

doi:10.3969/j.issn.1003-3114.2023.04.023

引用格式:顾潮,肖婷婷,丁飞,等.基于模块化交通流组合预测模型[J].无线电通信技术,2023,49(4):761-772.

[GU Chao,XIAO Tingting,DING Fei,et al. Traffic Flow Combination Prediction Model Based on Modularization [J]. Radio Communications Technology,2023,49(4):761-772.]

基于模块化交通流组合预测模型

顾潮¹,肖婷婷²,丁飞¹,周启航¹,赵芝因¹

(1.南京邮电大学物联网学院,江苏南京210003;

2.南京信息工程大学电子与信息工程学院,江苏南京210044)

摘要:短时交通流预测是智能交通系统的核心能力组件之一,为城市交通管理、交通控制和交通引导提供智能决策支撑。针对交通路网交通流呈现的非线性、动态性和时序相关性,提出一种基于模块化的交通流组合预测模型ICEEMDAN-SSA-BiGRU。采用改进的基于完全自适应噪声集合经验模态分解(Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise,ICEEMDAN)方法对交通流非线性时间序列进行分解,获取本征模态分量(Intrinsic Mode Functions,IMF);利用双向门控循环单元(Bi-directional Gate Recurrent Unit,BiGRU)挖掘交通流量序列中的时序相关性特征;基于动态自适应 t 分布变异方法改进麻雀搜索算法(Improved Sparrow Search Algorithm,ISSA),实现对BiGRU网络权值参数的迭代寻优,避免了短时预测结果陷入局部最优;基于公开PeMS数据集对短时交通流预测性能进行评估与验证。实验结果表明,所提组合模型的短时交通流预测性能优于10个传统模型,改进后的交通流量预测平均绝对误差(Mean Absolute Error,MAE)指标接近10.98,平均绝对值百分误差(Mean Absolute Percentage Error,MAPE)指标接近10.12%,均方根误差(Root Mean Square Error,RMSE)指标接近12.42,且在不同数据集下所提模型具有较好的泛化性能。

关键词:短时交通流预测;完全自适应噪声集合经验模态分解;本征模态分量;麻雀搜索算法;双向门控循环单元

中图分类号:TP391

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1003-3114(2023)04-0761-12



Traffic Flow Combination Prediction Model Based on Modularization

GU Chao¹,XIAO Tingting²,DING Fei¹,ZHOU Qihang¹,ZHAO Zhiyin¹

(1. School of Internet of Things,Nanjing University of Posts and Telecommunications,Nanjing 210003,China;

2. School of Electronic and Information Engineering,Nanjing University of Information Science and Technology,Nanjing 210044,China)

Abstract: Short-term traffic flow prediction is one of the core competence components of intelligent transportation system, which provides intelligent decision support for urban traffic management, traffic control and traffic guidance. In this paper, Iceemdan-Isa-Bigr, a modular combined traffic flow prediction model, is proposed based on the nonlinear, dynamic and temporal correlation of traffic flow in the traffic network. Firstly, an Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) method based on fully adaptive noise set empirical Mode decomposition was used to decompose the nonlinear time series of traffic flows and obtain the intrinsic mode component. Secondly, bidirectional gated cycle unit is used to explore temporal correlation characteristics of traffic flow sequence. Then, based on the dynamic adaptive distributed variation method, the SSA algorithm is improved to achieve iterative optimization of the weight parameters of Bi-directional Gate Recurrent Unit (BiGRU) network, which avoids short-term prediction results falling into local optimal. Finally, the short-time traffic flow prediction performance is evaluated and verified based on open PeMS data set. Experimental results show that the short-time traffic flow prediction performance of the proposed combined model is better than that of the 10 traditional models. The improved Mean Absolute Error (MAE) index is close to 10.98, the Mean Absolute Percentage Error (MAPE) index is close to 10.12%, and the Root Mean Square Error (PMSE) index is close to 12.42. The proposed model has bet-

收稿日期:2023-03-16

ter generalization performance under different data sets.

Keywords: short-term traffic flow prediction; complete ensemble empirical mode decomposition with adaptive noise; intrinsic mode function; sparrow search algorithm; BiGRU

0 引言

随着城市化的发展,城市人口和机动车数量逐渐增加,道路利用率趋于饱和状态,尤其是早晚高峰时段,形成了城市“交通病”。为解决这一问题,研究人员提出智能交通系统(Intelligent Traffic System, ITS)^[1],通过计算机、通信、传感、控制、电子等技术手段,集成道路、汽车、驾驶员等各种资源。大数据与人工智能技术的发展,ITS 已经开始向数据驱动型 ITS 转变^[2]。其中,短期交通流预测是未来 ITS 的核心能力组件之一,为交通管理、交通控制和交通引导提供决策依据,为智慧出行提供智力支持。

现有交通流预测模型主要包括两类:统计模型和机器学习模型^[3]。统计模型可分为线性理论模型和非线性理论模型。线性理论模型包括自回归移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA)及其变体^[4-6]。非线性理论模型包括 K 近邻非参数回归(K-Nearest Neighbors, KNN)^[7]、卡尔曼滤波器(Kalman Filter, KF)^[8]、支持向量机(Support Vector Machine, SVM)^[9]和人工神经网络(Artificial Neural Network, ANN)^[10]。ITS 的建设与发展,针对道路感知能力正在获得有力提升,如地磁线圈、雷达检测器、视频监控以及车辆 GPS 数据等,为交通流预测提供了大量基础数据。由于数据多源、异质以及区域离散分布,采用传统线性和非线性方法进行短时交通流预测存在技术挑战。基于机器学习的深度学习^[11-12]适应于大样本量交通数据集的特征提取与分类识别,目前受到业界的广泛关注。考虑到深度网络相较于单层网络更能挖掘物理量的时空特征,李静宜等人^[13]对长短期记忆(Long Short-Term Memory, LSTM)神经网络架构进行深度设计,采用 3 层 LSTM 神经网络结构,融入遗传算法对 LSTM 层数、Dense 层数、隐藏层神经元数和 Dense 层神经元数进行寻优,能更好捕获路网交通流的波动特性,从而实现更准确的短时交通流预测精度。

