UCLA
Technical Reports

Title

Participatory Sensing for Community Data Campaigns: A case study

Permalink

https://escholarship.org/uc/item/95t603tj

Authors

Acker, Amelia
Lukac, Martin
Estrin, Deborah

Publication Date

2010-11-30

Participatory Sensing for Community Data Campaigns: A case study

November 2010

Amelia Acker
UCLA Information Studies
BOX 951520, 207 GSEIS Bldg
Los Angeles, CA 90095-1520
aacker@ucla.edu

Martin Lukac
Nexleaf Analytics
2356 Pelham Ave
Los Angeles, CA 90064-2212
martin@nexleaf.org

Deborah Estrin
Center for Embedded Network Sensing
UCLA 3563 Boelter Hall
Los Angeles, CA 90095-1596
destrin@cs.ucla.edu

ABSTRACT:

Participatory Sensing is a process whereby individuals and communities use mobile phones and web services to observe, analyze, and present personal and environmental artifacts, events and experiences. In this technical report we describe a community data campaign that made use of smartphone based participatory sensing for environmental needs assessment. Community organizers defined the content of the participatory sensing campaign. 68 individuals participated over the course of 6 weeks, uploading over 450 mini-surveys, including over 700 images.

KEYWORDS: *participatory sensing, community-data campaigns, participatory research, mobile data collection*

ACM CLASSIFICATION KEYWORDS:
mobile applications, social media

1. PARTICIPATORY SENSING

Participatory sensing offers communities and individuals a rapid, cost-efficient method for making robust observations through directed data campaigns. By harnessing common smart phone utilities such as internet access, camera, and GPS, communities can incorporate mobile sensing applications (apps) into everyday life activities—participants can self-report observations while going to school, work, home, or on errands. Campaign organizers can then access that data in raw and analyzed form in almost real-time, analyze collected observations, and use the data to gauge and determine specific needs within their neighborhood environments. Information gathered from community data campaigns can then be used as case making evidence to shape priorities of community organizations, local public health institutions, and public policy initiatives.

At a civic level, the goals of participatory sensing with mobile phones are: to facilitate engagement within a community by enabling participants to document their neighborhood's strengths and needs; and to encourage the distribution and decentralization of collected data so that participants can access, analyze, and make observations about the data. Web-based data management, mapping, visualization, and social networking provide important complements to these mobile technologies. Through web-based assets community based organizations and their members can themselves configure, aggregate, compare, and interpret data obtained through systematic *community-data campaigns.* Putting sensors into people's hands can enable participants to understand the significance of their individual data sensing contributions at personal and community levels. Finally, this approach provides communities with a unique avenue of engagement with the ordinary and everyday—where individuals can act critically and conscientiously on the data submitted and collected in a campaign to reach policy makers and make changes in local and personal environments.

This article describes a community data campaign that took place during the Spring of 2010 in the East Los Angeles neighborhood of Boyle Heights. We were approached by a collaborative of community based organizations (CBOs) to provide Boyle Heights community members with CENS developed applications on GPS-enabled mobile phones. Community organizers recruited local residents to carry phones with them as they went about their day for either 1 or 2 days and to use the CENS developed application to respond to 5 surveys that collected information about their neighborhood, daily route, work, home and school environments. We discuss the development and deployment of this community data campaign pilot. We detail the participation in the data campaign in Boyle Heights and then outline the advantages, disadvantages, and implications for practice in future projects involving participatory sensing in community based participatory research and needs assessment. Because the data collected belongs to the community members and organizers we do not present the content of the data uploaded, nor draw conclusions from that data regarding participant behavior or experiences.

We first discuss two areas of related work t in Section 2. In Section 3, we describe the design process, workflow and implementation of the data campaign. We discuss the findings of the data campaign and our fieldwork in Section 4. Section 5 concludes with implications for future work.

2. RELATED WORK

We identified two primary areas of related work: participatory research and experience sampling methods, and mobile data collection.

