Published in final edited form as:

Conf Proc IEEE Eng Med Biol Soc. 2018 July ; 2018: 6044–6047. doi:10.1109/EMBC.2018.8513602.

In-Home Sleep Apnea Severity Classification using Contact-free Load Cells and an AdaBoosted Decision Tree Algorithm

Clara Mosquera-Lopez,

Department of Biomedical Engineering, Oregon Health and Science University, Portland, OR, 97239 USA

Joseph Leitschuh,

Department of Biomedical Engineering, Oregon Health and Science University, Portland, OR, 97239 USA

John Condon, Chad C. Hagen, Cody Hanks, and Peter G. Jacobs [Member, IEEE]

Department of Biomedical Engineering, Oregon Health and Science University, Portland, OR, 97239 USA

Abstract

We present a method for automated diagnosis and classification of severity of sleep apnea using an array of non-contact pressure-sensitive sensors placed underneath a mattress as an alternative to conventional obtrusive sensors. Our algorithm comprises two stages: i) A decision tree classifier that identifies patients with sleep apnea, and ii) a subsequent linear regression model that estimates the Apnea-Hypopnea Index (AHI), which is used to determine the severity of sleep disordered breathing. We tested our algorithm on a cohort of 14 patients who underwent overnight home sleep apnea test. The machine learning algorithm was trained and performance was evaluated using leave-one-patient-out cross-validation. The accuracy of the proposed approach in detecting sleep apnea is 86.96%, with sensitivity and specificity of 81.82% and 91.67%, respectively. Moreover, classification of severity of the sleep disorder was correctly assigned in 11 out of 14 cases, and the mean absolute error in the AHI estimation was calculated to be 3.83 events/hr.

I. Introduction

Obstructive sleep apnea (OSA) is a common medical condition that affects nearly 15% of the population [1] and it is often associated with other medical conditions including cardiovascular diseases, obesity, and diabetes. Despite its prevalence, a large portion of clinically significant cases remains undiagnosed [2]. OSA is characterized by the repeated occurrence of events in which breathing slows or completely stops during sleep time. Currently, a definitive diagnosis of OSA is made by the identification of relevant symptoms including daytime sleepiness, snoring, fatigue, among others and by analyzing the patient's airflow, movement, and brain activity data collected overnight during a polysomnography (PSG) study. Signals from PSG are used to quantify the occurrence of apnea and hypopnea events. The index calculated as the average of events over the total sleep time is referred to

as the Apnea-Hypopnea Index (AHI) and it is the gold standard parameter to diagnose sleep apnea and to rate its severity. An AHI < 5 events/hr is considered normal as supported by epidemiological evidence of minimal health effects; an AHI between 5–15 events/hr is regarded as mild sleep apnea, an AHI between 15–30 events/hr as moderate, and an AHI > 30 events/hr as severe [3].

The PSG study is an intrusive test that requires the use of multiple sensors worn by the patient during bed time to measure the patient's physiological variables used when identifying respiratory events for AHI calculation. These sensors likely interfere with normal sleep patterns and affect the patient's sleep quality by introducing additional sources of discomfort and sleep disturbances [4]. Home sleep apnea test (HSAT) provides a reduced set of sensors that are worn on the face and chest for use at home in a patient's own bed. However, HSAT remains obtrusive and without a technician supervising the patients in their home, patients are responsible for the application and maintenance of adequate sensor position, contact, and signal quality throughout the recording. There is therefore a need for easy-to-install unobtrusive sensors that can be deployed to continuously monitor the patient's sleep and automatically detect alterations in their sleep patterns. In this paper, we present a novel measurement system consisting of an array of multiple load cells (LC) arranged on a rectangular metallic plate that is placed underneath the patient's mattress. The system includes signal processing and machine learning (ML) algorithms that are able to extract the relevant breathing and movement-related features necessary for identifying apnea events. The features are then used by a ML algorithm to quantify AHI and thereby sleep apnea severity.

The rest of this paper is organized as follows: Section II briefly reviews previous work on sleep monitoring using LCs and automated scoring of sleep tests. Section III presents our classification approach using data from LCs. Section IV discusses experimental results, and Section V concludes the paper.

II. Related Work

Recent technological advances have made possible the development of PSG devices for sleep monitoring and sleep disorders diagnosis. Therefore, much research is being devoted to develop diagnostic equipment that are easy to deploy and use by patients and healthcare providers. From the patient standpoint, new non-contact sensing technologies such as pressure-sensitive LCs have the potential to make sleep testing more comfortable and seamless. On the other hand, automated scoring algorithms embedded in the diagnostic device have allowed fast diagnosis of OSA [5]. In this section we provide a brief review of research works that relate to the system and methods we present in this paper.

