ohmage: An Open Mobile System for Activity and Experience Sampling

John Hicks, Nithya Ramanathan, Hossein Falaki, Brent Longstaff, Kannan Parameswaran, Mohamad Monibi, Donnie H. Kim, Joshua Selsky, John Jenkins, Hongsuda Tangmunarunkit, Deborah Estrin
University of California, Los Angeles

Advances in technology and infrastructure have positioned mobile phones as a convenient platform for real-time assessment of an individuals health and behavior, while offering unprecedented accessibility and affordability to both the producers and the consumers of the data. In this paper we address several of the key challenges that arise in leveraging smartphones for health: designing the complex set of building blocks required for an end-to-end system, motivating participants to sustain engagement in long-lived data collection, and interpreting both the data and the quality of the data collected.

We present ohmage, a mobile to web platform that records, analyzes, and visualizes data from both prompted experience samples entered by the user, as well as continuous streams of data passively collected from sensors onboard the mobile device. In order to address the system design and participation motivation challenges, we have incorporated feedback from hundreds of behavioral and technology researchers, focus group participants, and end-users of the system in an iterative design process. ohmage additionally includes rich system and user analytics to instrument the act of participation itself and ultimately to contextualize and better understand the factors affecting the quality of collected data over time. We evaluate the usability and feasibility of ohmage using data from 3 studies with a variety of populations including young moms and recent breast cancer survivors. More than 85% of the diverse set of participants who responded to exit surveys claim they would use ohmage for further personal behavior discovery.

Categories and Subject Descriptors: H.1.2 [**Information Systems**]: User/Machine Systems – Human Factors; H.4 [**Information Systems Applications**]: Miscellaneous; H.5.2 [**User Interfaces**]: Evaluation/Methodology, User Centered Design; J.4 [**Social and Behavioral Sciences**]: Psychology

General Terms: Design, Human Factors, Experimentation

Additional Key Words and Phrases: Experience Sampling Method (ESM), In Situ Data Collection, Mobile Computing, Wireless Health Monitoring

1. INTRODUCTION

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 helping an individual patient reduce stress in their life, or by an epidemiologist studying a population's 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, valid, reliable, and feasible tools to self-monitor behaviors are required to address the limitations of current methodology. To avoid these sources of error and delay, a technique called ecological momentary assessment (EMA) was developed to monitor affect, cognitions, and behavior in real time in a person's natural environment [Shiffman et al. 2008]. EMA has been shown to greatly increase the validity and reliability of patient reported data: Data is collected in real time, therefore is not subject to recall and retrieval biases [Schwarz 2007]; data is harder to fake [Collins et al. 2003; Hufford 2007]; errors are identified early [Collins et al. 2003]; and data is instantaneously and automatically entered into a secure central database, minimizing human error.

However, EMA has traditionally relied on paper diaries to log events in real time, automated telephone systems to record data, and more recently PDA or wireless mobile devices to log 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 ohmage to enable rapid prototyping of end-to-end EMA systems. 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 ohmage, a researcher can ask a participant, for example, to monitor diet, 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, ohmage makes it possible to measure a participant's timely adherence to the data collection process and to configure when and why a participant is queried to collect data. Both are necessary to ensure high quality and unbiased data collection. ohmage aims to make it easy for doctors and researchers to conduct rapid experience sampling studies to study individuals or populations *in situ*.

ohmage includes four system mechanisms to facilitate rapid prototyping of EMA data collection. Specific contributions include: survey authoring including 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 snapshot of each users data.

The design of ohmage is driven by interviews with key behavioral researchers, focus group sessions with potential users of a mobile self monitoring system like ohmage, and surveys of breast cancer survivors who have used ohmage in a research study to self-monitor 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).

We have deployed ohmage in several pilot studies to evaluate how well ohmage meets these design goals. These pilot studies include: internal technical pilots to evaluate the validity and reliability of automated physical activity monitoring, a research study of breast cancer survivors (described above), and a pilot study with young mothers to evaluate the validity and reliability of the prompted survey and automated data collection using ohmage.

For each study we also present a set of expressive analytics developed to evaluate the success of a deployment in terms of: the engagement of the participants over time, the participants satisfaction rates with the study, and the quality of collected data over time. These analytics are designed to give researchers insight into the quality of collected data, both in real time and over the lifetime of a study. The metrics we present are derived from both an analysis of data collected from each participant, as well as an analysis of the participant's interactions with the smartphone itself.

2. SYSTEM OVERVIEW

ohmage employs a number of system components in concert to collect, return, and analyze participatory data. In this section we present a high level overview of ohmage from both the perspective of the researcher designing and running a study and the participant collecting and uploading data. We define the pieces that construct a study and how the system components work together to achieve the goals outlined in the introduction.

2.1 Motivating Example: Monitoring the Diet, Stress, and Exercise of Young Mothers

Researchers at UCLA wanted to design a study to monitor 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. To evaluate the validity and reliability of the measures through comparison with biomarkers, researchers wanted to collect the data over a six month period from a number of participants large enough to show statistical significance.

There are multiple approaches to collecting these data. One could interview participants weekly or monthly about their behaviors, ask the participants to keep a daily log either on paper or electronically, or ask the participants to complete daily online surveys. None of these leverage the power of modern day smartphones.

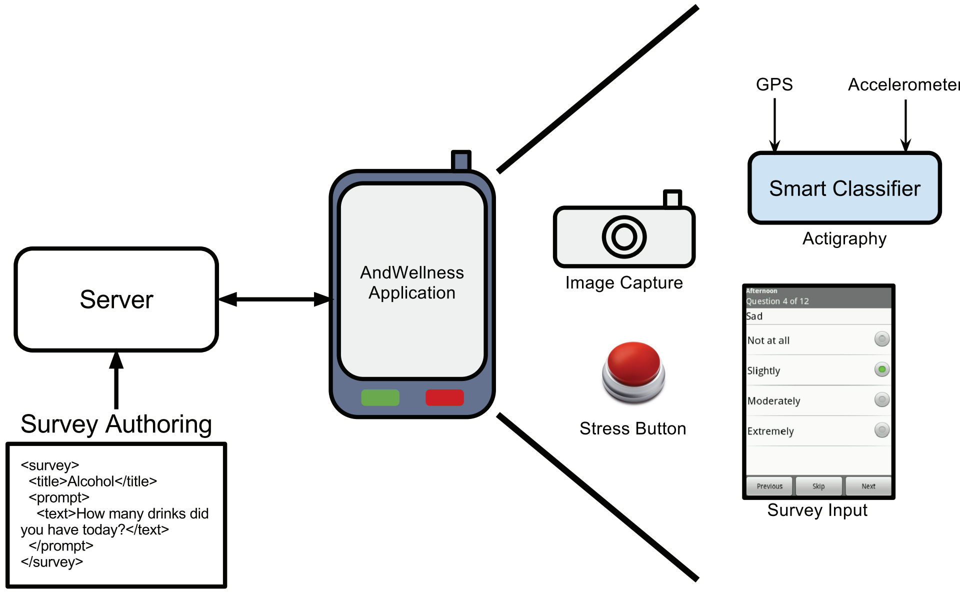


Fig. 1. Surveys are initiated by uploading the validated survey XML configuration to the server. Participants can then login to specific campaigns with the Android based ohmage application to automatically download all relevant surveys and immediately begin collecting data. Data collection includes survey responses, image captures, stress/food button pushes, and actigraphy or mobility traces.

We proposed the use of ohmage for this study. Surveys are launched at preconfigured times and location throughout the day, and are answered with a few taps on the device's touch screen. The surveys are either triggered at predefined times and locations throughout the day, or initiated in response to certain events. The responses are packaged up and returned over cellular or Wi-Fi channels and displayed and monitored in real-time. This method is able to reduce or eliminate the delay between event and measurement and is designed to be non-intrusive enough to maintain participant engagement over time.

