

电动自行车道路交通安全事故严重性影响因素分析*

肖向良

(湖南省交通科学研究院有限公司, 湖南 长沙 410015)

摘要: 电动自行车已成为中国居民重要的代步工具,但其带来的交通安全问题也非常突出。文中基于电动自行车事故历史数据,采用决策树方法构建电动自行车事故严重程度模型,分析伤害严重程度与骑行者、违法行为、道路、天气环境等的交互效应。结果表明,影响事故伤害严重程度最重要的因素是骑行者年龄与违法行为,年龄超过 55 岁和在没有非机动车道的道路上非机动车不靠车行道右侧行驶、非机动车逆向行驶、非机动车超速、横过机动车道时不下车推行等违法行为为事故更易导致严重伤害;容易导致高伤害严重程度的交通情境有 5 种,并提出相关改善建议。

关键词: 交通安全;电动自行车;交通事故;决策树;伤害严重性

中图分类号:U491.3

文献标志码:A

文章编号:1671-2668(2020)06-0032-05

电动自行车作为一种介于人力自行车和摩托车之间的交通工具,具有价格低廉、行驶速度较快、便捷、环保、可达性高等优点,深受中国居民喜爱。但电动自行车速度快、安全防护能力弱、稳定性差等特点导致其较易发生交通事故,且一旦发生事故,相比自行车,其事故伤害严重程度更大。目前关于摩托车和非机动车的安全研究多针对传统自行车和摩托车,而电动自行车在驾驶群体、运行速度、行驶距离等方面均不同于传统自行车及摩托车,其交通安全特征属性及相应的安全管理政策应具有自身特性。对交叉口处电动自行车与传统自行车闯红灯行为的调查结果表明,电动自行车用户闯红灯的概率显著高于传统自行车;对合肥 2009—2011 年 205 名传统自行车及电动自行车用户的住院记录的分析发现,电动自行车用户受重伤的比例达 1/3,而传统自行车用户仅 17%。而目前少有对电动自行车安全问题的研究,尤其缺少对电动自行车事故伤害严重程度影响因素的研究。

交通事故伤害严重性通常划分为离散的等级程度,常采用统计回归模型进行伤害严重程度建模。而统计回归模型都有其模型假设和因变量与自变量间既定的函数关系,同时自变量间的相互关联也困扰统计回归模型的应用。事故伤害严重程度是由人一车一路一环境多种因素共同导致,伤害致因错综复杂,与统计回归模型的假设不符。决策树方法可在不给出事故严重程度和影响因素间函数关系及不

考虑自变量间相互关联的情况下,清晰、高效地发掘二者间的内在关联,还能处理变量的交互作用,分析多种因素组合对事故伤害严重性的共同影响。该文基于 2014—2016 年湖南省电动自行车交通事故记录,基于决策树方法构建电动自行车事故伤害严重程度模型,分析事故伤害严重程度与骑行者、违法行为、道路及天气环境等之间的交互效应,识别易导致高伤害严重程度的交通情境,为降低电动自行车事故伤害严重程度措施制定提供依据。

1 数据准备

道路交通事故数据来自公安部“六合一”交通管理综合应用平台,交通事故信息涵盖事故特征、驾驶员特征、车辆特征、事故发生时间及环境特征。筛选 2014—2016 年湖南电动自行车与机动车碰撞事故数据(不含简易事故,因其信息记录项少)共 2 484 起,事故信息记录包括骑行者事故伤害严重程度、性别、年龄、职业、户口性质、事故发生时间、星期、季节、天气、照明、路口、道路类型及违法行为 13 类。

从不同角度对交通事故进行分析,结果(见表 1)显示:1) 事故伤害严重程度。骑行者受重伤与死亡事故的比重达 15.34%。2) 受伤人群特征。男性比例(57.69%)略高于女性(42.31%);老年人群(大于 55 岁)占比较大,达 35.1%;受伤人员户口性质多为非农业(86.15%)。3) 事故发生位置。事故多发生于一般城市道路,占比为 56.28%;发生在路段的

