

# 基于多源信息融合的配网故障 外部原因识别

谭媛, 张文海, 王杨  
(四川大学电气工程学院, 成都市 610065)

**摘要:** 配网故障原因的准确识别对于缩短故障查找时间、提高供电恢复速度和供电可靠性有重要意义。根据责任归属可将配网故障原因分为内部原因和外部原因, 内部原因指设备绝缘弱化、过电压等电气相关原因, 而外部故障通常由于天气、动物或人类活动引起。由于外部原因导致的故障是多种因素共同作用的结果, 为此提出融合线路参数、天气、时间等非电气量信息及动作电流、故障相数等电气量信息的故障外部原因识别方法, 首先对5种典型外部故障原因的特点及相关影响因素进行分析, 构建识别模型, 然后使用无监督学习训练得到深度信念网络各层的最优参数, 利用有监督学习对全局参数进行微调, 得到基于深度信念网络的配网故障外部原因识别模型, 最后利用西部某地区的实际故障数据对算法的准确性进行了验证, 结果显示识别准确率可达94.82%, 证明了方法的正确性。

**关键词:** 配网故障; 外部原因; 原因识别; 深度学习; 多源信息融合

## Distribution System External Fault Causes Identification based on Multi-Source Information Fusion

TAN Yuan, ZHANG Wenhai, WANG Yang

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

**ABSTRACT:** Accurate identification of distribution network fault causes is of great significance to shorten the time of fault search and improve the speed of power supply recovery and power supply reliability. According to the attribution of responsibility, the causes of distribution network fault can be divided into internal causes and external causes. Internal causes refer to electrical related causes such as equipment insulation weakening and overvoltage, while external faults usually refer to weather, animal or human activities. Faults caused by external causes are the result of a combination of many factors. In this paper, a fault external cause identification method is proposed, which integrates the non-electrical information such as line parameters, weather and time, and the electrical information such as fault current and phase number. Firstly, the characteristics and related influencing factors of five typical external fault causes are analyzed to build the basis of identification model. Then unsupervised learning training is used to obtain the optimal parameters of each layer of deep belief network, and supervised learning is used to fine-tune the global parameters, and the classification model of external causes of distribution network faults based on deep belief network is obtained. Finally, the accuracy of the algorithm is verified by using the actual fault data of a certain area in western China. The result shows that the recognition accuracy can reach 94.82%, which proves the correctness of the method.

This work is supported by National Key R&D Program of China (No. 2020YFF0305800) and Sichuan Natural Science Foundation (No. 2022NSFSC0234).

**KEYWORDS:** distribution system fault; external fault causes; fault cause identification; deep learning; multi-source information fusion

中图分类号: TM 773

文献标志码: A

文章编号: 1000-7229(2023)03-0077-08

DOI: 10.12204/j.issn.1000-7229.2023.03.008

## 0 引言

10 kV 配电线路作为配电网的重要组成部分, 量

大、面广、线长、运行环境复杂, 故障跳闸率居高不下, 据统计, 80%以上的停电均是由配网线路故障引起的<sup>[1]</sup>, 配网线路故障已成为影响供电连续性和可靠性的重要因素。配网线路发生故障后, 主要依靠人工巡视查找故障点, 然而该过程费时耗力, 相比将故障设备从系统中安全隔离以及故障处理的时间, 停电时

基金项目: 国家重点研发计划项目(2020YFF0305800); 四川省自然科学基金(2022NSFSC0234)

<http://www.cepc.com.cn>

间大部分消耗在查找故障原因和故障点上。目前,针对配网故障定位已开展大量研究<sup>[2-4]</sup>,相继提出了大量定位算法,取得了一定的效果。然而针对配网故障原因识别的研究相对较少<sup>[5-9]</sup>,实际应用中,有效的故障原因识别将对缩短故障排查时间、提高供电可靠性有积极意义<sup>[5]</sup>。

配网故障原因根据责任划分可分为内部原因和外部原因,其中内部原因包括设备质量欠佳、设备老化击穿等电气相关原因,而外部原因引发的故障通常指由于天气、动物或人类活动等外部事件引起的故障,具体原因主要包括车辆、动物、雷击、施工、树木5大类。据西部某地区近7年配网线路故障统计数据可知,由外部原因引发的配网故障占比超过50%。现有方法通常基于故障录波数据对由动物、雷击、车辆、树木原因引发的故障进行识别<sup>[6-9]</sup>。该类方法一定程度上实现了故障原因识别,其本质上利用了不同故障原因时,“故障模型”的差异在波形特征上的反映,然而实际故障模型非常复杂,故障波形包含的故障信息有限,且不同故障原因造成的故障波形也可能存在相似,识别准确度受限。实际分析发现,树木<sup>[10]</sup>、动物<sup>[11]</sup>、雷击引起的故障通常还与天气、时间等非电气量有关<sup>[12]</sup>,因此融合多种信息进行故障原因识别将有利于提升识别准确度。

