Classifying obstructive sleep apnea using smartphones

Mamoun Al-Mardini [a,*], Fadi Aloul [a], Assim Sagahyroon [a], Luai Al-Husseini [b]

[a] American University of Sharjah, Sharjah, United Arab Emirates
[b] Arab German Sleep Medicine, Amman, Jordan

ARTICLE INFO

ABSTRACT

Obstructive sleep apnea (OSA) is a serious sleep disorder which is characterized by frequent obstruction of the upper airway, often resulting in oxygen desaturation. The serious negative impact of OSA on human health makes monitoring and diagnosing it a necessity. Currently, polysomnography is considered the gold standard for diagnosing OSA, which requires an expensive attended overnight stay at a hospital with considerable wiring between the human body and the system. In this paper, we implement a reliable, comfortable, inexpensive, and easily available portable device that allows users to apply the OSA test at home without the need for attended overnight tests. The design takes advantage of a smatrphone's built-in sensors, pervasiveness, computational capabilities, and user-friendly interface to screen OSA. We use three main sensors to extract physiological signals from patients which are (1) an oximeter to measure the oxygen level, (2) a microphone to record the respiratory effort, and (3) an accelerometer to detect the body's movement. Finally, we examine our system's ability to screen the disease as compared to the gold standard by testing it on 15 samples. The results showed that 100% of patients were correctly identified as having the disease, and 85.7% of patients were correctly identified as not having the disease. These preliminary results demonstrate the effectiveness of the developed system when compared to the gold standard and emphasize the important role of smartphones in healthcare.

1. Introduction

Obstructive sleep apnea (OSA) is a common disorder affecting 2–4% of the adult population [1], and it is considered to be the most prevalent sleep disorder. There are three types of sleep apnea, namely: central, obstructive, and mixed sleep apnea. Central Sleep Apnea (CSA) occurs when the brain sporadically fails to trigger the breathing muscles in the chest, while obstructive sleep apnea (OSA) occurs when air is physically blocked from flowing into lungs during sleep intermittently. Mixed Sleep Apnea (MSA) is a combination of both central and obstructive sleep apnea [2]. Of the three, OSA is the most common while CSA and MSA are significantly rarer. In most cases, the patient is unaware of the breath stoppages because the body does not trigger a full awakening. As per the American Academy of Sleep Medicine (AASM), OSA is defined as a cessation in the airflow lasting for more than two breaths [2].

The National Sleep Foundation (NSF) reported that for adults to function healthily they should obtain seven to eight sleeping hours every night [3]. Frequent obstructions of airflow during this period have a considerable influence on the performance of the human during the daytime. OSA causes excessive sleepiness, non-restorative sleep, high blood pressure, cardiovascular diseases, memory loss problems, erectile dysfunction, personality changes, and depression [4,5]. Besides daytime tiredness, OSA patients may experience job impairment and automobile accidents.

Attended overnight Polysomnography (PSG) is considered the gold standard for OSA diagnosis. To get an overnight OSA test, the patient should stay in a specialized sleep laboratory for more than one night, with 22 wires attached to his/her body in order to record and analyze several neurologic and cardio respiratory signals. This brings great anxiety to the patients and they may not be able to sleep well during the night. Moreover, very few hospitals can accommodate the PSG test and it is rarely found in rural areas, which makes it unavailable for everyone and costly. Because of the aforementioned complications of using PSG, a need has arisen for portable devices with acceptable accuracy, high levels of usability, and depend on acquiring fewer biological signals [2,5–8].

Smartphone technology has become abundant among millions of people. At the same time, the availability and popularity of mobile health (mhealth) represented by health applications (Apps) have also increased. Mobile health applications are now used to

support the management of chronic diseases, such as: diabetes, heart disease, and hypertension. Additional examples of mobile health apps include helping people to avoid some bad habits that affect their health, such as: smoking cessation, body weight loss, and reducing alcohol consumption [9,10]. As the development of mHealth apps and the number of users of mHealth technologies increase, there is a need to understand the usability and learnability of these mobile devices and applications. For this reason, some studies focused on studying the usability and learnability of health apps on smartphones by applying theoretical models, such as: FITT (Fit between Individuals, Task, and Technology) [11] and Health-ITUEM (Health IT Usability Evaluation Model) [10]. Findings from these studies demonstrate the positive usability of smartphones in solving health with some differences in performance among devices.

In this paper, we focus on implementing a portable device that is inexpensive, reliable, and accurate when compared with PSG. The developed design makes use of the pervasiveness of smartphones by designing an Android application that is able to extract the biomedical readings from the patient and analyze them to screen and classify OSA. This screening is considered a preliminary test used to identify people at high risk of having the disease who may need further confirmatory diagnostic tests at a hospital [9].

In the developed system, we develop an algorithm that combines and analyzes the physiological readings to reliably infer if the patient suffers from OSA. This algorithm draws from conclusions made in published literature, experimental data and tests collected in this work, and finally the consultation of a physician expert who has been actively involved in this project. All readings are analyzed solely on the smartphone without the usage of any external resources. Finally, we examine our system's ability to screen the disease as compared to the gold standard. We performed the tests on 15 subjects, and the results showed that 100% of patients were correctly identified as having the disease, and 85.7% of patients were correctly identified as not having the disease.

The rest of this paper is organized as follows: Section 2 includes a literature review of related research. An overview of the pretest screening and its impact on the probability of having OSA is explained in Section 3. Methods for extracting the physiological signals will be outlined in Section 4. Section 5 explains the algorithms used to analyze the collected physiological signals. Section 6 presents the developed system architecture. Implementation of the application is explained in Section 7. Section 8 is a presentation and discussion of the results. Finally, Section 9 concludes this paper.

