**RESEARCH**

# Development of a support vector machine learning and smart phone Internet of Things-based architecture for real-time sleep apnea diagnosis

Bin Ma[1], Zhaolong Wu[1], Shengyu Li[2], Ryan Benton[2], Dongqi Li[2], Yulong Huang[3], Mohan Vamsi Kasukurthi[2], Jingwei Lin[4], Glen M. Borchert[5], Shaobo Tan[2], Gang Li[1], Meihong Yang[1*] and Jingshan Huang[2,5*]

## Abstract

**Background:** The breathing disorder obstructive sleep apnea syndrome (OSAS) only occurs while asleep. While polysomnography (PSG) represents the premiere standard for diagnosing OSAS, it is quite costly, complicated to use, and carries a significant delay between testing and diagnosis.

**Methods:** This work describes a novel architecture and algorithm designed to efficiently diagnose OSAS via the use of smart phones. In our algorithm, features are extracted from the data, specifically blood oxygen saturation as represented by SpO2. These features are used by a support vector machine (SVM) based strategy to create a classification model. The resultant SVM classification model can then be employed to diagnose OSAS. To allow remote diagnosis, we have combined a simple monitoring system with our algorithm. The system allows physiological data to be obtained from a smart phone, the data to be uploaded to the cloud for processing, and finally population of a diagnostic report sent back to the smart phone in real-time.

**Results:** Our initial evaluation of this algorithm utilizing actual patient data finds its sensitivity, accuracy, and specificity to be 87.6%, 90.2%, and 94.1%, respectively.

**Discussion:** Our architecture can monitor human physiological readings in real time and give early warning of abnormal physiological parameters. Moreover, after our evaluation, we find 5G technology offers higher bandwidth with lower delays ensuring more effective monitoring. In addition, we evaluate our algorithm utilizing real-world data; the proposed approach has high accuracy, sensitivity, and specific, demonstrating that our approach is very promising.

*Correspondence: yangmh@sdas.org; huang@southalabama.edu
[1] Shandong Provincial Key Laboratory of Computer Networks, Qilu
University of Technology (Shandong Academy of Science), Jinan, China
[5] College of Medicine, University of South Alabama, Mobile, AL 36688,
USA
Full list of author information is available at the end of the article

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 2 of 13

**Conclusions:** Experimental results on the apnea data in University College Dublin (UCD) Database have proven the efficiency and effectiveness of our methodology. This work is a pilot project and still under development. There is no clinical validation and no support. In addition, the Internet of Things (IoT) architecture enables real-time monitoring of human physiological parameters, combined with diagnostic algorithms to provide early warning of abnormal data.

**Keywords:** Sleep apnea, Obstructive sleep apnea syndrome, SpO2, Support vector machine

## Background

People in high intensity positions generally develop more sleep disorders [1]. Furthermore, sleep disorders often result in additional health problems since sleep is essential for our overall mental and physical health. Sleep apnea syndrome (SAS) is a wide spread, common sleep disorder affecting over 4% of men and 2% of women worldwide [2]. Notably, only 20% of people with SAS are actually aware they present this disorder [3].

The current tool for diagnosing SAS is polysomnography (PSG). PSG monitors breathing airflow [4], breathing events [5], snoring [6], blood oxygen saturation (SpO2) [7], electrooculography (EOG) [8], electroencephalography (EEG) [9], and electrocardiography (ECG) [10]. While effective in diagnosis, PSG carries several limitations. First, PSG is quite labor intensive as it requires a specific series of recording systems [11]. Second, PSG requires continual, hands-on supervision as patients must wear numerous body sensors. And third, PSG equipment carry significant costs with the core PSG Machines costing ~ $6000 [12–14].

Furthermore, the construction of a diagnosis model for SAS is non-trivial. The default employment of all features generally results in less than desirable performance, due to several features typically being irrelevant and/or redundant; additionally, poor feature selection can similarly also result in suboptimal performance. As such, finding an optimal set of features, without undue burden on the analysis, is essential.

SpO2 and ECG are the two features most widely studied in physiological signals collected by PSG. SpO2 represents the percentage of hemoglobin in the blood. In hypopneas status, the reduction of airflow causes a reduction in oxygen saturation. Typically, classifiers employ time frequency domain and time domain SpO2 features for apnea detection, and this strategy is the most widely accepted in the field [5]. However, there are a number of issues with existing solutions; they require the use of expensive equipment to collect the day and/or they require some offline processing in order to perform a diagnosis.

In order to detect apnea, a pulse oximeter, which is a kind of portable sensor, is utilized to record SpO2. The use of pulse oximeters carries several advantages. First, pulse oximeters produce clearer record details than ECGs

for evaluating hypopnea status [6]. Second, a regular pulse oximeter costs ~ $100 dollars, which is dramatically less expensive than a PSG machine. Third, combined with our proposed classification method and our smart phone-based system, it allows both on-line diagnosis and addresses many PSG-based issues with diagnosing obstructive sleep apnea syndrome (OSAS) (e.g. labor costs, computing resources, and time).

Internet of Things (IoT) strategies are widely utilized in an array of areas including sensor related applications [15–17] and detection systems [18–20]. Furthermore, combining IoT, cloud computing, and machine learning has previously proven effective for precise, real-time sleep apnea detection and diagnosis [21]. Cloud computing reduces the cost on servers, hardware, software licenses, and safety maintenance [22]. For instance, a sleep apnea monitoring system was proposed by Bsoul et al. [23]. The system, which utilized ECG readings, allowed for real-time monitoring; however, the diagnosis requires the processing to be done on the mobile phones, which could be problematic.

Recently, major research efforts in sleep apnea detection have focused on reducing costs through automation. One major focus of automation has been the use of preprocessing algorithms, which can be carried out on acquiring both testing and training data to improve diagnostic model accuracy and to reduce the complexity of the classification algorithms as well. The monitoring and diagnosis of SAS often include three major stages: feature extraction, feature selection, and pattern classification [24]. Publicly available physical databases are widely utilized for evaluating the efficiency of preprocessing operations [25–28].

In terms of prediction, Oliver and Flores-Mangas developed a detection algorithm in oximetry [29] that operated in real-time; however, they did not provide a comparison to standard PSG methodologies in terms of performance. Another approach, implemented by Heneghan et al. [30], used linear discriminative analysis to classify apnea using eight ECG and five SpO2 features. The sensitivity and specificity of their classifications were 51.4% and 87.3%, respectively; they noted that if either the ECG or SpO2 feature sets were missing, their model could still work, although at reduced performance. McNames and Fraser [25] and Raymond et al. [26] both

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 3 of 13

achieved accuracies of over 90% based on one-minute basis and one-minute segment analyses of the Apnea-ECG Database [31]. Notably, however, both methods required manual activities. McNames and Fraser used manual inspections to identify apneic episodes [25], while Raymond et al. relied on manual editing to improve performance [26].

Machine learning has been applied to many areas, such as medical diagnosis [32] and fault detection [33]. Importantly, unlike previously developed algorithms, we propose a Support Vector Machine (SVM) based, IoT framework embedded scheme for diagnosing sleep apnea. SVM methods have become a standard means for data classification primarily due to their simplistic numerical comparisons and optimal solution determination ease for a specific context [25]. It has been extensively used in biomedical signal classification applications, such as electromyography (EMG) [34], electroencephalogram (EEG) [35], and electrocardiogram (ECG) [36]. In addition, several approaches have also recently been developed for determining the optimal hyperplane of kernel-based, non-linear SVMs [37] in which a hyperplane is constructed that separates data representing two different groups (classes).

