

基于多源信息融合的输电线路故障诊断研究

葛亚明¹,周陈斌²,孟屹华²,沈蛟骁²,曹海欧¹,任旭超¹

(1. 国网江苏省电力有限公司,江苏 南京 210000;

2. 国网江苏省电力有限公司苏州供电分公司,江苏 苏州 215000)

摘要:随着新型电力系统的加速建设,输电系统的规模和复杂性不断增加,而以多源数据作为驱动源、满足准确率与低耗时要求的输电线路故障诊断算法亟待研究。提出一种基于改进NRBO-XGBoost算法的多源信息融合输电线路故障诊断方案。首先,通过对线路两侧保护测量电气量和动作开关量的分析,解耦出区内外故障场景下时/频域差动电流和差动电压、暂态极性和动作信号的关联性特征;其次,将解耦出的多源故障特征向量输入至XGBoost串行学习算法,并同时引入NRBO算法对XGBoost的训练参数进行全局优化;最后,基于改进NRBO-XGBoost算法的辨识输出结果,获取完整的输电线路区内外故障诊断模型。在PSCAD/EMTDC中搭建了IEEE-30标准节点输电系统模型,通过对4种典型场景下的案例测试,结果表明所提出的多源信息融合算法能够满足线路故障诊断准确率99%的要求,在诊断速度上相较于传统智能算法也体现出了一定的优越性。

关键词:输电线路;故障诊断;XGBoost算法;NRBO算法;多源信息融合

中图分类号:TM28 **文献标识码:**A **DOI:**10.19457/j.1001-2095.dqed25972

Research on Fault Diagnosis of Transmission Line Based on Multi-source Information Fusion

GE Yaming¹, ZHOU Chenbin², MENG Yihua², SHEN Jiaoxiao², CAO Haiou¹, REN Xuchao¹

(1. State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210000, Jiangsu, China; 2. Suzhou Power Supply Company of State Grid Jiangsu Electric Power Co., Ltd., Suzhou 215000, Jiangsu, China)

Abstract: With the accelerated construction of new power systems, the scale and complexity of transmission systems are constantly increasing. Therefore, it is urgent to study transmission line fault diagnosis algorithms that utilize multi-source data as driving sources and meet requirements for accuracy and low time consumption. A multi-source information fusion transmission line fault diagnosis method based on the improved NRBO-XGBoost algorithm was proposed. Firstly, by analyzing the measured electrical quantities and action switch quantities on both sides of the line protection, the correlation features of time/frequency domain differential current and differential voltage, transient polarity, and action signals under internal and external fault scenarios were decoupled. Then, the decoupled multi-source fault feature vectors were input into the XGBoost serial learning algorithm, and the NRBO algorithm was introduced to globally optimize the training parameters of XGBoost. Finally, based on the identification output of the improved NRBO-XGBoost algorithm, a complete transmission line fault diagnosis model for internal and external faults was obtained. An IEEE-30 standard node transmission system model was constructed using PSCAD/EMTDC. Through testing in four typical scenarios, the results demonstrated that the proposed multi-source information fusion algorithm achieves a line fault diagnosis accuracy of 99%, meeting the required threshold. Additionally, it exhibits certain advantages in terms of diagnosis speed compared to traditional intelligent algorithms.

Key words: transmission line; fault diagnosis; XGBoost algorithm; NRBO algorithm; multi-source information fusion

近年来,随着“新型电力系统”建设目标的深入推进,大规模光伏、海上风电等新能源经过长

基金项目:国家电网有限公司科技项目(J2023145)

作者简介:葛亚明(1984—),男,本科,高级工程师,主要研究方向为电网调度运行技术、电力系统继电保护技术,Email:93126956@qq.com

距离输电线路并入高压电网^[1]。此背景下,确保对于输电线路故障的快速识别、建立准确高效的线路区内外故障诊断模型^[2],是维持新型电力系统稳定运行的关键问题之一。

针对复杂场景下的线路故障诊断问题,借助人工智能算法,通过融合历史故障数据、训练生成故障诊断模型,是一种新的应对思路^[3-5]。发挥智能算法的自主学习能力,可以有效提升“非透明”故障工况下的诊断结果可靠性,降低对于某一类特定故障特征的“强依赖性”。目前,电力系统中广泛采用的人工智能学习算法包括:支持向量机(support vector machine, SVM)、卷积神经网络(convolutional neural network, CNN)、长短期记忆递归神经网络(long short term memory, LSTM)、随机森林(random forests, RF)等,应用场景覆盖变压器和电机本体的异常状态辨识^[6-7]、负荷预测^[8]和电力经济调度优化^[9]等。例如,文献[6]建立了基于油中溶解气体分析的SVM变压器故障诊断模型;为提高永磁直驱单片机的故障诊断准确率,文献[7]则引入SVM对电流时域特征进行分段提取;文献[8]考虑多种随机性因素的综合影响,联合CNN-LSTM构建了电动汽车充电负荷的预测模型,并同时引入RF进行误差修正;文献[9]采用SVM与最小二乘拟合相结合的思路,构建了电力现货市场价格的预测模型。

然而,传统智能算法的训练过程大多依赖循环迭代,训练参数繁多复杂,应用于线路故障诊断场景时,处理非线性故障暂态的性能欠缺,操作速度也有待提升。近年来,计算机领域热门的极限梯度提升算法(extreme gradient boosting, XGBoost)采用决策树并行计算的思想,并同时引入正则项惩罚机制调节过于复杂的树模型,具有对大规模数据集处理速度快、占用内存低和非线性辨识度高的优势^[10-11],更加符合线路故障诊断的实际需求。但XGBoost的性能也受自身参数如最大迭代次数、种群数等影响较大,期望辅助优化算法进行参数自适应。常用的优化算法如粒子群优化(particle swarm optimization, PSO)等,却容易陷入局部最优,同时缺乏合理增大种群范围的手段,获取全局最优解的能力也不足。

另一方面,当前针对线路故障诊断算法的研究,在进行样本训练时,多数仅基于单一类型的数据集,并未充分考虑多源异类信息的融合,诊断模型的鲁棒性可以进一步提升。此外,少数基

于多源信息融合的故障诊断算法,过度依赖算法自身对原始数据集的辨识和不良数据剔除能力,缺乏考虑线路故障特征的数据集解耦重构预处理过程,导致生成的诊断模型欠缺理论性支撑。

