Research on the correlation of dangerous driving behaviors based on naturalistic driving experiment

Xingcan Liang

1 Beijing Jiaotong University, Beijing, 100000, China

* e-mail: 18120818@bjtu.edu.cn

Abstract: In China, although road traffic safety has improved in recent years, the number of deaths caused by traffic accidents is still at a high level every year. Accidents due to drivers’ dangerous driving behavior account for a large proportion. This study designed and implemented a highway naturalistic driving experiment. The driving performance data and the observer record data obtained in the experiment were used to extract four dangerous driving behaviors: rapid acceleration, rapid deceleration, sudden steering and drowsy driving. The correlation between the three kinds of urgent behaviors and driving fatigue was analyzed qualitatively and quantitatively. Finally, the correlation model between the frequency of urgent behaviors and drowsy driving was built using Naive Bayes. The feasibility of using this method to identify drowsy driving behavior was discussed.

1. Introduction

With the rapid development of the economy, the number of motor vehicles has soared, but the frequency of road traffic accidents remains high. Statistics show that traffic accidents caused by human factors account for 80% to 90%. NHTSA pointed out that the most dangerous driving behaviors are drunk driving, drug-impaired driving, distracted driving, speeding, and drowsy driving. Although drowsy driving is not the leading cause of traffic accidents, the number of casualties in a single incident has been at a high level. Due to the seriousness of drowsy driving, many scholars have conducted a lot of research on it, most of which focus on the identification of drowsy driving and technology-based interventions. At the same time, rapid acceleration, rapid deceleration, and sudden steering are easily ignored by dangerous driving behavior researchers because they are ubiquitous in daily driving behavior and are not illegal in road traffic regulations. Relevant researches were mainly on the recognition of these dangerous driving behaviors. Drowsy driving may cause the handling of the vehicle to be different from the normal state, which may lead to changes in the frequency of the rapid acceleration, rapid deceleration, and sudden steering. However, there was very little research on the correlation between drowsy driving and these kinds of urgent behaviors.

This paper analyses the relationship between the three kinds of urgent behaviors and drowsy driving from the qualitative and quantitative perspectives and explores the feasibility of identifying the drowsy driving behavior based on the frequency of the urgent behaviors.

2. Naturalistic driving experiment and data acquisition

2.1. Experimental design

The driving behavior data in this paper were obtained in the naturalistic driving experiment. The experiment recruited ten drivers (nine males and one female) to perform driving tasks, one observer to
record driver fatigue. Because of the monotonous driving environment, drowsy driving often occurs on highways. Therefore, the route of the experiment was set on the Wuhan-Xiayang section on the Han-Shi highway (Figure 1).

According to the driving performance indexes that need to be collected in the experiment, non-invasive data acquisition devices were installed on the vehicle, including HD cameras for the road environment (Figure 2) and smartphones placed above the dashboard (Figure 3). The goal is to collect the speed of the vehicle, the acceleration of each axis, the steering wheel angle, and record driver fatigue based on the Karolinska Sleepiness Scale (Table 1) during the experiment.

### Table 1. Fatigue level description corresponding to KSS scale value

| KSS scale value | Fatigue level                                      |
|-----------------|---------------------------------------------------|
| 1               | Extremely awake                                   |
| 2               | Very awake                                        |
| 3               | Awake                                             |
| 4               | A little awake                                     |
| 5               | Between awake and sleepy                           |
| 6               | Sleepy, this state of sleep is easy to control    |
| 7               | Very sleepy, this state of sleep has some control difficulty |
| 8               | Very sleepy, hard to control                      |
| 9               | Extremely sleepy, almost falling asleep           |

2.2. Data collection

This experiment aims to extract rapid acceleration, rapid deceleration, sudden steering data while driving and record driver fatigue. The index data collected in the experiment are shown in Table 2. After pre-processing the data, ten drivers’ complete driving behavior data were obtained.

3. Urgent behaviors extraction method

Urgent behaviors include rapid acceleration, rapid deceleration, and sudden steering. The most widely used method is to set the threshold and duration. If the corresponding driving behavior index values meet the given conditions, that kind of behavior is determined to have occurred. According to the existing algorithms and relevant industry regulations, the recognition conditions and duration set for the extraction of the urgent behaviors are shown in Table 3. The algorithm steps to extract rapid acceleration and rapid deceleration is shown in Figure 4; sudden steering is shown in Figure 5.
Table 2. Index data obtained in the experiment

| Index name           | Acquisition equipment | Sampling rate | Unit       | Data description                                                                  |
|----------------------|-----------------------|---------------|------------|-----------------------------------------------------------------------------------|
| Speed                | Smartphone            | ≥15Hz         | km/h       | Vehicle speed, non-direction                                                      |
| Acceleration_X       | Smartphone            | ≥15Hz         | G          | Acceleration in the direction of travel. A positive value indicates acceleration, and a negative value indicates deceleration |
| Acceleration_Y       | Smartphone            | ≥15Hz         | G          | Lateral acceleration, positive and negative means the steering direction is opposite |
| Steering wheel angle | Steering wheel angle sensor | >20Hz        | degree     | A positive value indicates turning right, a negative value indicates turning left   |
| KSS value            | Observer              | 5min          | -          | Driver fatigue level, according to the Table 1                                    |

