Journal of Southeast University (English Edition) Vol. 35, No. 4, pp. 502 − 508 Dec. 2019 ISSN 1003—7985

# Evaluation of driving behavior based on massive vehicle trajectory data

Sun Chao[1,2] Chen Xiaohong[1] Zhang H. Michael[1,3] Zhang Junfeng[2]

([1] Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai 201804, China)
([2] Shenzhen Urban Transport Planning Center Co., Ltd., Shenzhen 518021, China)
([3] College of Civil and Environmental Engineering, University of California Davis, Davis 95616, USA)

**Abstract:** Based on the driver surveillance video data and controller area network (CAN) data, the methods of studying commercial vehicles' driving behavior is relatively advanced. However, these methods have difficulty in covering private vehicles. Naturalistic driving studies have disadvantages of small sample size and high cost, one new driving behavior evaluation method using massive vehicle trajectory data is put forward. An automatic encoding machine is used to reduce the noise of raw data, and then driving dynamics and self-organizing mapping (SOM) classification are used to give thresholds or the judgement method of overspeed, rapid speed change, rapid turning and rapid lane changing. The proportion of different driving behaviors and typical dangerous driving behaviors is calculated, then the temporal and spatial distribution of drivers' driving behavior and the driving behavior characteristics on typical roads are analyzed. Driving behaviors on accident-prone road sections and normal road sections are compared. Results show that in Shenzhen, frequent lane changing and overspeed are the most common unsafe driving behaviors; 16.1% drivers have relatively aggressive driving behavior; the proportion of dangerous driving behavior is higher outside the original economic special zone; dangerous driving behavior is highly correlated with traffic accident frequency.

**Key words:** driving behavior; global positioning system (GPS) navigating data; automatic coding machine; self-organizing mapping (SOM)

**DOI:** 10.3969/j.issn.1003 − 7985.2019.04.013

According to statistics, more than 90% of traffic accidents are related to the drivers[1]. Driving behavior analysis is of great significance in improving the issue of road traffic safety, carrying out active safety management of vehicles and implementing energy conservation and emission reduction. Previous driving behavior evaluation methods mainly focused on operating vehicles (including operating buses and trucks). The research methods are relatively advanced, including the analysis of driver monitoring data, controller area network (CAN) data and traffic flow information for driving behavior analysis and safety control. Driving behavior research for non-operating vehicles mainly relies on the naturalistic driving test, but naturalistic driving test has shortcomings such as limited sample coverage, high requirements for equipment and so it is difficult to promote this method widely. Vehicle satellite positioning research has the advantages of being low cost, convenient operation and is not time-consuming. With the popularity of smart phones, the massive positioning data generated by navigation software can be of great use in driving behavior research, and the results can be more representative and accurate, therefore, making the study of citywide driving behavior study possible. This paper analyzes the desensitized GPS navigation data provided by a map software programme for driving behavior studies.

In the field of operating vehicles' driving behavior studies, little research has been done. Fatigue driving and the following experiment of 40 professional drivers proved that the degree of fatigue has a significant impact on headway distance[2]. Thiffault et al.[3] proved that drivers' fatigue degrees vary significantly in a monotonous environment compared with a complex environment. Guo et al.[4] used an analytic hierarchy process to obtain the variation law of CAN data corresponding to drivers' fatigue driving process. Analysis of operating vehicle traffic accidents proved that lack of safety awareness, risk perception and driving skill defects are the main causes of traffic accidents for operating vehicle drivers[5]. Based on CAN and video surveillance data, a relatively complete research method system is established.

In the field of non-operating vehicles' driving behavior studies, Zhang et al.[6] used naturalistic driving data to build a lane-changing decision-making model, and established lane utility function in off-ramp areas, respectively, in free flow, steady flow and congested flow. They also generalized lane-changing characteristic under traffic flow state evolution. By analyzing naturalistic driving da-

**Received** 2019-06-16, **Revised** 2019-10-19.

**Biographies:** Sun Chao (1985—), male, Ph. D. candidate, senior engineer; Chen Xiaohong (corresponding author), female, doctor, professor, Tongjicxh@163.com.

