Real-time Sleep Apnea Detection with One-channel ECG Based on Edge Computing Paradigm

4 authors:

Aleksandar Stojanovski
University "St. Kliment Ohridski" - Bitola
2 PUBLICATIONS **2** CITATIONS

SEE PROFILE

Eftim Zdravevski
Ss. Cyril and Methodius University in Skopje
157 PUBLICATIONS **1,437** CITATIONS

SEE PROFILE

Saso Koceski
Goce Delcev University of Štip
96 PUBLICATIONS **632** CITATIONS

SEE PROFILE

Vladimir Trajkovik
Ss. Cyril and Methodius University in Skopje
275 PUBLICATIONS **1,429** CITATIONS

SEE PROFILE

Real-time sleep apnea detection with one-channel ECG based on edge computing paradigm

Aleksandar Stojanovski, Eftim Zdravevski *, Saso Koceski, and Vladimir Trajkovik

Faculty of Computer Science and Engineering, Sts. Cyril and Methodius University, Skopje, Macedonia
Goce Delcev University, Stip, Macedonia

Abstract. Sleep apnea is a disorder that causes people to stop breathing multiple times during their sleep, when untreated. It can be diagnosed trough polysomnography (PSG), which is a time consuming, expensive and must be performed in special laboratories. Due to its complexity, different alternatives to PSG have been developed. This paper presents a system based on the edge-computing paradigm for detection and alerting of sleep apnea events, using data from a single-channel ECG sensor. A framework for automated feature selection is used for the extraction and selection of the important features. Some ECG signal specific features were also added to the generic framework. We have evaluated several machine learning algorithms for sleep apnea detection based on the generic features and the ECG-specific features on a dataset containing 70 recordings, available in the PhysioNet database. The obtained results show that the combination of generic features and ECG-specific features improve the detection accuracy to up to 82% with a small set of about 20 computationally efficient features.

Keywords: Sleep apnea, PhysioNet, ECG, QRS, feature extraction

1 Introduction

Obstructive sleep apnea (OSA) is a sleep-related breathing disorder that causes upper airway occlusion during sleep. The reported prevalence is 4% in adult men and 2% in adult women [1, 2]. Currently there is no home appliance that is capable of diagnosing sleep apnea and this is why it often is undiagnosed. The most common effect of sleep apnea is excessive daytime sleepiness. It can also increase the risk of developing cardiovascular (CV) diseases [3, 4, 5, 6, 7].

The relationship between OSA and CV disease has been examined by a large number of community-based studies with smaller number of severe cases of OSA. On the other hand, clinically based studies with individuals of higher OSA severity record only a small number of events and are not able to provide enough variables for the models

* This work was partially financed by the Faculty of Computer Science and Engineering at the Sts. Cyril and Methodius University, Skopje, Macedonia. We also acknowledge the support of Microsoft Azure for Research through a grant providing computational resources for this work.

[8]. This is mostly caused by the inability to perform a long-term follow-up on the patients [2]. Treatment using nasal continuous positive airway pressure (NCPAP) in the early stages of the disease could reduce adverse health effects [9].

The degree of severity of the syndrome is measured by using the apnea-hypopnea index (AHI). The AHI represents the number of apnea-hypopnea events per hour of sleep. The OSAH is classified as normal, mild, moderate or severe if belongs to the interval [0; 5), [5; 15), [15; 30), or over 30, respectively.

The sleep quality improvement and heart monitoring technologies are actively adopted by companies and the market for such devices and software is rising. Some of the commercially available products have the ability to achieve up to 97% accuracy in detecting paroxysmal atrial fibrillation (AF), the most common type of heart rhythm problem. The commercially available systems are designed to provide biometric feedback to patients and doctors.

With the emergence of the edge computing paradigm and the development of cheap and portable ECG devices, a new opportunity arises to detect sleep apnea in real-time without the need of patient hospitalization. In this paper we propose an edge-computing architecture for real time detection of sleep apnea. We evaluate the feasibility of the architecture with experiments from a real dataset from PhysioNet by developing different machine-learning based models for feature extraction and classification. The main challenge we are addressing is identifying computationally efficient features that can be computed on devices with lower computational power and battery capacity, such as portable ECG devices or smart phones.

