9-2015

Candy crushing your sleep

Kasthuri JEYARAJAH
Singapore Management University, kasthurij.2014@phdis.smu.edu.sg

Meeralakshi RADHAKRISHNAN
Singapore Management University, meeralakshm.2014@phdis.smu.edu.sg

Steven C. H. HOI
Singapore Management University, CHHOI@smu.edu.sg

Archan MISRA
Singapore Management University, archanm@smu.edu.sg

DOI: https://doi.org/10.1145/2800835.2804393

Candy Crushing Your Sleep

Kasthuri Jayarajah
School of Information Systems
Singapore Management Univ.
kasthurij.2014@phdis.smu.edu.sg

Meera Radhakrishnan
School of Information Systems
Singapore Management Univ.
meeralakshm.2014@phdis.smu.edu.sg

Steven Hoi
School of Information Systems
Singapore Management Univ.
chhoi@smu.edu.sg

Archan Misra
School of Information Systems
Singapore Management Univ.
archanm@smu.edu.sg

UbiComp/ISWC '15 Adjunct, September 7-11, 2015, Osaka, Japan.
2015 ACM. ISBN 978-1-4503-3575-1/15/09$15.00
http://dx.doi.org/10.1145/2800835.2804393

Abstract
Growing interest in *quantified self* has led to the popularity of lifelogging applications. In particular, health and wellness related applications have seen an upsurge with the advent of wearables such as the Fitbit. In this paper, we focus on the *quality of sleep* that directly impacts the overall wellness of individuals. In particular, in this work, we present a first of its kind study that (1) unobtrusively quantifies the quality of sleep and (2) seeks to identify attributing aspects of our daily lives such as an individual's usage of apps throughout the day and his/her physical environment that may affect sleep. We use real life, in-situ smartphone data from 400+ undergraduate students over an observation period of 15 months, and present our initial observations.

Introduction
The availability of a multitude of sensors and applications have enabled smartphones to be more than mere communication devices. Lifelogging applications such as BeWell [8] have become increasingly popular amongst users in that they objectively account for the user's activities and behavior (e.g., sleep duration, walk duration, etc.) and provide insights into the user's life, and in certain cases, recommend behavioral and lifestyle changes.

The quality of sleep of individuals, over short and long term, have direct health and wellness implications [1]. Further, prior studies [12, 13] have shown strong correlation between the quality of sleep and the contexts an individual is subjected to over the course of the day. Such contexts include, but are not limited to, the physical environment of the user, his/her interactions with social groups, and applications a user accesses (which relates to the content a user consumes). The availability of sensing and logging on the smartphone allows us to unobtrusively monitor such contexts, and understand which of these contexts may affect the quality of sleep. An equally interesting problem is of identifying if the amount or quality of sleep a person gets the prior night affects his behavior the next day.

In this work, we address the former problem and make the following contributions:

1. provide a set of heuristics for quantifying sleep duration and disturbances during sleep for deriving quality of sleep based on the PSQI index [4],
2. present a first of its kind study to understand the effects of App usage on sleep considering both the category of the App and the temporal effects (e.g., recency of usage), and
3. present preliminary evidence that users' App usage and other physical contexts can indeed influence the quality of sleep.

In the sections to follow, we provide a brief summary of related work, describe the methodology and the dataset used followed by our analyses. We conclude with a discussion on the limitations of our current study and plans for future work.

Related Work

In this section, we briefly describe recent works in monitoring and prediction of sleep quality, in the Ubicomp domain.

