Social fMRI: Measuring and Designing Social Mechanisms using Mobile Phones

by

Nadav Aharony

S.M., Massachusetts Institute of Technology (2008)
B.Sc., Technion - Israel Institute of Technology (2004)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2012

Author ______________________________
Program in Media Arts and Sciences
December 16, 2011

Certified by ______________________________
Alex P. Pentland
Toshiba Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by ______________________________
Mitchel Resnick
LEGO Papert Professor of Learning and Research
Academic Head
Program in Media Arts and Sciences

Social fMRI: Measuring and Designing Social Mechanisms using Mobile Phones

by

Nadav Aharony

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on December 16, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences

Abstract

A key challenge of data-driven social science is the gathering of high quality multi-dimensional datasets. A second challenge relates to the design and execution of social experiments in the real world that are as reliable as those within a controlled laboratory, yet yield more practical results. We introduce the Social Functional Mechanism-design and Relationship Imaging, or "*Social fMRI*" - an approach that enhances existing computational social science methodologies by bridging rich data collection strategies with experimental interventions.

In this thesis, we demonstrate the value of the Social fMRI approach in our *Friends and Family* study. We transformed a young-family residential community into a living laboratory for 15 months, through a very fine-grained and longitudinal data collection process combined with targeted experimental interventions. Through the derived dataset of unprecedented quality, the Social fMRI approach allows us to gain insights into intricate social mechanisms and interpersonal relationships within the community in ways not previously possible.

This thesis delivers the following contributions: (1) A *methodology* combining a rich-data experimental approach together with carefully designed interventions, (2) a *system* supporting the methodology - implemented, field-tested, and released to the world as an open-source framework with a growing community of users, (3) a *dataset* collected using the system, comprising what is, to date, the richest real-world dataset of its genre, (4) a very large set of *experimental findings* that contribute to our understanding of important research questions in computational social science in addition to demonstrating the methodology's potential. Among the results described in this thesis are the design and evaluation of a novel mechanism for social support in a health-related context, the observation that the diffusion of mobile applications relies more on the face-to-face interaction ties than on self-perceived friendship ties, and a gained understanding of the evolution of modeling and prediction processes over time and varying sample sizes.

Thesis Supervisor: Alex P. Pentland
Title: Toshiba Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Social fMRI: Measuring and Designing Social Mechanisms using Mobile Phones

by

Nadav Aharony

The following people served as readers for this thesis:

Thesis Reader________________________________

Moshe E. Ben-Akiva
Professor of Civil and Environmental Engineering
Department of Civil Environmental Engineering, MIT

Thesis Reader________________________________

Stuart Madnick
Professor of Information Technology
MIT Sloan School of Management

Contents

List of Figures

List of Tables

Acknowledgements

The research described in this thesis is a result of many collaborations and there are many people that I would like to thank. First and foremost is my PhD advisor and thesis supervisor, Alex (Sandy) Pentland. I would also like to thank my thesis readers, Stuart Madnick and Moshe Ben-Akiva for their helpful comments and insights.

Many collaborators were involved in the different components of this work, some of them took part in more than one component and investigation direction. In conducting the Friends and Family Living Laboratory study: Cory Ip, Wei Pan, Cody Sumter, Sai Moturu and Inas Khayal. In the data processing and analysis: Yaniv Altshuler, Scott Greenwald, Cody Sumter, Michael Fier, Kai Kunze, Bruno Lepri, Jacopo Staiano, Yves-Alexandre de Montjoye and Wei Pan. In the development of the Funf project: Alan Gardner, Cody Sumter, and Wei Pan. I would also like to thank MIT Housing for their support in conducting the Friends and Family study, in particular Jason McNight, Michael Collins, and Dennis Collins. This work would not be possible without support from Media Lab sponsors. In particular, I would like to thank Chris Wren, Daniel Sandler, Mike Lockwood and Aaron Brown from Google for their support and technical advice, as well as Motorola and Bank of America. In addition, I thank the undergraduate students who contributed to this research work: An Li, Yanping Chen, Jay Baxter, Christopher Graves, Connie Chan, Stephen Zhang, Itai Turbahn, Morgan Krey, Tammy Ngai, Harley Zheng. Finally, I want to give special thanks to the experiment participants of the Friends and Family study, without whom none of this work would be possible.

I would also like to thank the past and present members of the Human Dynamics group, who have been my peers and friends over the last several years: Nicole Freedman, Ben Waber, Yaniv Altshuler, Manuel Cebrian, Yves-Alexandre de Montjoye, Coco Krumm, Cody Sumter, Iyad Rahwan, Inas Khayal, Daniel Olguin Olguin, Taemie Kim, Anmol Madan, Lanthe Chronis, Bruno Lepri, and Ankur Mani.

There are many other people at the Media Lab - students, faculty, and staff, who have helped formulate my thoughts and ideas, and made my time here unforgettable. I want to thank Yadid Ayzenberg, Scott Greenwald, Eyal Shachar, Amit Zoran, Roy Shilkrot, Oren Zuckerman, Jamie Zigelbaum, Marcelo Coelho, Sajid Sadi, Andy Lippman, David Reed, Chris Csikszentmihlyi, Whitman Richards, Mitch Resnick, Hiroshi Ishii, Pattie Maes, Linda Peterson, Paula Aguilera, Jon Ferguson, Joost Bonsen, Peter Pflantz, and Will Glesnes.

My final acknowledgement goes to my family for their ongoing support: my parents Gideon and Lea Aharony, and most of all, my wife, Maya. This thesis is dedicated to you.

1

Introduction

Picture an imaging chamber, one so big that it could be placed around an entire community. What if we could, with permission, record and display nearly every facet of behavior, communication, and social interaction among its members as they live their everyday lives? This potential would afford rich insights into humanity - how societies operate, how real-world relationships form and change over time, and how behavior and choices spread from one person to another. We could diagnose the health of a community, and of its individuals. We could also conduct proactive investigations, applying carefully designed experimental interventions within the community. This would let us isolate experimental variables and test effects of policies and social mechanisms before, during, and after their application. We could even measure the effects of feeding this information back to the members of the community themselves.

We have built the beginnings of what we call "The Social fMRI". To build it, one doesn't need to build a huge chamber, a-la Fuller's city dome (Figure 1-1), just a bunch of modern smartphones. Using our mobile sensing software, we transformed a residential community into a living laboratory for over 15 months. Many signals were collected from each participant, altogether comprising what is, to date, the richest real-world dataset of its kind.

Figure 1-1: "Geodesic Dome City", Richard Buckminster Fuller (1968)

In the following sections we develop the concept and the motivation for the methodology being developed, as well as review the goals, key contributions, and general outline of this thesis.

1.1 Approach: 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. This amounts to creating a social and behavioral sensing platform . *Social* **f***MRI* 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.

The general framework of the Social fMRI idea is a combination of a longitudinal living-laboratory or "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 [Eagle and Pentland, 2006], adding much more richness while incorporating a strong element of active interaction and carefully directed stimulation of the study population.

We wanted to create an experimental framework that would enable high-dimensional and multi-modal data collection from many users over extended periods of time. The core of the approach is that the data is being collected at the level of the individual. The idea behind this user-centric approach is that it enables access to all aspects and dimensions of the person's life, collected via the end-users, with their informed consent. The core of the data collection system is a mobile phone, which acts as a wearable ubiquitous sensor for its owner's behavior and social patterns. The transition from simpler "feature phones" to computationally advanced and sensor-packed smartphones, is turning them into fully fledged behavioral sensors that allow inference of activity and context, and enables us to start investigating connections between individual and social behavior. Much of this information is not available at the service provider level. The phone data collection is augmented by several other collection components that augment the dataset with additional dimensions of information. The phone as well as PC based tools allow communication back with the user population as well as the design and deployment of interventions and sub-experiments.

1.2 Motivation

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. describe the potential of computational social science to increase our knowledge of individuals, groups, and societies, with an unprecedented breadth, depth, and scale [Lazer, 2009]. Computational social science combines the leading techniques from network science [Barabasi and Albert, 1999, Watts and Strogatz, 1998, Newman, 2003] with new machine learning and pattern recognition tools specialized for the understanding of people's behavior and social interactions [Eagle and Pentland, 2006, Eagle et al., 2009].

The pervasiveness of mobile phones has made them popular scientific data collection tools, as social and behavioral sensors of location, proximity, communications, context, and even things like mood, personal traits, and state of health. Eagle and Pentland [Eagle and Pentland, 2006] coined the term "Reality Mining" to describe collection of sensor data pertaining to human social behavior. While existing works have demonstrated results for modeling and inference of social network structure and personal information out of mobile phone data, most are still proofs of concept in a nascent field. The work of the "*Data Scientist*" is still that of an artisan, using personal experience, insight, and sometimes even gut feeling in order to extract meaning out of the abundance of data and noise.

As the field of computational social science matures, there is need for more structured methodologies. Ones that would assist the researcher or practitioner in designing data collection campaigns, understanding the potential of collected datasets, and estimating the accuracy limits of present analysis strategy vs. alternative ones. While there will always be a need for the expert and experienced data artisan to push the envelope of what we can learn, with the exponential increase in accumulated data and the rise of a big-data ecosystem, there is an imperative need to create a more accurate science and engineering of data *collection*, *processing*, and *analysis*. The work presented in the thesis is a building block in this larger effort.

1.2.1 Improving Our Understanding of People and Society

We can group some of the main areas under active investigation of social networks and complex systems in the *real world* under two categories. The first includes the study of the structure of networks, their properties, and how they come to be, and the other is related to the flow of information and influence through such networks:

- ***Properties of Networks***

  There has been a lot of work pertaining to understanding properties of networks, and human networks in particular. These include understanding intricate community structure and internal organization of social networks. Existing works on understanding and modeling structural properties of social networks include, for example [Newman and Park, 2003, Richards and Macindoe, 2010b, Richards and Macindoe, 2010a, Barabasi and Albert, 1999, Watts and Strogatz, 1998, Clauset et al., 2008, Ahn et al., 2010], research looking at how networks and groups might evolve over time [Palla et al., 2007]. Much of this work involves modeling and analysis based on pre-existing and available network structures and simulation, not targeted experimental data collection.

  In real life, an edge between two individuals is a superset of many factors (some might even say many types of social-networks). Often, the network projection that is exposed to researchers compounds these factors in a way that is hard to discern.For example, a phone network, like those studied by many researchers these days, compounds a great number of social network types (business relationships, friendships of varying levels, social groups and affiliations, etc.) and a lot of noise (e.g. people calling automated answering services or ordering food delivery). This agglomeration might clutter up the analysis and possible conclusions. Phone-based mobile sensing experiments like Reality Mining [Eagle and Pentland, 2006] and Social Evolution [Madan, 2010] have shown the potential of constructing reliable social networks using mobile phone sensing, combined with self-reported individual and community properties. They have demonstrated the fact that connections between people span multiple network modalities (e.g. one spanned by physical proximity vs. another spanned via a phone-call network), and that people's behaviors and interaction differ by strength and type of ties, however they have not attempted an extensive mapping of the "ground truth" of people's affiliations, grouping, and sub-networks in life. A deliberate collection of information from as many network modalities as possible, combined with collection of as many of people's formal and informal affiliations and groupings, would give researchers a strong boost in their work towards building the complex models that describe real world social networks.

- ***Network Effects and Diffusion***

  Another body of work in the study of social networks relates to the spread of ideas, disease, fashions, media, decisions, habits, and many kinds of behaviors. These works cross many fields outside the traditional area of epidemiology: social science [Burt, 1987], economics [Avery and Zemsky, 1998], marketing [Aral and Walker, 2010], spread of innovation [Rogers, 2003], online social networks [Onnela and Reed-Tsochas, 2010], or health policy [Christakis and Fowler, 2007], are just a few examples. Interesting models like Centola's complex contagion [Centola and Macy, 2007] begin to offer a more intricate model of influence, as dependent on multiple neighboring nodes and not just one "influencer". As discussed by Onnela et al. in context of Facebook data, domain-limited results might be harder to generalize for the physical world. This motivates the collection of data from within the physical world for investigation of real-world decision making, rather than online representations or proxies for the decision making process. Experiments like the Social Evolution study show the potential in combining mobile phone collected data with submitted information on the subject's state (e.g. political opinion, or physical symptoms) for modeling social diffusion [Madan et al., 2010a]. One of the great challenges in collecting data in the real world with regards to social influence and decision making is the desire to log all of the instances of the desired activity (e.g. deciding to do action X), together with as much as the information on the exposure of an individual to what caused them to make their behavioral choice. For example, a person choosing to purchase a certain product, might have heard about it from a friend, saw it mentioned on TV, or read about it online. While it is impossible to collect *all* instances and sources of exposure, using a rich-data collection approach we might be able to find types of actions that we are able to target more comprehensively than before.

Better understanding of the issues described above will contribute to active research in a multitude of domains, including network science, the study of complex systems, the social sciences, economics, health, and policy making. The Social fMRI approach as well as the dataset generated through the work described in this thesis could add unique contributions in all of these research areas. In addition to the large number of signals collected, the fine-grained collection of data over time allows for observations of the order and evolution of events. This opens the door to making strong claims on cause and effect . In addition, conducting intentionally designed interventions, while doing the rich data collection before, during and after the intervention, could improve real-world testing of policies of and social mechanisms.

1.2.2 Complementing Existing Methodologies

The Social fMRI is a ubiquitous computing approach for conducting computational social science experiments in a way that deals with three key issues with existing methodologies: First is the "offline" nature of many of existing datasets, which are based on offline analysis of previously collected data. This turns them into purely observational studies, and makes it hard to test cause and effect. In most cases, the investigators do not have the option to go back to the study participants and ask them follow-up questions that are based on the analysis. This leads to the second limitation of many existing datasets, which is the fact that they are usually constrained in the types of data that is made available through them. Key academic works in computational social science have used relatively constrained domain data - be it call and SMS records like in [Gonzalez et al., 2008, Eagle et al., 2010, Hidalgo and Rodriguez-Sickert, 2008] , online social networking data as in [Onnela and Reed-Tsochas, 2010], or financial transaction data as in [Krumme et al., 2010]. Finally, as mentioned by Lazer, Pentland et. al [Lazer, 2009], designing scientifically valid experimental interventions with subjects living their everyday routine is very challenging for a wide range of reasons, which we aim to alleviate.

1.3 Thesis Goals

A key goal of this thesis is to present the Social fMRI approach and demonstrate its utility. In particular, in this thesis we aim to demonstrate merits of a *rich-data* experimental paradigm. We define rich-data collection defined terms of number of signals collected (dimensionality), resolution of collection (e.g. raw values vs. aggregate features), sampling rate, and generally the amount of information collected, in an information theoretic sense. Rich-data collected longitudinally over significant periods of time allows us to investigate a broad range of scientific questions from a diversity of domains in a single data-collection campaign. It allows us to discover ties between variables that we would not have though to collect in a traditional experimental design, and it allows us to go back to the data in hindsight and ask questions that we might not have thought about asking originally, in addition to other benefits.

In this thesis we give strong emphasis to doing things in the real world. This includes doing data collection in the real world, or "in the wild"; An approach for conducting *in-situ* experiments vs ex-situ laboratory testing; Insights and quality assessment models for designing real-world data collection initiatives; A system developed not as a proof-of-concept but as a field-deployable system, which is already being reused and repurposed in other experiments.

A final emphasis of the work presented in thesis is on creating platforms that can be reused and extended. This includes the creation of a unique dataset that can serve a platform for researchers and many potential investigations, the development of an experimental methodology that can be reproduced in other settings, and the development of a mobile data collection platform for other researchers, developers, and end-users.

1.4 Contributions

This thesis makes the following key contributions:

1. **The Methodology:** We present the Social fMRI methodology, combining highly dimensional social observatories with carefully designed intervention stimuli for exploring social mechanism design and social dynamics in naturalistic settings. We have developed a system supporting this methodology, and a longitudinal experiment as implementation and testing of the proposed approach. We show the benefits of a rich-data experimental approach, and developed techniques and methods to assist with the design and implementation of data-rich data collection and analysis campaigns.

2. **The System:** We designed, implemented, and released freely to the world a scalable and extensible mobile-phone centric social and behavioral sensing and data collection system. The system was deployed in the field with 140 experimental subjects, for the duration of 15 months. In Oct 2011 it was launched as an open source extensible sensing and data processing

framework for mobile devices. Its core concept is to provide an open source, reusable set of functionalities, enabling the collection, uploading, and configuration of a wide range of data types. It is packaged in a multi-tiered way, to support the needs of end-users, researchers, developers of mobile applications who might want to integrate these features into their apps, and low-level developers that might want to extend and test new algorithms and methodologies for social and behavioral sensing and inference.

3. **The Dataset:** We have assembled a social and behavioral dataset collected under naturalistic settings, in situ, of a family housing community, for over 15 months. The dataset includes over *1,000,000 hours*, or *117 years* of continuous human behavior of 140 co-located individuals over the 15 month period. The data contains over 25 signals logged via each subject's mobile phone, as well as survey questioners financial transaction information for all subjects. It also include Facebook behavior logs for more than 80 individuals for the same time period, and more. To the best of our knowledge, the study in which the data was gathered is the most comprehensive mobile phone experiment performed to date in academia. The dataset generated is probably the largest and richest dataset ever collected on a residential community. This dataset will be shared with the academic community as much as privacy and subject protection considerations would allow.

4. **Specific Experimental Results:** We present several specific findings contributing to solving important research questions in Computational Social Science:

   (a) **Designing Mechanisms for Social Support in Health:** We conducted a fitness intervention and found results that contribute to our understanding of social incentives and motivation in the real-world. Results suggest that:

      - Social factors have an effect on the physical activity behavior, motivation, and adherence over time.
      - Social incentives, and particularly our novel Peer-Reward mechanism encouraging social influence among participants, support higher activity returns per dollar invested in the system.
      - Finally, results support the notion of a complex-contagion-like effect [Centola and Macy, 2007] related to pre-existing social ties between participans.

(b) **Understanding Role of Different Network Types:** We demonstrate several properties of different network modalities connecting the same individuals.

- Our observations suggest that the diffusion of mobile apps relies more on the face-to-face interaction ties than on self-perceived friendship ties.
- We see very different properties for different network types of same individuals.

(c) **Understanding Evolution of Learning Accuracy for Phone Collected Data:** We conducted an investigation of the improvement of modeling and prediction processes over time and sample size.

- We demonstrated characteristics of incremental learning of multiple social and individual properties from raw sensing data collected from mobile phones, as the information is accumulated over time.
- Furthermore, we proposed a method for advance prediction of the maximal learning accuracy possible for the learning task at hand, using just the first few measurements. This information can be useful in many ways, including:
  - Informing real-time resource allocation for data collection, for an ongoing data collection campaign.
  - Estimating limits and time needed for desired accuracy of a given method.
  - Early evaluation of modeling and learning strategies.
- We show that the characteristics of incremental learning along the sample-size axis resembles that of the evolution along the time-axis, and can be fitted as well to the form of a Gompertz function.
- We propose a method for approximating the maximal amount of information (or accuracy) that can be achieved with large sample sets, as well as an approximation to the accuracy that an given group size would result in.

(d) **Modeling:** Using the rich data collected through the study to better understand and predict properties and behaviors of individuals, properties of dyadic ties, and network properties. We also uncover a structure of connections and correlations between different properties and collected signals. These include initial studies relating to sleep, mood, and personality in the context of socialization and network ties.

1.5 Thesis Outline

This thesis is divided into three parts: The first deals with methodology, and includes information about the Social fMRI approach, the technical system that incorporates its ideas, and the living laboratory experiment that implements it. The second part includes results and analysis, and the third part includes a final discussion, a look ahead, and the thesis' conclusion.

Chapter 1 - Introduction: We give an overview of what we want to accomplish in this work, and what is presented in the thesis. We briefly present the research approach and the motivation for doing this work. We highlight the goals of the thesis and its key contributions, and outline the remainder of the document.

Part I - Methodology

Chapter 2 - The Social fMRI: We discuss the Social fMRI approach. We review relevant methodologies and sample some of the previous work related to data collection for understanding human social behavior - from traditional methodologies to approaches currently used in computational social science. The goal of this review is to show where our approach work fits within the landscape, and to highlight some of its advantages and potential in this context. In addition, we discuss hypotheses and research questions we hope to contribute to.

Chapter 3 - Funf: Open Source Sensing Framework: We present our data collection system for mobile phones, called Funf. We review related systems, detail the principles of the Funf framework, describe our experimental system deployed during the 15 months-long Social fMRI study, and finally go over the modular system architecture of the Funf Open Sensing Framework.

Chapter 4 - Funf: Open Source Sensing Framework: We describe the Friends and Family living laboratory study. We review experimental design, methodology, and procedure, including its sub-components, timeline, and aspects of privacy and subject protection. In addition we give an overview of the community in which the study was conducted and its subject pool.

Part II - Experimental Results

Chapter 5 - Virtual Imaging Chamber: Investigating a Living Community: We give an overview of the Friends and Family dataset and some basic statistics about the data and the different social network modalities that can be spanned for the same people from the collected interactions and reported data. We then 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 5.4. In Section 5.5 we continue with looking at the social fabric by investigating network effects on decisions and choice.

Chapter 6 - Investigating Social Mechanisms: Out of several interventions conducted over the past year and planned for the upcoming months, in this chapter 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 [Centola and Macy, 2007] like effect related to pre-existing social ties between participants. In addition to the main analysis of the experiment's direct results, a great amount of work has gone into the careful experimental design of the intervention so that the in-situ experiment would be comparable to the reliability and intentionality of ex-situ laboratory experiments. In the chapter we touch on these aspects as well.

Chapter 7 - Learning about Learning: Tools for The Practitioner As the field of computational social science matures, there is need for more structured methodology. One that would assist the researcher or practitioner in designing data collection campaigns, understanding the potential of collected datasets, and estimating the accuracy limits of current analysis strategy vs. alternative ones. Such methodology would assist it in the process of maturing from a field of craft and exploration into a more rigorous field of science and engineering. In this chapter we investigate the properties of learning and inference of real world data collected via mobile phones over time, and over varying sample sizes. We look at the dynamic learning process over time, and how the ability to predict individual parameters and social links is incrementally enhanced with the accumulation of additional data. We develop several models that predict social and individual properties from sensed mobile phone data in the Friends and Family dataset, including detection of life-partners, ethnicity, and whether a person is a student or not. Then, for this set of diverse learning tasks, we investigate how the prediction accuracy evolves over time, as new data is collected. Finally, based on gained insights, we propose a method for advance prediction of the maximal learning accuracy possible for the learning task at hand, based on an initial set of measurements. This has practical implications, like informing the design of mobile data collection campaigns, or evaluating analysis strategies. We then use the same approach to evaluate the connection between possible learning accuracy and the size of a subject sample within the community, and again are able to model learning accuracy functions along the axis of sample size.

Part III - Finale

Chapter 8 - Conclusion: Closing remarks and summary.

Part I

Methodology

— *"The study of anything as complex as the social life of an entire community requires the use of most of the methodological tools available to the social scientist."*

Leon Festinger (1919-1989)

2

The Social fMRI

The proposed and demonstrated Social fMRI approach combines two main components: The first is a "virtual imaging chamber", which is composed by the integration of a multitude of ubiquitous sensors and data sources for collecting a very rich and highly dimensional dataset on a group of subjects, as they live their everyday lives. The second component is the "functional discovery" component, which includes the delivery of carefully designed stimuli, or interventions, to the measured set of subjects. This entails a methodology for doing experimental designs in the real world, where the specific intervention is tailored to fit the lives and characteristics of the individuals and community.

This chapter includes a review of relevant methodologies and sample some of the previous work related to data collection for understanding human social behavior - from traditional methodologies to approaches currently used in computational social science. The discussion is framed around some metrics that help us compare between works in this broad space. The goal of this review is to show where our approach work fits within the landscape, as well as highlights some of its advantages and potential in this context. Following the review, the general principals of the Social fMRI approach are presented, followed by a discussion.

2.1 Ubiquitous Social Observatories

It is a challenging task to compare experiments and datasets across disciplines, populations, and research topics. Many different fields are interested in understanding social behavior in the context of investigating people and their behavior - these include sociology, psychology, and in particular their social-psychology "offspring", public health and medicine, economics, marketing, as well as many disciplines. Nevertheless, As we review the history of experimental work in these fields, we can notice some emerging trends in methodologies and experimental approaches, as well as the evolution of those trends as technology advances. I shall refer to these types of experiments as *'social observatories'*, as their main purpose is to observe and collect data about target individuals and groups in order to answer research questions. These questions might be related to the subjects, or to the act of collection itself, like the MyLifeBits project " [Gemmell et al., 2002, Bell and Gemmell, 2007].

This discussion aims to lay out the landscape of social observatories and enabling technologies and approaches along three cross-disciplinary axes: First, is the *duration* of data collection - which could range from just a few minutes of survey or interview, all the way to the full lifetime of a person, or even several generations. The second axis of comparison is the *size* of the subject pool, typically referred to as the "N" of the study - which has great relevance to the statistical power of the experiment and the strength and generalizability of the conclusions that can be drawn from it. Finally, we add another dimension of comparison, which is related to the a rough notion of the amount of information that is available in the data - some rough notion of "*throughput*" or "*richness*" of the study from the data sense. The idea is that dataset throughput or richness (these terms are used interchangeably in the document) 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 sampling at 1KHz still does not collect a lot of information). This thesis uses this definition of richness as a qualitative term, although it could be quantitatively defined based on information theoretic and signal processing terminology, this would be a body of work on its own, and it out of the scope of the current context.

Figure 2-1 gives a high-level qualitative overview of social observatories and datasets, comparing them along axes of sample size (represented by circle size), duration, and throughput/richness. This illustration highlights the potential of ubiquitous technologies for the design of "social observatories" and the collection of very rich datasets. To get a sense of orientation in Fig. 2-1: An "omniscient" or "god's eye" view would have maximal richness (as it knows everything, all the time) and maximal duration (i.e. all of a person's lifetime), and so it would be on the top-right most corner of the space (*9* in Fig. 2-1). Web browsing done in "anonymous" or "incognito" mode, where no cookies or cross-session information is available, would give a data-collector on the server side very low richness to the individual browsing the web - mostly limited to the information directly collected on the browsed website, and only for the duration of the person's browsing session. Therefore, it would be very close to the origin of the axes, at the bottom-left corner of Fig 2-1.

The remainder of this section will review the different areas of the space spanned between the anonymous single web query all the way to the omniscient observer.

2.1.1 Traditional Approaches

At the bottom of the diagram are traditional sociology studies as well as many of the corporate "donated" datasets, which we discuss in section 2.1.2. Eagle reviews traditional social science methodologies and sample datasets [Eagle, 2010], and notes that traditional social science research commonly suffers from characteristic issues: First, most of the data is collected via self-reports, which are known to be plagued by missing data and inaccuracies. Second, there is an absence of extensive longitudinal data. Finally, the scale of studies is limited due to the cumbersome and demanding collection methodologies. On our axes of comparison, these properties put most of these studies close to the bottom of the richness scale and time-span scale, and a relatively small sample size.

