35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Detection of Apnoea from Respiratory Time Series Data Using Clinically Recognizable Features and kNN Classification

Anirudh Thommandram[1], J. Mikael Eklund, *SMIEEE*[1] and Carolyn McGregor, *SMIEEE*[2]

Abstract— **Apnoea is a sleep related breathing disorder that is common in adults and can be described as a temporary closure in the upper airway during sleep. A system using time series analysis of one minute epochs of respiratory impedance signals to detect apnoea is described. An algorithm has been developed using MATLAB for extracting clinically recognizable features from the respiratory impedance signal. One minute samples are classified using kNN classification of the feature set. The output of the system has been shown to detect apnoeic episodes in eight eight-hour patient records collected from the PhysioNet database. The specificity of the classifier is 88.1% and the sensitivity is 95.7%. ROC analysis was performed and the area under the ROC curve is 0.9604. Future research will include testing the classifier in a much larger dataset and also a novel method for the presentation of classification results to physicians.**

I. Introduction

Apnoea is a sleep related breathing disorder that is common in adults. Apnoea that goes undiagnosed can be a risk factor for some cardiovascular diseases. Therefore, early detection of apnoea in patients is essential. Currently, classification of apnoeic episodes is performed using polysomnography, which is an expensive process involving the patient spending the night in a sleep lab connected to devices recording several physiological signals and being monitored by medical professionals.

In this paper, we present a system for the automatic detection of sleep apnoea using k-Nearest-Neighbour (kNN) classification of clinically recognizable features. The system extracts these features from only the respiratory impedance (RI) signal of the patient, which can be captured from standard bedside monitors in real-time. We report our analysis results of using this approach as the sensitivity and specificity of the algorithm compared to apnoea datasets annotated by human experts.

II. Background and Related Work

A. Sleep Apnoea

Sleep related breathing disorders are common in the adult population and has been reported to be found in 4% of men and 2% of women [1]. The most prevalent type of sleep related breathing disorder in adults is obstructive sleep apnoea (OSA), accounting for about 84% of cases [2]. OSA can be defined as a temporary closure of the upper airway during sleep when air is prevented from entering the lungs [3]. This obstruction produces a negative pressure in the airway and results in ineffective airflow, where the amount of oxygen inhaled decreases and the concentration of CO_2 increases. Episodes of OSA are typically accompanied by a drop in blood oxygen saturation which leads to a central nervous system activation, triggering the airway to open without the patient even knowing. This phenomenon has been observed to repeat over 600 times in a single night in patients with severe sleep apnoea [4]. In other cases, the activation causes the patient to wake up in order to breathe. The disjointed sleep pattern due to OSA can lead to excessive daytime sleepiness, poorer cognitive performance and depression. Also, OSA that goes undiagnosed has been discovered to be a factor in the development of hypertension, congestive heart failure and even stroke [5, 6].

A typical sleep study involves recording multiple channels of various bio-signals requiring many sophisticated devices and electrode attachments to patients as well as specialised attending personnel. The cost and availability of these resources present an opportunity for low cost and more accessible screening methods. Opportunities abound to support the monitoring of patients using techniques such as we propose while they sleep in their bed in their own homes.

B. Related Work

Most automated detection schemes involve black box modeling systems such as artificial neural networks. While some of these systems perform well, their adoption in practice is very low, possibly because of the lack of traceable rules inherent to black box systems [7]. Obstructive sleep apnoea and hypopnoea can be automatically detected by extracting wavelet-based features in electrocardiogram (ECG) recordings and finding patterns using a feed forward neural network [8]. The ECG waveform is the most used signal in the development of apnoea detection algorithms. Although apnoea is a respiratory condition, it is widely recognised that the effects can be seen in the ECG. Mendez et al. demonstrated the importance of using time-variant or time-frequency approaches for correctly managing the non-stationarities in the signals, typical of apnoea episodes [9].

A signal representing the respiratory rate can be generated from an ECG signal by detecting peaks in the waveform. These signals, called ECG derived respiratory signals (EDR) are a vital part of apnoea detection algorithms. Avci et al used wavelet decompositions of EDR signals used the wavelet detail components as features for classification through a nonlinear auto-regressive type artificial neural network. They achieved an accuracy of 93.3% in the subject-based assessment [10]. Correa explored the value of the power spectrum of several EDR signals in classifying apnoea. Power spectral density was calculated for each epoch and central, mean and peak frequencies were obtained. They used a threshold based decision process to classify the

[1]A. Thommandram and J.M. Eklund are with the Department of Electrical, Computer and Software Engineering, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada.