Sleep Quality Monitoring: Prior work in this area have focused on detecting and capturing sleep disturbances using sensors on the smartphone such as the microphone (to sense sleep disturbances such as coughing and snoring) [9, 7, 5], light (to sense darkness) [9, 7, 5], and the accelerometer (for movement of the phone indicating the user's awaken state) [9]. Further, in [5], the authors use the duration of phone usage to detect the onset of a user falling asleep.In StudentLife[14], they infer the bedtime, wake up time and sleep duration of college students from their phone sensor data and daily activities over a 10 week long period. Similar to [5], in our work, we use App usage-based heuristics to approximate the user's duration of sleep and disturbances. Contrary to these works, in addition to monitoring sleep quality, our work aims to understand factors that affect the quality of sleep. We use the PSQI scoring method to derive sleep quality indices similar to that of [7].

Sleep Quality Prediction: SleepMiner[2] seeks to predict sleep quality based on contextual information such as human postures, positions and ambience (sound and light). This work is the closest to our work, but differs in that we focus on user behavior such as his/her app usage, social interactions and physical environment such as work/home that may affect sleep quality. We believe that the two works are complementary. To the best of our knowledge, our work is the first of its kind to understand the effects of Apps used (content consumed) on sleep quality, unobtrusively.

Methodology

In this work, we pose the problem of labeling the quality of a sleep episode as a classification task. We hypothesize that the quality of sleep depends on the independent variables: (1) user's App usage, (2) physical environment, and (3) social interactions. In the following subsections, we describe our approach and the dataset used.

Overall Methodology

We expect the independent variables under consideration to affect sleep quality in a number of ways. For instance, the Apps used by the user during the day can be of different categories such as news, gaming or social Apps. More use of social Apps may mean that the user communicates with his friends/family more which is a sign of wellbeing. Also, one could reason that gaming Apps require more alertness from the user as opposed to a News or infotainment App. Further, the requirement for such alertness closer to sleeping time, may or may not influence sleep differently than such requirements over the entire day. Similarly, we suspect that environmental changes, for example, whether a user spent more or less time at work may influence sleep at different levels. Also, the interactions a user has with his peers/friends in the physical world may also impact sleep quality – for example, more time spent with a friend in the evening may elevate the user's happiness level thereby resulting in a more comfortable sleep. We go through the following key steps:

Step 1: Feature Extraction – For the three independent variables, we further identify features (as listed in Table 1) and extract features for each (user, day) pair.

Step 2: Feature Selection – We perform a principal component analysis to identify which among the features are useful in explaining the variance observed and for reduction of dimensionality.

Step 3: Classification – We classify the quality of an episode i of user u as $Q_{i,u} \in S$ where S is the set of discrete quality levels. $S = \{Good, Bad\}$ and $S = \{VeryGood, Good, ModeratelyGood, Moderately Poor, Poor, VeryPoor\}$ for the cases of binary and multi-class classification, respectively. We use a J4.8 classifier (decision tree) for this purpose as our primary goal is to which features are associated with changes in sleep quality. We consider the episodes at a *daily* granularity, i.e., behavior over day i affects quality of sleep for day i, and the evaluations in this work do not account for cumulative effects. We defer this for future work.

Feature Class	Features
App Usage	Total Usage
	Usage by Category(5)
	Usage 1 hour prior to sleep
	Usage 4 hours prior to sleep
Physical Environment	Work (in-campus) time
	Outside-class time
	group-context outside-class

Table 1: Feature class and corresponding features extracted

Dataset Description

We use in-situ, continuous smartphone data from over 400 users over a period of 15 months (January 2014 to March 2015), who are opt-in participants of LiveLabs [10], as part of a larger, live testbed effort. The participants are students of the Singapore Management University of whom 54% are males, 34% are females and 12% are unknown. Among these 14% are freshmen, 16% are sophomores, 20% are junior students and 50% are senior students. The students belong to six different schools – Social Science (9.8%), Economics (11.42%), Accountancy (13.76%), Business (32.98%), Law (5.71%) and Information Systems (25.97%). In total, across all users,

Feature Class	No. of Observations
Sleep Quality	26,718
App Usage	26,718
Physical Environment	37,772

