

## 2.2 Participatory Sensing (PART)

The vision of Participatory Sensing is of distributed data collection and analysis spanning the personal, urban, and global scale, often using “everyday” technologies like cell phones, in which participants make key decisions about what, where and when to sense. Previously called Urban Sensing, the area was renamed to emphasize its wide applicability and strong conceptual grounding in user participation.

The area targets technologies and applications that transform our capacity to help individuals, families, and communities monitor and improve their own health behaviors, adopt sustainable practices in resource consumption, and participate in civic processes. Each of these three touchstone topics—health (Figure 1), sustainability and civic engagement—is being explored in real-world deployments, such as AndWellness (Google and NIH-funded); the Personal Environmental Impact Report and What’s Invasive; and Mobilize in the Los Angeles Unified School District, Boyle Heights Neighborhood Collaborative and Remapping LA, respectively. In addition to these application-driven experimental deployments, the area has conducted technology-focused research around topics necessary for complete, robust participatory sensing systems, including: participant recruitment and reputation, task planning, and sensing campaign management; configurable triggers for experience sampling; human activity classification based on mobile phone sensors; integration with environmental monitoring assets; and data visualization. Privacy challenges are being addressed within a holistic ethics framework that emphasizes principles of participant primacy, data legibility, longitudinal engagement, and parsimony.



Figure 1. Diagram depicting general mHealth system use cases.

Since the introduction of the term by CENS in 2006, the area of participatory sensing (along with urban sensing) has generated a body of multidisciplinary work spanning many universities, including UCLA, Dartmouth, Columbia, MIT, CMU. It has also inspired work outside computer science in design, urban planning, and the arts, while becoming a driver application for other research topics, including a recent NSF Future Internet Architecture award to UCLA. The more technical aspects of our PS innovations are included in the research section of this report.

### *Application Drivers & Pilot Deployments*

Ongoing collaborations target the three focus areas, including joint projects with the Semel Institute’s Global Center for Children and Families at the UCLA School of Medicine, the National Park Service, the Los Angeles Unified School District (LAUSD), UCLA REMAP, and others. The group continue to increase the scale of public use of these technologies. In Spring 2010, Google seeded a partnership with LAUSD to incorporate participatory sensing on Android phones into computer science and mathematics classrooms. That program, Mobilize, will now expand across the district through a new five-year NSF award (<http://mobilizingcs.org>). It builds on experience from community data collection, for example, a collaboration that began last year with the Boyle Heights Neighborhood Collaborative in Downtown Los Angeles mapped, recorded, and accumulated data on community member circulation and related condition. This was a unique, active and participatory approach that supported the Boyle Heights Planning for Place project in developing its plan for a healthy community. Now, Nokia and CENS are supporting the exploration of new approaches to community case-making and storytelling using this data as a case study, which will in turn impact architecture research. Work continues on an ARRA-funded NIH Challenge Grant to develop an innovative real-time assessment of behavioral exposures for cardiovascular disease (CVD) in young overweight mothers. Other health

science projects include exploratory work around supporting cancer survivorship research, HIV behavior survey with the Center for HIV Identification, Prevention, and Treatment Services, and a collaboration with UCSF on “mHealth” architecture.

Applications continue to take on a wider variety of diverse populations. While in the longer term scalability of any individual deployment/experiment will raise additional challenges, our experience to date has shown that the more important dimension for us to focus on is supporting a diverse mix of dynamic PS efforts, rather than any individual effort at very large scale. For the most part the latter more traditional notion of scaling is likely to be well addressed by mechanisms used in other areas of distributed systems and web applications (cloud solutions, schema-less data stores, etc.) Whereas a focus on an increase in the range of applications and projects is generating highly valuable feedback on system feature set, robustness requirements and optimization targets; participant pool coordination, planning and management needs; and the importance of careful user experience and interface design for successful deployment, and other issues that are more unique to this particular domain.

Platforms used by the group are becoming more mature and general. CENS is generalizing the codebase from AndWellness to also support the LAUSD Mobilize Program, as well as adding visual analytics capabilities using the popular statistics package R. AndWellness includes four system mechanisms to facilitate rapid prototyping of personal data collection. Specific systems contributions include survey authoring; a composable and extensible trigger framework that makes it easy to launch survey data collection based on time, place, or a user’s activity; a phone top ‘button’ that allows a participant to capture a quick emotion (such as a ‘stress button’ to document stress events)—and the time and location surrounding that event—without having to go through the burden of answering an entire survey; low-power data collection services (e.g. location, acceleration, mobility) to facilitate contextual and automated data collection without draining the battery and without interrupting the user; and a toolkit of generic visualizations that provide a quick snapshot of each user’s data.

Rapid campaign authoring, deployment and management tools, are enabling new data collection campaigns to be quickly created and deployed through the assembly and customization of a pre-existing web services and user interface components. Additionally, through collaboration with other institutions, CENS will integrate other emerging platforms and standards for data collection, such as Open Data Kit (ODK).

In addition to smartphone software, CENS is also developing SMS (text message) based systems on top of the UNICEF RapidSMS framework, to enable participation by the still large numbers of people who do not own smartphone or do not have a supported smartphone platform.

On the mobile devices themselves, research continues to better understand handset usage models as they inform resource management mechanisms; especially with respect to power consumption, where improved knowledge will support the goal of being able to continually run participatory sensing applications on everyday handsets. Additionally, we continue to explore embedding local processing to tighten the feedback loop with users. In general, as our deployments expand, we dedicate more effort to user interface improvements and usability study.

#### *Privacy, Ethics, Law and Policy*

Based on campaign deployment experience and ongoing consideration of privacy concerns, a major focus of research and implementation that emerged in 2009-2010 was the Personal Data Vault (PDV): a logically isolated secure repository for participatory sensing data that is controlled by the handset owner. Conceptual and technical work on this continues, including a version based on standard web protocols and another more experimental distributed version. In both cases, the PDV receives participatory sensing data as it is collected and selectively distributes it to third party applications according to a set of sharing rules created by the user. The groups collaboration with Prof. Jerry Kang of the UCLA School of Law has yielded a law review paper and initial legal approach that could provide additional protection for the data contained in the PDV. The implementation will be integrated into CENS applications in 2011.

The PDV is one of several examples of CENS participatory sensing research that is influenced through interaction with ethics education and research that aims to promote the participatory principles and user empowerment fundamental to this area. This work is in its third and final year of funding from the NSF Ethics Education in Science and Engineering, and is centered a participant-observer study of CENS research that aims to develop educational materials promoting ethics considerations in the development of participatory sensing systems, as well as original work in Information Studies on participatory practices in data collection. In addition to academic publications, reports, and popular articles in this area, a new interdisciplinary undergraduate course is being offered this spring that will explore the topics in depth. The project generated a second grant by the same team that will generate multimedia curriculum materials on a several different ethics topics.

### *Future Work*

During the coming year we will continue to focus on expanded real-world deployments with a larger variety and number of users, higher stakes uses in real communities, and more robust, shared systems (such as the PDV and AndWellness core platform) to support data collection campaigns. These deployments will provide the systems scaffolding and practical opportunities to incorporate technology research in campaign deployment, management, recruitment, incentive, data processing and the other areas listed above. They will also provide opportunity for formal user studies and more concrete understanding of how to achieve maximum impact in the real world.

After an assessment of progress at a Summer 2010 “mini-retreat” the group has also started to focus on methodology for case-making and storytelling based on participatory sensing data. Acknowledging that data (and visualizations of that data) are often understood by stakeholders in participatory sensing through the articulation, sharing, and correction of narratives told using the data, the group is researching how to better support this in the collection and analysis process. Assistance in case-making and storytelling will help to motivate ongoing collection in communities and makes it easier for them to leverage the data that they have collected towards desired change. This work includes collaboration with visual and media arts students and faculty, and will culminate in automated data summary approaches that complement the work already being done in data visualization and analysis.

## PART 01 Inferring Everyday Mobility States using Mobile Phones

- Deborah Estrin, Faculty\*
- Hossein Falaki, Graduate Student
- Brent Longstaff, Graduate Student, PI\*
- Nithya Ramanathan, Faculty

\* Primary Contact

### Overview

Previously, CENS developed Ambulation, an Android application which uses a mobility classifier to sense users' activities, i.e., driving, biking, running, walking, or being still. Ambulation was designed to help caregivers evaluate mobility-affecting chronic diseases such as MS and Parkinson's. This year, the classifier, which we call Mobility, has been integrated into our new health application, AndWellness. Mobility data are useful in many campaigns, either as the primary data in campaigns similar to Ambulation, or as metadata in other health-related applications. We improved the energy-efficiency of the classifier and changed to a modular design, allowing multiple AndWellness campaigns to run simultaneously on a phone. Finally, we moved the classifier to the server side of the application to allow modification and use of the classifier even after data has been collected during AndWellness campaigns.

### Approach

Originally, the classifier operated continuously, sampling the accelerometer and GPS all the time. The requirements of AndWellness allowed for a lower granularity of data, so we modified the sensor sampling to support duty cycling. We also changed the GPS adaptive sampling algorithm. This algorithm was originally developed on the Nokia N95, which had a more energy-efficient accelerometer than our Android devices do. We switched to using Wi-Fi for adaptive GPS sampling with a new service, WiFiGPSLocation, which caches Wi-Fi signatures and their associated GPS coordinates, allowing phones to recognize their location by the local Wi-Fi beacons. If they have already cached the GPS location associated with the current Wi-Fi beacons, WiFiGPSLocation can use the remembered value instead of turning on GPS. This way, GPS is only required when the user is moving or in an unfamiliar place. Since people tend to frequent the same places (e.g., work and home), this reduces the amount of required GPS sampling considerably.

### System Description

We updated the activity classification system to become a service usable by multiple applications on the mobile phone. Mobility is now a separate component, implemented as a content provider. It samples sensor data, classifies the user's mobility mode, and stores it, along with the raw data, for applications to access. AndWellness campaigns can retrieve this data via the content provider. Mobility deletes old data after a reasonable timeout (e.g., a week), so all currently-running applications which require mobility data have time to retrieve them. We divided sensor sampling into separate components as well. GPS sampling is now handled in the WiFiGPSLocation service, which offers a synchronous method for Mobility to retrieve the data. We also created a service to handle accelerometer sampling, to provide a simpler interface. Various components of AndWellness, including Mobility, can use this without incurring additional sampling requirements above that of the most demanding component by itself. It also simplifies application design, allowing for synchronous requests for data.

A more recent change was to move the mobility classifier to the server, where it can be run on the uploaded user data. This allows greater flexibility for testing new classifiers, since the classifier can be updated without modifying the phone application.

### Accomplishments

The mobility classifier has become a service, Mobility, and its implementation has improved. Most importantly for users, its improved sampling strategy saves energy so it can now run for a longer period of time without requiring the phone to be charged. It has been decoupled from a specific app, and now is a service available to any other

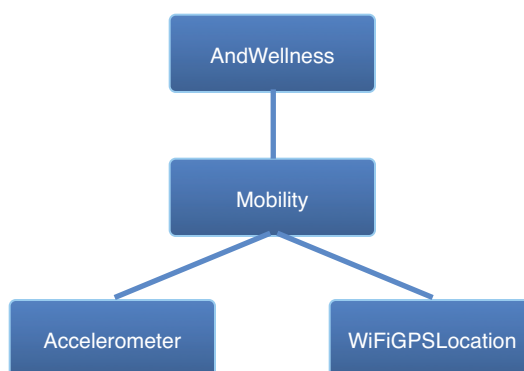


Figure 1. Android service diagram. The accelerometer service and WiFiGPSLocation service sample the accelerometer and GPS, respectively. Mobility gets the samples from them and provides them to other applications. The user activates Mobility through AndWellness, and AndWellness retrieves the stored Mobility data.

participatory sensing application which requires mobility data or metadata. It incorporates the modular design of using separate sampling services instead of directly interfacing with the phone's sensor API. The new server-side classifier allows us to try new classifiers on existing data, and update the model without redeploying applications to users' phones.

#### **Future Directions**

As we deploy the mobility classifier in more AndWellness campaigns, we will continue to tune it to improve accuracy. We are currently training a new classifier model to replace the existing one, which can be inaccurate on newer phones in some situations, such as when the user is running. We may also implement a multi-model democratic classifier or use semi-supervised learning on the server if necessary to make the classifier to perform well across different users and phones.

## PART 02 Sensing Everyday Places and Paths

---

- Deborah Estrin, Faculty\*
- Donnie Kim, Graduate Student, PI\*
- Younghun Kim, Graduate Student
- Mani Srivastava, Faculty

\* Primary Contact

### Overview

As mobile devices have become capable of locating themselves almost all the time, a variety of mobile applications have emerged that seek to continuously track a user's location context. For instance, geo-reminders allow us to set and receive a to-do list whenever we enter or leave a particular place. Social applications plan to provide services for seamlessly sharing whereabouts, querying users that are presently located at an art gallery, and inferring hotspots by the frequency of physical visits by users. Tracks generated by humans also provide useful information for map building, traffic estimation, and ride sharing. Moreover, automatically detected visit and travel behaviors can help studies of human spatial and temporal behavior, and support research for urban planning, sustainability, epidemics, and health care. Interestingly, all these applications can benefit from continuously understanding and keeping track of location as people normally do: places and paths. By automatically learning the places that one visits throughout one's daily life, noticing when one enters and leaves these places, and remembering paths one travels between them, we can unleash many interesting applications.

### Approach

The key challenges we face to provide such service are 1) accurately detecting places closer to our semantics, 2) automatically parsing travel paths from day-long location traces, and 3) minimizing energy consumption. We overcome these challenges by designing a robust place detection algorithm, a sensitive movement detector, and an on-demand path tracker. A place detection algorithm attempts to automatically find places (colloquial representations of locations such as "my office" or "5th floor cafe") that carries a semantic meaning to an individual user. Semantic places are directly inferred from pervasive radio signals by periodically scanning neighboring beacons. To reduce energy consumed during a stay at a place, scans are suspended while a movement detector detects no movement from a more energy-efficient inertial sensor. A path is defined as a set of time series coordinates that interconnects places. Paths are tracked by acquiring periodic position fixes from position systems only when traveling between places.

### System(s) Description and/or Experiments

As SensLoc runs in the background of the mobile device, places are gradually learned as a user visits them and spends a substantial amount of time. A new place is learned by saving its place signature whenever a visit to an unknown place is detected, and sometime later in the day asking the user to confirm and tag a name, such as "home", "Fred's office", or "Organic foods @westwood". A user can recall the place by looking at the visit time, presented as enter and leave time, and the associated geographic coordinate, plotted on a map, provided as a hint. Revisited places are recognized using previously saved place signatures. Entrance to and departure from selected places are notified to applications requesting the *place detection* service. When a user leaves a place, *path tracking* (if enabled) is initiated until the user arrives at another place. Any positioning system available on the device can be used including GPS or systems supported by energy-efficient mechanisms to track paths. If *path recording* is requested, paths are saved, and provided to various applications requesting the service. Unrecorded path tracking can also provide real-time current positions to navigation and location-based search applications with minimum delay by periodically updating the user's current position. This is also when real-time positions are most likely used (e.g., when I'm mobile), and quick responses are most appreciated (e.g., when I'm lost).

Figure 1 presents the overall architecture of SensLoc. The system consists of three main building blocks to provide its service while reducing its energy requirements: place detector, movement detector, and path tracker. The place detector regularly scans neighboring radio beacons to detect place visits when the radio environment stabilizes indicating an entrance. Once an entrance is determined, the place detector consults with the place database to recognize the place and triggers the movement detector to find opportunities to sleep. If no movement is detected, the movement detector signals the place detector to sleep, and awakens it when a movement is detected again. When the place detector senses that the surrounding radio environment is changing, it declares a place departure, saves the visit history, turns off the movement detector, and powers on the path tracker. Path tracking is initiated and records the path to the path database (if enabled) until the next place visit. Path tracker can also hint the place detector to sleep when the user is traveling at high speeds, and unlikely to approach a place anytime soon. We use Wi-Fi access points (APs) to sense places, accelerometer to detect movements, and GPS to track paths.

To evaluate our framework, we gathered three different data sets from both real-life and scripted-tours. Five individuals collected data for a week and two people for four weeks as they went about their normal lives. A scripted-tour data set comprised of 50 visits to 25 different places people go often near a campus. Each volunteer also kept a written diary of places they visited with enter and exit times. Using these data sets, we evaluate SensLoc's effectiveness in detecting place visits, tracking travel paths, and its overall energy consumption during a daily operation. While the performance and cost indeed depends on a user's surrounding and travel patterns, we show that SensLoc consistently outperforms previous place learning techniques, promptly tracks paths, and saves significant energy.