[2]C. McGregor is with the Faculty of Business and Information Technology, University of Ontario Institute of Technology.

Corresponding author: mikael.eklund@uoit.ca.

minute sample as apnoea or not apnoea. They found that spectral analysis of the R-wave area of ECG can be a useful indicator of apnoea [11]. Spectrum analysis on ECG has been used in various ways in the past in attempts to find underlying patterns in the data that cannot be noticed with simple visual analysis. Liu et al. used the ECG to generate a heart rate variability (HRV) signal and performed spectrum analysis on it. An interesting finding was that the modulation in HRV was a little earlier than the start and the end of a sleep apnoea event by about 5 seconds. This can be useful for real time monitoring and intervention [12].

Yilmaz et al. attempted to not only classify apneic episodes, but also to determine the sleep stage of a patient at any point in time. They also use the ECG R-R interval (RRI) data and extract 4 features that are put through a support vector machine (SVM) algorithm. They achieved an accuracy of 76% for sleep stage scoring and 87% accuracy for the detection of apnoea [13]. Ghunaimi et al. used Statistical Signal Characterization (SSC) to determine apneic episodes. They took the Hilbert transform of the R-R interval signal, which produces an analytical signal that is useful for calculating instantaneous attributes of a series at any point in time. They computed the amplitude mean, amplitude deviation, period mean, and period deviation over a 5 minute moving window and chose optimal threshold values based on receiver operator characteristic (ROC) analysis [14]. Schluter et al. developed a decision tree classifier towards an approach for automatic sleep stage scoring and apnoea detection by using rules formulated for sleep technicians for manual scoring. They used derivative dynamic time warping (DDTW) to perform pattern matching to detect the same shapes in the signal as a sleep technician would detect [15].

It has been hypothesized that simple features can still be sufficient for accurate detection if extracted from a variety of signals. Belal et al. used a combination of HR, RR and SpO_2 features and fed them to a neural network to find correlations [16]. Xie et al. examined the performances of apnoea detection using ECG and saturation of peripheral oxygen (SpO2) signals, individually and in combination. They used the ECG features proposed by referenced literature and focused on feature designs of the SpO2 due to the strong reflection of arterial oxygen saturation on the airflow fluctuation. They found the best results came when they used 31 SpO2 features and 8 ECG features and put them through a combination of three classifiers with the final decision made by majority voting [17]. Using 39 features to get a good accuracy may seem excessive when compared to the process a human expert takes to perform the same classification. Isa et al. assessed the performance of several classification methods using different ECG feature sets proposed in literature. They found a higher overall accuracy when using only 3 features than they did with a system that used 8 features. This shows that a large number of features are not necessary to achieve success. Another finding of their research is that particular classification techniques work better on particular types of feature sets [18].

Apart from the ECG, the RI waveform is believed to be very valuable for apnoea detection as it is directly correlated to breathing effort. Lee et al. recognized that RI signals also contain fluctuations caused by the beating of the heart. These fluctuations are misinterpreted as breaths by bedside monitors and cause inaccurate alarms [19]. They developed a cardiac filter that involved resampling the RI at the frequency of the ECG to reduce the effect of the periodic fluctuations caused by the heart. Some research has moved in a different direction, taking the respiratory measurement from locations other than the chest. Ansari et al. discovered that it is possible to extract a reliable RR using signal processing from impedance measured across the arm [20]. While this method resolves the cardiac effect, it is prone to movement artifacts. Yen et al. describe a method for detecting apneic events in patients while they were titrated for continuous positive airway pressure (CPAP). They obtain an RI signal using a forced oscillation technique which applies an oscillatory pressure signal to the respiratory system and measuring nasal pressure and airflow. This airway impedance value was compared to a fixed threshold to classify apnoea [21].

Our approach is unique in that we are using only the RI waveform to detect apnoea and no ECG, HR, or SpO_2 signals to provide a more detailed view of events. Obtaining an RI signal can be done in several ways. One way is to measure the impedance between ECG leads, but another method is to use a strap across the chest of a patient and measuring expansion and contraction through a strain gauge type device. These devices can be manufactured very cheaply and have application for home based monitoring removing reliance on the skilled placement of the ECG leads. Through this study we will ascertain the feasibility of creating an apnoea detection system that produces sufficient accuracy using minimal cost hardware and sensors.