**Foundation items:** The National Natural Science Foundation of China (No. 71641005), the National Key Research and Development Program of China(No. 2018YFB1601105).

**Citation:** Sun Chao, Chen Xiaohong, Zhang H. Michael, et al. Evaluation of driving behavior based on massive vehicle trajectory data[J]. Journal of Southeast University (English Edition), 2019, 35(4): 502 − 508. DOI: 10.3969/j.issn.1003 − 7985.2019.04.013.

ta, Wang et al. [7] comfirmed that: 1) A vehicle forward collision avoidance warning system can lower car following interval time significantly; 2) Multi-lane changing proportion of high-grade roads is higher[8]; 3) Chinese domestic drivers tended to be more aggressive than foreign drivers while changing lanes[9]. Olsen et al. [10] used American naturalistic driving data while conducting the study of drivers' lane changing behavior. Bone et al. [11] studied the traits of aggressive and distracted drivers. Paaver et al. [12] proved that with proper education and intervention, new drivers' risk driving behavior can drop significantly. Ahmed et al. [13-14] established the highway risk evaluation model with parameters such as traffic volume, speed, occupancy, weather and road geometric shape. Li et al. [15] proved that the visual environment of a curved section in a mountain has a significant impact on drivers' perception-response time. Non-operating vehicle test data has high precision, but the data acquisition cost is too high, and test results are not representative enough.

In GPS data research field, Grengs et al. [16] conducted an in-depth analysis for four-week GPS positioning data from 78 users, and put forward the research methods of individual travel characteristics based on GPS and GIS. Zhao et al. [17] carried out behavior research on elderly driving behavior using GPS data. Liao et al. [18] carried out driving trajectories modelling based on GPS data. Nowadays, GPS positioning data is mainly used for highway driving behavior recognition and dangerous driving behavior detection based on operating vehicle supervision platform data in China. However, the data sample of the current research is relatively small[18] or mainly focuses on operating vehicle drivers[19]. Relative research exploited the use of GPS navigation data, yet further use such as the study of general motor vehicle drivers and application in traffic safety can be done in-depth. This paper analyzes the driving behavior characteristics by high-precision mass vehicle GPS data to reveal drivers' driving behavior features and common unsafe behavior, so that we can take more directive measures to improve traffic safety.

## 1 Data Introduction and Research Methods

### 1.1 Raw data and data screening

Raw data contains the vehicle-navigation data of a one-month period in Shenzhen, as describted in Tab. 1.

**Tab. 1** Raw data sample

| Date | Time | ID | Status | Longitude/(°) | Latitude/(°) | Speed/(km · h$^{-1}$) | Direction/(°) |
|---|---|---|---|---|---|---|---|
| 2017-05-09 | 20:54:44 | 17435251342634322685 | 11 | 114.394377 | 22.747687 | 40 | 242 |

Note: Status 11 means navigating software is on; 10 means navigating software is off; Direction is the heading direction angle of vehicle towards north.

The total volume of data is about 500 GB. Original data contains lots of invalid data. Positioning offset is a common problem and some users' data lasts a quite short period of time, which is not worthy of studying. To tackle these problems, the record data corresponding to 10 users with the highest number of record in units of each 5 min is chosen to ensure the validity of data, which means that for each hour period, 120 samples are chosen at most. Data from the morning peak (8:00—9:00), night (21:00—22:00) and deep night (0:00—1:00) are selected for data analysis.

### 1.2 Research method

#### 1.2.1 Auto encoder

In this research, the input and output nodes of network are consistent with vectors of GPS transportation data:

$$\boldsymbol{X} = \{\delta, v_a, v_s, a_s, a_a^+, a_s^+, a_a^-, a_s^-\} \tag{1}$$

where $\delta$ is the proportion of over-speeding time; $v_a$ is the average speed of vehicle; $v_s$ is the standard deviation of vehicle speed; $a_s$ is the standard deviation of vehicle acceleration; $a_a^+$ is the average acceleration where $a > 0$; $a_a^-$ is the average acceleration where $a < 0$; $a_s^+$ is the standard deviation of vehicle acceleration where $a > 0$; $a_s^-$ is the standard deviation of vehicle acceleration where $a < 0$.

