CMPE 492

Senior Project Final Report

01.06.2016

Sleep Quality Monitoring with the Smart Watch

Advisor

Dr. Bert Arnrich

Students

A. Müge Kurtipek / 2011400159

Akın Toksan / 2010400186

1. INTRODUCTION AND MOTIVATION

Fitness and health trackers have come into fashion for several years already. Some people are getting more aware about their health since there are multiple behavioral factors that are ingrained in a person's daily routine such as long and tiring studying hours, monthly loans, high level of stress at work, malnutrition, limited sleep time and so on.

Speaking of the limited sleep time, decreasing sleep quality (SQ) is one of the biggest issues in modern life, which is affected by industrial disasters, and medical and other occupational errors, besides the other factors that we mentioned above [1]. It is actually considered as a public health problem [2]. Almost 70 million US adults have sleep disorder [1]. Lack of a good sleep or other sleep pathologies, which are caused by disrupted biological (circadian) rhythms can lead to several diseases such as bipolar disorder, major depressive disorder [3], heart disease, cancer, and obesity [4]. Other than that, sleep quality (affected by circadian rhythms) is also involved in structuring a person's mood, level of concentration and digestion [5].

For that matter, monitoring a person's sleep quality and making proper changes to improve it is crucial. There are some effective solutions to diagnose the sleep in terms of quality and other aspects such as Polysomnography (PSG). But PSG has comprehensive setting including cables, which require various sensors like EEG (Electroencephalography) to monitor brain waves during the sleep. It is also very expensive which is another factor to limit the usage only for clinical area. Another solution for monitoring the sleep is actigraphy. It is cheaper and a watch-like device or a smart watch collects the data from the user. People can use it at home, without spending a night necessarily in a clinic, unlike PSG.

Sleeping cycle of a person has three stages in turn: Light sleep, deep sleep, REM (rapid eye movement) sleep. REM is the stage when the person dreams during the sleep, so intense brain activities occur in this stage. Light sleep is the stage when the person calms down, relaxes his/her muscles, lowers the heartbeat and breathing rate. Deep sleep is the stage when the hormones are triggered to regulate body growth, to repair cells and to restore of energy. The main criteria to calculate the sleep quality is not the total time of the sleep but the distribution of these different sleep stages [6]. A longer deep sleep period and a proper wake-up during the light sleep stage contribute to a better sleep. Also, waking up at the proper time in light sleep stage helps for a better mental and physical health [7].

Our objective for the project is to detect physical and biological changes occurred during the sleep and their relation with the sleep quality. There are different parameters and biological characteristics that affect the sleep quality such as body movements, heart rate, snoring, breathing rate and so on. Hence, for our project, we have used an off-the-shelf smart watch for data collection with its built-in sensors.

Besides, we have also asked user some questions about his/her sleep quality in the next morning in order to have a ground truth for our interpretation of the collected sensor data. We have used Pittsburgh Sleep Quality Index [9] questionnaire to see the sleep quality score of the previous night and have trained our system with these scores.

In the upcoming sections, we present the state of art, our approach to the problem, the methods we have used and some possible future work.

2. STATE OF ART

As a state of art we selected 3 main papers and studies to have a better understanding about the topic and the current solutions.

2.1. Intelligent Sleep Stage Mining Service with Smartphones

In this paper [10], researchers are presenting Sleep Hunter, which is a mobile application for sleep stage detection. They are using actigraphy [11] approach to identify the light sleep, REM and deep sleep. The application then can be used to calculate the sleep quality of the user and to smart call service, which means waking up the user at the best time.

Apart from brain activities, they are focusing on distinguishable and detectable physical activities that a user does during different sleep stages. For instance, large body movements and fast heart beat are more likely to be seen during the light sleep, short movements like arm and leg trembling are more likely to be seen during the deep sleep. Because the user dreams during REM sleep movements like reaching out to air and activities like talking during the sleep are more likely to be seen during the REM stage. Also apart from the physical activities of a person, the sleep environment can affect the sleep quality. That is why they are collecting data from the microphone, accelerometer and light sensor of the smart phone.

2.1.1. System Design

The authors' approach to detect the sleep stages starts with classifying different body movements that can be done in different stages. To be able to classify the movements, they found some threshold values by making experiments on 10 volunteer subjects. They asked the subjects to perform different body movements while they have their smart phones under their pillows. Then, they analyzed the magnitudes of the accelerometer data and found some threshold values to classify the movements. They followed a similar approach ("first conduct the experiment, find the thresholds and then apply") to classify sound data and light sensor data, too.

2.1.2. Evaluation

After finding the classifying threshold values for each type of data they moved on to the phase of building the detection model. For that purpose, the authors proposed linear-chain Conditional Random Field (CRF) to integrate the features. They estimated the weights of each features with making calculations on the training data then they decide on the ones, which gave them the best results.

As an experiment they collected 90 night sleep data of people from different ages. During these nights, they wanted the subjects to also wear a device named Zeo, which is a sleep quality detection system based on EEG. They took on the Zeo results as a ground truth. Then they made 10-fold cross validation and came up with the result in **Table 1** below.

