

基于小波包与长短时记忆融合的铁路旅客流量预测模型^①

成 强

(重庆交通大学 信息科学与工程学院, 重庆 400074)

通讯作者: 成 强, E-mail: 13648392624@163.com

摘 要: 短期铁路客运需求量的实时精准预测可以为实时调整客运服务结构提供依据. 铁路旅客流量数据具有时变性、非线性和随机波动性等特点, 传统的预测模型无法精准的预测短期内的客流量. 本文提出一种基于小波包分解与长短时记忆融合的深度神经网络预测模型 (WPA-LSTM), 首先用小波包分解将原始客运量时间序列分解重构成多个不同尺度的低频和高频序列, 然后分别针对各个子序列进行 LSTM 模型训练和预测, 最后将各子序列的预测值叠加作为 WPA-LSTM 模型的输出. 采用某高铁 367 天的日旅客流量数据对模型进行实验验证, 并与季节性模型和基于经验模态的长短时记忆融合模型进行对比, 实验结果表明, WPA-LSTM 模型可有效提高铁路旅客流量预测的精度.

关键词: 深度学习; 小波包分析; 长短时记忆神经网络

引用格式: 成强. 基于小波包与长短时记忆融合的铁路旅客流量预测模型. 计算机系统应用, 2018, 27(7): 121-126. <http://www.c-s-a.org.cn/1003-3254/6391.html>

Hybrid Model Based on Wavelet Packet and Long Short-Term Memory for Railway Passenger Traffic Volume Prediction

CHENG Qiang

(School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: Real-time accurate prediction of short-term railway passenger demand can provide basis for real-time adjustment of passenger service structure. Railway passenger flow data has characteristics such as time-varying, nonlinear and stochastic volatility. Traditional forecasting models cannot predict the short-term passenger traffic volume accurately. This study proposed a hybrid deep learning model based on Wavelet Packet Analysis and Long Short-Term Memory (WPA-LSTM). Firstly, the original passenger volume time series is decomposed into several low-frequency and high-frequency sequences with different scales by wavelet packet. Then, the LSTM model training and prediction are carried out respectively for each sub-sequence. Finally, the prediction values of each sub-sequence are superimposed as the output of WPD-LSTM model. The model was validated by the daily passenger flow data of 367 days of one high-speed railway. The model was compared with the seasonal model and the empirical model. The experimental results show that WPD-LSTM model can effectively improve the accuracy of railway passenger traffic forecasting.

Key words: deep learning; Wavelet Packet Analysis (WPA); Long Short-Term Memory (LSTM)

随着政企分开以及中国铁路总公司的成立, 铁路改革的头重脚轻, 重点就是要加强对于短期市场变化的系统已经迎来改革发展“新常态”. 客运经营作为铁路应对能力, 及时获取相关客运需求信息, 实时调整客运

^① 收稿时间: 2017-10-11; 修改时间: 2017-11-03; 采用时间: 2017-11-08; csa 在线出版时间: 2018-06-27

服务机构,而这这就要求我们提高对于短期客运量的预测精度。

短期客运量预测是以月度或者日度客运量变化为预测对象,研究短期内客运量的波动情况。由于短期客运量受多种随机因素影响,相比长期客运量具有明显的非线性、随机性和波动性^[1],这在一定程度上增加了预测难度。在类似客运量时间序列预测方面,小波分析和神经网络相结合的方法收到广泛关注,而深度学习^[2]的飞速发展又为客运量时间序列的预测提供了新的研究思路和方法。

深度学习的概念由 Hinton 等于 2006 年提出。深度学习神经网络提供了高维数据空间和低维嵌套结构的双向映射,有效解决了大多数非线性降维方法所不具备的逆向映射问题。杨祎玥等^[3]将小波分解和循环神经网络 RNN 应用在时间序列的预测中,探索了水文时间序列的预测精度。但是小波分解仅仅将低频特性持续分解,而且循环神经网络 RNN 存在梯度消失和梯度爆发的缺点;然而小波包分析 (Wavelet Packet Analysis, WPA) 不仅对低频部分进行分解,而且还对高频部分进行分解,且长短时记忆神经网络 LSTM^[4]是在 RNN 的基础上的改进模型,可以有效处理梯度消失和梯度爆发的问题。

