

基于 T-BOX 大数据的驾驶行为分析

电子电气与物理学院 电子信息工程 1901 班：张心怡

指导教师：方卫东

摘要：驾驶人作为汽车的直接操控者，其驾驶行为直接影响着道路交通安全。本文以 T-BOX 采集的数据为基础，将车辆行驶轨迹进行可视化还原并绘制了车速热力图，展示了行车路线、行车速度等驾驶信息。通过分析不良驾驶行为，将其以每日行程为单位，表征为行驶车速、疲劳驾驶和急变速三个特征。并进一步细分为六个具体指标，同时设计算法以实现不良驾驶行为的自动识别和判定。其次，基于不良驾驶行为特征提取结果，利用 K-means++ 聚类算法对驾驶行为进行聚类，从而进行倾向性分析。本文分别使用熵权法(Entropy Method)与层次分析法(Analytic Hierarchy Process, AHP)计算各级指标权重，最终采用两者结合的组合赋权法确定各级指标权重并构建驾驶行为评分模型。最终，使用评分模型对部分实例进行了评分，并根据倾向性判别结果对评分进行了有效性验证。根据最终的验证结果表明，该评分模型能够对汽车驾驶人的驾驶行为进行有效的综合评估。

关键字：T-BOX；驾驶行为分析；K-means++；组合赋权法；综合评价

Driving behavior analysis based on T-BOX big data

Abstract:As the direct controller of the car, the driver's driving behavior directly affects road traffic safety. Based on the data collected by T-BOX, this paper visualizes and restores the vehicle driving trajectory and draws a vehicle speed heat map, showing driving information such as driving route and driving speed. By analyzing the bad driving behavior, it is characterized as three features of driving speed, fatigue driving and sharp shifts in terms of daily trips. It is further subdivided into six specific indicators, and the algorithm is designed to automatically identify and determine the bad driving behavior. Second, based on the results of bad driving behavior feature extraction, the K-means++ clustering algorithm is used to cluster the driving behaviors so as to perform propensity analysis. The Entropy Method and Analytic Hierarchy Process (AHP) were used to calculate the weight of indicators at each level, and the combined weighting method was used to determine the weight of indicators at each level and construct the driving behavior scoring model. Finally, the scoring model was used to score some of the instances, and the validity of the scores was verified based on the propensity discriminant results. According to the final validation results, the scoring model can provide an effective comprehensive assessment of the driving behavior of car drivers.

Keywords: T-BOX; driving behavior analysis; K-means++; combined assignment method; comprehensive evaluation

1. 目的意义

随着国民经济水平不断增长，人民的生活水平也在不断提高，汽车已经在人们的生活中扮演着必需品的角色。频发的交通事故给人民的生命安全造成了极大的威胁。当今车联网技术高速发展，交通问题的改善有赖于该类新兴技术的支持。

交通安全与驾驶行为息息相关。研究驾驶行为并提出科学合理的综合评价方法不仅有助于减少交通事故的发生和加强社会秩序稳定，而且有利于引导驾驶员养成科学良好的驾驶习惯并提出改进驾驶行为的针对性意见。驾驶行为倾向性判定和驾驶行为综合评价能为智能驾驶辅助系统提供科学准确的安全风险评估，对保护生命和财产具有重要的显示意义。驾驶行为分析也为保险行业的车险业务提供了技术基础，帮助车险公司进行客户筛选、费率定价和理赔等工作，从而为车辆车险理赔提供支持。

2. 行车轨迹可视化

在进行驾驶行为研究之前，需要先对原始数据进行预处理，并分析影响驾驶行为的影响因素以及驾驶行为研究方案。本文使用 python 调用第三方库 Folium 绘制车速和加速度的热力图来进行轨迹数据的可视化。

图 2-1 为使用 Folium 工具进行车速和轨迹可视化还原后的 L**36145 车辆在 2019 年 3 月间车速热力图。



图 2-1 车辆 L**36145 2019 年 3 月间的速度热力图

3. 不良驾驶行为评价指标及判定

结合现有实验数据进行分析，本文最终以一天的行程为单位行程，使用行驶车速、疲劳驾驶行为和急变速行为三个不良驾驶行为作为表征参数。并继续对其细分，分为了平均车速、超速次数、连续驾驶时长、疲劳驾驶次数、急加速次数和急减速次数六个二级指标。选取的指标为第 5 章构建驾驶行为综合评分模型提供了层次结构和理论依据，具体的不良驾驶行为指标选取如表 3-1 所示。

