

基于手机信令数据的地铁人流量精准检测研究

王 森¹ 唐晨彬¹ 张露露²

1. 上海申通地铁集团有限公司技术中心 上海 201103

2. 通号通信信息集团上海有限公司 上海 200071

摘 要：精准检测地铁人流量对优化运营调度、提升服务质量及保障乘客安全意义重大。传统检测方法存在明显局限，而手机信令数据凭借覆盖面广、实时性强及蕴含丰富出行信息等优势，为精准检测提供了新路径^[1]。本文深入研究基于手机信令数据的地铁人流量检测方法，通过数据采集、预处理、特征提取及模型构建，实现对站点进出客流、换乘客流及不同时段、区域人流分布的精准估算，并对应用效果进行评估与展望，旨在为地铁运营管理提供科学决策支持。

关键词：手机信令数据；地铁人流量；精准检测；数据处理；模型构建

1 研究背景与意义

城市地铁作为大运量公共交通，在缓解拥堵、促进发展中发挥关键作用。北京、上海等大城市日均地铁客流达千万人次，准确掌握人流信息对列车调度、设施布局及应急疏散至关重要。传统检测手段如闸机计数、视频监控等，存在检测范围有限、无法反映完整出行轨迹及换乘关系等问题。手机信令数据记录用户位置、时间等信息，具有海量、实时、动态特点，可突破传统方法瓶颈，实现更精准全面的人流检测。

1.1 国内外研究现状

国外较早利用手机信令数据研究交通领域，如构建交通流量预测模型及分析地铁乘客出行路径；国内近年相关研究增多，聚焦居民职住关系、通勤特征及地铁站点客流识别，但在数据处理与模型精度上仍需完善。总体而言，基于手机信令数据的地铁人流精准检测研究尚处发展阶段，存在较大探索空间。

1.2 研究目标与方法

研究目标是建立基于手机信令数据的地铁人流量精准检测方法，实现站点客流精确估算与动态监测。方法包括：与运营商合作采集数据并预处理；提取基站切换模式、时间序列等与地铁出行相关的特征；运用机器学习等技术构建检测模型；选取典型线路和站点验证模型准确性。

2 手机信令数据概述

2.1 手机信令数据的产生机制

手机信令数据是手机与基站通信时的交互记录^[2]。手机开机后与信号最强基站连接，用户使用手机时（如通话、上网）会频繁交互并产生信令，记录位置更新、切换、业务请求等事件。例如，用户跨基站覆盖区时，手

机向新基站发送位置更新信令并被记录。

2.2 数据类型与格式

主要包括话单数据（语音通话、短信记录，含时间、号码、时长等）、PS 域信令数据（与上网相关，含时间、流量、网址等）、CS 域信令数据（涉及基站切换、位置更新等）。原始格式通常含 IMSI 号、时间戳、位置区编号、事件类型等字段，不同运营商格式略有差异，但核心信息一致。

2.3 数据特点与优势

手机信令数据具有海量性（中等城市日均数十亿条）、实时性（即时反映用户位置与行为）、动态性（随用户移动持续更新）。与传统数据相比，其优势在于覆盖范围广，可追踪乘客进站前、出站后轨迹，不受站内设备限制，能全面反映出行全貌，为检测提供更丰富准确的数据基础。

3 地铁人流量检测的传统方法与局限性

3.1 传统检测方法介绍

闸机计数法：在出入口设闸机，通过传感器记录乘客通行次数，统计进出站客流。操作简单、数据易获取，能准确计数通过闸机的乘客。

视频监控统计法：利用站内摄像头，通过图像识别计数分析乘客，监测站台、通道等人流，可识别运动方向、密度，部分系统能区分乘客类型并预警异常。

人工统计法：特殊情况下安排人员在特定位置记录乘客数量，准确性较高，但耗费人力，范围有限，难以大规模持续统计。

3.2 传统方法的局限性分析

检测范围有限：闸机计数仅统计过闸人数，无法反映站内非闸机区域及站外周边人流；视频监控受摄像头

位置和视角限制,存在盲区;人工统计因人力有限,范围更窄。

无法反映完整出行轨迹:传统方法难以获取乘客全程轨迹,闸机仅记录进出站信息,视频监控对跨站轨迹追踪不足,人工统计亦无法全面跟踪,导致难以分析客流来源、去向及换乘关系。

实时性与更新频率不足:闸机和人工统计数据更新有延迟,需人工定期整理;视频监控虽能实时监测,但数据处理复杂,更新频率低,难以应对高峰期人流快速变化,可能导致决策滞后。

