

高速公路实时事故风险研判模型及可移植性

游锦明, 方守恩, 张兰芳, 折欣

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

摘要: 通过 G15 沈海高速公路南通段上布设的高清卡口过车数据对路段上发生的实时事故风险进行研究. 采用配对案例对照方法, 结合基于随机森林的参数选取方法对 3 个子路段上的事故分别建立了支持向量机模型. 结果表明, 基于高清卡口采集的高分辨率过车数据构建的支持向量机模型相对既有研究中的模型而言其性能较优; 对 3 个子路段分别构建的支持向量机模型进行可移植性分析发现各支持向量机模型均具有一定的可移植性, 经过参数重新标定后可直接应用于至邻近道路对其实时事故风险状态进行研判, 并有着相对较高的预测精度.

关键词: 高速公路; 实时研判; 事故风险; 支持向量机 (SVM); 可移植性

中图分类号: U491

文献标志码:

Real-Time Crash Risk Prediction Models and Transferability Analysis on Freeways

YOU Jinming, FANG Shouen, ZHANG Lanfang, SHE Xin

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China)

Abstract: The paper aims to investigate the real-time crash risk based on the High Definition Monitoring System data on G15 Freeway in Nantong, China. Matched case-control method and parameter filtering method based on random forest were utilized to build SVM (support vector machine) models for the crashes on three sub-segments respectively. Results show that the SVM models based on high definition data collected by High Definition Monitoring System show better performance than those in existing studies. The transferability research was also conducted to verify the transferability of the proposed SVMs and results indicate that the models can be transferred to a certain extent. They could be applied in real-time crash prediction process on road segments nearby after the calibration of the parameters in the models and the transferred models have relatively higher

prediction accuracy.

Key words: freeway; real-time prediction; crash risk; support vector machines (SVM); transferability

随着我国交通行业的快速发展, 公众在交通出行时对高速公路交通安全的要求越来越高, 交通安全领域已由传统的被动安全改善向主动安全风险防控发展. 然而由于我国高速公路信息化发展迟滞, 多数高速公路尚未完善配置密集的监控检测设备用以信息采集, 使得调度管理人员无法对高速公路进行全方位把控, 部分高速公路路段处于监控盲区, 从而导致该部分路段的实时交通运行安全无法得到保障. 与此同时, 我国高速公路信息化正处于高速发展阶段, 公安部、交通部等主管部门纷纷出台相关规划文件, 在新型智能交通快速发展的背景下, 提出了主动风险防控的需求, 各地高速公路交通管理部门也在同步进行信息化建设, 逐渐加密完善监控检测、信息发布等设备布设, 提升高速公路交通安全服务水平.

为了实现全方位的高速公路实时事故风险研判, 基于我国高速公路现状信息化管理下的稀疏高清卡口交通流数据, 对实时事故风险研判模型进行构建, 并开展模型的可移植性研究, 从而为国内高速公路相关交通安全管理部门在实时主动交通安全风险分析方面提供借鉴.

1 研究综述

在实时事故风险预测方面, 国外专家开展了大量的研究, 尤其以美国、欧洲、日本等发达国家为主, 由于其高速公路信息化较早、信息化程度较高, 许多专家学者基于其海量密集的高速公路运营数据, 开

收稿日期: 2018-04-27

基金项目: 国家重点研发计划(2016YFC0802701)

第一作者: 游锦明(1992—), 男, 博士生, 主要研究方向为道路安全与环境工程. E-mail: 1210702@tongji.edu.cn

通信作者: 方守恩(1961—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为道路安全工程. E-mail: fangsek@tongji.edu.cn

展了广泛的事故先兆研究工作。

早期高速公路交通流数据以线圈数据为主,专家学者通过采用不同的模型方法在交通流状态参数与事故结果之间建立起因果关联,如 Logistic 模型^[1-3]、神经网络模型^[4-6]、贝叶斯方法^[7-8]等。2012 年随着美国高速公路上布设了新型的自动车辆识别系统(AVI),使得旅行速度的实时测度成为了可能,该指标能够反映车辆在某个路段的平均行驶速度。Abdel-Aty 等^[7]通过结合该系统所采集的数据和路段上的事故数据进行建模,其 Logistic 回归模型能够鉴别 69% 的事故伴随着 42.01% 的误报率,同时提出了一种基于贝叶斯改进方法的 Logistic 回归模型,模型能够提升 3.5% 的事故预测精度,但是误报率也提升为 45.83%。由于大量的事故误报会影响驾驶人对管控评价系统的认同度从而降低管控系统的有效性,因此模型还有待继续改进。

