Towards Sleep Apnea Screening with an Under-the-mattress IR-UWB Radar Using Machine Learning

Abdul Q. Javaid, Carlo M. Noble*, Russell Rosenberg* and Mary Ann Weitnauer

Georgia Institute of Technology, Atlanta, Georgia, USA

*Neurotrials Research Inc., Atlanta, GA 30332-0250, USA

Email: aqjavaid@gatech.edu, maweit@gatech.edu.

Abstract—In this work, we apply machine learning to investigate the effectiveness of an Impulse Radio Ultra-Wide Band (IR-UWB) radar panel, in an under-the-mattress configuration, for detecting apnea events in subjects known to have obstructive sleep apnea (OSA). We consider a collection of features, some novel and some inspired by features that worked well for sleep apnea detection using other types of sensors (i.e., not IR-UWB). To extract the features, we collected a total of 25 hours of data from four subjects as they slept through the night. The data included digitized samples of the IR-UWB radar return signal and the scored polysomnograph (PSG), which is the gold standard and measures a large number of physiological parameters in a well-equipped sleep laboratory. Normal and apnea epochs were extracted from the IR-UWB data corresponding to normal and apnea epochs in the PSG data. Statistical features were derived from these extracted epochs and a Linear Discriminant classifier was trained. Using cross-validation, we found that the classifier had an accuracy of around 70% in detection of apnea and normal epochs. The novel aspect of this project involves processing and investigation of different methods for feature extraction on data obtained from real apnea subjects and suggests that the radar, when paired with other under-the-mattress sensors might provide an effective screening device in a convenient form factor.

I. Introduction

Sleep apnea is a disorder that involves involuntary cessation of breathing during sleep for 10 or more seconds [1]. Sleep apneas are categorized into two main types: obstructive sleep apnea (OSA) and central sleep apnea (CSA) [1]. OSA occurs when the muscles such as soft palate and tongue temporarily obstruct the upper airway causing a temporary pause in breathing. A 90% or greater decrease in nasal pressure or breathing amplitude is termed as "Obstructive Apnea" (OA) while a 30% or greater drop in breathing amplitude is termed as "Obstructive Hypopnea" (OH) [1]. In CSA, the brain fails to send signals to the breathing muscles and a pause in breathing takes place without any narrowing of the upper airway. OA and OH are the most common presentations of sleep disordered breathing (SDB) in adults, with an estimated prevalence of 5-15% [2]. Sleep apnea is associated with growing number of health problems which include high blood pressure, stroke, irregular heart rate and depression [3], [4]. The metrics to quantify severity of the disorder include Apnea-Hypopnea Index (AHI), which is the number of apnea and hypopnea events in one hour [1]. The standard approach for diagnosis of sleep apnea is polysomnography (PSG), which monitors sleep and respiration by measuring various physiological parameters, including Electrocardiogram (ECG), blood oxygen saturation (SpO_2), respiratory effort (thoracic and abdominal) and nasal airflow [1]. However, PSG is an expensive procedure and requires patients to sleep in technician-attended and well-equipped labs. Moreover, large number of sensors are attached to the patient at various points on the body for monitoring of these many physiological parameters and may cause discomfort and also vary the sleeping behavior of the patient.

This paper investigates the feasibility of applying machine learning algorithms to the signal from an extremely low-power IR-UWB radar to detect episodes of OA and OH in adults. An IR-UWB radar transmits small duration pulses and if focused on human torso, the reflected pulses contain information about breathing and cardiac rate [5]. The IR-UWB radar signal in this study has been severely analog filtered to just the respiration band (cut-off: 0.7 Hz), to reduce panel cost, so the times of pulse arrivals are not measured, but rather the Fourier components of rhythmic motions associated with breathing are captured in this band. As such, an apnea event manifests itself in the received signal. Its extremely low-interference, non-ionizing nature and convenient form factor of a panel placed under a mattress makes IR-UWB radar a promising component of a home screening tool for OSA [6], [7], which is the motivation for this research. This paper explores 14 moment-based features and time-domain principal components-based feature that exploits the start and stop times of apnea periods as indicated by the scoring of the PSG.

II. Other Apnea sensing Technologies

Research on home based sleep apnea monitoring can be divided into three main categories based on phsyiological signals under consideration: (1) Heart rate variablility (HRV) analysis and respiratory signal derived from ECG (2) Respiratory signals, more commonly known as 'Respiratory Effort' signals, obtained via nasal sensors, chest straps (respiratory plethysmography) or radars (3) Nocturnal sound analysis by capturing snoring via microphones. The first two methods (except radar) still require sensors being attached to the subject and may pose challenges in hygene and skin irritation. The third method sometimes utilizes a contact microphone to detect snores while a non-contact microphone requires to be placed at an optimal position to avoid loss in efficiency of the system [8]. HRV, derived from ECG, is the most widely explored physiological parameter and has been used in [9], [10] to detect sleep apnea.

