

# 基于图填补神经网络的配电网稀疏量测数据推演方法

李企洲<sup>1</sup>, 李 梁<sup>1</sup>, 赵 健<sup>1</sup>, 高 源<sup>1</sup>, 孙 洲<sup>2</sup>, 陈 峰<sup>3</sup>

- (1. 上海电力大学 电气工程学院, 上海 200090;
2. 国网浙江省电力有限公司嵊州市供电公司, 浙江 绍兴 312499;
3. 国网浙江省电力有限公司电力科学研究院, 杭州 310014)

**摘 要:** 针对配电网量测设备部署不全、数据传输丢失等导致的量测数据稀疏问题, 提出一种基于 GINN (图填补神经网络) 的配电网稀疏量测数据推演方法, 以解决配电网现有量测数据精度低、稀疏度高的问题。首先, 基于 GINN 的量测特征编码器模块, 提取配电网节点量测数据的功率、电压等潮流特征, 并利用 Transformer 网络建立了节点量测不同潮流特征之间的关联。其次, 基于 GINN 的图编码器模块, 建立了配电网节点间的拓扑连接关系, 并利用 GCN (图卷积网络) 实现了节点量测潮流特征的传递与更新。然后, 通过两个模块分别捕捉的节点量测数据不同潮流特征间的关系和节点间的拓扑关联, 利用稀疏的量测数据实现了缺失数据的推演补齐。最后, 采用 IEEE 14、30、57 和 118 节点系统开展仿真测试, 验证了所提方法的有效性。

**关键词:** 配电网; 稀疏量测; 图填补神经网络; 数据补齐

**DOI:** 10.19585/j.zjdl.202603009

开放科学(资源服务)标识码(OSID):



## An inference method for sparse measurements in distribution networks based on a graph imputation neural network

LI Qizhou<sup>1</sup>, LI Liang<sup>1</sup>, ZHAO Jian<sup>1</sup>, GAO Yuan<sup>1</sup>, SUN Zhou<sup>2</sup>, CHEN Feng<sup>3</sup>

- (1. College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China;
2. State Grid Shengzhou Power Supply Company, Shaoxing, Zhejiang 312499, China;
3. State Grid Zhejiang Electric Power Co., Ltd. Research Institute, Hangzhou 310014, China)

**Abstract:** Incomplete deployment of measurement devices and data transmission losses can result in sparse measurements in distribution networks. To address this issue, this paper proposes an inference method for sparse measurements based on a graph imputation neural network (GINN). The proposed method aims to improve the accuracy and reduce the sparsity of existing measurements. First, a GINN-based measurement feature encoder module is designed to extract power flow features such as power and voltage from nodal measurements. A transformer network is employed to model cross-feature correlations among different power flow features. Second, a GINN-based graph encoder module explicitly encodes topological connectivity between distribution network nodes. By incorporating a graph convolutional network (GCN), this module enables the propagation and updating of nodal power flow features. Subsequently, by leveraging two modules to capture the correlations between different power flow features of node measurements and the topological correlations across nodes, the missing data is inferred and completed using the sparse measurements. Finally, simulation tests are conducted on IEEE 14-, 30-, 57-, and 118-bus systems to validate the effectiveness of the proposed method.

**Keywords:** distribution network; sparse measurement; GINN; data completion

## 0 引言

在现代电力系统中, 数据的采集、分析和应用发挥着至关重要的作用。高质量的数据不仅是

电力资源优化配置的基础, 还能为系统的运行维护提供可靠依据。配电网作为电力系统的重要组成部分, 其量测数据是保障系统安全稳定运行的关键基础。配电网的实时状态监测和状态估计都离不开完整准确的配电网数据<sup>[1-3]</sup>。

然而, 由于量测设备部署不全、数据传输丢失, 配电网存在量测数据缺失的问题。目前配电

网的量测数据主要来源于配电自动化系统,但由于配电网规模庞大,节点和支路众多,难以实现配电自动化量测设备的全面部署<sup>[4-6]</sup>。此外,在量测数据采集和传输的过程中,容易受到通信设备故障、数据拥塞和网络安全等问题的影响,导致数据传输丢失<sup>[7]</sup>。上述因素共同导致了配电网量测数据的稀疏性,进而影响对配电网运行状态的全面监测,降低了配电网运行的安全性和可靠性。因此,有必要研究配电网的稀疏量测数据推演方法,即利用稀疏的量测数据推演得到缺失的量测数据,实现配电网的全面感知。

现有的数据推演方法可分为基于统计学、基于机器学习和基于状态估计三类<sup>[8-9]</sup>。基于统计学的方法通常通过数据的历史分布模式和相关性来推演补齐缺失数据,常见的方法包括均值插值、线性插值和回归分析等<sup>[10-11]</sup>。基于机器学习的方法一般是先对缺失数据集进行聚类或分类,以识别数据中的潜在模式和关系,随后通过训练模型对数据进行推演<sup>[12-13]</sup>。然而,上述方法仅从数据本身出发进行补全和校验,未能充分考虑配电网的电气潮流特性,导致推演结果的准确性较低。

基于状态估计的方法通过利用电力系统的潮流模型和已知量测数据,结合神经网络和数学优化技术实现缺失数据的推演<sup>[14-15]</sup>。文献[16]提出一种基于改进生成式对抗网络的数据推演方法,该方法利用生成器生成可能的缺失数据,并通过判别器验证生成数据的真实性,从而使推演的数据能够更好地反映量测真实的时序特性。文献[17]基于长短期记忆网络,利用已知数据建立对缺失数据的映射。通过记忆和更新历史数据的动态特性,实现了对缺失数据的推演。文献[18]基于深度神经网络,通过相关性分析提取强相关量测,利用海量历史数据建立状态估计模型,推演得到缺失的量测数据。但是,基于状态估计的方法忽略了配电网的拓扑结构信息,使得数据推演的精度难以进一步提高。此外,在面临量测稀疏的情况时,可能会因为无法满足量测数据的冗余度要求而导致推演算法失效。

配电网的拓扑结构由节点和连接线路组成,构成了一种天然的图。由于拓扑结构约束,配电网节点状态量之间存在明显的空间耦合性,这使得节点的量测数据呈现出非欧氏空间的特性。由

于传统的数据推演方法在处理这种非欧氏空间数据上存在一定困难,图神经网络成为一种有效的工具,能够捕捉和利用节点之间的拓扑关系,从而解决由拓扑结构引起的非欧氏空间数据的问题<sup>[19-24]</sup>。针对上述问题,本文将在考虑量测稀疏的场景下,提出一种基于GINN(图填补神经网络)的配电网稀疏量测数据推演方法,其主要创新和贡献如下:

1)提出了一种基于GINN的数据推演方法。从图的角度建模配电网的缺失数据推演问题,GINN通过量测特征编码器和图编码器两个模块,分别捕捉节点量测数据不同潮流特征间的关系和节点间的拓扑关联,实现了缺失数据的推演补齐。

2)通过GINN的量测特征编码器模块,利用Transformer网络建立了节点量测不同潮流特征之间的非线性关系。该方法能够在稀疏量测数据条件下充分挖掘特征间的关联性,提高了数据推演的精度。

3)通过GINN的图编码器模块,建立了配电网节点间的拓扑连接关系,并利用GCN(图卷积网络)实现了节点量测潮流特征的传递与更新。该方法利用拓扑结构数据增强了节点间的关联性,克服了传统方法忽略拓扑数据所带来的推演精度下降问题。

## 1 问题建模

### 1.1 配电网稀疏量测数据推演

定义一张图 $G=\{V,E\}$ 用于表示配电网,其中 $V=\{v_1,\dots,v_N\}$ 和 $E$ 分别表示配电网中节点和支路数的集合。对于配电网的节点 $m$ ,用 $z_m=[z_m(k_0),\epsilon,z_m(k_2),\epsilon,\dots,z_m(k_n)]$ 表示 $k_0$ 到 $k_n$ 时刻节点 $m$ 的量测数据,包括节点电压幅值、节点注入有功功率和无功功率等。 $\epsilon$ 表示节点 $m$ 在某时刻的量测数据缺失。配电网稀疏量测数据推演的目标就是利用稀疏的量测数据推演补齐缺失的数据,得到全部的量测数据 $z_m=[z_m(k_0),z_m(k_1),z_m(k_2),z_m(k_3),\dots,z_m(k_n)]$ 。配电网稀疏量测数据推演的示意图如图1所示。

### 1.2 基于状态估计的配电网稀疏量测数据推演模型

由于配电网的量测数据来源于不同的配电自动化系统量测设备,导致量测数据在时间和空间

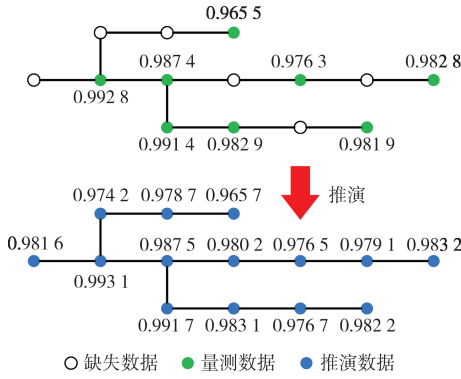


图1 稀疏量测数据推演示意图

Fig.1 Schematic diagram of sparse measurement inference

上存在差异,所以需要综合利用多源的稀疏量测数据进行时序的缺失数据推演。多源量测矩阵 $z$ 的表达式为:

$$z = \begin{bmatrix} z_1(k_0) & z_1(k_1) & z_1(k_2) & \cdots & z_1(k_n) \\ z_2(k_0) & z_2(k_1) & \epsilon & \cdots & \epsilon \\ z_3(k_0) & \epsilon & z_3(k_2) & \cdots & z_3(k_n) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ z_N(k_0) & \epsilon & \epsilon & \cdots & z_N(k_n) \end{bmatrix} \quad (1)$$

式中:  $z_N(k_n)$ 为 $k_n$ 时刻节点 $N$ 的量测数据;  $\epsilon$ 为当前时刻该节点的量测数据缺失。

如式(1)所示,当部分节点的量测数据缺失时,配电网的量测数据在时空分布上表现出较高的稀疏性。

为了获得配电网的节点电压幅值状态向量 $V$ 和电压相角状态向量 $\delta$ ,此时的配电网稀疏量测数据推演即为求解与量测向量加权残差平方和最小的理论值向量对应的电压幅值和相角,其表达式为:

$$\left\{ \begin{array}{l} [V|\delta] = \arg \min_{V, \delta} [z - h(V, \delta)]^T W [z - h(V, \delta)] \\ \text{s.t. } z = [V_i, P_i, Q_i, P_{ij}, Q_{ij}] \\ P_i = V_i \sum_{j=1}^N V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \\ Q_i = V_i \sum_{j=1}^N V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \\ P_{ij} = V_i^2 g_{ij} - V_i V_j g_{ij} \cos \delta_{ij} - V_i V_j b_{ij} \sin \delta_{ij} \\ Q_{ij} = -V_i^2 b_{ij} - V_i V_j g_{ij} \sin \delta_{ij} + V_i V_j b_{ij} \cos \delta_{ij} \end{array} \right. \quad (2)$$

式中:  $W$ 为量测权重矩阵;  $h(\bullet)$ 为非线性量测函数,用于实现状态量到量测量的转换;  $N$ 为注入

节点总数;  $V_i$ 为节点 $i$ 的电压;  $P_i$ 和 $Q_i$ 分别为节点 $i$ 的注入有功和无功功率;  $P_{ij}$ 和 $Q_{ij}$ 分别为支路 $ij$ 的注入有功和无功功率;  $G_{ij}$ 和 $B_{ij}$ 分别为节点导纳矩阵中元素 $ij$ 的实部和虚部;  $\delta_{ij}$ 为节点 $i$ 和 $j$ 之间的相角;  $g_{ij}$ 为支路 $ij$ 的电导;  $b_{ij}$ 为支路 $ij$ 的电纳。

基于状态估计的数据推演方法对数据质量的依赖性强,要求输入量测具有冗余度。面对量测数据稀疏的情况时,该方法受到缺失数据的误导,容易产生较大的推演误差,使得推演的精度下降。另外,针对大规模和复杂的配电网,该方法的计算复杂度较高且计算效率低下,难以满足实时性要求。

实际上,状态向量 $x$ 与历史量测数据存在数学上的函数映射关系,可以利用神经网络实现函数映射关系的拟合并构建其概率分布模型,通过输入部分量测数据,推演出全部节点的状态信息。

$$x = \int_z F(x|z) F(z) dz \quad (3)$$

式中:  $F(x|z)$ 为状态向量 $x$ 在给定量测 $z$ 的条件下的条件概率密度函数;  $F(z)$ 为量测 $z$ 的边缘概率密度函数。

配电网稀疏量测数据推演的输入变量为非欧式空间数据,既包含了节点的电压、注入有功和无功功率等量测数据,也包含了表示节点间拓连接关系的邻接矩阵数据。输出变量为节点的电压幅值和电压相位。传统的数据推演方法忽略了配电网的拓扑结构数据,这使得推演的精度有限,推演结果与真实值存在一定的差距。面对非欧式空间数据的挑战,图神经网络表现出强大的应用潜力,能够有效捕获和处理节点之间复杂的空间关系,为提高数据推演的精度提供有力工具,有助于实现配电网的精准态势感知。

