

基于分布式学习的 5G 容量预测方法

5G Capacity Prediction Method Based on Distributed Learning

王哲¹,张千坤²,张鹏³,崔赞龙⁴,陈任翔²,刘西西²(1. 中国铁塔股份有限公司广东分公司,广东广州 510220;2. 中讯邮电咨询设计院有限公司,北京 100048;3. 中国联合网络通信集团有限公司,北京 100033;4. 中国联通苏州分公司,江苏苏州 215000)

Wang Zhe¹, Zhang Qiankun², Zhang Peng³, Cui Yunlong⁴, Chen Renxiang², Liu Xixi² (1. China Tower Co., Ltd. Guangdong Branch, Guangzhou 510220, China; 2. China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China; 3. China United Network Communications Group Co., Ltd., Beijing 100033, China; 4. China Unicom Suzhou Branch, Suzhou 215000, China)

摘要:

在移动网络中,每个小区都有独特的特性,每个小区训练的容量模型无法在其他小区使用。针对这个问题,提出一种基于分布式学习的5G小区容量预测方案,以联邦学习的架构为基础,联合不同小区的局部训练结果,完成全局模型的训练,从而达到增加模型泛化能力的目的。同时为减少计算量,提出一种基于仿射传播聚类的小区聚类算法,将不同小区的KPI进行归一化作为聚类的输入,将小区划分为不同的类别进行模型训练,从而减少局部模型的数量,提高整体训练的效率。

关键词:

容量预测;容量模型;小区聚类;联邦学习;分布式
doi:10.12045/j.issn.1007-3043.2026.01.007
文章编号:1007-3043(2026)01-0029-06
中图分类号:TN929.5
文献标识码:A
开放科学(资源服务)标识码(OSID):



Abstract:

In mobile networks, each cell has unique characteristics, and the capacity model trained in each cell can not be used in other cells. In view of the above problems, a 5G cell capacity prediction scheme based on distributed learning is proposed. Based on the framework of federated learning, the local training results of different cells are combined to complete the training of the global model, so as to achieve the purpose of increasing the generalization ability of the model. At the same time, in order to reduce the amount of calculation, a community clustering algorithm based on affine propagation clustering is proposed, and the KPIs of different communities are normalized as the input of clustering, and the communities are divided into different categories for model training, thereby reducing the number of local models and increasing the efficiency of overall training.

Keywords:

Capacity forecast; Capacity model; Cell clustering; Federated learning; Distributed

引用格式:王哲,张千坤,张鹏,等. 基于分布式学习的5G容量预测方法[J]. 邮电设计技术,2026(1):29-34.

1 概述

随着5G网络的发展和部署,数据业务流量成倍增长。流量是用户业务量在网络中的直接体现,也是运营商收入的基础^[1]。如何在有限投资下充分挖掘用户业务量需求,精准预测未来网络容量发展趋势,越

来越成为运营商关注的焦点。在运营商网络中,用户和业务分布的不均衡会导致网络中存在大量热点区域,甚至会出现因小区容量受限导致用户潜在业务需求被抑制的情况^[2]。同时,针对现网已有的容量瓶颈,采用传统方式进行扩容无法真正满足未来需求,而大规模扩容又会导致投资收益率低,因此对小区容量的精准预测就变得越来越重要^[3]。

5G小区容量本质上是时间序列数据,因此对其预

收稿日期:2025-11-10

测问题可以转化为时间序列预测建模问题。目前小区容量预测的方法主要基于业务模型^[4]和话务模型^[5]的抽象模型,或通过系统平均频谱效率所得,一般基于单业务模型考虑业务数据经过各个传输协议层处理后到达物理层的实际速率需求,再结合空口平均吞吐量,从而得到单业务的用户流量。然后根据用户的话务模型分布比例特征,获得综合业务的用户容量。以往的方法主要基于统计和概率分布等数学理论知识对小区容量进行建模和预测。这些方法属于经验模型,通过有限的参数对容量进行建模,不依赖整个数据集。随着移动蜂窝网在广度和深度上的扩展,容量特性已不足以用线性变化来表征,文献[6]分析了许多蜂窝网络的容量特性,将小区容量分为2个部分:可预测的小区容量和不可预测的小区容量,可预测的小区容量具有自相关性。文献[7]提出一种从时间和空间角度来预测容量的方法,实验结果表明,该方法与以往的线性经验模型相比,预测精度有显著的提高。