2013 年 Yu 等^[9]采用分类回归树和贝叶斯 Logistic 回归对指标选取后,构建了支持向量机模型,采用受试者曲线下面积(AUC)对不同核函数的支持向量机模型进行评判,AUC 值均大于 0.7,能够较优地对事故风险进行预测。同年 Yu 等^[10]采用随机效应 Logistic 回归模型分析工作日和周末事故数据差异、道路线形数据和远程微波交通检测器采集的交通流数据(RTMS),结果发现工作日事故更易发生在拥堵区域而周末事故往往发生在自由流条件下。2013 年 Ahmed 等^[11]采用随机梯度推理机器学习方法,结合 RTMS 和 AVI 的交通流数据对事故风险进行建模,模型最终达到了 89% 的事故预测精度伴随着 6.5% 的误报率。Shi 等^[12]通过对未处理的 AVI 原始数据对事故风险进行研究,采用多层随机参数模型和负二项模型进行建模来提升模型对于复杂数据和样本间的非匀质性的适应能力,研究表明低速、高速度离散性和高流量会显著提升事故风险。这些研究均表明了采用高分辨率的交通流参数对事故预测的有效性。

国内由于高速公路信息化程度较低,在实时事故风险预测方面的研究较少。哈尔滨工业大学赵新勇^[13]结合支持向量机和决策树提出了基于交通流参数的事故多发点交通安全事件风险研判模型,选取高速公路追尾事故作为分析的目标交通安全事件,采用了支持向量机算法建立了追尾事故风险预测模型。孙剑等^[14]基于上海市快速路上的线圈交通流数据和事故数据,采用贝叶斯网络(BN)对事故风险进行建模,基于高斯混合模型和最大期望算法对

贝叶斯网络模型参数进行估计,分别对事故点前后 2 组检测器和 4 个时间段的 8 组交通流数据进行建模,结果表明使用事故点上下游各一检测器、基于事故前 5~10 min 内的交通流数据的模型效果最优,其事故预测精度达到了 76.94%。次年贾丰源等^[15]对模型进行了改进,采用随机森林对变量进行筛选,最终筛选变量后建立的贝叶斯网络模型预测效果更佳,其事故预测准确率达到 82.78%。

研究表明事故风险预测模型经过参数重训练后能够应用于其他高速公路实时交通安全评价^[16],但是由于国内外驾驶人驾驶习惯的差异性和不同区域的交通流特性不同,直接使用现有的模型实际上并不可取,模型的可移植性分析显得尤为重要。Sun 等^[17-18]基于上海市快速路延安高架和南北高架的交通流检测器数据,翔实地研究了不同时间段、不同惩罚因子等模型参数对于实时事故风险预测模型的影响,并进行了相互的可移植性研究,验证了模型的有效性和普适性。然而该部分研究依然基于间距为 300~500 m 的检测器,上下游共 4 个检测器,与国外研究保持一致。而国内高速公路上检测设备较为稀疏,不能支撑模型研判,这使得高密度的模型无法在这种条件下进行应用验证。另一方面,许多研究指出速度离散性对于事故风险有着显著影响,但是由于现有交通流数据多数来自于集计的线圈数据,速度离散性指标并没有得到充分研究^[19]。因此开展基于稀疏检测器的高速公路实时事故风险研判研究,探索将已信息化高速公路路段的研判模型应用至相邻路段,并进行模型可移植性研究,为新信息化路段的实时交通安全风险防控提供决策依据。

建模技术路线如图 1 所示。

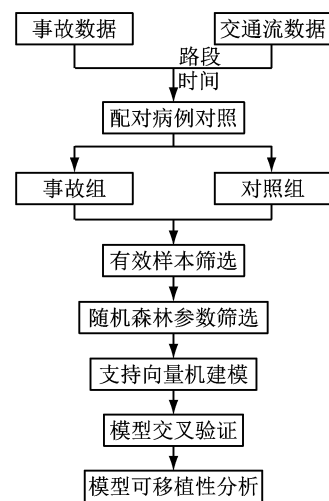


图 1 建模技术路线

Fig.1 Architecture of the study

2 数据准备

数据来源于 G15 沈海高速南通段,该段全长约 100 km,设计速度 $120 \text{ km} \cdot \text{h}^{-1}$,双向 6 车道,数据主要为事故数据和高清卡口过车数据两大类。

