

DOI: 10.19783/j.cnki.pspc.230392

## 基于自适应图注意力网络的短期用户负荷预测

黄冬梅<sup>1</sup>, 陈欢<sup>2</sup>, 王宁<sup>3</sup>, 吴志坚<sup>3</sup>, 胡伟<sup>4</sup>, 孙园<sup>5</sup>

(1. 上海电力大学电子与信息工程学院, 上海 201306; 2. 上海电力大学电气工程学院, 上海 200090;  
3. 国网江苏省电力有限公司苏州供电分公司, 江苏 苏州 215004; 4. 上海电力大学经济与  
管理学院, 上海 200090; 5. 上海电力大学数理学院, 上海 201306)

**摘要:**为提高短期用户负荷预测精度,提出了一种基于自适应图注意力网络(adaptive graph attention network, AGAT)的短期用户负荷预测模型。首先,针对用户负荷存在规模小、波动性强的问题,通过门控机制结合多个大小不同的扩张卷积核来构造时序特征提取层,从多个尺度上提取用户负荷的高维时序特征。同时,考虑到不同用户负荷间潜在的动态相关性,使用马氏距离构造动态图学习层,生成动态图邻接矩阵。然后,采用图注意力网络根据动态图邻接矩阵将用户负荷的高维时序特征进行信息汇聚。最后,通过全连接层输出用户负荷预测值。为验证AGAT模型的有效性,采用UCI电力负荷数据集进行预测实验,分别与多种基线模型比较。实验结果表明,所提模型预测指标优于各基线模型,有助于提高短期用户负荷预测精度。

**关键词:**短期用户负荷预测; 自适应图注意力网络; 时序特征提取; 动态图学习; 图神经网络

### Short-term user load prediction based on an adaptive graph attention network

HUANG Dongmei<sup>1</sup>, CHEN Huan<sup>2</sup>, WANG Ning<sup>3</sup>, WU Zhijian<sup>3</sup>, HU Wei<sup>4</sup>, SUN Yuan<sup>5</sup>

(1. College of Electronic and Information Engineering, Shanghai University of Electric Power, Shanghai 201306, China;  
2. College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China; 3. Suzhou Power  
Supply Branch, State Grid Jiangsu Electric Power Co., Ltd, Suzhou 215004, China; 4. College of Economics and  
Management, Shanghai University of Electric Power, Shanghai 200090, China; 5. College of Mathematics  
and Physics, Shanghai University of Electric Power, Shanghai 201306, China)

**Abstract:** To improve the accuracy of short-term user load prediction, a model based on an adaptive graph attention network (AGAT) is proposed. First, to solve the problem of small scale and strong volatility of user load, a gating mechanism is used to construct a time sequence feature extraction layer combined with multiple expanded convolution kernels of different sizes, and the high-dimensional time sequence features of user load are extracted at multiple scales. Considering the potential dynamic correlation between different user loads, Mahalanobis distance is used to construct the dynamic graph learning layer and generate the dynamic graph adjacency matrix. Then the graph attention network is used to gather the information of the high dimensional time sequence features of the user load according to the dynamic graph adjacency matrix. Finally, the predicted user load is output through the fully connected layer. To verify the validity of the AGAT model, the UCI power load dataset is used for prediction experiments. The experimental results show that the prediction indices of the proposed model are better than those of various baseline models. This is helpful for the improvement of the accuracy of short-term user load prediction.

This work is supported by the National Social Science Foundation of China (No. 19BGL003).

**Key words:** short-term user load prediction; adaptive graph attention network; time sequence feature extraction; dynamic graph learning; graph neural network

## 0 引言

随着“双碳”目标的提出,电力体制改革不断

基金项目: 国家社会科学基金项目资助(19BGL003); 上海市科委地方院校能力建设项目资助(20020500700)

深入,电力市场更加多元化,新兴的售电公司与电力用户共同构成了电力零售市场<sup>[1-3]</sup>。掌握用户用电规律,精准预测用户用电负荷,为用户定制低碳科学的用电方案已成为售电公司的核心竞争力之一<sup>[4-5]</sup>。但是,传统的系统级负荷预测技术无法支撑售电公司应对偏差电量考核,用户级负荷预测技术的重要

性日益彰显<sup>[6-9]</sup>。

目前, 针对波动相对平滑、不确定性较弱的系统级负荷预测, 国内外学者基于统计学模型<sup>[10-14]</sup>和人工智能技术<sup>[15-19]</sup>已经取得了丰富的研究成果, 现有的负荷预测模型通常能够获得较高的预测精度。并且, 随着深度学习的不断发展, 深度学习技术以其对于大数据集强大的非线性拟合能力, 在负荷预测领域表现越来越突出。然而, 受限于用户级负荷随机性大、波动性强等特点, 用户级负荷预测的精度仍难以保证, 其仍面临着巨大的挑战。现有的针对用户级负荷预测的研究通常以系统级负荷预测方法为基础, 结合不同的数据挖掘方法开展探索性研究。文献[20-21]分别将集成经验模态分解和变分模态分解与深度学习结合, 提出了相应的用户负荷预测模型。文献[22]借助互信息与深度学习理论, 提出了一种基于最大相关最小冗余和长短期记忆网络的用户负荷短期预测模型。

上述用户负荷预测方法均只注重于挖掘单个用户用电负荷序列的时间相关性, 往往忽略了多用户用电行为之间存在的潜在关联。图神经网络作为一种新兴的深度学习算法, 能够有效学习图结构数据, 具备同时挖掘节点内的特征信息与节点间的相关性信息的优越能力, 在多个领域表现出了良好的预测性能, 也吸引了部分学者尝试将其应用于用户负荷预测领域。文献[23]提出了一种基于 K-means 聚类 and 自适应时空同步图卷积神经网络的居民用户集群负荷预测方法, 能够挖掘并利用不同居民用户用电负荷之间的时空相关性。文献[24]提出了一种新的图卷积循环神经网络, 结合图卷积网络和长短期记忆网络, 能够同时从具有相似消费模式的用户中提取空间和时间信息。文献[25]提出了一种图卷积网络和长短期记忆网络相结合的多信息融合模型, 并引入时空注意力模块提取负荷的动态时空相关性。但现有的基于图神经网络的短期用户负荷预测研究均只考虑用户用电行为之间的静态相关性, 未考虑到用户用电行为之间的相关性并不是一成不变的, 通常会随着时间推移而发生变化。单一的静态图将难以准确描述各用户用电行为之间的动态相关性。

