

# 基于深度学习的城市轨道交通 短时客流起讫点预测\*

侯晓云<sup>1</sup> 邵丽萍<sup>1</sup> 李 静<sup>1\*\*</sup> 黄 磊<sup>1</sup> 李雪岩<sup>2</sup>

(1. 北京交通大学经济管理学院, 100044, 北京; 2. 北京联合大学管理学院, 100101, 北京 // 第一作者, 硕士研究生)

**摘 要** 提出了一种基于门控循环单元(GRU)神经网络的城市轨道交通短时客流 OD(起讫点)预估模型。以实际数据为例, 引入同期天气数据, 对工作日的数据进行训练预测, 并与长短期记忆(LSTM)神经网络模型进行对比。预测结果表明: 相对于 LSTM 模型, GRU 模型不仅模型简单、收敛速度明显较快, 而且在预估误差和预测稳定性等方面也略优, 更适于短时客流 OD 的快速预测。

**关键词** 城市轨道交通; 短时客流起讫点预测; 门控循环单元神经网络; 长短期记忆神经网络

**中图分类号** U293.13

**DOI:**10.16037/j.1007-869x.2020.01.013

## Urban Rail Transit Short-time Passenger Flow OD Forecasting Based on Deep Learning Modeling

HOU Xiaoyun, SHAO Liping, LI Jing, HUANG Lei, LI Xueyan

**Abstract** A short-term passenger origin-destination (OD) forecasting model for urban rail transit based on gated recurrent unit (GRU) neural network is proposed. Based on the practical data, by importing the weather data during the same period, the training prediction of the working day data is conducted and compared with the long short-term memory (LSTM) neural network model. The results show that the convergence speed of GRU is obviously faster than LSTM, the prediction errors and stability are slightly better than LSTM. Therefore, GRU model is more suitable for short-term passenger flow OD prediction.

**Key words** urban rail transit; short-term passenger flow OD forecasting; GRU neural network; LSTM neural network

**First-author's address** School of Economics and Management, Beijing Jiaotong University, 100044, Beijing, China

在通勤出行的高峰时段, 一旦地铁发生故障导致停运, 就会有大量地下客流涌上地面。快速有序疏散人群是保证乘客安全和社会安定的关键。根据城市轨道交通路网短时客流 OD(起讫点)历史数据, 结合出行外部因素, 可以预测乘客出行终点, 为快速制定应急响应方案提供科学依据, 为城市轨道交通路网的日常运营和调度提供参考。

短时客流 OD 可以反映在一个较短时间范围内, 客流需求的时空分布特征<sup>[1]</sup>。目前, 对于城市轨道交通短时客流 OD 估计的研究仍在起步阶段。OD 估计问题实质上是一个时间序列的预测问题, 主要采用参数优化模型和迭代估计模型<sup>[2]</sup>。其中以卡尔曼滤波递推算算法<sup>[1,3]</sup>为基础的迭代估计模型多应用于在线客流 OD 估计中, 但是卡尔曼滤波递推算算法每次计算时需要面临大量的矩阵和向量运算, 存在运算效率方面的不足。随着深度学习的发展, 文献[4]利用前3个时间段的 OD 量基于 LSTM(长短期记忆网络)预估当前时段的客流 OD 量, 并用粒子滤波算法求解客流 OD 递归贝叶斯估计问题, 进一步对预估客流量进行校正。上述对于客流 OD 的预测大都基于历史的客流 OD 进行递推, 但是在现实生活中, 客流 OD 分布与出行外部因素如天气状况等有着密切关系, 天气直接影响着市民的出行选择, 文献[5]对站点进站客流量与同期天气进行相关性分析, 结果显示相关性显著。而现有文献大都忽略了出行外部因素对客流 OD 的影响。因此现有的客流 OD 动态估计在准确性方面仍具有进一步提升的空间。

本文提出了一种基于 GRU(门控循环单元)神经网络的城市轨道交通短时客流 OD 预测方法, 并

\* 国家自然科学基金“青年基金”项目(71103014); 国家级大学生创新创业训练计划项目(170140032); 北京市哲社办课题(14JGC095); 北京市交通委员会科技课题(B17M00080); 北京市交通行业科技课题(201905-ZHJC2)

