

基于 Attention-GAT-LSTM 的算法模型在新型电力系统中的应用探索

刘锦涛, 孙玉芹, 郭子涛, 王添翼, 程文

(上海电力大学数理学院, 上海 201306)

摘要: 精确的短期电力负荷预测对新型电力系统日发电计划的制订和实时调度至关重要, 为取得准确可靠的负荷预测结果, 针对真实用电负荷数据的时序性、不确定性等特征, 提出了一种基于 Attention-GAT-LSTM 的智能算法, 并应用在实际的新型电力系统中。在原始数据的处理中创新地结合了自注意力机制, 引入了数据处理单元附加权值, 并采用跳跃连接机制防止结果出现过拟合; 将处理后的数据传递到图注意力网络(graph attention network, GAT)进行空间节点的特征提取, 再传递到长短期记忆网络(long short-term memory, LSTM)进行时间特征的提取; 通过前向传播、反向传播和梯度下降方法, 使 LSTM 层的权重和偏置得到迭代更新, 有效地减少信息在迭代过程中的丢失并突出关键时间点信息。最后通过多种不同模型的对比分析, 验证了该方法在短期电力负荷预测(小时级)时具有较高的预测精度, 可以为新型电力系统的运行调度、规划建设提供数据支持。

关键词: 新型电力系统; 电力负荷预测; 图神经网络; 自注意力机制; 长短期记忆网络

Exploration of the Application of Attention-GAT-LSTM Algorithm Model in New Power Systems

LIU Jintao, SUN Yuqin, GUO Zitao, WANG Tianyi, CHENG Wen

(College of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 201306, China)

Abstract: Accurate short-term power load forecasting is essential for the formulation of daily power generation plans and the real-time dispatching of new power systems. In order to obtain accurate and reliable load forecasting results, this paper proposes an Attention-GAT-LSTM intelligent algorithm targeting the temporal and uncertain characteristics of real power load data, and applies it in practical new power systems. In raw data processing, this algorithm innovatively integrates a self-attention mechanism, introduces a data processing unit to assign weights, and employs a skip-connection mechanism to prevent overfitting. The processed data is transmitted to a graph attention network (GAT) for spatial node feature extraction, then passed to a long short-term memory (LSTM) network for temporal feature extraction. Through forward propagation, back propagation, and gradient descent methods, the weights and biases of the LSTM layer are iteratively updated, effectively minimizing information loss during iterations and highlighting key time-point information. Comparative analyses with various models demonstrate the method's high prediction accuracy in short-term (hour-level) power load forecasting, providing data support for the operational dispatching, planning and construction of new power systems.

Key words: new power system; power load forecasting; graph neural network; self-attention mechanism; long short-term memory network

0 引言

能源供给是社会运行的重要保障,而电力系统的安全稳定运行对国家经济发展和居民生活有重要影响。在党中央“碳达峰”、“碳中和”战略目标的推动下,大规模新能源接入电力系统,使传统电力系统的运行受到了较大考验。而电力负荷预测是电力系统稳定运行和规划的重要环节,对社会生产生活也有很大影响,因此电力负荷预测在学术界和工业界一直都有着较高的研究热度^[1-2]。历史电力负荷序列具有时序性、随机性、波动性和非线性特征,需要开发有效的模型以解决上述问题,为电力系统的安全可靠运行提供保障。如今工业界逐渐向工业互联网迈进,其必然会带来用电负荷特性的变化,使负荷特性变得更加复杂和多变,对于短期电力负荷预测提出了更高要求^[3-4]。

针对负荷数据的时序性特征,传统的电力负荷预测方法有诸如:线性回归法、时间序列法、趋势外推法、灰色预测法等,基本思想是利用历史电力负荷数据对未来负荷做预测,以此来充分捕捉时序性特征。但是这些方法对非线性电力负荷数据处理能力有限^[5-8]。

近年来,随着机器学习领域的研究不断深入,相应的模型也被广泛地应用在电力负荷预测领域,并取得了较好的成果。传统的机器学习算法包括随机森林、朴素贝叶斯、k均值聚类等,这些算法模型可以很好地捕捉负荷与影响因素间的非线性关系,解决相对复杂的数据特征关系等问题^[9-10],但它们在大数据集的表现并不好。随着经济的发展和工业产业的技术进步,对电力需求也来到了新的层面,对电力有关部门和企业而言,找寻更加精准、有效的预测方法是重中之重^[11]。深度学习算法在近年来取得了进一步发展并逐渐兴起,较为常见的有:卷积神经网络(convolutional neural networks, CNN)、循环神经网络(recurrent neural network, RNN)、长短期记忆网络(long short-term memory, LSTM)等。这些深度学习算法在电力负荷预测领域已经有了广泛的应用^[12-13],并且有相关研究人员通过组合模型的方式使得整体预测效果有所提升^[14-17]。

考虑到多用户用电行为之间潜在关联关系的存在,上述模型均只注重于挖掘用电负荷序列之间的

潜在时序关系,故限制了预测精度的进一步提升。近年来,图神经网络算法因为可以有效学习图结构数据,因此具备挖掘用电负荷序列之间的潜在空间关系,在交通预测领域有广泛的应用^[18]。也吸引了相关研究人员将其应用于用户负荷预测领域^[19-20],并通过结合其他算法的方式取得了较好的效果^[21-22]。

目前已有的相关研究不管是侧重于负荷的时间序列性关系的算法模型,还是侧重于用户节点间相关性的图神经网络算法模型都有局限性^[23]。为进一步挖掘各个用户负荷时序关系,同时可以挖掘其中潜在空间关系,本文提出了一种基于自注意力机制并结合 LSTM 和图注意力网络的算法模型。首先,通过自注意力机制编写一个数据处理器,将原始数据按照需要特征进行分类,根据不同需求可以添加诸如节假日、天气、气候等特征,为后续的模型训练提供更精细化的特征数据。其次,将特征数据经过线性变化层,对特征进行分析分类,挖掘潜在的用户关系信息,并通过图注意力网络层进行用户空间节点信息的挖掘提取。接着,将通过 LSTM 模块进行时序特征的提取,并做出时序预测。最后所有数据输入到全连接层,根据设定的时间步长需求来进行相应的负荷值预测。通过采用公开的真实电力负荷数据集进行验证,结果表明本文所提模型表现皆优于其他模型。

