

基于高速公路交易数据的出行模式分析与 差异化收费策略*

吕能超^{1,2}, 董新雨^{1,2,3}, 罗如意⁴, 曾岳凯⁵, 徐达^{1,2}, 周新聪^{2,3}

1. 武汉理工大学智能交通系统研究中心, 湖北 武汉 430063
2. 武汉理工大学交通与物流工程学院, 湖北 武汉 430063
3. 水路交通控制全国重点实验室, 湖北 武汉 430063
4. 湖北交投科技发展有限公司, 湖北 武汉 430034
5. 武汉中交交通工程有限责任公司, 湖北 武汉 430050

摘要: 基于高速公路交易数据, 选取10项表征用户个体特征和出行时空特征的指标构建用户特征模型。采用K-means、模糊C-means以及自组织映射算法对用户特征进行分类, 并应用于某路段的ETC数据。研究结果表明, 相比于K-means和模糊C-means, SOM模型在用户出行模式分类上具有更优效果; 将高速公路出行用户划分为六类具有合理性。基于分类结果, 针对性提出了个性化差异收费策略, 并通过数值仿真验证了策略的合理性。

关键词: ETC数据; 聚类算法; 出行模式分析; 差异化收费

中图分类号: U491.5 **文献标志码:** A **文章编号:** 2097-0137(2025)03-0129-10

Travel pattern analysis and differentiated tolling strategies based on highway transaction data

LV Nengchao^{1,2}, DONG Xinyu^{1,2,3}, LUO Ruyi⁴, ZENG Yuekai⁵, XU Da^{1,2}, ZHOU Xincong^{2,3}

1. Intelligent Transport Systems Research Center, Wuhan University of Technology, Wuhan 430063, China
2. School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China
3. State Key Laboratory of Maritime Technology and Safety, Wuhan 430063, China
4. Hubei Communications Investment Technology Development Company Limited, Wuhan 430034, China
5. CSHCC Traffic Engineering Company Limited, Wuhan 430050, China

Abstract: Highway transaction data were utilized to select 10 indicators representing user-specific characteristics and the spatiotemporal features of travel, forming the basis for constructing a user characteristic model. To classify highway user characteristics, the K-means, fuzzy C-means, and self-organizing map algorithms were applied to ETC data from a specific road segment. The results indicate that, compared to K-means and fuzzy C-means, the SOM model performs better in classifying user travel patterns and supports the reasonable classification of highway users into six categories. Based on these classification results, a personalized differential tolling strategy is proposed, and its feasibility is validated through numerical simulation.

Key words: ETC data; clustering algorithm; travel mode analysis; differentiated charges

* 收稿日期: 2024-12-19 录用日期: 2025-02-21 网络首发日期: 2025-04-08
基金项目: 国家重点研发计划(2023YFB4302600); 国家自然科学基金(52472366, 52072290);
湖北省重点研发计划(2024BAB051)
作者简介: 吕能超(1982年生), 男; 研究方向: 智能网联交通; E-mail: 333167@whut.edu.cn

全文阅读



ZR20240357

出行模式划分有助于识别用户的出行特征,同时通过合理地配置交通资源,能够提升交通系统的运行效率(严新平等,2021)。动态定价等策略被认为是控制高速公路用户数量的有效手段(Yao et al.,2021)。以往的研究大多利用调查数据(张苇冲等,2023),并根据用户特点对用户群体进行分类。固定交通摄像头、激光雷达等传感器虽然能够获取静态数据,但这些数据通常受地理位置限制,并且无法提供出行者的连续信息,限制了对用户出行模式的深入分析。GPS 轨迹数据、手机数据、IC 卡数据等自动采集的数据类型,能够用于分析出行者的行为。然而由于样本量通常较小且采集过程既耗时又昂贵,在实际应用中存在一定限制(闫晟煜等,2024)。

ETC 系统全天候不间断运行,覆盖全国高速公路网络,实时将采集到的数据传回路网中心(李君羨等,2021)。海量 ETC 数据被广泛应用于如收费站通行能力分析(向往等,2023),路段拥堵判别(王灵丽等,2022),车辆异常运行状态辨识(邹兵等,2022),分车型制定动态限速方案(吕能超等,2023),交通运行参数预测等(温惠英等,2024)。然而受限于设备和技术条件,数据采集过程中可能存在缺失等异常情况。此外,尚未建立完善的出行轨迹溯源机制,无法对出行者的行程和行为模式全面分析。

多种数据挖掘方法被用于发现用户潜在的习惯行为特征。Qian et al.(2018)提出一种基于车辆行为特征的电子收费客户细分方法。林培群等(2023)提出基于货运风险特征画像的货车运行风险等级识别模型。魏广奇等(2019)借助 K-means 聚类方法识别使用高速公路日常通勤的车辆。Yun et al.(2023)建立铁路常旅客的分类管理模型,有针对性地制定服务提升计划。Zong et al.(2022)建立基于 SOM 的客户分层模型,为企业制定有针对性的营销策略。Cao et al.(2023)采用包括传统 K-均值、模糊 C-均值和自组织映射(SOM, self organizing maps)模型等探索个人出行规律并对聚类模型进行比较和评价。现有研究建模过程中所采用的特征指标较为单一。此外,缺乏对不同算法在出行模式分析中效果的系统性比较。

目前,针对高速公路出行的研究相对较少(Chen et al.,2021)。ETC 系统提供了丰富的数据追溯出行者的行驶轨迹,这为从出行时空特征角度分析用户行为提供可能(张弛等,2023)。个人行为受出行频次、出行距离等多种因素的影响,反

