Detecting breathing and snoring episodes using a wireless tracheal sensor - a feasibility study

Marcel Młyńczak, Student Member, IEEE, Ewa Migacz, Maciej Migacz and Wojciech Kukwa*

Abstract—Objective: Sleep-disordered breathing is both a clinical and a social problem. This implies the need for convenient solutions to simplify screening and diagnosis. The aim of the study was to investigate the sensitivity and specificity of a novel wireless system in detecting breathing and snoring episodes during sleep. Methods: A wireless acoustic sensor was elaborated and implemented. Segmentation (based on spectral thresholding and heuristics) and classification of all breathing episodes during recording were implemented through a mobile application. The system was evaluated on 1,520 manually labeled episodes registered from 40 real-world, whole-night recordings of 16 generally healthy subjects. Results: The differentiation between normal breathing and snoring had 88.8% accuracy. As the system is intended for screening, high specificity of 95% is reported. Conclusions: The system is a compromise between non-medical phone applications and medical sleep studies. The presented approach enables the study to be repetitive, personal, and inexpensive. It has additional value in the form of well-recorded data which are reliable and comparable. Significance: The system opens unexplored possibilities in sleep monitoring and study enabling a multi-night recording strategy involving the collection and analysis of abundant data from thousands of people.

Index Terms—sleep breathing disorders, snoring, tracheal sound analysis, machine learning, smartphone application

I. INTRODUCTION

Snoring is the most common breathing disorder during sleep. It may be either episodic or habitual (HS - habitual snoring). It is the most significant single symptom of all other sleep breathing disorders, such as upper airway resistance syndrome (UARS) or obstructive sleep apnea (OSA). This is why snoring is the primary symptom asked about in every questionnaire focused on sleep-disordered breathing (SDB) diagnosis. Paradoxically, while the sound of snoring is the hallmark of OSA and each patient is asked about snoring intensity and frequency, polysomnography (the gold standard sleep study), does not analyze snoring events precisely in most cases [1].

HS is linked to cardiovascular complications like atherosclerosis [2], [3] and hypertension [4]. Other possible adverse effects of HS include daytime sleepiness [5], and progression of upper airway collapsability [6]. According to others, snoring is not associated with medical hazard in adults [7]. In children, habitual snoring without obstructive sleep apnea is a known risk factor for cardiovascular and neurobehavioral disorders. Children with HS experience cognitive and often heightened behavioral deficits similar to children with OSA, despite the absence of recognized intermittent hypoxia or repeated arousals [8]. Nighttime diastolic blood pressure is significantly higher in children with HS compared with controls after adjusting for age, sex, and body mass index [9]. HS in children is also associated with reduced flow-mediated vasodilation, which is a measure of endothelial dysfunction [10].

Above all, snoring is a social problem. With the most widely accepted estimate for the prevalence of chronic snoring around 40% in adult men and 20% in adult women [11], [12], millions of bed partners experience impaired sleep quality worldwide [13]. Social intolerance of loud, persistent snoring is the most serious reason for people to start treatment. Among many surgical and non-surgical modalities, the most reliable are palatal surgery and mandibular advancement devices (MAD) [7], [14]. Some snorers experience not only fluttering, but a full collapse, of the pharyngeal walls during inspiration. This leads to apnea, a pause in breathing longer than 10 seconds. Direct consequences of single apnea are: hypoxia, hypercapnia, arousal, and increased sympathetic activity. This in turn leads to higher risk of cardiovascular diseases. 13% of men and 6% of women have moderate to severe OSA [15], but less than 10% of these are diagnosed in most countries.

There is an enormous need for a simplified screening instrument capable of convenient and reliable diagnosis of OSA and snoring [16], [17]. As millions of children undergo surgery each year due to sleep-disordered breathing and thousands of adults undergo snoreplasty, there is a great need for a system to provide long-term data analysis, to both measure the severity of sleep breathing disorders and to monitor the effects of different treatment modalities. Both snoring and OSA, in children and adults, need to be measured easily and reliably in a home environment.

Different approaches to SDB screening, particularly using m-Health systems, have already been presented [18], [19]. The easiest way to record one’s breathing sounds during sleep is to use one of multiple smartphone applications that are prepared to analyze snoring [20]. So far, there are dozens of smartphone applications to record and analyze snoring, but all are based on ambient, built-in microphone recordings. By contrast, in-lab, medical sleep studies use medical sensors that record snoring directly from the patient. These studies are expensive and limited.

