

基于 CFSFDP-BP 的车辆出行群体 分类及识别方法

谢 坤¹, 董宏辉¹, 卢玲玉², 耿庆桥³, 李鹏辉¹, 董春娇¹

(1. 北京交通大学 综合交通运输大数据应用技术交通运输行业重点实验室, 北京 100044; 2. 中国城市规划设计研究院深圳分院 市政所, 广东 深圳 518040; 3. 交通运输部规划研究院 公路所, 北京 100028)

摘 要:以货车和私家车出行轨迹数据为基础,以出行开始时间、出行结束时间、总运行时长等 7 个出行指标作为特征量,提出了基于密度峰值聚类的(CFSFDP)出行群体分类方法,将货车和私家车分别划分为三类出行群体。以出行特征指标为输入,密度聚类算法的分类标签为输出,建立基于 BP 网络的车辆出行群体类别识别模型,实现对同一车型下的不同出行群体的快速识别。研究结果表明:提出的基于 CFSFDP-BP 的车辆出行群体分类及识别方法预测精度和可靠性较好,对货车出行群体类别的识别准确率为 0.991,对私家车出行群体类别识别准确率为 0.988,且对货车出行群体的识别效果优于私家车的识别效果。货车出行群体可划分为灵活型-低强度、传统型-中强度以及传统型-高强度三类;私家车出行群体可划分为通勤型-低强度、通勤型-中强度以及灵活型-高强度三类出行特征群体。研究结果能够为交管部门制定精细化的交通管理策略,提高交通运行效率提供支持。

关键词:道路交通;货车与私家车;出行特征;密度峰值聚类算法;出行群体识别

中图分类号:U491.1 **文献标志码:**A **文章编号:**1671-5497(2026)03-0725-09

DOI:10.13229/j.cnki.jdxbgxb.20240887

Classification and recognition method of vehicle travel groups based on CFSFDP-BP

XIE Kun¹, DONG Hong-hui¹, LU Ling-yu², GENG Qing-qiao³, LI Peng-hui¹, DONG Chun-jiao¹

(1. Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China; 2. Municipal Engineering Department, China Academy of Urban Planning & Design Shenzhen Branch, Shenzhen 518040, China; 3. Institute of Highways, Planning and Research Institute of the Ministry of Transport, Beijing 100028, China)

Abstract: Based on truck and private car trajectory data, a trip group classification method using Clustering by fast search and find of density peaks(CFSFDP) was proposed, in which seven trip indicators—including trip start time, trip end time, and total operation duration—were employed as feature variables. The trucks and private cars were respectively divided into three trip groups. A BP neural network-based

收稿日期:2024-08-08.

基金项目:国家重点研发计划项目(2023YFC3009601);2023年度国家社科基金重大项目(23&ZD138).

作者简介:谢坤(1982-),男,博士研究生.研究方向:交通安全,低碳交通. E-mail: xiekun@bjtu.edu.cn

通信作者:李鹏辉(1991-),男,副教授,博士.研究方向:交通安全,智能交通. E-mail: 9323@bjtu.edu.cn

vehicle trip group identification model was subsequently established, with trip characteristic indicators serving as inputs and the density clustering algorithm's classification labels as outputs, thereby enabling rapid identification of different trip groups within the same vehicle type. The results indicated that the proposed CFSFDP-BP-based trip group classification and identification method demonstrated satisfactory prediction accuracy and reliability. An identification accuracy of 0.991 was achieved for truck trip groups, while 0.988 was obtained for private car trip groups, with superior performance observed for truck trip group identification compared to private cars. Three distinct truck trip groups were identified: flexible-low intensity, traditional-medium intensity, and traditional-high intensity. Meanwhile, three private car trip groups were classified as commute-low intensity, commute-medium intensity, and flexible-high intensity. These findings are expected to support traffic management authorities in formulating refined traffic management strategies and improving traffic operational efficiency.

Key words: road traffic; trucks and private cars; travel characteristics; clustering by fast search and find of density peaks; identification of travel groups

0 引言

在车辆出行特征研究方面,已有研究多基于车牌数据、轨迹数据和卡口数据等,针对货车或者私家车中的一种,提取出行次数、出行时长和出行距离等一系列出行特征指标,利用机器学习方法从中挖掘车辆出行特征,缺乏针对不同群体出行特征的刻画方法。分类定量刻画货车和私家车出行特征,能够为城市交管部门制定精细化的交通管理策略,提高交通运行效率提供支持。

基于货车行驶轨迹数据构建出行链,分析其出行特征,然后基于聚类方法对出行特征群体进行划分是当前的研究热点^[1-4]。张文畅^[5]基于货车轨迹数据,通过提取违章次数、货运距离、超速次数、运输趟次、夜间驾驶时间等指标,采用 SOM-Kmeans 两阶段聚类方法进行聚类,根据聚类结果将货车群体划分五大类。薛倩等^[6]采用青岛市货车 GPS 数据,结合速度、时间和距离阈值确定停留点,对出行分布特征进行统计挖掘。王静^[7]基于货车 GPS 轨迹数据,对货车的货运业务点及主要基地仓库进行识别,构建车辆的货运出行链并识别计算对应链的相关特征,最后采用聚类方法基于出行链特征将货车出行群体分为了 3 类。Gan 等^[8]利用货车 GPS 轨迹数据,从中提取 OD 数据,挖掘货车的出行模式。

