

Socio-Technical Network Analysis from Wearable Interactions

Katayoun Farrahi^{*1}, Rémi Emonet^{†2} and Alois Ferscha^{‡1}

¹Pervasive Computing Institute, JKU Linz, Altenberger Strasse 69, A-4040 Linz, Austria

²Idiap Research Institute, Martigny, Rue Marconi 19, CH-1920 Martigny, Switzerland

Abstract

Wearable sensing platforms like modern smartphones have proven to be effective means in the complexity and computational social sciences. This paper draws from explicit (phone calls, SMS messaging) and implicit (proximity sensing based on Bluetooth radio signals) interaction patterns collected via smartphones and reality mining techniques to explain the dynamics of personal interactions and relationships.

We consider three real human to human interaction networks, namely physical proximity, phone communication and instant messaging. We analyze a real undergraduate community's social circles and consider various topologies, such as the interaction patterns of users with the entire community, and the interaction patterns of users within their own community. We fit distributions of various interactions, for example, showing that the distribution of users that have been in physical proximity but have never communicated by phone fits a gaussian. Finally, we consider five types of relationships, for example friendships, to see whether significant differences exist in their interaction patterns. We find statistically significant differences in the physical proximity patterns of people who are mutual friends and people who are non-mutual (or asymmetric) friends, though this difference does not exist between mutual friends and never friends, nor does it exist in their phone communication patterns.

Our findings impact a wide range of data-driven applications in socio-technical systems by providing an overview of community interaction patterns which can be used for applications such as epidemiology, or in understanding the diffusion of opinions and relationships.

*farrahi@pervasive.jku.at

†remi.emonet@idiap.ch

‡ferscha@pervasive.jku.at

1 Introduction

Agent based computer models are used to model complex systems of interacting agents. They are believed to be a key in revealing the hidden processes underlying our complex global systems [5]. Agent based models have already been applied to a range of applications ranging from financial markets [17] to artificial intelligence and game theory [6, 19]. In order to obtain realistic agent-based simulations, we need to exploit information from real human data. Reality Mining performs real-time knowledge mining from mobile phones as sensors, potentially providing many insightful details about human behavior. Large scale mobile phone data collections can be used to capture collective and social human behaviors ubiquitously both over large populations and long durations of time. In this paper, we perform a statistical analysis of a large mobile phone data collection to reveal patterns of human interactions.

It is important to have some understanding of how humans interact in the real world, not only for agent based modeling, but also for epidemiology, psychology, and sociology. If we consider the particular example of agents, we can apply the data-driven statistics on the interaction information to define the social network structure when creating a community of agents and use this information for information spreading. Secondly, we can consider the distributions of real-life human interaction information in order to evaluate how realistically an agent based system is mimicking the real world. Similar logic is applicable to many other applications.

Understanding the different types of human interactions and how they relate to each other in a community is the goal of this paper. We are considering interactions which can be sensed with a mobile phone. We

consider the Bluetooth sensor to give us physical proximity data. We also consider mobile phone communication data, including both call and SMS networks. We are investigating the relationships between these social interaction networks to discover if they are highly correlated, and to investigate what underlying properties can be found.

The contributions of this paper are as follows.

1. We analyze the overall connectivity of a real-life community, consisting of 72 individuals over a 9 month period to determine user interaction patterns with the entire community in terms of physical proximity, call, and instant messaging.
2. We study the collective behavior of users within their own community (as opposed to the entire community, mentioned previously) to see how much overlap exists between different types of interaction. We consider call versus physical proximity interactions and messaging versus call networks to understand the sensor data correlations.
3. We consider the monthly variations in both cases of overlapping networks and find that after 6 months of time, the cumulative average in the overlapping call to Bluetooth network approaches the overall mean, whereas this occurs only after 4 months in the SMS to call network, potentially indicating the duration of data required to study these networks.
4. We fit the real data based on distributions, generalizing this community's interactions for use in applications such as agent-based systems.
5. Finally, we consider five types of relationships, for example friendships, to determine whether the interaction types of these groups differ. One result of this analysis is that we find the physical proximity patterns of people who are mutual friends statistically significantly differs from those who are non-mutual friends, however their call patterns do not differ significantly.