In our algorithm, the values of SpO2 are divided into one-minute segments as a main subset of characteristic data, and the characteristics, including maximum, minimum and other features (shown in Table 1) of this SpO2 segment are utilized to diagnose apnea. From the aspect of IoT framework for monitoring and diagnosing sleep apnea, researchers have developed some frameworks. Rofouei et al. [38] presented a non-invasive, wearable neck-cuff IoT system for monitoring the sleep of people. Kumar et al. [39] proposed a framework to help patients with OSA. The framework used sensors within the patient environment as well as wearable devices to monitor and collected a variety of data. Data collected included, but was not limited to, room temperature, room humidity, patient blood pressure, heart rate, body temperature, and blood oxygenation. The collected data by the sensors were sent to the cloud layer via a mobile phone or similar device; the data then could be used for analysis. However, there is a lack of automatic diagnosis in current studies. Thus, we combine the IoT framework based sleep monitoring system with our algorithm for automatic diagnosis and remote warning. Finally, our utilization and preprocessing of SpO2 data for diagnosis improves the accuracy of detection and reduces the false positive.

This article is a significant extension of our previous work [40] with all sections markedly extended in this report. Substantial extensions (excluding minor modifications) include: (1) Background: expanded the introduction of PSG along with six additional references; added more explanations on SpO2 and oximeter; added more explanations on IoT along with four references; added more elaborations on recent research on sleep apnea monitor and detection along with nine references; added more about the application of machine learning along with two references. (2) Materials and Methods: added four citations to describe SVM; expanded the description of the SVM models, especially the detailed explanation of the ten features used in SVM; added more details of our IoT architecture including the introduction of three main components. (3) Results and Discussion: added more explanations of sensitivity, specificity, and accuracy; added a table to show the True Negative (TN), True Positive (TP), False Negative (FN), False Positive (FP); and False added more detailed explanations of the structure of our data, along with three figures containing participant amounts categorized by age, Body Mass Indexes (BMIs), and Apnea Hypopnea Indexes (AHIs); compared our algorithm with Deep Belief Network (DBN) and Adaptive Boosting (AdaBoost) using the same data

**Table 1  SpO2 statistical features**

| Feature Name | Description |
|---|---|
| $S_{max}$ | Maximum SpO2 in 1 min |
| $S_{min}$ | Minimum SpO2 in 1 min |
| $S_{mean}$ | Average SpO2 in 1 min |
| $S_{vari}$ | SpO2 Variance in 1 min |
| CorreC | Correlation coefficient of 1 min SpO2 sample data |
| ZCount | The count of the number of zero crossing points within a segment where $S_{mean}$ is the baseline |
| SpSlope | Slope of the regression line in 1 min |
| AbSlope | Absolute slope |
| Bias | Bias of the regression line of SpO2 in 1 min |
| Dmean | Delta index: 1 min average of the absolute differences between two successive mean values of SpO2 signal over 12 s intervals |

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 4 of 13

features along with one table and one figure to show the better performance of our algorithm; simulated and compared the packet loss and packet retransmission of both 4G and 5G network along with two figures.

The most significant innovations in this work are the IoT architecture and the diagnosis algorithm of sleep apnea based on SVM. Importantly, we proposed to use the SpO2 as the feature data for training and testing on SVM, DBN and AdaBoost respectively. Therefore, the best algorithm is selected as the diagnosis algorithm of sleep apnea.

The rest of this article is structured as follows. The Methods section introduces SVM, presents our IoT architecture, and describes our algorithm. The Results section reports experimental findings, and the Discussion section discusses our results and concludes with proposed future directions.

## Methods

In this study, we utilized data curated by St. Vincent's University Hospital at University College Dublin; we entitle this Sleep Apnea Database as the UCD Database [41]. SpO2 features were obtained from the database after preprocessing and then a SVM is used for diagnosing SAS. The general process is depicted in Fig. 1.

### Purpose and overview of our method

The method presented in this work provides a novel means of detecting likely apnea events through analyzing SpO2 data. To achieve this, SpO2 data are preprocessed and divided into single minute intervals (segments). Each

segment is then classified as either an apnea event, which is defined as an at least ten second cessation of breathing [42] or as an non-apnea event. Smaller interval segments (e.g. 30 s) were not employed due to the increased likelihood of interrupting an apnea event. As such, 1 min intervals were selected to label data in this study. We utilize MATLAB R2014b for preprocessing and find this preprocessing step reduces complexity and greatly enhances overall performance.

In addition, SpO2 signals are individually down sampled to one Hertz (Hz). SpO2 values < 50 are regarded as noise [43] (as recommended by domain experts). Next, ten features are computed from each segment which is showed in Table 1. To start, the SpO2 maximum, minimum, variance, mean, and correlation coefficient are computed; these are correspondingly designed as $S_{max}$, $S_{min}$, $S_{vari}$, $S_{mean}$, and *CorreC*. $S_{mean}$, for each segment, is used as a baseline to calculate *ZCount*, which is the number of times a zero crossing occurs. In addition, each segment employs linear regression to construct a representative line from which three features: the regression line slope (*SpSlope*), the bias (*Bias*), and the absolute value (*AbSlope*). Finally, a delta index, *Dmean* is used as another Sleep Apnea–Hypopnea Syndrome (SAHS) detecting parameter; this requires two steps. The first step requires that the mean of SpO2 signal be calculated for every 12 s interval within a segment. The second step requires that the absolute differences between consecutive mean within the segment be averaged; this is the *Dmean*. Once all the features for each segment is calculated, a classifier can be constructed.
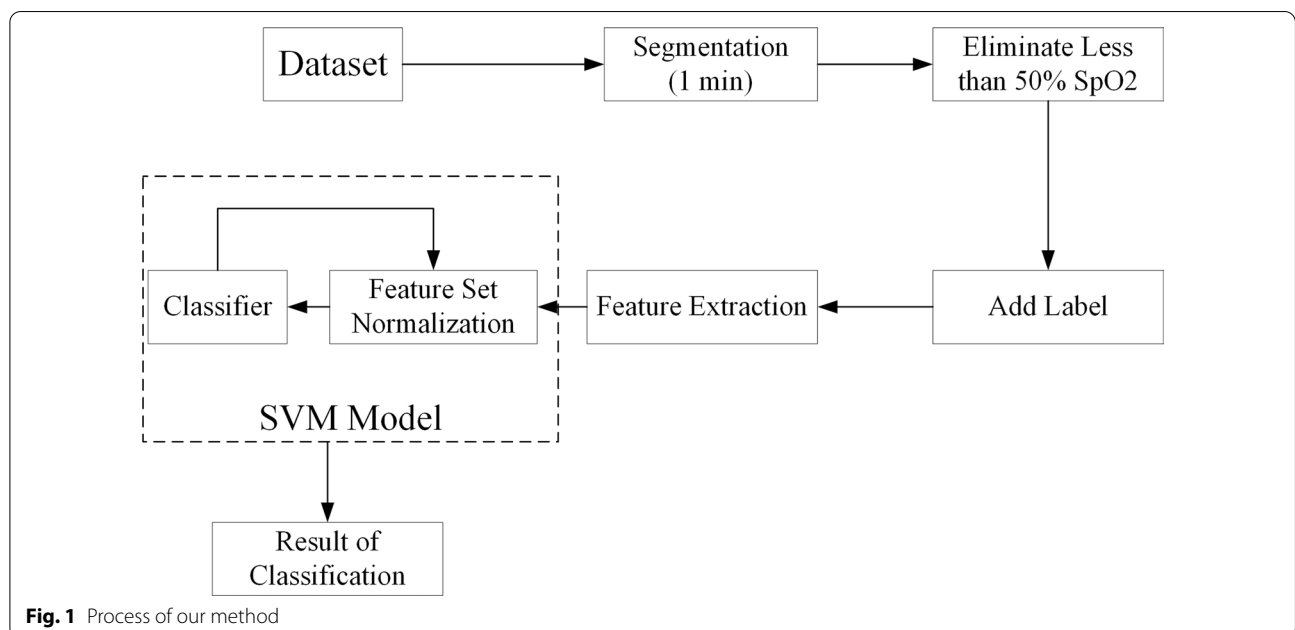


**Fig. 1** Process of our method

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 5 of 13

Our SVM classifier, which uses the Radial Basis Function (RBF) kernel, requires two principal user-provided parameter settings. The two user-provided parameters are $C$, which is set to 1,000, and $\sigma$, which is set to five. In our experiment, they are set by using trial and error. In addition, every feature is scaled to the range of 0–1 because SVMs, in general, are sensitive to feature dynamic ranges. We use an freely available Python package scikit-learn (version 0.19.0) to create the SVM and to test it; it is executed in the cloud in order to evaluate feature set performance with ten-fold cross-validation being used to train and then validate our methods.

