

基于 Wi-Fi 探针的车站通道走行时间 实时估计与客流预警 *

王 越¹ 胡 华^{2**} 陈 娟³

(1. 苏交科集团股份有限公司, 210019, 南京; 2. 上海工程技术大学城市轨道交通学院, 201600, 上海;
3. 北京交通大学经济管理学院, 100044, 北京//第一作者, 助理工程师)

摘要 目的:为建立车站通道走行时间实时估计模型, 提高城市轨道交通车站的运营效率与安全性, 以提升乘客的出行体验。方法:采用基于 Wi-Fi 探针的数据采集方式, 通过在车站通道中安装 Wi-Fi 探针收集信息数据, 包括乘客设备 MAC(媒体接入控制)地址、信号强度、距离嗅探器的距离等信息。通过 java 语言对采集的数据进行初始处理, 利用 Mysql 和 navicat premium 数据库的组合对数据进行深度清洗, 验证了采用 BPR 美国联邦公路局函数建立车站通道走行时间实时估计模型的可行性, 并采用蚁群聚类算法将延误时间作为对车站通道客流预警等级划分的依据。结果及结论:总结了 Wi-Fi 探针客流采集原理和原始数据清洗方法, 建立基于 BPR 函数的车站通道走行时间实时估计模型, 实现了车站通道乘客走行时间实时估计, 其模型的准确率可达 92.8%。将车站通道客流预警划分为畅通、基本畅通、拥挤、严重拥挤等 4 个等级, 利用上海轨道交通 11 号线江苏路站的 Wi-Fi 探针数据进行了实例验证与分析, 证明了模型的可适用性及预测精度。

关键词 地铁车站; Wi-Fi 探针; 通道走行时间; BPR 函数
中图分类号 U293.13

DOI:10.16037/j.1007-869x.2023.07.008

Station Passage Walking Time Real-time Estimation and Passenger Flow Early-warning Based on Wi-Fi Probes

WANG Yue, HU Hua, CHEN Juan

Abstract Objective: Model for station passage walking time real-time estimation is established to improve urban rail transit station operation efficiency and safety, to enhance the passenger travel experience. Method: By adopting a data collection method based on Wi-Fi probe, information data such as passenger device MAC (medium access control) address, signal

strength, distance from sniffer are collected by Wi-Fi probes installed in station passage. The collected data are initially processed by java language, and further cleaned in-depth by using a combination of Mysql and navicat premium database, which verifies the feasibility of using BPR (bureau of public roads) function to establish a station passage walking time real-time estimation model, and the ant colony clustering algorithm is used to classify the station passage passenger flow early-warning level based on delay time. Result & Conclusion: The Wi-Fi probe passenger flow collection principle and original data cleaning method are summarized, station passage walking time real-time estimation model based on BPR function is established, and real-time estimation of passenger walking time through station passage is realized, the model for which reaches an accuracy rate of 92.8%. The passenger flow early-warning is classified into four levels: smooth, generally smooth, crowded and severely crowded. The model applicability and prediction accuracy is proved by the practical case verification and analysis of Shanghai Rail Transit Line 11 Jiangsu Road Station Wi-Fi probe data.

Key words metro station s; Wi-Fi probe; passage walking time; BPR function

First-author's address JSTI Group, 210019, Nanjing, China

随着我国的轨道交通线网规模和客流量的迅猛增长, 车站运营效率及安全性亟待提升, 而车站乘客流的交通特性复杂, 难以建立通用性人流交通模型。1958 年伦敦交通局汉金等人首先计算出地铁站乘客平均速度为 $1.6 \text{ m/s}^{[1]}$ 。文献 [2] 对站内通道乘客流的密度-流量和密度-速度进行了研究,

* 中央高校基本科研业务费专项资金资助项目(2021CZ106); 上海市科委地方院校能力建设项目(19030501400); 国家自然科学基金面上项目(52072235, 71971136)

