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 Riekki
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

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.
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-technical 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.

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-technical 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 feedback.

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

<table>
<thead>
<tr>
<th>user id</th>
<th>start date</th>
<th>stop date</th>
<th>start time</th>
<th>stop time</th>
<th>duration</th>
<th>calories</th>
<th>hr min</th>
<th>hr max</th>
<th>hr avg</th>
</tr>
</thead>
</table>

Table 2: A sleep sample with main features in dataset

<table>
<thead>
<tr>
<th>user id</th>
<th>starting time</th>
<th>waking time</th>
<th>sleep span</th>
<th>sleep continuity</th>
<th>sleep efficiency</th>
<th>sleep feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>000001</td>
<td>2018-8-26 23:30:11</td>
<td>2018-8-27 06:30:20</td>
<td>25209000</td>
<td>4.5</td>
<td>92</td>
<td>sleep well</td>
</tr>
</tbody>
</table>

Table 3: Main investigated exercise features or variables

<table>
<thead>
<tr>
<th>Features/Variables abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Number of exercise performed in one day</td>
</tr>
<tr>
<td>$TD$</td>
<td>Total duration in day’s all exercise</td>
</tr>
<tr>
<td>$TR$</td>
<td>Total trimp in day’s all exercise</td>
</tr>
<tr>
<td>$TC$</td>
<td>Total calories consumption in day’s all exercise</td>
</tr>
<tr>
<td>$T$</td>
<td>Trimp in day’s last exercise</td>
</tr>
<tr>
<td>$TD_{Zone_i}$ $i=1,...,5$</td>
<td>Heart rate duration in zone $i$ in day’s exercise</td>
</tr>
<tr>
<td>$C$</td>
<td>Calories consumption in day’s last exercise</td>
</tr>
<tr>
<td>$D$</td>
<td>Duration in day’s last exercise</td>
</tr>
<tr>
<td>$TOS$</td>
<td>Span time between last exercise stop time to sleep starting time</td>
</tr>
<tr>
<td>$D_{Zone_i}$ $i=1,...,5$</td>
<td>Heart rate duration in zone $i$ in day’s last exercise</td>
</tr>
</tbody>
</table>

Table 4: Model Performance

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE on testing set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>2.753662</td>
</tr>
<tr>
<td>Model 2</td>
<td>2.724675</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.908299</td>
</tr>
</tbody>
</table>

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.
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 to 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