



交通运输系统工程与信息

Journal of Transportation Systems Engineering and Information Technology

ISSN 1009-6744, CN 11-4520/U

《交通运输系统工程与信息》网络首发论文

题目：基于数据驱动的交通事故伤害程度影响因素及其耦合关系研究
作者：胡立伟，吕一帆，赵雪亭，薛宇，张成杰，雷国庆，刘凡
收稿日期：2022-05-17
网络首发日期：2022-07-25
引用格式：胡立伟，吕一帆，赵雪亭，薛宇，张成杰，雷国庆，刘凡. 基于数据驱动的交通事故伤害程度影响因素及其耦合关系研究[J/OL]. 交通运输系统工程与信息. <https://kns.cnki.net/kcms/detail/11.4520.u.20220721.1715.004.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于数据驱动的交通事故伤害程度影响因素及其耦合关系研究

胡立伟*, 吕一帆, 赵雪亭, 薛宇, 张成杰, 雷国庆, 刘凡
(昆明理工大学, 交通工程学院, 昆明 650500)

摘要: 为准确识别影响山区高速公路交通事故伤害程度(TAIDME)的相关因素, 构建了随机森林朴素贝叶斯-耦合度模型(RFNB-CDM)对其进行研究。首先将 2016 年-2020 年云南省 1760 起山区高速公路事故数据进行处理, 进行综合考量后将涉及事故信息、道路信息、肇事机动车辆信息及驾驶人信息等 4 类的 18 个相关因素作为初始特征进行研究, 使用 RF 模型进行特征提取, 并得到各因素对于山区高速公路交通事故严重程度(TASME)的重要性排序; 其次将新特征输入 NB 模型, 对 TAIDME 的影响因素进行单因素分析; 为改进原有模型不能对影响因素之间的关系进行准确刻画缺点, 本文引入耦合度模型并进行实例验证分析。结果表明: RFNB 模型相较于 RF 和 NB 模型, 得到的预测结果更精确, 分类性能分别提升 5.56%和 14.79%。其中, 追尾碰撞、18:00-次日 6:00、事故车辆数 2 辆、下坡段、夜间无路灯照明、货运、大中型货车和直行匀速这 8 类因素存在时更易加重 TAIDME, 追尾碰撞和直行匀速这两类因素发生耦合作用时, 最易导致重大伤害事故; 道路表面干燥、路侧金属防护和中央绿化带隔离这 3 类因素存在时可降低 TAIDME, 路侧金属防护和中央绿化带隔离这两类因素发生耦合作用时 TAIDME 最低。研究结论可为山区高速公路交通事故预防及降低山区高速公路事故发生后的伤害程度提供理论依据与决策参考。

关键词: 交通工程; 影响因素; RFNB-CDM; 山区高速公路; 事故伤害程度
中图分类号: U491.31 **文献标志码:** A

Study on Influence Factors and Coupling Relationship of Traffic Accident Injury Degree Based on Data Driven

Hu Liwei , Lü Yifan , Zhao Xueting , Xue Yu , Zhang Chengjie , Lei Guoqing , Liu Fan
(School of Traffic Engineering, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to accurately identify the relevant factors affecting the traffic accident injury degree of mountainous expressway (TAIDME), a model named random forest naive bayes – coupling degree model (RFNB-CDM) was constructed. Firstly, 1760 accident data of mountainous expressway in Yunnan Province from 2016 to 2020 were processed. After comprehensive consideration, 18 related factors including accident information, road information, accident motor vehicle information and driver information were studied as initial features. RF model was used for feature extraction, and the importance ranking of each factor for the severity of traffic accidents (TASME) of mountainous expressway was obtained. Secondly, the new features are input into the NB model to conduct a single factor analysis on the influencing factors of TAIDME. In order to improve the shortcomings of the original model that cannot accurately describe the relationship between the influencing factors, this

收稿日期: 2022-05-17 修改日期: 2022-07-04 录用日期: 2022-07-06
基金项目: 国家自然科学基金/National Natural Science Foundation of China (61863019)。
作者简介: 胡立伟 (1978-), 男, 山东潍坊人, 教授, 博士。 *通信作者: liweihukm@kust.edu.cn

paper introduces the coupling degree model to make an example verification analysis. Among them, rear-end collision, 18:00 - the next day 6:00, the number of accident vehicles 2, downhill, no street lighting at night, freight, large and medium-sized trucks and straight uniform these eight kinds of factors are more likely to increase TAIDME, rear-end collision and straight uniform velocity these two kinds of factors coupling, the most likely to lead to major accidents. Road surface dryness, road side metal protection and central green belt isolation can reduce TAIDME, and when road side metal protection and central green belt isolation are coupled, the TAIDME is the lowest. The conclusion of this study can provide theoretical basis and decision-making reference for the prevention of traffic accidents and the reduction of the injury degree of mountain highway accidents.