因此, 针对上述问题, 为充分挖掘用户负荷的时序特征以及不同用户负荷间潜在的动态相关性, 本文提出了一种基于自适应图注意力网络(adaptive graph attention network, AGAT)的短期用户负荷预测方法。首先, 通过时序特征提取层提取所有用户负荷的时序特征。其次, 对于获得的负荷时序特征,

通过动态图学习层, 构建出自适应的动态图邻接矩阵。然后, 使用图注意力网络进行信息汇聚。最后使用全连接层进行输出映射, 预测接下来指定时刻的各用户的负荷值。基于真实数据集的实验结果表明, 相比于现有的各种基线模型, 所提模型能够有效提升用户短期负荷预测的精度。

## 1 图注意力网络

注意力机制的基本思想是基于现有的数据挖掘其潜在的相关性, 并以此突出所需的重要特征, 从而实现忽略无关信息而关注重点信息<sup>[26-28]</sup>。图注意力网络能够计算各中心节点与邻域节点的相似性, 再通过信息聚合, 实现对相似性越高的信息赋予越高的权值, 体现出各个邻域节点对中心节点影响的差异性<sup>[29-31]</sup>。

对于单个图注意力层的计算, 主要包括权重系数求取和特征聚合两部分。

输入为网络节点的特征向量集合, 如式(1)所示。

$$V = \{v_1, v_2, \dots, v_i, \dots, v_N\} \quad (1)$$

式中:  $v_i$  为节点  $i$  的特征向量;  $N$  为节点个数。

对于节点  $i$ , 可根据注意力机制计算出其与邻域节点的相似性, 通过 softmax 归一化后, 权重系数可表示为

$$a_{ij} = \frac{\exp(\text{LeakyReLU}(s(Kv_i \parallel Kv_j)))}{\sum_{j \in \mathcal{N}_i} \exp(\text{LeakyReLU}(s(Kv_i \parallel Kv_j)))} \quad (2)$$

式中:  $\mathcal{N}_i$  为节点  $i$  的邻域节点集合; LeakyReLU 为非线性激活函数;  $s$  为节点相似性计算函数, 本文采用全连接层进行计算;  $K$  为节点特征变换的权重参数矩阵;  $\parallel$  为拼接操作。通过进行加权求和, 节点  $i$  的新特征向量  $v'_i$  可表示为

$$v'_i = \sigma \left( \sum_{j \in \mathcal{N}_i} a_{ij} Kv_j \right) \quad (3)$$

式中,  $\sigma$  为 sigmoid 激活函数。

多头注意力机制有助于增强注意力层表达能力的鲁棒性, 分别使用多组相互独立的注意力机制对式(3)进行扩展, 并将得到的输出结果进行拼接, 可表示为式(4)。

$$v'_i = \parallel_{h=1}^H \sigma \left( \sum_{j \in \mathcal{N}_i} \alpha_{ij}^{(h)} K^{(h)} v_j \right) \quad (4)$$

式中:  $H$  为注意力机制的头数;  $\alpha_{ij}^{(h)}$  为第  $h$  组注意力机制计算的权重系数;  $K^{(h)}$  为对应的权重参数矩阵。

计算流程如图 1 所示, 其中注意力头数为 3。

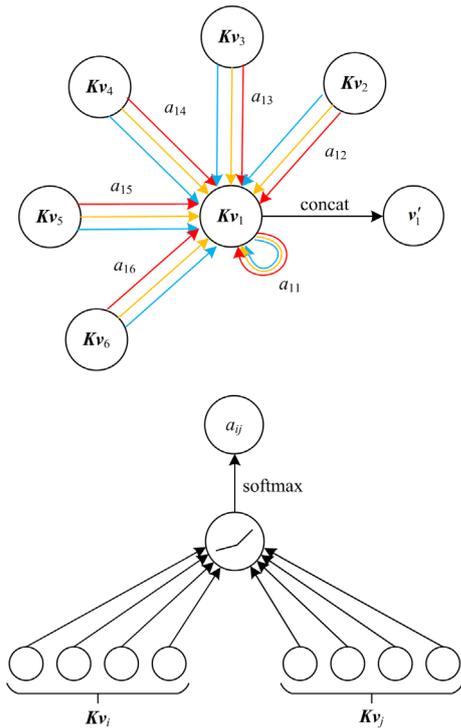


图 1 多头图注意力计算流程

Fig. 1 Multi-head graph attention calculation process

## 2 基于自适应图注意力网络的短期用户负荷预测模型

### 2.1 时序特征提取层

本节构造了一种采用结合门控线性单元和扩张卷积、适用于电力负荷的时序特征提取层(time sequence feature extraction layer, TFE)，以捕获负荷数据的时序特征。如图 2 所示，每个时序特征提取层有一个门控线性单元(gated linear unit, GLU)，每个 GLU 由两组组合扩张卷积(combinatorial dilated convolution, CDC)组成。

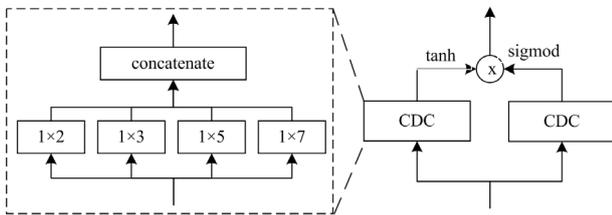


图 2 时序特征提取层

Fig. 2 Time sequence feature extraction layer

扩张卷积在标准卷积核中注入空洞以增大感受野，能够以较少的参数和非递归的方式处理长时间序列，有利于并行运算和缓解梯度爆炸的问题<sup>[32]</sup>。

一般情况下，扩张卷积的感受野大小公式为

$$L = d * (k - 1) + 1 \quad (5)$$

式中： $d$  为决定跳跃距离的扩展因子，通常以 2 的指数次幂增长； $k$  为原始卷积核大小。

与此同时，选择合适的卷积核大小对于卷积网络来说至关重要。卷积核过大可能无法捕获输入序列的短期特征，过小则可能难以发现长期特征。在图像处理中，inception 是一种广泛使用的多尺度卷积核策略，即为使用多个不同大小的卷积核并将其输出连接起来输出到下一层。本文构造了由 4 种不同尺寸的卷积核组成的组合扩张卷积层，其大小分别为  $1 \times 2$ 、 $1 \times 3$ 、 $1 \times 5$  和  $1 \times 7$ ，使得网络能获得丰富的感受野。

