DOI：10.13234/j.issn.2095-2805.2025.8.145 **中图分类号**：TM715 **文献标志码**：A

基于LSTM-GNN组合模型的电动汽车短期充电负荷预测

刘奕舟[1,2](中国电源学会学生会员)，李振华[1,2]，魏 伟[3]，李振兴[1,2]，徐艳春[1,2]，张 磊[1,2]

(1.三峡大学电气与新能源学院，宜昌 443002；2.智慧能源技术湖北省工程研究中心(三峡大学)，宜昌 443002；3.国网湖北省电力有限公司营销服务中心(计量中心)，武汉 443080)

摘要：在新能源时代背景下，针对电动汽车充电负荷对电网的影响这一热点问题，提出1种基于长短期记忆法-图神经网络LSTM-GNN(long short-term memory network and graph neural network)组合模型的电动汽车短期充电负荷预测方法，以解决现有模型的不足。首先，运用科普拉(Copula)理论对3种典型日负荷进行相关性分析，构建包含主要因素与充电站历史负荷的特征输入矩阵，并建立滑动窗口数据集。然后，利用LSTM与GNN对特征输入矩阵进行特征提取和图数据构建，图神经网络层完成特征聚合并输出预测负荷值及误差评价指标。算例结果表明，该模型在不同区域充电站能有效提取充电负荷关键特征，相较传统模型具有更好的预测效果和泛化能力。

关键词：负荷预测；深度学习；图神经网络；长短期记忆；电动汽车；充电负荷

Short-term Charging Load Forecasting for Electric Vehicles Based on Combined LSTM-GNN Model

LIU Yizhou[1,2], *Student Member*, *CPSS*, LI Zhenhua[1,2], WEI Wei[3], LI Zhenxing[1,2], XU Yanchun[1,2], ZHANG Lei[1,2]

(1. College of Electrical Engineering & New Energy, China Three Gorges University, Yichang 443002, China; 2. Hubei Provincial Engineering Research Center of Intelligent Energy Technology, China Three Gorges University, Yichang 443002, China; 3. Marketing Service Center (Measurement Center), State Grid Hubei Electric Power Co., Ltd., Wuhan 443080, China)

Abstract: In the context of new energy era, aimed at the hot issue of the impact of electric vehicle charging loads on power grid, a short-term charging load forecasting method for electric vehicles based on a combined long short-term memory network and graph neural network (LSTM-GNN) model is proposed to address the shortcomings of existing models. First, three typical daily loads are correlated using the Copula theory, a feature input matrix containing the main factors and historical loads of charging stations is constructed, and a sliding window data set is established. Second, LSTM and GNN are used for feature extraction and graph data construction of the feature input matrix, and the GNN layer completes the feature aggregation and outputs the predicted load value and error evaluation indexes. The results of an example show that the proposed model can effectively extract the key features of charging load at charging stations in different regions, and it has better forecasting performance and generalization capability than traditional models.

Keywords: Load forecasting; deep learning; graph neural networks (GNN); long and short-term memory (LSTM); electric vehicle; charging load

收稿日期：2024-06-19；**修回日期**：2024-07-11；**录用日期**：2024-07-18；**网络首发日期**：2024-07-29

基金项目：国家自然科学基金资助项目(52277012)；武汉强磁场学科交叉基金资助项目(WHMFC202202)

This work is supported by National Natural Science Foundation of China under the grant 52277012; the Interdisciplinary Program of Wuhan National High Magnetic Field Center under the grant WHMFC202202

随着全球气候变化和各国“碳达峰、碳中和”环保政策的推进，新能源汽车的普及率逐渐升高。电动汽车充电站在各城市的铺展普及，其产生的充电负荷对电网带来的波动成为研究重点。准确预测大规模电动汽车充电负荷[1]对于电力系统的规划、运营和管理至关重要，有助于有效平衡电网负荷，维持电网安全稳定运行。

电动汽车充电负荷预测相较于传统电力系统负荷预测更具挑战，因其变化性、随机性及复杂性较高。目前已有预测方法包括基于统计学的传统方法，如贝叶斯统计和回归分析；基于模型的方法，如蒙特卡洛模拟[2]和马尔科夫出行链[3]；基于机器学习的方法，如支持向量机 SVM(support vector machine)、聚类分析等。此外，结合传统时序模型和深度学习的方法(如自回归移动平均数 ARIMA (autoregressive integrated moving average model)与深度学习结合[4])也有应用，但传统模型在处理非平稳性和数据特征提取方面的能力有限。现代预测方法如图神经网络 GNN(graph neural networks)[5-6]、深度强化学习和长短期记忆 LSTM(long short-term memory)网络也逐渐被采用。LSTM 网络能有效捕捉时间序列中的长期依赖，卷积神经网络 CNN(convolutional neural network)[7]处理历史数据时能捕获本地化模式和特征，结合 *K*-Medoids 聚类算法和 CNN 的模型[8]及结合 LSTM、自注意力机制和图卷积网络 GCN(graph convolutional network)[9]的模型在预测中效果较为理想，但存在对高维特征提取不足等问题。近年来，GNN 在不同领域的应用[10]显示出其潜力，将其与 LSTM 结合应用于电动汽车充电负荷预测亦成为可能。

综上所述，考虑电动汽车历史充电负荷数据与气温、日期特征数据的复杂性，本文提出并建立了 1 种结合 LSTM[11]与 GNN[12-13]的 LSTM-GNN 组合预测模型。通过算例实验对改进前后及与其他传统典型模型的对比，本文方法能够较为精准地预测一定时间内 3 个不同区域电动汽车充电站的充电负荷。通过统计各误差评价指标与典型日预测结果得出，本模型的预测精度更高，为电网与电力系统运行提供了有效的预测数据支持，验证了本文方法的正确性和有效性。

1 理论基础

1.1 长短期记忆网络

LSTM 网络是 1 种深度学习模型，用于处理序列数据，特别是在处理具有长时依赖关系的序列数据时表现优越。LSTM 的基本思想是通过引入记忆单元(memory cell)和门控机制来控制信息流。这有助于模型更好地记住长期依赖关系，同时防止梯度消失和梯度爆炸。本文中 LSTM 模型设计为双向层数可调，其结构如图 1 所示。

