

DOI: 10.20189/j.cnki.CN/61-1527/E.202505010

融合 DS-Attention 的 Informer 频谱 预测改进方法

刘丁胤¹, 徐东辉¹, 方喜龙², 张 望¹

(1. 火箭军工程大学, 西安 710025;

2. 96751 部队, 梅州 514700)

摘 要: 为解决现有频谱预测模型在采集于真实环境中的具有非平稳特征的数据集上的预测性能较差的问题, 提出一种融合 DS-Attention 的 Informer 频谱预测改进方法。将 DS-Attention 可以防止过平滑化优点与 Informer 自注意力蒸馏机制能够压缩网络参数数量的优点相结合, 在真实环境采集的频谱数据集上进行广泛实验, 与 5 种典型模型进行预测性能对比。结果表明: 所提出的方法在 MAE、RMSE、 R^2 及 MAPE 等性能指标方面均取得最好结果。相较于传统神经网络模型 LSTM 与 TCN, 在 RMSE 上分别提升了 31.0% 与 56.0%; 相较于 Transformer 变体模型 Autoformer 与 Informer, 分别提升了 18.4% 与 9.0%, 验证了所提方法的有效性。

关键词: 认知无线电; 频谱预测; 非平稳时间序列; 深度学习;
Informer; DS-Attention

中图分类号: TN919.23

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2025)05-0107-10



听语音
与作者
聊科研

Enhanced Informer-Based Spectrum Prediction Method Incorporating DS-Attention

LIU Dingyin¹, XU Donghui¹, FANG Xilong², ZHANG Wang¹

(1. Rocket Force University of Engineering, Xi'an 710025, China;

2. Troops No.96751, Meizhou 514700, China)

ABSTRACT: An enhanced Informer-based spectrum prediction method incorporating De-Stational-Attention (DS-Attention) was proposed to address the performance limitations of existing spectrum prediction models when applied to real-world non-stationary datasets. The proposed approach synergistically combined the anti-over-stational capabilities of DS-Attention with the parameter-efficient characteristics of Informer's self-attention distillation mechanism. Extensive experiments performed on authentic spectrum measurement datasets demonstrate that the method achieves superior performance across multiple evaluation metrics, including MAE, RMSE, R^2 ,

收稿日期: 2025-03-18

修回日期: 2025-06-03

录用日期: 2025-07-11

基金项目: 军事类研究生资助课题 (JYKT910342023006)

第一作者: 刘丁胤 (1998—), 男, 硕士研究生, 主要从事认知无线电研究。E-mail: 13985436018@163.com

引用格式: 刘丁胤, 徐东辉, 方喜龙, 等. 融合 DS-Attention 的 Informer 频谱预测改进方法[J]. 火箭军工程大学学报, 2025, 39 (5): 107-116. LIU Dingyin, XU Donghui, FANG Xilong, et al. Enhanced informer-based spectrum prediction method incorporating DS-attention[J]. Journal of Rocket Force University of Engineering, 2025, 39 (5): 107-116.

and MAPE, outperforming five baseline models. Specifically, compared with conventional neural network models such as LSTM and TCN, the proposed method achieves RMSE improvements of 31.0% and 56.0%, respectively. When benchmarked against Transformer variants Autoformer and Informer, it achieves RMSE improvements of 18.4% and 9.0%, respectively. These results substantiate the effectiveness of the proposed approach.

KEYWORDS: cognitive radio; spectrum prediction; non-stationary time series; deep learning; Informer; DS-Attention

认知无线电能够有效提高电磁频谱利用率,通过频谱感知、频谱决策、频谱共享及频谱切换四大关键技术共同作用而实现认知用户 (secondary user, SU) 对频谱资源的充分利用^[1-2]。频谱预测作为一种辅助技术,在认知无线网络方面具有广泛的应用。例如,辅助频谱感知^[3]应用于动态频谱接入技术^[4]以及辅助频谱共享过程^[5]等。借助频谱预测可以实现主动管理频谱,并实现降低延迟、提高能效及吞吐量的目的^[6]。传统频谱预测方法通过模型驱动,在实际场景下预测性能受限。

数据驱动的方法主要包含机器学习方法及深度学习^[7]。其中,深度学习通过非线性建模能力可有效提取频谱数据中的复杂依赖关系。长短期记忆 (long short-term memory, LSTM) 网络在频谱预测中被广泛应用。早期研究通过基础 LSTM 实现信道状态预测^[8-9],后续改进结合田口法优化参数提升信号电平预测精度^[10],而卷积长短期记忆 (convolutional long short-term memory, ConvLSTM) 网络通过卷积结构进一步捕获时空多维特征^[11]。双向卷积长短期记忆 (bidirectional convLSTM, bi-ConvLSTM) 网络与序列到序列 (sequence to sequence, seq2seq) 的融合架构突破了单维度预测限制^[12],但上述 LSTM 衍生模型仍存在长期依赖捕捉能力受限的固有缺陷。

为应对长期频谱预测挑战,基于注意力机制的 Transformer 模型展现出了独特优势,但其平方级计算复杂度制约了实际应用。针对此问题,研究者提出了多维度改进方案。Reformer 方法通过哈希注意力降低计算开销^[13]; Autoformer 方法创新性地采用序列分解策略提升趋势周期建模能力^[14]; Informer 方法则通过注意力蒸馏实现参数压缩^[15]。尽管改进模型增强了全局建模能力,但在实际采集的非平稳频谱数据上仍存在性能瓶颈。值得注意的是,现实场景的频谱时序普遍具有非平稳特性,近期提出的 Non-stationary Transformers 通过平稳化处理与去平稳注意力的协同机制^[16],

能够在保持时序依赖的同时,缓解过度平稳化问题,为复杂环境下的频谱预测提供了新思路。

本文针对 Transformer 及其变体模型在预测过程中训练数据需求量大、预测输出效率低、在非平稳预测任务中性能降低等问题,提出了一种融合 DS-Attention 的 Informer 频谱预测改进方法。首先,归纳分析了当前频谱预测的相关方法;其次,描述了基于异构认知无线网络下的系统模型,并对现实世界所采集的频谱数据集进行了分析;再次,对本文所提出的融合 DS-Attention 的 Informer 网络模型的设计过程及原理进行解释;最后,基于现实通信环境中采集的数据集进行了仿真验证。结果表明,所提模型相比基线模型在各性能指标上具有较大提升。