Key words: traffic engineering; influencing factors; RFNB-CDM; mountain expressway; accident injury degree

0 引言

近几年,随着我国西南地区基础设施建设逐步前进,山区高速公路里程不断增加,在便捷出行的同时,也诱发了更多的交通事故。相关数据表明,由于山区高速公路道路线形复杂,行驶货车较多,因此发生交通事故后对车内人员造成的事故伤害程度也更加严重。因此,对山区高速公路交通事故伤害程度(Traffic Accident Injury Degree of Mountainous Expressway, TAIDME)的影响因素进行提取分析,可以有效保障山区高速公路的行车安全。

在交通事故严重程度的研究方面,国内外学者已经做了大量研究,但这些研究主要集中于城市道路或者一般的高速公路,对 TAIDME 的影响因素的研究并不多。Kateřina 等^[1]研究发现,男性和女性、不同年龄段之间(18-25 岁、25-64 岁及 65 岁以上)的交通事故成因存在显著差异;Lee 等^[2]研究发现,较大的降雨强度、较长的排水长度和较大的道路水深对事故严重程度有显著影响;戢晓峰等^[3]研究发现了影响二级公路事故规模最为关键的 4 个交通流参数;孙轶轩等^[4]研究发现,碰撞类型、道路属性、事故致因和驾驶人类别等因素对交通事故严重程度分类的影响较大;赵玮等^[5]研究发现,冰雪环境下本地驾驶员对自身驾驶能力的错误估计更易导致交通事故。这些研究可以为本文对 TAIDME 的影响因素分析研究提供理论依据。

如今,机器学习广泛运用于交通事故及其严重程度的分析与预测,但这些机器学习方法普遍有各自的局限性。房锐等^[6]证明了随机森林(Random Forest, RF)模型对于交通冲突可能性的预测准确率较高;李英帅等^[7]运用 RF 模型研究了电动自行车骑行者交通事故伤害程度影响因素的重要度排序;张轮等^[8]通过仿真实验表明,朴素贝叶斯(Naive Bayes, NB)分类器适用于高速公路交通事件的检测系统;Fang 等^[9]研究贝叶斯网络在交通事故严重性预测方面的性能,将其与回归模型预测的拟合度对比,发现了贝叶斯网络优于回归模型;张军等^[10]通过 AdaBoost 分类器反应不同交通状态,从而判断是否危险,实现实时交通事故预测;Athanasios^[11]利用 RF 模型和 Bayes 逻辑回归结合雅典城市的实时交通和天气数据进行研究,从而制定相应的交通战略以降低事故严重程度和发生可能性;张文钧等^[12]提出了一种双层 Bayes 模型:随机森林朴素贝叶斯(Random Forest Naive Bayes, RFNB),使用 RF 模型将原始特征进行降维,从而构建具有高分类性能的 NB 分类模型,实验证明,RFNB 模型明显优于现有的最先进的朴素贝叶斯分类模型和其他传统分类模型。

综上,针对国内外对目前对于 TAIDME 研究较少的现状,本文以云南省山区高速公路为例,综合事故信息、机动车驾驶人信息、道路信息及车辆信息等因素,运用随机森林朴素贝叶斯-耦合度模型(Random Forest Naive Bayes- Coupling Degree Model, RFNB-CDM)鉴别各影响因素对山区高速公路机动车交通事故伤害的影响程度,并对各因素的耦合关系进行了研究,改进了 RFNB 模型无法对影响因素之间的关系进行准确刻画的缺点,可为相关部门在预防山区高速公路交通安全事故的决策中提供理论支持。

1 数据说明与数据预处理

1.1 数据说明

本文数据源自 2016-2020 年云南省发生的共 1760 起山区高速公路事故数据,为研究山区高速公路交通事故伤害程度影响因素,在剔除不符合本文要求以及含有缺失值的数据后保留 1193 条作为研究对象。原始数据统计了事故时间、事故形态、事故参与车数、平面线形、纵断面线形、照明条件、道路表面状况等 23 类事故相关因素。

1.2 数据预处理

(1) 因变量设置

参考《道路交通事故信息调查》中对交通事故的分类与分级,根据本文对于交通事故伤害程度的研究目的,结合本文山区高速公路事故数据,最终确定将轻微事故、一般事故、严重事故及特重大事故共 4 类作为山区高速公路交通事故严重程度(Traffic Accident Severity of Mountainous Expressway, TASME)的划分,具体划分如下表 1 所示。

表 1 TASME 划分

Table 1 Classification of TASME

事故严重程度	事故性质	划分标准
I	轻微事故	死亡人数为 0, 且受伤人数为 0~1 人
II	一般事故	死亡人数为 0, 且受伤人数为 2~5 人
III	严重事故	死亡人数为 1~2 人, 或受伤人数为 6~9 人
IV	特重大事故	死亡人数>2 人, 或受伤人数>9 人