Table 3. Urgent behavior recognition conditions

| Urgent behavior      | Conditions                        | Duration |
|----------------------|-----------------------------------|----------|
| Rapid acceleration   | Acceleration_X≥3m/s²              | T≥2s     |
| Rapid deceleration   | Acceleration_X≤-3m/s²             | T≥2s     |
| Sudden steering      | Acceleration_Y≥0.5g, Speed≥80 km/h| T≥2s     |

Figure 4. Rapid acceleration and rapid deceleration extraction algorithm

4. Correlation analysis of urgent behaviors and drowsy driving

4.1. Qualitative analysis

The rapid acceleration and rapid deceleration behavior are caused by the driver’s poor control of the vehicle speed. When in good spirits, drivers can maintain vehicle speed better and predict the changes in the surrounding environment in advance. Therefore, the behavior of stepping urgently on the accelerator is not natural to occur. When drivers start to feel sleepy, the driver’s ability to perceive the surrounding environment will reduce, especially on the highway. And the driver is likely to loosen the accelerator due to laziness, which causes large acceleration fluctuations (Figure 6). Therefore, we assumed that the higher the degree of driver fatigue, the higher the frequency of rapid acceleration and rapid deceleration.
Figure 5. Sudden steering extraction algorithm

Figure 6. Comparison of acceleration value between awake state and fatigue state. (a) is the acceleration image when the driver is awake (KSS value = 1), and (b) is the acceleration image when the driver is fatigued (KSS value = 6).

The sudden steering behavior is caused by the driver’s poor control of the vehicle’s travel direction. On straight sections, when the driver is in good mental condition, he often adjusts the steering wheel appropriately. The steering wheel angle fluctuation is small. When the driver is fatigued, his ability to maintain the steering wheel angle is reduced. And it is easy to appear the problem that vehicle direction cannot be corrected in time, showing a significant fluctuation in the steering wheel angle value (Figure 7). Therefore, from a qualitative perspective, the degree of fatigue should be positively related to the frequency of sudden steering behavior. The higher the driver’s fatigue level, the greater the possibility of sudden steering behavior.

The frequency of rapid acceleration, rapid deceleration, and sudden steering extracted in the previous section are compared with the KSS value (Figure 8). The orange polyline indicates the KSS value of the driver, and the blue polyline indicates the smoothed frequency of the corresponding urgent behavior. It is found that the increase and decrease trends of the two are basically consistent.
4.2. Quantitative analysis of Pearson correlation coefficient

In order to analyze the correlation between the frequency of urgent behaviors and the KSS fatigue value, the correlation coefficients of all drivers are listed in Table 4. The data in the table show that there is a positive correlation between the frequency of urgent behaviors and fatigue value, and the association of rapid acceleration and rapid deceleration is higher than that of sudden steering.

| Driver number | Pearson correlation coefficient | Driver number | Pearson correlation coefficient |
|---------------|-------------------------------|---------------|-------------------------------|
|               | RA | RD| SS | RA | RD| SS | RA | RD| SS |
| 1             | 0.608 | 0.564 | 0.473 | 6             | 0.490 | 0.467 | 0.387 |
| 2             | 0.459 | 0.580 | 0.326 | 7             | 0.666 | 0.662 | 0.318 |
| 3             | 0.509 | 0.456 | 0.323 | 8             | 0.608 | 0.564 | 0.324 |
| 4             | 0.469 | 0.454 | 0.200 | 9             | 0.547 | 0.518 | 0.237 |
| 5             | 0.583 | 0.600 | 0.436 | 10            | 0.724 | 0.673 | 0.512 |

5. Drowsy driving recognition model based on urgent behaviors frequency

Given the urgent behaviors frequency discussed in the previous section has a positive correlation with drowsy driving, this section uses a Naive Bayes model to establish a drowsy driving identification model based on the frequency of urgent behaviors.

5.1. Naive Bayes model
Naive Bayes is a classification algorithm, which uses known prior probabilities, modifies the prior probabilities through samples, obtains the posterior probabilities, and finally uses the posterior probabilities for classification. Suppose that each instance \(x\) can be described by the conjunction of the attribute values, and the class label \(c\) takes a value from a finite set \(C\). The function of Naive Bayes is to give the most likely class token \(c(x)\), after given the attribute values \(<a_1, a_2, ..., a_m>\) describing the instance. The calculation formula for class \(c(x)\) is:

\[
c(x) = \text{argmax}\{P(a_1, a_2, ..., a_m|c)P(c)\}
\]

