Device-free Sleep Stage Recognition through Bed Frame Vibration Sensing

Zhizhang Hu
Carnegie Mellon University
Pittsburgh, PA
zhizhanh@andrew.cmu.edu

Emre Sezgin
Nationwide Children's Hospital
Columbus, OH
Emre.Sezgin@nationwidechildrens.org

Simon Lin
Nationwide Children's Hospital
Columbus, OH
Simon.Lin@nationwidechildrens.org

Pei Zhang
Carnegie Mellon University
Moffett Field, CA
peizhang@cmu.edu

Hae Young Noh
Carnegie Mellon University
Pittsburgh, PA
noh@cmu.edu

Shijia Pan
University of California, Merced
Merced, CA
span24@ucmerced.edu

ABSTRACT

Sleep disorder impairs people's health. To better analyze user's sleep quality, it is important to monitor people's sleep stages. Prior methods, such as polysomnography (PSG) and Fitbit, are often intrusive and having potential to change user's daily sleep routine. We present a non-intrusive approach to identify sleep stages through bed-frame vibrations. The proposed system detects the movement of users during their sleep and estimates their sleep stages via their movements induced vibration on the bed frame. Our system extracts features from the vibration signal and distinguishes sleep stages between rapid eye movement (REM) and non-rapid eye movement sleep (NREM). We evaluated our system in two ways 1) to understand the data distribution over different sleep stages, and 2) to explore the feasibility of distinguishing different movement characteristics over different sleep stages. The system achieved an area under curve (AUC) score of 0.84 when classifying REM sleep stages over NREM stages.

CCS CONCEPTS

• Information systems → Sensor networks; • Human-centered computing → Ubiquitous and mobile computing design and evaluation methods; • Computer systems organization → Sensor networks.

KEYWORDS

Non-intrusive sensing, Device-free human sensing, Sleep stage recognition, Sleep monitoring

ACM Reference Format:

Zhizhang Hu, Emre Sezgin, Simon Lin, Pei Zhang, Hae Young Noh, and Shijia Pan. 2019. Device-free Sleep Stage Recognition through Bed Frame Vibration Sensing. In *The 1st ACM International Workshop on Device-Free Human*

Sensing (DFHS'19), November 10, 2019, New York, NY, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3360773.3360883>

1 INTRODUCTION

Sleep stage recognition has gain ascending interest as an approach to assist in solving sleep disorder [11]. Lack or low quality of sleep has an impact on daily activities and may result in physical and mental issues. Researchers have explored the relation between sleep and other diseases, such as Parkinson's disease [4] and Alzheimer's disease [1]. It is established that patients with increased severity of obstructive sleep apnea may spend more time in REM than deep sleep, and in light sleep than wake [7]. As a result, monitoring sleep stages is important for doctors to screen and diagnose sleep related diseases.

The first standardized criteria for classifying sleep stages were established in 1968 by Allan Rechtschaffen and Anthony Kales (R&K scoring manual)[9]. In 2004, American Academy of Sleep Medicine(AASM) commissioned a revision of sleep scoring rules, covering not only sleep stages but also the scoring of arousals, respiratory events, sleep-related movement disorders, and cardiac abnormalities. These studies mainly divided sleep into two main stages: rapid eye movement (REM) and non-REM (NREM).

Clinically, the current gold standard for evaluation of sleep stages is polysomnography (PSG), which is usually conducted at a sleep lab. Subjects are asked to wear various physiological signal sensors such as electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), etc., as shown in Figure 1 (a). A trained expert gives a sleep stage every 30 seconds based on specific patterns in the obtained signals [14]. The conventional PSG approach often requires dedicated equipment, sleep labs, and specialized personnel, which makes the long-term continuous monitoring difficult if not impossible. Besides, sleep diagnosis often requires long waiting lists due to the limited resources [6].

