A Personalized Approach for Detecting Unusual Sleep from Time Series Sleep-Tracking Data

Zilu Liang
Artificial Intelligence Research Center
National Institute of Advanced Industrial Science and Technology
Faculty of Engineering, Univ. of Tokyo
Tokyo, Japan
zilu.liang@aist.go.jp

Mario Alberto Chapa Martell
Faculty of Engineering, Univ. of Tokyo
Tokyo, Japan
mchapa0300@gmail.com

Takuichi Nishimura
Artificial Intelligence Research Center
National Institute of Advanced Industrial Science and Technology
Tokyo, Japan
takuichi.nishimura@aist.go.jp

***Abstract*— Nowadays emerging sleep-tracking technologies such as Fibit make it possible for individuals to collect personal sleep data. However, people find it difficult to gain insights from these data without proper analysis. The objective of this study was to investigate the possibility of establishing a sleep analysis approach that helps people detect their unusual sleep pattern by considering their own sleep baselines instead of the population average. The proposed approach was consisted of two steps. In the first step, the dimension of time series sleep data was reduced using permutation entropy. Following that, univariate outlier detection techniques were applied to detect unusual sleep patterns. We tested our approach on a real sleep tracking data set consisting of 35 days of time series data tracked using a Fitbit Charge HR. Depending on the univariate outlier detection technique used, the identified unusual sleep differed. We found that permutation entropy of a sleep time series was strongly correlated to the time that the user went to bed and weekly correlated to minutes asleep, but was not correlated to minutes awake, awakening count and sleep efficiency. Based on the analysis results, we pointed out the directions for future study on personal sleep data analysis.**

Keywords—Personal informatics; sleep; health; wearable computing; time series mining; outlier detection; entropy

I. Introduction

Personal sleep-tracking is an emerging area that interests both academia and industry. Many sleep-tracking tools have been developed to help people collect personal sleep data on daily basis and in an unobtrusive manner [1,3-4]. Most of these tools are based on the principle of actigraphy that infers sleep stages according to the movement of the person. Previous studies suggested that actigraphy measurements were sufficiently accurate for healthy adults when validated against Polysomnograpy (PSG) measurements [2]. While collecting sleep data is becoming easier, it remains difficult for users to interpret their sleep data and to gain actionable knowledge for sleep improvement. For example, users cannot tell whether one night's sleep was usual or not, letting alone the causes for unusual sleep.

We defined unusual sleep as the sleep pattern that significantly deviate from an individual's sleep baseline. The protocol used in standard PSG test in clinical settings are not applicable to personal sleep tracking, as they were developed for diagnose sleep disorders, and the unusual sleep patterns of an individual are not necessarily sleep disorders. For example, 6-hour sleep may be considered as sleep deprivation for most people, but some people feel perfectly fine only sleeping 6 hours daily. On the other hand, 8-hour sleep should be considered unusual for people who habitually sleep 9 hours. Analysis on personal sleep data is a new field which requires the establishment of a new and systematic framework to satisfy individual's needs rather than simply borrowing existing protocols from clinical settings, because the purpose of such analysis is mostly to help users establish a personal sleep baseline and is not the diagnosis of sleep disorders.

The goal of this study is to propose a method for detecting unusual sleep from self-tracking sleep data by counting in personal differences. We adopted a general approach to define "unusual" as the statistical outliers from each user's own baseline sleep patterns. The reason for using time series data (minute-by-minute sleep data) instead of aggregate data (total sleep duration, number of awakenings, etc.) was that aggregate sleep summary data were derived from time series data using certain algorithms, which may increase uncertainty of the data. The proposed approach is consisted of two steps. In the first step the high-dimensional time series data were reduced to one dimensional data using Permutation Entropy (PE) [5], which represents the complexity of a time series data. We adopted permutation entropy (PE) for dimension reduction as the targeted time series were not equal length (i.e. sleep data could last between 6-9 hours). In addition, permutation distribution is robust to thermal drift in time series sleep data collected using wearable devices. In After computing the PE of all sleep time series, unusual sleep were detected using typical univariate outlier detection techniques. We will describe our method in detail in the following section.