## 2 基于GINN的稀疏量测数据推演

### 2.1 GINN原理

GINN通过量测特征编码器和图编码器两个模块,分别捕捉配电网节点量测数据不同潮流特征间的关系和节点间的拓扑关联,从而更好地填补缺失数据,使得填补的缺失数据在结构上更符合原始数据的特征。采用GINN实现配电网稀疏量测数据推演,基于配电网的拓扑结构学习相邻节点量测数据相关性特征,提高了数据推演生成数

据的精度。

## 2.2 基于GINN的配电网稀疏量测数据推演模型

考虑一个无向图  $G=\{V,E,A\}$  用于表示配电网的拓扑结构, 其中  $V(G)=\{v_1,\dots,v_N\}$  为所有节点数  $N$  的集合,  $E(G)$  为边数  $L$  的集合, 边  $e_{ij}$  连接节点  $v_i$  和  $v_j$ ,  $A$  为邻接矩阵。将配电网的稀疏量测数据作为 GINN 的输入,  $X\in\mathbb{R}^{n\times k}$  表示包含  $k$  个量测潮流特征的  $n$  个节点量测值的输入特征矩阵, 其中潮流特征包括节点注入有功和无功功率、线路有功和无功功率、节点电压幅值和相角等。  $X_{ij}$  表示  $X$  的第  $(i,j)$  项,  $X_{i\cdot}$  和  $X_{\cdot j}$  分别表示矩阵  $X$  的第  $i$  行和第  $j$  列。给定一个二进制缺失掩码  $M\in\mathbb{R}^{n\times k}$ , 其中  $M_{ij}=1$  表示  $X_{ij}$  有量测,  $M_{ij}=0$  表示  $X_{ij}$  处量测缺失, 对于缺失的节点, 用零向量初始化。数据推演的目标就是预测当  $M_{ij}=0$  时缺失的特征  $X_{ij}$ , 即量测缺失节点的电压幅值和相角。

本文提出的基于 GINN 的配电网稀疏量测数据推演模型主要包括量测特征编码器和图编码器两个关键部分, 量测特征编码器旨在挖掘节点量测数据不同潮流特征间的关系, 图编码器旨在建立配电网节点间的拓扑关联。通过两个模块的共同作用实现了节点量测潮流特征的提取与更新, 从而生成缺失的量测数据, 具体结构如图 2 所示。

量测特征编码器将包含缺失项并以零填补的输入特征矩阵  $X$  及其对应的缺失掩码  $M$  作为输入, 生成  $\mathbb{R}^{n\times d}$  中的稠密特征矩阵  $Q$  作为输出。具体地, 对于输入特征矩阵的第  $i$  行量测  $X_{i\cdot}$ , 通过一个嵌入层将每个特征  $X_{ij}$  投影到一个  $d$  维特征空间。由于传统的嵌入层无法直接处理连续的数值特征, 因此首先将每个连续特征转化为独热编码, 即 0 和 1 的二进制向量。然后, 将这些独热向量与原始特征一起输入, 经过一层前馈神经网络进行处理。且该层网络的参数在所有连续特征之间是共享的。

然后, 使用 Transformer 网络学习节点量测不同潮流特征之间的关联。由于特征之间没有固有的顺序, 因此不采用位置嵌入。Transformer 模块由 MHA(多头自注意力层)、FF(前馈层)和 RC(残差连接)组成, 之后进行 LN(层归一化)。将所有特征输入到具有两个堆叠层的 Transformer 中, 然后应用 PL(池化层)来计算 Transformer 输出的

均值表示。Transformer 层和池化层都使用缺失掩码  $M$  分别作为序列注意力掩码和池化掩码, 以确保只有  $M_{ij}=1$  的  $X_{ij}$  才有助于量测  $X_{i\cdot}$  的最终表示。具体公式如下:

$$R_{i\cdot}=LN((MHA(EB(X_{i\cdot}),M_{i\cdot})+EB(X_{i\cdot}))) \quad (4)$$

$Q_{i\cdot}=PL(LN((MHA(R_{i\cdot},M_{i\cdot})+R_{i\cdot})),M_{i\cdot})$  (5)  
式中:  $EB(\cdot)$  为对输入  $X_{i\cdot}$  执行嵌入操作;  $R$  为经过层归一化得到的中间特征表示。

图编码器由图生成器和 GCN(图卷积网络)编码器组成。图生成器以量测特征编码器的输出  $Q\in\mathbb{R}^{n\times d}$  为输入, 输出一个邻接矩阵  $A\in\mathbb{R}^{n\times n}$ , 实现了节点间拓扑连接关系的构建。计算流程如下:

$$Q^*=\text{ReLU}(QW) \quad (6)$$

$$S_{ij}=\frac{Q_{i\cdot}^*Q_{\cdot j}^{*\text{T}}}{\|Q_{i\cdot}^*\|_2\|Q_{\cdot j}^*\|_2} \quad (7)$$

$$S_{ij}^*=\begin{cases} S_{ij}, S_{ij}>\epsilon \\ 0, S_{ij}\leq\epsilon \end{cases} \quad (8)$$

$$A=D^{-1/2}S^*D^{-1/2} \quad (9)$$

式中:  $W\in\mathbb{R}^{d\times d}$  为可学习的权重矩阵;  $\text{ReLU}(\cdot)$  为激活函数;  $S_{ij}$  为稠密邻接矩阵;  $S_{ij}^*$  为稀疏对称邻接矩阵;  $A$  为基于稀疏对称邻接矩阵  $S_{ij}^*$  的归一化矩阵, 其目的是利用阈值  $\epsilon$  对含噪声的量测值进行剪枝;  $D$  为  $S^*$  的度矩阵。