As an alternative to PSG, attempts have been made to develop machine learning-based sleep stage classification. Reimer et al. utilized a chest strap and two accelerometers as a wearable-based system to recognize the sleep stages. With handcrafted features, their system achieved a 90.00 % accuracy for classifying Wake, REM and N3 stages [10]. Equipped with an accelerometer and photoplethysmography (PPG) sensor, wearable fitness trackers like Fitbit were validated to have a good agreement with PSG in monitoring sleep stages for healthy adolescents [3]. These prior works would

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DFHS'19, November 10, 2019, New York, NY, USA

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-7007-3/19/11...\$15.00

<https://doi.org/10.1145/3360773.3360883>

Photo (a) source: <https://bipolaraspliomom.wordpress.com/2011/07/15/are-you-still-awake-our-eeeg-story/>

Figure 1: An example scenario.

ask users to wear devices during their sleep. This requirement affects users' routine sleep habits, impeding accurate measurement of sleep stages and diagnosis of sleep disorders. Tomura et al. proposed one criterion for sleep stages and wake detection based on the macro change of heart rate variability (HRV) acquired from the filtered vibration waves [12]. Komine et al. developed one non-intrusive mattress type pressure sensor based sleep monitoring system, which has a capability of customizing pre-trained classification model into individual patient's exclusive model with estimated sleep stage data from Tomura's method [5]. This work illustrated the possibility of in-home sleep monitoring. However, the accuracy of data from Tomura's method can not be ideally comparable to which from traditional PSG labeled data, the need for PSG labeled data is still indispensable.

In this paper, we present a non-intrusive structural vibration-based sleep stage monitoring system for REM and NREM detection. The intuition is that the target's movements during their sleep induce the mattress to vibrate. This vibration propagates through the bed frame and is captured by the vibration sensor (geophone) mounted on the frame structure underneath the mattress, as an example shown in Figure 1 (b). Our system detects the 'events' – sleep movements induced impulsive vibration signals – and associate them to different sleep stages. The contributions of this work are as follows.

- We present a device-free sleep stage monitoring system through during-sleep motion induced bed frame vibration. The system is low-cost and friendly for in-home long time self-monitoring.
- We analyzed the recognition accuracy of our vibration-based approach and compared to the traditional approaches.
- We evaluated our system through real-world deployment with ground truth labeled by sleep study experts.

The remainder of this paper is structured as follows. Section 2 presents the methods plus the classification model utilized in this research. Section 3 describes the evaluation of system performance. Finally, conclusions and future work are discussed in Section 4.

2 METHODOLOGY

Our system consists of mainly four modules: data acquisition, event detection, feature extraction, and sleep stage learning model. The proposed system consists of five main modules: bed frame vibration

Figure 2: System overview.

Figure 3: A sensing node mounted on the bed frame. The circle points out the geophone in the 3-D printed mounting devices. The blue box conducts data acquisition. The Ethernet cable connects the data acquisition device to an edge device.

sensing, sleep motion event detection, ambient noise modeling, event feature extraction and sleep stage classification as shown in Figure 2. First, the vibration signal is fed into the event detection module from the sensor. Event detection module estimates whether the current signal contains an event. Once an event is detected, it is sent to the feature extraction module. The feature of the event signal will be extracted for sleep stage classification model. Finally, the classified sleep stage is the output. If the vibration signal does not contain events, it will be sent to ambient noise modeling module, which will be elaborated in section 2.2.

2.1 Data Acquisition

The data acquisition module consists of a vibration sensor (geophone), an ADC module, and a processor. The sampling frequency is set at 1000 Hz. The sensor is mounted to the bed frame, as shown in Figure 3. A mounting device is 3-D printed to ensure that the sensor is perpendicular to the mattress and capture the vibration. When the patient motion induced vibration propagates to the bed frame, the geophone converts the velocity of the bed frame motion into voltage. This voltage is then amplified by an adjustable opamp (up to 100 dB), and the amplified signal is then digitized by the analog-to-digital converter (ADC). The digitized signal is then sent to an edge device or a server through an Ethernet connection. We used a Raspberry Pi as an edge device in the implementation demonstrated in Figure 3 to organize and store the sensing data overtime. The deployment process is easy to operate, and the cost of this system could be as low as 30 dollars per bed, both of these characteristics make it being friendly for in-home monitoring.

