SADHealth: A Personal Mobile Sensing System for Seasonal Health Monitoring

Liam McNamara and Edith Ngai

***Abstract*—People's health, mood, and activities are closely related to their environment and the seasons. Countries at extreme latitudes (e.g., Sweden, U.K., and Norway) experience huge variations in their light levels, impacting the population's mental state, well-being and energy levels. Advanced sensing technologies on smartphones enable nonintrusive and longitudinal monitoring of user states. The collected data make it possible for healthcare professionals and individuals to diagnose and rectify problems caused by seasonality. In this paper, we present a personal mobile sensing system that exploits technologies on smartphones to efficiently and accurately detect the light exposure, mood, and activity levels of individuals. We conducted a 2-year experiment with many users to test the functionality and performance of our system. The results show that we can obtain accurate light exposure estimation by opportunistically measuring light data on smartphones, tracking both personal light exposure and the general seasonal trends. An optional questionnaire also provides insight into the correlation between a user's mood and energy level. Our system is able to inform users how little light they are experiencing in the winter time. It can also correlate light exposure data with reduced mood and energy, and provide quantitative measurements for lifestyle changes.**

***Index Terms*—Activity, health, light, mobile sensing, seasonality.**

I. Introduction

A PERSON'S environment, sunlight exposure and activity level significantly affects their health and well-being. Natural light synchronizes diurnal rhythms in physiology, sleep, muscle, and cardiovascular function. It can also elevate one's alertness, cognitive performance, and mood. Thus, it is common for people in the Nordic countries, who experience fewer hours of sunlight in the long and dark winters, to make use of light-boxes to supplement their sunlight. Seasonal effects, such as the "winter blues," cause changes not only in mood but in energy levels, sleep patterns, eating, and social behavior. In extreme cases, patients may suffer from seasonal affective disorder (SAD), which is a form of depression that can occur in the autumn and winter months. The symptoms include fatigue, lack of interest in normal activities, social withdrawal and weight gain. In the United States, 4%–6% of people are estimated to suffer from SAD. Another 10%–15% experience a milder form of winter-onset SAD [1].

Unfortunately, data collection in this area generally employs intermittent questionnaires to diagnose seasonal effects, which provides patchy, nonquantitative and subjective data. To investigate people's seasonal response, we need to automatically continuously collect quantitative and objective data. As mobile technology advances, smartphones nowadays possess many on-board sensors. We propose using smartphones as convenient tools for personal well-being/environment data collection. To the best of our knowledge, we are the first to explore the capabilities of a smartphone to monitor human light exposure and provide long-term monitoring of seasonal effects on human health, behavior, and well-being.

Measuring, understanding, and reflecting upon a patient's light exposure and activity levels is crucial when attempting to diagnose and remedy seasonal disorders. With access to this data, it becomes possible to monitor and analyze their health and well-being, so as to give early warnings for those who may need further diagnosis or medical treatment. Medical professionals and psychologists can use this quantitative and historical data for diagnosis. Users of light-boxes and light-rooms can use such a system to precisely monitor their light exposure and record any changes in their mood and behaviors [2].

A seasonality monitoring system should be able to provide user-friendly and stable long-term monitoring. It has to be unobtrusive and yet still provide accurate sensing measurements that lead to meaningful interpretation. To encourage participation in such a monitoring system, users need a high level of comfort, simplicity, and convenience of use. External wearable devices (such as light sensors or motion detection belts) as used in sports applications are not ideal, as they are designed to only function for short periods of time and may not be pleasant for frequent daily use. Energy-efficiency is another concern when constantly running the application on the phone. The mobile application should not significantly drain the battery or disrupt the normal operation of the phone.

This paper presents *SADHealth*, a personal sensing system that provides a smartphone application to collect data on activity levels and light exposure without any external sensing device. The system is assisted by a backend server to support data storage and analysis based on the collected data. The mobile sensing system exploits the capabilities of smartphones to provide sufficiently accurate data for monitoring seasonality effects. We observe that such monitoring only requires coarse-grained information to characterise the general trend of a user's environment and activities. Based on this, we optimize the sampling of on-board sensors to reduce energy consumption of the phone for long-term monitoring. We performed a 2-year study on the system's efficacy, reliability, and usability to ensure that long-term usage is viable.

Manuscript received January 31, 2015; revised December 02, 2015; accepted January 03, 2016. Date of publication February 18, 2016; date of current version March 23, 2018.

L. McNamara is with Swedish Institute of Computer Science, Stockholm SE-164 51, Sweden (e-mail: ljjm@sics.se).

E. Ngai is with the Department of Information Technology, Uppsala University, Uppsala SE-751 05, Sweden (e-mail: edith.ngai@it.uu.se).

Digital Object Identifier 10.1109/JSYST.2016.2521805

The main contributions of this work are: 1) design of an unobtrusive sensing system that does not require extra devices for light, activity and mood monitoring; 2) energy-efficient sensing strategies working with the cloud server to provide accurate monitoring and analysis of user states; and 3) implementation of the proposed system and a 2-year study to demonstrate its energy-efficiency, accuracy, and efficacy in supporting long-term studies on seasonal health.

This paper is organized as follows. Section II discusses related work. Section III describes an overview of our approach and Section IV presents the details of our system. The evaluation is covered in Section V followed by our conclusion and future work in Section VI.

II. Related Work

Wireless body area networks (WBANs) of sensors around the human body have been studied for medical, e-health and sport applications [3]. BikeNet [4] is a mobile sensing system that uses a number of sensors embedded into a cyclist's bicycle to gather quantitative data about the cyclist's rides and health. The system provides environmental data that can help users to find better or nonpolluted routes. It demonstrated that mobile wireless sensor networks can improve quality of life, including how we are impacted by our environment and how we can use data to regulate our activity patterns.

SociableSense [5] is a mobile sensing application that monitors the sociability of people in the workplace. This system utilizes accelerometer, microphone sensors, and Bluetooth to capture human behavior in office environments and measures their sociability with colleagues. The collected data can provide a deep quantified understanding of social dynamics in office environment, which may help companies to better manage employees and increase their productivity.

Apart from physical health, researcher's attention has been drawn to understanding human emotion and mental health. Affective healthcare [6] has been proposed as a mobile service that allows people to understand and recognize their emotions and their level of stress. iCalm [7] is a compact wearable sensing system for long-term monitoring of autonomic nervous system and motion data, including electrodermal activity, temperature, motor activity, and photoplethysmography. It is designed as a reliable, low power, and low-cost wearable system. Similarly, MoodMiner [8] and EmotionSense [9] both use mobile phones to assess an individual's mood. Various mobile sensors such as accelerometer, light sensor, sound sensor (microphone), and location sensor (e.g., GPS) work together with the call logs on smartphones to monitor human behavior and assess daily mood. Improving user's sleep quality was addressed in iSleep [10], where off-the-shelf smartphones use their microphones to monitor users while they sleep. Similar to our approach, they use decision trees to classify users' behavior.