Table 2: Granularity of dataset

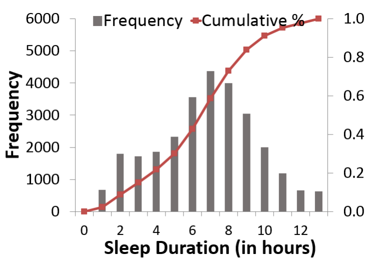

Figure 1: Distribution of Sleep Duration

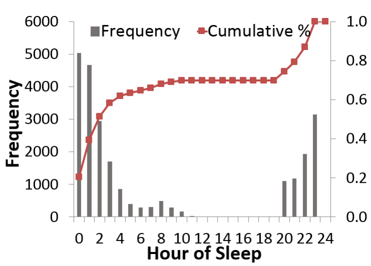

Figure 2: Distribution of Time of Sleep

there are records of 26,718 nights of sleep (see Table 2).

Feature Extraction

Sleep Quality

The Pittsburgh Sleep Quality Index (PSQI) is a cumulative score of 7 components as described in [4]. From accessibility logs from the smartphone, we deduce three of the seven components, namely, (1) sleep duration, (2) duration of disturbance during sleep and (3) the overall sleep efficiency, which we use to arrive at the sleep quality score of each sleep episode.

From the accessibility logs of the phones, we infer the longest periods of inactivity as representative of the approximate time a user was asleep. Further, short bursts of usage, during a period of inactivity is considered as disturbances during the sleep. For example, a user who's asleep could receive a notification through the Messaging App from a friend, and the user wakes up to respond. In addition, we also consider the time user spends in snoozing and resetting the alarm before finally waking up, also as a form of sleep disturbance. We describe below our heuristics in detail.

Computing duration of sleep: The longest period of inactivity during late evening to morning hours (of the following day) is likely the duration for which the user is asleep. However, we observe in our data that users tend to interact with their phones for short durations during this period resulting m chunks of reasonably long periods of inactivity separated by short periods of activity, T_{gap} (e.g., 15 minutes). For our purpose, we consider the top-3 $(m = 3)$ longest periods of inactivity, and merge two adjoining periods as a single *continuous* inactivity period, T_{sleep}, if the time separation between the pair of inactivity periods is less than or equal to 10 minutes $(T_{gap} = 10)$. As such, the total separation time becomes the time for which the user's sleep was *disturbed*, $T_{disturbance}$, whose maximum is 2 X T_{gap}. We plot the distribution of the duration of sleep episodes in Figure 1 where the x-axis is the duration in hours and the y-axis show the frequency and CDF. We observe that at least 60% of the episodes were of duration that is less than or equal to 7 hours. Further, we also infer the time at which the user went to sleep and woke up as the start time and end time of T_{sleep}, respectively. Figure 2 shows the distribution of the hour of the day at which users went to bed, as observed in our data.

Computing alarm snooze time: We observe instances of the Clock App being used multiple times during morning hours suggestive of the users' habit of snoozing (and resetting) the alarm consecutively (the details of how the App usage details are extracted is outlined in a later section). We consider *snoozing* as part of *disturbed* sleep. Among the 13,255 instances of "alarm between 3 AM to noon (pertaining to 306 users), for 9239 of the instances (70%), the user woke up after the first time the alarm set off. Out of the remaining 30%, 48% of the instances had a total "snooze" time of at least 30 minutes. We observed that over 47.5% of the users had 20% or higher fraction of individual sleeping instances that were *disturbed* (See Figure 3).

Computing sleep quality score: The PSQI rates each of the seven components on a scale from 0 to 3 where a higher score indicates poorer quality of sleep. We compute the overall sleep efficiency as the ratio of duration of sleep to the sum of sleep duration and sleep disturbance. To arrive at the final sleep quality score, we sum the three individual scores and scale by a factor of 7/3. As PSQI recommends, we use a quality threshold of 5 to label individual sleep episodes as *Good vs. Poor*.

