Characterizing Social Networks using the Sociometer

Tanzeem Choudhury
Intel Research Seattle
tanzeem.choudhury@intel.com

Alex Pentland
MIT Media Laboratory
sandy@media.mit.edu

Abstract

Knowledge of how groups of people interact is important in many disciplines, e.g. organizational behavior, social network analysis, knowledge management and ubiquitous computing. Existing studies of social network interactions have either been restricted to online communities, where unambiguous measurements about how people interact can be obtained, or have been forced to rely on questionnaires, or diaries to get data on face-to-face interactions. Survey-based methods are error prone and impractical to scale up. This paper describes our work in developing a computational framework to model face-to-face interactions within a community. We have integrated methods from speech processing and machine learning to demonstrate that it is possible to extract information about people's patterns of communication, without imposing any restriction on the user's interactions or environment. Furthermore, we analyze some of the conversational dynamics and present results that demonstrate distinctive and consistent turn-taking styles for individuals during conversations. Finally, we present results that show strong correlation between a person's turn-taking style during one-on-one conversations and the person's role within the network.

Contact:
Tanzeem Choudhury
Intel Research Seattle
1100 NE 45th Street (6th Floor)
Seattle, WA 98103

Tel: 1-206-545-2507
Fax: 1-206-633-6504
Email: tanzeem.choudhury@intel.com

Key Words: Social network analysis, organization behavior, machine learning

Acknowledgement: Thanks to Brian Clarkson for his help with the design of the sociometer. We would also like to thank Sumit Basu and Rich DeVaul for helpful discussions, and Ken Fishkin and James Landay for their feedback.

Support: This work has been partially supported by Center for Bits and Atoms NSF research grant number NSF CCR-0122419.

Characterizing Social Interactions using the Sociometer

Tanzeem Choudhury and Alex Pentland

Our decision-making is influenced by the actions of others around us. How people communicate also has an effect on information diffusion (Gladwell 2000). Although people heavily rely on email, telephone and other virtual means of communication, high complexity information is primarily exchanged through face-to-face interaction (Allen 1997). Knowledge of people's communication networks can also be used in coordinating collaboration between group members and for deeper understanding of organizational behavior. We believe that wearable sensor data combined with pattern recognition techniques can play an important role in sensing and modeling physical interactions.

Sensing and Modeling Human Networks

Prior work on modeling face-to-face network using sensors uses proximity (Want 1992), a weak approximation of the actual communication network. Our focus is to model the network based on conversations that take place within a community. To model real-world interactions, we need to collect data from real-world scenarios. We have conducted an experiment at the MIT Media Lab, where 23 people agreed to wear the *sociometer*. The sociometer is an adaptation of a wearable data acquisition board, designed by the electronic publishing and the wearable computing groups at the Media lab (Gerasimov 2002). It stores the following information - (i) identity of people wearing the sociometer (IR sensor - 17Hz) and (ii) speech information (microphone-8KHz). During the experiment the users wore the device both indoors and outdoors for six hours a day for 11 days. The participants were a mix of students, faculty and administrative support staff who were distributed across different floors of the building and across different research groups.

To protect the user's privacy we only extract speech features (spectral peaks, energy etc.) and never process the content of the speech. This is sufficient for our purposes, as we are interested in who people talk to and how, and not necessarily what they talk about. At the end of the data collection phase we conducted a survey to gauge the acceptance of the sociometer among the users: Figure 1 summarizes the results and shows users were generally accepting of the sociometer.