Activity recognition is used to deduce what kind of activities are performed by users from their accelerometer data. In [11], user activity was measured by using a single wearable accelerometer and eight activities, including walking, running, climbing, and sitting were classified using collected data. Ward *et al.* [12] studied how to perform activity recognition on assembly tasks using body-worn accelerometers and microphones. It described a method for continuous recognition of activities (sawing, hammering, filing, drilling, grinding, sanding, etc.) using microphones and three-axis accelerometers mounted at two positions on the user's arms. Activity and emotion recognition in [13] focuses more on mental disorders to support early diagnosis of psychiatric diseases. Real-time motion classification has also been investigated for wearable computing applications [14], which can retrieve information about user's status in real time. Context awareness has also been explored to analyze features for activity recognition [15]. A survey of various activity recognition papers is presented by Lara and Labrador [16]. They show that a range of different classification techniques can provide high (over 90%) classification accuracy. Such findings indicated to us that we should select an approach with low computational overhead that still maintains such performance.

Although mobile sensing has been widely investigated for healthcare and sports applications, none of the existing work has explored monitoring the seasonality effect on human behavior, mood, and well-being. To the best of our knowledge, this is the first work to explore the sensing capability of a smartphone to measure the light exposure and to study behavior and mood changes of individuals due to the effect of season/environment. Different from existing healthcare and sports applications, the monitoring of seasonality requires long-term but possibly coarse-grained sensing data. These unique properties motivate our investigation into optimizing the communication and sampling intervals of the on-board sensors to reduce energy consumption, while ensuring adequate data for observing user activity, sociability, mood, and environmental changes in seasonality.

III. Requirements

We shall now describe the requirements and design goals of our system. To participate in our system, users only need their smartphones and to have occasional Internet access for uploading their data to the backend server. Their smartphones can then report and access the data from the server through WiFi or mobile data networks. Our system incorporates a public Internet server, which is responsible for storing and processing the sensed data in a secure manner. The user transmits their data (*ID*, *environmental sensor data*, *time*, and *date*) to the server to be stored privately, as seen in Fig. 1. In this way, we reliably keep track of users data, including accelerometer, light, location, and mood readings without filling up phone storage. Collected data are split into 2-day long periods and saved locally on the SD card of the smartphone, then opportunistically transferred to the server through a device initiated authenticated HTTP push request.

A. Sensing Types

According to studies on seasonality [17], the light/dark cycle is the dominant force in synchronizing people's physiological diurnal rhythms with their external environment. There is considerable anecdotal and empirical evidence that diurnal rhythms

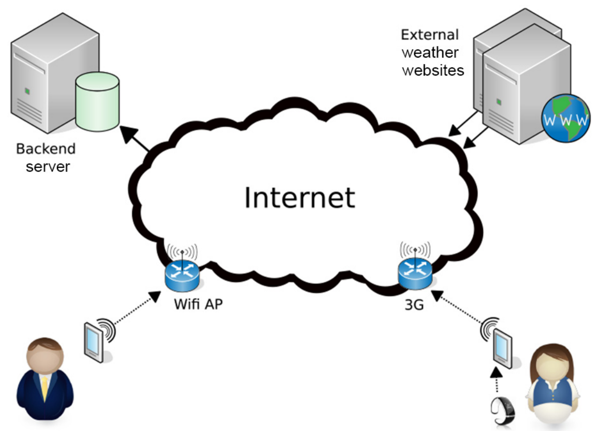

Fig. 1. Architecture diagram, showing device connectivity between all active components of the system. Two users with differing network connection types and one with an external fitness bracelet.

are a significant factor in determining mood, cognitive performance, health, energy, and well-being. Most people have experienced changes in their mood or behavior with the change of the seasons. When designing our system, we were particularly interested in measuring the amount of light users received at different times of the year and exploring how this seasonal change may interact with their sensed activity level and social behavior.

The recent advancements of embedded sensors in smartphones makes the development of our mobile sensing application on seasonality and health possible. Most current smartphones are equipped with location awareness (e.g., GPS), motion sensors (accelerometers), and light sensors, which allow us to the measure the location, activities, and social behavior of people. We also see the trend of embedding new sensor types into smartphones, such as air pressure, temperature, and humidity sensors in the Samsung Galaxy device series. This technology revolution changes the ways that we can make use of our phones and opens a new frontier of novel crowdsensing applications for the environment and personal health monitoring.

1) Light: Electronic light sensors can be used to measure light intensity in *lux*, (the SI unit of illuminance). It can measure the brightness of different light sources. Table I presents approximate lux values of different light sources for comparison. We use the phone's built in light sensor to measure this value. The light sensors on a smartphone were originally equipped for tuning the brightness of the screen. The physical placement of a phone depends on the user's behavior and habits, which presents new challenges in collecting accurate and reliable sensing data. For example, it can be hard to take measurements from the light sensor if the phone always has a cover in place or is constantly placed in a bag. This problem also occurs in activity sensors, such as the accelerometer.

2) Activity: Modern smartphones have accurate high-frequency triaxial accelerometers to allow more advanced interaction modalities. If the phone is carried by the user, these sensors can passively capture their movements. Acceleration recordings can then be *classified*, through machine learning

TABLE I
LUX VALUES FROM DIFFERENT LIGHT SOURCES

Lux count	Environment example
1	Full moon at a tropical latitude
80	Office building lights in hallway
320–500	Office lighting
1000	Overcast day;typical TV studio lighting
10 000–25 000	Full daylight (not direct sun)
32 000–130 000	Direct sunlight

techniques, into different activities, gaining an understanding of what person is doing and thus how lively they are being. Rather than permanently recording, intermittent sampling can be used to roughly capture a user's activity patterns. Classification works best when previously trained on the relevant user, we investigate if it can still be reasonably successful when not specifically trained on the operating user. The phone will not be able to capture the user activities accurately if the user always places the phone stationary on the desk rather than keeping it on their person. Turning numerical data into categorical classifications with human-understandable meaning are particularly useful for the users. Classification can also avoid problem of acceleration caused by cars, trains, etc., leading to large accelerations but not reflecting any physical exertion of the user. This technique can help giving synopsis of longer periods of time. For example, users can simply be shown with the number of hours they have spent performing vigorous activities, rather than opaque numbers. Activity classification has previously been performed using many different machine learning techniques [18], including decision trees, multilevel perceptrons, and support vector machines. We use decision trees generated with the C4.5 algorithm,[1] an efficacious but comparatively computationally nonintensive technique [16]. This ensures lightweight operation of the classification procedure and enables the use on users' resource constrained devices.