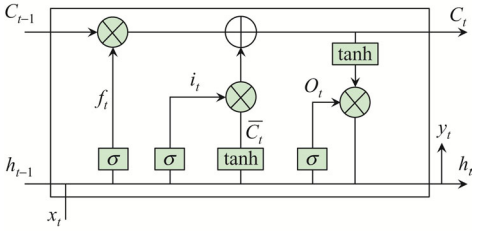


图 1 LSTM 网络结构

Fig. 1 Structure of LSTM network

LSTM 中的记忆单元是网络的关键组件，它负责存储和输出信息。记忆单元能够保留先前的信息，并且可以选择性地丢弃或添加新的信息。模型中的 3 个输入量，包括当前时刻输入样本 X_t，上一时刻的短期记忆信息 H_{t-1} 和上一时刻的长期记忆 C_{t-1}。信息的保留与舍弃则由 LSTM 内部的 3 个门来控制，分别为遗忘门、输入门及输出门。这些门通过 Sigmoid 激活函数产生 0~1 之间的值，用于控制信息的流动。门控机制的设计使得网络可以学习决定何时更新记忆单元中的信息、何时读取信息及何时忘记信息，LSTM 的表达式为

$$\begin{cases} I_t = \sigma(X_i \boldsymbol{W}_{xi} + H_{t-1}\boldsymbol{W}_{hi} + b_i) \\ O_t = \sigma(X_i \boldsymbol{W}_{xo} + H_{t-1}\boldsymbol{W}_{ho} + b_o) \\ F_t = \sigma(X_i \boldsymbol{W}_{xf} + H_{t-1}\boldsymbol{W}_{hf} + b_f) \end{cases} \tag{1}$$

式中：I_t 为输入门；O_t 为输出门；F_t 为遗忘门；$\boldsymbol{W}_x$ 和 $\boldsymbol{W}_h$ 分别为 3 个门的权重矩阵；b_i、b_o、b_f 分别为 3 个门的偏置项；σ 为 Sigmoid 激活函数。

1.2 图神经网络

GNN 是一类专门用于处理图结构数据的深度学习模型，与传统的深度学习模型相比，GNN 可以将特征量之间的复杂关系进行敏锐获取，并能执行相应复杂特征提取任务，适用于社交网络、生物网络、知识图谱等。

图神经网络的基本原理核心在于通过学习节点的特征表示向量，从而对节点和边的特征向量信息进行转移和聚集，节点可以通过聚合相邻节点的特征信息而改变其表示，其表达式为

$$h_i = f(x_i, \{h_j \,|\, v_j \in N(v_i)\}) \tag{2}$$

式中：x_i为节点 v_i的属性；h_j为 v_i的邻居节点的表示；f为聚合函数。

基于 GCN 的传统方法主要利用矩阵卷积运算，传统方法的优点为可通过局部信息更新节点特征向量，同时其主要依赖于节点的特征矩阵进行卷积操作，当特征维度较高时，会导致计算和存储成本增加，因此其缺点为较难处理高维和动态特征，GCN 引入了卷积运算来更新节点表示向量。节点 v_i的更新规则表示为

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(v_i)} \frac{1}{\sqrt{\deg(v_i)\cdot \deg(v_j)}} h_j^{(l)} \boldsymbol{W}^{(l)} \right) \tag{3}$$

式中：$\deg(v_i)$为节点 v_i的度；$\boldsymbol{W}^{(l)}$为权重矩阵。

1.2.1 图数据生成法

图数据生成过程如图 2 所示。使用上一层结果作为节点，计算节点特征矩阵的内积，并将其通过 softmax 函数计算的结果作为权重，得到邻接矩阵。

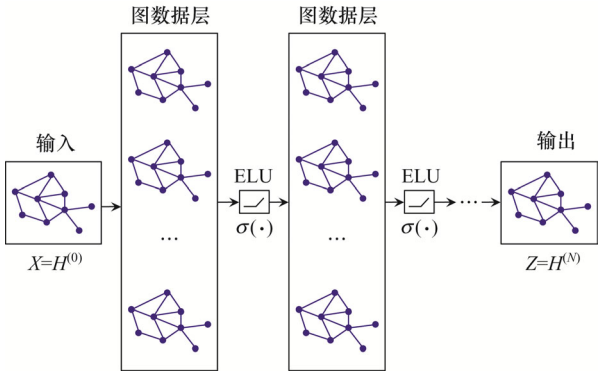


图 2　图数据生成过程

Fig. 2 Process of graph data generation

softmax 函数将含有 K 个实值的向量转换为 K 个总和为 1 的实值向量。向量的 K 个输入值可以是正数、负数、0 或大于 1，softmax 函数能将它们转换为 0~1 的值，因此它们可以被解释为概率。如果其中 1 个输入较小或为负，则 softmax 函数将其变成小概率，反之为大概率，但始终保持在 0~1 之间，其公式可表示为

$$\sigma(\vec{z})_i = \frac{\mathrm{e}^{z_i}}{\sum_{j=1}^{K} \mathrm{e}^{z_j}} \tag{4}$$

式中：z_i为输入向量的元素，可以取任何数值；e^{z_i}为标准指数函数，应用于输入向量的每个元素；$\vec{z}$为 softmax 函数的输入向量，由$(z_0, z_1, \cdots, z_K)$组成；公式底部的项为归一化项，确保函数的所有输出值总和为 1，从而构成有效的概率分布。

比较典型的步骤为将特征全图数据变为局部图数据，过程如图 3 所示。该过程主要分为 3 个步骤：第一步为在 B^2层选择节点 a；第二步为在 B^1层抽样 3 个邻居节点 c、f、j；第三步为在 B^0层抽样 6 个邻居节点 d、e、h、i、l、k。最终通过局部图数据更好地进行下一步分析预测。

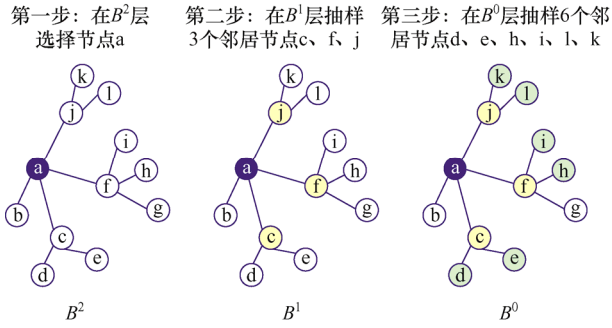


图 3　特征全图转局部图过程

Fig. 3 Process of transforming full feature graph to local graph

1.2.2 图采样与聚合法

目前提出的新技术方法也涉及图采样与聚合 GraphSAGE(graph sample and aggregate)方法和图注意网络 GAT(graph attention network)方法[12]，其优点是可以更加灵活地使用维数减特征的动态图系统，增强对模型的适应性。本文选择 GraphSAGE 法来实现图神经网络的应用。