III. Methodology

A. Data

The PhysioNet Apnoea-ECG database is used in this study. The database is comprised of 70 records of patient data, containing a digitized 100Hz ECG signal. 40 records were of apnoeic patients, 10 records were considered borderline and 20 were control records. Each recording has approximately eight hours of data and is accompanied with a set of reference apnoea annotations. The annotations were derived by human experts on the basis of simultaneously recorded related signals [22, 23]. There is an annotation for each minute of the recording to indicate the presence of apnoea during that minute. Although the database is primarily for ECG recordings, eight of the recordings also have respiratory effort signals. Chest and abdominal respiratory effort signals were obtained using inductance plethysmography and oronasal airflow was measured using nasal thermistors.

For our study, we focus solely on the respiratory effort signal measured across the chest present in eight of the records. This waveform is also called the respiratory impedance (RI) signal, as it is generated by measuring the impedance of a wire coil strapped around a person's rib cage. The signal value increases as the chest expands during inspiration and decreases as the chest contracts during expiration. A one hour sample of an apnoea record from PhysioNet showing the respiratory signal and annotations ('A' for apnoeic, '●' for non-apnoeic) is shown in Fig. 1.

Figure 1. One hour sample of respiratory impedance at the chest along with apnoea annotations.

B. Feature Extraction

A core aspect of any classification problem is constructing a meaningful feature set. Many approaches to automatic detection and classification of apnoea use purely statistical values such as spectral variances and entropy. Such features do not necessarily have any meaning to a human expert. While some of the systems perform quite well, there is much reluctance by medical professionals in adopting such systems simply because the reasoning behind the classifications is so different from their training. As a goal for our research, the features must be clinically recognizable. This refers to features that are not purely mathematical or statistical in nature. They are things that human experts observe when deciding if a signal is apnoeic.

A key step in selecting features is determining the breaths taken by the patient. To do this, the RI wave is used to define the interval between breaths from the peaks of the impedance waveform, which correspond to the maximum chest expansion in each breath. The peaks are found using an edge detection algorithm written in MATLAB. The signal is normalized and the locations and characteristics of the peaks are recorded. Four normalized one minute window samples are shown in Fig. 2. Samples 1 and 2 are annotated as non-apnoeic and samples 3 and 4 are marked as apnoeic by human experts.

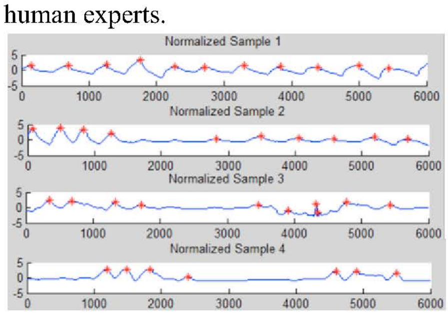

Figure 2. Normalized 1-minute window samples with breaths marked by peak detection algorithm.

Four clinically observable features were selected from the RI waveform: the stability of the peak-to-peak time, the stability of the heights of peaks, the presence of long pauses, and flat-lining. The stability of the peak-to-peak time is a measure of how consistent the breath times are. The value ranges between 0 and 100, with higher numbers representing a more stable peak-to-peak time. The stability of the heights of the peaks is a measure of how consistent the amplitude of breaths is. The presence of long pauses feature is included to detect instances where the breathing is very slow. The flat lining indication is a very strong indication of apnoea as it represents no effort in breathing. In this study, we try to determine how accurate we can build a classifier based on this modest feature set.

C. kNN Classification

A classification system takes a sample set of features and assigns the sample a label representing which class it believes the sample belongs to. The k-Nearest-Neighbour (kNN) algorithm is a very popular approach due to its simple nature and relatively robust performance. It is described as a lazy learning algorithm where there is a minimal training phase but costly test phase. This is because in the worst case, every training sample might contribute to the decision. kNN is a nonparametric classification technique, which means it makes no assumptions on the underlying data distribution. This is useful when dealing with real world parameters. It classifies objects based on the closest examples (neighbours) from the training set. The closeness is calculated as the Euclidean distance between the test and training samples in the feature space. The label that comprises the majority of the k nearest training samples is assigned to the test sample.

