

城市轨道交通短时客流不确定性预测模型*

郭 旷¹ 王雪梅¹ 张 宁²

(1. 同济大学浙江学院, 314051, 嘉兴; 2. 东南大学教育部 ITS 工程研究中心, 210096, 南京//第一作者, 高级工程师)

摘 要 以城市轨道交通实际运营客流数据为基础, 针对现有短时客流预测存在的问题, 从运营时段特征、客流类型及站点周边用地类型等影响因素出发, 剖析了短时客流存在不确定性的原因; 基于周期性差分自动平滑回归模型和支持向量机理论, 构建了短时客流预测组合模型, 捕捉短时客流的周期性特征和局部非线性特征; 为提高短时客流预测结果的可信度, 引入广义自回归条件异方差模型来构建短时客流不确定性预测模型。通过实例, 验证结果表明, 周期性差分自动平滑回归-在线支持向量机组合模型对于周期性强且稳定的客流具有优越的预测性能, 广义自回归条件异方差模型的短期客流不确定性预测结果更为准确可靠。

关键词 城市轨道交通; 短时客流; 不确定性预测

中图分类号 U293.13

DOI:10.16037/j.1007-869x.2020.01.006

Uncertainty Prediction Model of Short-term Passenger Flow in Urban Rail Transit

GUO Kuang, WANG Xuemei, ZHANG Ning

Abstract Based on the actual data of passenger flow collected from urban rail transit operation, and aiming at the existing problems in short-term passenger flow prediction, the causes of uncertainty in short-term passenger flow is analyzed according to the influencing factors such as operation period characteristics, passenger flow and land use types around stations. Based on the SARIMA (autoregressive integrated moving average) model and the SVM (support vector machine) theory, a prediction model of short-term passenger flow is constructed to capture the periodicity and local non-linearity characteristics of short-term passenger flow. To improve the reliability of the of short-term passenger flow, GARCH (autoregressive conditional heteroskedasticity) model is introduced to construct the uncertainty prediction model. Through actual case analysis, the testified results show that the SARIMA-OLSVM model has superior prediction performance for the strong periodical and stable passenger flow, and the GARCH model is more accurate and reliable for the short-term passenger flow uncertainty pre-

diction.

Key words urban rail transit; short-term passenger flow; uncertainty prediction

First-author's address Zhejiang College, Tongji University, 314051, Jiaxing, China

城市轨道交通短时客流在受到多种随机因素的影响下, 具有显著的随机性和非线性特点。根据城市轨道交通短时客流的随机特性, 城市轨道交通短时客流状态可分为平均状态和不确定性状态。其中, 短时客流不确定性为在统计时段内衡量客流平均状态的波动性, 其反映了短时客流时间序列中无法预期的部分对客流产生的影响。

目前, 城市轨道交通短时客流不确定性预测尚处在起步阶段, 相关研究不多。文献[1-6]均基于平均状态的乘客交通流运营历史数据进行建模和预测。但目前, 对基于不确定性状态数据的建模和预测仍不成熟, 尚未形成完整的城市轨道交通短时客流不确定性预测的流程。文献[7-8]分析了不同因素对客流不确定性的影响, 其研究表明, ARIMA(自回归求和移动平均)-GARCH(广义自回归条件异方差)模型和 ARIMA-SV(随机波动率)模型均能较好地拟合客流波动情况。文献[10-15]基于各种模型研究了道路交通流的不确定性。城市轨道交通乘客交通流与道路交通流虽不尽相同^[9], 但就随机不确定性方面存在相似之处。

为了更准确地描述轨道交通客流状态, 本文在国内外现有短时客流预测的研究成果基础之上, 以城市轨道交通 AFC(自动售检票)系统实际数据为依据, 充分探讨客流的影响因素, 构建城市轨道交通短时客流预测模型及表征其随机不确定性特征的预测模型, 并对短时客流不确定性预测过程中的核心模型进行分析和优选。

* 浙江省教育厅纵向科研项目(Y201534856)