2.1 事故数据

采用的事故数据为 2016 年 1 月 1 日至 2016 年 10 月 31 日期间 G15 沈海高速公路南通段所登记录入的事故数据.公安部交通事故登记录入了事故发生时间、事故地点、事故类型、事故形态、死伤人数等信息,记录信息较全面,部分事故信息在记录时缺少方向信息,因此该部分事故数据无法直接定位到高速公路路段上进行后续分析应用,予以剔除.整体而言,共记录事故数据 5 924 起,按照事故形态划分如图 2.

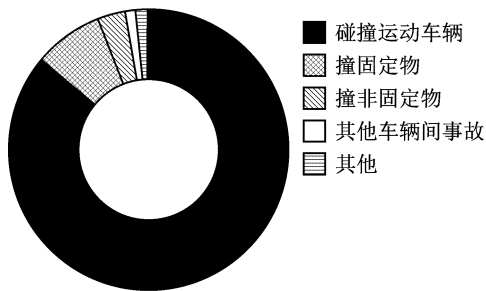


图 2 G15 沈海高速南通段事故形态统计

Fig.2 Detailed crash types of the crashes on G15 Freeway in Nantong

从图中可以发现,车辆间事故(含碰撞运动车辆、其他车辆间事故)占了总体约 88%,而碰撞运动车辆占了约 86%,说明该路段上大部分事故均为动态交通中的追尾碰撞事故,而对于单车事故而言由于偶然因素较多,因此着重研究车辆间事故的实时风险研判。

2.2 高清卡口交通流数据

采用的高清卡口交通流数据为 2016 年 1 月 1 日至 2016 年 10 月 31 日期间 G15 沈海高速公路南通段上公安部交通管理部门所布设的高清卡口系统所采集的过车数据.高清卡口过车数据相对于传统的线圈数据而言,能够更精准地采集单独车辆通过时的详细数据,可以准确分辨车辆所处车道信息、车速信息等,为进行速度离散性于交通安全的影响研究提供了数据支撑.该路段共布设 5 套高清卡口系统,其中一套 D122 为单行卡口,只能记录单侧通行方向车辆过车信息.其布设点位信息如图 3.

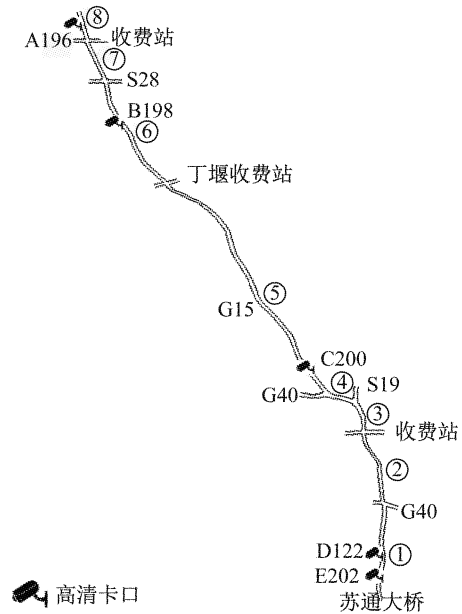


图 3 G15 沈海高速南通段高清卡口布设点位情况
Fig.3 Locations of the high definition monitoring system on G15 freeway in Nantong

