

城市快速路实时交通流运行安全主动风险评估

孙 剑, 孙 杰

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

摘要: 基于上海市快速路系统采集的线圈检测器数据和事故数据,应用贝叶斯网络(BN)模型对快速路实时交通流参数与事故风险进行建模分析,并利用可有效应对缺失数据的高斯混合模型和最大期望算法分别对 BN 模型输入和参数进行估计,进而主动评估快速路实时交通流运行安全风险,并对事故状态提前做出预警.分别对可能事故点前后 2 组检测器和 4 个时间段的 8 组交通流数据进行了建模,结果表明使用事故发生地点上下游各一个检测器在事故发生前 5~10min 内的交通流数据建立的 BN 模型效果最好,事故预测准确率为 76.94%.最后不仅与朴素贝叶斯分类、K 近邻、反向传播(BP)神经网络等经典事故风险估计算法进行了对比分析,还与现有的实时风险评估成果进行了对比,结果表明 BN 模型预测效果最好.

关键词: 城市快速路; 交通事故; 贝叶斯网络; 检测器数据; 主动风险评估

中图分类号: U121

文献标志码: A

Proactive Assessment of Real-time Traffic Flow Accident Risk on Urban Expressway

SUN Jian, SUN Jie

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

Abstract: Based on dual-loop detector data and accident data collected on Shanghai expressway, the Bayesian networks (BN) model was adopted for the modeling and analysis of real-time traffic flow parameters and accident risk on expressways. Gaussian mixture model and the expectation-maximization algorithm which could effectively deal with the missing data were also used in the parameters estimation of BN model. Then real-time traffic safety risk was evaluated, and accident warning could be carried out in advance. Different combinations of dual-loop detector data and time segments before accidents were used to develop the optimal accident risk estimation model by BN. The results show that

the BN model adopting the nearest detector data upstream and downstream of the accident site within 5 to 10 minutes before the accident performs the best and the accident prediction accuracy is up to 76.94%. At last, a comparative study was made of the classical accident risk estimation algorithms including naive Bayes classifier, K nearest neighbor and back propagation (BP) neural network as well as the existing real-time risk assessment studies. And the results show that the BN model obtains the best predictive results.

Key words: urban expressway; traffic accident; Bayesian network (BN); detector data; proactive assessment of accident risk

快速路是城市交通的主骨架,承担着大量的城市机动化交通.据统计,北京城市快速路系统总长度约 380 km,只占城市路网总长度的 8%,但却承担了城市日机动车出行量的 50%以上^[1];上海中心城区近 8%的快速道路承担超过 30%以上的车公里出行.另一方面,偶发性交通事故不仅造成人员及财产损失,同时亦是快速路交通拥堵的重要致因.尤其是在高峰时段,虽然事故严重程度低,但是由此引起的交通拥堵和延误会极大地浪费社会成本.而据美国联邦公路局的统计,由于交通事故/事件引起的拥堵约占总拥堵的 25%^[2].现有研究多关注于事故发生后的自动判别及拥堵蔓延,但是对于事故发生前的交通状态与事故发生的相互关系则关注较少.近年来研究表明,交通运行状态与事故风险存在着关联关系. Lee 等^[3]认为交通流在短期内的变化对事故的发生会有很大的影响.因此,使用事故数据和相对应的交通流检测器数据建立合适的模型可以预测快速路上发生事故的风险并且及时发现有可能发生交通事故的交通条件.这里,识别可能发生交通事故的交通条件的问题可以简化成将可能发生交通事故和不会发生交通事故的

收稿日期: 2013-09-02

基金项目: 国家自然科学基金(51278362); 教育部新世纪人才计划 (NCET-13-0425)

第一作者: 孙 剑(1979—),男,工学博士,教授,博士生导师,主要研究方向为交通仿真与实验,交通系统分析与优化.

