

5-26-2022

Transformer DGA fault diagnosis based on the random forest feature optimization and MAEPSO-ELM algorithm

Xuehui DING

School of Electric & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China

Hailin XU

Guangdong Electric Power Research Institute, Guangdong Power Grid Co., Ltd., Guangzhou 510080, China

Yingting LUO

Guangdong Electric Power Research Institute, Guangdong Power Grid Co., Ltd., Guangzhou 510080, China

Xin YANG

School of Electric & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China

Shenglong E

Guangdong Electric Power Research Institute, Guangdong Power Grid Co., Ltd., Guangzhou 510080, China

Follow this and additional works at: <https://jepst.researchcommons.org/journal>

Recommended Citation

DING, Xuehui; XU, Hailin; LUO, Yingting; YANG, Xin; and E, Shenglong (2022) "Transformer DGA fault diagnosis based on the random forest feature optimization and MAEPSO-ELM algorithm," *Journal of Electric Power Science and Technology*. Vol. 37: Iss. 2, Article 21.

DOI: 10.19781/j.issn.1673-9140.2022.02.021

Available at: <https://jepst.researchcommons.org/journal/vol37/iss2/21>

This Article is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology.

基于随机森林特征优选与 MAEPSO-ELM 算法的变压器 DGA 故障诊断

丁学辉¹, 许海林², 罗颖婷², 杨鑫¹, 鄂盛龙²

(1. 长沙理工大学电气与信息工程学院, 湖南 长沙 410114; 2. 广东电网有限责任公司电力科学研究院, 广东 广州 510080)

摘要:针对变压器故障智能诊断方法中输入特征不同影响诊断结果以及粒子群算法优化极限学习机准确率低的问题, 提出基于随机森林特征优选与多尺度协同变异粒子群极限学习机的变压器 DGA 故障诊断方法。首先, 基于故障样本 DGA 数据建立候选特征集, 采用随机森林算法计算各特征重要性评分并降序排列, 通过序列前向选择法筛选最优输入特征; 其次, 针对极限学习机参数选择困难的问题, 引入多尺度协同变异粒子群算法进行优化; 最后, 将之与 IEC 三比值法以及不同组合极限学习机诊断性能进行比较。实例表明所提方法诊断精度更高。

关键词:故障诊断; 随机森林; 多尺度协同变异粒子群; 极限学习机

DOI:10.19781/j.issn.1673-9140.2022.02.021 中图分类号: TM411 文章编号: 1673-9140(2022)02-0181-07

Transformer DGA fault diagnosis based on the random forest feature optimization and MAEPSO-ELM algorithm

DING Xuehui¹, XU Hailin², LUO Yingting², YANG Xin¹, E Shenglong²

(1. School of Electric & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China;

2. Guangdong Electric Power Research Institute, Guangdong Power Grid Co., Ltd., Guangzhou 510080, China)

Abstract: A transformer DGA fault diagnosis method is proposed based on the random forest feature optimization and multi-scale cooperative mutation particle swarm limit learning machine for the problems that different input characteristics effects the diagnosis results and the low accuracy of particle swarm algorithm optimization limit learning machine. Firstly, the candidate feature set is established on the basis of the DGA data in the fault sample. The random forest algorithm is utilized to calculate the feature importance scores and rank them in a descending order. The optimal input features are then selected by the sequence forward selection method. Next, aiming at the problem of difficult parameter selection of extreme learning machine, a multi-scale cooperative mutation particle swarm optimization algorithm is introduced for optimization. Finally, the method is compared for the diagnostic performance with the IEC three-ratio method and different combinations of extreme learning machines. An example shows that the proposed method has higher diagnostic accuracy.