1 系统模型与频谱数据分析

1.1 系统模型

本文所定义的长期频谱预测问题,是基于在不同位置及宽频段内采集的历史频谱数据,对未来多个时隙、多个频带的频谱占用情况进行预测。如图 1 所示,对于一个具有 I 个主用户及 J 个次用户所构成的异构认知无线网络,频谱传感器将获取的宽频带历史感知数据上传至频谱数据库并生成频谱图,频谱数据库将该数据用于长期频谱预测模块中进行离线训练,当频谱接入任务来临时,对未来频谱使用情况进行预测,以指导下一步的频谱接入。假设 K 个频谱传感器持续监测 F 个频段的使用情况,并将接收到的信号功率传输至该地区的频谱数据库。考虑到实际情况中难以对每个位置的频谱进行精确测量,因此,将感兴趣的位置区域划分为网格状,通过 (x, y) 网格创建插值频谱图,其中 $x \in \{1, 2, \dots, X\}$, $y \in \{1, 2, \dots, Y\}$ 。本文模型中频谱传感器呈稀疏分布,因此,选择逆距离加权的方法融合各频谱传感器的接收信号功率。对于给定空间位置 (x, y) 在 f 频段 t 时

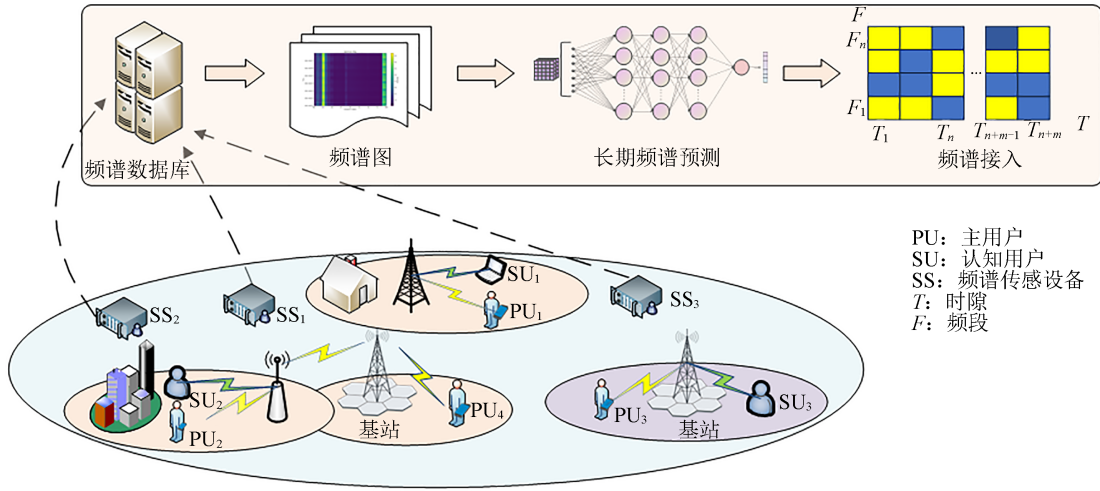


图1 系统模型

Fig.1 System model

隙内的插值频谱测量值 $p_{(f,t)}^{(x,y)}$ 可表示为^[17-18]

$$p_{(f,t)}^{(x,y)} = \sum_{k=1}^K \omega_k^{(x,y)} p_{(f,t)}^k \quad (1)$$

式中: $p_{(f,t)}^k$ 表示第 k 个频谱传感器在 f 频段下, t 时隙内的接收信号功率; $\omega_k^{(x,y)}$ 为第 k 个频谱传感器的权重系数, 其表达式为

$$\omega_k^{(x,y)} = \frac{(d_k^{(x,y)})^{-\alpha}}{\sum_{n=1}^K (d_n^{(x,y)})^{-\alpha}} \quad (2)$$

式中: $d_k^{(x,y)}$ 表示网格 (x, y) 到第 k 个频谱传感器的欧式距离; α 为路径损耗指数 (取值介于 2~6 之间)。因此, 插值后的频谱测量值可以表示为一个四阶张量:

$$\mathbf{P} \in \mathbb{R}^{X \times Y \times F \times T} \quad (3)$$

网格 (x, y) 以 L 表示, 即 L 表示以正方形单位面积定义的网格空间中某一特定位置 (x, y) 。因此, 四阶张量可以表示为

$$\mathbf{P} \in \mathbb{R}^{L \times F \times T} \quad (4)$$

式中: L 、 F 、 T 分别代表空间、频率、时间 3 个指标。而频谱预测本质上是时间序列预测问题, 即根据时间顺序所获得的历史数据 P_t , 预测 $P_{t+\Delta}$, 因此, 所定义的基于现实世界数据的频谱预测问题可表示为

$$[P_{t-n}, P_{t-n+1}, \dots, P_t] \rightarrow [P_{t+1}, P_{t+2}, \dots, P_{t+m}] \quad (5)$$

式中: n 、 m 分别表示历史观测值序号及预测步数。

1.2 频谱数据分析

由于模型生成的数据可能遵从某些统计规律而与实际情况不符, 为更加接近实际通信环境, 采用实际测量数据集。在 Electrosense^[19] 开放 API 上获取 600~800 MHz 频段的聚合频谱测量数据。

该数据集来源于位于西班牙马德里市中心 6 km 范围内 2 个传感器的实地测量数据。从 2021 年 6 月 1 日至 2021 年 6 月 8 日, 以 1 min 的时间间隔及 2 MHz 的频率间隔对 600~800 MHz 频段进行不间断测量, 得到的接收信号功率。

为明确频谱数据是否可以用于长期频谱预测工作, 需要对数据进行相关性分析。采用 Pearson 相关系数分别对频谱数据的时间、频率相关性进行分析。对于 2 个随机变量 x 、 y , 其 Pearson 相关系数表达式为