The main contribution of this paper is that we proposed a personalized approach for detecting unusual sleep by considering personal baselines rather than population average, thus making the detection more adapted to personal context.

Our approach is distinct from existing sleep quality evaluation approach. Through this work we hope to inspire future research on the establishment of personalized sleep analysis framework.

II. Related Work

A. Sleep-tracking Technologies

With the development of pervasive computing and wearable computing technologies, personal sleep tracking becomes more and more affordable. Currently a large number of commercial sleep tracking tools exist in the consumer market, ranging from sleep inducing (e.g. White Noise), smart waking (e.g. Smart Alarm Clock) to sleep pattern tracking (e.g. Fitbit and Jawbone). Most of these tools are based on the principle of actigraphy that provides reasonable accuracy for normal, healthy adult populations [2]. Total sleep hour is a popular metric for evaluating sleep quality in commercial sleep-trackers, and 8-hour is a widely adopted cutting-off point. Whereas it became easier for people to collect personal sleep data, these data were rarely systematically analyzed and users found it difficult to gain insights and actionable knowledge [6].

B. Outlier Detection in Data Mining

Outlier detection is a very important task in data mining. According to [7], an outlier "is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism". Existing outlier detection methods can be classified into six categories: statistical test [8], depth-based approach [9], deviation-based approach [10], distance-based approach [11], density-based approach [12], and high-dimensional approach [13]. Outlier detection techniques have been applied to many areas including fraud detection [14], intrusion detection [15], medical condition monitoring [16], and pharmaceutical research [17].

Very recently, outlier detection was applied to detect abnormal sleep pattern in [18], in which median absolute deviation (MAD) was used to detect abnormal bedtime, wake time and sleep duration. Since abnormal sleep patterns were detected based on the statistical characteristics of the datasets, personal differences can be counted in the analysis. However, only three sleep metrics were considered and much information in sleep time series was lost. Built on this work, we proposed a new approach for unusual sleep pattern detection from sleep time series data. We decided to directly conduct analysis on time series sleep data to avoid potential errors in aggregate sleep data which were usually derived from time series data.

In the proposed method, sleep time series data were first converted to permutation entropy (PE) which represents the information contained in each time series, and then typical univariate outlier detection methods were applied to find the PE outliers as unusual sleep.

III. Proposed Approach

This study aims to investigate the possibility of establishing sleep analysis protocol that counts in personal differences, assuming that individuals tracked their sleep using wearable devices and such data were sufficiently reliable. Instead of using a population-level average as cutting-off point for usual and unusual sleep (e.g. minutes asleep = 8 hours, or sleep efficiency ≥ 95%), we applied time series data mining techniques to adaptively identify unusual sleep by considering personal sleep baselines. In practice, the minute-by-minute time series sleep data collected by wearable devices (e.g., Fitbit) or embedded sensors (e.g., Beddit) is in fact the movement level of a user detected by accelerometer. We decided to work on minute-by-minute sleep time series data instead of daily aggregate sleep data due to the following two reasons: (1) time series data contains important tempo-spatial information; (2) aggregate data is derived from time series data and thus may contain algorithm-specific errors. Based on such personalized analysis, individuals would find it easier to make sense of their sleep-tracking data rather than simply being horrified by the comparison to population-level average. This is especially true for individuals who have distinct sleep patterns from other people. In what follows we will describe the problem formulation and the proposed algorithm in detail.

A. Problem Formulation

The goal of the data mining task was to detect unusual sleep nights from a set of sleep time series data. Unusual sleep referred to the sleep that significantly deviated from an individual's basic sleep pattern.