近年来,多源信息融合的思想在军事、交通、电力系统等方面的应用较多,在配网故障诊断方面的研究主要集中在故障定位上,然而在配网故障原因识别方面尚未有相关研究。

由于配电网故障的发生伴随着许多偶然性和必然性因素,受影响因素多(天气、温度、线路类型等),因此基于多源信息融合进行配网故障原因识别将有效提升识别准确性。如今机器学习快速发展,如逻辑回归算法<sup>[13]</sup>、神经网络算法<sup>[14]</sup>、支持向量机算法<sup>[15]</sup>,但这些方法的特征提取能力还存在不易获得全局最优解而导致模型泛化能力差、识别准确率低等不足<sup>[16]</sup>。而深度学习神经网络是一种更深层的机器学习方法,有更好的特征提取和样本学习能力,结合具有更多维度的输入样本数据,为融合多源信息实现配网故障原因识别提供了一种新手段。

本文针对由外部原因引发的故障,融合由电气量和非电气量构成的多源信息,涉及保护动作信息、天气、时间、线路参数等4方面信息,提出一种基于深度学习的多源信息融合配网故障外部原因识别方法,进行数据融合的信息更全面、更多源,并用实际故障数据对算法的准确性进行验证,证实该方法的正确性和可行性。

## 1 配网故障外部原因分析与特征提取

### 1.1 外部事件引发的配网故障原因分析

引发配网故障的外部原因主要包括车辆、动物、雷击、施工、树木。本文收集了2015年至2021年西部某区配网线路故障原因统计情况,如图1所示。其中,所有故障原因均已由专业运检人员现场核查确定,均为单一故障原因,故障原因不明或不确定的故障数据未纳入论文分析。该区共发生故障1723次,其中外部事件引发的故障903次,占比52.41%,各外部故障原因具体占比如图1所示。

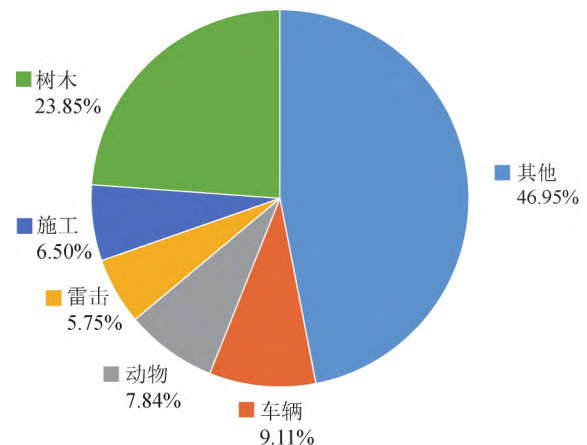


图1 某区2015—2021年不同故障原因占比  
Fig.1 Proportion of different fault causes in a district from 2015 to 2021

1) 车辆引起的故障。车辆引起的故障主要指车辆撞电线杆造成导线交缠引发相间短路故障以及施工车辆车臂挂线造成接地故障,常发生在白天,天气通常为多云或阴天,此类天气有利于车辆行驶及车辆施工,因此车流量大、车辆施工作业多,故车辆导致的配网故障也随之增多。

2) 动物引起的故障。动物攀爬电力设备导致绝缘距离缩短引起接地故障或相间短路故障,天气通常为多云或阴天,高温和强降雨时动物外出活动减少,因此动物造成的配网故障也大幅减少。动物引起的故障主要是开关柜和环网柜故障,柜内电力设备运行时产生热量,冬季吸引动物来取暖,因此动物导致的配网故障最多发生在冬季。

3) 雷击引起的故障。雷击导致设备过电压,从而导致设备绝缘击穿,引起接地故障或相间短路故障,该类故障通常发生在夏季的雷雨天气,主要集中在4到9月,其余时间发生雷击故障较少。

4) 施工引起的故障。施工引起的故障主要指土建施工挖伤电缆引发的故障,其中单相故障居多。该类故障主要发生在白天,天气通常为多云或阴天。炎

热天气和大雨天气时,现场施工较少,因施工导致的配网故障也随之减少。

5) 树木引起的故障。树枝触线造成接地或相间短路故障,该类故障常发生在夏季的雷雨天气。当树枝被雨水淋湿,绝缘能力迅速下降,在风力作用下容易造成碰线,从而引发故障。

### 1.2 故障特征信息提取

为进一步分析不同故障原因与相关因素之间的关系,分别从天气、时间、线路参数、保护动作信息四方面进行了梳理,各故障原因与各因素之间的相关性如图 2 至图 7 所示,相关性总结如表 1 所示。

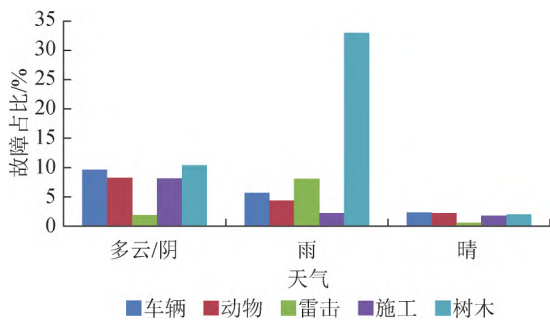


图 2 外部故障原因与天气相关性对比  
Fig. 2 Correlation analysis of external fault causes with weather

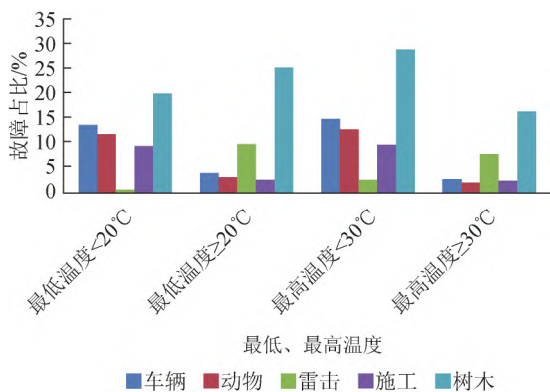


图 3 外部故障原因与最低、最高温度相关性对比  
Fig. 3 Correlation analysis of external fault causes with minimum and maximum temperature

