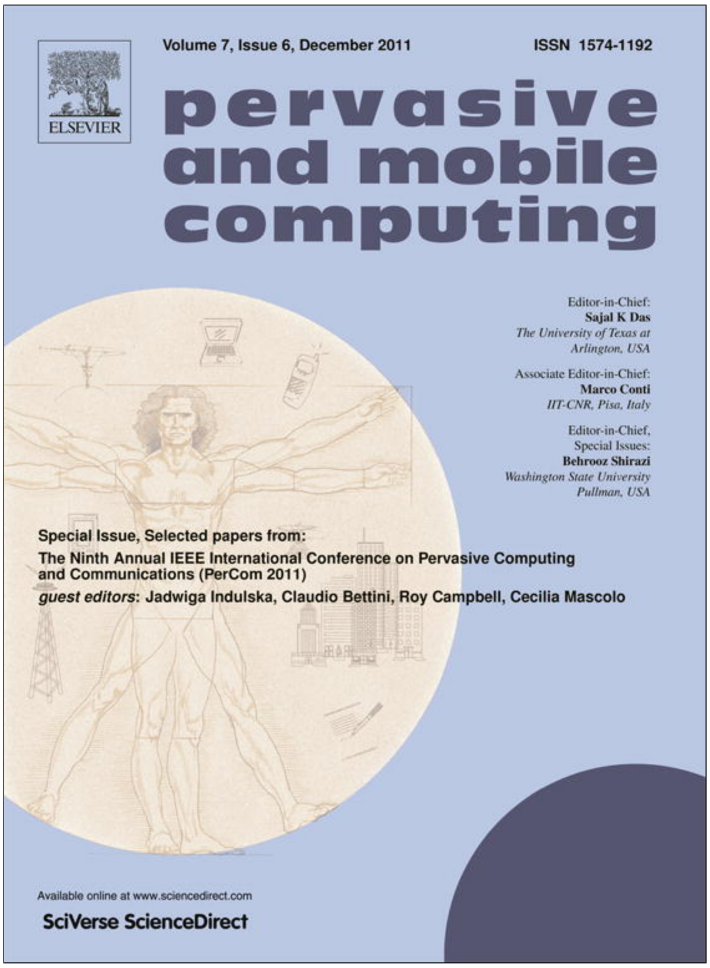
Volume 7, Issue 6, December 2011
ISSN 1574-1192
ELSEVIER
pervasive and mobile computing
Editor-in-Chief:
Sajal K Das
The University of Texas at Arlington, USA
Associate Editor-in-Chief:
Marco Conti
IIT-CNR, Pisa, Italy
Editor-in-Chief, Special Issues:
Behrooz Shirazi
Washington State University
Pullman, USA
Special Issue, Selected papers from:
The Ninth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom 2011)
guest editors: Jadwiga Indulska, Claudio Bettini, Roy Campbell, Cecilia Mascolo
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Pervasive and Mobile Computing 7 (2011) 643–659

Contents lists available at SciVerse ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Social fMRI: Investigating and shaping social mechanisms in the real world

Nadav Aharony [a], Wei Pan [a], Cory Ip [a], Inas Khayal [a,b], Alex Pentland [a,*]

[a] The Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

[b] Masdar Institute of Science and Technology, Abu Dhabi, UAE

ARTICLE INFO

Article history:
Available online 25 September 2011

Keywords:
Social computing
Mobile sensing
Rich data
Social health
Mobile health

ABSTRACT

We introduce the Friends and Family study, a longitudinal living laboratory in a residential community. In this study, we employ a ubiquitous computing approach, Social Functional Mechanism-design and Relationship Imaging, or Social fMRI, that combines extremely rich data collection with the ability to conduct targeted experimental interventions with study populations. We present our mobile-phone-based social and behavioral sensing system, deployed in the wild for over 15 months. Finally, we present three investigations performed during the study, looking into the connection between individuals' social behavior and their financial status, network effects in decision making, and a novel intervention aimed at increasing physical activity in the subject population. Results demonstrate the value of social factors for choice, motivation, and adherence, and enable quantifying the contribution of different incentive mechanisms.

1. Introduction

Imagine the ability to place an imaging chamber around an entire community. Imagine the ability to record and display nearly every facet and dimension of behavior, communication, and social interaction among the members of the said community. Moreover, envision being able to conduct interventions in the community, while measuring their effect—by both automatic sensor tools as well as qualitative assessment of the individual subjects. Now, think about doing this for an entire year, while the members of the community go about their everyday lives.

Utilizing pervasive computing devices and methodologies, we developed a mobile-phone-centric social and behavioral sensing system that we have deployed with 130 adult members of a young-family living community for over a year now. During this year we have amassed what is, to the best of our knowledge, an unprecedented longitudinal dataset, which we dub the *Friends and Family* dataset. The dataset includes continuous collection of over 25 phone-based signals—including location, accelerometry, Bluetooth-based device proximity, communication activities, installed applications, currently running applications, multimedia and file system information, and additional data generated by our experimental applications. In addition, we collect financial information through receipts and credit card statements, logs of Facebook socialization activities, daily polling of mood, stress, sleep, productivity, and socialization, as well as other health and wellness related information, standard psychological scales like personality tests, and many other types of manually entered data by the participants.

The data enable us to construct multiple network modalities of the community—such as the phone communication network, physical face-to-face encounters network, online social network, self-reported network, and more. We use these networks to investigate questions like how things spread in the community, such as ideas, decisions, mood, or the seasonal

* Corresponding author.
E-mail addresses: nadav@media.mit.edu (N. Aharony), panwei@media.mit.edu (W. Pan), coryip@media.mit.edu (C. Ip), ikhayal@media.mit.edu (I. Khayal), pentland@mit.edu, pentland@media.mit.edu (A. Pentland).

1574-1192/$ – see front matter
doi:10.1016/j.pmcj.2011.09.004

flu. Our high level goals include the investigation of "natural" and externally imposed social mechanisms related to behavior and decision making, together with designing and evaluating new mechanisms or tools for helping people make better decisions.

In this paper we describe three investigations conducted during the longitudinal study: The first deals with observations on *individual* social behavior. We use individuals' financial status and social behavior to provide new insights on the question raised by Eagle et al. [1] about the causality between the two. The second investigation looks at *network effects* on decision making by looking at how different social relationships might predict the spread of mobile applications. The third component we report on relates to the design and execution of experimental *interventions* while measuring their effect on individual and group behavior. Such interventions are a major component of the study and the Social fMRI approach, thus it accordingly receives prominence in the current discussion.

Out of several interventions conducted over the past year and planned for the upcoming months, in this report we focus on a fitness and physical activity intervention conducted between October to December of 2010. Using an experimental intervention within the Friends and Family study population, we test social mechanism-design principles. In particular, we propose a novel social mechanism in which subjects are rewarded based on their peers' performance and not their own. Results suggest that: (1) Social factors have an effect on the physical activity behavior, motivation, and adherence over time. (2) Social incentives, and particularly our novel Peer-Reward mechanism encouraging social influence among participants, support higher activity returns per dollar invested in the system. (3) Finally, results support the notion of a complex contagion [2] like effect related to pre-existing social ties between participants.

The contributions of the work described in this paper are as follows: We present the Friends and Family study and very rich dataset; We describe the field-proven system that has been deployed and used for over a year, which we intend to release as an open source platform for social and behavioral data collection and feedback; We conducted the fitness intervention and find results that contribute to our understanding of social incentives and motivation in real-world; We present analysis that shows that individuals' social interaction diversity correlates with their current income level, suggesting a contradictory social theory to the currently prevailing theory. Finally, we present analysis that shows a relationship between the number of mobile applications that two people share in common to the time they physically spend face-to-face. Our observations suggest that the diffusion of apps relies more on the face-to-face interaction ties than on self-perceived friendship ties.

In the remainder of this paper, we first review related literature and the contextual underpinning of our proposed vision (Section 2). We then introduce our approach (Section 3). In Section 4 we go into our methodology—the Friends and Family living laboratory and its characteristics. Next we describe our system architecture (Section 5). In Sections 6–11 we provide analysis and discussion that demonstrate the potential of the study dataset and the Social fRMI approach.

2. Related work and context

2.1. Ubiquitous social observatories

In recent years the social sciences have been undergoing a digital revolution, heralded by the emerging field of "computational social science". Lazer, Pentland, et al. [3]. describe the potential of computational social science to increase our knowledge of individuals, groups, and societies, with an unprecedented breadth, depth, and scale. [3] highlights challenges in terms of scientific approach for observation and intervention when dealing with real people in their living environments, including issues of subject privacy, monitoring, and altering of environments during the discovery process.

Fig. 1 gives a high-level qualitative overview of social observatories and datasets, comparing them along axes of sample size, duration, and a very rough notion of "throughput" or the amount of information in the datasets. The idea is that dataset throughput is a function of the data dimensionality (number of different signals collected), its resolution (e.g. raw or aggregate), sampling rate (how often data is collected), and unique information in it (an accelerometer sensor lying on a desk for a week does not collect a lot of information). This diagram illustrates the potential of ubiquitous technologies for the design of "social observatories" and the collection of very rich datasets. At the bottom of the diagram are traditional sociology studies as well as many of the corporate "donated" datasets. Leading traditional datasets include, for example, the Framingham Heart Study [4], which stands out for its duration and a subject pool of several thousands, however its "throughput" is low as subjects were sampled approximately once in three years.