Microwave radars (doppler and IR-UWB) have gained a lot of popularity in measuring vital signs. The main differences between the doppler and the UWB radars are the bandwidth, power, and the band of operation. The IR-UWB radar signal can be spread throughout the 3 to 10 gigahertz band [11], which spans many other licensed and unlicensed bands. The doppler radar must be used in either a licensed band or an ISM (industrial, scientific and medical) band, and makes more interference on other devices because of its more narrow bandwidth and also it is more vulnerable to the interference from these other devices. Because of its extremely wide bandwidth, the UWB radar can be operated at a much lower average power than the doppler radar, for the same signal-to-noise ratio in the radar receiver, and it has higher time resolution which translates to more sensitivity to small periodic movements of the reflecting surfaces. A doppler radar has been used in [12] to detect different types of breathing disorders that include apneas. Similarly two microwave radars are used in [13] to detect sleep apneas. Doppler radar is used in [14], [15], in a device placed on the bedside table named 'Sleepminder', developed by Biancamed for detection of sleep and wake patterns along with apnea events. The novelties in our work include the sensor location under the mattress and some novel features.

IR-UWB radar is known to correlate well with chest straps, for measuring the changes in respiratory amplitude [16], since they both measure chest displacement. However, since airflow is important for distinguishing OA from OH [17], we are investigating combining the radar with microphone data in our on-going work. This paper reports our interim results based only on the radar sensor.

III. METHODS

A. Measurement Setup

The data for the project was collected at a sleep research laboratory (Neurotrials Research Inc.) [18] under a protocol approved by the Georgia Institute of Technology (GT) Institutional Review Board (IRB). The IR-UWB radar system developed by Sensiotec Inc. [19] is placed under the mattress. The radar transmitted pulses are 13 ns long, centered at 4.2 GHz. At the receiver, the reflected signal is time-gated and down-converted to baseband and then hardware-filtered into a respiration band (low pass, cut-off frequency 0.7 Hz) and heart band (cut-off frequencies: 0.5 - 6 Hz), respectively. This paper considers only the respiration band. Next, the outputs of each filter band are sampled at 128 Hz and quantized for subsequent digital signal processing. PSG data was also collected simultaneously. The data from the IR-UWB radar and the PSG data are both time stamped. A specialist from Neurotrials scored the PSG and marked all the normal and apnea epochs. Fig. 1 shows the block diagram of the set up for data collection.

B. Signal Processing

1) Preprocessing of Respiration Signal: In order to facilitate the extraction of features, the quantization noise in time domain respiratory signal is removed by filtering twice with a 20 tap triangular filter. To avoid changing the amplitude and shape of the signal, only the central tap is assigned a
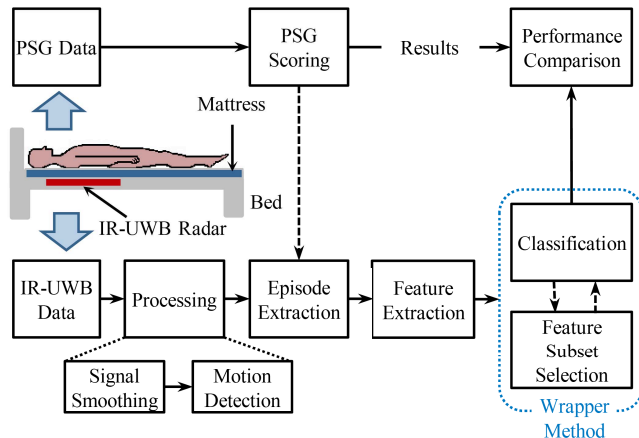

Fig. 1. Block diagram of the set up. The IR-UWB radar was placed under the mattress. The PSG and IR-UWB radar data was time-synchronized.

weight of two while the remaining taps have unity weights. An example of the respiration signal produced by the IR-UWB radar after the application of triangular filter is the blue oscillating waveform shown in Fig. 2. The signal shows changes in both amplitude and frequency during apnea.

