Accident Analysis and Prevention 94 (2016) 59–64

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Utilizing the eigenvectors of freeway loop data spatiotemporal schematic for real time crash prediction

Shou'en Fang[a], Wenjing Xie[a], Junhua Wang[a,*], David R. Ragland[b]

[a] School of Transportation Engineering, Tongji University, 4800, Cao'an Highway, Shanghai, China
[b] School of Public Health, UC Berkeley, Safe Transportation Research and Education Center (SafeTREC), University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 3 September 2015
Received in revised form 7 March 2016
Accepted 12 May 2016
Available online 31 May 2016

Keywords:
Real time accident model
Crash precursor
Spatiotemporal schematics
Road safety

ABSTRACT

The concept of crash precursor identification is gaining more practicality due to the recent advancements in Advanced Transportation Management and Information Systems. Investigating the shortcomings of the existing models, this paper proposes a new method to model the real time crash likelihood based on loop data through schematic eigenvectors. Firstly, traffic volume, occupancy and density spatiotemporal schematics in certain duration before an accident occurrence were constructed to describe the traffic flow status. Secondly, eigenvectors and eigenvalues of the spatiotemporal schematics were extracted to represent traffic volume, occupancy and density situation before the crash occurrence. Thirdly, by setting the vectors in crash time as case and those at crash free time as control, a logistic model is constructed to identify the crash precursors. Results show that both the eigenvectors and eigenvalues can significantly impact the accident likelihood compared to the previous study, the proposed model has the advantage of avoiding multicollinearity, better reflection of the overall traffic flow status before the crash, and improving missing data problem of loop detectors.

1. Introduction

Recent advances in Intelligent Transportation System (ITS) allow traffic safety studies to extend from historic data-based analyses to real-time applications. This study employs the real-time capability of Advanced Traffic Management and Information Systems (ATMIS) for safety enhancement. In the last 20 years, researchers have developed crash prediction models to relate crash risks to some real-time traffic flow parameters collected from loop detectors such as traffic occupancy and vehicle speed variances. Madanat and Liu (1995) proposed the concept of "Real-time Incident Likelihood Prediction" to analyze the crash similarity by modelling the crash data on freeway with the detected flow data and weather data. Hughes and Council (1999) were the first to use loop detector data to explore the relationship between freeway safety and operations during peak periods. They proposed that real time detected speed variance was significant in crash occurrence and macroscopic measures were not suitable for real time safety analysis.

Since then, quite a few research studies followed Oh et al. (2001) employed a Non-Parametric Bayesian Classification Model to estimate accident likelihood. This study demonstrated that speed variation detected 5 min before accident could be used to estimate an accident effectively by using loop detector data and by setting different threshold level different percentage of accidents can be identified. Lee et al. (2002) proposed the concept "Crash Precursor" for traffic flow characteristics observed prior to crash occurrence. An aggregate log-linear model was developed which proved that variation of speed and traffic density were statistically significant predictors of crash frequency. In the coming year, Lee et al. (2003) developed a probabilistic real-time crash prediction model with a rational method to select crash precursors and optimal observation time slice durations. The study found that the difference between the speed at the upstream detector and the speed at the downstream detector was significantly higher.

Abdel-Aty et al. (2004) proposed that 5 min speed data before the accident was too short to be applied to real time traffic management. They utilized data from 7 loop stations which were closest to the accident site. Meanwhile, 6 periods of 5 min data of each loop before the crash were collected and case-control logistic regression method was used, and the model could achieve more than 69% crash identification. Abdel-Aty et al. (2005) and Abdel-Aty and Pande (2005) applied Bayesian classifier based methodology and probabilistic neural network (PNN) to predict crash likelihood. This study suggested that coefficient of variation measured in 5 min time slices of 10–15 min prior to the crash time in some upstream loops

* Corresponding author.
E-mail addresses: fangsek@tongji.edu.cn (S. Fang), 100931@tongji.edu.cn (W. Xie), benwjh@163.com (J. Wang), davidr@berkeley.edu (D.R. Ragland).

http://dx.doi.org/10.1016/j.aap.2016.05.013
0001-4575/

affected crash occurrence most significantly. Abdel-Aty et al. (2005) suggested that utilizing hazard ratio values about 15 min before the accident occurrence could effectively reduce the impending risk. His method analyzed the flow data 90 min before and after the accident occurrence for four loop detectors upstream and two loop detectors downstream. Pande and Abdel-Aty (2006) also used Kohonen clustering algorithm, classification tree, multilayer perceptron (MLP), and normalized radial basis function (NRBF) neural networks to achieve 75% prediction of rear-end crashes.