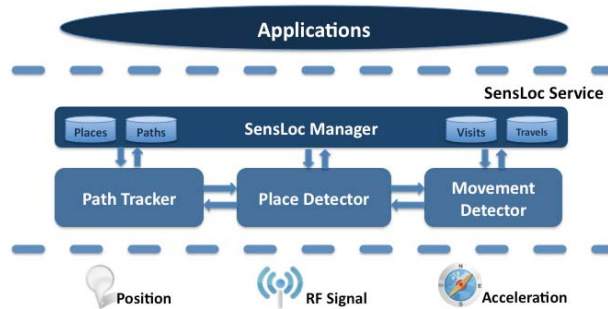


Figure 1

### Accomplishments

Our results show that SensLoc can both semantically and energy-efficiently provide location context to applications by using a combination of acceleration, Wi-Fi, and GPS sensors to find semantic places, detect user movements, and track travel paths. Place visits and path travels are inferred from raw sensor data, which is energy-efficiently achieved by leveraging our tendency to spend about 90% of the time indoors and 10% in a vehicle or at outdoors. Precision and recall of detecting semantic places are both improved compared to the previous state-of-the-art PlaceSense [1] approach by additionally exploiting signal strength changes of the surrounding beacons and adapting parameters to the neighboring beacon density. The accuracy gains are particularly noticeable when a user's routine includes back-to-back visits to nearby indoor places (e.g., rooms on different floors) that shares even a single strong beacon. SensLoc's enhanced place detection algorithm also improves the detected place entrance and departure times by over 2.3 times the precision of previous approaches. However, at some places where beacon signals are weak and unstable, PlaceSense, which only considers the presence of beacons, detects places more robustly. Path tracking is only initiated when a user is traveling between places, which allows us to achieve highly efficient duty cycling of positioning systems (e.g., GPS 8.3% active time), and still covers 95% of the travel distance. This not only saves energy but also boosts the overall quality of the collected position estimates. Lastly, the average power consumption of SensLoc is about 54.8 mW, which is 6.2 times less than that of collecting GPS periodically. On average, accelerometer, Wi-Fi, and GPS are activated for about 20-22, 2-4, and 1-2 hours everyday, respectively.

### Future Directions

We believe we have solved some of the major practicality issues with continuous location tracking, and illustrated that an approach with a holistic and semantic point of view may provide a realistic solution for many applications. Our results also suggest that there is still more room for improvement to push the place detection performance even further. Adaptive approaches intermixing several place learning techniques based on the radio environment and the application needs may allow us to cover the remaining 5% places that are challenging. Using more energy-efficient sensors may also reduce the energy cost. For example, cell tower information, which almost comes for free, can replace Wi-Fi scans, if mobile service providers become less reluctant in disclosing cell tower information and more platforms provide common APIs to scan every neighboring cell towers. However, we think most research should focus on developing an application stack with a well-defined set of APIs, and create a feedback loop with the users that could tell us what is really important to address. The outcome of these field studies will expose application demands and provide nuances to tune the system for particular uses or situations.

[1] Donnie H. Kim, Jeffrey Hightower, Ramesh Gorvindan, and Deborah Estrin, in Proceedings of the Eleventh International Conference on Ubiquitous Computing (UbiComp'09), (Orlando, Florida), Oct. 2009.

## PART 03 Bodies in Motion

---

- Jay Chien, Graduate Student
- Elizabeth Do, Undergrad
- Nabil Hajj Chehade, Graduate Student
- William Kaiser, Faculty
- Claire Lochner, Undergrad
- Greg Pottie, Faculty, PI\*
- Tremaine Rawls, Undergrad
- Natali Ruchansky, Undergrad

\* Primary Contact

### Overview

The focus of the Bodies in Motion project is to infer human activity using data from accelerometers worn on the body. It encompasses machine learning (for making inferences from the data), optimal placement of the sensors, training regimens, support tools for determining the feature set and labelling the data, development of models useful for classifying and searching the data, and an experimental program to collect the data. The project origin in CENS is from supplemental funding granted for the medical implications of sensors worn on the body. In the previous year, extensive work had been done in creating a signal processing toolbox based upon Naive Bayes classifiers to make use of data from one or two triaxial accelerometers. A feature set had been created based on segmenting the data into fixed four second windows, with independent decisions on the activity taken in each window. This is important on its own, as human activity level is a strong predictor of future conditions such as diabetes. Also in the prior year a set of experimental questions had been formulated with the view to considering what would happen if more sensors could be placed on the body. The overall goal is more precise classification of activities, so as to be useful in such domains as rehabilitative therapy and sports training. In the reporting period, experiments were carried out in the summer undergraduate research program, leading to considerable evolution of the tools and experimental methods. These were carried over into a senior undergraduate design class, leading to even greater changes in the toolbox and opening new research directions that will be pursued in the coming year. The success potentially creates a new model for how to cycle between summer undergraduate projects, bringing research to the classroom, and advancing graduate student research.

### Approach

The hard problem in machine learning is constructing the set of features that should be used in classification. These features collectively specify a model. Once this is done, there are many methods that can be used to actually perform the classification. Our general approach is to create an intuitive model that can be used for each of classification, data search, and explanation. For example, some classifiers give good accuracy but obscure the connection to the physical actions. However, for sports training or rehabilitative therapy, it is also necessary to know in what precise way an activity is done so that corrective instructions can be given, or a diagnostic decision made. A tree-based system decomposes activities into classes using intuitively understood features. There is potentially more work in constructing it, but the resulting classifier is then easy to train for a given individual, has low computational cost, and can serve all our desired purposes. To construct it however, extensive experimentation is required.

### System Description and Experiments

The Personal Activity Monitor (PAM) device consists of a triaxial accelerometer that can be sampled at a fixed rate, storage, and a USB port for getting power and uploading the data. Twelve were acquired for the summer program, and an additional sixty for the design class in the fall with an enrollment of 18 students. Straps were also acquired for attaching the sensors to the body over clothing. MATLAB based software is then used for tagging the data and classifying the activities.

In the summer, there were two major experimental phases. In the first, the students devised a set of scripted activities to be pursued at various locations around the campus. One student would wear the PAM devices while the other would record the activities being done. The objectives included determining what features should be used for classifying the data and how much training time would be required. In the second, two PAM devices were used for classifying office activities, including sitting, standing, fidgeting, walking, and slouching. The general experimental procedures developed in the summer served as the basic pattern for the design class in the fall. In the senior design class, a wide variety of experiments were pursued, using more sensors, including classification of proper form in weightlifting and tennis serves, different types of walking, first responder activities, training time requirements for everyday activities, and 3-D visualization of motions. For the course, it became apparent that there were

considerable hardware reliability issues, and need for improvements in the support tools for feature selection, synchronization, and data labeling.

### **Accomplishments**

In the summer program, four undergraduates were supervised by two graduate student researchers to conduct a series of experiments. The undergraduates were empowered to make decisions on how experiments were to be conducted and features selected. A breakthrough realization was that organization of the large set of activities into a decision tree would both reduce the dimensionality of the feature space required for bifurcation of the classes, and reduce the required training time. This greatly improved efficiency in selecting features and collecting data. Moreover, since the Naive Bayes classifier already implicitly assumes independence of features, the conditional independence assumption required for tree construction imposes little penalty in classification accuracy. Further, thinking about activities in such a hierarchical fashion leads to an intuitively understood model for the activities. Very high classification accuracies were obtained for the office activities by first dividing them among the low-energy and high-energy activities, and then working on the subgroups. Training of only a few minutes was required to then accurately classify unscripted activities. This is in contrast to many studies in which good clinical results are obtained only to have disappointing results afterwards. The exciting outcome is that even extremely simple classifiers can do a good job if first a short data set is collected for each individual and the features and decision tree are carefully selected. A refereed conference paper resulted. All four undergraduate researchers have applied to grad school, and two applied for NSF fellowships.

The use of multiple PAM devices in the design class revealed shortcomings in hardware reliability, and also the need for better synchronization methods. The TA (Jay Chien, who had also been involved in the summer research program) devised some methods, and the students in the class suggested better experimental approaches. Jay developed a tool to enable visualization of how the features separated the classes, which soon became the workhorse of all the groups. The projects also revealed limitations of the tree classifier system, which had a fixed decision period of 4 seconds. This was not nearly small enough for getting at the details of the phases of walking or tennis serves. Thus a variable sampling window size capability was added. This led to a new view of how we should construct the tree, with different sampling window sizes for different branches as more detailed classifications are made. The 3-D visualization project was also very fruitful. The idea of the student group was to use the classifiers to drive the visualization. An unforeseen benefit is that this also enables creation of motions that a subject can try to copy, and thus the tool potentially could be used to train new physical activities. This class was a first attempt to have undergraduate students in a regular class perform research. It succeeded very well: the students proposed audacious projects which have helped set the research agenda for the coming year. Currently there are also student groups pursuing projects that will extend the toolsets in a graduate embedded systems class.

This leads to a new model for the interaction of research and teaching. Summer students under the supervision of graduate student researchers devise and carry out an experimental program which stretches prior tools and procedures. This testing hardens the experimental system to the point that it can be used in a design class. The class poses scaling and ease of use issues which lead to further developments, and proposes experiments that lead to new research questions. These are pursued by the graduate student researchers in planning for the next summer session. The graduate class can also lead to thesis proposals for the students who have pursued. In the process many more American students are exposed to research and encouraged to pursue graduate studies, while enabling experiments on a scale that would not be possible for the graduate students working alone. Thus, both research and instructional needs are advanced.

### **Future Directions**

Work is proceeding in four general directions. First, we are working with the device manufacturer to improve reliability, not only for use by students but in a series of major clinical studies that will shortly be starting. Second, refinements to the tree classifier are being made to improve its ease of use, particularly in largely automating the tedious process of labeling data. Third, we are investigating use of alternative classifiers at different branches of the tree. One student is currently on exchange at the Australian Centre for Field Robotics in Sydney to work with machine learning experts there. Fourth, we are looking into how to properly frame context. We can for example suppose that the likelihood of particular activities is governed by time of day (a high level) and by the activities most recently pursued. This is naturally modeled with a Markov structure. Potentially, use of context information can improve the reliability of classifiers, and also reduce the energy required to gather a statistically significant sampling of the activities. We anticipate another set of experiments with summer interns this year, and further experiments in undergraduate and graduate design classes. A fifth direction that may be launched in the coming year is work relating the models implicit in our tree to the problem of search of sensor data for making medical inferences. The tree structure is compelling in that it leads to an easily understood model; the features used for classification may also serve as tags for enabling search.

## PART 04 Adaptive Battery Management on Smartphones

- Deborah Estrin, Faculty, PI\*
- Hossein Falaki, Graduate Student\*
- Ramesh Govindan, Faculty, PI

\* Primary Contact

### Overview

Modern mobile phones are not single-purpose devices anymore. But rather, they are multi-functional programmable computers. Users run a plethora of applications in addition to voice calling on their smartphones. Significant diversity in usage habits combined with the diversity in hardware platforms makes battery life of smartphones unpredictable. If batteries lasted long enough, unpredictability would not have been a serious concern. However, the linear battery improvements are no match for the rate of new features and applications on smartphones. As a result, the average battery life time of smartphones gets shorter on each new generation of smartphones in the market.

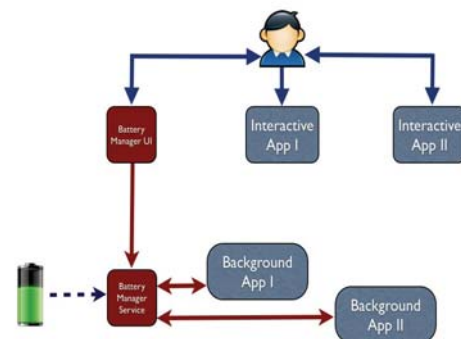
In addition, many pervasive computing applications, such as those developed at CENS, have components that continuously run in the background. Such applications are particularly power consuming. Unlike traditional fully interactive applications, users do not have direct control over the resource consumption of background tasks. Therefore, many users are usually startled by their smartphone's short battery life when running CENS applications.

In this work we introduce a new system to give users control over their phone's battery life. We do so by CENS applications, which run in the background and consume significant power, adaptive to usage and context.

### Approach

Four approaches to managing power consumption to minimize user surprises can be considered:

- Most systems leave battery management entirely to applications. Most commercial mobile platforms such as Android and Symbian follow this approach.
- The opposite of the previous approach is managing energy as a primary system resource. Applications can consume only what is allocated to them by the OS. ECOSystem and Cinder are re- search operating systems that follow this approach.
- A third approach to guaranteeing a reasonable battery life is limiting the amount of work that applications can do without the user's direct control. For example, the iPhone OS does not allow arbitrary background jobs by limiting multitasking.
- Our approach that we propose in this work combines the best of the first and second approaches.



Architecture for adaptive battery management on smartphones. The Battery Manager Service monitors resource usage and energy consumption and signals background applications.

Managing energy consumption at the highest layer, i.e., applications, is more effective. The application can make better choices to trade off accuracy and performance with energy. On the other hand, applications lack a global view of the system. Specially, effective battery management requires knowledge of other applications workload, something that applications do not expose to others for a good reason.

### System(s) Description and/or Experiments

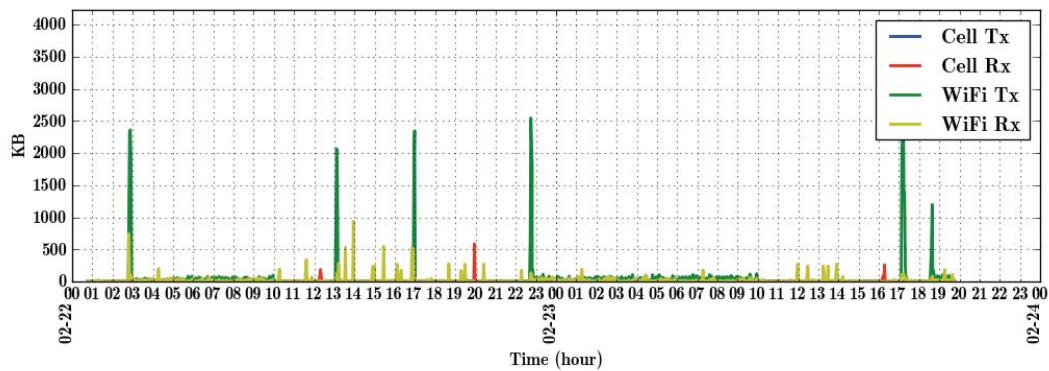
#### *SystemSens (Power Monitor Service)*

We continued improving and extending SystemSens, the logging software tool that was developed at CENS to log detailed system related information on Android smartphones. SystemSens can now run on any Android phone and report on a wider range of operating system events. We also implemented a web interface to visualize SystemSens traces. Each user can login to this service and monitor resource consumption on her phone through a number of time graphs. The SystemSens service builds models that related consumption of different resources to battery level changes on the phone of each individual. The model building process is repeated every night for each user and the resulting model is sent back to the SystemSens Client on the smartphone. The client then uses this model to regulate resource consumption of background services to meet user specified battery deadlines.

### WiFiGPSLocation Service

WiFiGPSLocation is an Android service to simplify duty-cycling of the GPS receiver when a user is not mobile. It runs as an Android Service on the phone and defines a simple interface that all other CENS applications can use to get the updated location of the user. To duty-cycle the power consuming GPS receiver uses the WiFi RF fingerprint to detect when a user is not moving to turn off GPS, and when the user starts moving to turn it back on.

We implemented this service to be adaptive based on the signals from SystemSens. Therefore, any other location-based CENS application, such as Andwellness mobility, that uses this service will be adaptive as well.



Example of a SystemSens network traffic graph during two days of the smartphone of an example user

### Accelerometer Service

As another example of an adaptive service that reports its resource consumption to SystemSens and complies with the energy consumption signals that SystemSens sends, we implemented a services to manage the Android accelerometer. Several other CENS applications such as Andwellness Mobility and SleepSens use this service.

### Future Directions

The next step of our research is evaluating the performance of this system. Our efforts to this point have been focused on the implementation of this system. We will evaluate our system using real users and improve its performance if needed. We also plan to release our system to other researchers who face similar problems with managing smartphone battery life.

## PART 05 Estimating Energy Consumption using Inertial Sensing on a Smartphone

- Adar Emken, Staff
- Donna Spruijt-Metz, Faculty
- Gaurav Sukhatme, Faculty, PI\*
- Harshvardhan Vathsangam, Graduate Student\*

\* Primary Contact

### Overview

Regular physical activity (PA) is beneficial for health, including reduced obesity and risk for cardiovascular disease, type 2 diabetes, and several forms of cancer. Walking is the most common type of activity among people who are physically active. Characterizing energy expenditure due to walking ubiquitously and in real-time would be a vital tool to assess physical activity levels and determine the effectiveness of interventions. Inertial sensors on cell-phones (such as accelerometers and gyroscopes) represent a promising opportunity in characterizing intensity of walk because of their sensing capability, portable form-factor and increasingly powerful computing capabilities. The availability of a mobile computing platform also opens up the possibility of developing machine learning algorithms that learn user models.

### Approach

The goal of this project was to determine whether on-body inertial sensing with mobile phones together with probabilistic machine learning techniques could be used to characterize energy expenditure as measured by rate of oxygen consumption ( $VO_2$ , mL/min). We did this for one type of activity, steady-state treadmill walking. Steady-state treadmill walking was chosen because of the ease of data collection over a range of walking speeds. At the time of this study, existing mobile phones did not have the required sensor suite or sampling capability as was desired. Hence we used an on-body inertial sensor interfaced to a laptop or a mobile phone. Our approach involved modeling joint probabilities of streaming data to train a linear probabilistic model, Bayesian Linear Regression (BLR). We aimed to explore whether probabilistic techniques could be used to improve predictions and whether alternative sensors like gyroscopes could be used to predict energy expenditure.



