Study of Phylogenetic for Computational Analysis of Sleep Apnea Syndrome for Patient (Healthcare & Treatment) Using Machine Learning (Robot Vision)

Romil Rawat, Vinod Mahor and Shrikant Telang

August 1, 2021

Study of Phylogenetic for computational analysis of Sleep Apnea syndrome for Patient (Healthcare & Treatment) using Machine Learning (Robot Vision)

Abstract:

The reconstruction of the Tree of Life is a classical and Complex problem in evolutionary biology that has benefited from numerous branches of mathematics, including probability, information theory, Proof Theory, combinatorics, and geometry. Additionally, advances in computer technology Design and architecture such as parallel, Centralized and distributed computing, and programs that exploit them efficiently in combination with the continual development of faster search strategies promise to make even larger phylogenetic problems increasingly tractable. However, the NP-completeness of the phylogeny problem represents a fundamental limitation in efforts to unearth the tree of life. Modern DNA sequencing technologies are producing a deluge of new genetic data, transforming how we view the Tree of Life and how it is reconstructed.Sleep Quality Analysis is currently accessed through different means, from simple commercial devices that measure activity through an accelerometer or more complex ones that use oxymetry, to medical devices such as [1]polysomnography(PSG) that can be used indirectly to estimate sleep quality. which is an expensive procedure involving much effort for the patient. Multiple systems have been proposed to address this situation, including performing the examination and analysis in the patient's home, using sensors to detect physiological signals that are automatically analyzed by algorithms. However, the precision of these devices is usually not enough to provide clinical diagnosis. Therefore, the objective of this review is to analyze already existing algorithms that have not been implemented on hardware but have had their performance verified by at least one experiment that aims to detect obstructive sleep apnea to predict trends.

Index Terms—Algorithms review sleep apnea, Algorithms, databases, phylogenetic, combinatorics , Proof Theory, evolutionary biology, medical

Introduction:

The scope of phylogenetic analysis [29] has increased greatly in the last decade, with analyses of hundreds, if not thousands, of taxa becoming increasingly common in our efforts to reconstruct the tree of life and study large and species rich taxa. As such, these results support recent suggestions that taxon number in and of itself might not be the primary factor constraining phylogenetic accuracy and also provide important clues for the further development of divide-and conquer strategies for solving very large phylogenetic problems.

Advancements in medical science and technology, medicine and public health coupled with increased consciousness about nutrition and environmental and personal hygiene have paved the way for the dramatic increase in life expectancy globally in the past several decades. However, increased life expectancy has given rise to an increasing aging population, thus jeopardizing the socio-economic structure of many countries in terms of costs associated with elderly healthcare and wellbeing. In order to cope with the growing need for elderly healthcare services, it is essential to develop affordable, unobtrusive and easy-to-use healthcare solutions. Smart homes, which incorporate environmental and wearable medical sensors, actuators, and modern communication and information technologies, can enable continuous and remote monitoring of elderly health and wellbeing at a low cost. Smart homes may allow the elderly to stay in their comfortable home environments instead of expensive and limited healthcare facilities. Healthcare personnel can also keep track of the overall health condition of the elderly in real-time and provide feedback and support from distant facilities. We need to create a comprehensive review on the state-of- the-art research and development in smart home based remote healthcare technologies and to make the concerning people aware with precision if any alarming situation is spotted so that instant action can be taken thus ensuring reduction of casualty. Quality of service in healthcare has always been under constant criticism in the modern era, as it is a very touchy subject. Health monitoring specially for elderly people is a concern and as most people in the modern times are job holders and have so hectic life. It is difficult to manage to keep a constant watch on the elderly of the house. Keeping a nurse or housekeeper is also a very costly issue nowadays. In this situation, remote health monitoring based on IOT can help to solve the problem.

The simulation protocol used was modelled on that followed by Bininda-Emonds to examine the scaling of accuracy in very large phylogenetic trees. For each run, a model tree of 4,096 taxa need to be generated according to a stochastic Yule birth process using the default parameters of the fundamental objective of biology is to produce an accurate tree of life for the world's 1.7 million known species. Before this tree can be completed, biologists face three computational challenges. They must develop the tools to analyze the large data sets and determine relationships for all living organisms, train people to collect the data and use the computational tools once they are developed, and engineer an infrastructure that can support the data collection and computational systems.

