Smartphone-Based Intelligent Sleep Monitoring

Pansheng Fang[1,2](✉), Zhaolong Ning[3], and Xiping Hu[1]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
{ps.fang,xp.hu}@siat.ac.cn

[2] School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China

[3] School of Software, Dalian University of Technology, Dalian 116620, China
zhaolongning@dlut.edu.cn

Abstract. The sleeping quality is one of the most important factors to judge people's health status, and has drawn increasing attention of the public recently. However, the quantified results of sleeping quality can generally be achieved in labs with the help of high precision instrument, such as Actigraphy or professional graph like Polysomnography (PSG), and are thus not available for the general public. In this paper, we construct a novel way of sleep-scoring system implanted in the iSmile app. iSmile first collects the sounds recorded by smart phone recorder, then classifies the sound frames with a light weight decision tree algorithm. Based on the number and the average amplitude of sleep-related events, we score the users' sleeping quality in three aspects (respectively cough-score, snore-score and talk-score) using Pittsburgh Sleep Quality Index (PSQI) and Pediatric Sleep Questionnaire (PSQ). During users' sleeping period, iSmile also collects data from the accelerator sensor to predict the users' mood (presented in valence and arousal) and recommend smart alarm sounds to help improve their mood. For the experiment, we involved 5 participants (20 nights in total) and achieved high precision of predicting sleep events (above 89%), with the users' valence and arousal improved by 14.57%. From succinct chart of sleeping score on the App UI, users can see the visualized results of their sleeping quality.

Keywords: Sleep scoring · Sleep event detection · Microphone

1 Introduction

Sleeping plays a significant role in human's live for human beings spend up to one-third of their lifetime sleeping. The effect of sleep related problems is wide and profound. According to the statistic provided by the Institute of Medicine (US), about 50 to 70% of US citizens suffer from chronic sleep problems [1]. Moreover, insufficient sleep can result in fatigue and frequent sleepiness at daytime, such as nodding off while driving and accidentally falling asleep at work [1, 2]. Clinical studies also reveal that sleep has something to do with many severe diseases such as diabetes, obesity and

V. C. M. Leung et al. (Eds.): 5GWN 2019, LNICST 278, pp. 43–59, 2019.
https://doi.org/10.1007/978-3-030-17513-9_4

depression [2, 3]. Obviously, the sleeping-related problems deserves more attention from the society, and people need to view the visualized judgement of their sleeping quality.

People also need to know about the sleeping events that reducing their sleeping qualities. The research in [12] indicates that snoring loudly can be a signal of sleep apnea. Frequent Coughing and snoring can be the symptom of Chronic Obstructive Pulmonary Disease (COPD) [13], and sleep-talking is relevant to poor sleeping quality [14, 15].

Though the public have started to focus on their sleeping quality, and specific sleeping events, they are not accessible to visualized results of sleeping quality, which are only achievable in labs through high precision instruments to analyze electrical signals with Polysomnography (PSG) [5, 6, 9]. Normal citizens also lack robust self-judging standards. The Pittsburgh Sleep Quality Index (PSQI) [7] and Pediatric Sleep Questionnaire (PSQ) [8] provide workable self-scoring standards, but some of the items in the questionnaires (such as “How often have you had trouble because you cough or snore loudly during the past month?” and “Select the frequency of snoring”) are actually not workable when users live alone. Even though the users have a sleeping partner, the self-judgement can be not accurate because people measure the loudness and the frequency. Moreover, wearable devices such as Actigraphy, are always utilized to detect the valence and arousal of users. However, these devices bring unnecessary costs and uncomfortable feelings to users wearing them [12].

The meteoric adoption of smartphones places a rich sensor platform in the pockets, purses, and backpacks of many people. Interestingly, many people choose to use their phones as an alarm clock, placing these sensors in proximity of the bed. A recent study by the Pew Internet and American Life project found that 44% of mobile phone owners (83% of teens) sleep with their phones on or near their beds [10]. This advance in technology and associated change in behavior offers the possibility of cheaply and effectively tracking people’s daily sleep behaviors without the need for additional hardwares or for a significant change in behavior.

In this paper, we presented iSmile, a workable system to score people’s sleeping quality automatically with the detection of sleeping events using smart phone microphones, and recommend a smart alarm to improve users’ mood. The App is easy to use: users just need to set the alarming time when they go to bed, and place the mobile phone somewhere near their heads, and the App will start to collect data for scoring. The App records the sound from the built-in microphone of mobile phones, extracts the features, and detects sleeping-related events such as coughing, snoring and sleep-talking. Simultaneously, the App will collect data from the accelerator sensors of mobile phones to monitor the movement rate during the users’ sleeping period. Before the alarm sounds, raw data can be uploaded to the cloud platform, and after the sever handles the raw data, a recommended alarm will be played at the mobile phones and a chart showing users’ sleeping quality can be seen on the App UI.

When it comes to the designing of the system, a series of problems may occur. First of all, we need to distinguish different events correctly. Luckily, the work of [11] has proved that sleeping events of different people have similar features in variance, RMS (root mean square value) and energy spectrum distribution. Secondly, we need to find ways to eliminate the influence brought by the ambient noise. Lastly, we need to link the data collected with reliable scoring standards to finish the visualization of the score. The framework of the whole system will be introduced later in the third part.