本文按照事故有无死亡结合受伤人数来划分 TAIDME, 将轻微事故和一般事故等未造成人员死亡的事故统称为非重大伤害事故, 将严重事故和特重大事故等造成人员死亡的事故统称为重大伤害事故, 具体划分如下表 2 所示。

表 2 TAIDME 划分

Table 2 Classification of TAIDME

事故伤害程度	事故性质	划分标准
I	非重大伤害事故	死亡人数为 0, 且受伤人数为 0~5 人
II	重大伤害事故	死亡人数≥1, 或受伤人数为>6 人

经过对事故数据统计分析后,得到 2016-2020 年云南省 TAIDME 的分布情况,如图 1 所示。

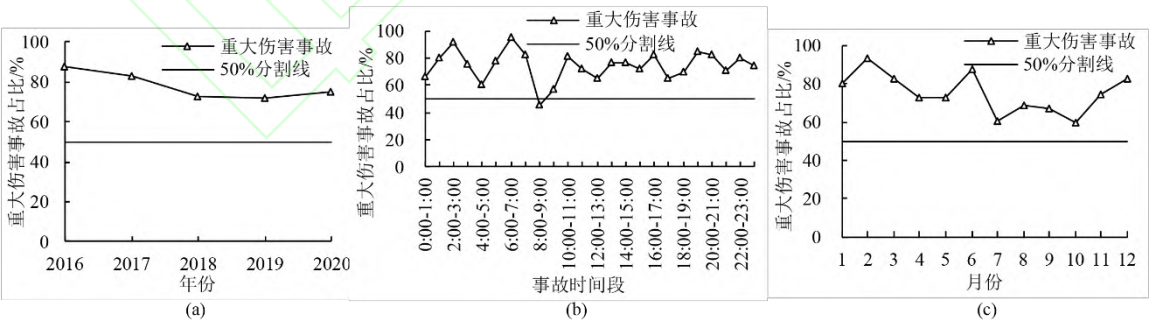


图 1 TAIDME 分布情况

Fig. 1 Distribution of TAIDME

由图 1 可知, 2016-2020 年云南省山区高速公路发生交通事故时造成的重大伤害事故明显高于非重大伤害事故; 2016 年、2017 年重大伤害事故占比更高; 每年 2 月、6 月和 12 月发生事故时造成的伤害更为严重; 重大伤害事故的时刻分布在 2:00~3:00、6:00~7:00、10:00~11:00、13:00~14:00、16:00~17:00、19:00~20:00 和 22:00~23:00 呈现 7 个高峰。

(2) 自变量设置

如图 2 所示，在对事故发生时人、车、路、环境以及包括事故本身在内的各方面信息进行综合考量后，选取 18 个因素作为初始自变量，并将其分类为事故信息、道路信息、肇事机动车车辆信息及驾驶人信息 4 类，具体划分如表 3 所示。