1 模型架构

1.1 数据处理单元

为保证原始数据的利用最大化和读取效率最大化,编写了一个专用的数据处理单元,结合了自注意力机制为其处理后的数据进行权值赋予,对后续模块的运行提供便利。

考虑到负荷数据大多以日、周、月、年为统计周期,因此数据处理单元是以此为参考,从原始数据中截取后再进行运行并赋权值的。其结构简化图如图 1 所示。

将原始数据按顺序添加进 A_n 中,再根据具体时间尺度需求决定需要的输入序列或矩阵的规模。对完成切割的数据进行分类处理,再结合自注意力机制进行数据隐藏节点的挖掘并赋予相应的自注意力权值,为后续模块更好地处理节点间空间关系、时间关系提供基础。

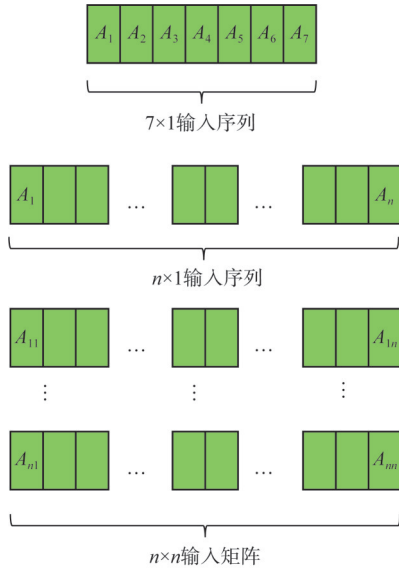


图1 数据处理单元输入简化示意图

Fig. 1 Simplified schematic diagram of input for data processing unit

其中自注意力机制是通过动态权重分配机制实现输入序列的上下文建模。核心在于建立序列元素间的依赖关系强度评估体系：对于给定的输入序列，其内部各元素的重要性权重并非均等，而是通过可学习的关联度计算模块生成注意力权重矩阵。该权重矩阵量化了每个元素对当前处理位置的影响程度，使模型能够自主聚焦于关键特征区域。相较于传统序列处理方法，自注意力机制通过权重筛选显著提升了特征选择的针对性和计算效率^[24]。

在自注意力机制运行过程中，输入数据的序列元素都会生成相对应的查询向量(query)、键向量(key)、值向量(value)。通过计算查询向量和所有键向量的点积得到相似度分数，再经过归一化函数将相似度分数变成注意力权重，然后根据这些权重对值向量进行加权求和，得到包含重要信息的新向量。新向量包含了当前处理元素所需关注的潜在相关性信息，这些信息能对后续的数据处理效率和效果有所提升。

自注意力机制根据分析输入序列的不同特征来分配权重的大小，重要特征的内容会被分配更大的权重，相对应的其他内容的权重则较小，以加快模型运行。运行结构图如图2所示。

注意力状态的计算公式如式(1)~(4)所示。

$$S_{ii} = O \tanh(Ph_i + Qh_i + b) \quad (1)$$

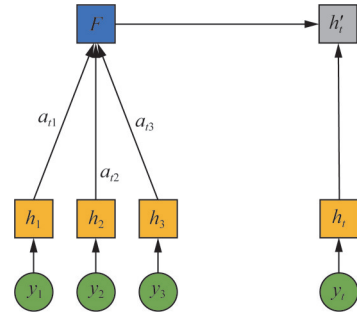


图2 单元结构

Fig. 2 Unit structure

$$a_{ti} = \frac{e^{S_{ti}}}{\sum_{k=i}^t e^{S_{tk}}} \quad (2)$$

$$F = \sum_{i=1}^t a_{ti} \times h_i \quad (3)$$

$$h'_i = f(F, h_i, y_i) \quad (4)$$

式中： y_i 为输入序列； h_i 为输入序列 y_i 的状态值；下标 t 为某具体时刻； a_{ti} 为历史负荷数据集 h_i 对输入序列 y_i 的权重值； h'_i 为最终的特征向量； F 为注意力权重值与状态值的乘积； O 、 P 、 Q 、 b 均为模型的学习参数^[24]。

1.2 图注意力网络 GAT

图神经网络(graph neural networks, GNN)是一种新兴的深度学习模型，专门用于分析和学习图结构数据。图结构数据由节点和连接这些节点的边组成，为了挖掘它们之间潜在的复杂关系，采用迭代节点及其邻居节点的信息来更新节点表示的方式^[25-27]。GNN的核心优势在于其能够捕捉节点间的复杂关系和依赖性。

在原有的GNN模型的基础上，引入自注意力机制构成图注意力网络(graph attention network, GAT)。本文在此引入了GATConv层，即图注意力卷积层，是图神经网络中的一个变体，它通过引入自注意力机制来学习节点间的重要性。在GATConv层中，每个节点会根据其邻居节点的特征动态地调整对邻居的关注度，这样可以更好地捕捉图中的结构信息。每个GATConv层可以有多个注意力头(heads)，这增加了模型的表达能力。通过堆叠多个GATConv层，模型可以学习节点特征的更高层次表示。在层之间使用跳跃连接，这是一种将输入直接与输出相加的机制，有助于信息在深层网络中的流动，尽可能避免出现梯度消失问题。局部时空信息传递图可以将当前时刻节点与其前一