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

式中: $\bar{x}$ 、 $\bar{y}$ 分别为随机变量的均值; r_{xy} 取值范围为 $[-1, 1]$, 取值越接近 1 (或 -1), 则表明随机变量正 (负) 相关性越明显。该数据集分为室外采集和室内采集 2 个部分, 对数据 600~800 MHz 频段内频谱使用情况 & 频率相关性进行分析, 对数据持续 1 000 min 的时间相关性进行分析, 其时间-频谱及时间、频率相关性可视化图像如图 2 所示。

由图 2 (a)、(d) 可知, 室内外频谱使用情况主要集中于以 625 MHz 及 800 MHz 为中心频率的频段范围内。由图 2 (b)、(e) 可知, 在时间维度上, 无论是短时间内还是在较长时间范围内, 二者均存在极强的相关性, 其值接近于极值 1, 呈现明显正相关。由图 2 (c)、(f) 可知, 频率相关性呈现点状分布, 虽然在频率维度上的相关性略差于时间维度, 但存在某些频率点如 625 MHz、650 MHz、690 MHz、750 MHz 等处呈现明显相关

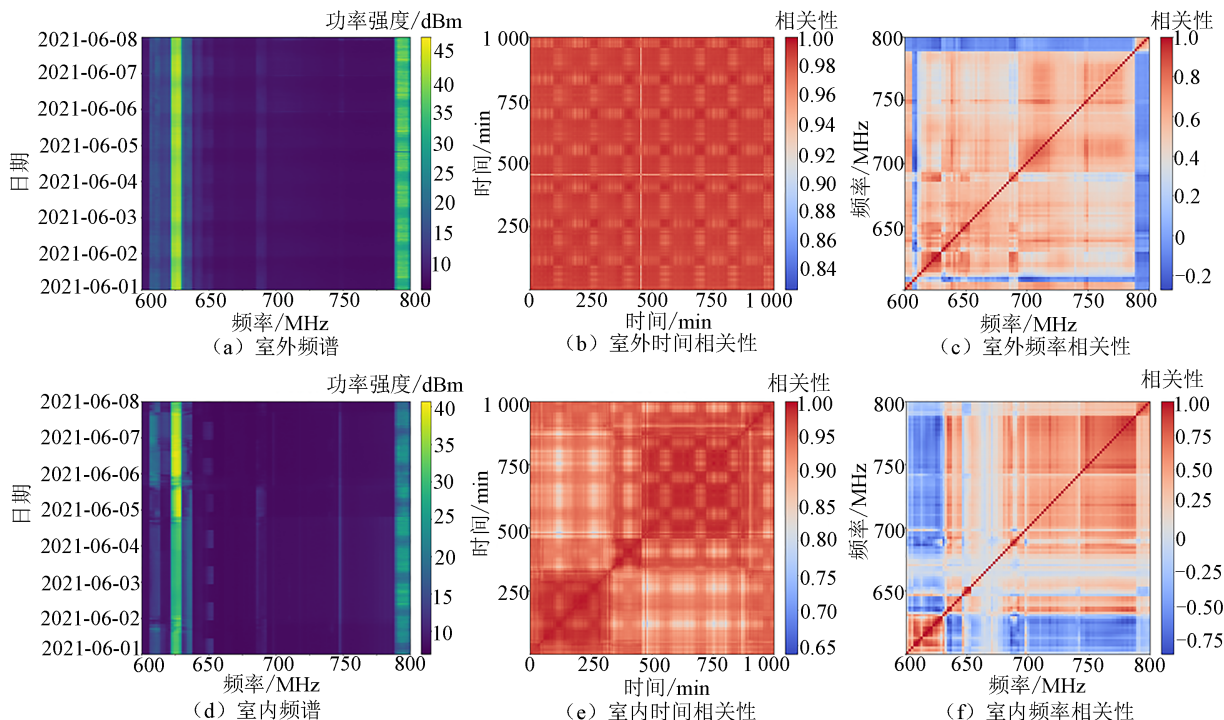


图 2 室内外频谱数据频谱图及相关性分析

Fig. 2 Indoor and outdoor spectrum data spectrum diagrams and correlation analysis

性。由此可知，该数据存在明显的长时间依赖性与频率相关性，即过去的频谱使用情况对未来的结果将会产生一定程度上的影响，可以用于长期频谱预测。

2 网络模型

2.1 Informer 模型压缩网络参数原理

Transformer 架构因其强大的长距离依赖建模能力，在时序预测任务中展现出显著优势。然而，

标准 Transformer 模型的高计算复杂度与资源需求制约了其在资源受限场景下的实际应用。针对上述缺陷，Informer 模型创新性地采用稀疏注意力机制与自注意力蒸馏技术，有效降低了计算复杂度与参数量，通过单步输出预测序列的架构设计，减少误差传递的同时提升推理效率^[15, 20-21]，其模型结构如图 3 所示。图中， $[P_{t-n}, P_{t-n+1}, \dots, P_t]$ 为输入的历史频谱观测数据。Informer 通过自注意力蒸馏机制降低数据维度和网络参数量。在 Informer 的编码器架构中，具有 2 个可以压缩数据维度的

