Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1814608, 12 pages
http://dx.doi.org/10.1155/2016/1814608

Research Article

Dynamic Recognition of Driver's Propensity Based on GPS Mobile Sensing Data and Privacy Protection

Xiaoyuan Wang,[1,2] Jianqiang Wang,[2] Jinglei Zhang,[1] and Jingheng Wang[3]

[1] *School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255049, China*
[2] *State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China*
[3] *Shandong Zibo Experimental High School, Zibo 255000, China*

Correspondence should be addressed to Xiaoyuan Wang; wangxiaoyuan@sdut.edu.cn
and Jianqiang Wang; wjqlws@tsinghua.edu.cn

Received 22 March 2016; Revised 9 August 2016; Accepted 29 August 2016

Academic Editor: Huaguang Zhang

Driver's propensity is a dynamic measurement of driver's emotional preference characteristics in driving process. It is a core parameter to compute driver's intention and consciousness in safety driving assist system, especially in vehicle collision warning system. It is also an important influence factor to achieve the Driver-Vehicle-Environment Collaborative Wisdom and Control macroscopically. In this paper, dynamic recognition model of driver's propensity based on support vector machine is established taking the vehicle safety controlled technology and respecting and protecting the driver's privacy as precondition. The experiment roads travel time obtained through GPS is taken as the characteristic parameter. The sensing information of Driver-Vehicle-Environment was obtained through psychological questionnaire tests, real vehicle experiments, and virtual driving experiments, and the information is used for parameter calibration and validation of the model. Results show that the established recognition model of driver's propensity is reasonable and feasible, which can achieve the dynamic recognition of driver's propensity to some extent. The recognition model provides reference and theoretical basis for personalized vehicle active safety systems taking people as center especially for the vehicle safety technology based on the networking.

1. Introduction

The influence of drivers' physiology and psychology characteristics on traffic safety is mainly represented as the driver's propensity [1]. Driver's propensity is the attitude experience of the drivers for the real traffic conditions affected by various dynamic factors, as well as the preference drivers show that suits with decision or behavior value. It is the dynamic measure of driver's emotional preference characteristics in driving process which guides the drivers' intent and affects vehicle handling behavior directly. It is a core parameter to compute driver's intention and consciousness in safety driving assist system, especially vehicle collision warning system. It is also an important influence factor to achieve the Driver-Vehicle-Environment Collaborative Wisdom and Control macroscopically, especially the vehicle safety controlled technology. Its types can be divided into radical type, common type, and conservative type. The previous studies were mainly focused on drivers' psychological characteristics measurement on relatively static and macrocosmic factors and the traffic safety effect; few research works have been conducted on microcosmic, dynamic measurement and computation of driver's emotion state in view of vehicle active safety. Xie and Wang [2] constructed a simple traffic model including car-following, lane-changing, and overtaking with consideration of driver's emotion under simplified road condition and discussed the influence of the change of cognition emotion on driving strategy; Wu and Hu [3, 4] studied the driver's behavior characteristics caused by anger, identified the state of angry driving, and researched the influence on driver's physiology, psychology, and traffic security affected by angry emotion, but there was no research concerning the microscopic and dynamic characteristics of the time-varying emotions. Set about affective computing, Lin and

Feng [5] analyzed and judged whether the drivers were in the aggressive driving state or not using the speech emotion recognition, facial expression recognition, vehicle driving state detection methods and so on and used the vehicular intelligent aided system to make decision of the security alert or safety anticollision for vehicles, to improve driving safety factor. In order to uncover whether emotional auditory stimuli can affect risky behavior in hazardous situations, Di Stasi et al. [6] organized that forty-nine volunteers rode a motorcycle in a virtual environment and went through a number of preset risky scenarios, some of which were cued by a kind of sound (beep, positive emotional sound, or negative sound). Results showed that the beep reduced frequency of accidents in upcoming risky situation, while the emotional cues did not. Likewise, the beep induced the drivers to decrease their speed and focus their gaze on relevant areas of the visual field, while the emotional sounds did not. These results suggest that auditory warning systems for vehicles should avoid using emotion-laden sounds, as their affective content might diminish their utility to increase driving alertness. Trick et al. [7] discussed implications for both basic research on attention-emotion and applied research on driving. The researchers focus on the effect of fleeting emotions on hazard perception and steering while driving. Taubman-Ben-Ari [8] studied the effects of positive emotion priming on the willingness to drive recklessly. The research shows that positive emotions of a relaxing nature, as well as thinking about the meaning of life, lowered the willingness to engage in risky driving. Many risk factors such as speed, drowsiness, drugs and alcohol consumption, and state of the car have been identified and have permitted the development of prevention policies. In contrast psychological factors remain poorly studied, particularly emotional state. For this phenomenon, M'bailara et al. [9, 10] researched the relationship between accident and emotional state. The results revealed that emotional reactivity is significantly associated with the drivers' responsibility, suggesting that emotional hypo- or hyperactivity is a significant source of accidents. Arnau-Sabatés et al. [11] studied those emotional abilities as predictors of risky driving behavior. The risky driving attitudes and emotional abilities of 177 future driving instructors were measured. The results demonstrate that risky attitudes correlate negatively with emotional abilities. The results obtained can be used to formulate the corresponding prevention programs to reduce risky driving behaviors. Chan and Singhal [12] analyzed the effects of emotional distraction on driving. This purpose was achieved using a driving simulator and three different types of emotional information: neutral, negative, and positive emotional words. Participants also responded to target words while driving and completed a surprise-free recall task of all the words at the end of the study. The findings suggest that emotional distraction can modulate attention and decision-making abilities and have adverse impacts on driving behavior.

In order to fill the vacancy of related research about microcosmic, dynamic measurement and calculation of driver emotion, Zhang et al. [13, 14] from the perspectives of vehicle safe driving support system, especially in the automotive collision warning system driver's intention, emotion, and other psychological effects associated with coupling awareness computing core research scientific issues, collaborative applications car GPS, car laser radar and other dynamic data acquisition board video sensor system, variable data capture vehicles environment, freedom of driving, the car following a complex and multilane vehicle under the cluster grouping of specific traffic scene when the driver becomes law emotion and collaborate deduction, especially for emotional dynamic measure, online real-time identification and characterization of key scientific issues, such as exploratory research. However, the study also found that, for the LiDAR data acquisition, due to its precision optical characteristics LiDAR is expensive, its installation is relatively complex, seismic immunity is weak, and it is prone to some error so the penetration rate is not high and experiments and subsequent data process are more cumbersome.

With GPS users increasing, especially the popularity of mobile GPS, researchers make it possible to excavate breadth and depth of traffic information GPS data. GPS has the advantages of low cost, high penetration, being easy to carry, zero invasive effect essentially for drivers, antivibration, and strong anti-interference. Data is paid more and more attention by traffic researchers. However, GPS data (such as location, trajectory) is easy to leak travelers hobbies, behavior patterns, habits, and other personal privacy. Date of GPS is often used by malicious attackers to detect and analyze the current position of the traveler, the last visited location, home address, working place, income levels, health status, political affiliation, etc.; it even can cause threat to personal safety of travelers. Thus, many scholars at home and abroad strengthen the protection of privacy on the GPS basic research in the promotion of GPS applications at the same time.