对于私家车出行特征挖掘,大多数研究基于轨迹数据和卡口数据等,提取相关特征指标,识别通勤车辆并进行出行特征的分析。魏广奇等^[9]基于高速公路收费流水数据,利用 k -means++ 聚类方法识别高速公路日常通勤车辆,从出行时间、

出行空间、车牌构成等方面分析通勤车辆的出行特征。杨帅等^[10]基于车辆单次出行轨迹信息,提出一种基于 Apriori 算法的通勤车辆识别方法,并采用 k -means 聚类算法进行对比分析,通过设置 7 种车辆出行特征指标,分析路网车辆构成。Zhou 等^[11]基于 GeoLife 数据集,将轨迹点数据序列转换为位置的时间序列,构建行为模型挖掘个人出行模式。Zhao 等^[12]基于原始智能卡交易数据,提取乘客出行特征指标,并将统计分析方法和无监督聚类方法结合,对乘客的出行模式进行挖掘。Zhang 等^[13]基于出租车轨迹数据,采用多元空间点模式分析空间相关性,识别出兴趣点从而挖掘兴趣点周围出租车出行模式,最后利用核密度估计方法估计了出租车的服务范围。

也有部分学者通过对多源交通大数据进行挖掘,分析车辆的出行特征,利用聚类方法将其划分为不同出行群体并且构建识别模型,从而对划分的出行群体进行识别。汪涛^[14]基于轨迹熵分析车辆的行为特征,并利用 K -means 聚类算法,将车辆划分为私家车、出租车、公交车以及其他车辆。黄正国^[15]基于车牌识别数据,通过提取车辆出行总天数、车辆使用强度以及车辆在点序列上重复出现的累积次数等指标,采用 K -means 聚类算法分别对公交车、出租车以及通勤车的出行属性进行了挖掘。梁泉等^[16]基于北京市公共交通刷卡和线站数据,对公共交通多源数据进行关联匹配并提取出行链,采用多层规划理论构建了个体出行知识图谱,提取了出行天数、出行空间均衡度等 7 类特征指标,并且构建了面向公共交通乘

客分类的BP神经网络模型用以准确识别不同类别乘客。马晓磊等^[17]基于北京市私家车的OBD数据,通过提取平均出行距离、出行频次等出行指标,并且利用密度峰值聚类算法进行聚类,依据聚类分析结果将驾驶人分为4类,最后构建隐马尔可夫模型对划分群体进行识别。

本文以货车和私家车出行轨迹数据为基础,以出行开始时间、结束时间、总运行时长等七个指标作为特征量,建立基于密度峰值聚类的(CFSFDP)车辆出行群体分类方法,提出基于BP神经网络的车辆出行群体识别模型,将货车和私家车出行群体进行分类,并研究各类出行群体的出行特征。

1 基于密度峰值聚类的车辆出行群体分类方法

为了更加精细地掌握车辆出行特性,结合出行指标数据集,分别对货车和私家车的出行群体进行划分,进而挖掘不同出行群体的出行特征。由于同一车型的出行指标数据集没有标签,表现出来的只有特征数据,故采用聚类算法对其进行分类。目前,聚类算法主要有以下几类:基于划分的聚类(如K-means聚类),基于密度的聚类(如DBSCAN聚类),基于层次的聚类(如综合层次聚类),基于网格的聚类,基于图论的聚类以及基于深度学习的方法等^[15]。

Rodriguez等^[18]于2014年在《Science》上提出了基于密度峰值的聚类算法(Clustering by fast search and find of density peaks),即CFSFDP聚类算法。该算法可以一次性找出聚类中心,且对数据形状的适应性较高。与其他基于密度的聚类算法相比,该算法需要输入的指标更少,不需要迭代,运行效率更高。因此,本文采用CFSFDP聚类算法,用于货车与私家车的出行群体分类。

CFSFDP聚类算法的核心参数包括局部密度 ρ_i 与相对距离 δ_i 。局部密度 ρ_i 可以通过截止内核算法和高斯算法获得。截止内核算法将数据点 x_i 看作中心点,将截断距离 d_c 视为半径,统计其邻域内包含数据点的个数。计算公式如式(1)和(2)所示:

$$\rho_i = \sum_{j \in I_s, j \neq i} \chi(d_{ij} - d_c) \quad (1)$$

$$\chi(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \quad (2)$$

式中: d_{ij} 为 x_i 与 x_j 之间的欧式距离; d_c 为截断距离,需在聚类前研究设定。

高斯算法计算的是所有数据对象到该点的高斯距离之和,如式(3)所示:

$$\rho_i = \sum_{j \in I_s, j \neq i} \exp\left(-\left(\frac{d_{ij}}{d_c}\right)^2\right) \quad (3)$$

相对距离 δ_i 计算公式如式(4)所示。由式(4)可知,若 x_i 的局部密度全局最大, δ_i 为 x_i 数据点与出行特征指标数据集中距离最大的数据点的距离值;否则, δ_i 为在所有局部密度大于 x_i 的数据点中,与 x_i 距离最小的数据点的距离值。

$$\delta_{q_i} = \begin{cases} \min_{q_j, j < i} \{d_{q_i, q_j}\}, & i \geq 2 \\ \min_{j \geq 2} \{d_{q_i, q_j}\}, & i = 1 \end{cases} \quad (4)$$