因此,本文利用长短时记忆神经网络 LSTM 的时间序列预测能力,结合小波包分析方法,建立了一个基于小波包分析的长短时记忆神经网络的铁路客运量时间序列预测模型。

1 小波包分析

1.1 小波包的定义

定义子空间 U_j^n 是函数 $u_n(t)$ 的闭包空间,而 U_j^{2n} 是函数 $u_{2n}(t)$ 的闭包空间,并令 $u_n(t)$ 满足下面的双尺度方程^[5]:

$$\begin{cases} u_{2n}(t) = \sqrt{2} \sum_{k \in Z} h(k)u_n(2t-k) \\ u_{2n+1}(t) = \sqrt{2} \sum_{k \in Z} g(k)u_n(2t-k) \end{cases} \quad (1)$$

式中, $g(k) = (-1)^k h(1-k)$, 即两系数也具有正交关系,当 $n = 0$ 时,可以直接给出:

$$\begin{cases} u_0(t) = \sum_{k \in Z} h(k)u_0(2t-k) \\ u_1(t) = \sum_{k \in Z} h(k)u_0(2t-k) \end{cases} \quad (2)$$

与在多分辨分析中, $\phi(t)$ 和 $\varphi(t)$ 满足双尺度函数方程:

$$\begin{cases} \phi(t) = \sum_{k \in Z} h(k)\phi(2t-k) \\ \varphi(t) = \sum_{k \in Z} g(k)\varphi(2t-k) \end{cases} \quad (3)$$

相比较,很显然 $u_0(t)$ 和 $u_1(t)$ 分别退化为尺度函数 $\phi(t)$ 和小波基函数 $\varphi(t)$ 。序列 $\{u_n(t)\}$ (其中 $n \in Z_+$) 称为由基函数 $u_0(t) = \phi(t)$ 确定的正交小波包。或称序列 $\{u_n(t)\}$ 为关于序列 $\{h(k)\}$ 的正交小波包。

1.2 小波包对时间序列的分解特性

小波分析是把时间序列 S 分解成低频信息 a_1 和高频信息 d_1 两部分。在分解中,低频 a_1 中失去的信息由高频 d_1 捕获。在下一层的分解中,又将 a_1 分解成低频 a_2 和高频 d_2 两部分,低频 a_2 中失去的信息由高频 d_2 捕获,如此类推下去,可以进行更深层的分解。小波包分析则不然,它不仅对低频部分进行分解,而且还对高频部分进行分解。因此,在对于实验数据的处理上,并没有将节假日和周末客流进行单独区分并预测,而是将完整数据构成的时间序列进行小波包的分解以及重构。图 1 为小波包对一维时间序列的分解特性示意图。在图 1 中, A 表示低频, D 表示高频,末尾下标的序号数表示小波包分解的层数 (也即尺度数)。分解具有关系:

$$S = A_1A_2A_3 + A_1A_2D_3 + A_1D_2A_3 + A_1D_2D_3 + D_1A_2A_3 + \dots + D_1A_2D_3 + D_1D_2A_3 + D_1D_2D_3 \quad (4)$$

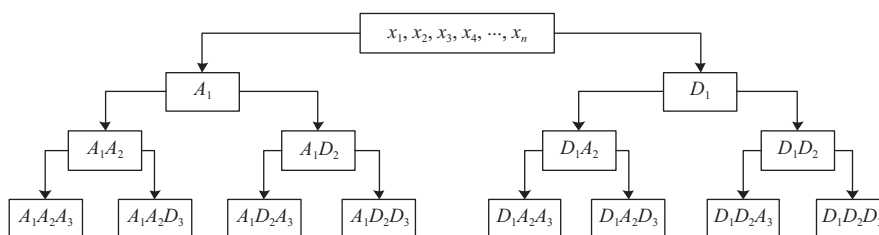


图 1 小波包对时间序列的分解

1.3 小波包算法

为了小波包简单地就说就是一个函数簇, 它定义为正交尺度函数 $u_0(t) = \phi(t)$ 确定的函数簇. 设 $g_j^n(t) \in U_j^n$, 则 $g_j^n(t)$ 可表示为 (其中 j 是尺度参数, k, l 是平移参数, n 是频率参数, 且 $j, k, l \in \mathbb{Z}, n \in \mathbb{Z}_+$)^[6]:

$$g_j^n(t) = \sum_l d_l^{j,n} u_n(2^j t - l) \quad (5)$$

小波包分解算法:

由 $\{d_l^{j+1,n}\}$, 求 $\{d_l^{j,2n}\}$ 与 $\{d_l^{j,2n+1}\}$, 即:

$$\begin{cases} d_l^{j,2n} = \sum_k a_{k-2l} d_k^{j+1,n} \\ d_l^{j,2n+1} = \sum_k b_{k-2l} d_k^{j+1,n} \end{cases} \quad (6)$$

小波包重构算法:

由 $\{d_l^{j,2n}\}$ 与 $\{d_l^{j,2n+1}\}$, 求 $\{d_l^{j+1,n}\}$, 即:

$$d_l^{j+1,n} = \sum_k [h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1}] \quad (7)$$

2 深度循环神经网络

深度循环神经网络 (Deep Recurrent Neural Network, DRNN) 本质上是一种通常意义上的深度神经网络 (多层感知机), 其特点是每层都有时间反馈循环, 并且层之间是叠加构成的. 每次神经网络的更新, 新信息通过层次传递, 每层神经网络也获得了时间性上下文信息.

然而传统的 DRNN 模型, 在训练过程中, 梯度下降更加倾向于按照序列结尾处的权值的正确方向进行更新. 也就是说, 越远的序列输入的对权值的正确变化所能起到的“影响”越小, 所以训练的结果就是往往出现偏向于新的信息, 即不太能有较长的记忆功能.

由此引出本文将要介绍的 LSTM (Long Short-Term Memory, 长短时记忆) 递归神经网络, 该算法使用 LSTM 结构替换了一般的递归神经网络的隐藏层节点, LSTM 结构增加了输入门 (Input gate)、输出门 (Output gate)、遗忘门 (Forget gate) 和一个内部单元 (Cell). 其中 LSTM 结构如图 2 所示.

输入门: 表示是否允许输入层的信息进入到该隐藏层节点. 门开的则允许输入层输出信号进入, 门关则不允许信号进入, 记为 τ . 输出门: 表示是否将当前节点的输出值输出到下一层. 门开的则允许该隐藏层节点的信号输出, 门关则不允许信号输出, 记为 ω . 遗忘门: 表示是否保留当前隐藏层节点存储的历史信息. 门开

的则保留当前隐藏层节点存储的历史信息, 门关则不保留当前隐藏层节点存储的历史信息, 记为 φ . s_c^t 表示 t 时刻存储的信息值, s_c^{t+1} 表示 $t+1$ 时刻存储的信息值, 在模型中 t 时刻的输出值将作为 $t+1$ 时刻的输入进行预测, 输出层与 DRNN 的输出层一致. 基于长短时记忆结构的递归神经网络模型结构如图 3 所示.

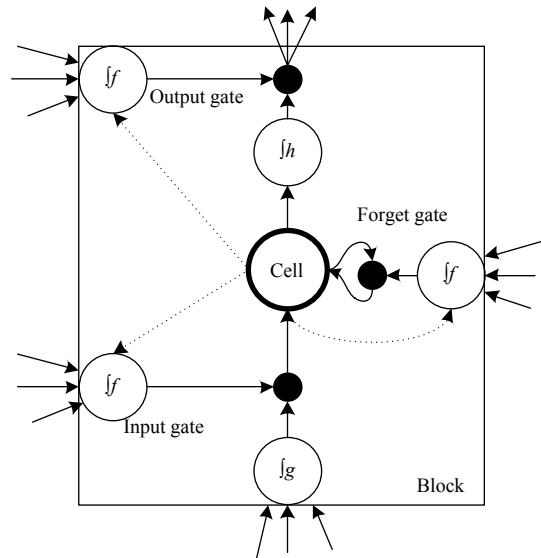


图 2 LSTM 结构

3 预测模型结构

利用小波包分析方法对铁路旅客流量时间序列进行预处理, 进行多尺度小波包分解和单支重构, 可获得不同的高频和低频序列, 对预处理的数据进行相空间重构后作为 LSTM 模型的训练数据, 建立一个基于小波包分析的长短时记忆神经网络的铁路旅客客流量时间预测模型, 称为 WPA-LSTM 模型.

