



基于TCN-MHSA模型的高速铁路 基站信号场强预测方法

俞晓波^{1,2}, 彭 艺¹

(1. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650500;
2. 中国铁路昆明局集团有限公司 昆明通信段, 云南 昆明 650011)

摘 要: 高速铁路基站信号场强是衡量基站信号强度的指标, 其质量直接影响无线网络覆盖基站的效果。对基站信号场强进行预测, 能够帮助铁路通信系统维护单位实时掌握基站运行状态, 助力运维人员提前制定网络优化计划, 排查并解决基站潜在隐患。文章提出一种基于时域卷积网络 (temporal convolutional network, TCN) 和多头自注意力 (multi-head self-attention, MHSA) 机制相结合的基站信号场强预测模型。该模型基于Abis接口数据, 采用TCN作为特征提取器, 利用其膨胀因果卷积机制扩大感受野, 高效捕获多尺度时间依赖关系, 然后引入MHSA机制, 在捕捉不同子空间特征、并行处理数据的同时, 根据数据不同维度的重要程度选择性提取信息, 从而增强模型对场强变化趋势的预测能力。试验结果显示, 在Abis接口数据集上, 该模型的平均绝对百分比误差 (mean absolute percentage error, MAPE)、均方误差 (mean-square error, MSE)、均方根误差 (root mean square error, RMSE)、平均绝对误差 (mean absolute error, MAE)、决定系数 (R -square, R^2) 指标均优于卷积神经网络 (convolutional neural network, CNN)、长短期记忆 (long short-term memory, LSTM) 网络、双向长短期记忆 (bidirectional long short-term memory, Bi-LSTM) 网络模型, 表明其具有良好的预测性能, 可为铁路通信网络质量监测及其后期网络优化提供决策支持。

关键词: 高速铁路; 时域卷积网络; 多头自注意力机制; Abis接口; 场强; 预测

中图分类号: U238; U285.21

文献标志码: A

文章编号: 1000-128X(2025)05-0149-08

0 引言

铁路综合数字移动通信系统 (global system for mobile communications-railway, GSM-R) 是铁路运输领域的核心通信基础设施^[1]。该系统除了提供基础调度通信功能, 在高速铁路场景中, 还是列车运行控制系统车-地信息传输的重要基础承载网络^[2], 网络信号场强作为评估 GSM-R 网络通信质量的核心指标^[3], 其质量直接影响车-地信息传输的电路交换数据链接稳定性, 因此分析网络信号覆盖情况、提前发现弱场区域, 可为网络优化决策提供重要的参考手段^[4]。

传统场强预测方法主要基于Hata模型^[5], 但建立符合实际铁路运行条件的电波传播模型, 需结合大量实际场强测试数据以保障建模效果^[6]。近年来, 基于深度神经网络的预测模型能从数据中自动学习相关特征, 具有良好的预测精度和鲁棒性, 受到众多研究者的关注。文献[7-8]提出利用人工神经网络预测 GSM-R 系统场强, 通过输入基站接收机信息与提取的环境特征, 无须构建复杂传播模型, 即可实现接收功率预测; 文献[9]采用广义回归神经网络进行场强覆盖预测, 分析了训练集构成和平滑因子对预测结果的影响; 文献[10]将神经网络训练任务分成多个子任务并行处理, 最终通过功能集成实现优于单一神经网络的预测性能; 文

基金项目: 国家自然科学基金资助项目 (62461030); 云南省基础研究重点项目 (202401AS070105)

通信作者: 俞晓波, 男, 主要从事铁路通信设备维护方面的研究; E-mail: 18206867540@163.com

引用格式: 俞晓波, 彭艺. 基于TCN-MHSA模型的高速铁路基站信号场强预测方法[J]. 机车电传动, 2025(5): 149-156.DOI:10.13890/j.issn.1000-128X.2025.05.016.

Citation: YU Xiaobo, PENG Yi. Signal field strength prediction method for high-speed railway base stations based on the TCN-MHSA model[J]. Electric drive for locomotives, 2025(5): 149-156.DOI:10.13890/j.issn.1000-128X.2025.05.016.

献[11]提出基于遗传算法和反向传播 (back propagation, BP) 人工神经网络的 GSM-R 无线网络通信场强预测模型, 相比传统 Hata 模型和单一 BP 神经网络模型, 预测精度显著提高; 文献[12]采用粒子群优化算法优化径向基函数神经网络参数, 构建 GSM-R 无线场强预测模型。

虽然神经网络模型在场强预测领域取得了一定进展, 但预测精度仍有待提升, 而且当前研究的建模数据多源自基站基础数据, 适用场景为基站初期建设规划^[13-15], 针对后期运维阶段的 GSM-R 网络场强变化预测研究仍较为欠缺。因此, 本文针对上述问题, 结合深度学习方法与接口监测系统采集的 Abis 接口数据, 提出一种基于时域卷积网络 (temporal convolutional network, TCN) 和多头自注意力 (multi-head attention, MHSA) 机制相结合的高速铁路基站信号场强预测模型。该方法有效融合了 TCN 和 MHSA 的优点, 即 TCN 的膨胀因果卷积通过间隔采样, 可高效捕捉序列不同尺度的时序特性; 在此基础上引入 MHSA 机制, 可自适应选择对预测更重要的时间步特征。二者结合使模型可准确捕捉基站场强变化规律与关键局部突变

