

DOI: 10.13382/j.jemi.B2206162

多策略改进 MPA 算法与 HKELM 的变压器故障辨识*

谢国民¹ 刘东阳¹ 刘 明²
(1. 辽宁工程技术大学电气与控制工程学院 葫芦岛 125105; 2. 朝阳师范高等专科学校数学计算机系 朝阳 122000)

摘 要:为解决目前变压器故障诊断精度低的问题,提出一种多策略改进海洋捕食者算法(MPA)与混合核极限学习机(HKELM)的变压器故障辨识方法。首先通过核主成分分析法(KPCA)对高维线性不可分的变压器故障数据进行降维,获取特征支持数据;然后通过伯努利混沌映射、改进阶段转换判据、最佳候选者等策略综合改进 MPA,加强全局开发能力;最后使用改进的 IMPA 算法对 HKELM 的参数寻优,构建变压器故障诊断模型。为验证模型有效性,分析比较常用算法优化的 HKELM 的 4 种变压器故障诊断模型。其中 IMPA-HKELM 的诊断精度为 94.7%,相比于另外 3 种基础算法优化的模型,诊断精度分别提升了 5.4%、8%、10.7%。结果表明,提出模型有效提升了故障诊断的分类性能,并实现了较高的故障诊断精度。

关键词: 故障诊断;油浸式变压器;伯努利混沌映射;混合核极限学习机;核主成分分析

中图分类号: TN06 **文献标识码:** A **国家标准学科分类代码:** 470.40

Transformer fault identification based on multi-strategy improved MPA algorithm and HKELM

Xie Guomin¹ Liu Dongyang¹ Liu Ming²
(1. Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China;
2. Department of Mathematics and Computer Science, Chaoyang Teachers College, Chaoyang 122000, China)

Abstract: For the purpose of tackling the difficulties of the low accuracy of transformer fault diagnosis, a transformer fault identification method based on multi-strategy improved ocean predator algorithm (MPA) and hybrid kernel extreme learning machine (HKELM) has been proposed. Firstly, kernel principal component analysis (KPCA) is applicable to decrease the dimension of high-dimensional linear inseparable transformer fault data and it is also used to obtain feature support data. Then, the MPA is comprehensively improved through strategies such as Bernoulli chaotic mapping, improved stage transition criterion, and best candidate to strengthen the global development ability. Finally, the improved IMPA algorithm is used to optimize the parameters of HKELM and construct the transformer fault diagnosis model. Aiming to validate the validity of the model, four transformer fault diagnosis models of HKELM optimized by common algorithms are analyzed and compared. The diagnostic accuracy of IMPA-HKELM is 94.7%, compared with the other three basic algorithms, the diagnostic accuracy is improved by 5.4%, 8% and 10.7% respectively. The results show that the proposed model effectively improves the classification performance of fault diagnosis and achieves higher fault diagnosis accuracy.

Keywords: fault diagnosis; oil-immersed type transformer; Bernoulli chaotic map; hybrid kernel extreme learning machine; kernel principal component analysis

0 引 言

电力变压器作为输变电系统的枢纽设备,其运行状态决定着电力系统的可靠性^[1]。若不能及时对故障做出

诊断并处理,就会造成无法预估的经济损失^[2]。因此快速准确的对变压器故障做出诊断意义重大。目前普遍运用于电力系统的是油浸式电力变压器,在其发生故障时绝缘油会分解出多种碳氢气体,因此油中溶解气体分析法(dissolved gas analysis, DGA)成为主要的变压器诊断

方法^[3]。基于 DGA 法的传统诊断方式依托大量经验,且要在气体浓度高的情况下才会效果明显,应用有一定的局限性^[4]。随着 AI 技术的发展,例如支持向量机、BP 神经网络、贝叶斯网络、极限学习机等^[5-8]的机器学习方法广泛的运用到变压器故障诊断中。文献[9]通过 ISSA 算法来确定最优的支持向量机参数,相比改进前诊断精度有所提高;神经网络是一种仿生模型,有较强的鲁棒性和容错能力,多层前馈型神经网络在故障诊断的大数据样本下性能更好。结合 BP 算法的 BP 神经网络具有优化计算和识别分类能力,文献[10]采用堆叠多个残差网络模块的方式加深 BP 神经网络的深度,提高诊断性能;文献[11]依托智能算法优化贝叶斯网络进行诊断。

虽然上述方法提高了诊断精度,但神经网络需要的数据样本多,模型较为复杂;贝叶斯网络需制定的规则繁多;支持向量机虽算法简单,但分类精度有待提高。极限学习机^[12](extreme learning machine, ELM)是一种单隐含层前馈神经网络,因其隐含层无需迭代、学习速度快且泛化能力强的优点,更适用于故障预测。混合核极限学习机(HKELM)则是将核函数线性结合为混合函数后引入 ELM 中,实现特征向高纬度的映射,具有良好的分类性能。HKELM 的诊断精度与正则化系数 C 、核参数、混合权重和原始数据维度密切相关。因此如何有效降低数据维度,搜索最优正则化系数、混合权重和核参数是研究关键^[13-14]。智能优化算法被广泛运用于寻优搜索中,能有效解决机器学习最优参数问题,文献[15]运用人工鱼群优化算法对极限学习机的权值及阈值进行搜索;文献[16]运用 QPSO 算法优化核极限学习机,提高分类性能和速率;文献[17]采用粒子群算法寻优混合核函数的参数优化极限学习机,提升学习机性能。虽然智能算法的结合提升了故障诊断的精度,但基础智能算法搜索不平衡、易陷入局部最优的缺点,导致对学习机各参数的寻优速度慢,离最优参数的偏差较大。使得寻优参数后的学习机疲于处理高维复杂的变压器故障数据,严重影响变压器故障诊断精度。

对此,本文提出一种改进的海洋捕食者算法(improved marine predators algorithm, IMPA)优化 HKELM 的变压器故障诊断方法。首先通过核主成分分析法^[18](kernel principal component analysis, KPCA)对非线性的高维变压器故障数据进行降维,并保留主要特征。其次改进海洋捕食者算法,加强搜索能力,提升寻优精度。最后通过 IMPA 寻优最佳混合核极限学习机参数,构成诊断模型。并与基于 PSO、GWO、MPA 优化算法的 HKELM 诊断模型进行比较实验,验证本文提出的模型分类精度与效率更优。

1 变压器故障数据特征提取与降维

当变压器发生故障,绝缘油会裂解产生 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 和 C_2H_6 这 5 种特征气体^[19]。在不同故障类别时,5 种气体的浓度和比例关系各不相同。为完整体现变压器故障特征,选取 5 种气体的浓度和不同气体之间比值关系作为特征数据^[20],如表 1 所示。