Hillis, director of Texas' School of Biological Sciences, is working on approaches to reconstruct the history of thousands of species and, with this history, to interpret biological information on the behavior, physiology, or molecular evolution of the species. Hillis' primary limitation is the inability to handle the large, complex data sets that are becoming more and more common within phylogenetic analysis.In practice, a modern phylogenetic analysis consists of finding the pattern of relationships that best accounts for the DNA or RNA sequence data, given a particular model of evolution. Evaluating a potential solution to a phylogenetic problem is computationally very intensive, and the number of possible solutions to a phylogenetic problem[29] increases precipitously as more organisms are added. This puts the phylogeny problem into the class of computational problems that are NP-complete; the solution time goes up exponentially as a function of the number of sequences in the analysis. The rapid increase of the size of phylogenetic data sets is largely due to discoveries in molecular biology and genetics that are producing gene sequences of more and more organisms. These large sets of sequences create phylogenetic problems that are very difficult to solve--even with the world's most powerful supercomputers.

Because of the difficulty of the problem, there is no provably optimal algorithm for solving it; instead, there are several heuristic approaches. Researchers are looking at the performance of an approach based on genetic algorithms, which can be easily scaled to take advantage of large pools of computer resources. They are intrinsically parallel algorithms, and the project has used for the Advanced Computation Center to refine and test the approach. Researchers are fascinated by analyzing large trees, because of the associated grand computational challenges that entail problems from theoretical computer science as well as from parallel computing. “Evolutionary biologists are currently generating a molecular data avalanche that is even hard to analyze on the most powerful supercomputers“, says the 34 year old scientist. “The challenge for computer science is to develop programs and methods for calculating evolutionary trees and to discover knowledge in the mass of molecular data.”

More than 60 different sleep disorders, divided into seven categories, have been identified by the International Classification of Sleep Disorders. Sleep-related breathing disorders are the second category which includes central sleep apnea, obstructive sleep apnea (OSA) and sleep-related hypoxemia and hypoventilation [1]. OSA is the most common disorder in this group and is characterized by partial or complete obstruction and recurrent collapse of the upper airway, affecting ventilation during sleep. The symptoms of this disorder are excessive daytime sleepiness caused by non-restorative sleep. It is estimated that an OSA prevalence in the general adult population ranges from 6% to 17%, considering an apnea-hypopnea index (AHI) of greater than or equal to 15 events/hour, with males being more affected than females, and this prevalence becomes more relevant with increasing age [2]. Polysomnography (PSG) is the gold standard for OSA diagnosis measuring multiple sensors to record the breath airflow, respiratory movement, oxygen saturation (SpO_2), electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), electrocardiogram (ECG) and body position [3]. OSA is diagnosed if the patient has reported the indicated symptoms and presents 5 or more obstructive respiratory events per hour of sleep during a PSG recording [1]. Alternatively, OSA can be diagnosed if a frequency of obstructive respiratory events greater than or equal to 15 events/hour is detected, independent of associated symptoms. OSA severity can be defined as mild ($5 \leq AHI < 15$ e/h), moderate ($15 \leq AHI < 30$ e/h) or severe (AHI S 30 e/h).

The analyzed algorithms would be suited to the third and fourth level devices. However, a more suitable categorization system based upon employed sensors in these out-of-center devices was presented by Collop et al.

[6], named SCOPER, being divided into six categories: sleep; cardiovascular; oximetry; position; effort; respiratory.

The four main tasks for the automatic recording and analysis of sleep were identified in a review performed by Penzel and Conradt [7] in 2000, specifically: replacement of conventional paper chart recorders; technicians should be able to introduce additional notes and observations that were performed during the nocturnal recording; the systems should support sleep evaluation including snoring, respiration and oxygen saturation for the diagnosis of sleep-related breathing disorders; a final report of the analysis should be produced at the end. A review of computer-aided approaches for OSA diagnosis was presented by Alvarez-Estevez and Moret-Bonillo [8] covering papers. The review analyzes sleep apnea-hypopnea syndrome screening approaches, apnea event detection and classification methods, comprehensive diagnostic systems, and commercial approaches. However, the inclusion criteria restricted the analysis to papers that used, at least, a subset of the PSG signals proposed in the AASM manual for scoring of sleep and associated events at the time of publication, excluding the nonstandard biomedical signals for the OSA diagnosis such as ECG or pulse wave analysis. Multiple reviews have been performed with focus on devices for home detection [6], [9], [10] and some papers focused completely on smartphone applications, as for example [11], [12]. However, a review of hardware is out of scope of this article.

