



# 基于分解集成与深度学习模型的5G传输 网络流量预测技术研究

郑才凤<sup>1</sup>, 魏涛<sup>2</sup>, 卢鸣<sup>2</sup>, 陈锋<sup>1</sup>, 沈力<sup>1</sup>, 蔡恬恬<sup>1</sup>, 夏诚<sup>1</sup>, 刘琛<sup>1</sup>, 柳鑫成<sup>1</sup>

(1. 中国移动通信集团设计院有限公司湖北分公司, 湖北 武汉 430021;

2. 中国移动通信集团湖北有限公司, 湖北 武汉 430021)

**摘要:** 随着越来越多的设备和应用依赖于5G网络, 准确的网络流量预测对运营商优化资源分配与保障网络安全管理具有重要意义。针对高度动态与复杂多变的5G传输网络流量难以预测的问题, 提出了基于奇异谱分解(SSD)和基于贝叶斯算法改进的双向长短期记忆(BOA-BiLSTM)网络的新型预测技术, 命名为SSD-BOA-BiLSTM。首先, 利用SSD对原始数据进行分解得到有限个奇异谱分量(SSC)。然后, 采用BOA-BiLSTM对SSC分别进行预测。最后, 叠加所有分量预测值得到最佳预测结果。结果表明, 所提预测技术在5G传输网的核心环、汇聚环与接入环流量的拟合系数分别为97.83%、90.47%和96.28%, 显著优于其他5种流行模型, 为5G流量管理提供了有效方法。

**关键词:** 预测; 奇异谱分解; 贝叶斯算法; 双向长短期记忆网络

**中图分类号:** TN929

**文献标志码:** A

**doi:** 10.11959/j.issn.1000-0801.2025298

## 0 引言

5G技术凭借超高速率、海量连接和超低时延的特点, 正推动物联网、智慧城市和远程医疗等领域的创新发展<sup>[1]</sup>。传输网络流量城域网作为5G体系中的关键基础设施, 采用核心层、汇聚层和接入层3层架构, 共同支撑城市范围内的数据高效传输。随着5G应用场景不断扩展和用户规模快速增长, 网络流量呈现出高度动态与复杂多变的特点。因此, 构建精准高效的流量预测技术, 对实现网络资源的灵活调度和运维成本优化具有重要意义<sup>[2-3]</sup>。

传统的网络流量预测方法, 如高斯模型<sup>[4]</sup>、自回归移动平均(ARMA)<sup>[5]</sup>等, 虽然在一定程

度上能够揭示流量的变化趋势, 但往往面临精度不足、时间复杂度高等问题。尤其在处理具有非线性、非平稳和自相似性等复杂特性的网络流量时, 传统方法的预测效果往往难以达到理想状态<sup>[6]</sup>。近年来, 人工智能技术迅猛发展, 已成功应用于图像识别<sup>[7]</sup>、自然语言处理<sup>[8]</sup>、语音识别<sup>[9]</sup>等多个领域, 如径向基函数(RBF)神经网络<sup>[10]</sup>、长短期记忆(LSTM)网络<sup>[11]</sup>、支持向量机(SVM)<sup>[12]</sup>等模型。此类模型在针对流量变化的非线性、非平稳信号方面具有很大的优势, 其预测精度有一定的提高, 但同时也会导致计算复杂度的增加。与单一人工智能预测模型相比, 组合预测模型在网络流量预测中拥有更高的准确性和适应性, 成为流量预测的主要方向。



目前组合预测模型普遍采用“分解—预测”框架,通过信号分解方法与人工智能模型相结合以提高精度。例如,朱业胜<sup>[13]</sup>采用EMD分解流量序列并建立残差GRU模型,虽提升了预测效果,但未解决EMD模态混叠与噪声干扰问题。赵玉婷等<sup>[14]</sup>提出基于EEMD与ARMA的模型,在缓解模态混叠方面有所改进,但EEMD计算量大且端点效应仍存在。相比之下,奇异谱分解(singular spectrum decomposition, SSD)<sup>[15]</sup>作为一种较新方法,计算效率高、能有效避免模态混叠。在5G传输网络流量预测中,SSD可有效去除噪声并保留关键特征,有助于构建更具针对性的子模型,从而提高整体预测准确性。

现有预测模型大多依赖历史信息对当前值进行预测。以深度学习模型LSTM为例,该类模型虽能够有效捕捉时间序列中的长期依赖关系,但其单向结构限制了模型对未来潜在信息的利用。为克服该局限,深度学习模型双向长短期记忆(bidirectional LSTM, BiLSTM)网络<sup>[16]</sup>在LSTM基础上引入反向隐含层,使其能够同时学习历史与未来的上下文信息。具体而言,BiLSTM网络通过前向隐含层处理历史数据,反向隐含层处理未来数据,并在输出层融合双向信息,从而显著提升了预测准确性<sup>[17]</sup>。例如,曹成远等<sup>[18]</sup>在蛋白质残基相互作用预测任务中使用BiLSTM网络,相较传统LSTM模型精度提升超过10%,体现出其在时序建模中的优越性能。然而,BiLSTM网络的性能很大程度上受学习率、隐含层神经元数量等超参数的影响。传统超参数优化方法往往计算开销大,且易陷入局部最优。针对该问题,本文采用贝叶斯优化算法(Bayesian optimization algorithm, BOA)<sup>[19]</sup>对BiLSTM网络超参数进行自动调优。BOA基于高斯过程建立目标函数的概率代理模型,并借助采集函数指导参数搜索方向,能够在较少迭代次数内获得接近最优的超参数组

合,从而提升模型适应性与预测精度。本文所提出的SSD-BOA-BiLSTM组合预测方法,融合了SSD、BOA和BiLSTM网络三者的优势。该技术在提高预测准确度的同时,也加强了对不同流量动态模式的适应能力。

## 1 基本原理

### 1.1 奇异谱模态分解

SSD自从2014年被Bonizzi等<sup>[20]</sup>提出后,已经应用到诸多工程学领域。SSD在SSA(singular spectrum analysis, SSA)的基础上<sup>[21]</sup>,通过分析待分解信号的频域信息来实现参数的自动化选择,这也使得SSD成为一种优秀的自适应分解算法<sup>[22-23]</sup>。其基本流程如下。

#### (1) 构建轨迹矩阵

若时间序列为 $x(n)$ ,其嵌入维数为 $M$ ,长度为 $N$ 。构建序列 $x(n)$ 为一个 $M \times N$ 的矩阵 $X$ ,则矩阵 $X$ 的第 $i$ 行可表示为 $X_i = [x_i, x_{i+1}, \dots, x_N, x_1, \dots, x_{i-1}]$ ,其中 $i = 1, 2, \dots, M$ 。若时间序列 $x(n) = [1, 2, 3, 4, 5]$ ,  $M = 3$ ,则相应矩阵 $X$ 为:

$$X = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \end{bmatrix} \quad (1)$$

其中,左边3列形成的矩阵为轨迹矩阵,将矩阵 $X$ 中右下角的3个元素搬到左上角,得到一个新的轨迹矩阵 $\hat{X}_N \times K$ :

$$\hat{X} = \begin{bmatrix} & & & & 1 \\ & & & & 2 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & \\ 3 & 4 & 5 & & \end{bmatrix} \quad (2)$$

此时 $\hat{X}$ 为新的规矩阵,可以有效增强信号中的震荡成分。

#### (2) 自适应选择嵌入维数 $M$

SSD通过自适应准则确定 $M$ 。计算残余分量 $v_j(n)$ 功率谱密度(PSD), $v_j(n)$ 的表达式如下。

$$v_j(n) = x(n) - \sum_{k=1}^{j-1} v_k(n), v_0(n) = y \quad (3)$$

其中,  $v_{kn}$  为已分解出的分量。

$$\text{PSD} = \int V_j(n) e^{-j2\pi f n} dn \quad (4)$$

根据 PSD 确定最大频率  $f_{\max}$ 。当迭代次数  $j=1$  时, 则归一化频率为  $f_{\max}/F_s \leq 10^{-3}$ , 残差将作为重大分量, 此时将嵌入维数设置为:  $M=N/3$ 。若迭代次数  $j>1$ , 则设置嵌入维数为  $M=1.2 \times \frac{F_s}{f_{\max}}$ 。

(3) 重构分量。

每个高斯函数代表一个谱峰, 其表达式为:

$$v_{(f,\theta)} = \sum_{i=1}^3 A_i \exp\left(-\frac{(f-\mu_i)^2}{2\alpha_i^2}\right) \quad (5)$$

其中,  $A_i$  表示第  $i$  个高斯函数的幅值;  $\alpha_i$  表示其带宽;  $\mu_i$  表示其位置点;  $\theta = [A\alpha]^T$  表示参数矢量, 并且满足  $A = [A_1, A_2, A_3]$ ,  $\alpha = [\alpha_1, \alpha_2, \alpha_3]$ 。

第一个函数为主谱峰, 频率表示为  $f_{\max}$ ; 第二个函数为次谱峰, 频率表示为  $f_2$ ; 第 3 个函数记录主次谱峰之间的任意频率, 即:

$$u_1 = f_{\max}, u_2 = f_2, u_3 = \frac{f_{\max} + f_2}{2} \quad (6)$$

为了获得模型参数  $A_i$ , 首先给定拟合的初始值, 再对模型进行加权最小二乘拟合, 即:

$$\begin{cases} A_1^0 = \frac{1}{2} \text{PSD}(f_{\max}), \alpha_1^0 = f_{\max} \\ A_2^0 = \frac{1}{2} \text{PSD}(f_2), \alpha_2^0 = f_2 \\ A_3^0 = \frac{1}{4} \text{PSD}(f_3), \alpha_3^0 = 4|f_{\max} - f_2| \end{cases} \quad (7)$$

参数  $A_i$  最优值根据 Levenberg-Marquardt 算法确定。主峰带宽  $\Delta f = 2.5\alpha_1$ ,  $\alpha_1$  为设定的初始值, 并继续进行迭代, 重构第  $j$  个 SSC。为保证迭代收敛性, 所以第二次迭代需要满足:

$$\sum_{i=1}^N [g_2^j(i)(v^j(i) - g_2^j(i))] > 0 \quad (8)$$

否则, 第一次运行的估计值  $g_1^j(n)$  将被当作  $g^j(n)$  的估计值。最后,  $g^j(n)$  和剩余时间序列  $v^j(n)$  的差值用尺度因子  $\hat{a}$  来进行调整, 即:

$$\hat{a} = \min_a \|v^j(n) - a\hat{g}^j(n)\|_2^2 \quad (9)$$

其中,

$$\hat{a} = \frac{g^T v}{g^T g}, \tilde{g}^j(n) = \hat{a} g^j(n) \quad (10)$$

(4) 迭代终止条件

第  $j$  次迭代后的残余信号  $\text{res}_n^j$  为:

$$\text{res}_n^j = y - \sum_{i=1}^j \text{SSC}_i \quad (11)$$

计算  $\text{res}_n^j$  与原始信号的归一化均方误差 (NMSE):

$$\text{NMSE} = \frac{\sum_{i=1}^N (\text{res}_i^j)^2}{\sum_{i=1}^N y_i^2} \quad (12)$$