由于交通流的内在变化规律是复杂的,通过深度网络捕获长期的历史信息占用了大量的训练时间和内存。因此研究人员对交通流时间序列进行分解,简化预测模型的结构,全面有效地提取特征。Rilling 等人^[14]提出了经验模态分解(Empirical Mode Decomposition, EMD),将信号中不同尺度的趋势或波动逐步分解为一系列具有不同频率的本征模态分量(Intrinsic Mode Functions, IMF)。理论上,非线性和随机性的信号可以被分解;然而,传统的 EMD 分解不完整,会引起混合和假模态的问题。因此,研究人员在 EMD 基础上进行改进与优化,取得了很好的效果。例如,Wei 等人^[15]将 EMD 和反向传播神经网络(Back Propagation Neural Network, BPNN)相结合,选择与原始数据高度相关的模态以提高预测效率,表现出显著的性能。同样,Chen 等人^[16]采用集成经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)对交通流时间序列进行分解,去除高频模态的同时引入 LSTM,对左侧重构模态进行预测。因为每个本征模态分量在时间序列中都起着重要作用,过早放弃某些分量会导致交通流特征信息不足。针对这一问题, Lu 等人^[17]通过 XGBoost 方法对经过完全自适应噪声集合经验模态分解(Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN)后的各车道交通流 IMF 进行预测,有效提取了其中的先验特征。Wang 等人^[18]将 CEEMDAN 与最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)相结合,有效地提高了非线性非平稳公路交通流量的预测精度。Huang 等人^[19]提出了通过 K-means 算法将 CEEMDAN 分解的交通流 IMF 进行聚类,并结合 BiLSTM 进行预测,有效降低了交通流量序列中的波动性与非平稳性。虽然上述方法模态混合问题在一定程度上得到了解决,但残余噪声和伪模态仍然存在。此外,训练数据规模较小或内存使用较高,依然存在难以捕捉交通流的深层变化特征的问题,尚需要进一步研究。

从交通流量序列的非线性、非平稳性及时序相关性特征考虑,现有研究未将三者全面考虑,且残余噪声和伪模态仍然存在。对此,本文将 CEEMDAN 能够细化交通流量时间序列非平稳性优势^[20]、麻雀搜索算法(Sparrow Search Algorithm, SSA)收敛速度快和寻优能力强的特点^[21]、双向门控循环单元(Bi-directional Gate Recurrent Unit, BiGRU)深度挖掘交通流序列中的时序相关性的优点^[22]三者结合,提出一种改进完全自适应噪声集合经验模态分解(Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, ICEEMDAN)和改进麻雀搜索算法(Improved Sparrow Search Algorithm, ISSA)优化 BiGRU 的组合模型。

1 基于 CEEMDAN 理论

EMD 分解是一种经典的自适应方法,用于解决非线性和非光滑信号问题。在交通流数据中,由于受到外部因素的影响,采集到的数据会包含更多的噪声,因此需要进行信号处理和去噪。EMD 分解能够将信号分解成多个 IMF,每个模态分量表示一个具有不同幅度和频率的波形,将数据分解成多个模态分量可以更好地观察到不同时间尺度上地变化趋势和周期性。但是 EMD 分解得到的模态分量存在模态混合,模态混合问题地出现导致时频分布不准确,使得一些模态分量失去了物理意义。EEMD 在原始信号的基础上加入高斯白噪声,解决了模式混合问题。但分解后的高斯白噪声无法消除,分解的完整性较差,导致算法重构误差较大。为了克服这些问题,Torres 等人^[23]提出了 CEEMDAN 旨在每个分解过程中加入自适应高斯白噪声来改进 EEMD,提高了 EEMD 分解的完整性的同时减少了重构误差及降低了计算成本。基于 CEEMDAN 分解的交通流时间序列数据的实现步骤如下:

步骤 1 对原始交通流量时间序列数据 x 第 j 次添加高斯白噪声,则交通流时间序列可表示为:

$$x^{(j)} = x + \beta_j v^{(j)}. \quad (1)$$

步骤 2 采用 EMD 方法对交通流数据进行分解,得到第一个模态分量 $IMF_1^{(j)}$,均值为 IMF_1 以及第一个残差分量 Res_1 :

$$IMF_1 = x - IMF_1^{(j)}, \quad (2)$$

$$Res_1 = x - IMF_1. \quad (3)$$

步骤 3 根据步骤 2 可以计算得出第二个模态分量 $IMF_2^{(j)}$,均值为 IMF_2 和第二个残差分量 Res_2 :

$$IMF_2 = \frac{1}{N} \sum_{j=1}^N E_1(Res_1 + \beta_1 E_1(v^{(j)})), \quad (4)$$

$$Res_2 = Res_1 - IMF_2. \quad (5)$$

步骤 4 以此类推,计算第 k 个残差分量 Res_k 和第 $k+1$ 个模态分量 IMF_{k+1} :

$$Res_k = Res_{k-1} - IMF_k, \quad (6)$$

$$IMF_{k+1} = \frac{1}{N} \sum_{j=1}^N E_1(Res_k + \beta_k E_k(v^{(j)})). \quad (7)$$

步骤 5 重复上述步骤,直到残差分量无法继续分解,即为单调函数或不超过两个极值点,最终残差分量可表示为:

$$Res = x - \sum_{k=1}^N IMF_k. \quad (8)$$

式中: x 为原始交通流的时间序列数据, IMF_k 为分解后得到的第 k 个 IMF, $v^{(j)}$ 为分解过程中在 $j(j=1, 2, \dots, N)$ 时刻添加的满足标准正态分布的高斯白噪声信号, β_j 为原始交通流时间序列数据 x 分解各阶段的信噪比, $E_k(\cdot)$ 表示 EMD 算法分解的第 k 个 IMF。

上述函数共同构成了原始信号在不同时间尺度上的特征,清晰地显示了原始交通流时间序列的变化趋势,有效地降低了预测误差。

2 SSA-BiGRU 短时交通流预测理论

2.1 SSA 原理

SSA 是基于麻雀觅食和反捕食行为的群体智能优化算法^[24],其中,麻雀种群中的个体被分为 3 种不同类型:生产者、觅食者和侦查者。生产者具有高能量储备,强大的探索能力和广阔的探索空间,并负责为整个种群寻找食物丰富的觅食区域。当麻雀发现捕食者时,生产者需要引领其他个体到达安全区域,以避免捕食者的攻击。假设麻雀搜索算法的种群规模为 N ,且在 d 维空间寻找最优解,则生产者的位置更新方程如下:

$$X_{id}^{iter+1} = \begin{cases} X_{id}^{iter} \times \exp\left(\frac{-i}{\alpha \times \max Iter}\right), & R_2 < ST \\ X_{id}^{iter} + Q \times L, & R_2 \geq ST \end{cases}, \quad (9)$$

式中: $iter$ 表示当前迭代次数, X_{id}^{iter+1} 表示第 i 只麻雀在第 d ($d=1, 2, \dots, dim$) 维的位置, $\alpha \in (0, 1]$ 为随机数, $\max iter$ 表示最大迭代次数, $R_2 \in [0, 1]$ 和 $ST \in [0.5, 1]$ 分别表示警戒值和安全阈值, Q 是服从正态分布的随机数, L 是一个 $1 \times dim$ 的行向量, 初始化所有元素为 1。