表 3-1 不良驾驶行为指标选取

一级指标	二级指标	评价方式
行驶车速	平均车速	日平均车速过大
	超速次数	车辆超速频率过大
疲劳驾驶行为	连续驾驶时长	车辆日连续驾驶时长过大
	疲劳驾驶次数	多次违反疲劳驾驶法规要求
急变速行为	急加速次数	车辆加速过快
	急减速次数	车辆减速过快

4. 基于聚类的驾驶行为倾向性分析

本文采用了一种改进的 K-means 聚类算法——K-means++ 聚类算法进行驾驶行为的倾向性分析，并使用 PCA 降维进行聚类结果的可视化，提出语义将聚类结果分为稳健型、一般型和激进型三个类别。

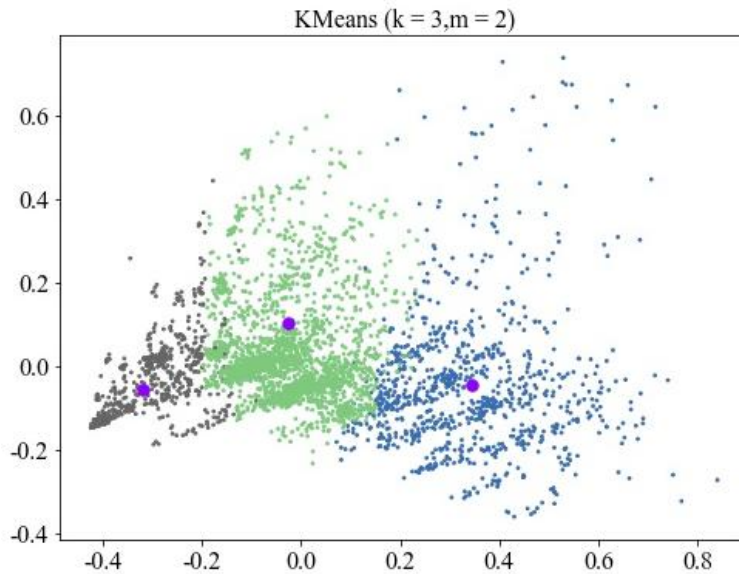


图 4-2 PCA (m=2)降维聚类效果图

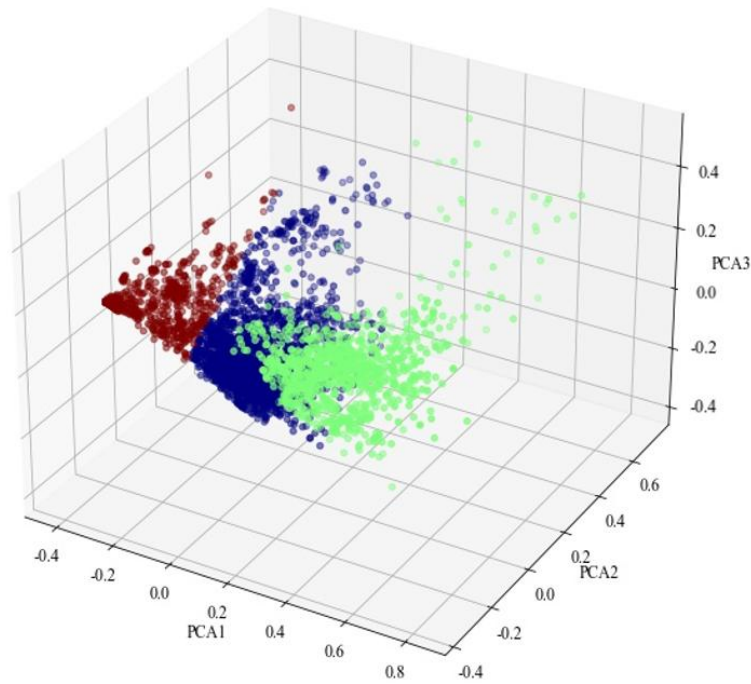


图 4-3 PCA (m=3)降维聚类效果图

5. 驾驶行为评分模型构建

5.1 驾驶行为评分模型权重确定

本文在研究主观赋权法和客观赋权法的基础上，采用了组合赋权法来确定指标权重。使用熵值法计算权重 a_j 。根据层次分析法得到权重 e_j 。最终得出组合赋权法的权重，具体权重见表 5-1。

