other users, so as to promote further contributions. For instance, 4X may notify users about an event of interest (e.g., free food; or the sighting of a bird species) or relevant conditions of interest (e.g., an available table by the power outlets at the coffee shop) to draw them to locations where data gaps exist but the currently collected data still aligns with their interests. In this way, 4X implements a directed data collection approach that offsets *the "cost" of deviation with the "value" of the information provided to the users.* This allows systems to present opportunities that are likely to be in people's interests and that may fit within their routines and goals, while still finding opportunities for data collection to advance desired system goals.

Multiple prior works have explored technical frameworks for building context-aware and participatory sensing systems. Sensr [**74**], AWARE [**37**], Ohmage [**123**], and PartS [**94**] provide end-to-end frameworks for building general purpose participatory sensing platforms for collecting, managing, and analyzing prompted self-reports and sensor data streams, and for enabling researchers to run user studies with the platforms. In contrast to these technical frameworks for implementing data collection strategies, our contribution is a novel *conceptual framework for designing a multi-stage data collection process that becomes more directed as data of interest is collected*. 4X is a necessary complement to these existing technical frameworks because building a multi-stage data collection protocol requires making non-trivial design decisions about how and when to enact certain data collection mechanisms depending on where people are, what data is available, and what people care about that other works have not considered or evaluated.



There is free food here! What kind of food it is?

(a) users opportunistically contribute new landmark locations and contribute at tracked locations they pass when notified along their daily routine.

There is free pizza nearby! Would you like to go?

(b) collected data is used to draw users to target locations of interest outside of their daily routine (1), where they are asked for additional contributions if they go (2).

Are there any open spots on the bike rack here?

(c) users are asked to contribute at locations they pass en route to the target location.

User Responds: "No more pizza left here."

(d) once a location has a sufficient amount of data, users are directed to other locations in need of data.

Figure 3.1. The 4X framework scaffolds data in four stages. People first contribute opportunistically, marking new landmarks for tracking and responding when queried for information as they pass existing landmarks (eXplore). Collected data then draws people to target locations outside of their path, where they can be queried for additional contributions (eXpand). People may also be queried en route to the target location for contributions (eXploit). As data at a location fills (indicated by the size and the darkness of circles), users are directed to other regions and locations in need of contributions (eXterminate).

3.3. 4X Framework

We propose a hybrid, multi-stage data collection approach for participatory sensing called 4X (eXplore, eXpand, eXploit, eXterminate) that collects low-effort, low-fidelity data opportunistically, and then uses these data to direct users to locations of interest to make additional contributions that build data fidelity and coverage; see Figure 3.1. Each stage of 4X aims to create more opportunities for data collection than opportunistic approaches without causing significant disruption to the user as directed approaches might, allowing for the building of fidelity and coverage. In this section, we describe the stages of the 4X framework and highlight key design decisions associated with implementing each stage. [2]

[2]We use the term "stages" to refer to the various data collection mechanisms that 4X uses as data becomes available and is collected, but do not mean to imply that only a single stage can occur at a time. For example, given data collected thus far, one user may contribute additional data opportunistically (via eXplore) while another is drawn from their routine to the location to contribute data (via eXpand). So while it may be useful to think about the data

3.3.1. eXplore: Collecting Opportunistic Contributions

The eXplore stage (Figure 3.1a) allows users to contribute opportunistically, both as a primary responder to indicate the presence of an event or location (e.g., there's someone giving away free food here) and as a secondary responder to contribute additional data when notified *at-location* during their daily routine (e.g., there is someone giving away free food here–can you tell us what kind of food it is?). Primary responder contributions act as landmarks, allowing for more data and interest to build up around specific locations and events. Subsequent contributions from secondary responders scaffold the data collection to increase data fidelity.

During the eXplore stage, contributions are collected opportunistically from those who can contribute conveniently along their existing routine. To limit disruption to users' existing routines, participatory sensing applications that instantiate the eXplore stage should make it easy for users to make low-effort contributions. This can be supported by low-effort interaction techniques for primary and secondary responders that allow contributions to be made in seconds (e.g., [**128, 131**]). Collected low-fidelity contributions from individual users can then be scaffolded and combined to build data fidelity across multiple contributions and users. Together, these techniques allow individual users to contribute conveniently while still building the needed data fidelity to meet system goals, so that the eXplore stage (and 4X more generally) can collect convenient contributions from larger groups of casual users rather than relying on small groups of dedicated volunteers.

collection process as proceeding in stages, the 4X framework does not preclude the possibility of multiple stages occurring at the same time.

3.3.2. eXpand: Directing Users to Places of Interest

The eXpand stage (Figure 3.1b) uses *collected data to present opportunities at nearby locations that are likely to be in people's interests and that may fit within their routine and goals.* When the collected data aligns with a user's interest, eXpand sends an *at-distance notification* (e.g., there is free food in a nearby building, would you like to go there?) in order to build data fidelity at the *target location of interest*. Subsequently collected data can then be used to draw in other interested users. In this way, instead of only relying on opportunistic contributions from eXplore to build data fidelity, eXpand seeks additional, directed contributions from those who do not immediately pass by target locations.

Unlike directed approaches that indiscriminately notify users to make contributions outside of their routine, 4X *selectively notifies* users about nearby locations only when data of interest is available to motivate them to deviate from their routine. Providing users with data that they find valuable helps to offset the cost of deviating from their routine, which limits perceived disruption since users value the information and may wish to go out of their way to the location of interest. Through selective notifications, eXpand creates new contribution opportunities as a byproduct of surfacing existing data of interest back to the user while limiting the perceived disruption since users are receiving information of interest.

Implementing the eXpand stage of 4X into a participatory sensing application requires (1) building a *user model* that captures individual users' interest in the collected data, and their willingness to go out of their way; (2) setting a *notification selection criteria* that determines whom and when to notify a user based on their interests (i.e., their user model); and (3) setting a *question selection criteria* to determine what additional information to solicit from users when they do decide to visit locations upon being notified. We discuss each of these components below:

(1) The **user model** determines how likely a user is to go out of their way when presented with information about locations and events of interest to them. A user model may consider (a) the kind of location (e.g., free food events; coffee shops) and the types of data about a location (e.g., food type; seating availability) that a user is interested in; (b) the amount or specificity of available information for each kind of location needed to draw in a user (e.g., private seating versus private seating by the outlets at a coffee shop); and (c) any contextual factors that may affect a user's likelihood to deviate from their routine, such as their schedule or how far they would be willing to go out of their way to visit a location of interest. To create a user model, a 4X system designer may ask users to directly express what information they find valuable and what they would want to be notified about, or train a model using machine learning to predict how likely a user is to go out of their way based on observations of their past decisions to act (or not) on data of interest presented to them.

(2) The **notification selection criteria** is used to decide whether to notify a user about a location or event of potential interest. Based on the collected data and the user model, this criteria considers the likelihood of users to go out of their way and determines how much available information is sufficient to notify a user, and at what distance. While it is generally advisable to notify users with information that is interesting or valuable to them, the notification selection criteria can be conservative and specify only sending information that is extremely valuable to the user (e.g., free food that a user really likes), or be aggressive and specify sending all information of potential interest. Likewise,

the criteria may specify sending information of interest to users only when they are a short walk from the location, or sending information when a much larger deviation is needed. These decisions broadly affect how many data collection opportunities are made available, and how valuable or disruptive users may find the notifications sent via eXpand.

(3) The **question selection criteria** is used to decide what piece of information to solicit from a user when they do decide to go to a target location of interest that is suggested by eXpand. The question selection criteria can be used to prioritize gathering information that is valuable to users generally (e.g., more users want to know if there is seating near power outlets than near windows), which provides direct value to those interested and opens up future data collection opportunities from those who visit the location. The criteria can also be used to prioritize data collection goals that are valued by the system or that better meet the needs of a particular subset of the users (e.g., collecting information about accessibility features).

3.3.3. eXploit: Creating Contribution Opportunities En Route

The eXploit stage increases data coverage by collecting data at places that a user now passes *en route* to a nearby eXpand target location and *around the target location* itself that a user would previously not have gone to (see Figure 3.1c). Like eXpand, opportunities for additional contribution in eXploit are presented as a byproduct of the user choosing to go to a target location. As these contributions are collected opportunistically like eXplore, the design decisions discussed there to make contributions convenient and low-effort similarly apply here.

While en route to a location, eXploit can request contributions from a user about tracked locations they pass to increase data coverage in regions between the user's original location and their target location. Since the effort to make a contribution is low and since users are already going to a location of interest based on information that the system provided to them, we expect that asking for these additional contributions will not be perceived as disruptive by the users.

After a user has reached the target location, eXploit can also ask users to make contributions in the region around the target location to further build data coverage. For example, birders contributing to the eBird [**120**] citizen science initiative may be willing to contribute information about other birds in the area after being notified about a bird they are interested in seeing at a nearby location. In this situation, it is advisable to have the contribution opportunities align with things that users are interested in contributing to, since some (small) deviation may still be involved.

In summary, we hypothesize that through the eXpand and eXploit stages, 4X increases the number of contribution opportunities and the actual number of contributions relative to the eXplore stage alone, without increasing disruption since eXpand only occurs when data of interest is available and the interactions for data contribution during the eXploit stage remain low-effort. Together, these two stages can increase both data fidelity and coverage, as users contribute data at, en route to, and around specific locations of interest.

3.3.4. eXterminate: Shifting Focus to New Regions and Locations

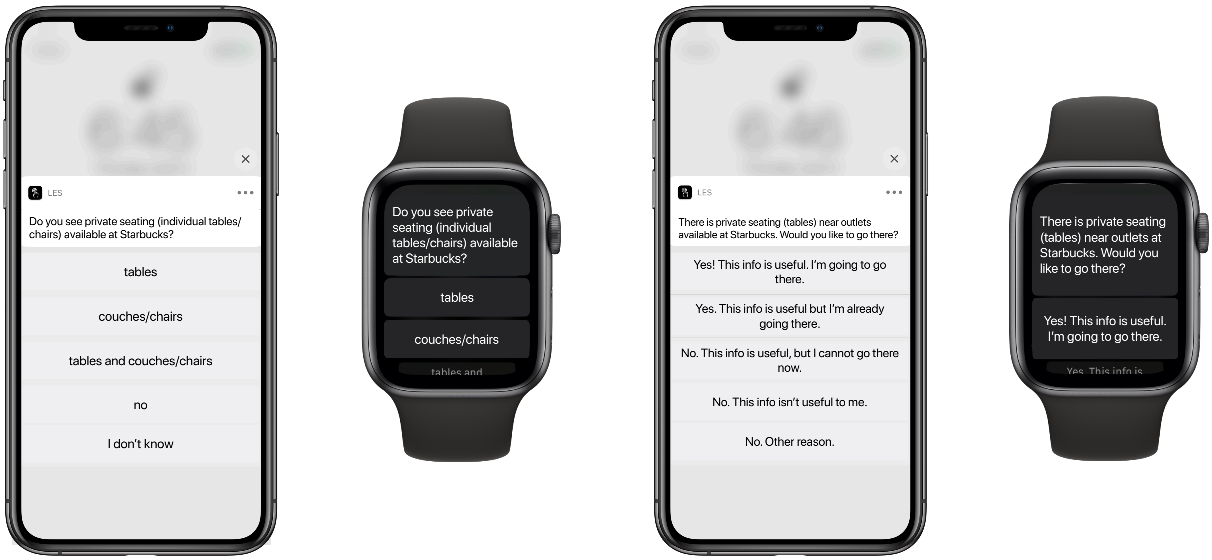
Finally, the eXterminate stage is used to determine if tracking for a location should restart or cease based on the current state of the collected data (see Figure 3.1d). We consider two different conditions for eXterminate: *data staleness* and *data fidelity thresholds*.

When eXterminate detects that data at a certain location has become stale and may no longer be valid, data scaffolds are cleared and data collection at the location begins anew. In cases where users want to know about dynamically changing data at different locations, eXterminating when the data is stale helps 4X to ensure that users are provided with up-to-date information that may motivate them to deviate from their routine. To determine that data has become stale, a participatory sensing system may request information from users to help determine whether data has become invalid. For example, a system tracking free food events may ask “Is there still food available?” to determine whether to continue tracking.

If instead eXterminate identifies that the collected data has reached a data fidelity threshold for a location (i.e., most or all of the desired information has already been collected), tracking at the location stops in order to promote data collection efforts at other locations. This helps a 4X system to work towards addressing other pressing data needs rather than over-emphasizing data collection efforts at places where they are no longer needed. To implement such a threshold, applications may consider including a flag in the data scaffold which indicates to eXterminate that enough data at a location for a certain time period has been collected. A simple example of this is checking to see if all questions in the data scaffold have already been answered by users.

3.4. LES: Campus Data Collection with 4X

Having detailed the 4X framework, we now present *LES* (Low-Effort Sensing), a location-based participatory sensing system built for the iPhone and the Apple Watch that instantiates the 4X framework to collect dynamic information about local places and events around college campuses from mobile crowds. In this section, our goal is to illustrate how the conceptual 4X framework might be implemented into a practical application. We describe (1) what LES is; (2)



(a) Queries for additional data contributions contain any needed context to answer the question, and the possible categorical responses for the query.

(b) Notifications about nearby places of interest are sent to users when the available data aligns with their data interests (as reported in the pre-installation survey).

Figure 3.2. LES interfaces for contributing data and for receiving data of interest. Users respond to simple queries about locations they are in or are about to pass (a), and receives valuable information about nearby places of interest that they care about (b). All responses can be made through contextual notifications on the device's lock screen without needing to open the installed application.

how LES instantiates each stage of the 4X framework; and (3) the technical implementation of LES.

3.4.1. What is LES?

LES is an application for sharing dynamic information about places and events on college campuses. With LES, students can contribute information about coffee shops, libraries, gyms, and free food events such as seating or equipment availability and what kind of free food is being given out. Users are notified and asked to contribute data about tracked locations as they are about to pass the location using a low-effort interaction that can be performed from the lock screen of their phone. They can also receive information of interest to them about nearby locations (e.g., open seating by the power outlets at their favorite coffee shop) when it is available. Unlike prior work on TASKer [**68, 69**] that studied collecting information about campus events by paying students to make contributions, LES is designed to collect convenient, low-effort contributions on users' existing routes, and to use people's interest in events and situations (e.g., free ice cream on a hot day; open seats at a coffee shop) to promote contributions at places outside of their current route.

3.4.2. Instantiating 4X in LES

In this section, we detail how we instantiate each stage of the 4X framework in LES for collecting data about locations and events on college campuses.

3.4.2.1. eXplore in LES. To instantiate the eXplore stage of 4X, LES notifies users to make opportunistic contributions when they are about to walk past locations tracked by the application. To limit disruption when making these contributions, LES implements contextual notifications that appear on the device's lock screen and can be responded to without needing to unlock the device or open the application when the user is about to pass by a tracked location; see Figure 3.2. Each notification contains the relevant contextual information derived through earlier responses by other LES users, a query for additional information about the location in the notification body, and the categorical responses to answer the query when the notification is opened; see Figure 3.2a. In addition to the relevant categorical responses for the query, we include the "I don't know" response as a way for users to say that they received the notification, but were unable to answer it; this response was included based on early user testing where users said they wanted a way to still respond to the notification if they could not answer, instead of simply ignoring it. Such contribution techniques have been proven to be low-effort in prior works such as Slide to X [**128**] and Twitch [**131**].

To create higher-fidelity data from low-effort contributions, LES uses *information scaffolds* that deconstruct the desired, richer information into low-fidelity components with corresponding questions that can be easily responded to with low-effort contributions. LES includes data scaffolds for a variety of locations on and around college campuses, including libraries, gyms, coffee shops, and free food events. These locations and their associated questions were informed by needfinding surveys of target users where respondents indicated what places around the campus they would like to know information about, and what information they would like to know. Rather than only capturing coarse information about the tracked locations (e.g., that private seating at a coffee shop is available), we chose to have our information scaffolds capture more detailed information (e.g., that private seating near the windows at a coffee shop is available) that could be useful for meeting the different information needs of users during the eXpand stage.

As an example, LES would use the following scaffold and question breakdown for collecting data about coffee shops with the question ordering being 1, 1a, 1b, etc.

There is *private seating [tables/couches/chairs]* near *[outlets, windows]* and/or *shared seating [communal tables]* near *[outlets, windows]* at Starbucks.

(1) Do you see *private seating (individual tables/chairs)* available at Starbucks? [Tables, Couches/Chairs, Tables and Couches/Chairs, No]
   (a) Are any of these near outlets? [Yes, No]
   (b) Are any of these near the windows? [Yes, No]
(2) Do you see *shared seating (communal tables)* available at Starbucks? [Yes, No]
   (a) Are any of these near outlets? [Yes, No]

(b) Are any of these near the windows? [Yes, No]

If any LES user responded “No” to either questions 1 or 2, then the following sub-questions would not be asked to other users. When a user is presented with a notification, all information is included in the notification to help direct them to what specific information is needed by the system. For example if there were private tables available, but not near outlets, a user passing the tracked Starbucks would be asked, “There is private seating (tables) available at Starbucks. Can you tell us if there are any near the windows?”

3.4.2.2. eXpand in LES. Once LES has collected some data opportunistically, it begins the eXpand stage by selectively notifying users when their interests align with the collected data. LES does this in two phases: it first sends an initial notification to users letting them know of information about a nearby location of interest and asks if they would like to go to the location (e.g., “There is private seating (tables) near outlets available at Starbucks. Would you like to go there?”; see Figure 3.2b), and then sends a second notification to users who do decide to go to the target location to request additional data there (e.g., “There is private seating (tables) available at Starbucks. Can you tell us if there are any near the windows?”).

As discussed in the previous section, implementing the eXpand stage of 4X requires forming a user model, setting a notification selection criteria, and setting a question selection criteria. We discuss how we implemented these components in LES and the associated design decisions below:

(1) **User model:** LES implements a simple model of user interest that focuses on what data users find valuable (e.g., a user wants to know about private seating at a coffee shop). To develop this model, we first conducted needfinding to determine the kinds of campus locations and types of data about these locations that students are generally interested

in. We then asked each LES user to fill out a pre-installation survey that asked them, for each type of data available for each kind of location, if they would like to receive notifications when such data is available (see each of the study sections for more details on the pre-installation survey).

(2) **Notification selection criteria:** Once LES has collected some data from opportunistic contributions, it selectively notifies users if the collected data matches their user model (i.e., if the data matches what the user indicated what they wanted to be notified for in the pre-installation survey). Although 4X can notify users at various distances during eXpand based on level of their interest, we chose to notify nearby users who are within a set distance radius of the target location (e.g., 300 meters; see Technical Implementation for more details).

(3) **Question selection criteria:** LES prioritizes collecting the types of data that users are more interested in knowing about; such information is more valuable to users, and also more likely to lead to additional data contribution opportunities from the larger group of users who may wish to go to the location should the collected information be of interest to them. To do this, LES rearranges the information scaffolds presented earlier to first ask questions that collect the data that most users of the system would be interested in knowing, based on the pre-installation surveys used to build the earlier user models. For instance, if we found that more users expressed a preference in knowing about shared seating rather than private seating, LES would ask question 2 and its sub-questions before asking question 1.

3.4.2.3. eXploit in LES. To implement the eXploit stage of 4X in LES, we consider how LES may notify and request additional contributions from users en route to a nearby eXpand target location. To collect additional data en route, LES notifies users to make contributions using the same interaction technique as used for eXplore when they pass tracked en route locations. To provide users with opportunities to make contributions en route, LES includes additional locations like bike racks and parking lots that users may frequently encounter while on their way to an eXpand target location.

While eXploit can also be used for requesting contributions around a location of interest that a user is going to, we chose not to implement this use case in LES because (a) the locations of interest were mostly spread out across the university campus; and (b) in our domain, users interested in one location did not often have an interest in visiting nearby locations.

3.4.2.4. eXterminate in LES. To instantiate the use of eXterminate to handle situations when data has become stale, LES uses *timed refresh cycles* and *verification questions* to signal when data scaffolds are no longer valid and should be cleared. Timed refresh cycles act as the base reset condition for data scaffolds to ensure that the presented data is still correct and up-to-date, and trigger based on how long the current data has been in the scaffold. If any data has been in the scaffold for longer than a specified threshold time, the scaffold is fully cleared. Then depending on the type of location, tracking of the location will either restart at places that continue to generate new data (e.g., coffee shops; libraries) or cease for events that are no longer happening (e.g., free food events). When no further questions can be asked from a data scaffold, LES uses a verification question to see if the current data is still correct. If a user responds that the data is no longer correct (e.g., no more free food left), LES clears the data scaffold and either restarts or stops tracking similarly to the timed refresh cycle.

While eXterminate may also use thresholds for data fidelity to shift data collection focus to other locations, we chose not to implement this feature in LES since there was not a need to promote or distribute data collection at all locations evenly. In other words, we let LES collect data at places where people naturally passed and where people were interested in knowing about the data.

3.4.3. Technical Implementation

LES consists of a client application written in Swift for iOS and Watch OS, and a back-end built using Node.js and MongoDB. The iOS front-end handles real-time outdoor and indoor location tracking using geolocation and Estimote Bluetooth iBeacons respectively, and the generation of contextual notifications using Apple's UserNotifications framework. When near a tracked region, the front-end presents users with a notification containing the current information about the location, the query for the next piece of information, and the possible answers to the query (see Figure 3.2a); similarly, the front-end will notify users about nearby places of interest within the distance threshold specified by the notification selection criteria (see Figure 3.2b). Responses to notifications are sent to the back-end that handles all tasks related to building data scaffolds, generating notification contents and queries, and syncing information with all users' applications in real-time.

As implemented, LES sets a notification radius of 300 meters (approximately 1.5 blocks in the deployed city) for eXpand, so that the radius will be large enough to require users to make some deviation from their routine, but not so far that the users would never go out of their way. In order to avoid spamming users with multiple notifications when they pass through areas with multiple tracked locations in close proximity, we additionally set a 10 minute notification interval between notifications so that once notified about a data collection opportunity via eXplore or eXpand, the user will not receive another notification within the threshold.

3.5. Study 1: Comparing 4X to Opportunistic and Directed Approaches

We present in the following sections two user studies of LES that demonstrate the advantages of the hybrid, multi-stage 4X framework over purely opportunistic and directed data collection approaches (Study 1), and the extent to which collected data can be used by 4X to promote additional data contributions from interested users who go out of their way (Study 2). In both studies, we are interested in understanding how the process of providing people with specific information that is of interest to them as it is dynamically acquired can better meet data collection goals while avoiding unnecessary disruption to users.

In Study 1, we compare 4X to an Opportunistic data collection approach and a Directed data collection approach to study how 4X might better meet the needs of users (e.g., convenience and low disruption) than a Directed approach, and provide more opportunities to meet the needs of the system (e.g., data collection goals) than an Opportunistic approach. Specifically, our Opportunistic approach aims to meet user needs by only asking for contributions when users can conveniently contribute along their existing routine. We expect that this will lead to some collected data, but also miss opportunities to collect more data when people would have been willing to visit locations outside of their routine. Meanwhile, our Directed approach aims to meet system data collection goals by asking users to contribute data at any tracked location they might be nearby, which they may or may not pass as part of their existing routine. This approach can potentially lead to more contributions from a larger set of users, but at the cost of greater disruption due to more notifications and deviation from users' routines. We expect that 4X will provide users with more data contribution opportunities than the Opportunistic approach by also notifying them about nearby tracked locations, but will be less disruptive than the Directed approach, since requests to travel to locations outside of their routines are only made when collected data suggests that they may be interested in knowing. To summarize, we hypothesize that:

H1: 4X creates more opportunities for data collection over an Opportunistic data collection approach without increasing disruption.