### IoT system architecture

We now propose a sleep apnea diagnosing scheme using the framework of the IoT (Fig. 2). This scheme is comprised of three primary components: a smart phone monitoring system, a portable device, and a medical cloud monitoring center. SpO2 data are read by the portable device from the human body surface, which is then transmitted to a mobile phone via a low-power consumption Bluetooth module. After data are uploaded, our smart phone monitoring system preprocesses physiological data then uploads this data into a medical cloud-based monitoring center which carries out three tasks: long-term patient data storage, analysis, and visualization. It also provides relevant analysis to help medical staff with their diagnoses allowing medical staff to more efficiently provide effective treatment for patients with sleep apnea.

(1) *Portable device design* As SpO2 signals must be directly collected for data analysis, a portable medical oximeter was used to monitor SpO2. Using a portable medical oximeter brought us the following benefits. First, this device is simple to operate and convenient to wear. The patient only needs to wear a fingertip blood oxygen probe. Second, the device is seamlessly integrated with four functional modules: (a) single chip microcomputer module, (b) multiple sensors, signal processing module, (c) power management module and (d) wireless communication module. And also, it can measure SpO2, breathing rate, and pulse rate in real-time while the patient is sleeping.

(2) *Smart phone monitoring system* We also designed an application to help monitor the status of the patient. The application was implemented in Java using the Android Studio 3.0 environment. It was debugged on a smart phone (Huawei version 10) using the Android operating system (version 8.0). The two interfaces of the application correspond to two functions. The first function is to monitor and display the history data including the waveform and digital form of physiological signal. Another function is to provide users with real-time diagnose. Note that in the backend, the smart phone handles the storage, upload, and download of data. Also note that the application realizes data acquisition and preprocessing, including format conversion and related calculation.

(3) *Cloud monitoring center* Specifically, the center manages the following tasks: (a) storage, (b) analysis, and (c) visualization. By using the center, the following advantages are obtained. First, in terms of database scalability, the platform allows multiple medical devices to simultaneously connect and ensures the stability and reliability of two-way communication between devices and the cloud. Second, it provides different network device access schemes such as 2G, 3G, 4G, 5G and Wi-Fi, as well as device side Software Development Kit (SDK) of
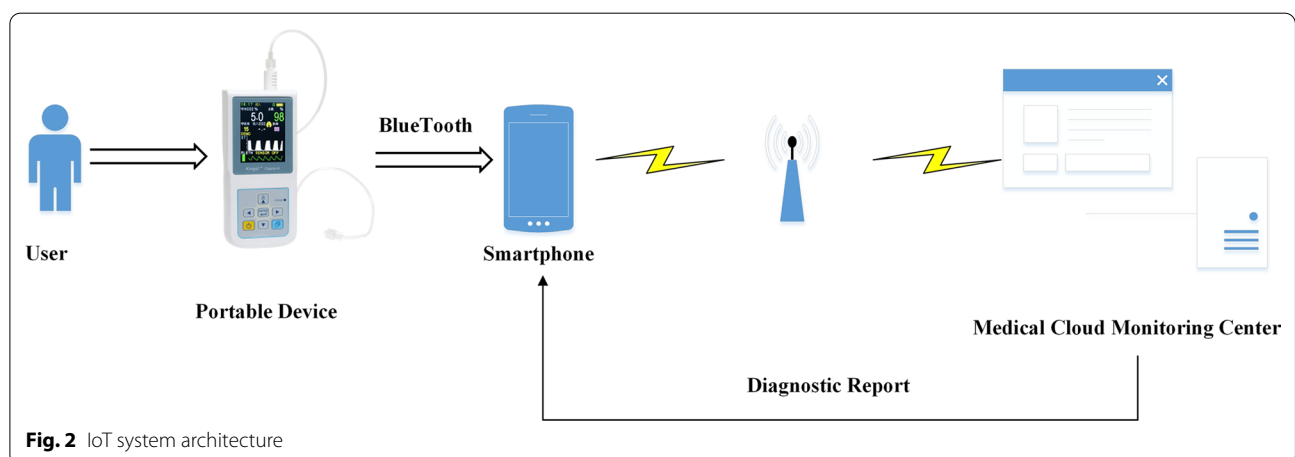


**Fig. 2** IoT system architecture

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 6 of 13

Message Queuing Telemetry Transport (MQTT), Constrained Application Protocol (CoAP), Hypertext Transfer Protocol (HTTP)/Hypertext Transfer Protocol Secure (HTTPS) and other protocols. The ability to handle multiple network schemes and protocols not only meets the real-time demand of long connection, but also meets the low power demand of short connection. Finally, in terms of security, it supports the custom data symmetric encryption channel on Transmission Control Protocol (TCP)/ User Datagram Protocol (UDP), which is suitable for power-sensitive and resource-constrained medical devices.

### Support vector machine learning

SVM learning is a widely utilized method of binary classification [44]. To create the model, an SVM exploits the information found within the training set. Once created, the model is later used to classify individual instances within the test data. During training, an SVM identifies hyperplane that best separates two classes:

$$w^T X_i + w_0 = 0$$

In particular, an SVM assumes the "best" hyperplane is a maximum margin hyperplane, which can be treated as an optimization problem. Given a training dataset $X = (X_1, X_2, \ldots, X_T)$, with associated class $y_i \in \{-1, 1\}$, 1 for normal and $-1$ for abnormal, for $i \in (1, \ldots, T)$, a hyperplane is obtained through the following mathematical constraints:

$$\begin{aligned}
&\underset{w, w_0}{\text{Minimize}} \quad \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{m} \xi_i \\
&\text{Subject to } w^T X_i + w_0 \geq 1 - \xi_i \ \text{if} \ X_i \in y_1 \\
&\qquad\qquad\ w^T X_i + w_0 \geq -1 + \xi_i \ \text{if} \ X_i \in y_2 \\
&\qquad\qquad\ \xi \geq 0
\end{aligned}$$

Margin distance equals $2/\|w\|$, and structural risk minimization involves a slack variable $\xi_i$ to be introduced in order to reduce overfitting while $C$ is a user-defined invariant. The nearest coordinate in the training data is used to calculate the distance to the hyperplane. The larger the distance, the better the classification effect with these specific data points referred to as support vectors that satisfy the rule:

$$y_i(w^T \cdot X_i + w_0) = 1$$

where $w$ represents the hyperplane's normal vector, and $w_0$ hyperplane's function bias of the. After which, normal vector $w$ can be calculated using the training dataset:

$$w = \sum_{i=1}^{n} \alpha_i y_i X_i$$

where $\alpha_i$ is the optimizing Lagrange multiplier. If $\alpha_i$ does not equal zero, a set of sample coordinates represent a support vector. Notably, $\alpha_i$ equaling zero will not disrupt the training model. In addition, the classification of $X_i$ is based on the function $h(X_i)$:

$$y_i = sign(h(X_i)) = sign(w^T \cdot X_i + w_0)$$

### Results and discussion

We have evaluated the specificity, sensitivity, and accuracy of our SVM classifier in terms of diagnosing SAS. We have also analyzed efficiency of the system when built upon IoT structure. The experimental environment employed was MATLAB R2014b, Intel (R) Core (TM) i7-6700 CPU @ 3.40 GHz.