We introduce a novel sleep study device which utilizes

M. Młyńczak is affiliated with the Institute of Metrology and Biomedical Engineering, Faculty of Mechatronics, Warsaw University of Technology, Poland. E. Migacz and W. Kukwa are connected with the Department of Otorhinolaryngology, Medical University of Warsaw, Poland. M. Młyńczak is affiliated with the Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland.

Corresponding author: W. Kukwa wojciechkukwa@gmail.com

Manuscript received July 20, 2016; revised 2016.
a smartphone as a signal processing center, into which the application is loaded and which could help the physician to diagnose, and a wireless sensor which records and transmits the breathing signals acquired from the trachea.

The aim of this study is to investigate the accuracy, particularly the sensitivity and specificity, of a novel method of differentiating between normal breathing and snoring episodes, using acoustic sensor and artificial intelligence techniques.

A preliminary version of this work was reported at the ERS International Congress in London, in 2016 [21].

II. MATERIALS & METHODS

A. Participants

The participants in the study were 16 generally healthy subjects (aged 25-75, 10 males, 6 females). Patients were randomly selected among students and staff of the Warsaw University of Technology. No medical history was collected from the participants. 40 whole-night recordings were analyzed. Subjects were asked to sleep in the most natural way. All were informed about the aim of the study (we complied with the World Medical Association Declaration of Helsinki regarding ethical conduct of research involving human subjects).

B. Sensor

We prepared a wireless acoustic sensor to measure sounds registered during sleep from the upper respiratory tract at the level of the trachea. The digital MEMS (micro-electromechanical system) microphone unit, enabling 16-bit registration resolution and digital adjustment of amplification, sensitivity, and subrange, was chosen. The housing was designed and delivered. Figure 1 presents the concept of the prepared system.

Preprocessing consisted of band-pass linear filtration to extract the spectral sub-band strictly connected with breathing and snoring frequencies. The filter was designed to have a low-pass cut-off frequency 3.5 kHz and a high-pass frequency of 150 Hz.

The segmentation section was intended to mark the beginnings and ends of consecutive respiratory episodes (episode-to-episode strategy). The algorithm began by calculating a sum of signal spectral features (from 20 frequency ranges divided from an analytic point of view. Figure 2 shows the positioning of the sensor on the neck.

C. Algorithms

The data stream was transmitted to the smartphone, where all algorithms ran in real time. The flow of the algorithm is presented in Figure 3.

Preprocessing consisted of band-pass linear filtration to extract the spectral sub-band strictly connected with breathing and snoring frequencies. The filter was designed to have a low-pass cut-off frequency 3.5 kHz and a high-pass frequency of 150 Hz.

The segmentation section was intended to mark the beginnings and ends of consecutive respiratory episodes (episode-to-episode strategy). The algorithm began by calculating a sum of signal spectral features (from 20 frequency ranges divided

Fig. 1. The concept of the system (the acoustic sensor with wireless connection to a smartphone).

Fig. 2. The positioning of the wireless sensor.

Fig. 3. Flow chart of the smartphone algorithm. ANN: Artificial neural network.
identically in the frequencies passed from the preprocessing filter). Then, an adaptive threshold based on 10-second segments of the signal (with 50% overlap) was calculated, enabling selection of those signal portions which exceeded the threshold.

Some heuristics were implemented to make the segmentation more robust:

- Duplicates created as a result of the overlap were removed.
- Episodes that lasted longer than 3 seconds were split in two (the smallest value of the acoustic signal envelope found between 30% and 70% of the episode duration was used as the break point).
- Episodes lasting less than 0.4 second were treated as artifacts or insignificant speckles and removed from the analysis.

Apnea sections were deducted from initial segmentation at once. A classical strategy was employed, based on periods between the determined episodes lasting over 10 seconds [22]. The improvement, regarding ineffective breaths, was added to the algorithm and it consisted of removing episodes lasting less than 0.5 second and whose time distance to at least one neighboring episode was more than 6 seconds.

At the final stage of segmentation, the quality of the recording was assessed to exclude those segments, which could provide little or no significant information from an analytics point of view. The assessment could be summarized by several rules identifying poor signal quality:

- less than 10 episodes in one minute of recording (probably bad segmentation or signals quality),
- more than 5 episodes with a duration greater than 0.5 seconds (probably a large number of clicks in the signal),
- signal envelope does not exceed an arbitrary threshold (probably low signal amplitude),
- break of at least 25 seconds between two consecutive segmented episodes (probably an error in the recording of a specific signal segment).