接下来, 基于邻接矩阵  $A$  的拓扑连接关系, 图编码器采用 GCN 层作为基本模块实现特征的提取与更新。GCN 能够直接处理图结构数据, 通过卷积操作将节点特征与其邻居节点的特征进行聚合, 从而捕捉图结构中的局部和全局信息。另外, GCN 可以有效地利用配电网的拓扑结构信息, 通过消息传递机制实现节点特征的更新, 计算效率较高, 适合处理大规模的图结构数据。与传统的图神经网络方法相比, GCN 的计算复杂度较低, 能够在保证性能的同时提高计算效率。具体地, 图编码器包含  $L$  个 GCN 层, 它将来自图生成器的  $A$  和来自量测特征编码器的  $Q$  作为输入, 并输出一个新的特征矩阵  $H$ , 该矩阵可用于填充任务。每个 GCN 层的操作基于如下方程:

$$H^{(l+1)}=f(H^{(l)},A) \quad (10)$$

$$f(H^{(l)},A)=\text{ReLU}(AH^{(l)}W^{(l)}) \quad (11)$$

式中:  $A$  为由图生成器计算得到的稀疏加权邻接

矩阵;  $H^{(0)} = Q$ , 为量测特征编码器的输出矩阵;  $W^{(l)} \in \mathbb{R}^{d \times d}$  为第  $l$  层的可训练权重矩阵。

最后, 根据量测特征编码器和图编码器的输出特征矩阵  $Q$  和  $H$ , 得到填充补齐后的节点特征矩阵  $\hat{Y}$ 。

$$\hat{Y} = \text{ReLU}((Q // H)W_1)W_2 \quad (12)$$

式中:  $W_1 \in \mathbb{R}^{2d \times d}$  和  $W_2 \in \mathbb{R}^{d \times k}$  为可训练权重矩阵。

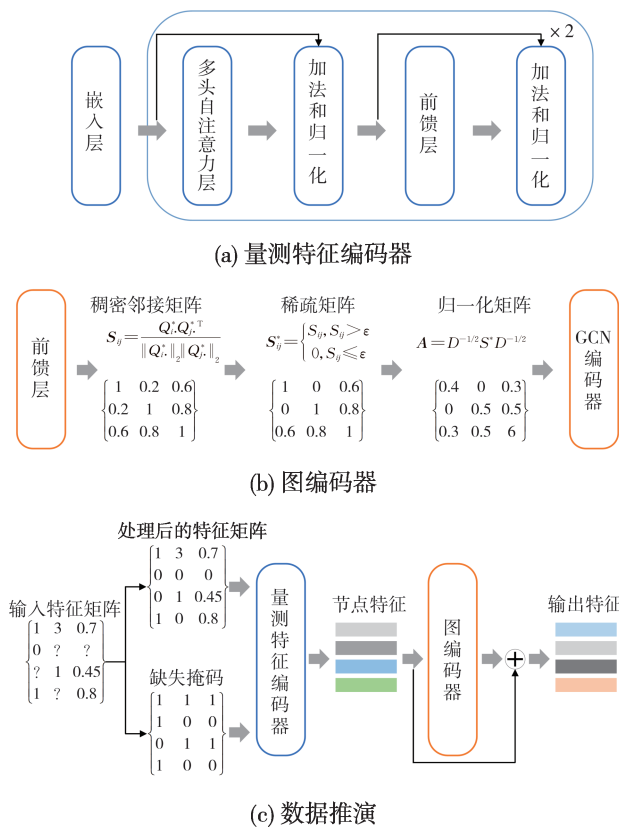


图 2 图填补神经网络结构  
Fig.2 Structure diagram of GINN

### 2.3 算法流程

本文提出了一种基于GINN的配电网稀疏量测数据推演方法, 该方法可以深度挖掘数据中的结构信息, 利用稀疏的量测数据实现缺失量测的推演补齐。数据推演的流程如图3所示。

1) 数据采集与预处理。首先, 通过仿真获得潮流数据作为实际值, 并在其基础上加入非高斯噪声生成量测数据。同时, 获取配电网的拓扑结构数据, 并建立用于表示拓扑连接关系的邻接矩阵。由于配电网通常是单相、两相、三相混合的情况, 配电网的数据通常存在不对称性。为了解

决数据的不对称性, 引入了相位归一化和相位补偿技术。相位归一化用于对每个相的数据进行标准化处理, 确保其值在同一量纲下。相位补偿则通过调整数据中的不平衡性, 解决由不同相混合造成的问题, 使得模型能够更好地处理不对称数据。接着, 基于量测数据和拓扑数据构建图数据库。为了减少数据量纲差异对模型训练的影响, 并防止过拟合, 对图数据进行归一化处理。最后, 将图数据集按照4:1的比例划分为训练集和测试集, 其中训练集用于训练GINN模型, 测试集用于评估GINN模型的推演性能。

2) GINN模型构建。设置GINN的网络结构和模型参数, 包括模型的层次结构、Transformer和GCN层数、激活函数、优化器、学习率等。

3) GINN模型离线训练。根据输入的历史图数据, 训练GINN数据推演模型, 输出配电网的节点电压幅值和相角数据。在每次训练中, 将训练集数据输入模型, 通过前向传播计算模型输出, 并根据推演值和真实值计算损失。通过反向传播计算损失梯度, 并使用随机梯度下降法迭代更新模型参数。训练完成后, 保存优化后的GINN模型。

4) 在线推演。根据配电网的实时量测数据和拓扑数据构建图数据, 并将其输入到训练好的GINN数据推演模型中, 即可推演出配电网的缺失数据。

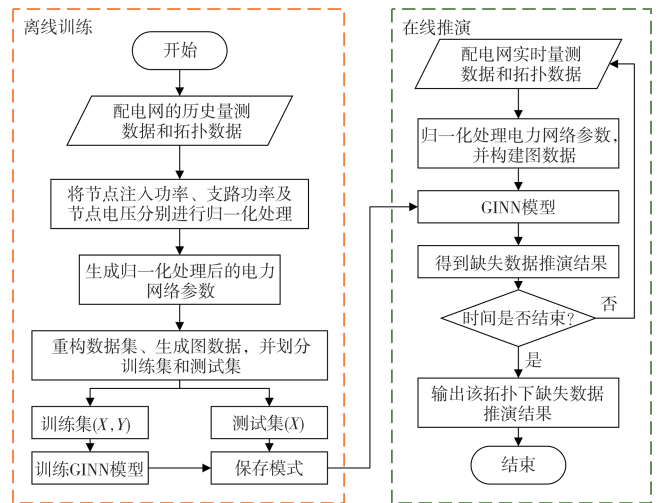


图 3 图填补神经网络数据推演流程  
Fig.3 Flowchart of GINN measurement inference

### 3 算例测试分析

#### 3.1 算例设计

本文使用电力系统仿真软件包MATPOWER进行仿真生成数据,并使用Python进行机器学习模型的训练和测试。对整个仿真期间的负荷数据进行潮流计算,并记录各种潮流结果,包括节点注入有功功率、无功功率、节点电压幅值、相角和线路潮流功率,并将其作为潮流真值。同时,在生成的数据中引入非高斯噪声,以模拟真实世界的的数据情况。本文采用IEEE 14、30、57和118节点系统进行实验,这些系统的主要差异体现在网络规模、节点数和线路数量上,从而影响了数据推演方法的效果。选择不同规模的IEEE标准测试系统是为了对比它们在不同网络规模和复杂性下的表现,从而验证本文所提方法的广泛适用性和鲁棒性。对于IEEE 14、30、57节点系统各自生成39 444组输入、输出张量样本。对于IEEE 118节点系统,生成17 520组输入、输出张量样本。将样本集的80%作为训练集,20%作为测试集。