### Metrics

We have used sensitivity, specificity, and accuracy as metrics to evaluate classification performance. They are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuary = \frac{TP + TN}{P + N}$$

where TP and TN represent the number of detected apneic segments and normal segments, respectively. FN and FP stand for the number of misidentified normal and apneic segments, respectively. Negative (N) and Positive (P) are the number of segments without and with apneic events. Sensitivity refers to the percentage of correctly classified apnea epochs, specificity as normal epoch percentage correctly classified, and accuracy as percentage of total segments correctly classified.

### Datasets

To evaluate our algorithm, we, as mentioned earlier, utilized the UCD Database that contains overnight, Jaeger-Toennies system-collected polysomnograms (in their entirety) obtained from 25 participants believed to exhibit disordered breathing during their sleep [41]. The Jaeger-Toennies system monitors all sleep related data including EEG, EOG, EMG, ECG, SpO2, etc. In our experiment, we only utilize the SpO2 data of these 25

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 7 of 13

**Table 2 Physiological properties of subjects in UCD database**

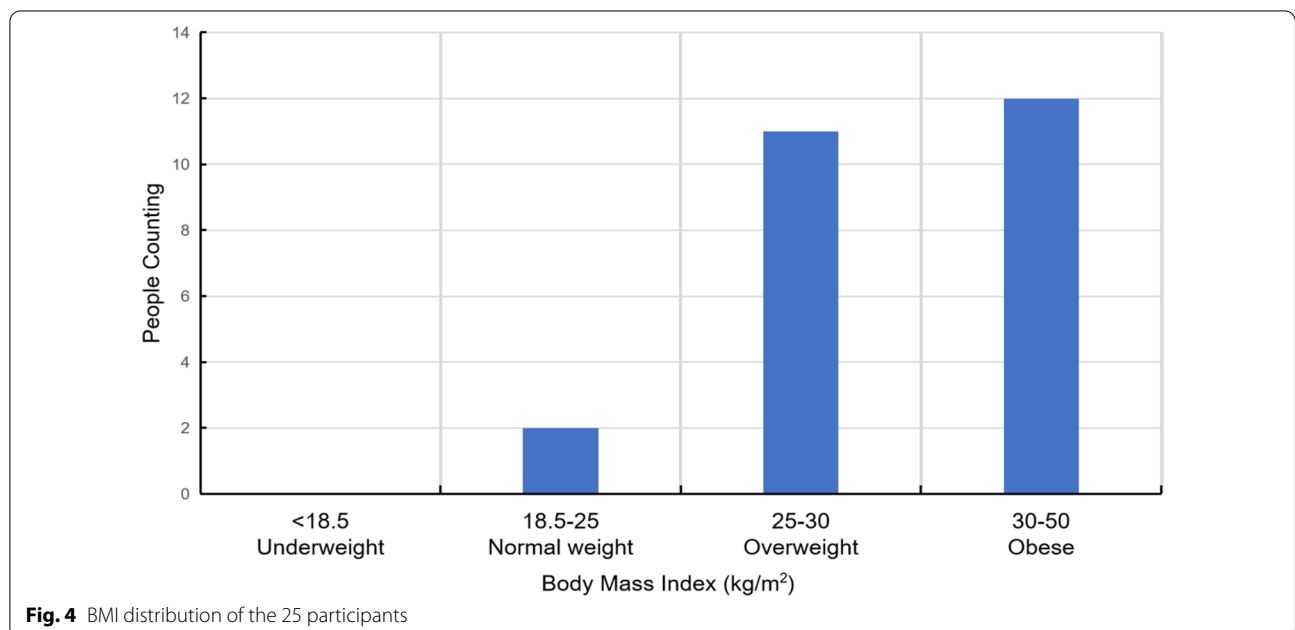|  | Mean value | Range |
|---|---|---|
| Age (years) | $50 \pm 10$ | 28–68 |
| Body Mass Index (kg/m$^2$) | $31.6 \pm 4$ | 25.1–42.5 |
| AHI | $24.1 \pm 20.3$ | 1.7–90.9 |



**Fig. 3** Age distribution of the 25 participants
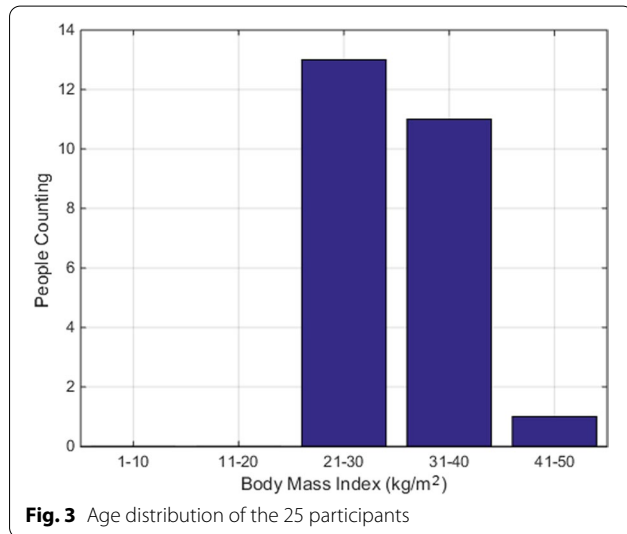
participants. All participants were adults that consented to have their SpO2 monitored by fingertip pulse oximetry. The participant physiological properties are shown in Table 2.

There were 21 males and four females with ages ranging from 28 to 68. As shown in Fig. 3, we find people with 37−54 years of age typically had more SAS. Similarly, the BMIs range from 25.1 to 42.5 kg/m$^2$, the majority of which seem to have high BMIs. While high BMIs appear to associate with a high risk of sleep apnea, this could be just co-incidence based on the data distribution. Details of the BMI are provided in Fig. 4. Moreover, patient AHIs range from 1.7 to 90.9 as shown in Fig. 5. In addition, all diagnostic information was annotated by experts to facilitate the classification of training data. With the help of domain experts, we treat every minute as a segment and assign each segment with a label, which is apnea or non-apnea.

In our experiment, we used the ten statistical features of SpO2 in order to learn and train. It allows us to verify that these features of SpO2 data can be used for the diagnosis of SAS. In all, the dataset has 1457 apnea events and 2278 non-apnea events.

### Performance comparison between DBN and AdaBoost

We also tried DBN and AdaBoost using the same data set and features. Table 3 and Fig. 6 show the performance of the classification. We tried the DBN by using the default setting of it in TensorFlow (version 1.6), an open-sourced machine learning package. We used scikit-learn (version 0.19.0) to try AdaBoost. By using trial and error, the
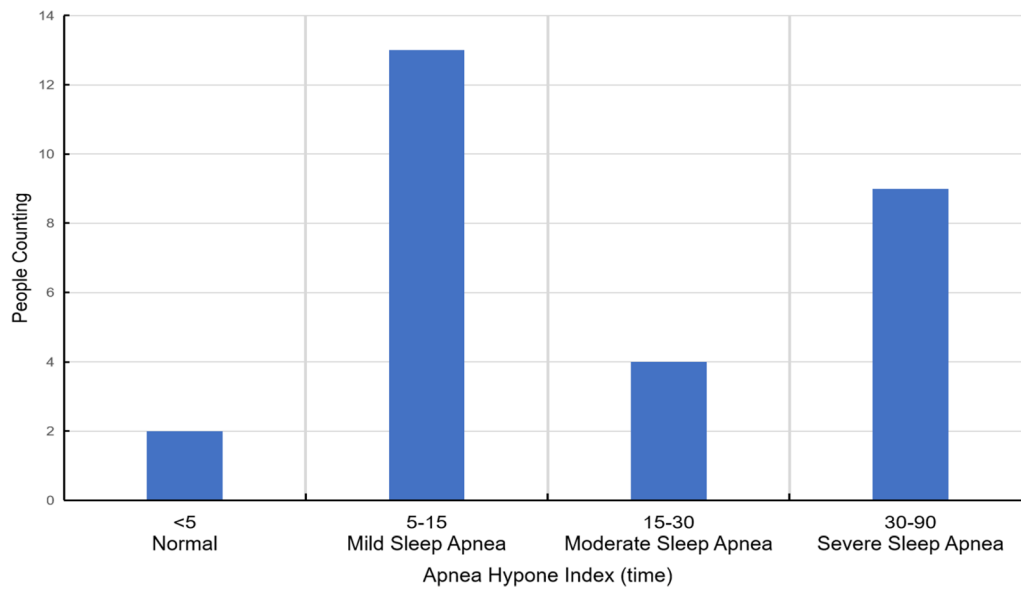


