

不同天气下快速路交织区交通冲突机理

周 备¹, 周千喜¹, 马壮林¹, 蔡凌霄², 路 熙³

(1. 长安大学 运输工程学院, 西安 710064; 2. 贵州省发展和改革委员会 贵州省民营经济和区域经济发展中心, 贵阳 550081; 3. 交通运输部 科学研究院, 北京 100029)

摘 要: 针对不同天气条件下城市快速路交织区交通冲突机理进行研究, 提出了一种基于车辆边框信息的交通冲突识别算法, 采用超阈值极值理论确定不同天气条件下的严重冲突阈值, 并结合机器学习模型和可解释性分析方法, 系统探讨了天气对交通冲突特征的影响。通过高精度轨迹数据, 解决了传统冲突指标提取方法精度不足、阈值选取主观性强以及天气因素影响研究缺失的问题。研究结果表明, 晴天与雨天的严重冲突阈值分别为 1.75 s 和 1.85 s, 雨天严重冲突比例 (39.1%) 显著高于晴天 (9.7%)。通过可解释性分析发现: ①天气条件显著调节交通冲突影响因素; ②晴天时, 车辆空间位置 (尤其是被冲突车辆是否位于匝道或交织车道) 是影响冲突严重程度的主导因素; ③雨天时, 影响因素的作用强度显著增大, 且交通流特征成为主导因素。本文研究成果可为交织区的差异化管理和安防防控提供理论依据。

关键词: 交通运输规划与管理; 交通冲突; 天气影响; 模型可解释性; 碰撞时间

中图分类号: U491.3 **文献标志码:** A **文章编号:** 1671-5497(2026)09-2329-11

DOI: 10.13229/j.cnki.jdxbgxb.20250105

Mechanisms of traffic conflicts in expressway weaving areas under different weather condition

ZHOU Bei¹, ZHOU Qian-xi¹, MA Zhuang-lin¹, CAI Ling-xiao², LU Xi³

(1. School of Transportation Engineering, Chang'an University, Xi'an 710064, China; 2. Guizhou Provincial Private Economy and Regional Economy Development Center, Guizhou Provincial Development and Reform Commission, Guiyang 550081, China; 3. China Academy of Transportation Science, Ministry of Transport, Beijing 100029, China)

Abstract: Aiming at the mechanism of traffic conflicts in urban expressway weaving areas under different weather conditions, a traffic conflict identification algorithm based on vehicle bounding box information was proposed. The peak over threshold extreme value theory was employed to determine severe conflict thresholds under distinct weather conditions, and machine learning models combined with interpretability analysis methods were utilized to systematically investigate the impact of weather on traffic conflict characteristics. By leveraging high-precision trajectory data, the limitations of traditional conflict metric extraction methods, such as insufficient accuracy, subjectivity in threshold selection, and the lack of

收稿日期: 2025-02-11.

基金项目: 国家自然科学基金青年基金项目 (52102404); 陕西省自然科学基金基础研究计划项目 (2024JC-YBMS-359); 中央高校基本科研业务费专项资金项目 (300102343204).

作者简介: 周备 (1986-), 男, 副教授, 博士. 研究方向: 交通安全. E-mail: bzhou3@chd.edu.cn

通信作者: 马壮林 (1980-), 男, 教授, 博士. 研究方向: 城市交通. E-mail: zhuanglinma@chd.edu.cn

exploration into weather-related influences were addressed. The results demonstrate that the severe conflict thresholds for sunny and rainy conditions are 1.75 s and 1.85 s, respectively, with the proportion of severe conflicts in rainy conditions (39.1%) being significantly higher than in sunny conditions (9.7%). Through interpretability analysis, the following findings are revealed: ① Weather conditions significantly modulate the influencing factors of traffic conflicts; ② Under sunny conditions, spatial vehicle positions (particularly whether the conflicted vehicle is located in a ramp or weaving lane) dominate the severity of conflicts; ③ In rainy conditions, the intensity of influencing factors increases markedly, and traffic flow characteristics emerge as the predominant factors. These findings provide a theoretical foundation for differentiated management and safety prevention and control in weaving areas.

Key words: transportation planning and management; traffic conflict; weather impact; model interpretability; time to collision

0 引言

快速路交织区作为连接快速路与其他道路的关键节点,因其车道变化和行驶方向差异所导致的频繁变道、加减速等行为,将显著增加交通安全风险,因此亟需深入研究其安全机理。

针对交织区安全问题,现有研究主要采用两类数据:传统事故数据和替代性安全评价指标,如碰撞时间(Time to collision, TTC)。在传统事故数据分析领域,曹倩等^[1]提出了考虑风险异质性的交通事故风险评估方法,通过融合道路类型和事故密度,提高了传统评估的准确性。Wang等^[2]则结合事故数据、道路几何特征和天气因素,构建了多层次贝叶斯逻辑回归预测模型。

相较于事故数据的时空局限性,交通冲突数据凭借其采集便捷性、样本丰富性等优势,逐渐成为研究焦点,其与事故的关联性已经获得部分验证^[3]。相关研究主要依赖实测车辆轨迹数据,仅少数采用仿真方法^[4]。例如,Yuan等^[5]、李佳硕等^[6]、马菲^[7]、卢启慧^[8]以及 Onelcin等^[9]的研究中,均使用无人机拍摄或运动相机拍摄的方式获取原始轨迹数据。在从原始轨迹数据中提取交通冲突指标及交通流参数时主要依托3类技术:计算机视觉算法^[10]、运动识别软件^[8]或算法结合人工识别的方式^[9]进行。而阈值确定方法包含:①经验法,如Xia等^[10]直接采用 $TTC=4$ s 阈值;②累计频率法,马菲^[7]以15%分位数划分冲突等级,卢启慧^[8]选用85%分位数,赵涛等^[11]验证了该方法的适用性;③极值理论法,穆松成^[3]构建了侵占后时间(Post encroachment time, PET)和改进碰撞时间(Modified time to collision, MTTC)超阈值模型,