Previous reviews have been conducted, examining algorithm implementations as presented by Faust et al. Analyzing ECG based algorithms, or classification techniques such as the review developed by Pombo et al. [14], concluding that the selection of effective features mainly influences the accuracy of the classifier models. However, there was no reference found in the cutting edge reviews to making a global survey of algorithms for OSA detection by evaluating multiple source sensors. Therefore, the focus of this work is in the analysis of the performance of different algorithms and methods that use signals from multiple source sensors but have not been implemented in hardware, to detect OSA.

1) PROPOSED WORK

a) Research Questions

- What are the data acquisition and data transmission in the domain of remote health care for people?
- How to develop a fast reaction intelligent automated system which will help to monitor host remotely?
- How to develop a system that can change the service mode into a pervasive way, and trigger the healthcare?
- How to interpret the Real time data, so that sudden and quick action can be performed.
- Development of Low cost, easily accessible and operated system.
- Easily wearable System.

b) Objective

- To understand the current trends and research in sleep apnea.
- Algorithms and techniques understanding of Work done so far.
- Database and datasets used.
- Making an automated system which will help to monitor host remotely is our primary objective.
- Making an alarm or reaction system which will react whenever there is an alarming situation.
- Providing a way to remotely monitor the temperature, pulse, counting the bowel discharge in a day and also the mount of sleep of the patient via Thinkspeak.
- Analyzing the collected data using the built in Matlab of the Thinkspeak sever to detect future hazards.
- Sending alarming messages via e-mail and twitter to the concerning authority or people if any abnormality is detected.
- Contributing in the field of IOT to pave a way for future project in the technological development.
- To develop a system can change the service mode into a pervasive way, and trigger the healthcare service based on patients' physical status rather than their feelings. In order to realize the pervasive healthcare service, a remote monitoring system is essential.
- A pervasive monitoring system that can send patients' physical signs to remote medical applications in real time.

- The system is mainly composed of two parts: the data acquisition part and the data transmission part. The monitoring scheme (monitoring parameters and frequency for each parameter) is the key point of the data acquisition part, and we designed it based on interviews to medical experts.
- Multiple physical signs (blood pressure, ECG, SpO2, heart rate, pulse rate, blood fat and blood glucose) as well as an environmental indicator (patients' location) are designed to be sampled at different rates continuously. Four data transmission modes are presented taking risk, medical analysis needs, demands for communication and computing resources into consideration.

c) Standard Databases

- Describe what arrangements will be made for data collection or source consultation, with special reference to access, and relations with participants or other bodies
- accessing international databases libraries to conduct a systematic review for previous researches
- PhysioNet
- Apnea-ECG database
- Thinkspeak
- MIT-BIH database;
- St. Vincent's database
- CHUS database
- Fantasia database.
- (UCI) machine learning repository.
- Data collection in a non-invasive way, including the use of cameras to estimate sleep stages and heart beat rate.

d) Proposed Algorithms for Current Research

- **Machine learning algorithms (SVM)-**Machine learning to generate a neural network algorithm to detect OSA from overnight pulse oximetry tracings
- **Pattern Matching Algorithms (Classification Algorithm)-**Sleep apnea is a common sleep disorder that causes pauses of breathing during sleep. Time and frequency analysis of heart rate variability (HRV), ECG-derived respiration, photoplethysmography, and other signals were proposed for minute-by-minute apnea classification algorithms for the whole night of sleep
- **Cryptographic Algorithms**-Creating Secure Architecture
- **Image Processing Algorithms-**Use of image and speech processing to estimate the apnea-hypopnea index.