2. Related work

Rapid advances in technology have enhanced the capability of smartphones and added powerful features to them. Nowadays, smartphones play a significant role in different areas including communication, social, education, and healthcare. Kailas et al. [12] mentioned that there are already more than 7000 documented applications concerned with healthcare, and with the increase in deployment of mobile phones, the role of smartphones in healthcare applications is expected to become significantly more pronounced in the coming years. In this section, we first highlight the role of smartphones in healthcare by providing some application examples, followed by a review of some attempts related to the utilization of mobiles in sleep apnea related research. The key features that give smartphones the advantage over the existing healthcare systems include portability, ease of use, and availability among people. Furthermore, smartphones may be used when clinicians are far away from their patients, or in rural areas where computer is not applicable [8,9].

2.1. Smartphones in healthcare

To detect abnormal Cardiovascular Disease (CVD), Jin et al. [13] and Oresko et al. [14] developed a cell phone-based platform that continuously records the electrical activity of the heart over a period of time (ECG) and analyze it in real-time. The phone-based platform was successful in classifying heartbeats into five categories with a prediction accuracy of more than 90%. Moreover, on CVD, Chen et al. [15] highlighted the fact that Heart Rate (HR) and Heart Rate Variability (HRV) can be accurately assessed from acoustic recordings of heart sounds using only a cell phone and a hands-free kit. The authors have successfully developed a heart rate estimator (and heart sound locator) with a sensitivity of 92.1% and a positive predictivity of 88.4% for detecting heart sound.

Moving to respiratory diseases, Zhang et al. [16] developed a breathing bio-feedback prototype using a smartphone and a breathing sensor. Bio-feedback is considered as an alternative technique in which one can control his/her body functions to achieve the desired results, such as reducing pain. The system acquires breathing using a respiratory sensor, and transmits the signal to a smartphone using Bluetooth technology. The connected sensor provides information (feedback) for the body (bio), and the smartphone receives the breathing signals and generates a visual breathing feedback display to the user. Up to now, only a small range experiments have been done to test the performance of the prototype system.

Another study by Scully et al. [17] measured the respiratory rate using the embedded camera in the smartphone. The developed system used the smartphone to record and analyze the varying color signals of a fingertip placed in contact with the camera. The authors have shown that mobile phone cameras have the potential to measure respiratory rate with a very close results compared with the commercial products.

A recent study addressed the detection of Sudden Infant Death Syndrome (SIDS) [18]. This study used the smartphone's built-in microphone, accelerometers, and other sensors to monitor the infant's heartbeat and respiration. The developed application is adaptive in the sense it learns from its environment and fine tune its parameter accordingly.

2.2. Obstructive sleep apnea (OSA) portable devices

HealthGear [19] is one of the first Windows-based mobile applications that uses blood oximetry to detect sleep apnea. The application was tested on 20 participants and the OSA detection algorithms identified with 100% accuracy all 3 cases of known OSA. Cao et al. [20] also used an oximeter combined with an accelerometer in their research to diagnose sleep apnea and other pulmonary diseases. Their work found that body posture can provide complementary information for analyzing the respiratory movement. An experiment to recognize apnea and hypopnea was conducted and the proposed algorithm was applied to recognize apnea and hypopnea in real time. The apnea and hypopnea were identified and the number of the apnea/hypopnea episodes (AHI) was counted automatically. Alqassim et al. [21] presented a proof of concept article discussing the feasibility of using smartphones and their built-in sensors in detecting sleep apnea.

Rofouei et al. [22] extended the work done in [19,20] and developed a non-invasive, wearable neck-cuff system capable of real-time sleep monitoring and visualization of physiological signals. In this work, only one subject was compared against the gold standard, and it was correctly diagnosed.

The previous studies used threshold-based methods on the data collected by the biomedical sensors to diagnose OSA. On the other hand, some researchers applied data mining techniques, such as [6], where the authors followed data mining approaches to build

an accurate classifier able to detect sleep apnea in a real-time fashion. The technique was tested on 70 participants and an accuracy of 93% was achieved. However, Tseng et al. [23] developed an Android smartphone application to diagnose OSA based on some prediction rules derived by employing decision tree algorithms to a large clinical data set without using bio-sensors. The Application was tested on 540 participants and an accuracy of 96.85% was achieved.

Signal processing techniques have been also used in diagnosing OSA. The authors in [5] proposed an inexpensive OSA screening technique, where they developed an accelerometer-based system by placing an accelerometer on the suprasternal notch. They then analyzed the data using signal processing techniques implemented on a microcontroller. The algorithm was tested on a sample test data and was able to identify and classify the majority of the sleep-apnea test data. However, it requires further refinement and testing on actual sleep-apnea patients.

3. Pretest probability of obstructive sleep apnea

Reviews reported by health agencies have revealed that portable devices used to diagnose sleep disorders result in high relative error rates compared to Polysomnography (PSG). These error rates are either high false-negative (a result that wrongly indicates the absence of a disease) or false-positive (a result that wrongly indicates the presence of a disease). Due to the high false-negative and false positive ratios for portable devices that diagnose sleep apnea, it is recommended to apply a pretest on patients. In essence, the pretest is simply a questionnaire that includes some questions related to OSA symptoms. The responses give an indication for the possible existence of OSA and it is part of the overall diagnostic process. Only those with a high pretest probability for sleep apnea should use a portable device for the diagnostic. The importance of a high pretest probability becomes clearer if studies that are focused on the general population are compared with studies focused on clinical populations. A high pretest probability reduces the number of false-positive diagnoses [24,25]. Later in this paper we will show how the pretest is integrated in the proposed algorithm. Following is a list of symptoms collected from the complaints of obstructive sleep apnea (OSA) patients which are used to assess the possibility of having OSA:

- Loud and irregular snoring.
- Observed or reported nocturnal cessation of breathing.
- Excessive daytime sleepiness.
- Unspecific mental problems such as fatigue, low performance, cognitive impairment.
- Movement during sleep.
- Morning dizziness, general headache, dry mouth.
- Impaired sexual functions.
- Obesity.