E-mail: sunjian@tongji.edu.cn

交通状态进行分类,分类模型的输入数据是事故发生前或正常的道路交通状态,输出结果则是发生事故的风险,进而分辨出是否可能发生事故.本研究中道路交通状态以事故地点附近的线圈检测器采集到的基本交通流数据(流量、速度和占有率)表示.利用快速路实时交通流检测数据进行事故风险的主动预测,一方面可以通过车内导航设备或者可变信息板(VMS)对驾驶员进行预警,防患于未然;另一方面,对于管理部门优化救援设施及警力布置等也具有重要的参考意义.

本文首次利用我国快速路交通流检测数据和事故视频监控系统以及事故救援报警信息整理的快速路事故数据,分析了快速路交通流运行状态和事故潜在风险的关系,通过改进的贝叶斯网络(Bayesian networks, BN)模型对快速路事故发生可能性进行了实时预估,并与已有的经典算法(包括朴素贝叶斯分类、K近邻、反向传播(BP)神经网络)进行了对比分析.

1 研究综述

目前,国外在高/快速路的事风险预测特别是基于历史数据的回归统计分析方面已经有了很多研究.而针对实时交通流运行风险的评估这一问题则是近几年理论研究的热点,研究的重点聚焦于如何利用不同的建模方法来分析和预测高/快速路的事风险. Abdel-Aty 和 Pande 等^[4-9]分别利用匹配案例对照 Logistic 回归模型(matched case-control logistic regression)^[4], 分割模型(split models)^[5], 概率神经网络(probabilistic neural network, PNN)^[6], 多层感知器/径向基神经网络(multi-layer perception/radial basis function neural network)^[7], 广义估计方程(generalized estimation equation)^[8]和分类树(classification trees)^[9]来预测事故发生的风险. Oh 等^[10]分别使用贝叶斯统计(Bayesian statistics)和概率神经网络^[11]对交通流运行风险进行了分析; Lee 等^[3]使用集计对数线性模型(aggregate log linear model)确定可能会导致事故发生的交通流状态.上述研究中,研究人员使用了不同的建模方法对事故发生的可能性进行了研究,但其中数据采集的地点和时间以及规模都不尽相同.如 Abdel-Aty 和 Pande 等^[4-9]所做研究使用的数据是从美国奥兰多一条长为 13.2 km 的 4 号州际公路采集的 1999—2002 年中的 1 528 条事故数据;而 Oh

等^[10-11]只采用了 52 条事故数据进行分析.总体来看,上述的建模方法可分为 2 类,统计学方法、人工智能数据挖掘方法.

贝叶斯网络作为贝叶斯公式的扩展,是不确定知识表达和推理领域最有效的理论模型之一,在工程领域有着广泛的应用.而在交通事故分析、交通流量预测、交通事件检测方面,贝叶斯网络也表现出比较好的效果^[12-14].因此本文使用贝叶斯网络 BN 模型对快速路实时交通流运行风险进行评估.同时考虑到,目前各类方法缺少同样数据环境下的对比分析,本文选择了经典的朴素贝叶斯分类(naive Bayes, NB)、K 近邻(K-nearest neighbor, KNN)和反向传播神经网络(back propagation neural network, BPNN)三种分类模型进行了对比分析.

2 研究路段和数据

2.1 研究路段

本文主要研究快速路主线基本段交通运行状态对事故风险的影响.考虑到快速路的线形以及出/入口匝道的影 响,本文在上海市延安高架上选取四段线形和匝道类型基本一致的主线路段.路段均以被隔开的连续两处匝道命名,4 段主线路段分别为西向延东立交入口匝道至茂名路上匝道段,东向虹许路上匝道至娄山关路上匝道段,东向延西立交入口匝道至凯旋路上匝道段,东向江苏路上匝道至华山路 上匝道段.上述 4 个路段均为 3 车道,线圈检测器的间隔为 300~500 m.

2.2 数据

本文研究的事故数据由事故视频监控系统和事故报警后救援车辆数据中提取.由于研究需要将每条事故信息和相应地点和时间的交通检测信息对应起来,同时由于事故的偶发性特征,因此在 4 个基本路段获取大量的事故数据异常困难.经过交通流检测数据和事故数据的质量控制和无效信息剔除后,目前采用了 2010 年 7 月研究路段上发生的 71 条两车相撞或多车相撞的事故数据.每条事故数据的信息包括事故的发生时间、事故地点、事故类型、事故时的天气和位置等信息,不同地点的事故数据数量和比例见表 1.