2 Related Work

The system design is inspired by some existing sleeping monitor apps (such as Sleep miner [19]) and smart alarm app (such as Sleep cycle [16] and Snail sleep [17]). The system presented in EAST [20] gives a method to extract features from accelerator sensors of mobile phones to predict users' moods when they wake up. ISleep [11] presents long-term sleep scoring based on the PSQI (Pittsburgh Sleep Quality Index), which is a widely-used subjective sleep quality assessing questionnaire. It also provides short-time sleep quality scoring ways, but only with the help of Artigraphy.

In addition to the PSQI, PSQ (Pediatric Sleep Questionnaire) also provide ways to assess sleeping quality from the aspect of snore/cough loudness and rate. To classify the severity of snore and other sleeping events [21, 22]. Authors in [21] show that the snoring intensity is measured as low at 90 $\pm$ 4 dB at 500 Hz, and high at 96 $\pm$ 4 dB at 500 Hz. The authors in [22] use the mean maximum decibel level to classify snoring as mild (40–50 dB), moderate (50–60 dB), or severe (>60 dB), and concludes that louder snoring can be an indication of more severe Obstructive Sleep Apnea (OSA). To handle other sleeping events like coughing and sleep-talking, we can apply the same way according to their average intensity and numbers.

The accurate detection should be based on robust knowledge of the acoustics of sleeping events. Authors in [23] did a research on 16 subjects and found that snoring frequency ranges from 137 Hz to 1243 Hz according to 200 samples. The author in [24] provides theoretical possibilities to distinguish snoring sound according to its acoustics indicating that snoring frequency is below 2000 Hz with peak power usually below 500 Hz.

We construct the iSmile smart scoring system making good use of the acoustic features of sleeping events, and manage to classify sleeping events such as snoring, coughing and sleep talking with an accuracy above 90%. iSmile uses our original scoring standard based on PSQI and PSQ to score users' sleeping quality nightly in three aspects: snore, cough and sleep talks. We also deploy smart alarm sounds designs to improve the mood of users.

3 System Overview

iSmile has combined the microphone sensing data part with the acceleration data processing part in the previous version of iSmile app, which improves the performance in distinguishing body movements from other sleep-related events.

The overall architecture of iSmile is shown in Fig. 1. iSmile is divided into two parts: the newly-implemented acoustics processing part and the accelerator signal processing part, which is accomplished in previous versions of iSmile [20]. iSmile read continuous acoustic signals at 16 kHz from mobile phone built-in microphones, and senses the accelerator sensor to acquire acceleration in x, y z direction.

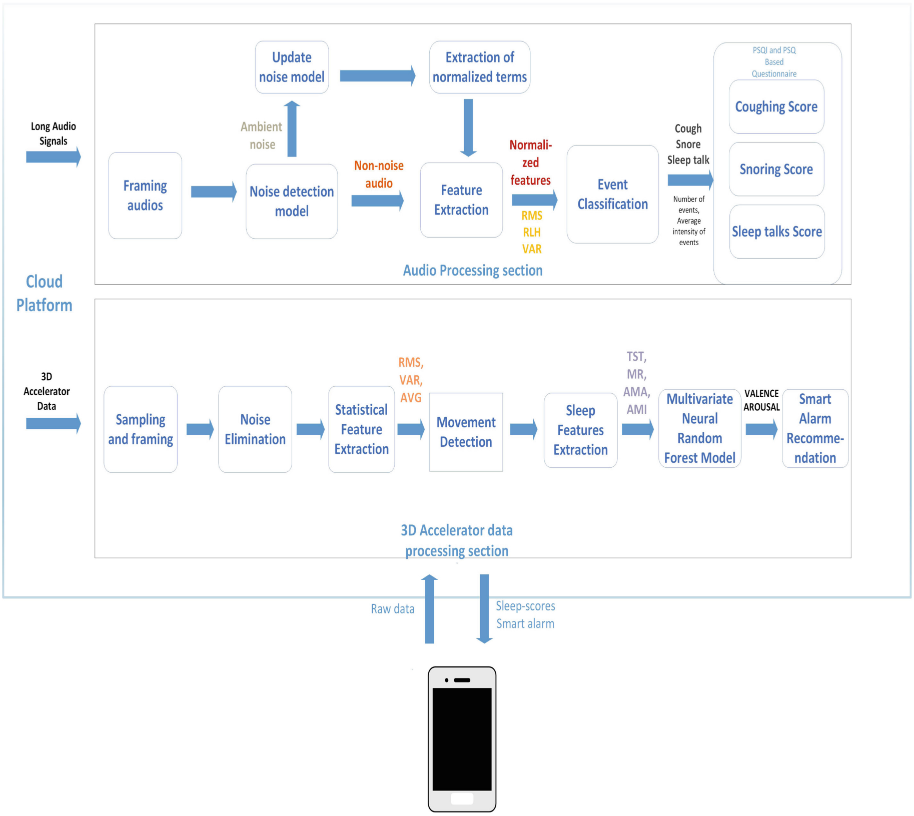


Fig. 1. The overall architecture of iSmile.

In the next step, raw data will be uploaded to the cloud platform. Then the audios are segmented into short pieces of about 5 s and sent to the acoustics part, the acceleration data will be processed in the other part.

In the acoustic processing part, the short audios are first fragmented into frames (20–50 ms for each). To reduce the energy loss of the edge of each frame, we apply a hamming window for each frame. They are fed to a noise detection fragment. Next, three features of the audio will be extracted from the non-noise frames to detect sleeping events. The noise audios will be used to update the noise detection model. After all the audios are processed, iSmile will score the users' sleeping quality according to the number and the average intensity of each sleeping event.

In the acceleration data processing part, four features will be extracted from the raw accelerator data. We apply an algorithm named neural random forest defined in [25] to predict the valence and arousal, representing the intrinsic attractiveness and measures how calming or exciting the information is respectively, of the users, and use a KNN-based algorithm to get recommended alarm sounds.