对于IEEE 14、30、57和118节点系统,分别采用64、110、216和562个量测值。作为量测配置情况的示例,IEEE 14节点系统32个量测的位置如下:节点注入有功和无功量测取自节点{2,4,8,11,12,14},即 $2 \times 7 = 14$ 个量测;节点1的电压幅值和相角量测,即 $2 \times 1 = 2$ 个量测;支路有功和无功潮流量测为支路1—2、2—3、2—5、4—6、4—7、6—11、6—13、12—13,即 $2 \times 8 = 16$ 个量测。

本文测试环境为PC机,处理器为Intel(R)Core(TM)i7-8750H CPU@2.20 GHz,内存为8.0 GB。编程环境采用Python软件实现,所提方法通过Pytorch模块构建GINN模型。设置训练迭代1 000轮,每次迭代使用的小批量数据的大小为64,初始学习率设置为0.001,学习率每120轮下降一次。

#### 3.2 评价指标

使用以下评价指标将本文所提方法与其他配电网稀疏量测数据推演方法进行比较,其计算公式如下:

$$C_{\text{MAE}} = \frac{1}{mm} \sum_{k=1}^m \sum_{i=1}^n |\hat{x}_k^i - x_k^i| \quad (13)$$

$$C_{\text{RMSE}} = \sqrt{\frac{1}{mm} \sum_{k=1}^m \sum_{i=1}^n (\hat{x}_k^i - x_k^i)^2} \quad (14)$$

式中: $C_{\text{MAE}}$ 为平均绝对误差数值; $C_{\text{RMSE}}$ 为均方根误差数值; $m$ 为测试样本总数; $n$ 为推演电压或相角值的总数; $\hat{x}_k^i$ 为实际值; $x_k^i$ 为推演值。

#### 3.3 数据推演结果

将 $n$ 个测量值作为输入, $m$ 个电压幅值和 $m$ 个相角作为推演输出训练配电网稀疏量测数据推演模型,该模型可用于实际配电网的离线训练。训练好的模型可以在给定测量值的情况下实时推演配电网的当前状态。电压幅值以标么值(p.u.)的形式表示,相角用角度( $^\circ$ )表示。

将本文提出的用于配电网稀疏量测数据推演的GINN与WLS(加权最小二乘法)、KF(卡尔曼滤波法)、CNN(卷积神经网络)和LSTM(长短期记忆神经网络)进行比较。

表1所示为WLS、KF、CNN、LSTM和所提的GINN算法在IEEE标准节点系统下电压幅值推演的RMSE(均方根误差)和MAE(平均绝对误差)比较,单位为 $10^{-2}$  p.u.。评价指标值表明,所提的GINN模型捕捉了输入量测值与推演的电压幅值之间的真实关系,具有更好的推演效果。

表1 各算法在IEEE标准节点系统下的电压幅值推演性能对比

Table 1 Comparison of voltage amplitude inference performance of each algorithm under IEEE standard bus system

| 算法   | IEEE 14 |         | IEEE 30 |         | IEEE 57 |         | IEEE 118 |         |
|------|---------|---------|---------|---------|---------|---------|----------|---------|
|      | MAE     | RMSE    | MAE     | RMSE    | MAE     | RMSE    | MAE      | RMSE    |
| WLS  | 1.912 5 | 2.669 4 | 3.120 3 | 4.682 5 | 3.961 1 | 5.647 3 | 1.940 9  | 2.548 2 |
| KF   | 1.895 1 | 2.563 7 | 3.048 2 | 4.387 5 | 3.871 3 | 5.400 3 | 1.902 8  | 2.483 4 |
| CNN  | 1.474 1 | 2.237 6 | 2.901 7 | 4.236 4 | 3.466 1 | 5.158 9 | 1.452 8  | 1.942 5 |
| LSTM | 1.871 5 | 2.438 4 | 2.873 6 | 4.145 8 | 3.540 2 | 5.116 5 | 1.363 5  | 1.820 7 |
| GINN | 0.152 0 | 0.275 4 | 0.287 5 | 0.470 4 | 0.287 7 | 0.464 0 | 0.120 8  | 0.176 8 |

表2所示为WLS、KF、CNN、LSTM和所提的GINN算法在IEEE标准节点系统下电压相角推演的RMSE和MAE比较,单位为角度( $^\circ$ )。由表2可知,本文所提方法的数据推演结果更接近实际状态值。

为了验证对缺失量测数据的推演效果,图4—7分别给出了所提的GINN与WLS、CNN和LSTM算法在IEEE标准系统缺失量测节点处电压幅值和相角的推演结果,即缺失量测节点的电压

表2 各算法在IEEE标准节点系统下的电压相角推演性能对比

Table 2 Comparison of voltage phase angle inference performance of each algorithm under IEEE standard bus system

| 算法   | IEEE 14 |         | IEEE 30 |         | IEEE 57 |         | IEEE 118 |         |
|------|---------|---------|---------|---------|---------|---------|----------|---------|
|      | MAE     | RMSE    | MAE     | RMSE    | MAE     | RMSE    | MAE      | RMSE    |
| WLS  | 0.851 9 | 0.840 2 | 0.964 7 | 1.510 5 | 1.495 0 | 2.114 6 | 1.978 2  | 2.567 5 |
| KF   | 0.764 2 | 0.808 1 | 0.827 7 | 1.493 9 | 1.415 8 | 2.030 8 | 1.931 3  | 2.337 1 |
| CNN  | 0.460 2 | 0.653 0 | 0.729 8 | 1.255 3 | 1.252 8 | 1.854 4 | 1.633 9  | 2.079 4 |
| LSTM | 0.309 6 | 0.492 8 | 0.610 2 | 1.162 2 | 1.151 7 | 1.723 6 | 1.028 8  | 1.232 8 |
| GINN | 0.057 9 | 0.125 1 | 0.168 3 | 0.513 5 | 0.149 7 | 0.255 9 | 0.122 4  | 0.180 4 |