\*\* 通信作者

利用上海市 2016 年 3 月份交通 IC 卡刷卡数据,结合同期天气数据进行训练预测,并与 LSTM 方法在收敛速度、误差以及稳定性等方面进行了对比,验证 GRU 神经网络在短时快速预测方面的优异性。

1 短时客流 OD 预测模型构建

1.1 神经网络

递归神经网络(RNN)因为可以连接先前的信息到当前的任务上,所以非常适用于处理时间序列<sup>[6-7]</sup>。但是 RNN 训练时会随机在每个时间步长中重写其短期记忆,且易于产生梯度消失的问题<sup>[8]</sup>。为了使模型能够学习长期信息,文献[9]提出了 LSTM 的概念。不同于重复模块中仅有一层 tanh 的 RNN,LSTM 的每个重复模块中有 4 个互相交互的层,为一层 tanh 和三种门。RNN 重复模块和标准 LSTM 重复模块分别如图 1 和图 2 所示。

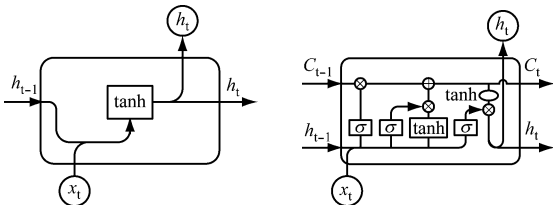


图1 RNN 重复模块 图2 标准 LSTM 重复模块

在实际的使用中,LSTM 有很多种变体。其中一种改动比较大的变体是 GRU<sup>[10]</sup>。GRU 把忘记门和输入门合二为一,把细胞状态 C 和隐藏状态 h 混合。GRU 模型比标准 LSTM 更加简单,其重复单模块如图 3 所示。

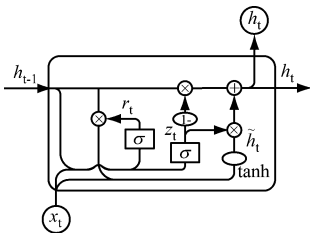


图3 GRU 重复模块

1.2 模型构建

1.2.1 相关变量的定义

定义城市轨道交通网络中共有  $N$  个站点,将一天划分成  $M$  个时间段, $O_{k,i}$  表示第  $k$  时间段  $i$  站点的进站客流量 ( $i = 1, 2, 3, \dots, N; k = 1, 2, 3, \dots, M$ ),  $D_{k,i}(j)$  表示第  $k$  时间段在  $i$  站点进站且在  $j$  站点出站的客流量,  $w_k$  表示第  $k$  时间段的天气数据。其中,  $O_{k,i} = \sum D_{k,i}(j)$ 。

1.2.2 构建神经网络

根据上述模型原理,根据实践经验,本文构建了十层网络。为了对抗过拟合,在每一层间加入了 Dropout 层<sup>[11]</sup> 和 Batch Normalization 层,后者能够有效地对抗内协变量移位(Internal Covariate Shift)问题。选用 PReLU (带参数的 ReLU) 作为激活函数。在预测模型构建中,设置步长为  $s$ ,以得到的  $k$  时段客流量作为输入变量,去预测  $k + 1$  时段的客流量,依此循环往复,滚动预测直到一天结束。

2 数据预处理与分析

2.1 数据预处理

本文的交通 IC 卡数据来源为上海市 SODA (开放数据创新应用) 比赛开放的 2016 年 3 月份数据,每天的数据量大约 800 MB,数据记录大约为 1 000 万条。源数据内容如表 1 所示。

表 1 中,进站时的票价为 0,出站时的票价为实际消费金额,因此通过票价即可判断乘客进出站状态。在进行数据分析之前,应先将源数据按卡号、进出站时间及票价进行 OD 配对处理。

在 OD 配对处理过程中发现两种数据异常情况:① 只有进站数据,没有出站数据;② 只有出站数据,没有进站数据。每天的异常数据约有 50 万条,仅占总记录的 2%。故本文剔除无法配对的数据异常。