H2: 4X is less disruptive than a Directed data collection approach and sends significantly fewer notifications.

3.5.1. Method and Analysis

3.5.1.1. Participants. We recruited 95 undergraduate and graduate students of a mid-sized U.S. university through mailing lists, social media, and word of mouth. Participant ages ranged from 18 to 28 (M = 20.30, SD = 1.93), with 72 female and 23 male participants. The study took place over 14 days, during which participants completed a pre-study survey, used LES as a part of their daily lives, and completed a post-study survey. We compensated participants $20 for their time spent on surveys and installing LES, but did not incentivize their behavior during the study (i.e., no monetary incentive was provided for task completion).

3.5.1.2. Procedure. The 95 study participants were randomly assigned to one of three study conditions: Opportunistic (32), Directed (32), and 4X (31).[3] We chose a between-subjects design instead of a within-subjects design to: (1) avoid any carryover effects on our measures when users switched between conditions since each condition only had subtle, hard-to-notice

[3]We removed 1 participant from the 4X condition because they received far more notifications than should have been allowed due to a temporary technical malfunction. This left us with log data from 94 participants (30 for 4X; 32 each for Opportunistic and Directed). 9 of these users did not fill out the post-study survey (2 for 4X; 3 for Opportunistic; and 4 for Directed), leaving us with survey responses from 86 participants (29 each for Opportunistic and 4X, and 28 for Directed).

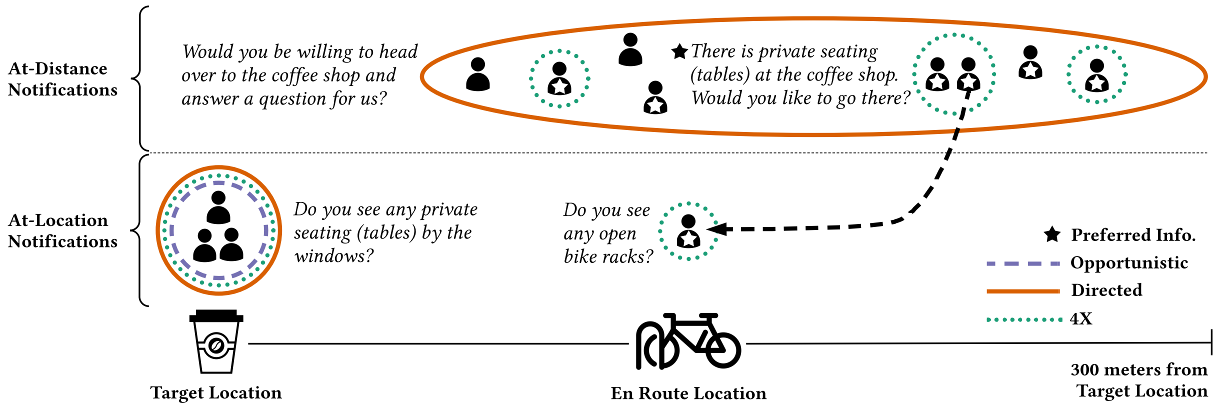


Figure 3.3. Illustration of the notification policies for the Opportunistic, Directed, and 4X conditions in Study 1. In all conditions, users receive at-location notifications that ask them to contribute additional information when they are about to pass by a tracked task location. Users in the Directed and 4X conditions also receive at-distance notifications up to 300 meters away from the task location. The Directed condition sent these notifications regardless of whether preferred information is available (information is included when available), whereas the 4X condition only sent these notifications if preferred information is available. A star on a person indicates that preferred information of interest is available for that user at the time of notification. Finally, 4X also asks users to make contributions en route when they decide to go to the target location.

differences in the notification policy; (2) control and mitigate weekly variability in routines; and to (3) avoid study fatigue associated with a necessarily longer within-subjects study.

Figure 3.3 summarizes the notification policies of each condition, detailing when users would receive notifications to contribute additional information based on the condition they are in. In all three conditions, users receive *at-location* notifications as they pass by landmark locations (eXplore). Users in the 4X condition additionally receive *at-distance* notifications that ask them to go to landmark locations when there is data of interest about the location available (eXpand); they are also asked to contribute data about other locations *en route* should they decide to go (eXploit). Users in the Directed condition receive at-distance notifications similar to 4X if data of interest is available, but would also receive a generic contribution notification (e.g., “We need some information about the Starbucks nearby. Would you be willing to head over and answer a question for us?") anytime they were near a landmark location with no data of interest available. In both the 4X and Directed conditions, all at-distance notifications are sent when users are within 300 meters of a landmark location. In all conditions, outside of the notification interval set by LES to avoid notification spamming, notifications are sent based on the notification policies of the 4X framework and the data collection approaches used in the Opportunistic and Directed control conditions.

We pre-populated each instance of LES with seven landmark locations around campus and the surrounding area (three coffee shops, two workspaces, and two gyms). We selected categories and locations with dynamically changing status that students were interested in knowing about (e.g., when tables or gym equipment become available) based on previous needfinding of university students. In addition, one of the authors hosted free food events around campus every other day at various times over the course of the study, which served as additional landmark locations. Data scaffolds for each of these locations start out empty, and can go up to five levels deep as users contribute reports to build up data fidelity. Scaffolds were cleared of information either after four hours, or when users reported that the data is no longer correct (eXterminate).[4] Beyond these landmark locations, we added six locations of bike racks and parking lots where users in the 4X condition were asked if free spaces were available when they passed by en route.[5]

To present information of interest to users, we collected through a pre-study survey each users' high-level preferences over the kind of information they were interested in being notified about. For example, a participant who is generally interested in private tables near windows at

[4]Four hours was used as the refresh time since we could not account for scheduling and mobility patterns of users in this study, and we wanted enough time for data to build in fidelity so that it could be used in the eXpand stage.

[5]Per day, a user may be asked to contribute information at multiple locations, but never more than once per location (within a data refresh cycle) and at only one location en route to avoid over-notifying users.

a coffee shop but not in communal tables can specify that they are interested in private tables, and in sitting by the window. Whenever preferred information is available, it is included in any notifications that users in the 4X and Directed conditions receive. For example, a user interested in coffee shop seating may receive a notification at-distance that reads: “There is private seating (tables) near outlets at Starbucks. Would you like to go there?” Users in the Opportunistic condition were not notified of information they may care about, but can access such information via a “For You” page within LES; this page was also available for 4X and Directed condition users. In cases where no preferred information is available, the Directed condition simply asked users if they would be willing to head to a nearby landmark location to contribute some information; see Figure 3.3.

We used user responses to at-distance notifications as an ecological momentary assessment (EMA) [**119**] to assess (a) whether they found the information presented to them useful; and (b) whether they were going out of their way based on the information presented or if they just happen to be already going to a landmark location; see Figure 3.2b. Users can also specify reasons for not going, such as having a scheduled event.

Beyond EMAs, users completed a post-study questionnaire following the 14 day usage period, which asked them to reflect on the disruption and value of LES during their usage, recall times when they were asked for contributions and elaborate on why they did or did not respond, and recall times when they were notified to go to locations outside of their routine and reflect on why they may or may not have gone. We coded these open-ended survey responses along the dimensions of (1) ease of contribution; (2) the value users received in getting notified about data of interest; and (3) the disruptiveness of the notifications received. Codes were aggregated into counts for each condition to present overall trends related to these dimensions. We were interested in these particular dimensions as they directly map back to the key design goals of 4X, namely (1) broadening participation with low-effort contributions; (2) increasing data fidelity and coverage by drawing users from their routines with data of interest; and (3) limiting disruption when asking for users to deviate or make additional contributions.

3.5.1.3. Measures and Analysis. We measured how disruptive users found each condition by asking how frequently they felt disrupted by LES notifications over the two-week study period on a 5-point Likert Scale (1: Never, 5: Always) in the post-study survey. To show that 4X creates additional data collection opportunities, we compared the number of notifications sent to users in the 4X condition at-location and at-distance.

We used the EMAs and post-study surveys to better understand why users decided to contribute or not, both at-location and at-distance, across the three conditions. To track specific cases where users in the 4X and Directed conditions went out of their way and contributed additional data, we used the EMA responses and log data to determine, respectively, (a) people's decisions to go out of their way; and (b) whether they went to locations they were notified about and made additional data contributions. We count all user responses at-location as actual data contributions, with the exception of when they select the "I don't know" option.

3.5.2. Results

Results show that LES effectively scaffolded low-effort contributions to build higher-fidelity data. During the 14 day deployment, the 94 users in Study 1 made a total of 705 data contributions. Of these, 224 contributions (31.77%) were subsequent contributions after the initial that were used to build data fidelity. While the initial contributions informed users when seating became available at libraries and when free food was available, subsequent contributions allowed users

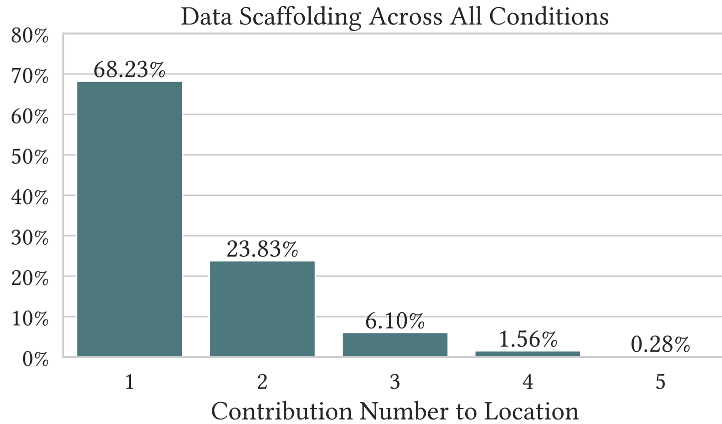


Figure 3.4. Percentage of users' contributions that built up data fidelity by scaffolding together multiple contributions beyond the first. 31.77% of all contributions made by users increased data fidelity beyond the first contribution (e.g., private tables available at a coffee shop), leading to higher-fidelity data from the scaffolded low-fidelity contributions (e.g., private tables near outlets available at a coffee shop).

to know that there was private seating near windows and that there was free pepperoni pizza. Figure 3.4 shows the frequency at which user contributions built up data fidelity by integrating multiple contributions beyond the first.

Across all conditions, users were notified a total of 2586 times with an average number of notifications per user per day of 1.18 for Opportunistic, 1.42 for 4X, and 3.26 for Directed. Table 3.1 breaks down the number of notifications sent in each condition, both at-location and at-distance. Users responded with information (i.e., made a valid contribution) 51.27% of the time when they were notified at-location to contribute additional data (53.32% for Opportunistic; 49.78% for 4X; and 50.25% for Directed). In post-study surveys, users cited ease of interaction as their primary reason for contributing at-location. Across all conditions, 47 out of 86 users (54.65%) mentioned ease as the main reason for contributing while at-location. For example, P17 (4X) wrote: *"I responded because I was skateboarding by [coffee shop], and it wasn't too*

Table 3.1. Breakdown of contributions and notifications sent by condition in Study 1.

	Opportunistic	4X	Directed
Valid Contribution Origin			
At-Location	281	220	187
At-Distance (Willing to Deviate)	–	1	5
At-Distance (Already Going)	–	1	10
Total Valid Contributions	281	222	202
Notifications Sent			
At-Location	527	446	402
At-Distance	–	152	1059
Total Notifications Sent	527	598	1461

hard to slow down, glance through the windows, and respond on my phone." Some users also mentioned contributing because they felt that their contribution would be helpful to others using LES, like P22 (Opportunistic): *"It was simple enough to respond, and I went into [coffee shop] to look for myself anyway. I figured it would be good to share this information with other people since I would want to know myself."*

3.5.2.1. H1: 4X Creates Additional Contribution Opportunities Without Adding Disruption Over the Opportunistic Approach. Our results show that 4X created more data collection opportunities without adding disruption over the Opportunistic approach. 4X created 34.09% more data collection opportunities by presenting users with information they wished to be notified about; beyond the 446 notifications sent at-location (1.06 notifications per user per day on average), 4X sent an additional 152 notifications at-distance (0.36 per user per day on average); see Table 3.1. Despite sending more total notifications than the Opportunistic condition (598 vs. 527), users did not find 4X to be more disruptive; the average reported disruption was 2.08 (SD = 0.63) for 4X and 2.10 (SD = 0.61) for the Opportunistic condition. A Mann-Whitney U Test showed that the distributions of the reported disruption were not significantly different (n_1 = 29, $n_2 = 28$, U = 412.000, p = 0.925).

A plausible explanation for why users did not find 4X to be more disruptive than the Opportunistic condition despite the additional notifications is that the 4X users generally found LES to be more valuable to them than the Opportunistic users because they received notifications with information of interest to them. 12 out of 29 users (41.38%) in the 4X condition mentioned how they liked being presented with data of value to them, and some, like P21, indicated how it would help them decide where to go: *"I liked how easy it was to be helpful and the idea that I can check to see if there are free places for me to do work before I actually went to those places."* In contrast, users in the Opportunistic condition did not receive notifications with information of interest to them, which some users wished they did. For example, P26 said: *"I would have liked if the notifications I received were instead to inform me of something that I wanted rather than asking me for inputs. Especially if there was free food somewhere, I would have liked to receive a notification about it."*

Despite creating additional contribution opportunities and presenting data of interest to users, in Study 1 only four users in the 4X condition deviated from their routes, which yielded only one additional data contribution. We identified three partial confounds that may have led to this outcome, which we will address later through Study 2. First, Study 1 was conducted during a period of the school year where many users had exams and other scheduling constraints, which limited the ability of users to go out of their way. In 35.53% of responses to at-distance notifications, 4X users indicated that scheduling constraints prevented them from visiting the location even though they found the information presented useful. Second, users experienced frigid and stormy weather conditions throughout the study period, which likely lessened their desire to go out of their way to any location. Third, from the post-study survey results, we found that in some cases the pre-study survey did not sufficiently capture users' notification preferences and thus sent them information that was not valuable to them. For example, while the pre-study survey allowed users to specify that they are generally interested in coffee shops, they could not specify that they only cared for a particular coffee shop and preferred not to be notified about others. Examples such as this contributed to the 18.42% of notifications sent to 4X users who indicated that the information presented was not useful to them. In Study 2, we will show how removing these confounds and improving our notification preference survey led to a significant number of additional contributions from 4X users who go out of their way upon receiving at-distance notifications containing data of interest to them.

3.5.2.2. H2: 4X is Less Disruptive than the Directed Approach. Having demonstrated some of the potential advantages of using 4X over the Opportunistic approach, we turn to compare 4X to the Directed approach. We found a significant difference in reported disruption; the average reported disruption was lower for 4X (M = 2.08; SD = 0.63) than for the Directed condition (M = 2.54; SD = 0.86). A Mann-Whitney U Test showed that the distributions of reported disruption were significantly different ($n_1 = n_2 = 29$, U = 280.500, p = 0.023).

A primary cause for the increased disruption in the Directed condition is the number of additional notifications sent. The Directed condition sent users nearly 7 times more at-distance notifications than 4X did (1059 vs. 152) and 2.5 times more notifications overall (1461 vs. 598); see Table 3.1. On average, the users in the Directed condition received a total of 3.26 notifications per user per day versus 1.42 for the 4X condition. 16 out of 29 users in the Directed condition (55.17%) mentioned that they did not like the quantity of at-distance notifications they received without any data of personal interest. P5 noted: *"Realistically, I'm not going to go out of my way* *to answer a question. I know this is out of self-interest and it deters from everyone's experience, but being honest."* This suggests that while directed approaches can reach more users, in practice they might not lead to the kind of system that is desirable or sustainable due to the increased disruption and lack of value for attracting users to make additional data contributions without providing additional incentives.

3.5.2.3. Summary of Study 1 Results. In summary, results of study 1 demonstrate that we can use the idea of flexible coordination to enable a hybrid approach that captures more data collection opportunities that may arise as situations on the ground change, and finds good opportunities within people's changing state of interest in a way that meets data collectors' needs (e.g. without causing much disruption).

3.6. Study 2: 4X Yields Additional Contributions

Study 1 demonstrated some of the potential advantages of 4X over Opportunistic and Directed approaches, namely that 4X can create additional data contribution opportunities without increasing disruption. However, due to potential confounds with respect to scheduling constraints, weather, and unexpressed preferences, only a few users went out of their way in response to notifications sent at-distance and contributed additional information. We designed Study 2 to address these confounds, and additionally measure the extent to which collected data can be used to promote additional data contributions from interested users who go out of their way. Specifically, Study 2 provides evidence for the following hypotheses:

H3: 4X collects additional data from users who respond to at-distance notifications and go out of their way to visit places of interest, while still being minimally disruptive.

H4: Users are more likely to go out of their way when presented with information that is more valuable to them.

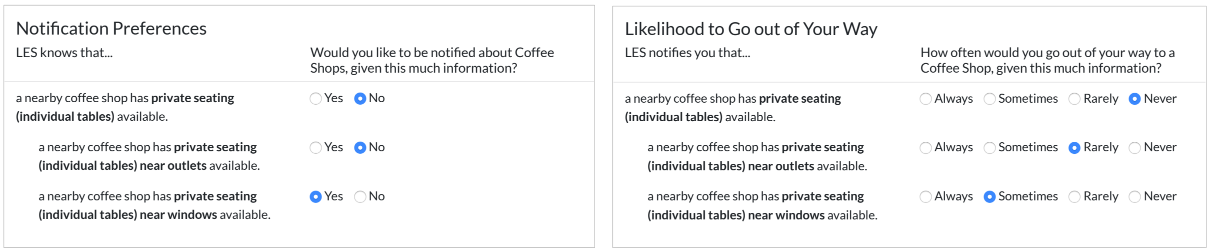
3.6.1. Method and Analysis

3.6.1.1. Participants. We recruited 18 undergraduate and graduate students of a mid-sized U.S. university through mailing lists, flyers, social media, and word of mouth. Participant ages ranged from 19 to 31 (M = 23.93, SD = 3.69), with 8 female and 9 male participants (one preferred not to specify). The study took place over 14 days, during which participants completed a pre-study survey, used LES as a part of their daily lives, and completed a post-study survey. We compensated participants $30 for their time spent on surveys and installing LES, but did not incentivize their behavior during the study (i.e., no monetary incentive was provided for task completion).

3.6.1.2. Procedure. All 18 users were assigned to a single condition in which they use the 4X version of LES (identical to Study 1) for 14 days.[6] We ran Study 2 during the university's summer session, when scheduling constraints are less restrictive and the weather is favorable compared to Study 1. In the absence of scheduling and weather confounds, we expect to see more users going out of their way to locations of interest and contributing additional data at target locations and en route.

To better capture users' notification preferences, we designed an improved notification preference survey that provides users with finer-grained control over what notifications they would like to receive (or not); see Figure 3.5a. Unlike the survey from Study 1, which only allows users to specify individual properties about locations that may be of interest to them

[6] All 18 users used LES; 16 completed the post-survey.



(a) Notification Preferences **(b)** Interest Preferences

Figure 3.5. Notification and interest preferences for Study 2. In this example, the participant wishes to be notified about private seating at coffee shops only when they are near windows (a), and reports a stronger interest in seating near windows than near outlets (b).

(e.g., private tables, windows), this improved survey allows users to specify, for any state of information LES has in its data scaffold, whether they would like to be notified or not. For example, Figure 3.5a demonstrates how this allows a user to specify that they only want to know about private tables when LES knows that these tables are near windows, which wouldn't have been possible in Study 1. Additionally, users can now select specific locations they would like to be notified about (e.g., notify me about Starbucks, but not Peet's Coffee). With an improved understanding of users' notification preferences, we expect to send fewer notifications at-distance with information that users do not find useful.

To study the extent to which people's interest in the data may influence their decisions to go out of their way and make additional data contributions, we designed an information preference survey that asks users to state how likely they think they are to visit target locations (i.e., Always, Sometimes, Rarely, or Never) when information of interest is presented to them, assuming no scheduling conflicts; see Figure 3.5b. By collecting both notification preferences and information preferences, we are able to measure the likelihood of users actually going out of their way when notified with information based on their notification preferences, given their reported degree of interest in acting on the data based on their information preferences.

Similar to Study 1, we used user responses to at-distance notifications as an ecological momentary assessment (EMA) [**119**] to assess whether they found the information presented to them useful and whether they were going out of their way based on the information presented or if they happen to be already going. Beyond EMAs, users completed a post-study survey following the 14-day usage period, which asked them to: (1) reflect on what they found valuable or not valuable about the information presented to them; and (2) why they decided to go out of their way and contribute additional data (or not). We followed the same procedure as Study 1 to code and analyze these qualitative responses.

3.6.1.3. Measures and Analysis. Similar to Study 1, we measured the number of times users went out of their way and made additional data contributions by examining logged location data to see if users actually went out of their way and made contributions. We then analyzed post-study survey responses to gain a deeper understanding of why they decided to go out of their way and make additional contributions. To evaluate whether people were more likely to go out of their way when presented with information of interest, we used users' information preferences to compute, for each level of interest (i.e., Always, Sometimes, Rarely, or Never), the proportion of times users went to the location when notified.

Similar to Study 1, we collected user reports of perceived disruption. To evaluate the extent to which 4X avoids over-notifying users by only sending notifications at-distance with information of interest, we also compared the number of notifications that 4X sent to the number of notifications that would have been sent had we notified all users within 300 meters of a target location–as the Directed condition did in Study 1–by simulating the Directed condition notification policy with location data collected during the study.

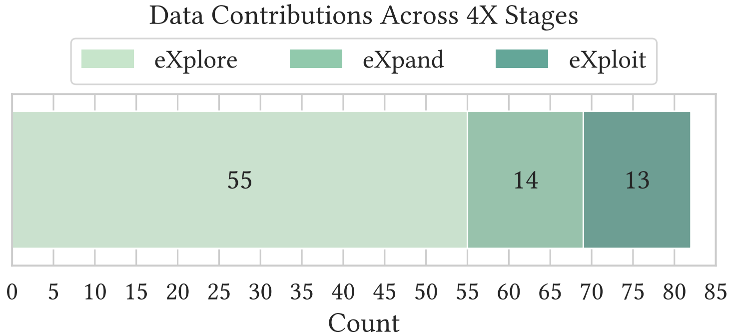


Figure 3.6. Total number of contributions in Study 2 by 4X stage. 4X yields 49.09% more data from directing users to locations of interest via eXpand and eXploit over data collected opportunistically via eXplore.