2.2 The Construction of a Study: The Campaign, Survey, and Prompt

Each study is called a *campaign*, which defines the entirety of the data collection parameters. The typical campaign queries participants for certain types of information. Each individual query is called a *prompt*, and our five current prompt types are: number, timestamp, hours before now, single choice and multiple choice. An example single choice prompt from the above example is: "Did you exercise today?" with the possible answers being "Yes" and "No". The prompts are grouped into *surveys*, which are defined as a group of prompts that are displayed to the participant together. When a participant responds to a survey, they answer each of the survey's prompts in turn before submitting the survey as a whole.

Each campaign can also collect *actigraphy* data, using the onboard GPS and accelerometer to infer each participant's current activity and return a sample once a minute along with location data. The campaign author can decide whether to

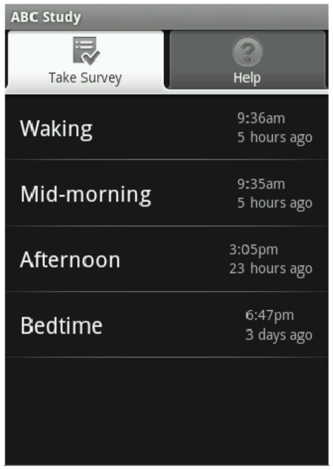


Fig. 2. The Android application survey selection screen. The surveys listed can be launched by the participant at anytime.

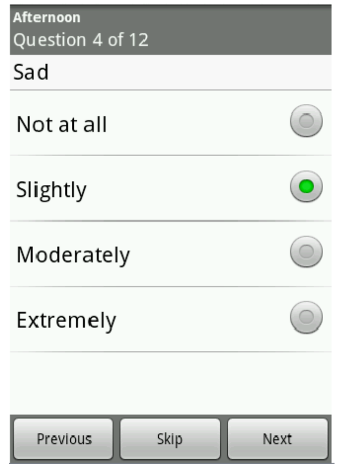


Fig. 3. An example of a single prompt as part of a survey. The participant can select any of the pre-configured responses, or submit a non-response by selecting skip.

enable this feature to collect further data without interrupting the participant, or to disable this feature to reduce energy usage and hence lower the burden of device maintenance.

Finally, each campaign includes *triggers*. ohmage's system of triggers reminds participants to respond to surveys. Each trigger is attached to a single survey and are based on time, location, and/or sensor readings. Each individual participant can modify their 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 keeps participants from forgetting 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.

Once a researcher has defined a survey, they author it using our XML schema discussed in detail in Section 3.1. The configuration file is first uploaded to the ohmage server, where it is validated and attached to the campaign. The same file is then retrieved by both the phone and visualization components for configuration.

2.3 Responding to Surveys on the Android Mobile Device

After a campaign is completely configured on the ohmage server, a participant can use the Android mobile application to login and automatically download all survey information. The main application screen is shown in Figure 2, which displays a list of surveys that can be launched. The surveys are either self-initiated by the participant in response to an event or are set to launch when triggered. On survey launch, the survey prompts are displayed one at a time in turn. Each prompt screen is layed out based on the individual prompt type. An example is shown in Figure 3, which shows a prompt of type single choice. The participant can either decide to answer the question or click the skip button to issue a non-response.[1] After completing a survey, the participant clicks submit and the survey is packaged and securely and reliably uploaded to the server when network access is available.

2.4 Monitoring the Campaign in Real-time

After participants start responding to surveys and uploading sensor data, both researchers and participants can login to the browser based front end. The campaign overview portion of the front end leverages the wireless nature of today's devices by showing the researcher real-time user statistics and upload information. Researchers can follow up with participants who are failing to upload consistently and check for problems with hardware, software, motivation, or lack of technical expertise. Additionally, both researchers and participants can login to the system to visualize their uploaded data. Researchers can use tools to perform basic analysis which can help to confirm initial hypothesis and discover correlations. Participants can view their data to set and meet goals and to gain a sense of accountability for their behavior.

3. SYSTEM DESCRIPTION

This section presents a detailed description of the components important to our described design goals. They include the overall survey framework, the visualization sub-system, and resource and energy efficient services.

3.1 A Survey Framework to Collect Data from Participants

The ohmage survey framework supports five prompt types, conditional branching in surveys, and repeating sets of prompts. Researchers can author a survey using our defined XML schema.

3.1.1 *Standard Survey Prompt Types.* As previously defined, surveys consist of a list of prompts, each querying the participant for a response. Each of the prompts can be defined as one of several prompt types. The prompt type indicates to the smartphone how to display the prompt to the participant and indicates to the server what type of information the prompt collects. We currently have defined a set of five common prompt types: number, timestamp, hours before now, single choice, and multi choice. A number represents any sort of float or integer and usually represents a measurement. A timestamp and hours before now both indicate time. Single choice and multi choice present an array of possible responses to the participant, who can select one or more than one respectively.

3.1.2 *Survey Branching.* The survey branching feature reduces the number of prompts shown to the participant by allowing each prompt to be hidden based on the response to previous prompts. For example, if a participant answered “No” to “Did you exercise today?”, they would not need to respond to the next prompt “For how long did you exercise?”. Instead, the study author can attach an optional *condition* field to the prompt. The condition field contains simple conditions such as “A == Yes” or composite conditions like “choice = 3 and amount > 10”. In the previous example, the author would add a condition such as “prompt exercise

[1]Focus group participants desired the ”skip” functionality as a form of basic privacy.

today == Yes", which would only display the prompt if the participant responded "Yes" to the previous prompt.

Researchers also requested the ability for basic looping of a set of prompts, a feature we call "repeatable sets".

3.1.3 *Repeatable Sets.* Repeatable sets contain a list of prompts. The conditions for prompts inside a repeatable set may only refer to other prompts within the set. Once the list of prompts is traversed, a "termination question" is displayed to determine if the participant wants to complete an additional iteration of the set, or if they want to continue with the remainder of the survey. An example repeatable set might be a list of prompts asking the participant about each meal of the day. The termination question would simply ask "Did you eat another meal today?". Responding "Yes" would cause the survey to repeat the set of meal prompts.

3.1.4 *Researcher Oriented Survey Authoring.* We introduced the primary components of the survey authoring framework, including prompts, conditions, and repeatable sets. To use these components and actually design a campaign with a set of surveys, the author must create a XML configuration file conforming to a provided schema. The server validates the configuration when it is loaded into the system using a large suite of tests. Some examples of these tests are: testing for prompt id uniqueness across a campaign, checking that each prompt defines valid properties based on its type, and verifying that conditions only reference prompts that came before it.

3.1.5 *Buttons as Lightweight Survey Abstractions.* During our focus groups, participants repeatedly expressed concern over the ease of responding to surveys. Based on this, we designed two experimental features. The first is a stress button, a button/icon that is placed on the phone's homescreen to be readily accessible. When the button is tapped, a stress event is recorded, which consists of a survey with a single prompt that is automatically responded to. No prompts are presented to the user. These surveys are time and location stamped, and so can be used to determine stressful locations or times throughout a typical day. We hope that users will find this feature useful for quickly noting events that can be used as reference points for later reflection.

The second experimental but more straightforward feature is the food button. This is also presented as an icon on the homescreen for quick accessibility. Tapping this icon will launch a short survey with only two prompts, the first asking for a picture of the meal that the participant is about to consume, and the second asking how hungry the user is at that moment before the meal. We hope that a short survey with a easily accessible link on the phone will encourage users to provide such in-the-moment information.

3.2 Utilizing Triggers to Maintain Long-term Engagement

The triggers in ohmage are created and managed using a highly customizable component called the *Trigger Framework* which is a part of the ohmage mobile application. This framework performs end-to-end trigger management from providing trigger authoring mechanisms to notifying the user to take surveys and handling the user responses. For simplicity of control and configuration, we centralized all functionality relating to notifications to the user as well as launching of survey reminders in this framework. Through our experience with triggers, feature requests from researchers, and feedback from participants, we discovered there are a number of corner cases and configuration parameters that the trigger framework needs to handle.

One of the main design goals of the trigger framework is to provide a uniform experience in creating and managing triggers of any supported type, such as *location based triggers* and *time based triggers*. Users of the mobile application can create triggers of any supported type and associate one more surveys using the graphical user interface. When a trigger fires, the notification manager in the trigger framework sends a notification to the user to take the associated surveys. Notifications are independent of the trigger type, and are accompanied with an audio and vibrating alert. All aspects of notifications, including frequency, lifetime of a notification, duration and sound of alerts, and suppression of notifications are configurable through the notification manager's GUI.