1 短时客流不确定性预测流程

短时客流不确定性预测流程如图 1 所示。图 1 中:数据采集环节主要获取运营客流数据;在短时客流不确定性预测环节,着重研究短时客流不确定性影响因素、短时客流预测及其不确定性预测模型;验证成果环节则着重评估短时客流预测的准确性和可靠性。

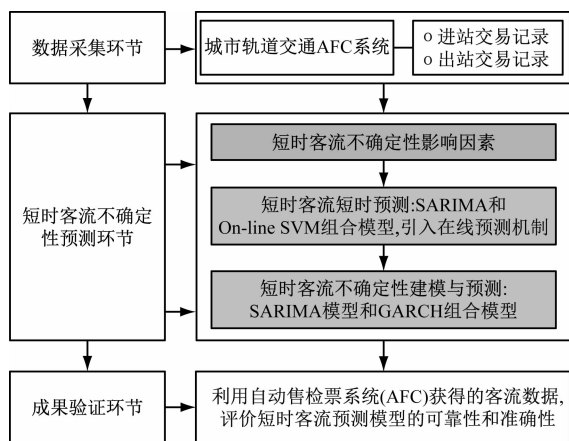


图 1 短时客流不确定性预测流程

2 短时客流不确定性预测模型的构建

2.1 短时客流不确定性影响因素

客流的影响因素极为复杂,一直是研究客流不确定性的最大障碍。由图 1 可知,短时客流不确定性影响因素是有效构建整个短时客流预测理论的基础。本文从城市轨道交通历史运营客流数据等资料出发,对运营时段(如工作日、双休日、节假日、平峰时段及高峰时段等)、客流类型(按持有票种类型分类,如普通单程票、普通储值票、员工票、学生票、纪念票、老人票及团体票等)、站点附近用地类型(如大型交通枢纽、商业中心、办公或住宅等)等多个影响因素进行分析,探讨各因素对短时客流不确定性的影响程度与内在联系,剖析短时客流不确定性影响因素。

2.2 短时客流预测一般模型

时间序列模型是短时客流预测模型中最广泛的应用之一,对于稳定的客流状态尤为适用。其中 SARIMA 模型(周期性差分自动平滑回归模型)就能较好地描述客流数据的固有周期性特征。但短时客流预测非线性特点显著,一般时间序列模型无法实现对其局部特征的实时捕捉。

在标准支持向量机模型中,根据输入输出样本集 $(x_i, y_i) (i = 1, 2, \dots, m)$,其中 $x_i \in X \subseteq R^n, y_i \in Y \subseteq R^n, m$ 是训练样本数,可使用函数 $\Phi(x_i)$ 将低维输入空间的非线性回归推广到高维特征空间的线性回归。具体有:

$$f(x) = w^T \Phi(x_i) + b \quad (1)$$

式中:

$f(x)$ ——样本 (x_i, y_i) 根据模型(1)输出的预测值,

w ——特征空间 F 的向量,

$\Phi(x_i)$ ——输入空间 x 推广到向量的函数,

b ——参数。

标准支持向量机虽然具有推广性好,预测精度高等优点,但其本身是离线模型。随着时间的推移,预测精度会逐渐下降,如遇客流突发状况,则预测精度更会明显降低。若将在线学习方法引入支持向量机理论,训练连续更新的实时客流数据,获得实时的预测函数,再来预测客流数据,则可提升短时客流预测模型的精度。

因此,本文提出基于短时客流不确定性影响因素分析,同时考虑到 SARIMA 模型能很好地捕捉时间序列的固有周期性特征,提取 SARIMA 模型的自变量为输入特征,运用到 On-line SVM 模型(在线支持向量机模型)中,通过 SARIMA 和 On-line SVM 组合模型(以下简为“OLSVM 模型”)来进行短时客流预测,描述客流的周期性及非线性特点。