本文使用的交通流运行数据采集于快速路上铺设的线圈检测器,为 5min 间隔的数据.采集到的交通信息包括 5min 内通过该检测器的车辆流量、所有车辆的平均速度以及平均的时间占有率.由于选取

表1 事故数据统计表

Tab.1 Summary of accident data

事故地点	编号	事故数	所占比例/%
延安高架西向延东立交入口匝道至茂名路上匝道	a	23	32.4
延安高架东向虹许路上匝道至娄山关路上匝道	b	17	23.9
延安高架东向延西立交入口匝道至凯旋路上匝道	c	11	15.5
延安高架东向江苏路上匝道至华山路上的匝道	d	20	28.2
总计		71	100

的快速路主线路段长度较短,为了降低出/入口匝道的对主线交通流的影响,每条事故数据对应的交通信息只来自于事故发生地附近的4个检测器——事故上下游最近的检测器各2个。检测器空间布置如图1所示。4个检测器也可以分为2组,一组是事故发生地上下游最近的2个检测器(图1中上游1,下游1),另外一组是所有4个检测器(图1中4个检测器)。

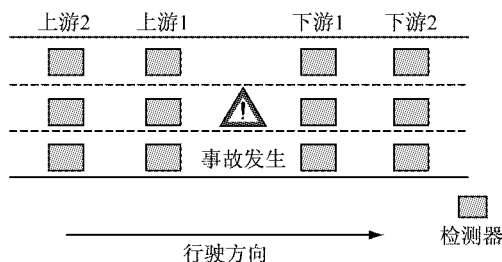


图1 事故发生点检测器示意图

Fig.1 Arrangement of detectors at the crash site

在时间轴上,事故发生前后的一段时间可以分为5个5min的时间片段^[15]。这些片段是:时间段0:事故发生后0~5 min;时间段1:事故发生前0~5 min;时间段2:事故发生前5~10 min;时间段3:事故发生前10~15 min;时间段4:事故发生前15~20 min。由于时间段0处于事故发生后,故只适合作事故检测,而非事故风险预估。此外交通管理中心必须要有一段识别事故即将发生和采取对应措施的反应时间,故只有通过利用时间段2、3和4中的交通检测信息建立的模型对主动安全管理才有实际的价值,所以下文中建立的时间段1的模型只具有参考意义。

事故数据和交通流数据获取后还需要一一对应,每条事故数据对应一系列的交通流检测数据,包括4个检测器在4个时间段内的数据。通过2组检测器和4个时间段的搭配,可以得到8种组合的输入数据。本文分别对这8组输入数据进行建模,获得

不同模型间的比较。每组的名称使用Pred X_Y这样的格式来命名,X代表代表检测器的数量,而Y代表了时间段。如Pred2_3表示使用2个检测器在时间段3内的交通检测信息来建模和预测。

由于研究是对可能发生事故和不会发生事故的交通状态进行分类,所以同时需要收集未发生事故条件下快速路的交通检测信息。与事故数据类似,非事故数据也需要有不同的交通条件,如高峰时刻和平峰时刻的交通,因此非事故数据是在不同的交通条件下随机选择的。根据Pande等^[15]的研究,为了能够基本表现实际事故发生与否的交通状态比例,同时满足本方法建模的数据需求,随机生成了非事故数据702条,使得事故数据与非事故数据的比例接近1:10。

3 贝叶斯网络

3.1 模型介绍

贝叶斯网络是一个带有概率注释的有向无环图。这个图模型能表达大的变量集合的联合概率分布(物理的或贝叶斯的),可以分析大的变量之间的相互关系,进而利用贝叶斯定理揭示的学习和统计推断功能,实现预测、分类、聚类、因果分析等数据挖掘任务^[16]。一个贝叶斯网络包括一个有向无环图(directed acyclic graph, DAG)和一个条件概率表集合。DAG中每一个节点表示一个随机变量,可以是直接观测变量或隐藏变量,而有向边表示随机变量间的条件依赖;条件概率表中的每一个元素对应DAG中唯一的节点,存储此节点对于其所有直接前驱节点的联合条件概率。

