

DOI: 10.13382/j.jemi.B2103878

基于 FEEMD-SAPSO-BiLSTM 组合模型的 短时交通流预测*

殷礼胜 魏帅康 孙双晨 何怡刚

(合肥工业大学 电气与自动化工程学院 合肥 230009)

摘要:为了提高短时交通流的预测精度和预测速度,基于交通流量序列的不平稳性和随机性,提出了快速集合经验模态分解(fast ensemble empirical mode decomposition, FEEMD)和自然选择自适应变异粒子群算法(selection adaptive particle swarm optimization, SAPSO)优化双向长短时记忆网络(bidirection long short-term memory, BiLSTM)相结合的预测模型。首先,利用FEEMD将原始不平稳的交通流量序列分解成多个较平稳的固有模态分量(intrinsic mode function, IMF)和残差分量(residue, Res),并滤除掉噪声部分,提高建模精度;其次,引入复合多尺度排列熵(composite multiscale permutation entropy, CMPE)检测交通流量子序列的随机性并根据随机性的相近程度对其进行聚类重组,简化模型的构建,提高预测精度;然后,对重组后的子序列使用BiLSTM进行预测,并利用SAPSO优化BiLSTM的权值和阈值,进一步提高组合模型的预测精度和预测速度;最后,将各子序列预测值叠加得到最终的预测值。实验结果表明,FEEMD-SAPSO-BiLSTM组合模型的均方根误差比FEEMD-PSO-BiLSTM和SAPSO-BiLSTM组合模型分别降低了22.9%和54.3%,收敛速度方面,FEEMD-SAPSO-BiLSTM明显快于FEEMD-PSO-BiLSTM模型。因此在预测短时交通流上,提出的组合模型提高了预测精度和预测速度,达到了期望的预测效果。

关键词: 短时交通流;快速集合经验模态分解;自然选择自适应变异粒子群;双向长短时记忆网络

中图分类号: U491.14;TN98 **文献标识码:** A **国家标准学科分类代码:** 580.2099

Short-term traffic flow forecast based on FEEMD-SAPSO-BiLSTM combined model

Yin Lisheng Wei Shuaikang Sun Shuangchen He Yigang

(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

Abstract: In order to improve the prediction accuracy and speed of short-term traffic flow, based on the instability and randomness of the traffic flow sequence, fast ensemble empirical mode decomposition (FEEMD) and natural selection adaptive mutation particle swarm optimization algorithm (SAPSO) are proposed to optimize the two-way Predictive model combined with long and short-term memory network (BiLSTM). Firstly, using FEEMD to decompose the original unsteady traffic flow sequence into multiple stable intrinsic modal components (IMF) and residual components (Res), and filter out the noise part to improve modeling accuracy; secondly, introducing composite Multi-scale permutation entropy (CMPE) to detect the randomness of traffic flow sub-sequences and regroups them to simplify model construction and improve prediction accuracy; then, using BiLSTM to predict the reorganized subsequences, and use SAPSO to optimize the weights and thresholds of BiLSTM to further improve the prediction accuracy and prediction speed of the combined model; finally, the prediction values of each sub-sequence are superimposed to obtain the final prediction value. The experimental results show that the root mean square error of the FEEMD-SAPSO-BiLSTM combined model is 22.9% and 54.3% lower than the FEEMD-PSO-BiLSTM combined model and the SAPSO-BiLSTM combined model, respectively. In terms of convergence speed, the FEEMD-SAPSO-BiLSTM model is obviously faster than FEEMD-PSO-BiLSTM model. Therefore, in predicting short-term traffic flow, the proposed combined model improves the prediction accuracy and speed and achieves the desired prediction effect.

Keywords: short-term traffic; fast ensemble empirical model decomposition; selection adaptive particle swarm optimization; bidirection long short-term memory

收稿日期:2021-01-10 Received Date: 2021-01-10

* 基金项目:国家自然科学基金资助项目(62073114,51577046,61673153)、省基金项目(JZ2021AKZR0344)资助

0 引言

随着我国汽车制造技术的成熟和交通路网的逐步完善,以及汽车保有量的极速上升,进而对交通规划的需求增大,而短时交通流预测作为其路径规划的基本参考之一,近年来成为学者研究的热点^[1]。短时交通流具有较强的不平稳性和随机性,各学者对此特点提出很多预测模型和算法^[2-3]。赵亚斌等^[4]采用非线性动态惯性权重粒子群算法优化支持向量机(support vector machine, SVM)的组合模型预测短时交通流,该方法通过改进型粒子群改善了模型的局部寻优能力,提高了预测精度,但忽略了交通流量序列的不平稳性和随机性的特点,缺少了对交通流量序列的分解处理。张阳等^[5]采用改进小波包分解与长短时记忆网络相结合的预测模型,该模型通过小波包分解极大的降低了原始交通流序列的不平稳性,但缺乏对预测模型的优化算法,并且分解的子序列较多,没有对其进行聚类,导致误差较高。陈如清等^[6]采用奇异谱分析方法先对原始交通流数据进行处理,去除噪声成分,降低交通流量序列的随机性,然后用烟花差分进化算法优化后的极限学习机进行预测。该方法虽然采用奇异谱分析法降低了交通流量序列的随机性,并采用优化算法提高模型的预测精度和泛化能力,但忽略了交通流量序列的不平稳性。

从短时交通流序列的不平稳性、随机性的特点来看,以上研究,有的忽略了交通流量序列的不平稳性,或者缺乏对分解后子序列的聚类处理,有的则没有针对预测模型的优化算法。本文针对这些研究的不足,为进一步提高交通流量的预测精度和预测速度,从交通流量序列的分解、子序列的聚类、预测模型的优化这 3 个方面综合考虑,拟提出一种基于快速集合经验模态分解(fast ensemble empirical mode decomposition, FEEMD)和自然选择自适应变异粒子群算法(selection adaptive particle swarm optimization, SAPSO)优化双向长短时记忆网络(bidirection long short-term memory, BiLSTM)的组合预测模型(FEEMD-SAPSO-BiLSTM)。首先,提出的 FEEMD 通过将原始不平稳的交通流量序列分解成多个较平稳的子序列,同时既可以去噪又可以保留有用信号的完整性。其次,引入的 CMPE 复杂度检测方法通过对交通流量子序列进行聚类重组,可减少要预测的子序列数量,降低因子序列较多产生的误差累计。SAPSO 通过对 BiLSTM 的权值和阈值的更新优化,加速预测模型的训练,提高模型的预测精度和速度。最后将各子序列的预测值叠加得到最终的预测值,并用该模型对合肥市望江东路与马鞍山路交叉口由南向北方向交通流量进行预测和对比分析,通过与 FEEMD-PSO-BiLSTM、SAPSO-BiLSTM 组合模型