WPA-LSTM 预测模型结构如图 4 所示.

首先将时间序列数据进行小波包分解和单支重构成多条序列, 然后分别输入 LSTM 模型进行预测, 最后将预测值叠加成最终预测值^[7].

WPA-LSTM 模型建立过程如下:

(1) 数据准备. 先选取历史旅客流量日数据, 对数据进行归一化处理, 并且保存其最大值和最小值.

(2) 小波分解和单支重构. 选择合适的小波包函数和分解初度, 对旅客流量时间序列进行多尺度小波包分解和单支重构, 获得该序列的近似系数和各细节系数, 然后对这些系数分别进行单支重构, 获得一条能够

描述原始序列趋势变化的低频序列和保留不同信息的高频序列。

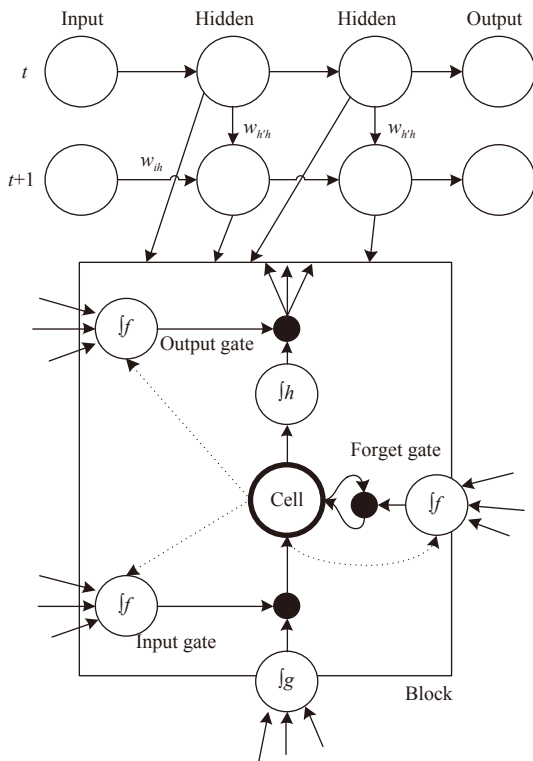


图3 基于长短时记忆结构的递归神经网络模型结构

(3) 相空间重构. 设置相空间重构的延迟数值, 分别对低频序列和各个高频序列进行相空间重构, 以此形成相应的训练数据和测试数据。

(4) LSTM 模型训练. 设定 LSTM 的学习速率、网络层数和每层的节点数目, 利用相空间重构的多组训练数据分别对 LSTM 模型进行训练。

(5) 预测客流量. 用训练好的 LSTM 模型对子序列的测试样本进行预测, 并将各子序列的预测值进行叠加, 从而得到最终的预测值, 这样就实现了对铁路旅客流量时间序列的建模和预测。

4 实例分析

以某高铁上, S1 到 S2 方向的 2015 年 12 月 28 日到 2017 年 1 月 15 日, 共计 367 天的日旅客流量数据为例, 来研究基于小波包分析的 LSTM 神经网络预测模型。然后, 又利用了自回归积分滑动平均模型^[8] (ARIMA) 和基于经验模态分解^[9,10] (EMD) 的 LSTM 神经网络模型进行预测, 并将其预测结果与基于小波包分析的 LSTM 神经网络模型进行比较。其中, 训练数据为 2015 年 12 月 28 日到 2016 年 11 月 3 日的 295 组经过相空间重构的数据, 测试数据为 2016 年 11 月 4 日到 2017 年 1 月 15 日, 共计 72 组数据。