觅食者始终跟随生产者获取高质量的食物并增加能量储备。一些觅食者监视生产者并与其竞争食物。当觅食者的能量储备低时, 它们将离开群体自己寻找食物以生存。觅食者的位置更新方程式为:

$$X_{id}^{iter+1} = \begin{cases} Q \times \exp\left(\frac{X_{id}^{iter} - X_{id}^{iter+1}}{t^2}\right), & i > \frac{n}{2} \\ X_{id}^{iter+1} + |X_{id}^{iter+1} - X_{id}^{iter+1}| A^+ \times L, & \text{otherwise} \end{cases}, \quad (10)$$

式中: X_{id}^{iter} 表示当前全局最差位置, n 表示种群中个体数量, X_{id}^{iter+1} 表示由生产者发现的全局最佳位置, A 是一个 $1 \times dim$ 维度的行向量, 其元素被随机分配为 1 或 -1, A^+ 是 A 的逆矩阵, 即 $A^+ = A^T (AA^T)^{-1}$ 。

在麻雀种群中, 一些个体扮演侦察者的角色。这些个体能够探测到捕食者的威胁并向其他个体发出警报以避免危险。在模拟实验中, 假设这样的个体占总种群的 10%~20%, 它们的初始位置被随机分配。侦察者的位置更新方程为:

$$X_{id}^{iter+1} = \begin{cases} X_{id}^{iter} + \beta \times (X_{id}^{iter+1} - X_{id}^{iter}), & f_i \neq f_g \\ X_{id}^{iter} + K \times \left(\frac{X_{id}^{iter} - X_{id}^{iter}}{|f_i - f_w| + \varepsilon} \right), & f_i = f_g \end{cases}, \quad (11)$$

式中: X_{id}^{iter} 表示当前全局最优位置; β 是步长控制因子, 是一个随机数, 其遵循平均值为 0、方差为 1 的正态分布; K 是一个在 $[-1, 1]$ 的随机数, f_i 表示当前个体的适应度值 (目标函数值), f_g 和 f_w 分别表示当前全局最优和最差适应度值, ε 是一个非常小的数, 以避免分母为 0 的情况。

2.2 BiGRU 原理

递归神经网络以序列数据为输入, 按序列演化为方向进行递归, 所有循环单元都由链连接^[25]。Cho 等人^[26]提出了门控循环单元 (Gate Recurrent Unit, GRU), 该方法具有较少的收敛参数, 其本质使用 GRN 模块单元代替了 RNN 的隐藏单元, 有效地解决了传统 RNN 由于短期记忆而导致的梯度消失问题。GRU 结构如图 1 所示。

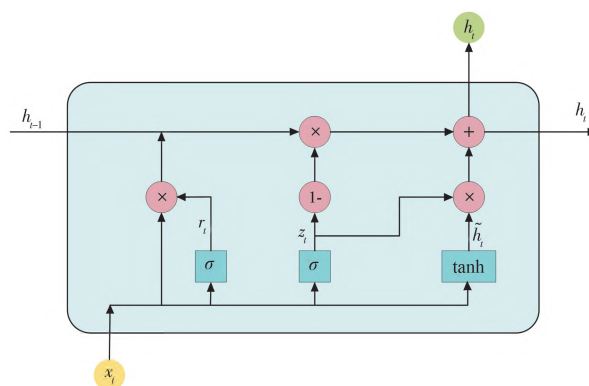


图 1 GRU 网络结构

Fig. 1 GRU networks structure

GRU 模块单元通过当前节点的输入 x_t 和上一节点的输出 h_{t-1} 计算复位门控状态 r_t 和更新门控状态 z_t :

$$r_t = \sigma(w_r \cdot [h_{t-1}, x_t]), \quad (12)$$

$$z_t = \sigma(w_z \cdot [h_{t-1}, x_t]), \quad (13)$$

式中: x_t 为 t 时刻输入数据, h_{t-1} 为前一时刻的隐藏状态信息, $[\]$ 表示 w_r 和 w_z 两个向量的连接, 即权重矩阵, 表 $\sigma(\cdot)$ 示 sigmoid 函数。

得到信号后, 首先通过 h_{t-1} 和 r_t 相乘得到复位内存数据, 然后与 x_t 进行拼接, 最后使用 tanh 激活函数将数据缩放到 $[-1, 1]$, 得到当前隐藏状态信息 \tilde{h}_t :

$$\tilde{h}_t = \tanh(w_h \cdot [r_t \cdot h_{t-1}, x_t]), \quad (14)$$

式中: $\tanh(\cdot)$ 表示双曲正切激活函数。

由于遗忘和记忆是同时进行的, 将之前得到的更新门 z_t 作为遗忘门, $1-z_t$ 作为输入门, 最终得到 GRU 的输出为:

$$h_t = z_t \cdot \tilde{h}_t + (1-z_t) \cdot h_{t-1}, \quad (15)$$

式中: h_t 为 GRU 网络在 t 时刻的输出, 门控信号 $z_t \in [0, 1]$, 其值越接近 0, 数据越被遗忘; 相反, 它越接近 1, 保留的数据就越多。

考虑到 GRU 状态是单向传递的, 这种情况下神经网络的映射输出仅基于时间数据的正向信息。本文采用 BiGRU 模型将正反向传播机制^[27]与 GRU 相结合, 使其较单向 GRU 模型挖掘更多交通流量序列信息。如图 2 所示, BiGRU 模型通过正、反向 GRU 隐藏状态信息 \vec{h}_t 和 \overleftarrow{h}_t 线性计算得到最终的交通流量预测值 y_t :

$$y_t = w_t \vec{h}_t + v_t \overleftarrow{h}_t + b_t, \quad (16)$$

式中: w_t 为正向 GRU 网络输出层权值系数, v_t 为反向 GRU 网络输出层权值系数, b_t 为 t 时刻双向门控循环单元中双隐态对应的偏置项。

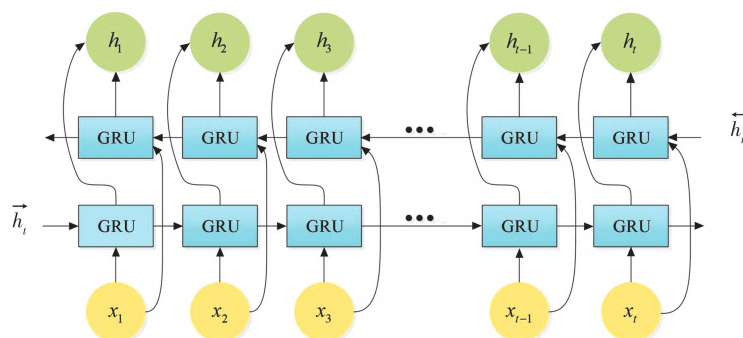


图2 BiGRU网络结构

Fig.2 BiGRU networks structure

3 ICEEMD-ISSA-BiGRU 短时交通流预测模型构建及算法

3.1 ICEEMDAN 原理

考虑到 CEEMDAN 方法仍然存在残留噪声和伪模态,使得误差在迭代过程中逐步积累,影响预测模型的训练和性能。本文提出了 ICEEMDAN,该算法有两个改进点:① 估计信号叠加噪声的局部均值,并将其与当前残差分量的平均值差异定义为主模态,减少了分解模式分量中存在的残留噪声;② 提取第 k 个模态分量,使用 $E_k(v^{(j)})$ 替换白噪声,减少模态重叠。具体步骤如下^[24]:

步骤1 设 $x^{(j)} = x + \beta_0 E_1(v^{(j)})$, 计算第一个残差分量:

$$Res_1 = M(x^{(j)}), \quad (17)$$

式中: $M(\cdot)$ 为模态分量的局部均值。

步骤2 计算第一个模态分量 IMF_1 :

$$IMF_1 = X - Res_1. \quad (18)$$

步骤3 第二个残差分量为 $Res_1 + \beta_1 E_2(v^{(j)})$ 的均值,第二个模态分量 IMF_2 定义为:

$$IMF_2 = Res_1 - Res_2 = Res_1 - M(Res_1 + \beta_1 E_2(v^{(j)})). \quad (19)$$

步骤4 同理,第 k ($k = 3, \dots, K$) 个残差分量 Res_k 表示为:

$$Res_k = M(Res_{k-1} + \beta_{k-1} E_k(v^{(j)})). \quad (20)$$

步骤5 最终得到 ICEEMDAN 算法的第 k 个 IMF_k :

$$IMF_k = Res_{k-1} - Res_k. \quad (21)$$

步骤6 执行步骤4,计算第 $k+1$ 个残差分量。

3.2 改进 SSA 优化 BiGRU 网络算法

BiGRU 网络采用传统的梯度下降法迭代网络参数,在梯度下降过程中易出现预测精度不足的问题^[28],因此本节通过两方面的改进以提升组合模型预测精度及收敛速度 k :

① 引入寻优能力及迭代速度在新型群体智能算法中较优的麻雀搜索算法对双向门控循环网络进行参数择优;

② 基于动态自适应 t 分布变异方法对 SSA 进行改进以缓解麻雀搜索算法易陷入局部最优解的问题。

在群体智能优化算法中,变异可以提高种群的多样性,帮助脱离局部极值。其中,高斯变异算子和柯西变异算子是比较常见的变异算子。研究表明,高斯分布和柯西分布具有不同的概率分布特性,高斯变异算子具有更好的局部开发能力,而柯西变异算子具有更好的全局探索能力^[29]。文中提出的动态自适应 t 分布是一种典型的标准统计分布,结合了高斯变异算子和柯西变异算子的优点,其形状随自由参数 n 的值而变化。当 $n=1$ 时, t 分布是柯西分布;当 $n \rightarrow \infty$ 时, t 分布类似于高斯分布。

由于 t 分布算子的变性能与变异比例因子和自由度参数相关^[30],所以本文在变异比例因子和 t 分布自由度参数中都引入了迭代次数。麻雀的变异比例因子和改进后生产者的位置公式如下:

$$\sigma = \frac{2(e^{-\frac{Iter}{\max Iter}})}{e-1}, \quad (22)$$

$$x'_i = x_i + x_i \sigma_i t(Iter), \quad (23)$$

式中: e 是自然对数的底数, $Iter$ 是当前迭代次数, $\max Iter$ 是最大迭代次数, x_i 和 x'_i 分别表示变异前和变异后第 i 只麻雀的位置, $t(Iter)$ 是以 $Iter$ 为自由度参数的 t 分布。

如图3所示,在迭代的开始阶段,雀群需要在广阔的搜索空间中寻找目标。此时,突变尺度因子 σ 很大, t 分布接近于柯西分布,使其具有更好的全局

探索能力。随着迭代次数的增加, t 分布的变异逐渐演变为高斯分布的同时突变尺度因子 σ 适当地减小, 使得雀群缓慢地接近最优的全局解。可以看出, 通过将迭代次数引入到自由度参数和变异比例因子 σ 中, 动态自适应 t 分布变异实现了对变异幅度的非线性自适应控制。最终, 动态自适应 t 分布变异方法提高了算法的全局和局部探索能力。由此得出 ISSA 优化 BiGRU 网络步骤如下:

步骤 1 初始化 BiGRU 和 ISSA 的超参数, 定义第 q 个 BiGRU 交通流量子网络待优化参数集合 W_q , 当前迭代次数 $Iter=1$, 当前权值参数编号 $h=1$, 最大迭代次数 $\max Iter$;

步骤 2 随机初始化 N 只麻雀的属性, 即为 N 只麻雀随机赋予 0~1 的随机数;

步骤 3 采用式(12)~式(16)对交通流量模态分量进行预测, 由式(24)计算结果与真实值间的 MAE, 即 ISSA 迭代过程中的适应度函数 f_i , 以修正得到第 q 个 BiGRU 预测子网络的第 h 个权值参数:

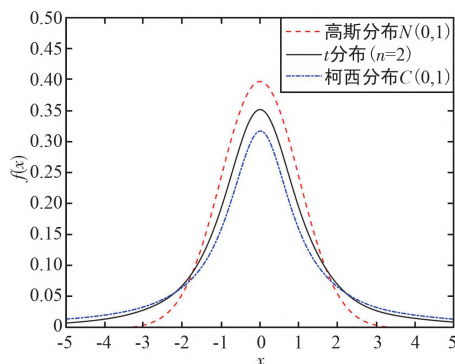
$$f_i = \sqrt{\frac{1}{n} \sum_{i=1}^n (himf_i^h - ximf_i^h)^2}, \quad (24)$$

式中: $himf_i$ 为 IMF 中第 i 个真实模态分量值, $himf_i^h$ 为训练第 h 个权值时输出的第 i 个交通分量预测值;

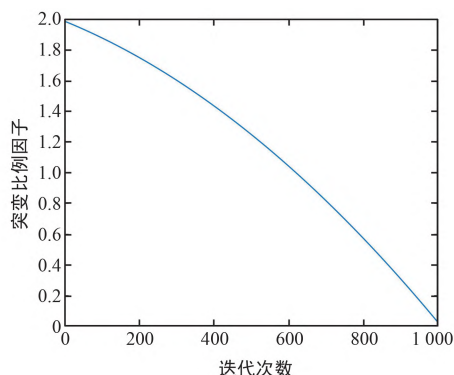
步骤 4 根据式(23)更新生产者的位置, 并由式(10)和式(11)更新觅食者和侦察者的位置;