** 通信作者

总结出通道乘客交通流的特性;文献[3]通过实地观测,构建了大型活动下室内通道行人流模型;文献[4]根据研究需要选取特定的时段进行了行人特性观测,总结出人流在平直通道中运动速度-密度曲线。上述研究多是通过视频记录和人工观测的方法获得客流相关数据,数据调查和提取过程重复且难以确保数据采集的一致性、准确性和普遍性,而对于走行时间的实时估计研究相对较少。目前,Wi-Fi 探针技术获取客流量数据日渐成熟,通过乘客的 AP(无线接入点)的 MAC(媒体接入控制)地址,从微观层面可实时获取乘客的位置,通过进一步数据挖掘后可得到地铁通道的客流量^[5]。

本文总结了 Wi-Fi 探针技术采集客流信息的原理和原始数据的清洗方法,在 Wi-Fi 探针技术实时监测客流量的基础上,研究了基于 BPR 函数的地铁车站通道乘客走行时间实时估计模型,并对车站通道客流做出预警等级划分,同时利用上海轨道交通 11 号线江苏路站的 Wi-Fi 数据进行了实例验证与分析。

1 Wi-Fi 探针数据获取及清洗

1.1 Wi-Fi 探针技术原理

Wi-Fi 探针技术是运用 WLAN(无线局域网)技术实现对开启或连接 Wi-Fi 的设备进行采集,即在无线接入的同时能够判定接入设备的位置。如图 1 所示,Wi-Fi 探针的客流数据采集原理为:探针高频率周期性地向四周广播发送 Beacon 帧(信标帧),通知附近的 Wi-Fi 设备(如手机、电脑)AP 的存在;Wi-Fi 设备也会周期性的地发送 probe 帧(探测帧),其中包含了设备的 MAC 地址、Rssi(信号强度)等信息^[5]。当 AP 接收到 Wi-Fi 设备发送的 probe 帧后,就获取了这个设备的 MAC 地址。因此,只要在 Wi-Fi 探针覆盖区域内的设备打开 Wi-Fi,探针就能收集到该设备的信息。

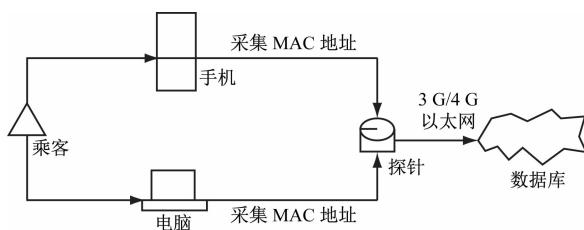


图 1 Wi-Fi 探针的客流数据采集原理

Fig. 1 Passenger flow data collection principle of Wi-Fi probe

1.2 Wi-Fi 探针数据的基本属性

Wi-Fi 探针通常会根据探针的探测半径及通道

长度以一定的间隔布设在通道的中间顶部位置,Wi-Fi 探针在获取到设备信息后可生成相应的数据文本(见表 1)。探针 ID(身份标识号),可用于区别数据由哪个探针设备采集到;由 Wi-Fi 设备的 MAC 地址(例如 9a:21:6a:7b:62:6a),可得知监测范围内开启 Wi-Fi 功能的无线终端数量;设备与 AP 布设位置的距离可通过 Range(手机距离探针的距离)这个指标来反馈,例如 14.0 m;另外数据的日期和时间也会储存为文本形式。

表 1 Wi-Fi 探针获取数据的主要属性

Tab. 1 Main attributes of data acquired by Wi-Fi probes

字段名	含义	类型	示例
MAC	采集到的设备 MAC 地址	文本型	9a:21:6a:7b:62:6a
Rssi	设备对于探针的信号强度	数值型	-30 dbm
Range	手机距离嗅探器的距离	数值型	14.0 m
ID	嗅探器设备 ID	文本型	0 010f377
Time	采集到 MAC 的时间	时间型	Sat Jun 04 22:45:28 2019