近年来,随着机器学习和人工智能的快速发展,深度学习逐渐成为容量预测的热门方向^[8-10]。大多数深度学习的训练都是在IDC完成的^[11-12],需要大量的数据和海量的算力,在训练阶段会消耗大量的资源,需要几天时间才可以完成训练。为了解决模型训练消耗大量资源的问题,联邦学习(Federated Learning)^[13-14]作为一种新的分布式机器学习方式被提出,通过聚合各个点位上报的局部训练结果来完成全局

模型的训练。当前使用比较多的为基于迭代模型平均的算法,在每次训练迭代中,小区可基于本地的训练数据,从中央服务器下载全局模型进行训练并上报中间训练结果,从而更新全局模型,一直循环迭代。为了进一步减少资源消耗,本文以联邦学习的分布式架构为基础,对5G容量进行预测。首先,本文提出一种小区聚类的方法,获取每个省份的时序基站侧数据,并进行数据清洗得到有用数据,采用仿射传播聚类^[15]的方式对这些时序数据归一化并进行聚类,将小区级模型提升至“聚类”级别,减少区域模型的数量;其次,提出一种联邦时间序列预测方法,将LSTM算法^[16-17]应用到联邦学习当中,在每个区域采用LSTM算法训练一个局域模型;最后,采用FedAvg算法^[18],随机选择 m 个区域,对这 m 个区域的梯度更新进行平均以形成全局更新,同时用当前全局模型替换未采样的区域。

2 基于仿射传播的小区聚类方法

小区聚类的整体框图如图1所示。

选取某运营商所有省份的KPI信息,主要包括小区名称、数据日期、RRC连接平均用户数_忙时、小区PRB利用率_忙时、经度、纬度、PDCP层用户面流量GB_忙时、方位角、网格区域分类。其中网格区域有78类,将其编号为0~77。选择RRC连接平均用户数_忙时、小区PRB利用率_忙时和PDCP层用户面流量GB_忙时3个指标,对同一小区的所有时间段的RRC连接