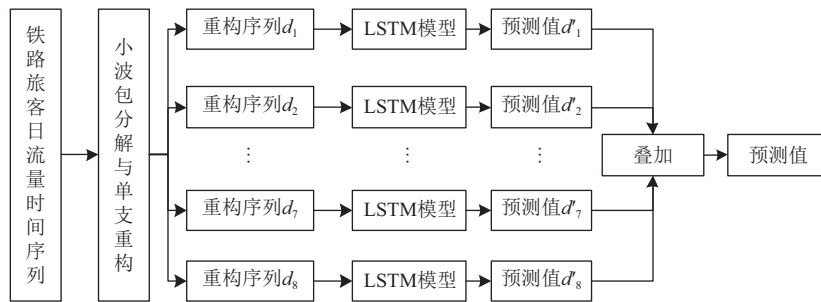


图4 WPA-LSTM 预测模型结构

4.1 小波包分解结果

首先, 对数据进行归一化处理, 以提高数据在训练过程中的收敛速度. 在这里使用 Max-Min 归一化方法, 对于序列中的最大值 x_{max} 和最小值 x_{min} , 对序列中所有元素 x_i 有:

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (8)$$

再挑选合适的小波包分解参数, 将原始信号合理的分解到恰当的多个频段上. 这里使用 Daubechies 小

波 1(db1), 对其进行小波包分解和单支重构, 其中单支重构后的情况如图 5-图 7 所示。

其中, 第一条曲线为原始序列曲线, 另外八条为 db1 小波包分解并进行重构后的 8 个不同尺度的序列. 这些序列能够捕捉到原始序列轻微的扰动, 将原始序列中突发性、局部性、隐藏性的信息保留在其中。

4.2 预测结果分析

各模型预测的结果如图 8 所示, 预测性能指标如表 1 所示。

通过对比拟合程度 (R 值) 可知, WPA-LSTM 模型最优, 其次是 ARIMA 模型, 最后是 EMD-LSTM 模型; 通过对比归一化均方误差 (NMSE) 可知, WPA-LSTM 模型最优, 其次是 ARIMA 模型, 最后是 EMD-LSTM 模型; 通过对比平均绝对百分比误差 (MAPE) 可知, WPA-LSTM 模型最优, 其次是 ARIMA 模型, 最后是 EMD-LSTM 模型; 综合考虑上述三项指标, WPA-LSTM 的预测精度最高。