2.1 Participatory Research and Experience Sampling Methods

Participatory sensing is part of a trajectory of *participatory research* in the social sciences. Participatory research is defined by positioning research participants as co-investigators in partnership with more 'traditional' data collection experts, such as university researchers [1]. Historically, participatory research has been primarily used in qualitative and survey-based social science projects. But increasingly the health sciences, urban and environmental planners have turned towards community based participatory research to gather data by the subjects most affected by research outcomes [1-3].

Participatory research projects involve community residents throughout the entire research cycle [2]. Participatory sensing offers many advantages to traditional community based participatory research, such as needs assessment and biosocial population surveys.

NGOs and the social sciences have a history of using mobile recording devices, such as cameras, PDAs, and phones (using SMS and MMS) for community needs assessment [2-8]. Many participant-driven needs studies have used cameras to capture data [4], [9-11]. For example, in the late 1990s Wang et al used documentary photography and photo-elicitation interviews to create Photovoice for participatory needs assessment in rural villages in Yunnan, China. Village women workers were able to use instamatic cameras to document their community experiences as workers, mothers, and community members. The data collected from this photovoice project, included participant interviews for embedded context, documented the community's needs and assets, and participants were able to identify specific concerns and promote health objectives alongside public health officials and NGO workers.

Experience Sampling Methods (ESM) include self-report surveys, diary studies as well as Ecological Momentary Assessment (EMA) [12]. Studies that use ESM place the responsibility of data capture upon the participants *in situ*, and because the data is captured in the moment, there is little need for participants to try and recall events. This quality of self-report (combined with minimal contact with the researchers) can raise the ecological validity of *in situ* observational studies.

In the late 1990s there was an increase in PDA-based ESM studies that ran on PalmOS (e.g. ArcSight). Since then, many ESM projects have made use of voice-mail, SMS and MMS diary studies [9], [11], [13], [14]. Steadily, different utilities embedded in mobile devices have been harnessed in ESM studies. For example, in clinical studies in psychology, many researchers use the diary study method using electronic mobile devices that could capture audio, photographs and text [6], [15], [16].

Research in psychology and HCI has shown that in the last decade people have become increasingly comfortable using electronic mobile devices for self-report [14], [15]. Piasecki, et al have shown that using mobile devices such as mobile phones for EMA can confront the most common limitations of self-report with paper diaries (backlogging and noncompliance) and raise the ecological validity of the participant's experience because mobile diary surveys follow along the participant's "stream of daily experience" [15]. Similar to Photovoice and photo-elicitation studies, EMA with mobile phones allows participants to answer with photo, text and GPS coordinates. It gives participants the opportunity to capture everyday experiences in an unobtrusive way; because of participants' familiarity with phones, the system is easy to learn and to collect data.

2.2 Mobile Technology for Data Collection

The use of non-voice data applications with mobile phones is increasingly popular in the United States. The latest Pew mobile internet access report has shown that the use of mobile data applications has significantly grown in the last year—including taking photographs, accessing the internet, and recording video [17].

Many participant-driven mobile data collection systems to capture user feedback and context have been successfully implemented in many projects in the last decade, including Open Data Kit (ODK) and MyExperience. MyExperience enables studies of user feedback and evaluation of mobile technology [18]. The Open Data Kit tool suite has been applied to a range of uses in developing countries—from community health worker management, to post-conflict population surveys, to home-based HIV testing and counseling in developing regions of the world [8]. These systems, and others like them, can be used to support the sort of community data campaigns described here. As extensible data collection systems become interoperable across mobile platforms, participant-driven data campaigns will become increasingly easy to implement enabling broader use of experience sampling methods on community member's personal mobile phones. In this paper we focus on our experience using such systems, and on the implications for future use and design.

3. CAMPAIGN DESIGN PROCESS

This section discusses the design and implementation process of the Boyle Heights community data campaign.