Figure 4: Example motion event detected by the vibration sensing system.

2.2 Event Detection

We define an ‘event’ as the signal segment that has signal energy significantly higher than that of the ambient noise signal [8]. The event detection module scans through the time series signal with a sliding window and determines whether a sliding window contains or is part of an event.

The raw sensor data streams have first applied a standardization to remove the reference voltage offsets. Then the system calculates local signal energy within a time window. The signal energy is defined as the sum of the squared values within a time window.

The size of the sliding window is heuristically decided to be around three quarters of a second, which is based on the empirical discovery that an unconscious activity exerts force to the bed frame for around three-quarters of a second. This window size design is small enough to achieve a sub-second time resolution.

Additionally, in order to have confident event detection, the system will compare the window signal energy to a noise model for anomaly detection. This noise model follows a Gaussian distribution [8], whose parameters will be updated based on non-event data continuously. If the current window signal’s energy value is beyond the threshold of N standard deviations from the mean of the noise, the system considers it as an event or part of an event. This threshold statistically reduces the false positive rate to below 1%. If next window’s energy also lays outside the threshold, the system labels the potential event as an event.

Limited by device’s function, there are situations of signal saturation. Since the local energy of the saturated section is larger than threshold, signal saturation has trivial impact on event detection. In the future, this problem will be fixed with updated devices.

Figure 4 presents the zoomed visualization at one six-second interval of the vibration data and two events detected.

2.3 Feature Extraction

We conduct feature extraction on each detected event signal. Since the detected events may have different duration, we first calculate the average power (AP) of the signal segment as a feature, which is the signal energy divided by the signal length. Figure 5 shows the frequency spectrum plot of one event signal S generated from the Fast Fourier Transform (FFT) in the frequency domain ranges from 0 Hz to 100 Hz. We empirically notice that most of the event signal’s frequency spectrum shows six significant peaks, which are

Figure 5: An example event signal’s spectrum plot.

marked with triangles in the example signal in Figure 5. We select six frequency peaks with the highest amplitude and use their peak value ($Freq_{pV}$) and location ($Freq_{pL}$) as features.

To compare the event signal’s frequency component distribution, we further normalize the event signal segment by its signal energy based on Equation 1.

$$S_{Norm} = \frac{S}{\sqrt{\sum_{i=1}^N S_i^2}} \quad (1)$$

where S denotes the signal in the time domain and S_i denotes data point in the signal segment. We conduct FFT analysis on S_{Norm} as well, and calculate the frequency domain signal energy for each 5 Hz frequency band ($NormFreq_{5hz}$). Since our system samples at 1000 Hz, we can obtain 100 $NormFreq_{5hz}$ values between 0 and 500 Hz. The final feature array is summarized in Equation 2.

$$F = [AP, Freq_{pV,1}, \dots, Freq_{pV,6}, Freq_{pL,1}, \dots, Freq_{pL,6}, NormFreq_{5hz,1}, \dots, NormFreq_{5hz,100}] \quad (2)$$

Since the last 100 features are extracted from the normalized signal, the value distribution of the first 13 features is highly skewed compared to the last 100. Additionally, the scale of the features is significantly uneven between the first 13 features and the last 100 features. Therefore, a log transformation was performed to the first 13 features. Then, the transformed data was standardized with a Min-max scaler. Due to the limited motion during the sleep, i.e., the number of events detected, the length of our feature set F is often larger than the number of data entries, and when this occurs, the classifier is prone to over-fitting [13].

To tackle this problem, a principal component analysis (PCA) was performed. We observed from the cumulative explained variance that the top 17 principal components are capable of keeping more than 99 % of the original variance. As a result, we use these top 17 components as features to build the learning model.