1.3 Wi-Fi 客流量数据处理及相关性分析

利用探针采集和统计客流量数据时,由于设备本身的采集特性,加之地铁车站采集环境、工作人员和乘客的不确定行为等都会影响探针采集客流数据的准确性,其主要问题和解决方案见表 2。

表 2 Wi-Fi 数据存在的问题及解决方案

Tab. 2 Problems with Wi-Fi data and solutions

问题	解决方案
在地面和高架站,可能采集到非地铁站内乘客的 Wi-Fi 设备数据	根据通道宽度以及 Range 值剔除距探针较远的 MAC 地址,排除地面站外等部分非通道内设备
采集到地铁站内工作人员的设备和车站内部固有设备	建立车站工作人员和固定设备 MAC 地址库,在输入时剔除
乘客在站内产生徘徊、滞留等行为	剔除相同且时间差大于等于 10 min 的 MAC 地址
多次重复采集相同终端设备的数据	对于单个探针或组合探针的数据最后进行 MAC 地址去重,且保留信号强度最强的一个
部分乘客可能没有携带 Wi-Fi 设备,携带设备的 Wi-Fi 功能没有打开或者携带多台设备	除了 MAC 地址外其他关键属性相同的 AP 保留一个

在站内,由于相邻探针的检测范围存在重叠性,因此探针客流量之间可能存在相关性,应在单个探针或组合探针数据清洗的基础上,对各个探针的客流量和实际客流量以及不同探针的客流量数

据进行相关性分析,从而获得能用于估计实际客流量的独立关联探针组。

常见的相关系数有皮尔逊相关系数、Gamma 相关系数、Kendall's Tau-b 相关系数、Spearman 相关系数等。其中皮尔逊相关系数又称积差相关系数,其相关系数的绝对值的大小表示集合的相关度的强弱。它反映了 2 个变量 x 与 y 之间的紧密程度,常用于线性相关分析。其值介于 -1 到 1 之间,绝对值越接近 1,则相关度越高;越接近 0,则相关度越低^[6]。经试验后发现,其适用于客流间的相关性分析^[7]。

2 通道走行时间实时估计模型与客流预警分级

2.1 通道内乘客流特性分析

行人流特性是指行人作为群体表现出的宏观特性。行人流特性的研究基本上是以行人流的 3 个参数即速度、密度及流量之间的关系研究为主。本文主要研究流量与时间之间的关系。为了反映拥挤状态下通道乘客量和走行时间的关系,选取高峰时段 Wi-Fi 探针清洗后的客流量数据(以下简称“Wi-Fi 客流量”)和乘客实际走行时间数据建立散点图,见图 2。

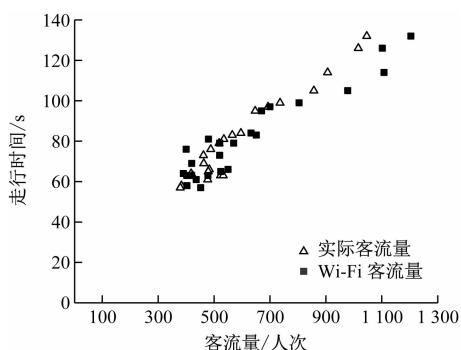


图 2 走行时间和客流量散点图

Fig. 2 Scatter plot of passenger walk time and flow volume

当客流量较少时,空间较大,乘客可采用较高的步行速度;当客流量增加时,由于密度增加,乘客可利用空间渐渐减少,致行走速度下降,而走行时间则随之增加;当乘客密度达到拥挤的临界水平,行走变得困难时,速度将进一步下降,走行时间变长。由图 2 可知,走行时间和 Wi-Fi 客流量有着明显的线性关系。因此可基于 Wi-Fi 客流量来估计乘客走行时间。