3.6.2. Results

3.6.2.1. H3: 4X Gathers Additional Contributions from At-Distance Notifications with Minimal Disruption. 4X yielded 49.09% more data by directing users to target locations of interest, resulting in data contributions made at that target locations and en route to them. Figure 3.6 shows that users in Study 2 made a total of 82 data contributions, 55 of which are from at-location contributions from eXplore and 27 of which (32.91%) are additional contributions that resulted from users responding to at-distance notifications and making additional contributions at target locations via eXpand (14 times) and en route via eXploit (13 times). These additional contributions helped to increase data fidelity at target locations by adding 25.46% more data above the 55 eXplore contributions, and additionally expanded data coverage at locations en route.

Similar to Study 1, users found 4X to be minimally disruptive; the average reported disruption is 1.88–a number between Never Disruptive and Rarely Disruptive–on a 5-point Likert scale (SD = 0.72). Analyzing the number of notifications sent at-distance, we found that 4X only sent 140 notifications (0.56 per user per day) versus the 1051 at-distance notifications (4.17 per user per day) that would have been sent by a Directed approach. In other words, by sending targeted notifications that are well-matched to people's notification preferences, 4X avoided unnecessarily notifying users at-distance who would not be interested in deviating, while still yielding additional data contributions from interested users beyond what can be collected via an Opportunistic approach.

Users generally contributed additional data whenever they went to the target location in response to the data they received (14 out of 17 times for eXpand; 13 out of 14 times for eXploit). Users noted that making additional contributions at the target locations they went to with LES was easy to do and required little effort: *"I picked up cold snacks, and then it asked me what type of cold snack. It was pretty natural to ask a question about something I specifically came for. It was very low-effort to respond, since I already had the knowledge."* (P16). Similarly, users also felt that it was easy to make contributions en route when LES asked about bike racks and parking lots: *"I went to [the gym] and then was asked about the bike racks. I responded because it was easy for me to respond as it was on my way, and I already had my phone in my hand."* (P2). These responses show the benefit of using low-effort interaction techniques in 4X systems to collect additional contributions en route to and at target locations that users are drawn to.

Users further highlighted reciprocity as a reason for contributing at eXpand and eXploit locations. For instance, P10 noted: *"I feel like when I get something from others, I would like to give back."* Beyond providing data back to other data collectors, some users wanted to better know how their contributions were being used by others: *"I wish I knew how my contributions affected others – there were many times I passed [a coffee shop] and was prompted, and I wish I knew how useful my reports were for others, I think it would have motivated me more to maybe step in and check instead of just peer through the windows and guess if I knew someone was*

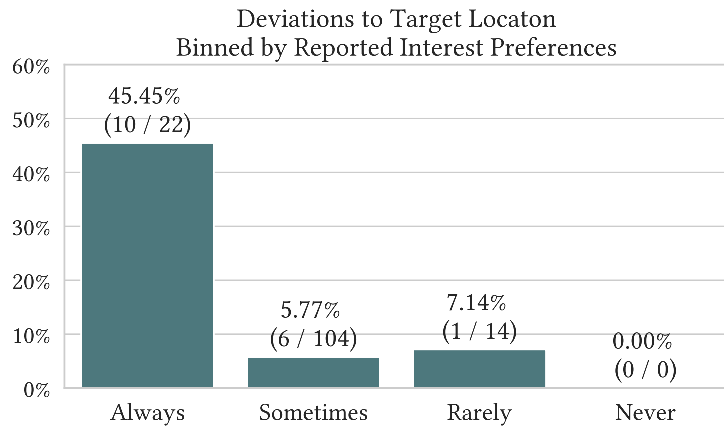


Figure 3.7. Users were more likely to go out of their way to locations when presented with information that is more valuable to them.

relying on the information." (P8). Along with providing users with data of interest to motivate additional contributions, 4X systems may also consider informing users about how other data collectors value their contributions as another way to motivate continued use of the system.

3.6.2.2. H4: Users are More Drawn to Locations of Interest When the Information is Valuable to Them. As we expected, users were more likely to go out of their way when they found the information presented to them to be valuable. In 16 of the 17 cases in which users went out of their way in response to an eXpand notification, users reported that they would "Always" or "Sometimes" go out of their way. Moreover, Figure 3.7 shows that when users are presented with the information they find most valuable (i.e., Always), they go out of their way 45.45% of the times versus less than 10% of the time when information is less valuable (i.e., Rarely and Never).

While this result shows that more valuable data has more drawing power, it is worth noting that there are many more cases in which users receive information that is moderately valuable (104 vs. 22), through which six people still went out of their way. From a design perspective, this suggests that while we should generally aim to collect and present information that users value most, presenting users with moderately valuable data may still be useful and can also lead to additional contributions.

Users reported going to nearby locations of interest that LES notified them about because they generally found the information to be useful. Compared to Study 1, in which 18.42% of eXpand notifications were found to not be useful, only 2.86% of eXpand notifications in Study 2 were found to not be useful. This result suggests that our revised notification preference survey may have better captured users' preferences and thus provided LES with a better user model to use when notifying users about nearby locations. From post-study responses, 9 out of 14 users who went out of their way specifically highlighted that they did so because they found the information they received to be valuable. For example, P16 wrote: *"I was interested in ice cream, as it had been pretty hot and muggy one of the afternoons. I received the notification while at my desk in [building] to visit [another building], and this was interesting enough to stop my work for 15 minutes."* In other words, providing users with data that they generally find useful can be a viable way to draw them to nearby locations of interest outside of their regular routine.

User responses to post-survey questions highlight notification timing as an important factor that influenced their decision to go out of their way and their perceived value of eXpand notifications from LES. P10 noted: *"It depends on timing. In the morning the cafe notification is valuable. In the afternoon after 5 or 6 pm gym availability is more important."* For free food, timing was most often dependent on when the user had last eaten and if they were occupied with work: *"One time I got a notification that there was food. Critically, this notification arrived in the afternoon when I am typically hungry because I do not ever pack enough food for lunch. I was happy to jump up from my desk and find the food because I am also usually less productive in the afternoon."* (P5). These reports suggest that eXpand in 4X may be even more effective in gathering additional contributions and reducing disruption by accounting for such contextual factors when deciding to notify users.

When users did not go out of their way, they primarily cited scheduling conflicts and existing plans as reasons for not going. In the post-survey, 14 out of 16 users noted scheduled meetings to be a reason why they could not go to the location they were notified about. Eight users noted that they were sometimes in a rush (e.g., to catch a train) and thus did not have time to respond. Three users mentioned that they did not want to respond to a notification because they were not interested in the data being collected, suggesting that users' personal interest in the data may influence their willingness to contribute.

Even if users did not go out of their way, some still found value in being provided information about locations of interest through LES as it helped them better plan when they might go to certain locations that they cared about. For example, P14 mentioned how they found the notifications for the gym useful even though they could not go because of work commitments: *"With [the gym], I was glad when I would see open spots, but I was usually at work and could not workout at that time. But at least I would know that the gym is generally empty for when I wanted to go to the gym."* (P14). This suggests that even when users cannot immediately act upon the data they receive, 4X systems can still provide value to the user as the information could help in planning future trips to locations.

3.6.2.3. Summary of Study 2 Results. In summary, results of study 2 demonstrate how 4X can use multiple data collection strategies based on when and which data collection opportunities become available to better meet the desired data collection goals without disrupting data contributors by staying within their interest. Our results also highlight the importance of modeling people's degree of interest in information, incorporating people's willingness in deviation given collected data when enacting different data collection strategies.

3.7. Discussion and Future Work

We propose *4X*, a framework for multi-stage data collection processes that determine effective data collection opportunities by reasoning about changing states of the world, people's locations, and their willingness in deviating from their routine based on our knowledge of the world. We demonstrated the effectiveness of 4X that uses multiple data collection strategies that are tailored towards people's dynamically changing interests and situations on the ground to better achieve desired data collection goals while still presenting opportunities that are likely to be in people's interests and that may fit within their routine and goals. In the rest of the section, we revisit the core ideas behind 4X, discuss how they may generally inform the design of participatory sensing systems, and present directions for future work.

3.7.1. Keeping Users' Data Interest at the Foreground

The most important aspect of any 4X system is keeping users' interests and goals in mind since 4X uses collected data as an incentive for enabling directed sensing to gain additional contributions. Our study results showed that this approach created additional data contribution opportunities (Study 1) and led to increased data contributions both at locations of interest via eXpand and en route via eXploit (Study 2). Unlike the directed approach which makes requests indiscriminately, having user models allowed 4X to selectively notify users only when data of interest is available, which reduced over-notification and limits disruption. In other words, 4X was able to increase data fidelity and coverage in ways that aligned with, and advanced, the interests and goals of users.

Having accurate user models in 4X is essential for selectively notifying users to provide information of interest while avoiding unnecessary disruption during the eXpand stage. In LES, we first conducted needfinding to determine the kinds of data that users are generally interested in about locations and events around campus. We then used a direct elicitation strategy (i.e., a questionnaire) to learn what specific information and events each user was interested in knowing about. While this approach was generally effective for increasing data collection opportunities without increasing disruption, results from our study suggest that including contextual factors such as users' schedules or how their data interests vary throughout the day may better capture when users are likely to go to a nearby location given some information of interest (as opposed to being interested, but not going). Having a more accurate user model that incorporates such contextual factors would allow a 4X system to reduce disruption from notifications that a user is unlikely to act upon, and may even increase contributions in cases where the model identifies situations where a user is willing to deviate further from their routine in response to information about locations and events that are particularly valuable to them. To build such richer user models, future 4X systems may complement direct elicitation by training machine learning models that predict a user's likelihood of going out of their way to a target location based on (1) contextual factors such as the user's schedule and their current activities [**39, 63, 7, 32**]; (2) their current and future routes [**82, 91, 105**] and associated cost of diversion to a target location [**59, 60**]; and (3) the value of the certain, known information about a target location to a user at different times during the day (e.g., line length at a coffee shop in the morning versus late in the afternoon).

While better user models can potentially reduce disruption and increase contributions for some sensing initiatives, the larger issue is that the collected data may not be motivating enough for users to deviate from their routine. For instance, volunteers helping to monitor city infrastructural issues such as potholes and broken streetlights would have no interest in visiting such locations. One approach to collecting such data using 4X is to leverage the eXploit stage to collect mundane data as a byproduct of a more interesting sensing initiative. As shown in Study 2, LES was able to gather information about bike rack and parking space availability that was unrelated to users' direct interest. This can allow a 4X system that, for example, primarily collects and shares information about events of interest in a city to also collect information about city infrastructural issues that users might encounter en route.

While we do not expect 4X to be useful in all cases, in some cases there may be clever ways to *transform* the collected data so that it becomes incentivizing and useful to data contributors. For example, recent work on Habitsourcing [**93**] creates immersive interactions and narratives that reference objects in the physical environment to support a user's habit-building practice (e.g., going for a run) while collecting sensing data as a byproduct of their habit building practices. Future Habitsourcing systems may use collected information about locations and events (however mundane) to shape immersive story narratives [**110**] and interactions that incentivize users to deviate from their routes as part of their habit-building practice, and while doing so, contribute additional data where it is desired. In future work, we are broadly interested in exploring methods such as this for transforming collected data in ways that support interactions and experiences that are valuable to the user, even if the collected data in its raw form is not.

3.7.2. Using Dynamic Data Collection Processes

4X is a dynamic process for collecting data that uses different data collection mechanisms depending on what data is available to the system and what available information users are interested in. Because 4X leverages the benefits of both opportunistic and directed approaches, it enables the collection of high-fidelity and high-coverage data while being minimally disruptive to users and providing them with data of interest. We show through Study 2 that users made 25% more contributions from eXpand to increase data fidelity and 24% more contributions from eXploit to increase data coverage. By not fixating on a single approach (i.e., opportunistic or directed), 4X is able to *flexibly work with the data and interest available* to achieve a big gain in data contributions while still meeting user needs that would not have been possible with either approach alone.

Dynamic data collection processes such as 4X can be used to scaffold data and motivation concurrently so that larger and larger groups of users can be drawn out of their way to contribute to participatory sensing efforts. We describe below a general strategy for achieving this–which 4X supports–that we call *incentive chaining*. Initially, a small subset of all users, who require the least amount of information to be motivated, are the ones most likely to be drawn out of their routines and make contributions at the target location they were drawn to. Their contributions would then increase data fidelity and motivate other users, who may have needed slightly higher incentives, to now be drawn to the target location and so on. For example, users may pass by an area and opportunistically report an event such as a pop-up concert. This information may then draw in relatively nearby users who are generally interested in any kind of music, where they specify that the concert is for rock music. Then, users who are interested in rock music may be drawn to the location from further away and make additional contributions that may draw in fans of the specific bands playing at the concert from even further away. Thus, incentive chaining and similar strategies that are enabled by dynamic data collection processes such as 4X can expand the set of users who could initially contribute to encompass a much broader set of users as data and motivations scaffold, but only doing so when users can be effectively motivated (e.g., there is now enough data for the user to be drawn in). In large-scale deployments, we expect that a small amount of additional contributions made by a few initial users drawn to locations of interest could lead to a snowball effect of significantly more opportunities created. Studying these incentive chaining effects at scale are a good avenue for future empirical evaluations of 4X-like dynamic data collection processes.

To allow for incentive chaining-like effects to occur, all sensing initiatives need ways to collect the initial pieces of data so that the data collection protocol can dynamically transition from stage to stage. However, some sensing initiatives may find it challenging to obtain these initial opportunistic contributions if, for example, users' normal routines do not coincide with tracked locations. While we used LES to collect all pieces of data for our sensing domain, other initiatives may consider using gamification mechanics [**129**] or even machine sensors [**115, 87**] to complement apps such as LES so that particular pieces of data contributed across applications can together support a broader sensing initiative. In other words, the stages of 4X can be enacted across sensing applications, whereby some applications (and their associated mechanisms) make contributions opportunistically, which then empowers other applications to use eXpand and eXploit to further build data fidelity and data coverage. In this way, dynamic data collection processes can easily be extended to integrate multiple forms of participatory sensing while still reaping the benefits of transitioning between data collection stages when users' interests and the collected data align.

3.7.3. Enabling Community-Based Data Collection

4X systems can be thought of as community-based data collection processes where data contributions from certain community members are of value to other community members, who then become motivated to make further contributions that benefit others in the community and so on. In Study 2, we found that collected data from some users successfully influenced others to deviate from their routine and make additional contributions. In this way, 4X creates cross-community interactions between data collectors since some contributions are being used to inspire contributions from others in the data collection community.

Seeing 4X as a community-wide data collection approach allows us to consider new opportunities for how we may design the data collection process at the community-level rather than only at the individual-level. For instance, in LES we rearranged information scaffolds based on the overall collected user preferences so that data of greatest interest to the community at each location would be collected first. Similarly, we can consider rearranging the information scaffolds based on the needs of smaller sub-communities at different times during the day. For example, those interested in coffee shops may want to know about how long the line at certain shops are during their morning commute, but be more interested in seating availability later in the day. In other words, by explicitly recognizing the needs of different sub-communities within the sensing initiative and tailoring the data collection to their changing needs, 4X is able to find and create more situations where users would be willing to deviate from their routines since the collected data would better align with the sub-community's data needs and interests at different times during the day.

Building 4X systems that continually consider the needs of different sub-communities during data collection requires crafting and tuning more sophisticated notification selection criteria and question selection criteria that can dynamically adjust to the needs of the community based on the information that is available and the people available who may be interested and willing to go out of their way. Manually tuning these selection criteria can become challenging and ineffective when community members have diverse data needs, and when decisions about who to ping now affect opportunities to meet others' needs later. Instead, we are interested in future work that create *Adaptive 4X* systems that can automatically reason and adapt their notification policies based on the value of soliciting certain contributions for the community and of the value of the collected data to people in the community. To best collect data valued by the community without being overly disruptive to any individual, such systems may use decision theory to make decisions about whether to request a data contribution now, or to wait for better opportunities when users may be closer to other locations of interest to them where additional data contributions would be particularly valuable [**76**]. To best leverage data collectors' efforts and support users' needs across the community, such systems may adaptively notify more or less users at different times so as to best use the community's data collection efforts to meet the changing data needs of the community (e.g., notify more people who may contribute about coffee shops in the morning). We expect these systems to make decisions that better align with the varying interests of users, which allows for user needs to be better supported while also advancing data collection goals.

CHAPTER 4

Opportunistic Supply Management Framework

While Hit-or-Wait and 4X evaluate the quality of opportunities with a single individual's routine, good opportunities can be relative to other opportunities that may arise across multiple people's routines. In this chapter, we will discuss the core challenge in designing community-level mechanisms that evaluate the quality of opportunities across multiple people's routines.

4.1. Introduction

The growth of mobile devices in recent years has helped to bring about physical crowdsourcing systems [**9, 126, 124**] that help connect people to solve local, communal problems. These systems need *community-level mechanisms* that can effectively manage the recruitment of volunteers in a way that meets the needs of volunteers and system goals to ensure long-term viability. While commercial physical crowdsourcing systems (e.g. Uber, Lyft, TaskRabbit, and Instacart) can use market mechanisms and financial incentives to engage workers whenever help is needed, existing approaches in volunteer-based systems are limited in how they strike the balance between the needs of volunteers and system goals. Most volunteer-based systems provide flexibility to volunteers by letting volunteers decide when and which tasks they contribute to [**53, 46, 9**]. This flexibility allows volunteers to meet their own needs, for example contributing to tasks that suit their schedules and routines, that are convenient for them, or that are of their interests. But the flexibility provided to volunteers and the opportunistic nature of their contributions collected can make it hard to meet the desired system goals [**43, 140**]. For example,

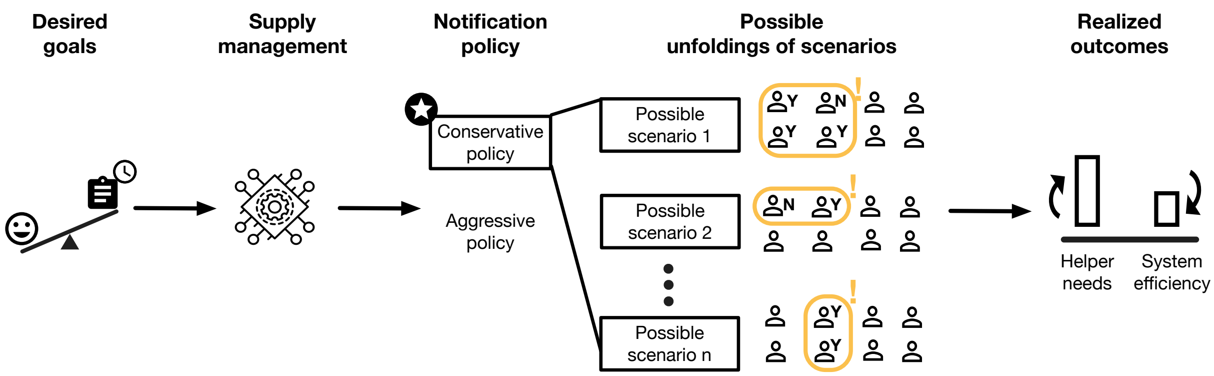


Figure 4.1. Each volunteer-based system has to manage the balance between system efficiency and the needs of volunteers, and each system may have their own desired balance. The opportunistic supply management framework allows the system to optimize the desired balance between system efficiency and the needs of volunteers by reasoning about different community-scale task notification policies that opportunistically decide who to send tasks across the community. As volunteers may or may not become available or accept tasks, supply management simulates over the possible unfoldings of adopting a task notification policy and chooses an optimal policy that can best achieve the desired goals.

a community-based peer-to-peer delivery system may passively wait for volunteers to show up and complete tasks on their own terms but the system may fail to complete tasks in a timely manner. In contrast, systems may direct volunteers to specific tasks to meet the system goals (e.g. by asking to fulfill an urgent task that is out of a volunteer's existing route), but this approach fails to provide much needed flexibility to volunteers because the systems coordinate contributions with the assumption that volunteers are committed to participate [**98, 114**]. This approach is less likely to attract volunteers than those systems that consider the needs of volunteers [**27**]; as a result, volunteers may not even complete the assigned or directed tasks.

As an alternative approach to these existing approaches, we consider in this dissertation a flexible coordination approach by which individuals do not need to commit to do specific tasks but the contributions can still be coordinated to collectively achieve desired system goals. To enable a flexible coordination approach that can effectively manage the recruitment of volunteers in a way that meets the needs of volunteers and system goals, we must address the core challenge in setting task notification policies that decide when, where, and to whom to notify of tasks. Existing technical solutions largely assume people's availability and participation when coordinating contributions across the community to meet desired system goals. However, by maintaining flexibility, we cannot be sure who will be available to contribute and whether they will actually contribute, thereby making it hard to know if a policy would be effective without knowing how it might, across a community of volunteers, lead to good outcomes that are aligned with the goals of the community. For example, an effective system must manage the tradeoffs imposed by being too aggressive in recruitment—which can be overly disruptive and result in a low task pickup rate—and being too restrictive in recruitment—which can involve too few volunteers, overburden the ones that are involved, and leave a disproportionately large number of task demands unfulfilled. Prior work in on-the-go crowdsourcing, in which volunteers are recruited to tasks they can contribute to conveniently in their existing routines, demonstrated that managing this tradeoff is not at all obvious [**77**]. Results showed that small changes in notification radius can lead to order of magnitude increases in the number of people reached, but at the same time affected other factors such as interest in ways that lead to order of magnitude decreases in the likelihood of task pickup. These early results suggest a general need for methods capable of reasoning about such trade-offs to generate policies which notify just enough people to meet demand with a well motivated and minimally disrupted crowd.

To overcome this challenge, we propose *opportunistic supply management framework*, a general architecture for a model-based, principled way of optimizing task notification policies for engaging contributors to achieve desired global outcomes. Supply management follows community members' routines and integrates models that describe how task notification policies affect the available supply of volunteers and their likelihood to accept tasks, and how that in turn affects system efficiency and the needs of volunteers. Using these models, Supply Management simulates the possible outcomes that may result from adopting a task notification policy and chooses an optimal policy for a given situation (or set of situations) that best achieves intended system goals and desired volunteer experiences in expectation. With this approach, Supply Management can reason about opportunities that may unfold across people's routines, take into consideration people's availability and willingness to help, and devise custom-tailored strategies that adapt to changing situations without ever imposing on what each individual must do. Unlike existing task recommendation mechanisms that only consider how each individual can best contribute to the system, supply management considers how to leverage volunteer efforts across the community to best meet system goals in ways that still ensure good volunteer experiences by not overburdening or disrupting potential volunteers.

We evaluate the effectiveness of opportunistic supply management in setting task notification policies and analyze the decisions it makes through a simulation study and a 4-week field deployment study ($N = 26$) in a peer-to-peer delivery setting. Our results demonstrate how opportunistic supply management can (1) help system designers arrive at policies that identify "goldilock zones" that effectively balance system and volunteer needs; (2) prioritize and promote specific goals, for example to avoid overdisrupting volunteers while attempting to complete tasks; and (3) make dynamic adjustments when a policy becomes ineffective, should tasks be completed more or less quickly than was predicted.

The rest of the chapter is organized as follows. We first review related work to motivate the need for community-scale mechanisms for managing the supply of opportunistic contributions. We then introduce the opportunistic supply management framework, and discuss how it can be used to model, simulate, and optimize the choice of task notification policies to achieve desired tradeoffs. We present the methods and results of two studies, and conclude with a discussion of future directions in developing enabling technologies for community-wide coordination that consider simultaneously the needs of community members and the goals of the system.