**Fig. 4** BMI distribution of the 25 participants

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 8 of 13



**Fig. 5** AHI distribution of the 25 participants

**Table 3 Performance comparison of SVM, DBM, and AdaBoost**

| Classifier | % Sensitivity | % Specificity | % Accuracy |
|---|---|---|---|
| SVM | 87.6 | 94.1 | 90.2 |
| DBN | 60.36 | 91.71 | 85.26 |
| AdaBoost | 72.64 | 87.18 | 83.64 |

Table 3 corresponds to Fig. 6

number of weak classifiers of AdaBoost was set to 10, 20, 30, and 40. Among them, 30 weak classifiers obtained the best experimental results. The performance using 30 weak classifiers was used to compare with our algorithm. The sensitivity, specificity, and accuracy of our SVM algorithm is 87.6%, 94.1%, and 90.2%, separately, which all surpass the DBN and AdaBoost. Based on sensitivity, specificity, and accuracy, which are common metrics



**Fig. 6** Performance comparison of SVM, DBM, and AdaBoost

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 9 of 13
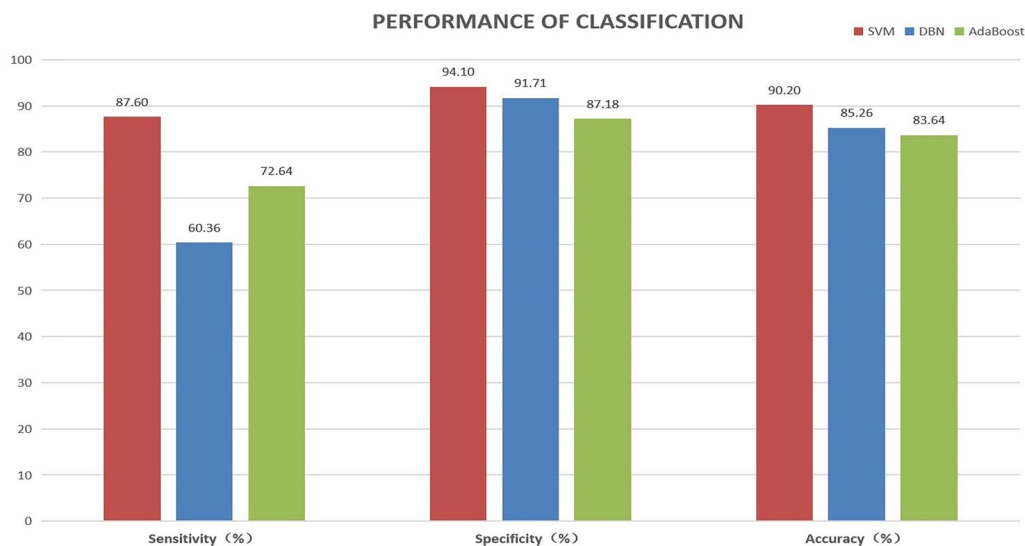
to evaluate algorithm performance, it appears that the SVM algorithm is more suitable to diagnose SAS than deep learning algorithm like DBN and AdaBoost. With respect to DBN, we believe part of the reason the SVM performs better is that deep learning networks typically require large amount of data; the problem domain used is rather small, in terms of instances, for deep learning. It also should be noted we used the default parameters for DBN; tuning DBNs tends to be time-consuming compared to SVMs. With respect to AdaBoost, the use of multiple weak classifiers does not guarantee an optimal solution. In case, it can be the tree approach of divide and conquer does not result in the best approximations for the problem; it can also be that the default parameters for the AdaBoost algorithm could be better optimized. That said, Radial Basis SVMs tend to be a good choice for many problems, giving good results.

### 5G and 4G network simulation result

To evaluate the performances of different networks when user numbers increase, we simulated the data transmission between the server and the client using 4G and 5G network. In the 4G network simulation, we set the bandwidth, delay variation, and packet loss rate to 20 Mbps, two to forty milliseconds, and 0.1%, separately. In the 5G network simulation, we set the bandwidth, delay variation, and packet loss rate to 1000 Mbps, one to five milliseconds, and 0.1%, separately. The packet loss comparison result is shown in Fig. 7. It shows that the range of 5G packet loss is lower than the range of 4G packet loss. We also simulated the number of packets retransmitted when a mobile user transmits data in 4G and 5G Network. Results are shown in Fig. 8. As shown in the figure, the number of packets retransmitted in 4G network
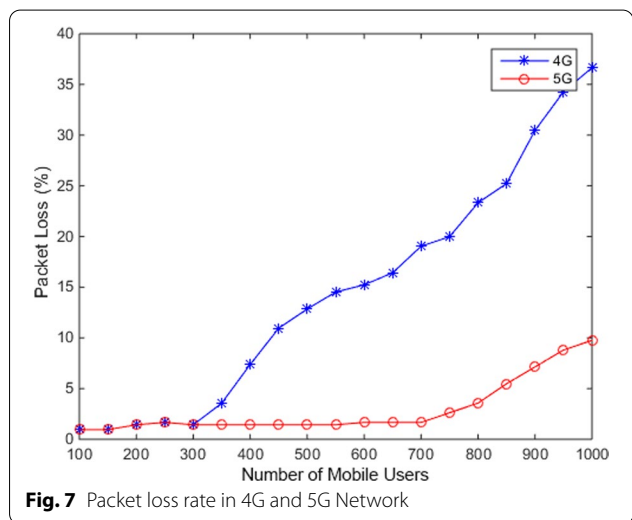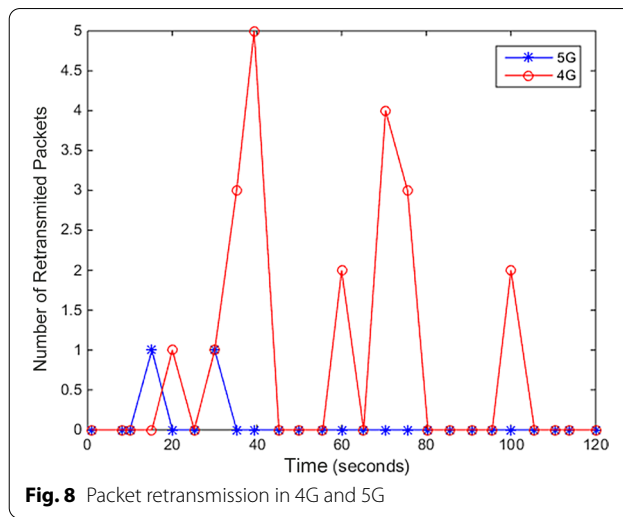


**Fig. 8** Packet retransmission in 4G and 5G

is higher than that in 5G network. Therefore, the 5G technology can effectively provide much larger data capabilities, and lower delays guarantee bandwidth availability for all users. We recommend users to use 5G network when using our application to ensure the transmission of data. Moreover, next generation 5G mobile networks are better able to provide IoT service and cloud computing service in streaming applications to mobile users and guarantee a higher Quality of Service (QoS) with significantly increased bandwidth.