点, 提升预测准确性, 为场强预测提供可靠的数据支撑。

1 基于 TCN-MHSA 的场强预测模型

由于机车移动电台所处的无线电波传播环境复杂, 其接收信号电平不仅会由于无线信道传播特性产生慢衰落, 还会由于大量多径信号的相长与相消干扰形成快衰落。为有效捕捉电平信号的时域衰落特征并实现场强预测, 本文采用对时序特征提取优异的 TCN 模型^[16], 并引入 MHSA 机制^[17]进行优化, 构建了基于 TCN-MHSA 的神经网络模型。该网络模型采用膨胀因果卷积对输入数据进行特征提取, 即膨胀因果卷积通过指数级扩展的膨胀因子, 在保持参数数量恒定的前提下显著扩大感受野, 使模型能有效捕获从局部到全局的多尺度时间依赖关系; 引入的 MHSA 机制可根据历史场强数据自动分配注意力权重, 强化关键特征信息, 提升模型对重要特征的敏感度。TCN-MHSA 模型由 TCN 层、MHSA 层和预测层组成, 整体框架如图 1 所示。

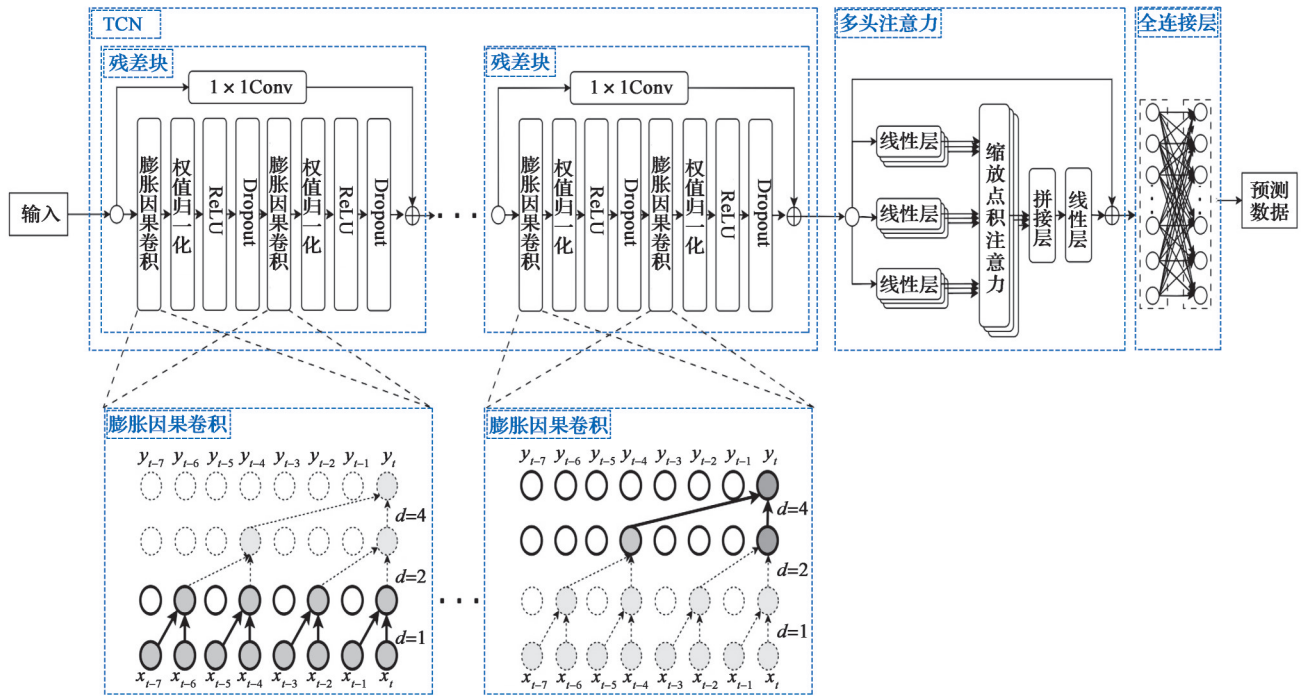


图 1 场强电平预测模型整体框架

Fig. 1 Overall framework of the field strength level prediction model

1.1 时域卷积网络层

时域卷积网络通过引入因果卷积保证模型的因果性, 即输出仅依赖当前与过去输入, 不依赖未来输入, 避免输入时序数据泄漏未来信息, 并通过残差连接保证梯度稳定, 其膨胀因果卷积结构和残差连接分别如图 2 和图 3 所示。

在 TCN-MSHA 模型中, 将离散的时间序列数据及其相关特征变量数据表示为

$$\mathbf{X} = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_{t-1}^1 \\ x_1^2 & x_2^2 & \cdots & x_{t-1}^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^d & x_2^d & \cdots & x_{t-1}^d \end{bmatrix} \quad (1)$$

式中: $\mathbf{X}$ 为输入矩阵; d 为输入矩阵的特征维度。

将 $\mathbf{X}$ 作为模型输入, 可得膨胀因果卷积表达式为

$$\mathbf{Y}_{i,t} = \sum_{j=1}^N \left(f(j) \cdot \mathbf{X}_{t-(N-j)d}^i \right) \quad (2)$$

式中: $\mathbf{Y}_{i,t}$ 为第 i 条 t 时刻的输出; $\mathbf{X}_{t-(N-j)d}^i$ 为第 i 条

$t-(N-j)d$ 时刻的输入; $f(j)$ 为卷积核权重; N 为卷积核大小; d 为膨胀因子, 通过调整 d 控制卷积核的采样间隔, 逐层扩大感受野, 从而捕捉长期依赖关系。