The remainder of this paper is organized as follows: in section 2 we provide a brief review of related works for sleep apnea detection and then, in section 3, we present the proposed approach. Next, in 4 the the system architecture is presented. Subsequently, in section 5 we we describe the dataset used for testing of the proposed approaches and discuss the obtained results. Finally, in section 6 we conclude the paper.

2 Related work

The traditional models for sleep studies include recordings of encephalography (EEG), electro-oculography (EOG), electromyography (EMG), electrocardiography (ECG), respiratory effort, oxygen saturation and oronasal airflow [10]. These studies are complex and require sleeping in specially equipped rooms in hospitals. There are improvements by the introduction of specialized wearable sensors, like Zephyr BioHarness [11], for monitoring and tracking multiple biometrics. Most of these sensors are used for monitoring. Nonetheless, statistical models or machine learning algorithms can be obtained on the obtained signals aiming to provide diagnostics aid.

In the last decade machine learning algorithms are often used for processing of large volumes of bio-physiological data[12, 13]. However, there are quite a few examples in the literature where machine learning is used for sleep apnea detection. Authors of [14] present a systematic review of classification techniques for prediction and detection of sleep apnea. In [15] a ubiquitous sensor system for sleep monitoring is presented. An approach for sleep apnea detection from an ECG signal based on feed-forward artificial neural networks is presented in [16].

The approach proposed in this paper uses machine learning and edge-computing technologies for real-time analysis and detection of sleep apnea. Our approach has the ability to monitor multiple parameters via continuous ECG processing and issue notifications and alarms.

3 Machine Learning Based Apnea Detection

We developed our approach using the automatic feature engineering and selection framework presented in [17, 18]. The proposed framework can be used regardless of the number of sensors, their type or their body placement location in case of body-worn sensors.

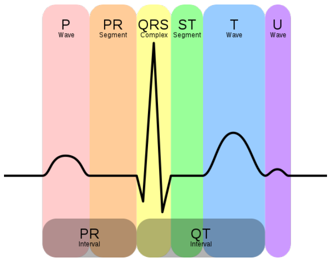

Fig. 1. Electrocardiogram (ECG) wave features and Intervals defined for PQRST wave.

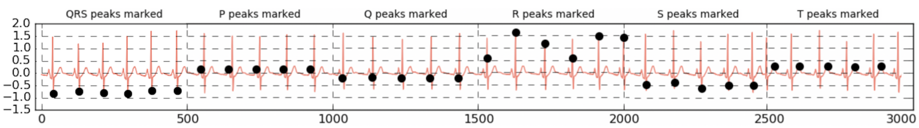

Fig. 2. QRS, P, Q, R, S and T characteristic peaks during 30 seconds of filtered ECG signal

3.1 Generic domain-independent feature extraction and selection

The framework uses a robust process of feature extraction that is executed in several steps. First, the data is treated like a stream and preprocessed by generating sliding window batches with a duration of 60 seconds without overlapping. From this series, multiple statistical measurements are calculated, such as minimum, maximum, range, mean, standard deviation, skewness, kurtosis, energy per sample, equal width histograms, quantile-based features, correlation-based features, linear and quadratic fit coefficients, etc.

Furthermore, from the original time series, *first derivatives* time series is generated and *delta series* based on the relative deviation from the mean value of the readings within one window. Additionally, series derived from Fast Fourier Transformation (FFT), frequencies, amplitudes and magnitudes are generated. For all of these series, multiple statistical features are extracted, similar to the ones computed from the original ECG time series. For the detailed feature engineering process we refer the reader to [17, 18].

3.2 Generating additional Electrocardiogram (ECG) wave features

An electrocardiogram (ECG) is a graph depicting the electrical activity of the heart. An ECG wave is a periodic wave. During each period, the wave is consisted of a P wave, QRS complex and a T wave. The importance of the analysis of the ECG signal has been recognized in the literature and so is the extraction of features [19, 20]. The important ECG wave features are shown in Fig. 1.

To extract the features from the ECG signals we are using the Python Online and Offline ECG QRS Detector based on the Pan-Tomkins algorithm [21, 22]. We use the detector with slight modification to detect the QRS, PR, QRS, QT distances between peaks as features. We also include the Beats per minute (BPM) as a feature. The characteristic peaks of the ECG signal obtained during a non-apnea segment of 30 seconds are marked in Fig. 2.