Each element of $\boldsymbol{X}$ is the result after standardization.

$$X_i = \frac{X_i - \overline{X_i}}{\delta_i} \tag{2}$$

where $\delta_i$ is the standard deviation of $\boldsymbol{X}$; $\overline{X_i}$ is the average value of $\boldsymbol{X}$.

The training set of the auto encoder is defined as

$$\boldsymbol{V} = \{X_1, X_2, ..., X_n\}^{\mathrm{T}} \qquad n \in M \tag{3}$$

where $X_i$ is a summary sample of vehicle data with a unique ID. The model is a five-layer self-coding network, and it is used for compressing information and extracting core characteristics for later analysis. The auto encoder has the advantage of reducing noise with minimum loss of raw data.

The activation function used in this model is tanh function:

$$f(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}} \tag{4}$$

#### 1.2.2 Self-organizing Mapping

Self-organizing mapping (SOM) is an unsupervised training neural network that automatically clusters input patterns through self-training. SOM neural network classifies the input mode set by finding an optimal reference vector set, and each reference vector is a connection weight vector corresponding to an output unit.

Compared with the traditional pattern clustering method, the cluster center formed by SOM can be mapped to surface or plane while keeping the topology unchanged.

Specific implementation steps are shown in Fig. 1.

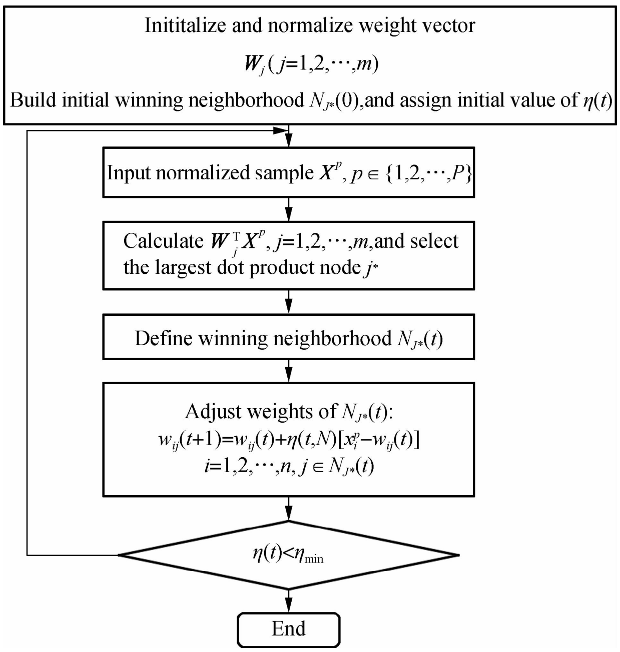


**Fig. 1** SOM computing process

## 2 Driving Behavior Evaluation Research

### 2.1 Driving behavior classification

The auto encoder is used to reduce the noise of raw data, and then SOM is used to classify driving behaviors as safe, relatively safe, relatively aggressive and aggressive, as shown in Fig. 2. Take speed, acceleration, and lateral offset as input values, and by using SOM and cluster analysis, the proportion distribution of different dangerous driving behaviors is carried out and shown in Tab. 2. The proportions are calculated based on raw data characteristics and clustered by SOM, so no thresholds are set manually, and $\eta$ is fixed according to the classified effectiveness. Relative thresholds setting is shown in Tab. 2.