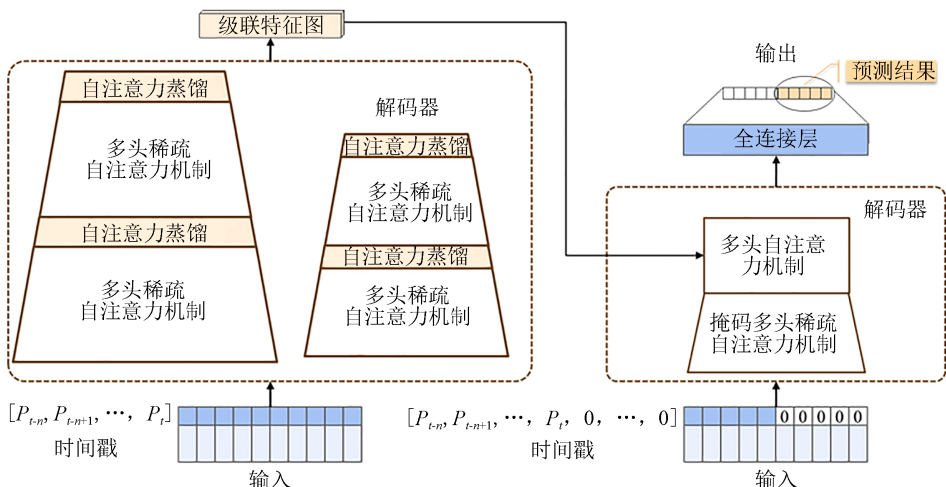


图 3 Informer 模型结构

Fig. 3 Structure of Informer model

编码结构, 其具体结构如图 4 所示。

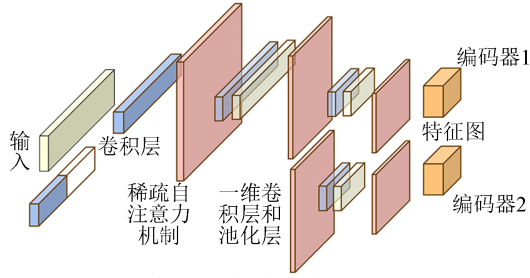


图 4 编码器自注意力蒸馏机制

Fig. 4 Self-attention distillation mechanism of the encoder

Informer 采用自注意力蒸馏对重点特征进行提取, 编码器 1 在稀疏注意力机制之后分别添加一层一维卷积层和池化层, 对特征进行下采样, 使得其输出相较于输入在长、宽维度上折半, 对输入序列进行压缩, 有效对稀疏注意力机制进行降维。编码器 2 具有和编码器 1 相同的结构, 但其输入数据为时间序列后半部分, 即只接收输入序列的一半数据, 并且去掉一层自注意力蒸馏层, 目的是在最终输出时与编码器 1 的输出维度对齐, 以便对二者输出的特征图进行拼接并传输至解码器当中。二者结合不仅能够降低数据维度, 还可以减少网络参数量。

最后, Informer 通过特定的生成式解码器一次性输出预测结果, 克服了 Transformer 的动态解码预测过程所带来的低效输出问题。这种生成式结构专为长时间序列预测任务设计, 也为多步频谱预测奠定基础, 避免了 Transformer 动态解码一步预测所带来的误差累积。

2.2 DS-Attention 原理

该实验中的混凝土箱型梁桥在爆炸荷载作用下, 仍面临传统 Transformer 在非平稳真实数据中的性能退化问题, 现有改进方案多采用序列平稳化处理以降低预测难度, 但该策略与真实场景中普遍存在的非平稳特性相矛盾, 难以满足频谱预测等需适应动态变化环境的实际需求。为此, 文献[16]提出, 通过标准化操作增强序列平稳性以改善可预测性^[22], 继而设计去平稳注意力模块以重构原始时序依赖关系。该协同机制在抑制过度平稳化的同时, 能够有效保留非平稳特征的关键时间依赖性, 显著提升复杂场景下的预测鲁棒性。

标准化模块用以处理由序列均值和标准差引起的非平稳特性, 假设输入序列为

$$\mathbf{X} = [x_1, x_2, \dots, x_S]^T \in \mathbb{R}^{S \times C} \quad (7)$$

式中: S 、 C 分别为输入序列长度和变量数。其标准化过程为

$$x'_i = \frac{1}{\sigma_x} \odot (x_i - \mu_x) \quad (8)$$

$$\mu_x = \frac{1}{S} \sum_{i=1}^S x_i \quad (9)$$

$$\sigma_x = \sqrt{\frac{1}{S} \sum_{i=1}^S (x_i - \mu_x)^2} \quad (10)$$

式中: x'_i 为归一化之后的序列; μ_x 、 σ_x 为序列的均值及标准差; $\odot$ 表示逐元素积。

平稳化后的注意力计算为

$$Q' = \frac{Q - 1\mu_Q^T}{\sigma_x}, K' = \frac{K - 1\mu_K^T}{\sigma_x} \quad (11)$$

$$Q'K'^T = \frac{1}{\sigma_x^2} (QK^T - 1(\mu_Q^T K^T) - (Q\mu_K)1^T + 1(\mu_Q^T \mu_K^T)1^T) \quad (12)$$

$$\text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) = \text{softmax} \left(\frac{\sigma_x^2 Q'K'^T + 1(\mu_Q^T K^T) + (Q\mu_K)1^T - 1(\mu_Q^T \mu_K^T)1^T}{\sqrt{d_k}} \right) \quad (13)$$

式中: μ_Q 、 $\mu_K \in \mathbb{R}^{d_k \times 1}$, $1 \in \mathbb{R}^{S \times 1}$ 分别为注意力机制查询向量 Q 和键向量 K 的均值; σ_x 为原始序列的标准差; $(Q\mu_K)1^T, 1(\mu_Q^T K^T)$ 表示对矩阵 $\sigma_x^2 Q'K'^T$ 每一行添加偏移。根据 softmax 函数的行平移不变原理, 对输入矩阵的同一行添加任意常数, 输出不变, 因此, 平稳化后的注意力计算可简化为

$$\text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) = \text{softmax} \left(\frac{\sigma_x^2 Q'K'^T + 1(\mu_Q^T K^T)}{\sqrt{d_k}} \right) \quad (14)$$

为了恢复数据的非平稳特征以防止过平稳化, DS-Attention 引入去平稳因子。令 $\sigma_x^2 = \tau$, $K\mu_Q = \Delta$, 将去平稳因子代入 softmax 计算中得到去平稳注意力, 可表示为

$$\text{Attn}(Q', K', V, \tau, \Delta) = \text{softmax} \left(\frac{\tau Q'K'^T + 1\Delta^T}{\sqrt{d_k}} \right) V \quad (15)$$