Naive Bayes assumes that given the instance class tag, the instance attribute values are independent of each other. That is, given the instance class label, the observed joint probability is the product of the probability of each attribute value, which is:

\[
P(a_1, a_2, ..., a_m|c) = \prod_{i=1}^{m} P(a_i|c)
\]

Substituting it into the equation, we can get the classification formula of Naive Bayes classifier:

\[
c(x) = \text{argmax}\{P(c)\prod_{i=1}^{m} P(a_i|c)\}
\]

In the formula, \(a_i\) is the \(i\)-th attribute value of \(x\); the probabilities \(P(c)\) and \(P(a_i|c)\) can be easily estimated by calculating the frequency of occurrence of different types of markers and attribute value combinations in the training instance set, the specific formulas are as follows:

\[
P(c) = \frac{\sum_{j=1}^{n} \delta(c_j, c)}{n} \quad P(a_i|c) = \frac{\sum_{j=1}^{n} \delta(a_{ji}, a_i) \delta(c_j, c)}{\sum_{j=1}^{n} \delta(c_j, c)}
\]

Where \(n\) is the number of training instances; \(c_j\) is the class label of the \(j\)-th training instance; \(a_{ji}\) is the \(i\)-th attribute value of the \(j\)-th training instance; \(\delta(c_j, c)\) is a binary function, its value is 1, if \(c_j = c\), otherwise is 0.

5.2. Construction of identification model based on Naive Bayes

The structure of the Naive Bayes classifier mainly includes class nodes and attribute nodes, which are represented by \(C\) and \(A\), respectively. The structure of the model constructed in this paper is shown in Figure 9. The class node \(C\) in the figure represents the degree of fatigue. The other three nodes \(A_1\), \(A_2\), and \(A_3\) represent the three attributes of rapid acceleration frequency, rapid deceleration frequency, and sudden steering frequency, respectively.

![Figure 9. Naive Bayes model structure](image)

For the input of the model, this paper set three level for the frequency of urgent behaviors. The specific settings are shown in Table 5. Therefore, when inputting the attribute value of the training instance set, only input the level value corresponding to the urgent behavior frequency. For the class label of the training instance set, set KSS fatigue values 1 to 4 as awake state (\(C = 1\)), fatigue value 5 as the critical state (\(C = 2\)), and 6 and above as fatigue state (\(C = 3\)).

| Urgent behavior | Level 1 | Level 2 | Level 3 |
|-----------------|---------|---------|---------|
| Rapid acceleration | \(n \leq 3\) | \(3 < n \leq 5\) | \(n > 5\) |
| Rapid deceleration | \(n \leq 3\) | \(3 < n \leq 5\) | \(n > 5\) |
| Sudden steering | \(n \leq 2\) | \(2 < n \leq 4\) | \(n > 4\) |
5.3. Verification of identification model

From the data of ten drivers collected in the naturalistic driving experiment, the data of the first nine drivers are selected as training examples, and the data of the tenth driver are used as test example. This article used the Naive Bayes class in MATLAB software for model verification.

The classification accuracy rate obtained by the test is 72.63%. After comparison, the model’s recognition accuracy rate of awake state reaches 79.17%, the recognition accuracy rate of the critical state is 63.81%, and the recognition accuracy rate of fatigue state is 66.92%. It can be seen that the model has the highest recognition rate for the driver’s awake state, followed by the fatigue state, and the recognition rate for the critical state is relatively low. Most of the misjudgments of the critical state judge the critical state as being fatigued, and under the principle of taking precautions, which will help to prevent drowsy driving. In summary, the identification model established in this paper has a particular ability to identify drowsy driving, which can provide a new idea for drowsy driving state detection technology.

6. Conclusion

This paper designed and implemented a naturalistic driving experiment to collect driving behavior data during vehicle driving. Next, the threshold method was used to extract the rapid acceleration, rapid deceleration and sudden steering behavior during the experiment. Then, we compared the correlation between the frequency of the three kinds of urgent behaviors and driver fatigue value, from the qualitative and quantitative perspectives. Result shows there are significant positive correlation between them. Finally, a drowsy driving identification model based on the frequency of urgent behaviors was established, and the recognition accuracy rate was 72.63%, indicating that it is feasible to use this method to identify drowsy driving.

References

[1] Wang M S, Jeong N T, Kim K S, et al. (2016) Drowsy behavior detection based on driving information. J. International Journal of Automotive Technology., 17(1):165-173.

[2] Songkroh A, Fooprateepsiri R, Lilakiataskun W. (2014) An intelligent risk detection from driving behavior based on BPNN and Fuzzy Logic combination. In: International Conference on Computer and Information Science. IEEE Computer Society. 105-110.

[3] Ren H J, Xu T, Li X. (2014) Driving Behavior Analysis Based on Trajectory Data Collected with Vehicle-mounted GPS Receivers. J. Geomatics and Information Science of Wuhan University., 39(06):739-744.

[4] Liu J, Li W, Zhang Y, Yang Y. (2016) Method and device for recognizing sudden turn of vehicle. P., China, CN105799709A, 2016.07.27.