

DOI: 10.19783/j.cnki.pspc.230936

基于特征提取与 INGO-SVM 的变压器故障诊断方法

包金山¹, 杨定坤², 张靖¹, 张英^{1,3}, 杨稼荣¹, 胡克林¹

(1. 贵州大学电气工程学院, 贵州 贵阳 550025; 2. 重庆邮电大学先进制造工程学院, 重庆 400065;
3. 贵州电网有限责任公司电力科学研究院, 贵州 贵阳 550002)

摘要: 针对使用支持向量机(support vector machine, SVM)对变压器进行故障诊断时有效特征提取困难、模型参数难以选择的问题, 提出一种基于特征提取与 INGO-SVM 的变压器故障诊断方法。首先, 使用核主成分分析(kernel principal component analysis, KPCA)方法对构建的 21 维待选特征进行特征融合和低维敏感特征提取。其次, 使用佳点集、随机反向学习和维度交叉学习等策略对北方苍鹰优化算法(northern goshawk optimization, NGO)进行改进。通过 2 个典型测试对改进北方苍鹰优化算法(improved northern goshawk optimization, INGO)进行性能测试, 验证了 INGO 算法的优越性。然后, 基于 KPCA 提取的低维敏感特征, 使用 INGO 对 SVM 的参数进行组合寻优, 建立基于 KPCA 特征提取与 INGO-SVM 的变压器故障诊断模型。最后, 对不同变压器故障诊断模型进行实例仿真对比实验。结果表明: 所提方法故障诊断精度高、稳定性好, 更适用于变压器的故障诊断。

关键词: 变压器; 故障诊断; 支持向量机; 核主成分分析; 北方苍鹰优化算法

Transformer fault diagnosis method based on feature extraction and INGO-SVM

BAO Jinshan¹, YANG Dingkun², ZHANG Jing¹, ZHANG Ying^{1,3}, YANG Jiarong¹, HU Kelin¹

(1. College of Electrical Engineering, Guizhou University, Guiyang 550025, China; 2. College of Advanced Manufacturing Engineering, Chongqing University of Posts & Telecommunications, Chongqing 400065, China;
3. Electric Power Research Institute of Guizhou Power Grid Co., Ltd., Guiyang 550002, China)

Abstract: It is difficult to extract effective features and select model parameters when using a support vector machine (SVM) for transformer fault diagnosis. A transformer fault diagnosis method based on feature extraction and an improved northern goshawk optimization (INGO) algorithm optimized SVM is proposed. First, kernel principal component analysis (KPCA) is used to conduct feature fusion and low dimensional sensitive feature extraction for the 21 dimensional candidate feature. Secondly, strategies such as good point set, random opposition-based learning, and dimensional cross learning are used to improve the northern goshawk optimization (NGO) algorithm. The performance of the INGO algorithm is tested using two typical test functions, verifying its superiority. Then, based on the low dimensional sensitive feature extracted by KPCA, INGO is used to optimize the parameters of the SVM, and a transformer fault diagnosis model is established based on KPCA feature extraction and INGO-SVM. Finally, simulation and comparative experiments are conducted on different transformer fault diagnosis models. The results show that the proposed method has high accuracy and good stability in fault diagnosis, and is more suitable for transformer fault diagnosis.

This work is supported by the National Natural Science Foundation of China (No. 52177016).

Key words: transformer; fault diagnosis; support vector machine; kernel principal component analysis; NGO

0 引言

变压器是电网系统中的重要组成部分, 其在运

行过程中发生故障会给电网系统造成巨大的经济损失。因此, 建立高效的变压器故障诊断模型, 辅助运维人员及时发现设备内部的异常信息, 对维护变压器的长期稳定运行具有重要意义^[1-3]。

在我国, 电网系统中使用的大型变压器以油浸式电力变压器为主^[4]。油浸式变压器在运行过程中出现发热、放电等现象时, 绝缘材料会裂解生成一

基金项目: 国家自然科学基金项目资助(52177016); 贵州省科技计划项目资助(黔科合支撑[2021]365); 贵州大学自然科学特别科学研究基金项目资助(2021-45)

系列气体溶解在变压器油中,且油中溶解气体含量与变压器的运行状态具有很强的相关性。因此,通过油中溶解气体分析(dissolved gas analysis, DGA)技术可以对变压器的运行状态进行诊断^[5]。基于 DGA 技术的变压器故障诊断方法分为传统方法和人工智能方法。传统方法有罗杰斯比值法^[6]、IEC 三比值法^[7]和大卫三角法^[8]等,这些方法在实际使用过程中存在编码缺失、边界过于绝对等缺陷,导致诊断效果不佳。

近年来,我国发布了《中国电力大数据白皮书》、《新一代人工智能发展规划》等文件,确定了我国电力系统及其设备的运行、维护正朝着人工智能方向跃进^[9]。在这样的时代背景下,人工神经网络^[10-11]、模糊逻辑算法^[12-13]、极限学习机^[14-15]、支持向量机^[16-17]等人工智能诊断方法得以快速发展。人工智能方法在一定程度上提高了变压器的故障诊断精度,但也存在一些不足。其一,在特征选取方面:大部分人工智能诊断方法直接将 DGA 气体浓度数据作为模型输入的特征量,而非采用和变压器运行状态关联特性更强的 DGA 气体浓度比值数据作为特征量^[18-19]。其二,在机器学习算法方面:人工神经网络的模型复杂,在训练阶段需要大量的样本数据且存在收敛速度慢和容易陷入局部最优的缺陷;极限学习机训练速度快但稳定性较差;模糊逻辑算法对样本数据要求较高;支持向量机能够很好地解决变压器故障诊断这类小样本、非线性问题,但分类精度与惩罚因子和核函数参数的选取密切相关^[20-22]。