The current version of ohmage supports time and location based triggers implemented using the trigger framework. The following sub-sections describes these trigger types in detail.

3.2.1 *Time Based Triggers.* Time based triggers are used to prompt the participant to take one or more surveys at a specific time of the day. For instance, a time trigger can be set to remind the participant to take some surveys during the lunch time everyday. This can be further customized to be triggered on any combination of the days of the week such as weekends. Time triggers can also be set to prompt the participant at a random time between two different times of day. On each day, these random triggers go off at different times with in the given interval. These combinations allow the researcher to either randomly sample a mood or behavior from a participant throughout a day, or to query a participant about event that happen at specific times of the day.

To help in calculating participant response times, whenever a survey response is submitted, the details of all the triggers which prompted that survey are also uploaded. The trigger time stamp is also added to this information, which can be used in data analysis to figure out how long the participant has taken to answer the surveys after the trigger went off. This helps in understanding the usefulness of the triggers and also to improve the trigger notification mechanisms if the surveys are not answered on time.

3.2.2 *Location Based Triggers.* Location based triggers are used to remind the participant to take surveys at a specific location. In our implementation, a location is defined as a circular region and it can be represented by its center's geographical coordinate along with the radius. Furthermore, these locations are assigned meaningful names such as *Home* and *Work*. Users of the mobile application can manage these locations (and the associated meaningful names) through a graphical user interface. This interface displays a map to the user on which any location can be marked and the radius can be adjusted visually.

Location based triggers can also be configured to go off only between two specific times of the day at the specified location. Thus, these triggers with location and

Table I. The seven fields comprising our generic data point format. The required fields together give a complete description of the context of the data point, without needing additional configuration information. The optional fields each give additional information which can be leveraged in various displays.

	Data Type	**Description**
label:	UTF-8 string	Data point label (required)
value:	Float or UTF-8 string	The data point value (required)
timestamp:	ISO 8601 timestamp	Data point creation time (required)
type:	UTF-8 string	count, measurement, event, category (required)
unit:	UTF-8 string	The value's unit (optional)
location:	List of floats	GPS coordinates (optional)
metadata:	JSON Object	Additional context for the data point (optional)

time can go off at most once per day. This is particularly useful when the participant can enter a location multiple times but only requires to be prompted the very first time. Without this time constraint, a location trigger will go off every time the participant enters the specified location. This behavior can be undesirable if the participant re-enters a location very soon after leaving it.

3.3 System Monitoring and Data Visualization

As described in Section 2.4, we have designed a browser based campaign front end to both allow the researcher to monitor the study and participant statistics, and the researcher and participants to view and analyze collected data in real time.

When designing these visualization toolkits, we found that different campaigns and surveys have disparate types of data and each researcher has their own focus for visualizations. These requirements led to the creation of customized visualizations per campaign. However, creating a new set of visualizations from scratch for every campaign is not scalable nor usable by those less technically oriented. Instead, we leverage the fact that all customizations for campaigns are authored using the XML authoring tools. In the XML schema, each prompt in a survey can only have one of a limited number of prompt types. This led to a generic data point format and a mapping from each of the prompt types, that is number, timestamp, hours before now, single choice, and multiple choice, into the data point format. Accordingly, any visualizations that are created for the generic format will automatically work for all campaigns, and all campaigns can instantly use the current set of visualizations. This data point abstraction is described below.

3.3.1 *Description of Generic Datapoint Format.* Our data point abstraction format meets three requirements. First, the format is amenable enough to represent all the required data types. As described in Section 3.1, ohmage surveys have five measurement oriented prompt data types: number, timestamp, hours before now, single choice, and multi choice. A number can represent either a count (i.e., How many meals did you have today?) or a measurement with a unit (i.e., What is the current air temperature?). The timestamp and hours before now types represent events (i.e., When did you wake up today?). The final two prompt types, single choice and multi choice, represent categories (i.e., What types of food did you eat today? Choose from the list below.) Mobility data, which is separate from the survey data, represents a category describing a participant's current activity. These four measurement types are *display types*, and the mapping from prompt type to display type is defined in the individual survey configurations.

Our second requirement states that 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. As seen in Table I, the information returned for each data point gives a complete description of that point including its time of creation, unit, and display type which allows for a number of pre-defined visualizations.

The final requirement is that the format is extendable to all types of metadata. The participant may want to label data points or survey with additional contextual information such as images, audio, or simply entered text. This metadata is associated with other data, but does not stand up as a full data point on its own. To this end, we have used an extensible metadata field, which contains an array of JSON objects containing the two fields *type* and *value*. An example type is *image link*, which indicates the value is a link to a compressed image. A hypothetical data viewer could then display the image as part of a metadata display.

3.4 Energy Efficient Services

3.4.1 *Location Service.* As identified in several related research projects [Constandache et al. 2009; Abdesslem et al. 2009; Kjærgaard et al. 2009], we also found that we cannot continuously run the GPS receiver to track user's location. This is due to the limited capacity of smartphone batteries[2] and high energy consumption of GPS on phones. Several solutions have been proposed to reduce the power consumption of location based applications.

All proposed solutions are based on using another lower-energy consuming sensor to detect if the user is stationary and turn off the GPS sensor. For example, [Zhuang et al. 2010] increase the interval that GPS remains off by detecting when the user is stationary using the accelerometer. It also suggests using GPS velocity to infer how long the user will stay within the tolerable location error range. RAPS [Paek et al. 2010] uses a combination of accelerometer and GSM cell tower signal strength to duty-cycle GPS. In addition RAPS adapts the sampling rate of GPS based on the applications' accuracy requirements. Similarly [Lin et al. 2010] uses multiple sensors to reduce energy consumption for location when the user is mobile. It switches to other modes of location sensing (such as Wi-Fi triangulation) based on application accuracy requirement and accuracy of each sensing technology at any location. In a more semantic approach SensLoc [Kim et al. 2010] identifies significant places of the user and uses a combination of accelerometer, Wi-Fi, and GPS to efficiently detect these places and the paths that a user takes between them.

For ohmage, we decided to use a simple cache of frequently visited Wi-Fi access point IDs to identify when the user is immobile and turn off the GPS receiver. We chose this simple algorithm to avoid many unexpected complexities that come with more sophisticated and experimental techniques. We found that, based on the requirements of ohmage, this simple technique is effective enough to reduce energy

[2] The capacity of Lithium-ion batteries that are shipped with most smartphones ranges from 1000 to 1500 mAh.

consumed by GPS.

We implemented our algorithm as an Android service that can be accessed by multiple entities through the Android IPC, thus allowing all other applications to "piggyback" on our service. This location service offers a synchronous interface to other interested applications. It also allows applications to manage the parameters that affect its accuracy and energy consumption, such as sampling frequency. The location service periodically scans for available Wi-Fi access points. Access points (AP) with "weak" signal strength are filtered and the remaining AP IDs are stored in a cache along with the corresponding GPS location. If GPS is not available at that location the cache entry reflects that. Each entry also includes the number of times the Wi-Fi signature has been visited and the time stamp of the last encounter. A Wi-Fi signature is considered "significant" if it has been visited more than three times within 10 minutes. At any point, if the currently visible Wi-Fi signature matches a significant cache entry, the GPS receiver is turned off and the corresponding location object (if it exists) is returned to applications that request location. Otherwise, the service keeps the GPS on to serve location queries of other applications.

3.4.2 *Mobility Inference.* Mobility is an important indicator of a patient's wellness, especially when his/her condition affects ambulation (e.g., undergoing physical therapy or suffering from Parkinson's disease). ohmage has an optional mobility component that automatically tracks the user's mode of locomotion or transport. It can distinguish whether a user is still, walking, running, biking, or driving as long as the user keeps the mobile device on him/her. The user's activity is inferred by a classifier that uses as features accelerometer data as well as the speed as reported by the GPS. Our algorithm analyzes one-second windows of accelerometer data sampled at 20-30 Hz. The variance and fft coefficients are calculated on the magnitude of acceleration of each sample within the one-second window. We use the Euclidean distance of the X, Y, and Z axes to measure the magnitude of acceleration. This way has the advantage of producing a value that is agnostic to the direction the phone is facing. This allows the user to carry it in any position. We trained a C4.5 decision tree classifier for the first version of the application, as in [Ryder et al. 2009]. It functions indoors even when the phone cannot use GPS because the accelerometer is enough to recognize being still, walking, or running. GPS is required to detect outdoor activities like biking and driving.