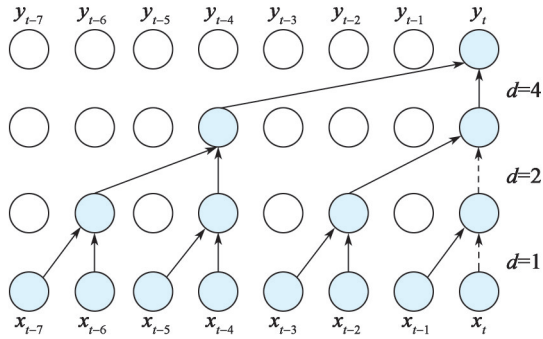


图2 膨胀因果卷积结构

Fig. 2 Dilated causal convolution structure

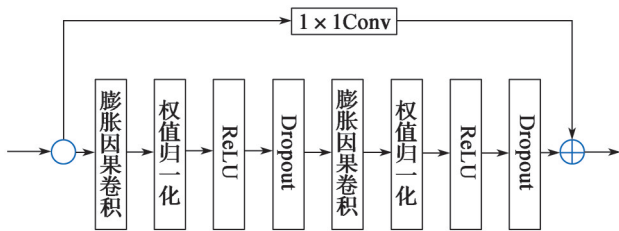


图3 残差连接示意

Fig. 3 Residual continuity

为加速训练收敛, 防止过度拟合, 引入非线性因素增强模型的表达能力。在卷积操作后添加权重归一化层、激活层和Dropout层, 其连接可表示为

$$D_i = \text{Dropout}(\text{ReLU}(\text{WeightNorm}(Y_{i,t}))) \quad (3)$$

式中: D_i 为经过权重归一化、激活层和Dropout操作后的输出; $Y_{i,t}$ 为第 i 层卷积。

为加强特征提取效果, 需累加膨胀因果卷积层。由于网络层次变深, 易出现梯度消失问题, 因此引入残差连接。每个卷积层的输出与对应的输入相加形成残差块, 通过激活函数得到该层输出为

$$O_i = \text{ReLU}(\text{res}(D_i, W_i) + \text{Conv}_{1 \times 1}(X_i)) \quad (4)$$

式中: O_i 为最终输出 O 的第 i 条输出数据, $O \in \mathbf{R}$; $\text{ReLU}(\text{res}(D_i, W_i) + \text{Conv}_{1 \times 1}(X_i))$ 为残差连接, 其中 X_i 为输入序列, W_i 为可变换参数矩阵, $\text{Conv}_{1 \times 1}(X_i)$ 通过 1×1 卷积实现与残差输出的特征维度匹配。

1.2 MHSA层

为提升模型对时序重要特征的敏感度, 弥补TCN在序列动态变化建模的不足, 本文在TCN层输出后引入MHSA机制, 将输入序列的特征向量映射至多个独立的子空间, 通过计算序列中每个位置与其他所有位置的注意力权重, 学习输入序列内部的依赖关系, 二者协同作用显著增强了模型对复杂时间模式的表征能力。这种层次化的特征提取和建模机制, 既保留了局部特征的精细结构, 又实现了全局时间依赖的有效建

模, 最终提升模型对复杂时间序列的建模精度。

假设TCN的输出数据表示为 $\mathbf{X} \in \mathbf{R}^{B \times T \times D}$, 其中 B 为输入流量数据的样本数, T 为时序长度, D 为特征维度。通过不同的可学习参数矩阵对 $\mathbf{X}$ 进行线性变换, 可分别映射得到自注意力机制所需的查询 (Query)、键 (Key)、值 (Value) 矩阵, 其数学形式如下:

$$\begin{cases} Q_i = \mathbf{X}W_i^q \\ K_i = \mathbf{X}W_i^k \\ V_i = \mathbf{X}W_i^v \end{cases} \quad (5)$$

式中: Q_i 、 K_i 、 V_i 按特征维度划分为 h 个头, 具体为 $Q_i = [Q_1, Q_2, \dots, Q_h]$, $K_i = [K_1, K_2, \dots, K_h]$, $V_i = [V_1, V_2, \dots, V_h]$; W_i^q 、 W_i^k 、 W_i^v 为可学习参数矩阵, $W_i^q \in \mathbf{R}^{D \times d_k}$, $W_i^k \in \mathbf{R}^{D \times d_k}$, $W_i^v \in \mathbf{R}^{D \times d_k}$, 其中 d_k 为每个头的维度。

单头自注意力计算公式为

$$\text{Attention}_i(Q_i, K_i, V_i) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i \quad (6)$$

式中: $\text{Attention}_i(\cdot)$ 为对第 i 个头注意力的计算。

通过求解多次注意力可得到多个不同输出, 拼接多个输出从而得到最终输出, 其表达式为

$$Z_{\text{multithread}}(Q, K, V) =$$

$$\text{Concat}(\text{Attention}_1, \text{Attention}_2, \dots, \text{Attention}_i) W^o \quad (7)$$

式中: $Z_{\text{multithread}}(Q, K, V)$ 为经线性变换得到的最终输出矩阵; Attention_i 为第 i 头的注意力输出; $\text{Concat}(\cdot)$ 为拼接函数; W^o 为输出的可变换参数矩阵。

