

基于贝叶斯分类的公路交通事件分类方法研究

李沛祺

(广东利通科技投资有限公司 广州 510700)

摘要 高速公路监控员的主观因素往往导致交通事件现场救援组织标准的不一致,进而降低了交通事故应急响应的效率。文中提出了一种基于贝叶斯分类的交通事件分类方法。该方法使用贝叶斯算法对交通事件数据进行分类并得到每个分类的概率,通过引入条件概率,改进传统的人为判断得出救援方案的分类方法,在一定程度上降低了人为主观判断对交通事件现场救援效率的影响。实验结果表明,该方法对交通事件的分类与实际情况相符。

关键词: 公路工程;交通事件;贝叶斯分类

中图分类号 U412

Research on Highway Traffic Event Classification Method Based on Bayesian Classification

LI Peiqi

(Guangdong Leatop Technology Investment Co.,Ltd.,Guangzhou 510700,China)

Abstract Personal judgments of highway management organization staffs often lead to different standards in rescue, which in turn reduces the efficiency of emergency response to traffic accidents. This paper propose a traffic incident Bayesian classification. This method uses Bayesian algorithm to classify traffic incident data and obtain the probability of each classification. By introducing conditional probability and improving the traditional method of classifying rescue plans based on subjective judgments, it partially eliminates the impact of subjective judgments on the efficiency of on-site rescue for traffic incidents. Experimental results demonstrate that this method aligns with the actual situation of traffic incident classification.

Key words Highway engineering, Traffic incidents, Bayesian classification

0 引言

从认知交通事故层面而言,事件作为事故的前兆特征,被定义为在特定时间下某种环境下的某种特定状态。当高速公路网发生交通事件时,管理部门需从事件交通影响程度的角度掌握情况,对事件关联路段进行流量管控与疏导,以避免拥堵的发生及扩散。在粤港澳大湾区的高速公路网络中,横贯珠江口、连通两岸的几条跨江通道堪称整个网络的核心要道,除了已经通车的虎门大桥、南沙大桥、黄埔大桥、港珠澳大桥,还有计划2024年建成通车的深中通道和黄茅海大桥以及计划2028年建成通车的狮子洋通道。有研究表明,过江通道对于网络化空间构建具有非常重要的意义^[1]。

珠江口公路跨江通道多数已处于饱和状态,高峰期、节假日交通较为拥堵。大多数学者从规划设计角度出发,结合区域高速公路网结构,对环珠江口两岸区域跨江通道的规划提出布局策略,如刘翰宁等^[2]、黄俊^[3]、黄宏敏等^[4]提出了适应分段定供给、差异化定功能的发展建议。但跨江通道从规划到建设往往需要几年时间,从现实意义出发,按

交通事件的影响程度对其进行合理分类,以指导相应现场救援应急方案的研究非常必要。

1 交通事件分类研究及发展现状

交通事件是高速公路监控管理的核心载体,包括计划性事件和突发事件,其中国内外研究对交通事故的分类方法比较一致。美国联邦公路局FHWA编写的《高速公路管理和运营手册》从通行能力、运行管理、应急管理3个角度进行了定义,分别为引起通行能力降低或者需求增加的活动、道路上发生的阻碍正常交通流量的活动;对正常交通运营产生不利影响的、未计划的、随机发生的活动;对公共安全或生命财产损失有影响的或有潜在影响的交通运输活动^[5]。

国内有学者从交通事故造成人员伤亡或车辆损坏的程度,对交通事件进行分类^[6-7],这与公安机关交通管理部门以交通事故的严重程度将交通事故分为轻微事故、一般事故、重大事故、特重大事故4种类型^[8]的思路是一致的。然而,高速公路运营单位对交通事件的管理比交警部门更加

作者简介:李沛祺(1994—),本科,助理工程师,研究方向为智慧交通。

精细,若直接套用此分类方法,则难免存在以小包大、以偏概全的问题。2022年12月施行的《广东省高速公路运营管理办法》将交通事件分为自然灾害、交通事故、公共卫生事件^[9]。有学者从高速公路运营管理的角度出发,研究高速公路交通事件处置等级划分^[7]。高速公路应急指挥调度围绕交通事件的演变趋势展开,在事件的演化过程中采取人为干预,如根据应急预案提前通知交警、路政、拯救、医疗、消防。

本文提出的交通事件分类依据参考广东省某高速公路运营单位印发的《突发事件应急管理办法》,交通事件分为9大类、100小类,如图1所示。交通事件包括计划性交通事件和突发性交通事件,其中计划性事件分为两大类,包括16小类,突发性事件7大类,包括84小类。

