Social Evolution: Opinions and Behaviors in Face-to-Face Networks

by

Anmol Madan

S.M., Massachusetts Institute of Technology (2005)
B.E., University of Pune, India (2003)

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

September 2010

Author ____________________
Program in Media Arts and Sciences
September, 2010

Certified by ____________________
Prof. Alex (Sandy) Pentland
Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by ____________________
Prof. Pattie Maes
Associate Academic Head
Program in Media Arts and Sciences

Social Evolution: Opinions and Behaviors in Face-to-Face Networks

by

Anmol Madan

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on September, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences

Abstract

Exposure to new ideas and opinions, and their diffusion within social networks, are important questions in education, business, and government. However until recently there has been no method to automatically capture fine-grained face-to-face interactions between people, to better model the diffusion process. In this thesis, we describe the use of co-location and communication sensors in 'socially aware' mobile phones to model the spread of opinions and behaviors of 78 residents of an undergraduate residence hall for an entire academic year, based on over 320,000 hours of behavior data.

Political scientists (Huckfeldt and Sprague, APSR, 1983) have noted the problem of mutual causation between face-to-face networks and political opinions. During the last three months of the 2008 US presidential campaigns of Barack Obama and John McCain, we find that political discussants have characteristic interaction patterns that can be used to recover the self-reported 'political discussant' ties within the community. Automatically measured mobile phone features allow us to estimate exposure to different types of opinions in this community. We propose a measure of 'dynamic homophily' which reveals surprising short-term, population-wide behavior changes around external political events such as election debates and Election Day. To our knowledge, this is the first time such dynamic homophily effects have been measured. We find that social exposure to peers in the network predicts individual future opinions ($R^2 \approx 0.8, p < 0.001$). The use of mobile phone based dynamic exposure increases the explained variance for future political opinions by up to 30%.

It is well known that face-to-face networks are the main vehicle for airborne contagious diseases (Elliott, Spatial Epidemiology, 2000). However, epidemiologists have not had access to tools to quantitatively measure the likelihood of contagion, as a function of contact/exposure with infected individuals, in realistic scenarios (Musher, NEJM, 2003), since it requires data about both symptoms and social interactions between individuals. We use of co-location and communication sensors to understand the role of face-to-face interactions in the contagion process. We find that there are characteristic changes in behavior when individuals become sick, reflected in features like total communication, temporal structure in communication (e.g., late nights and weekends), interaction diversity, and movement entropy (both

within and outside the university). These behavior variations can be used to infer the likelihood of an individual being symptomatic, based on their network interactions alone, without the use of health-reports. We use a recently-developed signal processing approach (Nolte, Nature, 2008) to better understand the temporal information flux between physical symptoms (i.e., common colds, influenza), measured behavior variations and mental health symptoms (i.e., stress and early depression).

Longitudinal studies indicate that health-related behaviors from obesity (Christakis and Fowler, 2007) to happiness (Fowler and Christakis, 2008) may spread through social ties. The effects of social networks and social support on physical health are well-documented (Berkman, 1994; Marmot and Wilkinson, 2006). However, these studies do not quantify actual face-to-face interactions that lead to the adoption of health-related behaviors. We study the variations in BMI, weight (in lbs), unhealthy eating habits, diet and exercise, and find that social exposure measured using mobile phones is a better predictor of BMI change over a semester, than self-report data, in stark contrast to previous work.

From a smaller pilot study of social exposure in face-to-face networks and the propagation of viral music, we find that phone communication and location features predict the sharing of music between people, and also identify social ties that are 'close friends' or 'casual acquaintances'. These interaction and music sharing features can be used to model latent influences between participants in the music sharing process.

Thesis Supervisor: Prof. Alex (Sandy) Pentland
Title: Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Social Evolution: Opinions and Behaviors in Face-to-Face Networks

by

Anmol Madan

The following served as a reader for this thesis:

Thesis Reader ______________________________

Prof. Tanzeem Choudhury
Professor of Computer Science
Dartmouth College

Social Evolution: Opinions and Behaviors in Face-to-Face Networks

by

Anmol Madan

The following served as a reader for this thesis:

Thesis Reader______________________________

Prof. David Lazer

Professor of Political Science and Computer Science, Northeastern University

Director, Program on Networked Governance, Harvard University

Acknowledgements

I would like to thank my thesis committee members for their invaluable support. Sandy Pentland motivated, and advised me to take bold steps in this scientific realm. Tanzeem provided valuable technical insight and discussions, and David guided me with the social and human perspectives of my work. I would also like to thank Devon Brewer, for his insights with study design.

I would like to acknowledge my wonderful collaborators at every stage. Various undergraduates played an invaluable role in the experimental data collection process. I would especially like to mention the efforts of Nevada Sanchez, Rita Chen, Thatcher Clay, Michael Snively for the first wave, and those of Iolanthe Chronis, Chris Palmer, Xiao Fan, Mu He, Gabe Tobon, Paul Kominers and Erons Ohienmhen for the second, larger wave. The influenza and epidemiology work was done with the help of Manuel Cebrian, and the social health results were developed with the help of Sai Motoru. The privacy review work was a collaboration with Ben Waber, Margaret Ding and Paul Kominers. I would also like to thank Kate Farrahi, Sune Lehmann and others for their feedback. It has been a privilege to work amongst the Human Dynamics members, viz., Ankur Mani, Ben Waber, Coco Krumme, Cory Ip, Daniel Olguin Olguin, Galen Picard, Iolanthe Chronis, Joost Bonsen, Manuel Cebrian, Nadav Aharony, Riley Crane, Sai Moturu, Taemie Kim, Wei Pan and Wen Dong. I look forward to future collaborations with all of you.

A big shout-out goes to my friends and roomates for making life delightful, whether in Cambridge, Boston, Pune or Seattle. Aaron and daycrew, Peggy and nightcrew, thanks for making student life easier. I'm indebted to my father, sister, May and amazing extended family in Delhi and Toronto, for sharing this fascinating journey.

Dedicated to my father, and sister Aditi, for their incredible patience and support!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

As citizens of the information age, we leave pervasive digital traces of our idiosyncratic behavior in emails, social networking sites, mobile phone call records, automatic teller machine (ATM) transactions, and metropolitan train systems. As researchers in the emerging field of Computational Social Science (CSS), our interests lie in building quantitative models of large-scale human social systems from these digital traces, and answering key supporting questions related to privacy, data-ownership, user-feedback and behavior manipulation.

The diffusion of information, ideas, opinions and media in a social network and the influence of individual nodes on the diffusion process are important questions in the social sciences. In this thesis, we use mobile phone sensors to model the spread of opinions and behaviors in face-to-face networks, in the social contexts of political opinions, health and weight management, and epidemiology.

The contributions towards this goal, discussed in detail in the following chapters, are as follows:

- Devise an experimental methodology, and engineer a mobile platform to capture face-to-face interactions and the adoption of various behaviors for a tight-knit social community

- In the context of political opinions, this approach captures underlying micro-variations in homophily, and can be used to identify political discussants and explain future opinion changes for individuals
- In the context of spatial epidemiology, this approach captures behavior variations that allows us to identify symptomatic individuals, and provides insight into the mechanism that links physical symptoms, behavior change and mental health symptoms
- In the context of social health, this approach shows how exposure to different aspects of unhealty diet and lifestyle can predict future BMI and weight changes for individuals embedded in the network
- In the context of social media, our approach helps understand the evolution of friendship ties and viral music propagation

1.1 The Role of Face-to-Face Networks in Social Diffusion

Why is it important to study the propagation of ideas and behaviors in social networks? Models of the diffusion of ideas, innovations, recommendations and media have been extensively studied in different fields of social science literature, including but not limited to management science (i.e., diffusion of innovations and viral marketing), political science (i.e., political influence on voting behavior), public health policy (i.e., social influence and support mechanisms for health and wellness) and spatial epidemiolgy (i.e., modeling the spread of contagious diseases and epidemics). In this section, we discuss selected prominent works across these different social sciences, that are relevant for this thesis.

In management science, the idea of diffusion of innovations gained popularity with the work of Rogers [105], who identified distinct classes of innovators, early adopters, mainstream majority, and laggards in the adoption of ideas. The network interactions between these classes were used to explain the adoption 'S-curves' of new products and opinions.

More recently, in the workplace context, Aral and Van Alystne [4] compared corporate email communication with objective performance, experience and demographic data about the employees of a hiring firm. They found that complex discussion topics have a very different underlying diffusion mechanism than simple declarative event news and factual information. Strong ties, equivalence of tenure and path length were significant predictors of the diffusion of discussion topics, while these factors had little effect on the propagation of factual news and information though the organization. This access to novel information was a better predictor of productivity than traditional human capital variables and was significantly related to additional revenue generated by the person, by approximately $70 per 'additional word seen', in their dataset.

In follow-up work, Aral and Van Alstyne [5] studied the role of network diversity and channel bandwidth, in access to novel information. They found that structural diversity was correlated with better performance– in economic terms, one standard deviation increase in non-redundant information was associated with $4600 of additional revenue, a ten percent improvement on projects completed, and a project duration reduction of fifteen days per person per month. In thier dataset, a diverse network had more channels, but most of these channels were low-bandwidth, weak ties.

In political science, studies have proposed two competing models of social influence and contagion [22]. The social cohesion model suggests that influence is proportional to tie strength, while the structural equivalence model [48] proposes that influences exist across individuals with similar roles and positions in networks. In both cases, social scientists agree that measuring flows of political influence within a social network is complex because (a) flows of information and influence are bi-directionally causal, and it is hard to isolate one-directional effects, i.e., from ego to alter (b) it is hard to accurately discriminate between change in political opinion due to social influence, and change due to common external social attributes and events and (c) it is hard to distinguish between short-term and long-term influence effects.

Huckfeldt and Sprague [55] studied the interdependence of an individual's political opinions, their political discussant network and geo-political context and demographics during the 1984 presidential elections. They found a social dissonance effect in the propagation of political opinions, i.e. voters showed a remarkably high accuracy in their perception of discussant's opinons where the discussants held similar opinions (approximately 90 percent accuracy), but for the discussants that held different political preferences than the main respondent, this perceptive ability was only about 55 percent aaccurate.

With regard to network structure and influence, Huckfeldt and Sprague found an 'inverse U' relationship with tie-strength, i.e. discussant effects are stronger for less intimate relationships like acquaintances and frequent contacts than they are for close friends. In thier words, this result is explained as, "individuals choose a close friend or spouse based on other important social factors (common interests, employment etc.) but find people outside strong ties that they resonate with in terms of political dialogue and opinions".

In public health literature, the hypothesis that social support is a key factor of individual health and wellness is broadly accepted [75, 60]. Recent work suggests that face-to-face networks are also an important vehicle for the propagation of healthy behaviors and combating unhealthy behaviors in public health and preventive social healthcare. Christakis and Fowler [27] studied the spread of obesity over 32 years of the Framingham heart study dataset. They found that an individual's likelihood of becoming obese increased substantially if a friend, sibling or spouse became obese. Using time-lagged variables and longitudinal analysis they determine that the observed weight changes represent induction, although the statistical validity of this finding has been recently questioned [33]. Christakis and Fowler have reported similar effects for the cessation of smoking, where it was found that the likelihood of cessation increased substantially for an individual if a spouse, sibling or friend gave up smoking, and clusters of participants gave up smoking [28], and also for the spread of happiness. In all these cases, the propagation of behaviors was attributed to social contagion on the basis of longitudinal analysis.

Apart from tie nature and strength, the structural properties of the underlying network have been shown to be important in accelerating or limiting the diffusion process, measured using agent-based simulations and emperical behavioral experiments in sociology. Mason and Goldstone [76] found that different network topologies affect system-wide performance for complex tasks (e.g., adoption of a new innovation with a community), as well as the balance between exploration and exploitation. For a complex problem-solving task with many local maxima but one global maximum, a small-world network structure leads to fastest solution, as participants are forced to invest in exploring the entire problem space due to the interaction limitations imposed by the small world network. Lazar and Friedman [65] reported similar results in an agent-based simulation of complex problem solving task within a contrained network environment. Hence, constraining information access and diffusion within a network may improve group performance for complex tasks.

The cascading effects of social diffusion also play a role in herding behavior in economics [8] and speculative bubbles in financial markets [108] .

1.1.1 The Problem of Informant Inaccuracy

So while the underlying social questions are important, why has it been hard to answer them in the past? The limitations stem from the tools used to collect data, and the associated problem of informant inaccuracy. Social scientists have traditionally relied on surveys and self-report measurements to model diffusion and influence phenomena, but the accuracy of self-reported face-to-face interaction data is relatively poor.

In a broad review of self-report literature, Bernard et. al. [16] found that about half of what informants report is inaccurate in some dimension. For example, in health care studies, around 25 percent to 50 percent of health related issues were unreported if they were over 9-12 months old. In social interaction studies, recall of interactions with specific individuals (close friends, people most time spent with, etc.) had approximately 50 percent accuracy. Brewer and Webster [19] measured the recall of friends amongst college students, who lived in a common university residence hall. During interviews with students, subjects were first asked to recall friends (unassisted) with the residence hall, and then asked to recognize friends from the complete list of residents. On average, subjects forgot about 20 percent of their friends.

This limited ability to recall is due to time omission (i.e., memory of events and actions decays with time) and telescoping effects (i.e., individuals tend to under-estimate the time dimension). In some contexts, poor recall is a function of how individuals like to be perceived, e.g., under reporting sick leave, unemployment levels, drug-use or missed voting opportunities. In other social situations, poor recall is due to our selective memories, as when efficiently building memories, humans rarely remember precise details of interactions. Instead, interactions and information about external events are encoded in relevant cognitive interpretations, e.g., a water-cooler conversation. Sociologists have used 'conceptual variables' [16] to capture beliefs and attitutes as supported by our memory structures, but such variables have not demonstrated predictive power. Demographic variables have been found to be uncorrelated with the ability to recall [19].

1.1.2 The Causality Problem

To manipulate future outcomes, it is important to understand the underlying causal mechanism, especially if the phenomenon involves complex interdependencies in social variables and manipulations are in the form of interventions or public policy decisions. However, conclusive causal explanation of behavioral change due to social induction in social networks is a remarkably hard problem. An important area of work has been around trying to conclusively disambiguate social induction from homophily and common confounding factors in networks. Associative mixing [87], temporal clustering and longitudinal analysis [27] are a few of the approaches that have been used to explain social induction as the underlying causal mechanism of adoption in networks.

An early example of misleading conclusion of social induction was Coleman's study [34] where the rapid adoption of Tetracycline was originally attributed to the social influence of early-adopter practitioners, considered thought leaders in the field. However, follow-on analysis almost forty years later [38] found that the adoption was entirely explained by the product characteristics and aggressive marketing techniques used by the drug company salesmen. Cohen-Cole and Fletcher [33] have questioned the statistical validity of the longitudinal regression analysis used by Christakis and Fowler [27] on the Framingham Heart studies. Applying similar longitudinal analysis to the National Add Health dataset indicates that acne, height and headaches are contagious (known to be impossible), which implies confounding in the original result.

On the other hand, several promising statistical and experimental works have resulted in more conclusive causal results. For voter behavior, Nickerson [88] used a placebo-treatment experiment design to isolate the influence of a spouse from the effects of common social attributes on voting behavior. Similar voting behavior between married couples is often attributed to homophily– shared values, neighborhood, education, children, exposure to mass media and local events. These contextual variables can be easily confused with actual influence that a person exerts on their spouse to vote for a candidate. In Nickerson's experiment, one-person in a two-person household was targeted by a controlled exogenous shock and the indirect boost in turnout on their spouse was compared to a placebo group that recieved a non-political shock, and a control group that received no shock. The treatment group transmitted roughly 60 percent of their increased propensity to vote to their untreated spouse. Salganik and Watts [107] used a replicated worlds experiment design to measure peer influence in the popularity of music tracks with internet users, and found unpredictable rankings of music tracks in each world, unrelated to the actual quality of music, but determined by the cascading effect of social influence.

In support of statistical methods used to disambiguate between causal mechanisms of adoption, Aral et. al [6] used a dynamic matched sample estimation framework to compare the relative role of homophily and social influence in the adoption of an online product across a global instant messaging network, and found that homophily explains more than 50% of perviously-assumed behavior contagion amongst network ties. The Phase Slope Index (PSI) is a recently proposed statistical method based on the use of cross-spectra of time series data as a noise-resistent measure of information flux between signals. This method is used to better understand the temporal relationship between physical symptoms, mental symptoms and measured behavior variations in Chapter 4.

1.2 Sensor-based Computational Social Science

Social interactions in the real world present a great avenue for understanding user behavior. Long term monitoring has been implemented using a variety of technologies including video [106], smartphones [41, 42, 73, 82, 1], wearable sensing devices [100, 63, 92, 91, 26], and web-based social media and electronic data [18, 110, 118, 69, 70, 3, 85].

The movement to model face-to-face interactions in social networks using sensing technologies, began with the work of Choudhury and Pentland [23], who devised the 'Sociometer' to understand the network structure, and detect when people were in conversation. Choudhury and Basu [25] modeled turn-taking behavior in face-to-face conversations, and found that the 'influence' of each participant in joint turn-taking, was correlated with betweenness centrality of the person, estimated from the social network.

There are four billion mobile phones worldwide, which can be used as ubiquitous sensors of location, proximity and communication. Eagle, Dong and Pentland [41, 39] coined the term Reality Mining, and used mobile phone sensors to learn network structure in the MIT community. Gonzalez et. al. [51] have shown that call detail records can be used to characterize temporal and spatial regularity in human mobility patterns better than random walk or Levy flight simulations. Other examples of the use of mobile phones to map human interaction networks include the CENS participatory sensing project at UCLA [1], and the mHealth and Darwin projects at Dartmouth [7, 82].

The use of sensors and electronic badges to understand social network evolution and face-to-face interactions has also evolved, since electronic badges often incorporate infrared(IR) sensors that capture the direction of face-to-face proximity. Olguin, Waber and Kim [92, 93, 91] have shown the importance of measurement and feedback using sociometric badges in several contexts, viz., understanding the link between communication patterns and task completion at a data-center, improving the productivity of call center employees by modifying their break structure, and improving the performance of groups in co-located and distributed meetings, by providing individuals feedback on thier personal communication style. Using similar sensor hardware, Wyatt et. al. [116] found a link between speaking style and social tie strength, suggesting that people change their conversation patterns depending on the type/strength of the relationship with the other person. Similar to [25], they also found a correlation between individual speaking style, and network (closeness) centrality. These result suggest that micro-interaction style may be a function of the role or influence of the participating individuals.

Communication patterns can be gathered through phone calls, text messages, e-mail, instant messaging and web-based social networks. Vocal analysis has been used to capture nonlinguistic communication and social signaling in different contexts [24, 72, 92, 100, 93].

1.3 Stochastic Models of Social Networks

A social network [112] is a mapping of social ties on a network graph, with actors (or nodes) connected by relations (or ties). Clustering methods have been proposed to identify different sets of nodes in the network structure, usually based on a distance metric computed from network structure [54, 83] Clustering is related to community structure in networks, and various methods have been devised to identify boundaries of these existant communities in network data [50, 10, 98, 62].

For modeling networks, Exponential Random Graph Models (ERGMs) are a popular choice because they can be used to answer simple as well as complex network questions, by defining the probability of an observed network as a function of different structural characteristics such as density, reciprocity, or cliquing [104]. Various extensions have been proposed to overcome limitations of basic ERMSs, e.g., hidden and latent ERGMs [53] to overcome thier inability to represent temporally-varying networks, and curved-exponential families [56] to reduce degeneracy at the cost of more complicated features. Classic ERGMs assume that ties/edges are binary (i.e., equal weights) and time homogenous, but this is not true of real-world sensor data or the underlying social ties. Wyatt et. al. [115] have extended ERGMs to incorporates multi-dimensional social ties, and a curved exponential family model to model the temporal inhomogeniety. When this representation was applied to sensor-based face-to-face behavioral data, they found that samples drawn from the fitted model explained emperical link density and path length distributions better than time-homogeneous ERGMs.

Markov Logic Networks [103] combine probabilistic modeling with first-order logic, and exponential formula weights represent the likelihood of the logic fomula being true in the observed dataset. Bayesian models like Latent Dirchlet Allocation [17], Author-Topic [109, 77] and Group-Topic [111] models have been applied to social network analysis when content or contextual data, that can be deconstructed into topics, is present.

Compartment models like the Susceptible, Infectious, Recovered (SIR) model and its variants are used to model population dynamics and epidemic curves using differential equations in epidemiology. Recent compartmental approaches have taken into account mobility variations, but only in a heuristic way and not at the individual level [20, 35, 43].

1.4 Discussion

In this chapter, we outlined the need for the subsequent contributions of this thesis, and provided background for various aspects of our work. Modeling how people adopt ideas and behaviors in face-to-face networks is an important question in social sciences. Progress has been limited, however, by problems of informant inaccuracy and challenges of disambiguating causality. A movement has been forming around the use of sensors and smartphones as media to capture face-to-face interactions. Meanwhile, a variety of statistical and Bayesian modeling tools are available to model social network interactions.

Chapter 2

Methodology

How can we generate fine-grained interaction data required to model face-to-face interactions in social contexts? Other researchers have used large-scale consumer data provided by mobile network operators to understand human behavior [51]. The approach used in this thesis was to devise a 'community microscope' to help understand how people interact continiously, over a long period.

This approach has several advantages: (a) it is the only way to probe participants for different self-reported opinions and behaviors, data which are unavailable from mobile operators (b) it make it possible to capture proximity, co-location and other interaction modalities beyond Call Data records (CDRs) provided by mobile operators (c) it allows participation of a complete tight-knit set of participants from a single community, across four different mobile-operators in the study.

On the other hand, deploying an experimental platform within a community is limited by scale, cost and effort.

2.1 Experiment Design: A Community Microscope

The experiment was designed as a long-term longitudinal study with seventy residents of an undergraduate residence hall (referred to as an undergraduate dormitory in North America), that serves as the primary residential, cooking, social activity and sleeping quarters for the residents. This residence hall was the smallest undergraduate dormitory at the university. The participants in the study represent eighty-percent of the total population of this hall, and most of the remaining twenty-percent are spatially isolated. The dormitory is known within the university for its pro-technology orientation and the decision of students to reside within the dorm is determined via self-selection, by both incoming students and incumbent residents. The students were distributed roughly equally across all four academic years (freshmen, sophomores, juniors, seniors), about 54% of the students were male, and predominantly engineering, mathematics and science majors. The study participants also included four graduate resident tutors that supervised each floor. The participants used data collection Windows Mobile devices as their primary phones, with their existing voice plans. Students had data access on these phones due to pervasive wireless (wifi) on the university campus and in the surrounding metropolitan area.

The overarching experiment was designed to study the adoption of political opinions, diet, exercise, obesity, eating habits, epidemiological contagion, depression and stress, dorm political issues, interpersonal relationships and privacy. A total of 320,000 hours of human behavior data was collected in this experiment. In addition to mobile phone sensors, participants completed daily, weekly, and monthly respondent surveys related to these dependent variables. The experiment deployment timeline is described in Table 2.1. Additional information about the weekly and monthly surveys is available here [73].

2.1.1 User Privacy Considerations

A key concern with such long-term user data collection approaches is securing personal privacy for participants. This study was approved by the Institutional Review Board (IRB). As financial compensation for completing monthly surveys and using data-collection devices as

Table 2.1: Experiment Deployment Timeline

iPhone pilot deployment with two other dormitories	Spring 2008
Engineering the Windows Mobile data collection platform	Summer 2008
First-half of the experiment, which included the US presidential elections	Fall 2008
Second-half of the experiment, which included the daily physical symptoms sub-experiment	Spring 2009
Post-processing and data-cleanup	Summer and Fall 2009

their primary phones, participants were allowed to keep the devices at the end of the study. The sensing scripts used in the platform capture only hashed identifiers, and collected data is secured and anonymized before being used for aggregate analysis. To minimize missing data from daily symptom reports for the influenza-related analysis, participants were compensated $1 per day that they completed the on-device symptom survey. Additional privacy-related discussion and participant quotes are available in Chapter 7.

2.2 Mobile Sensing Platform

With the above goals in consideration, the mobile phone based platform for data-collection was designed with the following long-term continuous sensing capabilities, based on Windows Mobile 6.x smartphones. Daily captured mobile sensing data was stored on-device on read/write SD Card memory. On the server side, these logs files were merged, parsed and synced by an extensive Python post-processing infrastructure, and stored in MySQL for analysis. This sensing software platform for Windows Mobile 6.x has been released under the LGPLv3 open source license for public use [84].

The mobile phone based platform for data-collection had the following sensing capabilities, and the platform architecture is shown in Figure 2-1.