The sociometer interfered with normal interactions	Never (7)	Few times (14)	Most of the time (0)	Always (0)
Will wear an audio recording device only if transcription is not done	Don't care (6)	Yes (12)	Do not like wearing it, even if transcription is not done (3)	Would never wear any data collection device (0)
How comfortable was the sociometer?	Did not notice it was there (5)	Comfortable (7)	Somewhat Uncomfortable (9)	Uncomfortable (0)

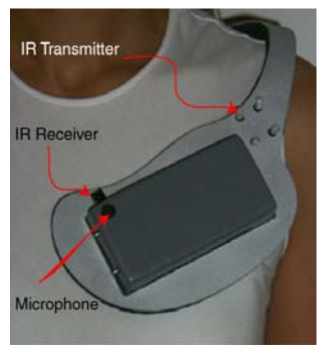

Figure 1: (a) Exit survey - number in parenthesis indicates how many users chose that response. (b) the "sociometer"

Detecting Conversations

To detect conversations, we need to reliably segment speech regions from the raw audio. As the first step, we extract spectral features that discriminate well between speech and non-speech regions (Basu 2002). A hidden Markov model is trained to detect voiced/unvoiced regions using the features. This method works very reliable even in noisy environment with less than 2% error at 10dB SNR. The downside of this approach is that all speech and not just the user's speech are detected. However, we can use the energy of the speech signal to segment the user's speech from the rest. When two people are nearby and are talking, although it is highly likely that they are talking to each other, we cannot say this with certainty. Results presented in (Basu 2002) demonstrate that we can detect whether two people are in a conversation by relying on the fact that the speech of two people in a conversation is tightly synchronized. We can reliably detect when two people are talking to each other by calculating the mutual information of the two voicing streams, which peaks sharply when they are in a conversation as opposed to talking to someone else. This measures works very well for conversations that are at least one minute in duration. During the data collection stage we asked the participants to fill out a daily survey providing a list of their interactions with others. Our algorithms detected 82% of the pairs that interacted based on the survey data. We also obtained hand-labeled ground truth from a subset of the users. Four participants labeled two days of their data in five-minute chunks (12 hours each). For the hand-labeled dataset, our performance accuracy in detecting conversations was

63.5% overall and 87.5% for conversations greater or equal to one minute. The conversations missed by our method were often in high-noise, multiple-speaker situations.

Learning the Social Network

Once we detect the pair-wise conversations we can identify the communication that occurs within the community and map the links between individuals. The link structure is calculated from the total number interactions each person has with others (interaction with another person that account for less than 5% of the person's total interactions are ignored). To get an intuitive picture of the interaction pattern within the group of people who were equipped with sociometer, we visualize the network diagram by performing multi-dimensional scaling (MDS) on the geodesic distances between the people (Figure 2). The nodes are colored according to physical closeness of office location. From this we see that, people whose offices are in the same general space seem to be close in the communication space as well.