以及 BiLSTM、SVM 模型的预测值的比较,以验证该文提出的组合模型在预测精度和预测速度上的优越性。

1 基于交通流量序列的快速集合经验模态分解

1.1 传统集合经验模态分解(EEMD)的缺陷与不足

在处理交通流量时间序列问题上,由于数据不平稳性的特点,导致预测模型的预测精度不高,性能不好,为了挖掘交通流量数据间隐藏的关系,使数据变得平稳些,需要对原始交通流量序列进行分解。经验模态分解(EMD)适用于处理这类信号,但是 EMD 在分解过程中会因为极值点分布不均匀而导致严重的模态混叠现象。针对这类现象,EEMD 被提出来,它通过对原始交通流量序列多次添加强度相同但序列不同的白噪声,以补充有缺失的信号,并对新信号进行分解^[7]。因进行多次迭代,其筛选固有模态分量(Intrinsic Mode Function, IMF)过程非常耗时,且每次迭代得到的各 IMF 分量也都混入了特定尺度下的大量噪声成分,若直接把某个特定尺度下的 IMF 分量完全去除,很有可能在去噪的同时也滤掉了一些对预测交通流量有用的成分,影响后续交通流序列分析的准确性。

1.2 改进的交通流量序列 EEMD 分解算法

基于以上分析,为了解决在去除交通流量某个尺度上的 IMF 分量时丢掉对预测有用的关键信息问题,本文采用了一种用于分解交通流量序列的改进方法:快速集合经验模态分解。在 EEMD 的基础上,通过加设阈值的方法剔除交通流量某些尺度上无用的 IMF 分量,在去除信号噪声的同时也保证了有用信号的完整性^[8]。

算法 1 FEEMD 对交通流量序列的分解算法

1) 在原始交通流量序列 $x(t)$ 上加入白噪声序列 $n(t)$,得到的 $X(t)$ 为加入白噪声序列后的交通流量序列^[9]。即:

$$X(t) = x(t) + n(t) \quad (1)$$

2) 找出式(1) $X(t)$ 的极大值和极小值,并采用三次样条差值法获得上下包络线 $U(t)$ 和 $L(t)$,再取均值得到交通流均值序列 $m(t)$:

$$m(t) = \frac{U(t) + L(t)}{2} \quad (2)$$

3) 提取信号的局部信号特征,即求出 $X(t)$ 与 $m(t)$ 的差值:

$$h(t) = X(t) - m(t) \quad (3)$$

对新的时间序列 $h(t)$ 进行判断,看是否满足如下条件:

(1) 极值点的个数与序列过零点的个数是否相等或

相差 1;

(2) 在任意的时刻点, 上下包络线的平均值是否为 0。

条件(1)表示两个极点之间不包含其他极点, 这样保证了两个极点之间序列的平稳性。条件(2)表示交通流子序列在 x 轴上下两个方向的幅值相等, 保证了分解的准确性。若满足条件, 继续进行下一步操作, 否则重复步骤 2)、3)。

4) 用上述方法得到第一个交通流 IMF 分量, 记为 $c_1(t)$, 再用交通流量序列 $X(t)$ 减去 $c_1(t)$ 作为新的交通流序列, 通过上述的步骤 1)~3), 可以得到 IMF2, 以此类推, 得到第一次加入白噪声后的交通流量时间序列 $X(t)$ 的分解子序列:

$$X(t) = \sum_{j=1}^m c_j(t) + r_m(t) \quad (4)$$

式中: $c_j(t)$ ($j=1, 2, \dots, m$) 为 $X(t)$ 分解后的第 j 个交通流 IMF 分量; $r_m(t)$ 为余项, m 为分解层数。

5) 对各个交通流 IMF 分量进行门限阈值处理, 软限幅函数为:

$$\eta_j(t) = \begin{cases} 0, & |c_j(t)| \leq l \\ c_j(t) \operatorname{sgn}(|c_j(t)| - l), & |c_j(t)| > l \end{cases} \quad (5)$$

根据文献[9]对 IMF 分量加设阈值筛除噪音的方法可知, l 是交通流序列软限幅函数的阈值, $l = \sigma \sqrt{2 \lg M}$, σ 是噪声的方差, M 是含噪信号的长度。当交通流量序列分解后的第 j 个 IMF 分量的值 $\leq l$ 时, 舍掉该分量, 反之则保留该分量。 $\eta_j(t)$ 为经过阈值处理后的第 j 个短时交通流 IMF 分量, 此时的 IMF 分量筛除掉了某些时间尺度上存在的噪声信号, 使对交通流量预测有用的 IMF 分量比例更大, 预测交通流量的性能更好。

6) 每次在原始交通流量序列 $x(t)$ 上加入不同的白噪声序列 $n_i(t)$ ($i=1, 2, \dots, n$), 重复步骤 1)~5), 得到:

$$X_i(t) = \sum_{j=1}^n \eta_{ij}(t) + r_{in}(t) \quad (6)$$

式中: $\eta_{ij}(t)$ 为第 i 次加入白噪声的原始交通流时间序列的第 j 个 IMF 分量; $r_{in}(t)$ 为第 i 个剩余交通分量; n 为叠加白噪声的次数。

