

浅论基于手机信令数据的人员出行停驻点分析

姚 鑫¹ 朱海明²

1. 天津港保税区规划国土和建设交通局 天津 300000

2. 天津城建设计院有限公司 天津 300000

摘要: 分析手机信令数据在城市规划和交通规划中已经得到了广泛的应用, 人员停驻点分析是进一步分析人员出行链, 描述人员出行特征和规律中最关键的一环。通过对手机信令数据预处理, 采用DBSCAN算法可以很好的得到手机停驻点, 并对DBSCAN算法的优缺点进行分析。最后, 总结了论文的发现和局限性, 并展望了未来的研究方向, 包括个性化交通规划、多模态出行分析、城市紧急事件应对、城市设计与空间规划等。

关键词: 手机信令; 停驻点分析; 交通规划

引言

1 背景介绍

随着全球城市化的不断加速, 城市面临着一系列复杂的挑战。随着城市人口规模不断增长, 机动车保有量的逐年提高, 导致城市交通系统的负担不断加重, 交通拥堵和交通事故频发, 给城市交通规划工作者提出新的课题和挑战。伴随着计算机技术的发展, 大数据分析逐渐成为交通规划工作者的重要工具和手段, 通过大数据分析能够弥补传统调查的不足, 获取更加真实、全面的城市交通特征。手机信令因其同时具有时间、空间属性, 成为城市规划领域应用最为广泛的数据。因此手机信令数据中人员停驻点分析, 是后续人员职住空间分析、出行强度、出行时间、出行空间分布等核心特征的前提, 对其展开研究具有实际的应用价值。

2 方法与数据收集

2.1 数据来源

本研究采用电信运营商脱敏后的某城市一个月的信令数据进行分析, 手机信令数据记录了用户在不同时间段内通信基站位置的变化, 从而描绘了用户的移动轨迹和行为模式。在此基础上, 经过一个月的统计, 通过识别日间、夜间停留最多的区域, 再辅以人员个体属性信息, 可以得出人员就业地(就学地)、居住地等交通规划最为重要的属性信息, 并以此为基础进行通勤空间分布、时间分布等指标的统计; 停驻点分析则是在此研究的基础上进一步细化, 通过对人员停驻点的判断, 从而得出人员一次出行的起终点, 从而完善人员的出行链, 更加完善人员一天的活动规律, 进而为城市规划和交通设计提供定量支持。

2.2 数据预处理

首先, 检查数据是否按照预期的格式进行存储, 例

如日期、时间、地理坐标等是否符合标准格式。再识别并处理缺失值, 可以使用插值、填充默认值等方法。检测并处理异常值, 这些异常值可能会影响后续分析的准确性。在数据中, 可能会存在重复的记录, 需要去除这些重复数据, 确保分析的准确性。将日期时间列转换为统一的日期时间格式, 以便后续的时间序列分析。然后过滤外地用户, 可以使用前面提到的方法之一, 根据手机号码归属地或基站位置进行过滤。如果数据来自不同的数据源, 需要将它们整合成一个统一的数据集, 确保数据的完整性和一致性。在数据预处理的最后, 可以进行简单的数据可视化, 例如绘制数据分布、查看缺失值情况等, 以帮助更好地理解数据的特点。大致过程可以分为几下几步:

第一步: 记录合并, 并计算停留时间。首先将在连续时间内重复出现在相同位置的记录合并为一条记录。合并规则为: 合并后开始时间=首条记录的开始时间; 合并后结束时间=末条记录的结束时间; 合并后停留时长=末条记录的结束时间-首条记录的开始时间。

第二步: 漂移处理, 并更新停留时间。由于部分人员可能处于几个基站的边界, 存在手机所属基站的漂移的情况, 主要表现在两个临近网格反复交替出现2次以上(如ABAB...), 且距离 $\leq 2.5\text{km}$ (考虑到基站最大服务半径约 1.5km), 需要对该类记录进行去漂移处理, 处理后新的网格位置判定规则如下: (1) 如果A或B为居住或工作地点, 将其它漂移点都统一调整为居住或工作地点所在的区域; (2) 如果A或B都不是居住或工作地点, 将累计停留时间相对较多的网格位置作为新的网格位置。

第三步: 去掉漂移数据以后, 再次进行连续记录合并处理, 并更新停留时间。

2.3 停驻点识别

停车点的识别主要包括以下几步：

第一步：在数据预处理后的前提下，初步根据停留时间（一般取大于10分钟）计算停车点。



图1 典型数据停车点分析示意

第二步：为过滤距离过近的出行，以单个用户为研究对象，选取不同的距离作为聚类分析距离阈值（考虑到网格为300m*300m，因此，最小取300m），对各停留点进行聚类。