设定 NMSE 的下限阈值 0.5%, 当计算结果小于下限阈值, 终止迭代。

假设 SSD 最终分解出  $k$  个 SSC, 此时的残余信号为  $\text{res}_n^k$ , 则有:

$$y = \sum_{i=1}^k \text{SSC}_i + \text{res}_n^k \quad (13)$$

奇异谱分解流程如图 1 所示。

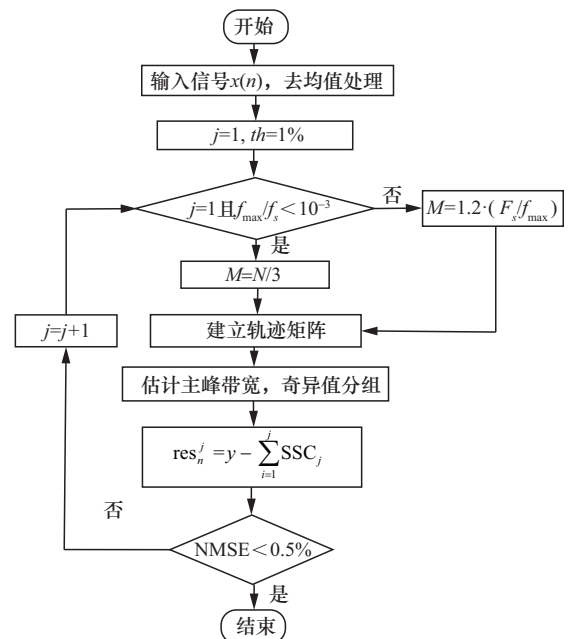


图1 奇异谱分解流程



## 1.2 贝叶斯优化算法

BOA为深度学习模型提供新的超参数优化方式<sup>[24]</sup>。其主要原理是根据已知黑盒函数的输出概率寻找函数的最优值。BOA的优势在于只需要为目标函数设定输入和输出,不需要了解其数学属性和内部结构,其迭代次数少、效率高,算法也很稳定。该算法超参数寻优取值为:

$$X^* = \arg \min_{x \in X} f(x) \quad (14)$$

其中,  $X^*$  为最佳参数集,  $x$  为模型的超参数集合,  $X$  为先验样本集合,  $f(x)$  为需要优化的目标模型。

## 1.3 双向长短期记忆网络

LSTM网络是从递归神经网络演变而来的深度学习网络模型<sup>[25]</sup>,可以改善递归神经网络模型中的一些缺点,如梯度消失<sup>[26]</sup>。LSTM网络共包含4种结构:记忆单元、遗忘门 $f_t$ 、输入门 $i_t$ 和输出门 $O_t$ ,LSTM内部结构如图2所示。

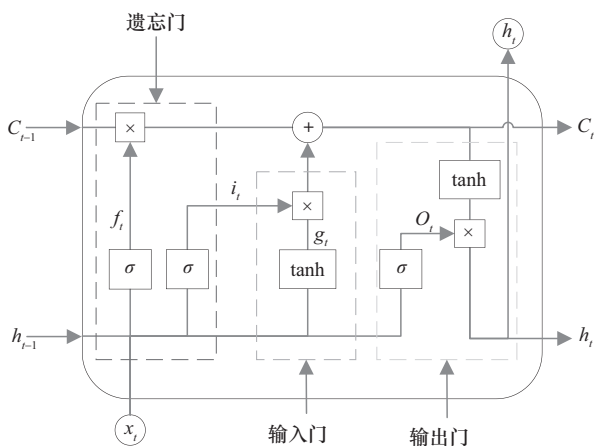


图2 LSTM内部结构

由图2可知,LSTM网络与递归神经网络相比,增加了具有3个门组件的单元状态。LSTM结构中的遗忘门 $f_t$ 负责确定该网络中剩余的信息百分比,其计算如下。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (15)$$

其中,  $x_t$  是输入序列,  $h_{t-1}$  是前一时刻的状态记忆,  $\sigma(\cdot)$  是S形激活函数,  $f_t$  是遗忘门的状态。输

入门负责选择性地存储单元状态下的新信息,其计算如下。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (16)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \sqrt{2} \quad (17)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (18)$$

其中,  $\tilde{C}_t$  是细胞状态候选,  $C_t$  是新的细胞状态,  $\tanh(\cdot)$  是双曲正切函数,  $W_i$  是输入门的权重矩阵,  $W_c$  是单元状态的权重矩阵,  $b_i$  是输入门的偏置,  $b_c$  是细胞状态的偏差,并且它是输入门的状态。输出门 $O_t$ 负责确定输出信息的当前状态,计算如下。

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (19)$$

$$h_t = O_t \cdot \tanh(C_t) \quad (20)$$

其中,  $W_o$  是输出门的权重矩阵;  $b_o$  是输出门的偏置;  $O_t$  是输出门的状态。

BiLSTM网络结合了前向和后向信息的双向长短期记忆模型,即它可以在两个方向上处理信息,克服了LSTM神经网络数据信息不足的缺点。正向和反向都有隐藏层,这些隐藏层可以在给定时间内一起提取正向和反向关键信息,有助于提高模型的预测精度。BiLSTM网络结构如图3所示。

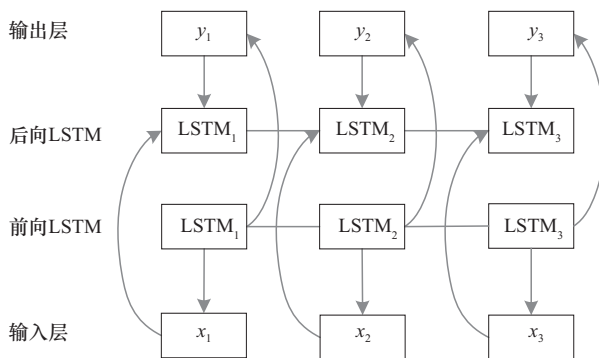


图3 BiLSTM网络结构

## 1.4 基于贝叶斯算法优化的BiLSTM网络模型

由于BiLSTM网络的性能与参数的选择密切相关,为了提高BiLSTM网络的预测能力,避免

陷入局部最优, 引入贝叶斯算法对 BiLSTM 网络的参数进行优化。将均方误差作为适应度函数, 每次迭代时均选择最小误差时的参数, 可以得到使 BiLSTM 网络预测性能最强时的参数。BOA-BiLSTM 的具体步骤如下。

(1) 对 5G 传输网络流量数据做预处理, 输入待预测序列, 划分训练集和测试集。将 BiLSTM 网络的学习率和隐含层节点数作为变量。

(2) 初始化 BOA 参数, 最大迭代次数设置为 50, 建立 BOA-BiLSTM 预测模型。

(3) 判断是否满足目标值, 如果达到, 停止迭代, 否则继续迭代, 得到最佳适应度值的 BiLSTM 网络参数。

(4) 模型预测及评价, 对模型 5G 传输网络流量进行预测, 并将预测值反归一化, 得到最终预测值。BOA-BiLSTM 流程如图 4 所示。

### 1.5 评价指标

本文采用均方根误差 (RMSE)、平均绝对误差 (MAE) 和平均绝对百分比误差 (MAPE) 来

定量分析预测模型性能, 它们的表达式分别如下。

$$E_{\text{RMSE}} = \sqrt{\frac{\sum_{i=1}^N (y(i) - y_d(i))^2}{N}} \quad (21)$$

$$E_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |y(i) - y_d(i)| \quad (22)$$

$$E_{\text{MAPE}} = \frac{1}{N} \sum_{i=1}^n \left| \frac{y(i) - y_d(i)}{y_d(i)} \right| \quad (23)$$

其中,  $N$  为预测样本点数,  $y_d(i)$  为待预测时间序列各点的实际结果,  $y(i)$  为模型的预测结果。RMSE、MAE 和 MAPE 值越小, 说明预测能力越强。

## 2 基于改进 BiLSTM 网络的 5G 传输网络流量预测技术研究

### 2.1 数据描述

为了验证所提出技术的有效性和稳定性, 对

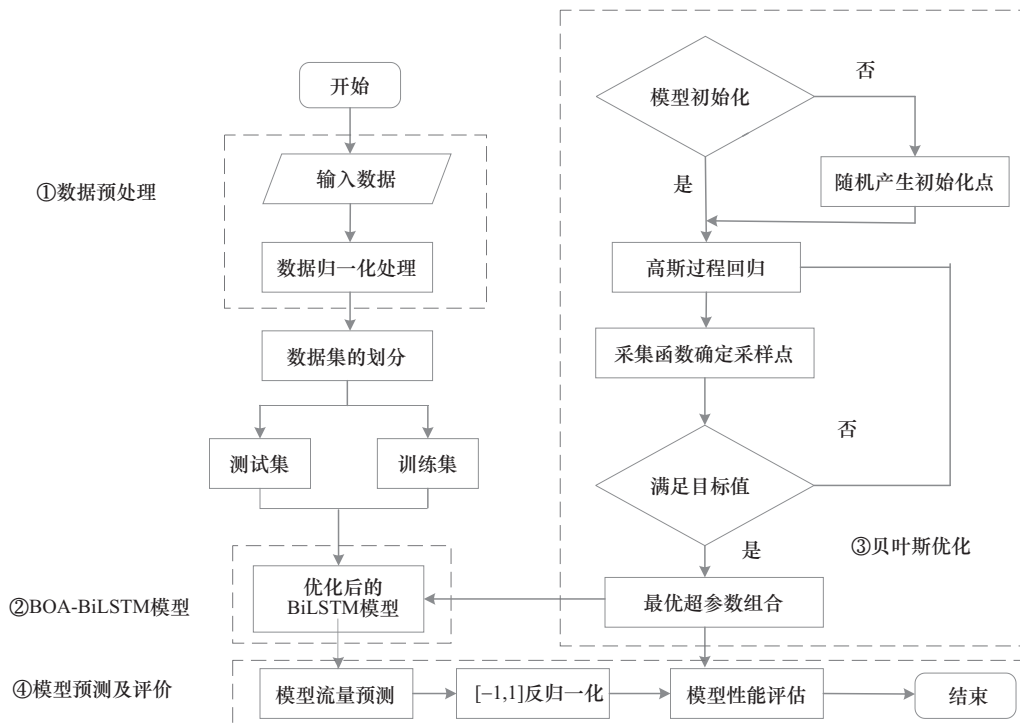


图4 BOA-BiLSTM 流程



5G 传输网络流量序列进行预测。选取的湖北移动 5G 传输网络流量序列位于孝感监测点, 以数据来自湖北移动运营商公开的部分数据集为例。选取 2024 年 11 月 1 日到 2025 年 2 月 28 日连续 120 天的 5G SPN 传输网络流量作为实验数据, 其中前 84 天的数据是训练数据, 后 36 天数据是测试数据。核心环、汇聚环以及接入环流量分别如图 5、图 6 与图 7 所示。