门控线性单元是一种简单的门控机制<sup>[33]</sup>，在本文中负责对组合扩张卷积层进行选择输出，包含一个输出门，给定输入  $X$ ，输出为

$$h = g(f_1(X) \cdot W_1 + b) \odot \sigma(f_2(X) \cdot W_2 + c) \quad (6)$$

式中： $W_1$ 、 $W_2$ 、 $b$  和  $c$  为可学习的参数矩阵； $f_1$ 、 $f_2$  为组合扩张卷积； $\odot$  为哈达玛积； $g$  为负责输出的 tanh 函数。

### 2.2 动态图学习层

现有的基于图神经网络的短期用户负荷预测研究通常使用皮尔逊系数等方法生成用户相关性矩阵作为静态图输入，只考虑了不同用户负荷间整体的静态相关性。然而，不同用户负荷间的相关性往往会随时间的变化而波动，使得单一的静态图邻接矩阵难以准确确定各节点在不同时间段上的邻域节点。

以两个用户某一年的真实电力负荷数据为例，如图 3 所示。可以看出，两个用户在工作日具有较高的相关性，但在休息日相关性较低。经过计算，两个用户负荷这一年的皮尔逊系数为 0.60，属于中等相关，但工作日和休息日的皮尔逊系数分别为 0.81(强相关)和 0.28(弱相关)，说明单一的静态相关系数无法准确描述负荷用户间的动态相关性。同时，尽管通过实时计算皮尔逊系数也能得到一个动态图邻接矩阵，但对于高随机性的用户负荷，其适用性难以保证。因此，本文提出一个用于捕获时间序列数据之间隐藏关系的动态图学习层(dynamic graph learning layer, DGL)，自适应地学习生成合适的动态图邻接矩阵。

现有的图构造方法通常使用距离度量来衡量节点之间的相似性。距离度量中最常见的是欧氏距离。但对于图结构数据，欧式距离并不是一个很好的度量指标，需要一个能随任务和特征进行自适应的距

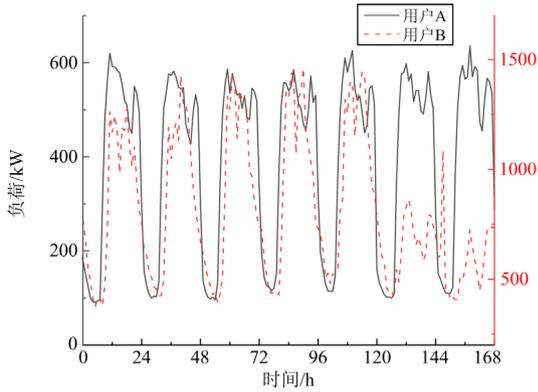


图3 两个电力用户一周的真实电力负荷

Fig. 3 Actual load of two power users in a given week

离度量。马氏距离是一种基于协方差的距离度量,它修正了传统欧式距离中维度尺度不一致且相关的问题。马氏距离的公式可表示为

$$D(\mathbf{x}_i^t, \mathbf{x}_j^t) = \sqrt{(\mathbf{x}_i^t - \mathbf{x}_j^t)^T \mathbf{M} (\mathbf{x}_i^t - \mathbf{x}_j^t)} \quad (7)$$

式中:  $\mathbf{x}_i^t$  和  $\mathbf{x}_j^t$  为时序特征提取层输出的  $i, j$  节点  $t$  时刻时序特征变量,  $i, j \in 1, 2, \dots, N$ ;  $\mathbf{M}$  为协方差矩阵。当协方差矩阵为单位阵时,则代表各维度独立同分布,此时马氏距离就变回了欧氏距离。本文将  $\mathbf{M}$  设置为动态图学习层中的一个可训练的权重参数。然后,通过马氏距离求取相关性。

$$G_i(\mathbf{x}_i^t, \mathbf{x}_j^t) = \exp\left(-D(\mathbf{x}_i^t, \mathbf{x}_j^t) / (\sqrt{2}\sigma_D)\right) \quad (8)$$

$$\mathbf{G}'_i[i, :] = \text{softmax}(G_i[i, :]) \quad (9)$$

式中:  $G_i(\mathbf{x}_i^t, \mathbf{x}_j^t)$  为  $t$  时刻  $\mathbf{x}_i^t$  和  $\mathbf{x}_j^t$  的相关性;  $\sigma_D$  为标准差;  $\mathbf{G}'$  为各节点间的归一化相关性矩阵。最后,对  $\mathbf{G}'$  进行二值化处理。

$$A_i[i, j] = \begin{cases} 1 & G'_i[i, j] \geq \tau \text{ or } i = j \\ 0 & G'_i[i, j] < \tau \end{cases} \quad (10)$$

式中:  $A_i$  为  $t$  时刻的动态图邻接矩阵;  $\tau$  为相关性阈值,控制着各节点的邻域节点个数,通过多次对比分析,本文取  $\tau = \frac{1}{N-1}$ 。

### 2.3 模型结构

本文将结合了时序特征提取和动态图学习的图注意力网络命名为 AGAT,其流程如图 4 所示。

首先,对输入的负荷数据使用  $1 \times 1$  卷积进行嵌入。然后,构造堆叠时序特征提取层,以此提取所有用户负荷的时序特征。需要指出的是,为避免时序特征提取层堆叠过深可能导致的退化问题,本文使用跳跃连接将每层时序特征提取层的输出进行拼接,并输入至动态图学习层,以此构建出自适应的

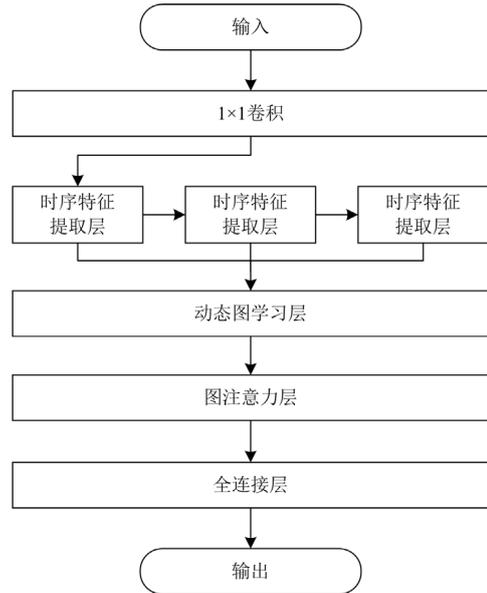


图4 AGAT 流程图

Fig. 4 Flowchart of the AGAT

动态图邻接矩阵。然后,使用图注意力网络进行信息汇聚。最后使用全连接层(fully connected layer, FC)进行输出映射,预测下一时刻各用户的负荷值。

