

基于注意力机制和卷积神经网络的 配电网拓扑辨识

杨秀¹, 蒋家富¹, 刘方¹, 田英杰², 李凡², 吴裔²

(1. 上海电力大学 电气工程学院, 上海市 杨浦区 200090;

2. 国网上海市电力公司电力科学研究院, 上海市 虹口区 200437)

Distribution Network Topology Identification Based on Attention Mechanism and Convolutional Neural Network

YANG Xiu¹, JIANG Jiafu¹, LIU Fang¹, TIAN Yingjie², LI Fan², WU Yi²

(1. Shanghai University of Electric Power, Yangpu District, Shanghai 200090, China;

2. Electric Power Research Institute of SG Shanghai Electric Power Company, Hongkou District, Shanghai 200437, China)

ABSTRACT: In view of the frequent changes of distribution network topology and the difficulty of obtaining the topology structure in real time, a distribution network topology identification method based on attention mechanism and convolutional neural network (ACNN) is proposed. The convolution neural network (CNN) is used to mine the relationship between measurement information and distribution network topology, and learn its mapping rules; Considering the problem of insufficient number of advanced measurement devices such as phasor measurement unit (PMU) and micro phasor measurement unit (μ PMU) installed in the current distribution network, the attention mechanism is integrated into the hidden layer of convolutional neural network to enhance the robustness of the model; The dimension of feature data set is reduced by random forest algorithm to reduce the time and space complexity of the model; Finally, numerical examples are carried out based on IEEE 33 node distribution network and PG & E69 node distribution network to verify the feasibility and superiority of the method, and to test the possibility of topology identification using fewer features. The results show that the proposed method has good superiority, robustness and strong generalization ability, Distribution network topology identification can be realized when only a small amount of time section measurement data is provided, and it is also applicable to radial network and ring network.

KEY WORDS: distribution network; attention mechanism; convolutional neural network; random forest; topology identification

摘要: 针对当前配电网拓扑变化频繁, 拓扑结构实时获取困难等问题, 文章提出基于注意力机制和卷积神经网络的配电网拓扑辨识方法。首先利用卷积神经网络挖掘量测信息和配电网拓扑结构之间的关系, 学习其映射规则; 考虑当前配电网中同步相量测量装置 (phasor measurement unit, PMU) 和微型同步相量测量装置 (micro phasor measurement unit, μ PMU) 等高级量测设备安装数量不足导致获取量测数据质量不高的问题, 在卷积神经网络隐藏层中融入注意力机制, 以增强模型鲁棒性; 通过随机森林算法对特征数据集进行降维, 降低模型时、空复杂度; 最后, 分别基于 IEEE 33 节点配电网和 PG&E69 节点配电网开展算例分析, 以验证方法的可行性和优越性, 并检验利用更少特征进行拓扑辨识的可能性。结果表明: 所提方法具有良好优越性和鲁棒性, 泛化能力强, 在仅提供少量时间断面量测数据情况下便可实现配电网拓扑辨识, 且对于辐射网和含环网络同样适用。

关键词: 配电网; 注意力机制; 卷积神经网络; 随机森林; 拓扑辨识

DOI: 10.13335/j.1000-3673.pst.2021.2538

0 引言

随着社会不断发展, 人类对电能供应要求不断提高。配电网作为现代社会最重要的能源基础设施之一, 也是保障用户高质量供电的“最后一公里”, 使得其稳定、安全运行对于保障用户可靠用电具有极其重要的影响^[1]。

随着新型电力系统不断发展建设, 分布式电源装机规模不断扩大, 电动汽车等新型负荷在用户侧

基金项目: 上海市科委电力人工智能工程技术研究中心项目 (19DZ2252800); 上海市科委青年扬帆计划 (21YF1414600); 上海市教委青年教师培训计划 (ZZDL20001)。

Project Supported by Electric Power Artificial Intelligence Engineering Technology Research Center of Shanghai Science and Technology Commission (19DZ2252800); Project Supported by Shanghai Sailing Program (21YF1414600); Shanghai Youth Teacher Training Program (ZZDL20001)。

占比日趋增加^[2]。为应对“源-荷”双侧不确定性导致的潮流高频幅波动和越限,配电网馈线开关动作将愈加频繁,导致网络结构也随之变化,运行方式愈加多样化、复杂化,给配电网安全稳定控制及经济运行分析带来全新挑战^[3-4]。精确的拓扑结构是配电系统进行潮流计算、状态估计、整定计算等分析应用的基础,也是规划运行、安全稳定控制的前提。然而,相较于输电网,配电网中监控及量测设备数量远不能满足需求,导致配电网结构大多数情况下难以实时准确获取^[5],且目前尚未进行拓扑校验工作,使得开展配电网拓扑辨识方法研究,基于当前量测数据实现拓扑精确辨识成为亟需解决的关键问题。

目前,国内外学者已对配电网拓扑辨识展开初步探索。根据辨识方法原理主要可分为传统辨识方法和人工智能辨识方法2类。在配电网拓扑传统辨识方法研究方面,文献[6]在传统矩阵法基础上进行改进,利用对称消去法降低邻接矩阵的阶数,并将对称性应用于求连接矩阵过程中,从而减少拓扑辨识计算量,但其对量测数据质量要求较高;文献[7]利用深度优先算法搜索 bridge 线路,基于混合整数二次优化获取初步拓扑结构,并采用树生成算法将无向图生成树,通过遍历树获取与潮流匹配评价指标最近的结构作为最终拓扑,原理清晰且易于实现,但由于忽略线损且对数据质量要求高,一定程度上限制了其适用性;文献[8]在传统互信息方法基础上构建双层拓扑辨识算法;文献[9]利用量测数据近似实时估计动态雅各比矩阵实现拓扑辨识;文献[10]根据电压幅值、相角量测量提出一种投票拓扑检测方法,但均需要足够数量的 μ PMU 或 PMU 等高级量测设备支撑,受限于设备成本昂贵,短期内尚难以满足实际需求;文献[11-12]基于功率方程和电网潮流计算方法建立物理模型,引入线性回归和雅各比矩阵进行迭代求解,实现基于量测数据的拓扑识别,由于要求各节点均配置有量测装置,对量测数据质量要求较高,一定程度上限制了其工程应用性;文献[13-15]均采用相关性方法判别馈线所属,再利用电压幅值随着潮流方向逐渐降低,确定负荷所属位置,但由于相关性阈值难以确定导致识别误差较大,且对未来分布式电源不断接入潮流呈现双向流动网络适应性尚且不足;文献[16]基于电压曲线相似性理论建立拓扑辨识算法,效果较为明显,但由于要求数据为连续量测数据,现实量测中通讯及测量装置难以做到,一定程度上限制其推广应用。此外,文献[7,11-16]适用对象仅为辐射状网

络结构,未来随着分布式可再生能源渗透率不断提升,以及对供电质量要求不断提高,“钻石型”、“花瓣型”等新型配电网发展建设成为趋势,配电网含环网络运行成为保障可靠供电的重要手段。然而,上述研究尚无法满足含环网络配电网拓扑辨识需求,且相关研究尚且鲜见。