2.2 走行时间与 Wi-Fi 客流量关系模型

本文采用 BPR 模型来确定乘客在通道内的走

行时间。BPR 函数是 1966 年由美国联邦公路局基于大量调查数据提出的一种路阻函数,其用来研究行程时间随车流量的变化而变化的关系,其求解速度快,具有广泛适用性^[8]。将车流量引申为人流量即时间随人流量变化而变化的关系,建立 BPR 函数的数学模型:

$$t = t_0 \left[1 + \alpha \left(\frac{q}{c} \right)^\beta \right] \quad (1)$$

式中:

q —通道 Wi-Fi 探针测得的客流量;

t —通道 Wi-Fi 探针测得的客流量为 q 时乘客的走行时间;

t_0 —乘客以自由流状态通行时,通过通道所用的走行时间;

c —地铁车站通道的通行能力;

α, β —待定参数。

$$t_0 = \frac{s}{v_0} \quad (2)$$

式中:

v_0 —乘客在自由流状态下在地铁通道内的走行速度,取 1.6 m/s^[1];

s —地铁车站通道长度。

将式(1)两边取对数得:

$$\ln \left(\frac{t}{t_0} - 1 \right) = \ln \alpha + \beta \ln \left(\frac{q}{c} \right) \quad (3)$$

$$\text{令 } Y = \ln \left(\frac{t}{t_0} - 1 \right), X = \ln \left(\frac{q}{c} \right), m = \ln \alpha, n = \beta,$$

则可将原非线性关系转化为线性关系: $Y = m + nX$ 。利用上述由 Wi-Fi 探针获取的 q, t ,以及 t_0 与 c 可得到对应的 X, Y 值。使用 SPSS 分析软件的线性回归功能标定 m, n ,即可反推出 BPR 模型中的待定参数 α, β ,从而确定 t 值。

2.3 通道客流预警分级

地铁车站拥堵状态等级的划分与聚类思想一致,都是将同一属性相同或近似值的样本数据进行归纳的过程^[9]。在基于蚁群觅食行为的聚类分析中,将数据视为具有不同属性的蚂蚁,而将聚类结果视为食物源,所不同的是,此时认为存在多个食物源。这样各个蚂蚁通过一定的概率实现移动,并聚集在不同的食物源而实现聚类。

如图 3 所示,本文采用蚁群聚类算法,将地铁车站通道拥堵状态的等级划分问题转化为各指标值的聚类问题来解决,将拥挤度数据点视为具有若干

属性的蚂蚁,将拥挤度分级中心视为食物源,以地铁通道 Wi-Fi 探针测得的客流量数据样本进行聚类分析。通过实时估计的走行时间与自由流状态下的走行时间的差值得到延误时间作为对车站通道预警等级划分的依据。