映了其出行决策和模式的差异性。这些差异性特征为差异化收费政策的制定奠定了基础(Ding et al.,2021)。因此,本文基于个体特征和出行时空特征等选取了 10 个指标构建高速公路用户特征模型。通过比较多个聚类算法,选择最优算法进行用户分类。并针对具有不同特征的用户群体,提出了差异化收费策略,旨在更好地引导交通流量分布、并提高道路使用效率。

1 出行模式划分模型

1.1 高速公路用户特征模型

结合个体特征和出行时空属性构建高速公路用户特征模型,如图 1 所示。其中:(1)车辆类别。高速公路行驶着 16 种不同类型的车辆,分为客车、货车和专项作业车三类,并根据核定载人数、车长和最大允许总质量等属性进一步细分。(2)出行开始、结束时段。通过提取所有车辆的行程起止时间并统计每小时的车辆数量,识别高峰和低谷期。出行时间分布如图 2 所示,根据分布规律将全天划分为不均等的 0:00-6:00, 6:00-10:00, 10:00-12:00, 12:00-15:00, 15:00-19:00 和 19:00-24:00 时间段。(3)出行距离。结合所建立的拓扑

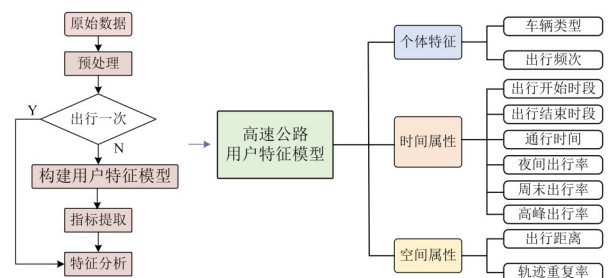


图 1 高速公路用户特征模型

Fig. 1 Characteristics model of highway users

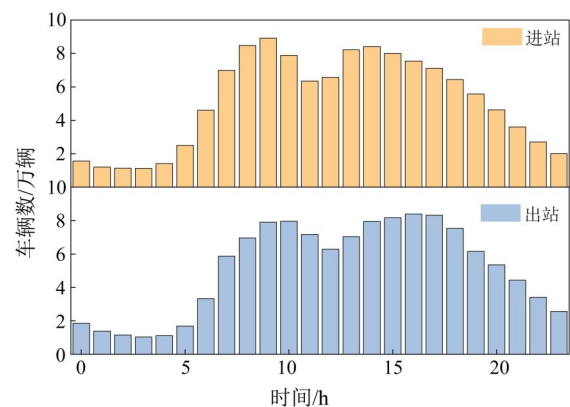


图 2 出行开始、结束时间分布图

Fig. 2 Distribution of trip start and end times

结构模型, 利用人工测距以及通过原始数据中的桩号等信息计算距离。(4)通行时间。车辆的行驶时间可分为两种, 第一种情况是行驶区段内包含服务区, 车辆可能会在服务区停留, 导致离开和驶入的时间差并不完全反映真实的行程时间。第二种情况是目标区段内不含服务区, 此时可通过计算车辆的行程时间差获得实际通行时间。(5)高峰出行率。高峰出行率是评估交通状况的重要指标, 高峰时段包括早高峰7:00-9:00和晚高峰17:00-19:00。(6)周末出行率。本文将周末出行定义为出行的开始时间和结束时间均在周末。(7)夜间出行率。车辆在0:00-5:00的行驶定义为夜间出行。该指标不仅反映了交通流量的时段分布, 还能揭示出行习惯和潜在的交通风险。(8)轨迹重复率是用户出行轨迹中重复路径的比例。通过计算用户在调查期内同一路径重复出现的次数与总出行次数之比来衡量。该指标有效反映用户的出行偏好。(9)出行频次。出行频次指用户在调查期内的所有出行次数。本文将用户从收费站进站到出站的全过程定义为一次出行。

1.2 自组织映射聚类算法

K-means算法计算复杂度低, 且对样本量的要求较小。但需要预设聚类数目和初始聚类中心, 设定不同会导致聚类结果不同(徐进等, 2022)。模糊聚类通过隶属度将样本点联系到多个聚类类别上, 聚类结果的清晰度受到模糊度因子 m 的影响。本文将 m 初始值设置在1.5~2.0之间(Ni et al., 2023)。基于模型的聚类算法利用特定模型进行聚类, 如基于神经网络的SOM算法等。由于高速公路用户的出行特征数据丰富但精度较低且空间分布没有明显规律, 本文采用K-means算法、模糊C-means算法和SOM算法聚类并评价, 选择最优算法进行用户分类。

SOM是一种无监督学习的神经网络模型(Zhang et al., 2020), 由一个二维网格组成, 每个节点表示一个神经元。输入数据通过与每个节点的连接进行传递, 节点间的连接权重会根据输入数据的特征进行调整, 使得相似数据在网格中聚集在相邻的节点上, 从而实现数据的无监督聚类, 其步骤如下:

(1)初始化: 从数据集中提取个体属性、时空间属性作为输入层。通过设定参数和权值, 初始化网络中的节点及连接权重。

(2)选择输入数据: 从数据集中随机选择一个样本 x 作为输入。计算到每个输出节点的距离

$d_i = \|x - w_i\|$, 其中 w_i 为第 i 个节点的权重向量。

(3)竞争: 选取距离最小的节点作为获胜节点。 $W^* = \arg \min \|x - w_i\|$ 是获胜节点。

(4)初始参数设置: 学习率过高会导致模型不稳定, 过低则会减缓收敛速度。综合考虑数据量和聚类指标的特性, 本文将初始学习率设置为0.1, 并逐渐减小; 将网格尺寸设置为 10×10 , 以充分捕捉数据的基本结构。

