

基于时依等比例风险回归模型的换道时长影响因素

李林波, 李 杨, 邹亚杰

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

摘要: 研究了基于时依等比例风险回归模型的换道时长(LCD)影响因素。利用开源 HighD 数据集, 提取了 560 条完整的换道(LC)轨迹, 对整体换道时长分布进行了非参数估计; 引入时依等比例风险回归模型, 对当前车辆换道时长的影响因素进行分析。结果表明: 大部分驾驶员在 3~8 s 内完成换道行为; 共有 6 个变量显著影响换道时长, 其中最为关键的为当前车辆自身车速, 同时当前车道的前车与目标车道的后车对换道时长的影响也非常显著。

关键词: 换道时长(LCD); 生存分析; 时依等比例风险回归模型; HighD 数据集

中图分类号: U491

文献标志码: A

Factors Affecting Lane-changing Duration Based on Time-dependent Proportional-hazards Regression Model

LI Linbo, LI Yang, ZOU Yajie

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

Abstract: Based on the time-dependent proportional-hazards regression model, factors affecting lane-changing duration(LCD) are studied. Five hundred and sixty complete lane-changing(LC) trajectories are extracted from the newly opened HighD dataset, and the overall distribution of lane-changing duration is estimated using the non-parametric estimator. Then, the time-dependent proportional-hazards regression model is introduced to analyze the factors affecting LCD. It is found that most drivers finish LC in 3~8 s. There are six variables, of which the most important one is current vehicle's speed. Additionally, the preceding vehicle in current lane and the following vehicle in target lane have significant effects on LCD.

Key words: lane-changing duration(LCD); survival analysis; time-dependent proportional-hazards regression model; HighD dataset

换道行为是指车辆(当前车辆)在行驶过程中由当前车道换至目标车道的行为, 不仅涉及与前方车辆的交互影响, 还受到目标车道前后车的影响。在现实交通场景中, 由换道行为导致的交通事故居高不下, 探索车辆换道行为的影响机理和决策机制迫在眉睫^[1-3]。目前关于换道行为的研究大致可以划分为车辆换道决策过程建模^[4]、车辆换道行为影响^[4]、车辆换道轨迹规划和追踪建模^[5-7]、车辆换道轨迹预测^[8]以及换道时长影响因素^[9-16]。换道时长描述了当前车辆由当前车道中心线驶至目标车道中心线所需要花费的时间。

对于换道时长的研究主要集中于换道时长分布以及换道时长影响因素。研究表明, 换道时长普遍在 1~16 s 内, 换道时长和时间段、道路类型、车流密度、车辆类型、换道方向、驾驶风格以及当前车辆与周围车辆的交互影响有关。符铎等^[17]的研究表明, 不同交通状况下高速公路换道时长服从对数正态分布, 换道时长与车辆换道影响区域内的交通状况存在显著联系。张颖达等^[18]指出, 强制换道的平均换道时长稍大于自由换道, 影响 2 类换道行为换道时长的因素不同, 各因素的影响程度也不同。Toledo 等^[9]采用多变量回归模型对 NGSIM(Next Generation Simulation)数据集集中的换道时长进行了回归建模, 车流密度、换道方向以及周边车辆干扰会显著影响换道时长的大小。Wu 等^[16]引入了半参数等比例风险回归模型对换道时长进行建模, 不同时间段的换道时长存在显著的差异性。Vlahogianni^[10]采用全参数的加速失效模型对车辆超车时长进行了建模, 当前车辆与当前车道前车的速度差、当前车辆与当前车道前车的距离、目标车道后车车速以及驾驶员性别等变量显著影响换道时长。Yang 等^[11]采用多元回归模型对换道时长进行研究, 获得的结果

收稿日期: 2021-04-10

基金项目: 国家重点研发计划(2018YFE0102800)

第一作者: 李林波(1974—), 男, 副教授, 博士生导师, 工学博士, 主要研究方向为交通规划、交通拥挤管理等。

E-mail: llinbo@tongji.edu.cn

通信作者: 李 杨(1996—), 男, 博士生, 主要研究方向为智能交通、交通规划。E-mail: cc960719@tongji.edu.cn



论文
拓展
介绍

与 Toledo 等^[9]的结果基本一致。以上研究均依靠换道过程中某一时刻的状态信息(如与周边车辆的相对位置、相对速度、相对加速度等)来探究换道时长的影响因素,而未曾考虑这些状态信息在换道过程中的动态特征。因此,本研究中试图将时依等比例风险回归模型^[19-21]引入对换道时长的建模中。作为生存分析类模型,该模型不仅可以研究影响因素与生存时间和结局的关系^[10,22-23],还能融入换道过程中的动态信息,从而完善对于换道时长机理的解析,以便更深入地把握车辆换道行为的内在特性。相比于 NGSIM 数据集^[24],本研究中将采用具有更多车流轨迹数据与更低测量误差的 HighD 数据集^[25]进行分析,从而提供高质量的数据支撑。