式中: $\{q_i\}_{i=1}^N$ 表示 $\{\rho_i\}_{i=1}^N$ 的一个降序排列的下标集合,如式(5)所示。

$$\rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N} \quad (5)$$

对所有数据点的局部密度 ρ_i 和距离 δ_i 计算完成后,以 ρ 为横轴, δ 为纵轴,将二元点集 $\{(\rho_i, \delta_i)\}_{i=1}^N$ 对应做出聚类决策图,选取聚类中心。若无法通过决策图一次性选出聚类中心时,根据决策值 γ_i 进行选择。决策值 γ_i 为局部密度 ρ_i 和距离 δ_i 的乘积,如式(6)所示,选取原则为将决策值较大的点视为聚类中心。

$$\gamma_i = \rho_i \delta_i \quad (6)$$

为了对车辆出行群体进行分类,首先计算车辆出行数据集中每个数据点的局部密度 ρ_i 与相对距离 δ_i ;然后,以局部密度 ρ_i 为横坐标轴,以相对距离 δ_i 为纵坐标轴绘制聚类决策图,根据绘制的聚类决策图,选取局部密度 ρ_i 与相对距离 δ_i 均较大的点作为初始聚类中心;最后,对车辆出行数据集中剩余非聚类中心点进行聚类处理。主要步骤如下:

(1)输入数据集 $S = \{x_1, x_2, \dots, x_N\}$

将出行指标数据集作为算法的输入参数,其中包含出行距离、总运行时长、出行次数、平均停留时间、出行开始时段、出行结束时段以及平均行程时间等特征。

(2)计算输入指标集的距离矩阵

采用欧式距离,货车与私家车距离矩阵分别计算,如式(7)所示:

$$d_{ij} = \sqrt{\sum_{l=1}^m (x_{il} - x_{jl})^2} \quad (7)$$

式中: m 为数据的维度; x_{il} 是数据点 x_i 在第 l 个维度下的取值。

(3) 计算截断距离 d_c

截断距离 d_c 的计算原则为: 以 d_c 为半径画圆, 要求落在 d_c 圆区域内的平均点数占总点数的 1%~2%。

(4) 绘制决策图并选取初始聚类中心

采用高斯方法计算局部密度, 采用式(4)计算相对距离; 根据计算结果绘制决策图。通过决策图来选取聚类中心, 仅能选出一个。因此, 利用式(6)计算决策值, 选取初始聚类中心。

(5) 输出聚类结果

将数据集中其余非聚类中心点进行聚类, 并将最终聚类结果输出 $C = \{c_1, c_2, \dots, c_m\}$ 。

算法对车辆出行群体进行分类的基本原理有以下两点: 一是聚类中心的局部密度值比较大; 二是聚类中心比其局部密度值更大的点较远。

2 基于神经网络的车辆出行群体类别识别模型

利用密度聚类算法对货车与私家车的出行指标分别聚类以后, 可以得到同一车型不同出行群体, 从而掌握其出行特点。聚类算法虽然能快速地对车辆出行群体进行划分, 但是由于其运行数据样本量少、且鲁棒性较低等局限性, 在交通大数据环境下, 使用该方法对车辆的出行群体进行识别效率较低。因此, 有必要以出行特征指标为输入, 以密度聚类算法的分类标签为输出, 建立车辆出行群体类别识别模型, 实现对同一车型下的不同出行群体的快速识别。

由于特征指标与车辆出行群体特征标签之间具有非线性关系, 因此选择具有很强的非线性映射能力, 且能够实现误差反向传播的 BP 网络进行车辆出行群体的识别。BP 网络拓扑结构由输入层、隐含层以及输出层组成, 各层神经元仅与相邻神经元之间相互连接, 从而组成了具有层次结构的前馈型神经网络系统。

当出行特征指标信息正向传播时, BP 网络将出行特征指标信息通过隐含层传送到输出节点, 经过非线性变换产生车辆出行群体类别标签。如果输出与密度聚类算法的分类标签相差很大, 则转入误差反向传播过程。当误差反向传播时, BP 网络通过隐含层将输出误差逐层反向传播到输入

层, 同时将误差值分配给各层的神经单元, 以获得调整各单元权重的依据。然后, BP 网络通过调整输入层和间接层节点之间的连接强度以及间接层节点和输出节点之间的连接强度和阈值, 沿梯度减少误差。在重新学习和训练后, 确定与最小误差相对应的网络参数并停止训练。

基于 BP 神经网络的车辆出行群体类别识别模型计算流程如下:

(1) 初始化网络权重

将 BP 网络的隐含层设置为 m 层, 隐含层的神经元数量设置为 n , 神经元之间的网络连接权重初始化为很小的随机数。

(2) 向前传播输入

根据划分的训练集提供神经网络的输入层, 利用公式(8)计算得到每个神经元的输出。

$$O_j = \frac{1}{1 + \exp\left[-\left(\sum_i \omega_{ij} O_i + \theta_j\right)\right]} \quad (8)$$