The pervasiveness of mobile phones has made them ubiquitous social sensors of location, proximity and communications. Because of this, mobile phone records from telecom companies have proven to be particularly valuable. For example, Gonzales et al. show that cell-tower location information can be used to characterize human mobility and that humans follow simple reproducible mobility patterns [5]. Eagle et al. find that the diversity of individuals relationships is strongly correlated with the economic development of communities [1]. These and other corporate "donated" datasets are usually characterized by having, on one hand, information on very large numbers of subjects, but on the other hand, this information is constrained to a specific domain (email messages, financial transactions, etc.), and there is very little, if any, contextual information on the subjects themselves. This is why, although their sample size may be in the millions, they are relatively low on the throughput axis in Fig. 1). As example, in [6] each sampling point was an aggregated 15-day call summary of anonymous phone users. In addition, domain-limited results might be harder to generalize for the physical world, as discussed by Onnela et al. in context of Facebook data [7]. Finally, there is the "*offline*" nature of most existing datasets which

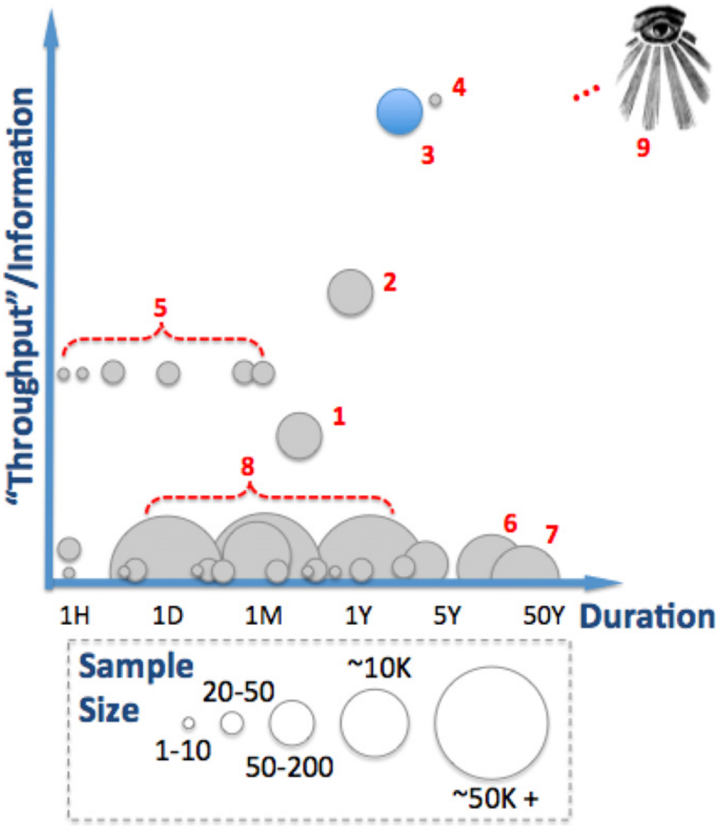

Fig. 1. Qualitative overview of social science "observatories" and datasets, along axes of data collection duration, qualitative"throughput", and the size of the subject sample. (1) Reality Mining [8], (2) Social evolution [11], (3) Friends and Family dataset, (4) Rich-data pioneers [15,16], (5) Sociometric Badge studies [14], (6) Midwest field station [17], (7) Framingham Heart Study [4], (8) Large call record datasets [5,1,6], (9) "Omniscient"/all-seeing view.

are based on previously collected data, making it harder to test cause and effect using these datasets. Nevertheless, these datasets are yielding significant contributions to our understanding of society, and one might imagine that by increasing the dimensionality and throughput, such datasets could lead to even further advancements.

An alternative approach is a bottom-up one, of collecting data at the level of the individual. Eagle and Pentland [8] defined the term"Reality Mining" to describe collection of sensor data pertaining to human social behavior. They show that using call records, cellular-tower IDs, and Bluetooth proximity logs, collected via mobile phones, the subjects' social network can be accurately detected, as well as regular patterns in daily activity [8,9]. This initial study was then expanded in Madan et al. [10], who conducted a similar experiment and show that mobile social sensing can be used for measuring and predicting the health status of individuals based on mobility and communication patterns. They also investigate the spread of political opinion within the community [11]. Other examples for using mobile phones for social sensing are those by Montoliu and Gatica-Perez [12] and Lu et al. [13]. Most of these were of an observational nature, and have not performed controlled experimental interventions for exploring social mechanism. Other types of sensor-based "social observatories" are the Sociometric Badges by Olguin et al. that capture human activity and socialization patterns via a wearable sensor badge [14]. A key aspect of the Sociable Badges is that they have been deployed in studies where sensor feedback was given to the corporate participants [14].

2.2. Physical activity sensing and feedback

In this paper we focus on a specific problem from the domain of health and wellness: Studies have shown a great increase in obesity and related chronic medical conditions over the last several decades. Physical activity has been shown to help alleviate the burden of obesity and other health conditions [18,19]. Over the past two decades, the accelerometer has been established and refined as a tool for tracking physical activity [20–22]. Accelerometry-based sensors have been found to provide more accurate estimates than other widely-used proxies for energy expenditure [21]. Although there is some error associated with using accelerometers to track energy expenditure in free-living situations, a significant relationship between accelerometer output and energy expenditure has nevertheless been established [20]. Several studies in the ubiquitous computing literature have targeted this important problem domain. Ubifit [23] is one of the most extensive works investigating ways to encourage physical activity. Additional projects include "Fish'n'Steps" [24] and "Houston" [25], among others. These works investigate diverse aspects of the problem, such as user interface, goal setting, or techniques for using the accelerometer for accurate expenditure measurement and activity detection.

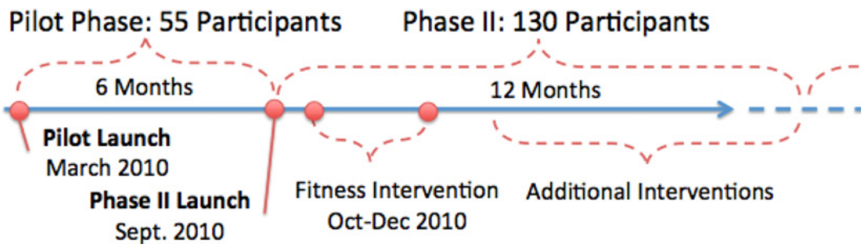

Fig. 2. High level timeline for the Friends and Family study.

Of particular relevance are those studies that involve social components [24–28]. It has long been established that social support is a resource for behavioral change and an indicator for health [29], however here is still much to be learned about the fine-grained *social mechanisms* related to physical activity behavior, as well as how to leverage such insights in *designing better socially-aware interventions and mechanisms* for encouraging healthy behavior change.

For activity measurement, relevant works are those using unaugmented phone-based activity detection [26,30], whereas the majority of studies to date used additional measurement devices that need to be carried by subjects. In the consumer world, a growing number of activity measuring mobile applications such as CardioTrainer [31] use the phone's accelerometer, combined with visualization and other feedback to help users increase their physical activity levels. Most applications aim to provide a step count measurement, and ask the users to hold the phone in a certain orientation while exercising in order to deliver accurate measurements.

3. The social fMRI

In the medical realm, Magnetic Resonance Imaging, MRI, is considered one of the most comprehensive diagnostic techniques available, and Functional MRI, fMRI, is one of the leading tools used for studying the brain through response to carefully designed stimuli. Analogously, we define Social Mechanism-design and Relationship Imaging, or *Social MRI*, which allows detailed sensing and imaging of social systems through the use of mobile phones, credit cards, social media, and telecommunications for social and behavioral sensing platform. *Social fMRI* takes it a step further—allowing for specifically designed stimuli and interventions to the social system, while measuring the individual and collective response. Just as fMRI helps us understand the inner workings of the brain, we hope that the Social fMRI approach could help us understand the inner workings of social systems and the way humans interact and react to one another. More than just an aspiration, in this paper we show a proof of concept as to how this could be done.

The general framework of the Social fMRI idea is a combination of a longitudinal living-laboratory/social-observatory type of study, coupled with a supporting system infrastructure that enables the sensing and data collection, data processing, and also a set of tools for feedback and communication with the subject population. The Social fMRI implements and extends the ideas of the Reality Mining approach [8], by (1) adding much greater data richness and dimensionality, combined with (2) a strong element of active interaction and carefully designed experimental stimulation of the study population.

4. Methodology

4.1. Living laboratory: the "Friends and Family" community

Starting March 2010, we initiated a living laboratory study conducted with members of a young-family residential living community adjacent to a major research university in North America. All members of the community are couples, and at least one of the members is affiliated with the university. The community is composed of over 400 residents, approximately half of which have children. The residence has a vibrant community life and many ties of friendship between its members. We shall refer to this residence as the "*Friends and Family*" community.

This study involves a relatively different subject population when compared to previous ubiquitous computing observatory studies. For example, colleagues and co-workers in Reality Mining [8], and undergraduates in [10]. The Friends and Family community includes a much more heterogeneous subject pool, and provides a unique perspective into a phase in life that has not been traditionally studied in the field of ubiquitous computing—married couples and young families.

As depicted in Fig. 2, a pilot phase of 55 participants launched in March 2010. In September 2010 phase two of the study included 130 participants, approximately 64 families. Participants were selected out of approximately 200 applicants, in a way that would achieve a representative sample of the community and sub-communities. One of the reasons for keeping the number below 150 is that these numbers fit well with Dunbar's social evolutionary theory regarding the number of people humans are able to maintain a relationship with [32]. Throughout the study we ask about social closeness between all participants in the study, and numbers larger than Dunbar's number could become quite tedious for subjects. We refer to experiments in our scale as "Dunbar scale" experiments. The research goals of the longitudinal study touch on many aspects of life, from better understanding of social dynamics to health to purchasing behavior to community organization. The two high-level themes that unify these varied aspects are: (a) how people make decisions, with emphasis on the social aspects involved, and (b) how we can empower people to make better decisions using personal and social tools.

