

基于预制架空长输热力蒸汽管道系统的一种热负荷预测方法

黄永福 王超 冯棕 李庆哲

中国电力工程顾问集团中南电力设计院有限公司 湖北武汉 430071

【摘要】文章提出了一种深度学习的热负荷预测方法,旨在提高热负荷预测的精度。强调了国家发改委和能源局发布的《全国煤电机组改造升级方案》中对整合供热资源和推进热网互联的重视,以及智慧供热技术的发展需求。针对工业园区和特殊工况下热负荷预测的研究不足,通过MIC分析筛选出与热负荷相关的气象和经济因素,构建了特征集。采用BiLSTM和Transformer网络的理论基础,并提出了PatchTST-BiLSTM混合预测模型。通过实测数据的算例分析,验证了该方法具有高预测精度。最新的气象因素和经济因素数据,以1小时为时间分辨率进行训练、验证和测试。选用了均方根误差(RMSE),平均绝对误差(MAE)和平均绝对误差百分比(MAPE)作为评价指标。实验结果表明,所提方法的预测精度高于传统的Transformer和PatchTST模型,以及其它几种机器学习和深度学习算法。最终得出结论,结合MIC相关性分析和PatchTST-BiLSTM深度学习技术的热负荷预测方法能更准确地计算热网负荷需求变化,提高供热效率,对于智慧供热领域具有一定的参考价值。

【关键词】热负荷预测;深度学习;智慧供热

1 项目背景

国家发改委、能源局发布《全国煤电机组改造升级方案》,强调整合供热资源,支持热网及老旧管网改造,推进热网互联,发挥热电联产供热能力。鼓励合理发展长输供热,扩大供热范围。长输蒸汽管网建设快速发展,覆盖距离增长,技术和能耗要求提高。智慧供热利用大数据、物联网、AI技术提升管理效率,减少能耗,推动供热行业智能化升级。但当前研究多集中于城市供热网络,对工业园区及特殊工况研究不足,需加强验证分析。

2 本文主要研究方法

为进一步提高热负荷预测精度,本文提出一种基于PatchTST-BiLSTM混合深度学习热负荷预测方法。通过最大互信息素分析热负荷、气象和经济因素的相关性,筛选特征并构建特征集。提出了一种Transformer的变体网络(PtachTST)与双向长短时记忆网络(BiLSTM)的混合预测模型。该模型结合了Transformer变体网络PtachTST和双向长短时记忆网络BiLSTM。PtachTST通过Token嵌入和多头自注意力机制提取局部信息和时序依赖关系,而BiLSTM则对特征向量进行双向建模。实测数据的算例分析验证了该方法的高预测精度。最后,基于实测负荷数据开展算例分析,实验结果证明了所提方法具有更高的预测精度。

3 基于MIC的相关性分析

MIC通过将两个变量间的关系离散在二维空间中,并基于网格划分的互信息计算来获取两个负荷之间的相关性。热负荷作为一种典型的时间序列,其内部变化规律复杂,不仅与用户的用能行为密切相关,还与气象因素、经济因素等存在着复杂的线性和非线性相关关系。相较于常用的Pearson相关系数和Spearman相关系数,MIC可以同时计算各变量间的线性和非线性相关性,且在保持计算复杂度低的基础上具有更高的鲁棒性。MIC的计算公式如式(1.1)所示。

$$MIC(x, y) = \max_{a \times b < S} \left(\frac{\int \int p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \log_2 \min(a, b)}{\log_2 \min(a, b)} \right) \quad (1.1)$$

本文利用MIC分析热负荷与不同气象因素以及经济因素的相关性,从而筛选重要特征,为后续输入特征集构建奠定基础。

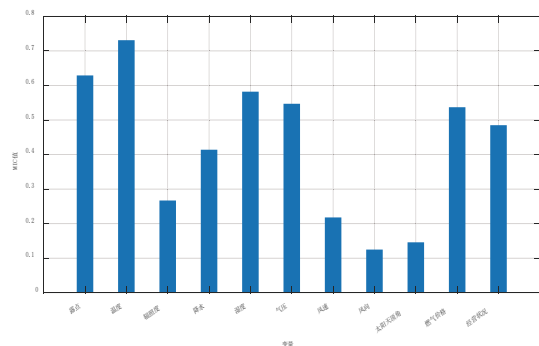


图1.1 MIC相关性指标

由图1.1可知，温度对热负荷的影响程度最大，其MIC为0.732。这是因为热负荷属于气象敏感负荷，当温度变化较大时，人类的用能行为也会发生较大变化。在经济因素中，燃气价格和经营状况也与热负荷具有较强相关性，其MIC值均高于0.3。基于上述分析，本文选择上述MIC均高于0.3的因素纳入输入特征。