在基于人工智能的配电网拓扑辨识方面,随着以深度学习为代表的人工智能技术在算法和算力方面得以突破,在电力系统优化分析中也得以广泛应用^[17],通过“数据+模型”混合驱动方式开展配电网拓扑辨识研究成为新方向。高泽璞等^[18]利用台区、配变、馈线、用户电表之间的关系构建知识图谱对配网进行拓扑辨识,一定程度上满足应用需求,但尚难以适用于结构变化频繁的配网;裴宇婷等^[19]基于深度神经网络构建拓扑辨识方法,辨识精度较高,若进一步解决参数调节复杂及容易陷入局部最优等问题,将更加满足工程适用性。孙伟等^[20]基于图卷积神经网络建立微电网拓扑辨识方法,将网络结构转化为线图,基于图卷积神经网络构建分类器从而实现拓扑辨识,其通过直接对图结构数据进行计算大幅度提升计算速率,若进一步从节点特征聚焦角度入手提升辨识精度,将更具备实际应用潜力。此外,根据配电系统实际运行特性,对于某区域确定的配电网,在正常运行情况下的拓扑结构为有限集,而系统潮流变量中蕴含与拓扑结构有关的丰富信息,可通过深度学习等方法充分挖掘量测数据与配网拓扑之间关联关系,可实现在有限量测数据和不确定信息下的拓扑辨识^[21]。

针对传统配电网拓扑辨识方法工程适用性方面的不足及人工智能技术应用新方向,本文将注意力机制和卷积神经网络结合,提出基于深度学习的配电网拓扑辨识方法。首先,基于随机森林算法进行特征选择,通过特征降维降低模型的时间、空间复杂度,并引入结合注意力机制的卷积神经网络充分挖掘量测数据与配网结构之间的联系;然后,利用注意力机制完成特征权重分配,在避免算力损耗的同时提升模型的鲁棒性,从而实现仅需关键节点断面量测数据情况下即可在线快速辨识出当前配网拓扑,且不仅适用于辐射网络,对弱环网也同样适用;最后,通过 IEEE 33 节点配电网和 PG&E69 节点配电网算例开展仿真分析,验证本文所提方法的可行性与优越性。

1 配电网拓扑辨识框架

1.1 主框架设计

本文拟构建的配电网拓扑辨识框架如图1所示,

主要包含 3 个部分：训练数据集的构建、离线训练和在线应用。其中，训练数据集构建主要包括 3 个阶段：量测数据获取、数据预处理以及特征筛选。首先，通过配电网 SCADA 系统获取历史断面量测数据及对应拓扑结构；然后对量测数据进行相应预处理及特征筛选；最后，将筛选出的特征类别与对应拓扑结构一

一对应，构建训练数据集。离线训练阶段主要将训练数据集输入到已构建的待训练配电网拓扑辨识模型中，进行模型自训练，学习量测数据集与拓扑结构之间的映射关系，最终得到训练后的模型。在线应用阶段，只需将实时断面量测数据输入到训练好的模型中，便可对当前配网拓扑结构进行辨识。

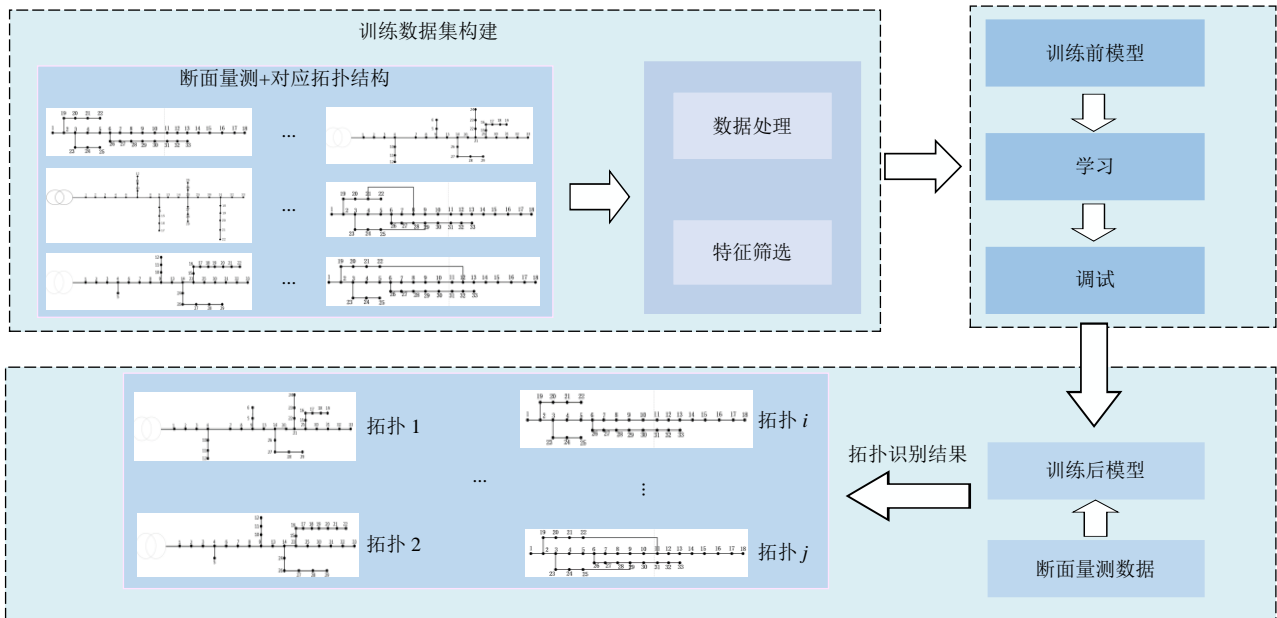


图 1 配电网拓扑辨识总体框架

Fig. 1 Overall framework of distribution network topology identification

1.2 数据获取及处理

1.2.1 数据获取

电力系统的潮流方程为

$$\begin{cases} P_i - U_i \sum_{j \in i} U_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) = 0 \\ Q_i - U_i \sum_{j \in i} U_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) = 0 \end{cases} \quad (1)$$

式中： P_i 、 Q_i 和 U_i 分别代表节点 i 的注入有功功率、注入无功功率和电压幅值； θ_{ij} 为节点 i 、 j 间的电压相角差。对于某确定配电网，只需确定上述 4 组变量中的任意 2 组便可通过求解潮流方程确定该配电网当前状态。

实际配电网中，电网结构庞大，量测装置有限，目前容易获取的量测数据主要有节点电压幅值、节点注入功率等量测数据，同时由于量测设备精度及通信原因，当前量测数据精度不高，且难以实现时间同步性。PMU 和 μ PMU 等高级量测设备应运而生，不仅能做到量测数据同时标注，且数据量测误差也大幅度降低，PMU 总矢量误差(total vector error, TVE)可达到 0.05%以下，而 μ PMU 甚至可达到 0.01%以下，但受限于成本，在实际配电网中运用还相对较少，仅在关键节点处安装。对此，本文选用较容易得到的断面量测数据：节点电压幅值 U

和节点注入有功功率 P ，作为初始特征类别。通过 SCADA 系统采集负荷和 DG 不同情况下节点电压幅值和节点注入有功功率的同时，记录对应网络拓扑，设定拓扑标签，构成训练数据集。

1.2.2 数据归一化

目前我国电网软硬件设施达到世界一流水平，配电网具有较强稳定性，负荷间距离较近使得线损相对较小，相邻节点电压幅值差异不明显。对于同一节点，在不同拓扑结构下，其电压幅值标幺值较接近，且变化不明显，若采用原始量测数据直接对配电网拓扑辨识模型进行训练，会增加训练难度，且影响拓扑辨识效果。为此，本文将电压幅值量测数据进行归一化处理，以增加训练集中数据差异性。归一化计算方法为

$$u_{\text{norm}} = \frac{u - u_{\min}}{u_{\max} - u_{\min}} \quad (2)$$

式中： u 和 u_{norm} 分别为该节点该归一化前后的节点电压幅值； u_{\min} 和 u_{\max} 分别为该节点处历史电压幅值量测数据的最小值和最大值。