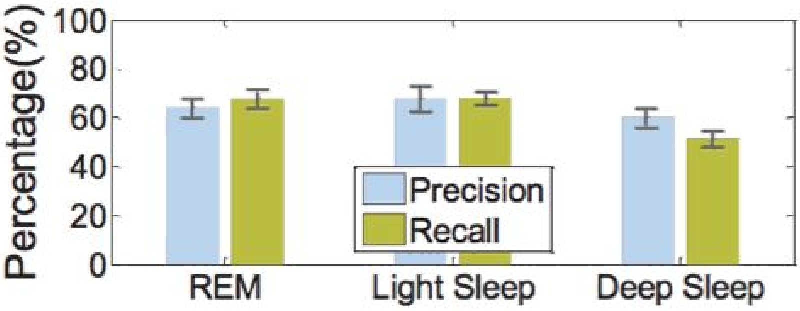

Table 1. The results of cross validation

To test their sleep stage detection system, the authors collected 30 more night sleep data and also asked the subjects to wear Zeo again. The results of the test data are in **Table 2.** The accuracy of the Sleep Hunter is found much higher than the existing commercial actigrahpy based products.

Ground Truth	Predictions: REM	Predictions: Light Sleep	Predictions: Deep Sleep	Recall
REM	538	206	39	68.71%
Light Sleep	246	630	77	66.11%
Deep Sleep	61	108	174	50.73%
Precision	63.67%	66.74%	60.00%	64.55% Accuracy

Table 2. Performance of sleep stage detection

2.1.3. Shortcomings

Since they build the application Sleep Hunter for Android smart phones, the user needs to sleep with the phone under his/her pillow. It is increasing the false positive detections of movements such as the movements of the person the user sleeps with. However, we have used smart watches to detect the sleep quality and it will enable us to reduce these kinds of false positive movement detections since the watch will be on the user's wrist.

The other shortcoming of the paper is the lack of biological activity monitoring during the sleep. As they also mentioned in the paper, heart rate is also another significant determinant of sleep stages. Since they used smartphones they were not able to collect heart rate data. But during our experiments we have used smart watches and we have also taken heart rate changes during the sleep time.

2.2. SleepTight: Low-burden, Self-monitoring Technology for Capturing and Reflecting on Sleep Behaviors

In the second paper [12], a low-burden and self-monitoring system named SleepTight is presented to the readers. Authors state that manual tracking of a health behavior increases the user's awareness and engagement with the issue. However, keeping the user to provide the needed input for a long-term is challenging. So, they implemented an app which is making use of the Android's widgets to keep user both interested about the app and to lower the data capture burden. (Figure 1)

Since SleepTight is a manual tracker, user should provide enough meaningful data to the system. So authors designed the system as a sleep diary, which makes it easier for the user to access the information about his/her behavior and encourage the self-reflection about his/her sleep.

The main goal of the app is to collect the data about *multiple behavioral factors* that any system could not achieve automatically, such as eating behaviors, tobacco, caffeine and alcohol consumption and medication. Another goal of the authors was to lower the capture burden to prevent the user from giving up the self-reflection completely.

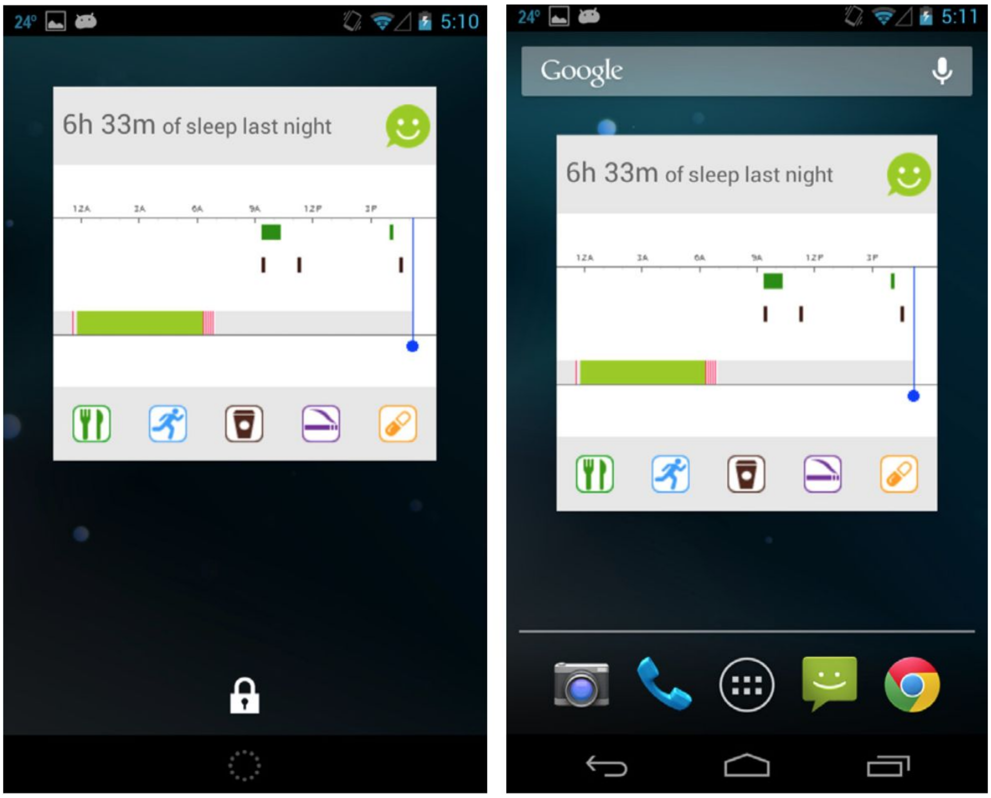

Figure 1. Lock screen and home screen for SleepTight

2.2.1. System Design

The main advantage of the system is that it is easily accessible through the lock screen and the home screen of the smart phone. So user can provide data without actually opening the app.