本文所用天气数据来源于美国国家海洋和大气管理局官方网站,包含风向、风速、温度、露点温度及气压等信息,记录间隔为 30 min。鉴于天气数据集中存在一定量的缺失值,本文选取其中的风速、温度及露点温度等重要字段较为完整的数据,

表 1 2016 年 3 月份上海市交通 IC 卡数据摘录

卡号	日期	时间	线路	车站	票价/元	计价方式
3102664781	2016-03-01	22:03:05	3 号线	曹杨路站	4	非优惠
3102664781	2016-03-01	11:38:03	3 号线	虹口足球场站	3	优惠
3102664781	2016-03-01	10:51:52	11 号线	枫桥路站	0	非优惠
3102664781	2016-03-01	21:43:07	3 号线	虹口足球场站	0	非优惠
602141128	2016-03-01	08:35:04	1 号线	莘庄站	0	非优惠

对缺失的数据使用拉格朗日插值法进行补填。最后,采用相邻数据取均值的方法,将天气数据的时间间隔处理成 15 min,与客流量数据相匹配。

## 2.2 数据分析

本文选取上海地铁人民广场站作为研究站点。根据交通 IC 卡数据情况,选取 5:30—23:00 为每日研究时段,以 15 min 为时间段,将每日的数据划分成 70 个时间段。对其一周内的人民广场站(起点)进站客流量进行分析,结果如图 4 所示。由图 4 可见,工作日与周末的进站客流量存在明显差异。基于上述分析,本文在研究客流 OD 的时候,把工作日与周末区分开来,并以工作日为例进行详细分析。

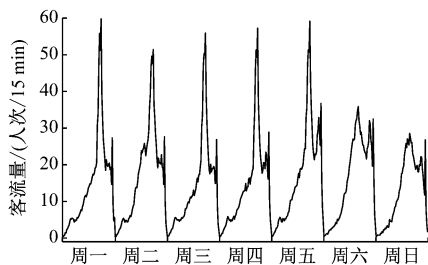


图4 2016年3月份第2周人民广场站进站客流量趋势图

此外,图5中的人民广场站进站量和莘庄站的出站量之间存在明显的正相关关系。可见,起点进站客流是影响终点出站客流的重要基础数据。

## 3 实例分析

本文以人民广场站—莘庄站和上海火车站—通河新村站两区间双向(即4个OD对)客流为例,对模型进行实例验证。

根据实际情况,结合同期处理好的天气数据,将实例数据输入构建好的 LSTM 模型和 GRU 模型,不断调整参数进行测试。测试发现,lr(学习率)为 0.01、epoch(迭代次数)为 400 次时,4 个 OD 对的拟合效果较好,Timestep(步长)为 3、4、5、6 的 LSTM 模型计算结果和 GRU 模型计算结果均趋于稳定,各 OD 对的收敛结果相近。本文仅展示人民广场站—莘庄站(由人民广场站进站且由莘庄站出站)的客流量迭代计算结果均方误差,如图 6 所示。

### 3.1 结果分析

由图 6 可见:① GRU 模型计算结果较 LSTM 模型计算结果先收敛,且差别显著。这得益于 GRU 运算简单。对于步长为 4 的情况,GRU 模型的计算结果在 epoch 约为 200 时就基本趋于稳定,而 LSTM 模型要在 epoch 约为 300 时才开始趋于稳定。② 图 6 c)中,当步长为 5 时,两模型的收敛曲