Given D days of time series sleep-tracking data $\boldsymbol{X} = \{X_{\boldsymbol{d}}\}_{d=1}^{D}$ where $X_d = \{x_d(i)\}_{i=0}^{T_d}$, sampled at equal intervals with $x_d(i) \in \{1,2,3\}$, find $\boldsymbol{X}_{unusual} = \{X_{\boldsymbol{u}}\}_{u=1}^{U} \subset \boldsymbol{X}$ that satisfies (1).

$$|X_u - X_0| > t\xi \tag{1}$$

where $X_{\boldsymbol{0}}$ is the median of the time series data set, $|X_u - X_0|$ represents the distance between $X_{\boldsymbol{0}}$ and $X_{\boldsymbol{u}}$, ξ is the natural variation of the data set and t is a threshold parameter.

B. Proposed Unusual Sleep Detection Algorithm

Our proposed approach for detecting unusual sleep was consisted of two steps: dimension reduction and outlier detection.

1) Dimension reduction.

There are two general approaches for time series analysis: curve based approach and feature based approach. Personal sleep-tracking data from wearable devices are not equal length, as individuals not necessarily sleep the same hours every day. It is thus not possible to directly compare sleep time series based on signal curve. We adopted feature based approach instead, which reduced high dimensional time series sleep data to low dimensional features. We selected permutation entropy (PE) as the key feature to represent the complexity of a sleep time series, as PE not only allows the comparison of time series of varying lengths but also is robust against observation or dynamical noise in signal, which is a common problem in accelerometry of wearable sensors.

Given a time series $= \{x(i)\}_{i=0}^{T}$, partitioned into subsequences of a fixed length m and with time delay τ, we can denote the new series as $X' = \{[x(i), x(i+\tau), x(i+2\tau), \ldots, x(i+(m-1)\tau)\}_{i=0}^{T'}$ with a total of $T' = T-(m-1)\tau$ elements. The parameter m is called embedding dimension, which mapped the original time series into an m-dimensional space with time delay τ. For each subsequence $x' \in X'$, the m values are put into sorted order to compute the permutation of indices from 0 to (m-1). Each possible permutation of length m is called an ordinal pattern. The permutation distribution (PD) of a time series is obtained by counting the frequencies of the distinct observed ordinal patterns of the elements $x' \in X'$. Let $\prod(x')$ be the permutation that an element $x' \in \mathbb{R}^m$ undergoes when being sorted, the PD of X' could be calculated by (2).

$$p_\pi = \frac{\#\{x' \in X' | \prod(x') = \pi\}}{T'} \quad (2)$$

The PE of order $m \geq 2$ is defined as the Shannon entropy of the probability distribution P [19], represented by (3).

$$H(P) = -\sum_{\pi \in S_m} p_\pi \, log p_\pi \quad (3)$$

where S_m is the set of all m-permutations.

The information content of a PD depends crucially on the choice of the embedding dimension. Too small embedding dimensions narrow the representational power of the distribution; too large embedding dimensions dilute the estimation of the distribution. We used the Minimum Entropy Heuristic (MinE) to automatically choose an embedding dimension with optimal representational entropy as proxy for representational power. As is shown in (4), MinE chooses the embedding dimension with the lowest average, normalized PE.

$$arg\, min_m \sum_{P \in \wp} e_P\,(m) \quad (4)$$

and

$$e_P(m) = \frac{\sum_{\pi \in S_m} p_\pi \log_0 p_\pi}{\log(m!)} \quad (5)$$

where $\wp$ is a set of permutation distribution of embedding dimension m, $\log_0(x)$ is 0 if $x = 0$ and $\log(x)$ otherwise.

2) Unusual sleep detection

We take a general approach to define "unusual sleep" as the statistical outliers in the sleep time series data set. After converting the sleep time series data to one-dimenional PE, it is possible to apply a wide variety of established univariate outlier detection methods.

