

文章编号: 2095-2163(2023)11-0180-05

中图分类号: U491.1

文献标志码: A

考虑上下客点与 POI 关系的网约车临时停靠点研究

邢雪, 翟娅奇

(吉林化工学院 信息与控制工程学院, 吉林 吉林 132022)

摘要: 针对当前网约车数量急剧增加的现状, 为避免网约车在商圈周围随意停靠造成交通拥堵, 引导网约车停靠在预先划分的区域将缓解交通压力, 因此对网约车临时停靠点的选址策略进行深入研究。利用滴滴出行成都市局部区域网约车公开数据集和城市热门 POI 数据, 使用 K-means 聚类算法, 分析 POI 数据和网约车上下客点数据之间的耦合关系, 从而提出网约车临时停靠点选址策略。研究表明: 网约车上客需求与附近 POI 数据密度正相关, 即 POI 密度越大, 网约车乘车需求越大。综上提出的结合上下车需求和 POI 聚类的网约车临时停靠点的选址方法, 可以为减轻交通拥堵提供数据支持, 具有实际意义。

关键词: 停靠点选址; 城市网约车停靠; K-means 聚类; POI 分析; 智慧交通

Study on the temporary parking point of online car hailing considering boarding point and POI

XING Xue, ZHAI Yaqi

(School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin Jilin 132022, China)

Abstract: In view of the rapid increase in the current number of online car hailing, in order to avoid the traffic congestion caused by the random parking of online car hailing around the business district, guiding online car hailing to park in the pre divided area will ease the traffic pressure. Therefore, the site selection strategy of online car hailing temporary stops could be studied in depth. Using the public data set of Didi Travel's local regional network taxi hailing in Chengdu and the city's popular POI data, K-means clustering algorithm is used to analyze the coupling relationship between POI data and online taxi boarding and alighting point data, thus proposing the location strategy of online taxi hailing temporary stops. The research shows that the demand for online car hailing is positively related to the nearby POI data density, that is, the greater the POI density is, the greater the demand for online car hailing is. To sum up, the proposed site selection method for temporary bus hailing stops based on boarding and alighting requirements and POI clustering can provide data support for reducing traffic congestion, which is of practical significance.

Key words: locatin selection of online car hailing; online car hailing of the city; K-means clustering; POI; smart transportation

0 引言

当前, 许多国家致力于交通运输向数字化、网络化、智能化发展, 努力构建智能交通系统^[1]。网约车的出现极大缓解了居民出行的困难, 并迅速成为城市居民出行的主要方式之一^[2]。网约车在便民的同时, 也给交通治理带来了一定的困难, 而城市繁华商圈周围路段也常存在拥堵、秩序混乱等问题, 网约车违规停车、等待乘客上下车更是加剧了交通拥堵^[3-6]。因此, 评估网约车上下客点的交通特性并

建立模型对网约车临时停靠点进行选址具有重要现实意义。

针对车辆停车需求问题, 目前已经有很多学者对此进行研究。杨圣文等学者^[7]对现有停车场进行调查, 对不满足停车需求现状的停车场利用建立的路网通行能力模型进行泊车路段的优化。Ren 等学者^[8]考虑到学生步行的可达性, 利用 k-means 的迭代聚类方法来分配学生和为校车停车站点进行选址, 以获取最短的学生上学的通勤时间。Vdovychenko^[9]通过分析车辆的行驶数据来对车辆

基金项目: 吉林省自然科学基金(20210101416JC)。

作者简介: 邢雪(1983-), 女, 博士, 副教授, 主要研究方向: 基于人工智能理论的智能交通关键技术。Email: xingx@jlicet.edu.cn; 翟娅奇(1996-), 女, 硕士研究生, 主要研究方向: 基于人工智能理论的智能交通关键技术。

收稿日期: 2022-11-17

的停靠点进行分类,随后对停靠点的分类与车辆运行路线流量之间的相关性进行建模,并进行实例分析。王国娟等学者^[10]根据设定的商业区停车管理水平评价指标建立基于模糊综合评价法的停车管理水平评价模型,研究表明构建的模型能够对商业区路内停车管理水平进行准确评价。王震邦等学者^[11]均衡考虑多个停车需求,设计基于交替方向乘子法(ADMM)分布式优化的停车匹配模型,实验证明模型优于其他常用模型,并有较优的收敛效果。郑慧敏^[12]对停车需求的影响因素进行分析,并提出基于滴滴出行 OD 数据的停车需求预测方法,利用模型进行公共停车场的选址分析证明了模型的实际意义。除此之外,双层目标模型^[13-14]、粒子群算法^[15]、DBSCAN^[16]等多种方法也被应用在停车选址的研究上。

