

# 考虑数据失衡的新型配电网两阶段拓扑辨识

陈碧云<sup>1</sup>, 吕怡博<sup>1</sup>, 梁志坚<sup>2</sup>, 张勇军<sup>3</sup>, 徐旗<sup>1</sup>, 付天旺<sup>1</sup>

(1. 广西电力系统最优化与节能技术重点实验室(广西大学), 广西 南宁 530004; 2. 广西大学电气工程学院, 广西 南宁 530004; 3. 华南理工大学电力学院, 广东 广州 510640)

**摘要:** 在新型电力系统快速发展的背景下, 配电网中各类分布式电源、储能、电动汽车、柔性负荷的接入占比不断增加, 运行方式日趋复杂多变, 其拓扑的精确辨识也更有难度。针对现有配电网量测数据采集周期较长、辨识方法对数据不平衡敏感而导致辨识精度不高的问题, 提出了一种两阶段的新型配电网拓扑辨识方法。首先, 采用两层堆叠的图卷积网络生成系列标签分类器, 再用卷积神经网络提取量测时间序列的特征, 并结合多标签分类学习实现第一阶段的初步辨识。其次, 对初步辨识获得的初始拓扑中状态为“阴性”(开断)的支路进行全状态空间搜索, 并通过潮流匹配模型, 筛选出耗散值最小的状态样本, 实现“假阴性”二次辨识。最后, 在改进的IEEE33节点含新能源配电网上进行仿真验证。结果表明, 所提模型和方法能有效解决数据失衡的问题, 并具有更高的辨识精度。

**关键词:** 配电网; 拓扑辨识; 多标签分类; 数据失衡

## Two-stage topology identification of a new-type distribution network considering data imbalance

CHEN Biyun<sup>1</sup>, LÜ Yibo<sup>1</sup>, LIANG Zhijian<sup>2</sup>, ZHANG Yongjun<sup>3</sup>, XU Qi<sup>1</sup>, FU Tianwang<sup>1</sup>

(1. Guangxi Key Laboratory of Power System Optimization and Energy Saving Technology (Guangxi University), Nanning 530004, China; 2. School of Electrical Engineering, Guangxi University, Nanning 530004, China; 3. School of Electric Power, South China University of Technology, Guangzhou 510640, China)

**Abstract:** Along with the rapid development of new-type power systems, the proportion of access to various types of distributed power sources, energy storage, electric vehicles, and flexible loads in the distribution network is increasing. The operation mode is becoming more complex and changeable, and the accurate identification of their topology is more difficult. There is a problem in that the data collection cycle of distribution network measurement is long and the identification method is sensitive to data imbalance, resulting in low identification accuracy. Thus this paper proposes a two-stage topology identification method of a new-type distribution network. First, a two-layer stacked graph convolutional network is used to generate a series of label classifiers, and then the convolutional neural network is used to extract the features of the measured time series, and the preliminary identification of the first stage is realized by combining multi-label classification learning. Second, the branches with a state of "negative" (breaking) in the initial topology obtained by preliminary identification are searched for the whole state space. The state samples with the smallest dissipation value are screened out through a power flow matching model to achieve a secondary identification of "false negative". Finally, simulation verification is carried out on the modified IEEE33 node power distribution network with new energy. The results show that the proposed model and method can effectively solve the data imbalance problem and have higher identification accuracy.

This work is supported by the National Natural Science Foundation of China (No. 52177085).

**Key words:** distribution grid; topology identification; multi-label classification; data imbalance

## 0 引言

随着新型电力系统建设步伐的加快, 配电网中

**基金项目:** 国家自然科学基金项目资助(52177085); 广东省基础与应用基础研究基金项目资助(粤桂联合基金-面上项目)(2021A1515410009)

各类可再生分布式电源、电动汽车、柔性负荷、储能等接入占比日益增高, 配电网的运行方式愈加复杂多变<sup>[1-3]</sup>, 如何利用数量有限、样本不均衡、噪声敏感的数据, 精确辨识配电网的真实拓扑是实现透明新型配电网、开展配电网自愈控制、负荷管理等智能化应用的重要基础。

现有的辨识方法主要有基于机理模型和基于数

据驱动两大类。其中,前者对量测数据的数量、质量和时间同步性的要求较高<sup>[4-5]</sup>,需要配置足够的高级量测设备做支撑,但目前大部分配电网区很难完全达到所要求的硬件配置水平<sup>[6]</sup>。而后者主要是针对节点或支路信息进行数据挖掘,应用深度神经网络、支持向量机等方法<sup>[7-10]</sup>,通过学习拓扑类别和输入特征的先验关系,获得拓扑辨识的模型并用于动态拓扑辨识,这一类方法在样本数量充足且分布均匀的条件下可以获得较好的辨识效果,而当存在不同类别量测数据分布不均时,噪声会对辨识精度产生较大的影响。

为提高两类方法的适用性和精度,研究学者提出将机理模型与数据驱动模型混合的方法。文献[11]提出了一种数据驱动的拓扑结构估计方法,基于直接连接的节点间相关性构造“信息-学习”的方法来重建电网拓扑结构。文献[12]对历史电压幅值数据作偏相关分析来消除节点数据间的多重共线性,再通过树生成的方式重现网络拓扑,但该方法应用于大规模配电网复杂度较高。文献[13]将数据驱动作为对配电网状态估计方法的补充,通过对实时数据的聚类 and 辨识,估算潜在拓扑及其概率,过滤部分低概率拓扑样本,以弥补了传统状态估计方法在不收敛或无观测状态下无法计算的缺点。文献[14]提出一种联合数据和模型驱动的方法,将第一步基于数据驱动的回归方法得到的初步拓扑估计通过数模驱动的牛顿拉夫逊法迭代和潮流计算进一步校正,但这种方法较依赖迭代初始状态的选取。由于配电网的节点电压往往与多个支路的状态有关,其量测数据具有典型的非欧式特征<sup>[15]</sup>,文献[16]将基于节点的拓扑图转化为描述边信息的线图,通过神经网络实现微电网中各支路状态辨识,以便更好地解决非欧数据问题<sup>[17-18]</sup>,但该方法需要获知已有的全拓扑结构。