We selected two widely used outlier detection techniques to detect unusual sleep and compared their results in the next section. The first technique was the S_n estimator proposed in [20], a widely used deviation-based approach for detecting global outliers. The S_n is an alternative of the robust scale estimator Median Absolute Deviation (MAD) for skewed distribution.

$$S_n = c * med_i\{med_j |x_i - x_j|\} \quad (6)$$

where c is a constant of value 1.1926. For a given sleep PE x_i, we first calculated the median of the distance between x_i and all other nights' PE x_j to obtain $med_j|x_i - x_j|$. If $med_j|x_i - x_j|/S_n \geq \tau$, we considered x_i as an outlier and thus the corresponding sleep was identified as unusual. The threshold τ was set to 2 based on previous study [21].

The second technique was Local Outlier Factor (LOF) proposed in [12], a widely used density-based approach for detecting local outliers. The local reachability distance (lrd) of a sleep PE point x_i was calculated using (7).

$$lrd_k(x_i) = 1/\left(\frac{\sum_{x_j \in kNN(x_i)} r\text{-}dist_k(x_i, x_j)}{Card(kNN(x_i))}\right)\} \quad (7)$$

where

$$r - dist_k(x_i, x_j) = \max\{\mathrm{k-distance}(x_j), \mathrm{dist}(x_i, x_j)\} \quad (8)$$

Then the LOF of x_i was calculated using (8).

$$LOF_k(x_i) = \frac{\sum_{x_j \in kNN(x_i)} \frac{lrd_k(x_j)}{lrd_k(x_i)}}{Card(kNN(x_i))} \quad (8)$$

IV. Evaluation Results

We applied the proposed approach to a sleep-tracking dataset for preliminary evaluation. In this section we present the analysis result, investigate the characteristics of the proposed approach, and summarize implications for future study.

A. Data Collection and Cleaning

The data set that we used for evaluating the proposed method was collected for 35 days during August and September 2015 using a Fitbit HR Charge. Ethics approval was obtained from University of Melbourne (HREC #1339336) to conduct self-tracking experiment on a healthy female who did not have diagnosed sleep problems. The subject was randomly selected and we did not intend to draw any gender specific conclusions. We were granted access to intra-day minute-by-minute sleep-tracking data through Fitbit Partner API which is not open to general public. Sleep-tracking data was sampled at constant interval of 60 seconds. Each time series last between (412, 553) minutes with standard deviation of 40 minutes. For the purpose of comparison, we also retrieved aggregate sleep data including minutes asleep, minutes awake, awakening count, bed time, and sleep efficiency. Bed time refers to the time a user went to bed and it was converted to numerical values, e.g., 20:00 → 2000 and 22:30 → 2230. Missing data entries were removed and thus there were 32 sleep-tracking time series in the dataset, which satisfied the 30-day requirement for personal data analysis [21].

B. *Selecting Parameter Values*

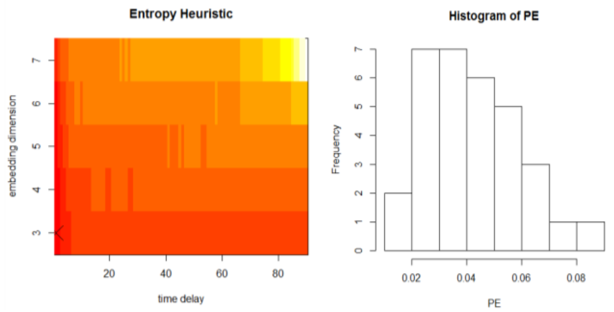


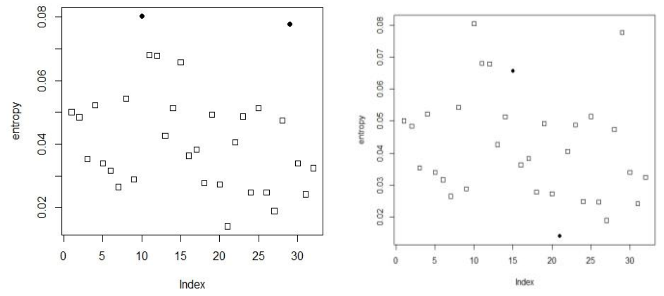
Fig. 1. (Left) Visualization of the average entropy estimates for varying time delays and embedding dimensions on the collected sleep-tracking time series dataset ; (Right) Histogram of PE.

