

# 基于 VMD-LSTM 轨道交通客流预测模型

黄海超, 陈景雅, 孙睿

(河海大学土木与交通学院, 江苏 南京 210098)

**摘要:** 客流量预测是城市智能交通系统的重要组成部分。为实现客流量的准确预测, 首先采用变分模态分解(VMD)将时序客流数据分解成不同时间尺度下的本征模态函数(IMF), 降低数据噪声对客流预测模型的影响, 再结合长短时记忆神经网络(LSTM)进行预测, 提出 VMD-LSTM 预测模型。采集明尼苏达州州际轨道交通客流数据对模型进行验证。结果表明: 相对传统 LSTM 预测模型, VMD 改进 LSTM 使平均绝对百分误差(MAPE)减少 8.38%, 均方根误差(RMSE)减小 256.99, 有效提高 LSTM 神经网络的预测精度与鲁棒性。

**关键词:** 轨道交通; 客流预测; 变分模态分解; 长短时记忆神经网络; 深度学习

**中图分类号:** U293.5

**文献标志码:** A

**本文引用格式:** 黄海超, 陈景雅, 孙睿. 基于 VMD-LSTM 轨道交通客流预测模型[J]. 华东交通大学学报, 2021, 38(1): 95-99.

## Rail Transit Passenger Flow Prediction Model Based on VMD-LSTM

Huang Haichao, Chen Jingya, Sun Rui

(College of Civil Engineering and Transportation, Hohai University, Nanjing 210098, China)

**Abstract:** Passenger flow prediction is an important part of urban intelligent transportation system. In order to realize accurate prediction of passenger flow, variational mode decomposition was adopted to decompose the time series into intrinsic mode function in different time scales, the long short-term memory neural network of deep learning was used to predict, and the VMD-LSTM prediction model was proposed. Data of Minnesota interstate subway passenger flow were collected to validate the model. The results show that compared with the traditional LSTM prediction model, the average absolute percentage error and the root mean square error decreases by 8.38% and 256.99% respectively after improved by VMD, the prediction accuracy and robustness of LSTM neural network are improved effectively.

**Key words:** rail transit; passenger flow prediction; variational mode decomposition; long short-term memory network; deep learning

**Citation format:** HUANG H C, CHEN J Y, SUN R. Rail transit passenger flow prediction model based on VMD-LSTM [J]. Journal of East China Jiaotong University, 2021, 38(1): 95-99.

由于轨道交通安全、准时和环保等特点, 已逐渐成为人们出行的首选交通方式, 基于数据挖掘技术的客流量预测是城市智能交通系统的重要组成部分, 不仅对人们出行和智能交通管控有着重要的指导意义, 也有助于相关部门提前进行运营规划和调度。

客流量预测模型主要分为以差分自回归移动平均模型<sup>[1]</sup>(ARIMA)和灰色模型(GM)为代表的参数模型与以支持向量机<sup>[2]</sup>(SVM)和神经网络<sup>[3]</sup>为代表的非参数模型。但传统模型收敛慢, 时效性低, 很难满足交通大数据情况下智能管控需求<sup>[4]</sup>。基于深

收稿日期: 2020-11-14

基金项目: 国家自然科学基金项目(52078190)

作者简介: 黄海超(1996—), 男, 硕士研究生, 研究方向为神经网络在智能交通领域的应用。E-mail: hhc123@hhu.edu.cn。

通信作者: 陈景雅(1967—), 女, 教授, 研究方向为互通立交安全评估。E-mail: 13805151397@139.com。

度学习的长短时记忆神经网络 LSTM (long short-term memory)不仅预测性能、收敛时间优于传统参数<sup>[5]</sup>与非参数<sup>[6]</sup>预测模型,偏差分布波动与预测鲁棒性也明显优于传统预测模型<sup>[7]</sup>;针对非平稳的客流量数据,经验模态分解<sup>[8]</sup>(EMD)、集成经验模态分解<sup>[9]</sup>(EEMD)被用于降低样本噪声干扰并有效提高预测精度。变分模态分解(variational mode decomposition, VMD)作为最新提出的分解算法,虽然在交通领域应用较少,但在电力荷载<sup>[10]</sup>、期货<sup>[11]</sup>等领域已证明其对非平稳信号良好的处理能力。