Golob et al. (2004) used cluster analyses to prove that mean volume, median speed and temporal variations in volume and speed are the key traffic flow elements affecting safety. Golob et al. (2008) also tried to use a single loop data to evaluate safety performance. He converted the 36 redundant traffic flow parameters into 8 factors to develop a logit model for accident severity, type of collision, collision lane and number of vehicles involved. Oh et al. (2005) employed a Nonparametric Bayesian estimated accident likelihood, although they used only 52 accident samples and 5 min standard deviation of speed as indicators. Xu et al. (2012) divided the traffic flow into 5 states, and evaluated the safety performance associated with each state. This helps the safety researches better understand the traffic flow state before the crash. Vlahogianni et al. (2012) also developed neural network models by analyzing different variables associated with secondary accident likelihood and suggested that traffic speed/duration of the primary accident, hourly volume, rainfall intensity and number of vehicles involved in the primary accident are the top associated factors. Hossain and Muromachi (2012) pointed out some of the major shortcomings of the previous models including location of detectors, variable space and modelling methods which result in the implemented scenario being impractical. And he addressed the aforementioned shortcomings by proposing a Bayesian belief net based framework to develop real-time crash prediction models.

AVI (Automatic Vehicle Identification) data had been utilized in trial to analyze the crash potential in recent years (Ahmed et al., 2012; Ahmed and Abdel-Aty, 2012). These studies found that AVI systems could provide a measure of the risk of a crash in real time, with the accuracy 75.93% for rear-end crashes. Moreover, when fusing the AVI data and RTMS (Remote Traffic Microwave Sensors) data, the model with the data fusion framework had a higher estimation accuracy, robustness and reliability.

Meanwhile, transferability and robustness of real-time freeway crash risk assessment were emphasized in recent works (Shew et al., 2013). Xu et al. (2014) developed a Bayesian updating approach with data from the I-880N freeway in 2002 and 2009 and from the I-5N freeway in 2009 to improve transferability, and found it effective.

Generally, these findings point to the potential use of detector data and crash data in the field of traffic safety. However, none of these models have been implemented in practical scenario so far, which may result from the some of the major shortcomings of the existing models as follows.

(1) Most of the previous research took 5 min volume, speed and occupancy data as variable in Logit model. However, the upstream loop data in the earlier period and downstream loop data in the later period are usually correlated. The correlation coefficients between loops next to each other are shown in Table 1. High correlation between nearby loop data was found. Also, as the time slice gets longer, the relativity gets severer. We believe that the correlation between parameters influent the models estimates. This can cause multicollinearity.
(2) Large bunch of loop data in the series stations fail to explain the likelihood while just some of the loop data variables are significant which is too abstract to directly reflect the traffic flow status. Thus it is hard to be applied in the real traffic management.
(3) Most of the accident precursor study utilized 6 loop detectors (4 loops upstream and 2 loops downstream) or 7 loop detectors (the nearest detector to the accident with the other 4 loops upstream and 2 loops downstream) (Abdel-Aty et al., 2005). This kind of dataset is not necessary the best choice to reflect traffic follow characteristics as accident precursors since none of the research showed that the speed, volume and occupancy of the 1st and 6th loop significantly impacted crash likelihood.
(4) Most of the previous study utilized the single slice time period data before the accident occurrence as the precursor in modelling. While the crash record time in the accident database is always not so accurate that the significant variable in certain time slice in one location may not reflect the real traffic situation before the accident happens.

This study addresses the aforementioned shortcomings by proposing a freeway loop data schematic eigenvectors based framework to develop real-time crash prediction models. Firstly, traffic volume, occupancy and density spatiotemporal schematics in certain duration before accident occurrence were constructed to describe the traffic flow status. Secondly, the eigenvectors and eigenvalues of the spatiotemporal schematics were extracted to represent traffic volume, occupancy and density situation before the crash occurrence. Thirdly, by setting the vectors in crash time as case and those at crash free time as control, a logistic model is constructed to identify the crash precursors.