为克服传统智能算法应用于线路故障诊断时的缺陷和不足,本文首先从时域电气量、频域电气量和保护动作开关量多维度出发,对线路区内外故障情形下的特征差异进行解耦分析,建立理论支撑;其次,在XGBoost算法中引入牛顿-拉夫逊优化(Newton-Raphson-based optimizer, NRBO),借助NRBO的陷阱规避和种群替换搜索操作^[12],提升不同线路故障场景下的参数自适应能力;最后,通过NRBO-XGBoost改进算法融合多源故障特征,训练生成高鲁棒性的线路故障诊断模型。基于IEEE-30标准节点的辨识测试结果,验证了所提故障诊断算法和模型的可靠性。

1 输电线路多源故障特征分析

1.1 时域故障特征

在时域中,以II型等效输电线路为例^[13],对线路区内外故障情形的时域特征进行分析。

首先,考虑图1所示的区外故障情形。图中: $u_m(t)$, $u_n(t)$ 和 $i_m(t)$, $i_n(t)$ 分别为线路 m, n 侧保护测量的时域电压和电流; $i_{m1}(t)$, $i_{n1}(t)$ 分别为线路 m, n 侧流过阻感参数的分支电流; R_f 为故障点的过渡电阻; R_{mn} , L_{mn} , C 分别为线路的电阻、电感和对地电容; Z_{mn} 为输电线路的总阻抗。

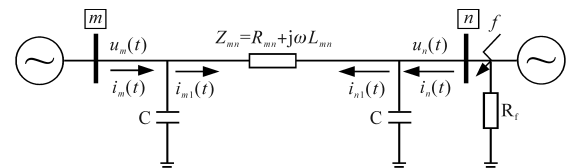


图1 线路区外故障时域等效电路

Fig.1 Time-domain equivalent circuit for line external faults

定义母线指向线路为正方向,依据图1可列写时域方程如下:

$$\begin{cases} i_{m1}(t) = i_m(t) - C \frac{du_m(t)}{dt} \\ i_{n1}(t) = i_n(t) - C \frac{du_n(t)}{dt} \end{cases} \quad (1)$$

线路区外故障时, $i_{m1}(t)$ 和 $i_{n1}(t)$ 幅值相等、方向相反,满足 $i_{m1}(t) + i_{n1}(t) = 0$ 。进一步,定义线路两侧的时域差动电流和差动电压为

$$\begin{cases} i_{cd}(t) = i_m(t) + i_n(t) \\ u_{cd}(t) = u_m(t) + u_n(t) \end{cases} \quad (2)$$

结合式(1)和式(2)可得:

$$i_{cd}(t) = C \frac{du_{cd}(t)}{dt} \quad (3)$$

式(3)描述的电气量关系符合电容电气特性。因此,输电线路区外故障的时域差动电压和差动电流故障特征可以用电容模型来表示。

其次,考虑图2所示的区内故障情形。

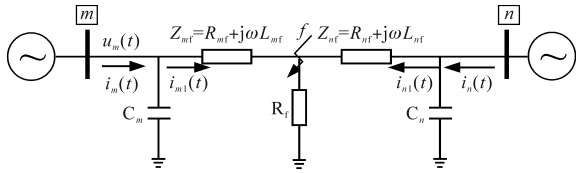


图2 线路区内故障时域等效电路

Fig.2 Time-domain equivalent circuit for line internal faults

图2中: Z_{mf}, Z_{nf} 分别为故障点与 $m(n)$ 侧间的线路阻抗; $i_f(t)$ 为故障点电流,满足 $i_f(t) = i_{m1}(t) + i_{n1}(t)$,其他参数含义参照图1。可列写时域方程如下:

$$\begin{cases} u_m(t) = R_{mf}i_{m1}(t) + L_m \frac{di_{m1}(t)}{dt} + i_f(t)R_f \\ u_n(t) = R_{nf}i_{n1}(t) + L_n \frac{di_{n1}(t)}{dt} + i_f(t)R_f \end{cases} \quad (4)$$

忽略较小的容抗电流^[13],近似认为保护测量电流与流过阻感参数电流相等,即 $i_m(t) \approx i_{m1}(t)$, $i_n(t) \approx i_{n1}(t)$ 。联立式(2)~式(4)可得:

$$\begin{cases} u_{cd}(t) = R_{eq}i_{cd}(t) + L_{eq} \frac{di_{cd}(t)}{dt} \\ R_{eq} = Z_{nf}R_{mf}/Z_{mn} + Z_{mf}R_{nf}/Z_{mn} + 2R_f \\ L_{eq} = Z_{nf}L_{mf}/Z_{mn} + Z_{mf}L_{nf}/Z_{mn} \end{cases} \quad (5)$$

式中: R_{eq}, L_{eq} 分别为时域等效电阻和等效电抗系数,故障发生后为常数。

由式(5)可知,当输电线路发生区内故障时,时域差动电压和差动电流具备阻感故障特征,与区外故障对应的频域电容特征存在差异。

1.2 频域故障特征

在频域中,可以采用精确的分布参数线路模型,从本质上消除电容电流的影响,而不会对计算量和采样频率有过高要求^[14]。

区外故障时的频域故障附加状态如图3所示。图中: $\Delta u_m(s), \Delta u_n(s)$ 和 $\Delta i_m(s), \Delta i_n(s)$ 分别为线路 m, n 侧的频域故障分量电压和电流; $\Delta U_f(s)$ 为故障点附加频域电压; $Z_m(s), Z_n(s)$ 分别为两侧系统的频域阻抗;单位线路分布阻抗和导纳分别满足 $z_l(s) = r + sl$ 和 $y(s) = g + sc$ 。

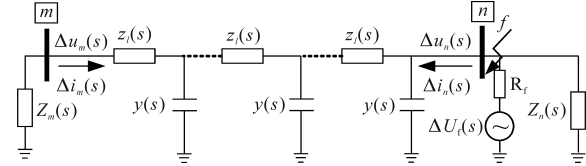


图3 线路区外故障频域附加等效电路

Fig.3 Superimposed equivalent circuit for line external faults in the frequency domain

依据图3,可列写区外故障频域方程如下:

$$\begin{cases} \Delta u_m(s) = \Delta u_n(s)ch\gamma(s)D - \Delta i_n(s)Z_c(s)sh\gamma(s)D \\ \Delta i_m(s) = \frac{\Delta u_n(s)}{Z_c(s)}sh\gamma(s)D - \Delta i_n(s)ch\gamma(s)D \\ Z_c(s) = \sqrt{r + sl/sc} \\ \gamma(s) = \sqrt{(r + sl)sc} \end{cases} \quad (6)$$