2 Related Work

This work best fits in the body of work on human to human interaction analysis from wearable sensors, especially Reality Mining [2]. Several different approaches and applications of group detection have been proposed based on Bluetooth proximity sensors. SocialNet [20] is an application that runs on portable devices and uses patterns of collocation sensed by Bluetooth over time, to infer shared interests between

users. The Group Discovery Co-location (GDC) algorithm [12], was developed to combine user meeting frequency and duration for group detection and was validated on one month of smart phone data carried by 141 students. GroupUS [1] is a probabilistic relational model based on Latent Dirichlet Allocation for group detection, more specifically, interaction type and social context detection. This model was validated on a set of 40 individuals over the course of a year. A different approach to group detection has been presented by Roggen et al. [18], where the wearable sensor of interest is the accelerometer and the approach is to cluster similar patterns in physical movement. This approach is of interest where the subjects are assumed to be moving and not likely feasible for indoor applications. Though various approaches have been investigated and group detection can be useful for a variety of applications, our focus in this paper is not on the particular task of group detection though we are also considering human interaction data obtained by Reality Mining.

Human interaction dynamics have been studied using other types of wearable sensing badges, typically combining several sensor features to obtain short-term, personal information about face-to-face interactions. The Sociometric badge [15] is an improved form of the Communicator badge [14], designed to automatically capture individual and collective patterns of behavior. The sociometric badge collects communication patterns of individuals by monitoring physical proximity, physical activity levels, and conversational times to quantify social interactions, group behavior, and organizational dynamics. It has also been used in the context of personal healthcare management [13]. The personal digital assistants designed by Wyatt et al. [21], contain 8 different sensors to capture human conversations over time. Human interaction analysis based on multi-modal sensing badges can provide rich details about human interactions as they are very personal in nature, however they are not feasible for large scale studies both due to the privacy considerations of the subjects and due to the requirement that subjects perpetually carry an additional device.

A few other works relating to human interaction data analysis focus on the data mining aspect [3, 7]. Farrahi et al. [3] consider the problem of human activity data mining based on probabilistic topic models to mine interaction patterns based on both location and physical proximity features. They additionally evaluate their methods by considering prediction of missing data. Hwang et al. [7] discover group patterns from moving object trajectory data obtained by mobile devices. Their model is formulated based on a trajectory approach to discover group patterns. Finally, the most

closely related work to ours is by Palla et al. [16], where the authors investigate the time dependence of overlapping communities on a large scale, uncovering basic relationships characterizing community evolution. The focus in [16] is on evolution and the time dynamics of a small group, whereas our paper targets a series of other issues including the relationship between individuals' social networks, and the connection of an individuals' interaction patterns with their relationships.

3 Sensor Data and Characteristics

The data we are using for this study was collected at MIT, and has previously been investigated for the purpose of measuring the shifts in individual habits, opinions, health, and friendships [9, 10, 11]. The participants were given Windows Mobile 6.x mobile phones to use as their primary phones collecting their call records, SMS logs, Bluetooth co-location and WLAN data. Relationship data was collected via monthly surveys, in which individuals identified their relationships based on five categories with the other participants of the study. The community chosen for data collection was a tightly knit community and the campaign started at the beginning of the academic semester, throughout which relationships were expected to develop. The experiment was designed as a long-term longitudinal study with eighty residents of an undergraduate residence hall that served as the primary residential, cooking, social activity, and sleeping quarters for the residents. The participants in the study represent eighty percent of the total population of this hall, and most of the remaining twenty percent were spatially isolated.

During the 2009 academic year, the dataset consists of 270 days, 3.15 million scanned Bluetooth devices, 61,100 logged call data records, and 47,700 logged SMS messages. Out of these events, 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. We provide some details about the features which we use in this paper.

- Interaction with Bluetooth sensors. The software scanned for Bluetooth wireless devices in proximity every 6 minutes. In our analysis we do not force symmetry as this can be an additional source of error; we take as input the data collected directly by the sensors and consider undirected ties.
- Communication. The software logged call and SMS details on the device every 20 minutes, including information about missed calls and calls not completed.

- Relationships. Participants were given monthly relationship surveys with a list of all the other individuals in the dataset. For each monthly survey, participants identified other residents that were their close friends, political discussants, social acquaintances, and whether they shared facebook and blog information, identical to those used here [8], since the goal of the data collection was to analyze diffusion patterns in opinions and relationships [9].