Although the original classifier usually performed well, classifier accuracy varies from user to user. In order to allow the classifier to adapt to different users, we implemented an online semi-supervised learning algorithm. As shown in Fig 4, [Longstaff et al. 2010] found that the democratic co-learning [Zhou and Goldman 2004] method was able to improve classification up to an accuracy of about 90%. By using this in ohmage's mobility classifier, we can improve the performance when the classifier does not initially perform well for some users.

4. IMPLEMENTATION

ohmage currently has three separate components: the server that accepts incoming data and responds to requests for data, the Android mobile application which sends data to the server, and the browser based front end which requests data from the

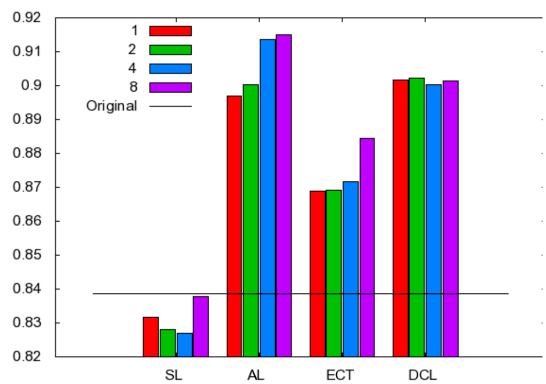


Fig. 4. Classification accuracy by method and number of iterations

server for display. Each component is connected through an HTTP based API and communicate through a well defined JSON based protocol. This openness and modularity allow both the phone and the front end application to be easily swapped out for other devices that implement our protocols. Our standard ohmage server and mobile components have implementations that may be of interest due to privacy, security, user interface design, IRB compliance issues, or simply as examples for future implementations.[3]

4.1 The Core ohmage Server Code

The ohmage server is the least noticed component, in that it has no direct interactions with the users. However, the server is the keystone that securely stores incoming participant data and correctly retrieves data when requested. The server APIs define its interactions with data collection devices and campaign monitoring tools. We describe the server's running environment and give a high level overview of its architecture.

4.1.1 *Server Environments.* The ohmage server is implemented in the Java 6 language, which will execute in any environment that has a Java VM implementa-

[3]We are currently using all of our components as a basis for an upcoming, student based data collection campaign in conjunction with LAUSD.

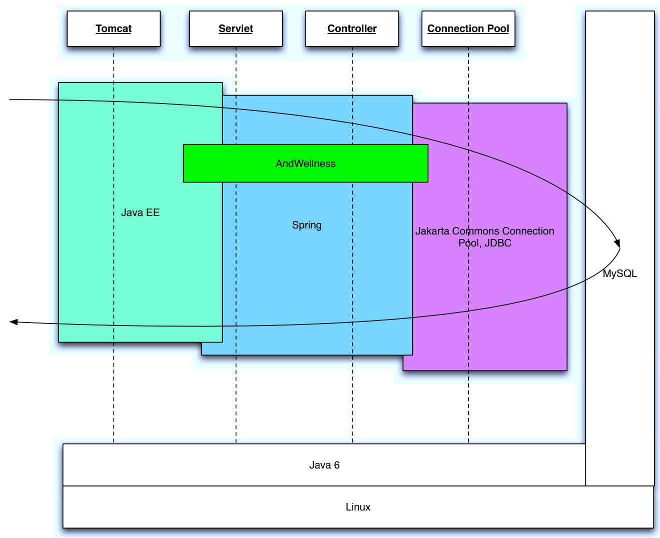


Fig. 5. A high level representation of the ohmage server architecture. ohmage lives within a number of Java systems, but dependencies on each are minimized to increase code portability. All API requests are mapped by Tomcat to a specific servlet which each have one implementing controller. The controllers are instantiated through Spring configurations and each consist of a list of request validators and request services which are executed in order.

tion. Our code is hosted in Tomcat 6.0 and uses JavaEE, but dependencies on the JavaEE API are minimized to allow the server to easily run in any Java Servlet environment. Collected data are stored in a local MySQL 5.1 database. Each campaign is designed to run in a separate virtual machine, maintaining a strict separation of data between studies to satisfy IRB privacy and security requirements.

4.1.2 *CentOS for Stability and Security.* We host the server code on the latest CentOS 5.5 server distribution. CentOS has an extremely long product cycle, which eliminates the need to handle frequent kernel updates and gives us enterprise class server stability. The system is locked down through the standard linux firewall, iptables, only allowing incoming connections on port 80 and 443 for HTTP access, and port 22 for server access and control. CentOS comes with SELinux automatically enabled, which gives a fine-grained resources access control for enhanced security. All campaign domains have an active SSL certificate and all communication to and from the server uses full SSL encryption.

4.1.3 *Modular Server Code.* The server code itself has been architected to be robust and modular. All of our defined APIs are called by sending HTTP POSTs to specific URLs. As displayed in Figure 5, incoming HTTPs calls are mapped to servlets implementing the feature by Tomcat based on a specific URL. Each servlet is mapped to a controller based on configuration in web.xml, the main Tomcat configuration file. Different implementations for each feature can be hot-swapped in by making a few XML edits.

Within each servlet/controller pair we use the Spring framework, a Java based toolkit that uses XML configuration files to wire together modules of Java code. Spring allows us to instantiate each controller as a list of validators and services. The controller then passes the request to each validator and each service in turn. As long as the validator or service implement our standard interfaces, they can again be swapped in and out through simple XML edits.

Finally, each service has direct access to the database connection pool. The database connection is abstracted by JDBC, allowing us to swap in different database implementations as necessary.

5. USE CASES, DATA, AND USER FEEDBACK

In this section we present a set of expressive analytics developed to evaluate the *success* of a typical deployment, as defined in the introduction. These analytics include:

—The aggregated and over time battery level of each participant's smartphone. High battery levels correlate with the quality of returned mobility metrics and sustained low battery levels correlate with the perceived burden of the study on the participant.

—The ratio of aggregated and over time capture rate of survey and mobility data relative to the requested capture rate.

—The aggregated and over time rate of participant interaction with the smartphone, where an *interaction* is defined as one on-off cycle of the main phone display. A higher interaction rate is an indicator of participant use of the smartphone in their day to day routines.

—An overview of the participants daily mobility data, including geographic span and hours spent both active and sedentary.

5.1 Study Overview

ohmage has collected data from three initial use cases. All are designed to collect survey and mobility data and to set baselines for various health, system, and participation based metrics. All three studies have at least one daily survey, collect continuous mobility data, and end with exit surveys to query each participant about their overall experience. Two of the studies also include SystemSense, a separately installed data collection software module to monitor overall participant interaction with the smartphone itself [Falaki et al. 2011; Falaki et al. 2010].

Our initial case study, code named *ABC*, measures the emotions and behaviors of breast cancer survivors. ABC includes four surveys triggered throughout the day, continuous mobility monitoring, and a single, in-person exit survey. We have data from 30 participants, each running ohmage on loaned phones for five consecutive days. ABC does not include SystemSense.

Our second case study, code named *Moms*, targets young mothers for 6 months to monitor diet, stress and exercise. Moms again includes four daily triggered surveys, collects continuous mobility data, and queries each participant with an exit survey similar to ABC. However, Moms also includes SystemSense to monitor participant interaction with the phone.