步骤 5 更新全局最优适应度值 $f_g = f_i$ 和历史最优位置 $Xb_d^{Iter} = x_i'$ 。此时若 $Iter < \max Iter$, 则令 $Iter = Iter+1$ 返回步骤 4 执行; 否则, 将全局最优位置所对应的麻雀所赋予的随机数作为第 q 个 BiGRU 网络中第 h 个权值参数的最优权值参数;

步骤 6 若 h 未达到权值参数总数, 令 $h=h+1$ 后返回步骤 2 继续执行; 否则, 输出 BiGRU 预测子网络中的最优参数集合 W_q^* 。



(a) 柯西分布、 t 分布和高斯分布概率密度曲线



(b) 突变尺度因子曲线

图 3 概率密度分布与突变尺度因子关系对比图

Fig. 3 Comparison diagram of the relationship between probability density distribution and abrupt scale factor

4 短时交通流预测组合建模

交通流量数据具有较强的非线性、非平稳及时序相关性, 使用单一的预测方法难以达到理想的效果。因而采用 ICEEMDAN 算法对交通流量时间序列数据进行分解, 以 BiGRU 的权值参数为优化对象, 结合 ISSA 确定 BiGRU 权值参数的最优值, 构建 ICEEMDAN-ISSA-BiGRU 组合模型, 对城市短期交通流进行准确预测。模型结构框图如图 4 所示。

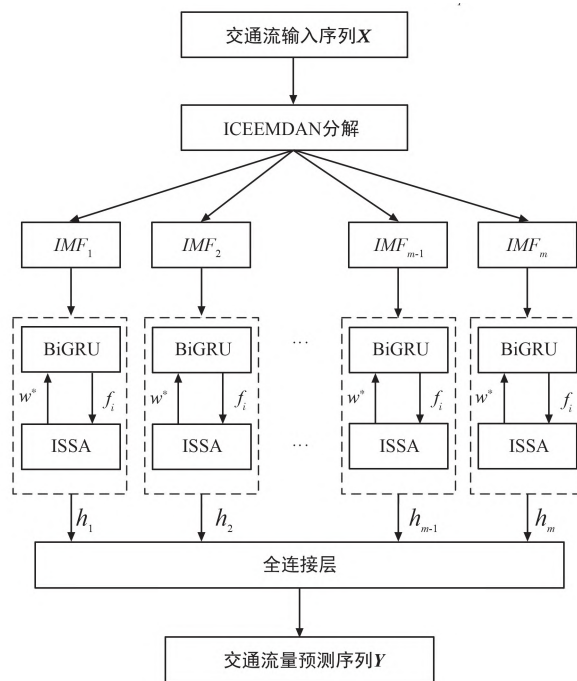


图 4 ICEEMDAN-ISSA-BiGRU 短时交通流预测模型结构

Fig. 4 ICEEMDAN-ISSA-BiGRU structure of short-term traffic flow prediction model

具体步骤如下:

步骤1 采用3.1节步骤将待预测交通流量序列 $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$ 分解,得到 m 组体现路网趋势性、周期性的交通流量模态分量 IMF;

步骤2 采用3.2节步骤分别对各 BiGRU 网络权值反复训练,最终得到最优参数集合 W^* ,及各性能最优的 BiGRU 预测子模型;

步骤3 采用步骤2训练完成的 BiGRU 预测子模型分别对交通流测试集模态分量进行预测,得到交通流量测试集模态分量的预测序列 $\mathbf{H}_q = (h_1, h_2, \dots, h_n)^T$,其中 n 表示交通流量数据总数;

步骤4 由式(8)可知,交通流量真实值由各交通模态分量及残差分量于等系数下相加而成,因此,总交通流量预测序列 \mathbf{Y} 由各交通模态分量的预测序列 \mathbf{H}_q 相加得到,即 $\mathbf{Y} = \sum_{q=1}^m \mathbf{H}_q = (y_1, y_2, \dots, y_n)^T$ 。

5 实验与分析

5.1 数据选取和实验参数设置

本文采用美国加州 Irvine 市 405-N 高速公路车流量探测器(编号:VDS-1209092)获得的 PeMS 数据集^[30],选择 PeMS 数据集中 2019 年 5 月 1—7 日样本数据,该样本数据中包括每 5 min 的交通流量、

速度和占有率等。数据集被分为两部分:① 选择 2019 年 5 月 1—6 日的数据当作训练集,训练本文模型;② 选择 2019 年 5 月 7 日的交通流量数据调整模型参数。随后,使用训练好的模型预测 2019 年 5 月 8 日的交通流量,并将预测值与 2019 年 5 月 8 日的真实值进行比较,评价模型性能。

考虑到作为预测模型输入的各种参数的维数不同且变化较大,使得模型的学习率可能会受到数据范围的限制而减缓收敛速度,进一步降低模型的稳定性和准确性。本文在模型数据输入之前,将数据集中各类数据的最大值和最小值作为参考值对数据集进行归一化处理:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (25)$$

式中: x_{norm} 为归一化数据, x 为原始数据, x_{\min} 和 x_{\max} 为各数据维度所对应的最大值和最小值。通过归一化处理,每个数据都被映射到一个 $[0, 1]$ 的值。

本文在 64 位 Windows 10 操作系统上使用 Matlab R2021b 进行所有实验,主要硬件包括 3.6 GHz 的 CPU 和 32 GB 的内存。根据前期研究及实验修正,制定仿真参数如表 1 所示。

表 1 仿真参数设置

Tab. 1 Simulation parameter setting

仿真参数	参数值	ISSA 仿真参数	参数值
学习率 lr	0.001	麻雀种群数 N	50
最大迭代次数 $Iter$	500	最大迭代次数 $\max Iter$	100
模态分量总数 m	5	侦察者的比例 $SD/\%$	20
隐藏神经元个数 N_{hy}	80	安全阈值 $ST/\%$	70

5.2 模型评价指标

采用 MAE、MAPE 和 RMSE 作为模型的评价标准,其表达式为:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i|, \quad (26)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right|, \quad (27)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2}, \quad (28)$$

式中: N 为测试数据, x_i 和 \hat{x}_i 为 t 时刻的真实值和预测值。以上 3 个评价准则取值越低,取得的预测效

果越优。

5.3 改进麻雀搜索算法对比分析

表 2 给出了应用于 SSA 和 ISSA 的 6 个基准测试函数以验证 ISSA 的优越性,实验函数包括 3 种单峰函数 $F_1 \sim F_3$ (用于局部优化能力测试) 和 3 种多峰函数 $F_4 \sim F_6$ (用于全局优化能力测试)。本节对每个测试函数进行 20 次实验,计算出 SSA 和 ISSA 测试结果的最优值、平均值、最差值和标准差。其中,平均值反映了算法的收敛精度,标准差反映了算法的稳定性。