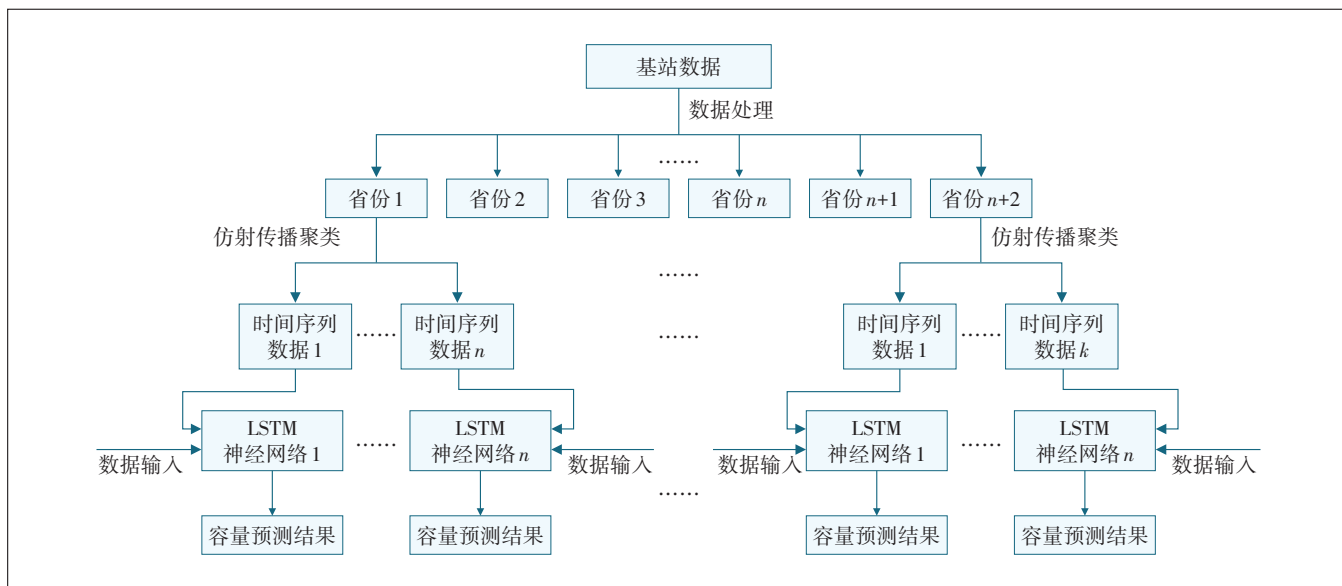


图1 小区聚类整体框图

平均用户数_忙时、小区 PRB 利用率_忙时、PDCP 层用户面流量 GB_忙时求均值得到样本数据。其次对这些样本值进行归一化处理:

$$x' = \frac{x - \min}{\max - \min} \quad (1)$$

其中, x 对应同一小区的所有时间段的 RRC 连接平均用户数_忙时、小区 PRB 利用率_忙时、PDCP 层用户面流量 GB_忙时求均值得到的样本数据。

第 i 个小区的样本值可以表示为 $\Psi_i = [\text{RRC}_i, \text{PRB}_i, \text{PDCP}_i]^T$ 。

采用仿射聚类传播算法,通过计算样本相似度矩阵 S 来对小区样本点实现聚类,将 2 个小区之间欧式距离平方的负值表示为相似度,如式(2)所示:

$$S(i, j) = -\|\Psi_i - \Psi_j\|_2, \forall i, j \in \{1, 2, \dots, N\}, i \neq j \quad (2)$$

其中, N 为小区的数量。

小区样本之间两两互相传递吸引度和可信度 2 类信息, $R(i, j)$ 表征小区 i 到小区 j 的吸引度,小区 j 可作为小区 i 聚类中心的概率, $A(i, j)$ 表征小区 j 到小区 i 的归属度,小区 i 可作为小区 j 聚类中心的概率,则:

$$R(i, j) = S(i, j) - \max_{k \neq j} \{A(i, k) + S(i, k)\} \quad (3)$$

$$A(i, j) = \min \left\{ 0, R(j, j) + \sum_{k \neq \{i, j\}} \max[0, R(k, j)] \right\} \quad (4)$$

对 $R(i, j)$ 和 $A(i, j)$ 初始化,将迭代次数 n 的值设为 0,计算 2 个小区之间的相似度 $S(i, j)$,更新 $R(i, j)$ 和 $A(i, j)$,并计算:

$$k = \arg \max_{j \in \{1, 2, 3, \dots, N\}} \{A(i, j) + R(i, j)\} \quad (5)$$

如果 $k = i$,则第 i 个小区作为聚类中心,否则 k 将被选为聚类中心。不断迭代,直到达到最大迭代次数或者聚类中心不再变化时结束。

聚类效果如图 2 所示。

3 联邦时序预测

3.1 LSTM 神经网络

本文采用长短时记忆网络(LSTM)对聚类后的小区训练时序模型(见图 3)。LSTM 本质上是循环神经网络的一种,可通过增加门限来解决循环神经网络短期记忆的问题,并可有效利用长距离的时序信息。