由图中可见,全长约 100 km 的高速公路路段只配置了 5 套高清卡口设备,2 个高清卡口最近的距离为 5 km,最远的距离长达 33 km,与国外布设的线圈布设密度(0.5 mile,约 800 m)相比,相差甚远.目前在既有研究中,已经证明了基于单检测器的交通流数据用以实时事故风险研判的可靠性和有效性^[20],因此,基于单个高清卡口的交通流数据对同路段的事故进行研判研究.以高速公路出入口和立交枢纽作为节点,将该路段划分为 8 个子路段,则共有 4 个子路段上布设了高清卡口,因此最终该 4 个路段上发生的事故能够通过时间和空间配对方法匹配到相应的交通流数据.如若事故发生在丁堰收费站和 S28 互通立交之间,则匹配到子路段⑥上布设的高清卡口设备 B198,根据事故发生时间确定其对应时间前 5 至 10 min 的高清卡口过车数据作为事故发生前的交通流状态.然后将该时间片段内的过车数据针对其时空分布抽取对应的各车道的分布参数(Q_n, E_{Q_n}, D_{Q_n})、速度分布参数($E_{V_n}, D_{V_n}, R_{V_n}$)等共 18 个指标,其中 Q_n, E_{Q_n}, D_{Q_n} 分别代表第 n 车道 5~10 min 内的流量、车辆时间分布均值,车辆时间分布离散度, $E_{V_n}, D_{V_n}, R_{V_n}$ 分别代表第 n 车道 5~10 min 内的速度均值、速度方差、速度极差。

2.3 配对案例对照方法

采用对照组和案例组 4 : 1 的比例对驾驶人、车辆特性及道路环境等因素进行控制变量研究.对照组根据事故发生的时间进行选定,分别选取其前 14

d 相应时间、前 7 d 相应时间、后 7 d 相应时间、后 14 d 相应时间的交通流数据作为对照组,未能匹配到高清卡口过车数据的样本及匹配到异常值过车数据的样本在最终样本确定时予以剔除,最终共有 3 个路段上的事故数据和过车数据完成配对,如表 1。

表 1 配对后有效样本信息

Tab.1 Sample information after spatial-temporal matching

卡口编号	路段编号	有效事故样本数/个	有效对照样本数/个
A196	⑧	55	224
B198	⑥	70	261
C200	⑤	48	141

确定有效样本后,将根据各子路段的样本进行分别建模,从而为后续的可移植性研究作好铺垫。

3 支持向量机建模

支持向量机(SVMs)作为现阶段最为流行的监督式机器学习算法之一,被广泛应用于文本分类、图像识别、语音识别等领域,在分类问题、回归问题上均有着良好的表现。由于传统的广义线性模型假设自变量与因变量之间有着显著的线性关系,而当自变量表现出很强的非线性特征时,那么采用广义线性模型去对参数进行估计时便会产生显著的偏差^[21],而支持向量机在解决小样本、非线性及高维模式识别中表现出许多特有的优势,因此基于当前样本的小样本特点,选取支持向量机 C-SVC 模型作为实时事故风险研判的模型手段。

在机器学习领域,样本变量的选取决定了机器学习的效率。而随机森林可以以自然的形式对分类回归问题中的变量进行排序从而筛选出影响结果的重要变量,因此为了构建高效的机器学习模型,提升模型计算效率和避免过拟合现象,选用随机森林进行参数筛选。针对 3 个子路段构建的样本分别筛选关键参数,用以后续的支持向量机建模。各子路段参数筛选后的参数概述信息如表 2 所示。

研究经验表明,采用径向基核函数作为 C-SVC 模型高维映射变换函数的分类预测模型的分类效果普遍优于基于多项式、Sigmoid 核函数的 C-SVC 模

表 2 随机森林参数筛选后各子路段交通流参数统计
Tab.2 Summary statistics of each sub-segment after traffic parameter filtering based on Random Forest

卡口编号	指标	均值	方差	1/4 分位数	3/4 分位数
A196	E_{Q8}	59.71×10^{-4}	6.26×10^{-4}	58.71×10^{-4}	62.67×10^{-4}
	D_{Q8}	4.77×10^{-4}	2.45×10^{-4}	2.85×10^{-4}	6.68×10^{-4}
	E_{V3}	60.39	20.78	50.88	78.44
	D_{V3}	25.88	10.11	16.94	36.10
B198	R_{V3}	110.45	26.59	98.00	124.00
	E_{Q8}	58.56×10^{-4}	10.41×10^{-4}	58.85×10^{-4}	62.36×10^{-4}
	D_{Q8}	4.52×10^{-4}	2.64×10^{-4}	2.57×10^{-4}	6.60×10^{-4}
	E_{V3}	63.28	21.14	56.38	78.56
C200	D_{V3}	29.53	11.19	22.48	39.69
	E_{Q1}	48.13×10^{-4}	30.55×10^{-4}	31.00×10^{-4}	57.00×10^{-4}
	E_{Q8}	60.17×10^{-4}	7.46×10^{-4}	58.69×10^{-4}	63.59×10^{-4}
	D_{Q8}	4.36×10^{-4}	2.59×10^{-4}	2.36×10^{-4}	6.43×10^{-4}
	E_{V3}	70.55	23.14	59.48	89.39
	D_{V3}	27.85	12.52	16.32	40.47