2) Motion Detection & Removal: The radar signal is very sensitive to motion and shows clipping to the ADC maximum and minimum levels whenever the subject makes large muscle movements. The part of the IR-UWB data that is motion corrupted is detected and removed. All the maximas and minimas in the input time domain signal are compared against pre-defined thresholds. Let Max_i and Min_i denote the i-th maxima and minima, respectively. If a maxima Max_i is above the upper threshold, then the portion of the signal that includes Max_i, Max_{i-1} and Max_{i+1} is labeled as motion corrupted. Similarly, if a minima Min_i is below the lower threshold, then the portion of the signal that includes Min_i, Min_{i-1} and Min_{i+1} is labeled as motion corrupted. The reason for including one previous and one next maxima or minima is to make sure that the complete portion encompassing motion is removed. The clean signal that falls between the two motion corrupted parts is standardized by calculating the z-score of each data point in that portion. The reason for separate standardization of data that are separated by motion corrupted regions is to avoid error in feature extraction stage as the radar signal's amplitude is different for different postures and varies across population, especially across genders.

3) Extraction of Normal & Apnea Epochs: Epochs or episodes of 60 second duration, that include the marked apnea event, are extracted from the IR-UWB radar data. Let n denote the number of samples in an epoch ($n = 128 \times 60$, since the sampling rate is 128 Hz). More specifically, the 60 seconds in each epoch are composed of $t_s - 20s$ to $t_s + 40s$, where t_s represents the apnea start time which is obtained from the PSG. Also the 60 second duration epochs corresponding to time intervals of normal sleep are extracted by using information from the PSG. All the epochs corresponding to both OA and OH are combined into one class, labelled as $Apnea$, and the normal epochs are assigned class label $Normal$, thus making our task one of binary classification.
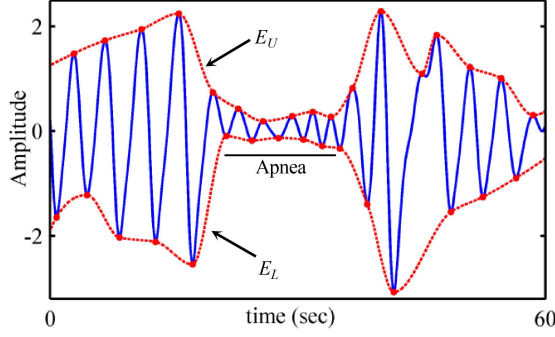
Fig. 2. Time domain respiration signal from IR-UWB radar with an obstructive apnea event. The upper and lower envelopes are shown in red.

C. Feature Extraction

The features are extracted from both the time and frequency domains for each epoch. We define f_k to be the k-th feature ($k = 1,2,...,15$). The process of feature extraction is divided into two phases. Phase I consists of first- and second-order moments of various quantities extracted directly from each episode. Phase II contains two more complicated features which are extracted after the application of pricincipal component analysis and signal processing algorithms.

1) Phase I: The features f_1 through f_{13} are calculated to capture the change in amplitude and frequency of respiratory signal in apnea epochs [1], [9]. Variances are calculated using the unbiased formula for sample variance.

- f_1: The Mean Absolute Deviation (MAD) of respiratory signal. In [20], authors apply this feature to R-R intervals extracted from ECG signals for apnea and normal epochs classification.

- f_2: The Mean Absolute Deviation of the maxima or peak values of respiratory signal.

- f_3: Number of times the signal crosses the mean.

- f_4: The variance of the time duration (number of samples) between the points where the signal crosses the mean.

- f_5: The variance of the number of samples between a local maxima and the next local minima in an epoch.

- f_6: The variance of the difference in amplitude of local maxima and next local minima in an epoch.

- f_7: The inter-quartile range, i.e., the difference between the 75th and the 25th percentiles of number of samples between mean crossings in each epoch. This feature is also extracted from R-R intervals of ECG data in [20].

- f_8: The sum of Power Spectral Density (PSD) values in frequency range [0 - 0.5 Hz] for each episode. The PSD values are obtained by taking the magnitude squared of the Discrete Fourier Transform (DFT) over the length of the epoch and then dividing by the length of the epoch. The PSD values for the ECG-derived respiration signal are used as features in [9].

In a normal breathing epoch, the signal properties such as peak-to-peak difference, frequency etc. are expected to remain constant throughout the length of the epoch. However, these properties will show variation in an apnea epoch. In order to capture this change, each $60s$ epoch is divided into small portions of t_d duration ($t_d = 10s$ is chosen in this paper, as the minimum duration of an apnea event is $10s$ [1]). Let these small portions be denoted by ϵ_i and the variance of the respiration signal in each small portion be represented by $\sigma_{\epsilon_i}^2$, where $i = 1, 2, ..., 6$. Following three features are extracted from these small epochs:

- f_9: The variance of $\sigma_{\epsilon_i}^2$ values across the small epochs ϵ_i, i.e., $var([\sigma_{\epsilon_1}^2, \sigma_{\epsilon_2}^2, \cdots, \sigma_{\epsilon_6}^2])$.