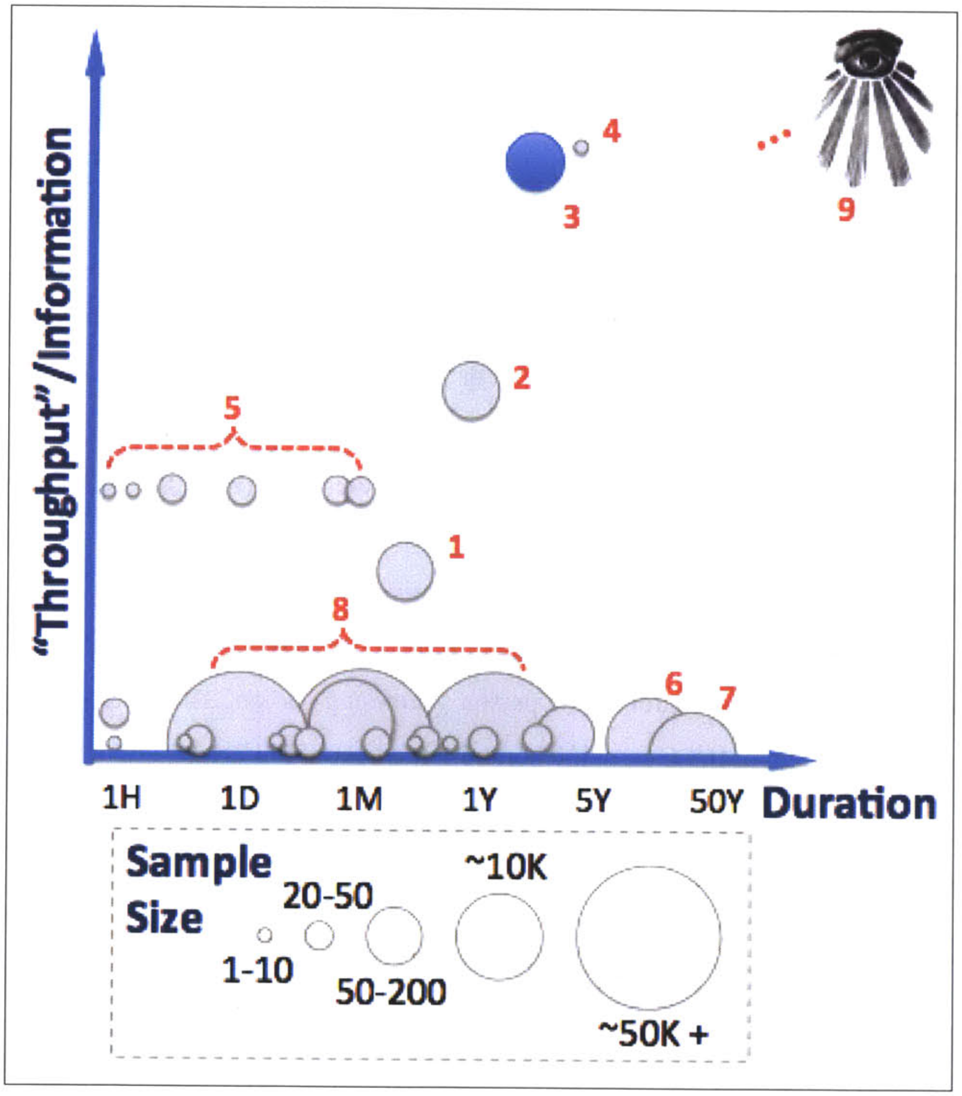

Figure 2-1: Qualitative Overview of social science "observatories" and datasets, along axes of data collection duration, qualitative"throughput", and the size of the subject sample. (1)Reality Mining [Eagle and Pentland, 2006], (2) Social Evolution [Madan et al., 2011], (3) Friends and Family Dataset, (4) rich-data pioneers [Roy et al., 2006, Gemmell et al., 2002] (5) Sociometric Badge studies [Olguín et al., 2009], (6) Midwest Field Station [Barker and of Kansas. Midwest Psychological Field Station, 1968], (7) Framingham Heart Study [Dawber, 1980], (8) Large Call Record Datasets [Gonzalez et al., 2008, Eagle et al., 2010, Hidalgo and Rodriguez-Sickert, 2008] , (9) "Omniscient"/All-Seeing View.

Examples of representative studies are those by [Cooley, 1909, Davis et al., 1941, Burt, 2004], and Festinger's work at MIT's student housing [Festinger et al., 1950]. Some of the larger scale and longer duration traditional data collection initiatives include, for example, Barker's work at the "Midwest Midwest Psychological Field Station" in Oskaloosa, Kansas, between 1949-1972 [Barker and of Kansas. Midwest Psychological Field Station, 1968] and the Framingham Heart Study [Dawber, 1980] (*(6)* and *(7)* in Fig. 2-1). The Framingram Heart Study stands out for its duration and a subject pool of several thousands, however its "throughput" and data-richness is low, as subjects were sampled approximately once every three years, generating a maximum of seven data samples per person.

2.1.2 Corporate and Government "Data Donations"

As Lazer, Pentland et al. describe the potential of computational social science [Lazer, 2009], they also highlight some of the 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.

Because of these and other challenges in gathering copious amounts of data from large numbers of subjects, many research works have used social datasets gathered by commercial entities like telecommunications operators [Gonzalez et al., 2008, Hidalgo and Rodriguez-Sickert, 2008, Cebrian et al., 2010, Eagle et al., 2010, Onnela et al., 2007, Cebrián et al., 2010], financial institutions [Krumme et al., 2010, Bollen et al., 2010], online social networking providers [Onnela and Reed-Tsochas, 2010], and government agencies, to investigate social science questions.

The pervasiveness of mobile phones the world over 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 [Gonzalez et al., 2008]. Eagle et al. combine nationwide communication logs with census data at the community level, and find that the diversity of individuals relationships is strongly correlated with the economic development of communities [Eagle et al., 2010]. 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 (Label *(8)* in Fig.2-1). As example, in [Hidalgo and Rodriguez-Sickert, 2008] 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 [Onnela and Reed-Tsochas, 2010]. Finally, there is the *"offline"* nature of most existing datasets which 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.

2.1.3 Reality Mining

An alternative approach is a bottom-up one, of collecting data at the level of the individual. Eagle and Pentland [Eagle and Pentland, 2006] 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 at the individual level, the subjects' social network can be accurately detected, as well as regular patterns in daily activity [Eagle and Pentland, 2006, Eagle et al., 2009]. This initial study was then followed in Madan et al. [Madan et al., 2010a], who conducted the Social Evolution study, 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 [Madan et al., 2011]. Other examples for using mobile phones for social sensing are those by Montoliu et al. [Montoliu and Gatica-Perez, 2010], Lu et al. [Lu et al., 2010], and from CENS center, e.g. Campaignr by Joki et al. [Joki et al., 2007] and additional works as described in [Abdelzaher et al., 2007]. Most of these were of an observational nature, and have not performed controlled experimental interventions for exploring social mechanism.

If we want to map these types of experiments on our 3-axis landscape, we see, for example that both the Reality Mining and Social Evolution studies had around the order of 100 subjects, while other similar studies are either around the same size or smaller in scale. The duration of most studies are from several months to a bit under one year. The original Reality Mining study collected only a limited number of signals (approx five), and was a proof of concept for the great potential in this type of approach. However, even those signals, collected continuously over the duration of the study, as well as the auxilery information about the differnt subjects (e.g. their role and lab affiliation), gives a tremendous boost to the study on the richness axis, compared to the datasets collected before it (*(1)* in Fig. 2-1). The Social evolution study took this trend further, increasing both duration and the number of signals collected (*(2)* in Fig. 2-1). The Friends and Family living laboratory study, the core data collection effort described in this thesis, continues the trend and increased both duration (15 months), and number of signals (tens of signals per subject), and is illustrated around *(3)* in Fig. 2-1.

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 [Olguín et al., 2009, Olguín Olguín, 2007]. A key aspect of the Sociable Badges is that they have been deployed in studies where sensor feedback was given to the corporate participants [Olguín et al., 2009, Kim et al., 2008].. While being rich in terms of throughput, resolution, and number of signals collected, these studies are limited in the sense of rich-data because the badges are constrained to an office related context and are only worn during work hours. The do not collect information throughout the subject's daily lives, and they do not collect auxilery information such as financial behavior, out-of-work social life, etc. This why we mark these types of experiments around the region of *(5)* in Fig. 2-1.

2.1.4 Rich-Data Pioneers

There are several notable pioneers in the collection of rich data from everyday life, including Gordon Bell and the "My Life Bits" project ([Bell and Gemmell, 2007, Gemmell et al., 2002], Steve Mann "The Human Cyborg" [Mann and Niedzviecki, 2001], Thad Starner [Starner and Rhodes, 2004], and Roy's Human Speechome Project [Roy et al., 2006] (see Fig. 2-2). Though operating on a small N (usually N=1), they have challenged what is possible to automatically collect on a person's life, pushing the envelope of rich data collection (illustrated around *(4)* in Fig. 2-1).

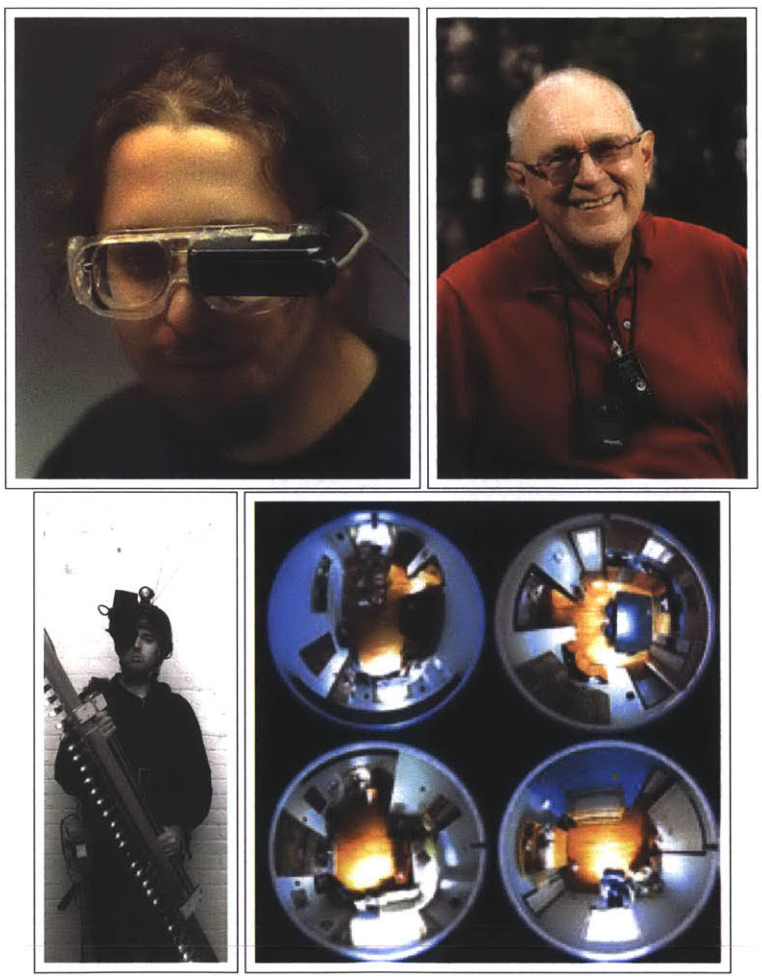

Figure 2-2: Rich data pioneers. Clockwise from top left:Thad Starner, Gordon Bell, Steve Mann, The Human Speechome project

2.2 Benefits/Strong points of the Social fMRI Approach

As discussed in previous sections, data collection at large is many times a questions of resources and accessibility to experimental subjects and information. For many given services, in particular digital communication and Internet services, data might be collected at different points along the path from the individual through the various infrastructure providers, all the way to the specific service provider. Each of these points of collection has its own benefits and challenges, and might lend itself better to one of the three axes over another.

Governmental and medical record data may be accumulated throughout a person's life, giving it a benefit on the duration axis, but it might have very low richness as far as what is known on the individual in between data-sampling, which may be years apart. Corporate service and infrastructure provider data might get full information on all users of that particular service. The concentration of all user data in one place affords very large sample sizes relatively easily, and in context of that service the data is very rich. However but there is no information about many other aspects of the subjects' life. Corporations have but narrow "peeping hole" access into the individual's life, and when such data is shared with researchers it usually loses further richness and level of detail.

By performing the data collection at the individual level, we are actually collecting it at the source. At the point of the individual, there is the accumulation of all the different services they use, behaviors, and so on. However it is much more challenging, and requires a great deal of logistics and recruitment efforts. It is very hard and costly to conduct very large scale experiments. This is especially true for highly fine grained experiments, where subjects are given phones or other sensors, that require installation or support, and constant engagement with the research team as the study progresses. Nevertheless, the currently manageable numbers of 100-200 seem to be useful enough for answering many research questions with sufficient confidence and statistical power. Another reason for keeping the number somewhere around or below 150, when detailed ground-truth social information is required, might be grounded in our evolution as a species. Number in this area fit well with Dunbar's social evolutionary theory regarding the number of people humans are able to maintain a relationship with [Dunbar, 1993]. Throughout the our Social fMRI investigation we want to understand the subject's own perspective about their relationships with the other subjects in their community. To do so, we ask, for example, about social closeness between all participants in the study. Every participant has to answer about different aspects of their relationship with every other participant, and subject pool sizes much larger than Dunbar's number could become quite tedious for participants. We shall refer to experiments in our scale as "*Dunbar scale*" experiments, and argue that this size fits well with studies that aim to measure a physical community in situ.

The Social fMRI is complementary to the existing dataset acquisition approaches. We hope that it will be possible to bridge between the highly rich and detailed but smaller sample size Dunbar scale experiments, and the larger sample but low richness datasets collected at the service provider side. Additional data collection efforts might also fall in between these two experimental approaches - for example experimental software that can be downloaded through a mobile phone's application store and invites potential users to download it and "donate" data to scientific initiatives. On one hand, it will be much less rich than a physically collocated Social fMRI experiment, but it would be much richer in information than the corporate datasets. On the other axis, while being much smaller in size than the full set of service customers in the latter type of dataset, it would have the potential to be orders of magnitudew larger than a physically collocated Social fMRI study.

2.3 Social fMRI Research Goals

There are many different research questions that we would like to try and answer at large. They can be unified under two high-level themes of investigation: First is the attempt to understand *how people make decisions*, with the emphasis on the *social* aspects involved in decision making. The second theme is investigating ways to *empower people to make better decisions* using personal and social tools. These two themes could be applied in a broad range of contexts that touch on many aspects of life - from health and wellness to a better understanding of social dynamics; from purchasing behavior to community organization. A key objective is to do all of this in the real world, as people live their everyday life, and not under artificial laboratory conditions.

2.3.1 Specific Research Questions

There are many different research questions that we wanted to investigate, and are able to investigate in a Social fMRI type of study. They can be unified under two high-level goals: First is the attempt to understand *how people make decisions*, with the emphasis on the *social* aspects involved in decision making. The second theme is investigating ways to *empower people to make better decisions* using personal and social tools. These two themes could be applied in a broad range of contexts that touch on many aspects of life - from health and wellness to a better understanding of social dynamics; from purchasing behavior to community organization. A key objective is to do all of this in the real world, as people live their everyday life, and not under artificial laboratory conditions.

Under the umbrella of the two high-level investigation goals, our Social fMRI study was driven by a combination of specific hypotheses, as well as an exploratory aspect - related to the social science questions and also the methodological approach.

There is a set of hypotheses and specific investigation goals that were the driving factors for initiating a Social fMRI longitudinal study. These are derived from our scientific motivations in Section 1.2.1. Following are some examples or preconceived hypotheses that we wanted to "cover" as much as possible with the data collection planning:

2.3.1.1 Properties Of Networks

As discussed in Section 1.2.1, there is a lot that we still do not understand about the structure of human networks, as well as the differences between multiple network modalities and social tied spanning even the same set of individuals. Attempting to answer these questions would drive our experiment design choices. Some specific research questions:

- Investigating the hypothesis that there are inherent differences in characteristic properties between different network modalities.
  Early works show face-to-face and phone communication networks demonstrate different characteristics and network measures [Madan, 2010, Olguín et al., 2009], however previous works have not attempted a targeted collection of a large number of network

modalities for the same individuals. A study goal targeting the collection of as many networking modalities as possible for its subject pool might help shed better light on the characteristics of different network modalities. Network modalities which we are able to collect (and have collected in the experiment described in Chapter 4) include (but not limited to):

Networks of mobile-phone communication:

- Network spanned by phone calls.
- Network spanned by SMS messages.
- Network spanned by phone contact book.

Networks spanned by sensing of Physical Proximity / Face-to-Face interaction:

- Relative proximity network spanned by Bluetooth discovery scans (similarly to [Eagle and Pentland, 2006, Madan, 2010, Paulos and Goodman, 2004]).
- Proximity networks spanned by common co-location with similar wifi access points.
- Proximity networks spanned by common co-location with similar mobile cell-towers.
- Proximity networks spanned by predefined distance between absolute GPS coordinates.

Online Social Network Activity:

- Facebook declared friendship network (A is "friend" of B)
- Networks spanned by interaction records of facebook activity (e.g. messages between users, "likes", etc.)
- Networks spanned by virtual "co-location" - i.e. two users are both linked to a common entity, like a group or event.

Self-perceived networks:

- Network spanned by reported closeness of each participant on every other participants.
- Additional networks spanned by specific survey questions (network spanend by pairs who said yes to "Would you feel comfortable asking participant X for a loan of $10").

For each of these networks we could perform basic property analysis and focus on characteristics that fit "Dunbar Scale" networks. These include standard properties like degree distribution, centrality measures [Newman, 2003], network motifs [Milo et al., 2002], or leadership, bonding, and diversity (LDB) analysis [Richards and Macindoe, 2010a].

- Investigating the stability of these networks over repeated questioning.
- Investigating the evolution of these networks over time (as many snapshots are gathered)
- Is there a notable difference between networks of communication vs networks of influence?
- How well can we infer the user's self perceived networks from the automatically sensed networks?

2.3.1.2 Network Effects and Diffusion

In this area specific questions are:

- Do social and cultural groups create boundaries that affect spread and diffusion?
- How do different communication channels play different roles in the spread of information? (e.g. face-to-face vs. phone communication vs. online interaction)
- How much of our behavioral adoption is driven by our perceived social ties vs. our exposure to individuals who we do not necessarily know or consider friends?
  Madan et al.'s inital result suggest [Madan et al., 2011] that in some cases, behavior is driven by exposure more than it is by perceived social ties. We want to follow up on this and investigate whether ties self-perceived friendship or closeness between two people are correlated with common behavior or adoption more than physical proximity exposure (where such closeness does not exist).

In order to be able to investigate diffusion of behavior, we must make sure to include the collection of decision points and behavioral data, in particular ones that might be influenced by social ties (e.g. choice of shopping, food, or entertainment activities)/

2.3.1.3 Network Evolution

There are many interesting questions that are in the context of network formation and evolution. A key one is "How do people choose their friends?". It might be related to the ideas of homophily [McPherson et al., 2001], propinquity [Festinger et al., 1950] or other factors. Even within these factors, there are many "competing" ones - does being neighbors influence the formation of a tie more than being in the same academic department? how does sharing religion, ethnicity or national/cultural background come into effect? What about other features like personality types or shared hobbies? In order to be able to answer these questions, we would need long enough data collection. Ideally, we would have a set of participants who are new to the community and are joining it during the period of collection, so that we can trace their behavior as they integrate and "settle down" in the new community.

2.3.1.4 Social and Behavioral Modeling

There are many potential models that would be useful - like inferring and predicting social and behavioral traits based on automatically sensed data, or detecting special events or other occurrences. Just a few examples: Can we infer personal traits? how hard is it to learn the identity of married couples? Are there unique patterns of people with or without children? Can we infer a person's occupation? income level? Can we detect social gatherings? Can we classify them? Can we detect extreme events (e.g. fire alarm) what are inherent correlations among sensed signals, or among different individual and network properties? Using a rich data collection combined with broad collection of "ground truth" information, can help us build models that we might be able to later test on larger but lower richness datasets.

2.3.1.5 Social Mechanism Design

A key component of the Social fMRI is the ability to conduct targeted interventions to help us uncover subtleties of existing social mechanisms, as well as try out new mechanisms and understand what components work or not. We want to try to leverage social and personal tools to empower people to make better decision in their life. Could we design experiments to help us understand or even quantify social capital [Hampton and Wellman, 2003, Putnam, 2000]. We want to design interventions and stimuli that help us isolate our test variables, but in a way that is safe for the subjects well-being. An example of an intervention might be as simple as showing users parts of their own logged data and measuring the effect of seeing this information on their future behavior. More intricate interventions could be policy changes in the community or conducting collective games whose' rules impose social components that help us investigate relevant questions,

2.3.2 Discussion: Potential of Rich Data As an Experimental Approach

The challenges faced by field studies of this type, as well as the approached taken to deal with them, are just as true today as they were in the 1940's, when first described by Leon Festinger [Festinger et al., 1950], p.9:

> "Field studies such as this one which attempt to describe, quantify, and experiment in the community they investigate, present intriguing methodological problems. The study of anything as complex as the social life of an entire community requires the use of most of the methodological tools available to the social scientist: informants, participant and nonparticipant observation, informal and standardized interviewing, sociometry, and field experimentation. These techniques all supplement one another. They gather data on different aspects of the community and way of life. It is only the integration of these diverse data that permits us to piece together a coherent and comprehensive picture of the community under study."

While challenges and the multipronged approach are similar, the technology has come a long way since Festinger's days. Using digital collection, processing, analysis, visualization, and communication techniques we are able to augment the traditional toolset of the social researcher.

If we evaluate our individual research questions independently, and the data collection we would like to do in the hope of shedding new light on it, we would see that a significant overlap in the desired collection effort. We would like to attempt to answer as many of these research questions as we can, without harming the integrity of the study, and while carefully designing the various collection and intervention components so that they do not interfere with one another. At the very least, if not find definitive answers to our current questions, we hope to at least collect data that would help investigate these questions in the future.

The long duration of such an experiment also affords us the ability to try out various approaches, evaluate them while the study is ongoing, and modify or add questions or components as needed. Ideally, an experiment as this should be run in a way that allows data to be collected as the study is ongoing, in order to find and fix bugs and do initial evaluation of the data and hypotheses. By collecting auxiliary data that is relatively low-cost as far as effort required, we are able to explain more of the variability in the data. There might be signals or pieces of information that we did not aniticipate in advance to be relevant for evaluating a specific hypothesis. However due to the fact it was collected, it allows us to find hidden or unexpected connections between different variables, or latent variables that affect the outcome in a way we did not anticipate.

The dataset itself could be used to answer questions that we have not thought about, or even questions from different domains. The original Reality Mining dataset has been used in dozens of publications already, including epidemiological modeling of mobile phone viruses [Yoneki et al., 2008] testing of new information security mechanisms [Altshuler et al., 2010], peer-to-peer routing algorithms [Miklas et al., 2007], modeling daily routines [Farrahi and Perez, 2008], identifying individuals based on location traces [De Mulder et al., 2008], and many others. Sociable badge datasets of face-to-face interactions have been used to create epidemiological models for the workplace [Waber et al., 2010]. A longitudinal Social fMRI study, that includes richer data collection than these and most other existing field studies, has the potential to contribute to many investigations the researchers have not originally imagined.

To summarize, the immense effort and costs associated with a longitudinal study in the real world, motivate us to try and cast a wide net, and get as much utility out of the invested resources. Rich data collection gives us the potential to answer many different research questions that we know in advance that we want to ask, or domain areas that we hope might be answered with the data, seeding questions for future analysis and experimentation. It allows us to go back in hindsight and answer new questions we might not have specifically targeted in advance, and finally is the potential for future researchers to be able to answer new questions we could not even conceive.

—Nothing is built on stone; all is built on sand, but we must build as if the sand were stone.

Jorge Luis Borges

3

Funf: Open Source Sensing Framework

As truly ubiquitous wearable computers, todays mobile phones are quickly becoming the primary source for social, behavioral, and environmental sensing and data collection. Phone-based sensor data is used for enabling a broad range of research projects and application domains: research in areas of health and wellness [Madan et al., 2011,Estrin and Sim, 2010], environmental sensing, e.g. [Joki et al., 2007], sociological and psychological investigations, e.g. [Eagle and Pentland, 2006, Killingsworth and Gilbert, 2010], support for disaster and crisis response, e.g. [Anokwa et al., 2009,Banks and Hersman, 2009], as well as event triggers that can be used for building richer and more personalized mobile experiences, are just a few examples for uses of this kind of data. The mobile phone sensing platform is the central component in implementing the "virtual imaging chamber" idea of the Social fMRI.

Googles Android platform [Google, 2011] has allowed researchers and developers to do more things with mobile devices than ever before. However, there is still a great gap between having API calls for accessing on-phone sensors and information, to having a "deployable" system or an end user application that fully utilizes this access. Such additional components include, for example, specialized mechanisms for privacy preserving data collection (e.g. encoding human readable text so that its useful for scientific analysis but does not expose the original information), delay tolerant communication with a back-end server (data collected locally when server not available, and uploaded in background whenever users connect to the network) [Fall, 2003], smart data collection algorithms that maximize battery life, the ability to remotely configure the data collection settings, encrypting data files on the phone's memory card for added protection, and so on. At the time leading to our Social fMRI field experiment, the Friends and Family Study, there was no available software system that supported these needs, or an open source one that we could adjust and modify for our needs. Therefore, we needed to develop the study's data collection framework on our own.

Figure 3-1: Funf Logo

The first iteration of the sensing framework, named "Funf", was used in the Friends and Family study. It was deployed in the field with the study's 140 participants, through a span of 15 months. Through this time we iterated on the framework and its settings, optimizing performance, fixing issues, and adding features. With this, we decided that we want to share our system and experience with the community, and turn it into an open source framework for researchers, developers, and end-users. The second iteration of Funf was launched as an open source framework in early October 2011. The Funf Open Sensing Framework (logo in Figure 3-1) is an extensible sensing and data processing framework for mobile devices. It currently supports Android OS based devices, though some of the architecture could be posted to additional platforms in the future. The core concept is to provide an open source, reusable set of functionalities, enabling the collection, uploading, and configuration of a wide range of data types. Together with the Funf framework, which is aimed at developers and coders, we deployed "Funf Journal", an Android application for researchers, self-trackers, and anyone interested in collecting and exploring information related to the mobile device, its environment, and its user's behavior. It is built using the Funf framework and makes use of many of its built-in features, and is freely available in the Android Market. We hope to foster a community around the Funf and the ideas of code sharing and re-use, where new development efforts would go towards extending a common platform rather than creating redundant functionality.

This chapter reviews related data collection software for mobile phones, reviews the ideas of the Funf framework, describes our experimental system deployed during the 15 months-long Social fMRI study, and finally the modular system architecture of the Funf Open Sensing Framework.

3.1 Motivation and Goals

3.1.1 Mobile Data Collection Platforms

There are several mobile data collection systems that have been developed over the last few years. A partial list of recent systems includes: Jigsaw [Lu et al., 2010], SystemSens [Falaki et al., 2011], the Nokia-based system by [Do et al., 2011], and Paco [Evans, 2011]. Most of these systems, as well as many other smaller scale applications are developed for specific data collection experiments or initiatives, and are usually tailored at the needs of that specific project. At the time we started to work on the Friends and Family study, there was no system available that answered the data collection requirements of the study, let alone an open source and free one. Therefore, we needed to develop the study's data collection framework on our own.

3.1.2 The Need

Our problem was a common one in the field - and many peers that we have talked with faced similar problems, and had to develop their own framework from scratch since nothing suitable was available for use. There is a great overlap between the needs of most if not all of such projects. Many of the system requirements for different and cross-disciplinary data collection initiatives are very similar. All of them need to deal with accessing sensors, storing data on the mobile device, and transferring it to a back-end of some sort. All would benefit from the ability to do remote configuration and debugging. All share the need for dealing with battery, processing, storage, and bandwidth limitations. This leads to an incredible waste of resources, duplicating efforts that go towards redundant functionality. Based on the experience gained through the Friends and Family study field deployment, we decided that to release Funf as an open source and free framework that would help future initiatives.

When considering limitations of existing platform, we see that most are not open and freely available, at least not at the time of writing. For the ones that are "open source", they were not designed as a framework for others to build on, but usually as merely open sourcing the code of the respective project "as is". They are usually not well documented, if at all. Many times, prototype-level code, which is the case for many existing academic systems, is not robust enough to be easily reused and repurposed. Finally, we have not seen any existing framework that offers access to as many data signals as our initial Funf version had supported.

Rather than having different developers and research teams re-invent the wheel, we are aiming to share the tools and experience we have already gained through our past deployments, and turn our system into an open source framework for researchers, developers, and end-users.

3.1.3 Choice of Operating System

One of the issues with developing for mobile platforms is the great diversity of platforms ([Oliver, 2009]) and the volatility of the mobile operating system (OS) market. Because of this, previously developed systems, developed for OS platforms that turned out to be short-lived, quickly become obsolete. Examples of these are the Nokia phone based system developed for the original Reality Mining study [Eagle and Pentland, 2006], or the Social Evolution study's "Windows Mobile" based framework [Madan, 2010]. In addition, not all mobile OS platforms allow open access to developers, and some greatly limit the things that developers could do with a standard application (e.g. in Apple's iOS). Because of this, some of the frameworks require "rooting" the phone in order to gain administrator access to it, which helps deal with the internal access issues of the sensing software, but also limits the potential installation base of the software.

Google's Android operating system seems to provide a "sweet spot" for scientific research: It offers great developer access to its internals, more than any other mobile OS. It also offers a comprehensive inter-application communication and privacy framework [Google, 2011]. Android is currently the most popular smartphone platform, and has been growing rapidly in recent years. There are currently over 500 different android device models. In Nov 2011, there were over 200 million active Android devices around the world, with more then 550,000 new devices activated daily [Whitney, 2011]. Android currently has the largest smartphone market share, with 47% in Nov 2011 [ComScore, 2011], and it is still on an upwards trend. It is also expanding to tablets, TVs, and other devices. These considerations led to our choosing it as our framework for our initial Friends and Family deployment, as well as for the base of the Funf Open Source release as far as cost-effectiveness and the expected longevity of our development efforts. However, there is no reason why Funf could or should not be extended or ported to other operating systems in the future.