The parameters involved in calculating PE were decided using entropy heuristic that was described in previous section. We investigated the change of permutation entropy as the embedding dimension m was set to {3, 4, 5, 6, 7} and time delay τ was set $\{1 \leq \tau \leq 90 | \tau \in R\}$ (since on average a normal cycle of human sleep last approximately 90 minutes). The entropy heuristic showing Figure 1 (left) suggested that m and τ should be set to 3 and 1 respectively. Under these parameter values, the PE distribution of the data set is shown in Figure 1 (right). For LOF, the number of neighbors k was heuristically set to 5.

C. *Unusual Sleep Detection*

The output of the S_n technique was the identified unusual sleep, while that of the LOF technique was a LOF score. The former discovered 2 usual sleep time series characterized by the largest PE. As for the latter, we sorted the LOF score in decreasing order and only kept the top two time series as identified unusual sleep for the purpose of comparison. Interestingly, one of the unusual sleep time series discovered using LOF was the one with the lowest PE. Figure 2 shows the entropy of identified unusual sleep represented by dots. The raw time series sleep data are plotted in Figure 3, with the identified unusual ones highlighted in red.

We also compared the result of the proposed method with typical population-level approaches that are one-fit-all in nature. The two baseline approaches used minutes asleep (8-Hour Rule) and sleep efficiency (95% Rule) respectively for evaluating sleep quality. The cutoff threshold were 8 hours for minutes asleep (good sleep quality: minutes asleep ≥ 8 hour) and 95% for sleep efficiency (good sleep quality: sleep efficiency ≥ 95%). As is illustrated in Figure 4, the baseline approaches generated many false alarms because personal sleep context was not considered. The subject concluded that her sleep needs was only approximately 7 hours and longer sleep did not improve day time functioning. Similarly, sleep efficiency lower than 95% usually did not cause problem either. Most of the unusual sleep patterns identified by the existing approaches were considered neither informative nor helpful by the subject. In comparison with population-level approaches, the proposed approach identified unusual sleep within the context of the subject's baseline and thus provided more targeted information.



(1) (2)

Fig. 2. Entropy of identified unusual sleep (represented by dots). (1) Detection result using S_n; (2) detection result using LOF.

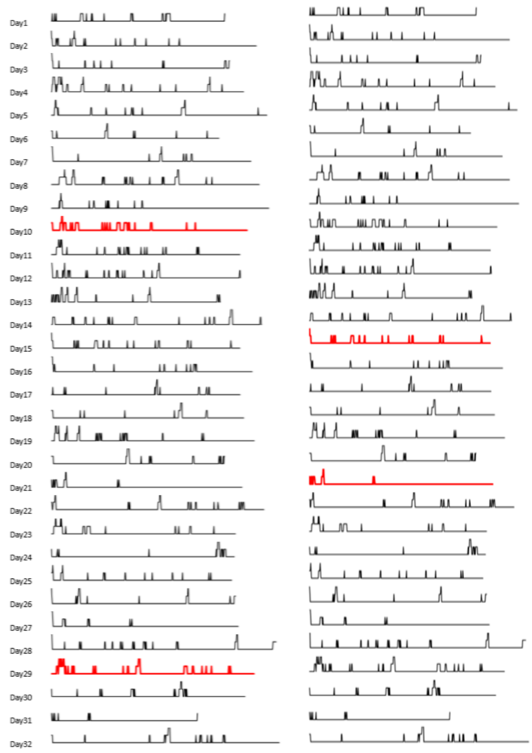


Fig. 3. Raw time series data of sleep. The identified unusual sleep were highlighted in red. (left) Unusual sleep detected by S_n technique, (right) unusual sleep detected by LOF.