幅值-时间变化曲线和电压相角-时间变化曲线。图4—7表明，所提出的GINN推演结果更接近潮流真值，其推演的状态更接近实际状态。

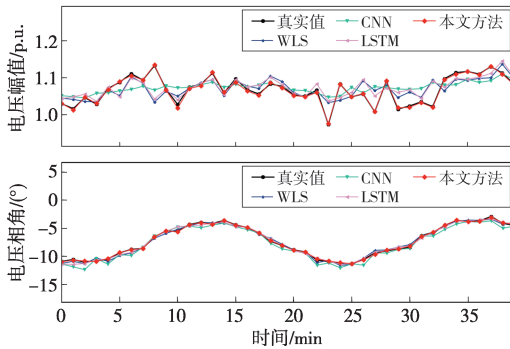


图4 IEEE 14节点系统缺失量测节点10的数据推演结果  
Fig.4 Inference result of missing measurement node 10 in the IEEE 14-bus system

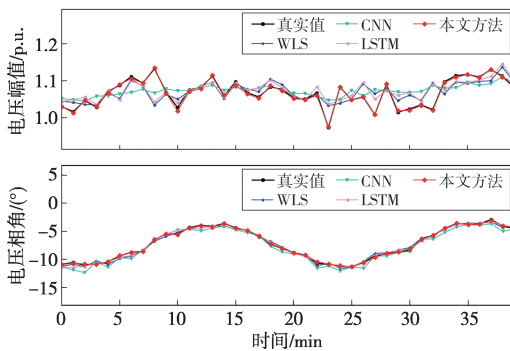


图5 IEEE 30节点系统缺失量测节点20的数据推演结果  
Fig.5 Inference result of missing measurement node 20 in the IEEE 30-bus system

### 3.4 抗差性能测试

为了验证数据推演模型对异常输入的抵抗能力，本文对不同数据异常情况下模型的推演性能进行了测试。将本文所提的数据推演模型分别在

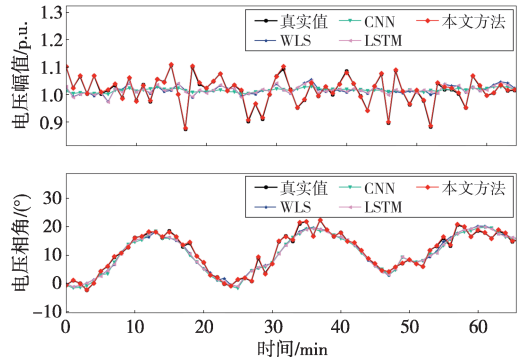


图6 IEEE 57节点系统缺失量测节点28的数据推演结果  
Fig.6 Inference result of missing measurement node 28 in the IEEE 57-bus system

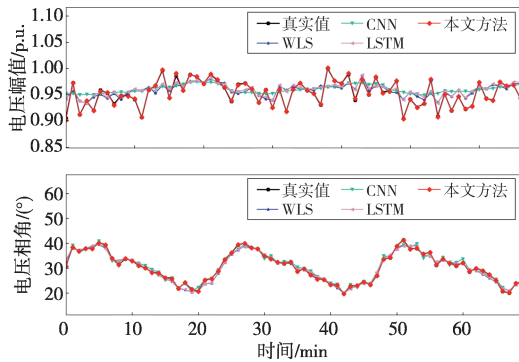


图7 IEEE 118节点系统缺失量测节点59的数据推演结果  
Fig.7 Inference result of missing measurement node 59 in the IEEE 118-bus system

不同实时量测噪声和不同量测缺失情况下测试推演算法的性能。

#### 3.4.1 噪声干扰时的性能

本文所提出的数据推演模型的训练环境为1%非高斯噪声。为了验证本文所提模型对异常数据的抗干扰能力，测试该模型面对2%、3%、4%的非高斯噪声输入时的推演精度，并与其他方法进行对比。以IEEE 30节点系统为例，采用电压幅值平均绝对误差和电压相角平均绝对误差作为评价指标，不同数据推演算法的推演精度如表3所示。

从表3可以看出，本文所提方法在非高斯噪声增大时依旧保持较高的推演精度。在1%非高斯噪声的训练环境中，所有方法均具有较好的推演性能。面对更强的噪声干扰时，本文方法仍具备更优的推演精度。以4%的非高斯噪声测试为例，与CNN相比，本文方法的电压幅值的推演精度提高

表 3 非高斯噪声实验结果

Table 3 Experimental results of non-Gaussian noise

| 算法   | 非高斯噪声占比<br>/ % | 电压幅值误差<br>/ p.u.      | 电压相角误差<br>/ (°) |
|------|----------------|-----------------------|-----------------|
| GINN | 1              | $0.25 \times 10^{-2}$ | 0.15            |
|      | 2              | $0.58 \times 10^{-2}$ | 0.74            |
|      | 3              | $0.91 \times 10^{-2}$ | 1.61            |
|      | 4              | $1.57 \times 10^{-2}$ | 2.38            |
| WLS  | 1              | $3.15 \times 10^{-2}$ | 0.98            |
|      | 2              | $3.44 \times 10^{-2}$ | 2.12            |
|      | 3              | $3.96 \times 10^{-2}$ | 3.81            |
|      | 4              | $4.46 \times 10^{-2}$ | 5.29            |
| KF   | 1              | $3.09 \times 10^{-2}$ | 0.94            |
|      | 2              | $3.32 \times 10^{-2}$ | 1.93            |
|      | 3              | $3.77 \times 10^{-2}$ | 3.74            |
|      | 4              | $4.21 \times 10^{-2}$ | 5.18            |
| CNN  | 1              | $2.90 \times 10^{-2}$ | 0.72            |
|      | 2              | $3.21 \times 10^{-2}$ | 1.75            |
|      | 3              | $3.48 \times 10^{-2}$ | 3.29            |
|      | 4              | $4.10 \times 10^{-2}$ | 4.90            |
| LSTM | 1              | $2.57 \times 10^{-2}$ | 0.56            |
|      | 2              | $2.82 \times 10^{-2}$ | 0.96            |
|      | 3              | $3.23 \times 10^{-2}$ | 1.82            |
|      | 4              | $3.49 \times 10^{-2}$ | 2.77            |