4 System Design

In this section, we introduce the structure of the audio processing part of iSmile in details. First of all, we present the method to detect ambient noise and the way to update the noise and the noise model. Then we introduce the features to classify different sleeping events. Afterwards we discuss the classification of sleeping events, and the measures to calculate normalized features terms. At last we will introduce the standard of scoring users' sleeping quality.

4.1 Ambient Noise Detection

Apparently, the ambient noise generated by various sources will affect the detection of sleeping events. At first, we need to distinguish the ambient noise from all the recorded audios. From our perspective, the ambient noise refers to the noise of the recording devices, the appliances that work continuously such as fans and air-conditioners, and other noise generated by outside environment such as rains, and winds.

The work of [11] indicates that ambient noises differentiate from sleeping events with relatively low variance of normalized standard deviation of each frame defined as follows:

$$\overline{std} = \frac{std_i - std_{mean}}{std_{mean} - std_{min}} \tag{1}$$

where std_i refers to the STD of i^{th} frame of a noise audio piece which captures the stability of an acoustic signal frame, while std_{mean} and std_{min} denotes the mean and min standard deviation of a T-second acoustic signal window.

The authors in [11] also indicate that the value 0.5 can be set as the threshold to judge whether an audio piece is an ambient noise. However, our observation of more than 18,000 labeled noise audios shows that in extremely silent environment, even a very small fluctuation of recorded audio can cause the dramatical rise of variance (shown in Fig. 2). In this case, mis-classification of ambient noises will happen frequently. In order to fix this problem, we make some improvements in the noise detection model. Based on our observation of noise audios' features, we find that the intensity (shown by RMS) is much lower than sleeping events (shown in Fig. 3). So, we add RMS as another feature in noise detection model. From Fig. 3, we can see that more than 95% of the ambient noise frames have a RMS below 15.0, so 15.0 can be the threshold of judging whether audio pieces are ambient noises. In other words, if a 5-second frame is detected to have a $\overline{std}$ lower than 0.5 or RMS lower than 15.0, it will be recognized as ambient noise. With this method, we exclude the disturbance of slight fluctuation of sounds in silent environment.

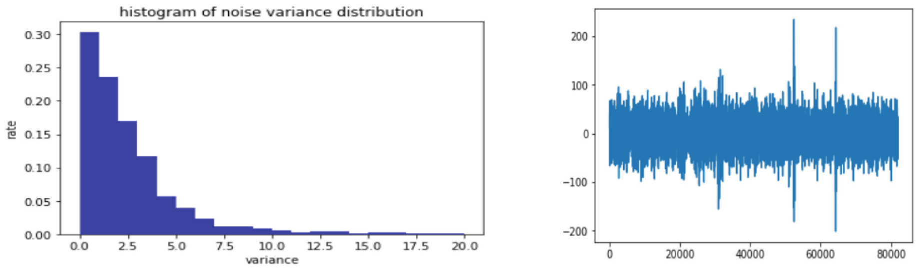


Fig. 2. a. Histogram of noise variance distribution. b. Noise waveform in silent environment

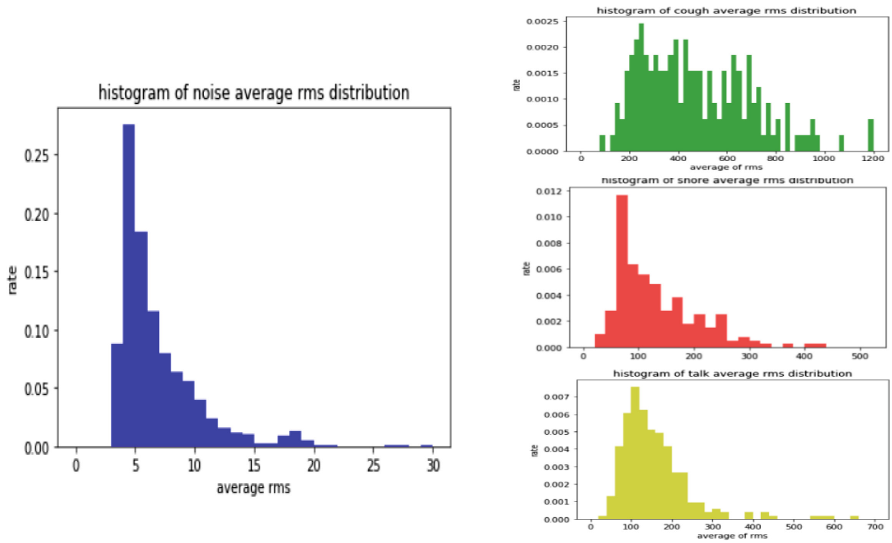


Fig. 3. The distribution of $\overline{\mathrm{RMS}}$ of ambient noise (left) and sleep-related events (right).

4.2 Sleeping Event Features

- To classify the sleeping events, we choose three features to predict sleeping events based on the key characteristics of each sleep-related events.