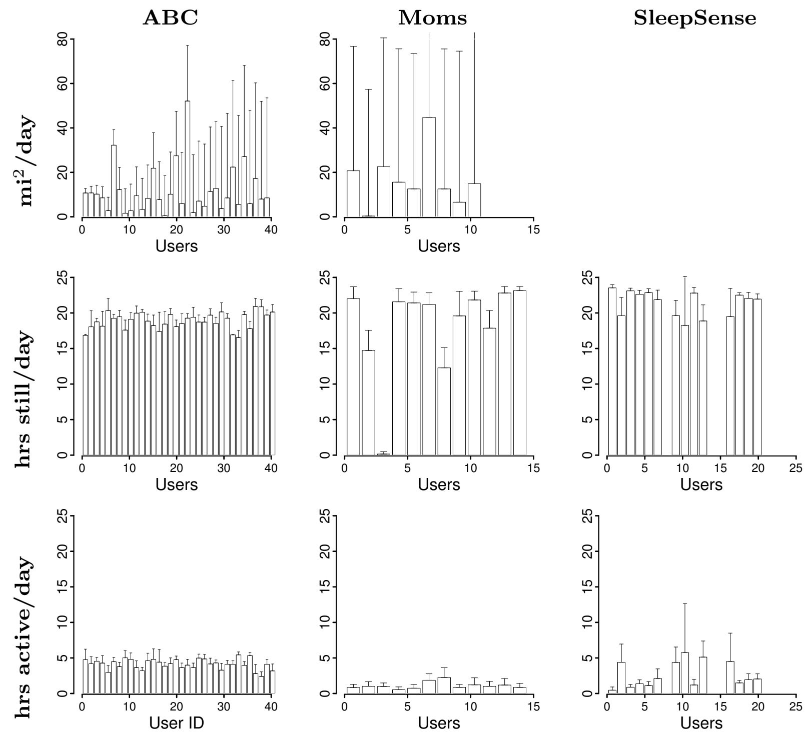


Fig. 6. Mobility metrics from three user studies. The first row presents the number of square miles a participant visits in a single day. The second row shows the hours spent still per day and the third row shows the hours spent active per day. Unfortunately, a software issue prevented all of the SleepSense study and a small number of participants in the Moms study from returning GPS coordinates, making their squares miles per day calculation impossible. In addition, some participants in the SleepSense study elected to disable mobility data, leaving them with no hours still or hours active data.

Our final study, code named *SleepSense*, focuses on the sleeping habits of participants. SleepSense only includes a single daily survey triggered when the phone first senses movement in the morning. SleepSense also allows the participant to disable mobility monitoring if they have concerns about battery life (and about half of the participants did). This study also runs SystemSense and we include data from 16 participants.

5.2 Methodology

In total, about 60 participants have completed one of the three studies, each using ohmage for five to seven consecutive days. Each study includes a survey component. In Moms and ABC, each participant is reminded to complete a short survey 4 times a day. The trigger times of the reminder are customizable by each individual participant. For SleepSense, the participant is notified to complete a survey once a day when the phone first detects motion in the morning. Upon completion, the survey is timestamped and returned to a central server. For the purpose of analysis, survey completion rate is defined as the number of surveys completed in a day versus

Table II. Summary of exit surveys from three user studies. Each participant is interviewed after completing the study about their satisfaction with the study, overall usability issues, and whether they would use the ohmage software in the future to monitor other behaviors.

ABC	Moms	SleepSense
Overall, how was your experience of using the mobile phone to fill out surveys (1 is little effort)?		
1.96	3.35	N/A
Did you have trouble keeping your phone charged?		
4 Yes / 21 No	2 No / 1 Yes	15 Yes / 8 No
Would you use ohmage for personal behavior discovery?		
21 Yes / 5 No	11 Yes / 0 No	20 Yes / 3 No
On a scale of 1-7, how helpful did you find the reminders (7 is very helpful)?		
5.5	5.86	N/A
What is the highest number of surveys acceptable in one day?		
3.81	4.22	N/A
What was your average number of minutes to complete a survey?		
1.83	0.96	0.78

the expected number of surveys.

All three studies also collect continuous mobility data, although this feature can be disabled in the SleepSense study to reduce battery consumption. Once a minute, the phone detects motion and speed to determine whether the participant is still, walking, running, biking, or driving. This data is timestamped, aggregated on the phone, and returned to a central server every few hours. Similar to the survey completion rate, the mobility capture rate is defined as the expected number of mobility data points over the actual received number.

The Moms and SleepSense studies include a separate monitoring software module called SystemSense to record the number of participant interactions with the phone, battery level, and battery charge rate. These numbers are aggregated once an hour on the phone and returned to the server.

Finally, all three studies end with an exit survey asking each participant six questions about their interactions with the phone during the study. These questions, and their answers averaged per survey, are presented in Table II.

5.3 Mobility Results

We present data from each study aggregating overall participant mobility and exit survey responses. The mobility analysis aggregates data into 24 hour chunks as the standard time period, balancing overall trends with bursty, within day activity. Also, a single day is a natural period of time to measure human behaviors, which tend to cycle every 24 hours. Unfortunately, most participants did not upload 100% of the expected mobility data per day, mostly due to either battery issues or starting or ending the study mid-day.

Figure 6 summarizes these results. Each row displays one mobility metric while the columns display results from each of the three studies. The first row presents the number of square miles visited by each participant in a single day, giving a sense of the geographic mobility of each participant. As might be expected, this metric has the most variance among participants.

The second and third rows show the estimated number of hours sedentary and the number of hours active per participant per day, where hours active is calculated

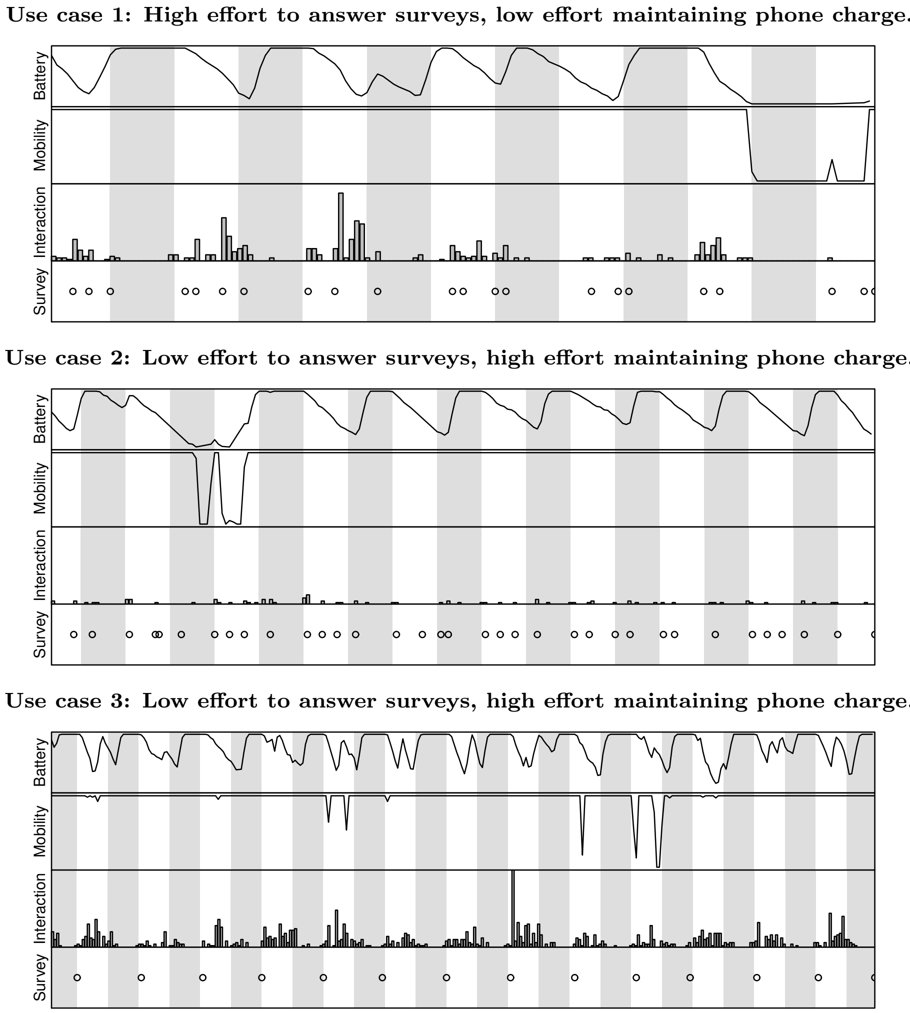
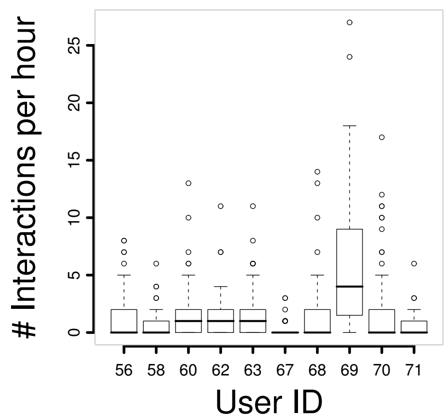