- f_{10}: The variance of absolute difference of $\sigma_{\epsilon_1}^2$ across ϵ_i, i.e., $var([|\sigma_{\epsilon_1}^2 - \sigma_{\epsilon_2}^2|, |\sigma_{\epsilon_2}^2 - \sigma_{\epsilon_3}^2|, \cdots, |\sigma_{\epsilon_5}^2 - \sigma_{\epsilon_6}^2|])$.

- f_{11}: The DFT is calculated for the small epochs and the frequency $\mathbf{f}_{\epsilon_i}$ corresponding to highest peak is detected in the spectrum of each small epoch ϵ_i. The variance $\sigma_{\mathbf{f}}^2$ of the frequency associated with the highest peak across these small epochs is used as a feature, i.e., $\sigma_{\mathbf{f}}^2 = var([\mathbf{f}_{\epsilon_1}, \mathbf{f}_{\epsilon_2}, \cdots, \mathbf{f}_{\epsilon_6}])$.

The upper envelope E_U of the time domain respiratory signal in each epoch is estimated by joining the maximas using cubic interpolation and the lower envelope E_L is obtained by cubic interpolation of minimas as shown in Fig. 2. The envelope difference, $E_D = E_U - E_L$, is divided into small epochs ε_i of duration t_d ($t_d = 10s$). The variance $\sigma_{\varepsilon_i}^2$ is calculated for each small epoch and the following features are derived:

- f_{12}: The variance of $\sigma_{\varepsilon_i}^2$ across the small epochs ε_i, i.e., $var([\sigma_{\varepsilon_1}^2, \sigma_{\varepsilon_2}^2, \cdots, \sigma_{\varepsilon_6}^2])$.

- f_{13}: The variance of absolute difference of $\sigma_{\varepsilon_i}^2$ across the small epochs ε_i, i.e., $var([|\sigma_{\varepsilon_1}^2 - \sigma_{\varepsilon_2}^2|, |\sigma_{\varepsilon_2}^2 - \sigma_{\varepsilon_3}^2|, \cdots, |\sigma_{\varepsilon_5}^2 - \sigma_{\varepsilon_6}^2|])$.

2) Phase II: Two more features (f_{14}, f_{15}) are extracted by processing the envelope difference, E_D, for each epoch. The last feature, f_{15}, is designed to capture the step-like changes in the envelope at the beginnings and ends of apnea periods. Let N_1 and N_2 be the number of epochs in apnea and normal sets available for training. Let the E_D signals for apnea and normal epochs be represented by the rows of matrices $\widetilde{A}$ and $\widetilde{B}$ ($\widetilde{A} \in \mathbb{R}^{N_1 \times n}$, $\widetilde{B} \in \mathbb{R}^{N_2 \times n}$). We create a matrix $\widetilde{\mathbf{X}}$ by stacking $\widetilde{A}$ on top of $\widetilde{B}$ i.e., $\widetilde{\mathbf{X}}^T = [\widetilde{A}^T \widetilde{B}^T]^T$ ($\widetilde{\mathbf{X}} \in \mathbb{R}^{(N_1 + N_2) \times n}$). Each epoch or row in $\widetilde{\mathbf{X}}$ is denoted by $\widetilde{\mathbf{X}}_i$. We define the average energy E_N of all normal epochs as

$$E_N = \frac{1}{N_2} \sum_{i=1}^{N_2} \left(\sum_{d=1}^{n} ((\widetilde{B}_i[d])^2) \right), \quad (1)$$