## 3 算例分析

### 3.1 数据集描述及设置

算例所用数据集为一个处理过后的UCI公开电力负荷数据集<sup>[34]</sup>。该数据集涵盖了2011年初至2014年底葡萄牙321个用户的电力负荷,采样周期为1h。数据集以3:1:1的比例划分出训练集、验证集和测试集。

在负荷预测过程中,不同的特征数据量纲不一定相同。采用最大最小归一化对所有数据进行处理,避免量纲对预测结果的干扰,计算过程如式(11)所示。

$$X'_t = \frac{X_t - X_{\min}}{X_{\max} - X_{\min}} \quad (11)$$

式中:  $X_t$  为时间序列  $t$  时刻的值;  $X_{\max}$ 、 $X_{\min}$  为时间序列的最大值和最小值。

算例模型采用单步预测策略,使用前7天的数据作为输入来预测下一时刻的负荷。在模型参数设置方面,通过对预测精度和预测效率的综合考虑和多次相关对比分析,算例模型参数设置如下:时序特征提取层层数设为3,卷积核数量设为16,图注意力层层数设为1,注意力头数设为8, batch size 设为64, dropout 设为0.2,训练轮数为100,学习

率设为 0.001, 使用 Adam 算法作为网络优化器, 损失函数为 MSE。

### 3.2 评估指标

本文采用平均绝对值误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)与平均百分数误差(mean absolute percentage error, MAPE)这 3 个在短期电力负荷预测中常被使用的评价指标对模型性能进行评估, 各指标计算方法如式(12)~式(14)所示。

$$M_{AE} = \frac{1}{T} \sum_{t=1}^T |\hat{y}_t - y_t| \quad (12)$$

$$R_{MSE} = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{T}} \quad (13)$$

$$M_{APE} = \frac{1}{T} \sum_{t=1}^T \left| \frac{\hat{y}_t - y_t}{y_t} \right| \times 100\% \quad (14)$$

式中:  $T$  为时间序列长度;  $y_t$  为第  $t$  时刻的实际值;  $\hat{y}_t$  为第  $t$  时刻的预测值。  $M_{AE}$  反映了绝对误差的平均值;  $R_{MSE}$  是预测值和实际观测之间平方差均值的平方根, 较  $M_{AE}$  对异常值更敏感; 而  $M_{APE}$  可以理解为  $M_{AE}$  的归一化结果。这 3 个评价指标都是值越小代表预测精度越高、效果越好。

由于所用数据集共包含 321 个电力用户, 若只考虑单个用户负荷的预测评估指标, 将难以准确评估模型预测效果。因此本文将从所有用户负荷相加得到的总负荷的评估指标和所有用户负荷评估指标的平均值两个角度对实验结果进行评估。

### 3.3 多种预测模型对比

为验证所提模型预测效果, 设置了 6 种基线模型作为对比, 分别为: 朴素预测法( naive forecast, NF); 差分自回归移动平均模型( autoregressive integrated moving average model, ARIMA); 卷积长短期记忆神经网络( convolutional neural networks and long-short term memory networks, CNN-LSTM); 针对时间序列改进的结合 CNN 和 LSTM 的长短期时序网络( long-and short-term time-series network, LSTNet)<sup>[35]</sup>; 时空图卷积网络( spatio-temporal graph convolutional networks, STGCN)<sup>[36]</sup>; 多元时间序列图神经网络( multivariate time-series graph neural networks, MTGNN)<sup>[37]</sup>。其中前 4 种基线模型在负荷预测领域已被广泛应用, 模型预测性能已被验证。而 STGCN 和 MTGNN 作为新兴的图神经网络模型, 在交通预

测、知识图谱等领域表现良好, 可为负荷预测提供参考。

表 1 和表 2 分别为不同模型预测的总负荷的评估指标和所有用户负荷的评估指标的均值。图 5 和图 6 分别为不同模型预测的总负荷某日的预测值对比图和随机选取的某用户负荷某日的预测值对比图。图 7 为不同模型的用户负荷归一化预测值的  $R_{MSE}$  箱线图。

表 1 不同模型预测的总负荷的评估指标

Table 1 Evaluation indicators of total load predicted by different models

模型类型	$M_{AE}/kW$	$R_{MSE}/kW$	$M_{APE}/\%$
NF	30 053.84	47 568.06	3.76
ARIMA	21 556.11	32 196.81	2.52
CNN-LSTM	15 065.55	24 261.27	1.77
LSTNet	14 288.44	23 055.20	1.68
STGCN	12 083.21	16 854.01	1.52
MTGNN	11 192.69	17 910.11	1.46
AGAT	10 155.78	15 805.35	1.33

表 2 不同模型预测所有用户负荷的评估指标均值

Table 2 Average of evaluation indicators of all user load predicted by different models

模型类型	$M_{AE}/kW$	$R_{MSE}/kW$	$M_{APE}/\%$
NF	266.52	432.95	11.76
ARIMA	143.84	206.35	7.62
CNN-LSTM	124.68	187.04	7.10
LSTNet	117.63	183.43	6.94
STGCN	110.69	173.39	6.68
MTGNN	110.13	172.72	6.65
AGAT	105.42	167.49	6.35