Fig. 1: An example recording procedure for a single participant. The yellow box indicates sensor mounting.

### System(s) Description and/or Experiments

#### Hardware

We captured motion using a modified version of the Sparkfun 6DoF Inertial Measurement Unit (IMU) worn on the right iliac crest. The sensor provided 6 sensor streams conveying triaxial acceleration (Freescale MMA7260Q tri-axial accelerometer) and triaxial rotational rates (2 Invensense IDG300 gyroscopes). The use of sensors in all three axes allowed movement capture in all three planes – sagittal, frontal and transverse. Data were sampled at 100 Hz and transmitted via Bluetooth (RN-41 Bluetooth module) to a nearby PC. We also capture data with the HTC Nexus one running Android 2.2.1. Rate of oxygen consumption ( $VO_2$ , mL/min) was used as the representation of energy expenditure. This was measured using the MedGraphics Cardio II metabolic system with BreezeSuite v6.1B (Medical Graphics Corporation). The metabolic system outputs data at the frequency of every breath.

#### Participants

Eight healthy adults (four men, four women) participated in the study. Each participant was asked to walk at 5 speeds (2.5, 2.8, 3.0, 3.3, and 3.5 mph) on a motorized treadmill for 7 minutes of recording time per speed. Following a transition between speeds,  $VO_2$  readings were allowed to stabilize for 2 minutes prior to data collection.

#### Pre-processing

Each sensor stream from the IMU was passed through a lowpass filter with 3dB cutoff at 20 Hz. This frequency was chosen keeping in mind that everyday activities fall in the frequency range of 0.1-10 Hz. Each stream was divided into ten-second epochs. Within each epoch, the 1024 point normalized FFT was extracted to obtain frequency information. An FFT was chosen to represent data because steady state treadmill walking is periodic and periodicities can be represented well with an FFT. The  $VO_2$  values from the MedGraphics metabolic system that fell within each epoch were averaged and matched appropriately. These represented the ground truth on which to train data. Thus

data for each user consisted of a sequence of epochs, each containing features from the IMU and the average rate of oxygen consumption for that epoch.

#### *Algorithmic Mapping*

We model the mapping from FFT features to energy expended with Bayesian Linear Regression (BLR). BLR adopts a probabilistic approach to linear regression by using priors over regression weights. For training the algorithm, a fraction of the data was uniformly sampled and partitioned into training data, the remaining fraction constituting test data. Different models were trained with the same training data but with different feature vectors. RMS error was calculated as a measure of accuracy. This was repeated over 10 trials for different randomly sampled data and results averaged. This was repeated for training data percentages from 10% to 90% and constituted a per-subject measure of performance. For reference, this algorithm was compared against standard linear regression.

#### **Accomplishments**

Our study showed that while standard linear regression results varied heavily with the amount of training data available (due to overfitting or bias), BLR showed consistent performance for the same data. Larger amounts of training data improved BLR predictions. While employing BLR, accelerometer and gyroscope data simultaneously improved prediction accuracy. Among accelerometer features using X-axis acceleration showed the next lowest prediction error. This was because the X-axis was aligned with the direction of forward movement. All three gyroscopic axes showed comparable errors. Using gyroscope features in Y and Z axes showed lower errors than corresponding acceleration Y and Z axes. This is due to gyroscopes only capturing dynamic rotational movement free from gravitational bias. Gyroscopes were capable of providing equivalent if not better results for energy prediction from treadmill walking. Additionally, combining accelerometer and gyroscope information reduced prediction errors.

#### **Future Directions**

We plan to expand our work in a number of directions. We are currently working on developing generalized models that are applicable across a range of physiological parameters such as height, weight, BMI, gender and age. This involves undertaking a larger study and collecting movement information for treadmill walking. We are developing a generalized linear model similar to Bayesian Linear Regression with informative initial conditions based on physiological parameters.

Another important issue is that we restrict ourselves to steady-state walking on a level plane. While our results are promising, further work is needed to generalize to overground walking in free-living conditions. For this we plan to study the performance of our algorithms for overground walking. If successful, this has the potential to vastly benefit the field of human calorimetry by providing accurate VO<sub>2</sub> values while offering the convenience and cost-effectiveness of inertial sensor based activity monitoring. We also plan to undertake similar analyzes for other kinds of cyclic activities still operating under the steady-state condition.

We are also currently working on porting all algorithms based on inertial data to a Google Nexus S phone and are developing a front-end interface. We plan to make this publicly available in the near future.

## PART 06 Energy-Efficient Rate-Adaptive GPS-based Positioning for Smartphones

- Ramesh Govindan, Faculty
- Jeongyeup Paek, Staff

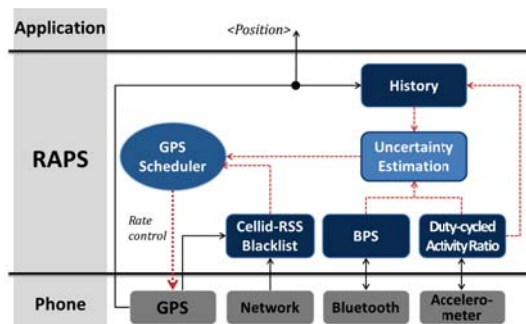
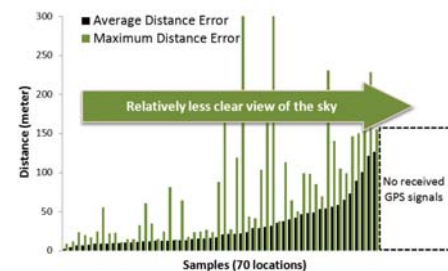
### Overview

Many emerging smartphone applications require position information to provide location-based or context-aware services. In these applications, GPS is often preferred over its alternatives such as GSM/WiFi based positioning systems because it is known to be more accurate. However, GPS is extremely power hungry. Hence a common approach is to periodically duty-cycle GPS. However, GPS duty-cycling trades-off positioning accuracy for lower energy. A key requirement for such applications, then, is a positioning system that provides accurate position information while spending minimal energy.

In this project, we develop RAPS, rate-adaptive positioning system for smartphone applications. It is based on the observation that GPS is generally less accurate in urban areas, so it suffices to turn on GPS only as often as necessary to achieve this accuracy. RAPS uses a collection of techniques to cleverly determine when to turn on GPS. It uses the location-time history of the user to estimate user velocity and adaptively turn on GPS only if the estimated uncertainty in position exceeds the accuracy threshold. It also efficiently estimates user movement using a duty-cycled accelerometer, and utilizes Bluetooth communication to reduce position uncertainty among neighboring devices. Finally, it employs celltower-RSS blacklisting to detect GPS unavailability (e.g., indoors) and avoid turning on GPS in these cases. We evaluate RAPS through real-world experiments using a prototype implementation on a modern smartphone and show that it can increase phone lifetimes by more than a factor of 3.8 over an approach where GPS is always on.

### Approach

The key insight that motivates our work is the observation that, when used by pedestrians in urban areas, GPS can exhibit errors in the range of 100m. GPS inaccuracy in urban “canyons” is well known, but we have found that, even in relatively benign environments such as college campuses or residential neighborhoods, GPS can exhibit this kind of inaccuracy, especially for pedestrian smartphone usage. Location-based applications will have to deal with this level of error using application-specific methods, such as map matching or map-snapping. So we ask: if applications can tolerate this position error, why not trade off some position accuracy for reduced GPS energy usage? A simple way to do this is to periodically duty-cycle GPS. This trades-off positioning accuracy for lower energy. However, the key challenge in this periodic GPS duty-cycling is to decide on a suitable time period; for almost any choice, there exists a user mobility pattern that will result in unbounded position error.



estimated distance traveled approaches a user-specified accuracy bound, our system (called RAPS, the rate-adaptive positioning system) activates GPS. This decision to activate GPS is delayed if the user’s current average activity level, as measured by a duty-cycled accelerometer, is inconsistent with historical activity at that position and time-of-day. Similarly, RAPS delays activation if the identifier and the signal strength from the currently active cell-

To avoid this, a thread of research has examined different heuristics to cheaply determine a change of position so that GPS can be selectively activate. Our work follows in this thread, but makes two contributions. First, it introduces novel techniques for cheaply inferring whether and when GPS activations are necessary. More important, rather than merely considering each technique in isolation, we design a complete system that uses a collection of techniques in concert to reduce energy usages.

At the core of our approach is a method to estimate user velocity from a history of previously measured velocities at the same location and the same time-of-day; intuitively, this velocity estimation leverages consistency in user behavior. When the

tower indicates that previous activation attempts at locations with comparable identifier and signal strength information failed frequently. Finally, RAPS delays GPS activation if it learns, over Bluetooth, of a more recent position fix completed by an opportunistic contact.

**System(s) Description and/or Experiments**

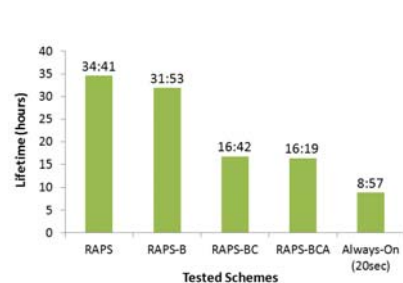
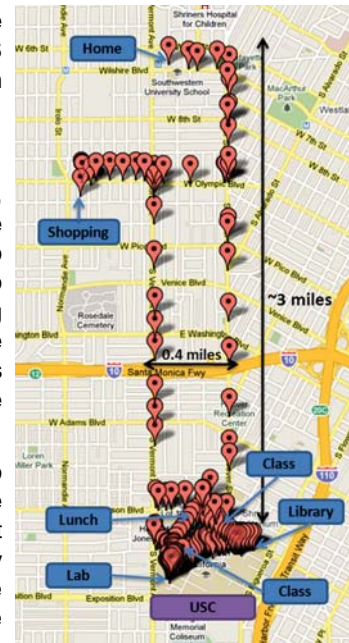
RAPS uses a collection of techniques to cleverly determine when to turn on GPS, and when not to. First, it uses a duty-cycled accelerometer to efficiently estimate user movement. RAPS detects whether the user is moving or not, and also measures the activity ratio, the fraction of time that the user is in motion between two position updates. The movement detection is used to prevent RAPS from activating GPS when the user has been stationary. The activity ratio is used to estimate the current velocity based on historical correlations between velocity and activity; this lets RAPS activate GPS only if the estimated uncertainty in position exceeds the accuracy threshold.

Second, RAPS stores the space-time history of user movements to estimate when to activate GPS. Whenever RAPS gets a new position update, it calculates the average velocity relative to the previous position, and associates this velocity and the recent activity ratio with the previous space-time coordinate. It uses averages of the velocity and activity ratio to estimate, during a subsequent visit to the same location at the same time-of-day, the likely user velocity. This in turn enables RAPS to estimate positioning uncertainty, allowing it to selectively activate GPS.

Third, RAPS employs celltower-RSS blacklisting to detect GPS unavailability (e.g., indoors) and avoids turning on GPS in these places. Whenever it succeeds or fails in obtaining new position update, it records the current celltower ID and the received signal strength (RSS) information and associates with that success or failure. Then, when it determines that it is time to activate GPS, it checks the celltower-RSS table for the historical probability of GPS availability and defers GPS activation if it believes that GPS is not likely to be available, thus avoiding unnecessary energy usage.

Finally, RAPS utilizes Bluetooth communication to reduce position uncertainty among neighboring devices. Whenever it receives a new position update, it broadcasts this information to Bluetooth peers so that they can update their position without activating GPS themselves. If a device receives a position update from a peer that has greater uncertainty than its own estimate, it replies with its more accurate information. Eventually, all devices in the neighborhood synchronize to the position information with least uncertainty.

We have implemented these techniques in a complete RAPS prototype on the Nokia N95 smartphones and have experimented with it on our university campus.



**Accomplishments**

Our evaluation reveals that RAPS has over 3.8x longer lifetime than a scheme in which GPS is always on, and about 1.9x longer lifetime than a periodic GPS scheme with comparable error rate. We also break down the contributions of each technique in RAPS towards these performance gains. Finally, we demonstrate that RAPS can be easily adapted to work atop a WiFi-based positioning system, WPS. In general, our evaluation is encouraging and suggests that RAPS can obtain substantial energy savings without sacrificing the accuracy of periodic GPS.

**Future Directions**

A lot of work remains to be done before RAPS can become truly practical. Our parameter settings (for example, the accelerometer duty-cycling) can be optimized. Although we believe the system has the right incentives for the adoption of Bluetooth synchronization, privacy and security considerations need to be considered. Finally, we intend to perform a more careful evaluation of the parameter space, which is exceedingly large: performance critically depends on user consistency and phone handling behavior, on environmental conditions, GPS visibility, celltower placement density, and so forth.

## PART 07 Named Data Networking: Applications

- Rahul Rao Basava, Graduate Student
- Alex Bonomo, Undergrad
- Jeff Burke, Faculty, PI\*
- Derek Kulinski, Undergrad
- Alessandro marianantoni, Staff\*
- Chenni qian, Graduate Student
- Zhenkai zhu, Graduate Student

\* Primary Contact

### Overview

This research is conducted within the larger NSF project Named Data Networking (NDN) which implements an inter-campus collaboration with 10 US campuses lead by CENS. Such collaborative research wants to create the Internet of the Future, replacing today focus on where – addresses and hosts – with what – the content that users and application care about.

At CENS and REMAP (REMAP is a joint collaboration research center between the UCLA Engineering School and UCLA School of Theater, Film and TV) we'll develop 3 (or more) applications as use case using the new network approach. Instrumented environment and participatory sensing are the two broader fields of investigation that describe the research context and link to past and ongoing activities at CENS.

Within the 3-year plan, the goal is to have a video conference application, a lighting control application and a Personal Data Cloud (PDC) system running over the NDN testbed (which connects the involved campuses) and by doing so to have addressed conceptual and technical issues that the new approach challenges us.

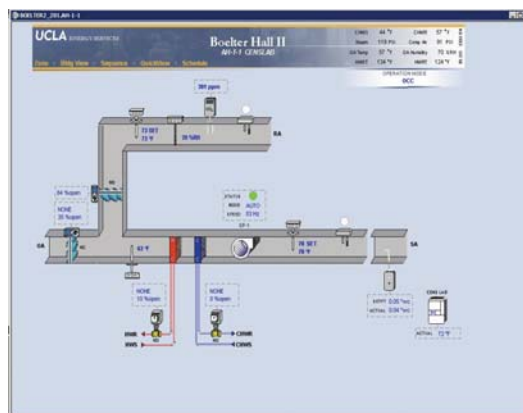


Figure 1. UCLA Building Management System Siemens InSight, detail of the HVAC system at CENS.

### Approach

Many internet applications will expand requiring always more network resources in order to provide high definition content, sensing and control and a consistent substrate to implement the Internet of Things. The buildings of the future will require control systems able to scale complexity and optimize data network traffic. International collaborations between corporations are developing standards that will push intelligence in appliances and devices at levels far beyond today's requirements.

One of the research challenges is to define a naming paradigm, perhaps a well defined grammar that is able to generate name hierarchies and suggest specific algorithms to parse and operate with the naming.

Background studies on Building Management Systems (BMS) highlight the relationship between physical space and systems in buildings. Based on a RESTful architecture scheme, the project Integrated Sensor Stream Storage System (IS4) [4] has been taken as a reference to build the name-space hierarchy for our approach.

The team members want the project results to have impact also outside of the academia. To do so several approaches are initiated, among those: create a visual representation through an architectural lighting installation, that will support the communication of the results to a broader audience. Data for the project will be provided by UCLA F&M through the UCLA BMS.

Another aspect regards the collaboration with corporation that are working actively and already gave a strong impact on the development of actual systems in the above mentioned fields. Particularly a conversation on media rich communication for environment has been initiate with Crestron, even if the company will have a more passive role since their interest on not divulgating their methods and technologies. Another attempt is in process thanks to contacts at Siemens AG, which is a sponsor of CO2morrow (AM past project in EU).

Such research process involves many participants and includes great deal of integration and communication between the parts so to maintain a cohesive path. To support the process that traces a common path back from multiple investigation interests, integration of works, open questions until the final unifying goals, I'm remotely attending Tarek

UIUC groups' application meetings and I also integrate security protocol requirements from UCI into the design of the lighting application.

### Systems Description

*AUDIO CONFERENCE TOOL (to be expanded to video) student: Zenkai Zhu*

An audio conference tool that is able to discover all existing conferences and allows rendezvous where participants can find each other and they can be enumerated over the network. It is based on the Mumble application and extended to be able to run over NDN.

Focus on: implement and run the conference tool over the testbed sites.

*PERSONAL DATA CLOUD (PDC) student: Derek Takeda*

It is a privacy oriented structured application for storing data from various self monitoring devices (such as mobile phones). It is a cloud-inspired NDN based version on the Personal Data Vault (PDV).