表 5-1 3 种方法的各层驾驶行为指标权重对比

准则层	指标层	熵权法权重		层次分析法权重		组合赋权法权重	
行驶车速 θ_1	平均车速 θ_{11}	0.313	0.313	0.057	0.286	0.185	0.299
	超速次数 θ_{12}	0.001		0.229		0.115	
疲劳驾驶行为 θ_2	连续驾驶时长 θ_{21}	0.585	0.682	0.107	0.143	0.346	0.413
	疲劳驾驶次数 θ_{22}	0.097		0.0036		0.066	
急变速行为 θ_3	急加速次数 θ_{31}	0.002	0.005	0.381	0.571	0.192	0.288
	急减速次数 θ_{32}	0.002		0.190		0.096	

5.2 驾驶行为综合评分模型

根据 5.1 节中组合赋权法确定的权重结果，本文以 100 分为满分对标准化后的不良驾驶行为特征进行评分，提出评分准则如下：

$$score = \sum_{j=1}^n (1 - x'_{ij}) \times w_j \times 100 \quad (5-1)$$

其中 x_{ij}' 指第 3 章中不良驾驶行为特征聚类分析中标准化后的数据值， w_j 指的是 5.1 节使用组合赋权法计算得出的权重。

5.3 综合评分实例分析

图 5-1 为第 3 章中对驾驶行为的倾向性判别结果。其中稳健型数据有 732 条，一般型数据有 2373 条，激进型数据为 897 条。图 5-2、图 5-3 和图 5-4 为使用评分模型分别对三种类型的驾驶行为进行评分，根据评分结果绘制的直方图。横坐标为评分数值，纵坐标为对应评分分值区间的数量。可以看出激进型的驾驶行为评分结果多分布于 69 分以下，一般型的驾驶行为评分结果多分布于 69-80 分之间，稳健型的驾驶行为评分结果多分布于 80-99 分之间。由此，可以得出结论，本评分模型具有一定的有效性。