2.3 短时客流不确定性预测模型

城市轨道交通短时客流状态分为平均状态和不确定性状态。现有的短时客流预测模型通常仅对乘客交通流平均状态进行预测,其预测的结果的可信度较低。

由于不确定性状态可用二阶随机波动特征的时间序列来表示,而具有二阶随机波动特征的时间序列可通过 GARCH 模型(广义自回归条件异方差模型)进行描述,因此,可采用 GARCH 模型来构造短时客流的不确定性模型,并预测短时客流。

GARCH 模型通过计算预测结果的方差,来计算对预测准确度的影响,认为过去的实际观测数值与数值误差的平方决定了条件方差,其表达式为:

$$\varepsilon_t = \sqrt{h_t} e_t \quad (2)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (3)$$

式中:

α_o ——常数项, $\alpha_o > 0$;
 e_t —— $e_t \sim \text{IIDN}(0, 1)$, 并且对所有的 t, e_t 与 $\{\varepsilon_i\}$ 独立;
 h_t —— t 时的条件方差, 即 $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$, 其中, Ψ_{t-1} 为 $t-1$ 时的条件信息;
 α_i ——滞后样本方差 ε_{t-i}^2 的系数, $\alpha_i \geq 0$;
 β_j ——滞后条件方差 h_{t-j} 的系数, $\beta_j \geq 0$ 。

3 实例分析

3.1 短时客流不确定性影响因素分析

本文以南京地铁 1 号线 3 个车站为例, 选择 2016 年 2 月 29 日至 2016 年 4 月 3 日期间的进站客流数据进行短视客流不确定性预测验证。将进站客流按照 15min 来分段统计客流量。由于南京地铁的日常运营时间为 6:00—23:00, 可知每个站点有 2 380 个样本数据。本研究中:运营时间因素选取 2 种不同特征日(工作日和双休日);站点按周边土地性质不同, 分别选取珠江路站(其周边为办公区和居住区)、新街口站(其周边为商业中心)、南京南站(其周边为大型交通枢纽);客流类型按一卡通与一票通等 2 种不同的票卡划分。

不同站点不同运营时间的短时客流时间分布曲线如图 2 所示。由图 2 可见:双休日客流较工作日客流的波动性强、非线性特征明显;珠江路站、新街口站和南京南站的客流量曲线各不相同, 其中南京南站和新街口站的客流波动性更大。

不同客流类型短时客流时间分布曲线如图 3 所示。由图 3 可见, 不同类型的客流时间分布存在较大差异, 其中持一卡通出行的客流曲线有明显的早晚高峰, 其出行时间较集中, 客流相对稳定, 而持一票通出行的客流曲线波动性较强, 出行时间较分散, 具有较强的不确定性。

3.2 基于 OLSVM 模型的短时客流预测结果评价

本文以 SARIMA 模型为基准, 选取预测客流量的绝对平均误差 E_{MA} 、绝对平均百分比误差 E_{MAP} 、均方根误差 E_{RMS} 作为 OLSVM 模型预测结果评价指标, 选取 3 种不同站点及不同运营时间的数据进行分析预测, 得到预测结果评价指标如表 1 所示。

由表 1 不难发现, 工作日客流预测精度较双休日高。这是由于双休日客流中随机客流占比较大, 其所表现出的波动性强、非线性明显等客流特征对模型的预测精度产生了影响。由表 1 还可看出, 珠