4.2. Related Work

As social or crowd computing systems seek to meet the needs of volunteers together with high-quality services, there's a need for community-level mechanisms that can balance system efficiency and volunteer needs under changing situations and conditions (e.g. volunteer availability or demands). The traditional solution for commercial services is markets; by using prices to align incentives, the systems can then recruit workers to meet system needs in ways that align with their self-interest. For example, on-demand services like Uber or Lyft use dynamic pricing mechanisms [**20, 19, 79**] to accommodate demand changes and manage supplies in a way that meets system goals. Some volunteer-based systems have also adapted market mechanisms by using *scrips* [**40, 72, 73**] and *virtual currencies* to motivate participation and ensure system efficiency. While useful in some cases, these mechanisms do not explicitly reason about the needs and wants of contributors that extend beyond monetary rewards when attempting to achieve desired system goals. In contrast, our work contributes a community-based mechanism that can reason explicitly about the experience of community members when attempting to achieve system goals as situations and conditions change. We envision that this approach will support the creation of new configurations of volunteer- and community-based systems and provide new tools for effectively managing such communities.

A core feature of our supply management framework is that it provides a community-level flexible coordination mechanism for *opportunistic coordination among volunteers who may or may not become available, or always provide help when asked.* This is in contrast to prior work where volunteer availability and participation are known a priori (e.g. volunteers are committed to participate or assumed they will be), and thus systems can use approaches that directly coordinate contributions to achieve desired outcomes. For example, some systems coordinate what each individual must do by either planning ahead of time [**22**] or providing just-in-time, step-by-step instructions [**83, 45**]. Specific to physical crowdsourcing, prior work focus on optimizing effective task assignments to minimize deviations from volunteer routines while maximizing system efficiency [**98, 23, 69, 114**], which largely assume that individuals will accept tasks when asked. But, these prior techniques break down in real-world settings like on-the-go crowdsourcing where any given volunteer may or may not go near task locations, and may or may not accept the tasks. This may result in assigned tasks being unfulfilled or disrupting people who are not willing or able to help. Instead of assuming the availability of any particular volunteer, opportunistic supply management models policies that describe the *conditions* under which to recruit volunteers, and reasons about the effectiveness of such policies in coordinating contributions opportunistically [**59, 60**] given the many scenarios that might occur during execution time (e.g., who is actually available, whether tasks are accepted or not).

There is a large body of literature within CSCW that studies how task suggestions and recommendations can promote effective contributions from volunteers (e.g., [**97, 125, 77, 32**]), but much of this literature focuses on how each individual can best contribute to the system and not on how coordinating contributions across a community can best meet system goals in ways that ensure the needs of volunteers. Specific to on-the-go crowdsourcing, recent work on Hit-or-Wait [**76**] introduced a decision-theoretic approach for deciding when to route a (best) task for a user to contribute to along their route in an online manner. While effective for determining a task for an individual to contribute to, this approach fails to consider how to best leverage volunteer efforts across the community. As a point of contrast, supply management can decide which subset of community members to recruit in a way that optimizes the balance between desired system efficiency and volunteer needs (e.g. avoid overburdening volunteers with high-effort tasks or with significant detours in existing physical crowdsourcing systems [**45, 117, 121**]). Rather than only consider overburden and disruption on an individual basis (e.g. in social Q&A [**135**]), community-level mechanisms such as supply management can also consider overburden and disruption across community members, which in turn allows systems to have a larger decision space to decide how to engage volunteers to best achieve the desired system and volunteer needs.

To enable community-level flexible coordination mechanisms in volunteer-based settings where there's uncertainty in availability and participation and where situations are changing, our technical approach takes inspiration from Horvitz's work on flexible computation [**55, 61**], which suggests principles and ideas on ways to reason about strategies that can achieve optimal outcomes under uncertainty. Flexible computation provides ways to select an optimal strategy or sequence of strategies for using scarce computational resources that generate useful bounded-resource solutions that are tailored towards varying situations. As flexible computation treats computation power as a scarce resource, by analogy, our approach treats volunteer attention and contributions as scarce resources, which differ from computation power in that individuals may have varying availability and willingness depending on their situations, may or may not decide to help when they are asked to help, may want to be minimally disrupted, and have dual needs in quality of services as a requester and user experience as a volunteer. To preserve opportunistic nature in participation and meet the desired goals, our technical approach also reasons about uncertainty in availability and participation and desired goals, and surfaces task needs to volunteers across the community at opportune moments when the volunteers can conveniently contribute to meet the goals. Instead of assuming future situations are fixed, our technical approach generates strategies that are custom-tailored to varying situations where potentially available resources are uncertain, and that can still provide optimal solutions to reach the desired outcomes with available resources.

In modeling volunteers and simulating policies to analyze their effectiveness, our technical approach bears resemblance to the use of decision-theoretic methods in online crowdsourcing to optimally control and design the allocation of tasks to workers in workflows [**28, 92, 138**]. While similar in some respects, a core difference is that prior approaches largely assume that workers can be recruited to contribute to tasks as needed. In our setting, people are notified of tasks as opportunities arise, and even then, people may or may not accept tasks presented to them. This difference requires us to consider and incorporate richer models of the changing conditions of the participants and their effects on system outcomes. Moreover, these models are used not only to optimize system efficiency, but also for factoring in volunteer experience into our optimization, which has largely been ignored in prior works. In this way, we leverage and extend existing modeling and optimization techniques from AI, not for promoting efficiency as usual but also for addressing core issues of import to social computing and CSCW [**80**], where the priorities are around building and maintaining communities [**84**] and supporting long term community development, particularly through peer production [**10**] and via volunteer-based systems.

4.3. Opportunistic Supply Management: A Framework for Community-Level Flexible Coordination Mechanisms

We begin this section by presenting *community-based, peer-to-peer delivery* as an illustrative example through which we highlight general challenges in setting task notification policies to effectively meet system goals and volunteer needs. We then introduce the opportunistic supply management, a framework for community-level flexible coordination mechanisms that optimizes community-wide task notification policies to achieve desired volunteer needs and system efficiencies. Finally, we demonstrate how we might use this framework to manage the supply of volunteers in a community-based peer-to-peer delivery application.

4.3.1. Community-Based Peer-to-Peer Delivery

As an illustrative example, we consider a community-based peer-to-peer package delivery service that seeks to leverage volunteers' existing routes to effectively deliver packages from pick up locations to drop-off locations. In this domain, people request deliveries and a system notifies people who pass by the item pick-up location and might be able to help. To engage potential volunteers, the system sets a *task notification policies* that determines the conditions under which to notify volunteers to tasks to balance system goals (e.g., the rate at which tasks are completed) and volunteer needs (e.g. avoiding over-disruption and over-burden). For instance, a task notification policy may set conditions to notify only nearby volunteers within a certain radius, or only people who have not already helped earlier today. Enacting a policy, the system makes potential volunteers aware of tasks that need to be completed when they are in conditions matching the policy. Potential volunteers can decide whether or not to help, and tasks get completed opportunistically as volunteers become aware of tasks and decide to help. While only some of these volunteers may accept tasks, others notified may also experience disruption in being asked to help when they are unwilling or unable.

As a system designer, we are interested in devising task notification policies that govern who to recruit across the entire community to optimize the desired balance between system efficiency and volunteer needs. While there may be certain tradeoffs that must be made between the two, we are interested in identifying policies that find goldilocks zones, where we might be able to, for example, complete tasks reasonably quickly with minimal disruption. Depending on the wishes of the community, we might also want to identify policies that can effectively prioritize system efficiency or volunteer needs more than the other as desired, while still considering both.

4.3.2. Core Obstacles in Setting Notification Policies to Balance System Efficiency and volunteer Needs

We use the peer-to-peer delivery example to illustrate three general challenges in setting task notification policies to balance system efficiency and volunteer needs. First, while at a high-level we would generally choose more aggressive policies when we value task completion and more conservative policies when we value low disruption, the best policy is dependent not only on our goals but also on the situation on the ground. For instance, even when task completion is prioritized over disruption, setting a conservative policy may be more effective still if it leads to completing tasks quickly enough but without disrupting many people. In other words, an effective approach for choosing a policy must consider *both* our goals and the situation on the ground; there is not a single best policy for all situations, nor across all outcomes we might hold.

Second, the best policy for a given situation may not correspond to our general intuitions, which can make manually choosing policies ineffective. For instance, one may think when the demand is low, it is always better to use a more conservative policy because it doesn't disrupt many people and can still complete some tasks. But in actuality, a more aggressive policy can outperform a more conservative one in some scenarios. For example, when the task pickup rate is high even when we reach out to a larger pool of potential volunteers (e.g., ping people who are farther away to help pickup and deliver packages), choosing an aggressive policy can often complete tasks quicker by tapping into a larger supply of volunteers, but without having to disrupt many people because once all demand is met, it stops notifying people anyway and thus keeps disruption low overall across the community. In other words, choosing an effective policy thus requires not only considering our goals and situations, but also the specific conditions around people's availability and willingness to help and how they may affect outcomes in non-intuitive ways.

Third, uncertainty in people's availability and participation may lead to significantly different outcomes that are valued differently. On any given day, the same policy may lead to significantly more notifications (e.g., if more people happen to go near the package center) and significantly fewer task pickups (e.g., if many people happen to be preoccupied that day). Choosing a policy by only considering the average-case scenario may be ineffective when certain outcomes, such as significant over-disruption, can incur a disproportionate cost on a community. In other words, simply considering the average case scenario is insufficient for making policy determinations because the inherent uncertainty in the domain implies that many possible scenarios are likely to unfold that may be valued very differently based on the community's goals. Choosing an effective policy requires taking such uncertainty into account, so that we can take into consideration the set of possible outcomes that may unfold instead of fixating on a single, likely outcome.

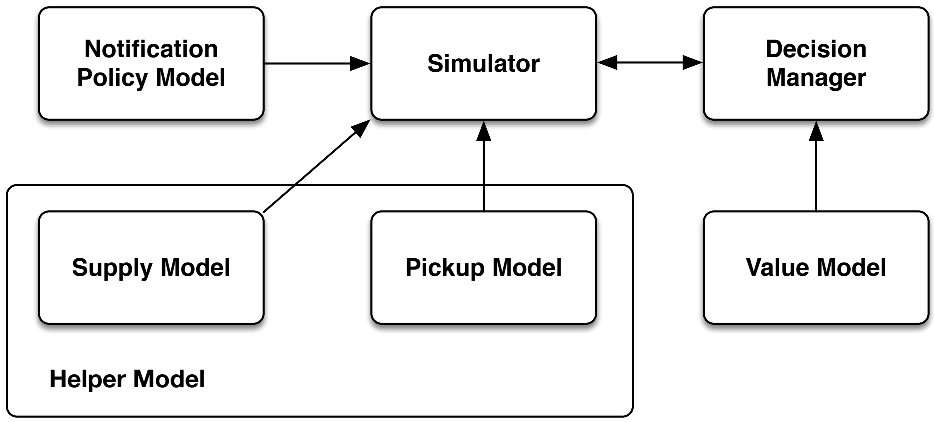


Figure 4.2. An architectural diagram of the opportunistic supply management framework for modeling and optimizing the choice of task notification policies. The framework consists of (1) *models* of notification policies, volunteers (i.e. supply and pickup), and values; (2) the *simulator*, which uses the models to simulate and evaluate the effectiveness of policies; and (3) the *decision manager*, that determines the optimal policy.

4.3.3. Opportunistic Supply Management Framework

To overcome the three core challenges in setting task notification policies, we introduce the *opportunistic supply management* framework, which provides a general decision-theoretic architecture for modeling and optimizing the choice of task notification policies; see Figure 4.2. Supply management simulates over and compares possible outcomes that can be reached by enacting a task notification policy, and chooses an optimal policy for a given situation (or set of situations) that best achieves intended goals for the system and for the community. Unlike existing technical approaches that assume that people will adhere to the chosen strategy and thus reach the desired goals, supply management models uncertainty in people's availability and willingness to help under different strategies. In other words, by incorporating uncertainties into the reasoning, supply management does not assume that people will follow through with the strategy, but instead works with the uncertainties to devise a policy that best addresses the desired goals. Supply management's model-based, simulation-based approach allows the system to reason about how the world will unfold, accounting for uncertainties, and devise custom-tailored strategies for varying situations that can best meet the desired goals. This allows the system to address the three core challenges in choosing policies in ways that (1) adapt to changing situations; (2) takes into consideration people's availability and willingness to help; and (3) accurately reasons about eventual outcomes across possible unfoldings.

Our supply management framework consists of three core components: (1) *models* of notification policies, volunteers, and goals; (2) the *simulator*, which uses the models to simulate and evaluate the effectiveness of policies; and (3) the *decision manager*, that determines the optimal policy. We introduce each in turn.

The supply management framework contains the following *models*:

- The *notification policy model* determines when, where, who, and how to notify a user about a task. A notification policy model may consider (a) at what distance to notify a user; (b) how often to notify a user; and (c) other contextual factors, such as a user's schedule [**63, 7, 32**], their current and future routes [**82, 91, 105**], and associated cost of diversion to the pickup location [**59, 60**], weather, and others that may affect users' likelihood of accepting a task. A notification policy can be represented as simple as a radius-based policy that notifies a user when they are within a certain radius from a task pickup location, or as a more complex policy, such as a decision tree based on contextual factors (e.g., notifies if the user is close by and has not been notified in a while and the deviation is less than 200 meters).
- The *supply model* predicts how many potential users a system is able to reach given a notification policy. Given a set timeframe, it considers how likely people are to

meet the conditions of the notification policy and thus be notified of the task (e.g. how likely people are to be within a certain radius of a pickup location for delivery). Given that supply can fluctuate across the course of a day based on people's routines, system designers may wish to model such fluctuations explicitly should they affect the timeliness of task completion and the degree of disruption a notification policy may cause. To better estimate supply, system designers may also explicitly model human mobility patterns (e.g. [**47, 96, 15**]).

- The *pickup model* predicts how likely a user is to accept a task when notified under the conditions of the notification policy. For example, we may expect that the pickup rate would be higher when people are nearby the pickup location, and when completing the task would be convenient for them (e.g., if they are likely going towards the drop-off location already).
- The *value function* is used to encode goals system designers and stakeholders care about, such as system efficiency and volunteer needs, and evaluate the outcomes with respect to the measures over encoded goals. For instance, system designers may encode the rate at which tasks are completed as a proxy measure for system efficiency. They may also encode the number of notifications being sent to volunteers as a proxy measure of disruption to volunteers, or how many times a volunteer helped in the past as a proxy measure of overburden. Depending on the kind of communities or services system designers and stakeholders may want to promote, they may value competing goals differently, for example by using parameters as weights across measures. To better capture what goals stakeholders may care about and how to value competing goals,

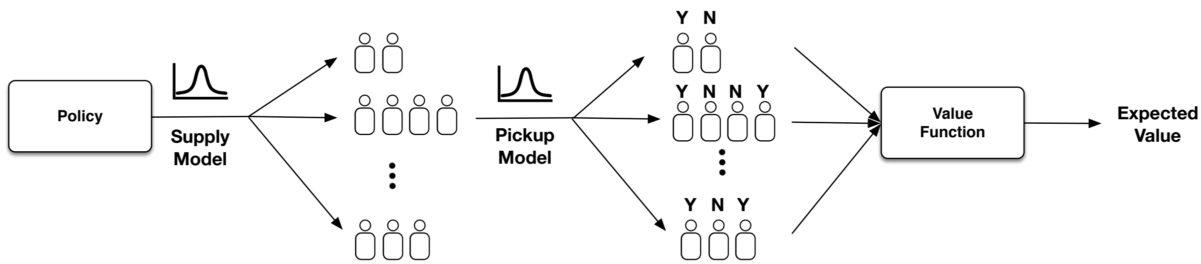


Figure 4.3. An example workflow of the opportunistic supply management framework for a given policy. The supply management framework simulates over possible unfoldings of how many people may show up and who will accept or decline tasks. Given the possible outcomes and the intended goals, the supply management framework evaluates the value of enacting the policy and compares across policies to choose the optimal one.

system designers can survey and elicit desired goals from different stakeholders [**141**] or use a participatory design method to engage end-users [**88**], respectively.

Given a policy from the policy model, the *simulator* uses the supply model and pickup model (e.g. trained on historical data of volunteers' past routes and previous task pickups or rejections) to simulate and evaluate the desirability of outcomes across possible scenarios that might unfold using the value function; see Figure 4.3. Each scenario results from drawing a sample from the supply and pickup models, which unfolds how many volunteers may be notified of tasks, and how many of these volunteers will accept and reject tasks.

The *decision manager* determines the optimal policy by using the value function to evaluate the results of the simulator for different policies. By evaluating the value function over outcomes across simulated scenarios, the decision manager is able to determine the effectiveness of a policy across many possible unfoldings, and not just with respect to a single, most likely scenario based on average number of people show up and average number of notifications being sent. This provides a more accurate measure of the effectiveness of policies given the inherent uncertainty.

Using the supply management framework, we can (1) find goldilocks zones that optimize the desired balance for intended system goals and volunteer needs; (2) encode and promote specific goals; and (3) make dynamic adjustment when previous decisions become ineffective. For (1) and (2), based on the value function that reflects the relative importance of various goals a certain community may have, the decision manager can find the best policy for a given situation (e.g., amount of demand) based on what it is that the community prioritizes. For example, if the community prioritizes system goals much more than volunteer needs, the decision manager may optimize the choice of policy as demand increases by effectively shifting from conservative policies to more aggressive ones.

For (3), we can use the decision manager to dynamically adjust the decisions we make over the course of the day when its choice of policy becomes ineffective. This can happen when—due to uncertainty in availability and willingness to help, or inaccuracies in our models— fewer people show up than expected, or tasks are completed more quickly than expected. In such cases, dynamic adjustment allows the system to incorporate new information about situations on the ground and recompute what would be the optimal policy for the rest of the day. We hypothesize that having such flexibility to devise custom-tailored solutions over the course of the day allows the system to better meet the intended goals, and can help to nudge policies in the right direction when models are inaccurate.

4.3.4. Applying Supply Management to Peer-to-Peer Delivery

We illustrate in this section how we might use the supply management framework to manage the supply of volunteers for peer-to-peer delivery. To model policies, we may for example consider a policy model that consists of a class of *radius-based notification policy* that notifies potential volunteers when they come within a certain radius around the pick up location. The radius affects both the likelihood of pickup and the number of people notified. A task notification policy with a small radius reaches fewer people, but since they are closer to the pickup location, the people reached may be more likely to pick up and then deliver packages along their route. A task notification radius with a larger radius will reach (and disrupt) more potential volunteers, but many of these volunteers may be less able or willing to go out of their way to pick up a package. To capture such differences empirically, we can build supply and pickup models that predict people's likelihood to be available to help. To evaluate outcomes, we can encode measures of task completion and disruption (e.g., how quickly tasks are picked up; how many people are notified and potentially disrupted) into the value function.

To find goldilocks zones that optimize the balance for intended system goals and volunteer needs, the supply management framework simulates different radius-based notification policies and determines the one that best balances system goals and volunteer needs using the value function. To *prioritize and promote specific outcomes*, such as not overly disrupting volunteers who are unlikely able to help, the value function may encode a cost for disruption (e.g. number of notifications being sent) so as to penalize outcomes that achieve high task completion rates but do so at the cost of notifying and disrupting many volunteers who could not help. As tasks may complete more or less quickly should more or less people be available and decide to help than was predicted, we can also *make dynamic adjustments* over the course of the day to recompute policies given what tasks were completed and situations that may arise over the rest of the day. For example, while running a policy determined by supply management, should actual outcomes be different than predicted by mid-day (e.g., less than predicted number of packages were picked up), supply management may choose a more aggressive policy for the rest of the day by recomputing the policy based on the models given conditions at the middle of the day.

4.4. Simulation Study

We conducted a simulation study in a community-based delivery setting to understand (1) how supply management finds the goldilocks zone that can optimize the balance of intended system goals and volunteer needs; (2) how prioritizing different goals affects the decisions supply management makes and how it finds the goldilocks zones in different contexts to promote those goals; and (3) how dynamic adjustment can lead to better decisions even when models are inaccurate. A simulation study is particularly useful for understanding the performance of supply management and the rationale behind its decisions because it allows us to experimentally vary the choice of value functions and the model accuracy.

4.4.1. Modeling

As we discussed in the previous section, the supply management requires *models* of notification policies, volunteers, and goals. We discuss how we implemented these components for simulations.

4.4.1.1. Notification Policy. We implement a simple, radius-based notification policy that notifies potential volunteers when they come within a certain radius around the pick up location. We consider two policies: *at-location policy* and *at-distance policy*. The at-location policy notifies people when they enter the pickup location. Since this policy notifies people when they are already at the pickup location, people will be more likely to help. However, this policy may not be able to reach many people and as a result it may not have the supply needed to get enough tasks completed or completed in a timely manner. The at-distance policy notifies people when they come within 100 meters (which equates to approximately one street block in the location where the study took place) of a pickup location. Since this policy notifies people who are less than a block away, this policy can reach people who are still likely to help, but at the same time it also starts notifying more people who are less likely to help because they are farther from the pickup location. As this policy captures more people who are close enough but not necessarily going to the pickup location, this may result in getting more tasks done in shorter time than when applying the at-location policy.

4.4.1.2. Supply and Pickup Model. We build a supply model at a population: (1) considering the likelihood that any given person might show up during one of three time windows (morning: [7am-9:59am], lunch: [10am-12:59pm], and afternoon: [1pm-3:59pm]) within a notification radius; (2) based on the number of users, sampling from this distribution to get a distribution of the number of people that may be reached by the notification radius. Similarly, we generate a pickup model based on the likelihood of people within a certain radius accepting a task when notified. We trained a population-based model with more coarse-grained, 3-hour time windows (morning, lunch, and afternoon) instead of using an individual-based or more fine-grained time windows (e.g. hourly window) for two reasons. First, even though an individual-based model can better capture individual differences in routines and behaviors, it will require us to collect large training datasets to train an accurate model. Second, we wanted to capture major commute cycles or routines and the fine-grained models may be less accurate without large datasets due to the variability of people's routines within fine-grained time windows (e.g. some people may go to work around 8am while others around 9am).

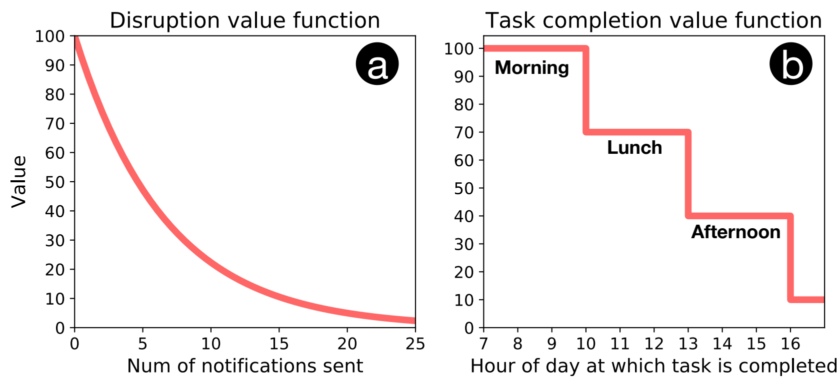


Figure 4.4. Value functions for disruption and task completion. (a) presents the disruption value function that computes the disruption value as a function of the number of notifications being sent. We set the disruption value function so that the disruption value decreases drastically when more than 15-20% of users are notified. (b) presents the task completion value function that computes the task completion value as a function of the hour of the day at which the task is being completed. We set the task completion value function so that the task completion value decreases across four time-blocks at the same decreasing rate.