Besides from the multiple behavioral factors taken from the user as an input, system also collects data about user's sleep behavior manually, such as time to go to bed, minutes to fall asleep, waking time, and finally the subjective sleep quality "very poor" and "very good". (Figure 2)

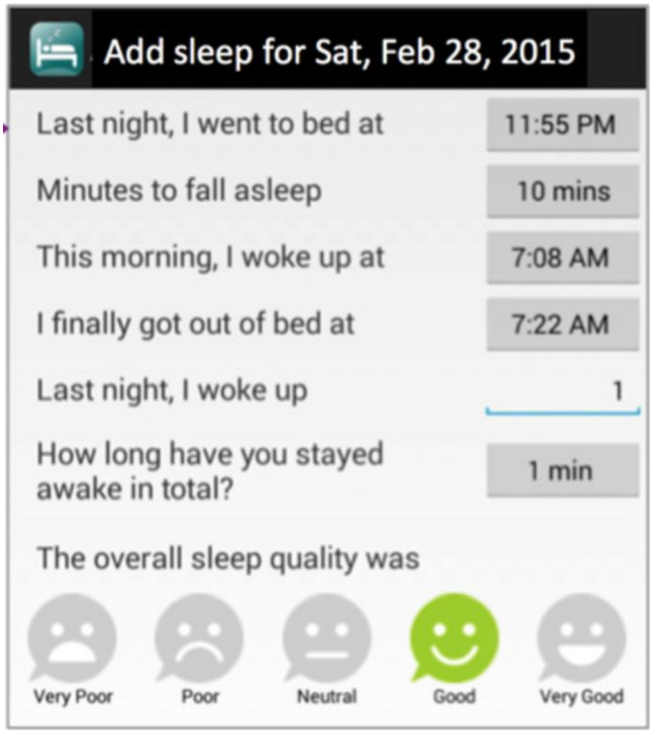

Figure 2. Survey to collect data in the next morning

2.2.2. Evaluation

Authors wanted to evaluate the effects of the widgets. So they separated the subjects into 2 groups to compare the two versions of the application: the full system (including lock screen and home screen widgets) and the app-only system (with no widgets)

Then, they made two face-to-face sessions with the subjects: one before they used the app and one after they used it for 4 weeks.

After the study, it turns out that 91% of the subjects used the home screen widget but only the 18% used the lock screen. In the sessions, 82% of the subjects expressed that they want to set some goals about their sleep such as waking up and going to bed at a certain time to have a regular sleep cycle, reducing the amount of time to get out of the bed after they wake up, feeling rested when they wake up and so on.

2.2.3. Shortcomings

The app does not detect the sleep quality by using any sensors but asks it to the user to keeping the logs of the behavior. So, the data collection is all manual and requires the full attention of the user to give accurate inputs of the behavior. Also, app does not give any advice about the sleep behavior of the patient but rather focuses on the comparison of different sleep night due to multiple behavioral factors, and leave the decision to the user.

Manually tracked data also construct a ground truth in our approach to detect the sleep quality.

2.3. Will You Have a Good Sleep Tonight? Sleep Quality Prediction with Mobile Phone

In the third paper [12], the authors are searching the relation between sleep quality and human context. By human context they mean human's physiological data, social data and surrounding environment. So, they are asking themselves whether the sleep quality of a person is depend on the physical activities of the person during the day or the environmental changes like light, temperature, pressure and social contact.
Their ground truth data is the daily version of the Pittsburg Sleep Quality Index (PSQI) [9]. They collected data from smartphone sensors like accelerometer, photonic sensors and microphone with the mobile application SleepMiner. Then with comparing the results of the collected data with the ground truth thresholds, they identify the relations of human context in sleep quality.

2.3.1. System Design

In the SleepMiner mobile application, they collected 4 types of data. First one is sound data, which is collected from the microphone of the phone in specified time periods. Second one is localization data, which is collected from GPS module of the smart phone. Third one is the light sensor data, which is collected by built-in light sensors of the smart phone. And the last one is acceleration data, which is also triggering other sensors to start sense. To identify the human context they also get logs of the daily phone usage of the user and the text and call records.

2.3.2. Evaluation

They made experiments on 15 subjects with 30 days of sleep data. They asked subjects to answer the PSQI questions every morning when they fully wake up and carry their smart phones with them, running SleepMiner application during the day. After analyzing the data, they found and report the correlation between daily activities, social and environmental and the sleep quality of a person. Their assessment is in **Table 3**.

Factors used	Accuracy	Precision	Recall	F1-score
Previous day	59.42	67.50	79.41	72.97
+ Attributes	73.91	84.85	82.35	83.58
+ Social	65.22	71.05	79.41	75.00
All	78.26	87.50	82.35	84.85

Table 3. Performance of SleepMiner %

2.3.3. Shortcomings

In the paper they focused on the effects of environmental conditions in one's sleep quality. However they are not looking the sleep patterns of the user while he/she is sleeping. The main shortcoming of the paper is not taking into consideration of sleep data. In our study, we have collected sensor data during the sleep.