Focus on: implementation of PDC over the NDN testbed, replication of data across multiple PDV under a granular Access Control List, data encryption.

*LIGHTING SYSTEM (to be expanded for BMS) Student:Chenni Qian*

It uses and integrates lighting control and building management system principles. The lighting interface is based on a Linux embedded network system (Gumstix) to be coupled with Philips lighting controllers and LED fixtures.

Focus: integration of lighting control and building control over NDN testbed.

*NDN INTEGRATION WITH BUILDING MANAGEMENT SYSTEM (BMS)*

focus: to receive and send over NDN testbed data streams related to resource consumption and production, Heating, Ventilating, and Air Conditioning (HVAC) from the UCLA BMS with support from UCLA Facilities & Management.

### Accomplishments

First stage definition of the syntax rules to create naming addresses, in regard of Functional Recommendations for Internet Resource Locators", RFC 1736, and Functional Requirements for Uniform Resource Names, RFC 1737.

The design documents and communication protocols with recommended security protocol from UCI for the 3 applications is at a stage that students are developing initial prototypes.

### Future Directions

Research more background studies on naming in relationship to our NDN project and possible expansion regarding "Internet of things".

Because of personal research interest my agenda includes development of a meaningful document on the cultural and perhaps social implication of the NDN approach.

Description of the naming in regards of: RFC 5234, as December 2010 the IETF provides definition language for communication protocols.

Specific milestones for the applications:

- Mid-March : requirements to receive data stream from UCLA BMS
- End-March : Architectural Lighting should be able to distribute code and bootstrap new controllers out-of-the-box
- End of June : integration BMS with NDN
- June: The conference tool expands including video.
- End of July : PDC Serve as back-end for UIUC applications



Figure 2. NDN application interfacing UCLA BMS, Illustration by Dominic Foog and Ren Rong.

More specific milestones after August 2011 are in the overall NDN schedule from Jim Thorton "ndn-sched.pdf". A general milestone by the end of the year is to run an evaluation of the functionality of the application over the testbed.

## PART 08 Architectural Lighting Application of Named Data Networking

- Jeff Burk, Faculty, PI
- Deborah Estrin, Faculty, PI
- Alessandro Marianantoni, Staff
- Chenni Qian, Graduate Student\*
- Mani Srivastava, Faculty\*

\* Primary Contact

### Overview

Modern building systems, such as architectural lighting, serve both functional and aesthetic purposes. This project explores lighting control system that employs Name Data Networking (NDN) protocols to create expressive instrumented environment. NDN is a novel content-centric networking architecture that treats data contents instead of IP address as the primary entity. It decouples location from identity, security and access. NDN requests and retrieves data content by name, making network design more flexible and scalable. Motivated by the advantages of NDN, we design NDN-based architectural lighting applications to enhance the performance of actual lighting control systems. Thus facilitating control in a secure way across networks and providing more intuitive ways of addressing and controlling lighting fixtures.

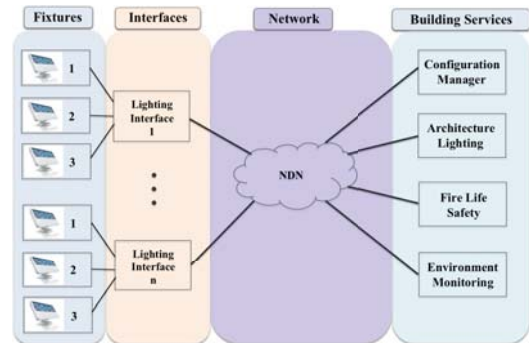


Figure 1. NDN-based Lighting Control System Diagram

### Approach

Many future Internet applications will expand the vision of ubiquitous computing to high definition content and interactivity, sensing and control, distributed processing, and user interfaces, at scales and complexity far beyond today's applications. Future buildings are likely to be constructed with digitally-controlled, addressable lighting and environmental systems, access control systems with a variety of presence, flow, and identity sensing, touch-activated networked displays and projection, paging or sound systems, and video recording or broadcasting capabilities.

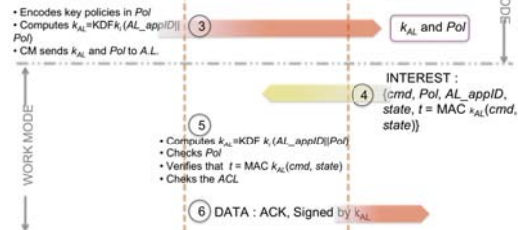


Figure 2. NDN Lighting System Security Diagram

The application agenda aims to develop a lighting application system for buildings that is able to address the following key research aspects:

**Security:** create a protocol over NDN capable of authorizing entities for controlling the fixtures. In other words, it is requested to implement a system to handle multiple keys with multiple sinks and entities that need to be authorized. The protocol is also able to efficiently sign and verify packets.

**Naming strategies:** create a naming topology for lighting installations, which potentially will extend to building management systems. We analyze the relationship between building systems' spatial structures and fixtures' meaningful deployment. This lead us to the design of two naming strategies for objects in buildings: spatial organization, and mixed space-system organization.

### System(s) Description and/or Experiments

#### Subsystem Description

The NDN based lighting control system consists of Fixtures, Interfaces, Network, and Building Services as shown in Figure 1. Fixtures are lighting facilities used to illuminate or create lighting environment. In this application, we use ethernet controlled Philips ColorBlast 12 LED wall-washing lights as lighting fixtures. Interface is a computing unit that is embedded in the fixture power supply unit to take care of the communication between the fixtures and NDN network. We use an ARM Cortex-A8 based embedded Linux platform, Gumstix Overo Air COM, as the interface. Network is the NDN network, the NDN testbed and its subset at UCLA in our implementation. Building Services are applications in the building automation control systems that control and monitor the building infrastructures such as

lighting or HVAC systems. In our application, we consider four building services: Configuration Manager (C.M.), Architecture Lighting (A.L.), Fire Life Safety (F.L.S.) and Environment Monitoring (E.M.).

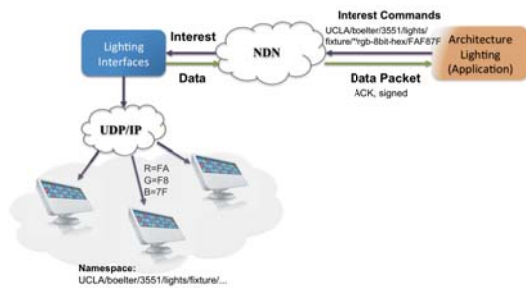


Figure 3. NDN-based Lighting Control System Diagram

Configuration Manager is responsible for maintaining and assigning namespace to lighting fixtures and other applications in NDN network. Architecture Lighting is an application in control of building lighting infrastructures. It expresses control commands to achieve both functional and aesthetic lighting purposes. Fire Life Safety includes fire prevention, suppression, and emergency evacuation and response. It is a building service with the highest priority that overrides other commands in case of fire in order to guarantee fire life safety aims. Environment Monitoring is a read only application that allows building residents to monitor lighting status and energy consumption in the building.

#### Naming Strategy

Every data in NDN network is identified by a name. Each object can have multiple names according to different naming strategies. In our implementation, we consider two naming strategies – spatial organization and mixed space-system organization.

##### (1) Spatial organization:

```
<campus>/<building::Name_id>/room/<room::Name_id>/lights/fixture/<fixture::Name_id>/rgb-8bit-hex/<fixture::HEX_value>
```

##### (2) Mixed space-system organization:

```
<campus>/<building::Name_id>/lights/fixture/<fixture::Name_id>/map/<x::int,y::int>/rgb-8bit-hex/<fixture::HEX_value>
```

##### (3) Naming generalization:

Naming address ::= <location>/<system>, where

```
<location> ::= <campus>/<building id>/room/<room id>
                | <campus>/<building id>/floor/<floor id>
                | <campus>/<building id>/floor/<floor id>/hallway
                | <campus><building id>
<system> ::= fixture/<fixture id>/<light_value>
                | HVAC/thermostat/<thermostat id>
<light_value> ::= rgb-8bit-hex/<8bit_hex_value>
                | rgb-16bit-hex/<16bit_hex_value>
                | hls-8bit-hex/<8bit_hex_value>
<building id> ::= boelter | melnitz | royce hall | ...
```

#### Security Model

The design of key exchange, access control, and authentication scheme is illustrated in Figure 2. In our system, the Configuration Manager distributes secret keys to each lighting fixture. It also encrypts key policies and updates Access Control List (ACL) for building services such as architecture lighting. Take architecture lighting as an example of application. After the authorization has been granted, the A.L. sends lighting control command embedded in Interest packet to the target fixture. The fixture then authenticates the identity of the architecture lighting. If authentication is successful, the fixture sets fixture RGB/HLS accordingly and sends a signed Data packet back to the application.

The security protocols involved in the authorization and authentication implementation include symmetric encryption algorithm, Message Authentication Code, collision resistant hash function, and Key Derivation Function.

#### Accomplishments

We have cross-compiled CCNx code base to ARM Cortex-A8 based embedded Linux platform, Gumstix Overo Air COM. We empirically proved that running CCNx on embedded platform is feasible.

On the fixture side, we implemented fixture discovery program based on light-specific UDP/IP protocols. The program is running in the background in specified periods. The result of the fixture discovery is maintained in the Configuration Manager to update namespaces and ACL for both lights and applications.

We developed architecture lighting application that sends lighting control commands in the Interest packet to lighting interfaces, gumstix in our case, over NDN. The lighting interfaces then parse the interest command and send RGB/HLS control information to the actual fixtures via UDP. The result is an illuminated environment with the lighting strategies specified by the architecture lighting applications. An illustration of the interest/data exchange mechanism is shown in Figure 3.

### **Future Directions**

In the coming year, our objective is to fully develop a working NDN-based lighting control system. In the short term, we pursue comprehensive implementation of building services with sophisticated security model, and be able to distribute code and bootstrap new controllers out of the box. Later in the year, we would like to expand the system to larger scales and semi-permanent setups, and extend to incorporate dynamic control from other data sources and/or Building Information Modeling examples. Challenges exist in scalable namespace design, avoiding namespace to be arbitrarily long; security design and implementation; and creative dynamic control schemes development.

## PART 09 Personal Data Cloud

- Rahul Basava, Graduate Student
- Alex Bonomo, Undergrad
- Jeff Burke, Faculty, PI\*
- Deborah Estrin, Faculty, PI
- Derek Kulinski, Graduate Student\*

\* Primary Contact

### Overview

The Personal Data Cloud is conceptually based on the Personal Data vault project. A Personal data Vault (PDV) is a secure repository of data. The Personal Data Cloud (PDC) refers to all PDV instances controlled by a single user. The PDV instances collect, store and distribute user's data. The user can apply access controls, filters and other data altering operations before letting a particular PDV instance transfer information to another PDV instance. All the communicating PDV instances are tailored for a particular application and have the ability to understand and work with the data exchanged between them. This project deals with the design and implementation for such a PDV which can securely store and transfer data to other PDVs over the Named Data Networking (NDN).

### Objectives

- User managed secure repositories for time and location traces controlled by the person whose traces they are.
- Granular and configurable sharing with provision for archiving data.
- Audit trail enforcement and API Controlled access by user authorized third-party services.
- Migration from provider to provider over lifetime.
- Deploy the project to run over the NDN network to use the security and delivery efficiencies which are built in features provided by NDN.

### Technical Challenges

- Support efficient federation of queries across PDVs.
- Defining "auditing" and designing enforcement mechanisms.
- Virtualization to isolate user vaults and application data processing in cloud providers.
- Design Security and Privilege model in context of medical and legal frameworks.
- Limits circumstances for legal discovery.
- Creating standards for best practices and auditing.
- Enable a marketplace of 'certified' applications that adhere to the standards.

### Approach

Every entity like a user's phone or desktop is considered to be a PDV. Such PDVs will be primarily considered as sources of information generated for a particular user. The users also lease or own PDVs that may be maintained by trusted organization (Providers).

Implementation is planned for two versions for the PDC, one for the cloud based providers and the other for mobile devices developed on the android operating system. These versions will share many common security and



Figure 1. Overview of Personal Data Cloud

transmission components built on top of NDN. The two versions will differ in their ability to process information and data storage function. Also the cloud version is expected to have a very comprehensive user interface for managing data. Other function like archiving, processing and filtering will be provided to each version based on the need of specific application.

The Andwellness project is the first application that will be using the PDV framework. The Android app of Andwellness will be changed to work with the android PDC version and the server will have a proxy application that will form a intermediary between the Andwellness server and the NDN transmission system. A Cloud PDC will be developed to apply filter and restrict information sent to the Andwellness servers. The filters will initially work based on time and location information.

### **System Description and a Use case**

The PDV instance that is sending data is called a 'Publisher' (PUBLISHER) and the PDV instance receiving the data is called a 'Receiver' (RECEIVER). A PDV instance can be both a publisher and a receiver. For example a PDV instance on a mobile phone can send data. Whereas a PDV instance in the cloud (say an instance managed by a PDV provider) could receive data from a mobile phone and in turn send data to another PDV instance after processing the data.

Since this implementation is deployed to work over the NDN, the communication is by nature connectionless and individual data records are transmitted independently from PUBLISHER to RECEIVER. The communication is inherently disruption tolerant and secure.

#### *Transferring from Mobile phone to Personal Cloud- A use case*

This use case describes how a user interacts with the PDV instances when sending data from a mobile phone (PUBLISHER) to a secure repository (RECEIVER) in the cloud managed by a PDV Provider.

1. The user downloads the Mobile application from a site whose identity is vouched for by some certificate authority like Verisign. The application gathers data about the user. This application runs a lightweight PDV instance to provide caching and disruption tolerance. The instance may also act as a store to facilitate local processing.
2. The cloud PDV instance has a user interface (UI) which allows the user to register and provides necessary information for connecting and gathering information from the applications that the user owns.
3. The user can have multiple identities based on public keys, digital signatures or username password pairs.
4. The user selects "Send Data" in the mobile application. The mobile application asks for data destination URL. The user provides the URL of the cloud PDV instance to which it has to connect for transferring data. The URL can be generated by the cloud PDV and shown to the user in the cloud PDV instance UI.
5. The mobile PDV instance initiates a stream creation process. The user types in the random token (generated by the cloud PDV) to authenticate with the cloud PDV instance. The instances automatically exchange stream related information which includes, name of the stream, key for encryption and key for signing.
6. The cloud PDV instance issues regular interests to the new stream and pulls data to its persistent store.

### **Progress**

The implementation is in the initial stage and we are exploring the available package of NDN. The security model for transmission and the transmission mechanisms have been finalized. The Andwellness Framework (version 2.0) is in the process of being changed to work with the Mobile version of the PDC.

### **Future Directions**

Developing the prototype implementation of the PDC for Andwellness framework is the main target. This will be followed by doing a survey for Sleep study using the new platform. The implementation will have a android PDV and a cloud based PDV with provision for time and location based filters which are user configurable.

The PDC will be extended to implement other applications.

## PART 10 Personal Data Vaults

- Jeff Burk, Faculty, PI
- Deborah Estrin, Faculty, PI\*
- Ramesh Govindan, Faculty, PI
- Mark Hansen, Faculty, PI
- Shuai Hao, Graduate Student
- Nilesh Mishra, Graduate Student
- Min Mun, Graduate Student\*
- Katie Shilton, Graduate Student

\* Primary Contact

### Overview

The increasing ubiquity of the mobile phone is creating many opportunities for personal context sensing, and will result in massive databases of individuals' sensitive information incorporating locations, movements, images, text annotations, and even health data. In existing system architectures, users upload their raw (unprocessed or filtered) data streams directly to content-service providers and have little control over their data once they "opt-in".

We present Personal Data Vaults (PDVs), a privacy architecture in which individuals retain ownership of their data. Data are routinely filtered before being shared with content-service providers, and users or data custodian services can participate in making controlled data-sharing decisions. Introducing a PDV gives users flexible and granular access control over data. To reduce the burden on users and improve usability, we develop three mechanisms for managing data policies: Granular ACL, Traceaudit and Rule Recommender.