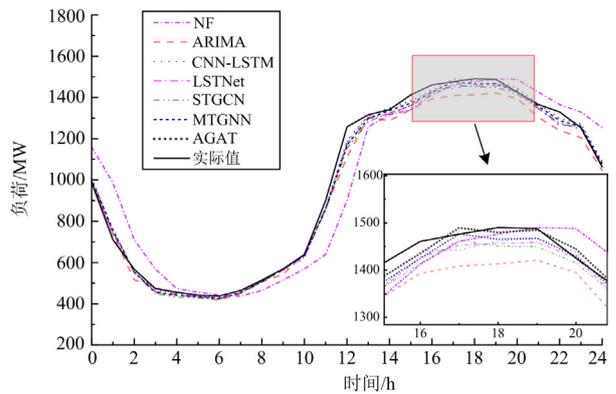


图 5 不同模型的总负荷预测值对比图

Fig. 5 Comparison of total load predicted values of different models

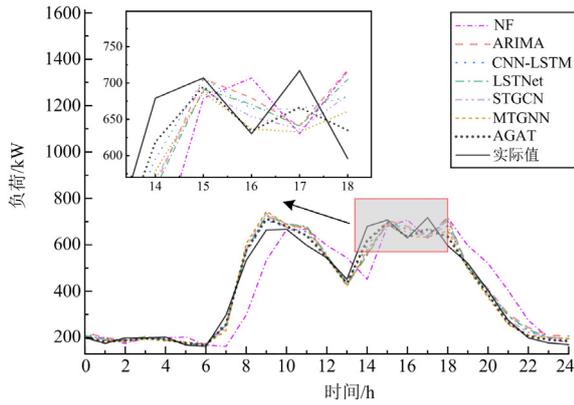


图 6 不同模型的用户负荷预测值对比图

Fig. 6 Comparison of user load predicted values of different models

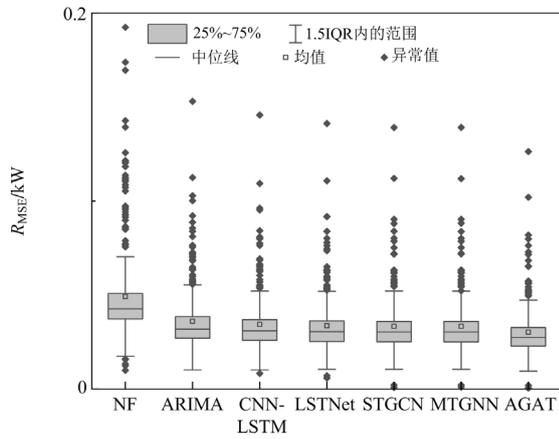


图 7 不同模型的用户负荷预测值  $R_{MSE}$  箱线图

Fig. 7  $R_{MSE}$  boxplot of user load predicted for different models

通过实验结果可以看出, 本文所提模型相比其他基线模型表现良好。传统的负荷预测统计模型如NF和ARIMA模型, 难以处理高波动性的用户负荷数据, 预测结果相对较差。而CNN-LSTM虽然结合了CNN和LSTM来捕捉用户负荷数据的时间依赖性, 但效果仍并不理想; 而LSTNet在CNN和LSTM的基础上进行了改进, 并使用了Skip-RNN以捕捉更长期的信息, 并充分利用序列的周期特性, 但是在捕捉空间依赖性方面还存在不足; STGCN通过时域卷积和空域卷积来挖掘多维时间序列的时空依赖性, 预测精度高于LSTNet, 但它依赖于用户间的图邻接矩阵, 最优图难以确定; MTGNN考虑了时间序列图结构的生成, 自适应地从数据中抽取稀疏图邻接矩阵, 预测结果优于STGCN, 但未考虑不同时间序列间的相关性可能会随时间改变而改变, 不能很好地捕获时空相关性, 预测结果仍然稍逊于本文所提模型。

### 3.4 网络设计分析

为探究模型各网络层对预测精度的影响, 进行相关消融实验, 结果如图8和图9所示。其中DGL-GAT和TFE-GAT代表动态图学习层和时序特征提取层分别与图注意力层的组合模型, 各模型超参数设置相同。

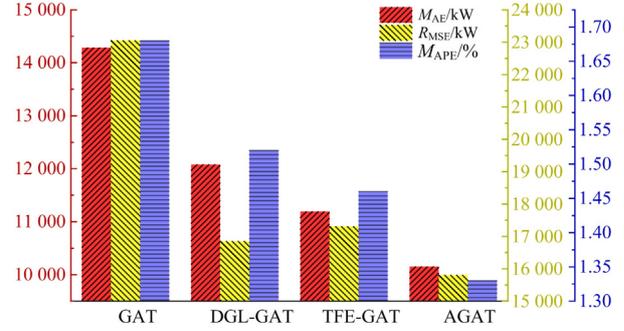


图 8 不同模型预测总负荷的评估指标

Fig. 8 Evaluation indicators of total load predicted by different models

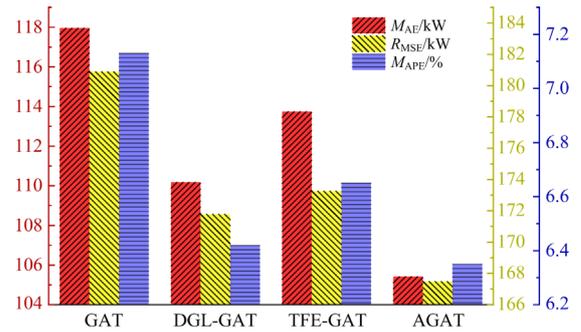


图 9 不同模型预测所有用户负荷的评估指标均值

Fig. 9 Average of the evaluation indicators of all user load predicted by different models

从实验结果可以看出, DGL-GAT和TFE-GAT的预测性能均优于单独的GAT模型, 而结合了时序特征提取层和动态图学习层的AGAT模型的3个预测指标均低于其他模型, 说明了时序特征提取层和动态图学习层的有效性, 且二者间具有良好的兼容性, 组合使用能进一步提高短期用户负荷预测精度。

为进一步验证动态图学习层的有效性, 选取了3个用户两周的真实负荷, 如图10所示。并分别展示了使用皮尔逊系数以当前时刻前一天内的数据计算得到其动态相关性的绝对值曲线和经过归一化处理的AGAT模型预测过程中的相应动态相关性曲线, 如图11所示。