4.4.1.3. Value Function. Figure 4.4 shows the value function for disruption (left) and task completion (right) respectively. The disruption value function models how over-notifications may cause disruption. We set the function to $a * e^{-bx}$, where $a = 100$ and $b = 0.15$. Due to potential user fatigue in 4-week long field deployment, we sought to set the disruption value function as conservative as possible in our studies by decreasing the value drastically after a certain threshold. We set the threshold as 15-20% of participants being notified because our initial dataset showed that, on average, 15.8% (SD: 4.9%) of participants were notified with the most conservative policy. We manually tuned the exponential decay function in a way that the value drops drastically after the threshold.

For the task completion value function, we model a decrease in value for later pickups across four time-blocks: morning, lunch, afternoon, and thereafter (100, 70, 40, and 10, respectively). We set a terminal value to 10 if a task that is not completed by the end of the day (e.g. 4pm in our studies), assuming the task will be completed sometime in the future. For simplicity, we assume all the delivery requests are posted at the beginning of the day (before 7am). We envisioned the kind of services that are useful for non-urgent deliveries, which is still good to be delivered on the same day and maybe earlier for the sake of requester experience but it's not a huge deal if it does not get delivered. Therefore, we weighted earlier deliveries higher than later deliveries to reflect the requester experience by decreasing the task completion value at the same rate across the time blocks. However, for other urgent-deliveries such as food deliveries, rather than decrease the task completion value at the same rate, we may decrease the task completion value more drastically after a certain threshold (e.g. 2-3 hours) to better capture the severity of the delay on the requester experience.

To relatively weight the value of disruption ($v(D)$), and the value of task completions ($v(C)$), we model the value function as $w1 * v(D) + w2 * v(C)$. By choosing these weights we can set the relative importance of disruption and task completion.

4.4.2. Generating Synthetic Datasets

We consider a realistic scenario in which we have 25 participants in our system. Based on prior work [**77**] and our own preliminary studies, we set the task pickup rate at 75% for at-location and 25% for at-distance, and set the supply rate as 15% for at-location and 60% for at-distance. We assume people will show up uniformly at random across the day.

4.4.3. Study Procedure

As noted previously, we use our simulations to study how supply management (I) finds goldilocks zones that optimize the balance for intended system goals and volunteer needs; (II) prioritizes and promotes specific outcomes; and (III) makes dynamic adjustment when previous decisions become ineffective. For all three simulation studies, we simulate policies using the provided models for 20,000 trials. We consider a range of demands from 1 to 10 tasks, and compute policies for each level of demand to evaluate the performance and understand decisions supply management makes as the demand changes. For the value function, we set equal weights for disruption and task completion, except for simulation (II) where we consider three conditions: a) volunteer focused ($w1 = 1$, $w2 = 0.5$), b) balanced ($w1 = 0.75$, $w2 = 0.75$), and c) system focused ($w1 = 0.5$, $w2 = 1$).

For simulation (III), we consider the opportunity to make a dynamic adjustment after the first time block. We evaluate the effectiveness of dynamic adjustment for varying degrees of inaccuracy in the supply model by setting the actual supply distribution of the at-distance policy in the first time block to 0%, 10%, 20%, 30%, 40% off from the distribution of the trained supply model. For simplicity, we assume that actual pickup distribution for at-distance, actual pickup and supply distributions for at-location are identical to their corresponding trained models.

4.4.4. Measures and Analysis

We measure the performance of the supply management framework and other fixed policies by considering the *expected value* of enacting a policy or policies. The expected value provides a measure of how good each policy is with respect to the goals we care about, across the distribution of possible outcomes that can arise given the uncertainty in people's mobility and decisions. To compute the expected value of a given policy, we draw a sample from the supply and pickup models for each trial, compute the value of simulated outcomes by using the value function, and compute an average value across all trials (20,000 trials in our studies).

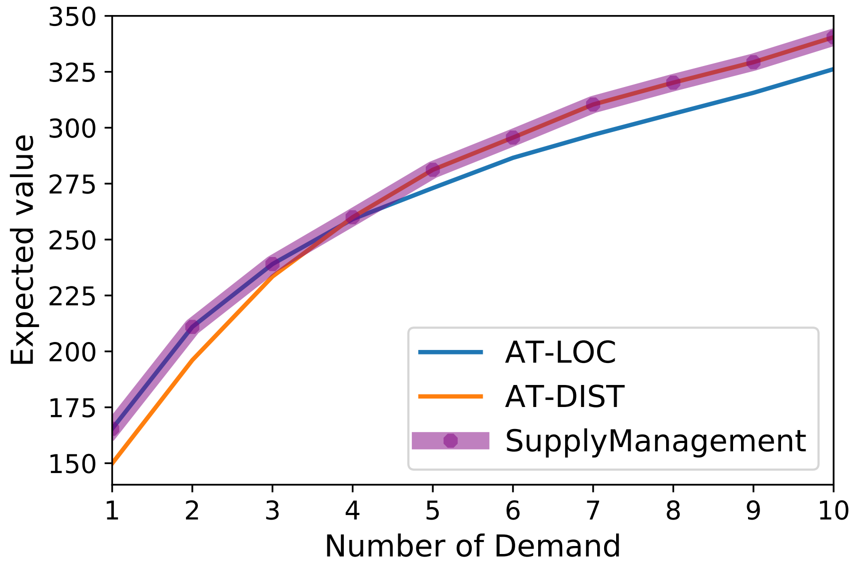


Figure 4.5. Supply management chooses the best policy that achieves the desired goals in balancing system efficiency and the needs of volunteers, while other fixed policies, the at-location policy (blue line) and the at-distance policy (orange line), would make trade-offs depending on the changing demand. Supply management (highlighted purple line) always chooses the policy with the highest expected value based on the changing demand. For example, supply management chooses the at-location policy when the number of task demand is less than 5, and chooses the at-distance policy when the number of task demand is more than 5.

4.4.5. Simulation Results

4.4.5.1. Simulation I Results: Finding Goldilocks Zones. Our results show that supply management can find goldilocks zones that can best meet the desired goals that are tailored to varying demands. Figure 4.5 shows that the opportunistic supply management chooses the best policy to balance system goals and volunteer needs for every value of demand, while fixed policies' performance varies with the changing demand. For example, when the demand was less than 4, at-location policy outperforms at-distance policy because it can get things done without having to disrupt too many people. However, as the demand increases, at-distance outperforms at-location

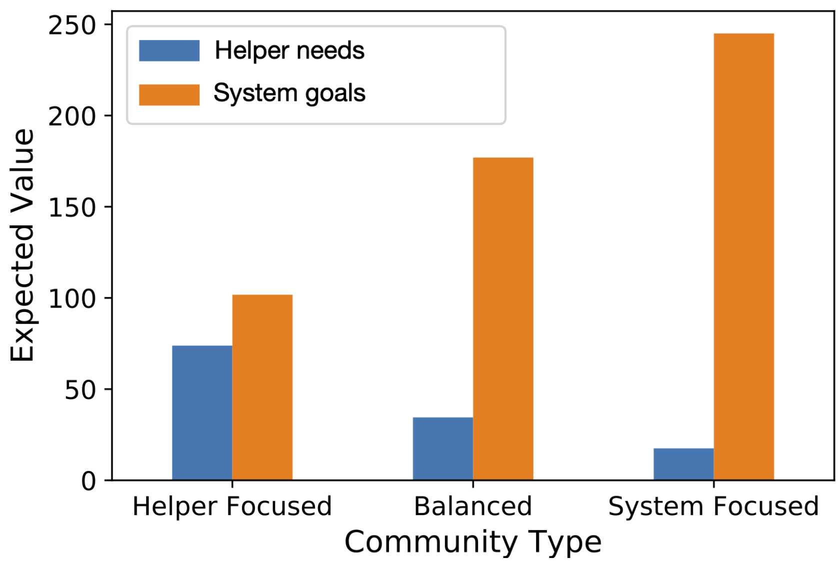


Figure 4.6. Supply management makes decisions that can maximize different outcomes that different types of community may care about. Shown are dimensions of disruption as volunteer needs (blue bar) and task completion as system goals (orange bar) being prioritized for three different types of community, namely (1) volunteer-focused, (2) balanced, and (3) system-focused. While setting a value function that is system-goal focused leads to higher task completion value when compared to a value function that is balanced or volunteer-focused, the value function that is system-goal focused also leads to highest disruption (i.e. lowest value).

policy because higher disruption is compensated by its ability to complete more tasks to meet the desired system goals. This shows that supply management can find an optimal policy as demand changes to maximize the outcomes that the system cares about.

4.4.5.2. Simulation II Results: Promoting Different Goals. Supply management makes decisions that can maximize different outcomes that different types of communities may care about. Our results show that setting a value function that is system-goal focused leads to higher task completion value when compared to a value function that is balanced or volunteer focused (see Figure 4.6). However, we also see that the system-focused has the highest disruption (e.g. lowest

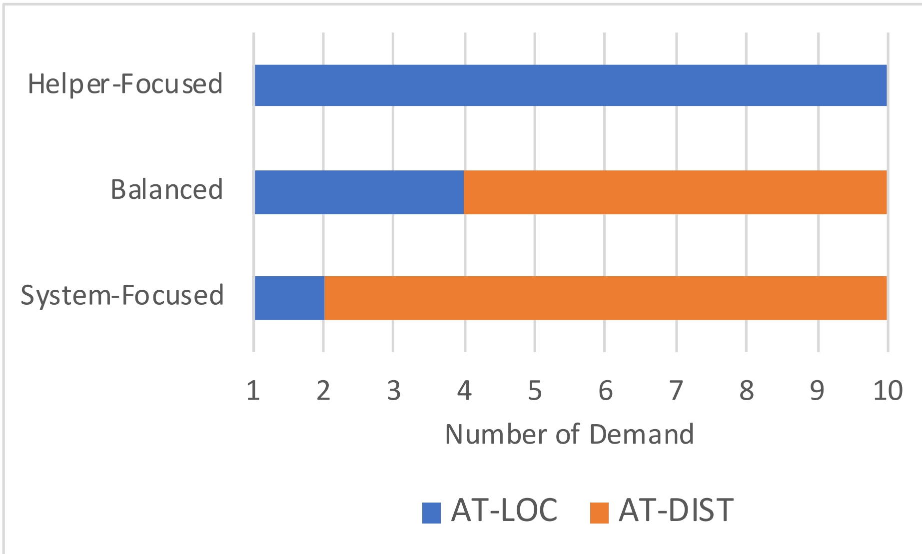


Figure 4.7. Supply management makes different decisions depending on varying demands and the goal orientation of the community, namely volunteer-focused, balanced, and system-goal focused. In this simulation, supply management chooses the at-location policy (blue) for the volunteer-focused community regardless of demands because it cares so much about disruption. In contrast, supply management chooses the at-location policy until demand is 2, and starts choosing the at-distance policy (orange) to get things done faster for the system-goal focused community.

value), compared to a value function that is more balanced or volunteer-focused. This means that supply management can allow stakeholders and system designers of a community to encode the outcomes that they care about (e.g. [**141, 88**]), and supply management can make decisions that can best meet those goals.

Figure 4.7 illustrates how the supply management framework makes optimal decisions with respect to both the situation at hand (e.g., how much demand needs to be filled) and with respect to the goal orientation of the community (e.g., volunteer focused, balanced, or system focused). In this case, for the volunteer-focused community, because it cares so much about disruption, it chooses at-location regardless of demands. In contrast, for the system-focused community, it chooses at-location until demand is 2, and starts using at-distance to get more things done faster. For the balanced community, supply management chooses at-location until demand is 4, and chooses at-distance afterward. By making effective decisions based on situations and goal orientation, the supply management framework is able to identify policies that are simultaneously aligned with the goals of a community, and properly tailored for the situation on hand.

4.4.5.3. Simulation Results III: Dynamic Adjustment. Our results show that dynamic adjustment outperforms non-adjustment by gaining 9.19% more values when models are accurate. These gains come from two types of scenarios. First, in scenarios where the actual number of supplies and pickups differ from the average case as predicted by the model (i.e. the rare probability events), it is better for supply management to adjust to a less aggressive policy (at-location) when significantly more tasks are completed than expected and a more aggressive policy (at-distance) when significantly less tasks are completed than expected. Second, when actual situations are as expected in the average case, having the ability to make adjustments can still lead to achieving better outcomes because supply management now has more flexibility to use a combination of policies to better meet the goals. These findings illustrate that how incorporating new information about situations on the ground and providing flexibility to devise custom-tailored solutions allows the supply management framework to produce even better results.

Our results also show that dynamic adjustment continues to outperform non-adjustment when models are inaccurate; see Figure 4.8. As more noise is added to the model, we see that dynamic adjustment (with inaccurate models) continues to outperform non-adjustment (with the correct model). By re-optimizing based on conditions in the middle of the day, dynamic adjustment (even with a less accurate model) can still shift policies in a more effective direction in ways

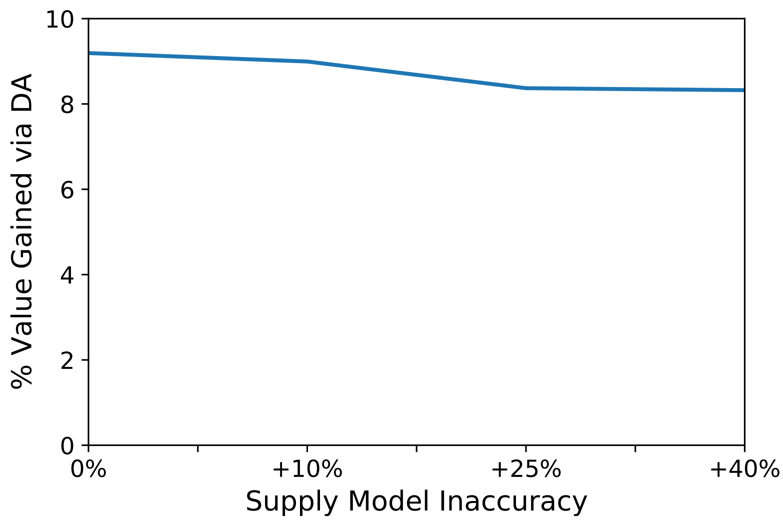
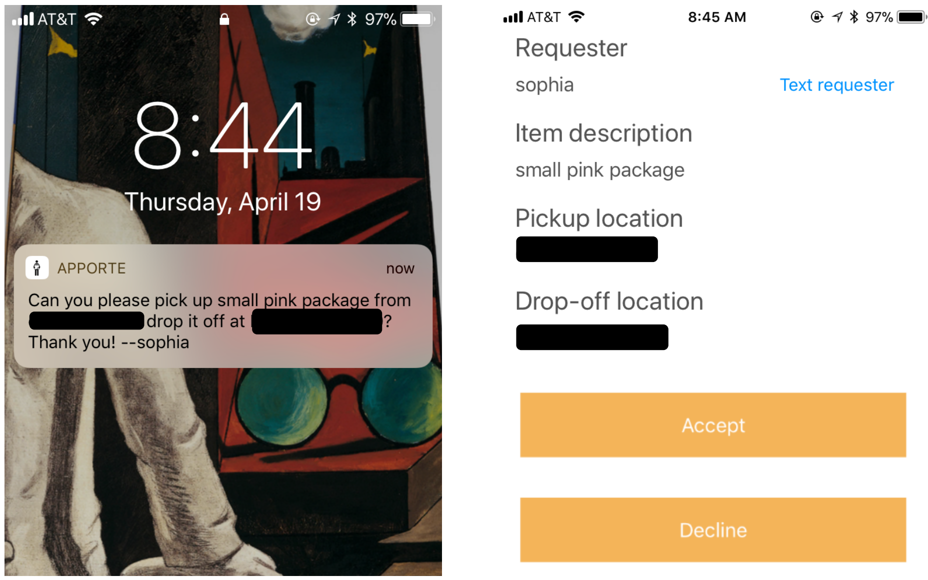


Figure 4.8. For varying degree of noises in the model, the percent value gained via dynamic adjustment with regard to the value gained with non-adjustment with perfect model.

that using a fixed policy throughout the day (even when it is optimized by supply management) cannot.

4.5. Study 2: Field Deployment

While our simulation study demonstrated the advantages of supply management over fixed policies and supply management's ability in promoting different goals by making different decisions, we could not observe whether supply management's chosen policy would be the best policy in the real-world and whether real-world outcomes and users' perceptions reflect the encoded goals. To complement our simulation study, we conducted a 4-week long field deployment of supply management to understand (a) how supply management performs (i.e. choosing the best policy) in an actual deployment; and (b) whether real-world outcomes and



(a) **(b)**

Figure 4.9. Apporte: a peer-to-peer delivery mobile application. (a) an example screenshot of a task notification that a potential volunteer receives when they pass by a pick-up location. (b) an example screenshot of an app when a potential volunteer clicks a notification. The potential volunteer can see the details of a request and can click "Accept" if they decide to help and click "Decline" otherwise. The volunteer can also text the requester if they need to coordinate with the requester.

users' perceptions are consistent with the encoded goals that supply management seeks to promote.

4.5.1. Apporte: Peer-to-peer Delivery Application

We developed a prototype, Apporte, a peer-to-peer delivery mobile application where users can request item deliveries and have the system notify people who pass by the item pick-up location. A user can post a request by providing an item description, a pick-up location, and a drop-off location. A potential volunteer who passes by the pick-up location will receive a notification asking if they can help deliver the item (Figure 4.9a). Once the potential volunteer clicks the notification, they can see the details of the request and texts the requester if they need to coordinate. If the potential volunteer decided to help, they can click “Accept”, otherwise click “Decline” (Figure 4.9b).

4.5.2. Participants

We recruited 26 undergraduate, graduate students, and staff members of a mid-sized U.S. university through mailing lists, social media, flyers, and word of mouth. Participant ages ranged from 19 to 37 (M = 24.75, SD = 4.04), with 15 female and 7 male, 1 genderqueer, 1 nonbinary.[1] The study took place over 4 weeks, during which participants completed a pre-study survey, used our application as a part of their daily lives, and completed a post-study survey. We compensated participants with a $65 gift card for their time spent on surveys and installing our prototype, but did not incentivize their behavior during the study (i.e., no monetary incentive was provided for task completion).

4.5.3. Modeling and Data Collection

4.5.3.1. Modeling. We use the same models for notification policies, supply, and pickup model as we did in the simulation. To implement the notification policies in the deployment, we used a bluetooth low energy beacon with the broadcasting signal power of -20dBm (which translates into approximately 3.5 meters in distance) and 200ms as the advertising interval for *at-location*, and we used a geo-fence with 100 meters as the radius for *at-distance*. To build population-based supply and pickup models, we used the first two weeks (10 weekdays) of the study as a training phase during which we collect data on people’s mobility and task pickup decisions within the

[1] 26 installed the app, one person dropped out of the study because her iPhone did not have consistent internet access, and one person did not continue using the app after the installation. 24 completed the post-study survey.

area where the study deployment takes place over the subsequent two weeks (see details in the next section). Based on empirically observed frequencies, we used beta distributions to model the likelihood of a user entering a notification radius across three time-blocks (morning, lunch, and afternoon), and the likelihood that a user accepts a task when pinged within that radius. This allows us to model noise in our estimates, whereby the simulator can first draw a probability of task pickup or supply from the beta distribution, and then simulate unfoldings based on that probability.

In our deployment, we wanted to create a setup in which volunteer experience is prioritized over system goals. This allows us to see whether the chosen policy can support system goals while ensuring low disruption to our participants (in noticeable ways). This setup also mitigates the risk of fatigue and notification blindness during the study itself, which can affect our findings should participants drop out. To do this, we set a lower weight of 0.15 for task completion and kept the weight for disruption at 1. Together with the trained supply and pickup models, this value function led supply management to choose at-location as the optimal policy for the deployment.

4.5.3.2. Data Collection Setup for Training Data. During the 2-week training phase, we exposed each participant to both policies (i.e. within-subjects design) on different days and weeks to help account for large individual variation in task pickup rates and large daily and weekly variation in both task pickup rate and mobility patterns. For instance, on day 1 of week 1, we collected pickup and supply data for at-location, and on day 2 of week 1 collected that of at-distance; switching between the two until the end of week 2. By doing this, we aimed to minimize individual, weekly, and daily variation in both pickup and supply data. During weekdays, task requests are made at 6:55am everyday and task request notifications are sent between 7am and

3:59pm. Task requests are not made during weekends. To prevent over-notifying participants, we also set a notification interval to 3 hours.

In the at-location condition, it is straightforward to collect supply data: the number of people that showed up within the range of the at location setting is the supply for at-location and the number of people crossing the geo-fence is the supply for at-distance. In the at-distance condition, we cannot directly observe (what would have been) the at-location supply because some participants may have gone to the pickup location because they were notified (at distance) and not because they were already going there. To eliminate this potential confound, we followed up with the participants who accepted at-distance notifications at the end of each day and asked them whether they had already planned to go to the pickup location or were just going there for the pickup. We count them as available supply for at-location if and only if they responded that they were already planning to go to the pickup location.

4.5.4. Study Procedure

In the latter 2 weeks of the deployment, all 26 participants used the supply management version of our prototype. Participants consented to enrollment and received instructions that asked them to pick up an item from a local coffee shop that was chosen as our pick-up location (Figure 4.10a) and to drop off at a specified drop-off location.

We used a pre-study survey to assess where the participants spent most of their time on campus in order to set a drop-off location that would be on the way for most of the participants. Based on the survey data, we chose an intersection in front of our school's engineering building as a drop-off location because most of our participants pass by that location as part of their everyday routine. To reduce the cost of coordinating with requesters to hand over items—which

(a) (b)

Figure 4.10. Pick-up location (a) and Drop-off location (b).

may affect volunteers' willingness to help but is not the main focus of this study—we made a collection box (Fig 4.10b) so that the participants could easily drop off items without having to directly coordinate with requesters.

Task requests were made at 6:55am every day and task request notifications were sent between 7am and 3:59pm. Since we did not want the package size to affect willingness to help, we only requested packages that were small enough to be carried in one hand. We requested 4 tasks, which was the median number of task pickups during the 2-week training phase, every day for 10 days. To prevent over-notifying participants, we set the interval between task notifications (per user) to 3 hours.