7) 由式(6)分解得到的各 IMF 的均值作为最终交通流量分解结果:

$$h_j(t) = \frac{1}{n} \sum_{i=1}^n \eta_{ij}(t) \quad (7)$$

最终原始交通流时间序列被分解成:

$$x(t) = \sum_{j=1}^m h_j(t) + r_m(t) \quad (8)$$

式中: $h_j(t)$ ($j=1, 2, \dots, m$) 为分解得到的第 j 个 IMF 分量; $r_m(t)$ 为残差分量。

2 交通流量子序列的复合多尺度排列熵 (composite multiscale permutation entropy, CMPE) 计算

原始交通流量序列经过 FEEMD 分解后, 分量较多, 建模时间变长。为了减少要预测的子序列数量, 降低因累计误差带来的影响, 减少整体预测时间, 本文引入复合多尺度排列熵这一概念。排列熵是一种衡量时间序列随机性的方法, 可以量化 IMF 分量复杂程度, 然后合并复杂程度相似的 IMF 分量, 达到减少 IMF 分量的目的。CMPE 通过对各个尺度上的粗粒化交通流序列求排列熵值, 然后对这些排列熵值求平均的方式获得最终的 CMPE 值, 避免了在较大尺度因子下, 多尺度排列熵丢失交通流序列有用信息的问题, 从而得到较准确的交通流量序列复杂度^[10-11]。

2.1 排列熵方法

将交通流量子序列 $\{x_i\}_{i=1}^n$ 进行相空间重构, 得到重构矩阵^[12]:

$$\mathbf{x} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(m-1)\tau} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n-(m-1)\tau} & x_{n-(m-2)\tau} & \cdots & x_n \end{bmatrix} \quad (9)$$

式中: m 、 τ 分别表示嵌入维数、延迟时间, 根据文献[11]的 C-C 算法, 得出当 $\tau=2$, $m=4$ 时, 交通流量各子序列变化更加规律, 分类更加明晰^[11]。将重构矩阵中每一个重构分量的元素按照数值大小做升序排列: $x_{i+(k_1-1)\tau} \leq x_{i+(k_2-1)\tau} \leq \cdots \leq x_{i+(k_m-1)\tau}$, k_1, k_2, \dots, k_m 表示各元素所在列的索引。设每种符号序列出现的概率分别为 P_1, P_2, \dots, P_g ($\sum_{i=1}^g P_i = 1$), 其中 $g \leq n - (m-1)\tau$ 。按照香农熵的形式可得此交通流量子序列的排列熵值:

$$H(m, \tau) = - \sum_{i=1}^g P_i \ln P_i \quad (10)$$

将式(10)进行归一化, 即:

$$H_p = H(m, \tau) / \ln(m!) \quad (11)$$

2.2 复合多尺度排列熵

CMPE 算法通过复合粗粒化的方式优化多尺度排列熵单一粗粒化方式的不足, 可最大程度保留交通流序列所蕴含的变化信息, 能较为准确地评估交通流子序列的复杂程度。

算法 2 复合多尺度排列熵算法

1) 对交通流量子序列 $\{x_i\}_{i=1}^n$ 进行粗粒化序列构造:

$$y_{a,l}^{(\omega)} = \frac{1}{\omega} \sum_{i=(l-1)\omega+a}^{l\omega+a-1} x_i \quad 1 \leq a \leq \omega, 1 \leq l \leq n/\omega \quad (12)$$

式中： ω 为尺度因子。当 $\omega = 1$ 时，粗粒化序列即为原始交通流量子序列；当 $\omega > 1$ 时， $y_{a,l}^{(\omega)}$ 表示在尺度因子 ω 下的第 a 个粗粒化序列， j 代表第 a 个粗粒化序列的第 j 个点。

2) 利用 2.1 节的排列熵方法计算在尺度因子 ω 下的每个粗粒化序列的排列熵值 $H_{pei}, 1 \leq i \leq \omega$ 。

3) 对步骤 1)、2) 得到的 ω 个排列熵值求平均，得到交通流量子序列在尺度因子 ω 下的复合多尺度排列熵值：

$$H_{CMPE} = \frac{1}{\omega} \sum_{i=1}^{\omega} H_{PEi} \quad (13)$$

根据上述方法，计算出交通流量各个子序列的复合多尺度排列熵值，计算结果如表 1 所示。以 CMPE 的值来表示子序列的随机性大小，其值越接近 1，子序列的随机程度越大，复杂程度越高；越接近 0，随机性越小，序列越规则。

CMPE 值代表了子序列的复杂程度，复杂度越接近的子序列，相互之间的相关性更强，对最终交通流量预测结果的影响基本相同，可将其叠加预测。当子序列的 CMPE 值相差小于 0.1 时，可认为其复杂度基本相似。其次，即使子序列的 CMPE 值相差较大，但变化较为规律，则其对预测结果的影响也较为规律，也可归为一类。所以据此，重组可按照以下两个条件进行：

(1) $|H_{CMPEi} - H_{CMPEj}| < 0.1$ ，其中 H_{CMPEi} 和 H_{CMPEj} 是任意两个交通流量子序列的 CMPE 值；

(2) 子序列的 CMPE 值近似成线性排列。

根据以上合并规则，将 9 个交通流量子序列进行聚类重组，重组结果如图 1 所示。

表 1 交通流量各子序列 CMPE 值

Table 1 CMPE value of each sub-sequence of traffic flow

| 交通流量子序列 | CMPE 值 |
|---------|--------|
| IMF1 | 0.98 |
| IMF2 | 0.95 |
| IMF3 | 0.90 |
| IMF4 | 0.70 |
| IMF5 | 0.55 |
| IMF6 | 0.41 |
| IMF7 | 0.21 |
| IMF8 | 0.19 |
| Res | 0 |