从用户A-B的相关性图可以看出, 动态图学习层输出的动态相关性曲线相较于皮尔逊系数更为平

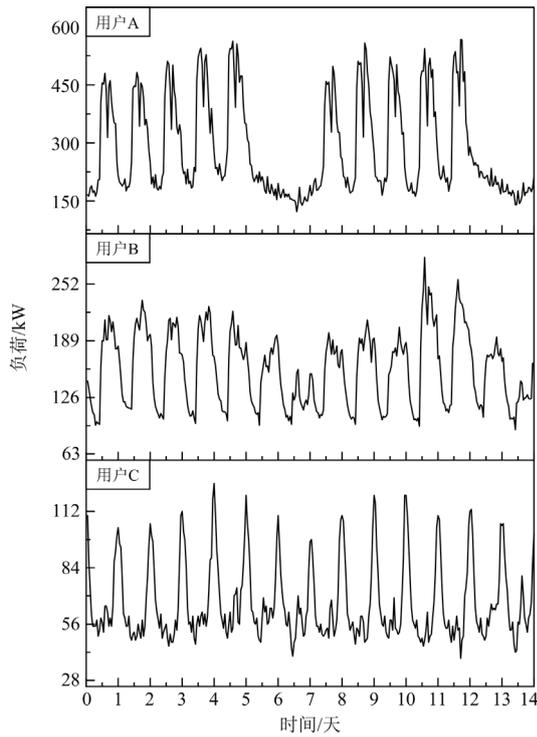


图 10 3 个电力用户两周的负荷

Fig. 10 Load for three power users over two weeks

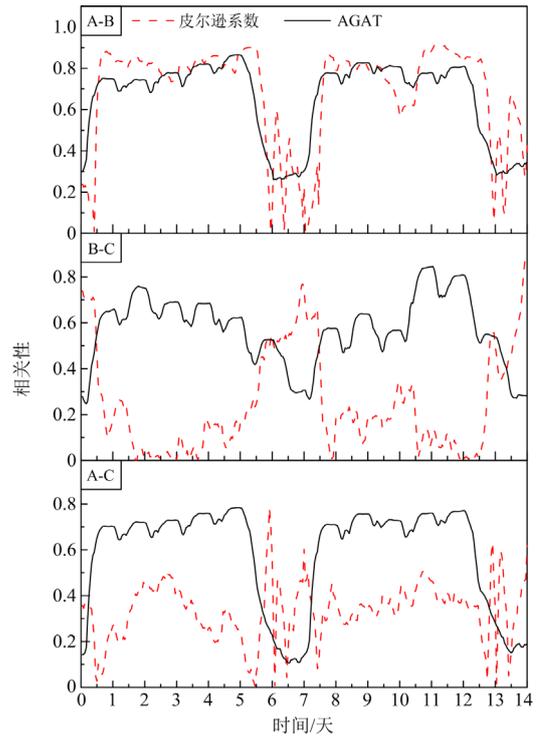


图 11 电力用户负荷数据的动态相关性曲线

Fig. 11 Dynamic correlation curve of power user load data

滑，时间滞后性较低。同时，由于用户 C 的峰值时间较用户 A、B 略微靠后，且保持时间较短，使得皮尔逊系数计算得到的用户 A-C、B-C 相关性曲线准确度相对较差。而动态图学习层输出的动态相关性曲线则更为合理，验证了动态图学习层的优越性。

### 3.5 不同预测时间对比

为了进一步验证模型对于不同预测时间尺度的预测效果，挑选了 LSTNet、STGCN、MTGNN 这 3 个预测效果较好的模型与 AGAT 模型使用相同的输入数据分别进行了不同预测时间尺度下的用户负荷预测实验，结果如表 3 和表 4 所示。

表 3 多种模型在不同预测时间尺度上预测的总负荷的评估指标

Table 3 Evaluation index of the total load predicted by multiple models on different forecast time scales

模型类型	评估指标	预测时间尺度						
		1 h	3 h	6 h	12 h	1 天	3 天	7 天
LSTNet	$M_{AE}/kW$	14 288.44	21 379.54	26 529.10	29 906.02	31 710.41	42 268.71	51 384.37
	$R_{MSE}/kW$	23 055.20	32 719.98	43 769.08	50 069.28	50 531.23	64 857.20	83 407.84
	$M_{APE}/\%$	1.68	2.47	3.08	3.55	3.72	4.76	5.42
STGCN	$M_{AE}/kW$	12 083.21	19 903.27	23 782.74	26 516.53	28 050.03	38 621.35	50 179.03
	$R_{MSE}/kW$	16 854.01	31 607.09	38 910.88	42 960.62	45 776.37	62 448.82	82 568.16
	$M_{APE}/\%$	1.52	2.35	2.76	3.13	3.27	4.39	5.33
MTGNN	$M_{AE}/kW$	11 192.69	18 927.12	24 014.93	25 021.29	27 030.05	37 805.52	50 661.08
	$R_{MSE}/kW$	17 910.11	29 568.95	38 868.32	41 593.82	45 318.71	60 756.19	82 443.15
	$M_{APE}/\%$	1.46	2.29	2.83	2.91	3.21	4.24	5.39
AGAT	$M_{AE}/kW$	10 155.78	17 328.79	21 961.72	24 507.58	26 447.05	36 967.45	50 054.45
	$R_{MSE}/kW$	15 805.35	28 709.30	36 003.72	40 470.71	44 507.34	59 357.51	82 245.79
	$M_{APE}/\%$	1.33	2.15	2.57	2.74	3.04	4.18	5.28

表4 多种模型在不同预测时间尺度上的所有用户负荷的评估指标平均值

Table 4 Average value of evaluation indexes of all user load predicted by multiple models on different forecast time scales

模型类型	评估指标	预测时间尺度						
		1 h	3 h	6 h	12 h	1 天	3 天	7 天
LSTNet	$M_{AE}/kW$	117.63	151.02	167.48	176.60	181.25	212.41	247.51
	$R_{MSE}/kW$	183.43	237.55	257.26	276.74	285.18	333.72	391.73
	$M_{APE}/\%$	6.94	8.84	9.53	9.94	10.35	10.96	11.27
STGCN	$M_{AE}/kW$	110.69	150.46	163.86	168.61	177.09	205.44	245.38
	$R_{MSE}/kW$	173.39	235.30	254.43	269.42	280.45	326.84	391.21
	$M_{APE}/\%$	6.68	8.58	9.27	9.58	10.04	10.58	11.21
MTGNN	$M_{AE}/kW$	110.13	152.40	161.10	169.56	175.89	207.38	244.39
	$R_{MSE}/kW$	172.72	233.32	251.12	271.65	281.02	328.26	390.53
	$M_{APE}/\%$	6.65	8.67	9.18	9.60	9.96	10.75	11.17
AGAT	$M_{AE}/kW$	105.42	140.48	158.06	166.85	172.24	204.44	244.26
	$R_{MSE}/kW$	167.49	218.46	247.58	265.96	277.59	325.04	389.31
	$M_{APE}/\%$	6.35	8.03	8.85	9.41	9.65	10.47	11.12

从实验结果可以看出, AGAT模型相比其他模型表现良好。以总负荷的  $M_{APE}$  为例, 在预测时间尺度分别为 1 h、3 h、6 h、12 h、1 天、3 天和 7 天的预测测试中, AGAT模型预测指标较其他模型至少降低 8.90%、6.11%、6.88%、5.84%、5.30%、1.42% 和 0.94%。说明AGAT模型对于短期预测的效果较好, 而随着预测时间尺度的增加, 各模型预测误差也在增大。因此, 本文所提模型更适合短期预测, 在长时间尺度上的预测仍有待研究。