The first one is root mean square (RMS), which presents the intensity of a non-noise audio piece. Suppose f is an acoustic piece that consists of n samples of acoustic amplitude $a_1, a_2, a_3, \ldots, a_n$. iSmile segments hours-long audios into 5-second-long frames, and each of them contains 80000 acoustic samples collected at 16 kHz. The RMS of the short audios frame f is given by:

$$\mathrm{RMS}(f) = \sqrt{\frac{a_1^2 + a_2^2 + a_3^2 + \ldots + a_n^2}{n}} \tag{2}$$

The second feature is based on the observation of power spectrum of different sleeping events. Using the FFT algorithm [26], we can have a convenient observation of the energy distribution of an audio piece. The study in [27] indicates that snoring has scattered energy content in the higher spectral sub-bands (>500 Hz.), and we believe that sleep-related events such as coughing and sleep talking have their unique features in terms of spectrum. From Fig. 4, we can find that three kinds of sleeping events differ a lot from each other in their spectrum. To be specific, although all of their peak power at around 200 Hz, coughing has the highest peak energy of more than 1400, while the peak energy of snoring and sleep talking are respectively approximately 60 and 160. Moreover, snoring has the most scattered sub-peaks distribution ranging from 1200 Hz to around 5000 Hz in the spectrum, and in the spectrum of talking only several sub-peaks with small intensity appear. In terms of coughing's energy spectrum, sub-peaks appear at 3000 Hz and 3800 Hz with considerable intensity. In our implantation, we define the ratio of low-frequency energies to high-frequency energies (RLH) as the second feature. The RLH of a non-noise audio f can be calculated as follows:

$$\mathrm{RLH}(f) = \frac{RMS(f_{low})}{RMS\left(f_{high}\right)}, \tag{3}$$

where $RMS(f_{low})$ and $RMS\left(f_{high}\right)$ denote the intensity of the energy of low frequency part and high frequency part of an audio frame respectively. We set 1200 Hz as the dividing point of high frequency and low frequency, and from Fig. 4 we can infer that snoring audios should have the lowest RLH while talking audios should have the highest one. In order words, the decreasing in RLH might be an indication of snoring events.

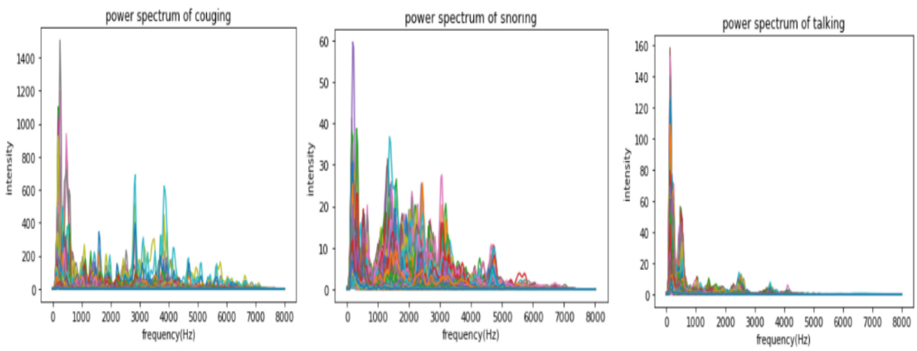


Fig. 4. The power spectrum of different sleep-related events

The third feature is variance (VAR), which reflects how far the amplitudes of acoustic signals within the frame are spread out. For instance, coughing is caused by the sudden and dramatic vibration of the vocal cord, so coughing audios should have higher variance than those of snoring and talking, which are both produced by relatively periodic vocal cord vibration.

4.3 Normalizing Measures

After an audio piece is detected as ambient noises, iSmile calculates three features for each of the frames. In order to estimate the current noise, iSmile computes the mean and standard deviation (*mean(RMS)*, *std(RMS)*, *mean(RLH)*, *std(RLH)*, *mean(VAR)* and *std(VAR)*) for each feature as the normalized terms. Then, each newly calculated distribution feature $(Feature_{new})$ will be used to update the current corresponding feature $(Feature_{current})$ according to an Exponential Moving Average (EMA) algorithm as follows:

$$Feature_{current} = Feature_{cureent} + \beta \times (Feature_{new} - Feature_{current}) \tag{4}$$

In iSmile, the parameter βis set to be 0.4. The EMA algorithm makes sure that the normalized terms can adapt itself automatically in different environments.

4.4 Classification of Sleeping Events

In Sect. 4.2, we defined RMS, RLH, VAR as three features to distinguish different sleep-related events. To improve the robustness of the classification and reduce the influence of different external environments, we will normalize the three features with the Feature terms defined in Sect. 4.3 to acquire the features vectors (FV: $(\overline{RMS}, \overline{RLH}, \overline{VAR})$) for classification. For example, $\overline{rms}$ of a non-noise audio is calculated in the following way:

$$\overline{RMS}(f) = \frac{RMS(f) - std(f)}{mean(f)} \tag{5}$$

Similarly, $\overline{RLH}$, *and* $\overline{VAR}$ are calculated in the same way.

Figure 5 shows the distribution of three normalized features. It is plotted from the 761 labeled data collected from 5 subjects using Meizu Note 5 during a one-week experiment. Specially, the class “*others*” refers to the sounds other than three sleep-related events, such as body moving (Fig. 6).

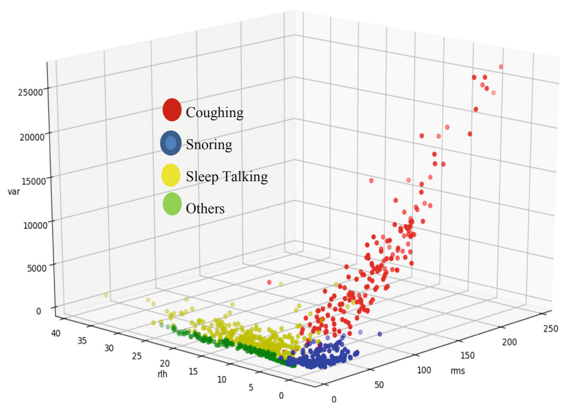


Fig. 5. The sound feature vectors of different events in feature space.