1.2.3 数据缺失值处理

由于数据采集设备和通信线路容易发生故障导致采集量测数据时部分数据丢失，影响后续数据

分析与处理。对此,本文对于缺失部分数据的训练样本,直接丢弃并重新采集数据;对于缺失部分数据的测试样本,则进行数据填补,以保证拓扑辨识可正常执行。

同一拓扑下的量测样本,受潮流方程约束,相邻节点电压分布规律趋同,仅存在电压幅值差异,如图2所示,节点A、B、C是相邻节点,其电压幅值曲线变化规律相似,仅存在大小差异,可通过相邻节点电压幅值连续断面量测数据,生成该节点缺失数据。

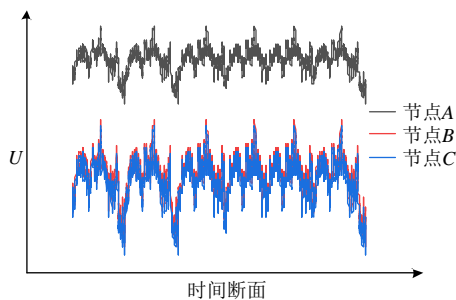


图2 电压曲线相似性

Fig. 2 Similarity of voltage curve

故 i 节点 t 时刻的缺失数据可用式(3)计算:

$$u_{i,t} = u_{i,t-1} + \frac{\sum_{j=1, j \neq i}^n (u_{j,t} - u_{j,t-1})}{n-1} \quad (3)$$

式中: $u_{i,t}$ 表示节点 i 在 t 时刻的电压幅值; n 表示与节点 i 同一条支路上的节点总数。可通过下式计算插补准确率(n_{pre}),从而对缺失处插补后的有效性和精度进行评估。

$$n_{pre} = 1 - \left| \frac{x - x_{rel}}{x_{rel}} \right| \times 100\% \quad (4)$$

式中: x 为插补后的值; x_{rel} 为真实值。通过选出多个时刻的各节点量测数据,将某时刻的部分接单量测数据去除,采用本方法对其进行填补,发现填补数据与原始数据误差不明显,精度维持在 95% 以上,且数据量越多,插补后数据与真实数据间的差异越小,即准确率越高。

1.2.4 特征贡献度计算

考虑到同时使用节点电压幅值和节点有功注入等量测数据作为训练特征数据集容易造成模型计算复杂度和空间复杂度较高,需要对数据集做降维处理,同时要为后续研究提供特征重要度的排序情况。考虑到随机森林的独特结构和机制,本文采用随机森林方法分别计算各特征类别对配网拓扑辨识的特征贡献度,实现特征类别筛选,达到降维效果,同时可以解决特征集由于空间相关性造成的耦合问题。随机森林由多个决策树组成的,其计算

特征重要度^[22]公式为

$$n_{importance} = \frac{1}{N} \sum_{i=1}^N |e_{OOB2} - e_{OOB1}| \quad (5)$$

式中: N 代表森林中决策树的数量;对每棵决策树,选择相应的袋外数据(out of bag, OOB)计算袋外数据误差,记作 e_{OOB1} ,随机对袋外数据所有样本的特征 X 加入噪声干扰(可随机改变样本在特征 X 处的值),再次计算袋外数据误差,记作 e_{OOB2} 。该数值能够说明特征的重要性,因为加入随机噪声后,袋外数据准确率大幅度下降(即 e_{OOB2} 上升),即该特征对于样本预测结果影响显著,进而说明该特征重要程度较高。

2 结合注意力机制的一维卷积神经网络分类器的构建

一般而言,配电网拓扑结构可以用导纳矩阵表征,而导纳矩阵影响系统潮流变化,使得潮流变量与配网结构间存在某种联系或规则。实际工程中,配网支路开关数量有限,故对于某特定配电网,其改变后的结构种类数为有限集,可通过深度学习等方法挖掘量测数据和配网拓扑结构之间的关系^[23],将配电网拓扑辨识转换为机器学习多分类问题。

2.1 一维卷积神经网络

CNN 是最早由 LeCun^[24]等受哺乳动物视觉皮层感受野启发提出的一种深度学习框架,也是一种有效用于特征提取和模式识别的前馈神经网络^[25]。一维卷积神经网络(one-dimensional convolutional neural networks, 1D-CNN)由输入层、卷积层、池化层、全连接层及输出层组成。其中,卷积层的功能是对输入数据进行特征提取,各卷积层根据卷积核工作方式有规律的处理输入数据;池化层通过舍弃次要特征突出重要特征的处理过程,实现数据降维和特征再提取功能。随着技术不断成熟,卷积神经网络凭借其参数共享和连接稀疏性,使得其调参容易,特征获取能力强,且不易陷入局部最优,应用在分类任务上表现出更加出色的效果。

2.2 注意力机制原理

注意力机制借鉴人类大脑的信号处理机制,通过快速对所有特征进行扫描,获得需要重点关注的特征类别,并根据特征类别的关键程度分配相应注意力权重,使大脑可利用有限资源处理庞大信息^[26]。在卷积神经网络中应用体现在生成每张特征图(feature map)的重要性存在差异,注意力机制的核心目标是通过计算得到每张 feature map 间的重要性差异,根据其重要度分配计算资源,并利用执

行效果反向指导 feature map 权重更新，最终高效、准确地完成任务。注意力模块实现原理如图 3 所示。

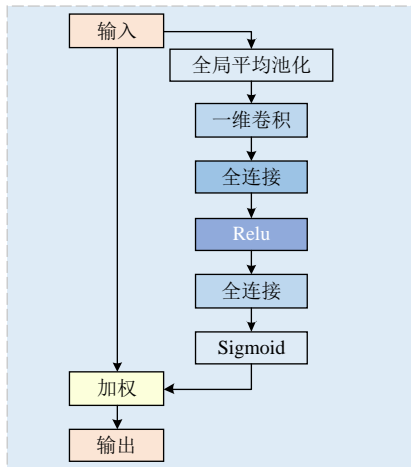


图 3 注意力机制原理

Fig. 3 Schematic of attention mechanism

具体实施方法如下^[27]：首先，对卷积得到的每张 feature map 进行全局平均池化操作，将每个 feature map 挤压成一个实数，如式(6)所示，将每张 feature map 挤压后的实数组合成一个向量，即每个特征的权重；得到权重向量后，通过全连接、Relu 激活函数、sigmoid 激活函数如式(7)所示，将权重赋予每个特征类别，得到加权后的 feature map；最后，通过式(8)指导 feature map 向有利于识别任务的方向更新。

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(ij) \quad (6)$$

$$s = \sigma(W_2 \delta(W_1 z)) \quad (7)$$

$$x_c = s_c u_c \quad (8)$$

式中： H 和 W 代表 feature map 的长和宽； u_c 代表卷积后的结果； z_c 代表每张 feature map 的重要度； s_c 为所有 feature map 的权重向量； σ 代表 Relu 激活函数； δ 代表 Sigmoid 激活函数； W_1 和 W_2 为 2 个不同的全连接操作。

2.3 结合注意力机制的一维卷积神经网络构建

针对配电网拓扑辨识问题，本文设计的结合注意力机制的一维卷积神经网络结构，如图 4 所示。其中：卷积神经网络的输入层神经元个数设置为输入特征个数；输出层神经元个数设置为拓扑结构类别个数；在输入层之后增加 1 层注意力层，通过该隐藏层将计算资源按照特征类别重要度进行分配。