4.2. Study data collection

One of the key goals of the Social fMRI idea is the collection of multi-modal and highly diverse range of signals from the subject population. We wanted to gather data on numerous network modalities, so that their properties and interrelation could be better understood. We applied a user centric, bottom up approach utilizing the following components:

4.2.1. Mobile phone sensing platform

This is the core of the study's data collection. Android OS based mobile phones are used as in-situ social sensors to map users activity features, proximity networks, media consumption, and behavior diffusion patterns. The mobile phone platform is described in more detail in the next section. We did not sponsor phone plans or data plans—users received a mobile phone that fit their desired provider, and they were responsible to port their existing account to it or open a new account. The condition was that the study phone be their primary phone for the duration of the study.

4.2.2. Surveys

Subjects complete surveys at regular intervals, combining web-based and on-phone surveys. Monthly surveys include questions about self perception of relationships, group affiliation, and interactions, and also standard scales like the Big-Five personality test [33]. Daily surveys include questions like mood, sleep, and other activity logging.

4.2.3. Purchasing behavior

Information on purchases is collected through receipts and credit card statements submitted at the participants discretion. This component targets categories that might be influenced by peers, like entertainment and dining choices.

4.2.4. Facebook data collection application

Participants could opt to install a Facebook application that logs information on their online-social network and communication activities. About 70% of subjects opted to install.

4.3. Subject protection and privacy considerations

The study was approved by the Institutional Review Board (IRB) and conducted under strict protocol guidelines. One of the key concerns in the design of the study was the protection of participant privacy and sensitive information. For example, data is linked to coded identifiers for participants and not their real world personal identifiers. All human-readable text, like phone numbers and text messages are captured as hashed identifiers, and never saved in clear text. Collected data is physically secured and de-identified before being used for aggregate analysis. A second important consideration was to be as unobtrusive as possible to the subjects' life routines. Participants are able to keep the phone at the end of the study. For mandatory out-of-routine behavior that asked of participants, like filling out surveys, subjects are compensated (e.g. $10 for completing the monthly survey). Participation in intervention or sub-experiment on top of the main study components is completely optional, and interventions are carefully designed with the interests of the participants in mind.

5. System architecture

5.1. Mobile phone platform

The phones run our software platform, which periodically senses and records information such as cell tower ID, wireless LAN IDs; proximity to nearby phones and other Bluetooth devices; accelerometer and compass data; call and SMS logs; statistics on installed phone applications, running applications, media files, general phone usage; and other accessible information. Over 25 different types of data signals are currently collected. The system also supports integration of user-level apps, like an alarm clock app we developed, for additional data collection and interventions. Additionally, the phone system includes a survey application. Sample screenshots are shown in Fig. 3. The system continuously runs as a background service, and has a set of triggers that make sure it restarts when the phone turns on or after the service is terminated. The main service is responsible for scheduling the different data collection actions. The configuration is set so that battery-intensive actions (e.g. GPS scans) are performed in intervals allowing usefulness while minimizing battery drain. A remote configuration capability allows for fine-tuning the system, with a goal of enabling a minimum of 16 h between charges. We are working toward releasing the software, named "Funf", as an open source framework [34].

5.1.1. Probes

Each type of signal collected by the system is encapsulated as a conceptual "probe" object. The *probes* terminology is used rather than "sensors" as probes include traditional sensors such as GPS or accelerometer, but also other types of information

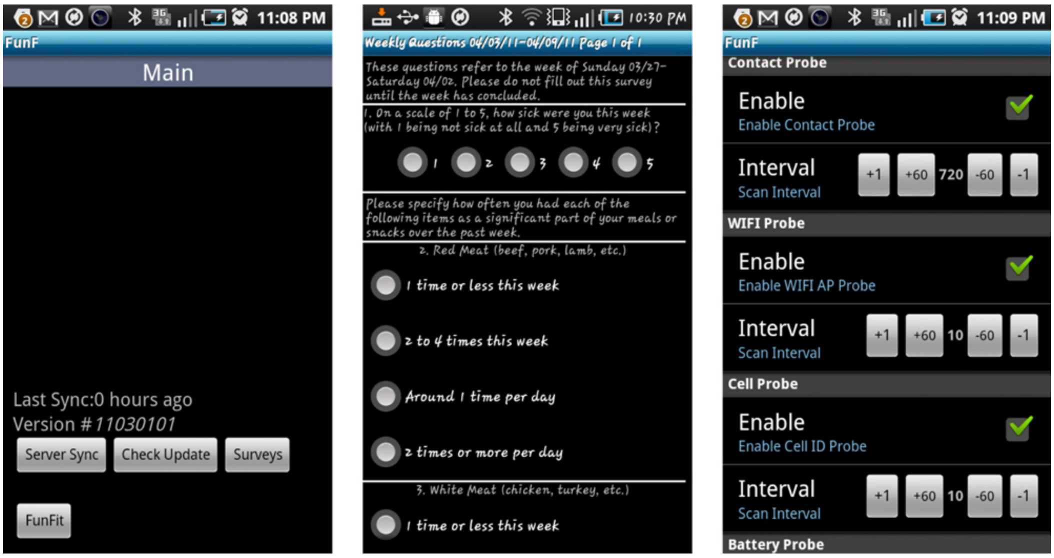

Fig. 3. Sample screenshots: Sync-state and version display (left), survey (center), and probe preferences debug screen.

Table 1
Signals collected via the mobile phone platform during the Friends and Family study. *Interval* is the maximum time between consecutive scans. N/A is marked for probes where interval definition is not applicable. *Opportunistic* describes whether the probe also uses an opportunistic strategy.

Signal	Interval	Opportunistic?	Notes
Wifi scan	5 m	Y	Access-points in range
Bluetooth scan	5 m	N	Bluetooth devices in range
Cell tower scan	5 m	Y	Current cell tower ID
GPS location	N/A	N[b]	2-states: 15 m outdoors/30 m indoors
Accelerometer	2 m	N	Detailed in Section 9.2
Installed aps	12 h	N	Currently installed applications
Running apps	30 s	N	Currently running applications
Call log	12 h	N	Event log and statistics (hashed)
SMS log	12 h	N	Event log and statistics (hashed)
Contact list	12 h	N	Periodic for tracking changes
Power state[a]	N/A	Y	Signals related to phone power (battery, charging state, etc.)
Screen state	N/A	Y	Triggered on state change (on/off)
Media scans[a]	24 h	N	Set of scans for media content (video/audio/images)
File scans[a]	7d	N	Set of file/directory scans.
Browser scans[a]	12 h	N	History and bookmark scans (hashed)
Alarm clock	N/A	Y	App that logs alarm clock usage
Phone info[a]	3 h	N	Set of signals on phone state (e.g. device ID, os version, timezone, etc.)
Network info[a]	3 h	N	Set of signals on mobile network state (e.g. current operator, data mode, etc.)
Funf info	3 h	Y	Info on installed Funf software version
Probe config	3 h	Y	Snapshot of probes' current setup

[a] Set of multiple signals.
[b] Opportunistic in current software version.

not traditionally considered as collected by sensors, like file system scans or logging user behavior inside applications. Using the modular probe architecture, it is very easy to add new probes to the system, or swap existing probes with an improved version. All probes support a common set of behaviors, and each defines a set of configuration parameters that control it, and the format of its output. Probes can be configured locally on the device or remotely through the back-end server. There are two main ways strategies for implementing and operating probes: A *proactive* probe strategy explicitly requests data to be collected at a certain time, and might need to turn on phone resources (e.g. turn GPS on if it was off), which might add direct battery costs. Probes supporting this strategy usually include a definition for a periodic execution, with a max interval between executions. An *opportunistic* strategy registers the probe as a listener for collecting different messages sent as broadcast within Android. These could be built-in messages like battery state or screen on/off state changes, or 3rd-party custom messages, like an alarm clock app that triggers a message every time the user sets an alarm or presses the "snooze" button. A probe might use either or both of these strategies for its data collection. Table 1 lists the different signals collected by the different probes in the Friends and Family study, with typical interval values. The wifi probe is an example for a probe that uses both strategies—it performs a wifi scan once every five minutes, but it also opportunistically listens to wifi scan results initiated by other processes in the system. The most updated version of the Funf software framework [34] includes a yet greater number of probes in addition to those listed in Table 1.

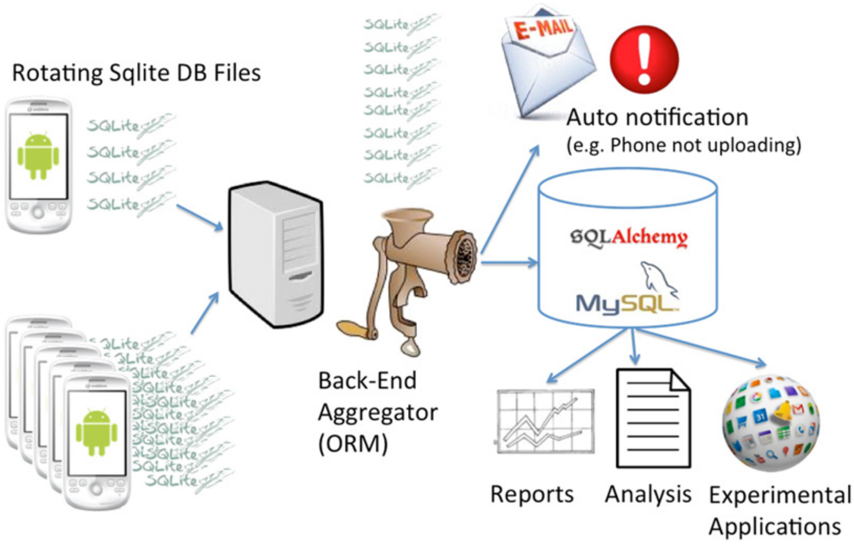

Fig. 4. Overview of back-end data-flow.