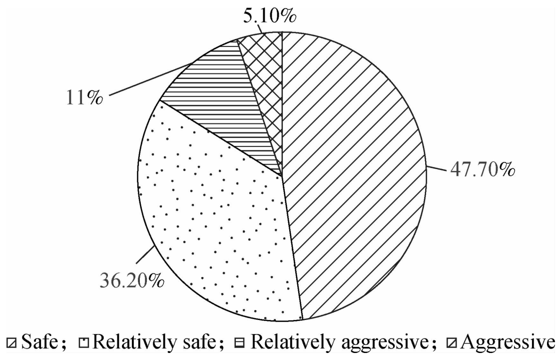


**Fig. 2** Percentage of various drivers

**Tab. 2** Thresholds settings for various dangerous driving behaviors

| Dangerous driving behavior | Indicators | Calculate method | Thresholds |
|---|---|---|---|
| Over speeding | Speed | Thresholds judgement | Based on road speed limit |
| Rapid speed changing | Absolute value of acceleration time percentage | Based on time percentage of $\lvert a\rvert>4$ m/s$^2$ and SOM is used to calculate rapid speed change proportion | $\lvert a\rvert>4$ m/s$^2$ |
| Rapid turning | Speed and angular speed | Thresholds judgement | $v<40$ km/h: angle change in 3 s period is greater than 90°; $v>40$ km/h: angular speed is greater than 5 °/s |
| Rapid lane changing | Direction angle average value, direction standard deviation | SOM classification | Average value and standard deviation of angular speed are [0.239, 72.69], [0.061, 49.5], [0.188, 31.82], [0.092, 12.65] |

Note: $a$ is the acceleration; $v$ is the vehicle speed.

The proportion of different dangerous driving behaviors is shown in Tab. 3.

**Tab. 3** Proportion distribution of different dangerous driving behaviors %

| Over speeding | Rapid speed change | Rapid turning | Rapid lane changing |
|---|---|---|---|
| 18 | 14 | 12 | 23 |

Driving direction is used to calculate rapid lane changing and rapid turning proportion. The proportion of rapid turning and rapid lane changing is calculated separately at the intersection area and road area, and the proportion is given based on SOM classification results. 16.1% of Shenzhen drivers are aggressive or relatively aggressive drivers, and over speeding and rapid lane changing the are the main dangerous driving behaviors.

### 2.2 Time distribution feature of driving behavior

Three time periods were extracted to carry out the Shenzhen driving behavior characteristics study, and the chosen time periods are morning peak (8:00—9:00), night (21:00—22:00) and deep night (0:00—1:00). Relative proportion is shown in Tab. 4. Traffic flow is the highest during morning peak, and night period represents normal driving behavior. During deep night, drivers are tired, so that fatal and serious injury accidents with injuries are likely to occur.

**Tab. 4** Driving behavior of different time periods

| Time period | Averages peed/ (km · h$^{-1}$) | Speed std/ (km · h$^{-1}$) | Proportion of low speed vehicles/% | Time proportion of low speed driving/% | Dangerous driving time proportion/% |
|---|---|---|---|---|---|
| 0:00—1:00(deep night) | 63.9 | 10.2 | 0 | 0.9 | 8.8 |
| 8:00—9:00(morning peak) | 55.0 | 13.4 | 1.9 | 6.0 | 8.1 |
| 21:00—22:00(night) | 57.7 | 11.9 | 0 | 2.0 | 7.0 |

In Tab. 4, the average speed is the average value of all vehicles; speed std is the average value of all vehicles' standard variation; the proportion of low speed vehicles is the proportion of vehicles travelling with a low speed($v$ < 10 km/h) more than 40% of total travelling time; the time proportion of low speed travelling is the time proportion of low speed travelling ($v$ < 10 km/h); dangerous driving time proportion is the time percentage of $|\text{speed} \times \text{acceleration}| > 100$[20] (km/h, m/s$^2$).

Conclusions can be drawn from Tab. 4. First, drivers drive the fastest in deep night, during which dangerous driving time proportion is the highest, and drivers tend to be more aggressive. Secondly, drivers travel at the lowest speed during the morning peak period when the low speed vehicle proportion is the highest. Drivers accelerate and decelerate frequently, and the speed variation is the highest. Thirdly, drivers' behaviors at 21:00—22:00 are the safest.