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 是一种常用的聚类算法，适用于基于密度的空间数据聚类。在手机信令数据中，可以使用DBSCAN来进行停车点分析，即识别出人群在城市中的停留区域。数据一般处理步骤：首先对数据进行预处理分析，再计算每个用户之间的距离，可以使用欧氏距离、曼哈顿距离等。这些距离将用于判断数据点之间的密度。选定合适的 ϵ 和 MinPts 参数，使用选定的 ϵ 和 MinPts 参数对数据进行 DBSCAN 聚类（一般取300米）。聚类的结果将包括核心点、边界点和噪声点。根据聚类结果，识别出核心点，这些点代表了停车点。每个核心点所属的簇即为一个停车点区域。最后将停车点区域在地图上进行可视化，以便观察人群在城市中的停留模式。



图2 典型数据停车点聚类分析示意

聚类后，新的停留点位置确定规则如下：

(1) 如果同一聚类内某停留点为居住或工作地，则将居住地或工作地作为新的停留点；

(2) 如果同一聚类内所有停留点均为非居住地和非工作地，则将停留时间最长的停留点作为新的停留。

依据上述分析，计算人员的停车点主要依赖停靠时间和聚类阈值半径，为了更加准确的分析人员的日均出行次数（出行强度），对不同出行时间和聚类半径得出的结果进行了相关分析，从统计结论来看停车15分钟是变化的拐点，即停车时间大于15分钟之后对于出行率的影响程度逐渐变缓；从聚类半径来看，300米即已经趋向平稳（城市内基站网格大小一般为300m*300m），故停车时间大于15分钟，聚类半径距离大于300米，对于识别

手机出行停车点较为合理。

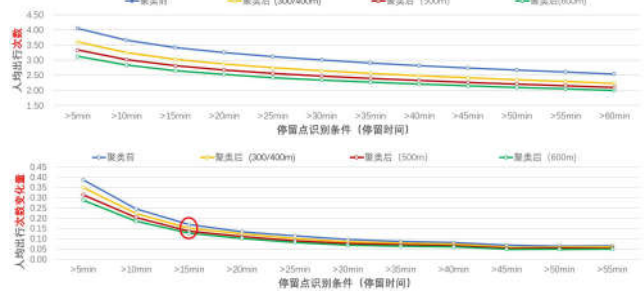


图3 人均出行次数、变化量随停留时间、聚类半径的变化图

2.4 DBSCAN算法的优点

DBSCAN能够识别出任意形状的簇，不仅限于凸形状，因此适用于复杂的城市地理环境和人群分布；可以自动排除噪声点（即那些没有足够邻居的数据点），从而有效地应对真实数据中的干扰；与一些传统的聚类算法（如K均值）不同，DBSCAN不需要事先指定簇的数量，它能够根据数据的密度分布自动判断；DBSCAN能够适应不同密度的区域，即使在不同部分的数据点密度不一致的情况下也能有效聚类；虽然DBSCAN的参数（ ϵ 和 MinPts）需要设置，但对于不同数据集和问题，可以根据实际情况进行调整；相对于一些其他聚类算法，DBSCAN的计算复杂度相对较小，使其适用于大规模数据集；DBSCAN可以帮助识别人流活动的热点区域，从而在城市规划中找到适合增设交通设施、商业服务设施的地点，从而为城市规划和交通设计提供精准数据支持。

2.5 停车点主要应用

(1) 城市交通热点区域识别。通过分析停车点分析，可以识别出交通拥堵的热点区域和高峰时段，从而指导交通规划者优化道路规划。通过调整道路布局、增设交通信号灯等措施，可以有效减少交通拥堵，提高道路通行效率。

(2) 公共交通优化。停车点分析可以揭示人们的出行偏好和主要出行路径。基于这些信息，可以优化公共交通线路的设置，确保公共交通服务更好地满足人们的出行需求，从而减少私人汽车的使用，缓解城市交通压力。

(3) 优化城市空间布局。通过了解人群流动的热点区域和流向，城市规划者可以更好地决定不同区域的用途规划，例如商业区、住宅区和休闲区的划分。

(4) 制定应急响应策略。通过监测停车点分析，可以更早地识别出现人员聚集的迹象，如景区人员滞留、体育场人员疏散等。这有助于制定更快速、高效的交通应对策略，以减少交通干扰、保障人员出行安全。

3 研究局限与展望

3.1 研究局限

尽管手机信令数据在停驻点分析和交通规划研究中具有很大的潜力,但也存在一些局限性需要后续考虑。目前存在的困难主要包括以下几方面。

3.1.1 采样偏差

手机信令数据的采集通常基于移动设备的连接情况,可能存在采样偏差。一些人群可能不使用移动设备或不连接网络,因此可能被排除在数据分析之外,导致对停驻点分析的不完全描述。

