

DOI: 10.13382/j.jemi.B2003575

基于自适应 VMD-Attention-BiLSTM 的交通流组合预测模型^{*}

殷礼胜 孙双晨 魏帅康 田帅帅 何怡刚
(合肥工业大学 电气与自动化工程学院 合肥 230009)

摘 要:针对短时交通流量序列的非平稳性和随机性的特征,为提高短时交通流预测精度和收敛速度,提出一种基于自适应变分模态分解(VMD)和结合注意力机制层的双向长短时记忆网络(BiLSTM)的组合预测模型。首先,使用自适应变分模态分解将时空交通流量序列分解为一系列有限带宽模态分量,细化了交通流信息,降低了非平稳性,提升了建模的精确度;其次,利用结合注意力机制的双向长短时记忆网络挖掘分解后交通流量序列中的时空相关性,从而揭示其时空变化规律,从而进一步提升了建模精确度,并且利用改进 Adam 算法进行网络权值优化,以加速了预测网络的训练收敛速度;最后,将各模态分量预测值叠加求和作为最终交通流预测值。实验结果表明,使用模态分解的预测模型预测性能明显优于未使用模态分解的预测模型,同时自适应 VMD-Attention-BiLSTM 预测模型相较于 EEMD-Attention-BiLSTM 预测模型,均方根误差降低了 47.1%,该组合预测模型提升了预测精度,并且能够快速预测交通流量时间序列。

关键词:短时交通流预测;自适应变分模态分解;双向长短时记忆网络;注意力机制

中图分类号: U491.1 **文献标识码:** A **国家标准学科分类代码:** 580.2099

Traffic flow combination prediction model based on adaptive VMD-attention-BiLSTM

Yin Lisheng Sun Shuangchen Wei Shuaikang Tian Shuaishuai He Yigang
(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

Abstract:In view of the non-stationary and random characteristics of the short-term traffic flow sequence, in order to improve the short-term traffic flow prediction accuracy and model training speed, this paper proposes a combined prediction model based on adaptive variational modal decomposition (VMD) and bi-directional long-term memory network (BiLSTM) combined with attention mechanism. Firstly, the spatial and temporal traffic flow sequence is decomposed by the adaptive VMD method to a series of modal components with limited bandwidth, which can refine the traffic flow information, reduce non-stationarity, and improve the accuracy of modeling. Secondly, the spatio-temporal correlation in the short-time traffic flow sequence after decomposition is mined by BiLSTM combined with attention mechanism to reveal its spatio-temporal variation rules, which further improves the modeling accuracy. In addition, in order to accelerate the training convergence speed of the prediction network, the network weight optimization is carried out by the improved Adam algorithm. Finally, the predicted value of each modal component is superimposed as the predicted value of the final traffic flow prediction value. The experimental results show that the prediction performance of the model using modal decomposition is obviously better than that of the model without modal decomposition, and the RMSE of the self-adaptive VMD-Attention-BiLSTM prediction model is reduced by 47.1% compared with that of the EEMD-attention-BiLSTM prediction model. The combined prediction model improves the prediction accuracy and can quickly predict the traffic flow time series.

Keywords:short-term traffic flow prediction; adaptive variational modal decomposition; bi-directional long-term memory network; attention mechanism

0 引言

近年来,随着科学技术的发展和交通拥堵问题的出现,智能交通系统(ITS)应运而生,其研究也成为了当今热门研究课题之一^[1]。实时、准确短时交通流量预测是实现 ITS 对交通流控制和诱导的基础。短时交通流具有较强的随机性和非平稳性等特征,各专家针对其特点提出了很多智能预测模型和算法^[2-3]。殷礼胜等^[4]运用集合经验模态分解(EEMD)和最小二乘支持向量机(LSSVM)的组合模型,通过将非平稳的交通流序列分解为稳定固定模态分量,再分别用 LSSVM 建模预测,提升了模型预测的准确度,但忽略了路网的时空相关性对交通流预测的影响。徐先峰等^[5]运用卷积神经网络(CNN)和双向长短时记忆网络(BiLSTM)的组合模型,通过卷积神经网络和双向长短时记忆网络分别捕获路网的空间特征和交通流量时间周期特征,提升了模型预测的精确度,但忽略了交通流量序列的随机性以及模型训练收敛时间较长的问题。杜圣东等^[6]提出一种序列到序列时空注意力的预测模型,利用 LSTM 编码、解码器和注意力机制,自适应地挖掘学习交通流量时间序列中时空相关的内在规律性,但是忽略了城市路网交通的复杂性及交通流量序列的非平稳性对模型预测精度的影响。

从交通流量时间序列的非平稳性和随机性的内在规律角度出发,以上研究通过降低交通流量时间序列非平稳性或深度挖掘交通流量时间序列中的时空相关性都提升了预测模型的准确度,但未将降低交通流量序列非平稳性和深度挖掘交通流量序列时空相关性结合考虑。综上所述,本文尝试结合降低交通流量序列非平稳性和深度挖掘交通流量序列时空相关性两方面优势,提出了一种基于自适应差分模态分解(VMD)和结合注意力机制的双向长短时记忆网络的组合预测模型。首先,使用自适应差分模态分解将交通流量时间序列分解为一系列有限带宽模态函数分量,将交通流量时间序列分解成不同尺度下的交通流量平稳序列,充分体现交通流量时间序列数据趋势性分量、周期性分量及随机性分量等;在此基础上,分别为重组的各模态分解分量构建结合注意力机制的双向长短时记忆网络预测模型,为了加速预测网络的训练和全局最优解的寻优,引入了改进的 Adam 算法进行网络权值优化;最后,将各模态分量预测值叠加求和作为最终交通流预测值。通过将本文预测模型与 LSTM 模型、BiLSTM 模型、Attention-BiLSTM 模型和 EEMD-Attention-BiLSTM 预测模型的预测值的对比,以验证本文模型在预测精度上的提升。

1 基于交通流量时间序列的自适应变分模态分解

1.1 差分模态分解

VMD 作为一种新的信号分解方法,VMD 可把复杂信号分解为 K 个有限带宽本征模态函数(band limited intrinsic mode function, BIMF)。相较于 EEMD 的递归筛选模式,VMD 在获取各分量信号时,通过迭代搜寻变分模型的最优解来确定每个模态函数的中心频率和带宽,最后得到 K 个有限带宽本征模态函数,表现出更好的噪声鲁棒性^[7]。