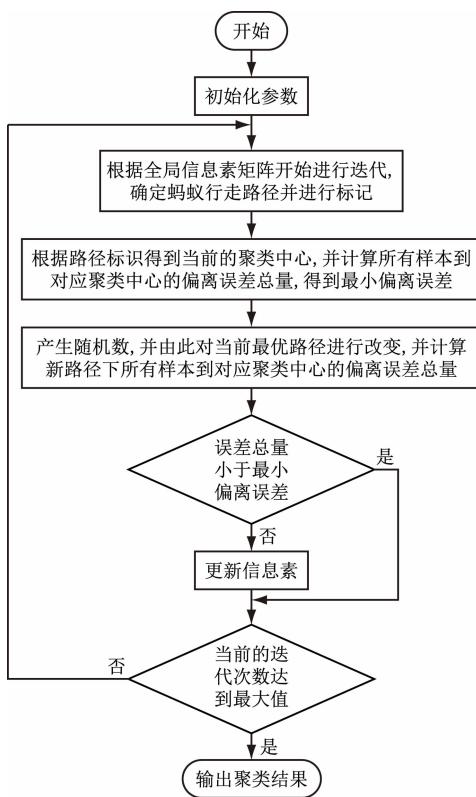


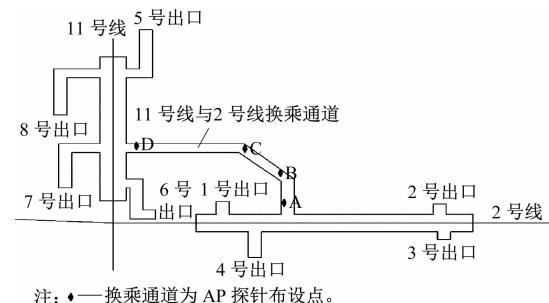
Fig. 3 Ant colony algorithm clustering process

3 实例验证

3.1 人流数据采集与分析

本文以上海轨道交通 11 号线江苏路站换乘通道为研究对象。江苏路站换乘通道的宽为 14.4 m, 长为 80 m, 自由流状态下的走行时间为 50 s。在地铁江苏路站换乘通道主要布设 4 个探针设备(见图 4), 分别为探针 A、B、C、D, 均布设在通道中间位置。选取 05:00—23:00 时间段, 按照 3 min 间隔采集的 Wi-Fi 探针测得的客流量为数据依据(共 340 个样本); 选取 DATASKY 品牌的 DS006 型嗅探设备, 调试好时间与储存模式; 测试人员到达指定位置获取人流数据。

根据清洗规则, 采用 java 语言调用整合关键标签得到客流量数据库; 使用 Mysql 和 navicat premium 数据库的组合根据上文清洗规则去除站内工作



注: ●—换乘通道为 AP 探针布设点。

图4 江苏路站换乘通道探针布设位置

Fig. 4 Probe placement locations of Jiangsu Road Station interchange passage

人员、长期逗留人员、非通道人员的数据。由于换乘通道的宽为 14.4 m, 因此选择 Range 值小于 15 m 和 8 m 的探针数据与实际客流量数据进行对比, 并对 A、B、C、D、AB、BC、ABC 分别进行处理。由于 A、D 位置处于通道两端, 所以误差较大。B 探针在通道中部, 所以效果较好。

通过对比发现, 当探针的 Range 值 ≤ 15 m 时没有 Range 值 ≤ 8 m 的效果好, 图 5 为探针 B 的 Range 值 ≤ 15 m 客流量与实际客流量对比, 图 6 为探针 B 的 Range 值 ≤ 8 m 客流量与实际客流量对比。当 B 探针的 Range 值 ≤ 15 m 时, 真实客流与实

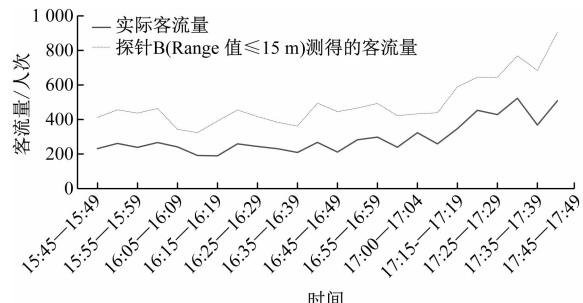


图5 探针B的Range值≤15m客流量与实际客流量对比

Fig. 5 Passenger flow of probe B Range ≤ 15 m and the actual passenger flow

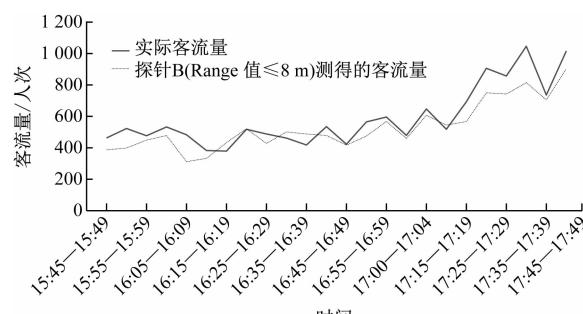


图6 探针B的Range值≤8m客流量与实际客流量对比

Fig. 6 Passenger flow of probe B Range ≤ 8 m and the actual passenger flow

际客流有着趋势上的相似,但是在量值上相差过大;当 Range 值 ≤ 8 m 时探针数据与实际数据较为吻合,但相似度为 93.3%,所以考虑将 AB 组合数据进行处理的效果会更好些,于是探究相邻探针的相关性,并将组合探针客流量数据进行数据处理。