型,因此采用径向基核函数作为样本高维映射核函数。在 C-SVC 模型的构建时,引入了常数 C 作为惩罚系数控制损失的大小。模型求解中 C 可作为调节参数,用于影响训练模型的性能。此外,径向基核函数中 γ 参数也是模型训练前需输入的常数,该参数的数值也会明显影响模型的性能。所以应用 SVM 方法解决分类问题还需解决 SVM 模型参数及核函数参数的寻优问题,得到分类效果最佳的一组 C、 γ 参数。目前在核函数参数选取方面,机器学习领域尚未有统一规则,往往根据历史经验,给定一定的参数范围空间 $C \in [2^0, 2^{16}]$ 、 $\gamma \in [2^{-16}, 2^0]$,按照梯度,通过编程枚举的方式对不同参数组合下的模型预测效果进行对比,结合十重交叉验证法得出最优的参数对(C, γ)并建立对应的 SVM 分类器。将各路段对应的样本按照训练集与测试集的比例为 70%、30%的比例分配,以训练集训练模型,最后采用测试集的预测精度(包含事故预测精度和总体预测精度,其中事故预测精度指在所有事故样本中成功被预测的事故样本的比例,总体预测精度指所有样本成功被预测的样本比例)和 AUC 作为支持向量机分类器的分类性能评价指标。3 个子路段对应的支持向量机分类器相关参数和分类性能如表 3。

表 3 各子路段支持向量机分类器参数及分类性能

Tab.3 Parameters and performance of each SVMs

卡口编号	$C/10^{-2}$	$\gamma/10^{-4}$	事故预测精度/%	原始事故预测为事故样本数量	原始事故样本数量	总体精度/%	AUC
A196	28.72	89.70	83.33	10	12	97.62	0.98
B198	5.44	9.77	76.47	13	17	95.96	0.94
C200	151.57	30.00	100.00	16	16	91.23	0.95

从表中可以看出,从事事故预测精度、总体精度、AUC各指标角度对分类器性能进行评价,都可以发现构建的3个基于高清卡口的支持向量机分类器分类性能较优,同比国内外其他研究中构建的模型而言,模型分类精度有着较程度的提升,这也间接地表明了采用高分辨率的交通流数据用以实时事故风险研判的可靠性,同时,也为模型的可移植性拓展研究提供了数据支撑。

4 模型可移植性分析

为提升模型的普适性,需对模型的可移植性进行分析,从而使模型能应用于其他路段。由于实时事故风险研判需要大量的历史事故数据和交通流数据进行支持,而往往对于新纳入信息化发展规划中的路段而言缺乏了历史的交通流数据用以进行研判支撑,通过模型可移植性分析,可以将已信息化路段构建的研判模型转移至周边道路进行应用,使得模型具备推广应用能力。考虑到国内高速公路的逐步信息化进程,对G15上已经能够进行实时事故风险研判的3个路段进行交互性的可移植性分析,即采用某个既定路段所构建的模型对其他相邻路段的交通流状态进行判别,从而达到实时事故风险研判的目的。例如,采用A196对应路段构建的支持向量机模型分别对B198路段的样本和C200路段的样本分别进行预测,采用事故预测精度和总体预测精度对分类器性能进行评价。各段交互预测结果如表4。

表4 各路段支持向量机模型可移植性分析

Tab.4 Transferability analysis of the SVMs on each sub-segment

模型	预测精度/%		
	A196 样本	B198 样本	C200 样本
A196	83.33,97.62	71.43,93.96	47.92,86.77
B198	100.00,97.85	76.47,95.96	72.92,93.12
C200	100.00,96.06	92.86,96.07	100.00,91.23

注:逗号前为事故预测精度,逗号后为总体预测精度。

从表中可以看出,3个支持向量机分类器的可移植性效果都较佳,尤其以C200处建立的支持向量机分类器而言,其建立的支持向量机分类器能够较好地A196路段处和B198路段处的事故样本和总体样本进行预测,说明其分类器的参数集合更具有普适性,可移植性更优。