由图 5、图 6 与图 7 可知, 核心环、汇聚环和接入环流量自动分成 12、9 与 8 组。其中核心环流量大致分布在 40 000~45 000 Mbit/s, 汇聚环流量集中分布在 17 500~25 000 Mbit/s, 接入环

流量集中分布在 5 500~7 500 Mbit/s。

## 2.2 预测模型建立

5G 传输网络流量具有非线性和非平稳性特点, 预测难度较大。为了提高预测精度, 本文提出一种基于 SSD、BOA-BiLSTM 网络的 5G 传输网络流量预测技术, 命名为 SSD-BOA-BiLSTM。技术的主要思想是, 先用 SSD 对 5G 传输网络流量进行分解, 其次, 用 BOA-BiLSTM 预测模型对分解后的各 SSC 进行预测, 最后将各 SSC 的预测结果进行重构, 得到最终预测值。SSD-BOA-BiLSTM 预测模型如图 8 所示, 其具体步骤如下。

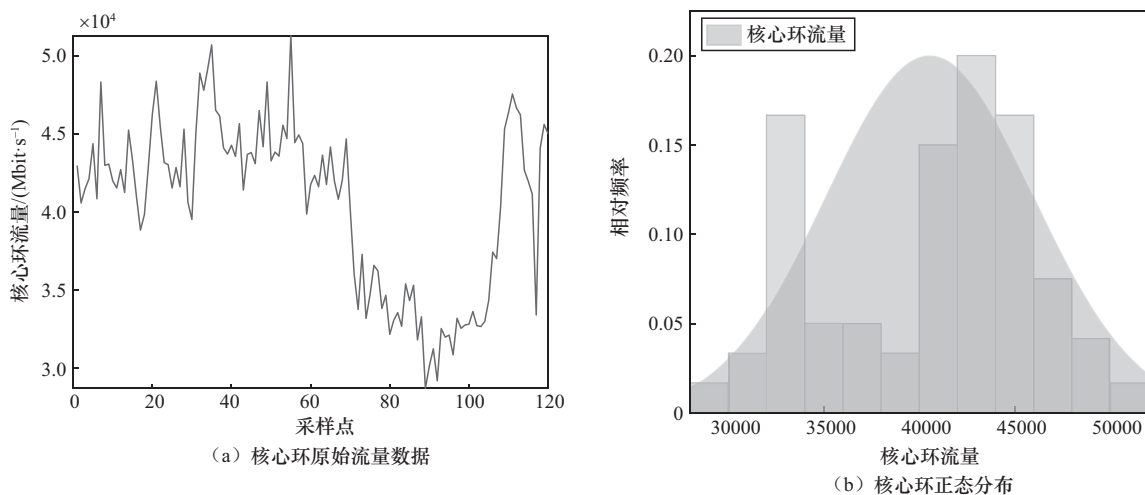


图5 核心环流量

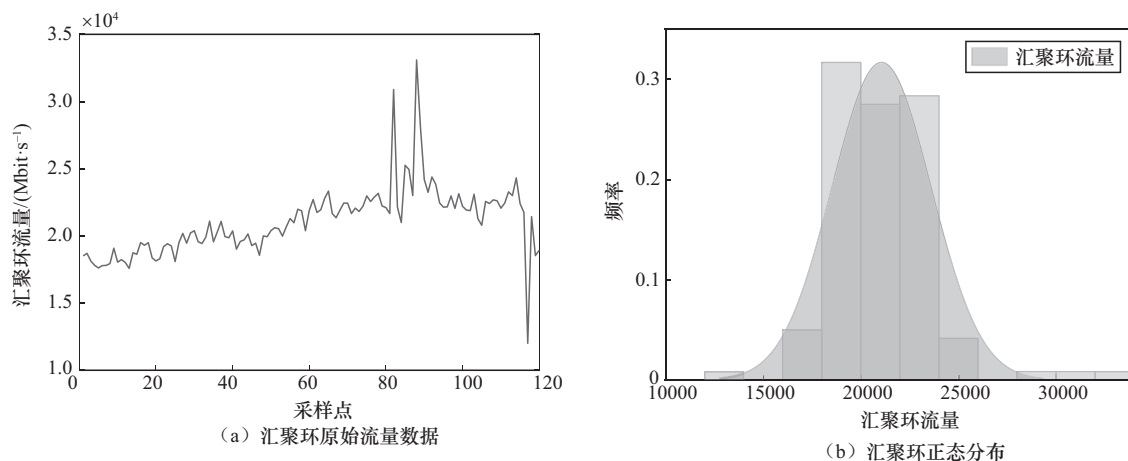


图6 汇聚环流量

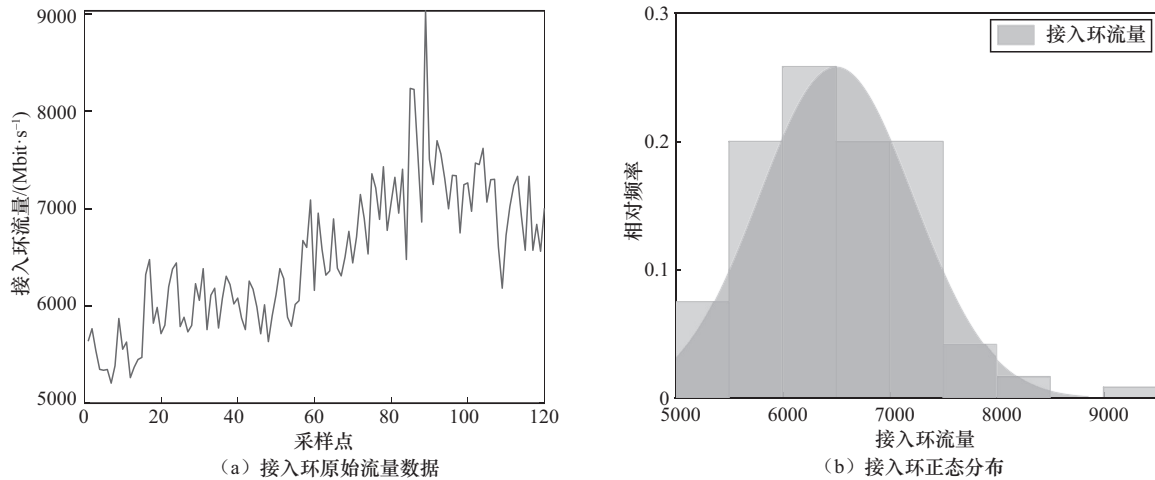


图7 接入环流量

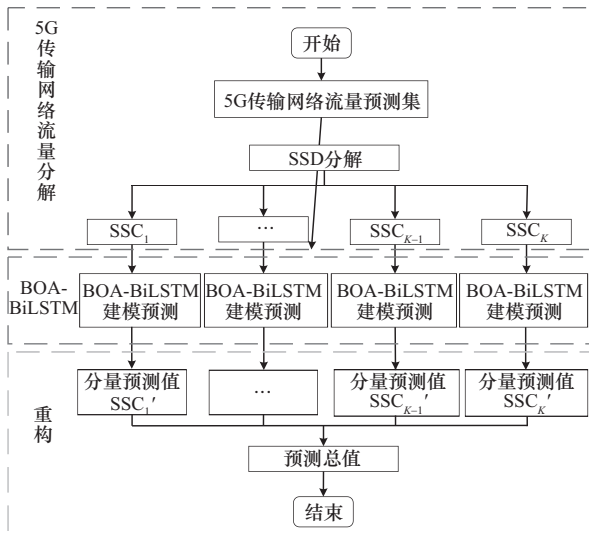


图8 SSD-BOA-BiLSTM 预测模型

(1) 对 5G 传输网络流量中存在的缺失数据用前后数据的平均值对单个缺失数据进行插值，用 3 次样条方法对连续缺失数据进行插值，消除异常数据点。

(2) 用 SSD 对预处理后的 5G 传输网络流量进行分解，得到多个 SSC。将分量按照 7:3 的比例划分为训练集和测试集，训练集用于对预测模型的性能进行训练，测试集用来验证预测模型的预测效果。

(3) 使用 BOA-BiLSTM 对 SSD 分解后得到的各 SSC 进行预测，得到各 SSC 预测结果。

(4) 将所有 SSC 的预测结果进行重构，得到 5G 传输网络流量的最终预测值。

### 2.3 实验参数设置及评价标准

为了验证 BOA-BiLSTM 预测模型的有效性，将 SSD-BOA-BiLSTM 预测模型分别应用于 5G 传输网 SPN 核心环、汇聚环、接入环流量预测实例中。本文采用 BOA 对 BiLSTM 参数优化时的参数设置见表 1。核心环、汇聚环、接入环流量 BOA-BiLSTM 预测结果分别如图 9、图 10 和图 11 所示。

表1 BOA-BiLSTM 参数设置

| 主要参数        | 参数值                   |
|-------------|-----------------------|
| BiLSTM1 神经元 | 130                   |
| BiLSTM2 神经元 | 82                    |
| 最大迭代次数      | 50                    |
| 学习率范围       | $[10^{-3}, 1]$        |
| L2 权重衰减系数范围 | $[10^{-10}, 10^{-2}]$ |

由图 9、图 10 与图 11 可知，采用 BOA-BiLSTM 模型对核心环、汇聚环与接入环流量进行预测，流量的流速峰值真实值与实际值较为接近。另外，选取未来 10 天的流量进行预测，预测的结果如图 11 (b) 所示，从而有效实时监控 5G 传输网络流量的流速峰值，避免网络故障。