3. METHODS

In our study we have used Samsung Gear S smart watch. To monitor the physical activities like turning from one side to another, leg jerking, body trembling we have used built-in acceleration sensor of the watch. To monitor the biological activities like change in heart rate in different sleep stages we have used built-in heart rate sensor.

We have collected data from smart watch's sensors in every 10 minutes for 80 seconds during the sleep. We determined the time periods considering the performance of the battery. We have collected data from both accelerometer and heart rate sensor with frequency of 20 Hz.

3.1. Experiments

3.1.1. Data Collection

We recorded total 15 nights sleeping data from 2 different subjects. In one night of a record, we had x, y, z values for acceleration; alpha, beta, gamma values for gyroscope and heart rate values with frequency of 20 Hz.

We collected sensor data for 80 seconds in every 10 minutes due to battery issues. So our data was not continuous. Each night's raw data was approximately 12 MB large. When we combined one night data it was about 1.5 hours long for a 8 hours of sleep.

3.1.2. Pittsburgh Sleep Quality Index (PSQI)

In order to label each night with its sleep quality, we used Pittsburg Sleep Quality Index (PSQI). [13] During collecting our training and test data we asked the subjects to fill the survey when they wake up with indicating date and their names. At the end, we combined the data and the labels from the survey with using these dates. For subjects' convenience, we had an online questionnaire on Google Forms [14] so that they could easily fill out from their phones when they woke up.

#	Questions	Answer
1	When did you go to bed last night?	
2	What about the time (in minutes) you approximately take to fall asleep?	
3	When did you get up in this morning?	
4	How long (in hours) was your sleep duration in last night?	
5	What is ratio between the two durations of sleep and of staying in bed? (can be obtained from answers 1, 3 and 4.)	
6	The sleeping trouble maybe because (Yes/No)	
6a	Have pain	
6b	Wake up in the middle of the night or early morning	
6c	Have to get up to use the toilet	

6d	Cannot breathe comfortably	
6e	Cough or snore loudly	
6f	Feel too cold	
6g	Feel too hot	
6h	Have sad dreams or even nightmares	
7	Have you taken medicine to help sleep? (Yes/No)	
8	Within the past week, how many times did you have trouble in staying awake in driving or eating, or have trouble in engaging social activity?	
9	How would you rate the overall sleep quality during the last week?	

Table 4.a. PSQI questionnaire to score SQ

#	Formula
#2	< 16min: 0, 16-30min: 1, 31-60min: 2, or > 60min: 3
#3	> 7am: 0, 6-7am: 1, 5-6am: 2, or < 5am: 3
#5	> 85%: 0, 75% ~ 84%: 1, 65% ~ 74%: 2, or < 65%: 3
#6	s = sum of scores from #6a, #6b,..., up to #6h where Yes and No yield 0 and 1, respectively. If s, 0: 0, 1 or 2: 1, 3 ~ 5: 2, or ≥ 6: 3
#7	Yes: 1 and No: 0
#8	nill: 0, once or twice: 1, or more than twice: 2
#9	very good: 0, fairly good: 1, fairly bad: 2, or very bad: 3

Table 4.b. Scoring criteria table

Once a subject filled the **Table 4.a** we were able to calculate his/her sleep quality score with the scoring criteria in **Table 4.b**. Higher the score coming from the **Table 4.b**, lower is the sleep quality.

Good Sleep	Average Sleep	Bad Sleep
PSQI score 0-3	PSQI score 4-8	PSQI score 9-16

Table 5.a. Sleep Quality Classification with PSQI results

Beside from the PSQI scoring, we also used subjective user opinion about the sleep as an alternative label. It was the answer of the question 9 in PSQI survey **(Table 4.b)**.

Good Sleep	Bad Sleep
Answer for question #9 is 0 or 1	Answer for question #9 is 2 or 3

Table 5.b Sleep Quality Classification with Subject's Opinion

3.1.3. Extracting Feature Sets

First, we used acceleration mean, acceleration variance, gyro mean and gyro variance values as a baseline to classify sleeps in terms of quality, since almost every paper that we referenced do so.

In the paper "The role of actigraphy in the study of sleep and circadian rhythms" [15], it is stated that sleep quality is directly affected by the distribution of the sleep stages. Each sleep stage has its characteristic physical and biological activities. For example, large body movements like body rollover and body trembling are more likely to occur during light sleep, small body movements like arm trembling, leg jerking are more likely to occur during deep sleep. Also heart rate can help us to classify these stages. Heart rate can be higher during REM sleep than the deep sleep because of the dreaming. Also from an audial perspective, speaking during the sleep and mumbling are more related with dreaming so more likely to occur during the REM sleep. Snoring is observed mostly during the light and deep sleep.

So, beside the baseline features which are indicating that there is a movement during the sleep, it was also important to classify these movements and to count their occurrences. That's why, before calculating the SQ score, we made experiments to identify the movements. We also included heart rate mean and heart rate variance as features.