3.1.4 Design Goals

With Funf, we consider two main scenarios at the developer level: First, is a developer who does not have an existing implementation, and has need for a data collection framework. In this scenario, we want to do our best so that Funf is the easier and more efficient option vs. developing new code from scratch. Either Funf already supports all desired functionality and it is mostly a matter of configuration and setup, or, Funf enables easier integration of new functionality when compared to developing an entire system from the ground up. We wanted to make sure Funf's architecture is both extensible and modular, as well as provides a core set of built-in functionality to cover the majority of use-cases.

The second scenario is that of a developer who already has an existing system. Existing systems have different specialties and strong points. One might support high-quality and efficient activity recognition functionality, while another have more efficient location logging, or a specialization in certain type of sensing - like social interaction, or health related sensing. In this case we have defined, together with peers from the mobile sensing community, an inter-system communication protocol - the *Open Probe Protocol*, or *OPP*. OPP allows supporting systems to communicate with one another, and treat one another as a data source or a data client, without the need to go into the internals of either system. OPP leverages Android's inter-application communication protocols.

Another design decision was that we did not want to enforce a specific back-end architecture or implementation. Organizations and individual developers have existing back-ends or back-end technology preferences, and we did not want to enforce our own proclivities. Funf defines a back-end communication protocol, and any back-end implementing this protocol could support Funf server communication. We provide a Python based back-end example, and we already received a community code contribution of a PHP based server. In addition, Funf's modular architecture allows for the protocol itself to be modified or replaced completely.

Existing data collection systems might not be as respecting of user privacy as possible, either deliberately or by negligence. For example, it is very easy for nearly any app, asking the right permission, to access the user's phone numbers or text messages and do as it wishes with them. This includes saving them in plain text on the memory card, or sending them to any server in the world, given the right permissions. Funf's default settings include the same privacy preserving mechanisms approved by MIT's institutional review board (IRB) for increased subject protection. For example, by default no human-readable text is ever saved, and it is instead encoded with a one-way hash encoding.

We have a firm belief that data collection at the level enabled by Funf is so privacy sensitive that it requires openness and disclosure. Android's built-in permission framework somewhat describes what type of things an app might do, but in many cases the explanation is not clear enough to users and leaves much room for different app behaviors. Funf itself is open, and Funf's licensing terms (LGPLv3 [http://www.gnu.org/licenses/,]) make it so that anyone changing its internal workings would also have to disclose those changes publicly. We also hope that any apps using Funf, or any apps doing similar data collection, will be open or at least disclose exactly what they are collecting and what would be done with the information (and as new uses for the data arise, the end-user would be involved in consenting to them).

3.2 Who Is Funf For?

The interaction with Funf is designed around a multi-tiered model. In the outer-most tier, no technical or software development skills are needed, while in the inner-most tier, core-level developers can add new functionality and enhance the framework, or easily package their inference algorithms for use by others. Funf is currently available only for Android-based devices.

Figure 3-2 illustrates the four tiers of interaction with the Funf framework, which are detailed in the following subsections.

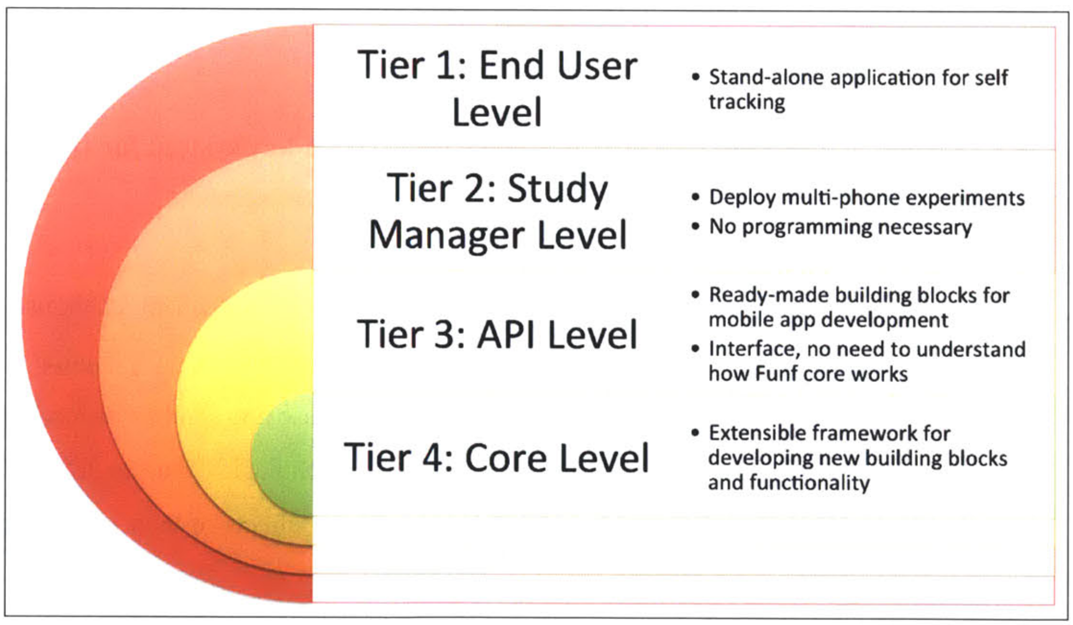

Figure 3-2: Illustration of the multiple tiers of interaction with the Funf framework.

3.2.1 Interaction Tier 1 - End-user level: Standalone App for End users

At the outer-most layer we provide a standalone mobile application for easy collection of data from a single phone. This application allows any user to collect and explore information related to their mobile device, its environment, and the user's behavior. It is well aligned with the burgeoning trend of self-tracking and the "quantified-self" movement [quantifiedself.com, 2011]. The appliction can be used by students and researchers to easily conduct automatic sensing and data gathering using mobile phones. For example, it could be used by university students who want to collect a dataset about their GPS location, physical, or social activity, to be used for class projects etc.

"Funf Journal" is the first application released for this tier and it is already available for free download on the android application store, the Android Market. It enables manual configuration of all data collection features, exporting the data in multiple ways (e.g. via email attachment). A simple pc-based script for visualization of some of the data signals is provided as a sample of its potential, and users can import their data as comma separated values (.csv) text files to any analysis software.

3.2.2 Interaction Tier 2 - Study manager level: Turn-key system for field study managers and researchers

Researchers and others interested in a multi-phone experimental deployment, could use Funf as a turn-key system for deploying and conducting field experiments, very similar to the Friends and Family study described in Chapter 4. They could use it to conduct data collection experiments ranging from health and wellness, social and psychological studies, tracing what apps users are running on their phones, or testing the battery effects of their own 3rd-party Android apps.

The Funf Journal app supports a networked mode where a server is set up to receive automatic data uploads from multiple deployed devices, as well as remote configuration of phone-side data collection settings via the server. Investigators would install Funf Journal on the devices they want to collect data from, set up the data-collection back-end server, remotely configure the desired sensors and data collection behavior, and set up optional components like user surveys. Full and up-to-date instructions and tutorials can be found on the Funf website, at http://funf.media.mit.edu, in particular the "getting started" guide at [Aharony et al., 2011a].

3.2.3 Interaction Tier 3 - API-level: Ready-made building blocks for mobile application development

The Funf framework is available as a pre-compiled java library file (.jar), which exposes an application interface (API) for a set of functionalities and building blocks that can be easily integrated with 3rd-party applications. This allows developers to build Android applications that leverage the capabilities and services of the Funf framework through its 3rd-party developer API, without the need to go into the Funf internals. This would allow developers to save time and focus on the key parts of their app while Funf takes care of things like logging and uploading of proprietary app data and generic phone-sensor data. Detailed API instructions, documentation, and tutorials, can be found at the Funf developer website at [Aharony and Gardner, 2011].

3.2.4 Interaction Tier 4 - Core-level: Extensible framework for development of new building blocks, tools, and algorithms

Core-level developers can go under the hood and use features that are outside the scope of the 3rd-party API. They could leverage the Funf framework's modular architecture and implemented features to focus on the new innovations that they care about - for example a new feature extration or activity detection algorithm. They could also contribute new features and capabilities to the framework, or add new sensor probes and improve performance of existing ones. Detailed Funf codebase documentation and tutorials can be found at the Funf developer website at [Aharony and Gardner, 2011].

3.3 Funf Key Features

Funf incorporates a set of built in data collection tools and functionalities. Key features are listed below. Detailed and up-to-date of the features, specific probes, and other components can be found at the Funf developer website [Aharony and Gardner, 2011]:

- Remote configuration from a back-end server. The phone-side application can be configured to routinely check a remote server and download any configuration updates.

Configuration files can be defined as a formatted text file.

- Many built-in data "probes" (see below).
- Automatic or manual data upload. Automatic upload is done via built-in mechanisms for server communication for data upload and synchronization.
- Delay tolerant implementation ([Fall, 2003]) - When Internet connection is not available, caches data locally until server connection is restored. The data cache size can be constrained by configuration so that it does not take over the devices storage space, but is ultimately limited only by the amount of free space on the device's storage memory.
- One-way hashing and encryption for sensitive data: Does not save any textual data or phone numbers in human-readable format. The framework supports encoding sensitive information by different coding methods, including one way hashing (or cryptographic hashing [Schneier, 1995]) which is a one-way encoding that is not practically reversible, as well as standard 2-way encryption which would allow data to be decrypted given the correct keys.
- Encryption of locally stored database files, and other privacy enhancing features.
- 3rd-party data input API allowing any app to leverage the Funf framework for collecting and storing arbitrary app data.
- Modular probe architecture allows for adding core data probes and modifying existing probe behavior.
- Basic survey system for manual data collection. Surveys can be defined as a text file which is synchronized with the device as part of the remote configuration protocol. A sample survey screen-shot is illustrated in Fig 3-3.
- Various optimizations for prolonging battery life (for example a state machine that adapts GPS logging frequency based on whether the device is indoors or outdoors) and dealing with everyday use-cases (e.g. when SD card is not available because user is copying music files to it).
- Field proven - Deployed for over 15 months with over 100 users in an MIT living laboratory experiment.

3.3.1 Funf Probe Architecture

Probes are the basic data collection objects used by the Funf framework. Each probe is a contained unit responsible for collecting a specific signal or type of information. The *probes* terminology is used rather than "sensors", as probes encompass traditional sensors such as GPS or accelerometer, but also other types of information not traditionally considered as collected by sensors, like file system scans, call-log extraction, or the logging of 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. Adding and updating new probes is described in detail in the developer documentation [Aharony and Gardner, 2011]. All probes support a common set of behaviors (like client registration, scheduling, data transmission protocols, etc.), 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. In addition, probes can be interconnected to one another so that the one probe's output is the input of the other, the latter acting as a client of the first. Funf allows the creation of a hierarchy or even a network structure of probes.

Funf includes a set of built-in probes, as well as a modular architecture allowing the addition of new probes by 3rd party developers. For example, one might be developing an app and would like to log what screens users are spending time on, and other information that might help the developer debug and improve the app. One option to do that would be to develop the totally independent code modules for logging the data, saving it to disk, and then sending it back to the developer's severs, while dealing with a whole range of issues such as protecting user privacy, added battery consumption, memory storage and so on. Alternatively, the developer might write a simple Funf probe that logs all needed information, and leverages the Funf framework's existing modules and experience in dealing with all these issues. Other built-in probes could be leveraged to improve the issue analysis.

There are two main 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. The built in wifi probe is an example for a probe that uses both strategies - it can perform a wifi scan for nearby access-points once every ten minutes, but it also opportunistically listens to wifi scan results initiated by other processes in the system.

Examples of implemented probes currently part of Funf: GPS, Location (based on Android's developer API [Google, 2011]), WLAN, Accelerometer, Bluetooth proximity scans, Cell tower ID, Call log, SMS log, Browser history, Contacts, Running apps, Installed apps, Screen on/off state, Music/Image/Video file scan, Battery status, and more. There are currently more than 35 built-in probes in the system (compared to the original 25 that were implemented during the initial study deployment). A full list of currently implemented probes can be found on the developer documentation site, at http://code.google.com/p/funf-open-sensing-framework/wiki/BuiltinProbes.

3.4 Funf: Initial Field Deployment Version

This section will review the system architecture and deployment setup of the initial version of the Funf system, which was designed for the Friends and Family living laboratory study, and deployed for over 15 months with the study's participants. The system was installed on people's primary mobile phone (given to them by the investigators), and was continuously running and logging data as long as the phone was turned on. During the field deployment, data collection was calibrated were calibrated to last at least 16 hours on a single charge with average usage of the phone (including applications, surfing the web, etc.), so that people could go to work and return home in the evening, though they would usually last for more.

3.4.1 Phone Data Collection

the Funf software was installed on all study phones, and periodically sensed and recorded 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 were collected in this initial field deployment. The system supported integration of user-level apps for additional data collection and interventions. Out of several applications we explored, only one was used in the actual deployment, an alarm clock application that logged study participant's actions (like alarm setting time or number of "snooze" actions). We ended up not using the other applications, like our own version of the Android Market app which recorded user app searches and button clicks, mostly since we did not want to overload the study with compounding or competing interventions. Additionally, the phone system included a survey application. Sample screen-shots are shown in Figure 3-3. 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 hours between charges.

3.4.2 Data Formats and Server Communications

Phone data is is saved in SQLite file format. Every 3 hours the system closed the current db file and started writing to a new file, 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

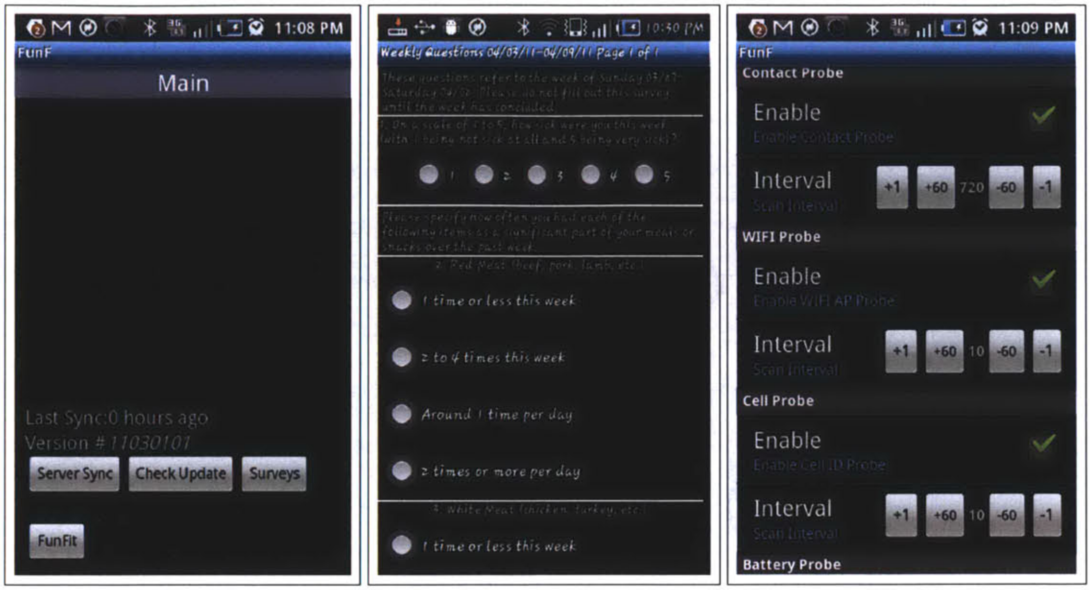

Figure 3-3: Sample screen-shots: Sync-state and version display (left), survey (center), and probe preferences debug screen.

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. Figure 3-4 shows the data flow diagram for the Funf phone-side system.

3.4.3 Back-end

All individual phone-generated SQLite files successfully uploaded to the server were collected at an incoming file directory, to be validated for corruption and processed into a central MySQL database by the back-end software. The software also generated email reports to investigators about the status of the deployed 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 as well as analysis. Figure 3-5 illustrates the back-end data flow.

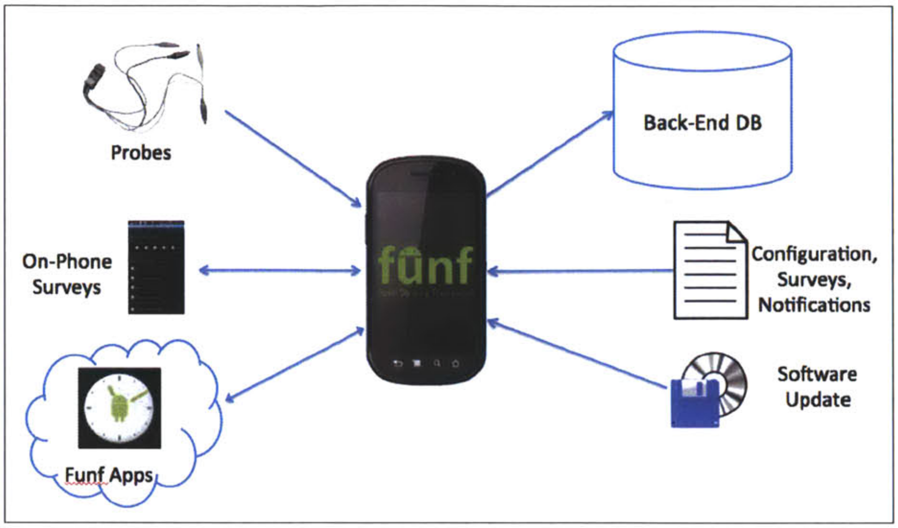

Figure 3-4: Data flow diagram for the Funf initial experimental deployment

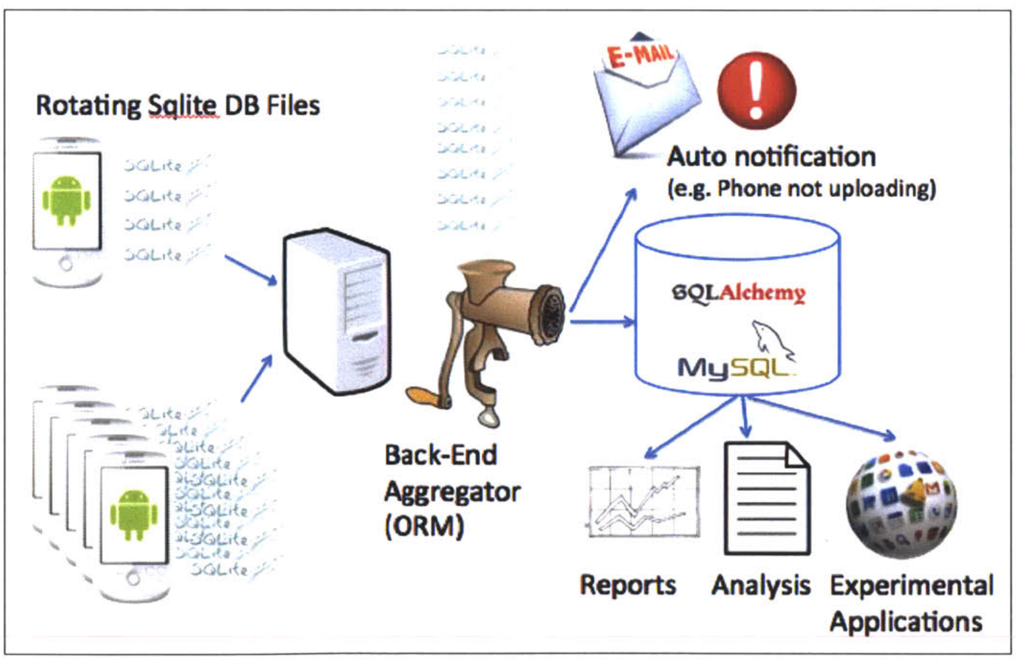

Figure 3-5: Overview of back-end data-flow.

3.5 Funf: Open Source Release Version

The Funf open source release version was developed starting April 2011. It included a re-design and modularization of the initial version, while reusing relevant code and applying some of the practicals lessons learned during the Friends and Family field deployment. A "sneak-peak" of the framework and a test-app was released in May 2011. A full release of the framework, including source code, was done October 5th, 2011. Figure 3-6 gives a high-level illustration of the Funf framework and data flow.

The Funf home page is currently available at: http://funf.media.mit.edu ([Aharony et al., b]), and the codebase and developer documentation is available at http://code.google.com/p/funf-open-sensing-framework/ ([Aharony et al., a]). The website, codebase, and documentation are presently actively maintained and kept up-to-date, and so we direct the reader to the websites for detailed information about using and developing with Funf.

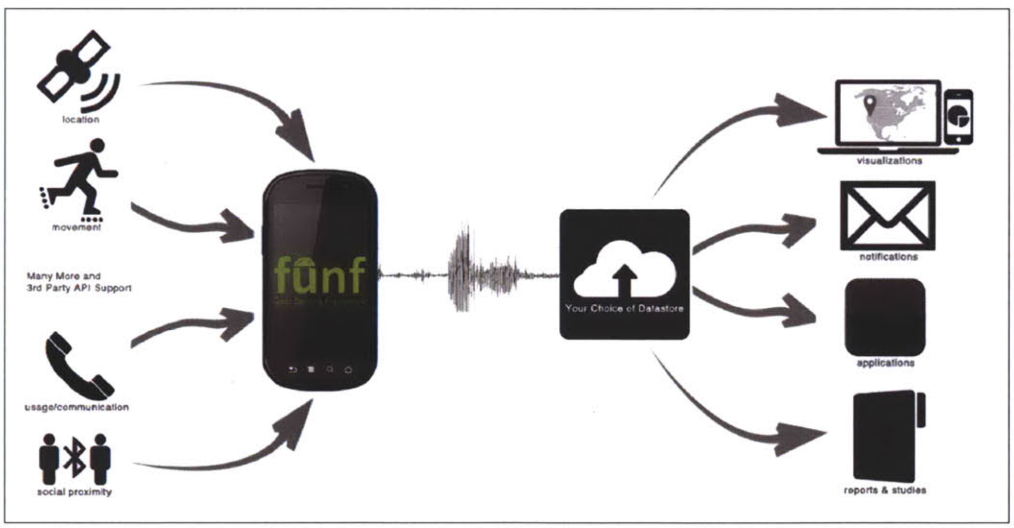

Figure 3-6: High Level overview of the Funf Open Sensing Framework

3.6 Funf Journal Application / Self Quantifiers, etc.

Funf Journal is an Android application built using the Funf framework, which makes use of many of its built-in features. It can be used by researchers, self-trackers, or anyone interested

Figure 3-7: Overview of the Funf Journal App data-flow.

in collecting and exploring information related to the mobile device, its environment, and its user's behavior. Funf Journal is probably the easiest way to get started with the Funf framework and its functionality, and can be used as an automated journal of a users' life. Figure 3-7 depicts a high level workflow, where the users can select which probes they want to collect data from and how often, and the Funf Journal app manages the data collection. When user desires (or with automatic sync turned on), the data will be sent to the user's selected destination, where it can be converted to a single SQLite database or comma separated value file, which in turn can be visualized or used as the user desires.

Funf Journal allows the user or researcher to configure data collection parameters for over 30 different built-in data probes, including all phone sensors, as well as additional data types and high-level probes that generate inferences and new data based on the output of sensor data. New probes are continuously added to the system, and as the core functionality is enhanced, new versions of the Funf Journal apps are enhanced as well. The app supports importing and exporting of probing configurations, as well as remote configuration (with user permission), allowing users and researchers to create "snapshots" of configuration settings that are aimed at different experiments or data collection goals.

Based on the configuration, the probe data is automatically collected the application, as the user carries the phone with them and uses it in every day life. The application stores the data on the phone in an encrypted format. The encryption key is generated based on a user selected password. This app is a private app for the user, and not part of an ongoing MIT data collection initiative - All of the collected data remains in the users' domain - their phone, and wherever they might choose to export it data to.

The data can be extracted from the phone in one of several ways - by manually exporting it via email or any other Android service that supports file transfer, by manually copying it out of the device's memory card, or by setting up a server and configuring Funf to automatically upload the data to it - and if Internet access is not available, the app will accumulate data on the device's memory card and wait until it is back online. When a server link is set up, the data collection configuration could also be performed remotely - the app will check the server for new configuration instructions for download.

Once the data is extracted from the phone, we provide a set of desktop utilities that allow decryption of the data, demonstrate some examples of visualizing and looking into the collected data, and also produce a clean database (in SQLite format) that you can use for further analyzing and exploring your data as well as for importing data into other applications or services.

Figure 3-8 depicts screen-shots the Funf Journal app.

3.7 Vision

Our longer term vision for Funf is to see Funf, or a related community driven effort, enable an ecosystem that is based based on shared standards for data collection, and have that framework become a de-facto open standard for doing mobile data collection.

One of the key reasons for the success of the Internet is considered its "hourglass" protocol architecture [Deering, 2001]. As Figure 3-9 shows, at its narrow "waist" is the Internet Protocol, or IP, which is common to all Internet services. On one hand, a single protocol maximizes interoperability of services. On the other hand, making the common mandatory layer narrow, meaning only in a certain point in the stack, affording for a lot of flexibility in the other layers. This allows, on the bottom of the hourglass, supporting a great deal of underlying network technologies - fiber optics, wireless radios, copper cables, and so on. At the top of the hourglass, there is also very great flexibility on the types of service that can be run on top of the Internet, as is well known. The common IP layer allows any of these services to run on any of the underlying networks.

In our vision, we see an analog hour-glass architecture for Funf. As depicted in Figure 3-10, the narrow waist contains the probe architecture as well as data protocols and formats. At the bottom of the stack are the different data sources that would feed into Funf as probes, and in particular we want to promote the more advanced types of probes that would add inference and sense-making to the collected data. This could be a place for machine learning researchers, algorithm developers, and systems researchers, to add new functionality and improve the existing. On the other end, we hope to enable a wide range of applications, experiments, and other uses of the data, and allow social scientists and non-programmer researchers, practitioners, and users, to enjoy the broad range of data collection possibilities.

Developers and researchers in lower layers would enjoy access to users who would test and collect data with their algorithms, giving them feedback as well as a short path to deploy their models in real-world applications. They would be able to focus on their area of interest and expertise rather than develop the entire infrastructure that Funf already implements. They would also be able to use Funf for easy collection of test and validation data as they develop their models. In a distant future, we can imagine different challenges and competitions in the classroom, encouraging innovation for new models and improving existing ones (e.g. competition to do maximal data collection with minimal battery impact). On the other end of the hourglass, social researchers would be able to easily conduct data collection experiments, and gain increasingly better functions and features. This could be a place for companies to innovate on either side of the hour-glass, essentially creating a mobile data ecosystem and marketplace.

3.8 Chapter Conclusion

The Funf framwework's codebase was officially released launched on October 5th, together with the Funf Journal application. Since then the project has received news coverage, and interest among researchers and developers. By December 2011, there were over 550 downloads of the Android application, and over 70 downloads of the developer libary by researchers and developers.

We have been contacted by individuals coming from a broad range of backgrounds: researchers, students, experienced "quantified selfers", and casual end-users who are interested in collecting and exploring their own data. So far all responses have been very positive all around.

Funf has received two Google Research awards, one in December 2010 for $75K, and a second one in December 2011 for and additional $78K to support its further development and expansion.

Funf currently has a growing community of users and developers, and we hope that it will reach a state of sustainability over time. We hope to foster a community around these ideas, where new development efforts would go towards extending a common platform rather than creating redundant functionality.

As mentioned earlier, in the following chapter we dive into one instance of using the Funf system in the real world, where Funf was used as the main component of a Social fMRI longitudinal experiment, the Friends and Family living laboratory study.