### Apnea versus non-apnea

An illustration of the normal and apnea conditions is shown in Fig. 9. As can be seen, apnea has a larger effect on oxygen saturation in the blood, which significantly impacts SpO2 levels; the minimum, mean, and variance statistical features derived from SpO2 for apnea and
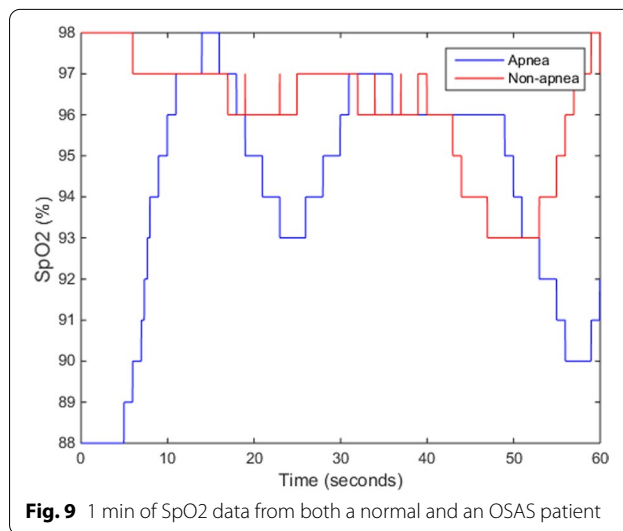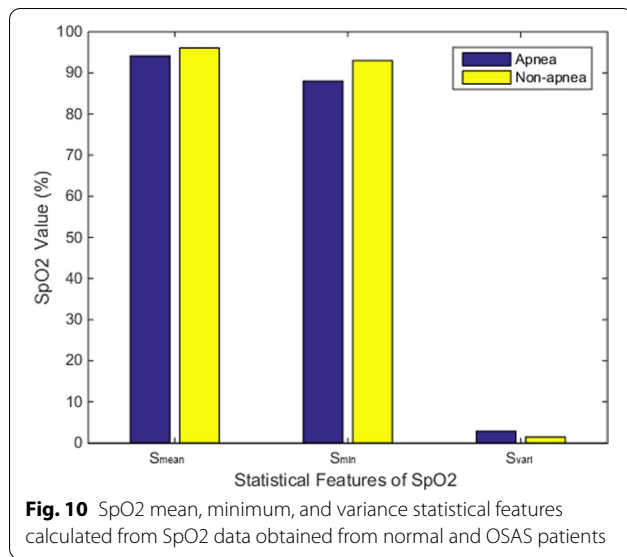


**Fig. 7** Packet loss rate in 4G and 5G Network



**Fig. 9** 1 min of SpO2 data from both a normal and an OSAS patient

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 10 of 13



**Fig. 10** SpO2 mean, minimum, and variance statistical features calculated from SpO2 data obtained from normal and OSAS patients

normal patients are shown in Fig. 10. In addition, our analysis clearly indicates our algorithm can accurately recognize apnea events occurring during sleep using the information derived from SpO2 data.
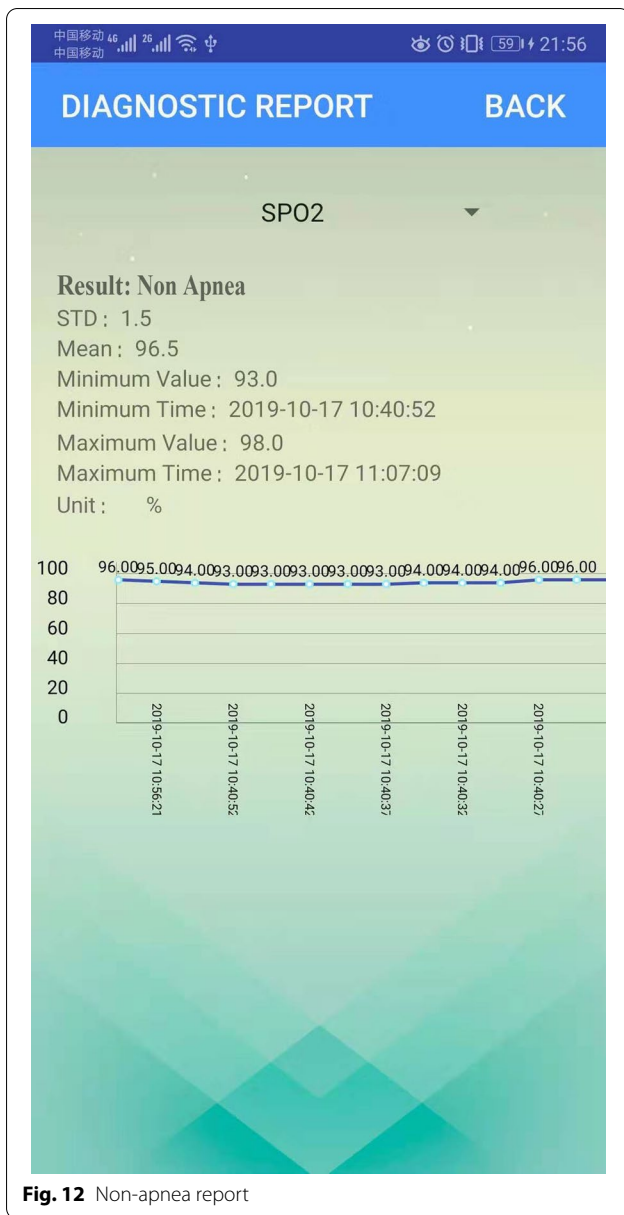
## The system test

The principle functions of our system are the storage of the patient physiological data, the preprocessing of the data, and the visualization of the data and results. To achieve this, the Bluetooth module permits a medical device to transmit specific physiological parameters to a smart phone. Next, the smart phone performs a variety of tasks. This includes the adaptive processing of the physiological parameters, the filtering out of noise (e.g. frequency interference or specious signals), and the collection of several physiological parameters including SpO2, breath rate, pulse rate, and End-tidal carbon dioxide (ETCO2). Next, the pulse waveform of ETCO2 and oxygen are displayed in real-time directly on the phone (as shown in Fig. 11). The Bluetooth and mobile phone transmission rate during result display is highly stable at 3.6 KB/s. Resultant data are stored in a cloud database allowing better maintenance. Smart phones can access diagnostic reports located in the cloud at any time, and medical clinicians can monitor patient status remotely. Figure 11 shows the statistical analysis and diagnostic result display of physiological data from 1 min; these include the minimum and maximum values, the mean, and the standard deviation. Furthermore, diagnostic reports can be obtained at any time; examples for non-apnea and apnea patients are shown, respectively, in Figs. 12 and 13.