3.1.2 隐私问题

手机信令数据涉及个人位置和行动信息,引发隐私担忧(本文研究均采用脱敏后的安全数据)。在进行数据分析时,必须严格遵循隐私保护法律和规定,确保数据的匿名性和安全性。

3.1.3 数据稀疏性

尽管手机信令数据规模庞大,但在某些区域或时间段可能存在数据稀疏性问题,导致分析结果的不够准确。特别是在偏远地区或人流较少的时间段,数据的可用性可能受限。

3.1.4 位置误差

手机信令数据反馈的是当前基站的位置,而基站有一定的服务半径(在人员相对稀疏的区域覆盖范围较大),这可能导致基站位置与实际人员位置存在一定的偏差,此外处于两个基站边缘地区人员,存在基站来回切换的情况,也会对数据分析造成一定的困扰。

3.1.5 社会差异

不同人群的手机使用习惯和移动模式可能存在差异,例如不同年龄、职业、收入阶层等。这可能导致分析结果偏向某些特定人群,影响对整体停驻点分析的理解。

因此尽管手机信令数据在停驻点分析和交通规划研究中具有巨大潜力,但研究人员需要认识到其局限性,并结合其他数据源和方法,以确保研究结果的准确性和可靠性。

3.2 未来研究方向

3.2.1 个性化交通规划

利用手机信令数据分析,可以实现更加个性化的交通规划。未来可以开发智能交通规划系统,根据个人出行习惯和行为模式,为每个人提供最佳的出行方案,从而减少交通拥堵和时间浪费。

3.2.2 多模态出行分析

未来研究可以进一步深入探讨不同出行方式(步行、自行车、公共交通、私家车等)之间的关系,以及如何将这些模式结合起来,实现更加高效的多模态出行。

3.2.3 城市紧急事件应对

手机信令数据可以用于识别城市中的突发事件,如交通事故、自然灾害等,从而实现更加精准的紧急事件应对和资源调配。

3.2.4 城市设计与空间规划

基于手机信令数据的停驻点分析,可以为城市规划者提供重要的参考,帮助他们更好地设计城市空间,布局公共设施,优化商业区和居住区。

3.2.5 人工智能与预测模型

结合人工智能和机器学习技术,未来可以开发更精准的人群流动预测模型,帮助交通规划者更好地预测和应对交通拥堵等问题。

结束语

通过手机信令数据的人员出行停驻点的分析,不仅可以较为精准的识别城市人员的职住空间,而且可以得出人员一天中出行的所有停驻点,呈现出人员完整的出行链条,可以避免传统调查中,受各种因素影响导致出行记录填写不完整,出行强度低估的问题,但与传统调查相比同样存在的一定的不足,比如因人员隐私保护缺少相关经济属性,无法进行下一步的深入挖掘,也包括人员出行方式的识别,仍需要进一步研究。但从整体来看,基于手机信令数据的人员出行停驻点分析,对于全面研究人员一天的出行活动和规律打开了一扇大门,可以为后续制定更加精细化的城市规划方案、城市治理方案提供相对准确和直观的数据支撑,在城市规划和交通规划中具有一定的实际应用价值。

参考文献

- [1]史立凯,韩竹斌,胡城峰.基于手机信令数据的城市交通需求预测方法研究[J].运输经理世界,2022(25):58-60.
- [2]尹怡晓,郭煜东,郝萌等.基于手机信令数据和Needleman-Wunsch序列匹配算法的都市圈出行方式识别[J].交通运输研究,2023,9(01):40-50.DOI:10.16503/j.cnki.2095-9931.2023.01.005.
- [3]刘鹏,林航飞.基于手机信令数据的职住地识别方法[J].综合运输,2022,44(05):14-17+33.
- [4]李建邨.基于手机信令数据的职住地获取研究[D].东南大学,2019.DOI:10.27014/d.cnki.gdnau.2019.004315.
- [5]吴雨佳,尹伟石,孟品超.基于手机信令数据的用户位置预测方法研究[J].长春理工大学学报(自然科学版),2022,45(05):130-137.
- [6]Yingna Z,王悦,胡昊宇, et al.基于手机信令大数据的京津冀城市群人口时空分布与流动特征分析[J].地域研究与开发,2023,42(03):161-167+180.
- [7]张海涛.基于手机信令数据的交通模式识别与预测研究[D].南京邮电大学,2022.DOI:10.27251/d.cnki.gnjdc.2022.001287.