GraphSAGE 结构由采样与聚合卷积 SAGE-Conv(sample and aggregate convolution)和 ELU 激活函数共同构成。SAGEConv 方法主要是用来构建 GraphSAGE 算法的卷积层，并以上一层的输出作为节点。其中 SAGEConv 的算法原理为

$$h_{N(\mathrm{dst})}^{(l+1)}=\mathrm{aggregate}\left(\left\{h_{\mathrm{src}}^{l},\forall_{\mathrm{src}}\in N(\mathrm{dst})\right\}\right) \tag{5}$$

$$h_{\mathrm{dst}}^{(l+1)}=\sigma(\boldsymbol{W}\cdot\mathrm{concat}(h_{\mathrm{dst}}^{l},h_{N(\mathrm{dst})}^{(l+1)})+\boldsymbol{b}) \tag{6}$$

$$h_{\mathrm{dst}}^{(l+1)}=\mathrm{norm}(h_{\mathrm{dst}}^{l+1}) \tag{7}$$

式中：$h_{N(\mathrm{dst})}^{(l+1)}$ 为目标节点 N 在第 l+1 层的特征；h_{src}^{l} 为源节点在第 l 层的特征；N(dst)为目标节点 N 的集合；$\forall_{\mathrm{src}}$ 为任意源节点；$\boldsymbol{W}$ 为权重矩阵；h_{dst}^{l} 为目标节点在第 l 层的特征；$\boldsymbol{b}$ 为偏置向量；$\mathrm{concat}(h_{\mathrm{dst}}^{l},h_{N(\mathrm{dst})}^{(l+1)})$ 为将聚合后的特征与第 l 层的目标节点和第 l+1 层的目标节点拼接在一起；$h_{\mathrm{dst}}^{(l+1)}$ 为目标节点在第 l+1 层的特征。式(5)~式(7)为对式(2)的改进。

其中源节点特征 feat_src 和目标节点特征 feat_dst 需要根据图类型被指定。接着经过 ELU 激活函数层，ELU 激活函数是对 ReLU 激活函数的改进版本，解决了 ReLU 在负数区域出现的梯度截断问题，其公式为

$$\mathrm{ELU}(x)=\begin{cases}x & x>0\\ \alpha(\mathrm{e}^{x}-1) & x\leqslant 0\end{cases} \tag{8}$$

相较于以往的 GCN 方法，节点 i 与邻居节点 j 可以通过使用不同的特征维度，映射到 1 个完全不同的特征空间。最终将节点自身的特征与聚合得到的邻居表征结合，更新节点的最终表征。GraphSAGE 法通过对式(2)节点更新规则的改进，使得每个节点的表征能够包含其局部邻域的信息，从而更好地捕捉数据集中的特征。

2 LSTM-GNN 组合模型

本文对充电站多位特征向量及时序序列特征进行分析后建模，提出了 LSTM-GNN 组合模型，主要目标是挖掘历史数据中的有用特征信息，最终根据历史充电负荷信息和多种特征影响因素预测未来 24 h 的充电负荷。

2.1 模型结构及预测流程

本文的组合模型 LSTM-GNN 主要由输入层、LSTM 层、全连接层 1、GraphSAGE 层、全连接层 2 和输出层构成，模型网络结构如图 4 所示。

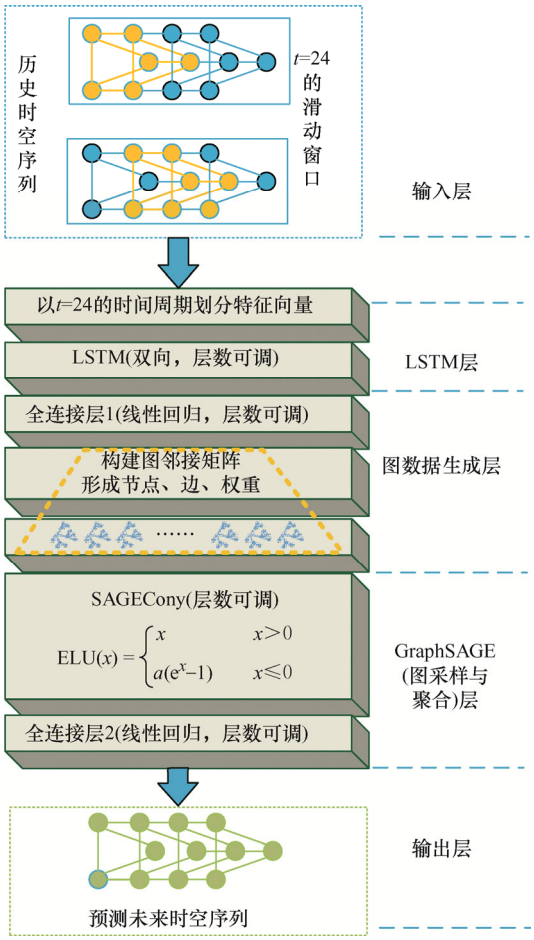


图 4 模型网络结构

Fig. 4 Network structure of model

本文提出的组合模型将历史时空序列以特征矩阵形式进行输入，构建 t=24 的滑动窗口数据集，随后经过 LSTM 层对时序特征进行处理，并通过全连接层完成低维特征与高维特征之间的转换；在图数据生成层中完成对图数据的构建，形成局部图与图邻接矩阵，最后经过 GraphSAGE 层训练得到最终输出结果，设置在最后的全连接层输出预测未来时空序列。

LSTM-GNN 组合模型的预测流程可分为 3 个主要步骤，分别为原始数据分类、构建图数据和生成负荷预测。预测流程如图 5 所示。按照 3 个不同区域对滑动窗口数据集进行划分，然后按周期以

24 组数据为 1 组构建特征输入矩阵，同时对特征因素进行相关性分析，通过构建图邻接矩阵 $\boldsymbol{A}$ 形成时空图数据，在模型中训练的目的是通过前 24 h 历史充电负荷来进行未来 1 h 充电负荷预测，最终通过均方误差损失函数生成 3 个区域充电站电动汽车负荷预测值。

