SEAMLESS HUMAN LIFE MONITORING AND TRACKING ALL-DAY LONG

by

Ali Yavuz KAHVECİ

B.S., Computer Engineering, Boğaziçi University, 2011

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computer Engineering
Boğaziçi University
2015

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my thesis supervisor Prof. Cem Ersoy for his patience, encouraging comments, guidance and measurable support. I have learnt a lot of things from him which will lead me in my future life.

I would like to thank Hande Alemdar for her continuous support during my thesis. To work with her was a great opportunity for me. I will never forget her voluntary help in my troublesome times. Also, I would like to thank Behçet Meyveci for his invaluable contributions to the preliminary studies.

I would like to thank to my colleague Yüksel Arslan, İbrahim Hökelek, Fatih Kara, and my supervisor Hikmet Aşmer from TÜBİTAK for their support and enjoyable environment they created during my studies. Special thanks also to my friends Akif, Ali, Asil, Beytullah, Recep, Süleyman, and Yusuf for their patience and collaboration during my data collection process which is the main part of my thesis.

I would want to thank to my fiancé, Tuğçe Özkaptan, who encouraged me, fed me and made me feel I am not alone. Her existence right beside me give me an incredible energy in the hardest times of my thesis. I would also want to thank my lovely sisters, Hilal Kurumahmut and Selcen Haskılıç and my father Muhsin Kahveci for keeping their support on me all the time and their encouraging talks during my long hardworking hours.

Finally, I want to thank sincerely to my mother, Nezahat Kahveci. It is the biggest chance in my life to have such a mother. Therefore, I would like to dedicate my thesis to her.

This research is supported by Boğaziçi University Research Fund (BAP) under the grant number 8684.

ABSTRACT

SEAMLESS HUMAN LIFE MONITORING AND TRACKING ALL-DAY LONG

Human life monitoring systems utilizing wireless sensors networks (WSNs) and/or smart phones became a hot topic for the evaluation of the life quality. A daily life of a human can be divided into three main parts as outside, home and sleep. There are various systems that monitor the lifestyle of a human either inside or outside the home. Yet, the challenge is to develop a system that covers all the activities of a person for 24 hours. Considering the fact that people spend one-third of their lives sleeping, sleep is another important activity to monitor. While sleep studies mainly focus on the sleep quality of a person, the effects of life style and ambient factors on the sleep quality are usually neglected. In this thesis, we propose a seamless human life monitoring system that covers 24 hours of a person's life including the sleep activity. The proposed system utilizes a WSN and a smart phone and collects life-log and sleep data from multiple users. The WSN collects nocturnal ambient and sleep data via various sensors. On the other hand, the applications running on the smart phone collect daily performed activities with their durations and locations. In the data collection phase, we employed nine people for fifteen days. The system is designed to provide the unobtrusiveness and respect the privacy of the users. By using the collected data, we extracted the sleep behavior and the life style choices of the users. In order to figure out the factors affecting the sleep quality of a person, we applied three feature selection algorithms namely the decision tree, the correlation coefficient, and the sequential feature selection to the collected data. The results indicate that two features namely the *Sleep Onset Latency* and the *Leisure Activity Duration* are reported as important features in all of three algorithms for their effects on the sleep quality.

ÖZET

GÜN BOYU KESİNTİSİZ İNSAN YAŞAMI TAKİBİ VE GÖZLEMİ

Kablosuz ağ algılayıcıları (KAA) ve akıllı telefonlar kullanılarak geliştirilen *Yaşam Takip* sistemleri yaşam kalitesinin çıkarılmasında sıkça kullanılmaktadır. İnsanların 24 saatlik yaşamı ev içi, ev dışı ve uyku olmak üzere üç ana kısma ayrılabilir. Kişinin ev içi veya ev dışı yaşamını takip eden ve yaşam kalitesini değerlendiren birçok çalışma mevcuttur. Fakat, asıl zor olan kişinin 24 saati boyunca gerçekleştirdiği eylemlerin takibini yapan bir sistemin geliştirilmesidir. İnsanların hayatlarının üçte birini uyuyarak geçirdiği düşünüldüğünde, uyku da ihmal edilmemesi gereken önemli bir eylemdir. Uyku çalışmalarında genellikle kişinin uyku kalitesine odaklanılırken, uyku kalitesini etkileyen etkenler göz ardı edilebilmektedir. Ayrıca, kişinin yaşam tarzının uyku kalitesine önemli etkisi olduğu düşünülmektedir. Bu amaçla, kişinin uyku da dahil 24 saatlik yaşamını kapsayan kesintisiz bir eylem takip sistemi geliştirdik. Önerilen sistem bir KAA ve bir akıllı telefon kullanarak, insanlardan yaşam günlüğü ve uyku verilerini toplamaktadır. KAA çeşitli algılayıcıları kullanarak gececil çevresel değerleri ve uyku verilerini toplamaktadır. Diğer taraftan, akıllı telefon üzerinde çalışan uygulamalar kişinin gerçekleştirdiği eylemleri ve bulunduğu yerleri süreleriyle beraber kayıt altına almaktadır. Deneysel veri toplama aşamasında dokuz kişi on beş gün boyunca sistemimizi kullanmıştır. Sistem geliştirilirken kullanıcılara rahatsızlık vermemesi ve onların özeline saygılı olması hedeflenmiştir. Kişinin uyku kalitesini etkileyen öğelerin bulunması amacıyla, toplanan verilere üç ayrı öznitelik seçim yöntemleri uygulanmıştır. Deneylerin sonucunda, *Uyku Gecikmesi* ve *Serbest Etkinlik Süresi* öznitelikleri bütün yöntemler tarafından uyku kalitesini etkileyen önemli etkenler olarak raporlanmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

acc^T_{mov}	Accelerometer movement threshold
mov_{max}	Maximum movement count
mov^T_{min}	One-minute movement threshold

LIST OF ACRONYMS/ABBREVIATIONS

API	Application Programming Interface
dSQI	Daily Sleep Quality Index
ECG	Electrocardiography
EEG	Electroencephalogram
GPS	Global Positioning System
GPIO	General Purpose Input/Output
MCU	Micro Controller Unit
NREM	Non-Rapid Eye Movement
PIR	Passive Infra-Red
PSQI	Pittsburgh Sleep Quality Index
PSG	Polysomnography
RAM	Random Access Memory
REM	Rapid Eye Movement
RR	Respiration Rate
SFS	Sequential Feature Selection
SOL	Sleep Onset Latency
SWS	Slow Wave Sleep
USB	Universal Serial Bus
UTC	Universal Time Clock
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network
WSN	Wireless Sensor Network

1. INTRODUCTION

"Life Monitoring" became important in many real-world applications for the computer science community in the recent years. With the remarkable advances on wireless sensor networks (WSNs) and smart phone technologies, it became relatively easy to figure out what a person is doing in a specific time period. In order to monitor a certain activity, many sensors such as accelerometers, global positioning system (GPS) receivers, pressure sensors, and passive infrared (PIR) sensors could be utilized. Although there are many life monitoring solutions both for inside and outside the home separately, it is very hard to find a solution covering activities both inside and outside the home. This thesis proposes a seamless daily human life tracking approach by making use of a WSN together with a smart phone. This approach divides the daily life into three parts as outside, home and sleep, and determines the activities and ambient factors affecting sleep quality.

WSNs attracted researchers with their low cost, increased computational power, decreased power consumption and portability for a variety of applications. Life tracking solution for healthcare of elderly in a home setting is a good example of WSN usage for smart homes [3]. As mentioned in [4], a wide range of sensors could be taken into consideration for inferring the activities of a person. Whereas, since the mobility of WSN solutions are restricted, the coverage area is limited to home or near-home environment. Hence, the outside life of a subject in the experiment of life monitoring with WSNs stays unknown.

On the other hand, smart phone-based life tracking solutions are capable of activity extraction outside the home since almost everybody has a smart phone equipped with many useful sensors. Although there are many smart phone operating systems, due to its ease of use, highly sophisticated Application Programming Interface (API) and allowance to manipulate its file system, Android smart phones are preferred. Daily life monitoring and tracking applications as in [5–7] generally use accelerometers and GPS receivers data for the inference of outside activities. However, when the person goes home and transits to a relatively stationary situation, the smart phone-based life tracking algorithms fail since those algorithms mainly depend on mobility and many smart phone users put away their phones at home.

Both of the life tracking solution approaches have their own deficiencies but they could be complementary to each other. In this regard, we propose a solution by using both of them wherever they are necessary and effective in real life.

When it is considered as Seamless Human Life Monitoring and Tracking All-day Long, the research focus evolves to life-logging of a person. The concept of life-logging was first proposed early in 1945 by Vannevar Bush [8]. Life-logging concept is basically collecting any useful information from a person and his/her environment that would be important to change and/or determine his/her daily life routine. Life-logging studies can be implemented in many fields like healthcare domain by using many technologies like WSNs and smart phones. As mentioned in [4], studies about life-logging area focus mainly on healthcare of the elderly since the civilized societies are getting older and it becomes harder to acquire a caregiver to every elderly human who is in danger by living alone in a non-hospital environment. Due to inevitable healthcare service necessities, remote and automatic monitoring of the human wellbeing and detection and warning systems should be taken into consideration in order to help the elderly increase their quality of life and live independently with minimum support.

While studying life-logging systems, the privacy of the subjects and the unobtrusiveness of the system should not be disregarded. Privacy issue is a key concept if it is desired to be used in the field excessively. Even though Kinect for Windows [9] is a promising life-logger system, it takes visual data from the subjects and his/her environment by violating the privacy issues. On the other hand, Sony Smartwatch [10] can be used in life-logging systems to achieve real-time results both by informing the subject and sending information about the subject to the server-side. However, the subjects are required to wear the Smartwatch even in their homes. Hence, it becomes an obtrusive equipment to live with.

Considering the life-logging concept, sleep study is an area of research that should not be disregarded. As stated in [11], sleep which occupies about one-third of our life time is undoubtedly an important activity to monitor. Sleep deprivation due to sleep-related disorders or other factors could introduce severe physical effects, cognitive impairments, and mental health complications. During sleep, the body restores metabolic functions and stores information collected on the day [12]. Moreover, Miwa *et al.* [13] describes sleep as refreshment from mental and physical fatigue and lack of sleep decreases mental concentration and immunity. Proper sleep is essential for healthy activity and mental function in all aspects of daily life. Hence, sleep is a crucial actor on a person's daily life and wellbeing. While good quality sleep can increase human's enthusiasm and helps concentrate better on his/her work during the day, an insufficient sleep may trigger many health problems. When the mental and/or physical condition of the sleeper and the ambient factors are changed, the functionality of the subject is impaired, causing problems such as somnolence, decrease of reaction time, loss of coordination that may cause traffic accidents, irritability, cardiovascular disease (CVD) such as hypertension, arrhythmia and others affecting the life quality of the subjects [14, 15].

Since sleep constitutes a big portion of our lives and is a major part of health and well-being, monitoring the quality of sleep can aid in the medical diagnosis of a variety of sleep and psychiatric disorders and can serve as an indication of several chronic diseases [16]. One should consider monitoring and investigating sleep behavior of a person in order to get a complete and informative life-log.

There are many sleep quality studies. One of them is Pittsburgh Sleep Quality Index (PSQI) [17] which is a self-rated questionnaire. It assesses the sleep quality and disturbances over a one-month time interval. In the questionnaire, there are sleep-related questions scoring a subject's sleep quality between 0 and 21. PSQI score and sleep quality is inversely proportional since increase in the PSQI score refers decrease in the sleep quality. Although PSQI has an important role in the sleep studies, it is filled monthly and the questions are subjective. Thus, the sleep score of a person is not necessarily an accurate indicator of the sleep quality.

The most accurate results on the sleep quality can be taken from sleep labs established in hospitals. A patient spends a night in the sleep lab and gives information via the sensors connected to him/her. Furthermore, a night watch monitors the sleeper through the night. The technique applied in these labs is polysomnography (PSG) which is the gold standard test for assessing disorders related to sleep. In PSG, multiple sensors are connected to the subject to record parameters during sleep and to determine if the patient suffers from a disorder [18]. However, too many sensors connected to the body may affect the measurements. Furthermore, PSG is an expensive exam that requires hospital admission for the night. The number of people who undergoes this exam is few compared to the potential number of subjects who could benefit from sleep disorder monitoring. The unfamiliar environment of the sleep laboratory as well as the equipment and sensors attached to the subject's body can disturb the subject's quality of sleep – a phenomenon referred to as the first-night effect [19].

Every sleep study focuses on the detection of sleep stages which are rapid eye movement (REM) sleep and non-REM (NREM) sleep. The American Academy of Sleep Medicine (AACM) further divides NREM sleep into three distinct stages [20]. In NREM stage 1, sleeper is in between sleep and wakefulness. The muscles are active and the eyes roll slowly. In NREM stage 2, sleeper becomes gradually harder to awaken. NREM stage 3 is called as slow-wave sleep (SWS). The sleeper is less responsive to the environment. In REM sleep, most muscles are paralyzed. Even though the sleeper exhibits electroencephalogram (EEG) waves similar to waking state, he/she is harder to arouse than any other state. The oxygen consumption by the brain is higher than when the sleeper is awake. Most of the dreams are seen in this stage. A healthy adult reaches REM approximately every 90 minutes with the latter half of sleep being more dominated by this stage. The sleep proceeds in cycles of REM and NREM usually four or five times per night. Greater amount of NREM stage 3 happens earlier in the night. REM sleep comprises about 20 to 25% of the total sleep in typical healthy adults [21].

In the sleep study literature, NREM stage 2 and NREM stage 3 are referred as deep sleep while NREM stage 1 and REM sleep represent light sleep. The quality of sleep is often an indicator of deep sleep duration. In our thesis, we focus on the detection of sleep stages as light and deep sleep. As a start of the sleep study, there are some straightforward inferences. Guazzelli *et al.* [22], propose that the total amount of time which spent in deep sleep decreases with age, while an increase in wakefulness has also been observed. Since the age has been known to have a significant effect on the sleep structure, we run our experiments on people with similar ages to accomplish a model free from age-related matters.

Creating a sufficient sleep stage detection algorithm could be achieved by isolating the sleeper from negative influences caused by devices used in the data collection process. Unobtrusiveness of the devices used to collect sleep data from the subject is a key factor to accomplish a high performance and comfortable sleep stage detection system. Hence, wireless body-area networks and wearable devices decrease the overall system performance. Systems proposed in [13, 23] place limitations on the subject's movement and it is more difficult to use in an ambulatory environment.

Another important factor that should be taken into consideration is the cost of the system. Although Actigraph [24] is a successful device to detect sleep stages, it is not preferable to utilize for the systems targeting low-cost because its price is as high as $1000. Furthermore, since the micro-controllers in WSNs operate with battery, the power consumption is an important constraint. While developing a wireless sensor network, the developers should consider the power consumption of the sensors.

In this thesis, we design, develop and deploy a seamless human life tracking system by making use of WSNs and smart phone applications. The ultimate system is deployed in real houses. In total, nine people used the system for fifteen days in their own houses.

Our contribution can be summarized with the following items;

- The system is designed to be fully unobtrusive and respectful to the privacy concerns of the subjects. In this manner, no wearable devices that would probably disturb the subject are preferred. Moreover, no data that could contain any private information such as video or audio are sampled and collected. The overall system is designed considering minimum dependence to the users.

- The system collects activity and location data from the subjects during the day including in-house activities. When the subject goes to bed, sleep data collection process begins. Both sleep-related and environmental data are collected during the person's sleep. Environmental data include noise level, temperature, humidity, light level and atmospheric pressure. The WSN system consists of five different sensors. For the life-logging study, four different Android applications are utilized.
- For the data collection process, a WSN kit and an Android-based smart phone are enough. The experiment continues fifteen days for one person. In order to accomplish the data diversity, nine different subjects are used in the experiments. The subjects are of similar age with the same gender, male. The subjects are chosen with similar ages in order to eliminate age-related versatility in sleep quality since the time spent in deep sleep decreases with age as mentioned in [22].
- Unlike many studies, our system runs in the people's own environment in which we can not intervene in case of any problem. As a result of this, there are some erroneous and/or inadequate data from some users. Furthermore, we are able to test the system's robustness and lifetime of the batteries in the field.
- In the feature selection phase, the sleep data is classified as good, bad or moderate by detecting the sleep stages and their durations during the night. After that, daily activity details and GPS data are used together with sleep quality data. At the end, we are looking for the factors that affect the overall quality of sleep.
- In the preliminary studies, the goal was to detect sleep stages as light and deep sleep from the accelerometer and the pressure mat data. The field experiments of this preliminary study were handled by two subjects for thirty nights. The former studies are explained in detail later in the thesis.

The system deployed to people requires five sensors, two micro-controllers with transceiver modules, one base station collecting data from the sensors, one computer that runs as an online data collector and a smart phone with a GPS module. The system collects data from the subject and his/her environment regarding the privacy issues. The feature selection process takes place after the data collection phase is completed.

The organization of the thesis is as follows; in Chapter 2, we provide a literature survey on life-logging and sleep quality studies. In Chapter 3, the system design and its components are explained in detail. In Chapter 4, information about the preliminary experiments of the thesis is made available. In Chapter 5, we use feature selection techniques to find out the factors affecting the sleep quality of a person. The conclusion and possible additions and modifications for the future studies take place in Chapter 6.

2. LITERATURE SURVEY

Human life tracking and monitoring studies have found a huge application field in computer science. Almost all of them have utilized WSNs and smart phone technologies. In this chapter, we explain the state of the art studies in two subsections life-logging and sleep study. Their proposals, advantages and disadvantages are explained by comparing with our study.

2.1. Life-logging Domain

Smart Reality Testbed [25] which is a life-logging platform is developed together with an interactive visual analytics tool to help users gaining insights about their daily life. In this study, 50 university students who have Android smart phones were asked to use the application for a month. The students are in different departments and of similar ages. The application collects GPS sensor data and application usage data. For processing the logged data, clustering technique is the most preferred descriptive modeling technique. Android application that collects life-log data from the smart phone is partly based on Funf framework which is an open-source library set developed by Massachusetts Institute of Technology. The application collects positioning, social, motion, environment, device and interaction information. The GPS sensor is sampled 4 seconds in every 2 minutes. The application usage data is triggered with screen-on, screen-off and app-switch events. The time passed between these events constitutes life-log data to be collected. In order to eliminate jitter and jiggle of the correct location, fixation filter is used. The visualization tool has three parts as information panel, timeline view, and map view. The visualization tool enables users to see and examine their daily life pattern. At the end of the experiment, it has been revealed that most people use their smart phone more than three hours a day and in some part of the day, users use their phone continuously for an hour. Most of the time, users use social networking applications like Facebook. In their study, they mostly focus on smart phone usage data whereas we want to examine daily behaviors of people by using the smart phone data.

In another study [26], the main focus is the health care monitoring of the elderly patients. Data acquisition system consists of handheld and PC devices namely MobileHub and HomeHub. With this solution, collection of vital information about the patients can be handled remotely. The goal of this study is to enhance the health care service by means of cost and effectiveness. The system collects and monitors activities and vital signs like blood pressure, blood sugar, heart rate, etc. In case of an emergency, caregivers are alerted by the system. Remote patient monitoring consists five services as follows; data acquisition, store/forward and visualize service, life tracking service, behavior monitoring and lifestyle guidance service. The experiments take place in the environments called Living Lab. With Living Labs in different cities of the country, the type of patient environment is homogeneous. The HomeHub runs on a Linux environment and has more computational power that MobileHub. On the other hand, MobileHub does not restrict the patient in home. The system collects information about patient's blood pressure, blood sugar, oxygenization level, ECG (Electrocardiography) signals, patient movement, medication, eating habit and activities. This study is not a cost-effective solution for our study and we aim unobtrusiveness whereas there are many on-body wearable sensors in this study which prevents getting real sleep performance by disturbing the patient.

In [27], the research is mainly on supporting elderly people. In order to accomplish, wireless optical oscillo-sensor, WSN, smart phone devices are used. Optical oscillo-sensor is used for estimation of elderly people's state inside the bed. WSN is for predicting the state of elderly in the room. The smart phone is used for destination of movement types outside the home. The collected data is stored in a structured platform which is called Informationally Structured Space. The robot partner is composed of Iphone, servo motor, mobile base and a microcontroller. WSN is composed of sensors equipped with wall, floor, ceiling, furniture and home appliances. WSN measures the environmental data and human motions. WSN has Kinect [9], SunSPOT [28], and wireless optical oscillo-sensor. For the bed part, the human states are classified into sleep, roll-over, cough and none. The smart phone logs human movement acceleration, angular velocity, and movement distance. The robot partner takes place when passive perception does not return any successful state. In the visualization system, human behavior, human state and life log are demonstrated. The main problem of this study is the violation of personal privacy. Taking the visual information of the room seriously violates the human's privacy. Moreover, the activities outside cannot be logged. Only movement-related information can be collected by the current system.