5.2. Data formats and server communications

Phone data is saved in SQLite file format, with files rotated every three hours to reduce data loss due to file corruption and to allow periodic data upload to the back-end. Since many participants do not have a mobile data service plan, the system was designed in a "delay tolerant" way: In the absence of network access, the phone accumulates the collected database files locally. When server connection is made (for example a participant connects to WiFi to browse the web), the system attempts to upload files. Once uploaded, files are also encrypted and saved in a backup directory on the phone, until data integrity is confirmed. In addition, the system downloads from the server parameter configurations files that define which data signals to collect, how often, and additional configuration parameters. It also checks for new versions of itself and notifies the user to update when new versions are available. Finally, the system downloads any new surveys that should be presented to the user.

5.3. Back-end

The server-side back-end processes all incoming SQLite files and inserts them into a central MySQL database. It sends email reports to investigators about status of phones and alerts of any issues. Additional services provide data for interventions and personal data visualization for participants. An object-relational-mapper (ORM) enables representing all data as code objects which simplifies development of applications that use the data (see Fig. 4).

6. Demonstrating the social fMRI approach

In the remainder of the article we present analysis and results from several components of the Friends and Family study. These are but initial forays into the study's comprehensive dataset, which serve (or had served) to inform the design of subsequent components and sub-experiments in the longitudinal study, formulate directions for further analysis, and demonstrate the potential of the Social fMRI's data-rich methodology. These components are aligned with the study's high-level goals of understanding social mechanisms related to behavior choices and decision making, as well as designing and evaluating new tools and mechanisms to help people make better decisions. We start by investigating behavior choices based on individual properties in Section 7. In Section 8 we continue with looking at the social fabric by investigating network effects on decisions and choice. The third and most significant discussion relates to ab active intervention in for investigating mechanisms of social support and behavior choices as well as evaluating a proposed social incentive mechanism of our own. The intervention is discussed from Section 9 onward. These examples make use of a variety of signals collected during the study, including physical collocation of participants, self-reported social closeness, app installation patterns, financial information, and physical activity sensing.

7. Individual behavior: social interaction diversity and financial status

The discovery of the strong correlation between social interaction patterns and the wealth and economical development of a community has attracted significant attention [1]. A *current challenge* is to understand the causality of this finding. Researchers tend to believe that a diverse set of social relationships brings benefits such as increased information or external opportunities, among others [35,36]. This approach follows a long line of classical social science literature: Granovetter's weak tie theory [37] and Burt's theory of structural holes [38], to name two.

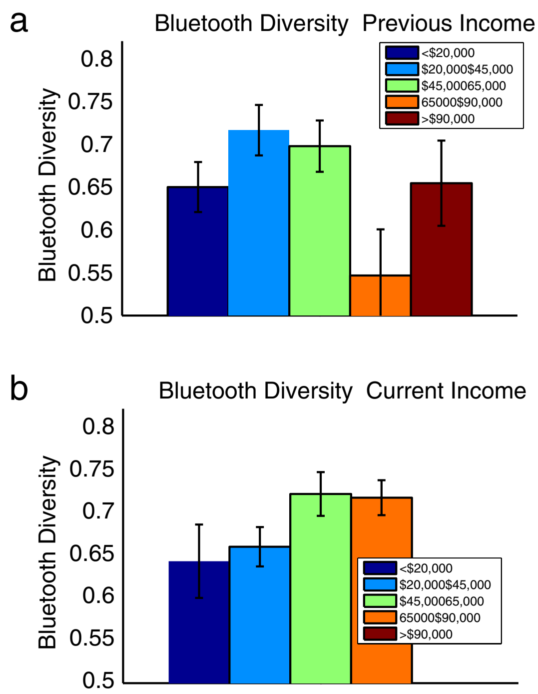

Fig. 5. We show here the mean Bluetooth interaction diversity $D_{\text{call}}(i)$ and its standard error for individuals in different income categories. The top plot is based on previous household income, and the bottom plot is based on current household income. There exists borderline positive correlation between current household coarse income and call diversity ($r = 0.32, p < 0.10$), and the correlation is much stronger within native English speakers in the participant pool ($r = 0.53, p < 0.06$). However, there is no correlation between previous estimated household income and face-to-face interaction diversity ($r = -0.28, p > 0.60$).

The Friends and Family study provides a unique opportunity for investigating this causal relationship. We are able to examine, on an individual-level, relationships between one's financial status (household income) and their interaction diversity by taking both the survey data and the phone sensed data into consideration. The richness of the study also allows us to observe changes in correlation rather than a one-time measure of correlation.

The prevailing causality explanations imply the following reasoning: If successful individuals are suddenly deprived of their incomes, as many participants in this study who left industry jobs to attend graduate school, naturally they will continue to keep their diverse interaction behavior. Their previous success suggests that they understand their social diversity and its benefits, and their future success still relies on their continuous diversity interaction. Since many of them came back to graduate school from relatively high-paying jobs, there are considerable income changes among participants. However, we surprisingly discover that users' social diversity patterns (as defined in [1]) correlate only with their *current* income, as illustrated in Figs. 5 and 6.

Thus, these observations suggest the opposite: Individuals will quickly lose their diversity in social interaction when their financial status gets worse; Individuals will quickly gain their interaction diversity when their financial status improves. We suspect that a stronger behavioral and psychologically oriented mechanism plays an important role in the other direction of causality, that is: Individuals' social diversity patterns are influenced by their financial status. We hypothesize that as good financial status ensures people safer and more satisfying living conditions [39], they naturally feel more confident [40] and secure in exploring new social potential [41,42]. While we provide a new perspective and some supporting evidence for this complicated causality problem, there is still a need for further evidence and experimentation to cross examine our theory as well as other related social theories. In the context of our current discussion, this result demonstrate how Social fMRI types of studies can provide novel perspectives to long-lasting debates in social science.

8. Social fabric and its influence on decision making

Individual decision making is not performed in a vacuum. People are embedded in a social-fabric, and social influence has observed effects on personal choice and behavior. As we set out to design social mechanisms that would support positive and desired behavior change, the Social fMRI approach enables us to gain a better understanding of social-fabric effects on decision making in an uninterrupted setting.

One of the signals collected in the Friends and Family study is the set of mobile applications, or "apps", installed on each phone. Most of these apps are downloaded from the "Android Market", which, around the time of the pilot phase of the study, Spring 2010, had approximately 100,000 apps to choose from (and approximately 450,000 apps in July 2011) [43]. We can treat the act of installing any app on the phone as a decision made by the user, one that is accurately and completely measured by the phone sensing software. Because of the large space of possible apps to install, we can attempt to look at

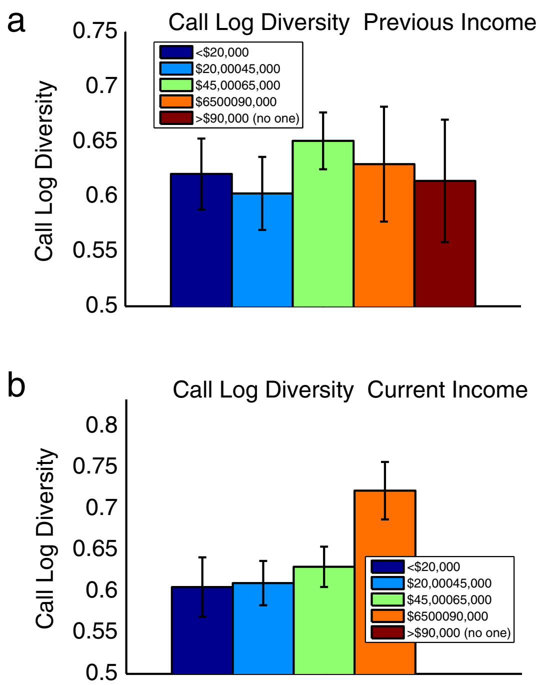

Fig. 6. We show here the mean call diversity $D_{\text{call}}(i)$ and standard error for individuals in different income categories. The top plot is based on reported previous household income, and the bottom plot is based on reported current household income. There exists positive correlation between current household coarse income and call diversity ($r = 0.28$, $p < 0.08$). However, there is no correlation between previous estimated household income and call diversity ($r = 0.003$, $p > 0.80$).

any patterns that link the fact that two (or more) subjects have common apps installed on their phone to various social measures, and attempt to determine any connection between the social ties to this decision making, if such exists. For this analysis we use data from the pilot phase of the study, collected for 55 subjects from March to early July.

8.1. General statistics

During the three months of the measurement, the 55 participants have installed around 870 unique apps (not counting any apps that come bundled with the phone or the OS version). For this analysis, we only look at app installations and ignore un-installations. We first demonstrate statistics for all of the apps in the study: In Fig. 7(a), we plot the distribution of number of users installing each app. We discover that our data corresponds very well with a power-law distribution with exponential cut. This is normal considering we have a limited number of subjects in this phase. We also plot the distribution of number of apps installed per user in Fig. 7(a), which fits well with an exponential distribution, and suggests that most users will only install a limited number of apps. The implication of this finding is that it might be more difficult to promote apps to users if they have already tried many apps previously.