A. Sleep Monitoring Using Load Cells

Brink et al. [6] employed LCs to measure basic sleep physiology parameters including heart rate, respiration rate, and to detect body movements during sleep. Austin et al. [7] and Beattie et al. [4], [8] studied the use of load cells placed under the supports of a bed to distinguish periods of sleep and wakefulness (with sensitivity of 0.808 and specificity of

0.812), detect respiration signals and measure its rate (with mean error of 0.18 breaths/min), and identify clinically relevant breathing disturbances such as apneas and hypopneas.

B. Automated Scoring of Sleep Studies

Apart from costly commercial automated PSG scoring systems and the previously discussed work by Beattie et al. in [4], several studies have aimed at providing tools for computer-assisted scoring of sleep studies using a limited amount of signals. Gutierrez-Tobal et al. [9] developed a boosted algorithm that classifies the severity of OSA from airflow recordings using spectral features from the frequency band 0.025–0.050 Hz and time-domain central tendency statistics, achieving accuracies between 81.00 – 86.50% when classifying the severity of the sleep apnea disease from minimal to severe. On the other hand, AHI estimation has been done using electrocardiogram (ECG) signals [10], [11]. For instance, Khandoker et al. [12] recently employed wavelet descriptors obtained from 5-second short-term ECG signals to detect individual apnea and hypopnea events with accuracies of 94.72% and 79.77%, respectively. Unlike LC-based systems, most of these methods required the deployment of sensors on the patient's body.

While there are a plethora of smartphone based applications which claim to detect sleep apnea, many of these are not scientifically proven [13]. Recently, Nandakumar et al. [14] used ultrasound emitted from the speaker of a smartphone and measured by the onboard microphone to capture chest movement to classify AHI severity achieving an average error of 1.9 events/hr.

III. Methodology

A. Data Acquisition System

Our sensing device was built by integrating four ultra-low profile 50 kg capacity half-bridge DLC902 LCs (Detail-Tec, China) into a custom 3D printed enclosure subassembly. The four half-bridge LCs were wired in a full bridge configuration with resistors to balance the signal and to enable positive bias in the unloaded state. The subassemblies were then affixed in a rectangular arrangement to a 1/8" aluminum plate. The spacing was such that two cells would register the weight on either side of the body across the shoulders at an 18" spacing. The two other cells would target the weight in the hip region by being 18" apart and 20" from the other pair. The vertical dimension of the cells was under 12 mm to ensure that their presence was undetectable through a standard mattress. The output of each subassembly was input into an pre-amplifier circuit and signal conditioning on a custom built PCB. The pre-amplifier circuit was designed to accept up to 8 channels and apply a six-pole Butterworth low-pass filter of 20 Hz to each channel. The analog signals were then digitized using the AD7689 (Analog Devices) A/D converter. Data was relayed to an Odroid-U3 processor (Samsung Ltd, South Korea) via a SPI bus. The Odroid formatted the data and then sent it to a Google Drive (Google, Mountain View CA) cloud location for offline processing. The sampling frequency employed to generate the LC signal in this system was 250.0 Hz. Figure 1 shows the installation schema of our acquisition system.

B. Subjects and Sleep Studies

Fourteen participants were recruited from the Oregon Health and Science University (OHSU) sleep clinic to participate in this study approved by the OHSU institutional review board. Table I summarizes participants' demographic information. Data was collected during one or two nights using a type III Cadwell ApneaTrak™ HSAT system, in conjunction with the LC-based system described in III-A. Participants were instructed to install the sleep apnea detection system by themselves on their own bed in their home by placing the plate between their mattress and the bed box spring (or between the mattress and the ground). While the target was to collect two nights of HSAT and LC data per subject, for some subjects, there were technical problems with the ApneaTrak™ that prevented two nights of data collection, and so only one night of data was collected for these subjects.

The AHI was visually scored from HSAT signals by an expert technician who followed the American Academy of Sleep Medicine guidelines [15] to annotate all apnea and hypopnea events. Annotations made by the technician were used as ground truth for training, tuning, and performance evaluation of the classification algorithms. Figure 2 shows an example of the waveforms corresponding to two minutes of data recorded from the array of load cells, airflow, and abdomen sensors, as well as the technician annotations where distorted respiratory events are identified. The technician determined which parts of the HSAT data were interpretable during a given night, and the total duration varied across participants and across nights.