2.2.1 Device Selection

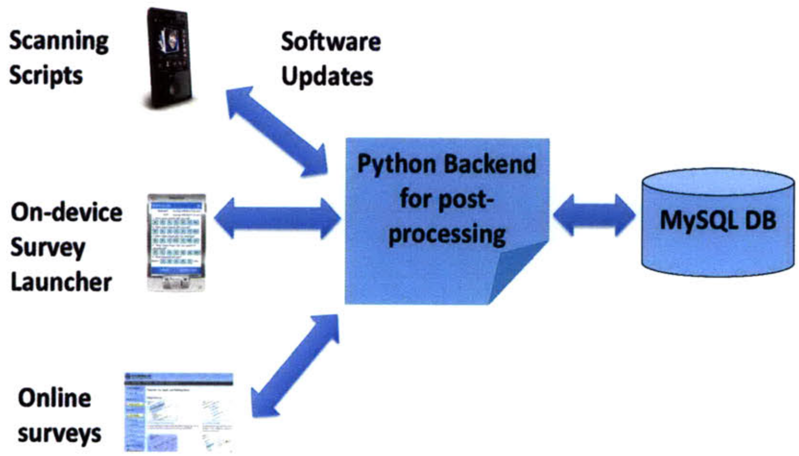
The platform was based on Windows Mobile 6.x devices, since they could be deployed with all four major American operators. Software was written using a combination of native-C and managed-C#. The software-sensing package was supported for six different handset models in the Windows Mobile product range. All supported devices featured WLAN, EDGE and SD Card storage, and most featured touch screens and flip-out keyboards. The HTC Tilt, a popular GSM phone in our experiment is shown in Figure 2-1.

2.2.2 Proximity Detection (Bluetooth)

The software scanned for Bluetooth wireless devices in proximity every 6 minutes (a compromise between sensing short-term social interactions and battery life, refer [42]). The Windows Mobile phones used in our experiment were equipped with class 2 Bluetooth radio transceivers, with practical indoor sensing range of approximately 10 feet. Scan results for two devices in proximity have a high likelihood of being asymmetric, which is accounted for in analysis. Due to API limitations of Windows Mobile 6.x, signal strength was not available during scans. The format for captured Bluetooth logs is shown in in Table 2.2.

2.2.3 Approximate Location (802.11 WLAN)

The software scanned for wireless WLAN 802.11 Access Point identifiers (hereafter referred to as WLAN APs) every 6 minutes. WLAN APs have an indoor range of approximately 125 feet and the university campus had almost complete wireless coverage. Across various rooms and common areas within the undergraduate residence, over 55 different WLAN APs with varying signal strengths can be detected. The format for logged WLAN APs is shown in Table 2.2.



(a) Platform Architecture and Data Sources

(b) HTC Tilt: the most popular WiMo device in our deployment

(c) On-device survey launcher screenshot

Figure 2-1: Data Collection Platform

2.2.4 Communication (Call and SMS Records)

The software logged Call and SMS details on the device every 20 minutes, based on recent events. These logs included information about missed calls and calls not completed. The format for logged call records and SMS messages is shown in Table 2.2.

Table 2.2: Logged Data Formats

Bluetooth	UTC timestamp remote device MAC hash
WLAN	UTC timestamp AP MAC hash AP ESSID Signal Strength 0-100
Calls	UTC start timestamp UTC end timestamp remote phone number hash incoming vs. outgoing flag missed call flag roaming flag
SMS	remote phone number hash incoming/outgoing

2.2.5 Battery Impact

The battery impact of long-term mobile sensing has been previously discussed [42]. In our case, periodic scanning of Bluetooth and WLAN APs reduced operational battery life by about 10-15 percent. Depending on the device models and individual usage patterns, the average usable battery life was between 14-24 hours. Windows Mobile 6.x phones have relatively poorer battery performance to begin with, and WLAN usage for web browsing by the user had significantly more impact on battery life than our sensing scripts. (e.g., using 802.11 WLAN for web browsing for 4-5 hours continuously on some phone models drained batteries completely). If available, users were provided with extended batteries for their phones. Over-the-air data uploads to the server were disabled for most of the experimental deployment due to WLAN battery considerations.

2.2.6 Post-Processing and Database

Daily captured mobile sensing data was stored on-device on read/write SD Card memory. On the server side, these logs files were merged, parsed and synced by an extensive Python post-processing infrastructure, and stored in a MySQL database for analysis.

This sensing software platform for Windows Mobile 6.x has been released under the LGPLv3 open source license for public use, and is available for download here [84].

2.3 The Dataset

Between September 2008 and June 2010, the dataset consists of 3.15 million scanned bluetooth devices, 3.63 million scanned WLAN access-points, 61,100 logged call data records, and 47,700 logged SMS messages. Of these, 2.08 million scanned bluetooth devices belong to other experiment participants, and 11,289 calls and 9533 SMS messages are exchanged with other experiment participants.

Various interesting aspects of the dataset are illustrated in the remaining figures in this chapter. Figure 2-2 and Figure 2-4 show the number of sample counts for Bluetooth, Calls and SMSs for each day of the Fall and Spring semester respectively. A slight reduction in total daily counts is seen in early December 2008, which was also end-of-semester. Similarly, Jan 2009 is the independent activities period, when many participants were away on interships and travel.

For both semesters, there is weekly temporal structure in the dataset, evident in Figure 2-3 and Figure 2-6. In the Fall 2008 months, there is a peak in SMS communication for the few hours on Friday evenings, and a corresponding dip in Bluetooth proximity during the same period. A similar but weaker structure with SMS messaging is seen in the Spring semester.

The daily temporal structure for the academic year reflects the peak activity times for the residents, shown in Figure 2-7. Participants tend to sleep very early in the mornings, rather late by the standards of working adults.

While network graphs are often used to study social networks, sensor data is multi-dimensions and dynamic, and these complex relationships between nodes cannot be easily summarized in a single plot. Figure 2-7 illustrates this property, with four different networks plotted for all participants. Two of these networks are based on weekly bluetooth physical proximity, a third is based on phone calling patterns, and the fourth is from self-reported close friend ties. It is seen that each network captures a different 'slice' of community interaction, and ties are not maintained across different modalities. This is further evident in the individual centrality measures for participant nodes, shown in Figure 2-8. Participant centality varies dramatically across different interaction modalities. For example, while some participants appear central in the self-report network, their social importance is not reflected in the bluetooth proximity or calling networks.

Figures 2-9 – 2-13 provide more information about the self-report networks for the same period. The 'close friends' network from six monthly sociometric surveys is compared to the 'willing to share Facebook photos and Tweets' network during the same period. The average number of close friends per individual ranges from 7-11 during the study period (increasing with time), and the average number of 'willing to share share Facebook photos and tweets' ties is about 3–4 times the number of close friends, during the same period. More information about the network measures plotted in Figure 2-10 and 2-11 is available here [112].

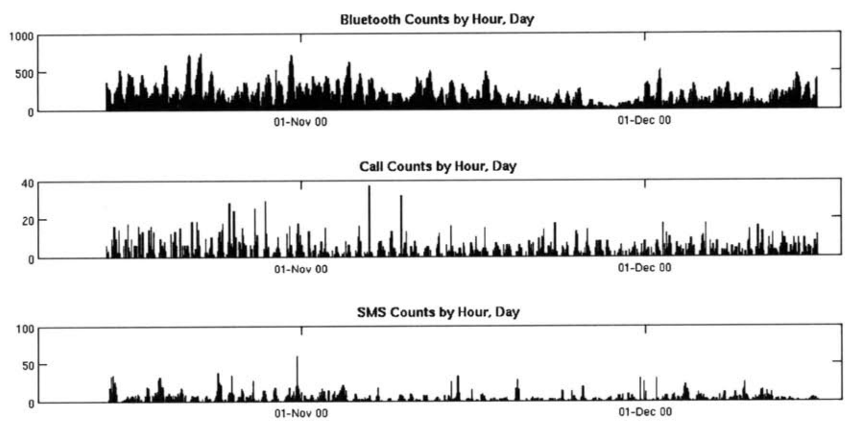


Figure 2-2: Total sample counts for Bluetooth (top), Calls (middle) and SMS (bottom) for the Fall 2008 semester. End-of-semester exams were in the first half of December 2008.

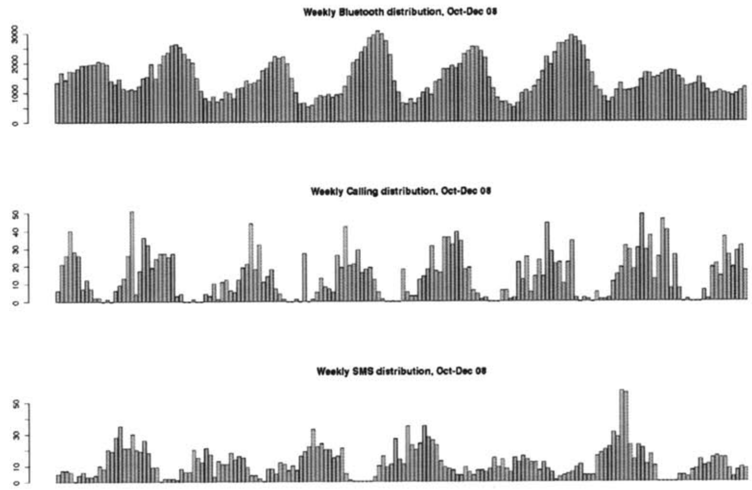


Figure 2-3: Weekly distribution of Bluetooth (top), Calls (middle) and SMS (bottom) for the Fall 2008 semester. Note the peak on Friday evenings for SMS and the corresponding drop in Bluetooth counts, lasting for a few hours

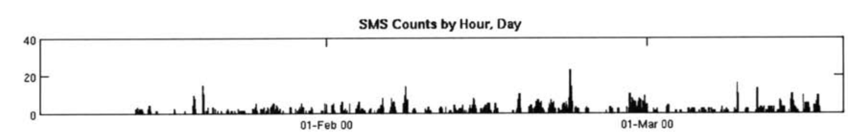


Figure 2-4: Total sample counts for Bluetooth (top), Calls (middle) and SMS (bottom) for the Spring 2009 semester. January 2009 was an independent activities period, and the semester started in February.

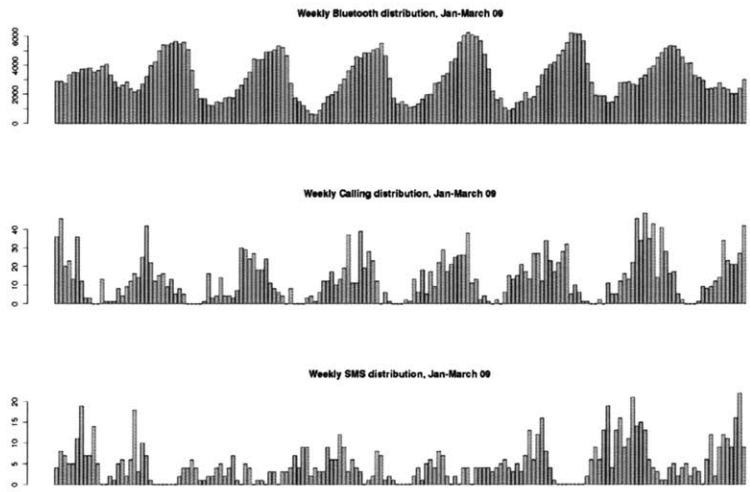


Figure 2-5: Weekly distribution of Bluetooth (top), Calls (middle) and SMS (bottom) for the Spring 2009 semester.

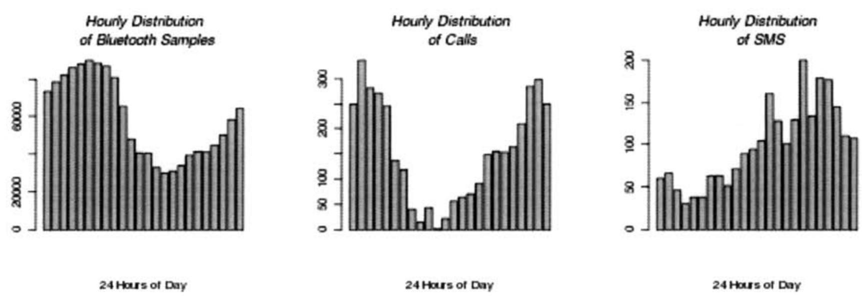


Figure 2-6: Hourly distribution of counts for Bluetooth (left), Call (middle) and SMS (right) for the entire academic year. The number of calls drops sharply in the early morning hours, which reflects the late bedtimes for the undergraduate students.

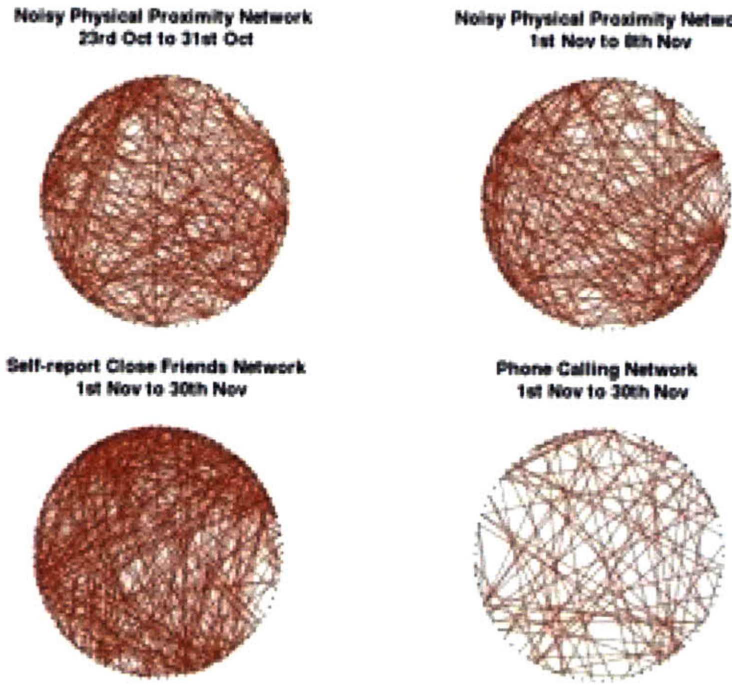
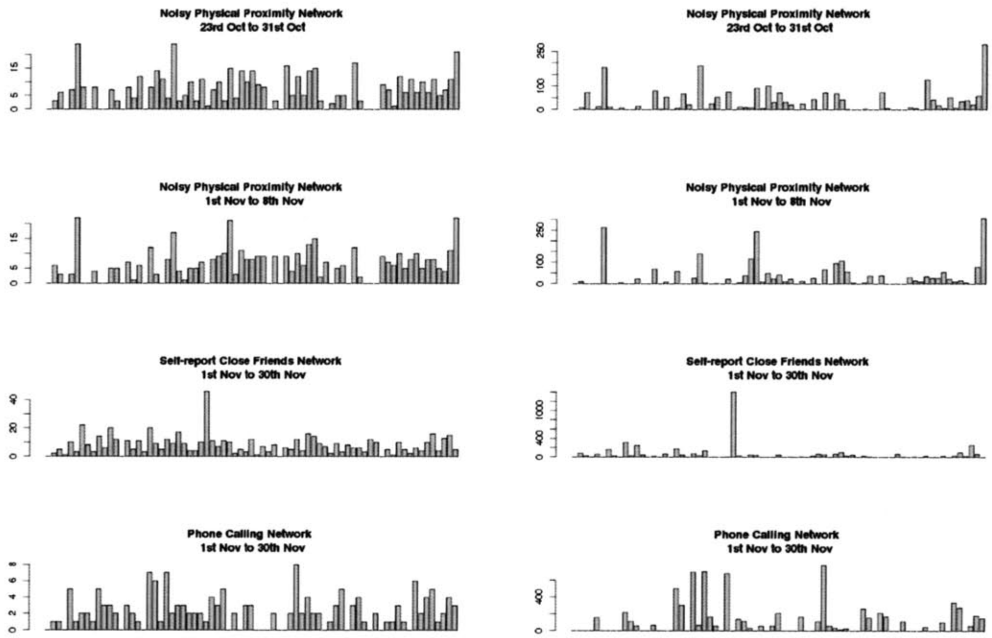


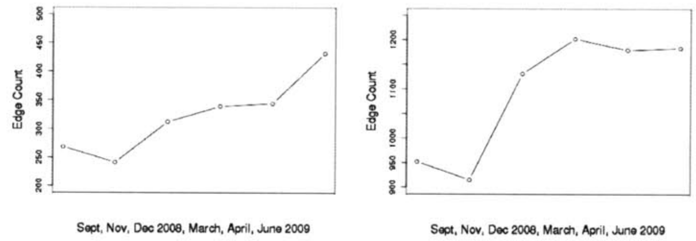
Figure 2-7: Four different interaction networks plotted for all participants. Node locations are invariant. Top left: Bluetooth physical proximity network, week of 23rd-31st October 2008. Top right: Bluetooth physical proximity network for the week of 1st-8th November 2008. Bottom left: Self-report close friends network, based on the survey completed in the first week of November. Bottom right: Phone calling network from 1st-30th November. Each modality captures a different 'slice' of social interactions.



(a) Degree centrality

(b) Betweenness centrality

Figure 2-8: Network centrality measures for individual nodes for the four different networks shown in Figure 2-7. The relative importance of a node in the community varies with the interaction modality. For example, nodes that seem central in the self-report network don't have the similar positions in bluetooth proximity and phone calling networks.



(a) 'Close friend' ties (b) 'willing to share Facebook photos and Tweets' ties

Figure 2-9: The number of close friend ties, compared to the number of 'willing to share Facebook photos and Tweets' ties for each of the six monthly sociometric surveys. The number of ties increase from Sept 2008 to June 2009. Overall, there are about 3–4 times as many ties for sharing Facebook/Twitter access, than for close friends.

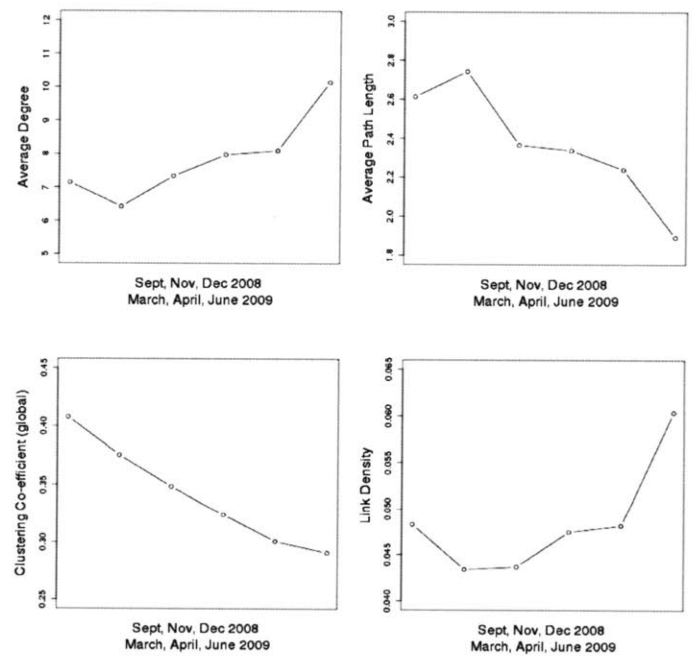


Figure 2-10: Network properties for the self-report close-friends network, for six monthly sociometric surveys between Sept 2008 and June 2009. The average number of (undirected) ties per node increases from approximately 7 (Sept 2008) to about 11 (June 2009) during the course of the experiment. The average (minimum) path length, correspondingly reduces from 2.6 hops to 1.9 hops in the same period.

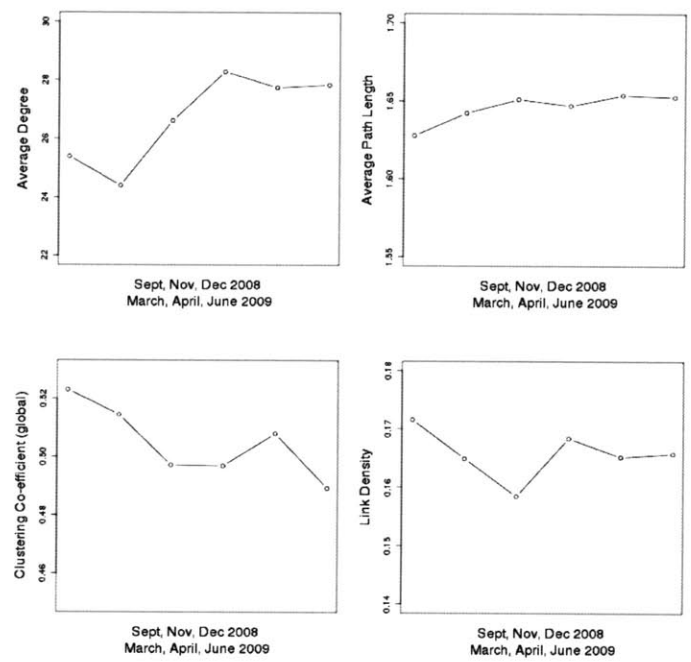


Figure 2-11: Network properties for self-reported 'willing to share Facebook photos and Tweets' networks, for six monthly sociometric srveys from Sept 2008 to June 2009. The average number of (undirected) ties per node increases from approximately 24 (Sept 2008) to about 27 (June 2009) during the course of the experimentThe average (minimum) path length, is around 1.6 hops.

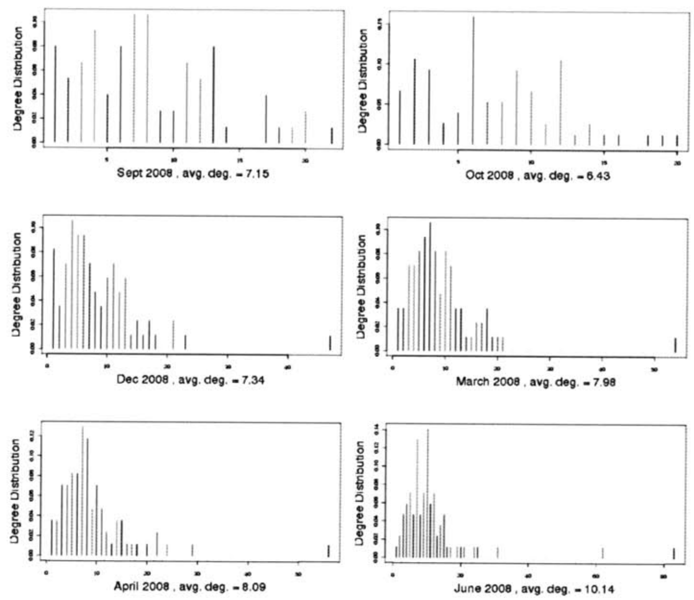


Figure 2-12: Degree Distribution for self-report close friend networks from Sept 2008 to June 2009

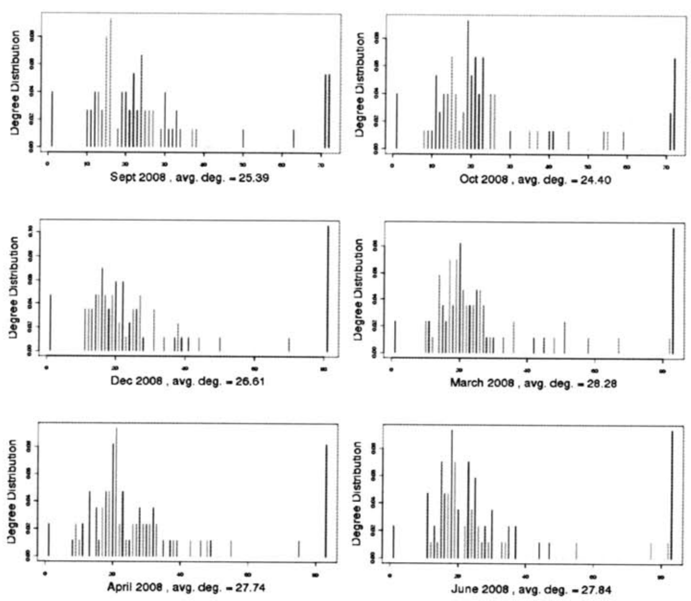


Figure 2-13: Degree Distribution for self-report 'willing to share Facebook photos and Tweets' from Sept 2008 to June 2009

2.4 Discussion

The data collection platform used for the study is described in detail in this chapter. We hope that this 'social sensing' approach will be used in other research studies, and also in applications that provide the user feedback, with the intention of inducing positive behavior changes. Any future mobile sensing platform should be designed along the following guidelines:

- It should be designed for comfortable, longterm, continuous use, ranging from several weeks to months.
- It should be meant to be used as the users always-on, primary mobile device, so that day-to-day use is intuitive, and not cumbersome. User's are more likely to remember to carry and charge a primary device.
- Inbuilt security, encryption and removal of personal identifiers are essential, as these devices collect sensitive personal information that should be secured.
- Support for user feedback, and the ability to visualize his/her own data, and potentially induce positive behaviors are important. Experience sampling could be used for active/reinforcement learning.

Our ability to understand social interactions will also improve with advancements in sensing hardware. For example, global positioning systems (GPS) and assisted cellular triangulation are better localization technologies than 802.11 WLAN APs used in our study. The resolution of face-to-face proximity measurements will improve if Bluetooth signal strength or Infra-Red (IR) sensors become common on smartphones.

Chapter 3

The Evolution of Political Opinions

We study the evolution of political opinions during the last three months of the 2008 US presidential campaign, and the role of face-to-face interactions, in this chapter.