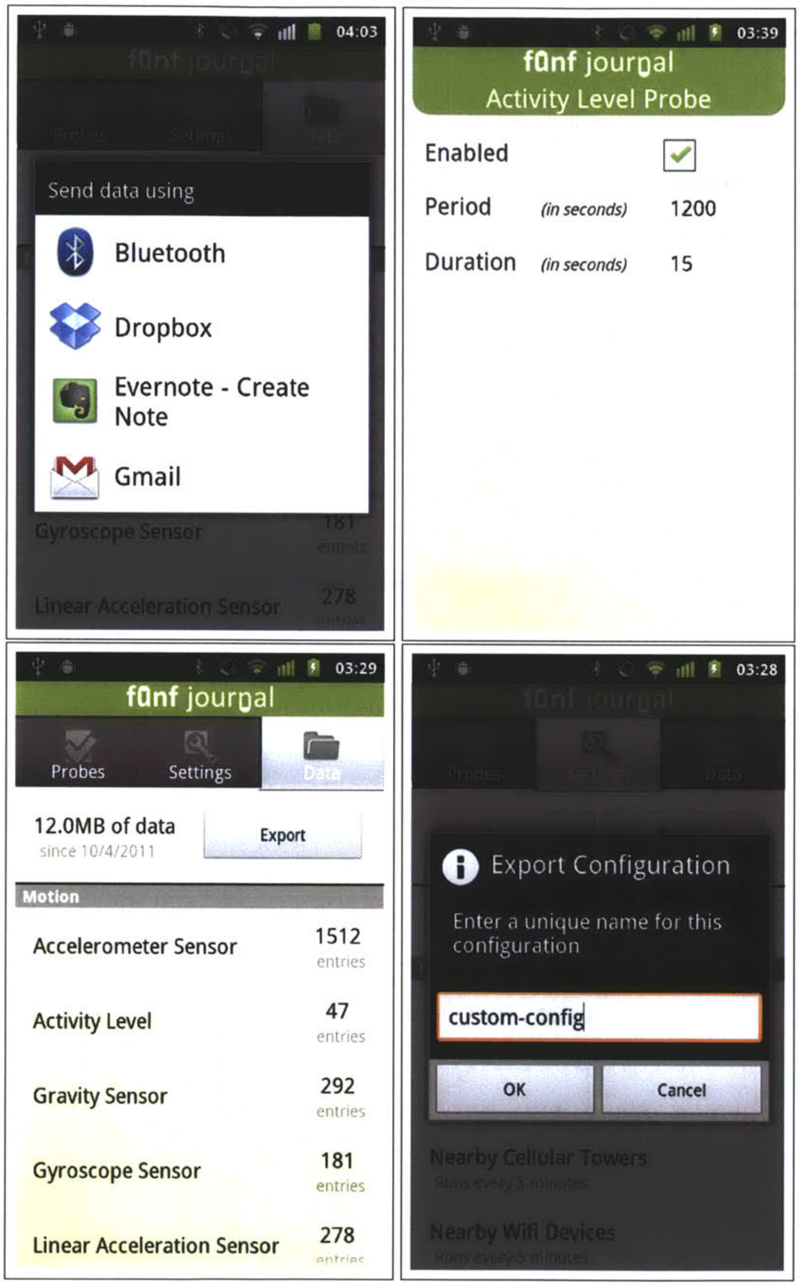

Figure 3-8: Funf Journal App screen-shots. Top-left: Manual data export screen. Top-right: Configuration settings for a specific probe (accelerometer activity levels), Bottom-left: Status of the currently collected data on the phone, with per-probe statistics, Top-right: Configuration export screen, allowing existing operational settings to be exported and imported.

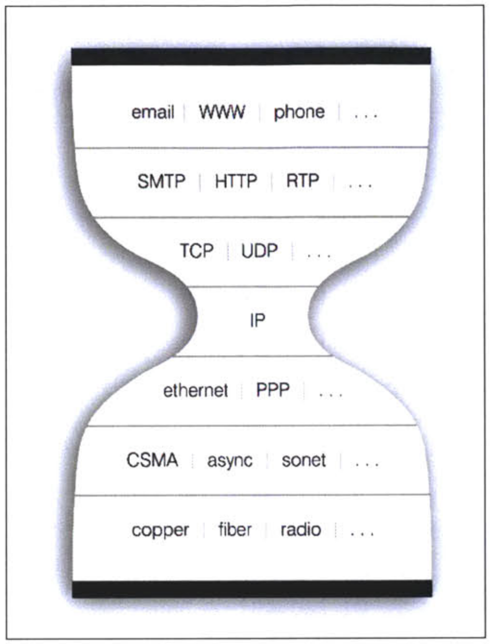

Figure 3-9: Internet "hourglass" model [Deering, 2001]

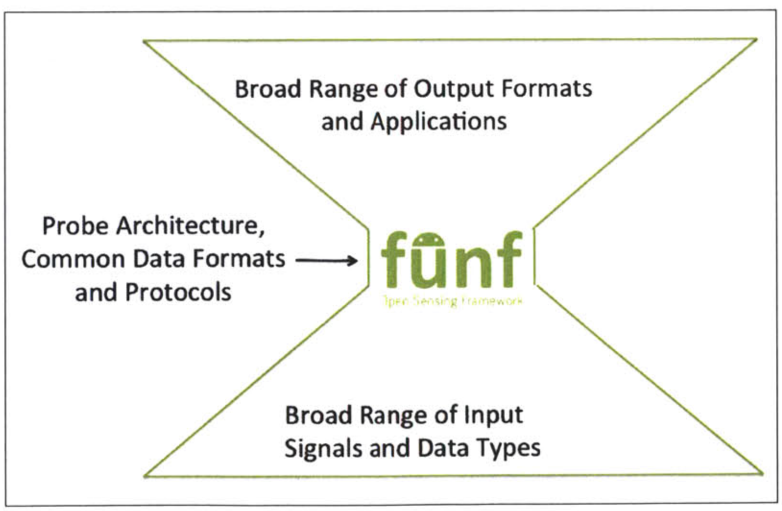

Figure 3-10: Funf "hourglass" model approach

— *"Studying a group in its natural setting has been universally characterized by the great diversity of research methods it employs."*

Leon Festinger

4

Friends and Family: A Social fMRI Experiment

The "Friends and Family" living laboratory study is the first Social fMRI experiment to be conducted. To the best of our knowledge, it is the most comprehensive mobile phone experiment performed to date in academia. The dataset generated from the study is probably the largest and richest dataset ever collected on a residential community. The study was conducted over a period of 15 months between March 2010 and June 2011. The study was conducted with members of a young-family residential living community adjacent to MIT. All members of the community are couples, and at least one of the members is affiliated with the university. Over the course of the study, 140 individuals were part of the subject pool. Due to the significant length of the study, some subjects had graduated while others joined the residence community, and this was a planned part of the study. At its peak, the study had close to 130 active participants over a period of 7 months during the 2010-2011 academic year.

The study included a broad range of methodological tools, including observations, automated sensing and data collection, regular questioners, interviews, sociometric surveys, and targeted intervention experiments. The main data collection and intervention tool was a mobile-phone-centric social and behavioral sensing system deployed to the study participants. The dataset includes continuous collection of over 25 phone-based signals - including location, accelerometery, Bluetooth-based device proximity, communication activities, installed applications, currently running applications, multimedia and file system information, and additional data generated by our experimental applications. each of the signals was sampled at a resolution appropriate to its characteristics. In addition, we collected financial information through receipts and credit card statements, performed logging of Facebook-based 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. This chapter reviews the Friends and Family experimental design, methodology, and procedure.

4.1 The "Friends and Family" Community

The longitudinal field study was conducted at a residential community for couples and families located near the MIT campus. All members of the community are couples, and at least one person in the couple is affiliated with the university (primarily as a graduate student, although some might be undergraduate students, postdoc affiliates, or other visiting researchers). 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.

4.1.1 Layout and Characteristics of the Community Environment

The community residence is composed of a single high-rise building and several low-rise buildings. The low-rise buildings have larger apartments which are only allocated to residents with children. The high-rise has smaller apartments, which house mainly couples without children, although high-rise apartments are also allocated to residents who have a single child. As depicted by the illustrative Figure 4-1, all buildings encircle a courtyard in which there is a large children's playground area as well as a barbecue cooking area and picnic tables. These serve as a socialization and social rendezvous place nearly year-round. The residence also includes community facilities, such as as two lounges, a laundry room, small exercise room, and and indoor children's playroom. All of these offer opportunities for social interaction among the neighbors. In addition, community volunteers organize various social events - from regular weekly "coffee-hours", to special larger scale events that are organized approximately once a month and include things like holiday celebrations, community barbecues, cooking competitions and so on. Other informal activities might be resident organized exercise groups or parent-children activities. While some residents are avid attendees of numerous social activities such as those described, other might irregularly attend just a few of them, or not at all.

4.1.2 Population Characteristics

The population itself is quite diverse, with a significant proportion of non-US nationals, and a mix of cultures, religions, occupations, and areas of study/research (for the students). Approximately half of the residents have children, usually of a young age. Section 5.2 describes more detailed demographics on the subset of residents that participated in the study.

4.1.3 Subject Pool Discussion

This study involved a qualitatively different subject population when compared to previous ubiquitous computing observatory studies, such as the populations of colleagues and co-workers in Reality Mining [Eagle and Pentland, 2006], and undergraduate students in [Madan

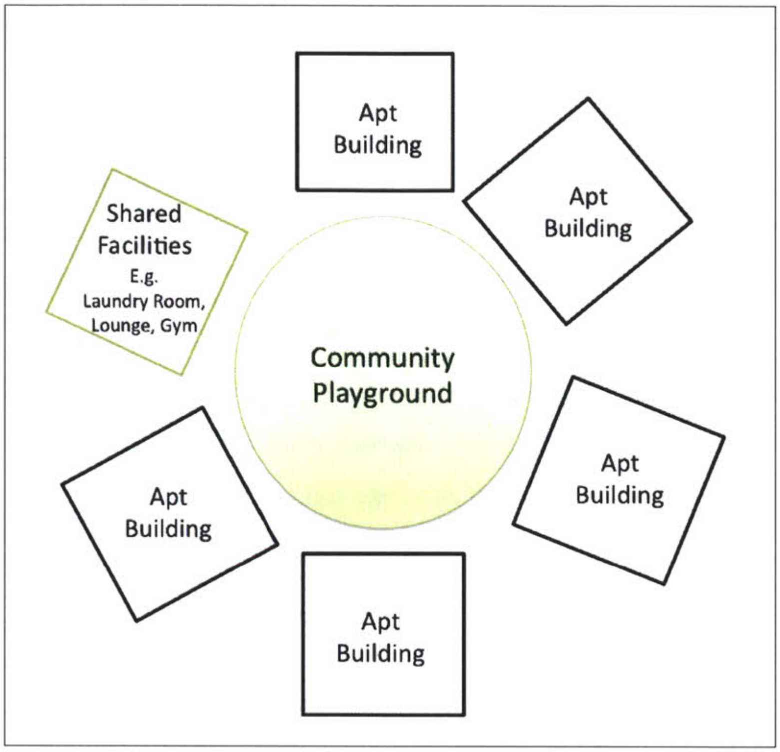

Figure 4-1: An illustrative depiction of the Friends and Family community's physical layout

et al., 2010a]. The Reality Mining study was conducted with a cohort of colleagues and co-workers. This means that the characteristics of the community are weighted towards these types of relationships, for example, life partners or childhood friends. This would also impact the collected social interaction data, in particular the face-to-face interactions, which are more centered around the work-day rather than weekends and after-hours, and any temporal ties that extend into the weekend and after-hours are actually very significant [Eagle et al., 2009]. This population and its characteristics fit the goals of the Reality Mining study, and enabled investigation of things like team dynamics and job satisfaction, as well as other work-place and organizational related questions. However, it is somewhat limited to those domains, especially in the context of face-to-face interactions that part of a person's family and private life.

The undergraduate community in the Social Evolution study was perfect for the study's goals of investigating social diffusion in a tightly knit community, as the chosen undergraduate dorm was very small and cohesive, and allowed investigators to capture a significant percentage of the dorm life and of individuals' social ties in the university [Madan, 2010]. The graduate families population of the Friends and Family community is quite different than the undergraduate population: They are, by large, autonomous family units, responsible for the economic, health-related, and social well-being a of themselves and their children. The residence is structured not as a dorm but more like a residential condominium. One difference that might be between this community and a typical high-rise or suburban community is that although quite heterogeneous with respect to the context of a university, it is not as diverse as one might found at a randomly selected urban community. Nearly all residents are in a similar situation - taking a break from their routine lives and careers in order to go to graduate school (most married graduate students are in a phase in life that is older than the average single graduate students). They also share the common affiliation with the university, which certainly adds to the interpersonal trust and affinity among participants. This is a very positive situation from the perspective of our research goals, as it allows us access to a very vibrant community with active community life, but it is largely self-organized rather than externally imposed or scaffolded (as in many standard university dorms).

The internal diversity within the community is a fruitful ground for our field investigation. Residents could be grouped in many different ways based on their demographic affiliation - cultural, ethnic, occupation, field of study, age, physical apartment proximity, and so on. This enables us to investigate questions related to social groups and social contexts in the real world, like whether these different grouping features significantly influence the evolution of social ties, the spread of information within the community, or even whether they can be automatically inferred from the collected sensor data about communication and behavior patterns?

The specific community in which the study was conducted is very symbolic: It is in this same community, more than 60 years ago, that Leon Festinger and his colleagues conducted their groundbreaking field study of social pressures in informal groups [Festinger et al., 1950]. The residence buildings at the time of Festinger's study were several hundred feet away from the current buildings, and they were set up as temporary housing before the permanent housing buildings were constructed. The population was more homogeneous than the Friends and

Family population, since at the time the main occupants were World War II veterans and their families, returning to school after the war. However many characteristics of the community existed even then.

4.2 Experiment Design

The Friends and Family study embodies the Social fMRI ideas. We set out to investigate a community *in-situ*, using mobile phones and many other investigation and data collection tools in order to create a uniquely rich and highly dimensional image of the community, over the period of 15 months. These combine to create a virtual imaging chamber. Baseline observations, mainly collected via electronic means but also augmented by additional observations, help us gain an understanding of the natural state of the community. However this collection is not purely passive or stationary, but far from it. The field study design includes a key aspect of interactivity. This includes interviews and participant self-reporting, which helped us get an insights into both ground-truth measures, as well as self-perceived view of the world (as it has been shown that self-reports do not necessarily establish factual ground-truth [Bernard et al., 1984, Marsden, 1990]). The interactivity component also includes a series of carefully designed stimuli, or interventions, that allow us to experimentally evaluate social mechanisms and at least partially control some of the experimental variables. The study was driven by specific hypotheses and research questions, combined with exploratory investigation, all under the research goals described in section 2.3.

The Friends and Family study was designed in two main phases. The first phase, or pilot phase, was launched in March 2010 and continued for 6 months until September. The study's second phase, or main phase, ran for 9 months throughout the 2010-2011 academic year, between September 2010 and June 2011. Figure 4-2 gives a high-level overview of the study timeline.

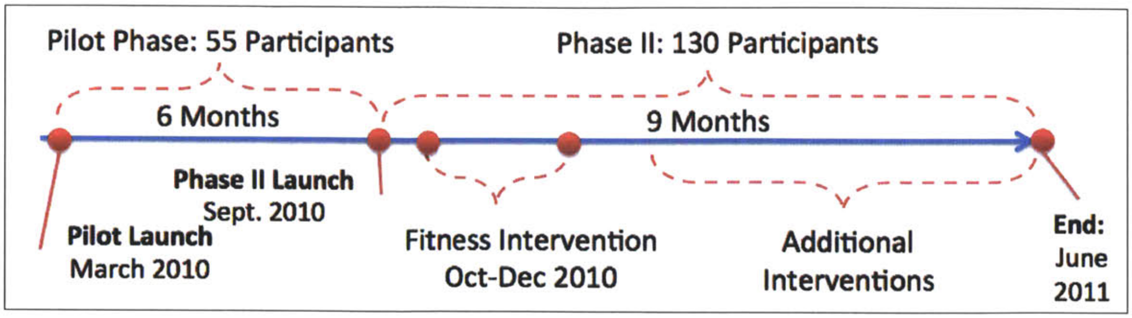

Figure 4-2: High level timeline for the Friends and Family study.

4.3 Sub Experiments and Interventions

In order to be able to dive into specific investigations as well as develop different aspects of the methodology, several interventions and sub-experiments were conducted with the study population. All sub-experiments were done on an opt-in basis. The main interventions performed during the study include:

4.3.1 Funfit: Social mechanism investigation in the context of wellness and social incentives

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. Subjects were divided into three experimental conditions, allowing us to isolate different incentive mechanisms related to monetary reward, the value of social information, and social pressure/influence. This is the most extensive intervention performed within the study, and is described in detail in Chapter 6

4.3.2 FunDashboard: Visualizing personal data

As part of the study, we wanted to see the effect of showing participants their own data, and investigating whether people actually use this, whether it made any observable behavioral change in their lives or in the adherence to the study requirements.

4.3.3 AppStimulus: Measuring choices and network effects

Mobile applications installed on the phone offer a very useful window into a type of decision making that could be influenced by social ties. We are able to capture the full data on app installations and un-installations on the study's mobile phones, including time of installations. After one year of passively observing app installations, in March 2011 we introduced an active stimulus to the system - each participant received budget of up to $30 they had to spend solely on mobile apps, with a limited time frame. We wanted to see how this affects user behavior, if at all, and whether we can trace any social influence and network effects in the data. /tododo we keep this, pointer to some basic analysis?

4.3.4 FunFatigue: Focused sleep study

While the study had daily surveys as well as other measured sensor signals that might help us infer and approximate sleep times, a detailed sleep study with accurate bio-signal measurements could provide deeper insights into the sleep patterns of individuals, as well as help us validate our sensor approximations. This sub-experiment, led by Sai Moturu, used a consumer device by Zeo Inc., which allows us to capture sleep data from users in an unobtrusive manner, presenting us with great opportunities to study sleep-related research questions in a natural environment - something not possible previously. The study involved 10 couples, each being measured for two weeks with the device and filling out additional a daily sleep-related questioner.

4.4 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 inter-relation could be better understood.

Had we designed a separate "traditional" experiment for each of the research goals and questions and listed the required signals and variables to answer each, there would be a lot of overlap in the required data collection and methodology. For example, many of the specific questions related to real world social networks would call for the collection of data related subject's self-perceived social network. Multiples ones would benefit from the information on physical proximity and face-to-face interactions between subjects, and so on. Many of the signals collected could be used for more than one investigation, and probably for new analysis which we have not even thought about. The immense effort and costs associated with a longitudinal study such as the Friends and Family study add to our motivation to cast a wide net and attempt to capture as many signals and data points as we can during the duration of study, in particular ones that can be automatically collected and are thus "cheap" as far as effort required by study participants or investigators.

We applied a user centric, bottom up approach, utilizing the components listed below. Figure 4-3 illustrates the key sources of data that was continuously collected throughout the study.

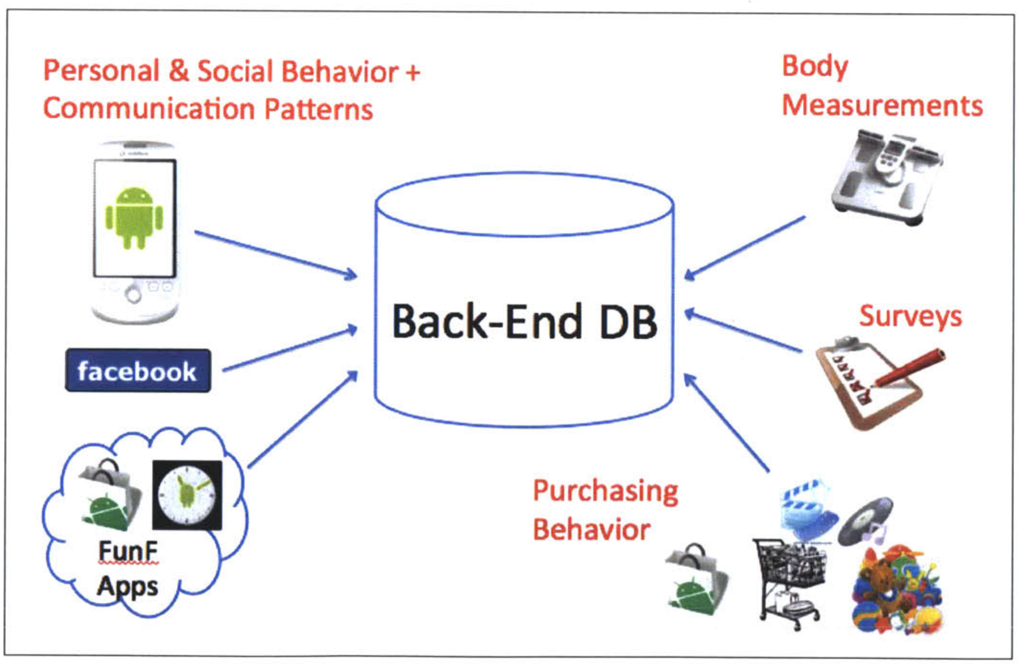

Figure 4-3: Overview of the main data collection components for the Friends and Family Social fMRI study.

4.4.1 Mobile Phone Sensing Platform (Funf Software)

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 sensing system is based on the first generation of the Funf sensing framework, described in Chapter 3, section 3.4. 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.

4.4.1.1 Sensing Configuration

Figure 4.1 describes the probes active for most of the study's duration, with their typical collection settings. Some of the settings were adjusted during the study, due to various reasons. Some of the changes were related to new probes being added, or existing ones enhanced, or debugged, as the study avanced. The majority of these changes happened during the pilot phase of the study, and the configuration has been mostly stable for the second and main phase of the study.

4.4.2 Surveys

Subjects complete surveys at regular intervals, combining web-based and on-phone surveys. Monthly surveys were administered roughly once a month, and include questions about self perception of relationships, group affiliation, sociL interactions, and also standard psychological scales like the Big-Five personality test [John and Srivastava, 1999]. Daily surveys include questions like mood, sleep, and other activity logging. For some periods of the study, weekly surveys were given which asked some activity questions in a weekly perspective (e.g. asking about how many times they ate red meat during the past week).

Signal	Interval	Opportunistic?	Notes
Wifi Scan	5m	Y	Access-points in range
Bluetooth Scan	5m	N	Bluetooth devices in range
Cell Tower Scan	5m	Y	Current cell tower ID
GPS Location	N/A	N**	2-states: 15m outdoors/30m indoors
Accelerometer	2m	N	Detailed in Section 6.2.2
Installed Aps	12h	N	Currently installed applications
Running Apps	30s	N	Currently running applications
Call Log	12h	N	Event log and statistics (hashed)
SMS Log	12h	N	Event log and statistics (hashed)
Contact List	12h	N	Periodic for tracking changes
Power State*	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*	24h	N	Set of scans for media content (video/audio/images)
File scans*	7d	N	Set of file/directory scans.
Browser Scans*	12h	N	History and bookmark scans (hashed)
Alarm Clock	N/A	Y	App that logs alarm clock usage
Phone Info*	3h	N	Set of signals on phone state (e.g. device ID, os version, timezone, etc.)
Network Info*	3h	N	Set of signals on mobile network state (e.g. current operator, data mode, etc.)
Funf Info	3h	Y	Info on installed Funf sofware version
Probe Config	3h	Y	Snapshot of probes' current setup

* Set of multiple signals ; ** Opportunistic in current software version

Table 4.1: 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.

4.4.3 Purchasing Behavior

The study included a data collection on the the diffusion of purchasing behavior, which could be a good measure of how people change behaviors and learn from one another. Some purchasing categories are could be social by nature, or influenced by social interactions. One example is peoples choice of restaurants, movies, or live shows. Another example is child-related purchases, as parents learn of new and better products through their friends and neighbors. In order to enable the study of purchasing behavior and how people consult, recommend, and influence each other, participants were asked to submit traces of purchase transactions.

During the pilot phase of the study, we asked participants to optionally submit purchase receipt, as they see fit. They were allowed to submit any purchase receipt, if they wanted to. They could also black-out any transactions or items that they were not interested in sharing with the investigators. In particular, the were asked to submit receipts related to discressionary spending, such as spending on movies, music shows, restaurants, or food delivery, which we hypothesized to have a higher chance to for some social component related to them. As compensation, participants would receive a reimbursement of 10% of the receipt value, up to $50 per household, per month. The idea of requesting itemized receipts was so that we could have access to the specific purchases - brands, type of food (e.g. healthy/unhealthy), which movies people saw, and so on. However, this turned out to be a heavy burden on both participants and investigators. While few participants submitted receipts diligently, others submitted very sporadically or not at all, making the data very spotty for investigated at community trends. We also realized how hard and time consuming it would be to scan, digitize, and encode the individual receipts, and did not have the resources to do it at scale.

Based on this gained understanding, for the main phase of the study we decided to change the request from individual receipts to monthly credit card and/or bank statements. We also made this component mandatory for all incoming participants. Nevertheless, they still had the choice of removing any transaction that they did not want to share with the investigators, essentially allowing them to submit completely redacted and empty statements.

4.4.4 Facebook Data Collection Application

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

4.4.5 Body Measurements

A subset of the participants who elected to take part in the fitness intervention (which is detailed in Chapter 6) were given body measurement scales (Omron Body Composition Monitor and Scale HBF-510) that can measure body weight, body-fat percentage, visceral fat, body mass index (BMI), nd skeletal muscle percentage. The participants were required to submit weekly measurements during the fitness intervention, and were also asked to optionally continue submitting the measurements for the remainder of the study.

Additional Data Sources There are several other auxiliary sources of data might provide useful information during the analysis of the dataset. These include, for example:

- Information about external events: These can be gathered from the residence social calendar, the MIT events calendar, MIT academic calendar, weather reports for the area, etc.

- Logs of web access to the participant information: During the fitness intervention as well as the data dashboard intervention, each participant had their own unique data URL, and our server logged the occurrence of page-views and any related information (e.g. whether it was accessed via the mobile phone or the PC)

- Qualitative Interviews with participants were conducted during the pilot phase in order to helpe with the design of future interventions and questioners.

4.5 Protocol

4.5.1 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. A second key consideration was how to make the study as least obtrusive as possible to participants' daily lives. The protocol submitted to the IRB included details of all the types of data collected and the potential discomforts to participants. It was also presented to the participants in multiple occasions and formats - including during the recruitment process, the subject participation consent forms, on the participant's information website, and in a short booklet given to participants at the beginning of the main phase of the study.

Two boundaries of protection were initially set up: The first is between the investigators and the outside world. As in similar experiments, no subject identifiers would be shared or published, and all research publications will include only aggregate and de-identified information and results. The second boundary is between the participants and investigators. All all human-readable text was captured as hash encoded identifiers, and never saved in clear-text. 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. Table 4.2 illustrates the idea of one-way hashing, as was also explained to study participants. Collected data is physically secured and de-identified before being used for aggregate analysis.

Following are examples of some of the measures taken to protect subject privacy, as included in the IRB proposal:

- Data will be linked only to coded identifiers for the participants, and not to any personal identifiers at any point that it is captured or transmitted.
- Content of SMS messages and all other textual information will be encrypted and not saved in human-readable form, and will not be published in any way that could be personally identifiable.

Example	Plain Text	Hashed (e.g. MD5)
Phone Number	"617-555-5555"	e7019cecff8ba9a08180c59a7a74365da3057873
Phone Number	"617-555-5551"	f171639a6cf7748fc3f8fa85641791f6b5967e00
Name	"John Smith"	e7019cecff8ba9a08180c59a7a74365da3057873
Text Message	"yes"	fb360f9c09ac8c5edb2f18be5de4e80ea4c430d0
Text Message	"no"	fd1286353570c5703799ba76999323b7c7447b06
Text Message	"Sorry Im in a meeting"	b11ad70951165 15dc8abf12cf29d2b96305e2725, 46498c44efba3ae0f9c033d521f8c64769e17ecf, af10ef20dd9060bbeead0afbc55381a66af442ef, 86f7e437faa5a7fce15d1ddcb9eaeaea377667b8, bd7580126b941404db395e16ff95491f65006476

Table 4.2: Example of one-way hashing of numbers and clear-text.

- Content of phone calls is not captured. Only information about the phone call event is recorded in an encoded manner.
- Image and media files are not captured or copied outside of the user's device. Only usage and feature statistics are logged (number of files, size).
- Incoming and outgoing phone numbers will only be used in encoded form and for statistical purposes only. Phone numbers will not be saved in human-readable form and will not be published in any way that could be personally identifiable.
- Raw audio is not stored. Only features like energy level will be extracted from the audio stream and stored. *Note: Even though this was approved in the protocol, it was not eventually collected during the study*
- For credit card statements or purchase receipts submitted for partial reimbursement, participants will have a choice whether or not to submit any statement or receipt, or to cross any item out before submitting the receipt.
- The subjects have the right to withdraw from the study at any given time and have their data destroyed.

A second important consideration was to be as unobtrusive as possible to the subjects' life routines. As compensation, 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 interventions or sub-experiments on top of the main study components is completely optional, and interventions were carefully designed with the interests of the participants in mind.

4.5.2 Participant Expectations and Compensation

Participants were informed of the researchers' requirements and expectations from them as well as the compensation they would receive for fulfilling the study's components. For the main phase of the study, an information booklet was prepared for the participants, which included an overview of the study, investigators' expectations of the participants, what participants should expect from the investigators, compensation information, and privacy considerations. Following are excerpts from the participant handbook regarding the requirements and expectations for the study participants.

4.5.2.1 Expectations from Participants

We have the following expectations of participants in the FunF study:

(a) You must use the study phone as your primary mobile phone for the duration of the study.

(b) You must keep your study phone ON whenever possible so that it can collect data.

(c) You should carry your phone with you whenever possible.

(d) You should charge your phone nightly so you will not have to turn it off to conserve battery.

(e) You should report problems with your phone as soon as possible!! You can report these by email to friendsnfamily@media.mit.edu.

