Driving Fatigue Prediction Model considering Schedule and Circadian Rhythm

Qi Zhang, Chaozhong Wu, and Hui Zhang

1Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, China
2Engineering Research Center of Transportation Safety, Ministry of Education, Wuhan, China

Correspondence should be addressed to Hui Zhang; zhanghuiits@whut.edu.cn

Received 29 December 2019; Accepted 5 February 2020; Published 19 March 2020

Guest Editor: Zeyang Cheng

Driver fatigue level was considered an accumulated result contributed by circadian rhythms, hours of sleep before driving, driving duration, and break time during driving. This article presents an investigation into the regression model between driver fatigue level and the above four time-related variables. With the cooperation of one commercial transportation company, a Naturalistic Driving Study (NDS) was conducted, and NDS data from thirty-four middle-aged drivers were selected for analysis. With regard to the circadian rhythms, commercial drivers operated the vehicle and started driving at around 09:00, 14:00, and 21:00, respectively. Participants’ time of sleep before driving is also surveyed, and a range from 4 to 7 hours was selected. The commercial driving route was the same for all participants. After getting the fatigue level of all participants using the Karolinska Sleepiness Scale (KSS), the discrete KSS data were converted into consecutive value, and curve fitting methods were adopted for modeling. In addition, a linear regression model was proposed to represent the relationship between accumulated fatigue level and the four time-related variables. Finally, the prediction model was verified by the driving performance measurement: standard deviation of lateral position. The results demonstrated that fatigue prediction results are significantly relevant to driving performance. In conclusion, the fatigue prediction model proposed in this study could be implemented to predict the risk driving period and the maximum consecutive driving time once the driving schedule is determined, and the fatigue driving behavior could be avoided or alleviated by optimizing the driving and break schedule.

1. Introduction

Driving fatigue is a major safety issue in transportation, which has been identified to be associated with an increased risk of traffic collisions on roads because fatigued drivers tend to be unfocused with reaction time increases and impaired driving performance [1]. Deaths caused by road traffic accidents has risen to 1.35 million in 2018 [2]. Generally, accidents that are directly or indirectly caused by fatigue driving account for 30%–45% [3]. Operating vehicles when drivers are fatigued, they are endangering themselves and others. 31% of highway vehicle drivers admitted to driving while they were unable to keep their eyes open [4]. Approximately, each year 100,000 fatigue driving accidents reported in the United States cause 1,550 deaths and 71,000 injuries [5]. Based on nationwide statistics in Canada, about 20% of drivers nodded off or fallen asleep while driving and kill about 400 Canadians each year [6, 7]. In China, the traffic accident mortality rate caused by fatigue driving is twice that caused by other reasons [8]. With regard to the frequency of fatigue-caused accidents, it accounts for 10%–20% of all road traffic accidents in Europe [9]. To reduce the risk of traffic accidents and prevent fatigued driving [10], most countries have implemented their own hours of driving regulations [11–16]. However, those regulations differ from country to country and mainly include two aspects: the maximum duration on-duty and rest break during driving (Figure 1).

With the aim to detect or predict the fatigue status of drivers, more measurements including contextual, contact, or contactless physiological features; driver behavior; vehicle maneuver; and environment [17, 18], and more complex mathematical algorithms [17, 19, 20] were proposed.
Whether in simulation experiments or field study, devices such as eye trackers and electroencephalograph recorders were equipped to collect data that will be used in algorithms [21–23]. All these studies are processing new algorithms, based on probability or statistics models and specific indicators, driving fatigue detection and prediction models were established through data mining. These studies have demonstrated that driving fatigue is detectable and predictable. However, those algorithms take long terms in driving fatigue data processing and optimal parameters training; even their accuracy of fatigue detection reached 80–90%, not ready to be implemented in real-time fatigue detection and prediction, especially in the case of new samples. This makes such algorithms more suitable for detection rather than for prediction, especially in the field driving environment, and it is difficult to obtain accurate physiological data. And obtained algorithms mostly tend to deal with the subtle driver factors during the driving process, rather than analyzing the fatigue value change throughout a given schedule. So, it is necessary to develop a fatigue prediction model for a whole driving process.