GPS, especially the popularity of mobile phones and portable GPS, makes the concept of synergy based on vehicle road traffic management and control possible and provides secure innovative ideas for the dangerous state detection based on driver and vehicle dynamic security technology of shared control by vehicle network. This technology can make use of nonvehicle personnel (such as traffic police and the driver relatives) of the vehicle to reach remote sharing and control interventions (such as advice to or warning of the driver, and the implementation of open warning light on the vehicle, limiting-velocity, traffic controls, pull-over, and other remote auxiliary controls), to solve the traffic safety problem when the driver is in danger. Meanwhile, real-time network-based storage and sharing driver status information can provide reference for accident analysis and tracking accountability to reduce dangerous driving behaviors of the drivers.

To meet the needs of vehicle road collaborative innovation technology, especially in shared controlled vehicle study security innovation vehicle networking, the GPS semantic mining and privacy preserving are collaboratively considered, and travel time is selected as measure basis; drivers' tendency dynamic identification theory is studied in this paper. We get dynamic data of human vehicle and environment which is corresponding to different propensity (such as radical type, common type, and conservative type) driver through the psychological test, virtual driving experiment, and real vehicle experiment, analyze data and extract the travel time as the parameter to build dynamic recognition model of driver's propensity based on support vector machine, and use experimental data for model parameter calibration, verification, and analysis.

2. Dynamic Recognition of Driver's Propensity

2.1. Support Vector Machine. Support vector machine (SVM) is a kind of new machine learning method proposed by Vapnik [15–17], based on the statistical learning theory, has a complete foundation of statistical learning theory and excellent learning performance, and is the youngest content of statistical learning theory and the most practical part. Contrasting with the heuristic neural network learning method and the implementation, SVM has a more rigorous theoretical and mathematical foundation; local minimum problem does not exist. The technology can solve the practical problems of small sample, high dimension, nonlinear and local minimum points, and so forth and keep good generalization capability in the small sample conditions, successfully applied in signal processing, regression analysis, function approximation, and other fields [18–20].

2.1.1. The Basic Idea of Support Vector Machine. The basic idea of SVM method is based on structural risk minimization (structural risk minimization, SRM) principle. Through a specific nonlinear mapping, the sample space is mapped to a high dimension as feature space of infinite dimensional (Hilbert space). In the feature space, to find the optimal classification or regression linear hyperplane, the plane is taken as the classification decision surface, so as to solve the problems in sample space of nonlinear classification and regression.

The learning process and model selection phase are two important aspects of the SVM algorithm. Reference [21] "Tikhonov, Ivanov and Morozov regularization for Support Vector Machine Learning" introduces the learning method of training support vector machine in consideration of structural risk minimization, comparing the advantages and disadvantages of three kinds of regularization algorithm; this paper chooses the appropriate regularization algorithm to finish the learning process of SVM based on achievements of [21] for achieving unification of the algorithm effectiveness and operability. Reference [22] "In-Sample and Out-of-Sample Model Selection and Error Estimation for Support Vector Machines" details common methods of SVM model selection phase, introducing the difference between in-sample and out-of-sample and application condition of the two methods. This paper perfects the SVM model selection phase of the dynamic recognition of driver's propensity based on travel time in order to make the model more reasonable and reliable.

The decision function of SVM is only determined by a few support vectors; the computational complexity depends on the number of support vectors, rather than the dimension of the sample space, which avoids the "dimension disaster in some sense."

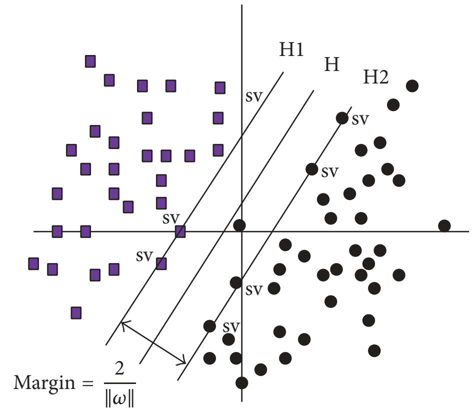


Figure 1: Optimal classification plane diagram of support vector machine.

2.1.2. Support Vector Machine Model. If there is a hyperplane such that two kinds of data can be classified and the distance between the date and the hyperplane is the largest, the plane is the optimal hyperplane.

(1) Linear Optimal Classification Hyperplane. The training sample consists of two categories of a given group (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$, where $x_i \in R^n$ and $y_i \in \{-1, 1\}$, if x_i belongs to the first category, so y_i is labeled as positive ($y_i = 1$); otherwise, $y_i = -1$ and $i = 1, 2, \ldots, n$. Specific ideas are shown in Figure 1. In Figure 1, circular and square, respectively, represent two types of samples, H is classification hyper plane, H1 and H2, respectively, represent the various types and are samples nearest to H and the plane is parallel to H, and the distance between them is called classification interval (margin). The so-called optimal classification face is required to correctly separate two kinds with farthest classification interval. The sample point of H1 and H2 is support vector.

If the sample is linearly separable, there is a hyperplane H:

$$\omega \cdot x + b = 0. \tag{1}$$

Make

$$\begin{aligned} \omega \cdot x_i + b &\geq 1, \quad y_i = 1, \\ \omega \cdot x_i + b &\leq -1, \quad y_i = -1. \end{aligned} \tag{2}$$

The formula (2) is normalized, the linearly separable data meet:

$$y_i \left[(\omega \cdot x) + b \right] \geq 1, \quad i = 1, 2, \ldots, n. \tag{3}$$

According to the optimal hyperplane definition, classification interval can be expressed as

$$\rho = \min_{\{x_i, y_i = 1\}} \frac{\left| \omega \cdot x_i + b \right|}{\|\omega\|} + \min_{\{x_j, y_j = -1\}} \frac{\left| \omega \cdot x_j + b \right|}{\|\omega\|} = \frac{2}{\|\omega\|}, \quad i = 1, 2, \ldots, n. \tag{4}$$

Make the max optimal $2/\|\omega\|$, so $(1/2)\|\omega\|$ or $(1/2)\|\omega\|^2$ is the min. A linear support vector machine is transformed into the problem of solving the following two convex programing problems:

$$\min \frac{1}{2}\|\omega\|^2$$
$$\text{Constraint conditions}: y_i\left[\left(\omega \cdot x_i\right)+b\right] \geq 1, \tag{5}$$
$$i = 1,2,\ldots,n.$$

The optimal solution can be obtained by the following Lagrange function:

$$L(\omega,b,\alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{n}\alpha_i\left[y_i\left(\omega \cdot x_i + b\right) - 1\right], \quad i = 1,2,\ldots,n, \tag{6}$$

where $\alpha_i \geq 0$ $(i = 1,2,\ldots,n)$ is the Lagrange multiplier.