表 1 变压器故障特征

Table 1 Fault characteristics of transformer				
20 维特征				
H_2	CH_4	C_2H_2	C_2H_4	C_2H_6
$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{H_2}$	$\frac{C_2H_4}{H_2}$	$\frac{C_2H_6}{H_2}$	$\frac{C_2H_2}{CH_4}$
$\frac{C_2H_4}{CH_4}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{C_2H_6}$	$\frac{C_2H_2}{C_2H_6}$	$\frac{C_2H_4}{C_2H_6}$
$\frac{C_2H_4}{C_1 + C_2}$	$\frac{CH_4}{C_1 + C_2}$	$\frac{C_2H_6}{C_1 + C_2}$	$\frac{C_2H_2}{C_1 + C_2}$	$\frac{H_2}{C_1 + C_2}$

其中, $C_1 + C_2 = CH_4 + C_2H_2 + C_2H_4 + C_2H_6$ 。

虽然选取 20 维的故障数据能够最大程度反应变压器故障特征,但是高维数据线性不可分的特点会降低 HKELM 的分类精度。在此采用核主成分分析法(KPCA)进行数据降维处理。KPCA 是通过核函数,将数据投射至高维度,变为线性可分数据,再通过主成分分析(PCA)搜索其线性子空间投射实现降维的方法^[21]。本文选取国家电网某省电网公司提供的变压器故障数据,按故障类别随机选取 250 组故障数据进行实验。

由变压器故障数据组成矩阵 $X_{N \times m}$,表示有 N 组 m 维数据,利用径向基(RBF)核函数计算其核矩阵 K ,将矩阵 K 中心化后得到 $\tilde{K}$,具体如式(1)~(3)所示:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (1)$$

$$K = XTX = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \cdots & k(x_N, x_N) \end{bmatrix} \quad (2)$$

$$\tilde{K} = K - L_N K - K L_N + L_N^T K L_N \quad (3)$$

其中, σ 为高斯核参数, L_N 表示其中所有元素均由 $1/N$ 的 $N \times N$ 矩阵构成,接下来用式(4)计算 $\tilde{K}$ 的特征值 λ 和特征向量 v ,进行特征值分解,最后用式(5)计算不同特征值下的故障诊断贡献率并按降序排列,进行主成分选取和降维。

$$\tilde{K}v = (XX^T)v = \lambda v \quad (4)$$

$$(\lambda_1 + \lambda_2 + \cdots + \lambda_s) / (\lambda_1 + \lambda_2 + \cdots + \lambda_n) \geq 85\% \quad (5)$$

贡献率计算结果如图 1 所示,贡献率之和达到 85% 以上的主成分不仅能保留数据的大部分特征,还能去掉冗余特征。图 1 可知前 10 维主成分贡献率之和已达到 85%,满足条件且有效降低了维度。因此选取前 10 个主元为降维后的维度,其贡献率统计如表 2 所示,贡献率之和为 85.100 6%。

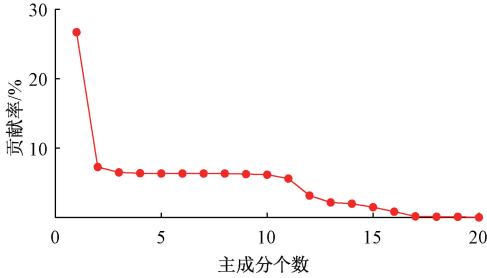


图 1 主成分贡献率折线图

Fig. 1 Principal component contribution rate curve diagram

表 2 贡献率统计表

Table 2 Contribution Rate Statistics

主成分编号	第 1 主元	第 2 主元	第 3 主元	第 4 主元	第 5 主元
贡献率/%	26.787	7.364 3	6.587 4	6.463 8	6.357 2
主成分编号	第 6 主元	第 7 主元	第 8 主元	第 9 主元	第 10 主元
贡献率/%	6.3572	6.3572	6.3572	6.2713	6.198

降维后的降维矩阵表达式如式(6)所示。

$$Y = \tilde{K}^T v = \tilde{K}^T \left[\frac{1}{\sqrt{\lambda_1}} v_1, \frac{1}{\sqrt{\lambda_2}} v_2, \dots, \frac{1}{\sqrt{\lambda_{10}}} v_{10} \right] \quad (6)$$

2 IMPA-HKELM 模型

变压器故障依据 DGA 方法能够构建多维的故障特征,它呈现出高维度、分类难度大的特点。KPCA 处理虽可以降低维度,但若要进一步准确辨识故障类别,需要尽可能最优的分类器参数,对算法的寻优能力要求很高。因此选择将 MPA 算法进行多策略改良,用于选取最优正则化系数和核参数,相比于文献[15-17]应用基础算法寻优更能挖掘变压器 DGA 的数据特征,提升 HKELM 的分类效果。

2.1 海洋捕食者算法

海洋捕食者算法依据海洋中优胜劣汰的原则演化而来,算法初期由式(7)进行环境初始化,保证初始解在搜索范围中均匀分布^[22]。

$$P_0 = lb + r \cdot (ub - lb) \quad (7)$$

式中: ub 、 lb 为解的上、下限, r 是 0~1 内均匀分布的随机数, P_0 为初始猎物矩阵。其次要确定精英矩阵 E , 其由适应度最优的个体“顶级捕食者”构成,即最优解位置。其次要确认猎物矩阵 P , 海洋捕食者算法的寻优主要依

靠以上两个矩阵。

由于 MPA 中捕食者和猎物间的速度差,搜索阶段分为高速比 ($t < t_{\max}/3$)、等速比 ($t_{\max}/3 < t < 2t_{\max}/3$) 和低速比 ($t > 2t_{\max}/3$) 这 3 个阶段,其中 $t_{\max}$ 为最大迭代次数。

1) 阶段 1 为高速比阶段,捕食者速度远远小于猎物,猎物以布朗运动移动位置,捕食者在此阶段进行算法的初期探索,如式(8)所示:

$$P_{i+1} = P_i + p \cdot R \otimes s_i \quad (8)$$

$$s_i = R_B \otimes (E_i - R_B \otimes P_i), i = 1, \dots, n$$

式中: S_i 表示猎物 P 和捕食者 E 位置变化时的步长, R_B 为示布朗运动的随机数向量,随机数由正态分布产生; p 为常数 0.5; R 向量为 0~1 间的随机数。 P_i 和 E_i 表示迭代到第 i 个猎物和第 i 个顶级捕食者。