1 换道时长模型建构

1.1 换道时长生存分析

生存分析是对生存时间的分布特征进行描述,并对影响生存时间的主要因素进行分析的一种统计方法。将生存分析应用到换道时长分析中,所对应的起始时间为换道的起始时刻,即换道起点时刻;截止时间也可称为失效时间,指换道结束时刻。广义生存时间为换道时长,指换道起点时刻至结束时刻所经历的时长,与之对应的基础公式如下所示:

$$\begin{aligned} f(t) &= \lim_{\Delta t \rightarrow 0} \frac{P_1(t, t + \Delta t)}{\Delta t} \\ F(t) &= \int_0^t f(t) dt \\ S(t) &= 1 - F(t) = e^{-H(t)} \\ h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P_2(t, t + \Delta t)}{\Delta t} = \frac{f(t)}{S(t)} \\ H(t) &= \int_0^t h(t) dt = -\ln S(t) \end{aligned}$$

式中: $P_1(t, t + \Delta t)$ 为车辆在 $(t, t + \Delta t)$ 内完成换道的概率; $P_2(t, t + \Delta t)$ 为时间 t 后车辆在 $(t, t + \Delta t)$ 内完成换道的概率; $f(t)$ 为各时刻车辆完成换道的概率; $F(t)$ 为 $f(t)$ 的分布函数; $S(t)$ 为车辆换道时长大于 t 的概率; $h(t)$ 为时间 t 后车辆完成换道的概率; $H(t)$ 为累积危险率函数。事实上,只要得知其中任意一个函数,就可以推导出其余公式。

1.2 基于时依等比例风险回归模型的换道时长建模

生存分析类函数大致可以划分为非参数类模型、半参数类模型以及参数类模型^[22]。非参数类模型在估计生存函数时对生存时间的分布没有要求,根据样本观测值提供的信息,用公式计算出每一个

时间点的生存函数、死亡函数、风险函数等。参数类模型根据样本观测值来估计假定的分布模型中的参数,从而获得生存函数、死亡函数、风险函数等,常用的分布有指数分布、Weibull 分布、对数正态分布、对数 Logistic 分布以及 Gamma 分布等。半参数类模型不需要对生存时间的分布做出假定,但是却可以通过一个模型来分析生存时间的分布规律以及危险因素对生存时间的影响。在以上 3 类模型中,半参数类模型在学界得到普遍使用。由于参数类模型对于模型的假设有着较高的要求,因此当假设不满足时效果较差。非参数类模型未作任何假设,因而难以解释各自变量对于因变量的影响程度,模型的可解释性较差。半参数类模型则结合了参数类模型以及非参数类模型的特点,一方面放宽了对生存分布的假设,另一方面又保留了具有可解释的参数部分^[22],因而有着更加稳健的表现。

半参数类模型中最为著名的为比例风险回归模型^[20],该模型并非直接基于生存函数而是基于危险率函数进行建模,由基准危险率方程以及一个参数方程组成,如下所示:

$$h(t|x_i) = h_0(t) \exp\left(\sum_{i=1}^n b_i x_i\right)$$

式中: $h_0(t)$ 为基本危险率函数; x_i 为自变量(影响换道时长的因素); b_i 为自变量对应的偏回归系数。

该模型中未对 $h_0(t)$ 作任何假定,是一个非参数部分的建模,回归系数 b_i 则是需要根据样本进行估计的参数,是一个参数部分的建模。该模型中仅是以某一时刻的变量信息 x_i 对因变量进行建模,然而在对因变量进行回归分析的过程中,自变量有可能也随着时间而变化,因此必须考虑自变量的时变特性。Fisher 等^[19]提出了考虑自变量时变特性的时依等比例风险回归模型,表达式如下所示:

$$h(t|x_i(t)) = h_0(t) \exp\left(\sum_{i=1}^n b_i x_i(t)\right)$$

$$S(t|x_i(t)) = \exp\left(-\int_0^t h_0(t) \exp\left(\sum_{i=1}^n b_i x_i(t)\right) dt\right)$$

式中: $S(t|x_i(t))$ 为生存函数; $x_i(t)$ 为随时间变化的自变量。通过极大似然函数对以上参数进行估计,极大似然函数的表达式如下所示:

$$U(\beta) = \sum_{i=1}^n \delta_i(t) \left(x_i(t) - \frac{\sum_{j \in R_i} \exp\left(\sum_{k=1}^n b_k x_j(t)\right) x_j(t)}{\sum_{j \in R_i} \exp\left(\sum_{k=1}^n b_k x_j(t)\right)} \right)$$

式中: $\delta_i(t)$ 为二元变量,若车辆在 t 时刻完成换道,其值为1,反之为0; R_i 为不同观测时刻的风险集合, $R_i(t) = \{t, x_i(t), \delta_i(t)\}$ 。不同于时不变等比例风险回归模型,在每个风险集中,自变量的值均在变化。

2 数据说明及轨迹提取

HighD数据采集于2017年至2018年间,采集地点为德国高速公路,采集时长为16.5 h,采集路程为45 000 km,总计包含11万辆车的信息。使用先进的计算机视觉算法、神经网络及贝叶斯平滑算法提取

出的轨迹数据为每秒40帧,数据集中每条轨迹的坐标误差不到10 cm,这主要得益于该数据集以4K(4 096×2 160)格式采集的缘故^[25]。

为了探索更为复杂情形下的换道时长机理,对换道过程中同时包含5辆车的换道轨迹进行提取与分析,如图1所示。首先,仅保留小汽车的换道轨迹,剔除卡车的换道轨迹,剔除不同时含有周边4辆车信息的小汽车换道轨迹;其次,以车辆在纵向方向上的速度为主要参考指标,同时结合车辆的加速度推断车辆换道的起始点以及截止点,如图2所示;最后,进行人工筛选以确定最终的每一条轨迹,确保同时包含5辆车的信息以及在目标车道换道过程中周边车辆的状态信息无任何的缺失。