4.5.5. Measures and Analysis

To evaluate the supply management's choice of policy in the real world, we compared the *expected value* of supply management's chosen policy with the other policy, as a baseline, with respect to a supply model built using the actual, realized supply. This allows us to know, given how many people actually showed up, whether the policy selected by supply management is indeed optimal for the deployment. We followed the same procedure of the simulation study to

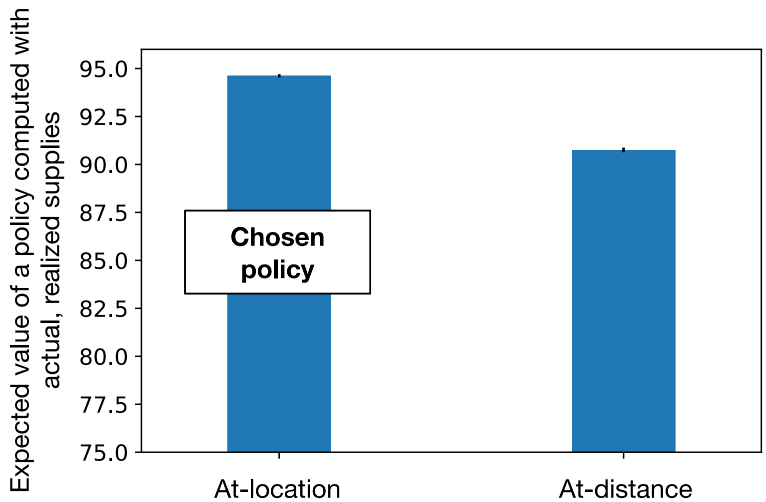


Figure 4.11. Expected values of each policy computed with actual, realized supplies. Supply management chose the at-location policy during the deployment and the results show that the average value of using the at-location policy is higher than using the at-distance policy in practice. Error bars indicate standard error of the mean.

compute the *expected value* of enacting a policy, but to provide a fair comparison between the chosen policy vs. the other policy, we used actual supplies but still used the trained pickup model since we are able to only observe actual pickup rates for the chosen policy. To help us understand users' perceptions of disruption, we measured how disruptive users found the application during the deployment period (the last two weeks) on a 5-point Likert scale (1: Not at all disruptive, 5: Very disruptive) by using a post-study survey.

4.5.6. Results of Field Deployment

4.5.6.1. The chosen policy is more effective than the other policy in practice. Results show that supply management's chosen policy outperformed the other policy when compared using a supply model trained based on actual, realized supplies. The expected value of using the chosen policy (i.e. at-location) is 94.62, and the expected value of using another policy (i.e. at-distance) is 90.06; see Figure 4.11. This indicates that supply management chose the best policy for our deployment study, not only in simulation with respect to its trained model (where it always produces the best outcome), but also in practice with respect to how many people actually showed up during the deployment.

While this result shows that supply management chose the at-location policy as the best policy in the current deployment setup, it is worth noting that supply management may choose different policies in other setups since choosing other policies may be an optimal decision. For example, supply management may choose a more aggressive policy (i.e. at-distance) if the number of task demands is very high because the cost of task incompletion will outweigh the cost of disruption; similar to what we showed in the simulation results (Figure 5). Supply management may also choose the at-distance policy to get things done faster if task completion is highly prioritized over disruption, even with the same number of task demands; as shown in Figure 7. To sum up, there's no single policy that is optimal across all possible situations, but supply management allows the system to find the policy that is best suited for the situation at hand and the goal orientation of the community without having to manually choose a policy in practice.

4.5.6.2. The real-world outcomes were in accordance with the encoded goals. In accord with encoded goals, results show that users were minimally disrupted while the system was still getting some tasks done. On average, only 12.5% (3 out of 24) of participants received notifications on any given day, and this still led to completing 60% of all tasks (2.4 out of 4).[2] Post-study survey responses indicate that the perceived cost of disruption over the deployment

---

[2] As a point of reference, it is worth pointing out that during the training data collection phase the at-distance policy notified 13.8 participants on average (SD: 2.71) while the at-location policy only notified 3.8 participants (SD: 1.17).

period was low; the average rating was 1.5 (SD: 0.59) on a 5-point Likert scale (1: Not disruptive at all, 5: Very disruptive). In other words, by setting a value function that prioritized low disruption over task completion, supply management produced actual, realized outcomes that did just that by steering the system in ways that supported users and their contributing in the way that was intended.

In open-ended post-survey responses, our participants reiterated that they did not find the application disruptive during the deployment period because they either did not receive any notifications or rarely received notifications. One participant said they did not receive any notifications despite being in the region: *"Not very disruptive, I didn't get any notifications the last 2 weeks."* Another participant mentioned how they expected to receive more notifications based on their experience over the first 2 weeks but did not get as much: *"Not disruptive at all – I remember I was surprised that I didn't get more notifications when I was in the area."* One participant elaborated on how their perceived disruption changed from the data collection period to the deployment period: *"The app was much more location-specific. That is, I only received requests when I was exactly in front of [the pickup location] or just under the [train station] platform. Earlier in the study I would receive requests when I was still in my apartment, which was a little more disruptive. In the last two weeks of the study this was less disruptive."*

In summary, these results indicate that the encoded goals that the system tried to prioritize—keep disruption low but still getting some tasks done—are reflected in the real-world outcomes such as the number of notifications sent and users' perceived cost of disruption.

4.6. Discussion

In this chapter, we introduced *opportunistic supply management*, a general decision-theoretic framework that provides an architecture for modeling and optimizing task notification policies for engaging on-the-go volunteers. We demonstrated the effectiveness of opportunistic supply management through a simulation study and a field deployment, which highlighted the use of opportunistic supply management to (1) find goldilocks zones that balance system efficiency and volunteer needs, (2) promote specific goal orientations of a community, and (3) dynamically adjust policies when actual outcomes are better or worse than predicted. The implication of our results is that supply management provides a community-level flexible coordination mechanism that enables effective coordination among volunteers without imposing what each individual must do, in a way that fully considers people's potential experiences in helping (and being asked to help) and the outcomes of the system.

In the rest of this section, we discuss (1) the framework's applicability to other domains; and (2) limitations and future work.

4.6.1. Applicability to Other Domains

While we studied supply management in the context of community-based on-the-go crowd-sourcing system, the general framework can be applied to other settings (1) to balance between the needs of volunteers and the system goals in volunteer-based social and crowd computing systems; (2) to promote worker experience in commercial, on-demand services; or (3) to manage help-seeking and help-giving in workplaces and learning communities.

4.6.1.1. Balancing Between Volunteer Needs and System Goals in Other Volunteer-based Systems. In volunteer-based systems, we envision how community-level flexible coordination mechanisms like supply management can be used to effectively engage volunteers across the entire community to meet the system goals while still accounting for the needs of volunteers. For example, in existing collective action platforms where systems identify a set of potential volunteers given a cause or a social problem [**118**], using frameworks like supply management allows the systems to model number of volunteers and the likelihood of volunteers reacting to a cause given a recruiting strategy [**118**], the likelihood of having a successful campaign given the required critical mass [**22**], as well as volunteer needs such as forming new social connections with other participants [**118**]. Depending on organizers and platform designers' needs, supply management can choose a recruiting strategy that can best balance both the success of campaigns with the needs of volunteers. Also, in social micro-volunteering systems that rely on volunteers' existing social network and crowdsourcing micro tasks [**14**], we could further reason about the desired balance between the system efficiency—that is contingent on a volunteer's posting to their social networks—with the volunteer's own needs in preserving their social capital with less-frequent posting.

In the absence of clear financial incentives, we also envision ways to explicitly model motivational factors and use these models to reason about people's likelihood of participation. For example, user models may incorporate factors such as personal interests or relevance in tasks [**27, 5, 120**], convenience threshold in participation [**131, 50**], or interpersonal bonds [**109**], or self-worth in completing tasks [**10**]. We could also better quantify users' intrinsic and extrinsic motivations for varying tasks by using a motivation scale (e.g. [**106**]) in building the user models. Our value functions may encode the needs of volunteers over the varying degree of intrinsic and extrinsic motivation, such as personal interests (e.g. topic similarities between their own interests and tasks) or potential social benefits (e.g. increasing tie strength [**44**] or interaction frequencies [**26, 36**]), and encode system values over quantity and coverage of contributions [**50, 43**]. Notification policies may further specify *how to present* tasking opportunities to a volunteer to better tailor to their motivational needs. For example, the system may highlight potential social benefits a volunteer may gain through the task, or highlight the cause or uniqueness of their contributions [**8**] depending on their extrinsic or intrinsic motivations.

4.6.1.2. Promoting Worker Experience in Commercial, On-demand Services. We envision how system designers may use the supply management framework to meet the needs of crowd workers, such as increasing wages and developing skills, while also meeting the needs of a platform and of requesters. One core obstacle for crowd workers to earn higher wages is requesters who exploit crowd labor by not accepting valid tasks or offering a low price per task. To overcome this obstacle, existing worker support tool helps crowd workers find better tasks by looking at requester's credentials (e.g. hourly pay, fair score, and reward score in Amazon Mechanical Turk [**116**]). While worker support tools are useful in filtering out bad requests, we envision approaches that can reduce bad task requests in the first place by using supply management as a simulation tool to highlight the consequences of exploitation and requesters' credentials on potential task outcomes, and to suggest possible actions that can be taken on a requester's side (e.g. lower the required experience or increase the price) to attract more or qualified crowds to improve task outcomes. This may result in a virtuous cycle where requesters care more about their credentials and thus receive high quality task outcomes, and crowd workers earn higher wages with fewer requesters trying to exploit their labor.

In addition to increasing wages to improve crowd worker experience, researchers also sought to train crowd workers (e.g. training for subjective judgment tasks like identifying accessibility issues [**52**]) and help them develop relevant skills. While there exist methods for skill development such as providing peer or expert feedback [**24, 33**] and offering mentorship [**122**], we envision ways to help crowd workers develop skills through "learn by doing" [**31**]. For example, we may model the rate at which workers learn a task type, and reason about the effect of task completion on the worker's skill development, as well as the effect of the worker's current skill level on the quality of tasks.

We also envision how mechanisms like supply management can provide *more flexibility* to workers in commercial, on-demand services to provide better worker experience while still providing good quality of services. While current on-demand or gig economy services provides some flexibility to workers as they can decide when and where they want to be "online" and complete tasks, the workers have to accept the tasks most of the time that an algorithm assigns to them [**89**]—which reduces the "online flexibility" that may affect worker experience [**137**]. Unlike existing task assignment algorithms for such services that mostly assume people will accept the tasks (e.g., [**90, 136**]), using frameworks like supply management can provide more "online flexibility" to workers as supply management incorporates uncertainty in availability and participation while still making sure that it meets the desired quality of services. This can allow on-demand services to improve worker experience as the systems offer more flexibility to workers in how and when they work, while still providing a good quality of services.

4.6.1.3. Managing Opportunistic Help-seeking and Help-giving in Workplaces and Learning Communities. In workplaces and learning communities, we envision that community-level mechanisms such as supply management can be used to effectively engage individual volunteers and manage efforts and attention among a pool of potential volunteers as their availability, ability, and willingness to contribute change over time. In such settings, we may explicitly model people's expertise [**97, 135, 125**], interruptibility [**39, 64**] and availability [**57**], locations, and schedules, and use these models to inform people's likelihood of being available to provide effective help. Our value function may encode system values over the timeliness of help and the quality of help provided, as well as over volunteer's experiences in providing or receiving help, along with costs for disruption or overburden. Using these models, supply management can effectively promote help-seeking and help-giving in ways that are considerate of community and system goals.

4.6.2. Limitations and Future Work

Our field deployment focused on understanding whether supply management's chosen policy was indeed the best policy (among a set of policies) in practice, and whether the real-world outcomes and users' perceptions reflected the intended goals that the system sought to promote. As a result, we chose to design a study that compared the supply management's chosen policy against the other policy, and that could provide quantitative and qualitative evidence of real-world outcomes with the chosen policy. While we think this design was appropriate for addressing our immediate research questions, it does have some limitations. In the rest of the section we discuss limitations in (1) lack of baseline approaches; (2) building and updating richer models; (3) study setup; and (4) measuring other impacts of notifications on volunteers' lives.

In our field deployment, we did not compare supply management to other baseline approaches such as manually choosing policies to uncover how much better supply management performs against other approaches in practice. In the future, we would like to conduct a longitudinal, between-subjects study in which users in one condition would use a prototype with supply management, while those in another condition would use a prototype with fixed policies. This longitudinal study would help us better understand the relative strengths or potential weaknesses of the supply management in practice.

While our studies used only simple, radius-based task notification policies, we may represent a policy as a decision tree that specifies conditions over the predicted cost of diversion [**59, 60**], contextual factors such as the user's current activities and schedules [**63, 32, 7, 82, 91**], and a footfall of a task location, and the user's prior responses to notifications [**77**]. As future work considers richer policy models, we may wish to employ more advanced optimization and approximation techniques (e.g., [**100**]) that can discover an effective policy without exhaustively comparing all policies.

Our field deployment used population-based models and did not update our models due to the short duration of field deployment. For example, we used the disruption value function that models how over-notifications may cause disruption to the community as a whole, and it did not take into account individual differences in the level of tolerable disruption. In the future, we may better capture the needs of volunteers at an individual-level with a value function that computes and aggregates the expected value per individual to account for different levels of tolerable disruption. The supply management framework can easily accommodate individual-based models and there are no technical challenges in doing this. While in our studies we allowed system designers to specify how they want to model the desired goals in the value function, but we may also train the value function based on explicit user feedback on their experience and satisfaction, which can be collected via methods like ecological momentary assessment [**119**]. To overcome the potential challenge in balancing between eliciting new information and disrupting users with notifications that sample their experience, we may use sampling methods that reason about the cost of interrupting people with the benefits of information gain in modeling [**71, 112**].

Our field deployment setup is limited in that it took place in a medium-size American university with undergraduate, graduate students, and staff members as study participants, and it treated all tasks as identical. In the future, we may partner with existing communities (e.g. community-based timebanking services [**9**]) in which community members are already helping each other with deliveries or other physical tasks in their community. To deploy supply management to an actual community, we should take into consideration task properties, such as task urgency, a task location, a task requester, as well as an estimated task time. Future work may investigate how and why the supply management makes certain decisions depending on intrinsic properties of tasks, and how a system may communicate their reasoning to volunteers to assist volunteers' decision making processes in whether or not to participate. Deploying systems in actual communities will help us gain insights into other real-world challenges and surface different stakeholder needs that we would otherwise not have been able to capture in a more controlled setting.

Beyond volunteers' perceived disruption, a future study may also examine how the choice of different notification strategies affect volunteer's routines and lives. For example, we may compute a detour distance from their planned routes, feelings of overburden, social-connectedness as a result of completing or rejecting tasks per notification policy. A longitudinal, between-subjects study will provide insight into volunteers' experiences and perception in providing help as a part of their daily lives.

4.7. Conclusion

This chapter introduces the idea of *community-level flexible coordination mechanisms* that coordinate opportunistic contributions across the entire community in a way that meets both the needs of volunteers and system goals. To enable this idea, we propose *opportunistic supply management*: a decision-theoretic framework for modeling, simulating, optimizing community-wide task notification policies that govern when, where, and to whom to notify of tasks across the community. From a simulation study and a field deployment, we found that supply management finds a goldilocks zone that optimizes the desired balance between help and system needs based on the situation on the ground and the goal orientation of a community, and supply management also chooses an optimal policy that steers user behaviors and real-world outcomes that are in accord with the intended goals.

Our work provides volunteer-based social and crowd computing system designers with insights on designing and implementing community-level flexible coordination mechanisms such as supply management into their applications. These insights are important for moving beyond systems that rely on coordinating contributions at an individual level towards systems that coordinating contributions across the entire community, and that can better meet the needs of volunteers and quality of services by having larger decision spaces in deciding who to engage with which tasks. We envision that future volunteer-based social and crowd computing systems that build on supply management framework may intelligently enact notification policies to govern how community members may help one another and prioritize different goals that they care about, so that the efforts of the community members are implicitly coordinated towards the intended goal of a community that they want to be part of as a member.

CHAPTER 5

Discussion

In this chapter, we first revisit core principles for designing flexible coordination systems; overview our technical approach for overcoming general challenges in realizing the idea of flexible coordination; discuss our technical frameworks for developing flexible coordination systems; and consider how we can design flexible coordination systems in other domains. Then, we discuss the benefits of a mixed-method approach in designing and deploying flexible coordination systems. Finally, we discuss some ethical concerns in designing and deploying flexible coordination systems.

5.1. Flexible Coordination Revisited

Volunteer-based physical crowdsourcing systems help connect people to solve local and communal problems that are difficult to achieve with a small group of dedicated volunteers. These volunteer-based systems need to attract enough willing volunteers who can make useful contributions to ensure long-term viability. However, current approaches are limited in that they either (a) provide volunteers flexibility to attract people but struggle to meet desired system goals with uncoordinated contributions; or (b) they directly coordinate contributions to meet desired system goals at the cost of taking away volunteer flexibility and autonomy.

To overcome such shortcomings, this thesis introduces the idea of *flexible coordination*, or ways to provide volunteers flexibility while still coordinating useful contributions. A flexible coordination system surfaces opportunities for volunteers to contribute that are within volunteers' routines or that support their goals, while that are still useful for achieving system goals. Unlike existing approaches that give volunteers full control over when and how they want to engage with opportunities, flexible coordination follows people's routines and proactively suggests opportunities when volunteers' needs align with that of a system. To effectively coordinate contributions without ever imposing on their routine or requiring that they accept a task that is suggested to them, flexible coordination preserves flexibility by accounting for uncertainty in people's availability and participation but still coordinates contributions that are within people's routines to achieve optimal outcomes.

To design effective flexible coordination systems, we need to address two core challenges. First, we do not have a fixed set of opportunities to optimize over because opportunities may dynamically arise depending on how people's routines are carried out. Second, the quality of opportunities is relative to an individual's and other's routines, and how our knowledge of the world may change; therefore, evaluating the quality of opportunities in isolation can be ineffective.

We first demonstrated how we may identify good opportunities within a single user's routine by building up an understanding of a user's future trajectories and an understanding of our knowledge of the world and people's interests and goals. We also demonstrated how we may identify good opportunities that may arise across multiple people's routines by considering other people's routines and their uncertain availability and engagement. As the set and quality of opportunities may dynamically change depending on how people's routines are carried out, rather than optimize over a fixed set of opportunities, we optimize and reason across possible unfoldings of opportunities that may arise.

5.1.1. Design Principles for Flexible Coordination Systems

Designing flexible coordination systems involves identifying good opportunities that can advance and balance individual and collective goals without ever imposing what each individual must do. In what follows, we discuss the core principles that we used for designing flexible coordinations systems.

Maintain user flexibility and autonomy but also surface good opportunities that can help advance people's goals. A core principle for designing a flexible coordination system is to preserve user flexibility and autonomy while still surfacing opportunities that can best advance people's individual and collective goals. This is in contrast to other intelligent systems that act as an algorithmic overlord that prescribes or imposes what each user must do (e.g. in services like Uber where a system assigns what each driver must do) [**89**]. For example, a flexible coordination system follows people's changing needs and goals, and surfaces opportunities to people that would otherwise have been unnoticed by people. While doing so, the system still provides people flexibility and autonomy to decide whether or not they want to act on the presented opportunities. By doing this, a flexible coordination system embraces the uncertainty in people's actions as a result of providing flexibility to people, but nevertheless still achieves globally effective outcomes that are in accord with people's goals without any explicit coordination. For example, Hit-or-Wait models the uncertainty in people's future trajectories, which affect which opportunities may arise within a user's routine, and finds good opportunities to elicit contributions that are convenient for them and are useful for achieving desired system goals. Likewise, supply management reasons about the uncertainty in people's engagement across the community, simulates possible unfoldings of scenarios when enacting a policy, and chooses an optimal community-wide policy that can best engage people with opportunities even when we do not know who will become available and be willing to help.

By maintaining flexibility, flexible coordination systems allow people to just go about their days, focus on their current goals and the task at hand, and contribute to other activities that suit their routine, that are convenient for them, or that are of interest to them. By suggesting opportunities to people within their routines, flexible coordination systems allow people to effectively advance and balance their goals without ever prescribing what they must do.

Simultaneously advance and balance individual and collective goals. A flexible coordination system needs to simultaneously achieve individual and their community goals. To do this, our technical frameworks for flexible coordination provide ways to encode and model goals that people and their communities care about and find ways to advance and balance both goals. For example, the supply management framework allows system designers and stakeholders to explicitly model volunteer's individual needs (e.g. low disruption) and collective goals (e.g. timely completion of task). Such modeling reflects volunteers' needs in focusing on their daily activities and goals, but at the same time, their other needs in wanting to help their community to effectively achieve collective goals when opportunities arise. Instead of choosing between the two goals, a flexible coordination system needs to support both goals as situations on the ground and people's interests and goals change.

Avoid overburdening people with opportunities (or information). A flexible coordination system should not overburden people with opportunities or information. A flexible coordination system treats people's attention and contributions as limited resources and devises bounded-resources solutions that can achieve effective outcomes given these limited resources. Instead of notifying whichever opportunity a person may come across, a flexible coordination system needs to continuously reason about opportunities that may arise within people's routines, evaluate the quality of opportunities within or across people's routines, and only notify people of the best opportunities to contribute to the desired system goals. By doing this, a flexible coordination system only engages people with opportunities when they can conveniently contribute within their routines and that are most useful for achieving desired system goals.

Reduce the cost of planning and coordination. A flexible coordination system also needs to reduce the planning and coordination burden on the people's side. This has the potential to increase the number of people volunteering to help others in a community. For example, we demonstrated how we can increase the task efficiency to provide a high quality of services by implicitly steering contributions towards where they are most needed by the system. We also demonstrated how we can also maintain community health to sustain participation by deciding which volunteers and how to engage them with different tasks so as not to overload or overdisrupt certain volunteers. Instead of putting the planning and coordination burden on the volunteers, we allow computers to take over planning and coordination challenges required for completing tasks towards meeting desired collective goals. However, unlike existing approaches that prescribe or pre-determine what each individual must do to better coordinate their contributions, a flexible coordination system needs to be able to coordinate contributions opportunistically even when it is uncertain how people's routines may unfold and which opportunities may become available.

5.1.2. Technical Approach for Flexible Coordination Systems

To design effective flexible coordination systems, this thesis introduces an overarching technical approach that builds an understanding of how people's routines may unfold, reasons about opportunities that may become available within their routines, and sets conditions under which to surface an opportunity to contribute by comparing across possible opportunities that may arise within or across people's routines.

In what follows, we reflect on how our technical approach builds upon and extends solutions in interruptibility and opportunistic planning, as well as solutions that use decision theory in other domains.

5.1.2.1. Interruptibility. In order to surface good opportunities to people, our technical approach builds upon prior work in interruptibility that investigates the disruptiveness of interruptions and strategies to reduce the costs of disruptions. One important factor that affects the disruptiveness of notifications is the moment of interruption. For example, prior approaches monitor (1) attentional states of users such as cognitive load; (2) current activities [**57, 39**]; and (3) transitions between activities [**54, 38**] to find good opportunities to interrupt. Likewise, our approach also sought to find opportunities so that people's routines can be minimally disrupted.