C. Predictive modeling

Figure 3 shows a flow diagram of the proposed classification approach.

1) Signal Conditioning and Feature Extraction—The signal from each of the four LCs in the array were combined and then filtered using a Chebyshev type II high-pass filter in order to remove the signal's baseline drift. The resulting high-passed signal was processed with a low-pass filter to eliminate frequencies greater than 5.0 Hz and other high frequency harmonics. Unlike [8], we did not remove movement artifacts from the signal since they carried significant information about the presence of distorted sleep patterns. Before extracting spectral features from the signal, we also standardized the signal to have unit standard deviation.

We employed the Fast Fourier Transform (FFT) to estimate the power spectral density (PSD) of an entire night of data. Several statistics were computed from the magnitude of Fourier coefficients from non-overlapping frequency bands of size 0.02 Hz. The extracted features include the maximum, mean, standard deviation, skewness, and kurtosis. These descriptive statistics convey useful information about the central tendency, dispersion, and shape of the distribution of Fourier coefficients. In addition, we extracted features from the signal in the time domain such as mean, median, variance, skewness, kurtosis, energy, and entropy of the signal's amplitude and absolute amplitude. However, these features were found to not be as relevant to classification accuracy as the spectral features.

In order to select a subset of descriptors for classification from the pool of time-domain and spectral features, we used analysis of variance (ANOVA) and F-test to identify the features with the highest F-scores between the feature and the target categories (p-value < 0.01). This resulted in six frequency bands (0.06–0.08 Hz, 0.36–0.38 Hz, 0.96–0.98 Hz, 1.18–1.20 Hz, 1.40–1.42 Hz, and 1.68–1.70 Hz) being selected. The skewness and kurtosis from those bands were found to be the most pertinent features for sleep apnea diagnosis. Figure 4 shows the distribution of these features across *Normal*, *Mild*, and *Moderate/Severe* sleep apnea.

2) Two-stage Classification Algorithm—We propose a two-stage classifier that splits the classification problem into two phases: i) An AdaBoosted ensemble of decision trees that detects OSA (*Abnormal* vs. *Normal*), and ii) a linear regression model trained on *Abnormal* samples only to then estimate the AHI and categorize the OSA as *Mild* or *Moderate/Severe* according to the predicted AHI value. Therefore, abnormal cases with AHI values between 5–15 events/hr are classified as *Mild*, and cases with AHI > 15 events/hr are considered *Moderate/Severe*. Moderate and severe OSA cases were considered a single category since both cases are clinically managed using similar therapeutic approaches, being positive airway pressure the current gold standard treatment [16], [17]. For subjects with recordings from two nights, the worst case scenario is selected for the final classification of the disease severity. Due to the known problem of night to night variability in disease severity, our algorithm uses the more severe night as reference to characterize the presence and severity of OSA. The hyper-parameters of our ML models including number of trees in the ensemble, learning rate, and linear regression regularization factor, were selected through grid search while optimizing the classification accuracy based on results from individual nights. Figure 5 shows the structure of the dominant decision tree in the ensemble, and (1) is the resulting linear regression model for AHI estimation.

$$AHI_{predicted} = 20.43 - 19.04X[2]^2 - 8.85X[3] - 11.96X[5] \quad (1)$$

IV. Results and Discussion

We assessed our approach on 23 recordings. We used leave-one-patient-out cross-validation to train and evaluate the algorithm. For subjects with two nights of data both nights were held out from the training set in order to better estimate the generalization performance of our classification algorithm. The average duration of the interpretable portions of studies in this data collection was 5.15 ± 2.94 hours per night. Table II shows the results for the categorization of OSA severity and the estimated AHI values for patients who tested positive for OSA. Our two-stage classification method reached a correct classification rate of 86.96% in the first stage when signals are analyzed to detect the presence of the sleep disorder with sensitivity of 81.82%, specificity of 91.92%, positive predictive value (PPV) of 90.00% and negative predictive value (NPV) of 84.61%. After the second stage, when AHI estimations are calculated to rate OSA severity, our method correctly classifies 78.57% of the cases with underline average AHI estimation error of 3.83 events/hr, and Pearson's correlation

coefficient of 0.84 between actual and predicted AHI values. Our algorithm tends to overestimate the AHI of patients belonging to the *Mild* group.

We observed that for patients with recordings from two consecutive nights, the AHI has low variability. However, the data collected from *Patient 12* shows a significant increment in the AHI calculated for the second night (corresponding to 3 obstructive and 8 central apneas) when compared to the first night from 2.17 events/hr to 27.04 events/hr, which is not captured by our algorithm that assigns both recordings to the *Normal* category. This high variation could be explained by the short duration of the recording collected during the second night.