After the distances between the peaks and the BPM are calculated, we extract the median, mean, maximum value, minimum value, standard deviation and the skewness from the series of distances between consecutive peaks from the same type. Thus, we obtain $8 \times 6 = 48$ ECG-specific features. Then, these features are added to the full feature set along with the generic features.

3.3 Feature selection

After all of the features are generated, for each feature the framework estimates the *feature importance*. All estimations are performed using a Random Forest classifier with 1000 trees and using its feature importance estimates. In addition to the importance, the framework also calculates the concept distribution drift sensitivity of each feature, as described in [17]. The features that are selected for the classification need to have high importance and low drift sensitivity. A grid search with Random Forest is performed to select the optimal combination of features from the calculated set.

With the feature selection the number of features is reduced to obtain more robust models and to shorten the model building and recognition time. After the feature selection process, using several machine learning algorithms, we generate classification models using the reduced feature sets.

A prediction model is defined by the feature subset, classification algorithm, and algorithm parameters. The evaluated Classification algorithms include: Random Forest, Extremely Randomized Trees, Support Vector Machines (SVM), Nave Bayes, Ada Boost, Logistic regression, kNN.

The following section describes how the proposed approach for sleep apnea detection can be put into context of an end-to-end production system deployed in a cloud.

4 System architecture

In this section we describe the architecture of the system for sleep apnea detection that uses the proposed framework for feature extraction and machine learning, but also leverages the principles of edge computing. The system is composed of two parts that work together: sensor communicating with mobile application, and web-based application deployed in the cloud. The general system architecture is shown in Fig 3.

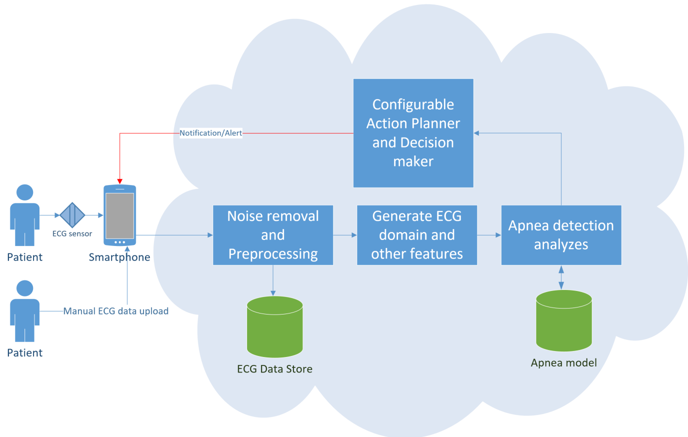

Fig. 3. General system architecture

4.1 Edge component - sensor and mobile application

During the night, when the patient is sleeping, the ECG sensor is attached to the patient's body and collects ECG data. In the off-line mode, the smart phone computes features, and utilizes the already built models to analyze and detect sleep apnea. If apnea is detected, this is recorded and also an alarm (e.g. visual, sound or vibrating) can be issued so the patient wakes up and puts a CPAP mask or other specially designed oral appliances, so he/she continue sleeping normally. In an on-line mode, the smart phone transfers all or some subset of collected data to the cloud, so it can be used for recalibrating the models and evaluating the performance of the classification models. Occasionally, when updates of the classification model are available, they could be pushed to the smart phone, so the performance can be improved in the future.

4.2 Cloud application

The cloud application is the location where the machine learning based processing and analyses is performed. The sleep apnea detection models are trained based on the PhysioNet data, but also on new data that is collected from patients. Using a personalized patient-centric approach, models can be individually calibrated to tune decision thresholds and optimize detection performance for each specific patient. After enough ECG data has been collected the system can upgraded by using more robust deep learning approaches to achieve long-term learning and better results. The cloud application also allows high-level management and monitoring of the whole system, such as: monitoring predictive performance on user-level, configuring notification system, meta-analysis of the whole performance, issuing updates of the models, etc.

5 Results

5.1 Methodology

For evaluating the approach, we use the PhisioNet database [23], an online database that contains a large collection of physiologic signal recordings. In our experiments we use 70 sleep recordings that contain a single channel ECG with sampling rate of 100Hz. There are 283 hours of train data, which is split into two distinct subsets for training and validation of 172 and 111 hours respectively, and 288 hours of independent test data.