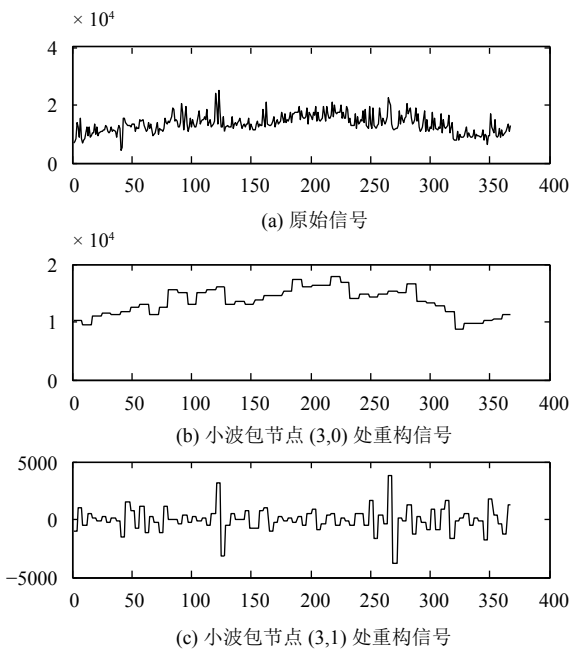


图5 旅客流量原始序列与 db1 分解后的信号对比图 1

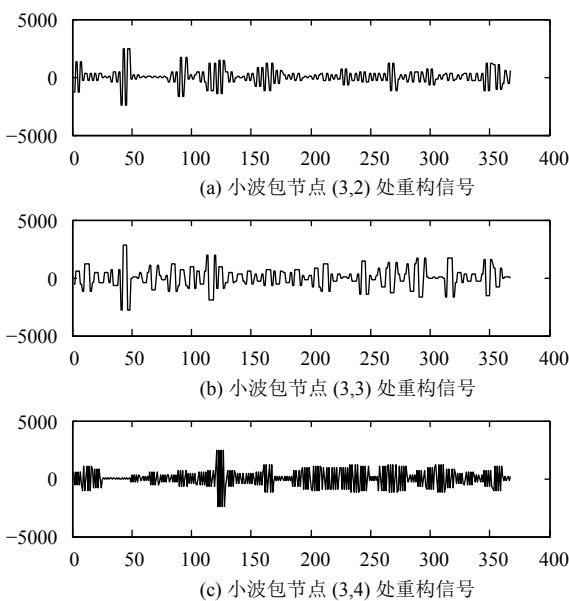


图6 旅客流量原始序列与 db1 分解后的信号对比图 2

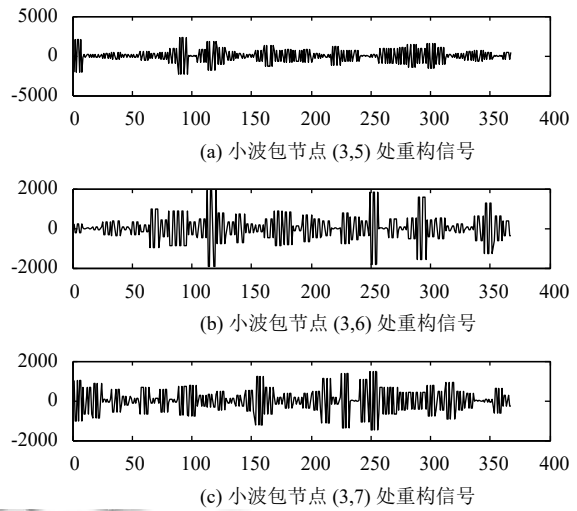


图7 旅客流量原始序列与 db1 分解后的信号对比图 3

4.3 误差分析

首先, 进行分日期的绝对预测误差对比, 如图 9 所示. 可以看出, ARIMA 模型在对 2017 年元旦附近, 特别是 2016 年 12 月 30 日这一天和 2017 年 1 月 2 日这一天的预测误差较大, 最大值为 7000 人; EMD-LSTM 模型也是在 2017 年 12 月 30 日这一天的预测误差最大, 最大值甚至达到了 7800 人, 而且 EMD-LSTM 模型的绝对误差明显多数高于其他两个模型; 而 WPA-LSTM 模型的绝对误差相对稳定, 在预测时段内没有特别剧烈的变化, 分布相对较平稳。

其次, 分析误差分布. 分别以 $\pm 10\%$ 和 $\pm 20\%$ 作为阈值上下限, 研究不同模型预测结果中过低预测 (低于阈值下限) 和过高预测 (高于阈值上限) 的日期数量占总日期数量的比值, 如表 2 所示。

由表 2 可以看出, 采用 $\pm 10\%$ 作为阈值时, 三种预测模型的过低预测和过高预测基本相当; 当采用 $\pm 20\%$ 作为阈值时, ARIMA 模型和 WPA-LSTM 模型的过低预测和过高预测基本相当, 而且明显优于 EMD-LSTM 模型. 综上所述, WPA-LSTM 模型的预测精度最高, 误差结构相对较好。

5 结论与展望

将基于小波包分析和长短时记忆神经网络融合的深度学习预测模型用于铁路旅客日流量的预测中, 实验结果证明, 比利用季节性模型 ARIMA 和基于经验模态分解的长短时记忆神经网络预测模型有着更好的逼近能力和容错性, 有效降低了预报误差, 提高了模型的预测精度, 具有较强的预报和泛化能力。

