Large-scale crowdsourcing systems such as SeeClickFix, FixMyCity, and FixMyStreet use everyday citizens to sense the physical environment, by asking them to report potholes, graffiti and non-functioning street lights in the community [8, 9, 10]. The data benefits researchers and city planners who cannot otherwise collect it without resorting to expensive surveys or machine sensing systems. While many similar systems exist for measuring physical space [1, 6, 7], all exhibit a tradeoff between user participation, data coverage, and data fidelity. This tradeoff happens because the such systems depend entirely on users contributing opportunistically (user chooses when they wish to contribute data) and do not allow for users to be directed to contribute the data which could be used to fill gaps in coverage. This could be remedied slightly by using a group of dedicated volunteers who contribute more regularly and will fill some gaps in coverage. However, the size of this group is significantly smaller than all those who could potentially contribute to the sensing effort, reducing the scale of the system. Furthermore, the interaction itself (taking out a phone and filling a survey) is not low-effort, and would therefore only be completed by individuals who are very dedicated to the sensing goal, reducing the pool of able contributors even further. This, therefore, poses two questions: (1) is it possible to encourage on-the-go crowds (individuals on a set route) to contribute lower-fidelity data points that, with scaffolding, yield high-fidelity data while ensuring complete data coverage (e.g. the description of a building facades from data points on architectural style, number of windows/doors, etc.) and (2) will the use of low-effort contributions (ones that do not require completion of a time consuming form) increase user participation in sensing tasks.

Current crowdsourcing systems exhibit low user participation due to their time consuming and opportunistic nature, which leads to poor data coverage of the physical space. These systems interact with users by either automatically measuring data about the environment though machine sensors like GPS [4, 5], or by asking the user to complete a form or report [8, 9, 10]. Machine sensors allow users to contribute data without any required input. These sensors, however, have limited applicability since they cannot measure attributes that require human interpretation or are subjective, such as identifying interesting architecture. Additionally, they rely on the user's existing mobility and are not guaranteed to achieve complete data coverage. The former is easily solved by systems that rely on user input, but current implementations, such as FixMyStreet, require users to use their smartphone to complete a time intensive form [10]. Such systems do not gain high user participation and are often only used by those who are very interested in the sensing goal the system is attempting to measure due to their time requirement. Moreover, all existing systems rely purely on opportunistic contribution for data collection. Opportunistic systems, such as the applications Waze and Foursquare, allow users to contribute information on optimal travel routes and places of interest when it is convenient for them [2, 3]. However, these systems have no means of compelling users to execute specific tasks, leading to a lack of coverage.

My project will focus on researching a new form of interaction technique between a researcher and on-the-go crowds which will (1) guarantee high data coverage, (2) fill data gaps, and (3) allow for the scaffolding of small pieces of low-fidelity data into higher-fidelity data to meet sensing goals proposed by said researcher. Most importantly, the previous goals will be met without needing crowds to significantly deviate from their existing route or requiring them to fill time-consuming surveys on their smartphones to contribute data, thus increasing user participation. From this, I specifically seek to answer the following research questions:

- RQ 1. Does the use of wearable devices or contextual touch, like the Apple Watch or 3D Touch on iPhone 6S/6S+, facilitate and encourage low-effort contributions for on-the-go crowds?
- RQ 2. Can on-the-go users be encouraged to contribute data by incentivizing them with information about opportunities of interest to them?

RQ 3. How might I direct individual on-the-go users to contribute specific, small pieces of information that can be scaffolded into higher fidelity information while ensuring that data coverage goals are met?