时刻、后一个时刻的自身节点相连,以此构建出可以传递前中后 3 个时刻的时空信息的结构,结构如图 3 所示。

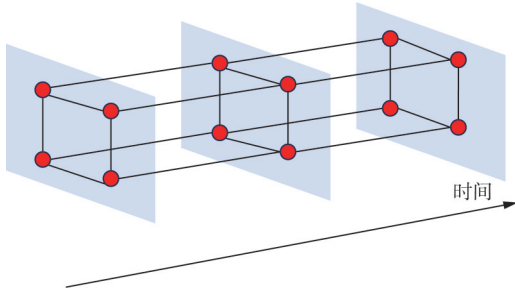


图 3 局部时空信息传递结构图

Fig. 3 Diagram of local spatiotemporal information transmission structure

通过自注意力机制为每个邻居节点分配不同的权重,构成邻接矩阵,如图 4 所示。这些权重表明了邻居节点对当前节点的重要性,增强了对图结构数据中不同的节点间关系的处理能力。GAT 模型通过学习这些权重,能够更加灵活地处理图中的节点关系。适用于那些节点间关系具有不同重要性的情况,允许模型动态地学习这些关系的重要性。

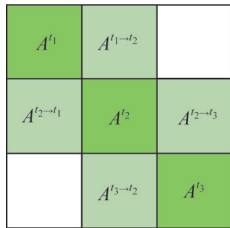


图 4 邻接矩阵

Fig. 4 Adjacency matrix

1.3 长短期记忆模型 LSTM

长短期记忆网络(long short-term memory, LSTM)是一种循环神经网络(RNN)的变体,专门设计用于捕捉时间序列数据中的长期依赖关系。由于传统 RNN 网络在长序列数据集上运行时,常出现梯度相关问题,因此 LSTM 网络通过引入遗忘门、输入门和输出门的门控机制来获得更好的效果。LSTM 能够维持和更新细胞状态,可以很好地捕捉数据的时序特征,从而在时序预测领域表现出色。LSTM 结构单元如图 5 所示。

1) $t-1$ 时刻,遗忘门决定细胞状态 $C_{(t-1)}$ 中需要舍去的信息。遗忘门读取当下隐藏层状态 $h_{(t-1)}$ 和 t 时刻的输入序列,并输出 0-1 值,1 表示完全保留,

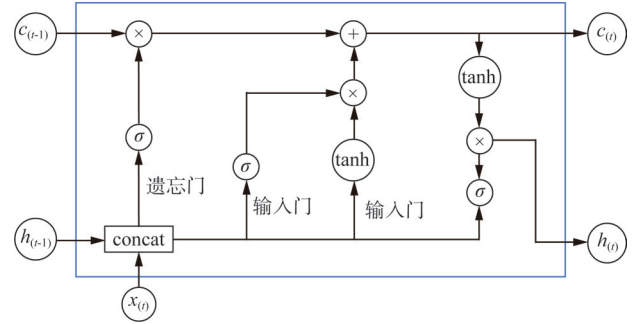


图 5 LSTM 结构示意图

Fig. 5 LSTM structure diagram

0 表示完全丢弃。

$$f_{(t)} = \sigma(W_f \cdot [h_{(t-1)}, x_{(t)}] + b_f) \quad (5)$$

式中: $f_{(t)}$ 为 t 时刻遗忘门状态; $\sigma(\cdot)$ 为激活函数; W_f 、 b_f 分别为遗忘门的权重与偏置^[28-30]; $x_{(t)}$ 为 t 时刻的输入序列。

2) t 时刻的输入序列 $x_{(t)}$ 传入到输入门,由输入门决定需要保存的信息。由 tanh 层生成该时刻记忆单元的临时状态 $\tilde{C}_{(t)}$,再更新得到新细胞状态 $C_{(t)}$ 。

$$i_{(t)} = \sigma(W_i \cdot [h_{(t-1)}, x_{(t)}] + b_i) \quad (6)$$

$$\tilde{C}_{(t)} = \tanh(W_c \cdot [h_{(t-1)}, x_{(t)}] + b_c) \quad (7)$$

$$C_{(t)} = f_{(t)} \times C_{(t-1)} + i_{(t)} \times \tilde{C}_{(t)} \quad (8)$$

式中: $i_{(t)}$ 为 t 时刻输入门状态; $h_{(t-1)}$ 为 $t-1$ 时刻的隐藏层状态; $x_{(t)}$ 为 t 时刻的输入序列; $C_{(t)}$ 为新细胞状态; W_i 、 b_i 分别为输入门的权重和偏置; W_c 、 b_c 分别为新细胞状态的权重和偏置; $\tanh(\cdot)$ 为激活函数^[28-30]。

3) t 时刻,细胞状态传递到输出门,由输出门决定需要输出的信息。输出序列为 $h_{(t)}$,同时也是下一个隐藏层中遗忘门的输入序列。

$$o_{(t)} = \sigma(W_o \cdot [h_{(t-1)}, x_{(t)}] + b_o) \quad (9)$$

$$h_{(t)} = o_{(t)} \times \tanh C_{(t)} \quad (10)$$

式中: $o_{(t)}$ 为 t 时刻输出门状态; W_o 、 b_o 分别为输出门的权重和偏置; $h_{(t)}$ 为输出序列^[28-30]。

在本模型中, GATConv 层的输出被用作 LSTM 层的输入,这样可以将图结构信息和时间序列信息结合起来,用于时空预测任务。在整个模型最后通过一个全连接层(fully connected layer, FC)输出为最终的预测结果。