针对各分类器对各样本集的错误分类的样本进行时间、天气分析可以发现:

(1) 对于A196对应路段的样本而言,采用

B198路段构建的SVM错误地将7起正常样本划分为事故样本,其中3起样本为雨天环境下的样本;采用C200路段构建的SVM错误地将6起正常样本划分为事故样本,其中3起为雨天环境下的样本。总体而言,错分类样本中有5起正常样本是B198-SVM和C200-SVM的错误划分,这5起样本中有3起样本都发生在国定假日期间。

(2) 对于B198对应路段的样本而言,采用A196路段构建的SVM错误地将20起样本错误划分,其中9起样本为雨天环境下的样本,12起为夜间发生的事故样本;采用C200路段构建的SVM错误将9起样本划分,其中4起为雨天环境下的样本,3起为夜间发生的事故样本。总体而言,错分类样本中有8起样本是A196-SVM、B198-SVM和C200-SVM3个分类器的错误划分,8起中有2起样本发生在国定假日期间。

(3) 对于C200对应路段的样本而言,采用A196路段构建的SVM错误地将25起样本错误划分,其中4起样本为雨天环境下的样本,5起为夜间发生的事故样本;采用B198路段构建的SVM错误将16起样本划分,其中4起为雨雪天环境下的样本,4起为夜间发生的事故样本。总体而言,错分类样本中有4起样本是A196-SVM、B198-SVM和C200-SVM3个分类器的错误划分。

总体而言,排除掉夜间发生的样本和雨雪天气环境下的分类错误样本,采用邻近路段构建的支持向量机分类器对当前路段进行研判能够非常接近当前路段所构建的支持向量机分类器性能,由于各个支持向量机性能也未能够实现完美地将样本进行分类,这也间接地表明了通过对交通流状态的判别能够有效地研判出多数事故高危状态,但是部分事故的发生受突发因素影响更加直接,如驾驶人分心驾驶疲劳等,该部分事故就目前运营采集数据而言暂无法获取,因此未来跟随着汽车技术发展能够采集到更高粒度的实时车辆级数据后,可以开展全方面的事故与交通流及驾驶人、道路环境等因素之间的内在关联研究。

从错误分类样本的发生的共性角度而言,夜间视野不佳、驾驶人疲劳等因素和雨天低能见度和路面湿滑等因素,均对交通安全有着较大的影响,但往来说,这些因素对单车事故影响较大,在流量较低情况下易导致单车事故发生,而在交通流参数上表征不甚明显,为了提高模型研判精度及可移植性效果,一方面由于当前样本量较小,后续可以随着样本

量的积累提取到更多的样本信息用以构建支持向量;另一方面为了后续对实时事故风险研判进行更加深入的研究时,可以引入高分辨率的实时气象数据进行综合分析,提高模型研判精度,并达到系统性的事故风险研判目的.同时,随着高速公路信息化的不断开展,可以逐步开展基于相邻路段、基于交叉路段的可移植性研究,提升模型的可移植性.除此之外,通过对可移植性的研究,可以考虑结合多个路段对应的分类器模型对某个路段的交通流状态进行综合研判,使得研判结果更加准确有效.

5 结语

通过采用 G15 沈海高速公路南通段上的高清卡口检测的高分辨率过车数据对实时事故风险进行研判,结合随机森林进行参数筛选,采用支持向量机对事故状态进行实时分类预测,结果表明基于高清卡口数据构建的实时事故风险研判模型能够较好地对事故状态进行预测,有着较高的事故预测精度,模型性能较优;基于对 3 个子路段分别构建的支持向量机模型进行可移植性分析,结果表明各支持向量机模型均具有一定的可移植性,通过对模型参数标定后可直接将模型应用到邻近道路对其实时事故风险状态进行研判,并有着相对较高的预测精度.除去夜间、雨雪天气环境下产生的错误判别样本后,模型可移植性效果更优.

