Appears: Workshop on Multimodal Sensors and Mobile Phones for Social Computing, at ICMI-MLMI 2009
Cambridge MA

SocialCircuits: The Art of Using Mobile Phones for Modeling Personal Interactions

Iolanthe Chronis
MIT Media Laboratory
20 Ames St.
Cambridge MA 02139
+001 617 253 5956
Ianthe@media.mit.edu

Anmol Madan
MIT Media Laboratory
20 Ames St.
Cambridge MA 02139
+001 617 253 5956
anmol@media.mit.edu

Alex (Sandy) Pentland
MIT Media Laboratory
20 Ames St.
Cambridge MA 02139
+001 617 253 3818
pentland@media.mit.edu

ABSTRACT

We describe SocialCircuits, a platform capable of measuring the face-to-face and phone-based communication network of a real-world community. This platform uses commodity mobile phones to measure social ties between individuals, and uses long and short term surveys to measure the shifts in individual habits, opinions, health, and friendships influenced by those ties.

We also describe the flagship experiment using this platform, a year-long study of an entire university undergraduate dormitory. Lastly, we discuss some of the key challenges we met in building and deploying the platform, including mobile phone hardware and software selection, privacy considerations, community selection and recruitment, and techniques for minimizing data loss.

Categories and Subject Descriptors
H.0 [Information Systems], I.2.1 [AI Applications]

General Terms: Algorithms, Measurement, Human Factors

Keywords: Experiment Design, Social Network Analysis, Mobile Computing

1. INTRODUCTION

The most important interactions that we have in our lives with others are those that occur face-to-face. In the past, specialized electronic sensors and badges have been used to measure this face-to-face interaction. For instance, the Sociometric badge [6] was designed to identify human activity patterns, analyze conversational prosody features and wirelessly communicate with radio base-stations and mobile phones. Sensor data from these badges has been used in various organizational contexts to automatically predict employees' self-assessment of job satisfaction and quality of interactions.

ICMI-MLMI'09, Workshop on Multimodal Sensor-Based Systems and Mobile Phones for Social Computing. November 6, 2009 Cambridge MA, USA.

Cell phones are a ubiquitous and natural part of our modern social lives. They provide a convenient tool for measuring social connectivity features related to phone calls and text messages. Sensor-enhanced smart phones can even determine personal location and nearness of friends. Several recent projects have used pervasive, mass-market mobile phones as active social sensors. Eagle and Pentland [1] coined the term Reality Mining, and used mobile phone Bluetooth transceivers, phone communication logs, and cellular tower identifiers to identify the social network structure, recognize social patterns in daily user activity, infer relationships, identify socially significant locations, and model organizational rhythms.

We have built a scalable and reusable platform that transforms smart phones into an advanced social sensor capable of capturing the relationships and influences within a dense community. This platform is needed because there are not many such data sets for the research community to use. Experiments of this type have a high engineering and deployment costs, and researchers often do not have the skills to build and deploy the platforms.

During the academic year of 2009-2010, we deployed our platform on an undergraduate community. During the course of the experiment we collected over 3 million co-location samples, over 60,000 phone calls, and over 20,000 text message samples, forming the world's largest dataset capturing face-to-face diffusion and social influence behaviors [3, 4]. In this paper, we describe our platform capabilities, illustrate the new types of research questions it can help answer, and provide guidelines for other researchers interested in following our approach.

2. SYSTEM CAPABILITIES

In this section, we describe our data-collection platform designed to collect dense, long-term social interaction data in naturally-occurring communities. Our platform is based on the Windows Mobile 6.x mobile operating system. Participants use experiment devices as their primary mobile phones by transferring their existing voice plans.

This platform below has been successfully tested with all four major US mobile operators, and 6 different smart-phone handsets. The source code for this Windows Mobile platform is available online (http://mob.media.mit.edu). The specific capabilities of our data collection platform are outlined below:

Detect Bluetooth wireless devices in proximity: Bluetooth and other wireless-radio based co-location techniques have been used to identify the nodes and edges in the social network [1].

Detect Wi-Fi (WLAN 802.11b) access point identifiers: Since most urban areas have a high density of Wi-Fi access points, these identifiers can be used to infer homogeneity and entropy of location and proximity patterns, e.g. is there a cluster of users who tend to visit similar locations frequently?

Capture Phone and SMS logs: The temporal and frequency features extracted from communication logs can be used to infer strength and type of social connection.