此外,由于设备质量、环境和运维技术的差异性,配电网各类故障出现的概率是不均匀的,相应地,某些故障类别导致的拓扑样本数量较少,尤其是高阶故障对应的拓扑样本数十分有限,与正常运行状态的拓扑样本数量相比,出现数据失衡问题<sup>[19]</sup>,随着新型配电网中源荷随机特性的增强,失衡趋于严重。数据失衡会使辨识结果偏向体量较大的类别,而忽略小样本类别<sup>[20]</sup>,现有数据驱动的拓扑辨识方法较少考虑数据失衡对辨识结果的影响。因此,本文提出了一种基于初步拓扑辨识和潮流匹配的双层配电网拓扑辨识模型。首先,针对现有辨识模型在未知拓扑结构以及不同拓扑结构间存在数据失衡的情况下辨识误差较大的问题,构建了基于多标签学

习的初步拓扑辨识模型,将单标签中大类别细分为多标签的学习,并选择合适的损失函数来解决非欧数据问题。其次,对于多标签分类层出现的“假阴”结果,增加基于潮流匹配模型的二次辨识,进一步提高辨识模型的整体泛化能力。

## 1 两阶段辨识模型设计

### 1.1 数据失衡问题

数据失衡又称样本比例失衡,是指不同类别的样本量差异非常大,或者少数样本代表了业务的关键数据,在有监督的机器学习任务中,表现为样本标签的分布不均匀。各类别数据失衡的程度可用失衡比  $p_{imb}$  来表示,如式(1)所示。

$$p_{imb} = \frac{N_j}{N_{\max}} \quad (1)$$

式中:  $N_{\max}$  为样本集中数量最多类别的样本数量;  $N_j$  为第  $j$  类样本的数量。

配网中各支路的阳性(闭合)样本分布占比  $c$  可以定义为

$$c = \frac{N_{\text{pos}}^i}{N} \times 100\% \quad (2)$$

式中:  $N$  为拓扑样本总数;  $N_{\text{pos}}^i$  表示支路  $i$  为阳性的样本数量。

### 1.2 两阶段辨识模型框架

针对以上问题,本文采用图1所示的两阶段模型进行拓扑辨识。

第一阶段:基于 MLC-GCN-CNN 的初步辨识。首先,采用两层堆叠的图卷积网络(graph convolutional network, GCN),由标签的相关系数矩阵和词嵌入向量矩阵生成系列标签分类器,再用卷积神经网络(convolutional neural networks, CNN)提取量测时间序列的特征,并结合多标签分类学习(multi-label classification, MLC)辨识出初始拓扑结构。

第二阶段:基于潮流匹配的假阴性二次辨识。对初始拓扑结构中辨识为“阴性”(开断)的支路进行全状态空间的搜索,应用潮流匹配模型找出耗散值最小的状态样本,得到二次辨识的拓扑结果。

其中,第一阶段以数据驱动模型代替经典拓扑辨识的混合整数线性规划模型,解决了物理模型的时效性问题,且通过 MLC 将数据样本的类别失衡转化为支路阴性状态频率的失衡,可以提升监督学习模型的泛化能力<sup>[15]</sup>。此外,针对非欧式数据特征,采用 GCN 来生成系列标签分类器,可以进一步弱化数据失衡对辨识精度的影响。

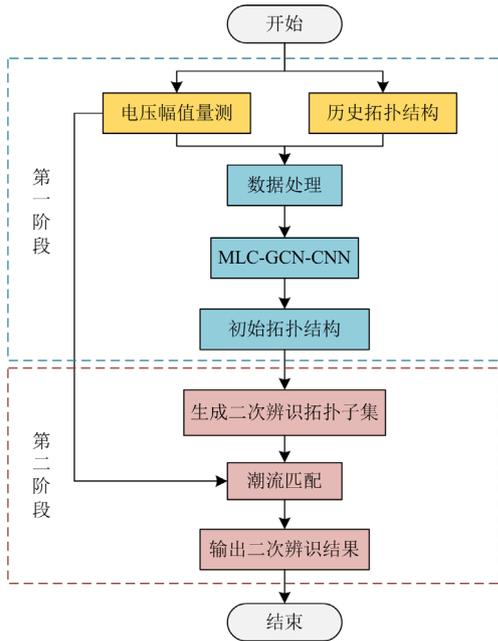


图1 两阶段辨识总框架

Fig. 1 Two-stage identification framework

## 2 两阶段拓扑辨识算法

### 2.1 MLC-GCN-CNN 初步辨识方法

对于有  $n$  条支路的配电网, 其多标签状态可表示为  $L = [l_1, l_2, \dots, l_j, \dots, l_n]$ , 其中

$$l_j = \begin{cases} 0 & \text{第 } j \text{ 条支路断开} \\ 1 & \text{第 } j \text{ 条支路闭合} \end{cases} \quad (3)$$

如图2所示, 为清晰地表征具有非欧量测数据特征的配电网的点与点、点与支路、支路与支路之间的关系, 选择配网中有直接连接关系的量测点之间的连接状态为标签, 首先依据历史数据, 生成表征标签之间信息传递的无向关联矩阵(邻接矩阵)  $\mathbf{A}$ , 再生成描述网络节点与支路关系的词嵌入向量矩阵(对角矩阵)  $\mathbf{A}_{\text{imp}}$ , 然后将  $\mathbf{A}$  和  $\mathbf{A}_{\text{imp}}$  作为输入, 经过两层堆叠的 GCN, 由非线性激活函数  $\delta(\cdot)$  在节点间进行信息传递, 得到  $D \times C$  维的分类特征集<sup>[21]</sup>。

$$H^{l+1} = \delta(H^l, \mathbf{A}) \quad (4)$$

式中:  $D$  为特征的维数;  $C$  为类别数;  $H^l$  为第  $l$  层节点特征;  $\delta(\cdot)$  为激活函数<sup>[22]</sup>。