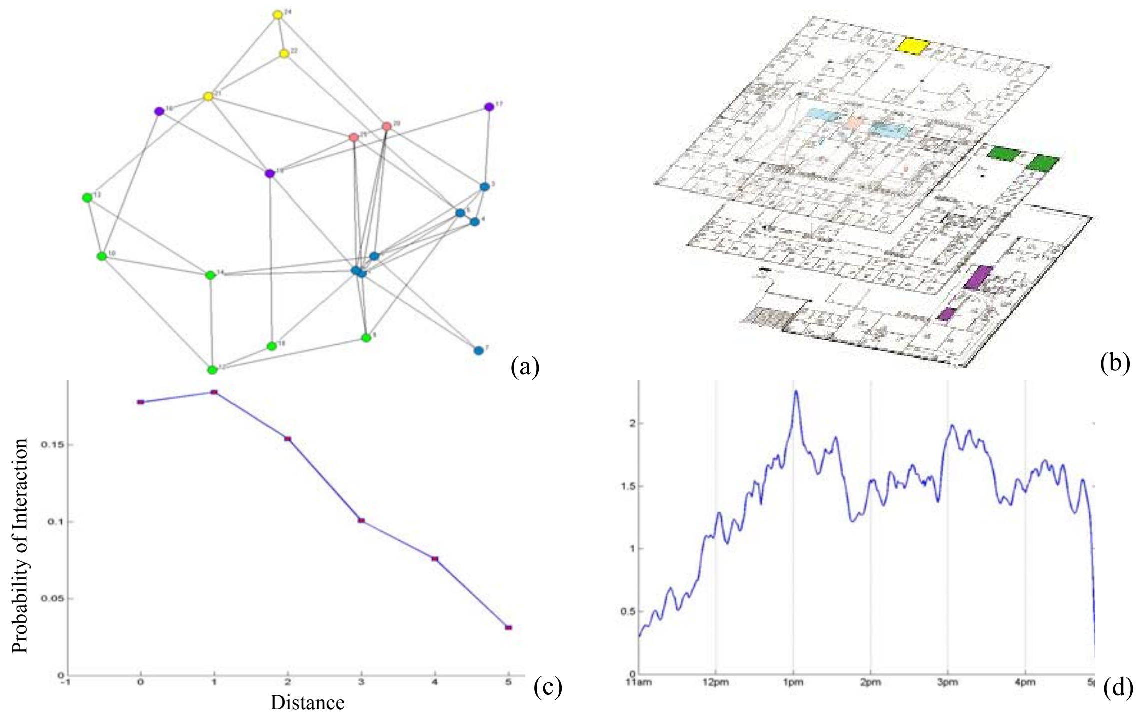

Figure 2: (a) Network structure based on MSD of geodesic distances (b) layout of subjects across the building. (c) probability of communication as a function of distance (d) average speech activity throughout the day

Structural layout is known to affect communication within an organization or community (Allen 1997). Figure 2(c) shows the probability of communication as a function of distance between offices. We grouped distances into six different categories – (i) office mates (x-axis 0) (ii) 1-2 offices away (x-axis 1) (iii) 3-5 offices away (x-axis 2) (iv) office on the same floor (x-axis 3) (v) offices separated by a floor (x-axis 4) (vi) office separated by two floors (x-axis 5). We have also calculated the average talking pattern throughout the day based on the fraction of time that speech was detected from a wearer's device (for every one-minute unit of time) as shown in Figure 2(d). This result is quite intuitive, as talking peaks during lunch time and also in the late afternoon when students often take breaks and when the weekly Media Lab student tea is held. These types of measurements of network behavior are much harder to do using surveys or self-report, but can easily be extracted from the analysis of the sensor data.

Centrality and Turn-taking Dynamics

Next we analyze the dynamics of the interactions, focusing on the turn-taking patterns of individuals and how they differ from each other. We start by defining a "turn" – for each unit of time we estimate how much time each of the participants speaks, the participants who has the highest fraction of speaking time is considered to hold the "turn" for that time unit. For a given interaction, we can easily estimate how a pair participating in the conversation transitions between turns. We use the speaker segmentation output within conversations to estimate the turn-taking transition probability. Because most of the conversations in the dataset are between pairs, we transition between two states: speaker A's turn and speaker B's turn. We selected eighty conversations which were on average 5 minutes long to compute the individual turn-taking dynamics. In selecting the conversations we made sure that we had at least four different conversation partners for each individual and multiple conversation instances for the same conversational pair, which applied to a subset of the participants only. When two people are interacting it is plausible that average turn-taking dynamics will affect each other and the resulting turn-taking behavior for that interaction will be a blend of the two transition matrices. If someone affects our average pattern a lot we may adapt to the behavior of that person's 'average conversation partner', if we are not affected at all we will probably maintain our average dynamics completely, or the resulting interaction behavior may be somewhere in between the two extremes. We model the transition probability of specific interaction as a combination of the individuals' turn taking styles, modeled by a two-dimensional "influence" parameter (for details see Choudhury 2003). By learning the influence parameters we can measure how much one person affects another's turn-taking behavior. We discovered an interesting and statistically significant correlation between a person's influence score and their betweenness centrality, the correlation value was 0.90 (p-value < 0.0004, rank correlation 0.92). '*Betweenness centrality*', measures how much control an individual has over the interaction of others who are not directly connected; people with high betweenness are often perceived as leaders (Freeman 1977). It appears that a person's interaction style is indicative of her role within the community based on centrality measure.