We start with political opinions, since literature suggests that strong ties and social context play an important role in determining them [55]. (Online purchasing behavior is a counter-example, where both online recommendations and face-to-face interactions play an important role [67]). It has been shown that strong, cohesive ties between people lead to high interpersonal influence and faster diffusion. These strong ties are reflected in co-location and communication patterns of users in our dataset [48] . Weak or 'bridging' ties, that allow for short path-lengths while maintaining high clustering, are harder to discriminate using our mobile sensing approach [113].

We find that self-reported *political discussants* have characteristic interaction patterns and political discussant network ties can be recovered from sensor data. Mobile features can be used to estimate unique individual exposure to different opinions, and help discover surprising patterns of dynamic homophily related to external political events, such as election debates and election day in 2008. To our knowledge, this is the first time such dynamic homophily effects have been measured. Using sensor features and estimated exposure and past opinions, it is possible to predict future opinions for individuals ($R^2 \approx 0.8, p \approx 0.001$), and measured exposure increases explained variance by up to 30% over that of survey responses of past opinions alone.

3.1 The Political Opinions Dataset

The dataset used in this analysis is from two sources, for a period of approximately three months, from September 2008 to November 2008.

3.1.1 Mobile Phone Data

The mobile phone interaction dataset consisting of approximately 450,000 bluetooth proximity scans, 1.2 million WLAN access-point scans, 16,900 phone call records and 17,800 SMS text message events. The average duration of phone calls is approx 138 seconds, and 58 percent of measured interactions were during weekdays.

3.1.2 Political Opinions Survey Instrument (Dependent Variables)

The dependent political opinions were captured using three monthly web-based surveys, once each in September, October, and November 2008 (immediately following the presidential election). The monthly survey instrument was based on establised political science literature, and consisted of the questions in Table 3.1. This survey instrument was identical to that used by Lazer and Rubineau [66], who measured the monthly political opinions of students across different universities (during the same 2008 election period) and studied the co-evolution of political opinions and self-report friendship networks.

It is well known that shifts in political opinions are gradual [55]. Approximately 30% of the participants changed their opinions for each of the dependent questions during the 3 month observation period. All opinion changes were limited to 1-point or 2-points on the respective 4/7-point scales. The correlation coefficient for dependent variables representing political opinions is shown in Appendix A.

Table 3.1: Political Survey Instrument used to capture different political opinions. All responses were constructed as Likert scales.

Survey Question	**Possible Responses**
Are you liberal or conservative?	7-point Likert scale Extremely conservative to extremely liberal
How interested are you in politics	4-point Likert scale Not interested to very interested
What is your political party preference?	7-point Likert scale Strong Democrat to strong Republican
Which candidate are you likely to vote for? (Sept and Oct) Which candidate did you vote for? (Nov)	Choice between leading Republican and Democrat nominees Choice between Barack Obama and John McCain
Are you going to vote in the upcoming election? (Sept and Oct) Did you vote in the election? (Nov)	4-point Likert scale Yes or No

Due to the demographics, most participants in the study lean towards democrats on the preferred party question and towards liberals on the liberal-conservative question. For some of the following analysis, to minimize the effects of these imbalanced classes, political party preferences responses were redistibuted into three classes, and liberal-conservative classes are redistributed into 4 classes. The resulting distribution of responses across these different classes of political opinions is summarized in Table 3.2. For the political party preferences case, Independents and Republicans are grouped together due to the small number of samples in each. More information about such binning of Likert scales is available here [59].

For each monthly survey, participants also identified other residents that were political discussants, close friends or social acquaintances at that time. Baseline information including race, ethniciy, political opinions of the person's parents and religious affiliations was also collected before the start of the experiment. These additional surveys were identical to those used by Lazer and Rubineau [66].

3.2 Analysis

3.2.1 Individual Exposure to Diverse Opinions

Can we measure our social exposure to diverse ideas and opinions? Threshold and cascade models of social diffusion [52, 61] assume that all individuals in a population have a uniform exposure, or that the underlying distribution of exposure to different opinions is known. While exposure to different opinions is dynamic and characteristic for every individual, it has previously not been incorporated into threshold and cascade models.

Dynamic exposure to different opinions can be estimated for each participant, on a daily or hourly basis. Contact between two individuals is a function of different extracted features, e.g., physical proximity counts (bluetooth), phone call and SMS counts, and total duration of proximity or phone conversation, or other measures of tie-strength. These features represent

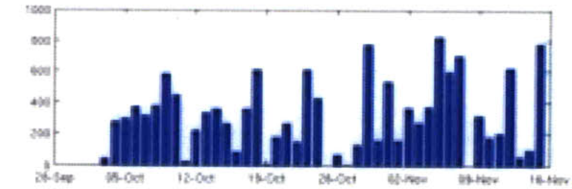
Table 3.2: Redistribution of Responses into Classes (due to majority of Democrats and Liberals

Response Type	**Combined Samples in Sept.**	**Combined Samples in Nov.**
	Are you Democrat or Republican?	
Strong Democrat	24	25
Moderate Democrat Slight Democrat	33	31
Independent Slight Republican Not very strong Republican Strong Republican	15	12
	Are you Liberal or Conservative?	
Extremely Liberal	11	6
Liberal	32	33
Slightly Liberal	12	14
Moderate middle of road Slightly Conservative Conservative Extremely Conservative	18	15

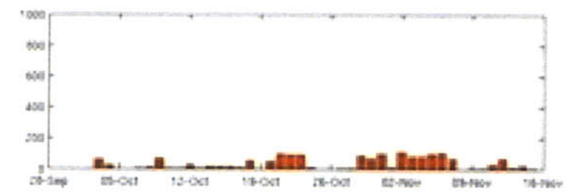
the time spent with others having different opinions in classes, at home, and in phone communication.

We propose two measures of social exposure. Normalized exposure, N_i represents the average of all opinions a person is exposed to on a daily basis, weighted by the amount of exposure to different individuals and their self-reported opinions, where opi_j represents the opinion reponse for person j for a particular question, $contact_{ij}$ is the bluetooth proximity counts between i and j (tie-strength), and $Nbr(i)$ is the set of neighbors for i in the interaction network.

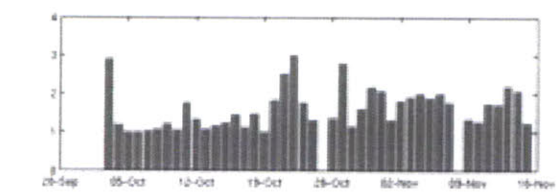
$$N_i(t) = \sum_{j \in Nbr(i)} contact_{ij} \cdot opi_j / \sum_j contact_{ij}$$



(a) Cumulative exposure to Democrats. Y-Axis is Bluetooth proximity counts.



(b) Cumulative exposure to Rpublicans Y-Axis is Bluetooth proximity counts.



(c) Normalized exposure for the Preferred party response. Y Axis

Figure 3-1: Characteristic daily normalized and cumulative exposure for one resident during the election period, from first week of Oct to mid November 2008. Contact is Bluetooth physical proximity. X-Axis is days for all graphs. As seen, this individual had much more exposure to democratic opinions than republican opinions during this period. Incidentally, this person did not show an opinion shift for the preferred-party response during the study (not shown).

Cumulative exposure, C_i to a particular political opinion O, represents the magnitude of a particular opinion that a person is exposed to on a daily basis, and is a function of the amount of contact with different individuals and their self-reported opinion, where $contact_{ij}$ is is the bluetooth proximity counts between i and j (a proxy for tie-strength), and $Nbr(i)$ is the set of neighbors for i in the interaction network. $contact_{ij}$ is can also be estimated from other mobile interaction features, like counts for calling, SMS, and 802.11 WLAN co-location.

$$C_{iO}(t) = \delta_j \cdot \sum_{j \in Nbr(i)} contact_{ij}$$

where $\delta_j = 1$ only if person j holds opinion O, and 0 otherwise.

Figure 3-1 shows cumulative and normalized exposure for one participant during the election campaign period. A limitation of measuring such social exposure from face-to-face interactions, is that it is not possible to model exposure to external sources, e.g., television news, press, or political weblogs. However, it is seen in the following sections, that using sensor-based measures of social exposure alone, we find dynamic patterns of homophily and can identify political discussants. Since self-reported political interest is postively correlated with seeking political information in mass media [55], it could be used as a control in the future.

3.2.2 The Emergence of Dynamic Homophily

Homophily, or the idea of "birds of a feather flock together", [64] is a fundamental and pervasive phenomenon in social networks, and refers to the tendency of individuals to form relationships with others that have similar attributes, behaviors or opinions. Homophily commonly occours in emperical network data, and has been extensively studied as a sociological phenomenon. McPherson and Smith [79] provide an indepth review of homophily literature. The emergence of homophily during network formation has be explained using agent based models, and in economics [36] by incorporating chance, choice, and tie formation costs.

In the context of political opinions, the emergence of homophily amongst randomly assigned individuals would be expected over several months or years [66, 79]. In emperical studies, homophily within a community has been shown to remain stable once network structure is no longer evolving. However, these emperical observations of homophliy in political networks have been limited to survey-based reports of social ties. In this section, we show that our social sensing approach uncovers short-term dynamic homophily artefacts, to our knowledge, not previously known in the social sciences.

As a starting point, static measures of homophily can be estimated using self-report social ties in September 2008 (start of study) and November 2008 (election day). Let w_i be the relative fraction of individuals within the community with opinion i. H_i is the homophily index, defined as the averaged ratio of same-type ties to total ties for individuals with opinion type i. IH_i is Coleman's inbreeding-homophily index, which measures the amount of bias with respect to baseline homophily as it relates to the maximum possible bias (i.e., the term $1 - w_i$). More information about these homophily indices is available here [36, 79].

$$w_i = N_i/N$$

$$H_i = s_i/(s_i + d_i)$$

$$IH_i = (H_i w_i)/(1 - w_i)$$

When the network structure is invariant, the relationship between homophily index H_i and relative fraction w_i reflects the type of homophily in the self-report friendship networks of September and November 2008. If $w_i > w_j$ implies $H_i > H_j$, then the parameters satisfy relative homophily. If $H_i = w_i$ for all i, then the parameters satisfy baseline homophily.. If $H_i > w_i$ for type i, then the parameters satisfy inbreeding-homophily, i.e. the tendency of friendships to be biased towards own types beyond the effect of relative population size. Finally, in some cases, the opposite may be true, such that $H_i < w_i$ for type i, wherein the paramters satisfy heterophily and people of type i are biased towards having different-type friendships.

In terms of the Coleman index, *inbreeding homophily* for type i exists if and only if $IH_i > 0$, and *inbreeding heterophily* for type i if and only if $IH_i < 0$. The inbreeding homophily index is 0 if there is pure baseline homophily, and 1 if a group completely inbreeds.

The relationship between homophily index H_i, inbreeding homophily IH_i and relative fraction w_i is shown in Figures 3-2 – 3-9 for the political interest responses and the preferred party dependent variables, in September and November 2008, for self-reported close-friend networks and self-reported political discussant networks.

For political interest with respect to close friend networks, relative homophily is observed in both Sept. and Nov. 2008, and tends closer to towards baseline homophily in Nov 2008. In Nov 2008, the 'somewhat interested' class (Figure 3-3, in blue) shows fairly high inbreeding hetrophily, i.e., a tendency for forming different-type relationships. This effect is not seen for the other classes.

For political interest with respect to political discussant networks, we see some relative homophily in both Sept. and Nov. 2008, although the 'very interested' class shows a tendency for inbreeding-homophily and the 'somewhat interested' class shows a tendency for inbreeding-hetrophily (Figure 3-5, red and blue respectively). In simpler terms, the 'very interested' participants have more ties with other people with high-interest in politics, and the 'somewhat interested' participants have more diverse ties with other classes of participants than should be expected.

For party preferences with respect to close friend networks, relative homophily is observed as well. Here, the 'strong democrats' show a slight tendency towards inbreeding-homophily, while the 'moderate democrats' show a tendency for inbreeding-hetrophily (Figure 3-7, red and blue respectively).

Finally for party preferences with respect to political discussants, relative homophily is stronger in Nov 2008 than Sept 2008. The 'strong democrats' show a tendency towards inbreeding-homophily while the 'moderate democrats' show a tendency towards inbreeding-hetrophily.

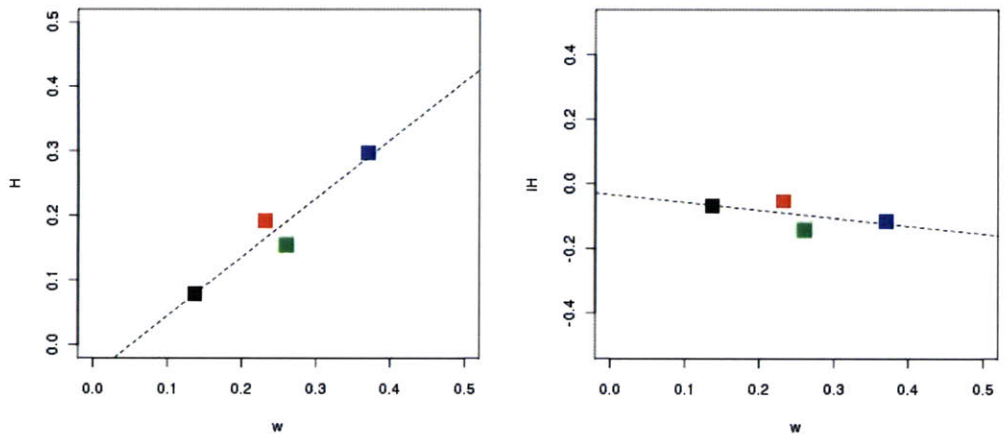


Figure 3-2: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Political Interest response for self-reported close friend ties in September 2008 (start). Red = "Very Interested", Blue = "Somewhat Interested", Green = "Slightly Interested" and Black = "Not at all Interested". X-Axis is w_i in both graphs.

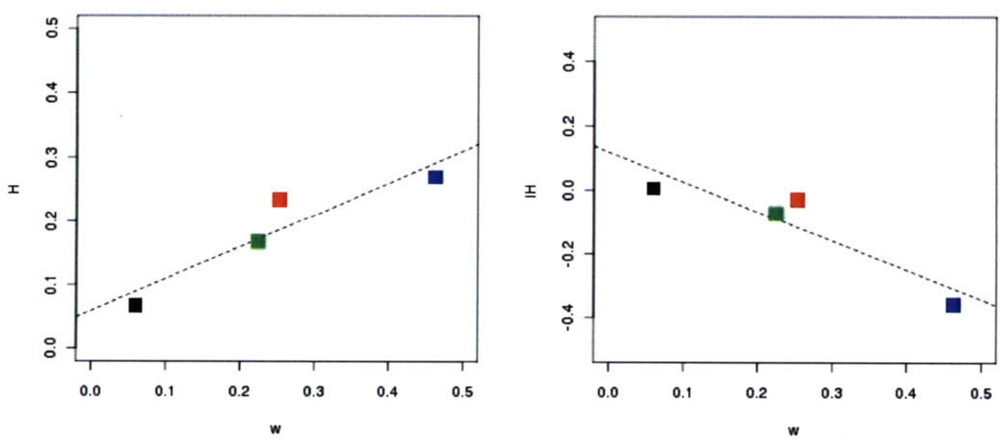


Figure 3-3: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Political Interest response for self-reported close friend ties in November 2008 (post-elections). Red = "Very Interested", Blue = "Somewhat Interested", Green = "Slightly Interested" and Black = "Not at all Interested". X-Axis is w_i in both graphs.

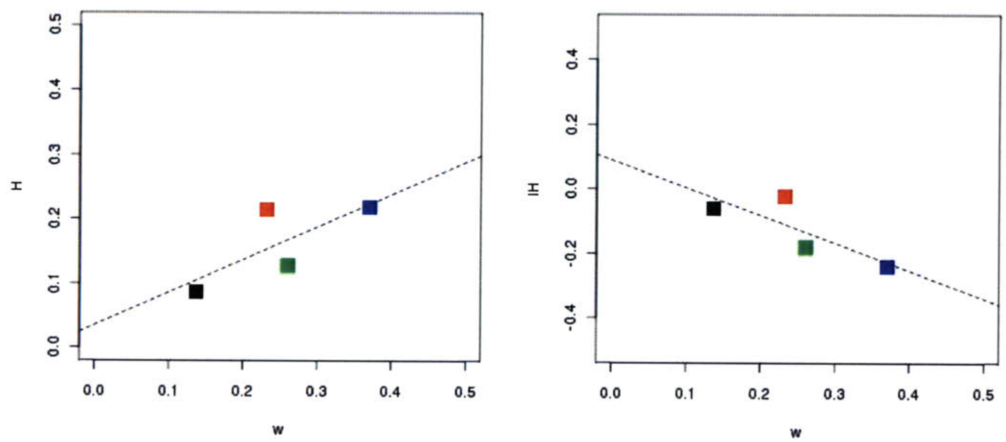


Figure 3-4: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Political Interest response for self-reported political discussant ties in September 2008 (start). Red = "Very Interested", Blue = "Somewhat Interested", Green = "Slightly Interested" and Black = "Not at all Interested". X-Axis is w_i in both graphs.

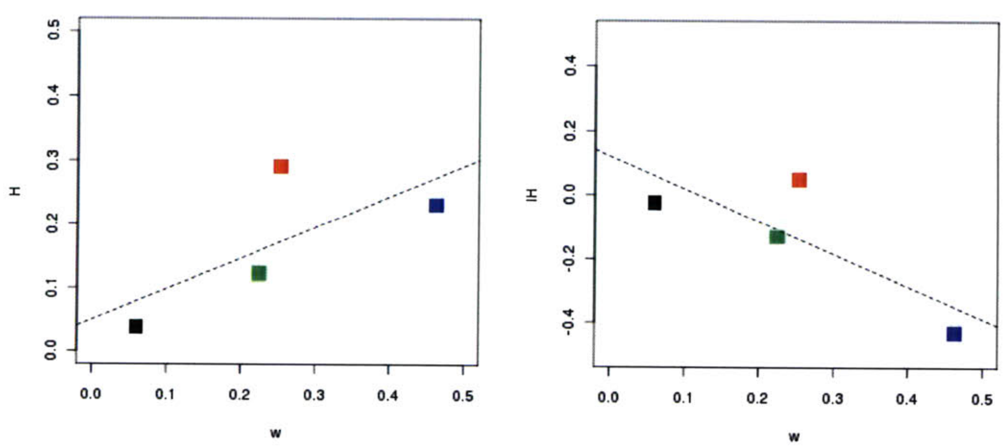


Figure 3-5: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Political Interest response for self-reported political discussant ties in November 2008 (post-elections). Red = "Very Interested", Blue = "Somewhat Interested", Green = "Slightly Interested" and Black = "Not at all Interested". X-Axis is w_i in both graphs.

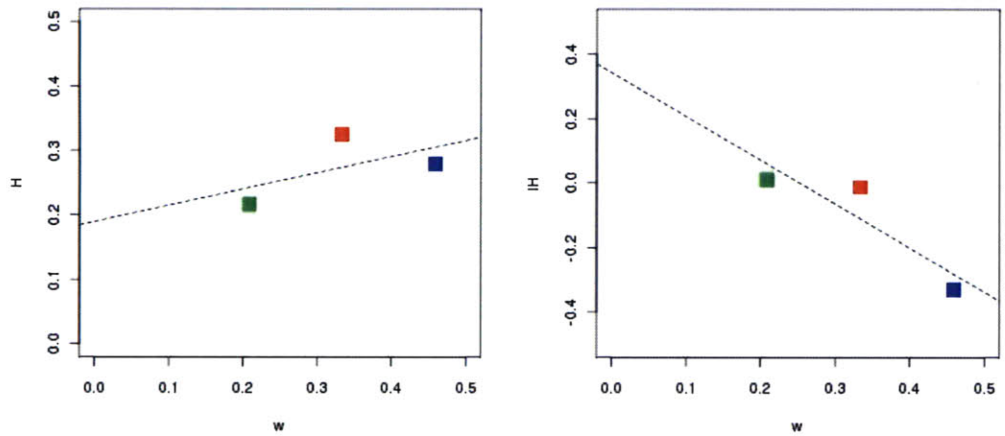


Figure 3-6: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Preferred Party response for self-reported close friend ties in September 2008 (start). Red = "Strong Democrat", Blue = "Not Strong Democrat, Leaning towards Democrats", and Green = "Independent, Leaning Towards Republican Party, Not Strong Republican, Strong Republican", X-Axis is w_i in both graphs.

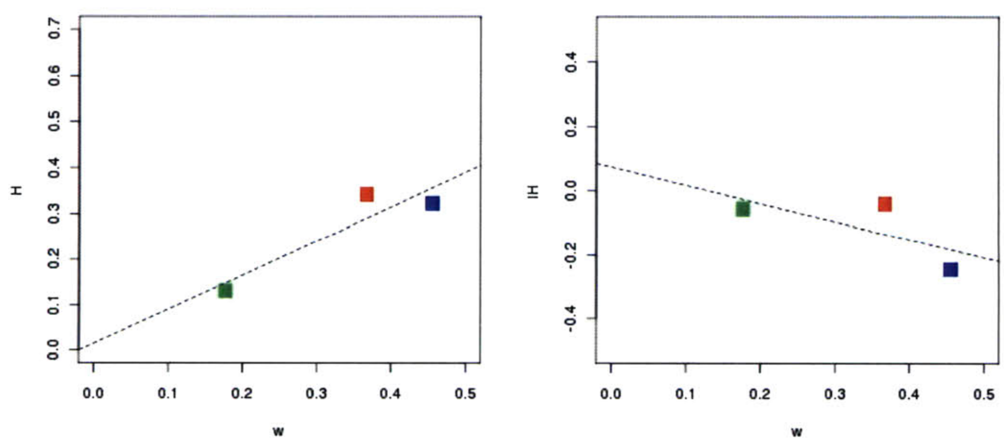


Figure 3-7: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Preferred Party response for self-reported close friend ties in November 2008 (post-elections). Red = "Strong Democrat", Blue = "Not Strong Democrat, Leaning towards Democrats", and Green = "Independent, Leaning Towards Republican Party, Not Strong Republican, Strong Republican". X-Axis is w_i in both graphs.

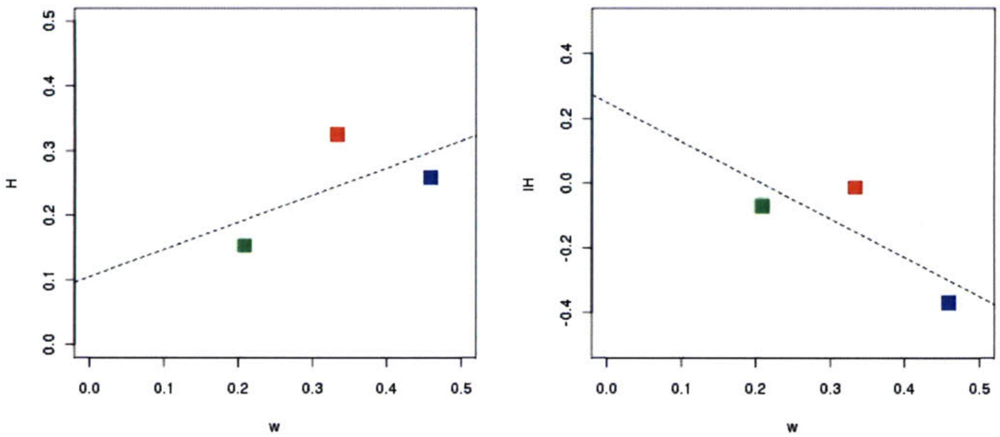


Figure 3-8: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Preferred Party response for self-reported political discussant ties in September 2008 (start). Red = "Strong Democrat", Blue = "Not Strong Democrat, Leaning towards Democrats", and Green = "Independent, Leaning Towards Republican Party, Not Strong Republican, Strong Republican", X-Axis is w_i in both graphs.

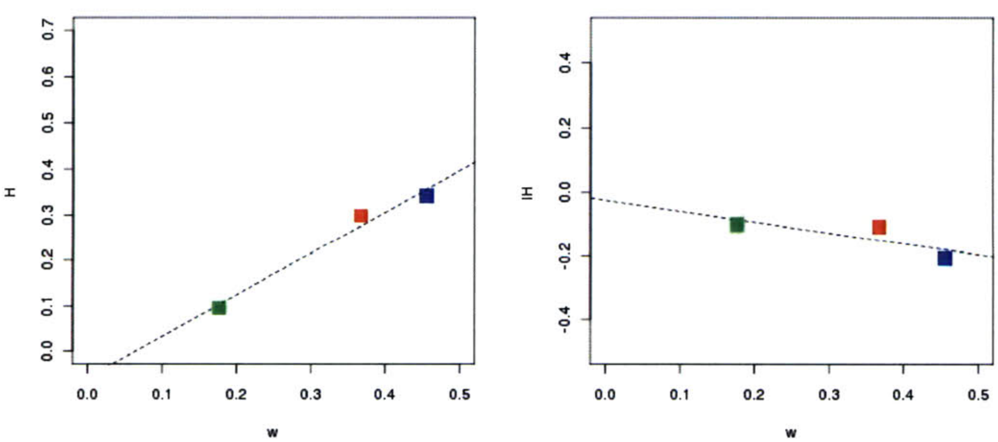


Figure 3-9: Homophily Index H_i (left) and Coleman's inbreeding Homophily index IH_i (right) for the Preferred Party response for self-reported political discussant ties in November 2008 (post-elections). Red = "Strong Democrat", Blue = "Not Strong Democrat, Leaning towards Democrats", and Green = "Independent, Leaning Towards Republican Party, Not Strong Republican, Strong Republican". X-Axis is w_i in both graphs.