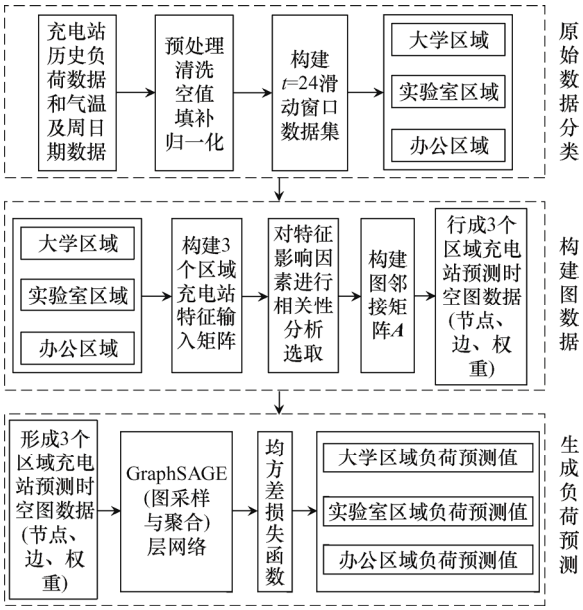


图 5　模型预测流程

Fig. 5 Flow chart of model prediction

2.2 参数设置

网络模型参数按照 LSTM 层、全连接层 1、图数据生成层、SAGEConv 层和全连接层 2 进行设置，其设置见表 1。LSTM 层采用双向结构，能够更全面地捕捉到序列数据中的上下文信息；图数据生成层使用 softmax 函数，因为在节点分类任务中，图神经网络的输出通常是每个节点所属类别的概率分布，使用 softmax 函数可以将输出的原始得分转换为概率，从而确定每个节点的类别；GraphSAGE 层激活函数使用 ELU 函数，因为 ELU 函数与 ReLU 函数相比，其在负区间的输出为指数衰减，使得它能够有效缓解梯度消失问题；对于 SAGEConv 层输入特征，在预处理阶段，首先通过特征选择和提取技术来减少特征的数量，通过预测特点从而选择最重要的特征，最终输出特征数确定为 100 个；对于 SAGEConv 层输出特征，通常根据具体任务需求来确定，选择 128 个输出特征是为了确保模型具有足够的表达能力，但同时又不会引入过多的参数，避免过拟合；全连接层实际上执行的是 1 种线性映射，与线性回归的本质相同，因此可以直接用于回归任务，因此全连接层 1 与 2 分别设置为 1 层与 2 层线性回归，且全连接层 2 为通过 LSTM-GNN 预测模型输出的单个时刻电动汽车充电负荷值，故神经元个数设置为 1。

表 1　LSTM-GNN 网络模型参数

Tab. 1 Parameters of LSTM-GNN network model

参数	设置
LSTM 层	双向结构，层数 1 层
全连接层 1	使用 1 层线性回归
图数据生成层	softmax(x*x_T)
GraphSAGE 层激活函数	$\mathrm{ELU}(x)=\begin{cases}x & x>0\\ \alpha(\mathrm{e}^{x}-1) & x\leqslant 0\end{cases}$，$\alpha$ 取值为 1
SAGEConv 层	输入特征数为 100，输出特征数为 128，层数 2 层
全连接层 2	使用 2 层线性回归，神经元个数为 1

在实际应用中，超参数的设置往往也需要通过实验和调优来确定。可以通过交叉验证、网格搜索等方法找到最优的输出特征数。对于本文实验中进行对比的其余模型，通过特征处理与实验调优，超参数设置情况及依据见表 2。

表 2　各对比模型网络模型参数

Tab. 2 Network model parameters for each comparative model

模型	设置情况及依据
CNN-LSTM 模型	层数共 5 层，CNN 卷积层设置 2 层，针对多阶输入特征；CNN 池化层大小 2×2，缩减特征图维度，神经元个数 1；LSTM 层，双向结构，层数 1 层，避免梯度消失或爆炸
LSTM-CNN 模型	同 CNN-LSTM 模型
SVM 模型	核函数类型为高斯核(RBF)，适用于非线性数据；正则化参数 C 取值为 200，通过网格搜索验证
GNN 模型	层数共 3 层，参考 LSTM-GNN 模型中设置；隐藏单元数为每层中神经元数量 128，通过实验确定
LSTM 模型	层数共 3 层，LSTM 为双向结构；单元数为 128，通过实验确定；时间步长为 24，依据为预测时间间隔

通过以上介绍，可总结得出本文 LSTM-GNN 组合模型设计的创新点主要在于：第一，结合了 LSTM 和 GNN 的优点，首先使用 LSTM 层获取时序数据的长程依赖，然后使用图神经网络学习上一步的时空关联信息；第二，对于 LSTM 层和 GNN 层均为可调节层数的超参数设置；第三，图神经网络层 GNN 使用的是 GraphSAGE 结构，其中的 2 层 SAGEConv 层的层数也为可调节超参数，使用 LSTM 结合图神经网络层，既保持了经典时序 LSTM 在时序数据集上的重要作用，又发挥了图神经网络 GNN 在时空关联上的作用，更好地提升了预测结果。

3 算例验证

选择美国 ACN-DATA 公开数据集[13-14]中 3 个分别位于加州理工大学、NASA 实验室和办公区的电动汽车充电站历史充电负荷数据集进行验证实验，周日期信息和是否双休日则由实际日期得到，气温信息由当地气象局历史数据得到。

3.1 原始数据集分析设置

基于数据集中 3 个不同典型区域(大学、实验室、办公区)的电动汽车充电站历史负荷信息具有各自不同的特点，参考现有文献对于电动汽车充电负荷预测特征因素的分析，选取了历史气温、周日期信息、是否双休日信息及充电站历史充电负荷共同构成特征输入矩阵。

3.1.1 原始数据集预处理

基于 LSTM-GNN 组合模型的电动汽车短期充电负荷预测方法，将采样间隔时间设为 1 h。使用 Python 中 Pandas 包对原始数据中无用信息进行清洗处理，原始数据集文件为 JSON 格式，清洗处理后数据集文件为 CSV 格式。仅保留每辆电动汽车开始充电时间、结束充电时间及充电电量信息；同时使用 Python 对数据集进行重新分类整理，将所有电动汽车充电负荷信息进行统计并依次划分，使用均值法将每日 24 h 充电负荷进行拆分并分配至每小时，以显示每 1 日 24 h 中每小时的充电负荷。清洗前后数据效果对比如图 6 与图 7 所示。



