Assessing Wellness by Sensing Everyday Activities and Interactions

Tanzeem Choudhury, Gaetano Borriello
Intel Research Seattle
1100 NE 45th Street, Suite 600
Seattle, WA 98105 USA
(tanzeem.choudhury,gaetano.borriello)@intel.com

Jonathan Lester
Dep't of Electrical Engineering
University of Washington
Box 352350, Seattle, WA 98195 USA
jlester@u.washington.edu

ABSTRACT
In this paper, we describe our vision of using wearable sensors to simultaneously track people's physical and social activities. We believe this will enable us to assess the overall well-being of individuals more easily. We present our initial work in capturing and modeling the social network, interactions, and physical activities automatically using multi-modal sensors.

INTRODUCTION
The overall well being of individuals depends on both their physical and emotional health/states. Just as the activities we perform impact our physical health, the social interactions we have affect our emotional health. Automatic sensing of the state of people's physical health is a very active research area. However, progress in sensing social well-being has been limited, especially when it comes to sensing face-to-face interactions. The conversations and interactions that happen in person are a critical component of one's social life. Furthermore, the structure and dynamics of social networks are of great importance in understanding the spread of disease [1, 2]. We are developing a system that can simultaneously track both physical activities and social interactions, using sensing and communications tools pulled from the wearable and ubiquitous computing communities, coupled with machine learning techniques, to unobtrusively study the physical and social behavior of large populations over extended periods of time.

SENSING AND MODELING SOCIAL INTERACTIONS
In our earlier work on automatic modeling of face-to-face interactions, we collected an auditory database of social exchanges among twenty-three people using wearable sensors (66 hours of data per person over two weeks) [3]. The users wore the sensors both indoor and outdoor and their movement and interactions were completely unconstrained. We applied speech and conversation detection methods to the auditory streams. These methods automatically locate the conversations, determine their participants, and mark which participant was speaking when. The overall accuracy of our algorithms in detecting conversation was 87%. Figure 1 shows a social network map generated from the data collected from the participants in this project.

In our current research, we are extending this work to further model people's roles and relationships within the network and to explicitly capture the interplay of the physical environment and the social networks. Specifically, we are trying to answer questions such as: How do people's interaction patterns relate to their positions within social networks? How do locations serve as hubs or bridges between social groups, channeling the evolution of interpersonal networks?

SENSING AND MODELING PHYSICAL ACTIVITIES
The task of inferring activities presents a difficult problem, due to the need for carefully crafted sensing inputs and accurate models. Recently the topic of human activity inference has received a great deal of attention in the ubiquitous computing (ubicomp) field. A majority of the work in the ubicomp community has focused on inferring activity from multiple sensors of a single modality, typically accelerometers, placed on several locations of the body [4, 5]. While reasonable, the use of multiple sensors placed at specific locations on a person can be quite
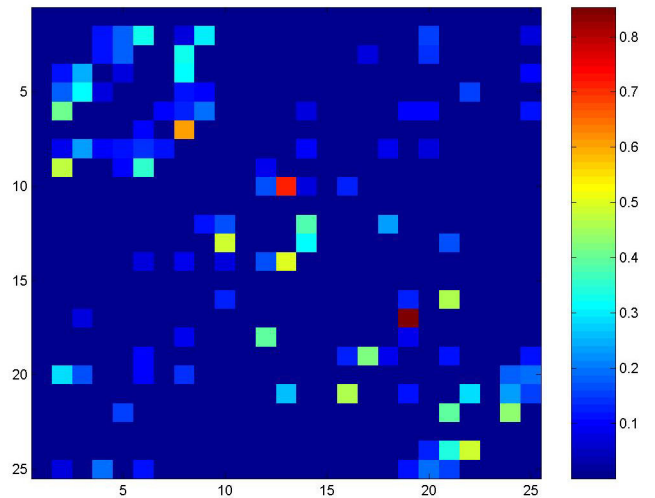