2) 阶段 2 为等速比阶段,此时猎物与捕食者速度相同,对猎物分为两部分展开搜索,一部分进行布朗运动搜索,另一部分进行 Lévy 运动开发搜索,如式(9)~(11)所示:

$$P_{i+1} = \begin{cases} P_i + p \cdot R \otimes s_i, i = 1, \dots, n/2 \\ P_i + p \cdot CF \otimes s_i, i = n/2, \dots, n \end{cases} \quad (9)$$

$$s_i = \begin{cases} R_L \otimes (E_i - R_L \otimes P_i), i = 1, \dots, n/2 \\ R_B \otimes (R_B \otimes E_i - P_i), i = n/2, \dots, n \end{cases} \quad (10)$$

$$CF = (1 - t/t_{\max})^{2a/t_{\max}} \quad (11)$$

其中, R_L 是服从 Lévy 分布的随机向量, p 为常数 0.5, CF 是自适应参数。

3) 阶段 3 为低速比阶段,这时工作倾向于局部开发,捕食者进行 Lévy 运动,猎物模仿捕食者 Lévy 轨迹进行规避,如式(12)、(13)所示:

$$P_{i+1} = E_i + p \cdot CF \otimes s_i \quad (12)$$

$$s_i = R_L \otimes (R_L \otimes E_i - P_i), i = 1, \dots, n \quad (13)$$

另外,海洋中掺杂各种变化,会影响个体行为, FADs 即为最普遍的影响因素,可看作局部最优。 FADs 影响式(14)能帮助摆脱局部最优。

$$P_{i+1} = \begin{cases} P_i + CF \cdot S_0 \otimes U, r \leq P_F \\ P_i + [P_F(1 - r) + r](P_{r1} - P_{r2}), r \geq P_F \end{cases} \quad (14)$$

其中, P_F 是 0.2 的固定值, U 为一个二进制的数组,由随机二进制向量构成, r 为 0~1 之间随机数, P_{r1} 、 P_{r2} 为随机挑选的猎物个体。

综上,MPA 算法兼顾全局与局部的开发,且自身的 FADs 能一定程度摆脱局部最优,选择 MPA 能够更高效的寻优所需参数,准确快速的识别故障数据特征,达到变压器故障辨识的目的。但因其自身随机因子、阶段间机械转换的影响,会导致优化的不均衡。并且因后期猎物的模仿行为,自身的 FADs 影响公式不足以应对局部最优问题。在此对其进行规则和策略上的优化。

2.2 MPA 算法改进

1) Bernoulli 映射

Bernoulli 映射是一种普遍应用的混沌映射,主要用于优化初期提供随机的混沌序列^[23],传统的随机数生成器初始化 MPA 时遍历性不足,导致影响后续搜索;引入 Bernoulli 映射可获得更好随机性和均匀度的混沌序列,提升寻优效果,Bernoulli 映射表达式如式(15)所示:

$$Z_{k+1} = \begin{cases} Z_k/1 + \rho & Z_k \in (0, 1 - \rho] \\ Z_k - 1 + \rho/\rho & Z_k \in (1 - \rho, 1) \end{cases} \quad (15)$$

其中, Z_k 是第 k 代序列时的值, ρ 是控制量; Z_k 和 ρ 两个变量对混沌序列生成有决定作用,根据实验得知 ρ 在 0.5 左右能保持较好的遍历性,设置初始值是分布在 0~1 的随机数。

2) 改进 MPA 阶段转换判据

在 MPA 算法中,优化过程的 3 个阶段分别包含总迭代次数的 1/3。显然,通过改变迭代的总数,每一阶段的迭代次数也会改变。但在相同的优化中增加总迭代次数,无法保证结果的好坏。如本应迭代 50 次后满足后期局部开发的条件,但由于总迭代次数设置过大,每一阶段迭代次数分量也增加,导致 50 次迭代后依旧停留在初期搜索,严重影响迭代速度。因此,在优化过程中引入式(16)、(17)函数和随机数 r 更改 MPA 更新规则。

$$F_1(t) = (t_{\max} - t)/t_{\max} \quad (16)$$

$$F_2(t) = 1/e^{(t/100)} \times F_1 \quad (17)$$

其中, t 表示当前迭代次数, $t_{\max}$ 表示总迭代次数。在每次迭代中, F_1 和 F_2 的值都使用上述公式确定。 F_1 和 F_2 函数的图像如图 2 所示。

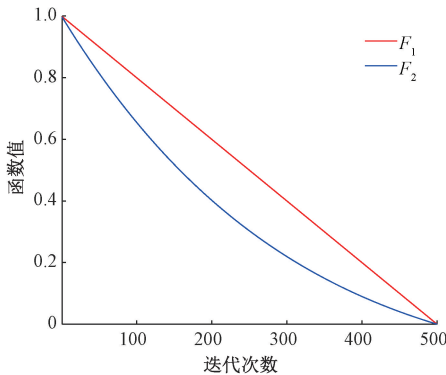


图 2 转换判据函数图像

Fig. 2 Transformation criterion function image

由图可知,纵坐标为 0~1,对应随机数 r 的值,横坐标为迭代次数。随着迭代次数的增加, F_1 从 1 线性递减到 0, F_2 从 1 指数递减到 0。在算法迭代初期,随机数 r ($0 < r < 1$) 的值更容易小于 F_2 , IMPA 执行优化过程中的第 1 部分,使算法具有较强的全局探索能力。在算法迭代

后期,随机数 r 的值更容易大于 F_1 , IMPA 执行优化过程中的第 3 部分,增加了算法的开发能力。当 r 小于 F_1 大于 F_2 时, IMPA 执行优化过程中的第 2 部分,用于平衡算法的开发和勘探策略。此时逐渐由勘探策略转向开发策略。

改进后的 MPA 迭代判据摆脱了机械式的转换;加强全局与局部寻优的适应性,平衡了 MPA 搜索;灵活的判据使得前后期的搜索性能最大化,提升了 MPA 寻优能力。

3) 最佳候选者策略

在猎物的引导运动中,无论是布朗运动还是 Lévy 飞行,都只考虑了一个位置。这个新位置的确定取决于步长,而步长是随机取决于所选择的运动和捕食者和猎物之间的差距。其中,适应度值并没有被考虑在内。因此,本文提出一种新的改进机制。在引导运动中为猎物选取几个新的最佳候选位置,它们的适应度值被考虑到猎物下一个位置的更新中。

选取规则为:取猎物中适应度排名前 10% 作为最佳候选解,这些个体不但包含最优解的特征,并且求其反向解,能够剔除局部最优造成的不利影响。

选取的最佳候选解利用凸透镜反向学习策略^[24]求其反向解。凸透镜成像原理如图 3 所示,反向解公式如式(18)所示:

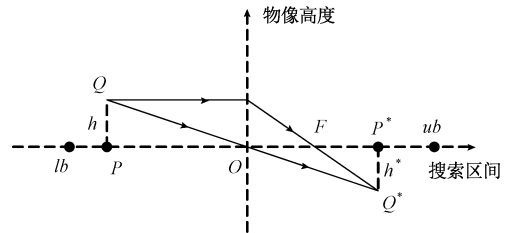


图 3 凸透镜成像原理

Fig. 3 Convex lens imaging principle diagram

$$P^* = \frac{ub - lb}{2} + \frac{ub - lb}{2\eta} - \frac{P}{\eta} \quad (18)$$

在搜索区间 $[lb, ub]$ 中求出反向解 P^* , h/h^* 记为 η , 它是反应物像之间比例的收缩因子,多为固定值,限制了搜索能力。在此引入双曲正切函数,该函数在 2、3 象限随自变量增大从 0 增至 1,改进后初期 η 值大,能更好的全局搜索,在后期 η 值小,能锁定局部精准搜索反向解。如式(19)所示:

$$\eta = d_1 - (d_1 - d_2) \tanh(\pi t/t_{\max}) \quad (19)$$

其中, d_1, d_2 为 η 的最大最小值。

最后采用贪婪选择策略对比更新前后的猎物的适应度值,保留适应度高的个体。最优候选者策略能锁定最优解范围,并且反向解更新后提高种群多样性,有效避免 MPA 陷入局部最优解。

2.3 改进后 IMPA 算法性能测试

传统 MPA 算法三阶段转换死板,十分影响算法性能;GWO 算法易陷入过早收敛的现象,复杂情况精度低;PSO 算法全局搜索能力较弱,导致种群多样性低,无法跳脱粒子聚集点。IMPA 初始化采用混沌映射提高种群多样性,其转换判据和最优解的优化利于摆脱局部最优,相比前 3 种算法搜索性能更佳。

为验证 IMPA 算法的性能,选取标准测试函数对其进行测试,并选择 MPA、GWO、PSO 这 3 种优化算法进行性能对比。选取测试函数如表 3 所示,测试函数 F_1 是标准的单峰函数,仅有一个极小点,适用于测试算法的收敛速度; F_2 则是多峰函数,在搜索空间中遍布许多极小点,可用来测试算法的全局寻优能力。

表 3 测试函数
Table 3 Test function

测试函数	搜索范围	维度
$f_1 = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	$[-10, 10]$	30
$f_2 = -20\exp\left(-0.2\sqrt{(1/n)\sum_{i=1}^n x_i^2}\right) - \exp\left[(1/n)\sum_{i=1}^n \cos(2\pi x_i)\right] + 20 + e$	$[-100, 100]$	30

由图 4、5 可知,对于 F_1 测试函数的测试结果,IMPA 的收敛速度快于其他优化算法,展现了优异的收敛速度。而 F_2 测试函数中,IMPA 能够搜索到最佳适应度值,并且迭代次数更少。这是由于 IMPA 算法中阶段转换判据的改良,和最佳候选者策略跳脱局部最优的作用,加强了收敛速度和寻优能力。因此通过多策略改进后的 IMPA 算法,收敛速度和跳脱局部最优的能力都更突出,寻优能力更强。

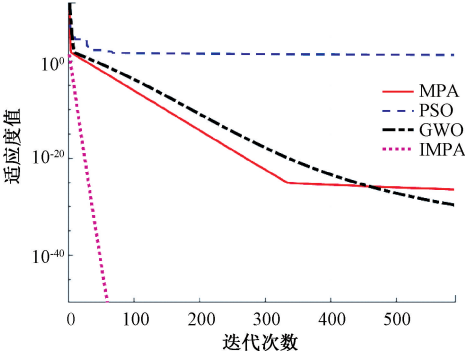


图 4 F_1 函数测试结果

Fig. 4 F_1 Function test results

2.4 混合核极限学习机

目前变压器故障辨识主要依托学习机进行,而变压

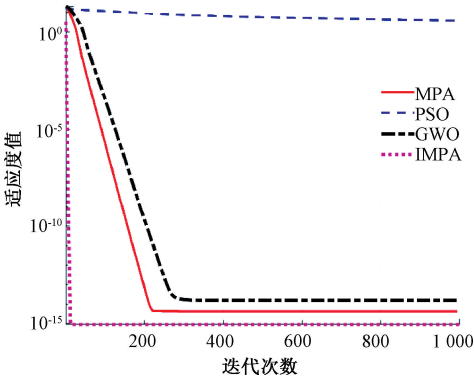


图 5 F_2 函数测试结果

Fig. 5 F_2 Function test results

器的故障数据复杂、数量多且部分数据可靠度低,这使故障辨识难度加大。因此除了学习机参数要选择最优,自身分类性能也是关键因素。本文选择学习速度快、泛化能力强的核极限学习机 (KELM),它是一种单隐含层前馈神经网络,通过引入核函数改善了原本迭代次数多速度慢的缺点,减少运算量并且提高效率,具有很好的非线性回归和分类效果,输出模型如式 (20) 所示:

$$f(x) = h(x)\beta = H\beta$$
 (20)

其中, x 为样本数据; $f(x)$ 是输出; $h(x)$ 是隐含层输入; H 表示特征映射矩阵,由核函数映射样本数据得来; β 则是将隐含层数据输出至输出层的向量。

$$\beta = H^T(HH^T + I/C)^{-1}T$$
 (21)

其中, C 为正则化参数; I 为单位矩阵; T 为训练集目标向量。核极限学习机的矩阵模型如式 (22) 所示:

$$\begin{cases} \Omega = HH^T \\ \Omega_{i,j} = h(x_i)h(x_j) = K(x_i, x_j) \end{cases}$$
 (22)

用核矩阵 Ω 替代极限学习机的 HH^T 矩阵, $K(x_i, x_j)$ 为核函数矩阵,学习机模型如图 6 所示。

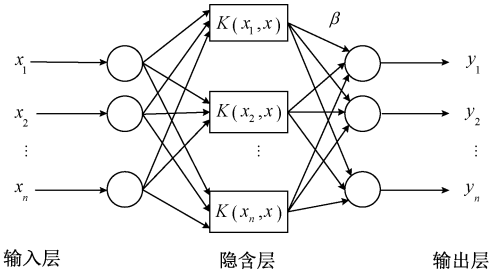


图 6 KELM 模型结构拓扑图

Fig. 6 KELM model structure topology diagram

核函数将输入的数据映射至高维度的隐含层空间,其中 $K(x_i, x_j)$ 为核函数矩阵, $x_1, x_2, \dots, x_n$ 为训练集的输入向量。因此可得输出模型如式 (23) 所示:

$$f(x)=\begin{bmatrix}K(x,x_1)\\\vdots\\K(x,x_N)\end{bmatrix}^T\left(\frac{I_0}{C}+\boldsymbol{\Omega}\right)^{-1}T=\begin{bmatrix}K(x,x_1)\\\vdots\\K(x,x_N)\end{bmatrix}^T\boldsymbol{\beta}\quad(23)$$