3) Sociability: Smartphones are able to recognize people's location and motion by the built-in location sensor and accelerometer. Location sensors can measure the position of users in order to understand how much time they spend travelling and visiting different places. We record the location of users every hour using either the phone's passive provider (last known location) or network provider (based on nearest mobile network station and WiFi stations) depending on their availability, to save the battery of the smartphones. GPS will not be manually enabled by our app due to its high energy costs and also considering GPS's inability to function indoors (though if otherwise turned on, it will be used).

Another factor that could be taken into consideration as a factor of sociability is *phone usage*. Smartphones facilitate socialisation for their users in different ways like making phone calls, text messaging, social network interactions, etc. The number of times phone has been used is recorded per day as auxiliary data for sociability measurement. The integrated data from the location sensor and phone usage count help us to monitor phone interaction. A more in-depth system could even monitor the user's Internet which has been shown to give insight into a user's depression [19].

[1] Specifically, the J48 implementation in the Weka machine learning suite.

4) Mood: Since emotion/mood sensors are not available, a questionnaire related to mood and energy levels has been included in our mobile application for users to optionally fill in. The optional questionnaire alert pops up two times a day and only during the day time. The questions are:

1) How is your *mood*?
2) How is your *sleep*?
3) How is your *energy level*?
4) How *social* have you been feeling?

The SADHealth questionnaire has been inspired by The Seasonal Pattern Assessment Questionnaire (SPAQ) [20], which has been summarized into four fundamental questions. Users enter their responses using a simple visual slide-bar, which is converted to nominal values. This questionnaire facilitates the regular monitoring of the user's self-reported mental health state, allowing future reflection of how a person was feeling and how their moods changed over time. When combined with the light and activity data, this could help identifying patterns in a user's life.

B. Design Goals and Major Challenges

Since the SADHealth system aims for *long-term monitoring*, it focuses on user comfort and practicality of daily use compared with many other short-term sensing applications for sports or medical diagnosis. We intend to exploit the full potential of smartphones to design an *unobtrusive* system that does not rely on any external wearable sensors.

The medically sensitive nature of this type of system requires some important comments to be made. The collection of personal data (e.g., location and activities) implies a requirement for the system to protect privacy and ensure that data is not leaked to unauthorized sources. The collected datasets have not been released to the public and participants we recruited were asked to give consent to their data being used in subsequent research. A more subtle consideration comes from users that may be in medically precarious situations that should not rely on our offering providing definite protection against any medical problems. This holds for any such system, which should be careful about promising protections or outcomes and should focus on the possibility of information collection enabling a user to change their own action or for it to be provided to a trained healthcare professional as means for much more informed diagnosis.

Furthermore, *energy-efficiency* is a major concern when it comes to sensing on smartphones. Continuous high-frequency sampling will consume battery, computation, and storage of the resource-constrained devices. Our architecture should ensure that the mobile application will not overload the phones or disrupt their normal operation. The energy consumption should be minimized to avoid battery drain and frequent charging of the phones, which could discourage user participation. A naive strategy is to reduce how often the sensors are sampled. However, this may lead to incomplete data that results in inaccurate observations.

While reducing the energy-consumption of the smartphones, the SADHealth system should provide *sufficiently accurate* data for reporting findings and enabling meaningful conclusions. It is essential to deal with missing or inaccurate data due to the limited capability of the on-board sensors or improper placement of the phones. Energy-efficiency and data accuracy have to be carefully balanced to provide adequate data for analysis. The data collection scheme should capture the key features of user behavior, health, and well-being due to seasonality change. The sensing data can then be stored and analyzed on the backend server.

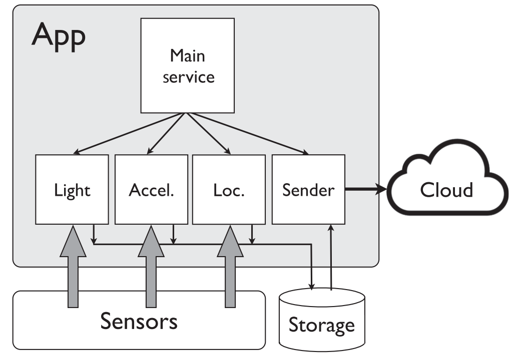

Fig. 2. Software system diagram depicting the main logical units of the smartphone application and their interaction.

IV. System Implementation

The mobile app in our SADHealth system is an Android Java application built using SDK version 4.2.2. The program was implemented as a foreground Service to avoid the Android OS terminating it. We ensure that the CPU schedules the event loop all the time through the *wake lock* feature. The internal structure of the app and how it interacts with the external hardware and our Internet cloud server is displayed in Fig. 2. We make use of the light sensor, accelerometer, and location sensors on the phone to collect data for the app. The sensing data are temporarily stored locally on the SD card before being sent in bulk to the cloud.

A. Energy-Efficient System Design

We now describe the basic aspects of the app as follows.

1) *Processing unit*—The Main Android service, which is always running, is basically a set of timers. It works as a scheduler for all the other services to control when to start or stop them. The main service involves no complex arithmetic calculation and has only very few I/O operations.
2) *Memory*—The application consists of several services and activities, but only the main service resides permanently in the phone's main memory while the application is running. Hence, the app does not perform much memory allocation and occupies only little memory.
3) *Data storage*—The recorded data from different sensors are stored in the phone's external storage. These files are labelled with a numeral extension indicating their collection time. Once the recorded data is successfully uploaded to the server, local copies of the files will be removed from the phone to free the storage.

4) *Data transfer*—SADHealth offers both manually initiated and regular automatic mechanisms for uploading the recorded data to the server. It also provides an automatic uploading feature to make the data transfer process more energy-efficient and user-friendly. It determines suitable periods to upload the data to server based on the following conditions:
   a) there is at least 1 day of data stored in the external memory;
   b) there is WiFi connection available with sufficient signal strength;
   c) there is over 70% battery power remaining in the phone.

   If all of the above conditions are satisfied, the app will upload the data automatically.
5) *Sensors*—Sensing measurements are intermittently scheduled by the main service. The sampling rate of a sensor varies based on its sensing type (described in Section IV-B). Nevertheless, no sensor performs continuous sampling.

B. *Energy-Efficient Mobile Sensing*

1) Light Readings: It is important to make sure that the light readings are accurate and frequent enough to gain an representative view of the user's environment. We have two different strategies for recording light sensor data to achieve this.