2) Available Algorithms For OSA Detection:

The analyzed algorithms were divided into five categories depending on the source sensor: pulse oximetry; ECG; respiration; sound; combined approaches.

Multiple algorithms have been proposed to detect OSA but have not been implemented on a developed system, in some cases probably due to the high degree of complexity used, thereby making it hard for the algorithm to be efficiently executed on systems with limited resources. Nevertheless, these issues can be resolved with the advance of technology. The aim of this work is to review papers that have presented algorithms based on the analyses of pulse oximetry, ECG, sounds, respiration since these fields seem to be the more promising approaches.

a) Based on Pulse Oximetry

Classical oximetry analysis includes the oxygen desaturation index (ODI), cumulative time spent below a defined saturation threshold (e.g., time below 90% = T90), number of falls in the SpO2 value below the defined baseline and signal variability, usually the delta index. A threshold approach was presented by Jung et al. [20] taking three points into consideration.

b) Based on ECG

Analysis of ECG waveforms and ECG-derived hear trate are commonly used to detect sleep-related breathing disorders. Lin et al. [36] separated the ECG signal into four spectral com- ponents (alpha, beta, delta and theta) using multi-resolution wavelet transforms and the wavelet coefficients were used as the training input for a four layer NN implemented with simple neural computing elements.

c) Based on Respiration

Oronasal airflow is one of the most direct indicators of breathing disorders and was used by Koley and Dey [12] to detect OSA. First the signal was filtered with a low pass Butterworth filter, to remove artefacts that usually affect this signal as a result of the system being used to make the measures, and normalized to avoid subject-dependent variations. The resulting signal was then segmented and three time domain features were extracted, the area covered by the respiration signal (measuring the total volume of air flow), variance (sensitive to respiration amplitudes that measure the air inflow rate to outflow rate) and the upper 90th percentile (decrease during apnea events). These **Based on Sound**

The breathing process produces characteristic sounds that can be used to detect the presence of disorders. This principle was used by Rosenwein et al. [10]. Suspected breathing disorder periods were detected when the result of the energy envelope of the audio signal with the average value subtracted was negative. Then six features were calculated from the suspected period, breathing and non-respiratory rates, duration of last respiratory event, variation of respiratory energy, average ventilation and mean energy value. These features were used as inputs to a binary-random forest classifier and the produced output was classified by an adaptive threshold produced for each subject's score distribution. Breathing sounds were also the base of the algorithm presented by Almazaydeh et al. [11] where voice activity detection (VAD) was used to classify respiratory signals. The sounds were first filtered and segmented, and were then applied to the FFT. These segments were further analyzed by the VAD algorithm which compared them against the threshold, determined by comparing the signal value against noise. The output identifies whether the segment was a normal breath or a breathing cessation (silence). A second threshold was then used to classify the silence as either apnea or normal. Recorded respiratory sounds were used to extract spectrum features, using the FFT, by Praydas et al. [11].

d) Based on Combined Approaches

Usually, the pulse oximeter provides both the SpO2 and heart rate signals, but it is also common for SpO2 only to be considered in the OSA detection algorithm. A different approach was presented by Zamarro´n et al. [21] where a combination of these two signals was used. The frequency spectrum of the signals was interpolated to get the spectral amplitude at equally distributed frequencies between 0 to 0.1 Hz and then the data was averaged. The algorithm looks for peaks on the apnea-related frequency band of both signals to classify OSA. An algorithm that uses both EEG and oximetry was presented by A´ lvarez et al. [9]. The PSD was applied to each recording using Welch's method where the data was divided into overlapping segments, the FFT was applied and the result was averaged. Afterwards spectral features were computed to detect OSA. Peak amplitude of the apnea-related frequency band and median frequency (to summarize the spectral content) were the selected features for the SpO2 PSD. Relative power on selected EEG bands (delta and alpha) and spectral entropy (disorder quantifier computed based on the Shannon entropy) were the features chosen for the EEG analysis.

e) Based on machine learning

Machine learning–based algorithms that use only information readily available from the clinical record have been shown to be useful screening tools for OSA in adults (17). In the current study, machine learning enhanced the test performance of isolated nocturnal pulse oximetry, but the authors did not evaluate a combined approach, using oximetry results in combination with clinical data, questionnaires, or urinary biomarkers.