To obtain the optimal solution for the above problem as a classification function,

$$f(x) = \operatorname{sgn}(\omega \cdot x + b) = \operatorname{sgn}\left[\sum_{x_i \subset \mathrm{SV}} \alpha_i y_i \left(x_i \cdot x\right) + b\right]. \tag{7}$$

(2) The Generalized Optimal Classification Hyperplane. The optimal hyperplane is discussed that the linear problems can be divided, in the training sample set linear inseparable case; some training samples cannot satisfy condition (1) and then can join a relaxation factor in conditions $\xi_i \geq 0$, which is

$$y_i\left[(\omega \cdot x) + b\right] \geq 1 - \xi_i. \tag{8}$$

The objective function is to find the minimum value of $(1/2)\|\omega\| + c\sum_{i=1}^{n}\xi_i$, where c is the penalty function; a larger c indicates more punishment misclassifications.

For the nonlinear problem, make the nonlinear problem $Q(x) : R^n \to Z$, Z is a high dimensional inner product space called the feature space, and $Q(x)$ is the feature mapping. Then one constructs the generalized optimal hyperplane in Z. One does not need to consider its exact form structure and only needs to carry on the inner product computation in the high dimensional space; kernel function $K(x_i, y_i)$ can be introduced. As long as the kernel function is satisfying the conditions of the inner product, it corresponds to a transformation space.

The nonlinear decision function is constructed in the input space:

$$y(x) = \operatorname{sgn}(\omega \cdot Q(x) + b) = \operatorname{sgn}\left[\alpha_i y_i K\left(x_i, x\right) + b\right]. \tag{9}$$

Type $K(x_i, x_j) = Q(x_i) \cdot Q(x_j)$ is called the kernel function; $0 \leq \alpha_i \leq c$ $(i = 1,2,\ldots,n)$ is the Lagrange multiplier.

Learning machine that can construct the decision-making function is called support vector machine. The structure of support vector machine is given in Figure 2.

(3) Kernel Function. One of the features of SVM is the introduction of kernel function. The low dimensional space vector set is usually difficult to classify, the solution is to be mapped into a high dimensional space, but this approach increases the complexity of calculation, and the kernel function cleverly solves the problem. At present, the kernel function is studied mainly in the following forms:

(1) The linear kernel function: $K(x, x_i) = x \cdot x_i$.
(2) Polynomial kernel function: $K(x, x_i) = [(x \cdot x_i) + 1]^d$.
(3) Gauss radial basis kernel function: $K(x, x_i) = \exp\{-|x, x_i|^2/2\sigma^2\}$.
(4) Neural network kernel function: $K(x, x_i) = \tanh(\nu(x \cdot x_i) + c)$.

d is the number of polynomials; σ is width parameter to RBF function; and ν and c are constants.

2.2. Experiment Design

2.2.1. Psychological Test. According to the questionnaire and the method in [13, 14], investigation and evidence collection to the driver show the driver's propensity types preliminarily. The content of psychological questionnaire reflects the psychological characteristics of drivers. 28 options questions in the table have been given scores in accordance with incremental numbers referring to the classic psychological scale. The larger the value, the greater the possibility that selecting this option on behalf of the driver is conservative. Scores of 28–48 are radical types, 49–68 are the common types, and 69–88 are the conservative type. In order to explain the content of the questionnaire, but being limited by the space, the following lists part of the subject and the option of the psychological questionnaire.

Part of the Subject and the Option of the Psychological Questionnaire

(1) Gender

Male (1), Female (2)

(2) Age

<30 years old (1), 30–40 years old (2), 40–50 years old (3), >50 years old (4)

(3) Driving years

<5 years (1), 5–10 years (2), 10–15 years (3), >15 years (4)

(4) Driving speed will exceed the speed limit or not when there is no other vehicles interference

Often (1), Occasionally (2), Never (3)

(5) The driver often follows the vehicle ahead or not

Yes (1), Maybe (2), Not usually (3)

(6) The driver always wants to overtake

Yes (1), Maybe (2), Not usually (3)

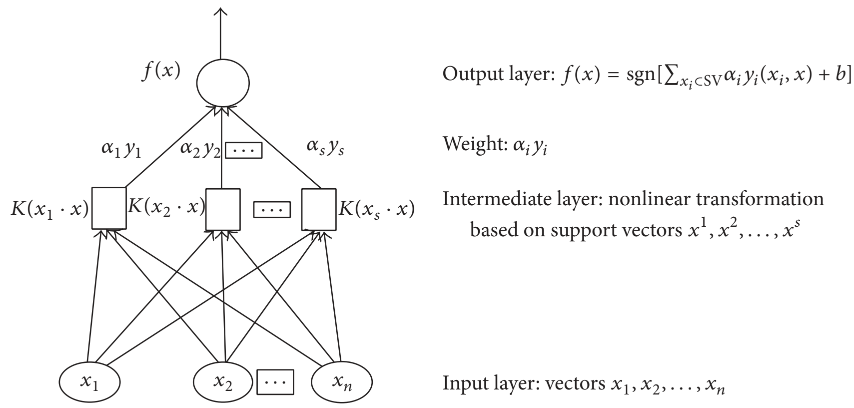


Figure 2: The structure of support vector machine.

(7) The driver often makes urgent acceleration or deceleration

Yes (1), Maybe (2), Not usually (3)

(8) Drivers may feel angry sometimes when they are overtaken

Yes (1), Maybe (2), Not usually (3)

(9) The driver will overtake when the headway is narrow

Yes (1), Maybe (2), Not usually (3)

(10) The mood will be agitated when there is a traffic jam

Yes (1), Maybe (2), Not usually (3)

(11) The driver will speed up to pass the intersection during signal alternating

Often (1), Occasionally (2), Never (3)

(12) The driver will speed up to go through corners

Yes (1), Maybe (2), Not usually (3)

(13) Drivers will improve the speed imperceptibly when there are friends in their cars

Yes (1), Maybe (2), Not usually (3)

(14) Drivers change driving route at the same time that they make a turn light

Yes (1), Maybe (2), Not usually (3)

The solving process of each subject's scoring average, standard deviation, coefficient of correlation is as follows. Organize that 50 drivers do the psychological questionnaire in accordance with the requirements, choose the options that are the closest ones to the feelings, thoughts, and behavior mostly, getting each subject's score of each driver in the psychological test. Using the first subject as instance, we can get the scoring average by using the first subject's total divided by headcount; we can get the standard deviation by using the score average according to the standard deviation formula and then with the score average and standard deviation we can get the coefficient of correlation.

Coefficient of internal consistency (Cronbach's Alpha coefficient) mainly reflects the reliability of the relationship between the subjects within the test, reviewing whether each subject of the test measures the same content and trait. Coefficient of internal consistency being greater than 0.7 means that the reliability of the scale is higher.

Shandong Jiaoyun group is commissioned to investigate through questionnaire in 50 drivers. Take the test scores as a sample, use SPSS17.0 software for statistical analysis to evaluate the reliability (stability of psychological measurement tools) and validity (effectiveness of measurement tool) of the questionnaire [15], and internal consistency coefficient is $0.836 > 0.8$, which indicates that the scale has high reliability homogeneity. Table 1 corresponds to mean, standard deviation, and the total score Pearson correlation coefficient in each item; about 90% problems of the score and total score in the 0.05 and 0.01 levels are significantly correlated, indicating that the scale has good content validity and can carry out test.