由上式可知核函数的确定决定预测结果,而单一核函数搜索有限制。*Poly* 核函数为全局核函数, *RBF* 核函数为局部核函数,将两个核函数运用线性结合的方式,构成新的混核函数,使 *KELM* 拥有全局和局部两方面的优秀分类性能, *Poly* 与 *RBF* 函数式如式(24)所示:

$$K_{Poly}(x_i,x_j)=(x,x_i+c_1)^d$$
$$K_{RBF}(x_i,x_j)=\exp(-\|x_i-x_j\|^2/\sigma^2)$$
(24)

其中, σ 、 c_1 、 d 为 *Poly* 核函数和 *RBF* 核函数的核参数,混核函数由二者线性结合得到:

$$K_H(x_i,x_j)=s_1K_{Poly}(x_i,x_j)+s_2K_{RBF}(x_i,x_j)$$
$$s_1+s_2=1$$
(25)

将混核函数代入输出函数可得 *HKELM* 模型,混合核函数提升了模型的分类速度和准确度,其中参数 C 、 σ 、 c_1 、 d 、 s_1 由对 *HKELM* 的影响很大,由 *IMPA* 对参数进行寻优,提升分类性能。

3 基于 IMPA-HKELM 变压器故障诊断模型

变压器故障数据复杂,传统变压器故障诊断方法有局限性并且诊断精度低。因此本文提出基于 *KPCA* 降维后故障数据建立的 *IMPA-HKELM* 诊断模型,通过 *IMPA* 算法对 *HKELM* 的正则化系数 C 、混合核参数、混合权重进行寻优,获取最优分类模型,寻优步骤如下:

步骤 1) 获取降维后变压器故障数据,划分为训练集与测试集。

步骤 2) 初始化 *IMPA* 参数:设定种群数量为 50;维度为 5;最大迭代次数为 100。

步骤 3) 设置 *HKELM* 各参数,取正则化系数 C 、核参数和混合权重 σ 、 c_1 、 d 、 s_1 。

步骤 4) 计算捕食者个体适应度,按大小排列,适应度最高的个体作为当前顶级捕食者(最优个体),即最优解位置。

步骤 5) 按照 *IMPA* 更新公式对海洋捕食者个体进行位置更新。

步骤 6) 计算更新后捕食者的适应度值,并于更新前个体比较,适应度高则更新为新最优个体,否则保留更新前最优个体位置。

步骤 7) 判断是否达到最大迭代次数,若未达到则返回步骤 3);若达到最大迭代次数,则输出当前最优解,即 C 、 σ 、 c_1 、 d 、 s_1 的最优值,寻优结束。

故障诊断流程如图 7 所示。

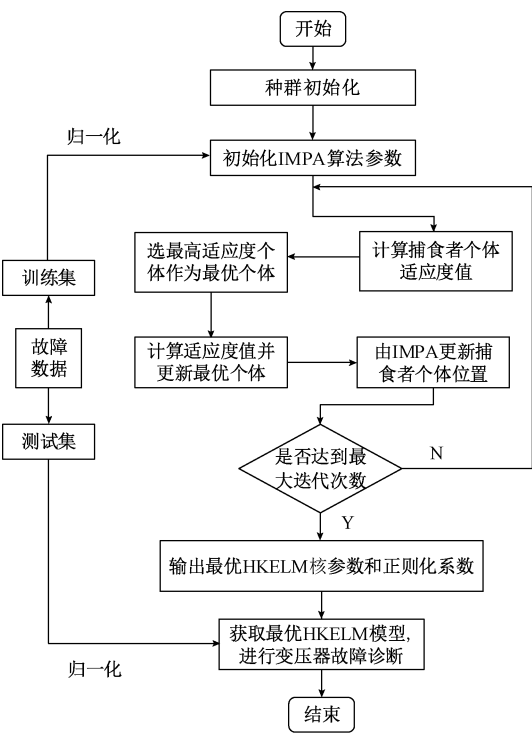


图 7 故障诊断流程

Fig. 7 Fault diagnosis flow chart

4 故障诊断实验分析

将从国家电网某电网中选取的 250 组变压器故障数据按照 7 : 3 比例划分为训练集和测试集,每种故障类型下的训练集和测试集样本数相同,对故障状态编号如表 4 所示。

表 4 数据分类
Table 4 Data classification

	正常	中低温 过热	高温 过热	低能 放电	高能 放电
状态编号	1	2	3	4	5
训练集数目	35	35	35	35	35
测试集数目	15	15	15	15	15

4.1 不同优化算法的 HKELM 分类结果比较

首先使用 *KPCA* 降维前的测试集数据进行实验。选择 *IMPA*、*MPA*、*GWO*、*PSO* 这 4 种智能算法优化 *HKELM*,构成变压器故障诊断模型。再将测试集数据归一化后输入诊断模型,实验结果如图 8~11 所示。

在 *KPCA* 降维前的变压器数据下, *IMPA* 相比于 *MPA* 提升了较高的分类精度,尤其提高了高温过热和中低温过热两类故障的准确率,并且 *IMPA* 的分类精度也远远优于 *GWO*、*PSO* 算法的分类模型。

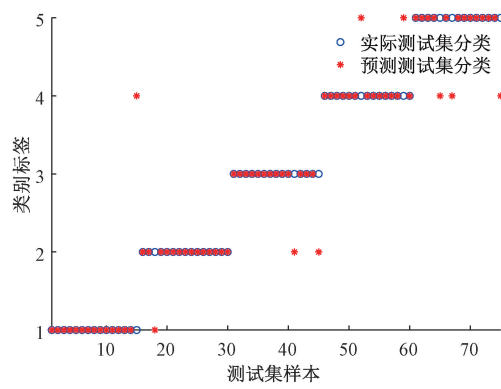


图 8 KPCA 降维前 IMPA 优化 HKELM 分类图
Fig. 8 IMPA optimized HKELM classification graph before KPCA dimensionality reduction

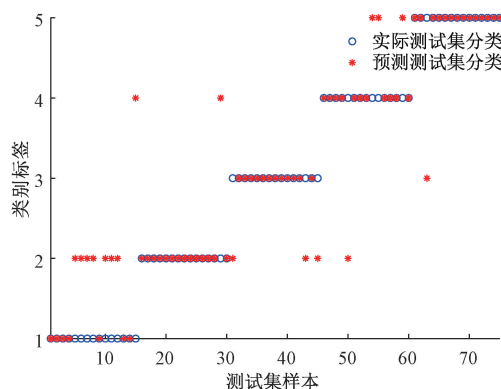


图 11 KPCA 降维前 PSO 优化 HKELM 分类图
Fig. 11 PSO optimized HKELM classification graph before KPCA dimensionality reduction