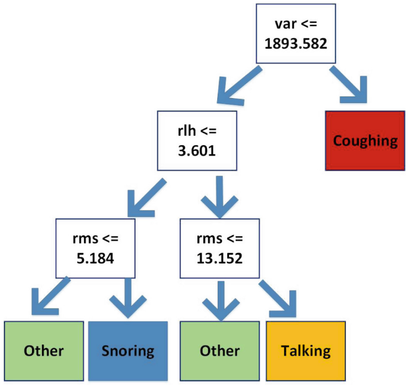


Fig. 6. The decision tree for event detection.

We apply a decision tree algorithm [28] as the classification model to detect sleeping events. The dotted rectangles indicate the splitting features, and the leaf nodes denote the classification results. The splitting features and thresholds are determined based on the information gain calculated by entropy. Specifically, the entropy of a node T is given by

$$\mathrm{Entropy}(\mathrm{T}) = -\sum\nolimits_{j} p(j) \cdot \log(p(j)), \tag{5}$$

where $p(j)$ is the relative rate of class j at node T. In iSmile's case, $j = 3$. After splitting node T into k nodes $(T_1, T_2, \ldots T_k)$, we can calculate the information gain by

$$\mathrm{G} = \mathrm{Entropy}(\mathrm{T}) - \left[\sum\nolimits_{j=1}^{k} \frac{n_j}{n} \mathrm{Entropy}(T_j)\right], \tag{6}$$

where n_j denotes the number of samples in node T_j. For each splitting, the system chooses the split that maximizes the information gain.

Next, we will describe the classification process in details. First, the non-noise audios are split into two groups according to $\overline{VAR}$, which reflects the extent of stability of acoustic signal. As a result, apparently audios with high $\overline{VAR}$ can be detected as coughing.

Then, the audios in the low $\overline{VAR}$ group are further split into two groups based on $\overline{RLH}$. As introduced in Sect. 4.2, low $\overline{RLH}$ reflects high dominant frequency, so low $\overline{RLH}$ audios are mostly classified as snoring and only a small number of audios in this group are classified as other noises because they have low $\overline{RMS}$. On the other hand, audios in high $\overline{RLH}$ group can be further categorized into two groups according to $\overline{RMS}$. While the low $\overline{RMS}$ audios are detected as other noise, the high $\overline{rms}$ audios are judged to be sleep-talking.

4.5 Scoring Users' Sleeping Quality

To realize the quantification of sleeping quality, we design a light-weight sleeping quality scoring standard based on two widely accepted criteria, one is PSQI, which provides the items of sleep-related events, and the other is PSQ, which gives the idea to scoring the quality from the angle of the intensity and appearing frequency of sleeping events. Table 1 lists the metrics in PSQI and PSQ that iSmile uses to calculate the score of sleeping quality. In our design, sleeping quality scores will be calculated from two aspects for different sleeping events. One is the event-intensity part, which shows whether the users snore, talk or cough loudly when sleeping, and the other is the event-rate part, which indicate whether users suffer from these problems frequently at nights. After the calculation of each of the two aspects is finished, we will add the scores of two parts and multiply the sum by 0.5 to acquire the scores for an event. Table 2 lists the items in iSmile's scoring items in details.

Table 1. Metrics from PSQI and PSQ iSmile uses

PSQI B4	Cough or snore loudly
PSQI B3	Cannot breathe comfortably
PSQ A2-a3	Usually/always snores
PSQ A4	Snore loudly

Table 2. The scoring criteria of iSmile

Scoring aspects	Detailed aspect	Rate
Coughing	The intensity of coughing	50%
	The frequency of coughing	50%
Snoring	The intensity of snoring	50%
	The frequency of snoring	50%
Sleep taking	The intensity of sleep talking	50%
	The frequency of sleep talking	50%

In terms of the event-intensity part, from 761 labeled sleeping events collected from 5 subjects, we can see the distribution of average RMS of each sleeping events. The results shown in Fig. 4 indicates that the average RMS of more than 95% of coughing is above 200, and less than 5% of that is above 800. If the average intensity of a coughing event is below 200, it will be scored as 100, which means the user snores slightly, while the ones whose intensity is between 200 to 800 receive score at 66.7, the remaining received only 33.3 points because it means users are coughing very loudly. Similarly, for snoring events, the average RMS below 50 will be scored at 100, and the remaining at 50. For talking, the average RMS below 120 and the remaining will be scored at 100 and 50, respectively.

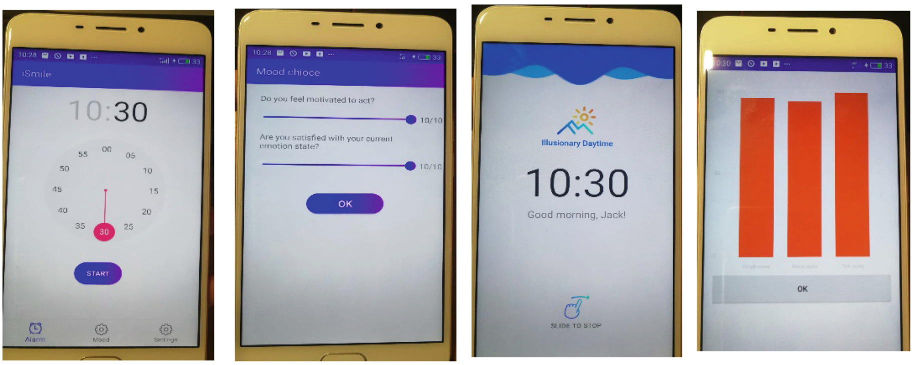
For event appearing frequency scoring, if a 5-second-long audio pieces is detected as sleeping events, users will receive 50 points for the relative event items. If no events are detected, users will be scored at 100.