(5)更新权重: 采用梯度下降法更新各节点参数, 使得获胜节点更接近输入数据。 $w_i(t+1) = w_i(t) + \alpha(t)h_{ci}(t)(x - w_i(t))$, 其中 $w_i(t)$ 是节点 i 在时间 t 的权重。 $\alpha(t)$ 表示学习率。 $h_{ci}(t)$ 是邻域函数, 表示节点间影响程度, 本文使用高斯函数。

(6)重复步骤(2)~(5)直到达到收敛状态。

1.3 分类模型评价指标

选取了三个评价指标评估聚类模型的效果。其中, 评价指标误差平方和SSE表示所有样本点到其各自聚类中心的距离平方和。轮廓系数SC取值范围为 $[-1, 1]$, 值越大聚类效果越好。戴维斯-布尔丁指数DBI取值范围为 $[0, \infty)$, 其值越小聚类效果越好(彭艺等, 2024)。

2 预处理及数据集构建

2.1 数据获取

原始数据时间跨度为2023年1-6月, 字段示例见表1。此外, 收集了高速公路收费站、门架及互通节点的地理信息, 用于构建空间拓扑模型。

表1 原始数据字段示例
Table 1 Example of raw data fields

字段	数据类型	样例	描述
Vehicle ID	个人属性	鄂ABC123	区分标识
Vehicle type	个人属性	1	车辆类型
Total fee	个人属性	10403	出行计费
Entry station	空间属性	河南信阳西站	驶入站名
Exit station	空间属性	湖北小河站	驶出站名
Entry time	时间属性	01-05T07:27:47	驶入时间
Exit time	时间属性	01-05T18:57:56	驶出时间

2.2 数据处理

由于设备限制、相邻车道干扰以及恶劣天气等因素的影响, ETC系统产生的数据完整性较差并存在噪声干扰。因此, 数据处理流程主要包括数据清洗和聚合, 如图3所示。

本文采用Z-score方法对用户特征属性进行标

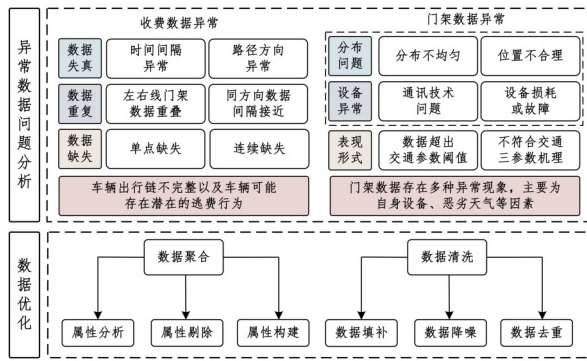


图3 数据处理流程图

Fig. 3 Data processing flowchart

准化处理(戢晓峰等, 2023)。对于数据失真问题, 通过获取基础信息并结合数据分布统计法进行修正。对于重复数据, 删除完全重复的数据以消除冗余信息; 对于部分相似的数据采取合并处理。同时, 利用交通流参数的时空分布规律和周期性显著的特征填补缺失数据。异常数据示例见表2。经过数据处理, 原始数据量减少至 18 820 345 条, 剔除约 1.37%(257 222 条)数据。

2.3 出行链构造

2.3.1 精准关联方法 车辆的通行卡标识唯一, 可利用该字段区分不同车辆。对于ETC流水数据, 选择ETC卡号作为初始关联的媒介; 对于MTC流水数据, 则使用CPC卡号。并通过“经过时间”(entry-time)将车辆的单次行程与多条流水记录关联并排序, 实现车辆轨迹的精准还原。

2.3.2 模糊关联方法 针对入口时间或出口时间缺失的情况, 以已知的时间信息为基准, 根据拓扑模型中节点的顺序, 采用滑动窗口方法向前搜索 n 小时内的流水数据, 通过模糊关联追踪符合条件的车辆行驶轨迹。

基于两种关联方法获得的轨迹数据, 若仍存在进站或出站信息缺失的情况, 可通过预测方法进行补充, 以重构完整的车辆出行链。如可利用构建的拓扑模型, 结合车辆的观测数据, 选择邻近节点填充缺失的位置信息; 同时, 根据门架与收费站间的历史平均车速推算相应的行程时间, 以填补时间数据的空缺。基于上述数据溯源方法构建的完整出行链, 如表3所示。表中 Passing-ID 为按顺序标识车辆经过的门架或其他监测设备。

表2 异常数据展示

Table 2 Display of anomalous data

Obumac-ID	Entry station	Entry time	Exit time	Vehicle ID	类型
67****45	456317A	01-05T07:27:47	01-05T18:57:56	鄂ABC123	正确
67****45	456317A	01-05T18:57:56	01-05T07:27:47	鄂ABC123	数据失真
67****45	456317B	01-05T07:27:47	01-05T18:57:56	鄂ABC123	
67****45	456317A	01-05T07:27:47	01-05T18:57:56	鄂XXX123	数据重复
67****45	456317A	01-05T07:27:47	01-05T18:57:56	鄂ABC123	
67****45	456317A	NULL	01-05T18:57:56	鄂ABC123	数据缺失
67****45	NULL	01-05T07:27:47	01-05T18:57:56	鄂ABC123	

表3 完整出行链

Table 3 Complete trip chain

Obumac-ID	Passing-ID	Entry time	Exit time	Sequences
42002CF5	44630C	2022-05-01T10:24:49	2022-05-01T10:40:42	1
42002CF5	44630F	2022-05-01T10:24:49	2022-05-01T10:56:05	2
42002CF5	446311	2022-05-01T10:24:49	2022-05-01T10:57:18	3
42002CF5	446312	2022-05-01T10:24:49	2022-05-01T11:00:59	4
42002CF5	446313	2022-05-01T10:24:49	2022-05-01T11:01:53	5
42002CF5	446314	2022-05-01T10:24:49	2022-05-01T11:05:12	6