(a)大学区域 (b)实验室区域 (c)办公区域

图 6 数据清洗处理前效果

Fig. 6 Effect of data cleaning before processing



(a)大学区域 (b)实验室区域 (c)办公区域

图 7 数据清洗处理后效果

Fig. 7 Effect of data cleaning after processing

3.1.2 特征因素相关性分析

采用科普拉(Copula)理论对原始数据集中提取 3 类典型日负荷数据，即工作日负荷、双休日负荷、高温日负荷进行相关性分析。

科普拉函数的一般形式为

$$C[u_1(x_1),u_2(x_2),\cdots,u_n(x_n)] \tag{9}$$

式中：$u_1(x_1),u_2(x_2),\cdots,u_n(x_n)$ 为各变量的累计概率函数；C 为科普拉函数。同时定义边缘分布函数 $V_1(x_1),V_2(x_2),\cdots,V_n(x_n)$ 作为函数 C 对各变量边缘分布作用的结果，得到等式为

$$V(x_1,x_2,\cdots,x_n)=C[u_1(x_1),u_2(x_2),\cdots,u_n(x_n)] \quad (10)$$

成功定义出本模型下科普拉函数后，引入斯皮尔曼秩相关系数[15]作为相关性分析评价指标，其公式为

$$\rho=12\int_0^1\int_0^1 C[V(x),W(y)]\mathrm{d}[V(x)]\mathrm{d}[W(y)]-3 \quad (11)$$

式中，$V(x)$、$W(y)$分别为随机变量 x、y 的边缘分布函数。整体相关性分析计算流程如图 8 所示。

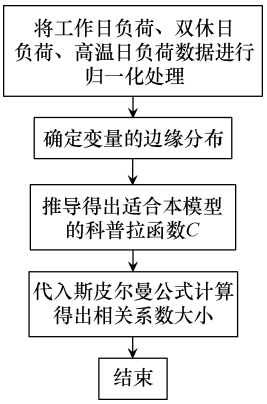


图 8　相关系数计算流程

Fig. 8 Flow chart of correlation coefficient calculation

通过以上公式和计算流程，最终可得到工作日负荷、双休日负荷、高温日负荷间的相关性系数，见表 3。可以看出，3 种典型日负荷之间相关性系数的数值均超过了 0.6，在以往对相关性系数及变量间相关性强度研究中，当相关性系数介于 0.6~0.8 之间即可认为变量间具有强相关关系[16-20]。因此本文所选取 3 种典型日所具有的日期、气温等特征因素，可以很好地代表负荷影响因素间的耦合信息，利用不同负荷间的耦合信息可以有效提高模型预测精度与效果。

表 3　工作日、双休日、高温日负荷之间相关性系数值

Tab. 3 Values of correlation coefficients among loads on weekday, weekend and high-temperature day

对比条件	相关性系数
工作日负荷-工作日负荷	1.000 0
工作日负荷-双休日负荷	0.688 7
工作日负荷-高温日负荷	0.721 2
双休日负荷-双休日负荷	1.000 0
双休日负荷-高温日负荷	0.708 9
高温日负荷-高温日负荷	1.000 0

3.1.3　特征输入矩阵的构建

在对气温、日期等量纲异化型数据进行分析后，为了提升模型收敛与训练速度，对本文中周日期和是否双休信息进行独热编码(one-hot encoding)，即使用 N 位状态寄存器来对 N 个状态进行编码，每个状态都有自己独立的寄存器位，且在任意给定时间内只有 1 个寄存器位是有效的，即只有一位是 1，其余都是 0。

处理好的数据集以 t=24 划分为滑动窗口数据集，按 24 组数据为 1 组，输入为 24 h 数据，输出为对未来 1 h 的预测值，为下步模型训练做好准备。同时设计由多种特征数据共同构成的输入特征矩阵，输入特征矩阵 $\boldsymbol{X}$ 为

$$\boldsymbol{X}=\left[\boldsymbol{w}_{(n+1)\times1}\quad \boldsymbol{T}_{(n+1)\times1}\quad K_{(n+1)\times1}\quad J_{(n+1)\times7}\right] \quad (12)$$

式中：每个训练样本 $\boldsymbol{X}$ 为$(n+1)\times10$ 阶的矩阵；$\boldsymbol{w}_{(n+1)\times1}$ 和 $\boldsymbol{T}_{(n+1)\times1}$ 分别为历史充电负荷和历史气温序列的列向量，其元素 W_j 和 T_j(j=1, 2, …, n+1)分别为当地历史充电负荷值和气温；K_i 为 1 位代码 0 或 1，0 代表当天为工作日，1 代表当天为双休日；$J_{(n+1)\times7}$ 为 7 位二进制代码，代表周一到周日，如 0000100 代表周五。模型的输入由若干个 24×10 的输入特征矩阵组成的输入向量构成。

3.1.4　数据的准备与划分

在对 3 个不同区域充电站历史充电负荷数据集进行处理与分析后，整个输入数据集按照 8:1:1 的比例划分训练集、验证集和测试集。

3.2　误差评价指标的选取

为了综合评估模型的效果，按照负荷预测里关于时间序列预测方面的惯用指标，综合考虑本预测模型的特点，合理考虑不同区域与不同典型日下负荷状态情况，并针对电动汽车充电负荷预测数据特征，选取了均方根误差 RMSE(root mean squared error)、平均绝对百分误差 MAPE(mean absolute

percentage error)和 R^2(*R*-squared)这 3 个评价指标来判断模型的预测效果。3 个指标(R_{MSE}、M_{APE} 和 R^2)计算公式分别为

$$R_{\mathrm{MSE}}=\sqrt{\frac{1}{q}\sum_{k=1}^{q}\left(w_k-\overline{w}_k\right)^2} \tag{13}$$

$$M_{\mathrm{APE}}=\frac{1}{q}\sum_{k=1}^{q}\left|\frac{w_k-\hat{w}_k}{w_k}\right|\cdot 100\% \tag{14}$$

$$R^2=1-\frac{\sum_{k=1}^{q}(w_k-\hat{w}_k)^2}{\sum_{k=1}^{q}(w_k-\overline{w}_k)^2} \tag{15}$$

式中：k 为预测各时刻，预测间隔为 1 h；w_k 为真实的 k 时刻的充电负荷；$\hat{w}_k$ 为 LSTM-GNN 模型预测的 k 时刻充电负荷；$\overline{w}_k$ 为充电负荷真实数据平均值；q 为测试数据集中数据点数。