where d denotes the sample index. The matrices $\widetilde{A}, \widetilde{B}$ and $\widetilde{\mathbf{X}}$ are normalized by E_N to obtain $A = \widetilde{A}/E_N$, $B = \widetilde{B}/E_N$ and $\mathbf{X} = \widetilde{\mathbf{X}}/E_N$. As apnea epochs are accompanied with a reduction in the energy of respiratory signal, E_N is only derived from normal epochs to enhance the difference between apnea and normal epochs. The value of E_N is obtained from training set and applied to validation and test sets. All the rows B_i in B are averaged to obtain an average envelope difference row vector B_{avg}. The feature f_{14} extracted in this phase is
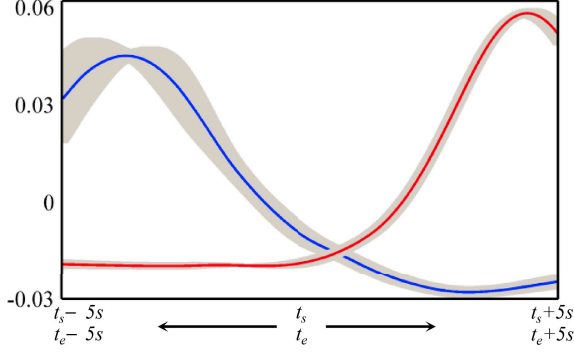
Fig. 3. Loading vectors (v_s^1 and v_e^1) obtained after application of Algorithm 1. The grey back ground indicates the loading vectors (selected in iterations of CV) superimposed on each other. The solid line (blue) represents the mean of all v_s^1 while the solid line (red) represents the mean of all v_e^1.

- f_{14}: For each epoch $\mathbf{X}_i$ in matrix $\mathbf{X}$, sum the squared distance between $\mathbf{X}_i$ and B_{avg}.

For the extraction of f_{15}, two matrices are generated from A using the apnea start time t_s and apnea end time t_e for each apnea event from PSG. The following steps are carried out:

- The E_D signals in A are aligned by apnea start time t_s and a 10-second portion from t_{s-5sec} to t_{s+5sec} is extracted for each signal. Let these extracted signals be represented by the rows of a matrix S_T ($S_T \in \mathbb{R}^{N_1 \times m}$, where m is the number of samples in the 10-second portion ($m < n$)).

- The E_D signals in A are aligned by apnea end time t_e and 10-second portion from t_{e-5sec} to t_{e+5sec} is extracted for each signal and stored as rows of a matrix E_T ($E_T \in \mathbb{R}^{N_1 \times m}$).

- Principal component analysis (PCA) is performed on both S_T and E_T matrices separately. Let V_s be the matrix whose columns $v_s^a[j]$ represent the principal component loading vectors (i.e., the right singular vectors) for S_T and V_e be the matrix whose columns $v_e^b[j]$ represent the loading vectors for E_T obtained after PCA, where a and b indicate the column number and j denotes the sample index. The columns in matrices V_s and V_e are arranged in descending order by the amount of variation they capture in the data set and are statistically orthogonal to each other [21].

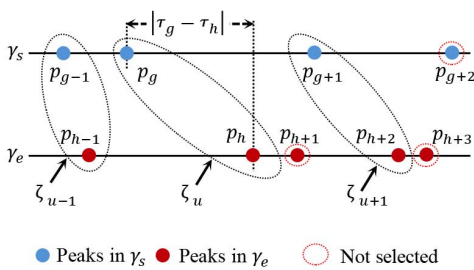The first principal component loading vector, shown in Fig.

Fig. 4. Formation of peak pairs ζ_u from correlation results γ_s and γ_e. The peaks with a dotted red circle do not fulfil the criteria for pair formation.

Algorithm 1

Inputs: v_s^1, v_e^1 and $\mathbf{X}$.
Outputs: The values for feature f_{15}.

1: **for** $i = 1 \rightarrow (N_1 + N_2)$ **do**
2:     **for** $\tau = -\infty \rightarrow \infty$ **do**
3:         $\gamma_s[\tau] = \sum\limits_{j=1}^{m} \mathbf{X}_i[j - \tau]V_s^a[j]$
4:         $\gamma_e[\tau] = \sum\limits_{j=1}^{m} \mathbf{X}_i[j - \tau]V_e^b[j]$
    end for
5:     Find peaks from γ_s and γ_e
6:     Make peak pairs ζ_u from γ_s and γ_e peaks
7:     $f_{15}[i]$ = Sum of peak values of the best pair
    end for
8: **return** f_{15}