1.3 预测层

为避免出现梯度消失问题, 引入残差连接, 前者对输入进行处理, 并映射为函数 $F(x)$, 后者直接将输入序列传递至网络末端, 与网络输出相加后得到最终结果, 表达式为

$$P' = O_i + M_i \quad (8)$$

式中: P' 为残差输出; O_i 为TCN处理后的输出; M_i 为MHSA机制层的输出。

为实现最终预测, 还需接入全连接层, 全连接层表达式为

$$P_{\text{pre}} = WP' + b \quad (9)$$

式中: P_{pre} 为最终预测输出; W 为可变换参数矩阵; b 为向量偏置。

2 试验及其试验结果分析

2.1 试验数据集及其预处理

本文数据来源于2024年12月中国铁路昆明局集团有限公司沪昆高速接口监测系统采集的基站Abis接口测量报告, 数据包括机车MSISDN号、占用小区及其小区广播信道号、上下行电平质量等信息, 测量报告如表1所示。

表 1 Abis接口测量报告
Table 1 Abis interface measurements

数据名称	含义
MSISDN	机车所用SIM卡号码
小区	机车当前占用小区名称
千米标	机车当前所处位置千米标
速度	机车当前运行速度
BCCH	机车当前所处小区的广播信道
RxLevelUp	机车当前占用小区上行接收电平值
RxLevelDown	机车当前占用小区下行接收电平值
RxQualUp	机车当前占用小区上行接收电平质量值
RxQualDown	机车当前占用小区下行接收电平质量值
TA	时间提前量
NBO电平	相邻小区下行电平值

数据集共含 65 839 条数据，其中 63 852 条用于模型训练，153 条用于单小区模型验证，1 814 条用于整条高速铁路小区模型验证。对数据集中部分下行场强电平数据进行可视化，如图 4 所示。由图 4 可知，实际采样的场强电平值波动范围较大，这一特征对精准分析其分布规律与变化趋势的预测模型带来挑战。

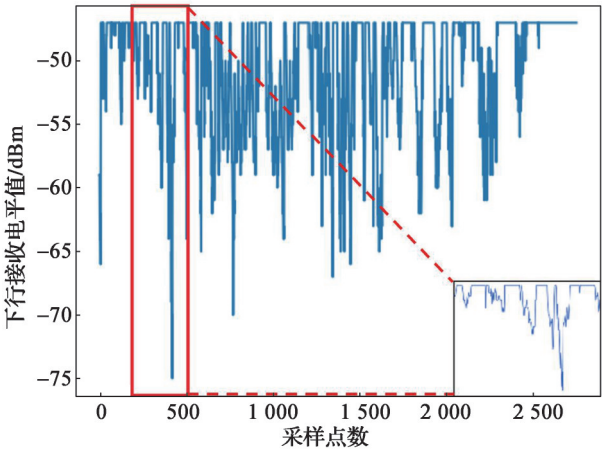


图 4 下行场强电平数据可视化
Fig. 4 Visualization of downlink field strength level data

为获取与下行场强电平值相关性强的关键特征变量，对数据集特征变量进行相关性分析，即首先剔除 MSISDN、小区、千米标等标识特征，再对数值型特征进行相关性分析，最终选定 RxLevelUp、TA、NB0 电平、RxLevelDown 作为关键变量，对应的相关性热力图如图 5 所示。

由图 5 可知，所选特征变量与场强电平存在不同程度相关性：①小区上行接收电平和下行电平场强的相关系数为 0.89，呈现高度相关性；②时间提前量、邻区电平值均与下行电平场强具有较高的相关性。通过特征变量选择，可去除冗余无关特征，仅将关键特征变量纳入下行场强电平预测模型训练，从而提升模型预测精度。

原始 Abis 接口数据存在缺失值，为保证试验数据

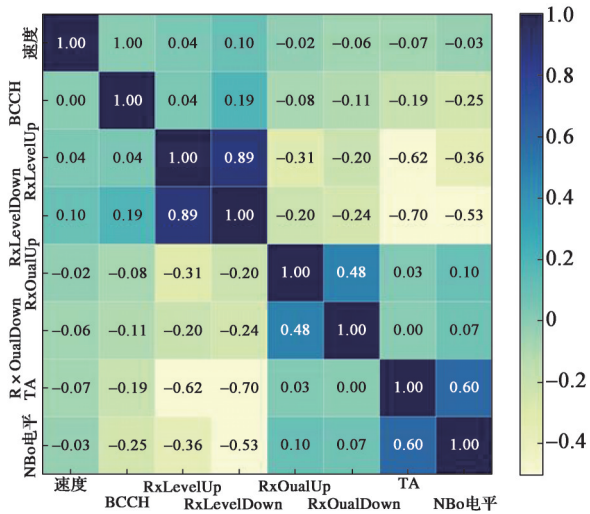


图 5 特征变量相关性热力图

Fig. 5 Correlation heatmap of characteristic variables

的完整性、提升预测模型的准确性与稳定性，对缺失值进行平均值插补填充；为消除特征间的量纲差异，均衡不同特征对模型的影响，同时加速模型训练、提升模型性能，对每个特征维度进行标准化操作，本文采用 StandardScaler 归一化方法，表达式为

$$Z_{\text{standardScaler}} = \frac{X - \mu}{\sigma} \tag{10}$$

式中： $Z_{\text{standardScaler}}$ 为归一化后的数据； X 为原始数据； μ 为特征均值； σ 为特征标准差。

特征数据处理完成后，将其输入 TCN-MHSA 模型，以预测基站场强。

2.2 试验环境与评价指标

本文提出的场强预测模型及其相关算法基于 Python 3.10 实现，深度学习框架选用 PyTorch 2.1。试验参数设置如下：样本批量为 32，训练周期为 50 轮；考虑到实际采样数据波动大、噪声多，为优化训练过程，采用平均随机梯度下降作为优化器。该优化器通过对多次迭代的参数取平均，可平滑更新轨迹，有效降低方差，使收敛过程更稳定，适配本试验的数据特点与训练需求。