3.3 实验结果分析

在进行实验训练时，为验证所提 LSTM-GNN 模型的有效性，分别选择了 SVM、GNN、LSTM、CNN-LSTM、LSTM-CNN 与所提 LSTM-GNN 模型进行未来 1 h 的电动汽车短期充电负荷预测，并通过分析预测误差、计算误差评价指标值进行对比。使用控制变量法，所有方法除使用模型不同外，输入特征矩阵数据、训练网络层数、特征维度、特征指标等完全相同。

3.3.1 误差评价指标对比分析

通过模型训练得到预测误差指标的各项数据见表 4，可以得到本文所提 LSTM-GNN 模型在 3 个不同区域充电站的预测评价指标均能达到最佳值，RMSE 和 MAPE 指标与传统模型 CNN-LSTM、LSTM-CNN、LSTM、GNN、SVM 等相比，均得到了较大幅度下降，而 R^2 指标则得到了较明显的提升，代表着预测负荷与真实负荷值拟合程度的提升。尤其在受工作日和双休日因素影响较大的办公区域体现最为明显；同时在预测时间 3~4 月加州天气波动情况较大和 7~8 月高温日情况下，所提出的预测模型仍有着较好的解释度。

表 4 3 个区域各模型预测误差指标对比

Tab. 4 Comparison of forecast error indexes across three regions among different models

区域	预测模型	RMSE/MW	MAPE	R^2
大学	CNN-LSTM	5.895 2	2.601 8	0.742 5
	LSTM-CNN	9.640 0	4.375 0	0.311 5
	LSTM	6.989 8	2.582 8	0.638 0
	GNN	5.242 0	3.417 8	0.704 3
	SVM	6.536 0	3.209 3	0.683 5
	LSTM-GNN	5.184 5	2.544 2	0.800 9
实验室	CNN-LSTM	11.994 8	8.895 4	0.920 8
	LSTM-CNN	14.316 0	8.803 6	0.887 1
	LSTM	24.321 0	9.263 4	0.674 2
	GNN	13.532 8	9.459 0	0.807 3
	SVM	15.667 3	11.287 0	0.864 8
	LSTM-GNN	11.750 9	8.704 3	0.924 0
办公	CNN-LSTM	2.095 4	2.991 0	0.365 7
	LSTM-CNN	2.110 7	3.324 0	0.356 5
	LSTM	2.151 1	2.864 8	0.331 6
	GNN	1.844 8	3.731 7	0.508 4
	SVM	2.200 5	3.829 1	0.300 5
	LSTM-GNN	1.771 8	2.760 2	0.546 5

3.3.2 预测结果对比分析

在对模型进行训练后，得到每个区域不同模型预测结果与实际负荷预测结果对比图，并从预测数据中随机选取了典型日预测结果数据进行分析，由于特征影响因素为气温、日期、是否双休日，且经过 3.1.2 节相关性分析，故典型日分别选取工作日、双休日、高温日 3 类，短期负荷预测结果取 48 h 预测结果作为参考。

大学区域典型日预测结果对比如图 9 所示，由图 9(a)和(b)可以看出，双休日负荷峰值小于工作日。在整体充电负荷较大的工作日和高温日夜间时段，LSTM-CNN、CNN-LSTM 和 LSTM 的预测结果与实际负荷结果相差较大；LSTM-GNN、GNN 和 SVM 模型的预测结果与真实负荷更加拟合。由图 9(c)可以看出，在负荷波动较大的高温日时段，白天 SVM 和 LSTM 的预测效果较差，同时 LSTM-GNN 和 GNN 的 MAPE 值均在 3%以内，且在夜间时段 LSTM-GNN 的部分点 R^2 达到 0.8，对于模型的解释度达到 80%，具有最好的拟合性，达到最佳

预测效果。

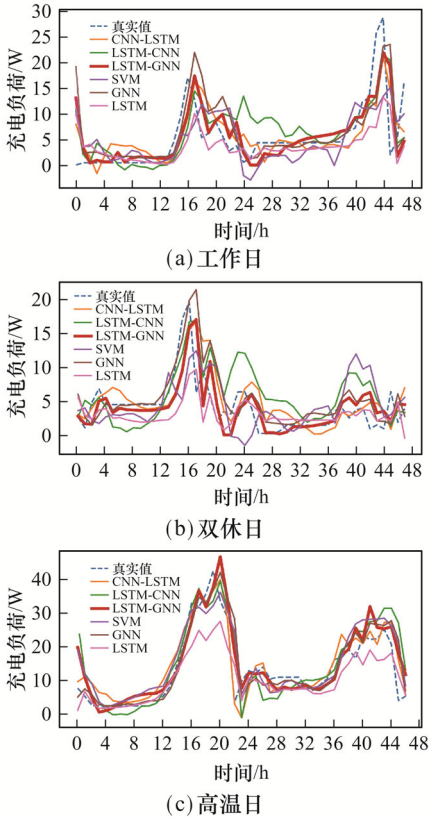
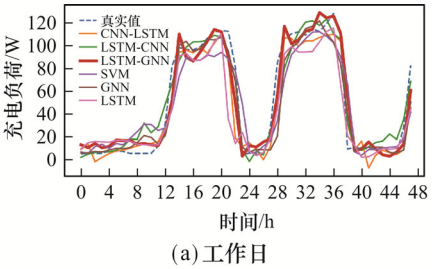


图 9　大学区域典型日预测结果对比

Fig. 9 Comparison of forecast results on typical days in college area

实验室区域预测结果对比如图 10 所示，可知实验室区域负荷在 3 种典型日的波峰波谷变动都较为规律。由图 10(b)和(c)可以看出，存在双休日夜间充电负荷较小和高温日负荷波动较大的情况。对于波峰负荷预测方面，LSTM 模型表现出最差的预测效果，与真实负荷曲线基本背离；其余模型整体都具有较好的拟合度，但 SVM 的部分点 MAPE 值超过了 10%，LSTM 的也接近 9%；LSTM-CNN 则在双休日夜间负荷波谷区具有较差的拟合度，同时 LSTM-GNN 的 RMSE 值仅为 11.750 9，且在高