2. Methodology

2.1. Step1 construction of spatiotemporal loop data schematic

Spatiotemporal loop data schematic is proposed to contain the traffic flow parameters of the turbulence stage prior to the crash time. In this study, a 6×6 matrix which focuses on both upstream and downstream section of crash location and contains 6 time slices (5 min or 2 min each) is proposed. Tables 2 and 3 show examples of spatiotemporal traffic flow schematics for 5-min data and 2-min data. Due to the eigenvector methodology proposed in this study, the spatio-temporal matrix should be square matrix. Then the time period cover by the matrix is depended on the number of loop detectors.

Abdel-Aty et al. (2004) suggested that it was not enough to take 5 min ahead of the accident if it came to real-time traffic management. In this research, 10 min was taken as the time offset prior to the accident occurrence. Therefore the time duration in the 6×6 schematic is 10–40 min before the crash. Additionally, it should be mentioned that loop detectors In California are installed approximately between 500 m and 800 m apart.

2.2. Step2 extraction of matrix eigenvectors and eigenvalues from the schematics

In linear algebra, an eigenvector of a square matrix is a vector that points in a direction which is invariant under the associated linear transformation. As eigenvectors can represent the character of the matrix, it is widely used as pattern recognition tool in numerous fields. Taking image recognition as a typical example, eigenvectors are used to represent the image pixel matrix, which is similar to the proposed method in this paper as follows. Eigenvectors of a square matrix represent the space characters in different directions. And eigenvalues is the degree of the stretch along the according direction. These figures are independent with each other from the

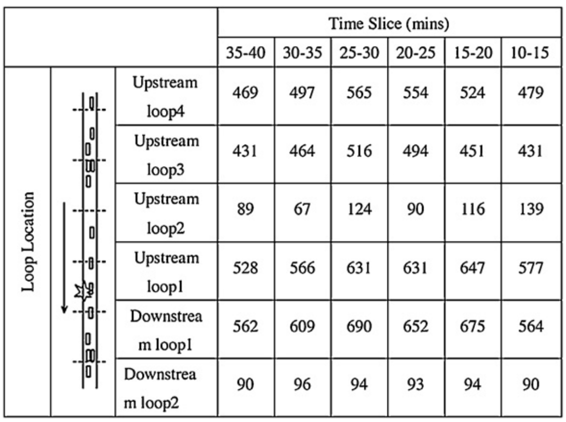
Table 1
correlation coefficients of traffic volume between the nearby loop data.

Loop ID / Time slice	601205 & 601206	601206 & 601207	601207 & 601208	601208 & 601209	601209 & 601210	601210 & 601211	601211 & 601212	601212 & 601213	601213 & 601214	601214 & 601215
30 s	76.43%	86.62%	70.68%	68.07%	79.42%	64.19%	74.78%	75.94%	84.44%	87.56%
1 min	89.00%	94.39%	85.97%	80.46%	87.78%	78.76%	85.26%	86.35%	92.52%	95.22%
2 min	85.43%	95.18%	84.60%	82.09%	86.74%	76.58%	82.73%	88.06%	93.55%	96.08%
5 min	95.85%	98.51%	95.43%	92.75%	94.12%	90.65%	91.78%	94.79%	98.38%	98.91%

Note: Samples used for correlation test is from I80, California, January 2010.

Table 2
6×6 spatiotemporal traffic flow matrix with 5-min data.