3 聚类结果

3.1 聚类方法比较

针对构建的用户特征模型采用K-means、模糊C-means和SOM算法进行聚类分析, 结果见表4。与其他两种方法相比, SOM算法在聚类过程中表现出较低的SSE、较高的SC和较低的DBI, 这些指标均表明其聚类效果优越, 且稳定性较高。因此, 选择SOM算法作为本文的聚类方法。

表4 不同聚类方法评价结果

Table 4 Evaluation results of different clustering methods

	K-means	模糊C-means	SOM
SSE	29 671. 52	9 280. 66	6 472. 53
SC	0. 36	0. 43	0. 52
DBI	1. 40	0. 82	0. 75

在数据预处理阶段, 将仅出行一次的用户筛选并单独作为第一类用户, 从后续聚类数据中剔除。肘部法通过观察SSE随k值的变化趋势确定聚类数。如图4~5所示, 随着k的增加, SSE逐渐下降, 直至出现“肘部位置”, 此时SSE的下降速率减缓, 对应的k值通常为最优聚类数。为验证该值的合理性, 结合轮廓系数法进行综合评估, 最终确定最优聚类数k = 5。

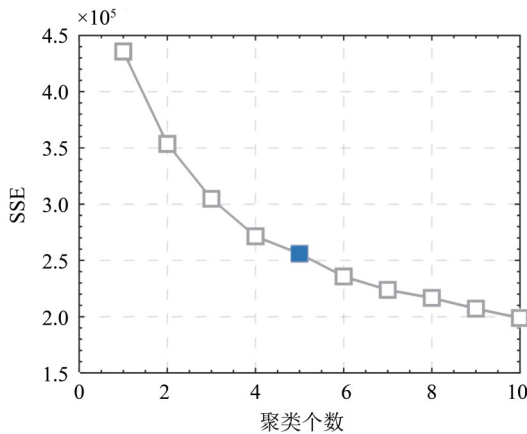


图4 肘部法聚类

Fig. 4 Elbow method clustering

3.2 用户出行特征分析

结合筛选出的第一类用户, 将出行用户划分为6类, 聚类结果如表5所示。统计结果显示, 研究群体的平均出行频率为6.55次, 说明大多数用户的出行频率较低且具有较高的随机性, 因此属于“第一类用户”, 即仅出行一次的用户。

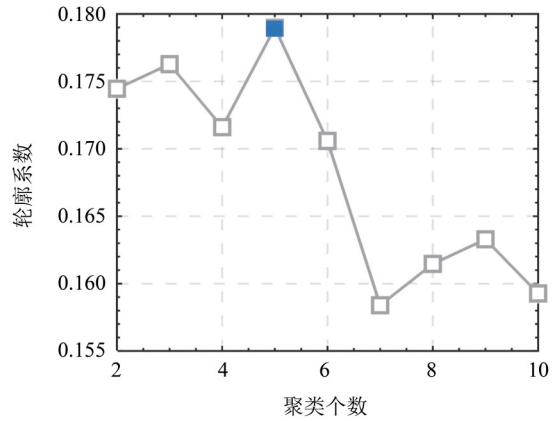


图5 轮廓系数法聚类

Fig. 5 Silhouette coefficient method clustering

第二类用户主要为客货混合用户, 其中客车用户偏多。以多次短距离出行为主, 出行时间和地点相对稳定, 通常在下午出行。将其定义为“高频低速下午出行用户”, 其特征为: ①出行强度较高。该类用户的平均月内总出行次数达到7.78次, 约有47.78%的用户出行次数超过3次。②出行距离较短。该类用户平均出行距离相对较短, 约有56.9%的用户的出行距离小于50 km。③行驶时间较长。由于包含较多货车, 这类用户的行驶速度较低, 因此行程时间相对较长。④出行时间稳定。该类用户主要集中在下午, 高峰期间较少出行。仅有8.17%选择在00:00-12:00间开始出行, 0.48%选择在此期间结束出行。相反, 约73.29%选择在12:00-19:00间开始出行, 75.3%在15:00-24:00间结束出行。

第三类用户主要为客货混合用户, 其中客车用户偏多。出行频次适中, 出行时间相对稳定, 通常在周末出行。将其定义为“低频远距离周末出行用户”, 其特征为: ①出行强度较低。用户的平均月内总出行次数仅为6.62次, 但单次出行距离较远。②周末出行率较高。进站、出站时间段均集中于时段5, 这表明该类用户出行时间稳定。此外, 该类用户在高峰时段和周末表现出较为频繁的出行活动, 而夜间出行相对较少; 有62.4%的用户选择周末出行。③空间稳定性较差。该类用户出行地点较为分散, OD重复率最低。

第四类用户主要为客、货车混合用户, 其中客车占比89.43%。这一类别的用户通常进行中长途出行, 其出行次数较多且稳定。将其定义为“中频时间稳定出行用户”, 其特征为: ①出行强度较高。平均总出行次数为7.29次, 有54.61%的用户出行次数在5次以上。用户出行主要集中在周

表 5 用户聚类结果示意
Table 5 Illustration of user clustering results

	聚类结果					
	类别 1	类别 2	类别 3	类别 4	类别 5	类别 6
客运车辆比例/%	19.35	59.46	91.78	89.43	65.72	5.34
货运车辆比例/%	80.41	40.34	8.16	10.51	34.15	94.53
进站时间段	3.74	4.69	4.78	2.10	1.95	3.57
出站时间段	4.00	5.05	4.92	2.24	2.20	3.33
行驶时间/s	17 036.73	5 382.02	2 291.52	1 993.12	3 525.87	39 745.93
距离/km	81.38	62.68	67.62	63.62	50.25	197.67
高峰出行率/%	22	22	25	30	32	15
周末出行率/%	20	24	28	25	22	20
夜间出行率/%	7	4	2	5	9	39
轨迹重复率/%	1	60	56	58	62	71
出行频次	1	7.78	6.62	7.29	10.29	6.34
样本数量	417 173	140 635	123 301	88 842	99 803	68 237

