Urban Traffic Incident Detection with Mobile Sensors Based on SVM

Bin Pan* (1)(2)(4), Hao Wu (1)(3)(4)
(1) State Key Laboratory of Rail Control and Safety, Beijing Jiaotong University, Beijing, China, 100044
(2) School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China, 100044
(3) Beijing Key laboratory of intelligent traffic data safety and privacy protection technology, Beijing Jiaotong University, Beijing, China, 100044
(4) Beijing Engineering Research Center of High-speed Railway Broadband Mobile Communications, Beijing Jiaotong University, Beijing, China, 100044

Abstract

Traffic accident detection is an important component in Intelligent Transportation System (ITS). Compared with freeway, urban accident detection is more complicated. The mean reason is that there are full of flow-disruptive entities (traffic signals, intersections and bus stops) in urban roads, which can disrupt the traffic flow in a similar way that an accident occurs. In this paper, an effective method for urban traffic accident detection has been designed to recognize the abnormal traffic state based on mobile sensors and Support Vector Machine (SVM). We regard the whole dynamic process of the vehicle passing through a road as an instance. Our results show that SVM is a good method to detect the urban traffic accident. In addition, we investigate three traffic variables (speed, acceleration and lane-changing) and their effect for urban traffic accident detection.

1. Introduction

In urban environment, frequent traffic accidents have affected the development of society seriously. The research on traffic incident detection is of great significance to improve the efficiency of road and ensure traffic safety. However, most of the research currently focused on the freeway. As an important component, urban roads have not get enough attention.

Nowadays, most of traffic data used in detecting incidents are still collected from fixed sensors, such as inductive loop detectors and cameras. The coverage of these sensors is extremely limited, and it is difficult to obtain traffic data beyond the coverage. With the development of vehicular ad hoc networks (VANET), each vehicle will equip with mobile sensors (such as smartphones [1][2], speed sensors, etc.) and On Board Units (OBUs). These mobile sensors will collect traffic data continuously, such as vehicle speed, acceleration and lane-changing state. These traffic data collected by mobile sensors will be sent to Road Side Units (RSUs) by OBU. Then RSUs forward the traffic data to central server for data processing. Vehicles can run anywhere in a road network theoretically, so on-board mobile sensors can provide a greater range of monitoring.

Compared with freeway, traffic accident detection in urban road will meet several challenges. First, the urban roads are uncontrolled-access roads. Different from freeway, there are no fixed entrances and exits in urban roads and the length of urban roads are not very long. In addition, urban roads have many entities, such as traffic signals, intersections and bus stops. These entities will interfere or block traffic flow, affecting the normal operation of the vehicles. So it will cause some interference to urban traffic accident detection, and it is more difficult to detect the traffic accident in the urban environment than in the freeway environment. Second, the traffic data collected by mobile sensors are not necessarily effective for urban traffic accident detection. Therefore, it is necessary to select the variables which are more accurate for the urban accident detection [3].

The rest of this paper is organized as follows. Section 2 discusses the related works on the urban traffic accident detection with mobile sensors. Section 3 describes our experiment in detail, which includes scenario assuming, parameter setting and data processing. Results are discussed in section 4. Finally, Section 5 summarizes the whole paper and gives a conclusion.

2. Related Work

Automatic incident detection (AID) is an important method for detecting potential traffic accidents or accidents occurred rapidly. With the development of control technology, computer technology and information technology, the number of AID algorithm has reached more than 20. In [4], the authors use machine learning methods to analyze information collected by vehicles to detect forward collisions in freeway control system. Based on fuzzy logic, a new traffic incident detection algorithm for freeway is provided in [5]. In [6], the authors propose a fast and effective approach to automatically detect traffic accident in a video. In [3], authors use the Kolmogorov-Smirnov test to evaluate whether each of the four traffic variables (speed, acceleration, lane-change ratio and travel time) has significantly different statistical characteristics in the presence of an incident. The research above all share a common feature, which is to identify the changed

variables in accident compared with the normal situation. For example, the most common traffic variable used in traffic accident detection is vehicle speed. When a traffic accident occurs, the common behavior of vehicles is to reduce the speed.

Most of the previous studies focused on the freeway accident detection using fixed sensors [7][8]. Of course, there are still some researchers obtaining traffic data by mobile sensors currently [9][10]. However, few researchers have studied the traffic accident detection in urban scenarios. In [3], the authors use the statistical method to analyze the effectiveness of several traffic variables in urban traffic accident detection.

This paper will use the machine learning method (SVM), to compare the accuracy rate of different traffic variables in urban accident detection. The traffic variables include vehicle speed, acceleration and lane-changing state.