将交通流量时间序列看成非平稳信号 f , 通过 Hilbert 变换,得到每个模态函数 $u_k(t)$, 得到其单边频谱。对各模态函数,通过对其对应的预估中心频率 $w_k(k=1,2,\dots,K)$ 的指数混叠 $e^{-jw_k t}$, 将每个模态的频谱调制到相应基频带。估算各模态信号带宽,可表示带约束条件的变分问题为

$$\begin{cases} \min_{\{u_k\}, \{w_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \\ \text{s. t.} \quad \sum_{k=1}^K u_k = f \\ \delta(t) = \begin{cases} 0, & t \neq 0 \\ \infty, & t = 0 \end{cases} \\ \int_{-\infty}^{+\infty} \delta(t) dt = 1 \end{cases} \quad (1)$$

式中: $\{u_k\} = \{u_1, \dots, u_K\}$, $\{w_k\} = \{w_1, \dots, w_K\}$, $\delta(t)$ 为脉冲信号。

通过引入拉格朗日乘法算子 $\lambda(t)$ 和二次惩罚因子 α , 将约束性变分问题转化为非约束性变分问题。

$$\begin{aligned} L(\{u_k\}, \{w_k\}, \lambda(t)) = \\ \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \\ \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \rangle \end{aligned} \quad (2)$$

通过交替更新 u_k^{n+1} 、 w_k^{n+1} 和 λ_k^{n+1} 来计算上式中的‘鞍点’。利用傅里叶等距变换,求得的模态分量 $\hat{u}_k^{n+1}$ 和中心频率 w_k^{n+1} 的最优解为:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha(w - w_k)^2} \quad (3)$$

$$w_k^{n+1} = \frac{\int_0^\infty w |\hat{u}_k(w)|^2 dw}{\int_0^\infty |\hat{u}_k(w)|^2 dw} \quad (4)$$

式中: $\hat{u}_k^{n+1}(w)$ 为当前剩余量 $\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w)$ 的维纳滤波; w_k^{n+1} 为当前模态函数谱的重心。

1.2 基于交通流量时间序列改进的自适应差分模态分解

由 VMD 原理可知,需要预先设置分解 BIMF 的个数 K ,其设置的合理性直接影响了最终分解结果^[8]。若设置 K 值较小,会造成模态分解不充分,得不到交通流量序列中的一些有限带宽的固有模态分量;若设置 K 值较大,会造成过度分解,干扰原交通流量序列中有效模态分量分析。

针对上述问题,本文尝试用一种基于信息熵的自适应差分模态分解算法。通过原交通流量序列和各模态分解余量间的信息散度差值来辨别相关程度,其相关性 E 计算如下:

$$E(Y_1, Y_2) = H(Y_1) - D_{KL}(Y_2 | Y_1) \quad (5)$$

式中: $H(Y_1)$ 表示原交通流量序列 Y_1 的熵; $D_{KL}(Y_2 | Y_1)$ 表示原交通流量序列 Y_1 与分解余量 Y_2 的 KL 散度和,其定义如式(6)。

$$D_{KL}(Y_2 | Y_1) = \sum_{x \in X} Y_2(x) \log \left(\frac{Y_2(x)}{Y_1(x)} \right) \quad (6)$$

由式(6)可知,原交通流量序列 Y_1 与分解余量 Y_2 的相关性越强, $D_{KL}(Y_2 | Y_1)$ 越小,则 $E(Y_1, Y_2)$ 越大。将分解余量与原交通流量序列进行相关性计算,并用下式进行归一化处理后,若 $\delta_i (i = 1, 2, \dots, K)$ 低于阈值时,则认为分解余量不再含有重要信息,原交通流量序列已被完全分解。

$$\delta_i = \frac{E_i}{\max(E_i)} \quad (7)$$

算法 1 改进自适应 VMD 具体算法步骤如下。

- 1) 初始化外循环变量 K 。
 - 2) 初始化模态函数集 $\{\hat{u}_k^1\}$ 、中心频率集 $\{w_k^1\}$ 、拉格朗日乘法算子 $\{\lambda^1\}$ 和 n 。
 - 3) 根据模态函数最优解式(3)和中心频率最优解式(4)更新 $\hat{u}_k^{n+1}$ 和 w_k^{n+1} 。
 - 4) 根据式(8)更新拉格朗日乘法算子,
- $$\lambda^{n+1}(w) = \lambda^n(w) + \tau \left[f(w) - \sum_k \hat{u}_k^{n+1}(w) \right] \quad (8)$$

式中: τ 为噪声容限参数。

5) 给定判别精度 $e > 0$, 若不满足式(9)的收敛条件,则 $n = n + 1$ 并返回步骤 3) 进行下一步内循环迭代;若满足则进行原交通流量序列与分解余量的相关性计算。

$$\sum_{k=1}^K \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < e \quad (9)$$

6) 分别用式(5)和(6)计算原交通流量序列和各模态分解余量的相关性 E 。

7) 根据式(7)进行归一化处理,若原交通流量序列与交通流量序列分解余量的归一化相关性值小于设定的阈值 δ , 则迭代结束并输出 K 个有限带宽的交通流量序列分量;若不满足,则 $K = K + 1$ 并返回步骤 2) 进行下一步迭代计算。

经改进自适应 VMD 分解后的交通流量时间序列与原交通流量时间序列相比具有更强的规律性。在交通流量时间序列进行分解的基础上,运用结合注意力机制的双向长短时记忆网络挖掘路网采样点之间的交通流量序列时空相关性,可提升预测精度。

2 结合注意力机制的双向长短时记忆神经网络

2.1 双向长短时记忆神经网络

交通流量序列实质是一种基于时间序列数据,其前后输入的交通流量数据是相互联系的,在 LSTM 中神经元的输出会在下一个时刻传递给自己^[9-10],其在前后相互联系的时间序列预测中具有独特的优势,且相较于传统的循环神经网络,可以有效解决梯度爆炸和梯度消失问题。典型 LSTM 单元^[11-13]主要由输入门、遗忘门和输出门构成,其结构如图 1 所示。

$$f_t = \sigma(\mathbf{w}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_f) \quad (10)$$

$$i_t = \sigma(\mathbf{w}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_i) \quad (11)$$

$$\hat{C}_t = \tanh(\mathbf{w}_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_c) \quad (12)$$

$$\mathbf{C}_t = f_t \cdot \mathbf{C}_{t-1} + i_t \cdot \hat{\mathbf{C}}_t \quad (13)$$

$$o_t = \sigma(\mathbf{w}_o \cdot [\mathbf{C}_{t-1}, \mathbf{x}_t] + b_o) \quad (14)$$