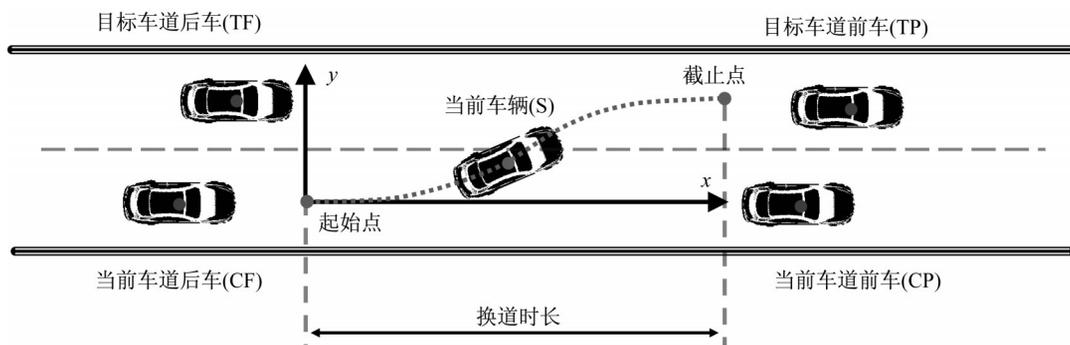


图1 换道轨迹示意图

Fig.1 Schematic diagram of LC trajectory

总计提取560条有效换道轨迹,HighD数据集的每个tracks文件中所包含的有效换道次数以及该数据集换道时长整体分布如图3所示。车辆的平均换道时长为5.69 s,中位数为5.55 s,标准差为1.35 s。

可以发现,中位数要略低于平均值,这说明多数车辆的换道时长是低于中位数的,部分车辆过长的换道时长拉大了整体的平均值。图4为换道轨迹集中的2条换道轨迹样本。各变量定义如表1所示。

表1 各变量定义

Tab.1 Definition of variables

| 车辆 | 变量 | 单位 | 定义 |
|--------|-------------------------|------------------|-----------------|
| | t_{LCD} | s | 换道时长 |
| 当前车辆 | Status | | 车辆是否完成换道状态(0或1) |
| | \dot{x}_S | $m \cdot s^{-1}$ | 当前车辆车速 |
| | \ddot{x}_S | $m \cdot s^{-2}$ | 当前车辆加速度 |
| 当前车道前车 | \ddot{x}_{CP} | $m \cdot s^{-2}$ | 当前车道前车加速度 |
| | Δx_{S-CP} | m | 当前车辆与当前车道前车距离 |
| | $\Delta \dot{x}_{S-CP}$ | $m \cdot s^{-1}$ | 当前车辆与当前车道前车速度差 |
| 当前车道后车 | \ddot{x}_{CF} | $m \cdot s^{-2}$ | 当前车道后车加速度 |
| | Δx_{S-CF} | m | 当前车辆与当前车道后车距离 |
| | $\Delta \dot{x}_{S-CF}$ | $m \cdot s^{-1}$ | 当前车辆与当前车道后车速度差 |
| 目标车道前车 | \ddot{x}_{TP} | $m \cdot s^{-2}$ | 目标车道前车加速度 |
| | Δx_{S-TP} | m | 当前车辆与目标车道前车距离 |
| | $\Delta \dot{x}_{S-TP}$ | $m \cdot s^{-1}$ | 当前车辆与目标车道前车速度差 |
| 目标车道后车 | \ddot{x}_{TF} | $m \cdot s^{-2}$ | 目标车道后车加速度 |
| | Δx_{S-TF} | m | 当前车辆与目标车道后车距离 |
| | $\Delta \dot{x}_{S-TF}$ | $m \cdot s^{-1}$ | 当前车辆与目标车道后车速度差 |