8.2. Network effects

8.2.1. Physical collocation network

We move on to investigate the network effect of app installations in our study community. To begin with, we look at the proximity co-location network of participants, which is inferred by using Bluetooth scan of devices in range. For each pair of users, we counted the number of co-location scans, and used this as a proxy for the actual time that they spend in a physical proximity to each other. For spanning the Bluetooth proximity network used in this analysis, we used data collected over the month of April, which was the period where all participants were physically on campus (March was the month of the university's Spring break, and from May onward some of participants had left for the summer). We removed the recorded Bluetooth counts between midnight and 7am every day, since devices in neighboring apartment might sense one another, which may be incorrectly recognized as social interactions. We generally saw that spouses have over 1000 co-located BT scan events after the removal. In addition to the Bluetooth scan count, for each pair of participants we also counted the number of common apps installed on both phones.

We divided the dataset into two groups by the threshold of 10 Bluetooth counts, which translates to at most 1 h of co-location for the month used in this analysis. Group 1 with Bluetooth hits ranging from 0 to 10,which we assume to be mostly strangers and distant acquaintances. Group 2 are pairs with Bluetooth hits ranging from 10 to 2000. Group 1 has a mean of 2.7253 apps in common between each participant pair. Group 2 has a mean of 4.9 apps in common. A 1-way ANOVA test shows a statistically significant result ($F = 74.48$, $p < 0.0000001$), and a K–S test shows strong significance

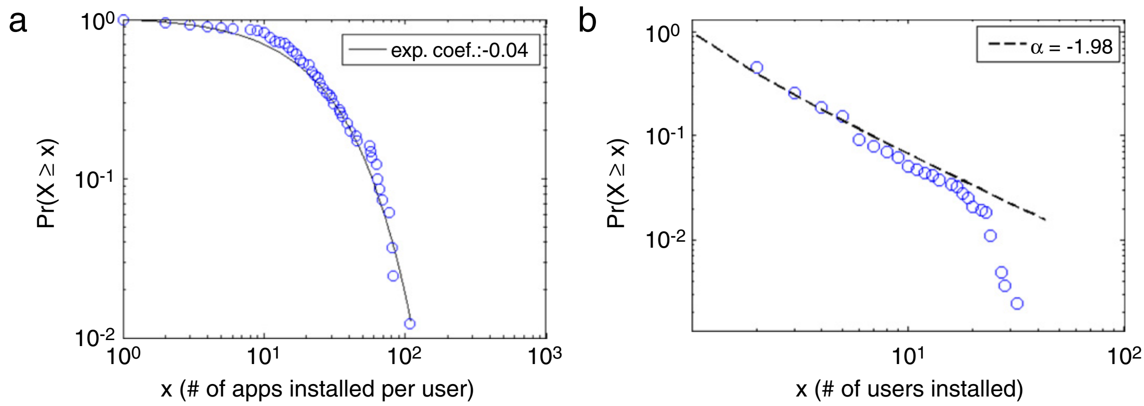

Fig. 7. App statistics. (a) Distribution of the number of users per app. (b) Distribution of the number of apps per user.

as well ($p < 1e - 18$). Both tests strongly reject the null hypothesis that the numbers of common apps are under the same distribution for both groups.

8.2.2. Self-report closeness network

The second network investigated is the self-report network. In the beginning of the study, each pilot phase participant was asked in a survey to label other participants on a closeness scale of 0–10. We then created an adjacency matrix based on all self reports, and calculated the common apps shared by every pair of participants. For each pair, the closeness measure in this result is defined as the average rating of both participants on one another. We again divided pairs into two groups: Group 1 includes pairs with closeness measure less than 1, and Group 2 is in the range (1, 10]. Therefore, Group 1 consists of strangers together with distant acquaintances, and closer relationships will all be included in Group 2. The mean number of shared apps for Group 1 and Group 2 was 4.76 and 4.05 respectively. ANOVA shows borderline significant difference in the numbers of common apps from both groups ($F = 4.97$, $p < 0.026$), but less strong than the BT proximity network. K–S also shows similar results ($p = 0.0045$). However, the mean number of common apps, suggests that the two groups share almost same number of apps, with the strangers group sharing even more common apps. We then tested the border threshold with other values between [0, 2], and notice little difference in the means and the two statistical tests.

8.3. Discussion

In conclusion, we discovered that people who spend more time in face-to-face interaction are more likely to share common apps. In fact, in our dataset, pairs with face-to-face interaction share on average two more common apps on their phones compared with pairs with little face-to-face interactions. Those face-to-face interactions might include group activities, religion-related interactions, time spent with significant others and many other possibilities. However, we also observed that the self-reported friendships do not result in an increase in the number of common shared apps. We believe our results provide strong evidence on app diffusion patterns: apps do spread via social interaction. In particular, the diffusion of apps relies more on the face-to-face interaction ties than the self-perceived friendship ties. Based on this initial analysis, we follow up with two threads of further investigation: The first is focusing on the specific mechanisms for understanding app installation choices, including additional network modalities, and demonstrating how the to combine of multiple network modalities to create an algorithm that predicts future app installation. This work is described in [44]. The second investigation direction relates to the generalization of app decisions to general decision making and the social mechanisms that affect it. Moreover, we are interested in learning how to use these insights to design better mechanisms for supporting decision making, in particular in the context of health and wellness. Following the passive observations of the study's pilot phase, we continued to design an active intervention that aims to help us better understand mechanism of social support and behavior choices.

9. Fitness-centered social intervention design

Between October–December 2010, an active intervention was carried out with the Friends and Family study pool. Its main goal was to explore the question of understanding social influence and motivation in the context of health and wellness activities. The intervention was presented to participants as a wellness game to help them increase their daily activity levels. 108 out of 123 active subjects at the time elected to participate. Subjects were divided into three experimental conditions: Control, Peer-View, and Peer-Reward (described below), allowing us to isolate different incentive mechanisms related to monetary reward, the value of social information, and social pressure/influence. Following an initial period where baseline activity levels were collected, all intervention subjects were given feedback on their performance in the form of a monetary reward, R, which was calculated as a function of their activity. Reward of up to \$5 was allocated every three days. Participants

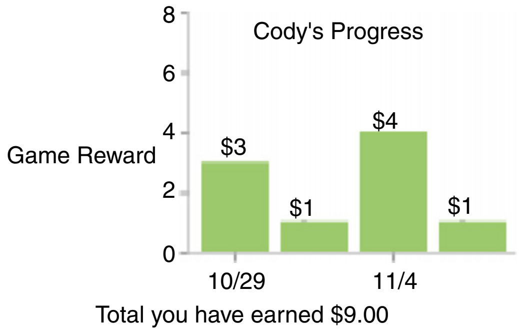

Fig. 8. Reward display for participants in the control condition.

were presented current, past, and total reward accumulated, as shown in Fig. 8, and could access their reward page via browser or the phone. Each group received a variant adapted to its condition. The game was not designed as a competition, and every subject had the potential to earn the maximal reward.

9.1. Experimental conditions

9.1.1. Control condition

All conditions have a baseline of self-monitoring. In the control condition, subjects saw only their own progress as visualized in Fig. 8. Also, reward given to the control subjects depended only on their own activity.

9.1.2. Experimental condition 1: "Peer-View"

In the first experimental condition, "Peer-View", subjects were shown their own progress and the progress of two "Buddies" also in the same experimental group. In turn, the subject's progress was visible to two other peers in the same experimental group. Each subject's reward still depended on his own activity.

9.1.3. Experimental condition 2: "Peer-Reward"

We propose a novel condition aimed at generating increased incentives for social influence and possibly the leveraging of social capital. This scheme is was chosen to closely match the theoretical framework developed in [45]. In this "Peer-Reward" mechanism, subjects were shown their own progress as well as that of two Buddies, but this time subjects' rewards depended solely on the performance of their Buddies. At the same time, their own performance reward was split between two other peers, to which the current user was a Buddy. If subject A had Buddies B and C, the maximal reward A could receive for this period is still $5 per three-day period: $2.5 from B and $2.5 from C. The Peer-Reward feedback page displays how much reward they received from their Buddies, as well how much reward they are earning for the people they are Buddy to.

9.2. Accelerometer-based activity measurement

Our investigation of social mechanisms does not require accurate activity classification and step measurement. We decided to implement a less accurate but more robust algorithm for estimating activity levels, which allowed for increased flexibility in the way participants could carry the phone, and reduced the obtrusiveness of the study. Accelerometer scans were sampled in a duty cycle of 15 s every 2 min. During the 15 s, raw 3-axis accelerometer measurements are sampled at 5 Hz rate and combined to compute the vector magnitude for each sample. The variance of the magnitude in each one-second block is then computed [21]. The score was calculated by giving one point for every second, thresholded to three states (1) "still" (2) "moderate activity" (3) "high activity". However, in this paper we are interested in general change in activity levels and therefore combine the two active levels. Participants were not constrained in the way they should carry the phone. [30] found this did not interfere with activity measurement and classification, and our tests suggest this as well. Participants were told that the more they carry the phone on their person, the more of their activity would be accounted for their game score.

9.3. Game reward calculation

Game reward was calculated every three days, using a reference window of the seven days preceding the current 3-day bin. Average and variance for daily activity count are calculated for the reference window, as well the daily average for the current 3-day bin. Reward depended solely on an individual's performance, and was mapped to be between $0.50–$5, in $0.50 increments between one standard deviation above and below the reference average value. Values above or below the bounds received max or min value, respectively. To avoid discouragement due to zero reward, we did not give less than 50 cents per reward period.