式中: V 为注意力机制的值向量; τ 、 μ 均通过多层感知机 (multilayer perceptron, MLP) 从未平稳化的原始序列及其均值与方差等统计信息当中学习得到, 该方式使得去平稳因子可以根据输入序列的不同特性进行自适应调整, 更好地捕捉数据当中的非平稳特性, 去平稳因子可表示为

$$\log \tau = \text{MLP}(\sigma_x, x) \quad (16)$$

$$\Delta = \text{MLP}(\mu_s, x) \quad (17)$$

DS-Attention 能够在保证预测性能的同时, 保持原始序列固有的非平稳特性, 将其恢复为时间依赖性, 以适用于现实场景下的频谱预测任务。

3 融合 DS-Attention 的 Informer 频谱预测算法

频谱预测是基于真实场景的时间序列任务, 用户接入频段进行通信带有随机性, 因此, 频谱数据必然具备非平稳特性。尽管 Informer 模型在长时间序列中已经表现出较好的预测性能, 但其并未考虑到时间序列非平稳特征对于长时间序列预测任务的影响。因此, 其在真实环境的频谱预测仍有待改进。为了实现真实场景下频谱数据预测性能的提升, 需要针对序列的非平稳特性进行处理。Passalis 等^[23]提出一种用于时间序列预测的深度自适应输入归一化层, 通过学习数据分布动态调整归一化方案, 解决了深度学习模型在处理非平稳时间序列预测时性能退化的问题; Yang 等^[24]通过因果卷积增强对复杂时间序列特征提取能力, 并结合门控循环单元优秀的提取时间特征的能力对数据中的长短期依赖关系进行再学习^[25], 有效弥补了 Informer 模型在处理非平稳和非线性时间序列数据时的不足; Yang 等^[26]提出一种基于平滑滤波注意力机制和差分注意力模块的深度学习模型 Diviner, 通过对非平稳特性进行建模, 捕获其中的稳定规律, 实现符合历史规律的非平稳

预测。

本文提出一种融合 DS-Attention 的 Informer 频谱预测改进模型, 由 DS-Attention 替换 Informer 中的稀疏注意力机制, 在避免过度提高计算复杂度及模型维度的同时, 保留频谱数据的非平稳特征, 防止过平稳化, 实现预测性能的提升, 改进模型的预测流程如图 5 所示。

频谱数据先经历嵌入组合之后再进行平稳化处理, 其中输入序列 $P = [P_1, P_2, \dots, P_T]^H \in \mathbb{R}^{T \times F}$, T 、 F 分别表示输入序列长度和变量数, 即时间步长与频率带宽。平稳化之后需要对输入序列进行嵌入组合, 由标量投射、局部时间戳嵌入组合而成。输入序列的标量值通过一维卷积转换成为特征维度向量 $[\mu_1, \mu_2, \dots, \mu_i]$, 表示为 $\{\alpha\mu_i^i\}$, 其中 α 为平衡标量投射与局部及全局嵌入之间大小的因子, 由于经历了标准化, 故取值为 1。局部嵌入用以表示输入序列各个位置的信息, 每个位置对应嵌入一个向量, 位置嵌入向量表示为 $PE(i) = [PE(0), PE(1), \dots, PE(L)]$ 。全局嵌入用以获取数据当中的全局时间信息, 全局嵌入是所有时间戳嵌入向量的综合, 表示为 $\sum SE$ 。经过平稳化和嵌入组合后的输入可表示为

$$P_t^{\text{feed}}[i] = \alpha\mu_i^i + PE(i) + \sum SE \quad (18)$$

将嵌入组合之后的频谱数据进行平稳化处理, 使其具备良好的预测性, 标准化过程可表示为

$$P_t^{\text{feed}'} = \frac{1}{\sigma_p} \odot (P_t^{\text{feed}} - \mu_p) \quad (19)$$

$$\mu_p = \frac{1}{T} \sum_{t=1}^T P_t^{\text{feed}} \quad (20)$$

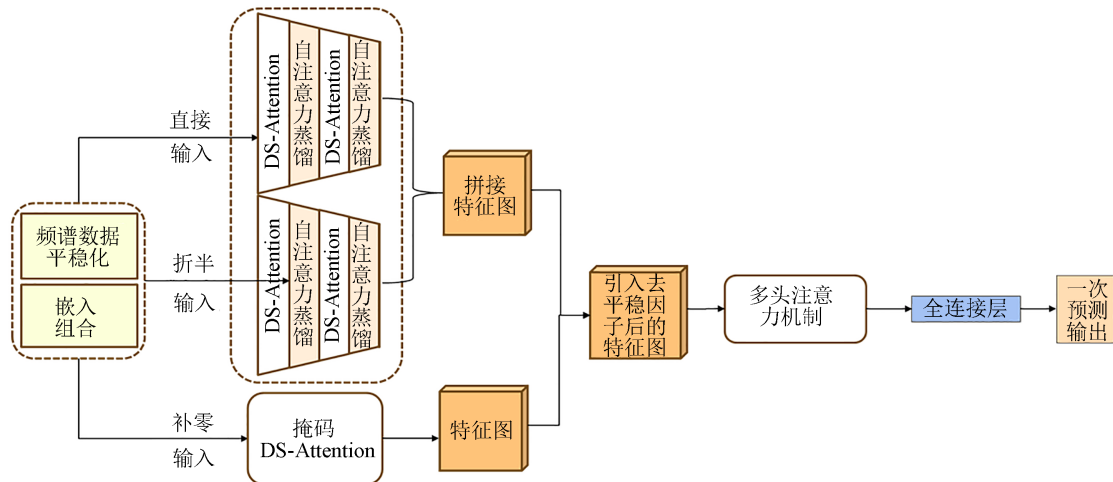


图 5 融合 DS-Attention 的 Informer 模型预测流程

Fig. 5 Prediction process of the Informer model incorporating DS-Attention

$$\sigma_p = \sqrt{\frac{1}{T} \sum_{t=1}^T (P_t^{\text{feed}} - \mu_p)^2} \quad (21)$$