同时, 用 CNN 网络提取电压幅值时间序列的  $D$  维特征集, 再与 GCN 模块输出的分类特征集进行加权点积运算得到  $x_{\text{GC}}$ , 运算结果与经全连接得到的  $x_{\text{CNN}}$  输入特征差异模块进行式(5)运算。

$$O_{\text{output}} = x_{\text{GC}} + x_{\text{CNN}} \quad (5)$$

经特征差异化模块计算得到各个标签的得分,

通过分别对比每个标签得分和阈值, 得到初步辨识结果。阈值在训练过程中通过对训练样本的标签得分与其真实标签对比筛选得到, 将模型计算输出标签得分大于阈值的置 1, 否则置 0, 得到初始拓扑状态。

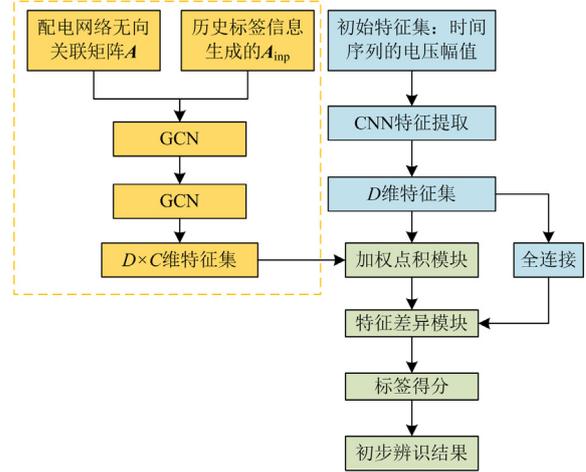


图2 基于 GCN 的多标签辨识算法流程图

Fig. 2 Flowchart of MLC-GCN-CNN

在单标签细分为多标签后, 类别间的不平衡将转化为标签状态次数的不平衡, 为解决这一失衡问题, 本文使用 ZLPR 损失函数, 通过对比各标签的得分平衡其权重<sup>[23]</sup>。

$$l_{\text{oss}} = \log \left( 1 + \sum_{i \in \Omega_{\text{neg}}} e^{s_i} \right) + \log \left( 1 + \sum_{j \in \Omega_{\text{pos}}} e^{-s_j} \right) \quad (6)$$

式中,  $\Omega_{\text{pos}}$ 、 $\Omega_{\text{neg}}$  分别为阳性和阴性的标签集合;  $s_i$  为第  $i$  类标签的得分。

### 2.2 潮流匹配

为了提升模型的泛化能力, 改善初步辨识结果可能存在“假阴”标签的问题, 采用潮流匹配模型进行假阴性二次辨识。配电网中节点的量测数据与拓扑结构密切相关, 一个节点量测的变化可以表征多回支路的状态变化, 在节点注入功率相同的情况下, 相似拓扑结构的量测值相差较小<sup>[24]</sup>。因此, 基于潮流匹配的二次辨识模型可以表示为

$$\begin{cases} \min & J(s) = [V^{\text{AMI}} - V^{\text{Est}}(s)]^T R^{-1} [V^{\text{AMI}} - V^{\text{Est}}(s)] \\ \text{s.t.} & V^{\text{Est}}(s) = f(L_{\text{AS}}, \mathbf{P}) \\ & \mathbf{P} = P^{\text{Est}} \\ & L_{\text{AS}} = g(L_i) \\ & L_i = N(V^{\text{AMI}}, La) \\ & L_i = \{L_1, L_2, \dots, L_b\} \end{cases} \quad (7)$$

式中：目标函数  $J(s)$  为耗散值，即电压幅值量测值与潮流估计值之差的加权平方和<sup>[24]</sup>； $\mathbf{P}$  为的节点注入功率的量测时间序列； $V^{\text{AMI}}$  为节点电压幅值的量测时间序列； $V^{\text{Est}}$  为潮流计算得到的电压幅值； $f(\cdot)$  为潮流方程； $L_i$  为第  $i$  个样本初步辨识得到的初始拓扑； $L_{\text{AS}}$  是由  $L_i$  拓展得到的一系列二次辨识拓扑子集； $La$  为历史标签信息； $N(\cdot)$  为 MLC-GCN-CNN 多标签学习网络。

1) 以初始辨识结果作为二次辨识的初始拓扑状态，提取其中状态辨识为 0 的标签，放入校核状态集  $\mathbf{S}_{\text{eq}}$ ，与各节点注入功率一起作为第二层输入，创建空集  $\mathbf{S}_{\text{ta}}$ 、 $\mathbf{S}_{\text{n}}$ ，分别用于存放各阶状态和各阶筛选保留的状态；

2) 创建空集  $\mathbf{S}_{\text{ta}}$ 、 $\mathbf{S}_{\text{n}}$ ，分别用于存放各阶状态和各阶筛选保留的状态，令  $k=0$ ，将初始状态分别存入  $\mathbf{S}_{\text{ta}(0)}$ 、 $\mathbf{S}_{\text{n}}$ ，并计算初始拓扑状态的耗散值；

2) 计算初始拓扑状态的耗散值，将初始状态放入  $\mathbf{S}_{\text{ta}(0)}$ 、 $\mathbf{S}_{\text{n}}$ ；

3)  $\mathbf{S}_{\text{ta}(0)}$  分别与  $\mathbf{S}_{\text{eq}}$  的不同拓展状态组合得到新的拓扑状态集  $\mathbf{S}_{\text{ta}(1)}$ ，即一阶子状态集， $k=1$ ；

4) 计算  $\mathbf{S}_{\text{ta}(k)}$  内每个新状态的耗散值，保留耗散值最小的 3 个对应状态放入  $\mathbf{S}_{\text{n}}$ ，删去  $\mathbf{S}_{\text{ta}}$  中被保留状态的对应元素；

5)  $\mathbf{S}_{\text{ta}(k)}$  的不同状态和  $\mathbf{S}_{\text{eq}}$  中的标签组成新的状态集，放入  $\mathbf{S}_{\text{ta}(k+1)}$ ， $k=k+1$ ；