目前研究大多集中于短时客流量预测<sup>[12]</sup>(15 min内),然而过短时间的预测不能为人们出行规划预留充足时间,提供有价值的出行指导。为此,文章创新性地将 VMD 引入客流量预测领域,降低客流量信号的非平稳性,结合深度学习的 LSTM 进行预测,提出 VMD-LSTM 预测模型,实现对每小时轨道交通客流量的准确预测,提高 LSTM 神经网络的预测性能。

### 1 VMD-LSTM 神经网络预测模型构建

#### 1.1 VMD 分解原理

VMD 是一种全新的自适应、完全非递归时频信号分解算法,通过构建变分模型,将原始时间序列  $s(t)$  分解为不同的具有有限带宽的分量  $u_k(t)$ ,对应的中心频率为  $\omega_k$ ,交替迭代更新寻找约束变分模型的最优解,在此过程中实现本征模态函数 IMF 的有效分离,克服了 EMD 存在端点效应和模态分量混叠的问题,其算法可表示为<sup>[13]</sup>

1) 对每一个  $u_k(t)$  进行 Hilbert 变换得到单侧频谱

$$\left[ \delta(t) + \frac{j}{\pi t} \right] * u_k(t) \tag{1}$$

式中:  $\delta(t)$  表示 Dirac 分布函数; \* 表示卷积运算。

2) 通过将每个模式的分析信号乘以估计的中心频率  $e^{-j\omega_k t}$ , 将频谱移到基带

$$\left\{ \left[ \delta(t) + \frac{j}{\pi t} \right] * u_k(t) \right\} e^{-j\omega_k t} \tag{2}$$

3) 通过对解调信号梯度的 L2 正则化进行高斯平滑估计,得到各模态函数的带宽,约束变分模型可表示为

$$\min \left\{ \sum_k \left\| \partial_t \left[ \left[ \delta(t) + \frac{j}{\pi t} \right] * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \tag{3}$$

$$\sum u_k = s(t)$$

式中:  $\partial_t$  表示求偏导;  $k$  为分量的个数。

4) 通过引入二次罚因子  $\alpha$  和拉格朗日乘子  $\lambda$ , 保证  $s(t)$  为原始时间序列在高斯噪声的影响下信号的重构精度,将约束变分问题转化为无约束变分问题

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[ \left[ \delta(t) + \frac{j}{\pi t} \right] * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| s(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), s(t) - \sum_k u_k(t) \rangle \tag{4}$$

5) 通过交替方向乘法(ADMM)更新分量  $u_k(t)$ , 相应中心频率  $\omega_k$  及拉格朗日乘子  $\lambda$  求解,并找到扩展拉格朗日函数表达式的最小值点,其更新方法如下

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{s}(\omega) - \sum_{i < k} \hat{u}_i^n(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{5}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega} \tag{6}$$

$$W_{x_i} X_t = \hat{\lambda}^n(\omega) + \tau(\hat{u}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)) \tag{7}$$

式(5)中:  $\hat{s}(\omega)$ ,  $\hat{u}_k^n(\omega)$ ,  $\hat{\lambda}(\omega)$  分别为  $s(\omega)$ ,  $u_k^n(\omega)$ ,  $\lambda(\omega)$  的傅里叶变换;  $\omega$  为频率;  $n$  为迭代的次数;  $\tau$  为拉格朗日乘子的更新参数。

6) 重复步骤 5), 满足精度  $\varepsilon$  时停止迭代,从而得到  $k$  个窄带 IMF 分量。迭代条件可表示为

$$\sum_k \frac{\left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2}{\left\| \hat{u}_k^n \right\|_2^2} < \varepsilon \tag{8}$$