* 基金项目:湖南省交通运输厅科技与创新项目(201340;201606)

事故数量约为路口事故数量的2倍。4) 事故发生时间。发生在9:00—16:59时段事故占事故总量的47.42%,发生在17:00—18:59时段事故也较多(14.45%);周末事故数量占比小(24.28%),仅为工作日的1/3;夏季事故占比最高(28.86%),冬季事故占比最小(20.09%)。5) 人员违法行为。由电动自行车骑行者违法行为导致事故占比达55.84%,

高于由机动车驾驶员违法行为导致事故占比(44.16%)。其中非信号控制交叉口电动自行车不让行是排名第一的违法行为(占比9.70%),其次是在没有非机动车道的道路上电动自行车不靠车行道右侧行驶(8.05%),在有非机动车道的情况下电动自行车未在非机动车道内行驶也占较大比例(7.65%),电动自行车超速行为所占比例较小(1.33%)。

表1 建模变量描述性统计

变量	变量描述	数量	比例/%	变量	变量描述	数量	比例/%
伤害程度	死亡	251	10.11	道路类型	一般城市道路	1 398	56.28
	重伤	130	5.23		城市快速路	174	7.00
	轻伤	1 867	75.16		一级公路	78	3.14
	仅财产损失	236	9.50		二级公路	160	6.44
发生时间	22:00 — 6:59	331	13.33		三级公路	229	9.22
	7:00 — 8:59	322	12.96		四级公路	174	7.00
	9:00 — 16:59	1 178	47.42		等外公路	122	4.91
	17:00 — 18:59	359	14.45	其他	149	6.00	
星期	19:00 — 21:59	294	11.84	性别	男	1 433	57.69
	工作日	1 881	75.72		女	1 051	42.31
	周末	603	24.28	年龄	<25	176	7.09
季节	春	674	27.13		25~44	818	32.93
	冬	499	20.09		45~54	643	25.89
	秋	594	23.91		55~64	538	21.66
	夏	717	28.86		65~74	280	11.27
天气	晴	1 611	64.86	≥75	29	1.17	
	阴	495	19.93	职业	工人	160	6.44
	雨	378	15.22		快递外卖	161	6.48
照明条件	白天	1 801	72.50		农民	590	23.75
	黄昏、黎明	123	4.95		自主经营者	502	20.21
	夜间无路灯照明	118	4.75		职员	213	8.57
	夜间有路灯照明	442	17.79		其他	858	34.54
路口/路段	路段	1 651	66.47	户口性质	非农业	2 140	86.15
	路口	833	33.53		农业	344	13.85
违法行为	在没有非机动车道的道路上电动自行车不靠车行道右侧行驶		200	8.05			
	电动自行车未在非机动车道内行驶		190	7.65			
	驾驶电动自行车在路段上横过机动车道时不下车推行		161	6.48			
	电动自行车不按照交通信号规定通行		180	7.25			
	电动自行车逆向行驶		174	7.00			
	非信号控制交叉口电动自行车不让行		241	9.70			
	电动自行车超速		33	1.33			
	电动自行车骑行者其他违法行为		208	8.37			
	机动车驾驶员违法行为		1 097	44.16			

2 决策树分析

2.1 决策树方法

决策树方法是一种非线性无参数的预测模型,其目标是将样本集通过某个特定属性(自变量)的不同特征分为多个相对同质的子样本集。CART(分类和回归树)算法是最常用的二分支决策树算法,当因变量为离散变量时生成分类树,而当因变量为连续变量时构成回归树。本研究的因变量为电动自行车事故伤害严重程度,为离散变量,故 CART 生成的