Fatigue is an accumulated result contributed by several factors; generally, fatigue is considered a suboptimal psychophysiological condition caused by sleep, rest, circadian effects, and daily activities [24]; with regard to driving, daily activities specifically refer to driving. Progressive decrement of driving performances proved the negative influence of consecutive driving, which was confirmed closely related to driving fatigue. By analyzing the driving log, Jovais et al. found that drivers will be fatigued and the accident risk will rise after driving 4 hours without rest or break [25]. After analyzing 1,924 driving events, Lin et al. found the accident risk in the first 4 hours is low. But the risk would increase more than 50% in the following 3 hours, and increased 80–130% during the final hour [26]. Forty drivers were recruited in the field driving experiments, and Ma analyzed their driving records and found that those drivers showed more driving errors and poor driving performance after consecutive driving 3.5 hours [27]. In a 6-hour simulated driving experiment, Jing et al. recorded four drivers’ physiological information and found that 235 minutes should be the consecutive driving time threshold [28].

In addition, taking enough sleep and rest during consecutive schedules are the two main effective methods to relieve fatigue in the field driving environment and simulators. In some driving simulation experiments, the neurocognitive measures of vigilance, reaction time, and driving performance were evidenced to impair after sleep deprivation [29, 30]. Other findings from field experiments also proved that drivers showed more tendencies to be drowsy after experienced sleep deprivation [31, 32]. In a statistical analysis of the hours of sleep for drivers who were involved in a representative sample of crashes, Tefft found that the shorter the drivers slept in the 24 hours before crashing, the more odds the drivers should be culpable for their crashes [33]. Besides the sleep before driving, the rest breaks between driving stages can also help drivers recover from driving fatigue. In the regulation of hours of service (HoS) [34], a term related to rest breaks was firstly included. The rule categorically stated that "Driving is not permitted if more than 8 hours have passed since the end of the driver’s last off-duty or at least 30 minutes since sleeper berth period." Chen and Xie analyzed 183 crash events and 398 noncrash events and found that the more the rest break was taken, the more the crash odds could be reduced [12]. After 4 hours of consecutive field driving, Yuhua et al. found that it was difficult for electrocardiograph signal return to a normal level unless male and female drivers rested at least 24 and 27 minutes, respectively [35]. The circadian rhythms have been proved to have an impact on drivers’ fatigue; their alertness and performance vary across the day driven by the circadian rhythm [36]. The circadian rhythm also caused a higher proportion of sleep-related accidents to occur, mostly in the early morning and early afternoon, mainly ranging in two periods (02:00–05:00 and 13:00–16:00) [37–42] between the thick black lines in Figure 2.

It could be concluded that the four time-related indicators, time of sleep before driving, rest time, circadian rhythm, and consecutive driving time, all have an influence on driving fatigue. Those algorithms completed in the previous research used the drivers’ parameters or the relevant parameters of the vehicles being driven; not all of the four time-related factors were taken into account, so these algorithms can only perform short-term fatigue detection of prediction. The purpose of this article is to study a newly developed fatigue prediction model; with more than one feature, the model is expected that the precision could be improved and as concise as possible, only considering the four mentioned variables, and could model the fatigue for the whole driving process.

2. Experimental Design and Data Processing

By cooperating with a commercial transportation company, one regular transportation schedule was finally selected, G70 (Han-Shi) Expressway from Wuhan to Xiangyang, China, which is more than 6 h of round-trip travel. Considering the time span of the experiment was more than 6 hours, all of the participants need to meet the following requirements: no sleep disorder, no job on shift or night routines, in good health, not on medication, and no serious disease in nearly five years; more male drivers were selected during this experiment as shown in Figure 3.
Two weeks before the experiments, one recruitment notice for participant recruiting was issued in the cooperate company, and any driver who was actively licensed and met the appeal requirements can register. All of the participants were also required to abstain from drinking alcohol, tea, and coffee within 72 hours before the driving schedule, and record their time of sleep within 24 hours before the driving schedule. Finally, a total of 50 participants who were actively licensed and in normal health, especially free from any sleep disorders [43, 44], were recruited from the company; they all signed informed consent agreement prior to their participation and got 500 yuan for his/her contribution, which was almost one time higher than their usual daily salary.