#### 1.2 LSTM 神经网络

LSTM 是深度循环神经网络的代表,通过门的开关实现时空记忆功能,能有效解决循环神经网络的梯度消失和爆炸问题。通过引入一个门控单元系统,可以控制何时忘记历史信息或使用新信息更新单元状态,使 LSTM 在解决非线性时间序列问题时效果极佳<sup>[14]</sup>,其过程可表达为

$$i_t = \sigma(W_{x_i} X_t + W_{h_i} H_{t-1} + W_{c_i} C_{t-1} + b_i) \tag{9}$$

$$f_t = \sigma(W_{x_f} X_t + W_{h_f} H_{t-1} + W_{c_f} C_{t-1} + b_f) \tag{10}$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{x_c} X_t + W_{h_c} H_{t-1} + b_c) \tag{11}$$

$$o_t = \sigma(W_{x_o} X_t + W_{h_o} H_{t-1} + W_{c_o} C_{t-1} + b_o) \tag{12}$$

$$H_t = o_t \tanh(C_t) \tag{13}$$

式中: $i, f, o, c$  分别为输入门  $i$ , 遗忘门  $f$ , 输出门  $o$ , 细胞状态  $c$  在  $t$  时刻的输出; $W_{x_i}, W_{x_f}, W_{x_o}, W_{x_c}$  分别为输入  $X_t$  与输入门  $i$ , 遗忘门  $f$ , 输出门  $o$ , 细胞状态  $c$  的权重; $W_{h_i}, W_{h_f}, W_{h_o}, W_{h_c}$  为隐藏层输出  $H_t$  与相应门的权重; $W_{c_i}, W_{c_f}, W_{c_o}$  为细胞状态输出  $C_t$  与相应门的权重; $b_i, b_f, b_o, b_c$  为偏置向量; $\sigma$  为激活函数。

### 1.3 VMD-LSTM 预测模型

VMD-LSTM 预测模型建立主要分为两部分。第一部分为数据预处理,首先将对原始数据中的重复、异常、缺省值进行清洗获得时序客流量数据。基于 VMD 分解原理,设定合适的超参数,将非平稳的每小时客流量数据分解为本征模态函数 IMF 和余量信号  $R(t)$ ,弱化样本噪声干扰;第二部分为分量预测,首先选择 LSTM 神经网络时间步长、输入层隐藏层神经元个数等基本参数,确定神经网络的训练方式与代价函数,获得网络拓扑结构。建立与分解分量个数相同的 LSTM 神经网络,对分量分别进行预测,获得仿真本征模态函数(SimIMF)最后将各预测值进行重构获得最终客流量预测值,并与真实客流量进行比较,评价模型性能。模型建立流程如图 1 所示。

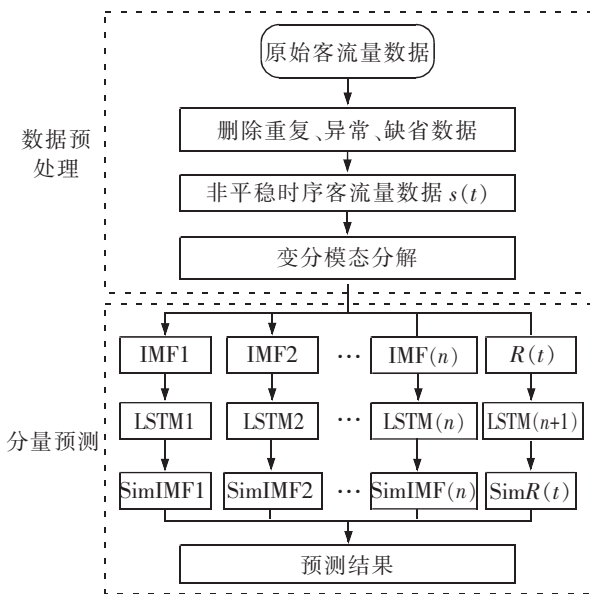


图 1 VMD-LSTM 神经网络模型流程图

Fig.1 VMD-LSTM neural network model flow chart

## 2 应用实例

### 2.1 数据预处理

采集明尼苏达州明尼阿波利斯至圣保罗 301 站台 2017 年每小时轨道交通客流量,明尼阿波利斯与圣保罗分别是明尼苏达州第一、二大城市,旅

客往来频繁,客流规模大;301 站台为圣保罗通往明尼阿波利斯及其他城市的主要站台,人员出行规律复杂,因此采用该站台作为研究对象。对轨道交通数据中的异常数据进行预处理。包括完全相同的记录、明显异常的记录及缺失的记录,数据清洗后共获得 10 605 条时间数据。由于数据样本较大,有效避免了机器学习过程中预测模型对个别样本的偏好,无需对样本进行抽样训练。