3 基于交通流量序列的 SAPSO-BiLSTM 理论

3.1 双向长短时记忆神经网络

考虑到交通流量序列在时间上前后的相关性，随机

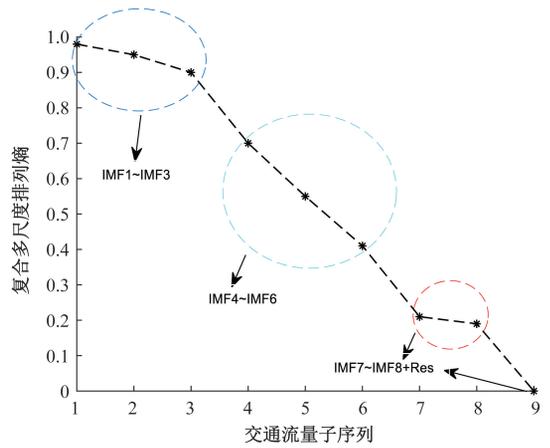


图 1 交通流量各子序列的聚类结果

Fig. 1 Clustering results of each sub-sequence of traffic flow

性强的交通流量因数据剧烈变化对预测精度造成的破坏，本文的预测模型采用双向长短时记忆网络。长短时记忆网络 (LSTM) 是众多循环神经网络中的一个改进型网络，传统循环神经网络因长期记忆能力不足的问题^[13]，当近期交通流量剧烈变化时，预测结果会偏向于近期的数据集，而忽略掉长期的数据。但 LSTM 具有长短期的记忆能力^[14]，在预测交通流量序列过程中能够真正有效地利用交通流量长短时的时序信息。双向 LSTM 由前向和后向两个单向的 LSTM 网络构成，结合了输入的交通流量前向和后向两个方向上的时序信息，对时序信息的前后关联性具有明显的优势^[15]。因为 BiLSTM 综合考虑了两个方向的交通流量序列，可使模型的预测精度得到进一步提升。BiLSTM 预测采用单步预测，BiLSTM 的输出为：

$$h'_t = w_d \cdot [h_t^{(f)}, h_t^{(b)}] + b_d \quad (14)$$

式中： w_d 和 b_d 分别为权值和阈值； $h_t^{(f)}$ 和 $h_t^{(b)}$ 是 BiLSTM 网络的输入。前向和后向 LSTM 的输入是经过分解处理后的交通流量数据集，输出分别为 $h_t^{(f)}$ 和 $h_t^{(b)}$ 。

3.2 SAPSO-BiLSTM

BiLSTM 权值和阈值的更新对交通流量序列的预测性能至关重要。本文采用了一种改进型粒子群算法，即自然选择自适应变异粒子群算法来优化 BiLSTM 的参数。SAPSO 可以改进粒子群算法 (PSO) 后期搜索能力低、容易发生早熟和陷入局部最优的缺陷。SAPSO 改进措施：在 PSO 的基础上^[16-17]，首先引入基于欧氏距离的区域划分，通过粒子种群和全局最优粒子的距离划分为界内区域和界外区域，加强粒子的优化效率。其次为了加强粒子的动态寻优能力，避免算法在参数寻优过程中陷入局部最优，引入自适应变异操作^[18]。而且在自适应变异的基础上引入遗传算法中的自然选择思想，在保证粒子良好的搜索能力前提下，通过对优良粒子的自然选

择来提高粒子群的收敛速度^[19]。

算法 3 SAPSO 优化 BiLSTM 算法

1) 初始化粒子群参数, 将 BiLSTM 网络的权值和阈值编码到种群中粒子的位置。

2) 根据适应度函数计算粒子适应度, 适应度函数

$fit = \frac{1}{l} \sum_{i=1}^l (\hat{y}_i - y_i)^2$, 其中 l 为交通流训练样本集数, \hat{y}_i 和 y_i 分别为 BiLSTM 中实际交通流量值和交通流量预测值。

3) 将种群中的粒子按照适用度优劣进行排序并均分为两部分, 适应度差的粒子被适应度好的粒子所取代。

4) 根据公式 $d_{ij} = \sqrt{\sum_{m=1}^D (P_{gm} - P_{jm})^2}$ 计算粒子与全局最优粒子的欧氏距离, d_{ij} 表示第 i 次迭代的第 j 个粒子与全局最优粒子的欧式距离, P_{gm} 表示在 D 维空间中全局最优粒子在第 m 维的分量, P_{jm} 表示第 j 个粒子在第 m 维的分量。按照距离由远及近排序, 当 $d_{ij} < d_{lim}$ 的粒子达到 43% 的比例后, 将 d_{lim} 作为界内区域和界外区域的边界。

5) 提出一种根据欧氏距离来动态调整变异概率的方法, 变异概率 $p = a \left(\frac{d_{max} - d_{ij}}{d_{max}} \right)$, a 为 $[0, 1]$ 的常数, 取 0.9, d_{max} 为粒子与全局最优粒子最大欧氏距离。 d_{ij} 越小, 则 p 越大, 即粒子离全局最优粒子越近, 变异概率越大, 因而可以更好的开发潜在区域。若粒子 q_{ij} 产生的随机数 $r > p$, 则进行变异操作: $\hat{x}_{ij} = x_{ij} + r(x_{ij} - x_{ik})$, x_{ij} 和 x_{ik} 分别为

粒子 q_{ij} 和界外区域粒子 q_{ik} 在空间中的位置, \hat{x}_{ij} 是粒子 q_{ij} 变异后的位置^[20]。

6) 判断是否达到最大迭代次数 100, 达到则输出最优解, 否则回到步骤 2) 继续迭代。

4 基于 FEEMD-SAPSO-BiLSTM 的交通流量组合预测模型构建与算法

4.1 交通流量序列数据集

本文的实验数据提取于合肥市望江东路与马鞍山路交叉口由南向北方向 2018 年 8 月 20 日至 24 日 8:00 ~ 13:50 的 350 个交通流数据样本, 时间间隔 5 min 记录 1 次。本文选取前 280 个数据作为训练样本, 后 70 个数据用来作预测样本。

4.2 FEEMD-SAPSO-BiLSTM 的交通流量组合预测模型

短时交通流具有的非线性和不平稳性的特点, 使得传统的单一预测方法难以得到较好的预测结果。由于 FEEMD 分解方法在分解非线性、不平稳时间序列方面效果显著, 并考虑到双向长短时记忆网络在预测短时交通流量上具有的能充分利用长短时的时序信息优势, 因此本文采用 FEEMD 和 BiLSTM 组合预测的方法对短时交通流进行预测, 并采用 SAPSO 算法对 BiLSTM 的参数进行优化处理, 对子序列利用 CMPE 评估复杂度并对其聚类重组。结构如图 2 所示。