6) 重复步骤 4)、5)，直至  $\mathbf{S}_{\text{eq}}$  为空；选取  $\mathbf{S}_{\text{n}}$  中耗散值最小的拓扑结构作为最终辨识结果。

将 A\* 搜索算法<sup>[25]</sup>与潮流计算相结合，首先用 A\* 搜索由初始拓扑生成二次辨识的候选拓扑子集，再对集合中的拓扑逐一进行潮流匹配，具体步骤如图 3 所示。

### 2.3 评价指标

多标签学习的准确率  $a_{\text{cc}}$  可以用每类标签预测结果的平均值表示<sup>[26]</sup>。

$$a_{\text{cc}} = \frac{1}{m} \sum_{i=1}^m \frac{|y^{(i)} \cap \hat{y}^{(i)}|}{|\hat{y}^{(i)}|} \quad (8)$$

式中： $m$  为标签样本总数； $y^{(i)}$  表示真实值； $\hat{y}^{(i)}$  表示预测值。在失衡样本中，若采用每类标签预测结果的均值来衡量多标签预测结果的准确性，指标往往会偏向体量大的类别，从而出现一种阳性指标“假高”的现象，因此，为平衡阳性样本与阴性样本的准确率，本文以灵敏度和特异性的几何平均准确率

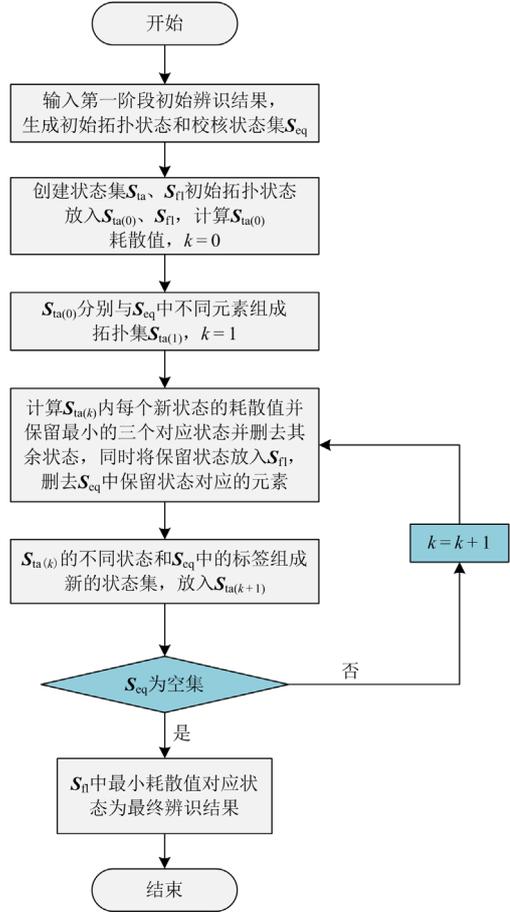


图 3 潮流匹配二次辨识流程图

Fig. 3 Flowchart of power flow matching

$G_{\text{-mean}}$  作为预测精度。

$$A_{\text{GM}} = \frac{1}{m} \sum_{i=1}^m G_{\text{-mean } i} \quad (9)$$

$$G_{\text{-mean } i} = \sqrt{R_{\text{eci}} \times S_{\text{pei}}} \quad (10)$$

式中， $R_{\text{eci}}$ 、 $S_{\text{pei}}$  分别为样本  $i$  的召回率和特异性。

$$R_{\text{eci}} = \frac{TP_i}{TP_i + FN_i} \quad (11)$$

$$S_{\text{pei}} = \frac{TN_i}{TN_i + FP_i} \quad (12)$$

式中， $TP_i$ 、 $TN_i$ 、 $FP_i$ 、 $FN_i$  分别为第  $i$  组样本中的真阳性、真阴性、假阳性、假阴性的数量<sup>[27]</sup>，含有多个样本的多标签分类的召回率、特异性应取其平均值。

$$A_{\text{Rec}} = \frac{1}{m} \sum_{i=1}^m R_{\text{eci}} \quad (13)$$

$$A_{\text{Spe}} = \frac{1}{m} \sum_{i=1}^m S_{\text{pei}} \quad (14)$$

### 3 案例分析

采用图 4 所示的改进 IEEE33 节点配电网对前述模型进行验证<sup>[28]</sup>。

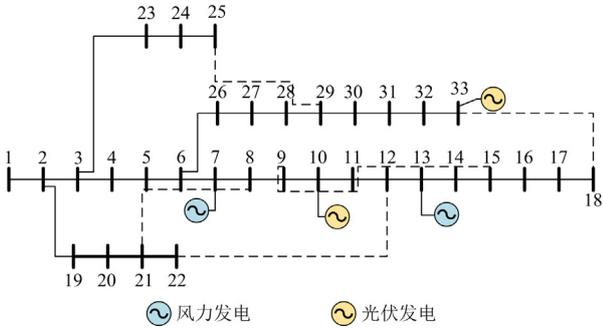


图 4 改进 IEEE33 节点配电网网络

Fig. 4 Improved IEEE33-node distribution network networking

图中分别在节点 7 和 13、节点 10 和 33 加入分布式风、光电源，构建新型配电网。通过联络开关的通、断状态组合，可生成 79 种拓扑，包括 43 种环网、35 种辐射网和 1 种孤岛模式<sup>[29]</sup>。其中，风、光的出力曲线如图 5<sup>[6]</sup>所示，负荷采用广西某地 158 天真实负荷数据，采样间隔为 0.5 h，如图 6 所示，仿真中采用 Matpower 潮流计算，共计生成 15 168 组数据样本，每组样本包含 33 节点的电压幅值和注入功率，采用 Pytorch 搭建辨识模型，使用的处理器为 Intel(R) Core(TM)i7，15 168 组数据样本训练时间为 197 s，单次辨识平均耗时为 0.02 s。