3.1 The Interaction Data

Considering an interaction network where the users are represented by nodes and interaction events are represented by links between nodes, then we can compute the number of links in the network for both the Bluetooth interaction and phone communication of the real data collected. We only consider the number of events occurring over the duration of a month. Figure 1 shows the link count of the data over time. We can see a jump in links at October, which is when most of the users started using the devices. There is a decrease in the Bluetooth interactions in December, due to the exam period and Christmas holidays. The phone communication however is mostly constant over time.

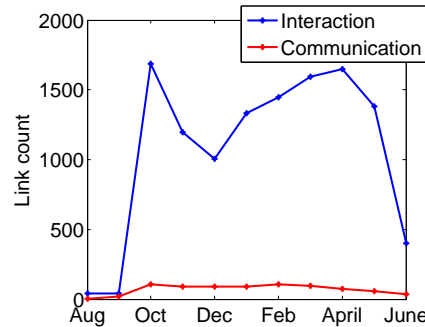


Figure 1. Link count for the interaction and communication networks. Number of unique pairs of users with at least one event logged by the sensors within the duration of the month. The interaction network shows most of the data collection occurred between October to June, which is the mobile sensed data interval chosen for experiments.

3.2 The Relationship Data

Now we consider a network such that users are represented by nodes however links represent relationships between users. We can then count the number of links

over time for the five types of relationships. These relationships are, close friendships, socialize, political discussants, share facebook photos, and share blog information. In Figure 2 we visualize a summary of the relationship data captured where symmetric ties are counted as 1 link and asymmetric ties are also counted as 1 link. We can see there are fewer overall close friendships and political discussants perhaps due to being more personal in nature. There are more pairs of individuals that share facebook information than other types of relationships.

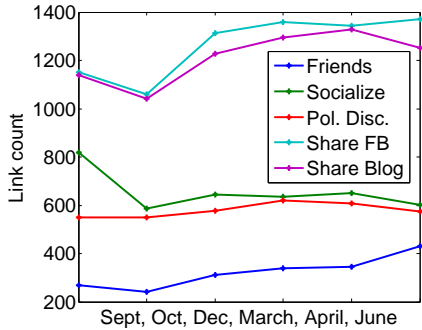


Figure 2. Friendship link count over time for five types of relationships, friends, socialize, political discussants, share facebook, share blog. Close friendships are probably the most personal relationship with the fewest links. Sharing facebook details has the most number of links at all times. All relationship types increased throughout the school semester, except for socialize, which decreased.

4 Experiments and Results

4.1 Smartphone Data

For the experiments and analysis in this section, we are considering 72 of the students from the undergraduate community due to lack of data for the other 8 students. We only include students that have at least 3 data entries recorded of any sort over the duration of the study. The time period we are considering is from 10.2008 to 06.2009 since in Figure 1 we noted this was the period with the significant amount of data collected. The data considered is not symmetrized. In this paper, we do not consider call directions (incoming versus outgoing). Missed calls are also considered as they are recorded in the call logs. All of the statistics are computed on the number of events logged and not on the duration of the events logged.

We now consider the overall ‘amount of interaction’ within the community, and potentially consider whether the community effect is indeed apparent in the data collected. This analysis should reveal the overall participant integration based on three sensors, further validating the sensed data corresponds to a community. We compute for each user the number of unique individuals in the community from which an interaction was ever detected over the number of individuals overall in the community (i.e. 72). The average of this measure shows the percentage of Bluetooth, call, and SMS interaction within the overall community. In Figure 3, we show the distribution of this measure on a log-log scale for the three types of sensor data. The log-log scale is chosen to show the SMS and Bluetooth on the same scale. We see that the Bluetooth network is highly connected. Several users have close to 100% connectivity within the community, which means their Bluetooth sensors sensed almost everyone else’s device at least once within the 9 months of the study and they have interacted with almost everyone else at least once. The SMS network is the least connected. The call network contains more connections than SMS. It could be that since these students are living together, they communicate less by phone. For the call distribution, we see that the probability of having greater than 10% call interaction within the community drops significantly. Less than 10% call interaction within the community, however, is highly probable.

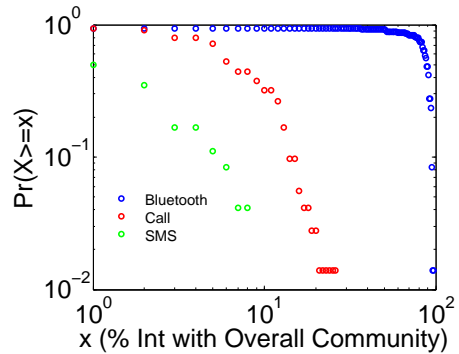


Figure 3. The distribution of user interaction connectivity within the community of 72 students, considering call activity, sms activity, and Bluetooth interactions. The distribution is plot on a log-log scale. The x-axis corresponds to the percentage of interaction, more specifically, the number of unique individuals from the community with which at least one interaction was recorded per user averaged over the 9 months over the total population.