以上研究已对车辆临时停车选址的可行性进行了论证,但是以网约车 OD 数据为出发点,并结合城市热门 POI 数据点对乘客上下车需求的分析,挖掘网约车停靠的交通特性,从而作为网约车临时停靠点的选址依据的相关研究较少。针对上述问题,通过对网约车 OD 数据和爬取到的城市热门 POI 数据进行相关性分析,研究网约车停靠需求的交通特性,在此基础上提出基于 POI 聚类的网约车临时停靠点的选址策略,并基于成都市滴滴出行数据集进行实例验证。

1 数据采集与特征分析

1.1 数据采集

POI(Point of Interest)被称作兴趣点,通常指互联网电子地图中的点类数据。应用 python 编写程序在高德地图上爬取成都局部区域的所有餐饮相关 POI 数据,提取时间范围为 2022-10-09~2022-10-10。爬取数据时选取的 POI 大类是餐饮服务,中类包括:餐饮相关场所、中餐厅、快餐厅、休闲餐饮场所、咖啡厅、茶艺馆、冷饮店、糕饼店、甜品店,字段信息包括 POI 的具体名称、类别、城市、地址、经纬度。

网约车 OD 需求数据是指乘客在乘坐网约车过程中的上下客数据,包括当前订单信息、上下客时刻时间戳以及上下客时刻位置经纬度。采用成都市局部区域网约车轨迹数据进行网约车 OD 数据提取,字段信息包括订单 ID、时间戳、当前位置经度、当前位置纬度。根据轨迹数据进行 OD 需求的提取过程如图 1 所示。

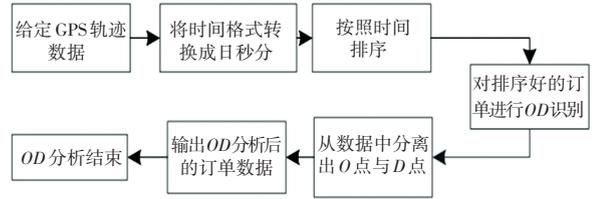
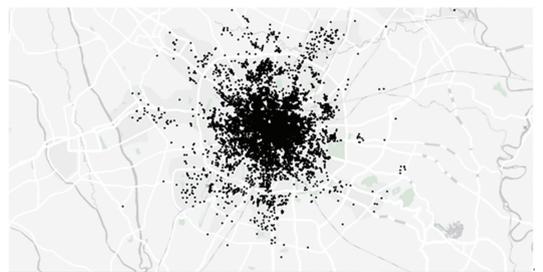


图 1 OD 数据获取示意图

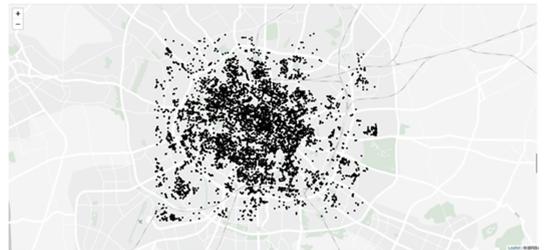
Fig. 1 Schematic diagram of OD data acquisition

1.2 OD 需求分布与城市 POI 耦合关系分析

基于高德地图,爬取得到餐饮相关 POI 位置分布,以及对应区域的网约车上下客需求点的分布,如图 2 所示。



(a) OD 需求数据分布情况



(b) 餐饮相关 POI 分布情况

图 2 成都市 OD 需求数据及餐饮相关 POI 分布情况

Fig. 2 Distribution of OD demand data and catering related POI data in Chengdu

近年来,网约车数量在城市的快速扩张极大地加强了城市交通站点不可达区域之间的通达性,改善了城市区域之间的联系。以成都市为例,对 OD 需求数据与餐饮相关 POI 数据(数据量分别为 104 148 条和 61 101 条)进行分析,研究网约车 OD 数据和城市餐饮 POI 分布之间的关系,乘客的网约车乘车需求顺应城市 POI 分布的空间特征,OD 需求密度分布以城市中心为起点向城市四周辐射开,网约车 OD 分布受城市 POI 点的影响,POI 分布与网约车 OD 点分布密度大的区域构成了城市繁华地带,并且 POI 与 OD 数据分布呈现明显的组团集聚分布形式,区域繁华程度随着远离城市中心逐渐递减。基于上述结合网约车需求量与 POI 的空间分布关系的分析,可以得出网约车 OD 点与 POI 分布