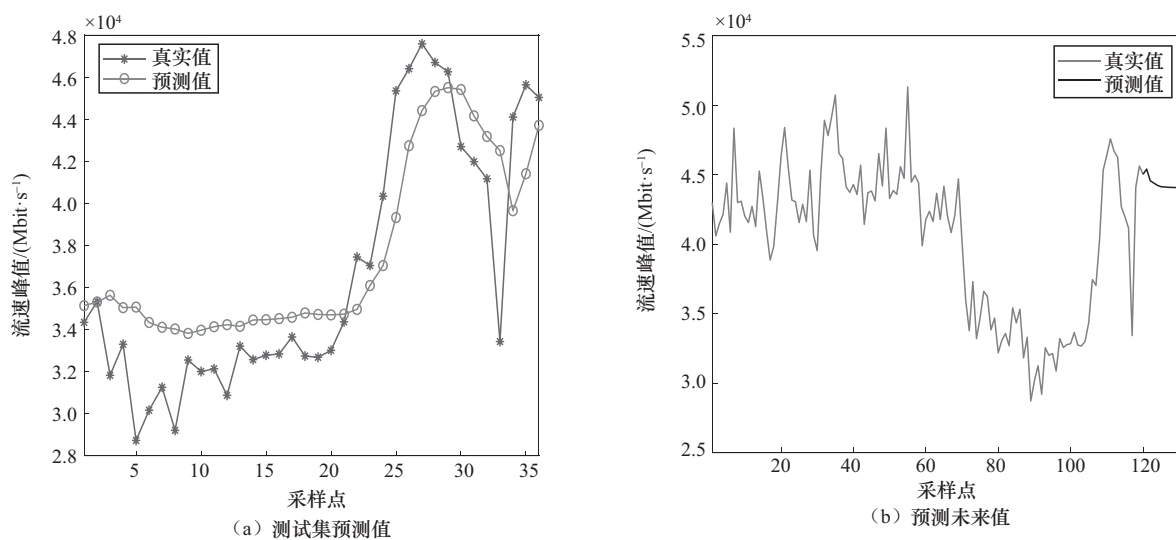


图9 核心环流量BOA-BiLSTM预测结果

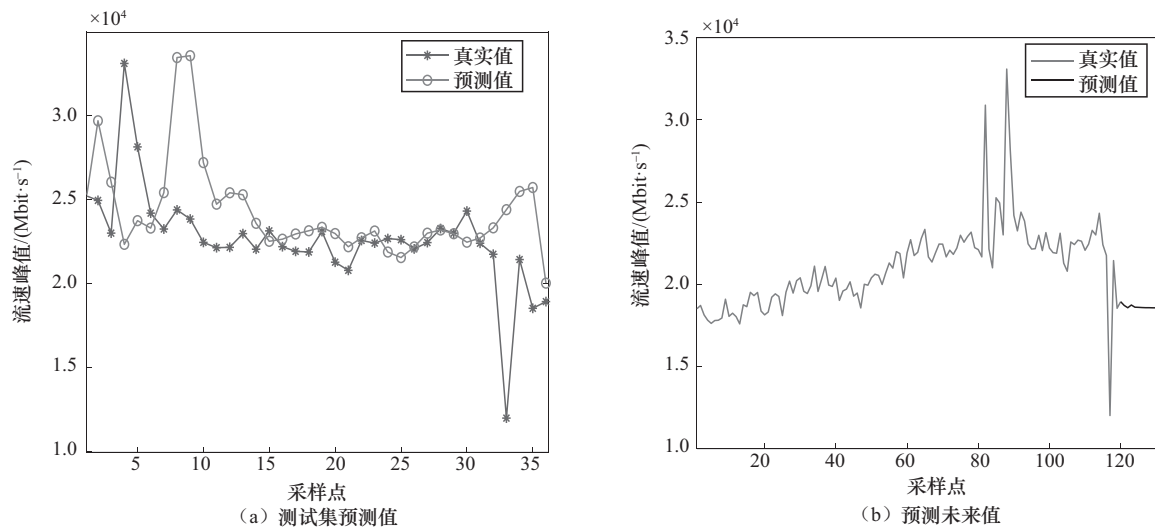


图10 汇聚环流量BOA-BiLSTM预测结果

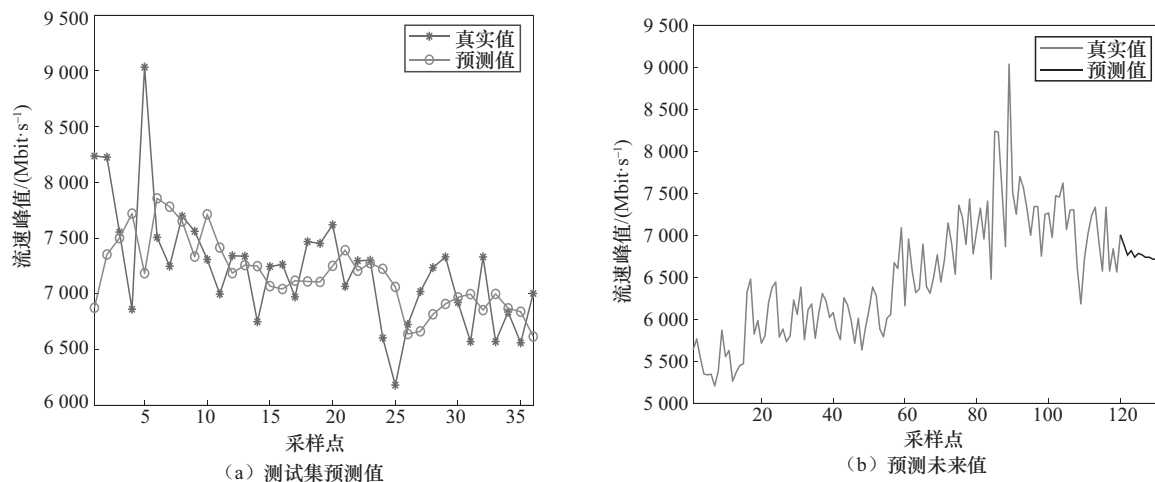


图11 接入环流量BOA-BiLSTM预测结果

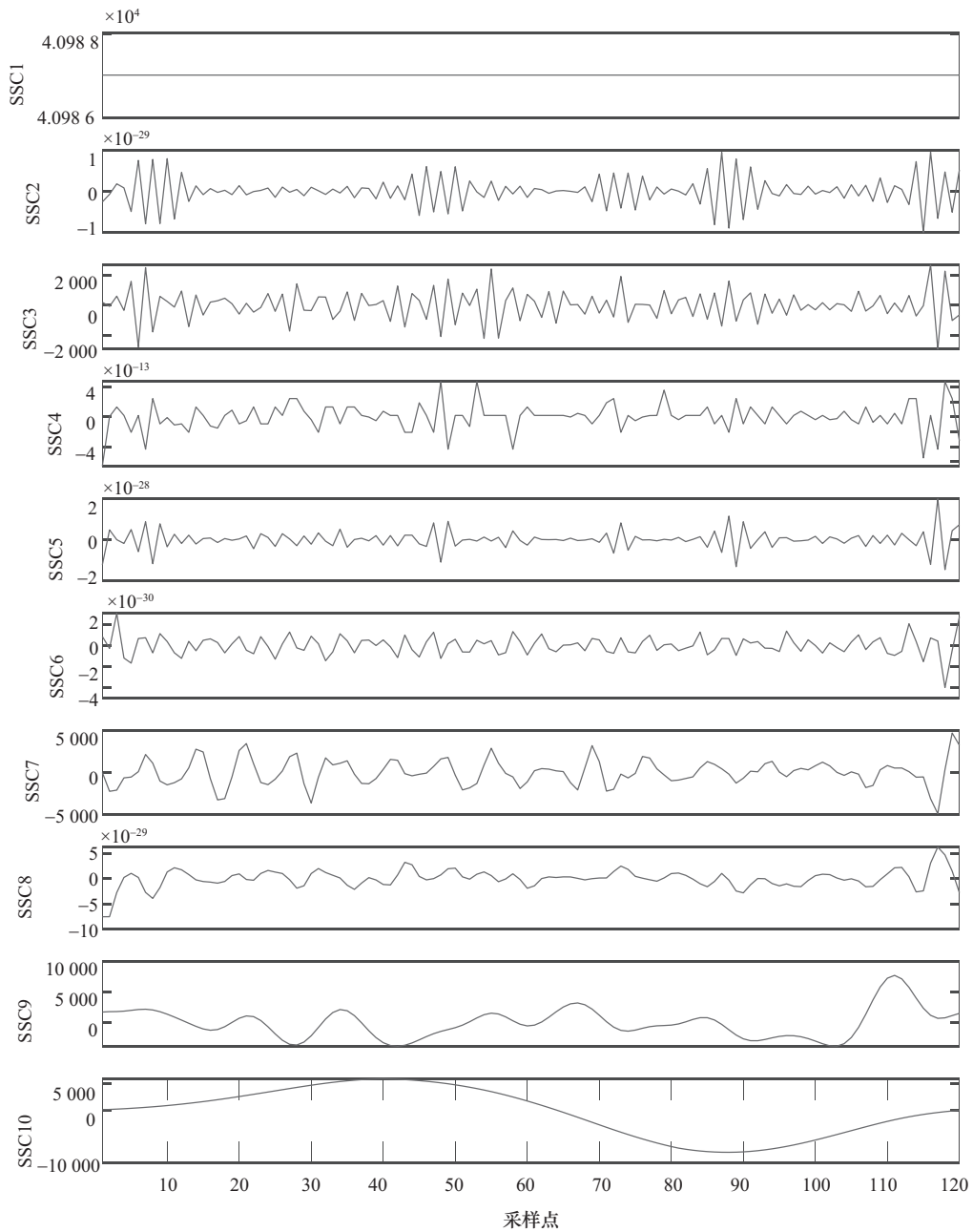
### 3 仿真实验与结果分析

#### 3.1 基于奇异谱分解的预处理

根据所提出的组合预测模型, 由 SSD 对核心环、汇聚环、接入环流量进行分解。分解的子分量被命名为 SSC1、SSC2、 $\dots$ 、SSC $n$ 。分解原始 5G 传输网络流量的结果如

