Characterizing Social Interactions using the Sociometer

Tanzeem Choudhury
Intel Research Seattle
1100 NE 45th Street, Seattle, WA 98105
tanzeem.choudhury@intel.com

Alex Pentland
MIT Media Laboratory
20 Ames Street, Cambridge, MA 02139
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.

Author Keywords
Social network analysis, organization behavior, machine learning

ACM Classification Keywords
H5.3 Group and Organization Interfaces.

INTRODUCTION
Our decision-making is influenced by the actions of others around us. How people are organized also has an effect on information diffusion [7]. Who are the people we talk to? How actively do we participate in the conversations? Can we identify the individuals who talk to a large fraction of the group or community members? Although people heavily rely on email, telephone and other virtual means of communication, high complexity information is primarily exchanged through face-to-face interaction[1]. Knowledge of people's communication networks can also be used in improving context-aware computing environments and in coordinating collaboration between group members. We believe that wearable sensor data combined with pattern recognition techniques can play an important role in sensing and modeling physical interactions.

SENSING & MODELING HUMAN NETWORKS
Prior work has focused on proximity-based models of face-to-face network [3, 9], 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 the hoarder board, a wearable data acquisition board, designed by the electronic publishing and the wearable computing groups at the Media lab [6]. The sociometer 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 had the device on them for six hours a day for 11 days, and wore the device both indoors and outdoors. The participants were a mix of students, faculty and administrative support staff. Subjects were distributed across different floors of the building and across different research groups.

The comfort, aesthetics, and placement of sensors, are important issues to consider in wearable design when it comes to greater user acceptance and reliable sensor measurements. In designing the sociometer, we followed the wearability criteria specified in [5], which explores the interaction between the human body and a wearable, and provides a guideline on the shape and placement of wearables that are not obtrusive and do not interfere with the natural movement of the human body.
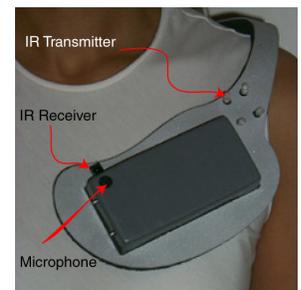

Figure 1: The Sociometer

| 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)

Table 1: Exit survey of the users. The number in parenthesis indicates how many users chose that response

We are aware that subject's privacy is a concern for any study of human interactions. 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: Table 1 summarizes the results and shows users were generally accepting of 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 proposed in [2] that discriminate well between speech and non-speech regions. An HMM 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. However, the downside of this 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 [2] 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. However, the survey data had only 54% agreement between subjects (where both subjects acknowledged having the conversation) and only
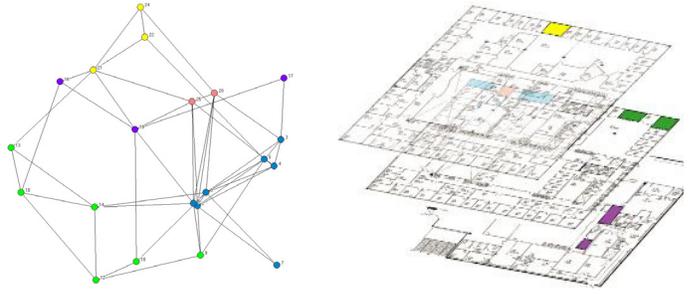
Figure 2: (a) Network structure based on MDS of geodesic distances. (b) layout of the subject across the building

29% agreement in the number of conversations. Consequently, we did not feel confident in comparing our algorithms against survey data alone. 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 measure all 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). This type of visualization is commonly used in social network analysis [10]. 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.

Effects of Distance on Face-to-Face Interaction
Structural layout is known to affect communication within an organization or community [1, 8].
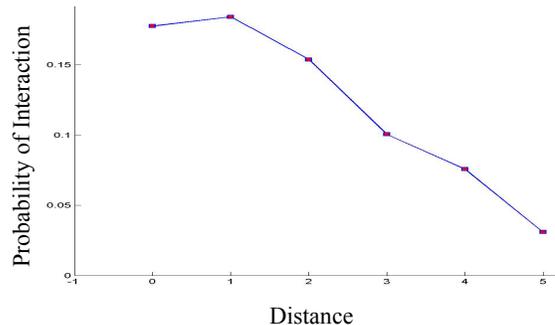
Figure 3: Probability of communication as a function of distance.
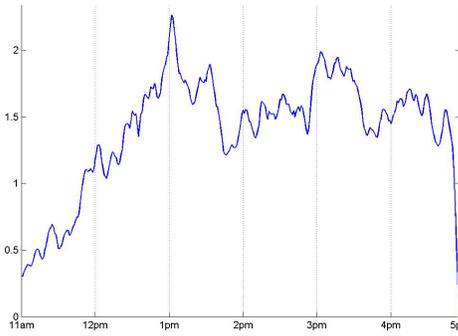
Figure 4: Speech activity over the course of the day averaged over all subjects across all days.

Figure 3 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).