构建一个指定领域的贝叶斯网络包括三个方面:①能够描述该领域的变量及其取值范围;②结构的学习,学习变量间的依赖关系,并以图形化的方式表示出来;③参数的学习,学习变量间的分布函数,获得局部条件概率分布表。建立一个贝叶斯网络是上述三个过程迭代、反复地交互进行^[16]。

由于交通流数据可能存在缺失状况,本文使用可以有效应对缺失数据的基于高斯混合模型的最大期望(expectation-maximization, EM)算法来构造一个贝叶斯网络,高斯混合模型可以描述输入的变量,而EM算法则用于参数的学习^[17]。

3.2 高斯混合模型

高斯混合模型(Gaussian mixture model, GMM)是一种对数据的“真实分布”进行模拟和逼近

的半参数表达模型,它将数据的概率密度函数用多个高斯函数的凸函数(多个高斯函数的线性组合)来表示,在有充分样本的前提下能模拟任意形状分布^[17]. GMM是由多个子高斯分布组成的,每个子高斯分布由均值 μ_i 和协方差矩阵 \mathbf{v}_i 来决定其几何特征.可以用一组交通流数据中的不同变量表示一组 T 维矢量: $\mathbf{X}=\{x_1, x_2, \dots, x_T\}$,这组数据的分布未知,但可以由 M 个高斯密度函数的加权和来近似模拟,表达式如下:

$$p(x_i | \lambda) = \sum_{i=1}^M w_i N(x_i, \mu_i, \mathbf{v}_i) \quad (1)$$

式中: λ 为模型参数的集合, $\lambda=\{\omega_i, \mu_i, \mathbf{v}_i\}$; w_i ($i=1, \dots, n$)为混合权重,并满足: $\sum_{i=1}^M w_i = 1$; $N(x_i, \mu_i, \mathbf{v}_i)$ 为均值为 μ_i ,协方差矩阵为 \mathbf{v}_i 的高斯密度函数:

$$N(x_i, \mu_i, \mathbf{v}_i) = \frac{\exp\{-\frac{1}{2}(x_i - \mu_i)' \mathbf{v}_i^{-1} (x_i - \mu_i)\}}{(2\pi)^{d/2} |\mathbf{v}_i|^{1/2}} \quad (2)$$

使用GMM要通过学习获得参数 λ ,这就需要对GMM进行训练使得该模型性能与样本数据达到最佳的拟合,研究表明最大似然估计可以很好地进行参数估计. T 维矢量 \mathbf{X} 对 λ 的似然函数为:

$$p(x | \lambda) = \prod_{i=1}^T p(x_i | \lambda) \quad (3)$$

最大似然估计的目的是找到一组新的参数 λ^{new} 使其满足 $p(x | \lambda^{\text{new}}) \geq p(x | \lambda)$.而参数 λ^{new} 可以通过最大期望(EM)算法获得.

3.3 最大期望算法

最大期望EM算法是一种从不完整数据或有数据丢失的数据集(存在隐含变量)中求解概率模型参数的最大似然估计方法.

此时引入辅助函数:

$$Q(\lambda, \lambda^{\text{new}}) = \sum_{i=1}^T \sum_{j=1}^M p(i | x_i, \lambda) \cdot \log[w_j N(x_i, \mu_j, \mathbf{v}_j)] \quad (4)$$

式中,一组数据由 M 个高斯密度函数表示. $p(i | x_i, \lambda)$ 为由样本得到的每个子高斯分布的后验概率,并满足:

$$p(i | x_i, \lambda) = \frac{w_i N(x_i, \mu_i, \mathbf{v}_i)}{\sum_{i=1}^M w_i N(x_i, \mu_i, \mathbf{v}_i)} \quad (5)$$