通过平均剩余寿命图优化阈值选取。

国内外研究虽从交通流和冲突视角揭示了交织区安全机理,仍存在3方面局限:①传统冲突指标提取方法依赖人工干预,存在覆盖不全、精度不足的缺陷;②累计频率法阈值受样本量制约,经验阈值法忽视动态交通状态影响;③既有成果多聚焦晴天场景,缺乏雨雪等复杂天气下的系统性对比分析。针对上述问题,本文以城市快速路交织区轨迹数据为研究对象,提出了基于车辆边框信息的交通冲突识别算法。采用超阈值极值理论对不同天气条件下的交通冲突数据进行阈值选取;通过机器学习方法深入分析了影响冲突严重程度的核心特征及其在不同天气条件下的差异性,以期快速路交织区的差异化管理提供理论支撑。

1 冲突指标提取

本文使用的数据集来源于CitySim^[12],该数据集主要由中佛罗里达大学智能安全交通实验室(UCF-SST)主导维护,同时也与西南交通大学、东南大学等展开合作,提供中国部分地区采集的交通轨迹数据。本文使用的数据为该数据集中的亚洲快速路A交织段数据,具体位于四川省成都市武侯区天府大道和南三环交叉处的天府立交,数据采集时间为2022年5月中旬。调查路段总长度为260 m,其中,交织段长度为160 m,道路限速为80 km/h。调查区域俯瞰图详见图1。数据中,无人机累计数据记录时长为313 min,其中晴天数据记录时长为213 min,雨天数据记录时长为100 min。由于无人机每次飞行时间有限,因此数据分多次采集,采集时间段主要覆盖早上8:00~9:30,以及晚上17:30~18:30。为确保轨

迹数据的精度, CitySim 使用了五步处理流程, 包括视频稳定化、目标过滤、多视频拼接、目标检测和跟踪、增强的错误过滤。其轨迹识别的平均精度达到 97.6%^[12], 可以满足后续数据分析及建模需求。



图1 调查区域俯瞰示意图

Fig. 1 Overview diagram of the study area

为衡量交通冲突, 学界已提出多种替代性安全评价指标, 例如 TTC、MTTC、避免碰撞减速度 (Deceleration rate to avoid a crash, DRAC) 等。本文研究对象为城市快速路交织区, 主要冲突类型为交织车流之间产生的小角度换道冲突和车辆换道引发其他车辆制动避险行为导致的追尾冲突。此外, 本文旨在利用高精度轨迹数据精确识别潜在碰撞的即时风险, 因此选择 TTC 作为核心度量指标, 并着重于改进其计算精度以适应交织区的冲突特点。该指标在时间层面上, 反映了当前时刻下后车速度大于前车时, 若双方维持原有的速度和行驶轨迹不变 (即假定驾驶员不采取任何避险措施), 从冲突发生到预计碰撞所需要的时间。以追尾冲突为例 (见图 2), 假设 t_1 时刻后车速度 v_1 大于前车速度 v_2 , 两车处于相同车道且行驶方向相同, 其车辆间空间距离为 l_{d1} , 若两车保持 t_1 时刻的行驶状态, 则可以得到经过 Δt 的时间长度, 在预计的 t_2 时刻两车会发生追尾碰撞。

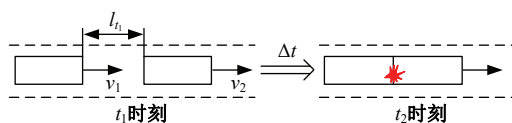


图2 追尾冲突 TTC 计算示意图

Fig. 2 Schematic diagram of rear-end conflict TTC calculation

此处 Δt 的时间长度就是 t_1 时刻两车的 TTC, 其计算公式为:

$$S_{TTC} = \Delta t = \frac{l_{d1}}{v_1 - v_2}, \quad v_1 > v_2 \quad (1)$$

式中: S_{TTC} 为距离碰撞时间, s; l_{d1} 为 t_1 时刻后车车头与前车车尾之间的空间距离, m; v_1 为 t_1 时刻后车的瞬时速度, m/s; v_2 为 t_1 时刻前车的瞬时速度, m/s。

传统 TTC 计算方法主要基于前、后车辆的相对速度和直线距离进行估算。然而, 此类方法存在两个主要局限性: ①假设车辆在潜在冲突过程中保持恒定速度和方向, 忽略了驾驶员行为和道路条件导致的速度和方向动态变化; ②难以有效识别小角度换道等复杂交织区冲突情况, 影响了风险评估的准确性。虽然 DRAC 和 MTTC 等指标能反映规避难度或提供更精细的轨迹预测, 但加速度数据在实际采集中标定和噪声处理较为复杂。相比之下, TTC 直接关联碰撞的时间临近性, 且可以通过精确的几何定位和相对速度信息进行计算。因此, 本文重点是克服传统 TTC 在几何表征上的局限性, 利用高精度轨迹数据中的车辆边框信息, 开发一种能够更准确计算复杂交互场景下 TTC 的方法, 从而提升风险识别的精度。

本文采用的高精度轨迹数据不仅包含时间戳、车辆标识符、车辆的中心位置坐标、行驶方向和速度等基础信息, 还包括车辆在平面投影上的矩形边框 4 个顶点坐标。基于这些丰富的数据特征, 使用 Python 编程语言开发了一种结合车辆几何特征和运动学参数的 TTC 提取算法。该算法的输入为原始轨迹数据, 输出为 TTC 值, 其核心思路是通过精确计算潜在碰撞点位置来获取更准确的 TTC 值。算法的关键步骤如下所示。

(1) 时间维度数据分组: 将原始轨迹数据按时间戳划分, 通过时间戳对齐多车状态, 解决异步数据流问题, 确保车辆对的运动参数在相同时间片段内可比。

(2) 运动参数提取: 对每个时间戳内的车辆对 (车辆 A 和 B), 基于原始轨迹数据, 提取其边框几何信息和运动参数, 计算相对速度向量。

(3) 碰撞点检测: 采用双向参考系转换方法进行点-边碰撞检测。首先, 以车辆 B 为参考系, 计算车辆 A 某个顶点的运动轨迹与车辆 B 某个边界线的交点, 若交点位于该边界线内, 则记录碰撞点坐标; 随后转换参考系, 以车辆 A 为基准, 逆向计算车辆 B 某个顶点的运动轨迹与车辆 A 边界线的交点。