具有耦合关系,考虑网约车上下车点与餐饮相关 POI 进行网约车临时停靠点的选取具有一定的现实意义。

2 网约车 OD 需求与餐饮相关 POI 关系研究

采用 K-means 聚类算法对 OD 数据进行聚类,使空间上距离很近的 POI 尽可能被归为一类,形成广义上的 POI。在此基础上,研究广义上的 OD 需求点与周围一定半径范围内的 POI 数据之间的关系,为网约车的临时停靠点的选址提供依据。

2.1 基于 K-means 的 OD 数据聚类

K-means 算法是一种非监督的聚类算法,其主要思想是对于给定的 OD 需求点数据集,按照各个 OD 点之间的距离大小进行区分,将总的 OD 需求数据划分成 K 簇。划分后的 OD 数据的空间分布特点为簇内的 OD 点尽可能紧密相邻,而簇与簇之间的距离尽可能地大。设数据集为 $D = \{x_1, x_2, x_3, \dots, x_n\}$,表示 n 个数据点构成的集合, K-means 算法的实现过程详见如下:

Step 1 确定 K 个初始 OD 数据聚类中心 $C_1(0), C_2(0), C_3(0), \dots, C_K(0)$ 。

Step 2 对所有 OD 数据,求其到 K 个聚类中心的距离,并将数据点 x_n 归类到与聚类中心点距离最小的簇内 $D_m(m = 1, 2, 3, \dots, K)$,迭代 Q 次,判定公式如下:

$$\|x_n - C_j(Q)\| < \|x_n - C_i(Q)\| \quad (1)$$

其中, $i, j = 1, 2, 3, \dots, Q$ 。

Step 3 每次迭代,更新聚类中心。推得的公式为:

$$J_j = \sum_{x \in D_j(Q)} \|x - C_j(Q+1)\|^2 \quad (2)$$

当 J_j 取得最小值时:

$$C_j(Q+1) = \frac{1}{N_{jx \in D_j(Q)}} \sum x \quad (3)$$

其中, N_j 是 $D_j(Q)$ 的样本个数。

Step 4 利用 Step 2 和 Step 3 对聚类中心进行迭代更新,如果点的位置变化很小,可以判定达到稳定状态,迭代结束。

2.2 OD 需求与 POI 数据的相关性分析

使用 K-means 聚类算法对网约车 OD 需求数据进行聚类,以各个类簇的质心作为广义的 OD 需求点,并以广义的网约车 OD 点为中心,半径为 R 画圆作为广义 OD 点的吸引范围,对于半径选取,一般认为是 400~800 m^[17],目前应用较为广泛的是 800 m,

以 800 m 作为覆盖半径,统计区域内所有 POI 数据的个数,对成都市局部区域进行网约车 OD 点和餐饮相关 POI 的统计见表 1。使用线性回归来拟合通过 K-means 算法聚类得到的各个类簇中的网约车 OD 数据与广义 OD 点吸引范围内的 POI 数据之间的数量关系。以广义 OD 需求点吸引范围内的 POI 数量为 x ,类簇中的 OD 需求点数量为 y ,可得:

$$y = ax + b \quad (4)$$

代入成都市数据可得 x 和 y 之间的函数关系如图 3 所示, x 和 y 满足 $y = 1.705\ 014\ 42x + 152.182\ 346\ 1$,对应的拟合度 R^2 为 0.964 786 13。结果表明,OD 需求数据与 POI 数据的数量关系上存在线性正相关关系。

表 1 成都市局部区域 OD 数据与 POI 统计情况表

Tab. 1 OD data and POI statistics of local regions in Chengdu

| POI 分组 | 广义网约车 上车点个数 | POI 总数 | 网约车上车 需求总数 | 平均每个上车 需求点的需求数 |
|---------------|----------------|-----------|---------------|-------------------|
| [1,200) | 16 | 2 219 | 8 155 | 453 |
| [200,400) | 12 | 3 769 | 7 733 | 644 |
| [400,600) | 7 | 3 686 | 6 959 | 869 |
| [600,800) | 10 | 7 439 | 13 274 | 1 206 |
| [800,1 000) | 8 | 8 107 | 13 981 | 1 553 |
| [1 000,1 500) | 16 | 18 933 | 36 599 | 2 152 |
| [1 500,2 000) | 3 | 5 579 | 9 936 | 3 321 |