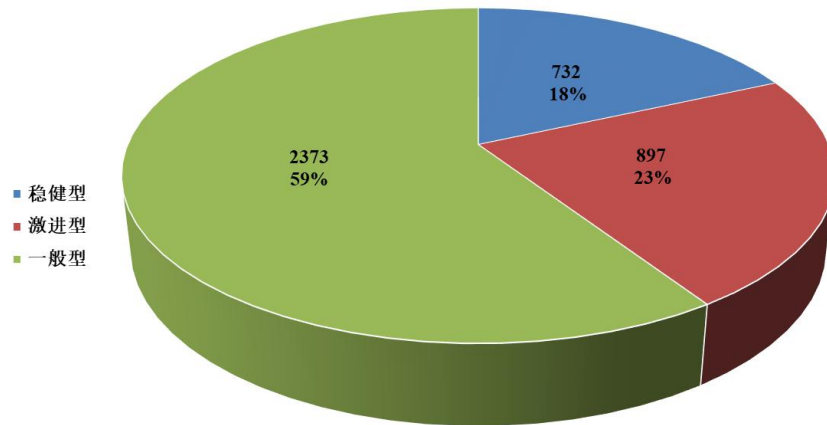


图 5-1 驾驶行为倾向性判别结果

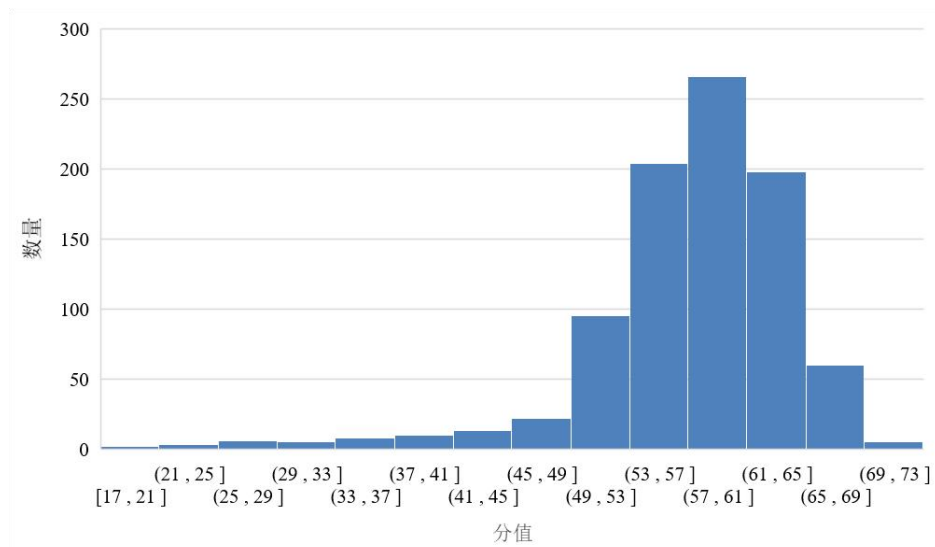


图 5-2 激进型驾驶行为评分直方图

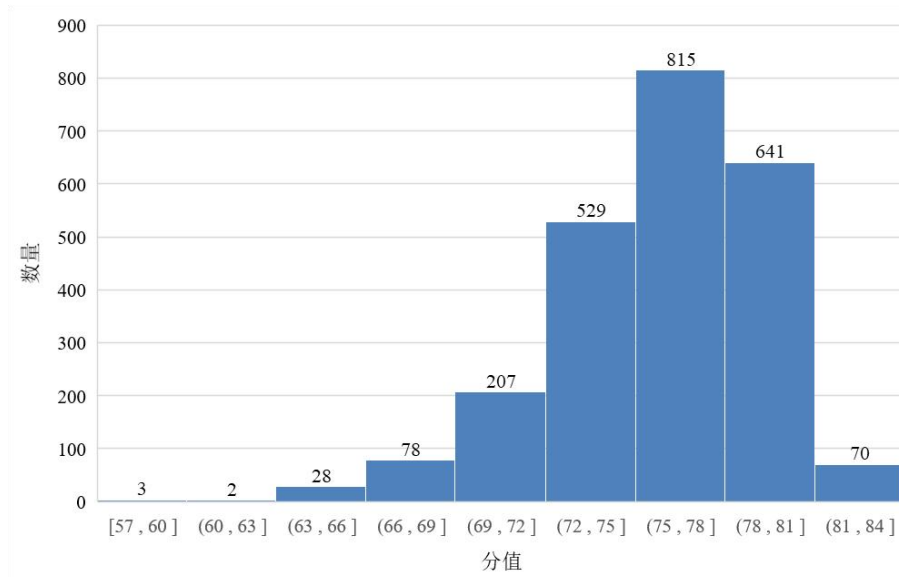


图 5-3 一般型驾驶行为评分直方图

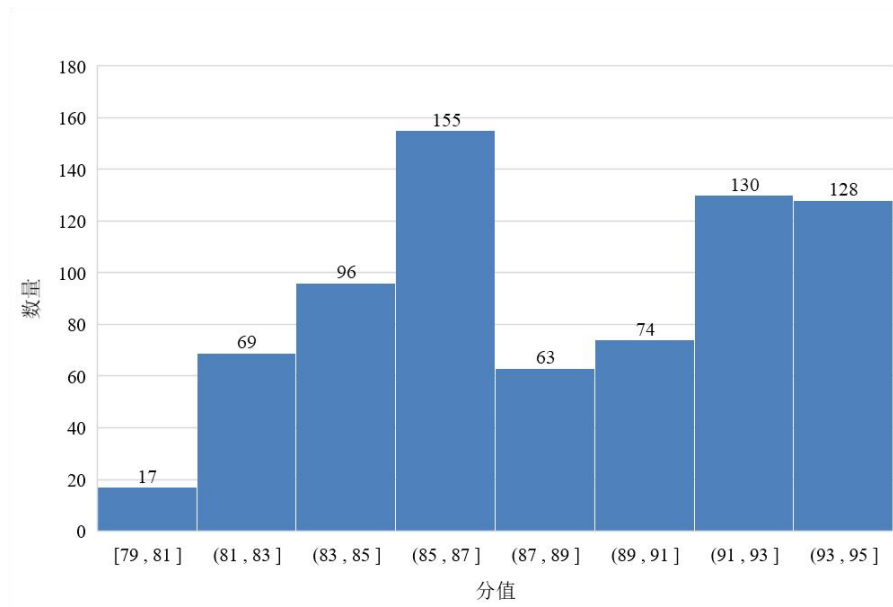


图 5-4 稳健型驾驶行为评分直方图

6. 总结

本文对车辆行驶轨迹进行可视化还原和车速热力图的绘制，展示了行车路线、行车速度等驾驶信息。通过对驾驶行为评价指标的研究，本文将不良驾驶行为归纳为三个维度：行驶车速、疲劳驾驶、急变速行为。并设计了不良驾驶行为识别算法，能够检测超速、疲劳驾驶、急加速、急减速等不良驾驶行为的指标。利用 K-means++ 算法对驾驶行为数据进行聚类分析，以此来判断驾驶者的行为倾向性，并将其归类为稳健型、一般型和激进型。使用熵值法和 AHP 层次分析法组合的组合赋权法确定各级指标权重，构建驾