注:进出站时间段以小时为单位,因交通流的连续性及波动性,时间点并不严格落在整数小时上。

末。②出行距离较远。平均出行距离为 63.62 km。车辆平均行驶时间最短,仅为 33.2 min。③空间稳定性较高。用户轨迹的重复率主要分布在(0.4, 0.6]范围内,约占 53.3%。④时间稳定性最高。车辆进站时间段主要集中在时段 2,约占 63.2%;车辆出站时间段也主要集中在时段 2,约占 56.9%。

第五类用户是客、货车混合用户,货车占比 34.15%。在所有类别的用户中,出行频次最高,主要以短距离出行为主,出行 OD 较为稳定。将其定义为“高频短距离高峰时段出行用户”,其特征为:①出行强度大。这类用户的出行次数最多,平均每月总出行次数为 10.3 次,有 57.11%的用户出行次数超过 3 次。然而,他们的周末出行率较低。②出行距离最短。这类用户的平均出行距离为 50.25 km,67.71%的用户出行距离在 50 km 以内。③空间稳定性较高。这类用户的出行 OD 稳定性仅次于第五类,61.27%的用户轨迹重复率 > 0.5。④高峰出行率高。车辆的出行时间相对分散,但在高峰时段出行的次数较多,早高峰出行尤为突出。但他们的速度仍比第二类用户高。这说明货车的比例对速度的影响更大一些。

第六类用户是客、货车混合用户,其中货车为主,占比 94.53%。这类用户主要进行多次中长途出行,出行开始和结束时间相对固定,OD 稳定性最高。将其定义为“中频长距离地点稳定用户”,其特征为:①出行强度低。这类用户的整体

出行频次较低,周末出行率最低,但部分车辆频繁出行。②出行距离最远。该类用户的平均出行距离为 197.7 km,92.67%的用户出行距离超过 100 km。③空间稳定性最高。这类用户的出行 OD 稳定性最高,约有 64.83%的用户轨迹重复率大于 0.5。④出行时间分散。车辆的出行时间相对分散,早高峰出行率最低,其更倾向于在夜间出行。通过分析历史平均速度和车辆属性,这类车辆经常在收费站停靠进行休息。

3.3 用户时空特征描述

如图 6 所示,周一和周五的交通流量显著高于其他工作日。而在双休日中,周日的交通流量普遍高于周六。从断面交通流量的日变化趋势来看,无论是 85 分位值还是 15 分位值均呈现出高度相似的 V 形变化特征,反映了日常通勤活动的显著规律性。不同日期的交通流量按小时划分如图 7 所示,其变化呈现出两个显著高峰:早高峰(8:00-10:00)和晚高峰(16:00-19:00)。此外,工作日和周末的交通流量差异较小。并且自北向南行驶的车流量高于反方向(如图 8~9 所示),即靠近中心城区方向的流量明显多于远离中心城区的方向。

高速公路通勤的车辆以省内本地车牌为主,占比 89.85%(见图 10)。受汉孝同城效应的影响,孝感车牌占比仅次于武汉车牌,达到了 4.29%。在非本省籍通勤车辆中,南阳籍车牌占比最高,其次为深圳和长沙。

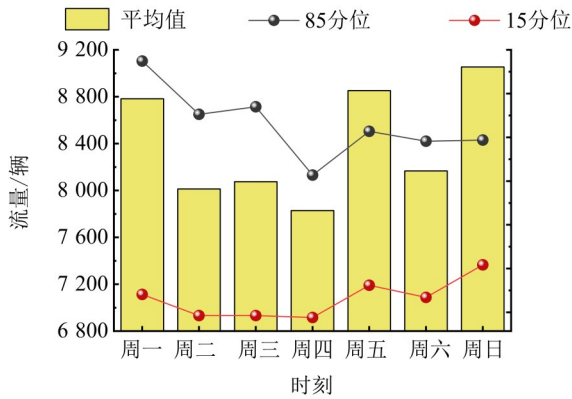


图6 按周分布的交通流量

Fig. 6 Distribution of traffic counts by week

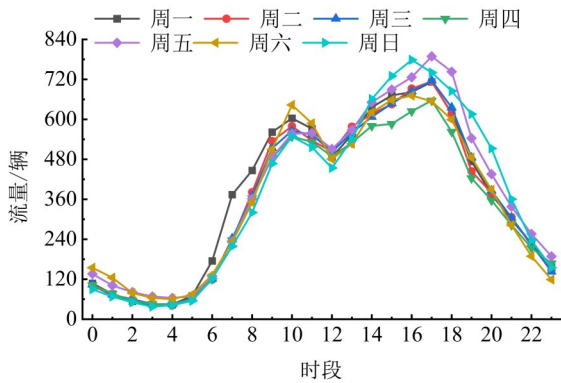


图7 按小时分布的交通流量

Fig. 7 Distribution of traffic counts by hour

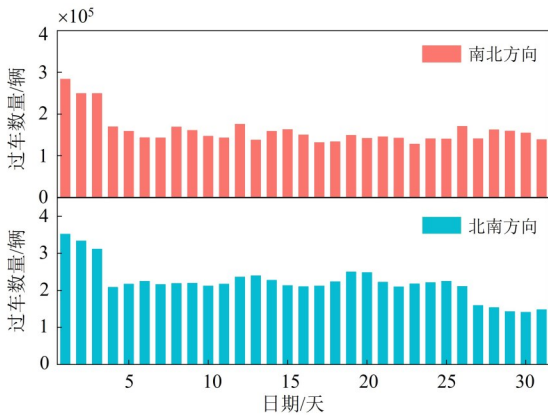


图8 双向车流量数据

Fig. 8 Bidirectional traffic

本地区的车辆通常进行短距离、短时间的省内出行,集中在关键通道,例如武汉西站到永安站、军山站到金口站。而长距离通勤主要驶向与本路段相连接的互道路段,如S13武监高速、S78蕲嘉高速、S15汉蔡高速和S17研孝高速。省外及跨省出行主要通过G42沪蓉高速、G50沪渝高速、G70福银高速和G4213麻安高速等。过境车辆的主要OD包括沪蓉高速—汉蔡高速、四环线—沪渝高