## 2.3 Driving behavior spatial distribution characteristics

There is a large difference between the infrastructure inside and outside the original Shenzhen Economic Special Zone, as well as the drivers' driving behavior. This paper takes Shekou area (in the original economic special zone) and East Shajing area (outside the original economic special zone) as examples to compare driving behavior in Shenzhen. The roads in the two chosen areas are mainly low-grade urban roads, as high-grade roads often cut across many administrative districts. High-grade roads mainly function for accessibility, while low-grade roads mainly function for flexibility. Continuity of traffic flow on low-grade roads is weak, which can better reflect regional travel characteristics. The chosen areas are shown in Fig. 3, and a driving behavior comparison is given in Tab. 5.


**Fig. 3** Typical traffic area

**Tab. 5** Driving behavior comparison between areas inside and outside the special economic zone

| Time period | Area | Average speed/(km · h$^{-1}$) | Speed std/(km · h$^{-1}$) | Proportion of low speed vehicles/% | Time proportion of low speed driving/% | Dangerous driving time proportion/% |
|---|---|---|---|---|---|---|
| 0:00—1:00 | East Shajing | 25.7 | 14.0 | 14.5 | 23.5 | 3.9 |
| (deep night) | Shekou | 33.8 | 14.7 | 13.2 | 21.5 | 7.6 |
| 8:00—9:00 | East Shajing | 16.8 | 11.2 | 47.0 | 48.0 | 2.0 |
| (morning peak) | Shekou | 19.4 | 11.8 | 41.7 | 42.6 | 4.0 |
| 21:00—22:00 | East Shajing | 21.3 | 12.1 | 29.4 | 33.1 | 3.0 |
| (night) | Shekou | 21.4 | 10.1 | 40.1 | 41.8 | 4.3 |

Comparing and analyzing the driving behavior inside and outside the original economic special zone, it can be summarized that the average driving speed of Shekou area is always higher than that of East Shajing. Low speed driving proportion in Shekou area during the morning peak and deep night is also lower than that in East Shajing, indicating that the infrastructure conditions in Shekou area are better than East Shajing's. The Shekou area has a higher proportion of dangerous driving time, which is more risky for driving.

## 2.4 Driving behavior analysis of main lane

This section analyzes the driving behavior of Beihuan Avenue, Longgang Avenue, Binhe Avenue and Meiguan Road. Results are shown in Tab. 6.

According to the analysis, the overspeed phenomenon of Beihuan Avenue and Meiguan Road during the off-peak hours is the most serious, and the dangerous driving time proportion is the highest. The random lane changing behavior of Binhe Avenue and Beihuan is more serious. The relationship between accidents with average speed and dangerous driving time are as shown in Fig. 4, Fig. 5 and Tab. 7.

According to the analysis, dangerous driving proportion during the late night period shows a good correlation

**Tab. 6** Driving behavior of different roads %

| Road section | Time period | Over speeding percentage | Dangerous driving time proportion | Random lane changing proportion |
|---|---|---|---|---|
| Beihuan Avenue | 0:00—1:00 | 17.0 | 10.0 | 17.5 |
| | 8:00—9:00 | 9.0 | 3.3 | 14.0 |
| | 21:00—22:00 | 10.1 | 7.1 | 14.0 |
| Longgang Avenue | 0:00—1:00 | 12.4 | 7.3 | 17.0 |
| | 8:00—9:00 | 8.9 | 4.8 | 12.2 |
| | 21:00—22:00 | 10.8 | 6.9 | 15.0 |
| Binhe Avenue | 0:00—1:00 | 14.5 | 8.5 | 21.0 |
| | 8:00—9:00 | 7.9 | 3.4 | 14.0 |
| | 21:00—22:00 | 8.0 | 4.5 | 14.4 |
| Meiguan Road | 0:00—1:00 | 15.9 | 7.9 | 11.0 |
| | 8:00—9:00 | 5.8 | 2.8 | 16.9 |
| | 21:00—22:00 | 9.4 | 5.8 | 12.5 |