是分类树。分类树生成过程是将样本集根据各分类属性(自变量)的不同属性不断分支最终得到因变量同质性最高的“纯”叶子节点的过程。

CART 算法使用 G_{ini} 值来衡量节点(样本集)内同质性,其值越大,样本的同质性越低。通过计算该节点内因变量为某一值的样本数占样本总数的比值来确定同质性,节点 t 的 G_{ini} 值为:

$$G_{\text{ini}}(t) = \sum_{i \neq j} p(j|t)p(i|t) \quad (1)$$

$$p(j|t) = \frac{p(j,t)}{p(t)} \quad (2)$$

$$p(j, t) = \frac{\pi(j)N_j(t)}{N_j} \quad (3)$$

$$p(t) = \sum_j p(j, t) \quad (4)$$

式中: i 和 j 分别为因变量的某一取值; $\pi(j)$ 为样本因变量取值为 j 的先验概率; $N_j(t)$ 为节点 t 内因变量取值为 j 的样本数; N_j 为因变量取值为 j 的样本总数。

此外, G_{ini} 值可用于衡量分类树内各自变量对因变量的重要程度, 这也是分类树方法输出的一个重要结果。自变量 x_j 的重要程度 (V_{IM}) 为:

$$V_{\text{IM}}(x_j) = \sum_{t=1}^T \frac{N_t}{N} \Delta G_{\text{ini}}(S_{x_j, t}) \quad (5)$$

式中: T 为分类树的节点总数; N_t 为节点 t 的样本数; N 为样本总数; $\Delta G_{\text{ini}}(S_{x_j, t})$ 为在节点 t 处根据自变量 x_j 分支后, 子节点 G_{ini} 值之和与父节点 t G_{ini} 值的差, 即 G_{ini} 值的减少量。

由式(5)可知, 自变量的重要程度实际上是各自变量根据样本数加权后在整个分类树种中 G_{ini} 值降低量的总和。也就是说, 在分类树模型中, 自变量提升样本同质性的效果越显著, 其重要程度越高。

2.2 基于决策树的电动自行车事故严重性影响因素分析

采用 CART 方法探究电动自行车事故伤害严重程度与骑行者特征、违法行为、道路特征及天气环境特征的交互效应。原始的事故伤害严重程度分为死亡、重伤、轻伤、仅财产损失四类, 根据以往研究成果, 构建分类树模型时采用二分类因变量能获得最好的预测精度, 故将伤害严重程度合并为严重(死亡与重伤)和非严重(轻伤与仅财产损失)两类。借助 SPSS 统计分析工具进行分类树模型构建。

将约 50% 的事故数据作为训练样本用于生成决策树, 另外 50% 作为检验样本对决策树的预测效果进行检验。最终模型对训练样本的总体预测正确率为 88.60%, 对检验样本的总体预测正确率为 86.80%, 对全部样本集的预测正确率为 87.70%, 预测精度高, 表明构建分类树能很好地拟合事故伤害严重程度与上述因素的关联。