In our implementation, the system will record the numbers of audios detected to be sleep-related events (let it be *num*), and before the alarm sounds, all of the sleeping quality scores will be summed together and divided by *num* as the final results.

5 Implementation and Evaluation

5.1 Implement on Android Devices

ISmile is facilitated by an emotion-aware cognitive system [20], to demonstrate the system performance. Since iSmile needs to collect raw data from the built-in microphone and accelerator sensor continuously. But for some reasons, nearly all of the Android devices manufactures will shutdown the process to save the battery life. We have tried hard to address this problem, including setting the Developer Mode and adding iSmile to the device whitelist. We even try the Doze model of the Android operating system [29], a new feature for Android 6.0, which reduces battery consumption by deferring background CPU and network activity for apps when the device is unused for long periods of time. Eventually we select Meizu Note 5 to conduct the experiments. The UI of iSmile is shown in Fig. 7.



Setting alarms Filling mood state Alarm sounds Showing

Fig. 7. Android UI of iSmile

The iSmile app takes up around 14 MB to install. When users run the iSmile app, they first enter the time setting page, where they can set the alarming time. After they press the *start* button, the system will start to collect record audios and acceleration data continuously. Before the alarm sounds, the raw data will be uploaded to the cloud platform, where the backend process will calculate the raw data to get the recommended music and the users' sleeping scores. Then on the Android side, recommended music will be played and users can see a chart showing their sleeping scores of coughing, snoring and sleep-talking. Users can record their current mood states, which will be uploaded to the backend process for later evaluations.

5.2 Experiments and Evaluation

In the previous version of iSmile, we have finished the design of smart-alarm recommendation. In that part, 8 subjects (three females and five males) were involved for ten days. 4 of the 8 subjects, labeled as control group, were given random alarms, and the other four, set as test group, were given the recommended alarms. Their self judgements of moods will be recorded every day, when they wake up. The results (shown in Figs. 8 and 9) reveals that the valence and arousal of the second group is raised by $\sqrt{10.33\% \times 10.33\% + 10.27\% \times 10.27\%} = 14.57\%$ compared with the first group.

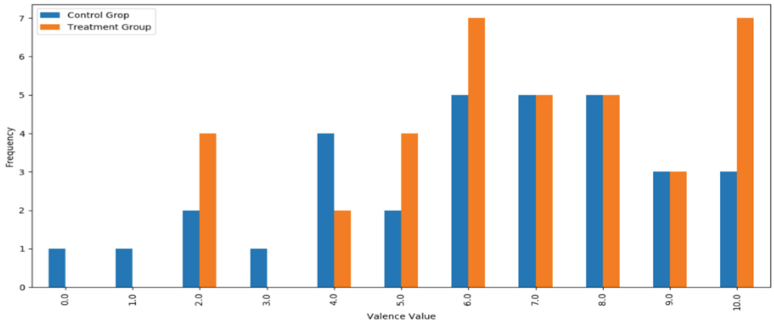


Fig. 8. Comparison results of valence values between the treatment group and the control group. (The higher, the better.)

To evaluate our newly-implanted audio processing part, we include 5 subjects (all males) and record audio of 20 nights in total. All of the 5 subjects are required to run the app and set an alarming time every night before they sleep. Later we listen to the audios recorded, and manually judge whether the system make the right detection of sleeping events. Table 3 shows the event detection results of data collected in the form of Audio Detection Accuracy (ADA).

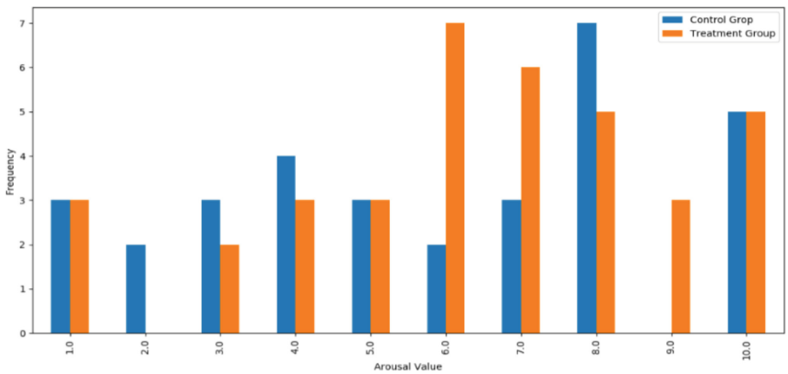


Fig. 9. Comparison results of arousal values between the treatment group and the control group. (The higher, the better.)

Table 3. The detection accuracy in the experiment

Subject	Coughing (ADA)	Snoring (ADA)	Sleep-talking (ADA)
1	6/6 (100%)	7/7 (100%)	17/28 (60.7%)
2	5/5 (100%)	5/5 (100%)	12/17 (70.6%)
3	8/8 (100%)	6/6 (100%)	4/6 (66.7%)
4	5/5 (100%)	4/5 (80%)	6/6 (100%)
5	3/3 (100%)	6/6 (100%)	3/3 (100%)
Total	27/27 (100%)	28/29 (96.6%)	42/60 (70%)

As shown in Table 3, over 20 nights of sleeping, iSmile has achieved high accuracy detection of over 95% in terms of coughing and snoring. Only 1 snoring audio is misclassified as other sleeping events. The relatively low detection accuracy of sleep-talking is explainable, since people can't really control themselves when they are sleep-talking. Audio signals of sleep-talking differ a lot from each other even though they are all produced by the same individual. Given more time, we could collect more audio data of sleep talking in order to increase its detection accuracy. If the forth class (other sounds) are included in the calculation of ADA, iSmile actually achieved an accuracy of 89.19% (132/148).