(4) TTC 计算: 根据式 (1), 基于碰撞点坐标和相对运动参数计算相应的 TTC 值, 得到车辆间的相对交通冲突。

(5) 迭代计算: 如图 3 所示, 分别对车辆 A 和车辆 B 的所有顶点重复步骤 (3) 和 (4), 对所有可能的碰撞点进行计算, 每对车辆需计算 32 种顶

点-边组合(4顶点 $\times$ 4边 $\times$ 2参考系方向),获取所有可能的候选TTC集合。

(6)最小值选取:选择最小TTC值作为该车车辆对的冲突风险指标。对于跨越多个时间戳的同一车辆对,对每个时间戳重复上述计算,选择最小的TTC值作为最终评估结果。

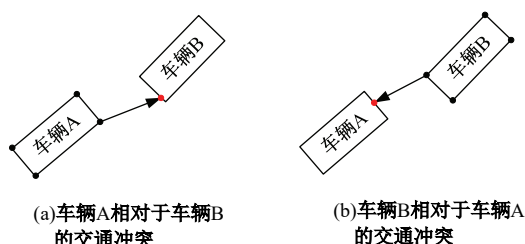


图3 碰撞点检测示意图

Fig. 3 Schematic diagram of collision point detection

本文算法具有以下优势:①考虑了车辆实际几何形状,避免了点源简化导致的误差;②能够有效识别追尾、换道和侧向碰撞等多种类型的交通冲突;③通过考虑车辆运动的连续性,确保了分析结果的时空一致性。基于车辆轨迹数据,经过TTC提取算法处理后分别得到不同天气条件下(晴天和雨天)的TTC值,考虑到较小的TTC值更能反映实际的交通冲突风险,仅保留 $TTC \leq 6$ s的数据用于后续分析。其中,晴天条件下冲突次数为8 755次,雨天条件下冲突次数为8 534次。

为检验晴天和雨天条件下TTC值分布的统计学差异,进行了双样本Kolmogorov-Smirnov(KS)检验。该检验是一种非参数检验方法,用于比较两个独立样本是否来自相同的分布。经过KS检验,晴天与雨天的TTC分布存在显著差异($D=0.1034, p<0.001$,其中, D 为两组累积分布函数的最大偏差),表明天气条件对TTC分布具有显著的统计学影响。

图4为不同天气条件下TTC值的核密度估

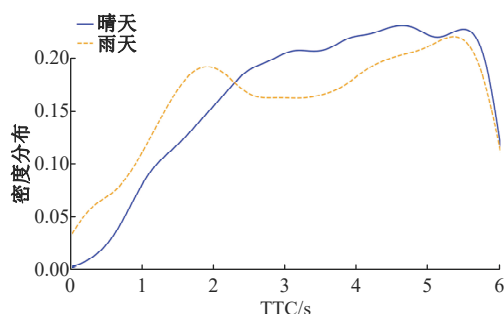


图4 晴天及雨天TTC值密度分布图

Fig. 4 Density distribution of TTC values in sunny and rainy conditions

计图,旨在揭示TTC值在晴天和雨天的分布差异。横轴表示TTC值,较小的TTC值意味着更高的潜在碰撞风险;纵轴表示特定TTC值的样本密度,反映了数据的集中程度。由图4可以看出,晴天时,TTC值较大的区域密度较高,而雨天时,TTC值较小的区域密度更大。这一现象表明,不同天气条件下车辆的冲突风险存在显著差异,晴天的交通环境相对安全,而雨天的驾驶风险更高。这种分布差异为理解交通冲突的风险特征提供了重要依据,并为后续设定严重冲突和一般冲突的阈值提供了数据支持。

2 超阈值极值理论阈值提取

2.1 超阈值极值理论

极值理论(Extreme value theory, EVT)是统计学和概率论的一个分支,其核心在于探索样本数据中极端数值(极大值或极小值)所遵循的概率统计规律。相比于经典统计学,极值理论主要聚焦于随机样本中异常值的分布特性,揭示一般事件分布背后极端事件的统计规律。因此,极值理论提供了更为精确和有针对性的极端事件分析工具。

超阈值极值模型通过选取特定阈值后,对所有超过设定阈值的数据进行极值建模。超阈值模型由于仅考虑数据尾部的概率分布情况,因此通常使用广义帕累托分布(Generalized pareto distribution, GPD)对阈值极值数据进行分布拟合。假设 $X_1, X_2, \dots, X_n$ 是来自总体分布函数 F 的一系列独立样本,定义阈值为 u ,若 $X_i > u$,则称 X_i 和 $X_i - u$ 分别为超阈值和超出量; $F_u(x)$ 为随机变量 X_i 超过阈值 u 的超出量的分布函数,如下所示:

$$F_u(x) = \Pr\{X - u > x, X > u\} = \frac{1 - F(u+x)}{1 - F(u)}, x > 0 \quad (2)$$

极值统计领域研究表明^[13],当阈值选取合适时,分布函数 $F_u(x)$ 可近似采用广义帕累托分布,如下所示:

$$F_u(x) \approx G(x; u, \sigma, \xi) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{x}{\sigma}\right), & \xi = 0 \end{cases} \quad (3)$$

式中: σ 为尺度参数,且 $\sigma > 0$; ξ 为形状参数,且 $-\infty < \xi < +\infty$ 。

2.2 阈值选取

超阈值模型的建立需要考虑总体样本的分布情况,不同阈值选取直接关系到极值样本的容量范围,并显著影响针对极值样本进行广义帕累托分布参数估计后结果的准确性。

超阈值模型的阈值选取方法可大致归结为数值法和图解法两大类。数值法主要包括均方差法、阈值自动选取法和多重阈值选取法等;而图解法则主要包含Hill图、平均残差寿命图和阈值稳定性分析图等。本文采用图解法中的平均剩余寿命图和阈值稳定性分析图来提取合适的阈值。图解法通过图形展示数据的内在规律和特征变化,以直观展示数据特征,确定合适的阈值,从而确保模型参数估计的准确性。阈值选取方法的原理如下所示。