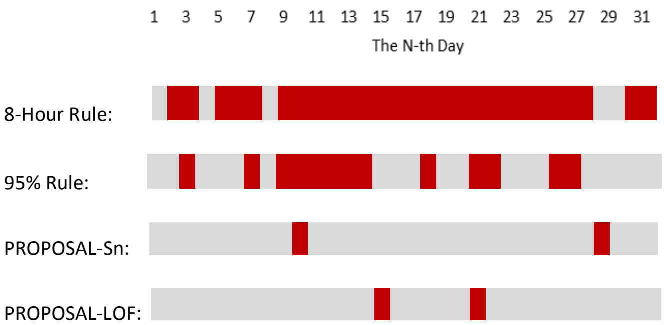


Fig. 4. Identified unusual sleep using existing and proposed approaches. Red blocks indicate unusual sleep and grey blocks indicate usual sleep.

D. PE v.s. Aggregate Sleep Metrics

As is mentioned before, dimension reduction on time series sleep data using Permutation Entropy (PE) has several benefits. However, it is difficult to interpret the mining results without understanding the physical meaning of PE in the context of personal sleep. We assumed that higher PDE indicated higher level of restlessness during sleep, which was characterized by long minute awake and high awakening count. To test the hypothesis, we investigated the relationships between the PE of a sleep time series and the corresponding aggregate sleep data on the same day, as is shown in Figure 5. It is demonstrated that PE was strongly correlated to bed time and weakly correlated to minutes asleep, but was not correlated to minutes awake, awakening count and sleep efficiency.

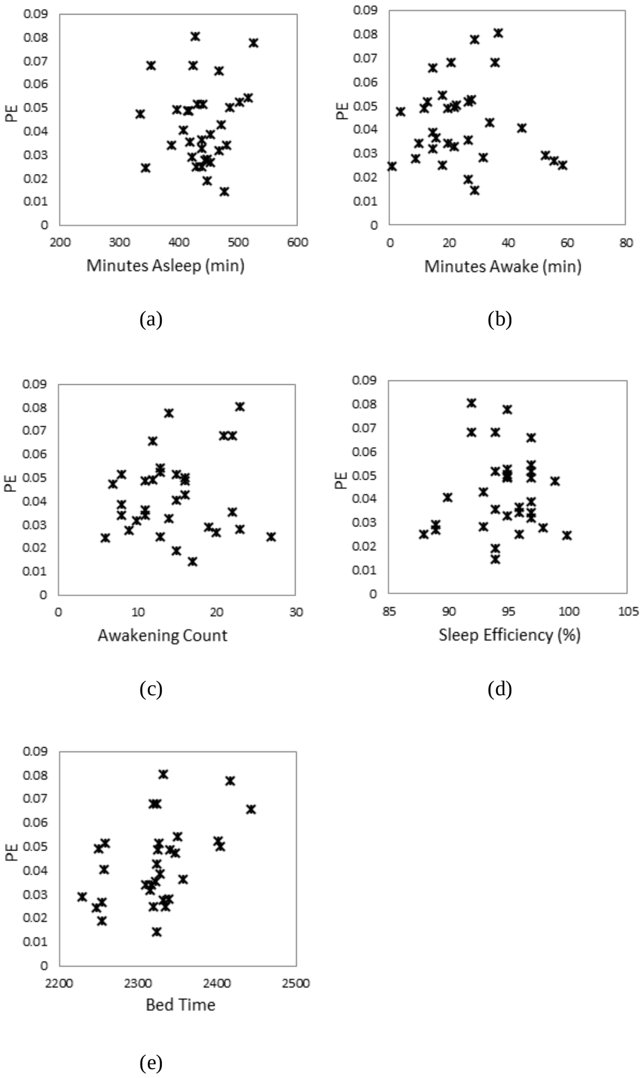


Fig. 5. The relationships between PE and aggregate sleep metrics. (a) PE v.s. minutes asleep (min); (b) PE v.s. minutes awake (min); (c) PE v.s. awakening count; (d) PE v.s. sleep efficiency (%); (e) PE v.s. bed time.