式中: ω_{ij} 为上一层神经单元 i 到本神经单元 j 的网络权重; O_i 为上一层神经单元 i 的输出; θ_j 为本单元的偏置。 O_j 称为激活函数, 引入激活函数是为了增加 BP 网络模型的非线性能力。

激活函数是 BP 网络运行的重要参数之一, 常用的激活函数主要有 identity、tanh、logistic、relu 激活函数四类。identity 适合于潜在行为是线性(与线性回归相似)的任务; logistic 函数用于分类器时效果更好; tanh 函数解决了 logistic 中心不为 0 的缺点, 但存在梯度易消失的缺点; relu 函数是一个通用的激活函数, 是针对 logistic 函数和 tanh 的缺点进行改进的, 因此本文在 BP 网络运行过程中使用 relu 激活函数。

(3) 反向传播误差

经过步骤(2), BP 网络模型最终在输出层得到实际的车辆出行群体类别标签, 根据公式(9)与密度聚类算法的分类标签相比较得到每个输出单元 j 的误差。

$$E_j = O_j(1 - O_j)(T_j - O_j) \quad (9)$$

式中: T_j 为输出单元 j 的预期输出。

将计算出的误差从后向前逐层进行传播, 误差计算公式如式(10)所示:

$$E_j = O_j(1 - O_j) \sum_k \omega_{jk} E_k \quad (10)$$

(4) 调整网络权重与神经元偏置

误差计算完成后, 根据误差计算结果调整神

经网络的权重和神经元的阈值。调整公式如式(11)所示:

$$\omega_{ij} = \omega_{ij} + (l) O_i E_j \quad (11)$$

利用式(12)对神经元偏置进行调整。

$$\theta_j = \theta_j + (l) E_j \quad (12)$$

式中: l 为学习率,通常取0~1之间的常数。本文将迭代次数设置为1 000次,学习率为0.1。

(5)网络运行结束

对于每个输入样本来说,若网络最终输出误差小于可接受的范围,或者迭代次数 t 达到了BP网络运行前提前设定的阈值,则对下一个样本进行训练;如超过预期可接受范围,则将迭代次数增加1次,转入步骤(2),继续训练当前样本。

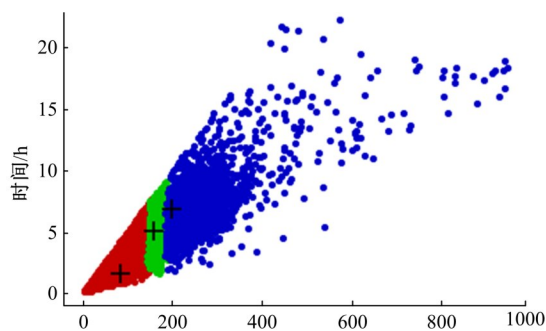
3 实验验证

3.1 车辆出行群体分类及识别

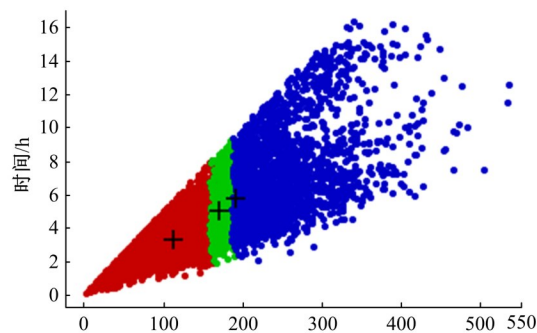
本文所用数据为两种车型的数据,一种为北京市2019年12月23日的货车数据,包括7 795辆货车的轨迹段。另一种为北京市2019年9月23日的私家车GPS数据,包括7 794辆私家车的轨迹段。将出行指标,包含出行距离、总运行时长、出行次数、平均停留时间、出行开始时段、出行结束时段以及平均行程时间7个字段的数据集作为输入,采用CFSFDP密度峰值聚类算法,对货车与私家车分别进行聚类。

结合聚类决策图以及公式(6)计算结果,最终货车和私家车出行指标数据集分别选取了3个聚类中心,即将货车与私家车分别聚成了3类出行群体,结果如图1(a)和(b)所示。聚类结果中车辆占比如图1(c)和(d)所示。货车聚类结果中类别1、2、3的占比分别为64.62%(5 037辆)、13.61%(1 061辆)和21.77%(1 697辆),即类别1的货车最多,类别2的货车最少。私家车聚类结果中类别1、2、3占比分别为60.87%(4 744辆)、10.69%(833辆)和28.44%(2 217辆),类别1的私家车最多,类别2的私家车最少。

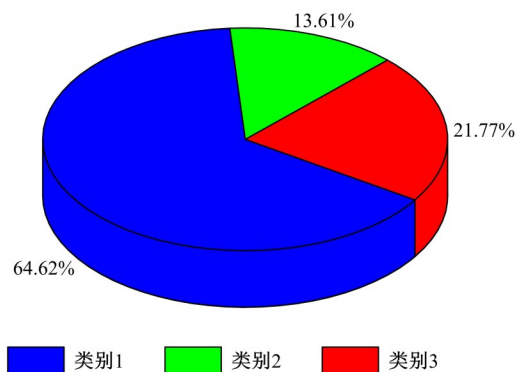
将聚类结果的70%作为训练集,30%作为测试集,建立基于BP网络的车辆出行群体识别模型,并根据混淆矩阵以及精确率、召回率、 F_1 等指标对模型进行评价,结果由表1所示。由混淆矩阵可知,货车出行群体中,类别1、2和3分别有1 516辆、319辆和483辆被准确识别。私家车出行群体中,类别1、2和3分别有1 411辆、241辆和