表 2 测试函数
Tab. 2 Test function

函数类型	公式	维度	上界和下界	最优解
单峰函数	$F_1(x) = \sum_{i=1}^n x_i^2$	30	$[-100, 100]$	0
	$F_2(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$	30	$[-100, 100]$	0
	$F_3(x) = \max\{ x_i , 1 \leq i \leq n\}$	30	$[-100, 100]$	0
多峰函数	$F_4(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	$[-500, 500]$	-12 569
	$F_5(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	$[-600, 600]$	0
	$F_6(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	$[-5.12, 5.12]$	0

从表 3 可以看出,ISSA 算法在所有测试函数中均成功获得最优解。具体的,在最优值、平均值、最差值和标准差方面都有较好的表现,表现出 ISSA 优化的准确性和稳定性。由此得出 ISSA 通过加入动态自适应 t 分布变异,有效地增强了局部和全局寻优能力,同时提高了算法的收敛精度、稳定性和收敛速度。

表 3 SSA 和 ISSA 的测试结果
Tab. 3 SSA and ISSA test results

测试函数	算法	最优值	平均值	最值	标准偏差
F_1	SSA	4.8801×10^{-35}	8.2877×10^{-20}	8.6216×10^{-19}	2.1765×10^{-19}
	ISSA	0	2.2508×10^{-196}	6.2383×10^{-195}	0
F_2	SSA	1.3876×10^{-34}	3.5179×10^{-19}	3.2669×10^{-18}	7.9483×10^{-19}
	ISSA	0	2.7370×10^{-218}	8.2109×10^{-217}	0
F_3	SSA	4.6777×10^{-21}	2.1774×10^{-12}	4.1323×10^{-11}	7.9168×10^{-12}
	ISSA	0	2.9173×10^{-105}	4.2388×10^{-104}	9.8060×10^{-105}
F_4	SSA	-11 897.79	-7 980.75	-3 539.17	2 687.95
	ISSA	-12 653.87	-12 163.34	-10 673.76	787.21
F_5	SSA	0	3.7007×10^{-18}	1.1102×10^{-16}	1.9929×10^{-17}
	ISSA	0	0	0	0
F_6	SSA	0	1.8948×10^{-15}	5.6843×10^{-14}	1.0204×10^{-14}
	ISSA	0	0	0	0

5.4 交通流预测评估与分析

为了验证所提出的组合预测模型 ICEEMDAN-ISSA-BiGRU 性能的优越性,在同一数据集下构建 10 个基础预测模型,并与 ICEEMDAN-ISSA-BiGRU 进行比较分析。表 4 给出了用于交通流预测对比的典型模型清单,其中模型 1~5 与本文模型进行消融对比实验,模型 6~10 与本文模型进行不同模型性能对比实验。

表 4 典型交通流预测模型
Tab. 4 Typical traffic flow prediction model

模型	模型介绍	缩写
模型 1	循环神经网络	GRU
模型 2	双向门控循环单元	BiGRU
模型 3	基于 SSA 优化 BiGRU 模型进行预测	SSA-BiGRU
模型 4	基于本文改进 SSA 算法优化 BiGRU 模型进行预测	ISSA-BiGRU
模型 5	CEEMDAN 和本文改进的 SSA 算法优化 BiGRU 模型进行预测	CEEMDAN-ISSA-BiGRU

续表 4

模型	模型介绍	缩写
模型 6	自回归移动平均模型	ARIMA
模型 7	反向传播网络	BPNN
模型 8	长期记忆递归神经网络	LSTM
模型 9	CEEMDAN 对交通流量数据分解,GWO 算法优化 LSTM 模型进行预测	CEEMDAN-GWO-LSTM
模型 10	CEEMDAN 对交通流量数据分解,K-means 优化 BiLSTM 模型进行预测	CEEMDAN-K-means-BiLSTM
本文方法	本文改进的 CEEMDAN 对交通流量数据分解,本文改进的 SSA 算法优化 BiGRU 模型进行预测	ICEEMDAN-SSA-BiGRU

表 5 为不同模型的交通流预测性能对比,为了更直观地展示不同模型之间的误差情况,将各典型预测模型性能指标以曲线的方式进行对比展示,如图 5 所示。可以看出,本文模型的 MAE、RMSE、RMSE 的计算效果最好,预测精度最高。

① 由图 5(a) 及表 5 可以看出,BiGRU 通过增加双向性能和门控机制可以更好地捕捉序列中的上下文信息及长期依赖关系,获得更充足的时序特征,其预测曲线较 GRU 更加贴合原始交通流量曲线,预测精度显著提高,表明本文选择双向门控循环单元的优势。

② 由图 5(b) 及表 5 可以看出,通过改进的自适应麻雀搜索优化算法代替梯度下降法训练所得的最佳网络参数 BiGRU 网络相较于 SSA-BiGRU 预测模型更为准确,其 MAE 降低了 3.89,MAPE 降低了 4.38%,RMSE 降低了 14.41,验证了改进的自适应麻雀搜索优化算法对提高预测模型精度的有效性。

③ 由图 5(c) 及表 5 可以看出,与 CEEMDAN-SSA-BiGRU 相比,本文模型的 MAE 降低了 3.31,MAPE 降低了 5.33%,RMSE 降低了 3.55。验证了 CEEMDAN 改进后能够有效地减少信号中的噪声,提高信号的可靠性,同时减少了分解模态分量之间的干扰和重叠,提高了分解效果和准确度。

④ 由图 5(d) 及表 5 可以看出,在单个模型中,不同模型的预测精度值相似。同时,考虑到交通流时空特征的 BPNN 模型与只考虑交通流时空特征的 ARIMA、LSTM 和 GRU 模型相比,没有明显优势。与使用原始交通流数据的单一模型相比,基于分解原理的组合模型在 MAE、MAPE 和 RMSE 上均表现出较好的改进,SSA-BiGRU 模型的 MAPE 和 RMSE 分别比 BiGRU 提高了 4.47% 和 12.21。

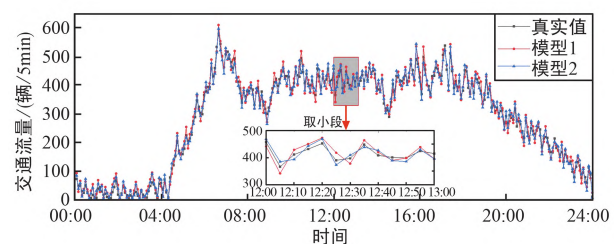
⑤ 由图 5(e) 及表 5 可以看出,本文组合模型弥补了 CEEMDAN-GWO-LSTM 和 CEEMDAN-K-means-BiLSTM 对非平稳性考虑的缺失以及对 CEEMDAN 中残留噪声和伪模态的忽视,因此预测