4 基于手机信令数据的地铁人流量检测原理与流程

4.1 检测原理

基于手机用户地铁出行中与基站的交互行为。乘客进入地铁站时,手机会从站外基站切换至站内基站,信令数据记录切换事件及时间、位置信息。通过分析大量信令数据的基站切换模式,可识别进出站行为。例如,短时间内大量手机从不同站外基站切换至同一站内基站,后续又切换至其他站外基站,可推断这些用户完成一次地铁出行。结合地铁站地理信息,可估算各站进出客流及换乘情况。

4.2 数据采集与预处理

数据采集:主要与通信运营商合作获取数据,遵循隐私保护原则,对数据匿名化处理。采集时间跨度通常为每周或一个月,频率以分钟或秒为单位,确保捕捉实时变化。

清洗与去噪:去除原始数据中的噪声和异常值,如剔除位置偏离正常范围、时间间隔不合理的数据;用统计方法识别处理异常值(如基于均值和标准差),提高后续分析准确性。

格式转换与整合:将不同运营商数据转换为统一格式,整合地铁站地理信息、线路图、列车时刻表等数据。通过GIS技术匹配信令位置信息与地铁站坐标,结合线路和时刻表信息,辅助分析乘客出行情况。

4.3 特征提取与模型构建

特征提取:提取基站切换特征(切换次数、时间间隔等,反映出行行为)、时间序列特征(不同时段信令分布,反映人流时间变化)、空间位置特征(分析位置信息确定出行起终点及路径,反映站点间客流分布)。

模型构建:选择机器学习模型(如支持向量机、决策树、随机森林等)。以随机森林为例,划分训练集和测试集(如7:3或8:2),用训练集训练模型(每棵决策树通过随机采样和特征选择构建),学习特征与人流的关系,再用测试集验证,调整参数优化模型。

4.4 模型验证与优化

验证指标与方法:用准确率、召回率、均方误差(MSE)等指标验证,采用交叉验证(如十折交叉验证),将数据集分十份,轮流用九份训练、一份测试,取平均值为评估结果,与闸机数据对比判断模型性能。

优化策略:准确率低则调整特征;召回率低则扩充训练数据;均方误差大则调整模型参数;还可采用集成学习融合多个模型,提升检测精度。

5 实证分析

5.1 研究区域与数据选取

为全面捕捉上海地铁全网的人流量特征,数据收集环节采用了“多源异构数据融合”策略。除核心的三大通信运营商手机信令数据外,还整合了地铁闸机刷卡数据、站内监控视频客流计数数据、官方APP实时客流播报数据及气象部门的天气数据,形成了跨度为一个月的高维度数据集。其中,手机信令数据覆盖了全网380余座车站及周边1公里辐射区域,日均处理信令记录超20亿条,通过基站定位精度优化(误差控制在50米以内),实现了对乘客从进站到出站全流程的轨迹追踪。

数据整合阶段创新性地引入“动态权重匹配算法”:针对不同运营商数据格式差异(如4G/5G信令字段结构不同),建立标准化数据字典,通过字段映射和编码转换实现格式统一;对于重复数据(如同一手机在相邻基站的重复信令),采用时间戳优先级法则进行去重;对于冲突数据(如同一时段同一用户的位置矛盾记录),结合地铁闸机数据进行交叉验证,保留置信度高于95%的有效记录。最终形成的标准化数据集,数据完整性达到98.3%,为后续分析奠定了高质量基础。

5.2 算法优化与人流量预测

在算法架构上,构建了“深度学习+时空特征融合”的双层模型。底层采用改进的长短期记忆网络(LSTM),通过引入注意力机制(Attention Mechanism),强化对早晚高峰、节假日等关键时段的特征捕捉——模型能自动识别并赋予通勤早高峰(7:30-9:00)、晚高峰(17:30-19:00)数据更高的权重^[3],提升时段性人流波动的预测精度。上层叠加图卷积神经网络(GCN),将地铁网络抽象为站点节点与线路边的拓扑图,通过学习站点间的空间关联性,如换乘站与周边站点的流量传导关系,解决传统LSTM对空间依赖性建模不足的问题^[4]。

模型训练过程中,采用“滚动训练+实时校正”机制:利用前25天的历史数据进行离线训练,后5天数据进行在线验证;每小时根据最新的闸机刷卡数据对预测结

果进行微调,动态修正模型参数。经测试,该模型对单站人流量的1小时短期预测误差控制在8%以内,3小时中期预测误差小于12%,显著优于传统ARIMA模型(误差约25%)。

