

文章编号: 2095-2163(2021)06-0026-05

中图分类号: U231+.92

文献标志码: A

上海轨道交通车厢满载率实时感知与估计

赵源

(同济大学道路与交通工程教育部重点实验室, 上海 201804)

摘要: 本文通过利用 WIFI 探针设备, 将其布设在上海轨道交通列车车厢内探测客流, 同时收集同日同车次的车厢称重数据、断面客流数据进行数据分析与融合, 得到列车车厢满载率的估计值, 从而建立起更为精确的地铁车厢满载率数据。

关键词: 轨道交通; 车厢满载率; WIFI 探针

Real-time perception and estimation of full load rate of Shanghai metro carriage

ZHAO Yuan

(Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai 201804, China)

[Abstract] This paper uses WIFI probe equipment to detect passengers flow in Shanghai metro carriages. At the same time, it collects the weighing data and cross-sectional passengers flow data of the same day and the same carriage for data analysis and fusion, and obtains the estimated value of the full load rate of the carriages, so as to establish more accurate metro carriages full load rate data.

[Key words] rail transit; full load rate; WIFI probe

0 引言

近年来, 城市人口快速增加, 轨道交通已经成为大众出行首选的主要交通方式, 具有安全性高、速度快、环保、容量大等多方面优势, 越来越受到广大乘客的欢迎。随着城市轨道交通的发展与信息技术手段的进步, 城市轨道交通的智慧运营已经成为了城市轨道交通发展的重要方向, 其中地铁车厢满载率是关系到运营管理的重要数据之一。现有车厢满载率一般采用车辆称重或者票卡清分的方式而得到, 计算得到的满载率以分级的方式呈现, 表现方式较为粗犷, 获取的数据精度不足, 不足以支撑智慧化、精细化的运营管理。如何利用科学可靠的技术实现地铁车厢内乘客均匀度的智能调控, 促进智慧地铁的发展, 成为当前研究的重要课题。

车厢满载率是通过车厢额定载客量与实际载客量的比值来确定。现有一些关于车厢满载率的研究论文。如: 文献[1]中提出, 通过对车辆定员概念进行分析, 将车辆定员与拥挤度进行探讨, 得出车厢定员标准应为 6 人/平米; 文献[2]介绍了中国和美国对于满载率的定义, 运用对比分析和统计学方法, 对公交车辆额定乘载人数进行修正, 建立了新的满载率计算模型, 从而实现公交线路车辆满载率理论

定义的修正; 文献[3]在基于视频图像分析的基础上, 利用卷积神经网络提出了一种地铁列车车厢拥挤度识别方法; 文献[4]在北京地铁 4 号线的断面区间车厢进行实地调研, 从而获取车厢立席密度, 进一步得到车厢满载率与乘客满意度之间的关系, 建立了乘客舒适度模型函数等等。

WIFI 探针是一种新型采集客流量的设备, 其可以检测到利用智能设备连接上 WIFI 用户的 MAC 地址等信息, 这些信息可记录一定的身份识别功能且具有唯一性。WIFI 探针具有扩展性较强且信号传输不易受非视距误差及多径衰落影响等优点。为解决当前现有的地铁车厢满载率数据精度不够的难题, 本论文拟将使用 WIFI 探针在地铁车厢内进行实地客流量检测, 通过与同日同车次的车厢称重数据、断面客流数据进行数据融合, 得到列车车厢满载率的估计值。

1 现状研究

1.1 客流数据采集技术研究

国外的轨道交通发展起步较早, 关于客流数据的采集方面有了较为先进的研究, 主要是将军用视频数据的采集与跟踪思路引入行人数据采集管理中。国内外客流数据的采集方法主要分为: 人工调

基金项目: 上海轨道交通车厢满载率实时感知与估计(19)GP027。

作者简介: 赵源(1978-), 男, 博士研究生, 高级工程师, 主要研究方向: 轨道交通运营安全与行车组织优化。

收稿日期: 2021-01-22

查统计法、自动检票计票统计法、闭路监视器观察法、图像处理技术法、无线电波监测粗略预测法。

中国轨道交通发展起步较晚,但随着科技的不断进步,在轨道交通客流采集方面,也取得了较大的进步。图像技术是目前适用范围最广的技术之一,在车站的使用频率最高。但由于图像识别设备价格较高,若在大范围内使用,将带来高成本的压力。红