We experimented with several k values and found that k=44 results in the optimal performance. Different distance functions were also evaluated before deciding on using inverse Euclidean distance weighted average, in which the class of each of the k nearest points is multiplied by a weight proportional to the inverse of the distance between that point and the test point. Since we only had 8 records with a respiratory effort signal, a simple 66% split of the data into training and test sets would result in poor training of the classifier. To improve the performance of the classifier, 10-fold cross-validation is applied. This technique involves dividing the full set into 10 approximately balanced subsets. Then, the kNN algorithm is applied in 10 iterations using the subset as the training set and the rest of the samples as the test set. The 10-fold cross-validated accuracy is calculated as the average of the 10 resulting accuracies. To implement the kNN algorithm, we used the open-source machine learning software known as WEKA. This Java based software is widely used by many researchers and scientists.

IV. Results And Discussion

A total of 3,947 epochs were processed by the kNN classifier. The accuracy, defined as the percentage of correct classifications was found to be 91.2%. The specificity was 88.1% and the sensitivity was 95.7%. We also performed receiver operator characteristic (ROC) analysis. The ROC curve is generated by plotting the true positive rate against the false positive rate and is shown in Fig. 3. The area under the ROC curve (AROC) is often used as a measure of the performance of a classifier. The area measures discrimination, which represents the ability to correctly classify a randomly chosen pair of apnoeic and non-apnoeic samples. The AROC of the kNN classifier was 0.9604.

While the results of this study are promising, it is important to note that the size of the sample set was very small. We made the most of the annotated data by not simply splitting it into training and test but rather use k-fold cross-validation with 10 folds. It was a goal of the study to determine whether high detection accuracy could be achieved using only the respiratory impedance waveform. If combined with some ECG analysis, the resulting accuracy could be boosted even more.

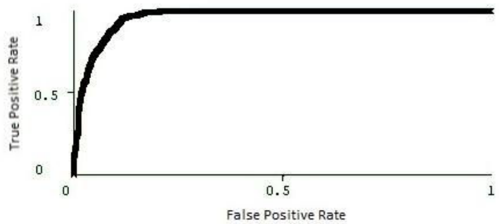

Figure 3. ROC curve for performance of algorithm.

Also, it was discovered that the value for the k parameter in the kNN classification algorithm can have a large impact on the accuracy of the system. Originally, small values of k (less than ten) were used but it was found that the classification was susceptible to noisy samples and the accuracy suffered. Generally, a small k value means that noise will have a greater influence on the classification. But simply choosing a very large value of k makes the classification computationally expensive. One method to choose a k value is to take the square root of the number of samples. Our optimal value of k=44 falls under the square root of the 3,947 samples so it was sufficiently robust to noise and also offers reasonable computational cost.

V. Conclusion and Future Work

This paper presents a system for the automatic detection of sleep apnoea using k-Nearest-Neighbour (kNN) classification of clinically recognizable features. We report our analysis of using this approach compared to the manual annotation by human experts using multiple physiological data streams. Our results show that it is possible to accurately detect apnoea from only the RI waveform and clinically recognizable features using the algorithms presented here. This enables a very feasible alternative to polysomnography and diagnosis of sleep apnoea can be made without the use of expensive machinery or specialised personnel. This can drastically reduce the number of patients with sleep apnoea that go undiagnosed.

In future work we will perform a study on a larger data set to further validate the feature set used to detect apnoeic episodes. While high specificity and sensitivity are the main goals of designing an apnoea detection algorithm, we are also working on a novel way to present classification results to physicians. Due to the clinically observable nature of the feature set, any classification can be easily deconstructed into traceable steps and checks. These checks will be part of the report presented to physicians at the end of the test.

References

[1] T.Young,M.Palta, J.Dempsey, J. Skatrud, S.Weber, andS.Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," New Engl. J. Med., vol. 328, no. 17, pp. 1230–1235, 1993.

[2] Malhotra, A., & White, D. P. (2002). Obstructive sleep apnoea. *Lancet*, 360(9328), 237–45. doi:10.1016/S0140-6736(02)09464-3

[3] AASM. (n.d.). The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.