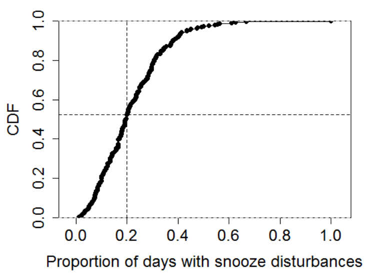

Figure 3: CDF of proportion of disturbed sleep episodes for individual users.

App Usage
In this sub-section, we describe how we extracted app usage related features from raw accessibility logs and screen ON/OFF times. For each $(day, user)$ pair for which we have a sleep time record, we consider the apps the user had used from the time he/she woke up the previous day till the time he/she slept, and not necessarily according to the calendar definition of a "day". To illustrate, for example, if a user had woken up on 1st January, 2014 at 8 AM, and went to bed on 2nd January, 2014 at 1 AM, then the usage of apps is considered between these two time points (and not the app usage during 1st January, 2014 alone) for $day =$1st January, 2014. As such, the extraction of app usage has a dependency on the "time to bed" feature from the previous sub-section.

We consider usage over three time windows: (1) the entire day ($totalusage$), (2) four hours prior to sleeping ($4hbefore$), and (3) one hour prior to sleeping ($1hbefore$). We hypothesize that the effect of app usage on sleep quality, if any, would be more significant as it is closer to sleep time.

For each window, we further segment the usage by category of the App. We hypothesize that certain categories of Apps (e.g., Games) would have a more significant effect on sleep quality than others (e.g., Lifestyle apps).

Extracting app usage duration: The LiveLabs database consists of two data sources that we used to infer the individual App usage durations. First, the "Profile State" dataset consists of changes of device profile which captures 17 different state changes to the device. This includes timestamped state changes such as device display ON ($state = 5$) and device display OFF ($state = 6$) among others. By collecting pairs of display ON and OFF times, we infer the times for which a user's device was "awake".

Second, during the times for which the device was "awake", we consult the "Accessibility" dataset to infer durations of App usage. The "Accessibility" dataset consists of timestamped events such as when a user touches the screen, types in a text box, clicks a button, etc [1]. In effect, this is a dataset consisting of the user's active interactions with the device (and does not consider Apps and services that run in the background such as Email sync). We batch consecutive interactions with the same App together, and consider the difference between the first and the last interaction as the duration of use for that App.

Extracting app category: The accessibility logs contain the name of the App, its package name, and the class associated with the accessibility event. However, this does not provide any labeling of the category of the App. To this end, we wrote a Python-based crawler to scrape the "genre" of the app based on its package name, off the Google PlayStore webpage, for a total of 3600+ unique Apps used by the LiveLabs participants. For example, for an App with package name *com.rovio.angrybirdsrio*, the crawler downloads the page `https://play.google.com/store/apps/details?id=com.rovio.angrybirdsrio`, and the html tag *span* with attribute *itemprop* equal to *genre* is scraped using the Beautiful Soup library. The PlayStore categorizes Apps as one of 26 general categories or 18 game categories.

As most users do not use Apps belonging to all 44

[1] `http://developer.android.com/reference/android/accessibilityservice/AccessibilityService.html`

categories, we grouped the App categories into ten representative categories. For the analyses presented in the remainder of the paper, we use only the four-most popular categories amongst our participants which are, namely, *Reading*, *Social*, *Entertainment*, *Information* and *Games*.