E. Discussions

1) Unusual sleep is not equivalent to bad sleep quality. Since our definition of “unusual” was the statistical outliers that were either much larger or much smaller than the medians, there are two types of unusual sleep, either extremely good sleep (positive unusual) or extremely bad sleep (negative unusual). In other words, unusual sleep does not necessarily mean bad sleep. From Figure 3 it is not hard to notice that the identified unusual sleep either represented very restless sleep (Day 10, 29) or very well-rested sleep (Day 21). According to our previous study, users tend to pay more attention to negative results [22], but the proposed approach does not differentiate the two cases. It is necessary to tweak the approach to only detect negative unusual sleeps.

2) Limitations of this study. Mining personal sleep-tracking data is nontrivial task. There are two fundamental limitations of this study. First, there is no ground truth for rigid evaluation of the proposed approach. In fact the usual sleep baseline for each individual is hard to obtain because there is no methodological support from the sleep research community. One solution could be replacing objective ground truth with subjective judgement as was done in [23], where the statistical analysis results from personal data were validated against the subjective perception of the data owners. Second, the issue of data quality was not considered. Commercial wearable devices have inherient limitations. The collection of self-tracking data was not done in a controlled manner and many human errors could be involved. Therefore, the identified outliers were in fact a mixture of measurement errors, contextual outliers and real sleep outliers (which could be further classified into good outliers and bad ones). We intend to address the above limitations in our future work.

In addition to the fundamental limitation, this study also has the following methodological limitations. First, the PE that we adopted for dimension reduction purpose could only characterize the information of the signal shape and thus the time stamp information was missing. In practice, however, the timing and duration of a certain event during sleep may have clinical significance. For example, awakenings happened in midnight was considered more abnormal than awakenings happened in the morning [24]. Second, our approach based on information theory could only capture limited features of a time series data. Aggregate, time-domain, and frequency-domain features may also contain important information that could help better differentiate unusual sleep from usual ones.

V. Conclusion

In this paper we proposed an approach for personalized analysis on sleep-tracking time series data. The proposed approach was consisted of two steps, dimension reduction using permutation entropy and unusual sleep detection using statistical approaches. We applied the proposed approach to a sleep time series data set collected from a healthy female. The results showed that the proposed approach could successfully identify unusual sleep time series characterized by either extremely large or small permutation entropy, representing extremely bad or good sleep respectively. However, it was hard to evaluate the accuracy of the proposed approach due to the lack of ground truth. In the next step, we plan to apply the proposed approach to more datasets and compare the analysis results with users' subjective judgement. We also intend to improve the comprehensiveness of the proposed approach by considering more features of personal sleep-tracking data as well as addressing the data quality issue.

ACKNOWLEDGMENT

This study was supported by NEDO's Artificial Intelligence Research Project, Australia Endeavour Research Fellowship, and Microsoft Bizspark. We would like to thank Dr. Tang Jie of City University of Hong Kong, Prof. James Bailey, Prof. Lars Kulik and Dr. Yuxuan Li of University of Melbourne for very inspiring and helpful discussions on this work.

REFERENCES

[1] T. Hao, G. Xing, and G. Zhou. "iSleep: unobtrusive sleep quality monitoring using smartphones", In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (SenSys '13). ACM, New York, USA, Article 4, 14 pages, 2013.

[2] T. Morgenthaler, C. Alessi, L. Friedman, et al. "Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007", SLEEP, vol.30, no. 4, p.519-529, 2007.