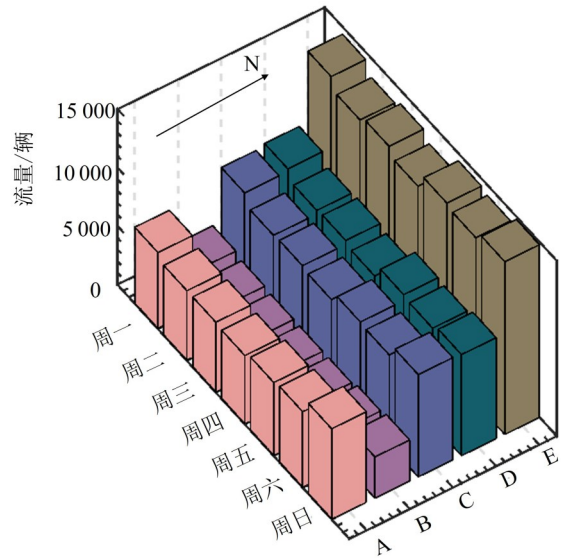


图9 典型区段切片流量统计

Fig. 9 Typical sliced flow statistics for sectors

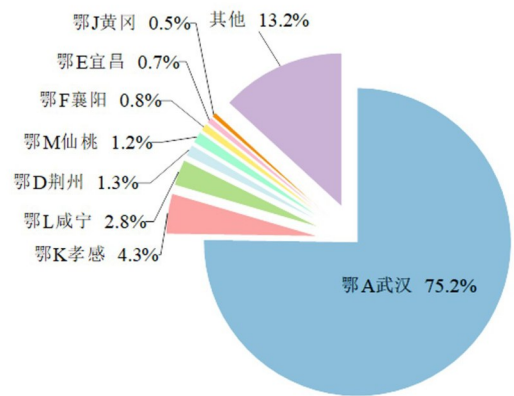


图10 省内和省外车牌地域分布

Fig. 10 Intra-province and out of province license plates

速、福银高速—研孝高速以及S337—麻安高速。

相较于人工收费用户,ETC驾驶员在高速公路上的累计行驶时间和里程更长。私家车司机倾向于选择ETC支付,而货车司机更偏好使用MTC

支付。超过 71.9% 的 ETC 支付车辆为客车。不同车型的流量和支付方式, 如图 11~12 所示。

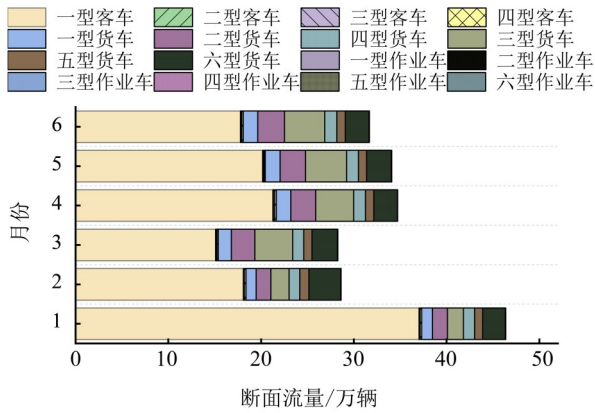


图 11 不同车型断面流量统计

Fig. 11 Cross-sectional flow for different vehicle types

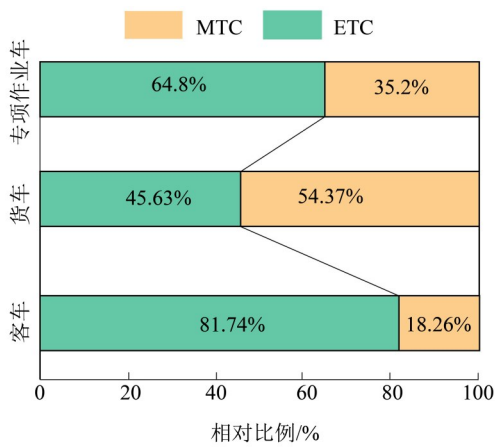


图 12 不同车型支付方式

Fig. 12 Payment methods for different vehicle types

3.4 差异化收费

3.4.1 拟定策略 目前, 高速公路的定价主要考虑出行需求、交通流量和历史价格等因素(Zhang et al., 2019)。然而, 现有的收费模式过于单一, 难以实现交通资源的高效利用。本文提出了一种基于出行模式的差异化收费方法, 具体措施包括: 针对出行频次较高的用户设定较低的收费标准; 在夜间及平峰时段提供优惠政策, 以促进错峰出行; 采用基于距离的梯度定价方法, 并对具有较强替代性的路段设定较低的收费标准; 针对定期通勤用户, 提供特定路段的通行费用折扣; 根据不同车型实施更加细化的定价策略等。针对 6 类不同特征用户的差异化收费建议见表 6。