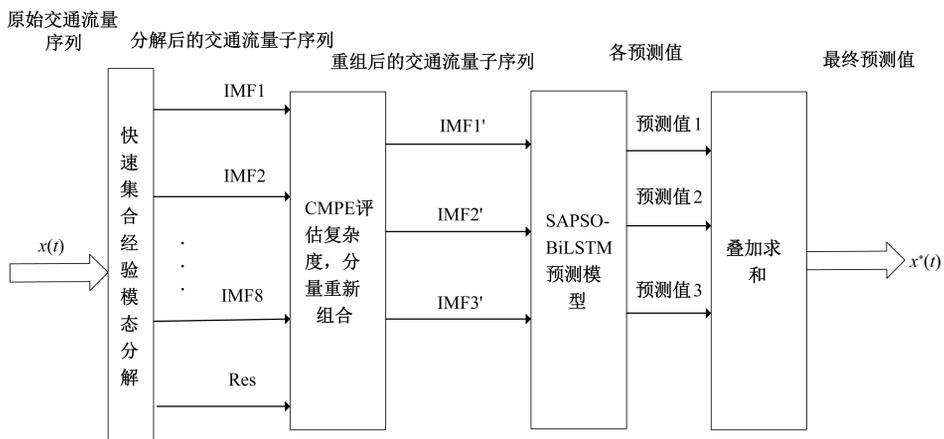


图 2 FEEMD-SAPSO-BiLSTM 组合预测模型

Fig. 2 FEEMD-SAPSO-BiLSTM combined forecasting model

4.3 基于 FEEMD-SAPSO-BiLSTM 组合预测算法

本文采用快速集合经验模态分解和自然选择自适应变异粒子群结合双向长短时记忆网络组合预测算法对短时交通流序列进行预测, 算法流程如图 3 所示, 步骤

如下。

1) 为了降低交通流量时间序列的非平稳特性, 首先调用算法 1, 用 FEEMD 对原始序列进行分解, 得到多组 IMF 分量和残差分量。

2) 调用算法 2 用复合多尺度排列熵对分解得到的各子序列进行聚类分组, 将排列熵值相近的子序列进行聚类重组:

$$x_{new1} = [x_{j1}, x_{j2}, \dots] + [x_{k1}, x_{k2}, \dots] \quad (15)$$

其中, $j \neq k$, 且 $j, k \leq 9$, x_{j1}, x_{j2}, \dots 为交通流量第 j 个子序列的数据, x_{k1}, x_{k2}, \dots 为交通流量第 k 个子序列的数据。

3) 将交通流量序列划分为训练样本和测试样本, 取前 280 个数据作为训练集, 后 70 个作为测试集。

4) 将交通流量的训练样本输入到 BiLSTM 训练模型, 对交通流量时间序列进行时间特征的挖掘。并调用算法 3, 采用 SAPSO 对 BiLSTM 的参数进行优化, 设置粒子种群规模数目为 60, 最大迭代数为 100。得出粒子每个位置对应 BiLSTM 的权值和阈值的最优解。