式中: f_t 为遗忘门; i_t 为输入门; o_t 为输出门; $\mathbf{x}_t$ 为交通流量时间序列向量; $\mathbf{h}_t$ 为隐含层输出向量; σ 为标准 sigmoid 激活函数; $\mathbf{C}_t$ 为记忆细胞; $\hat{\mathbf{C}}_t$ 为候选细胞信息向量; $\mathbf{w}_f, \mathbf{w}_i, \mathbf{w}_o, \mathbf{w}_c, b_f, b_i, b_o$ 和 b_c 为各权值矩阵和偏移量。

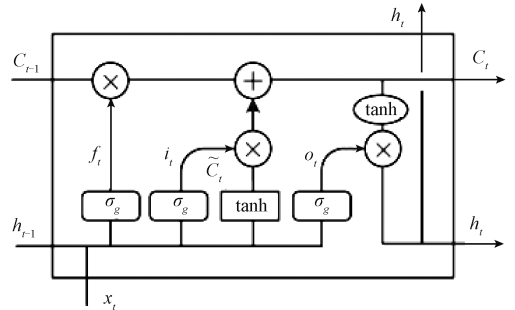


图 1 长短时记忆网络结构

Fig. 1 LSTM structure diagram

双向 LSTM 网络^[14-15]是由向前和向后两个单向

LSTM 网络堆叠而成,其中向前 LSTM 表示输入从开始到结束的前向交通流量序列进行预测,向后 LSTM 表示输入从结束到开始的反向交通流量序列进行预测。由于综合考虑了两个方向的交通流量序列信息,使模型预测的精度得到了提升,BiLSTM 的输出为:

$$h'_t = w_y \cdot [h_t^{(f)}, h_t^{(b)}] + b_y \quad (15)$$

式中: $h_t^{(f)}$ 和 $h_t^{(b)}$ 为前向和反向 LSTM 网络的输出向量; w_y 和 b_y 为权值和偏移量。

2.2 基于交通流量序列时空特性的注意力机制层

基于时空的交通流量历史数据集,其远近不同的采样节点在空间上与预测节点之间存在重要性差异。结合注意力机制^[16-19]的 BiLSTM 网络能够捕捉时空数据集中多采样点之间的时空相关性,避免了预测网络模型将不同路段的节点同样对待,忽略对时空中重要采样节点信息的特别关注。

将自适应 VMD 模态分解后体现交通流趋势分量、周期性分量和随机性分量等特征的模态分量重组,输入注意力机制层,通过计算概率的方式对关键交通流特征分量分配足够的权值^[14],从而突出关键信息的影响。

2.3 改进 Adam 算法

由于结合注意力机制的双向长短时记忆网络结构复杂,导致在处理大量交通流量时间序列数据时计算目标函数梯度不够准确,从而影响了预测的结果,因此需要采用优化算法进行权值优化。Adam 算法通过计算目标函数梯度的一阶、二阶矩估计,从而自适应设定相应的学习率^[20]。其中 Adam 算法计算双向长短时记忆网络模型参数 θ_t 每步迭代更新量为:

$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{n}_t} + \varepsilon} \quad (16)$$

式中: α 为学习率, $\hat{m}_t$ 为修正后一阶矩估计; $\hat{n}_t$ 为修正后二阶矩估计; ε 为常量 10^{-8} 。

本文尝试以 Adam 为基础,将 Adam 的当前迭代点处的搜索方向和上两个时刻迭代点处搜索方向加权求和作为下一步迭代的搜索方向来改进 Adam,以加快预测模型收敛速度,其中改进 Adam 算法计算网络模型参数 θ_t 每步迭代更新量为:

$$\theta_t = \theta_{t-1} - \left[\alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{n}_t} + \varepsilon} + \gamma \cdot (p_{t-1} - p_{t-2}) \right] \quad (17)$$

式中: γ 历史迭代所占权重; p_{t-1} 和 p_{t-2} 表示为上两个时刻迭代点处迭代更新量。

算法 2 改进 Adam 具体算法流程步骤如下。

1) 初始化学习率 $\alpha = 0.01$, 初始化历史迭代所占的权重 $\gamma = 0.5$ 。

2) 初始化一阶矩估计和其指数衰减速率分别为

$m_0 = 0$ 和 $\beta_1 = 0.9$, 二阶矩估计和其指数衰减速率分别为 $n_0 = 0$ 和 $\beta_2 = 0.9$ 。

3) 初始化改进 Adam 算法迭代更新量 $p_0 = 0$ 。

4) 计算目标函数梯度 g_t :

$$g_t = \nabla_{\theta} f_t(\theta_{t-1}) \quad (18)$$

式中: $\nabla_{\theta} f_t(\theta_{t-1})$ 表示第 t 个目标函数关于 θ 的偏微分。

5) 根据式 (19) 和 (20) 分别计算一阶矩估计 m_t 和二阶矩估计 n_t 。

$$m_t = m_{t-1} \cdot \beta_1 + g_t \cdot (1 - \beta_1) \quad (19)$$

$$n_t = n_{t-1} \cdot \beta_2 + g_t^2 \cdot (1 - \beta_2) \quad (20)$$

6) 根据式 (21) 和 (22) 分别计算一阶矩估计的偏差修正 $\hat{m}_t$ 和二阶矩估计的偏差修正 $\hat{n}_t$ 。

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (21)$$

$$\hat{n}_t = \frac{n_t}{1 - \beta_2^t} \quad (22)$$

7) 根据式 (23) 更新算法每步迭代更新量。

$$p_t = \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{n}_t} + \varepsilon} + \gamma \cdot (p_{t-1} - p_{t-2}) \quad (23)$$