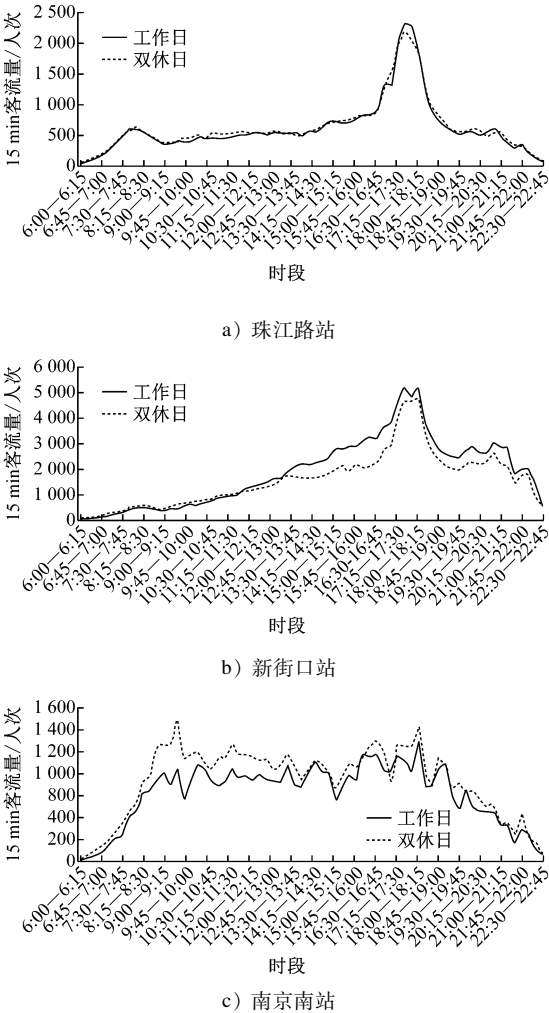


图 2 不同站点不同运营时间的短时客流时间分布曲线

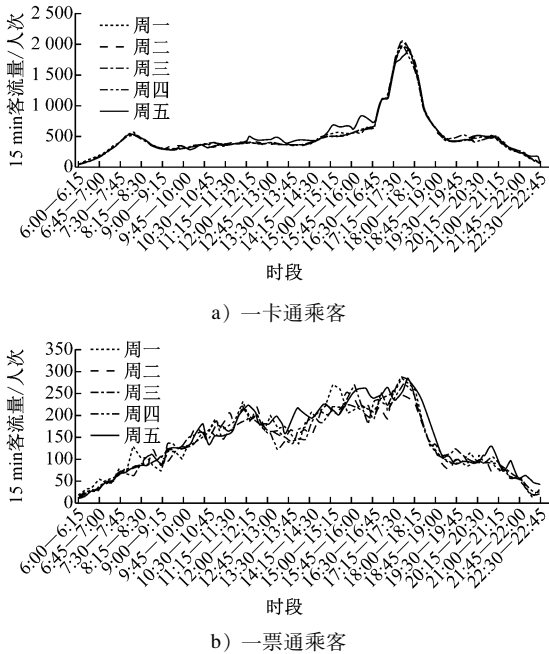


图 3 不同客流类型短时客流时间分布曲线

江路站(周边为办公区和居住区)客流预测精度最高,新街口站(周边为商业中心)和南京南站(周边为大型交通枢纽)相对较差,这是由于珠江路站点客流以通勤客流为主,且持一卡通出行的客流比例高,客流相对稳定,故预测模型能较容易地捕捉其规律。

由表 1 的评价结果, OLSVM 模型的预测性能整体优于 SARIMA 模型,尤其 OLSVM 模型对于双

休日客流的预测精度较高,对新街口站(周边为商业中心)和南京南站(周边为大型交通枢纽)的客流预测更准确。这说明 OLSVM 模型能很好地捕捉客流的局部非线性特征,实时地适应客流的变化。值得一提的是,在珠江路站的工作日客流预测中,SARIMA 模型表现出了最优的性能。这说明 SARIMA 模型对于周期性强且稳定的客流具有优越的预测性能。

表 1 SARIMA 和 OLSVM 两种模型的短时客流预测结果评价

站点名称	特征日	OLSVM 模型预测结果			SARIMA 模型预测结果		
		$E_{MSE}/\text{人次}$	$E_{MA}/\text{人次}$	$E_{MAP}/\%$	$E_{MSE}/\text{人次}$	$E_{MA}/\text{人次}$	$E_{MAP}/\%$
珠江路站	工作日	45.0	34.5	6.8	44.3	33.9	6.4
	双休日	65.2	53.3	14.3	157.5	132.2	30.4
新街口站	工作日	96.8	63.6	9.8	106.2	76.3	10.5
	双休日	275.7	259.8	11.3	562.5	449.4	22.9
南京南站	工作日	196.8	113.6	10.6	211.1	122.9	11.5
	双休日	248.7	203.8	11.1	278.9	223.5	20.2