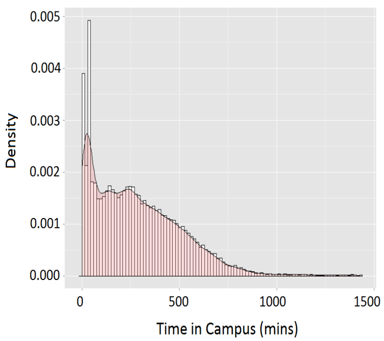

Figure 4: Distribution of Time in Campus

Physical Environment

To understand the effects of changes in the user's physical environment on sleep quality, we identified and considered two physical aspects of our daily lives: (1) working and (2) socializing. In the context of students, we proxy working as the time they spend in-campus. As our observation period spans 15 months, we have data from both session time (Jan to April and Aug to Nov) and break time (May to July and December). Likewise, we proxy socializing as the amount of time the students spent with their friends/peers outside class hours.

Work time: As part of the LiveLabs testbed effort [10], an indoor localization system has been operational campus-wide since August, 2013. This captures the location on-campus of any device that connects to the campus WiFi, with a location accuracy of 6 - 8 meters, approximately every 2-3 minutes. With this level of granularity users can be localized to logical sections within the buildings with reasonable accuracy. A *section* is a semantic location such as a seminar room or meeting room. In total, there are 167 such *sections* across the campus.

For each user for whom we have sleep quality and App usage data, we infer the time spent in campus as the difference in time between the first and last record of indoor location for each day. Here, we hypothesize that people spending lesser time in campus (or spending less time at work) results in better quality of sleep. In Figure 4, we plot the density (y-axis) of the amount of time students spent on-campus (x-axis) with a bandwidth parameter set to 15 minutes. As this contains both session and break times, we observe two peaks: one occurs during break times near zero, and the other around 6 hours during session times. Interestingly, we also observe that this distribution has a non-negligible long tail suggesting that there are times when students stay on-campus for very long times.

Social time: We use the state-of-the-art group detection system, *GruMon*[11], to extract durations and locations at which the students spent time with friends/peers. We segment the groups as *small* (2-3 students), *medium* (4-7 people) and *large* (greater than 7). In particular, we look at the periods for which the students were alone or in groups (of various sizes), *outside of class hours*. Class hours and non-class hours are understood based on the semantics of the location – i.e., if a student is in a seminar room or class room, then the student is assumed to be *in-class*. We hypothesize that students having more social time would have a much better sleep quality than those spending most of the time alone in campus.

Feature Selection

We performed a principal component analysis (PCA) to understand, out of the 24 app usage and physical environment features, which features were important in explaining the variance to both (1) reduce dimensionality and (2) select features that were important to the analysis.

To this end, we applied PCA on the App usage features and physical environment features, separately.

App usage features: We first perform PCA on the six total usage features: (1) total app usage time, (2) total social app usage time, (3) total reading app usage time,

	Proportion of variance explained
PC1	0.34
PC2	0.17
PC3	0.17
PC4	0.16
PC5	0.16
PC6	0.0

Table 3: Proportion of variance explained by principal components for total app usage features.

(4) total game app usage time, (5) total entertainment app usage time, and (6) total information app usage time. We specified a value of six for the number of components (assuming each feature is independent) and observed the proportion of variance explained by each factor. Table 3 lists the proportion of variance explained by each of the 6 orthogonal components. We observe that only five of the components are needed to explain the variance in entirety, and the proportions explained by the 5 components (specified as PC1 through PC5) are significant.

Further, we observed the factor loadings of the 6 features after varimax based rotation (See Table 4). Interestingly, we observe that the total app usage and social app usage features load heavily (0.85 and 0.99, respectively) on the first component, and the remaining 4 features each load to heavily (1.0) on separate components. We believe that the reason for both total and social app usage features to load heavily on the same component is because app usage is heavily biased by social app usage – most students use social apps heavily in comparison with other categories of apps. This is consistent with previous studies on app usage [3].

We repeated the same for the $4hbefore$ and $1hbefore$ features, separately, and found that the results from the PCA analysis were consistent with the above finding. Henceforth, we dropped the total app usage feature from feature set (resulting in 5 features per time window).