4. Methods to extract physiological signals

In the developed system, we use the smartphone's microphone to extract respiratory effort by placing it on the throat (where the obstruction happens). An external microphone connected via a wire to the smartphone is used to record the respiratory efforts. Oxygen saturation level is usually captured by an external Bluetooth-based pulse oximeter, which measures the oxygen level in the blood. An oximeter is a sensor that is placed on a thin part of the patient's body, usually a fingertip or earlobe, or in the case of an infant, across a foot. Body posture and movement can be detected by continuously taking the directions of the body along all axes. An accelerometer can be used to detect these movements.

Researchers typically use an external accelerometer for this purpose. However, in the developed design, we take advantage of the built-in accelerometer in the smartphone. The smartphone will be placed on the patient's arm, and body movement will be detected by calculating the displacement in the three axes (X, Y, and Z).

5. Methods to analyze collected physiological signals

An apnea event is detected when there is a cessation of airflow for more than 2 breaths or at least 10 s [2]. Based on this definition, we need to detect the obstruction of the airflow, and the accompanied desaturation reflected in the blood. In the developed system, we analyze three physiological signals to screen the obstructive sleep apnea (OSA). The analysis for these signals will be explained in the following subsections.

5.1. Oxygen saturation

In the developed implementation, a pulse oximeter is used to measure oxygen level and provide continuous data transmission of a 4 byte data packet sent every second. In the developed system, we use the oxygen desaturation index (ODI) which is defined as the average number of events per hour [26]. An event is detected if the oxygen level is below the average by 4%, and lasts at least 10 s. The following steps are followed to detect an apnea event using SpO2 readings [19]:

- Read oxygen level every second.
- Initial calibration phase.
  - The level of oxygen differs between individuals, therefore we need to find the average in the first few minutes, and record it.
  - An event is detected if the oxygen level is 4% below the average calculated in the previous step. Therefore, a threshold is defined as $0.96*Average$. $X = 0.96 * N$, where X is the threshold and N is the average.
- Moving calibration phase:
  - After the occurrence of an apnea event, the oxygen saturation does not return to the same level just before the event. Therefore, we need to recalculate the average and the threshold again. $X\_New = 0.96 * N\_New$, where X_New is the new threshold and N_New is the new average.
- If the oxygen level is detected to be lower than the defined threshold, then we have to be sure that the event lasted at least 10 s.
- By the end of the sleeping time, ODI is calculated. OSA is confirmed if one of the following conditions is met:
  - ($ODI \geqslant 10$) or
  - ($ODI \geqslant 5$) && (high pretest probability).

Fig. 1 shows the oxygen level in the blood for one of our tested subjects, which is extracted from the oximeter every second. When the patient started to snore, a decrease in the oxygen saturation followed. After a few seconds, the oxygen saturation reached a level less than the threshold value that lasted for more than 10 s.

5.2. Body movement

Based on research reported by [22,27], body movement may cause variation in the pulse oximetry readings. Therefore, any change in the oxygen level that is associated with body movement is eliminated. In the developed application, the oxygen saturation and body movements are synchronized; both readings are
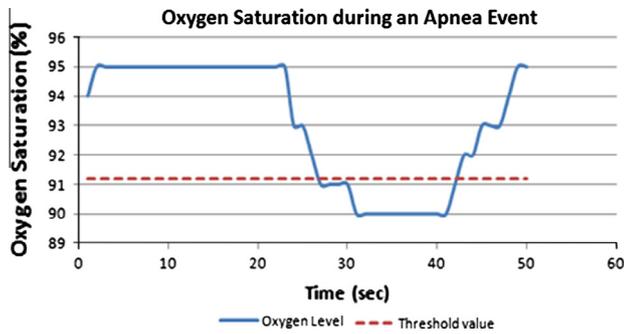
Fig. 1. Oxygen saturation during an apnea event.

recorded at the same time. This synchronization allows us to eliminate the oxygen saturation readings, when movements occur.

The 3-axis accelerometer integrated in the smartphone is used to measure the 3-axis accelerations of the subject body along the directions of the three axes. Body positions are categorized into two types: motion and rest. To classify motion and rest, we need to detect the displacement of the three axes (X, Y, and Z). We may consider the displacement of each axis separately to detect motion or use a concept called Signal Vector Magnitude (SVM) proposed in [20] which reflects the motion intensity of the body. SVM is the square root of the sum of the 3-axis accelerations, the amplitude of which mirrors the motion amount of the body. Therefore, rather than dealing with each axis separately to detect motion, we use the SVM concept to combine the values of all axes and detect the motion. In the developed system, we measure the SVM value of the body in different positions, and notice the difference between them. For example, the SVM value of the body when the patient is lying on his/her back ranges from $\{-1, 4\}$. However, the value ranges from $\{9, 15\}$ when the patient is laying on his/her side. Therefore, if the difference is more than 5, then a movement is detected.

5.3. Respiratory effort

The patient's airflow is recorded by attaching a microphone to the patient's throat. Frequency of breathing is different among people but experimentally has been proven to fall between 200 and 800 Hz [27]. Extracting this range of frequencies allows us to exclude noise and analyze the breathing signal only. Signal processing functions are required to analyze this range of frequencies. The correctness and accuracy of these functions are first verified using Matlab software running on an external server, and then ported to the smartphone Android environment. The following steps are used to detect apnea event using respiratory effort readings:

- Read respiratory signal every second.
- Initial calibration phase:
  - The energy of the breathing signal differs between individuals, so we need to find the average in the first few minutes, and record it.
  - An event is detected if the energy is below 90% of the average calculated in the previous step. Therefore, a threshold is defined as $0.9 * \text{Average}$. $X = 0.9 * N$, where X is the threshold and N is the average.
  - Threshold value can vary due to the differences in the devices being used and their sensitivity [1].
- If the energy of the breathing signal is detected to be lower than the defined threshold, then we have to be sure that the event lasts at least 10 s.