8) 根据式 (24) 进行网络模型参数 θ_t 更新,重复以上步骤直到最大迭代次数。

$$\theta_t = \theta_{t-1} - p_t \quad (24)$$

3 基于交通流量序列时空特性的改进自适应 VMD-Attention-BiLSTM 组合预测模型

交通流量序列具有较强非平稳、随机性的特征,单一的预测模型很难得到较好的预测效果。由于自适应变分模态分解在非平稳交通流量序列分解的优势和结合注意力机制的长短时记忆网络在深度挖掘基于时空特性的路网变化规律的优势,最终得到较为精准的预测值。其具体训练及预测步骤如下。

1) 将预测点相邻路网范围内交通流量数据作为输入。假设交通网络由 n 个采样节点组成,使用时间步长为 t 的数据预测时间 T 的交通流量,输入可表示为:

$$X_T = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} = \begin{bmatrix} x_{T-t}^1 & x_{T-t+1}^1 & \cdots & x_{T-2}^1 & x_{T-1}^1 \\ x_{T-t}^2 & x_{T-t+1}^2 & \cdots & x_{T-2}^2 & x_{T-1}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{T-t}^n & x_{T-t+1}^n & \cdots & x_{T-2}^n & x_{T-1}^n \end{bmatrix} \quad (25)$$

2) 调用改进的自适应变分模态分解算法 1 对交通流量序列进行平稳化处理,将其分解为 K 个体现交通流量时间序列的趋势性特性、周期性特性及随机性特性的平稳模态分量 $BIMF_i^v (i = 1, \cdots, K)$, 降低交通流量时间序列中的非平稳性。

3) 将 $BIMF_i^n (i = 1, \cdots, K)$ 通过注意力机制层, 通过计算概率的方式对不同采样点交通流时空特征分量分配足够的权值, 从而捕捉路网中不同采样节点与预测采样节点的时空相关性。Attention 层在 t 时刻的权值系数 α_t 和输出 p_t 为:

$$e_t = \tanh(w_e v_t + b_e) \tag{26}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^n \exp(e_i)} \tag{27}$$

$$p_t = \sum_{i=1}^n \alpha_i v_i \tag{28}$$

式中: e_t 为 t 时刻输入交通流序列 v_t 的注意力概率分布值; w_e 和 b_e 为权值和偏置量。

4) 将通过注意力机制层的交通流量时间序列输入双向长短时记忆网络训练模型, 对交通流量时间序列进行时间特征的挖掘, 并调用改进 Adam 算法 2 对双向长短时记忆网络进行最优参数选取并加快预测模型训练速度。

5) 将双向长短时记忆网络输出通过全连接计算出最终的预测结果 y_t :