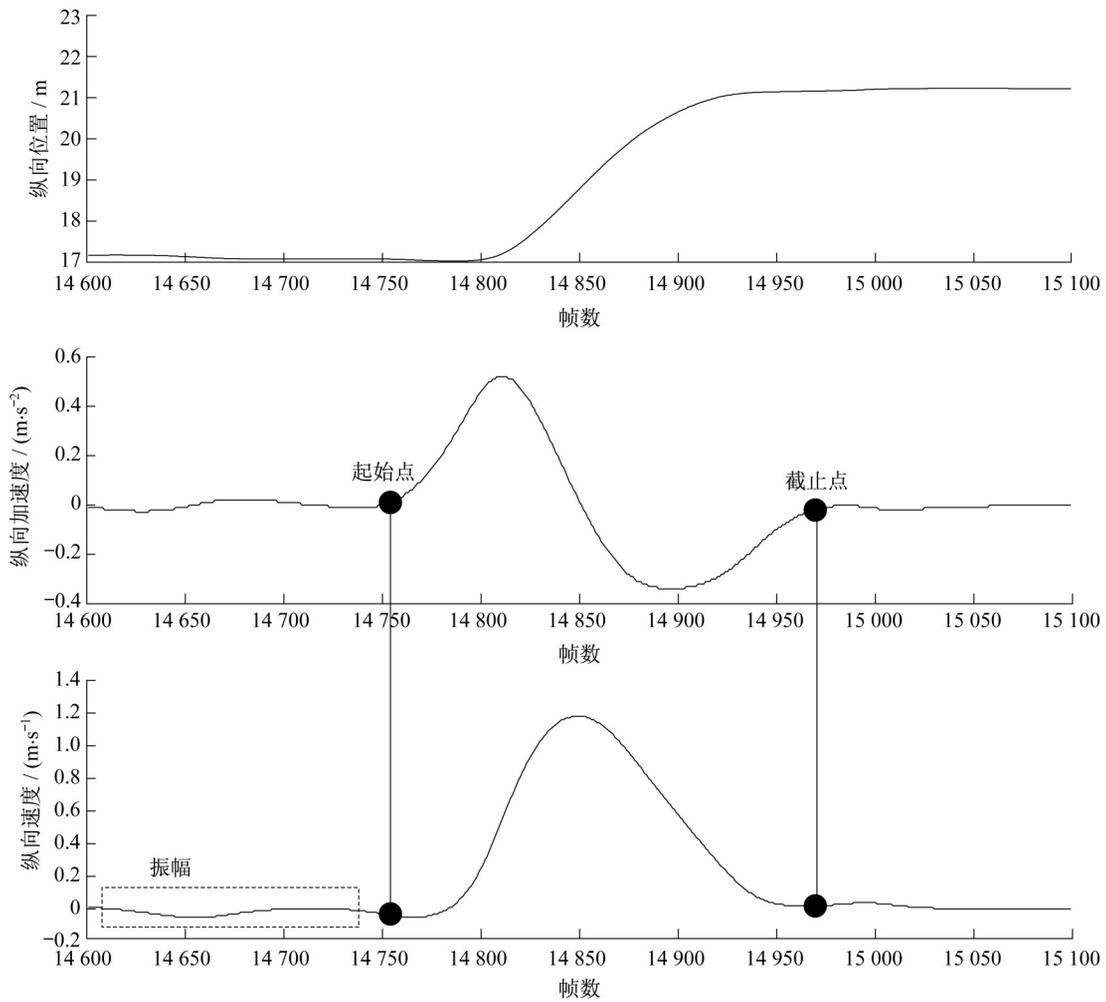


图2 某条换道轨迹起始点和截止点示意图(数据采集间隔为0.025 s)

Fig.2 Schematic diagram of starting point and ending point of a LC trajectory(data acquisition interval is 0.025 s)

3 模型验证及分析

3.1 整体分布结果分析

对换道时长进行回归建模,采用非参数方法KM(Kaplan-Meier)和NA(Nelson-Aalen)方法^[26]对换道时长的整体部分进行估计,以获得换道时长生存函数的一个初步分析。估计方法如下所示:

$$S_{KM}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i}$$

$$H_{NA}(t) = \prod_{t_i < t} \frac{d_i}{n_i}$$

式中: d_i 为在时刻 t 完成换道的车辆数; n_i 为在时刻 t 前完成换道的车辆数。

图5为基于整体生存函数以及累积风险率函数的换道时长分布。可以发现,生存函数曲线在3 s至8 s之间时陡然下降,意味着大部分车辆在3 s至8 s

之间完成换道。同时,引入中位生存时间 α_{MST} ,其定义为平均50%的换道时长失效的时间,这意味着每辆车都有50%的机会完成换道操作。采用KM方法所得的中位生存时间为4.975 s(置信区间下限为4.900 s,上限为5.050 s)。在3~8 s时,生存曲线快速下降,而在8~12 s时曲线平缓下降。累积危险率函数同样也是如此,先在3~8 s时快速上升后在8~12 s时平缓上升。从表2可以得出,近15%的车辆于4 s内完成换道,近79%的车辆于6 s内完成换道,近96%的车辆于8 s内完成换道,而换道时长大于8 s的车辆仅占4%。

3.2 回归模型结果分析

由于HighD数据集的采集间隔为0.025 s,以0.100 s为间隔进行重新集聚(取其平均值)。若每条轨迹换道时长为 T ,则每条轨迹共有观测记录 $T/0.100$ 条。每条观测记录均包含当前时间点周边