Background Scan Manager: This component initiates background scans for Bluetooth and Wi-Fi at user-specified periods. The default time interval between scans is 5 minutes. As discussed in the section 4, the choice of scan interval also impacts the phone's battery life. The scan manage module also ensures that Bluetooth and WLAN 802.11 radios on the phone are activated as necessary.

On-Device Survey Launcher: The platform supports launching single-screen daily surveys, as soon as subjects turned their phones on in the morning. During four months of our study, this survey application asked questions about common flu symptoms that might have been exhibited over the previous 24 hours.

On-Device User Feedback Engine: Daily surveys also had the ability to pull images and data from the central server, allowing use of feedback about other subjects' responses or changing world conditions. For two months of our study, we used this functionality to learn about the effect of social feedback on daily health habits. An example is shown in Figure 1.

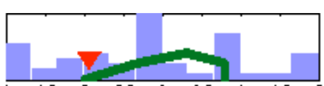

Fig 1. *This image was included in an on-device health survey. The red arrow represents the past responses of subject, the green represents the subject's closest friends (as determined by experimental data) and the blue represents the entire community.*

Over-the-Air Application Updater: The platform supports a native updater application that is designed to fetch compressed installer files (.cab format) from a remote server. This component enables seamless remote deployment of software updates, bug fixes, and new experiment modules. The client downloads updates opportunistically, when 802.11 WLAN access and adequate disk-space are available.

Custom Music Player: In order to study the diffusion of music, a custom music player allows participants to play, share, rate and search through the music library. Participant in our deployments have access to over 1500 independent music tracks of different genres, sourced under the Creative Commons license. All client application events are logged on the server.

Flexible Integration with Web-Based Surveys: Our platform integrates with commercial, off-the-shelf web tools for launching surveys for a large number of participants. A flexible, python-based back-end post-processing infrastructure efficiently parses through exported comma-separated value (CSV) survey responses, and generates detailed MySQL tables per user, used for statistical analysis. In the experiment described in section 3, monthly questionnaires spanning tens of pages were used to collect training labels related to subjects' relationships, health, music tastes, political opinions, attitudes, etc.

3. EXAMPLE APPLICATION: THE ADOPTION OF POLITICAL OPINIONS

This mobile platform was successfully deployed in a real-world setting with an university undergraduate dormitory for an academic year, with a total of over sixty-five participants. The participants represent eighty percent of the total population of the dormitory—the remaining twenty percent of residents declined to participate in this study citing privacy concerns. The undergraduate dormitory is known for its pro-technology orientation and tight-knit community.

The research goal of this experimental deployment was to model the adoption of opinions and social behaviors within this community. Literature across many social sciences suggests that our opinions and behaviors diffuse over our existing social networks. However, to date there has been no method to automatically capture fine-grained social interactions between people and then use the data to better model the diffusion process. It has been shown that surveys and other human-intensive methods are not scalable and can be highly inaccurate due to recollection biases.

In addition to mobile phone data, as training labels, political opinions were sampled in this community three times – September, October, and then November (within 3 days of the presidential election). The survey was designed as a Likert scale and included the following questions:

- *'Are you liberal or conservative?'*
  7-point scale, from 'extremely conservative' to 'extremely liberal'

- *'How interested are you in politics?'*
  4-point scale, from 'not interested' to 'very interested'

- *'What is your political party preference?'*
  7-point scale, from 'strong Democrat' to 'strong Republican'

Using our platform, it is possible to model the dynamic exposure to different opinions for every individual. Contact between two individuals is a function of different mobile phone features— e.g. time spent together during the day or in classes, time spent socializing in the evenings or late at night, phone calls and SMS's exchanged, the count of interactions or the total duration of interaction. It is possible to estimate two types of exposure based on these mobile phone features:

Normalized exposure represents the average of all opinions a person is exposed to on a daily basis, *weighted by the amount of exposure* to different individuals and their self-reported opinions. *Cumulative exposure* represents the magnitude of a particular opinion that a person is exposed to on a daily basis, and is a function of the *amount* of contact with different individuals and their self-reported opinions.