频谱数据经历嵌入及平稳化后, 分别通过直接输入、折半输入以及补零输入的方式进入编码器 1、编码器 2 和解码器当中, 融合 DS-Attention 的 Informer 频谱预测模型以去平稳化注意力代替稀疏注意力。去平稳注意力表达式为

$$\text{Attn}(Q', K', V, \tau, \Delta) = \text{softmax}\left(\frac{\tau_p Q' K'^T + 1 \Delta_p^T}{\sqrt{d_k}}\right) V \quad (22)$$

$$Q' = Q - 1\mu_Q^T / \sigma_p, K' = K - 1\mu_K^T / \sigma_p \quad (23)$$

式中: τ_p 、 Δ_p 为去平稳因子, 可通过 MLP 从输入的频谱数据及其统计特性中学习得到, 当频谱数据非平稳特性随环境动态变化时, 去平稳因子根据不同的输入频谱数据自适应调整。

融合 DS-Attention 的 Informer 频谱预测算法在输入数据处理上保留嵌入操作, 能够获取有利于时间序列预测的时间与位置信息, 在嵌入之后进行平稳化操作, 使得整体数据具备良好的可预测性。DS-Attention 引入的去平稳因子重新纳入非平稳特性, 网络模型从原始序列中学习数据本身具有的非平稳特征, 防止过平稳化。此外, 去平稳因子通过 MLP 学习得到, 计算复杂度较低。将 Informer 当中的稀疏自注意力机制替换为 DS-Attention, 保留了 Informer 的简单网络结构与编码器当中的自注意力蒸馏机制, 实现了网络模型的压缩, 降低网络参数量, 且去平稳因子也能克服 Informer 不能捕获频谱预测数据当中非平稳特性的缺陷。最后, 在输出部分保留 Informer 一次输出预测结构, 避免了 Transformer 动态解码输出预测结果过程的低效及误差累积。该模型通过 DS-Attention 与 Informer 网络结构的融合, 实现了压缩网络参数与生成式输出预测结果, 且在处理动态变化的非平稳频谱预测时更加高效灵活, 提升了在实际通信场景下采集频谱数据集上的预测性能。

4 实验与结果分析

4.1 实验设置

本实验在配备 Intel Core i7-13700H CPU、NVIDIA RTX 4060 GPU 和 16 GB 内存的计算机上进行, 使用 Python 3.8.0、PyTorch2.4.1 进行模型训练。

为了定量分析融合 DS-Attention 的 Informer 模型频谱预测性能, 采用了时间序列预测任务当中被广泛认可的性能评价指标, 包括均方根误差 RMSE、平均绝对误差 MAE 以及平均绝对百分比误差 MAPE, 同时, 增加决定系数 R^2 衡量模型的拟合优度, 其表达式分别为

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - \hat{p}_i)^2} \quad (24)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |p_i - \hat{p}_i| \quad (25)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{p_i - \hat{p}_i}{p_i} \right| \quad (26)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (p_i - \hat{p}_i)^2}{\sum_{i=1}^n (p_i - \bar{p})^2} \quad (27)$$

式中: p_i 、 $\hat{p}_i$ 、 $\bar{p}$ 分别为真实值、预测值以及均值。

4.2 结果分析

时间序列数据的非平稳特征指的是数据的统计特性, 如方差、均值、协方差等随时间变化而改变。真实通信环境下, 由于用户数量的变化、设备的开关机以及环境的改变等情况, 频谱监测数据信号强度将随时间改变, 进而导致其均值随时间变化。同时, 在一些特殊的时段或区域, 可能会出现大量的干扰信号, 导致频谱监测数据的方差随之波动。因此, 频谱数据通常具备非平稳特征。为进一步验证该模型在非平稳数据集上的有效性, 对频谱数据集的非平稳特征进行分析, 对 620~630 MHz 频段上的数据均值与方差随时间进行可视化, 将数据集按时间等分为 7 个时间段, 计算该频段内各频率频谱测量值均值与方差, 结果如图 6 所示。可以看出, 在 620~630 MHz 频段上, 该频谱监测数据集的均值与方差均随时间发生变化, 其中 622 MHz、628 MHz 等几个频点上, 数据方差随时间变化尤为明显, 表明该数据具有非平稳特性。

为验证改进模型的有效性, 选取自回归模型 (autoregressive integrated moving average, ARI-MA)、传统神经网络模型 LSTM、时序卷积 (temporal convolutional network, TCN) 及 Transformer 改进变体模型 Autoformer 与 Informer 等作为基线进行对比分析。根据对 Electrosense 数据集的分析可知, 在 620~630 MHz 频谱占用率最高, 使用情况更为复杂, 因此, 将该频段 10 MHz 的范