## 4 结论

针对用户负荷随机性大、波动性强的特点和不同用户负荷间潜在的动态相关性, 本文构建了一种基于自适应图注意力网络的短期用户负荷预测模型, 得出以下结论:

1) 实验对比了多种时序预测基线模型, 所提模型预测效果较好, 预测指标均优于其他模型;

2) 时序特征提取层和动态图学习层均能有效提高GAT模型的预测性能, 且二者间具有良好的兼容性, 组合使用能进一步提高短期用户负荷预测精度;

3) AGAT模型在 1 h至 7 天的预测时间尺度上均具有良好的预测性能, 能够有效挖掘用户负荷数据的时空依赖性, 有助于提高短期用户负荷预测精度。

后续将继续优化网络, 并考虑气象、电价等外部因素对用户负荷的影响, 进一步提高用户负荷预测性能以及在更长时间尺度上的用户负荷预测。

## 参考文献

[1] 梁志远, 谢宇霆, 胡秀珍, 等. 碳中和驱动下电力市场零售平台交易模式研究[J]. 电测与仪表, 2022, 59(8): 47-57.

- LIANG Zhiyuan, XIE Yuting, HU Xiuzhen, et al. Research on the trading mode of electricity market retail platform driven by carbon neutrality[J]. Electrical Measurement & Instrumentation, 2022, 59(8): 47-57.
- [2] 赵文杰, 杨萌, 刘军会, 等. 基于双层随机规划的电力零售商短期决策模型[J]. 智慧电力, 2022, 50(8): 38-45.
- ZHAO Wenjie, YANG Meng, LIU Junhui, et al. Short-term decision-making model for electricity retailer based on bilevel stochastic programming[J]. Smart Power, 2022, 50(8): 38-45.
- [3] 王辉, 沈怡迪, 徐浩成, 等. 基于电力用户细分市场的售电公司最优销售定价策略[J]. 智慧电力, 2021, 49(7): 44-51.
- WANG Hui, SHEN Yidi, XU Haocheng, et al. Optimal pricing strategy for power sale companies based on power consumer segment market[J]. Smart Power, 2022, 50(8): 38-45.
- [4] 李雅婷, 唐家俊, 张思, 等. 考虑多重不确定性因素的售电公司购售电决策模型[J]. 电力系统自动化, 2022, 46(7): 33-41.
- LI Yating, TANG Jiajun, ZHANG Si, et al. Decision-making model of electricity procurement and sale for electricity retailers considering multiple uncertain factors[J]. Automation of Electric Power Systems, 2022, 46(7): 33-41.
- [5] 刘瑾, 赵晶, 冯瑛敏, 等. 基于梯度提升决策树的电力物联网用电负荷预测[J]. 智慧电力, 2022, 50(8): 46-53.
- LIU Jin, ZHAO Jing, FENG Yingmin, et al. Power load forecasting in power internet of things based on gradient boosting decision tree[J]. Smart Power, 2022, 50(8): 46-53.
- [6] 张宇, 李成鑫, 喻理, 等. 偏差电量考核机制下售电公司优化调度研究[J]. 计算机仿真, 2022, 39(8): 300-306.
- ZHANG Yu, LI Chengxin, YU Li, et al. Research on

- optimal dispatching of electricity retailers considering daily deviation penalty[J]. *Computer Simulation*, 2022, 39(8): 300-306.
- [7] 刘晟源, 章天晗, 林振智, 等. 数据赋能低压配用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.
- [8] 王晨, 叶江明, 何嘉弘. 基于不同分布下 GARCH-M 族模型的短期用户负荷预测[J]. *电力工程技术*, 2022, 41(5): 110-115.  
WANG Chen, YE Jiangming, HE Jiahong. Short-term user load forecasting based on GARCH-M family model with different distributions[J]. *Electric Power Engineering Technology*, 2022, 41(5): 110-115.
- [9] 杨胡萍, 余阳, 汪超, 等. 基于 VMD-CNN-BIGRU 的电力系统短期负荷预测[J]. *中国电力*, 2022, 55(10): 71-76.  
YANG Huping, YU Yang, WANG Chao, et al. Short-term load forecasting of power system based on VMD-CNN-BIGRU[J]. *Electric Power*, 2022, 55(10): 71-76.
- [10] ZHU Jizhong, DONG Hanjiang, ZHENG Weiye, et al. Review and prospect of data-driven techniques for load forecasting in integrated energy systems[J]. *Applied Energy*, 2022, 321: 119269.
- [11] 牛牧童, 廖凯, 杨健维, 等. 考虑季节特性的多时间尺度电动汽车负荷预测模型[J]. *电力系统保护与控制*, 2022, 50(5): 74-85.  
NIU Mutong, LIAO Kai, YANG Jianwei, et al. Multi-time-scale electric vehicle load forecasting model considering seasonal characteristics[J]. *Power System Protection and Control*, 2022, 50(5): 74-85.
- [12] ALAMIN M A, HUQUE M A. Comparison of ARIMA and SVM for short-term load forecasting[C] // 2019 9th annual information technology, electromechanical engineering and microelectronics conference, March 13-15, 2019, Jaipur, India: 1-6.
- [13] FAN Guofeng, ZHANG Liuzhen, YU Meng, et al. Applications of random forest in multivariable response surface for short-term load forecasting[J]. *International Journal of Electrical Power & Energy Systems*, 2022, 139: 108073.
- [14] 程志友, 汪德胜. 基于机器学习与疫情关联特征的短期负荷预测[J]. *电力系统保护与控制*, 2022, 50(23): 1-8.  
CHENG Zhiyou, WANG Desheng. Short-term load forecasting based on machine learning and epidemic association features[J]. *Power System Protection and Control*, 2022, 50(23): 1-8.
- [15] 欧阳福莲, 王俊, 周杭霞. 基于改进迁移学习和多尺度 CNN-BiLSTM-Attention 的短期电力负荷预测方法[J]. *电力系统保护与控制*, 2023, 51(2): 132-140.  
OUYANG Fulian, WANG Jun, ZHOU Hangxia. Short-term power load forecasting method based on improved hierarchical transfer learning and multi-scale CNN-BiLSTM-Attention[J]. *Power System Protection and Control*, 2023, 51(2): 132-140.
- [16] MADHIARASAN M. Accurate prediction of different forecast horizons wind speed using a recursiveradial basis function neural network[J]. *Protection and Control of Modern Power Systems*, 2020, 5(3): 48-56.
- [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] 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.
- [19] 孔祥玉, 马玉莹, 艾芊, 等. 新型电力系统多元用户的用电特征建模与用电负荷预测综述[J]. *电力系统自动化*, 2023, 47(13): 2-17.  
KONG Xiangyu, MA Yuying, AI Qian, et al. Review on power consumption characteristics modeling and load forecasting of power consumption for diverse users in new power system[J]. *Automation of Electric Power Systems*, 2023, 47(13): 2-17.
- [20] 刘友波, 吴浩, 刘挺坚, 等. 集成经验模态分解与深度学习的用户侧净负荷预测算法[J]. *电力系统自动化*, 2021, 45(24): 57-64.  
LIU Youbo, WU Hao, LIU Tingjian, et al. User-side net load forecasting method integrating empirical mode decomposition and deep learning[J]. *Automation of Electric Power Systems*, 2021, 45(24): 57-64.
- [21] 李文武, 石强, 李丹, 等. 基于 VMD 和 PSO-SVR 的短期电力负荷多阶段优化预测[J]. *中国电力*, 2022, 55(8): 171-177.  
LI Wenwu, SHI Qiang, LI Dan, et al. Multi-stage optimization forecast of short-term power load based on VMD and PSO-SVR[J]. *Electric Power*, 2022, 55(8): 171-177.