Figure 1. Network map based on proximity information. x-axis and y-axis values represent subject ID number, and the color value indicates, for each subject, the proportion of their total interactions with each of the other subjects.
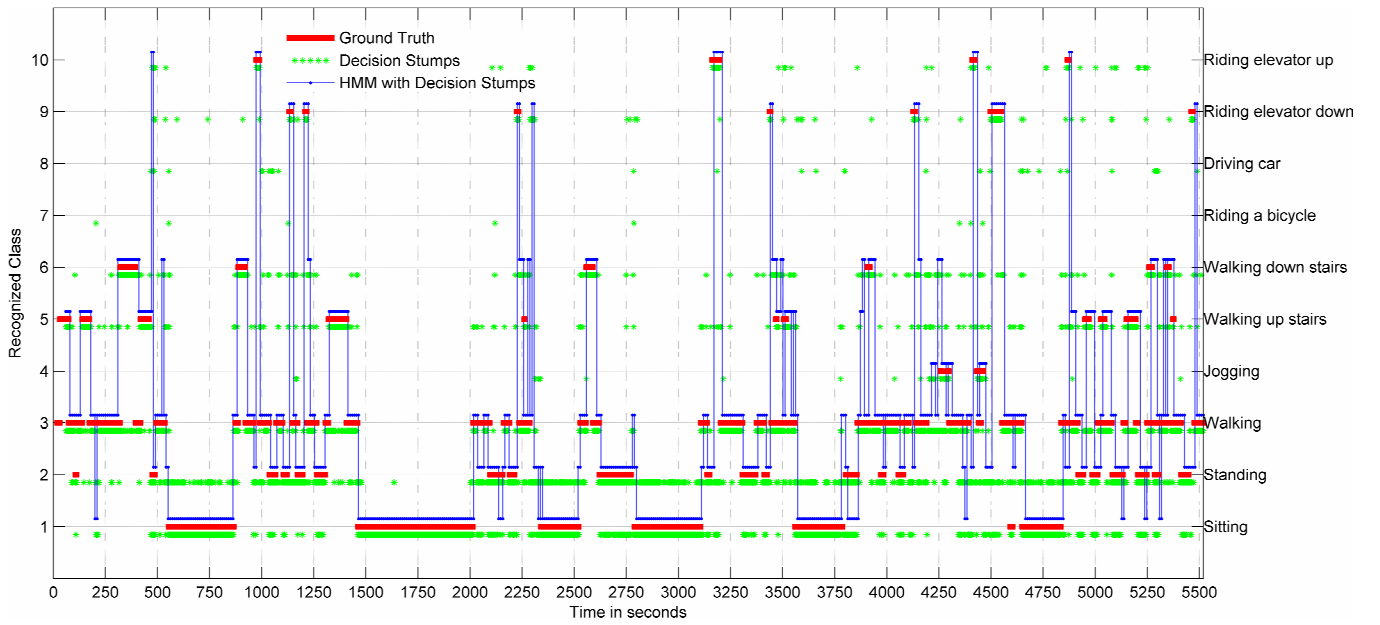
Figure 2. Output of the decision stumps classifier (at 4Hz), HMM with probabilities as inputs (using a 15 second sliding window with 5 second overlap), and the ground truth for a continuous hour and half segment of data.

obtrusive and presents a large barrier for user acceptance. Also, current sensing technology has not yet reached the point where inexpensive sensors can be readily placed in clothing. As a result, the use of a single sensor placement is more preferable; especially if such a sensor placement can be integrated into a device that people already carry, such as cell phones or wrist-watches.

Work by Bao and Intille [**4**] has shown that an appropriate sensor subset (two locations), does not affect the recognition scores significantly (by less than 5%) compared to a system with five sensors; whereas the use of a single sensor reduced the average accuracy by 35%. Our current hypothesis is that the information lost by using a single sensor placement, versus multiple sensor placements, can be offset by the use of multiple sensing modalities. Furthermore, multiple modalities will be better suited to capture the rich perceptual cues that are present in the environment, which a single modality may be unable to capture. The use of multiple sensor modalities has already shown promise in earlier activity recognition experiments [**6**].

To capture the diverse cues from movement, sound, light, etc., about an ongoing activities, we have built a very small sensing unit (2.53 sq. in.), shown in Figure 3, that includes eight different sensors: accelerometer, audio, IR/visible light, high-frequency light, barometric pressure, humidity, temperature, and compass. Using these sensors, we have collected a large annotated dataset of various human activities (over 35 hours) from two volunteers over a period of a few weeks.

From these sensor readings we have used a combination of discriminative decision stumps and generative HMM
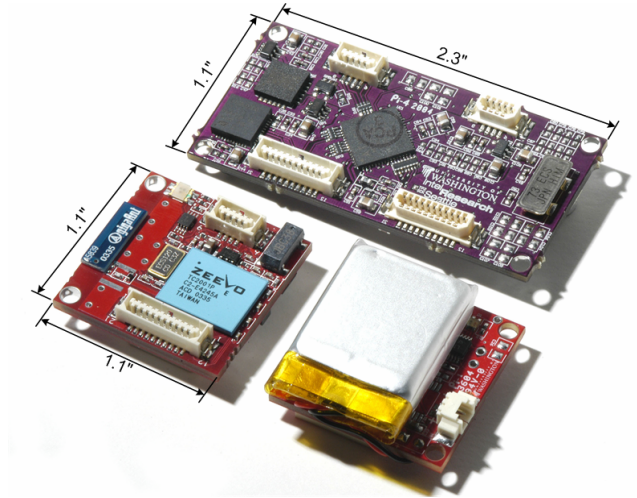