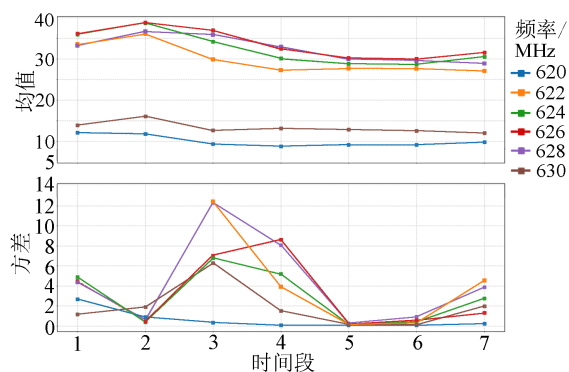
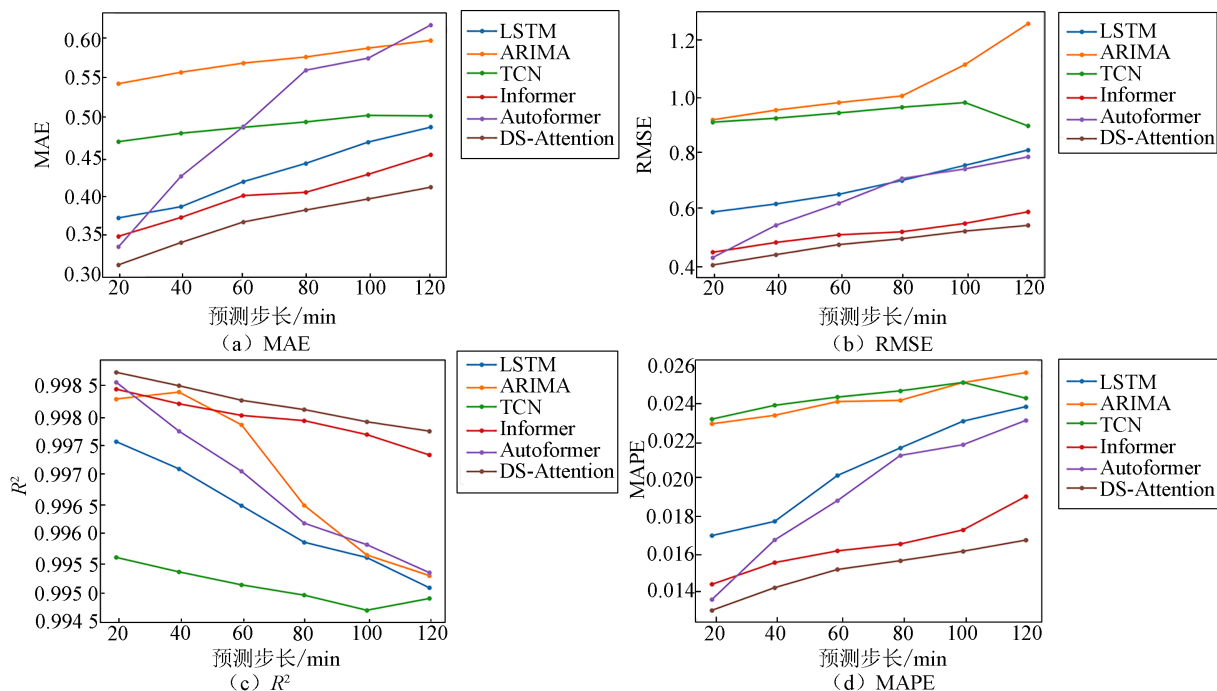


图 6 620~630 MHz 频段数据非平稳特征

Fig. 6 Non-stationary features of the data in the 620~630 MHz frequency band

围设置为预测带宽，预测步长设置为从短期预测 (20 min) 每间隔 20 min 逐步增加到长期预测 (120 min)，回顾时间窗口设置为 100 min，各模型仿真结果如图 7 所示。可以明显看出，所提模型在各项评价指标上的表现均为最优，并且在 120 min 时长的长期预测上也能保持良好的性能。其中，ARI-MA 本身是基于统计学原理进行预测，但是模型过于简单，难以处理数据当中的复杂非线性关系，在处理真实通信环境下的预测问题时性能较差，只适合短期频谱预测，在步长增加至 100 min 时，其 RMSE、MAE、MAPE 均有明显增加。而 TCN 是基于卷积网络提取时间特征，根据频谱

数据的时间依赖性来进行预测，但是真实环境采集的数据在频率上也有复杂的关联，本文实验针对 10 MHz 带宽上的多频点预测，TCN 的 R^2 相较于其他模型表现不佳，表明其拟合能力不足。LSTM 作为处理长时间序列预测的常用方法，其性能良好，但是在训练的过程中，由于其循环结构导致训练时间较长，且会随预测步长增加进一步增长，性能下降也较快，更适合处理中等长度的预测问题。Autoformer 能够有效捕捉长距离的相关性，适用于长期预测，但是其预测性能数据依赖性强，要想实现预测性能的提升，需要大量数据进行训练，在回顾窗口较小而预测步长较长时，预测性能较差，在预测步长为 100 min 时，其 MAE 达到 0.57。Informer 作为 Transformer 的变体结构，在降低计算复杂度上提升较大，但是其在真实环境采集的数据集上性能仍有待提升。所提模型在预测步长为 20 min 时 RMSE 为 0.4，而 LSTM、TCN 的 RMSE 分别为 0.58 与 0.91，所提模型相较于二者分别提升了 31.0% 与 56.0%。而 Transformer 变体模型 Autoformer、Informer 的 RMSE 为 0.49 与 0.44，所提模型分别提升了 18.4% 与 9.0%，预测性能均有明显提升。以 Transformer 为架构的 Autoformer 与 Informer 长期预测任务当中需要大量数据支撑，当预测步长为 120 min 时，

图 7 各模型 MAE、RMSE、 R^2 、MAPE 性能指标对比Fig. 7 Comparison of MAE, RMSE, R^2 , and MAPE performance indices of different models

型作为当前较为先进的时间序列预测模型, 具备优秀的性能, 但在克服序列非平稳特征上仍有不足, 其 MAE 分别为 0.617 4 与 0.451 5, 而所提模型为 0.41, 性能分别提升了 33.6% 与 9.2%。Informer 模型, 在预测步长 20~120 min 时, 其平均 MAE、RMSE、MAPE 分别为 0.399 9、0.517 9 与 0.016 1。而所提模型分别为 0.366 9、0.481 1 与 0.015, 相较于 Informer 整体分别提升 8.2%、7.1% 与 6.8%。

5 结论

针对多频段长期频谱预测在实际通信场景当中的预测性能提升问题进行研究, 提出了一种融合 DS-Attention 的 Informer 频谱预测改进方法, 通过引入去平稳因子, 提升模型在非平稳时间序列上的预测性能, 推进频谱预测算法在真实复杂的通信场景下的部署工作。通过仿真实验, 同 5 种典型模型进行对比, 结果表明, 所提出的模型各性能指标方面均取得最优性能, 主要结论如下。

1) 所提模型克服了以 Transformer 为基本架构模型需要大量训练数据支撑预测性能的劣势, 在长期预测任务中预测性能显著提升, 可以通过回顾较少的时间窗口获取优秀的预测性能。当预测步长为 120 min 时, 所提模型的 MAE 相较于 Autoformer 与 Informer 分别提升 33.6% 与 9.2%。

2) 所提模型能够克服 Informer 模型处理非平稳时间序列预测任务性能较差的劣势, 对比各模型其性能均有所提升。在预测步长 20~120 min 时, 相较于 Informer 模型, 所提模型的平均 MAE、RMSE、MAPE 分别提升 8.2%、7.1% 与 6.8%, 表明其在非平稳时间序列预测任务中更具优势。