Real-time social Reminder and Reminiscence System (R^3S) [29] is an assistive technology that harnesses mobile and wearable technologies to address the unmet needs specifically relating to personalized information, social contact and perceived safety. The goal of the study is to answer the question of "by using the technology, is it possible to improve the independence of a person with dementia and enable them to live longer within their own home?" The life-logging component of the system, miLifeCom, is composed of an Android smart phone and Sony Smartwatch [10] wearable display. The user should wear the smart phone around the neck. The application records an image in 30 seconds intervals together with location and accelerometer data. When a known-person's image is taken, the image is processed and the information about that person like name, relationship status is displayed on the smartwatch. Additionally, to record user's heart rate, a Bluetooth sensor is connected to the user's chest. For the experiments, six people wear the setup for 2.5 hours taking around three hundred pictures. As a result of the experiment, fast and accurate interactions are identified within 5 seconds. Slow interactions take more. The main deficiency of this study is that the system does not provide wide range of data. It only takes pictures for image processing. Moreover, the sensors prevent users from their daily routine activities. The garment worn by the users is uncomfortable and disturbing.

In [30], life-log means photography, physiological signals, activity time, and location data. By wearable gadgets, daily information of a user can be acquired and logged. There are two kinds of wearable gadgets for low rate and high rate data. Universal Time Clock (UTC) is used for synchronization of data coming from different gadgets. In total, there are four gadgets for audio, position, video and motion and bio-information. GPS is utilized to retrieve the exact time and location information. With the usage of acceleration sensor, user's movement status can be obtained. The collected data can be transferred to the server or other device by ZigBee. The boards mainly consist of a microcontroller unit (MCU) with WPAN, GPS module and related sensors like accelerometer, ECG analog sensor, camera module, etc. Wearable gadget for high data rate is composed of embedded processor, RAM, flash memory, video I/O interface. It supports WLAN and WPAN for communication.The main problem of this study is that unobtrusiveness of the gadgets is disregarded. Furthermore, by taking photos and videos, the privacy of the user and his/her companies are violated.

Kim *et al.* [31] defines personal life-logging as data logging activity recording and tracing one's personal real life event. Life-log research is an effort to analyze such real life data logs to create informative and digitized stories of person's life. The goal of this study is to design the back-end life-log system that can support and facilitate life-logging practice and personalized analysis on their records. The system is composed of lightweight neck-held camera and a smart phone with the embedded GPS. Moreover, magnetometer, accelerometer, thermometer and ambient light sensor data values are sampled. The camera's sampling time differs according to environmental changes. The camera takes more pictures when environmental changes increase. In the experimental phase, a person's journey from his house to the work office is taken. After the journey, post-processing of camera images is done on the server computer. With image processing, the people encountered during the journey can be retrieved if they are known by the user.The problems for this study are the violation of privacy issues due to picture taking and obtrusiveness of wearable sensor set as camera and smart phone.

2.2. Sleep Study Domain

Iwamoto *et al.* [32] examines sleep stages by measuring respiration count with the usage of Doppler sensor. For ground truth, ECG signals are used. The Doppler sensor is placed below the bed. For sleep stage determination, the fact that number of respiration increases during REM sleep is used. In the experiment, respiration count over 18.1 per minute is assumed as REM sleep. It is assumed as light sleep if respiration count is between 16.7 and 18.1 per minute. It is assumed as deep sleep if respiration count is below 16.7 per minute. According to experiment result, the accuracy of Doppler sensor is 90.62%.The problem of this study is that the subjects are assumed as healthy. The system cannot get correct results for people with sleep apnea disorder. Moreover, if the sleeper lies down on his/her face the respiration count would be miscalculated.

The main focus in [33] is to release the subject from cables of sleep-related sensor. In this regard, a WSN consisting of four different sensors is used. The sensors are breathing belt, temperature sensor, oxygen saturation sensor and heart rate sensor. The collected data are sent to a central switchboard unit using SimpliciTI technology. The center unit arranges the data and transmits to the end-device using Bluetooth technology. The problem of this study is that, even though it is argued that releasing the subject from the cables makes the system unobtrusive; the sensors are attached to the subject's body. Hence, the system is still obtrusive.

In [13], SenseWear Pro2 Armband is used to detect roll-over movements during sleep. It is worn around the right upper arm and measures acceleration in two axes, skin temperature, near-body temperature, heat flux and galvanic skin response. The success rate of roll-over detection algorithm is 84.2%. for the ground truth, a video camera is used. From this point, the sleep quality is calculated. As a result of the early experiments, sleep quality and sleep duration are inversely proportional. For the final experiment, three healthy and 2 major depressive disorder (MDD) patients are examined. The subjects wear SenseWear for about one year. According to experiment results, even though the difference in sleep duration between healthy subjects and MDD patients is small, average sleep quality in MDD patients is significantly lower. As a result, it appears MDD reduces the sleep quality. There are two problems of this study; SenseWear supplies acceleration data from two axes. Third dimension is disregarded. Moreover, the sensor is obtrusive since it is attached to the body and disturbs the subject.

In another study [34], the system is composed of pressure sensor array, airflow sensor, signal multiplexing, resistance-voltage transducer and microcontroller with analog to digital converter (ADC), multiplexer and serial communication to data logger. The body movements and respiration count are calculated. The system is tested with five subjects producing eight full night sleep totaling 39 hours of sleep. By using the collected data, body movement rate, apnea level, respiration rate (RR) calculation, RR variability parameters are calculated. Body movement detection algorithm is implemented detecting large amplitude changes during 2-second window in at least three signals to determine if a movement event is occurred. Although it is a very successful study and its calculations are very close to real values, the main problem is the mobility issue. The system cannot be moved by the user to another place. Thus, it becomes very similar to sleep rooms in hospitals.

The focus of study in [35] is the unobtrusive assessment of movements in bed using data from load cells installed under each support of a bed. Load cell data classifies movements into two classes as body movement and leg movement. Load cells that are strain gauge transducers convert applied force into a resistance change. The output data are proportional to force. Six load cells with *100kg* capacities are used. Seventeen users participated to the study giving one night data. The sleep data constitutes a trajectory. The points in the trajectory are used to approximate the body center of mass trajectory. Three features are extracted from the trajectory as the distance between initial and end points, the length of the trajectory and the variance of the trajectory. Using these features, user's leg movements are scored as periodic leg movement (PLM). In this study, the sleep quality has been disregarded. PLM score gives information about the user's possible health problems but does not give general sleep quality of the user.

Samy *et al.* [36] describe an unobtrusive framework for sleep stage identification based on a high-resolution pressure-sensitive e-textile bed sheet. From this study, a set of sleep related biophysical and geometric features from the bed sheet is extracted. Two phase classification procedure for awake-NREM-REM stage identification is performed. The bed sheet is a matrix of pressure sensors generated by the intersections of conductive buses. For the performance evaluation, three classifiers including K-nearest neighbor (k-NN), Naive Bayes and Support Vector Machine (SVM) are used. The pressure images are divided into thirty-second-epochs for identifying the sleep stages. From the biophysical information, the following features are extracted; respiration rate (RR), its variability, leg movement, body movement, posture and body orientation. For the experiment, seven subjects participated in the study, supplying a full-night sleep data. The results are shown that Naive Bayes classification method outperforms the others and the overall performance of this study is 71.1% according to PSG analysis. The problem of this study is that the performance of their system is evaluated by using PSG analysis as a ground truth. However, the environmental factors and subject's current well-being are disregarded although these factors significantly affect the sleep quality.

The study in [37] describes the derivation of metrics by means of a location mapping model to determine characteristics of sleep patter including bed time, rise time, sleep disturbances and time in bed. The main contribution is the development, validation and implementation of an ambient night time behavior monitoring system. This system is used in sixteen independent living apartments in which elderly people reside. Each of these apartments is fitted with various sensors including PIR sensor for motion detection, contact sensor on all windows and doors, sensors on all light switches, temperature sensors in each room, brightness sensors in each room and sensors of electricity and heath usage. Moreover, PSQI (Pittsburg Sleep Quality Index) questionnaire is filled by the residents periodically. Additionally, a wrist-worn tri-axial Actigraph [24] is used for night time activity detection. A sleep disturbance is determined by a change in location between the bed time and rise time. Time in bed is calculated as the duration between bed time and rise time removing the sleep disturbance periods. Time outside the home is calculated as the duration between exterior door closing and re-opened. "Fell asleep", "woke up" and "actual sleep" metrics are derived from Actigraph data. As mentioned in the study, wrist actigraph compliance was found to be disturbing. Two participants take the Actigraph off for many nights making the night invalid. This behavior of residents shows that the obtrusiveness of the wearable sensor prevents the study from getting real sleep data.

Another study [38] proposes a system that enables monitoring sleep stage of a person remotely. The system can estimate sleep stage in real-time from heart beat and body movement data acquired by an unconstraint-typed pressure sensor. The instant sleep stage can be remotely displayed in tablets of related people. For the sleep stage estimation, Database-based Compact Genetic Algorithm (DcGA) is developed. The main role of DcGA is to investigate whether each frequency of the heart beat data should be passed or filtered according to the past data. The problem of this study is that it needs past data to estimate sleep stage from every person. This system would achieve better results as the past data increase. At the early stages, the performance of DcGA would be poor.

In [39], the effectiveness of three methods in sleep stage detection is investigated. These methods are fixed-duration five-second epoch (M1), three consecutive one-second sub-epoch (M2) and three non-consecutive 1-second sub-epoch (M3). EEG signal of the PSG data from ten healthy adult women is used. Spectral frequency analysis is used as an extraction feature from the EEG signal. As a result of the experiments, M3 is proved to be the most effective and suitable method for detecting sleep stages. M3 performs *91%*, M2 *82%* and M1 *70%* according to PSG analysis. In this study, the data is collected in the sleep room of a hospital. The data, though, would not show true story of the sleep since the sleep set is a disturbing and affects subject's sleep quality in a negative way.

In [40], a multimodality sleep condition inference system is proposed. The system is based on heart rate sensor, passive infrared (PIR) sensor and wireless microphone as audio sensor. The system is a distributed client-server system. In the client side, the sensed data can be collected and sent to the server. The experiments held with one person for thirteen days. This, study strictly violates the privacy of the subject by collecting audio data of the environment while we collect only the noise level of the room. Moreover, the concentration of the study is to detect whether the subject is awake or in sleep. The sleep stages or sleep quality measurements are neglected.

Sleep As Android [41] application is readily available on the *Google Play Market*. By using this application, the users are able to see their sleep behaviors. Moreover, the light sleep and deep sleep ratios can be examined. Another important feature is that this application can wake up the users in their light sleep intervals. The data collection and principles are very similar to our sleep experiments. However, there are many problems of this applications. Firstly, the application cannot infer if a user gets up and leaves the bed for some necessities like toileting through the night. The users have to manually pause the application if they do so. Secondly, the unlocked version of the application is not free. So, with the locked version, researchers are not allowed to get the raw sensor data. Another problem of this application is that the users should put their phone right under their heads to get the most accurate results. On the other hand, most of the people put their phones as far as possible while sleeping. Users would not prefer putting their phones under their pillows. With the problems mentioned above, this application is not suitable for our goals.

3. SEAMLESS LIFE MONITORING SYSTEM ARCHITECTURE

In this thesis, we propose a seamless life monitoring system which collects data from the user's 24-hour life. The proposed system consists of two main parts as the WSN and the smart phone. In the WSN, Arduino platform [1] and Xbee transceivers [2] are utilized. In the network, there are two separate microcontrollers equipped with five sensors. For the smart phone, Android operating system is used and four applications are utilized for life-logging data collection. The system can be divided into four different parts as data collection, pre-processing of collected data, post-processing and feature extraction and sleep quality estimation. With the analysis of the collected data, the sleep quality of a person is extracted and the factors affecting the sleep quality is examined. The system overview is given in Figure 3.1.

3.1. Wireless Sensor Network

The wireless sensor network (WSN) is utilized for sleep study data collection. In WSN, we have two separate microcontrollers; one for the ambient sensing, the other one for the sleep data collection. The first microcontroller is equipped with light, humidity and barometric pressure sensors. The second one is equipped with pressure mat and accelerometer sensors. The sensed data are transfered to the base station via Xbee modules connected to the Arduino microcontroller. The base station consists of an Xbee module and a computer. The connection between them is provided via a serial port. The wireless network setup can be seen in Figure 3.2.

3.1.1. Arduino Platform

Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software [1]. In recent years, the Arduino platform is preferred in a wide variety of fields by many researchers and engineers because it has

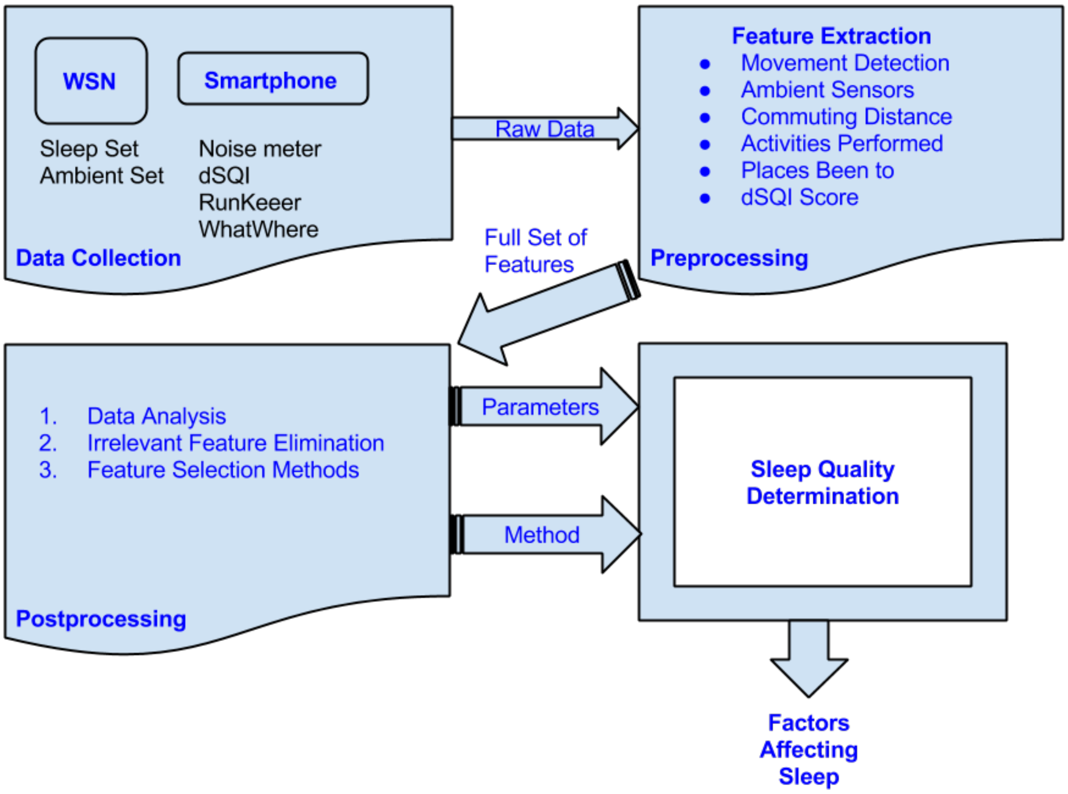


Figure 3.1. Seamless life monitoring system overview.

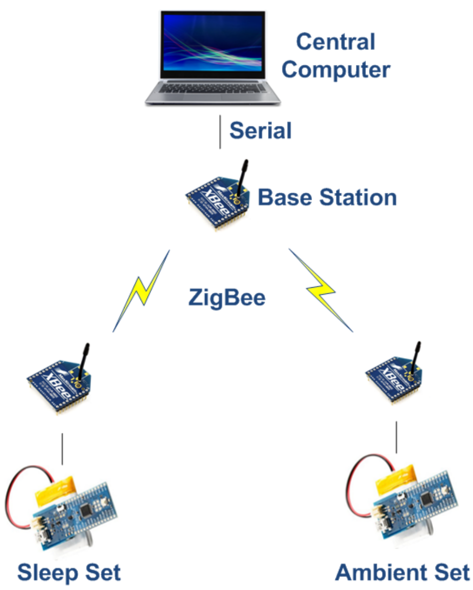


Figure 3.2. Wireless Sensor Network setup for data collection.

many GPIO pins and an easy application development environment. The board can be operational via a mini USB connection or a battery. Furthermore, Arduino has an open-source software component and uses Arduino Programming Language which is similar to C++. By using Arduino IDE, one can write a code, compile it and upload to the board via FTDI programmer interface [42]. The compiled binary is called as sketch.

In the Arduino family, we preferred Arduino Fio since it is compatible with the Xbee transceiver module. The microcontroller in the Arduino Fio is ATmega328P which operates at $3.3V$ with $8MHz$. Arduino Fio has fourteen digital GPIO pins, eight analog pins, an on-board resonator, a reset button, battery connection pin, mini USB connection for power and charge, power switch, ready-to-use socket for Xbee modules. The details can be seen in Figure 3.3. The analog pins provide ten bits resolution. Thus, the sensor readings are between 0 and 1023. It has a $32KB$ of flash memory for storing the sketches. The DC current per I/O pin is $40mA$. The DTR pin is directly connected to the sleep pin of the Xbee module. The Xbee module is the most power consuming module in our system. Its power consumption can be controlled during runtime by switching to the active mode when only it needs to transmit data. Otherwise, it stays in the sleep mode.

3.1.2. Xbee Module

For the wireless communication, Xbee Series 1 RF module [2], which uses ZigBee networking protocol for point-to-multipoint or peer-to-peer networking, is utilized. The frequency band operates around $2.4GHz$ with a $30m$ communication range. It supports serial interface data rate between $1200bps$ and $250kbps$. Since the bottleneck for the WSN sytem is power, the power consumption of the module should be minimized. The wireless communication module is the most power consuming part. In the Xbee module, the operation voltage is between $2.8V$ and $3.4V$ which is also compatible with Arduino Fio. In the transmission mode, the current drawn is $45mA$ and in the receive or idle mode, it is $50mA$. Thanks to its support for sleep modes, the power consumption of the Xbee module can be minimized and hence the lifetime of the battery can be

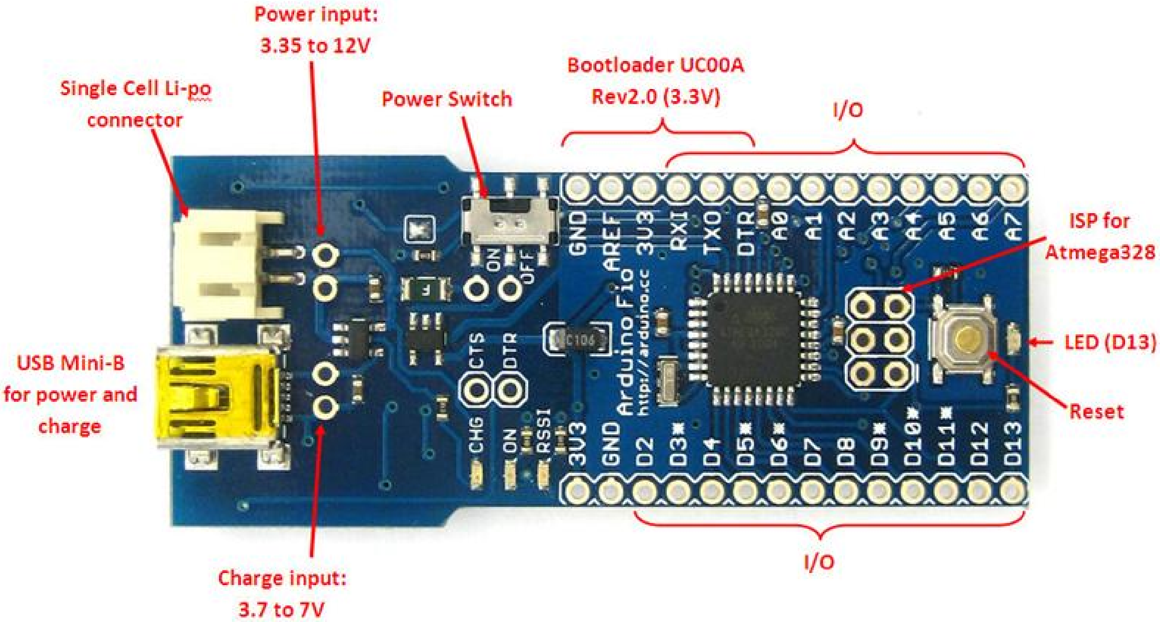


Figure 3.3. Arduino Fio [1].

extended since the current drawn goes down to 10µA in the sleep mode. The Xbee module used in our system can be seen in Figure 3.4.