其中细胞状态 C_t 表示长时累计信息, h_t 表示隐层

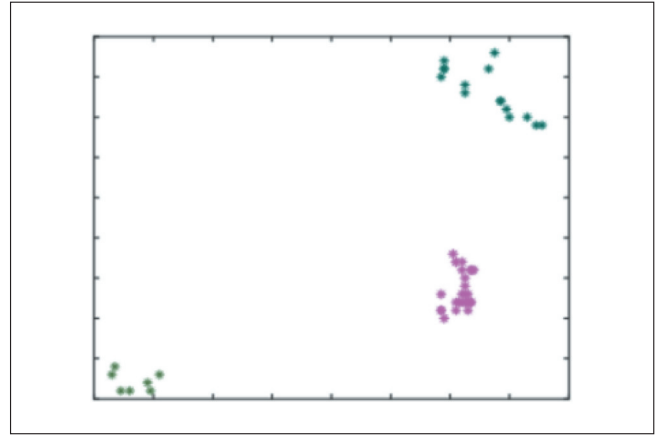


图 2 聚类效果

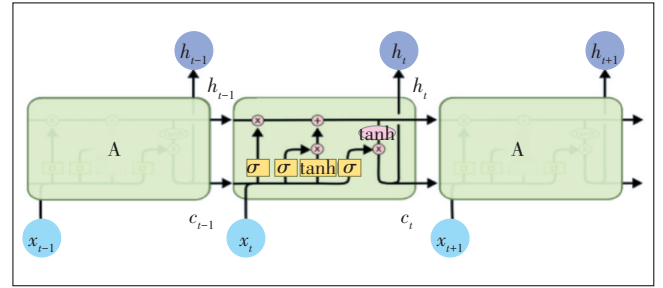


图 3 LSTM 神经网络

状态,代表短时的存储单元的输出。输入端有 3 个参数,细胞状态 C_{t-1} ,隐层状态 h_{t-1} , t 时刻的输入向量 x_t ,输出端有 2 个参数:细胞状态 C_t 和隐层状态 h_t 。

LSTM 中存在 4 个神经网络层,以“门”的结构形式存在,主要包括输入门、输出门、遗忘门和更新门。其中,输入门控制不同时刻输入 LSTM 单元的外界信息,如式(6)所示。遗忘门控制 LSTM 单元在不同时刻丢弃信息,如式(7)所示。更新门控制 LSTM 单元更新内部存储信息,如式(8)和式(9)所示。输出门控制不同时刻 LSTM 单元的输出信息,由 2 个部分组成,如式(10)和式(11)所示。

$$i_t = \text{sigmoid}(W_i \times [h_{t-1}, x_t] + b_i) \quad (6)$$

$$f_t = \text{sigmoid}(W_f \times [h_{t-1}, x_t] + b_f) \quad (7)$$

$$\tilde{C}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (8)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (9)$$

$$o_t = \text{sigmoid}(W_o \times [h_{t-1}, x_t] + b_o) \quad (10)$$

$$h_t = o_t \times \tanh(C_t) \quad (11)$$

其中, W 和 b 表示相应的权重系数矩阵和偏置向量。

3.2 联邦学习设计

FedAvg 是联邦学习的经典算法,其详细的架构和工作流程如图4所示。

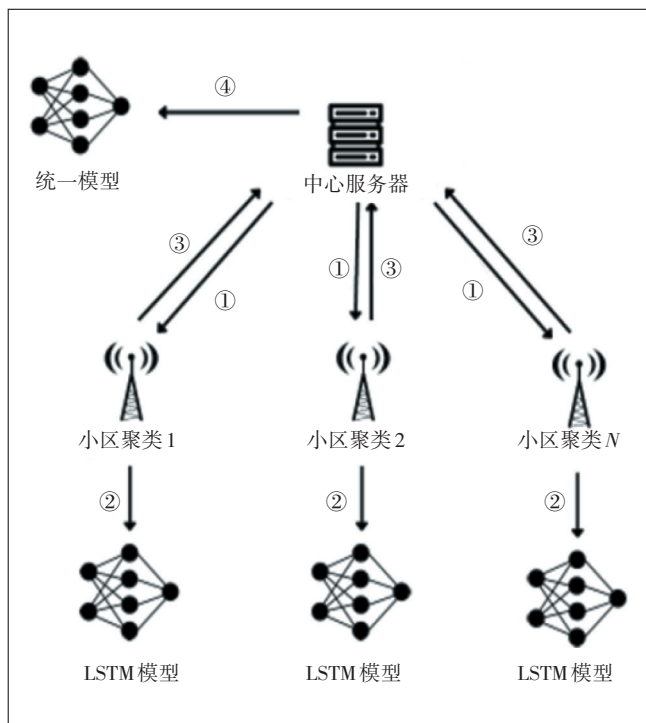


图4 联邦学习框架

FedAvg 将训练分为多个通信回合,每循环一次中心服务器都会通过本地服务器训练的模型聚合获得全局模型,然后用测试集对全局模型进行验证,来验证当前联邦学习的进度。每个通信回合包含以下4个工作步骤。