(1) 平均残差寿命图

对于一组给定的样本 $X_1, X_2, \dots, X_n$, 针对初始阈值 u_0 , 假设超出该阈值的样本服从广义帕累托分布, 定义平均残差寿命 $e(u_0)$ 为:

$$e(u_0) = E(X - u_0 | X > u_0) \quad (4)$$

根据广义帕累托分布的均值计算方法可以得到:

$$e(u_0) = \frac{\sigma_{u_0}}{1 - \xi} \quad (5)$$

式中: σ_{u_0} 为超出阈值 u_0 的样本所服从广义帕累托分布的尺度参数。

根据广义帕累托分布的基本性质, 若超出初始阈值 u_0 的所有样本服从广义帕累托分布, 则阈值逐步增大时, 对于任意一个 $u > u_0$, 超出阈值 u 的所有样本同样服从广义帕累托分布, 由此计算得出:

$$e(u) = \frac{\sigma_u}{1 - \xi} = \frac{\sigma_{u_0} + \xi(u - u_0)}{1 - \xi} = \frac{\sigma_{u_0}}{1 - \xi} - \frac{\xi}{1 - \xi} u_0 + \frac{\xi}{1 - \xi} u \quad (6)$$

由于在合理阈值区间内, 其形状参数 ξ 保持相对稳定, 因此, 当 $u > u_0$ 时, $e(u)$ 是一个关于阈值 u 的线性函数。将 $(u, e(u))$ 作为坐标值所绘制的图即为平均残差寿命图。若阈值 u 选取得当, 平均残差寿命图应呈现近似线性趋势, 由此可判断合适阈值的选取范围。

(2) 阈值稳定性分析图

仅依靠平均残差寿命图进行阈值选取具有一

定的主观性, 同时也可能存在阈值选择范围较大的问题, 因此引入阈值稳定性分析图进一步精确阈值选择的范围以保证选取的客观性。

对于任意阈值 u , 可以对大于阈值范围的极值样本进行广义帕累托分布的参数估计, 从而得出不同阈值取值下的形状参数和尺度参数。若极值样本服从广义帕累托分布, 则其形状参数 ξ 应较为稳定。同时, 尺度参数 σ_u 为 u 的线性函数值, 即:

$$\sigma_u = \sigma_{u_0} + \xi(u - u_0) \quad (7)$$

若 $\xi \neq 0$, σ_u 会随着阈值 u 的变化而改变。因此, 引入修正后的尺度参数:

$$\sigma_u^* = \sigma_u - \xi u = \sigma_{u_0} + \xi(u - u_0) - \xi u = \sigma_{u_0} - \xi u_0 \quad (8)$$

式中: σ_u^* 为修正后的尺度参数。

该参数修正了阈值变化带来的线性趋势, 使得修正后的尺度参数在合理阈值区间内保持大致恒定。

因此, 对于合适的阈值 u , 参数估计得到的形状参数 ξ 和修正后的尺度参数 σ_u^* 应是基本不变的。基于上述原理, 可以通过绘制不同阈值下广义帕累托分布的形状参数和修正后的尺度参数变化图来判断是否存在一个合理的阈值区间。

晴天条件下, TTC 阈值稳定性分析图及平均残差寿命图如图 5 和图 6 所示。雨天条件下, TTC 阈值稳定性分析图及平均残差寿命图如图 7 和图 8 所示。

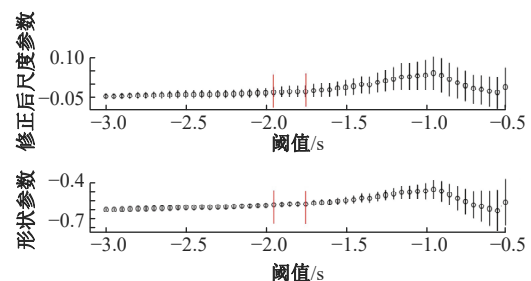


图5 晴天TTC阈值稳定性分析图

Fig. 5 Stability analysis of TTC threshold in sunny conditions

根据晴天下的TTC平均残差寿命图, 在阈值区间 $[-6.0, -1.0]$ 内, 图形均呈现近似线性趋势, 难以判断合理阈值; 根据修正后尺度参数和形状参数稳定性分析图得到的阈值区间均为 $[-1.95, -1.75]$, 取交集的最大值 -1.75 作为晴天部分TTC的划分阈值。

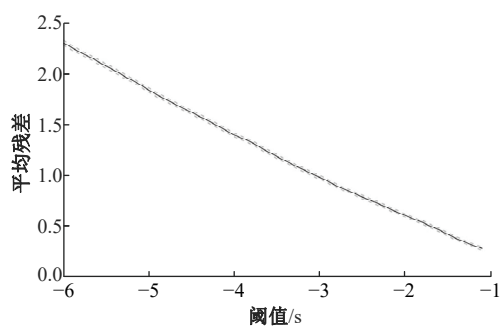


图 6 晴天 TTC 平均残差寿命图

Fig. 6 Mean residual life plot of TTC in sunny conditions

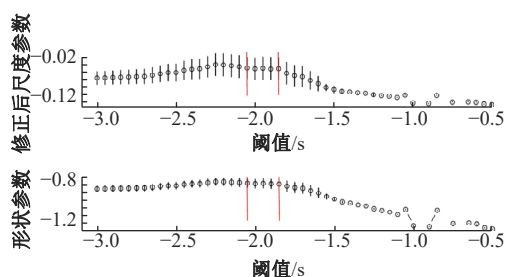


图 7 雨天 TTC 阈值稳定性分析图

Fig. 7 Stability analysis of TTC threshold in rainy conditions

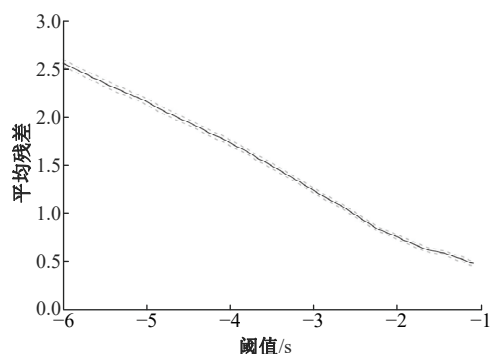


图 8 雨天 TTC 平均残差寿命图

Fig. 8 Mean residual life plot of TTC in rainy conditions

雨天下的 TTC 数据,根据平均残差寿命图,在阈值区间 $[-6.0, -1.0]$ 内,图形均呈现近似线性趋势;根据修正后尺度参数和形状参数稳定性分析图得到的阈值区间均为 $[-2.05, -1.85]$,取交集的最大值 -1.85 作为雨天部分 TTC 的划分阈值。