While these standard measures of static homophily provide an approximate sense of the segregation and clustering amongst different classes of opinions, they provide very little knowledge about the dynamic nature of the underlying phenomena. We propose a better measure of homophily based on social exposure.

Unlike self-reported ties, social exposure can be used to estimate dynamic homophily patterns on the daily timescale. Dynamic homophily estimated from normalized exposure is given as,

$$\Delta_i(t) = \left| O_i - \sum_{j \in Nbr(i)} contact_{ij} / \sum_j contact_{ij} \right|$$

$$H(t) = \sum_i \Delta_i(t)/n$$

where $\Delta_i(t)$ is the difference between the individual's opinions and the opinions he/she is exposed to, $H(t)$ is a daily measure of dynamic homophily for the entire community, and O_i are an individuals political opinion responses, on the full-range of the 4 or 7-point scale, i.e., for the political interest response, O_i ranges from 1 ("Very Interested") to 4("Not at all interested") and for the preferred party response, O_i ranges from 1("Strong Democrat") to 7("Strong Republican"). Unlike the static homophily measures above, O_i is not based on the redistributed classes.

Daily variations in $H(t)$ are due to changes in mobile phone interaction features, that capture how participants interact with others. A negative slope in $H(t)$ implies that residents have more social exposure to individuals sharing similar opinions, in comparison to the previous day or week. Similarly, an upward slope implies that residents have decreasing social exposure with others having similar opinions.

This daily metric reveals surprising behavior during the final election debate and around election day. For a few days around the election day and final debates, participants show a higher tendency overall to interact with like-minded individuals. Statistical tests to support this are summarized in Table 3.3, and this effect is shown in Figures 3-10 - 3-13.

Statistical validation of these dynamic homophily variations is done using repeated-measures ANOVA, by which dynamic homophily values for different aspects of political opinions are compared during three relevant periods (conditions). For each period, the average dyamic homophily for the 5-day period per participant was estimated. This analysis was first done for all participants, and then repeated for freshmen-only, who had only been in the community for a month before start of the study, and where stronger effects are observed. Full ANOVA tables, condition means, standard deviations, sample counts and supporting information to validate normality assumptions are available in Appendix A [1].

The three experimental conditions (periods) chosen for validating the main effect were:

- Baseline Period (1st condition): 4th October to 10th October 2008
- Final election debate Period (2nd condition): 12th October to 18th October 2008
- Election period (3rd condition): 1st November to 7th November 2008

[1] in additional to the normality of residuals assumptions validated in Appendix I, repeated measures ANOVA may also be susceptible to carry-over effects (i.e., results from the previous phase carry-over into the next phase) and order effects (i.e., the sequence of conditions or treatments affects results)

Table 3.3: Statistically significant variations in Dynamic Homophily around the final election debate period (15th Oct 2008) and election day (4th Nov 2008) period. Dynamic homophily is calculated using bluetooth proximity (phone calling and SMS are not significant for any self-reported political opinions)

Opinions Evaluated for main effects over three periods (conditions)	Result Summary
Political Interest for all participants	Significant effect, higher tendency to interact with like-minded individuals during debate and final election period as compared to baseline period, $F - value = 8.49, p < 0.0004$ (see Figure 3-10)
Political Interest for freshmen only	Significant effect, higher tendency to interact with like-minded individuals during debate and final election period as compared to baseline period, $F - value = 3.43, p = 0.04$ (see Figure 3-13)
Party preference for all participants	Not a significant effect, $F - value = 0.87, p < 0.42$ (see Figure 3-11)
Liberal-conservative tendency for all participants	Significant effect, higher tendency to interact with like-minded individuals during debate and final election period as compared to baseline period, $F - value = 6.26, p < 0.003$ (see Figure 3-12)

These results are illustrated in more detail in the figures below. Figure 3-10 shows $H(t)$ for political interest for all participants, where daily network structure is estimated on the basis of Bluetooth proximity counts. The first dip in this graph corresponds to the period of the final election debate during the campaign, 14th Oct 2008. The difference between the three conditions is statistically significant ($F - value = 8.49, p < 0.0004$).

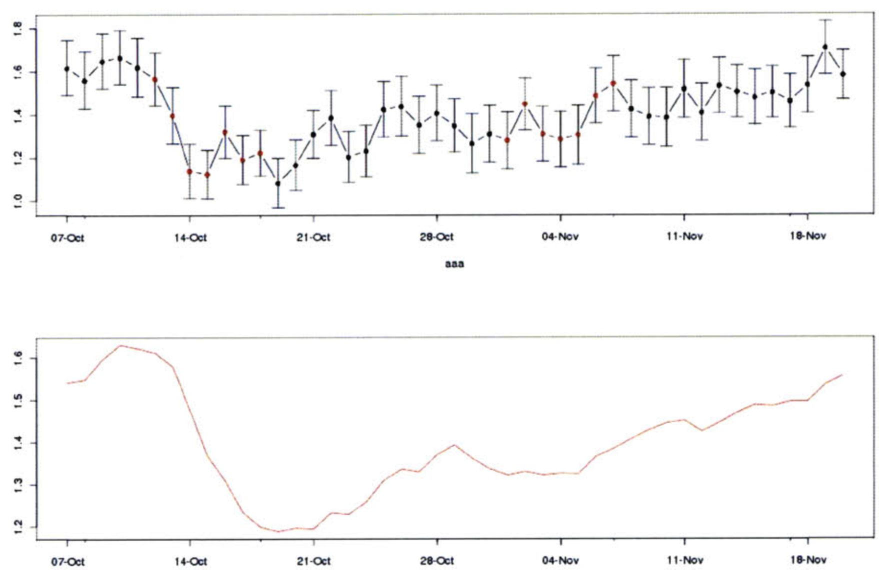


Figure 3-10: Dynamic homophily of **political interest responses** (using bluetooth proximity) for all participants. Top: actual values and standard error bars. Bottom: Moving average. Notice the decline, i.e. tendency to interact with others having similar opinions, lasting for a few days, around Oct 15th 2008, which was the last presidential debate.

Figure 3-13 shows $H(t)$ for political interest only for freshmen, based on daily bluetooth proximity networks. The dynamic homophily effects for freshmen, who only had a month to form ties in this community at this point, are visually pronounced, and a second dip is seen around 4th November 2008 (Election day, $F - value = 3.43, p = 0.04$).

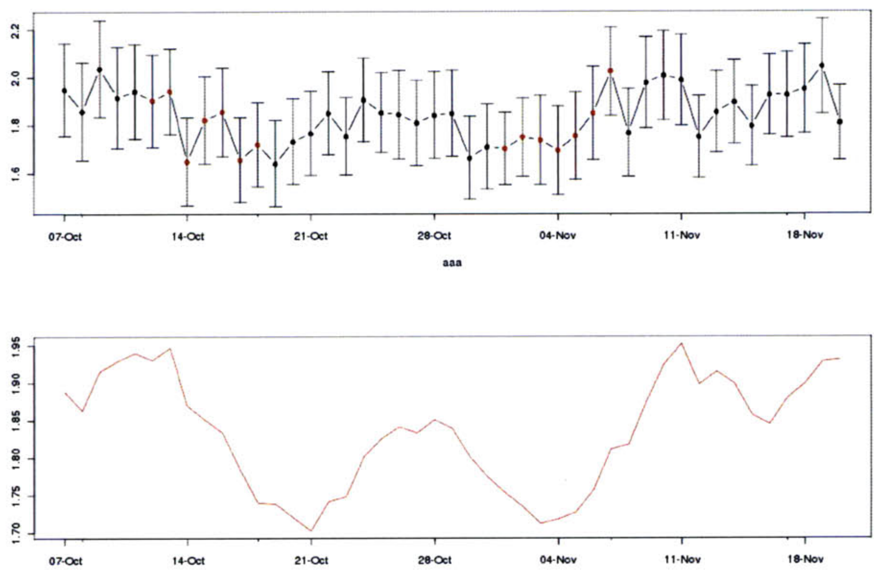


Figure 3-11: Dynamic homophily of **preferred party responses** (using bluetooth proximity) for all participants. Top: actual values and standard error bars. Bottom: Moving average. Participants show a higher tendency to interact with others having similar opinions, lasting for a few days, around election events.

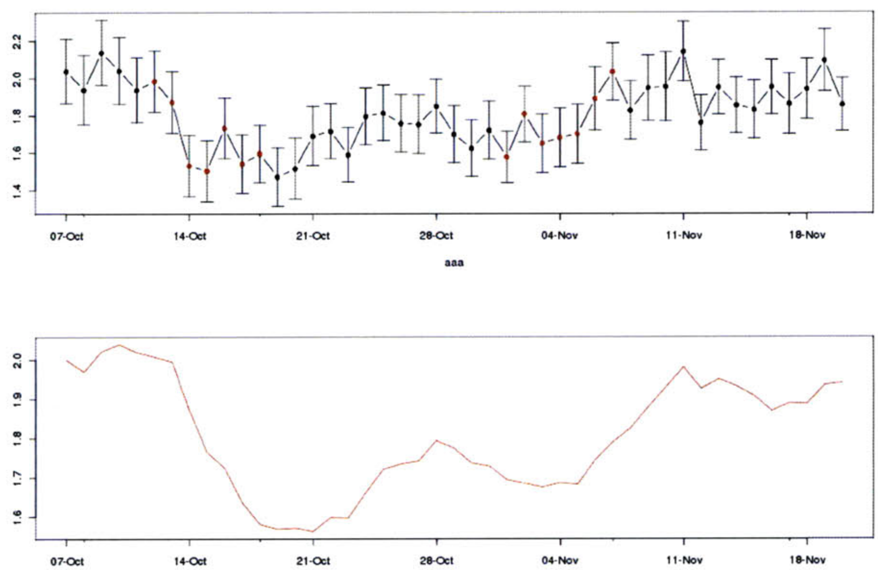


Figure 3-12: Dynamic homophily of **liberal-conservative responses** (using bluetooth proximity) for all participants. Top: actual values and standard error bars. Bottom: Moving average. Participants show a higher tendency to interact with others having similar opinions, lasting for a few days, around election events.

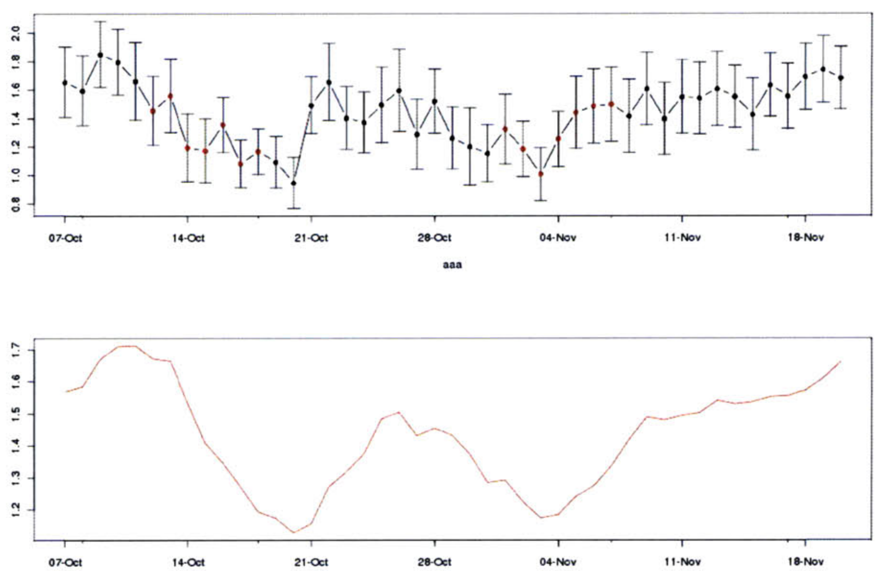


Figure 3-13: Dynamic homophily of **political interest responses** (using bluetooth proximity) only for **Freshmen** . There are two periods of decline, each lasting for a few days. The first is around Oct 15th (last presidential debate) and the second is around 4th Nov, Election Day.

These behavior changes related to external events are seen in bluetooth proximity data, but not in calling and SMS interactions, as shown in Figure 3-14 and Figure 3-15 (freshmen only). This suggests that exposure to different opinions based on physical proximity plays a more important role than exposure to opinions via phone communication. Similar results are also observed for the preferred party responses and liberal-conservativ responses with respect to phone communication.

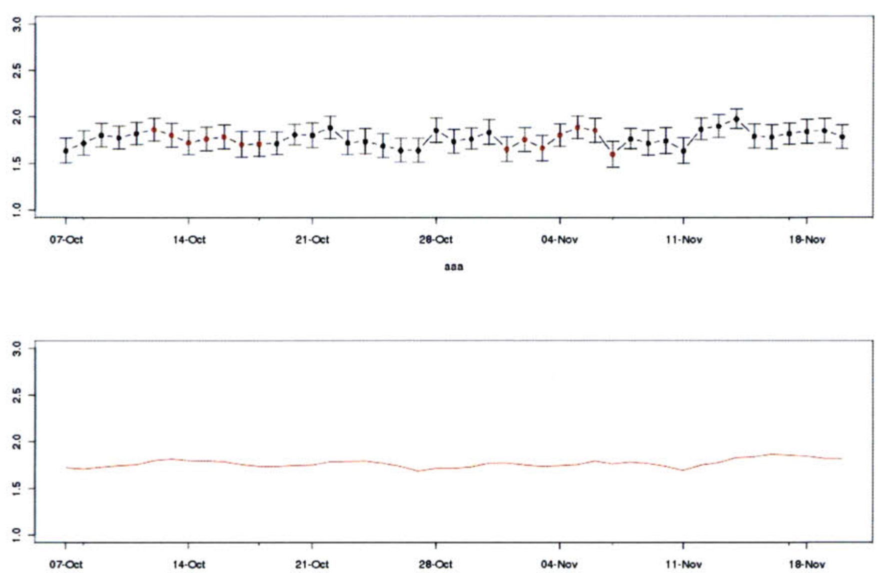


Figure 3-14: Dynamic homophily based on the daily phone-calling network for political interest responses estimated for all residents shows no variation related to election events.

Dynamic homophily graphs for the entire year are shown in Appendix A.

The clustering co-efficient [11] and average path length [112] are commonly used measures to study the clustering and reach of nodes within networks. These measures calculated for daily Bluetooth interaction networks during the same period do not show equivalent significant variations around these periods of interest, which is quite surprising.

As seen in Figures 3-10 - 3-16, there is a tendency of individuals to interact with like-minded individuals around important external events. The effect is significant for political interest and liberal-conservative opinions, but not for particular party preference. It seems

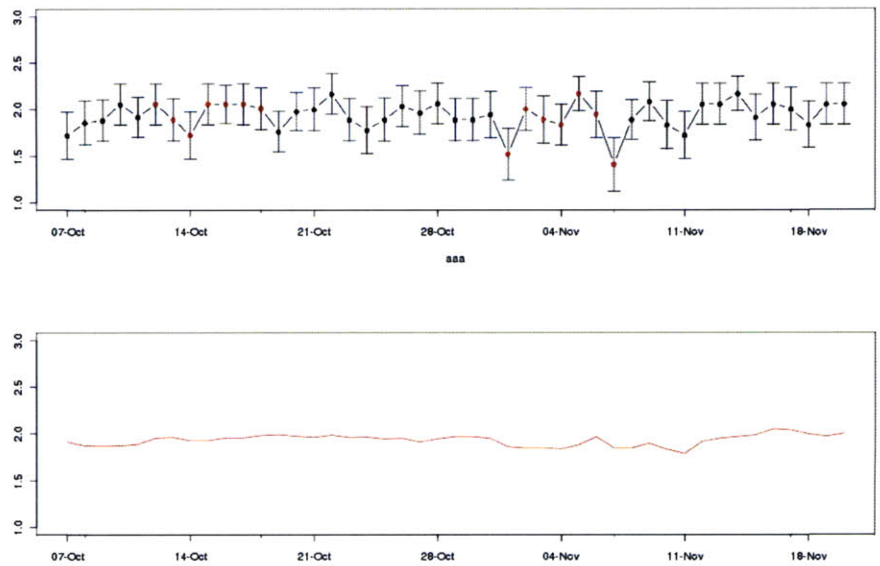


Figure 3-15: Dynamic homophily based on the daily phone-calling network for political interest responses for freshmen alone shows no variation related to election events.

to last only for a few days. This dynamic homophily effect is only observed in Bluetooth co-location networks, and not in calling or SMS networks. This measure captures dynamic patterns of homophily related to global political events from mobile phone sensor data. To our knowledge, this is the first time such an effect has been quantitatively measured. Since the effects last for a few days, they are not tied to a single meeting or discussion alone.

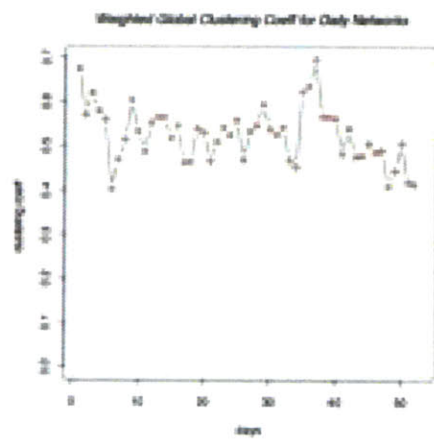


Figure 3-16: Weighted clustering coefficient (Y-axis) for daily Bluetooth interaction networks during the same period, and the X-axis represents days.

3.2.3 Inferring Political Discussants

What are the behavioral patterns of political discussants? In monthly self-reported survey responses, only 39.6 percent of political discussants are also close friends. Similarly, it is found that having similar political opinions does not increase the likelihood that two individuals will be political discussants in this dataset.

While these political discussants do not fit the mould of 'close friends' or individuals with similar political opinions. we find that it is possible to identify political discussants from thier interaction patterns. Classification results based on mobile phone interaction features – total communication; weekend/late-night communication; total proximity; and late-night/weekend proximity, that characterize a political discussant are shown in Table 3.4. Two different approaches are used for comparison, an AdaboostM1 based classifier [47] and a Bayesian network classifier[29], where each input sample represents a possible tie, and both show similar results. As the classes are severely unbalanced, cost-sensitive approaches are used in both case.

In this analysis, political discussants are treated as unidirectional ties. Precision and recall of the discussant class are similar if self-reported the trainling labels are converted to bi-directional ties.

Table 3.4: Classification Results: Identifying Political discussants based on exposure features

Meta-cost AdaboostM1 (individual classifiers are decision stumps)
5-fold cross validation

	Precision	Recall	F-Measure
Non-discussants	0.87	0.62	0.72
Political discussants	0.35	0.67	0.46

Cost-sensitive Bayesian Network classifier
5 fold cross-validation and K2 hill-climbing structure learning

	Precision	Recall	F-Measure
Non-discussants	0.87	0.61	0.72
Political discussants	0.35	0.70	0.46

3.2.4 Exposure and Future Opinions

Exposure based features described in the previous section can be used as a feature to train a linear predictor of future opinions. The coefficients used in a linear model of opinion change include normalized exposure during the period, the persons opinion at the start of the study (September 2008), and a constant term that represents a linearly increasing amount of media influence as we get closer to the election date (Nov. 2008). For the various political opinion questions, regression values are in the $R^2 = 0.8, p < 0.01$ region. Using exposure based features explains an additional 15% - 30% variance across different political opinion questions. The effects for freshmen are approximately twice as strong as compared to the entire population, similar to the variations in dynamoc homophily related to external events. In the context of social science literature, this is a fairly strong effect. Also, since the evolution of political opinions is quite gradual, this approach may be expected to explain more variance for opinions and habits which evolve faster, e.g. food eating habits.

Table 3.5: Least squares regression results for the opinion change model. The dependent variable in all cases is the self-reported political opinion in November. The independent regression variables are averaged opinion of self-reported close friends relationships and political discussants (I and II), normalized bluetooth exposure (III), and normalized exposure combined with past opinion (IV). As seen, automatically captured mobile phone features substantially outperform self-reported close friends or political discussants.

Political Opinion	**I** *Self-reported Discussants*	**II** *Self-reported Close Friends*	**III** *Normalized Exp. Only*	**IV** *Normalized Exp. & Sept Opinion*
Preferred Party	not sig.	not sig.	0.21**	0.78***
Liberal or Conservative	not sig.	not sig.	0.16*	0.81***
Interest in Politics	not sig.	0.07*	0.24**	0.74***
Preferred Party (freshmen only)	not sig.	not sig.	0.46*	0.83*
Interest in Politics (freshmen only)	not sig.	not sig.	0.21**	0.78***

All values are R^2, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

3.3 Discussion

In this chapter, we propose a novel application of face-to-face proximity and communication sensing, for modeling the spread of political opinions. Using automatically captured sensor data, we can estimate social exposure to different opinions for individuals on a daily basis. Around notable political events in this dataset, individuals show a tendency to spend more time with peers who share similar opinions. This effect only lasts only for a few days, and is reflected in a measure of dynamic homophily we propose, but not captured by existing measures of homophily. To our knowledge, this is the first time such a dynamic homophily effect has been discovered in empirical data. For these external events, the community shows more interactions with like-minded individuals for political interest, but less so for party choices. In self-report networks, static homophily analysis shows that people with high-interest and strong democratic party support tend to form in-breeding relationships, while those with moderate political interest and moderate democratic party support tend to have more diverse ties than would be expected.

Using automatically estimated exposure to different opinions, we predictive future political opinions for individuals, and it is found that normalized proximity accounts for up to 30% of the variance in their political opinions in November 2008. While only some political discussants are close friends, and they are not peers with similar demographic background or baseline political opinions, political discussants show a characteristic interaction style in physical proximity and phone communication features, which can be used for classify them.

There are however, several limitations of this work. The simple linear model of opinion change could be improved with dynamic/latent Bayesian models.The relationship between between observed changes in opinions, and the diversity/frequency of contact with potential change-inducers, needs to be explored further. The underpinnings of the observed dynamic homophily effects, and why they are not reflected in other network measures, should also be studied in more detail.

Chapter 4

Face-to-Face Interactions and Epidemiological Behavior Change

Face-to-face interactions are the primary medium for propagation of airborne contagious disease [97]. An important question in behavioral epidemiology and public health is to understand how individual behavior patterns are affected by, and intertwined with, physical and mental health symptoms. Until recently, epidemiologists have not had access to sensing and modeling capabilities to quantitatively measure behavioral changes experienced by symptomatic individuals in real-world scenarios [37]. Such research requires simultaneously capturing symptom reports, mobility patterns and social interactions amongst individuals continuously over long-term duration. In this chapter, study the link between physical respiratory symptoms, influenza, stress, mild depression and automatically captured behavioral features. This is an important problem for several reasons.

Quantitatively understanding how people behave when they are infected would be a fundamental contribution to epidemiology and public health, and can inform treatment and intervention strategies, as well as influence public policy decisions. On one hand, clinical epidemiology has accurate information on the evolution of the health of individuals over time but lacks realistic social interaction as well as spatio-temporal data [44]. On the other hand, current research efforts in theoretical epidemiology model the rate of infection in a population whose behavior is stationary over time and do not account for individual changes [95]. For instance, if a person infected with influenza continues his habitual lifestyle instead of isolating himself, he could pose a bigger risk to others in proximity. Based on our analysis and results, policymakers can recommend social interventions that minimize such risk.

On the modeling front, compartmental epidemiological models (e.g., the Susceptible, Infectious, Recovered or SIR model) commonly assume that movement and interaction patterns for individuals are stationary during infection, i.e., that individuals will continue their typical behavioral patterns when sick. More recent epidemiological models accommodate reduced mobility variations to to fit epidemic curves, but in a heuristic way due to lack of data at the individual level [20, 35, 43], which possibly limited their prediction accuracy during the 2009 H1N1 influenza epidemic [78]. To our knowledge, we provide the first quantitative results on this important measurement based on mobile sensing. Our results, described below, can be plugged into the SIR model by specifying the number and frequency of contacts that individuals will likely have when going from the S(usceptible) to the I(infected) state, and therefore improving prediction accuracy. Furthermore, predicting likelihood of symptoms from behavior could lead to a possible early-warning system and intervention by medical experts, with the associated savings in economic and human resources [30].

The experimental context is two months of our experimental deployment, from February to April 2009. Individuals were surveyed on a day-to-day basis for symptoms of airborne contagious diseases like common colds, influenza and gastroenteritis. We find that there are characteristic changes in behavior when individuals are sick, reflected in automatically captured features like their total communication, communication patterns with respect to time of day (e.g. late night, early morning), diversity of their network and entropy of movement within and outside the university. Due the pervasiveness of the mobile phones, this approach can be scaled to large-scale models of epidemiological contagion in the future. We then use a recently developed signal processing approach [49] to shed light on the information flux between physical symptoms, behavior changes and stress based on temporal information flux gathered by our mobile sensors.