Figure 3.4. Xbee module including its antenna [2].

In order to enable the Xbee modules for communication, there are some properties that should be set at the beginning. The baud rate, flow control, data bits, parity, stop bits should be adjusted to communicate with the Xbee module from the X-CTU programmer interface [43]. After the connection is provided, the modem should be configured according to the network settings. In order to create a network, each Xbee module including the base station should have the same Personal Area Network (PAN)

ID and Channel (CH). By setting a unique PAN ID, only legitimate nodes can communicate with each other. For the packet transmission, the Destination Address Low (DL) of the Xbee modules on Arduino Fio boards, should be set to the 16-bit Source Address (MY) of the base station. The list of parameter values for the setting explained above can be seen in Table 3.1. Moreover, for power saving issues, the Sleep Mode (SM) should be set to 1 (PIN HIBERNATE). The modem configurations tab of X-CTU software can be seen in Figure 3.5. Lithium polymer batteries are preferred since they are compatible with Arduino Fio boards. They work at $3.7V$ with capacities between $900mAh$ and $5000mAh$. Arduino Fio boards have a built-in charger mechanism for those batteries over USB.

Table 3.1. X-CTU Xbee modem configurations.

PC Settings	**Values**	**Modem Configurations**	**Values**
Baud Rate	57600	Sleep Mode	1 (PIN HIBERNATE)
Flow Control	NONE	Channel	C
Data Bits	8	PAN ID	3137
Parity	NONE	Destination Address High	0
Stop Bits	1	Destination Address Low	1 (Base Station)
		MY (Source Address)	Preferred

3.1.3. Sensors Employed in the System

In our system, we utilize five sensors namely pressure mat, accelerometer, photocell, humidity and barometric pressure. The pressure mat and the accelerometer sensors are used together on a single board to collect sleep data. The ambient board is composed of photocell, humidity and barometric pressure sensors. This board collects environmental data during the experiment starting from when the subject goes to bed until he/she is awake. Sensor values are digitized to discrete values between 0 and 1023. In order to achieve unobtrusiveness, we preferred ambient sensors instead of wearable sensors. For the privacy concerns, we do not collect visual and/or audio data. The pictures of the sensors are given in Figure 3.6.

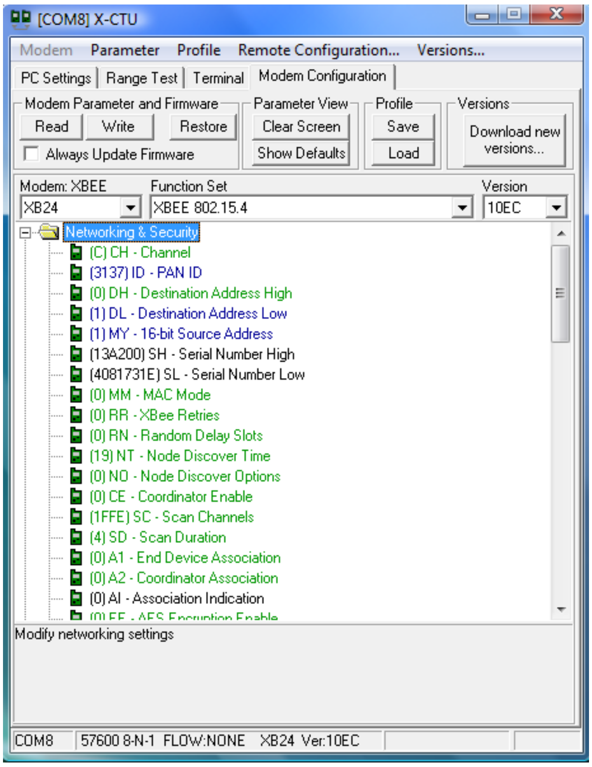


Figure 3.5. X-CTU Xbee module configuration interface.

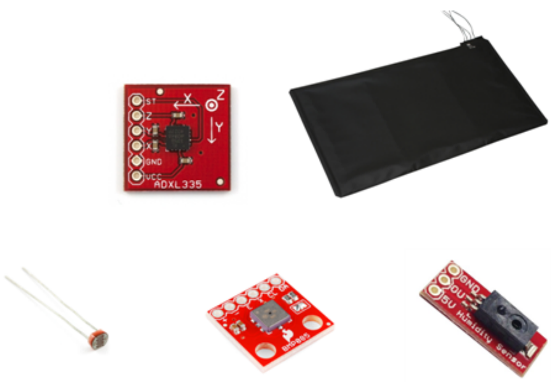


Figure 3.6. Accelerometer, Pressure Mat, Photocell, Barometric Pressure, Humidity.

3.1.3.1. Pressure Mat. The pressure mat is used to sense whether the subject is lying on the bed. In the system, the pressure mat sensor data is analyzed every thirty seconds. If the value is above the threshold, then the board sends accelerometer readings. The pressure mat is utilized to sense whether the subject leaves the bed during the night as well. If the subject leaves the bed, then the board stops sending the accelerometer data.

3.1.3.2. Accelerometer. As the accelerometer sensor we utilize ADXL-335 which can supply readings in three dimensions (x, y, z). It has ±3g sensitivity. Since converting the raw data into acceleration values (m/s^2) is time and power consuming, the raw data is sent to the base station. The conversion is handled on the computer. The accelerometer is strapped on the pressure mat in order to get the most accurate readings. The board starts sampling the accelerometer 300 times for 30 seconds if the pressure mat reading is above the threshold value.

3.1.3.3. Photocell. The photocell is used to collect the ambient light level during the sleep activity. The photocell values are sent to the base station in every one minute. Since converting raw sensor readings to lux unit values is complex and time-consuming, raw sensor readings are sent to the base stations.

3.1.3.4. Humidity. As the humidity sensor we utilize HIH-4030 [44]. The humidity values are sent to the base station every one minute interval. The raw sensor readings are converted to the relative humidity on the board.

3.1.3.5. Barometric Pressure. As the barometric pressure and the temperature sensors BMP-085 [45] is used. From this sensor, we are able to retrieve the temperature as well. The raw sensor readings are converted to the meaningful temperature ($°C$) and pressure values (atm). After the conversion, the values are sent to the base station. The sensor readings are sent to the base station with one-minute period.

3.1.4. Constraints and Limitations of WSN

One of the most important factors when designing and implementing a WSN is the lifetime of the batteries. Since the boards in the field do not have a direct connection to the electricity; sensor and communication device selection becomes critical in terms of their energy consumption levels. The selection of the battery in terms of its energy storage capability plays an important role for the lifetime of the system. The problems occurred during the field tests can be classified into three parts as the battery lifetime, wiring problems and sensitivity of the pressure mat.

3.1.4.1. Battery Lifetime. Since the boards in the network get its energy only from the battery, the system lifetime is limited to the battery lifetime. In this regard, selecting a low power consuming board is the key factor for increasing the system survivability. For these concerns, we chose Arduino Fio platforms, Xbee modules and lithium polymer batteries. The system lifetime can be significantly increased by using batteries in parallel. Moreover, Arduino Fio boards allow on the fly recharging using its USB interface. Thus, we do not have to remove the empty battery for recharging. There are three main factors that affect the lifetime of the batteries.

First, we can decrease the sampling rate of the sensor readings to minimize the power consumption. However, we have to consider the possible data loss in lower sampling rates. In the lights of examination in [4], we set the sampling rate to $10Hz$ for all the sensors to balance the trade-off between energy consumption and data loss.

Second, we are able to decrease the energy consumption of the wireless communication module by changing the sleep mode of the module in the runtime. The Xbee module wakes up only when it needs to send new sensor reading. Otherwise it stays in sleep mode to minimize unnecessary energy consumption. This is an important optimization since the wireless communication is the most power consuming component. In the preliminary tests, we observed that putting the Xbee module in sleep mode significantly increases the battery lifetime. The ambient sensor set can last twice longer than the sleep sensor set since it can use the sleep mode feature more often. In the current system, the ambient sensor set samples the data, sends to the base station and puts the Xbee module into sleep mode for sixty seconds. This approach extends the lifetime of the batteries significantly. For the bed setup, an event-based data collection system is adapted. The board checks the pressure mat sensor readings every thirty seconds. If the value is above the predefined threshold value, this means that the subject is on the bed and we start sampling the accelerometer data for thirty seconds. After sending these sensor readings to the base station, the board checks the pressure mat value again. If the reading is below the threshold, the board puts the Xbee module into sleep for thirty seconds. The only drawback of the sleeping mode is that the re-association with the network can take as long as $300ms$, though the wake up time for the Xbee modules is $13.2ms$ [4]. In order to eliminate the data loss possibility during the wake-up of the Xbee module, the board waits for $300ms$ to make sure that the Xbee module is ready to send data.

Third, the duty cycle of the board is another important factor that determines the battery lifetime. In [4], the battery lifetime of the boards are examined. Their experiments show that an Arduino Fio board can operate 24 hours with $1800mAh$ battery while the sampling rate is $10Hz$, the duty cycle is 10% and the Xbee module is in idle mode. On the other hand, the same setup can operate 145 hours when the sampling rate is $10Hz$, the duty cycle is 10% and the Xbee module is in the sleep mode. When considering that our sleep set runs 6-7 hours a day, the bed sensor can last at least a week and the ambient sensor set lasts even longer. Therefore, we set the duty cycle of the board to 10%.

3.1.4.2. Wiring Problems. Since Arduino boards are shipped without connectors which are used to connect the sensors to the board.The post wiring of these connectors becomes an important issue for the system reliability. While the ambient sensor set is stationary and out of subject interaction, the subjects lie on the pressure mat. Thus, the connections between the sensors and the board could be lost during the night. Our field tests showed that the connectors can easily be broken while the subject moves in his/her sleep. So, soldering should be strong and flexible enough to connect the sensors and the board. For the benefit of the system performance, the soldering of the sensors to the board should be handled cautiously or by professionals if possible.

3.1.4.3. Pressure Mat Problems. The most important part of the current system is the collection of sleep data. Thus, the positioning of the pressure mat and the accelerometer plays an important role. Since we attach the accelerometer on the pressure mat, the pressure mat's placement becomes an important parameter for the success of the system. Our field tests have showed that depending on the subject's mobility during his/her sleep, the position of the pressure mat changes with time possibly resulting damaged cables or poor accelerometer readings. Moreover, the pressure mat should be placed in a way that the accelerometer is right under the subject's shoulder. In other placements of the accelerometer, we observed poor sensor readings so that the sleep stage detection could not be performed.

3.1.5. Sleep Setup Design

We utilize two Arduino Fio boards to setup a wireless sensor network testbed. The first one collects the accelerometer data triggered by the pressure mat readings from the bed during sleep. The setup of this sensor can be observed in Figure 3.7. The shown sleep board is ready-to-use in the field. The pin connection details are given in Figure 3.8.

To detect whether a subject is lying on the bed or not, we propose a threshold-based mechanism whose main procedure can be seen in Figure 3.9. To reduce the power consumption, this mechanism runs in every thirty seconds interval by going from sleep mode to the active mode. Even the board is in active mode, the Xbee module can be put into sleep mode. When the pressure mat detects a person is on the bed, then three dimensional accelerometer values are sent to the base station. The accelerometer is sampled at $300Hz$ to get a detailed movement spectrum.

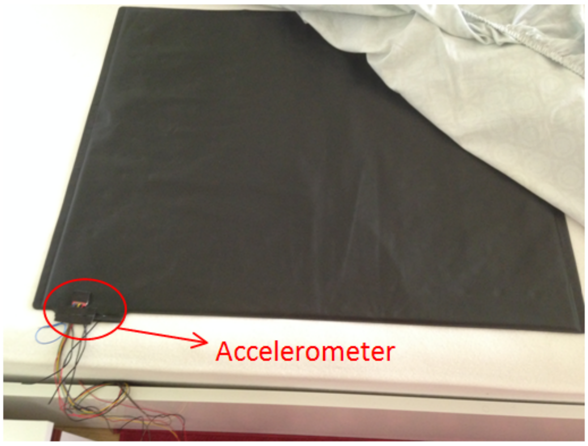


Figure 3.7. Sleep board overview on the bed.

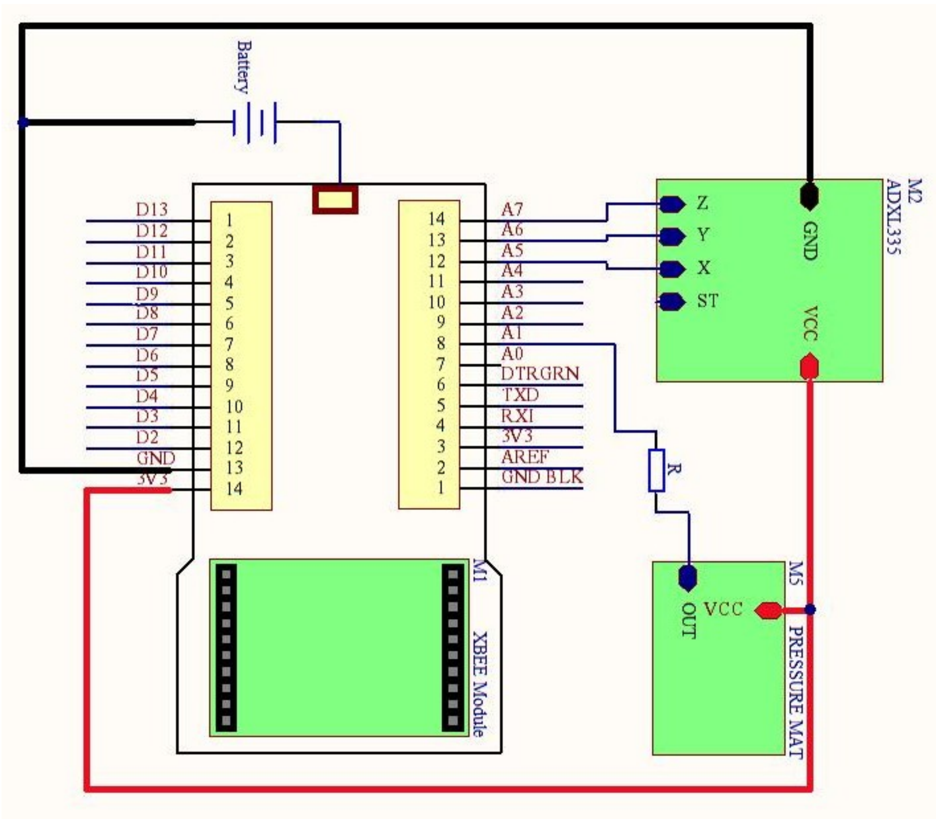


Figure 3.8. Sleep board schematic with the sensors connected.

loop
  $pressureMatValue \leftarrow analogRead(A0)$
  if $pressureMatValue \geq threshold$ **then**
    $xbeewake()$
    $transmitValue(pressureMatValue)$
    for $i := 0, \ldots, 300$ **do**
      $x_acc \leftarrow analogRead(A5)$
      $y_acc \leftarrow analogRead(A6)$
      $z_acc \leftarrow analogRead(A7)$
      $transmitValue(x_acc)$
      $transmitValue(y_acc)$
      $transmitValue(z_acc)$
    end for
  else
    $xbeesleep()$
    $delay(30sec)$
  end if
end loop

Figure 3.9. Sleep board data sense and transmit procedure.

The second board is utilized for collecting ambient values during the sleep via photocell, humidity and barometric pressure sensors. This board is placed in the same room near to the subject's bed in order to collect ambient data closer to the sleeping person. The setup of this board can be seen in Figure 3.11. The pin connections of the board and the configurations of the sensors can be seen in the schematic given in Figure 3.10.

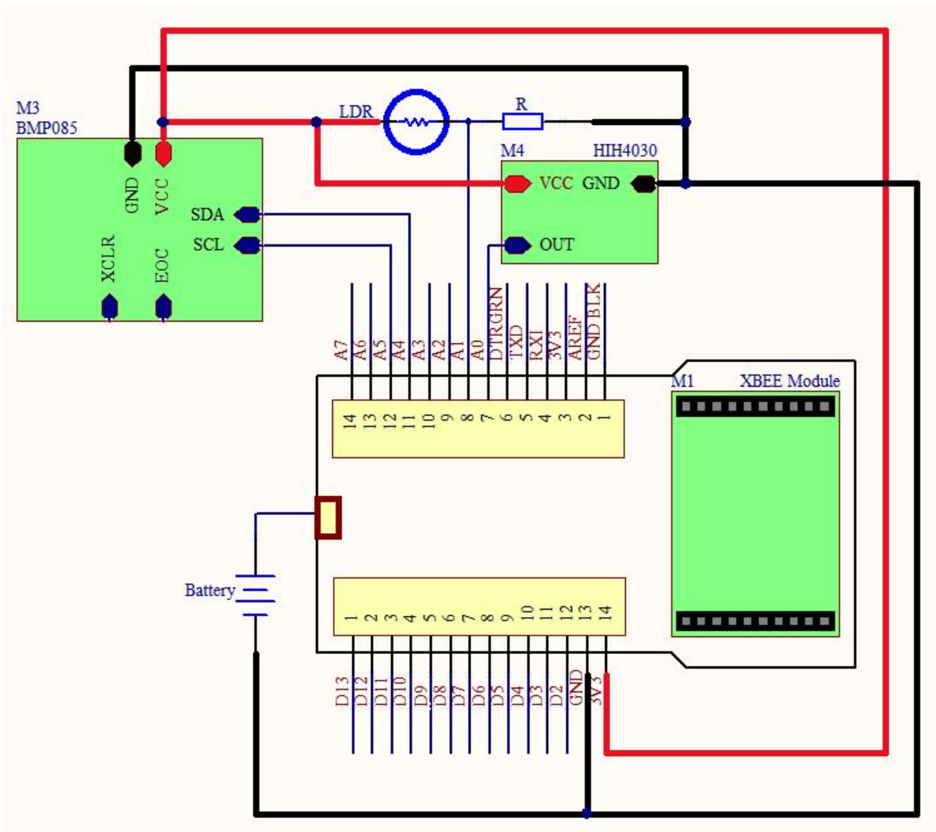


Figure 3.10. Ambient board schematic with the sensors connected.

The working principle of the ambient board is in a duty cycled manner. When the board is switched on, the loop starts running. The sensor values are sampled and transmitted to the base station for further analysis. After the transmission, the Xbee module is put into sleep and the main process sleeps for one minute. When the Xbee module is awakened and the same sampling and transmission procedures are repeated. The main flow procedure of this operation can be seen in Figure 3.13.

The base station of the WSN is serially connected to the central computer via

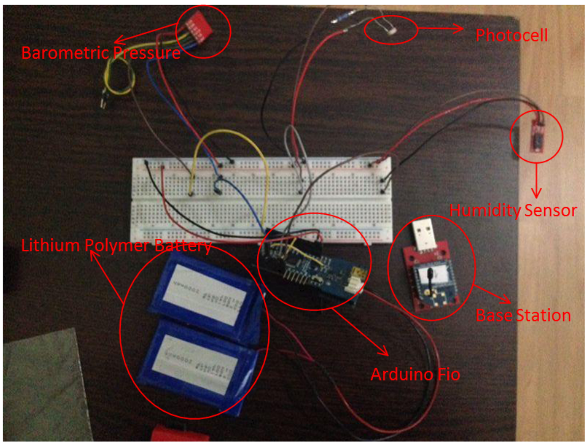


Figure 3.11. Ambient board overview near the bed.

the Xbee Explorer USB interface. On the computer side, an application collects and logs the incoming sensor readings from the base station for post-processing. The base station which is connected to the computer and receiving data can be seen in Figure 3.12. The red led means the base station is working and the green led means the base station is receiving data from other boards.

Figure 3.12. Base station connected to central computer.

3.2. Smart Phone Applications

Smart phones equipped with many useful sensors have simplified research and field tests in many application areas. In the life tracking field, mostly preferred sensors

loop
  $rawHumidity \leftarrow analogRead(A0)$
  $rawLight \leftarrow analogRead(A1)$
  $bmpSDA \leftarrow analogRead(A3)$
  $bmpSSCL \leftarrow analogRead(A4)$
  $relativeHumidity \leftarrow calculateHumidity()$
  $temperature \leftarrow calculateTemperature(bmpSDA, bmpSSCL)$
  $pressure \leftarrow calculatePressure(bmpSDA, bmpSSCL)$
  $xbeewake()$
  $transmitValue(relativeHumidity)$
  $transmitValue(rawLight)$
  $transmitValue(temperature)$
  $transmitValue(pressure)$
  $xbeesleep()$
  $delay(1min)$
end loop

Figure 3.13. Ambient board data sense and transmit procedure.

are the accelerometer and GPS sensors. The reasons of preferring smart phones over WSNs are that people carry their smart phones while this is not the case for the existing WSN. Moreover, smart phones offer higher computational power and storage capabilities. Hence, we integrated four smart phone applications into our system. We developed the two of them and the other two are downloaded from the *Google Play Market*.