综上所述,本文提出一种基于特征提取与 INGO-SVM 的变压器故障诊断方法。首先,结合 IEC 和 IEEE 推荐的常用特征选取方法对特征量进行扩充,得到 21 维待选特征,并使用核主成分分析方法对 21 维待选特征进行特征融合和低维敏感特征提取;其次,使用佳点集、随机反向学习和维度交叉学习等策略对北方苍鹰优化算法进行改进,解决原始北方苍鹰优化算法容易陷入局部最优和收敛精度低的缺陷;然后,基于核主成分分析得到的低维敏感特征,使用改进北方苍鹰优化算法对支持向量机的惩罚因子和核函数参数进行组合寻优,获得基于特征提取与 INGO-SVM 的变压器故障诊断模型。最后,通过实例仿真对比实验,验证了所提方法的有效性。

1 基于 KPCA 的变压器故障特征提取

在传统的人工智能诊断方法中,通常直接将 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6 这 5 种气体浓度数据作为模型输入的特征量。但 5 种气体浓度数据包含的故障信息并不完整,导致模型容易陷入局部最优,

降低了故障诊断结果的可靠性。文献[23]表明,油中溶解气体浓度比值数据与变压器运行状态之间的关系更加密切。但以 5 种气体浓度数据为基础的传统比值法,因选取的特征量数量较少,无法准确反映气体数据与变压器故障类型之间的关联特性。因此,结合 IEC 和 IEEE 推荐的特征选取方法对特征量进行扩充,得到 21 种待选特征如表 1 所示。

表 1 变压器故障待选特征集

编号	特征量	编号	特征量	编号	特征量
S1	H_2	S8	C_2H_2/C_2H_4	S15	$C_2H_6/总烃$
S2	CH_4	S9	C_2H_2/C_2H_6	S16	$C_2H_6/总气$
S3	C_2H_6	S10	C_2H_4/C_2H_6	S17	$C_2H_4/总烃$
S4	C_2H_4	S11	$H_2/总烃$	S18	$C_2H_4/总气$
S5	C_2H_2	S12	$H_2/总气$	S19	$C_2H_2/总烃$
S6	CH_4/H_2	S13	$CH_4/总烃$	S20	$C_2H_2/总气$
S7	C_2H_2/CH_4	S14	$CH_4/总气$	S21	$总烃/总气$

在表 1 中, S6—S21 为增加的 16 种浓度比值特征,其中,总气 = $H_2 + CH_4 + C_2H_2 + C_2H_4 + C_2H_6$,总烃 = $CH_4 + C_2H_2 + C_2H_4 + C_2H_6$ 。为使得样本数据在保留数据原有信息的前提下减少数据间的量级差异,对上述 21 种待选特征按照式(1)进行 Z-score 标准化处理。

$$S_i^* = \frac{S_i - \mu_i}{\delta_i} \quad (1)$$

式中: S_i 为第 i 个待选特征的原始值; S_i^* 为 S_i 标准化处理后的值; μ_i 、 δ_i 分别为第 i 个待选特征的均值和标准差。

进一步,针对待选特征之间的信息冗余问题,使用 KPCA 方法对 Z-score 标准化处理后的 21 种待选特征进行特征融合和低维敏感特征提取,得到的 Pareto 图如图 1 所示。

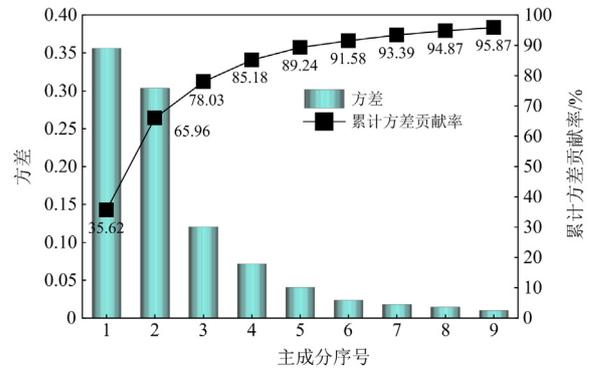


图 1 核主成分分析

Fig. 1 Kernel principal component analysis

由图 1 可知, 前 9 个主成分已经包含故障样本的绝大部分信息, 累计方差贡献率达到 95.87%, 超过 95%。因此选择降维后的低维敏感特征维数为 9。

2 北方苍鹰优化算法及其改进

2.1 北方苍鹰优化算法

北方苍鹰优化(northern goshawk optimization, NGO)算法是由 Mohammad Dehghani 等人于 2021 年提出的一种新型元启发式智能优化算法^[24], 算法灵感来源于北方苍鹰的狩猎策略。在 NGO 算法中, 北方苍鹰种群用矩阵 \mathbf{X} 表示为

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_i \\ \vdots \\ \mathbf{X}_N \end{bmatrix} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & & \vdots & & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & & \vdots & & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix}_{N \times m} \quad (2)$$

式中: N 为种群规模; m 为求解问题的维度; \mathbf{X}_i 为第 i 个北方苍鹰的位置; $x_{i,j}$ 为 \mathbf{X}_i 的第 j 维分量。

NGO 算法的迭代寻优过程分为猎物识别阶段和追逐及逃生阶段, 具体过程如下所述。

阶段一: 猎物识别阶段, 该阶段属于算法的全局搜索阶段。在此阶段, 北方苍鹰在目标搜索空间内随机选择一个猎物并迅速对其进行攻击, 其数学模型为

$$\mathbf{P}_i = \mathbf{X}_k, k=1,2,\dots,i-1,i+1,\dots,N \quad (3)$$

$$x_{i,j}^{\text{new},P1} = \begin{cases} x_{i,j} + r(p_{i,j} - lx_{i,j}), & F_i < F_p \\ x_{i,j} + r(x_{i,j} - p_{i,j}), & F_i < F_p \end{cases} \quad (4)$$

$$\mathbf{X}_i = \begin{cases} \mathbf{X}_i^{\text{new},P1}, & F_i^{\text{new},P1} < F_i \\ \mathbf{X}_i, & F_i \leq F_i^{\text{new},P1} \end{cases} \quad (5)$$

式中: \mathbf{P}_i 为第 i 个北方苍鹰的猎物位置; $p_{i,j}$ 为 \mathbf{P}_i 的第 j 维分量; r 为 $[0,1]$ 范围内分布的随机数; l 为取值为 1 或 2 的随机整数; F_i 、 F_p 分别为第 i 个北方苍鹰及其猎物的目标函数值; $\mathbf{X}_i^{\text{new},P1}$ 为第一阶段更新后第 i 个北方苍鹰的新位置; $x_{i,j}^{\text{new},P1}$ 为 $\mathbf{X}_i^{\text{new},P1}$ 的第 j 维分量; $F_i^{\text{new},P1}$ 为第一阶段更新后第 i 个北方苍鹰的目标函数值。