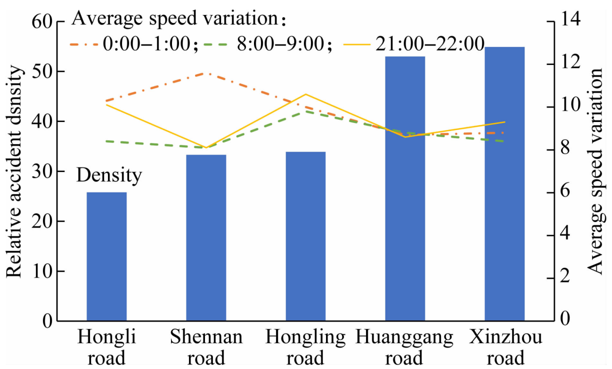


**Fig. 4** Relationship between average speed variation and accident density

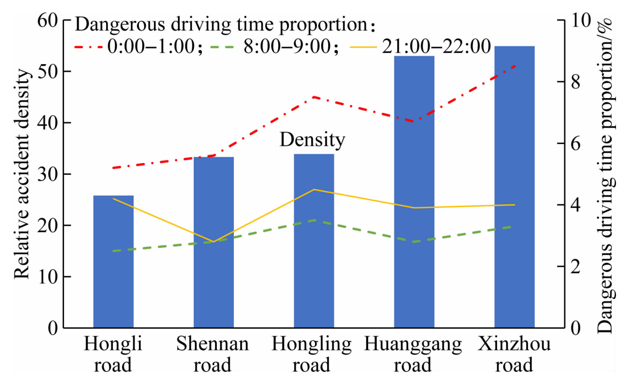


**Fig. 5** Relationship between dangerous driving time proportion and accident density

**Tab. 7** Dangerous driving vehicles proportion-accident density correlation

| Road section | Dangerous driving vehicles proportion/% | | | Relative accident density/ |
|---|---|---|---|---|
| | 0:00—1:00 | 8:00—9:00 | 21:00—22:00 | (times · (year · km) $^{-1}$) |
| Hongli Road | 5.2 | 2.5 | 4.2 | 25.8 |
| Hongling Road | 7.5 | 3.5 | 4.5 | 33.9 |
| Huanggang Road | 6.7 | 2.8 | 3.9 | 53.0 |
| Shennan Road | 5.6 | 2.8 | 2.8 | 33.3 |
| Xinzhou Road | 8.5 | 3.3 | 4.0 | 54.9 |

with relative accident density, indicating that the proportion of dangerous driving behavior during the late night period can be used as a good measure of road safety.

### 2.5 Driving behavior comparison of normal roads and accident-prone roads

To analyze the difference between the driving behavior of normal roads and accident-prone roads, the Beihuan-Huanggang road section and Beihuan-Caitian road sections (see Fig. 6) are taken as examples, and the relative indices are shown in Tabs. 8 to 10.

According to analysis, at deep night, the travelling speed of accident-prone roads is significantly higher than that of normal roads, and sharp turns, rapid acceleration/deceleration behavior proportions are significantly higher than those of normal road sections.


**Fig. 6** Location distribution of normal and accident-prone roads

**Tab. 8** Speed comparison between normal and accident-prone roads km/h

| Time period | Weekdays | | Weekends | |
|---|---|---|---|---|
| | Normal section | Accident-prone | Normal section | Accident-prone sections |
| 0:00—1:00 | 57 | 60 | 54 | 62 |
| 8:00—9:00 | 39 | 40 | 29 | 12 |
| 21:00—22:00 | 19 | 19 | 25 | 45 |

**Tab. 9** Acceleration proportion distribution comparison between normal and accident-prone roads %

| Acceleration/($m \cdot s^{-2}$) | Weekdays | | Weekends | |
|---|---|---|---|---|
| | Normal section | Accident-prone sections | Normal section | Accident-prone sections |
| (−3, −2] | 1.7 | 2.1 | 1.4 | 2.4 |
| (−2, −1.38] | 2.4 | 3.2 | 2.5 | 3.7 |
| (−1.38, 1.38) | 90.7 | 89.4 | 91.2 | 88.3 |
| [1.38, 2) | 3.2 | 3.3 | 2.9 | 3.4 |
| [2, 3) | 2.0 | 2.0 | 2.0 | 2.2 |