4.1 Link Between Physical Symptoms, Behavior Changes and Stress

Sensor-based approaches may allow us to make a fundamental contribution to epidemiology, in understanding the link between physical symptoms, mental health symptoms, and their expression in social interactions and behavior. The intertwined relationships between these is not well understood, due to limitations of existing clinical diagnosis and public health tools, but plays an important role in clinical diagnosis, treatment and management of chronic conditions.

In medical literature, substantial evidence has been found for an association between stress and increased illness behavior, and less convincing but provocative evidence was found for a similar association between stress and infectious pathology: Introverts, isolates, and persons lacking social skills may also be at increased risk for both illness behaviors and pathology [32]. Various medical conditions that involve activation of the immune system are associated with psychological and neuroendocrine changes that resemble the characteristics of depression. Recent studies have presented empirical evidence on the relationship between the behavioral effects of immune activation and depressive symptomatology, characterized by reduced locomotor, exploratory, and social behavior [117].

The association between psychosocial stress and susceptibility to upper respiratory tract infection has also been investigated in people with a history of recurrent common colds and flu. Several dimensions of psychosocial stress, including exposure to stressful experiences, stress-prone personality traits, and signs of emotional disturbance have been investigated in people with a history of recurrent common colds and flu. Experts conjecture that stress depletes local immune protection, increasing susceptibility to colds and flu. Alternatively, psychological disturbances could develop in response to frequent illness [40].

4.2 The Dataset

The analysis in this section combines data from two sources, daily survey instrument (completed on the mobile phone) and automatically captured interaction data as explained in Chapter 2. The dataset described here corresponds to the period of 1st February to 15th April 2009, the peak influenza months in New England.

4.2.1 Baseline and Daily Survey Instruments

A baseline survey, daily mobile phone survey instrument and post-experiment survey were designed under the supervision of experienced epidemiologist [1] to generate training labels for this section.

For meaningful analysis, it is important to isolate any confounding effects due to immunization prior to the start of the experiment. The participants completed a baseline survey a few days before start of the study. Approximately 20 participants reported received influenza immunization via a flu-shot or flu-mist spray, and they are not considered in the analysis in the next section.

The survey instrument was designed to capture daily symptom reports from the participants. This survey instrument was designed to elucidate syptom self-reports related to common contagious conditions– common colds, influenza, gastroenteritis, and fever possibly due to these conditions or other factors. In addition, to understand the link between physical symptoms and behavior change, the survey also included two questions related to mental health, related to stress and early onset of depression. The question text, possible responses and order is shown in Table 4.1.

This daily survey intrument was launched each day automatically on the mobile phone platform. The launcher application launched a foreground survey dialog at 6am everyday that asked the user to respond to the six survey questions. After three reminders, the device UI was unusable until the user completed the survey. In the experiment deployment, users

[1] Dr. Devon Brewer

Table 4.1: Symptom Survey Questionnaire. All questions were Yes/No responses

Survey Question (as shown on mobile phone)
Do you have a sore throat or cough?
Do you have a runny nose, congestion or sneezing?
Do you have a fever?
Have you had any vomiting, nausea or diarrhea?
Have you been feeling sad, lonely or depressed lately?
Have you been feeling stressed out lately?

were paid $1 USD for every completed daily survey as participation incentive. The survey launcher invoked the following daily questions:

The design of the survey questionnaire and subsequent labeling of self-report responses was supervised by a trained epidemiologist. A total of 2994 survey responses were generated using the smartphone-based survey launcher described in the previous section within the relevant date range, of which 2099 responses were from individuals who had not been immunized, with an approximate survey completion rate of 63% overall during the study. These responses were converted into 48-hour windows, since individuals take up to one day to report a symptom. This approach also reduces the impact of uncompleted survey responses, since there is a higher likelihood that a participant will have completed at least one survey during a 48-hour period. A few samples were dropped because of missing sensor data during the period, e.g., due to lost or broken mobile phones. With these steps, a total of 2283 sets of good quality mobile behavior sensor data and dependent variable reports were obtained.

For analysis, mobile symptom reports have to be converted into syndrome conditions. It is especially important to distinguish between symptoms that represent common colds and allergies versus CDC-defined influenza [57] which has a characteristic signature reflected in runny nose, sore throat and fever symptoms. Due to our limited expertise in this area, selected combinations of self-reported symptoms were labeled as CDC-defined influenza by a medically trained epidemiologist (see Acknowledgements). In our dataset, twelve such cases of influenza, on average lasting 5-7 days and each affecting a distinct individual were observed. The respiratory symptoms not identified as influenza cases by our expert, are considered common colds or seasonal allergies.

4.2.2 Mobile Interacton Data

The phone sensor data during this period consists of 1,424,000 bluetooth samples, 201,000 WLAN AP samples, 15,700 call data records amd 11,269 SMS records.

4.3 Analysis

4.3.1 Mobile Behavioral Features

The following features were extracted from mobile phone sensor data over 48-hour window sizes, with 50 percent overlapping windows. The window size was chosen for epidemiological reasons, as individuals take up to one day to report a symptom. The features chosen represent statistics of whom we talk to and where we are, i.e. the total number of interactions, the diversity of interactions, and the entropy of our behaviors. Such observational data has been shown to reflect important aspects of individual and collective behavior, like friendships and individual job satisfaction [42, 90]. For days with reduced activity, the entropy features capture the higher predictability of an individual's behavior, and was a better feature than total number of bluetooth devices or WLAN APs observed.

Total Communication

This is the total number of phone calls and SMS exchanged, both with other participants as well as third parties. This measure includes incoming and outgoing communication.

Late night and Early Morning Communication

Call and SMS communication between 10pm and 9am on weekdays, with both other participants and non-participants.

Communication Diversity

The number of unique individuals reflected in phone and SMS communication within the 48-hour period.

Physical Proximity Entropy with Other Participants

This is the entropy of distribution of Bluetooth proximity with other participants.

$$H_p = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

where $p(x_i)$ is the empirical probability of Bluetooth proximity with the remote device x_i belonging to another participant, within the particular time-window, i.e., $p(x_i)$ is the ratio of the number of times the remote device x_i was scanned divided by total count of scanned devices in the 48-hour period.

Physical Proximity Entropy with Other Participants Late Night and Early Morning

Similarly, this is the entropy of the distribution of Bluetooth proximity with other participants in the study, but only during late-night and early morning periods.

Physical Proximity Entropy for Bluetooth Devices Excluding Experimental Participants

Similarly, this is the entropy of distribution of Bluetooth proximity. However, all Bluetooth devices in discoverable mode scanned on the phone are considered in this case. This feature reflects variations in interactions with 'familiar strangers', i.e., bluetooth beacons that the user is often in proximity to, say at the bus-stop or in the classroom [81].

WLAN Entropy based on University WLAN APs

This is entropy for the distribution of WLAN access points scanned within the given period. Only WLAN APs belonging to the university are considered.

$$H_w = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

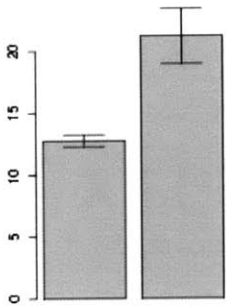
where $p(x_i)$ is the empirical probability of scanning a WLAN AP x_i within the particular time-window. Similar to Bluetooth physical proximity above, $p(x_i)$ is the ratio of the number of times the WLAN AP x_i was scanned divided by the total count of scanned WLAN APs in the 48-hour period.

WLAN Entropy based on external WLAN APs

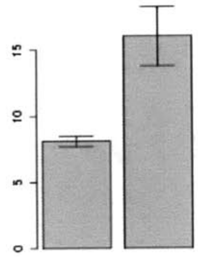
Similarly, this is entropy for the distribution of WLAN access points scanned within the given period. Only WLAN APs external to the university are considered.

4.3.2 Behavioral Effects of Low Intensity Symptoms (Runny Nose, Sore Throat and Cough)

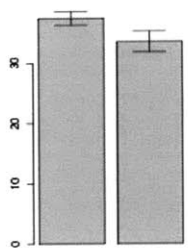
A sore throat or runny nose report may either be a symptom of CDC-defined influenza or simply an independent respiratory condition due to common colds or allergies.



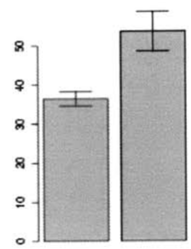
(a) Total communication increases ***



(b) Late-night early morning communication increases **



(c) Overall bluetooth entropy decreases *

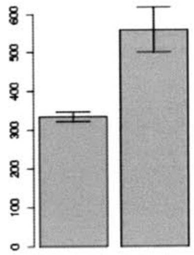


(d) Total WLAN APs detected increase **

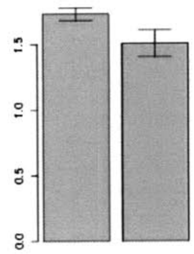
Figure 4-1: Behavior effects of runny nose, congestion, sneezing symptom, n=587/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

For the runny nose condition (n=587/2283), participants show increased total communication as well as increased late night early-morning communication. Additionally, total counts of Bluetooth proximity and measures of WLAN entropy increases, which is perhaps counter-intuitive. P-values are generated using unbalanced t-tests assuming unequal variance.

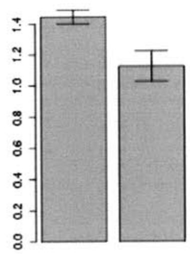
For sore-throat reports, Bluetooth-based entropy with respect to other residents in the study dormitory increases. This again, is slightly counter-intuitive, but may be explained if participants are spending more time indoors and hence have a higher likelihood of interacting with other participants, than they would if they were spending time in classes and activities. It is also found that WLAN based entropy measures, both with respect to university WLAN



(a) Bluetooth entropy with respect to other dorm residents increases ***



(b) WLAN entropy with respect to university WLAN APs reduces *

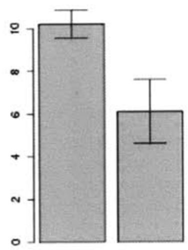


(c) WLAN entropy with respect to external WLAN APs reduces **

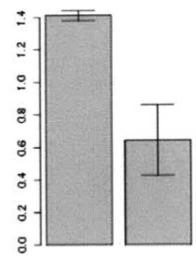
Figure 4-2: Behavior effects of sore throat and cough symptom, n=393/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

APs and external WLAN APs decrease with sore-throat reports, indicating more predictable movement patterns for the individual.

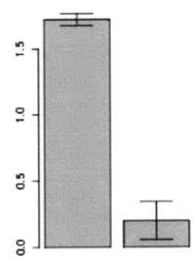
4.3.3 Behavior Effects of Higher-Intensity Symptoms (Fever and Influenza)



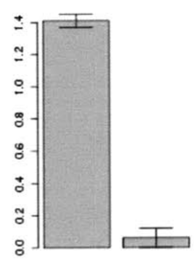
(a) Late night early morning calls and SMS decrease **



(b) Late night morning Bluetooth counts and entropy decrease*



(c) WLAN based entropy with respect to university WLAN APs decreases ***

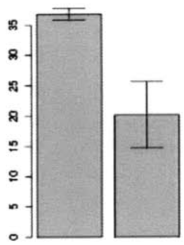


(d) WLAN Entropy with respect to external WLAN APs decreases ***

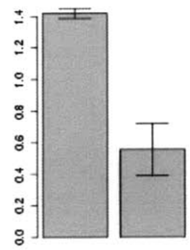
Figure 4-3: Behavior effects of fever, n=36/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

For more intense conditions like a fever or CDC-defined influenza, participants have lower activity and entropy levels, and this is captured using mobile sensors. Due to the severity of these symptoms, the number of reported cases in our dataset is lower than that of low intensity symptoms (runny nose, sore throat/cough). The number of rate of infection amongst participants and study cohort sizes, however, are comparable to Phase I clinical trials [102].

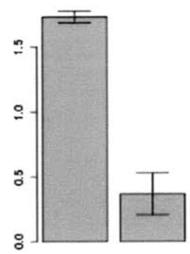
For fever, variations are observed in the late night and early morning behavior. Phone communication, Bluetooth proximity counts, and Bluetooth entropy all show a decrease for the late night early morning window. WLAN-based entropy measures with respect to the university WLAN APs as well as external WLAN APs both reduce dramatically.



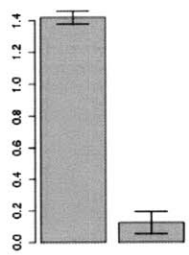
(a) Total Bluetooth interactions and entropy decrease **



(b) Late night early morning Bluetooth entropy with respect to other participants decreases **



(c) WLAN based entropy with respect to university WLAN APs decreases ***

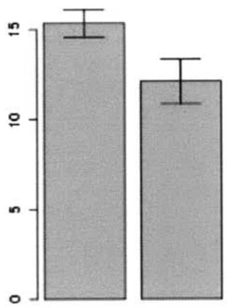


(d) WLAN Entropy with respect to external WLAN APs decreases ***

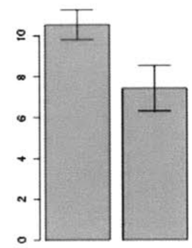
Figure 4-4: Behavior effects of CDC-defined influenza, n=54/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

Similar effects are seen for days labeled as CDC-defined Influenza, as overall bluetooth entropy, bluetooth entropy with regard to other dorm residents and WLAN based entropy features decrease. This is also expected because fever is a known influenza symptom.

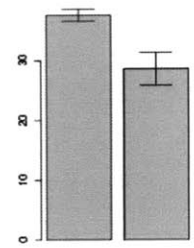
4.3.4 Behavioral Effects of Stress and Mental Health Symptoms



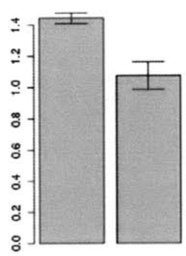
(a) Total communication decreases *



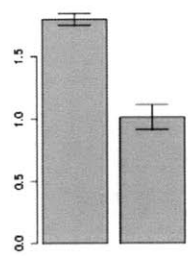
(b) Late-night early morning communication decreases *



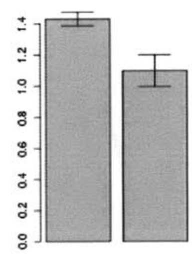
(c) Overall Bluetooth entropy decreases *



(d) Late night early morning Bluetooth entropy with respect to other experiment participants reduces **



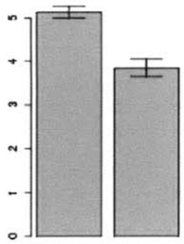
(e) WLAN based entropy with respect to university WLAN APs decreases ***



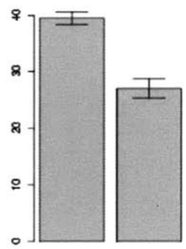
(f) WLAN Entropy with respect to external WLAN APs decreases ***

Figure 4-5: Behavior Changes with self-reported sad-lonely-depressed responses n=282/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

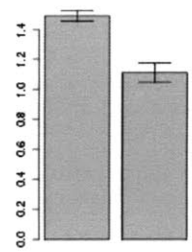
In addition to the physical symptoms described in the above section, the on-device mobile questionnaire also includes two daily questions related to stress levels and sadness, loneliness or depression. As discussed in the previous section, the link between behavior change, physical symptoms and stress is not very well understood. Measuring these self-report variables alongside symptom data allows modeling the covariance and potentially causation across the three sets of variables. With both often-stressed and sad-depressed-lonely responses in our dataset, participants show a consistent tendency to isolate themselves, reflected in various sensor modalities.



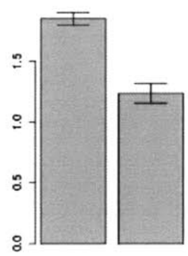
(a) Communication diversity decreases **



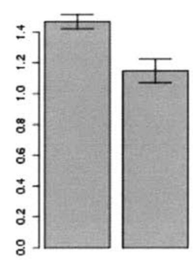
(b) Overall Bluetooth entropy decreases **



(c) Late night early morning Bluetooth entropy with other experiment participants reduces **



(d) WLAN based entropy with university WLAN APs decreases ***



(e) WLAN Entropy with external WLAN APs decreases ***

Figure 4-6: Behavior Changes with self-reported often-stressed responses n=559/2283, *: $p < 0.05$ **: $p < 0.01$ ***: $p < 0.001$

For the often-stressed response, participants communication diversity decreases, both overall Bluetooth based entropy and Bluetooth entropy with respect to other residents during late-night early morning hours decreases, and WLAN based entropy decreases both with respect to university WLAN APs and external WLAN APs.

For the sad-lonely-depressed responses, a similar tendency to isolate themselves is observed. Total communication decreases and communication during late-night early morning decrease, overall Bluetooth entropy and Bluetooth entropy with respect to other residents decreases.

4.3.5 Symptom Classification using Behavioral Features

It is evident that there are characteristic behavioral changes associated with respiratory symptoms, fever, influenza, stress and depression. With this is in mind, is it useful to train a classification scheme that identifies when individuals are likely to be symptomatic from behavioral features alone. There are two key considerations with regard to designing such a classification scheme.

First, consider how such a classification system would be used in a scenario where the user has the mobile sensing application installed on their personal phone. When this application detects uncharacteristic variations in behavior, it could predict the likelihood that the user is infected with a known symptom and potentially inform a nurse, family member or healthcare professional. Such proactive healthcare is especially useful for conditions with risk of under-reporting by patients (e.g., mental health, elderly healthcare). With this goal in mind, the classification model should have asymmetric misclassification penalties.

A second consideration is due to correlations amongst dependent symptoms. While behavior variations with respect to symptoms are reported individually in the previous section, in reality, self-reported symptoms are correlated. Figure 4-7 shows the correlations between these variables, reordered using K-nearest-neighbor clustering based on effect size. Four main clusters that emerge are: stress + depression; runny nose + sore throat; fever + influenza; and runny nose + sore throat + fever + influenza.

Given these considerations and unbalanced class sizes, classification is done using a Bayesian-network classifier with MetaCost, a mechanism for making classifiers cost-sensitive [96]. Structure learning for the network is performed using K2 hill climbing and the results are based on 4-fold cross-validation.

Recall, Precision and F-measure for the symptoms class as a function of increasing misclassification penalty for the symptoms class are plotted in Fig 4-8(a) - 4-8(e), for different symptom clusters. Recall from the trained classifier is also compared with random assignment of priors averaged over 1000 simulated runs, to demonstrate improvement over 'chance'. The X-Axis represents increasing MetaCost misclassification penalties, and the

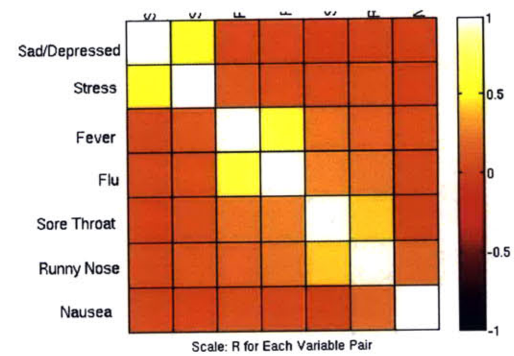
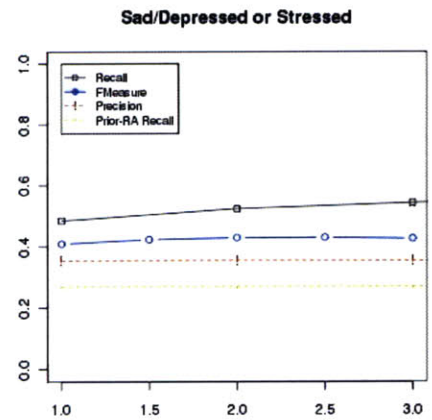
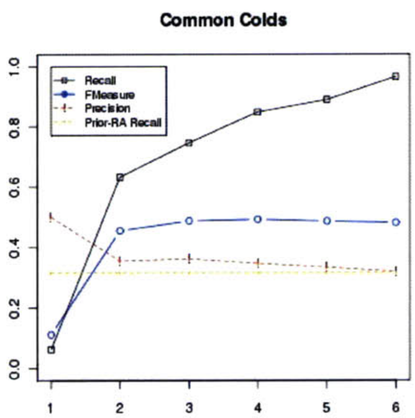


Figure 4-7: K-nearest neighbour reordering of correlations between the different dependent symptoms, to illustrate the dependent symptom variables that are closely related to each other

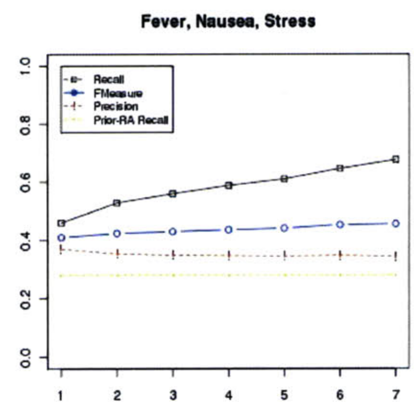
Y-Axis shows Recall, Precision, F-Measure (for symptom class) and Recall based on chance. As seen, Recall of the symptom class improves substantially for the sore throat + runny nose + congestion cluster, and also for influenza and fever clusters. Overall prediction accuracy (not shown) is not a useful quality metric due to unbalanced classes, and ranges between 60-80%.



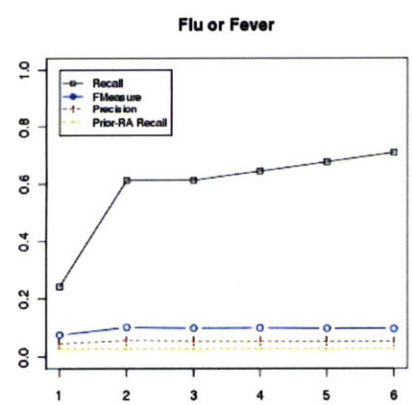
(a) Sad-Depressed-Stressed Symptoms



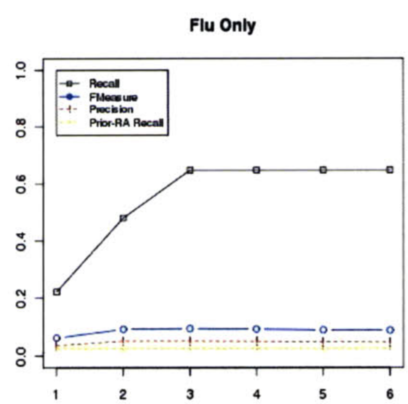
(b) Sore-Throat, Cough, Runny Nose, Congestion, Sneezing Symptoms



(c) Fever, Nausea, Stress Symptoms



(d) Flu and Fever Symptoms



(e) Flu only (as per CDC definition)

Figure 4-8: Classification results, recall for different symptoms ranges from 0.6 to 0.9 for the symptom class. Y-Axis shows Recall, Precision, F-Measure (for symptom class) and Recall based on chance. X-Axis is the MetaCost misclassification penalty used with the Bayesian Network classifier.

4.3.6 Temporal Flux Between Behavior, Stress and Physical Symptoms

There is extensive medical and health policy interest in understanding the temporal link between behavior change, stress and physical symptoms. In econometrics, the Granger causality test is a technique for determining whether one time series is useful in forecasting another– A time series X is said to Granger-cause Y if it can be shown, usually through a series of F-tests on lagged values of X (and with lagged values of Y also known), that those X values provide statistically significant information about future values of Y. Unfortunately, Granger causality tests have been shown to have poor noise immunity [89]. In this section, we use a recently proposed spectral method, Phase Slope Index to gain insight into the temporal relationships between signals of behavior features, stress and physical symptoms.

The Phase Slope Index (PSI) Method

PSI [49] is a recently proposed spectral estimation method designed to measure temporal information flux between time-series signals. The method is based on the knowledge that the phase slope of the cross-spectrum of two signals can be used to estimate information flux between these signals in the time domain. Independent noise mixing does not affect the complex part of the coherency between multivariate spectra, and hence PSI is considered more noise immune than Granger analysis. PSI has been used to make causal inferences for brain cell activation and other domains, and is calculated as,

$$\Psi_{ij} = \Upsilon\left(\sum_{f \in F} C_{ij}^{*}(f) C_{ij}(f + \delta f)\right)$$

where C_{ij} is the complex coherency. When the input signals are distributed across multiple epochs, then this estimate is normalized by its standard deviation, calculated using the Jackknife method [49].

Results

Our approach to using PSI for measuring information flux is based on validating causal links consistently across multiple participants in our dataset. This approach is first validated on two simulated time series of varying sequence lengths (n, representing number of continuous samples available per user) that have a partial-causal relationship between them, and additive noise. The leading time series has x symptom days. The follower time series has y lagged symptom days and z days of additive uniform noise, where the lag between the two series for symptom days is 1 or more days. The scatter-plot in Figure 4-9 shows the ability of PSI to recover causal structure (normalized PSI coefficient > 0) across different ranges of parameters for the simulated signals. The X and Y-axes represent n and x, and each point is averaged over 1000 runs with $y = x/3$ and $z = x/3$, (these values are such that would be intuitively expected for symptoms in our dataset), and all points above the plane of Z $= 0$ represent correctly-estimated PSI values. It is important to note that the method recovers the correct direction of information flux for 97.6% of the samples over the surface of the simulated signals.