EM算法中,如果 $Q(\lambda, \lambda^{\text{new}}) \geq Q(\lambda, \lambda)$,则 $p(x | \lambda^{\text{new}}) \geq p(x | \lambda)$.最后得到估计的参数如下:

$$w_i^{\text{new}} = \frac{1}{T} \sum_{i=1}^T p(i | x_i, \lambda) \quad (6)$$

$$\mu_i^{\text{new}} = \frac{\sum_{i=1}^T p(i | x_i, \lambda) x_i}{\sum_{i=1}^T p(i | x_i, \lambda)} \quad (7)$$

$$\mathbf{v}_i^{\text{new}} = \frac{\sum_{i=1}^T p(i | x_i, \lambda) (x_i - \mu_i)(x_i - \mu_i)^T}{\sum_{i=1}^T p(i | x_i, \lambda)} \quad (8)$$

EM参数学习方法包含两个步骤:① E-step:利用当前网络结构和参数计算每个子高斯分布的期望后验概率;② M-step:基于E-step参数的期望值,计算出新的最大可能的参数分布 λ^{new} ,用来替换原有的 λ ;重复过程①和②,直到所估计的参数达到指定的迭代次数或达到局部最优.

3.4 贝叶斯网络分类器构建

本文使用Matlab来实现基于高斯混合模型的EM算法,学习和训练贝叶斯网络的结构和参数,从而构建出一个合适的贝叶斯网络分类器.整个贝叶斯网络分类器的构建分为三个步骤.

首先是数据准备阶段,主要是输入事故数据与非事故数据以及相应的交通流数据,使用随机的方法确定训练样本和测试样本,为保证模型检验的可靠性,训练样本和测试样本的比例为1:1.

其次是分类器的训练,根据训练数据和基于高斯混合模型的EM算法来构建贝叶斯网络分类器.

最后是分类器的应用,使用分类器对测试数据进行分类,其输入是已经构建好的分类器的推断引擎和测试数据,输出是测试样本属于各类别的后验概率,最后根据这一概率进行分类.具体分类时如果根据测试样本属于某类别的概率较大者则属于该类别,会导致事故结果的分类准确率较低,经过观察得到的概率结果,如果确定一个合适的阈值,若该数据属于事故数据的概率大于该值,则认为是事故数据;反之,认为是非事故数据.

3.5 模型检验

本文使用的事故数据与非事故数据的比例接近1:10,相对于现实中事故的发生概率(现实中事故率小于0.001%)^[15],这个比例是比较高的,但是依然不能仅使用整体的分类率来判断模型的好坏.因为即使模型把所有事故都分类成非事故,由于事故数据相对较少,整体的分类准确率也可以高达90.8%.所以本文分别使用事故数据的分类准确率,非事故数据的分类准确率以及整体的分类准确率这3个标准来检验模型的有效性.其中事故数据的分类准确率为分类正确的事故数据数占总事故数据数的

比例;同样的,非事故数据的分类准确率为分类正确的非事故数据占总非事故数据的比例.考虑到事故的发生会带来严重的影响,本文的判断指标优先考虑模型对事故数据的分类率.

由于训练样本和测试样本是将同一数据样本随机分成的比例一样的两部分,所以每次针对同一整体数据样本建模得到的模型分类性能都不会完全一致,因此本文对 8 组数据分别进行建模时,对每组组合都分别测试 10 次.这样可以得到每组数据的模型

分类结果的平均值和标准差,以此来判断哪组数据建立的模型性能较优.

4 模型结果

基于延安西路 4 个主线路段的 71 条事故和 702 条非事故数据,对分成的 8 类交通流组合分别建立贝叶斯网络模型后分类,得到如下结果,见表 2.

表 2 不同组合下的分类结果统计表

Tab.2 Classification results of different model combinations

模型	时间段	事故分类准确率/%		非事故分类准确率/%		整体分类准确率/%	
		平均值	标准差	平均值	标准差	平均值	标准差
Pred4_1	1	71.94	10.57	75.77	2.38	75.41	1.66
Pred2_1	1	80.28	5.04	69.11	3.53	70.16	3.08
Pred4_2	2	64.72	9.79	80.17	3.70	78.73	2.65
Pred2_2	2	76.94	7.25	71.20	1.98	71.74	1.38
Pred4_3	3	72.22	12.55	73.91	4.49	73.76	3.56
Pred2_3	3	75.83	5.28	66.77	5.12	67.62	4.35
Pred4_4	4	67.50	6.22	75.40	3.42	74.66	2.71
Pred2_4	4	72.78	6.31	67.83	2.52	68.29	1.98
平均		72.78		72.52		72.55	