Physical environment features: We performed PCA on the six physical environment features: (1) time in-campus, (2) time outside class, (3) time alone outside class, (4) time in small groups outside class, (5) time in medium-sized groups outside class and (6) time in large groups outside class. We observed that only four orthogonal components can explain more than 98% of the variance. In Table 5, we list the factor loadings of the features on the four components (RC1 through RC4). Interestingly, we observe that the four features related to the social context load heavily on four orthogonal components. We also note that the time in campus and time outside classes load on the same factor as time alone outside classes. Henceforth, we club these three features together and call them collectively as "time in campus".

	RC1	RC2	RC3	RC4	RC5
Total app usage	**0.85**	0.37	0.27	0.19	0.17
Social app usage	**0.99**	-0.08	-0.04	-0.03	-0.05
Reading app usage	0.04	0.00	0.01	0.00	**1.00**
Game app usage	0.08	**1.00**	-0.01	0.00	-0.01
Entertainment app usage	0.08	0.00	**1.00**	0.00	0.00
Information app usage	0.06	0.00	0.00	**1.00**	0.00

Table 4: Factor loadings of total app usage features.

	RC1	RC2	RC3	RC4
Time in campus	**0.54**	0.44	0.34	0.30
Time outside classes	**0.63**	0.51	0.40	0.40
Time alone outside classes	**0.97**	0.19	0.12	0.07
Time in small groups outside classes	0.23	**0.95**	0.15	0.09
Time in medium-sized groups outside classes	0.16	0.16	**0.97**	0.09
Time in large groups outside classes	0.11	0.10	0.09	**0.98**

Table 5: Factor loadings of physical environment features.

Experimental Evaluation

Here, we present results from our preliminary evaluations for the cases of binary and multi-class classification. We normalize duration measurements of features (App usage

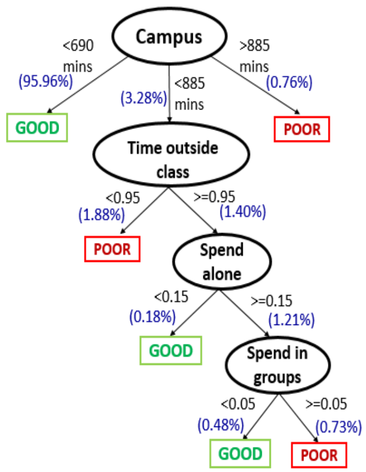

Figure 5: Sleep Quality vs. Physical Environment

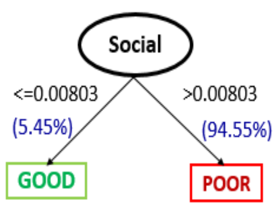

Figure 6: Overall App Usage by Category

and physical environment) by total usage or total time spent in campus, respectively, prior to classification.

Binary classification: For our experimental evaluation we used the J4.8 decision tree classifier in Weka [2] to determine which among the set of features would be correlated with sleep quality by classifying into two different classes 'Good' and 'Poor'. We ran the classifier separately on the four sets of features: (1) physical environment features, (2) App usage features over the day, (3) App usage features 4 hours before sleep, and (4) App usage features 1 hour before sleep.

Figure 5 shows the decision tree obtained after running the J4.8 classifier with physical environment features. We make the following observations:

1. Spending fewer hours in the campus is associated with good quality sleep (recall of 0.981); we note that students generally spend long hours in campus towards the end of the semester working on projects and studying for exams which could be indicative of stress in students.

2. In students staying longer than 11hours, we observe poor quality of sleep if they spend most of their time (95%) studying in classes/seminar rooms. Further, in cases where students spent more than 95% of their time outside classes, those who spent their time alone for less than 15% of the time were associated with good quality sleep.

3. In general, students who spent extremely long hours in campus (e.g., $\geq$ 885 mins) were observe to have suffered from bad quality of sleep.