2.2.2. Vehicle Experiment. In the vehicle experiment, the research perspective is based on the vehicle safety controlled technology, respecting and protecting the driver privacy as a precondition. The experiment roads travel time is taken as the characteristic parameter obtained through GPS. The dynamic recognition model of driver's propensity is established based on support vector machine. In this paper the permeability of 20%, 25%, 30%, 35%, 40%, 45%, and 50% under the driver's propensity dynamic identification is studied, space is limited, and take the permeability of 20% as an example.

(1) Experimental Equipment. In city road environment, dynamic Driver-Vehicle-Environment information acquisition system (as shown in Figure 3, including SG299-GPS noncontact multifunction speedometer, BTM300-905-200 laser range sensor, HD camera, MiniVcap monitoring system,

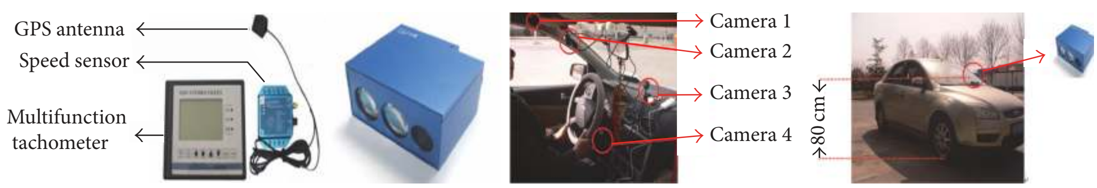


Figure 3: Dynamic Driver-Vehicle-Environment information acquisition system.

Table 1: Corresponding average value, standard deviation, and the Pearson correlation coefficient of total score for each question.

Title	Average value	Standard deviation	Correlation coefficient
1	1.30	0.470	0.710**
2	2.40	0.821	0.397**
3	2.15	0.933	0.518**
4	1.90	0.718	0.640**
5	1.75	0.716	0.451**
6	2.05	0.759	0.649**
7	1.90	0.718	0.580**
8	1.80	0.523	0.440**
9	1.80	0.768	0.133
10	2.00	0.562	0.575**
11	2.05	0.686	0.238
12	1.75	0.550	0.255
13	1.80	0.768	0.412*
14	2.00	0.795	0.751**
15	1.90	0.788	0.467**
16	2.00	0.795	0.575**
17	2.00	0.562	0.575**
18	1.95	0.605	0.585**
19	2.15	0.813	0.387*
20	1.95	0.686	0.669**
21	1.95	0.759	0.387*
22	1.75	0.716	0.536**
23	1.90	0.718	0.648**
24	1.85	0.671	0.614**
25	2.00	0.795	0.594**
26	1.80	0.616	0.552**
27	1.95	0.510	0.424*
28	2.40	0.995	0.750**

Note: the correlation coefficient table superscripts $*$ that correspond to the project were significant at 0.05 level; the correlation coefficient table superscripts $**$ that correspond to the project were significant at 0.01 level.

HD video camera, and notebook computer) is used to collect and process experiment data. The paper introduces the main instruments of dynamic Driver-Vehicle-Environment information acquisition system in this experiment, the experiment mainly uses SG299-GPS noncontact multifunction speedometer of dynamic Driver-Vehicle-Environment information acquisition system to collect driver's travel time for data processing, and the collected data of other instruments

Figure 4: Experimental route.

is not applied to the modeling. In addition, software SPSS17.0, Ulead VideoStudio 10, and so forth, are used [13, 14].

(2) Experimental Conditions. The experimental conditions are sunny weather and dry road surface; time is a normal working day at 8:30–10:30. Liantong Road is a two-way road with six lanes. According to the field survey, the traffic flow is 800 vehicles per hour of Liantong Road in Zibo City, Shandong province. Road traffic is stable but does not reach the congestion state. The experimental samples are 50 taxi loaded GPS, according to the traffic flow rules, to make experiments with no influence for traffic flow as the premise. Considering the influence of signal intersection, the queue length, the turning of the vehicle, and other external environmental factors on the accuracy in the experiment, the starting point and ending point of experiments in the experimental section are selected that the distance between the two ends of intersection is 100 meters; flag poles were labeled in Figure 4. The target segment length is 1200 meters.

(3) Experiment Content. According to the theory method in [13, 14] for vehicle test, the vehicle are running on the specified route, to ensure the normal operation of the instruments in the experimental process; dynamic comprehensive information collection system can in real time record the relevant experimental data.

The driving route is Nanjing Road-Liantong Road-Century Road-Huaguang Road (red and blue arrows in Figure 4) near Shandong University of Technology in Zibo, Shandong province, blue arrow part is the experimental route, and experimental road section is between purple

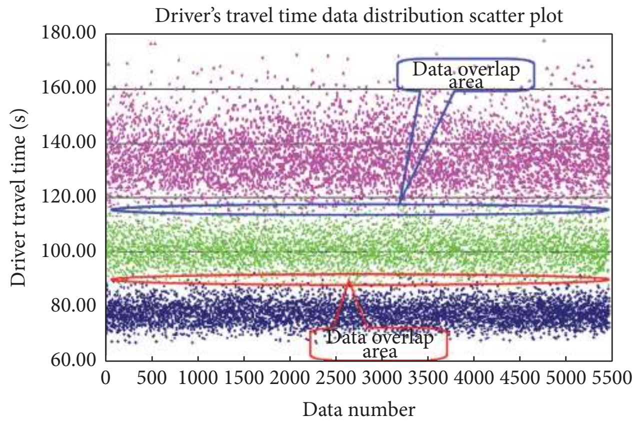


FIGURE 5: Driver's travel time data scatter diagram.

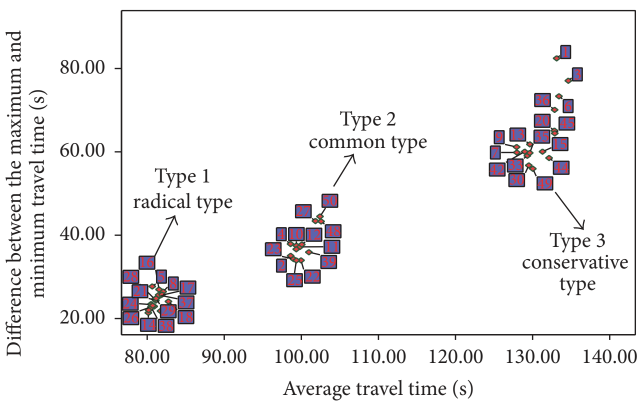


FIGURE 6: Drivers clustering scatter diagram.

flagpoles (target section of the research). The route of the vehicle equipping the dynamic Driver-Vehicle-Environment comprehensive information acquisition system is shown in Figure 4. When the target vehicle moves from Nanjing Road into the Liantong Road and through the start point marked by flagpoles, start time is being recorded, and when it passes the end point, the overtime is being recorded, so as to obtain the target vehicle travel time through the experimental section. We used HD camcorder, pedal force tester, laser range finder, GPS, gyroscope, and noninvasive test equipment to collect and record the target vehicle and its (internal and external) relevant dynamic sensor information, to prepare for subsequent research.