Features
Acceleration Mean (accMean)
Acceleration Variance (accVar)
Gyroscope Mean (gyroMean)
Gyroscope Variance (gyroVar)
Heart Rate Mean (hrMean)
Heart Rate Variance (hrVar)
Movement Percentage (movement)
Non-movement Percentage (nonMovement)
Body Rollover Percentage (bodyRollover)
Body Trembling Percentage (bodyTrembling)
Arm Reaching Percentage (armReaching)
Arm Trembling Percentage (armTrembling)

Table 6. Feature Sets

3.1.4. Body Movement Detection

We have designed an experiment and main objective of the experiment was to find threshold values for different movements in order to identify their counts in the raw data in further steps. We decided on measuring 4 different on-the-bed body movements namely “body rollover (turning sideways)”, “body trembling”, “arm reaching to air” and “arm trembling”. We conducted the experiment on 2 subjects and ask them to repeat each 4 movements 10 times while wearing the smart watch. After collecting the data from them we plotted the acceleration magnitude and spectrogram of the movements to see the differences between them.

Figure 3 shows “body rollover” movement of a subject and **Figure 4** shows “arm trembling” movement for the same subject. As can be seen from the plots, the acceleration magnitude patterns look slightly different (bottom part of the plot, under the spectrogram). Body rollover last longer with a lower acceleration magnitude.

From the spectrogram, we can see the energy differences of these movements. Since arm trembling is more sudden and shorter movement, it has a higher frequency.