4 算法理论介绍

4.1 双向长短时记忆网络

在序列建模任务中，以长短时记忆网络（LSTM）为代表的循环神经网络具有较强的时间依赖关系捕捉能力。相较于传统递归神经网络（RNN），LSTM通过引入输入门和遗忘门以及输出门而得到改进。从而解决了RNN在训练过程中的长期依赖性问题，避免了梯度爆炸现象。LSTM中各变量的计算规则如下所示。

$$\begin{cases} u_t = \sigma(W_{xu}x_t + W_{hu}h_{t-1} + b_u) \\ f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ g_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \\ c_t = c_{t-1} \odot f_t + g_t \odot u_t \\ h_t = \tanh(c_t) \odot o_t \end{cases} \quad (1.2)$$

式中： u_t ， f_t ， o_t 分别表示 t 时刻输入门，遗忘门和输出门； W_{xu} ， W_{xf} ， W_{xo} ， W_{xg} 为 x_i 对应的权重矩阵； W_{hu} ， W_{hf} ， W_{ho} ， W_{hg} 为 h_{i-1} 对应的权重矩阵； b_u ， b_f ， b_o ， b_g 为偏置矩阵； x_i 为 t 时刻的输入特征矩阵； h_t 为 t 时刻的隐藏状态； c_t 为中间状态值； σ 为sigmoid激活函数； $\odot$ 表示元素按元素乘积。LSTM的基本网络结构与内部结构如图1.2和图1.3所示。

LSTM的基本网络结构与内部结构如图1.2和图1.3所示。

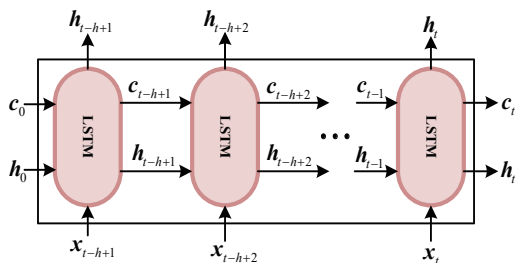


图1.2 LSTM网络结构

单向LSTM只能捕捉负荷序列从过去到未来的时间依赖性，而无法推导出反向的特征信息。本文采用BiLSTM对负荷序列进行时间特征提取，以解决单向LSTM的固有

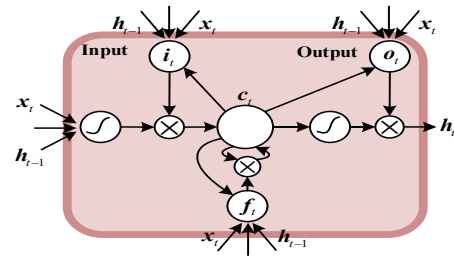


图1.3 LSTM内部结构

缺陷。图1.4描述了BiLSTM的结构，它由正向和反向单向LSTM组成。BiLSTM可以向前和向后训练。与单向LSTM相比，BiLSTM可以充分利用输入信息，更全面地进行时间特征提取。