阶段二: 追逐及逃生阶段, 该阶段属于算法的局部搜索阶段。当北方苍鹰对猎物进行攻击后, 猎物会尝试逃跑, 北方苍鹰会继续对其进行追捕。假设本次狩猎范围的半径为 R , 北方苍鹰与猎物之间追逐过程的数学模型为

$$x_{i,j}^{\text{new},P2} = x_{i,j} + R(2r-1)x_{i,j} \quad (6)$$

$$R = 0.02(1-t/T) \quad (7)$$

$$\mathbf{X}_i = \begin{cases} \mathbf{X}_i^{\text{new},P2}, & F_i^{\text{new},P2} < F_i \\ \mathbf{X}_i, & F_i \leq F_i^{\text{new},P2} \end{cases} \quad (8)$$

式中: $\mathbf{X}_i^{\text{new},P2}$ 为第二阶段更新后第 i 个北方苍鹰的新位置; $x_{i,j}^{\text{new},P2}$ 为 $\mathbf{X}_i^{\text{new},P2}$ 的第 j 维分量; t 为当前迭代次数; T 为最大迭代次数; $F_i^{\text{new},P2}$ 为第二阶段更新后第 i 个北方苍鹰的目标函数值。

2.2 北方苍鹰优化算法的改进

2.2.1 佳点集种群初始化

在原始 NGO 算法中, 北方苍鹰种群成员在目标搜索空间内随机初始化。这种初始化方法难以保证种群的多样性, 从而影响算法后续的寻优性能。为提高 NGO 算法的初始种群质量, 提出一种基于佳点集理论的种群初始化方法。

佳点集是由我国数学家华罗庚等人提出的数学理论^[25], 其基本原理与构造如下。

① 设 G_s 是 s 维欧式空间中的单位立体。

② 如果 $d \in G_s$, 存在点集 $P(k)$

$$P(k) = \{(\{d_1^k \cdot k\}, \{d_2^k \cdot k\}, \dots, \{d_s^k \cdot k\}), 1 \leq k \leq N\} \quad (9)$$

式中: $k=1,2,\dots,N$; $\{d_s^k \cdot k\}$ 表示取小数部分。

若 $P(k)$ 的方差 $\varphi(n)$ 满足

$$\varphi(n) = V(d, \varepsilon) n^{-1+\varepsilon} \quad (10)$$

式中, $V(d, \varepsilon) n^{-1+\varepsilon}$ 为只与 d 和 ε (ε 为任意正数) 有关的常数, 则称 $P(k)$ 为佳点集, d 为佳点。

③ 一般情况下, 采用分圆域法求取佳点集, 取 $d_s^k = 2 \cos(2\pi k/p), 1 \leq k \leq s$, p 表示满足圆域的最小素数。

使用佳点集策略改进后的种群初始化公式为

$$x_{i,j}^0 = l_{b_j} + (u_{b_j} - l_{b_j}) \cdot \{d_j^i \cdot i\} \quad (11)$$

式中: $x_{i,j}^0$ 为初始种群中第 i 个北方苍鹰的第 j 维分量; u_{b_j} 和 l_{b_j} 分别为北方苍鹰在第 j 维分量上的上界和下界; d_j^i 为使用分圆域法获得的第 i 个佳点在第 j 维分量上的取值。

2.2.2 随机反向学习策略

随机反向学习策略(random opposition-based learning, ROBL)是由 LONG 等人提出的一种算法改进策略^[26], 该策略的核心思想是根据当前个体在目标搜索空间内生成一个随机反向个体, 以丰富种群多样性, 从而提高算法的搜索能力。将 ROBL 应用于 NGO 算法的第一阶段, 在迭代前期, 可以增强算法的全局寻优能力; 在迭代后期, 种群的多样性还可以提高算法避免陷入局部最优的能力。

但若在每次迭代时, 对所有个体执行随机反向学习操作, 必然会增加算法的计算时间。故通过设置动态选择概率 ρ 对此进行调节。

$$\rho = \omega_1 + (\omega_2 - \omega_1) \times (1 - t/T) \quad (12)$$

式中: ω_1 为动态选择概率下限, 取 $\omega_1 = 0.2$; ω_2 为动态选择概率上限, 取 $\omega_2 = 0.5$ 。

结合动态选择概率 ρ , 使用随机反向学习策略对第一阶段的北方苍鹰个体进行位置更新, 更新公式为

$$x_{i,j}^{\text{new},P1} = l_{b_j} + r(u_{b_j} - x_{i,j}), r < \rho \quad (13)$$

2.2.3 维度交叉学习策略

NGO 算法通过 N 个北方苍鹰并行迭代搜索目标问题的最优解。在北方苍鹰种群中, 如果将目标函数值较高的个体称为精英个体, 则不同精英个体落入最优解领域的维度分量各不相同。为充分共享精英个体的优势维度信息, 增强算法的局部开采能力, 提出一种维度交叉学习策略。同时, 为了更好地平衡算法的全局搜索和局部开发能力, 对狩猎范围半径 R 进行非线性改进。改进后, 迭代前期北方苍鹰种群具有较大的狩猎范围, 有利于增强算法的全局搜索能力; 迭代后期北方苍鹰种群具有较小的狩猎范围, 有利于增强算法的局部开发能力。改进后的狩猎范围半径公式为

$$R = 0.2(1 - t/T)^{2t/T} \quad (14)$$

进一步, 将 NGO 算法第二阶段的北方苍鹰按目标函数值排序, 平均分为两组。目标函数值较低的一组北方苍鹰个体按式(6)和式(14)进行位置更新。目标函数值较高的一组北方苍鹰个体使用维度交叉学习的方法进行位置更新, 更新公式为

$$x_{i,j}^{\text{new},P2} = x_{i,j}L + e_{i,j}(1-L) \quad (15)$$

式中: $e_{i,j}$ 为第 i 个北方苍鹰交叉学习对象的第 j 维分量; L 取值为 0 或者 1, 其表达式如式(16)。

$$L = \begin{cases} 0, & r \leq p_{\text{dim}} \\ 1, & r > p_{\text{dim}} \end{cases} \quad (16)$$