2.2.1 决策树模型结果

利用事故数据使用 CART 算法生成分类树(见图 1), 该决策树的最大深度为 5 层, 共进行 8 次分

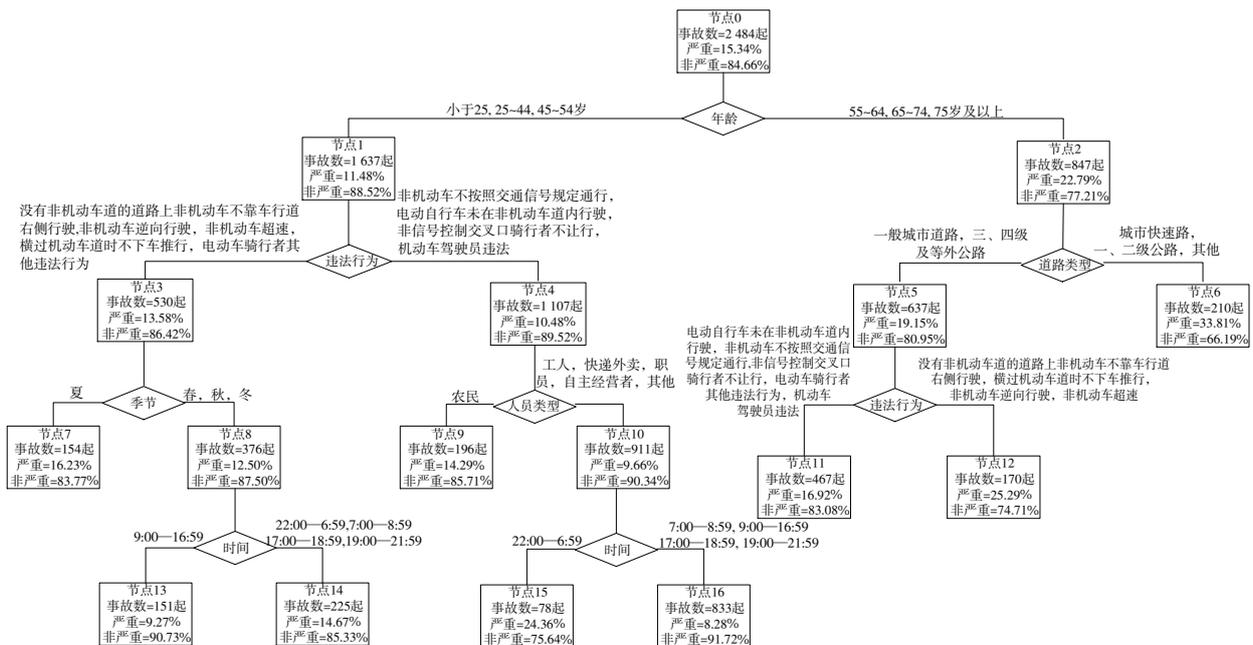


图 1 决策树模型结果

类, 最终产生 9 个叶节点, 叶节点的严重事故比例涵盖 8.28%~33.81%, 其中叶节点 6、7、11、12、15 的严重事故占比高于所有事故严重伤害平均占比。

2.2.2 决策树节点分析

节点 0 根据年龄分为子节点 1 和 2。显然节点

2 分支年龄在 55 岁以上的严重事故率显著高于年龄在 55 岁以下的节点 1 分支。

节点 1 根据违法行为分为子节点 3 和 4。相比其他违法行为, 由在没有非机动车道的道路上电动自行车不靠车行道右侧行驶、电动自行车逆向行驶、

电动自行车超速、横过机动车道时不下车推行及电动车骑行者其他违法行为导致的事故更易出现严重伤害(分别为13.58%、10.58%)。节点3再根据季节划分为叶节点7和子节点8。夏季发生的事故导致严重伤害的可能性(16.23%)高于其他3个季节(12.50%)。节点8根据时间划分为两个叶节点13和14。发生在9:00—16:00时段的事故导致严重伤害的可能性(9.27%)低于其他时段(14.67%)。回到节点4,根据人员类型划分为叶节点9和子节点10。相比骑行者类型为工人、快递外卖、职员、自主经营者及其他人员,骑行者类型为农民更易出现严重伤害事故(分别为14.29%、9.66%)。节点10继续划分为两个叶节点15和16。发生在22:00—6:59时段的事故导致严重伤害的可能性(24.36%)远高于其他时段(8.28%)。

节点2根据道路类型划分为子节点5及叶节点6。发生在城市快速路,一、二级公路及其他公路上的事故导致严重伤害的可能性(33.81%)远高于发生在一般城市道路,三、四级及等外公路上的事故(19.15%)。节点5根据违法行为继续划分为两个叶节点11和12。由在没有非机动车道的道路上电动自行车不靠车行道右侧行驶、横过机动车道时不下车推行、电动自行车逆向行驶及超速导致的事故更易出现严重伤害(25.29%)。