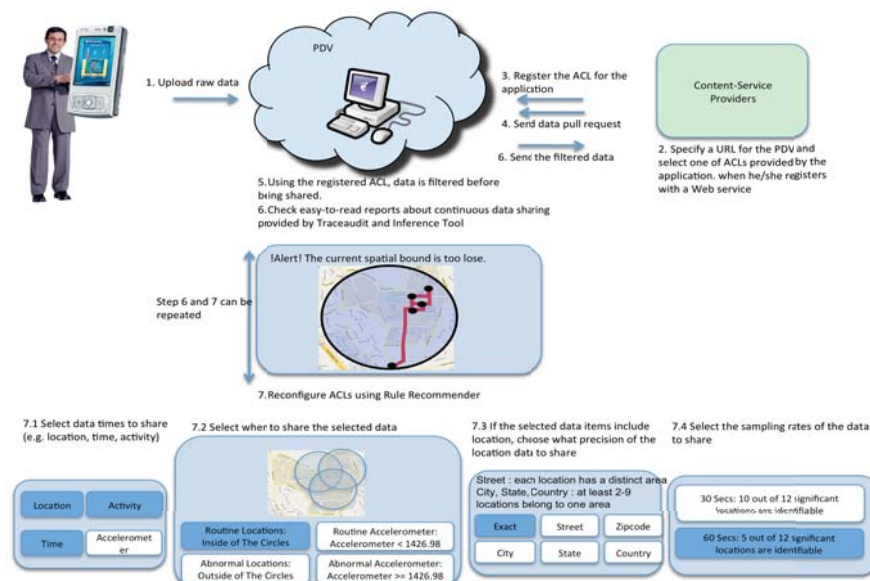


Figure 1 System usage scenario

### Approach

The centerpiece of our approach to privacy for personal participatory sensing data is an individually controlled secure data repository that we call Personal Data Vault (PDV). The PDV decouples the capture and archival of personal data streams from the sharing of that information. Instead of individuals sharing their personal data streams directly with services, we propose the use of secure containers to which only the individual has complete access. The PDV facilitates the selective sharing of subsets of this information with various types of Access Control Lists (ACLs). Rather than relying on third parties to correctly control the sharing, we argue that the owner of the data can actively participate in making controlled data sharing decisions. Similar architectures [1-4] have been introduced for location-based online social networks. We build on this earlier work to create a complementary platform for exploring the functional performance and legal implications of supporting privacy preserving personal participatory sensing applications using PDVs. Our system offers three significant privacy benefits over prior work. First, using Granular ACLs, our system provides users with fine-grained control on data. Users (or trusted guardians on behalf of users)

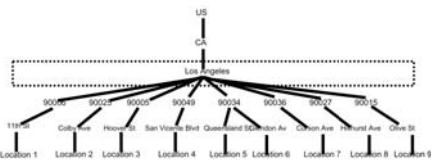


Figure 2 Location tree of significant locations to probe the implication of sharing location data at each of the aggregation levels

can define not only who has access to what data but also the granularity of data for sharing, which gives users more flexibility in expressing sharing decisions. Second, Traceaudit logs and displays transactions and transformations of users' data and enables users to track who visibly viewed their data, how frequently they viewed the data, and how the data is used. It also gives users easy-to-read reports about continuous data sharing by visualizing the consequences of data sharing, using currently and previously shared information. Users can be alerted to update their sharing decisions when their sharing policies may have been judged by the system to be different from their intention; alternatively, users can periodically review these reports in a manner analogous to reviewing

credit reports. Third, Rule Recommender pre-calculates constraint values for a set of pre-defined privacy policies, makes the computed constraints available for display, and facilitates the policy re-configuration

### System Description and Experiments

A Granular ACL contains an entity and a set of filters. The entity specifies the type and name of third parties accessing data from the PDV. The filter has a list of constraints that define data to be shared: Each constraint is described by type and its attributes. To be very specific, we use the JavaScript Object Notation (JSON) to format the ACLs. When a third party, such as a content-service provider, requests data at the PDV, the PDV first determines the filters associated with the applicable entry to decide what granularity of data to send. Our prototype PDV currently supports the following three constraints: bound for limiting data space to be shared, precision for governing the precision of data value to be disclosed and Frequency for ensuring that the shared data does not exceed a specified sampling rate.

The prototype PDV system permits two types of Traceaudit mechanisms: local Traceaudit to log operations performed inside PDV and Traceaudit for third-party applications to track actions occurring outside PDV (i.e. inside third-party applications that read data from the PDV). The Traceaudit mechanism for third party applications currently provides data access and usage logs for traceability, but is limited to data shared by the PDV directly with third-party applications, and does not track data derived by the applications themselves.

The Rule Recommender provides a high-level interface for setting sharing policies. Specifically, it pre-computes constraint values (ACLs) for a set of pre-defined high-level user intentions). An example of the pre-defined high-level user intention is that users wish to share data of their regular/routine patterns or exceptional/abnormal behavior. A routine or anomalous behavior may be defined by any contextual information provided by users. We deal with location or numeric application-specific data like sleep quality for Waking-survey. Another example is that people may want to share location data at coarser levels like street, city, state and country so that their current or visited locations are not identifiable.

We demonstrate how our system facilitates the fine-grained data sharing using the two policy management mechanisms and real data traces collected from Ambulation and Waking-survey.

One example is the following. After using Ambulation for a while, a user reviewed a report provided from the Local Traceaudit. She found that an unusual visit to the health center, which she didn't expect to share with her friends, was disclosed. She felt uncomfortable with her current sharing settings and decided to change the sharing policies with help from the Rule Recommender. The Rule Recommender showed pre-computed constraint values for a set of pre-defined privacy policies as shown in Figure 3. She chose "Routine Locations" and "Exact" options first to share exact location when she is at routine locations (Figure 3 (a) and (b)). Then, she picked "Abnormal Locations" and "City" options to share location at a city level when she is at abnormal locations (Figure 3 (c) and (d)). She did so because she found that location data at a zip code level was not coarse enough to represent her intention that her visits to unusual significant locations should not be identifiable. Figure 2 shows the location tree constructed for her unusual significant locations. From this tree, the system found that when location data was aggregated to a city level or coarser, more than two of significant locations always belonged to one area. Using the new privacy policies, she didn't expose her visit to the sensitive location and made sure that other significant locations were unidentifiable by sharing them at city level. At the same time, she was able to track her location traces in the area where she spent the majority of time everyday.

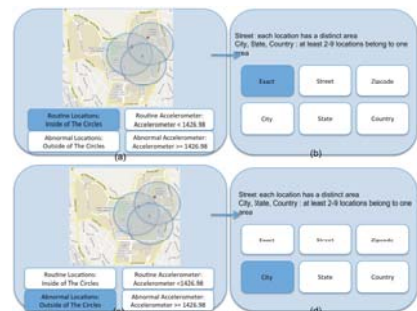


Figure 3 Sample Rule Recommender user interfaces

### **Accomplishments**

We present the PDV architecture, an individually controlled secure data repository decoupling the capture and archiving of personal data streams from the sharing of that information. To illustrate the feasibility of PDV, we developed a prototype of the architecture and demonstrated how our system facilitates the fine-grained data sharing using the two policy management mechanisms and real data traces collected from Ambulation and Waking-survey. We also evaluated the system resource requirements of the core PDV functions, storage and selective sharing, as well as policy management services that are essential to support PDV use. This work is published in the 6th International Conference on emerging Networking EXperiments and Technologies (CoNEXT) 2010[5].

### **Future Directions**

The most critical focus for future work is to evaluate and improve the usability of the system. In addition, Extending Traceaudit to a richer information flow traceability, especially for derived data, is part of the future work.

[1] R. Cceres, L. Cox, H. Lim, A. Shakimov, and A. Varshavsky. Virtual individual servers as privacy-preserving proxies for mobile devices. In Proc. of 1st ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds (MobiHeld), 2009.

[2] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, and D. Starin. Persona: An online social network with user-defined privacy. In SIGCOMM, 2009.

[3] S. Seong, J. Seo, M. Nasielski, D. Sengupta, S. Hangal, and et al. PrPI: A Decentralized Social Networking Infrastructure. In ACM Workshop on Mobile Cloud Computing and Services: Social Networks and Beyond, 2010.

[4] A. Tootoonchian, S. Saroiu, Y. Ganjali, and A. Wolman. Lockr: Better privacy for social networks. In CoNEXT, 2009.

[5] M. Mun, S. Hao, N. Mishra, K. Shilton, J. Burke, D. Estrin, M. Hansen, R. Govindan, "[Personal Data Vaults: A Locus of Control for Personal Data Streams](#)," To appear in The 6th International Conference on emerging Networking EXperiments and Technologies (CoNEXT) 2010, Philadelphia, PA, November 30, 2010



for offloading, and provide considerable opportunities to exploit parallel processing. In this work, we develop a runtime called Odessa that automatically and adaptively determines how best to use these techniques.

### System Description

Motivated by the lessons from the earlier measurement study, Odessa adaptively exploits pipelining, data-parallelism and offloading to improve performance and accuracy of these applications. The Odessa runtime runs on the mobile device; this enables Odessa to transparently improve application performance across different mobile platforms when the mobile is disconnected from the server. The design the Odessa runtime has three goals, in order of decreasing importance:

- It must simultaneously achieve low makespan and high throughput in order to meet the needs of mobile interactive-perception applications.
- It must react quickly to changes in input complexity, device capability, or network conditions. This goal ensures that transient changes in makespan or throughput are minimized or avoided.
- It must have low computation and communication overhead.

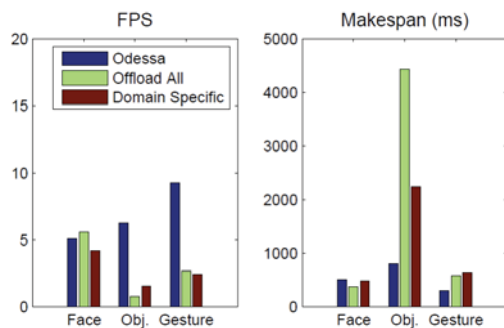


Figure 2. Main Result on Netbook (FPS: Frames Per Second)

Prior approaches for offloading frame the problem using a discrete or graph optimization formulation. For this approach to be effective, accurate estimates of stage execution time are required on both the mobile device and the server, which are often obtained by offline profiling. However, the results in our measurement study show that the execution time can vary significantly and cannot easily be modeled offline. Odessa uses a greedy algorithm that periodically acquires information from a lightweight application profiler to estimate the bottleneck in the current configuration. Then, its decision engine uses simple predictors based on nominal processor frequencies, and a recent history of network measurements, to estimate whether offloading or increasing the level of parallelism of the bottleneck stage would improve performance. Surprisingly, this greedy and incremental approach works very well to improve makespan and throughput, and incurs

negligible overhead. Rarely, Odessa's decision may need to be reversed because its estimators may be off, but it has a built-in self-correcting mechanism to maintain stability.

### Accomplishments

We have explored the design of a runtime, called Odessa, that enables interactive perception applications on mobile devices. The unique characteristics of the applications drive many of the design decisions in Odessa, whose lightweight online profiler and simple execution time predictors help make robust and efficient offloading and parallelization decisions. Our evaluation of Odessa shows that it can provide more than 3x improvement in performance compared to application configurations by domain experts as in Figure 2. Additionally, Odessa can adapt quickly to changes in scene complexity, compute resource availability, and network bandwidth.

### Future Directions

Much work remains, including exploring the performance of Odessa under a broader range of applications, extending it to take advantage of the public cloud, and exploring easy deployability.

## PART 12 MobiProg: Adaptive Programming System for Cloud-Enabled Smartphone Applications

- Dero Gharibian, Graduate Student
- Ramesh Govindan, Faculty, PI\*
- Michael Gray, Graduate Student
- Nupur Kothari, Staff
- Taehee Lee, Graduate Student
- Todd Millstein, Faculty, PI
- Nilesh Mishra, Graduate Student
- Luis Pedrosa, Graduate Student\*
- Jeff Vaughan, Staff

\* Primary Contact

### Overview

With the advent of Smartphones as an emerging class of personal Internet capable devices, mobile applications, or apps, are rapidly becoming the corner stone of what defines the user's experience. While many of these apps will perform simple tasks on the phone, a new class of cloud-enabled applications is becoming more popular. These apps augment the smartphone's capabilities, allowing them to leverage the large datasets and the computational power that can be harnessed in large-scale computing infrastructures known as server clouds.

Many of these cloud-enabled mobile applications delegate all or most of the business logic to the cloud, reducing the smartphone to a thin client. However, there is a vast universe of unexplored possibilities where the smartphone collaborates more closely with the cloud to enhance the user's experience. Such applications can be described in abstract as a function  $F$ , as shown in Figure 1. This function can depend on time-varying inputs from the phone ( $I(t)$ ), as well as data-sets residing in the cloud ( $D(t)$ ). The MobiProg project aims to create a new framework for the development of such cloud-enabled smartphone applications.

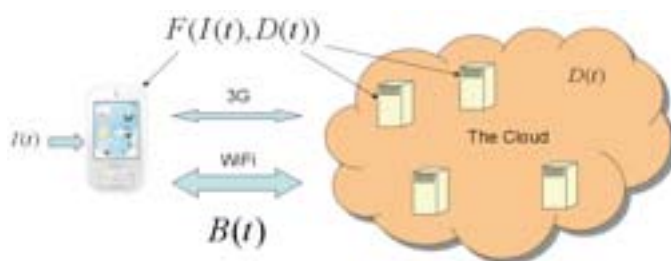


Figure 1: Abstract problem description.

Using this framework, developers implement a single application in a way that is agnostic to running on the phone or in the cloud. Once the application is implemented using this framework, a run-time system partitions the application, automatically executing the relevant portions on the phone or in the cloud so to maximize a given performance metric (e.g. processing time or battery life). This approach not only removes the burden of manual partitioning from the developer but also accounts for situations where the optimal partitioning decision depends on the runtime context, such as radio technology in use (GSM, CDMA, or WiFi) and the amount of charge remaining in the battery.

### Approach

The MobiProg framework allows applications to be implemented following a simple data-flow-like pattern, i.e. applications are written as a linear pipe-line of components. Furthermore, since individual components are implemented using the Java language, the Java Runtime Environment automatically handles most issues regarding component-level portability, namely between the server and the Android smartphone platform, our initial development target. By enforcing such constraints on the developer, the partitioning algorithm is drastically simplified, being reduced to making the decision of which components execute locally or remotely. The run-time system then invokes each component on the right platform, automatically marshalling any data to where it is needed.

Given a run-time such as this, the key to executing applications in an optimal way lies in the quality of the decisions made by the partitioning algorithm. Such a decision becomes particularly challenging given the many factors that equate into it. From a high-level point of view, one or more consecutive components should be executed remotely if their local processing cost is greater than that of sending the data to the remote system, executing there, and retrieving the results. Already the decision depends on such dynamics as the kind of radio in use, as this directly impacts the cost of sending and receiving data over the network. However, the harder question is how does one predict the cost of executing a component on a particular platform without actually performing the execution.

Previous frameworks have tackled this issue by making the simplifying assumption that such a prediction can be made simply by averaging the cost of previous executions. While this assumption may suffice for some components, it entirely disregards several key factors such as algorithmic complexity and the influence of the input data on the algorithms resource usage. To account for these factors, MobiProg employs a prediction framework where a set of predictor functions can be leveraged to more accurately derive component costs for each execution and, consequently, make better partitioning decisions.

We have implemented several prediction functions with varying degrees of generality and precision. The simpler functions, for example a linear regression of experimental profiling data vs. input size, may provide less accurate predictions but can operate with virtually any input data type. To further improve this accuracy, these metrics can be combined with application specific heuristics, for example a word counting function to predict the complexity of speech recognition algorithms. Finally, a more general approach that we have also developed tries to retain a high level of accuracy while keeping the generality of the simpler approaches by using static analysis to heuristically determine the algorithmic complexity of the component. Using this final approach, each component's code is statically analyzed off-line, producing a signature indicating its input complexity dependencies. Armed with these signatures, the run-time system can analyze concrete input data and get a quick yet precise estimate of the cost of executing the component with that specific input.

### **System Description**

An initial run-time system has been implemented for both the server platform and the Android smartphone platform. While the current implementation does not yet perform fully automatic partitioning, it does allow us to manually test several partitioning configurations and validate assumptions. A few key applications have also been ported to use this system, namely a speech recognition and translation application, a bar-code scanner, an optical character recognition application, and a Sudoku solver.

The prediction framework has also been implemented with several initial prediction functions. These prediction functions already include some generic heuristics based on input size and average executions times, as well as some application specific heuristics for the speech recognition component (e.g. word counting and signal-to-noise ratio).

Finally, the static analysis based complexity heuristic is under active development and can already accurately detect the complexity of some simple hand crafted test-cases. Additional corner cases are being handled as the system is being tested with publicly available third-party libraries.

### **Accomplishments**

While a full experimental evaluations has not yet been completed, a few initial tests have been conducted on the proof-of-concept system. Preliminary findings seem to suggest that there is in fact room for improvement in static partitioning schemes, as system context and dynamics can have a significant effect in the outcome of the partitioning decision. Experimental profiling data also shows that component complexity, as one would expect, can also vary drastically based on input.

Further experimentation is needed to evaluate the final system, specifically to quantify how much of an improvement, if any at all, can be achieved in relation to statically partitioned applications and state-of-the-art dynamic portioning schemes.

### **Future Directions**

In the upcoming year we expect to finish the development of the proof-of-concept framework prototype and to complete the effort of porting key initial applications. A more thorough experimental evaluation will also be conducted and the work is expected to be published in peer-reviewed venue within this time-frame.