1) *Periodic sensing*—The proximity sensor is measured every 10 min to determine if anything is covering the front of the phone (e.g., blocking the light sensor). If the proximity sensor is blocked then the sampling of the light will be skipped. Otherwise, the light sensor measures the ambient light for 5 s and only stores the maximum value from this period. This mechanism avoids false zero readings being taken when the phone is covered.
2) *User activated sensing*—Activation of the light sensor is also triggered by the user unlocking the phone screen. Whenever the phone is unlocked, the light sensing service will be called and the maximum light reading in a 5-s timeframe will be stored. This strategy provides the most accurate light measurements as the phone is exposed to light in the same environment as the user. With user-activated sensing, we can also ensure that the phone is not covered or placed inside a bag or pocket.

2) Accelerometer Readings: Most people carry their phones with them a significant proportion of the time. Our app records the phone's accelerometer readings to try and understand the activity level of users. For example, we would like to know when and how much time a user spent sitting stationary or running in a day. Since most human significant activities last for a non-negligible period of time, we believe that periodic sampling over short-time intervals is enough to characterize the activity levels of a person's day. This is different from many existing accelerometer applications for measuring detailed gestures or movement changes in sports, which require high-frequency and continuous sampling.

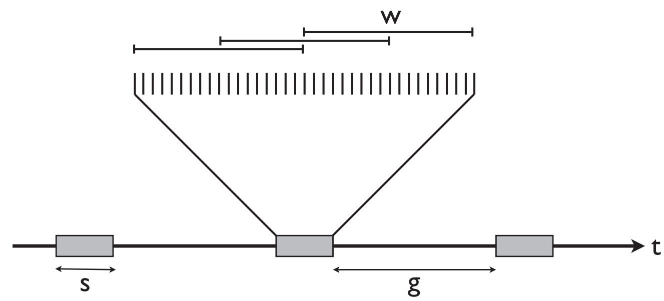

Fig. 3. Acceleration sensing: when the app is running, the acceleration is sampled for short periods s $= 10$ s, with gaps of g $= 110$ s. During this sampling period, many separate acceleration samples are taken at 50 Hz. Multiple overlapping windows of samples are then taken (length $w = 256$ from the set of readings taken in this sampling period).

TABLE II
SAMPLING RATE AND CAUSE OF SAMPLING EVENT FOR DIFFERENT SENSING TYPES IN THE SADHEALTH SYSTEM. LIGHT EXPOSURE IS SAMPLED BOTH PERIODICALLY AND WHENEVER THE SCREEN OF THE PHONE IS UNLOCKED

Sensor type	Sample type	Sampling rate
Light	Periodic On Unlock	5 s every 600 s –
Proximity	Periodic	2 s every 600 s
Acceleration	Periodic	10 s every 120 s
Screen unlock	On Unlock	–
Location	Periodic	Every hour

Periodic sampling here allows simpler and cheaper operation, while still capturing the general trends at a coarse-grain..

Fig. 3 depicts how the samples are taken. The sampling rate of the accelerometer has been set to 50 Hz to record raw acceleration for each axis separately, without any noise filtering. Moreover, to make sure stored acceleration data is independent of the gravity, linear acceleration data are used. Acceleration data measurements are timestamped and stored locally in a file, to be processed later for activity recognition analysis. This classification technique is based on the Weka system and its J-48 decision tree.

The sensing activities that SADHealth performs are listed in Table II. We also specify whether the light sample was caused by periodic sensing or user activated sensing (unlocking the phone). Compared with continuous sampling, periodic sampling of light reduces the sensor operation time to less than 1%. Equally, the accelerometer only samples 8% of the time. Such a large reduction in sampling produces less data and reduces the energy expenditure of the phone significantly. The reduced data size is particularly important for long-term monitoring, since the phones have limited storage and may not want to (or have the opportunity to) transfer large data files.

C. *Data Processing and Communication*

Data processing is performed at multiple different points in the system, as shown in Fig. 4. Many types of raw data are collected from the smartphone (depicted on the left). In particular, the voluminous acceleration data are preprocessed by feature extraction, which reduces its size significantly. The processed

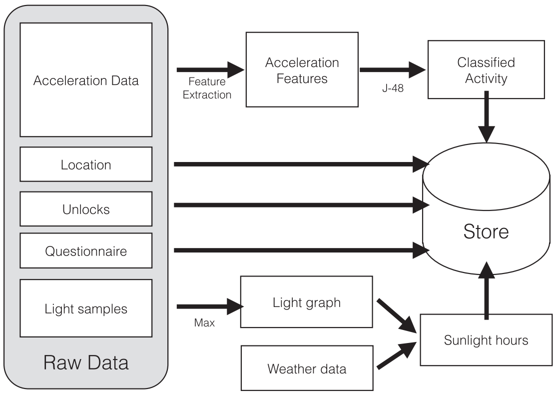

Fig. 4. Data flow diagram showing differing stages of data collection and storage.

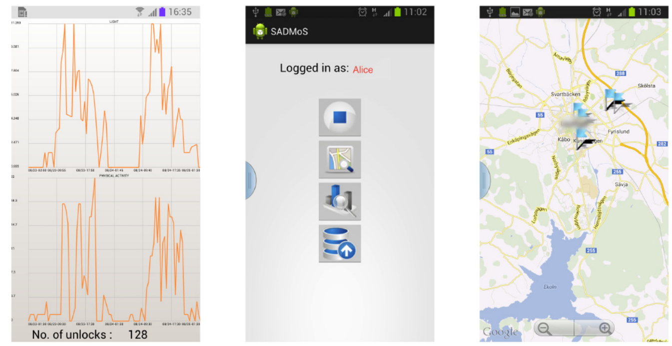

Fig. 5. Application GUI screenshots showing the menu page, data visualization, and location information. (a) User login page with simple start/stop, map, visualization, and upload buttons. (b) User activity and light exposure (inverted colour for clarity). (c) Recent history of user locations.

data can then be communicated over the network to the backend server.

Data pre-processing through feature extraction and format conversion is necessary before feeding the activity recognition tool with our recorded data. Each partially overlapping window from the acceleration sampling procedure is separately processed. For each component of the acceleration (X, Y, and Z dimensions), the following statistical features are extracted:

1) mean;
2) standard deviation;
3) energy of the sequence: $\sum_i i^2/w$, where i represents each reading and $w = 256$, and the window length;
4) Pearson's correlation between each pair of acceleration components ($X{-}Y$, $X{-}Z$, and $Y{-}Z$).

All these above will produce 12 features from each sample window, which can then be processed by the activity classification.