#### 3.1 多标签特征分析

假设配电网在  $t = 50$  s 时发生图 7 所示的拓扑变化，各节点电压幅值变化如图 8 所示。

由图可见，当支路状态发生变化时，不仅支路两端的节点电压发生变化，网络中所有节点电压均

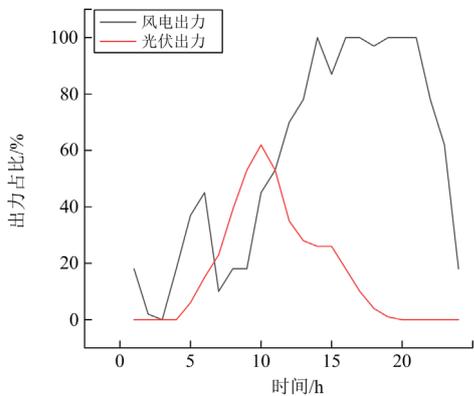


图 5 24 h 风电及光伏出力占比分布

Fig. 5 The 24 h wind power and photovoltaic output proportion distribution

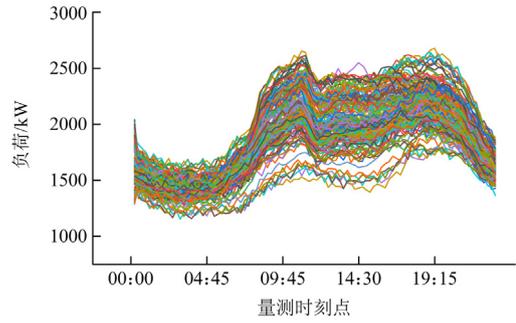


图 6 广西某地 158 天负荷曲线

Fig. 6 The 158-day load curve in a certain place in Guangxi

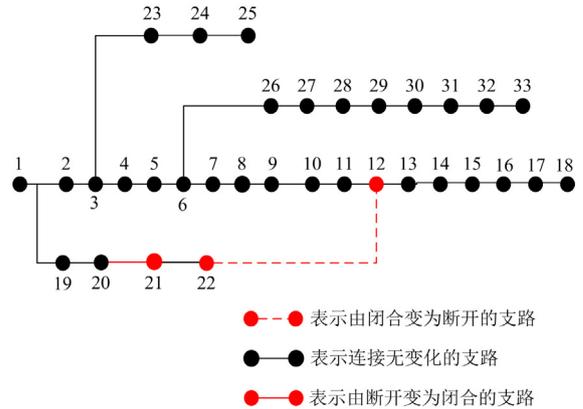


图 7 拓扑变化场景图

Fig. 7 Topology change scenario diagram

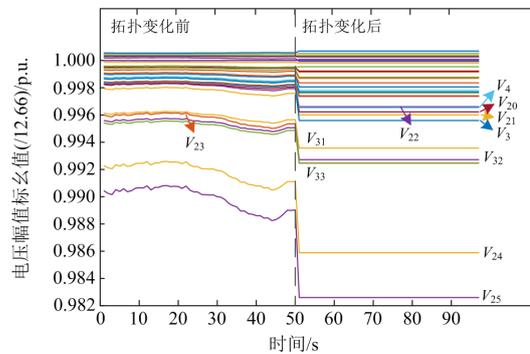


图 8 在图 7 场景下节点电压幅值变化图

Fig. 8 Node voltage change diagram of the Figure 7

有不同程度的波动，即每个节点的电压幅值均与多回支路的状态有关。

#### 3.2 不同损失函数输出得分对比分析

为了对比标签间存在的失衡情况，根据式(2)计算上述样本集中各个标签代表的支路中样本阳性样本占比，其分布如图 9 所示。

可以看到，大部分标签的阳性样本占比在 90% 以上，与阴性样本之间存在显著的失衡。少量标签的阳性占比在 30%~40%之间，失衡情况相对较轻，

这部分标签对应的支路为联络线，状态改变比较灵活，状态分布相对均匀。

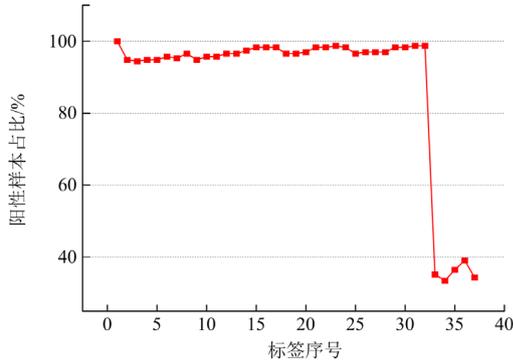


图 9 样本集各支路阳性样本占比分布

Fig. 9 Distribution of positive samples in each branch of the sample set

取常用的 BCEWithLogitsLoss(以下简称 BCEL 函数)损失函数与式(9)的损失函数进行对比,用测试集中 20 个连续的样本时间序列进行计算,得到迭代过程中两种损失值的变化,如图 10 所示。图 11 为分别使用两种损失函数的情况下,基于 MLC-GCN-CNN 初次辨识的标签得分情况。