**Tab. 10** Sharp turn proportion comparison between normal and accident-prone roads %

| Time period | Normal sections | | Accident-prone sections | |
|---|---|---|---|---|
| | Sharp turn | Normal driving | Sharp turn | Normal driving |
| Weekdays | 3 | 97 | 6 | 94 |
| Weekends | 3 | 97 | 4 | 96 |

## 3 Conclusions

1) Based on the large volume of positioning data, this paper analyzes the driving behavior of Shenzhen drivers, gives the spatial and temporal distribution characteristics of driving behavior, and analyzes the driving behaviors of important passages and the correlation between dangerous driving behaviors and accidents. The method used in this paper is proved to be effective in analyzing GPS trajectory data, and provides a new way to study driving behavior. Rapid lane changing is the most common traffic violation in Shenzhen.

2) 16.1% drivers' behavior is relatively aggressive, and they are more accident-prone, so that more traffic management measures should be taken upon them.

3) During deep night, drivers tend to over-speed; and during the morning peak, due to traffic congestion, drivers tend to accelerate and decelerate frequently.

4) Comparison between the driving behavior on similar roads inside and outside the original economic special zone indicates that due to the difference of development levels, the infrastructure level inside is higher than that outside, which creates the difference in dangerous driving proportion.

5) Dangerous driving behavior during deep night is highly correlated to accident densities.

## References

[1] Tang Z Z, Zhang T J, He Y. *Road safety assessment*[M]. Beijing: China Communications Press, 2008. (in Chinese)

[2] Wu B, Yang Z Z, Xie J, et al. Vehicle speed distribution characteristics on work zone section of freeway[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2016, **16**(2): 219 − 224, 231. DOI: 10.16097/j.cnki.1009 − 6744.2016.02.033. (in Chinese)

[3] Thiffault P, Bergeron J. Fatigue and individual differences in monotonous simulated driving[J]. *Personality and Individual Differences*, 2003, **34**(1): 159 − 176. DOI: 10.1016/s0191 − 8869(02)00119 − 8.

[4] Guo S, Teng J, Guo X, et al. Research on fatigue driving of commercial vehicle drivers based on multi-source real data [C]//*China Intelligent Transportation Annual Meeting*. Guangzhou, China, 2014: 85 − 96. (in Chinese)

[5] Tao J. The deep analysis of traffic accidents drivers' causes in operating vehicles[J]. *Shanxi Science and Technology of Communications*, 2018, **253**(4): 155 − 158. (in Chinese)

[6] Zhang L F, Chen C, Zhang J Y, et al. Modeling lane-changing behavior in freeway off-ramp areas using naturalistic driving data[J]. *Journal of Tongji University* (*Natural Science*), 2018, **46**(3): 318 − 325, 333. (in Chinese)

[7] Wang X S, Zhu M X, Xing Y L. Impacts of collision warning system on car-following behavior based on naturalistic driving data [J]. *Journal of Tongji University* (*Natural Science*), 2016, **44**(7): 1045 − 1051. (in Chinese)

[8] Wang X S, Li Y. Characteristics analysis of lane changing behavior based on the naturalistic driving data[J]. *Journal of Transport Information and Safety*, 2016, **34**(1): 17 − 22. DOI: 10.3963/j.issn 1674 − 4861.2016.01.001. (in Chinese)

[9] Wang X S, Yang M M. Cut-in behavior analyses based on naturalistic driving data[J]. *Journal of Tongji University* (*Natural Science*), 2018, **46**(8): 1057 − 1063. DOI: 10.11908/j.issn.0253 − 374x.2018.08.008. (in Chinese)

[10] Olsen E C B, Lee S E, Wierwille W W, et al. Analysis of distribution, frequency, and duration of naturalistic lane changes[J]. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2002, **46**(22): 1789 − 1793. DOI: 10.1177/154193120204602203.