## PART 13 AndWellness I: System Summary

---

- Deborah Estrin, Faculty, PI\*
- Hossein Falaki, Graduate Student
- John Hicks, Staff
- John Jenkins, Graduate Student
- Brent Longstaff, Graduate Student
- Mohamad Monibi, Staff
- Kannan Parameswaran, Graduate Student
- Nithya Ramanathan, Faculty, PI\*
- Josh Selsky, Staff

\* Primary Contact

### Overview

Introducing smartphones to the health care process is a “natural and necessary” component of 21st Century global health that makes good on the promise of personalized medicine regardless of individual economic status/developing nation status. The “always-on” and “always-worn” status of smartphones makes it possible to collect real-time behavioral signatures as indicators of health and disease with resolution spanning from the individual to the population. These signatures, akin to genetic fingerprints or cardiac rhythms, provide up-to-the-minute indicators of individual and population health status, which can be utilized by a doctor who is helping an individual patient reduce stress in their life, or by an epidemiologist who is studying a populations cardiovascular disease risk factors. By personalizing behavioral interventions and diagnoses for health management and research, smartphones facilitate equal access to the benefits of personalized medicine globally.

Engaging participants through self-monitoring in general has been shown to be critical to succeed in behavioral intervention programs [Donovan and Marlatt 2005; Marlatt and Gordon 1985]. But more affordable, engaging, valid, reliable, and feasible tools to self-monitor behaviors are required to address the limitations of current methodology. To minimize error in self-reporting behaviors, a technique called ecological momentary assessment (EMA) was developed to monitor affect, cognitions, and behavior in real time in a persons natural environment [Shiffman et al. 2008]. EMA has been shown to greatly increase the validity and reliability of patient reported data. Smartphones can significantly increase the power of EMA by providing information on contextual, spatial, or temporal associations to behaviors, and with reduced burden on the user. Yet, phones have not been sufficiently utilized to track and measure health behaviors through space and time [Story et al. 2009].

We have developed AndWellness in order to enable rapid prototyping of end to end EMA systems, and to continue to engage participants in data collection. We extend the definition of EMA, which traditionally is limited to survey based data collection, to include data collected from the numerous monitoring devices that are now available both on the phone (e.g. GPS, accelerometer, camera) and off the phone (e.g. wireless heart rate monitors, blood pressure cuffs). Using AndWellness, a researcher can ask a participant, for example, to monitor exercise, stress, and blood pressure several times a day using a combination of human-in-the-loop measurements and automated monitoring; view all the data together; and then update the monitoring ‘prescription’ as needed. Further, AndWellness makes it possible to measure a participants timely adherence to the data collection process, and configure when and why a participant is queried to collect data. Both are necessary to ensure high quality and unbiased data collection. AndWellness aims to make it easy for doctors and researchers to conduct rapid experience sampling studies to study individuals or populations in situ.

### Approach

The design of AndWellness is driven by interviews with key behavioral researchers, focus group sessions with over 60 diverse potential users of a mobile self monitoring system like AndWellness, and surveys of breast cancer survivors and young moms who have used AndWellness in several research studies to self-monitor diet, exercise, mood, and stress. Through all of these intensive interactions, we have observed a desire for individual customizability (of interest to both researchers and end users), minimal user burden (especially of interest to end users), and validity and reliability of measures (especially of interest to researchers). As a participatory sensing application, AndWellness provides customizable spatial-, social-, temporal-, and mobility-triggered reminders, assessments, and interventions that can be designed by the user, and are relevant to a wide array of behavior change objectives.

### System(s) Description and Experiments

AndWellness includes five system mechanisms to facilitate rapid prototyping of participatory health data collection, storage, and analysis: 1) low-power data collection services (e.g. mobility) for automated data collection without

draining the battery; 2) Survey authoring; 3) a toolkit of generic visualizations that provide a quick snapshot of each users data; 4) a composable and extensible trigger framework that makes it easy to launch survey data collection based on time, place, or a users activity; and 5) a phone top 'button' that allows a participant to capture a quick emotion (such as a 'stress button' to document stress events) – and the time and location surrounding that event – without having to go through the burden of answering an entire survey.

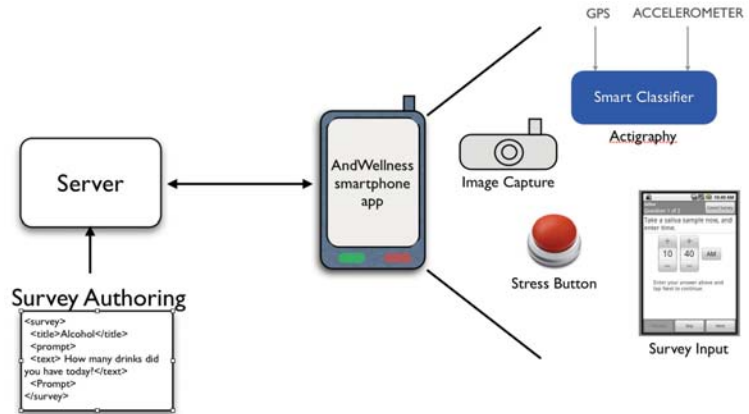


Figure 1. Data flow from phone to server. Data collection includes survey responses, image captures, stress/food button pushes, and actigraphy traces.

### Mobility

We have a mobility classification module which classifies a participant's activities into 'still', 'walk', 'run', 'bike', and 'drive' using features computed on the phone based on continuously sampled GPS and accelerometer data. A number of research areas are based on improving the customizability, power consumption, and adaptability of this classifier, discussed in other sections. In this section, we discuss our efforts in improving achieving high accuracy classification for the AndWellness deployments. C4.5 decision tree algorithm is used for classification. Our experience demonstrates that this decision tree must be trained using data collected on the same platform that will be used to collect data during classification (e.g. Android, or Nokia), to accommodate differences in accelerometer hardware and driver software. For example, when the decision tree classifier is trained with data collected on a Nokia platform, and then run on data collected on an Android platform, classification for a set of scripted activities is approximately 75%, with classification of 'run' having the worst performance. This accuracy increases to 90% or more, when the decision tree classifier is trained and then run on data collected on the Android platform.

### Researcher Oriented Survey Authoring.

To design a campaign with a set of surveys, the author creates an XML configuration file conforming to a provided schema. The server validates the configuration when it is loaded into the system. To increase the modularity and generalizability of the system, this XML configuration file is also used by the data visualization system to dynamically create visualizations. Each prompt type is mapped to a specific type of visualization. The web-based visualization client can then use the XML file to dynamically create visualizations for each survey, providing quick out-of-the-box visualizations for any campaign. The data point API provides a general interface to access data.

### Datapoint API

Our data point abstraction format meets three requirements. First, the format is amenable enough to represent all prompt data type and mobility data. Using the datapoint API, data consumers are able to obtain all necessary context about the data without requiring separate or additional configuration information. That is, each individual data point is meaningful on its own. This allows a data consumer to query for multiple streams of data without needing to first query for survey or campaign specific information. Finally, the format is extensible to all types of metadata using JSON objects.

### Trigger Framework

The reminder framework in AndWellness supports reminders that are based on time, activity, or location, and can be extended to other modalities. This reminder service provides programatic and graphical based authoring and configuration of reminders: The main design goal for this framework is to provide a uniform experience in creating and managing triggers of any supported type.

### Buttons as Lightweight Survey Abstractions.

Participants in our focus groups frequently expressed the concern that surveys would be difficult to answer, and requested a lightweight method to record a specific event quickly. This notion of a desktop 'button' has evolved, and has become especially important to capture events that are unplanned, such as stressful events, or are not on a regular schedule, such as eating. We are exploring two buttons for deployment. First, a stress button, which records time and location when pressed. Second, a food button, which launches a short survey including the option to take a picture.

### *Evaluation*

Main metrics for evaluation are battery usage and accuracy. We have outlined a battery testing methodology to address hurdles presented by using different phones with different batteries, usage scenarios, and configuration when comparing energy usage. Validity and accuracy evaluated through a number of behavioral health studies, discussed in the next AndWellness section.

### **Future Directions**

In the coming year, we plan to complete the design and implementation of the AndWellness system, and release the code as open source. This release will enable any researcher, with minimal technical capabilities on staff, to install, author, configure, and initiate an EMA campaign using Android smartphones. The AndWellness server application will additionally be serving as the back-end for a number of the major smartphone and participatory sensing efforts undertaken at CENS, including Mobilize and What's Invasive.

## PART 14 AndWellness II: 3 Pilots to assess diet, stress, exercise, energy, mood, and sleep in diverse populations

---

- Deborah Estrin, Faculty\*
- Hossein Falaki, Graduate Student
- John Hicks, Graduate Student
- John Jenkins, Graduate Student
- Brent Longstaff, Graduate Student
- Mohamad Monibi, Staff
- Kannan Parameswaran, Graduate Student
- Nithya Ramanathan, Faculty\*
- Josh Selsky, Staff

\* Primary Contact

### Overview

We have deployed AndWellness with a number of pilot studies to begin to evaluate how well AndWellness meets its design goals of engaging participants while providing valid and reliable data. These pilot studies include several internal technical pilots to evaluate validity and reliability of the automated exercise monitoring, research study of breast cancer survivors, and a pilot study with young mothers to evaluate the validity and reliability of the survey and automated data collection using AndWellness.

### Approach: Construction of Study

In a research study, AndWellness facilitates the collection of survey data collected by prompting the participant to answer questions on the phone, and actigraphy data which is continuously collected without human input by sampling the onboard GPS and accelerometer devices onboard the phone to infer the participant's current activity. All responses are time-stamped and geo-tagged, and uploaded to a server using a delay-tolerant reliable transport protocol when network access is available. Surveys can have one of five types of questions, and surveys can branch to different questions based on a participant's response. Finally, participants can be prompted to complete surveys using the ANdWellness reminder framework. Reminders are based on time, location, and/or sensor readings. Each individual participant can modify the trigger settings on their mobile device. When a trigger fires, the Android notification system rings or vibrates and displays a small message to the participant. The flexibility of the trigger system along with personal trigger customization reminds participants to collect data while conforming to the participant's daily schedule. Using our trigger framework, a number of parameters can be configured using the GUI, including the time and location of trigger, and the number and duration of notifications.

### System(s) Description and/or Experiments

We have three ongoing studies that are currently utilizing AndWellness.

#### *Cardiovascular disease risk factors and behaviors in young moms*

Behavioral researchers at UCLA are using AndWellness to study cardiovascular risk factors in young mothers. Basic measurement parameters included a participant's daily exercise routines, their diet, and their stress and mood levels throughout the day measured by survey; and a participant's mobility recorded continuously using our automated mobility classifier. In-person assessments will occur at baseline, 3, and 6 months, and consist of retrospective self-reports and minimally invasive biomarkers to evaluate validity and reliability of phone measures.

Randomization will occur in two stages. First Moms will be randomized to a self-monitoring condition where participants carry smartphones running ANDWellness (n=45) or a control condition with no self-monitoring and no phones (n=15). Moms in the self-monitoring condition will be randomized to receive probes in one of three schedules (n=15 fixed time schedule, 15 random, 15 location-based), to occur 4 times a day. Probes will remind participants to complete a survey on diet, stress, or exercise. Additionally participants will have the option of recording diet and stress events using the phone's desktop buttons. ANDWellness will assess compliance to the survey reminders. The strategy for evaluating the reliability and validity will take place over two 3 month periods. In the first 3 months of carrying mobile phones, the phone data collection will establish the base rate for all four assessed behaviors. During the second 3 month period, ANDWellness will provide real-time feedback comparing Mom's current behavior with base-rates established in the first 3 month period for all assessed behaviors and compliance to the phone's reminders. The primary aims of the study are to first, examine the validity and reliability of phone data on assessed behaviors, through comparison with biomarkers and retrospective surveys; and second to preliminarily examine trends and correlations in assessed behaviors and compliance to the phone associated with the type of probe schedule (fixed, random, and location-based), and real-time feedback.

We have 15 participants that have taken part in a 1-week pilot version of this study, and have over 90% compliance in survey data collection. In this 1-week pilot, participants are additionally asked to complete a 30 minute scripted activity involving sitting, walking, and running for 10 minutes each, in order to evaluate the validity and reliability of the mobility classifier. Current results for two participants of over 90% accuracy are promising, but more data is needed before conclusions can be drawn.

#### *SleepSens*

SleepSens is a study of participant's sleep habits that asks participants to answer an eight-question survey when they first wake up in the morning. It also uses a sleep detection algorithm that approximates when a person goes to sleep and when a person wakes up by detecting phone movements. It uses this wake detection to send a reminder to the participant to take the survey. SleepSens is able to approximate a person's sleep patterns through its internal sleep and wake detection components, by continuously sampling the accelerometer onboard the phone. Once it is approximated that the user is awake, SleepSens will put itself to sleep until a hard-coded time later in the day. This prevents SleepSens from consuming unnecessary battery power. We have approximately 20 participants all of which have more than 90% compliance with filling out the daily survey.

#### *ABC*

AndWellness is currently employed in a study to measure the mood and energy in young breast cancer survivors. The study includes four surveys triggered throughout the day: The sleep survey is triggered in the morning and queries about the previous night's sleep and prompts participants to measure their stress levels by swabbing their cheek with a q-tip and saving the sample; the mood survey is triggered mid-morning, mid-afternoon, and in the evening, and includes 10 mood and stress prompts; the exercise and substance use prompts are also grouped in the evening survey. We have approximately 20 participants, with more than 90% compliance with the survey data collection.

#### **Accomplishments**

We have conducted focus groups with potential participants, and interview and pilots with participants, receiving overwhelmingly positive feedback about the design and implementation of AndWellness:

- Focus groups with over 60 diverse potential users of a mobile self monitoring system like AndWellness
- Three pilots ongoing with over 50 participants, studying diet, stress, exercise, mood, energy, and sleep.
- Interviews and surveys with over 30 participants that have used AndWellness.

Almost every participant asked in our interviews has said that that they would find AndWellness useful in their personal behavior management; most participants appreciated the reminders and configurability provided by the trigger framework.

We have ongoing collaborations with several Universities outside of UCLA (Arizona State University, Northwestern Medical University, and UC San Francisco), and numerous requests from people domestically and internationally to use the AndWellness platform.

#### **Future Directions**

In the coming year, our three ongoing pilots are expected to complete, and at least three more planned pilots will be completed. The AndWellness software platform will be used as the basis for the Mobilize project in collaboration with the Exploring Computer Science curriculum at local high schools in Los Angeles Unified School District, therefore development resources will be focused on ensuring that AndWellness meets requirements of both Mobilize as well as ongoing and future behavioral studies. Finally, AndWellness will be released as open source software.

## PART 15 Mobilize Project

- Deborah Estrin, Faculty, PI\*
- Ian Fellows, Staff
- Neal Fultz, Staff
- Valerie Hajdik, Staff
- Mark Hansen, Faculty, PI
- Jeroen Ooms, Staff
- Hongsuda Tangmunarunkit, Staff\*

\* Primary Contact

### Overview

The vision of Participatory Sensing is of distributed data collection and analysis spanning the personal, neighborhood, and community scales. In Mobilize, student learners design, deploy, and analyze the results of data campaigns; while their fellow students act as data campaign participants, making key decisions about what, where and when to make observations using mobile devices to support the subject of the campaign. The Mobilize software system is designed to engage students in an integrated and iterative learning experience; one in which they choose the topic of the campaign, design the details of the experiment and participant experience, and most importantly, develop their computational thinking skills as they explore the datasets generated by their campaigns.

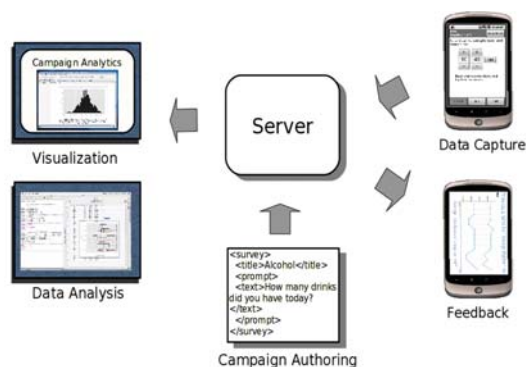


Figure 1. Mobilize software functionality

### Approach

A data campaign is a coordinated data gathering exercise. A campaign consists of one or more surveys, each of which is a collection of prompted inputs, or questions. The answers to these questions provide insight into the subject of the campaign, organizing the capture of “facts” and personal observations from participants. To gather both widespread and timely responses, Mobilize distributes a custom mobile phone application for participants, ensuring that all responses are automatically timestamped, geocoded and uploaded for analysis and visualization.

The Mobilize software system supports expressive campaign “authoring,” mobile phone-based data capture and realtime feedback, privacy respecting data management, and extensible data exploration.

#### *Campaign Authoring*

A campaign author creates a configuration file (using the Mobilize XML specification) that describes the surveys, the survey content and the conditions around which they are delivered to participants. A campaign consists of one or more surveys. Each survey comprises of a set of questions/actions called prompts. Properties of individual prompts and their ordering are specified in the configuration file. After the campaign configuration is completed, it is then uploaded to the server and validated. A web application will allow students to publish their campaigns for use by other students and to manage active campaigns by displaying various participation statistics. These displays are themselves authorable by the students as they share a common language with the data analysis tool we will describe below.