采用 VMD 算法对非平稳交通量时间序列数据进行分解时,由于 VMD 在轨道交通客流量预测领域应用较少,缺乏先验知识确定超参数。综合考虑重构误差与中心频率分布,二次罚因子  $\alpha$  取 1 000,分解模态数  $K$  取 11。分解得到 10 个 IMF 分量以及一个余量信号  $R(t)$ ,分解结果如图 2 所示。分量频率由高到低进行排列,更加清晰的呈现原始客流量数据在不同时间尺度下的变化规律,其中高频信号表示客流量变化的细节,含有大量样本噪声;低频部分则表征客流量的趋势,频率最低的余量信号  $R(t)$  也称为趋势项,表征全年轨道交通客流量的波动趋势。

### 2.2 神经网络模型构建

选择前 10 个月(8 880 条时间数据)作为训练集,最后两个月(1 725 条时间数据)用于测试模型性能。时间步长取 4,隐含层数取 1,隐含层神经元个数取 200;迭代次数取 250。初始学习率设为 0.01,同时为防止学习率过大,模型来回震荡,采用衰减法动态调整学习率,每迭代 50 次学习率衰减 50%。损失函数采用均方根误差,通过 AdamOptimizer 优化器对模型进行优化,得到 LSTM 神经网络的拓扑结构。

根据以上参数,建立 11 个 LSTM 神经网络,对分解获得的分量分别进行训练、预测,根据 VMD 算法对预测结果进行重构,得到最终每小时客流量预测值。并绘制总体预测残差,如图 3 所示,可以看出 LSTM 神经网络仅在初期前 1%数据范围内预测误差较大,随后残差迅速变小并趋于稳定,表现出良好的鲁棒性。

同时为验证 VMD-LSTM 神经网络模型的性能,分别与 LSTM 神经网络、随机森林(RF)、支持向量机(SVM)进行对比。对比模型与原模型采用相同的参数设置,RF 预测模型利用经验公式确定子树数目为 103,SVM 预测模型采用自适应算法搜索最佳超参数  $c, g$ ,预测结果如图 4 所示。