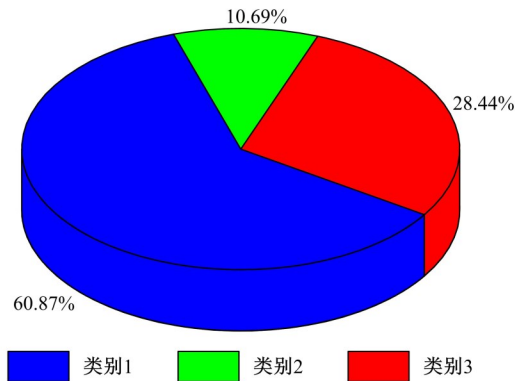
(a)货车聚类结果



(b)私家车聚类结果



(c)货车占比



(d)私家车占比

图1 聚类结果及车辆占比

Fig. 1 Proportion of vehicles in clustering results

658辆被准确识别。

由表1可知,货车出行群体识别准确率为0.991,召回率为0.991,精确率为0.991, F_1 值为0.991。

表 1 货车群体识别模型评估结果
Table 1 Evaluation results of truck group identification model

车型	数据集	准确率	召回率	精确率	F_1
货车	训练集	0.989	0.989	0.989	0.989
	测试集	0.991	0.991	0.991	0.991
私家车	训练集	0.990	0.990	0.990	0.990
	测试集	0.988	0.988	0.988	0.988

私家车出行群体识别准确率为 0.988,召回率为 0.988,精确率为 0.988, F_1 值为 0.988。识别结果表明,本文提出的方法预测精度和可靠性较好,且对货车出行群体的识别效果优于私家车的识别效果。

3.2 车辆群体出行特征分析

对识别出的货车与私家车 3 类出行群体分别从出发/结束时段、总运行时长、平均停留时间、平均行程时间、出行次数和出行距离等方面进行出行特征分析。由图 2 可知,货车 3 种类别出行开

始时段均集中分布在 0:00~1:00,出行结束时段均集中分布在 22:00~23:00。其中,类别 1 的出行开始时段在 7:00~8:00 时间段存在一个小高峰,出行结束时段在 16:00~17:00 时间段存在高峰现象。与类别 2 和 3 相比较,类别 1 出行群体的出行时间选择偏好更加灵活,类别 2 和类别 3 的出行时间选择更为固定。

私家车 3 种类别出行开始时段主要集中在上午,出行结束时段则具有差异性。其中,类别 1 的出行结束时段主要集中分布在晚高峰即 17:00~18:00 时间段,即该类别的出行主体主要为通勤出行。类别 2 的出行结束时段主要集中在 18:00~19:00 以及 22:00~23:00,并且两个时间段的出行量分布较为均衡。类别 3 的出行结束时段分布存在两个高峰,分别是 17:00~18:00 以及 22:00~23:00,大部分出行分布在 22:00~23:00 时间段,由此可以看出,类别 3 的出行群体更倾向于夜间回程。