(4) Data and Analysis. Figure 5 is the experimental data of different drivers travel time scatter distribution. Scatter obvious agglomeration in three different regions, according to the clustering results, with three kinds of color rendering in order to understand.

Figure 6 is the result of clustering analysis using the discrete data.

According to the clustering results, the driver may be categorized into three categories: radical type, common type, and conservative type. The classification results and the

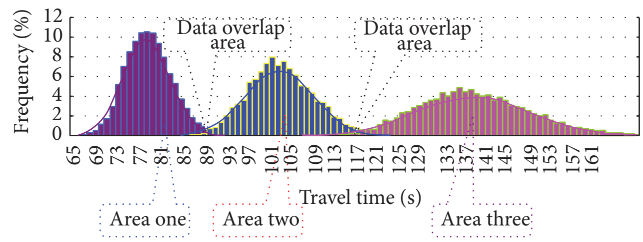


FIGURE 7: The driver's travel time frequency distributions and fitted curve.

psychological test results are matched basically determined by the clustering analysis.

In Figure 6 "drivers clustering scatter diagram," explanation of drivers clustering scatter plot is as follows: draw the driver's travel time scatter diagram, as shown in Figure 6, and the points gather for three categories. The points near 80 s are for category 1 "radical driver"; the points near 100 s are for category 2 "common driver"; and the points near 130 s are for category 3 "conservative driver." Three driver types corresponding to three kinds of points are annotated on the diagram in Figure 6.

The explanation of the relationship between travel time and psychological parameters is as follows: psychological test score can reflect the driver's propensity, and the length of the travel time can also reflect the driver's propensity. According to the psychological test score driver can be categorized into radical type, common type, and conservative type, and, according to the clustering analysis of travel time parameter, the data points of travel time gather into three categories, the points are near 80 s, 100 s, and 130, and the three categories are on behalf of the radical driver, the common driver, and the conservative driver. So travel time and psychological state are consistent.

The data is distributed intensively in three different regions near travel time of 80 s, 100 s, and 130 s, which is observed in Figure 5. A small amount of data scatters in the overlap regions. Collect different travel time frequency distribution and data fitting curve in three different regions, as shown in Figure 7.

Three regional and overall data description statistics are given in Table 2.

The above analysis shows that identifying and distinguishing the driver's propensity is possible by using the travel time as parameter, choosing and constructing reasonable model.

2.2.3. Virtual Driving Experiment. Use a driving simulator of FORUM8 Japan Co. production to simulate the experiment in virtual driving experiment (in Figure 8); the driving simulator for UC-Win road ver.4 is used in the simulation software. The software has driving simulation function and can simulate various types of driving conditions with the peripheral equipment, output characteristic index of different vehicle travel time, velocity, displacement, acceleration, the vehicle position, the steering wheel operation amount, accelerator pedal, and brake pedal force.

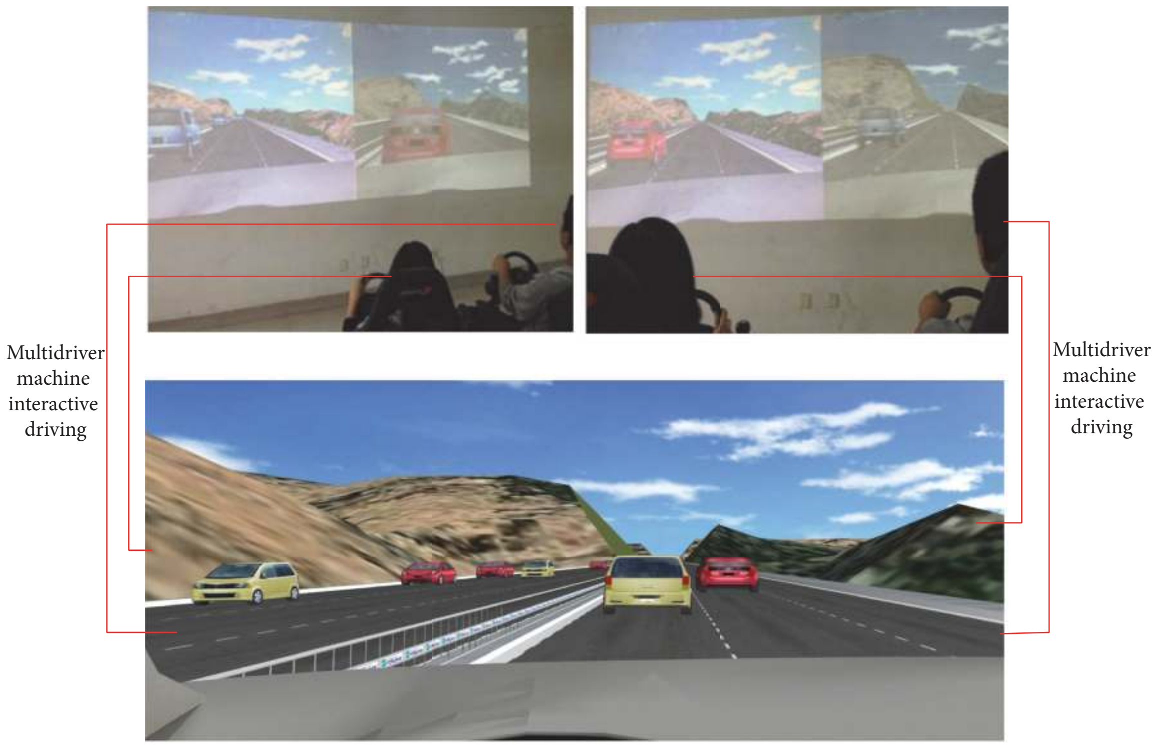


Figure 8: Virtual driving experiment.

Table 2: Three regional description statistics.

Data area	Description statistics: Maximum (seconds)	Minimum (seconds)	Average (seconds)
Area one	96.93	65.55	77.87
Area two	129.66	83.40	101.30
Area three	189.44	103.90	133.37
Data area	189.44	65.55	104.83

In the verification experiment, to set virtual reality scene of city traffic, road traffic standards in China, and the corresponding road facilities before the experiment, give the driver driving simulator training and finally obtain the corresponding experimental data. Only the driver's dynamic travel time is considered in this paper.

2.3. Recognition of Driver's Propensity Based on Support Vector Machine. This paper, based on the support vector machine, in different permeability (due to limited space, this paper takes 20% permeability as an example to elaborate) collects the characteristic parameter (travel time) of drivers in certain experimental sections to recognize driver's propensity.

As a preliminary study, driver orientation is categorized into radical type, common type, and conservative type in this paper, belonging to the SVM identification problem of the multiclass case. Using the Directed Acyclic Graph proposed by J. Platt to establish the recognition model of driver's propensity, the algorithm is based on one versus one, for k samples, comprising $k(k-1)/2$ nodes. Each node is one-versus-one classifier. Design classification function $f_{ij}(x)$ is used to discriminate the two classes i and j samples; if $f_{ij}(x) > 0$, then determine the samples x belong to i class. Obviously, $k = 3$ for the recognition of driver's propensity; the recognition model based on a directed acyclic graph is shown in Figure 9.

From the clustering data, experimental data samples of 30 pilots are randomly selected for model calibration.