Changes Speech Activity throughout the day

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 4. 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 Measure

Centrality measures are extensively used in social network analysis to understand an individual's involvement within the community. There are various centrality measures based on the number of links, as well as the importance of the links in maintaining the connectedness of the network. One particular measure is the '*betweenness centrality*', which measures how much control an individual has over the interaction of other individuals who are not directly connected. People with high betweenness are frequently viewed as leaders [4]. The betweenness centrality of our network is shown in Figure 5. Participants were assigned ID numbers 2-11 and 13-25. ID#15's device had technical failure and ID#23 did not participate regularly, so their data are ignored in further analysis. Some notable aspects of the betweenness measure are: (i) ID 8 was assigned to the author, who communicated with the subjects to coordinate data collection and hence the high betweenness measure may be biased (ii) ID 7 was an undergraduate student working with the author, note ID 7 mainly interacts with ID 8 and has a betweenness of 0 (iii) ID 16 and 17 are our least communicative subjects, and have low betweenness (iv) ID 4, although very tightly integrated within her group, she is
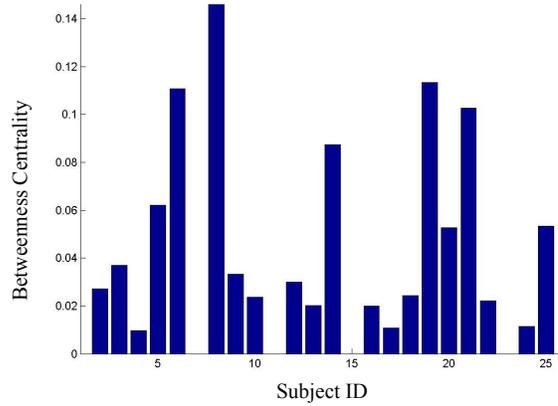

Figure 5: Betweenness centrality of the network

mostly isolated from other groups and consequently has a low betweenness score.

TURN-TAKING DYNAMICS

Next we analyze some of the dynamics of the interactions. We primarily focus on the turn-taking patterns of individuals and how they differ from each other. We use these individual dynamics to later estimate how much an individual's overall pattern changes during her interaction with specific individuals. 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.
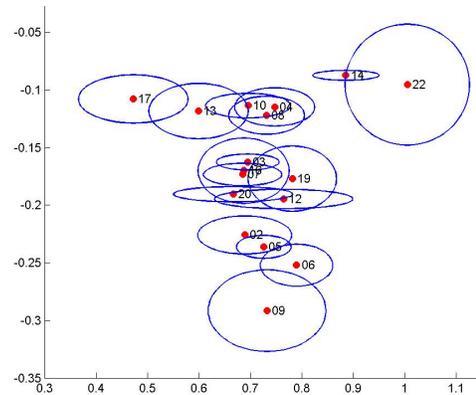

Figure 6: Multidimensional scaling of the average turn-taking transition tables. Each individual's mean is given by the red circle and the ellipse around shows the variance in speaker's style over different conversations.

3

Once we have estimated the turn-taking transition probabilities for the individuals we can measure how similar or dissimilar they are from each other. Figure 6 shows the output of multidimensional scaling of the transition probabilities using a Euclidean distance metric, which shows that individuals have distinctive turn-taking styles and that these turn-taking patterns are not just a noisy variation of the same average style. We show the turn-taking styles of seventeen individuals, because these seventeen are the ones who were participants in the eighty conversations selected. Later we will use ten of these subjects to do further analysis, those who have had at least four different conversation partners, in order to estimate their mean behavior and how they change or are influenced by other people's interaction behavior.

Estimating Influences from Turn-taking Dynamics

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. Now 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 centrality, the correlation value was 0.90 (p-value < 0.0004, rank correlation 0.92). It appears that a person's interaction style is indicative of her role within the community based on centrality measure.
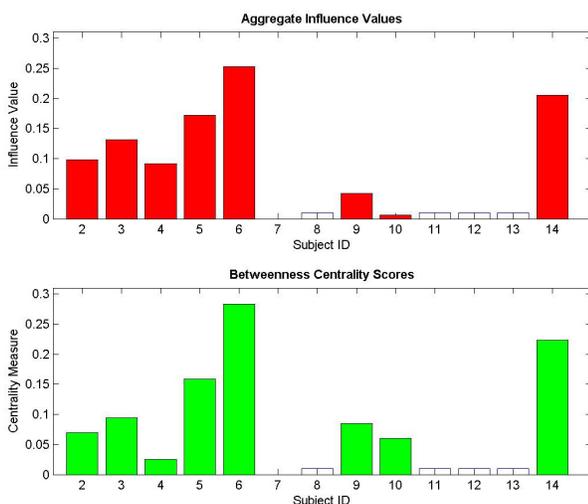
Figure 7: Influence value and centrality measure for a subset of the participants. Note, a participant with high influence also has high centrality.

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 and consistent turn-taking styles for individuals.   We have presented 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

1. Allen, T., *Architecture and Communication Among Product Development Engineers*. 1997, Sloan School of Magement, MIT: Cambridge. p. pp. 1-35.
2. Basu, S., *Conversation Scene Analysis*, in Dept. of Electrical Engineering and Computer science. Doctoral theis MIT. 2002, p. 1-109.
3. Borovoy, R., *Folk Computing: Designing technology to support face-to-face community building*. Doctoral theis MIT. 2001.
4. Freeman, L.C., *A set of measures of centrality based on betweenness*. Sociometry, 1977. **40**: p. 35-41.
5. Gemperle, F., et al., *Design for Wearability*. 1998, Institute for Complex Engineered Systems, Carnegie Mellon University, http://www.ices.cmu.edu/design/wearability/files/Wearability.pdf
6. Gerasimov, V., T. Selker, and W. Bender, *Sensing and Effecting Environment with Extremity Computing Devices*. Motorola Offspring, 2002. **1**(1).
7. Gladwell, M., *The Tipping Point: How little things make can make a big difference*. 2000, New York: Little Brown.
8. Menzel, H., *Review of studies in the flow of information among scientists*. Columbia University Bureau of Applied Social Research, 1960.
9. Want, R., et al., *The Active Badge Location System*. ACM Transactions on Information Systems, 1992. **10**: p. 91-102.
10. Wasserman, S. and K. Faust, *Social Network Analysis Methods and Applications*. 1994: Cambridge University Press.