3.3 基于 GARCH 模型的短时客流不确定性预测结果评价

本文采用 SARIMA 模型、SARIMA-GARCH 组合模型(以下简称 GARCH 模型)分别对 3 种不同站点的客流进行可靠性分析。短时客流预测的可信度可根据置信区间宽度 D_{CI} 、无效覆盖量 Q_{KN} 和无效覆盖率 R_{KP} 等 3 个指标进行评估,并以 R_{PI} 表示 GARCH 模型动态 D_{CI} 值与 SARIMA 模型固定 CI 值的比例。3 种不同类型客流的预测结果评价如表 2 所示。

表 2 SARIMA 模型和 GARCH 模型短时预测可靠性评估

站点名称	SARIMA 模型预测结果		GARCH 模型预测结果		$R_{PI}/\%$
	Q_{KN}	$R_{KP}/\%$	Q_{KN}	$R_{KP}/\%$	
珠江路站	13	3.73	16	4.65	72.2
新街口站	21	6.28	18	5.19	73.5
南京南站	28	8.14	25	7.23	63.4

注: $R_{PI} = \text{GARCH 模型的动态 } D_{CI} \text{ 值} / \text{SARIMA 模型的固定 } D_{CI} \text{ 值} \times 100\%$

由表 2 可知:① 由于设定的置信水平为 95%,故 R_{KP} 的值越接近 5%,则模型预测的可靠性越高;由于 GARCH 模型的 R_{KP} 值更接近于 5%,所以其预测可靠性更优。② 从表 2 中的 R_{PI} 值可知,GARCH 模型预测的 D_{CI} 值基本均小于 SARIMA 模型的固定 D_{CI} 值。

可见,相比于 SARIMA 模型,GARCH 模型在保证 Q_{KN} 和 R_{KP} 的基础上,能获得较多更窄的预测置信区间,也就是说,GARCH 模型短期客流不确定性预测结果更为准确可靠。

4 结语

短时客流是反映城市轨道交通服务水平和系统运行状态评价的重要指标,能反映客流实时的变化规律。根据可靠的短时客流预测结果,可提前制定或调整运营计划(如列车调度、发车间隔调整和工作人员排班等),以适应客流的波动,并为决策者提供关键依据。客流预测值一旦大于预设的临界值,决策者可提前制定调整计划,提前安排工作人员疏导拥挤客流,以防紧急事件如踩踏事故等的发生。

本文以乘客实际进出站数据为基础,根据短时客流的不确定性影响因素,构建了城市轨道交通短时客流不确定性预测模型,并对预测过程中的模型选择进行探讨。

实例验证结果表明,SARIMA 模型对于周期性强且稳定的客流具有优越的预测性能,OLSVM 模型可以捕捉客流的局部非线性特征,GARCH 模型的短期客流不确定性预测结果更为准确可靠。