式中, p_{dim} 为维度交叉学习概率, 取 $p_{\text{dim}} = 0.3$ 。

2.3 INGO 寻优性能测试

为测试 INGO 算法的改进效果, 本文从 CEC2005 中选择 2 个典型测试函数对其进行寻优性能测试, 并与 NGO、GWO 和 PSO 算法进行对比分析。2 个典型测试函数的函数表达式如式(17)和式(18)所示, 其余函数信息如表 2 所示。

$$f_1(x) = \max_i \{|x_i|, 1 \leq i \leq n\} \quad (17)$$

$$f_2(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \quad (18)$$

表 2 测试函数信息

Table 2 Information of test functions

函数	函数类型	维度	取值范围	最优解
f_1	Schwefel	30	[-100,100]	0
f_2	Rastrigin	30	[-5.12,5.12]	0

在仿真实验过程中, 为确保实验的公平性, 各优化算法设置以下公共参数: 种群规模 $N = 30$, 最大迭代次数 $T = 1000$ 。同时, 为了避免单次运行导致的结果偏差, 本文在 Matlab 上对每个测试函数独立运行 20 次, 得到平均最优适应度变化曲线分别如图 2、图 3 所示。

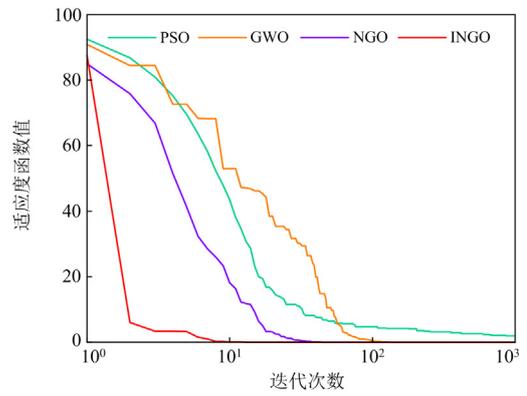


图 2 $f_1(x)$ 函数寻优过程

Fig. 2 $f_1(x)$ function optimization process

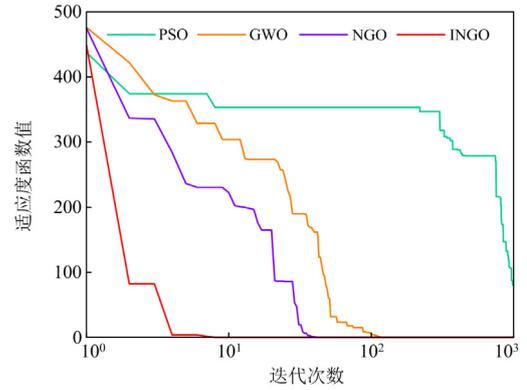


图 3 $f_2(x)$ 函数寻优过程

Fig. 3 $f_2(x)$ function optimization process

INGO 及其对比算法的寻优结果统计如表 3 所示。表 3 中, Mean 为平均最优适应度函数值, Best 为 20 次寻优过程的最优适应度函数值, Sad 为标准差。其中, Mean 和 Best 用于检验算法的搜索性能, Sad 用于检验算法的稳定性, 表中加粗部分表示各统计量的最优结果。

表 3 寻优结果对比

Table 3 Comparison of optimization results

函数	统计值	INGO	NGO	GWO	PSO
f_1	Best	5.89×10^{-116}	3.05×10^{-79}	2.21×10^{-16}	1.21
	Mean	2.75×10^{-112}	2.02×10^{-76}	1.05×10^{-14}	1.64
	Sad	6.49×10^{-112}	2.25×10^{-76}	1.21×10^{-14}	0.25
f_2	Best	0	0	0	78.64
	Mean	0	0	0.33	107.64
	Sad	0	0	1.10	36.67

由图 2、图 3 及表 3 可以看出，在两个测试函数寻优过程中，INGO 算法在收敛速度、寻优精度和稳定性方面明显优于 NGO、GWO 和 PSO 算法，验证了本文所提改进方法的优越性。

3 基于特征提取的 INGO-SVM 模型

3.1 支持向量机

标准的 SVM 是一种线性二分类模型^[27]，基于给定训练样本集，在样本空间构造一个最优分类超平面： $\mathbf{w}^T \mathbf{x} + b = 0$ ，将两类不同类别的样本分开。最优分类超平面的求解过程在数学上可以描述为一个二次凸优化问题，可表示为

$$\begin{cases} \min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}^T\|^2 \\ \text{s.t. } y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, i = 1, 2, \dots, m \end{cases} \quad (19)$$

式中： \mathbf{w} 为超平面法向量； b 为偏置量； \mathbf{x}_i 为输入故障样本数据； y_i 为 \mathbf{x}_i 对应的故障标签。

但变压器的故障诊断属于线性不可分问题，需要引入核函数和软间隔。其中，核函数的作用是将样本数据从原始空间映射到一个高维特征空间，使得样本数据在高维特征空间线性可分；而软间隔的作用是通过引入惩罚因子和松弛变量，允许超平面的划分在一些样本上出错。此时，求解广义最优分类超平面的目标函数和约束条件更新为

$$\begin{cases} \min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}^T\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s.t. } y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, 2, \dots, m \end{cases} \quad (20)$$

式中： C 为惩罚因子； ξ_i 为松弛变量； $\phi(\mathbf{x}_i)$ 为将 \mathbf{x}_i 映射到高维特征空间的特征向量。

使用拉格朗日乘子法和对偶原理得到式(20)的对偶问题为

$$\begin{cases} \max_a \sum_{i=1}^m a_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m a_i a_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s.t. } \sum_{i=1}^m a_i y_i = 0, 0 \leq a_i \leq C, i = 1, 2, \dots, m \end{cases} \quad (21)$$