在构建神经网络中，激活函数为首要考虑的参数。激活函数种类多样，常见的有 Relu、Sigmoid 和 tanh 激活函数。考虑到本文一维卷积神经网络应用主要功能是多分类，对此选用 tanh 函数作为卷积神经网络隐藏层的激活函数。

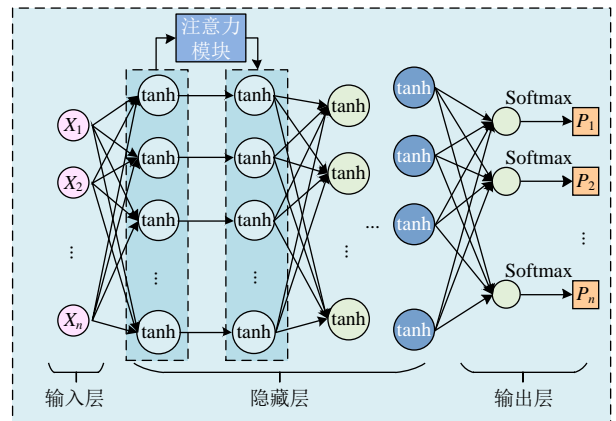


图 4 结合注意力机制的一维卷积神经网络结构

Fig. 4 Structure diagram of ACNN

池化层选用最大池化函数，同时考虑到网络结构中各节点对于拓扑辨识的重要性存在差异，因此在卷积层充分挖掘特征生成 feature map 后，引入注意力机制，对每张 feature map 根据其对拓扑辨识任务的重要性赋予权重，以提高模型辨识准确率和鲁棒性。输出层结果通过归一化指数函数 Softmax 输出，其输出为属于拓扑数据库中每种拓扑结构的概率，且所有概率之和为 1，因此卷积神经网络的输出是一个 n 维向量，向量中第 i 个元素代表该样本属于第 i 类拓扑的概率。损失函数选用 categorical_crossentropy 函数。应用损失函数计算实际输出与期望输出之间的距离，网络模型训练目标为最小化 categorical_crossentropy 损失函数。

网络的隐藏层数及每层神经元个数采用试错法确定，并采用十折交叉验证方法检验模型分类能力以验证隐藏层数和每层神经元个数设置的合理性。十折交叉验证方法将数据集随机分成 10 份，每次取其中 9 份做训练集对模型进行训练，另外 1 份作为测试集，对模型训练结果进行测试，取 10 次测试集分类正确率的平均值作为最终测试结果。

2.4 量测噪声处理

本文方法以配电网实时量测数据为基础，量测设备精度将显著地影响拓扑识别准确性，目前受限于量测设备及通信条件，数据中出现噪声不可避免，若不对噪声进行处理，直接训练不仅会加重计算任务，同时会造成计算误差叠加过大。若数据仅含少量噪声，卷积神经网络得益于本身结构特性，具有天然滤波作用，可对少量噪声进行滤波^[21]，且将注意力机制融入卷积神经网络将进一步提高模型抗噪能力；若数据中含有大量噪声，将显著降低算法收敛速度，甚至对训练生成模型也会有很大的副作用，例如模型过度学习数据集中噪声造成过拟合等。因此，需要在构建训练数据集之前，对量测

数据进行降噪处理。

2.5 模型评价指标

任何机器学习的模型效果通常有特定指标来进行评价。在使用机器学习时,通常首先将数据集按照一定比例进行拆分获得训练集和测试集,然后用训练集对模型进行训练,最后再用测试集对模型的效果进行测试。由于本文训练模型功能为多分类,可采用准确率(accuracy, ACC)来分析模型总体分类效果,并采用混淆矩阵对模型分析单个样本类别分类效果,并直观展示模型在单个类别分类的效果,在得到混淆矩阵基础上计算其查准率(precision, PRE)和召回率(recall, REC),其计算公式分别为

$$n_{ACC} = \frac{T}{N} \quad (9)$$

$$n_{PRE} = \frac{n_{TP}}{n_{TP} + n_{FP}} \quad (10)$$

$$n_{REC} = \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (11)$$

式中: T 是所有分类中正确的样本数量; N 是所有样本总数量; n_{TP} 表示分类为正的样本数量; n_{FP} 表示分类为正的负样本数量; n_{FN} 表示分类为负的正样本数量。此处的正类和负类是相对的,以多分类为例,假设有 4 个类别: A、B、C、D,在做分类任务时,当设置 A 为正类时,剩余 B、C、D 则均可看作负类。

2.6 未知拓扑处理机制

模型实现拓扑辨识过程中,可能会遇到当前拓扑不包含在已知拓扑库中的情况,因此进行拓扑辨识之前需要计算待辨识数据样本与每个训练集样本之间的相关性系数^[28],并设定相关性阈值,若待辨识数据与训练集样本最大相关性系数低于该阈值,则认为该拓扑结构为未知拓扑结构,应人工采集当前量测数据和对应拓扑结构,输入模型中进行增量训练。

3 算例分析

为了验证本文方法的有效性和优越性,分别选取 IEEE 33 节点配电网和 PG&E69 节点配电网为研究对象进行测试。

3.1 模拟量测数据

模型训练中所有节点的电压幅值、注入有功功率等量测数据由 MATLAB 软件仿真获取。考虑到现实中会有分布式电源存在,选择部分节点接入分布式能源。通过仿真生成不同场景下的量测数据构建数据集,其中场景生成主要基于风电、光伏发电及节点注入负荷的场景变化。为模拟真实环境,分别对风机、光伏以及负荷进行不确定性建模。其中

风速采用威布尔(Weibull)分布、光照强度采用贝塔(Beta)分布、负荷默认服从正态分布^[29],通过蒙特卡洛模拟得到的出力概率分布如图 5 所示。

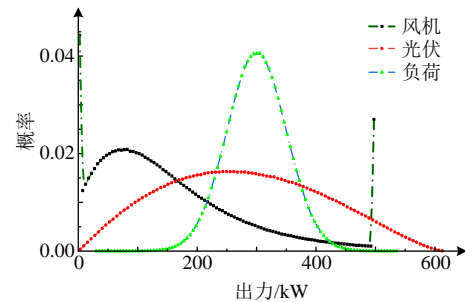


图 5 出力概率分布

Fig. 5 Output probability distribution

3.2 IEEE 33 节点配电网

IEEE 33 节点配电网初始拓扑结构如图 6 所示,主要包括 32 条支路,5 条联络线。风、光分布式电源的接入位置参考文献^[30]。在配电系统正常运行情况下通过改变联络线开关状态生成 28 种拓扑结构,其中 20 种为辐射网,8 种为环网。对每种拓扑结构下通过接入负荷及风光不确定性模型生成仿真场景集合,得到 3000 组样本数据,样本总数为 84 000 组。每组样本数据为 33 个节点的电压幅值和注入有功功率,以及对应拓扑结构标签。随机选取其中 70% 作为训练集,30% 作为测试集。

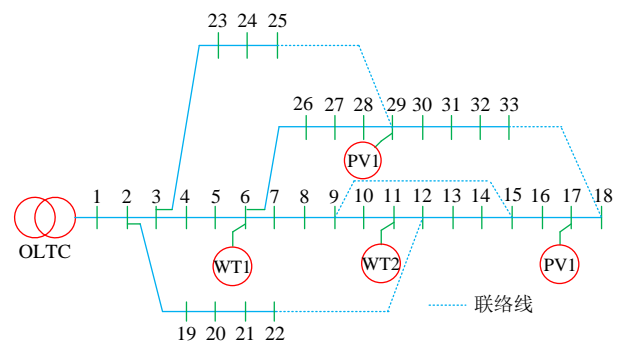


图 6 IEEE33 节点配电网初始拓扑

Fig. 6 IEEE33 node distribution network initial topology

通过随机森林算法计算节点电压幅值、注入有功功率对拓扑辨识的贡献度,并按照贡献度大小进行排序,如图 7 所示。可知,除根节点外,其他所有节点电压幅值对拓扑辨识的贡献度均高于节点注入有功功率,与配网自身结构原理分析结论相同,即配电网节点的注入有功功率只与该节点负荷大小有关,和拓扑结构无关,与本文分析结论相互印证。对此,后续模型训练可舍弃节点注入有功功率,采用节点电压幅值作为特征集,以降低模型训练的计算复杂度和空间复杂度。