根据阈值确定结果,将晴天和雨天的 TTC 数据中每一个冲突事件分类为一般冲突或严重冲突。具体各组数据的冲突分类情况见表 1。从表 1 中可以看出,不同天气条件下的 TTC 阈值和严重冲突数量及比例存在显著差异,反映了天气条件对交织区交通冲突特性的显著影响。同时,

雨天严重冲突的数量(3 333 起)和比例(39.1%)远高于晴天(850 起,9.7%),这表明雨天条件下路面湿滑、能见度降低等因素显著增加了交通安全风险,导致冲突发生更频繁且更为严重。

表 1 冲突严重程度分类表

Table 1 Table of conflict severity classification

数据划分	TTC 阈值/s	严重冲突数量	比例/%
晴天	1.75	850	9.7
雨天	1.85	3 333	39.1

3 冲突严重程度影响因素分析

3.1 特征信息概述

将晴天和雨天部分的 TTC 使用不同阈值进行严重冲突/非严重冲突划分后,将提取到的交通流特征信息、车辆特征信息、路段特征信息等与每一起冲突事件一一对应,用于快速路交织区交通冲突严重程度影响因素分析。同时,为进一步细化影响因素分类,将交织区内的交织路段分为了 5 个区间,各区间长度信息如表 2 所示。其中, W_1 区间位于交织段最上游,按照路段行车方向,下一个区间为 W_2 区间,并以此类推直到最后一个区间。

表 2 二级区间分类信息表

Table 2 Classification information for secondary intervals

所处路段位置	区间编号	区间长度/m
交织段	W_1	30
	W_2	30
	W_3	40
	W_4	30
	W_5	30

最终,提取出的冲突严重程度影响因素共 46 个,主要包括路段车流特征信息、冲突发生区间信息、相关车辆信息以及环境因素这 4 类对冲突严重程度有潜在影响的变量,相关信息详见表 3。其中,车辆行驶方向角度一阶差分标准差反映了车辆行驶方向变化的波动性和不稳定性,若该指标较大,则说明车辆行驶角度波动较大,即车辆轨迹不稳定;若该指标较小,则说明车辆行驶角度波动较小,即车辆轨迹较为稳定。

3.2 LightGBM 模型建立及超参数优化

LightGBM (Light gradient boosting machine) 是一种基于决策树的梯度提升框架,主要用于分类和回归任务。该框架具有以下主要特点:①采

表 3 变量相关信息

Table 3 Variable information

变量	单位及描述	变量	单位及描述
前 3 min 路段车流速度均值	m/s	前 3 min 前方区间车流速度均值	m/s
前 3 min 路段车流速度标准差	—	前 3 min 前方区间车流速度标准差	—
前 3 min 路段车流密度均值	pcu/km	前 3 min 前方区间车流密度均值	—
前 3 min 路段车流密度标准差	—	前 3 min 前方区间车流密度标准差	—
前 3 min 路段车流量均值	pcu/h	前 3 min 前方区间车流量均值	pcu/h
前 3 min 路段车流量标准差	—	前 3 min 前方区间车流量标准差	—
前 3 min 自身区间车流速度均值	m/s	前 3 min 后方区间车流速度均值	m/s
前 3 min 自身区间车流速度标准差	—	前 3 min 后方区间车流速度标准差	—
前 3 min 自身区间车流密度均值	pcu/km	前 3 min 后方区间车流密度均值	pcu/km
前 3 min 自身区间车流密度标准差	—	前 3 min 后方区间车流密度标准差	—
前 3 min 自身区间车流量均值	pcu/h	前 3 min 后方区间车流量均值	pcu/h
前 3 min 自身区间车流量标准差	—	前 3 min 后方区间车流量标准差	—
瞬时路段车流量	pcu/h	冲突车辆行驶方向角度一阶差分标准差	—
瞬时路段车流密度	pcu/km	被冲突车辆行驶方向角度一阶差分标准差	—
瞬时路段车流速度	m/s	冲突车辆位于匝道和交织车道	0-否;1-是
主线至主线流量	pcu	冲突车辆位于主线外侧车道	0-否;1-是
主线至匝道流量	pcu	冲突车辆位于主线中间车道	0-否;1-是
匝道至主线流量	pcu	冲突车辆位于主线内侧车道	0-否;1-是
匝道至匝道流量	pcu	被冲突车辆位于匝道和交织车道	0-否;1-是
交织比	%	被冲突车辆位于主线外侧车道	0-否;1-是
交织段车流密度	pcu/km	被冲突车辆位于主线中间车道	0-否;1-是
交织段与上游段车流密度差	pcu/km	被冲突车辆位于主线内侧车道	0-否;1-是
交织段与下游段车流密度差	pcu/km	是否晚高峰	0-否;1-是

用基于直方图的决策树学习算法,显著提升了大规模数据处理效率;②实现了Leaf-wise的树生长策略,相比传统的Level-wise方法具有更高的准确性和计算效率;③支持直接处理类别型特征,无需进行独热编码转换;④提供灵活的超参数调节机制,便于针对具体应用场景进行优化。基于上述特点,LightGBM 特别适用于处理高维特征和大规模数据集的机器学习任务。

为充分对比晴天和雨天冲突严重程度影响因素的异同,针对晴天和雨天 TTC 数据分别建模。在模型训练过程中,首先,将晴天 TTC 数据和雨天 TTC 数据分别按照 7:3 的比例分为训练集和测试集。然后,在训练集上使用递归特征消除(Recursive feature elimination, RFE)进行特征筛选,以 LightGBM 为基模型,逐步剔除对曲线下面积(Area under the curve, AUC)值贡献最小的特征,最终保留关键特征子集。在筛选后的特征空间上,使用随机搜索算法进行 LightGBM 模型超参数调优。针对每个模型,进行 200 次随机搜索,每次搜索时均使用 5 折交叉验证。最后,使用测试集数据检验模型预测效果。获取的各模型最优超参数如表 4 所示。