**Fig. 11** Physiological parameters in real-time
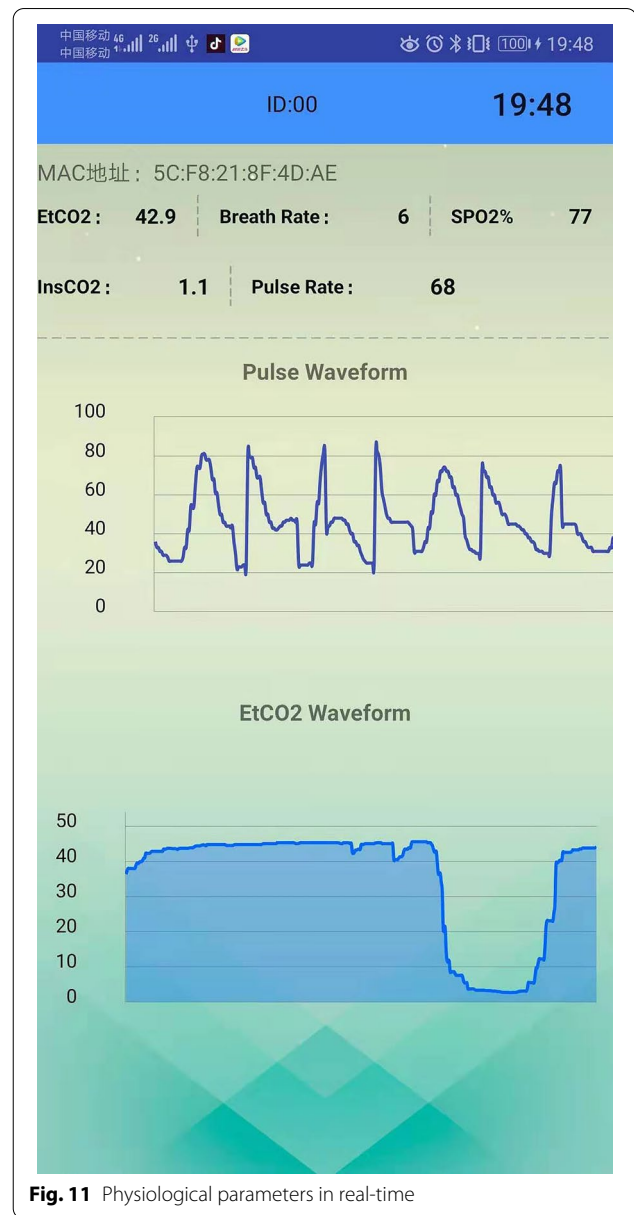
## Conclusions

OSAS represents a serious breathing disorder that occurs while a patient is sleeping. While the premiere method of OSAS diagnosis is currently PSG, it is both costly and complicated to use. What's more, PSG requires significant turnaround time for obtaining a diagnosis. In light of this, we have developed a method to detect apnea in real-time. To achieve this, we have developed a system to assess SpO2 statistical features using a SVM classifier, deployed on an IoT-based architecture and cloud computing. This novel methodology confers several advantages over current diagnostic tools. Firstly, cloud computing provides both a low cost, always accessible source of storage and

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 11 of 13



**Fig. 12** Non-apnea report



**Fig. 13** Sleep apnea syndrome report

a set of reliable computing resources essential for physiological data analysis. Secondly, our IoT architecture-based sleep monitoring system provides easy to use, remote monitoring facilitating real-time diagnosis. Thirdly, our algorithm can be employed for detecting significant, potentially serious apnea events before consulting with a medical professional for an initial diagnosis. Fourthly, an experimental evaluation of our system's ability to diagnose OSAS using pre-existing, real-world, clinical data confirm that our algorithm can efficiently diagnose sleep apnea as evidenced by its successfully achieving 87.6%, 90.2%, and 94.1% sensitivity, accuracy, and specificity rates, respectively. Notably, the sensitivity,

specificity, and accuracy of our algorithm well surpasses the performance of the existing, widely used DBN and AdaBoost tools.

Of note, in the near future, 5G networks will allow our IoT based application to operate even more efficiently through experiencing less packet loss and higher packet retransmission. In addition to this, we also plan to explore our tool's utility in classifying apnea severity and for predicting future apnea, both by using the current features and by utilizing additional features, diagnostic record integration, disease histories, and more detailed patient profiles. Moreover, in this study, our method was applied to a small-scale data in which

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 12 of 13

the BMIs of the participants are mostly high. In the future, we will apply it to large-scale data to optimize the performance of our algorithm; we would also like to expand the study to include a range of BMI. Also, we will compare our algorithm with more algorithms and determine the impact on performance as the parameters of the different algorithms change.

## Abbreviations

AdaBoost: Adaptive boosting; AHIs: Apnea Hypopnea Indexes; BMIs: Body Mass Indexes; CoAP: Constrained application protocol; DBN: Deep belief network; ECG: Electrocardiography; EEG: Electroencephalography; EMG: Electromyography; EOG: Electrooculography; ETCO2: End-tidal carbon dioxide; FN: False negative; FP: False positive; Hz: Hertz; HTTP: Hypertext transfer protocol; HTTPS: Hypertext transfer protocol secure; IoT: Internet of Things; MQTT: Message queuing telemetry transport; N: Negative; OSAS: Obstructive sleep apnea syndrome; P: Positive; PSG: Polysomnography; QoS: Quality of service; RBF: Radial basis function; SAHS: Sleep apnea–hypopnea syndrome; SAS: Sleep apnea syndrome; SDK: Software development kit; SpO2: Blood oxygen saturation; TCP: Transmission control protocol; TN: True negative.

## Authors' contributions

BM, SL, and ZW created the system framework and composed the draft of the manuscript. CB, DL, GMB, RB, ST, YH, conducted the experiments. DL, JL, MVK, MY, RB, ST, and YH also participated in preparation of the manuscript. JH conceived the study and revised the manuscript. All authors read and approved the final manuscript.

## Availability of data and materials

The apnea related data used in this research can be downloaded at: https://physionet.org/content/ucddb/1.0.0/. St. Vincent's University Hospital / University College Dublin Sleep Apnea Database is an open-access database. This database contains 25 full overnight polysomnograms with simultaneous three-channel Holter ECG, from adult subjects with suspected sleep-disordered breathing.

## Ethics approval and consent to participate

Not applicable. All data have been obtained from publically available sources. See "Availability of data and materials" for greater details.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

[1] Shandong Provincial Key Laboratory of Computer Networks, Qilu University of Technology (Shandong Academy of Science), Jinan, China. [2] School of Computing, University of South Alabama, Mobile, AL 36688, USA. [3] College of Allied Health Professions, University of South Alabama, Mobile, AL 36608, USA. [4] Ocean School, Fuzhou University, Fuzhou, China. [5] College of Medicine, University of South Alabama, Mobile, AL 36688, USA.

Published: 17 December 2020

## References

1. Yacchirema DC, Sarabia-JáCome D, Palau CE, Esteve M. A smart system for sleep monitoring by integrating IoT with big data analytics. IEEE Access. 2018;6:35988–6001.
2. Obstructive sleep apnea (OSA): Practice essentials, background, pathophysiology. https://emedicine.medscape.com/article/295807-overview. Accessed 14 Sep 2019.
3. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. Sleep. 1997;20(9):705–6.
4. Lloberes P, Sánchez-Vidaurre S, Ferré À, Cruz MJ, Lorente J, Sampol G, Morell F, Muñoz X. Effect of continuous positive airway pressure and upper airway surgery on exhaled breath condensate and serum biomarkers in patients with sleep apnea. Archivos de Bronconeumología (English Edition). 2014;50(10):422–8.
5. Dafna E, Tarasiuk A, Zigel Y. Automatic detection of whole night snoring events using non-contact microphone. PLoS ONE. 2014;8(12):e84139.
6. Güder F, Ainla A, Redston J, Mosadegh B, Glavan A, Martin TJ, Whitesides GM. Paper-based electrical respiration sensor. Angew Chem Int Ed. 2016;55(19):5727–32.
7. Babbar A, Siddiqi F, Faisal M, Bajwa A, Agarwal A, Patel M, Cury J, Seeram V, Jones LM, Louis M, et al. Use of non-invasive mechanical ventilation to prevent hypoxia during bronchoscopy in patients with sleep apnea or at high risk of sleep apnea. Am J Respir Crit Care Med. 2017;195:A2576.
8. Morselli LL, Temple KA, Leproult R, Ehrmann DA, Van Cauter E, Mokhlesi B. Determinants of slow-wave activity in overweight and obese adults: roles of sex, obstructive sleep apnea and testosterone levels. Front Endocrinol. 2018. https://doi.org/10.3389/fendo.2018.00377.
9. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, Harrod CG. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of sleep medicine clinical practice guideline. J Clin Sleep Med. 2017;13(3):479–504.
10. Song C, Liu K, Zhang X, Chen L, Xian X. An obstructive sleep apnea detection approach using a discriminative hidden Markov model from ECG signals. IEEE Trans Biomed Eng. 2016;63(7):1532–42.
11. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea. Am J Respir Crit Care Med. 2002;165(9):1217–39.
12. AlGhanim N, Comondore VR, Fleetham J, Marra CA, Ayas NT. The economic impact of obstructive sleep apnea. Lung. 2008;186(1):7–12.
13. Hillman DR, Murphy AS, Antic R, Pezzullo L. The economic cost of sleep disorders. Sleep. 2006;29(3):299–305.
14. PSG machine—Google Shopping—find the best prices and places to buy. https://www.google.com/search?newwindow=1&tbm=shop&sxsrf=ACYBGNTmBf_C_-Jcw1u7MRq6e1uWvlG7fg%3A1580283034432&psb=1&x=0&y=0&q=PSG+Machine&oq=PSG+Machine&aqs=products-cc. Accessed 14 Sep 2019.
15. Zhang R, Hao F, Sun X. The design of agricultural machinery service management system based on Internet of Things. Procedia Comput Sci. 2017;107:53–7.