模型超参数通过网格搜索确定，搜索空间如下：TCN 中卷积核大小为 2、3、4；TCN 网络层数为 1、2、3；MHSA 机制的头数为 2、4、8；隐藏层维度为 32、64、128；学习率为 0.001、0.005、0.010。网格搜索参数与验证损失关系的箱线图如图 6 所示。

由图 6 可知，学习率 0.005 对应的箱体位置低、范围窄，中位数显著低于 0.001、0.010，是收敛速度与稳定性的平衡点；隐藏层维度为 64 时，箱体中位数低于 32，并且异常值更少，整体分布比 128 更紧凑；TCN 网络层为 2 时，箱体中位数低于 1 层，比 3 层更稳定；卷积核为 4 时，箱体中位数最低，分布最紧凑，明显优于卷积核 2、3；MHSA 头数为 8 时，箱体中位数略低

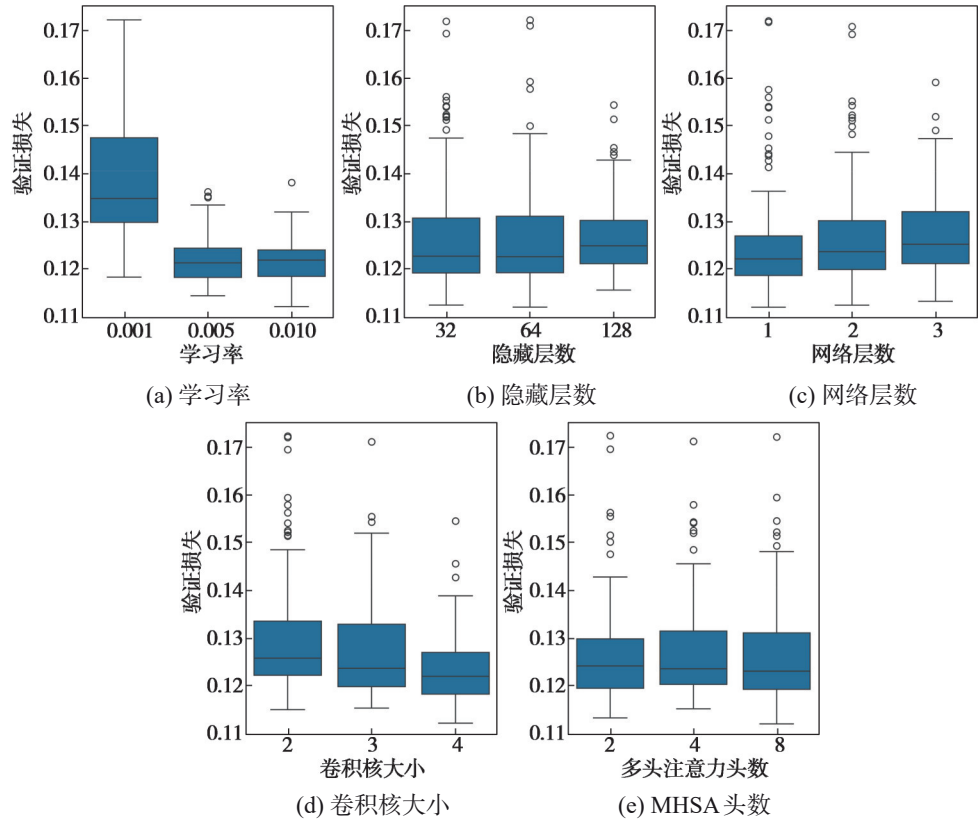


图6 参数与验证损失关系箱线图

Fig. 6 Box plot of the relationship between parameters and validation loss

于2、4，并且异常值更少。经过搜索，最终超参数设置如下：隐藏层维度为64，TCN网络层数为2，卷积核为4，MHSA机制的头数为8，学习率为0.005。

为评估本文方法的可靠性与有效性，本文采用5个主要评估指标进行试验验证，分别为均方误差（mean squared error, MSE）、均方根误差（root mean squared error, RMSE）、平均绝对误差（mean absolute error, MAE）、平均绝对百分比误差（mean absolute percentage error, MAPE）、决定系数（R-square, R^2 ），具体计算公式为：

$$e_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \tag{11}$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \tag{12}$$

$$e_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \tag{13}$$

$$e_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \tag{14}$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y} - y_i)^2} \tag{15}$$

2.3 试验分析

为验证本文模型在基站场强预测中的优越性，选取3种经典神经网络模型进行对比试验，分别为长短期

记忆（long short-term memory, LSTM）网络、双向长短期记忆（bidirectional long short-term memory, Bi-LSTM）网络和卷积神经网络（convolutional neural network, CNN）。此外，为满足消融试验需求，将TCN单独纳入试验。考虑到实际应用中需分析单个小区与多个小区的电平情况，试验分为单个小区电平预测与一个无线闭塞中心（radio block center, RBC）管辖区段内多小区电平预测两部分。