In the initial stages of the collaboration, we (UCLA/CENS) created a working document that described the smart phone data collection applications that would be used to support the community data campaign. Representatives from the Boyle Heights Planning for Place (BHPP) committee defined the content of all data collection and survey application questions. After 1-2 rounds of feedback with committee representatives and group-wide discussion between UCLA and the Boyle Heights CBOs, we converged on a general description of all 5 data collection apps and a detailed definition of the first application, known as BH-Home which was a survey that featured needs assessment of the home environment (food preparation and consumption, personal routine, and home repair). After an initial round of local testing the survey was loaded onto an Android phone and delivered it to the committee to test. We received a number of small suggestions for improved understandability of question wording, ordering, and presentation, and designed a

second version that offered bilingual, Spanish or English, options. After completing this first app, we had further detailed discussions about the specific questions and wording for the remaining four applications.

The additional four applications also documented personal/community environments: BH-Route traced the participants commute for up to 3 hours and asked questions about transportation (e.g. walking, bus, private car); BH-School that asked about school conditions; BH-Work documented work conditions; and BH-Neighborhood assesses the participants feelings towards the surrounding neighborhood. Shortly thereafter the remaining applications were implemented and made available to the BHPP committee for testing.

After a second round of participant feedback from several recruited community members, the final applications were completed. We separately began discussing a basic website interface for the BHPP organizers to access the collected data, and how the phones would be shared through the community.

After three rounds of testing, the deployment began by training a group of CBO representatives. These representatives then trained the community participants involved in the campaign how to use the phones and upload surveys. The phones were retrieved and delivered by trainers week-by-week to new participants. The initial distribution, retrieval, and deployment was carried out solely by community organizers involved in the data campaign. We monitored the state of uploads and the deployment and on one occasion went to help with resetting some of the phones.
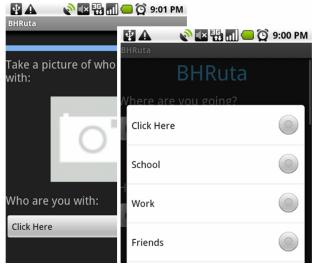

Figure 1. Screen shots from BH-Route

At the completion of the deployment we discussed with CBO representatives the form of the summary data statistics that would be most useful to the groups planning analysis efforts and made those available over shortly thereafter.

Four stages comprised the data campaign: planning the system (designing surveys, system, feedback); building the system (architecture, updating the phones); pretesting (iterative, more feedback); deployment, analysis, and dissemination (circulation of phones, collection and analysis of data).

The 5 phone apps all used the same code base. The questions, multiple choice responses, layout, and upload parameters were automatically loaded at compile time from configuration files.

For the tracking deployment, we created a data collection website which provided a variety of views of the data. The main display showed a grid containing the number of surveys for each day. Selecting a day would bring up all the survey entries for that day, including the pictures and a static map showing the location. The website also provided a map view of the data for each application for the deployment.
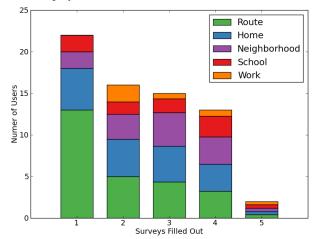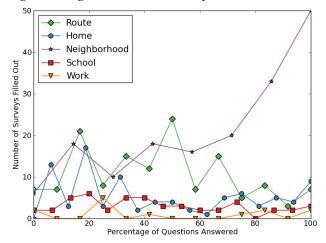

Figure 2. Usage levels across all surveys.

Figure 3. Completeness per survey across apps.

4. PARTICIPATION DATA

This section discusses three levels of findings. There were a totally of 68 participants who filled out 462 surveys total.

4.1 Survey Completeness

Figure 2 features total survey uploads across all users, including users who uploaded multiple surveys across unique apps. The route app was the most popular survey across all five apps with a 75% participation rate (51 users), whereas the work app had the lowest participation rate of 16% (11 users). The average user uploaded 6.85 surveys (e.g. 3 route uploads, 2 home uploads, 1

neighborhood uploads, 1 school upload); the average user uploaded 2.4 unique surveys (e.g. path, home, school).