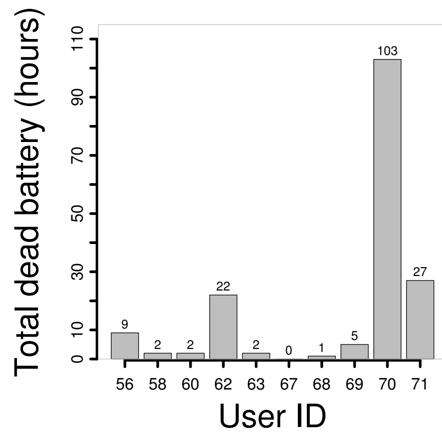
Fig. 7. Study participation data from three selected participants. From top to bottom the graphs display: battery charge level, mobility data return rate versus the expected rate, number of interactions with the phone, and a dot for every recorded survey completion. The shaded areas indicate night time and the white areas indicate day time.

as time spent walking, running, and biking with the phone carries (in handbag, pack, or pocket). Both metrics give an interesting overview of the various study populations.

The exit survey results are summarized in Table II. The survey responses are generally positive, most participants indicating that using the mobile phone to answer surveys was easy and convenient, and that they would continue to use ohmage in the future for other behavior discoveries. Most appreciated the utility of the survey reminders, and indicated that they completed each survey in one or two minutes.



(a) Interactions per hour for Moms participants



(b) Dead battery time for Moms participants

Fig. 8. Compares interactions per hour and total dead battery time for the Moms participants.

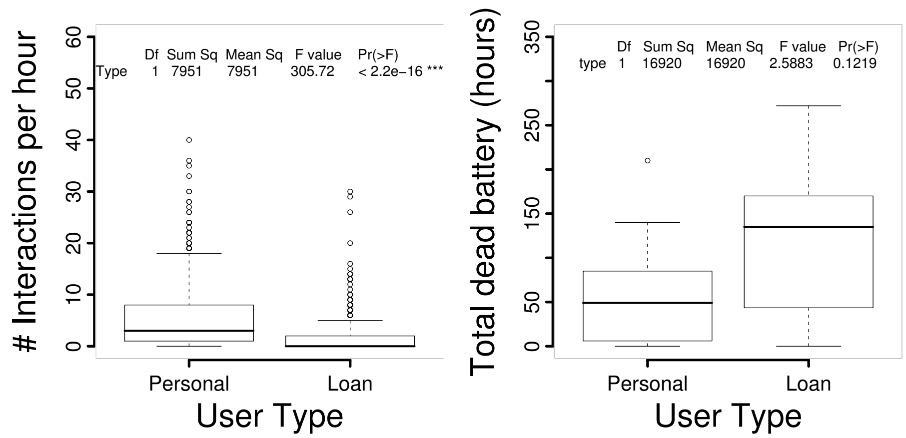
5.4 Participation Analytics

Similar to measuring usability, ohmage collects metrics to allow the researcher to analyze *participation quality* over time for each participant. Three example participants are shown in Figure 7. From the top down we see: battery charge level, mobility data return rate, number of interactions with the smart phone, and a small circle for every survey completion. These metrics give the researcher an in-depth view into each participant's usage of the ohmage system, allowing the researcher to pinpoint and regulate data quality. Specific participant's with questionable data can be manually examined, or the metrics can be used to automatically infer quality of data over time.

As mentioned, the first two rows of data are battery charge and mobility return rate. It's clear that the quality of mobility data drastically decreases when the battery charge level gets too low so data collecting during those periods can be marked as such.

The third row conveys the amount of participant interaction with the phone. Participant 1 interacts with the phone much more than just answering surveys, indicating she is more familiar and comfortable using the phone in her day to day life and could be why she found it easier to maintain the battery charge. Participant 2 only interacts with the phone to complete a survey, indicating the phone is not otherwise a part of her daily routines. However, participant 2 has a more consistent survey response rate than participant 1 and indicates the surveys are low effort to complete. Finally, participant 3 is a power user with her own phone, but indicates the additional power draw from the mobility data noticeably increased the burden of the study.

While per participant analytics give insight into each specific user's participation style and quality, aggregated metrics start to analyze possible correlations. Figure 8 aggregates the number of interactions per hour and total dead battery time in



(a) Mean and 75% confidence of interactions (b) Mean and 75% confidence of total dead battery hours

Fig. 9. Compares interactions per hour and total dead battery time for personal versus loaned phones.

hours for participants from the Moms study. In this way, the researcher can, at a glance, see the overall interaction style and potentially the quality of all participants. Participants can easily be categorized as power users, such as user 69, or survey only users, such as user 67.

Continuing to aggregate, Figure 9 summarizes the number of interactions per hour and total dead battery hours by whether the participant used his or her own smartphone. This provides some evidence for the observation that participants using their own phones are already accustomed to their use and maintenance, and hence are able to return higher quality data. The first plot shows the personal phone users have a noticeably higher number of interactions per hour, but the second plot shows a significant decrease in the total number of dead battery hours for personal phone users. These both highlight the importance of user familiarity with the survey hardware and software and start to validate the use of participation metrics to predict and measure data quality.

Participation and data quality can also be assessed by the survey completion rate and mobility capture rate. Mobility capture rate per day generally reflects the time a participant is able to keep their phone charged and the ohmage application running. This metric directly indicates mobility data quality, as missing mobility data must be inferred. Low mobility capture rates also indicate a lack of engagement with the study. Per user survey and mobility return rates are aggregated in Figure 10.

5.5 Discussion and Design Goals

New experience sampling systems record metrics that allow researchers to assess and improve user participation over time in real-time. This insight attempts to solve two lingering problems from older systems. First, participants might not understand the instructions or experience device failure, problems that the researcher

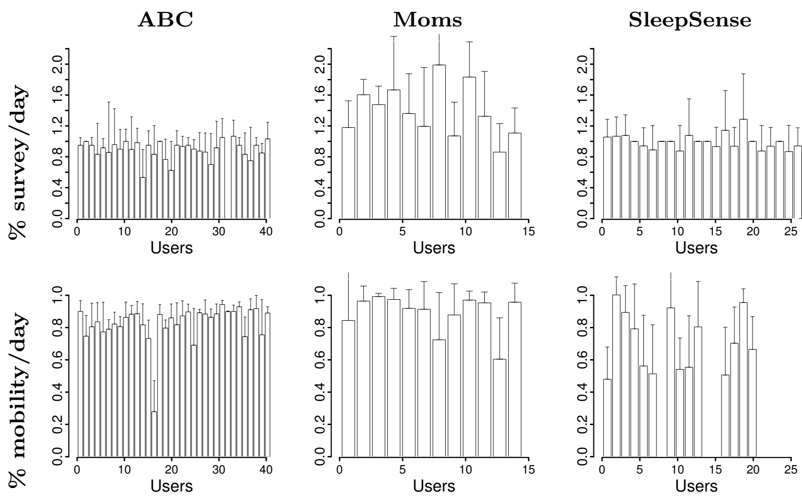


Fig. 10. Evaluation metrics from all three user studies. The three columns present the average survey completion rate and the average mobility data capture rate per participant. A number of users in the SleepSense study elected to turn off mobility data collection.

might not otherwise notice until the end of the study. Since finding and retaining participants is both time and resource intensive, obtaining high quality data from each is important. The real-time nature of these participation quality metrics can give researchers specific actionable responses.