2.3.1 单个小区电平的预测

图7展示了5种预测模型在高速铁路单个小区电平预测中的趋势变化，表2进一步汇总了各模型的关键指标。图7采用双轴坐标系对比基站场强数据的实际值与预测值：横轴（x轴）为按空间顺序排列的小区编号，纵轴（y轴）为基站下行场强电平值。其中，“蓝色折线”表示实际测量的基站场强数据（反映真实网络状

表2 不同模型对高速铁路单个小区场强电平的预测精度对比

Table 2 Comparison of prediction accuracy across models for single-cell field strength level on high-speed railway					
模型	MAE	MSE	MAPE	RMSE	R^2
CNN	0.237 4	0.094 00	0.728 3	0.306 3	0.900 5
LSTM	0.289 8	0.125 80	0.997 0	0.354 6	0.866 6
Bi-LSTM	0.285 4	0.122 26	0.999 4	0.350 2	0.870 0
TCN	0.224 1	0.088 25	0.733 6	0.297 1	0.906 4
TCN-MHSA	0.213 0	0.075 20	0.729 5	0.274 2	0.920 2

态),“红色折线”表示模型预测的基站场强数据(展示预测效果)。

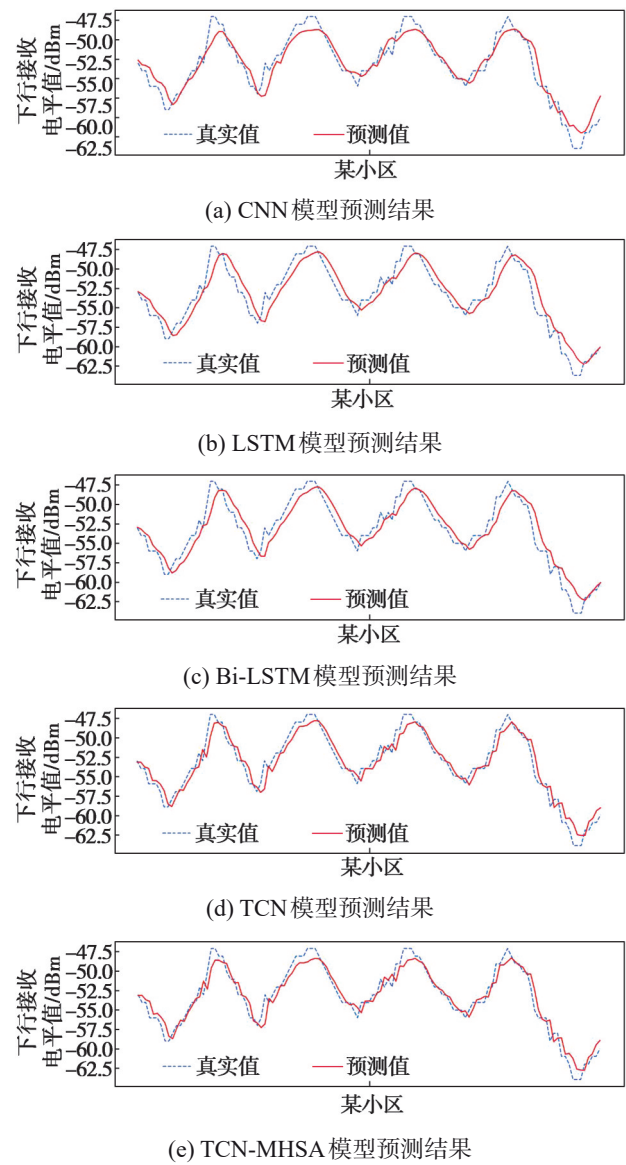


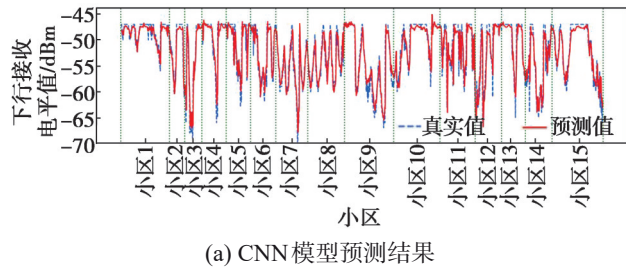
图 7 各模型单小区场强电平预测结果

Fig. 7 Prediction results of single-cell field strength level for models

2.3.2 一个RBC管辖区段内多小区电平的预测

图 8 展示了 5 种预测模型在高速铁路某 RBC 管辖区段内多小区电平预测中的趋势变化,表 3 进一步汇总了各模型的关键指标。

由图 7 和图 8 可知,5 个模型对基站信号电平变化趋势的预测整体表现良好,但均存在小幅波动,这一现象源于电平数据本身具有强烈的跳变性,降低了模



(a) CNN 模型预测结果

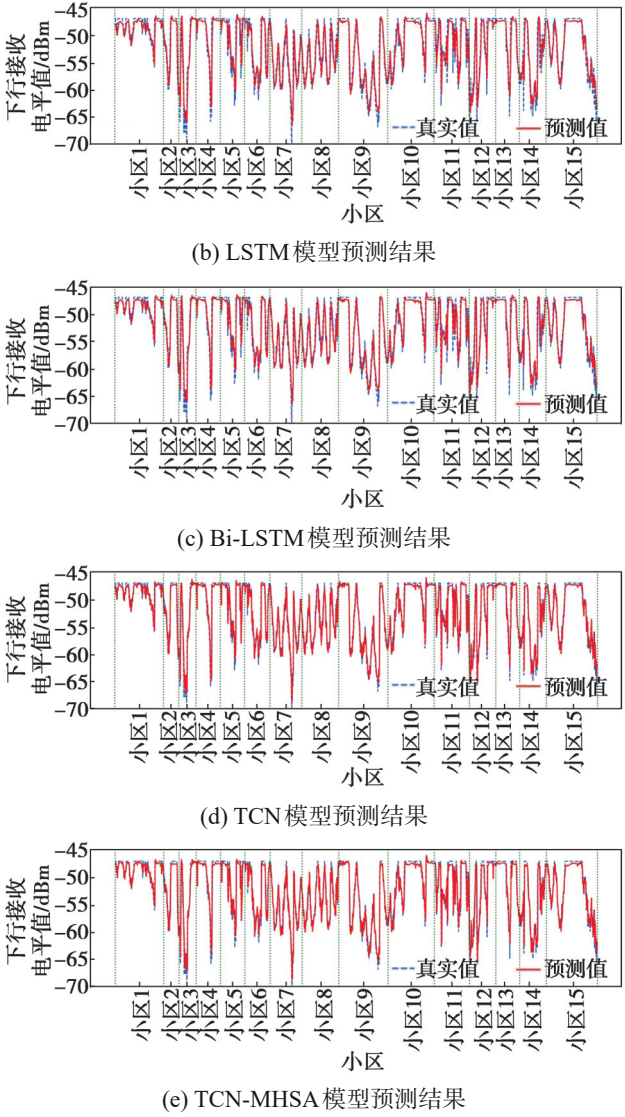


图 8 各模型多小区场强电平预测结果
Fig. 8 Prediction results of multi-cell field strength levels for models

表 3 不同模型对多小区场强电平预测精度对比