- [22] 钟劲松, 王少林, 冉懿, 等. 基于互信息和 LSTM 的用户负荷短期预测[J]. 电力建设, 2022, 43(7): 96-102. ZHONG Jinsong, WANG Shaolin, RAN Yi, et al. Short-term consumer load forecasting based on mutual information and LSTM[J]. Electric Power Construction, 2022, 43(7): 96-102.
- [23] 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/OL]. 电网技术: 1-11[2023-01-31]. <http://kns.cnki.net/kcms/detail/11.2410.TM.20221013.1648.003.html>. 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/OL]. Power System Technology: 1-11 [2023-01-31]. <http://kns.cnki.net/kcms/detail/11.2410.TM.20221013.1648.003.html>.
- [24] ARASTEHFAR S, MATINKIA M, JABBARPOUR M R. Short-term residential load forecasting using graph convolutional recurrent neural networks[J]. Engineering Applications of Artificial Intelligence, 2022, 116: 105358.
- [25] XIE Jiangjian, ZHONG Yujie, XIAO T, et al. A multi-information fusion model for short term load forecasting of an architectural complex considering spatio-temporal characteristics[J]. Energy and Buildings, 2022, 277: 112566.
- [26] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // Advances in Neural Information Processing Systems 30 (NIPS 2017), 2017, Long Beach, USA: 5998-6008.
- [27] 狄曙光, 刘峰, 孙建宇, 等. 基于改进 ABC 和 IDPC-MKELM 的短期电力负荷预测[J]. 智慧电力, 2022, 50(9): 74-81. DI Shuguang, LIU Feng, SUN Jianyu, et al. Short term power load forecasting based on improved ABC and IDPC-MKELM[J]. Smart Power, 2022, 50(9): 74-81.
- [28] 樊江川, 于昊正, 刘慧婷, 等. 基于多分支门控残差卷积神经网络的短期电力负荷预测[J]. 中国电力, 2022, 55(11): 155-162, 174. FAN Jiangchuan, YU Haozheng, LIU Huiting, et al. Short-term load forecasting based on multi-branch residual gated convolution neural network[J]. Electric Power, 2022, 55(11): 155-162, 174.
- [29] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J]. arXiv preprint arXiv: 1710.10903, 2017.
- [30] 奚鑫泽, 邢超, 覃日升, 等. 基于多层特征融合注意力网络的电能质量扰动识别方法[J]. 智慧电力, 2022, 50(10): 37-44. XI Xinze, XING Chao, QIN Risheng, et al. Power quality disturbance recognition method based on multi-layer feature fusion attention network[J]. Smart Power, 2022, 50(10): 37-44.
- [31] 张亮, 安军, 周毅博. 基于时间卷积和图注意力网络的电力系统暂态稳定评估[J]. 电力系统自动化, 2023, 47(7): 114-122. ZHANG Liang, AN Jun, ZHOU Yibo. Transient stability assessment of power system based on temporal convolution and graph attention network[J]. Automation of Electric Power Systems, 2023, 47(7): 114-122.
- [32] LI Yangfan, LI Kenli, CHEN Cen, et al. Modeling temporal patterns with dilated convolutions for time-series forecasting[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 16(1): 1-22.
- [33] LIU Shaohua, WANG Yisu, SUN Jingkai, et al. An efficient spatial-temporal model based on gated linear units for trajectory prediction[J]. Neurocomputing, 2022, 492: 593-600.
- [34] Electricity consumption dataset[EB/OL]. [2022-10-14]. <https://github.com/laiguokun/multivariate-time-series-data>.
- [35] LAI Guokun, CHANG Weicheng, YANG Yiming, et al. Modeling long-and short-term temporal patterns with deep neural networks[C] // The 41st international ACM SIGIR conference on research & development in information retrieval, July 8-12, 2018, Ann Arbor, USA: 95-104.
- [36] WANG Xiaoyang, MA Yao, WANG Yiqi, et al. Traffic flow prediction via spatial temporal graph neural network[C] // Proceedings of the Web Conference 2020, April 20-24, 2020, Taipei, China: 1082-1092.
- [37] WU Zonghan, PAN Shirui, LONG Guodong, et al. Connecting the dots: multivariate time series forecasting with graph neural networks[C] // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, July 6-10, 2020, Virtual Event, USA: 753-763.

收稿日期: 2023-04-11; 修回日期: 2023-06-05

作者简介:

黄冬梅(1964—), 女, 教授, 博士生导师, 研究方向为海洋与电力时空信息技术; E-mail: dmhuang\_dl@163.com

陈欢(1999—), 男, 硕士研究生, 研究方向为电力负荷预测; E-mail: 18758673328@163.com

孙园(1980—), 男, 通信作者, 副教授, 主要研究方向为数据分析、挖掘和建模。E-mail: combmathe@shiep.edu.cn

(编辑 周金梅)