外传感器、压力传感器也是常用的客流采集技术,但其只适合低密度客流量。比如,安装在电梯处或者公交车的入口处,并不适用地铁车站环境。目前,中国采用的轨道交通客流采集技术^[5]有:超声波定位采集技术、自动检票机采集技术、人工采集技术等。各种客流采集技术的优缺点详见表 1。

表 1 客流采集技术对比表

Tab. 1 Passengers flow collection technology comparison table

采用方法	优点	缺点
压力板式客流采集技术	定位较准确、实施成本低、周期短	系统部件易损坏,后期调试保养费用高
超声波定位采集技术	系统搭载框架携带方便、低功耗	定位误差大,设备的成本大
红外定位采集技术	定位精度高、可信度高	设备成本高、工作量大、系统易受环境影响
图像处理技术	图像识别可信度高、信息提取速度快	采集精度不高、稳定性欠缺
自动检票机采集技术	有效采集进出站整体客流数据	无法定位站内分布的客流密度信息
人工采集技术	设备投资耗费低、客流统计指标全面	人工调查工作量大,后期处理成本工作复杂

1.2 数据挖掘算法研究

数据挖掘^[6-9]是指从大量的数据中,通过算法发现其中有用信息的过程。国外对于数据挖掘技术的研究起步较早,已有了较为广泛的应用。如电信、金融、体育、政府管理、医疗服务等领域。国内虽起步较晚,但发展较快,近几年的研究主要集中在算法方面。数据挖掘的理论基础可以归结到以下几个方面:

- (1) 数据归约;
- (2) 数据压缩;
- (3) 模式发现;
- (4) 概率理论;
- (5) 微观经济观点;
- (6) 归纳数据库。

1.3 客流数据预测研究

客流预测分为短时客流预测和长时客流预测,在长时间的研究成果中,形成了各自所适用的研究模型,其中 ARIMA 模型和 SVM 模型的使用较为广泛。而 ARIMA 模型被广泛适用于客流的预测,运算步骤可以归纳为 4 个阶段:参数的预测、序列规则选择、模型优化及校验。客流的准确预测是提高轨道交通运营安全的关键,其中对客流数量的预测属于交通量预测的重要组成部分。客流预测与预测模型主要有神经网络模型、计量模型、非线性系统理论模型等。

1.4 上海轨道交通列车车厢满载率研究现状

目前有关轨道交通列车车厢满载率的研究,大都是在获得客流量的基础上,通过计算得到列车的

满载率。获得客流量的方法,除了传统的 AFC 计数装置和人工客流调查以外,基于高清视频识别、红外传感器、热敏传感器等自动化客流采集技术日趋成熟。但是,由于高成本、覆盖范围小、精度低等问题而难以得到推广和应用。WIFI 探针、蓝牙等技术已被用于各大商场客流数据采集以及车站站厅客流采集及拥堵点分析,但单项技术获取的数据量精度仍然较低。现有针对轨道交通列车满载率的研究方法,主要有车厢称重数据及票务清分数据。车厢称重数据获取的满载率较为粗犷,精度不高;票务清分数据是根据 AFC 刷卡数据所获取的断面客流量通过清分模型得到的,其准确性也有待进一步提升。

目前上海轨道交通列车车厢实际满载率的感知还处在研究阶段,数据精度不高,尚具有一定的提升挖掘空间。

2 车厢满载率估计方法

2.1 票务清分系统模型

票务清分系统主要是通过清分模型,将运营过程中的票款收入,按照一定的规则分配到各运营线路上。票务清分系统模型通过乘客的进出站进行分配,得到乘客的出行路径,进而得到线路的断面客流。

清分方法的发展主要经历了以下几个阶段:

- (1) 人工分账法;
- (2) 出站确认法;
- (3) 基于乘客出行路径的清分方法。

其中,基于乘客出行路径的清分方法主要包括

3 种:最短路径清分法、K 短路径清分法和多路径选择概率清分算法。乘客出行路径的清分方法的思路及适用情况详见表 2。

2.2 车厢称重系统

车厢称重系统是通过安装在车厢中的称重传感器,来检测车厢中客流的一种装置。称重传感器是一种将质量信号转变为可测量的电信号输出的装置。车厢称重系统采集客流的基本原理为:车厢称重系统通过称重传感器,将车厢内采集到的数据进行存储;当轨道交通列车运行到站台时,称重系统自动连接站台区域覆盖的无线网络,并将所采集到的信息进行“车-地”传输;再通过线路(既有传输系统中百兆以太网通道)传输至线路侧 OCC(运行控制