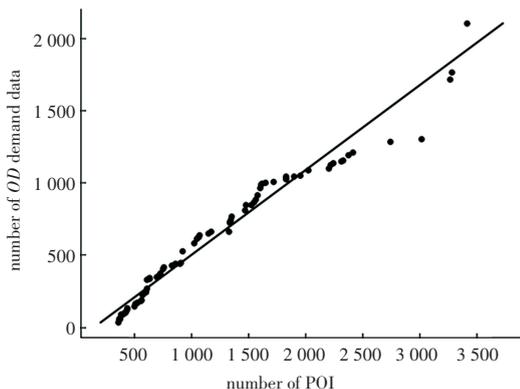


图 3 成都市 OD 需求与 POI 数量的关系曲线

Fig. 3 Relation curve between OD demand and POI quantity in Chengdu

3 结合 OD 点和 POI 的网约车临时停靠点的选址及实现

3.1 网约车临时停靠点选址

上述分析表明 OD 需求数据与餐饮相关 POI 数量之间存在线性正相关性,基于 OD 需求点和餐饮相关 POI 数据进行聚类,以得到的聚类结果作为网

约车临时停靠点的选址是具有合理性的, 网约车的临时停靠点的选址策略如下:

Step 1 获取相应城市区域的餐饮相关 POI 数据和网约车轨迹数据。

Step 2 对网约车轨迹数据进行 OD 提取, 并基于 K-means 聚类算法对 OD 数据进行聚类, 得到广义上的 OD 需求点。

Step 3 计算广义上的 OD 需求点吸引半径内的 POI 的数量, 筛选出吸引范围内 POI 数量大于等于某个特定值的 OD 需求点, 将其作为初始候选的网约车临时停靠点。

Step 4 对候选的网约车临时停靠点吸引范围内的 POI 进行聚类, 得到各个初选网约车临时停靠点吸引范围内的 POI 的聚类中心, 即是最终的网约车临时停靠点的选址。

3.2 实例验证

以成都市为例, 验证分析选址策略的有效性。

根据选址策略可得:

(1) 获取相关数据: 爬取获得成都市局部区域餐饮相关 POI 数据、共 61 101 个数据点; 使用滴滴全样本轨迹数据作为实验数据, 数据主要来自于 2016 年 11 月的成都市二环局部区域轨迹数据, 共计 706 万条, 对 11 月 1 日 08:00~17:00 时间内的轨迹数据进行 OD 数据提取操作后, 获取到共计 104 147 条 OD 数据。

(2) 基于 K-means 聚类算法进行聚类得到 500 个聚类中心, 即广义上的 OD 需求点。

(3) 以聚类得到的 OD 需求点为圆心, 统计 800 m 范围内 POI 数量大于 100 的 OD 需求点, 将其作为初选的网约车临时停靠点, 总共筛选获得 318 个初候选点。

(4) 对初始候选点吸引范围内的 POI 进行聚类, 得到的聚类中心即是网约车临时停靠点的最终选址。

当前网约车上车点分布如图 4 所示, 得到的拟选的网约车临时停靠点如图 5 所示。由图 4 和图 5 可以看出, 餐饮商家和 OD 点呈现边缘分散、中间密集分布规律, 分布密集的区域主要位于城市中心区域和南部区域, 由 K-means 聚类得到的广义的网约车上下车点的分布与城市网约车上下车的餐饮相关 POI 的热门区域相重合, 表明广义上的网约车上下车点位于人流量密集区域; 通过对广义的网约车上下车点根据范围内的 POI 数量进行筛选, 完成对临时网约车停靠点选址的优化。优化后的网约车临