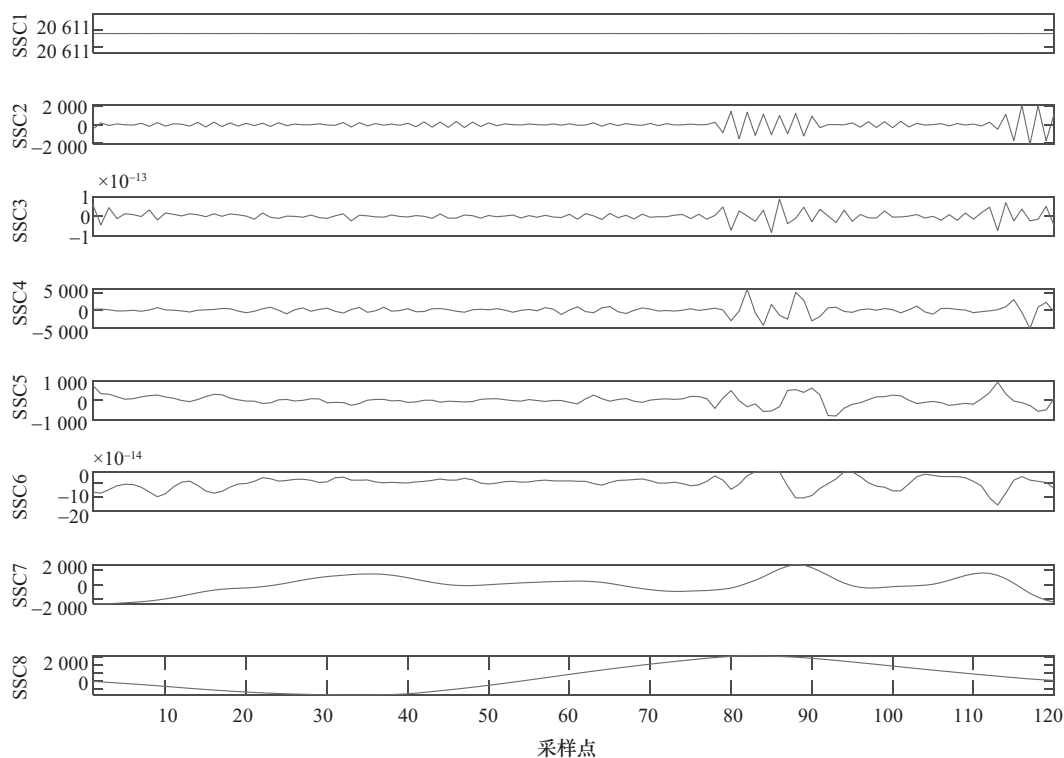
图 12 所示, 显示了经过分解之后获得的若干个子分量。

由图 12 可知, 核心环流量分解为 1 个直流分量和 9 个 SSC 子分量, 汇聚环流量分解成 1 个 SSC 与 7 个 SSC 子分量, 接入环流量分解成 1 个 SSC 与 8 个 SSC 子分量。由于 SSC 为直流分量, 信号较为稳定, 因此不需要参与预测。

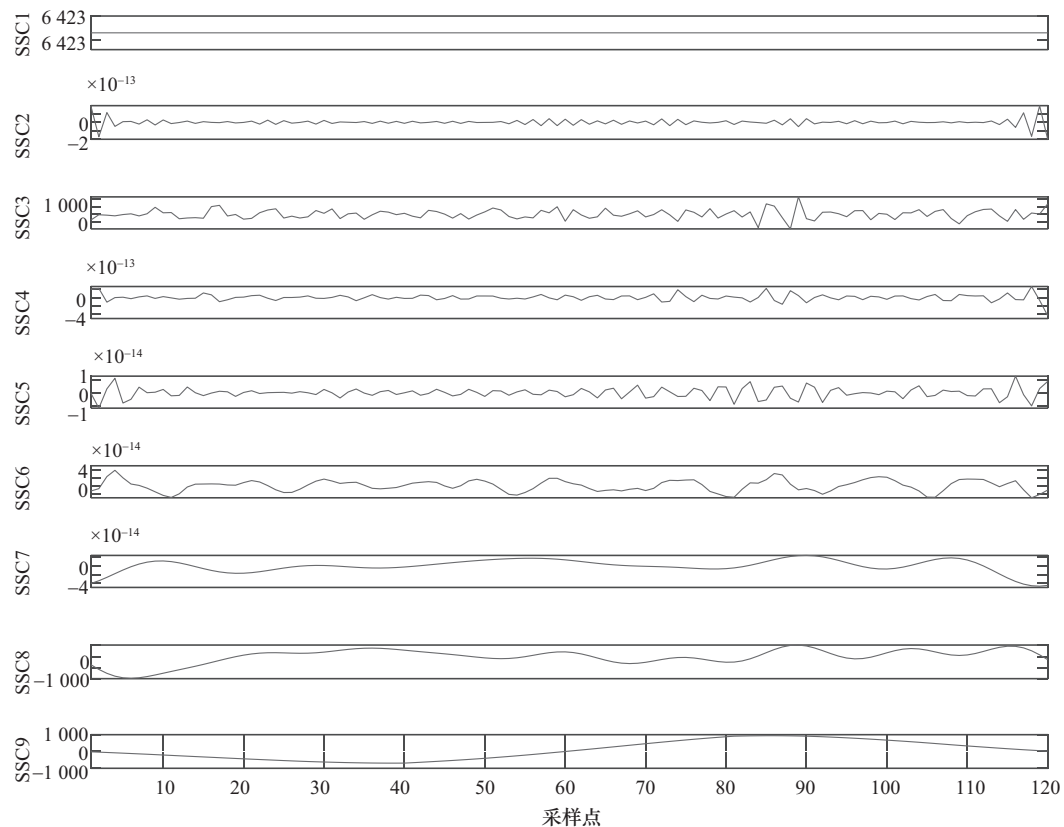


(a) 核心环流量SSD分解的结果

图 12 SSD 原始 5G 传输网络流量的结果



(b) 汇聚环流量SSD分解的结果



(c) 接入环流量SSD分解的结果

图12 SSD原始5G传输网络流量的结果(续)

### 3.2 实验结果与分析

通过奇异谱分解分别得到5G传输网络流量核心环、汇聚环与接入环的子分量之后,使用BOA-BiLSTM模型进行预测各个子分量,重构所有的子分量的预测结果,得到5G传输网络流量最终预测值。SSD-BOA-BiLSTM预测结果如图13所示。

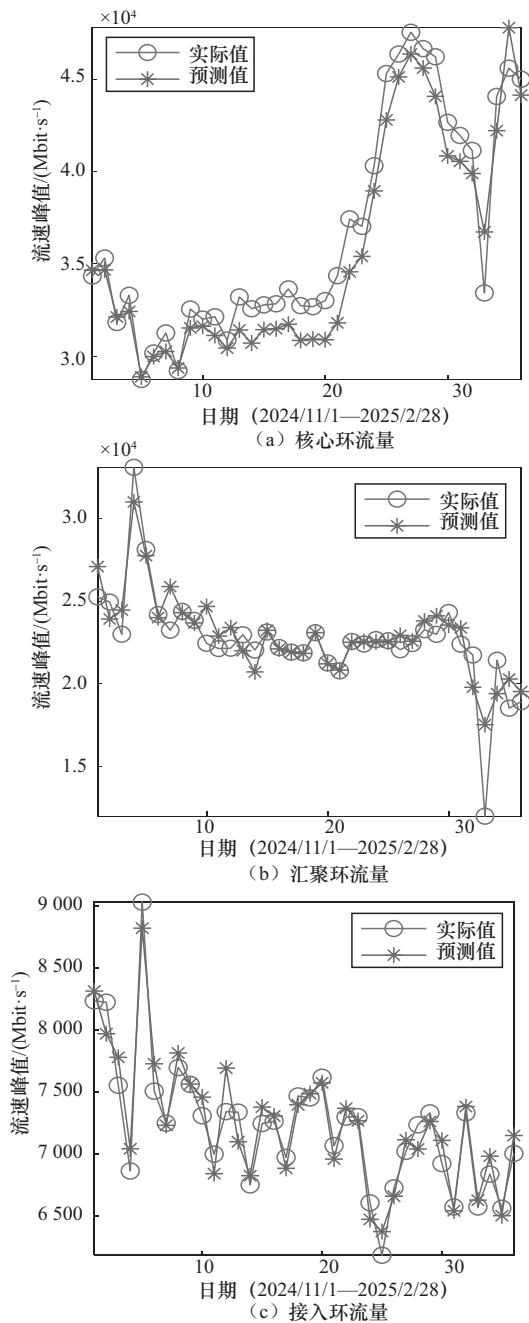


图13 SSD-BOA-BiLSTM预测结果

为了证明SSD-BOA-BiLSTM的优越性,选取了5种比较流行的对比模型,使用核心环、汇聚环、接入环流量数据进行对比实验。对比模型分别为BP、LSTM、BiLSTM、BOA-BiLSTM和VMD-BOA-BiLSTM。

图14显示了核心环、汇聚环、接入环流量数据集在2024年11月1日到2025年2月28日的5G传输网络流量预测结果。

为评估所提出模型的5G传输网络流量预测效果,比较了传统神经网络与基于深度学习的预测模型,单一深度学习预测模型有LSTM、BiLSTM和BOA-BiLSTM,分解集成组合预测模型有VMD-BOA-BiLSTM和SSD-BOA-BiLSTM。本文从定量角度给出了各个模型具体的预测性能指标。各个预测模型的RMSE、MAE、MAPE具体数据见表2。为直观展示各个指标,图15给出了核心环、汇聚环和接入环5G传输网络流量各个预测模型的RMSE、MAE、MAPE雷达图。

由表2和图15可知,在核心环、汇聚环和接入环的流量预测中,SSD-BOA-BiLSTM模型的RMSE、MAE和MAPE 3项误差指标均最小,与原始数据的拟合度最优。单一预测模型中,BP和LSTM表现较差,以核心环流量预测为例,BP模型的RMSE和MAE均较大,其中RMSE超过5000,MAE接近5000;而BiLSTM模型的RMSE约为3872,误差较BP和LSTM有明显降低。所提出的BOA-BiLSTM模型在BiLSTM基础上进一步优化,RMSE降至3479.6106;VMD-BOA-BiLSTM模型表现更优,RMSE和MAE较BOA-BiLSTM分别降低392.9294和351.994。经SSD分解后,环路流量分量得到更充分的拆分,使得预测误差最小、拟合度最高,充分验证了该模型能融合SSD、BOA与BiLSTM的优势,通过简化子序列为精准预测奠定基础。核心环、汇聚环、接入环流量各模型的预测误差小提琴图分别如图16所示。