3) Discussion

Different approaches (Phylogenetic Design) have been followed with the aim of detecting OSA. From the overall analysis of this review it is recognized as future directions for the research to produce more robust OSA diagnosis tools by implementation of the presented algorithm in efficient hardware, produce more research with deep learning classifiers, capable of self-learning the features, and validate the achieved results of the algorithms by independent research groups using publicly available databases so that the results can be reproduced. This has special interest for home diagnostic devices since they could be used as a first OSA diagnosis tool, leading to a considerable reduction in the diagnostics cost and waiting time for access to a sleep study. However, these devices are more susceptible to data errors caused by factors not con- trolled at the home of the subject. Therefore, an adaptation of the proposed algorithms to a real world environment in efficient hardware is the major challenge identified. The main gaps in the current state of the art are the algorithms capable of self-learning the features.

References

[1] M. Sateia, "International classification of sleep disorders-third edition: highlights and modifications," Chest, vol. 146, no. 5, pp. 1387–1394, Nov. 2014.
[2] C. Senaratna et al., "Prevalence of obstructive sleep apnea in the general population: A systematic review," Sleep Med. Rev., vol. 34, no. 1, pp. 70–81, Aug. 2017.
[3] L. Almazaydeh, K. Elleithy, and M. Faezipour, "Detection of obstructive sleep apnea through ECG signal features," in Proc. IEEE Int. Conf. Electro/Inf. Technol., Indianapolis, IN, USA, May 2012, pp. 5559–5562.
[4] T. Penzel, "Home sleep testing," in Principles and Practice of Sleep Medicine, 6th ed. Amsterdam, The Netherlands: Elsevier, 2016, pp. 1610–1614.
[5] A. Chesson, R. Berry, and A. Pack, "Practice parameters for the use of portable monitoring devices in the investigation of suspected obstructive sleep apnea in adults," SLEEP, vol. 26, no. 7, pp. 907–913, 2003.
[6] N. Collop et al., "Obstructive sleep apnea devices for out-of-center (OOC) testing: Technology evaluation," J. Clin. Sleep Med., vol. 7, no. 5, pp. 531–548, Oct. 2011.
[7] T. Penzel and R. Conradt, "Computer based sleep recording and analysis,"Sleep Med. Rev., vol. 4, no. 2, pp. 131–148, 2000.
[8] D. Alvarez-Estevez and V. Moret-Bonillo, "Computer-assisted diagnosis of the sleep apnea-hypopnea syndrome: A review," Sleep Disorders, vol. 2015, Jul. 2015, Art. no. 237878.
[9] S. Pullano et al., "Medical devices for pediatric apnea monitoring and therapy: Past and new trends," IEEE Rev. Biomed. Eng., vol. 10, no. 1, pp. 199–212, Sep. 2017.
[10] W. Flemons et al., "Home diagnosis of sleep apnea: A systematic review of the literature," Chest, vol. 124, no. 4, pp. 1543–1579, 2003.
[11] M. Al-Mardini, F. Aloul, A. Sagahyroon, and L. Al-Husseini, "On the use of smartphones for detecting obstructive sleep apnea," in Proc. IEEE 13th Int. Conf. Bioinformat. Bioeng., Chania, Greece, Nov. 2013, pp. 1–4.
[12] M. Al-Mardini, F. Aloul, A. Sagahyroon, and L. Al-Husseini, "Classi- fying obstructive sleep apnea using smartphones," J. Biomed. Informat., vol. 52, pp. 251–259, Dec. 2014.
[13] O. Faust, R. Acharya, E. Ng, and H. Fujita, "A review of ECG based diagnosis support systems for obstructive sleep apnea," J. Mech. Med. Biol., vol. 16, no. 1, pp. 1–25, 2016.
[14] N. Pombo, N. Garcia, and K. Bousson, "Classification techniques on computerized systems to predict and/or to detect Apnea: A systematic review," Comput. Methods Programs Biomed., vol. 140, pp. 265–274, 2017.
[15] T. Penzel, G. Moody, R. Mark, A. Goldberger, and J. Peter, "The apnea- ECG database," in Proc. Comput. Cardiology, Cambridge, MA, USA, 2000, pp. 255–258.
[16] A. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Com- ponents of a new research," Circulation, vol. 101, no. 23, pp. 215–220, 2000.
[17] S. Quan et al., "The sleep heart health study: Design, rationale, and methods," Sleep, vol. 20, no. 12, pp. 1077–1085, 1997.
[18] Y. Ichimaru and G. Moody, "Development of the polysomnographic database on CD-ROM," Psychiatry Clin. Neurosci., vol. 53, no. 2, pp. 175–177, 1999.
[19]D. Dean et al., "Scaling up scientific discovery in sleep medicine: The national sleep research resource," Sleep, vol. 39, no. 5, pp. 1151–1164, 2016.
[20]D. Jung et al., "Real-time automatic apneic event detection using noctur- nal pulse oximetry," IEEE Trans. Biomed. Eng., vol. 65, no. 3, pp. 706–712, Mar. 2018.
[21] D. A´ lvarez, R. Hornero, D. Aba´solo, F. Campo, and C. Zamarro´n, "Non- linear characteristics of blood oxygen saturation from nocturnal oximetry for obstructive sleep apnoea detection," Physiol. Meas., vol. 27, no. 4, pp. 399–412, 2006.
[22]J. Marcos, R. Hornero, D. A´ lvarez, F. Campo, and M. Aboy, "Automated detection of obstructive sleep apnoea syndrome from oxygen saturation recordings using linear discriminant analysis," Med. Biol. Eng. Comput., vol. 48, no. 9, pp. 895–902, 2010.
[23] J. Marcos, R. Hornero, D. A´ lvarez, F. Campo, and C. Zamarro´n, "As- sessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry," Med. Eng. Phys., vol. 31, no. 8, pp. 971–978, 2009.
[24] J. Morales et al., "Sleep apnea hypopnea syndrome classification in SpO2 signals using wavelet decomposition and phase space reconstruction," in Proc. IEEE 14th Int. Conf. Wearable Implantable Body Sensor Netw., Eindhoven, The Netherlands, May 2017, pp. 43–46.
[25] D. A´ lvarez, R. Hornero, J. Marcos, and F. Campo, "Feature selection from nocturnal oximetry using genetic algorithms to assist in obstructive sleep apnoea diagnosis," Med. Eng. Phys., vol. 34, no. 8, pp. 1049–1057, 2012.
[26]D. Morillo and N. Gross, "Probabilistic neural network approach for the detection of SAHS from overnight pulse oximetry," Med. Biol. Eng. Comput., vol. 51, no. 3, pp. 305–315, 2013.
[27] L. Almazaydeh, M. Faezipour, and K. Elleithy, "A neural network system for detection of obstructive sleep apnea through SpO2 signal features," Int. J. Adv. Comput. Sci. Appl., vol. 3, no. 5, pp. 7–11, 2012.