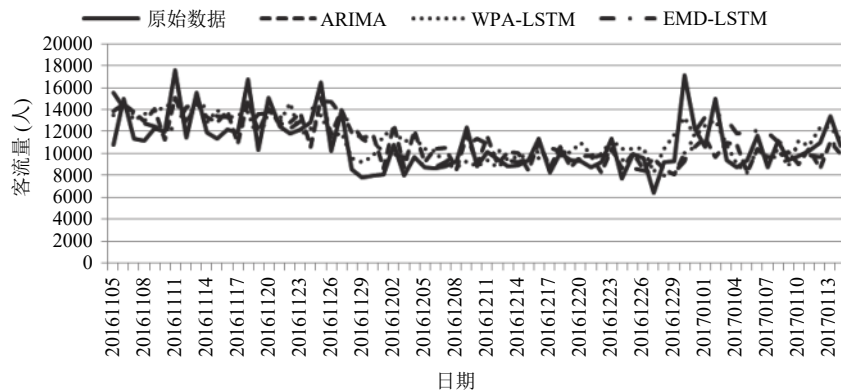


图8 各模型预测结果对比

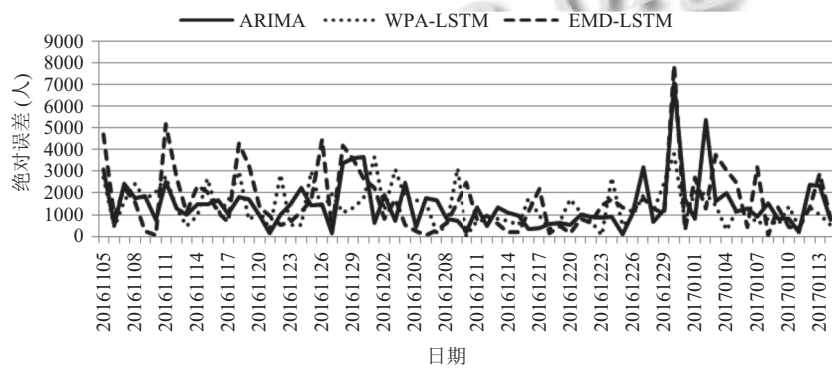


图9 预测值绝对误差对比

表1 各模型的预测性能指标

模型	R	NMSE	MAPE
ARIMA	0.6506	0.6069	0.1378
WPA-LSTM	0.7832	0.4178	0.1292
EMD-LSTM	0.5414	0.8118	0.1516

表2 各模型误差分布 (单位: %)

模型	阈值为±10%		阈值为±20%	
	过低预测 比例	过高预测 比例	过低预测 比例	过高预测 比例
ARIMA	38.03	19.72	12.68	4.23
WPA-LSTM	43.66	14.08	14.08	2.82
EMD-LSTM	32.39	18.31	25.35	5.63

参考文献

- 褚鹏宇, 刘澜. 基于变权重组合模型的铁路客流量短期预测. 计算机工程与应用, 2017, 53(4): 228-232, 262. [doi: 10.3778/j.issn.1002-8331.1506-0091]
- 罗向龙, 焦琴琴, 牛力瑶, 等. 基于深度学习的短时交通流预测. 计算机应用研究, 2017, 34(1): 91-93, 97.
- 杨祎玥, 伏潜, 万定生. 基于深度循环神经网络的时间序列预测模型. 计算机技术与发展, 2017, 27(3): 35-38, 43. [doi: 10.3969/j.issn.1673-629X.2017.03.007]

10.3969/j.issn.1673-629X.2017.03.007]

- 陈巧红, 孙超红, 余仕敏, 等. 基于递归神经网络的广告点击率预估研究. 浙江理工大学学报(自然科学版), 2016, 35(6): 880-885. [doi: 10.3969/j.issn.1673-3851.2016.11.014]
- 孙奎培, 王朝英. 小波分析和小波包在图像消噪中的应用. 通信技术, 2009, 42(1): 285-287.
- 田苗苗. 小波分析和小波包在图像压缩中的应用. 安徽科技学院学报, 2009, 23(5): 32-34.
- 任崇岭, 曹成铉, 李静, 等. 基于小波神经网络的短时客流量预测研究. 科学技术与工程, 2011, 11(21): 5099-5103, 5110. [doi: 10.3969/j.issn.1671-1815.2011.21.031]
- 刘建军, 廖闻剑, 彭艳兵. 两种时间序列模型在客流量预测上的比较. 计算机工程与应用, 2016, 52(9): 228-232, 270. [doi: 10.3778/j.issn.1002-8331.1406-0024]
- 秦喜文, 邢婷婷, 董小刚, 等. 基于EEMD的公路客流量波动性与周期性研究. 东北师大学报(自然科学), 2016, 48(1): 44-48. [doi: 10.16163/j.cnki.22-1123/n.2016.01.011]
- 何九冉, 四兵锋. EMD-RBF组合模型在城市轨道交通客流预测中的应用. 铁道运输与经济, 2014, 36(10): 87-92. [doi: 10.3969/j.issn.1003-1421.2014.10.017]