9.4. Discussion of experiment design considerations

One of the great advantages of the Social fMRI and other ubiquitous living-laboratory approaches is the ability to conduct interventions and structured experiments with the study population, as they live their everyday life. In contrast to most fitness-related studies who recruit participants specifically for the fitness study and many times pick participants who actively want to increase their physical activity [24,25,23], we faced similar challenges to those discussed in [46] for working with general populations in the wild. The sub-experiment had to be tailored to the nature of the subjects and the community, and be unobtrusive as well as attractive enough that the study population would want to opt-in.

We had to consider a range of attitudes toward physical activity. The intervention was thus designed as an non-competitive game, where each person is judged based on their own performance and performance change. A previously non-active participant could gain the same reward as a highly active one, while the highly active person would need to work harder. We also had to assume subjects might talk to each other and share information about the game. This is one of the reasons we made sure every participant would have potential to earn the same reward amount. Additional practical considerations included the fact that not everyone had data-plans, and data upload could be delayed. Since we needed it for the reward calculation, we added feedback to users about their data upload state, and also designed the accelerometer and reward three-day bins in a way that would allow for backlogged data to reach the server in time.

By creating a network structure rather than closed team structure for the social interventions (A receives reward for B and C's performance, while D and E receive reward for A's), we are able to disambiguate and focus on the dyadic and asymmetric relationship of the person doing the activity vs. the person receiving the reward, motivated to convince the first.

10. Preparatory analysis

10.1. Self reported closeness

For the social conditions allocation, we wanted to leverage our knowledge of the subjects' network. We decided to use the network of self-perceived closeness since this network is explicit in participant's minds (as opposed to the Bluetooth collocation network, for example), and this was desirable for the experimental conditions. Each participant rated every other participant on a scale from 0 (not familiar) to 7 (very close). Basic analysis for the intervention participants network shows that it is a fully connected graph except one user. On average, each participant knows 14 other participants. Each participant has, on average, 7 explicit friendship ties (closeness >2) in the study pool. The mean distance between any two participant is 1.9.

10.2. Experimental condition allocation

Based on the closeness and marriage ties information, we designed an allocation algorithm to pair each participant in Peer-See and Peer-Reward with two buddies within their group. We wanted to ensure that at least some participants are paired with existing friends, while keeping the desired network topology and avoiding reciprocal pairings. Due to the sparsity of the friendship network, our division to disjoint experimental groups, and our enforced constraints, we formulated an integer programming optimization problem that attempts to prioritize closer friends as buddies: For participants $p_1, \ldots, p_N$ in either the Peer-See or the Peer-Reward group, we match each participant p_i with two buddies. p_i will monitor the progress, and in the Peer-Reward case, get paid for the progress of her buddies. We set $b_j^i = 1$ if p_j is p_i's buddy, 0 otherwise. Due to symmetry, each participant also naturally has two other participants to monitor her.

The participant self-reported closeness is captured in an adjacency matrix $\mathbf{R}$, where $\mathbf{R}_{i,j}$ captures the closeness between p_i and p_j. Another matrix $\mathbf{M}$ captures the marriage status between two participants, where $\mathbf{M}_{i,j} = 1$ if p_i and p_j are married/partners, 0 if not. Our goal is to allocate closest friends for each participant as buddies, while keeping the designed network topology. This can be written as a binary integer programming formulation:

$$\arg \max_{b_j^i, i,j \in \{1,\ldots,N\}} \sum_{i=1,\ldots,N} \sum_{j=1,\ldots,N} b_j^i \mathbf{R}_{i,j} \tag{1}$$

subject to: $\forall i$, $\sum_{j=1,\ldots,N} b_j^i = 2$,
$\forall i \in \{1, \ldots, N\}$, $b_i^i = 0$,
$\forall i, j \in \{1, \ldots, N\}$, $b_j^i + b_i^j < 2$,
$\sum_{i=1,\ldots,N} \sum_{j=1,\ldots,N} \mathbf{M}_{i,j} b_j^i = 0$.

The first constraint forces each participant to have exactly two buddies. The second constraint is set so that participants cannot be buddy of themselves. The third constrain prohibits two participants from being buddies of each other. The last constraint means that participants cannot have their spouses as buddies. This decision eliminates the unique and complicated effects resulting from marriage ties, and it ensures that our fitness peer monitoring topology is purely constructed on regular friendship ties.

The integer programming problem cannot be solved directly, and we apply an iterative approach: In each iteration, we randomize initial values and use the branch-and-bound algorithm to search for reasonable results, and we select the best solution among all iterations to match individual with their buddies for both social condition groups in our experiment.

Table 2
Number of subjects in each condition.

Condition	Initial	Dropped	Total
Control	18	2	16
Peer-See	45	5	40
Peer-Reward	45	4	41

Table 3
Dates and days of periods used for analysis.

Period	Dates	Days
1	Oct 5–Oct 27	1–23
2	Oct 28–Nov 15	24–42
3	Nov16–Dec 5	43–62

Table 4
Average accelerometer score by time of day. The average score per reading is much lower during the night and highest in the afternoon, as expected.

Time of day	Average accelerometer score per reading
Midnight-6AM	0.23
6AM-Noon	1.29
Noon-6PM	2.34
6PM-Midnight	1.31

11. Post-intervention analysis

11.1. Subject pool

Eleven subjects were removed from the study pool over the course of the intervention (due to prolonged technical issues that prevented reliable activity tracking, long durations of out of town travel, or dropping out of the longitudinal study entirely). Their data has been removed from the analysis, except for cases of analyzing peer effects for their Buddies. For details on the final number of subjects in each study condition, see Table 2.

11.2. Intervention periods for analysis

For analysis of changes in activity level through the intervention, we divided the intervention into three periods: the baseline period before the beginning of the intervention was officially announced, the first 19 days of the intervention, and the second 20 days of the intervention. The periods are summarized in Table 3. For this analysis, the days after the intervention begins are broken up into two periods, and we focus on the latter one to account for any novelty effects and allow us to take a first look at the persistence of any change in behavior. Another timing aspect that should be noted is that when considering the experiment periods in weather and school-year contexts, we can assume that physical activity becomes more challenging as the experiment advances due to the North American winter conditions. In addition, for period 3 we can expect increased end-of-semester workload and stress for the student participants in the subject group.

11.3. Normalized activity values

For analysis purposes, we normalized activity levels to the span of a single sample. For example, a normalized "daily average activity" is calculated by summing all accelerometer samples for the day and then dividing by the total count of accelerometer readings for the day. This gives us the average activity level per reading for that day. This allows us to easily do things like compare between normalized average activity levels in different times of day. It is trivial to convert a normalized value to actual time: For example, a normalized daily average value of 1.0 for an experimental group represents an average activity of 96 min per member.

11.4. Aggregated activity levels

One would reasonably assume that accelerometer readings would not be uniformly distributed throughout the day. A visual inspection of the distribution of non-zero readings indicated that the day should be split into four quarters of six hours each, starting at midnight, in order to explore the difference in average accelerometer score per reading. Table 4 confirms that activity varies greatly throughout the day, and that it correlates with general intuition about the times of high and low activity.

We refer to the a day's worth of accelerometer measurement for one person as a "person-day". For a single person, a complete day's worth of data was 720 accelerometer score readings, since accelerometer scans were taken in

Table 5
Pairwise K–S comparison of activity level of the three experimental conditions pre- and post-intervention. The groups which are being compared are listed in the first column. "Group 1 mean" refers to the group listed first and "Group 2 mean" refers to the group listed second. *D* is the K–S statistic.

Groups tested	Group 1 mean	Group 2 mean	*D*	*p*-value
Pre-intervention (Period 1)				
Cntrl vs. PS&PR	1.162	1.241	0.3261	0.046*
Cntrl vs. PSee	1.162	1.266	0.3478	0.078
Cntrl vs. PRew	1.162	1.216	0.3043	0.164
PSee vs. PRew	1.266	1.216	0.2609	0.316
Post-intervention (Periods 2 and 3)				
Cntrl vs. PS&PR	1.207	1.328	0.3718	0.001***
Cntrl vs. PSee	1.207	1.341	0.4193	0.001***
Cntrl vs. PRew	1.207	1.316	0.3590	0.007**
PSee vs. PRew	1.341	1.316	0.1026	0.976

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.005$.

two-minute intervals. Data was considered "missing" for an interval if there was no accelerometer score logged for that interval. As also assumed in [8], we attribute most missing data to the phone being off, usually during night-time. As the current analysis deals primarily with daily average activity levels and change in daily average activity across time and experimental condition, we precluded person-days that did not have sufficient information for generating a reliable average score for the day. We observed that for days that had more than 50% of the possible readings, the missing datapoints were relatively uniformly spread across the day, while in days with fewer than 50% of possible readings, they were not uniformly distributed and could not reliably be used. When a person's day had fewer than 50% of the possible readings, that day was not used for the analysis and calculation of averages. Removed measurements account for less than 5.4% of the available measurements.