参考文献

- [1] HUANG Y C, PAN H. Short-term prediction of railway passenger flow based on RBF neural Network[J]. The Fourth International Joint Conference on Computational Sciences and Optimization, 2011(4):594.
- [2] 姜桂艳, 常安德, 牛世峰, 等. 基于 BP 神经网络的交通数据序列动态可预测性分析方法[J]. 北京工业大学学报(自然科学版), 2011(7):1019.
- [3] 冯金巧, 杨兆升, 孙占全, 等. 基于小波分析的交通参数组合预测方法[J]. 吉林大学学报(工学版), 2010, 40(5):1220.
- [4] CHEN Q, LI W Q, ZHAO J H. The use of LS-SVM for short-term passenger flow prediction [J]. Transport, 2011, 26(1):5.
- [5] WANG X M, ZHANG N, ZHANG Y L, et al. Forecasting of Short-Term Metro Ridership with Support Vector Machine Online Model [J]. Journal of Advanced Transportation, 2018(1678):1.
- [6] WEI Y, CHEN M C. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks[J]. Transportation Research Part C Emerging Technologies, 2012, 21(1):148.
- [7] 徐世鹏. 城市轨道交通站点客流不确定性机理及预测研究[D]. 南京:东南大学, 2016.
- [8] 徐世鹏, 张宁, 邵星杰. 城市轨道交通站点客流不确定性研究[J]. 都市轨道交通, 2015, 28(3):12.
- [9] CHEN M C, WEI Y. Exploring time variants for short-term passenger flow[J]. Journal of Transport Geography, 2011, 19(4):488.
- [10] KAMARIANAKIS Y, KANAS A, PRASTACOS P. Modeling traffic volatility dynamics in an urban network[J]. Transportation Research Record Journal of the Transportation Research Board, 2006, 1923(1):18.
- [11] SOHN K, KIM D. Statistical model for forecasting link travel time variability [J]. Journal of Transportation Engineering, 2009, 135(7):440.
- [12] TSEKERIS T, STATHOPOULOS A. Real-time traffic volatility forecasting in urban arterial networks[J]. Transportation Research Record Journal of the Transportation Research Board, 1964(1):146.
- [13] GUO J, WILLIAMS B M. Real-time short-term traffic speed level forecasting and uncertainty quantification using layered kalman filters[J]. Transportation Research Record Journal of the Transportation Research Board, 2010(1):28.
- [14] TSEKERIS T, STATHOPOULOS A. Short-term prediction of urban traffic variability: stochastic volatility modeling approach [J]. Journal of Transportation Engineering, 2010, 136(7):606.
- [15] MARTCHOUK M, MANNERING F, BULLOCK D. Analysis of freeway travel time variability using bluetooth detection[J]. Journal of Transportation Engineering, 2011, 137(10):697.
- [16] VLADIMIR N, VAPNIK. The nature of statistical learning theory[J]. Technometrics, 1996, 8(4):1564.

(收稿日期:2018-05-18)

(上接第 17 页)

较长。

3) 由于建筑物的存在,当偏离比逐渐减小时,地表沉降槽宽度和地层滑移角逐渐增加,沉降曲线呈非对称分布。

4) 在地铁隧道近距离穿越筏板基础过程中,当 $L/D \geq 1.5$ 时,盾构法施工隧道与建筑物之间距离为安全距离。这是由于筏板基础自身刚度较大,能对控制最大沉降差起到较好调节作用。

参考文献

- [1] 钱七虎. 迎接我国城市地下空间开发高潮[J]. 岩土工程学报, 1998(1):112.
- [2] 任建喜, 杨锋, 贺小俐, 等. 地铁隧道暗挖施工引起的桥桩基础

变形规律与控制技术[J]. 城市轨道交通研究, 2016(9):110.

- [3] 姜忻良, 贾勇, 赵保建, 等. 地铁隧道施工对邻近建筑物影响的研究[J]. 岩土力学, 2008(11):3047.
- [4] 李超人, 周传波, 陈东, 等. 双线盾构施工引起地表及建筑物沉降规律研究[J]. 城市轨道交通研究, 2017(6):108.
- [5] 魏新江, 魏纲, 丁智. 盾构施工与邻近不同位置建筑物相互影响分析[J]. 岩土力学, 2007, 28(增刊1):505.
- [6] 李涛, 陈慧娴, 刘波, 等. 双线隧道盾构施工对临近高层建筑物的影响分析[J]. 湖南科技大学学报(自然科学版), 2013, 28(4):43.
- [7] 中华人民共和国建设部, 中华人民共和国国家质量监督检验检疫总局. 建筑地基基础设计规范: GB 50007—2011[S]. 北京: 中国建筑工业出版社, 2012.

(收稿日期:2018-04-27)