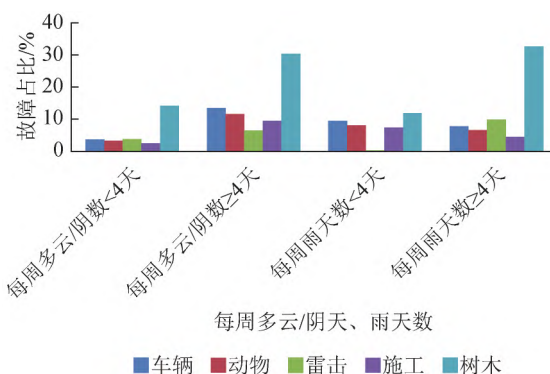


图 4 外部故障原因与每周多云/阴天、雨天数相关性对比  
Fig. 4 Correlation analysis of external fault causes with number of cloudy and rainy days per week

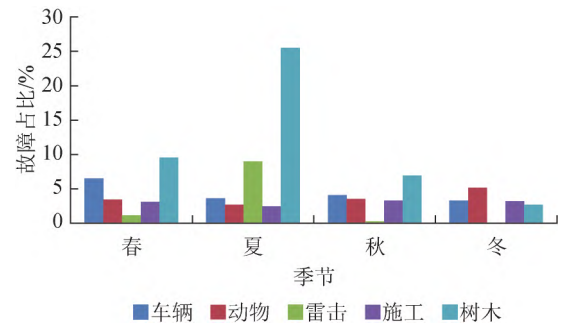


图 5 外部故障原因与季节相关性对比  
Fig. 5 Correlation analysis of external faults causes with season

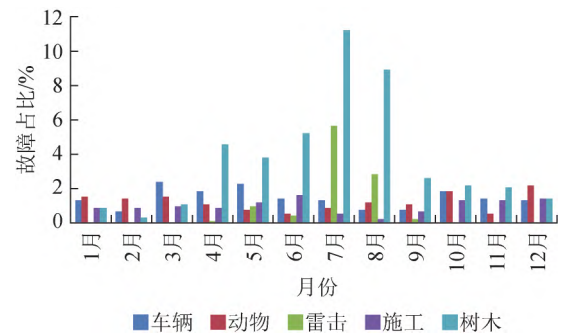


图 6 外部故障原因与月份相关性对比  
Fig. 6 Correlation analysis of external faults causes with month

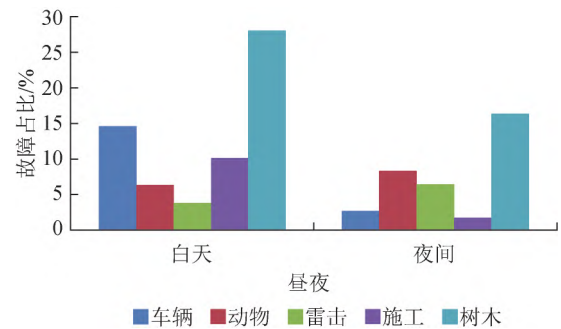


图 7 外部故障原因与昼夜相关性对比  
Fig. 7 Correlation analysis of external faults causes with time of day

1) 天气信息: 天气、最高温度、最低温度、每周多云/阴天数、每周雨天数。

由图 2 至图 4 可知,施工、车辆、动物引起的故障在天气为多云或阴天、每周多云/阴天数超过 4 天、环境最低温度低于 20℃ 时易发生;雷击和树木引起的故障在天气为雨天、每周雨天数超过 4 天、环境最低温度高于 20℃ 时易发生,其中树木引起的故障在环境最高温度低于 30℃ 时易发生,而雷击引起的故障在环境最高温度高于 30℃ 时易发生。

2) 时间信息: 季节、月份、昼夜。

如图 5 至图 7 所示,施工和车辆引起的故障都是人为导致的,故集中在白天;雷击和树木引起的故障常发在夏季 7 月最多;动物引起的故障常发生在冬季,

表 1 故障原因与特征信息关系  
Table 1 Relationship between fault causes and feature information

| 故障原因 | 故障相数     | 动作电流               | 重合闸是否成功 | 线路地区特征 | 线路类型   | 季节 | 月份 | 昼夜 | 天气   | 最低温度 | 最高温度 | 每周多云/阴天数 | 每周雨天数 |
|------|----------|--------------------|---------|--------|--------|----|----|----|------|------|------|----------|-------|
| 车辆   | 单相、两相、三相 | 多为 2 000 A~7 000 A | 是/否     | 农村、城区  | 架空线    |    |    | 白天 | 多云/阴 | <20℃ | <30℃ | ≥4       | <4    |
| 动物   | 多为两相、三相  | 多为 2 000 A~9 000 A | 是/否     | 农村、城区  | 架空线、电缆 | 冬  | 12 |    | 多云/阴 | <20℃ | <30℃ | ≥4       | <4    |
| 雷击   | 单相、两相、三相 | 多为 2 000 A~6 000 A | 是/否     | 农村     | 架空线    | 夏  | 7  |    | 雨    | ≥20℃ | ≥30℃ | <4       | ≥4    |
| 树木   | 单相、两相、三相 | 多为 2 000 A~7 000 A | 是/否     | 农村     | 架空线    | 夏  | 7  |    | 雨    | ≥20℃ | <30℃ | <4       | ≥4    |
| 施工   | 多为单相     | 多为 2 000 A~7 000 A | 否       | 城区     | 电缆     |    |    | 白天 | 多云/阴 | <20℃ | <30℃ | ≥4       | <4    |