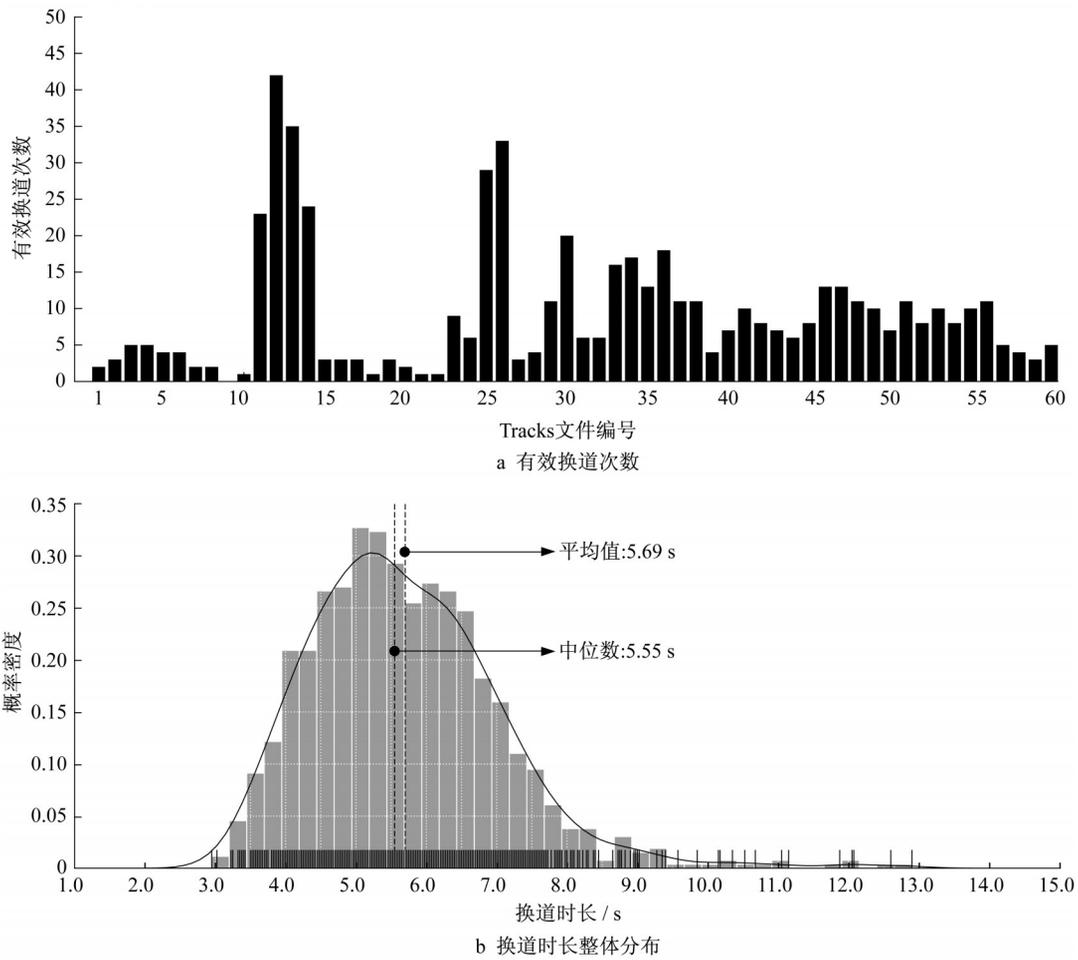


图 3 HighD数据集的每个 tracks 文件中有效换道次数和换道时长整体分布

Fig.3 Effective LC counts for each track file in HighD dataset and overall distribution of LCD