3.2.1. Daily Sleep Quality Index (dSQI)

This application contains questions inspired from PSQI questionnaire. While the PSQI score is calculated monthly, the dSQI score is calculated in a daily basis. The subjects are asked to answer the questions every morning after they wake up. The questions are about the details of sleep for the last night. The importance of this application is to log the subjective assessment. The purpose of this application is to collect information that cannot be inferred by the sensory sleep set. The questions are given in Figure 3.14. With the answers given, the subject's dSQI score is calculated. The daily score indicates the feelings of the subject about his/her sleep quality. The scoring mechanism is explained in detail in Appendix A. The score ranges between 0 and 20 in which 0 means the best quality of sleep while 20 means the worst quality. The user interface of the application can be seen in Figure 3.15.

3.2.2. WhatWhere

We developed an acitivity-logger application called *WhatWhere* to collect information about the users' places and activities. Once the application is started, it sends notifications periodically to the user to enter his/her current place and activity. The probable activity and place names are taken from a configuration file. Additionally, the time interval for notification can be read from the configuration file as well. The details of this file can be seen in Appendix B. If the file is not found from the file system, then the application runs with the default values. The default time interval for notifications is thirty minutes. The user interface of this application is given in Figure 3.16. The answers are recorded with a time stamp. Hence, we can infer what the subject is doing and where he/she is at a specific time interval. Also, relation between *WhatWhere* and *Run Keeper* applications can be built.

$Q1)$ What time did you go to bed last night?

$Q2)$ What time did you wake up this morning?

$Q3)$ How long(in minutes) did it take you to fall asleep last night?

$Q4)$ How many hours of actual sleep did you get last night?(This may be different than the number of hours you spent in bed.)

$Q5)$ How many times did you wake up last night?

$Q6)$ Last night, did you have trouble sleeping because you ...

- $Q6a)$ cannot get to sleep within 30 minutes?(Y/N)
- $Q6b)$ wake up in the middle of the night or early morning?(Y/N)
- $Q6c)$ have to get up to use the bathroom?(Y/N)
- $Q6d)$ cannot breathe comfortably?(Y/N)
- $Q6e)$ cough or snore loudly?(Y/N)
- $Q6f)$ feel too cold?(Y/N)
- $Q6g)$ feel too hot?(Y/N)
- $Q6h)$ had bad dreams?(Y/N)
- $Q6i)$ have pain?(Y/N)

$Q7)$ Last night, did you take medicine to help you sleep(prescribed or "over the counter")?(Y/N)

$Q8)$ Yesterday, did you have trouble staying awake while driving, eating meals, or engaging in social activity?(Y/N)

$Q9)$ How would you rate your sleep quality last night?

− Very good − Fairly good

− Fairly bad − Very bad

$Q10)$ Yesterday, how much of a problem has it been for you to keep up enough enthusiasm to get things done?

−No problem at all −Only a very slight problem

−Somewhat of a problem −A very big problem

Figure 3.14. dSQI questions.

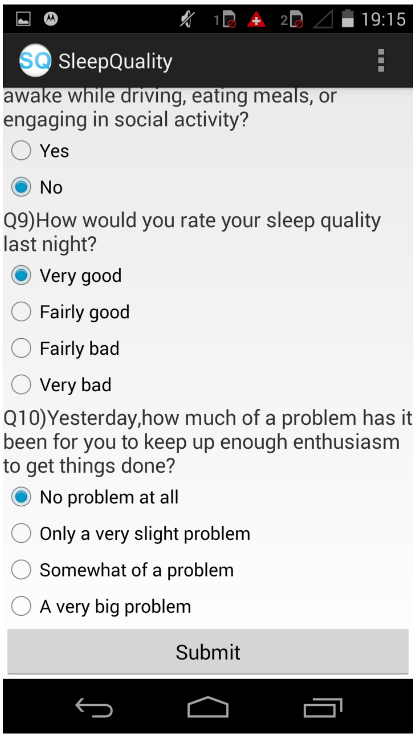
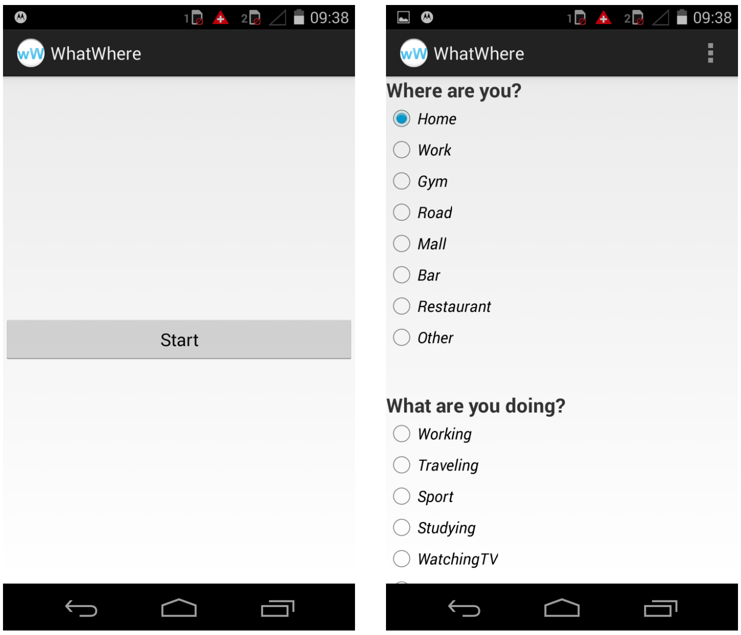


Figure 3.15. User interface of daily Sleep Quality Index (dSQI).



(a) Starting WhatWhere application.

(b) Multiple choices about place and activity.

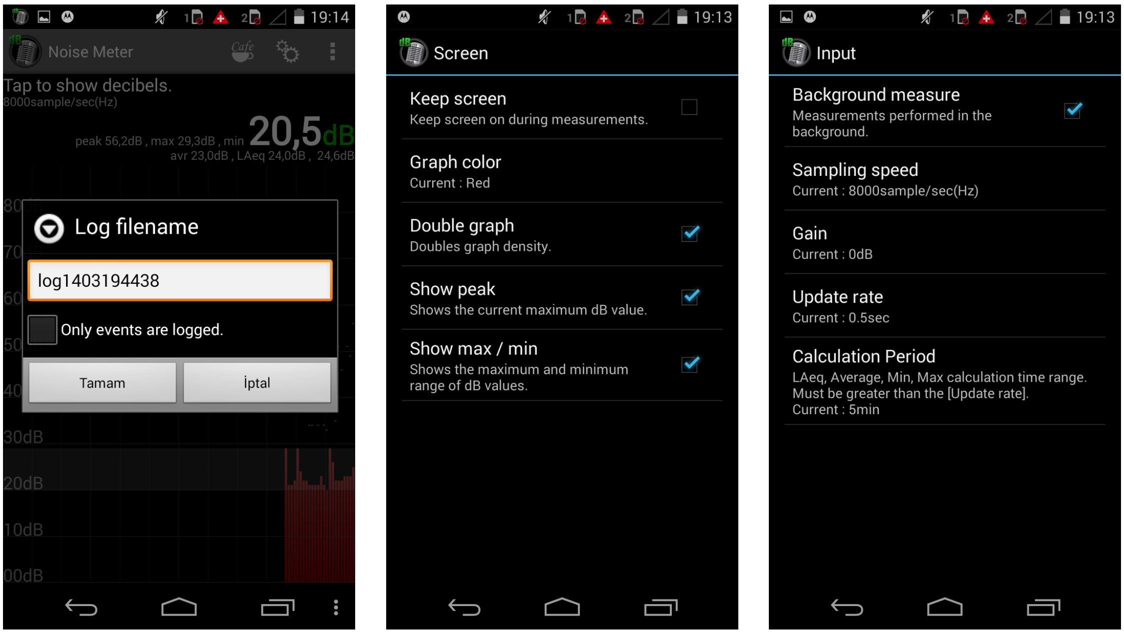
Figure 3.16. *WhatWhere* screens.

3.2.3. Noise Meter

Noise Meter [46] application can be downloaded from the *Google Play Market*. It is preferred for logging the ambient noise level during sleep since the noise higher than $40dB$ which is equivalent to the noise level of a library affects the quality of sleep in a bad way. Furthermore, the noise level gives inference about whether the sleeper snores. The best part of this application is that the log files can be downloaded from the phone's file system for further analysis. Before starting to use the application, there are some configuration settings to be done. First we have to uncheck "Keep Screen" option. Otherwise, the screen is kept on during the measurements and the battery consumption increases enormously. Moreover, "Background measure" option should be checked for the application to keep measurements after the screen is turned off. Additionally, the user is able to change some other configurations. For the current usage; sampling speed is $8000Hz$, gain is $0dB$, update rate is 0.5 seconds and calculation period is five minutes. The images and the details about the application can be seen in Figure 3.17. Additionally, the general noise levels for different real life situations can be examined in Table 3.2.

Table 3.2. Known noise levels

Noise Level	**Place or Situation**
$0dB$	Threshold of hearing
$10dB$	Soundproof room
$20dB$	Very calm room
$30dB$	Whispering words
$40dB$	Library
$50dB$	Interior noise in car
$60dB$	Large stores, talking
$70dB$	Noisy office
$80dB$	Traffic on a busy roadway
$90dB$	Train sound

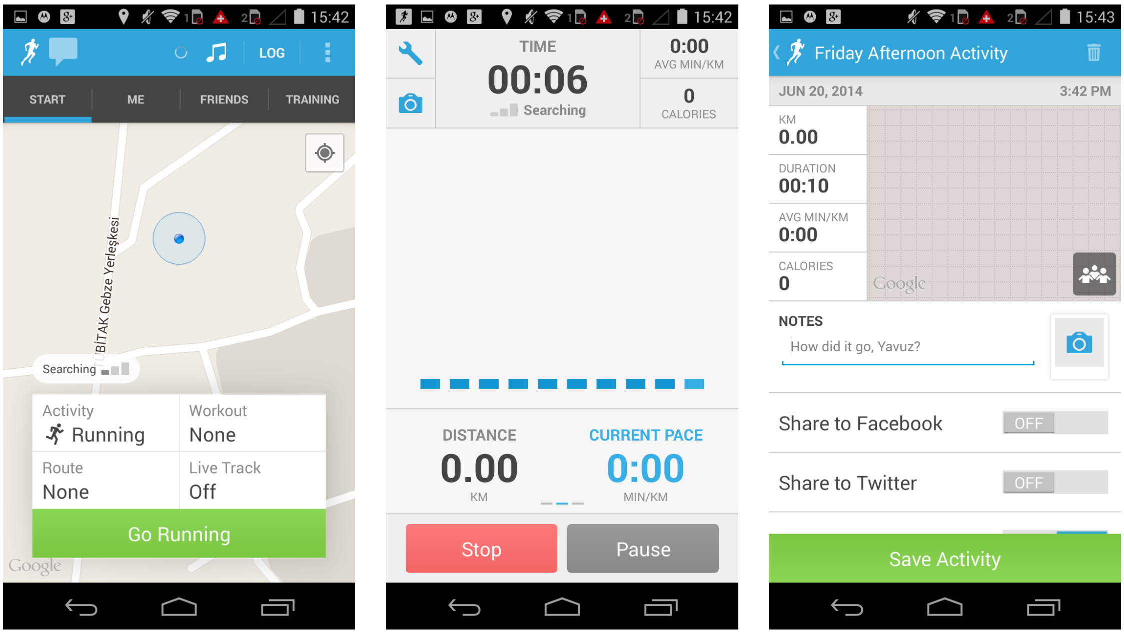


(a) Starting data collection. (b) Screen settings. (c) Input settings.

Figure 3.17. *NoiseMeter* screens.

3.2.4. RunKeeper

The *Run Keeper* [47] application can be downloaded from the *Google Play Market*. This application is selected to track the subject when he/she is outside. It is important to state that we are not seeking for activity details. The subject starts the *Run Keeper* in the morning with the activity label “Other”. If he/she becomes stationary like working, the application can be paused. At end of the day, the subject stops the application and saves the current tracking data. This application gives us the total commuting distance, the places stopped by, and the time passed outside. Moreover, the log files can be downloaded from the server site as GPX (GPS Exchange Format) and/or Google Earth files for further analysis. The screen details of the *Run Keeper* application is given in Figure 3.18.



(a) Starting GPS tracking. (b) Pausing/Stopping GPS tracking. (c) Saving GPS tracking.

Figure 3.18. *RunKeeper* screens.

4. PRELIMINARY EXPERIMENTS

The evolution of the thesis can be divided into three main phases. In the first phase, we focused on the activity tracking inside the home. For this, we utilized the WSN composed of many sensors. With this study, we were aiming to recognize the activities and their durations in a day. In the second phase, we focused on the detection of the sleep behavior and quality of a person. To achieve this, we made some modifications in our WSN. Moreover, we added smart phone applications into the system to get more information about the sleep-related issues. In the last phase, we added two more smart phone applications to get more information about the outside life of a person.

4.1. Life Tracking Studies in Home

In this phase, we used a WSN composed of many sensors. The sensors included in this study are contact sensor, force sensor, sonar distance sensor, pressure mat and ez430-Chronos [48] watch. The sensor values are sampled and transmitted to the base station via Arduino board and Xbee modules.

With the usage of such a system, we were planning to answer the following questions:

- How many hours a person stays outside?
- How many hours a person sleeps?
- How much time a person spends on the couch?
- What is the behavior of a person inside the home?
- What is the sleep behavior of a person?

On the other hand, the system cannot infer many information. For instance, the activity cannot be inferred while a person is sitting on the couch. He/she might read a book, watch TV or study his/her lessons. It is very important to decide whether we need the couch sitting data for life tracking studies. To recognize the activity performed in a certain place like a couch, we have to equip the couch with many sensors as mentioned in [4]. However, using such a system is highly user-dependent. Furthermore, increasing the number of sensors in the system increases the complexity and decreases the robustness. After the experimental studies, we have figured out that we needed to evolve our system to a more user-friendly and low complex one.

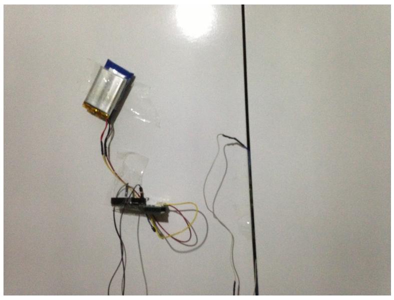
4.1.1. System Deployed

The experiments were held in a real house. The contact sensor is placed on the side of wardrobe door. The sonar sensor is placed on the inner side of the bedroom door. Te force sensor is placed under the couch in the room. The pressure mat is placed under the bed. Lastly, Chronos watch is worn on the wrist of the subject. The pictures of the sensors and their placement is given in Figure 4.1 In the system, there are two different WSNs since they communicate over different protocols. One is composed of Xbee modules and ez430-Chronos watch communicates with its own USB dongle. The default programs brought with the development kit of Chronos watch are able to show the data coming from the watch but not able to log the data. For this reason, a programming language called Processing Programming Language [49] is used. It is a Java-based programming language, development environment, and an online community. Simply, with the written program, called as sketch, we intervene into the serial communication port of the connected USB dongle. With this feature, we could use the accelerometer data for further studies on sleep.

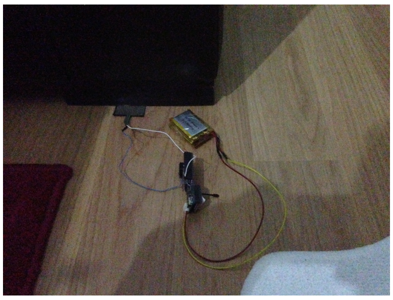
4.1.2. Lessons Learned from the Deployments

In order to give examples about the experiment results, we added Chronos watch data of two separate sleeps that can be inspected in Figure 4.2. The graph in the figure shows raw accelerometer data coming from the Chronos watch. When the Chronos watch data graph is examined, it can be easily seen that the movement of the subject in the second day during sleep is higher than the first day. It can be said that the subject gets a better sleep the first day. The reason of this change according to the

(a) Sonar sensor placed on the door



(b) Contact sensor placed on the wardrobe door.



(c) Force sensor placed under the couch.

(d) ez430-Chronos watch

Figure 4.1. Sensors deployed in the room.

subject's briefing is that the subject went to the gym in the first day. As a result, we can infer that physically-tired healthy people get a sleep with less movement.

On the other hand, the Chronos watch has many problems. During the experiments, we encountered loss of connection during sleep. The probable cause is the loss of line of sight. The subject may have put his arm under the pillow. Furthermore, while the watch is sending data more frequently at the early stages of sleep, the density of data sent to the computer decreases with time. Additionally, the watch disturbs the subjects in their sleep since it is a wearable sensor and violates the unobtrusiveness of the system. Due to problems mentioned above, we decided to change the sensor used for collecting sleep information. In the next phase, we switched to a single accelerometer sensor attached on the side of the pressure mat.

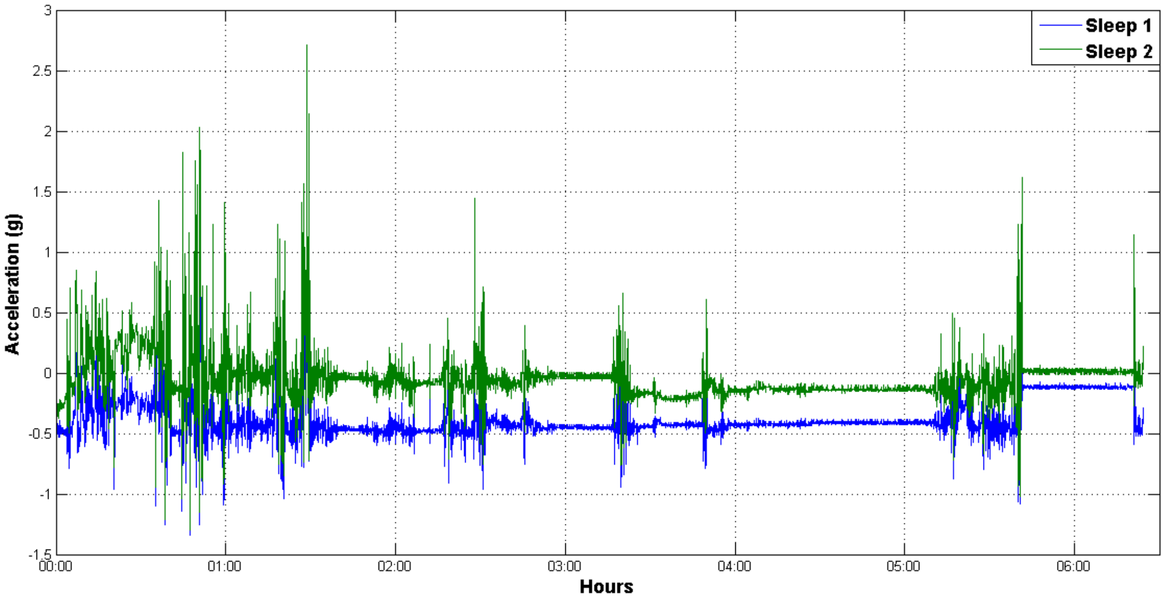


Figure 4.2. Chronos watch accelerometer data of two separate sleeps.

For the sensor readings, we added a diagram for a 24-hour collected data that can be examined in Figure 4.3. In the diagram, the pressure sensor gives an idea about the subject's sleep behavior. The force sensor data shows the time and duration of the subject's sitting on the couch. The sonar sensor gives information about the subject's entrance to the room and exit from the room. The distance sensor gives information about the times the subject changes his clothes. By inspecting the graph, we can infer the mobility of the subject inside the home. The duration spent on the couch, his sleep/wake up time and the duration, how many times he changes clothes can be followed with the collected sensor data. The example in Figure 4.3 starts data collection around 21:00. The subject sits on the couch for two and a half hours. Around 23:30, he goes to the bed. During this period, the subjects enters and leaves the room four times and opens his wardrobe four times. He sleeps for almost nine hours. When the subject wakes up, he leaves the room probably for toileting or bath for a short time and when coming back, he changes his clothes. During the day, he leaves the room two times. The last times is for bath or toileting since after this event, he opens the wardrobe door and goes to the bed.