During the schedule, participants could take a break in one service area [1, 8] when they feel extremely fatigued and consider continue driving will lead to high collision risk, and the driver’s face and the surrounding environment were recorded by several cameras (Figure 4(a)); a vision-based lane departure warning device Mobileye C2-270 was used to record the lane position data at a sampling rate of 8 Hz (Figure 4(b)), and all of the equipment were installed in one automatic transmission private vehicle (Figure 4(c)).

To obtain data with a large change in the fatigue level, driving schedules started around 09:00, 14:00, 21:00 were selected. However, 16 participants were excluded for various reasons: three drivers failed to complete the schedule, four of them drank alcohol within 24 to 72 hours before their schedules, and nine of them failed to fully cover neither of the two circadian rhythm peak periods. Finally, 34 drivers (middle-aged, mean = 47.8, SD = 5.1, and held the driving license for an average of 18.2 years with a standard deviation of 6.5 years) were selected, and the statistics of their schedules are shown in Figure 5. 24 participants whose schedules started around 09:00 were divided into the morning groups and further divided into three subgroups considering self-reported time of sleep within 24 hours before their schedules (time of sleep was recorded by one wearable device: smart bracelet) [45]. Every five participants whose schedules started around 14:00 or 21:00 were divided into the afternoon group and the night group.

The Karolinska Sleepiness Scale (KSS) [46] was used for evaluating subjective sleepiness because of its validity and reliability, and its scores range from 1 to 9, where 1 indicates extremely alert and 9 indicates extremely sleepy, even falling asleep; the higher the level is, the more fatigued the driver is. The self-reported KSS results of participants were recorded every five minutes by one experiment recorder and used as the subjective fatigue level measurement in this study. After the driver’s KSS data were obtained, all KSS data were converted into fatigue value. In this study, the fatigue value is used to substitute the tiredness scores and defined as a nonunit constant, just for data processing and modeling. The fatigue value is transformed from a 9-grade KSS [47] to a 150 unmeasured value considering the actual performance of drivers and the related results in sleep studies [48], the transform process of the conversion between the KSS level and the overall fatigue value and the fatigue value changing data of the four factors are shown in Figure 6.

By analyzing the recorded video of the 34 selected participants, the driver’s consecutive driving time, rest time, and the circadian rhythm range during the schedules could be obtained; the real-time fatigue value could be linearly summed to give overall tiredness scores [49], and the basic relationship between the four variables is shown in

\[ F(t) = F_{t_d} + F_{t_s} - F_{t_c} + F_{t_r}, \]

where \( F(t) \) is the total fatigue value in this study, \( F_{t_d} \) is the fatigue value caused by consecutive driving, \( F_{t_s} \) is the fatigue value after sleep, \( F_{t_c} \) is the fatigue value relieved by rest, and \( F_{t_r} \) is the fatigue value caused by circadian rhythms.

At first, the KSS levels were transferred to the fatigue value and expanded to 1 Hz by cubic spline interpolation, as shown in Figure 6 (fatigue value). To facilitate the subsequent model establishment, all time-related parameters were counted in seconds. Limited by the experimental conditions, circadian rhythm data were directly cited from exiting chronobiology studies [24, 36]. At the beginning of the experiment, the fatigue value was assumed only caused by circadian rhythm and time of sleep; after removing the fatigue value of the circadian rhythm (\( F_{t_c} \)) from the beginning, the remaining fatigue values were considered to be only caused by different times of sleep (\( F_{t_s} \)). After removing the fatigue value of circadian rhythm and time of sleep from the total fatigue value, the remaining fatigue values were considered to be caused by driving and rest (\( F_{t_d} + F_{t_r} \)). Because all of the experiments were carried out in a similar environment, the influence of various factors on driver fatigue, such as traffic flow, weather, and light, was ignored.
Driving performance is another valuable measurement to confirm the drivers’ fatigue status; in this study, the standard deviation of lateral position (SDLP) is used as a verification indicator parameter [47, 50, 51]; it reflects the driver’s ability to avoid unintentional lane departure or lane crossing. The lane line position was collected by Mobileye at a frequency of 10 Hz, the average of SDLP data in one second was processed into new data after filtering and used as a verification indicator, and the SDLP is computed by

\[
SDLP = \sqrt{\frac{\sum_{i=1}^{n} (d_i - d_{avg})^2}{n}},
\]

where \(d_i\) denotes the \(i\)th lane position for this horizontal curve segment. The lane position refers to the distance from the center of the vehicle to the right edge of each lane. \(d_{avg}\) denotes the average lane position and \(n\) denotes the sample size of the lane position.