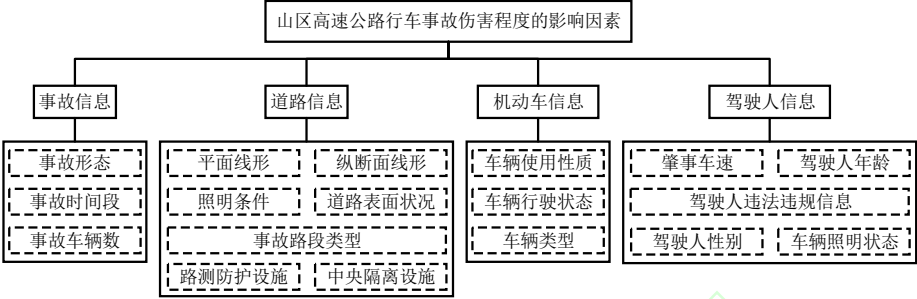


图 2 TAIDME 的影响因素

Fig. 2 Influencing factors of TAIDME

表 3 TAIDME 的影响因素划分

Table 3 Division of influencing factors of TAIDME

类别	变量名	分类	赋值	类别	变量名	分类	赋值	
事故信息	事故时间段	6:00-12:00	1	肇事机动车车辆信息	中央隔离设施	金属防护	2	
		12:00-18:00	2			混凝土防护	3	
		18:00-次日 6:00	3			绿化带	4	
	事故形态	正面碰撞	1			无防护	5	
		追尾碰撞	2			车辆使用性质	私用	1
		侧面碰撞	3	客运	2			
		侧翻	4	货运	3			
		刮撞行人	5	肇事机动车驾驶人信息	车辆行驶状态		直行匀速	1
		其他	6				直行加速	2
	事故车辆数	1 辆	1			直行减速	3	
		2 辆	2			变道及转弯	4	
		≥3 辆	3			静止及停车	5	
平面线形	直线段	1	其他			6		
	弯道段	2	车辆类型		大中型客车	1		
	纵断面线形	上坡段			1	大中型货车	2	
		下坡段			2	小型客车	3	
		无坡度			3	小型货车	4	
照明条件	白天	1			摩托车	5		
	夜间有路灯照明	2	肇事车速		无法计算	1		
	夜间无路灯照明	3			<60km/h	2		
	黄昏或黎明	4			60-100km/h	3		
道路表面状况	干燥	1			≥100km/h	4		
	潮湿	2	驾驶人违法违规行为	有违法违规行为	1			
	结冰	3		无违法违规行为	2			
事故路段类型	普通路段	1	驾驶人年龄	(18, 25]岁	1			
	特殊路段	2		(26, 64]岁	2			

路侧防护设施	不明	1	驾驶人性别	≥65 岁	3
	金属防护	2		男	1
	混凝土防护	3		女	2
	植物防护	4	车辆照明状态	打开	1
	无防护	5		未打开	2
	不明	1		其他	3

2 随机森林朴素贝叶斯-耦合度模型构建

随机森林朴素贝叶斯-耦合度模型(RFNB-CDM)的运行过程分为三阶段：第一阶段为构建 RF 模型对于原始特征进行特征选择，得到各特征对于 TAIDME 的重要度排序，并筛除不重要的特征，以提升模型分类性能；第二阶段为在 RF 特征选择的基础上构建 NB 模型，分析并获得对 TAIDME 影响较大的因素；第三阶段为基于 RFNB 构建耦合度模型，最终得到上述因素影响 TAIDME 的耦合关系。模型运行的具体流程如图 3 所示。

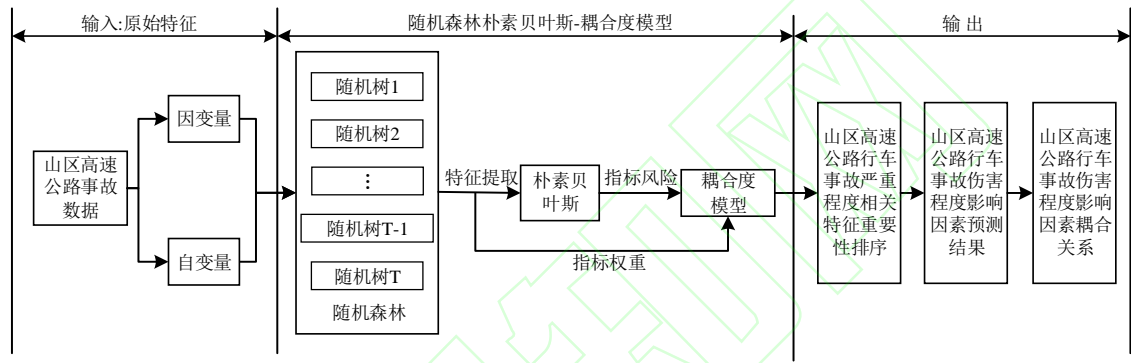


图 3 RFNB-耦合度模型运行示意图

Fig. 3 Operation diagram of RFNB model

2.1 随机森林构建

RF 模型^[14]是一种机器学习算法，是由多颗决策树组成的一种组合分类模型，多用于分类、回归以及降维，且精度较好。将 RF 作为第一层模型进行特征提取，以捕获单词组合的高层特征，在 RStudio 中建立 RF 模型，模型训练目标为事故伤害程度，为尽可能降低误差，需求出最佳分割点 k_w 的值，运用式

$$k_w = \sqrt{m} \quad (1)$$

式中： m 的值为数据集变量个数。

由式(1)得出 k_w 的值为 4，经实验筛选得到最佳 ntree 值为 620。将事故伤害程度设置为目标变量进行训练，将事故数据分为训练集与测试集，并将它们七三分，缺失值设置为 na.pass，进行 RF 模型构建。

2.2 朴素贝叶斯模型构建

本文使用 NB 模型作为 RFNB 模型的第二层，NB 算法是 bayes 算法中实用性较高的一种算法，具有较好的统计学习能力。

假设 C 为预测特征属性 X 的类， $D = \{(X_1, C_1), (X_2, C_2), \dots, (X_N, C_N)\}$ 为数据集，根据贝叶斯定理，可知最高后验概率，当后验概率 $P(C_i | X_i)$ 的值大于阈值 λ 时，则判定为重大伤害事故发生，定义为 1，反之为不发生，定义为 0。

$$Y = \begin{cases} 1, & \frac{P(X_i|C_i)P(C_i)}{P(X_i)} \geq \lambda \\ 0, & \frac{P(X_i|C_i)P(C_i)}{P(X_i)} < \lambda \end{cases} \quad (2)$$