性能更加突出,其 MAE、MAPE 和 RMSE 指标性能相对较优的 CEEMDAN-K-means-BiLSTM 分别下降了 4.36、5.77% 及 4.13。结果表明,本文提出的组合预测模型在交通流预测方面具备较好的性能。

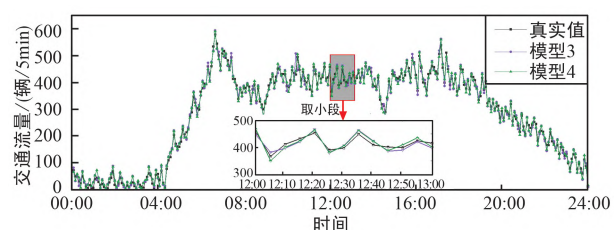
表 5 不同模型的交通流预测性能对比

Tab. 5 Comparison of traffic flow prediction performance of different models

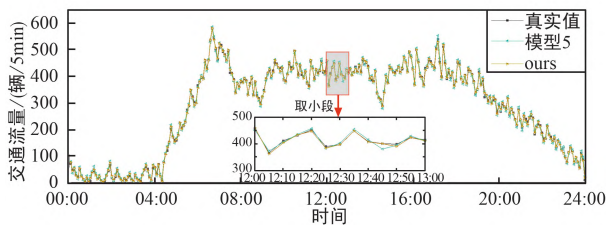
模型	MAE	MAPE/%	RMSE
模型 1	34.02	37.31	40.32
模型 2	29.04	28.22	36.76
模型 3	22.56	23.75	24.55
模型 4	18.67	19.37	20.14
模型 5	14.11	15.45	15.97
模型 6	33.74	36.25	50.66
模型 7	34.12	37.19	49.86
模型 8	34.96	38.32	50.98
模型 9	16.23	17.37	19.34
模型 10	15.34	15.89	16.56
本文方法	10.98	10.12	12.42



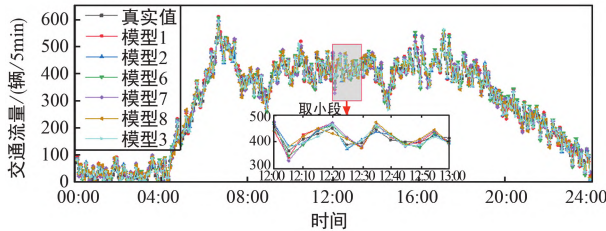
(a) GRU 与 BiGRU 预测曲线对比



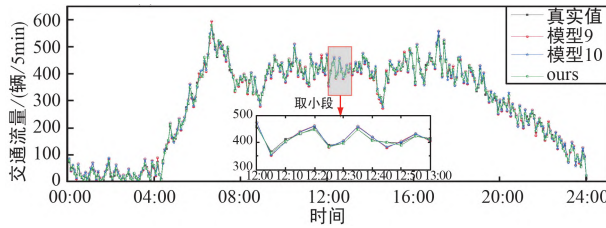
(b) SSA 改进前后组合模型性能曲线对比



(c) CEEMDAN 改进前后组合模型性能曲线对比



(d) 组合模型与单一模型预测性能曲线对比



(e) 不同组合模型预测性能对比

图 5 不同模型的短时交通流预测性能对比

Fig. 5 Comparison of short-time traffic flow prediction performance of different models

5.5 泛化性能实验对比与分析

为验证本文组合模型稳定的泛化能力,选取美国洛杉矶高速公路公开的 METR-LA 数据集^[31],再次与 CEEMDAN-GWO-LSTM 和 CEEMDAN-K-means-BiLSTM 预测模型进行对比验证,其预测性能指标如表 6 所示。可以看出,本文模型在该数据集下展示出较优的预测性能,与针对优化线性与非线性分量却忽略交通流量非平稳性的 CEEMDAN-GWO-LSTM 相比,其 MAE、MAPE 和 RMSE 下降了 3.62、3.59% 以及 2.72。与考虑了非线性及非平稳性但未优化 CEEMDAN 中的残留噪声和伪模态的 CEEMDAN-K-means-BiLSTM 相比,其 MAE、MAPE 和 RMSE 下降了 2.69、1.85% 以及 0.81。综合利用 PeMS 数据集和 METR-LA 数据集的交通流预测实验分析可知,本文组合模型在不同数据集下,预测精度均优于传统模型,具备良好的泛化能力。

表 6 METR-LA 数据集下预测模型的误差对比

Tab. 6 Error comparison of prediction models in METR-LA dataset

模型	MAE	MAPE/%	RMSE
CEEMDAN-GWO-LSTM	19.38	16.63	31.64
CEEMDAN-K-means-BiLSTM	18.45	14.89	29.73
ICEEMDAN-ISSA-BiGRU	15.76	13.04	28.92

6 结束语

本文提出了一种模块化的短时交通流预测组合模型,该模型利用 ICEEMDAN 方法将原始交通流时间序列数据分解为一定数量的 IMF 分量,克服了传统 CEEMDAN 方法存在的残留噪声和伪模态问题;基于动态自适应分布变异方法 ISSA 算法,避免传统 SSA 算法在短时预测时陷入局部最优;在 BiGRU 网络中,通过 ISSA 算法对 BiGRU 的权重参数进行迭代寻优。基于 PeMS 数据集进行短时交通流预测实验,结果表明,所提组合模型中各功能模块均能有效工作,组合模型性能优于传统模型。同时结合 METR-LA 数据集进行泛化性能评估,实验结果得出所提模型在不同的数据集均表现出更优的泛化性能。

参考文献

[1] QURESHI K N, ABDULLAH A H. A Survey on Intelligent Transportation Systems[J]. Middle-East Journal of Scientific Research, 2013, 15(5): 629-642.

[2] KAFFASH S, NGUYEN A T, ZHU J. Big Data Algorithms and Applications in Intelligent Transportation System: A Review and Bibliometric Analysis[J]. International Journal of Production Economics, 2021, 231: 1-15.

[3] 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 Part C: Emerging Technologies, 2019, 108: 12-28.

[4] LU S, ZHANG Q, CHEN G, et al. A Combined Method for Short-term Traffic Flow Prediction Based on Recurrent Neural Network [J]. Alexandria Engineering Journal, 2021, 60(1): 87-94.

[5] LIU B, TANG X, CHENG J, et al. Traffic Flow Combination Forecasting Method Based on Improved LSTM and ARIMA[J]. International Journal of Embedded Systems, 2020, 12(1): 22-30.