由表 2 可看出,贝叶斯网络分类器的总体准确率和非事故准确率比较高,与此同时事故的分类准确率也比较理想,平均为在 70%左右,最高事故分类准确率超过 80%.且结果显示模型 Pred2_1 和模型 Pred2_2 具有比较好的分类准确率.如果考虑到时间段 1 的数据的模型不具有实际价值,那么模型 Pred2_2 的事故分类率结果最好,即采用事故地点最近的 2 个检测器在事故发生前 5~10min 内的交通流数据建立的贝叶斯网络模型在分辨是否会发生事故方面具有最好的效果.

同时需要注意的是,通过贝叶斯网络建立的分类型器使用 2 个检测器的数据比使用 4 个检测器的数据对事故数据的分类效果普遍要好,而对整体分类和非事故数据的分类则显示了相反的结果.这可能是所选路段的进出匝道对主线交通流造成的干扰造成的.

此外,需要指出的是事故是否发生是根据 BN 模型计算的各交通检测信息所对应的事故发生后验概率来判断的,而判断时需要确定合适的阈值,本文为确定此值,针对模型 Pred4_1 和 Pred2_1 的数据结果做了相应的敏感性分析以得出最优的阈值.由敏感性分析结果可知,事故数据的分类准确率随着阈值的增大慢慢减小,而非事故数据的分类准确率则会增大.但当阈值大于 0.07 时,事故数据的分

类准确率降幅较大,而非事故数据的分类准确率的增幅不太明显,故选取阈值为 0.07.

5 结果对比

5.1 不同算法结果对比

为了进一步验证方法的有效性,本文选取了几种经典的分类算法与贝叶斯网络进行了对比.包括朴素贝叶斯分类(NB)、K 近邻(KNN)和反向传播神经网络(BPNN)算法.这些算法已经被广泛应用于包括交通事故预测^[18]、交通流量预测^[19]等的研究中.

基于本文所用的数据集,表 3 给出了本方法与 NB、KNN 和 BPNN 算法在分类准确率方面的比较结果.数据的准备工作均类似于第 2.2 节所述.为了简化结果显示,表中只列出了分别针对 8 组数据中(除了时间段 1 的 2 组)分类结果最好的模型组合.由于事故分类的准确性对研究的意义更大,而非事故数据的分类准确率带来的影响远小于事故数据,所以选择事故分类准确率最高的那组模型为最优模型.

由表 3 可知,贝叶斯网络和朴素贝叶斯分类器表现出了较高的事故分类准确性,而 K 近邻和 BP 神经网络算法对事故的分类效果不明显.由于事故数据的特殊性、事故数据与非事故数据容量的不平

衡以及总样本量偏少,K 近邻和 BP 神经网络算法的样本学习不够充分,无法获得较高的分类准确率.而基于贝叶斯原理的贝叶斯网络和朴素贝叶斯分类则能够比较好地处理这种数据.

表 3 不同算法的结果对比
Tab.3 Comparison of the results by several different methods

方法	事故分类 准确率/%	非事故分类 准确率/%	整体分类 准确率/%	所选数据
BN	76.94	71.20	71.74	Pred2_2
NB	67.50	77.49	76.55	Pred4_3
KNN	29.40	98.30	92.10	Pred4_2
BPNN	44.44	97.46	93.70	Pred4_4

5.2 与现有研究结果对比

Abdel-Aty 和 Pande^[6]使用 PNN 模型对事故和非事故的交通流数据进行分类,得到了 72.5%的事故分类率和 62.5%的整体分类准确率;Pande 等^[15]也根据事故发生前不同时间段以及不同数量检测器的交通流数据使用 logistic 回归和分类树确定是否会发生事故,结果表示最好的模型分别可以分辨出 61.7%和 61.9%的事故.由此可见,利用事故发生前的交通流数据对事故风险进行预测的准确率不会达到特别高的结果.