表 4 超参数调优结果

Table 4 Hyperparameter optimization results

参数	含义	晴天 TTC	雨天 TTC
		模型取值	模型取值
subsample	每棵树训练时随机选择的样本比例,防止过拟合	0.6	1.0
num_leaves	单棵树的最大节点数,控制树的复杂度	100	31
n_estimators	模型迭代的次数	500	1 500
min_child_samples	一个节点上最小的样本数量	20	30
max_depth	树的最大深度,限制树的生长,防止过拟合。设置为-1时,深度不受限制	-1	5
learning_rate	学习率,控制每次迭代更新的步长	0.001	0.1
colsample_bytree	每棵树随机采样的特征比例	0.6	0.6

为评估模型的预测性能,选用准确率和 AUC 作为主要评估指标。准确率是指模型正确预测的样本数占总样本数的比例,用于衡量模型总体预测的准确性;AUC 值用于评估模型对不同分类阈

值的综合判别能力,其值越接近 1,表示模型在正负样本区分上的性能越优越。上述两个指标均为越大越好,能够全面反映模型在预测严重交通冲突时的性能表现。

使用最优超参数训练的晴天 TTC 预测模型的准确率为 0.799 4, AUC 值为 0.753 5;雨天 TTC 预测模型的准确率为 0.836 0, AUC 值为 0.880 6,说明两个模型预测效果均较好,可以用于下一步的模型预测结果解释。

3.3 SHAP 值影响因素分析

为解决机器学习模型可解释性问题,Lundberg 等^[14]提出了基于 Shapley 值的沙普利可加解释(SHapley additive explanations, SHAP)分析方法,该方法可对回归和分类模型的预测结果进行可视化解释,并已在相关研究中得到广泛应用^[15]。SHAP 方法基于合作博弈论中的 Shapley 原理,将特征变量视为博弈参与者,量化

各特征对模型预测结果的贡献度,从而实现单个预测结果的直观解释。

SHAP 方法的核心思想是通过构建可解释的替代模型阐释复杂模型的预测机制。该方法计算特征在所有可能组合中的边际贡献,既可从全局角度评估特征对整体预测的平均影响,也可从局部角度分析特征对单个样本预测的具体贡献。图 9 和图 10 分别为晴天和雨天 LightGBM 模型的 SHAP 值分析结果(彩图参见电子版,后文同)。图中,各变量按照重要性从高到低排列,重要性排序依据是各变量 SHAP 值的平均绝对值。横轴表示 SHAP 值大小,正值表示该变量的特征值提高了严重冲突的概率,负值表示减小了严重冲突的概率;纵轴表示不同的特征变量。图中每个点代表一个样本,点的颜色从蓝到红表示特征值从低到高的变化,色条宽度则反映了样本的分布密度,有效揭示了特征值与其预测结果之间的关系。