After segmentation came parameterization, in which acoustic signal parameters were calculated for all remaining episodes. The ones presented below come from the analysis of the full set of analyzed methods, reduced using the mRMR (minimum Redundancy Maximum Relevance) feature selection method [23]. The main parameters were:

- the average and standard deviation of the signal’s absolute values,
- the three first maxima of the frequency spectrum estimated using 20th order Burg’s AR modeling method,
- the average and standard deviation of the signal’s values on the mel scale,
- the ratios of the expected value to the minimum and of the maximum to the minimum – calculated for a sound episode extended by 5 seconds before and after each segment,
- the parameters of 8th order Linear Prediction Coding modeling,
- statistical parameters specified on the basis of historical data.

In the next step of the process, the input vector containing the sound signal parameters was fed to the classification stage. There was an assembly of three independently trained and differently composed multi-layer artificial neural networks. Normal breathing was marked as ‘0’, and snoring as ‘1’ (each network had one output neuron).

The output data obtained from each classification module were fed to an inference module. The final output was established using a voting strategy. When the classification outputs of all neural networks were consistent, the process was completed and output was stored. Otherwise, the result closest to ‘0’ or ‘1’ was treated as the final classification.

The real-time analysis was conducted so that first recorded minute was processed during registration of the second minute and so on. The scheme is presented in Figure 4.

![Real-time analysis schematic](image-url)

**Fig. 4.** Real-time analysis schematic.

Dividing the analysis process into small portions of recorded signal made it possible to obtain a result for the whole period of sleep immediately after waking up.

All algorithms were implemented in Java as an Android application. Registrations were carried out on the three types of devices: HTC One M8, HTC Desire C and Samsung Galaxy Tab 2.

In order to prepare the Artificial Neural Network modules and check the accuracy of the system, we examined acoustically and labeled 1,525 episodes, of which 520 were snoring ones, 1,000 indicated normal breathing, and 5 were marked as “uncertain”. 60% of the database was assigned to a training set and the remainder to testing (608 episodes, including 207 snores, were used for accuracy evaluation).

Labeling of the episodes was provided by four experts with a peer-review strategy. The dominant answer is treated as the final single reference. In case of draw, the vote for main reviewer prevails.

**D. Methods for result analysis**

The analysis of results may include various parameters. The number of pairs of neighboring episodes (assumed as inspiration and expiration) determined the number of full breaths, as in the equation:

\[ N_B = \left( \frac{n_e}{2} \right) + n_e \mod 2 \]

where: \( N_B \) is the number of breaths and \( n_e \) is the number of detected episodes.

The time between two respiratory episodes, separated by a single respiratory event, determined the respiratory rate (RR).
Snoring episodes may be defined as ‘separate’ snoring episodes or as a group of ‘aggregate’ ones. The number of these was calculated as in the equation:

\[ N_S = n_{SS} + \sum_{i=1}^{k} \left( \frac{n_{CS}(i)}{2} \right) + n_{CS}(i) \mod 2 \]

where: \( N_S \) is the number of snores, \( n_{SS} \) is the number of separate snores, \( n_{CS}(i) \) is the number of aggregate (collected) snores in the \( i \) group, and \( k \) is the number of groups.

### III. RESULTS

Our algorithm to determine breathing episodes based on spectral thresholding and heuristics worked very well for the entire recordings, except for body position changes, which had to be removed from analysis. Figures 5, 6, 7, and 8 present the portions of recorded signals with segmentation and classification results for normal breathing and snoring, respectively.

Proposed system achieved mean 88.8% accuracy in the differentiation between normal breathing and snoring. As it is intended mainly for screening, high specificity of 95% is reported. Relatively large Cohen’s Kappa, which equalled 0.7775, and was included to measure inter-rater agreement, removing the part of the agreement occurring by chance, should also be reported.

The confusion matrix of classification results and overall evaluation of the accuracy (based on comparison with manual labeling provided by four experts with a peer-review strategy) are provided in Table I and II, respectively.