3, is selected each from V_s and V_e as it captured maximum variation in the data sets S_T and E_T. Cross-correlation is performed between v_s^1 and the i-th signal $\mathbf{X}_i$ in $\mathbf{X}$ and the result is denoted by γ_s. Similarly, cross-correlation γ_e is performed between v_e^1 and the i-th signal $\mathbf{X}_i$. In general, the cross-correlation γ of any discrete signal x with a discrete signal y is given by

$$\gamma[\tau] = \sum_{j=-\infty}^{\infty} x[j - \tau]y[j], \qquad (2)$$

where τ represents the correlation lag. Once the cross-correlation results γ_s and γ_e are obtained for the i-th signal $\mathbf{X}_i$, peaks are detected in each of these results. Each peak, $p_w(\eta_w, \tau_w)$, is characterized by peak number w, amplitude value η and a position or lag value in terms of τ. Peak pairs are then formed by choosing one peak from γ_s and one peak from γ_e as shown in Fig. 4. For a g-th peak $p_g(\eta_g, \tau_g)$ selected from γ_s, a peak $p_h(\eta_h, \tau_h)$ is selected from γ_e to form a pair such that p_h is the first peak in γ_e which is located immediately after p_g and before p_{g+1} in terms of position. If this condition is satisfied, then a peak pair ζ_u is formed. If no p_h from γ_e satisfies the above criteria, then p_g is dropped and p_{g+1} is considered for pair formation. The peaks which are used once in pair formation do not participate in any other pair while the peaks which do not take part in pair formation are ignored. Once all possible peak pairs are formed, the pair that has maximum separation between the constituent peaks is selected and the sum of peak values is stored as feature f_{15}, i.e.,

- $f_{15} = \{\eta_g + \eta_h : \eta_g \in \zeta_g, \eta_h \in \zeta_h, max|\tau_g - \tau_h|\}$.

The above process is repeated for remaining signals in $\mathbf{X}$ and is described in Algorithm 1. The vectors v_s^1 and v_e^1 are obtained from training data and are applied to validation and test data for extraction of f_{15}. The above features $f_1 - f_{15}$, extracted from normal and apnea epochs, are standardized by subtracting mean and dividing by standard deviation before being used with a classifier.

D. Classification

1) Classification Algorithm.: The Linear Discriminant (LD) classifier [22] is considered for the classification task in this study. The LD classifier assumes that the class conditional

densities follow multivariate Gaussian distributions and the covariance matrix (Σ) is the same for each class. Let the p-th feature vector belonging to class c ($c = 1, 2$ in binary classification) be denoted by $\mathbf{x}_{p,c}$. If there are N_c number of observations present for each class and N is the total number of observations for all classes, then the prior probability π_c for each class will be $\pi_c = \frac{N_c}{N}$ and the conditional mean vector μ_c for each class will be $\mu_c = (\sum_{p=1}^{N_c} \mathbf{x}_{p,c})/N_c$. The covariance matrix is given by

$$\Sigma = \frac{1}{N - C} \sum_{c=1}^{C} \sum_{p=1}^{N_c} (\mathbf{x}_{p,c} - \mu_c)(\mathbf{x}_{p,c} - \mu_c)^T. \quad (3)$$

For a given object with feature vector $\mathbf{v}$ in the validation or test set, the class label is determined by calculating the discriminant value δ_c for each class. The class with the highest discriminant value is then assigned to the object. The discriminant value for the LD classifier is given by

$$\delta_c = \log(\pi_c) + \mathbf{v}^T \Sigma^{-1} \mu_c - \frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c. \quad (4)$$

2) Training & Testing: The total number of apnea epochs (N_1) in this paper, both OA and OH, were more than twice the number of normal epochs (N_2) as data was collected from subjects with very high AHI in this initial study. So the following steps involving cross-validation [23], as shown in Fig. 5, are carried out for data division to make sure that approximately equal number of apnea and normal epochs are available for training:

- *Step* 1: Let the total number of normal, OA and OH epochs be represented by r, s and t respectively ($s + t = N_1$ and $r = N_2$). The sets of both OA and OH epochs are each randomly partitioned into two subsets. Let s_1 and s_2 be the number of epochs in the two subsets obtained from OA set and let t_1 and t_2 be the number of epochs in the two subsets obtained from OH set such that $s_1 + t_1 \approx r$, i.e., $s_1 \approx t_1 \approx \frac{r}{2}$, $s_2 = s - s_1$ and $t_2 = t - t_1$.