Loop Location	Time Slice (min) 35–40	30–35	25–30	20–25	15–20	Oct-15
Upstream loop4	469	497	565	554	524	479
Upstream loop3	431	464	516	494	451	431
Upstream loop2	89	67	124	90	116	139
Upstream loop1	528	566	631	631	647	577
Downstream loop1	562	609	690	652	675	564
Downstream loop2	90	96	94	93	94	90

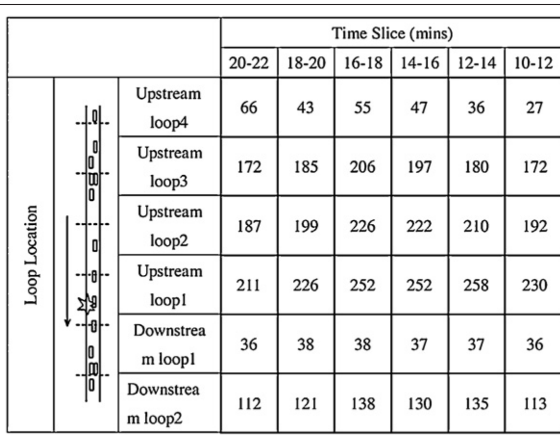


Loop Location	Time Slice (mins) 35-40	30-35	25-30	20-25	15-20	10-15
Upstream loop4	469	497	565	554	524	479
Upstream loop3	431	464	516	494	451	431
Upstream loop2	89	67	124	90	116	139
Upstream loop1	528	566	631	631	647	577
Downstrea m loop1	562	609	690	652	675	564
Downstrea m loop2	90	96	94	93	94	90

Note: Traffic flow is counted by unit cars.

Table 3
6×6 spatiotemporal traffic flow matrix with 2-min data.

Loop Location	Time Slice (min) 20–22	18–20	16–18	14–16	Dec-14	10-Dec
Upstream loop4	66	43	55	47	36	27
Upstream loop3	172	185	206	197	180	172
Upstream loop2	187	199	226	222	210	192
Upstream loop1	211	226	252	252	258	230
Downstream loop1	36	38	38	37	37	36
Downstream loop2	112	121	138	130	135	113



Loop Location	Time Slice (mins) 20-22	18-20	16-18	14-16	12-14	10-12
Upstream loop4	66	43	55	47	36	27
Upstream loop3	172	185	206	197	180	172
Upstream loop2	187	199	226	222	210	192
Upstream loop1	211	226	252	252	258	230
Downstrea m loop1	36	38	38	37	37	36
Downstrea m loop2	112	121	138	130	135	113

Note: Traffic flow is counted by unit cars.

view of linear algebra. With this advantage, the multicollinearity between loop data can be well avoided.

The spatiotemporal schematic was designed as square matrix so that a suit of eigenvectors and corresponding eigenvalues of the matrix can be easily extracted. The following equation defines eigenvectors and eigenvalues,

$$Ax = \lambda x \quad (1)$$

where A is the square matrix, x is eigenvector and λ is the eigenvalue. The value λ can be obtained by solving the following equation.

$$|\mathrm{A} - \lambda E| = 0 \quad (2)$$

where, E is an identity matrix.

The number of the eigenvalues is equal to the size of the matrix. For every root of the aforementioned equation, according to the defining equation, one eigenvector will be attained.

Table 4
Example of the variable.

V_{E1}	V_{E2}	V_{E3}	V_{E4}	V_{E5}	V_{E6}
0.8738	0.0510	−0.036 − 0.005i	−0.036 + 0.005i	−0.0075 + 0.006i	−0.0075 − 0.006i

V_{CE1}	V_{CE2}	V_{CE3}	V_{CE4}	V_{CE5}	V_{CE6}
2391.56	33.98	0.1 + 32.55i	0.10 − 32.55i	13.62 + 3.23i	13.62 − 3.23i

According to the descending sequence of the eigenvalues, corresponding eigenvectors are arranged. For example, the eigenvectors and eigenvalues of Table 2 in sequence are as follows,

Eigenvalue =	2391.56	33.98	0.1 + 32.55i	0.10 − 32.55i	13.62 + 3.23i	13.62 − 3.23i
	−0.46 + 0.00i	−0.26 − 0.22i	−0.26 + 0.22i	0.42 − 0.11i	0.42 + 0.11i	−0.12 + 0.00i
	−0.41 + 0.00i	−0.51 + 0.05i	−0.51 − 0.05i	−0.65 + 0.00i	−0.65 + 0.00i	−0.67 + 0.00i
Eigenvector =	−0.09 + 0.00i	0.27 − 0.34i	0.27 + 0.34i	−0.16 + 0.00i	−0.16 − 0.00i	0.45 + 0.00i
	−0.54 + 0.00i	0.52 + 0.00i	0.52 + 0.00i	0.48 + 0.06i	0.48 − 0.06i	0.09 + 0.00i
	−0.56 + 0.00i	−0.15 + 0.25i	−0.15 − 0.25i	0.18 + 0.04i	0.19 − 0.04i	0.46 + 0.00i
	−0.08 + 0.00i	0.03 + 0.27i	0.03 − 0.27i	−0.30 − 0.01i	−0.29 + 0.00i	−0.33 + 0.00i

In this matrix, every column represents an eigenvector for the traffic flow schematic in Table 2.