According to the Physionet paper [24], all sleep recordings are taken from 32 subjects. From them 25 are male and 7 are female. There is only 1 recording for 4 of the subjects, 22 subjects contributed with two recordings each, 2 subjects contributed with 3 recordings each, and 4 subjects contributed 4 recordings each. In total there are 70 recordings which were divided into a equally-sized training and test sets.

The selection procedure for the training and test sets is the following. The recordings were ordered based on the number of minutes with apnea. From those 70 recordings, from each set is chosen pair of recordings which is randomly picked. One of these recordings is assigned to the training set, the other recording is assigned to the test set. With this approach, the distribution of apnea durations was approximately equal in the training and test sets, each of them containing about 35 recordings.

Each minute of recordings is scored by experts, on the basis of the respiratory and oxygen saturation signals, using amplitude criteria for airflow and desaturation. No differentiation between apnea and hypopnea events is made when events of disordered breathing were scored [25].

The apnea recordings were arranged in three groups:

- Group A (apnea): recordings with strong, clear occurrence of sleep apnea (more than 100min). With this criterion there are forty recordings that belong to this group.
- Group B (borderline): recordings with some degree of sleep apnea (between 5 and 100 minutes). The recordings revealed either mild apnea or obstructive snoring in otherwise healthy subjects. With this criterion there are 10 recordings that belong to this group.
- Group C (control): recordings of healthy subjects. These subjects did not have sleep apnea (they may have fewer than 5 min of apnea) nor habitual snoring. There are 20 recordings that belong to this group.

Considering that the labels from group B were very rare, we performed the following mapping. For patients of group B the minutes during which an apnea was detected

(Group B) were considered to be of apnea recordings, otherwise normal. Likewise, group C were considered as normal (non-apnea). Thus, a binary classification problem was obtained, with a goal to classify whether a minute of recording corresponds to a period when apnea was present.

5.2 Discussion

Throughout the experiments we used the training and testing subsets of the dataset. In particular, we divided the training subset in training and validation subsets, which were used for feature selection and classifier parameters fine-tuning. After that we evaluated the best feature sets for each classifier on the independent test subset.

In Table 1 the maximum accuracy is presented, as obtained for each classifier on both the validation and test subset when using generic and ECG-specific features in the feature selection and model training. In Table 2 we present the classification accuracy of each classifier when using only the generic features.

Table 1. Maximally obtained accuracies on the Validation and Test subsets based on all features

Classification algorithm	Validation set	Test set
AdaBoost	0.761505	0.685554
ExtremelyRandomizedTrees	0.814421	0.773241
kNN	0.783841	0.730383
LogisticRegression	0.797482	0.720582
NaiveBayes	0.756558	0.714725
RandomForest	0.804377	0.777649
SVM	**0.825663**	**0.783680**

Table 2. Maximally obtained accuracies on the Validation and Test subsets without the ECG specific features

Classification algorithm	Validation set	Test set
AdaBoost	0.760006	0.685554
ExtremelyRandomizedTrees	0.814421	0.773241
kNN	0.772748	0.720640
LogisticRegression	0.789387	0.720583
NaiveBayes	0.756558	0.681668
RandomForest	0.804377	0.767442
SVM	**0.822965**	**0.780259**

It can be observed that the best results are obtained by using the SVM classifier. For the best accuracy the parameters of the SVM classifier are $C = 100$ and $\gamma = 0.0001$, while using radial based function kernel. The best results are consistent when testing on both the validation and the test. Similarly, Random Forest and Extremely Randomized Trees performed considerably well. This is important as their models are very lightweight and could be encoded as a set of if-then-else rules, thus applicable for computation on mobile devices. The results are about 1% worse than the ones in the PhysioNet approach [24], which uses a specialized algorithm for sleep apnea detection that is more computationally intensive.

Interestingly, even with thousands of iterations and different feature sets evaluated on the validation set, the optimal feature sets contained no more than 20 features. This is important because shows that even a small number of computationally efficient generic features provide substantial results in sleep apnea detection. We reason that domain specific features were not as useful as one might expect is because there are a lot of subjects, with multiple hours of recordings during which the electrode of the ECG device might have been detached. Even if this was not the case, the Pan-Tomkins algorithm sometimes could not detect the characteristic peaks during the window, thus limiting the ability to derive valuable domain-specific features. In such cases, some or even all domain specific features were encoded as 0. To improve upon this, a future work is needed where different detection thresholds could be tuned within the Pan-Tomkins algorithm. It is worth noting that the most useful features selected for the best models contained about 20 features, among which were: minimum, maximum, mean, skewness, kurtosis, standard deviation and percentiles of first derivatives, raw ECG signal, and beats-perminutes signal. All of these are computationally efficient and can be computed in real time on the mobile devices.