(f) You should keep your phones Bluetooth, Wifi, and GPS turned ON. Because of peculiarities of the Android OS, you may occasionally get messages related to these services, but always select to keep or turn it on.

These services will not be transmitting constantly, but rather we keep them on in order to do the periodic scanning.

(g) You must complete the weekly and monthly surveys.

(h) You are expected to submit your credit card statements as outlined above.

(i) Google Voice and similar services: Google Voice diverts phone communication information (calls, text messages, etc.) from the phone to the Internet. Therefore, for users of the Google Voice service, the information we collect from the phones will be incomplete. Since we do not wish to restrict your activities, you are welcome to use services like Google Voice. However, we are currently working on an application that will allow us to access Google Voice information, and we will ask you to allow our app to log this information once it is ready.

You will receive email notification of any changes or additions to the expectations above. If we feel you are not meeting the above expectations, we reserve the right to remove you from the study.

What Participants Should Expect

We sincerely appreciate your participation in this study and understand that there is some inconvenience involved with it, which we want to minimize.

It is almost certain that there will be some bugs with the phone software, and the most reliable and efficient way for us to deal with these is for you to report bugs to us as soon as you notice them.

We are also interested in any feedback you have to give us about any element of the study. Accordingly, we hold weekly office hours in the lounge, when you can come by to talk about whatevers on your mind, study-related or not! You can also email us anytime at friendsnfamily@media.mit.edu.

We will try to minimize email broadcasts to participants. Once the study is up and running for the semester, we will try to keep emails to one per week with relevant reminders and updates. In return, we ask that you read this weekly email carefully.

4.5.3 Participant Compensation

The general approach to the study compensation was along these lines: Participants who remained in the study for a minimum duration (six month for subjects who joined during the pilot and nine months for participants who joined in the main phase of the study), while fulfilling the study's requirements, would be able to keep their study phone (worth at time of purchase between $300-$500). In addition, increase compliance over the long experiment duration (over 1 year), we wanted to compensate for anything that took participants out of the ordinary, in particular for the burden of filling out the many surveys. Finally, many of the optional components had additional incentives.

During the main phase of the study, the monthly compensation breakdown was as follows: Each participant would receive $20 for completing all weekly and daily surveys. If they had not filled out all of them, they would only get pro-rated amounts (with a minimum of filling out at least 50% of the daily surveys in order to be compensated). They would also receive $10 for completing the monthly survey. An additional $10 per household was given for credit card and bank statements submitted, leading to a total of $70 per family for compliance with the core components of the study.

4.5.4 Recruitment

4.5.4.1 Recruitment Strategy

There were several requirements for the study participant pool. First, as in similar face-to-face studies, like [Olguín et al., 2009], [Eagle and Pentland, 2006], and [Madan, 2010], there was a requirement for geographical density so that we would be able to sense relevant face-to-face interactions that are part of the subject's lives. Since one of the key study goals revolves around understanding real life groups and affiliations and their effect on behavior and choice, we wanted to make sure to sample different types of groups, in a way that would capture multiple groups in the subject's lives, as well as enough member of certain group types within the community so that we would be able to gain meaningful insights on the group. This might include, for example, a significant portions of residents of a certain floor, a religious group, countries of origin, and so on. Another condition for the study was sufficient level of English, so that the subjects can correctly answer the surveys and interact with the research team as needed. All students and affiliates of the university posses the required level of English, however not all student partners do. Another condition was that the residents expect to remain in the residence for a minimum time during the study's duration (6 month for the pilot phase, and 9 months for the second phase). There was also preference to recruit both members of a couple, although in some cases only one household member was recruited - for example in a case where the couple was part of a key group we wanted to "cover", but one of the couple members did not want to participate.

4.5.4.2 Recruitment Procedure

Recruitment was done in two phases - first for the pilot phase starting March 2010, where the desire was to recruit around 50 subjects. For the second phase of the study, the goal was to increase the number of participants to around 130, also considering that some of the pilot participants would have graduated and left the community. Recruitment began about 1-2 months before each phase started. The study was advertised in the following ways: Emails to the community's public messaging mailing list which many residents subscribe to, fliers on put on residents' doors, announcements made during one of the community's monthly "town-hall" meeting. In addition a special info-session was held with the investigators, to explain about the study and address any questions or potential concerns. In order to be considered for participation, interested residents were asked to fill a "pre-participation" survey that gathered demographic information as well as other questions related to the study requirements, such as group affiliations, level of English, information about how long they lived in the residence and planned to live there, and so on.

Out of the 400 residents in the community, more than 200 applied to participate, out of which

140 were selected for the two phases combined (56 in the first phase, and 84 in the second). As some of the first phase participants left the residence and the study, at its peak the study had 130 concurrent participants.

Part II

Experimental Results

—"*In theory, there is no difference between theory and practice. But in practice, there is.*"

Yogi Berra

5

Virtual Imaging Chamber: Investigating a Living Community

The Friends and Family study has given us unprecedented vantage point into the life of a vibrant community and its members. The great number of diverse signals collected per person and about the community as a whole allow us to interpolate a multi-dimensional "image" of the living community throughout the study's duration. In this chapter we review some of the insights already learned in several different investigation trajectories including decision making and influence, social economics, social psychology, health, and more.

This chapter starts with an overview of the Friends and Family dataset and basic statistics about the data and the different social network modalities that can be spanned for the same people from the collected interactions and reported data. We then present analysis and results from several components of the Friends and Family study. These are but initial forays into the study's comprehensive dataset, some of which were done as the study was ongoing, and served to inform the design of subsequent components and sub-experiments. Aside from the direct research questions that we can answer, these first investigations help us 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 5.4. In section 5.5 we continue with looking at the social fabric by investigating network effects on decisions and choice. We then discuss two other investigation trajectories currently ongoing - Looking into the connection between socializing patterns and wellness related topics, like sleep, stess, or mood, and finally, investigating connections between personality traits of individuals and the social networks that form around them. 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, among many others.

5.1 Friends and Family Dataset

In this section we give an overview of the accumulated Friends and Family study dataset. The phone-gathered data encompasses over *1,000,000 hours*, or *117 years* of continuous human behavior. For comparison, the original reality mining study had accumulated over 350,000 hours, or 40 years of behavior data [Eagle and Pentland, 2006]. However, a key difference to remember is the aspect of richness: About 5 signals were collected via the original Reality Mining study (Bluetooth devices in proximity, cell tower IDs, call logs, built-in application usage, and phone-status), whereas in the Friends and Family Study more than 25 signals were automatically collected per person, with some of them having multiple parameters and dimensions collected per signal.

5.1.1 Phone Data Statistics

The accumulated size of the database files uploaded from the study phone devices adds up to *over 60 Gigabytes*. The data is composed of over *30 million* individual probe scan events, where some events capture a multiple of data signals. These include:

- 14 million accelerometer scans, covering 53 years of 2-minute activity-level summaries
- 20 million Wifi scans, which in turn accumulated 243 million total scanned device records.
- 5 million Bluetooth proximity scans, which in turn accumulated 16 million total scanned device records.
- 200,000 phone calls.
- 100,000 text messages (SMS).
- 200K GPS coordinate-level location records.
- Over 16,000 submitted daily surveys that include information on things like daily mood, stress, and socialization levels.
- For the Facebook component:
  - Data on 99 opted-in users, including a total of 2300 friendship ties.
  - Statistics on 15.5K emails, 4637 comments, 2300 social events, 1237 likes, 480 social groups.

5.2 Getting to Know the Neighbors

In this section we review the demographics and profile related information, as reported by the participants. The subject pool was composed of 51% males and 49% females. As far as affiliation, 54% (70 participants) were students, 43% were partners of students, and 2% were employees or post-doctorate researchers (some of the couples were both students at the institute). Figure 5-1 depicts the distribution of the number of couples with hildren in the subject pool. Figure 5-2 depicts the subject age distributio. Figures 5-3 and 5-4 depict the reilgious and ethnic distributions respectively (the actual religion and ethnicity details are encoded). These statistics show the great diversity of the subject pool.

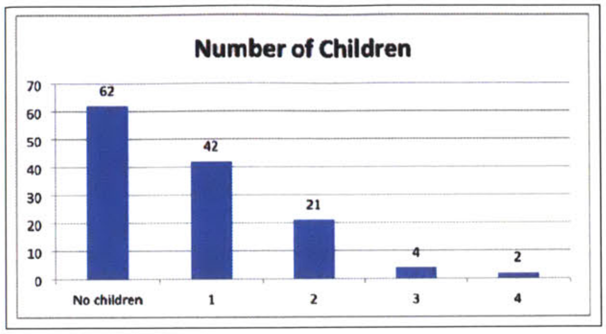

Figure 5-1: Distribution of children for study participants.

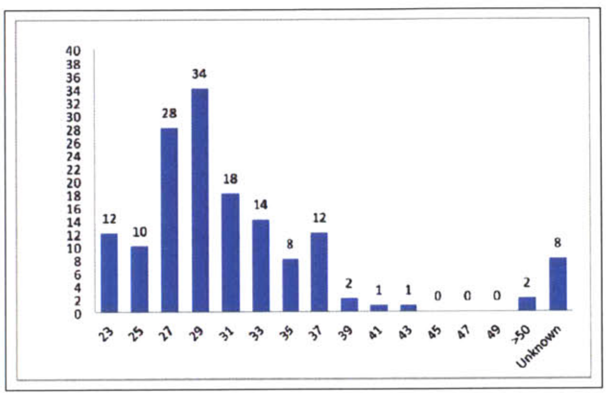

Figure 5-2: Distribution of age for study participants.

5.3 One Community. Many Networks

As discussed in earlier chapters, the Friends and Family dataset incldues data on multiple network modalities, and allows us to span different social network representations for the same individuals, based on different modes of interaction. Figure 5-5 shows an illustrative example of four networks with very different characteristics spanned for the same individuals - a face-to-face interaction network spanned by using collected Bluetooth scans, a phone call network , spanned by subject's call records, a Facebook interaction network spanned from our Facebook component logs, and finally, a self reported closeness network, spanned from

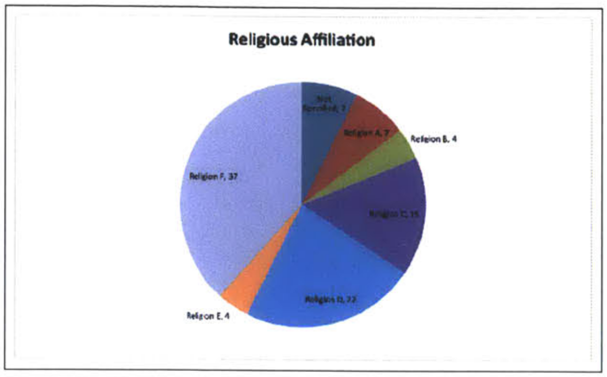

Figure 5-3: Religious affiliation for study participants (religion name encoded)

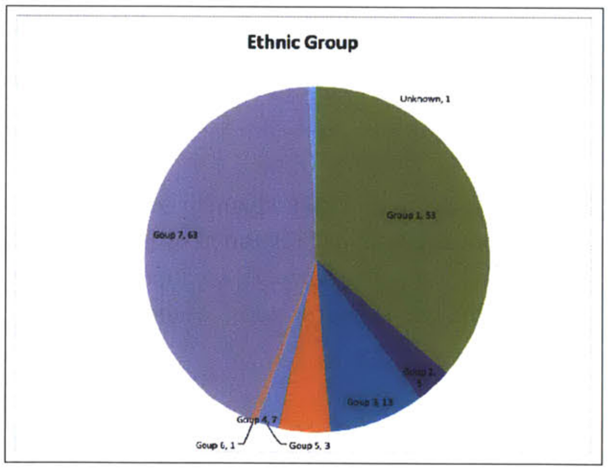

Figure 5-4: Ethnic affiliation for study participants (ethnicity name encoded)

manually entered individual surveys. Even though these are the same individuals, it is clearly seen that these networks have very different characteristics, and the nodes have different roles. A node central in one (e.g. node titled "fa10-01-54" is central in the call network, but not in the other ones). In the next sections and in the following chapters we show some initial investigations regarding differences between these networks and how one might affect the subject's lives more than another.

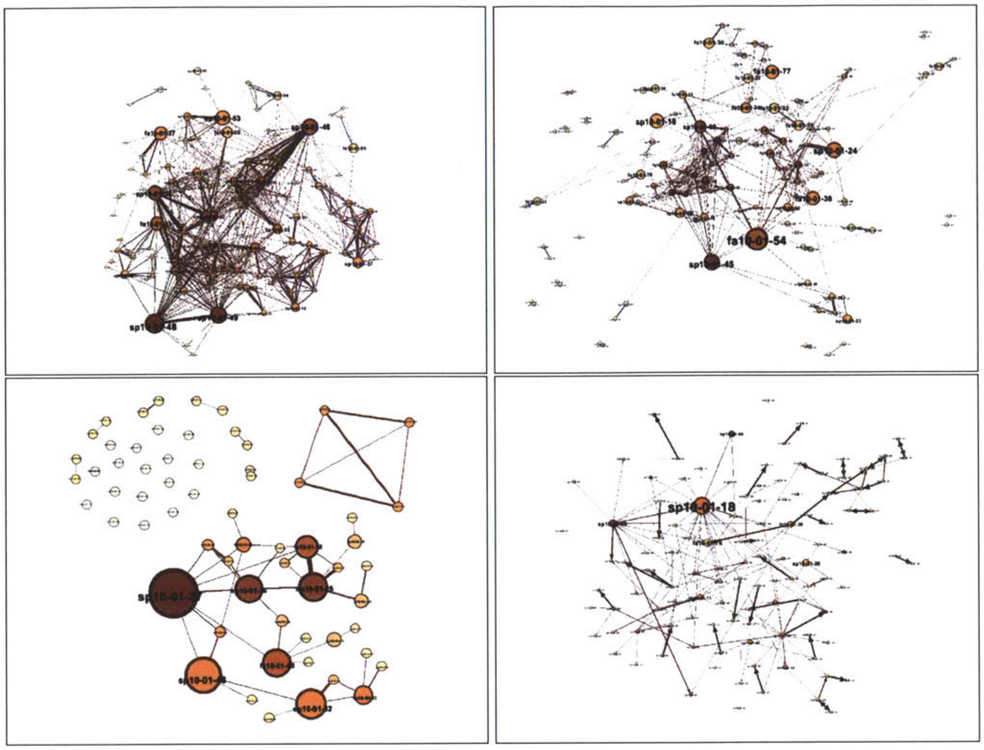

Figure 5-5: Examples of multiple network types spanned for the same set of individuals. From top left, clockwise: Physical proximity network (spanned by Bluetooth scans), phone call network (spanned by call logs), Facebook interaction network (spanned by interaction between study participants on Facebook), and finally, the self reported closeness network, spanned from manually entered individual surveys.

5.4 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 [Eagle et al., 2010]. 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 [Bruggeman, 2010, Page, 2008]. This approach follows a long line of classical social science literature: Granovetter's weak tie theory [Granovetter, 1973] and Burt's theory of structural holes [Burt, 1995], to name two.

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 [Eagle et al., 2010]) correlate only with their *current* income, as illustrated in Fig. 5-6 and Fig. 5-7.

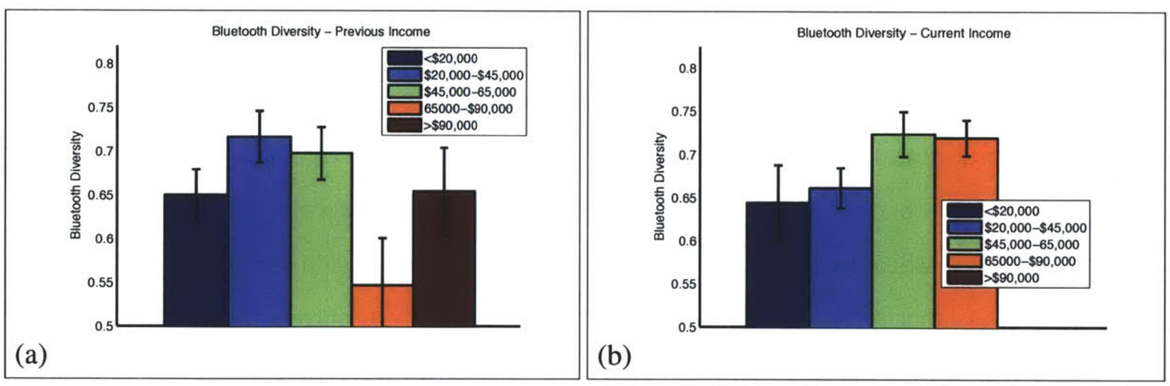

Figure 5-6: We show here the mean Bluetooth interaction diversity $D_{\text{call}}(i)$ and its standard error for individuals in different income categories. The left plot is based on previous household income, and the right plot is based on current household income. There exists borderline positive 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 ($r = -0.28, p > 0.60$).

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 behav-

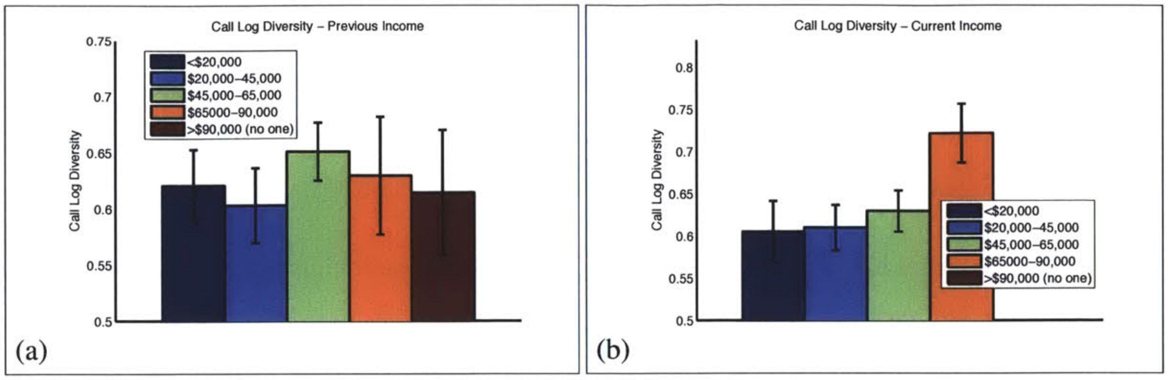

Figure 5-7: We show here the mean call diversity $D_{\text{call}}(i)$ and standard error for individuals in different income categories. The left plot is based on reported previous household income, and the right plot is based on reported current household income. There exists positive correlation between 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 ($r = 0.003, p > 0.80$).

ioral 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 [Frijters et al., 2004], they naturally feel more confident [Paridon et al., 2006] and secure in exploring new social potential [Clydesdale, 1997] [Pong and Ju, 2000]. 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.

5.5 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) [Androlib, 2011]. 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 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.

5.5.1 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 Figure 5-8(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 Figure 5-8(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.

5.5.2 Network Effects

5.5.2.1 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

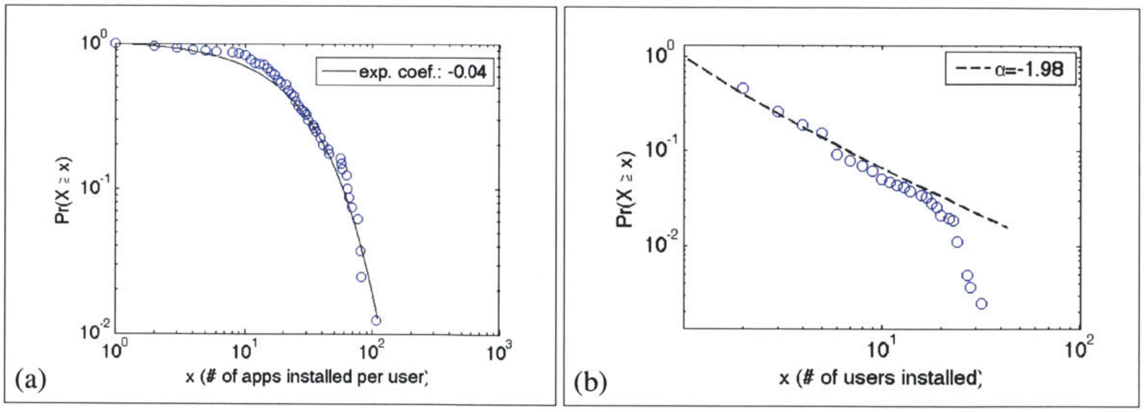

Figure 5-8: App statistics. (a) Distribution of the number of users per app. (b) Distribution of the number of apps per user.

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 hour 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 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.

5.5.2.2 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.

5.5.3 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 [Pan et al., 2011]. 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 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.

5.6 Conclusion

These are just a few examples from the initial analysis performed on the data. They help demonstrate the great potential of the dataset to create a virtual imaging chamber for a living community, helping us analyze different aspects of the community and individual's lives. Additional ongoing work includes our studies of the connection between sleep, mood, and socialization in a healthy population [Moturu et al., 2011b, Moturu et al., 2011a], where we find a statistically significant bidirectional relationship between sleep and mood, as well as correlations in sleep and mood among couples. These results also match known results for healthy populations and demonstrate how this type of automated data collection techniques can be applied to this and other public health areas. As this dataset is so rich and vast, we are currently initiating collaborations with peer researchers from a multitude of domains to help investigate different research directions. While the current chapter discussed mainly passive observations, in the next chapter we touch on the aspect of active interventions, and targeted experimental designs for an ongoing living laboratory.

—*"If you want to build a ship dont herd people together to collect wood and dont assign them tasks and work, but rather teach them to long for the endless immensity of the sea."*

Antoine de Saint-Exupery, the Wisdom of the Sands

6

Investigating Social Mechanisms in the Real World

One of the key components of the Social fMRI approach is the "*functional*" aspect of applying carefully designed stimuli into the previously observed system, and measuring their effect. In the case of social and psychological investigation of how a group of people operate, the most interesting types of stimuli are those that are aimed at social dynamics and social mechanisms. The goal of these interventions is to (1) learn more about the innate social mechanisms in humans and societies, and possibly allow us to quantify qualitative ideas like "social capital" in a real world environment. (2) Engage in social mechanism design - develop and test new mechanisms that allow us to improve outcome for individuals and the community at large (like health-related goals, for example).

In this chapter we review the most significant intervention performed during the Friends and

Family study - The "FunFit" fitness centered social mechanism design experiment. Aside from the main analysis of the experiment's direct results, a great amount of work has gone into the careful experimental design of the intervention so that the in-situ experiment would be comparable to the reliability and intentionality of ex-situ laboratory experiments. In the chapter we touch on these aspects as well.

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 [Centola and Macy, 2007] like effect related to pre-existing social ties between participants.

Aside from the specific research goals of the study, we hope that this work will contribute to an emerging methodology of *design and engineering for social support*

6.1 Physical Activity Sensing and Feedback

In this intervention 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.Proposed explanations link the problem to societal changes that promote less physical activity and increased food consumption [Hedley et al., 2004, Hill and Peters, 1998, Chang and Lauderdale, 2005], while Madan and Pentland show this might be due to norming effects of exposure rather than friendship per se [Madan et al., 2011, Madan et al., 2010b]. Physical activity has been shown to help alleviate the burden of obesity and other health conditions [Bauman, 2004, Blair et al., 2001, HASKELL, 1994]. Over the past two decades, the accelerometer has been established and refined as a tool for tracking physical activity [Bouten et al., , Eston et al., 1998, Troiano et al., 2008]. Accelerometry-based sensors have been found to provide more accurate estimates than other widely-used proxies for energy expenditure [Eston et al., 1998]. Although there is some error associated with using accelerometers to track energy expenditure in free-living situations, especially associated with their shortcomings in measuring sedentary activity and static exercise such as weight training [Bouten et al., , Hendelman et al., 2000], a significant relationship between accelerometer output and energy expenditure has nevertheless been established [Bouten et al.,]. Several studies in the ubiquitous computing literature have targeted this important problem domain. Ubifit [Consolvo et al., 2008, Consolvo et al., 2009] is one of the most extensive works investigating ways to encourage physical activity. Additional projects include Fish'n'Steps [Lin et al., 2006] and Houston [Consolvo et al., 2006], and additional works as reviewed in [Consolvo et al., 2008] and [Maitland and Siek, 2009]. 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.

Of particular relevance are those studies that involve social components [Lin et al., 2006, Consolvo et al., 2006, Anderson et al., 2007, Toscos et al., 2008, Foster et al., 2010]. It has long been established that social support is a resource for behavioral change and an indicator for health [Berkman and Glass, 2000], 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 [Anderson et al., 2007, Saponas et al., 2008], 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 [Worksmartlabs, 2011] use the phone's acclerometer, 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.

6.2 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 were presented current, past, and total reward accumulated, as shown in Figure 6-1, 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.

6.2.1 Experimental Conditions

6.2.1.1 Control Condition

All conditions have a baseline of self-monitoring. In the control condition, subjects saw only their own progress as visualized in Figure 6-1. Also, reward given to the control subjects depended only on their own activity.

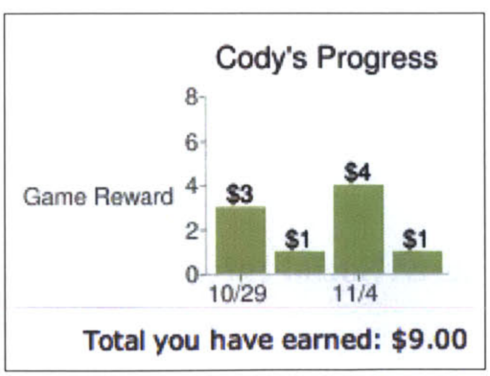

Figure 6-1: Reward display for participants in the Control condition.

Figure 6-2: Sample excerpts from reward page shown to participants in the Peer-See (left) and Peer-Reward (right) conditions.

6.2.1.2 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. A sample of the feedback presented to a Peer-View subject is shown in Figure 6-2.

6.2.1.3 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. 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. A sample of the feedback presented to a Peer-View subject is shown in Figure 6-2. 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.

6.2.2 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 seconds every 2 minutes. During the 15 seconds, raw 3-axis accelerometer measurements are sampled at 5Hz rate and combined to compute the vector magnitude for each sample. The variance of the the magnitude in each one-second block is then computed [Eston et al., 1998]. 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. [Saponas et al., 2008] 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.

6.2.3 Game Reward Calculation

Game reward was calculated every three days, using a reference window of the seven days preceding the current 3-day bin.

6.2.4 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 [Lin et al., 2006, Consolvo et al., 2006, Consolvo et al., 2008], we faced similar challenges to those discussed in [Maitland and Siek, 2009] 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 towards 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 diadic and asymmetric relationship of the person doing the activity vs. the person receiving the reward, motivated to convince the first.

6.3 Preparatory Analysis

6.3.1 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.

6.3.2 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 $b^i_j = 1$ 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:

$$\arg \max_{b^i_j, i,j \in \{1,...,N\}} \sum_{i=1,...,N} \sum_{j=1,...,N} b^i_j \mathbf{R}_{i,j} \tag{6.1}$$

$$\begin{aligned} \text{subject to: } \forall i, \sum_{j=1,...,N} b^i_j &= 2, \\ \forall i \in \{1, ..., N\}, b^i_i &= 0, \\ \forall i, j \in \{1, ..., N\}, b^i_j + b^j_i &< 2, \\ \sum_{i=1,...,N} \sum_{j=1,...,N} \mathbf{M}_{i,j} b^i_j &= 0. \end{aligned}$$

The first constraint forces each participant to have exactly two buddies. The second constraint is set so that participants can not 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.

6.4 Post-Intervention Analysis

6.4.1 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 6.1.

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

Table 6.1: Number of subjects in each condition

6.4.2 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 6.2. 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.

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

Table 6.2: Dates and days of periods used for analysis.

6.4.3 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 minutes per member.

6.4.4 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 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 6.3 confirms that activity varies greatly throughout the day, and that it correlates with general intuition about the times of high and low activity.

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

Table 6.3: 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.

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 two-minute intervals. Data was considered "missing" for an interval if there was no accelerometer score logged for that interval. As also assumed in [Eagle and Pentland, 2006], 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 accross the day, while in days with fewer than 50% of possible readings, they were not uniformally 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 measurments.

6.4.5 Activity Levels by Condition

Table 6.4 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.

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

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$

Table 6.4: 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.

6.4.6 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 6.5 and 6.6 present information on reward efficiency for this dataset, based on actual monetary reward paid. Table 6.6 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 6.5 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 minutes of physical activity per week, per participant.

As the underlying differences between the two social conditions were not clearly apparent, in Table 6.6 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.

6.5 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.

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

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 [Centola and Macy, 2007], 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.

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**

** $p < 0.01$

Table 6.6: 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.