$$y_t = \sigma(w_g h'_t + b_g) \tag{29}$$

式中: w_g 和 b_g 为权重和偏置; h'_t 为双向长短时记忆网络

输出。

6) 利用构建的 Attention-BiLSTM 预测模型对测试集交通流量时间序列分解的各分量进行预测, 通过全连接层得到交通流量时间序列各分量预测值。

7) 将交通流量各个分量时间序列进行叠加得到交通流预测值。

Attention-BiLSTM 网络预测模型和自适应 VMD-Attention-BiLSTM 组合预测模型流程如图 2 和 3 所示。

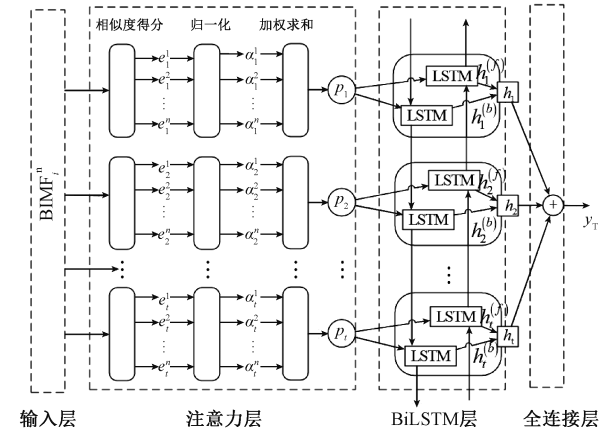


图 2 Attention-BiLSTM 网络预测模型
Fig. 2 Attention-BiLSTM network prediction model

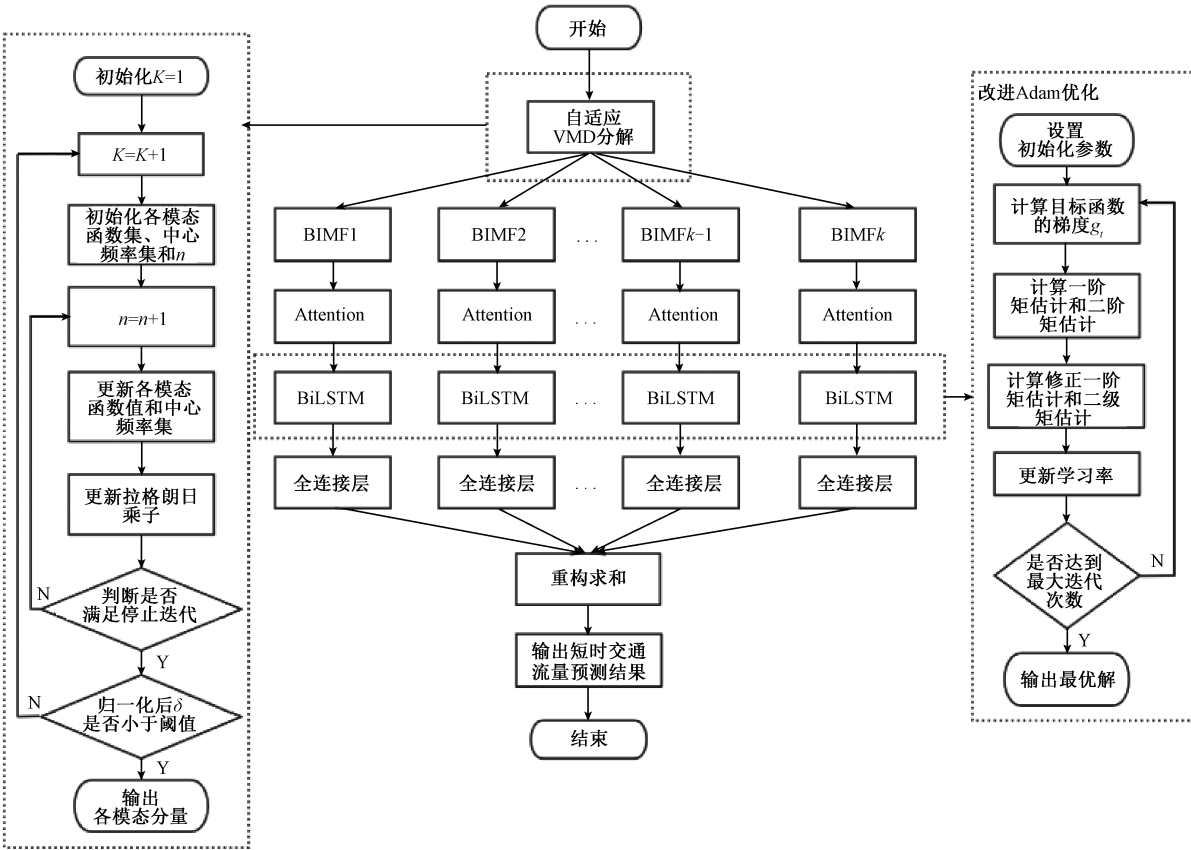


图 3 自适应 VMD_Attention_BiLSTM 组合预测模型流程
Fig. 3 Flow chart of adaptive VMD_Attention_BiLSTM combined prediction model

4 工程应用及结果分析

4.1 数据描述与评价指标

本文选取美国交通研究数据实验室提供的高速公路交通数据集,数据集共由 4 500 个采样节点构成,选取的观测节点编号为 S292、S293、S294、S343、S344、S345、S346、S713、S714,其中以 S345 作为预测观测点。观测点位置示意图如图 4 所示。取样时间为 2016 年 3 月 1~24 号 7:00~19:00 时段交通流数据作为实验数据,共有 9 个采样节点,交通流量序列采样间隔为 5 min,共有 3 456 个数据样本,将前 19 d 的交通流量数据作为训练集,后 5 d 的交通流量数据作为测试集。



图 4 观测点位置示意图

Fig. 4 Schematic diagram of the location of the observation point

为评估模型预测效果,本文采用平均绝对误差(mean absolute error,MAE)、平均绝对值百分误差(mean absolute percentage error,MAPE)和均方根误差(root mean squared error,RMSE)作为数据评价标准,其定义如下:

MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \tag{30}

MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% \tag{31}

RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \tag{32}

式中: $\hat{y}_t$ 为实际交通流量序列观测值; y_t 为交通流量序列预测值。

4.2 预测模型构建与训练

本文使用 TensorFlow 中 Keras 框架搭建实验所需模型,并且在 python3.6 环境中进行训练和预测。预测模型超参数设置如下:差分模态分解中的惩罚因子 α 设置为 2 000,噪声容限参数 τ 设置为 0.3,改进自适应差分模态分解算法 δ 阈值设置为 0.02,改进 Adam 算法历史迭代所占权重 γ 设置为 0.5,搜索终止条件最大迭代次数设置为 50,BiLSTM 隐含层神经元个数设置为 64。

预测模型优化目标为均方误差函数,目标函数为:

L_{MSE} = \frac{1}{n} \sum_{k=1}^n (\hat{y}_t - y_t)^2 \tag{33}

式中: $\hat{y}_t$ 为实际交通流量序列观测值; y_t 为交通流量序列预测值。

运用自适应 VMD 算法对交通流序列进行分解,得到 8 个 BIMF 分量,其各个分量余量与原交通流序列的归一化相关性 E 如表 1 所示。其中前 7 个分量归一化值大于 0.02。如图 5 所示,自适应 VMD 可将原交通流序列较好的分解为 7 个 BIMF 分量,其中心频率分别分布在 0.1~1 Hz 和 100~1 500 Hz,使各分量得到了有效的分解,避免了模态分解不充分和分解过度的问题。

表 1 各 BIMF 余量与原交通流序列的归一化相关性值

Table 1 The normalized correlation value of each BIMF margin and the original traffic flow sequence

BIMF0	BIMF1	BIMF2	BIMF3	BIMF4	BIMF5	BIMF6	BIMF7
1.000	0.887 2	0.789 0	0.746 6	0.799 1	0.789 8	0.757 8	0.016

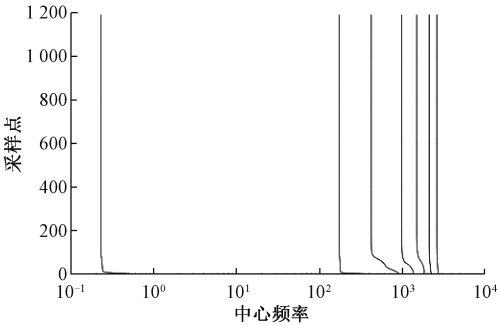


图 5 各分量中心频率分布

Fig. 5 Center frequency distribution diagram of each component

改进自适应 VMD 分解原交通流量序列分解为 7 个较为平稳的时间序列,其分解后如图 6 所示,首行为原始

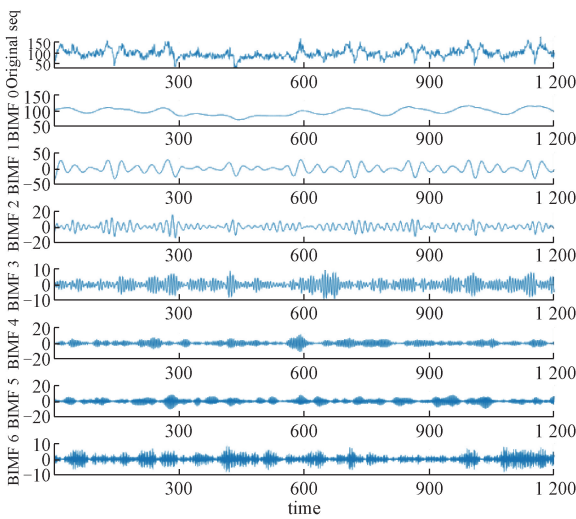


图 6 基于自适应 VMD 分解的交通流子序列

Fig. 6 Traffic flow subsequence based on adaptive VMD decomposition

交通流量序列,余下各分量依次从低频分量到高频分量。从分解的 BIMF 分量的特征可知,低频分量主要的交通流的趋势项,主要显现了其长期走势;一些较低频率分量具有较强的正弦波动特征,可认为是交通流的周期性分量;最后高频分量可认为是交通流的随机影响部分。