#### *Data Capture*

A campaign participant downloads the Mobilize mobile application to their Android smartphone (in the future and contingent on available resources, additional smartphone types may be supported) and then downloads the XML for the particular campaign(s) in which they will participate. Participants can activate the application anytime to answer and submit survey questions. In addition to manual survey activation, a trigger mechanism is also provided. A trigger automatically launched a survey based on some certain criteria such as time and location. Mobilize currently supports relatively simple triggers, but these will be extended in future versions of the system. One specific class of triggers are exposed in the mobile phone application and allow participants to adjust their timing or frequency.

#### *Feedback*

Mobilize accommodates various forms of participant feedback. At a basic level these mechanisms are designed to sustain user engagement and participation. Our feedback schemes draw on data analysis metrics such as ranking among peers, campaign adherence, and progress towards goals. The first version of the system will have a very simple form of user feedback, however this function will be significantly enhanced in the second year of development.

### *Data management*

The Mobilize data server maintains the data uploaded for each campaign in an SQL database. The database schema for a campaign is defined by the XML that is authored by the campaign owner, a student or teacher. Within a campaign, data are associated with the particular user that captured them. Moreover, when initially uploaded a participant's data are not made accessible for a configurable period of time until the individual releases the data for sharing. This allows participants, who we consider to be the owners of their data, to decide to delete data that they do not want to share.

### *Data exploration*

While the campaign is running and after the campaign participants have completed their activities, the student learner campaign authors will be able to download campaign data and will be provided with a rich, but scaffolded, set of tools with which to explore the data. During the course of the campaign, high level participation statistics will be generated automatically to support campaign management and engagement (such as number of responses, spatial coverage of the current responses, time-series analysis of data contribution levels, etc.), but all detailed exploration of data will be created by the student learners themselves, rather than be pre-scripted by the system. This design is essential to the pedagogical aims of Mobilize and will be explained further in the discussion of the curriculum.

## **Systems Description**

The system consists of the following three components: Server, Web Interface, Mobile phone, and Data analysis support.

### *Server*

The server component provides user authentication and authorization; supports upload, validation and storage of campaign configurations and survey responses and images; provides APIs for retrieving campaign configuration, survey responses and data points; and supports aggregate campaign data for feedback, data analysis and visualization.

The standard communication protocols, HTTP and TLS/SSL, are used to securely transport data between components. All text-based data (e.g., survey responses) is encoded in JSON format and send over HTTPS as content message. An image is sent as binary content over the HTTPS protocol. The server is designed to be as stateless as possible, so authentication information must be sent with each call to the server.

Two different authentication schemes are designed to authenticate clients. Thin-client authentication is designed for web browser clients. After a web user is authenticated using password, a temporary token is generated and sent to the browser to be stored as a secured cookie for later authentication. This token is also stored locally on the server. After tokens expire, log-in is required to generate new tokens for accessing the server. Thick-client authentication is designed for a thick client (e.g. phone application) that can keep local state. In this mechanism, a user's password is securely stored locally on the phone. The same user's credential is also created and stored on the server. This allows users to logon to a phone to interact with the application offline. The campaign configuration file is validated against the database schema. The campaign configuration is then used to validate incoming data uploads (e.g. survey responses) from the phone application.

These server functions are implemented in Java 6. They running on top of Tomcat 6.0, MySQL 5.1 database, and CentOS operating system.

### *Mobile Phone Application*

The mobile phone application supports loading of campaign configurations; trigger functionality, including a GUI for users to adjust their trigger configuration such as changing trigger times, and the generation of notifications to remind participants to take actions; survey prompts based on menus, free text, and image input; buttons that allow users to initiate short surveys; mobility data capture in the background based on sampling of GPS, wifi, and accelerometer devices in the phone; and the upload of the captured data to the server. It also supports user login on the phone so that phones can be shared among data campaign participants.

The Android phone application is designed to read a campaign configuration and render appropriate prompt display according to the configuration. The campaign schema is shared between the phone and the server. A prompt type and properties associated with each prompt need to be specified in the configuration. This information tells the phone application on how to render and handle each prompt. Seven prompt types are currently supported.

Campaign configuration is an XML file. It contains properties and description of a campaign. A campaign configuration contains one or more surveys. A survey can contain prompts and repeatable sets of prompts with a rendering order defined by their top-down appearance within a survey and in conjunction with conditions. A prompt corresponds to a question or an action in a survey. Prompts should be designed to reflect facts, status, condition or observation of the object and its environment focused in the campaign study. A prompt configuration consists of 4 parts: prompt

metadata (e.g., ID), web-based visualization properties (e.g., display type), phone-based prompt display properties, and branching condition (optional).

The phone application is implemented in Java and runs on the Android platform.

#### *Data access, analysis and visualization*

This component provides the key functions for student learners to access the data, and to support their development and exploration of computational thinking and analysis skills. It supports secure login, campaign management, data management, and basic campaign monitoring. Through a web interface authorized users download campaign data for local data analysis and insert additional web-based data visualizations that are created through their analysis. The analysis document can be uploaded to the system to be shared to other students.

The local analysis component consists of R and Deducer. R provides software environment for statistical computing, data analysis and graphics. On top of R, Deducer is used to provide an intuitive graphical user interface for students to perform basic data manipulation and analysis tasks without programming getting in their way. To fully support the Mobilize curriculum, Deducer will be extended and customized e.g., basic geo analysis, text mining and image analysis will be added.

Campaign analytics and data visualization aimed to help campaign authors monitor and run their campaigns are done at the server. We are using RApache, an Apache web server with embedded R interpreter, to perform actual statistical analysis and graphics. An RApache web application will handle campaign analysis requests through its interactions with the Mobilize server.

#### **Accomplishments**

Describe specific accomplishments (e.g., findings, deployments, system performance) during the reporting period.

#### **Future Directions**

Mobilize 1.0 will be released in June 2011 and used during our Professional Development and Summer Student Internship program during the summer of 2011. The initial version will have all of the functionality described, but in relatively simple forms. During the summer and fall of 2011 we will use our pilot usage experience to prioritize the expansion and deepening of features. Likely areas of focus for version 2 include: supporting richer forms of feedback and game mechanics in PS data campaigns, extending the support for continuous data streams such as activity (from GPS, wifi, and accelerometers on the phone), and communication (from audio and application statistics gathered on the phone), and adding tie in to social media applications. By pursuing our implementation through multiple cycles of iteration and use, we will prioritize our development to meet the needs of the Mobilize classrooms and learning objectives.

While the vision for Mobilize is that students will create their own campaigns, focused on issues of concern to them in their school and neighborhoods, we will seed this process by providing example canonical applications. These canonical applications serve three important purposes. First, it gives us concrete end to end tests of our functionality and performance during the months of software development, testing and integration. Second, it provides concrete areas in which to gather background data sets and organize specific lessons for the first year of the new curricular unit in ECS. And third, it provides a concrete experience for students and teachers to use the technology of Participatory Sensing, and based on this experience to begin to generate ideas for new and different campaigns.

We define more than one canonical application in order to test and demonstrate different aspects of the PS authoring, data capture and analysis system; and to provide a range of examples to students and teachers in order to communicate the breadth of potential applicability.

## PART 16 Boyle Heights Planning for Place: Mobile Participatory Sensing Project

---

- Amelia Acker, Graduate Student\*
- Deborah Estrin, Faculty, PI\*
- Martin Lukac, Staff

\* Primary Contact

### Overview

Participatory Sensing is a process whereby individuals and communities can use ever-more-capable mobile phones and web services to observe, analyze, and respond to the personal and environmental factors that impact community health. 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 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. In 2010, we partnered with community organizers in the neighborhood of Boyle Heights in East Los Angeles.

### Approach

In May of 2010 the Building a Healthy Boyle Heights Collaborative (BHBH) launched its planning phase for The California Endowment's Building Healthy Communities Initiative (TCE – BHCI). Over 30 organizations and 5,000 community residents participated in the ten month planning process. CENS partnered with the BHBH collaborative to develop mobile sensing application surveys to document environmental conditions and community assets. Community residents collected local data to inform the sequencing of TCE-BHCI outcomes and the development of the overall community logic model. Four community based organizations (CBOs) of the BHBH Collaborative, including East LA Community Corporation, InnerCity Struggle, Union de Vecinos, Violence Intervention Program.

Community residents participated in a community-wide mobile data campaign with participatory sensing applications. The data collection was grounded in community residents' and organizations' experiences of their personal environments. To complement the data generated through town hall discussions, working group meetings, and other in-person meetings, the BHBH Community Engagement Community developed multiple surveys to collect community level data (home, neighborhood, work, school, route). The mobile surveys prompted users to give semi-structured responses; participants could submit a variety of multiple-choice answers, phone photograph submissions, label photographs, free text responses, and GPS coordinates. The five surveys encouraged participants to document and reflect upon their living and working conditions; meals and food preparation; transportation and commuting patterns; sense of community and support; and experiences of public and private spaces in their neighborhood.

The goals of the Boyle Heights data campaign was to facilitate community engagement by enabling residents to document their neighborhood's strengths, needs and activities; and to encourage the distribution and decentralization of collected data so that participants can access, analyze, and make observations about the collected 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.

### Project Deployment and response

#### *System Description*

CENS developed mobile five applications, each acting as a mini-survey that the BHBH team had designed:

- BH-Ruta traced the participants path and asked questions about transportation (e.g. walking, bus, private car).
- BH-Escuela asked about school conditions.
- BH-Trabajo asked participants to survey work conditions.
- BH Casa asked about home conditions.
- BH-Vecindario asked for observations about the neighborhood.

The project team acquired 30 Android phones from Verizon wireless for the project and developed a basic website interface to access the collected data in real-time. We also prepared documentation with screenshots, detailed instructions, and consent forms for participants during the deployment and data collection.

### *Survey Response*

There were a total of 68 participants who filled out 462 surveys total. The BH-Ruta app was the most popular survey across all five apps with a 75% participation rate (51 users), whereas the BH-Trabajo 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. The BH-Vecindario 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.

Each application 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. 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.

### *Resident reactions*

After the data had been collected, the community based organizers carried out follow-up interviews and met with the participants to share the data and discuss their experience. The participants were asked to give their overall perceptions of having participated in the process. Participants talked about their experiences using the phones, walking through their community with the phones, explaining to other people what they were doing, what they would have done differently and what they learned. The campaign data was shared and participants were able to analyze the data in their own words, ask questions about different patterns, make connections, raise more questions, offer reactions and suggestions for future work.

### **Accomplishments**

The response rate and findings of the Boyle Heights data campaign provides evidence that participatory sensing can be successfully applied to community data campaigns. According to a white paper authored by our BHBH collaborative partners, the mobile sensing project was a helpful tool for community engagement in involving community members in deeper dialogue and discussion on issues related to healthy homes and neighborhoods in Boyle Heights. Our collaborative partners understood and have emphasized that the partnership with CENS to be equitable, which is essential in participatory based research with collaborations between community organizers and university researchers.

### **Future Directions**

Putting sensors into people's hands can enable participants to understand the significance of their individual data sensing contributions at personal and community levels. 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.

In the next year we plan to continue our partnership with the BHBH collaborative by reiterating and improving our approach to community data campaigns. Future participatory sensing campaigns will incorporate pretesting, participant feedback (during deployment), and real time campaign management and analytics through a web-based portal.

## PART 17 Remapping LA: Cultural Civic Computing in Los Angeles

- Manu Basal, Graduate Student
- Jeff Burke, Faculty, PI\*
- James Dellemonico, Staff
- Taylor Fitz-Gibbon, Staff
- Diana Ford, Undergrad
- Fabian Wagmister, Faculty, PI

\* Primary Contact

### Overview

Since the creation of its participatory sensing area (previously urban sensing), CENS has collaborated with the UCLA Center for Research in Engineering, Media and Performance (REMAP) on its “civic cultural computing” project Remapping LA, a collaboration with California State Parks and others to contribute to the renewal of Downtown L.A. by enabling communities to power the processes of imagining, specifying, and designing technology for their neighborhoods and public spaces. The project focuses on the 32-acre Los Angeles State Historic Park (LASHP), the first urban state park in California. Activities having included (1) Technical assistance for UCLA courses on “engaged media production,” in which students produce media and collect data using mobile, web, and other tools—including participatory sensing systems—that explore Los Angeles neighborhoods; (2) community workshops and participatory media art and technology projects with a variety of groups, from the youth Anahuak Soccer Federation to the Gay and Lesbian Elders Housing; (3) experimental media artworks on the history and geography of Los Angeles. Each of these projects leverages participatory sensing technologies and concepts as documented in this and past annual reports. Collaboration with REMAP provides a cultural perspective on the use of mobile technologies, and has generated its own unique systems and applications research.



Figure 1. Map of parcels near the LASHP.

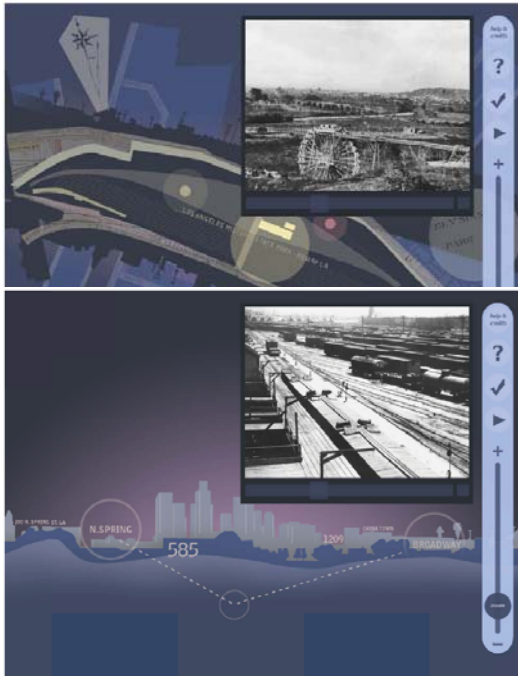
In 2010, for the first time and despite the economic crisis, the California Department of Parks and Recreation committed over \$100k to the innovative research, teaching, and public programs of Remapping LA. This Interagency Agreement will support (1) the launch of a community media web portal where contemporary and historical media about Los Angeles can be contributed and viewed, (2) continued development and operation of research technology infrastructure that includes the UCLA/Cisco WiFi network operated by REMAP and CENS, and (3) the collaborative creation of public digital media experiences for the Park by student teams enrolled in Spring and Summer courses. The results will contribute to construction and interpretive planning for the Park as well as the broader state-of-the-art in hybrid physical and digital spaces and participatory media systems.

### Approach

This year, Remapping LA involved a combination of multidisciplinary coursework, community engagement, and software development. Students from a variety of departments participated in the Engaged Media Production course, which developed four place-based media projects for the Park. The LASHP Users project investigated and document park use: visit duration, time of day, frequency, and purpose; paths take to the park and modes of transportation. The Home as History project used local homes as a point of departure for exploring neighborhood history. Ghost stories mapped local folklore in the areas adjacent to the park, and the Natural Environment project focused on documentation of the ecology of the area. These projects influenced the research team’s focus for the year. In particular, the Home as History project inspired increased, systematic documentation of the neighborhood, including creation of GIS layers that tie local parcels to historical and cultural context. (See Figure 1.) This data, along with historical media, is being used in a new tablet application, Hilos, that allows park visitors to explore media about the park and surrounding neighborhoods through GPS and compass-based interfaces to location and orientation, respectively.

### System(s) Description and/or Experiments

An alpha version of the Hilos application was developed for the Apple iPad, though the first public version will likely be ported to Android tablets. The application is a “thin client” to a remote database of geotagged and keyworded media. (See Figures 2 & 3 for interface examples.) Categories are selected by walking to a particular location in the



Figures 2 & 3. Hilos location-based tablet application.

Park associated with a historical event. Additional media may then be loaded by lifting the tablet up to the horizon and looking around. A graphic representation of the city skyline is aligned with the visitor's view using the tablet's compass sensor, and touching buildings of interest provides access to additional media. A further view is planned that explores connections between the particular parcel, the park, and the rest of Los Angeles. The system will gather usage data and other metrics to further inform the design process.

The Hilos application is complemented by an SMS interface developed by NexLeaf Analytics, a non-profit formed by CENS researchers. The RapidSMS-based software tool will enable easy authorship of text messaging based interaction with interpretive content about the Park.

In addition to location-based media at the Los Angeles State Historic Park, REMAP and its collaborators also just received a MacArthur Digital Media & Learning award to bring youth media to 2200 Los Angeles Metro buses with over 7 million riders per week. While this project, Out the Window, <http://www.out-the-window.org/> starts with video loops, it hopes to eventually take advantage of the embedded systems on the buses that have GPS and WiFi capability, to play back location-based media authored by the youth.

### Accomplishments

Remapping LA secured California State funding for the first time, developed a new generation of location-based media applications based on the experience from previous years, and expanded its geotagged media documentation of the area around the Los Angeles State Historic Park. Out the Window further expanded the team's efforts to incorporate the public transit system in Los Angeles. The work has been presented in a number of local, national, and international talks and dialogues.

### Future Directions

In the Summer of 2011, the first public deployment of the tablet application will take place at the Los Angeles State Historic Park, and video, location-based banner images, and text message interaction in Out the Window will run for at least 30 days across the Metro systems. These deployments will inform the next stage of research into location-based media, and provide further opportunities for connection to the data collection efforts that other CENS projects are involved in.