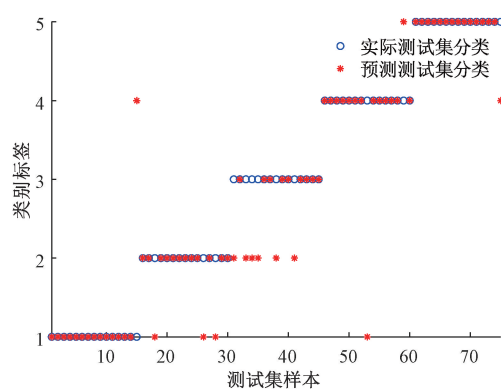


图 9 KPCA 降维前 MPA 优化 HKELM 分类图
Fig. 9 MPA optimized HKELM classification graph before KPCA dimensionality reduction

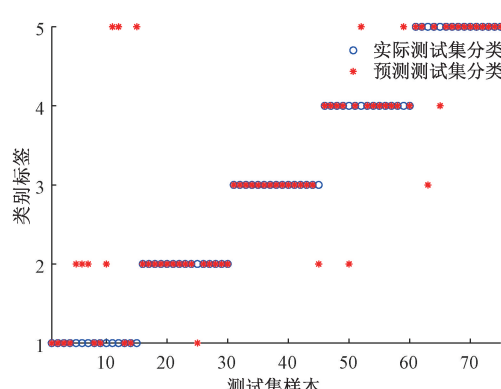


图 10 KPCA 降维前 GWO 优化 HKELM 分类图
Fig. 10 GWO optimized HKELM classification graph before KPCA dimensionality reduction

应用 4 种算法优化降维后的 KPCA-HKELM 模型,对核参数、混合权重 s_1 和 C 进行寻优,得到分类结果如图 12~15 所示。

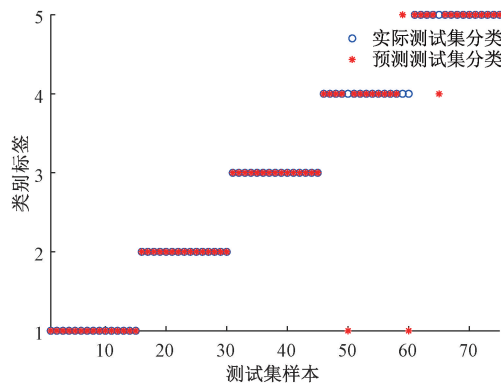


图 12 KPCA 降维后 IMPA 优化 HKELM 分类图
Fig. 12 IMPA optimized HKELM classification graph after KPCA dimensionality reduction

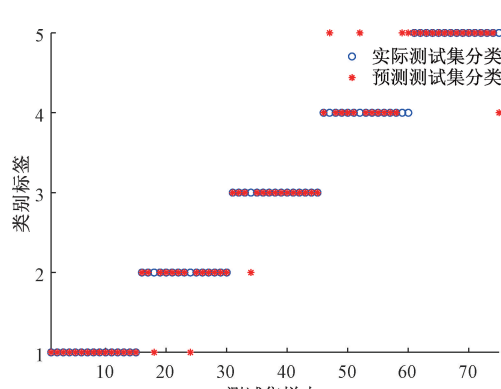


图 13 KPCA 降维后 MPA 优化 HKELM 分类图
Fig. 13 MPA optimized HKELM classification graph after KPCA dimensionality reduction

但是由于高维度的故障数据,导致 IMPA-HKELM 模型的分类并未达到理想状态,5 种故障类别下的准确率均有提升空间。

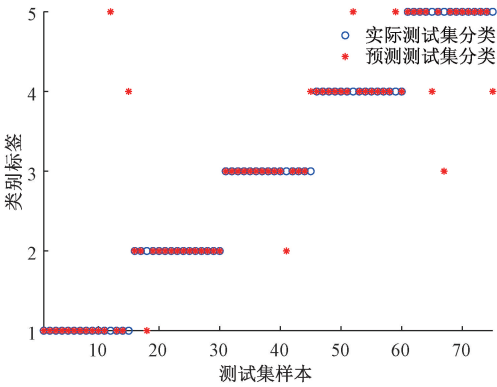


图 14 KPCA 降维后 GWO 优化 HKELM 分类图
Fig. 14 GWO optimized HKELM classification graph after KPCA dimensionality reduction

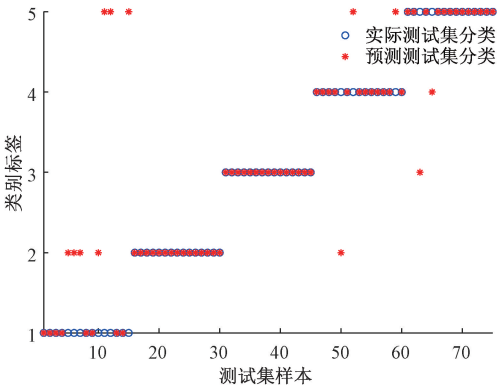


图 15 KPCA 降维后 PSO 优化 HKELM 分类图
Fig. 15 PSO optimized HKELM classification graph after KPCA dimensionality reduction

由图 12~15 可以看出,KPCA 能在保留数据大部分特征的情况下,降低数据维度。不同优化算法的分类模型应用 KPCA 降维后数据,分类准确率均有较大提升。KPCA-IMPA-HKELM 模型的精度更达到 94.7%,其中正常、中低温过热和高温过热故障类型无错误诊断。因此用 KPCA 法处理故障数据对提升诊断准确率十分有效,4 种分类模型诊断准确率对比如表 5 所示。

表 5 诊断准确率对比分析

	故障诊断准确率/%			
	IMPA	MPA	GWO	PSO
KPCA 降维前	88.0	82.7	81.3	77.3
KPCA 降维后	94.7	89.3	86.7	84.0

4.2 不同学习机模型分类结果比较

HKELM 作为一种前馈型神经网络,在 4.1 节中结合 IMPA 构建的模型具有良好的分类性能。为验证 HKELM 的故障辨识优势,选取两种前馈型神经网络:BP 神经网络、

极限学习机(ELM)和一种反馈型神经网络:ELMAN 神经网络,结合 IMPA 算法进行比较,诊断结果如表 6 所示。

表 6 诊断准确率对比分析

	故障诊断准确率/%			
	IMPA-HKELM	IMPA-BP	IMPA-ELM	IMPA-ELMAN
KPCA 降维前	88.0	82.7	84.0	69.3
KPCA 降维后	94.7	92.0	93.3	86.7

由结果可知,前馈型神经网络的诊断准确率均高于反馈型神经网络,且本文的 IMPA-HKELM 模型诊断准确率比其他前馈神经网络模型更高,因此本文的诊断模型能够实现有效准确的变压器故障辨识。