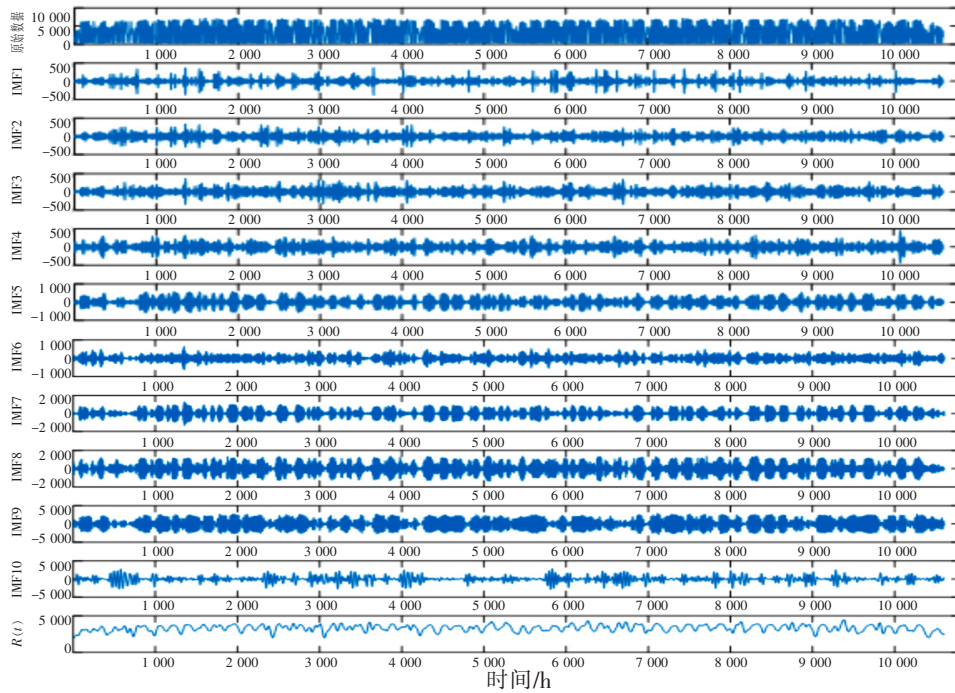


图 2 客流量数据 VMD 分解结果

Fig.2 VMD decomposition results of passenger flow sequence

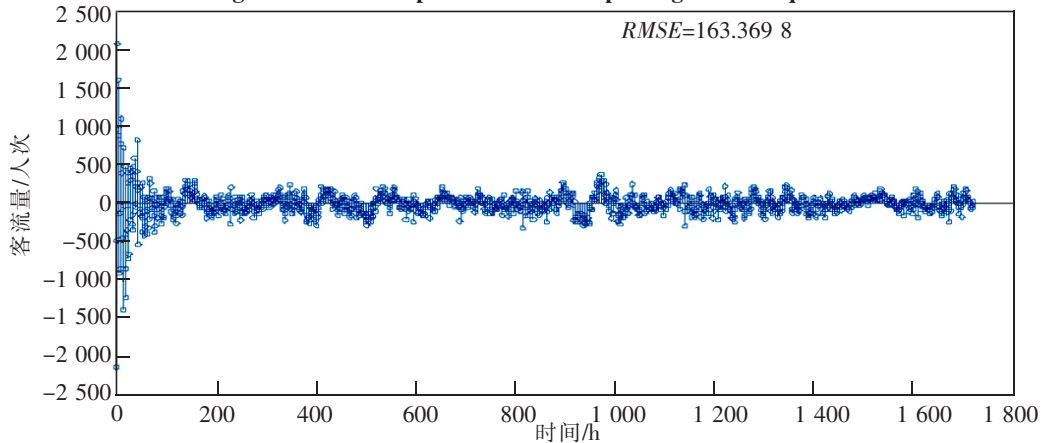


图 3 预测残差

Fig.3 Predicting residual

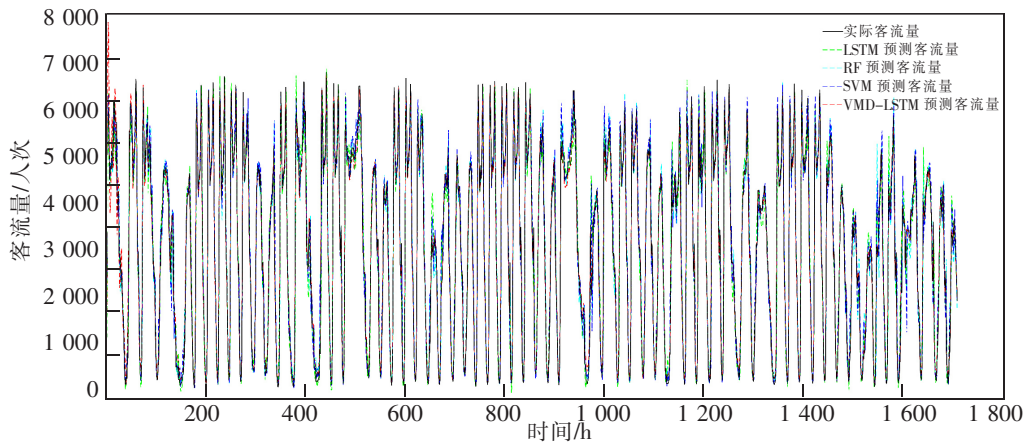


图 4 神经网络模型对比

Fig.4 Comparison of neural network models

平均绝对百分误差(MAPE)体现了真实值与预测值的相对偏差,可直接衡量预测结果的好坏,常被用于评价预测模型的优劣,但不能直观反映真实值与预测值的差值。均方根误差(RMSE)可直接体现真实值与预测值的绝对差值,且特大或特小误差对指标影响明显,可有效弥补 MAPE 的不足,相关系数(R)可用于表征真实值与预测值变化趋势的相似程度。分别采取以上指标对各预测模型性能进行评价,结果如表 1 所示。

表 1 经典预测模型性能对比

Tab.1 Performance comparison of classical prediction models

参数	LSTM	RF	SVM	VMD-LSTM
MAPE	15.22%	15.42%	13.48%	6.84%
RMSE	420.36	434.42	410.36	163.37
R	97.64%	97.72%	97.79%	99.67%

结合图 4 和表 1 可以看出,在轨道交通客流量数据样本足够大的情况下,所有预测模型均有极高的预测精度。传统的时间序列预测模型 RF,SVM 与基于深度学习的 LSTM 预测误差相差不大。但 LSTM 神经网络克服传统预测模型每次加入新数据,需要重新训练的缺陷,收敛速度也有较大提高,体现神经网络良好的综合性能。