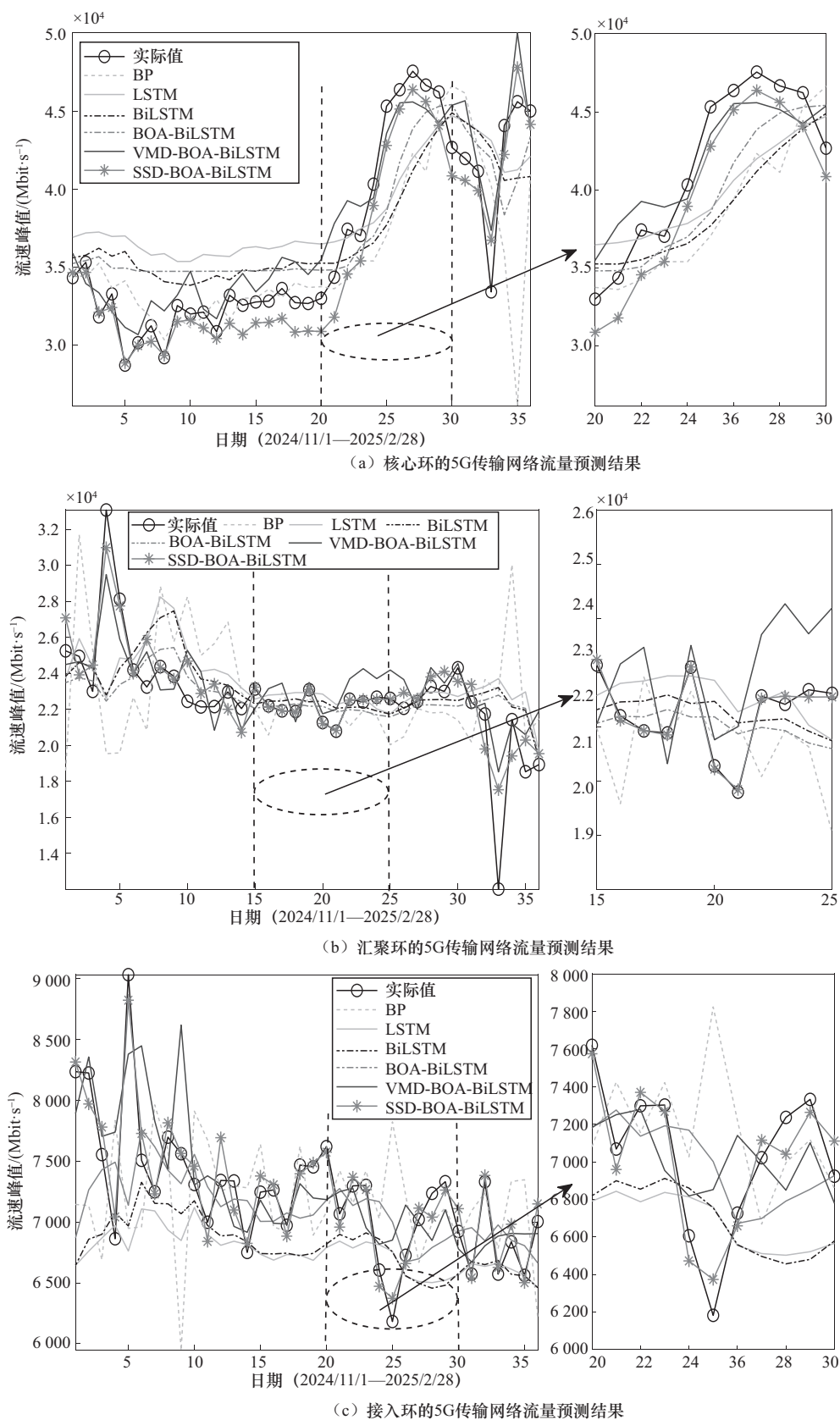


图14 所提模型与对比模型的5G传输网络流量预测结果比较

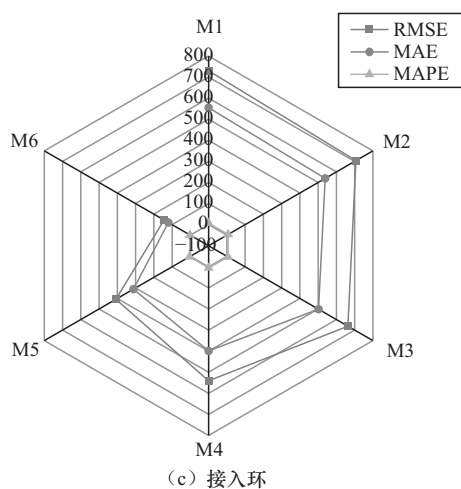
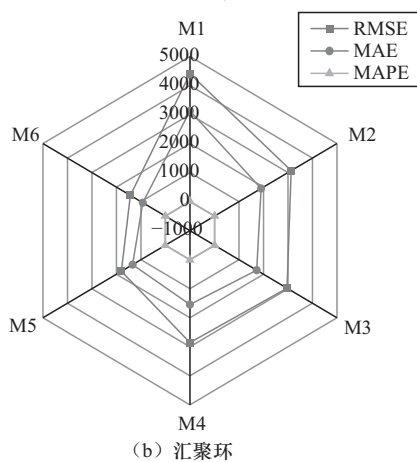
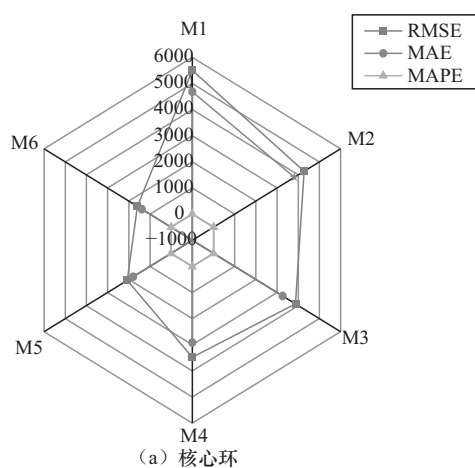


图15 各模型评价雷达图

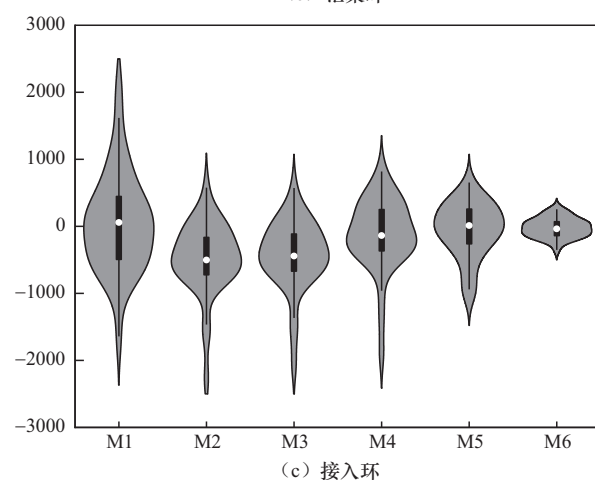
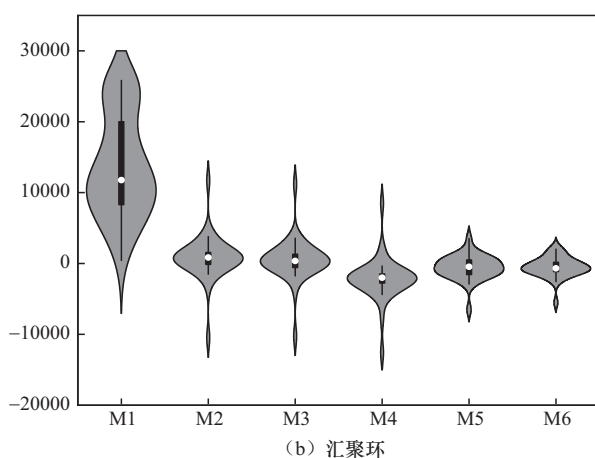
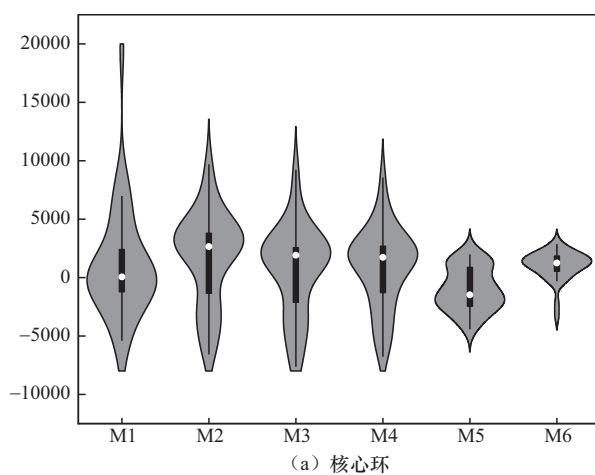


图16 各模型的预测误差小提琴图

从图16可以看出,“白点”表示中位数,M1-M6模型的中位数分布不同,其中M6模型的中位数大致分布在0附近。“黑矩形”表示25%~75%四分位距,反映数据的离散程度,其中M6模型的误差离散程度最小,进一步印证了SSD-

BOA-BiLSTM模型误差集中且离散程度最低的显著优势。箱线图的“须”,定义数据的合理分布区间,超出此范围的点通常被视为潜在异常值。泰勒图可分析所提模型的可靠性和有效性,通过相关系数、均方根误差和标准差之间的关系,显



表2 各模型评价结果

| 地点  | 编号        | 模型                    | RMSE               | MAE                | MAPE           |
|-----|-----------|-----------------------|--------------------|--------------------|----------------|
| 核心环 | M1        | BP                    | 5 503.800 0        | 4 682.600 0        | 13.180         |
|     | M2        | LSTM                  | 4 282.200 0        | 3 820.600 0        | 10.878 0       |
|     | M3        | BiLSTM                | 3 872.540 0        | 3 264.533 9        | 8.959 3        |
|     | M4        | BOA-BiLSTM            | 3 479.610 6        | 2 912.539 9        | 8.158 8        |
|     | M5        | VMD-BOA-BiLSTM        | 2 072.200 0        | 1 793.300 0        | 4.978 00       |
|     | <b>M6</b> | <b>SSD-BOA-BiLSTM</b> | <b>1 593.500 0</b> | <b>1 384.500 0</b> | <b>3.745 7</b> |
| 汇聚环 | M1        | BP                    | 4 404.400 0        | 3 123.100 0        | 14.220 0       |
|     | M2        | LSTM                  | 3 131.434 9        | 1 903.416 9        | 9.284 6        |
|     | M3        | BiLSTM                | 2 962.300 0        | 1 723.900 0        | 8.384 2        |
|     | M4        | BOA-BiLSTM            | 2 854.400 0        | 1 544.100 0        | 7.522 7        |
|     | M5        | VMD-BOA-BiLSTM        | 1 798.300 0        | 1 342.700 0        | 6.580 7        |
|     | <b>M6</b> | <b>SSD-BOA-BiLSTM</b> | <b>1 430.400 0</b> | <b>926.580 0</b>   | <b>4.644 9</b> |
| 接入环 | M1        | BP                    | 729.658 5          | 554.702 6          | 7.599 4        |
|     | M2        | LSTM                  | 707.320 0          | 539.120 0          | 7.128 9        |
|     | M3        | BiLSTM                | 663.027 0          | 502.281 7          | 6.659 9        |
|     | M4        | BOA-BiLSTM            | 540.830 0          | 398.860 0          | 5.371 8        |
|     | M5        | VMD-BOA-BiLSTM        | 408.060 0          | 311.872 3          | 4.305 7        |
|     | <b>M6</b> | <b>SSD-BOA-BiLSTM</b> | <b>144.019 1</b>   | <b>120.361 3</b>   | <b>1.659 9</b> |