61.7%, 电压相角的推演精度提高 51.4%。

### 3.4.2 量测数据缺失时的性能

为了验证本文所提模型对数据缺失的适应性, 构建了 4 个数据缺失比例不同的数据集, 百分比分别为 5%、10%、15%、20%, 以测试数据缺失逐渐增多时的推演性能。以 IEEE 30 节点系统的数据推演结果为例, 推演结果如图 8 所示。

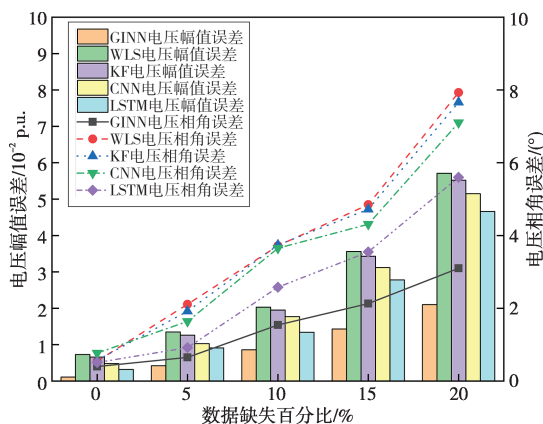


图 8 不同数据缺失比例下的数据推演结果

Fig.8 Inference result of data with different missing proportions

由图 8 可知, 量测数据的完整性对数据推演的精度起着重要作用。随着数据缺失比例的逐渐增大, WLS、KF、CNN 和 LSTM 的推演精度大幅下降, 而本文方法的推演精度下降趋势与其他方法相比较为缓慢。当数据缺失比例最高时, 相较于 WLS、KF、CNN 和 LSTM 方法, 本文所提方法在电压幅值的推演精度方面分别提高了 63.2%、62%、59.2% 和 57.7%。此时, WLS、KF、CNN 和 LSTM 的电压相角误差较未缺失数据时分别增加了 1 341%、1 243%、822% 和 998%, 而本文方法的电压相角误差较未缺失数据时仅增加了 675%。实验结果表明, 本文方法在面对量测数据缺失情况时表现出更强的鲁棒性。

### 3.5 推演效率对比

为了验证不同方法的数据推演效率, 将本文方法与其他方法进行对比。在小规模系统中, WLS 的计算效率较高, 但在大规模系统中, 计算复杂度急剧上升, 导致推演时间显著变长。而一般的基于状态估计的方法, 如 CNN, 通过并行计算和端到端训练实现高效处理, 在大规模系统中实现了较快的数据推演。本文所提方法相比于一般的基于状态估计的方法, 可以发掘数据中的结构信息, 通过图结构进行高效信息传递和并行计算, 因而在大规模系统中实现了更快的数据推演。表 4 给出了本文方法与其他方法在不同的 IEEE 节点系统下的计算时间对比。

表 4 各算法的推演效率

Table 4 Inference efficiency of various algorithms

| 测试系统      | 推演耗时/s  |         |         |
|-----------|---------|---------|---------|
|           | WLS     | CNN     | GINN    |
| IEEE 14   | 0.003 1 | 0.001 3 | 0.001 8 |
| IEEE 57   | 0.007 5 | 0.004 1 | 0.005 5 |
| IEEE 118  | 0.014 8 | 0.008 7 | 0.009 6 |
| IEEE 2383 | 0.308 6 | 0.027 4 | 0.014 2 |

由表 4 可知, 随着系统规模的增大, WLS 的推演耗时受系统规模影响较大, 尤其是在 IEEE 2383 节点系统中, 该方法已经难以满足数据推演在线应用的实时性要求。但本文算法的推演耗时并没有明显增长, 基本不受系统规模的影响。以 IEEE 2383 节点系统为例, 本文算法的推演效率较 WLS 和 CNN 分别提高了 95.39% 和 48.18%。综

上所述,本文算法大规模配电网数据推演中展现出了卓越的计算效率和实时性,具有更广泛的应用前景。

#### 4 结语

为了应对配电网量测设备部署不全、数据传输丢失等导致的量测稀疏问题,本文提出了一种基于GINN的配电网稀疏量测数据推演方法。GINN通过量测特征编码器和图编码器两个模块,分别捕捉节点量测不同特征间的关系和节点间的拓扑关联,实现了缺失量测数据的推演补齐。该方法利用历史量测数据和状态来训练推演模型,使其能够通过历史量测数据来推演出任意时刻缺失的数据。然后,基于多个IEEE基准系统进行仿真测试,测试结果表明,本文方法在量测稀疏条件下具有较高的推演精度和计算效率。与传统的数据推演方法相比,本文方法不仅提高了推演精度,还具有更强的鲁棒性和可解释性。

本研究也存在一些局限性。首先,仿真测试所使用的配电网系统为标准测试系统,实际应用中可能面临更多的不确定性和复杂性。另外,模型的性能在极端稀疏的条件下仍需进一步提升。未来的研究应在更多实际场景下对本文方法进行验证,以提升模型的泛化能力。同时,探索结合其他先进的机器学习技术来优化模型的结构和参数,进一步提高推演精度和效率。