式中： \mathbf{a} 为拉格朗日乘子向量； a_i 为拉格朗日乘子； $K(\mathbf{x}_i, \mathbf{x}_j)$ 为核函数。

核函数的选择对于 SVM 的分类效果影响显著，径向基核函数^[28](radial basis function, RBF)的分类效果优于其他核函数，因此，本文选择 RBF 核函数作为 SVM 的核函数。RBF 核函数的表达式为

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (22)$$

式中， γ 为核函数参数。

最后，结合 KKT 条件求解得分类决策函数为

$$f(\mathbf{x}) = \sum_{i=1}^m a_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (23)$$

综上所述，SVM 的分类性能与惩罚因子 C 和核函数参数 γ 的设定密切相关。因此，采用 INGO 算法对 SVM 参数进行组合寻优。

3.2 基于有效特征提取的 INGO-SVM 模型流程

所提基于特征提取与 INGO-SVM 的变压器故障诊断方法，主要包括模型输入特征提取和模型参数优化两个方面。首先，采用 KPCA 方法对模型输入特征进行提取。然后，使用 INGO 算法对 SVM 的惩罚因子 C 和核函数参数 γ 进行组合寻优，构建最优参数组合 SVM 模型，根据最优参数组合 SVM 模型对测试集数据进行故障诊断。

在使用 INGO 算法对 SVM 的参数进行优化时，定义适应度函数为 SVM 对训练集样本数据进行 k 折交叉验证返回的平均分类准确率。基于特征提取与 INGO-SVM 的变压器故障诊断流程如图 4 所示，具体步骤描述如下。

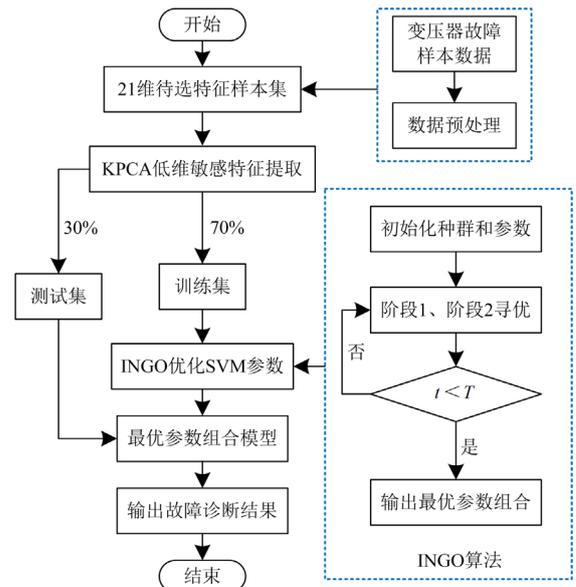


图 4 变压器故障诊断流程

Fig. 4 Flow chart of transformer fault diagnosis

步骤 1: 收集变压器故障样本数据。

步骤 2: 数据预处理。以 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6 这 5 种 DGA 气体浓度数据为基础, 计算得到 21 维待选特征样本矩阵, 使用 Z-score 标准化方法对样本数据进行标准化处理, 得到 21 维待选特征样本集。

步骤 3: 有效特征提取。使用 KPCA 方法对 21 种待选特征进行特征融合和低维敏感特征提取, 并按比例将样本数据划分为训练集和测试集。

步骤 4: 模型训练。基于训练集数据, 使用 INGO 算法对 SVM 的惩罚因子 C 和核函数参数 γ 进行组合寻优, 建立最优参数组合 SVM 模型。

步骤 5: 故障诊断。将测试集数据输入最优参数组合 SVM 模型, 输出对应的变压器故障类型。

4 变压器故障诊断实例分析

本文从文献[29]中收集到具有明确结论的 260 组变压器样本数据, 剔除其中与同类型样本数据数值差异较大的 9 个异常样本后, 将剩下的 251 组样本作为数据集 1, 用于对模型的准确性和稳定性进行验证。每组样本数据中包括变压器油中 5 种溶解气体(H_2 、 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6)含量及其对应的样本类型。变压器的样本类型分为: 高能放电(H-D)、低能放电(L-D)、局部放电(P-D)、高温过热(H-T)、中低温过热(LM-T)及正常状态(N-C), 对其编号为 1~6。结合 IEC 和 IEEE 推荐的常用特征量选取方法对特征量进行扩充, 得到 21 种待选特征, 并采用 KPCA 方法对待选特征进行特征融合和低维敏感特征提取, 获得最终的特征样本集, 将最终的特征样本集按 7:3 的比例分为训练集和测试集, 作为模型的输入。数据集 1 的故障类型编号及样本数据分布如表 4 所示。

表 4 故障类型编号及样本数据分布(数据集 1)

Table 4 Type number and sample data distribution (Dataset 1)

故障类型	H-D	L-D	P-D	H-T	LM-T	N-C
类型编号	1	2	3	4	5	6
训练集	30	31	29	30	31	25
测试集	13	13	13	13	13	10

4.1 特征优选对 INGO-SVM 模型的性能影响

为验证基于 KPCA 有效特征提取方法的优越性, 与传统特征选取方法进行仿真对比实验。基于数据集 1, 将各方法选取的特征量输入 INGO-SVM 模型进行故障诊断。仿真实验过程中, 各参数设置如下: 惩罚因子 c 的搜索范围为(0.1,100); 核函数参数 γ 的搜索范围为(0.1,5); 种群规模 $N=30$; 最大迭代次数 $T=200$; 交叉验证的折数 $k=10$ 。

诊断对比结果如表 5 所示。由表 5 可以看出, 基于 KPCA 特征提取与 INGO-SVM 模型的故障诊断精度最高, 故障诊断精度为 92.00%; 而使用 ICUSU 法、关键气体法、IEC 三比值法和 Doernenburg 法推荐的特征量作为 INGO-SVM 模型输入时的故障诊断精度依次是 84.00%、78.67%、77.33%和 70.67%。结果表明基于 KPCA 的特征提取方法能够显著提高变压器的故障诊断精度。