12 月最多。

3) 线路信息: 类型、地区特征。

施工引起的故障常发生在城区电缆区域, 树木和雷击引起的故障常发生在农村架空线区域, 车辆引起的故障发生在城区和农村的架空线区域, 动物引起的故障常发生在城区电缆区域和农村架空线区域。

4) 保护动作信息: 故障相数、动作电流、重合闸是否成功。

故障相数包括单相、两相、三相, 施工引起的故障相数常为单相, 动物引起的故障常为两相或三相。施工对设备损伤通常很严重, 重合闸动作会失败。动物引起的故障通常发生在开关柜和环网柜内, 为弧光短路故障, 电流通常很大。

## 2 基于多源信息融合的配网故障外部原因识别方法

### 2.1 深度学习理论

深度信念网络( deep belief network, DBN) 作为深度学习的主要实现方法之一, 其模型通过受限玻尔兹曼机( restricted Boltzmann machine, RBM) 的顺序堆叠来构造, 通过对 RBM 进行逐层的无监督训练, 再利用反向传播算法对整个网络进行有监督训练, 通过梯度下降法, 使整个网络的误差最小<sup>[17]</sup>, 达到网络的整体最优。其结构如图 8 所示, 包括输入层( 状态信息输入)、RBM 层( 状态信息训练与学习) 和输出层( 信息输出)。

### 2.2 数据样本获取

本文用于深度信念网络的原始数据集是从西部某电力公司 13 个站, 143 条 10 kV 线路中获取的, 从 2015 至 2021 年的 10 kV 配网真实故障样本数, 共计 1 723 条。故障原始数据由电力公司运检部的配网故障统计中获得, 由于故障数据集是人为记录的, 存在数据缺失、重复、错误等信息, 因此作者进一步从多个部门( 配网、调度、运维) 以及保护装置、监控平台等设备中对每一条故障记录数据进行了核查和筛选, 从原始 1 821 条数据中筛选出 1 723 条准确的且包含完

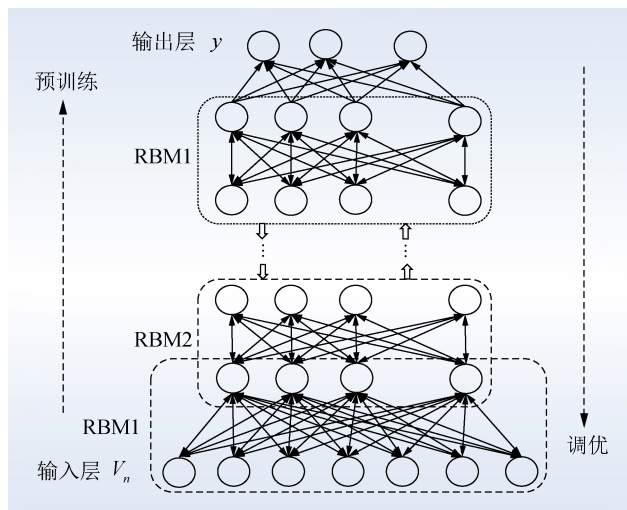


图 8 深度信念网络架构  
Fig. 8 Structure of DBN

整故障信息的故障数据, 基于这些数据开展本文的相关研究。

要利用深度信念网络对故障数据进行分析, 关键要对故障事件的多源特征信息进行提取。根据 1.2 节可知, 本文所运用的多源信息, 主要包含 4 类信息, 即每个故障事件发生时的非电气量信息( 包括天气信息、时间信息、故障线路信息) 和电气量信息( 包括保护动作信息)。其中, 天气信息通过气象平台获取, 时间信息通过配网、调度和运维的故障事件记录获取, 故障线路信息通过电力系统设备资产运维精益管理系统获取, 保护动作信息通过保护装置和调度监控平台获取。

### 2.3 数据样本预处理

#### 2.3.1 输入数据预处理

上述的多源信息涉及多种不同的数据格式, 为了提高学习精度, 加速模型的收敛速度, 首先需要对上述多源信息进行数据预处理。将样本数据分为两种类型, 第一种是非数值型数据, 如线路类型、季节等, 针对这一类型数据, 本文使用深度学习常用的 one-hot<sup>[18]</sup> 编码方式处理该类输入数据; 第二种是数值型数据, 如月份、电流等, 对这些数据采用最小最大归一化方法<sup>[19]</sup> 对数据进行预处理, 并将特

征数据  $x_i$  缩放到  $[0, 1]$  的范围内。通过上述数据预处理可以获得输入的特征向量  $X_i = \{x_1, x_2, x_3, \dots, x_i\}$ 。再结合权重向量矩阵  $T_i$  对样本数据进行进一步处理,从而使得主要相关指标增强,不相关的指标减弱。权重向量矩阵  $T_i$  如式(1)所示,  $r_{ij}$  为对应特征数据的权重值:

$$T_i = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1l} \\ r_{21} & r_{22} & \dots & r_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ r_{il} & r_{i2} & \dots & r_{il} \end{bmatrix} \quad (1)$$