- By the end of the sleeping period, apnea/ hypopnea index (AHI) which is defined as the average number of events per hour [2] is calculated. OSA is confirmed if one of the following conditions is met:
  - ($AHI \geqslant 10$) or
  - ($AHI \geqslant 5$) && (high pretest probability).

Fig. 2 shows the energy of the respiratory effort signal for one of the subjects in our work during one minute time window. The spikes in the graph represent the snoring periods. As we can see, the snoring period is followed by an obstruction of the airway, where the energy values are less than the threshold value which is represented by the red line.

5.4. Determining the final diagnostic

At this stage, we combine the results of the analysis of the physiological signals (oxygen saturation, body movement, and respiratory effort). The analysis of the oxygen saturation and body movement is used in calculating the ODI index. On the other hand, AHI is computed using the respiratory effort analysis. Finally, we find the average of the calculated ODI and AHI and then make a decision for the existence of OSA as illustrated in the flowchart of Fig. 3. After finding the average, the severity of OSA is reported based on the following categories [1]:

- Mild: $5 \leqslant Average < 15$.
- Moderate: $15 \leqslant Average < 30$.
- Severe: $Average \geqslant 30$.

The collection and analysis of the three physiological signals allow us to find the influence of each physiological signal on the final diagnostic. Each signal can be used in determining the existence of an apnea event, therefore, it can be used to get the final diagnostic. In the developed system, we have determined the final diagnostic based on three proposed data sets, which are:

- Partial data set 1: based on oxygen saturation and body movement.
- Partial data set 2: based on respiratory effort.
- Complete data set: based on all physiological signals (oxygen saturation, body movement, and respiratory effort), i.e., a combination of partial data set 1 & 2.

6. System architecture

The developed system consists of four hardware components, which are: a smartphone as a central processing point, an oximeter to record oxygen saturation, a microphone to record respiratory
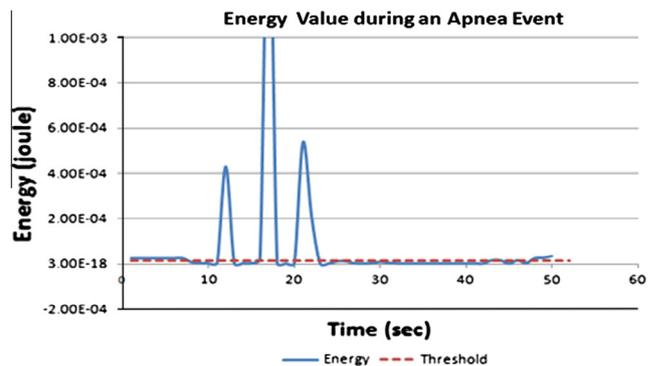

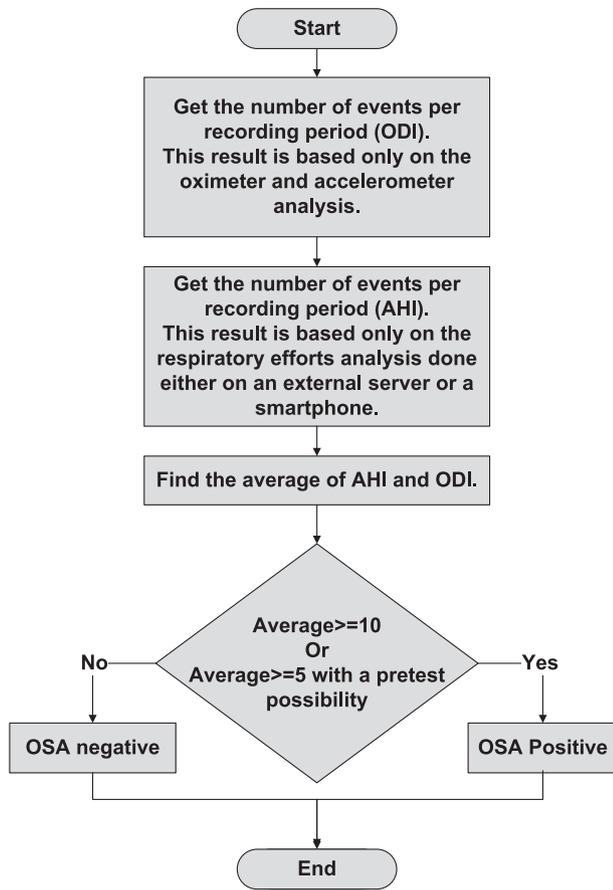**Fig. 2.** Energy value during an apnea event.

Fig. 3. A flow chart explaining the steps needed to determine the final diagnostic.

Fig. 4. System architecture.

Fig. 5. Main activities of the developed application.

effort, a built-in accelerometer to record body movement, and a Bluetooth interface that is used as a gateway between the smartphone and the oximeter. On the other hand, the system makes use of two software environments, which are: Android Software Development Kit (SDK) and Matlab. Fig. 4 shows the hardware components of the developed system.

7. Implementation

The developed application consists of four main activities: a DataCollector activity, a Pretest activity, a Diagnose activity, and a Report activity. An activity is an application component that provides a Graphical User Interface (GUI) with which a user can interact in order to trigger an action, such as dial the phone, send an email, or take a photo. An application usually consists of multiple activities that are loosely bound to each other [28]. Fig. 5 shows the main Activities of the application and how they interact with each other.