中心);由 OCC 综合前置网络交换机,上传至 COCC(地铁线网指挥中心)综合前置网络交换机,进入分析服务器。称重系统客流数据采集过程如图 1 所示。

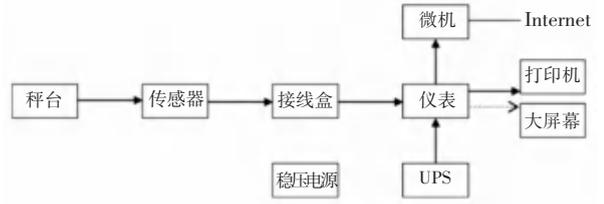


图 1 称重系统客流数据采集过程

Fig. 1 Passengers flow data collection process of weighing system

表 2 乘客出行路径的清分方法对比表

Tab. 2 Comparison table of classification methods for passengers walking routes

清分方法	思路	适用情况
最短路径清分法	假设乘客按照最短路径出行	线网规模不复杂
K 短路径清分算法	将 OD 间的客流分配给路径阻抗最小的 K 条路径	线网规模较复杂
多路径选择概率清分法	综合考虑乘客出行路径影响因素,利用阻抗结合清分规则筛选出有效路径,最后得到客流分配比例	线网规模复杂

上海轨道交通车厢称重系统采集到的数据内容包括:列车编号、车厢号、车厢舒适程度、车厢温度、

车厢载重、时间等信息,见表 3。车厢称重系统将车厢的拥挤度情况分为 3 类:舒适、较拥挤、拥挤。

表 3 上海轨道交通车厢称重数据

Tab. 3 Weighing data of Shanghai metro carriages

列车编号	车厢号	舒适程度	车厢温度单位(摄氏度)	车厢载重单位(吨)	数据日期	时间
17015	170851	舒适	20	39.8	20201201	14:30:00
17015	170862	舒适	20.1	43.2	20201201	14:30:00
...

3 WIFI 探针采集技术

3.1 WIFI 探针功能及工作原理

WIFI 探针采用 WLAN(无线局域网)技术实现,对于开启或连接 WIFI 的设备进行数据采集。采集原理为:AP 周期性地向四周发送 BEACON(信标)帧,通知周围的 WIFI 设备(如手机、笔记本电脑等) AP 的存在;WIFI 设备也会周期的发送 PROBE(探测)帧,其中包含 WIFI 设备的 MAC 地址(Medium Access Control)、信号强度(RSSI)等信息。当 AP 检测到 WIFI 设备传送的 PROBE 帧,即记录 WIFI 设备传送的信息。因此,在 WIFI 探针区域内打开或连接 WIFI 则可收集 WIFI 设备信息。

在工作状态下,对于开启 WIFI 功能或连接上 WIFI 的电子设备,WIFI 探针能检测其发送的信号且记录生成数据文本,其中包括 WIFI 设备 MAC 地

址、RSSI 和数据抓取的日期和时间等基础信息。

3.2 WIFI 探针数据分析

将 WIFI 探针数据进行清洗,得到数据形式见表 4。

表 4 WIFI 探针采集数据表

Tab. 4 WIFI probe collection data sheet

WIFI_TIME	MAC	RSSI
14:40:00	38.29.5A.AD.32.03	-64
14:40:02	3C.A5.81.FA.F5.69	-44
...

其中,MAC 表示 WIFI 探针采集的 MAC 地址(即手机 MAC 地址);RSSI 代表 WIFI 探针探测到的该帧的信号强度,可用于表征源设备与 WIFI 探针间的距离(距离越近,信号强度值越趋近于 0);WIFI_TIME 指 WIFI 探针抓取智能设备 WIFI 信号的时间。

4 基于多源数据融合的地铁列车车厢满载率估计方法

4.1 长短时神经网络满载率预测模型

通过以上的研究,实现了 WIFI 探针数据的基本分析。为了准确可靠地实现轨道交通列车车厢客流的预测,需建立基于采集的客流数据与实际数据间的预测模型。本文利用模型实现车厢客流实时预测,获取短时车厢内客流的动态情况。

长短时记忆神经网络(LSTM)^[10]是循环卷积神经网络(RNN)中的一种,常用于解决一般循环卷积网络中处理长期依赖梯度弥散的问题。根据站内客流量具有短时性与短期依赖性的特点,本文基于循环神经网络的机理,建立了长短时记忆神经网络的客流预测模型。