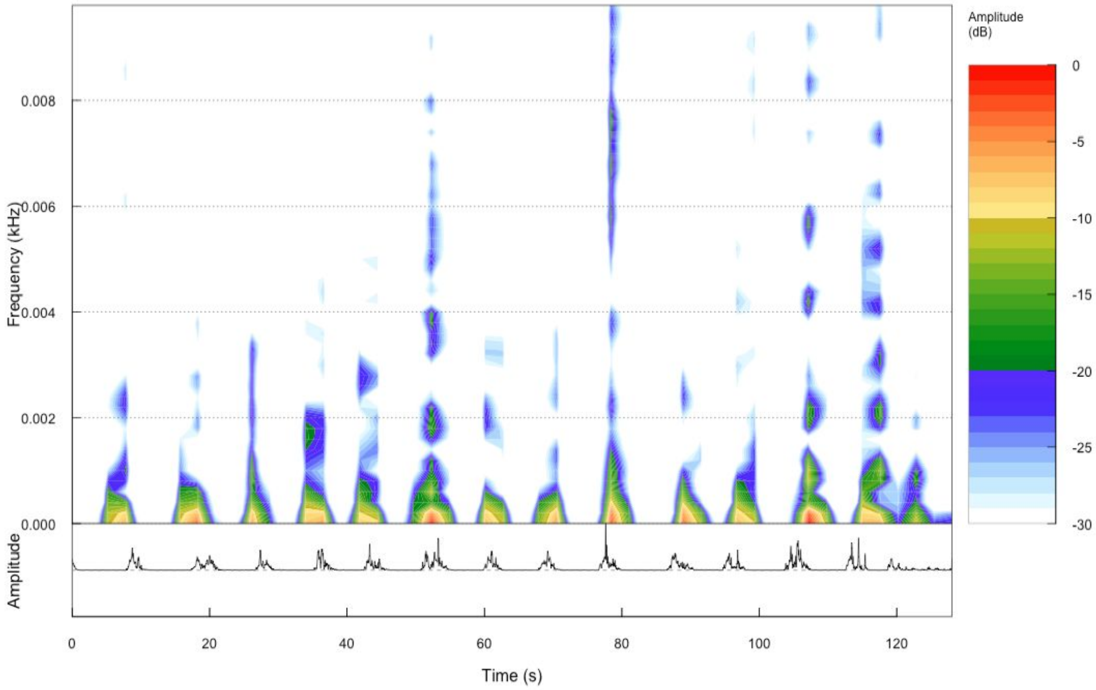

Figure 3. Spectrogram and plot for "Body Rollover" movement

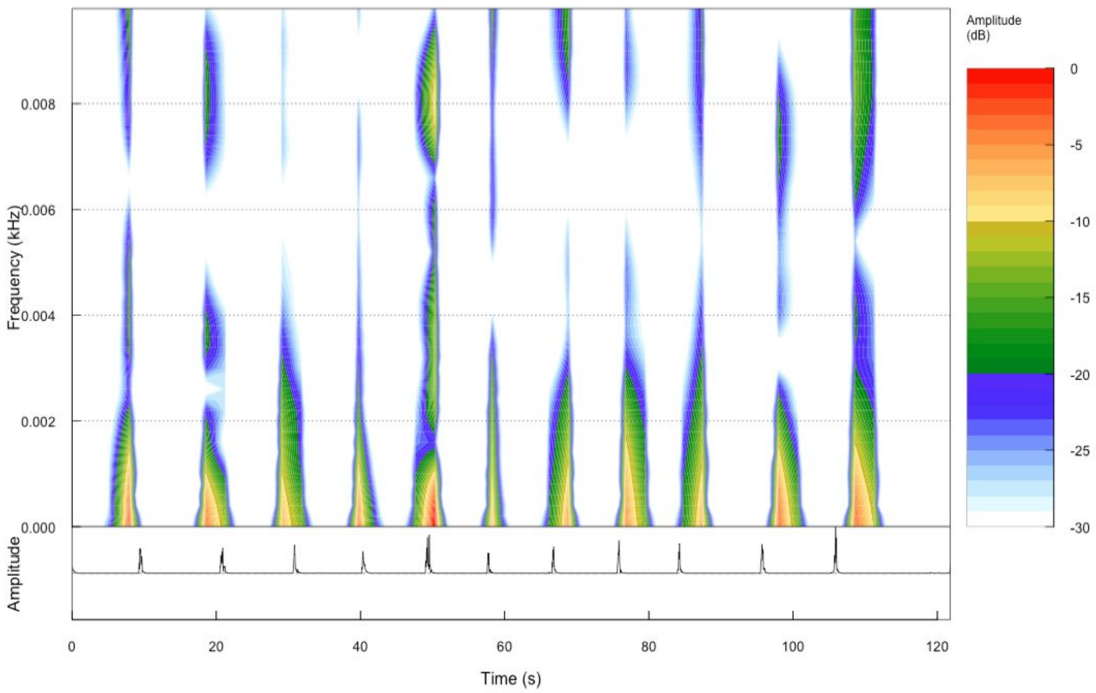

Figure 4. Spectrogram and plot for "Arm Trembling" Movement

We used windowing for deeper analysis of the movement differentiating. We divided raw data into 5-second windows with 50% overlapping. We labeled each window with the labels a1 (arm reaching), a2 (arm trembling), b1 (body rollover), b2 (body trembling) and n (non movement). With using R software, we draw a decision tree from the data to see the thresholds (**Figure 5**). This tree helped us through the process of classifying the movements. According to the thresholds in the decision tree, we counted the different movement occurrences for each night and we added these counts as features to our classifier to compute the overall sleep quality of the night.

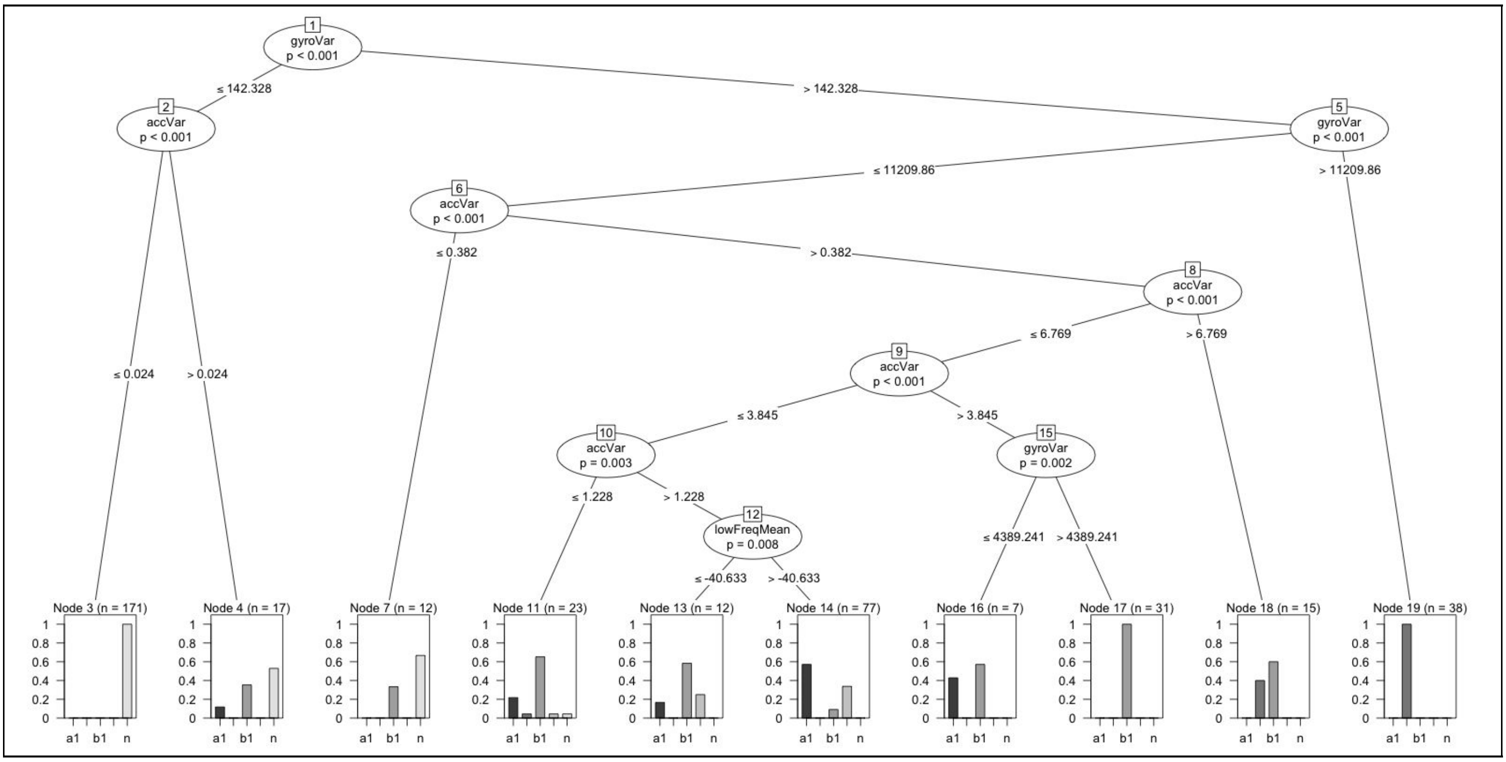

Figure 5. Decision Tree for Movement Detection

4. RESULTS

After creating the feature set in **Table 6,** we analyzed each night data one by one and found the values of each feature for each night. At the end, to train and test our system, we had 15 rows of data. Each row was representing a night and in each row we had the feature values as well as the two types of labels of that night's sleep.