In order to apply PSI to our dataset, the subset of participants that show both physical symptoms and stress and depression related responses are considered. There is however, a trade off to be made between using data from fewer participants with longer sequences and hence more reliable estimates, versus using data from more participants with shorter sequences, and better validation across participants. Hence, PSI was estimated for two sets of data– sequences of minimum length 40 days and minimum length 60 days, shown in Figure 4-10.

Each approach generates slightly different directed links and normalized coefficients. The twelve largest PSI coefficients across both methods on the basis of a combined ranking score are listed in descending order in Table 5.1 and illustrated in Figure 4.3.6. An example insight is that 'often-stressed' is useful in forecasting proximity, communication and WLAN behaviors, which suggests that individuals realize and report that they are stressed before it is reflected in their behavior. Another insight is that in two cases Bluetooth interaction features are used to forecast WLAN features– this suggests that a behavior change is reflected

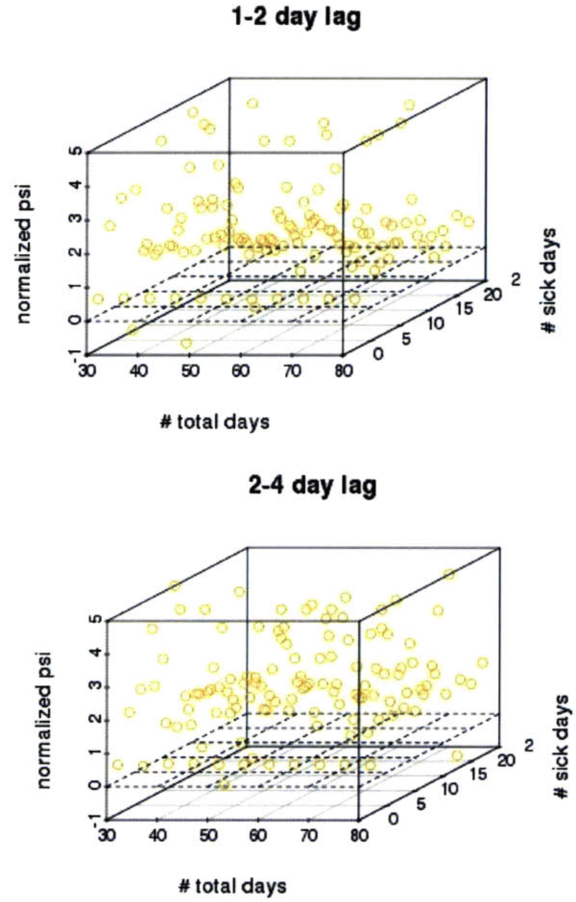
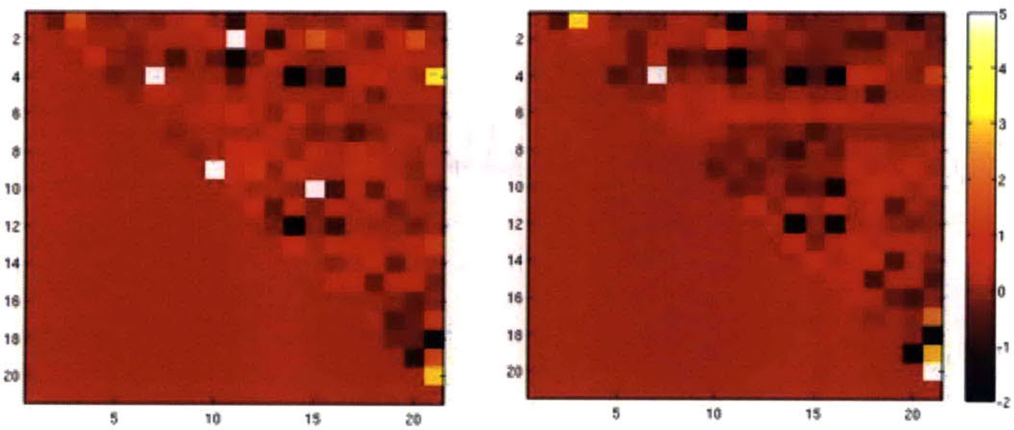


Figure 4-9: PSI evaluation on simulated data. Z-axis is the estimated PSI value, across a wide range of total days (n) and sick days(x), with additive noise. Points above the Z=0 plane (97.6%) represent correctly estimated direction of information flux, for the simulated data.

in face-to-face interactions with others before it is reflected in the movement patterns of the individual.

Table 4.2: PSI Results ordered by combined scores from sequences of min n=40 and min n=60

Source	**Follower**
Runny nose	WLAN entropy with external APs
Sad-depressed-lonely	Sore throat-cough
Often stressed	Total Bluetooth proximity counts
Communication diversity	Late-night early morning Bluetooth proximity counts
Often stressed	Communication diversity
Often stressed	Late-night early morning Bluetooth proximity counts
Bluetooth entropy with other residents	External WLAN entropy
Runny nose	Total WLAN counts
Often stressed	WLAN entropy with university APs
Bluetooth proximity counts with other residents	External WLAN entropy
Late-night early morning communication	Overall Bluetooth entropy
Sad depressed lonely	Bluetooth entropy



(a) Min Length 40 (b) Min Length 60

Figure 4-10: PSI co-efficients for two sets of sequences based on participant data. List of features: 1=sad-depressed-lonely 2=often-stressed 3=sore-throat 4=runny-nose 5=fever 6=nausea 7= influenza 8=total communication 9=latenight/early morn comm. 10= communication diversity 11=total Bluetooth proximity 12=overall Bluetooth entropy 13=Bluetooth proximity with other residents 14=Bluetooth entropy with other residents 15=late-night/early morn Bluetooth proximity with other residents 16=late-night/early morn Bluetooth entropy with other residents 17=WLAN counts 18=external WLAN counts 19=overall WLAN entropy 20=WLAN entropy with university APs 21=WLAN entropy with external APs

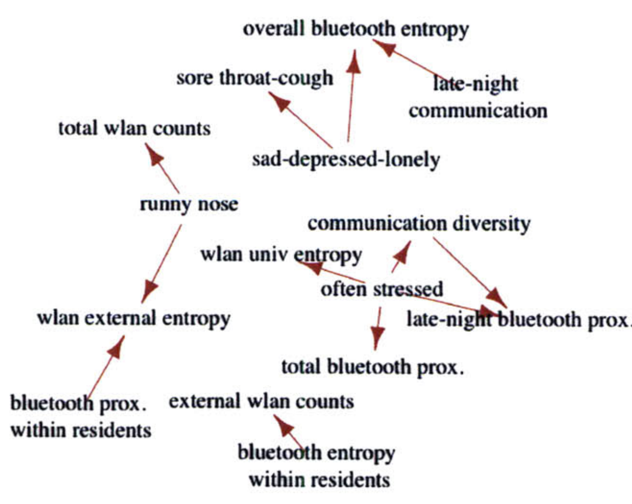


Figure 4-11: Highest-ranked PSI relationships across both data subsets. Directed ties represent temporal flux.

4.4 Discussion

In this chapter, we describe a novel application of sensing face-to-face interactions. We use mobile phones as an active sensing and prediction platform to identify behavior changes reflected in mobile phone sensors, when individuals suffer from common colds, influenza, fever, stress and mild depression. We show that it is possible to determine the health status of individuals using information gathered by mobile phones alone, without having actual health measurements about the subject. Given the pervasiveness of the mobile phone, this opens an avenue for modeling of epidemiological contagion in social networks without the need for medical health reports. We hope our findings can have a direct impact in the fields of spatial and behavioral epidemiology.

There are limitations of this work, that should be improved upon in the future. The statistical tests linking features to symptoms assume that the samples are independent, which is not an entirely correct assumption, since we have repeated measures from a limited set of individuals. This would be improved with a repeated-measures approach, similar to that used in Chapter 3 with political opinions, that will also allow us to understand how much of the observed variation in symptomatic behavior is individual dependent. Our work does not account for confounding behavior changes due to external events, e.g., exams or the end of semester. On the modeling front, the Bayesian classifier used does not incorporate stochastic information about symptoms or behaviors from previous days– this temporal structure could be better modeled using the latent Markov family of models. As mentioned earlier, there is also extensive interest in augmenting compartmental epidemiological models with parameters that represent empirical behavior changes for symptomatic individuals. In addition to classifying observed symptoms from behavior, we are also interested in forecasting when individuals are likely to be sick in future analysis.

The understanding of behavior change in social health will also benefit from the continuous evolution of mobile sensing technologies. The mobile platform used did not support Bluetooth signal strength, which could provide a better measure of physical proximity. WLAN-based location sensing could be replaced with GPS and other location technologies in the future.

Nonetheless, these results have the potential to truly impact healthcare and the medical system. The development of mobile predictive health tools will improve the doctor-patient interaction model, such that health-workers and nurses can use diagnostic information for early detection of conditions, ultimately leading to better healthcare for individuals and lower costs for providers and insurers..

Chapter 5

Face-to-Face Networks and Social Health

What is the role of face-to-face interactions in the diffusion of health-related behaviors- diet choices, exercise habits, and long-term weight changes? In this chapter, we use mobile co-location and communication sensing to model the diffusion of social health-related behaviors via face-to-face interactions.

According to the World Health Organization [114], our society is currently in the midst of a global obesity epidemic, with over a billion overweight and over 300 million clinically obese adults worldwide. This increasing trend is attributed to lifestyle changes in our society, including increased consumption of energy-dense, nutrient-poor foods with high levels of sugar and saturated fats, and reduced physical activity.

Physiological sensing is being increasingly used to study health [71, 94, 80]. However, the focus on social determinants of health is limited. Recent work suggests that obesity and other health-related lifestyle decisions spread through social networks, and in particular long-term face-to-face networks may play an important role. Longitudinal studies based on the Framingham Heart study social network indicate that health-related behaviors from obesity [27] to happiness [45] can spread through social ties. This work has generated greater interest in the study of peer effects on health [46]. Further, the effects of social networks and social support on physical and mental health and the powerful role that they can play in health promotion are well documented [13, 14, 15, 60, 31, 75].

While these studies clearly indicate the importance of social determinants on health, there is limited work studying real-world interactions and their impact on health at a closer level. The latest sensing technologies provide us with the capabilities to collect such fine-grained data and gather new insights, and understand how actual face-to-face interactions lead to changes in behavior. For example, to what extent are eating habits of an individual influenced by those of their spouse, roommate, close-friend or casual acquaintance? Is the diffusion of social behaviors a characteristic of the *person being influenced*, or the *influencer* or simply the *context* of the relationship? Is the underlying causal process *social contagion* or is it practically impossible to disambiguate *homophily* and *confounding* due to the limitations imposed by measurement and estimation techniques? These are some examples of open questions in the context of social contagion of health behaviors. Understanding the role of social network interactions in the ongoing obesity epidemic will enable the design of novel technologies and better interfaces to control adverse spreading, and facilitate social support real-time interventions for positive reinforcement of a healthy lifestyle.

Similar to the analysis of political opinions in Chapter 3, popular epidemiological models do not incorporate social exposure effectively. Cascade and threshold models treat interactions between nodes as a point-estimate, related to tie-strength, and not as dynamic, multi-dimensional interactions. SEIR (susceptible-exposed-infected-recovered) models are an extension of SIR epidemiological models that include an exposure state, but only as an averaged incubation period parameter common to the sample population and not as a unique value for every node. In this paper, exposure to different opinions for individuals as measured using mobile phone sensors is used to explain BMI changes.

The analysis in this chapter is based on social interaction data and health related behavior self-reports, collected as part of study described in Chapter 2. We find that social health behaviors in participants are correlated with the behaviors of peers that they are exposed to over continuous long-term durations. Automatically captured social interactions can be

Table 5.1: Monthly Social Health Survey Instrument

Survey Question	Possible Responses
Current weight and height (weighing scales provided in common areas)	Actual numeric values
Servings of salads consumed, on average per week	0 to 6 servings
Servings of fruits and vegetables consumed, on average per week	0 to 6 servings
Self reported level of healthiness of diet	5-point Likert scale Very Unhealthy to Very Healthy
Instances of aerobic exercise (20 mins or more), on average per week	0 to 6 times
Instances of active sports, on average per week	0 to 6 times

used to estimate exposure for individuals and subsequently predict future BMI changes amongst participants. To better understand the diffusion mechanism, we contrast the role of exposure to various behaviors – peers that are obese, are inactive, have unhealthy dietary habits and those that display similar weight changes in the observation period. These results suggest that it is possible to design self-feedback tools and real-time interventions in the future.

5.1 The Social Health Dataset

In the following analysis, social interaction data for the entire 2009 spring semester is considered. The mobile phone dataset for this period includes of 20609 phone calls, 11669 SMS messages and 2,291,184 scanned bluetooth devices, which includes communication with non-residents as well.

For training labels, participants completed social health related survey instrument for the months of March, April and June, which included the following questions:

As with the survey instrument in Chapter 4, this was also designed under the supervision of an epidemiologist [1]. Participants also listed their close friends and social acquaintances

[1] Dr. Devon Brewer

(binary responses) while completing each monthly survey. The histograms of BMI changes and weight changes for all participants from March to June is plotted in Figure 5-1. Figure 5-1 also shows the Pearson's correlations between the dependent variables.

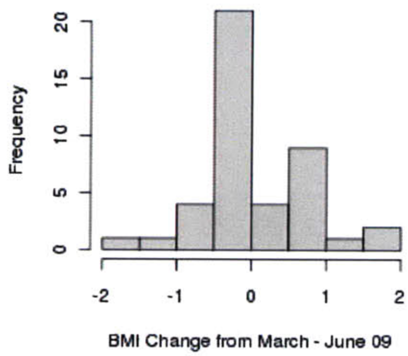
5.2 Results

The effect of social influence on behavior has been well established in the literature, and behavior is at the root of the obesity problem with dietary choices and exercise habits playing a significant role. Christakis and Fowler's recent work [27] on the effect of social ties on obesity has generated greater interest in the study of peer effects on health [46]. While these studies have used networks of self-identified social contacts, they lacked data about real face-to-face interactions that occur on a more regular basis. With the belief that this information is useful in studying this phenomenon, we analyze the effect that friends, acquaintances and face-to-face interactions have on weight change in our study population.

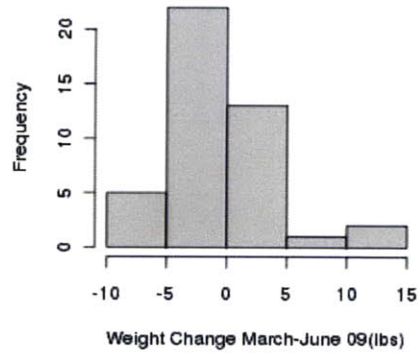
5.2.1 Features that Reflect Exposure

Mobile phone interactions reflect exposure to opinions and behaviors in physical proximity and phone communication. For each individual, we compute exposure as the number of alters reflected the interaction modalities listed below, conditional to specific eating and exercise behaviors described in the next section.

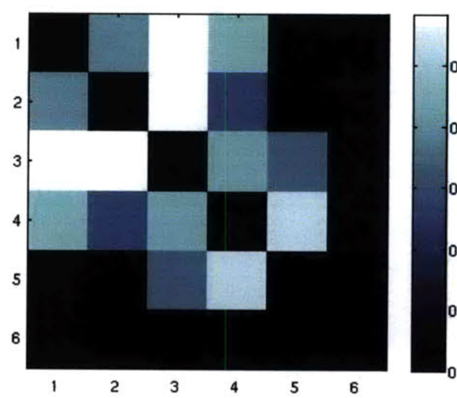
- Total Bluetooth exposure: alters reflected in Bluetooth co-location data
- Late-Night & Early-morning Bluetooth exposure: alters reflected in Bluetooth co-location data between the hours of 9am and 9am the next morning
- Weekend Bluetooth exposure: alters reflected in Bluetooth co-location data only for Saturdays and Sundays



(a) Histogram of BMI changes across all participants from March to June 09. More participants lost weight than gained weight, during the spring semester.



(b) Histogram of Weight changes across all participants from March to June 09 (in lbs)



(c) Pearson correlation coefficient matrix for all dependent self-report variables for June 09, all $p < 0.01$. From top to bottom, left to right, the variables are salads per week ($mean = 1.5, sd = 1.4$), veggies and fruits per day ($mean = 1.9, sd = 1.3$), healthy diet category ($mean = 3.8, sd = 1.1$), aerobics per week ($mean = 2.1, sd = 2.1$), sports per week ($mean = 0.8, sd = 1.5$), BMI change compared to March 09 ($mean = 0.11, sd = 0.68$). BMI Change does not show a strong correlation with either the eating or exercise habits. A healthy diet shows some correlation with the playing sports.

Figure 5-1: Characteristics of self-reported dependent variables related to BMI change, weight change, diet and exercise habits

- Total Phone and SMS exposure: alters reflected in phone communication and SMSs exchanged; both incoming and outgoing communication are included. Duration of calls is not considered
- Weekend Phone and SMS exposure: alters reflected in phone communication and SMS exchanged only for Saturdays and Sundays

In addition to in the following analysis, we also considered the interaction counts for the ego-alter pair as exposure features. However, simply the number of auto-detected alters conditioned by their behaviors outperformed the more complex features in our dataset.

5.2.2 Exposure and its Impact on Body Mass Index

Body Mass Index (BMI) is a commonly used metric to estimate healthy body weight, based on an individual's height. It is equal to the mass in kilograms divided by the square of an individual's height in meters. Individuals with a BMI of 30 or over are categorized as obese while those who have a BMI between 25 and 30 are considered overweight.

An individual's change in BMI is used as the dependent variable, and role of various exposure-based independent variables described above is estimated using linear regression. As mentioned earlier, BMI is a better indicator of healthy weight as it takes into account individual differences in physical stature. Approximately similar results as described below were also obtained while using an individual's weight change (in lbs) as a dependent variable (see Table 5.3.

Exposure to Overweight and Obese Peers

As per the definition of obesity, participants with $BMI >= 30$ are considered obese in our dataset. The independent variables used in this analysis were the number of obese persons that had actual interactions with the individual in question in the form of Total Bluetooth exposure and Late-Night & Early-morning Bluetooth exposure. A significant correlation that explained about 17% of the variation was found as reported in Table 5.2.

For comparison, this analysis was compared to the use of independent variables where as the number of self-reported close friends and social acquaintances that were obese. However, no significant correlation was found between the self-reported independent variables and the dependent variable measuring change in an individual's BMI from March to June 2009.

As the dataset contains only a small number of obese people, hence the linear regression analysis was repeated using exposure features to include both overweight and obese individuals ($BMI >= 25$) as opposed to just obese individuals. In this case, Total Bluetooth exposure and Late-Night & Early-morning Bluetooth exposure to individuals that are overweight or obese explained about 25% of the variation (see Table 5.2).

Similar to above, when this analysis was repeated using exposure to self-reported close friends and social acquaintances that were either overweight or obese, no significant correlation was found between the self-reported independent variables and the dependent variable measuring change in an individual's BMI from March to June 2009.

Exposure to Peers with Unhealthy Diets and Poor Exercise Habits

So far, we have looked at exposure features that focused on the physical aspects indicative of the health of peers. In this section, exposure features indicative of healthy or unhealthy diets and poor exercise habits are considered.

As explained previously, in monthly survey responses, users self-reported their diet on a 7-point Likert scale ranging from "Very Unhealthy" (1) to "Very Healthy" (7). Based on the distribution of responses, a response of 3 or less on this scale is considered as unhealthy eating behavior for this analysis.

BMI Change for the period of March to June 2009 was once again the dependent variable and a similar pattern of results as earlier were observed. Total Bluetooth exposure and Late-Night & Early-morning Bluetooth exposure to peers with unhealthy eating habits explained approximately 17% of the variation in the dependent variable. The exposure to close friends and social acquaintances with unhealthy eating habits did not show significant correlation with BMI change in this period.

We now attempt to understand the role of exposure to individuals who tend to be less physically active. Total activity is the sum of self-reported responses for aerobics per week and sports per week, from the survey responses in the previous section. Based on the distribution of responses, an individual is considered inactive if the total activity is less than or equal to 3. The results were again consistent with the previous section, where Total Bluetooth exposure and Late-Night & Early-morning Bluetooth exposure to peers who were physically inactive explained about 23% of the variance in BMI change.

Exposure to Peers Who Had Substantial Weight Gain in the Same Period

Finally, we looked to see if there was a correlation between BMI change for an individual from March-June 2009 and exposure to peers who gained substantial weight during the same period. Only individuals who gained more than 4 pounds were considered.

The Total Bluetooth exposure and Late-Night & Early-morning Bluetooth exposure features show the most significant correlation to BMI change of all our analyses, and these features together explain about 35% of the variability in the independent variable.

Consistent with the above analysis, exposure to exposure to close friends and social acquaintances who gained weight did not show significant correlation. it is also interesting to note that when the same analysis was repeated using individuals who lost more that 4 pounds, none of the features showed significant correlations.

Table 5.2: Least square regression results: BMI Change

Features	R-Squared	p-value
Exposure to Obese Individuals	0.168	0.009
Exposure to Overweight and Obese Individuals	0.251	0.001
Exposure to Individuals That Eat Unhealthy	0.167	0.009
Exposure to Individuals That Are Inactive	0.246	0.001
Exposure to Individuals That Gained Weight	0.349	<<0.0001

Table 5.3: Least square regression results: Weight Change

Features	R-Squared	p-value
Exposure to Obese Individuals	0.174	0.003
Exposure to Overweight and Obese Individuals	0.259	0.0009
Exposure to Individuals That Eat Unhealthy	0.086	0.06
Exposure to Individuals That Are Inactive	0.252	0.001
Exposure to Individuals That Gained Weight	0.373	<<0.0001

5.3 Discussion

In this chapter, we study the impact that face-to-face interactions have on BMI and weight changes in our community. Our approach allows us to compare the role of social exposure to different types of peers– those that are obese, overweight, have unhealthy dietary habits, and inactive lifestyles.

We find that exposure measured using bluetooth proximity to peers that are overweight or obese and to peers that have unhealthy dietary habits or inactive lifestyles, is correlated with weight changes in an individual as opposed to exposure to close friends and social acquaintances, with the same health-related behaviors. The largest correlations observed are with from exposure to peers with weight gains (> 4 lbs) during the same period. In all cases, we find that exposure measured via self-reported close-friend and acquaintance relationships is not statistically significant. These results are intuitive, and reiterate that we are affected by the behaviors of people that we interact regularly with.

As with the other analyses in this thesis, it should be noted that these results are based on small population of students, and it an open question of how it applies to real-world communities. However, they provide a starting point for the discussion on the importance of studying social networks based on real world interactions. We see statistically significant results indicating that face-to-face interactions might actually have a much larger effect on individual behaviors affecting health. In Chapter 3, a similar effect is seen with political opinions, where social exposure explains the evolution of political opinions better than self-report data.

Chapter 6

Social Ties and Viral Media

In this chapter, we study the role of face-to-face interactions in social relationships, and the spead of viral media (music).

Viral media has become an important part of our popular culture. A growing research area is modeling how ideas and 'memes' propagate though society. Probabilistic topic modeling and information retrieval approaches [17, 77, 111] have been used to measure social diffusion for weblogs and online forums. By tracking hyperlinks and named entities [2], researchers have modeled information cascades through online social media networks. The propagation of information 'memes' [68] has been shown to reflect the daily 'news cycle'. Viral recommendations for products show a long-tailed distribution, and social influence is more relevant in the adoption of niche products [67]. Salganik et. al. [107] measured social influence for online music in eight simultaneous 'worlds' and found highly unpredictable and completely varying rankings of music tracks in each instance, unrelated to the actual quality of music – evidence of social influence of previous listeners.

In this chapter, we study a new dimension of viral media propagation, i.e., the link between face-to-face interactions and the social diffusion process. We show that it is possible to recover social network ties and predict the sharing of music using mobile social interaction features alone. This mechanism seems to work differently for strong versus weak ties. The pilot deployment and analysis in this section preceded, and informed, the primary study detailed in Chapter 2.

6.1 Methodology and Dataset

A variant of the mobile phone platform described in Chapter 2 was deployed with seventeen residents of three floors of a similar undergraduate dormitory for one month, as a pilot experiment to validate platform capabilities before the launch of the primary study. Data was captured from two sources; long-term social interaction data in the form of WLAN IDs, call and sms logs were captured using mobile phones; and in addition, the consumption and propagation of music with timestamps was logged on a central server. Data was discarded from two participants due to logging errors.

A custom-built music player was installed on the mobile phones, which allowed participants to listen, share, rate and search through music tracks provided on thier device. Overall, participants had access to over 1500 independent music tracks from a wide assortment of genres during the course of the study. All music-related events were logged on the server-side, and user-ratings were also captured, as a potential control for track quality. To send a track to any other participant, participants simply click on the 'share' button in the mobile phone application and select the recipient. Before sending, and optionally anytime during playback, participants provided a quality 'rating' for the music track. The player interface and 5-star rating system was similar to that used by the popular iTunes [58] mobile application, and is shown in Figure 6.1.