The DataCollector activity is considered to be the main activity in the application. This activity is the one presented to the user when the application is launched. It is responsible for starting the process of collecting the physiological signals from the patient. Pretest activity is simply a questionnaire that assesses the possibility of an individual having the disease. The answers are then forwarded to the Diagnose activity to be used in the Analysis process. The Diagnose activity is responsible for analyzing the readings collected in the DataCollector activity, and the information passed by the Pretest activity. The Report activity is the last page displayed to the user, in which all the related details are summarized.
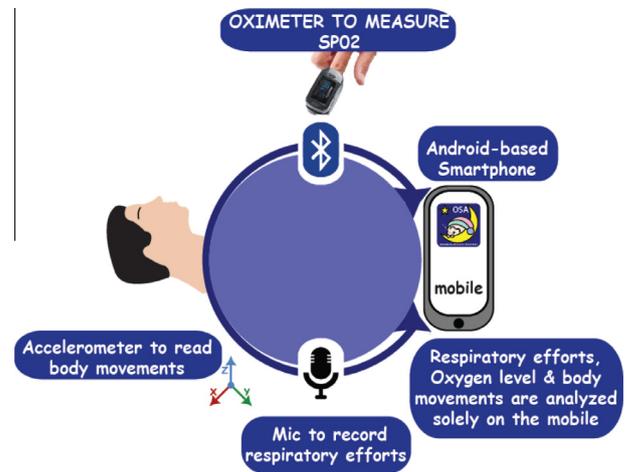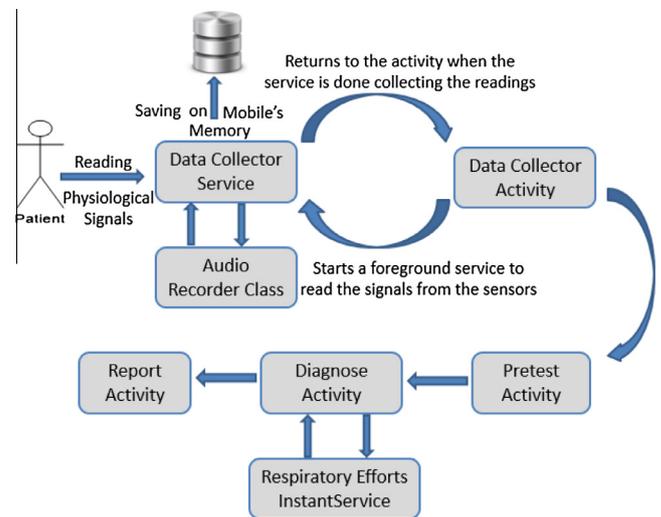
Processing the recorded respiratory effort requires filtering the signal and calculating the energy every second. The Android SDK does not contain a specific library for signal processing functions. Therefore, we have implemented the filtering functions in Android, in order to use them in extracting the required frequencies. Since the frequencies lie between 200 and 800 Hz, we have implemented and used a band pass Infinite Impulse Response (IIR) filter. Fig. 6 shows the respiratory effort for one of the subjects in our work during one hour. The spikes in the graph represent the snoring periods, and the dotted black line represents the average threshold value that is calculated during this hour.

Since the application must run for a long time, during which we use the I/O ports and memory excessively, we must take the smartphone's performance under consideration. Choosing the right algorithm and data structure should take high priority in order to improve performance and utilize the available resources. Following is a list of all points that we have taken under consideration in developing the application to enhance the performance and utilize the resources.

7.1. Running the application for long periods

Running the application for a long time may impact the performance of the application and it may become unresponsive. In the
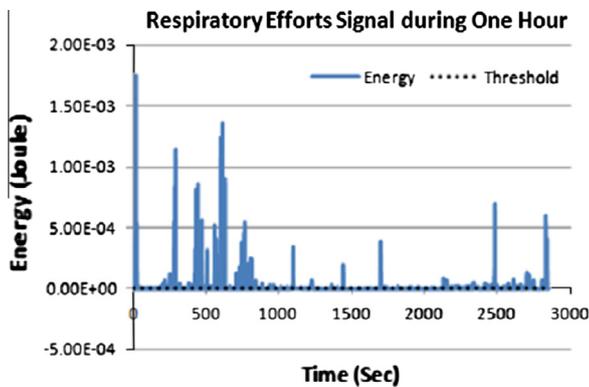
Fig. 6. Respiratory effort for one of the patients during one hour.

developed application, the smartphone collects the physiological readings from the patient for long periods (6–8 h) without user interaction. By the end of this period, the application will most likely become unresponsive when the user tries to interact with the Graphical User Interface (GUI), such as by pressing a button. This may happen when long operation takes place in the main thread, which prevents the application to process any GUI events in the application. For this reason, we have implemented an Intent-Service which handles asynchronous requests (expressed as Intents) off the main thread by way of Intent requests. Each intent is added to the IntentService's queue and handled consecutively. This allows the application to run for a long time with a decrease in the processing overhead. Moreover, if one of the operations takes a long time, it will not affect the response of the application and it will not hang. Running the application for a long time also consumes battery life. However, implementing services, avoiding creating unnecessary objects, and using static over virtual variables have helped in elongating battery life. Therefore, the patient can use the smartphone during the test without plugging it to any power resource, which makes the patients feel more comfortable. Collecting the physiological signals during the sleeping period (6–8) hours consumes (30–40)% of the battery of the used phone (Samsung-Galaxy SII). On the other hand, processing the physiological signals consumes (40–50)% of the battery life. Note that the phone can be connected to the power supply during the physiological signal analysis, since no data reading are required from the patient.

7.2. Processing Large Files

Recording the respiratory effort during the sleeping period (6–8 h) results in a file with an average size of 4 GB. To analyze the file, we read it in a vector format, and then perform the required analysis (filtering and energy calculations). This method leads to Out of Memory problem. Therefore, we split the recorded file into segments of one or two hours, and then read and analyze each file separately.

Breaking the recording file into chunks may have an impact on the number of apnea events. An apnea event may start at the end of the recording file, and end at the beginning of the next file. If each file is analyzed separately, then this apnea event will not be considered in the calculation. For this reason, if an apnea event is started at the end of a file, we check if it continues in the next file and satisfies the conditions of an apnea event (i.e., that it lasts at least 10 s).

7.3. Avoid creating unnecessary objects

Allocating memory is a very important design consideration in this application's performance. It is recommended to avoid creating objects in the application, since object creation is never free. Allocating more objects in the application enforces the process of periodic garbage collection. Fewer objects create mean less-frequent garbage collection, which has a direct impact on user experience [29].