模型的输入神经元个数设置为 33,即 33 个节点的电压幅值。由于训练数据有限,为了防止优化

函数陷入局部最优从而偏离真正的全局最优解，本文选用浅层网络。经测试，隐藏层层数选择4层比较合适，层数再减小将会影响分类精度，再增大分类效果增加不明显且会显著增加模型复杂度，每层神经元组合结合经验和少量试错后，设置为16、16、64和64。

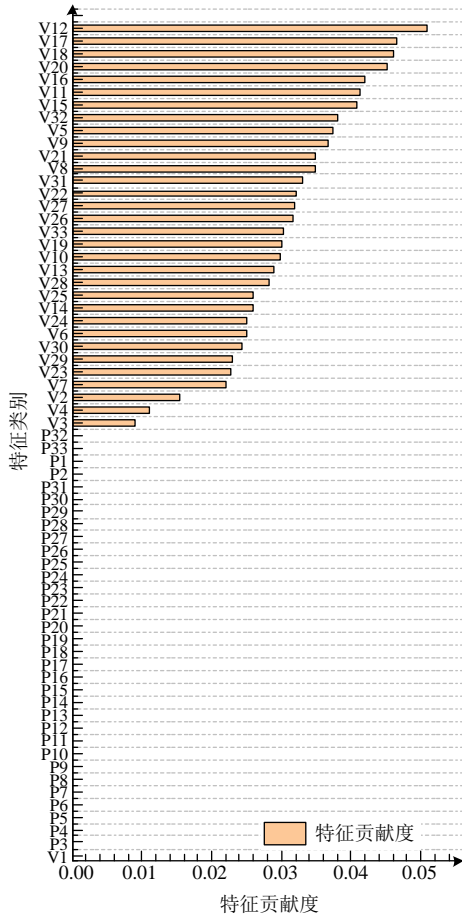


图7 特征贡献度排序

Fig. 7 Feature contribution ranking

将设计完成的模型投入训练，同时为验证本文方法的优越性，还将其与其他常用机器学习算法如CNN、深度神经网络(deep neural networks, DNN)、极限梯度提升(eXtreme gradient boosting, Xgboost)、支持向量机(support vector machines, SVM)进行比较，同理经多次试验后深度神经网络层数设置4层，XGBoost的最大深度设置为5，其训练结果如表1所示。

智能算法	准确率/%	查准率	召回率	测试集时间/s
DNN	94.00	0.6714	0.5858	1.5920
SVM	96.59	0.8925	0.7132	1.1251
CNN	99.31	0.9414	0.9757	0.9328
XGBoost	96.85	0.8632	0.6943	0.9920
ACNN	99.85	0.9875	0.9604	0.8910

由表1结果可知，所有分类器均取得较理想的分类效果，从评价指标和测试集的验证时间分析，CNN、ACNN、SVM和XGBoost相较于DNN分类效果更佳，且CNN和ACNN也优于XGBoost和SVM。从模型结构进行分析，首先由于CNN和ACNN网络层之间连接稀疏及权值共享机制大幅度减少神经网络需要训练参数的个数，相较于DNN和XGBoost，极大地降低了模型计算量。其次，CNN和ACNN含有卷积层和池化层，通过卷积层对数据进行提取特征，模型主要对数据的特征进行处理，而DNN、SVM及XGBoost则对数据本身进行处理，同时CNN和ACNN的池化层会对无关变量进行削减从而降低计算量。综上避免了类似于DNN、SVM及XGBoost因计算量过大造成模型训练过拟合、训练时间过长等缺陷。但由于XGBoost在损失函数处理上使用二阶导数，并加入正则项，可以有效缓解模型过拟合，同时在特征粒度上采用并行优化使得模型拟合更快，精度更高。SVM计算的复杂度取决于其支持向量的数目，而不是样本空间的维数，其次，模型的分类机制是少数支持向量决定最终结果，在此过程中会忽略大量冗余样本，某种意义上可以大幅度降低模型的计算复杂度，从而提升模型拟合速率，提高其分类准确率。

但表1数据中ACNN和CNN的各项性能差异不大，无法体现ACNN的优越性。从2个模型结构角度进行分析，ACNN相较CNN增加注意力机制模块，注意力机制模块对于模型性能的提升主要在于其抗噪能力。由于训练数据集来源于仿真，数据质量较好，无法体现出ACNN的优越性。考虑到性能优良、量测误差较好的PMU设备和μPMU设备的量测总矢量误差分别为0.05%和0.01%，其他普通量测设备总矢量误差更大，因此在数据中分别加入0.01%、0.05%、0.5%和1%的噪声模拟配网真实量测数据，其模型训练结果如表2所示。

智能算法	噪声水平/%			
	TVE=0.01%	TVE=0.05%	TVE=0.5%	TVE=1%
DNN	92.93	89.95	85.90	78.68
SVM	95.36	90.21	86.72	79.34
CNN	99.23	98.05	89.52	85.47
XGBoost	96.42	94.83	90.07	83.39
ACNN	99.77	99.26	93.34	89.69

由表2结果可以看出，在一定范围内，随着训练数据集中噪声水平的不断增加，ACNN模型分类

效果的优越性逐步得以体现，这是因为注意力机制能够将模型的算力集中于重要的特征类别，因此避免了特征重要度较低的特征类别因噪声带来的大量算力消耗，从而提升模型的抗噪能力。

为进一步验证模型有效性，对模型计算归一化混淆矩阵，其热力图结果如图8所示，其中第*i*行第*j*列元素含义为：真实拓扑标签是*i*，预测拓扑标签是*j*的概率。由图示结果可以看出：对角线上的元素几乎全为1，非对角线上的元素基本为0，也印证了本文方法在分类方面的有效性。

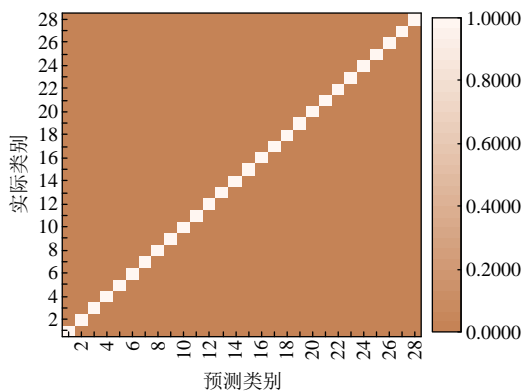


图8 IEEE 33 节点算例混淆矩阵热力图
Fig. 8 IEEE 33 node example confusion matrix thermodynamic diagram

为探究在量测数据无误差情况下拓扑分类无法完全准确的原因，将拓扑错误数据样本进行采集，发现其对应拓扑结构较相似，取出典型如图9所示，若邻近节点7上负荷为0或很小，而节点间距离又较短，线路损耗忽略不计时，则以下2个拓扑结构的电压特征极其相似，易出现拓扑辨识错误。在实际配网中，通常拓扑结构变化间隔时间较长，可在间隔时段内多次采集量测数据进行综合分析，当7号节点负荷增加后，2个拓扑结构量测数据间差异性将得以体现，即可实现辨识。