本文得出的结果优于现有其他研究的建模方法,除了 BN 模型优良的分类性能外,还可能是由于所用数据的规模大小有所不同.使用较大规模的数据可能会造成数据的过拟合,从而得不到特别好的结果.但总的来说,这种对比在一定程度上也可以反映出利用贝叶斯网络对事故风险进行预测的有效性和先进性.

6 结语

本文旨在研究快速路上实时交通流数据与交通运行事故风险的关系,利用实时快速路检测器数据可动态预估交通流运行的事故风险,从而实现主动交通安全管理,为有效缓解快速路交通安全问题提供了新的理论方法.主要结论如下:

(1) 本文利用基于高斯混合模型的 EM 算法构建的贝叶斯网络模型针对不同的检测器和时间段组合分别建立了 8 组不同的模型.结果表明使用事故发生前 5~10 min 的上下游各一个检测器的交通信息数据建立的贝叶斯网络表现最好,事故预测准确率在 76.94%.

(2) 贝叶斯网络与朴素贝叶斯分类、K 近邻、BP

神经网络算法的结果对比表明,贝叶斯网络分类算法对事故预测准确度较高,但对非事故数据也有一定的误报率.

(3) 进一步的研究将继续采集事故和相对应的交通流参数,同时考虑如何改变样本组合、输入参数组合以及测试新的分类方法来进一步提高事故数据预测的准确率.

参考文献:

[1] WANG Dianhai, CHEN Xuewen, YANG Shaohui. Analysis of traffic bottleneck on Beijing urban expressway[M]. Beijing: China Communications Press, 2007.

[2] Federal Highway Administration. Traffic congestion and reliability trends and advanced strategies for congestion mitigation [R]. Washington D C: US Department of Transportation, Federal Highway Administration, 2005.

[3] Lee C, Hellinga B, Saccomanno F. Real-time crash prediction model for application to crash prevention in freeway traffic[J]. Transportation Research Record, 2003(1840): 67.

[4] Abdel-Aty M, Uddin N, Pande A, et al. Predicting freeway crashes from loop detector data by matched case-control logistic regression[J]. Transportation Research Record, 2004, 1897: 88.

[5] Abdel-Aty M, Uddin N, Pande A. Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways[J]. Transportation Research Record, 2005, 1908: 51.

[6] Abdel-Aty M, Pande A. Identifying crash propensity using specific traffic speed conditions [J]. Journal of Safety Research, 2005, 36(1): 97.

[7] Abdel-Aty M, Pande A. Classification of real-time traffic speed patterns to predict crashes on freeways[C/CD]//83rd Annual Meeting of the Transportation Research Board. Washington D C: Transportation Research Board, 2004.

[8] Abdel-Aty M, Abdalla F M. Linking roadway geometrics and real-time traffic characteristics to model daytime freeway crashes: generalized estimating equations for correlated data [J]. Transportation Research Record, 2004, 1897: 106.

[9] Pande A, Abdel-Aty M. Assessment of freeway traffic parameters leading to lane-change related collisions [J]. Accident Analysis & Prevention, 2006, 38(5): 936.

[10] Oh C, Oh J S, Ritchie S G, et al. Real-time estimation of freeway accident likelihood[C/CD]//80th Annual Meeting of the Transportation Research Board. Washington D C: Transportation Research Board, 2001.

[11] Oh C, Oh J S, Ritchie S G. Real-time hazardous traffic condition warning system: framework and evaluation [J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(3): 265.

[12] De Oña J, Mujalli R O, Calvo F J. Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks [J]. Accident Analysis & Prevention, 2011, 43(1): 402.