### IV. DISCUSSION

There is no physical nor mathematical definition of snoring. What we hear as snoring, it was so labeled – snoring is “in the ear of the beholder” [24]. There are different methods of studying breathing and other physiological parameters during sleep. The gold standard in sleep studies is polysomnography (PSG). The study measures multiple signals, among them airflow, breathing effort, oxygen saturation, and snoring. According to the scoring manual of the American Academy of Sleep Medicine, three equivalent methods of snoring detection exist. They are: acoustic sensor (microphone), nasal pressure transducer (cannula), and piezoelectric vibration sensor [1], [25]. These sensors do not measure snore events in the same manner and exhibit important differences in sensitivity and positive predictive value (0.79 and 0.94, respectively, for overhead audio sensor; 0.78 and 0.92 for piezoelectric sensor; and 0.37...
and 0.82, respectively, and 0.55 and 0.67, respectively, for different nasal cannulas), as was finally proved by Arnardottir et al. in 2015 [1].

The study of Arnardottir et al. showed clearly that the audio-based recordings of snoring are superior to other methods recommended by the AASM. Similarly to the team of Arnardottir we used an audio-based sensor, but our recordings came directly from the trachea and not from an ambient microphone. We feel the tracheal recordings may be beneficial and be recorded; the only thing that can be recorded is loud snoring. In the home environment does not allow measurement of the ambient sound of breathing was acquired and the recordings were simultaneous with full PSG in 50 patients. Data from 10 patients were used to develop the program and those of 40 patients were used to validate it. The sensitivity and specificity of the system showed that it could be used effectively in a controlled laboratory setting.

A smartphone application is the easiest way to record ones snoring, but it has some impassable limitations. First, there are hundreds of different smartphones, having different built-in microphones. This makes it impossible to prepare a tool to similarly measure and analyze breathing sounds on each of those phones. Second, the use of an ambient microphone in the home environment does not allow measurement of the sounds of regular breathing, which is too quiet to be properly recorded; the only thing that can be recorded is loud snoring. Third, sleep position (supine, prone, lateral) influences the characteristics of the sound [34]. This information could be correlated with respiratory-related analysis and improve the inference part of the study [35]. Finally, various environmental sounds, including the bed partner’s snoring, would affect recordings from an ambient microphone.

There are several studies using tracheal recordings in de-
Limitations of the study

The system was not validated against full polysomnography, the gold standard in sleep studies. So far, accuracy has only been checked against manual scoring of the audio recordings. It should be noted that the experts sometimes could not agree on labeling. This could have an impact on overall accuracy. As in every other study where audio recordings were analyzed by experts, there was an inter-scorer disagreement in around 17% of studied snoring/breathing episodes [36]. This is due to the fact that no objective criteria of snoring exist. At the same time, the accuracy of manual scoring is superior to any automatic systems used in polysomnography.

Based on this problem, we decided to maximize the specificity and allowed the sensitivity to be at the level of 76.8%, particularly considering home-based applications.

Criteria to assess signal quality were introduced arbitrarily, due to the lack of standard. They could be adaptively established during analysis. We did not focus on that in the paper.

There were no sensor loosening events during our measurement. However, it seems that the additional parameter in algorithm to alarm subject about the loosening should be added to deal with such possibility.

V. Conclusion

The system enables classifying between normal breathing and snoring. The combination of a smartphone and an external sensor is in our opinion a perfect compromise between non-medical phone applications and medical sleep studies. This approach enables the study to be repetitive, personal and inexpensive (as with smartphone applications) while relying on well-recorded, reliable, comparable data (like a medical sleep study).

This opens new possibilities in sleep monitoring and study, enabling the collection of huge amounts of data from thousands of people. Multi-night recordings finally make ideas like the “Human Sleep Project” proposed by Till Roenneberg possible [37]. The results are available immediately after recording and could be shared with a doctor.

Further projects with a multichannel wireless sensor for smartphone owners are already underway and the results will be presented soon.

Author Contribution

M. Młyńczak and W. Kukwa prepared the concept and gathered all parts of the publication. M. Młyńczak is the main author of the algorithms elaborated in Materials & Methods section and the calculations presented in Results. W. Kukwa and E. Migacz prepared the Introduction and Discussion. Each author contributed a part to the system and was engaged in labeling the data and stating the final Conclusions.

Acknowledgment

The sensor and mobile application were designed and prepared by the Clebre Company. The presented system was described in an EPO Patent Application, submitted in April of 2016.