Each survey had a different number of questions; users had a choice to partially complete surveys. Figure 3 shows the percentage of questions filled out per survey submitted. The neighborhood app was the most completed survey. The other four apps were maximally filled out 7.2% of the time. This is expected for a survey with many questions, however it is not possible to say whether this was due to the length of the survey or the content of the survey. Moreover, participants were asked to fill out as many questions as they wanted to; if more complete responses were requested, the numbers might have been different.
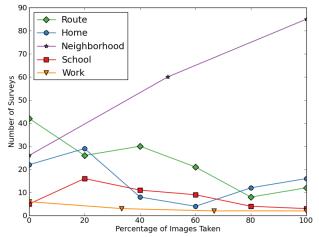

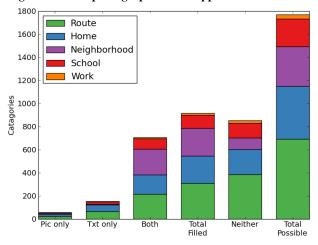**Figure 4. Use of photographs across apps.**

Figure 5. Were pictures labeled consistently?

4.2 Use of photographs across surveys

Each app asked participants to submit multiple photographs as part of survey completion. 60 participants took at least 1 and 11.25 pictures on average per user. 8 participants did not take any photographs. There were 765 images out of 1771 possible: representing 43.20% maximum contribution of images. Figure 4 is comprised of the total images submitted as part of survey uploads across all apps.

Each time a picture could be taken in a survey, there was the opportunity to label it either with a drop down or with free form

text. Figure 5 plot shows how many times just the picture was taken, just the question was answered, or both.

4.3 Time and Location

The phones were passed between CBOs at the beginning of the week and handed out on Tuesday or Wednesday. Some CBOs left the phones with an individual for the entire week while others had the phones passed between participants. Figure 6 shows that the bulk of survey submissions occurred on first few days that the participants carried the phones; survey submissions tapered off at the end of the loaning period as expected since they were instructed to use the applications for a period of only 1-2 days.
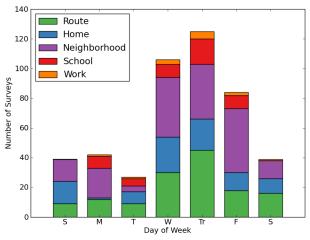

Figure 6. Survey submission by day of the week, broken down by app.
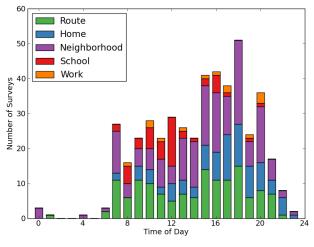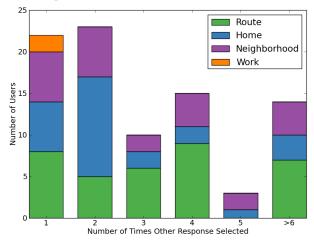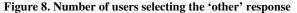

Figure 7. Survey submission by hour of day, broken down by app.

Figure 7 shows the survey submission by time of upload. The average time for survey submission was approximately 2:30PM. The bulk of neighborhood surveys were submitted towards the end of the day (as people came home from school, work, errands).

For most questions, the users had the option to select an 'Other' response and type in their own response. The percentage of

'other' responses were as follows for each of the campaigns: Route: 28.21%, Home: 16.22%, Neighborhood: 27.00%, Work, 1.69%. All the questions in the school application were pictures or free text. Figure 8 shows that 22.1% of the users selected an 'other' response 6 or more times.
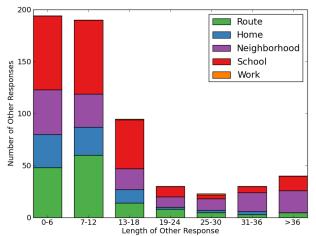

Figure 8. Number of users selecting the 'other' response

Figure 9. Length of response typed into survey when 'other' response was selected.