7.4. Static over virtual

In the developed application, we defined the variables and methods as static. This allows us to access them directly without the need for creating objects. Invocations of static variables and methods are about 15–20% faster when compared to virtual [29].

7.5. Detecting false results

The application reads from three sensors, in the case that the application detects abnormal readings such as zero oxygen level or no respiratory noise, the application will automatically stop and alert the user to rerun the test.

8. Results and discussion

We performed the test on 15 subjects, 8 of which had already been diagnosed with obstructive sleep apnea (OSA), and 7 who were healthy with no symptoms. The subjects included 14 male and 1 female with different ages. Selecting subjects was conducted in collaboration with a medical sleeping center, and in consultation with the director of the center. The selected patients covered all OSA types (mild, moderate, and severe) in the test. This selection criteria allowed us to measure the capability of the application in classifying all sleep apnea types. The results showed that the apnea/hypopnea index (AHI) values obtained from the developed system are close to the values obtained using the gold standard as shown in Fig. 7. Furthermore, all suspected patients were correctly diagnosed as having the disease. On the other hand, the healthy participants were all diagnosed correctly as healthy, except for one participant who had a pulmonary disease that led to the only misdiagnosing case. In the following subsections, we will present the impact of the proposed data sets on the final diagnostic. After that, we will discuss the accuracy factors that assess the reliability of the developed system.

8.1. The impact of each physiological signal on the final diagnostic

In this subsection, we will assess the accuracy of the classification using partial data sets. The oxygen saturation and body
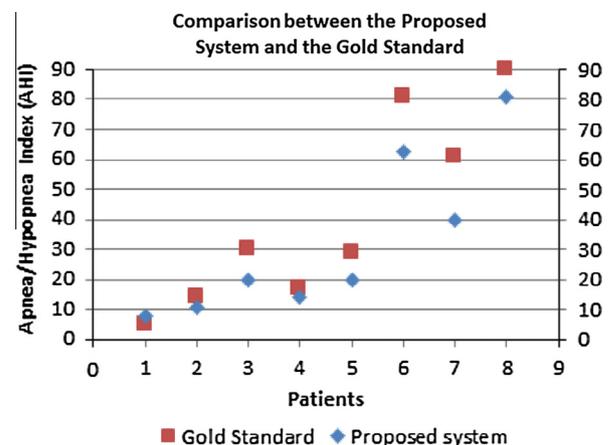

Fig. 7. Comparison between the developed system and the gold standard.

movement are considered one set (partial data set 1), and respiratory effort represents the other set (partial data set 2). While the complete data set represents the case when all physiological signals are used. Table 1 lists the results obtained from each set besides the results of the gold standard.

As seen in Table 1, all data sets correctly diagnosed the patients as having OSA. However, the severity of patient 3 was incorrectly identified in all data sets. Furthermore, the severity of patient 4 was incorrectly identified in partial data set 2. The values obtained using partial data set 1 (i.e. oxygen level and body movement) are closer to the gold standard values than the values obtained using the complete data set (i.e. oxygen level, body movement, and respiratory effort). This indicates that oxygen level and body movements are more critical to the detection of OSA. Future work will evaluate the effectiveness of using single parameters in detecting OSA. Table 2 shows the error[2], which is defined as the difference between the squares of the actual and expected values. The sum of these errors is used in calculating the root mean square error (RMSE) as seen in Eqs. (1) and (2):

$$RMSE = \sqrt{MSE} \tag{1}$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m} E_i \tag{2}$$

RMSE value is important in comparing the partial data sets statically. The smaller the RMSE, the more accurate the system is considered to be. The RMSE value of the developed system is found to be 32.5, and it has been noticed that partial data set 1 out performs partial data set 2.

8.2. Accuracy factors

When evaluating the accuracy of a clinical test, the terms sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are used. Sensitivity and specificity do not depend on the disease prevalence in the population of interest. On the other hand, PPV and NPV are dependent on the pervasiveness of the disease in the population of interest. The following terms are fundamental to the understanding of factors that affect accuracy of clinical test [30,31]:

- True positive: Sick people correctly diagnosed as sick.
- False positive: Healthy people incorrectly diagnosed as sick.

- True negative: Healthy people correctly diagnosed as healthy.
- False negative: Sick people incorrectly diagnosed as healthy.

8.2.1. Sensitivity
"Sensitivity" evaluates the percentage of actual positives which are correctly identified as such (the percentage of sick people who are correctly diagnosed as having the disease). Sensitivity is defined as shown in Eq. (3).

$$Sensitivity = \frac{Number\ of\ True\ Positives}{Number\ of\ True\ Positives + Number\ of\ False\ Negatives} \tag{3}$$

8.2.2. Specificity
"Specificity" evaluates the percentage of negatives which are correctly identified as such (the percentage of healthy people who are correctly diagnosed as not having the disease). Specificity is defined as shown in Eq. (4).

$$Specificity = \frac{Number\ of\ True\ Negatives}{Number\ of\ True\ Negatives + Number\ of\ False\ Positives} \tag{4}$$

In the developed system, we measured the number of true positives, false positives, true negatives, and false negatives. After that we calculate the sensitivity and specificity as shown in Eqs. (5) and (6):

- Number of true positives: 8 sick people correctly diagnosed as sick.
- Number of false positives: 1 healthy individual incorrectly diagnosed as sick.
- Number of true negatives: 6 healthy people correctly diagnosed as healthy.
- Number of false negatives: 0 sick people incorrectly diagnosed as healthy.

$$Sensitivity = \frac{8}{8+0} = 100\% \tag{5}$$

$$Specificity = \frac{6}{6+1} = 85.7\% \tag{6}$$

The values 100% and 85.7% for the sensitivity and specificity, respectively, mean that 100% of patients were correctly identified

Table 1
Comparison between the proposed data sets against the gold standard.