#### 参考文献

- [1] NAYAK S, DWIVEDI D, BABU K V S M, et al. Data imputation using self attention based model for enhancing distribution grid monitoring and protection systems [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73:1-11.
- [2] SHAHBAZIAN R, GRECO S. Generative adversarial networks assist missing data imputation: a comprehensive survey and evaluation [J]. IEEE Access, 2023, 11: 88908-88928.
- [3] 张新鹤,何桂雄,梁琛,等.基于分割区域的配电网异常线损数据辨识与修正[J].浙江电力,2023,42(10):90-100.  
ZHANG Xinhe, HE Guixiong, LIANG Chen, et al. Identification and correction of abnormal line loss data in distribution networks based on segmented regions [J]. Zhejiang Electric Power, 2023, 42(10):90-100.
- [4] 张汪洋,樊艳芳,侯俊杰,等.基于集成深度神经网络的配电网分布式状态估计方法[J].电力系统保护与控制,2024,52(3):128-140.  
ZHANG Wangyang, FAN Yanfang, HOU Junjie, et al. Distribution network distributed state estimation method based on an integrated deep neural network [J]. Power System Protection and Control, 2024, 52(3):128-140.
- [5] LEI C, BU S Q, WANG Q G, et al. Observability defense-constrained distribution network reconfiguration for cyber-physical security enhancement [J]. IEEE Transactions on Smart Grid, 2024, 15(2):2379-2382.
- [6] LI Y Y, SONG L D, HU Y, et al. Load profile inpainting for missing load data restoration and baseline estimation [J]. IEEE Transactions on Smart Grid, 2024, 15(2):2251-2260.
- [7] 吴莉艳,孙开元,陈坤,等.基于CNN-LSSVM的电力系统虚假数据攻击检测[J].浙江电力,2024,43(11):90-96.  
WU Liyan, SUN Kaiyuan, CHEN Kun, et al. Detection of false data injection attacks against power systems using a CNN-LSSVM model [J]. Zhejiang Electric Power, 2024, 43(11):90-96.
- [8] 周远翔,林孟龙,陈健宁,等.基于自注意力生成对抗网络的电力设备在线监测缺失数据填补[J].高电压技术,2023,49(5):1795-1809.  
ZHOU Yuanxiang, LIN Menglong, CHEN Jianning, et al. Missing data imputation for online monitoring of power equipment based on self-attention generative adversarial networks [J]. High Voltage Engineering, 2023, 49(5):1795-1809.
- [9] HABIB B, ISUFI E, VAN BREDA W, et al. Deep Statistical Solver for Distribution System State Estimation [J]. IEEE Transactions on Power Systems, 2024, 39(2):4039-4050.
- [10] 赵洪山,寿佩瑶,马利波.低压台区缺失数据的张量补全方法[J].中国电机工程学报,2020,40(22):7328-7337.  
ZHAO Hongshan, SHOU Peiyao, MA Libo. A tensor completion method of missing data in transformer district [J]. Proceedings of the CSEE, 2020, 40(22):7328-7337.
- [11] APRILLIA H, YANG H T, HUANG C-M. Statistical load forecasting using optimal quantile regression random forest and risk assessment index [J]. IEEE Transactions on Smart Grid, 2021, 12(2):1467-1480.
- [12] WANG Y, WANG Y Q, SUN Y H, et al. Resilient dynamic state estimation for multi-machine power system with partial missing measurements [J]. IEEE Transactions on Power Systems, 2024, 39(2):3299-3309.
- [13] XU D L, XU J J, QIAN C, et al. A pseudo-measurement modelling strategy for active distribution networks considering uncertainty of DGs [J]. Protection and Control of Modern Power Systems, 2024, 9(5):1-15.

- [14] 赵友国,刘尚伟,王冠中,等.基于正交分解的电力系统状态估计可观性分析[J].浙江电力,2021,40(7):1-5.  
ZHAO Youguo, LIU Shangwei, WANG Guanzhong, et al. Observability analysis of power system state estimation based on orthogonal decomposition [J]. Zhejiang Electric Power, 2021, 40(7): 1-5.
- [15] WANG Y J, GU J, YUAN L. Distribution network state estimation based on attention-enhanced recurrent neural network pseudo-measurement modeling [J]. Protection and Control of Modern Power Systems, 2023, 8: 31.
- [16] 王守相,陈海文,潘志新,等.采用改进生成式对抗网络的电力系统量测缺失数据重建方法[J].中国电机工程学报,2019,39(1):56-64.  
WANG Shouxiang, CHEN Haiwen, PAN Zhixin, et al. A reconstruction method for missing data in power system measurement using an improved generative adversarial network [J]. Proceedings of the CSEE, 2019, 39(1): 56-64.
- [17] 王子馨,胡俊杰,刘宝柱.基于长短期记忆网络的电力系统量测缺失数据恢复方法[J].电力建设,2021,42(5):1-8.  
WANG Zixin, HU Junjie, LIU Baozhu. Recovery method for missing measurement data of power systems based on long short-term memory networks [J]. Electric Power Construction, 2021, 42(5): 1-8.
- [18] 俞文帅,张晓华,卫志农,等.基于深度神经网络的电力系统快速状态估计[J].电网技术,2021,45(7):2551-2561.  
YU Wenshuai, ZHANG Xiaohua, WEI Zhinong, et al. Fast state estimation for power system based on deep neural network [J]. Power System Technology, 2021, 45(7): 2551-2561.
- [19] 杨隽雯,尚磊,叶欣智,等.考虑配电网故障重构的电压薄弱节点辨识方法[J].电力工程技术,2025,44(1):39-49.  
YANG Juanwen, SHANG Lei, YE Xinzhi, et al. A voltage weak node identification method considering fault reconstruction in distribution networks [J]. Jiangsu Electrical Engineering, 2025, 44(1): 39-49.
- [20] 呼士召,刘仕琦,江健健,等.配电网多负荷场景下移动储能两阶段经济运行策略[J].智慧电力,2025,53(1):45-53.  
HU Shizhao, LIU Shiqi, JIANG Jianjian, et al. A two-stage economic operation strategy for mobile energy storage considering different load scenarios in distribution network [J]. Smart Power, 2025, 53(1): 45-53.
- [21] 蔡木良,范瑞祥,贺贵东,等.基于时间规律聚合电动汽车的有源配电网源荷储协同调控[J].智慧电力,2025,53(1):54-61.  
CAI Muliang, FAN Ruixiang, HE Guidong, et al. Source-load-storage coordinated scheduling for active distribution network based on aggregation of electric vehicles with time pattern [J]. Smart Power, 2025, 53(1): 54-61.
- [22] 许超,李永刚,张书伟,等.考虑配电网三相电压特征的IHPO-CSSVM电压暂降源识别[J].电力需求侧管理,2025,27(1):101-106.  
XU Chao, LI Yonggang, ZHANG Shuwei, et al. Voltage sag source identification using IHPO-CSSVM with consideration of three-phase voltage characteristics on the distribution network [J]. Power Demand Side Management, 2025, 27(1): 101-106.
- [23] 张祥龙,袁兆祥,董树锋,等.考虑节点碳排放强度的输配协同多目标最优潮流[J].电力建设,2025,46(7):123-132.  
ZHANG Xianglong, YUAN Zhaoxiang, DONG Shufeng, et al. Multi-objective optimal power flow of integrated transmission and distribution network considering node carbon emission intensity [J]. Electric Power Construction, 2025, 46(7): 123-132.
- [24] 陈彬,廖锦霖.极端天气下多灾害地区配电网防灾能力提升技术综述及展望[J].电力建设,2025,46(1):107-121.  
CHEN Bin, LIAO Jinlin. Technology for improving distribution network disaster-prevention capabilities for intercurrent natural disasters areas under extreme weather: review and prospect [J]. Electric Power Construction, 2025, 46(1): 107-121.

收稿日期:2025-01-05;修回日期:2025-02-01

作者简介:

李企洲(2000),男,硕士研究生,研究方向为配电网状态推演。

赵健(1990),男,博士,教授,研究方向为智能配用电、灵活性资源调度、分布式发电管控。(通信作者)

(本文编辑:徐 晗)