1.4 跳跃连接

跳跃连接是一种深度学习网络中用于解决梯度消失和梯度爆炸问题的技术:它通过在网络中添加

直接连接,允许信号绕过一些层直接传递。这种结构的核心思想是让网络学习输入和输出之间的残差,即网络的输出是输入加上一个残差函数的输出,这样即使网络很深,梯度也能通过这些连接直接反向传播回早期层,避免了深层网络训练中的典型问题。在残差网络中,每个残差块包含若干卷积层和一个跳跃连接,如果输入和输出的维度不一致,会通过 1×1 的卷积来调整维度以匹配,然后进行相加操作。这样的设计不仅有助于梯度的传递,还有助于特征的融合,使得深层网络能够利用从浅层网络直接传递过来的信息,捕捉到更丰富的特征,从而提高了网络的性能和训练效率。跳跃连接的运行原理如图6所示。

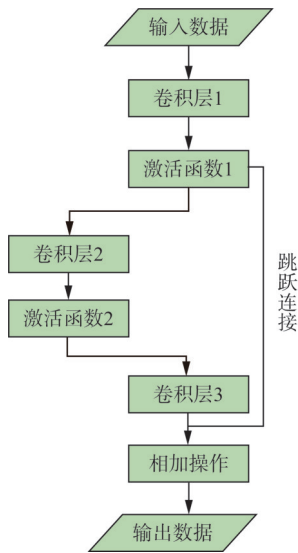


图6 跳跃连接运行原理图

Fig. 6 Skip connection operation principle diagram

输入数据进入第一个卷积层,通过一个激活函数。经过激活函数的数据进入第二个卷积层,再通过一个激活函数。最后,数据进入第三个卷积层。跳跃连接是指从第一个激活函数的输出直接跳到第三个卷积层之后的相加操作。第一个激活函数的输出矩阵(通过跳跃连接)和第三个卷积层的输出矩阵,通过逐个元素相加,形成最终的输出数据。

通过这样的方式,即使所用的网络结构较为复杂,所用信息也能通过这些跳跃连接直接传递,减少了信息丢失,使得深层网络的训练变得更加容易。

1.5 AG-LSTM模型结构

基于AG-LSTM算法模型的运行流程如图7

所示。

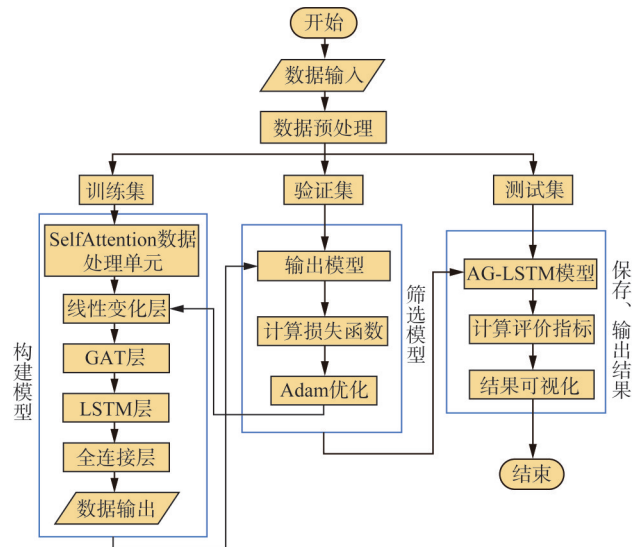


图7 模型整体运行流程图

Fig. 7 Overall model workflow diagram

首先,通过使用自注意力机制来编写self attention数据处理单元,通过此单元处理数据中每个节点的特征,将原始序列转换成矩阵形式以提升整体运行速率,在运算过程中根据使用需求调整单元尺寸并进行权值赋予,使模型能够关注不同部分的重要性,为后续图神经网络能更好地捕捉节点间复杂关系打下基础。通过这种方式,模型不仅考虑了节点自身的特征,还考虑了节点之间的相互作用,从而提高了模型的性能。

其次,构建AG-LSTM模型。历史负荷数据在经过数据处理单元处理后,节点被赋予了不同权重,再输入进GAT图注意力网络层,对节点的特征关系进行捕捉。在GAT层前添加一个线性变化层,用于将输入特征与GAT层的输出进行结合,形成跳跃连接,这种方式有助于防止数据在多GAT层的运行中损失重要信息,并且可以很好地防止梯度消失、梯度爆炸等问题的出现。本文中的GAT层主要构成为多个GATConv层,其使用自注意力机制来更新图数据中每个节点的特征表示。在模型运行之初,建立一个空模块列表来存储多个GATConv层,在首个图注意力层接收原始输入特征,并将其维度从输入维度映射至隐藏层维度。若该层采用多头自注意力机制,其输出特征维度将扩展至隐藏层维度乘以头数。随后的图注意力层进一步处理前一层的输出,通过特征维度的降维操作,

将特征维度重新压缩至隐藏层维度,以维持信息的紧凑性并为后续层提供处理后的特征。此外,模型中引入了跳跃连接,该连接通过一个线性变换层(输入线性)将原始输入特征与图注意力层的输出特征相结合。这一机制有助于在多层图注意力处理过程中保留原始信息,从而增强模型对输入数据的理解能力,并可以很好地防止梯度消失、梯度爆炸等问题的出现。

然后,经过图注意力层处理的数据被送入 LSTM 层,以捕捉其中的各种时序关系。LSTM 层接收来自 GAT 层的输出,由遗忘门决定是否有信息应从细胞状态中移除,并整合新的输入信息;由输入门负责决定新输入信息是否全部应被纳入细胞状态;输出门负责根据细胞状态和当前输入信息,形成输出状态。通过前向传播,LSTM 层接收来自前一时间步的隐藏状态和当前时间步的输入数据,以此更新其隐藏状态和细胞状态。在多轮次的训练过程中,通过反向传播和梯度下降方法,不断迭代 LSTM 层的各个参数,从而提高模型的性能。过程如图 8 所示。