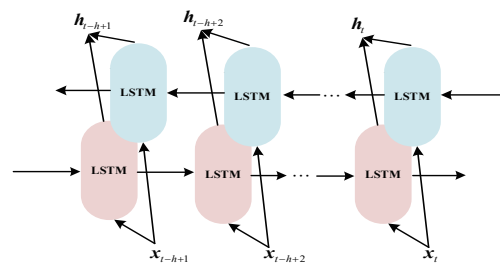


图1.4双向LSTM网络结构

如图1.3所示，BiLSTM的隐藏状态 h 由正向LSTM和反向LSTM共同决定，具体如下：

$$\tilde{h} = [\vec{h} \oplus \overleftarrow{h}] \quad (1.3)$$

4.2 Transformer网络

由于本文只用到了Transformer编码器结构，因此不再对解码器结构进行说明。图1.5展示了典型的Transformer编码器结构。

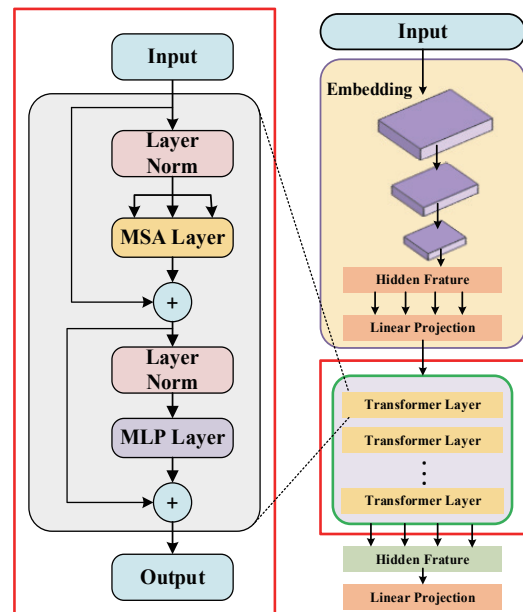


图1.5 Transformer编码器结构

5 PatchTST-BiLSTM热负荷预测模型

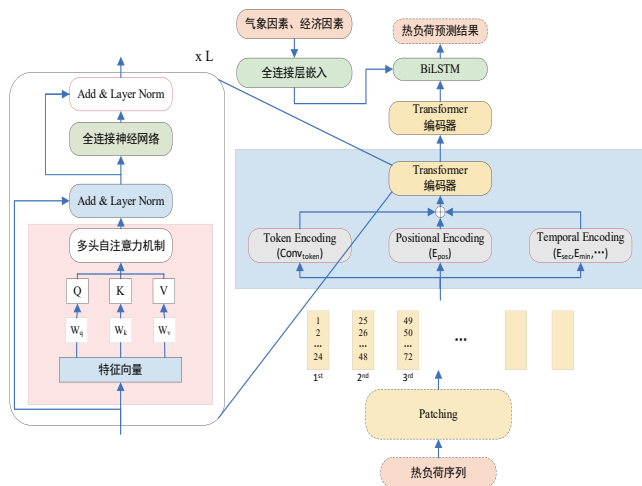


图1.6负荷预测模型结构

图1.6展示了PatchTST-BiLSTM的模型架构。首先，对原始负荷序列进行Patching，Patching将原始负荷序列划分为若干子序列。然后，将每一个序列视为一个token，并进行三种编码操作：Token Encoding、Positional Encoding以及Temporal Encoding。将编码后彼此相互独立的token输送至Transformer编码器中，基于Transformer的多头自注意力机制建立全局依赖关系，通过位置前馈全连接神经网络对每个token中的局部信息进行非线性特征提取。最后，将Transformer编码器输出的特征向量进一步通过BiLSTM进行双向时间依赖关系捕捉，经全连接层投影至预测窗口维度输出预测结果。

6 仿真分析

6.1 实验数据介绍与处理

本研究的实验数据来自最新的某测点实际热负荷数据、温度、压力以及相关气象因素和经济因素。时间分辨率为

1小时。训练集、验证集、测试集的比例为7:2:1。为了全面地评价模型预测结果，选取均方根误差（RMSE），平均绝对误差（MAE），平均绝对误差百分比（MAPE）作为评价指标。实验平台的硬件配置为Intel(R)Core(TM)i7-10870H CPU@2.20GHz 2.21GHz处理器，显卡NVIDIA GeForce RTX 2060 GDDR6@6GB和256G内存；软件采用Windows11操作系统，Matlab 2023b仿真实验平台。本文采用滑动窗口的方式构建输入特征集，滑动窗口原理如下所示：（见图1.7）

在超参数选择中，分为模型结构超参数和训练算法超参数。所提方法的主要模型结构超参数包括多头自注意力的头数，BiLSTM的层数与神经元个数。训练算法的主要超参数包含了学习率，批次大小以及最大轮次。由于超参数寻优空间较大，单纯网格搜索耗时且计算量大，不适合超短期预测。本文结合人工经验和网格搜索确定超参数，即大部分超参数由人工经验确定，少量留给机器进行网格搜索。最终，在验证集中确定的最优超参数配置如下。

在PathTST-BiLSTM负荷预测模型中，多头自注意力机制的头数为16，BiLSTM的层数为2，神经元个数分别为64与128，学习率为0.01，批次大小为32，最大轮次为100。

6.2 预测模型评价指标

为了验证和评估本文所提预测模型以及其他对比模型的预测性能，将采用以下平均绝对误差（MAE）、均方根误差（RMSE）和决定系数R2四个模型评价指标进行预测结果的综合评估：

MAE通过计算预测值与实际值之间差异的绝对值然后取平均来衡量模型预测的平均偏差大小。MAE公式如下：

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|^2 \quad (1.4)$$

RMSE是衡量预测值与实际值之间差异的常用指标，它反映了预测误差的标准差。RMSE的值越小，表示模型的预测值与实际值之间的差异越小，模型的预测性能越好。RMSE的计算公式如下：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \quad (1.5)$$