式中: $Z_c(s)$ 为线路阻抗传播系数; D 为线路长度; $\gamma(s)$ 为线路分布参数传播系数。

效仿时域式(2),定义频域差动电流和电压如下:

$$\begin{cases} \Delta i_{cd}(s) = \Delta i_m(s) + \Delta i_n(s) \\ \Delta u_{cd}(s) = \Delta u_m(s) + \Delta u_n(s) \end{cases} \quad (7)$$

结合式(6)和式(7),区外故障时的线路两侧频域差动电压和差动电流满足:

$$\frac{\Delta u_{cd}(s)}{\Delta i_{cd}(s)} = Z_c(s) \frac{ch\gamma(s)D + 1}{sh\gamma(s)D} \quad (8)$$

基于某输电线路的实际参数,文献[11]详细研究了式(8)右端项 $Z_c(s)[ch\gamma(s)D + 1]/[sh\gamma(s)D]$ 的幅频特性,指出其跟随角频率 $\omega = 2\pi f$ 的幅频变化规律与分布电容的频域特性一致。因此,当线路发生区外故障时,频域差动电压和差动电流故障特征也可以用电容模型来描述。

另一方面,线路区内故障时的频域故障附加状态如图4所示。

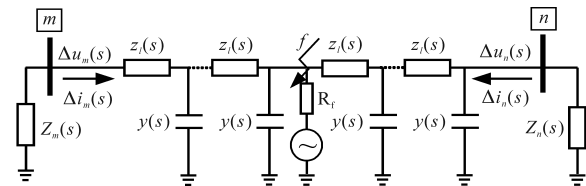


图4 线路区内故障频域附加等效电路

Fig.4 Superimposed equivalent circuit for line internal faults in the frequency domain

设故障点与 m 侧的距离为 d ,则线路两侧的频域电压和电流满足:

$$\begin{cases} \Delta u_m(s) = -Z_m(s)\Delta i_m(s) \\ \Delta u_n(s) = -Z_n(s)\Delta i_n(s) \end{cases} \quad (9)$$

结合式(8),式(9)可以改写为

$$\begin{cases} \frac{\Delta u_{cd}(s)}{\Delta i_{cd}(s)} = -\frac{Z_n(s) + kZ_m(s)}{1+k} \\ k = \frac{\Delta i_m(s)}{\Delta i_n(s)} \approx \frac{Z_n(s) + z_l(s)(D-d)}{Z_m(s) + z_l(s)d} \end{cases} \quad (10)$$

式中: k 为两侧频域电流的偏移参数。

考虑到阻抗角的一致性, k 可近似视为实数,且受频率变化影响较小。因此,式(10)可以简化为下式:

$$\frac{\Delta u_{cd}(s)}{\Delta i_{cd}(s)} = -\frac{Z_n(s) + kZ_m(s)}{1+k} = -sL'_{eq} - R'_{eq} \quad (11)$$

式中: R'_{eq}, L'_{eq} 分别为频域等效电阻和电抗系数,故障发生后为常数。

由式(11)可知,当线路发生区内故障时,频域差动电压和差动电流符合阻感故障特征,与区外故障对应的电容故障特征存在差异。

1.3 暂态极性故障特征

只要故障前相间电压不为零,则故障点电压就会在短路瞬间突变,产生瞬态不稳定电弧^[15]。根据电磁波理论,不稳定电弧产生的暂态信号将从故障点向外端传输。当输电线路发生区内外故障时,两侧保护处的暂态电压和电流极性方向分别如图5a和图5b所示。

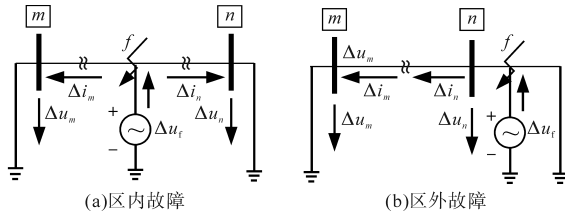


图5 区内外故障时线路两侧暂态电压和电流极性方向

Fig.5 Polarity of transient voltage and current on both sides during internal and external faults

由图5可见,对于线路区内故障,两侧保护处的暂态电压和电流极性方向相同(均为正极性);而对于区外故障,两侧保护处的暂态电压和暂态电流极性方向相反(一侧为正极性,另一侧为负极性)。

以能量集中度较好的Db6小波为母小波^[15],综合考虑分解效率和暂态特征识别率的平衡,采用4层小波变换提取暂态极性,由Mallat算法将信号分解到不同频带后,利用频带的尺度系数和小波系数进行重构,得到特定频带的信号。

假设故障电压、电流在 j 尺度上的重构系数分别为 $C_{j,k}^u$ 和 $C_{j,k}^i$,极性计算公式为^[16]

$$\begin{cases} \theta = \arccos \frac{C_{j,k}^u C_{j,k}^i}{|C_{j,k}^u| |C_{j,k}^i|} \\ 90^\circ < \theta < 270^\circ \Rightarrow \text{正极性} \\ -90^\circ < \theta < 90^\circ \Rightarrow \text{负极性} \end{cases} \quad (12)$$

通过交换线路两侧保护的极性判断结果,可以区分区内、外故障。当两侧保护均为正极性,则对应区内故障;反之,若一侧为正极性,另一侧为负极性时,则对应区外故障。

1.4 保护动作开关量故障特征

虽然利用上述时/频域和极性故障特征能够识别区内、外故障,但无法细化内部故障区域。为实现这一目的,可以结合保护开关量故障特征。

以距离保护I段为例,当线路内部故障时,线路两侧距离I段的动作情况与故障位置相关。如图6所示,以距离保护为例,一般情况下,距离I段保护范围整定为线路全长的80%^[17],则可将线路分为3个保护区域:① S_1 为线路近 m 侧保护区段;② S_2 为线路 m, n 两侧保护的公共保护区段;③ S_3 为线路近 n 侧保护区段。 $S_1 \sim S_3$ 分别占线路总阻抗的20%,60%,20%。