Figures 6,7 and 8 illustrate the decision trees for the overall, $4hbefore$ and $1hbefore$ cases, respectively. We observe that (1) certain categories of Apps positively influence quality of sleep, and (2) using Apps close to bed time (e.g., $1hbefore$) adversely affects the quality, as hypothesized. We make the following observations:

1. Low usage of social Apps (e.g., less than 1% of the time) shows association with good quality of sleep (recall of 0.998); one reason could be that those who spend less time on social Apps may actually be those who spend more time socializing in the physical world, maintaining healthy relationships.

2. The Gaming, Reading and Social Apps tend to show positive or negative association with sleep quality depending on the proportion of time spent and the time of usage.

3. Although, using social Apps over the entire day shows negative association with sleep quality, we observe that more usage correlates with good sleep quality when used within 4 hours before sleep.

4. App usage 4 hours and 1 hour prior to sleep has contradicting effects on sleep quality. Using Reading and Gaming Apps within 1 hour to bed correlates with poor quality of sleep whereas using them slightly earlier correlates with good quality sleep. This finding is consistent with recent work in psychology [6].

In Table 6, we summarize the precision and recall values of classifying "Good" quality of sleep. Overall, we observe very high recall and average precision.

[2] http://www.cs.waikato.ac.nz/ml/weka/

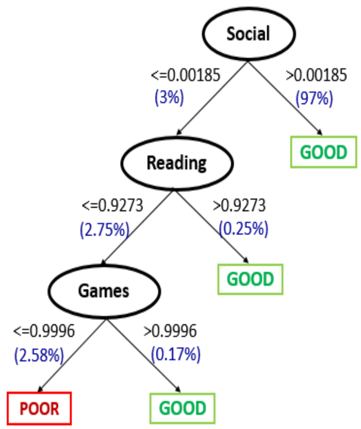

Figure 7: App Usage Within 4 hours before sleep

Figure 8: App Usage Within 1 hour before sleep

	Precision	Recall	F-Measure
Physical Environment	0.607	0.981	0.75
Overall App Usage	0.636	0.998	0.777
$1hbefore$	0.603	1	0.752
$4hbefore$	0.609	0.983	0.752

Table 6: Precision, Recall, F-measure values of *Good* sleep quality classification.

Threats to validity: Other than the contexts considered here, it is reasonable to assume that more obvious sleep-related factors such as the time a person goes to bed may affect sleep quality. We compute the deviation in time to bed, for each user, each day, from his *mean time to bed* and regressed over sleep quality. However, we observed that the deviation in time to bed only explains 0.215% of the variance in sleep quality.

Multi-class classification: Based on the final sleep scores (0,2.3, 4.67, 7.0, 9.33 and 16.33) that we obtained, we quantize sleep quality into 6 levels. In all the cases, the size of the decision tree increased and the overall F-measure value decreased. But for the classes 'Good' and 'Moderately Poor', the F-measure was higher compared to other classes and thus, suggesting that the decision tree could better classify sleep quality for these two classes.

Discussion & Conclusion

We briefly describe the limitations in the current study and our plans for future.

Limitations: One of the key advantages of our methodology is the ability to understand sleep quality *unobtrusively*. However, this also suffers from a number of drawbacks that can affect the validity of our analysis. The measurement of sleep time is based on the assumption that users check their phones before going to sleep and check their phones as soon as they wake up. This assumption may be less accurate for demographics other than students. Another drawback is that sleep disturbances can be measured only if a user interacts with the phone when he wakes up in the middle of the night.

Future work: Currently, we aggregate the sleep episodes across all users. However, a more precise evaluation would be person-centric as each individual is likely to be fundamentally different. We plan to perform person-centric analyses to identify which factors affect sleep quality for each individual user. Further, other factors such as the amount of physical activity a user gets during the day, the places a user visits outside of home and work, and socializing are known to influence the positive affect of people and hence may contribute towards better sleep quality. We intend to include these factors in future work. Preliminary analyses on the secondary problem, whether sleep quality of prior night affects an individual's behavior the following day, were inconclusive. However, we intend to investigate the problem further.