3. Methodology

3.1 Scenario and Assumptions

In this study, we assume that every vehicle can continuously collect their own real-time traffic data by on-board sensors. These data include identity, location, time, speed, acceleration, lane-changing state and other information. OBU broadcasts the traffic data that can reflect traffic conditions periodically. These traffic data are collected by RSU and uploaded to the central server for processing. In these data, we are concerned about the vehicle speed, acceleration and lane-changing state. In this study, we only consider the influence of traffic lights in urban roads.

3.2 Traffic variables

In this paper, we consider three traffic variables for urban traffic accident detection. The three traffic variables and their traffic characteristics are as follows.

- **Vehicle speed**: When the vehicle travels near the accident point position, it will have a process of slowing down or even stopping until the vehicle passes through the accident point. After that, the vehicle will accelerate to its normal speed. This would be quite easy to observe if an incident occurred on a freeway. However, the vehicle will slow down or even stop at the traffic lights (red light) in the urban road. These traffic controls can cause flow disruption in a similar way that an incident would cause.

- **Vehicle acceleration**: When the vehicle travels near the accident point, it will have a relatively large deceleration. Having passed through the accident point, it will have a relatively large acceleration. However, when the vehicle comes to traffic lights (red light), it will also have a deceleration and acceleration process. It is not easy to distinguish the impact of traffic lights or traffic accidents.

- **Lane-changing**: If there is a traffic accident, some vehicles will choose to change lanes to pass through the accident point. After passing through the accident point, the vehicle will choose to change to other lanes if the current lane is congested. However, if the vehicle comes to the traffic lights (red light), it will also execute the lane-changing operation and to find a parking space. In order to quantify the lane-changing state, in this study, we use the value "0" to represent no lane-changing, the value "1" to represent changing to the left lane, the value "2" to represent changing to the right lane.

3.3 Experiment

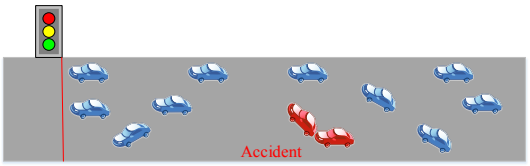

Figure.1 Accident scenario in urban road

In this study, we assume a three-lanes urban road scenario with a traffic light as shown in Fig.1. We adopt a microscopic traffic simulator VISSIM, in order to describe the mobility of the vehicles more carefully. The normal speed of vehicles are from 48km/h to 58km/h. Near the accident point, we perform the deceleration operation to the vehicles, which speed are set from 20km/h to 25km/h. Traffic lights cycle time is set to 60s, and red time and green time are 30s respectively. In order to facilitate the subsequent work for data processing, we select the 400m section, which includes an accident points and traffic lights for analysis. Assuming that the accident vehicles occupy two lanes, there is only one lane left for other vehicles' running. The volume of vehicles (veh/h) in the network is set to 600, 800 and 1000. The traffic data that we need, which is speed, acceleration and lane-changing state, are recorded with 1-second sampling period.

3.4 Data processing

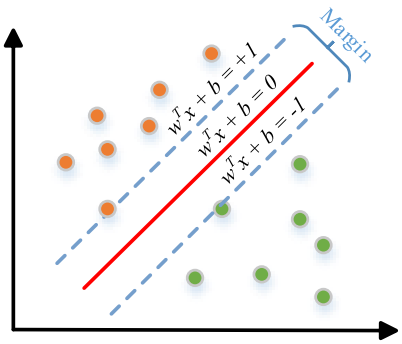

Figure.2 Support Vector and Margin

SVM is a relatively new machine learning pattern classifier model that uses the samples of "training data" to define an optimal boundary between classes. Training vectors are chosen to lie closest to the class boundary and are called support vectors. Given a training set of instance-label pairs (x_i, y_i), i = 1,…,n, y_i is either -1 for a non-incident and 1 for an incident and indicates the class to which the point x_i belongs [11].

The general mathematical form of the linear SVM is $f(x) = w^T x + b$, w is the normal vector of the hyperplane, and b is a variable (Fig.2). The linear SVM function finds the two closest points in each data set and creates a hyperplane between them to separate the two data sets. SVM achieves nonlinear classification by mapping the input vectors into a higher-dimensional feature space through the kernel function ϕ until the data becomes linear. A method called Soft Margin is used to split the classes to solve the problem that we can't found a hyperplane to separate the two classes. The problem of finding the optimal hyperplane can be turned into an optimization problem below, which is a trade-off between a large margin and a small error penalty.