When multiplying a matrix with its eigenvector, the matrix just multiplies the eigenvector by a constant (i.e. the eigenvalue). Any vectors input to the spatiotemporal matrix will be transformed disorderedly while eigenvectors will steadily magnify or shrink. As a matrix can provide the distribution of elements within the matrix and their correlations, these mathematical variables (eigenvectors and eigenvalues) are able to represent the features of the spatiotemporal loop data schematic.

2.3. Step 3 processing the variables for case control study

Eigenvectors and eigenvalues for the traffic flow schematics were taken as the variables for further case control modelling. Considering the fact that the vector mainly represents the direction in multidimensional space, the angel between the eigenvector and the standard vector [1 1 1 1 1 1] could be an easier model input variable than a vector. This cosine value can be calculated as follows

$$\cos(\mathrm{a}, \mathrm{b}) = \frac{a \cdot b}{|a| \cdot |b|} \tag{3}$$

where, a, b are the eigenvector and the standard vector respectively.

Moreover, an ideal traffic flow with low crash risk would involve all the vehicle moving with the same speed and same gap. In this situation all the values in the speed, occupancy and volume schematics are the same and their corresponding eigenvectors are [1 1 1 1 1 1]. Hence, [1 1 1 1 1 1] is reasonable choice as the standard vectors for 6×6 loop data schematic. Taking the traffic volume matrix in Table 2 as an example, the input variables for further modelling are listed in Table 4. Defining variables, 'V' means volume (similarly 'S' means speed and 'O' means occupancy). 'E' means the eigenvalue, 'CE' means the cosine value between the eigenvector and standard vector. For example, V_{E1} represents the cosine value between the 1st volume matrix eigenvector and standard, and V_{CE1} represents the eigenvalue of the 1st eigenvector.

2.4. Step 4 logistic model

Logistic regression analysis is commonly used in accident modelling research. For the case control study (accident occurrence 1 and crash free 0) in this paper, a logistic regression method is applied. The general form for the probability accident model follows,

$$\mathrm{P}(y_i = 1|x_i) = p_i = \frac{e^{\alpha+\beta x_i}}{1 + e^{\alpha+\beta x_i}} \tag{4}$$

And the odds ratio is calculated as,

$$\frac{p_i}{1 - p_i} = e^{\alpha+\beta x_i} \tag{5}$$

where

$x_i = \textit{input variables series}$,

$\alpha = \textit{constant in logistic model}$,

β = coefficients for independant input variables series.

3. Data

Interstate freeway accidents which occurred in California in 2012 collected from the California Statewide Integrated Traffic Records System (SWITRS) were used in this study. The corresponding loop data was collected in Caltrans Performance Measurement System (PeMS). PeMS data includes average speed, traffic volume, and average occupancy for every thirty seconds at all loops located upstream and downstream of the corresponding accident. Accident data and relevant traffic flow information from the loops were linked according to the time indicator and the location indicator, which is the post mile. SWITRS and PeMS have different post mile fields—absolute post mile and California post mile respectively. A computer program was composed to match them.

1902 accident records and the corresponding loop data in four freeways (I5, I10, I405 and I15) which had the highest crash frequency in their districts were initially collected as the case in this

Table 5
Accident data.

Freeway	District	Length (miles)	Records
I5	3	26.7 × 2	189
I10	7	46.8 × 2	739
I405	12	24 × 2	316
I15	8	239.6 × 2	658
Total		337.1 × 2	1902

Table 6
Final selected samples.

	5 min		2 min	
	case	control	case	Control
Final Complete Samples	153	601	124	562

Table 7
Modelling results of 6 × 6 spatiotemporal schematics.