根据上文所述皮尔逊相关性研究方法,分别对探针 A 与 B,A 与 AB,B 与 AB 做相关性分析,研究其相关程度。使用 SPSS 分析软件求得皮尔逊相关系数,分别为 0.92、0.88、0.87。故探针 AB 组合数据与 A、B 探针数据之间具有强相关性。

通过分析发现,探针 AB 组合的数据与实际客流量数据相似度达 97.3%,图 7 将探针 AB 组合的 Range 值 ≤ 8 m 客流量与实际客流量对比。由图 7 可知,与实际客流量相关性最高的是探针 AB 组合。选择 AB 组合 Range 值小于 8 m 的数据作为研究人流特性的数据,并进一步剔除该探针组合的 Wi-Fi 客流量数据。

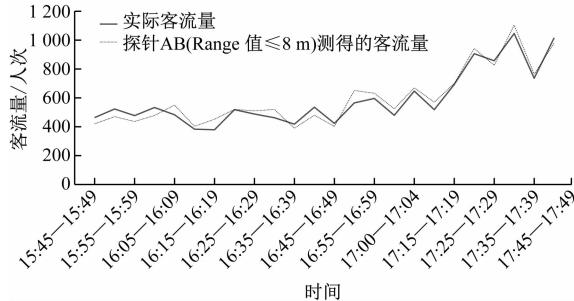


图 7 探针 AB 的 Range 值 ≤ 8 m 客流量与实际客流量对比

Fig. 7 Passenger flow of probe AB Range ≤ 8 m and the actual passenger flow

3.2 模型确定与验证

本研究选择 40% 的测试数据为模型代入数据。由图 8 可知: Y 的实际累积概率和预期大多数点落在对角线上,所以残差正态,故方程有意义;通过系数拟合,得到 m 和 n 的值。将 m 和 n 代入式(1),得到车站换乘通道内乘客流走行时间和流量的函数

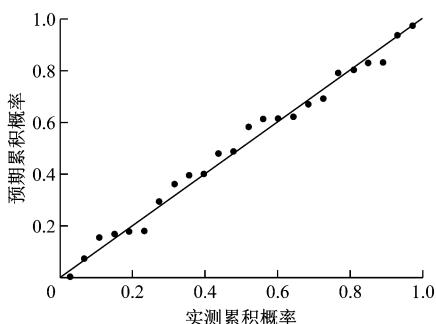


图 8 回归标准化残差的正态图

Fig. 8 Normal plot of regression-standardized residuals

关系为:

$$t = t_0 \left[1 + 2.683 \left(\frac{q}{c} \right)^{1.742} \right] \quad (4)$$

采用上海轨道交通 11 号线江苏路站换乘通道的实际数据与探针清洗数据后代入式(4),得到江苏路站换乘通道的走行时间实时估计模型(见式(5)),模型曲线如图 9 所示。

$$t = 50 + 3.844 316 599 \times 10^{-4} q^{1.742} \quad (5)$$

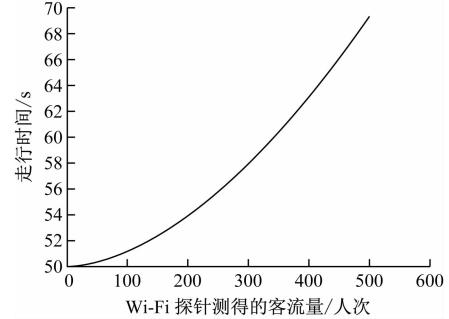


图 9 江苏路站换乘通道走行时间和客流量的模型曲线

Fig. 9 Model curves of passenger walk time and flow volume in the Jiangsu Road Station interchange passage