We also evaluated the performance of the algorithm when using HSAT signals from nasal cannula and abdomen belt sensors as inputs instead of the LC signal. Performance indicators corresponding to this experiment are presented in Table III. The most relevant frequency bands changed slightly, but our algorithm is flexible enough to support different input feature vectors with minimal impact on performance. We found that the LC-based sensors yield higher accuracy in detecting OSA than airflow and abdominal sensors.

There are a few limitations associated with our sensing device. Like other commercially available HSAT systems, it does not provide signals to accurately capture patients' sleep stages. Moreover, additional research is needed to develop algorithms to find signatures that allow the disambiguation of breathing and movement-related signals when the patient shares the bed, as well as automated methods to detect the valid portions of the recordings, thus discarding the segments of the signal recorded while the patient has moved outside the range of the sensing device.

Despite the discussed shortcomings of the system at its current stage, the potential of LC-based sensing systems to accurately detect and diagnose OSA severity in unattended settings is demonstrated by the high correlation between the results of the manual scoring of HSAT studies and the score estimated by automatically processing pressure signals from load cells.

V. Conclusion

In this paper, we described the design and evaluation of a system for unobtrusive sleep monitoring that is able to accurately detect and predict the severity of OSA using an array of non-contact LC conveniently deployed underneath patients' mattresses. We tested our two-stage prediction algorithm on a set of 23 recordings (overnight sleep studies on 14 subjects), reaching an accuracy of 86.96% in detecting OSA (sensitivity = 81.82%, specificity = 91.67%). The overall diagnosis of the OSA severity was accurate in 78.57% of the test cases. Severity classification was done by estimating the AHI which was found to be strongly correlated with the score manually assigned by an expert technician (*Pearson's* $R=0.84$). Participants in the study were able to install the system themselves with minimal help from clinicians. Our results confirm the potential of LC-based systems equipped with ML algorithms to seamlessly monitor sleep disorders in a home environment, and to reduce the cost and improve reproducibility of sleep studies.

Acknowledgments

Research supported by the NIH Grant NIH/NIHLB 2R01HL098621-04A1.

References

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol*. 2013; 12(5):757–761.
2. Gibson G. Obstructive sleep apnoea syndrome: underestimated and undertreated. *Br Med Bull*. 2004; 72(1):49–65. [PubMed: 15798032]
3. A. A. of Sleep Medicine. Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep*. 1999; 22(5):667–689. [PubMed: 10450601]
4. Beattie ZT, Hagen CC, Pavel M, Hayes TL. Classification of breathing events using load cells under the bed. *Proc. of IEEE Eng Med Biol Soc; Minneapolis, MN, USA: IEEE; 2009*. 3921–3924. Conference Proceedings
5. Aurora RN, Swartz R, Punjabi NM. Misclassification of osa severity with automated scoring of home sleep recordings. *Chest*. 2015; 14(3):719–727.
6. Brink M, Miller CH, Schierz C. Contact-free measurement of heart rate, respiration rate, and body movements during sleep. *Behavior Research Methods*. 2006; 38(5):11–21.
7. Austin D, Beattie ZT, Riley T, Adami AM, Hagen CC, Hayes TL. *Proc of IEEE Eng Med Biol Soc. San Diego, CA, USA: IEEE; 2012*. Unobtrusive classification of sleep and wakefulness using load cells under the bed. Conference Proceedings
8. Beattie ZT, Jacobs PG, Riley TC, Hagen CC. *Proc of IEEE Eng Med Biol Soc. Milan, Italy: IEEE; 2015*. A time-frequency respiration tracking system using non-contact bed sensors with harmonic artifact rejection. Conference Proceedings
9. Gutierrez-Tobal G, Alvarez D, Del Campo F, Hornero R. Utility of adaboost to detect sleep apnea-hypopnea syndrome from single-channel airflow. *IEEE Trans Biomedical Eng*. 2016; 63(3):636–646.
10. Lee W, Nagubadi S, Kryger MH, Mokhlesi B. Epidemiology of obstructive sleep apnea: a population-based perspective. *Expert Rev Respir Med*. 2008; 2(3):349–364. [PubMed: 19690624]
11. Jung DW, Hwang SH, Lee YJ, Jeong D, Park KS. Apneahypopnea index prediction using electrocardiogram acquired during the sleep-onset period. *IEEE Trans Biomedical Eng*. 2016; 64(2):295–301.
12. Khandoker AH, Gubbi J, Palaniswami M. Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings. *IEEE Trans Information Tech in Biomed*. 2009; 13(6):1057–1067.
13. Behar J, Roebuck A, Domingos JS, Geder E, Clifford GD. A review of current sleep screening applications for smartphones. *Physiological Meas*. 2013; 34(7):29–46.
14. Nandakumar R, Gollakota S, Watson N. *Proc ACM Mobile Syst, App, and Services. Florence, Italy: ACM; 2015*. Contactless sleep apnea detection on smartphones. Conference Proceedings
15. Berry RB, Brooks R, Gamaldo CE, Harding SM, Lloyd RM, Marcus CL, Vaughn BV. The American Academy of Sleep Medicine, Guideline. 2016. The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications.
16. Epstein LJ, Kristo D, Strollo PJ, Friedman N, Malhotra A, Patil SP, Ramar K, Rogers R, Schwab RJ, Weaver EM, Weinstein MD. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J of Clinical Sleep Med*. 2009; 5(3):263–276. [PubMed: 19960649]
17. Spicuzza L, Caruso D, Di Maria D. Obstructive sleep apnoea syndrome and its management. *Ther Adv Chronic Dis*. 2015; 6(5):273–285. [PubMed: 26336596]