- *Step* 2: In this step, 5-fold cross-validation (CV) is used to train and test the classifier using the normal epochs (r) and apnea epochs ($s_1 + t_1$). In 5-fold CV, the apnea epochs and normal epochs are each divided into 5 folds. Four folds from each class are used in training the classifier and the remaining fold from each class is used as validation set until each fold from both classes is used once for validation. The feature subset selection [24] is also done in this step. The performance of the classifier with all possible feature subsets is also analyzed on the validation fold [25]. The feature subset with least number of features giving the maximum accuracy for the validation fold is then selected from all possible subsets. Since 5-fold CV is employed, the feature subset selection will yield 5 subsets for each of the 5 iterations of CV. The subset that provides best value for accuracy is chosen from these 5 subsets. The values of sensitivity, specificity and accuracy [26] obtained with the complete set of features and with feature subsets are averaged across the 5 iterations of 5-fold CV.
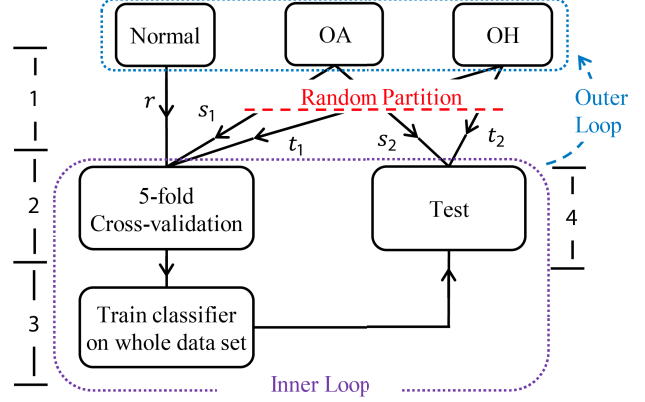

Fig. 5. Flow chart for division of data. $s_1 \approx t_1 \approx \frac{r}{2}$ and $s_1 + t_1 \approx r$.

- *Step* 3: The classifier is trained on the complete data set available in *Step* 2 with the complete bag of features and also with the best feature subset selected from the 5 subsets obtained in 5-fold CV.

- *Step* 4: The classifier is used on the remaining OA and OH epochs, $s_2 + t_2$. Only the value of sensitivity for detecting apnea epochs is calculated.

The above steps, *Step* 1 through *Step* 4, are repeated 10 times making the whole process as 10×5-fold CV. The values of sensitivity, specificity and accuracy obtained in 5-fold CV process of *Step* 2 and sensitivity values in test phase of *Step* 4 are averaged over 10 iterations (outer loop in Fig. 5).

IV. RESULTS & DISCUSSION

Data was collected from 4 subjects (3 male and 1 female, 48 ± 6.9 years, 210 ± 20.5 lbs and AHI 49 ± 29) who were previously diagnosed with sleep apnea. Full night recordings (6 to 7 hours) were obtained and after pre-processing, motion detection and removal, 476 OA, 392 OH and 361 normal epochs were extracted from the recordings. The data was partitioned according to Fig. 5. The results of classification between *Apnea* and *Normal* classes are summarized in Table I. The classification results with all features is indicated by "All" in Table I. The selected feature subsets have different number of features in each iteration. The average number of features rounded to nearest integer in the selected subsets is shown in Table I. The values for sensitivity, specificity and accuracy in the table are for the data in validation folds of cross-validation phase and averaged through the iterations of 10×5-fold CV. Similarly the sensitivity for epochs in *Step* 4 (listed in the last column of Table I) is averaged by the number of iterations in the outer loop.

The results indicate that the classfier when used with complete bag of features has an overall sensitivity of 64.6% and specificity of 64% during the cross-validation steps and a sensitivity of 66% for the epochs in the test folds of *Step* 4 in Fig. 5. There is an increase in overall sensitivity, specificity and accuracy when the LD classifier is used after feature subset selection. The number of features selected is almost half the total number of features as the feature subset with lowest number of features is selected during feature selection.

TABLE I.   CLASSIFICATION RESULTS FOR APNEA (OA & OH) AND NORMAL EPOCHS

| No. of f_k | Avg. results cross-validation (validation folds) | | | Test data |
	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)
All	64.6	64	65.6	66
7	71.2	70.8	73.1	67

V. CONCLUSION

This paper explored the effectiveness of an under-the-mattress IR-UWB radar for detection of sleep apnea events using a machine learning approach. 15 features were considered, including several novel ones, and an LD classifier was trained and tested with 5-fold cross-validation on over 1,200 60-second epochs collected from 4 subjects. The system showed a classification accuracy of 73% which is not considered good enough for a radar-only device.

The on-going work is focused on addition of other under-the-mattress sensors (microphones) to the IR-UWB radar system, to give more information relating to airflow, which is expected to improve accuracy.

ACKNOWLEDGEMENTS

REFERENCES

[1] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, C. L. Marcus, R. Mehra, S. Parthasarathy, S. F. Quan *et al.*, "Rules for scoring respiratory events in sleep: update of the 2007 aasm manual for the scoring of sleep and associated events," *J Clin Sleep Med*, vol. 8, no. 5, pp. 597–619, 2012.