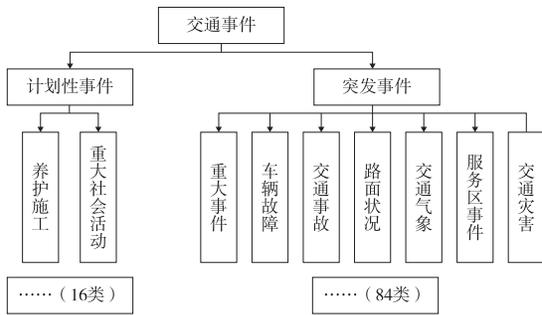


图1 交通事件分类

2 基于朴素贝叶斯的交通事件分类方法

2.1 朴素贝叶斯算法的应用

朴素贝叶斯分类广泛应用于分类算法中,主要基于贝叶斯定理和属性条件独立性假设来计算概率,属于有监督学习算法。在现实生活中,朴素贝叶斯算法被广泛应用于文本分类、情感分析、信用评估、医学诊断、推荐系统等领域。贝叶斯分类通过先验概率以及特征之间的条件概率(或其他相关概率值)来计算后验概率值,从而进行分类。

2.2 使用条件概率进行分类

给定某个由 B_1, B_2, \dots, B_n 属性表示的数据点,通过计算数据点来自类别 A_i 的概率,应用贝叶斯准则可以得到式(1):

$$p(c_i|x,y) = \frac{p(x,y|c_i)p(c_i)}{p(x,y)} \quad (1)$$

其中, x, y 表示特征变量, c_i 表示分类, $p(c_i|x,y)$ 即表示在特征为 x, y 的情况下分入类别 c_i 的概率。因此,结合条件概率和贝叶斯定理可以得出如下结论:

如果 $p(c_1|x,y) > p(c_2|x,y)$, 则分类属于 c_1 ;

如果 $p(c_2|x,y) > p(c_1|x,y)$, 则分类属于 c_2 。

可以通过计算其他3个概率值得出结果。由于朴素贝叶斯分类算法假设特征之间相互独立,因此在事件分类时可以利用特征间的独立性简化计算公式。该算法简单易懂,在事件分类问题中的分类效率和效果相对稳定。

2.3 模型选用分析

本文选用朴素贝叶斯分类算法对交通事件进行分类,考虑了以下问题。(1)朴素贝叶斯模型的原理植根于古典数学理论,确保了其分类效率的稳定可靠。(2)其在小规模数据上表现优异,且能有效应对多分类任务。鉴于交通事件数据难以通过公开渠道获取,朴素贝叶斯算法的特性使其更适合处理此类数据。(3)其对缺失数据不太敏感,针对每个项目通常也只会相对较少的特征数,交通事件的特征属性并不明显,对不同类型事件的应急处理方案没有统一标准。(4)算法比较简单,对结果解释容易理解。综上,本文合理选用朴素贝叶斯分类算法对交通事件进行分类。然而,尽管在理论上属性之间相互独立,但通常情况下这一假设并不成立,因此有学者通过考虑关联性对朴素贝叶斯算法进行改造。

3 实验分析

3.1 问题描述

本文以2021年6月至7月发生在某路段路面的历史交通事件作为训练样本数据集,剔除未通知拯救、急救或消防的轻微事件,得到的训练集如表1所列。

表1 交通事件数据库(训练样本集合)

事件id	事件位置	占用车道	通知拯救	通知医疗	通知消防	事件分类
1	沙田互通-海鹏岛收费站	无车道占据	否	否	否	车流量大
2	东涌互通-海鹏岛收费站	占据路肩	是	否	否	追尾
3	海鹏岛收费站-沙田互通	占据第三车道	是	是	是	追尾
4	沙田互通-海鹏岛收费站	无车道占据	否	是	否	车流量大
5	东涌互通-海鹏岛收费站	无车道占据	否	否	否	车流量大
6	海鹏岛收费站-沙田互通	无车道占据	否	否	否	车流量大
7	东涌互通-海鹏岛收费站	无车道占据	否	否	否	车流量大
8	沙田互通-海鹏岛收费站	占据路肩	是	是	是	追尾
9	沙田互通-海鹏岛收费站	无车道占据	否	否	否	车流量大
10	东涌互通-海鹏岛收费站	无车道占据	否	否	否	车流量大
11	沙田互通-海鹏岛收费站	占据路肩	是	是	是	翻车
12	沙田互通-海鹏岛收费站	占据路肩	是	否	否	追尾