In addition, total sunlight hours from a user's location is also gathered from external weather website[2] to be stored on the data server. This enables an awareness of how much sunlight a user *could* have been exposed to, and can ameliorate the situation where the phone was unable to capture the true light exposure the user experienced (e.g., bad placement or not carrying the phone). After gathering all the data from the smartphones and weather website, the cloud server will perform data analysis on seasonal health and provides long-term storage of the data. It can also summarize the data from different mobile users and observe the change of their health states across seasons. For example, a general relation of light exposure across the seasons can be explored by aggregating the light data collected from experiment participants. It may raise the awareness of users about their light exposure, activities, and seasonal health. As a suggestion, 10–15 min of sun exposure daily could be recommended for users who have spent too little time outdoor.

Similarly, the cloud server can apply data mining techniques to explore the correlation of mood, activity level, and sociability with seasonality. In long term, sequential patterns and prediction could be made to analyze people's health trend and provide early warning of potential risks of health issues for the users. The aggregated data in the server can be presented to individual users and medical professionals for health monitoring, diagnosis and intervention, or medical research purposes.

[2]Weather website [Online]. Available: http://www.wunderground.com/

D. Mobile Application Interface

The SADHealth mobile application is designed to be a simple, informative, and user-friendly application with little required knowledge to operate. A simple login system using third party OpenId authentication system via the user's Google account means people need not even register to begin using the app.

Fig. 5(a) shows the user interface of the app after login. The four buttons provide different functions: starting/stopping the app (top), displaying the location data (second), showing the light and accelerometer data (third), and uploading data to the server (bottom), respectively. A user can visualize his light and activity data as shown in Fig. 5(b). These graphs allow the user to keep track of his light exposure and activity level over time. Similarly, he can see the places that he has visited recently on a map, see Fig. 5(c). From the above data, the user can easily observe the trend of his/her behavior and environment change. Personal sensed data can be downloaded and potentially further analyzed by classification and data mining methods to give more formal interpretation.

V. EVALUATION

The design and implementation of the system evolved gradually as we developed and gained understanding of the important features and technical limitations. Some initial experiments were performed with separate light-sensing hardware before the pure smartphone-based application was built. We shall now describe the most important and informative results. Initial investigations focused on a group of eight volunteers in Sweden using Samsung Galaxy S3 and S4 smartphones running Android OS version 4.1 or above. The experiment was initiated during April 2013 with the idea to monitor users for a long period and observe the cycle of seasonal effects on real people. The initial aims of this early phase of the experiment was to ensure the usability of the app and the accuracy of the collected data. We uploaded the SADHealth mobile app on Google Play store[3] for users to download. It has been downloaded nearly 500 times and used in over 10 countries. Once we were confident about the functionality of the app, a larger scale evaluation was performed with the aim of collecting data from larger group of users over a sufficiently long time to observe seasonal effects on human life. New users were added as the experiment continued and some users optionally left. Overall there was around 50 different users that participated in the experiment at some point with a maximum peak of 30 users in the summer of 2013.

Using a third party application to estimate smartphone sensor's current draw, we measured how expensive using the sensing hardware is. The proximity sensor drew 1.3 mA for 2 s, the light sensor 0.2 mA for 5 s and the acceleration sensors were 0.2 mA for 10 s. Though comparatively small numbers, they will all add up when being used constantly on a mobile device and will contribute to the device not being able to enter sleep mode.

A. Light Data

The most important question to answer when using phones to monitor people's environment is whether the collected data reflect reality. With phones potentially being poorly positioned or absent from the user, they could introduce misrepresentations of the user's environment. Therefore, we conducted a preliminary experiment to determine the correlation between a phone's light sensor and the ground-truth. The ground-truth was recorded by subjects continually wearing an external light-sensing device, the *HOBO* sensor.[4] This small sensor can be attached to person's clothing and set to continuously monitor and record the exposed light levels. Subjects wore the HOBO sensor for a month, with it continually measuring the ambient light every 10 s. The potential range of lux values detected by the phone and HOBO sensor are 76 000 and 32 000, respectively. Both light sensors can easily detect levels up to a bright sunny day.

Fig. 6 displays a plot of the maximum lux measurement made during each 2-h period over the course of a week period, both curves follow the same pattern. Each day in the time period has a peak in the morning when the subject travels to work/university and is exposed to the strongest sunlight. The Spearman's coefficient between the datasets is $\rho = 0.78$ indicating a correlation between both mechanisms of light monitoring. The phone approach performs well at catching the peaks of daytime sunlight (and most of the troughs); however, it occasionally misses the total magnitude during the midrange values of early morning and late afternoon.

It appears that phones are indeed able to gain an accurate picture of the user's exposed light levels. It would not be expected to gain perfect correlation, as both methods are affected by different biases, and can only indirectly measure the true light exposure a person experiences.

The SADHealth app light measurements are taken both periodically and when the user unlocks their phone. We expect that

[3]Google Play Website [Online]. Available: http://play.google.com/store

[4]HOBO Data Loggers [Online]. Available: http://www.onsetcomp.com/

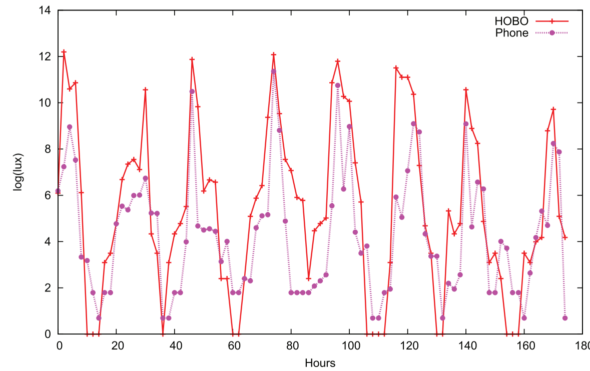

Fig. 6. A week's worth of light readings from the continuously recorded HOBO sensor and the intermittent readings from SADHealth. Correlation coefficient $\rho = 0.78$.

light readings taken during unlocking the screen will be most accurate, since it ensures that the phone screen is not covered and is carried by a user. For the sake of explanation, we shall now discuss a detailed analysis of light measurements from a period of 48 h. We find that there are 337 light samples collected in this period. Out of these 337 samples, 128 samples (38%) were collected when the users were unlocking their phones. The result showed that the phone can collect a large amount of accurate light measurements with user-activated sampling. Apart from user-activated sampling, periodic sampling is performed to obtain adequate data in a regular basis. In periodic sampling, light data are automatically taken only when the proximity sensor is unblocked, to avoid invalid measurements. Our app takes light measurements every 10 min when unblocked (as detected by the proximity sensor). Over the period of 48 h, the proximity sensor was sensed 287 times and detected 78 times (27%) where the phone was blocked. Hence, 209 times (73%) led to successful readings of the light data. The result indicated that periodic sensing can provide useful light measurements most of the time. Periodic sampling could be complemented by user activated sampling to obtain adequate light measurements comparable to an external wearable light sensor.