Name	Partial Data Set 1 (Oxygen + Body Movements)		Partial Data Set 2 (Respiratory Effort)		Complete Data Set (Oxygen + Body Movement + Respiratory Effort)		Gold Standard	
	AHI	Severity	AHI	Severity	Severity	AHI	Severity	AHI
Patient 1	8.1	Mild	7.9	Mild	Mild	8	Mild	5.2
Patient 2	11	Mild	10.5	Mild	Mild	10.8	Mild	14
Patient 3	20.57	Moderate	19	Moderate	Moderate	19.8	Severe	30.2
Patient 4	15.45	Moderate	14.5	Mild	Moderate	15	Moderate	17
Patient 5	21.2	Moderate	17.9	Moderate	Moderate	19.6	Moderate	29
Patient 6	65.6	Severe	60	Severe	Severe	62.8	Severe	80.7
Patient 7	41.3	Severe	38.1	Severe	Severe	39.7	Severe	60.6
Patient 8	91.8	Severe	70.4	Severe	Severe	81.1	Severe	89.9

Table 2
Root mean square error (RMSE) for all data sets.

Name	Partial Data Set 1	Partial Data Set 2	Complete Data Set	Gold Standard Error[2]			
	AHI	AHI	AHI	AHI	Partial data set 1	Partial data set 2	Complete data set
Patient 1	8.1	7.9	8	5.2	8.41	7.29	7.8
Patient 2	11	10.5	10.8	14	9	12.25	10.2
Patient 3	20.57	19	19.8	30.2	92.7	125.44	108.2
Patient 4	15.45	14.5	15	17	2.4	6.25	7.8
Patient 5	21.2	17.9	19.6	29	60.84	123.2	88.4
Patient 6	65.6	60	62.8	80.7	228	428.49	320.4
Patient 7	41.3	38.1	39.7	60.6	372.49	506.25	436.8
Patient 8	91.8	70.4	81.1	89.9	3.61	380.25	77.4
Total					777.45	1589.42	1057
RMSE					27.9	39.9	32.5

as having the disease, and 85.7% of patients were correctly identified as not having the disease. On the other hand, 14.3% of patients that did not have the disease were incorrectly identified.

Besides the sensitivity and specificity measured above, we calculated the accuracy of the severity classification (Eq. (7)), i.e., how accurately the system classifies the sick people into the correct severity (mild, moderate, and severe). Among the 8 patients that diagnosed as having the disease, 7 of them were classified correctly.

$$Accuracy = \frac{7}{8} = 87.5\% \qquad (7)$$

8.2.3. Positive predictive value

The positive predictive value (PPV) of a test is the percentage of positive test results that are true positives. This percentage is dependent on the prevalence of the disease in the population of interest and is useful to clinicians since it answers the question: "How likely is it that this patient has the disease given that the test result is positive?" [30]. The PPV is defined as shown in Eq. (8).

$$PPV = \frac{Number\ of\ True\ Positives}{Number\ of\ True\ Positives + Number\ of\ False\ Positives} \qquad (8)$$

8.2.4. Negative predictive value

The negative predictive value (NPV) of a test is the percentage of positive test results that are true negatives. This percentage is dependent on the prevalence of the disease in the population of interest and is useful to clinicians since it answers the question: "How likely is it that this patient does not have the disease given that the test result is negative?" [30]. The NPV is defined as shown in Eq. (9).

$$NPV = \frac{Number\ of\ True\ Negatives}{Number\ of\ True\ Negatives + Number\ of\ False\ Negatives} \qquad (9)$$

In the developed system, we measured the positive predictive value and the negative predictive value for the tests done on the 15 persons. The results are reported in Eqs. (10) and (11):

$$PPV = \frac{8}{8+1} = 99.9\% \qquad (10)$$

$$NPV = \frac{6}{6+0} = 100\% \qquad (11)$$

9. Conclusion and future work

Obstructive sleep apnea (OSA) is a potentially serious sleep disorder, which is characterized by repetitive pauses in breathing during sleep. Several types of sleep apnea exist, but the most common type is obstructive sleep apnea. Polysomnography (PSG) is the gold standard in diagnosing OSA, but it requires an attended overnight test in the hospital, and imposes high cost and discomfort to the patients. In an effort to relieve patients from the requirements dictated by PSG, researchers have experimented with various home-bound and inexpensive techniques for detecting OSA. Due to the popularity of smartphones and their powerful computational capabilities, we propose an OSA detection technique that makes use of the built-in sensors in the smartphone. This detection platform extracts and analyzes physiological signals including oxygen saturation, body movement, and respiratory effort with the final output being a reliable screening and classification of OSA. Following a comprehensive review of published work in this area, we proposed one approach that builds upon previous published work. To ascertain the diagnostic result, we compared our results with the gold standard and were in continuous consultation with a sleep-disorder specialist as well. The smartphone application developed performed as expected in recording and extracting the physiological signals from the patients and analyzing the collected data solely on the smartphone without the need for external resources. Hence, a self-contained, reliable and portable solution. Upon examining our system's ability to screen the disease as compared to polysomnography, we tested the application on 15 subjects. The results showed that 100% of patients were correctly identified as having the disease, and 85.7% of patients were correctly identified as not having the disease. These results demonstrate the effectiveness of the system as compared to the gold standard and emphasize the important role that smartphones can play in healthcare in the future. Currently, the application is limited to only diagnose obstructive sleep apnea. In the future, we plan to modify and add to the application to assist in diagnosing central and mixed sleep apnea as well. Further, an overlap between OSA symptoms and another disease might lead to misdiagnosis. Future work should address this issue by studying the symptoms of the other disease, and identify how to discriminate between it and OSA.

References

[1] Berger M, Oksenberg A, Silverberg S, Arons E, Radwan H, Iaina A. Avoiding the supine position during sleep lowers 24 h blood pressure in obstructive sleep apnea (OSA) patients. J Hum Hypertens 1997;11(10):657–64.

[2] Quan S, Gillin JC, Littner MR, Shepard JW. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. Sleep 1999;22(5):662–89.