An important consideration in the experiment design was sourcing content for the music application. Participants become aware of new music [99] from both external media (e.g. radio, internet, etc.) and through recommendations in their personal network. By choosing independent music across different genres, it was possible to ensure that the artists and albums distributed through the study music service were not featured in mass media or were otherwise familiar to the participants. This eliminated potential confounding effects, since the only way participants could be exposed to other artists or music tracks was though

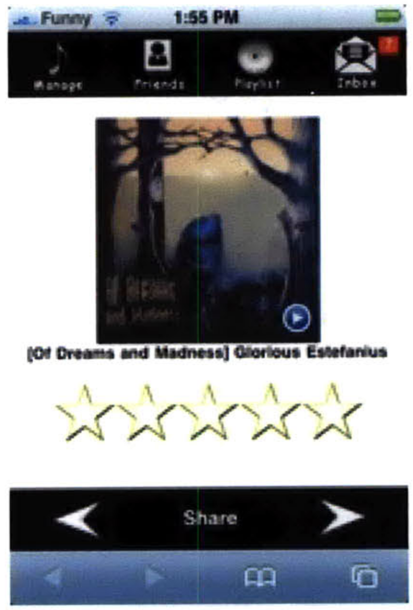


Figure 6-1: UI Screenshot of music player application installed on mobile phones, in addition to the background data-collection scripts. The four top tabs are Manage, Friends, Playlist and Inbox (which shows new tracks received). Users could play, rate and share tracks with other participants. Users were assigned a limited number of music tracks (80) at the start, and when they shared a track, a randomly selected new track was added to their library. Quality ratings (from 5 stars) were required for every music track that was shared. To share a track, user's could select any participants from the experiment using a dropdown (alphabetical) list. A total of 1500 indie tracks from all genres were available in the service. 111 songs were shared and 1234 songs were played during the 30 day period.

their peers– either via explicit sharing or social exposure in common areas etc. To avoid any potential copyright violations, the content was sourced under the Creative Commons license or with explicit permission from the independent artists for experimental use.

The social interaction dataset for the entire month consisted of 3499 unique call events (making a call, receiving a call or missed call), 350 short message (SMS) events (sending or receiving), and 570,000 snapshots of WLAN identifiers. The average length of a call is 122 seconds, 663 call events are during off-peak hours (i.e. after 11pm and before 9am), and 1154 call events are from weekends (i.e. either Saturday or Sunday). Over the entire month, 111 songs were shared and 1234 songs were played by users on mobile phones.

The following features were extracted for every possible dyad of participants and used in the subsequent analysis of relationships and sharing behavior. After eliminating missing data, interactions and dependent variables from 210 possible dyads was used in the following analysis.

Communication features

The communication features used in analysis were:

- Total phone communication and SMS communication, i.e., features that represent the strength of ties
- Off-peak communication (after 11pm and before 8am) and weekend communication (Saturday and Sunday of the week), which represent the nature of ties
- Incoming versus outgoing communication, which has been shown to reflect importance of an individual

Location-based features

For some of the analysis, the following co-location features were also used:

- the Jensen Shannon divergence [29] between distributions of the first hundred most-frequently observed WLAN IDs between individuals. This feature represents the tendency of individuals to visit the same set of locations both within and outside the universty (e.g., classrooms, coffee shops, etc.).

6.2 Social Relationships

Is there a interaction signature that identifies our close friends or casual acquaintances? It has been shown that co-location patterns are closely related to existing social ties, as well as the future ties [42]. In this analysis, we evaluate the the role of phone communication patterns and viral media in predicting social relationships.

User self-assessments of mutual-relationships from the sociometric survey instrument ('friend', 'acquaintance', or 'dont know this person') were used as training labels. The phone communication features were correlated with the user-stated relationship ($R^2 \approx 0.37$, $p < 0.01$). The music sharing between individuals can be considered as an 'active message probe', that reveals the underlying network structure, including the weak-ties, as music propagates through the social network. With this in mind, when the number of music tracks shared between individuals was included as an additional feature, the linear correlation improved slightly ($R^2 \approx 0.44$, $p < 0.01$)

The communication and music sharing features can also be used to discriminate between different types of relationships, i.e., 'friends' versus 'acquaintances'. The total communication and total number of shares between individuals are positively correlated with both 'friends' and 'acquaintance' types of relationships. However, off-peak communication and SMS communication features were positively correlated only with the 'friend relationships, and not with the 'acquaintance relationships. The linear separability between the 'friend class of dyads and the 'non-friend' class is shown in Figure 6.2.

Using phone communication features, it is possible to recover 'friend' relationships from amongst all possible dyads. The classification results for predicting relationships using a cost-sensitive BayesNet classifier and SVM classifier with 5-fold cross validation are shown in Table 6.1. The first row is the 'non-friends class and the second row is the 'friends class. As the classes are unbalanced unbalanced (since only 59 of 210 dyads are self-reported 'friends'), a cost-sensitive approach is used for model training, and the misclassification penalties for the 'friends class : non-friends class were chosen as 3:1. The BayesNet classifier outperforms SVM approach in overall accuracy, but recall for the friends class is slightly better with the

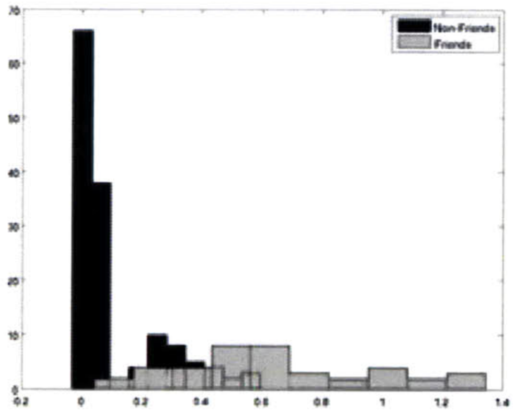


Figure 6-2: Histogram of 'friend relationships vs. values predicted using phone communication features. X-axis values are a linear prediction based on the input features, and the Y axis represents the freqency of relationships in that bin. As seen, the 'friends' class is concentrated towards the left of the figure, while the 'non-friends' class has a flatter distribution, that is shifted right. For a simple linear classifier, 'friends' can be visually separated by drawing a vertical line at x=0.2.

SVM model.

In Table 6.2, we apply the idea that music sharing represents an 'active probe' that reveals the underlying network structure as it propagates, to classification of social ties within the community. With the use of music shares as an additional feature, the classification accuracy improves.

Table 6.1: Relationship classification accuracy with only phone communication features (total, late night, and weekend calls and SMSs). 59 of 210 dyads belonged to the 'friends' class. Ratio of misclassification penalties was 1 (non-friends) : 3 (friends).

Model	Overall Accuracy	Class	Precision	Recall	F-Measure
Cost-sensitive BayesNet with 5-fold CV	87.3%	Non-Friends	0.88	0.97	0.92
		Friends	0.84	0.525	0.646
Cost-sensitive Support Vector Machine with 5-fold CV (polynomial kernel)	83%	Non-Friends	0.88	0.89	0.891
		Friends	0.615	0.6	0.608

Table 6.2: Relationship classification accuracy with phone communication features combined with music sharing features (i.e., 'active probe' of the network). 59 of 210 dyads belonged to the 'friends' class. Ratio of misclassification penalties was 1 (non-friends) : 3 (friends).

Model	Overall Accuracy	Class	Precision	Recall	F-Measure
Cost-sensitive BayesNet with 5-fold CV	90.1%	Non-Friends	0.89	0.98	0.94
		Friends	0.92	0.6	0.73
Cost-sensitive Support Vector Machine with 5-fold CV (polynomial kernel)	89.6%	Non-Friends	0.94	0.92	0.93
		Friends	0.74	0.8	0.77

6.3 Face-to-Face Interactions and Propagation

What is the relationship between viral propagation of music and face-to-face interactions? Do people share music mostly with their close friends, or equally with everyone in the community? For this analysis, communication and location features were correlated with observed sharing behaviour ($R^2 \approx 0.43$, $p < 0.01$). The most significant predictors of music sharing were total calls and total off-peak duration, SMS communication and co-location based on WLAN APs.

It is interesting to note that dyadic sharing behaviour shows higher correlation with automatically captured communication and location features than self-reported relationships ($R^2 \approx 0.37$, $p < 0.01$) for mutually acknowledged friends. This shows that social interactions and exposure automatically captured using mobile phone sensors is a better predictor of the likelihood of two individuals sharing music than their own self-assessments. The likelihood of sharing did not vary significantly with the direction of music sharing for participants.

The music tracks shared with other users during this study can be broken down into two distinct types:

- About 70% of the total shares were between mutually-acknowledged 'friends'. For this pairs of individuals, the correlation of location and communication features with propagation is even higher. This represents diffusion within strong ties.
- The remaining 30% of shares were between strangers or weak ties, i.e., people who are in the same classes or floor, but were not close-friends. For this subset of dyads, the location and communication features are not significantly correlated with sharing. Weak ties [113] are know to play an important role in information diversity.

The two types of sharing highlight the strengths and weaknesses of our approach. Face-to-face interaction features work well to predict transmission probability for cohesive, 'strong' ties. However by themselves, they are less useful in identifying 'weak' social ties or the transmission probabilities associated with such ties. Other approaches like mapping email interactions or social network sites may be more useful. The Author-Recipient-Topic (ART) model and Latent Dirchlet Allocation (LDA) are examples of approaches that have been used to identify roles, relationships and group membership from email interactions [17, 77].

Similar to the classification of social relationships above, it is possible to classify music sharing based on the social interaction pattern between individuals. Instead of using a regressive model, we redistribute the observations of music sharing between participants into distinct classes determined from the underlying distribution. This can be implemented in several ways. With a two-class model (sharing and no-sharing), without using any self-report relationship data, the two-class prediction accuracy using a cost-sensitive Bayesian-Net classifier is 71.5% (For the sharing class, $Precision = 0.69$, $Recall = 0.426$, and $F - Measure = 0.527$).

Since for a majority of possible dyads, no tracks were shared, or less than three tracks were shared, it may be better to use a three-class model for prediction. In this case, the three classes are (a) no sharing (b) less than 3 tracks shared (c) greater than 3 tracks shared. Classification accuracy across all three classes with this approach, using 5-fold cross-validation and a BayesianNet classifier (not cost-sensitive) is 69%.

In the previous section, we find that music sharing patterns can be used as an 'active probe' to recover social relationships. Based on this result, it is possible to reconstruct classification as a heirarchial model, where we first infer 'close friend' ties in the data, and then use the predicted ties as well as communication and location features to predict the music sharing class (2 classes, music shared or not shared). The cross-validated classification accuracy with this approach, using similar methods as above, is 72%.

6.4 Dynamic Bayesian Model of Social Influence

Social influence is the ability of a person to manipulate the propagation process, by inducing others to adopt or reject a new opinion or music track. But how can we estimate the tendency of participants to influence the music preferences of others? Our approach to measuring social influence is based on the latent-state influence model, proposed by Dong [39]. The latent-state influence model is a tractable approximation for hidden Markov modelling of multiple interacting stochastic processes. In a Hidden Markov model of n interacting processes, the number of latent states is product of the number of latent states per process, which implies that an impractical number of model parameters have to be learnt as the number of interacting chains increases. In the corresponding influence model, the number of model parameters is reduced as the latent state distributions for time $t+1$ are based on a linear combination of the latent states for time t. The static weights for this linear combination are the influence values, and reflect the coupling between the interacting Markov chains.

$$P(s_{t+1}^i / s_t^1, s_t^2, ... s_t^n) = \sum_{j=1:n} \alpha_{ij} P(s_{t+1}^i / s_t^j)$$

where s_t^i is the latent state of markov chain i at time t, α_{ij} is the influence of chain i on chain j, for n chains with k latent states per chain.

The forward backward algorithm for latent state estimation and the maximum likelihood algorithm for estimation of model parameters of the influence model are derived from the equivalence between the influence model and corresponding hidden Markov model, and can be found here. [39]f.

The music consumption (i.e., playing songs) and propagation dynamics (i.e., sharing tracks with others) between participants over time can be modelled as n interacting Markov chains using the influence model. Each participant represents a chain, and the observed variable is a function of captured interactions with other participants or media consumption. The inter-chain influences then represent the 'social influence' between the nodes.

Figure 6.4 shows the influence values for sixteen participants based on observed media consumption. The observed variable is the number of times a participant played one of the three most popular tracks. Two latent states ($k = 2$) are assumed per chain and represent the level of 'activation' for the participant. Each chain evolves with a time-step equal to one day. Self-influences are absent from this graph because the playback sequences per person for the three most-popular tracks are sparse. In the future, it would be interesting to explore how these influence values are related to transmission probability.

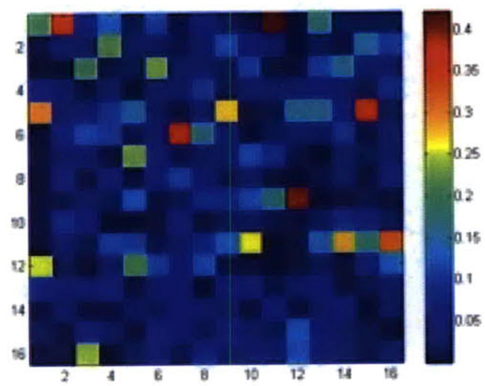


Figure 6-3: Social influence matrix (α_{ij}) for 16 participants based on their music consumption. The observed state for each chain is the number of times the three most popular tracks are played (that day) by the participant. The time-step for all chains in 1-day, with data from 30 days used in the model. Two latent states are assumed per chain and represent the level of activation for the participant. The influence parameters represent inter-chain dynamics for the entire 30-day period.

6.5 Discussion

This pilot deployment was our first step towards designing the year-long, community study discussed in Chapter 2. This smaller study provided insight on the design of the mobile platform, user adoption and usage, and the different types of social diffusion questions that could be answered using this approach.

We find that passive phone communication features (i.e., calling records and SMS logs) are correlated with stated relationships. These linear correlations increase when we consider music sharing as active probes in the network. Overall, it is possible to correctly identify close to 90 percent of self-reported close friends in the community, using both phone communication and music sharing patterns.

With regard to the sharing of music, about 70 percent of the sharing behavior between dyads is highly correlated with captured social interaction features, which is consistent with the theory of influence due to social cohesion and strong ties. The remaining 30 percent of sharing behavior is between weak ties or strangers. Music consumption (playback) for individual nodes can be used to estimate the 'social influences' between people.

Chapter 7

Conclusion

In this thesis, we quantify the role of face-to-face interactions and behaviors in the contexts of different social diffusion phenomena, based on over 320,000 hours of behavior and interaction data. With political opinions, we use social exposure to reveal patterns of dynamic homophily, and show that it is possible to identify discussants and estimate future opinions from face-to-face interactions alone. In the context of epidemiology, we find that changes in interaction patterns are a powerful predictor of physical and mental health symptoms and syndromes, and can be used to explore the temporal relationship between behavior change and physical and mental symptoms. With diet and obesity, we estimate the effects of different types of social exposure on BMI changes over an entire semester. Finally, we show the face-to-face interaction and phone communication patterns can be used to identify social ties, predict the sharing of music, and compare the relative 'social influence' of individuals.

We now discuss various implications of this work, viz., the underlying cognitive mechanism that may explain why face-to-face interactions seem to play such an important role, the potential privacy risks, and real-world applications that such tools may enable.

7.1 The Cognitive Link Between Social Exposure and Behavior Change

In the contex of political opinions and social health, we find that exposure to different types of individuals is much more important than self-report data, for explaining the adoption of a new opinion or behavior. Beyond the effect of informant inaccuracy, perhaps this is related to the construct of social signaling in our behavior. It has been shown that mood and emotion contagion occour in social settings and at the workplace [12]. Pentland and others [101, 21] have proposed that such social signaling is an important component of our social decision-making abilities, and is effectively captured in vocal prosody and physical gestures and movement. It is hence, an open question, whether such social learning is also an avenue for adoption of opinions and habits, which may explain why face-to-face interactions and social sensing play such an important role.

7.2 Privacy Implications

As seen in the preceeding chapters, modeling human social behavior holds a lot of promise for our digital society. However, the legal and ethical boundaries around data ownership and user privacy with regard to such data are still unclear. We briefly discuss current legislative standards towards using such data in both workplace and consumer settings [74].

At the workplace, behavior modeling approaches have the potential to increase productivity, but there is also potential for disproportionate loss of privacy. With regard to data ownership, the European Union (EU) has more stringent data privacy policies than the United States. According to the EU Directive on Privacy and Electronic Communications for public-sector employees, storage of individuals communication data is usually only permitted if user have provided thier explicit consent. Similarly, the privacy rights of private-sector employees are protected by Article 8 of the European Convention on Human Rights. The United States, on the other hand, has a relatively lax policy concerning employee monitoring. As long as employees do not have a reasonable expectation of privacy (i.e., they have been previously informed of possible monitoring), companies are permitted to access electronic communications, such as e-mail, stored on company servers and also allowed to monitor employees through phone, computer, and video surveillance. Hence, to deploy such social sensing technologies at the workplace, explicit employee consent is an important consideration.

On the consumer front, social interaction data in various forms is collected by mobile phone operators, banking institutions and other agencies. Although the respective companies own this data, the Federal Communications Commission (FCC), through the Telecommunications Act of 1996 Section 702, maintains fairly strict requirements on how the data is used and shared. Companies are required to ensure the privacy of the data, and may only disclose the data to business affiliates who provide secondary services necessary for the telecommunications services being provided. Upon written request from the customer, they are also required to disclose the customer's information to any party specified, which is an important clause for researchers and developers interested in building applications with such interaction data. Banks maintain similar privacy policies with their interaction data, as required by Regulation P: Privacy of Consumer Financial Information (12 CFR 216) of the Federal Reserve Board.

Anonymization cannot be guaranteed for social network data by the removal of personal identifiers alone. Backstrom et. al. [9] have shown passive and active attacks whereby it is possible to identify original participants using embedded nodes. Another approach for passive de-anonymization, based on using a known auxiliary graph, was shown by Narayan [86].

There is also risk that sensing technologies may unintentionally harm non-participants. Consider an example a bluetooth transciever to log physical proximity. Individuals who are not part of the experiment may object that their unique Bluetooth identifiers are automatically logged by the system. However the data being collected is in some sense, public information and non-users are free to set their Bluetooth devices to non-discoverable mode (the default setting on most new phones and laptops, where Bluetooth communication is active but the unique identifier is not continually broadcast). The legal precedent of such unintentional monitoring can be compared to that for phone calls, i.e., in the case of Smith v. Maryland, it was shown that there is no legitimate expectation of privacy with respect to information such as the recipient and duration of a call. The Bluetooth argument is possibly analogous, i.e., if you are broadcasting your Bluetooth identifier; there is little expectation of privacy with respect to your Bluetooth identity.

During our experimental deployment, two types of user reactions with regard to participant privacy were observed. A majority of residents trusted the IRB evaluation process and felt that any privacy risks related to participant had been alleviated. A small minority of dormitory residents, however, were actively concerned and declined to participate in the deployment. Fortunately, this minority represented a spatially isolated section of the building. Privacy-related quotes from some concerned non-participants, and some participant responses, are given in Table 7.1.

These conversations illustrate various privacy perspectives related to our work. It is important to outline technical, public policy and legal guidelines for data collection, storage and analysis.

Table 7.1: Specific quotes from non-participants participants who were concerned about possible privacy implications of the study. In some cases, participants directly responded to the concerns of their peers, and these comments are also reported below.

On the data collection process:
"privacy aside, I personally have problems with people who don't live here leaving things in the dorm. Especially on a long-term basis, especially without permission, especially if they're trying to 'study' us."

On the topic of participant privacy:
"So, just because I do something in a lounge where people can see it doesn't make it legal for people to film me without permission and use it in a study. ... See also, the Fourth Amendment."
"You do realize the Fourth Amendment applies to what the government can't do to citizens, not what citizens can't do to other citizens?"

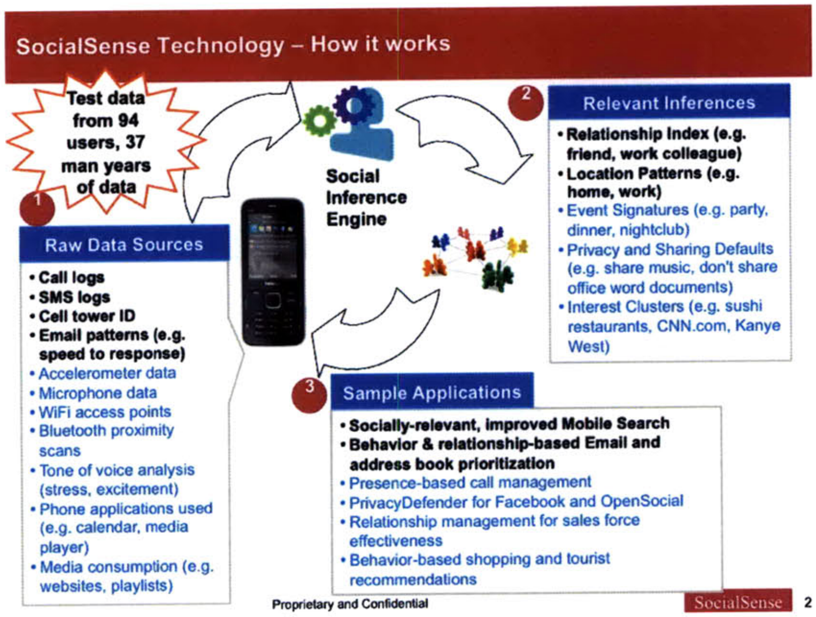
On the topic of informed consent:
" 'A quick poll of a cross-section of the dorm' does not constitute permission. A significant fraction of xxx residents have a problem with this. Please do not place any devices in xxx"
"I actually have a really big objection to this. Your presenting this as fait accompli is really problematic

On the topic of bluetooth co-location sensing:
"What's the big deal? I've been recording all blue tooth activity from the ceilings of public spaces in the dorm for the past 9 years and posting all the data on xxxx. If you are concerned with who is recording your bluetooth devices, this is the perfect opportunity to change your privacy settings; god-forbid you walk past someone somewhere sometime who is recording you"

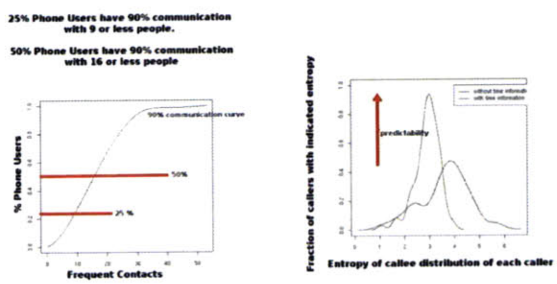
7.3 Real-World Mobile Applications

The future of such computational social science technologies lies in using the models and knowledge, to give feedback to users, and influence public policy decisions, to induce people towards positive behaviors and a better life. Hence, we have prototypes several mobile phone applications that perform social inferences based on user behavior, and provide feedback to the user directly.

In Figure 7-1, some of these examples are shown. Figure 7-1(a) is a general inference mechanism and API to enable 'socially aware' applications. A specific example of an application built on top of such APIs, is an automated 'speed-dialer' that predicts whom you are likely to call next based on past location and phone communication patterns, as shown in 7-1(b). Finally, we expect that the epidemiology related results could be used as part of a predictive webservice for patients and doctors, and a screenshot of such a webservice is shown in Figure 7-1(c).



(a) A conceptual social inference engine and 'API' to export different aspects of the user's behavior for 3rd party applications



(b) An example application, that predicts whom the user is likely to call next, based on

(c) An behavior-based feedback example UI based on the epidemiology results, to give users and health professionals feedback on thier chronic conditions

Figure 7-1: Sample (behavior-based) feedback applications

Appendix A

Supporting Information for Political Analysis

A.1 ANOVA analysis for Dynamic Homophily

A.1.1 Political Interest for All Residents

Table A.1: Mean, SD and n for Dynamic Homophily estimates, as a function of period (for three conditions: Baseline, Final Debate and Election Day)

Period	Mean	SD	Samples
Baseline	1.61	0.941	60
Final Debate	1.22	0.76	60
Election Day	1.45	0.84	60

Table A.2: Source Table for (Period) Repeated-Subjects ANOVA

Source	Df.	SS.	MS	F-value	p
Period	2	4.32	2.18	8.49	0.00035 ***
Residuals	118	30.35	0.25		

Hence the null hypothesis (i.e., three conditions have equal means) is rejected. The residual normality assumption required for ANOVA holds true (See figure).