2.4 Sleep Stage Learning Model

The intuition that the detected motion-induced vibration events can be distinguished by the selected features is that in different sleep stages, the strength and range of the motion may vary, which results in different characteristics of the vibration signals. We use a scalable machine learning system for tree boosting named XGBoost. XGBoost is a shorten name for extreme gradient boost developed

Figure 6: Average number of events detected in labeled epochs.

by Chen and Guestrin [2]. XGBoost is a regularized extension of traditional boosting ensemble techniques that belongs to the classification and regression tree (CART) family. The prediction of a new classification entry after each iteration is made in an additive way, meaning that the predictions are made from weak classifiers that are continually being improved over the previous error. Therefore, the classifiers affected by data set bias could be improved, which is suitable for out imbalanced data set. Higher weights will be given to incorrectly classified samples at the next step, forcing the classifier to improve the performance in the following iterations. The final classifier is the strongest one as it includes the combined improvement of all the previous modeled trees. So far, the conventional gradient boosting machine (GBM) lacked a robust regularization factor, which made them vulnerable to overfitting. XGBoost overcomes this defeat by providing a stronger regularization framework that constrains overfitting, which could also select the critical features.

3 EVALUATION

We evaluate our system with real-world deployment and data collected from four patients (one night each) at the National Children’s Hospital. The data was gathered in the sleep laboratory at the hospital in 08/2018 over three nights. Time interval from first day’s 5 p.m. till second day’s 7 a.m. is defined as one night. The data is de-identified, and only the vibration data epoch that is labeled by sleep experts are used in this work in consideration of the Data Usage Agreement to the hospital and to protect the patients’ privacy. The dataset is prepared by extracting the 6-second vibration signal around the time of epochs that experts’ labeled. Five types of labels are provided including ‘W’, ‘R’, ‘N1’, ‘N2’, ‘N3’, which are corresponding to wake, REM, and three stages of NREM. In the process of learning model tuning, over 10,000 XGBoost models with different hyperparameters were established and validated with five stratified cross validation. To tackle the inherent imbalance, bootstrap and oversampling techniques were conducted in model tuning. The models were deployed in Google Colab environment based on Python 3.6. The grid-search technique was implemented for model tuning, of which the evaluation metric was average AUC score after stratified cross validation.

Table 1: Typical Confusion Matrix of Leader Model

Actual \ Pred	NREM	REM	Error Rate
NREM	26	6	0.1875
REM	1	7	0.125

Table 2: Confusion Matrix of Test Data

Actual \ Pred	NREM	REM	Error Rate
NREM	7	1	0.125
REM	1	1	0.5

3.1 Motion Event Detection Analysis

We applied the event detection algorithm on four patients/nights of data provided. We observe that the motion variation in N1 stage has a significant difference over patients. Overall, the average number of event detected during ‘W’ P1 had the highest number of ‘R’ epochs (up to 39) and demonstrates a low motion activity over the R stage. P2 had the lowest activity level during N2, and their activity levels in other sleep stages were stable. P3 had zero ‘R’ epoch, and their motion was mainly concentrated in N2 and N3. It is believed that with more patients enrolled and a long-term monitoring, we would be able to find the statistically significant correlations between their motions over different sleep stages and their conditions.

3.2 Sleep Stage Recognition Analysis

Since the data is imbalanced and the task is binary classification, we use the AUC score as the validation metric. Table 1 depicts the typical confusion matrix of the leader model with an average AUC at 0.84 from stratified five-fold cross validation. A random stratify split test data set was utilized for the performance test. In the typical validation performance, the model achieved an accuracy of 0.825. However, due to the limitation of size and data’s inherent imbalance, the performance on the test set shows an accuracy of 0.8. The model has the capability to handle the imbalanced data, forcing the prediction of the less weighted class to have a stronger performance. This result demonstrates the possibility to classify the sleep stage between REM and NREM with highly imbalanced bed frame vibration data from multiple users.

4 CONCLUSION

This paper presents a device-free sleep monitoring system, which employs a single geophone sensor and one sensor node. The system detects the user’s in-bed motions in a contact-less approach through bed frame vibration sensing, and conduct sleep stages recognition using XGBoost. Our system has multiple advantages over the conventional contact-based setup, including 1) non-intrusiveness, 2) easy to deploy for in-home long-term monitoring, 3) require a relatively short training phase. Stratified cross-validation and prediction test achieve an area under curve (AUC) score of 0.84 when recognizing users’ sleep stages over three participants’ data.