Figure 1. The multi-modal sensor board (top) along with a Bluetooth iMote (lower left), and a USB rechargeable Lithium-Polymer battery board (lower right).

classifiers to classify ten basic activities. The current classifications scheme uses the outputs from the decision stumps classifiers as inputs to the HMM; allowing us to leverage the good discriminative qualities of the decision stumps classifiers with the temporal smoothness of the HMM. On their own, the decision stumps classifiers have an overall accuracy of 86%, when combined with the HMM our overall accuracy increases to 92% and, as shown in Figure 2, our classifications are much smoother.

2

PRIVACY CONCERNS AND SENSING PLATFORMS

Preserving the privacy of users is of critical importance when it comes to using sensors to track individual well being, especially when it involves the recording of audio. One approach to protecting privacy is to have the low level hardware perform most of the feature processing and extraction, so instead of recording/transmitting raw signals only more abstract measures like mean, variance, frequency bands, etc. are transmitted. This prevents the data from being intercepted and listened to; however, the information we can extract from such features is not necessarily benign. While these features abstract the data away from something we can intuitively understand, these features and the result of the inference we perform can still be a large privacy concern. For example, one could imagine a situation where the knowledge that two individuals are talking to one another is just as damaging as an audio recording of such a conversation. A well designed wearable platform maybe be more suited to tackle this problem than sensors embedded in the environment, as the user will have ultimate control in deciding whether to record or transmit the data. As presented in Table 1, our early surveys show that a well-designed wearable can be used unobtrusively to collect data and that people are more receptive to transmitting conversational information if the content can't be decoded (i.e., transcribed). Although we realize that more thorough and complete studies of user privacy need to be conducted,

The wearable interfered with normal interactions	Never (33%)	Few times (67%)	Most of the time (0%)	Always (0%)
Will wear an audio recording device only if transcription is not done	Don't care (29%)	Yes (57%)	Do not like wearing it, even if transcription is not done (14%)	Would never wear any data collection device (0%)

Table 1: Exit survey of twenty-one users from our earlier work on sensing social interactions. The number in parenthesis indicates what percentage of the users chose that response.

we believe that there are viable privacy solutions that can be incorporated into a wearable sensor-based wellness assessment system.

CONCLUSION

We have presented our initial work on building a sensor-based system for the assessment of physical and mental wellness. Currently, we are working on developing richer models of social interactions that can capture people's roles and relationships based on locations and more complex, compound activities, (e.g., making dinner, cleaning the house, etc.). We also aim to explore the usage models for this technology in managing personal wellness.

BIOGRAPHIES

Tanzeem Choudhury is a researcher at Intel Research Seattle and an Affiliate Professor of Computer Science and Engineering at the University of Washington. Her primary research focus is on machine learning and machine sensing of human and group behavior. She is currently working on sensor-based models of human activity and social networks.

Jonathan Lester is a graduate student in Electrical Engineering at the University of Washington. His main research focus is on sensing and activity inference in ubiquitous computing and the mapping of those computational requirements to various embedded platforms. He is currently working on systems for sensor-based models of human activity.

Gaetano Borriello is a part-time researcher at Intel Research Seattle and a Professor of Computer Science and Engineering at the University of Washington. His primary interests are in ubiquitous computing, multi-device interactions, and sensor/actuator systems.

REFERENCES

1. A. S. Klovghal, *Social networks and the spread of infectious diseases: The aids example*, Social Science & Medicine 21 (1985), no. 11, 1203-1216.
2. S. Eubank, H. Guclu *et al.*, *Modeling disease outbreaks in realistic urban social networks*, Nature 429 (May 2004), pp. 180-184.
3. T. Choudhury, "Sensing and modeling human networks," *Doctoral Thesis*, MIT, Cambridge, MA, 2003, pp. p. 1-105.
4. L. Bao and S. Intille, *Activity recognition from user-annotated acceleration data*, Proc. Pervasive 2004, pp. 1-17.
5. N. Kern, B. Schiele *et al.*, *Multi--sensor activity context detection for wearable computing*, Proc. EUSAI 2003, pp. 220-232.
6. P. Lukowicz, H. Junker *et al.*, "Wearnet: A distributed multi-sensor system for context aware wearables," *Proc. Ubicomp* 2002, pp. 361-370.