[4] Penzel, T., & Conradt, R. (2000, April). Computer based sleep recording and analysis. *Sleep medicine reviews*. doi:10.1053/smrv.1999.0087

[5] T. Young, P. Peprad,M. Palta, K.M. Hla, L. Finn, B. Morgan, J. Skatrud, Population- based study of sleep-disordered breathing as a risk factor for hypertension, Archives of Internal Medicine 157 (1997) 1746–1752.

[6] Himmelmann, a. (1999). Hypertension: an important precursor of heart failure. Blood pressure, 8(5-6), 253–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10803484

[7] McGregor, C., Catley, C., James, A., Padbury, J., 2011, Next generation neonatal health informatics with Artemis, Studies In Health Technology And Informatics, 169, p.115-119.

[8] Khandoker, A. H., Gubbi, J., & Palaniswami, M. (2009). Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings. IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society, 13(6), 1057-67. doi:10.1109/TITB.2009.2031639

[9] Mendez, M. O., Corthout, J., Van Huffel, S., Matteucci, M., Penzel, T., Cerutti, S., & Bianchi, a M. (2010). Automatic screening of obstructive sleep apnea from the ECG based on empirical mode decomposition and wavelet analysis. Physiological measurement, 31(3), 273–89. doi:10.1088/0967-3334/31/3/001

[10] Avcı, C., Delibaşoğlu, İ., & Akbaş, A. (2012). Sleep Apnea Detection Using Wavelet Analysis of ECG Derived Respiratory Signal, (February), 27–28.

[11] Correa, L. S., Laciar, E., Mut, V., Torres, A., & Jané, R. (2009). Sleep apnea detection based on spectral analysis of three ECG - derived respiratory signals. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2009, 4723–6. doi:10.1109/IEMBS.2009.5334196

[12] Liu, W. X. (2008). Defintion of Sleep Apnea Event by one minute HRV Spectrum Analysis, (2006), 2292–2294.

[13] Yilmaz, B., Asyali, M. H., Arikan, E., Yetkin, S., & Ozgen, F. (2010). Sleep stage and obstructive apneaic epoch classification using single-lead ECG. Biomedical engineering online, 9, 39. doi:10.1186/1475-925X-9-39

[14] Al Ghunaimi, B., Hossen, A., & Hassan, M. O. (2004). Screening of obstructive sleep apnea based on statistical signal characterization of Hilbert transform of RRI data. Technology and health care : official journal of the European Society for Engineering and Medicine, 12(1), 67–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15096688

[15] Schluter, T., & Conrad, S. (2010). An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection. 2010 IEEE International Conference on Data Mining, 1007-1012. IEEE. doi:10.1109/ICDM.2010.60

[16] Belal, S. Y., Emmerson, A. J., & Beatty, P. C. W. (2011). Automatic detection of apnoea of prematurity. Physiological measurement, 32(5), 523-42. doi:10.1088/0967-3334/32/5/003

[17] Xie, B., & Minn, H. (2012). Real-time sleep apnea detection by classifier combination. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society, 16(3), 469–77. doi:10.1109/TITB.2012.2188299

[18] Isa, S. M., Fanany, M. I., Jatmiko, W., & Murni, A. (2010). Feature and Model Selection on Automatic Sleep Apnea Detection using ECG, 357–362.

[19] Lee, H., Rusin, C. G., Lake, D. E., Clark, M. T., Guin, L., Smoot, T. J., Paget-Brown, A. O., et al. (2012). A new algorithm for detecting central apnea in neonates. Physiological measurement, 33(1), 1-17. doi:10.1088/0967-3334/33/1/1

[20] Ansari, S., Najarian, K., Ward, K., & Tiba, M. H. (2009). Extraction of Respiratory Rate from Impedance Signal Measured on Arm: A Portable Respiratory Rate Measurement Device. 2009 IEEE International Conference on Bioinformatics and Biomedicine, 197-202. IEEE. doi:10.1109/BIBM.2009.68

[21] Yen, F. C., Behbehani, K., Lucas, E. a, Burk, J. R., & Axe, J. R. (1997). A noninvasive technique for detecting obstructive and central sleep apnea. IEEE transactions on bio-medical engineering, 44(12), 1262–8. doi:10.1109/10.477699

[22] Penzel, T., Rg, G. B. M., Goldberger, M. A. L., & Peter, H. (2000). The Apnea-ECG Database - Computers in Cardiology 2000, 255–258.

[23] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, 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 101(23):e215-e220