4.3 不同诊断模型的多指标结果比较

为进一步验证本文所构建的诊断模型性能,除诊断准确率外,加入召回率和精确率两个指标进行验证。在多分类问题中,两个指标需要先计算每一个故障类型的召回率和精确率,再取其平均值,得到宏召回率和宏精确率,因本文是均衡数据集,宏召回率等于准确率。表 7 为 4 种算法优化 HKELM 的对比结果,表 8 为不同网络模型的对比结果。由结果可以看出,IMPA-HKELM 模型在准确率、精确率、召回率 3 方面均有良好结果,因此可以对变压器故障做出准确的辨识。

表 7 4 种算法优化 HKELM 的对比结果

Table 7 Comparison results of four algorithms for optimizing HKELM (%)			
	准确率	宏召回率	宏精确率
IMPA-HKELM	88.0	88.0	88.6
MPA-HKELM	82.7	82.7	84.9
GWO-HKELM	81.3	81.3	83.4
PSO-HKELM	77.3	77.3	84.2
KPCA-IMPA-HKELM	94.7	94.7	94.8
KPCA-MPA-HKELM	89.3	89.3	90.1
KPCA-GWO-HKELM	86.7	86.7	87.1
KPCA-PSO-HKELM	84.0	84.0	86.7

表 8 不同网络模型的对比结果

Table 8 Comparison results of different network models (%)			
	准确率	宏召回率	宏精确率
IMPA-HKELM	88.0	88.0	88.6
IMPA -BP	82.6	82.6	86.1
IMPA -ELM	84.0	84.0	86.9
IMPA -ELMAN	69.3	69.3	70.1
KPCA- IMPA-HKELM	94.7	94.7	94.8
KPCA- IMPA -BP	92.0	92.0	92.7
KPCA- IMPA -ELM	93.3	93.3	94.2
KPCA- IMPA -ELMAN	86.7	86.7	87.3

5 结 论

针对目前变压器故障数据复杂,诊断精度低的问题,提出了经核主成分分析法(KPCA)数据降维的改进海洋捕食者算法(IMPA)优化混合核极限学习机(HKELM)的变压器故障诊断模型。将 20 维的变压器故障数据通过 KPCA 进行特征提取,在保留主要特征的前提下,有效的降低了故障数据的维度,便于 HKELM 对数据的分类诊断,提升了诊断精度。因 MPA 算法搜索不平衡,寻优速度和精度都有待提高,通过伯努利混沌映射进行初始化;并对 MPA 的寻优阶段转换判据进行改进,加强寻优能力;最后由最佳候选者策略提升算法跳脱局部最优能力。应用测试函数验证 IMPA 的性能。并与其他智能算法对比,得知 IMPA 在寻优能力方面更优越。

通过对比实验结论得知,KPCA 降维前 IMPA-HKELM 模型的诊断精度为 88%,相比于 MPA-HKELM、GWO-HKELM、PSO-HKELM 分别提升了 5.4%、6.7%、10.9%,可知改进后的 MPA 优化算法寻优能力强,提升了 HKELM 的分类精度;KPCA 降维后,4 种模型的诊断精度又分别有所提升,因此 KPCA 处理后能够提升故障诊断的效果。然后选取多个前馈、反馈神经网络结合 IMPA 进行诊断精度对比,验证了 HKELM 的有效性。最后加入召回率、精确率两个指标进一步验证本文模型的性能,结果显示其 3 项指标均获得较好的结果,验证了本文变压器故障诊断模型的良好性能。

参考文献

- [1] 曹伟嘉,杨留方,徐天奇,等. 基于精英反策略麻雀搜索优化随机森林的变压器故障诊断[J]. 国外电子测量技术,2022,41(2):138-143.
- CAO W J, YANG L F, XU T Q, et al. Transformer fault diagnosis based on elite counterstrategy sparrow search algorithm optimized random forest [J]. Foreign Electronic Measurement Technology, 2022, 41(2): 138-143.
- [2] 代杰杰,宋辉,杨伟,等. 基于油中气体分析的变压器故障诊断 ReLU-DBN 方法[J]. 电网技术,2018,42(2):658-664.
- DAI J J, SONG H, YANG W, et al. Dissolved gas analysis of insulating oil for power transformer fault diagnosis based on ReLU-DBN [J]. Power System Technology,2018,42(2):658-664.
- [3] 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.

- [4] 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变压器故障诊断[J]. 高电压技术,2018,44(11):3664-3671.
- CHEN H, PENG H, SHU N Q, et al. Fault diagnosis of transformer based on LS-TSVM optimized by bat algorithm [J]. High Voltage Engineering, 2018, 44(11):3664-3671.
- [5] 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 T, QIAN Y, GUO C J, et al. Research on transformer fault diagnosis based on a beetle antenna search optimized support vector machine [J]. Power System Protection and Control, 2020,48(20):90-96.
- [6] 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 J, DING H H, MA X H, 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.
- [7] 张朝龙,何怡刚,杜博伦,等. 基于深度学习的电力变压器智能故障诊断方法[J]. 电子测量与仪器学报,2020,34(1):81-89.
- ZHANG CH L, HE Y G, DU B L, et al. Intelligent fault diagnosis method of power transformer using deep learning [J]. Journal of Electronic Measurement and Instrumentation,2020,34(1):81-89.
- [8] 魏金萧,周步祥,唐浩,等. 综合 RapidMiner 与改进粒子群极限学习机算法的变压器故障诊断[J]. 电力系统及其自动化学报,2019,31(3):133-138.
- WEI J X, ZHOU B X, TANG H, et al. Transformer fault diagnosis with the combination of RapidMiner-modified particle swarm optimization-extreme learning machine algorithm [J]. Proceedings of the CSU-EPSC, 2019,31(3):133-138.
- [9] 李黄曼,张勇,张瑶. 基于 ISSA 优化 SVM 的变压器故障诊断研究[J]. 电子测量与仪器学报,2021,35(3):123-129.
- LI H M, ZHANG Y, ZHANG Y. Research on transformer fault diagnosis based on ISSA optimized SVM [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(3):123-129.
- [10] 赵文清,严海,周震东,等. 基于残差 BP 神经网络的变