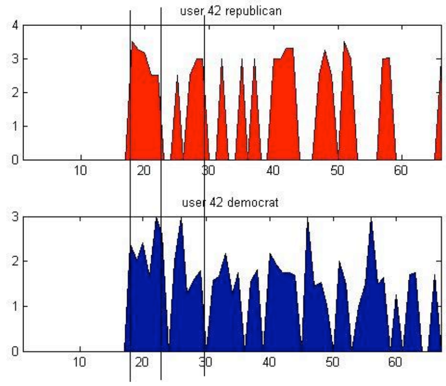

Fig 2. *Normalized face-to-face exposure calculated for one individual over 2 months. The upper graph (red) represents exposure to republicans and the lower graph (blue) represents exposure to democrats. The vertical axis represents the intensity of opinion (i.e. 0 = independent, 1= slight democrat, 2= democrat, 3=strong democrat, and similarly for republicans). The horizontal axis is time over days. Overlaid vertical lines represent the days of the election debates and the day of presidential elections (4th Nov).*

Modeling exposure to different opinions allows us to predict whether a person is likely to adopt or reject new opinions. Accounting for the normalized exposure to different opinions helps predict an individual's future opinions better than using his/her past opinions and community-wide influences. For the 'interest in politics', 'political party preference' and 'liberal or conservative' questions, factoring in automatically-captured exposure explains an additional 15% variance, additional 9% variance and additional 6% variance respectively (over using past opinions alone). Exposure to different political opinions seems to play a bigger role for freshmen. For freshmen, factoring in 'automatically-captured' exposure improves the explained variance by 22%, 25% and 30% respectively, for the three questions [4].

This example illustrates how our mobile platform can be used to model hitherto immeasurable aspects of human behavior. In the next section, we provide guidelines on how other researchers can devise similar experiments using our tools.

4. GUIDELINES FOR FUTURE REAL WORLD DEPLOYMENTS

In this section, we provide our perspective on key issues that must be addressed by other researchers, based on our lessons learnt over a year of deployment.

4.1 Mobile Phone Sensors

A key decision related to hardware selection is sensing capabilities.

Absolute Location -- GPS, cell tower triangulation, and Wi-Fi-based localization (e.g. Skyhook) are location-sensing technologies with varying resolutions available on mobile phones. Wi-Fi triangulation and signal strength can also be used to calculate location on a room-level resolution. However, due to the 5-minute scan interval in our platform, it is not possible to use popular location tracking techniques like particle filters or Kalman filters [2].

Co-Location (Proximity) – Most mobile phones are equipped with class 2 Bluetooth radio transceivers, which can detect Bluetooth devices within a maximum range of 10m. However, most commodity mobile phones do not provide signal strength information. Also, a device will not be detected unless the Bluetooth radio has to be set to the 'discoverable' mode, which is often disabled by default as a security feature. It is necessary to programmatically turn this feature on.

Context and Activity – 3-Axis accelerometers can be used to detect many physical activities, including walking, running, or sleeping [5], although the classification accuracy is reduced if there is no fixed body position for the mobile phone. Accelerometer data can also be used to detect whether the subject is carrying the mobile phone—if the phone hasn't moved for 13 continuous hours, then it is probably sitting on a desk and not being used by the subject.

4.2 Mobile Phone Usability

In order for subjects to use mobile devices over extended periods of time, the phone platform must be user-friendly. Battery life, physical dimensions and weight of the device, available device RAM, and operating system interface are all key factors that impact long-term usability. Running background scan code every 5 minutes on a phone decreases its battery life by approximately 15-20%, for instance, but it is essential that this decrease does not cause the total battery life to dip below the 16-hour level, where a fully charged phone will power off before the end of the day. Similarly, users have different preferences for phone ergonomics--people tend to leave heavy and bulky devices at home, instead of carrying them all the time. If a device has low RAM, when the phone launches scanning process every 5 minutes, it may slow down applications and UI response time—an effect seen with both iPhone and Windows Mobile devices. Finally, the OS itself plays a key role in the usability of the device, and subsequently, the level of engagement participants have in the long-term experiment. iPhone and Android devices have easy-to-use interfaces with thousands of available applications.

4.3 Platform Openness and Cross-Carrier Support

Today, not all phone platforms are equally open—the programmer often does not have access to the sensors and features that are required. For instance, an iPhone implementation of our code can only run on jailbroken iPhones due to the inability to run background processes with the official Apple SDK. In our case, although Windows Mobile devices did not have a compelling interface, they did have the platform API support and cross-carrier compatibility (both GSM and CDMA) required for a high-density deployment. In the future, we expect that both Android and iPhone devices will support CDMA operators, which would make them suitable for experiments of this nature. For initial deployment, we bought mobile devices and preloaded them with certificates and applications but this approach is not scalable past ~200 phones. In the future, we expect that it will be possible to deploy large-scale experiments by the installation of a single application through an AppStore.