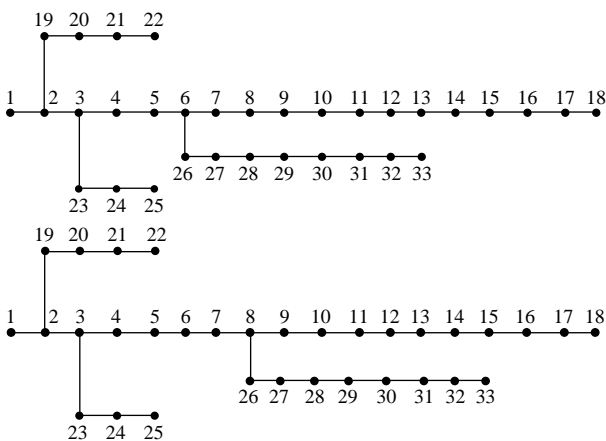


图9 IEEE 33 节点配电网拓扑改变图
Fig. 9 IEEE 33 node distribution network topology change diagram

理论上，用更少的特征来实现拓扑辨识将更能满足实际需求，也与在实际量测中仅在关键节点处布有量测装置的工程实际相符合。本文按照特征重要度大小依次减少特征类别，以用更少的特征类别对模型进行训练和测试，得到的测试集准确率与特征类别数关系如图10所示。

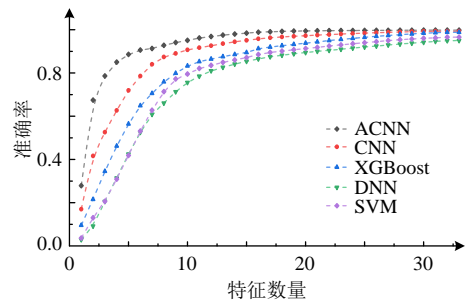


图10 特征数量与验证集准确率关系
Fig. 10 Relationship between feature number and accuracy of verification set

由图10可以看出，在最开始减少特征类别数量时，随着特征数量的减少，验证集的准确率下降缓慢甚至不下降，因为计算特征贡献度时，某些节点电压数据之间具有强相关性，随机森林算法将自动从高相关性的多个特征类别中选取一个作为具有实际贡献度的类别，其他特征类别的特征贡献度自动减小并排至序列尾部，这就造成当最开始减少特征重要度小的的特征类别时，前期正确率下降缓慢甚至保持不变，而随着特征重要度大的特征类比开始减少时，验证集准确率快速降低。

此外，在一定范围内，当特征数量减少时，ACNN的验证集准确率与其他4种模型的差距逐渐增大，因为注意力机制的加入导致模型的注意力集中于重要的特征类别上，因此特征重要度小的特征类别减少对模型的影响极小；而重要特征类别减少时，其准确率下降速度则陡然加快。当特征类别数量减至15时，验证集准确率依旧能达到90%以上，表明在实际应用中可以仅在关键节点处采集量测数据进行拓扑辨识，这也与电网实际情况相符，因此本文模型具有极强的实际应用价值。

3.2 PG&E69 节点配电网

标准 PG&E69 节点配电网的原始拓扑结构如图11所示，选择5个节点接入分布式电源，其接入位置参考文献[31]。

与IEEE33节点配电网算例类似，在保证配电系统正常运行前提下，通过改变联络线的开关状态获得28种拓扑结构，辐射状网络有20种，含环网络有8种。每种拓扑结构下，改变“源-荷”场景

生成 3000 组样本数据，共有 84 000 组样本数据。鉴于在上节已开展特征贡献度测试，此处仅考虑节点电压作为特征集，该特征集共有 69 维分别为每

个节点电压幅值。将特征集与对应的拓扑标签组合构成训练数据集。将训练数据集随机取 70% 作为训练集，30% 作为测试集。

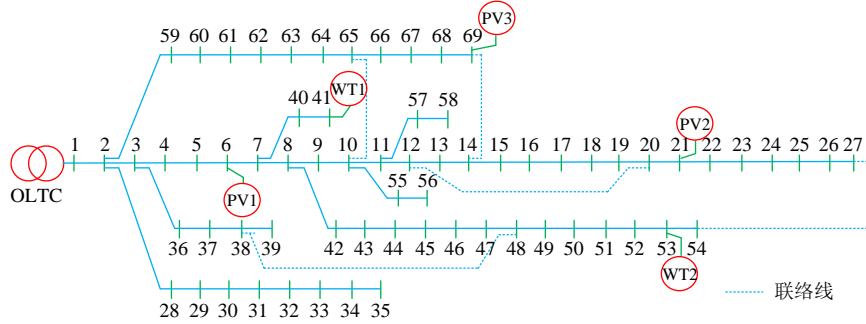


图 11 PG & E69 节点配电网初始拓扑
Fig. 11 Initial topology of PG & E69 node distribution network

为验证本文所提模型的泛化能力，仅更改模型输入层，将输入层由 33 维改成 69 维，中间隐藏层和输出均沿用上节模型，将训练集输入该模型进行训练，并用测试集对训练好的模型进行测试，同时与其他 4 种机器学习算法结果进行对比，其结果如表 3 所示。

表 3 PG&E69 节点配电网算法性能比较
Table 3 Comparison of algorithm performance on the PG&E69 node distribution network

智能算法	准确率/%	查准率	召回率	测试集时间/s
DNN	87.00	0.5726	0.4276	2.6914
SVM	92.76	0.7815	0.7862	2.0532
CNN	98.31	0.8635	0.9017	1.3784
XGBoost	93.15	0.7831	0.7880	1.8920
ACNN	99.84	0.9799	0.8982	1.0261

由表 3 数据可以看出，本文模型在 PG&E69 节点配电网中仍有较理想的拓扑识别效果，并未因节点数量的增加而如 DNN、XGBoost 等模型，辨识效果骤降，因此模型具备较好泛化能力。这同样由于卷积神经网络和注意力机制的特殊结构，上文中提到卷积神经网络具有卷积层、池化层，并且层与层之间的连接稀疏，导致了卷积神经网络能够节约大量算力从而集中于重要任务上，本文算法凭借注意力机制和卷积神经网络的特殊结构，使得其分类效果明显高于 DNN、SVM 和 XGBoost。但是由于 ACNN 和 CNN 的差距并不大，同理为了体现出注意力机制的抗噪能力，在量测数据中分别加入 0.01%、0.05%、0.5% 和 1% 的噪声，从而验证模型的优越性、鲁棒性，其测试结果如表 4 所示。

由表 4 结果可知，在一定范围内，随着数据集中噪声水平越高，本文模型与其他 4 种模型分类效果相比优越性越得到体现，证明了注意力机制的结合对卷积神经网络模型的鲁棒性提升具有巨大意义。同理，由于该模型为多分类模型，进一步计算

该系统的归一化混淆矩阵热力图如图 12 所示。

表 4 考虑量测噪声的 PG&E69 节点配电网测试集准确率
Table 4 Accuracy of PG&E69 node test set considering measurement noise

智能算法	噪声水平/%			
	TVE=0.01%	TVE=0.05%	TVE=0.5%	TVE=1%
DNN	86.93	84.95	78.90	71.68
SVM	92.13	88.61	83.30	78.50
CNN	98.18	96.65	89.48	83.21
XGBoost	92.67	90.98	86.31	81.99
ACNN	99.78	99.40	94.35	88.49