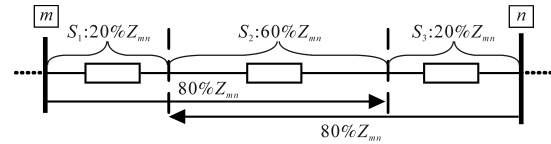


图6 线路两侧距离I段保护区

Fig.6 Protection range of the distance relay zone-I at two ends of the line

对于 m 侧距离I段,其保护范围为: S_1+S_2 ;而 n 侧距离I段的保护范围为: S_2+S_3 。综合考虑线路两端I段保护范围,依据两侧保护动作信号,进一步细化故障的所在区域,即:①仅收到 m 侧距离I段动作信号,则为区段 S_1 内部故障;②仅收到 n 侧距离I段动作信号,则为区段 S_3 内部故障;③同时收到两侧I段动作信号,则为区段 S_2 内部故障;④均未收到两侧I段动作信号,则为区外故障。

2 基于改进NRBO-XGBoost算法的线路故障诊断模型

2.1 多源信息融合线路故障诊断思路

第1节分析表明:线路两侧保护测量的电气量时域、频域和暂态极性以及保护动作开关量特征中均蕴含了线路的故障状态信息。通过收集电网运检平台存储的电气量/开关量历史故障数据,建立故障数据库并进行多源融合、训练学习,

生成用于线路故障诊断的模型。当线路发生新的故障后,将保护测量的故障特征向量输入至预先训练好的诊断模型,便可获取线路的实际区内、外故障状态。而为了确保故障诊断模型的精准度和高效率,多源数据融合算法的选择至关重要。

2.2 XGBoost 串行集成学习算法

XGBoost 算法是基于梯度提升决策树 (gradient boosting decision trees, GBDT) 的新型串行集成机器学习算法^[12]。相比于传统 GBDT 算法, XGBoost 在训练时充分利用了代价函数的一阶和二阶导数信息,同时在每一轮迭代训练时不使用全部数据集,而是采取选择性采样策略、自动填补缺失值,提高了算法的运行速度和工程实用性。XGBoost 的算法原理如图 7 所示,包括以下步骤:

①步骤 1:输入训练样本集并定义决策树的目标函数如下:

$$\begin{cases} L(\Phi) = \sum_j \sum_i l(y_i, \hat{y}_i^{(j)}) + \sum_j \Omega(f_j) \\ \Omega(f_j) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \end{cases} \quad (13)$$

式中: $l(y_i, \hat{y}_i^{(j)})$ 为损失函数; $y_i, \hat{y}_i^{(j)}$ 分别为第 i 个样本的实际值和前 j 个决策树对样本 i 的预测值; $\Omega(f_j)$ 为正则项,代表第 j 棵树的模型复杂度; T, ω 分别为当前树的深度和叶子节点的权重,用于惩罚过于复杂的树模型,避免过拟合; γ, λ 为控制正则化强度的参数,当二者均为 0 时,则变为普通的梯度提升树的目标函数。

②步骤 2:迭代求解第 j 棵树的输出,进而加和得到目标函数值:

$$\begin{cases} L^{(j)} = \sum_i l(y_i, \hat{y}_i^{(j)}) + \Omega(f_j) \\ \hat{y}_i^{(j)} = \hat{y}_i^{(j-1)} + \eta f_j(x_i) \end{cases} \quad (14)$$

式中: η 为收缩率,默认值设为 1。

③步骤 3:对损失函数进行二阶泰勒展开,引入残差项 R 并迭代学习到第 j 棵树,如下式所示:

$$\begin{cases} L^{(j)} = \sum_i [g_i f_j(x_i) + \frac{1}{2} h_i f_j^2(x_i)] + R + \Omega(f_j) \\ g_i = \frac{\partial l(y_i, \hat{y}_i^{(j-1)})}{\partial \hat{y}_i^{(j-1)}} \\ h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(j-1)})}{\partial (\hat{y}_i^{(j-1)})^2} \\ R = \gamma T + \frac{1}{2} \lambda \sum_j \omega_j^2 \end{cases} \quad (15)$$

④步骤 4:分裂树节点,根据设定条件进行迭代判断。记决策树 q 在叶子节点 j 上的取值为

$$I_j = \{i | q(x_i) = j\} \quad (16)$$

利用下式对树的节点进行拆分,其中下标“L”与“R”分别代表值划入左、右子树的目标函数最优值。此外,当损失函数减小的值大于 γ 时,需要对决策树进行拆分。

$$\begin{cases} L_{\text{split}} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \\ H_j = \sum_{i \in I_j} h_i \\ G_j = \sum_{i \in I_j} g_i \end{cases} \quad (17)$$

⑤步骤 5:停止节点分裂并生成决策树。

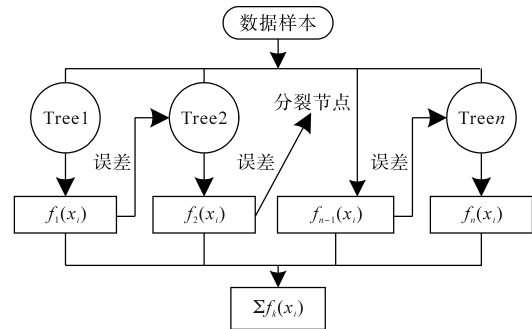


图7 串行集成学习算法

Fig.7 Serial integrated learning algorithm

2.3 基于NRBO的模型参数优化改进

考虑本文的线路故障诊断应用场景,进一步增强 XGBoost 在线路故障诊断中的适应性,引入 NRBO 算法从迭代次数、树的最大深度以及学习率 3 个关键维度,对 XGBoost 参数改进优化。

①步骤 1:确定适应度函数并初始化种群。首先,定义适应度函数如下:

$$\begin{aligned} \min f(x_1, x_2, \dots, x_n) \\ l_b < x_j < u_b \quad j = 1, 2, \dots, dim \end{aligned} \quad (18)$$

式中: $f(\cdot)$ 为适应度函数; $\mathbf{x} = (x_1, x_2, \dots, x_j, \dots, x_n)$ 为决策向量; dim 为问题的维数; u_b, l_b 分别为取值的上、下界。

本文中,选择 XGBoost 算法对故障特征测试集的辨识错误率作为适应度。

其次,在候选解的集合生成具有 N 个总体的初始随机种群:

$$x_j^n = l_b + rand(0, 1) \times (u_b - l_b) \quad (19)$$

式中: x_j^n 为第 n 个总体 ($n=1, 2, \dots, N$) 的第 j 个维度位置; $rand(0, 1)$ 为 0 到 1 之间的随机数。

基于式(19),得出种群矩阵如下:

$$X_n = (x_j^n)_{N \times dim} \quad (20)$$

②步骤2:依据牛顿-拉夫逊规则(NRSR)更新新矢量位置。基于牛顿-拉夫逊方法利用NRSR获取更新后矢量,加快收敛速度,操作如下:

$$\begin{cases} NRSR = randn \times \frac{(y_w - y_b)\Delta x}{2(y_w + y_b - 2x_n)} \\ y_w = rand(1,0) \times [mean(Z_{n+1} + x_n) + rand(1,0)\Delta x] \\ y_b = rand(1,0) \times [mean(Z_{n+1} + x_n) - rand(1,0)\Delta x] \\ Z_{n+1} = x_n - randn \times \frac{(X_w - X_b)\Delta x}{2(X_w + X_b - 2x_n)} \end{cases} \quad (21)$$