11.5. *Activity levels by condition*

Table 5 presents information about daily average activity levels in a pairwise comparison of the three experimental conditions pre- and post-intervention, using the K–S test, where *D* is the K–S statistic. For this analysis, the two post-intervention periods are combined into one. Ideally, in the pre-intervention period we expect the null hypothesis to be true. While in the comparison that compares the control group vs. both social conditions the result is statistically significant ($p < 0.05$), in the direct pairwise comparison the test does not exhibit statistical significance, as expected. Conversely, according to the experimental mechanism design hypothesis, we anticipate that the social conditions will do better than the control, and possibly exhibit difference properties when compared to each other. For all comparisons between the social conditions (independently and jointly) and the control group, K–S test shows statistical significance ($p < 0.01$ and $p < 0.005$). However the difference between the two social conditions comes out non-significant under this comparison, possibly due to the inclusion of novelty effects through combining both post-intervention periods. The difference between the two social conditions is not significant, possibly because higher activity levels are increasingly difficult to obtain (e.g., a 'headroom' effect). In any case, in this experiment we do not care about absolute activity levels as much as the *monetary efficiency* of the incentive mechanisms.

11.6. *Reward efficiency*

We are interested in the change in activity levels for each group rather than simple comparison of activity means. Furthermore, we want to evaluate the effectiveness of the exogenous money or energy injected into a system. We define "reward efficiency", η, which represents the activity change per dollar invested in the system. Reward efficiency for condition i is defined as:

$$\eta_i = \frac{\overline{a_{i,3}} - \overline{a_{i,1}}}{\overline{R_{i,3}}}$$

where $\overline{a_{i,k}}$ is the mean activity level for all participants in group i in period k, and $\overline{R_{i,k}}$ is the average reward per participant in group i in period k. Period 3 is used as the reference frame since we want to look at longer-term adherence. Tables 6 and 7 present information on reward efficiency for this dataset, based on actual monetary reward paid. Table 7 shows results of pairwise K–S testing of reward efficiency values (*D* is the K–S statistic), where all but one demonstrate statistical significance. In Table 6 we see that reward efficiency is more than doubled between the control condition and the Peer-See condition, and the efficiency of the Peer-Reward group is even more than the latter when comparing the conditions as a whole. In relative terms, we observe an average activity increase for Control, Peer-See, and Peer-Reward of 3.2%, 5.5%, and 10.4% respectively, counting in data from all times of day, days of week, sick-times, holidays, and so on. For the Peer-Reward condition, this comes down to an average increase of 84 min of physical activity per week, per participant.

Table 6
Reward efficiencies (η). Reward efficiency is defined as the amount of activity level increase per dollar of reward paid.

Condition	Activity change from Period 1 to Period 3	Reward in Period 3	Reward efficiency (Δ/\$)
Overall			
Control	0.037	\$3.00	0.012
Exp 1	0.070	\$2.77	0.0253
Exp 2	0.126	\$3.04	0.0416
Close Buddies (both Buddies score 3 or higher)			
Exp 1	0.118	\$2.68	0.0444
Exp 2	0.269	\$3.00	0.0896
Stranger Buddies (both Buddies score 2 or lower)			
Exp 1	−0.007	\$2.82	−0.0025
Exp 2	0.137	\$2.95	0.0464
Mixed Buddies (one close, one stranger)			
Exp 1	0.154	\$2.75	0.0560
Exp 2	0.053	\$3.12	0.0171

Table 7
Pairwise K–S comparison of the differences between reward efficiencies. *D* is the K–S statistic. All differences are statistically significant, except the difference between the two experimental groups when taken in their entirety.

Groups being compared	Group 1 reward efficiency	Group 2 reward efficiency	*D*	*p*-value
Overall				
Cntrl vs. PSee	0.0120	0.0253	1.000	0.001**
Cntrl vs. PRew	0.0120	0.0416	1.000	0.001**
PSee vs. PRew	0.0253	0.0416	0.429	0.432
Close Buddies (both Buddies score 3 or higher)				
PSee vs. PRew	0.0444	0.0896	1.000	0.002**
Stranger Buddies (both Buddies score 2 or lower)				
PSee vs. PRew	−0.0025	0.0464	1.000	0.001**
Mixed Buddies (one Close, one Stranger)				
PSee vs. PRew	0.0560	0.0171	1.000	0.001**

** $p < 0.01$.

As the underlying differences between the two social conditions were not clearly apparent, in Table 7 we dive into the social component. We divide the subjects according to their pre-reported closeness level with their Buddies. Although the overall comparison of the social conditions does not present statistical significance, the further grouping according to pre-existing relationships shows that the Peer-Reward condition achieves better results in two out of the three cases (close buddies and stranger buddies), while the Peer-See condition achieves better results for mixed buddies. For all these cases, we get statistical significance ($p < 0.01$). We see a complicated interaction element with regards to the Buddy closeness, which we touch on in the next section.

11.7. Discussion

In this analysis we begin investigating the effectiveness of the different motivation and influence mechanisms for encouraging increased physical activity in-situ. We focus on two key metrics: The first is differences in average activity levels, both across conditions and chronological periods of the experiment, and the second is the efficiency of the reward "investment" in the system.

When daily average activity levels are analyzed, they support the hypothesis that the social components of both experimental conditions, together and separately, lead to a statistically significant positive difference. Analysis of the difference of effect between the two socially involved experimental groups is more complex, and dividing the experimental groups based on pre-intervention closeness of the Buddy triads reveals different trends. When reward efficiency is analyzed, we again see a significant difference between the control group on the one hand and the two experimental groups, taken together, on the other.

Results confirm our notions that embedding the social aspects in this non-competitive game adds to physical activity performance and activity adherence over time, compared to the socially isolated control condition. An interesting question arises with respect to the social mechanisms. In the Peer-See group, there is social information that traverses the links between peered Buddies, but participants still receives a "selfish" reward. In Peer-Reward, both information and reward traverse the links between peers, and a potential for social influence as motivator. The intensity of pre-existing social relationships seems to play a factor, and results seem to support a complex contagion like phenomena, as described by Centola and Macy [2], especially when observing the interplay in triads where there is a "mix" of close and stranger peers. We have yet to investigate the communication patterns between the peers, and their subjective view of their condition, to try and understand if and how the social influence or pressure was exerted. We hope that by analyzing additional signals already collected, like the communication logs and co-location information, as well as related surveys administered to the participants, we will be able to shed more light on these underlying processes.

Had this intervention been conducted in springtime, one might expect a natural rise in physical activity as weather improves, which might have made it hard to separate the intervention's contribution. By going against the natural trend during winter, we challenge our experimental mechanisms. While results are not fully conclusive, they may suggest that while performance in the control and even Peer-See conditions deteriorates as time passes, the performance in Peer-Reward is slower to start but steadier in increase over time. The observations might support a hypothesis that the Peer-Reward condition induces social capital that takes time to build up, but once in place provides a more sustainable incentive structure than the direct monetary reward, or alternatively, a way to augment the exogenous monetary compensation with indigenous social capital, leading to a higher efficiency, and higher "return" on every Dollar invested in the system.

It is also important to mention that by design choice, we did not perform any external communication "scaffolding" to encourage social interaction. There were no mechanism within the study software for sharing results and promoting discussion—any such actions were done by participants on their own accord using their existing means of interaction. Related studies with social components [28,25] suggest that adding explicit communication mechanisms to the technical system might add to the social effects of the intervention.

12. Conclusion

In this paper we introduced the Friends and Family Study, which combines high-dimensionality and high-throughput social and behavioral sensing using ubiquitous mobile phones, together with experimental interventions. We described our Android phone centered system that has been deployed in the study for over a year now. We presented initial results of two analyses and a specific experimental intervention that demonstrates the great potential of the study dataset, its underlying technical system, and the of the general Social fMRI approach for measuring and experimenting with social mechanisms.

Through the individual behavior analysis, we showed that individuals' social interaction diversity correlates with their current income level, suggesting a contradictory social theory to the currently prevailing theory.

Through the network effect analysis, we demonstrated a relationship between the number of mobile applications that two people share in common to the time they physically spend face-to-face. Our observations suggest that the diffusion of apps relies more on the face-to-face interaction ties than on self-perceived friendship ties.

Through the fitness intervention example, we demonstrated challenges and benefits of leveraging our prior observations for the experiment design. We presented three key findings through this intervention: First, results support there is a statistically significant effect of social components on the real-world in-situ physical activity levels. Second, results show that our novel Peer-Reward social influence mechanism leveraging social capital can increase the efficiency of exogenous money and resources invested in the system. This could contribute to the design of future policies and intervention. Finally, we see a complex interaction effect related to pre-existing social ties inside the social experimental conditions. This could support hypothesis of a complex contagion like effect that should be further investigated. Immediate future work includes expanding the analysis of the existing data, as well as the design of new experiments based on these initial observations, particularly in the area of quantifying social capital and favor exchange. We hope that isolating and evaluating health related social mechanisms will become part of the toolbox for encouraging healthy behavior, combined with other components such as user interfaces, accurate measurement techniques, and individual goal setting.

In the same way that fMRI techniques help map the interconnections and mechanics of the human brain, we hope that our work will help advocate an evolution from mostly passive observatories to data-rich Social fMRI type of studies that can help further our understanding of the interconnections and mechanics of human society.

Acknowledgments

We would like to thank Sai Moturu, Cody Sumter, and Maya Orbach for their contribution toward the experimental deployment, surveys, and visualizations, and the undergraduate contributors for their help with the development of the sensing platform. We thank the reviewers for their thoughtful comments. Most importantly, we thank the Friends and Family Study participants. This research was partially sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053, and by AFOSR under Award Number FA9550-10-1-0122, by the MASDAR Institute of Science and Technology, and by generous gifts from P&G, Best Buy, Motorola, Google, and Samsung. Views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the sponsors.