线最平滑。曲线的平滑度在一定程度上反映了收敛的稳定性。

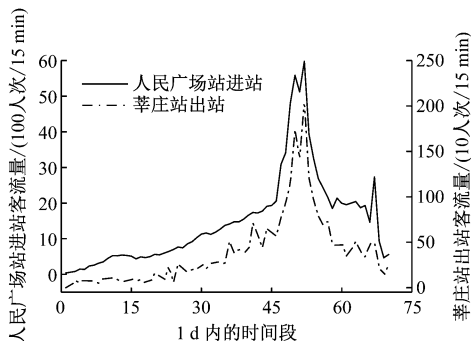


图5 人民广场站进站客流量与莘庄站出站客流量(人民广场站进站)变化关系

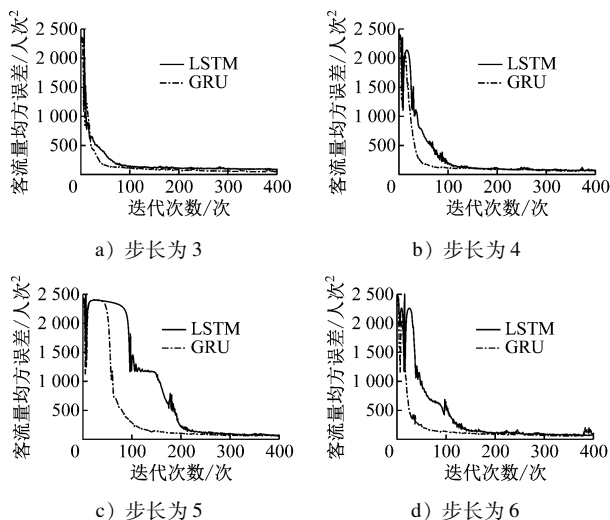


图6 人民广场站—莘庄站的 15 min 客流量迭代计算结果均方误差

### 3.2 模型评估与讨论

为了更加全面地比较不同模型的预测效果,本文选取  $E_{\text{RMS}}$  (均方根误差)、 $E_{\text{MA}}$  (平均绝对误差) 及  $E_{\text{MAP}}$  (平均绝对百分比误差) 作为模型评估指标。

$$E_{\text{RMS}} = \sqrt{\frac{1}{N} \sum (y_{\text{true}} - y_{\text{pred}})^2} \quad (1)$$

$$E_{\text{MA}} = \frac{1}{N} \sum |y_{\text{true}} - y_{\text{pred}}| \quad (2)$$

$$E_{\text{MAP}} = \frac{100}{N} \sum_{i=1}^N (|y_{\text{true}} - y_{\text{pred}}| / y_{\text{true}}) \% \quad (3)$$

式中:

$y_{\text{ture}}$ ——实际值;

$y_{\text{pred}}$ ——预测值;

$N$ ——预测样本数。

这三个指标均能反映预测值与实际值之间的误差大小。 $E_{\text{MA}}$  反映模型预测与真实值的误差情

况; $E_{RMS}$ 对异常值敏感,能反映模型的稳定性。若预测结果的 $E_{RMS}$ 、 $E_{MA}$ 值越小,则说明预测值和实际值越接近,模型的预测精度越高。

为了更好地评估两个模型,在学习率相同的条

件下,对迭代次数进行调整,并利用训练好的模型对4个OD对的测试集分别进行预测。总体性能评价指标取4个OD对应性能指标的均值。得到的预测误差见表2。

表2 当学习率为0.01时的LSTM模型和GRU模型预测误差

人次

迭代 次数/次	模型	步长为3时		步长为4时		步长为5时		步长为6时	
		$E_{MA}$	$E_{RMS}$	$E_{MA}$	$E_{RMS}$	$E_{MA}$	$E_{RMS}$	$E_{MA}$	$E_{RMS}$
300	GRU	5.660 7	8.814 8	5.765 2	8.900 9	5.662 5	8.896 5	5.771 1	9.019 4
	LSTM	5.782 1	8.827 1	6.760 7	11.239 6	6.358 9	9.662 9	7.219 6	13.032 7
350	GRU	6.217 9	9.771 2	5.927 7	8.978 0	5.706 3	8.764 3	5.847 3	9.354 7
	LSTM	5.803 6	8.955 0	5.954 5	8.858 4	6.465 2	10.618 5	6.733 9	11.004 1
400	GRU	5.899 1	8.988 7	5.846 4	9.241 6	5.622 3	8.715 4	5.634 8	8.905 2
	LSTM	5.942 0	9.162 8	6.479 5	9.971 0	5.891 1	9.197 9	6.534 8	10.873 1
450	GRU	6.085 7	9.413 1	6.537 5	10.163 5	6.155 4	9.830 2	6.526 8	9.464 5
	LSTM	6.517 9	9.949 7	6.391 1	9.642 3	6.207 1	9.454 0	6.760 7	10.743 9

