Prediction of Sleep Efficiency from Big Physical Exercise Data

Xiaoli Liu
Biomimetics and Intelligent Systems
Group, University of Oulu
Oulu, Finland

Satu Tamminen
Biomimetics and Intelligent Systems
Group, University of Oulu
Oulu, Finland

Topi Korhonen
Polar Electro Oy
Kempele, Finland

Juha Rönning
Biomimetics and Intelligent Systems
Group, University of Oulu
Oulu, Finland

Jukka Riekkö
Center for Ubiquitous Computing,
University of Oulu
Oulu, Finland

ABSTRACT

Physical exercise can improve sleep quality. However, how to perform physical exercise to achieve the best possible improvements is not clear. In this article, we build predictive models based on volume real data collected from wearable devices to predict the sleep efficiency related to users' daily exercise information. As far as we know, this is the first study to investigate insights of prediction of sleep efficiency from volume physical exercise data collected from real world.

CCS CONCEPTS

• **Information systems** → **Data analytics.**

KEYWORDS

Predictive models, physical exercise, sleep quality

ACM Reference Format:

Xiaoli Liu, Satu Tamminen, Topi Korhonen, Juha Rönning, and Jukka Riekkö. 2019. Prediction of Sleep Efficiency from Big Physical Exercise Data. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*, September 9–13, 2019, London, United Kingdom. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3341162.3347078>

1 INTRODUCTION

Approximately 30% of the general population experience sleep disruption, while 10% experience both sleep disruption

and daytime dysfunction consistent with a diagnosis of insomnia as defined by the National Institutes of Health (NIH). Sleep disruption potentially exacerbates various health problems, such as obesity, diabetes, and depression, and insomnia relates to development of hypertension, atherosclerosis, and acute myocardial infarction.

Physical activities and exercises may reduce the negative effects of stress on health-related outcomes (sleep quality, well-being, and affectivity) [11]. Dolezal et al. [3] conclude that the physical exercise is an effective feature for those who do not experience adequate sleep quantity or quality. Physical exercise is even considered to be a nonpharmacological therapy for insomnia, which is readily available.

However, there is limited research related to how physical exercise can improve sleep quality, in both prediction and evaluation perspectives. In addition, most research only involves a small amount of participants in the study, which restricts the generalizability of the results to the larger population. Our research addresses the critical challenges of prediction of sleep quality with analysing big physical exercise data collected from wearable devices. As far as we know, this is the first effort in the domain to predict the sleep efficiency based on users' exercise information using large datasets collected from real world.

The main contribution of this article is building predictive models based on volume real data collected from Polar Electro wearable devices to predict the sleep efficiency related to users' daily exercise information. Using the prediction results, we can provide corresponding exercise training plan to achieve good sleep quality. Our results show that increasing number of features will not dramatically improve the model accuracy, while collecting more samples related to each user will greatly improve the model performance.

The remainder of this article is organized as follows: Section 2 presents background and related work. Section 3 describes the datasets and data preparation. Predictive models are presented in Section 4. Conclusion and future work are described in Section 5.

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.

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

© 2019 Association for Computing Machinery.

ACM ISBN xxx-x-xxxx-xxxx-x

<https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx>

2 BACKGROUND AND RELATED WORK

Wearables have acquired wide popularity, because of their acceptable prices and easy-to-use features. Shin et al. review the state of the art and show that the research involving Wearable Activity Trackers (WAT) should consider both technological and non-technological aspects, because of the complex impacts of the devices on human behaviour [10]. A positive relationship between Wearable Fitness Technologies (WFT) and health benefits has been discovered and it can be suggested that WFT devices have potential to facilitate a change in health behaviour [5].

Technology advancement for processing big data collected from wearables presents a significant potential for exercise and sleep analytics. The monitoring of people during their daily life provides valuable insights into the behavioral patterns related to sleep. By correlating human activity sensed by wearables with sleep patterns, we pave the way to perform systematic and scalable analytics and provide insights to clinicians and individuals for the early diagnosis of sleep problems, which directly influence the quality of sleep of individuals. Furthermore, the development of pervasive computing solutions connects wearables to other sensors and actuators, allowing to perform the sleep diagnostics in the home of patients and to improve the sleep environment for patients. As an example, NIH developed the National Sleep Research Resources [1] for integrating heterogeneous data sources for clinical research in sleep.

However, most current research related to exercise and sleep focuses on statistical analysis. The effect of long-term moderate aerobic exercise on sleep, quality of life and mood of individuals with chronic primary insomnia were investigated in [8]. Myllymäki et al. [7] examined the effects of vigorous late-night exercise on sleep to test the recommendation that intensive exercising is not suggested within the last 3 h before bed time. Compared to previous research, we propose to build the predictive models to predict the sleep efficiency using exercise information based on big real datasets from a huge amount of real users of Polar Electro wearable devices.