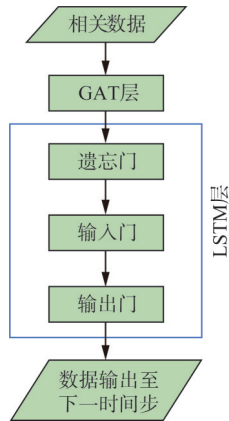


图 8 LSTM 层运行流程图

Fig. 8 LSTM layer workflow diagram

最后,LSTM 层的输出通过一个全连接层(FC)进行转换,以生成最终的预测结果。

2 数据处理

2.1 数据设置

算例所用数据集涵盖了 2011 年初至 2014 年底某地区 321 个用户的电力负荷^[31]。数据划分为训练集、验证集和测试集进行预测,比例为 6 : 2 : 2。本文为单步时间序列预测策略,即在 t 时刻对此前

时刻的历史负荷数据中提取特征,并通过 24×7 的时间窗口将这些数据重新构建为 AG-LSTM 负荷预测模型的输入数据。

2.2 数据归一化

经过归一化处理,可以加速损失函数的收敛速率并避免梯度的过大波动,同时也能确保计算结果的准确性。由于模型输入常常包括各种不同的单位制数值,因此为了减少这些差异对于预测效果的影响并且优化模型的效果与性能,必须执行归一化操作。本研究使用了 min-max 标准化的方法来实现。

$$X_i^* = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (11)$$

式中: X_i 为第 i 个原始数据; X_i^* 为 X_i 经过归一化后的值; $X_{\max}$ 和 $X_{\min}$ 分别为原始数据中的最大值和最小值。在得到归一化的结果后,为了准确反映预测结果准确性,需进行反归一化操作。

$$x^* = X_i^* (X_{\max} - X_{\min}) + X_{\min} \quad (12)$$

式中 x^* 为反归一化后的值。

2.3 模型评估指标

本文采用均方根误差 R_{MSE} (root mean square error, RMSE)、平均绝对百分误差 M_{APE} (mean absolute percentage error, MAPE) 和平均绝对误差 M_{AE} (mean absolute error, MAE) 3 个常用的评价指标来作为评价电力负荷预测模型的性能指标。计算公式如式(13)~(15)所示。

$$R_{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y(i) - y'(i))^2} \quad (13)$$

$$M_{\text{APE}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y(i) - y'(i)}{y(i)} \right| \times 100\% \quad (14)$$

$$M_{\text{AE}} = \frac{1}{n} \sum_{i=1}^n |y(i) - y'(i)| \quad (15)$$

式中: $y'(i)$ 为 i 时刻的预测值; $y(i)$ 为 i 时刻的真实值; n 为数据的数量。

3 模型算例分析

3.1 模型结构分析

对模型进行消融实验,以确定模型中各模块在运行中的有效性。经过 10 次测试后取平均值,并将结果可视化如图 9 与表 1。其中 LSTMonly 代表仅含有 lstm 结构; noGAT 代表 lstm 层和自注意力数据处理器组合模型,无图注意力网络层; noself 代表

图注意力网络层和lstm层的组合模型, 无自注意力数据处理器。

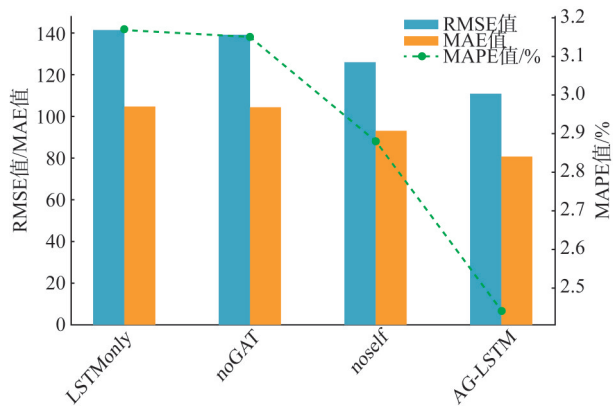


图9 不同模块评价指标对比

Fig. 9 Comparison of evaluation metrics for different modules

表 1 不同模块评价指标统计

Tab. 1 Statistical analysis of evaluation metrics for different modules

模型	$M_{APE}/\%$	R_{MSE}/kW	M_{AE}/kW
LSTMonly	3.17	141.42	104.89
noGAT	3.15	138.82	103.98
noself	2.88	129.91	95.66
AG-LSTM	2.44	111.13	81.01

从实验结果来看, 不论是只结合了自注意力机制、只结合了图学习层还是结合了自注意力机制与图学习层的预测性能相较于 LSTM 模型都有提高, 说明了图学习层和自注意力机制在长时间尺度下对时序预测模型提升的有效性, 同时也说明这些模块存在良好兼容的可能, 组合后对模型负荷预测的性能有提升。

同时, 为了验证自注意力数据处理器对运行效率的提升, 对无自注意力数据处理器的模型和 AG-LSTM 模型运行时间进行统计, 如表 2 所示。

表 2 运行时间统计

Tab. 2 Running time statistics

模型	运行轮次/次	单次平均耗时/s
noself	200	79.83
AG-LSTM	200	69.75

从统计结果可知, 自注意力数据处理器添加对于模型的运行效率有明显提升, 缩减了运行时间。

3.2 不同时间步长分析

为了寻找本模型效果最好的时间步长, 本文设计了 6 个不同的时间步长实验, 每次实验训练次数为 200 次, 针对每种时间步长经过 10 次重复实验取平均值, 采用 MAPE、RMSE、MAE 作为评价指标, 并将数据记录如表 3 所示。

从实验结果来看, 时间步长为 1 h 时表现最佳。随着预测时间步长的增大, 模型的各个评价指标数值呈现上升趋势, 当步长增加到 3 h 及大于 3 h 时, 模型性能骤降, 在 6 h 时对应的 MAPE 值仅为 13.16%, 此时表现最差, 说明本模型在预测时间步长较短时, 呈现出较好的效果。此优点也契合了本研究的目标, 针对历史用电负荷进行短期预测。以此为依据, 后续的其他实验也将以 1 h 时间步长进行。