a) 模型下发。根据本地服务器的数量即小区聚类的数量,中心服务器会通过网络向各个本地服务器发送上一次的全局模型,准备进行本地训练。

b) 本地训练。当本地服务器在接收到全局模型后,会利用本地的数据源对全局模型进行有监督的训练。

c) 模型上传。各个本地服务器会把完成本地训练的本地模型通过网络发送给中心服务器。

d) 模型聚合。中心服务器对多个本地模型的梯度参数进行聚合,获得新的全局模型。

假设第 k 个小区聚合第 t 个通信回合的模型参数表示为 $W_k^{(t)} = \{w_k^{(t), 1}, w_k^{(t), 2}, \dots, w_k^{(t), L}\}$, $w_k^{(t), l}$ 表示第 k 个小区聚类集合全局模型的第 l 层参数, $l = 1, 2, \dots, L$ 。 L 是本地模型的最大网络层数。每个

小区聚类集合的本地数据集为 $D_k = \{(x_{k, n}, y_{k, n})\}_{n=1}^{N_k}$, 其中 $x_{k, n}$ 表示小区聚类集合的时间信息, $y_{k, n}$ 表示小区聚类集合的容量信息, N_k 表示本地数据集的最大样本数量。本地模型更新的过程如式(12)所示:

$$W_k \leftarrow W_k - \eta \frac{1}{N_k} \sum_{n=1}^{N_k} \nabla F_k(W_k, x_{k, n}, y_{k, n}) \quad (12)$$

其中, ∇F_k 为本地更新的损失函数, $\frac{1}{N_k} \sum_{n=1}^{N_k} \nabla F_k(W_k, x_{k, n}, y_{k, n})$ 为第 k 个客户端在本地数据集 D_k 下计算的随机梯度, η 表示学习率。在每次本地更新阶段,客户端会对本地模型进行 t 次本地更新。客户端完成本地更新以后,会把本地的模型参数通过网络上传到服务器进行模型参数聚合,获得新的全局模型,聚合过程如式(13)所示:

$$W^{(t+1)} \leftarrow \sum_{k=1}^K \frac{N_k}{N} W_k^{(t+1)}, \quad t = 1, 2, \dots, T \quad (13)$$

其中, $W^{(t+1)}$ 表示第 $t+1$ 个通信回合通过模型聚合得到全局模型, T 表示最大的通信回合数。在模型分配阶段,服务器会把最新的全局模型分配给各个客户端,以进行本地更新,过程如式(14)所示:

$$W_k^{(t+1)} \leftarrow W^{(t+1)} \quad (14)$$

4 仿真实验

本次仿真选择某市的数据进行验证,表1为部分KPI指标,包括时间信息、基站和小区信息以及对应的容量信息。本次仿真是对小区下行容量进行预测,选择3000个采样点,验证算法的可行性。

首先对选取的3000个采样点进行归一化处理,再进行聚类,共得到6类小区信息数据,对这6类数据分别进行LSTM训练,以第1个聚类结果为例,将数据分为训练值、验证值和测试值,其中训练值用于训练本地模型,验证值用于验证模型在本地效果,测试值为验证经过联邦学习后统一模型的效果(见图5)。

聚类本身需要进行归一化处理,可以采用聚类后的值统一进行LSTM传播模型训练,这里采用3个小区的聚类样本进行本地化训练,可以得到3个LSTM传播模型,采用验证值来检测训练模型的效果,得到的结果如图6所示。