[11] Bone S A, Mowen J C. Identifying the traits of aggressive and distracted drivers: A hierarchical trait model approach [J]. *Journal of Consumer Behaviour: An International Research Review*, 2006, **5**(5): 454 − 464. DOI: 10.1002/cb.193.

[12] Paaver M, Eensoo D, Kaasik K, et al. Preventing risky driving: A novel and efficient brief intervention focusing on acknowledgement of personal risk factors[J]. *Accident Analysis & Prevention*, 2013, **50**: 430－437. DOI: 10.1016/j.aap.2012.05.019.

[13] Ahmed M, Yu R, Abdel-Aty M. Safety applications of automatic vehicle identification and real-time weather data on freeways[C]//*18th ITS World Congress*. Orlando, FL, USA, 2011.

[14] Ahmed M, Abdel-Aty M. A data fusion framework for real-time risk assessment on freeways[J]. *Transportation Research Part C: Emerging Technologies*, 2013, **26**: 203－213. DOI: 10.1016/j.trc.2012.09.002.

[15] Li Y, Chen Y R. Driver vision based perception-response time prediction and assistance model on mountain highway curve[J]. *International Journal of Environmental Research and Public Health*, 2016, **14**(1): 31. DOI: 10.3390/ijerph14010031.

[16] Grengs J, Wang X G, Kostyniuk L. Using GPS data to understand driving behavior[J]. *Journal of Urban Technology*, 2008, **15**(2): 33－53. DOI: 10.1080/10630730802401942.

[17] Zhao Y N, Yamamoto T, Morikawa T. An analysis on older driver's driving behavior by GPS tracking data: Road selection, left/right turn, and driving speed[J]. *Journal of Traffic and Transportation Engineering (English Edition)*, 2018, **5**(1): 56－65. DOI: 10.1016/j.jtte.2017.05.013.

[18] Liao L, Wu J, Zou F, et al. Trajectory topic modelling to characterize driving behaviors with GPS-based trajectory data[J]. *Journal of Internet Technology*, 2018, **19**(3): 815－824.

[19] Liu Y J, Zeng C, Wang S J, et al. An evaluation method of safety and energy-saving driving behavior based on satellite positioning data[J]. *Journal of Highway and Transportation Research and Development*, 2018, **35**(1): 121－128, 158. (in Chinese)

[20] Zhu X Y, Hu X B, Chiu Y C. Design of driving behavior pattern measurements using smartphone global positioning system data[J]. *International Journal of Transportation Science and Technology*, 2013, **2**(4): 269－288. DOI: 10.1260/2046－0430.2.4.269.

# 基于海量车辆轨迹数据的机动车驾驶员驾驶行为评价

孙　超[1,2]　陈小鸿[1]　张红军[1,3]　张俊峰[2]

([1]同济大学道路与交通工程教育部重点实验室，上海 201804)

([2]深圳市城市交通规划设计研究中心股份有限公司，深圳 518021)

([3]College of Civil and Environmental Engineering, University of California Davis, Davis 95616, USA)

**摘要：**基于监控数据及车辆总线数据的营运车驾驶行为评价研究方法已相对成熟，但难以推广至私家车，针对自然驾驶试验样本容量小、设备昂贵的缺点，提出了一种基于海量秒级 GPS 导航数据的驾驶行为评价方法. 运用自动编码机对数据进行降噪处理，结合行车动力学、SOM 自组织映射分类等方法给出超速、急变速、频繁变道、急转弯等行为的判定方法及阈值，计算不同驾驶特性驾驶员及危险驾驶行为比例. 在此基础上，分析得到深圳市机动车驾驶员驾驶行为时空分布特征、典型通道驾驶行为特征，并选取典型道路进行事故多发段与正常段的驾驶行为对比. 结果表明，深圳危险驾驶行为以频繁变道和超速为主，16.1% 的驾驶员驾驶行为偏冒进，原特区外相对原特区内危险驾驶行为比例更高，危险驾驶行为与事故密度呈高度相关.

**关键词：**驾驶行为；GPS 导航数据；自动编码机；自组织映射

**中图分类号：**U491