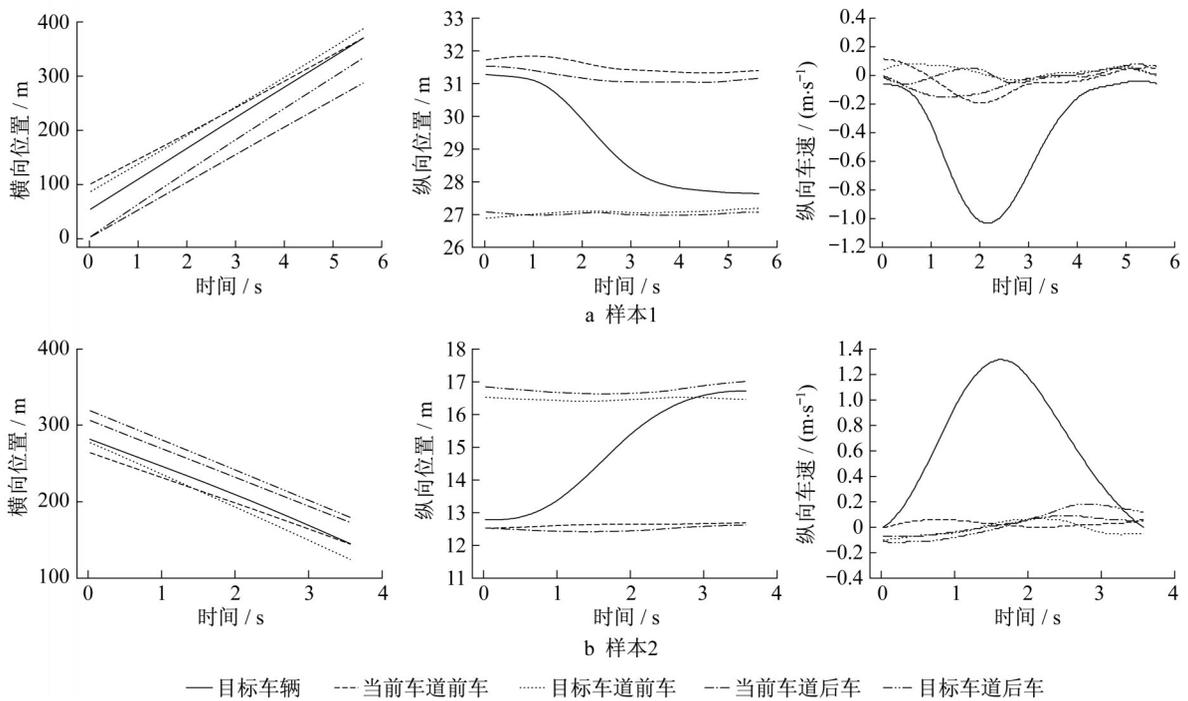


图 4 HighD数据集中2条换道轨迹样本

Fig.4 Two examples of LC trajectories in HighD dataset

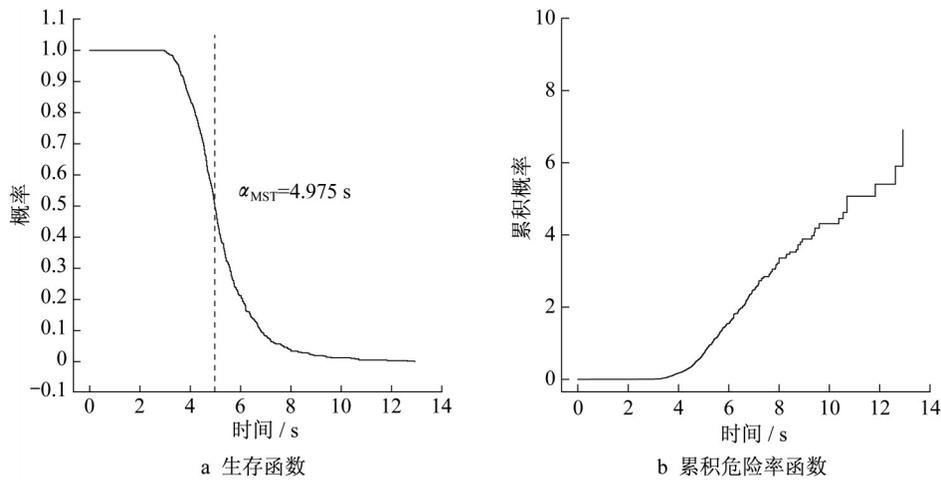


图5 基于整体生存函数和累积危险率函数的换道时长分布(KM和NA方法)

Fig.5 Distribution of LCD based on overall survival function and cumulative hazard function (KM and NA method)

表2 各换道时长下尚未完成换道操作车辆所占百分比
Tab.2 Percentage of at-risk vehicles at different LCDs

| 各换道时长下未完成换道操作车辆所占百分比/% | | | | | | |
|------------------------|--------|-------|-------|------|------|------|
| 0 s | 2 s | 4 s | 6 s | 8 s | 10 s | 12 s |
| 100.00 | 100.00 | 84.46 | 21.25 | 3.93 | 1.25 | 0.36 |

车辆的速度、加速度以及位置信息。图6为时依等比例风险回归模型的输入数据格式。以0.100 s为时间间隔输入各时刻周边车辆的状态信息,为防止模型出现过拟合以及提高模型的泛化性能,引入L1和L2正则化惩罚项,并将不同轨迹的ID变量设置为聚类变量。

| 每条观测记录 | ID | 起始时间点/s | 截止时间点/s | 时变变量1 | 时变变量2 | 时变变量3 | 其他变量 | 是否完成换道 |
|--------|-----|---------|---------|-------|-------|-------|------|--------|
| 1 | 1 | 0.0 | 0.1 | 0.1 | 5.0 | 11.0 | ... | False |
| 1 | 1 | 0.1 | 0.2 | 0.2 | 6.0 | 11.1 | ... | False |
| 1 | 1 | 0.2 | 0.3 | 0.3 | 6.5 | 11.2 | ... | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1 | 1 | 5.9 | 6.0 | 6.0 | 8.0 | 15.0 | ... | True |
| 2 | 2 | 0.0 | 0.1 | 0.3 | 7.0 | 12.0 | ... | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

图6 时依等比例风险回归模型数据输入格式

Fig.6 Data input format of time-dependent proportional-hazards regression model

时依等比例风险回归模型模拟结果如表3所示。取0.05的显著水平,共有6个变量显著影响换道时长,分别是 \dot{x}_s 、 \ddot{x}_{CP} 、 Δx_{S-CP} 、 Δx_{S-CF} 、 \ddot{x}_{TF} 、 Δx_{S-TF} 。

随着 \dot{x}_s 、 Δx_{S-CP} 、 Δx_{S-CF} 、 \ddot{x}_{TF} 、 Δx_{S-TF} 的增加,车辆有更短换道时长的概率将会更低,而随着 \ddot{x}_{CP} 的增加,换道时长增加的概率则会越高。以当前车辆的速度 \dot{x}_s 为例,其在模型中的回归系数为0.117,则 $\exp(0.117) = 1.124$ 。随着当前车辆速度每增加 $1 \text{ m}\cdot\text{s}^{-1}$,其危险率函数则会相应增加 $12.4\% = (1.124 - 1.000) \times 100\%$,其中1.000代表基准危险

率,因而当前车辆将会有更短的换道时长,反之,当前车辆将会需要更长的换道时长。可以发现,最为显著影响换道时长的3个变量分别为 \dot{x}_s 、 \ddot{x}_{CP} 、 \ddot{x}_{TF} ,目标车道前车的3个变量均不显著,当前车道后车仅有一个变量显著,而当前车道前车以及目标车道后车均有2个变量显著。因此,可以认为当前车道前车以及目标车道后车影响当前车辆换道时长的程度要高于当前车道后车以及目标车道前车。最后,需要对模型进行风险比例假设的检验(Schoenfeld检验),这是由于比例风险回归模型运用的前提是假设