表 3 不同时间步长评价指标统计

Tab. 3 Statistical analysis of evaluation metrics at different time steps

时间步长/h	$M_{APE}/\%$	R_{MSE}/kW	M_{AE}/kW
1	2.44	111.13	81.01
2	2.55	115.63	84.58
3	12.98	544.38	429.52
6	13.16	545.07	432.37
12	13.01	544.42	429.94
24	12.98	544.38	429.55

3.3 多种预测模型对比

为保证实验公平性, 本文将 CNN、KNN、Bilstm、RNN、AG-LSTM 模型所使用数据集保持一致, 评价指标选用 MAPE、RMSE、MAE。连续运行 10 次后取平均值, 分别预测出未来 7 个月的每小时负荷。其中 CNN 参数为: 卷积层 3 层, 通道数为 16、32、64, 卷积核大小为 3, 池化核大小为 2, 步长为 2; KNN 参数为: 邻居数量为 2 000; Bilstm 参数为: 隐藏层神经元数量为 2, 层数为 1 层, 批大小为 64, 学习率 0.001; RNN 参数为: 隐藏层神经元数量 64, 层数为 2 层, 批大小 64, 学习率 0.001。

图 10 为多种模型预测结果, 表 4 中统计了多种模型预测结果的评价指标。在预测时间步长为 1 h 的结果中, AG-LSTM 模型的 RMSE 值为 111.13 kW, 相较于 CNN 模型、KNN 模型、BiLSTM 模型、RNN 模型分别降低了 187.14 kW (62.74%)、

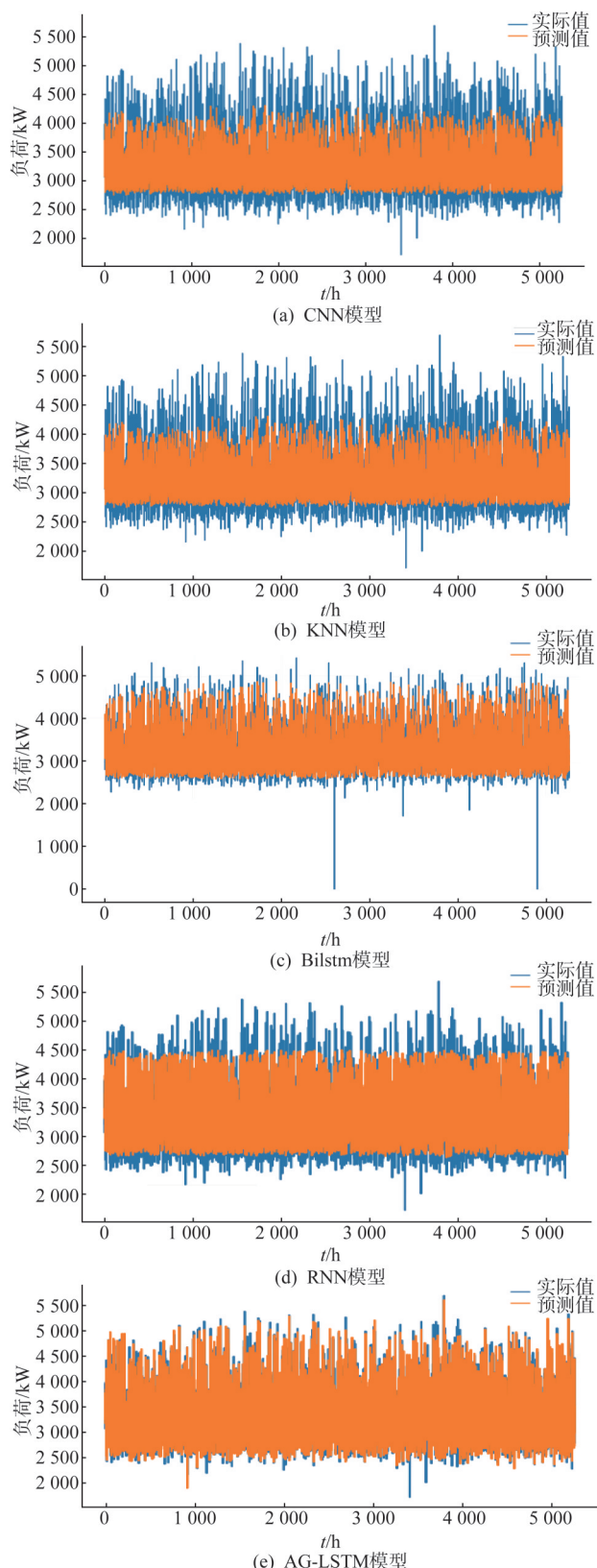


图 10 多种模型预测结果对比图

Fig. 10 Comparison chart of prediction results of multiple models

表 4 多种模型预测结果的评价指标统计

Tab. 4 Statistics of evaluation metrics of prediction results of multiple models

模型名称	$M_{APE}/\%$	R_{MSE}/kW	M_{AE}/kW
CNN	6.19	298.27	212.41
KNN	7.22	318.05	241.63
Bilstm	5.32	227.46	172.46
RNN	5.95	257.92	198.06
AG-LSTM	2.44	111.13	81.01

206.92 kW (65.06 %)、116.33 kW (51.14 %)、146.79 kW (56.91 %)。AG-LSTM 模型的 MAE 值为 81.01 kW，相较于 CNN 模型、KNN 模型、BiLSTM 模型、RNN 模型分别降低了 131.40 kW (61.86 %)、160.62 kW (66.47 %)、91.45 kW (53.06 %)、117.05 kW (59.10 %)。AG-LSTM 模型的 MAPE 值为 2.44 %，相较于 CNN 模型、KNN 模型、BiLSTM 模型、RNN 模型分别降低了 3.75 %、4.78 %、2.88 %、3.51 %。预测负荷结果显示，AG-LSTM 模型表现优异，准确度与其他成熟预测模型相比有明显提升。