[2] N. M. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proceedings of the American Thoracic Society*, vol. 5, no. 2, pp. 136–143, 2008.

[3] G. T. O'Connor, B. Caffo, A. B. Newman, S. F. Quan, D. M. Rapoport, S. Redline, H. E. Resnick, J. Samet, and E. Shahar, "Prospective study of sleep-disordered breathing and hypertension: the sleep heart health study," *American journal of respiratory and critical care medicine*, vol. 179, no. 12, pp. 1159–1164, 2009.

[4] D. J. Gottlieb, G. Yenokyan, A. B. Newman, G. T. O'Connor, N. M. Punjabi, S. F. Quan, S. Redline, H. E. Resnick, E. K. Tong, M. Diener-West *et al.*, "Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure the sleep heart health study," *Circulation*, vol. 122, no. 4, pp. 352–360, 2010.

[5] A. Lazaro, D. Girbau, and R. Villarino, "Analysis of vital signs monitoring using an ir-uwb radar," *Progress In Electromagnetics Research*, vol. 100, pp. 265–284, 2010.

[6] E. Staderini, "Uwb radars in medicine," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 17, no. 1, pp. 13–18, Jan 2002.

[7] V. Nguyen, A. Javaid, and M. Weitnauer, "Spectrum-averaged harmonic path (shapa) algorithm for non-contact vital sign monitoring with ultra-wideband (uwb) radar," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, Aug 2014, pp. 2241–2244.

[8] D. Pevernagie, R. M. Aarts, and M. De Meyer, "The acoustics of snoring," *Sleep medicine reviews*, vol. 14, no. 2, pp. 131–144, 2010.

[9] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 6, pp. 686–696, 2003.

[10] A. H. Khandoker, M. Palaniswami, and C. K. Karmakar, "Support vector machines for automated recognition of obstructive sleep apnea syndrome from ecg recordings," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 1, pp. 37–48, 2009.

[11] F. C. Commission *et al.*, "Revision of part 15 of the commissions rules regarding ultra-wideband transmission systems, first report and order, fcc 02," *V48, April*, 2002.

[12] Y. S. Lee, P. Pathirana, C. Steinfort, and T. Caelli, "Monitoring and analysis of respiratory patterns using microwave doppler radar," *Translational Engineering in Health and Medicine, IEEE Journal of*, vol. 2, pp. 1–12, 2014.

[13] M. Kagawa, K. Ueki, H. Tojima, and T. Matsui, "Noncontact screening system with two microwave radars for the diagnosis of sleep apnea-hypopnea syndrome," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 2052–2055.

[14] P. de Chazal, E. O'Hare, N. Fox, and C. Heneghan, "Assessment of sleep/wake patterns using a non-contact biomotion sensor," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 2008, pp. 514–517.

[15] A. Zaffaroni, P. de Chazal, C. Heneghan, P. Boyle, P. R. Mppm, and W. T. McNicholas, "Sleepminder: an innovative contact-free device for the estimation of the apnoea-hypopnoea index," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 7091–9094.

[16] J. C. Lai, Y. Xu, E. Gunawan, E. Chua, A. Maskooki, Y. L. Guan, K.-S. Low, C. B. Soh, and C.-L. Poh, "Wireless sensing of human respiratory parameters by low-power ultrawideband impulse radio radar," *Instrumentation and Measurement, IEEE Transactions on*, vol. 60, no. 3, pp. 928–938, 2011.

[17] W. T. McNicholas, "Diagnosis of obstructive sleep apnea in adults," *Proceedings of the American thoracic society*, vol. 5, no. 2, pp. 154–160, 2008.

[18] Neurotrials Research, Inc. [Online]. Available: http://www.neurotrials.com/

[19] S. Foo, "Ultra wideband monitoring systems and antennas," Mar. 27 2013, uS Patent App. 13/851,287.

[20] B. Yιlmaz, M. H. Asyalι, E. Arιkan, S. Yetkin, and F. Özgen, "Sleep stage and obstructive apneaic epoch classification using single-lead ecg," *Biomedical engineering online*, vol. 9, p. 39, 2010.

[21] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.

[22] B. D. Ripley, *Pattern recognition and neural networks*. Cambridge university press, 1996.

[23] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 111–147, 1974.

[24] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[25] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, no. 1, pp. 273–324, 1997.

[26] A. G. Lalkhen and A. McCluskey, "Clinical tests: sensitivity and specificity," *Continuing Education in Anaesthesia, Critical Care & Pain*, vol. 8, no. 6, pp. 221–223, 2008.