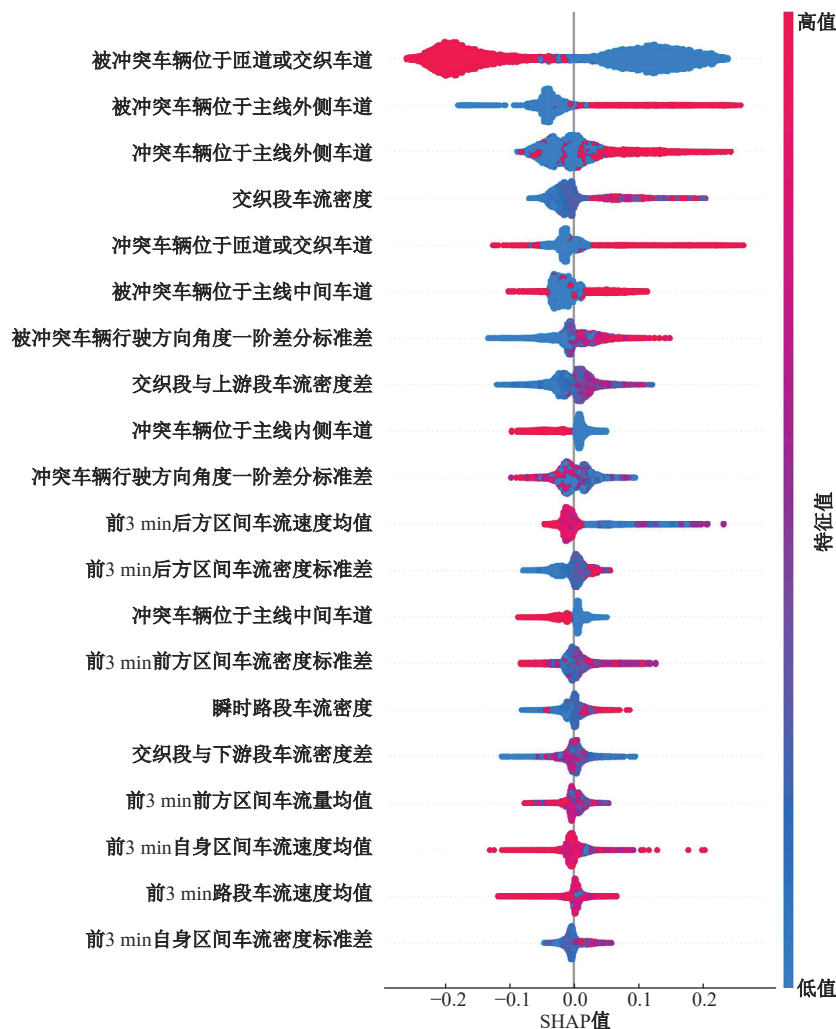


图 9 晴天 LightGBM 模型 SHAP 值分析图

Fig. 9 SHAP value analysis of the LightGBM model in sunny conditions

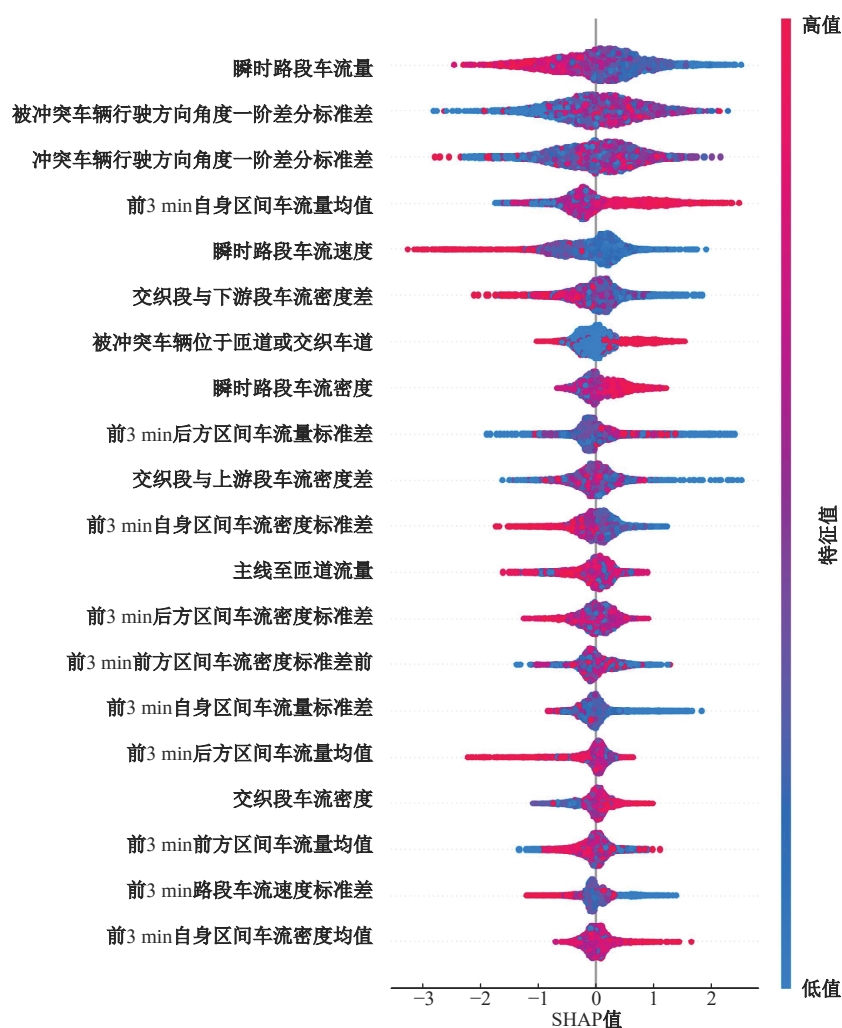


图10 雨天LightGBM模型SHAP值分析图

Fig. 10 SHAP value analysis of the LightGBM model in rainy conditions

由于建模涉及的影响因素较多,为了聚焦于对冲突严重程度影响最为显著的因素,并清晰呈现这些关键因素在模型中的重要性,图9和图10中仅展示了每个模型中最重要前20个影响因素。对比图9和图10的SHAP值分析结果,有以下主要发现:

(1)快速路交织区交通冲突的影响因素在晴天和雨天条件下呈现显著差异。SHAP值分析表明,在前20个最重要的影响因素中,有11个因素同时出现在两种天气条件下,表明这些因素具有稳定的影响作用。在这11个共同因素中,有4个因素表现出一致的作用效果,具体表现为:交织段车流密度较大时(红色样本),其对应的SHAP值多集中在正值区域,表明交织段车流密度与严重冲突风险呈正相关,这反映了该变量是影响交织区安全性的基础性因素,但其在不同天气条件下的重要性有一定差异,晴天时,该变量排名第4,

而在雨天条件下,该变量排名第17,表明晴天时该变量影响更为突出;瞬时路段车流密度较大时(红色样本),其对应的SHAP值多集中在正值区域,表示瞬时路段车流密度与严重冲突风险呈正相关,说明较高的瞬时路段车流密度会增加车辆间相互作用的概率,并提升交通冲突的风险;被冲突车辆行驶方向角度一阶差分标准差较大时(红色样本),其对应的SHAP值多集中在正值区域,由此可见,该变量与严重冲突风险呈正相关,表明车辆轨迹的不稳定性会显著增加严重冲突的风险;交织段与下游段车流密度差较大时(红色样本),其SHAP值主要分布在负值区,表明该变量与严重冲突风险呈负相关,这是由于较大的密度差有助于疏导交织区的交通流,使车辆能够更快地完成交织过程,从而降低持续暴露于高风险状态的时间。

(2)晴天条件下的特征影响模式。在晴天条

件下,车辆的空间位置特征展现出主导性影响。如图 9 所示,前 5 大影响因素中有 4 个与车辆空间位置直接相关,包括:被冲突车辆位于匝道或交织车道、被冲突车辆位于主线外侧车道、冲突车辆位于主线外侧车道,以及冲突车辆位于匝道或交织车道。具体分析表明:被冲突车辆位于匝道或交织车道时(红色样本),其对应的 SHAP 值多集中在负值区域,说明发生严重冲突的风险显著降低,这可能是因为驾驶员在这些区域会保持更高的警觉性,同时车辆在专用车道中的运行更为规范,相互干扰较少;当冲突车辆或被冲突车辆位于主线外侧车道时(红色样本),其对应的 SHAP 值多集中在正值区域,说明严重冲突风险显著提高,这主要归因于外侧车道是交织活动最为频繁的区域,车辆变道需求大,行驶轨迹复杂;冲突车辆位于匝道或交织车道时表现出的影响具有双向性(红色样本同时分布在正 SHAP 值和负 SHAP 值区域),这种异质性效应与具体的交织区构造、交通流组成以及驾驶员行为特征有关。

(3)雨天条件下的特征影响模式。雨天条件下的影响因素表现出两个显著特点:①各影响因素的 SHAP 值绝对量级显著大于晴天条件,这表明雨天环境极大加剧了各因素对交通冲突的影响程度;②交通流特征在影响因素中占据主导地位,这一点与晴天条件形成鲜明对比。具体分析发现:在 20 个主要影响因素中,仅有 1 个与车辆位置相关(被冲突车辆位于匝道或交织车道),其余 19 个均为交通流特征相关因素;瞬时路段车流量作为最重要的影响因素,与严重冲突风险呈负相关(红色样本多集中在负 SHAP 值区域),这是由于雨天条件下较小的车流量通常伴随着更大的速度差异和更不规则的车辆行为,而较大的车流量反而会促使驾驶员采取更谨慎的驾驶策略;车辆行驶方向角度一阶差分标准差与严重冲突风险呈正相关(红色样本多集中在正 SHAP 值区域),这反映了雨天条件下车辆轨迹稳定性对安全的重要影响;冲突前 3 min 自身区间车流量均值与严重冲突风险呈正相关(红色样本多集中在正 SHAP 值区域),这说明较高的持续交通需求会增加雨天条件下的安全风险,这是由于持续的高流量会增加驾驶员的心理负担,同时也提高了车辆间相互干扰的可能性;瞬时路段车流速度与严重冲突风险呈负相关(红色样本多集中在负 SHAP 值区