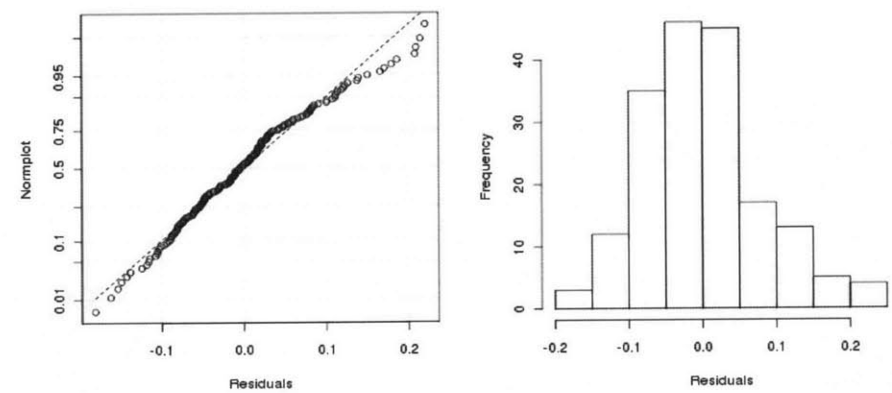


Figure A-1: Residuals normality assumption is validated for ANOVA analysis of Political Interest dynamic homophily for all participants

A.1.2 Preferred Party for All Residents

Table A.3: Mean, SD and n for Dynamic Homophily estimates, as a function of period (for three conditions: Baseline, Final Debate and Election Day)

Period	Mean	SD	Samples
Baseline period	1.922	1.457	61
Final Debate period	1.788	1.325	61
Election Day period	1.83	1.27	61

Table A.4: Source Table for (Period) Repeated-Subjects ANOVA

Source	Df.	SS.	MS	F-value	p
Period	2	0.582	0.2911	0.87	0.42
Residuals	118	40.2	0.33		

Since the p-value is NOT significant, the null hypothesis (i.e., three conditions have equal means) cannot be rejected. TThe residual normality assumption required for ANOVA holds true (See figure).

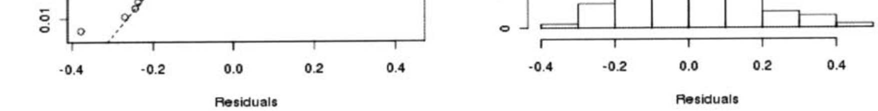


Figure A-2: Residual normality assumption is validated for ANOVA analysis of Preferred Party dynamic homophily for all participants

A.1.3 Liberal-Conservative for All Residents

Table A.5: Mean, SD and n for Dynamic Homophily estimates, as a function of period (for three conditions: Baseline, Final Debate and Election Day)

Period	Mean	SD	Samples
Baseline period	2.05	1.22	61
Final Debate period	1.644	1.13	61
Election Day period	1.80	1.08	61

Table A.6: Source Table for (Period) Repeated-Subjects ANOVA

Source	Df.	SS.	MS	F-value	p
Period	2	5.29	2.647	6.26	0.0026 **
Residuals	118	50.74	0.422		

Since the p-value is significant, the null hypothesis (i.e., three conditions have equal means) is rejected. TThe residual normality assumption required for ANOVA holds true (See figure).

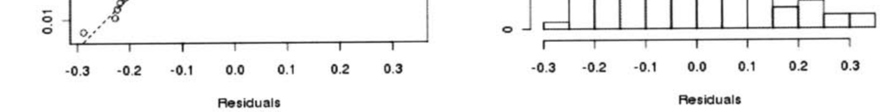


Figure A-3: Residuals normality assumption is validated for ANOVA analysis of Liberal-Conservative dynamic homophily for all participants

A.1.4 Political Interest for Freshmen Only

Table A.7: Mean, SD and n for Dynamic Homophily estimates, as a function of period (for three conditions: Baseline, Final Debate and Election Day)

Period	Mean	SD	Samples
Baseline period	1.69	0.925	18
Final Debate period	1.25	0.7	18
Election Day period	1.31	0.81	18

Table A.8: Source Table for (Period) Repeated-Subjects ANOVA

Source	Df.	SS.	MS	F-value	p
Period	2	2.03	1.01	3.43	0.043 *
Residuals	34	10.04	0.29		

Since the p-value is significant, the null hypothesis (i.e., three conditions have equal means) is rejected. The residual normality assumption required for ANOVA holds true (See figure).

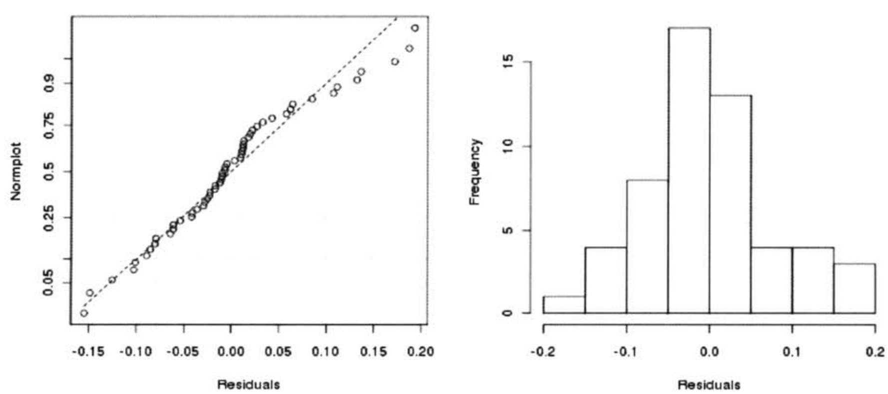


Figure A-4: Residuals normality assumption is validated for ANOVA analysis of Political Interest dynamic homophily for Freshmen only

A.2 Dynamic Homophily graphs for the entire year period

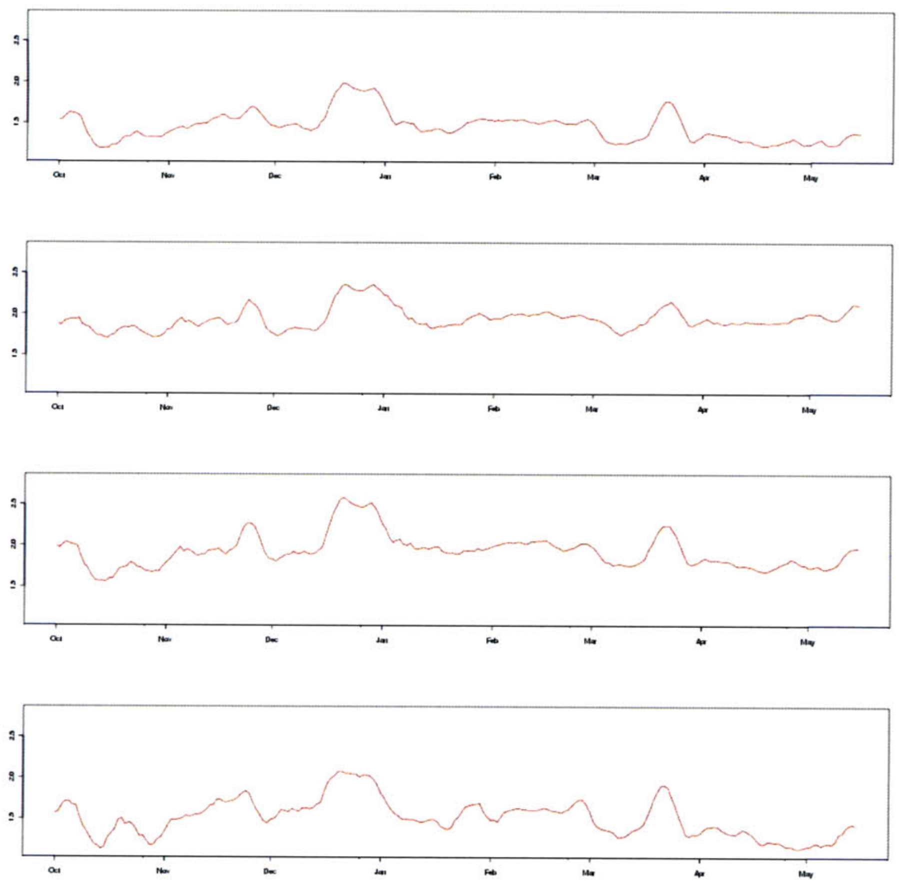


Figure A-5: Dynamic Homophily (moving average) for the entire year for four different cases based on bluetooth proxmity. From top to botton: (a) political interest for all residents (b) preferred party for all residents (c) liberal-conservative for all residents (d) political interest for freshmen only

Bibliography

[1] T. Abdelzaher, Y. Anokwa, P. Boda, J. Burke, D. Estrin, L. Guibas, A. Kansal, S. Madden, and J. Reich. Mobiscopes for human spaces. *IEEE Pervasive Computing - Mobile and Ubiquitous Systems*, 6(2), 2007.

[2] E. Adar, L. Zhang, L. Adamic, and R. M. Lukose. Implicit Structure and the Dynamics of Blogspace . In *Workshop on the Weblogging Ecosystem, 13th International World Wide Web Conference*, 2004.

[3] N. Agarwal and H. Liu. *Modeling and Data Mining in Blogosphere*, volume 1. Morgan & Claypool Publishers, 2009.

[4] B. Aral and V. Alstyne. Information, technology and information worker productivity: Task level evidence. *International Conference on Information Systems (best paper)*, 2006.

[5] B. Aral and V. Alstyne. Productivity effects of information diffusion in networks. *International Conference on Network Science 2007, International Conference on Information Systems 2007*, 2007.

[6] S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence based contagion from homophily driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences (PNAS)*, 106(51), 2009.

[7] S. Avancha, A. Baxi, and D. Kotz. Privacy in mobile technology for personal healthcare. *Submitted to ACM Computing Surveys*, 2009.

[8] C. Avery and P. Zemsky. Multidimensional uncertainty and herd behavior in nancial markets. *Am. Econ. Rev.*, 88:724–748, 1998.

[9] L. Backstrom, C. Dwork, and J. Kleinberg. Wherefore Art Thou R3579X? Anonymized Social Networks, Hidden Patterns, and Structural Steganography. *WWW Conference*, 2007.

[10] J. P. Bagrow and E. M. Bollt. Local method for detecting communities. *Phys. Rev. E*, 72(4):046108, Oct 2005.

[11] Barrat A., Barthlemy M., Pastor-Satorras R., and Vespignani A. The Architecture of Complex Weighted Networks. *Proceedings of the National Academy of Sciences*, 101:3747–3752, 2004.

[12] S. G. Barsade. The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior. *Administrative Science Quarterly*, 47(4):644–675, 2002.

[13] L. Berkman. Assessing the physical health effects of social networks and social support. *Annual Review of Public Health*, 5(1):413–432, 1984.

[14] L. Berkman. The role of social relations in health promotion. *Psychosomatic Medicine*, 57(3):245, 1995.

[15] L. Berkman, T. Glass, I. Brissette, and T. Seeman. From social integration to health: Durkheim in the new millennium. *Social Science & Medicine*, 51(6):843–857, 2000.

[16] Bernard H.R., Killworth P., Kronenfeld D., Sailer L. The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Reviews in Anthropology*, 1984.

[17] D. M. Blei, A. Y. Ng, M. I. Jordan, and J. Lafferty. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:2003, 2003.

[18] D. Boyd and N. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer Mediated Communication*, 13(1):210, 2007.

[19] D. Brewer and C. Webster. Forgetting of friends and its effects on measuring friendship networks. *Social Networks*, 21(4):361 – 373, 2000.

[20] D. Brockmann. Human mobility and spatial disease dynamics. *Review of Nonlinear Dynamics and Complexity*, 2009.

[21] M. Buchanan. Behavioural Science: Secret signals. *Nature*, 457, 2009.

[22] R. Burt. Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, 1987.

[23] T. Choudhury. *Sensing and Modeling Human Networks.* PhD thesis, Massachusetts Institute of Technology, 2003.

[24] T. Choudhury. Characterizing social networks using the sociometer. *Association of Computational Social and Organizational Science*, 2004.

[25] T. Choudhury and S. Basu. Modeling Conversational Dynamics as a Mixed Memory Markov Process. *Advances of Neural Information Processing Systems (17)*, 2004.

[26] T. Choudhury, M. Philipose, D. Wyatt, and J. Lester. Towards activity databases: Using sensors and statistical models to summarize peoples lives. *IEEE Data Eng. Bull*, 29(1):49–58, 2006.

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

[28] N. Christakis and J. Fowler. The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21):2249, 2008.

[29] Christopher Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

[30] J. Cohen, P. Neumann, and M. Weinstein. Does preventive care save money? health economics and the presidential candidates. *The New England journal of medicine*, 358(7):661, 2008.

[31] S. Cohen, B. H. Gottlieb, and L. G. Underwood. Social relationships and health. *American Psychologist*, 59:676–684, 2004.

[32] S. Cohen and G. Williamson. Stress and infectious disease in humans. *Psychological Bulletin*, 109(1):5–24, 1991.

[33] E. Cohen-Cole and J. M. Fletcher. Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis. *BMJ*, 337, 2008.

[34] J. Coleman, E. Katz, and H. Menzel. *Medical Innovation: A Diffusion Study*, volume 311. Bobbs-Merill Company, IN, 1966.

[35] V. Colizza, A. Barrat, M. Barthelemy, A. Valleron, and A. Vespignani. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. *PLoS Medicine*, 4(1):95, 2007.

[36] S. Currarini, M. Jackson, and P. Pin. An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4):1003–1045, 2009.

[37] D. M. Musher. How Contagious are Common Respiratory Track Infections. *New England Journal of Medicine*, 2003.

[38] V. den Bulte and G. Lilien. Medical Innovation Revisited: Social Contagion versus Marketing Effort. *American Journal of Sociology*, 2001.

[39] W. Dong and A. Pentland. Multi-sensor Data Fusion Using the Influence Model. *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks*, 2006.

[40] P. Drummond and B. Hewson-Bower. Increased psychosocial stress and decreased mucosal immunity in children with recurrent upper respiratory tract infections. *Journal of psychosomatic research*, 43(3):271–278, 1997.

[41] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[42] N. Eagle, A. Pentland, and D. Lazer. Inferring Social Network Structure Using Mobile Phone Data. *Proceedings of National Academy of Sciences*, 106(36):15274–15278, 2009.

[43] J. Epstein, D. Goedecke, F. Yu, R. Morris, D. Wagener, and G. Bobashev. Controlling pandemic flu: the value of international air travel restrictions. *PLoS One*, 2(5), 2007.

[44] R. Fletcher and S. Fletcher. *Clinical epidemiology: the essentials.* Lippincott Williams & Wilkins, 2005.

[45] J. Fowler and N. Christakis. Dynamic Spread of Happiness in a Large Social Network: longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, 337(dec04 2):a2338, 2008.

[46] J. Fowler and N. Christakis. Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais. *Journal of health economics*, 2008.

[47] Freund, Y. and Schapire, R.E. A short introduction to boosting. *Japonese Society for Artificial Intelligence*, 5:771–780, 1999.

[48] Friedkin N.E. *A Structural Theory of Social Influence.* Cambridge University Press, 1998.

[49] G. Nolte et al. Robustly Estimating the flow of direction of information in complex social systems. *Physical Review Letters*, 100, 2008.

[50] M. Girvan and M. E. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.*, 99:7281, 2002.

[51] M. Gonzalez, C. Hidalgo, and A.-L. Barabasi. Understanding Individual Human Mobility Patterns. *Nature*, 453:779–782, 2008.

[52] Granovetter M. Threshold models of collective behavior. *American Journal of Sociology*, 83:1420–1443, 1978.

[53] F. Guo, S. Hanneke, W. Fu, and E. P. Xing. Recovering temporally rewiring networks: A model-based approach. In *In ICML07*, 2007.

[54] M. Handcock, A. Raftery, and J. Tantrum. Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170:301–354, 2007.

[55] Huckfeldt, R. and Sprague, J. Discussant Effects on Vote Choice: Intimacy, Structure and Interdependence. *The Journal of Politics*, 53:122–158, 1991.

[56] D. R. Hunter and M. S. Handcock. Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, 15:565–583, 2006.

[57] J. Imboden, A. Canter, L. Cluff, and et al. Convalescence from influenza: a study of the psychological and clinical determinants. *Archives of Internal Medicine*, 108(3):393, 1961.

[58] A. Inc. itunes website. `http://www.apple.com/itunes/`.

[59] S. Jamieson. Likert scales: how to (ab) use them. *Medical Education*, 2004.

[60] I. Kawachi and L. Berkman. Social ties and mental health. *Journal of Urban Health*, 78(3):458–467, 2001.

[61] Kempe D., Kleinberg J., Tardos E. Maximizing the Spread of Influence in a Social Network. In *Proceedings of KDD 2003,*, Washington DC., 2003.

[62] G. Kossinets and D. J. Watts. Empirical analysis of evolving social networks. *Science*, pages 88–90, 2006.

[63] M. Laibowitz, J. Gips, R. Aylward, A. Pentland, and J. Paradiso. A sensor network for social dynamics. *Proceedings of the 5th international conference on Information processing in sensor networks*, page 491, 2006.

[64] P. Lazarsfeld and R. K. Merton. Friendship as a Social Process: A Substantive and Methodological Analysis. *Freedom and Control in Modern Society, Morroe Berger, Theodore Abel, and Charles H. Page, eds.*, 1954.

[65] D. Lazer and A. Friedman. The Network Structure of Exploration and Exploitation. *Administrative Science Quarterly*, 52(4), 2007.

[66] D. Lazer, B. Rubineau, N. Katz, C. Chetkovich, and M. A. Neblo. Networks and political attitudes: Structure, influence, and co-evolution. Working Paper Series rwp08-044, Harvard University, John F. Kennedy School of Government, Sept. 2008.

[67] J. Leskovec, L. Adamic, and B. Huberman. The Dynamics of Viral Marketing. *ACM Conference on Electronic Commerce (EC)*, 2006.

[68] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the Dynamics of the News Cycle. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2009.

[69] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins. Microscopic evolution of social networks. *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 462–470, 2008.

[70] H. Liu, J. Salerno, and M. Young. *Social Computing, Behavioral Modeling and Prediction.* Springer, 2008.

[71] B. Lo, S. Thiemjarus, R. King, and G. Yang. Body sensor network–a wireless sensor platform for pervasive healthcare monitoring. In *The 3rd International Conference on Pervasive Computing.* Citeseer, 2005.

[72] A. Madan and A. Pentland. VibePhones: Socially Aware Mobile Phones. In *Intl. Symposium of Wearable Computing*, 2006.

[73] A. Madan and A. Pentland. Modeling Social Diffusion Phenomena Using Reality Mining. In *AAAI Spring Symposium on Human Behavior Modeling*, 2009.

[74] A. Madan, B. Waber, M. Ding, P. Kominers, and A. Pentland. Reality Mining: The End of Personal Privacy? *Proceedings of 1st Engaging Data Forum*, 2009.

[75] M. Marmot and R. Wilkinson. *Social determinants of health.* Oxford University Press, 2005.

[76] W. Mason, A. Jones, and R. Goldstone. Propagation of innovations in networked groups. *27th Conference of Cognitive Science Society*, pages 1419–1424, 2005.

[77] A. Mccallum, A. Corrada-emmanuel, and X. Wang. The author-recipient-topic model for topic and role discovery in social networks: Experiments with enron and academic email. Technical report, 2004.

[78] D. G. McNeil. Models Projections for Flu Miss Mark by Wide Margin. *New York Times*, June 2009.

[79] M. McPherson, L. Smith-Lovin, , and J. M. Cook. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415–444, 2001.

[80] A. Milenkovic, C. Otto, and E. Jovanov. Wireless sensor networks for personal health monitoring: Issues and an implementation. *Computer Communications*, 29(13-14):2521–2533, 2006.

[81] S. Milgram. The Familiar Stranger: An Aspect of Urban Anonymity. *The Individual in a Social World: Essays and Experiments, Longman*, 1977.

[82] E. Miluzzo, C. Cornelius, A. Ramaswamy, T. Choudhury, A. Campbell, and Z. Liu. Darwin Phones: The Evolution of Sensing and Inference on Mobile Phones. *Mobisys 2010*, 2010.

[83] N. Mishra, R. Schreiber, I. Stanton, and R. Tarjan. Clustering social networks. *Lecture Notes in Computer Science*, 4863:56–67, 2007.

[84] MIT Media Lab. Social evolution project. `http://social.media.mit.edu`.

[85] S. Moturu. *Quantifying the Trustworthiness of Social Media Content: Content Analysis for the Social Web.* Lambert Academic Publishing, 2010.

[86] A. Narayanan and V. Shmatikov. De-anonymizing Social Networks. *IEEE Security and Privacy*, 2009.

[87] M. Newman. Associative mixing in networks. *Phys Rev Lett*, 89, 2002.

[88] D. Nickerson. Is voting contagious? Evidence from two field experiments. *American Political Science Review*, 2008.

[89] G. Nolte, A. Ziehe, N. Krmer, F. Popescu, K. robert Mller, I. Guyon, D. Janzing, and B. Schlkopf. Comparison of granger causality and phase slope index.

[90] D. Olguin, P. Gloor, and P. A. Capturing Individual and Group Behavior Using Wearable Sensors. In *AAAI Spring Symposium, Human Behavior Modeling*, Palo Alto, CA, 2009.

[91] D. Olguin Olguin, P. Gloor, and A. Pentland. Wearable Sensors for Pervasive Healthcare Management. *3rd International Conference on Pervasive Computing Technologies for Healthcare*, 2009.

[92] D. Olguin Olguin and A. Pentland. Social Sensors for Automatic Data Collection. In *Proceedings of the Fourteenth Americas Conference on Information Systems*, page 171, 2008.

[93] D. Olguin Olguin, B. Waber, T. Kim, A. Mohan, K. Ara, and A. Pentland. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics-B*, 2009.

[94] N. Oliver and F. Flores-Mangas. HealthGear: a real-time wearable system for monitoring and analyzing physiological signals. *Proc. Body Sensor Networks*, pages 61–64, 2006.

[95] O.S. Miettinen. *Theoretical Epidimiology: principles of occurence research in medicine.* Wiley New York, 1985.

[96] P. Domingos. MetaCost: A General Method for Making Classifiers Cost-Sensitive. In *Fifth International Conference on Knowledge Discovery and Data Mining, KDD-99*, 1999.

[97] P. Elliott, et al. *Spatial Epidimiology.* Oxford University Press, 2000.

[98] G. Palla, A. L. Barabasi, and T. Vicsek. Quantifying social group evolution. *Nature*, pages 664–667, 2007.

[99] Paul Felix Lazarsfeld and Bernard Berelson and Hazel Gaude. *The people's choice: how the voter makes up his mind in a presidential campaign.* Columbia University Press, 1944.

[100] A. Pentland. Socially aware computation and communication. In *Proceedings of the 7th international conference on Multimodal interfaces*, page 199. ACM, 2005.

[101] A. Pentland. To Signal is Human. *American Scientist*, 2010.

[102] H. Rang, M. Dale, J. Ritter, and P. Moore. *Pharmacology 5th Edition.* Churchill Livingstone, Edinburg, 2003.

[103] M. Richardson and P. Domingos. Markov logic networks. In *Machine Learning*, page 2006, 2006.

[104] G. Robins, T. Snijders, P. Wang, M. Handcock, and P. Pattison. Recent developments in exponential random graph (p*) models for social networks. *Social Networks*, 29(2):192–215, 2007.

[105] E. M. Rogers and E. Rogers. *Diffusion of Innovations, 5th Edition.* Free Press, 1962.

[106] D. Roy, R. Patel, P. DeCamp, R. Kubat, M. Fleischman, B. Roy, N. Mavridis, S. Tellex, A. Salata, J. Guinness, M. Levit, and P. Gorniak. The human speechome project. *Lecture Notes in Computer Science*, 4211:192, 2006.

[107] M. J. Salganik, P. S. Dodds, and D. J. Watts. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market . *Science*, 311, 2006.

[108] R. J. Shiller. *Irrational exuberance.* Princeton University Press, 2000.

[109] M. Steyvers, P. Smyth, M. Rosen-Zvi, and T. Groffiths. Probabilistic author-topic models for information discovery. In *The Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 306–315, 2004.

[110] F. Wang, K. Carley, D. Zeng, and W. Mao. Social computing: From social informatics to social intelligence. *IEEE Intelligent Systems*, pages 79–83, 2007.

[111] X. Wang, N. Mohanty, and A. Mccallum. Group and topic discovery from relations and text. In *In Workshop on Link Discovery: Issues, Approaches and Applications, in conjunction with the 10th International ACM SIGKDD Conference*, pages 28–35, 2005.

[112] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press: London, 1994.

[113] D. Watts and S. Strogatz. Collective dynamics of small world networks. *Nature*, pages 443–445, 1998.

[114] World Health Organization Global strategy Report on Diet, Physical Activity and Health. http://www.who.int/dietphysicalactivity/publications/facts/obesity/en/.

[115] D. Wyatt, T. Choudhury, and J. Bilmes. Discovering Long Range Properties of Social Networks with Multi-Valued Time-Inhomogeneous Models. *Proceedings of AAAI*, 2010.

[116] D. Wyatt, T. Choudhury, J. Bilmes, and J. Kitts. Towards Automated Social Analysis of Situated Speech Data. *Proceedings of Ubicomp*, 2008.

[117] R. Yirmiya, Y. Pollak, and et. al. Illness, Cytokines, and Depression. *Annals-New York Academy of Sciences*, 917:478–487, 1999.

[118] D. Zeng, F. Wang, and K. Carley. Guest Editors' Introduction: Social Computing. *IEEE Intelligent Systems*, pages 20–22, 2007.