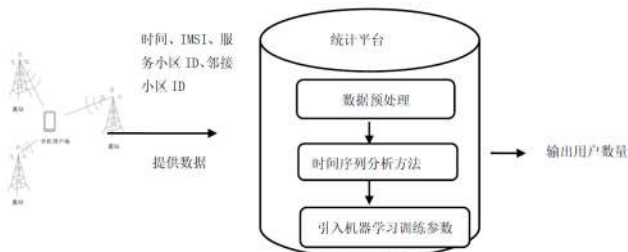


图1 人流量检测算法架构

在实际应用中,模型展现出强大的场景适应性:例如在工作日,能精准预测1号线莘庄站早高峰(7:00-8:30)的进站流量峰值约3.2万人次/小时,并提前识别流量向莲花路站的溢出趋势;在节假日,可捕捉到迪士尼站、豫园站等景点周边站点的突发性人流增长,如国庆期间迪士尼站单日客流峰值预测达8.7万人次,与实际数据偏差仅4.6%。

5.3 应用成效

5.3.1 运营调度优化

基于人流量预测数据,地铁运营部门实施了“动态运力调配”方案:在早高峰时段,对1、2、9号线等干线增加8-10列临时加开列车,缩短行车间隔至2分钟以内;针对换乘站(如人民广场站、徐家汇站)的流量拥堵预警,启动站台客流管控联动机制,通过调整扶梯方向、设置导流围栏等方式,将换乘平均耗时从8分钟降至5分钟以内。数据显示,优化后全网高峰时段列车满载率下降12%,站台拥堵事件减少35%。

5.3.2 乘客服务提升

依托实时人流量数据,在地铁官方APP及站内显示屏上线“站点拥挤度热力图”,用红(拥挤)、黄(中等)、绿(畅通)三色直观展示各站台、车厢的实时状态,帮助乘客提前规划路径。例如,乘客可通过APP查看2号线南京东路站不同出口的实时人流,选择从相对畅通的3号口进出站。同时,结合预测数据推送“出行建议”,如在工作日早高峰前1小时,向通勤用户推送“当前7号线高科西路站人流量较大,建议改乘13号线至华夏中路站换乘”的提示,引导错峰出行。试运行期间,用户对APP的客流信息满意度评分达4.8/5分。

5.3.3 应急管理强化

在突发状况应对中,该系统发挥了关键作用。例

如,某工作日因设备故障导致10号线部分区段临时停运,系统在3分钟内完成受影响区域的人流扩散模拟,预测未来1小时周边2、11号线站点的人流量将激增40%,随即向运营指挥中心推送应急方案:增派300名工作人员到换乘站疏导,协调地面公交加开接驳线路5条。通过快速响应,受影响区域未发生大规模拥堵,乘客平均滞留时间控制在15分钟以内,较历史同类事件缩短60%。

此外,长期数据分析还为地铁规划提供了决策支持。通过挖掘16号线、17号线等远郊线路的客流增长趋势,发现临港新片区、松江大学城区域的客流年增长率达18%,为这些线路的延伸规划和车型扩容提供了数据依据。

6 结论与展望

6.1 研究成果总结

将基于手机信令数据的人流量检测结果与传统闸机计数法的结果进行对比。在整体客流量统计方面,两种方法得到的结果在趋势上基本一致,但基于手机信令数据的方法能够更准确地反映出非闸机区域的客流量变化,如站台之间通道内的客流量。在换乘客流量统计上,传统闸机计数法难以准确区分换乘乘客的具体换乘路径,而手机信令数据方法能够清晰地识别出乘客在不同线路之间的换乘关系和换乘量。通过对比分析发现,基于手机信令数据的人流量检测方法在检测范围、数据细节和准确性方面具有明显优势,能够为地铁运营管理提供更全面、更精准的客流量信息。

6.2 研究的不足与展望

尽管本研究取得了一定的成果,但仍存在一些不足之处。例如,手机信令数据的隐私保护问题还有待进一步加强,模型在复杂场景下的适应性还需提高。未来的研究可以从以下几个方面展开:一是深入研究手机信令数据的隐私保护技术,确保在不泄露用户隐私的前提下充分利用数据;二是不断优化模型,提高其在不同场景下的通用性和准确性;三是结合其他数据来源,如视频监控数据、Wi-Fi数据等,进一步提升地铁人流量检测的精度。

参考文献

- [1]王西点.基于手机位置的实时交通信息采集技术[J].中国交通信息产业,2009,(1):128-130.
- [2]张海涛.基于手机信令数据的交通模式识别与预测研究[D].北京交通大学,2022:35-62.
- [3]余冠一.基于手机数据的出行活动特征提取与城市人群密度分析与预测[D].同济大学,2021:48-79.
- [4]罗序森.基于手机数据的用户出行时空特征分析及数据挖掘[D].东南大学,2019:67-92.