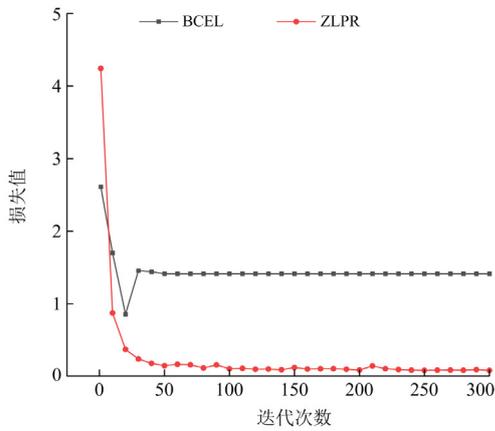


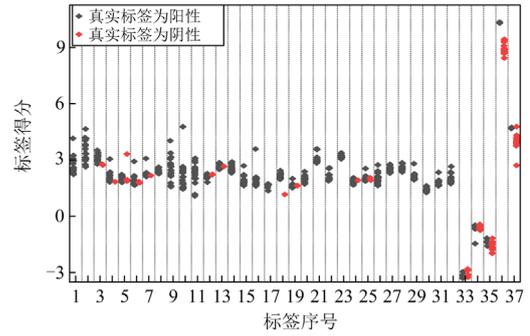
图 10 不同损失函数迭代过程

Fig. 10 Iterative process of different loss functions

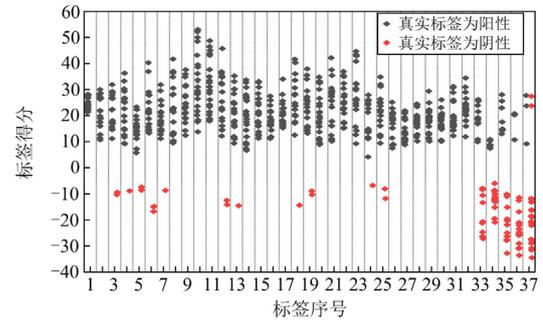
由图 10 可见,使用 BCEL 函数进行反向迭代,经过 50 次迭代后,损失函数值基本收敛到 1.4,整体损失偏大。而使用 ZLPR 函数,经过 50 次后,损失函数值则基本收敛到 0.15,整体损失较小,优于 BCEL 函数。

对比图 11 中两分图可见, BCEL 损失函数输出的各个标签得分未能清晰地划分阳性和阴性状态,难以选择合适的阈值作为输出判定,整体辨识效果较差。由 ZLPR 函数可知,该损失函数考虑了标签阳性、阴性的误差,在训练过程中阈值自适应为 0,

因而对含标签失衡的数据有较好的改善效果。



(a) BCEL 损失函数



(b) ZLPR 损失函数

图 11 不同损失函数的标签得分情况

Fig. 11 Label scores for different loss functions

### 3.3 敏感度分析

为模拟电网正常运行与故障运行时数据不平衡的情况,在上述样本集中的辐射结构、环网结构、孤岛结构各选择一个样本,设置失衡率分别为 5%、26%、50%。真实选取的网络运行状态中还需要考虑量测误差  $e_v\%$ ,  $e_v\%$  分别取 0.1%、0.5%、1%,对两个阶段的输出,辨识精度使用 2.3 节所述多标签中的几何平均值作为评价,测试集选取 5 个时间序列的量测数据。图 12 为计算第一阶段输出辨识结果的特异性和召回率,图 13 是两阶段的几何平均准确率。

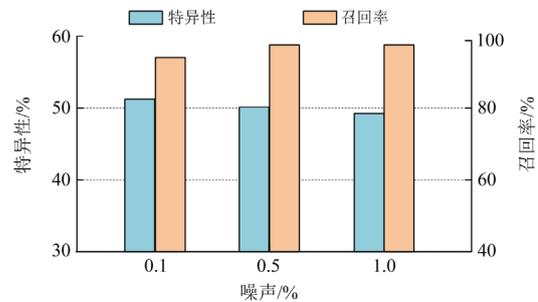


图 12 噪声对 MLC-GCN-CNN 算法的召回率和特异性影响

Fig. 12 Effect of noise on recall and specificity of MLC-GCN-CNN algorithms

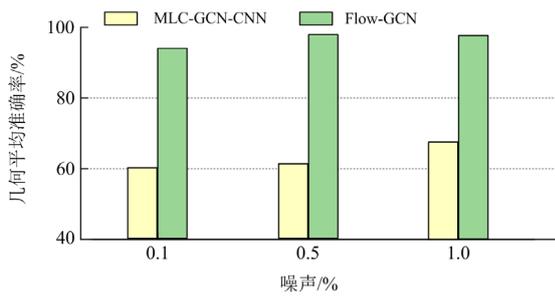


图 13 噪声对两辨识阶段的影响

Fig. 13 Influence of noise on two stages of identification

可以看到, 第一阶段中, 虽然表征阴性样本准确率的特异性指标偏低, 但由于较高的召回率保证了辨识结果中阳性样本的正确率, 使得模型在第二阶段并在保证阳性样本正确率的条件下, 对辨识错误的阴性样本标签加以修正, 因而展现出二次辨识后几何平均指标有较大提升, 整体辨识精度得到提高。

### 3.4 算法对比仿真

为进一步论证本文所提模型的有效性与实用性, 本节选取 DNN、SVM、CNN 单标签分类算法进行拓扑辨识的模型与 MLC-GCN-CNN 的多标签辨识方法作横向对比, 计算结果如图 14 所示。

可以看到, 将类别间的失衡转化为标签间的开合比失衡后, 即使加入较大程度的类别失衡数据, 标签间开合失衡比变化也较小, 对各个标签阳性占比分布影响不大, 只有部分支路有微小变化, 这是因为将大类细分分为多支路标签后, 即便较高的失衡比下也只是相当于各个标签闭合次数发生变化, 并且在总样本多的情况下这种变化不大。

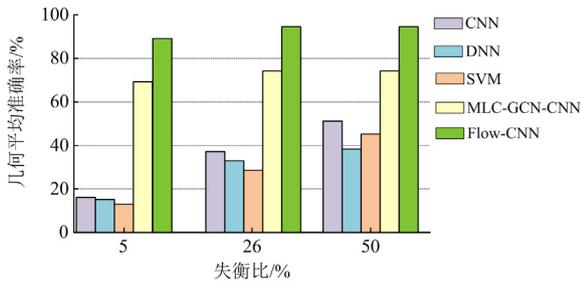


图 14 不同算法仿真结果对比

Fig. 14 Simulation results comparison of different algorithms

### 3.5 未知拓扑仿真

为进一步检验模型的泛化能力, 选取训练集中不存在的拓扑类型, 使用该结构下 1 个时间序列的样本进行测试。经初步辨识得到的拓扑与经过二次辨识后的拓扑如图 15 所示。