表3 时依等比例风险回归模型模拟结果

Tab.3 Simulation results of time-dependent proportional-hazards regression model

| 车辆 | 变量 | 回归系数 a | $\exp(a)$ | 标准差 | 统计值 | 显著性 |
|--------|-------------------------|----------|-----------|-------|--------|-------|
| 当前车辆 | \dot{x}_S | 0.117 | 1.124 | 0.011 | 10.826 | 0.000 |
| | \ddot{x}_S | 0.143 | 1.154 | 0.116 | 1.270 | 0.204 |
| 当前车道前车 | \ddot{x}_{CP} | -0.397 | 0.673 | 0.153 | -2.860 | 0.004 |
| | Δr_{S-CP} | 0.009 | 1.009 | 0.002 | 4.424 | 0.000 |
| | $\Delta \dot{x}_{S-CP}$ | -0.004 | 0.996 | 0.020 | -0.283 | 0.777 |
| 当前车道后车 | \ddot{x}_{CF} | 0.146 | 1.158 | 0.126 | 1.315 | 0.188 |
| | Δr_{S-CF} | 0.009 | 1.009 | 0.002 | 5.806 | 0.000 |
| | $\Delta \dot{x}_{S-CF}$ | -0.030 | 0.970 | 0.023 | -1.482 | 0.138 |
| 目标车道前车 | \ddot{x}_{TP} | 0.073 | 1.076 | 0.145 | 0.534 | 0.593 |
| | Δr_{S-TP} | 0.001 | 1.001 | 0.002 | 0.606 | 0.545 |
| | $\Delta \dot{x}_{S-TP}$ | -0.011 | 0.989 | 0.020 | -0.589 | 0.556 |
| 目标车道后车 | \ddot{x}_{TF} | 0.301 | 1.352 | 0.142 | 2.320 | 0.020 |
| | Δr_{S-TF} | 0.009 | 1.009 | 0.002 | 5.662 | 0.000 |
| | $\Delta \dot{x}_{S-TF}$ | -0.012 | 0.988 | 0.020 | -0.648 | 0.517 |

任意2个样本均不随时间变化。上述6个变量检验结果如表4所示,可以发现模型符合风险比例假设的前提。

表4 比例风险检验结果

Tab.4 Schoenfeld test results

| 变量 | 统计量 | 自由度 | 显著性 |
|-------------------|-------|-----|------|
| \dot{x}_S | 0.023 | 1 | 0.88 |
| \ddot{x}_{CP} | 1.080 | 1 | 0.30 |
| Δr_{S-CP} | 0.366 | 1 | 0.42 |
| Δr_{S-CF} | 0.093 | 1 | 0.76 |
| \ddot{x}_{TF} | 1.012 | 1 | 0.31 |
| Δr_{S-TF} | 0.023 | 1 | 0.88 |
| 总计 | 2.970 | 6 | 0.81 |

4 结语

采用时依等比例风险回归模型与HighD开源数据集,总计提取了560条完整的换道轨迹(换道时长平均值为5.69 s,中位数为5.55 s),对换道时长进行深入分析。引入非参数方法对该数据集中的整体换道时长分布进行估计,发现大部分车辆在3 s至8 s内完成了换道行为;引入时依等比例风险回归模型对换道时长进行建模,发现共有6个变量显著影响换道时长,对换道时长影响最大的3个显著变量分别为当前车辆自身的车速、当前车道前车加速度、当前车道后车加速度,同时当前车道前车以及目标车道后车对当前车辆换道时长的影响最为显著。显著影响换道时长的变量可作为换道决策的参考依据,甚至可以作为换道轨迹预测算法中的输入变量,同时换道时长的平均值、中位数、中位生存时间等指标均可用以衡量不同换道轨迹集间换道行为的差异性。受限于数据集,仅依据周边车辆的运动信息对

当前车辆换道时长的影响因素进行建模,后续研究可以采集驾驶员特征、天气状况、不同道路类型等数据做进一步的探索。

作者贡献说明:

李林波:相关概念及方案讨论,论文语言组织。

李 杨:数据处理及分析,实验设计及初稿撰写。

邹亚杰:论文审阅,结果分析及建议。

参考文献:

- [1] 吴兵,王文璇,李林波,等.多前车影响的智能网联车辆纵向控制模型[J].交通运输工程学报,2020,20(2):188.
WU Bing, WANG Wenxuan, LI Linbo, *et al.* Longitudinal control model for connected autonomous vehicles influenced by multiple preceding vehicle [J]. Journal of Traffic and Transportation Engineering, 2020,20(2):188.
- [2] 吴兵,罗雪,李林波.考虑驾驶风格的模糊控制跟驰模型[J].同济大学学报(自然科学版),2020,48(5):694.
WU Bing, LUO Xue, LI Linbo. Fuzzy control car-following model with the consideration of driving style [J]. Journal of Tongji University (Natural Science), 2020,48(5):694.
- [3] LI L, LI Y, NI D. Incorporating human factors into LCM using fuzzy TCI model [J]. Transportmetrica B: Transport Dynamics, 2021,9(1):198.
- [4] ZHENG Z. Recent developments and research needs in modeling lane changing [J]. Transportation Research, Part B: Methodological, 2014,60:16.
- [5] LUO Y, XIANG Y, CAO K, *et al.* A dynamic automated lane change maneuver based on vehicle-to-vehicle communication [J]. Transportation Research, Part C: Emerging Technologies, 2016,62:87.
- [6] YANG D, ZHENG S, WEN C, *et al.* A dynamic lane-changing trajectory planning model for automated vehicles [J]. Transportation Research, Part C: Emerging Technologies, 2018,95:228.