References

[1] N. Eagle, M. Macy, R. Claxton, Network diversity and economic development, Science 328 (5981) (2010). http://www.sciencemag.org/content/328/5981/1029.full.pdf, doi:10.1126/science.1186605. URL: http://www.sciencemag.org/content/328/5981/1029.abstract.

[2] D. Centola, M. Macy, Complex contagions and the weakness of long ties, The American Journal of Sociology 113 (3) (2007) 702–734. URL: http://www.jstor.org/stable/10.1086/521848.
[3] D. Lazer, A. Pentland, et al., Computational social science, Science 323 (5915) (2009) 721–723. http://www.sciencemag.org/content/323/5915/721.full.pdf, doi:10.1126/science.1167742, URL: http://www.sciencemag.org/content/323/5915/721.short.
[4] T. Dawber, The Framingham Study: The Epidemiology of Atherosclerotic Disease, Harvard U. Press, 1980.
[5] M.C. Gonzalez, C.A. Hidalgo, A.-L. Barabasi, Understanding individual human mobility patterns, Nature 453 (7196) (2008) 779–782. doi:10.1038/nature06958.
[6] C.A. Hidalgo, C. Rodriguez-Sickert, The dynamics of a mobile phone network, Physica A: Statistical Mechanics and its Applications 387 (12) (2008), doi:10.1016/j.physa.2008.01.073. URL: http://www.sciencedirect.com/science/article/B6TVG-4RM7MXN-5/2/3f55af67920a2a2763b8944d88c760d0.
[7] J.-P. Onnela, F. Reed-Tsochas, Spontaneous emergence of social influence in online systems, Proceedings of the National Academy of Sciences 107 (43) (2007). http://www.pnas.org/content/107/43/18375.full.pdf+html, doi:10.1073/pnas.0914572107. URL: http://www.pnas.org/content/107/43/18375.abstract.
[8] N. Eagle, A. Pentland, Reality mining: sensing complex social systems, Personal and Ubiquitous Computing (10) (2006) 255–268.
[9] N. Eagle, A.S. Pentland, D. Lazer, Inferring friendship network structure by using mobile phone data, Proceedings of the National Academy of Sciences 106 (36) (2009). http://www.pnas.org/content/106/36/15274.full.pdf+html, doi:10.1073/pnas.0900282106. URL: http://www.pnas.org/content/106/36/15274.abstract.
[10] A. Madan, M. Cebrian, D. Lazer, A. Pentland, Social sensing for epidemiological behavior change, in: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Ubicomp'10, ACM, New York, NY, USA, 2010, pp. 291–300. doi:10.1145/1864349.1864394.
[11] A. Madan, K. Farrahi, D.G. Perez, A. Pentland, Pervasive sensing to model political opinions in face-to-face networks, in: Pervasive Computing, 2011, pp. 214–231.
[12] R. Montoliu, D. Gatica-Perez, Discovering human places of interest from multimodal mobile phone data, in: Proc of 9th Int. Conference on on Mobile and Ubiquitous Multimedia, MUM,',',', 2010.
[13] H. Lu, et al. The Jigsaw continuous sensing engine for mobile phone applications, in: Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, SenSys'10, New York, NY, USA, 2010. doi:10.1145/1869983.1869992.
[14] D.O. Olguín, B.N. Waber, T. Kim, A. Mohan, K. Ara, A. Pentland, Sensible organizations: technology and methodology for automatically measuring organizational behavior, IEEE Transactions on Systems, Man, and Cybernetics, Part B 39 (1) (2009) 43–55.
[15] The human speechome project. URL: http://www.media.mit.edu/cogmac/projects/hsp.html.
[16] J. Gemmell, G. Bell, R. Lueder, S. Drucker, C. Wong, Mylifebits: fulfilling the memex vision, MULTIMEDIA'02, New York, NY, USA, 2002, doi:10.1145/641007.641053.
[17] R. Barker, Midwest psychological field station, in: Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior, Stanford University Press, 1968, URL: http://books.google.com/books?id=p1arAAAAIAAJ.
[18] A. Bauman, Updating the evidence that physical activity is good for health: an epidemiological review 2000-2003, Journal of Science and Medicine in Sport 7 (Suppl. 1–1) (2004) 6–19. doi:10.1016/S1440-2440(04)80273-1.
[19] S.N. Blair, Y. Cheng, S.J. Holder, Is physical activity or physical fitness more important in defining health benefits? Medicine & Science in Sports & Exercise 33 (6) (2001). URL: http://journals.lww.com/acsm-msse/Fulltext/2001/06001/Is_physical_activity_or_physical_fitness_more.7.aspx.
[20] C.V. Bouten, K.T. Koekkoek, M. Verduin, R. Kodde, J.D. Janssen, A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity, IEEE Transactions on Biomedical Engineering (1997), doi:10.1109/10.554760.
[21] R.G. Eston, A.V. Rowlands, D.K. Ingledew, Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities, Journal of Applied Physiology 84 (1) (1998) 362–371. http://jap.physiology.org/content/84/1/362.full.pdf+html, URL: http://jap.physiology.org/content/84/1/362.abstract.
[22] R.P. Troiano, et al., Physical activity in the united states measured by accelerometer, Medicine & Science in Sports & Exercise 40 (1) (2008) 181–188. doi:10.1249/mss.0b013e31815a51b3.
[23] S. Consolvo, et al., Activity sensing in the wild: a field trial of ubifit garden, in: Proceeding of CHI'08, ACM, New York, NY, USA, 2008, pp. 1797–1806. doi:10.1145/1357054.1357335.
[24] J.J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, H.B. Strub, Fish'n'steps: encouraging physical activity with an interactive computer game, in: Proc. of the International Conf. on Ubiquitous Computing, 2006.
[25] S. Consolvo, K. Everitt, I. Smith, J.A. Landay, Design requirements for technologies that encourage physical activity, in: Proceedings of CHI'06, ACM, New York, NY, USA, 2006, doi:10.1145/1124772.1124840.
[26] I. Anderson, et al., Shakra: tracking and sharing daily activity levels with unaugmented mobile phones, Mobile Networks and Applications 12 (2007) 185–199. doi:10.1007/s11036-007-0011-7.
[27] T. Toscos, et al. Encouraging physical activity in teens can technology help reduce barriers to physical activity in adolescent girls? in: Pervasive Computing Technologies for Healthcare, 2008, doi:10.1109/PCTHEALTH.2008.4571073.
[28] D. Foster, C. Linehan, S. Lawson, Motivating physical activity at work: using persuasive social media extensions for simple mobile devices, Design (2010). URL: http://eprints.lincoln.ac.uk/3153/.
[29] L. Berkman, T. Glass, Social integration, social support, and health, in: L. Berkman, I.E. Kawachi (Eds.), Social Epidemiology, 2000.
[30] S.T. Saponas, et al. ilearn on the iphone: real-time human activity classification on commodity mobile phones, Tech. Report, University of Washington, 2008.
[31] Cardiotrainer. http://www.worksmartlabs.com/cardiotrainer/about.php.
[32] R.I.M. Dunbar, Co-evolution of neocortex size, group size and language in humans, Behavioral and Brain Sciences 16 (4) (1993) 681–735. URL: http://www.bbsonline.org/Preprints/OldArchive/bbs.dunbar.html.
[33] O.P. John, S. Srivastava, The big-five trait taxonomy: history, measurement, and theoretical perspectives, Guilford, vol. 2, 1999, pp. 102–138. URL: http://books.google.com/books?hl=en&lr=&id=b0yalwi1HDMC&oi=fnd&pg=PA102&dq=The+Big+Five+Trait+Taxonomy&ots=748wRbZrNn&sig=NFkhOOsTchZWYPu4QztigxvSY2o.
[34] Funf: open sensing framework. URL: http://funf.media.mit.edu.
[35] J. Bruggeman, Network diversity and economic development: a comment, Arxiv Preprint. arXiv:1011.0208.
[36] S. Page, The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies, Princeton Univ. Pr., 2008.
[37] M. Granovetter, The strength of weak ties, The American Journal of Sociology 78 (6) (1973) 1360.
[38] R. Burt, Structural Holes: The Social Structure of Competition, Harvard Univ. Pr., 1995.
[39] P. Frijters, J. Haisken-DeNew, M. Shields, Money does matter! evidence from increasing real income and life satisfaction in east germany following reunification, The American Economic Review 94 (3) (2004) 730–740.
[40] T. Paridon, S. Carraher, S. Carraher, The income effect in personal shopping value, consumer selfconfidence, and information sharing (word of mouth communication) research, Academy of Marketing Studies Journal 10 (2) (2006) 107–124.
[41] T. Clydesdale, Family behaviors among early us baby boomers: exploring the effects of religion and income change, 1965-1982, Sociological Forum 76 (1997) 605.
[42] S. Pong, D. Ju, The effects of change in family structure and income on dropping out of middle and high school, Journal of Family Issues 21 (2) (2000) 147.
[43] Androlib statistics page. URL: http://www.androlib.com.
[44] W. Pan, N. Aharony, A. Pentland, Composite social network for predicting mobile apps installation, in: Proceedings of the 25th Conference on Artificial Intelligence, AAAI-11, San Francisco, CA, 2011.
[45] A. Mani, I. Rahwan, A. Pentland, Localizing externalities in social networks: Inducing peer pressure to enforce socially efficient outcomes, in: Proceedings of the Workshop on Information in Networks (WIN), New York, 2010.
[46] J. Maitland, K.A. Siek, Technological approaches to promoting physical activity, in: Proceedings of the 21st Annual Conference of the Australian Computer–Human Interaction, ACM, New York, 2009, doi:10.1145/1738826.1738873.