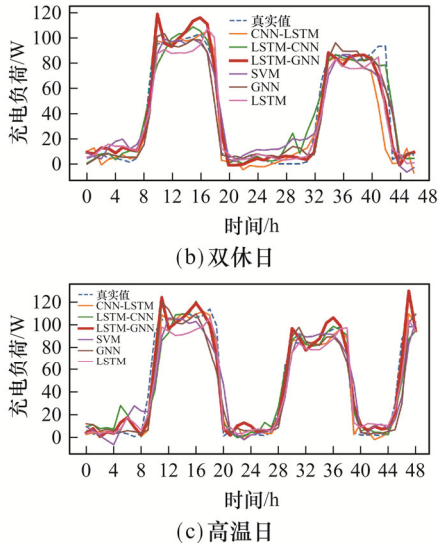


图 10　实验室区域典型日预测结果对比

Fig. 10 Comparison of forecast results on typical days in laboratory area

温日内部分点 R^2 达到 0.924，对模型的解释度超过 90%，具有最佳的预测效果。

办公区域预测结果对比如图 11 所示。由图 11(a)和(b)可以看出，办公区域负荷受工作日与双休日影响较大，休息日充电负荷整体处于较低水平。总体上只有 LSTM-GNN 与 GNN 预测结果与真实负荷曲线有着较高的拟合度，但在 MAPE 数值方面，GNN 的部分点超过了 3.5%，由此可见单一模型存在不确定性。由图 11(c)可以看出，在高温日负荷波峰区域，其他模型的预测精度均与真实值相差较大，只有 LSTM-GNN 的部分点 R^2 超过了0.5，在样本数量较少的情况下，只有 LSTM-GNN有着较高的模型解释度，达到最佳的预测效果。

LSTM-GNN 模型预测效果更好的原因在于：LSTM-GNN 模型更好地结合了不同模型的优点，避免了因异常数据与用户行为习惯差异导致的不能较好进行学习分析的问题。与单一的 LSTM、SVM 与 CNN-LSTM、LSTM-CNN 模型相比，LSTM-GNN 模型通过引入 GraphSAGE 算法，允许对节点特征进行灵活的采样和聚合，通过在每一层采样一部分邻居节点并聚合它们的特征，模型可以逐步汇聚全局信息，同时降低特征维度，使得模型更容易处

理高维特征和动态图结果，提高了模型的适用性。

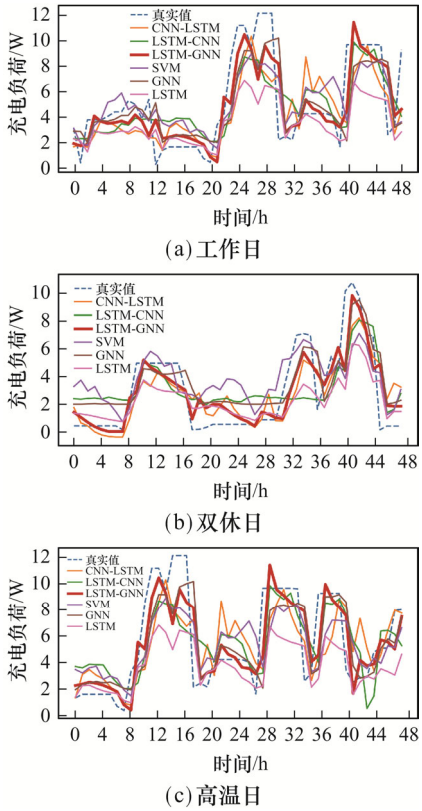


图 11 办公区域典型日预测结果对比

Fig. 11 Comparison of forecast results on typical days in office area

4 结论

本文首先分析现有电动汽车短期充电负荷预测方法的特点和存在的问题，提出 1 种 LSTM-GNN 组合预测模型；此外，考虑到不同区域电动汽车充电站的特征输入不同，提出使用 GraphSAGE 算法将特征向量进行分类与聚合，并通过图神经网络中的聚合框架进行训练，最终通过算例仿真输出预测结果，得出以下结论。

(1)从模型设计上，首先使用 LSTM 层获取长程依赖的时序数据，然后使用图神经网络学习上一步的时空关联信息，同时对 LSTM 层和 GNN 层均采用可调节层数的超参数设置，既保持了经典时序 LSTM 在时序数据集上的重要作用，又发挥了图神经网络 GNN 在时空关联上的作用，更好地提升了预测结果。

(2)考虑不同区域充电站用户充电差异，对原始数据集进行了预处理，同时根据 3 种典型日负荷进行了相关性分析和特征因素选取，设计了以历史气温、日期类型、周末因素与历史充电负荷构成的 10 阶特征输入矩阵。

(3)通过对算例实验结果中典型日预测结果的分析，在工作日、双休日与高温日，该模型的预测方法与 SVM、LSTM、GNN、CNN-LSTM、LSTM-CNN 模型相比，预测精度更高，各预测评价指标结果均更优，能够对电力系统的稳定运行起到参考作用。在后续研究中，还可考虑引入更多的特征因素，如充电站地理位置、出行规律、电动汽车电池 SOC 状态、不同时段电价等，通过丰富输入特征矩阵，更加完善总体充电负荷预测模型，以达到提升预测模型精度的目的。

参考文献：

[1] Lee Z J, Chang D, Jin Cheng, et al. Large-scale adaptive electric vehicle charging [C]// 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP). Anaheim, CA, USA, 2019: 863-864.

[2] 刘青，戚中译. 基于蒙特卡洛法的电动汽车负荷预测建模[J]. 电力科学与工程, 2014, 30(10): 14-19.
Liu Qing, Qi Zhongyi. Electric vehicles load forecasting model based on Monte Carlo simulation [J]. Electric Power Science and Engineering, 2014, 30(10): 14-19. (in Chinese)

[3] 陈丽丹，聂涌泉，钟庆. 基于出行链的电动汽车充电负荷预测模型[J]. 电工技术学报, 2015, 30(4): 216-225.
Chen Lidan, Nie Yongquan, Zhong Qing. A model for electric vehicle charging load forecasting based on trip chains [J]. Transactions of China Electrotechnical Society, 2015, 30(4): 216-225. (in Chinese)

[4] Gao Qiang, Zhu Tao, Zhou Weijiang, et al. Charging load forecasting of electric vehicle based on Monte Carlo and deep learning [C]// 2019 IEEE Sustainable Power and Energy Conference (iSPEC). Beijing, China, 2020: 1309-1314.