5) 将重构后的交通流量测试集序列输入到优化后的 BiLSTM 中, 并将各个子序列的预测结果进行叠加求和得出最终的交通流量预测结果。

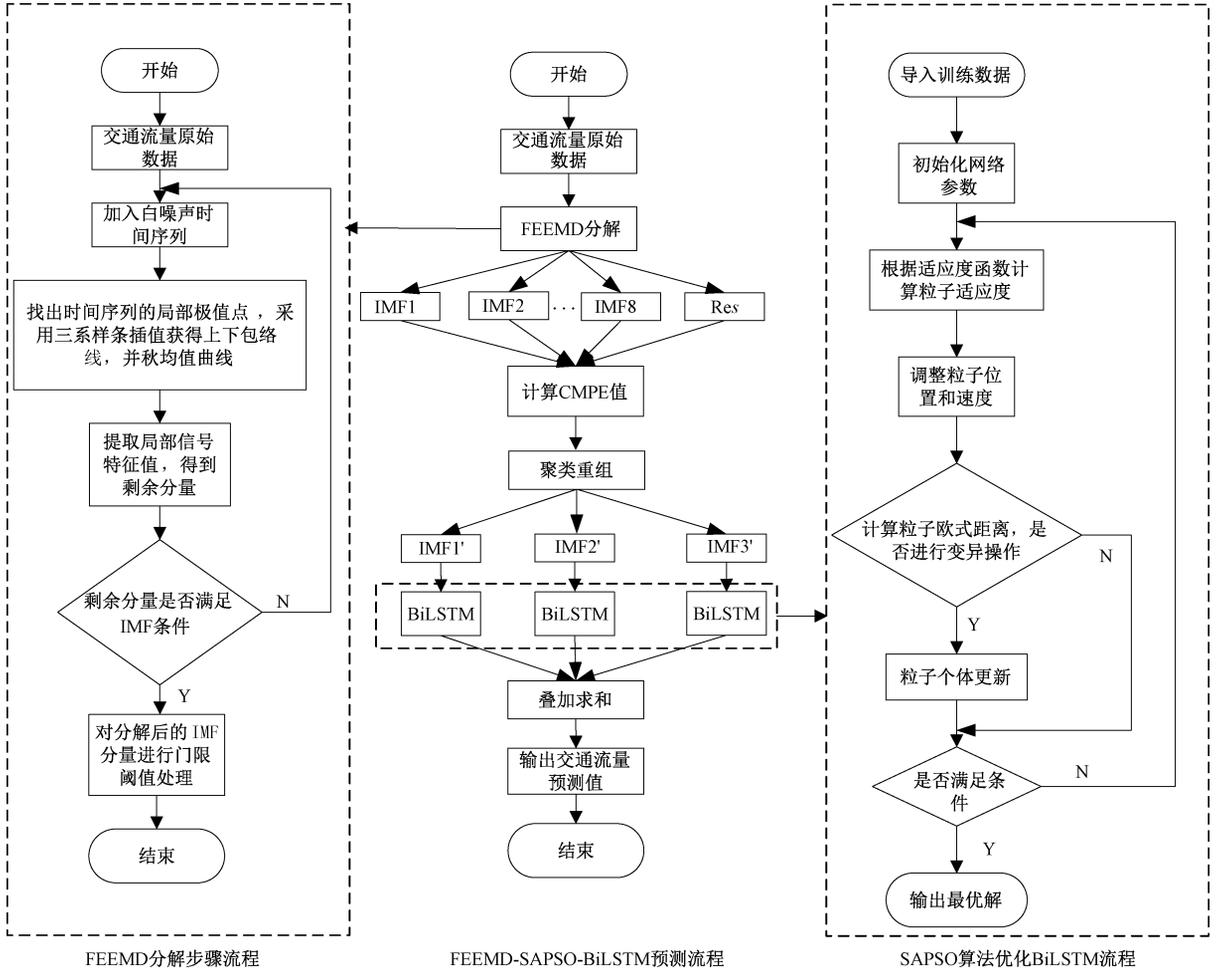


图 3 FEEMD-SAPSO-BiLSTM 组合预测模型算法流程

Fig. 3 FEEMD-SAPSO-BiLSTM combined forecasting model algorithm flow chart

5 实验结果与分析

5.1 交通流量时间序列数据集分解

经 FEEMD 分解产生的子序列如图 4 所示, 由图 4 可知, 分解出多个从高频到低频分布的模态分量, 其中 8 个 IMF 分量和一个残差分量 (Res)。图 5 所示是根据交通流量各子序列聚类重组后的示意图, 第 1 个频率较高的子序列是 IMF1 ~ IMF3 合并后的结果, 第 2 个子序列是

IMF4 ~ IMF6 合并的结果, 第 3 个则是 IMF7、IMF8 和残差分量的合并。

将重组后的交通流量子序列分别输入到 FEEMD-SAPSO-BiLSTM 预测模型进行预测, 将预测结果进行叠加, 即得到最终的交通流量预测结果, 如图 6 所示。

5.2 评判标准

为了分析各个模型的预测精度, 本文采用平均绝对误差 (mean absolute error, MAE) 和均方根误差 (root mean square error, RMSE) 两个评价指标来分析预测模型样本

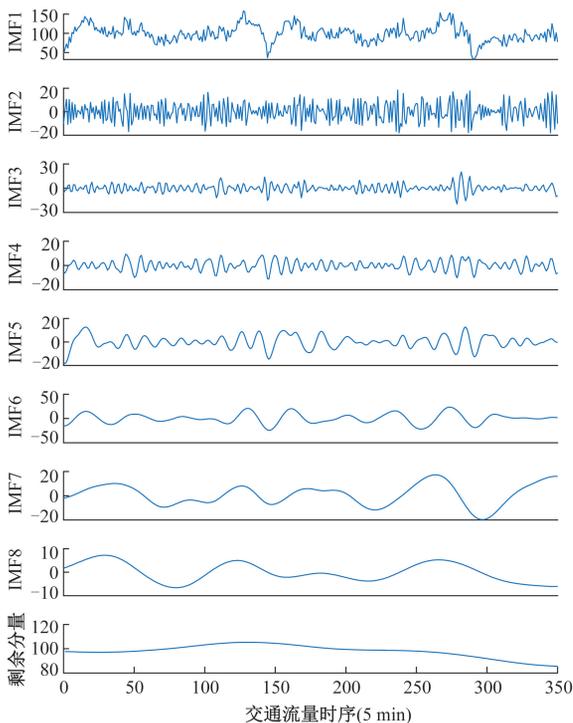


图 4 经 FEEMD 分解的子序列

Fig. 4 Subsequence decomposed by FEEMD

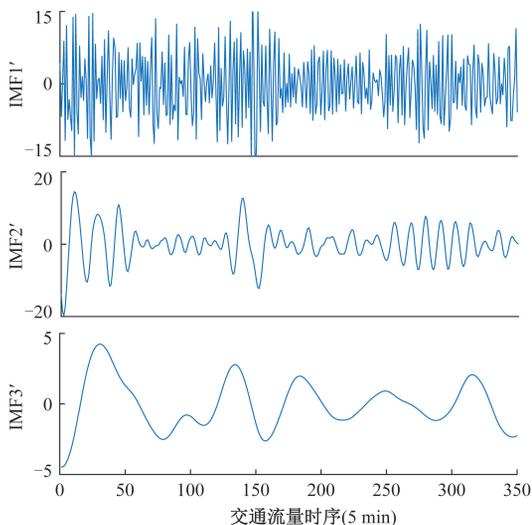


图 5 交通流子序列重组后的结果

Fig. 5 The result of traffic flow subsequence reorganization

的拟合程度。

平均绝对误差:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x(i) - x_r(i)| \quad (16)$$

均方根误差:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(i) - x_r(i))^2} \quad (17)$$

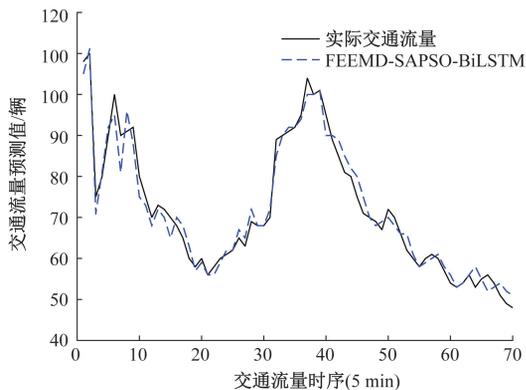


图 6 FEEMD-SAPSO-BiLSTM 组合模型预测结果

Fig. 6 FEEMD-SAPSO-BiLSTM combined model prediction results

式中: n 为交通流量数据的个数; $x(i)$ 为实验预测结果; $x_r(i)$ 为实际交通流量数据。

5.3 实验对比与分析

一个模型预测性能的优劣体现在预测精度和收敛速度上。为了验证本文的组合预测模型有较高的收敛速度,取一组交通流量数据样本分别输入到 FEEMD-SAPSO-BiLSTM 和 FEEMD-PSO-BiLSTM 模型中。图 7 所示为两个模型的迭代曲线示意图,由图 7 可知,SAPSO 迭代 14 次后就开始收敛,而 PSO 需要迭代 25 次后才能找到全局最优解,所以用改进后的粒子群算法优化双向长短时记忆网络的收敛速度更快,预测交通流量的性能更优。