域),这是因为较高的车流速度通常发生在交通运行较为顺畅的时段,此时车辆间的相互作用相对较少,且驾驶员有更充足的时间和空间应对潜在的冲突风险。

4 结束语

本文基于 CitySim 轨迹数据,对城市快速路交织区在不同天气条件下的交通冲突特征进行了系统分析。使用 Python 编程语言开发了基于车辆边框信息的 TTC 提取算法,采用超阈值极值模型确定了不同天气条件下的严重冲突阈值,并结合 LightGBM 模型和 SHAP 值分析方法,深入探讨了晴天和雨天条件下影响交通冲突严重程度的关键因素。主要研究结论如下:①天气条件对交通冲突影响因素具有显著的调节效应。通过 SHAP 值分析发现,在前 20 个最重要的影响因素中,11 个因素在两种天气条件下均表现出稳定的影响作用,而其余 9 个因素则呈现出明显的天气相关性。这种差异性表明,快速路交织区的安全管理策略需要充分考虑天气因素的影响。②晴天条件下,车辆的空间位置特征是决定冲突严重程度的主导因素。其中,被冲突车辆位于匝道或交织车道、被冲突车辆位于主线外侧车道、冲突车辆位于主线外侧车道,以及冲突车辆位于匝道或交织车道是对冲突严重程度影响最重要的几个因素。③雨天条件下的影响机制呈现出两个显著特征:各影响因素的 SHAP 值绝对量级普遍大于晴天条件,表明雨天环境强化了各因素对交通冲突的影响程度。此外,交通流特征取代空间位置特征成为主导性影响因素,这种转变反映了恶劣天气条件下交通系统运行特性的变化。基于上述结论,为进一步改善城市快速路交织区交通安全,应采取差异化的管理策略,根据不同天气条件采取不同的交通管理措施,同时,特别需要关注雨天条件下的流量控制和速度管理。

参考文献:

- [1] 曹倩,李志慧,陶鹏飞,等.考虑风险异质特性的路网交通事故风险评估方法[J].吉林大学学报:工学版,2023,53(10):2817-2825.
Cao Qian, Li Zhi-hui, Tao Peng-fei, et al. Traffic accident risk assessment method for road network consider risk heterogeneity[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(10):

- 2817-2825.
- [2] Wang L, Abdel-Aty M, Shi Q, et al. Real-time crash prediction for expressway weaving segments[J]. Transportation Research Part C: Emerging Technologies, 2015, 61: 1-10.
- [3] 穆松成. 基于交通冲突极值统计的事故预测模型研究[D]. 威海: 哈尔滨工业大学汽车工程学院, 2022.
- Mu Song-cheng. Research on crash prediction model based on extreme statistics of traffic conflicts[D]. Weihai: School of Automotive Engineering, Harbin Institute of Technology, 2022.
- [4] 姜雪. 城市快速路交织区交通冲突机理研究[D]. 大连: 大连交通大学交通工程学院, 2021.
- Jiang Xue. Study on traffic conflict mechanism in urban expressway weaving areas[D]. Dalian: School of Traffic Engineering, Dalian Jiaotong University, 2021.
- [5] Yuan R, Abdel-Aty M, Xiang Q. A study on diversion behavior in weaving segments: individualized traffic conflict prediction and causal mechanism analysis[J]. Accident Analysis & Prevention, 2024, 205: No. 107681.
- [6] 李佳硕, 郑展骥, 顾欣, 等. 考虑快速路交织区驾驶人强行变道行为的交通冲突机理分析[J]. 交通信息与安全, 2023, 41(3): 1-11.
- Li Jia-shuo, Zheng Zhang-ji, Gu Xin, et al. An analysis of the mechanism of traffic conflicts considering risky lane-changing behavior in weaving sections of expressways[J]. Journal of Transport Information and Safety, 2023, 41(3): 1-11.
- [7] 马菲. 基于轨迹数据的快速路交织区交通风险评估[D]. 济南: 山东大学齐鲁交通学院, 2024.
- Ma Fei. Risk assessment of expressway weaving sections based on trajectory data[D]. Jinan: School of Qilu Transportation, Shandong University, 2024.
- [8] 卢启慧. 基于交通冲突预测模型的城市快速路交织区安全评价方法研究[D]. 南京: 东南大学交通学院, 2024.
- Lu Qi-hui. Research on safety evaluation method of urban expressway weaving areas based on traffic conflict prediction model[D]. Nanjing: School of Transportation, Southeast University, 2024.
- [9] Onelcin P, Alver Y. A new lane change index for lane change conflicts at weaving segments[J]. Traffic Injury Prevention, 2023, 24(7): 559-566.
- [10] Xia Y, Qin Y, Li X, et al. Risk identification and conflict prediction from videos based on TTC-ML of a multi-lane weaving area[J]. Sustainability, 2022, 14: No. 4620.
- [11] 赵涛, 张宁, 王小超, 等. 基于图神经网络轨迹预测的合流区交通冲突预测方法[J]. 山东大学学报: 工学版, 2024, 54(2): 36-46.
- Zhao Tao, Zhang Ning, Wang Xiao-chao, et al. A traffic conflict prediction method for merging areas based on trajectory prediction with graph neural network[J]. Journal of Shandong University (Engineering Science), 2024, 54(2): 36-46.
- [12] Zheng O, Abdel-Aty M, Yue L, et al. CitySim: a drone-based vehicle trajectory dataset for safety oriented research and digital twins[J]. Transportation Research Record, 2023, 2678(4): 606-621.
- [13] 史道济. 实用极值统计方法[M]. 天津: 天津科学技术出版社, 2006.
- [14] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, CA, USA, 2017: 4768-4777.
- [15] 温惠英, 何梓琦, 李秋灵, 等. 高速公路货车换道冲突预测及其影响因素分析[J]. 吉林大学学报: 工学版, 2024, 54(10): 2827-2836.
- Wen Hui-ying, He Zi-qi, Li Qiu-ling, et al. Traffic conflict prediction and influencing factors analysis of truck lane change on expressway[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(10): 2827-2836.