Understanding Ambulatory and Wearable Data for Health and Wellness

Akane Sano  Rosalind W. Picard
Massachusetts Institute of Technology Media Lab Affective Computing Group
{akanes, picard}@media.mit.edu

Abstract
In our research, we aim (1) to recognize human internal states and behaviors (stress level, mood and sleep behaviors etc), (2) to reveal which features in which data can work as predictors and (3) to use them for intervention. We collect multi-modal (physiological, behavioral, environmental, and social) ambulatory data using wearable sensors and mobile phones, combining with standardized questionnaires and data measured in the laboratory. In this paper, we introduce our approach and some of our projects.

Introduction and Our Approach
Recently, we have so many devices to monitor our daily lives: pedometer, activity monitor, and sleep monitor etc. Many people wear them to quantify their personal behaviors; however, how can we use our collected data other than showing them in graphs? Our motivation to collect data is not only to visualize them but also to understand the meaning, recognize something internal behind the data (health condition or emotional states) and feedback them to users to help them to change their behaviors.

Fig.1 Our approach to understand ambulatory data.

In our research, we aim (1) to recognize human internal states and behaviors (stress level, mood and sleep behaviors etc), (2) to reveal which features in which data can work as predictors and (3) to use them for intervention. We collect multi-modal (physiological, behavioral, environmental, and social) ambulatory data using wearable sensors and mobile phones, combining with standardized questionnaires and data measured in the laboratory. In this paper, we introduce our approach and some of our projects.

Projects
Understanding long-term sleep behavior at home
Sleep is one of the major problems in modern society. Polysomnography is a gold standard to measure sleep; however, it requires people to visit a sleep laboratory and spend one night with many wires and sensors on the body which interfere with their sleep. Recently, many wearable devices have been commercialized as consumer versions of actigraphs (Fitbit, Jawbones, etc) to monitor daily activity and lack of movements that may correspond to sleep; however most of them are based only on accelerometer data and the accuracy to detect sleep stages is unclear. According to validated algorithms (Cole et al. 1992), accelerometer data has been used only to detect sleep and wake. Recent applications also utilize the amount of activity to infer greater likelihood of REM for alarm clock awakening. In our project, we combine electrodermal activity (EDA), skin temperature, and accelerometer and investigate how multi-modal data help us to understand sleep better. EDA refers to electrical changes in the surface of the skin activated by the sympathetic nervous system. It can be measured through skin conductance using electrodes on the skin surface (Boucsein, 1992). Traditionally, EDA has been measured with gelled electrodes on palms and fingers; however we have been using dry electrode based EDA measurement systems for more natural and comfortable long-term measurement. We compared sleep EDA and Polysomnography and confirmed that EDA shows peaks in NREM2 and SWS (Sano and Picard, 2011) (Fig.2). We have been working on sleep/wake detection using accelerometer, skin temperature and...
and electrodermal activity data. Combination of these three could improve the accuracy of sleep and wake detection.

Fig. 2 Comparison between EDA and sleep stages

Memory consolidation using multi-modal data
We also have worked on how sleep sensor data can help estimate sleep dependent memory consolidation. Past studies have shown that sleep can enhance memory consolidation: consistent and significant performance improvement on a Visual Discrimination Task (VDT) became proportional to the amount of sleep in excess of six hours, and subjects with an average of eight hours then exhibited a correlation in performance to a particular pattern of sleep stages derived from the EEG: percent of SWS (Slow Wave Sleep) in the first quarter of the night, times percent of REM (Rapid Eye Movement) in the last quarter (Stickgold et al. 2000). We wanted to see if the non-EEG data from the sensor gives prediction of memory consolidation. We collected EEG (electroencephalogram), EDA and ACC data during sleep and VDT performance before and after sleep for N=24 healthy adults for 2 nights each for a total of 48 nights of data. We extracted features and applied machine learning techniques from the sleep data to classify whether the participants showed improvement in the memory task. Our results showed 60-70% accuracy in a binary classification of task performance using EDA or EDA+ACC features, which showed significantly higher accuracy than the more traditional use of EEG-based sleep stages to predict VDT improvement (p< 0.05, Sano and Picard, 2013a).

Stress recognition using mobile phone and wearable sensors
In this project, we aim to find objective markers in mobile phone and wearable sensor data that correspond to stress as evaluated by the Perceived Stress Scale (PSS) (Cohen, Kamarck, and Mermelstein. 1983). We collected 5 days of data for each of N=18 healthy adults: a wrist sensor (ACC and EDA), mobile phone usage (call (CALL), short message service (SMS), COMM: CALL+SMS, location (MOB) and screen on/off (SCREEN)) and surveys (personality, stress, mood, sleep, tiredness, general health, alcohol or caffeinated beverage intake and electronics usage). We applied sequential forward floating selection (SFFS) to understand which feature from each modality is related to stress level. In addition, we used machine learning to classify whether the participants had high or low PSS stress scores. We found stress related features in mobile phone data as well as in the survey and personality data (Table 1) (Sano and Picard, 2013b).

Fig 3 Mean accuracy of classification (using 6 different classifiers, N=48 nights)

Table 1 Stress recognition results

<table>
<thead>
<tr>
<th>Classification accuracy</th>
<th>Modality</th>
<th>Best feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>87.5%</td>
<td>Post Survey</td>
<td>Often felt so sad or down that you had trouble functioning in school or personal life</td>
</tr>
<tr>
<td></td>
<td>Mobile phone usage (CALL, SMS, MOB, COMM, SCREEN)</td>
<td>Mean duration of calls between 9pm-12am, SD of % of SMSs between 9pm-12am, Mean of SD of mobility radius, SD of total missed transactions, SD of % of screen ons between 6-9pm</td>
</tr>
<tr>
<td>81.3%</td>
<td>Big Five Test</td>
<td>Neuroticism</td>
</tr>
<tr>
<td></td>
<td>Evening phone survey</td>
<td>SD of answer time</td>
</tr>
<tr>
<td>75%</td>
<td>MOB</td>
<td>Mean of SD of mobility radius</td>
</tr>
<tr>
<td></td>
<td>SCREEN</td>
<td>SD of % of screen ons between 6-9pm</td>
</tr>
</tbody>
</table>

References