In the overall data 7 in 15 was labeled as good sleep and 8 were average sleep. There weren't any bad sleep recorded. However in terms of good and average it was a balanced data.

Our data looks as in **Table 7**:

name	acc Mean	acc Var	gyro Mean	gyro Var	hr Mean	hr Var	arm Reaching	arm Trembling	body Rollover	body Trembling	non Movement	movement	Subject's Opinion Label	PSQI Label
Muge1	0,07855	0,15570	9,839	0,03427	68,55	128,09	4,5	2,0	25,2	0	1236,3	31,7	2	1
Muge2	0,06709	0,09803	9,854	0,02084	70,67	92,28	4,1	1,0	20,3	0	1100,6	25,4	2	3
Muge3	0,10386	0,25540	9,823	0,06021	64,73	29,20	5	3,0	32,3	0	1189,7	40,3	2	3
Akin4	0,11842	0,28358	9,822	0,07084	61,14	77,62	5,3	8,0	38,3	0	1231,4	51,6	3	4
Akin5	0,10495	0,22423	9,824	0,05770	57,11	161,33	5,1	0,0	37,5	0	1218,4	42,6	2	4
Akin6	0,12387	0,36827	9,836	0,11688	64,45	342,09	5,8	6,0	29,2	0	998	41	3	5
Akin7	0,08260	0,12743	9,739	0,01606	61,31	189,15	3,1	0,0	12,2	0	499,7	15,3	4	6
Akin8	0,11960	0,29399	9,791	0,09351	57,94	163,02	5	0,0	32,3	0	1191,7	37,3	2	4
Akin9	0,14541	0,43089	9,758	0,13882	60,53	102,56	2,9	1,0	23,6	0	823,5	27,5	3	3
Akin10	0,11330	0,29997	9,807	0,12002	71,50	591,22	3,8	0,0	29	0	880,2	32,8	2	4
Muge11	0,07894	0,13958	9,754	0,03733	65,91	66,72	5,2	0,0	28,6	0	1516,2	33,8	1	1
Muge12	0,09602	0,18839	9,808	0,05999	63,46	21,96	8,6	0,0	41,6	0	1399,8	50,2	1	0
Akin13	0,10025	0,21594	9,802	0,06607	62,39	229,12	18,7	0,8	41	4,1	1252,4	64,6	2	4
Muge14	0,07839	0,14953	9,773	0,02584	64,73	20,05	9,5	0,0	22,1	1,8	1207,6	33,4	3	5
Muge15	0,07428	0,15321	9,736	0,01717	64,13	56,42	14,7	2,0	25,7	4,2	1307,4	46,6	2	3

Table 7. Raw data of 15 nights

During our experiments we used binary logistic regression. PSQI Label of the sleep was the dependent variable and all the features in **Table 6** were the independent variables. We applied stepwise method of the binary logistic regression to find the features, which contribute to the results most. We made training and testing experiments in 3 different formats: without any cross validation (here test set is equal to the training set), with 5-fold cross validation and 1-night-out cross validation. The results can be found in **Table 8.** We also made the accuracy calculations of binary logistic regression according to Subject's Opinion Labels, but it didn't change the results since the labels are same with PSQI labels in binary case.

	Accuracy	Features
5-fold-cross-validation	62,6%	accMean + gyroVar + hrVar + movement + armReaching
1-night-out-cross-validation	60%	accMean + gyroVar + hrVar + movement + armReaching
without-cross-validation	73%	accMean + accVar +gyroMean +gyroVar + nonMovement+ bodyRollover + armTrembling + hrMean + armReaching + bodyTrembling

Table 8. Binary Logistic Regression Results with PSQI Label

PSQI Labels are between 0 and 16, and Subject's Opinion Labels are between 1 and 4. By reducing it to the binary (good or average) we calculated the binary logistic regression accuracies. To find the accuracy without reducing to the binary we used Linear Regression. Again we used stepwise method of linear regression to find the features contribute to the accuracy result most. **Table 9** shows the accuracy results according to Subject's Opinion Label for 3 different cross validation that we tried with the features, which gave the highest results. **Table 10** shows the accuracy results according to PSQI Label for 3 different cross validation that we tried with the features, which gave the highest results (Expanded versions of the feature set abbreviations can be found in Table 6).

The difference between the accuracy results of PSQI and Subject's Opinion Labels are not surprising. PSQI Label accuracy results are much more lower since PSQI has 17 prediction intervals for the calculated score can fall whereas Subject's Opinion Labels leads to only 4 intervals.