表 6 不同类别用户的差异化收费

Table 6 Differentiated tolling strategies for different user

用户类别	用户特征	收费策略
1	仅出行一次	分时段+分支付
2	高频低速下午出行	分频次+分路段+分支付
3	低频远距离周末出行	分时段+分路段+分支付
4	中频时间稳定出行	分时段+分路段+分支付
5	高频短距离高峰时段出行	分频次+分时段+分车型
6	中频长距离地点稳定	分路段+分车型

3.4.2 实例验证 借鉴典型省市定价经验, 结合“高频短距离高峰时段出行用户”特征, 设计差异化收费策略。选取京港澳高速湖北段某区间作为示例路段, 开展数值仿真模拟实验。根据收费站及门架的位置, 将路段划分为若干断面(见图 13)。

路段编号	实验路段	道路等级	车道数	设计速度 (km/h)	距离 (km)	平行路段	道路等级	车道数	设计速度 (km/h)	距离 (km)
1	A-B	高速公路	4/6	100	13.77	a-b	一级公路	4	80	13.77±5
2	B-C	高速公路	4/6	100	16.66	b-c	一级公路	4	80	16.66±5
3	C-D	高速公路	4/6	100	10.61	c-d	一级公路	4	80	10.61±5
4	D-E	高速公路	4/6	100	15.91	d-e	一级公路	4	80	15.91±5
5	E-F	高速公路	4/6	100	12.98	e-f	一级公路	4	80	12.98±5
6	F-G	高速公路	4/6	100	17.97	f-g	一级公路	4	80	17.97±5
7	G-H	高速公路	4/6	100	10.4	g-h	一级公路	4	80	10.4±5
8	H-I	高速公路	4/6	100	11.75	h-i	一级公路	4	80	11.75±5
9	I-J	高速公路	4/6	100	18.03	i-j	一级公路	4	80	18.03±5

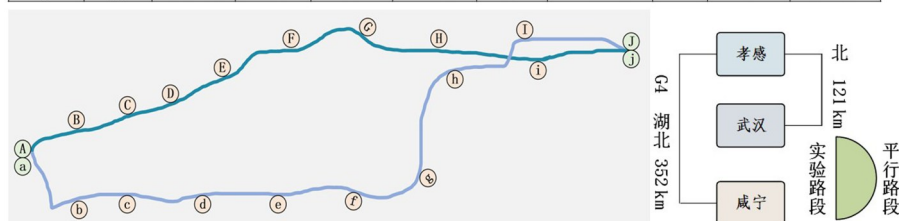


图 13 实验场景

Fig. 13 Experimental scenario

评价指标包括道路饱和度、流量变化率及路网平均负荷度。图14中, 研究区间各路段道路饱和度值均有上升, 道路利用率显著提高。同时所有路段的服务水平均维持在二级及以上。各路段流量变化如图15所示。高峰时段的每小时交通流量普遍下降, 其中晚高峰的降幅高于早高峰, 分别减少约7.8%和7.1%。平峰时段的每小时交通流量增加约7.3%。该策略有效诱导了高峰交通需求向通行费较低的平峰时段转移, 实现交通流量的时间均衡分布。图16展示了全高峰时段内路网平均整体饱和度的变化。饱和度呈现先增后减的趋势。路网的平均饱和度整体下降, 其中高峰时段的下降幅度最为显著, 并出现了时段前移现象。

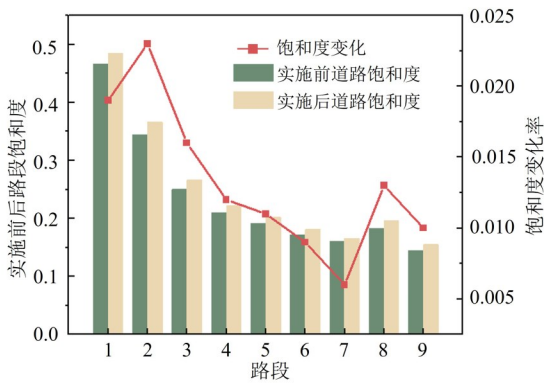


图14 路段饱和度变化

Fig. 14 Variation in segment saturation

4 结论

(1) 针对ETC数据提取涉及的失真、重复、缺失等异常问题, 构建了数据优化模型。该方法能够有效提升数据质量。同时, 基于优化后的数据提出了出行链溯源构造方法, 实现了对出行用户全链条行为的精准还原。

(2) 选取10项表征个体特征和出行时空特征的关键指标, 构建用户特征模型。采用K-means、模糊C-means和SOM算法进行分析, 发现SOM算法分类效果更优。同时, 将出行模式划分为六类能够全面刻画不同用户群体的特征。

参考文献:

戴晓峰, 孔晓丽, 陈方, 等, 2023. 基于ETC数据和A-BiLSTM神经网络的高速公路节假日短时交通流预测模型[J]. 交通信息与安全, 41(3): 166-174.
李君羨, 周一晨, 高志波, 等, 2021. 基于收费数据的高速公