	Variables	Estimate (β)	Std Err	Wals	Sig.	General Predicting Ratio	Accident Predicting Ratio
Spatiotemporal schematic model with 5-min data	V_{E3}	−0.015	0.007	4.974	0.026	86.05%	17.5%
	O_{CE1}	0.244	0.106	5.323	0.021		
	O_{CE3}	−9.389	4.162	5.09	0.024		
	O_{CE5}	−11.397	4.782	5.679	0.017		
	S_{CE1}	−0.29	0.099	8.598	0.003		
	S_{CE3}	91.676	29.527	9.64	0.002		
Spatiotemporal schematic model with 2-min data	V_{CE1}	0.273	0.165	2.74	0.098	80%	2.4%
	O_{E3}	1.101	0.57	3.731	0.053		

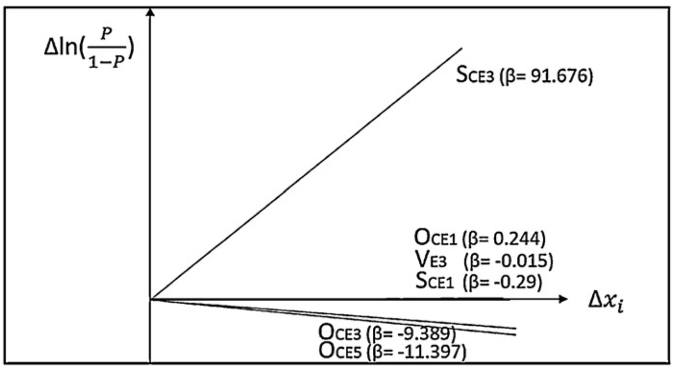


Fig. 1. Sensitivity analysis of model variables.

study as shown in Table 5. The loop data on the same prior duration of the crash occurrence on the same day 1, 2, 3 and 4 week before were selected as the control.

Due to the missing data deficiency of the loop systems, incomplete matrix with more than one row or one column was excluded while processing the loop data matrix. The final case and control samples are shown in Table 6.

4. Result

Six spatiotemporal schematic eigenvectors and eigenvalues based logistic models of speed, occupancy and volume were set up. Therefore there are 36 input variables for modelling.

The parameter estimates and related statistical summary of coefficients of models from 5-min slice matrix and 2-min slice matrix are shown in Table 7. It can be learned from the table that both the eigenvectors and eigenvalues can significantly impact the accident likelihood. Both of these two models works in crash risk predicting, but 5-min matrix works much better with general predicting ratio 86.05% and accident predicting ratio 17.5%. Compared with 5 min slice data, 2 min slice data may not well represent the traffic flow variation in a certain period before the crash.

The gradient (β) in Fig. 1 shows how changing variable values in the 6 by 6 model impacts the crash risk which is represented by the natural logarithm of accident odd ratio. In the 6 by 6 model, the most important risk factor is SCE3, which means the speed value affects the crash most. The conclusion can be much similar with most of the previous studies that the speed variation in the crash nearby section is one of the major risks to the crash. And the other two important protective factor are OCE3 and OCE5.

5. Discussion

This paper proposed loop data schematic eigenvectors based real-time crash prediction models with the general predicting ratio 86.05% and accident predicting ratio 17.5%. Although the predicting ratio of the model is not as high as some of the previous studies (For example, Abdel-Aty et al. (2004) identified 89% of crashes utilizing modern ITS devise data such as AVI and RTMSs.), the proposed model has the following advantages compared with the previous models:

(1) As the eigenvectors and eigenvalues are independent, the traffic index schematic based method avoids the multicollinearity problem and is more reliable.
(2) The speed, occupancy and volume schematics can reflect the overall traffic flow status before the crash occurrence rather than a single loop parameter.
(3) The model has certain forgiving ability with respect to the input. If just a few values are invalid in the schematic, (for example if some values in Table 2 become 0), the direction of the eigenvectors in Table 4 will just change slightly. Data collected with more recent technologies can be used to further improve the accuracy and adaptability of the predicting model. Although the proposed method has good robustness of missing data, data completeness will absolutely add to the accuracy of prediction.
(4) For the fact of uncertain crash time in the accident database, the proposed model reflects the traffic fluctuate situation in the whole road section in 30 min period. Therefore the model can cover 10 min recording error for the accident record.