通过权重向量矩阵加权后,计算得到新的特征向量  $X'_i$ :

$$X'_i = \{x'_1, x'_2, x'_3, \dots, x'_i\} \quad (2)$$

式中:  $x'_i$  表示加权后新的特征数据。

### 2.3.2 配网故障外部原因编码

故障原因识别是多分类任务,识别结果可以分为各种故障原因类型。下面对不同的原因类型进行编码,以 F1—F6 代来表示,如表 2 所示。

表 2 故障原因编码  
Table 2 Coding table of failure causes

| 故障原因 | 编码            | 代码 |
|------|---------------|----|
| 车辆   | (1,0,0,0,0,0) | F1 |
| 动物   | (0,1,0,0,0,0) | F2 |
| 雷击   | (0,0,1,0,0,0) | F3 |
| 施工   | (0,0,0,1,0,0) | F4 |
| 树木   | (0,0,0,0,1,0) | F5 |
| 其他   | (0,0,0,0,0,1) | F6 |

## 2.4 故障原因识别模型

### 2.4.1 故障原因识别流程

本文提出的基于多源信息融合的配网故障外部原因识别流程如图 9 所示。

主要步骤如下:

1) 采集配网故障发生时故障线路的电气量信息和非电气量信息。

2) 将多源信息融合后的原始数据样本进行编码、归一化处理得到故障样本数据集。

3) 将故障样本数据集划分为训练、微调样本数据集及测试样本数据集。

4) 将训练、微调样本数据集输入初始化的深度信念神经网络进行模型训练,得到具有最优参数的深度神经网络模型。

5) 将测试样本数据集输入训练完成后的模型,识别该样本数据对应的配网故障发生时的外部原因,实现故障原因的准确识别。



图 9 故障原因识别流程  
Fig. 9 Process of fault cause identification

### 2.4.2 模型训练过程

在上述故障识别原因流程的基础上,对深度神经网络模型的训练过程进行细化分析。假定输入层神经元为  $l$  个,将天气、时间、线路信息、保护动作信息等多源数据融合后,归一化处理得到的数据样本生成特征向量  $X_i$  输入神经网络,假定输出层神经元个数为  $j$ ,  $y_j (j = 1, 2, 3, \dots, m)$  表示输出层神经元的输出,隐藏层神经元  $k$  个,  $d_k$  表示隐藏层的输出。基于神经网络模型的训练过程如下:

#### 1) 参数初始化。

随机初始化权重矩阵  $V$ 、 $X$ , 设置采样模式  $P$  和训练次数为 1, 设置误差  $E$  为 0, 设置学习率  $\eta$  为介于 0 到 1 之间的小数,网络训练后的准确率  $E_{\min}$  设置为正数。

#### 2) 输入训练样本,计算 $n$ 个样本在各层的输出。

$$d_k = f_1 \left\{ \sum_{m=1}^n v_{lk} \text{sigmoid} [\text{Conv} [W \tanh(W + X_i r_i * h_{i-1})] \otimes X_i \oplus X_i] \right\} \quad (3)$$

$$y_j = f_2 \left( \sum_{m=1}^n w_{kj} d_k \right) \quad (4)$$

式中:  $r_i$  表示权重值;  $h_i$  表示隐藏层节点的状态向量;  $v_{lk}$  表示输入层和隐藏层的神经元之间的连接权重;  $w_{kj}$  表示隐藏层和输出层神经元的连接权重;  $f_1(x)$  和  $f_2(x)$  分别为隐藏层和输出层的激活函数。sigmoid 和 Conv 分别表示 sigmoid 函数、卷积函数; \* 表示卷积运算;  $\otimes$  表示克罗内积;  $\oplus$  表示矩阵加法。

### 3) 误差计算。

误差转移包括输入层  $X_i$  到输出  $y_j$ 。当输入  $n$  个训练样本, 神经网络的误差  $E$  可由式(5) 计算得到:

$$E = \frac{1}{n} \sum_{m=1}^n E(e) \quad (5)$$

式中:  $E(e)$  是样本  $m$  的训练误差, 计算过程如下:

$$E(e) = \frac{1}{2} \sum_{k=1}^n (d_k - y_j)^2 \quad (6)$$

### 4) 权重调整。

使用梯度下降算法更新各层神经元之间的连接权重, 其中  $v_{lk}$  和  $w_{kj}$  的梯度  $\Delta v_{lk}$  和  $\Delta w_{kj}$  的计算过程如下:

$$\Delta w_{kj} = \eta y_j (d_k - y_j) (1 - y_j) \quad (7)$$

$$\Delta v_{lk} = \eta y_j (1 - y_j) \sum_{m=1}^n y_j \Delta w_{kj} d_k \quad (8)$$

完成上述训练流程以后, 神经网络中各层神经元的参数得以调整与优化, 即可得到最优神经网络模型, 就可以记录、体现训练样本中故障发生时的多源融合信息与故障原因之间的关联关系。当再次发生故障时, 输入故障发生时相应的融合信息数据, 即可通过该神经网络模型的计算输出故障原因类型。