	Accuracy	Features
5-fold-cross-validation	73%	accMean + gyroMean + gyroVar + armReaching + bodyRollover + hrVar
1-night-out-cross-validation	80%	accMean + gyroMean + gyroVar + armReaching + bodyRollover + hrVar
without-cross-validation	93%	accMean + gyroMean + gyroVar + armReaching + bodyRollover + hrVar

Table 9. Linear Regression Results with Subject's Opinion Label

	Accuracy	Features
5-fold-cross-validation	33%	accMean + gyroMean + gyroVar + hrVar + movement + armReaching
1-night-out-cross-validation	33%	accMean + gyroMean + gyroVar + hrVar + movement + armReaching
without-cross-validation	53%	accMean + gyroMean + gyroVar + hrVar + movement + armReaching

Table 10. Linear Regression Results with PSQI Label

5. CONCLUSION AND DISCUSSIONS

After all, we've seen that it is possible to detect whether a person slept well or not by collecting data with a smartwatch. Our results can be improved by collecting more data from diverse group of people. Also some new features such as audio and light can be added to the system to have more precise results.

We saw that sleep quality has strong relations with the physical and biological activities that one perform during sleep. There exist more accurate state-of-art solutions for sleep quality monitoring but smart watch is a more easy-to-use device than PSG and EEG since it has no cables or any other attachments, and since it is attached to one's body it is more accurate than using a smart phone.

From our results, we can conclude that finding one's sleep quality in a more general way is easier. During our experiments we got higher accuracy values for the binary logistic regression, where the dependent variable was the label of the quality (Good or bad) and independent variables were the features in feature set. **(Table 6)** We got lower accuracy values for the linear regression, where dependent variable was the label of the quality in detailed (Score between 0-16) and independent variables were the feature set. From these results we concluded that it is harder to say one night's sleep quality with exact percentage, but it is easier to classify the night's sleep as Good, Average or Bad.

We believe that our project can be integrated into the daily life of smartwatch users who want to have better night sleeps and healthier lives. Following-up their sleep quality scores regularly and seeing the improvement in the scores, people can be more aware about how can they eat better, exercise better and have a better psychological state in general. And in the long run, this improvement in the public health can help societies to have more social and sharing members.

6. FUTURE WORK

As we mentioned in the results section, we could more or less detect the quality of a person. The main difficulty that we encountered was not having sufficient number of nights for data collection, because we had to share the smartwatch with other groups. As an improvement for the detection, more data should be collected from different people as much as possible to increase the accuracy of the system.

As an improvement, some new features can be added to the feature set such as light and audio data during the sleep. Also, we only checked for 4 types of movement. So, feature sets can be extended with more movement types.

Another obstacle for our project was to not collect a continuous data. Thus, we could somehow missed some of the movements during the sleep, since we collected data in every 10 minutes.

In the scope of wider study in the future, a more accurate ground truth can be referenced to instead of PSQI scoring. We used PSQI for calculating the subject's SQ to use as a ground truth to our own calculated SQ with the system. But it turns out that, PSQI scoring is not that significant to distinguish the sleep qualities. It was really hard to have a "bad sleep" score for a subject according to the questionnaire, despite the subject was really insistent on having a bad sleep.

Lastly, our system can be implemented into the smart watch with Tizen, so that the results can appear on the screen in real-time when the user wakes up. This would provide a better user experience and also would make it easier for the user to track the analysis of his/her previous sleeps.

7. REFERENCES

[1] Institute of Medicine. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Washington, DC: The National Academies Press; 2006.
[2] http://www.cdc.gov/features/dssleep/
[3] Harvey, A. Sleep and circadian rhythms in bipolar disorder: seeking synchrony, harmony, and regulation. American Journal of Psychiatry 165, 7 (2008), 820–829.
[4] Foster, R. G., and Wulff, K. The rhythm of rest and excess. Nature Reviews Neuroscience 6, 5 (2005), 407–414.
[5] Foster, R. G., and Kreitzman, L. Rhythms of life: the biological clocks that control the daily lives of every living thing. Yale University Press, 2005.
[6] Carpenter, J. S., and Andrykowski, M. A. Psychometric evaluation of the pittsburgh sleep quality index. Journal of psychosomatic research 45, 1 (1998), 5–13.
[7] Trinder, J., Kleiman, J., Carrington, M., Smith, S., Breen, S., Tan, N., and Kim, Y. Autonomic activity during human sleep as a function of time and sleep stage. Journal of sleep research 10, 4 (2001), 253–264.
[8] Towards Circadian Computing: "Early to Bed and Early to Rise" Makes Some of Us Unhealthy and Sleep Deprived

[9]Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ,The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research
[10]Weixi Gu, Zheng Yang, Longfei Shangguan, Wei Sun, Kun Jin, Yunhao Liu, School of Software and TNList, Tsinghua University, China ,Intelligent Sleep Stage Mining Service with Smartphones
[11] https://en.wikipedia.org/wiki/Actigraphy
[12] Choe, E., Lee, B., Kay, M., Pratt, W., Kientz, J., SleepTight: Low-burden, Self-monitoring Technology for Capturing and Reflecting on Sleep Behaviors
[13] Yin Bai, Bin Xu, Yuanchao Ma, Guodong Sun, Yu Zhao,Will You Have a Good Sleep Tonight Sleep Quality Prediction with Mobile Phone
[14]https://docs.google.com/forms/d/1VE4Ljw7VLW5YyMOPMCWMRdfjyh3jpyQTHwSNBOyiJtg/viewform
[15] Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., and Pollak, C. The role of actigraphy in the study of sleep and circadian rhythms. american academy of sleep medicine review paper. *Sleep 26*, 3 (2003), 342–392.