表 5 诊断结果对比

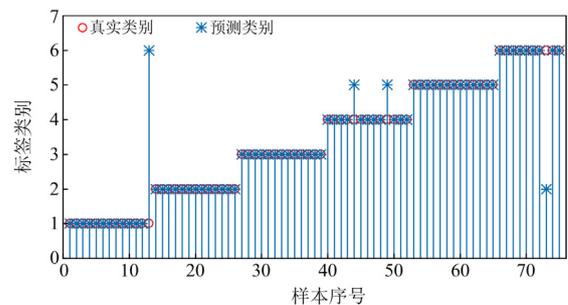
Table 5 Comparison of diagnostic results

特征选取方法	参数 (C, γ)	判断准确率/%
Doernenburg 法	(169.667,0.559)	70.67
ICUSU 法	(137.511,0.198)	84.00
IEC 三比值法	(129.217,0.759)	77.33
关键气体法	(198.221,0.982)	78.67
KPCA 特征提取	(40.084,3.113)	92.00

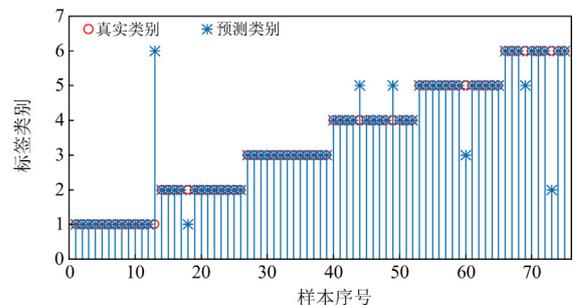
4.2 不同优化算法优化 SVM 的性能对比

为了验证 INGO-SVM 模型的优越性, 本文将数据集 1 经 KPCA 特征提取后的变压器样本数据分别输入 INGO-SVM、NGO-SVM、GWO-SVM 和 PSO-SVM 模型进行故障诊断。实验过程中参数设置与 4.1 节保持一致。故障诊断结果如图 5 所示。

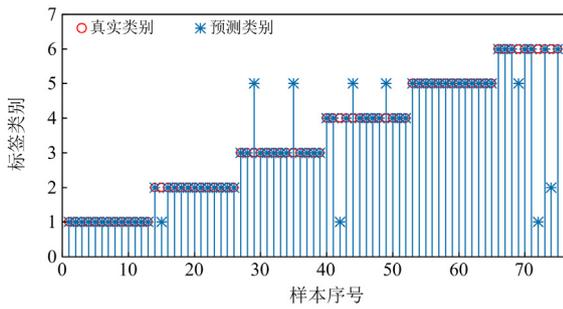
由图 5 可以看出, INGO-SVM 模型的诊断性能更加突出, 在 75 个测试样本中只有 4 个样本预测错误, 故障诊断精度达到 94.67%。而 NGO-SVM、GWO-SVM 和 PSO-SVM 模型分别在 75 个预测样本中有 7 个、9 个和 11 个预测错误, 故障诊断精度



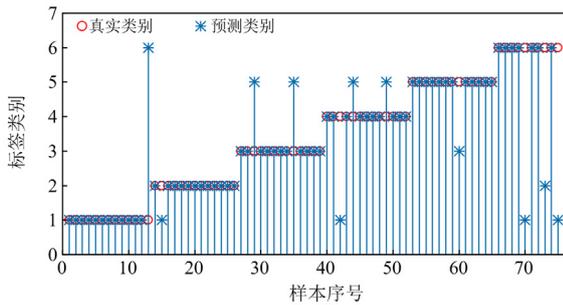
(a) INGO-SVM模型诊断结果: 94.67%



(b) NGO-SVM模型诊断结果: 90.67%



(c) GWO-SVM模型诊断结果: 88.00%



(d) PSO-SVM模型诊断结果: 85.33%

图 5 不同优化算法优化 SVM 的故障诊断结果

Fig. 5 Fault diagnosis results of SVM optimized by different optimization algorithms

分别为 90.67%、88.00%和 85.33%。仿真对比实验表明 INGO-SVM 模型的故障诊断精度最高。

为了进一步对 INGO-SVM 模型的稳定性进行检验, 将上述 4 种方法分别做 20 次独立仿真实验, 统计最高诊断精度、平均诊断精度和标准差如表 6 所示, 表中加粗部分为各统计量的最优值。

表 6 不同诊断方法诊断结果

Table 6 Diagnostic results using different diagnostic methods

模型	最高诊断率	平均诊断率	标准差
INGO-SVM	96.00	93.53	1.58
NGO-SVM	92.00	88.47	3.01
GWO-SVM	94.67	87.33	3.86
PSO-SVM	92.00	87.27	4.14

由表 6 可以看出, INGO-SVM 模型的最高诊断精度为 96.00%, 平均诊断精度为 93.53%, 标准差为 1.58%。INGO-SVM 模型的平均诊断精度分别比 NGO-SVM、GWO-SVM 和 PSO-SVM 模型高出 5.06%、6.2%和 6.26%, 同时标准差也是 4 个模型中最低的。实验结果表明所提方法与其余方法相比, 诊断性能最优且最稳定。

4.3 实例仿真分析

为了验证所提基于特征提取与 INGO-SVM 的

变压器故障诊断方法对不同变压器样本数据的适应能力, 从文献[30]中收集到上海电网 220 kV 及以上变压器油色谱数据共 198 组, 作为数据集 2, 用于对模型准确性和稳定性再次进行实例仿真分析。数据集 2 中每组样本数据包括变压器油中 5 种溶解气体含量及其对应的样本类型。其中, 变压器的样本类型分为: 高能放电(H-D)、低能放电(L-D)、局部放电(P-D)、高温过热(H-T)、中温过热(M-T)及正常状态(N-C), 对其编号为 1~6。数据集 2 的故障类型编号及样本数据分布如表 7 所示。

表 7 故障类型编号及样本数据分布(数据集 2)

Table 7 Type number and sample data distribution (Dataset 2)

故障类型	H-D	L-D	P-D	H-T	M-T	N-C
类型编号	1	2	3	4	5	6
训练集	14	14	14	15	14	68
测试集	6	6	6	6	6	29