We have already considered the connectivity in the overall network. Next, we investigate the connectivity within a user’s own networks by looking for correlations and relationships between the interaction and communication network (Section 4.2) and the call and SMS network (Section 4.3).

4.2 Social interactions: phone vs Bluetooth

We consider the relationship between the students who communicate with each other with their mobile phones and the students whom are in physical proximity, captured by Bluetooth. We only consider interaction with other students in the data collection campaign. The phone communication in this section is considered to be all of the phone activities, including calls and SMS. The implications of these results can be useful for applications relating to information spreading or disease spreading where the overall correlations between a community’s calling and face to face interaction activities need to be modeled.

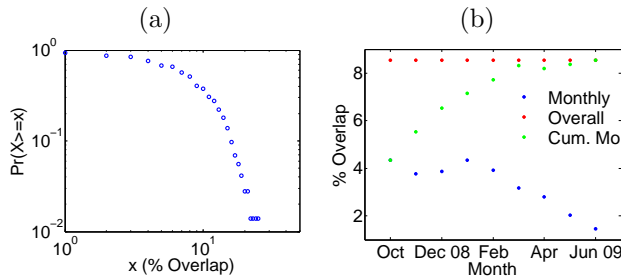


Figure 4. (a) Distribution of % overlap between the overall communication and Bluetooth networks on a log-log scale. (b) Monthly variations in the % overlap between the communication and Bluetooth networks averaged over all users.

In Figure 4 (a), we consider the percentage of overlap between the communication network, A , and the Bluetooth network, B . For each student, we compute the percentage of overlap within their community, which can be found using the relation $\cap(A, B) / \cup(A, B)$. We plot the distribution of the average user’s overlap on a log-log scale. The maximum overlap is 25.7%, meaning one user communicates by phone with about a quarter of the people they interacted with in the community. The minimum is no overlap (0%), meaning some users never called anyone within the community of people they interacted with. The average overlap between the mobile phone network and the Bluetooth interaction network is 8.55%. We can

see from Figure 4(a), that the probability of having less than 10% overlap is quite high, and drops significantly for great than 10% overlap. There is never more than 25.7% overlap between these networks.

We are now interested to know over what duration should data be collected and considered in order to get an understanding of the relationships between the interaction networks. In Figure 4(b) we plot the monthly variations in the overlap between the phone and Bluetooth networks. We plot three curves: (1) ‘overall’ is the mean overall average overlap between A and B computed over the 9 months from (a), (2) ‘monthly’ is only considering the interactions which occurred over the specified month, and (3) ‘cum mo’ is the cumulative monthly, and is all the interactions which occurred up until the specified month. We can see in a specific month, the overlap between these networks is much lower than the average. With the data we are using, we can see that the cumulative monthly approaches the overall monthly after about 6 months. This indicates that with our data, we approach the mean after about 6 months. Obviously more experiments on other data collections are necessary to obtain more conclusive results, though these results raise interesting questions and reason for further analysis.

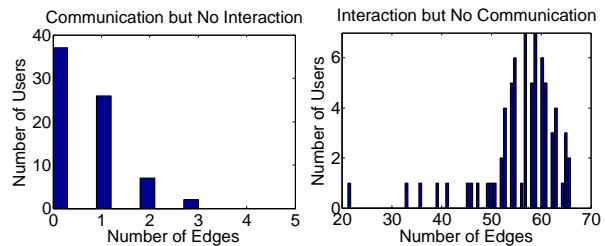


Figure 5. Histogram of the number of users which communicate by phone but never interact based on Bluetooth and vice versa.

In Figure 5, we plot the distribution of users which would not be accounted for in a network of individuals which both communicate by phone and have Bluetooth proximity recorded. The number of users which communicated by phone but never were in physical proximity is quite low. This never occurs for about half of the users. This occurred in one case for about 25 users. It may be that in these 25 cases, the individuals were either never in physical proximity and only communicate by phone, or when they have been in physical proximity, either they were not carrying their mobile phones or their Bluetooth has been turned off. As expected, the number of users which interacted but never communicated by phone is quite high.