Fig. 1.
Data acquisition system's installation schema.

Fig. 2.

Two-minute signal segments from load cell array, airflow, and abdomen belt sensors, with annotated respiratory events.

Fig. 3.

Flow diagram of the classification system. Dashed links are activated only if sleep apnea is detected.

Fig. 4.
Distribution of features extracted from relevant frequency sub-bands grouped by disease severity.

Fig. 5.
Structure of the dominant decision tree in the classification ensemble.

TABLE I

Patients' Overall Demographic Information

Description	Value
Age (years)	48 ± 21
Female/Male	11/3
Body mass index (Kg/m^2)	33.49 ± 7.83

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE II

Classification Results

ID	Night	Study duration * (hours)	Clinical Category		AHI		Severity	
			Actual	Predicted	Actual	Predicted	Actual	Predicted
1	1	7.76	Abnormal	Abnormal	9.63	11.20	Mild	Mild
	2	9.08	Abnormal	Abnormal	10.07	12.23		
2	1	6.93	Abnormal	Normal	17.46	<5	Moderate	Mild
	2	5.35	Abnormal	Abnormal	16.16	12.22		
3	1	3.56	Abnormal	Abnormal	12.49	11.96	Mild	Mild
4	1	8.15	Normal	Normal	3.61	<5	Normal	Normal
	2	5.37	Normal	Normal	1.51	<5		
5	1	4.83	Abnormal	Abnormal	31.42	28.93	Severe	Moderate/Severe
6	1	5.07	Normal	Abnormal	4.42	16.02	Normal	Moderate/Severe
	2	0.77	Normal	Normal	1.34	<5		
7	1	5.00	Abnormal	Abnormal	5.70	13.58	Mild	Mild
	2	1.40	Abnormal	Abnormal	5.07	8.02		
8	1	8.57	Abnormal	Abnormal	5.11	10.12	Mild	Mild
9	1	1.54	Normal	Normal	4.02	<5	Normal	Normal
	2	1.45	Normal	Normal	1.40	<5		
10	1	9.59	Normal	Normal	1.49	<5	Normal	Normal
	2	9.05	Normal	Normal	2.14	<5		
11	1	3.03	Normal	Normal	1.67	<5	Normal	Normal
12	1	2.82	Normal	Normal	2.17	<5		
	2	0.43	Abnormal	Normal	27.04	<5	Severe	Normal
13	1	4.26	Abnormal	Abnormal	18.87	18.99	Moderate	Moderate/Severe

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

ID	Night	Study duration * (hours)	Clinical Category		AHI		Severity	
			Actual	Predicted	Actual	Predicted	Actual	Predicted
14	1	8.70	Normal	Normal	0.71	<5	Normal	Normal
	2	5.88	Normal	Normal	0.18	<5		

* Valid portion of the study used for sleep test scoring and AHI calculation.

TABLE III

Comparative Accuracy Analysis LC vs. HSAT Input Signals

Input signal	Accuracy %	Sensitivity %	Specificity %	PPV %	NPV %
Contact-free LC	86.96	81.82	91.67	90.00	84.62
Airflow	82.61	72.73	91.67	88.89	78.57
Abdomen	78.26	72.73	83.33	80.00	76.92