4.2 实例验证

4.2.1 测试方案

实验测试是在上海轨道交通列车车厢内布置 WIFI 探针,使用 WIFI 探针采集列车运行过程中车厢中的客流量。由于列车在运行过程中,WIFI 探针会受到地面信号的干扰,为避免这种情况出现,选取地下车站进行测试。为此,选取上海轨道交通 11 号线上江苏路站-桃浦新村站为测试站点;由于列车车头及车尾处不易受到前后节车厢的干扰,因此选取列车车头处车厢及其邻近的两节车厢作为测试对象。测试共需 8 名工作人员参与:前两节车厢内分别为 3 个人,其中 2 人在同一节车厢的前后处,高举手持 WIFI 探针采集乘客手机的 MAC 地址,1 人在列车运行间隔的两站中通过人工计数统计当前车厢的客流量、拍照记录客流量及记录列车在每一站的停、发车时间;最后一节车厢内设置两人,在车厢的前后处高举手持 WIFI 探针,采集乘客手机的 MAC 地址。

4.2.2 源数据预处理

利用WIFI探针采集的源客流数据建立客流量

数据库,根据设计过滤的算法进行布置 4 探针点采集得到的源数据进行预处理。WIFI 探针 A 的源数据处理前后的对比如图 2 所示。

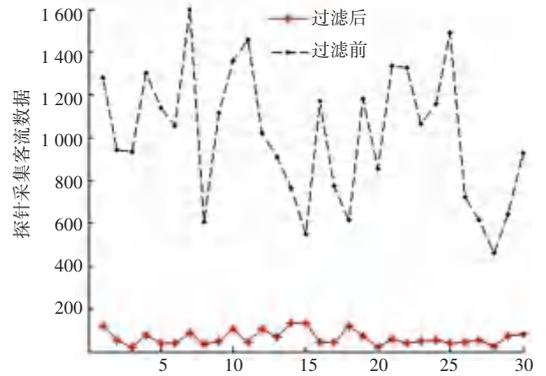


图 2 WIFI 探针 A 源数据过滤前后对比图

Fig. 2 Comparison of WIFI probe A source data before and after filtering

从图 2 中可见:过滤前,采集得到的客流数据量大,且曲线走势曲折,波动性较大,说明源数据并不符合实际客流数据的情况;过滤后,数据曲线在 0-200 之间波动且走势平缓,符合实际客流情况。由此说明了源数据预处理算法的有效性与必要性。同理,对其它 4 个探针进行统一操作得到预处理的数据结果见表 5。

表 5 多探针采集数据与实际客流量

Tab. 5 Multi-probe data collection and actual passengers flow

探针 a	探针 b	探针 c	探针 d	实际客流量
48	38	31	19	79
21	84	68	81	45
35	103	46	100	28
30	99	36	36	25
...

4.2.3 数据处理与分析

通过以上数据分析,采集客流数据的 80%作为网络的训练样本,20%实际客流量作为测试集,数据集样例见表 6。

表 6 数据集样例表

Tab. 6 Sample data set table

车站	车厢实际客流	列车发车时间	列车停站时间	WIFI 探针检测客流	断面客流
江苏路-隆德路	95	14:40:03	14:41:50	25	81
隆德路-曹杨路	98	14:42:36	14:44:00	21	84
曹杨路-枫桥路	120	14:45:13	14:46:28	22	90
枫桥路-真如	127	14:47:10	14:48:37	23	93
真如-上海西站	127	14:49:21	14:51:22	30	102
上海西站-李子园	125	14:51:57	14:53:45	22	62
李子园-祁连山路	115	14:54:29	14:56:22	28	56
祁连山路-武威路	110	14:57:00	14:58:45	22	40
武威路-桃浦新村	105	14:59:26	15:01:30	26	78

通过使用 WIFI 探针数据、当天同车次同一节车厢人工测试得到的实际客流,以及票务数据计算的满载率进行对比分析,如图 3 所示。

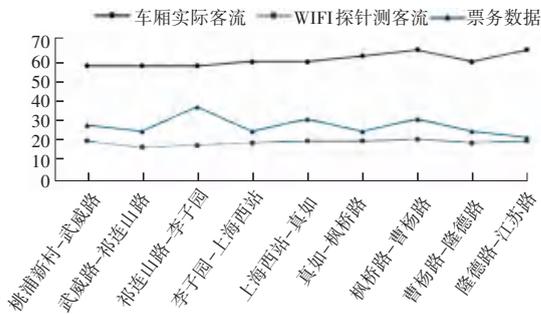


图 3 满载率对比分析图

Fig. 3 Full load rate comparative analysis chart

4.2.4 基于长短时神经网络客流量预测模型

本文的 LSTM 由输出层、隐含层与输出层组成。其中根据数据的特性设置,设置相应的初始参数见表 7。

表 7 LSTM 网络基本参数设置

Tab. 7 LSTM network basic parameter settings

参数	数值
最大训练次数	3 000
隐含层神经元数量	25
训练的目标误差	$1 * 10^{-6}$

将多源数据与实际客流量建立 LSTM 神经网络,在 python 环境中实现训练。神经网络训练过程如图 4 所示:

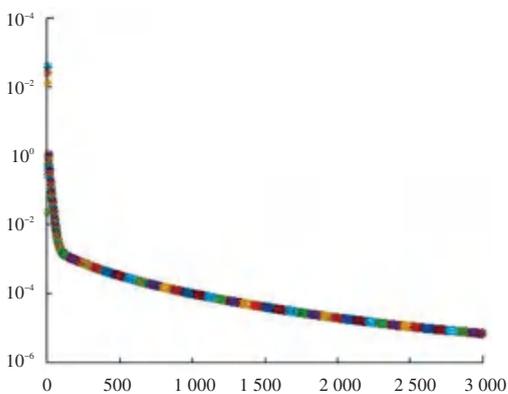


图 4 LSTM 训练误差曲线

Fig. 4 LSTM training error curve

从图 4 中可见,刚开始时,误差曲线下下降较快,在 500 次训练后,训练的数据明显减慢,在 3 000 次训练误差值基本接近预设定的误差值 $1 * 10^{-6}$ 。预测结果如图 5 所示。

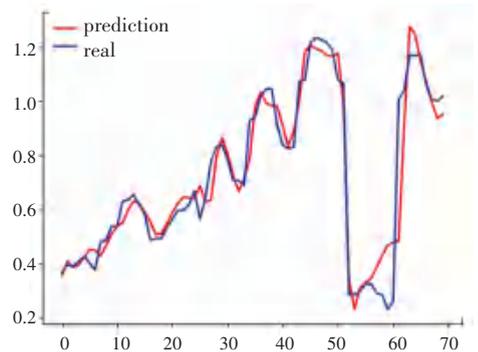


图 5 LSTM 模型预测结果图

Fig. 5 LSTM model prediction result

5 结束语

本文基于目前城市轨道交通车辆面临高峰客流而易产生拥挤等安全隐患及缺乏有效的客流量采集手段的背景下,引用新型的采集设备——WIFI 探针在车辆上进行客流数据的采集,详细说明了 WIFI 探针的原理及在轨道交通车厢内采集客流数据的布设方案。通过对比 WIFI 探针数据与同日同时段的票务清分数据,以及人工检测的实际客流数据,根据客流采集数据的特征,选择 LSTM 长短时记忆神经网络,建立客流数据的预测模型,数据的预测精度达到较高水平。

参考文献

- [1] 沈景炎. 关于车辆定员与拥挤度的探析[J]. 都市轨道交通, 2007(5): 14-18.
- [2] 刘欢,李文权. 城市公交调度中满载率问题的研究[J]. 交通运输工程与信息学报, 2008, 6(4): 104-109.
- [3] 张杏蔓,鲁工圆. 基于视频图像分析的地铁列车车辆拥挤度识别方法研究[J]. 交通运输工程,与信息学报, 2020, 18(3): 142-152.
- [4] 史芮嘉. 城市轨道交通系统输送能力利用率测算及优化研究[D]. 北京:北京交通大学, 2007.
- [5] 陈菁菁. 城市轨道交通客流检测技术的特征及其应用分析[J]. 城市轨道交通研究, 2018, 21(1): 137-142.
- [6] 曾异平. 数据挖掘方法研究:关联和趋势分析[D]. 成都:电子科技大学, 2003.
- [7] 黄庆. 大数据挖掘与数据处理方法[J]. 电脑迷, 2018(11): 89.
- [8] 祁友杰. 多源数据融合算法综述[J]. 航天电子对抗, 2017, 33(6): 37-41.
- [9] 赵传申,何顺刚,杨吉宏,等. 基于多分类-关联规则的数据流分类算法[J]. 计算机工程, 2010, 9: 38-40.
- [10] 王雁飞. 基于长短时记忆神经网络的异常行为识别研究[D]. 成都:四川师范大学, 2018.