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 [Foster et al., 2010, Consolvo et al., 2006] suggest that adding explicit communication mechanisms to the technical system might add to the social effects of the intervention.

6.6 Chapter Contributions

In this chapter we presented 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 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.

— *"Beware of the man who works hard to learn something, learns it, and finds himself no wiser than before."*

Kurt Vonnegut

7

Learning about Learning: Tools for The Practitioner

As discussed earlier, mobile phones are quickly becoming the primary source for social, behavioral, and environmental sensing and data collection. Todays smartphones are equipped with increasingly more sensors and accessible data types that enable the collection of literally dozens of signals related to the phone, its user, and its environment. A great deal of research effort in academia and industry is put into mining this raw data for higher level sense-making, such as understanding user context, inferring social networks, learning individual features, predicting outcomes, and so on. In many cases, this analysis work is the result of exploratory forays and trial-and-error. Adding to the challenge, the devices themselves are a limited platform, and any data collection campaign must be carefully designed in order to collect the right signals, in the appropriate frequency, and at the same time not exhausting the devices limited battery and processing power.

While many existing works have demonstrated results for modeling and inference of social network structure and personal information out of mobile phone data, most are still mainly proofs of concept in a nascent field. The work of the data scientist is still that of an artisan, using personal experience, insight, and sometimes gut feeling, in order to extract meaning out of the plethora of data and noise.

As the field of computational social science matures, there is need for more structured methodology. One that would assist the researcher or practitioner in designing data collection campaigns, understanding the potential of collected datasets, and estimating the accuracy limits of current analysis strategy vs. alternative ones. Such methodology would assist it in the process of maturing from a field of craft and exploration into a more rigorous field of science and engineering.

In this chapter we investigate the properties of learning and inference of real world data collected via mobile phones over time, and over varying sample sizes. We also show the potential of the Social fMRI in helping us do investigations such as this, and help us formalize generalizable methodologies. In particular, we look at the dynamic learning process over time, and how the ability to predict individual parameters and social links is incrementally enhanced with the accumulation of additional data. To do this, we use the Friends and Family dataset, which contains rich data signals gathered from the smartphones of 140 adult members of a young-family residential community for over a year, and is one of the most comprehensive mobile phone datasets gathered in academia to date.

We develop several models that predict social and individual properties from sensed mobile phone data, including detection of life-partners, ethnicity, and whether a person is a student or not. Then, for this set of diverse learning tasks, we investigate how the prediction accuracy evolves over time, as new data is collected. Finally, based on gained insights, we propose a method for advance prediction of the maximal learning accuracy possible for the learning task at hand, based on an initial set of measurements. This has practical implications, like informing the design of mobile data collection campaigns, or evaluating analysis strategies. We then perform a similar analysis for evaluating how social-network based learning improves as the size of a sample within a community increases. It is very intuitive that the more "coverage" of a community we have, the higher our prediction rates. However, there are many questions we can ask about the process itself. Does subject added to the data pool offer an equal marginal contribution to our prediction accuracy? Can we predict our incremental accuracy based on adding more people? Can we use this to decide on our sample size for a given experiment?

7.1 Understanding the evolution of learning over time

7.1.1 Methodology

7.1.1.1 Machine Learning Predictions

In order to evaluate learning over time, which is the main goal of the our current work, we needed a set of learning and prediction models to work with. These will be mostly illustrative models for the purpose of our discussion, which enable us to conduct our main analysis. In order to achieve our final goal of predicting participants' personal and social information, we utilized two approaches first is a machine learning approach, described in this section, and the second is a social network based prediction approach, described in the following section.

The first step in applying the machine learning methodology is to create features vectors for each participant in the study. Each feature vector contains information on the participants communication and phone usage patterns as were collected during the study. In order to cope with the huge amount of data collected during the study, we developed code using C and Pythons NetworkX software package library. Our code parsed the collected data, and extracted feature vectors for each participant. We extracted 32 different features within a specified time interval. Namely, we extracted following features for each participant:

- **Internet usage features:** We calculate the number of distinct searches performed using the phones browser the participant queried, and the number distinct bookmarks the saved by the user.
- **Calls pattern features:** We compute the total number of calls, the number of unique phone numbers each user was in contact with, and the total duration of all calls. We also

calculate the number of incoming/outgoing/missing calls and the total call durations according per call type.

- **SMS messages pattern features:** We compute the total number of SMS messages, the number of unique phone numbers each participant connected with via SMS, and the of total incoming/outgoing SMS messages.
- **Phone applications related features:** We count the number of applications installed and uninstalled on each device. We also compute the total number of currently running applications (originally sampled every 30 seconds).
- **Alarm clock features:** We count the number of alarm-clock alarms and the number of snooze presses for each participant that used our alarm clock app.
- **Location features:** We calculated the number of different cellular cell tower ids and the number different wifi network names names seen by the smart-phone. These features act as a rough indication of the number of different locations a participant visited during the time period.

Our next step was to extract all participant features for different time intervals. Using the extracted features we can build different classifiers that are able to predict the participants' personal information. We used the WEKA software [Hall et al., 2009] in order to test different machine learning algorithms. In our experiments we evaluated a number of popular learning methods: we used WEKA's C4.5 decision trees, Naive-Bayes, Rotation-Forest, Random-Forest, and AdaBoostM1. Each classifier was evaluated using the 10-fold cross validation approach, and in order to compare results between different classification algorithms, we used each classifiers Area Under Curve, or AUC measure (also referred to as ROC Area) and F-measure results. In order to obtain an indication of the usefulness of various features, we analyzed their importance using WEKAs information gain attribute selection algorithm.

Using the machine learning approach we built five different classifiers that predict the following: (1) the gender of the participant, (2) whether the participant is a student or not, (3) whether the participant has children or not, (4) whether the participant is above the age of 30, and (5) whether the participant is a native US citizen or not.

7.1.1.2 Social Network-based Predictions

Another method for predicting a participants personal information details is using the participants' different social networks. Using the data collected in the study. We can span different types of social networks between the participants, according to different interaction modalities. In the context of this discussion, we define the following social networks each derrived from a different interaction modality:

- **SMS Social Network:** We can construct the communitys SMS messages social network (see Figure 2) as a weighted graph $G_s =< V_s, E_s >$ according to the SMS messages the participants sent. Each weighted link $e = (u, v, w) \in E_s$ in this social network represent connection between two different phone numbers $u, v \in V$, while w is the strength of the link defined as the number of SMS message send between the two phone number . The SMS network also includes encoded phone numbers outside of the study which were contacted by more than one participant.
- **Bluetooth Social Network:** We can construct a weighted network graph $G_s =< V_B, E_B >$ of face-to-face interaction according to information collected about nearby Bluetooth devices. Each link $(u, v, w) \in E_s$ in this social network represent the fact that the two devices $u, v \in V_B$ encounter each other at least one time, while the w is the strength of the link, defined as the number of times the two devices encounter one another.
- **Calls Social Network:** Similarly to the SMS social network, we can construct a network based on the participants call graph $G_C =< V_C, E_C >$ according to the participants' phone calls. In this social network each link $(u, v, w) \in E_C$ represents the fact at least one call was made between two different phone numbers,$u, v \in V_C$, while w is the strength of the link. defined as the number of calls between u and v.

By using the social networks defined above, together with different graph theory algorithms, we can predict different types of personal and social information. In order to predict the participants' significant other we analyzed the Bluetooth social network. We predicted that each participants significant other is the person that the participant spent the most time with during the measured interval. Namely, let $u \in V_B$ then: $significant - other(u)) = v|(u, v, w) \in E_B and \forall (u', v', w') \in E_b w > w'$

In order to predict the subjects ethnicity we used the SMS social network (Figure 7-1). We used the Louvain algorithm for community detection [Blondel et al., 2008], which separates the graph into disjoint groups. At each iteration, we assume that we have information on the ethnicity of at least some of the nodes. The general idea is to then generate an ethnicity prediction for the members of each detected community based on the ethnicity of the majority of known nodes in that community. This is similar to the ideas of the label propagation approach in [Mislove et al., 2010] and [Xie and Szymanski, 2011].

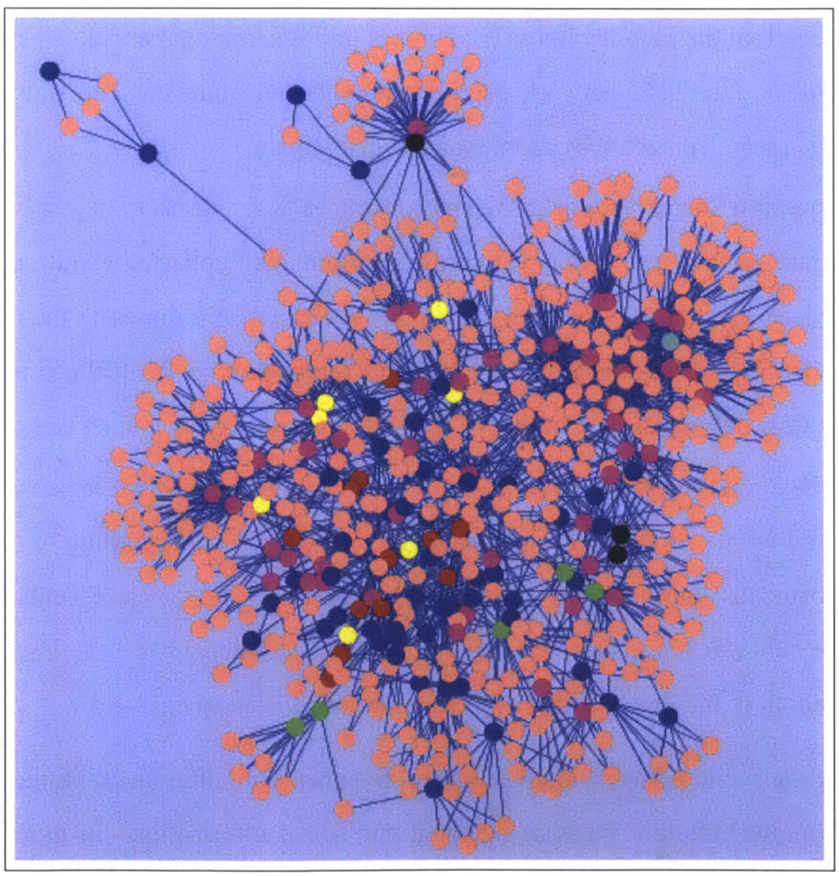

Figure 7-1: SMS-based social network graph created over 65 weeks (graph also includes unknown out-of-study nodes, which connect to at least two known in-study nodes). Different vertex colors represent different ethnicity.

7.1.1.3 Prediction of Accuracy Evolution over Time

As mentioned earlier, the goal of this analysis is to study the evolution of the learning process of personal features and behavioral properties along the time axis. For this analysis, we care less about the specific learned models and their generalizability, but rather care about using them to study and benchmark the evolution of the learning process as data accumulates. Understanding this process is of significant importance to researchers in a variety of fields, as it would provide approximation for the amount of time that is needed in order to "learn" these features for some given accuracy, or alternatively, what is the level of accuracy that can be obtained for a given duration of time.

In order to model this process we used the Gompertz function:

$$y(t) = ae^{be^{ct}} \tag{7.1}$$

This model is flexible enough to fit various social learning mechanisms, while providing the following important features:

(a) Sigmoidal advancement, namely the longer the process continues the more precise its conclusions will be.

(b) The rate at which information is gathered is smallest at the start and end of the learning process.

(c) Asymmetry of the asymptotes, implied from the fact that for any value of t, the amount of information gathered in the first t time steps is greater than the amount of information gathered at the last t time steps.

The Gompertz function is frequently used for modeling a great variety of processes (due to the flexible way it can be manipulated using the parameters a, b, and c), such as mobile phone uptake [Rouvinen, 2006], population in a confined space [Erickson et al., 2006], or growth of tumors [d'Onofrio, 2005] (see an illustration in Figure 7-2).

The applicability of the Gompertz function for the purpose of modeling the evolution of locally "learning" the preferences and behavior patterns of users was demonstrated in [Pan

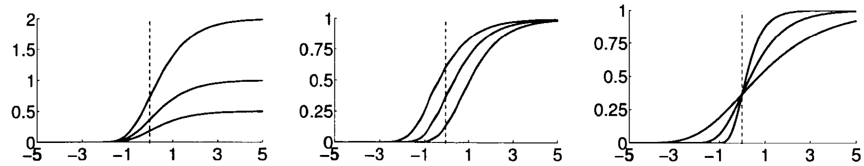

Figure 7-2: An illustration of the *Gompertz function*. The charts represent the following functions (from left to right) :$y = ae^{-e^t}$, $y = e^{-ae^t}$ and $y = e^{e^{-at}}$, for $a = \frac{1}{2}$, $a = 1$ and $a = 2$.

et al., 2011], where a prediction of the applications that mobile users would chose to install on their phones was generated using an ongoing learning process, and closely resembles the form of the Gompertz function.

For generating the models presented here, we have ran a Gompertz regression on data obtained from the predictors and classifiers developed using the methods described above. Each predictor/classifier was executed on data gathered between November 1th and November 30th, 2010. Starting from an input of a single day (November 1st), in each consecutive execution, another day of data was added to the input (so that iteration 1 was on data from November 1st, execution 2 had input of data two days, November 1 and 2 together, and so on, until an accumulation of 30 days in which the classifier ran on data from the entire month of November. Figure 7 - Figure 10 in the results section present the results of 4 of the classifiers that we have run.

7.1.2 Results

7.1.2.1 Machine Learning Classifier Results

Using the machine learning algorithms we succeed in predicting different personal information. Our prediction results vary according to the amount of data, the number of features, and the time periods for which the classifier ran on.

(a) **Gender prediction:** We predicted the gender of the participants. Our dataset included the gender information on 103 participants. Our decision tree classifier (J48) got AUC of 0.642 and F-measure of 0.611, where the most influential features where the number of Internet searches and the number of alarms. In general, female participants perform fewer search queries using their smartphones.

Prediction Model	Influential Features	AUC	F-Measure
Age	# searches	0.592	0.562
Has Children?	# missed calls	0.803	0.682
Gender	# searches	0.642	0.611
Is Student?	-	0.606	0.608
Is US-Native?	SMS Messages	0.828	0.806

Table 7.1: Predicting personal information results

(b) **US-natives prediction:** We tried to predict whether the origin of the participant is inside or outside the United States. Our dataset contained information about the origin of 86 participants. Our Nave-Bayes classifier got AUC of 0.828 and an F-measure of 0.806. Where the most influence features where: the number of incoming and outgoing SMS. In general, participants born outside the United States send and receive fewer SMS messages than US natives.

(c) **Have children prediction:** We tried to predict which of the participants in the study have children. Our dataset contained information about the children of 63 participants. Our Nave-Bayes classifier got AUC of 0.803 and an F-measure of 0.682, when using only four features: Number of missing calls, total number of application installed, distinct number of application installed, and number of alarms set. In general, participants that have children have more missed calls and fewer applications installed.

(d) **Is-student prediction:** We tried to predict which of the study participants are students (vs a different occupation). Our dataset contained information on about 88 participants, almost half of them are students. Our Rotation-Forest classifier gave AUC of 0.639 and an F-measure of 0.625.

(e) **Age prediction:** We tried to predict which of the study participants are above 30 years old or above. Our dataset contained information about 80 participants, out of them 34 were age 30 or above. Our decision tree classifier (J48) got AUC of 0.592 and an F-measure of 0.562, where the most influential features where the number of Internet searches and the number of calls. In general, participants above the age of 30 performed fewer search queries using their smartphones.

7.1.2.2 Social Network Based Prediction Results

We predicted that each participants significant other is the person that the participant spent the maximum time with during the study according to the Bluetooth social network graph. We ran this prediction on the face-to-face interactions Bluetooth graph that was created during time period of 30 days in November 2010 (Figure 7-3). Our prediction succeeded in classifying 65.6% of the couples (44 out of 67).

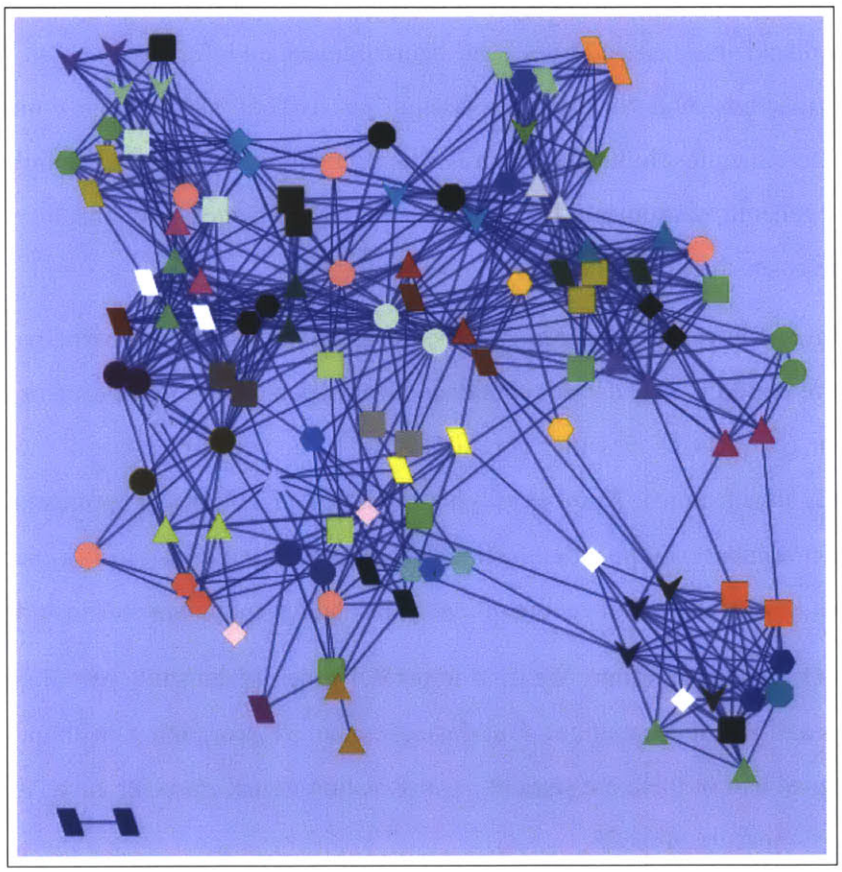

Figure 7-3: Bluetooth social network graph of face-to-face interaction during November 2010. Couples are marked with same shape and color, and each link represents at least 100 interaction.

The Louvain method for community detection partitioned the SMS social network into 13 disjoint groups (Figure 7-4). Using our method we succeeded in predicting the ethnicity of 60% of the participants (77 out of 128).

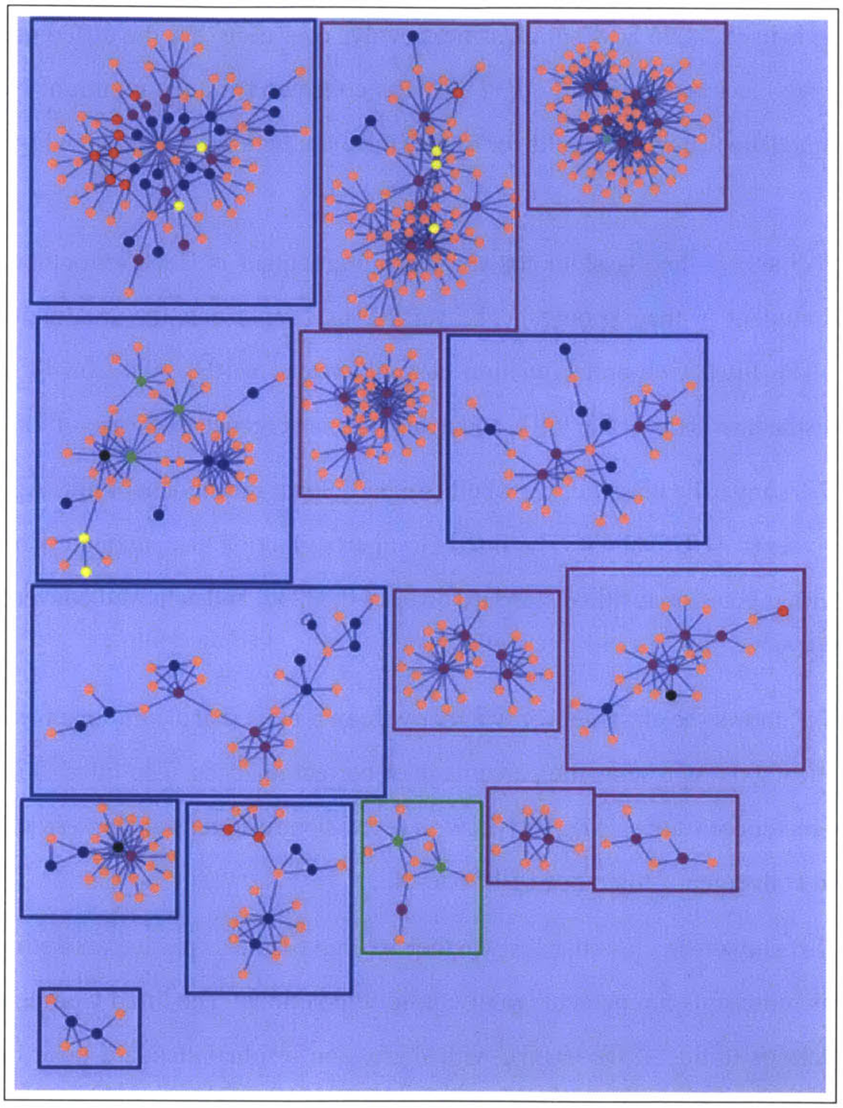

Figure 7-4: Partitioned SMS Social Network Using Louvain Algorithm: Each group is tagged with a different ethnicity according to the majority of known ethnicity of the group (similar to label propagation algorithms) (Blue: Asian, Purple: White, Green: Middle Eastern)

7.1.2.3 Incremental Learning over Time Results

We have shown in the previous sub-sections that using different communication patterns and social network graphs we are able to predict specific personal and social information. Our next step is to examine how our classifiers evolve over time. We ran different classifiers with increasingly accumulating daily data that was collected from the month of November 2010. We obtained the following results for four of the classifiers, as presented in Figure 7-5 - Figure 7-8.

Figure 7-5 shows the classifier for whether a participant is US born or not (e.g. an international student or their spouse). The vertical axis represents the area under curve (AUC) values. The fitted Gompertz function has parameters of (0.8, -0.4, -0.14), with regression residual standard error of 0.02591, and achieved convergence tolerance of 7.404e-06.

Figure 7-6 shows the classifier for whether a participant is a student or not. Again, the vertical axis represents AUC, values. The fitted Gompertz function has parameters of (0.69, -0.35, -0.06), with regression residual standard error of 0.02237, and achieved convergence tolerance of 4.095e-06.

Figure 7-7 shows the classifier for whether we can predict that a participants significant other. The vertical axis represents the percentage of correct matches. The fitted Gompertz function has parameters of (0.66, -0.78, -0.12), with regression residual standard error of 0.02762, and achieved convergence tolerance of 1.505e-06.

Figure 7-8 shows the classifier for whether we can predict a participants ethnicity. The vertical axis represents the percentage of correct predictions. The fitted Gompertz function has parameters of (0.68, -2.18, -0.05), with regression residual standard error of 0.06676, and achieved convergence tolerance of 5.568e-06.

Figure 7-14 demonstrates the correlations among the learning process dynamics of several features. It was calculated using the Pearson product-moment correlation coefficient (a measure of the linear dependence between two variables X and Y, giving a value between +1 and -1). The correlation is defined as the covariance of the two variables divided by the product of their standard deviations. In general, variables of correlation higher than 0.5 are usually considered strongly correlated.

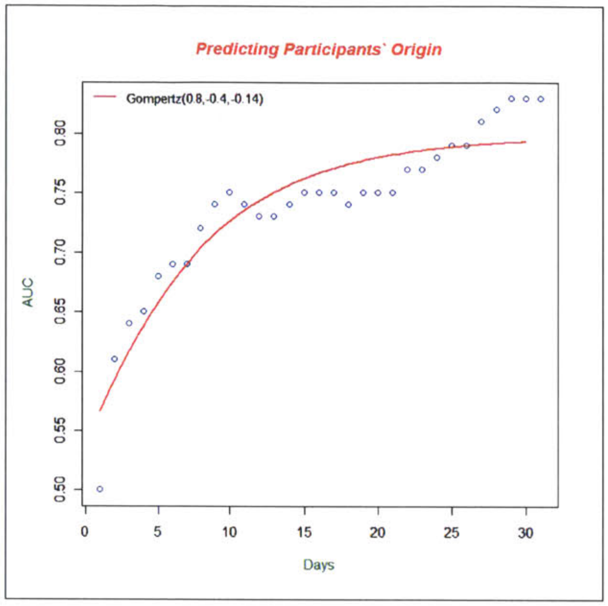

Figure 7-5: Participants' origin Naive-Bayes classifiers AUC results

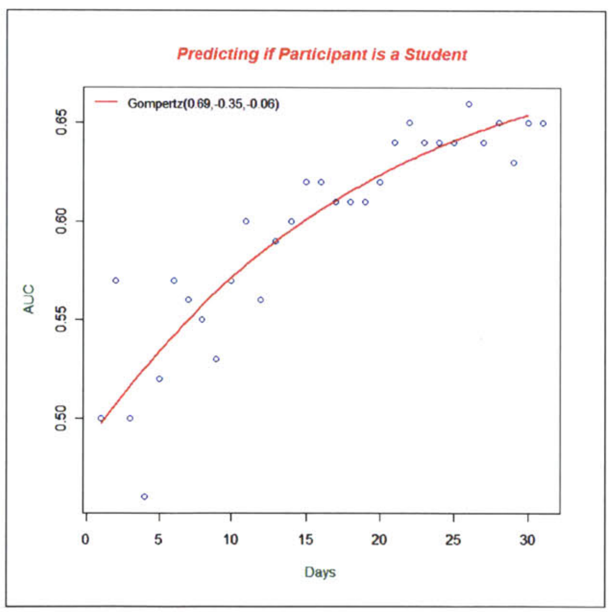

Figure 7-6: Predicting If the Participant is a Student over Time: Rotation-Forest Classifier AUC results

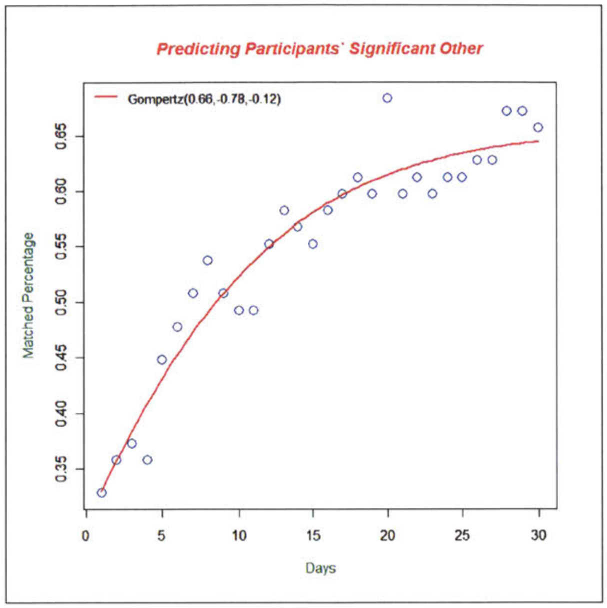

Figure 7-7: Predicting Significant Other over Time we chose the significant other as the node with the maximum strength.

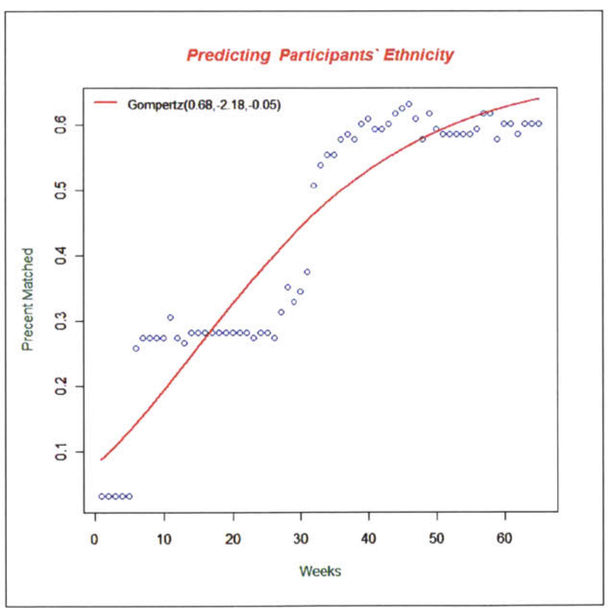

Figure 7-8: Predicting ethnicity using SMS social network over time (65 weeks) after every week we analyze the graph with the same method as described at 3.4 (Louvain Algorithm).