将江苏路地铁站实际走行时间测试数据的 60% 作为验证值代入式(5)得到走行时间估计值并进行对比,通过误差分析计算,该模型的准确率可达到 92.8%。Wi-Fi 探针测量的客流量可用于地铁通道走行时间的实时估计。当知道通道长度和通行能力以及在自由流状态下的走行时间,按 Wi-Fi 探针收集并按照清洗规则得到实时客流量数据,再根据 BPR 函数会得到通道相应的走行时间估计模型。该方法具有一定的适用性。

3.3 预警等级确定

将样本数据代入式(5),将得到通道各段的走行时间与自由流状态下的走行时间的求差值,继而得到各时间间隔内的乘客延误时间。以此进行聚类分级,划分为 4 个等级,即①畅通,②基本畅通,③拥挤,④严重拥挤。相应地,将算法中聚类中心数量设为 3,且将聚类中心从小到大排序为: L_1 、 L_2 、 L_3 。如图 10 所示得到的聚类结果为: $L_1 = 0.11$ min, $L_2 = 0.36$ min, $L_3 = 0.55$ min。

依据聚类分级的结果,通道内客流在早晚高峰期间出现了拥堵现象。根据数据分析及其分级结果,在 08:30 和 17:30 左右,通道内的客流状态达到极度拥堵状态。因此,工作人员应在该段时间及时做出响应,以确保人流有序进出站,防止客流过于拥挤而引发安全事故。

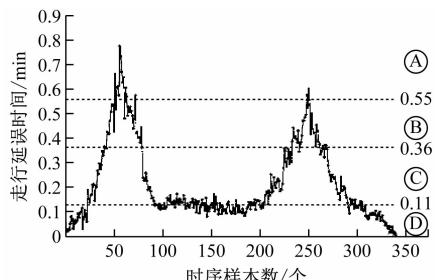


图 10 江苏路站换乘通道乘客走行延误时间聚类结果
Fig. 10 Clustering results for interchange passage passenger walk delay time in Jiangsu Road Station

4 结语

本文分析了 Wi-Fi 探针客流采集原理和原始数据清洗方法,在 Wi-Fi 探针测得的实时客流量的基础上研究了基于 BPR 函数的地铁车站通道乘客走行时间实时估计模型,并且对车站通道客流做出预警等级划分,并利用上海轨道交通 11 号线江苏路站的 Wi-Fi 数据进行了实例验证与分析。研究成果可为地铁车站通道行人交通建模、车站通道人流仿真等提供方法参考。在后期的研究中,对 Wi-Fi 数据的清洗挖掘还应做更进一步的深入研究,以便为客流特性研究以及客流组织提供更有效的依据。

参考文献

- [1] HANKIN B D, WRIGHT R A. Passenger flow in subways [J]. Journal of the Operational Research Society, 1958, 9(2): 81.
- [2] 娄永梅, 张炳森, 张宁, 等. 轨道交通车站内通道设施的行人流特性研究 [J]. 铁路通信信号工程技术, 2018, 15(9): 55.
LOU Yongmei, ZHANG Bingsen, ZHANG Ning, et al. Study on characteristics of pedestrian flow on passageway of rail transit stations [J]. Railway Signalling & Communication Engineering,

(上接第 43 页)