The changing light over the course of the experiment's whole 2-year period is presented in Fig. 7. There are two curves, the first shows lux sensed by experiment participants and the second plots the amount of sunlight hours as recorded by weather stations. This data was restricted to samples collected by users in the 100-km range of our university town, so that readings would not be confused by users in the southern hemisphere or environments with significantly different climates. Two peaks can clearly be seen in the data, with peaks around June/July and troughs around December/January. This naturally matches the annual solar variation. The curves are due to variation of the weather and the sampling process used to collect it (users may not venture outside on particular days). The data have been processed to allow clear visual presentation of it. First all readings were binned and the maximum light reading selected as that bin's representative value. This curve was still noisy, so Fig. 7 plots the exponentially smoothed time series with factor $\alpha = 0.2$ for clearer presentation. The experienced sunlight data were collected from weather underground and contains their

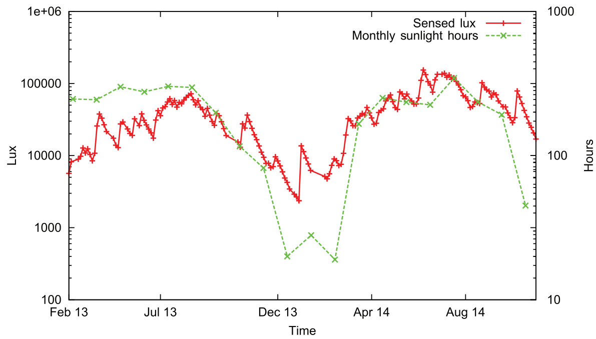

Fig. 7. Two measures of light experienced over a 2-year period 2013–2014. Plotting the light measurements (in lux) from experiment subject's phones and the total number of sunlight hours recorded by the nearest weather station.

data for the *Uppsala/Ultuna* weather station. Sunshine time is defined as the time when the direct solar radiation exceeds 120 W/m^2. The data were collected with an highly accurate Campbell–Stokes heliograph instrument used in weather stations. Obviously, the two curves are measuring fundamentally different, yet highly related, quantities. However, this serves as a sanity check and shows how people are still able to capture some lux even when sunlight hours are extremely low. This may be caused by people making a specific effort to travel outside around lunch time in such northerly latitudes.

B. Activity Data

Similar to the light data collection, we conducted some preliminary experiments to test the efficacy of activity monitoring and classification through mobile phones. We now present a comparison between continuous acceleration logging and periodic acceleration logging. Our experiment involved testing the classification accuracy when applied applied to a continuously logged dataset and when applied to a subset of periodically sampled data. Hence, both datasets were collected over the same time period while a user performs a specific physical activity, The set of activities are *running*, *biking*, *walking,* and *sitting*. accelerometer recording of the same smartphone's accelerometer. The SADHealth classification technique need not be limited to these four activity types, and can easily be extended. Also, the level of intensity may actually be of more interest to health professionals. Recorded activity sessions in both datasets were labelled manually for the ground truth. After preprocessing and feature extraction, the data were used as input to the activity recognition tool. The data was collected over the period May-Aug 2013 and the activities we performed by five different people. We do not differentiate between users when classifying, aiming for general-purpose nontailored behavior, rather than learning individual's patterns. After feature extraction, the continuous and periodic datasets contained 8612 and 689 instances, respectively, which represents a total 38.2 h of recorded and labelled activities.

Fig. 8 depicts the accuracy of classification, comparing continuous sampling with periodic sampling. The results suggest that both sampling methods can achieve high accuracy in activity classification. Our periodic sampling scheme significantly reduces the number of samples to only 8% when compared to continuous sampling. Overall, an activity prediction accuracy of 95.6% was achieved for periodic recordings, whereas activity recognition accuracy for continuous recordings was 97.9%, showing very similar performance.

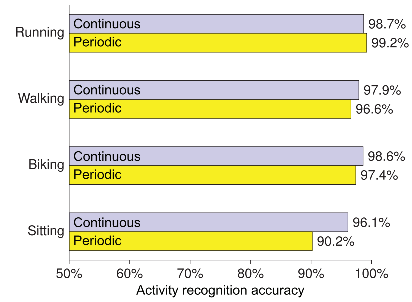

Fig. 8. Accuracy of activity recognition with J-48, comparing continuous and periodic recording. 10-fold cross-validation has been used for evaluation.

TABLE III
CONFUSION MATRIX IN PERCENTAGE FOR CONTINUOUS/PERIODIC SAMPLING

	Classification			
Actual	Running	Walking	Biking	Sitting
Running	98.7/99.2	0.9/0.8	0.2/0.0	0.2/0.0
Walking	0.3/0.0	97.9/96.6	0.4/0.4	1.4/3.0
Biking	0.1/0.0	0.8/0.0	98.6/97.4	0.5/2.6
Sitting	0.2/0.0	3.2/6.9	0.5/2.9	96.1/90.2

TABLE IV
PEARSON CORRELATION BETWEEN ANSWERS ON THE QUESTIONNAIRE

	Mood	Sleep	Energy	Social
Mood	–	0.44	0.75	0.60
Sleep	–	–	0.48	0.52
Energy	–	–	–	0.60

The confusion matrices for both continuous and periodic sampling are shown in Table III. The diagonals indicate the correctly classified activities in percentage, other elements show the percentages of misclassification. The classification is performed by a 10-fold cross-validation of the labelled activity samples. We also observe that both sampling schemes share a similar pattern of misclassification. For example, running is sometimes misclassified as walking due to the similarity of their movements and some sitting data is misclassified as walking.

C. Questionnaire Analysis

To gain a rough understanding of how the users classified their *mood*, *sleep*, *energy levels,* and *social feelings*, as described in Section III-A4. Table IV shows the Pearson correlation between each of the four aspects covered by the questionnaire. It can be seen that all aspects are positively correlated, some heavily. The correlation between mood and energy is the strongest amongst all pairs. Otherwise, social has a very strong connection to the other three. Sleep seems to be the least heavily correlated with the others.

It was observed that different individuals gave systemically different answers to these questions. Some users gave consistently higher or lower values, while some gave values varied to less or greater degrees. This type of noise and bias is expected in user-supplied subjective questions. Aggregate statistics of this data is not their primary use, rather this data is much more informative when considered on a per-user basis.