式中：Y 的值为重大伤害事故是否发生的判定结果。

通过归一化加权平均方法，C 对于 TAIDME 的影响程度可以定义为

$$Z_i = \sum_{i=1}^N P(X_i|C_i)P(C_i|X_i) \quad (3)$$

式中：Z_i 的值为 C_i 对于 TAIDME 的影响程度。

2.3 基于 RFNB 的 TAIDME 耦合模型构建

由 RF 确定各级指标的重要性系数 S，S 的值越大，说明该指标在 TAIDME 的影响因素中越重要，因此，可由下式求出其指标权重

$$w_{ij} = \frac{S_{ij}w_i}{\sum_i \sum_j S_{ij}}, \sum_j w_{ij} = 1 \quad (4)$$

式中：w_i 为一级指标权重，w_{ij} 为二级指标权重。

利用 NB 得出各因素的后验概率 P(C_i|X_i)，通过全概率公式确定各类二级风险指标对 TAIDME 的功效值 U_{ij}，则各类指标对整个 TAIDME 的有序贡献度 U_i 为

$$U_i = \sum_{j=1}^m w_{ij}U_{ij} \quad (5)$$

可求出各指标的耦合度 C_m 为

$$C_m = \left[\frac{U_1 U_2 U_3 \cdots U_m}{\prod (U_i + U_j)} \right]^{\frac{1}{m}}, C_m \in [0,1] \quad (6)$$

其中，当 C_m ∈ [0,0.3] 时，为低耦合；当 C_m ∈ (0.3,0.7] 时，为中耦合；当 C_m ∈ (0.7,1] 时，为强耦合。耦合度越高，代表耦合作用后导致 TAIDME 越高。

3 案例应用与分析

3.1 基于 RF 的 TAIDME 特征选择

将数据集输入 RF 模型，可得到按模型平均基尼指数^[15]降序(Mean Decrease Gini)进行排列的各特征变量的重要度排序；为提取重要度较高的特征变量，需对较为冗余的特征变量进行筛选，采用后向搜索方法，每次去除一个重要性最低的特征变量，最后获得一个特征变量最少、分类精度最高的特征集作为本文最终的特征指标体系，并采用 10 折交叉验证进行验证以保证模型稳定性。RF 特征选择结果如图 4 所示。最终所提取出的事故伤害特征指标集为：{事故形态，纵断面线形，事故车辆数，行驶状态，路侧防护设施，车辆类型，路表面状况，中央隔离设施，事故时间段，车辆使用性质，照明条件}。