Second, participants can miss entries and back-fill previous data at the end of the day or week. This second problem can be particularly insidious as there may be no malicious intent by the participant, but back-filling data side steps the point of gathering data in situ and is difficult for the researcher to detect. This problem can cast doubt on the quality of data from entire studies.

While ohmage cannot claim to directly solve either of these problems, we do attempt to decrease the burden of collecting data on participants to decrease dropout rates, as well as to define usability and participation quality metrics that will help researchers assess and improve data quality. Some design goals particularly stand out.

First, the process of survey design and creation should directly cater to reducing the burden on participants. The ABC and Moms studies, which query participants four times a day, allow participants to individually customize the times at which survey reminders trigger. The SleepSense study goes a step further, automatically triggering a survey when the participant first moved their phone in the morning. On the other hand, participants indicate that after completion, the system should give better feedback that the surveys were successfully collected by the server.

Second, experience sampling systems can benefit from defining and collecting contextual metrics from the collection platform itself. Analyzing survey completion rates and mobility capture rates (metrics from within the system) is helpful and informative, but does not always give context as to why the rates are high or low. SystemSense based metrics, that is measuring battery life and participant interaction with the phone, allows a much clearer understanding of the ongoing quality of participation. A short glance at Figure 7 immediately conveys information about the context of participation for the example participants.

Finally, exit surveys only capture a single snapshot of time. The exit surveys can easily be skewed by a single recent experience with the system. To better validate the direct value of participation quality metrics, future studies can integrate satisfaction surveys directly into the system itself. Going even further, future studies can measure additional metrics to monitor participant motivation and quality over time including, the time between survey trigger and actual response, the time it took each participant to complete a survey, and more in-depth monitor of participant interaction with the phone. These, combined with more integrated exit surveys, can fully validate a set of expressive analytics designed to monitor and contextualize participation quality over time.

6. RELATED WORK

Related work can be divided into two classes: previous experience sampling based studies published by clinical psychologists, and specific software systems designed to survey individuals from electronic devices.

6.1 Ecological Momentary Assessment

Psychologists have used variations of the EMA method for some time. Arguably one of the earliest and most low tech method is the paper diary. Participants are asked to log surveys in a notebook at predetermined times or whenever certain events occur. Lemmens et al. use this method to study how participants report alcohol consumption and compares weekly interviews to daily diary entries [Lemmens P 1988]. Diaries are cheap and reduce recall bias but they suffer from numerous problems. First, while diaries have a very low upfront cost, the post-study analysis including entering diary data and following up on illegible entries can be labor intensive and costly. Second, adherence to the entry schedule cannot be verified. Recent studies show that participants tend to fake entries, the most common method being to fill out many entries at once right before turning in the diary [Stone et al. 2003]. Finally, asking participants to log data multiple times a day without other motivation introduces adherence issues. Reminders to complete surveys, especially if a participant is falling behind, helps to engage them in the data collection process and boost adherence rates [Shiffman et al. 2008].

Later studies employ a wide range of techniques to enhance participant adherence to study protocols and to assess per user and overall protocol compliance. Leveraging existing infrastructure, Searles et al. deployed an automated telephone system and instructed participants to call in daily and respond to various audible prompts [Searles et al. 1995]. These systems eliminate problems with data entry as all data are recorded as integer touch-tone responses to precompiled questions. Responses are also time stamped, eliminating the ability to retroactively answer responses. Later studies work to optimize this method by utilizing the real-time nature of the data collection to contact participants who are failing to call in and by handing out mobiles to facilitate the ease of calling in [Collins et al. 2003].

Finally, recent studies employ various hand held electronic devices to combine both prompt triggers to remind the participants to complete a survey, and compliance checks via time stamps to ensure the assessments were completed at the expected times. Epstein et al. employed Palm Pilots to study the situations that trigger heroin and cocaine cravings and use in recovering participants. The study was designed to prompt the participant for survey responses randomly during the participant's waking hours, and also whenever use or cravings occurred. The PDA devices had no wireless connectivity and were collected and replaced often to download data and recharge batteries. While requiring participants to return to the lab once a week was burdensome, the devices themselves recorded data quickly and easily, and the study garnered an overall 75% response rate [Epstein et al. 2009].

Other studies make use of the wireless nature of smartphones by using simple SMS messaging. They improve self-monitoring of behavior by both sending reminders and collecting survey data via texts [Patrick et al. 2009], [Obermayer et al. 2004]. This technique is flexible as the entire study including the SMS sending schedule can be tweaked configured per user. Hence, texting allows many of the conveniences that come from wireless such as per user configuration of triggers, reminders and survey questions and the real-time viewing of returned data. However, collecting large amounts of data, continuous data, and other types of non-textual data is not possible.

6.2 Other Software Systems

Specific data collection systems include commercial software such as Pendragon Forms, Frontline Forms, and Nokia Data Gathering, which provide tools to organize surveys and collect data, but they are closed source and closed standards. Open source projects such as JavaRosa, RapidSMS, FrontlineSMS, and EpiHandy are more flexible to tune the system for particular uses, but are primary focused on collecting textual data. These systems work well for their intended audiences, mainly targeted data collection by an educated set of participants but don't incorporate reminders, triggers, or other methods to enhance the participation of users monitoring their daily life. Our system instead assumes non-technical participants and allows and motivates these participants to collect data about themselves and their environment. ohmage is designed to fit unobtrusively into a participants' daily life to encourage longer term data collection and lower drop out rates.

In addition, several research efforts have been focused on improving the quality of the user responses. The Context-Aware Experience Sampling (CASE) tool incorporates sensor data to trigger self-report surveys at specific moments of interests [Intille et al. 2003]. However, it was not designed to run on a participant's own personal device and does not offer the flexibility of specifying dynamic trigger conditions and generic actions. MyExperience extends context-triggered ES with passive logging of contextual information such as device usage and sensor readings [Froehlich et al. 2007]. By combining in situ qualitative data (*e.g.*, survey answers) and quantitative usage data of the smartphone (*e.g.*, phone calls), it provides different perspectives for researchers to look at the gathered data. It also lowers the learning curve for researchers to collect data by providing a lightweight XML-based configuration. Mobile Heart Health attempts to go further to link triggers to contextual data by continuously reading heart rate data from a wireless ECG [Morris and Guilak 2009]. To minimize participant interruption and target stressful events, the system waits for strong deviations in a participants heart rate before activating surveys and other automated coaching techniques. Finally, UbiGreen demonstrates an experience sampling application that combines automated sensing along with an innovative incentive feedback system to maintain motivation [Froehlich et al. 2009]. The participants of the study are given feedback directly on the phone, thereby eliminating the need to login to an online website and more tightly coupling participation and incentives to continue.

Reconexp combines a website where participants can review data they have provided during the day by using the ESM running on a hand held device [Khan et al. 2008]. In cases where participants are unable to respond, they may answer at a later time using a website. However, participants cannot initiate the queries by themselves and do not automatically synchronize the captured data with a remote server. Using a feedback mechanism to improve response quality has also been studied and proven to be effective. By showing participants their own collected information, it makes the information personally relevant and interesting and increases the value of the study to the participants [Hsieh et al. 2008]. To further improve responses by reducing the load on participants, researchers have found that collecting images is much quicker and easier than answering a multitude of survey questions. The images provide their own context, and can be used to jog participants memory for followup questions at a later time [Carter and Mankoff 2005]. Other systems use various bluetooth devices to capture even more types of data, further reducing the need for imprecise survey responses [Rasid and Woodward 2005].

7. CONCLUSION

As simple and unobtrusive as answering an SMS message, participants can complete an entire EMA survey on their smartphone. A short vibration or buzz serves as a reminder to complete a survey. In our experience, less than 10% of participants have expressed hesitation to this mode of data collection. The responses participants provide to the surveys on the phone may be approximate – at times the responses may be completely inaccurate. But the impact remains: For these several moments a day, an individual can evaluate their energy, mood, activity, or any other behavior without worrying about judgment. That person can reflect clearly and lucidly on where they are at, and perhaps where they want to be. It is this power, to unobtrusively yet intelligently remind us to live in the moment, that makes smartphones such a powerful tool for behavior change.