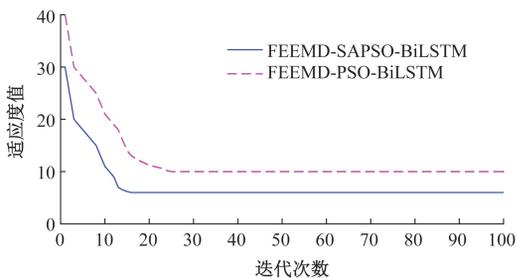


图 7 SAPSO 和 PSO 迭代曲线比较

Fig. 7 Comparison of SAPSO and PSO iteration curves

为验证本文提出的组合模型对提高交通流量预测精度的作用,分别用 FEEMD-SAPSO-BiLSTM 组合模型、FEEMD-PSO-BiLSTM、SAPSO-BiLSTM 组合模型以及 BiLSTM、SVM 模型对交通流量进行预测。各模型的预测结果如图 8 所示,相对误差如图 9 所示,表 2 为各模型平均绝对误差和均方根误差的对比。本文提出的改进型粒子群算法能够改善粒子群算法后期搜索能力低、陷入局部最优的缺点,由图 8、9 可知,本文提出的组合模型比 FEEMD-PSO-BiLSTM 组合模型拟合效果好,预测精度更高;另外 FEEMD-SAPSO-BiLSTM 组合模型相较于

SAPSO-BiLSTM 整体预测效果也更好,尤其当数据剧烈变化时,如不平稳段 20~25 段和 35~40 段。原因在于 FEEMD 分解算法结合 CMPE 量子子序列的复杂度进行聚类重组,解决了原始交通流量序列非平稳特点,并减少

子序列数量,简化模型的构建,降低了预测误差。并且还可知单个 BiLSTM 模型的预测精度也优于 SVM 模型。综上所述,本文提出的 FEEMD-SAPSO-BiLSTM 组合模型一定程度上提高了短时交通流的预测性能。

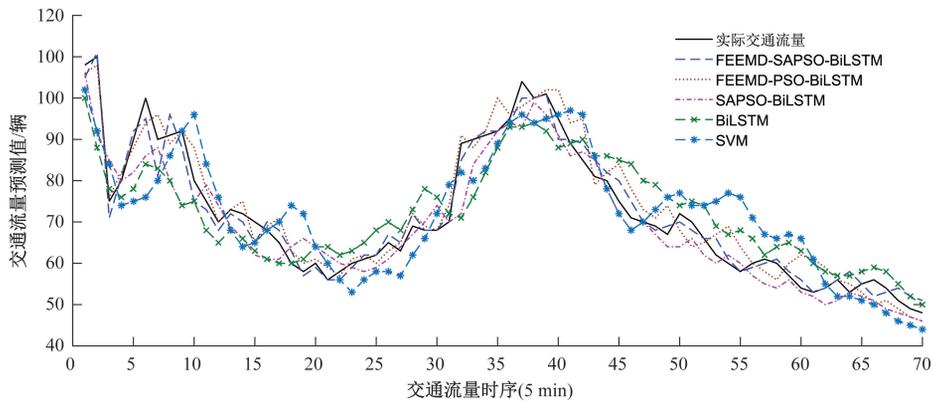


图 8 各模型预测结果对比

Fig. 8 Comparison of prediction results of various models

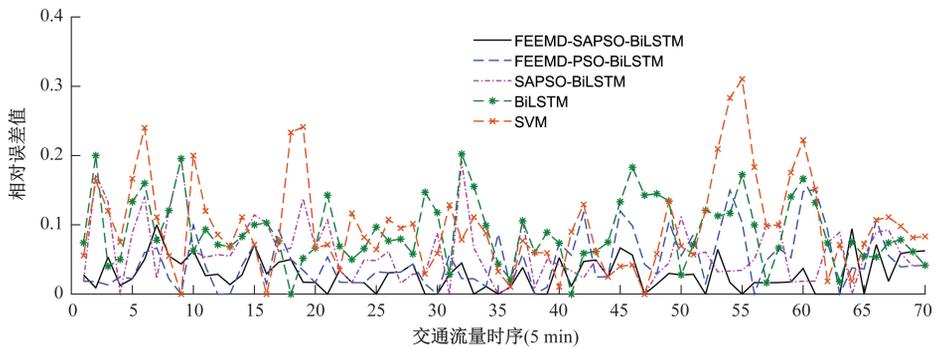


图 9 各模型预测结果的相对误差对比

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

表 2 各模型预测误差

Table 2 Forecast errors of each model

| 预测模型 | MAE | RMSE |
|--------------------|---------|---------|
| FEEMD-SAPSO-BiLSTM | 2.685 7 | 2.767 1 |
| FEEMD-PSO-BiLSTM | 3.300 0 | 4.196 9 |
| SAPSO-BiLSTM | 4.142 8 | 5.974 9 |
| BiLSTM | 6.628 5 | 7.863 1 |
| SVM | 7.336 1 | 8.373 4 |

6 结 论

为了降低短期交通流序列的非平稳性、随机性,提高交通流量的预测精度和预测速度,本文提出快速集合经验模态分解和自然选择自适应变异粒子群优化双向长短时记忆网络相结合的短期交通流预测模型。通过对实际的交通流量序列的预测,证实了本文所提出的 SAPSO 能更好的优化 BiLSTM 参数,一定程度上提高了预测精度

和预测速度,使用 FEEMD 对原始交通流量序列进行分解并对子序列通过 CMPE 评估随机性大小进行聚类重组的预测模型相较于未使用 FEEMD 的预测模型更能降低交通流序列的不平稳性,提高了预测精度。这种组合预测模型为短时交通流预测提供了一种新的方法,但本文未考虑路网节点的空间相关性以及天气等因素对路段的影响,因此这些将是本文未来的研究内容。