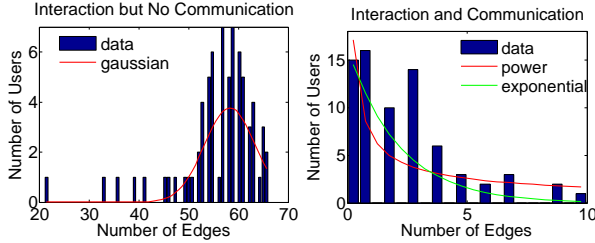


Figure 6. Data fit results. The distribution of users which have physical proximity but no communication data is fit to a first order gaussian; the distribution of users which both have physical proximity and communication is fit to both a power and exponential distribution.

In order to simulate the behavior of this community for other examples, for example to generate realistic agent behavior in a community, we require a distribution to sample from. Therefore, we fit the data to basic distributions. In Figure 6, we fit the distribution of the community which interacted but never communicated to a first order gaussian, $y = a \cdot \exp(-((x - b)/c)^2)$. We find a good fit to be $a = 3.4, b = 58.24, c = 6.8$ resulting in a root mean square error (rmse) of 1.44. We fit the distribution of the community which both interacted and communicated by phone to an exponential and power distribution. The exponential distribution is fit as $a \cdot \exp(b \cdot x)$ with $a = 16.3, b = -0.47$, $\text{rmse} = 3.92$; the power distribution is $a \cdot x^b$ with $a = 7.11, b = -0.63$, resulting in $\text{rmse} = 4.09$.

4.3 Phone interactions: calling vs SMS

Similar experiments performed in Section 4.2 are now conducted for the phone communication networks to find the relations between the call versus SMS data between users' own interaction networks. In Figure 7, we see the call and SMS networks are much more interconnected, several users have up to 100% overlap. Several users also have very little overlap, and this is mostly due to lack of SMS activity and not lack of phone activity as we can see from Figure 8. Figure 7(b) shows that after 4 months of activity, the cumulative average approaches the overall average, indicating potentially less data is necessary for studying the phone interaction correlations than the phone versus the physical proximity correlations. We also observe that in both networks (Figures 7(b) and 4(b)), the monthly curve decreases over time. Further analysis is necessary to make conclusive results but we speculate that people interact with many individuals at the

beginning of the semester and then proceed to interact more with a closer group as time progresses.

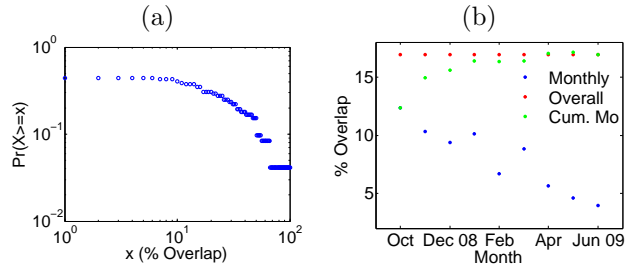


Figure 7. (a) Distribution of % overlap between the call and SMS networks on a log-log scale. (b) Monthly variations in the % overlap between call and SMS networks averaged over all users.

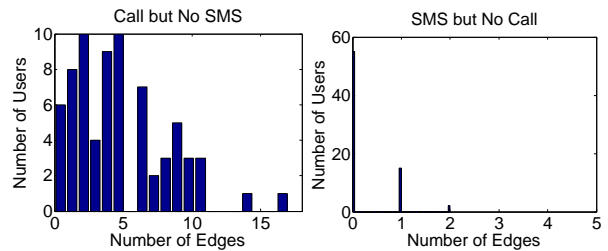


Figure 8. Histogram of the number of users which communicate by phone call but never SMS and vice versa.

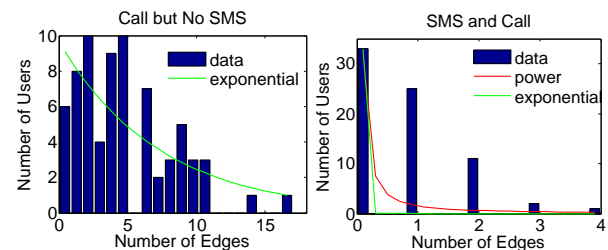


Figure 9. Data fit to basic distributions. The distribution of users which have call but no messaging data is fit to a gaussian; the distribution of users which interact with both types of phone communication is fit to both a power and exponential distribution.