可以看到, 初步辨识的结果中, 支路 18-33 和支路 8-21 均出现了“假阴性”, 而经过二次辨识后,

则排除了假阴性, 得到与真实拓扑一致的辨识结果。二次搜索的过程如图 16 所示。

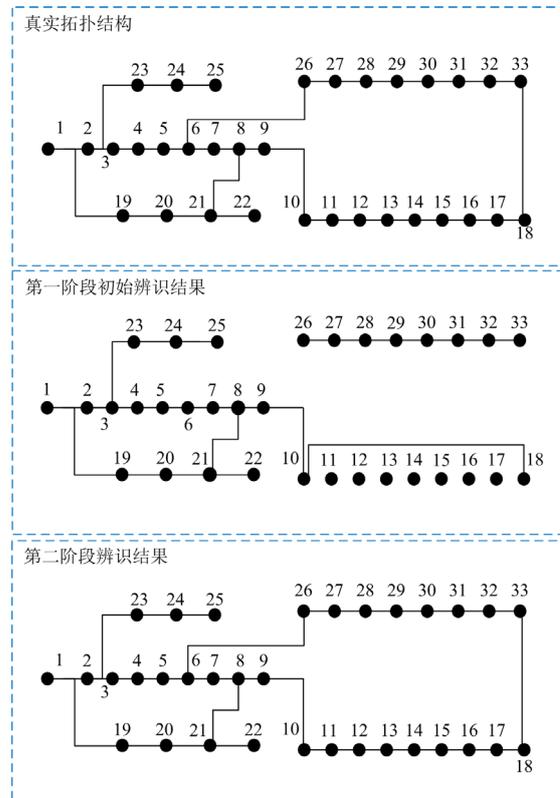


图 15 未知样本集真实结构与两阶段辨识输出

Fig. 15 Unknown sample set true structure and two-stage identification output

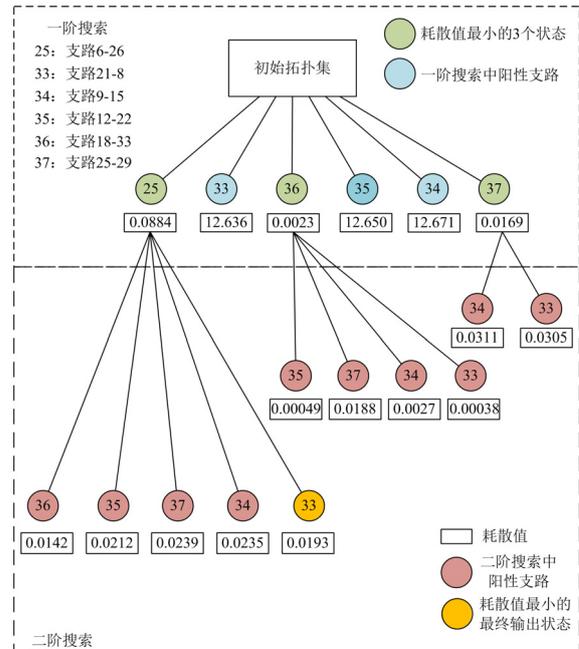


图 16 未知拓扑结构搜索过程耗散值树状图

Fig. 16 Tree view of dissipated values during unknown topology

## 4 结论

本文针对新型配电网不同运行方式下量测数据失衡的特点,提出了一种“初步拓扑辨识+潮流匹配”的两阶段配电网拓扑辨识模型,经过算例仿真结果表明:

1) 所提拓扑辨识模型在采样周期较短、量测误差较大且存在数据失衡的情况下,仍具有较高的辨识精度;

2) 对于历史数据中未出现的拓扑场景,所提模型依然能辨识出真实拓扑,具有较好的泛化能力。

综上所述,本文所提两阶段拓扑辨识模型在失衡条件下有较强的鲁棒性和泛化能力,所用量测采样周期短,更适用于当前变化频繁、拓扑复杂的配电网运行模式,具有较高的实际价值。