From Figure 9, it needs to build 3 two-class classifiers for the sample sets; for samples provided, select the linear kernel function

$$K(x, x_i) = x \cdot x_i. \quad (10)$$

Considering the theory of support vector machine, the classification lines of radical type (type 1) and conservative type (type 3), radical type and common type (type 2), and common type and conservative type are

$$\begin{aligned} f_{13}(v) &= 283.84 - 3.11t = 0, \\ f_{12}(v) &= 473.04 - 5.84t = 0, \\ f_{23}(v) &= -381.85 + 3.82t = 0. \end{aligned} \quad (11)$$

The classification line formula defined is $f_{mn}(v) = k_{mn}v + b_k = 0$ $(m = 1, 2,\ n = 2, 3,\ k = 1, 2, 3)$, and the classification

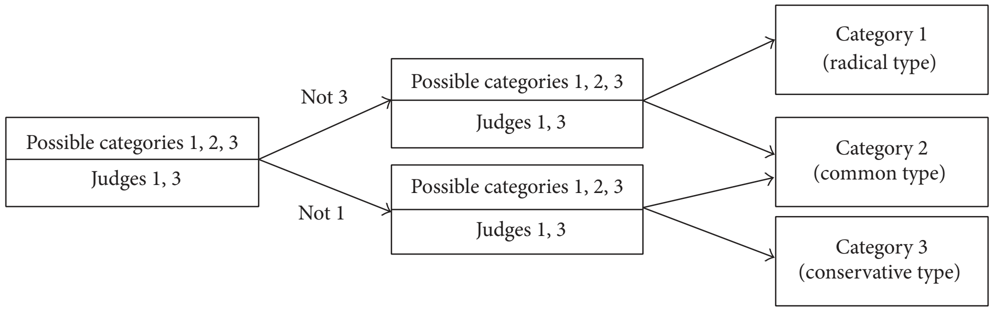


FIGURE 9: Recognition model of driver's propensity based on directed acyclic graph algorithm.

standard can be expressed as follows: if $f_{mn}(v) > 0$, the sample corresponding category is m, and if $f_{mn}(v) < 0$, the sample corresponding category is n. The recognition process corresponds to Figure 9, and the final samples can be obtained.

2.4. Model Verification. Select some experimental data corresponding to 20 drivers (different from the model calibration selected) for the model validation. Combined with the identification process, the classification line between the various types, and classification criteria established in Figure 9 to validate the driver's propensity model, some verified results are given from any one driver in Table 3.

The gaining process of percentage is as follows: use number 17 driver as an example to explain the different driving propensities' percentage calculation processes. Organize that number 17 driver does many effective tests and count travel time for each test, we can get the driver propensity by inputting the travel time to the driver propensity recognition model, assume that the driver propensity's result is radical driver, we can get radical sample rate by the number of radical driver experiments divided by the total number of experiments, and the other types' calculative processes are the same.

In the table $f_{12} = 0$ corresponding to the sample belongs to the support vector, which is located in the classification line. All data for the same experiment on the same experimental road of number 17 driver show the following: the ratio of sample belonging to the support vector is about 15%, identification results show that about 14% of the sample output is type 2 (common type), and the remaining 71% of the sample recognition results are type 1 (radical type). The driver's mental test score is 46, which belongs to the radical driver, consistent with the results of model verification. The above results show the driver's propensity type is mixed strategy, namely, a driver's propensity with a certain probability. The final validation results of 20 drivers are in Table 4, accuracy corresponding to percentage figures shows that each probability corresponds to a certain type.

Experiment results show that when only considering travel time and personal privacy protection, the dynamic recognition accuracy rate of driver's propensity is more than 70%. Model validation is basically consistent with the results of psychological tests.

TABLE 3: The validation results of part of real-time data of driver number 17.

Travel time (s)	f_{13}	f_{12}	Output result of model
74.42	>0	>0	Type 1 (radical type)
79.61	>0	>0	Type 1 (radical type)
79.76	>0	>0	Type 1 (radical type)
76.63	>0	>0	Type 1 (radical type)
76.06	>0	>0	Type 1 (radical type)
80.89	>0	>0	Type 1 (radical type)
82.44	>0	<0	Type 2 (common type)
72.79	>0	>0	Type 1 (radical type)
75.85	>0	>0	Type 1 (radical type)
73.87	>0	>0	Type 1 (radical type)
78.61	>0	>0	Type 1 (radical type)
75.01	>0	>0	Type 1 (radical type)
72.72	>0	>0	Type 1 (radical type)
85.99	>0	<0	Type 2 (common type)
74.93	>0	>0	Type 1 (radical type)
79.40	>0	>0	Type 1 (radical type)
80.63	>0	>0	Type 1 (radical type)
80.73	>0	>0	Type 1 (radical type)
67.86	>0	>0	Type 1 (radical type)
78.77	>0	>0	Type 1 (radical type)
78.38	>0	>0	Type 1 (radical type)
83.66	>0	<0	Type 2 (common type)
74.96	>0	>0	Type 1 (radical type)
80.02	>0	>0	Type 1 (radical type)
81.00	>0	=0	0
76.61	>0	>0	Type 1 (radical type)
77.45	>0	>0	Type 1 (radical type)
80.70	>0	>0	Type 1 (radical type)
...	...	...	...

Further increase the penetration rate to 25%, 30%, 35%, 40%, 45%, and 50% and, with a similar study, the correlation results obtained in Figure 10, Figure 11 (box plot), and Figure 12 (curve 1) show the following.

Obviously, in the case of different penetration rate, the accuracy of recognition is different about driver's propensity. When penetration increases from 20% to 40%, recognition

Table 4: Final validation results of 20 drivers.

Driver number	Model validation results (accuracy rate)	Psychological test results
36	Conservative type (71.6%)	76 (conservative type)
22	Common type (70.6%)	50 (common type)
7	Conservative type (70.3%)	72 (conservative type)
35	Conservative type (70.2%)	73 (conservative type)
25	Common type (71.7%)	53 (common type)
48	Common type (70.6%)	65 (common type)
49	Conservative type (70.3%)	80 (conservative type)
14	Radical type (71.6%)	40 (radical type)
2	Common type (70.6%)	61 (common type)
16	Conservative type (71.6%)	76 (conservative type)
24	Radical type (71.3%)	46 (radical type)
11	Common type (70.9%)	61 (common type)
21	Radical type (71.1%)	43 (radical type)
39	Common type (72.0%)	54 (common type)
10	Common type (71.2%)	53 (common type)
17	Radical type (71%)	46 (radical type)
45	Conservative type (70.5%)	75 (conservative type)
9	Conservative type (71.5%)	73 (conservative type)
5	Radical type (71.3%)	40 (radical type)
12	Common type (70.3%)	47 (common type)

accuracy rate is improved obviously, and when penetration increases from 40% to 50%, recognition accuracy rate is still improving but not obvious.