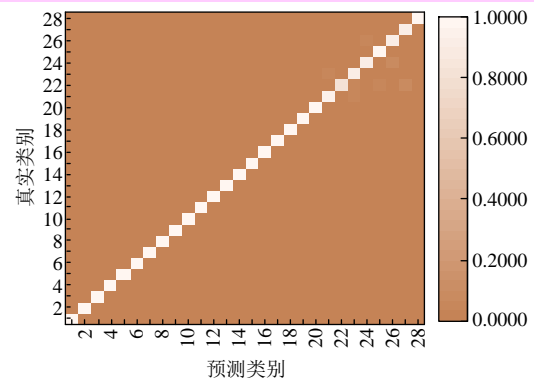


图 12 PG&E69 节点算例混淆矩阵热力图
Fig. 12 PG&E69 node example confusion matrix thermodynamic diagram

由图可以看出，对角线上元素基本为 1，非对角线上元素基本为 0，由此可得在 PG&E69 节点配电网中本文模型仍具有良好分类效果，证明了本文模型的良好泛化能力。与 IEEE 33 节点配电网算例类似，测试特征数量对模型分类准确率的影响，其结果如图 13 所示。5 种模型分类准确率变化情况与 IEEE 33 节点算例结果类似，随着特征数量减少，测试集的分类准确率下降速率先缓慢，后随着特征数量降低到一定程度，准确率下降速度突然加快，其原因于上节已经论述。对于本文模型，当特征数量为 13 时，验证集准确率为 90.12%，这使得模型在实际中具有十分广阔的应用前景。

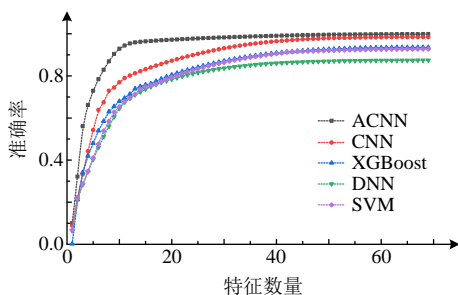


图13 特征数量与验证集准确率关系

Fig. 13 Relationship between feature number and accuracy of verification set

4 结论

本文基于卷积神经网络和注意力机制建立配电网拓扑辨识框架,采用随机森林智能算法计算节点电压幅值和节点注入功率对拓扑辨识的特征贡献度,筛选出构建训练数据集的特征类别,并基于数据集训练模型挖掘其与拓扑结构间的关系,实现映射规则学习,最后通过 IEEE 33 节点配电网和 PG&E69 节点配电网算例验证本文构建模型的有效性和优越性。基于本文研究成果得出如下结论:

1) 通过实际验证,节点电压量测数据对拓扑辨识的特征贡献度要远大于节点注入功率;

2) 和传统辨识方法相比,本文构建的拓扑辨识模型数据质量要求低,辨识速度快可实现在线辨识,且不仅适用于辐射网,同样适用于弱环网;

3) 引入注意力机制与卷积神经网络结合,通过算例验证,模型的性能、鲁棒性均得到有效提升,并具备更好的泛化能力;

4) 通过实际算例验证可进一步删减非关键特征,仅用关键节点量测数据即可完成配电网拓扑辨识,使得本文所提模型具备更广阔的应用前景。

未来在实际工程应用中,可将本文所建模型与 SCADA 系统互联,将采集的历史数据处理后与对应拓扑结构标签一并送入模型中进行训练,待模型训练好后,将实时量测数据送入模型以实现实时拓扑辨识,以支撑系统状态估计、安全稳定分析、实时调度控制等高级应用。

参考文献

[1] 何维国,王赛一,许唐云,等.城市韧性配电网建设与发展路径[J].电网技术,2022,46(2):680-690.
HE Weiguo, WANG Saiyi, XU Tangyun, et al. Construction and development path of the urban resilient distribution network[J]. Power System Technology, 2022, 46(2): 680-690(in Chinese).

[2] 祁琪,姜齐荣,许彦平.智能配电网柔性互联研究现状及发展趋势[J].电网技术,2020,44(12):4664-4676.
QI Qi, JIANG Qirong, XU Yanping. Research status and development prospect of flexible interconnection for smart distribution

networks[J]. Power System Technology, 2020, 44(12): 4664-4676(in Chinese).

[3] 张旭,王洪涛.高比例可再生能源电力系统的输配协同优化调度方法[J].电力系统自动化,2019,43(3):67-75.
ZHANG Xu, WANG Hongtao. Optimal dispatch method of transmission and distribution coordination for power systems with high proportion of renewable energy[J]. Automation of Electric Power Systems, 2019, 43(3): 67-75(in Chinese).

[4] 余贻鑫.面向21世纪的智能电网[J].天津大学学报(自然科学与工程技术版),2020,53(6):551-556.
YU Yixin. A brief description of the basics of the smart grid[J]. Journal of Tianjin University (Science and Technology), 2020, 53(6): 551-556(in Chinese).

[5] 王金丽,盛万兴,宋祺鹏,等.配电网电能质量智能监控与治理仿真[J].电网技术,2014,38(2):515-519.
WANG Jinli, SHENG Wanxing, SONG Qipeng, et al. Research on intelligent power quality monitoring and on-line simulation for distribution network[J]. Power System Technology, 2014, 38(2): 515-519(in Chinese).

[6] 郭帅文,燕跃豪,蒋建东,等.基于邻接矩阵的网络拓扑辨识算法[J].电力系统保护与控制,2018,46(12):50-56.
GUO Shuaiwen, YAN Yuehao, JIANG Jiandong, et al. Network topology identification algorithm based on adjacency matrix[J]. Power System Protection and Control, 2018, 46(12): 50-56(in Chinese).

[7] 刘超,杨扬,梁栋,等.基于AMI潮流匹配的中压配电网两阶段拓扑辨识[J].电力系统及其自动化学报,2020,32(3):123-128.
LIU Chao, YANG Yang, LIANG Dong, et al. Two-stage topology identification of medium-voltage distribution network based on power flow matching of AMI measurements[J]. Proceedings of the CSU-EPSA, 2020, 32(3): 123-128(in Chinese).

[8] DENG Wei, ZHANG Zhenyu, DUAN Jing, et al. Improved topology identification algorithm of distribution network mutual information[C]/Proceedings of 2019 IEEE Sustainable Power and Energy Conference. Beijing: IEEE, 2019.

[9] WANG Xiaozhe, BIALEK J W, TURITSYN K. PMU-based estimation of dynamic state Jacobian matrix and dynamic system state matrix in ambient conditions[J]. IEEE Transactions on Power Systems, 2018, 33(1): 681-690.

[10] ARGHANDEH R, GAHR M, VON MEIER A, et al. Topology detection in microgrids with micro-synchrophasors[C]/Proceedings of 2015 IEEE Power & Energy Society General Meeting. Denver: IEEE, 2015.

[11] ZHANG Jiawei, WANG Yi, WENG Yang, et al. Topology identification and line parameter estimation for non-PMU distribution network: a numerical method[J]. IEEE Transactions on Smart Grid, 2020, 11(5): 4440-4453.

[12] 宁家鑫,刘云霄,章家维,等.数据驱动的三相配电网拓扑与线路参数辨识[J].中国电机工程学报,2021,41(8):2615-2627.
NING Jiaxin, LIU Yuxiao, ZHANG Jiawei, et al. Data-driven topology and line parameter identification of three-phase distribution grid[J]. Proceedings of the CSEE, 2021, 41(8): 2615-2627(in Chinese).

[13] LUAN Wenpeng, PENG J, MARAS M, et al. Smart meter data analytics for distribution network connectivity verification[J]. IEEE Transactions on Smart Grid, 2015, 6(4): 1964-1971.