2.3 事故严重性影响因素分析

分类树方法不仅能以树状图形式展示电动自行车事故伤害严重性与各影响因素的交互效应,同时能获知各因素对事故伤害的重要程度(见表2)。

表2 各因素对事故伤害的相对重要性

因素	相对重要性/%	因素	相对重要性/%
驾驶者年龄	100	职业	42
违法行为	77	时间	40
道路类型	66	季节	30
户口性质	47	天气	3

(1) 骑行者年龄。骑行者年龄是影响电动自行车事故伤害严重性的最重要因素,也是决策树第一层的分类变量。老年骑行者更易导致严重伤害事故,这与以往传统自行车与摩托车事故伤害的研究结论一致。老年人群风险感知与避让能力较差、身体构造更脆弱,这些都使老年电动车骑行者更易出现严重事故。

(2) 违法行为。违法行为是重要性排名第二的事故伤害严重性影响因素。无论是年轻骑行者还是

老年骑行者,在没有非机动车道的道路上电动自行车不靠车行道右侧行驶、电动自行车逆向行驶、电动自行车超速、横过机动车道时不下车推行的违法行为都易导致严重伤害事故。

(3) 道路类型。道路类型是重要性排名第三的事故伤害严重性影响因素。发生在城市快速路,一、二级公路及其他公路上的事故更易出现严重伤害,对老年骑行者群体尤为显著。其原因在于该类型道路限速高,机动车行驶速度快,一旦与电动自行车发生碰撞事故(尤其是涉及老年骑行者),往往会导致严重伤害甚至死亡。

(4) 户口性质与职业。户口性质与职业分别是重要性排名第四、第五的事故伤害严重性影响因素。农民骑行者更易出现严重伤害事故,尤其是对于年龄小于55岁的骑行者群体。

(5) 时间与季节。时间与季节也与事故伤害严重性有关联。22:00—6:59时段发生的事故导致严重伤害的可能性最高,而发生在9:00—16:00时段的事故导致严重伤害的可能性最低;发生在夏季的事故更易导致严重伤害。这些情境尤其针对年龄小于55岁的骑行者群体。

(6) 天气。天气与事故伤害严重性的关联性较低,这主要是由于电动自行车骑行者多选择在天气良好的条件下出行。

3 严重事故高发情境及改善建议

3.1 严重事故高发情境分析

根据分类树挖掘事故严重程度较高的交通情境,即多种因素组合下严重事故高发情境。对于电动自行车事故,主要的严重事故高发情境包括:

(1) 事故发生在城市快速路和一、二级公路上,且骑行者年龄 ≥ 55 岁(310起,严重事故占比为33.81%)。

(2) 事故发生在一般城市道路和三、四级及等外公路上,骑行者年龄 ≥ 55 岁,事故发生前出现在没有非机动车道的道路上不靠车行道右侧行驶、横过机动车道时不下车推行、逆向行驶及超速行驶等违法行为(170起,严重事故占比为25.29%)。

(3) 事故发生在22:00—6:59时段,骑行者年龄 < 55 岁,事故发生前出现不按交通信号规定通行、未非机动车道内行驶、在非信号控制交叉口不让行及机动车驾驶员违法等违法行为(78起,严重事故占比为24.36%)。

(4) 事故发生在一般城市道路和三、四级及等外公路上,骑行者年龄 ≥ 55 岁,事故发生前出现未非机动车道内行驶、不按交通信号规定通行、在非信号控制交叉口不让行及其他违法行为、机动车驾驶员违法等(637起,严重事故占比为19.15%)。

(5) 事故发生在夏季,骑行者年龄 < 55 岁,事故发生前出现在没有非机动车道的道路上不靠车行道右侧行驶、逆向行驶、超速行驶、横过机动车道时不下车推行及其他违法行为(154起,严重事故占比为16.23%)。

上述交通情境多数发生于年龄 ≥ 55 岁的骑行者,与前文所述年龄是最重要的影响因素相吻合。识别这些严重事故高发情境,可为有针对性地实施电动自行车交通安全改善措施提供理论参考,有利于加强对电动自行车骑行者认知自身事故风险状况的预警,促使他们规避高风险区域、高风险时段并减少不安全骑行行为。