式中: $randn$ 为服从均值为0、方差为1的标准正态分布随机数; $mean$ 为取平均数函数; Δx 为自变量变化量; X_b, X_w 分别为当前所有迭代下的最佳位置与最差位置; Z_{n+1} 为人为引入的中间矢量位置; y_w, y_b 为由 Z_{n+1} 与 x_n 随机生成的两个新矢量位置。

结合式(20)和式(21),得到更新后矢量位置 x_n^{IT+1} 如下:

$$\begin{cases} X1_n^{IT} = x_n^{IT} - NRSR + a(X_b - X_n^{IT}) \\ X2_n^{IT} = X_b - NRSR + a(X_b - X_n^{IT}) \\ X3_n^{IT} = X_n^{IT} - \delta(X2_n^{IT} - X1_n^{IT}) \\ \delta = [1 - (\frac{2IT}{MAX_IT})]^5 \end{cases} \quad (22)$$

$$x_n^{IT+1} = r[rX1_n^{IT} + (1-r)X2_n^{IT}] + (1-r)X3_n^{IT} \quad (23)$$

式中: IT, MAX_IT 分别为当前迭代次数与最大迭代次数; r 为0到1之间的一个随机数。

③步骤3:陷阱规避操作(TAO)。对最佳位置 X_b 与位置 x_n^{IT+1} 进行组合,改变 x_n^{IT+1} 得到更好的矢量位置 X_{TAO}^{IT} ,避免参数陷入局部最优。若 $rand$ 值小于 DF ,使用以下公式更新 X_{TAO}^{IT} ,即

$$\begin{cases} X_{TAO}^{IT} = x_n^{IT+1} + \theta_1(\mu_1 X_b - \mu_2 X_n^{IT}) + \theta_2 \delta(\mu_1 \times mean(X_n^{IT}) - \mu_2 X_n^{IT}) & \text{if } \mu_1 < 0.5 \\ X_{TAO}^{IT} = X_b + \theta_1(\mu_1 X_b - \mu_2 X_n^{IT}) + \theta_2 \delta(\mu_1 \times mean(X_n^{IT}) - \mu_2 X_n^{IT}) & \text{otherwise} \\ X_n^{IT+1} = X_{TAO}^{IT} \end{cases} \quad (24)$$

式中: $\theta_1, \theta_2, \mu_1, \mu_2$ 为随机数。

θ_1 和 θ_2 分别限制在 $[-1, 1]$ 和 $[-0.5, 0.5]$, μ_1 和 μ_2 的生成原则为^[12]

$$\begin{cases} \mu_1 = 3\beta \times rand(0,1) + (1-\beta) \\ \mu_2 = \beta \times rand(0,1) + (1-\beta) \\ \beta = \begin{cases} 0 & \mu_1 \geq 0.5 \\ 1 & \text{其他} \end{cases} \end{cases} \quad (25)$$

式中: β 为一个随机的二进制数。

2.4 线路故障诊断流程

基于改进NRBO-XGBoost算法的线路故障诊断流程如图8所示。

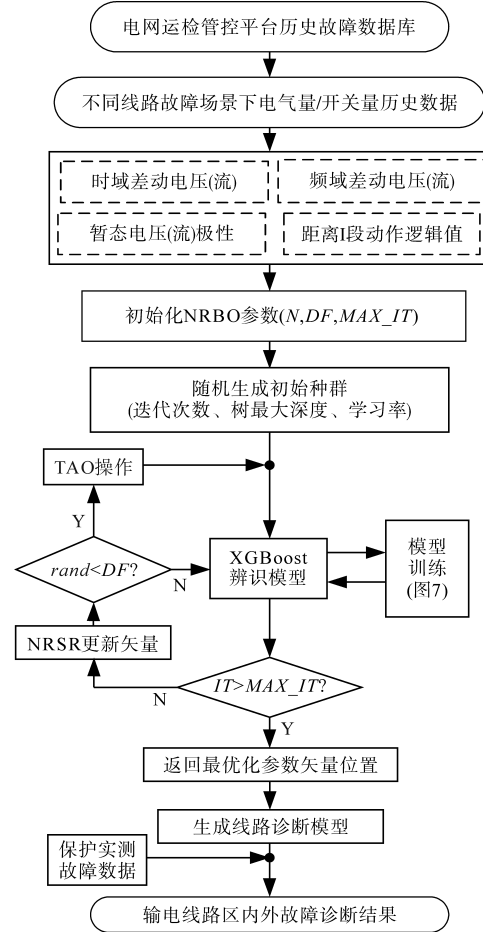


图8 基于改进NRBO-XGBoost算法的多源信息融合线路故障诊断流程

Fig.8 Fault diagnosis process based on NRBO-XGBoost algorithm and multi-source information fusion

线路故障诊断主要步骤如下:

1)从电网运检管控平台历史故障数据库中读取相关电气量和开关量信息,并进行预处理;

2)采用小波变换、时域滤波采样、频域故障分量提取等信号处理方法,提取区内外故障场景下时域差动电压(流)、频域差动电压(流)、线路两侧暂态电压与电流的极性、两侧距离保护I段的动作逻辑值等故障特征量;

3)将多源故障特征作为NRBO-XGBoost算法的输入,形成训练样本,并且通过迭代学习生成线路故障诊断模型;

4)在线路发生新的故障后,将保护实测的相关电气量和开关量故障数据输入诊断模型,得到线路实际故障状态的诊断结果。

3 案例分析

3.1 诊断测试场景

为了验证所提出线路诊断算法的可行性,在PSCAD/ETMPC中搭建如图9所示的IEEE-30标准节点测试输电系统,以线路Bus15—Bus24作为故障诊断对象,设置如下4种代表性故障工况:

- 1) 线路Bus15—Bus24首端 f_1 处区内故障;
- 2) 线路Bus15—Bus24中段 f_2 处区内故障;
- 3) 线路Bus15—Bus24末端 f_3 处区内故障;
- 4) 线路Bus15—Bus24区外 $f_5\sim f_8$ 处故障。

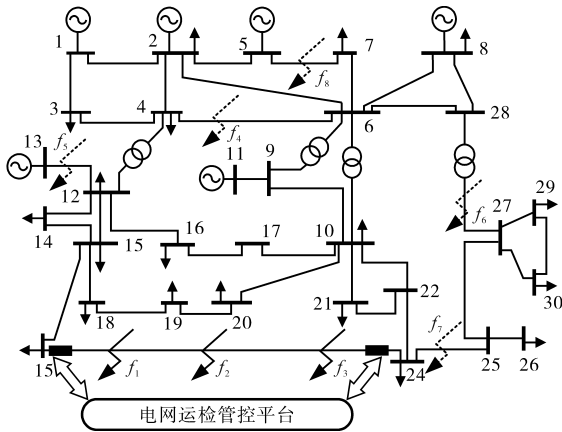


图9 IEEE-30标准节点测试输电系统

Fig.9 IEEE-30 standard node test transmission system

为了贴近实际工况,各场景下模拟不同故障位置/类型、过渡电阻以及3类典型故障性质,包括^[18]:雷击故障(故障相和零序电流为正弦波,直流含量高)、鸟害故障(故障相和零序电流为正弦波,直流分量和谐波含量低)、外破故障(故障相/零序电流波形畸变,衰减直流分量和谐波含量高)。Bus15和Bus24处的保护采样频率设置为2.4 kHz,线路两侧保护的采样故障信息传送到电网运检平台构建原始数据集,生成诊断模型并进行故障状态辨识测试。本文采用Matlab 2020a软件模拟电网运检平台的诊断模型生成过程,基础硬件环境为3.40-GHz-i5 CPU/8 GB RAM。

3.2 XGBoost-NRBO算法故障诊断结果

考虑验证集和实验集之间的独立性、随机性和多样性要求,将原始数据集不同故障状态下,线路Bus15—Bus24两侧保护提取的时/频域差动电压(流)、暂态极性和开关量故障特征集中以简单随机抽样选出70%样本为训练集,30%样本作为实验测试集,如表1所示。XGBoost-NRBO算法的初始化参数设置如下:最大迭代次数

$MAX_IT=10$,种群数 $N=6$, $DF=0.5$ 。

表1 训练集与测试集样本统计

Tab.1 Sample statistics for training and testing sets

线路故障状态	故障标签	训练集数目	测试集数目
线路首端内部故障	1	1 214	520
线路中段内部故障	2	1 214	520
线路末端内部故障	3	1 214	520
线路区外故障	4	643	275

将故障特征样本输入后,XGBoost-NRBO算法诊断结果的混淆矩阵如图10所示,NRBO的适应度曲线如图11所示。结合图10与图11可知,对于训练集,算法可以保证完全正确辨识;而在测试集的2 112个样本中仅有类别3的一个样本被错误地辨识为类别4;此外,NRBO算法在第7次迭代时适应度(即错误率)收敛接近于0,能够指向性地优化所生成线路诊断模型的结构与自学习策略。

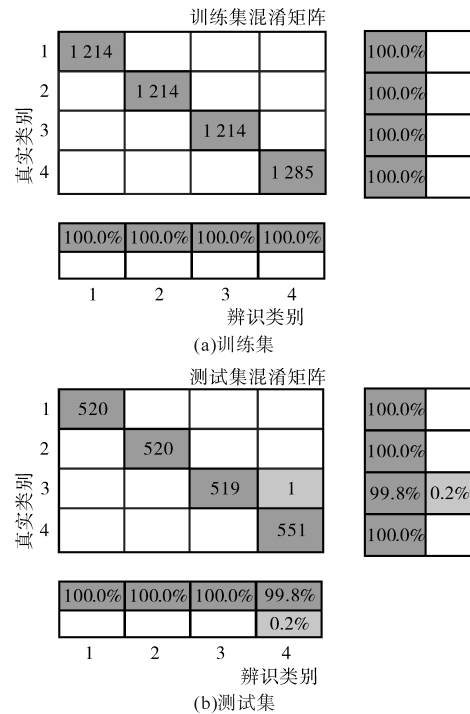


图10 XGBoost-NRBO算法辨识结果的混淆矩阵

Fig.10 Confusion matrix of identification results of the XGBoost-NRBO algorithm

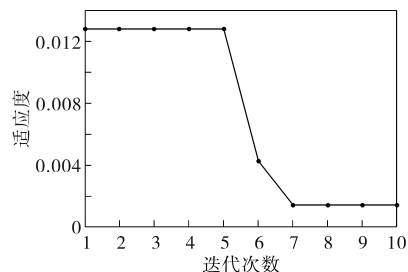


图11 NRBO适应度曲线

Fig.11 The fitness curve of the NRBO algorithm

3.3 与其他故障诊断算法的性能对比

为了进一步验证所提出的故障诊断算法相较于其他传统算法的优越性,设置对比实验。在相同故障特征原始数据集的条件下,将提出的XGBoost-NRBO故障诊断算法分别与SVM, PSO-SVM, NRBO-SVM和 PSO-XGBoost, 以及不使用NRBO参数优化的XGBoost进行对比。考虑相同的线路区内、外故障场景,比较指标包括准确率和诊断耗时。其中,对于使用参数优化的算法,优化迭代次数统一设置为10次,种群数目初始化为6;对于没有使用参数优化的算法,其基本参数使用随机数方式统一生成。每组对比实验重复5次,平均结果如表2所示。此外,图12~图14给出某次诊断过程的详细对比结果。

表2 各类故障辨识算法性能比较

Tab.2 Performance comparison of various fault identification algorithms

算法名称	算法耗时/s	诊断准确率/%
NRBO-XGBoost	21.647 1	99.95
PSO-XGBoost	24.029 0	98.23
XGBoost	1.874 1	94.07
NRBO-SVM	23.845 6	97.76
PSO-SVM	35.768 0	97.51
SVM	2.942 2	91.83