In the future, we will collect more sleep data of the same participant on different nights and different participants. Our prior work on footstep detection has reached a 0.99 accuracy, the event detection module in this system will be further developed to detect

and classify different classes of events. Various machine learning algorithms will be explored in next step, to establish a more comprehensive model with enhanced classification accuracy. Furthermore, we want to establish the correlation between the metrics measured by device-free sensing system and the sleep quality measured by experts and clinical equipment.

ACKNOWLEDGMENTS

We would like to thank Dr. Mark Splaingard and Tim Held at the Sleep Disorders Center at Nationwide Children's Hospital for their help and support with recruitment and data collection. This research was supported by the Highmark and National Science Foundation CMMI-1653550 funding.

REFERENCES

- [1] Donald L Bliwise. 2004. Sleep disorders in Alzheimer's disease and other dementias. *Clinical cornerstone* 6, 1 (2004), S16–S28.
- [2] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. ACM, 785–794.
- [3] Massimiliano de Zambotti, Fiona C. Baker, Adrian R. Willoughby, Job G. Godino, David Wing, Kevin Patrick, and Ian M. Colrain. 2016. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiology & Behavior* 158 (2016), 143 – 149. <https://doi.org/10.1016/j.physbeh.2016.03.006>
- [4] Ingeborg H Hansen, Mikkel Marcussen, Julie AE Christensen, Poul Jennum, and Helge BD Sorensen. 2013. Detection of a sleep disorder predicting Parkinson's disease. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 5793–5796.
- [5] Takahiro Komine, Keiki Takadama, and Seiji Nishino. 2016. Toward the next-generation sleep monitoring/evaluation by human body vibration analysis. In *2016 AAAI Spring Symposium Series*.
- [6] MO Mendez, Matteo Matteucci, Sergio Cerutti, AM Bianchi, and Juha M Kortelainen. 2009. Automatic detection of sleep macrostructure based on bed sensors. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 5555–5558.
- [7] Andrew Keong Ng and Cuntai Guan. 2012. Impact of obstructive sleep apnea on sleep-wake stage ratio. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 4660–4663.
- [8] Shijia Pan, Amelie Bonde, Jie Jing, Lin Zhang, Pei Zhang, and Hae Young Noh. 2014. Boes: building occupancy estimation system using sparse ambient vibration monitoring. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2014*, Vol. 9061. International Society for Optics and Photonics, 90611O.
- [9] Allan Rechtschaffen. 1968. A manual of standardized terminology, technique and scoring system for sleep stages of human subjects. *Public Health Service* (1968).
- [10] Ulrich Reimer, Sandro Emmenegger, Edith Maier, Zhongxing Zhang, and Ramin Khatami. 2017. Recognizing Sleep Stages with Wearable Sensors in Everyday Settings.. In *ICT4AgeingWell*. 172–179.
- [11] Avi Sadeh and Christine Acebo. 2002. The role of actigraphy in sleep medicine. *Sleep Medicine Reviews* 6, 2 (2002), 113 – 124. <https://doi.org/10.1053/smr.2001.0182>
- [12] S. Tomura, T. Harada, H. Sato, K. Takadama, and M. Aoki. 2015. Estimating surrounding symptom level of dementia patient by sleep stage. In *2015 9th International Symposium on Medical Information and Communication Technology (ISMICT)*. 190–194. <https://doi.org/10.1109/ISMICT.2015.7107526>
- [13] Guoyin Wang, Qing Liu, Andrzej Skowron, Yiyu Yao, et al. 2003. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing: 9th International Conference, RSFDGrC 2003, Chongqing, China, May 26-29, 2003, Proceedings*. Vol. 2639. Springer Science & Business Media.
- [14] Bülent Yılmaz, Musa H Asyali, Eren Arıkan, Sinan Yetkin, and Fuat Özgen. 2010. Sleep stage and obstructive apneic epoch classification using single-lead ECG. *Biomedical engineering online* 9, 1 (2010), 39.