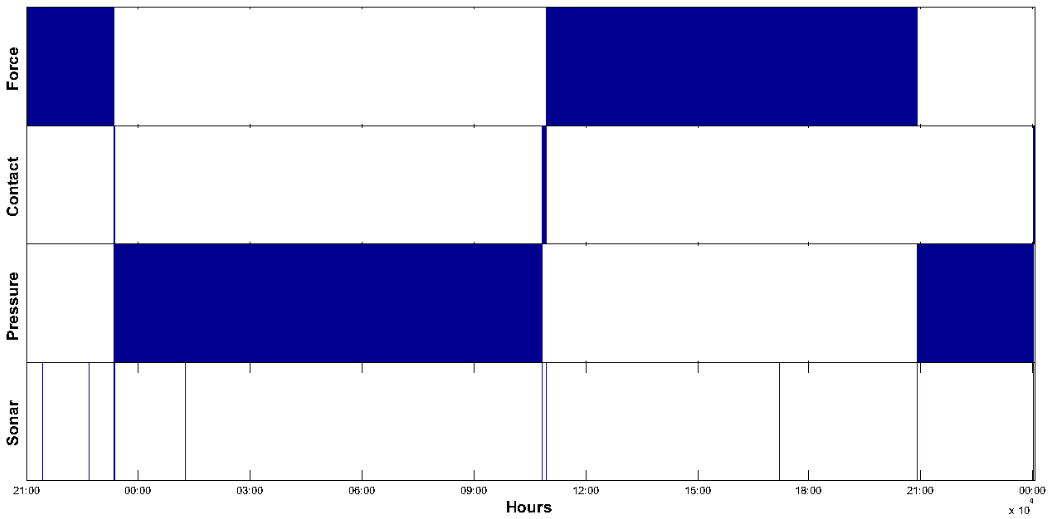


Figure 4.3. Force, sonar, pressure and contact sensor readings of a person for proving the capability of activity monitoring.

However, using such a WSN is hard to maintain because there are many sensors placed in very delicate positions. Using only the sensor data, we are not able to recognize the activities inside the home. We can get information about the place of the subject and the time passed on a particular place. As a result, we have decided to remove low priority sensors. Instead, we developed an Android application that logs the place and activity of a person periodically.

4.2. Sleep Quality Studies

Secondly, we focused on calculating the sleep quality of a person. For this reason, we added an accelerometer to the system. It is placed on the edge of the pressure mat aligned to the sleeper's shoulder. The field tests were performed by two subjects for one month duration. The sleep quality measurement is performed to achieve the sleeper's deep sleep and light sleep periods and durations. Due to the absence of Polysomnography (PSG) devices, we are not able to detect sleep stages in terms of REM and NREM. Furthermore, we developed an Android application for keeping subject's assessments about every sleep. This questionnaire scores the subject's daily-SQI inspired from PSQI [17]. In order to record the nocturnal noise level of the subject's room, we employed the *NoiseMeter* application on smart phone.

4.2.1. Sleep Quality Estimation

For the sleep quality estimation, the method in [50] is applied. In this method firstly, the collected raw accelerometer data is parsed and derivate operation which takes the difference of consecutive sensor readings from x, y, z dimensions, is applied. Secondly, the acceleration magnitude for each reading is calculated from the derivative of x, y, z data. Thirdly, from the acceleration magnitude, the movement count for each minute is calculated. In our case, individual magnitudes represent $100ms$ movement. So, in order to calculate one minute movement count, the data is divided into 600 sub arrays which represent one minute. If the magnitude value is greater than acc^T_{mov} threshold, then movement count is increased by one. Finally, the smoothening function is applied to the movement count array in order to eliminate noisy data. Smoothening function is handled with previous and latter two readings together with the current one. The method procedure is explained in Figure 4.4.

Since the analog-to-digital converter (ADC) in Arduino has 10-bit resolution, the raw sensor readings can be between 0 and 1023. Also, since the accelerometer sensitivity is $\pm 0.03 m/s^2$, the sensor values tend to change 4-5 out of 1023 due to the sleeper's respiration. The acceleration magnitude is calculated by taking the square root of the sum of squares of x, y, z values. The threshold value acc^T_{mov} can be calibrated according to the accelerometer's sensitivity, the sleeper's weight and the toughness of the bed. In our experiments, the threshold value acc^T_{mov} is $\sqrt{48}$ for the normal weighted subjects because the sensor readings for every dimension change 4 out of 1023 when the sleeper does not move. If the derivate magnitude is higher than the threshold then it is considered as movement and movement count is increased by one. The sleep analysis is handled according to the one-minute epochs. The maximum movement count mov_{max} is 600 and minimum value is 0. We set mov^T_{min} threshold value to 5 to decide whether there is a movement in a certain 1-minute period because as stated in [50], normal people tend to make five little reflexive movements during deep sleep cycles. This threshold should be calibrated according to the sleeper's condition. For example, if the

Taking the derivative of raw accelerometer data

loop
  for $i := 0, \ldots, n$ **do**
    $x_i^{drv} \leftarrow x_{i+1} - x_i$
  end for
end loop

Taking the magnitude of derivate x, y, z data

loop
  for $i := 0, \ldots, n$ **do**
    $m_i \leftarrow x_i^{drv} + y_i^{drv} + z_i^{drv}$
  end for
end loop

Calculate one minute movement

loop
  for $j := 0 \ldots, n/mov_{max}$ **do**
    for $i := j \times mov_{max}, \ldots, j \times mov_{max} + mov_{max}$ **do**
      if $m_i > acc_{mov}^{T}$ **then**
        $movementCount \leftarrow movementCount + 1$
      end if
    end for
    $movementArry_j \leftarrow movementCount$
  end for
end loop

Smoothen one minute movement counts

loop
  for $i := 0 \ldots, n/mov_{max}$ **do**
    $m_i^s \leftarrow (m_{i-2} \times 0,04) + (m_{i-1} \times 0,02) + m_i + (m_{i+1} \times 0,02) + (m_{i+2} \times 0,04)$
  end for
end loop

Figure 4.4. Feature extraction from the accelerometer data procedure.

Finding the SOL time from smoothened movement data

loop
  $solTime \leftarrow 0$
  for $i := 1, \ldots, length(m_i^s) - 10$ **do**
    $counter \leftarrow 0$
    for $j := 0, \ldots, 9$ **do**
      if $m_{i+j}^s \geq mov_{min}^T$ **then**
        $counter \leftarrow counter + 1$
      end if
    end for
    if $counter \leq 1$ **then**
      $solTime \leftarrow i$
      break;
    end if
  end for
end loop

Figure 4.5. SOL-10 method procedure.

sleeper has Parkinson's disease, the threshold value should be increased. At this point, we are able to extract deep sleep and light sleep cycles from the accelerometer data.

The sleep onset latency (SOL) is the time passed between a person goes to the bed and falls into sleep. The SOL time can be found by checking the movement count of every minute. For the SOL time calculation, we used SOL-10 method from [50]. In this method, we check consecutive ten-minute movement data to decide whether the person is in sleep. If there is no movement detected in a ten-minute period, the first minute of that period is set as the time the person falls into sleep. Hence, this minute represents the SOL time of that night. The procedure of SOL-10 method can be seen in Figure 4.5. The total sleep duration is calculated by subtracting the time passed in the bed from the SOL time. Moreover, if there is no movement detected during 30 minutes or more, then this period is considered as the deep sleep cycle. When this procedure is applied to the full night data, we can come up with the number of deep seep cycles. Also, by summing all the deep sleep cycle durations, we can calculate the total deep sleep duration. Finally, the difference between the total sleep duration and the total deep sleep duration gives us the total light sleep duration. In Figure 4.6, the result of this algorithm applied to one night sleep activity is given.

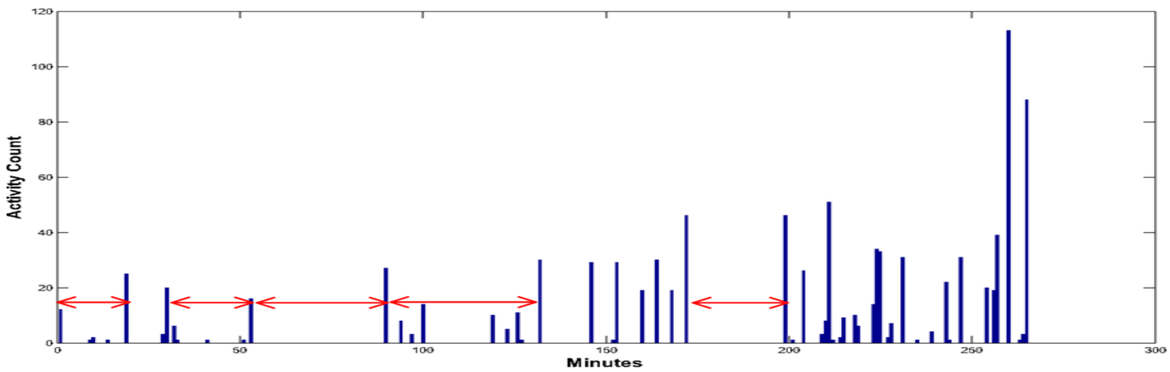


Figure 4.6. Sleep activity in one night (two sided arrows indicate deep sleep).

4.2.2. System Deployment and Data Collection

For the field tests, the sleep set is deployed in two different houses. The subjects collected data about one month. The ages of the subjects are 23 and 25. The sleep set is composed of the accelerometer attached to the pressure mat sensor, a central computer with a base station and an Android smart phone. The sleep set is utilized to inspect the subject's nocturnal sleep activities. By using this set, we are able to achieve the subject's deep and light sleep durations, the deep sleep ratio and the sleep onset latency (SOL). Especially, the deep sleep ratio is considered to be an indicator of the sleep quality. The more detailed analysis is given in Chapter 5. The data collection process provides 60-night sleep data. The sleep quality of each night is classified as *Good*, *Moderate* and *Bad* according to the dSQI scoring A. At the end of the experiment, there are no *Bad* quality sleep days. The reason could be that the subjects are healthy young men and do not suffer from any sleep-related diseases. Subject 1 has 21 *Good* quality sleeps and 9 *Moderate* quality sleeps. On the other hand, Subject 2 has 24 *Good* quality sleeps and 6 *Moderate* quality sleeps.

4.2.3. Inferences from the Experiments

The dSQI scores are calculated in a daily manner. The scoring can be found in Appendix A. The dSQI score between 0 and 5 is labeled as *Good* sleep quality. The score between 6 and 10 is labeled as *Moderate* sleep quality. The scores above 10 is labeled as *Bad* sleep. In Figure 4.7, the sleep parameters of 60 nights collected during the experiment from the subjects can be seen. The dSQI scores represent the perceived sleep quality for the sleepers. The dSQI score is directly related to the number of deep sleep cycles and the deep sleep ratio. When the number of deep sleep cycles and/or the deep sleep ratio are decreased, the dSQI score of that night increases. For example, in days 3, 27, 36 and 45, the deep sleep ratio and the number of deep sleep cycles are decreased. As a result, the dSQI scores of those nights are increased resulting in worse quality sleep. Additionally, the sleep onset latency (SOL) has an important effect on the dSQI score. For example, in days 3, 16, 20 and 27, increase in the SOL times results in increase in the dSQI scores. Hence, the SOL time has a negative influence on the sleep quality. A longer SOL time results in a worse sleep quality. The experiment shows us that the SOL time, the deep sleep ratio and the number of deep sleep cycles are closely related to the sleep quality of a person.

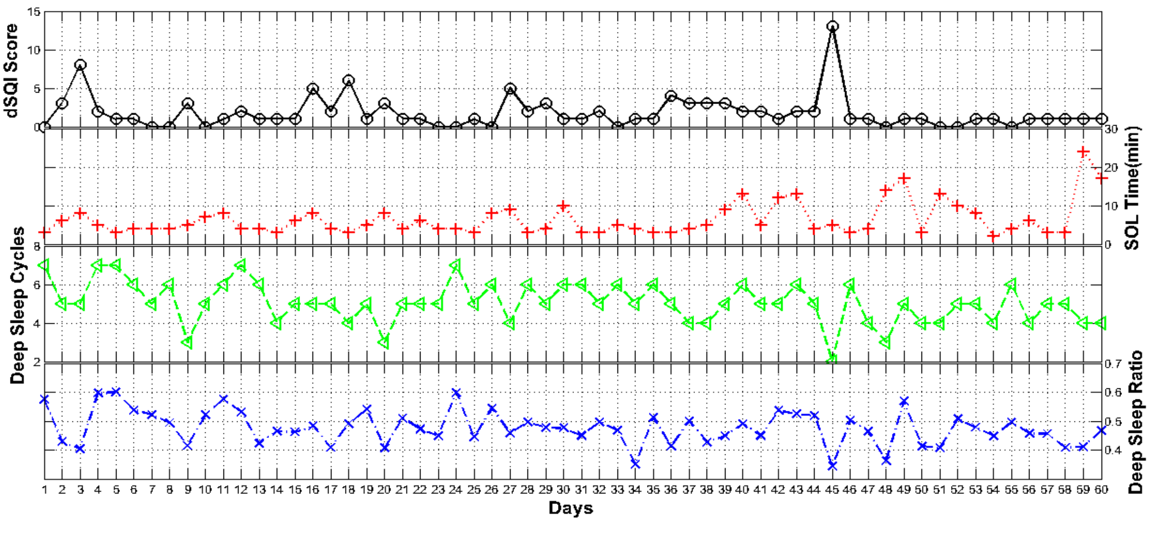


Figure 4.7. The sleep parameters and the dSQI scores of subjects during experiment.

5. SLEEP QUALITY DETERMINATION

In this chapter, the system setup explained in Chapter 3 is used to determine the factors affecting the sleep quality of a person. Nine subjects whose ages are between 25 and 29 are involved for the field experiments. At the end of the experiments, we report the results which explain how each feature affects the quality of sleep. The features are sleep parameters, noise level, temperature, humidity, atmospheric pressure, locations, the activities performed during the day, the total commuting distance during a day. The main steps of the experiment shown in Figure 5.1 will be explained in detail below.

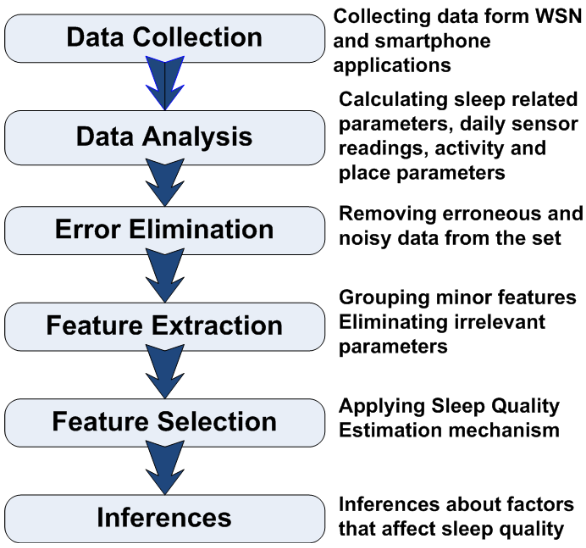


Figure 5.1. The main steps of the Sleep Quality Estimation mechanism.

5.1. Data Collection

For the data collection phase, the system which consists of a WSN and a smart phone is utilized. In the WSN part, sleep data and ambient factors are periodically sampled and transmitted to the base station when the subjects are in sleep. Besides, *RunKeeper*, *dSQI*, *WhatWhere* and *Noise Meter* applications run on the smart phone to collect the data when the subjects are outside as well. The sampling periods of the sensors and the applications are given in Table 5.1.

Table 5.1. Sampling periods of sensors.

Data Type	Sensor Type	Location	Sampling Period
Light Level	Photocell	Near bed	60 seconds
Relative Humidity	HIH-4030	Near bed	60 seconds
Temperature	BMP085	Near bed	60 seconds
Atmospheric Pressure	BMP085	Near bed	60 seconds
Pressure Mat	Arun PM3	Under bed	1 second
Accelerometer	ADXL335	On pressure mat	100 milliseconds
Noise Level	Noise Meter App	Near bed	1 second
GPS Location	Run Keeper App	In pocket	1 second
Activity and Place	WhatWhere App	In pocket	30 minutes

The sampled data from the sensor network is collected at the central computer via a data collector application. The smart phone applications save the sampled data on the phone's file system. These data are transfered offline to the central computer for further analysis. The sampled data are written to the text files in a format which could be easily parsed. The sample log files are given in Figure 5.2. The data collection process took fifteen days for each subject. Every subject is provided with the WSN equipment, a central computer collecting data, a smart phone installed with the necessary applications. Subjects are not manually involved during the data collection process. The data files can be directly used by the data analysis applications without any modification.

5.2. Data Analysis and Feature Extraction

We implemented the data analysis and feature extraction programs to interpret the collected data. By using these programs, we can extract information about the sleep-related values, the mean, the variance and the range of the ambient values, the activity and place durations, and the commuting distance. Since the central computer and the smart phone are connected to the Internet, they use the UTC time and hence,

```
Time       dB   peak dB LAeq
2014/12/01 23:57:59 20.7      50.1      27.7
2014/12/01 23:58:00 18.6      43.0      27.6
2014/12/01 23:58:00 23.5      61.2      27.5
2014/12/01 23:58:01 17.3      37.1      27.4
2014/12/01 23:58:01 21.3      42.0      27.2
2014/12/01 23:58:02 21.7      49.6      27.2
2014/12/01 23:58:02 23.2      64.1      27.1
2014/12/01 23:58:03 26.2      59.0      27.1
2014/12/01 23:58:04 18.9      44.8      27.0
```

(a) Noise meter log file.

```
23:56:58.124 HUM 73.98
23:56:58.124 LIG 1.00
23:56:58.124 TMP 20.50
23:56:58.124 ATM 0.9901
23:57:33.724 MAT 1023
23:57:33.724 ACC 509 527 417
23:57:33.819 ACC 506 525 415
23:57:33.943 ACC 504 527 415
23:57:34.037 ACC 505 521 413
23:57:34.146 ACC 501 520 418
23:57:34.240 ACC 505 522 421
23:57:34.333 ACC 517 530 389
```

(b) Sensor data collector log file.

```
<trk>
  <name><![CDATA[Other 11/18/14 7:06 am]]></n
  <time>2014-11-18T05:06:47Z</time>
<trkseg>
<trkpt lat="40.861771000" lon="29.303348000">
<trkpt lat="40.861694000" lon="29.303392000">
<trkpt lat="40.861621000" lon="29.303448000">
<trkpt lat="40.861643000" lon="29.303333000">
<trkpt lat="40.861710000" lon="29.303258000">
<trkpt lat="40.861793000" lon="29.303196000">
<trkpt lat="40.861866000" lon="29.303123000">
<trkpt lat="40.861935000" lon="29.303066000">
<trkpt lat="40.861987000" lon="29.302971000">
```

(c) Runkeeper log file.

```
06:47:04.597   Home        GettingReady
08:11:49.250   Work        Working
08:59:45.940   Work        Working
09:38:06.993   Work        Working
11:31:08.379   Work        Working
12:01:15.799   Restaurant  Eating
13:09:04.133   Work        Working
13:39:11.324   Work        Working
14:09:13.595   Work        Working
14:39:20.504   Work        Working
15:17:30.517   Work        Working
17:03:19.159   Road        GoingHome
```

(d) Whatwhere log file.

Figure 5.2. Raw sensor reading log files.

the timestamps of data in different applications are synchronized. Since nine subjects collects data for fifteen days, there should be a total of 135 days of data. However, there are some incomplete experiments caused by several reasons such as damaged cables, dead batteries, or user negligence. Thus, the total number of error-free days used in the experiment is decreased to 101 days. We use three different applications for data analysis.

The sensor analysis application parses the given log files and calculates the nocturnal mean, variance, and range of ambient sensor readings including temperature, humidity, and atmospheric pressure. Moreover, *Noise Meter* log files are parsed by this application as well. The output is the mean, variance, and range of nocturnal noise levels. Also, the accelerometer magnitude values for $100ms$ time slices are calculated for sleep data analysis. The output features of data analysis are the sleep and wake up times, the total sleep duration, the sunrise time, the sleep durations before/after sunrise, the mean, variance, and range of noise level, temperature, humidity, and atmospheric pressure values for the night.