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \\ &s.t.\ y_i\left(w^T\phi(x_i)+b\right) \geq 1-\xi_i \\ &\xi_i \geq 0, i=1,2,...,l \end{aligned} \tag{1}$$

Where ξ_i and C (error penalty) are the nonnegative slack variables, which measure the degree of misclassification of the data, x_i is the training/testing vector (input pattern), y_i indicates the class to which the point x_i belongs, and ϕ is a projection function from lower-dimension space into higher-dimension space.

And the most commonly used measurement for classification is the accuracy rate (AR):

$$AR = \frac{number\ of\ correctly\ classified\ cases}{total\ number\ of\ cases} * 100\% \tag{2}$$

In this study, given a train set of instance-label pairs (X_i, Y_i), i =1,…,n, n is the number of vehicles of one instance. In order to record the data of vehicles dynamically in the road, for a certain traffic variable, we consider the data x_{ij} generated per second as a characteristic value, starting from the vehicle comes into the road entrance. Where i=1,…,n, n is the number of vehicles, j=1,…,m, m is the total time x_i is in the section of urban road. All the characteristic values of a vehicle are considered as an instance X_i. Y_i is either -1 for a non-incident and 1 for an incident. So we can get a train set of instance-label pairs (X_i, Y_i), also named a sample. Since the total time of each vehicle traveling in the road is not the same, the number of characteristic values of each vehicle are not the same. We take the maximum number of characteristic values for all vehicles as standard. If the number of characteristic values for other vehicles is less than the standard, we add "0" in the final, until the length of the values is as long as the standard. Therefore, the number of characteristic values for all vehicles are the same.

4. Result and Discussion

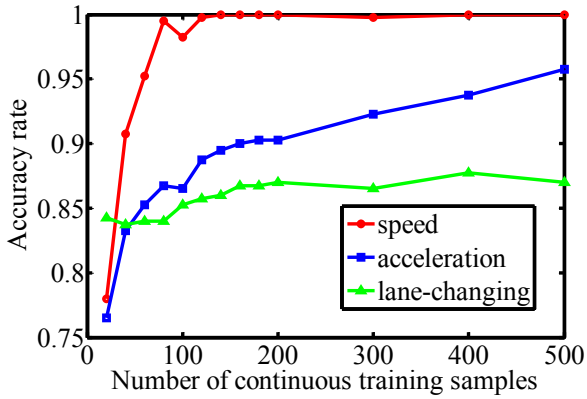

Figure 3. Accuracy rate for different variables

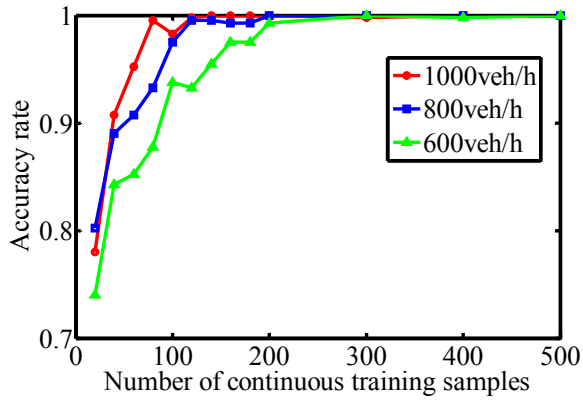

Figure 4. Accuracy rate for speed with different traffic volumes

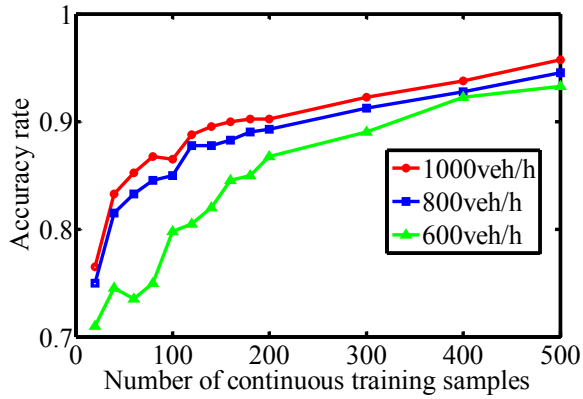

Figure 5. Accuracy rate for acceleration with different traffic volumes

There are already many software package for SVM. In this study, we use the LIBSVM designed by professor Lin Chih-Jen. In the simulation, we choose different numbers of continuous training samples, in which the samples of accident and normal situation have equal weighting, to validate the effect of different numbers of training samples on the accuracy rate. We select 400 vehicles as the testing samples, in which the samples of accident and

normal situation have 200 vehicles respectively, to ensure that the total time of these vehicles arriving traffic lights continuously could cover several traffic light cycles. We can get a model by training the training samples in LIBSVM. Input different testing samples to the model above, we can obtain the corresponding output. Finally, we can get the accuracy rate by (2) above.