The majority of the participants kept the length of the free responses short, indicating simply that the multiple-choice selection in the drop downs were not sufficient or expressive. However Figure 9 shows that 20% of the free responses were greater than 18 characters, i.e., a response greater than one to three words, indicating that a phrase or sentence was needed to provide a longer explanation.

5. EVALUATION

In this section we discuss the technical feasibility and usability of the survey apps for experience sampling and community needs assessment based on this community deployment.

5.1 Advantages and Disadvantages

Participatory data campaigns pursue transparency through participation, and as with any large collaboration, the more people participate, the greater diversity of viewpoints and motivations of community members [4]. Inherent advantages and disadvantages arise quickly during the design and implementation stages of campaign.

The frequency of survey uploads, including photograph and free-text submission is evidence of the systems 'adoptability'. The transparent and collaboratory nature of the campaign appears to have provided a compelling rationale for community members to participate. The collective design of survey apps likely contributed to the clear and comprehensible data collection experience. The participatory design also provided an analytic perspective towards data analysis and interpretation because the data collection instruments (i.e. survey questions) were determined by members of the same community from which participants were recruited.

Potential risks to participants who engage in a community data campaign may relate to documentation and the circulation of data—and the multiple levels of context in each survey upload, such as potential personal identification through photographs and geotags of routes. In addition, personal judgment may be brought to bear upon what is and is not appropriate to represent in photographs; this may be especially noticeable in repair questions, questions that feature or ask about people surrounding the participants, and especially the BH-School survey that assumes most users are minors. It would behoove researchers, facilitators, and trainers who introduce the system to new users to consider: images of persons who have not given explicit consent; personal judgment of the user (i.e. high school students); the material inequalities or incentives to participation (do phones get taken away or given back to organizers?); and the difficulty/time that can be spent parsing, classifying, and interpreting a high volume of photographs.

The weakest element of this PS data campaign experience, was the lack of predetermined data analysis practices and tools. The community organizers had not had access to raw data of this type before, and the process of data analysis was slower than that of data collection. This was likely made worse by the relatively small amount of resources available for this portion of the project (both money and time).

5.2 Implications for Future Work

In addition to providing the CBOs with community contributed data for their planning process, the engagement suggested several ideas for future enhancements.

Pretesting: It is common best practice to pretest surveys before large scale deployment in order to identify the best wording of questions and identify other bugs in one's instrument. Online tools make such pre-testing relatively easy to do and its clear from our experience that future deployments would benefit from a short phase for pre-testing the content of each survey before conducting a full scale experiment. In particular, we suggest the use of an initial pretest in which questions make heavy use of the 'other' option, and later surveys can adjust the list of multiple choices offered to participants based on what is provided in the pretest.

Real-time campaign management and analytics. The community campaign organizers in our study would have

benefited from the implementation and use of a tool for real time data-campaign management. Such a tool would assist community organizers at all stages of the process, the logistics of phone distribution and use, to the monitoring of participant engagement (by checking uploads via the password protected website) on a daily basis so that they could step in and offer help as needed. Perhaps most importantly, through seeing the data as they are contributed, the campaign organizers can begin to foster their familiarity with the data and to encourage more extensive analysis of campaign results.

Participant feedback: Our simple applications did not include any mechanism for real time feedback to participants on data uploaded. Both intuition, as well as results from prior work suggests strongly that implementation of some form of user feedback, engagement, or incentives are essential to engage users with their data and to incentivize them to contribute.

Our deployment experience shows that even a simple form of Participatory Sensing can be used for community data gathering with very little training. The process demonstrated the value of the workflow among facilitators, organizers and participants, which involved repeatable iterations (design and testing) and real time data analysis.

More generally, Participatory Sensing data campaigns can support community asset mapping and planning, and can promote visibility and transparency into environmental risks. There is significant economy of scale and scope in building a capacity for participatory sensing to support the broad area of preventive health and community centered interventions. Moreover, the systems readily provide detailed quantitative feedback on the effectiveness of the platform and individual interventions as a byproduct of their use by enabling a naturally iterative process of refinement, deployment and exploration. A community platform could simultaneously create a one of a kind infrastructure focused on health, education, activism, and youth development.