驶行为评分模型。使用评分模型对实例进行打分，并结合聚类结果，验证了评分模型的合理性。

主要参考文献

- [1] 中华人民共和国国家统计局. 国家数据-年度数据[EB/OL]. [2022]. 中华人民共和国国家统计局, <https://data.stats.gov.cn/search.htm?s=交通事故>.
- [2] 陈辉. 面向车联网大数据的驾驶行为分析[D].贵州大学,2020.
- [3] 张佳薇,郑岳涵,李明宝.组合赋权-模糊综合评价的安全驾驶行为分析[J].科学技术与工程,2022,22(23):10255-10261.
- [4] 杨光. 基于车辆行驶数据的驾驶员行为模型构建与应用[D].西安石油大学,2021.
- [5] 彭江琴. 基于 GID 的车联网保险 UBI 费率与驾驶行为评分研究[D].南京邮电大学,2016.
- [6] Wakasugi T. A study on warning timing for lane change decision aid systems based on driver's lane change maneuver[C]. International Technical Conference on the Enhanced Safety of Vehicles, 2005.
- [7] Ayuso M, Guillen M, Perez-Marin A M. Using gps data to analyse the distance travelled to the first accident at fault in pay-as-you-drive insurance[J]. Transportation Research Part C, 2016, 68:160-167.
- [8] Jun J, Ogle J, Guensler R. Relationships between crash involvement and temporal-spatial driving behavior activity patterns: use of data for vehicles with global positioning systems[J]. Transportation Research Record, 2019, 246-255.
- [9] 龚雪蕾,方菁,肖传浩,谭晓萍,廖爱梅.实现可持续发展目标（SDGs）中降低道路交通事故目标预测——中国、日本和柬埔寨比较研究[J].伤害医学(电子版),2019,8(04):14-21.
- [10] 唐涵润. 基于行车数据的驾驶人不良行为综合评价与仿真干预研究[D].东南大学,2021.
- [11] 张志鸿. 基于 OBD 数据分析的驾驶行为研究[D].长安大学,2017.
- [12] 孙一帆,张敬磊,王战古,于杰,王晓原,孔栋.基于模拟驾驶实验的不同驾驶倾向驾驶员的酒后驾驶行为[J].科学技术与工程,2018,18(07):246-251.
- [13] 程兴乾. 基于车联网的用户驾驶行为统计分析研究[D]. 北京工业大学, 2019.
- [14] R.Verma, B. Mitra and S. Chakraborty, "Avoiding Stress Driving: Online Trip Recommendation from Driving Behavior Prediction" 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom)[C], Kyoto, Japan, 2019, 1-10.
- [15] 王博文,王景升,朱茵,闫硕,薛曦初.基于 K-means++的驾驶人致因因素倾向性分析[J].科学技术与工

程,2021,21(36):15695-15699.

- [16] 刘应吉,曾诚,王书举,姚羽,刘梦雅.基于卫星定位数据的驾驶行为安全与节能评价方法[J].公路交通科技,2018,35(01):121-128+158.
- [17] 张龙雨.商用车驾驶员驾驶行为评价算法研究[D].重庆理工大学,2022.
- [18] 姚知涵.基于GPS轨迹数据的车辆分类方法研究[D].杭州电子科技大学,2022.
- [19] 姚文钦.基于CAN数据的渣土车驾驶行为分析系统研究与建设[D].贵州大学,2022.
- [20] 吴纪铎,刘爱松,赵梦海.T-BOX车载网联终端设计[J].汽车零部件,2022(02):11-16.
- [21] 谢卉瑜,边旭东,张亦弛.车载终端T-BOX技术分析及发展趋势研究[J].时代汽车,2022(01):162-163.
- [22] 许晓慧.驾驶愤怒行为特征分析及其干预措施的研究[D].东南大学,2018.
- [23] 张正.基于车联网驾驶行为评分的安全驾驶卫士分析与设计[D].南京邮电大学,2017.
- [24] 王子惠.基于UBI驾驶行为评分的车险定价研究[D].湖南大学,2019.