[14] 闫卫国,王高猛,林济铿,等.基于AMI量测信息的低压配电网拓扑校验方法[J].中国电力,2019,52(2):125-133.
YAN Weiguo, WANG Gaomeng, LIN Jikeng, et al. Method for LV distribution network topology verification based on AMI metering data[J]. Electric Power, 2019, 52(2): 125-133(in Chinese).

- [15] 肖勇, 赵云, 涂治东, 等. 基于改进的皮尔逊相关系数的低压配电网拓扑结构校验方法[J]. 电力系统保护与控制, 2019, 47(11): 37-43.
XIAO Yong, ZHAO Yun, TU Zhidong, et al. Topology checking method for low voltage distribution network based on improved Pearson correlation coefficient[J]. Power System Protection and Control, 2019, 47(11): 37-43(in Chinese).
- [16] 陈招安, 黄纯, 张志丹, 等. 基于 T 型灰色关联度和 KNN 算法的低压配电网台区拓扑识别方法[J]. 电力系统保护与控制, 2021, 49(1): 163-169.
CHEN Zhaoan, HUANG Chun, ZHANG Zhidan, et al. Topology identification method of a low voltage distribution network based on T-type grey correlation degree and KNN algorithm[J]. Power System Protection and Control, 2021, 49(1): 163-169(in Chinese).
- [17] 杨挺, 赵黎媛, 王成山. 人工智能在电力系统及综合能源系统中的应用综述[J]. 电力系统自动化, 2019, 43(1): 2-14.
YANG Ting, ZHAO Liyuan, WANG Chengshan. Review on application of artificial intelligence in power system and integrated energy system[J]. Automation of Electric Power Systems, 2019, 43(1): 2-14(in Chinese).
- [18] 高泽璞, 赵云, 余伊兰, 等. 基于知识图谱的低压配电网拓扑结构辨识方法[J]. 电力系统保护与控制, 2020, 48(2): 34-43.
GAO Zepu, ZHAO Yun, YU Yilan, et al. Low-voltage distribution network topology identification method based on knowledge graph[J]. Power System Protection and Control, 2020, 48(2): 34-43(in Chinese).
- [19] 裴宇婷, 秦超, 余贻鑫. 基于 LightGBM 和 DNN 的智能配电网在线拓扑辨识[J]. 天津大学学报(自然科学与工程技术版), 2020, 53(9): 939-950.
PEI Yuting, QIN Chao, YU Yixin. Online topology identification for smart distribution grids based on LightGBM and deep neural networks[J]. Journal of Tianjin University (Science and Technology), 2020, 53(9): 939-950(in Chinese).
- [20] 孙伟, 朱世睿, 杨建平, 等. 基于图卷积网络的微电网拓扑辨识[J]. 电力系统自动化, 2022, 46(5): 71-81.
SUN Wei, ZHU Shirui, YANG Jianping, et al. Topology identification of microgrid based on graph convolutional network[J]. Automation of Electric Power Systems, 2022, 46(5): 71-81(in Chinese).
- [21] 胡丽娟, 刁赢龙, 刘科研, 等. 基于大数据技术的配电网运行可靠性分析[J]. 电网技术, 2017, 41(1): 265-271.
HU Lijuan, DIAO Yinglong, LIU Keyan, et al. Operational reliability analysis of distribution network based on big data technology[J]. Power System Technology, 2017, 41(1): 265-271(in Chinese).
- [22] 杨晓梅, 郭林明, 肖先勇, 等. 基于可调品质因子小波变换和随机森林特征选择算法的电能质量复合扰动分类[J]. 电网技术, 2020, 44(8): 3014-3020.
YANG Xiaomei, GUO Linming, XIAO Xianyong, et al. Classification of multiple power quality disturbances based on TQWT and random forest feature selection algorithm[J]. Power System Technology, 2020, 44(8): 3014-3020(in Chinese).
- [23] XIAO Mengmeng, WANG Shaorong, ULLAH Z, et al. Topology detection in power distribution system using kernel-node-map deep networks[J]. IET Generation, Transmission & Distribution, 2020, 14(19): 4033-4041.
- [24] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [25] 万晓琪, 宋辉, 罗林根, 等. 卷积神经网络在局部放电图像模式识别中的应用[J]. 电网技术, 2019, 43(6): 2219-2226.
WAN Xiaoqi, SONG Hui, LUO Lin'gen, et al. Application of convolutional neural networks in pattern recognition of partial discharge image[J]. Power System Technology, 2019, 43(6): 2219-2226(in Chinese).
- [26] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72.
ZHANG Chenjia, ZHU Lei, YU Lu. Review of attention mechanism in convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(20): 64-72(in Chinese).
- [27] 谢可, 郭文静, 祝文军, 等. 面向电力物联网海量终端接入技术研究综述[J]. 电力信息与通信技术, 2021, 19(9): 57-69.
XIE Ke, GUO Wenjing, ZHU Wenjun, et al. Review on access technology of massive terminals of power internet of things[J]. Electric Power Information and Communication Technology, 2021, 19(9): 57-69(in Chinese).
- [28] 杨冬锋, 付强, 刘晓军, 等. 配电网动态拓扑与线路参数联合在线辨识方法[J]. 电力系统自动化, 2022, 46(2): 101-108.
YANG Dongfeng, FU Qiang, LIU Xiaojun, et al. Joint online identification method for dynamic topology and line parameters of distribution network[J]. Automation of Electric Power Systems, 2022, 46(2): 101-108(in Chinese).
- [29] 杨晓萍, 王李瑾. 基于概率潮流的含分布式电源配电网优化[J]. 太阳能学报, 2021, 42(8): 71-76.
YANG Xiaoping, WANG Lijin. Optimization of distributed power distribution network based on probabilistic load flow[J]. Acta Energetica Solaris Sinica, 2021, 42(8): 71-76(in Chinese).
- [30] DOROSTKAR-GHAMSARI M R, FOTUHI-FIRUZABAD M, LEHTONEN M, et al. Value of distribution network reconfiguration in presence of renewable energy resources[J]. IEEE Transactions on Power Systems, 2016, 31(3): 1879-1888.
- [31] 徐俊俊, 吴在军, 周力, 等. 考虑分布式电源不确定性的配电网鲁棒动态重构[J]. 中国电机工程学报, 2018, 38(16): 4715-4725.
XU Junjun, WU Zaijun, ZHOU Li, et al. Robust dynamic reconfiguration for distribution networks considering uncertainty of distributed generations[J]. Proceedings of the CSEE, 2018, 38(16): 4715-4725(in Chinese).



杨秀

在线出版日期: 2022-03-31。

收稿日期: 2021-12-31。

作者简介:

杨秀(1972), 男, 博士, 教授, 博士生导师, 研究方向为分布式发电与微网技术研究、HVDC 与 FACTS 的运行与控制等, E-mail: yangxiu721102@126.com;

蒋家富(1999), 男, 硕士研究生, 研究方向为大数据及智能配电网, E-mail: jeff6797@126.com;

刘方(1989), 男, 博士, 讲师, 通信作者, 研究方向为电力系统经济调度、人工智能技术在电力系统运行分析中的应用等, E-mail: liufang_SUEP@163.com;

田英杰(1969), 男, 高级工程师, 研究方向为电力大数据和电力人工智能等, E-mail: 13901712348@163.com;

李凡(1989), 男, 工程师, 研究方向为电力系统自动化, E-mail: tianda907@163.com;

吴裔(1987), 男, 工程师, 研究方向为电力系统自动化, E-mail: wuy@sh.sgcc.com.cn。

(责任编辑 温杰)