4 结论

面向“双碳”战略目标优化和推广新型电力系统是国家重点需求。结合国家对工业领域的数字化、智能化改革以及当下工业领域用电需求的不断细化，对于新型电力系统的末端感知能力要求也更高了。本文提出了 AG-LSTM 算法模型，结合真实案例中的电力数据，以验证是否可以在新型电力系统中有所作用，通过实验验证分析得出以下结论。

1) 数据预处理阶段的创新。在原始数据的处理中，创新地结合了自注意力机制，引入了数据处理单元附加权值，其作用是发掘原始电力负荷数据中不同用户间潜在的关系，即发掘出原始数据中隐藏的节点，并将此关系信息传递给下一模块，使 AG-LSTM 模型更好地抓住未经过筛选的原始数据中具有相似用电习惯的用户，为后续图注意力网络模块的运行打了很好的基础。

2) 算法运行中的创新。经过大量实验验证了数据处理单元可以有效加快后续模块的运行。以图注意力网络和长短期记忆网络为核心，AG-LSTM 模型可以有效处理提取数据中的时空特征，并结合跳跃连接实现较少的信息丢失和突出关键时间点的信

息,提高了模型性能。

3)锁定关键参数的创新。通过进行不同预测时间步长的实验,验证了AG-LSTM模型在短期电力负荷预测(小时级)领域的可靠性与准确性,针对短期电力负荷预测(小时级)这一痛点问题有较好的表现。

4)通过进行多种模型的对比分析实验,与CNN模型、KNN模型、BiLSTM模型、RNN模型相比,在小时级预测中本文所提出的AG-LSTM模型性能最佳。未来研究将考虑电价政策等电力系统外的复杂影响因素,并充分结合我国大电网系统统一调度优势,进一步提升模型的可靠性与准确性。

参考文献

- [1] 康重庆,夏清,张伯明.电力系统负荷预测研究综述与发展方向的探讨[J].电力系统自动化,2004,(17):1-11.
KANG Chongqing, XIA Qing, ZHANG Boming. Review of power system load forecasting and its development [J]. Automation of Electric Power Systems, 2004, (17): 1-11.
- [2] 李丹,贺帅,颜伟,等.考虑动态时间锚点和典型特征约束的年均日负荷曲线预测[J].中国电力,2024,57(11):36-47.
LI Dan, HE Shuai, YAN Wei, et al. Annual daily average load curve prediction considering dynamic time anchors and typical feature constraints [J]. Electric Power, 2024, 57(11): 36-47.
- [3] YU Chunnam, MIROWSKI P, HO T K. A sparse coding approach to household electricity demand forecasting in smart grids[J]. IEEE Transactions on Smart Grid, 2017, 8(2): 738-748.
- [4] STEPHEN B, TANG Xiaoqing, HARVEY P R, et al. Incorporating practice theory in sub-profile models for short term aggregated residential load forecasting[J]. IEEE Transactions on Smart Grid, 2017,8(4): 1591-1598.
- [5] 李栓,刘莉,刘阳.趋势外推法在电力负荷预测中的应用[J].沈阳工程学院学报(自然科学版),2005(2):64-65.
LI Shuan, LIU Li, LIU Yang. Application of trends extrapolation on the power load forecast [J]. Journal of Shenyang Institute of Engineering (Natural Science), 2005 (2): 64-65.
- [6] 李小燕.基于灰色理论的电力负荷预测[D].武汉:华中科技大学,2007.
- [7] 程超.基于时间序列法和回归分析法的改进月售电量预测方法研究[D].重庆:重庆大学,2016.
- [8] AL A M A, HOQUE M A. Comparison of ARIMA and SVM for short-term load forecasting [C]//Annual Information Technology, Electromechanical Engineering and Microelectronics Conference, March 13-15, 2019, Jaipur, India. New York: IEEE, 2019:1-6.
- [9] 吴潇雨,和敬涵,张沛,等.基于灰色投影改进随机森林算法的电力系统短期负荷预测[J].电力系统自动化,2015,39(12):50-55.
WU Xiaoyu, HE Jinghan, ZHANG Pei, et al. Power system short-term load forecasting based on improved random forest with grey relation projection [J]. Automation of Electric Power Systems, 2015, 39(12): 50-55.
- [10] YIN Linfei, SUN Zhixiang, GAO Fang, et al. Deep forest regression for short-term load forecasting of power systems[J]. IEEE Access, 2020(8): 49090-49099.
- [11] 刘晟源,章天晗,林振智,等.数据赋能低压配电系统精益化运行的关键技术与算法[J].电力系统自动化,2023,47(3):187-199.
LIU Shengyuan, ZHANG Tianhan, LIN Zhenzhi, et al. Key technologies and algorithms of data empowerment for lean operation of low-voltage power distribution and consumption system [J]. Automation of Electric Power Systems, 2023, 47(3): 187-199.
- [12] 曾国治,魏子清,岳宝,等.基于CNN-RNN组合模型的办公建筑能耗预测[J].上海交通大学学报,2022,56(9):1256-1261.
ZENG Guozhi, WEI Ziqing, YUE Bao, et al. Energy consumption prediction of office buildings based on CNN-RNN combined model [J]. Journal of Shanghai Jiao Tong University, 2022,56(9): 1256-1261.
- [13] 杨子民,彭小圣,熊予涵,等.计及邻近风电场信息与CNN-BiLSTM的短期风电功率预测[J].南方电网技术,2023,17(2):47-56.
YANG Zimin, PENG Xiaosheng, XIONG Yuhan, et al. Short-term wind power prediction based on information in neighboring wind farms and CNN-BiLSTM [J]. Southern Power System Technology, 2023, 17(2): 47-56.
- [14] 张未,余成波,王士彬,等.基于VMD-LSTM-LightGBM的多特征短期电力负荷预测[J].南方电网技术,2023,17(2):74-81.
ZHANG Wei, YU Chengbo, WANG Shibin, et al. Multi-featured short-term power load forecasting based on VMD-LSTM-LightGBM [J]. Southern Power System Technology, 2023, 17(2): 74-81.
- [15] 陈振宇,刘金波,李晨,等.基于LSTM与XGBoost组合模型的超短期电力负荷预测[J].电网技术,2020,44(2):614-620.
CHEN Zhenyu, LIU Jinbo, LI Chen, et al. Ultra short-term power load forecasting based on combined LSTM-XGBoost Model [J]. Power System Technology, 2020, 44 (2) : 614-620.
- [16] 陆继翔,张琪培,杨志宏,等.基于CNN-LSTM混合神经网络模型的短期负荷预测方法[J].电力系统自动化,2019,43(8):131-137.
LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model [J]. Automation of Electric Power Systems, 2019,43(8): 131-137.
- [17] 刘文杰,刘禾,王英男,等.基于完整自适应噪声集成经验模态分解的LSTM-Attention网络短期电力负荷预测方法[J].电力建设,2022,43(2):98-108.
LIU Wenjie, LIU He, WANG Yingnan, et al. Short-term power load forecasting method based on CEEMDAN and LSTM-Attention network [J]. Electric Power Construction, 2022, 43(2): 98-108.
- [18] YAO Huaxiu, WU Fei, Ke Jintao, et al. Deep multi-view spatial-temporal network for taxi demand prediction [J]. Statistics, 2018(2): 1-9.
- [19] XIE Jiangjian, ZHONG Yujie, XIAO Tong, et al. A multi-information fusion model for short term load forecasting of an