将测试集的差分模态分解分量分别输入相应的结合

注意力机制的双向长短时记忆网络预测模型中,运用结合注意力机制的长短时记忆网络针对各分量进行不同的时空规律挖掘预测,得到的曲线拟合效果在曲线拐角处具有一定的波动,但得到的预测结果曲线总体上具有较好的拟合效果,其预测结果如图 7 所示。

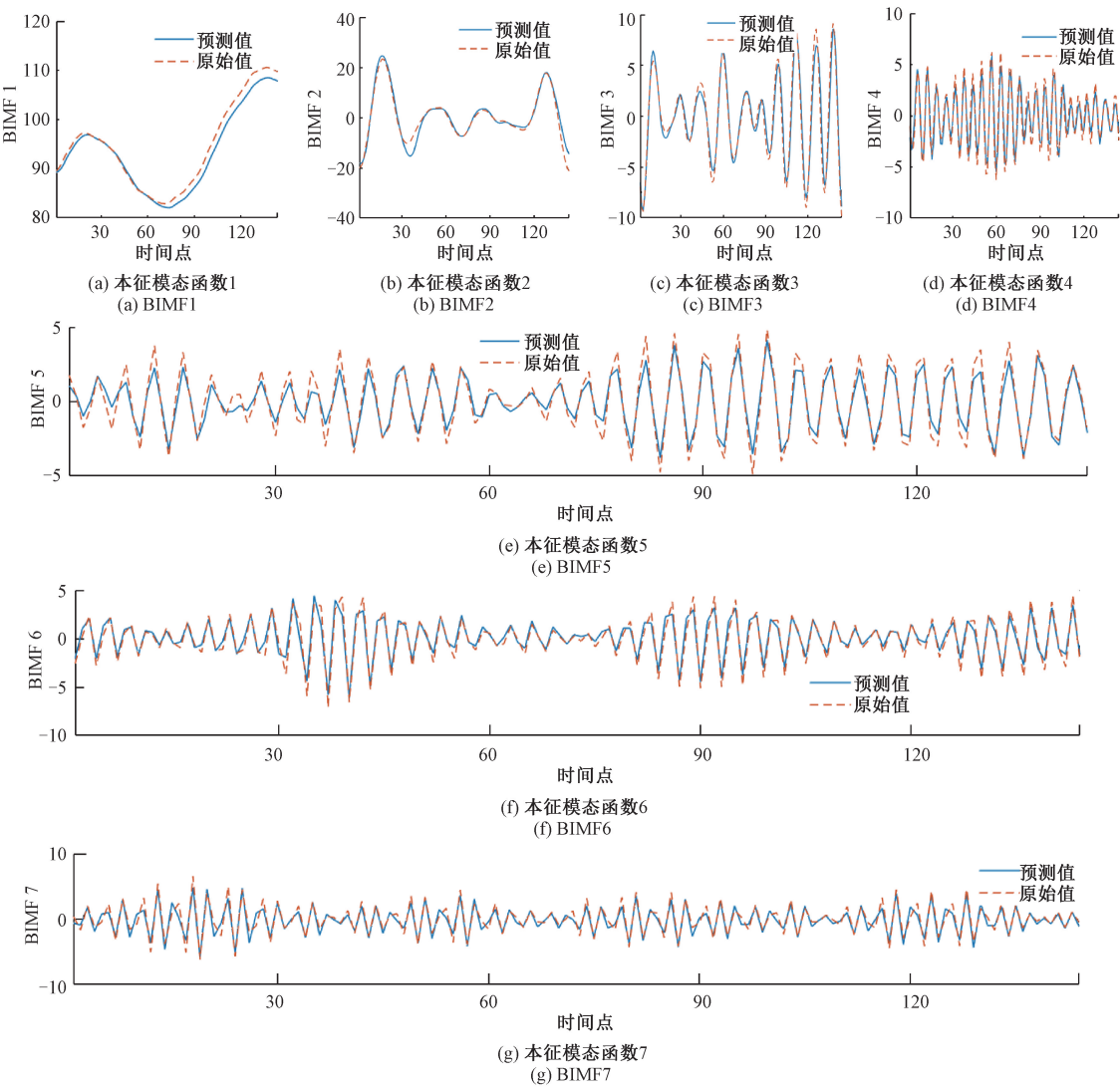


图 7 各分量预测结果曲线

Fig. 7 Curve of prediction results of each component

4.3 实验结果分析

为了验证本文提出模型的有效性,在原有训练数据集上建立 5 种模型,分别为 LSTM 模型、BiLSTM 模型、Attention-BiLSTM 模型、EEMD-Attention-BiLSTM 模型和本文预测模型。其中,各模型的评估指标如表 2 所示,预测值与实际值比较如图 8 和 9 所示。

表 2 不同预测模型的误差比较

Table 2 Error comparison of different prediction models			
预测模型	MAE	MAPE/%	RMSE
LSTM	8.69	8.96	11.15
BiLSTM	7.84	8.08	10.33
Attention-BiLSTM	4.27	4.61	5.44
EEMD-Attention-BiLSTM	3.34	3.79	4.25
VMD-Attention-BiLSTM	2.21	2.33	2.89

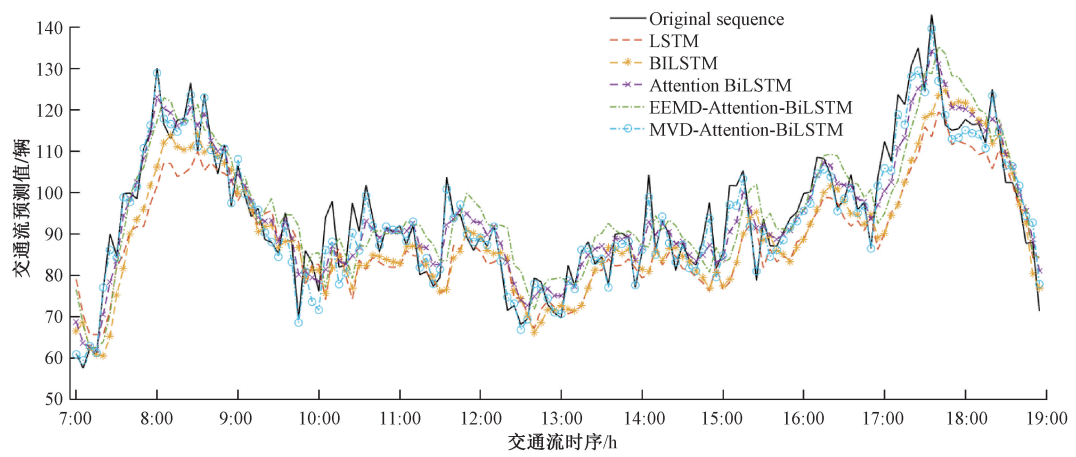


图 8 各预测模型预测结果和实际值的比较

Fig. 8 Comparison of prediction results and actual values of each prediction model