The scores of 5 subjects in 4 nights are presented in Figs. 10, 11 and 12. From the results, we can infer that all of the 5 subjects have different sleeping problems. For example, subject 5 coughs a bit more than the other four subjects. Except subject 2, the other four participants all have slight snoring problems. From the degree of sleep-talking, subject 1 receives the lowest scores, indicating that his brain seems to behave more actively at nights.

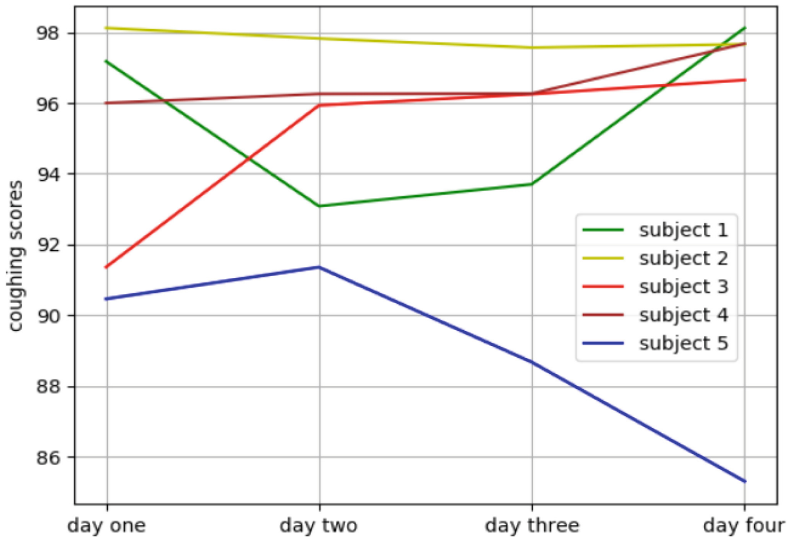


Fig. 10. Coughing scores of 5 subjects.

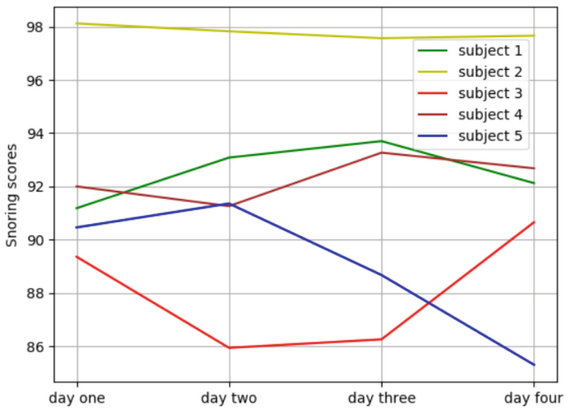


Fig. 11. Snoring scores of 5 subjects

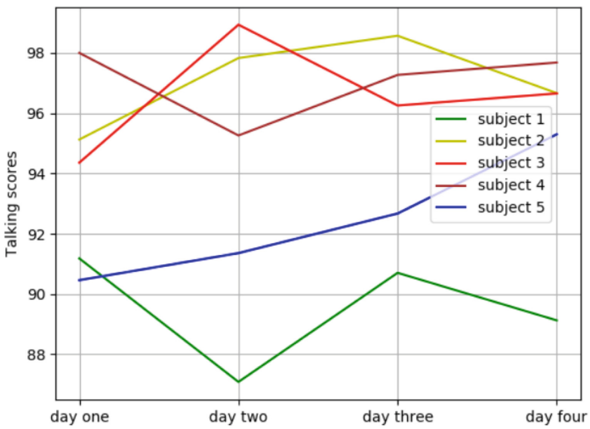


Fig. 12. Sleep-talking scores of 5 subjects

6 Conclusions and Future Work

In this paper, we introduce an original sleeping quality scoring app, named iSmile, which samples audios at 16 kHz from the smart phone built-in microphones to detect sleep-related events and scoring. ISmile also collect the data from smart alarm recommendations. To use this app, users just simply set the alarming time before they sleep, and the system will start to work. Before the alarm sounds, the raw data will be uploaded to the cloud platform. The backend program will calculate the sleeping scores and the recommended alarm sounds to improve the users' mood when they wake up. In the previous version of iSmile, we have achieved a 14.57% increase in users' mood in the form of valence and arousal. For the newly-deployed audio processing part, we have achieved a high accuracy detection of sleeping events for 5 subjects in 20 nights in total, and given the sleeping scores to show each individual's sleeping quality related to coughing, snoring as well as sleep-talking.

However, there are still some problems for us to handle. On one hand, due to the fact that backend program is deployed at the cloud platform, we need to upload all the raw data to the cloud site for later processing. In that case, there can be a huge data flow using this app. In other words, iSmile is not recommended for users without WI-FI currently. On the other hand, although we have achieved nearly 100% accuracy in detecting coughing and snoring, the ADA of sleep talking are much lower, which does great harm to the reliability of the final results of sleeping scores. What's more, the iSmile app has not been tested in many Android devices, but its stability still remains unknown. In the future, we will put much effort to address these problems and increase the system robustness. For example, we can deploy the processing part on the Android site to reduce the data flow, and we should collect more data of sleep talking to increase the judging accuracy. Last but not least, iSmile should be tested on different Android devices so that its performance can be judged in a more appropriate way.