- architectural complex considering spatio-temporal characteristics[J]. *Energy & Buildings*, 2022(277): 1 – 11.
- [20] WU Di, LIN Weixuan. Efficient residential electric load forecasting via transfer learning and graph neural networks [J]. *IEEE Transactions on Smart Grid*, 2023, 14(3): 2423 – 2431.
- [21] ARASTEHFAR S, MATINKIA M, JABBARPOUR M. Short-term residential load forecasting using graph convolutional recurrent neural networks [J]. *Engineering Applications of Artificial Intelligence*, 2022(116): 105358.1– 105358.10.
- [22] 董雷, 陈振平, 韩富佳, 等. 基于图卷积神经网络与 K-means 聚类的居民用户集群短期负荷预测[J]. *电网技术*, 2023, 47(10): 4291 – 4301.
- DONG Lei, CHEN Zhenping, HAN Fujia, et al. Short-term load forecasting of residential user groups based on graph convolutional neural network and K-means clustering [J]. *Power System Technology*, 2023, 47(10): 4291 – 4301.
- [23] 史含笑, 王雷春. 结合 LSTM 和自注意力机制的图卷积网络短期电力负荷预测[J]. *计算机应用*, 2024, 44(1): 311 – 317.
- SHI Hanxiao, WANG Leichun. Short-term power load forecasting bgraph convolutional network combining LSTM and self-attention mechanism [J]. *Journal of Computer Applications*, 2024, 44(1): 311 – 317.
- [24] ASHISH Vaswani, SHAZEER Noam, PARMAR Niki, et al. Attention is all you need[C]// *Advances in neural information processing systems* 30 (2017), December 4 – 9, 2017, Long Beach, America. Massachusetts: MIT Press, 2019: 1 – 6.
- [25] SU Zhong, ZHENG Guoqiang, HU Miaosen, et al. Short-term load forecasting of regional integrated energy system based on spatio-temporal convolutional graph neural network [J]. *Electric Power Systems Research*, 2024(232): 110427.1 – 110427.10.
- [26] BIAN Mengqing, CHU Xiaodong, YE Hua. Short-term load forecasting based on residual graph convolutional neural network [C] //2023 International Conference on Power System Technology (PowerCon), September 21 – 22, 2023, Jinan, China. New York: IEEE, 2023: 1 – 5.
- [27] NGUYEN K D Q, YANG C Q, VU V H, et al. A benchmarking evaluation of graph neural networks on traffic speed prediction [J]. *Lecture Notes in Computer Science*, 2022 (13725): 472 – 488.
- [28] 于伸庭. 基于长短期记忆网络和卷积神经网络(LSTM-CNN)的 PM2.5 浓度预测研究[D]. 上海: 上海交通大学, 2020.
- [29] 任建吉, 位慧慧, 邹卓霖, 等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测[J]. *电力系统保护与控制*, 2022, 50(8): 108 – 116.
- REN Jianji, WEI Huihui, ZOU Zhuolin, et al. Ultra-short-term power load forecasting based on CNN-BiLSTM-Attention [J]. *Power System Protection and Control*, 2022, 50(8): 108 – 116.
- [30] 庞传军, 张波, 余建明. 基于 LSTM 循环神经网络的短期电力负荷预测[J]. *电力工程技术*, 2021, 40(1): 175 – 180, 194.
- PANG Chuanjun, ZHANG Bo, YU Jianming. Short-term power load forecasting based on LSTM recurrent neural network [J]. *Electric Power Engineering Technology*, 2021, 40(1): 175 – 180, 194.
- [31] LAI guokun. Electricity consumption dataset [EB/OL]. [2017 – 04 – 09]. [https:// github. com/aiguokun/ultivariate-time-series-data](https://github.com/aiguokun/ultivariate-time-series-data).

收稿日期: 2025-02-26

作者简介:

刘锦涛(1999), 男, 硕士研究生, 从事电力大数据、电力负荷预测研究工作, grandmaster210123@163.com;

孙玉芹(1971), 女, 通信作者, 教授, 硕士生导师, 博士, 从事电力大数据、智能微电网研究工作, 2008000011@shiep.edu.cn;

郭子涛(2001), 男, 硕士研究生, 从事电力大数据研究工作。