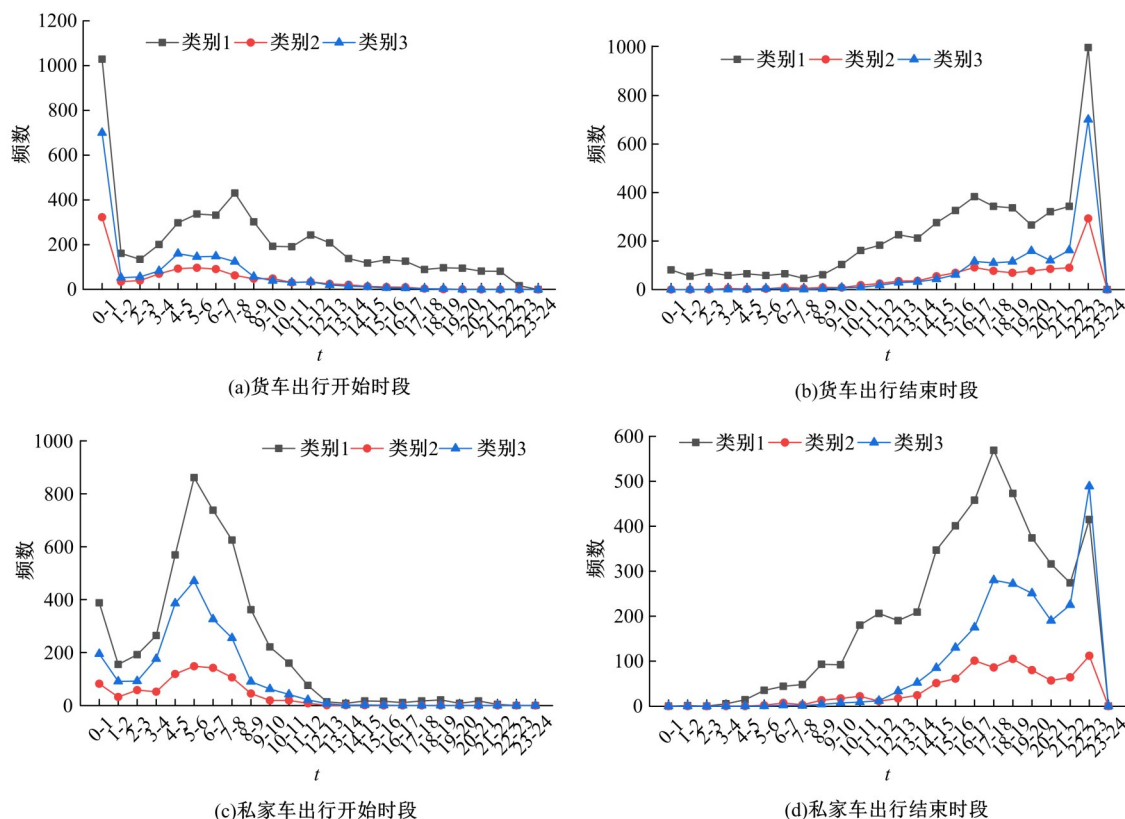


图 2 出行开始与结束时段分布特征

Fig. 2 Distribution characteristics of travel start and end time periods

通过绘制箱线图刻画货车与私家车 3 种类别群体的出行指标差异性,如图 3 所示,彩图参见电子版。黛青色箱体代表数据的 25%~75% 分布范围,位于箱体里的黑线代表数据的中位数,空心

矩形代表数据的均值,实心菱形代表数据的异常值。由图 3 可知,货车出行的 3 种类别,总运行时长范围分别为 1.8~3.5 h、3.9~6 h 和 5.8~8.5 h;平均停留时长范围分别为 1.5~3.9 h、1.7~3.5 h

和1.5~2.2 h。私家车的三种类别,总运行时长范围分别为2~4 h、3.9~5.9 h和5.4~8.4 h;平均停留时长范围分别为1.3~2.3 h、1.2~1.8 h