将数据集 2 经 KPCA 特征提取后的变压器样本数据输入 INGO-SVM 模型进行故障诊断, 其故障诊断结果如图 6 所示。

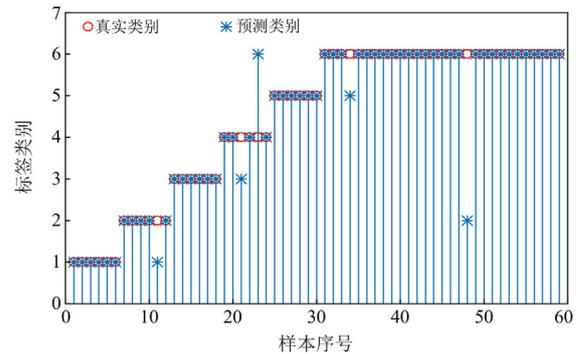


图 6 数据集 2 故障诊断结果

Fig. 6 Fault diagnosis results of Dataset 2

由图 6 可以看出, 所提故障诊断方法在数据集 2 的 59 个测试集样本中只有 5 个样本预测错误, 故障诊断精度达到 91.52%。进一步, 基于数据集 2 对模型进行 20 次独立仿真实验, 得到其平均诊断精度为 89.67%, 标准差为 2.37%。实验结果表明所提方法在保证故障诊断精度的同时具有很好的稳定性。

数据集 1 和数据集 2 的仿真结果表明, 所提方法故障诊断精度高、稳定性好, 更加适用于变压器的故障诊断。

5 结论

1) 用 KPCA 对构建的 21 种待选特征进行特征融合和低维敏感特征提取, 获得能够表达变压器故障有效信息的 9 个低维敏感特征, 避免了多维特征

之间的信息冗余和高维数据容易影响模型诊断精度的问题, 提升了变压器故障诊断模型的诊断精度。

2) INGO 算法克服了 NGO 算法存在的容易陷入局部最优和收敛精度低的缺陷, 与 NGO、PSO 和 GWO 算法相比, INGO 算法具有更快的寻优速度、更高的寻优精度和稳定性。

3) 实例仿真结果表明, 基于 KPCA 特征提取与 INGO-SVM 的变压器故障诊断模型在保证故障诊断精度的同时具有更高的稳定性, 更加适用于变压器的故障诊断。

参考文献

- [1] 杨威, 蒲彩霞, 杨坤, 等. 基于 CNN-GRU 组合神经网络的变压器短期故障预测方法[J]. 电力系统保护与控制, 2022, 50(6): 107-116.
YANG Wei, PU Caixia, YANG Kun, et al. Short-term fault prediction method for a transformer based on a CNN-GRU combined neural network[J]. Power System Protection and Control, 2022, 50(6): 107-116.
- [2] 朱保军, 咸日常, 范慧芳, 等. WRSR 与改进朴素贝叶斯融合的变压器故障诊断技术研究[J]. 电力系统保护与控制, 2021, 49(20): 120-128.
ZHU Baojun, XIAN Richang, FAN Huifang, et al. Transformer fault diagnosis technology based on the fusion of WRSR and improved naive Bayes[J]. Power System Protection and Control, 2021, 49(20): 120-128.
- [3] 王艳, 李伟, 赵洪山, 等. 基于油中溶解气体分析的 DBN-SSAELM 变压器故障诊断方法[J]. 电力系统保护与控制, 2023, 51(4): 32-42.
WANG Yan, LI Wei, ZHAO Hongshan, et al. DBN-SSAELM transformer fault diagnosis method based on dissolved gas analysis in oil[J]. Power System Protection and Control, 2023, 51(4): 32-42.
- [4] KHARE V, KHARE C J. Aspects of transformer in electricity generation: a review[J]. Journal of Engineering and Technology, 2018, 7(2): 66-84.
- [5] Mineral oil-impregnated electrical equipment in service guide to the interpretation of dissolved and free gases analysis: IEC-60599[S]. Geneva, Switzerland: International Electrotechnical Commission, 2007.
- [6] ROGERS R R. IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis[J]. IEEE Transactions on Electrical Insulation, 1978(5): 349-354.
- [7] IEC. Mineral oil-impregnated electrical equipment in service-guide to the interpretation of dissolved and free gases analysis: IEC 60599—2007[S]. Geneva, Switzerland: IEC, 2007.
- [8] DUVAL M. A review of faults detectable by gas-in-oil analysis in transformers[J]. IEEE Electrical Insulation Magazine, 2002, 18(3): 8-17.
- [9] 张国治, 陈康, 方荣行, 等. 基于 DGA 与鲸鱼算法优化 Logit Boost-决策树的变压器故障诊断方法[J]. 电力系统保护与控制, 2023, 51(7): 63-72.
ZHANG Guozhi, CHEN Kang, FANG Rongxing, et al. Transformer fault diagnosis method based on DGA and whale algorithm optimizing Logit Boost-decision tree[J]. Power System Protection and Control, 2023, 51(7): 63-72.
- [10] PATEKAR K D, CHAUDHRY B. DGA analysis of transformer using artificial neural network to improve reliability in power transformers[C]// 2019 IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), November 21-23, 2019, Chennai, India: 1-5.
- [11] 吴君, 丁欢欢, 马星河, 等. 改进自适应蜂群优化算法在变压器故障诊断中的应用[J]. 电力系统保护与控制, 2020, 48(9): 174-180.
WU Jun, DING Huanhuan, MA Xinghe, et al. Application of improved adaptive bee colony optimization algorithm in transformer fault diagnosis[J]. Power System Protection and Control, 2020, 48(9): 174-180.
- [12] KHAN S A, EQUBAL M D, ISLAM T. A comprehensive comparative study of DGA based transformer fault diagnosis using fuzzy logic and ANFIS models[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(1): 590-596.
- [13] 张鹏, 齐波, 李文璞, 等. 考虑部位和性能的电力变压器差异化评价方法[J]. 中国电机工程学报, 2019, 39(20): 6138-6147, 6195.
ZHANG Peng, QI Bo, LI Wenpu, et al. A differentiated evaluation method of transformers considering the part and its performances[J]. Proceedings of the CSEE, 2019, 39(20): 6138-6147, 6195.
- [14] 吕忠, 周强, 周琨, 等. 基于遗传算法改进极限学习机的变压器故障诊断[J]. 高压电器, 2015, 51(8): 49-53.
LÜ Zhong, ZHOU Qiang, ZHOU Kun, et al. Fault diagnosis of transformer based on extreme learning machine optimized by genetic algorithm[J]. High Voltage Apparatus, 2015, 51(8): 49-53.
- [15] 吴杰康, 覃炜梅, 梁浩浩, 等. 基于自适应极限学习机的变压器故障识别方法[J]. 电力自动化设备, 2019, 39(10): 181-186.
WU Jiekang, QIN Weimei, LIANG Haohao, et al. Transformer fault identification method based on self-adaptive extreme learning machine[J]. Electric Power Automation Equipment, 2019, 39(10): 181-186.
- [16] 叶远波, 李端超, 谢民, 等. 基于 SSA-SVM 的继电