通过FedAvg算法完成整个联邦学习后,将全局统一模型对这3个小区聚类样本的测试数据集进行测

表1 某市数据

数据日期	gnb_id	cell_id	省份	RRC连接平均 用户数_忙时	小区上行PRB 占用总数_忙时	小区上行PRB 可用总数_忙时	小区下行PRB 占用总数_忙时	小区下行PRB 可用总数_忙时	上行PDCP层用 户面流量GB_ 忙时	下行PDCP层用 户面流量GB_ 忙时
20220124	5849763	8	某市	4.953	20 098 750.76	589 732 416	30 384 227.96	1 376 217 570	0.033 917 051	0.858 068 408
20220125	5849763	8	某市	9.582	35 060 447.53	589 568 889	34 107 355.99	1 375 835 097	0.728 072 314	0.609 753 525
20220126	5849763	8	某市	20.225	22 717 297.75	589 568 889	32 007 071.55	1 375 835 370	0.040 543 574	1.061 819 795
20220127	5849763	8	某市	5.310	19 355 891.75	589 568 889	25 121 589.40	1 375 835 097	0.018 524 551	0.615 551 367
20220128	5849763	8	某市	15.134	21 694 191.73	589 568 889	27 473 875.71	1 375 835 643	0.037 731 787	0.716 904 727
20220129	5849763	8	某市	10.316	22 326 230.49	589 405 089	60 225 033.56	1 375 453 443	0.093 665 547	3.005 342 920
20220130	5849763	8	某市	13.157	20 993 866.03	589 568 616	33 591 704.65	1 375 835 643	0.066 823 711	1.314 392 627
20220131	5849763	8	某市	3.853	16 868 637.59	589 404 816	22 815 245.27	1 375 453 443	0.013 210 732	0.697 243 174

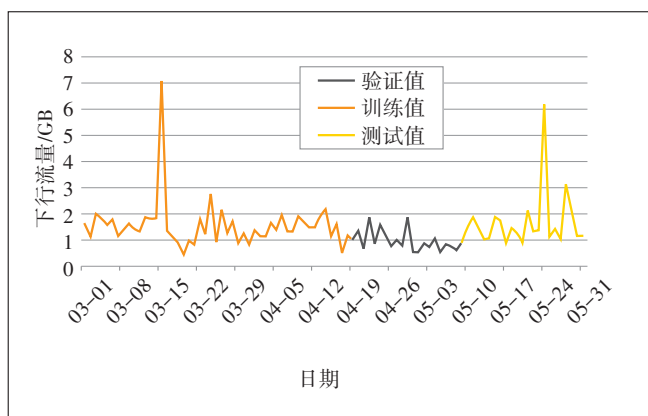


图5 样本数据分布

试,结果如图7所示。

本文通过回归模型指标NRMSE来表征模型训练的效果,即RMSE的归一化形式, RMSE采用欧氏距离来表示预测值与真实值的差距, NRMSE的计算公式为:

$$NRMSE = \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (15)$$

表2所示为小区聚类集合采用本地训练的LSTM神经网络模型以及采用联邦学习的模型得到的结果,从表2可以看出,联邦学习模型的精度接近于本地数据集单独训练的模型精度。

5 结论

随着5G网络的开发和部署,网络规划对容量预测的准确性、复杂度和通用性提出了更高的要求。考虑到不同的小区、不同的基站本地生成的模型不具备泛化性,本文采用联邦学习的方法,通过对小区进行聚类,减少本地模型的数量,并联合本地的局部神经网络模型生成统一的全局模型,减少了资源的开销,同时不需要将本地数据上传到集中服务器,减少了数据泄露的风险。基于联邦学习的时序容量预测可处理