由表2可以直观看出,当epoch为450次时,各项步长的拟合效果均开始降低。这说明在epoch大于400次时,模型为了更好的拟合训练集,出现了过拟合现象。当epoch为300~400次时,步长为3、4、5、6的最优预测结果( $E_{MA}$ 和 $E_{RMS}$ 最小)均由GRU模型得到,且GRU模型计算结果比LSTM模型计算结果更稳定,预测误差较小。最后,根据式(3)得出两个最优模型在高峰期间提前15 min预测结果的日均 $E_{MAP}$ 。由于高峰期间的客流量一直是交通领域的关注点,同时也不为0,因此,可以用式(3)对比模型性能。GRU模型预测结果的日均 $E_{MAP}$ 为15.3%, $E_{MAP}$ 最低为9.7%;LSTM模型预测结果的日均 $E_{MAP}$ 为15.9%, $E_{MAP}$ 最低为10.6%。

利用最优的LSTM模型和GRU模型对2016年3月份最后一星期数据进行预测,人民广场站—莘庄站的客流量预测结果与实测值比较如图7所示。

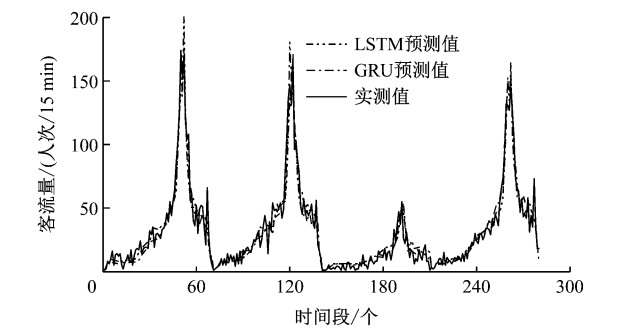


图7 人民广场站进站-莘庄站出站的OD对客流量的预测结果与实测值比较

由图7可见,LSTM模型和GRU模型的预测效果较接近,但LSTM模型的峰值预测存在较明显异常,而GRU模型的预测结果更稳定。

经过上述分析,在城市轨道交通OD短时客流预测研究中:①对于收敛速度,GRU模型快于LSTM;②在预测误差及稳定性评估方面,GRU模型略优于LSTM模型;③对于晚高峰期间提前15 min预测的 $E_{MAP}$ ,GRU模型的日均 $E_{MAP}$ 为15.3%,最低 $E_{MAP}$ 为9.7%,LSTM模型的日均 $E_{MAP}$ 为15.9%,最低 $E_{MAP}$ 为10.6%。

#### 4 结语

本文提出了一种基于GRU神经网络的城市轨道交通短时客流OD动态估计模型,并且将天气因素引入到短时客流OD预测中。在城市轨道交通短时客流OD预测研究中,模型简单、收敛速度更快的GRU与LSTM预测效果接近,且在预测稳定性方面略优于LSTM,验证了GRU神经网络在短时快速预测方面的优异性。

由于在实际的短时客流OD预测中,时间是一个非常关键的因素,因此,模型简单、收敛速度快的GRU更适用于实际的城市轨道交通运营分析。

目前,城市轨道交通的运营已经逐渐趋于网络一体化、智能化,本文的研究成果可以为乘客出行路径诱导、高峰时期客流预警限流、快速制定应急响应方案及需求响应式公交的实时调度等提供技术支撑。

#### 参考文献

[1] 陈志杰,毛保华,柏赞,等.基于多时间尺度的城市轨道交通短时OD估计[J].交通运输系统工程与信息,2017,17(5):166.