To evaluate the proposed forms of low-effort input in RQ1, I will create a prototype that allows users to (1) mark locations for tracking, (2) contribute information on the locations, (3) be notified when a location is nearby, and (4) be queried for more information on a nearby location. I will develop the application using an Apple Watch that can easily be interacted with while a user is on-the-go. Additionally, I will develop a second version to utilize 3D Touch on the iPhone 6S/6S+ that will afford users who do not own an Apple Watch a low-effort interaction with the application. Using these prototypes, I will conduct a small, 3-5 user usability study in the Delta Lab space with a test scenario of finding free food on campus to learn how potential users interact with the application and gauge the degree of effort required (see Appendix A for interview guide). The feedback received will be used to improve the interaction by minimizing effort needed to contribute data. This process will be completed within 2 weeks.

Following usability testing, I will brainstorm a testing scenario for application use. An example scenario would be a system which asks users to tag bridges and architecture they find interesting. This subjective task cannot be accomplished by machine sensors and affords follow-up questions, such as asking for the weight limit of the bridge or number of windows on a building's facade. Furthermore, a complete description of the location would require many users to contribute small pieces of low-fidelity data that could be scaffolded into higher-fidelity data. This will lead to the development of the scaffolding algorithm: an algorithm that not only can combine pieces of data, but will intelligently query and direct users for information to fill any gaps. The scenario and algorithm will be implemented into the prototype built earlier, along with any feedback gained from usability testing, and will take 1-2 weeks.

To then test how users can be encouraged to contribute data (RQ2) and directed to fill data gaps (RQ3) using the scaffolding algorithm above, I will recruit 15-20 undergraduate students through multiple student groups on campus from a variety of schools to conduct a small scale study. Recruitment will be done by flyering various parts of campus and using school-wide groups on social media to advertise the study. This will help reduce any potential bias that might exist in specific schools (computer science students more likely to use application, for example). I expect this process to take 2-3 weeks. Once users are recruited, I will conduct a 1-month long study to evaluate both the low-effort interaction at scale and the scaffolding algorithm. The study will begin towards the end of Winter Quarter and extend into Spring Quarter. Through this study, I will assess the quality of data scaffolding and data coverage obtained from using the scaffolding algorithm.

At the conclusion of the study, I will conduct in-person interviews to gather qualitative feedback on the interaction and usage which will inform me about how likely an individual would be to use the system on a day-to-day basis. The in-person interviews will be an open-ended conversation about when a user would/would not want to contribute data, in what circumstances would a user be willing to deviate from their route to contribute data, how the contribution experience could be improved (see Appendix B for example questions). From this, I want to learn how users might be better coordinated or incentivized to help fill data gaps and contribute additional data to a prior contribution. These findings, along with the results from above, will be summarized into a conference paper.

I have developed applications for iOS previously and have learned how to develop for the Apple Watch through Apple's tutorials. For usability and interaction testing, I will utilize skills from EECS 330: Human-Computer Interaction, which I am currently a TA for. If the need for artificial intelligence techniques in the scaffolding algorithms arises, I will rely on skills learned from EECS 348: Intro. to Artificial Intelligence and EECS 349: Machine Learning.

References

1. 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.
2. Waze. https://www.waze.com/, 2016.
3. Foursquare. https://foursquare.com/, 2016.
4. Eiman Kanjo. Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping. Mobile Networks and Applications, 15(4):562–574, 2010.
5. Susan E. Matthews. How google tracks traffic. http://www.theconnectivist.com/2013/07/how-google- tracks-traffic/, 2013.
6. Paul M Aoki, RJ Honicky, Alan Mainwaring, Chris Myers, Eric Paulos, Sushmita Subramanian, and Allison Woodruff. Common sense: Mobile environmental sensing platforms to support community action and citizen science. In Ubicomp, 2008.
7. Celia Evans, Eleanor Abrams, Robert Reitsma, Karin Roux, Laura Salmonsen, and Peter P Marra. The neighborhood nestwatch program: Participant outcomes of a citizen-science ecological research project. Conservation Biology, 19(3):589–594, 2005.
8. SeeClickFix http://en.seeclickfix.com/, 2016.
9. FixMyCity. http://www.fixmycity.de/, 2016.
10. FixMyStreet. https://www.fixmystreet.com/, 2016.