示对比模型预测目标数据的能力。5G 传输网络流量预测的泰特图如图 17 所示。

在图 17 中,用三角形符号表示的方法均方根偏差较低,标准差较低,相关系数值更接近于 1,说明实际值与预测值之间的相关性更加线性。三角形符号坐标位置最接近观测点的坐标位置,表明 SSD-BOA-BiLSTM 模型的预测性能是最好的。散点图能更直观地展示各模型的实际值与预测值之间的关系。各模型 5G 传输网络流量原始值与预测值的散点图如图 18 所示。从图 18 可以看出,

提出的 SSD-BOA-BiLSTM 模型的散点最为集中分布在趋势线两侧,表明提出的预测模型是可靠和准确的。

为验证 SSD-BOA-BiLSTM 模型与 5 个对比模型在预测性能上的差异,本节基于上述对比实验,采用 Diebold-Mariano (DM) 统计方法,对 5 个基准模型与 SSD-BOA-BiLSTM 模型展开检验,以判断它们之间是否存在显著差异。5G 传输网络流量的 DM (Diebold-Mariano) 统计检验结果见表 3。

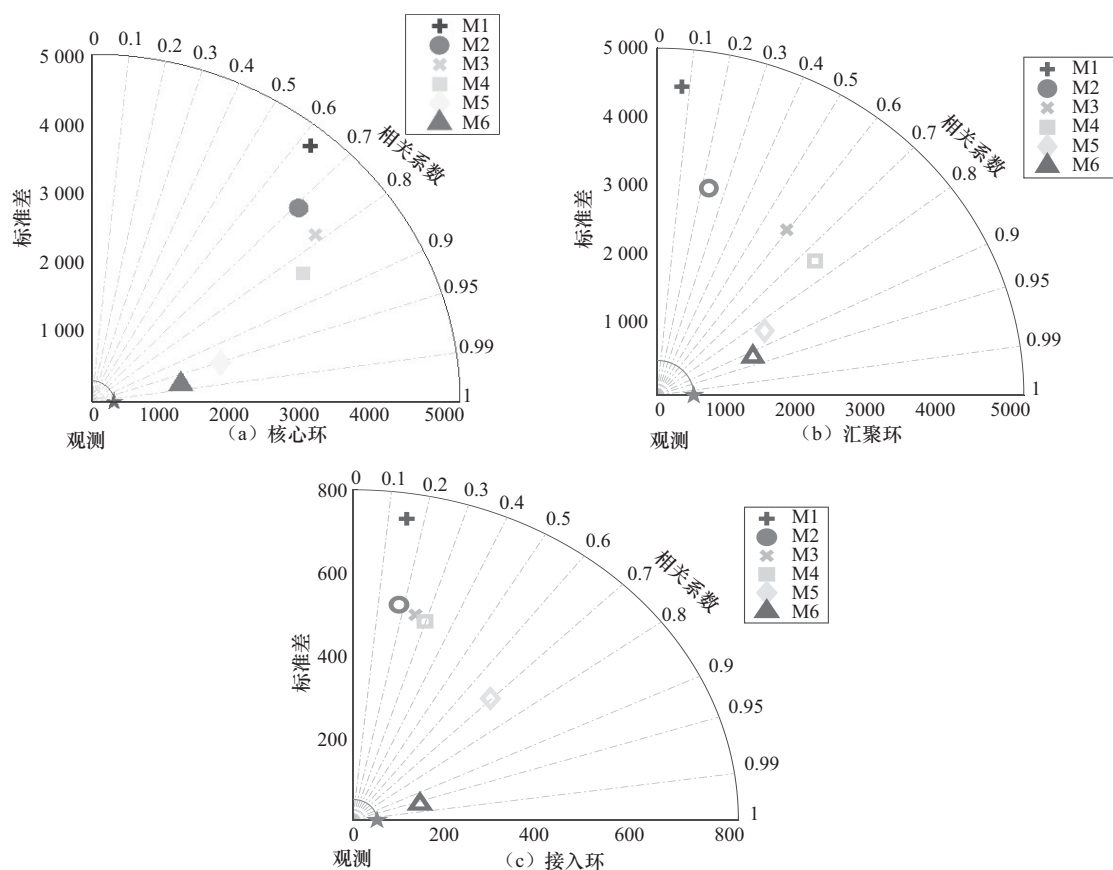


图17 5G传输网络流量预测技术的泰特图

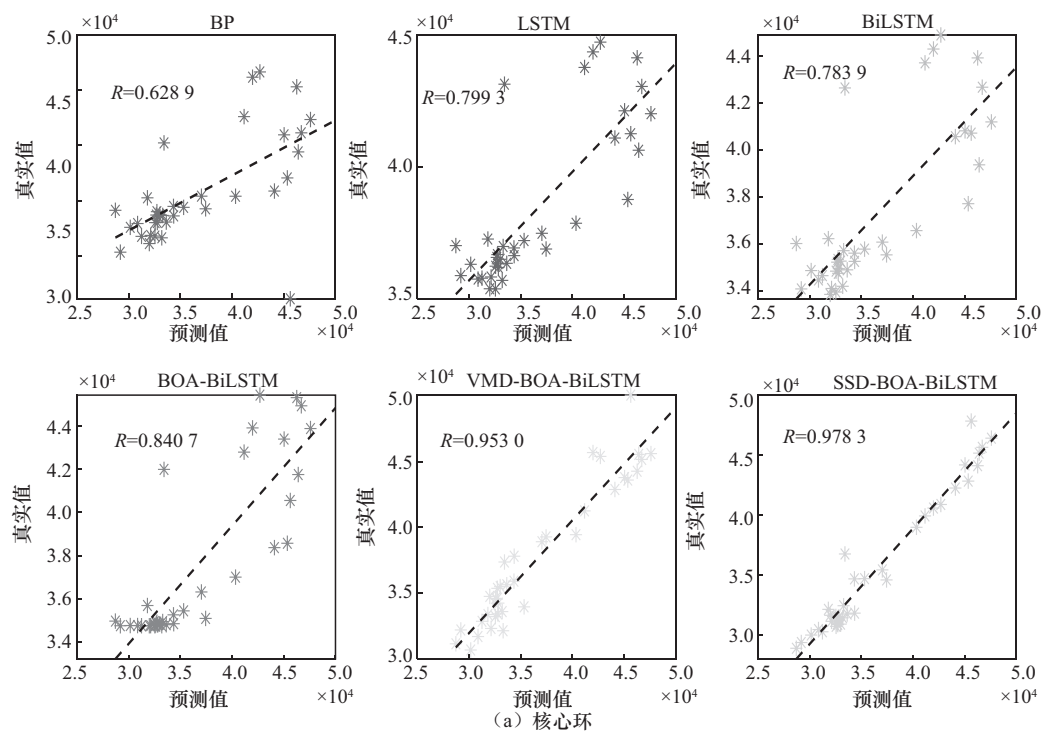
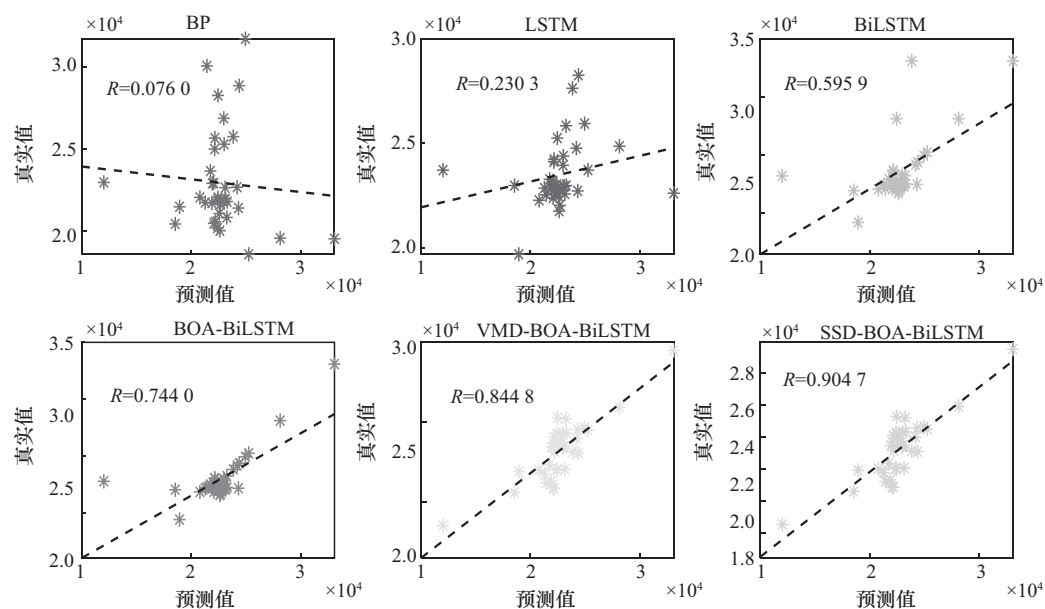
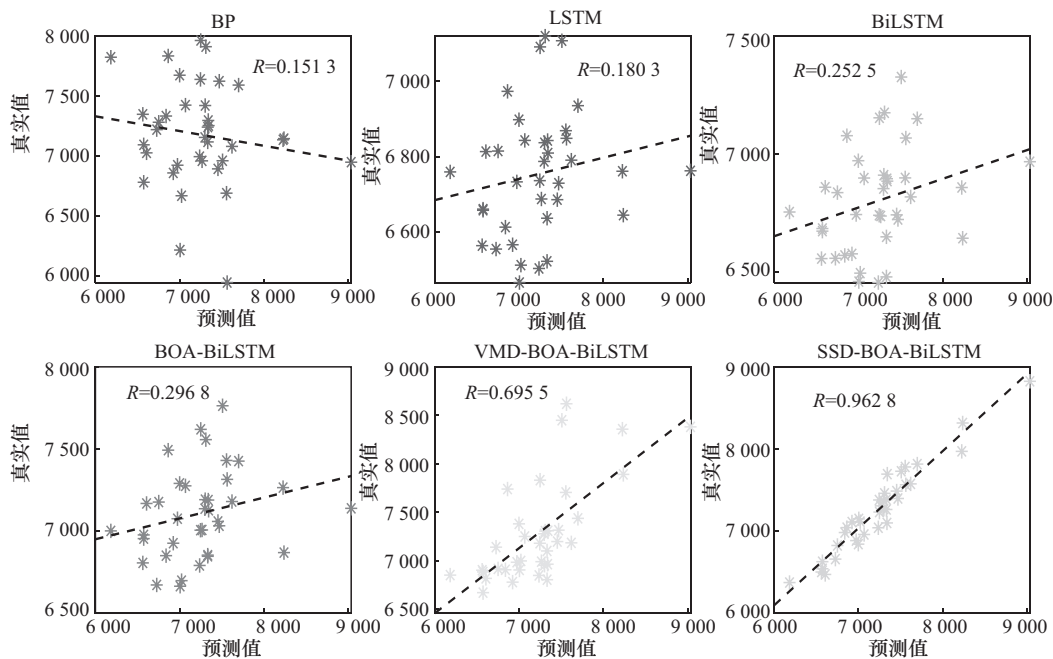