7.2 Discussion and Future Work

As reviewed in section 7.1.1.3, the Gompertz function is a well-known technique that has been used to model processes over time. Our analysis confirms that the evolving learning of social and individual features, as mobile phone sensing data accumulates over time, can also be fitted to the form of a Gompertz function. We see that this result is true for the prediction of different features, both social and individual, and for a set of different prediction methodologies, using a varying number of input signals, all collected via mobile phones in a field deployment.

Correlations between the evolution trends of the different learning process, as depicted in Figure 7-14, may imply underlying correlation between the raw data itself, and can hence be used as additional validation for correlated features and observations (such as the suggestion that people might have a higher tendency to marry within their own ethnic group, as has been widely observed [Kalmijn, 1998,McPherson et al., 2001]). In addition, this information could be used for informing the design of data collection configuration for an ongoing or future data collection initiative. For example, if we know of two features that are highly correlated in the same experiment, but one of them is very cheap to gather from a processing or battery power perspective, while the other is very expensive, we might decide that the cheaper one is sufficient (e.g. one requires just reading the phones built-in call-log database while the other requires battery-intensive GPS scanning). Alternatively, we might want to make sure that two correlated values are gathered in order to strengthen the result and help deal with noise.

We can take our findings further, and extrapolate, using the learned Gompertz functions, the learning behavior and limits over time. Figure 7-10 shows the result of this extrapolation, two years into the future (the original source data is just one month). We can now gain different insights. First, extrapolation can be used to predict our maximal expected accuracy. In addition we can estimate where we are on each signals estimated accuracy curve. We can then use this information to evaluate the analysis method, anticipate the timeline for increased accuracy, and understand when it is time to stop collecting/analyzing as we have reached a state of saturation. Another possible use is comparing different learning processes to one another, and using this information as part of the experiment or analysis management process. In addition, deviation from the expected curves might actually point at problems in

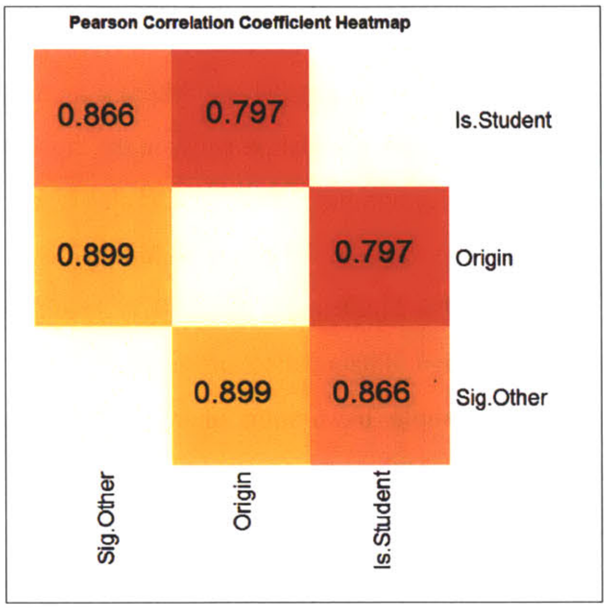

Figure 7-9: Pearson correlation between the learning process dynamics for three of the properties we predict. As might be expected, there are some strong correlation between the different evolution trajectories of the learning processes of the three features. However, notice that while some are very highly correlated (e.g. Origin Significant other), which might point out a strong correlation in the underlying data itself (i.e. people tend to get married more within the same ethnic group), other display lower correlation (e.g. Origin and Is student).

the data collection process.

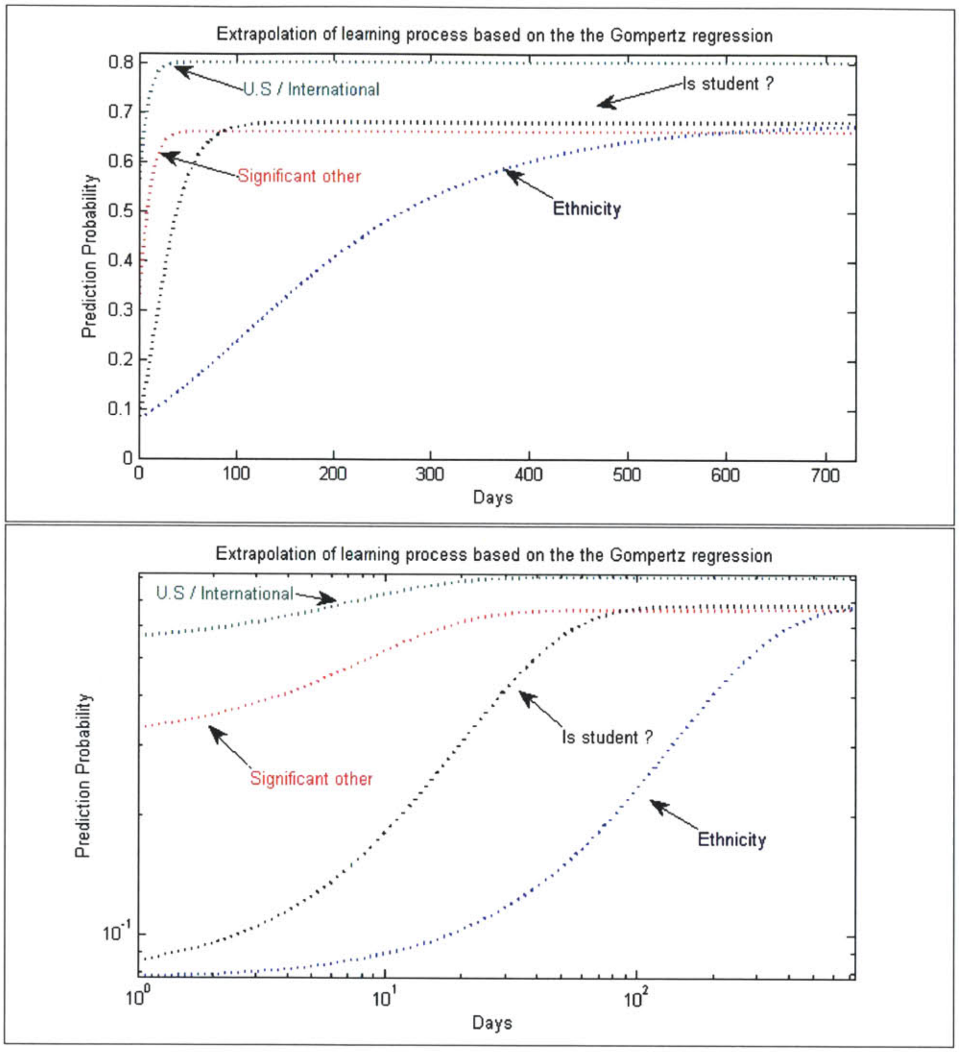

Figure 7-10: Extrapolation of the learning process based on the Gompertz regression for the four learning tasks, in linear scale (top) and log-log scale (bottom).

There are different reasons that might explain why there are the saturation limits in accuracy of our learning. For example, we can take the learning of user parameters which are based on Bluetooth proximity, as we have used in our prediction of significant-other ties and have also been used in [Madan et al., 2011, Paulos and Goodman, 2004, Eagle and Pentland, 2006, Aharony et al., 2011b, Aharony et al., 2010]. All of these analyses assume that the phone is an accurate proxy for its owner and is located where the owner is. It has been shown by Dey et al. [Dey et al., 2011] that people actually carry the phone with them much less than they might think. This discrepancy could account for some of the inaccuracies of trying to learn user parameters based on phone-sensed data. Based on our observations, we can suggest this approach as a mechanism for answering several important questions, such as:

(a) Given a social network, how easy would it be for someone who monitors the behavioral activities of its members to infer it?

(b) What kinds of social features are more difficult to learn than others?

(c) What is the highest level of prediction accuracy that can be reached in a reasonable amount of time?

This could, in-turn, inform the allocation of data collection, processing, and analysis resources, as well as investigator time. Aside from their academic importance, such questions may also have significant financial implications. Social information has become a valuable data on its own merit, of high and tangible value, as it is used by many marketing and advertising platforms for doing targeted advertising to maximize their advertisement hit rates.. Furthermore, this insight may also have broader implications in areas of defense and homeland security, due to the importance of social information for cyber criminals and terrorists:

(a) Selling to highest bidder (both "legit" bidders, advertisers, etc., or in the black market to other attackers) [Krishnamurthy and Wills, 2009].

(b) Bootstrapping other attacks e.g. using this as part of a complex "Advanced Persistent Threats" (APT) attack [Solutionary, 2011, Binde et al., 2011].

(c) Business espionage - e.g. analyzing a competitor's customer base and profile high-yielding customers for targeted marketing [Martin Brunner and Todt, 2010].

7.3 Understanding the evolution of learning as a factor of community coverage

7.3.1 Overview and Motivation

The previous analysis explored learning over time. We then initiated a follow-up investigation using a similar approach. This time, rather than investigate the increase in available information as it accumulates over time, we evaluate the learning performance as a factor of our "coverage" of a given community. That is, the number of subjects we have access to with respect to the actual size of the measured network - our sample size within the network.

Analogously to the previous exploration, we want to analyze the evolution of the process of learning personal features and behavioral properties, as we increase the size of the sample group. For this analysis, we again do not care as much about the specific learned models and their generalizability, but about using them to study and benchmark the evolution of learning accuracy. Understanding this process is of significant importance to researchers in a variety of fields, as it would provide approximation for the needed level of coverage (sample size vs community or network size), in order to "learn" specified features for some given accuracy. Alternatively, it could give the investigators an idea of the expected level of accuracy that can be obtained for a given socially-relevant data collection campaign.

In this section we describe ongoing work in this direction, and our new and promising results.

7.3.2 Procedure

7.3.2.1 Social Network Based Prediction

We used the SMS social network as described in 7.1.1.2, and also divided it using the Louvain community detection algorithm as before. In order to predict participant attribute a, we look at the community which the participant belong to. We predicted attribute a's value as equal to the majority value in the node's community. We tested 5 prediction models that work similarly to the social prediction ones we did earlier: Ethnicity, religion, has-children?, origin (is American?),and age (is over 30?).

Similarly to Figure 7-4 , Figures 7-11 and 7-11 show the

7.3.2.2 Incremental Evaluation

We then performed these classifiers in a set of incrementally increasing sample sizes - starting from a single subject, until covering 90% of the subject population community. For each incremental step with sample size k, we performed 100 random trials. In each trial k random

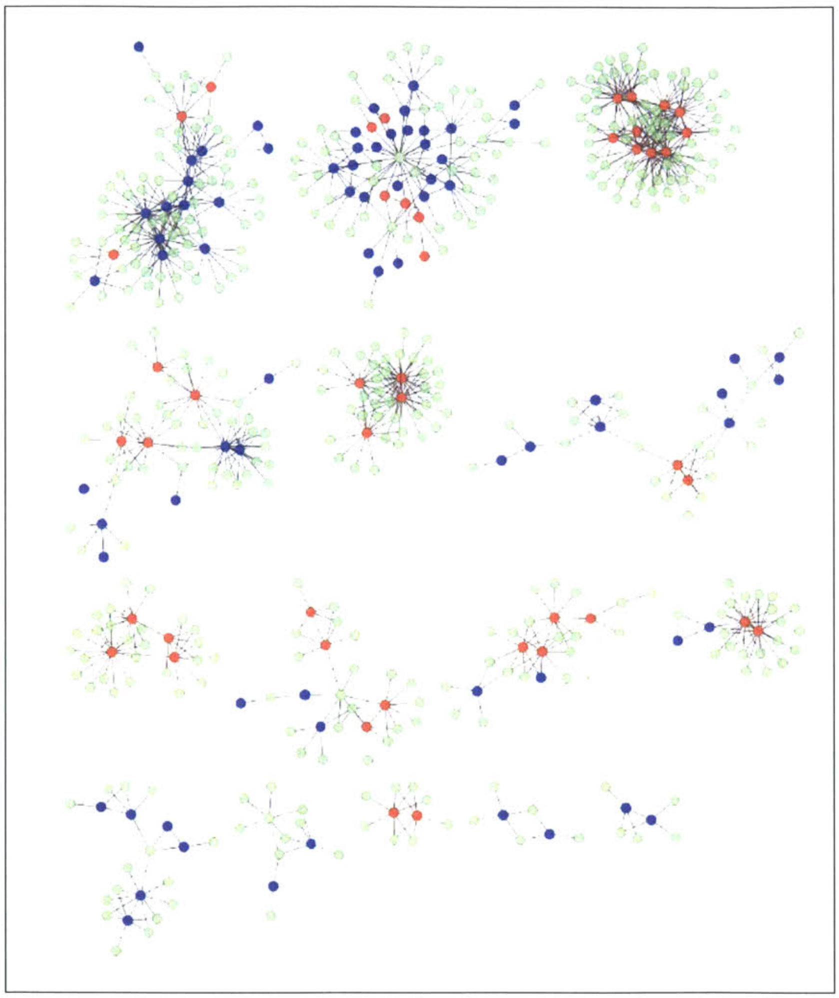

Figure 7-11: Partitioned SMS Social Network Graph created with data of over 65 weeks (each out-of-study node is connected to at least two in-study nodes), marking origin. Red: American native, Blue: Non-American (International)

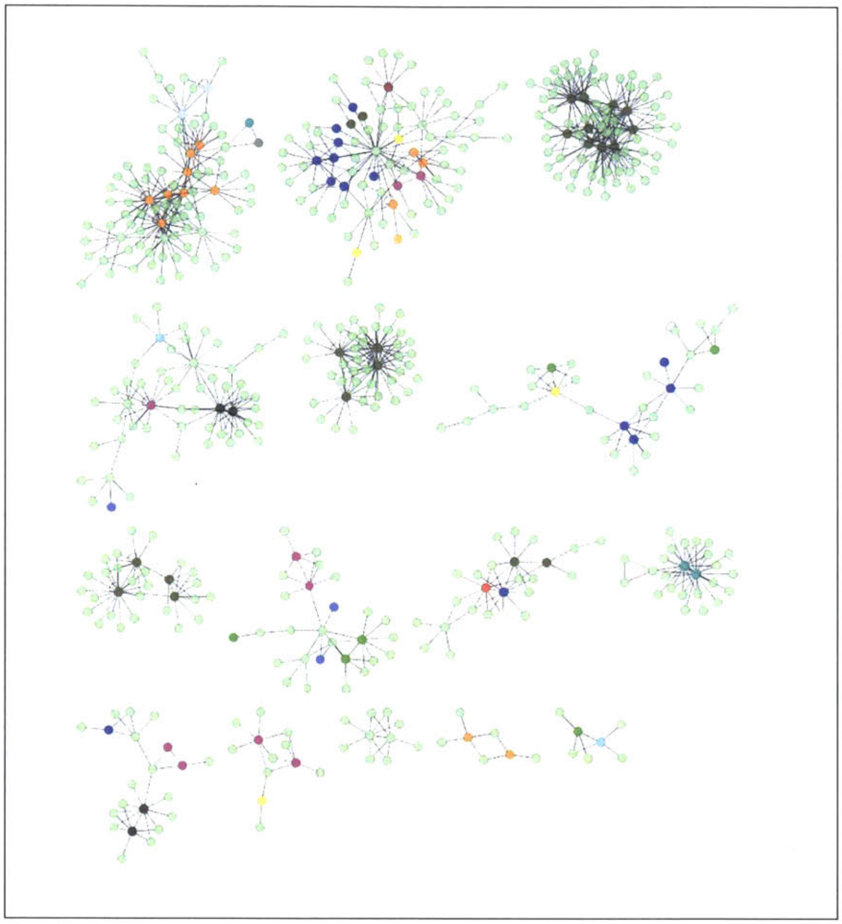

Figure 7-12: Partitioned SMS Social Network Graph created with data of over 65 weeks (each out-of-study node is connected to at least two in-study nodes). Different colors represent different reported religious affiliation.

users were picked as the training group, while the rest of the community was the test group, and calculated the average performance for sample size k. And so, we started with 100 random training groups containing a single user each (and the test groups that contained the rest of the community), in the next step we took 100 random groups of 2 users each, and so on, until testing 100 groups in the size of 90% of the users.

7.3.2.3 Gompertz Modeling

As earlier, we conducted a Gompertz regression to fit the best parameter values to the different learning processes. The regression yielded surprisingly accurate results:

- **Ethnicity** : Residual standard error: 0.02026, Achieved convergence tolerance: 3.167e-06
- **Religion:** Residual standard error: 0.02115, Achieved convergence tolerance: 8.747e-06
- **Children:** Residual standard error: 0.01901, Achieved convergence tolerance: 2.464e-06
- **Origin:** Residual standard error: 0.02515, Achieved convergence tolerance: 8.127e-06
- **Age:** Residual standard error: 0.02056, Achieved convergence tolerance: 6.189e-06

Figures 7-13 presents the results over time for the five predictors that we have run so far: origin, ethnicity, age, parenthood, and religion. .

Figure 7-14 depicts the Pearson correlation between the performance of the 5 predictors as sample size increases.

7.3.3 Discussion and Future Work

We see very similar behavior among the learning processes as sample size increases. We also see there is a limit, or "saturation" where additional community coverage does not increase accuracy of prediction, or increases it only marginally. We have shown that the dynamics of this process can be modeled using the Gompertz function, and hence can be further extrapolated in order to predict bounds on the overall ability to learn specific features.

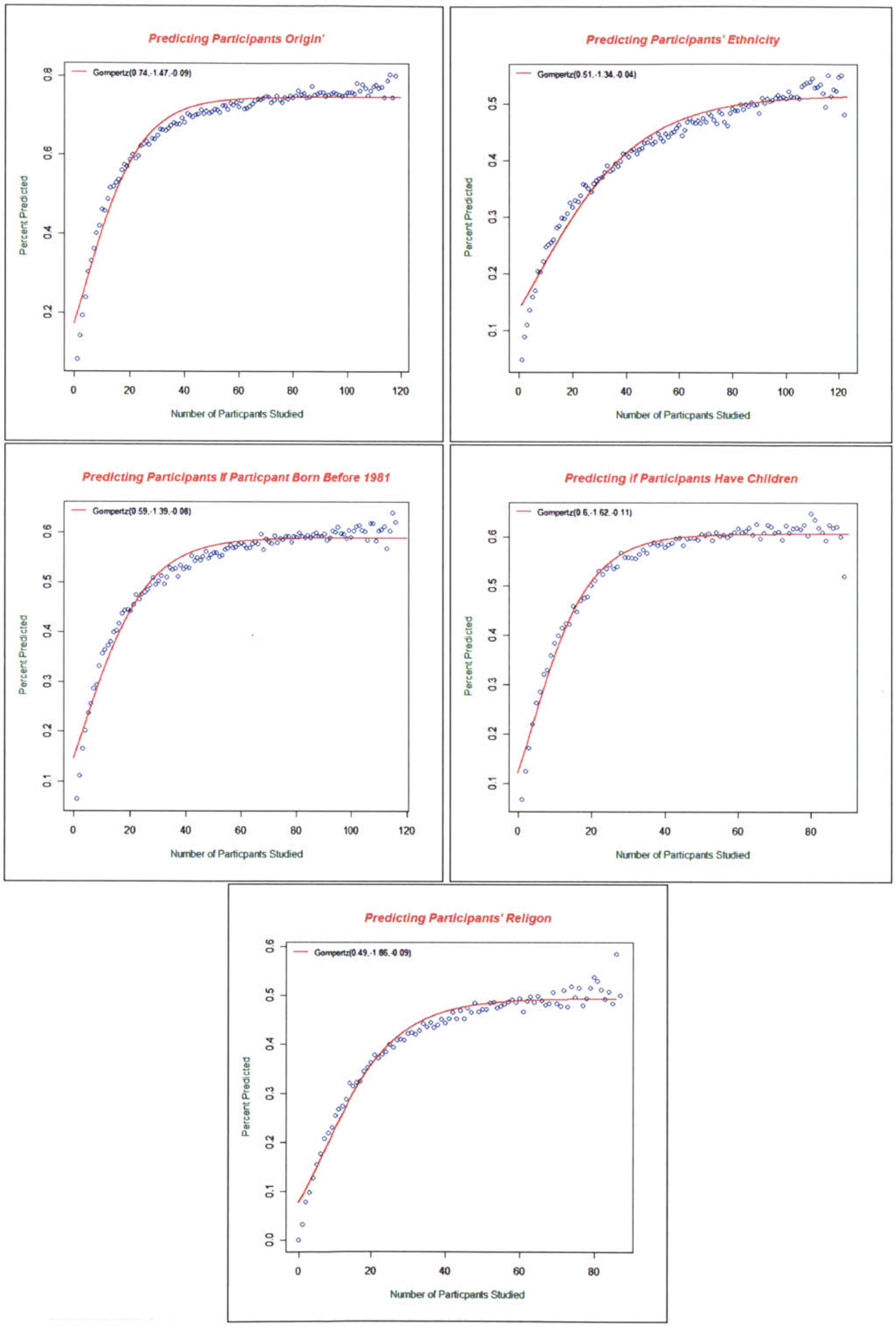

Figure 7-13: Learning process for predicting participants' Origin, Ethnicity, Age, Parenthood, and Religion, with data regarding groups of growing sizes. For each group's size, 100 random groups were generated, and their performance values averaged. Red line indicates the Gompertz regression.

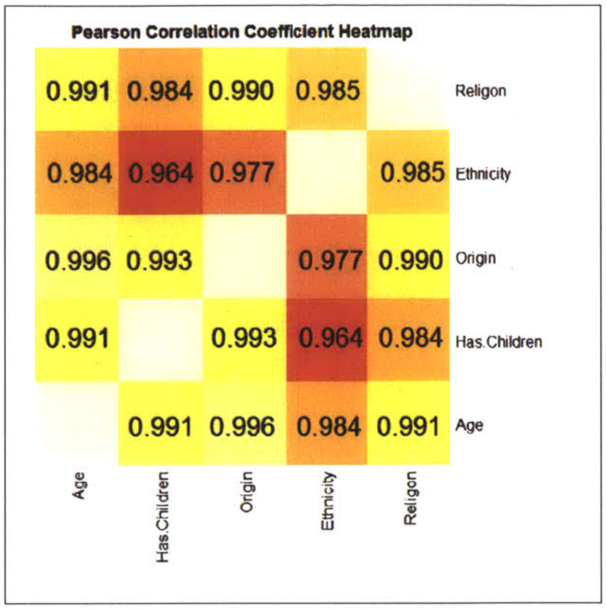

Figure 7-14: Pearson correlation matrix between the prediction vectors of the 5 classifiers.

This information can be used informing the design of the data collection configuration for an ongoing or future data collection experiment or initiative. This can be done by extrapolating the Gompertz function with the regression values found for several initial small sampling sets. Based on this extrapolation, an approximation for the maximal amount of information (or accuracy) that can be achieved with large sample sets, as well as an approximation to the accuracy that an given group size would result in. In ongoing and future work

In addition, correlations between the evolutions of the different learning process, as depicted in Figure 7-14, may imply underlying correlation between the features themselves as well as the raw data signals, and can hence be used as additional validation for correlated features and observations (such as the suggestion that people might have a higher tendency to marry within their own ethnic group, as observed in [Kalmijn, 1998, McPherson et al., 2001]).

There is a lot more to follow up on as we generalize these insights into practical rules. Right now the prediction is done on a single network modality in a single subject pool. It might be interesting to see how the underlying network affects the incremental sample size behavior - for example the number of communities and how tightly knit the network is, average node degree, and so on. We intend to test this with both the different network modalities of the

Friends and Family dataset, as well as other social network datasets.

It would also be interesting to compare how does the dynamics of the learning process as a function of the community's size correlate with the dynamics of the learning process over time, as well as generating two dimensional models of the learning processes (as a function of both time and sample size). We intend to conduct this study on the Friends and Family dataset in the coming months.

7.4 Contributions

Our main goal in this discussion was to investigate the learning processes over time and sample size, rather than evaluate the specific models and how they generalize. In future work we intent to come back to each of these models and evaluate it in detail. We are also continuing our investigation of the properties of learning and prediction of human and social constructs based on mobile phone gathered data.

The contributions of the work described in this chapter are the following:

(a) We demonstrated characteristics of incremental learning of multiple social and individual properties from raw sensing data collected from mobile phones, as the information is accumulated over time.

(b) We have shown that for different learning tasks, prediction methods, and input signals, the evolving learning of social and individual features, as mobile phone sensing data accumulates over time, can be fitted to the form of a Gompertz function.

(c) Furthermore, we proposed a method for advance prediction of the maximal learning accuracy possible for the learning task at hand, using just the first few measurements. This information can be useful in many ways, including:

- Informing real-time resource allocation for data collection, for an ongoing data collection campaign.
- Estimating accuracy limits and time needed for desired accuracy level of a given method.
- Early evaluation of modeling and learning strategies.

(d) We show that the characteristics of incremental learning along the sample-size axis resembles that of the evolution along the time-axis, and can be fitted as well to the form of a Gompertz function.

(e) We propose a method for approximating the maximal amount of information (or accuracy) that can be achieved with large sample sets, as well as an approximation to the accuracy that an given group size would result in. In ongoing and future work

(f) Finally, we presented new models for predicting social and individual features from raw mobile phone sensed data, which were developed as part of the methodological analysis.

We believe that this work is a real step towards the methodological toolkit for the computational social scientist and practitioners of the big-data ecosystem. It also shows the potential of Social fMRI's rich and finely grained data collection approach. While there will always be the need for the expert and experienced data artisan, with the exponential increase in accumulated data and the rise of a big-data ecosystem, there is an imperative need to create a more accurate science and engineering of data collection, processing, and analysis. Our work is a building block in this larger effort.

Part III

Finale

8

Conclusion

We are in the midst of a multipronged revolution. Technology is able to place tens of sensors into a single device that fits in the palm of a hand, with computing power that rivals that of supercomputers from a few years back, and a communication channel with bandwidth comparable to what was used to connect entire cities just a few decades ago. Our evolving culture makes us want to carry these devices with us around the clock, giving us the ability to connect to any human on the planet, as well as access to any and all of human knowledge, including the occasional game of hurling infuriated birds at fortified complexes built by shivering green pigs [Rovio, 2009]. Alongside, the coming of age of the mobile computing operating systems and their third-party development frameworks are giving developers and researchers unprecedented access to the richness of sensors, processing power, and the other capabilities of the mobile devices. Just as many high-end technologies make their way down from large corporations and institutions to the common man, end users now have the power to "quantify" themselves. The same tools that have been helping large organizations to data-mine, optimize, and improve themselves, is now available to consumers. They can collect, process, and make use of their life's data, with the potential to improve their lives and the lives of others. Data that has been previously available to service providers and advertisers about user mobility and behavior patterns can now be made available to the users themselves, and, with their consent, can be used by science to help us understand ourselves better. This premise enables us to change the way we do science, as well as the scale and scope of what constitutes a social science experiment.

8.1 Two Experiments, One Community, Sixty Five Years Apart

The specific community in which the study was conducted is very symbolic: It is in this same community, more than 60 years ago, that Leon Festinger and his colleagues conducted their groundbreaking field study of social pressures in informal groups [Festinger et al., 1950]. The residence buildings at the time of Festinger's study were several hundred feet away from the current buildings, and they were set up as temporary housing before the permanent housing buildings were constructed. The population was more homogeneous than the Friends and Family population, since at the time the main occupants were World War II veterans and their families, returning to school after the war. However many characteristics of the community existed even then - It was a vibrant community of graduate student families, where neighbors interacted with one another, organized in formal and informal groups and affiliations, and participated in shared activities. There are also similarities in the research goals and questions - from how communities operate, to the formation of friendship, and how "things" spread in a network. Festinger et al.'s study was one of the first studies that used adjacency matrices to represent social networks (see Fig. 8-1). While very trivial and common nowadays, at that time the investigators had to spend several pages of text to explain the ideas behind it and the different arithmetic properties that the adjacency matrix representation affords. Another quite amazing anecdote is that similarly to our present-day goals, the researchers were also interested to understand how things spread through a social network, and they did this by "planting" fake rumors in the community, and trying to follow up on their spreading patterns, as illustrated by them in Figure 8-2. There were differences as well. As the investigators report: "Only wives were interviewed. The reasoning behind this decision was the following: Since men were all deeply engrossed in their studies, careers, and part-time jobs, the women of the family usually bore the burden of social life; the women would be easier to contact than the men, who were busy in their classes and labs at all hours." (p.184). However, some of the more significant differences are in the great leaps that science has made with tools and technology. The richness of the data, the continuous collection of a great number of signals and data types, gives us an unprecedented, multi-dimentional image of community life. This combines together with traditional methods, as technology is still not able, and might never be completely able to fully replace the human. These methods are, however, a powerful tool that can be added to the arsenal of the investigator, or, to quote Festinger: "The study of anything as complex as the social life of an entire community requires the use of most of the methodological tools available to the social scientist."