3. Modeling and Verification

Five participants (three were from the morning group and two were from the other two groups) were randomly selected and used as validation, and their data were not used in the modeling process. In this study, the model validation is performed using the root mean squared error (RMSE) and the coefficient of determination (R-square). Although some of the circadian rhythm data are missing, it does not affect its modeling, which has already been mentioned above, and through the above data processing process, the fatigue value data of the four indicators were obtained and their prediction model could be established.

3.1. Modeling of Fatigue Value Contributed by Circadian Rhythm. According to the previous research work, it could be seen that the fatigue value caused by circadian rhythm shows a regular change with the time of the day [36, 49, 51, 52]; in this study, it is assumed to be the same among all participants as an underlying fixed variable. Firstly, the missing circadian rhythm data are interpolated by cubic spline interpolation, and then the effects of circadian rhythm shown in Figure 7 were best fitted with a sum of sin functions shown as

\[
F_{tc} = \sum_{i=1}^{8} a_i \sin(b_i t_c + c_i),
\]
where $t_c$ is the time of day, from 0 to 86400 s; $a_i, b_i, c_i$ ($i = 1, 2, 3, \ldots, 8$) are numerical constants and given at the end of this article; and $F_{tc}$ is the fatigue value caused by circadian rhythm. The SSE, $R$-square, and RMSE are 146, 0.9991, and 0.7437, respectively.

3.2. Modeling the Fatigue Value Caused by Hours of Sleep. At the beginning of the experiments, the initial fatigue value of all participants with different times of sleep could be obtained. The fatigue values of different sleeping times could be best fitted with the model as in equation (4), and shown in Figure 8.

$$F_{ts} = 954.9 e^{-0.721 t_s},$$

where $t_s$ is the time of sleep in hours within 24 hours before driving, $F_{ts}$ is the fatigue value with different times of sleep. The validation parameters are SSE: 180.8, $R$-square: 0.911, and RMSE: 2.497.

3.3. Modeling of the Fatigue Value Contributed by Consecutive Rest. In this study, each participant takes rest in the service area. After collecting the fatigue value data before and after the rest, the fatigue value relieving effects of different rest times could be estimated, which could be fitted with the model as in equation (5), and shown in Figure 9.

$$F_{tr} = \frac{1.216 t_r^2 + 1238 t_r + 2.417 \times 10^5}{t_r + 1.6 \times 10^6},$$

where $t_r$ is the time on rest in seconds, $F_{tr}$ is the fatigue mitigation value, and the three validation parameters are SSE: 681.0, $R$-square: 0.8592, RMSE: 4.497.

3.4. Modeling of the Fatigue Value Caused by Consecutive Driving. This study consists of three sets of experiments: morning group, afternoon group, and night group. In driving stages, different driving times and corresponding changes in the fatigue value are collected and presented in various colors in Figure 10. The analysis of its changing tendency can be obtained as follows:

$$F_{td} = 104.4 \sin\left(4.539 \times 10^{-3} t_d - 0.01652\right) + 2.922 \sin\left(5.053 \times 10^{-4} - 3.135\right),$$

where $t_d$ is the time of day, from 0 to 86400 s; $a_d, b_d, c_d$ ($i = 1, 2, 3, \ldots, 8$) are numerical constants and given at the end of this article; and $F_{td}$ is the fatigue value caused by circadian rhythm.

![Figure 7: Fatigue value changes caused by the circadian rhythm.](image)

![Figure 8: Fatigue value changes caused by time of sleep.](image)

![Figure 9: Fatigue value changes caused by rest.](image)

![Figure 10: Fatigue value changes after consecutive driving.](image)
Figure 11: Time variable model description.