- [13] Sun S, Zhang C, Yu G. A Bayesian network approach to traffic flow forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2006, 7(1): 124.
- [14] Zhang K, Taylor M A P. Effective arterial road incident detection: a Bayesian network based algorithm [J]. Transportation Research Part C: Emerging Technologies, 2006, 14(6): 403.
- [15] Pande A, Nuworsoo C, Shew C. Proactive assessment of accident risk to improve safety on a system of freeways [R]. San Luis Obispo: California Polytechnic State University, 2012.
- [16] 李艳美. 基于贝叶斯网络的数据挖掘应用研究[D]. 西安: 西安电子科技大学, 2008.
LI Yanmei. Study on application of data mining based on Bayesian networks [D]. Xi'an: Xidian University, 2008.
- [17] 陶建斌, 舒宁, 沈照庆. 基于高斯混合模型的遥感影像连续型朴素贝叶斯网络分类器[J]. 遥感信息, 2010 (2): 18.
TAO Jianbin, SHU Ning, SHEN Zhaoqing. Continuous Bayesian network classifier for remote sensing images based on improved gaussian mixture model [J]. Remote Sensing Information, 2010 (2): 18.
- [18] LÜ Y, Tang S, Zhao H. Real-time highway traffic accident prediction based on the k-Nearest neighbor method[C]//2009 International Conference on Measuring Technology and Mechatronics Automation. Zhangjiajie: ICMTMA, 2009: 547-550.
- [19] XIE Yuanchang, ZHANG Yunlong. A wavelet network model for short-term traffic volume forecasting [J]. Journal of Intelligent Transportation Systems, 2006, 10 (3): 141.

~~~~~

(上接第 858 页)

- [3] Hunt A. Wind-tunnel measurements of surface pressures on cubic building models at several scales[J]. Journal of Wind Engineering and Industrial Aerodynamics, 1982, 10(2): 137.
- [4] Raju K G R, Singh V. Blockage effects on drag of sharp-edged bodies [J]. Journal of Wind Engineering and Industrial Aerodynamics, 1975, 1(3): 301.
- [5] Awbi H B. Wind-tunnel-wall constraint on two-dimensional rectangular-section prisms[J]. Journal of Wind Engineering and Industrial Aerodynamics, 1978, 3(4): 285.
- [6] Laneville A, Trepanier J Y. Blockage effects in smooth and turbulent flows: The case of two-dimensional rectangular cylinders[J]. Journal of Wind Engineering and Industrial Aerodynamics, 1986, 22(2/3): 169.
- [7] Laneville A. Turbulence and blockage effects on two dimensional rectangular cylinders [J]. Journal of Wind Engineering and Industrial Aerodynamics, 1990, 33(1/2): 11.
- [8] Noda M, Utsunomiya H, Nagao F. Basic study on blockage effects in turbulent boundary layer flows[J]. Journal of Wind Engineering and Industrial Aerodynamics, 1995, 54/55: 645.
- [9] Parkinson G V, Cook N J. Blockage tolerance of a boundary-layer wind tunnel [J]. Journal of Wind Engineering and Industrial Aerodynamics, 1992, 42(1/3): 873.
- [10] 谢壮宁, 刘帅, 石碧青. 低矮房屋标准模型的风洞试验研究 [J]. 华南理工大学学报: 自然科学版, 2011, 39(6): 106.  
XIE Zhuangning, LIU Shuai, SHI Biqing. Investigation into wind tunnel test of standard low-rise building model [J]. Journal of South China University of Technology: Natural Science Edition, 2011, 39(6): 106.
- [11] Wang L, Liang S G, Tang H Q, et al. Investigation on wind tunnel blockage effect of super high-rise building [C]// The seventh International Colloquium on Bluff Body Aerodynamics and Applications. Shanghai: [s. n.], 2012: 565-572.
- [12] 周炬毅, 顾明. 单通道测压管路系统的优化设计 [J]. 同济大学学报: 自然科学版, 2003, 31(7): 798.  
ZHOU Xuanyi, GU Ming. Optimization of dynamic pressure measurement of single-channel tubing systems [J]. Journal of Tongji University: Natural Science, 2003, 31(7): 798.
- [13] Maskell E C. A theory for the blockage effects on bluff bodies and stalled wings in a closed wind tunnel [R]. London: Aeronautical Research Council Reports and Memoranda, 1963.
- [14] Peitzman F W. Low speed wind tunnel investigation to develop high altitude wall correction in the Northrop 7×10 foot low speed wind tunnel [R]. [S. l.]: AIAA, 1978.