3. Conclusion

Identification on driver's propensity is extremely important for research on the car driver assistance, particularly active safety warning system. Through psychological testing, virtual driving test, and real vehicle test, different types of bias (such as aggressive, general, and conservative types) of driver dynamic sensing data is obtained. For respecting and protecting the driver's privacy, select and analyze driver dynamic travel time data. Real-time identification of the driver's propensity is built with application of pattern recognition method supporting vector machine model, and real vehicle test data is used for parameters calibration and validation of the model. Verification results and psychological questionnaire test results obtained are consistent. The results show that the proposed method can identify the driver's propensity dynamic driver tendentious deduction.

In this paper, the problem of dynamic recognition of driver's propensity is studied in the different penetration, as space is limited, 20% penetration rate is taken as an example to study issues in detail, recognition accuracy rate is about 70%, and there is inaccuracy error in the recognition results. The main reasons for the error may be the following: first, the early data collection is affected by instruments and equipment, which leads to late identification; second, the characteristic parameter reasons, only taking the driver's propensity dynamic travel time as characteristic parameter, may have certain influence on calibrating the late driver's propensity recognition model, thereby affecting the recognition accuracy rate; third, the impact of permeability and different penetration rates also have certain effect on data acquisition, thus affecting the propensity model calibration and recognition accuracy rate.

In addition, there are still shortcomings in this research: (1) the age, gender, preferences, and other factors of the driver are not considered, for example, even though, for radical type female driver, the radical female driver influenced by psychological and physiological conditions may provide travel time that is different from that provided by radical male drivers; (2) a section of Liantong Road is selected for the experiments, which does not reflect other specific sections of road conditions; (3) fewer environmental considerations: experiment is conducted in good weather conditions, the driver's propensity changes with environment changes are not considered. Of course, the impacts of environmental complexity for the driver are not only these. Other factors, such

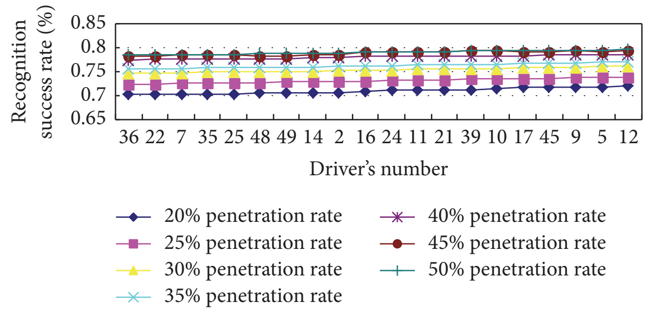


Figure 10: Recognition success rate with different penetration rate.

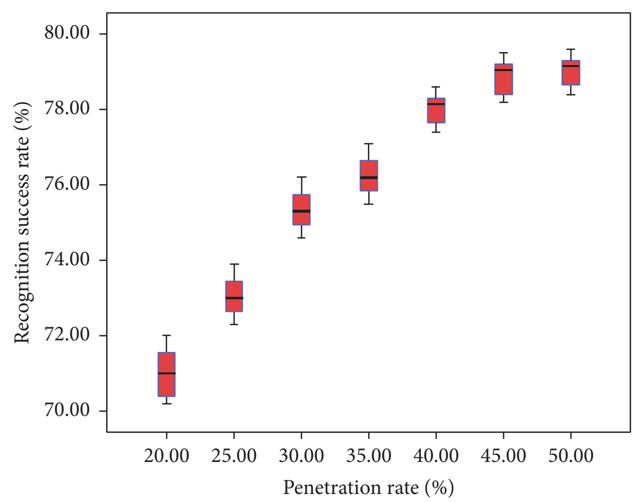


Figure 11: Recognition accuracy rate box plot trends under different penetration. Note that the middle of the black line is the median; the box is of 4 points' position (the bottom of the box is 25% quantile; the top of the box is 75% quantile), two extending lines represent extremes (the one below represents the minimum value; the one above represents the maximum value). By the median box plots trend, the improving trend of the recognition accuracy rate can be seen.

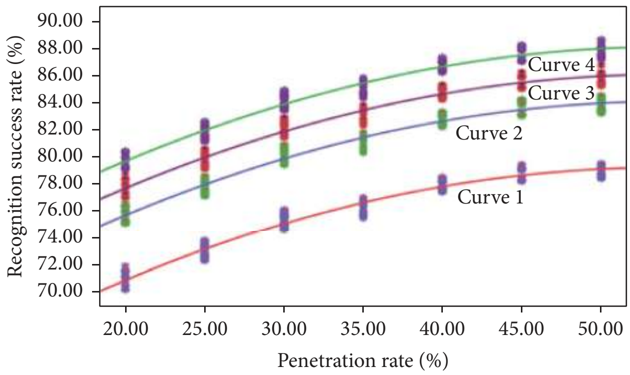


FIGURE 12: Recognition accuracy rate plot trends under different penetration.

as road capacity and service levels, proportion particularly the number of carts, and the sequential activities of moving targets in vehicle marshalling, will have impact on the driver's emotions, intentions, and behaviors to varying degrees, and further research is needed in these areas in future.

Future research can be done on this basis from the following aspects: improved driving dynamics tendency recognition algorithm, finely dividing types of driver's propensity; taking further identification studies by GPS other data and other factors (such as road conditions, age, gender, driving experience, and preferences) in the condition of respecting and protecting the privacy of the driver combined with travel time mentioned; based on GPS dynamic research and existing research findings, exploring car drivers tendentious mining law; deriving the intrinsic transfer law of driving tendency laws by GPS dynamic data law; and extending the dynamic identification on specific sections of the driver's propensity to the intersection.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study was supported by the State Key Laboratory of Automotive Safety and Energy under Project no. KF16232, the National Natural Science Foundation of China (Grants nos. 61074140, 61573009, and 51508315), Natural Science Foundation of Shandong Province (Grant no. ZR2014FM027), Social Science Planning Project of Shandong Province (Grant no. 14CGLJ27), and Project of Shandong Province Higher Educational Science and Technology Program (Grant no. J15LB07).

References

[1] X. Feng and C. Fang, "Cluster analysis of driver characteristics evaluating indicator," *Transportation Science and Technology of Heilongjiang*, vol. 165, no. 11, pp. 161–163, 2007.

[2] L. Xie and Z. Wang, "Driver's emotion model under simplied traffic pattern," *Acta Automatica Sinica*, vol. 36, no. 12, pp. 1732–1743, 2010.

[3] C. Wu and L. Hu, "Review on the study of Motorists' driving anger," *China Safety Science Journal*, vol. 20, no. 7, pp. 3–8, 2010.

[4] L. Hu, *Characteristics of Angry Driving Behaviors and Its Effects on Traffic Safety*, Wuhan University of Technology, Wuhan, China, 2011.

[5] Q. Lin and S. Feng, "Prevent aggressive driving from affective computing," *Lightweight Vehicle Technology*, vol. 278, no. 10, pp. 31–33, 2012.

[6] L. L. Di Stasi, D. Contreras, J. J. Cañas, A. Cándido, A. Maldonado, and A. Catena, "The consequences of unexpected emotional sounds on driving behaviour in risky situations," *Safety Science*, vol. 48, no. 10, pp. 1463–1468, 2010.

[7] L. M. Trick, S. Brandigampola, and J. T. Enns, "How fleeting emotions affect hazard perception and steering while driving: the impact of image arousal and valence," *Accident Analysis & Prevention*, vol. 45, pp. 222–229, 2012.