We fit the phone communication distributions, and visualize the results in Figure 9. We found the best fit

for the distribution of people who call but never SMS to be an exponential with $a = 9.64, b = -0.14$, $rmse = 2.45$. The fit for people who both call and SMS was best as a power distribution with $a = 1.54, b = -1.32$, $rmse = 6.39$.

4.4 Relationship-Specific Interactions

We are interested to know if there is a difference in the interaction and communication patterns of people with strong relationships versus people without relationships. This analysis is a starting point for work in prediction, i.e. in order to determine if certain relationships can be predicted from interaction patterns and vice versa. We would expect there to be a significant difference between these categories, for example, people who consider themselves to be mutual friends likely interact more than people who are not friends. To look at this in more detail, we pick three groups of relationship types where we consider the first month and the last month of the campaign time frame to obtain the groups: (1) **mutual** friends throughout the study, (2) **non-mutual** friends throughout the study, (3) **never**, or not friends throughout the study.

Note that the example with ‘friends’ was given, but this relationship type is then replaced with ‘socialize’, ‘political discussants’, and so on. We want to see if the interaction patterns of these three relationships differ significantly or not. We take the three set of user pairs’ interactions (number of events, unnormalized computed over the entire 9 months) to see if these distributions vary significantly across these groups based on f-tests. First (in Table 1, we present the results with the 3 groups combined to determine which relationship types contain significantly different interactions. Following this analysis in Table 2, we compute pairwise statistics on the significant cases to ‘filter’ the results presented to the interesting cases. The p-values for all possible relationship types resulting from f-tests are shown in Table 1. The statistically significant cases (considering a threshold of $p = 0.05$) are shown in bold, and occur for the friendship and the share facebook relationships only in the case of interaction with Bluetooth and not phone. None of the other relationships have statistically significant differences, which is surprising but is likely due to the tight community of students who are living together.

The p-value resulting from the anova test tells us if at least one of the groups differs significantly from the others. We then consider each possible combination of relationships to see which two groups differ. These results are shown in Table 2. Again, there are only two cases which differ significantly, which is again surprising. In terms of friendship, mutual versus non-mutual

Table 1. f-test results (p values) comparing the difference in the interaction patterns for the 3 groups: mutual friends, non-mutual friends, and never friends. Statistically significantly different distributions are shown in bold and only occur for the friendship and share Facebook relationships.

Relationship Types	Bluetooth Interactions	Phone Activity
Friendship	0.0469	0.663
Socialize	0.2673	0.8118
Political Discussants	0.1263	0.8448
Share Facebook	6x10-7	0.7235
Share Blog	0.1263	0.8448

Table 2. t-test results (p values) comparing the pairwise cases with significant differences in Table 1. We see the groups which had statistically significant differences in bold. They only occur for physical proximity and never for phone activity.

Friendship	Bluetooth Interactions	Phone Activity
Mutual vs Non-mutual	0.0284	0.1589
Non-mutual vs Never	0.12	0.4666
Mutual vs Never	0.0578	0.6064
Share Facebook		
Mutual vs Non-mutual	0.198	0.712
Non-mutual vs Never	0.0875	0.7506
Mutual vs Never	4.45x10-8	0.4571

friends have significantly differing physical proximity patterns, but not phone interactions. The relationship share facebook photos has very statistically significantly differing physical proximity patterns for the groups mutual versus never. None of the other cases have significantly differing distributions. We would have expected the phone activity patterns to differ more than the physical proximity patterns, but this was never the case.

5 Conclusions

Human behavior modeling based on Reality Mining promises to reveal a better understanding of collective social behavior, particularly human to human interactions. In this paper, we consider three types of interactions, sensed by mobile phones, namely physical proximity, call and SMS. Considering a real dataset of 72 individuals over 9 months, we investigate the rela-

tionship between these networks, the interaction patterns of users' with the entire community, of users' within their own community, and the potential predictive power of relationships based on human interactions. Our results reveal several interesting and unexpected behaviors, including the finding that mutual friends versus non-mutual friends' physical proximity patterns are statistically significantly different, though their phone communication patterns are not. Additionally, people who mutually share facebook photos versus people who mutually never share facebook photos have significantly differing physical proximity patterns, but people who asymmetrically share facebook photos have no significantly differing interactions with symmetric sharers, nor non sharers. The results of this analysis are aimed towards data-driven applications, such as agent based modeling, and raise several directions for future studies. For further contribution to data-driven applications, the results presented should be validated on other datasets with varying community structures, and on datasets with no community structure. We would then know whether these results are generalizable, and see to what extent the community-targeted data collection has an effect on the results.