## 3 算例分析

基于多源信息融合的配网故障外部原因识别模型由多层的神经网络组成, 模型的输入为故障发生时获取的多源信息数据样本(经预处理), 最后的输出为样本对应于各个故障原因的概率值, 概率值最大的状态即为识别出来的故障原因。

根据现有的研究可知, 影响深度信念网络算法识别准确率的主要因素包括训练样本的数量、神经网络层数和训练迭代次数等超参数的设置<sup>[20]</sup>。本文重点展示了神经网络层数和训练迭代次数对识别准确率的影响, 同时验证了算法的收敛性, 并将调试好后的神经网络算法与现有的其他四种算法进行了比较, 验证了本文提出的基于多源信息融合的配网故障外部原因识别方法的有效性。

为了更加准确地记录故障发生时配网的运行状况及外部环境状况, 准确利用多源信息识别配网故障原因, 本文将故障发生当日内变化较快的电气量信息按每半小时的时间频率进行采样收集, 对当日内变化较慢的非电气量信息按每小时的时间频率采样收集, 获取了实际运行网络中经处理后的 1 723 条 250 维数据样本。针对 DBN 的无监督训练、有监督训练以及验证过程, 根据已有经验, 将数据样本按照 7:2:1 的比例划分预训练集、微调集和测试集。同时设定预训练及微调过程的迭代次数为 100 次, 学习率设置

为 0.1。

### 3.1 神经网络层数对识别准确率的影响

隐藏层的数量是深度神经网络的关键参数之一, 对学习过程中样本的特征提取有着巨大影响, 但该参数的选定暂无准确的理论指引。通常说来, 随着隐藏层数量的增大, 网络的训练时间会增长, 学习效率会降低, 并且更容易出现过拟合的现象; 但隐藏层的数量过少, 会导致对样本的特征提取不足, 从而导致学习算法在测试集上的表现不佳。深度信念网络相较于传统神经网络, 在结构上更为复杂, 其网络层数一般在 3 层以上。根据本文的训练规模, 从 3 层开始验证了网络层数对识别准确率的影响, 结果如表 3 所示。

表 3 网络层数对识别准确率的影响  
Table 3 Influence of the number of network layers on recognition accuracy

| 层数 | 测试样本数 | 分类正确数 | 识别准确率/% |
|----|-------|-------|---------|
| 3  | 172   | 124   | 72.09   |
| 4  | 172   | 143   | 83.14   |
| 5  | 172   | 162   | 94.19   |
| 6  | 172   | 160   | 93.02   |
| 7  | 172   | 157   | 91.28   |

如表 3 中数据所示, 在当前学习任务规模下, 针对 DBN 的 5 层结构(250-150-90-30-6), 算法的识别准确率已经达到了较高的值, 说明此时方法的特征提取能力已经能够满足该学习任务, 且训练时间适中, 学习效率较高。即说明本文提出的基于多源信息融合的配网故障外部原因识别方法能够在已有故障数据的基础上完成对模型的训练, 进而在之后的故障发生时, 完成对故障原因的识别。

### 3.2 迭代次数对识别准确率的影响

对算法识别准确率影响较为重要的另一个因素是模型训练过程中有监督训练的参数迭代次数, 该次数对网络全局参数微调的效果有巨大影响, 同样会影响模型的特征提取能力。本文在 3.1 节的实验基础上, 选取相同的 5 层网络结构, 测试了参数微调迭代次数从 10 次到 200 次, 模型识别准确率的变化。

如图 10 所示, 随着有监督训练的迭代次数的增加, 模型的识别准确率逐步提升, 当迭代次数达到 80 次左右时, 模型已逐步收敛, 并在 90 次左右达到最大值, 此时的识别准确率为 94.51%。

### 3.3 算法有效性验证

从 3.1 节的实验结果分析中可以看出, 本文提出的方法可以准确有效地识别配网故障外部原因。为了进一步说明本文提出配网故障外部原因识别方法, 验证本文使用的核心算法——深度信念网络, 相较于

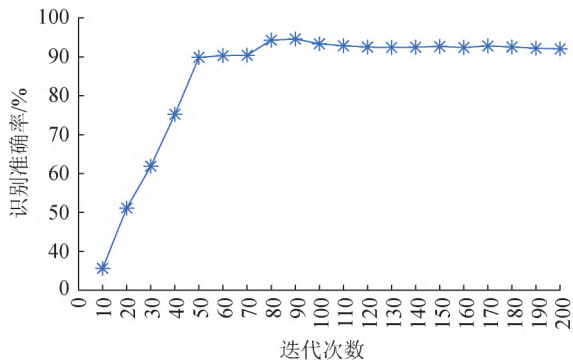


图10 迭代次数对识别准确率的影响  
Fig. 10 The effect of iteration number on recognition accuracy

其他机器学习和神经网络算法在识别准确率上的提升,本文将深度信念网络与逻辑回归算法(logistic regression, LR)、随机森林(random forests, RF)<sup>[21]</sup>算法、BP神经网络(artificial neural networks, ANN)算法<sup>[22-23]</sup>、卷积神经网络(convolutional neural networks, CNN)算法<sup>[24]</sup>进行了比较。在网络层数和迭代次数验证的基础上,设定模型网络结构为(250-150-90-30-6),迭代次数为90次,分别在随机选择25%、50%、75%、100%的训练样本集数量基础上对模型进行训练,在相同的测试集上验证模型识别准确率。结果如图11所示。