3.2 改善建议

(1) 加强对老年骑行者、特定违法行为的管控。骑行者年龄与事故前违法行为是影响电动自行车事故伤害严重程度最重要的因素。针对老年骑行者、特定违法行为(包括电动自行车逆向行驶、在没有非机动车道的道路上不靠车行道右侧行驶、横过机动车道时不下车推行等)的安全教育与执法应作为改善电动自行车交通安全的优先实施策略。

(2) 大力改善高等级道路慢行交通环境,加强治理力度。发生在城市快速路和一、二级公路等限速较高道路的电动自行车事故占比较大,且这些事故的严重程度高(尤其是对于老年骑行者)。建议:1) 电动自行车流量大的高等级道路尽可能开辟非机动车专用道,并与机动车道隔离,避免碰撞事故发生;2) 对于没有非机动车专用道的快速道路,严格执法,禁止电动自行车进入机动车道行驶,并开展安全教育。

(3) 加强电动自行车在特定时节下的出行管理。特定的时段与季节发生的事故中严重伤害事故比例较高,建议加强事故多发季节电动自行车的出行限制,强化对特定时段的执法。

4 结语

当前中国开始注重交通事故预测建模理论与实践研究,基于事故数据驱动的交通安全管理理念已逐步受到重视。电动自行车交通安全成为当前不容

忽视的安全问题。文中提取湖南省2014—2016年电动自行车交通事故数据,基于决策树方法构建电动自行车事故伤害严重程度模型,分析事故伤害严重程度与骑行者特征、违法行为、道路特征及天气环境等的交互效应,分析容易导致严重事故的交通情境,为电动自行车的安全管理提供参考。

参考文献:

- [1] 中国报告大厅.2016年我国电动自行车行业标准政策分析[EB/OL].[2016-04-13].<http://www.chinabgao.com/freereport/71818.html>.
- [2] KIM J K, KIM S, ULFARSSON G F, et al. Bicyclist injury severities in bicycle-motor vehicle accidents[J]. *Accident Analysis and Prevention*, 2007, 39(2): 238-251.
- [3] YAN X, MA M, HUANG H, et al. Motor vehicle-bicycle crashes in Beijing: Irregular maneuvers, crash patterns, and injury severity[J]. *Accident Analysis and Prevention*, 2011, 43(5): 1751-1758.
- [4] LIN M R, KRAUS J F. A review of risk factors and patterns of motorcycle injuries[J]. *Accident Analysis and Prevention*, 2009, 41(4): 710-722.
- [5] WU C X, YAO L, ZHANG K. The red-light running behavior of electric bike riders and cyclists at urban intersections in China: An observational study[J]. *Accident Analysis and Prevention*, 2012, 49(11): 186-192.
- [6] ZHANG Y Q, WU C X. The Effects of sunshields on red light running behavior of cyclists and electric bike riders[J]. *Accident Analysis and Prevention*, 2013, 52: 210-218.
- [7] HU F, LYU D L, ZHU J, et al. Related risk factors for injury severity of E-bike and bicycle crashes in Hefei[J]. *Traffic Injury Prevention*, 2014, 15(3): 319-323.
- [8] JUNG S, QIN X, OH C. Improving strategic policies for pedestrian safety enhancement using classification tree modeling[J]. *Transportation Research Part A (Policy & Practice)*, 2016, 85: 53-64.
- [9] TAVAKOLI K A, RABIEYAN R, BESHARATI M M. A data mining approach to investigate the factors influencing the crash severity of motorcycle pillion passengers[J]. *Journal of Safety Research*, 2014, 51: 93-98.
- [10] CHANG L Y, CHIEN J T. Analysis of driver injury severity in truck-involved accidents using a nonparametric classification tree model[J]. *Safety Science*, 2013, 51(1): 17-22.