4.4 Community Selection

For an experiment to measure social diffusion in a tight-knit community, participant density is more important than volume. In our case, it was important to account for as many factors as possible to why a particular subject changed their behavior, and hence necessary to capture as many of that subject's friends and acquaintances as possible

For this experiment, we chose a small dorm of around 90 students that was physically and socially distant from the rest of the campus. Over 65% percent residents reported that a majority of their friends and acquaintances lived inside the dorm. Other considerations might be how often the members of the community carry their cell phones, and whether they have concerns about privacy.

4.5 Community Preparation

Even after IRB approval, a few of our potential subjects were hesitant about participating in the experiment due to privacy concerns. Ironically, it was the very traits that we selected the community for that affected this behavior—high social cohesion and co-influence within the community meant that when a few people had concerns, those concerns spread to their network. The most important potential privacy concern in this community was participant re-identification based on mobile phone data and survey responses. To alleviate potential privacy concerns, steps were to taken to hash, anonymize and remove any personal identifiers from the data at each stage of the collection process.

Another concern was the social pressure applied to participate. The high density required for participation, coupled with the sensitivity of the questions asked during the monthly survey, made a few potential participants uncomfortable. We alleviated this concern by decoupling the survey responses from the act of participation—answering surveys became a lucrative component that was not mandatory for participation in the experiment. All of the subjects filled out the surveys anyway, and were more comfortable about doing so.

4.6 Long-Term Reliability

With any long-term experiment using complex equipment, there may be unforeseen technical issues. It is important to make sure that technical issues are rapidly addressed, without any significant data loss. Below are practical suggestions from our experience.

Get help on the ground: The experimenter cannot be around all the time to identify the various kinds of issues that tend to come up. They also may have trouble tracking down individual subjects who have software or hardware issues. We hired two of our subjects to help solve simpler technical issues and be a resource within the dorm for experiment-related questions.

Encourage participants to report problems: As much as we asked for bug reports, the subjects' time constraints and aversion to 'annoying' us would mean that a problem would go undetected for days, even if several subjects were encountering the same difficulty. This issue went away when we offered a 'bounty' on bugs—first report of a certain issue was paid $10.

Backup on-device data at multiple locations: In addition to syncing new data with the server every 24 hours, data was also backed-up on the phone so that every month or two, any missing data could be collected.

Use minimal number of different handsets: Even phones that run the same operating system may have crucial differences in the specifics of file system layout, registry keys, and sensor performance. Supporting more than one model inflates development time and may introduce costly bugs.

Avoid on-device flash memory: We used external memory cards to store the data. However, there is a finite number of writes to disk that are advisable for flash memory before risk of disk corruption. Writing to disk every 5 minutes may cause the disk to fail catastrophically without warning, after about 3 months.

Pay subjects incrementally: Paying subjects one dollar per day that they filled out on-phone surveys yielded a high response rate.

5. ACKNOWLEDGMENTS

We would like to thank our UROP engineering team—Xiao Fan, Griffin Chronis, Gabe Tobon, Mu He, Christopher Palmer, Sylvester Ohienmhen, and Paul Kominers. We would also like to thank Dr. David Lazer for assistance with experiment design.

6. REFERENCES

[1] Eagle, N., and Pentland, A. 2006. Reality Mining: Sensing Complex Social Systems. In Personal and Ubiquitous Computing, Vol 10, #4, 255-268.

[2] Liao, L, Patterson, D.J., Fox, D., and Kautz, H. 2007. Learning and Inferring Transportation Routines. In Artificial Intelligence, 171(5-6), 2007, 311-331.

[3] Madan A., and Pentland A. 2009. Modeling Social Diffusion Phenomena Using Reality Mining. In AAAI Spring Symposium on Human Behavior Modeling. Palo Alto, CA.

[4] Madan A., and Pentland A. 2009. Social Sensing to Model the Evolution of Opinions. In submission.

[5] Munguia Tapia, E., Intille, S., and Larson, K. 2007. Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. In Proceedings of the 11th International Conference on Wearable Computers (ISWC '07). Boston, MA.

[6] Olguin, D., Waber, B., Kim, T., Mohan, A., Ara, K., Pentland, A. 2009. Sensible Organizations: Technology and Methodology for Automatically Measuring Organizational Behavior. In IEEE Transactions on Systems, Man, and Cybernetics Part B. Vol. 29, February 2009.