[28] S. Mostafa, F. Mendonca, F. Morgado-Dias, and A. Ravelo-Garc´ıa, “SpO2 based sleep apnea detection using deep learning,” in *Proc. 21st Int. Conf. Intell. Eng. Syst.*, Larnaca, Cyprus, Oct. 2017, pp. 91–96.
[29] Stephen A. Smith , Joseph W. Brown, Cody E. Hinchliff,” Analyzing and Synthesizing Phylogenies Using Tree Alignment Graphs”, September 26, 2013https://doi.org/10.1371/journal.pcbi.1003223.

ABBREVIATIONS

.PSG – polysomnography. . **OSA -** obstructive sleep apnea. AHI-apnea-hypopnea index **SpO2-** Oxygen saturation.
.EEG – Electroencephalogram. **EOG** - electrooculogram **EMG –**electromyogram. **ECG** –electrocardiogram.
.AASM -American Academy of Sleep Medicine. **HRV**-heart rate variability. **PPG** -Photoplethysmography.
.ODI -oxygen desaturation index. **NN**- Neural Network. .**CTM-**Computational Theory of the Mind.
.LZC -Lempel Ziv Complexity. **SVM-** Support Vector Machine. .**VAD-**voice activity detection.
.FFT -Fast Fourier Transform. .**PSD-**Power Spectral Density.