- 保护装置状态评估方法研究[J]. 电力系统保护与控制, 2022, 50(8): 171-178.
- YE Yuanbo, LI Duanchao, XIE Min, et al. A state evaluation method for a relay protection device based on SSA-SVM[J]. Power System Protection and Control, 2022, 50(8): 171-178.
- [17] 方涛, 钱晔, 郭灿杰, 等. 基于天牛须搜索优化支持向量机的变压器故障诊断研究[J]. 电力系统保护与控制, 2020, 48(20): 90-96.
- FANG Tao, QIAN Ye, GUO Canjie, et al. Research on transformer fault diagnosis based on a beetle antennae search optimized support vector machine[J]. Power System Protection and Control, 2020, 48(20): 90-96.
- [18] 田凤兰, 张恩泽, 潘思蓉, 等. 基于特征量优选与ICA-SVM的变压器故障诊断模型[J]. 电力系统保护与控制, 2019, 47(17): 163-170.
- TIAN Fenglan, ZHANG Enze, PAN Sirong, et al. Fault diagnosis model of power transformers based on feature quantity optimization and ICA-SVM[J]. Power System Protection and Control, 2019, 47(17): 163-170.
- [19] 汪可, 李金忠, 张书琦, 等. 变压器故障诊断用油中溶解气体新特征参量[J]. 中国电机工程学报, 2016, 36(23): 6570-6578, 6625.
- WANG Ke, LI Jinzhong, ZHANG Shuqi, et al. A new characteristic parameter of dissolved gas in oil for transformer fault diagnosis[J]. Proceedings of the CSEE, 2016, 36(23): 6570-6578, 6625.
- [20] 邱海枫, 苏宁, 田松林. 改进支持向量机在电力变压器故障诊断中的应用研究[J]. 电测与仪表, 2022, 59(11): 48-53.
- QIU Haifeng, SU Ning, TIAN Songlin. Research on the application of improved support vector machine in power transformer fault diagnosis[J]. Electrical Measurement & Instrumentation, 2022, 59(11): 48-53.
- [21] 陈欢, 彭辉, 舒乃秋, 等. 基于蝙蝠算法优化最小二乘双支持向量机的变压器故障诊断[J]. 高电压技术, 2018, 44(11): 3664-3671.
- CHEN Huan, PENG Hui, SHU Naiqiu, et al. Fault diagnosis of transformer based on LS-TSVM optimized by bat algorithm[J]. High Voltage Engineering, 2018, 44(11): 3664-3671.
- [22] LI J, ZHANG Q, WANG K, et al. Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2016, 23(2): 1198-1206.
- [23] KIM S W, KIM S J, SEO H D, et al. New methods of DGA diagnosis using IEC TC 10 and related databases part 1: application of gas-ratio combinations[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2013, 20(2): 685-690.
- [24] DEGHANI M, HUBALOVSKY S, TROJOVSKY P. Northern goshawk optimization: a new swarm-based algorithm for solving optimization problems[J]. IEEE Access, 2021, 9: 162059-162080.
- [25] 闫少强, 杨萍, 朱东林, 等. 基于佳点集的改进麻雀搜索算法[J]. 北京航空航天大学学报, 2023, 49(10): 2790-2798.
- YAN Shaoqiang, YANG Ping, ZHU Donglin, et al. Improved sparrow search algorithm based on good point set[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(10): 2790-2798.
- [26] LONG W, JIAO J J, LIANG X M, et al. A random opposition-based learning grey wolf optimizer[J]. IEEE Access, 2019, 7: 113810-113825.
- [27] 吴广宁, 袁海满, 宋臻杰, 等. 基于粗糙集与多类支持向量机的电力变压器故障诊断[J]. 高电压技术, 2017, 43(11): 3668-3674.
- WU Guangning, YUAN Haiman, SONG Zhenjie, et al. Fault diagnosis for power transformer based on rough set and multi-class support vector machine[J]. High Voltage Engineering, 2017, 43(11): 3668-3674.
- [28] 康佳宇, 张沈习, 张庆平, 等. 基于ANOVA和BO-SVM的变压器故障诊断方法[J]. 高电压技术, 2023, 49(5): 1882-1891.
- KANG Jiayu, ZHANG Shenxi, ZHANG Qingping, et al. Fault diagnosis method of transformer based on ANOVA and BO-SVM[J]. High Voltage Engineering, 2023, 49(5): 1882-1891.
- [29] 田晓飞. 基于改进蝙蝠算法优化支持向量机的变压器故障诊断研究[D]. 成都: 西华大学, 2019.
- TIAN Xiaofei. Research on transformer fault diagnosis based on improved bat algorithm optimized support vector machine[D]. Chengdu: Xihua University, 2019.
- [30] 孙琛. 基于机器学习算法的变压器故障诊断[D]. 上海: 上海交通大学, 2019.
- SUN Chen. Transformer fault diagnosis based on machine learning algorithms[D]. Shanghai: Shanghai Jiao Tong University, 2019.

收稿日期: 2023-07-21; 修回日期: 2023-10-05

作者简介:

包金山(1996—), 男, 硕士研究生, 研究方向为电气设备故障诊断; E-mail: 193084634@qq.com

胡克林(1991—), 男, 通信作者, 博士, 特聘教授, 硕士生导师, 研究方向为电气设备状态监测与诊断、新型微纳传感技术。E-mail: klhu@gzu.edu.cn

(编辑 魏小丽)