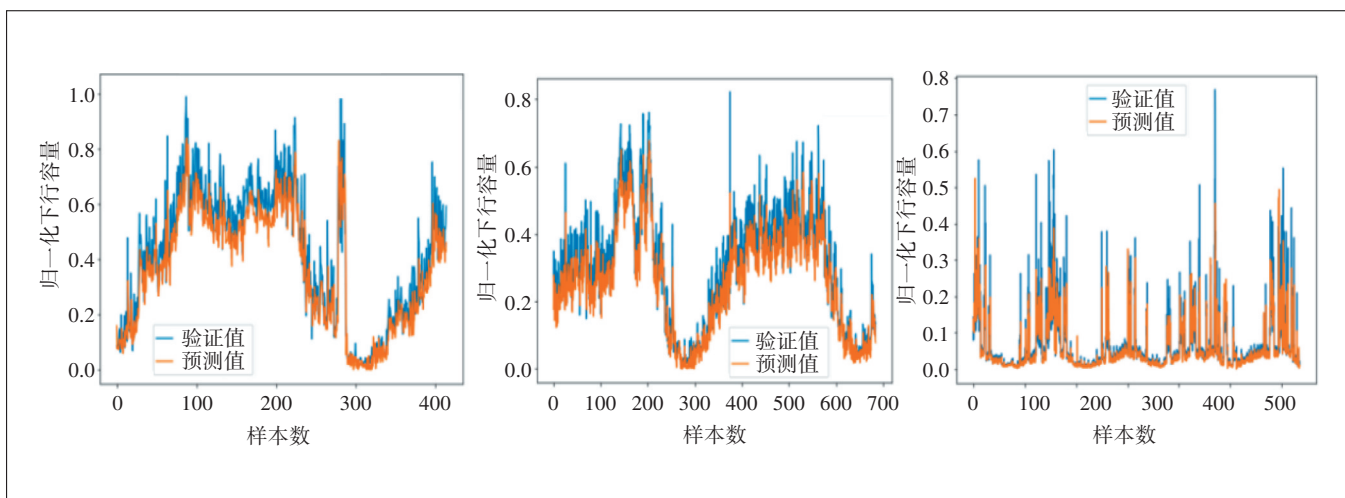


图6 验证值预测结果

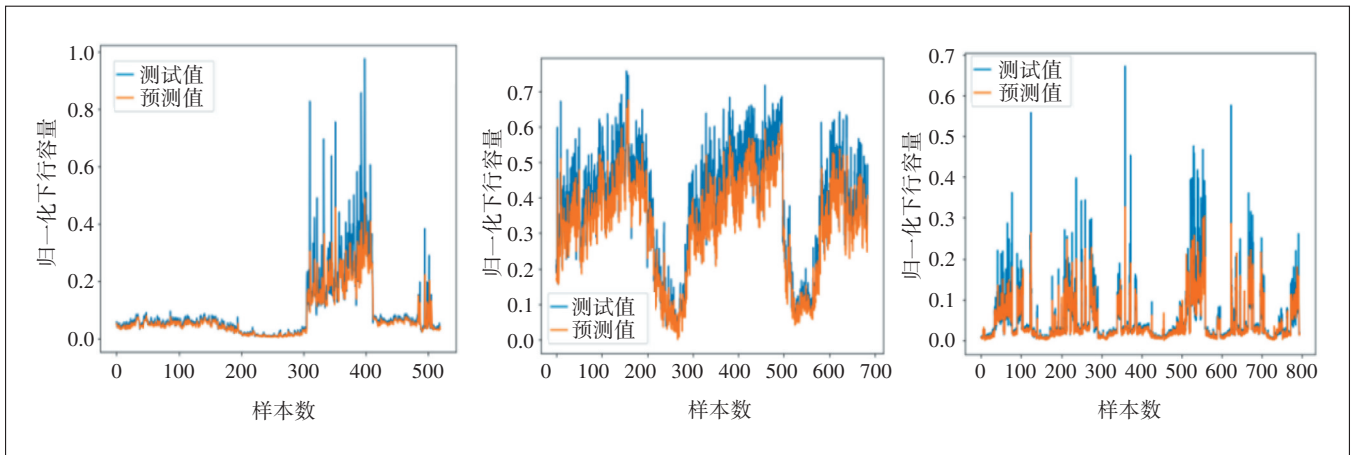


图7 测试值预测结果

表2 模型的评测结果

	小区聚类1	小区聚类2	小区聚类3
LSTM	0.831	0.323	0.834
联邦学习	0.883	0.307	0.839

大量的数据,为后续5G海量物联(mMTC)场景的容量预测奠定了基础。仿真结果表明,该方法可以达到和本地模型训练相近的精度,可以很好地应用于5G现网场景。

参考文献:

[1] 陈玲,高和,郑盛. 5G用户及业务预测方法研究[J]. 邮电设计技术,2020(8):11-14.
 [2] 汪洁云. 5G室内覆盖及数字化建设研究[J]. 中国设备工程,2023(5):38-40.
 [3] 刘永涛,王磊,李军,等. 面向5G网络演进的超密集组网规划方案研究[J]. 电信工程技术与标准化,2019,32(8):1-6.
 [4] 马占军. 5G网络业务模型及带宽需求分析[J]. 通讯世界,2020,27(6):64-65.
 [5] 康灵坡. 5G基站站址规划策略研究[J]. 电子世界,2021(16):9-10.
 [6] LI R P, ZHAO Z F, ZHENG J C, et al. The learning and prediction of application-level traffic data in cellular networks[J]. IEEE Transactions on Wireless Communications, 2017, 16(6): 3899-3912.
 [7] YANG L, GU X X, SHI H F. A novel satellite network traffic prediction method based on GCN-GRU[C]//2020 International Conference on Wireless Communications and Signal Processing (WCSP). Nanjing: IEEE, 2020: 718-723.
 [8] 周艳艳,郝园媛. 基于机器学习的运营商客户行为分析[J]. 科学技术与工程,2022,22(2):585-592.
 [9] 欧阳晔,杨爱东,孟凡语. 一种博弈论辅助的机器学习算法检测用户流失行为[J]. 电信科学,2020,36(6):79-89.

[10] 朱凡,王印琪. 基于k-means与神经网络机器学习算法的用户信息聚类及预测研究[J]. 情报科学,2021,39(7):83-90.
 [11] SHI W Q, ZHOU S, NIU Z S. Device scheduling with fast convergence for wireless federated learning[C]//ICC 2020 - 2020 IEEE International Conference on Communications (ICC). Dublin: IEEE, 2020:1-6.
 [12] Accessed [EB/OL]. [2025-05-25]. <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper>.
 [13] PAN J P, CAI L, YAN S, et al. Network for AI and AI for network: challenges and opportunities for learning-oriented networks [J]. IEEE network, 2021, 35(6): 270-277.
 [14] OPPO. 6G AI-cube intelligent networking [R/OL]. [2025-05-25]. [https://www.oppo.com/content/dam/oppo/en/mkt/newsroom/press/oppo-unveils-6g-white-paper/6G% 20AI-Cube% 20Intelligent% 20Networking.pdf](https://www.oppo.com/content/dam/oppo/en/mkt/newsroom/press/oppo-unveils-6g-white-paper/6G%20AI-Cube%20Intelligent%20Networking.pdf).
 [15] 田增山,张千坤,周牧,等. 基于CFR虚拟阵列天线的AOA室内定位[J]. 电子学报,2018,46(6):1468-1474.
 [16] TCHRAKIAN T T, BASU B, O'MAHONY M. Real-time traffic flow forecasting using spectral analysis [J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(2): 519-526.
 [17] YE Q, SZETO W Y, WONG S C. Short-term traffic speed forecasting based on data recorded at irregular intervals [J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1727-1737.
 [18] 温依霖,赵乃良,曾艳,等. 基于本地模型质量的客户端选择方法[J]. 计算机工程,2023,49(6):131-143.

作者简介:

王哲,工程师,硕士,主要从事网络运营管理、网络优化提质等工作;张千坤,工程师,硕士,主要从事移动增值业务相关咨询设计工作;张鹏,高级工程师,主要从事移动网络规划、移动通信新技术研究等工作;崔赞龙,工程师,主要从事移动网络规划、移动通信新技术研究等工作;陈任翔,高级工程师,学士,主要从事移动增值业务相关咨询设计工作;刘西西,工程师,硕士,主要从事移动增值业务相关咨询设计工作。