D. Correlation With Seasonality

We will now focus on a period in the dataset with the greatest darkness, the winter period from late 2013 to early 2014. The results are calculated for each day by picking each users' median hourly light exposure, and then taking the average of the days in each month. As expected Fig. 9 shows changing light exposure during summer (June to August) and winter (November to March). It can be observed that the median lux value is relatively high and increasing in value during the summer months, while the winter shows a decrease in the median lux value. This graph demonstrates that SADHealth can capture the change of light exposure for mobile users across the seasons despite using a relatively low sampling rate. users.

Fig. 9(a) focuses on the correlation of mood and light exposure data from October 2013 to March 2014. We analyzed the data from users who have completed the questionnaire from the app on their mood and sociability states. The solid line in the graph represents the light data (in lux) in log scale. The dotted line shows the mood data collected by the questionnaire in the app. We can see that the mood data show a similar shape as the curve of light exposure. Although the variation of mood is small, it may indicate a weak correlation of mood with seasonal change among our mobile participants. We believe that larger scale experiment will be able to generate more convincing observation about the general public. At the moment, we are able to show the change of mood and light exposure across seasons for individual users, and demonstrate that the data of individuals could be aggregated to draw a summary for the community. Similarly, Fig. 9(b) shows the sociability of users in different months. However, changes in light do not necessarily have a strong instantaneous correlation with the reported mental state, we do seem to find a gentle lasting effect on mood. To show this end, we plot the noninstantaneous correlation between light and mood responses. Fig. 10 shows the corresponding correlation when a delay offset is introduced to the comparison. For example, it plots the relation between experienced light and the reported mood *X* number of days afterwards. It can be seen that although none of the relationships have a magnitude above $\rho = 0.5$, energy and mood are positively correlated with light in the short term, whereas sociability and sleep are negatively correlated, this effect roughly lasts for a month. This data does not exclude confounding factors (e.g., time of year), but does allow some insight to how the factors change with relation to one another. Such plots could prove useful when examining whether a specific user is experiencing seasonal changes.

E. Energy Consumption

The long running nature of these experiments caused us to be concerned about the resulting impact on a phone's battery power consumption, so we shall now present some aggregate data on how much lifetime different phone models experienced while running our app, seen in Fig. 11. It should be noted that the devices were still being operated as user's normal day-to-day phone for making calls, texts, and using the Internet. The three most popular models owned by our users were the Samsung S3, Nexus 4, and Nexus S. We compared the lifetime of the phone's battery running SADHealth using Wakelocks, using Alarm Manager, and not running the SADHealth app at all. We observe that the phones running the app with Alarm Manager have longer battery lifetime than the ones running Wakelocks. The result also shows that the phones not running SADHealth app may have the battery lifetime 1.5 h more than the phones running the app. Given the multiple types of sensing data being monitored continuously on the smartphones, we think that the energy consumption resulted from the app is reasonable.

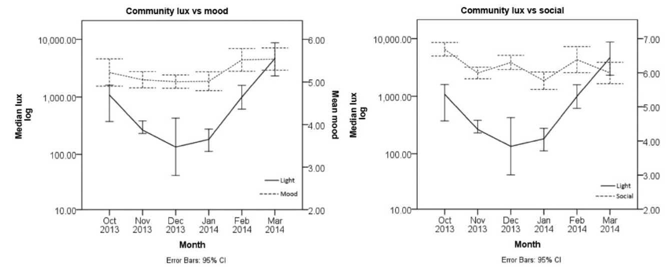

Fig. 9. Experienced light exposure plotted with the user reported mood during a winter period.

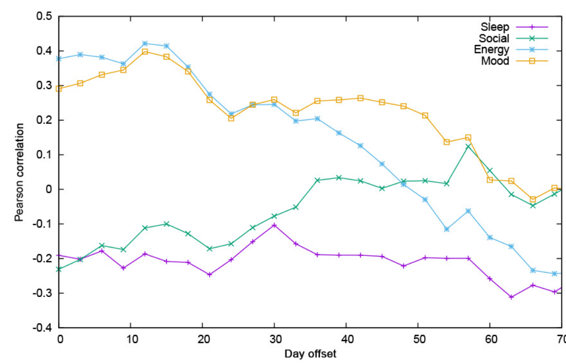

Fig. 10. Changing correlation between light and the four questionnaire results, with gradually increasing delay offsets introduced.

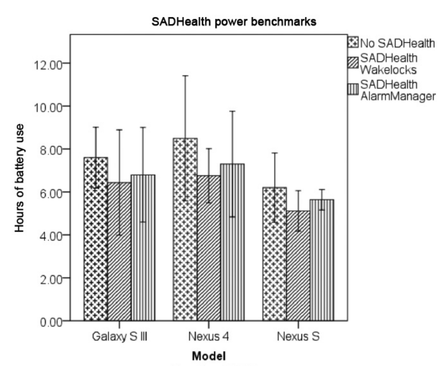

Fig. 11. Energy consumption of the SADHealth app on three different phone models when using three different software scheduling systems.

VI. Conclusion and Future Work

The relationship between people's light exposure, activity level and sociability is of particular interest far away from the equator. This paper presented a light exposure and activity level monitoring system for people concerned with their response to seasonal variations in their environments. An unobtrusive and lightweight Android application was developed which ensured that the collected data were accurate and useful in characterizing a person's environment. Furthermore, we presented data from a 2-year long experiment investigating some relationships between environment and behavior. Users of our system also have the personal benefit of a historical record of their activities, mood and light exposure to help inform their lifestyle choices.

Results showed that periodic and user activated sampling using the proximity and light sensors on the phone can provide adequate and high quality light measurements that are comparable to external sensors. We also demonstrated that intermittent sampling of accelerometer can achieve accurate activity classification and significantly reduce the number of samples for energy efficiency. The sociability and mood sensing questionnaire data although subjective and noisy showed some patterns of depression in winter months. It is not especially practical to consider the collective data, but it is more useful to examine individual behavior. Furthermore, positive correlation was seen between light exposure and mood/energy, while there was a negative correlation with sleep and socializing.

The next step is to extend the study to a greater number and more varied set of participants that are willing to collect data for long periods. Crucial to enabling this is creating an attractive and polished user-interface. To this end, we have begun working with graphic designers and UX specialists to make an improved version of the application. Another required feature is the ability for users to mannually share an informative synopsis of their recent light/activity data to social media (e.g., Facebook and Twitter). This will serve not only to advertise the app to new users but will promote engagement with existing users and encourage reflection on their recent behavior. It could even spur people to arrange exterior social activities to combat winter blues. We believe seasonality monitoring applications such as ours will form a meaningful part of health and wellness technology in the future.