- MA Longxiang, LIU Weining, JIANG Yajun, et al. Metro train-induced vibration influences on surrounding environments based on sliced finite element-infinite element coupled model [J]. Journal of Vibration and Shock, 2017, 36(15): 111.
- [6] LEI S, GE Y, LI Q. Effect and its mechanism of spatial coherence of track irregularity on dynamic responses of railway vehicles [J]. Mechanical Systems and Signal Processing, 2020, 145: 106957.
- [7] 郑培治, 祖炳洁, 杨建福, 等. 不同轨道随机激励下的车辆动力学性能仿真研究 [J]. 价值工程, 2015, 34(36): 217.
ZHENG Peizhi, ZU Bingjie, YANG Jianfu, et al. Study of the vehicle dynamics performance under different rail stochastic excitations [J]. Value Engineering, 2015, 34(36): 217.
- [8] SALCHER P, ADAM C, KUISLE A. A stochastic view on the effect of random rail irregularities on railway bridge vibrations [J]. Structure and Infrastructure Engineering, 2019, 15

2018, 15(9): 55.

- [3] 史建港. 大型活动行人交通特性研究 [D]. 北京: 北京工业大学, 2007.
SHI Jian'gang. Study on traffic characteristics of large-scale activity pedestrians [D]. Beijing: Beijing University of Technology, 2007.
- [4] 单庆超. 城市轨道交通行人流动建模及仿真 [D]. 北京: 北京交通大学, 2010.
SHAN Qingchao. Modeling and simulation of pedestrian movement in urban rail transit [D]. Beijing: Beijing Jiaotong University, 2010.
- [5] 陈菁菁, 江志彬. 基于 WiFi 啼探数据的地铁网络客流分析技术 [J]. 城市轨道交通研究, 2018, 21(5): 153.
CHEN Jingjing, JIANG Zhibin. Passenger flow analysis technology in metro network based on WiFi probing data [J]. Urban Mass Transit, 2018, 21(5): 153.
- [6] 彭海. 皮尔逊相关系数应用于医学信号相关度测量 [J]. 电子世界, 2017(7): 163.
PENG Hai. Application of Pearson correlation coefficient in correlation measurement of medical signals [J]. Electronics World, 2017(7): 163.
- [7] 王玺铭. 基于时空相关性分析的短时交通流量预测方法研究 [D]. 杭州: 浙江工业大学, 2020.
WANG Ximing. Research on short-term traffic flow forecasting method based on spatial-temporal correlation analysis [D]. Hangzhou: Zhejiang University of Technology, 2020.
- [8] Transportation Research Board Business Office. Highway capacity manual 2008 [M]. Washington, D. C: National Research Council, 2009.
- [9] 周继彪, 赵鹏飞, 董升, 等. 基于蚁群算法的地铁车站行人拥挤等级划分方法 [J]. 城市交通, 2019, 17(4): 105.
ZHOU Jibiao, ZHAO Pengfei, DONG Sheng, et al. Pedestrian congestion levels at subway stations with ant colony algorithm [J]. Urban Transport of China, 2019, 17(4): 105.

(收稿日期:2021-02-09)

(12): 1649.

- [9] 万鹏, 翟婉明, 王开云. 考虑轮对弹性时车辆运动稳定性分析 [J]. 铁道车辆, 2008, 46(6): 8.
WAN Peng, ZHAI Wanming, WANG Kaiyun. Analysis of running stability of vehicles with the consideration of wheelset elasticity [J]. Rolling Stock, 2008, 46(6): 8.
- [10] 杨晓璇. A 型地铁车辆车轮多边形形成机理初步研究 [D]. 成都: 西南交通大学, 2018.
YANG Xiaoxuan. Preliminary study on mechanism of wheel polygon of type A metro vehicle [D]. Chengdu: Southwest Jiaotong University, 2018.
- [11] 翟婉明. 车辆-轨道耦合动力学: 上册 [M]. 4 版. 北京: 科学出版社, 2015.
ZHAI Wanming. Vehicle-track coupled dynamics: Volume 1 [M]. 4th ed. Beijing: Science Press, 2015.

(收稿日期:2021-01-18)