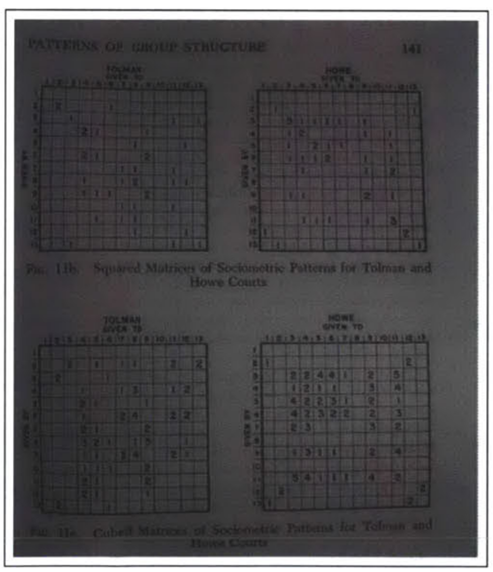

Figure 8-1: Adjacency matrix representation of a subset of the community in Festinger's study [Festinger et al., 1950].

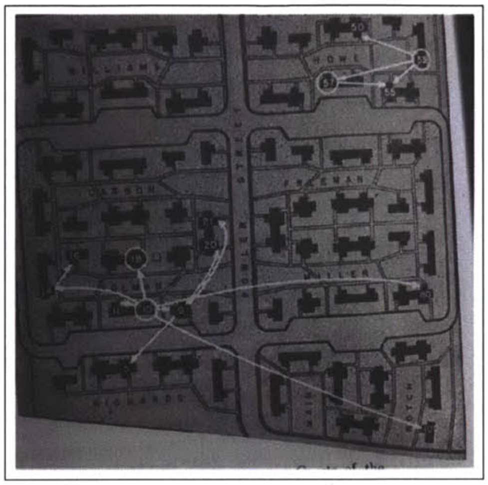

Figure 8-2: Illustration of the spreading of a rumor in Festinger's study [Festinger et al., 1950].

8.2 Closing Remarks

I could not begin to enumerate the lessons I have learned working on this multi-faceted project. A field study of this scope comes with a host of challenges. Aside from obvious technical and logistical challenges, it is important to note the human aspects of working with a living community. Even the act of deploying the experiment or a sub-component, handing out phones and training to over 100 people, can take days or even weeks - coordinating busy schedules, setting up meetings and then rescheduling over and over again due to changes and constraints. The study required us to deal with a broad range of attitudes, habits, personalities, and lifestyles. Our main goal was to make it as effortless on their lives, while maximizing the scientific validity and utility of the results. Personally, it took me time to adjust and accept that data collection in the real world is extremely noisy, and can never be perfect. At the beginning of the study, I wanted to try and keep track of the time each phone was broken and not collecting data, mark any trips and special events in people's lives so that we could later use it to account for some of the variability in the data. I quickly realized that this task is impossible to keep up over time. The good news is, however, that in many instances, this accounting is not a necessity. In some scenarios the automatically collected data itself can be used to detect and mark such anomalies. In many other cases, the salient trends rise up from the noise on their own, as we could also see in our investigation of learning over time described in Chapter

7. Through the study, we explored and improved different techniques to help preserve subject adherence over time, a challenge faced by any longitudinal study. We have found that the best solutions are those that tie the desired behavior (like having subjects carry the phone with them as much as possible, or filling out daily surveys month after month), to egocentric motivations of the participants - from scaffolded mechanisms like micro-payments and financial rewards, to games and competitions where a byproduct of a successful participation is to keep the phone charged, all the way to leveraging the subject's curiosity of viewing their own data about their lives.

In this thesis we presented four levels of contribution (also illustrated in Figure 8-3):

(a) At the base level, we introduced the *Social fMRI methodology*, which combines a rich-data experimental approach with carefully designed interventions.

(b) At the second level, we developed the *Funf system*, one instance of a Social fMRI system implementation. Funf was implemented, deployed in the field for over 16 months, and released to the world as an open source framework with a growing community of users, developers, and researchers, as well as received several awards.

(c) At the third level, we conducted the *Friends and Family Study* and compiled the *Friends and Family dataset*, the riches real-world dataset of its genre to date. This is one instance of a Social fMRI deployment using the Funf system.

(d) Finally, we presented a set of *experimental analyses and findings* that contribute to our understanding of important research questions in computational social science in addition to demonstrating the methodology's potential. Among the results described in this thesis we reviewed the design and evaluation of a novel mechanism for social support in a health-related context, the observation that the diffusion of mobile applications relies more on the face-to-face interaction ties than on self-perceived friendship ties, and a gained understanding of the evolution of modeling and prediction processes over time and varying sample sizes.

More than any specific contribution, we hope that the real contribution of our work will present itself in the way that others use these as platforms for future discovery and innovation. The Friends and Family dataset could be used for many future analyses and results, to be

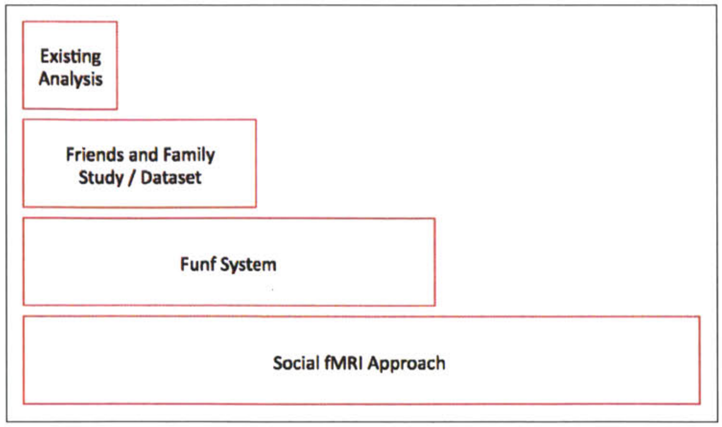

Figure 8-3: Summary of key contributions and they way they build on one another.

performed by us as well as by partners from the academic community. The Funf system can be used for many other deployments and experimental scenarios, and it is already being used by several other projects. Finally, the Social fMRI approach could be implemented in other means, with different systems, and hopefully expanded to its fullest potential. This work is just a step in a long and challenging road. In the same way that fMRI techniques help map the interconnections and mechanics of the human brain, I 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.

Bibliography

[Abdelzaher et al., 2007] Abdelzaher, T., Anokwa, Y., Boda, P., Burke, J., Estrin, D., Guibas, L., Kansal, A., Madden, S., and Reich, J. (2007). Mobiscopes for human spaces. *IEEE Pervasive Computing*, 6:20–29.

[Aharony and Gardner, 2011] Aharony, N. and Gardner, A. (2011). Funf developer site. http://code.google.com/p/funf-open-sensing-framework/.

[Aharony et al., 2011a] Aharony, N., Gardner, A., and Sumter, C. (2011a). Funf getting started guide. http://funf.media.mit.edu/gettingstarted.html.

[Aharony et al., a] Aharony, N., Gardner, A., Sumter, C., and Pentland, A. Funf: Open sensing framework, developer website. http://code.google.com/p/funf-open-sensing-framework/.

[Aharony et al., b] Aharony, N., Gardner, A., Sumter, C., and Pentland, A. Funf: Open sensing framework. http://funf.media.mit.edu.

[Aharony et al., 2011b] Aharony, N., Pan, W., Ip, C., Khayal, I., and Pentland, A. (2011b). Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, (0):–.

[Aharony et al., 2010] Aharony, N., Pan, W., Ip, C., and Pentland, A. (2010). Tracing mobile phone app installations in the "friends and family" study. In *Proceedings of the 2010 Workshop on Information in Networks (WIN'10)*, NY, NY.

[Ahn et al., 2010] Ahn, Y.-Y., Bagrow, J. P., and Lehmann, S. (2010). Link communities reveal multi-scale complexity in networks. *Nature*, 466:761+.

[Altshuler et al., 2010] Altshuler, Y., Dolev, S., Elovici, Y., and Aharony, N. (2010). Ttled random walks for collaborative monitoring. In *NetSciCom 2010 (Second International Workshop on Network Science for Communication Networks)*, San Diego, CA, USA.

[Anderson et al., 2007] Anderson et al., I. (2007). Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. *Mob. Netw. Appl.*, 12:185–199.

[Androlib, 2011] Androlib (2011). Androlib statistics page.

[Anokwa et al., 2009] Anokwa, Y., Hartung, C., Brunette, W., Borriello, G., and Lerer, A. (2009). Open source data collection in the developing world. *Computer*, 42(10):97 –99.

[Aral and Walker, 2010] Aral, S. and Walker, D. (2010). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*.

[Avery and Zemsky, 1998] Avery, C. N. and Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. *American Economic Review*, 88(4):724–48.

[Banks and Hersman, 2009] Banks, K. and Hersman, E. (2009). Frontlinesms and ushahidi - a demo. In *Information and Communication Technologies and Development (ICTD), 2009 International Conference on*, page 484.

[Barabasi and Albert, 1999] Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

[Barker and of Kansas. Midwest Psychological Field Station, 1968] Barker, R. and of Kansas. Midwest Psychological Field Station, U. (1968). *Ecological psychology: concepts and methods for studying the environment of human behavior*. Stanford University Press.

[Bauman, 2004] Bauman, A. (2004). Updating the evidence that physical activity is good for health: an epidemiological review 2000-2003. *Journal of Science and Medicine in Sport*, 7(1, Supplement 1):6 – 19. Physical Activity.

[Bell and Gemmell, 2007] Bell, G. and Gemmell, J. (2007). A digital life. *Scientific American*.

[Berkman and Glass, 2000] Berkman, L. and Glass, T. (2000). Social integration, social support, and health. In Berkman, L. and Kawachi, I. E., editors, *Social Epidemiology*.

[Bernard et al., 1984] Bernard, H. R., Killworth, P., Kronenfeld, D., and Sailer, L. (1984). The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology*, 13(1):495–517.

[Binde et al., 2011] Binde, B. E., McRee, R., and O'Connor, T. J. (2011). Assessing outbound traffic to uncover advanced persistent threat. Technical report, Sans Institute.

[Blair et al., 2001] Blair, S. N., Cheng, Y., and Holder, S. J. (2001). Is physical activity or physical fitness more important in defining health benefits? *Medicine & Science in Sports & Exercise*, 33(6).

[Blondel et al., 2008] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

[Bollen et al., 2010] Bollen, J., Mao, H., and Zeng, X.-J. (2010). Twitter mood predicts the stock market. *CoRR*, abs/1010.3003.

[Bouten et al.,] Bouten, C. V., Koekkoek, K. T., Verduin, M., Kodde, R., and Janssen, J. D. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans Biomed Eng*.

[Bruggeman, 2010] Bruggeman, J. (2010). Network diversity and economic development: a comment. *Arxiv preprint arXiv:1011.0208*.

[Burt, 1995] Burt, R. (1995). *Structural holes: The social structure of competition*. Harvard Univ Pr.

[Burt, 1987] Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, 92(6):1287–1335.

[Burt, 2004] Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2):349–399.

[Cebrian et al., 2010] Cebrian, M., Lahiri, M., Oliver, N., and Pentland, A. (2010). Measuring the collective potential of populations from dynamic social interaction data. *Selected Topics in Signal Processing, IEEE Journal of*, 4(4).

[Cebrián et al., 2010] Cebrián, M., Pentland, A., and Kirkpatrick, S. (2010). Disentangling social networks inferred from call logs. *CoRR*, abs/1008.1357.

[Centola and Macy, 2007] Centola, D. and Macy, M. (2007). Complex contagions and the weakness of long ties. *The American Journal of Sociology*, 113(3):pp. 702–734.

[Chang and Lauderdale, 2005] Chang, V. W. and Lauderdale, D. S. (2005). Income disparities in body mass index and obesity in the united states, 1971-2002. *Arch Intern Med*, 165(18):2122–2128.

[Christakis and Fowler, 2007] Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4):370–379.

[Clauset et al., 2008] Clauset, A., Moore, C., and Newman, M. E. J. (2008). Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101.

[Clydesdale, 1997] Clydesdale, T. (1997). Family behaviors among early us baby boomers: Exploring the effects of religion and income change, 1965-1982. *Soc. F.*, 76:605.

[ComScore, 2011] ComScore (2011). comscore reports november 2011 u.s. mobile subscriber market share.

[Consolvo et al., 2006] Consolvo, S., Everitt, K., Smith, I., and Landay, J. A. (2006). Design requirements for technologies that encourage physical activity. In *Proceedings of CHI '06*, New York, NY, USA. ACM.

[Consolvo et al., 2009] Consolvo, S., Landay, J. A., and McDonald, D. W. (2009). Designing for behavior change in everyday life. *IEEE Computer*, 42(6):86–89.

[Consolvo et al., 2008] Consolvo et al., S. (2008). Activity sensing in the wild: a field trial of ubifit garden. In *Proceeding of CHI '08*, pages 1797–1806, New York, NY, USA. ACM.

[Cooley, 1909] Cooley, C. H. (1909). *Social Organization: A Study of the Larger Mind.* Schocken, New York.

[Davis et al., 1941] Davis, A., Gardner, B. B., and Gardner, M. R. (1941). *Deep South.* The University of Chicago Press, Chicago.

[Dawber, 1980] Dawber, T. (1980). *The Framingham Study. The Epidemiology of Atherosclerotic Disease.* Harvard U. Press.

[De Mulder et al., 2008] De Mulder, Y., Danezis, G., Batina, L., and Preneel, B. (2008). Identification via location-profiling in gsm networks. In *Proceedings of the 7th ACM work-*

shop on Privacy in the electronic society, WPES '08, pages 23–32, Alexandria, Virginia, USA.

[Deering, 2001] Deering, S. (2001). Watching the weight of the protocol hourglass. http://www.iab.org/wp-content/iab-uploads/2011/03/hourglass-london-ietf.pdf. In *IETF 51*.

[Dey et al., 2011] Dey, A. K., Wac, K., Ferreira, D., Tassini, K., Hong, J. H., and Ramos, J. (2011). Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th international conference on Ubiquitous computing*, UbiComp '11, pages 163–172, New York, NY, USA. ACM.

[Do et al., 2011] Do, T., Blom, J., and Gatica-Perez, D. (2011). Smartphone usage in the wild: A large-scale study of applications and context. In *Proc. ACM Int. Conf. on Multimodal Interaction (ICMI)*, Alicante.

[d'Onofrio, 2005] d'Onofrio, A. (2005). A general framework for modeling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences. *Physica D*, 208:220–235.

[Dunbar, 1993] Dunbar, R. I. M. (1993). Co-evolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences*, 16(4):681–735.

[Eagle, 2010] Eagle, N. (2010). *Mobile Phones as Social Sensors*, volume The Handbook of Emergent Technologies in Social Research, chapter 22. Oxford University Press.

[Eagle et al., 2010] Eagle, N., Macy, M., and Claxton, R. (2010). Network diversity and economic development. *Science*, 328(5981).

[Eagle and Pentland, 2006] Eagle, N. and Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, (10):255–268.

[Eagle et al., 2009] Eagle, N., Pentland, A. S., and Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proc. Natl. Academy of Sciences*, 106(36).

[Erickson et al., 2006] Erickson, G., Currie, P., Inouye, B., and Winn, A. (2006). Tyrannosaur life tables: An example of nonavian. *Science*, 313:213–217.

[Eston et al., 1998] Eston, R. G., Rowlands, A. V., and Ingledew, D. K. (1998). Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *Journal of Applied Physiology*, 84(1):362–371.

[Estrin and Sim, 2010] Estrin, D. and Sim, I. (2010). Open mhealth architecture: An engine for health care innovation. *Science*, 330(6005):759–760.

[Evans, 2011] Evans, B. (2011). Paco. http://code.google.com/p/paco/.

[Falaki et al., 2011] Falaki, H., Mahajan, R., and Estrin, D. (2011). Systemsens: a tool for monitoring usage in smartphone research deployments. In *Proceedings of the sixth international workshop on MobiArch*, MobiArch '11, pages 25–30, New York, NY, USA. ACM.

[Fall, 2003] Fall, K. (2003). A delay-tolerant network architecture for challenged internets. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, SIGCOMM '03, pages 27–34, New York, NY, USA. ACM.

[Farrahi and Perez, 2008] Farrahi, K. and Perez, D. G. (2008). What did you do today?: discovering daily routines from large-scale mobile data. In *MM '08: Proceeding of the 16th ACM international conference on Multimedia*, pages 849–852, New York, NY, USA. ACM.

[Festinger et al., 1950] Festinger, L., Schachter, S., and Back, K. (1950). *Social pressures in informal groups; a study of human factors in housing*. Harper, Oxford, England.

[Foster et al., 2010] Foster, D., Linehan, C., and Lawson, S. (2010). Motivating physical activity at work: using persuasive social media extensions for simple mobile devices. *Design*.

[Frijters et al., 2004] Frijters, P., Haisken-DeNew, J., and Shields, M. (2004). Money does matter! evidence from increasing real income and life satisfaction in east germany following reunification. *The American Economic Review*, 94(3):730–740.

[Gemmell et al., 2002] Gemmell, J., Bell, G., Lueder, R., Drucker, S., and Wong, C. (2002). Mylifebits: fulfilling the memex vision. MULTIMEDIA '02, New York, NY, USA.

[Gonzalez et al., 2008] Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196):779–782.

[Google, 2011] Google (2011). Android developer site. http://developer.android.com/index.html.

[Granovetter, 1973] Granovetter, M. (1973). The strength of weak ties. *ajs*, 78(6):1360.

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explorations*, 11(1).

[Hampton and Wellman, 2003] Hampton, K. and Wellman, B. (2003). Neighboring in netville: How the internet supports community and social capital in a wired suburb. *City Community*, 2(4):277–311.

[HASKELL, 1994] HASKELL, W. L. (1994). Health consequences of physical activity: understanding and challenges regarding dose-response. *Medicine & Science in Sports & Exercise*, 26(6).

[Hedley et al., 2004] Hedley et al., A. A. (2004). Prevalence of overweight and obesity among us children, adolescents, and adults, 1999-2002. *JAMA: The Journal of the American Medical Association*, 291(23).

[Hendelman et al., 2000] Hendelman, D., Miller, K., Baggett, C., Debold, E., and Freedson, P. (2000). Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc*, 32(9 Suppl):S442–9.

[Hidalgo and Rodriguez-Sickert, 2008] Hidalgo, C. A. and Rodriguez-Sickert, C. (2008). The dynamics of a mobile phone network. *Physica A: Statistical Mechanics and its Applications*, 387(12).

[Hill and Peters, 1998] Hill, J. O. and Peters, J. C. (1998). Environmental contributions to the obesity epidemic. *Science*, 280(5368).

[http://www.gnu.org/licenses/,] http://www.gnu.org/licenses/, G. L.

[John and Srivastava, 1999] John, O. P. and Srivastava, S. (1999). *The Big-Five trait taxonomy: History, measurement, and theoretical perspectives*, volume 2, pages 102–138. Guilford.

[Joki et al., 2007] Joki, A., Burke, J. A., and Estrin, D. (2007). Campaignr: A framework for participatory data collection on mobile phones.

[Kalmijn, 1998] Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*, 24:395–421.

[Killingsworth and Gilbert, 2010] Killingsworth, M. A. and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006):932.

[Kim et al., 2008] Kim, T., Chang, A., Holland, L., and Pentland, A. (2008). Meeting mediator: enhancing group collaborationusing sociometric feedback. In *CSCW*, pages 457–466.

[Krishnamurthy and Wills, 2009] Krishnamurthy, B. and Wills, C. E. (2009). On the leakage of personally identifiable information via online social networks. In *Proceedings of the 2nd ACM workshop on Online social networks*, WOSN '09, pages 7–12, New York, NY, USA. ACM.

[Krumme et al., 2010] Krumme, C., Cebrián, M., and Pentland, A. (2010). Patterns of individual shopping behavior. *CoRR*, abs/1008.2556.

[Lazer, 2009] Lazer, A. Pentland et al., D. (2009). Computational social science. *Science*, 323(5915):721–723.

[Lin et al., 2006] Lin, J. J., Mamykina, L., Lindtner, S., Delajoux, G., and Strub, H. B. (2006). Fish'n'steps: Encouraging physical activity with an interactive computer game. In *Proc. of the International Conf. on Ubiquitous Computing*.

[Lu et al., 2010] Lu et al., H. (2010). 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.

[Madan, 2010] Madan, A. (2010). Social evolution: Opinions and behaviors in face-to-face networks. Master's thesis, Massachusetts Institute of Technology.

[Madan et al., 2010a] Madan, A., Cebrian, M., Lazer, D., and Pentland, A. (2010a). Social sensing for epidemiological behavior change. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, pages 291–300, New York, NY, USA. ACM.

[Madan et al., 2011] Madan, A., Farrahi, K., Perez, D. G., and Pentland, A. (2011). Pervasive sensing to model political opinions in face-to-face networks. In *Pervasive'11 (in press)*.

[Madan et al., 2010b] Madan, A., Moturu, S., and Pentland, A. (2010b). Social sensing: Obesity, unhealthy eating and exercise in face-to-face networks. In *Wireless Health '10*.

[Maitland and Siek, 2009] Maitland, J. and Siek, K. A. (2009). Technological approaches to promoting physical activity. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction*, New York. ACM.

[Mann and Niedzviecki, 2001] Mann, S. and Niedzviecki, H. (2001). *Cyborg: Digital Destiny and Human Possibility in the Age of the Wearable Computer*. Randomhouse (Doubleday). ISBN: 0-385-65825-7.

[Marsden, 1990] Marsden, P. V. (1990). Network data and measurement.

[Martin Brunner and Todt, 2010] Martin Brunner, Hans Hofinger, C. K. C. R. P. S. and Todt, S. (2010). Infiltrating critical infrastructures with next-generation attacks. Technical report, Fraunhofer-Institute for Secure Information Technology SIT Munich.

[McPherson et al., 2001] McPherson, M., Smith-Lovin, L., and Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.

[Miklas et al., 2007] Miklas, A. G., Gollu, K. K., Chan, K. K. W., Saroiu, S., Gummadi, K. P., and De Lara, E. (2007). Exploiting social interactions in mobile systems. In *Proceedings of the 9th international conference on Ubiquitous computing*, UbiComp '07, pages 409–428, Berlin, Heidelberg. Springer-Verlag.

[Milo et al., 2002] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827.

[Mislove et al., 2010] Mislove, A., Viswanath, B., Gummadi, K. P., and Druschel, P. (2010). You are who you know: inferring user profiles in online social networks. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 251–260, New York, NY, USA. ACM.

[Montoliu and Gatica-Perez, 2010] Montoliu, R. and Gatica-Perez, D. (2010). Discovering human places of interest from multimodal mobile phone data. In *Proc of 9th Int. Conference on on Mobile and Ubiquitous Multimedia (MUM',',')*.

[Moturu et al., 2011a] Moturu, S. T., Khayal, I., Aharony, N., Pan, W., and Pentland, A. (2011a). Sleep, mood and sociability in a healthy population. In *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '11)*, Boston, MA.

[Moturu et al., 2011b] Moturu, S. T., Khayal, I., Aharony, N., Pan, W., and Pentland, A. S. (2011b). Using social sensing to understand the links between sleep, mood, and sociability. In *IEEE International Conference on Social Computing (SocialCom 2011)*, Boston, MA.

[Newman, 2003] Newman, M. (2003). The structure and function of complex networks. *SIAM Review*, 45:167–256.

[Newman and Park, 2003] Newman, M. E. J. and Park, J. (2003). Why social networks are different from other types of networks. *Phys. Rev. E*, 68:036122.

[Olguín et al., 2009] Olguín, D. O., Waber, B. N., Kim, T., Mohan, A., Ara, K., and Pentland, A. (2009). Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 39(1):43–55.

[Olguín Olguín, 2007] Olguín Olguín, D. (2007). Sociometric badges : wearable technology for measuring human behavior. Master's thesis, Massachusetts Institute of Technology.

[Oliver, 2009] Oliver, E. (2009). A survey of platforms for mobile networks research. *SIGMOBILE Mob. Comput. Commun. Rev.*, 12:56–63.

[Onnela and Reed-Tsochas, 2010] Onnela, J.-P. and Reed-Tsochas, F. (2010). Spontaneous emergence of social influence in online systems. *Proc. Natl. Academy of Sciences*, 107(43).

[Onnela et al., 2007] Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A.-L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336.

[Page, 2008] Page, S. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton Univ Pr.

[Palla et al., 2007] Palla, G., Barabasi, A.-L., and Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(7136):664–667.

[Pan et al., 2011] Pan, W., Aharony, N., and Pentland, A. (2011). Composite social network for predicting mobile apps installation. In *Proceedings of the 25th Conference on Artificial Intelligence (AAAI-11)*, San Francisco, CA.

[Paridon et al., 2006] Paridon, T., Carraher, S., and Carraher, S. (2006). The income effect in personal shopping value, consumer selfconfidence, and information sharing (word of mouth communication) research. *Academy of Marketing Studies Journal*, 10(2):107–124.

[Paulos and Goodman, 2004] Paulos, E. and Goodman, E. (2004). The familiar stranger: anxiety, comfort, and play in public places. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, pages 223–230, New York, NY, USA. ACM.

[Pong and Ju, 2000] Pong, S. and Ju, D. (2000). The effects of change in family structure and income on dropping out of middle and high school. *Journal of Family Issues*, 21(2):147.

[Putnam, 2000] Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon Schuster, New York.

[quantifiedself.com, 2011] quantifiedself.com (2011). The quantified self; http://quantifiedself.com/about/.

[Richards and Macindoe, 2010a] Richards, W. and Macindoe, O. (2010a). Characteristics of small social networks. Tech Report 2010-033, MIT-CSAIL, Cambridge, MA.

[Richards and Macindoe, 2010b] Richards, W. and Macindoe, O. (2010b). Decomposing social networks. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 114 –119.

[Rogers, 2003] Rogers, E. M. (2003). *Diffusion of Innovations*. Free Press, 5 edition.

[Rouvinen, 2006] Rouvinen, P. (2006). Diffusion of digital mobile telephony: Are developing countries different? *Telecommunications Policy*, 30(1):46 – 63.

[Rovio, 2009] Rovio (2009). Angry birds.

[Roy et al., 2006] Roy, D., Patel, R., DeCamp, P., Kubat, R., Fleischman, M., Roy, B., Mavridis, N., Tellex, S., Salata, A., Guinness, J., Levit, M., and Gorniak, P. (2006). The human speechome project. In Vogt, P., Sugita, Y., Tuci, E., and Nehaniv, C., editors, *Symbol Grounding and Beyond*, volume 4211 of *Lecture Notes in Computer Science*, pages 192–196. Springer Berlin / Heidelberg.

[Saponas et al., 2008] Saponas et al., S. T. (2008). ilearn on the iphone: Real-time human activity classification on commodity mobile phones. Tech report, University of Washington.

[Schneier, 1995] Schneier, B. (1995). *Applied cryptography (2nd ed.): protocols, algorithms, and source code in C*. John Wiley & Sons, Inc., New York, NY, USA.

[Solutionary, 2011] Solutionary (2011). White paper: The advanced persitent threat (apt).

[Starner and Rhodes, 2004] Starner, T. and Rhodes, B. (2004). *Wearable Computers*, volume Encyclopedia of Human-Computer. Berkshire Publishing.

[Toscos et al., 2008] Toscos et al., T. (2008). Encouraging physical activity in teens can technology help reduce barriers to physical activity in adolescent girls? In *Pervasive Computing Technologies for Healthcare*.

[Troiano et al., 2008] Troiano et al., R. P. (2008). Physical activity in the united states measured by accelerometer. *Med Sci Sports Exerc*, 40(1):181–8.

[Waber et al., 2010] Waber, B. N., Pollock, E., Cebrián, M., Crane, R., Danon, L., and Pentland, A. (2010). Modeling corporate epidemiology. *CoRR*, abs/1008.3402.

[Watts and Strogatz, 1998] Watts, D. and Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

[Whitney, 2011] Whitney, L. (2011). Google: 200 million android devices now active worldwide. news.cnet.com.

[Worksmartlabs, 2011] Worksmartlabs (2011). Cardiotrainer. http://www.worksmartlabs.com/cardiotrainer/about.php.

[Xie and Szymanski, 2011] Xie, J. and Szymanski, B. K. (2011). Community detection using a neighborhood strength driven label propagation algorithm. *CoRR*, abs/1105.3264.

[Yoneki et al., 2008] Yoneki, E., Hui, P., and Crowcroft, J. (2008). Wireless epidemic spread in dynamic human networks. *Bio-Inspired Computing and Communication*.