- 压器故障诊断[J]. 电力自动化设备, 2020, 40(2): 143-148.
- ZHAO W Q, YAN H, ZHOU ZH D, et al. BP neural network based on residual winding of transformer fault diagnosis [J]. Electric Power Automation Equipment, 2020, 40(2): 143-148.
- [11] 全兆景, 乔征瑞, 李金香, 等. 基于 MPC 算法优化的贝叶斯网络变压器故障诊断[J]. 电子测量技术, 2021, 44(17): 41-45.
- TONG ZH J, QIAO ZH R, LI J X, et al. Transformer fault diagnosis based on MPC algorithm optimized by Bayesian network [J]. Electronic Measurement Technology, 2021, 44(17): 41-45.
- [12] 崔江, 唐军祥, 张卓然, 等. 基于极限学习机的航空发电机旋转整流器快速故障分类方法研究[J]. 中国电机工程学报, 2018, 38(8): 2458-2466, 2555.
- CUI J, TANG J X, ZHANG ZH R, et al. Fast fault classification method research of aircraft generator rotating rectifier based on extreme learning machine [J]. Proceedings of the CSEE, 2018, 38(8): 2458-2466, 2555.
- [13] ZHANG Y, GUO J, ZHOU Q, et al. Research on damage identification of hull girder based on probabilistic neural network (PNN) [J]. Ocean Engineering, 2021, 238: 109737.
- [14] 谢国民, 倪乐水, 曹媛. 基于 VSRP 与 β -GWO-SVM 的变压器故障辨识方法[J]. 高电压技术, 2021, 47(10): 3635-3641.
- XIE G M, NI L SH, CAO Y. Transformer fault identification method based on VSRP and β -GWO-SVM [J]. High Voltage Engineering, 2021, 47(10): 3635-3641.
- [15] 雷帆, 高波, 袁海满, 等. 基于 DGA 的粗糙集与人工鱼群极限学习机的变压器故障诊断[J]. 高压电器, 2017, 53(10): 124-130.
- LEI F, GAO B, YUAN H M, et al. Transformer fault diagnosis based on DGA rough set and artificial fish swarm extreme learning machine [J]. High Voltage Apparatus, 2017, 53(10): 124-130.
- [16] 谢佩军, 高婷婷, 叶宏武. 量子粒子群优化核极限学习机的船舶变压器故障诊断[J]. 系统科学与数学, 2021, 41(7): 1807-1816.
- XIE P J, GAO T T, YE H W. Marine transformer fault diagnosis based on quantum-behaved particle swarm optimization and kernel extreme learning machine [J]. Journal of Systems Science and Mathematical Sciences, 2021, 41(7): 1807-1816.
- [17] 李可军, 徐延顺, 魏本刚, 等. 基于 PSO-HKELM 的变压器顶层油温预测模型[J]. 高电压技术, 2018, 44(8): 2501-2508.
- LI K J, XU Y SH, WEI B G, et al. Prediction model for top oil temperature of transformer based on hybrid kernel extreme learning machine trained and optimized by particle swarm optimization [J]. High Voltage Engineering, 2018, 44(8): 2501-2508.
- [18] 王昱皓, 武建文, 马速良, 等. 基于核主成分分析-SoftMax 的高压断路器机械故障诊断技术研究[J]. 电工技术学报, 2020, 35(S1): 267-276.
- WANG Y H, WU J W, MA S L, et al. Mechanical fault diagnosis research of high voltage circuit breaker based on kernel principal component analysis and SoftMax [J]. Transactions of China Electrotechnical Society, 2020, 35(S1): 267-276.
- [19] 余松, 胡东, 唐超, 等. 基于 TLR-ADASYN 平衡化数据集的 MSSA-SVM 变压器故障诊断[J]. 高电压技术, 2021, 47(11): 3845-3853.
- YU S, HU D, TANG CH, et al. MSSA-SVM transformer fault diagnosis method based on TLR-ADASYN balanced data set[J]. High Voltage Engineering, 2021, 47(11): 3845-3853.
- [20] 杨森, 陈莎莎, 李光茂, 等. 基于变分模态分解与特征选择的变压器有载分接开关机械故障诊断[J]. 南方电网技术, 2019, 13(1): 39-47, 59.
- YANG S, CHEN SH SH, LI G M, et al. Mechanical fault diagnosis of transformer on-load tap-changer based on variational mode decomposition and feature selection [J]. Southern Power System Technology, 2019, 13(1): 39-47, 59.
- [21] 张贺娜, 乐燕芬, 施伟斌. 基于特征降维的核岭回归室内定位算法[J]. 仪器仪表学报, 2020, 41(10): 83-91.
- ZHANG H N, LE Y F, SHI W B. Kernel ridge regression based indoor location algorithm using feature reduction [J]. Chinese Journal of Scientific Instrument, 2020, 41(10): 83-91.
- [22] RADHAKRISHNAN R K G, MARIMUTHU U, BALACHANDRAN P K, et al. An intensified marine predator algorithm (MPA) for designing a solar-powered BLDC motor used in EV systems [J]. Sustainability, 2022, 14(21): 14120.
- [23] 葛唱, 钱素琴. 改进麻雀搜索算法的无人车路径规划[J]. 导航定位学报, 2022, 10(6): 107-111.
- GE CH, QIAN S Q. Path planning of unmanned vehicle based on improved sparrow search algorithm [J]. Journal

of Navigation and Positioning, 2022, 10(6): 107-111.

- [24] 何庆, 罗仕杭. 混合改进策略的黑猩猩优化算法及其机械应用[J]. 控制与决策, 2023, 38(2): 354-364.

HE Q, LUO SH H. Chimp optimization algorithm based on hybrid improvement strategy and its mechanical application [J]. Control and Decision, 2023, 38(2): 354-364.

作者简介



谢国民, 1991 年于辽宁工业大学获得学士学位, 2003 年于辽宁工程技术大学获得硕士学位, 2012 年于辽宁工程技术大学获得博士学位, 现为辽宁工程技术大学教授, 主要研究方向为电力系统信息检测与故障诊断。

E-mail: lngdxgm@163.com

Xie Guomin received his B. Sc. degree from Liaoning University of Technology in 1991, M. Sc. degree from Liaoning Technical University in 2003 and Ph. D. degree from Liaoning Technical University in 2012, respectively. Now he is a professor in Liaoning Technical University. His main research interests include power system information detection and fault diagnosis.



刘东阳(通信作者), 2021 年于辽宁工程技术大学获得学士学位, 现为辽宁工程技术大学研究生, 主要研究方向为电力系统信息检测与故障诊断。

E-mail: 1318904908@qq.com

Liu Dongyang (Corresponding author)

received his B. Sc. degree from Liaoning Technical University in 2021. Now he is a M. Sc. candidate of Liaoning Technical University. His main research interests include power system information detection and fault diagnosis.



刘明, 2005 年于大连交通大学获得学士学位, 2011 年于大连交通大学获得硕士学位, 现为朝阳师范高等专科学校副教授, 主要研究方向为数据挖掘、网络与分布式技术。

E-mail: 21972107@qq.com

Liu Ming received her B. Sc. degree from Dalian Jiaotong University in 2005, M. Sc. degree from Dalian Jiaotong University in 2011. Now she is an associate professor at Chaoyang Teachers College. Her main research interests include data mining, network and distributed technology.