References

1. Institute of Medicine, Colten, H.R., Altevogt, B.M. (eds.): Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. The National Academies Press, Washington, DC (2006)
2. Parish, J.M.: Sleep-related problems in common medical conditions. Chest J. **135**(2), 563–572 (2009)
3. Foley, D., Ancoli-Israel, S., Britz, P., Walsh, J.: Sleep disturbances and chronic disease in older adults: results of the 2003 national sleep foundation sleep in America survey. J. Psychosom. Res. **56**, 497 (2004)
4. Gu, Y., Zhan, J., Liu, Z., Li, J., Ji, Y., Wang, X.: Sleepy: adaptive sleep monitoring from afar with commodity WiFi infrastructures. In: IEEE Wireless Communications and Networking Conference (WCNC), 15–18 April 2018
5. Douglas, N.J., Thomas, S., Jan, M.A.: Clinical value of polysomnography. Lancet **339** (8789), 347–350 (1992)
6. Divo, C.M., Selivanova, O., Mewes, T., Gosepath, J., Lippold, R., et al.: Polysomnography and ApneaGraph in patients with sleep-related breathing disorders. ORL J. Otorhinolaryngol. Relat. Spec. **71**, 27–31 (2009)
7. Buysse, D., Reynolds, C., Monk, T., Berman, S., Kupfer, D.: The Pittsburgh sleep quality index (PSQI): a new instrument for psychiatric research and practice. Psychiatry Res. **28**, 193–213 (1989)
8. Chervin, R.D., Hedger, K.M., Dillon, J.E., Pituch, K.J.: Pediatric Sleep Questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. Sleep Med. **1**, 21–32 (2000)
9. Chesson, A., et al.: Practice parameters for the indications for polysomnography and related procedures. Sleep **20**, 406–422 (1997)
10. Pew Internet and American Life Project. http://pewinternet.org. Accessed 5 Dec 2018
11. Hao, T., Xing, G., Zhou, G.: iSleep: unobtrusive sleep quality monitoring using smartphones. In: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, Roma, 11–15 November 2013
12. Maimon, N., Hanly, P.J.: Does snoring intensity correlate with the severity of obstructive sleep apnea? J. Clin. Sleep Med. **6**, 475–478 (2010). [PubMed]
13. Scharf, S.M., Maimon, N., Simon-Tuval, T., et al.: Sleep quality predicts quality of life in chronic obstructive pulmonary disease. Int. J. Chron. Obstruct. Pulmon. Dis. **6**, 1–12 (2010)
14. Veldi, M., Aluoja, A., Vasar, V.: Sleep quality and more common sleep-related problems in medical students. Sleep Med. **6**, 269–275 (2005)
15. Manni, R., et al.: Poor sleep in adolescents: a study of 869, 17-year old Italian secondary school students. J. Sleep Res. **6**, 44–49 (1997)
16. Northcube: Sleep cycle alarm clock (2017). https://www.sleepcycle.com/. Accessed 5 Dec 2018
17. Seblong: Snail sleep (2016). http://www.seblong.com. Accessed 5 Dec 2018
18. Rofouei, M., et al.: A non-invasive wearable neck-cuff system for real-time sleep monitoring. In: BSN (2011)
19. Bai, Y., Xu, B., Ma, Y., Sun, G., Zhao, Y.: Will you have a good sleep tonight? Sleep quality prediction with mobile phone. In: BodyNets (2012)
20. Hu, X., et al.: Emotion-aware cognitive system in multi-channel cognitive radio ad hoc networks. IEEE Commun. Mag. **56**(4), 180–187 (2018)

21. Lamont, J., Baldwin, D.R., Hay, K.D., Veale, A.G.: Effect of two types of mandibular advancement splints on snoring and obstructive sleep apnoea. Eur. J. Orthod. **20**, 293–297 (1998)
22. Maimon, N., Hanly, P.J.: Does snoring intensity correlate with the severity of obstructive sleep apnea? J. Clin. Sleep Med. **6**(5), 475–478 (2010)
23. Agrawal, S., Stone, P., McGuinness, K., Morris, J., Camilleri, A.E.: Sound frequency analysis and the site of snoring in natural and induced sleep. Clin. Otolaryngol. Allied Sci. **27**(3), 162–166 (2002)
24. Perez-Padilla, J.R., Slawinski, E., Difrancesco, L.M., Feige, R.R., Remmers, J.E., Whitelaw, W.A.: Characteristics of the snoring noise in patients with and without occlusive sleep apnea. Am. Rev. Respir. Dis. **147**, 635–644 (1993)
25. Welbl, J.: Casting random forests as artificial neural networks (and profiting from it). In: Jiang, X., Hornegger, J., Koch, R. (eds.) GCPR 2014. LNCS, vol. 8753, pp. 765–771. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11752-2_66
26. Ganapathiraju, A., Hamaker, J., Picone, J., Skjellum, A.: A comparative analysis of FFT algorithms
27. Pevernagie, D., Aarts, R.M., De Meyer, M.: The acoustics of snoring. Sleep Med. Rev. **14**, 131–144 (2010)
28. Mitchell, T.: Machine Learning. McGraw-Hill, Boston (1997)
29. Sohil, A., Rastogi, S., Chawla, R.: Analysis for performance optimization of android applications. Int. J. Comput. Sci. Eng. Technol. **8**(3), 65–71 (2017)