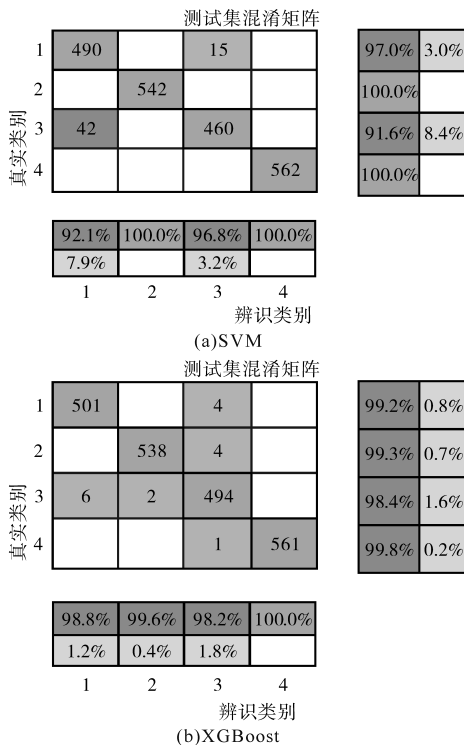


图12 SVM与XGBoost辨识算法测试集混淆矩阵

Fig.12 Confusion matrix of SVM and XGBoost identification algorithm test set

从图12中可以看到:当不使用参数优化算法时,SVM测试集的总诊断准确率为97.3%(图12a混淆矩阵对角线样本和/矩阵总样本),算法耗时为2.887 4 s,而XGBoost测试集的总诊断准确率为99.19%(图12b),算法耗时为1.842 3 s。另一方面,当分别使用PSO算法和NRBO算法优化诊断模型参数时,图13给出了XGBoost的故障诊断精度差异对比。由图13可见,PSO-XGBoost经过10次优化迭代后,其准确率依然难以收敛;而NRBO-XGBoost在第4次迭代后,其准确率就快速收敛,同时稳态值也较高。

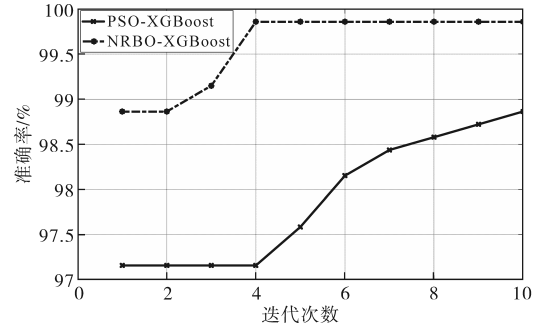


图13 NRBO与PSO参数优化效果比较

Fig.13 Comparison of parameter optimization effects between NRBO and PSO

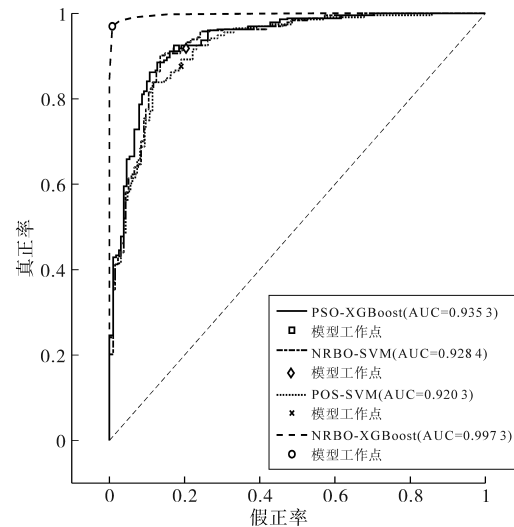


图14 受试者工作曲线

Fig.14 ROC curves

此外,为验证NRBO对XGBoost算法诊断输出准确性的提升,图14还对比了不同参数优化算法下的受试者工作曲线(receiver operating characteristic curve, ROC)。在图14中,横坐标为假正率,表示错误诊断为“正”类的区外故障,而纵坐标为真正率,表示正确诊断为“负”类的区内故障。ROC曲线下方面积(AUC)越大,对应输出诊断分类准确性越高。由图可见, NRBO-XG-

Boost算法的最大AUC为0.997 3,表明生成的故障诊断模型在区分“正”、“负”类线路故障样本时具有极高的准确性和可靠性。

综上,对比各类算法的诊断结果可知:

1)在大容量训练集条件下,XGBoost算法是一种基于决策树的集成学习算法,具有更强的非线性适应能力,可通过并行计算来有效地处理大量数据和特征。而传统的SVM计算复杂度随数据量增加而增加,因此XGBoost的故障诊断平均精度更高、速度更快;

2)相较于传统SVM算法,XGBoost算法对于参数变化的抗干扰能力更强,对数据预处理要求降低,能够处理存在部分缺失值的原始数据。相较而言,SVM对原始数据的缩放和归一化则要求更为苛刻,并且对原始训练集的质量和选择较为敏感。因此,在采用相同的参数优化方式后,NRBO-/PSO-XGBoost算法在更短的运算时间内能达到更高的辨识准确率;

3)由于NRBO相较PSO引入了TAO操作,综合改进后的NRBO-XGBoost算法的收敛速度更快、辨识准确率更高。

总体上看,本文提出的改进NRBO-XGBoost诊断算法充分结合了XGBoost算法和NRBO算法各自的优势,在精确度和运算效率两个核心指标上,相较于传统的SVM算法都有显著提升,能够有效服务于线路区内、外故障诊断的需求。同时,通过融合异类多源故障特征,也进一步确保了诊断结果的可靠性。

4 结论

本文提出一种基于改进NRBO-XGBoost的多源数据融合输电线路故障诊断算法。主要结论如下:

1)对输电线路区内、外故障的建模分析表明,线路两侧保护测量的时域和频域差动电压(流)和暂态极性电气量特征以及保护动作的开关量特征均能反映输电线路的故障状态;

2)采用NRBO优化的改进XGBoost算法对于包含电气量和开关量的多源故障特征输入具有较高的辨识度和抗干扰能力,生成的故障诊断模型能够服务于线路区内、外故障诊断的需求;

3)通过与其他传统智能诊断算法的性能仿真对比,验证了新型NRBO-XGBoost算法在线路故障诊断准确率和诊断速度上的优越性。

但需要说明,输电线路区内、外故障关联的状态指标还包含如温度信号、线路图像特征、重合闸跳闸信号等多源化数据。为了更为细化不同线路故障场景下的诊断区分度(如线路金属护层的绝缘损伤程度、具体的故障类型、故障点过渡电阻大小等),需要进一步完善多维度的故障特征分类体系,通过融合更多源的异类故障特征,进一步提升故障诊断结果的完整性和可靠性。