[3] National Sleep Foundation (NSF). <http://www.sleepfoundation.org/article/hot-topics/let-sleep-work-you> [accessed 11.25.13].

[4] Ancoli-Israel S, DuHamel ER, Stepnowsky C, Engler R, Cohen-Zion M, Marler M. The relationship between congestive heart failure, sleep apnea, and mortality in older men. CHEST Off Pub Am Coll Chest Physic 2003;124(4):1400–5.

[5] Bucklin CL, Das M, Luo SL. An inexpensive accelerometer-based sleep-apnea screening technique. In: Proceedings of the IEEE 2010 National on Aerospace and Electronics Conference (NAECON); July 2010. p. 396–99.

[6] Burgos A, Goi A, Illarramendi A, Bermudez J. Real-time detection of Apneas on a PDA. IEEE Trans Inf Technol Biomed 2010;14(4):995–1002.

[7] Hamada M, Masahiro I. Home monitoring using portable polygraphy for perioperative assessment of pediatric obstructive sleep apnea syndrome. Tokai J Exp Clin Med 2012;37(3):66.

[8] Gao R, Yang L, Wu X, Wang T, Lu S, Han F. A phone-based e-health system for OSAS and its energy issue. In: IEEE international symposium on Information Technology in Medicine and Education (ITME), vol. 2. August 2012. p. 682–86.

[9] Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC research notes, vol. 3(1). 2010. p. 250.

[10] Brown III W, Yen PY, Rojas M, Schnall R. Assessment of the health IT usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inf 2013;46(6):1080–7.

[11] Sheehan B, Lee Y, Rodriguez M, Tiase V, Schnall R. A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents. Appl Clin Inf 2012;3(4).

[12] Kailas A, Chia-Chin Chong, Watanabe F. From mobile phones to personal wellness dashboards. IEEE Pulse 2010;1(1):57–63.

[13] Jin Z, Sun Y, Cheng A. Predicting cardiovascular disease from real-time electrocardiographic monitoring: an adaptive machine learning approach on a cell phone. In: Proceedings of the IEEE annual international conference on Engineering in Medicine and Biology Society (EMBC)); September 2009. p. 6889–92

[14] Oresko J, Jin Z, Cheng J, Huang S, Sun Y, Duschl H, et al. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans Inf Technol Biomed 2010;14(3): 734–40.

[15] Chen T, Kuan K, Celi L, Clifford G. Intelligent heartsound diagnostics on a cellphone using a hands-free kit. Assoc Adv Artif Intell 2009.

[16] Zhang Z, Wu H, Wang W, Wang B, A smartphone based respiratory biofeedback system. In: Proceedings of the IEEE 3rd international conference on Biomedical Engineering and Informatics (BMEI), vol. 2. October 2010. p. 717–20.

[17] Scully C, Lee J, Meyer J, Gorbach A, Granquist-Fraser D, Mendelson Y, et al. Physiological parameter monitoring from optical recordings with a mobile phone. IEEE Trans Biomed Eng 2012;59(2):303–6.

[18] Siewiorek D. Generation smartphone. IEEE spectrum 2012;49(9):54–8.

[19] Oliver N, Flores-Mangas F. HealthGear: automatic sleep apnea detection and monitoring with a mobile phone. J Commun 2007;2(2):1–9.

[20] Cao Z, Zhu R, Que R. A wireless portable system with micro sensors for monitoring respiratory diseases. IEEE Trans Biomed Eng 2012;59(11):3110–6.

[21] Alqassim S, Ganesh M, Khoja S, Zaidi M, Aloul F, Sagahyroon A. Sleep Apnea Monitoring using mobile phones. In: IEEE 14th international conference on applications and services in e-health networking (Healthcom)). October 2012. p. 443–46.

[22] Rofouei M, Sinclair M, Bittner R, Blank T, Saw N, DeJean G, et al. A non-invasive wearable neck-cuff system for real-time sleep monitoring. In: Proceedings of the IEEE international conference on Body Sensor Networks (BSN)). May 2011. p. 156–61.

[23] Tseng MH, Hsu HC, Chang CC, Ting H, Wu HC, Tang PH. Development of an Intelligent App for obstructive sleep apnea prediction on android smartphone using data mining approach. In: Proceedings of the IEEE international conference on Ubiquitous Intelligence & Computing/Autonomic & Trusted Computing (UIC/ATC)). 2012. p. 774–79.

[24] Penzel T, Blau A, Garcia C, Schbel C, Sebert M, Fietze I. Portable monitoring in sleep apnea. Curr Resp Care Rep 2012;1(2):139–45.

[25] Collop N, Tracy S, Kapur V, Mehra R, Kuhlmann D, Fleishman S, et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. J Clin Sleep Med (JCSM), Off Pub Am Acad Sleep Med 2011;7(5):531.

[26] Chung F, Liao P, Elsaid H, Islam S, Shapiro CM, Sun Y. Oxygen desaturation index from nocturnal oximetry: a sensitive and specific tool to detect sleep-disordered breathing in surgical patients. Anesth Analg 2012;114(5):993–1000.

[27] Yadollahi A, Moussavi Z. Acoustic obstructive sleep apnea detection. In: Proceedings of the IEEE annual international conference on Engineering in Medicine and Biology Society (EMBC)); 2009. p. 7110–13

[28] Android Developers. <http://developer.android.com/guide/components/activities.html> [accessed 25.11.13].

[29] Android Developers. <http://developer.android.com/guide/components/processes-and-threads.html> [accessed 25.11.13].

[30] Lalkhen A, McCluskey A. Clinical tests: sensitivity and specificity. Cont Educ Anaesth Crit Care Pain 2008;8(6):221–3.

[31] Mangrulkar R, Gruber S. Patients and Populations: Medical Decision-Making, open.michigan. <http://open.umich.edu/education/med/m1/patients-pop-decision-making/fall2011> [accessed 13.06.13].