[5] 刘德澳, 董明刚, 叶威, 等. 基于图神经网络的非均匀风电场功率预测方法[J]. 太阳能学报, 2023, 44(11): 279-286.
Liu Deao, Dong Minggang, Ye Wei, et al. Power prediction method of non-uniform wind farm layout based on graph neural network [J]. Acta Energiae Solaris Sinica, 2023, 44(11): 279-286. (in Chinese)

[6] 薛阳, 燕宇铖, 贾巍, 等. 基于改进灰狼算法优化长短期记忆网络的光伏功率预测[J]. 太阳能学报, 2023, 44(7): 207-213.
Xue Yang, Yan Yucheng, Jia Wei, et al. Photovoltaic power prediction model based on IGWO-LSTM [J]. Acta Energiae Solaris Sinica, 2023, 44(7): 207-213. (in Chinese)

[7] 姚芳, 汤俊豪, 陈盛华, 等. 基于ISSA-CNN-GRU模型的电动汽车充电负荷预测方法[J]. 电力系统保护与控制, 2023, 51(16): 158-167.
Yao Fang, Tang Junhao, Chen Shenghua, et al. Charging load prediction method for electric vehicles based on an ISSA-CNN-GRU model [J]. Power System Protection and Control, 2023, 51(16): 158-167. (in Chinese)

[8] 季玉琦, 严亚帮, 和萍, 等. 基于 *K*-Medoids 聚类与栅格法提取负荷曲线特征的 CNN-LSTM 短期负荷预测[J]. 电力系统保护与控制, 2023, 51(18): 81-93.
Ji Yuqi, Yan Yabang, He Ping, et al. CNN-LSTM short-term load forecasting based on the *K*-Medoids clustering and grid method to extract load curve features [J]. Power System Protection and Control, 2023, 51(18): 81-93. (in Chinese)

[9] 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 by graph convolutional network combining LSTM and self-attention mechanism [J]. Journal of Computer Applications, 2024, 44(1): 311-317. (in Chinese)

[10] 李致远, 徐丙磊, 周颖仪. 基于图神经网络的账户余额模型区块链地址分类方法[J]. 通信学报, 2023, 44(9): 115-126.
Li Zhiyuan, Xu Binglei, Zhou Yingyi. Graph neural network-based address classification method for account balance model blockchain [J]. Journal on Communications, 2023, 44(9): 115-126. (in Chinese)

[11] 舒征宇, 翟二杰, 李镇翰, 等. 基于 Dropout 优化算法和 LSTM 的铅酸蓄电池容量预测[J]. 电源学报, 2023, 21(5): 173-181.
Shu Zhengyu, Zhai Erjie, Li Zhenhan, et al. Prediction of lead-acid battery capacity based on dropout optimization algorithm and LSTM [J]. Journal of Power Supply, 2023, 21(5): 173-181. (in Chinese)

[12] Scarselli F, Gori M, Tsoi A C, et al. The graph neural network model [J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61-80.

[13] Lee Z J, Sharma S, Johansson D, et al. ACN-sim: An open-source simulator for data-driven electric vehicle charging research [J]. IEEE Transactions on Smart Grid, 2021, 12(6): 5113-5123.

[14] Lee G, Lee T, Low Z, et al. Adaptive charging network for electric vehicles [C]// 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP). Washington, DC, USA, 2017: 891-895.

[15] 杨柳, 李明. 基于科普拉函数的电、热负荷及风电出力间的相关性分析与建模[J]. 吉林电力, 2021, 49(2): 1-5.
Yang Liu, Li Ming. Correlation analysis and modeling between electric-heat load and wind power output based on copula function [J]. Jilin Electric Power, 2021, 49(2): 1-5. (in Chinese)

[16] 罗凤章, 张旭, 杨欣, 等. 基于深度学习的综合能源配电系统负荷分析预测[J]. 高电压技术, 2021, 47(1): 23-32.
Luo Fengzhang, Zhang Xu, Yang Xin, et al. Load analysis and prediction of integrated energy distribution system based on deep learning [J]. High Voltage Engineering, 2021, 47(1): 23-32. (in Chinese)

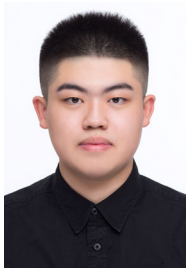
[17] 李振华, 兰芳, 钟悦, 等. 基于双骨架共绕技术的测量保护一体化电流传感器[J]. 高电压技术, 2022, 48(11): 4427-4436.
Li Zhenhua, Lan Fang, Zhong Yue, et al. Measurement-protection-integrated current sensor based on double-

bobbin co-winding technology [J]. High Voltage Engineering, 2022, 48(11): 4427-4436. (in Chinese)

[18] 王玉梅, 张家康. 基于卷积神经网络多判据融合的井下电网故障选线方法[J]. 电源学报, 2023, 21(5): 110-117.
Wang Yumei, Zhang Jiakang. Fault line selection method for mine power grid based on fusion of multiple criteria of convolutional neural network [J]. Journal of Power Supply, 2023, 21(5):110-117. (in Chinese)

[19] 李振华, 廖星锐, 童悦, 等. 基于动态重燃弧模型的 VFTO 与 VFTC 仿真及特性分析[J]. 电力系统保护与控制, 2023, 51(3): 79-88.
Li Zhenhua, Liao Xingrui, Tong Yue, et al. Simulation and characteristic analysis of VFTO and VFTC based on a dynamic reignition arc model [J]. Power System Protection and Control, 2023, 51(3): 79-88. (in Chinese)

[20] Guo Linqi, Erliksson K F, Low S H. Optimal online adaptive electric vehicle charging [C]// 2017 IEEE Power & Energy Society General Meeting. Chicago, IL, USA, 2018: 1-5.



刘奕舟

作者简介：

刘奕舟(1999—)，男，中国电源学会学生会员，硕士研究生。研究方向：电动汽车负荷预测，电能计量技术。Email：mythliu093@163.com。

李振华(1986—)，男，通信作者，博士，教授。研究方向：电磁测量与传感，电力设备状态监测技术。Email：lizhenhua1993@163.com。

魏伟(1989—)，男，博士，高级工程师。研究方向：电力计量技术，电力设备状态评估技术，碳计量技术。Email：1036243440@qq.com。

李振兴(1977—)，男，博士，教授。研究方向：电力系统继电保护。E-mail：lzx2007001@163.com。

徐艳春(1973—)，女，博士，副教授。研究方向：主动配电网电能质量检测与治理,电力系统优化运行与调度。E-mail：xyc7309@163.com。

张磊(1986—)，男，博士，副教授。研究方向：大规模新能源接入后电力系统优化运行控制。E-mail：leizhang3188@163.com。

(责任编辑：郭凯凯)