Fig.3 shows that vehicle speed is more effective than the other variables (acceleration and lane-changing). The accuracy rate almost closes to 100%, when the number of continuous training samples is about 100. But for acceleration, its growth rate for accuracy rate is not as fast as speed with the increase of the number of continuous training samples. When we make use of 500 train samples, the accuracy rate is about 95%. In addition, using the lane-changing variables has the worst result. Although the accuracy rate of lane-changing variable is higher than the other two variables when the number of samples is small, the accuracy rate has not increased too much with the increasing number of training samples. The accuracy rate has been maintained at about 85%.

Fig.4 shows the variation of the accuracy rate with different volumes of vehicles for speed variable. As we can see, the accuracy rate is increasing with the increase of traffic volumes. The main reason is that when the traffic volume is small, the vehicles can pass through the incident point without slowing down significantly. However, when the traffic volume is high relatively, the vehicles will slow down to a great extent or even stop. The effect of traffic volumes on the accuracy rate is very small, when the number of training samples is more than 200. This is because the enough samples will provide various characteristics values in different cases for the accident detection.

Fig.5 shows the variation of the accuracy rate with different volumes of vehicles for the acceleration variable. Like the conclusion of fig.4, the accuracy rate is increasing with the increase of traffic volume. However, the accuracy rates are still different when the number of training samples reaches 500. Therefore, in other words, more samples are needed for training in order to achieve the ideal accuracy rate.

5. Conclusion

In this paper, a traffic accident detection method has been designed to recognize the abnormal traffic state based on mobile sensors and SVM. The related real-time data are collected by RSU in VANET and then sent to the center server for data processing. For every vehicle, we treat a data value collected per second for some traffic variable in the road as a characteristic value. Then all of the characteristic values constitute an instance. Added with a label, we can get a sample. Our results show that SVM is a good method to detect the urban traffic accident, though the urban traffic detection is more difficult than the accident detection in freeway.

In addition, we investigate three traffic variables (speed, acceleration and lane-changing state) and their effect for urban traffic accident detection. The most effective variable is speed and the least effective variable is lane-changing state. In addition, the accuracy rate of accident detection with three variables respectively is similar, when the training samples are enough.

6. Acknowledgements

This paper is supported by New Century Excellent Talents in University (No. NCET-13-0657) and the State Key Laboratory of Rail Traffic Control and Safety Contract No.RCS2016ZT015, Beijing Jiaotong University. And it is supported in part by the Chinese National Programs for High Technology Research and Development 863 project (No.2015AA016005).

7. References

[1] A. Kovacheva, R. Frank, and T. Engel, “Luxtraffic: A collaborative traffic sensing system,” in *Local and Metropolitan Area Networks*, 2013, pp.1–6.

[2] S. Panichpapiboon and P. Leakkaw, “Traffic sensing through accelerometers,”*IEEE Transactions on Vehicular Technology*, pp. 1–1, 2015.

[3] C. Siripanpornchana, S. Panichpapiboon, and P. Chaovalit, “Effective variables for urban traffic incident detection,” in *Vehicular NETWORKING Conference*, 2015, pp. 190–195.

[4] D. Nejdet, “Traffic accident detection by using machine learning methods,” *Faculty of Economics Management Department*, 2012.

[5] B. Xie, Z. Hu, and H. Ma, “Fuzzy-logic-based traffic incident detection algorithm for freeway,” in *International Conference on Machine Learning and Cybernetics*, 2008, pp. 1254–1259.

[6] S. Xia, J. Xiong, Y. Liu, and G. Li, “Vision-based traffic accident detection using matrix approximation,” in *Control Conference*, 2015, pp. 1–5.

[7] V. L. Knoop, H. J. V. Zuylen, and S. P. Hoogendoorn, *Microscopic Traffic Behaviour near Incidents*, pp. 75-97, 2009.

[8] H. Teng and Y. Qi, “Detection-delay-based freeway incident detection algorithms,” *Transportation Research Part C Emerging Technologies*, vol. 11, no. 3, pp. 265–287, 2003.

[9] Q. Wang, H. Yu, T. Zhu, and G. Li, “The impact analysis of traffic incident and prediction model on travel time under incident condition,”in *International Conference on ITS Telecommunications*, 2013, pp. 32–37.

[10] V. C. Dang, M. Kubo, H. Sato, and A. Yamaguchi, “A simple braking model for detecting incidents locations by smartphones,” 2014, pp. 1–5.

[11] M. Motamed, “Developing A Real-time Freeway Incident Detection Model using Machine Learning Techniques.” Ph.D. dissertation, The University of Texas at Austin.