## 参考文献

- [1] 王利利, 王皓, 任洲洋, 等. 计及灵活资源调节潜力的高压配电网新能源接纳能力评估[J]. 中国电力, 2022, 55(10): 124-131.  
WANG Lili, WANG Hao, REN Zhouyang, et al. Evaluation of renewable energy accommodation capacity of high voltage distribution networks considering regulation potential of flexible resources[J]. Electric Power, 2022, 55(10): 124-131.
- [2] 廖小兵, 周自强, 乐健, 等. 考虑电池储能老化和需求侧响应的交直流混合配电网动态重构方法[J]. 电力系统保护与控制, 2023, 51(14): 43-54.  
LIAO Xiaobing, ZHOU Ziqiang, LE Jian, et al. Dynamic reconfiguration of AC/DC hybrid distribution network considering battery energy storage aging and demand response[J]. Power System Protection and Control, 2023, 51(14): 43-54.
- [3] 段建东, 孙茜, 陈鲁鹏, 等. 基于改进和声搜索算法的有源配电网优化调度[J]. 高压电器, 2021, 57(11): 1-9.  
DUAN Jiandong, SUN Qian, CHEN Lupeng, et al. Optimal dispatching of active distribution network based on improved harmony search algorithms[J]. High Voltage Apparatus, 2021, 57(11): 1-9.
- [4] 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.
- [5] ZHOU L, LI Q, ZHANG Y, et al. Consumer phase identification under incomplete data condition with dimensional calibration[J]. International Journal of Electrical Power and Energy Systems, 2021, 129: 106851.
- [6] 章振宇. 配电自动化系统建设效果综合评估[D]. 南昌: 南昌大学, 2019.
- [7] 任鹏哲, 刘友波, 刘挺坚, 等. 基于互信息贝叶斯网络的配电网拓扑鲁棒辨识算法[J]. 电力系统自动化, 2021, 45(9): 55-62.  
REN Pengzhe, LIU Youbo, LIU Tingjian, et al. Robust identification algorithm for distribution network topology based on mutual-information Bayesian network[J]. Automation of Electric Power Systems, 2021, 45(9): 55-62.
- [8] 潘毓笙, 秦超. 基于两阶段特征选择和格拉姆角场的配电网拓扑辨识方法[J]. 电力系统自动化, 2022, 46(16): 170-177.  
PAN Yusheng, QIN Chao. Identification method for distribution network topology based on two-stage feature selection and Gramian angular field[J]. Automation of Electric Power Systems, 2022, 46(16): 170-177.
- [9] 蒋玮, 汤海波, 祁晖, 等. 基于集成深度神经网络的配电网联络关系辨识技术[J]. 电力系统自动化, 2020, 44(1): 101-108.  
JIANG Wei, TANG Haibo, QI Hui, et al. Distribution network connectivity recognition based on ensemble deep neural network[J]. Automation of Electric Power Systems, 2020, 44(1): 101-108.
- [10] 唐捷, 蔡永智, 李其霖, 等. 基于三相表特征约束聚类的低压台区用户相序识别方法[J]. 电力系统自动化, 2022, 46(8): 60-67.  
TANG Jie, CAI Yongzhi, LI Qilin, et al. Phase identification of low-voltage transformer area based on constrained clustering considering the feature of three-phase meter[J]. Automation of Electric Power Systems, 2022, 46(8): 60-67.
- [11] ZHOU L, ZHANG Y, LIU S, et al. Consumer phase identification in low-voltage distribution network considering vacant users[J]. International Journal of Electrical Power & Energy Systems, 2020, 121: 106079.
- [12] LIU Y, WANG P. Partial correlation analysis based distribution network topology identification[J]. CSEE Journal of Power and Energy Systems, 2023, 9(4): 1493-1504.
- [13] ZHANG X, HUO M, LI H, et al. Probabilistic topology identification combining state estimation and data-driven approaches[C] // 2022 IEEE 3rd China International Youth Conference on Electrical Engineering (CIYCEE), 2022, Wuhan, China.
- [14] ZHANG J, WANG Y, ZHANG N. 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.
- [15] 邵晨颖, 刘友波, 邵安海, 等. 基于生成对抗网络与局部电流相量的配电网拓扑鲁棒辨识[J]. 电力系统自动化, 2023, 47(1): 55-62.  
SHAO Chenying, LIU Youbo, SHAO Anhai, et al. Robust identification for distribution network topology based on generative adversarial network and partial current phasor[J]. Automation of Electric Power Systems, 2023, 47(1): 55-62.
- [16] 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.
- [17] 王铮澄, 周艳真, 郭庆来, 等. 考虑电力系统拓扑变化的消息传递神经网络暂态稳定评估[J]. 中国电机工程学报, 2021, 41(7): 2341-2350.  
WANG Zhengcheng, ZHOU Yanzhen, GUO Qinglai, et al. Transient stability assessment of power system considering topological change: a message passing neural network-based approach[J]. Proceedings of the CSEE, 2021, 41(7): 2341-2350.
- [18] ZHANG Z, LI Y, ZHANG J, et al. A topology identification and impedance estimation method for distribution network with distributed generations[J]. IFAC-Papers OnLine, 2020.
- [19] 王建, 姚江宁, 刘泽青, 等. 恶劣天气下配电网故障统计分析及其概率分布拟合[J]. 电力系统保护与控制, 2022, 50(17): 143-153.  
WANG Jian, YAO Jiangning, LIU Zeqing, et al. Fault statistical analysis and probability distribution fitting for a power distribution network in adverse weather conditions[J]. Power System Protection and Control, 2022, 50(17): 143-153.
- [20] 肖华昕. 基于改进生成对抗网络的多类别不平衡学习研究[D]. 天津: 河北工业大学, 2021.
- [21] 杨敏航, 陈龙, 刘慧, 等. 基于图卷积网络的多标签遥感图像分类[J]. 计算机应用研究, 2021, 38(11): 3439-3445.  
YANG Minhang, CHEN Long, LIU Hui, et al. Multi-label remote sensing image classification based on graph convolutional network[J]. Application Research of Computers, 2021, 38(11): 3439-3445.
- [22] KIPF T, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv: 1609.02907, 2017.
- [23] SU J, ZHU M, MURTADHA A, et al. ZLPR: a novel loss for multi-label classification[J]. arXiv preprint arXiv: 2208.02955, 2022.
- [24] 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.
- [25] 路永鑫, 魏云冰, 赵启承, 等. 基于层次分析法和改进A\*算法的电力应急机器人路径规划[J]. 电力系统保护与控制, 2021, 49(9): 82-89.  
LU Yongxin, WEI Yunbing, ZHAO Qicheng, et al. Path planning of a power emergency robot based on an analytic hierarchy process and improved A\* algorithm[J]. Power System Protection and Control, 2021, 49(9): 82-89.
- [26] 王岳, 李雅文, 李昂. 科技资源文本层次多标签分类方法[J]. 计算机工程与应用, 2023, 59(13): 92-98.  
WANG Yue, LI Yawen, LI Ang. Academic resource text hierarchical multi-label classification[J]. Computer Engineering and Applications, 2023, 59(13): 92-98.
- [27] 李文楷, 刘原池, 刘子越, 等. 基于正样本-背景数据的校正混淆矩阵[J]. 海南大学学报(自然科学版), 2023, 41(3): 293-302.  
LI Wenkai, LIU Yuanchi, LIU Ziyue, et al. A calibrated confusion matrix based on positive and background data[J]. Natural Science Journal of Hainan University, 2023, 41(3): 293-302.
- [28] 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.
- [29] MA L, WANG L, LIU Z. Topology identification of distribution networks using a split-EM based data-driven approach[J]. IEEE Transactions on Power Systems, 2022, 37(3): 2019-2031.

收稿日期: 2023-05-04; 修回日期: 2023-07-31

作者简介:

陈碧云(1978—), 女, 通信作者, 博士, 副教授, 研究方向为智能电网与综合能源系统的规划与运行、电力大数据应用; E-mail: chenbiyu@gxu.edu.cn

吕怡博(1996—), 女, 硕士研究生, 研究方向为智能配电网、电力大数据应用; E-mail: lv\_yibo96@163.com

梁志坚(1969—), 男, 硕士, 副教授, 研究方向为新型电力系统保护与控制。

(编辑 姜新丽)