The activity analysis program parses *WhatWhere* log files. We can extract information about the time spent at home, outside, and bed. The time spent in bed is taken from the previous program's output. Furthermore, physical, mental, social, and leisure activity durations are calculated for each day. As listed in Appendix B, there are many activities that a person can perform during a day. For healthier analysis, we need to group the activities as physical, mental, social and leisure activities. *Sport*, *Traveling*, *Shopping*, *GettingReady*, *CommutingHome*, and *CommutingWork* activities are grouped as physical activities. Mental activities are composed of *Working* and *Studying*. Social activities consist of *Eating* outside and *HangingOut*. Lastly, leisure activities are comprised of *WatchingTV*, *Relaxing*, and *Reading*. Since *WhatWhere* is a notification-based application, the details of a day and the accuracy of the collected data is on subject's responsibility. The activity analysis program outputs more accurate and detailed results when the subject inputs more on *WhatWhere*.

For the sleep data analysis, the procedure explained in Figure 4.4, is executed. With this procedure, we are able to calculate the total sleep duration, the sleep onset latency, the deep sleep duration, and the number of deep sleep cycles for each subject's sleep.

The raw data including the commuting distance during a day for each person is taken from *RunKeeper* application's server site. Also, the GPS location log files can be downloaded for further analysis. All the durations in feature list are given in terms of minute. The full set of feature list can be seen in Table 5.2.

5.3. Inferences from Data Visualization

In this section, we plot collected data only for per subject and report only visually available inferences without applying any feature selection mechanism. Hence, the figures given in this section provide limited information about the relation between the sleep quality and the features. In the next section, we will apply feature selec-

Table 5.2. Full set of features.

Deep Sleep Ratio (%)	Sleep Duration (*min*)
Sleep Onset Latency (SOL) (*min*)	Number of Deep Sleep Cycles
Total Deep Sleep (*min*)	Sleep Duration Before Sunrise (*min*)
Sleep Duration After Sunrise (*min*)	Commuting Distance (*km*)
Mean Temperature (°C)	Mean Humidity (%)
Mean Atmospheric Pressure (*atm*)	Mean Noise Level (*dB*)
Variance of Temperature (°C)	Variance of Humidity (%)
Variance of Atmospheric Pressure (*atm*)	Variance of Noise Level (*dB*)
Range of Temperature (°C)	Range of Humidity (%)
Range of Atmospheric Pressure (*atm*)	Range of Noise Level (*dB*)
Time spent home (*min*)	Time spent outside (*min*)
Time spent in bed (*min*)	Physical Activity Duration (*min*)
Mental Activity Duration (*min*)	Social Activity Duration (*min*)
Leisure Activity Duration (*min*)	

tion algorithms which consider all the available data and hence provide more accurate information about this relation.

All the subjects employed in the data collection process are currently working. The activity and place durations of Subject 1 for sixteen days are given in Figures 5.3 and 5.5, respectively. The activity and place durations of Subject 1 for sixteen days are given in Figures 5.4 and 5.6, respectively. These durations are meaningful for one person's daily living, therefore they are reported only for Subject 1 and 5 separately. Note that the days 4, 5, 11, and 12 of Subject 1 and the days 1, 2, 5, and 6 of Subject 5 are weekend days. For instance, in Figure 5.3, it can be seen that mental activity durations for the days 4, 5, 11, and 12 are zero since Subject 1 does not work at weekends. Furthermore, place duration values for the same subject do not have significant differences between weekdays and weekends since Subject 1 is an outgoing person at weekends as shown in Figure 5.5. On the other hand, place duration values of Subject 5 for the days 1, 2, 5, and 6 are zero as shown in Figure 5.6 since Subject 5 does not go out at weekends. This fact can be also be seen from the commuting distances given in Figure 5.7. Subject 5 does not go out at weekends whereas the commuting distances of Subject 1 for the days 4, 11, and 12 are higher compared to the weekdays. The daily sleep parameters of Subject 1 and Subject 5 are shown in Figures 5.8 and 5.9, respectively. As seen in these figures, although a person is working during weekdays and not working during weekends, this does not show any obvious implication on the subject's sleep quality. Besides, some subjects prefer staying home at weekends, others can prefer going out, while some subjects sleep more than the others at weekends. However, these behaviors do not imply that those subjects get better quality sleep. As a result, we can say that the sleep quality is not only affected by the subject's lifestyle but also depends on other factors. Feature selection algorithms which consider all the available data and hence provide more accurate information about the factors affecting the sleep quality will be applied.

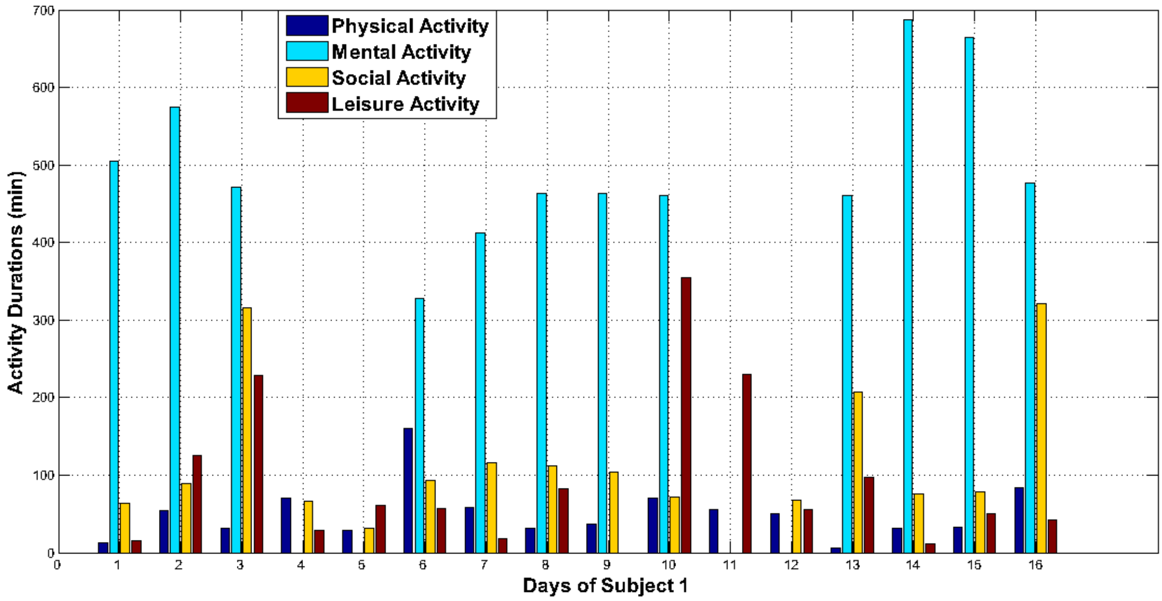


Figure 5.3. Activity durations of Subject 1 for 16 days.

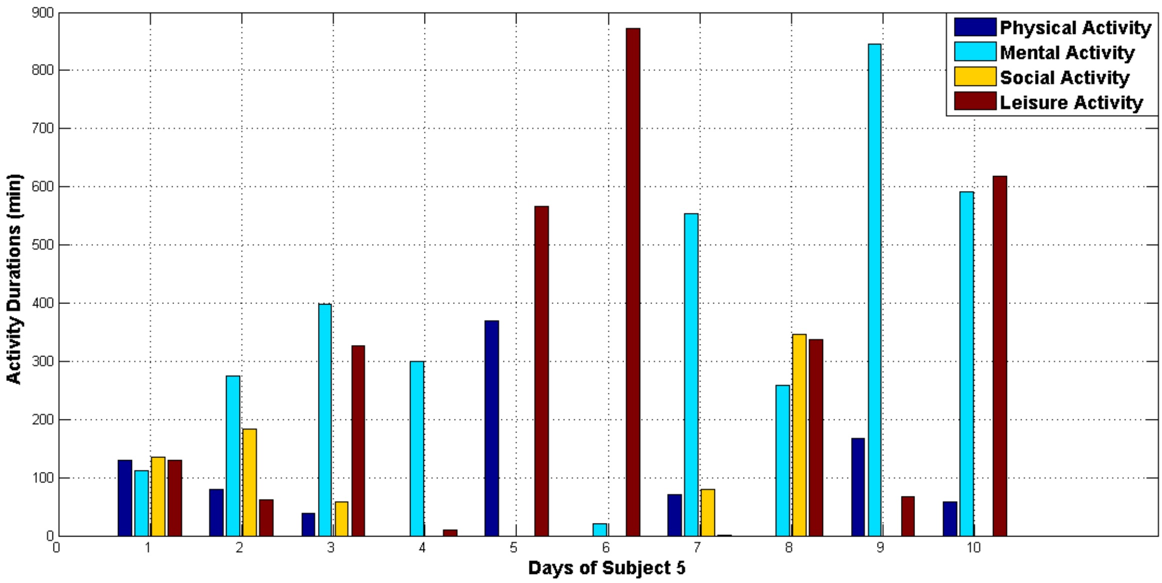


Figure 5.4. Activity durations of Subject 5 for 10 days.

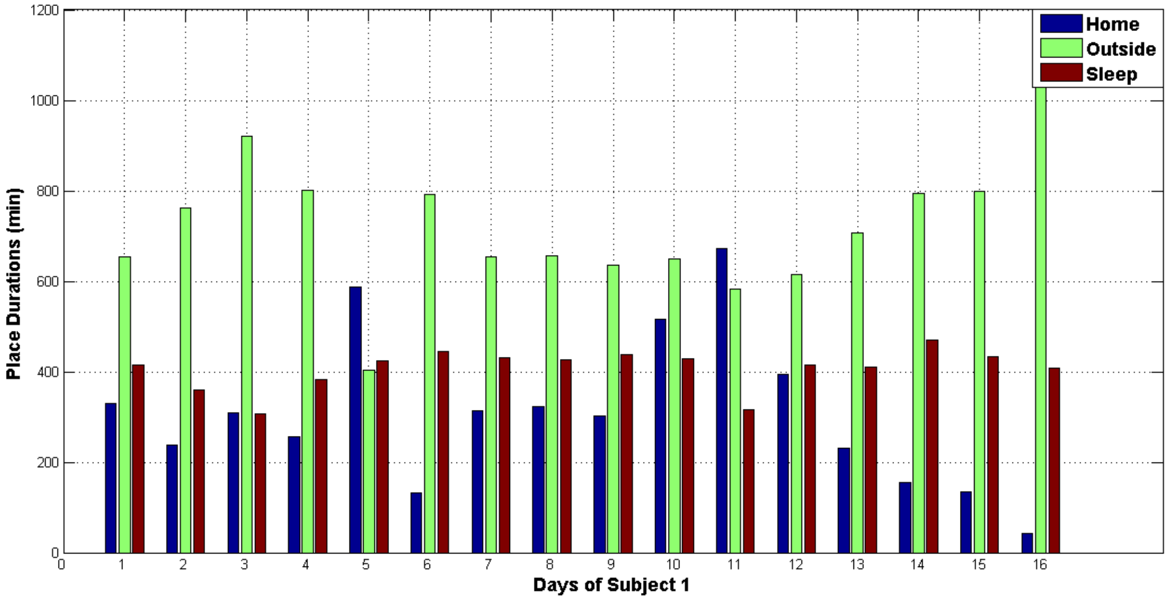


Figure 5.5. Place durations of Subject 1 for 16 days.

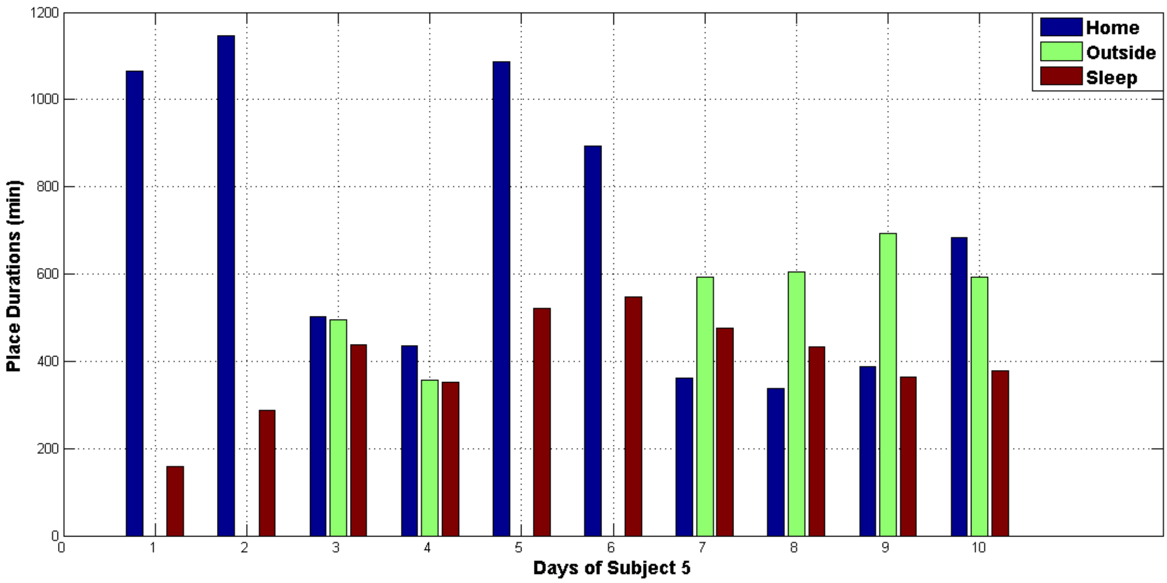


Figure 5.6. Place durations of Subject 5 for 10 days.

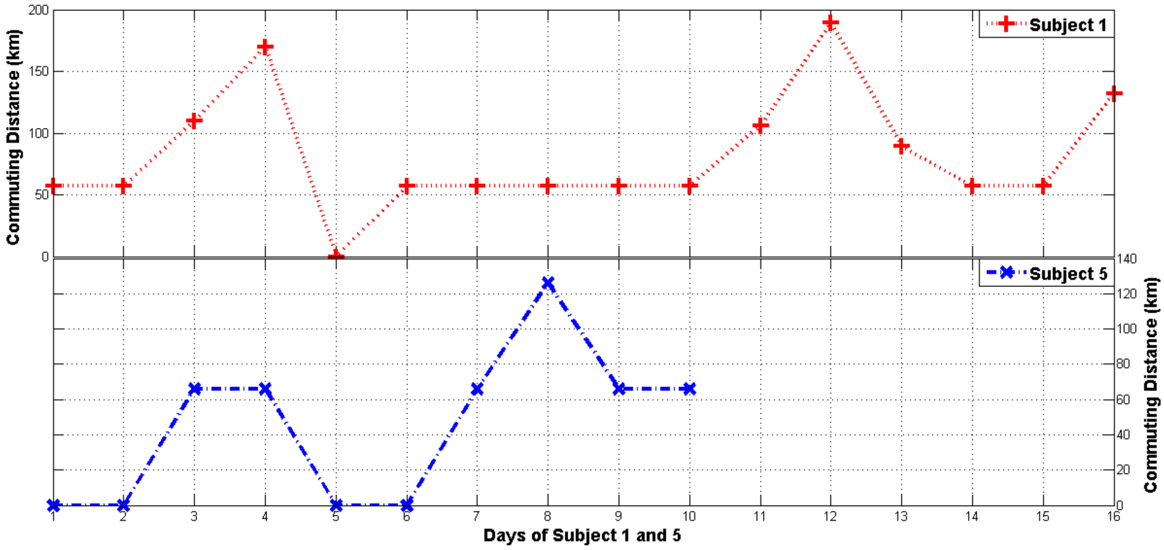


Figure 5.7. The commuting distances of Subject 1 and 5.

The visual representations of sleep parameters of Subject 1 and 5 are given in Figures 5.8 and 5.9, respectively. It can be observed that the deep sleep duration decreases when the sleep onset latency and/or the sleep duration after sunrise increase. For instance, in days 2 and 12, subject's SOL times are increased resulting in lower deep sleep durations. As other examples, in days 3, 10 and 11, the sleep durations after sunrise are increased because those days are the subject's off days. As a result, his deep sleep duration decreases dramatically in comparison to the previous days. Moreover, the SOL time is the time passed for a subject between going to bed and falling into sleep. A short SOL time generally indicates that the subject is quite ready to fall into sleep, most likely he/she gets tired with various activities during the day. These observations show that the deep sleep ratio is dependent on the sleep duration, the sleep onset latency (SOL) and the sleep duration after sunrise.

The deep sleep ratios of Subject 1 and 5 are given in Figure 5.10. This figure shows us that the sleep duration does not determine the sleep quality just by itself. For instance, on days 1, 2, and 5 for Subject 5 the deep sleep ratios are not higher compared to the other days, although, Subject 5 sleeps longer than the other days.

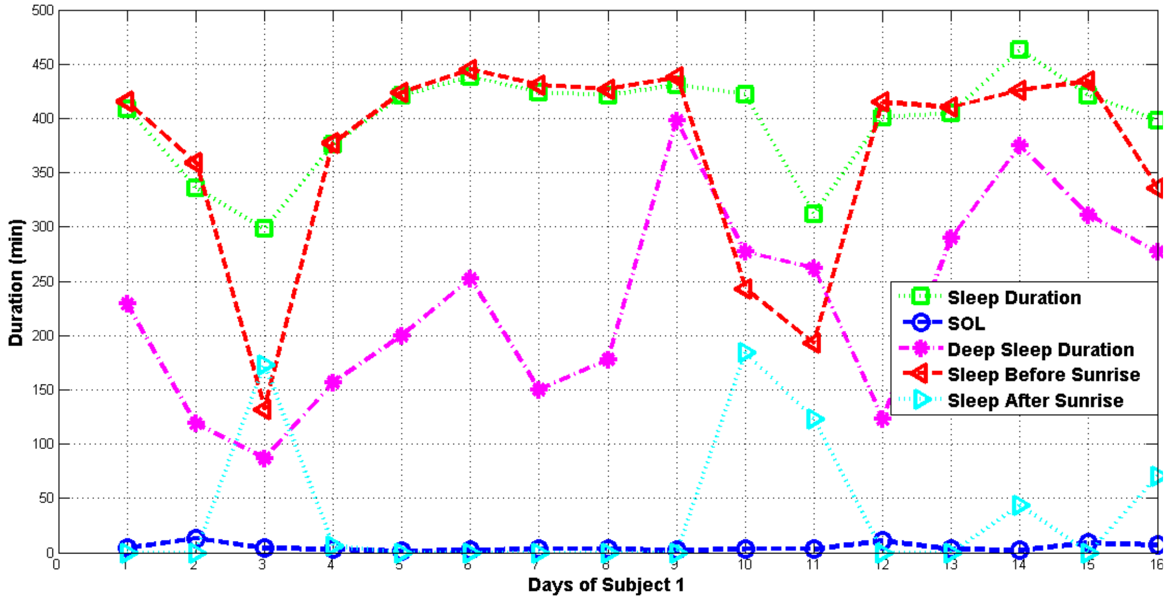


Figure 5.8. Sleep parameters of Subject 1 for 16 days.

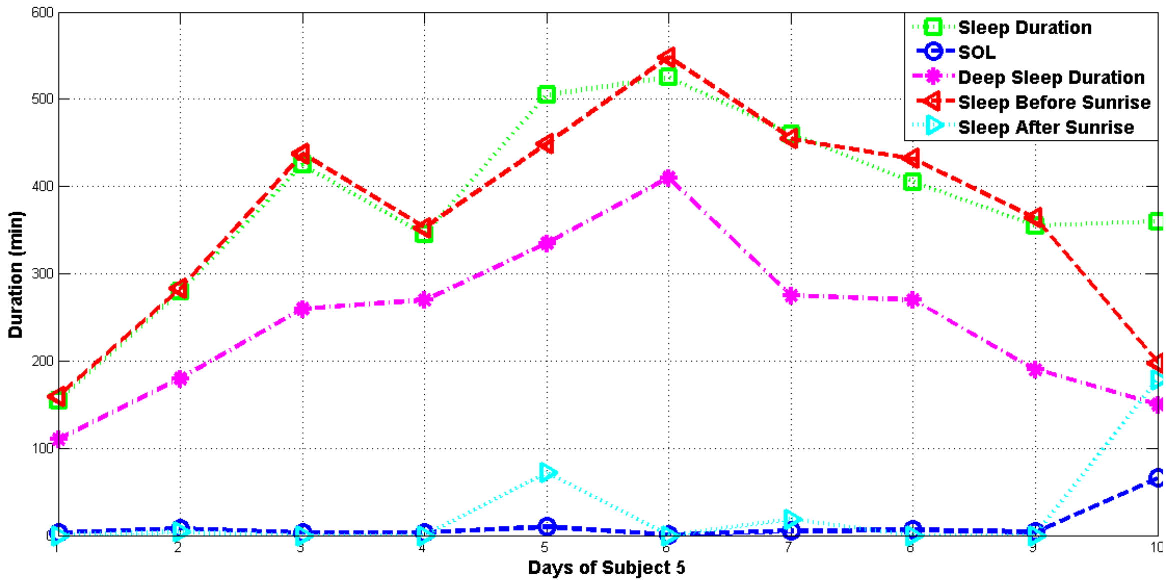


Figure 5.9. Sleep parameters of Subject 5 for 10 days.