For sleep apnea detection, a direct airflow sensor is the most valuable, which means that we should expect that extracted features that relate to respiratory signal should have the biggest influence. To capture the respiration signal from ECG we need to use a technique called EDR (ECG-Derived Respiration). Another technique includes observing the beat-to-beat variations in RR intervals or their reciprocals, which are primarily due to respiratory sinus arrhythmia (RSA) in most individuals. Therefore, it can be expected the BPM rate and the RR distance to be most influential. In Table 3 are presented the 10 most important ECG features that were identified in the study. Most of them are the first derivatives of the features that were extracted from the one channel ECG. According to [3], the first derivative emphasizes the high frequency components of the ECG and firmly establishes the fact that these are very real and significant components of the ECG.

Table 3 confirms that for sleep apnea biggest influence has the respiratory signal component of the ECG. Particularly respiratory rate features like BPM and RR. This are even more pronounced in the first derivative component of these two features, as it can be noted from Table 3.

In future works, in addition to the ECG data, some other information for the patients could prove useful (e.g. disease history, gender, age). It could be used for segmenting users and providing more personalized models, as in [26], or to use this nominal data as features in the models [27].

Table 3. The 10 most important ECG-specific features

Feature Name	Score
First derivative of ECG minimum	0.02036
First derivative of ECG BPM rate	0.02022
First derivative of ECG RR distance	0.01787
First derivative of ECG skewness	0.01703
First derivative of ECG percentile 60	0.01532
ECG minimum	0.01491
First derivative of ECG percentile 95	0.01373
Delta between ECG RR difference	0.01351
First derivative of ECG QT distance	0.01335

6 Conclusion

We can also observe that on most classification algorithms, the introduction of the ECG signal specific features improves the classification accuracy on the validation dataset, the test dataset or on both of them. The improvement in accuracy is not always that substantial and for some of the classifiers it is insignificant. This shows that the used general framework for feature generation and selection gives good classification performance for ECG signals even without prior domain knowledge. For some of the classifiers the results were the same which is a direct consequence of signal specific features generated from the ECG not been selected as top picks during the feature selection phase. Further investigation is needed to include even more domain specific features and further enhance the classification performance, which is on similar level as state-of-the-art approaches. The proposed system architecture uses the edge-computing paradigm to facilitate real-time detection of sleep apnea, calibration of models and collecting data for further analysis and improvement of models.

References

[1] Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., Badr, S.: The occurrence of sleep-disordered breathing among middle-aged adults. New England Journal of Medicine **328**(17) (1993) 1230–1235

[2] Young, T., Peppard, P., Palta, M., Hla, K.M., Finn, L., Morgan, B., Skatrud, J.: Population-based study of sleep-disordered breathing as a risk factor for hypertension. Archives of internal medicine **157**(15) (1997) 1746–1752

[3] Force, A.O.S.A.T., of Sleep Medicine, A.A., et al.: Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine **5**(3) (2009) 263

[4] Leung, R.S., Comondore, V.R., Ryan, C.M., Stevens, D.: Mechanisms of sleep-disordered breathing: causes and consequences. Pflügers Archiv-European Journal of Physiology **463**(1) (2012) 213–230