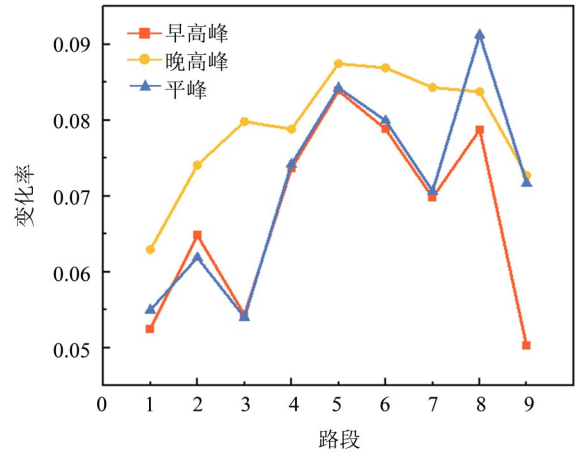


图15 路段流量变化率

Fig. 15 Segment traffic flow variation rate

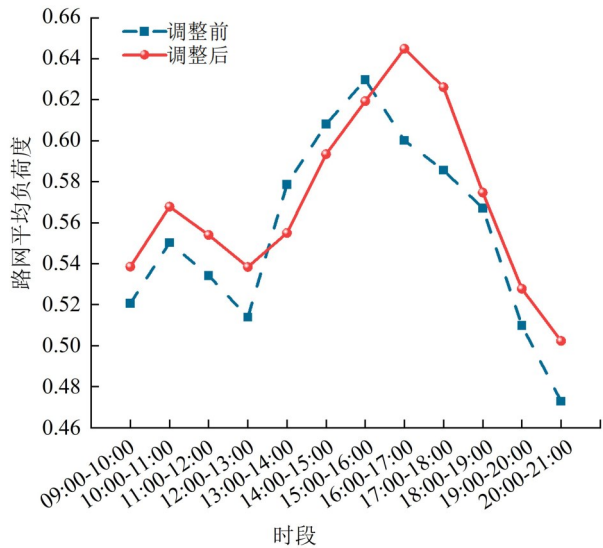


图16 路网平均负荷度变化

Fig. 16 Variation in the average network load factor

(3) 基于用户分类结果, 提出了差异化收费策略, 并通过数值仿真进行验证。结果表明, 实施策略后, 早、晚高峰时段的每小时交通流量分别下降约7.1%和7.8%, 而平峰时段增加约7.3%, 显著提升了交通系统的整体运行效率与服务水平。

路收费站通行能力分析[J]. 公路交通科技, 38(11): 106-116+125.
林培群, 龚敏平, 周楚昊, 2023. 面向运输风险识别的高速公路货车用户画像方法[J]. 华南理工大学学报(自然科学

- 学版),51(6):1-9.
- 吕能超,王玉刚,周颖,等,2023.道路交通安全分析与评价方法综述[J].中国公路学报,36(4):183-201.
- 彭艺,吴桐,杨青青,2024.下行RIS-NOMA的用户集群方法[J].中山大学学报(自然科学版中英文),63(1):128-136.
- 王灵丽,黄敏,薛田莉,等,2022.基于微观交通仿真的道路预约出行效益研究[J].中山大学学报(自然科学版中英文),61(6):89-98.
- 魏广奇,苏跃江,吴德馨,等,2019.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,19(3):237-244.
- 温惠英,何梓琦,胡宇晴,等,2024.高速公路新能源汽车碳排放量的测算及空间分布特征[J].华南理工大学学报(自然科学版),52(8):1-13.
- 向往,王超伦,于安康,等,2023.考虑高速公路主线收费通道引导信息面向对象的驾驶人换道行为研究[J].中国公路学报,36(8):257-270.
- 徐进,杨子邈,陈钦,等,2022.基于电子不停车收费数据的山区高速公路车速分布与车型分类研究[J].交通运输系统工程与信息,22(5):75-84+116.
- 闫晟煜,孙可欣,温福华,等,2024.收费模式重大调整后的高速公路运输量监测方法重构[J].交通运输工程学报,24(5):259-269.
- 严新平,褚端峰,刘佳仑,等,2021.智能交通发展的现状、挑战与展望[J].交通运输研究,7(6):2-10+22.
- 张驰,翟艺阳,王韩,等,2023.基于ETC数据的山区高速公路客货车行驶速度特征研究[J].中国公路学报,36(11):441-455.
- 张苇冲,杨涛,吕能超,2023.基于车辆位置与速度特征的驾驶行为模式分类方法[J].交通信息与安全,41(1):85-94.
- 邹兵,余志,何兆成,等,2022.基于车辆身份检测数据的城市道路使用特征[J].中山大学学报(自然科学版中英文),61(3):132-139.
- CAO R, CHEN X, JIA J, et al, 2023. Uncovering equity and travelers' behavior on the expressway: A case study of Shandong, China[J]. Sustainability, 15(11):8688.
- CHEN L W, CHEN D E, 2021. Exploring spatiotemporal mobilities of highway traffic flows for precise travel time estimation and prediction based on electronic toll collection data[J]. Veh Commun, 30:100356.
- DING W, WANG Z, CHEN J, et al, 2021. Potential trend discovery for highway drivers on spatio-temporal data[J]. Wirel Netw, 27(5):3407-3422.
- NI J, CHEN W, TONG J, et al, 2023. High-speed anomaly traffic detection based on staged frequency domain features[J]. J Inf Secur Appl, 77:103575.
- QIAN C, YANG M, LI P, et al, 2018. Application of customer segmentation for electronic toll collection: A case study[J]. J Adv Transp, 2018(1): 3635107.
- YAO E, WANG X, YANG Y, et al, 2021. Traffic flow estimation based on toll ticket data considering multitype vehicle impact[J]. J Transp Eng Part A Syst, 147(2):04020158.
- YUN J, JEONG S, KONG J, et al, 2023. Investigation of the policy effect of high-speed metropolitan rail service in seoul through latent class modeling analysis of stated preference survey [J]. Transp Res Rec J Transp Res Board, 2677(12):529-542.
- ZHANG J, WU J, HU B, et al, 2020. Intelligent fault diagnosis of rolling bearings using variational mode decomposition and self-organizing feature map[J]. J Vib Control, 26(21/22):1886-1897.
- ZHANG K, ZHOU Z, QI Y, et al, 2019. Dynamic pricing strategy for high occupancy toll lanes based on random forest and nested model [J]. IET Intell Transp Syst, 13(4):636-643.
- ZONG Y, PAN E, 2022. A SOM-based customer stratification model [J]. Wirel Commun Mob Comput, 2022(1): 7479110.

(责任编辑 王海蓉)