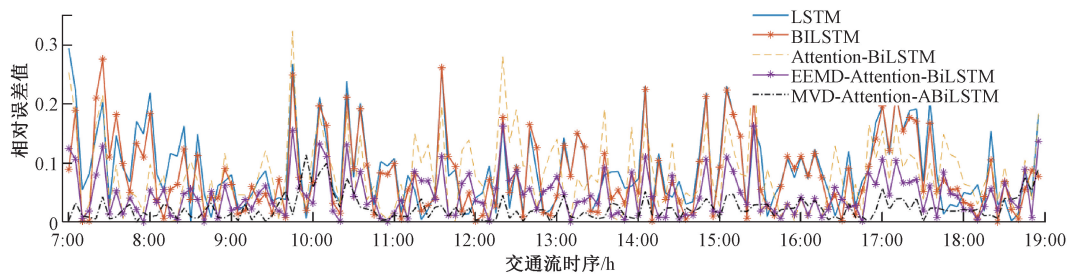


图 9 各预测模型的预测相对误差对比

Fig. 9 Comparison of prediction relative errors of various prediction models

如表 2 所示,LSTM 和 BiLSTM 模型的各项误差要明显大于经由模态分解后的预测模型,其中本文预测模型相较于 Attention-BiLSTM 预测模型,MAE 降低了 2.03,MAPE 降低了 2.28%,RMSE 降低了 2.55。对比使用 EEMD 模态分解的预测模型,使用改进自适应 VMD 模态分解的预测模型,其 RMSE 降低了 47.1%。根据图 8 和 9 可得,在短时交通流量时间序列数据急剧变化的时间段,13:00~16:00 时间段,使用差分模态分解的预测模型曲线拟合效果优于未使用模态分解的预测模型,且预测结

果拟合效果也优于 EEMD 模态分解的预测模型。为验证本文模型的稳定性的优劣,将美国加利福尼亚的高速公路数据集 PEMS4 和 PEMS8 分别输入相应的预测模型,其不同数据集的预测模型误差如表 3 所示。相比于 Attention-BiLSTM,本文预测模型在 PEMS4 和 PEMS8 数据集中平均绝对误差分别降低了 2.44 和 1.79,均方根误差降低了 26% 和 29%。本文模态分解预测模型在 3 个不同的数据集中均具有较好的预测精度和稳定性能。

表 3 不同数据集的预测模型误差对比

Table 3 Comparison of prediction model errors in different data sets

数据集	BiLSTM		Attention-BiLSTM		VMD-Attention-BiLSTM	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
PEMS4	8.36	13.47	5.90	6.10	3.46	4.83
PEMS8	6.25	12.58	5.46	6.14	3.67	4.76

如图 10 所示,改进 Adam 算法优化模型相较于 SGD 具有更快的收敛速度和搜寻全局最优解的能力,相比于 Adam 算法,改进 Adam 算法在第 5 代时就能找到全局最

优解,而 Adam 算法需要在第 8 代才能找到全局最优解,因此改进 Adam 算法的收敛速度更快且避免陷入局部最优解。

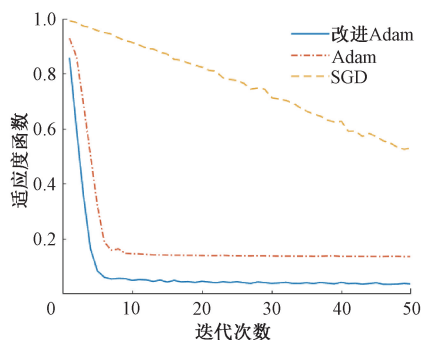


图 10 各算法迭代曲线

Fig. 10 Iterative curve of each algorithm

5 结 论

本文从交通流量序列非平稳和随机性特征出发,从提升交通流预测准确度角度,提出了基于改进自适应 VMD-Attention-BiLSTM 预测模型,结合时空特征交通流量序列进行预测仿真实验。通过实验仿真可知,其预测精确度明显优于未使用模态分解来降低交通流非平稳性特征的预测模型。相较于使用 EEMD 的预测模型,同样也在一定程度上提升了短时交通流预测准确度。这种组合模型为短时交通流量预测提供了一种新的想法,但本文并没有考虑天气和不同节假日因素对路段的影响,因此这将是本文研究的下一步内容。

参考文献

- [1] DARAGHMI Y A, YI C W, CHIANG T C. Negative binomial additive models for short-term traffic flow forecasting in urban areas [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15 (2): 784-793.
- [2] DAI G W, MA C X, XU X C. Short-term traffic flow prediction method for urban road sections based on space-time analysis and GRU [J]. IEEE Access, 2019, 7: 143025-143035.
- [3] 尹邵龙, 赵亚楠. 深度学习在城市交通流预测中的实践研究[J]. 现代电子技术, 2015(15):158-162.
YIN S L, ZHAO Y N. Practical research on deep learning in urban traffic flow prediction [J]. Modern Electronic Technology, 2015(15):158-162.
- [4] 殷礼胜, 唐圣期, 李胜, 等. 基于 EEMD-IPSO-LSSVM 的交通流组合预测模型[J]. 电子测量与仪器学报, 2019, 33(12):126-133.
YIN L SH, TANG SH Q, LI SH, et al. Combined traffic flow forecasting model based on EEMD-IPSO-LSSVM [J]. Journal of Electronic Measurement and Instrumentation,

2019, 33(12):126-133.