Our technical approach for flexible coordination systems extends existing techniques and approaches in interruptibility in two ways. First, while existing strategies evaluate the interruptible opportunities in isolation of other opportunities, our work provides ways to evaluate the opportunities across other opportunities that may arise within people's routines. While it is useful in some settings to evaluate the current opportunity in isolation (e.g. deciding whether or not to interrupt a meeting with an incoming call or defer the call [**58**]), we may need to evaluate the current opportunity with other opportunities that may become available as a person carries out their routine. For example, a virtual assistant may need to reason about when to remind a user of a task that the user expressed uncertainty about time (e.g. "Remind me to write a short email to John sometime tomorrow") [**111**]. In such settings, rather than evaluate the current opportunity in isolation based on a set of criteria (e.g. considering whether or not a person is near the computer), an intelligent virtual assistant may need to compare the quality of a current opportunity with other opportunities that may arise to find the better opportunity. For example, an intelligent virtual assistant may decide not to remind the user of a task even if the user is near a computer right now (e.g. if the user is preoccupied with other tasks) and wait for better opportunities in the future situations in which the user may go to a coffee shop and has nothing to do while waiting in line.

Second, unlike existing solutions that reason only about interruptible opportunities to better achieve individual goals (e.g. individual productivity), our technical approach reasons about opportunities to contribute towards collective goals while still helping people to advance their individual goals. As the success of achieving desired collective goals is dependent on other volunteers' uncertain availability and engagement, our technical approach provides ways to build up an understanding of many people's routines, reason about how setting different conditions to engage people across situations may unfold, and what outcomes they may collectively reach. By reasoning about opportunities across the community, our technical approach allows us to move beyond optimizing individual productivity to community-wide or organization-wide productivity where there's a need to coordinate efforts from community members without compromising individual goals.

5.1.2.2. Opportunistic Planning. Our technical approach draws inspiration from *opportunistic planning*, which considers the problem of choosing ideal opportunities for diversions on trips to a primary destination [**59, 67, 60**]. For example, prior work demonstrated how systems may opportunistically recommend unplanned waypoints (e.g. a rest stop or a refueling stop) depending on a user's need during the trip. To do this, existing solutions reason about potentially valuable opportunities and the costs associated with investing time in a diversion during a trip.

Analogously, our technical approach reasons about opportunities within or across people's routines and the costs associated with surfacing opportunities to users and the potential benefits of their contributions towards collective goals.

Unlike existing solutions that mostly assume there's a fixed set of opportunities to optimize over during a trip, our technical approach recognizes that people's routines are uncertain and opportunities may dynamically arise as people carry out their routines. Existing solutions assume that there's a fixed set of opportunities to optimize over because (1) a person's destination is known a priori; or (2) a system has to make a decision at a fixed moment in time given current opportunities (e.g. upon a user's request for recommendations). Rather than optimize among a fixed set of opportunities, our technical approach optimizes over immediate situations and possible future situations to resolve the unique challenge in settings where opportunities may dynamically arise. This allows people to carry out their routines as they wish but flexible coordination systems can still find good opportunities to dynamically coordinate their contributions under uncertainty.

5.1.2.3. Decision-Theoretic Approach. Given uncertainties about people's routines, and opportunities that may become available within their routines, how do we determine when and whom to surface opportunities? From the perspective of decision theory, decisions about when and whom to surface opportunities should be determined by the expected utility of actions. Our technical approach chooses an optimal decision that has the greatest expected value by taking into consideration the costs and benefits of surfacing opportunities and uncertainties in people's routines. For example, Hit-or-Wait considers when and which opportunities to surface by comparing the expected value of surfacing an opportunity now with the expected value of making a decision later if we wait. Supply management considers which community-scale policy set to govern who to engage with which opportunities by considering the uncertain engagement and availability of community members and simulating and comparing outcomes over all possible unfoldings of scenarios.

In modeling people's routines and guiding decisions about when to engage people with which opportunities, our technical approach bears resemblance to the use of decision-theoretic methods in online crowdsourcing to optimally control the allocation of tasks in workflows [**28, 92**]. Whereas efficiency is the primary reason for using decision-theory in earlier work, in our setting, the use of decision theory is further motivated by its ability to (1) provide seamless interactions [**56**] and (2) factor in volunteer experiences to simultaneously achieve their individual and collective goals. This enables flexible coordination systems that allow people to just go about their days, focusing on their routines, and make convenient contributions that seamlessly fit into their routines but that are still useful for achieving the desired system goals.

5.1.3. Technical Frameworks for Flexible Coordination Systems

Based on our overarching technical approach, we designed and developed technical frameworks that find opportunities within and across people's routines to achieve desired system goals while still respecting individual needs and goals people may have. While we only modeled simple individual goals in our studies, such as low disruptions and no deviation, we can easily model richer individual goals and preferences in our technical frameworks. For example, we may incorporate a richer user model that encodes people's individual preferences such as cost of diversion, level of commitment, and tolerable level of disruptions. Encoding rich individual preferences would allow us to make more custom-tailored strategies to meet their needs and have a richer decision space to devise solutions. For example, if the system knows someone is more committed and would not be bothered by system-initiated dialogs, the system may elicit more information from the user to reduce the uncertainty of their future routes (e.g. by asking their next destination). This allows us to make more informed decisions about when and which opportunities to surface to the user. Likewise, depending on people's perceived cost of diversion, the system may become more directed with some users who are more willing to deviate off of their existing routine (e.g. asking them to go out of their existing routes to search in places where no one searched) or become more opportunistic if people's perceived cost of diversion is high. In other words, incorporating richer user models will allow the system to make custom-tailored strategies that respect people's needs and constraints while maximally leveraging their efforts within their convenience and interests.

In addition, we can also incorporate a richer task model to capture different properties of tasks in our technical frameworks. While our supply management study only focused on simple, identical tasks (e.g. same pick-up and drop-off location for delivery tasks), in real-world settings, some delivery tasks may be more or less urgent (e.g. delivering food for lunch vs. delivering a package) and may require special expertise or access (e.g. picking up a book from a university library). Capturing these properties in a task model means that we may need to make some modifications to existing models, such as value function and a policy model. For example, a modified value function may need to compute the value of completing tasks by taking into consideration each task's urgency and deadlines. As there are more dimensions to consider in the value function, using heuristics to manually choose the best policy may become more ineffective and the benefits of our model-based, principled way of choosing policies will become larger.

5.1.4. Applicability to Other Domains

We demonstrated ways to realize the idea of flexible coordination to empower people and communities to achieve their goals with the help of intelligent systems that address the need for coordination and the need for flexibility. In workplaces, communities, and people's personal lives, we envision ways to use the idea of flexible coordination to design intelligent systems that transform how we get things done, help one another, and generally balance multiple goals (e.g., individual and collective goals; short-term and long-term goals) that demand our attention across our busy lives. This section articulates a few concrete directions.

Opportunistic Help-Seeking and Help-Giving in Workplaces. While this thesis developed ways to effectively advance collective goals through people's physical routines, in future work we may develop ways to advance help-seeking and collaboration in workplaces through people's work routines. Flexible coordination mechanisms for the workplace might help identify opportunistic moments for helping others and for collaborating on key tasks, while still ensuring that people can be productive and work towards their individual goals. We expect that many of the frameworks we developed in this thesis can translate and be used in the work setting, but where models of people's mobility patterns and routines are replaced with models of people's work schedules, routines, and needs. Once developed, these models can be used by individual-level coordination mechanisms to decide when to best recruit a person to seek or provide help (or to continue attending to their own tasks and goals), and by community-level coordination mechanisms to decide how to best orchestrate work and collaborations across a team or organization.

Distributed Sensemaking in the Community of Sensemakers. We may also extend the idea of flexible coordination to develop sensemaking platforms that help people reduce the cost of individual sensemaking and leverage their sensemaking efforts to help the community of future users. A core challenge in designing intelligent sensemaking support tools is to adapt to the dynamic and iterative nature of sensemaking processes [**113**]. Fully automated, pre-planned approaches that prescribe sensemaking actions to a user may fail because they encounter new information and mentally learn new concepts and relations in a dynamic manner. Fully manual, user-driven approaches that provide users the flexibility to explore the information space dynamically can be too time-consuming to navigate through the sensemaking process.

To overcome such shortcomings, we may use the idea of flexible coordination to develop a platform that follows the changing state of people's mental models of information space and provides guidance to move between stages based on the cost and benefits of current and future activities. For example, the system may suggest a user moving from an exploration stage to a comparison stage if the system knows when a user is not encountering new options. To do this, the system may encode the state of a user's mental models, model the likelihood of encountering new options and their effect on the mental models (e.g. using information-theoretic estimators such as Good-Turing [**103**]), and compute the expected value of gaining new information vs. the cost of conducting an additional search.

Personal Virtual Assistants to Balance Short-Term and Long-Term Goals. While this thesis focused on developing flexible coordination approaches for advancing and balancing individual and collective goals, future work may develop flexible coordination approaches for helping an individual advance and balance between their short-term and long-term goals (e.g., meeting an impending deadline and living a healthy life). Specifically, we may use the idea of flexible coordination to develop intelligent personal virtual assistants that can reason about various short-term and long-term tradeoffs in the decisions that a person might make, and that can surface such reasoning to users to support their everyday lives. For example, an intelligent personal assistant may consider the “privacy cost” of accidental information disclosure in notifications [**75**] by reasoning about the trade-offs between short-term benefits of immediate information access and long-term privacy cost (e.g., based on the content of the notification and the context in which the notification appeared).

5.2. A Mixed-Method Approach for Designing and Deploying Flexible Coordination Systems

While simulation studies are a common evaluation method in AI, they are less commonly used in CSCW and the design of social computing systems. In this section, we wish to highlight three core benefits of running simulations before deploying flexible coordination systems to actual users: (1) understanding boundaries and limitations, (2) identifying corner cases; and (3) understanding the implications of a given set of encoded values on outcomes.

First, system designers can better understand the capacity or capability of the service when being deployed in a community. For example, they may be able to understand how many delivery requests can be fulfilled or how fast they can be delivered in the best- or worst-case scenario. This understanding of boundaries may allow system designers to communicate the capabilities to users in order to set better expectations.

Second, system designers can identify some corner cases or failure cases before deploying the system to actual users. For example, system designers may find that, due to existing mobility patterns, the system inadvertently overburdens a certain group of volunteers or may not be able to provide services to requesters who live in certain areas. By identifying failure cases prior to the deployment, system designers may come up with solutions to resolve the issues or raise the potential issues to community members to devise potential solutions together.

Lastly, system designers can understand whether the system's decisions, based on the encoded values, will lead to outcomes that are in accord with the intended goals. For example, system designers can observe decisions systems may make based on the encoded values, possible unfoldings of people's actions based on the system decisions, potential outcomes as a result of people's actions, and whether or not these outcomes are in accord with the intended goals. By better understanding the implications of encoded values on potential outcomes, system designers can fine-tune the values in the value function in a way that can lead to desired outcomes. However, running simulations alone is not enough because people's behaviors still are uncertain. Instead, simulations should be accompanied by user studies to help system designers uncover unknown unknowns in user behaviors. For example, in Hit-or-Wait deployment, we found that people sometimes took unexpected routes or did not move to adjacent states (e.g. taking shortcuts or trespassing), in ways that were unexpected by our model. And, this caused the system to miss opportunities to notify people in other areas. This highlights how running user studies can help system designers uncover uncertain user behaviors that can happen in the open-world.

In summary, to effectively design and deploy flexible coordination systems we need to leverage both simulations as a pre-deployment tool and user studies as a discovery tool for unknown unknowns.

5.3. Ethics

When designing flexible coordination systems to help people achieve individual and collective goals, it is important to take into account ethical considerations and mitigate the risks of reinforcing inequality or bringing negative outcomes. Here, we discuss potential ethical concerns: fairness, potential negative uses, transparency, privacy, and access to power or authority.

5.3.1. Fairness

Flexible coordination systems may raise fairness issues around which groups of people might be able to benefit from services and who are being asked to contribute, both of which are largely dependent on people's mobility patterns. For example, people may not go to areas that they perceive as less safe, such as low socioeconomic status (SES) areas [**126**], while they may frequent other areas that they do not have any safety concerns (e.g. high SES areas). As a result, requesters in some areas may not be able to take advantage of the benefits of services due to the lack of potential volunteers while requesters in other areas may get high quality of services.

If people's inherent mobility patterns reinforce inequality, how might we improve flexible coordination systems that sought to surface opportunities within people's routines? One approach might be to identify people who are more willing to deviate from their routines and take a more directed approach to complete tasks in areas that are underserved.

5.3.2. Negative Uses

While the intended use case for flexible coordination systems is to allow people to better help each other in a community, the systems may be used for nefarious purposes. For example, people may go about their days and conveniently deliver packages to the next locations. However, these volunteers may be unknowingly contributing to the evil collective efforts, such as transporting drugs. As another example, while many people have a good intention to help other people find lost items within a community, few people may use the system to better identify areas that need search and retrieve the item without returning it to the owner. While it is outside of the scope of the current research, in the future, we need ways to mitigate these negative use cases when deploying actual on-the-go crowdsourcing systems.

5.3.3. Transparency

While in our deployment studies the systems did not explicitly communicate system decisions to users, future systems that leverage flexible coordination should be transparent about how and why the systems make certain decisions. Without the necessary transparency in system decisions, it can be difficult for people to decide whether or not to act on the presented opportunities, and it can be unclear to people how the opportunities may help achieve collective goals. As a first step toward this direction, in Hit-or-Wait's post-study interview, we visualized the reasoning behind Hit-or-Wait's decisions by highlighting how the system predicted people's likely routes and why the system decided to hit or wait based on previous search efforts across regions and a user's likely future routes. Our finding shows that communicating seemingly opaque system decisions to people made them value their contributions more. This is important because, by highlighting the uniqueness of people's contributions towards achieving collective goals, people may be more likely to continue to participate [**8, 108**].

5.3.4. Privacy

While flexible coordination provides clear benefits from encoding and modeling people's routines, there exist privacy concerns when collecting data for training models. To address some of the privacy concerns, we may collect data in a more privacy-preserving manner. For example, in our studies, we trained a population-based model for route prediction where we did not have to know which route data came from which users. While this may lead to less accurate models, our studies show that, even with less accurate models, our approaches can still steer users' actions towards the "right" direction to achieve desired collective goals. In addition, even with the less accurate model, systems like supply management may dynamically adjust their strategies when predicted outcomes differ from actual, realized outcomes. In the future, we may also compute the desired threshold of model accuracy for systems to be able to make good enough decisions to help people achieve their individual needs and collective goals. By understanding the necessary model accuracy for systems to achieve desired outcomes, system designers may investigate ways in which they can reach the desired threshold without compromising user privacy much (e.g. by combining various features that do not violate people's privacy but are still somewhat informative).

5.3.5. Who Controls the Value Function?

While our technical frameworks allow us to encode preferences and goals that volunteers may have, this can introduce unintended potential biases and raises questions about who should be in control of the value function for its use and effectiveness. First, we may elicit individual preferences such as which type of tasks and who volunteers may want to help, but optimizing for these preferences may introduce biases towards a certain group of people. Some volunteers may only want to help who they deem most need help; for example, elderly people who cannot go to grocery stores during the pandemic. However, optimizing for these preferences may disadvantage other groups of people who are also in need. For example, students who need to work part-time during the day while attending classes may not be able to go to the package center to pick up their packages within business hours. While the system may communicate who needs what kind of help and why, this also raises privacy concerns and whether or not the requesters want to reveal potentially sensitive information to receive help from the community. This suggests that encoding and optimizing the individual preferences alone may not be enough, and we need to carefully reason about what unintended consequences we may encounter.

Second, while we may elicit and encode individual preferences, it is unclear who should be in control of the value function to ensure the intended uses [**99**] and its effectiveness. For example, when many volunteers have different preferences and needs, it is not at all obvious how to encode all the preferences into a value function and who makes the final decision about what the configuration of the value function should look like. This becomes even more challenging when community members may not be able to reach consensus or system designers may not be able to fully grasp the "optimal" configuration for the community. Depending on the configuration of the value function, some groups of users may not be able to reach the goals they want. This suggests that we need ways to better understand the trade-offs between one configuration and the others, and how to make a final decision, either collectively or through an authority figure, to figure out what ends up in the value function and how they are weighed.

CHAPTER 6

Conclusion

This thesis introduces the idea of *flexible coordination* that follows people's routines and surface opportunities at moments when the interests of volunteers align with that of a system. Flexible coordination allows us to design volunteer-based physical crowdsourcing systems that provide flexibility to volunteers and coordinate useful contributions to achieve globally effective outcomes. This thesis has presented three frameworks that enable the idea of flexible coordination:

- **Opportunistic Hit-or-Wait:** a general decision-theoretic mechanism that intelligently controls decisions over when to notify a person of a task among many tasks that they can contribute to along their existing routes, in ways that reason both about system needs across tasks and about a helper's changing patterns of mobility.
- **4X:** a framework for multi-stage data collection processes that determine effective data collection strategies by reasoning about dynamically changing state of the world, people's changing interests and willingness in deviating off of their routine based on our knowledge of the world.
- **Opportunistic Supply Management:** a general decision-theoretic framework for modeling and optimizing the choice of task notification policies that meet the needs of helpers and system efficiency.

The fundamental idea of this thesis is to design intelligent systems to help people advance and balance their individual and collective goals without ever imposing what each individual must do. As a result, users have flexibility and autonomy to decide what they want to do and how they want to do it, while intelligent systems serve as a helpful assistant to monitor and surface opportunities to advance people's goals that would otherwise have been difficult for people to keep track of. Instead of making all or nothing trade-offs, flexible coordination systems help people simultaneously achieve both goals by following people's routines and identifying opportunities within their routines that are convenient for them and useful for achieving collective goals. This allows people to focus on their current goals and the task at hand while still contributing to other goals when opportunities arise. We believe that this seamless interaction and careful coordination will be a powerful model for designing intelligent systems to help people achieve multiple goals.

References

[1] SeeClickFix, 2019.

[2] Karl Aberer, Saket Sathe, Dipanjan Chakraborty, Alcherio Martinoli, Guillermo Barrenetxea, Boi Faltings, and Lothar Thiele. Opensense: Open community driven sensing of environment. In *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming*, IWGS '10, pages 39–42, New York, NY, USA, 2010. ACM.

[3] Elena Agapie, Jaime Teevan, and Andrés Monroy-Hernández. Crowdsourcing in the field: A case study using local crowds for event reporting. In *Third AAAI Conference on Human Computation and Crowdsourcing*, 2015.

[4] Florian Alt, Alireza Sahami Shirazi, Albrecht Schmidt, Urs Kramer, and Zahid Nawaz. Location-based crowdsourcing: extending crowdsourcing to the real world. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pages 13–22. ACM, 2010.

[5] Paul Aoki, Allison Woodruff, Baladitya Yellapragada, and Wesley Willett. Environmental protection and agency: Motivations, capacity, and goals in participatory sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 3138–3150, New York, NY, USA, 2017. ACM.

[6] Paul M. Aoki, R. J. Honicky, Alan Mainwaring, Chris Myers, Eric Paulos, Sushmita Subramanian, and Allison Woodruff. A vehicle for research: Using street sweepers to explore the landscape of environmental community action. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 375–384, New York, NY, USA, 2009. ACM.

[7] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K. Dey. Modeling and Understanding Human Routine Behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 248–260, New York, NY, USA, 2016. ACM. event-place: San Jose, California, USA.

[8] Gerard Beenen, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, and Robert E. Kraut. Using Social Psychology to Motivate Contributions to Online Communities. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 212–221, New York, NY, USA, 2004. ACM.

[9] Victoria ME Bellotti, Sara Cambridge, Karen Hoy, Patrick C Shih, Lisa Renery Handalian, Kyungsik Han, and John M Carroll. Towards community-centered support for peer-to-peer service exchange: rethinking the timebanking metaphor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2975–2984. ACM, 2014.

[10] Yochai Benkler, Aaron Shaw, and Benjamin Mako Hill. Peer production: A form of collective intelligence. *Handbook of collective intelligence*, 175, 2015.

[11] Michael S Bernstein, Joel Brandt, Robert C Miller, and David R Karger. Crowds in two seconds: Enabling realtime crowd-powered interfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 33–42. ACM, 2011.

[12] Michael S. Bernstein, Greg Little, Robert C. Miller, Björn Hartmann, Mark S. Ackerman, David R. Karger, David Crowell, and Katrina Panovich. Soylent: A word processor with a crowd inside. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 313–322, New York, NY, USA, 2010. ACM.

[13] Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, et al. Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 333–342. ACM, 2010.

[14] Erin Brady, Meredith Ringel Morris, and Jeffrey P Bigham. Gauging receptiveness to social microvolunteering. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1055–1064, 2015.

[15] Dirk Brockmann, Lars Hufnagel, and Theo Geisel. The scaling laws of human travel. *Nature*, 439(7075):462, 2006.

[16] Jeffrey A. Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B. Srivastava. Participatory sensing. In *Workshop on World-Sensor-Web (WSW '06): Mobile Device Centric Sensor Networks and Applications*, pages 117–134, 2006.

[17] Yung-Ju Chang, Gaurav Paruthi, Hsin-Ying Wu, Hsin-Yu Lin, and Mark W. Newman. An investigation of using mobile and situated crowdsourcing to collect annotated travel

activity data in real-word settings. *International Journal of Human-Computer Studies*, 102:81 – 102, 2017. Special Issue on Mobile and Situated Crowdsourcing.

[18] Cen Chen, Shih-Fen Cheng, Aldy Gunawan, Archan Misra, Koustuv Dasgupta, and Deepthi Chander. TRACCS: A Framework for Trajectory-Aware Coordinated Urban Crowd-Sourcing. In *HCOMP*, May 2014.

[19] Le Chen, Alan Mislove, and Christo Wilson. Peeking beneath the hood of uber. In *Proceedings of the 2015 Internet Measurement Conference*, pages 495–508. ACM, 2015.

[20] M Keith Chen and Michael Sheldon. Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform. In *Ec*, page 455, 2016.

[21] Yueyue Chen, Pin Lv, Deke Guo, Tongqing Zhou, and Ming Xu. Trajectory segment selection with limited budget in mobile crowd sensing. *Pervasive and Mobile Computing*, 40:123–138, 2017.

[22] Justin Cheng and Michael Bernstein. Catalyst: triggering collective action with thresholds. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1211–1221, 2014.

[23] Shih-Fen CHENG, CHEN CEN, Thivya KANDAPPU, Hoong Chuin LAU, Archan MISRA, Nikita JAIMAN, Randy Tandriansyah DARATAN, Ming Hui KOH, et al. Scalable urban mobile crowdsourcing: Handling uncertainty in worker movement. *ACM Transactions on Intelligent Systems and Technology*, 9(3):1, 2017.

[24] Chun-Wei Chiang, Anna Kasunic, and Saiph Savage. Crowd coach: Peer coaching for crowd workers' skill growth. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–17, 2018.

[25] Lydia B Chilton, Juho Kim, Paul André, Felicia Cordeiro, James A Landay, Daniel S Weld, Steven P Dow, Robert C Miller, and Haoqi Zhang. Frenzy: collaborative data organization for creating conference sessions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 1255–1264. ACM, 2014.

[26] James S Coleman. Social capital in the creation of human capital. *American journal of sociology*, 94:S95–S120, 1988.