Table 3 Comparison of prediction accuracy across models for multi-cell field strength levels

模型	MAE	MSE	MAPE	RMSE	R^2
CNN	0.249 4	0.148 1	1.624 6	0.384 9	0.852 8
LSTM	0.258 1	0.167 6	1.651 4	0.409 4	0.833 5
Bi-LSTM	0.245 3	0.151 2	1.628 8	0.388 8	0.849 8
TCN	0.254 4	0.137 5	1.424 8	0.370 8	0.863 3
TCN-MHSA	0.211 0	0.117 6	1.387 0	0.343 0	0.883 1

型对特征的准确提取能力,其中 TCN-MHSA 模型的预测趋势与真实值最为接近,明显优于独立使用 TCN 的预测效果。CNN 模型在预测平稳值时波动较大,原因在于其擅长通过局部卷积提取局部特征,但容易忽略图像中的全局依赖信息;LSTM、Bi-LSTM 模型虽然更关注全局依赖关系,但是局部特征预测的细致度明显不足。由表 2 和表 3 可知,本文提出的 TCN-MHSA 模型在所有评价指标上均表现最优,充分验证了该模型

的有效性与优越性。

由以上试验得出结论: 与CNN模型相比, TCN-MHSA模型采用膨胀因果卷积, 在满足场强时序数据因果性约束的条件下, 能有效捕捉长时序序列的中长期特征依赖关系, 因此对长期趋势的预测效果更优; 与传统的LSTM和Bi-LSTM模型相比, TCN-MHSA模型的局部电平预测精度更高, 这得益于其引入的膨胀因果卷积网络, 通过聚焦关键特征、利用膨胀卷积网络提升时间序列分析的准确性与上下文感知能力; 与单独TCN模型相比, TCN-MHSA模型对场强电平局部突变的预测效果更佳, 这与MHSA机制的引入, 为特征空间分配更多权重, 确保模型聚焦关键空间关联(如信号切换区域), 从而更精准捕捉电平序列的突变特征相关。

3 结束语

本文提出一种基于TCN和MHSA机制相结合的基站信号场强预测模型。相比于其他几种常用模型, TCN-MHSA模型之所以能更准确地捕捉基站信号场强的变化趋势, 核心在于其融合了TCN与MHSA机制的优点, 能更好地理解输入数据特点、提取电平数据的时序特征。因此, 该模型在MAPE、MSE、RMSE、MAE、 R^2 等指标方面相较于其他模型展现出更优的拟合效果, 可为GSM-R网络中基站信号场强监测与网络优化指导提供理论与应用参考。

参考文献:

- [1] 王凯燕. 铁路GSM-R场强覆盖预测模型研究[D]. 兰州: 兰州交通大学, 2023.
WANG Kaiyan. Research on prediction model of railway GSM-R field strength coverage[D]. Lanzhou: Lanzhou Jiatong University, 2023.
- [2] 杨锐, 郭桂芳, 赵武元. 铁路5G-R运维技术方案研究[J]. 中国铁路, 2021(8): 19-25.
YANG Rui, GUO Guifang, ZHAO Wuyuan. Research on technical scheme for operation and maintenance of 5G-R[J]. China railway, 2021(8): 19-25.
- [3] 曹峰, 郜新军. 高速铁路联调联试总体方案研究[J]. 轨道交通装备与技术, 2024, 32(1): 60-64.
CAO Feng, GAO Xinjun. Research on the overall scheme of high-speed railway joint commissioning and joint test[J]. Rail transportation equipment and technology, 2024, 32(1): 60-64.
- [4] 张涛, 高利民, 姜尚炎. 基于检测数据的GSM-R网络评估方法[J]. 中国铁路, 2023(11): 92-100.
ZHANG Tao, GAO Limin, JIANG Shangyan. GSM-R network evaluation method based on inspection data[J]. China railways, 2023(11): 92-100.
- [5] 钟章队, 艾渤, 刘秋研, 等. 铁路数字移动通信系统(GSM-R)应用基础理论[M]. 北京: 北京交通大学出版社, 2009.
ZHONG Zhangdui, AI Bo, LIU Qiuyan, et al. The basic theoretical foundation of railway digital mobile communication system (GSM-R)[M]. Beijing: Beijing Jiaotong University Press, 2009.
- [6] 韩斌杰, 杜新颜, 张建斌. GSM原理及其网络优化[M]. 2版. 北京: 机械工业出版社, 2009.
HAN Binjie, DU Xinyan, ZHANG Jianbin. GSM principle and network optimization[M]. 2nd ed. Beijing: China Machinery Press, 2009.
- [7] 丁珂. 基于误差反向传播神经网络的GSM-R场强预测[J]. 电气化铁道, 2022, 33(1): 67-70.
DING Xun. GSM-R field strength prediction based on error backpropagation neural network[J]. Electric railway, 2022, 33(1): 67-70.
- [8] 靳翔. 铁路GSM-R场强覆盖的预测与调整方法的研究[D]. 兰州: 兰州交通大学, 2021.
JIN Xiang. Research on the prediction and adjustment method of railway GSM-R field strength coverage[D]. Lanzhou: Lanzhou Jiatong University, 2021.
- [9] 关捷, 李国宁, 温宇钧. 基于GRNN的GSM-R场强覆盖预测算法[J]. 铁道标准设计, 2014, 58(2): 106-112.
GUAN Jie, LI Guoning, WEN Yujun. Prediction algorithm of GSM-R field intensity coverage based on GRNN[J]. Railway standard design, 2014, 58(2): 106-112.
- [10] 胡荣青, 张成挺, 任国伟, 等. 基于深度残差神经网络的5G信号室内分布预测[J]. 电子技术应用, 2024, 50(12): 77-81.
HU Rongqing, ZHANG Chengting, REN Guowei, et al. Prediction of indoor 5G signal distribution based on deep residual neural network model[J]. Application of electronic technique, 2024, 50(12): 77-81.
- [11] JIN X, ZHENG Y S. GSM-R wireless field strength coverage prediction algorithm based on GA-BP algorithm[C]// IOP conference series: earth and environmental science, 2020, 526: 012157. DOI: 10.1088/1755-1315/526/1/012157.
- [12] WANG K Y, ZHOU Q H. GSM-R wireless field strength coverage prediction algorithm based on PSO-RBF neural network algorithm[J]. Journal of physics: conference series, 2022, 2383(1): 012097. DOI: 10.1088/1742-6596/2383/1/012097.
- [13] 陈小平. 基于TA值的CTCS-3无线超时分析[J]. 铁路通信信号工程技术, 2021, 18(12): 32-37.
CHEN Xiaoping. Analysis of CTCS-3 radio timeout based on TA value[J]. Railway signalling & communication engineering, 2021, 18(12): 32-37.
- [14] 赵一颖, 姜志威. 基于实地模型的GSM-R仿真系统研究[J]. 铁道标准设计, 2021, 65(7): 178-182.
ZHAO Yiyang, JIANG Zhiwei. Research on GSM-R simulation system based on realistic geographic model[J]. Railway

- standard design, 2021, 65(7): 178-182.
- [15] 刘勇. 400 MHz 列车数字无线调度通信系统大尺度电波传播模型研究[D]. 北京: 中国铁道科学研究院, 2024.
- LIU Yong. Research on large-scale radio wave propagation model of 400 MHz train digital wireless dispatching communication system[D]. Beijing: China Academy of Railway Sciences, 2024.
- [16] BAI S J, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J/OL]. arXiv: 180.01271, 2018[2025-06-05]. DOI: 10.48550/arXiv.1803.01271.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing System. Long beach: ACM, 2017: 6000-6010. DOI: 10.65215/r5bs2d54.

Signal field strength prediction method for high-speed railway base stations based on the TCN-MHSA model

YU Xiaobo^{1,2}, PENG Yi¹

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 2. Kunming Communication Depot, China Railway Kunming Group Co., Ltd., Kunming, Yunnan 650011, China)

Abstract: Signal field strength measures the power of signals transmitted by base stations along high-speed railways and directly affects wireless network coverage. Predicting field strength enables railway communication maintenance enterprises to monitor the operational status of base stations in real time, and helps operation and maintenance personnel plan network optimizations and identify and mitigate potential hazards. This paper proposes a signal field strength prediction model for base stations, which combines a temporal convolutional network (TCN) with a multi-head self-attention (MHSA) mechanism. Using Abis interface data, this model employed TCN as a feature extractor, leveraging its dilated causal convolution mechanism to expand the receptive field and efficiently capture multi-scale temporal dependencies. The MHSA mechanism captured features across different subspaces, processed data in parallel, and selectively extracted information across different dimensions based on their importance, thereby enhancing the model's ability to predict field strength trends. Experimental results on Abis interface datasets show the TCN-MHSA model outperforms the convolutional neural network (CNN), long short-term memory (LSTM) network, and bidirectional long short-term memory (Bi-LSTM) network models in the following indicators: mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and R square (R^2). These results indicate the model's strong predictive performance can provide decision-making support for the quality monitoring of railway communication networks and future network optimization.

Keywords: high-speed railway; temporal convolutional network; multi-head self-attention mechanism; Abis interface; field strength; prediction