- [5] 徐先峰, 黄刘洋, 龚美. 基于卷积神经网络与双向长短时记忆网络组合模型的短时交通流预测[J]. 工业仪表与自动化装置, 2020(1):13-18.
XU X F, HUANG L Y, GONG M. Short-term traffic flow prediction based on the combined model of convolutional neural network and bidirectional long and short-term memory network [J]. Industrial Instrumentation and Automation, 2020(1):13-18.
- [6] 杜圣东, 李天瑞, 杨燕, 等. 一种基于序列到序列时空注意力学习的交通流预测模型[J]. 计算机研究与发展, 2020, 57(8):1715-1728.
DU SH D, LI T R, YANG Y, et al. A traffic flow prediction model based on sequence-to-sequence spatiotemporal attention learning [J]. Computer Research and Development, 2020, 57(8):1715-1728.
- [7] 张妍, 韩璞, 王东风, 等. 基于变分模态分解和 LSSVM 的风电场短期风速预测[J]. 太阳能学报, 2018, 39(1):194-202.
ZHANG Y, HAN P, WANG D F, et al. Short-term wind speed prediction of wind farm based on variational mode decomposition and LSSVM [J]. Acta Energia Sinica, 2018, 39(1):194-202.
- [8] 刘尚坤, 唐贵基, 王晓龙. 基于改进变分模态分解的旋转机械故障时频分析方法[J]. 振动工程学报, 2016(6):1119-1126.
LIU SH K, TANG G J, WANG X L. Rotating machinery fault time-frequency analysis method based on improved variational modal decomposition [J]. Chinese Journal of Vibration Engineering, 2016(6):1119-1126.
- [9] ZHAO Z, CHEN W, WU X, et al. LSTM network: a deep learning approach for short-term traffic forecast [J]. Iet Intelligent Transport Systems, 2017, 11(2):68-75.
- [10] 惠文珊, 李会军, 陈萌, 等. 基于 CNN-LSTM 的机器人触觉识别与自适应抓取控制[J]. 仪器仪表学报, 2019, 40(1):211-218.
HUI W SH, LI H J, CHEN M, et al. Tactile recognition and adaptive grasping control of robots based on CNN-LSTM [J]. Journal of Instrumentation, 2019, 40(1):211-218.
- [11] 陆文琦, 芮一康, 冉斌, 等. 智能网联环境下基于混合深度学习的交通流预测模型[J]. 交通运输系统工程与信息, 2020, 20(3):51-57.
LU W Q, RUI Y K, RAN B, et al. Traffic flow prediction model based on hybrid deep learning in intelligent network environment [J]. Transportation System Engineering and Information, 2020, 20(3):51-57.

- [12] 李锋, 陈勇, 向往, 等. 基于量子加权长短时记忆神经网络的状态退化趋势预测[J]. 仪器仪表学报, 2018, 39(7):220-228.
LI F, CHEN Y, XIAN W, et al. Prediction of state degradation trend based on quantum weighted long-term memory neural network [J]. Journal of Instrumentation, 2018, 39(7):220-228.
- [13] 魏昱洲, 许西宁. 基于 LSTM 长短期记忆网络的超短期风速预测[J]. 电子测量与仪器学报, 2019, 33(2):69-76.
WEI Y ZH, XU X N. Ultra-short term wind speed prediction based on LSTM long and short term memory network [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(2):69-76.
- [14] 周意乔, 徐昱琳. 一种基于双向 LSTM 的复杂环境下实时人体姿势识别方法[J]. 仪器仪表学报, 2020, 41(3):192-201.
ZHOU Y Q, XU Y L. A real-time human posture recognition method based on bidirectional LSTM in complex environment [J]. Chinese Journal of Scientific Instrument, 2020, 41(3):192-201.
- [15] WU Y, TAN H, QIN L, et al. A hybrid deep learning based traffic flow prediction method and its understanding [J]. Transportation Research, 2018, 90(MAY):166-180.
- [16] RAN X, SHAN Z, FANG Y, et al. An LSTM-based method with attention mechanism for travel time prediction[J]. Sensors, 2019, 19(4):861.
- [17] 闻畅, 刘宇, 顾进广. 基于注意力机制的双向长短时记忆网络模型突发事件演化关系抽取[J]. 计算机应用, 2019, 39(6):1646-1651.
WEN CH, LIU Y, GU J G. Extraction of evolutionary relations of emergencies based on the model of bidirectional long-short memory network based on attention mechanism [J]. Computer Applications, 2019, 39(6):1646-1651.
- [18] 孔子迁, 邓蕾, 汤宝平, 等. 基于时频融合和注意力机制的深度学习行星齿轮箱故障诊断方法[J]. 仪器仪表学报, 2019, 40(6):224-230.
KONG Z Q, DENG L, TANG B P, et al. Fault diagnosis

method of planetary gearbox in deep learning based on time-frequency fusion and attention mechanism [J]. Chinese Journal of Scientific Instrument, 2019, 40(6):224-230.

- [19] DO L N N, VU H L, VO B Q, et al. An effective spatial-temporal attention based neural network for traffic flow prediction [J]. Transportation Research, 2019, 108(10):12-28.
- [20] 赵会茹, 赵一航, 郭森. 基于互补集合经验模态分解和长短期记忆神经网络的短期电力负荷预测[J]. 中国电力, 2020, 53(6):48-55.
ZHAO H R, ZHAO Y H, GUO S. Short-term power load forecasting based on complementary ensemble empirical mode decomposition and long-short-term memory neural network [J]. China Electric Power, 2020, 53(6):48-55.

作者简介



殷礼胜, 1994 年于安徽大学获得学士学位, 2004 年于广西大学获得硕士学位, 2007 年于重庆大学获得博士学位, 现为合肥工业大学副教授, 主要研究方向为通信信道建模、交通流预测。

E-mail: yls20000@163.com

Yin Lisheng received his B. Sc. degree in 1994 from Anhui of University, M. Sc. degree from Guangxi University in 2004, Ph. D. degree from Chongqing University in 2007. Now he is an associate professor at Hefei University of Technology. His main research interests include communication channel modeling and traffic flow forecast.



孙双晨(通信作者), 2019 年于郑州轻工业大学获得学士学位, 现为合肥工业大学硕士研究生, 主要研究方向为通信信道建模、交通流预测。

E-mail: sensencen2050@163.com

Sun Shuangchen (Corresponding author) received his B. Sc. degree from Zhengzhou University of Light Industry in 2019. Now he is a M. Sc. candidate at Hefei University of Technology. His main research interest includes traffic flow forecast.