[8] O. Taubman-Ben-Ari, "The effects of positive emotion priming on self-reported reckless driving," *Accident Analysis & Prevention*, vol. 45, pp. 718–725, 2012.

[9] K. M'bailara, T. Atzeni, B. Contrand, E. Lagarde, and C. Galéra, "Emotional status and responsibility in road traffic accident," *European Psychiatry*, vol. 9, no. 28, pp. 252–260, 2013.

[10] Z. Jianyou, Z. Rui, Z. Yang, and H. Longfei, "Analysis of driver's adaptability of stress response in dangerous situation," *Procedia Social and Behavioral Sciences*, vol. 96, pp. 2173–2178, 2013.

[11] L. Arnau-Sabatés, J. Sala-Roca, and M. Jariot-Garcia, "Emotional abilities as predictors of risky driving behavior among a cohort of middle aged drivers," *Accident Analysis & Prevention*, vol. 45, pp. 818–825, 2012.

[12] M. Chan and A. Singhal, "The emotional side of cognitive distraction: implications for road safety," *Accident Analysis & Prevention*, vol. 50, pp. 147–154, 2013.

[13] X. Wang and J. Zhang, "Extraction and recognition methods of vehicle driving tendency feature based on driver-vehicle-environment dynamic data under car following," *International Journal of Computational Intelligence Systems*, vol. 4, no. 6, pp. 1269–1281, 2011.

[14] J. Zhang, X. Wang, X. Ban, and K. Cao, "Prediction method of driver's propensity adapted to driver's dynamic feature extraction of affection," *Advances in Mechanical Engineering*, vol. 5, Article ID 658103, 2013.

[15] V. N. Vapnik, *The Nature of Statistical Learning*, Springer, New York, NY, USA, 1995.

[16] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[17] X. G. Zhang, "Introduction to statistical learning theory and support vector machines," *Acta Automatica Sinica*, vol. 26, no. 1, pp. 32–42, 2000.

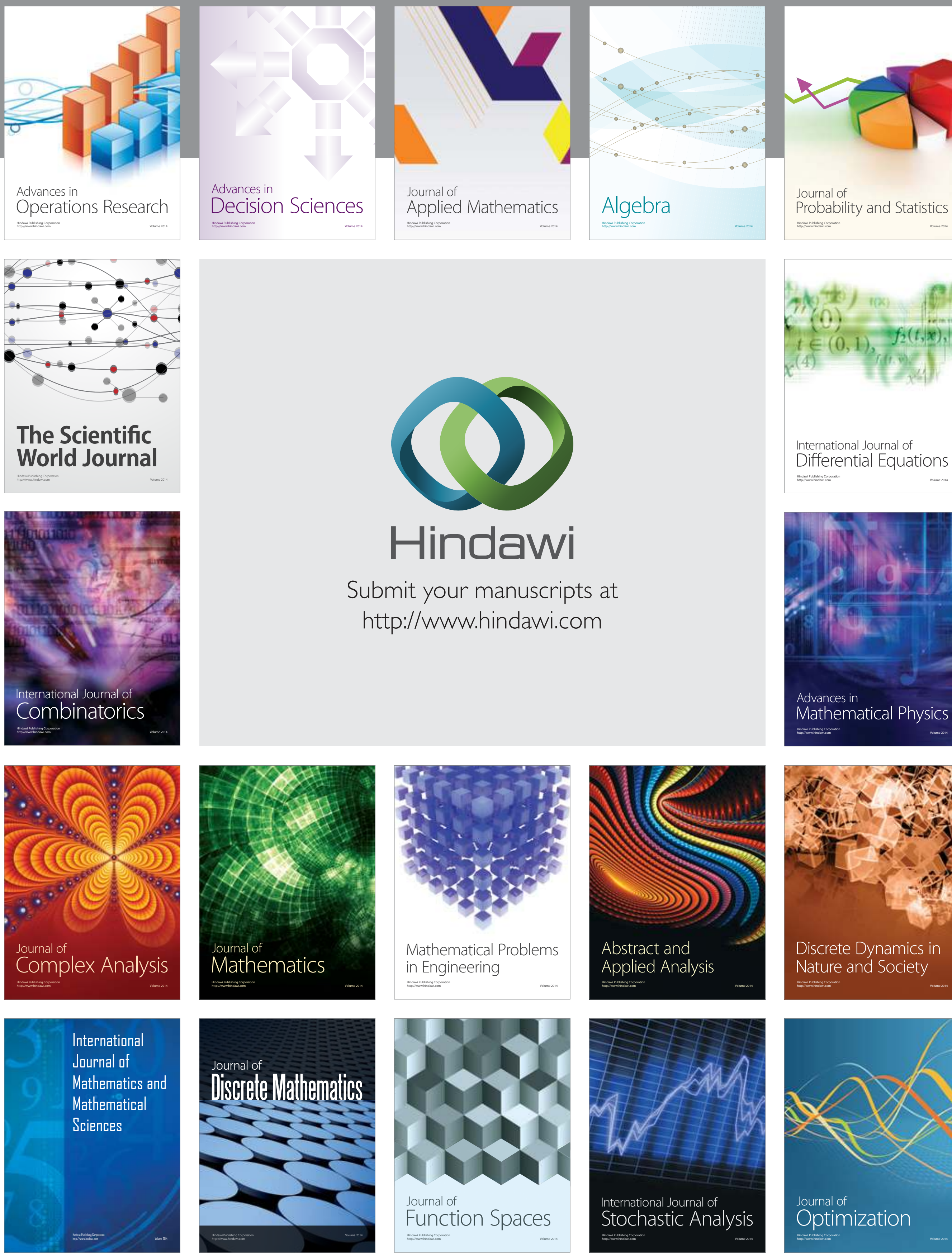
[18] Q.-H. Xu and J. Shi, "Aero-engine fault diagnosis based on support vector machine," *Journal of Aerospace Power*, vol. 20, no. 2, pp. 298–302, 2005.

[19] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[20] V. Vapnik, "Principles of risk minimization for learning theory," in *Advances in Neural Information Processing Systems*, pp. 831–838, 1991.

[21] L. Oneto, S. Ridella, and D. Anguita, "Tikhonov, Ivanov and Morozov regularization for support vector machine learning," *Machine Learning*, vol. 103, no. 1, pp. 103–136, 2016.

[22] D. Anguita, A. Ghio, L. Oneto, and S. Ridella, "In-sample and out-of-sample model selection and error estimation for support vector machines," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 9, pp. 1390–1406, 2012.


Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
Algebra
Journal of Probability and Statistics
The Scientific World Journal
Hindawi
Submit your manuscripts at
http://www.hindawi.com
International Journal of Differential Equations
International Journal of Combinatorics
Advances in Mathematical Physics
Journal of Complex Analysis
Journal of Mathematics
Mathematical Problems in Engineering
Abstract and Applied Analysis
Discrete Dynamics in Nature and Society
International Journal of Mathematics and Mathematical Sciences
Journal of Discrete Mathematics
Journal of Function Spaces
International Journal of Stochastic Analysis
Journal of Optimization