图18 各模型5G传输网络流量原始值与预测值的散点图



(b) 汇聚环



(c) 接入环

图18 各模型5G传输网络流量原始值与预测值的散点图(续)

由表3可知,核心环流量中,M1~M4模型的 $P$ 值均小于0.01,表明模型M6与这4个模型存在显著性差异的概率至少为99%;M5模型的 $P$ 值小于0.05,说明M6与M5存在显著性差异的概率至少为95%。汇聚环流量方面,除M2、M3模型 $P$ 值小于0.05外,其余模型 $P$ 值均小于

0.01,意味着M6与M1、M4、M5存在显著性差异的概率至少为99%。接入环流量中,各模型 $P$ 值均小于0.01,表明M6与M1-M5均存在显著差异。上述结果通过对5种预测模型的合理性检验,再次印证了所提模型的普适性与实用性。

表3 5G传输网络流量的DM检验结果

| 不同地区 | 核心环      |         | 汇聚环      |         | 接入环     |         |
|------|----------|---------|----------|---------|---------|---------|
|      | DM统计量    | P值      | DM统计量    | P值      | DM统计量   | P值      |
| M1   | 4.929 7* | 0.000 0 | 6.403 3* | 0.000 0 | 3.378 3 | 0.000 4 |
| M2   | 3.992 9* | 0.000 0 | 1.829 8* | 0.033 6 | 3.009 8 | 0.001 3 |
| M3   | 3.846 7* | 0.000 1 | 1.725 4* | 0.042 2 | 3.039 3 | 0.001 2 |
| M4   | 2.590 9* | 0.004 8 | 2.476 0* | 0.006 6 | 2.479 0 | 0.006 6 |
| M5   | 1.916 2* | 0.027 7 | 1.698 8* | 0.004 7 | 3.331 0 | 0.000 4 |

注:\*表示1%置信水平。

## 4 结束语

在“双碳”目标纵深推进与5G网络规模化部署的双重背景下,传输网络流量的精准预测成为优化网络架构、提升资源利用率、降低碳排放的关键支撑。本文构建的SSD-BOA-BiLSTM模型,通过SSD技术将非线性、非平稳的5G流量序列拆解为特征更明晰的子序列,借助BOA对BiLSTM网络的关键参数进行全局寻优,有效克服了单一深度学习模型在复杂序列预测中的局限性。实验结果显示,该模型在核心环、汇聚环、接入环3类场景中, RMSE、MAE、MAPE等误差指标均显著低于BP、LSTM等对比模型,且经DM检验证实其性能优势具有统计显著性,为网络负载均衡与动态资源调度提供了高精度的决策依据。尽管本技术已经显示出良好的性能,但仍可以从以下几个方向进行改进。

(1) 对双向长短期记忆网络预测模型结构的改进。尝试不同的网络架构和参数设置,以寻找更优的模型配置。

(2) 文中仅考虑单一因素5G传输网络流量的预测性能,但5G传输网络流量在实际运行中受网络结构、用户行为、工作日与节假日等因素影响。未来可以考虑多因素对5G传输网络流量的影响,进一步提升5G传输网络流量的稳定性和普适性。

## 参考文献:

- [1] 刘琳. 人工智能技术在5G时代的运用探讨[J]. 科技资讯, 2021, 19(31): 25-26, 32.
- [2] 孙晓楠. 网络流量测试与监控研究[D]. 北京: 北京邮电大学, 2012.
- [3] 雷霆, 余镇危. 一种网络流量预测的小波神经网络模型[J]. 计算机应用, 2006, 26(3): 526-528.
- [4] 李振刚. 基于高斯过程回归的网络流量预测模型[J]. 计算机应用, 2014, 34(5): 1251-1254.
- [5] 鲁晓帆. 基于ARMA模型的校园网网络流量预测模型的分析与研究[J]. 网络安全技术与应用, 2014(8): 98-99.
- [6] 王静超. 时间序列分析技术在网络流量监控中的应用研究[D]. 天津: 天津理工大学, 2014.
- [7] Cheng F C, Zhang H, Fan W J, et al. Image recognition technology based on deep learning[J]. Wireless Personal Communications, 2018, 102(2): 1917-1933.
- [8] Young T, Hazarika D, Poria S, et al. Recent trends in deep learning based natural language processing [review article][J]. IEEE Computational Intelligence Magazine, 2018, 13(3): 55-75.
- [9] Song Z J. English speech recognition based on deep learning with multiple features[J]. Computing, 2020, 102(3): 663-682.
- [10] 王俊松, 高志伟. 基于RBF神经网络的网络流量建模及预测[J]. 计算机工程与应用, 2008, 44(13): 6-7, 11.
- [11] 王亦然, 郭晓东. 基于人工智能的数据中心机房网络流量预测[J]. 信息系统工程, 2020(11): 110-111, 113.
- [12] Wang Q M, Fan A W, Shi H S. Network traffic prediction based on improved support vector machine[J]. International Journal of System Assurance Engineering and Management, 2017, 8(3): 1976-1980.
- [13] 朱业胜. 基于经验模态分解的网络流量预测方法的研究与实现[D]. 南京: 东南大学, 2020.
- [14] 赵玉婷, 努尔布力, 吾守尔·斯拉木. 基于时序EEMD网络流



- 量预测方法[J]. 计算机仿真, 2018, 35(11): 466-469.
- [15] 董雪, 赵生校, 陆艳艳, 等. 基于奇异谱分解和双重注意力机制的超短期光伏功率预测方法[J]. 激光与光电子学进展, 2023, 60(5): 0525001.
- [16] Liu X, Gherbi A, Li W B, et al. Multi features and multi-time steps LSTM based methodology for bike sharing availability prediction[J]. Procedia Computer Science, 2019, 155: 394-401.
- [17] 吴雨芯, 蔡婷, 张大斌. 基于层级注意力机制与双向长短期记忆神经网络的智能合约自动分类模型[J]. 计算机应用, 2020, 40(4): 978-984.
- [18] 曹成远, 吕强. 使用双向LSTM的深度神经网络预测蛋白质残基相互作用[J]. 2017(3): 531-535.
- [19] 刘小雄, 武燕, 史静平, 等. 一种故障诊断的贝叶斯优化算法研究[J]. 计算机应用研究, 2009, 26(1): 129-131.
- [20] Bonizzi P, Karel J M H, Meste O, et al. Singular spectrum decomposition: a new method for time series decomposition[J]. Advances in Adaptive Data Analysis, 2014, 6(4): 1450011.
- [21] Elsner J B. Analysis of time series structure: SSA and related techniques[J]. Journal of the American Statistical Association, 2002, 97(460): 1207-1208.
- [22] Yang H, Zhang A H, Li G H. A new singular spectrum decomposition method based on Cao algorithm and amplitude aware permutation entropy[J]. IEEE Access, 2021, 9: 44534-44557.
- [23] Xiao C A, Tang H S, Ren Y, et al. Fuzzy entropy assisted singular spectrum decomposition to detect bearing faults in axial piston pump[J]. Alexandria Engineering Journal, 2022, 61(8): 5869-5885.
- [24] Xu S L, Jiang X M, Huang J Z, et al. Bayesian wavelet PCA methodology for turbomachinery damage diagnosis under uncertainty[J]. Mechanical Systems and Signal Processing, 2016, 80: 1-18.
- [25] Chen J X, Cheng L S, Yu H, et al. Rolling bearing fault diagnosis and health assessment using EEMD and the adjustment Mahalanobis - Taguchi system[J]. International Journal of Systems Science, 2018, 49(1): 147-159.
- [26] Wang J J, Gao R X, Yan R Q. Integration of EEMD and ICA for wind turbine gearbox diagnosis[J]. Wind Energy, 2014, 17(5): 757-773.

#### [作者简介]

**郑才凤** (1996-), 女, 中国移动通信集团设计院有限公司湖北分公司工程师, 主要研究方向为有线传送网、人工智能。

**魏涛** (1972-), 男, 中国移动通信集团湖北有限公司高级工程师, 主要研究方向为有线网络规划设计。

**卢鸣** (1976-), 男, 中国移动通信集团湖北有限公司工程项目经理, 主要研究方向为有线网络规划设计。

**陈锋** (1983-), 男, 中国移动通信集团设计院有限公司湖北分公司高级工程师, 主要研究方向为有线传送网、有线通信设备工程。

**沈力** (1987-), 男, 中国移动通信集团设计院有限公司湖北分公司工程师, 主要研究方向为有线传送网、哑资源管理。

**蔡恬恬** (1987-), 女, 中国移动通信集团设计院有限公司湖北分公司高级工程师, 主要研究方向为有线传送网、有线通信设备工程。

**夏诚** (1992-), 男, 中国移动通信集团设计院有限公司湖北分公司工程师, 主要研究方向为有线传送网、有线通信设备工程。

**刘琛** (1983-), 女, 中国移动通信集团设计院有限公司湖北分公司工程师, 主要研究方向为有线传送网、有线通信设备工程。

**柳鑫成** (1987-), 男, 中国移动通信集团设计院有限公司湖北分公司高级工程师, 主要研究方向为有线传送网、传输线路管道工程。