和0.8~1.4 h。货车与私家车具有相似的出行规律特征,即第3类群体出行总运行时长最长,但是平均停留时间最短。

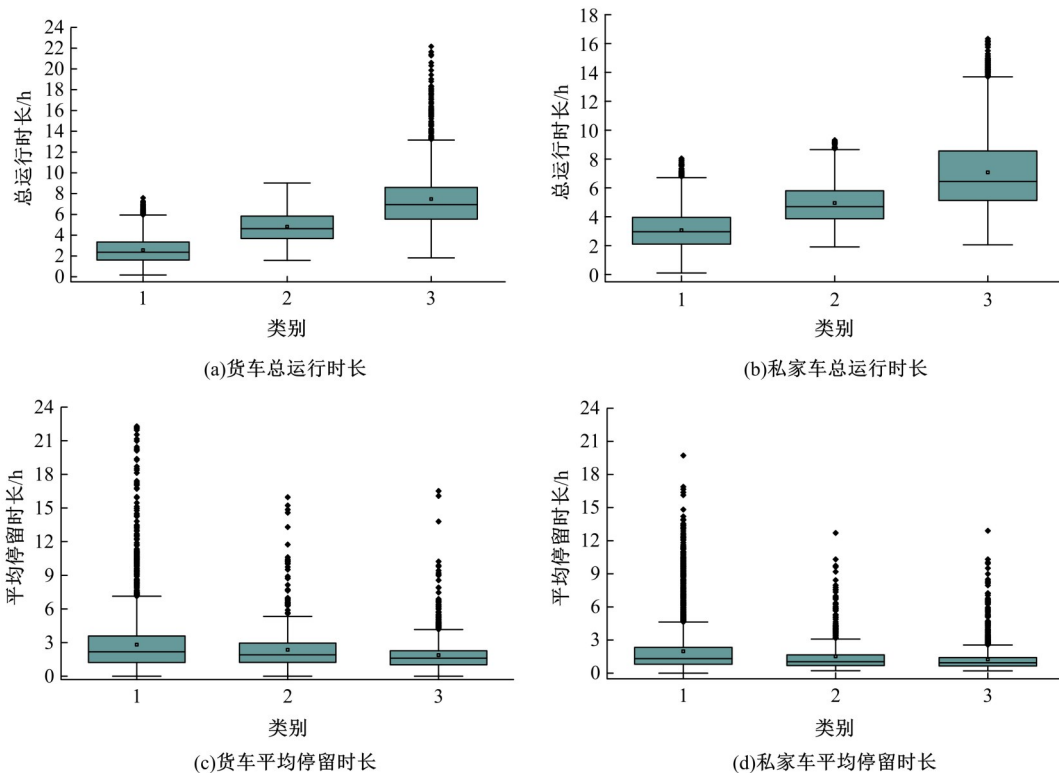


图3 总运行时长与平均停留时间

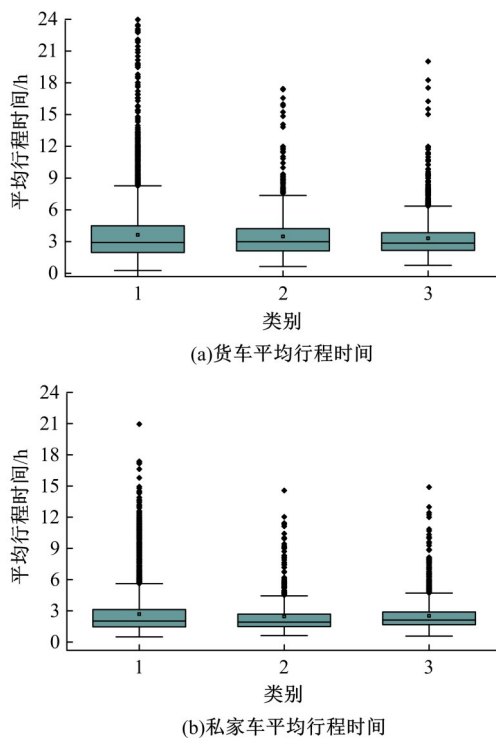
Fig. 3 Total travel time and average stay duration

由图4可知,货车3种类别平均行程时间分别为3.64 h、3.49 h和3.3 h;出行次数分别为3.84次、5.49次和6.58次。私家车3种类别平均行程时间分别为2.69 h、2.45 h和2.51 h;出行次数分别为4.82次、6.01次和6.33次。即,类别1的平均行程时间最长,平均出行次数最少。

货车三种类别的平均出行距离分为86.81、167.56和269.38 km。私家车三种类别的平均出行距离分别为98.92、175.42和250.76 km。即,类别1出行距离最短,类别3出行距离最长。

综上所述,货车类别1的总运行时长、出行次数以及出行距离最小,出行时间分布较为灵活;类别2的总运行时长、出行次数以及出行距离适中,出行时间分布比较固定;类别3的总运行时长、出行次数以及出行距离最高,出行时间分布比较固定。因此,将类别1、2、3分别定义为灵活型-低强度运输群体,传统型-中强度运输群体,以及传统型-高强度运输群体。私家车类别1的总运行时长、出行次数以及出行距离最小,出行时间分布具有明显的通勤特性;类别2的总运行时长、出行次

数以及出行距离适中,出行时间分布具有明显的通勤特性;类别3的总运行时长、出行次数以及出



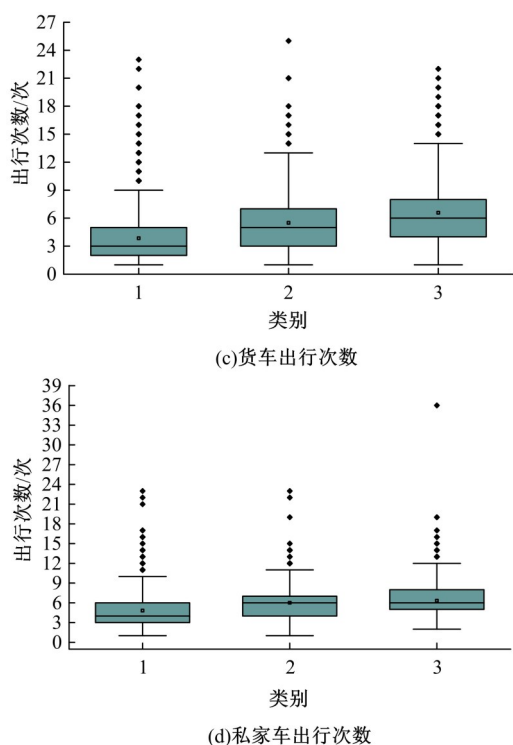


图 4 平均出行时间与出行次数

Fig. 4 Average travel time and number of private car trips

行距离最高,具有明显的夜间回程偏好。

因此,将类别 1、2、3 分别定义为通勤型-低强度出行群体,通勤型-中强度出行群体,以及灵活型-高强度出行群体。

4 结 论

为了分类定量刻画货车和私家车出行特征,本文以货车和私家车出行轨迹数据为基础,结合出行开始时间、结束时间、总运行时长等 7 个出行特征指标,建立了基于密度峰值聚类的(CFSFDP)车辆出行群体分类方法。针对货车和私家车分别确定了 3 个初试聚类中心,将货车与私家车划分为 3 类不同的出行群体。以密度聚类算法的分类标签为目标,建立了基于 BP 网络的车辆出行群体类别识别模型,实现了对同一车型下的不同出行群体的快速识别。研究结果表明:提出的基于 CFSFDP-BP 的车辆出行群体分类及识别方法预测精度和可靠性较好,对货车出行群体类别的识别准确率为 0.991,对私家车出行群体类别识别准确率为 0.988,且对货车出行群体的识别效果优于私家车的识别效果。货车出行群体可划分为灵活型-低强度、传统型-中强度以及传统型-高强度 3 类;私家车出行群体可划分为

通勤型-低强度、通勤型-中强度以及灵活型-高强度 3 类出行特征群体。研究结果能够为城市交通管理部门制定精细化的交通管理策略、提高交通运行效率提供支持。

参考文献:

- [1] Zhou T, Zhang J. Analysis of commercial truck drivers' potentially dangerous driving behaviors based on 11-month digital tachograph data and multilevel modeling approach[J]. Accident Analysis & Prevention, 2019, 132: No. 105256.
- [2] Du J, Aultman-hall L. Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: automatic trip end identification issues[J]. Transportation Research Part A: Policy and Practice, 2007, 41(3): 220-232.
- [3] 周荣贵, 高沛, 李雨璇, 等. 基于轨迹数据的高速公路小客车异常驾驶行为阈值[J]. 吉林大学学报:工学版, 2024, 54(9): 2581-2587.
Zhou Rong-gui, Gao Pei, Li Yu-xuan, et al. Abnormal driving behavior thresholds of highway minibuses based on trajectory data[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(9): 2581-2587.
- [4] 张健钦, 仇培元, 杜明义. 基于时空轨迹数据的出行特征挖掘方法[J]. 交通运输系统工程与信息, 2014, 14(6): 72-78.
Zhang Jian-qin, Qiu Pei-yuan, Du Ming-yi. Mining method of travel characteristics based on spatio-temporal trajectory data[J]. Journal of Transportation System Engineering and Information Technology, 2014, 14(6): 72-78.
- [5] 张文畅. 基于轨迹数据的货车司机货运特征画像及应用研究[D]. 成都:西南交通大学交通运输与物流学院, 2021.
Zhang Wen-chang. Research on the characteristic portrait of truck drivers based on trajectory data[D]. Chengdu: School of Transportation and Logistics, Southwest Jiaotong University, 2021.
- [6] 薛倩, 李旭, 王建春. 基于 GPS 数据的青岛市货运车辆出行特征分析[J]. 中阿科技论坛, 2023(5): 96-100.
Xue Qian, Li Xu, Wang Jian-chun. Research on the travel characteristics of freight vehicles in Qingdao based on GPS data[J]. China-Arab States Science and Technology Forum, 2023(5): 96-100.
- [7] 王静. 基于 GPS 数据的货运车辆出行特性研究[D].

- 北京:北京交通大学交通运输学院, 2022.
- Wang Jing. Research on travel characteristics of freight vehicles based on GPS data[D]. Beijing: School of Traffic and Transportation, Beijing Jiaotong University, 2022.
- [8] Gan M, Nie Y M, Liu X, et al. Whereabouts of truckers: an empirical study of predictability[J]. Transportation Research Part C: Emerging Technologies, 2019, 104: 184-195.
- [9] 魏广奇, 苏跃江, 吴德馨, 等. 基于高速公路流水数据的通勤车辆特征研究[J]. 交通运输系统工程与信息, 2019, 19(3): 237-244.
- Wei Guang-qi, Su Yue-jiang, Wu De-xin, et al. Trip characteristics of vehicle with commuting property based on highway ticket data[J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(3): 237-244.
- [10] 杨帅, 刘樟伟, 杨宇星, 等. 基于高快速路卡口数据的车辆出行特征分析[C]//中国城市交通规划年会, 青岛, 2018:No. 13.
- Yang Shuai, Liu Zhang-wei, Yang Yu-xing, et al. Analysis of vehicle travel characteristics based on toll-gate data of highways and expressways[C]//Annual Conference of Urban Transport Planning in China, Qingdao, China, 2018:No. 13.
- [11] Zhou K, Tian Z, Yang Y. Periodic pattern detection algorithms for personal trajectory data based on spatio-temporal multi-granularity[J]. IEEE Access, 2019, 7: 99683-99693.
- [12] Zhao J, Qu Q, Zhang F, et al. Spatio-temporal analysis of passenger travel patterns in massive smart card data[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(11): 3135-3146.
- [13] Zhang S, Tang J, Wang H, et al. Revealing intra-urban travel patterns and service ranges from taxi trajectories[J]. Journal of Transport Geography, 2017, 61: 72-86.
- [14] 汪涛. 基于历史行车轨迹集的车辆出行模式可视分析方法研究[D]. 合肥:合肥工业大学计算机与信息学院, 2021.
- Wang Tao. Research on visual analysis method of vehicle travel pattern based on historical driving track set [D]. Hefei: School of Computer and Information, Hefei University of Technology, 2021.
- [15] 黄正国. 基于车牌识别数据的车辆出行特征研究[D]. 成都:西南交通大学交通运输与物流学院, 2019.
- Huang Zheng-guo. Research on vehicle travel feature based on license plate recognition data[D]. Chengdu: School of Transportation and Logistics, Southwest Jiaotong University, 2019.
- [16] 梁泉, 翁剑成, 林鹏飞, 等. 基于个体出行图谱的公共交通通勤行为辨别方法研究[J]. 交通运输系统工程与信息, 2018, 18(2): 100-107.
- Liang Quan, Weng Jian-cheng, Lin Peng-fei, et al. Public transport commuter identification based on individual travel graph[J]. Journal of Transportation Systems Engineering and Information Technology, 2018, 18(2): 100-107.
- [17] 马晓磊, 姚李亮, 沈宣良. 基于OBD数据的驾驶人出行模式挖掘[J]. 交通信息与安全, 2021, 39(2): 70-77.
- Ma Xiao-lei, Yao Li-liang, Shen Xuan-liang. Drivers' travel pattern mining based on OBD data[J]. Journal of Transport Information and Safety, 2021, 39(2): 70-77.
- [18] Rodriguez Alex, Laio Alessandro. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.