参考文献

- [1] 聂程,王跃,雷万钧,等. 新能源并网输电电缆谐波谐振分析及抑制方法[J]. 电力系统自动化,2017,41(17):128-133.
NIE Cheng, WANG Yue, LEI Wanjun, et al. Harmonic resonance analysis and suppression method for transmission cables between renewable energy sources and power grid[J]. Automation of Electric Power Systems, 2017, 41(17): 128-133.
- [2] 程晓磊,王鹏,王渊,等. 基于大数据的配电网故障诊断预测模型设计[J]. 电气传动,2022,52(2):61-66.
CHENG Xiaolei, WANG Peng, WANG Yuan, et al. Design of fault diagnosis and prediction model for distribution network based on large data[J]. Electric Drive, 2022, 52(2): 61-66.
- [3] 焦京海,李杰. 基于改进残差网络的电缆终端缺陷识别方法[J]. 电气传动,2023,53(11):31-36.
JIAO Jinghai, LI Jie. Terminal defect identification method of cable based on improved residual network[J]. Electric Drive, 2023, 53(11): 31-36.
- [4] JANNATI M, VAHIDI B, HOSSEINIAN S H. Incipient faults monitoring in underground medium voltage cables of distribution systems based on a two-step strategy[J]. IEEE Transactions on Power Delivery, 2019, 34(4): 1647-1655.
- [5] 石立国,李延真,李元付,等. 机理-数据融合驱动的互联变流系统故障诊断方法[J]. 电气传动,2024,54(4):11-20.
SHI Ligu, LI Yanzhen, LI Yuanfu, et al. Mechanism-data fusion driven fault diagnosis method for interconnected converter systems[J]. Electric Drive, 2024, 54(4): 11-20.
- [6] 周晓华,冯雨辰,胡旭初,等. 基于秃鹰搜索算法优化SVM的变压器故障诊断[J]. 南方电网技术,2023,17(6):99-106, 116.
ZHOU Xiaohua, FENG Yuchen, HU Xuchu, et al. Transformer fault diagnosis based on bald eagle search algorithm optimized SVM[J]. Southern Power System Technology, 2023, 17(6): 99-106, 116.
- [7] 李梦宇. 基于多段特征提取的永磁直流电机故障诊断[J]. 自动化应用,2024,65(7):131-134.
LI Mengyu. Fault diagnosis of permanent magnet DC motor based on multi-segment feature extraction[J]. Automation Application, 2024, 65(7): 131-134.
- [8] 赵子莹,彭清文,邓铭,等. 考虑多因素影响与误差修正的充电站负荷预测[J]. 浙江电力,2024,43(4):21-28.
ZHAO Ziyun, PENG Qingwen, DENG Ming, et al. Load fore-

- casting for charging stations considering multiple factors and error correction[J]. Zhejiang Electric Power, 2024, 43(4): 21-28.
- [9] 张睿,杨瑞,董海雷,等. 基于PSO优化LS-SVM的电力现货市场价格预测[J]. 电气传动自动化, 2021, 43(6): 37-41.
ZHANG Rui, YANG Rui, DONG Hailei, et al. Power spot market price prediction based on PSO-optimized LS-SVM[J]. Electric Drive Automation, 2021, 43(6): 37-41.
- [10] 李国栋,周扬,李凯. 基于SARIMAX-XGBoost模型的区域能耗预测[J]. 电力信息与通信技术, 2022, 20(3): 26-33.
LI Guodong, ZHOU Yang, LI Kai. Area energy consumption prediction based on SARIMAX-XGBoost model[J]. Electric Power Information and Communication Technology, 2022, 20(3): 26-33.
- [11] 于飞,樊清川,宣敏. 结合振动特征优选和GWOA-XGBoost的电机轴承故障诊断[J]. 国防科技大学学报, 2023, 45(3): 99-107.
YU Fei, FAN Qingchuan, XUAN Min. Motor bearing fault diagnosis combining vibration feature selection and GWOA-XGBoost[J]. Journal of National University of Defense Technology, 2023, 45(3): 99-107.
- [12] SOWMYA R, PREMKUMAR M, JANGIR P. Newton-Raphson-based optimizer: a new population-based metaheuristic algorithm for continuous optimization problems[J]. Engineering Applications of Artificial Intelligence, 2024, 128: 107532.
- [13] 康小宁,吴招座,索南加乐,等. 参数识别纵联差动保护应用中输电线路模型适用频带研究[J]. 电力系统保护与控制, 2011, 39(19): 8-16.
KANG Xiaoning, WU Zhaozuo, SUONAN Jiale, et al. Research on suitable frequency band of transmission line model in parameter identification application of longitudinal differential protection[J]. Power System Protection and Control, 2011, 39(19): 8-16.
- [14] 索南加乐,王晨清,康小宁,等. 基于分布参数频域模型识别的纵联保护原理[J]. 中国电机工程学报, 2013, 33(10): 145-153, 7.
SUONAN Jiale, WANG Chenqing, KANG Xiaoning, et al. Principles of longitudinal protection based on distributed parameter frequency domain model identification[J]. Proceedings of the CSEE, 2013, 33(10): 145-153, 7.
- [15] 靳维,陆于平. 基于双端弱同步的配电网行波测距方法[J]. 电力自动化设备, 2018, 38(8): 102-109.
JIN Wei, LU Yuping. Fault location method based on traveling wave with dual-terminal weak synchronization for distribution network[J]. Electric Power Automation Equipment, 2018, 38(8): 102-109.
- [16] 靳维. 风电场不确定性弱馈源相互作用机理与继电保护的研究[D]. 南京:东南大学, 2022.
JIN Wei. Study on mechanism of uncertainty weak feeder interaction in wind farm and relay protection[D]. Nanjing: Southeast University, 2022.
- [17] 刘道兵,郭营营,李世春,等. 考虑三段式线路保护主保护范围的电网故障诊断解析模型[J]. 电网技术, 2023, 47(5): 1905-1911.
LIU Daobing, GUO Yingying, LI Shichun, et al. Analytical model of grid fault diagnosis with main protection range considering three-section line protection[J]. Power System Technology, 2023, 47(5): 1905-1911.
- [18] 李婷,张孝军,潘华,等. 基于运检管控平台多源信息融合的输电线路故障类型诊断方法[J]. 南方电网技术, 2023, 17(12): 109-118.
LI Ting, ZHANG Xiaojun, PAN Hua, et al. Fault type diagnosis method for transmission line based on multi-source information fusion on operation and maintenance control platform[J]. Southern Power System Technology, 2023, 17(12): 109-118.

收稿日期:2024-06-03

修改稿日期:2024-07-17