3 DATASETS AND DATA PREPARATION

Datasets

Many people are utilizing different kinds of wearables to track their physical activities and sleep quality. In general, user data are stored in repositories, which potentially support deep analytics for discovering novel insights based on these huge data sets. In this research, we utilize anonymous datasets from Polar Flow database continuously collected from users for eight months with Polar wearables. The data integrate physical activities and sleeping data in mobile health repositories from Polar activity tracker, Polar fitness

tracker, Polar running watches, and Polar sleep plus. Table 1 presents a typical sample record with main features including user id, exercise date, start time, stop time, duration, calories consumption and heart rate. Table 2 shows a typical sleep sample with main features, such as user id, sleep starting time, waking time, sleep span, sleep efficiency, sleep continuity, and sleep feed back.

Sleep efficiency is a general metrics used in sleep science and it refers to the actual percentage of sleep time. More specifically, it is the ratio of the total asleep time compared to the total amount of time spent in bed in a night. Poor sleep is usually associated with a sleep efficiency lower than 85%, normal sleep with an efficiency 85% or higher, and good sleep with an efficiency higher than 90%. Sleep efficiency depends on many factors and is considered as a highly individual variable. Afterwards, we will define good and bad sleep of each person by using his or her mean and standard deviation of sleep efficiency. Sleep continuity is an unique parameter which is only calculated by Polar devices and it describes how continuous the sleep is [2]. Sleep Plus utilizes intelligent algorithms based on wrist movements and positions to track the sleep timing, duration, and quality [6][9]. Sleep continuity is described in five scales. The higher the value is, the better the continuity of the sleep is. More information related to sleep efficiency and continuity can be found from Polar Electro Oy Sleep Tracking with Polar Sleep Plus.

Data preparation

Data preparation is carried out to get the related high quality data. Exercise data and sleep data are stored in different datasets. One user may perform more than one physical exercise in a day, and may not track the sleep status. Exercise and sleep records based on user id and date are integrated. Records where sleep efficiency is less than 60, exercise duration is less than 10 minutes, calories values are N/A, calories consumption per hour is less than 50 cal and higher than 2000 cal are filtered out. Missing values in resting heart rate are imputed with 55bpm. TRaining IMPulse (Trimp) is used to measure the quantify training load. Trimp is formulated by using the heart rate, gender, and duration of the exercise. Table 3 lists the main investigated features used for predicting the sleep efficiency. The main features include users' total daily exercise information, last exercise information, and when user does the last exercise before sleep (span time between last exercise stop time to sleep starting time). Features related to total exercise are sum of exercises performed in one day.

4 MODEL

Generalized Boosted regression Models (GBM) are selected to predict the sleep quality based on exercise information. Gradient boosting produces the predictive models through

Table 1: An exercise sample with main features in dataset

user id	start date	stop date	start time	stop time	duration	calories	hr min	hr max	hr avg
000001	2018-7-25	2018-7-25	2018-7-25 19:08:21	2018-7-25 19:45:43	2242250	333	115	176	158

Table 2: A sleep sample with main features in dataset

user id	starting time	waking time	sleep span	sleep continuity	sleep efficiency	sleep feed back
000001	2018-8-26 23:30:11	2018-8-27 06:30:20	25209000	4.5	92	sleep well

Table 3: Main investigated exercise features or variables

Features/Variables abbreviations	Description
N	Number of exercise performed in one day
TD	Total duration in day's all exercise
TR	Total trimp in day's all exercise
TC	Total calories consumption in day's all exercise
T	Trimp in day's last exercise
$TDZone_i$ ($i=1,...,5$)	Heart rate duration in $zone_i$ in day's exercise
C	Calories consumption in day's last exercise
D	Duration in day's last exercise
TOS	Span time between last exercise stop time to sleep starting time
$DZone_i$ ($i=1,...,5$)	Heart rate duration in $zone_i$ in day's last exercise

an ensemble of weak prediction models. The idea of gradient booting was first observed by Leo Breiman and subsequently developed explicit regression gradient boosting algorithms by Jerome H. Friedman[4]. GBM are good learners with high accuracy. The model can list the importance of each feature and draw marginal plots to visualize the relationship between predictors and outcome, which are important for us to know how the exercise affects the sleep and which are important features. Deep learning is hard to give those information.

We start with a simple hypothesis that sleep efficiency is affected by total exercise (TD , TR , TC) performed in the day. Model 1 is built using three features as inputs. 80% data has been used for training and 20% are used for testing. The number of trees is set to 10000, cross validation fold is 10, depth is 3 and shrink arranges from 0.005 to 0.1. Then, we consider to build the model 2 with all the features presented in Table 3. Those features include the last exercise information and TOS besides the total exercise information. The majority of users have records less than 10. That is because users need to track both the exercise and sleep at the same time during the tracking period. In addition, we may not be able to access sufficient data, because Polar Sleep PlusTM is an algorithm just introduced in 2017. Users with many exercise and sleep records are very interested. A predictive model is built for those users.

Table 4 presents the performance of three models on the testing set, respectively. Model 1 represents the GBM using three features as inputs. Model 2 represents GBM using features in Table 3. Model 3 is the GBM in the dataset, where the number of exercise and sleep days are larger than 200. This subset has 57 users, while each user has more than 200 records.