参考文献

- [1] 朱春辉. 基于深度学习的认知无线电频谱预测技术研究[D]. 西安: 西安电子科技大学, 2023: 5-9.
ZHU C H. Research on cognitive radio spectrum prediction technology based on deep learning[D]. Xi'an: Xidian University, 2023: 5-9.
- [2] 徐东辉, 刘思佚, 张文博, 等. 基于改进简单加权法的多制式通信切换策略[J]. 火箭军工程大学学报, 2025, 39(1): 41-49.
XU D H, LIU S Y, ZHANG W B, et al. Improved simple additive weighting-based multi-standard communication switching strategy[J]. Journal

- of Rocket Force University of Engineering, 2025, 39(1): 41-49
- [3] ALI A, HAMOUDA W. Advances on spectrum sensing for cognitive radio networks: theory and applications[J]. IEEE Communications Surveys & Tutorials, 2017, 19(2): 1277-1304.
- [4] HUANG X L, TANG X W, HU F. Dynamic spectrum access for multimedia transmission over multi-user, multi-channel cognitive radio networks [J]. IEEE Transactions on Multimedia, 2020, 22(1): 201-214.
- [5] XIN C S, PAUL P, SONG M, et al. On dynamic spectrum allocation in geolocation spectrum sharing systems[J]. IEEE Transactions on Mobile Computing, 2019, 18(4): 923-933.
- [6] ELTOM H, KANDEEPAN S, EVANS R J, et al. Statistical spectrum occupancy prediction for dynamic spectrum access: a classification[J]. EURASIP Journal on Wireless Communications and Networking, 2018, 2018(1): 1-17.
- [7] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [8] YU L, CHEN J, DING G R. Spectrum prediction via long short term memory[C]//2017 3rd IEEE International Conference on Computer and Communications (ICCC). Piscataway, NJ, USA: IEEE, 2017: 643-647.
- [9] YU L X, WANG Q L, GUO Y F, et al. Spectrum availability prediction in cognitive aerospace communications: a deep learning perspective[C]//2017 Cognitive Communications for Aerospace Applications Workshop (CCAA). Piscataway, NJ, USA: IEEE, 2017: 1-4.
- [10] YU L, CHEN J, DING G R, et al. Spectrum prediction based on Taguchi method in deep learning with long short-term memory[J]. IEEE Access, 2018, 6: 45923-45933.
- [11] SHAWEL B S, WOLDEGEGBREAL D H, POLLIN S. Convolutional LSTM-based long-term spectrum prediction for dynamic spectrum access[C]//2019 27th European Signal Processing Conference. Piscataway, NJ, USA: IEEE, 2019: 1-5.
- [12] WANG L, HU J, JIANG R D, et al. A deep long-term joint temporal spectral network for spectrum prediction[J]. Sensors, 2024, 24(5): 1498.
- [13] KITAEV N, KAISER L, LEVSKAYA A. Reformer: the efficient transformer[EB/OL]. (2020-02-18) [2025-04-25]. <https://arxiv.org/abs/2001.04451>.

- [14] WU H X, XU J H, WANG J M, et al. Auto-former: decomposition transformers with auto-correlation for long-term series forecasting[EB/OL]. (2022-01-07) [2025-04-25]. <https://arxiv.org/abs/2106.13008>.
- [15] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11106-11115.
- [16] LIU Y, WU H, WANG J, et al. Non-stationary transformers: exploring the stationarity in time series forecasting[J]. Advances in Neural Information Processing Systems, 2022, 35: 9881-9893.
- [17] Wang L, Hu J, Zhang C, et al. Deep learning models for spectrum prediction: a review[J]. IEEE Sensors Journal, 2024, 24(18): 28553-28575.
- [18] PAN G L, WU Q H, DING G R, et al. Deep stacked autoencoder-based long-term spectrum prediction using real-world data[J]. IEEE Transactions on Cognitive Communications and Networking, 2023, 9(3): 534-548.
- [19] RAJENDRAN S, CALVO-PALOMINO R, FUCHS M, et al. Electrosense: open and big spectrum data[J]. IEEE Communications Magazine, 2018, 56(1): 210-217.
- [20] 杨迪丹. 基于 Informer 的频谱长序列预测[D]. 西安: 西安电子科技大学, 2023: 37-43.
YANG D D. Long sequence prediction of spectrum based on Informer[D]. Xi'an: Xidian University, 2023: 37-43.
- [21] TRAN N T, JACK X. Fourier-mixed window attention: accelerating informer for long sequence time-series forecasting[EB/OL]. (2024-04-17) [2025-04-25]. <https://arxiv.org/abs/2307.00493>.
- [22] KIM T, KIM J, TAE Y, et al. Reversible instance normalization for accurate time-series forecasting against distribution shift[C/OL]. The International Conference on Learning Representations. (2022-01-29) [2025-04-25]. <http://openreview.net/forum?id=cGDAkQo1C0p>.
- [23] PASSALIS N, TEFAS A, KANNIAINEN J, et al. Deep adaptive input normalization for time series forecasting[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(9): 3760-3765.
- [24] YANG Y X, DONG Y L. Informer_casual_LSTM: causal characteristics and non-stationary time series prediction[C]//2024 7th International Conference on Machine Learning and Natural Language Processing. Piscataway, NJ, USA: IEEE, 2024: 1-6.
- [25] 陈昊, 郭文普, 康凯. 基于相位变换和 GhostNet-门控循环单元的自动调制识别方法[J]. 火箭军工程大学学报, 2024, 38(4): 86-92.
CHEN H, GUO W P, KANG K. Automatic modulation recognition method based on phase transformation and GhostNet-GRU[J]. Journal of Rocket Force University of Engineering, 2024, 38(4): 86-92.
- [26] YANG Y G, GENG S P, ZHANG B C, et al. Long term 5G network traffic forecasting via modeling non-stationarity with deep learning[J]. Communications Engineering, 2023, 2: 33.

(编辑: 于福兴)