时停靠点同时具备网约车需求量大、人流量密集等特点, 该选址策略对于缓解城市交通拥堵、规范网约车运行具有重要实际意义。

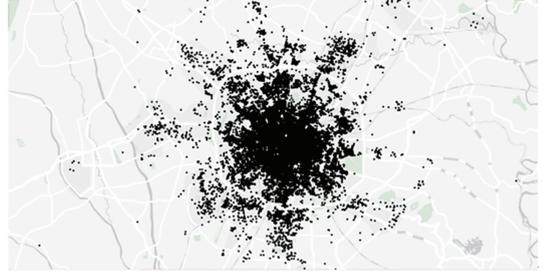


图 4 当前网约车上车点分布图

Fig. 4 Distribution of boarding points for current online car hailing

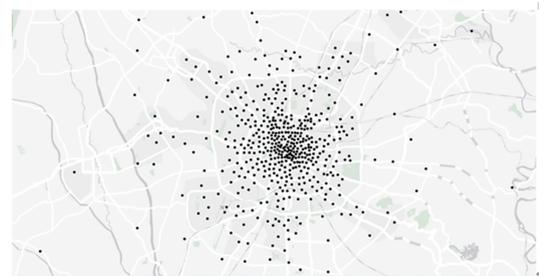


图 5 拟选网约车临时停靠点

Fig. 5 Proposed temporary parking point for online car hailing

4 结束语

针对网约车临时停靠点的选址策略问题进行初步研究, 考虑到以往的网约车运行特点和城市热门 POI 分布特点, 建立了考虑 OD 点与 POI 数据状态下的基于 K-means 聚类算法的网约车临时停靠点的选址策略。以城市中餐饮相关 POI 数据和网约车 OD 数据作为研究对象, 研究发现网约车 OD 数据与周边 POI 数据密度之间存在正相关关系, 利用模型进行网约车临时停靠点的选址。最后以成都局部区域的网约车 OD 数据和 POI 数据进行实例验证, 对 K-means 聚类算法得到的候选点进行筛选得到最终的网约车临时停靠点, 并论证该选址策略的合理性和可行性。未来将在以下方面进行深入研究:

(1) 除餐饮相关 POI 外, 可以研究其他类型的 POI 数据的聚类分析对于网约车临时停靠点选址的影响, 进一步建立考虑 OD 数据和 POI 的城市区域划分方式的优化模型。

(2) 对城市进行功能上的交通小区的划分, 并进行交通小区分类, 对于不同分类的交通小区提供不同的网约车临时停靠点的选址。

参考文献

- [J]. 中国公路学报, 2016, 29(6):1-161.
- [2] 第47次中国互联网络发展状况统计报告[EB/OL]. (2021-02-03) [2022-10-09]. http://www.cac.gov.cn/2020-09/29/c_1602939918747816.htm.
- [3] 邝先验,吴玉刚,陈奕希,等. 考虑上下客车辆路边临停的混合交通流模型[J]. 计算机工程与设计, 2021, 42(11):3286-3292.
- [4] 李玲. 交通新业态对成都市的社会影响与治理研究[J]. 中国新通信, 2021, 23(14):153-156.
- [5] 邓星,程望杰,彭良波,等. 基于POI数据的城市公共停车场复合开发模式研究—以武汉市为例[J]. 城市建筑, 2022, 19(10):65-69.
- [6] 傅思萍. 智慧城市智能停车诱导系统设计[J]. 宁德师范学院学报(自然科学版), 2021, 33(2):166-173, 185.
- [7] 杨圣文,李特,刘庆华,等. 城市建成区路侧停车规划模型及算法[J]. 重庆交通大学学报(自然科学版), 2022, 41(8):30-36.
- [8] REN Jingxuan, JIN Wenzhou, WU Weitao. A two-stage algorithm for school bus stop location and routing problem with walking accessibility and mixed load[J]. IEEE Access, 2019, 7(7):119519-119540.
- [9] VDOVYCHENKO V. Assessment of the influence of the time spent by vehicles at the stopping point of urban passenger transport on the level of conflict in the interaction of the route flow[J]. Technology Audit and Production Reserves, 2020, 3(2(53)):47-51.
- [10] 王国娟,吕文红,付守艳,等. 基于层次分析法的商业区路内停车管理水平评价[J]. 科学技术与工程, 2020, 20(21):8810-8816.
- [11] 王震邦,宋运忠. 考虑多方利益的大规模共享停车匹配优化策略[J]. 复杂系统与复杂性科学, 2023, 20(2):52-59.
- [12] 郑敏慧. 基于滴滴网约车OD数据的停车规划方法研究[D]. 北京:北京交通大学, 2018.
- [13] 何胜学,高蕾. 停车需求分布的双层规划模型及算法[J]. 交通运输系统工程与信息, 2019, 19(1):83-88, 96.
- [14] 杨晓芳,付强,牛兆雨. 基于停车诱导信息板的最优停车场推荐的双层目标模型[J]. 计算机应用研究, 2014, 31(10):3017-3019, 3040.
- [15] 张文会,苏永民,戴静,等. 居住区共享停车泊位分配模型[J]. 交通运输系统工程与信息, 2019, 19(1):89-96.
- [16] 李小龙,马英奇. 基于DBSCAN与LOF算法的共享汽车计价和停车点策略分析[J]. 新型工业化, 2021, 11(8):6-8.
- [17] HSIAO S, LU Jian, STERLING J, et al. Use of geographic information system for analysis of transit pedestrian access[J]. Transportation Research Record Journal of the Transportation Research Board, 1997, 1604(1):50-59.