Table 4: Model Performance

Model	RMSE on testing set
Model 1	2.753662
Model 2	2.724675
Model 3	1.908299

Root mean squared error (RMSE) is used to evaluate models' performance. The RMSE of model 1 on the testing set is 2.753662, which is quite near the RMSE of model 2 on the testing set. It seems that increasing the features will not significantly improve the model performance. One possible reason is that these features are not sufficient enough to enable a good model and other factors may also affect the sleep efficiency beside exercise. Another possible reason is that each user's records are less and it is hard for the model to capture enough information.

Figure 1: Performance on the training data set.

Figure 2: Performance on the testing data set.

The RMSE of model 3 on the testing set is 1.908299, which is greatly reduced compared to model 2 because each user has more than 200 records, which enable model 3 can capture more information for each user. Figure 1 and Figure 2 describe the real sleep efficiency compared to the predicted sleep efficiency by model 3 on the training and testing data set, respectively. The RMSE on the training set is 1.65839. The correlation between the prediction and the real on the training set is 0.7857589 and on the testing set is 0.6744432. The results also show that exercise and sleep are personalized. The question is that whether model 3 can be transfer to other users.

5 CONCLUSION

In this paper, we present the first effort in the domain to predict the sleep efficiency based on users' exercise information using large datasets collected from real world. Three GBM models have been built and our results show that increasing the number of features will not improve the mode accuracy. The model performance will be greatly improved, if we can collect more sample data for each user. In the future, we will take consideration of other affecting factors besides exercise

to predict the sleep efficiency. We will perform classification based on exercise according to users' sleep feedback.

ACKNOWLEDGMENTS

We thank Polar Electro Oy for providing the data and platform for experimentation. The first author also thank to Finnish Cultural Foundation for supporting this research.

REFERENCES

- [1] N. Buscemi, B. Vandermeer, C. Friesen, L. Bialy, M. Tubman, M. Ospina, T. P. Klassen, and M. Witmans. 2005. Manifestations and management of chronic insomnia in adults: Summary. *AHRQ Evidence Report Summaries* (June 2005), 1–10.
- [2] Patrick Celka, Kaisu Martinmaki, Topi Korhonen, Nuutti Santaniemi, and Jussi Virkkala. 2017. Sleep-wake detection and computation of sleep continuity from a wrist unit in children, adolescents and adults. *IFMBE Proceedings book series EMBEC NBC 2017* 65 (June 2017), 1049–1052.
- [3] Brett A. Dolezal, Eric V. Neufeld, David M. Boland, Jennifer L. Martin, and Christopher B. Cooper. 2017. Interrelationship between sleep and exercise: A systematic review. *Advances in Preventive Medicine* (October 2017). <https://doi.org/10.1155/2017/1364387>
- [4] Jerome H. Friedman. 2000. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics* 29 (2000), 1189–1232.
- [5] Abbey Lunney, Nicole R. Cunningham, and Matthew S. Eastin. 2016. Wearable fitness technology: A structural investigation into acceptance and perceived fitness outcomes. *Computers in Human Behaviour* 65 (December 2016), 114–120. <https://doi.org/10.1016/j.chb.2016.08.007>
- [6] K. Martinmäki, T. Korhonen, K. Säynäjäkangas, N. Santaniemi, T. Posio, P. Puolakanaho, and E. Pekonen. 2018. Measuring and estimating sleep quality. *Justia Patents* (August 2018).
- [7] T. Myllymäki, H. Kyröläinen, K. Savolainen, L. Hokka, R. Jakonen, T. Juuti, K. Martinmäki, J. Kaartinen, M. L. Kinnunen, and H. Rusko. 2011. Effects of vigorous late-night exercise on sleep quality and cardiac autonomic activity. *J. Sleep Res.* (2011), 146–153. <https://doi.org/10.1111/j.1365-2869.2010.00874.x>
- [8] G. S. Passos, D. Poyares, M. G. Santana, C. V. D'Aurea, S. D. Youngstedt, S. Tufik, and M. T. de Mello. 2011. Effects of moderate aerobic exercise training on chronic primary insomnia. *Sleep Medicine* 12 (2011), 1018–1027. <https://doi.org/10.1016/j.sleep.2011.02.007>
- [9] K. Pesonen and L. Kuula. 2018. The Validity of a New Consumer-Targeted Wrist Device in Sleep Measurement: An Overnight Comparison Against Polysomnography in Children and Adolescents. *J. Clin Sleep Med.* 14 (2018), 585–591. <https://doi.org/10.5664/jcsm.7050>
- [10] Grace Shin, Mohammad Hossein Jarrahi, Yu Fei, Amir Karami, Nicci Gafinowitz, Ahjung Byun, and Xiaopeng Lu. 2019. Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review. *Journal of Biomedical Informatics* 93 (May 2019). <https://doi.org/10.1016/j.jbi.2019.103153>
- [11] K. Wunsch, N. Kasten, and R. Fuchs. 2017. The effect of physical activity on sleep quality, well-being, and affect in academic stress periods. *Nature and Science of Sleep* 9 (April 2017), 117–126. <https://doi.org/10.2147/NSS.S132078>