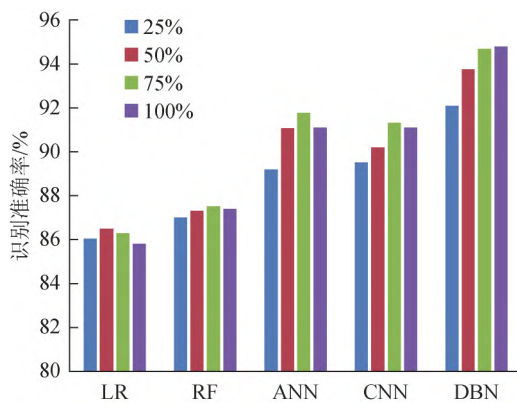


图11 不同算法对识别准确率的影响  
Fig. 11 The effect of different algorithms on the recognition accuracy

从图11的实验结果可以看出,LR算法、RF算法、ANN算法、CNN算法及DBN算法的最大识别准确率分别为86.51%、87.53%、91.81%、91.34%及94.82%。LR和RF算法总体的识别准确率低于基于神经网络的ANN、CNN和DBN算法,这是得益于神经网络在大维度样本数据上具有更好的特征提取能力。随着样本数据量的增大,基于神经网络的ANN、CNN和DBN算法的识别准确率逐渐提高。同时相比ANN和CNN算法,DBN算法在有监督的微调迭代基础上,引入无监督的预训练,加深了网络深

度,优化了BP网络在训练过程中易陷入局部最优的问题,提高了特征提取和泛化能力,避免了在更大数据样本训练过程中的过拟合现象导致识别准确率下降的问题,可以应用于故障情况更加复杂、数据规模更大的实际配电网中。综上所述,本文提出的基于多源信息融合的配网故障外部原因识别方法是准确、可靠且有效的。

## 4 结论

本文对车辆、动物、雷击、施工、树木5种典型配网故障外部原因的特点及相关影响因素进行了分析,总结出4种特征信息,包括天气、时间、线路信息和保护动作信息,将这些信息融合后作为神经网络的输入样本来训练模型,实现了对配网故障外部原因的分类识别。不同故障原因有区域特性,识别模型需结合当地的气象环境等因素进行建模分析。在大量数据样本的情况下,本文使用的深度信念网络相比传统机器学习分类算法,具有更高的识别准确性。以西部某地区配网实际1723组数据进行验证,识别准确率达94.82%,证明了本方法有助于快速识别故障原因,对缩短故障排查时间、提高供电可靠性有积极意义。

## 5 参考文献

- [1] 王玲,邓志,马明,等. 基于状态估计残差比较的配电网故障区段定位方法[J]. 电力系统保护与控制,2021,49(14): 132-139. WANG Ling, DENG Zhi, MA Ming, et al. A method for locating fault sections in distribution networks based on the comparison of state estimation residual errors [J]. Power System Protection and Control, 2021, 49(14): 132-139.
- [2] 白星振,宋昭杉,葛磊蛟,等. 含分布式电源的复杂配电网相间故障定位等效解耦模型[J]. 电力建设,2022,43(2): 45-53. BAI Xingzhen, SONG Zhaoshan, GE Leijiao, et al. An equivalent decoupling model for fault location in complex distribution network with distributed generation [J]. Electric Power Construction, 2022, 43(2): 45-53.
- [3] 刘健,张志华,黄炜,等. 分布式电源接入对配电网故障定位及电压质量的影响分析[J]. 电力建设,2015,36(1): 115-121. LIU Jian, ZHANG Zhihua, HUANG Wei, et al. Influence of distributed generation on fault location and voltage quality of distribution network [J]. Electric Power Construction, 2015, 36(1): 115-121.
- [4] 雷倩,吉兴全,文福拴,等. 利用暂态分量的含分布式电源配电网故障诊断[J]. 电力建设,2016,37(2): 42-49. LEI Qian, JI Xingquan, WEN Fushuan, et al. Fault diagnosis of distribution system with distributed generation employing transient component [J]. Electric Power Construction, 2016, 37(2): 42-49.
- [5] XU L, CHOW M Y. A classification approach for power distribution systems fault cause identification [J]. IEEE Transactions on Power Systems, 2006, 21(1): 53-60.
- [6] MINNAAR U J, NICOLLS F, GAUNT C T. Automating