## PART 18 Environmental Participatory Sensing

- Jameel Al-Aziz, Undergrad
- Eric Graham, Staff, PI\*
- Kyung Han, Graduate Student
- Jinha Kang, Graduate Student
- Dylan Vassallo, Undergrad
- Eric Yuen, Staff

\* Primary Contact

### Overview

Citizen science programs have collected information at local, regional, and continental scales to help answer diverse questions in the environmental sciences. As a result, hundreds of thousands of non-scientists have contributed to scientific research projects in recent years. We have the opportunity to greatly expand participation in such projects that engage the public in data collection by using digital technology and mobile phones.

Ultimately, mobile technologies have the capacity to greatly improve environmental literacy by supporting educational resources to broaden participation in a national citizen science field campaign that investigates the connection between climate and the living environment. To this end, CENS researchers have been working on two main environmental participatory citizen science projects that are leveraging digital technology and mobile connectivity to increase data quality, the recruitment, and the retention of participants..

What's Invasive! (WI; <http://whatsinvasive.com>) is a CENS citizen science campaign for locating invasive weeds. Any National Park Service official (or other authorized person) can create a park-specific weed list with descriptions and photos uploaded to our server. All Android and iPhones running the WI application will be able to download this list and the phone is ready to help in identifying and locating these species. All the data collected by the users of the phones are uploaded to the server and are immediately available to anyone for downloading as a CSV file.

Project BudBurst (BudBurst) is a national field campaign for citizen scientists designed to engage the public in the collection of important climate change data based on the timing of leafing and flowering of trees and flowers. It is run by the National Environmental Observatory Network (NEON), with whom CENS is collaborating Observations of plants made by an Android phone running the BudBurst Mobile app are sent to a CENS database and visualization is done on a CENS-hosted web page.

### Approach

Both projects have been worked on for multiple years by CENS and are based on mobile data collection, with WI having both an iPhone and an Android client whereas BudBurst currently only has an Android client. Work is focused on expanding both systems beyond their previous capabilities, with innovation occurring in terms of primarily sending location-specific information to the handsets and increasing location determining GPS capabilities through maps and updated accuracy information.

### System(s) Description and/or Experiments

#### *What's Invasive!*

The What's Invasive "experience" starts after the user has been authenticated. The application then determines which weed list is located closest to the user and automatically enables it as their default park. The user can then use the application to view a list of the top invasive plants and/or animals for the park. Should the user see an invasive plant/animal that is on the list, they can capture a geo-tagged photo and/or note observation. The user is also allowed to choose "Other" if they would like to upload data on an item that is not on the list. The observations are then uploaded to the What's Invasive server, where both the user and park staff can see all observations displayed and download them.



Figure 1. The two primary Environmental Participatory Sensing, Android-based citizen science projects: BudBurst Mobile and What's Invasive. An iPhone app of What's Invasive is not shown.

### *BudBurst Mobile*

The Project BudBurst android application has a new improved interface and several additional features. The purpose of the interface change is to provide sleek and user-friendly screens for easy navigation and to easily make observations. We also provide as much information as possible in the simplest way, for example, displaying the species or phenophase information by simply clicking the photos or showing multiple sets of information, such as species and phenophase information, a photo taken, date and notes in one screen in an intuitive way.

There are six main categories in the application: My Plant, Quick Capture, Local Lists, Plant Maps, Plant News and Floracaching. Basically, users can manage their plants and observations taken from My Plant, Capture and Local Lists. The application provides a variety of species lists for users to add different kinds of species from budburst, invasives, natives, blooming, tree lists by querying our local database as well as from the USDA. We have also made geolocated observations more accurate by combining the map feature enabling users to actually see their location and adjust their location correctly on the map with the GPS system.

### **Accomplishments**

#### *What's Invasive!*

The iPhone version of What's Invasive is now available on the App Store. Additionally, we added the following features to the system:

- Animal Observations, in parallel to the plant observations.
- Notes, as a new field of data uploaded.
- Improved UI, updated to be more Androidian.
- Default photo module, allows larger photos and retake possibilities.
- GPS Accuracy, added to the latitude and longitude data.
- Debug mode, added to allow users to report bugs and for the developers to receive better quality data.

#### *BudBurst Mobile*

Added scrape engine from USDA and the Quick Capture concept. Currently, the app is being used to map the trees on the UCLA campus.

### **Future Directions**

Mobile phone based tools have the potential to take citizen science to a new level, by combining the convenience of immediate data entry and photo sharing using targeted apps with field campaigns focused on environmental issues. We are partnering with a growing number of environmental groups that will be using the apps that we are developing. The result is a new type of connected citizen scientist – one who collects scientifically relevant data as part of his or her daily routine. We expect these efforts will revolutionize the way citizen science participants are recruited and retained.

Efforts to make mobile data collection even more engaging and to retain participants are resulting in our collaboration with the Game Lab at UCLA. The fusion of games with data collection, combined with online tools to allow discovery of patterns in data that are important to the individuals who collect the data, will reach out to a more diverse community of citizen scientists who represent a greater cross-section of society with different and varied motivations for participating in research.

## PART 19 Ethics in Personal Mobile & Participatory Sensing

- Jeffrey Burke, Faculty, PI\*
- Deborah Estrin, Faculty, PI
- Mark Hansen, Faculty, PI
- Katie Shilton, Graduate Student\*

\* Primary Contact

### Overview

The mobile phone network is emerging as the largest sensor network on the planet. Mobile phone users, however, are generally unaware of the dual uses of this network, in which their communication devices are also information gathering devices. What are the ethics of coordinating this network for research purposes? Can researchers achieve meaningful consent and active participation of mobile phone users? The three-year Ethics in Personal Mobile & Participatory Sensing research and education project has:

- Conceptualized a participatory approach to managing privacy in personal mobile sensing applications;
- Created curricula to teach participatory ethics to diverse STEM undergraduate and graduate students;
- Investigated design changes to encourage ethics in urban sensing, ubiquitous computing, and broader technology education communities.

This project achieved these goals by focusing on anti-surveillance values: ethics like privacy, consent, participation, equity, and forgetting, which help counter the social consequences of surveillance. Anti-surveillance values are in many ways more difficult to incorporate into design than values such as efficiency or productiveness. They are more difficult to quantify than efficiency, which can be estimated and measured. They are more difficult to define, stemming from moral and social philosophies rather than market demands. This project asked: What values are important to the design of mobile and participatory sensing systems? How do anti-surveillance values affect technology development? How do design setting and design participants influence anti-surveillance values? And how can designers as well as social scientists make anti-surveillance values an integral part of design?

### Methods

To analyze and understand anti-surveillance values in participatory sensing design, the lead graduate student pursued a qualitative research project employing ethnography and action research. The author spent three years embedded at CENS, an engineering laboratory developing ubiquitous sensing technologies. In this role, she intervened in projects, highlighting social values during design conversations. She also observed the effects of mentors and collaborators, interactions with clients, designers testing their own software, and institutional ethical mandates as implemented by UCLA's Internal Review Board (IRB). She conducted interviews with all students and staff involved with participatory sensing projects (n=30). She coded the interviews to find values that compete in the design of sensing infrastructures, challenges to anti-surveillance values, and techniques to address those challenges.



Figure 1. People and practices that deploy values levers in CENS design

### **Findings: Values Levers in Design**

Designers at CENS take anti-surveillance values seriously, but they face significant competing values such as efficiency, innovation and productivity. There are technical limitations on the projects and system features that designers can pursue. Students face stringent deadlines and pressure to publish their ideas quickly. The constant pressures of technical innovation combine to make a values-oriented design process difficult.

Nevertheless, anti-surveillance values were expressed and acted on by a variety of both practices and people at CENS. These agents were doing a particular kind of work in design: creating space for new conversations about ethics and values. They demonstrated values levers: practices which pry openings for discussion of anti-surveillance values during design. A wide variety of values levers were present during design at CENS. They included a diverse set of both practices and people within design. Among the most important were:

- Practices such as interdisciplinary design and self-testing which encouraged designers to focus on sensitive data
- Creativity that arose from working within values constraints
- The values reactions and feedback of users
- Values awareness procedures and advocacy by leadership
- Values awareness procedures implemented according to institutional rules
- Advocacy by a design team member dedicated to values work, and
- Funding for interdisciplinary teams and long-term development resources

The network of values levers in design led to the major research finding: that the structure of a design environment can encourage social values as design criteria. Actors and design practices at CENS used values levers to create space for conversations about, and action on, anti-surveillance values. Three design structures in particular: self-testing, interdisciplinary work, and a team member dedicated to values issues, proved quite effective at generating conversations about, and technological features based on, social impacts and social values. Two others, including the oversight of UCLA's Internal Review Board and user-designer feedback loops, hold promise, but need adjustment to be truly effective values levers.

### **Accomplishments**

These findings contributed to the design of lab-based approaches to ethics education, as well as a seminar curriculum. The lab-based approach focused on institutionalizing ethical decision-making at CENS through attention to values levers. For example, we implemented a process by which all participatory sensing pilot projects are evaluated by laboratory leadership for ethical and human subjects concerns. We also engaged with UCLA's IRB to address new challenges to privacy and consent posed by participatory sensing projects. We also created a seminar course entitled "Mobile Technologies: Participation and Surveillance," which was taught and evaluated during the spring of 2010.

The project team also created new technologies based on the anti-surveillance values examined in this project. Our research team outlined requirements for the Personal Data Vault (PDV), a secure repository designed to help participatory sensing participants manage their data. We also collaborated with Jerry Kang, a professor in UCLA's School of Law, to determine legal protections for participatory sensing data held in a Personal Data Vault. More information on the PDV can be found elsewhere in this annual report.

We refined our research questions and curricular ideas, and presented our project results, at several national workshops and conferences. Katie Shilton served as a planning committee member for the Ethical Guidance for Research and Application of Pervasive and Autonomous Information Technology (PAIT) workshop organized by Indiana University in March 2010. At this workshop, case studies from CENS participatory sensing projects served as discussion prompts. Other workshops and conferences during the last year at which we have presented our research include the annual iConference (Seattle, February 2011), the Annual Meeting of the Association for Library and Information Science Education (ALISE) (San Diego, January 2011), the Annual Meeting of the American Society for Information Science & Technology (ASIS&T) (Pittsburgh, October 2010), the Privacy Law Scholars Conference (Washington DC, June 2010), the Values in Design Workshop 2010 (New York, May 2010), and Computers, Freedom & Privacy (CFP) (San Jose, June 2010).

### **Future Directions**

The project will end in the summer of 2011. Plans include the presentation of a dissertation based on this work, as well as the construction of a technical report with recommendations for applying these findings to ubiquitous computing design. We will also construct a website to disseminate our curricular and research results.

## PART 20 Courage & Creativity: The Innovation of Ethics in Science and Engineering

---

- Amelia Acker, Graduate Student
- Jeff Burke, Faculty, PI\*
- Katie Shilton, Graduate Student, PI

\* Primary Contact

### Overview

'Courage & Creativity: The Innovation of Ethics in Science and Engineering' (C&C) is a multi-media curriculum project that explores the generative role of ethics in scientific and technological innovation. An interdisciplinary team made up of researchers affiliated with CENS, UCLA School of Film and Television, the UCLA Department of Information Studies, and Eyes of the World Media Group is creating two short-form documentaries, curriculum supplements, and a web portal to encourage STEM students to thoughtfully integrate their personal values and ethical motivations into science and engineering careers. The project illustrates cases in which ethical challenges, and choices fostered innovation in science research and engineering.

Our approach to the Courage & Creativity project comes out of our experience building in-situ education approaches for considering ethics in the development of emerging technologies. Ethics education currently struggles to portray the complexities of ethics embedded within scientific practice. Often, ethics in science and engineering education are seen and taught as punitive, full of constraints and regulations. However, constraints (such as don't cheat, don't falsify data) are only one function that ethics have in science and engineering. Our experience working directly with engineers on a previous EESE project, Ethics in Personal Mobile & Participatory Sensing, demonstrated that another role for ethics in STEM education is a focus on the creative and generative force that social values play in contributing to the development of new knowledge.

### Approach

This project addresses the contradiction between ethically-driven constraints and innovation by building and disseminating new understandings of how ethical inquiry in science and engineering has been an essential catalyst for new discoveries. It contributes new research on the work of formal and informal ethical inquiry in generating innovation in science and engineering. Based on this research, we are developing an educational website with curricular materials and case studies illustrated using original video and multimedia content. This material, including interviews, location footage, and motion graphics, will bring an STS perspective to infrastructure for undergraduate and graduate STEM ethics training.

The project focuses upon social values such as equity, fairness, openness and sharing in both segments to illustrate how people's ethical frameworks influence their creativity in science and engineering. Students who engage with the video and multimedia content will be encouraged to critically evaluate dominant perspectives of ethics in science, and contribute to a deliberative critique of the disciplinary, cultural, and institutional forces that shape innovation in science.

### Deliverables

Two short documentaries will use interviews with innovators mixed with cultural imagery and location footage of the process of scientific creativity to demonstrate the concrete, but seldom discussed, links between personal and community values and creativity. The documentaries explore how what people believe influences their creativity and impacts the work they do. The stories re-envision ethics as a source of inspiration instead of a set of binding constraints.

In addition to the two documentary segments, larger sections of the source materials and curricular materials will be made available online through an interactive website to invite dialogue and further exploration on the topics discussed. The final products will not only be usable in classrooms, but engaging enough to disseminate virally on Facebook and YouTube, or be incorporated into longer documentaries or television and radio segments. We hope they will be widely shared, remixed, and refined upon.

### Framing and Case Studies

#### *Conceptual Framework*

Within each segment and curricula materials, the project strives to highlight social values such as equity, fairness, openness and sharing in both segments to illustrate how people's ethical frameworks influence their creativity in science and engineering. Students who engage with the video and multimedia content will be encouraged to critically

evaluate dominant perspectives of ethics in science, and contribute to a deliberative critique of the disciplinary, cultural, and institutional forces that shape innovation in science.

### *DTC Consumer Testing*

The first case study looks at recent innovations in personal genomics and direct-to-consumer genetic testing. It considers values of equity and accessibility embedded in the field by such diverse influences as the dark histories of eugenics, the imagined futures of science fiction; ethical debates about the role of humans in nature; legislative decisions such as the 2008 Genetic Information Non-Discrimination Act (2008).

It examines how values reflected in these events support entrepreneurship and transform new understandings of civil rights, personal identity and the law. The documentary in turn explores how these social influences foster personal perspectives and beliefs among scientists, engineers, and entrepreneurs, and the way these perspectives form an important foundation for the innovation ecosystem. By focusing on the individuals behind advances in genomics, we show viewers how and why our subjects do science and engineering, and how creativity from ethics influences their work.

### *Open Source Movements*

For the second documentary, we will explore the ethics of openness at the center of the open source software movement, and the ways this openness has changed not only technology development, but even fields as diverse as music, robotics, literature, and gaming. We trace the story of how ethos of openness and sharing fostered creative “hacks” of copyright law, and how these novel licenses for sharing influenced diverse open source movements. What began as a method to create better software took on moral dimensions as supporters adopted arguments for freedom, autonomy, and alternative understandings of intellectual property.

### **Current Work and Future Directions**

Currently, we are in the initial production development of the first documentary segment on personal genomics. In March of 2011 we begin subject outreach and research on the second segment that will feature the ethics of openness in creative commons licensing and open source software (OSS). Presently we are in the design and development stages of compiling additional curricular materials.

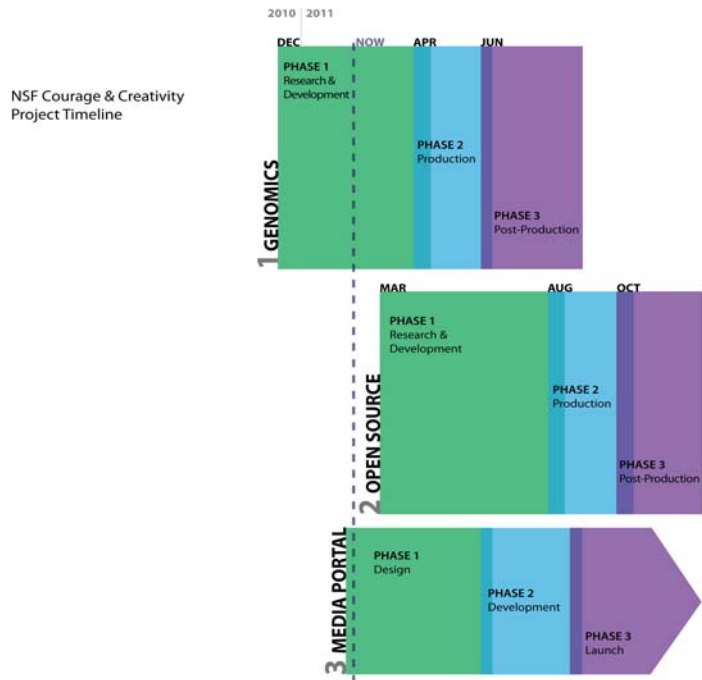


Figure 1. Timeline for Courage & Creativity project, 2010-2011.