(续表1)

13	沙田互通-海鸥岛收费站	无车道占据	否	否	否	车流量大
14	沙田互通-海鸥岛收费站	无车道占据	否	否	否	车流量大
15	沙田互通-海鸥岛收费站	占据第三、四车道、路肩车道	是	否	是	抛锚
16	沙田互通-海鸥岛收费站	无车道占据	否	否	否	车流量大
17	东涌互通-海鸥岛收费站	无车道占据	否	否	否	车流量大
18	东涌互通-海鸥岛收费站	无车道占据	否	否	否	车流量大

将字符串数据转换为整型,转换后的数据(保留后6列)如表2所列。

表2 转换后的训练样本集合

event_location	lane_occupied	rescue	emergency	firemen	event_type
1	1	0	0	0	1
2	2	1	0	0	2
3	3	1	1	1	2
1	1	0	1	0	1
2	1	0	0	0	1
3	1	0	0	0	1
2	1	0	0	0	1
1	2	1	1	1	2
1	1	0	0	0	1
2	1	0	0	0	1
1	2	1	1	1	3
1	2	1	0	0	2
1	1	0	0	0	1
1	1	0	0	0	1
1	3	1	0	1	3
1	1	0	0	0	1
2	1	0	0	0	1
2	1	0	0	0	1

然后,将转换后的数据导入 Matlab 中。

3.2 结果分析

假设在沙田互通-海鸥岛收费站区间发生一起两辆货车追尾事故,现场占据路肩,已通知拯救人员到场处理。输入样本[1,2,1,0,0],可得到如下的运行结果:

请输入类别未知的样本。示例:[1,2,1,1,1]

[1,2,1,0,0]

计算 $p(C_i)$ 的结果为:

$p_c=0.6667 \quad 0.2222 \quad 0.1111$

计算 $p(X|C_i)$ 的结果为:

$p_{xc}=0 \quad 0.0938 \quad 0$

计算 $p(C_i)*p(X|C_i)$ 的结果为:

$p=0 \quad 0.0208 \quad 0$

结论:当前类别未知的样本属于类别2的概率最大,其概率为0.020833。

求得事件分类为“追尾”的概率最大,与事件假设条件相符。

4 结语

本文对交通事件进行了科学合理地分类分级,有助于更好地处理事故,更有针对性地制定处理方案,可以在实际事件处理中辅助执法者判断事故责任,提高事故处理的效率和准确性,分析事故原因和规律,以预防类似事故的再次发生。通过统计各类交通事件的发生频率和影响程度,可以更有针对性地制定交通管理措施和规划方案,提高交通系统的安全性和效率。本文提出使用朴素贝叶斯分类算法来计算条件概率,改进传统的人为判断,从而得出救援方案的分类方法。实验结果证明,该分类方法具有有效性。此外,可以通过文本分析挖掘交通事件的信息以及事件之间的关联信息,进一步提高预测精度。

参考文献

- [1] 陈朝萌. 粤港澳大湾区物流网络发展研究[J]. 财富时代. 2019(8):149-153.
- [2] 刘翰宁,郭宝山,王伯文,等. 珠江口道路过江通道布局规划思考[J]. 交通与港航. 2021,8(4):46-50.
- [3] 黄俊. 粤港澳大湾区跨珠江通道布局优化探究[J]. 铁道建筑技术. 2021(8):93-97.
- [4] 黄宏敏,邹峻,赵国锋. 广州市骨架路网提升改善策略[J]. 城市道桥与防洪. 2022(8):10-12.
- [5] 郭敏. 什么是交通事件? 如何做好交通事件管理以降低二次事故发生概率?[EB/OL]. <https://mp.weixin.qq.com/s/X3zRT0CfRcg2XKfj-xWwQ>, 2019-03-11.
- [6] 崔瑾,赵海涛,张程玮. 高速公路隧道交通事件分类及响应措施[J]. 公路. 2022,67(7):321-325.
- [7] 陈元元,杜世贝,冯松宝,等. 高速公路交通事件评级机制研究[J]. 洛阳理工学院学报(自然科学版). 2022,32(2):46-52.
- [8] 广东省道路交通事故当事人交通违法行为分类表[S].
- [9] 广东省高速公路运营管理办法[EB/OL]. http://www.gd.gov.cn/zwgk/wjk/qbwj/yfl/content/post_4010733.html, 2022-09-09.