6 Acknowledgements

This work is supported under the FP7 ICT Future Enabling Technologies program of the European Commission under grant agreement No 231288 (SOCIONICAL).

References

- [1] T. Do and D. Gatica-Perez. Groupus: Smartphone proximity data and human interaction type mining. In *IEEE International Symposium on Wearable Computers (ISWC)*, San Francisco, USA, June 2011.
- [2] N. Eagle, A. Pentland, and D. Lazer. Inferring social network structure using mobile phone data. *Proc. of the National Academy of Sciences (PNAS)*, 106(36):15274–15278, 2009.
- [3] K. Farrahi and D. Gatica-Perez. Probabilistic mining of socio-geographic routines from mobile phone data. *IEEE Journal of Selected Topics in Signal Processing (J-STSP)*, 4(4):746–755, 2010.
- [4] A. Ferscha et al. Pervasive socio-technical fabric. *Proceedia Computer Science*, 7(0):88–91, 2011.
- [5] D. Helbing, S. Bishop, P. Lukowicz, and R. Conte. Futurict european flagship proposal. <http://www.futurict.ethz.ch/FuturIcT>.
- [6] P. Holme and G. Ghoshal. Dynamics of Networking Agents Competing for High Centrality and Low Degree. *Physical Review Letters*, 96(9):098701+, 2006.
- [7] S.-Y. Hwang, Y.-H. Liu, J.-K. Chiu, and E.-P. Lim. Mining mobile group patterns: A trajectory-based approach. In *Proc. of Advances in Knowledge Discovery and Data Mining (PAKDD)*, pages 713–718, 2005.
- [8] D. Lazer, B. Rubineau, N. Katz, C. Chetkovich, and M. A. Neblo. Networks and political attitudes: Structure, influence, and co-evolution. Working paper series, Harvard University, Sep 2008.
- [9] A. Madan, M. Cebrián, D. Lazer, and A. Pentland. Social sensing for epidemiological behavior change. In *Ubicomp*, Copenhagen, Denmark, 2010.
- [10] A. Madan, K. Farrahi, D. Gatica-Perez, and A. Pentland. Pervasive sensing to model political opinions in face-to-face networks. In *Pervasive*, San Francisco, USA, 2011.
- [11] A. Madan, S. Moturu, D. Lazer, and A. Pentland. Social sensing: Obesity, unhealthy eating and exercise in face-to-face networks. In *Proc. of ACM Wireless Health*, San Diego, USA, 2010.
- [12] S. Mardenfeld, D. Boston, S. J. Pan, Q. Jones, A. Iamntichi, and C. Borcea. Gdc: Group discovery using co-location traces. In *Social Computing Applications*, August 2010.
- [13] D. Olguin, P. Gloor, and A. Pentland. Wearable sensors for pervasive healthcare management. *PCT Healthcare*, 2009.
- [14] D. O. Olguin, J. A. Paradiso, and A. Pentland. Wearable communicator badge: Designing a new platform for revealing organizational dynamics. In *IEEE International Symposium on Wearable Computers (ISWC)*, Oct. 2006.
- [15] D. Olguin et al. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 39(1), Feb. 2009.
- [16] G. Palla, A. Barabasi, and T. Vicsek. Quantifying social group evolution. *Nature*, 446:664–667, Apr. 2007.
- [17] M. Raberto, S. Cincotti, S. M. Focardi, and M. Marchesi. Agent-based simulation of a financial market. *Physica A: Statistical Mechanics and its Applications*, 299(1-2):11, 2001.
- [18] D. Roggen, M. Wirz, D. Helbing, and G. Tröster. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *Networks and Heterogeneous Media*, 2011.
- [19] G. Szabó and G. Fáth. Evolutionary games on graphs. *Physics Reports*, 446(4-6):97–216, July 2007.
- [20] M. Terry, E. D. Mynatt, K. Ryall, and D. Leigh. Social net: using patterns of physical proximity over time to infer shared interests. In *Extended Abstracts on Human Factors in Computing Systems (CHI)*, pages 816–817, 2002.
- [21] D. Wyatt, T. Choudhury, and H. Kautz. Capturing spontaneous conversation and social dynamics: A privacy sensitive data collection effort. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Hawaii, USA, 2007.