Concluding remarks: In this work, we presented the first of its kind study, to the best of our knowledge, on understanding what factors affect sleep quality, from unobtrusive measurements. We focused on the influence of App usage and physical environment changes and identified whether the different aspects of daily life positively (or otherwise) correlate with quality of sleep. Although the analyses and results presented are preliminary, we believe that this work has the potential to open up new research questions in the space of personal health, quantified self, ubiquitous computing, and machine learning.

Acknowledgements

This material is based on research sponsored in part by the Air Force Research Laboratory, under agreement number FA2386-141-0002, and partially supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office, Media Development Authority (MDA). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

References

[1] Alvarez, G. G., and Ayas, N. T. The impact of daily sleep duration on health: a review of the literature. *Progress in cardiovascular nursing 19*, 2 (2004), 56–59.

[2] Bai, Y., Xu, B., Ma, Y., Sun, G., and Zhao, Y. Will you have a good sleep tonight?: sleep quality prediction with mobile phone. In *Proc. Body Area Networks* (2012).

[3] Böhmer, M., Hecht, B., Schöning, J., Krüger, A., and Bauer, G. Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In *Proc. HCI with mobile devices and services* (2011).

[4] Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., and Kupfer, D. J. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry research 28*, 2 (1989), 193–213.

[5] Chen, Z., Lin, M., Chen, F., Lane, N. D., Cardone, G., Wang, R., Li, T., Chen, Y., Choudhury, T., and Campbell, A. T. Unobtrusive sleep monitoring using smartphones. In *Proc. PrevasiveHealth* (2013).

[6] Fossum, I. N., Nordnes, L. T., Storemark, S. S., Bjorvatn, B., and Pallesen, S. The association between use of electronic media in bed before going to sleep and insomnia symptoms, daytime sleepiness, morningness, and chronotype. *Behavioral sleep medicine 12*, 5 (2014), 343–357.

[7] Hao, T., Xing, G., and Zhou, G. isleep: unobtrusive sleep quality monitoring using smartphones. In *Proc. Sensys* (2013).

[8] Lane, N. D., Mohammod, M., Lin, M., Yang, X., Lu, H., Ali, S., Doryab, A., Berke, E., Choudhury, T., and Campbell, A. Bewell: A smartphone application to monitor, model and promote wellbeing. In *Proc. PrevasiveHealth* (2011), 23–26.

[9] Min, J.-K., Doryab, A., Wiese, J., Amini, S., Zimmerman, J., and Hong, J. I. Toss'n'turn: smartphone as sleep and sleep quality detector. In *Proc. Human factors in computing systems* (2014).

[10] Misra, A., and Balan, R. K. Livelabs: Initial reflections on building a large-scale mobile behavioral experimentation testbed. *ACM SIGMOBILE Mobile Computing and Communications Review 17*, 4 (2013), 47–59.

[11] Sen, R., Lee, Y., Jayarajah, K., Misra, A., and Balan, R. K. Grumon: Fast and accurate group monitoring for heterogeneous urban spaces. In *Proc. Sensys* (2014).

[12] Tanaka, H., Taira, K., Arakawa, M., Masuda, A., Yamamoto, Y., Komoda, Y., Kadegaru, H., and Shirakawa, S. An examination of sleep health, lifestyle and mental health in junior high school students. *Psychiatry and clinical neurosciences 56*, 3 (2002), 235–236.

[13] Uezu, E., Taira, K., Tanaka, H., Arakawa, M., Urasakii, C., Toguchi, H., Yamamoto, Y., Hamakawa, E., and Shirakawa, S. Survey of sleep-health and lifestyle of the elderly in okinawa. *Psychiatry and clinical neurosciences 54*, 3 (2000), 311–313.

[14] Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., and Campbell, A. T. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proc. Ubicomp* (2014).