Figure 12: Comparison of the SDLP with the fatigue value of the three participants. (a) Subject with 7.4 hours of sleep. (b) Subject with 5.5 hours of sleep. (c) Subject with 3.8 hours of sleep. (d) Subject with 7.7 hours of sleep. (e) Subject with 7.8 hours of sleep.
3.5. Verification of the Fatigue Value Prediction Model Using SDLP. When drivers start driving, their time of sleep within 24 hours before driving could be collected, the time point when the driver starts driving as well as the whole circadian rhythm range in real time could be identified, and the two fatigue value parts are the basic fatigue value. Additionally, the consecutive driving time and rest time during driving could also be collected in real time. The driving fatigue value at any time point could be calculated using the collected data and established models, as shown in Figure 11.

SDLP data of the five selected participants could also be calculated; after extracting the corresponding time period of the five subjects’ schedules, the fatigue value of circadian rhythm during the schedule, and the fatigue value of consecutive driving, rest, and different time of sleep were estimated. The relationship between the two sets of data can be calculated by Spearman’s rho using

\[
\rho = \frac{\text{cov}(\sigma_{\text{FI}}, \sigma_{\text{sdlp}})}{\sigma_{\text{FI}} \sigma_{\text{sdlp}}},
\]

where \( \rho \) denotes the usual Pearson correlation coefficient but applied to the two ranked sets of data, \( \text{cov}(\sigma_{\text{FI}}, \sigma_{\text{sdlp}}) \) is the covariance of the two ranked sets of data, and \( \sigma_{\text{FI}}, \sigma_{\text{sdlp}} \) are the standard deviations of the two ranked sets of data.

The Spearman correlation coefficients of the 5 participants are (a): 0.9766, (b): 0.9549, (c): 0.9804, (d): 0.9661, and (e): 0.9035, respectively. Such values indicate that there is a strong correlation between the calculated fatigue value with the model and the SDLP, and the trends of the two sets are also consistent. Because most of the scheduled routes are highways, there were few lanes changing and interacting with other road users, so the change in SDLP could be seen as mainly caused by driver fatigue. For the participants from the morning group (a)–(c), in Figure 12, the trend of SDLP is the most moderate for the one with 7.4 sleep hours. The trends of SDLP of the two participants with 5.5 and 3.8 sleep hours are similar, especially during the period before having a break. In this study, 0.3 was considered as the SDLP threshold [23, 53], and the three participants exceeded the threshold during the later driving stage in their schedules. When 100 was seen as the participants’ fatigue value threshold [49, 51, 52], only the SDLP of the participants with 7.4 hours of sleep is less than the threshold. In addition, the shorter the participants’ sleep, the more the SDLP exceeds the threshold. For the afternoon group and the night group, the SDLP trends of the two participants are different from each other and the morning group. Compared with other participants, the participants in the afternoon group own the lowest fatigue value, but the SDLP exceeds the threshold at the earliest. However, there is also a turning point at around 8000 s for the two participants for the afternoon group and the night group. When the SDLP of the night participant approaches 0.3 for the first time, its SDLP fluctuates around 0.3. After taking rest in the service area, the SDLP increased significantly.

3.6. Maximum Predicted Consecutive Driving Time. After the model is established and verified, the maximum consecutive driving time of drivers who start driving at different common times could be obtained, as shown in Figure 13. Assuming that the driver starts driving at 09:00, 14:00, and 21:00, respectively, the driving duration before the predicted fatigue value of drivers exceeds the threshold could be calculated. In this study, the threshold was set to 100 by reviewing the physiology-related literature. Because the peak of the circadian rhythm is between 13:00 and 16:00, the fatigue value of the driver who starts driving at 14:00 shows an arched shape. Within the first 2.5 hours of consecutive driving, driving fatigue rises slowly, and then turns to fall. During the whole driving process, driving fatigue reaches the threshold at 2.05 hours for the first time. At the same time, its initial fatigue value is also the highest among the three groups. When the driver starts driving around 09:00, under the influence of low circadian rhythm fatigue value, the initial fatigue value is the lowest among the three kinds; the area between the two dotted lines is the fatigue value changing the range for drivers who have different times of sleep. Before reaching the fatigue threshold, it is almost always the lowest. Finally, the driver will exceed the threshold after 2.57 to 3.05 hours of consecutive driving. Assuming the driver starts driving around 21:00, throughout the whole driving process, its tendency is very familiar to the morning group and their slope is almost the same before the fatigue value threshold, but the fatigue rises more quickly and reaches the threshold after 2.25 hours of consecutive driving.