当采用 VMD 优化 LSTM 后,预测误差显著降低,VMD-LSTM 模型的预测精度相对三种对比模型总体有了较大提高。对比传统时间序列预测模型 RF,SVM,MAPE 分别减小了 8.58%,6.64%,RMSE 分别减小了 271.05,246.99;相比 LSTM 神经网络,MAPE 减小了 8.38%,RMSE 减小了 256.99,R 值也有一定的提高。这是因为 VMD 将含噪声的客流量数据分解成模态分量与余量信号,将含噪声的高频信号单独分离出来,从而弱化样本噪声对总体预测精度的干扰,有效提高神经网络的预测精度;同时使非平稳客流量数据转换为多个平稳时间序列,LSTM 神经网络对这些平稳信号进行预测,进一步提高模型的鲁棒性。

### 3 结论

1) 采用 VMD 将含噪声的轨道交通客流量数据分解成具有原始数据局部特征的不同频率信号,降低噪声对模型的干扰,提出 VMD-LSTM 神经网络预测模型

2) 采用明尼苏达州轨道交通数据验证 VMD-LSTM 预测模型性能,结果表明 VMD 改进 LSTM 能显著提高神经网络的预测精度与鲁棒性,预测误差相对传统预测模型更小。

### 参考文献:

- [1] LI L C,WANG Y G,ZHONG G,et al. Short-to-medium term passenger flow forecasting for metro stations using a hybrid model[J]. KSCE Journal of Civil Engineering,2018,22(5):1937-1945.
- [2] 孟歌,王洪业,李丽辉,等. 基于 EMD 的 SVR 方法在铁路客流预测中的应用[J]. 铁路计算机应用,2020,29(4):28-32.
- [3] ZHANG J,GOU W H. Research on railway passenger flow prediction method based on GA improved BP neural network [J]. Journal of Computer and Communications,2019,7(7):283-292.
- [4] 蒲悦逸,王文涵,朱强,等. 基于 CNN-ResNet-LSTM 模型的城市短时交通流量预测算法[J]. 北京邮电大学学报,2020,43(5):9-14.
- [5] LIU L J,CHEN R C,ZHU S Z. Impacts of weather on short-term metro passenger flow forecasting using a deep LSTM neural network[J]. Applied Sciences,2020,10(8).
- [6] 李艳,彭春华,傅裕,等. 基于 CNN-LSTM 网络模型的风电功率短期预测研究[J]. 华东交通大学学报,2020,37(4):109-115.
- [7] LIU D Q,WU Z,SUN S R. Study on subway passenger flow prediction based on deep recurrent neural network[J]. Multi media Tools and Applications,2020(prepublish):1-14.
- [8] 赵阳阳,夏亮,江欣国. 基于经验模态分解与长短期记忆神经网络的短时轨道交通客流预测模型[J]. 交通运输工程学报,2020,20(4):194-204.
- [9] 傅晨琳,黄敏,沙志仁. 基于 EEMD-BP 方法的城市轨道交通进站客流短期预测[J]. 铁道运输与经济,2020,42(3):105-111.
- [10] ZHANG G,LIU H C,LI P L,et al. Load prediction based on hybrid model of VMD-mRMR-BPNN-LSSVM[J]. Complexity,2020(4):1-20.
- [11] ZHU Q,ZHANG F,LIU S,et al. A hybrid VMD - BiGRU model for rubber futures time series forecasting[J]. Applied Soft Computing,2019,84:13.
- [12] 李梅,李静,魏子健,等. 基于深度学习长短期记忆网络结构的轨道交通站短时客流量预测[J]. 城市轨道交通研究,2018,21(11):42-46,77.
- [13] DRAGOMIRETSKIY K,ZOSSO D.Variational mode decomposition[J]. IEEE Transactions on Signal Processing,2014,62(3):531-544.
- [14] WAN H Y,GUO S N,YIN K,et al. CTS-LSTM:LSTM-based neural networks for correlated time series prediction [J]. Knowledge-Based Systems,2020,191.