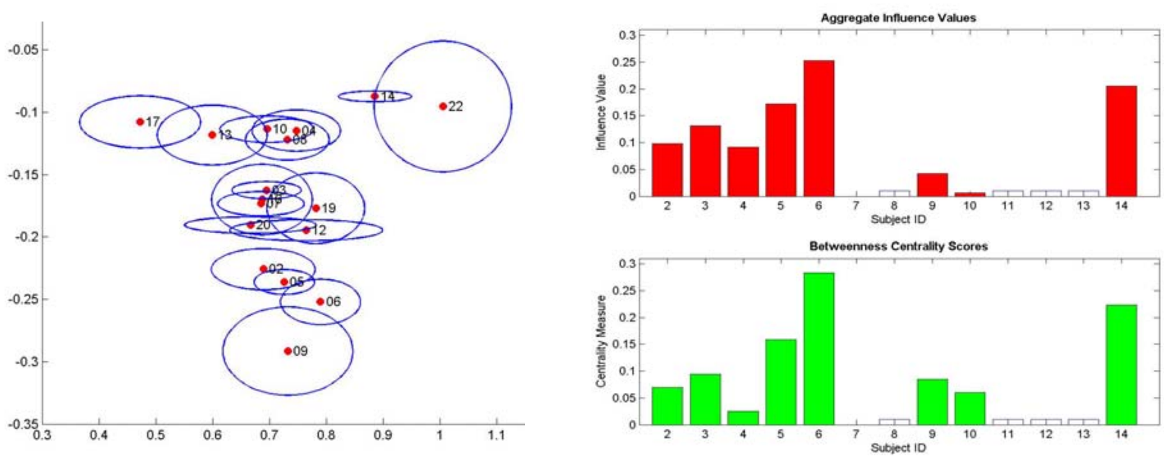

Figure 3: (a) MDS of average turn-taking styles. Each individual's mean is given by red circle and the ellipse around shows the variance over different conversations. (b) Influence value and betweenness score for a subset of individuals

Conclusion

In this paper we demonstrated the feasibility of learning social interactions from raw sensor data. We have presented a framework for automatic modeling of face-to-face interactions, starting from the data collection up to modeling the structure and dynamics of social networks by analyzing whom we talk to and how we talk to them. We believe better models of social network and organizational dynamics will facilitate efficient means of collaboration and information propagation. We have integrated methods from speech processing and machine learning to demonstrate that it is possible to extract information about people's patterns of communication without imposing any restriction on the user's interactions or environment. We have presented results that demonstrate distinctive turn-taking styles for individuals and new results that show strong correlation between a person's aggregate influence value and her centrality score. This indicates the possibility inferring a person's leadership role within the network from their conversational styles.

References

Allen, T. (1997). Architecture and Communication Among Product Development Engineers. *MIT Sloan School of Management WP # 165-97*, pages 1-35.

Basu, S. (2002). Conversation Scene Analysis. *MIT: EECS Ph.D Thesis,* pages 1-109.

Choudhury, T. (2003). Sensing and Modeling Human Networks. *MIT: Media Lab Ph.D. Thesis*, pages 1-106

Freeman, L. C. (1977). "A set of measures of centrality based on betweenness." *Sociometry* **40**: 35-41.

Gerasimov, V., T. Selker, et al. (2002). "Sensing and Effecting Environment with Extremity Computing Devices." *Motorola Offspring* **1**(1).

Gladwell, M. (2000). *The Tipping Point: How little things make can make a big difference*. New York, Little Brown.

Want, R., A. Hopper, et al. (1992). "The Active Badge Location System." *ACM Transactions on Information Systems* **10**: 91-102.