Ma *et al. BMC Med Inform Decis Mak* 2020, **20**(Suppl 14):298

Page 13 of 13

16. Zhao GX, Qi B. Application of the IOT technology in the intelligent management of university multimedia classrooms. In: Applied mechanics and materials: 2014: Trans Tech Publications Ltd; 2014. p. 2050–3.

17. Ma B, Li C, Wu Z, Huang Y, Zijp-Tan ACvd, Tan S, Li D, Fong A, Basetty C, Borchert GM et al. A PWM-based muscle fatigue detection and recovery system. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM): 3–6 Dec 2018; Madrid; 2018. p. 1013–6.

18. Ma B, Li C, Wu Z, Huang Y, van der Zijp-Tan AC, Tan S, Li D, Fong A, Basetty C, Borchert GM, et al. Muscle fatigue detection and treatment system driven by Internet of Things. BMC Med Inform Decis Mak. 2019;19(7):275.

19. Xu R, Zhang L, Zhao H, Peng Y. Design of network media's digital rights management scheme based on Blockchain technology. In: 2017 IEEE 13th international symposium on autonomous decentralized system (ISADS): 22–24 Mar 2017; Bangkok; 2017. p. 128–33.

20. Baojiang S, Qiaoqiao K. Design and application of Internet of Things technology based on WeChat control. In: 2018 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData): 30 July–3 Aug 2018; Halifax; 2018. p. 549–53.

21. Yacchirema D, Sarabia-Jácome D, Palau CE, Esteve M. System for monitoring and supporting the treatment of sleep apnea using IoT and big data. Pervas Mobile Comput. 2018;50:25–40.

22. Sun T, Wang X. Research of data security model in cloud computing platform for SMEs. Int J Secur Appl. 2013;7(6):97–108.

23. Bsoul M, Minn H, Tamil L. Apnea MedAssist: real-time sleep apnea monitor using single-lead ECG. IEEE Trans Inf Technol Biomed. 2011;15(3):416–27.

24. Haoyu L, Jianxing L, Arunkumar N, Hussein AF, Jaber MM. An IoMT cloud-based real time sleep apnea detection scheme by using the SpO2 estimation supported by heart rate variability. Future Gener Comput Syst. 2019;98:69–77.

25. McNames JN, Fraser AM. Obstructive sleep apnea classification based on spectrogram patterns in the electrocardiogram. In: Computers in cardiology: 24–27 Sept 2000; Cambridge; 2000. p. 749–52.

26. Raymond B, Cayton RM, Bates RA, Chappell M. Screening for obstructive sleep apnoea based on the electrocardiogram-the computers in cardiology challenge. In: Computers in cardiology. vol. 27. Cambridge; 2000. p. 267–70.

27. Burgos A, Goñi A, Illarramendi A, Bermudez J. Real-time detection of apneas on a PDA. IEEE Trans Inf Technol Biomed. 2010;14(4):995–1002.

28. Shinar Z, Baharav A, Akselrod S. Obstructive sleep apnea detection based on electrocardiogram analysis. In: Computers in cardiology: 24–27 Sept 2000; Cambridge; 2000. p. 757–60.

29. Oliver N, Flores-Mangas F. HealthGear: a real-time wearable system for monitoring and analyzing physiological signals. In: International workshop on wearable and implantable body sensor networks (BSN'06): 3–5 April 2006; Cambridge; 2006. p. 61–4.

30. Heneghan C, Chua C-P, Garvey JF, de Chazal P, Shouldice R, Boyle P, McNicholas WT. A portable automated assessment tool for sleep apnea using a combined holter-oximeter. Sleep. 2008;31(10):1432–9.

31. Penzel T, Moody GB, Mark RG, Goldberger AL, Peter JH. The apnea-ECG database. In: Computers in cardiology: 24–27 Sept 2000; Cambridge; 2000. p. 255–8.

32. Zhang Y, Balochian S, Bhatnagar V. Emerging trends in soft computing models in bioinformatics and biomedicine. Sci World J. 2014. https://doi.org/10.1155/2014/683029.

33. Tingren TF, Xie L, Wang. Multi-fault diagnosis of pneumatic control valve with sparse Bayesian extreme learning machine. 2015.

34. Subasi A. Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. Comput Biol Med. 2013;43(5):576–86.

35. Mehta SS, Lingayat NS. Biomedical signal processing using SVM. In: IET conference proceedings. Institution of Engineering and Technology; 2007. p. 527–32.

36. Fu K, Qu J, Chai Y, Dong Y. Classification of seizure based on the time-frequency image of eeg signals using HHT and SVM. Biomed Signal Process Control. 2014;13:15–22.

37. Salem O, Guerassimov A, Mehaoua A, Marcus A, Furht B. Anomaly detection in medical wireless sensor networks using SVM and linear regression models. Int J E-Health Med Commun. 2014;5(1):20–45.

38. Rofouei M, Sinclair M, Bittner R, Blank T, Saw N, DeJean G, Heffron J. A non-invasive wearable neck-cuff system for real-time sleep monitoring. In: 2011 international conference on body sensor networks: 23–25 May 2011; 2011. p. 156–61.

39. Kumar KMC. Internet of fitness things—a move towards quantified health: concept, sensor-cloud network, protocols and a new methodology for OSA patients. In: 2015 IEEE recent advances in intelligent computational systems (RAICS): 10–12 Dec 2015; 2015. p. 364–9.

40. Ma B, Wu Z, Li S, Benton R, Li D, Huang Y, Kasukurthi MV, Lin J, Borchert GM, Tan S et al. A SVM-based algorithm to diagnose sleep apnea. In: 2019 IEEE international conference on bioinformatics and biomedicine (BIBM-19): 2019; San Diego; 2019.

41. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000;101(23):E215–20.

42. Magalang UJ, Dmochowski J, Veeramachaneni S, Draw A, Mador MJ, El-Solh A, Grant BJ. Prediction of the apnea-hypopnea index from overnight pulse oximetry. Chest. 2003;124(5):1694–701.

43. Xie B, Minn H. Real-time sleep apnea detection by classifier combination. IEEE Trans Inf Technol Biomed. 2012;16(3):469–77.

44. Ai-mei D. Research and implementation of support vector machine and its fast algorithm. Int J Multimed Ubiquit Eng. 2014;9(10):79–90.

## Publisher's Note