[3] Z. Chen, M. Lin, F. Chen, et al. "unobtrusive sleep monitoring using smartphones," in Pervasive Computing Technologies for Healthcare (PervasiveHealth), pp.145-152, 2013.

[4] J.-K. Min, A. Doryab, J. Wiese, et al. "Toss 'n' turn: Smartphone as sleep and sleep quality detector," in Proc. of CHI 2014, 2014.

[5] A. M. Brandmaier, "pdc: An R package for complexity-based clustering of time series," Journal of Statistical Software, vol. 67, no. 5, 2015.

[6] Z. Liang, W. Liu, B. Ploderer, et al, "Designing intelligent sleep analysis systems for automated contextual exploration on personal sleep-tracking data", New Frontiers in Artificial Intelligence. (In press)

[7] D. Hawkins. Identification of Outliers. Chapman and Hall, London, 1980.

[8] V. Barnett and T. Lewis. Outliers in Statistical Data. John Wiley, 3rd edition, 1994.

[9] J. Turkey. Exploratory Data Analysis. Addison-Wesley, 1977.

[10] T. Palpanas, D. Papadopoulos, V. Kalogeraki, D. Gunopulos, "Distributed deviation detection in sensor networks", SIGMOD Record, vol.32, no.4, p.77-82, 2003.

[11] E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large dataset", In Proc. of 24th International Conference on Very Large Data Bases (VLDB'98), New York, NY, pp 392-403, 1998.

[12] M. Breuning, H-P. Kriegel, R. Ng, and J. Sander, "LOF: Identifying density-based local outliers", In Proc. of 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD'00), Dallas, Texas, pp 93- 104, 2000.

[13] C. C. Aggarwal and P. S. Yu, "Outlier detection in high dimensional data", In Proc. of 2001 ACM SIGMOD International Conference on Management of Data (SIGMOD'01), Santa Barbara, California, USA, 2001.

[14] T. Fawcett and F. Provost, "Adaptive fraud detection," Data Mining and Knowledge Discovery Journal, vol.1, no.3, p.291-316, 1997.

[15] G. Singh, F. Masseglia, C. Fiot, et al, "Data Mining for intrusion detection: from outliers to true intrusions," Advances in Knowledge Discovery and Data mining, vol. 5476, p.891-898, 2009.

[16] M. Hauskrecht, I. Batal, M. Valko, et al, "Outlier detection for patient monitoring and alerting," J. Biomed. Inform. Vol.46, no.1, p.47-55, 2013.

[17] M. F. Engels, L. Wouters, R. Verbeeck, and G. Vanhoof, "Outlier mining in high throughput screening experiments," J. Biomol. Screen, vol.7, no.4, p.341-351, 2002.

[18] K. Huang, X. Ding, J. Xu, et al, "Monitoring sleep and detecting irregular nights through unconstrained smartphone sensing," 2015.

[19] C. Bandt, B. Pompe, "Permutation entropy: a natural complexity measure for time series," Physical Review Letter, vol.88, no.17, 174102-1-174102-4, 2002.

[20] P. J. Rousseeuw and C. Croux, "Alternative to the median absolute deviation," Journal of the American Statistical Association, vol. 88, no. 424, p.1273-1283, 1993.

[21] R. V. Hogg, E. A.Tanis, and M. J. M. Rao. Probability and statistical inference. McMillan New York, vol. 993, 1977.

[22] Z. Liang, B. Ploderer, W. Liu, et al, "SleepExplorer: A visualization tool to make sense of correlations between personal sleep data and contextual factors", Personal and Ubiquitous Computing. (In press)

[23] A. Doryab, M. Frost, M. Faurholt-Jepsen, et al, "Impact factor analysis: combining prediction with parameter ranking to reveal the impact of behavior on health outcome", Pers. Uniquit. Comput, vol.19, p.355-365, 2015.

[24] M. M. Ohayon, "Nocturnal awakenings and comorbid disorders in the American general population", Journal of Psychiatric Research, vol.43, p.48-54, 2009.