Acknowledgment

The authors would like to thank their previous master students at Uppsala University for their contribution to the SADHealth project, particularly to N. Makvandian, K. Niroumand, and K. Goguev, and all participants in the experiments.

References

[1] L. N. Rosen *et al.*, "Prevalence of seasonal affective disorder at four latitudes," *Psychiatry Res.*, vol. 31, no. 2, pp. 131–144, 1990.

[2] D. H. Avery, D. Kizer, M. A. Bolte, and C. Hellekson, "Bright light therapy of subsyndromal seasonal affective disorder in the workplace: Morning vs. afternoon exposure," *Acta Psychiat. Scand.*, vol. 103, pp. 267–274, 2001.

[3] M. Chen, S. Gonzalez, A. Vasilakos, H. Cao, and V. C. Leung, "Body area networks: A survey," *Mobile Netw. Appl.*, vol. 16, no. 2, pp. 171–193, Apr. 2011.

[4] S. B. Eisenman, E. Miluzzo, N. D. Lane, R. A. Peterson, G.-S. Ahn, and A. T. Campbell, "The BikeNet mobile sensing system for cyclist experience mapping," in *Proc. 5th Int. Conf. Embedded Netw. Sensor Syst.*, 2007, pp. 87–101.

[5] K. K. Rachuri, C. Mascolo, M. Musolesi, and P. J. Rentfrow, "SociableSense: Exploring the trade-offs of adaptive sampling and computation offloading for social sensing," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, 2011, pp. 73–84.

[6] E. Vaara, I. Silvăşan, A. Ståhl, and K. Höök, "Temporal relations in affective health," in *Proc. 6th Nordic Conf. HCI Extending Boundaries (NordiCHI'10)*, 2010, pp. 833–838.

[7] R. R. Fletcher *et al.*, "iCalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 215–223, Mar. 2010.

[8] Y. Ma, B. Xu, Y. Bai, G. Sun, and R. Zhu, "Daily mood assessment based on mobile phone sensing," in *Proc. 9th Int. Conf. Wearable Implantable Body Sensor Netw. (BSN'12)*, Washington, DC, USA, 2012, pp. 142–147.

[9] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas, "EmotionSense: A mobile phones based adaptive platform for experimental social psychology research," in *Proc. 12th ACM Int. Conf. Ubiq. Comput.*, 2010, pp. 281–290.

[10] T. Hao, G. Xing, and G. Zhou, "iSleep: Unobtrusive sleep quality monitoring using smartphones," in *Proc. 11th ACM Conf. Embedded Netw. Sensor Syst.*, 2013, p. 4.

[11] N. Ravi, N. D. P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proc. 17th Conf. Innov. Appl. Artif. Intell. (IAAI)*, 2005, pp. 1541–1546.

[12] J. A. Ward, P. Lukowicz, G. Troster, and T. E. Starner, "Activity recognition of assembly tasks using body-worn microphones and accelerometers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1553–1567, Oct. 2006.

[13] D. Tacconi *et al.*, "Activity and emotion recognition to support early diagnosis of psychiatric diseases," in *Proc. 2nd Int. Conf. Pervasive Comput. Technol. Healthcare*, 2008, pp. 100–102.

[14] R. DeVaul and S. Dunn, "Real-time motion classification for wearable computing applications," Media Laboratory, MIT, Cambridge, MA, USA, Tech. Rep., 2001 [Online]. Available: http://www.media.mit.edu/wearables/mithril/realtime.pdf

[15] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proc. Smart Objects Ambient Intell. Innov. Context-Aware Serv. Usages Technol.*, 2005, pp. 159–163.

[16] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, Jul. 2013.

[17] C. Rastad, "Winter fatigue and winter depression prevalece and treatment with bright light," Ph.D. dissertation, Dept. Public Health Caring Sci., Uppsala Univ., Uppsala, Sweden, Mar. 2009.

[18] A. F. Dalton and G. O'Laighin, "Identifying activities of daily living using wireless kinematic sensors and data mining algorithms," in *Proc. 6th Int. Workshop Wearable Implantable Body Sensor Netw. (BSN)*, 2009, pp. 87–91.

[19] C. M. Morrison and H. Gore, "The relationship between excessive Internet use and depression: A questionnaire-based study of 1,319 young people and adults," *Psychopathology*, vol. 43, no. 2, pp. 121–126, 2010.

[20] A. Magnusson, "Validation of the seasonal pattern assessment questionnaire (SPAQ)," *J. Affective Disorders*, vol. 40, no. 3, pp. 121–129, 1996.

Liam McNamara received the Ph.D. degree in computer science from the University College London, London, U.K., in 2010, on the topic of content distribution in short-range wireless networks.

He is currently a Researcher with the Networked Embedded System Research Group, Swedish Institute of Computer Science (SICS), Stockholm, Sweden. He was a Postdoc Researcher with Uppsala University, Uppsala, Sweden, from 2011 to 2013. His research interests include mobile sensing, home Internet connection monitoring, and visual light communication.

Edith Ngai received the Ph.D. degree in computer science and engineering from Chinese University of Hong Kong, Shatin, Hong Kong, in 2007.

She is currently an Associate Professor with the Department of Information Technology, Uppsala University, Uppsala, Sweden. She was a Postdoc with Imperial College London, London, U.K., from 2007 to 2008. She is a Visiting Researcher at Ericsson Research Sweden, Stockholm, Sweden, from 2015 to 2016. She has conducted research with the University of California, Los Angeles (UCLA), Los Angeles, CA, USA; Simon Fraser University, Burnaby, BC, Canada; and Tsinghua University, Beijing, China. Her research interests include Internet-of-Things, network security and privacy, smart city, and e-health applications.

Dr. Ngai is a member of ACM. She has served as a TPC Member in leading networking conferences, including ICDCS, Infocom, ICC, Globecom, IWQoS, etc. She was a Program Chair of ACM womENcourage 2015, a TPC Co-Chair of the IEEE SmartCity 2015 and the IEEE ISSNIP 2015. She has served as a Guest Editor in special issue of the IEEE INTERNET-OF-THINGS JOURNAL, IEEE TRANSACTIONS OF INDUSTRIAL INFORMATICS, and *Mobile Networks and Applications (MONET)*. She is a Technical Committee Member of Computer Communications and an Associate Editor in EAI endorsed Transactions on Collaborative Computing. She is a VINNMER Fellow (2009) awarded by VINNOVA, Sweden. Her coauthored papers have received Best Paper Runner-Up Awards in the IEEE IWQoS 2010 and the ACM/IEEE IPSN 2013.