- transmission-line fault root cause analysis[J]. IEEE Transactions on Power Delivery, 2016, 31(4): 1692-1700.
- [7] BARRERA NÚÑEZ V, KULKARNI S, SANTOSO S, et al. Feature analysis and classification methodology for overhead distribution fault events[C]//IEEE PES General Meeting. July 25-29, 2010, Minneapolis, MN, USA. IEEE, 2010: 1-8.
- [8] 秦雪, 刘亚东, 孙鹏, 等. 基于故障波形时频特征配网故障识别方法研究[J]. 仪器仪表学报, 2017, 38(1): 41-49.  
QIN Xue, LIU Yadong, SUN Peng, et al. Study on the line fault root-cause identification method in distribution networks based on time-frequency characteristics of fault waveforms [J]. Chinese Journal of Scientific Instrument, 2017, 38(1): 41-49.
- [9] KULKARNI S, LEE D, ALLEN A J, et al. Waveform characterization of animal contact, tree contact, and lightning induced faults[C]//IEEE PES General Meeting. July 25-29, 2010, Minneapolis, MN, USA. IEEE, 2010: 1-7.
- [10] XU L, CHOW M Y, TAYLOR L S. Data mining and analysis of tree-caused faults in power distribution systems [C]//2006 IEEE PES Power Systems Conference and Exposition. October 29 - November 1, 2006, Atlanta, GA, USA. IEEE, 2007: 1221-1227.
- [11] CHOW M Y, TAYLOR L S. Analysis and prevention of animal-caused faults in power distribution systems[J]. IEEE Transactions on Power Delivery, 1995, 10(2): 995-1001.
- [12] DOOSTAN M, CHOWDHURY B H. Power distribution system fault cause analysis by using association rule mining [J]. Electric Power Systems Research, 2017, 152: 140-147.
- [13] XU L, CHOW M Y, TIMMIS J. Power Distribution Outage Cause Identification using Fuzzy Artificial Immune Recognition Systems (FAIRS) algorithm [C]//2007 IEEE Power Engineering Society General Meeting. June 24-28, 2007, Tampa, FL, USA. IEEE, 2007: 1-8.
- [14] CHOW M Y, YEE S O, TAYLOR L S. Recognizing animal-caused faults in power distribution systems using artificial neural networks[J]. IEEE Transactions on Power Delivery, 1993, 8(3): 1268-1274.
- [15] LI L N, CHE R F, ZANG H Z. A fault cause identification methodology for transmission lines based on support vector machines [C]//2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). October 25-28, 2016, Xi'an. IEEE, 2016: 1430-1434.
- [16] 贾京龙, 余涛, 吴子杰, 等. 基于卷积神经网络的变压器故障诊断方法[J]. 电测与仪表, 2017, 54(13): 62-67.  
JIA Jinglong, YU Tao, WU Zijie, et al. Fault diagnosis method of transformer based on convolutional neural network [J]. Electrical Measurement & Instrumentation, 2017, 54(13): 62-67.
- [17] 王功明, 乔俊飞, 关丽娜, 等. 深度信念网络研究现状与展望[J]. 自动化学报, 2021, 47(1): 35-49.  
WANG Gongming, QIAO Junfei, GUAN Lina, et al. Review and prospect on deep belief network [J]. Acta Automatica Sinica, 2021, 47(1): 35-49.
- [18] KARTHIGA R, USHA G, RAJU N, et al. Transfer learning based breast cancer classification using one-hot encoding technique[C]//2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). March 25-27, 2021, Coimbatore, India. IEEE, 2021: 115-120.
- [19] 石鑫, 朱永利. 深度学习神经网络在电力变压器故障诊断中的应用[J]. 电力建设, 2015, 36(12): 116-122.  
SHI Xin, ZHU Yongli. Application of deep learning neural network in fault diagnosis of power transformer [J]. Electric Power Construction, 2015, 36(12): 116-122.
- [20] DENIL M, SHAKIBI B, DINH L, et al. Predicting parameters in deep learning [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2. December 5-10, 2013, Lake Tahoe, Nevada. New York: ACM, 2013: 2148-2156.
- [21] 杨杰, 吴浩, 董星星, 等. 基于电流故障分量特征和随机森林的输电线路故障类型识别[J]. 电力系统保护与控制, 2021, 49(13): 53-63.  
YANG Jie, WU Hao, DONG Xingxing, et al. Transmission line fault type identification based on the characteristics of current fault components and random forest [J]. Power System Protection and Control, 2021, 49(13): 53-63.
- [22] 申元, 马仪, 孟见刚, 等. 基于BP神经网络的输电线路故障原因辨识研究[J]. 智能电网, 2017, 5(2): 134-141.  
SHEN Yuan, MA Yi, MENG Jiangang, et al. Research on transmission line fault reason recognition based on BP neural network [J]. Smart Grid, 2017, 5(2): 134-141.
- [23] XU L, CHOW M C, GAO X Z. Comparisons of logistic regression and artificial neural network on power distribution systems fault cause identification [C]//Proceedings of the 2005 IEEE Midnight-Summer Workshop on Soft Computing in Industrial Applications, 2005. SMCia/05. June 28-30, 2005, Espoo, Finland. IEEE, 2005: 128-131.
- [24] 李宇, 杨柳林. 基于卷积神经网络的配电网单相接地故障识别[J]. 电气工程学报, 2020, 15(3): 22-30.  
LI Yu, YANG Liulin. Identification of single-phase-to-earth fault in distribution network based on convolutional neural network [J]. Journal of Electrical Engineering, 2020, 15(3): 22-30.

收稿日期: 2022-04-18

作者简介:

谭媛(1994)女, 硕士, 工程师, 主要研究方向为配网故障原因识别, E-mail: 405402585@qq.com;

张文海(1989)男, 博士, 助理研究员, 高级工程师, 通信作者, 主要研究方向为配网故障诊断与电力扰动分析学, E-mail: zhangwh2079@scu.edu.cn;

王杨(1991)男, 博士, 研究员, 四川省人才计划入选专家, 主要研究方向围绕新型电力系统“高比例新能源、高比例电力电子”接入这一特征, 开展新型电力系统电能质量分析与控制、宽频振荡广域监测、溯源与抑制、非线性控制理论在新型电力系统中的应用等, E-mail: fwang@scu.edu.cn.

(编辑 白婉欣)