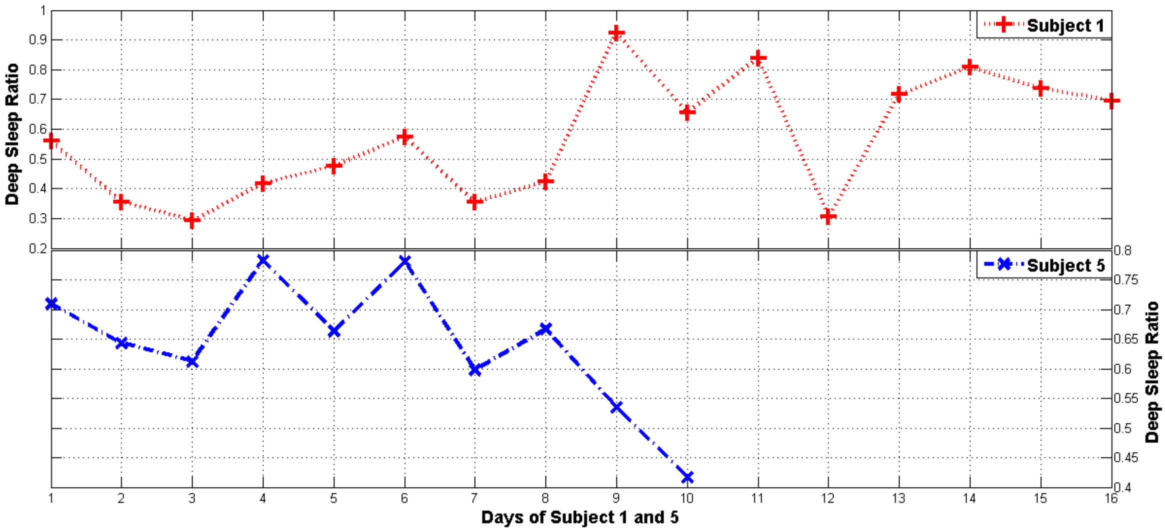


Figure 5.10. The deep sleep ratios of Subject 1 and 5.

The daily deep sleep ratios and the ambient values such as temperature, humidity, pressure, and noise level are given in Figure 5.11. The figure does not yield any obvious relation about the deep sleep ratio and the ambient factors. Hence, some technical methodologies such as machine learning algorithms should be used to find out the features affecting the sleep quality of a person.

The nocturnal ambient noise levels for two different subjects are given in Figure 5.12. It is known that Subject 5 is a snoring person and both subjects sleep in similar calm rooms. The visual data shows that the ambient noise level of Subject 5 is higher than that of Subject 1 and hence confirms that Subject 5 is a snoring person. Besides, the noise level fluctuates continuously during the night for snoring subjects. The higher noise levels at the start and the end of the night may indicate subject's movements before and after the sleep and an alarm clock set.

All the subjects participated in this study sleep in a calm room except Subject 5 since he is a snoring person and this results in a higher noise level. Moreover, as it can be seen in Figure 5.13, the average of daily noise levels does not change significantly.

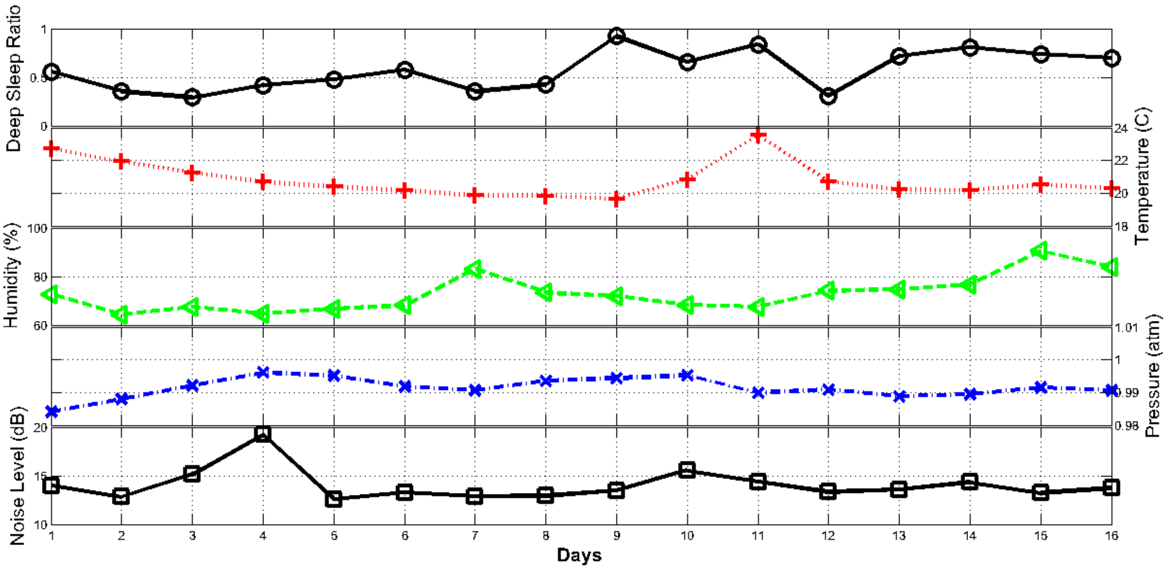


Figure 5.11. The deep sleep ratios and the ambient values of Subject 1 for 16 days.

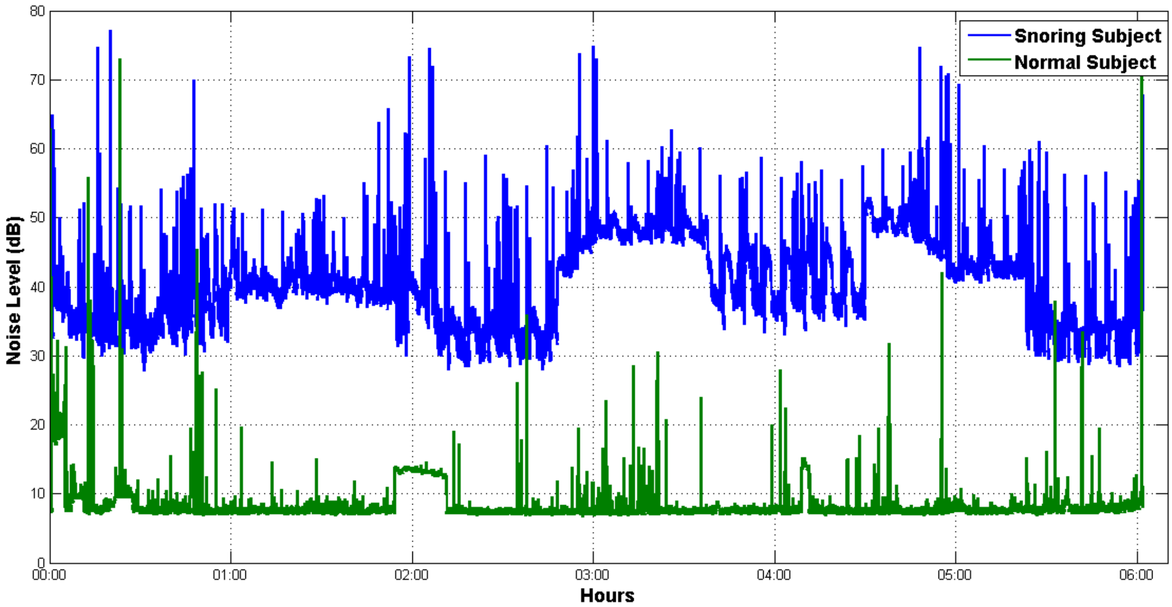


Figure 5.12. Snoring example of 2 different subjects for 1 day.

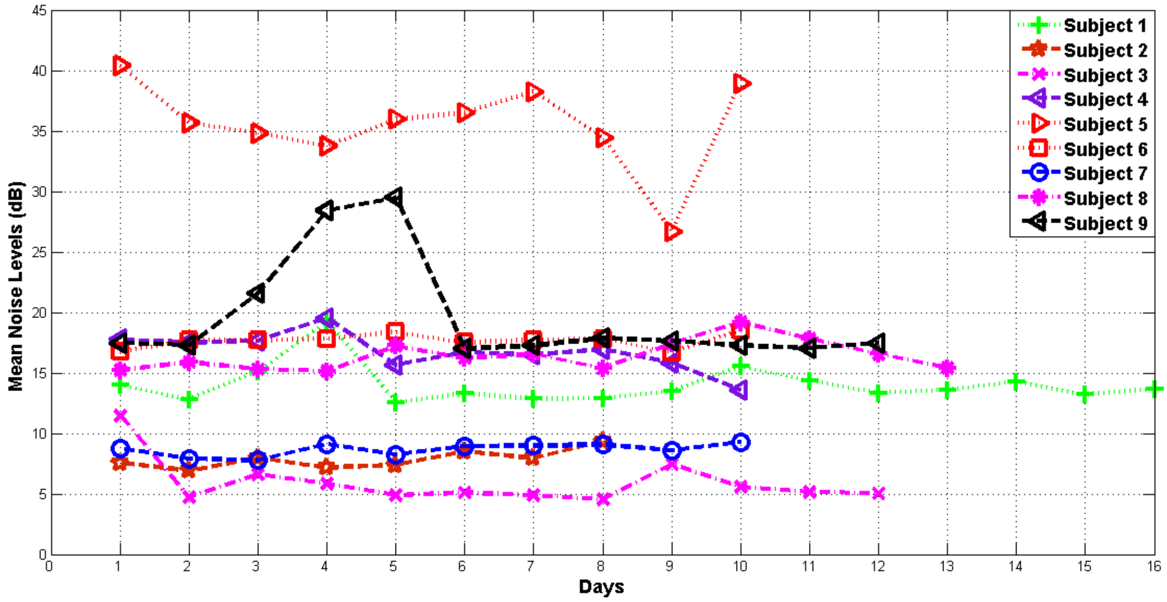


Figure 5.13. Mean noise levels(dB) of all subjects during experiments.

5.4. Feature Selection Algorithms

In this section, feature selection algorithms that are used to acquire the most important features for predicting the sleep quality of a person are explained. These algorithms include the decision tree, the correlation coefficient, and the sequential feature selection. Nocturnal dSQI score is calculated daily from the *dSQI* questionnaire application. The purpose of the feature selection is to find out the factors affecting the *dSQI* score and hence, the prediction target for all the algorithms is set to the nocturnal dSQI score.

5.4.1. Decision Tree for Feature Selection

A decision tree is constructed as a decision support tool to make predictions as new data arrives. The tree consists of rules and prediction results. The root node is the most important rule for classifying the data. The intermediary branches classify the data by making use of their rules. Each branch descending from a node represents the values of the feature in that node. The leaves represent the decision results for the data given to the decision tree. Decision trees are widely used in artificial intelligence, data mining and machine learning areas. The basic algorithm for decision trees is the greedy algorithm constructing the model in a top-down recursive divide-and-conquer manner [51]. Classification generation is completed when all the tuples in a subset belongs to the same class or the further separation can not be made.

In the decision tree, generally entropy-based measurement called as *information-gain* is used to select the attribute for the best separation of the training data. Moreover, *entropy* could be defined as the purity in a data set. For instance, the entropy is zero when all the data belongs to the same class. A simple decision tree algorithm yields only one decision tree for an example data set, though there may be multiple trees consistent with the data. An example implementation of decision tree is ID3 [52] which is biased in that it tends to prefer attributes with higher number of data. Hence, we preferred to use the C4.5 implementation [52] since this bias is suppressed by using an alternative measures called the *Information Gain Ratio* and the *Split Information*.

Accuracy of the decision tree might degrade in cases where there is a lot of noise associated with the training data and/or the number of training data is small. To eliminate this over-fitting problem, the C4.5 uses the *post-pruning* method [51]. The C4.5 creates an initial decision tree and performs *post-pruning* operation on this tree. This operation optimizes the tree by removing some nodes or subtrees from the original tree.

5.4.2. Correlation Coefficient Algorithm for Feature Selection

The correlation coefficient can be used to measure the inter-dependency of two random variables. The value of the correlation coefficient lies in between $+1$ and -1. If the coefficient value is $+1$, the variables are positively correlated at a maximum level. If the value is 0, then there is no correlation between them. If the value is -1, then they are negatively correlated at a maximum level. In this thesis, Pearson's correlation coefficient technique is used [53].

Assuming we have an *n-by-m* data matrix, the number of rows n represents the same type of distinct data and the number of columns m represents the different features of the data. After applying the correlation coefficient analysis to this data set, the result is the *m-by-m* matrix representing the correlation values of the features. Since, at the beginning, we have m features and hence the resultant matrix is an *m-by-m* matrix. The element of column i and row j in the resultant matrix is the correlation value between the features i and j. Additionally, the matrix is diagonally symmetric since the correlation between features i and j is represented both in $m(i,j)$ and $m(j,i)$. Moreover, the diagonal values are always 1 since every feature is maximally correlated by itself. The feature selection can be performed according to values in the matrix since those values are the covariance of the two features. If those features are highly correlated, then their covariance value is closer to ± 1.

5.4.3. Sequential Feature Selection

Sequential Feature Selection (SFS) can be defined as the elimination of features which yields no information for the model's classification of the data. SFS in the forward selection mode starts with an empty set of features and sequentially adds important features according to a specific evaluation function. This recursive selection stops when the selection criterion is reached. On the other hand, SFS in the backward selection mode starts with the full set of features and sequentially eliminates irrelevant features according to a specific evaluation function. The time complexity of elimination process is higher compared to the forward version of SFS [54]. At the end of the selection, the new feature set should perform the classification of the data with a minimum error [55]. The advantages of SFS are its simplicity and fast convergence time [56]. In our study, SFS in the forward selection mode is preferred since it is faster than the backward selection.

5.5. Implementation of Feature Selection Algorithms and Results

Our data set is composed of 101 separate days because there are some erroneous days that needed to be removed from the data set. When the dSQI scores are examined, it has been seen that there are no *Bad* sleep days in the data set. This shows us that the sleepers involved in the data collection do not suffer from sleep-related diseases and do not have any sleep disorders. Hence, the feature selection methods are applied to the data with two classes as *Good* and *Moderate*.

5.5.1. Decision Tree-based Feature Selection

For the feature selection part, we make use of CART [57] which is an implementation of C4.5 in MATLAB. The algorithm results in a five-node decision tree as shown in Figure 5.14 using the data of 101 days. The attributes utilized for creating the rules are *Leisure Activity Duration*, *Mean Noise*, *Variance of Temperature*, *Sleep Onset Latency*, and *Sleep Duration*. The *Leisure Activity Duration* is the most significant feature affecting the sleep quality of a person. The *Mean Noise* and the *Variance of Temperature* are the second significant features.

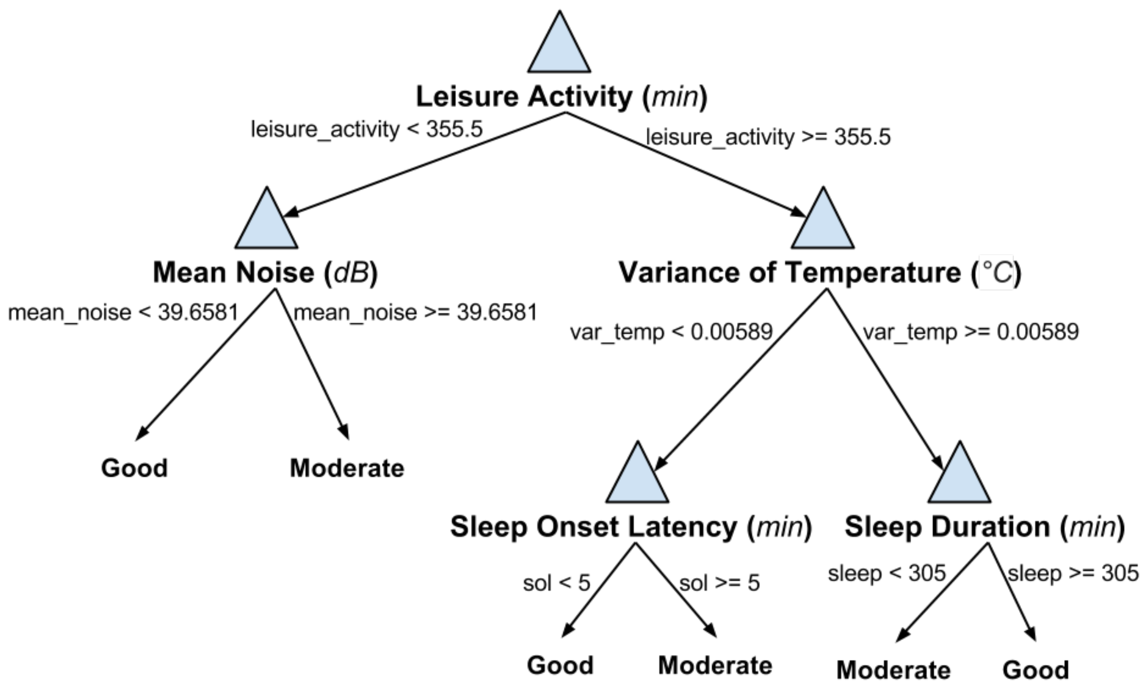


Figure 5.14. Decision tree for the sleep quality determination.

From the Figure 5.14, if the daily *Leisure Activity Duration* is less than *355.5 min*, the sleep quality of person can be classified as *Good* if the *Mean Noise* level is less than *39.65 dB*, otherwise it is classified as *Moderate*. When the duration of leisure activities such as watching TV and relaxation times increases; the perceived sleep quality of a person decreases. Thus, we can conclude that if a person does not engage in mental and/or physical activities, the quality of sleep for that night decreases. This situation explains the reason of increase in the dSQI score of Subject 8 on weekends since he prefers to stay home during weekends. Besides, the increase in the noise level of the room influences the sleep quality in a negative way. As an example, in our experiments we have detected that Subject 5 snores during his sleep. Also, the overall dSQI score is higher than the other participants. As indicated by the decision tree, his poor sleep quality is due to the ambient noise level in his bedroom.

According to the decision tree, if the daily *Leisure Activity Duration* is higher than *355.5 min*, the *Variance of Temperature* is higher than 0.06°C, and the *Sleep Duration* is higher than 305*min*, the dSQI score of a person can be classified as *Good*. The variance of ambient temperature turns out to be an important feature affecting the sleep quality of a person. Yet, the variance of temperature at the level of 0.006°C looks very small since we believe that humans are most likely not able to sense this much change in the temperature.

5.5.2. Correlation Coefficient-based Feature Selection

The input data is normalized to the range 0-1 because the order of magnitudes of the features are different. The input data matrix has 28 columns since we have 27 features from the experiments and one more column for the dSQI classes. Note that dSQI classes are obtained through the dSQI questionnaire application while the other features are obtained from the WSN and the smart phone. We used Pearson's correlation coefficient algorithm in MATLAB to obtain the correlation matrix whose size is *28-by-28*. The last column or the last row in the correlation matrix shows the level of correlation between the dSQI classes and the features. The correlation coefficients whose absolute values are higher than 0.1 are shown in Table 5.3. These are the eight most important features which affect the sleep quality.

Table 5.3. The most important features and correlation coefficients.

Feature Name	**Correlation Coefficient**
Sleep Onset Latency (SOL) (*min*)	0.3086
Commuting Distance (*km*)	0.2726
Leisure Activity Duration (*min*)	0.2555
Variance of Humidity (%)	0.2319
Sleep Duration After Sunrise (*min*)	0.171
Number of Deep Sleep Cycles	-0.1539
Social Activity Duration (*min*)	-0.1533
Deep Sleep Duration (*min*)	-0.1458

We observed that the *SOL* duration is the most important feature affecting the sleep quality with the correlation coefficient of 0.3086. This means that when the *SOL* duration increases, the dSQI score increases resulting worse sleep quality. Furthermore, the *Commuting Distance* has come out as the second important feature affecting sleep quality with the correlation coefficient of 0.2726. The increase in the commuting distance causes an increase in the dSQI score. Hence, we can say that the *Commuting Distance* has negative influence on the sleep quality of a person. Moreover, the *Leisure Activity Duration* is positively correlated with the dSQI score and its correlation coefficient is 0.2555. So, it has negative influence on the sleep quality of a person. Remember that *Leisure Activity Duration* is reported as the most important feature affecting the sleep quality in the decision tree as shown in Figure 5.14. On the other hand, the *Number of Deep Sleep Cycles* and the *Deep Sleep Duration* features are negatively correlated with the dSQI score. Thus, these features have positive influence on the sleep quality of a person.