4. Conclusions and Discussion

Real-time detection of driver fatigue is technically mature, and most of the vehicles have been equipped with different
kinds of relevant equipment to detect whether or not the driver is fatigued. But predicting the driver fatigue degree remains an important research topic. The purpose of this study is intended to predict driver fatigue only using time-related variables, such as consecutive driving time, consecutive rest, time of sleep within 24 hours before driving, and circadian rhythms during driving. Field experiments were conducted to obtain the fatigue-level indicators (KSS) and the driving performance (SDLP). The model in this study was developed from field experiments using an equipped car, which might be more believable than using a driving simulator. The strong correlation between the predicted fatigue value and SDLP collected from the Mobileye demonstrated that the developed method in this study is reliable. A major finding of this study is the establishment of models contributed by circadian rhythm, time of sleep, time of rest, and the detailed modeling of driving duration starting at different times. The analysis results indicated that when drivers reached a higher fatigue level, they would swing around when driving. And driving fatigue affects lane-keeping performance significantly. At the same time, even when the driver is at a low fatigue level, their driving performance will still reduce after driving for a consecutive time. In a word, the results demonstrate that it is possible to predict driver’s fatigue degree at a certain moment using the developed driving fatigue prediction model which only considers the four time-related variables. Instead of using KSS to verify the accuracy of the model, the driving performance indicator SDLP was used because the operational indicator standard is more uniform. Even though the KSS has been widely used, for different observers (the observer in this study did not change) and participants, their understanding of fatigue is difficult to maintain in a consistent standard. In our best case, the correlation between the calculated fatigue value and computed SDLP is up to 98%, showing that the model in this study could reflect the drivers’ fatigue change trend perfect. The maximum consecutive driving time calculated in this article using the model is really shorter than any existing regulation; even the longest driving time calculated in this article is 20% shorter than any existing regulation; even the longest driving time calculated in this article is 20% shorter. It is hard to say whether this is a problem with the model established in the study or the setting of the fatigue value threshold or some urgent improvements in the existing regulations; this will be further studied in future research.

The purpose of this study is to use the time-related variables to predict the time when the driver will be fatigued. In the data analysis process, the authors try to avoid using behavioral, physiological, and vehicle data for modeling. The other variables, including physiological variables (e.g., respiration), behavioral variables (e.g., head position), or vehicle information (e.g., speed), always need to face the problem of statistical frequency when dealing with them. Regardless of the size of the time window used for data processing, a time window is always required when getting an indicator. This results in the inability to perform real-time fatigue prediction. However, this study has several limitations. First, KSS and SDLP used in this study may limit the accuracy of the results. As for driving performance indicators, because of the limited funds and lack of equipment, only the SDLP was analyzed. In addition, the collection frequency of KSS is a little high (one point every 5 minutes), and this may help drivers keep awake [54]. What is more, in these field experiments, participants need driving on the highway for at least 4 hours, and it may become difficult for drivers to judge their drowsiness [55] after a period of driving. Second, age and gender are the two main important indicators that may affect influential driving performance, but they are not taken into account in this study. Some experimental conditions factors, such as weather and traffic flow, are not considered either. Even though the authors tried their best to ensure that the experiments are conducted in the same conditions, there is still no guarantee that all conditions are the same. Moreover, when drivers restart driving after rest in the service area, the model of driving time needs to be deeply explored. When drivers start driving again, their starting time is close to the starting time of the next set of experiment group. Moreover, for the experiments starting in the afternoon and night, the fatigue accumulated in the daytime before the experiments are only estimated using the hours of sleep, which may be inaccurate, and the fatigue of these drivers should be divided more meticulously in the future work.

Data Availability

The data used in this study are available from the corresponding author on request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61603282, U1764262, 51775396) and the Major Project of Technological Innovation of Hubei Province (2017CFA008).

Supplementary Materials

As shown in Figure 7, the black line is the best fit curve of the fatigue value of circadian rhythm and the curve could described by equation (3), where \( F(t) \) is the fatigue value caused by circadian rhythm; \( t_c \) is the time of day, 0 to 86400 seconds from 00:00 every day; and \( a_i, b_i, c_i \) \((i = 1, 2, 3, \ldots, 8)\) are numerical constants, and all of the parameters are shown in the below supplementary table. (Supplementary Materials)

References