6. CONCLUSIONS

This paper provides initial evidence that participatory sensing can be successfully applied in community data campaigns to address a range of civic and public health initiatives important to community interest groups. Participatory sensing for community data campaigns can be seen as part of a range of recent, innovative and minimally invasive experience sampling methods for community based needs assessment research, biosocial survey research, and participatory research projects.

7. REFERENCES

[1] M. Cargo and S. L. Mercer, "The Value and Challenges of Participatory Research: Strengthening Its Practice," *Annual Review of Public Health*, vol. 29, no. 1, pp. 325-350, 2008.

[2] C. R. Horowitz, M. Robinson, and S. Seifer, "Community-Based Participatory Research From the Margin to the Mainstream: Are Researchers Prepared?," *Circulation*, vol. 119, no. 19, pp. 2633-2642, May. 2009.

[3] Lindau, S. T. and McDade, T. W., "Minimally Invasive and Innovative Methods for Biomeasure Collection in Population-Based Research," in *Biosocial Surveys*, Washington, D.C: National Academies Press, 2008, pp. 251-277.

[4] C. Wang and M. A. Burris, "Photovoice: Concept, Methodology, and Use for Participatory Needs Assessment," *Health Education & Behavior*, vol. 24, no. 3, pp. 369-387, 1997.

[5] A. Stone, S. Shiffman, A. Atienza, and L. Nebeling, *The Science of Real-Time Data Capture: Self-Reports in Health Research*, 1st ed. Oxford University Press, USA, 2007.

[6] A. S. Green, E. Rafaeli, N. Bolger, P. E. Shrout, and H. T. Reis, "Paper or plastic? Data equivalence in paper and electronic diaries," *Psychological Methods*, vol. 11, no. 1, pp. 87-105, Mar. 2006.

[7] "MobileActive.org | A global network of people using mobile technology for social impact.." [Online]. Available: http://mobileactive.org/. [Accessed: 03-Oct-2010].

[8] Y. Anokwa, C. Hartung, W. Brunette, G. Borriello, and A. Lerer, "Open Source Data Collection in the Developing World," *Computer*, vol. 42, no. 10, pp. 97-99, 2009.

[9] M. Clark-IbáÑez, "Framing the Social World With Photo-Elicitation Interviews," *American Behavioral Scientist*, vol. 47, no. 12, pp. 1507 -1527, 2004.

[10] L. Deng and L. P. Cox, "LiveCompare: Grocery Bargain Hunting Through Participatory Sensing."

[11] B. A. T. Brown, A. J. Sellen, and K. P. O'Hara, "A diary study of information capture in working life," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 438-445, 2000.

[12] M. Csikszentmihalyi and R. Larson, "Validity and reliability of the Experience-Sampling Method," *The Journal of Nervous and Mental Disease*, vol. 175, no. 9, pp. 526-536, Sep. 1987.

[13] S. Carter and J. Mankoff, "When participants do the capturing: the role of media in diary studies," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 899-908, 2005.

[14] J. Brandt, N. Weiss, and S. R. Klemmer, "txt 4 l8r: lowering the burden for diary studies under mobile conditions," in *CHI '07 extended abstracts on Human factors in computing systems*, pp. 2303-2308, 2007.

[15] T. M. Piasecki, M. R. Hufford, M. Solhan, and T. J. Trull, "Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations, and barriers," *Psychological Assessment*, vol. 19, no. 1, pp. 25-43, Mar. 2007.

[16] N. Bolger, A. Davis, and E. Rafaeli, "Diary Methods: Capturing Life as it is Lived," *Annual Review of Psychology*, vol. 54, no. 1, pp. 579-616, 2003.

[17] A. Smith, *Mobile Access 2010*. Pew Internet & American Life Project, 2010.

[18] J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, and J. A. Landay, "MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones," in *Proceedings of the 5th international conference on Mobile systems, applications and services*, pp. 57-70, 2007.