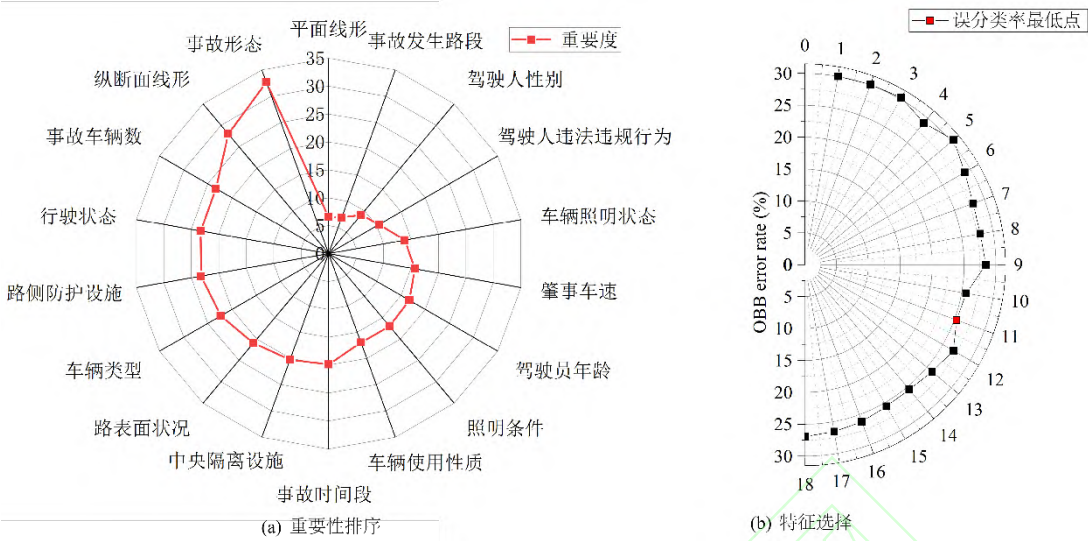


图 4 RF 特征选择结果

Fig. 4 RF Feature Selection Results

3.2 基于 NB 的 TAIDME 单因素影响程度分析

将新的特征指标集输入 NB 模型，可得各影响因素的后验概率 $P(C_i|X_i)$ ，并得到其对 TAIDME 的影响程度 Z_i ，如表 4 所示。

表 4 TAIDME 各影响因素的影响程度

Table 4 Influence Degree of TAIDME Factors

变量名称	分类	影响程度 /%	变量名称	分类	影响程度 /%
事故时间段	6:00-12:00	8.92	路侧防护设施	金属防护	66.16
	12:00-18:00	12.37		混凝土防护	1.32
	18:00-次日 6:00	46.78		植物防护	0.34
事故车辆数	1 辆	18.70		无防护	0.42
	2 辆	46.83	中央隔离设施	不明	0.04
	≥3 辆	28.88		金属防护	2.10
事故形态	正面碰撞	1.70		混凝土防护	0.48
	追尾碰撞	52.14		绿化带	64.38
	侧面碰撞	2.17		无防护	4.90
	侧翻	3.91	车辆使用性质	私用	11.73
	刮撞行人	8.36		客运	9.80
纵断面线形	其他	2.47		货运	57.59
	上坡段	27.86	车辆类型	大中型客车	0.20
	下坡段	54.72		大中型货车	42.87
照明条件	无坡度	11.91		小微型客车	33.33
	白天	27.59		小微型货车	2.73
	夜间有路灯照明	4.10		摩托车	0.21

道路表面状况	夜间无路灯照明	39.99	车辆行驶状态	直行匀速	68.46
	黄昏或黎明	29.56		直行加速	0.11
	干燥	69.44		直行减速	6.03
	潮湿	15.10		变道及转弯	0.54
	结冰	1.32		静止及停车	0.58
	不明	0.17		其他	0.04

相比于传统的 RF 模型(模型一)与 NB 模型(模型二)，本文采用的 RFNB 模型(模型三)分类性能更好、准确率更高，分类性能分别提升了 5.56%和 14.79%，三类模型分类结果对比情况如表 5 所示。

表 5 模型分类结果比较

Table 5 Model classification results

模型名称	分类精度/%	性能提升/%
模型一	74.79	5.56
模型二	68.78	14.79
模型三	78.95	—

3.3 基于 RFNB 的 TAIDME 耦合分析

根据表 4，将 18:00-次日 6:00、事故车辆数 2 辆、追尾碰撞、下坡段、夜间无路灯照明、道路表面干燥、路侧金属防护、中央绿化带隔离、货运、大中型货车和直行匀速等 11 类对 TAIDME 影响较高的因素作为事故伤害程度耦合的对象，分为事故信息、道路信息和肇事机动车辆信息三类，通过耦合度模型可确定各因子的指标权重及功效值，如表 6 所示。

表 6 TAIDME 的各因素风险指标权重及功效值

Table 6 Weight and Efficacy Value of Risk Indicators for TAIDME

类别	风险要素	编号	w_i	w_{ij}	U_{ij}
事故信息 A_1	追尾碰撞	1		0.093	0.521
	18:00-次日 6:00	2	0.280	0.071	0.512
	事故车辆数 2 辆	3		0.116	0.468
道路信息 A_2	下坡段	4		0.205	0.547
	夜间无路灯照明	5		0.062	0.400
	道路表面干燥	6	0.491	0.066	0.694
	路侧金属防护	7		0.081	0.662
	中央绿化带隔离	8		0.077	0.644
肇事机动车辆信息 A_3	货运	9		0.072	0.576
	大中型货车	10	0.229	0.064	0.429
	直行匀速	11		0.093	0.685