(下转第115页)

影响。

此外,在列车离站到进入候车阶段的一小段时间内,计算结果与仿真结果相差明显,而在候车阶段时两者结果又快速接近乃至重合。这是由于为了方便建模和计算,站台仿真模型将乘客的上下车速率平均化,并简化了行人在站台上行走的过程,对周期内各时段候车人数的准确度有一定影响,但并不影响周期内的站台最大人数,且各周期误差不具备累积性。因此,站台仿真模型验证了计算模型较高的有效性。

## 4 结语

本文对大客流期间站台连接设施处的瓶颈效应进行分析,考虑列车到站周期和滞留情况,以列车到站周期为基础,建立了侧式站台人数计算模型,并通过站台仿真模型来验证该计算模型的有效性。计算结果和仿真结果的相对误差较小,说明计算模型是有效的。

侧式站台人数计算模型能为岛式站台人数计算提供参考,可有效预测站台人数,有助于掌握突发性大客流的形成与发展机制,为车站的运营组织提供参考依据。

## 参考文献

[1] 金一兵. 模糊数学在预测车站最高聚集人数中的应用[J]. 铁

(上接第 58 页)

- [2] BERA S, RAO K V. Estimation of origin-destination matrix from traffic counts: the state of the art[J]. *EuropeanTransport*, 2011(49):2.
- [3] 姚向明, 赵鹏, 禹丹丹. 城市轨道交通网络短时客流 OD 估计模型[J]. *交通运输系统工程与信息*, 2015, 15(2): 149.
- [4] 蒋熙, 贾飞凡, 冯佳平. 基于 AFC 数据的城轨路网客流 OD 在线动态估计[J]. *交通运输系统工程与信息*, 2018, 18(5): 129.
- [5] 李梅, 李静, 魏子健, 等. 基于深度学习长短期记忆网络结构的地铁站短时客流量预测[J]. *城市轨道交通研究*, 2018(11):42.
- [6] HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities[J]. *Proceedings of the National Academy of Sciences*, 1982, 79(8):2554.
- [7] ELMAN J L. Finding structure in time[J]. *Cognitive Science*,

道运输与经济, 2004, 26(2):68.

- [2] 赵宇刚, 毛保华, 杨远舟. 城市轨道交通站台最高聚集人数计算方法研究[J]. *交通运输系统工程与信息*, 2011, 11(2):149.
- [3] 王永亮, 张星臣, 蒋洋, 等. 城市轨道交通侧式换乘站台最高聚集人数计算[J]. *交通运输系统工程与信息*, 2013, 13(5):167.
- [4] KEPAPTSOGLU K, KARLAFTIS M G. A model for analyzing metro station platform conditions following a service disruption[C]// *International IEEE Conference on Intelligent Transportation Systems*. Madeira Island, Portugal: IEEE, 2010: 1789.
- [5] 王刚. 城市轨道交通站台乘客聚集模拟模型研究[D]. 北京: 北京交通大学, 2012.
- [6] SERIANI S, FUJIYAMA T, HOLLOWAY C. Exploring the pedestrian level of interaction on platform conflict areas at metro stations by real-scale laboratory experiments[J]. *Transportation planning and technology*, 2017, 40(1): 100.
- [7] 王菲, 张全欣, 毛保华, 等. 基于时刻表协调的同台换乘聚集人数计算模型[J]. *交通运输系统工程与信息*, 2013, 13(3):163.
- [8] 商金涛, 陈峰. Vissim 在城市轨道交通车站客流仿真中的应用[J]. *城市轨道交通研究*, 2013, 16(1):54.
- [9] 周小妍. 广州地铁三号线(含三北线)各峰期运输组织的优化分析[J]. *企业技术开发*, 2015, 34(20):72.

(收稿日期:2018-03-14)

1990, 14(2):179.

- [8] BENGIO Y. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Transactions on Neural Networks*, 1994: 5.
- [9] HOCHREITER S, SCHMIDHUBER J. Long short-term memory. [J]. *Neural Computation*, 1997, 9(8):1735.
- [10] KYUNGHYUN C, BART V B, CAGLAR G, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// *EMNLP. The 19th Conference on Empirical Methods in Natural Language Processing*. Doha, Qatar:EMNLP, 2014.
- [11] GEOFFREY E H, NITISH S, ALEX K, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. *Neural and Evolutionary Computing*, 2012(6):35.

(收稿日期:2019-04-01)