R2是模型拟合优度的评价指标，R2的值通常介于0和1之间。其值越接近1，说明模型对数据的解释能力越强，预测的准确性越高。决定系数公式为：

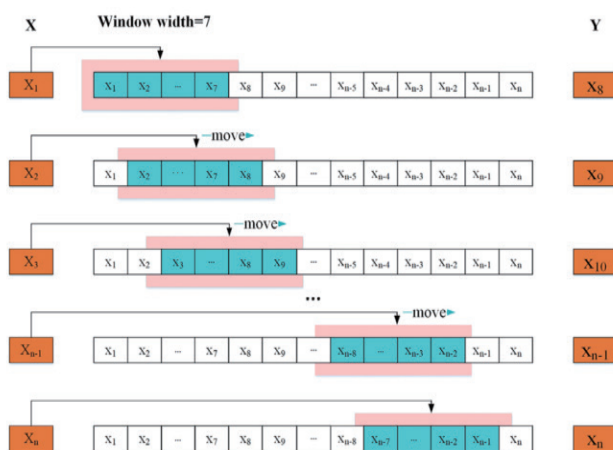


图1.7滑动窗口流程

$$R^2 = 1 - \frac{\sum_{t=1}^N (x_t - \hat{x}_t)^2}{\sum_{t=1}^N (x_t - \bar{x}_t)^2}$$

(1. 6)

6.3 仿真实验分析

图1. 8展示了本文所提预测模型的预测结果，可知本文所提模型的预测结果可以很好地贴合热负荷的实际值。此外，为证明所提方法的合理性，本文设置了两个对比实验，分别为传统的Transformer和PatchTST。

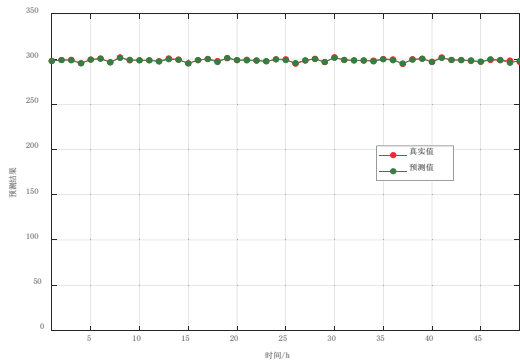


图1. 8本文预测结果

表1预测误差评价指标

Model	RMSE	R2	MAE
Transformer	247.984	0.821	187.045
BiLSTM	238.974	0.831	183.398
本文所提方法	219.171	0.857	174.954

据表1可知PatchTST具有更高的预测精度。说明了PatchTST通过保留负荷序列的局部信息可以更好地学习到负荷内部的变化规律，从而提高模型的预测性能。

为证明所提方法的优越性，选取了在负荷预测领域表现优异的机器学习和深度学习算法，分别是BiLSTM，时间卷积神经网络（TCN），支持向量机（SVM），极端梯度提升X（GBoost）。测试集的选取与消融实验相同，各模型的预测误差如表2所示。

表2预测误差评价指标

Model	RMSE	R2	MAE
本文模型	266.154	0.813	192.225
TCN	271.956	0.803	212.962
SVM	292.745	0.781	236.051

XGBoost	256.839	0.817	190.062
---------	---------	-------	---------

据表2可知，所提方法相较于其他对比模型具有更高的预测精度。此外，各对比模型的预测精度均低于消融实验中的模型1。证明了本文提出的预测方法无论是在模型结构还是方法设计上是合理且有效的。

7 结论

本文介绍了一种结合MIC相关性分析和PathTST-BiLSTM深度学习技术的热负荷超短期预测方法。该方法通过MIC分析确定了影响热负荷波动的关键气象和经济因素，并选取了主要气象因素作为预测模型的输入特征。通过这种方式，预测模型能更准确地计算热网负荷需求变化，提高供热效率。同时，该方法还量化分析了热负荷的自相关性，并利用滑动窗口技术构建了预测模型的输入特征集。在此基础上，提出了一个基于PathTST-BiLSTM的混合深度学习预测模型，其中PatchTST作为编码器进行特征提取，BiLSTM作为解码器进行双向建模，最终通过全连接层输出预测结果。仿真对比实验显示，该预测方法的精度高于现有深度学习模型，具有一定的参考价值。

参考文献:

[1] 李玉志, 刘晓亮, 邢方方, 温国强, 卢楠滢, 何慧, 焦润海. 基于Bi-LSTM和特征关联性分析的日尖峰负荷预测[J]. 电网技术, 2021, 45(7): 2719-2730.

[2] 王继东, 杜冲. 基于Attention-BiLSTM神经网络和气象数据修正的短期负荷预测模型[J]. 电力自动化设备, 2022, 42(4): 172-177.

[3] 梁珊. MIC与传统相关分析方法比较研究[J]. 山西青年, 2019, (9): 47, 49.