The results of the proposed traffic flow schematics eigenvector based model will help to predict the impending risk of accidents for real-time traffic control, which can help to achieve preemptive safety management and cut down accident rate in traffic safety management. Future research will be focused on the countermeasures to achieve optimal traffic flow schematics with minimum crash risks.

Acknowledgements

This research is supported by the National “Twelfth Five-Year” Plan for Science & Technology Support Project in China (2014BAG01B04), the National Natural Science Foundation (71301119) and the Shanghai Natural Science Foundation (12ZR1434100).

References

Abdel-Aty, M., Pande, A., 2005. Identifying crash propensity using specific traffic speed conditions. J. Safety Res. 36 (1), 97–108.
Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F.M., Hsia, L., 2004. Predicting freeway crashes from loop detector data by matched case-control logistic regression. TRR J. Transp. Res. Board 1897 (1), 88–95.
Abdel-Aty, M., Pande, A., Hsia, L.Y., Abdalla, F., 2005. The potential for real-time traffic crash prediction. ITE J. 75 (12), 69.
Ahmed, M.M., Abdel-Aty, M.A., 2012. The viability of using automatic vehicle identification data for real-time crash prediction. IEEE Trans. Intell. Transp. Syst. 13 (2), 459–468.
Ahmed, M.M., Abdel-Aty, M., Yu, R., 2012. Bayesian updating approach for real-time safety evaluation with automatic vehicle identification data. TRR J. Transp. Res. Board 2280 (1), 60–67.

Golob, T.F., Recker, W.W., Alvarez, V.M., 2004. Freeway safety as a function of traffic flow. Accid. Anal. Prev. 36 (6), 933–946.

Golob, T.F., Recker, W., Pavlis, Y., 2008. Probabilistic models of freeway safety performance using traffic flow data as predictors. Safety Sci. 46 (9), 1306–1333.

Hossain, M., Muromachi, Y., 2012. A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. Accid. Anal. Prev. 45, 373–381.

Hughes, R., Council, F., 1999. On establishing the relationship (s) between freeway safety and peak period operations: performance measurement and methodological considerations. In: Paper Presented at the 78th Annual Meeting of Transportation Research Board, Washington, DC.

Lee, C., Saccomanno, F., Hellinga, B., 2002. Analysis of crash precursors on instrumented freeways. TRR J. Transp. Res. Board 1784 (1), 1–8.

Lee, C., Hellinga, B., Saccomanno, F., 2003. Real-time crash prediction model for application to crash prevention in freeway traffic. TRR J. Transp. Res. Board 1840 (1), 67–77.

Madanat, S., Liu, P.-C., 1995. A Prototype System for Real-Time Incident Likelihood Prediction. Transportation Research Board-National Research Council.

Oh, C., Oh, J.-S., Ritchie, S., Chang, M., 2001. Real-time estimation of freeway accident likelihood. In: Paper Presented at the 80th Annual Meeting of the Transportation Research Board, Washington, DC.

Oh, J.-S., Oh, C., Ritchie, S.G., Chang, M., 2005. Real-time estimation of accident likelihood for safety enhancement. J. Transp. Eng. 131 (5), 358–363.

Pande, A., Abdel-Aty, M., 2006. Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways. TRR J. Transp. Res. Board 1953 (1), 31–40.

Shew, C., Pande, A., Nuworsoo, C., 2013. Transferability and robustness of real-time freeway crash risk assessment. J. Safety Res. 46, 83–90.

Vlahogianni, E.I., Karlaftis, M.G., Orfanou, F.P., 2012. Modeling the effects of weather and traffic on the risk of secondary incidents. J. Intell. Transp. Syst. 16 (3), 109–117.

Xu, C., Liu, P., Wang, W., Li, Z., 2012. Evaluation of the impacts of traffic states on crash risks on freeways. Accid. Anal. Prev. 47, 162–171.

Xu, C., Wang, W., Liu, P., Guo, R., Li, Z., 2014. Using the Bayesian updating approach to improve the spatial and temporal transferability of real-time crash risk prediction models. Transport. Res. C Emerg. Technol. 38, 167–176.