- [6] KUMAR P B, HARIHARAN K. Time Series Traffic Flow Prediction with Hyper-parameter Optimized ARIMA Models for Intelligent Transportation System [J]. Journal of Scientific & Industrial Research, 2022, 81(4): 408–415.
- [7] CAI P, WANG Y, LU G, et al. A Spatiotemporal Correlative K-nearest Neighbor Model for Short-term Traffic Multistep Forecasting [J]. Transportation Research Part C: Emerging Technologies, 2016, 62: 21–34.
- [8] ZHOU T, JIANG D, LIN Z, et al. Hybrid Dual Kalman Filtering Model for Short-term Traffic Flow Forecasting [J]. IET Intelligent Transport Systems, 2019, 13(6): 1023–1032.
- [9] TANG J, CHEN X, HU Z, et al. Traffic Flow Prediction Based on Combination of Support Vector Machine and Data Denoising Schemes [J]. Physica A: Statistical Mechanics and Its Applications, 2019, 534: 1–19.
- [10] MORETTI F, PIZZUTI S, PANZIERI S, et al. Urban Traffic Flow Forecasting Through Statistical and Neural Network Bagging Ensemble Hybrid Modeling [J]. Neurocomputing, 2015, 167: 3–7.
- [11] DING F, ZHU Y, YIN Q, et al. MS-ResCnet: A Combined Spatiotemporal Modeling and Multi-scale Fusion Network for Taxi Demand Prediction [J]. Computers & Electrical Engineering, 2023, 105: 1–17.
- [12] ZHENG H, LIN F, FENG X, et al. A Hybrid Deep Learning Model with Attention-based Conv-LSTM Networks for Short-term Traffic Flow Prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(11): 6910–6920.
- [13] 李静宜, 丁飞, 张楠, 等. 基于深度 LSTM 与遗传算法融合的短期交通流预测模型 [J]. 无线电通信技术, 2022, 48(5): 836–843.
- [14] RILLING G, FLANDRIN P, GONCALVES P. On Empirical Mode Decomposition and Its Algorithms [C] // IEEE-EURAS-IPWorkshop on Nonlinear Signal and Image Processing. Grado: IEEE, 2003: 8–11.
- [15] WEI Y, CHEN M C. Forecasting the Short-term Metro Passenger Flow with Empirical Mode Decomposition and Neural Networks [J]. Transportation Research Part C: Emerging Technologies, 2012, 21(1): 148–162.
- [16] CHEN X, CHEN H, YANG Y, et al. Traffic Flow Prediction by an Ensemble Framework with Data Denoising and Deep Learning Model [J]. Physica A: Statistical Mechanics and Its Applications, 2021, 565: 1–11.
- [17] LU W, RUI Y, YI Z, et al. A Hybrid Model for Lane-level Traffic flow Forecasting Based on Complete Ensemble Empirical Mode Decomposition and Extreme Gradient Boosting [J]. IEEE Access, 2020, 8: 42042–42054.
- [18] WANG Z, CHU R, ZHANG M, et al. An Improved Hybrid Highway Traffic Flow Prediction Model Based on Machine Learning [J]. Sustainability, 2020, 12(20): 1–22.
- [19] HUANG H, CHEN J, HUO X, et al. Effect of Multi-scale Decomposition on Performance of Neural Networks in Short-term Traffic Flow Prediction [J]. IEEE Access, 2021, 9: 50994–51004.
- [20] HAN T, LIU Q, ZHANG L, et al. Fault Feature Extraction of Low Speed Roller Bearing Based on Teager Energy Operator and CEEMD [J]. Measurement, 2019, 138: 400–408.
- [21] LI X, MA X, XIAO F, et al. Time-series Production Forecasting Method Based on the Integration of Bidirectional Gated Recurrent Unit (BiGRU) Network and Sparrow Search Algorithm (SSA) [J]. Journal of Petroleum Science and Engineering, 2022, 208: 2–13.
- [22] BAO K, BI J, GAO M, et al. An Improved Ship Trajectory Prediction Based on AIS Data Using MHA-BiGRU [J]. Journal of Marine Science and Engineering, 2022, 10(6): 804.
- [23] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise [C] // 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Prague: IEEE, 2011: 4144–4147.
- [24] COLOMINAS M A, SCHLOTTHAUER G, TORRES M E. Improved Complete Ensemble EMD: A Suitable Tool for Biomedical Signal Processing [J]. Biomedical Signal Processing and Control, 2014, 14: 19–29.
- [25] KARPATY A, JOHNSON J, LI F F. Visualizing and Understanding Recurrent Networks [J/OL]. (2015–06–05) [2023–03–12]. <https://arxiv.org/abs/1506.02078v2>.
- [26] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation [J/OL]. (2014–06–03) [2023–03–12]. <https://arxiv.org/abs/1406.1078>.

- [27] LIU D, BAO Y, HE Y, et al. A Data Loss Recovery Technique Using EMD-BiGRU Algorithm for Structural Health Monitoring[J]. Applied Sciences, 2021, 11(21):10072.
- [28] NIU D, YU M, SUN L, et al. Short-term Multi-energy Load Forecasting for Integrated Energy Systems Based on CNN-BiGRU Optimized by Attention Mechanism[J]. Applied Energy, 2022, 313(1):1-17.
- [29] ZHOU F, WANG X, ZHANG M. Evolutionary Programming Using Mutations Based on the t Probability Distribution[J]. Acta Electronica Sinica, 2008, 36(4):667-671.
- [30] Caltrans PeMS. Caltrans Performance Measurement System [EB/OL]. [2023-03-12]. <https://pems.dot.ca.gov>.
- [31] LI Y, YU R, SHAHABI C, et al. Diffusion Convolutional Recurrent Neural Network: Data-driven Traffic Forecasting [J/OL]. (2017-07-06) [2023-03-12]. <https://arxiv.org/abs/1707.01926>.

作者简介:



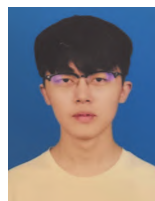
顾 潮 南京邮电大学硕士研究生。
主要研究方向:物联网与信息系统、车辆目标检测与边缘智能技术。



肖婷婷 南京信息工程大学本科在读。
主要研究方向:通信与信息系统、信号与信息处理。



丁 飞 南京邮电大学副教授。主要研究方向:智能网联车辆通信与网络技术、边缘智能与协同计算、智能系统与大数据。



周启航 南京邮电大学硕士研究生。
主要研究方向:车联网、表情识别、人脸检测。



赵芝因 南京邮电大学硕士研究生。
主要研究方向:物联网与信息系统、车辆目标检测与边缘智能技术。