Smartphones also make it easy to incorporate images, automated data collection, and programmable reminders, increasing self-accountability. Participants in our focus groups and studies have been especially excited about the ability to take pictures of food that they were eating, to help with portion control and nutrition monitoring. Participants used phrases such as “The picture doesn’t lie” to describe how images would be used to hold themselves accountable. This notion of accountability to self is a strong motivator for the use of ohmage: *Ground truth* is not as important as a *self* truth for behavior change. The use of approximate measures for behavior change is not only important for measuring diet related behaviors. The introduction of an accelerometer and GPS on most smartphones has made it possible for the phone to act as an actigraph. While the phone may not be able to distinguish between 20 minutes and 22 minutes of exercise, it can still serve as a general accountant for how mobile or sedentary a person is from one day to the next.

ACKNOWLEDGMENTS

This work was made possible with funding from the UCLA Center for Embedded Networked Sensing (NSF #CCF-0120778), the National Institutes of Health (NHLBI #5RC1HL099556), Northwestern University (#20094762), the UCLA Center for HIV Identification Prevention and Treatment Services, and gifts from Google and Nokia.

REFERENCES

Abdesslem, F., Phillips, A., and Henderson, T. 2009. Less is more: Energy-efficient mobile sensing with SenseLess. In *Proceedings of ACM SIGCOMM Mobiheld workshop*.

Carter, S. and Mankoff, J. 2005. When participants do the capturing: the role of media in diary studies. 899–908.

Collins, R. L., Kashdan, T. B., and Gollnisch, G. 2003. The feasibility of using cellular phones to collect ecological momentary assessment data: Application to alcohol consumption. *Experimental and Clinical Psychopharmacology 11,* 1, 73–78.

Constandache, I., Gaonkar, S., Sayler, M., Choudhury, R., and Cox, L. 2009. EnLoc: Energy-efficient localization for mobile phones. In *INFOCOM 2009, IEEE*. IEEE, 2716–2720.

Donovan, D. and Marlatt, G. 2005. *Assessment of addictive behaviors*. Guilford Press.

Epstein, D. H., Willner-Reid, J., Vahabzadeh, M., Mezghanni, M., Lin, J.-L., and Preston, K. L. 2009. Real-Time Electronic Diary Reports of Cue Exposure and Mood in the Hours Before Cocaine and Heroin Craving and Use. *Arch Gen Psychiatry 66,* 1, 88–94.

Falaki, H., Mahajan, R., and Estrin, D. 2011. SystemSens: A Tool for Monitoring Usage in Smartphone Research Deployments. In *Proceedings of the sixth ACM international workshop on mobility in the evolving internet architecture*. ACM.

Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., and Estrin, D. 2010. Diversity in smartphone usage. In *MobiSys '10: Proceedings of the 8th international conference on Mobile systems, applications and services*. ACM, New York, NY, USA.

Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B., and Landay, J. A. 2007. Myexperience: a system for in situ tracing and capturing of user feedback on mobile phones. In *MobiSys '07: Proceedings of the 5th international conference on Mobile systems, applications and services*. ACM, New York, NY, USA, 57–70.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., and Landay, J. A. 2009. Ubigreen: investigating a mobile tool for tracking and supporting green transportation habits. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*. ACM, New York, NY, USA, 1043–1052.

Hsieh, G., Li, I., Dey, A., Forlizzi, J., and Hudson, S. E. 2008. Using visualizations to increase compliance in experience sampling. In *UbiComp '08: Proceedings of the 10th international conference on Ubiquitous computing*. ACM, New York, NY, USA, 164–167.

Hufford, M. 2007. *The Science of Real-Time Data Capture: Self-Reports in Health Research*. Oxford University Press, 54–75.

Intille, S. S., Rondoni, J., Kukla, C., Ancona, I., and Bao, L. 2003. A context-aware experience sampling tool. In *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*. ACM, New York, NY, USA, 972–973.

Khan, V.-J., Markopoulos, P., Eggen, B., IJsselsteijn, W., and de Ruyter, B. 2008. Reconexp: a way to reduce the data loss of the experiencing sampling method. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. ACM, New York, NY, USA, 471–476.

Kim, D., Kim, Y., Estrin, D., and Srivastava, M. 2010. SensLoc: Sensing Everyday Places and Paths using Less Energy.

KJÆRGAARD, M., LANGDAL, J., GODSK, T., AND TOFTKJÆR, T. 2009. Entracked: energy-efficient robust position tracking for mobile devices. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*. ACM, 221–234.

LEMMENS P, KNIBBE R, T. F. 1988. Weekly recall and diary estimates of alcohol consumption in a general population survey. *Journal of Studies on Alcohol 49,* 2, 131–135.

LIN, K., KANSAL, A., LYMBEROPOULOS, D., AND ZHAO, F. 2010. Energy-accuracy trade-off for continuous mobile device location. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 285–298.

LONGSTAFF, B., REDDY, S., AND ESTRIN, D. 2010. Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 4th International ICST Conference on Pervasive Computing Technologies for Healthcare 2010*. IEEE, 1–7.

MARLATT, G. AND GORDON, J. 1985. *Relapse Prevention: Maintenance Strategies in Addictive Behavior Change*. Guilford Press.

MORRIS, M. AND GUILAK, F. 2009. Mobile heart health: Project highlight. *IEEE Pervasive Computing 8*, 57–61.

OBERMAYER, J., RILEY, W., ASIF, O., AND JEAN-MARY, J. 2004. College smoking-cessation using cell phone text messaging. *Journal of American College Health 53,* 2.

PAEK, J., KIM, J., AND GOVINDAN, R. 2010. Energy-efficient rate-adaptive GPS-based positioning for smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 299–314.

PATRICK, K., RAAB, F., ADAMS, M., DILLON, L., ZABINSKI, M., ROCK, C., GRISWOLD, W., AND NORMAN, G. 2009. A text message-based intervention of weight loss: Randomized controlled trial. *Journal of Medical Internet Research 11,* 1.

RASID, M. AND WOODWARD, B. 2005. Bluetooth telemedicine processor for multichannel biomedical signal transmission via mobile cellular networks. *Information Technology in Biomedicine, IEEE Transactions on 9,* 1 (March), 35–43.

RYDER, J., LONGSTAFF, B., REDDY, S., AND ESTRIN, D. 2009. Ambulation: a tool for monitoring mobility patterns over time using mobile phones. In *Computational Science and Engineering, 2009. CSE'09. International Conference on*. Vol. 4. IEEE, 927–931.

SCHWARZ, N. 2007. *The Science of Real-Time Data Capture: Self-Reports in Health Research*. Oxford University Press, 11–26.

SEARLES, J., PERRINE, M., MUNDT, J., AND HELZER, J. 1995. Self-report of drinking using touch-tone telephone: extending the limits of reliable daily contact. *Journal of Studies on Alcohol 56,* 4, 375–382.

SHIFFMAN, S., STONE, A. A., AND HUFFORD, M. R. 2008. Ecological momentary assessment. *Annual Review of Clinical Psychology 4,* 1, 1–32.

STONE, A. A., SHIFFMAN, S., SCHWARTZ, J. E., BRODERICK, J. E., AND HUFFORD, M. R. 2003. Patient compliance with paper and electronic diaries. *Controlled Clinical Trials 24,* 2, 182–199.

STORY, M., GILES-CORTI, B., YAROCH, A. L., CUMMINS, S., FRANK, L. D., HUANG, T. T.-K., AND LEWIS, L. B. 2009. Work group iv: Future directions for measures of the food and physical activity environments. *American Journal of Preventive Medicine 36,* 4, Supplement 1, S182 – S188. Measurement of the Food and Physical Activity Environments - Enhancing Research Relevant to Policy on Diet, Physical Activity, and Weight.

ZHOU, Y. AND GOLDMAN, S. 2004. Democratic co-learning. In *Proceedings of the 16th IEEE international conference on tools with artificial intelligence*. Washington, DC: IEEE Computer Society Press, 594–202.

ZHUANG, Z., KIM, K., AND SINGH, J. 2010. Improving energy efficiency of location sensing on smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 315–330.