注： w_i 与 w_{ij} 分别为一级指标和二级指标的权重值。

通过上述结果，可由耦合度模型得到 TAIDME 致因耦合的单因素耦合、双因素耦合模型的耦合度，TAIDME 影响因子耦合拓扑分析结果如图 5 所示。

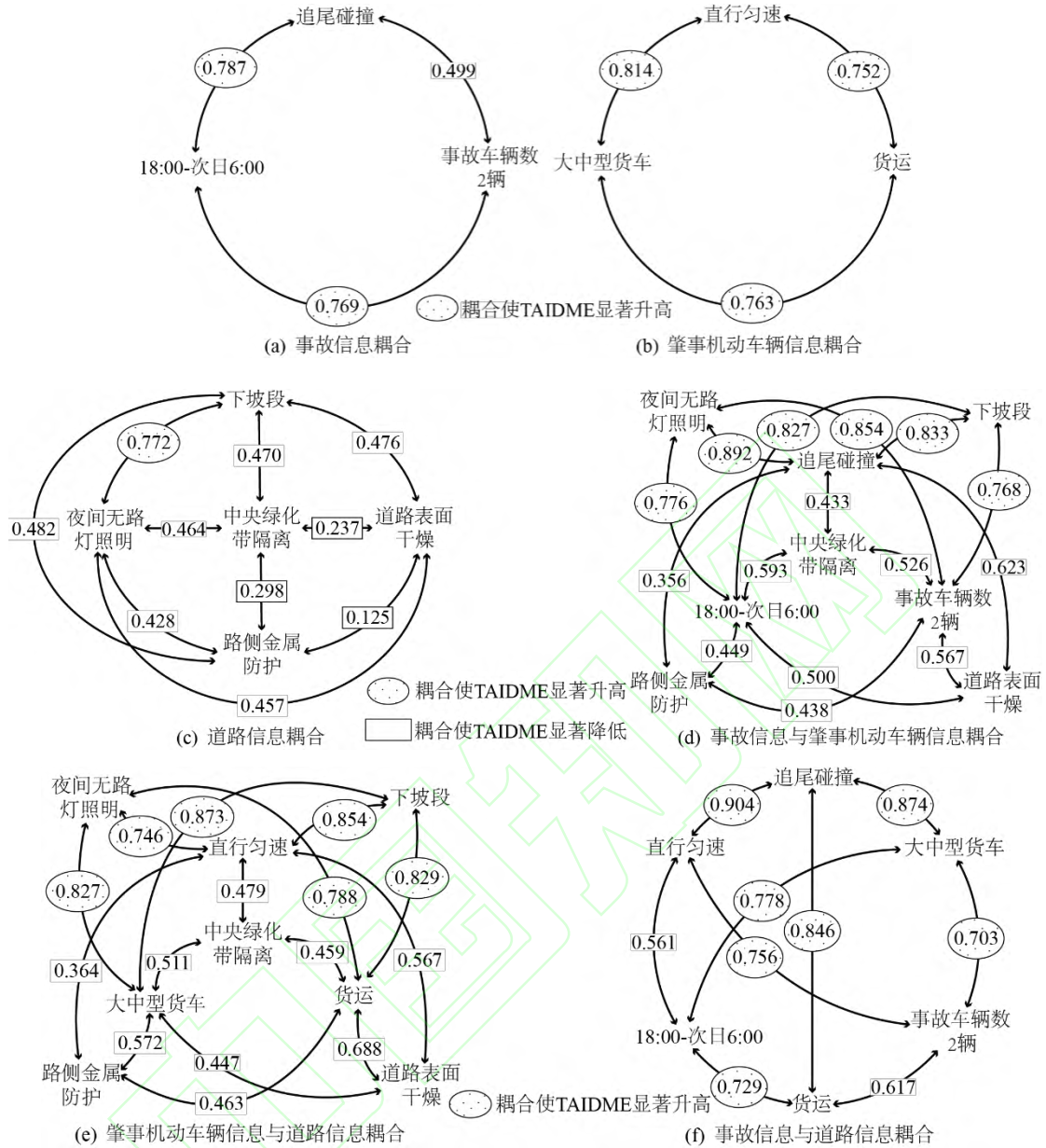


图 5 TAIDME 影响因素的单、双因素耦合拓扑分析结果

Fig. 5 Topological Analysis Results of TAIDME's Influence Factors were Coupled by Single and Double Factors

由图 5 可知，事故发生时存在追尾碰撞、18:00-次日 6:00、事故车辆数 2 辆、下坡段、夜间无路灯照明、货运、大中型货车和直行匀速这几类因素时，TAIDME 高，当直行匀速与追尾碰撞发生耦合作用时，TAIDME 最高；存在路侧金属防护、中央绿化带隔离和道路表面干燥这几类因素时，可有效降低 TAIDME，当路侧金属防护与道路表面干燥发生耦合作用时，TAIDME 最低。

4 结论

(1) 以云南省 1760 起山区高速公路交通事故数据为基础，构建了 RFNB-CDM，对 TAIDME 的影响因素进行分析，模型采用 RFNB 模型对 TAIDME 进行单因素分析，并提取影响程度较大的因素，采用耦合度模型分析各因素的耦合特性，量化各因素耦合作用时对 TAIDME 的影响程度。该方法可为预防山区高速公路事故、减轻事故伤害程度提供一定决策参考。

(2) RFNB 分析结果表明：18:00-次日 6:00、事故车辆数 2 辆、追尾碰撞等 11 类因素对 TAIDME 的影响较大；耦合分析结果表明：追尾碰撞、18:00-次日 6:00、事故车辆数 2 辆、下

坡段、夜间无路灯照明、货运、大中型货车和直行匀速等 8 类因素存在时, 事故造成的伤害程度更高; 路侧金属防护、中央绿化带隔离和道路表面干燥等 3 类因素存在时, 事故造成的伤害程度更低。

参考文献:

- [1] Bucsuházy K, Matuchová E, Zůvala R, et al. Human factors contributing to the road traffic accident occurrence[J]. Transportation Research Procedia, 2020, 45: 555-561.
- [2] Jonghak L, Junghyo C, Taekwan Y, et al. Traffic accident severity analysis with rain-related factors using structural equation modeling – A case study of Seoul City[J]. Accident Analysis and Prevention, 2018, 112: 1-10.
- [3] 戢晓峰, 吴亚欣, 袁华智, 等. 二级公路交通流特征对事故严重程度的影响[J]. 中国公路学报, 2020, 33(2): 135-145. [JI X F, WU Y X, YUAN H Z, et al. Influences of Traffic Flow Characteristics on Accident Severity on Secondary Roads[J]. China Journal of Highway and Transport, 2020, 33(2): 135-145.]
- [4] 孙轶轩, 邵春福, 赵丹, 等. 交通事故严重程度 C5.0 决策树预测模型[J]. 长安大学学报(自然科学版), 2014, 34(5): 109-116. [SUN Y X, SHAO C F, ZHAO D, et al. Traffic accident severity prediction model based on C5.0 decision tree[J]. Journal of Chang'an University(Natural Science Edition), 2014, 34(5): 109-116.]
- [5] 赵玮, 徐良杰, 冉斌, 等. 基于 Adaptive Lasso 及 RF 算法的冰雪天气交通事故分析[J]. 中国安全科学学报, 2017, 27(2): 98-103. [ZHAO W, XU L J, RAN B, et al. Analyzing traffic crash under iced and snow weather condition based on Adaptive Lasso and RF[J]. China Safety Science Journal, 2017, 27(2): 98-103.]
- [6] 房锐, 张琪, 胡澄宇, 等. 基于风险矩阵的干线公路弯道路段交通冲突风险评估模型[J]. 交通运输系统工程与信息, 2021, 21(2): 166-172. [FANG R, ZHANG Q, HU C Y, et al. Risk Assessment Model Based on Risk Matrix for Traffic Conflict on Arterial Highway Bend Section[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(2): 166-172.]
- [7] 李英帅, 张旭, 王卫杰, 等. 基于随机森林的电动自行车骑行者事故伤害程度影响因素分析[J]. 交通运输系统工程与信息, 2021, 21(1): 196-200. [LI Y S, ZHANG X, WANG W J, et al. Factors Affecting Electric Bicycle Rider Injury in Accident Based on Random Forest Model[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(1): 196-200.]
- [8] 张轮, 杨文臣, 刘拓, 等. 基于朴素贝叶斯分类的高速公路交通事件检测[J]. 同济大学学报(自然科学版), 2014, 42(4): 558-563+595. [ZHANG L, YANG W C, LIU T, et al. A Naive Bayesian Classifier-based Algorithm for Freeway Traffic Incident Detection[J]. Journal of Tongji University(Natural science), 2014, 42(4): 558-563+595.]
- [9] Fang Z, Hongguo X, Huiyong Z, et al. Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models[J]. Mathematical Problems in Engineering, 2013, 43: 402-411.
- [10] 张军, 胡震波, 朱新山. 基于 AdaBoost 分类器的实时交通事故预测[J]. 计算机应用, 2017, 37(1): 284-288. [ZHANG J, HU Z B, ZHU X S. Real-time traffic accident prediction based on AdaBoost classifier[J]. Journal of Computer Applications, 2017, 37(1): 284-288.]
- [11] Athanasios T. Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials[J]. Journal of Safety Research, 2017, 61:9-21.
- [12] 张文钧, 蒋良孝, 张欢, 等. 一种双层贝叶斯模型: 随机森林朴素贝叶斯[J/OL]. 计算机研究

- 与发展: 1-12[2021-09-01]. <http://kns.cnki.net/kcms/detail/11.1777.TP.20210201.1743.004.html>.
[ZHANG W J, JIANG L X, ZHANG H, et al. A Two-Layer Bayes Model: Random Forest Naive Bayes[J/OL]. Journal of Computer Research and Development: 1-12[2021-09-01]. <http://kns.cnki.net/kcms/detail/11.1777.TP.20210201.1743.004.html>.]
- [13] 中华人民共和国公安部. GA/T 1082-2013, 道路交通事故信息调查[S]. 中国标准出版社, 2013.
- [14] Seyed H A H, Seyed M H H. Traffic accident severity prediction using a novel multi-objective genetic algorithm[J]. Taylor & Francis, 2017, 22: 425-440.
- [15] 宋述芳, 何入洋. 基于随机森林的重要性测度指标体系[J]. 国防科技大学学报, 2021, 43(2): 25-32. [SONG S F, HE R Y. Importance measure index system based on random forest[J]. Journal of National University of Defense Technology, 2021, 43(2): 25-32.]