- [7] LI L, LI Y, NI D, *et al.* Dynamic trajectory planning for automated lane-changing[C]// Transportation Research Board 100nd Annual Meeting. Washington DC: TRB, 2021:1-21.
- [8] XIE G, GAO H, QIAN L, *et al.* Vehicle trajectory prediction by integrating physics- and maneuver-based approaches using interactive multiple models [J]. IEEE Transactions on Industrial Electronics, 2018,65(7): 5999.
- [9] TOLEDO T, ZOHAR D. Modeling duration of lane changes [J]. Transportation Research Record, 2007,1999(1): 71.
- [10] VLAHOIANNI E I. Modeling duration of overtaking in two lane highways [J]. Transportation Research, Part F: Traffic Psychology and Behaviour, 2013,20: 135.
- [11] YANG M, WANG X, QUDDUS M. Examining lane change gap acceptance, duration and impact using naturalistic driving data [J]. Transportation Research, Part C: Emerging Technologies, 2019,104: 317.
- [12] LI Y, LI L, NI D. Comprehensive survival analytical for lane-changing duration[C]// Transportation Research Board 100nd Annual Meeting. Washington DC: TRB, 2021:1-19.
- [13] WANG Q, LI Z, LI L. Investigation of discretionary lane-change characteristics using next generation simulation data sets [J]. Journal of Intelligent Transportation Systems, 2014, 18(3): 246.
- [14] OLSEN E C B, LEE S E, WIERWILLE W W, *et al.* Analysis of distribution, frequency, and duration of naturalistic lane changes [J]. Proceedings of Human Factors and Ergonomics Society Annual Meeting, 2002,46(22): 1789.
- [15] CAO X, YOUNG W, SARVI M. Exploring duration of lane change execution [C]// Australasian Transport Research Forum. Brisbane:[s.n.], 2013: 1-17.
- [16] WU J, ZHANG S, SINGH A K, *et al.* Hazard-based model of mandatory lane change duration[C]// CICTP 2017. Reston: American Society of Civil Engineers, 2018:805-811.
- [17] 符铎砂, 胡嘉诚, 何石坚. 基于交通状况及行驶速度的高速公路换道时间研究[J]. 公路交通科技, 2020,37(4): 133.
FU Xinsha, HU Jiacheng, HE Shijian. Study on expressway lane-changing time based on traffic condition and driving speed [J]. Journal of Highway and Transportation Research and Development, 2020, 37(4): 133.
- [18] 张颖达, 邵春福, 李慧轩, 等. 基于NGSIM轨迹数据的换道行为微观特性分析[J]. 交通信息与安全, 2015,21(6): 19.
ZHANG Yingda, SHAO Chunfu, LI Huixuan, *et al.* Speed characteristic analysis based on NGSIM trajectory data [J]. Journal of Transport Information and Safety, 2015,21(6): 19.
- [19] FISHER L D, LIN D Y. Time-dependent covariates in the Cox proportional-hazards regression model[J]. Annual Review of Public Health, 1999,20(1): 145.
- [20] COX D. Regression models and life tables [J]. Journal of the Royal Statistical Society, Series B: Methodological, 1972, 34(2): 187.
- [21] CHEN C, XIE Y. Modeling the safety impacts of driving hours and rest breaks on truck drivers considering time-dependent covariates[J]. Journal of Safety Research, 2014,51: 57.
- [22] WASHINGTON S P, KARLAFTIS M G, MANNERING F L. Statistical and econometric methods for transportation data analysis[M]. Boca Raton: Chapman and Hall/CRC, 2020.
- [23] 王文成, 杨小宝, 周映雪, 等. 基于Cox风险模型的城市道路交通拥堵持续时间分析[J]. 长安大学学报(自然科学版), 2015,35(Z1): 235.
WANG Wencheng, YANG Xiaobao, ZHOU Yingxue, *et al.* Cox's proportional hazard regression model for estimating urban traffic congestion duration [J]. Journal of Chang'an University (Natural Science Edition), 2015,35(Z1):235.
- [24] PUNZO V, BORZACCHIELLO M T, CIUFFO B. On the assessment of vehicle trajectory data accuracy and application to the Next Generation Simulation (NGSIM) program data [J]. Transportation Research, Part C: Emerging Technologies, 2011,19(6): 1243.
- [25] KRAJEWSKI R, BOCK J, KLOEKER L, *et al.* The HighD dataset: a drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems[C]// 2018 21st International Conference on Intelligent Transportation Systems (ITSC). [S.l.]:IEEE, 2018:1-8.
- [26] KAPLAN E L, MEIER P. Nonparametric estimation from incomplete observations[J]. Journal of the American Statistical Association, 1958,53: 457.