参考文献:

- [1] ABDEL-ATY M, UDDIN N, PANDE A. Predicting freeway crashes from loop detector data by matched case-control logistic regression[J]. *Transportation Research Record: Journal of the Transportation Research Board*, 2004, 1897(1): 88.
- [2] GOLOB T F, RECKER W, PAVLIS Y. Probabilistic models of freeway safety performance using traffic flow data as predictors[J]. *Safety Science*, 2008, 46(9): 1306.
- [3] XU C, LIU P, WANG W, *et al.* Evaluation of the impacts of traffic states on crash risks on freeways[J]. *Accident Analysis & Prevention*, 2012, 47(1): 162.
- [4] OH J, OH C, RITCHIE S G, *et al.* Real-time estimation of accident likelihood for safety enhancement[J]. *Journal of Transportation Engineering*, 2005, 131(5): 358.
- [5] ABDEL-ATY M, PANDE A. Identifying crash propensity using specific traffic speed conditions[J]. *Journal of Safety Research*, 2005, 36(1): 97.
- [6] VLAHOGIANNI E I, KARLAFTIS M G, ORFANOY F P. Modeling the effects of weather and traffic on the risk of secondary incidents[J]. *Journal of Intelligent Transportation Systems*, 2012, 16(3): 109.
- [7] AHMED M, ABDEL-ATY M, YU R. Bayesian updating approach for real-time safety evaluation with automatic vehicle identification data[J]. *Transportation Research Record: Journal of the Transportation Research Board*, 2012, 2280(1): 60.
- [8] HOSSAIN M, MUROMACHI Y. A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways[J]. *Accident Analysis & Prevention*, 2012, 45(8): 373.
- [9] YU R, ABDEL-ATY M. Utilizing support vector machine in real-time crash risk evaluation[J]. *Accident Analysis & Prevention*, 2013, 51(1): 252.
- [10] YU R, ABDEL-ATY M. Investigating the different characteristics of weekday and weekend crashes[J]. *Journal of Safety Research*, 2013, 46: 91.
- [11] AHMED M, ABDEL-ATY M. A data fusion framework for real-time risk assessment on freeways[J]. *Transportation Research Part C: Emerging Technologies*, 2013, 26(1): 203.
- [12] SHI Q, ABDEL-ATY M, YU R. Multi-level Bayesian safety analysis with unprocessed automatic vehicle identification data for an urban expressway[J]. *Accident Analysis & Prevention*, 2016, 88(1): 68.
- [13] 赵新勇. 基于多源异构数据的高速公路交通安全评估方法[D]. 哈尔滨: 哈尔滨工业大学, 2013.
ZHAO Xinyong. Freeway traffic safety evaluation method based on multi-source heterogeneous data[D]. Harbin: Harbin Institute of Technology, 2013.
- [14] 孙剑,孙杰. 城市快速路实时交通流运行安全主动风险评估[J]. *同济大学学报(自然科学版)*, 2014, 42(6): 873.
SUN Jian, SUN Jie. Proactive assessment of real-time traffic flow accident risk on urban expressway[J]. *Journal of Tongji University (Natural Science)*, 2014, 42(6): 873.
- [15] 贾丰源,孙杰,孙剑. 快速路交通流运行安全关键参数识别与评估[J]. *同济大学学报(自然科学版)*, 2015, 43(2): 221.
JIA Fengyuan, SUN Jie, SUN Jian. Key variables identification and proactive assessment of real-time traffic flow accident risk on urban expressway[J]. *Journal of Tongji University (Natural Science)*, 2015, 43(2): 221.
- [16] SHEW C, PANDE A, NUWORSOO C. Transferability and robustness of real-time freeway crash risk assessment[J]. *Journal of Safety Research*, 2013, 46(9): 83.
- [17] SUN J, SUN J, CHEN P. Using support vector machine models for real-time prediction of crash risk on urban expressways[J]. *Transportation Research Record: Journal of the Transportation Research Board*, 2015, 2432(1): 91.
- [18] SUN J, SUN J. Real-time crash prediction on urban expressways: Identification of key variables and a hybrid support vector machine model[J]. *IET Intelligent Transport Systems*, 2016, 10(5): 331.
- [19] THEOFILATOS A, YANNIS G. A review of the effect of traffic and weather characteristics on road safety[J]. *Accident Analysis & Prevention*, 2014, 72: 244.
- [20] YOU J, WANG J, GUO J. Real-time crash prediction on freeways using data mining and emerging techniques[J]. *Journal of Modern Transportation*, 2017, 25(2): 116.
- [21] LAO Y, ZHANG G, WANG Y, *et al.* Generalized nonlinear models for rear-end crash risk analysis[J]. *Accident Analysis & Prevention*, 2014, 62(1): 9.