[27] Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. Suggestbot: Using intelligent task routing to help people find work in wikipedia. In *Proceedings of the 12th International Conference on Intelligent User Interfaces*, IUI '07, pages 32–41, New York, NY, USA, 2007. ACM.

[28] Peng Dai, Mausam, and Daniel S. Weld. Decision-theoretic control of crowd-sourced workflows. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, AAAI'10, pages 1168–1174. AAAI Press, 2010.

[29] David Dearman, Timothy Sohn, and Khai N. Truong. Opportunities exist: Continuous discovery of places to perform activities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2429–2438, New York, NY, USA, 2011. ACM.

[30] Janis L. Dickinson, Benjamin Zuckerberg, and David N. Bonter. Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, 41(1):149–172, 2010.

[31] Mira Dontcheva, Robert R Morris, Joel R Brandt, and Elizabeth M Gerber. Combining crowdsourcing and learning to improve engagement and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3379–3388, 2014.

[32] Afsaneh Doryab, Victoria Bellotti, Alaaeddine Yousfi, Shuobi Wu, John M. Carroll, and Anind K. Dey. If it's convenient: Leveraging context in peer-to-peer variable service transaction recommendations. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3):48:1–48:28, September 2017.

[33] Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. Shepherding the crowd yields better work. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pages 1013–1022, 2012.

[34] Nathan Eagle. txteagle: Mobile crowdsourcing. *Internationalization, design and global development*, pages 447–456, 2009.

[35] Shane B. Eisenman, Emiliano Miluzzo, Nicholas D. Lane, Ronald A. Peterson, Gahng-Seop Ahn, and Andrew T. Campbell. Bikenet: A mobile sensing system for cyclist experience mapping. *ACM Trans. Sen. Netw.*, 6(1):6:1–6:39, Jan 2010.

[36] Nicole B Ellison, Charles Steinfield, and Cliff Lampe. The benefits of facebook âĂIJfriends:âĂİ social capital and college studentsâĂŹ use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

[37] Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. Aware: Mobile context instrumentation framework. *Frontiers in ICT*, 2:6, 2015.

[38] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. Investigating Episodes of Mobile Phone Activity As Indicators of Opportune Moments to Deliver Notifications. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pages 181–190, New York, NY, USA, 2011. ACM.

[39] James Fogarty, Scott E. Hudson, Christopher G. Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny C. Lee, and Jie Yang. Predicting Human Interruptibility with Sensors. *ACM Trans. Comput.-Hum. Interact.*, 12(1):119–146, March 2005.

[40] Eric J Friedman, Joseph Y Halpern, and Ian Kash. Efficiency and nash equilibria in a scrip system for p2p networks. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 140–149. ACM, 2006.

[41] Raghu K. Ganti, Nam Pham, Hossein Ahmadi, Saurabh Nangia, and Tarek F. Abdelzaher. Greengps: A participatory sensing fuel-efficient maps application. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 151–164, New York, NY, USA, 2010. ACM.

[42] Raghu K. Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *IEEE Communications Magazine*, 49(11):32–39, November 2011.

[43] Kapil Garg, Yongsung Kim, Darren Gergle, and Haoqi Zhang. 4x: A hybrid approach for scaffolding data collection and interest in low-effort participatory sensing. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), November 2019.

[44] Eric Gilbert and Karrie Karahalios. Predicting tie strength with social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 211–220. ACM, 2009.

[45] Cole Gleason, Dragan Ahmetovic, Saiph Savage, Carlos Toxtli, Carl Posthuma, Chieko Asakawa, Kris M Kitani, and Jeffrey P Bigham. Crowdsourcing the installation and maintenance of indoor localization infrastructure to support blind navigation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–25, 2018.

[46] Jorge Goncalves, Denzil Ferreira, Simo Hosio, Yong Liu, Jakob Rogstadius, Hannu Kukka, and Vassilis Kostakos. Crowdsourcing on the spot: altruistic use of public displays, feasibility, performance, and behaviours. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 753–762, 2013.

[47] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *nature*, 453(7196):779, 2008.

[48] Carlos Guestrin, Daphne Koller, Ronald Parr, and Shobha Venkataraman. Efficient solution algorithms for factored mdps. *Journal of Artificial Intelligence Research*, 19:399–468, 2003.

[49] Aakar Gupta, William Thies, Edward Cutrell, and Ravin Balakrishnan. mclerk: enabling mobile crowdsourcing in developing regions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1843–1852. ACM, 2012.

[50] Aaron Halfaker, Oliver Keyes, and Dario Taraborelli. Making peripheral participation legitimate: reader engagement experiments in wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 849–860. ACM, 2013.

[51] Kyungsik Han, Patrick C. Shih, Victoria Bellotti, and John M. Carroll. It's Time There Was an App for That Too:: A Usability Study of Mobile Timebanking. *International Journal of Mobile Human Computer Interaction*, 7(2):1–22, 2015.

[52] Kotaro Hara, Vicki Le, and Jon Froehlich. Combining crowdsourcing and google street view to identify street-level accessibility problems. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 631–640, 2013.

[53] Kurtis Heimerl, Brian Gawalt, Kuang Chen, Tapan Parikh, and Björn Hartmann. Communitysourcing: engaging local crowds to perform expert work via physical kiosks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1539–1548, 2012.

[54] Joyce Ho and Stephen S. Intille. Using Context-aware Computing to Reduce the Perceived Burden of Interruptions from Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 909–918, New York, NY, USA, 2005. ACM.

[55] Eric Horvitz. Reasoning under varying and uncertain resource constraints. In *AAAI*, volume 88, pages 111–116, 1988.

[56] Eric Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 159–166, New York, NY, USA, 1999. Association for Computing Machinery.

[57] Eric Horvitz, Paul Koch, Carl M Kadie, and Andy Jacobs. Coordinate: Probabilistic forecasting of presence and availability. In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, pages 224–233. Morgan Kaufmann Publishers Inc., 2002.

[58] Eric Horvitz, Paul Koch, Raman Sarin, Johnson Apacible, and Muru Subramani. Bayesphone: Precomputation of Context-Sensitive Policies for Inquiry and Action in Mobile Devices. In *User Modeling 2005*, pages 251–260. Springer, Berlin, Heidelberg, July 2005. DOI: 10.1007/11527886_33.

[59] Eric Horvitz, Paul Koch, and Muru Subramani. Mobile opportunistic planning: methods and models. In *User Modeling*, volume 4511, pages 228–237. Springer, 2007.

[60] Eric Horvitz and John Krumm. Some help on the way: Opportunistic routing under uncertainty. In *Proceedings of the 2012 ACM conference on Ubiquitous Computing*, pages 371–380. ACM, 2012.

[61] Eric J Horvitz, Gregory F Cooper, and David E Heckerman. Reflection and action under scarce resources: Theoretical principles and empirical study. In *IJCAI*, pages 1121–1127, 1989.

[62] Gary Hsieh, Ian Li, Anind Dey, Jodi Forlizzi, and Scott E. Hudson. Using visualizations to increase compliance in experience sampling. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, UbiComp '08, pages 164–167, New York, NY, USA, 2008. ACM.

[63] Kazushi Ikeda and Keiichiro Hoashi. Crowdsourcing go: Effect of worker situation on mobile crowdsourcing performance. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1142–1153. ACM, 2017.

[64] Shamsi T Iqbal and Eric Horvitz. Disruption and recovery of computing tasks: field study, analysis, and directions. In *CHI*, volume 7, pages 677–686, 2007.

[65] Shenggong Ji, Yu Zheng, and Tianrui Li. Urban sensing based on human mobility. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1040–1051. ACM, 2016.

[66] Hyunggu Jung, Victoria Bellotti, Afsaneh Doryab, Dean Leitersdorf, Jiawei Chen, Benjamin V Hanrahan, Sooyeon Lee, Dan Turner, Anind K Dey, and John M Carroll. 'masterful' matchmaking in service transactions: Inferred abilities, needs and interests versus activity histories. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1644–1655. ACM, 2016.

[67] Ece Kamar, Eric Horvitz, and Chris Meek. Mobile opportunistic commerce: mechanisms, architecture, and application. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 1087–1094. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[68] Thivya Kandappu, Nikita Jaiman, Randy Tandriansyah, Archan Misra, Shih-Fen Cheng, Cen Chen, Hoong Chuin Lau, Deepthi Chander, and Koustuv Dasgupta. Tasker: Behavioral insights via campus-based experimental mobile crowd-sourcing. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 392–402. ACM, 2016.

[69] Thivya Kandappu, Archan Misra, Shih-Fen Cheng, Nikita Jaiman, Randy Tandriansyah, Cen Chen, Hoong Chuin Lau, Deepthi Chander, and Koustuv Dasgupta. Campus-scale mobile crowd-tasking: Deployment and behavioral insights. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 800–812, New York, NY, USA, 2016. ACM.

[70] Salil S. Kanhere. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. In *2011 IEEE 12th International Conference on Mobile Data Management*, volume 2, pages 3–6, June 2011.

[71] Ashish Kapoor and Eric Horvitz. Experience sampling for building predictive user models: a comparative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 657–666, 2008.

[72] Ian A Kash, Eric J Friedman, and Joseph Y Halpern. Optimizing scrip systems: Efficiency, crashes, hoarders, and altruists. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 305–315. ACM, 2007.

[73] Ian A Kash, John K Lai, Haoqi Zhang, and Aviv Zohar. Economics of bittorrent communities. In *Proceedings of the 21st international conference on World Wide Web*, pages 221–230. ACM, 2012.

[74] Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. Sensr: Evaluating a flexible framework for authoring mobile data-collection tools for citizen science. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 1453–1462, New York, NY, USA, 2013. ACM.

[75] Yongsung Kim, Adam Fourney, and Ece Kamar. Studying preferences and concerns about information disclosure in email notifications. In *The World Wide Web Conference*, WWW '19, 2019.

[76] Yongsung Kim, Darren Gergle, and Haoqi Zhang. Hit-or-wait: Coordinating opportunistic low-effort contributions to achieve global outcomes in on-the-go crowdsourcing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 96:1–96:12, New York, NY, USA, 2018. ACM.

[77] Yongsung Kim, Emily Harburg, Shana Azria, Aaron Shaw, Elizabeth Gerber, Darren Gergle, and Haoqi Zhang. Studying the effects of task notification policies on participation and outcomes in on-the-go crowdsourcing. In *HCOMP*. AAAI, 2016.

[78] Stephen F. King and Paul Brown. Fix my street or else: Using the internet to voice local public service concerns. In *Proceedings of the 1st International Conference on Theory and Practice of Electronic Governance*, ICEGOV '07, pages 72–80, New York, NY, USA, 2007. ACM.

[79] Nikita Korolko, Dawn Woodard, Chiwei Yan, and Helin Zhu. Dynamic pricing and matching in ride-hailing platforms. *SSRN*, 2018.

[80] Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. *Building successful online communities: Evidence-based social design*. Mit Press, 2012.

[81] John Krumm. A markov model for driver turn prediction. In *SAE World Congress & Exhibition*. SAE International, apr 2008.

[82] John Krumm and Eric Horvitz. Predestination: Inferring destinations from partial trajectories. pages 243–260. Springer, 2006.

[83] Benjamin Lafreniere, Tovi Grossman, Fraser Anderson, Justin Matejka, Heather Kerrick, Danil Nagy, Lauren Vasey, Evan Atherton, Nicholas Beirne, Marcelo H Coelho, et al. Crowdsourced fabrication. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 15–28, 2016.

[84] Airi Lampinen and Barry Brown. Market design for hci: Successes and failures of peer-to-peer exchange platforms. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 4331–4343. ACM, 2017.

[85] Nicholas D. Lane, Shane B. Eisenman, Mirco Musolesi, Emiliano Miluzzo, and Andrew T. Campbell. Urban sensing systems: Opportunistic or participatory? In *Proceedings of the 9th Workshop on Mobile Computing Systems and Applications*, HotMobile '08, pages 11–16, New York, NY, USA, 2008. ACM.

[86] Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell. A survey of mobile phone sensing. *IEEE Communications Magazine*, 48(9):140–150, Sep. 2010.

[87] Gierad Laput, Walter S. Lasecki, Jason Wiese, Robert Xiao, Jeffrey P. Bigham, and Chris Harrison. Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds.

In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1935–1944, New York, NY, USA, 2015. ACM.

[88] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Ritesh Noothigattu, Daniel See, Siheon Lee, Christos-Alexandros Psomas, et al. Webuildai: Participatory framework for fair and efficient algorithmic governance. (CSCW), 2019.

[89] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1603–1612. ACM, 2015.

[90] Minne Li, Zhiwei Qin, Yan Jiao, Yaodong Yang, Jun Wang, Chenxi Wang, Guobin Wu, and Jieping Ye. Efficient ridesharing order dispatching with mean field multi-agent reinforcement learning. In *The World Wide Web Conference*, pages 983–994. ACM, 2019.

[91] Lin Liao, Donald J. Patterson, Dieter Fox, and Henry Kautz. Learning and inferring transportation routines. *Artificial Intelligence*, 171(5):311–331, April 2007.

[92] Christopher H Lin, Mausam Mausam, and Daniel S Weld. Dynamically switching between synergistic workflows for crowdsourcing. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[93] Katherine Lin, Henry Spindell, Scott Cambo, Yongsung Kim, and Haoqi Zhang. Habitsourcing: Sensing the environment through immersive, habit-building experiences. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 639–650. ACM, 2016.

[94] Thomas Ludwig, Julian Dax, Volkmar Pipek, and Dave Randall. Work or leisure? designing a user-centered approach for researching activity "in the wild". *Personal Ubiquitous Comput.*, 20(4):487–515, Aug 2016.

[95] Thomas W Malone and Kevin Crowston. The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)*, 26(1):87–119, 1994.

[96] Rosario N Mantegna and H Eugene Stanley. Stochastic process with ultraslow convergence to a gaussian: the truncated lévy flight. *Physical Review Letters*, 73(22):2946, 1994.

[97] David W McDonald and Mark S Ackerman. Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 231–240. ACM, 2000.

[98] James McInerney, Alex Rogers, and Nicholas R. Jennings. Learning periodic human behaviour models from sparse data for crowdsourcing aid delivery in developing countries. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, UAI'13, pages 401–410, Arlington, Virginia, United States, 2013. AUAI Press.

[99] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 220âĂŞ229, New York, NY, USA, 2019. Association for Computing Machinery.

[100] Patrenahalli M. Narendra and Keinosuke Fukunaga. A branch and bound algorithm for feature subset selection. *IEEE Transactions on computers*, 9(C-26):917–922, 1977.

[101] Jeffrey Nichols and Jeon-Hyung Kang. Asking questions of targeted strangers on social networks. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 999–1002, New York, NY, USA, 2012. ACM.

[102] Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. Platemate: crowdsourcing nutritional analysis from food photographs. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 1–12. ACM, 2011.

[103] Alon Orlitsky and Ananda Theertha Suresh. Competitive distribution estimation: Why is good-turing good. In *Advances in Neural Information Processing Systems*, pages 2143–2151, 2015.

[104] Sangkeun Park, Sujin Kwon, and Uichin Lee. Campuswatch: Exploring communitysourced patrolling with pervasive mobile technology. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):134:1–134:25, Nov 2018.

[105] Donald J. Patterson, Lin Liao, Dieter Fox, and Henry Kautz. Inferring High-Level Behavior from Low-Level Sensors. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, Anind K. Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *UbiComp 2003: Ubiquitous Computing*, volume 2864, pages 73–89. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[106] Lisa Posch, Arnim Bleier, Clemens M Lechner, Daniel Danner, Fabian Flöck, and Markus Strohmaier. Measuring motivations of crowdworkers: The multidimensional crowdworker motivation scale. *ACM Transactions on Social Computing*, 2(2):1–34, 2019.

[107] Reid Priedhorsky, Mikhil Masli, and Loren Terveen. Eliciting and focusing geographic volunteer work. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, pages 61–70, New York, NY, USA, 2010. ACM.

[108] Al M. Rashid, Kimberly Ling, Regina D. Tassone, Paul Resnick, Robert Kraut, and John Riedl. Motivating Participation by Displaying the Value of Contribution. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 955–958, New York, NY, USA, 2006. ACM.

[109] Yuqing Ren, F Maxwell Harper, Sara Drenner, Loren Terveen, Sara Kiesler, John Riedl, and Robert E Kraut. Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *Mis Quarterly*, pages 841–864, 2012.

[110] Mark Owen Riedl and Vadim Bulitko. Interactive Narrative: An Intelligent Systems Approach. *AI Magazine*, 34(1):67–67, 2013.

[111] Xin Rong, Adam Fourney, Robin N. Brewer, Meredith Ringel Morris, and Paul N. Bennett. Managing uncertainty in time expressions for virtual assistants. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 568–579, New York, NY, USA, 2017. ACM.

[112] Stephanie Rosenthal, Anind K Dey, and Manuela Veloso. Using decision-theoretic experience sampling to build personalized mobile phone interruption models. In *International Conference on Pervasive Computing*, pages 170–187. Springer, 2011.

[113] Daniel M Russell, Mark J Stefik, Peter Pirolli, and Stuart K Card. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 269–276, 1993.

[114] Adam Sadilek, John Krumm, and Eric Horvitz. Crowdphysics: Planned and Opportunistic Crowdsourcing for Physical Tasks. In *Seventh International AAAI Conference on Weblogs and Social Media*, June 2013.

[115] Manaswi Saha, Michael Saugstad, Hanuma Teja Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, and Jon Froehlich. Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 62:1–62:14, New York, NY, USA, 2019. ACM.

[116] Susumu Saito Carlos Toxtli Saiph Savage, Chun Chiang and Jeffrey P. Bigham. Becoming the super turker: Increasing wages via a strategy from high earning workers. In *Proceedings of the World Wide Web Conference (WebConf 2020)*, 2020.

[117] Daisuke Sato, Uran Oh, João Guerreiro, Dragan Ahmetovic, Kakuya Naito, Hironobu Takagi, Kris M Kitani, and Chieko Asakawa. Navcog3 in the wild: Large-scale blind indoor navigation assistant with semantic features. *ACM Transactions on Accessible Computing (TACCESS)*, 12(3):1–30, 2019.

[118] Saiph Savage, Andres Monroy-Hernandez, and Tobias Höllerer. Botivist: Calling volunteers to action using online bots. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 813–822, 2016.

[119] Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4(1):1–32, 2008. PMID: 18509902.

[120] Brian L. Sullivan, Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282 – 2292, 2009.

[121] Shuai Sun, Yan Li, Wayne ST Rowe, Xuezhi Wang, Allison Kealy, and Bill Moran. Practical evaluation of a crowdsourcing indoor localization system using hidden markov models. *IEEE Sensors Journal*, 19(20):9332–9340, 2019.

[122] Ryo Suzuki, Niloufar Salehi, Michelle S Lam, Juan C Marroquin, and Michael S Bernstein. Atelier: Repurposing expert crowdsourcing tasks as micro-internships. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2645–2656, 2016.

[123] H. Tangmunarunkit, C. K. Hsieh, B. Longstaff, S. Nolen, J. Jenkins, C. Ketcham, J. Selsky, F. Alquaddoomi, D. George, J. Kang, Z. Khalapyan, J. Ooms, N. Ramanathan, and D. Estrin. Ohmage: A general and extensible end-to-end participatory sensing platform. *ACM Trans. Intell. Syst. Technol.*, 6(3):38:1–38:21, Apr 2015.

[124] Rannie Teodoro, Pinar Ozturk, Mor Naaman, Winter Mason, and Janne Lindqvist. The motivations and experiences of the on-demand mobile workforce. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 236–247. ACM, 2014.

[125] Loren Terveen and David W McDonald. Social matching: A framework and research agenda. *ACM transactions on computer-human interaction (TOCHI)*, 12(3):401–434, 2005.

[126] Jacob Thebault-Spieker, Loren G. Terveen, and Brent Hecht. Avoiding the South Side and the Suburbs: The Geography of Mobile Crowdsourcing Markets. In *CSCW*, pages 265–275, New York, NY, USA, 2015. ACM.

[127] Anthony Tomasic, John Zimmerman, Aaron Steinfeld, and Yun Huang. Motivating contribution in a participatory sensing system via quid-pro-quo. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '14, pages 979–988, New York, NY, USA, 2014. ACM.

[128] Khai N Truong, Thariq Shihipar, and Daniel J Wigdor. Slide to x: unlocking the potential of smartphone unlocking. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3635–3644. ACM, 2014.

[129] Kathleen Tuite, Noah Snavely, Dun-yu Hsiao, Nadine Tabing, and Zoran Popovic. Photocity: Training experts at large-scale image acquisition through a competitive game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1383–1392, New York, NY, USA, 2011. ACM.

[130] Heli Väätäjä and Paul Egglestone. Briefing news reporting with mobile assignments: perceptions, needs and challenges. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 485–494. ACM, 2012.

[131] Rajan Vaish, Keith Wyngarden, Jingshu Chen, Brandon Cheung, and Michael S Bernstein. Twitch crowdsourcing: crowd contributions in short bursts of time. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3645–3654. ACM, 2014.

[132] Luis Von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326. ACM, 2004.

[133] Luis Von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum. recaptcha: Human-based character recognition via web security measures. *Science*, 321(5895):1465–1468, 2008.

[134] Jiangtao Wang, Yasha Wang, Daqing Zhang, Feng Wang, Yuanduo He, and Liantao Ma. Psallocator: multi-task allocation for participatory sensing with sensing capability constraints. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1139–1151. ACM, 2017.

[135] Ryen W White, Matthew Richardson, and Yandong Liu. Effects of community size and contact rate in synchronous social q&a. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2837–2846. ACM, 2011.

[136] Zhe Xu, Zhixin Li, Qingwen Guan, Dingshui Zhang, Qiang Li, Junxiao Nan, Chunyang Liu, Wei Bian, and Jieping Ye. Large-scale order dispatch in on-demand ride-hailing

platforms: A learning and planning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 905–913. ACM, 2018.

[137] Ming Yin, Siddharth Suri, and Mary L Gray. Running out of time: The impact and value of flexibility in on-demand crowdwork. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 430. ACM, 2018.

[138] Haoqi Zhang, Eric Horvitz, and David C Parkes. Automated workflow synthesis. In *AAAI*, 2013.

[139] Haoqi Zhang, Edith Law, Rob Miller, Krzysztof Gajos, David Parkes, and Eric Horvitz. Human computation tasks with global constraints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 217–226. ACM, 2012.

[140] Haiyi Zhu, Robert Kraut, and Aniket Kittur. Organizing without formal organization: Group identification, goal setting and social modeling in directing online production. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 935–944, New York, NY, USA, 2012. ACM.

[141] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. Value-sensitive algorithm design: method, case study, and lessons. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):194, 2018.

[142] Brian D Ziebart, Andrew L Maas, Anind K Dey, and J Andrew Bagnell. Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 322–331. ACM, 2008.

[143] John Zimmerman, Anthony Tomasic, Charles Garrod, Daisy Yoo, Chaya Hiruncharoenvate, Rafae Aziz, Nikhil Ravi Thiruvengadam, Yun Huang, and Aaron Steinfeld. Field trial of tiramisu: Crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1677–1686, New York, NY, USA, 2011. ACM.