5.5.3. Sequential Feature Selection

The MATLAB implementation of Sequential Feature Selection (SFS) in the forward selection mode is utilized to obtain the most significant features affecting the sleep quality. We prefer the forward selection mode because it is reported in literature as faster algorithm than the backward selection mode [54]. Since 10-fold cross validation technique is used for the selection criteria, the selection operation is performed ten times. The outcome of the SFS analysis is the set of features which classify the data set with the minimum error rate according to the dSQI classes. The resultant feature list is composed of *Sleep Onset Latency (SOL)*, *Deep Sleep Cycles*, *Deep Sleep Duration*, *Sleep Duration Before Sunrise*, *Sleep Duration After Sunrise*, *Leisure Activity Duration*, and *Variance of Humidity*. These results indicate that there are multiple features affecting the sleep quality of a person. According to the SFS analysis, the quality of sleep is affected by five sleep-related features. It is also affected by the daily activities of a person such as *Leisure Activity Duration* and the ambient factors of the sleep room such as *Variance of Humidity*.

To sum up, we performed three feature selection algorithms namely the decision tree, the correlation coefficient and the sequential feature selection to find the features which affect the sleep quality of a person for each algorithm is given in Table 5.4. The + entry means that this feature is selected as an important feature by the corresponding algorithm. As it can be seen from the table, only the *Sleep Onset Latency* and the *Leisure Activity Duration* features turned out to be significant for all the algorithms.

Table 5.4. Feature list affecting the sleep quality of a person.

Feature Name	**Decision Tree**	**Corr. Coef.**	**SFS**
Sleep Onset Latency (SOL) (*min*)	+	+	+
Number of Deep Sleep Cycles	-	+	+
Deep Sleep Duration (*min*)	-	+	+
Sleep Duration (*min*)	+	-	-
Sleep Duration Before Sunrise (*min*)	-	-	+
Sleep Duration After Sunrise (*min*)	-	+	+
Mean Noise (*dB*)	+	-	-
Variance of Temperature (°C)	+	-	-
Commuting Distance (*km*)	-	+	-
Leisure Activity Duration (*min*)	+	+	+
Social Activity Duration (*min*)	-	+	-

6. CONCLUSION

In the field of activity tracking studies, WSNs are preferred due to their low cost and low power consumption while smart phones are used because of higher computational power and portability for a variety of applications. The challenge is the development of a system which has an ability to track and monitor the inside and outside activities for all day long. In the literature, there are many systems which can track either inside or outside activities. To develop a seamless activity tracking system, sleep is another important activity which should not be disregarded since people spend one-third of their lives in sleep. The WSNs are generally preferred for collecting the environmental and sleep data while the smart phones are preferred for tracking the outside activities. Furthermore, systems developed for activity tracking studies should take the user's privacy issues and unobtrusiveness into account. Otherwise, people would not prefer such systems or become reckless for a careful usage.

In this thesis, we proposed a seamless activity tracking system which covers 24 hours of a person including his/her sleep behaviors. For the data collection phase, the system which consists of a WSN and a smart phone is utilized. In the WSN, there are two separate micro-controllers for collecting nocturnal ambient factors and sleep stage identification. On the other hand, in the smart phone there are four applications. *RunKeper* and *WhatWhere* are used to collect data about outside activities while *Noise Meter* and *dSQI* are used to collect sleep-related data. For the experiments, we employed nine people for fifteen days.

The main purpose of this study is to determine the factors which affect the sleep quality of a person. For this reason, we applied the decision tree, the correlation coefficient, and the sequential feature selection algorithms to the collected data. The target values are the dSQI classes which demonstrate a person's perceived sleep quality. Using the decision tree analysis, *Leisure Activity Duration*, *Mean Noise*, *Variance of Temperature*, *Sleep Onset Latency*, and *Sleep Duration* are reported as significant features affecting the sleep quality. According to the correlation coefficient analysis, *Sleep Onset*

Latency, *Leisure Activity Duration*, *Commuting Distance*, *Variance of Humidity*, *Sleep Duration Before Sunrise*, *Sleep Duration After Sunrise*, and *Social Activity Duration* turn out to be significant features. The sequential feature selection analysis yields that *Sleep Onset Latency*, *Number of Deep Sleep Cycles*, *Deep Sleep Duration*, *Sleep Duration Before Sunrise*, *Sleep Duration After Sunrise*, *Leisure Activity Duration*, and *Variance of Humidity* affect the sleep quality. As a result, only two features namely the *Sleep Onset Latency* and the *Leisure Activity Duration* are reported as important features in all three algorithms for their effects on the sleep quality.

In this thesis, we demonstrated that the sleep quality of a person depends on several ambient factors and daily activities by collecting 24-hour multi-modal data. Prior studies utilize only limited information such as accelerometer and pressure sensor data. However, we showed that the ambient factors and the daily activities should be considered as well. Our experiments demonstrated that several ambient parameters and life style choices are highly correlated with the sleep quality. In our study, we collected data from people of similar ages. Since people involved in the experiments do not have any sleep-related disorders, the collected data is not uniformly distributed. As opposed to similar studies in the literature which use only Actigraphy in sleep quality estimation, we took a holistic approach and showed that nocturnal ambient factors and activities performed during the day are also important in sleep quality together with sleep-related factors. As a future work, a more diverse participant set including the elderly and people suffering from sleep disorders can be studied for a more comprehensive evaluation of the determinants of the sleep quality.

APPENDIX A: DSQI SCORING SCHEME

PSQIDURAT (DURATION OF SLEEP)

IF Q4 > 7, THEN set value to 0

Q4 < 7 and ≥ 6, THEN set value to 1

IF Q4 < 6 and ≥ 5, THEN set value to 2

IF Q4 < 5, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIDISTB (SLEEP DISTURBANCE)

Each of Q6 subquestions scores 1 if answer is Yes, 0 otherwise

IF Q6b + Q6c + Q6d + Q6e + Q6f + Q6g + Q6h + Q6i = 0,

THEN set value to 0

IF Q6b + Q6c + Q6d + Q6e + Q6f + Q6g + Q6h + Q6i > 1 and < 2,

THEN set value to 1

IF Q6b + Q6c + Q6d + Q6e + Q6f + Q6g + Q6h + Q6i ≥ 3 and ≤ 5,

THEN set value to 2

IF Q6b + Q6c + Q6d + Q6e + Q6f + Q6g + Q6h + Q6i ≥ 6,

THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQILATEN (SLEEP LATENCY)

First, recode Q3 into Q3new thusly

IF Q3 ≥ 0 and ≤ 15, THEN set value of Q3new to 0

IF Q3 > 15 and ≤ 30, THEN set value of Q3new to 1

IF Q3 > 30 and ≤ 60, THEN set value of Q3new to 2

IF Q3 > 60, THEN set value of Q3new to 3

Check If Q3new = 2 and Q6a = Yes. If so, set Q3new = 3

Finally PSQILATEN = Q3new

PSQIDAYDYS (DAY DYSFUNCTION DUE TO SLEEPINESS)

IF Q8 = Yes then Q8 = 2 else Q8=0

Q10 is in the range 0-3

IF Q8 + Q10 = 0 THEN set value to 0

IF Q8 + Q10 $\geq$ 1 and $\leq$ 2, THEN set value to 1

IF Q8 + Q10 $\geq$ 3 and $\leq$ 4, THEN set value to 2

IF Q8 + Q10 = 5, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIHSE (SLEEP EFFICIENCY)

Diffsec = Difference in seconds between day and time of day Q1 and day Q2

Diffhour = Absolute value of diffsec / 3600

newtib = IF diffhour > 24, then newtib = diffhour – 24

IF diffhour $\leq$ 24, THEN newtib = diffhour

tmphse = (Q4 / newtib) * 100

IF tmphse $\geq$ 85, THEN set value to 0

IF tmphse < 85 and $\geq$ 75, THEN set value to 1

IF tmphse < 75 and $\geq$ 65, THEN set value to 2

IF tmphse < 65, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQISLPQUAL (OVERALL SLEEP QUALITY)

Q9

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIMEDS (NEED MEDS TO SLEEP)

Q7

If yes THEN set value 2 ELSE set value 0

PSQI SCORE

DURAT + DISTB + LATEN + DAYDYS + HSE + SLPQUAL + MEDS

TOTAL ≤ 4 **good sleep quality**

$4 <$ TOTAL < 9 **moderate sleep quality**

TOTAL ≥ 9 **poor sleep quality**

Minimum Score = 0 (better); Maximum Score = 20 (worse)

APPENDIX B: WHATWHERE CONFIGURATION FILE

Places:

- Home
- Work
- Gym
- Road
- Mall
- Bar
- Restaurant
- Other

Activities:

- GettingReady
- CommutingWork
- CommutingHome
- Working
- Traveling
- Sport
- Studying
- WatchingTV
- Shopping
- HangingOut
- Relaxing
- Reading
- Eating
- Other

TimerInterval 1800000

REFERENCES

1. Arduino, *Arduino Platform*, 2005, `www.arduino.cc/`, [Accessed January 2015].

2. Digi, *Xbee ZigBee Module*, 2005, `www.digi.com/xbee/`, [Accessed January 2015].

3. Alemdar, H. and C. Ersoy, "Wireless Sensor Networks for Healthcare: A survey", *Computer Networks*, Vol. 54, No. 15, pp. 2688–2710, 2010.

4. Ertan, H., H. Alemdar, O. D. Incel and C. Ersoy, "Designing A Wireless Sensing System for Continuous Behavior and Health Monitoring", *6th International Workshop on Ubiquitous Health and Wellness (UbiHealth 2012)*, Newcastle, UK, 2012.

5. Lockhart, J. W., T. Pulickal and G. M. Weiss, "Applications of Mobile Activity Recognition", *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pp. 1054–1058, ACM, New York, NY, USA, 2012.

6. Kose, M., O. Incel and C. Ersoy, "Performance Evaluation of Classification Methods for Online Activity Recognition on Smart Phones", *Signal Processing and Communications Applications Conference (SIU), 2012 20th*, pp. 1–4, IEEE, 2012.

7. Incel, O. D., M. Kose and C. Ersoy, "A Review and Taxonomy of Activity Recognition on Mobile Phones", *BioNanoScience*, Vol. 3, No. 2, pp. 145–171, 2013.

8. Bush, V., "As We May Think", *Atlantic Monthly*, Vol. 176, No. 1, pp. 641–649, 1945.

9. Microsoft, *Kinect for Microsoft*, 2010, `www.microsoft.com/`, [Accessed January 2015].

10. Sony, *Sony Smartwatch*, 2012, `www.sonymobile.com/`, [Accessed January 2015].

11. WHO, *Technical Meeting on Sleep and Health*, 2004, www.who.int/en/, [Accessed January 2015].

12. Drucker-Colin, R., "The Function of Sleep is to Regulate Brain Excitability in order to Satisfy the Requirements Imposed by Waking", *Behavioural Brain Research*, Vol. 69, No. 1–2, pp. 117–124, 1995.

13. Miwa, H., S.-I. Sasahara and T. Matsui, "Roll-over Detection and Sleep Quality Measurement using a Wearable Sensor", *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 1507–1510, 2007.

14. Terán-Santos, J., A. Jimenez-Gomez and J. Cordero-Guevara, "The Association between Sleep Apnea and the Risk of Traffic Accidents", *New England Journal of Medicine*, Vol. 340, No. 11, pp. 847–851, 1999.

15. Lattimore, J.-D. L., D. S. Celermajer and I. Wilcox, "Obstructive Sleep Apnea and Cardiovascular Disease", *Journal of the American College of Cardiology*, Vol. 41, No. 9, pp. 1429–1437, 2003.

16. Samy, L., M.-C. Huang, J. J. Liu, W. Xu and M. Sarrafzadeh, "Unobtrusive Sleep Stage Identification Using a Pressure-Sensitive Bed Sheet", *IEEE Sensors Journal*, Vol. 14, No. 7, pp. 2092–2101, 2014.

17. Buysse, D. J., C. F. Reynolds, T. H. Monk, S. R. Berman and D. J. Kupfer, "The Pittsburgh Sleep Quality Index: A New Instrument for Psychiatric Practice and Research", *Psychiatry Research*, Vol. 28, No. 2, pp. 193–213, 1989.

18. Deutsch, P. A., M. S. Simmons and J. M. Wallace, "Cost-effectiveness of Split-night Polysomnography and Home Studies in the Evaluation of Obstructive Sleep Apnea Syndrome", *Journal of Clinical Sleep Medicine AASM*, Vol. 2, No. 2, pp. 145–153, 2006.

19. Suetsugi, M., Y. Mizuki, K. Yamamoto, S. Uchida and Y. Watanabe, "The Effect of Placebo Administration on the First-night Effect in Healthy Young Volunteers", *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, Vol. 31, No. 4, pp. 839–847, 2007.

20. Silber, M., S. Acoli-Israel, M. Bonnet, S. Chokroverty, M. Grigg-Damberger and M. Hirshkowitz, *The Visual Scoring of Sleep in Adults*, 2007.

21. TDSM, *Natural Patterns of Sleep*, 2007, `healthysleep.med.harvard.edu/`, [Accessed January 2015].

22. Guazzelli, M., I. Feinberg, M. Aminoff, G. Fein, T. Floyd and C. Maggini, "Sleep Spindles in Normal Elderly: Comparison with Young Adult Patterns and Relation to Nocturnal Awakening, Cognitive Function and Brain Atrophy", *Electroencephalography and Clinical Neurophysiology*, Vol. 63, pp. 526–539, 1986.

23. Nassir, A. and O. Barnea, "Wireless Body-Area Network for Detection of Sleep Disorders", *Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, pp. 1–5, 2012.

24. ActiGraph, *Accurate and Reliable In-home Sleep Assessment*, 2011, `www.actigraphcorp.com/`, [Accessed January 2015].

25. Jeon, J. H., J. Yeon, S. goo Lee and J. Seo, "Exploratory Visualization of Smartphone-based Life-Logging Data using Smart Reality Testbed", *BigComp*, pp. 29–33, 2014.

26. Kozlovszky, M., L. Bartalis, B. Jokai, J. Ferenczi, P. Bogdanov, Z. Meixner, L. Nemeth and K. Karoczkai, "Personal Health Monitoring with Android-based Mobile Devices", *Information Communication Technology Electronics Microelectronics (MIPRO), 2013 36th International Convention on*, pp. 326–330, 2013.

27. Yoshihara, Y., D. Tang and N. Kubota, "Life Log Visualization System Based

on Informationally Structured Space for Supporting Elderly People", *2013 Second International Conference on Robot, Vision and Signal Processing*, Vol. 2, pp. 78–83, 2013.

28. Oracle, *SunSPOT*, 2004, `www.sunspotworld.com/`, [Accessed January 2015].

29. Burns, W., C. D. Nugent, P. J. McCullagh and H. Zheng, "Design and Evaluation of a Smartphone-Based Wearable Life-Logging and Social Interaction System", *2014 IEEE 27th International Symposium on Computer-Based Medical Systems, New York, NY, USA, May 27-29, 2014*, pp. 435–440, IEEE, 2014.

30. Ryoo, D.-W. and C. Bae, "Design of The Wearable Gadgets for Life-Log Services based on UTC.", *IEEE Trans. Consumer Electronics*, Vol. 53, No. 4, pp. 1477–1482, 2007.

31. Kim, P. H. and F. Giunchiglia, "Life Logging Practice for Human Behavior Modeling.", *SMC*, pp. 2873–2878, IEEE, 2012.

32. Iwamoto, T., H. Touyama and M. Matsumoto, "A Method for Determining Sleep Stages by Using Doppler Sensor", *SICE Annual Conference (SICE), 2013 Proceedings of*, pp. 2380–2385, 2013.

33. Nassir, A. and O. Barnea, "Wireless Body-Area Network for Detection of Sleep Disorders", *Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, pp. 1–5, 2012.

34. Pino, E., A. Dorner de la Paz, P. Aqueveque, J. Chavez and A. Moran, "Contact Pressure Monitoring Device for Sleep Studies", *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pp. 4160–4163, 2013.

35. Adami, A. M., A. G. Adami, T. L. Hayes and Z. T. Beattie, "A System for Assessment of Limb Movements in Sleep", *e-Health Networking, Applications Services*

(Healthcom), 2013 IEEE 15th International Conference on, pp. 419–423, 2013.

36. Samy, L., M.-C. Huang, J. Liu, W. Xu and M. Sarrafzadeh, "Unobtrusive Sleep Stage Identification Using a Pressure-Sensitive Bed Sheet", *Sensors Journal, IEEE*, Vol. 14, No. 7, pp. 2092–2101.

37. Kealy, A., K. McDaid, J. Loane, L. Walsh and J. Doyle, "Derivation of Night Time Behaviour Metrics using Ambient Sensors", *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*, pp. 33–40, 2013.

38. Tajima, Y., M. Nakata and K. Takadama, "Personalized Real-Time Sleep Stage Remote Monitoring System", *8th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 1–5, 2014.

39. Malaekah, E. and D. Cvetkovic, "Automatic Sleep Stage Detection using Consecutive and Non-Consecutive Approach for Elderly and Young Healthy Subject", *Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), 5th ISSNIP-IEEE*, pp. 1–6, 2014.

40. Peng, Y.-T., C.-Y. Lin and M.-T. Sun, "A Distributed Multimodality Sensor System for Home-Used Sleep Condition Inference and Monitoring", *1st Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare (D2H2)*, pp. 20–23, 2006.

41. Urbandroid, *Sleep As Android*, 2014, `play.google.com/`, [Accessed January 2015].

42. FTDI, *Future Technology Devices International (FTDI) Chip*, 2005, `www.ftdichip.com/`, [Accessed January 2015].

43. Digi, *X-CTU Software*, 2005, `www.digi.com/xctu/`, [Accessed January 2015].

44. Honeywell, *Humidity Sensors*, 2008, `sensing.honeywell.com/`, [Accessed January 2015].

45. Sensortec, *Digital Barometric Pressure Sensor*, 2009, `www.bosch-sensortec.com`, [Accessed January 2015].

46. JINASYS, *Noise Meter*, 2013, `play.google.com/store/`, [Accessed January 2015].

47. FitnessKeeper, *Run Keeper*, 2014, `play.google.com/store/`, [Accessed January 2015].

48. TI, *ez430-Chronos Watch*, 2010, `www.ti.com/tool/ez430-chronos/`, [Accessed January 2015].

49. Fry, B. and C. Reas, *Processing Programming Language*, 2004, `www.processing.org/`, [Accessed January 2015].

50. Natale, V., M. Drejak, A. Erbacci, L. Tonetti, M. Fabbri and M. Matroni, "Monitoring Sleep with A Smartphone Accelerometer", *Sleep and Biological Rhythms*, Vol. 10, No. 4, pp. 287–292, 2012.

51. Ratanamahatana, C. A. and D. Gunopulos, "Scaling up the Naive Bayesian Classifier: Using Decision Trees for Feature Selection", *Proc. Workshop Data Cleaning and Preprocessing (DCAP'02)*, at IEEE International Conference on Data Mining (ICDM'02), 2002.

52. Quinlan, J. R., "Induction of Decision Trees", *Machine Learning*, Vol. 1, No. 1, pp. 81–106, 1986.

53. Wu, W.-J. and Y. Xu, "Correlation Analysis of Visual Verbs Subcategorization based on Pearson's Correlation Coefficient", *International Conference on Machine Learning and Cybernetics (ICMLC)*, Vol. 4, pp. 2042–2046, 2010.

54. Jain, A. and D. Zongker, "Feature Selection: Evaluation, Application and Small Sample Performance", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 19, No. 2, pp. 153–158, 1997.

55. Marcano-Cedeño, A., J. Quintanilla-Domínguez, M. Cortina-Januchs and D. Andina, "Feature Selection using Sequential Forward Selection and Classification Applying Artificial Metaplasticity Neural Network", *36th Annual Conference on IEEE Industrial Electronics Society (IECON'2010)*, pp. 2845–2850, IEEE, 2010.

56. Kudo, M. and J. Sklansky, "Comparison of Algorithms that Select Features for Pattern Classifiers", *Pattern Recognition*, Vol. 33, No. 1, pp. 25–41, 2000.

57. Breiman, L., J. Friedman, R. Olshen and C. Stone, *Classification and Regression Trees*, Wadsworth and Brooks, Monterey, CA, 1984.