[5] Bradley, T.D., Floras, J.S.: Obstructive sleep apnoea and its cardiovascular consequences. The Lancet **373**(9657) (2009) 82–93
[6] Kohler, M., Stradling, J.R.: Mechanisms of vascular damage in obstructive sleep apnea. Nature Reviews Cardiology **7**(12) (2010) nrcardio–2010
[7] Chami, H.A., Fontes, J.D., Vasan, R.S., Keaney Jr, J.F., O'Connor, G.T., Larson, M.G., Benjamin, E.J., Gottlieb, D.J.: Vascular inflammation and sleep disordered breathing in a community-based cohort. Sleep **36**(5) (2013) 763–768
[8] Kendzerska, T., Gershon, A.S., Hawker, G., Leung, R.S., Tomlinson, G.: Obstructive sleep apnea and risk of cardiovascular events and all-cause mortality: a decade-long historical cohort study. PLoS medicine **11**(2) (2014) e1001599
[9] Dimsdale, J.E., Loredo, J.S., Profant, J.: Effect of continuous positive airway pressure on blood pressure: a placebo trial. Hypertension **35**(1) (2000) 144–147
[10] Flemons, W., Buysse, D., Redline, S., Oack, A., Strohl, K., Wheatley, J., Young, T., Douglas, N., Levy, P., McNicolas, W., et al.: Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. Sleep **22**(5) (1999) 667–689
[11] Hailstone, J., Kilding, A.E.: Reliability and validity of the zephyr bioharness to measure respiratory responses to exercise. Measurement in Physical Education and Exercise Science **15**(4) (2011) 293–300
[12] Lameski, P., Zdravevski, E., Koceski, S., Kulakov, A., Trajkovik, V.: Suppression of intensive care unit false alarms based on the arterial blood pressure signal. IEEE Access **5** (2017) 5829–5836
[13] Bhattacharyya, A., Pachori, R.B., Upadhyay, A., Acharya, U.R.: Tunable-q wavelet transform based multiscale entropy measure for automated classification of epileptic eeg signals. Applied Sciences **7**(4) (2017) 385
[14] Pombo, N., Garcia, N., Bousson, K.: Classification techniques on computerized systems to predict and/or to detect apnea: A systematic review. Computer methods and programs in biomedicine **140** (2017) 265–274
[15] Pombo, N., Garcia, N.M.: ubisleep: An ubiquitous sensor system for sleep monitoring. In: Wireless and Mobile Computing, Networking and Communications (WiMob), 2016 IEEE 12th International Conference on, IEEE (2016) 1–4
[16] da Silva Pinho, A.M., Pombo, N., Garcia, N.M.: Sleep apnea detection using a feed-forward neural network on ecg signal. In: e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on, IEEE (2016) 1–6
[17] Zdravevski, E., Lameski, P., Trajkovik, V., Kulakov, A., Chorbev, I., Goleva, R., Pombo, N., Garcia, N.: Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering. IEEE Access **5** (2017) 5262–5280
[18] Zdravevski, E., Stojkoska, B.R., Standl, M., Schulz, H.: Automatic machine-learning based identification of jogging periods from accelerometer measurements of adolescents under field conditions. PloS one **12**(9) (2017) e0184216
[19] Gottlieb, D.J., Yenokyan, G., Newman, A.B., O'Connor, G.T., Punjabi, N.M., Quan, S.F., Redline, S., Resnick, H.E., Tong, E.K., Diener-West, M., et al.: Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study. Circulation **122**(4) (2010) 352–360

[20] McNames, J., Fraser, A.: Obstructive sleep apnea classification based on spectrogram patterns in the electrocardiogram. In: Computers in Cardiology 2000, IEEE (2000) 749–752
[21] Sznajder, M., ukowska, M.: Python Online and Offline ECG QRS Detector based on the Pan-Tomkins algorithm (May 2017)
[22] Pan, J., Tompkins, W.J.: A real-time qrs detection algorithm. IEEE transactions on biomedical engineering (3) (1985) 230–236
[23] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation **101**(23) (2000 (June 13)) e215–e220 Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.
[24] Mietus, J.E., Peng, C.K., Ivanov, P.C., Goldberger, A.L.: Detection of obstructive sleep apnea from cardiac interbeat interval time series. In: Computers in Cardiology 2000. Vol.27 (Cat. 00CH37163). (2000) 753–756
[25] Penzel, T., McNames, J., De Chazal, P., Raymond, B., Murray, A., Moody, G.: Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. Medical and Biological Engineering and Computing **40**(4) (2002) 402–407
[26] Zdravevski, E., Lameski, P., Trajkovik, V., Pombo, N., Garcia, N.: Importance of personalized health-care models: A case study in activity recognition. Studies in health technology and informatics **249** (2018) 185188
[27] Zdravevski, E., Lameski, P., Kulakov, A.: Weight of evidence as a tool for attribute transformation in the preprocessing stage of supervised learning algorithms. In: Neural Networks (IJCNN), The 2011 International Joint Conference on, IEEE (2011) 181–188