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] 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.
- [4] 赵亚斌,白璘,武奇生,等. 基于 MPSO-RVM 的短期交通流预测方法研究[J]. 计算机应用研究, 2020, 37(S1):69-71.
ZHAO Y B, BAI L, WU Q SH, et al. Research on short-term traffic flow prediction method based on MPSO-RVM[J]. Application Research of Computers, 2020, 37(S1): 69-71.
- [5] 张阳,杨书敏,辛东嵘. 改进小波包与长短时记忆组合模型的短时交通流预测[J]. 交通运输系统工程与信息, 2020, 20(2):204-210.
ZHANG Y, YANG SH M, XIN D R. Short-term traffic flow prediction based on improved wavelet packet and long-short-term memory combined model [J]. Transportation System Engineering and Information, 2020, 20(2):204-210.
- [6] 陈如清,李嘉春,俞金寿. 基于 FWADE-ELM 的短时交通流预测方法[J]. 控制与决策, 2019, 11(3):1-8.
CHEN R Q, LI J CH, YU J SH. Short-term traffic flow prediction method based on FWADE-ELM [J]. Control and Decision, 2019, 11 (3): 1-8.
- [7] 李利品,党瑞荣,樊养余. 改进的 EEMD 算法及其在多相流检测中的应用[J]. 仪器仪表学报, 2014, 35(10): 2365-2371.
LI L P, DANG R R, FAN Y Y. Improved EEMD algorithm and its application in multiphase flow detection [J]. Chinese Journal of Scientific Instrument, 2014, 35(10): 2365-2371.
- [8] 高晓东,郑连勇,毕斌,等. 基于快速 EEMD 的自适应滤波算法在输电线路激光测距中的应用[J]. 电气自动化, 2018, 40(5):112-115.
GAO X D, ZHENG L Y, BI B, et al. Application of adaptive filtering algorithm based on fast EEMD in transmission line laser ranging [J]. Electrical Automation, 2018, 40(5): 112-115.
- [9] 何怡刚,白月皎,鲁力. 基于 DE-QPSO 算法的 MKRVM 对电容式 RF-MEMS 开关的寿命预测方法[J]. 电子测量与仪器学报, 2020, 34(12):66-75.
HE Y G, BAI Y J, LU L. Life prediction method of capacitive RF-MEMS switches based on DE-QPSO algorithm based on MKRVM [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(12):66-75.
- [10] 董治麟,郑近德,潘海洋,等. 基于复合多尺度排列熵与 FO-SVM 的滚动轴承故障诊断方法[J]. 噪声与振动控制, 2020, 40(2):102-108.
DONG ZH L, ZHENG J D, PAN H Y, et al. Rolling bearing fault diagnosis method based on compound multi-scale permutation entropy and FO-SVM [J]. Noise and Vibration Control, 2020, 40(2): 102-108.
- [11] 陈柯宇,高金凤,吴平. 基于多尺度加权排列熵的管道泄漏检测[J]. 测控技术, 2019, 38(2):118-122.
CHEN K Y, GAO J F, WU P. Pipeline leak detection based on multi-scale weighted permutation entropy [J]. Measurement and Control Technology, 2019, 38 (2): 118-122.
- [12] 钱伟,杨慧慧,孙玉娟. 相空间重构的卡尔曼滤波交通流预测研究[J]. 计算机工程与应用, 2016, 52(14): 37-41.
QIAN W, YANG H H, SUN Y J. Research on Kalman filter traffic flow prediction based on phase space reconstruction [J]. Computer Engineering and Applications, 2016, 52(14): 37-41.
- [13] CAI Y, LI B, WANG L Q. Application of feature extraction through convolution neural networks and SVM classifier for robust grading of apples [J]. Instrumentation, 2019, 6 (4):59-71.
- [14] 韩亮,黄谦,蒲秀娟,等. 使用深度长短期记忆网络的心冲击伪迹抑制方法[J]. 仪器仪表学报, 2020, 41(11):198-206.
HAN L, HUANG Q, PU X J, et al. A cardiac impact artifact suppression method using deep long - short memory networks [J]. Chinese Journal of Scientific Instrument, 2020, 41(11):198-206.
- [15] 宋予佳,张健,邢珺. 基于长短期记忆网络构建短时车流预测模型[J]. 公路, 2019, 64(7):224-229.
SONG Y J, ZHANG J, XING J. Construction of short-term traffic flow prediction model based on long and short-term memory network [J]. Highway, 2019, 64(7): 224-229.
- [16] ZHAO Z, CHEN W H, WU X M, et al. LSTM network: A deep learning approach for short-term traffic forecast [J]. IET Intelligent Transport Systems, 2017, 11(2):68-75.
- [17] 梁旭,刘才慧. 基于混合粒子群算法的在线检测路径规划[J]. 国外电子测量技术, 2015, 34(12):30-34.
LIANG X, LIU C H. Online detection path planning based on hybrid particle swarm algorithm [J]. Foreign

Electronic Measurement Technology, 2015, 34(12): 30-34.

- [18] 赵强,王敬东,刘云霄,等. 基于粒子群 TOA 室内定位的系统误差修正[J]. 电子测量技术, 2017, 40(8): 189-192.

ZHAO Q, WANG J D, LIU Y X. System error correction based on particle swarm TOA indoor positioning [J]. Electronic Measurement Technology, 2017, 40(8): 189-192.

- [19] 杜涵潇,汤旻安. 基于 CEEMD-RSVPSO-KELM 的用户侧微电网短期负荷预测[J]. 电测与仪表, 2020, 57(18): 69-76.

DU H X, TANG M AN. User-side microgrid short-term load forecasting based on CEEMD-RSVPSO-KELM [J]. Electrical Measurement and Instrumentation, 2020, 57(18): 69-76.

- [20] 付雪松,王建林,胡志雄,等. 基于动态权重 PSO 算法的眼科 OCT 设备横向分辨率检测[J]. 仪器仪表学报, 2019, 40(9): 142-150.

FU X S, WANG J L, HU ZH X, et al. Detection of horizontal resolution of OCT device in ophthalmology based on dynamic weight PSO algorithm [J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 142-150.

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 in 2004 from Guangxi of university, and Ph. D. degree in 2007 from Chongqing of university. Now he is an associate professor in Hefei university of technology. His main research interests include communication channel modeling and traffic flow forecast.



魏帅康 (通信作者), 2018 年于江苏海洋大学获得学士学位, 现为合肥工业大学硕士研究生, 主要研究方向为交通流预测、复杂系统建模。

E-mail: 13151729195@163.com

Wei Shuaikang (Corresponding author) received his B. Sc. degree from Jiangsu Ocean University in 2018. He is currently a M. Sc. candidate at Hefei University of Technology. His main research interests include traffic flow forecasting and complex system modeling.