Key words: fault diagnosis; random forest; multi-scale cooperative mutation particle swarm; extreme learning machine

收稿日期: 2020-02-03; 修回日期: 2020-10-21

基金项目: 广东电网有限责任公司科技项目(GDKJXM20173051)

通信作者: 杨鑫(1983-), 男, 博士, 副教授, 主要从事电力系统防雷接地、低温高压绝缘技术的研究; E-mail: yan_19830713@163.com

变压器承担着对电压的转换和对电能的分配,及时发现变压器内部潜在故障并准确诊断其故障类型对维护整个电网的安全运行与供电可靠性具有重要意义。

随着变压器油中溶解气体(DGA)在线监测技术不断发展,越来越多的基于DGA数据的人工智能算法被运用在变压器故障诊断领域并取得了不错的效果,如支持向量机、人工神经网络、贝叶斯分类器等^[1]。这些智能算法克服了传统IEC三比值法分类边界过于绝对、反映故障类型与表现特征之间关联关系准确率不足的局限性^[2],但诊断模型训练效果差,难以满足实际工程运用需求。

近年来,极限学习机(extreme learning machine, ELM)以其学习速度快、泛化能力优良、分类准确率高等特点在变压器故障诊断领域运用广泛,然而ELM对变压器故障分类时,随机产生的输入层权值和隐层阈值对故障诊断准确率会带来不良影响^[3]。常用粒子群优化(particle swarm optimization, PSO)算法求解ELM参数^[4],但该算法在搜索过程中易导致群体多样性丧失、早熟、陷入局部最优。为此,本文利用多尺度协同变异的自适应粒子群优化(multi-scale cooperative mutatingly self-adaptive escape PSO, MAEPSO)算法对ELM参数进行优化,建立多尺度协同变异粒子群极限学习机(MAEPSO-ELM)变压器故障诊断模型。MAEPSO算法的多尺度高斯变异机制,能促使整个种群以尽量分散的变异尺度对解空间进行更加详尽的探索,在强化算法全局寻优能力的同时兼顾局部精细化探索的能力,找到适应值更高的参数以提高ELM的诊断效果。

此外,目前基于DGA和智能算法建立的诊断方法大部分采用 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 、 C_2H_6 等主要特征气体及部分比值作为算法输入特征,在选择特征量时并没有充足的依据,而各种人工智能诊断方法会因为输入特征不同影响诊断结果^[5]。文献^[6]基于筛选器模式利用粗糙集方法对变压器故障输入特征进行约简;文献^[7]基于封装器模式利用SVM分类准确率评估变压器故障输入特征并通过遗传算法搜索适应度最高的输入特征集。筛选器模式独立于后续的智能分类算法,选取的输入特征集

准确度不高;封装器模式使用的智能搜索算法计算量大且需要提前设定输入特征维数。随机森林算法利用变量重要性评分(variable importance measure, VIM)值这一指标反映各输入特征对因变量变化所做贡献的重要性,在输入特征选择领域应用广泛^[8]。本文拟利用随机森林算法计算候选特征集各特征重要性评分并降序排列,结合优化的ELM,基于序列前向选择法逐一添加VIM值最高的输入特征计算其分类准确率,由分类准确率最大确定最优输入特征。该方法相比筛选器模式和封装器模式,具有计算量小且无需提前设定特征个数并能给出各输入特征对故障分类影响程度的优点。

综上所述,本文基于变压器DGA数据及相关运算属性构建候选特征集,建立基于随机森林特征优选与MAEPSO-ELM算法的变压器故障诊断方法,通过实例验证所提方法的有效性与正确性。

1 基于随机森林的特征选择算法

随机森林算法是利用bootstrap重采样技术从原始样本中抽取多个样本,对每个采样得到的样本集建立决策树进行训练,然后组合多棵决策树投票得到最终分类结果的集成决策树算法,同时该算法能给出各输入变量重要性评分,广泛应用于特征选择场景^[9]。随机森林算法对特征进行评估主要思想是通过未参与决策树训练的袋外(out-of-bag, OOB)数据施加扰动,再计算预测准确率的变化来获得特征重要性。

假定变压器故障数据训练集 (X, y) ,每个样本含 n 个输入特征 x_1, x_2, \dots, x_n ,随机森林算法通过bootstrap方法随机抽取 K 个子样本数据集,对应 K 个OOB数据集,特征重要性排序步骤如下:

- 1)初始化 $k=1$;
- 2)使用第 k 个子样本数据集时生成并训练决策树 T_k ,计算第 k 个OOB数据集的预测准确率 L_k^{OOB} ;
- 3)对OOB数据集中的特征 $x_i (i=1, 2, \dots, n)$ 添加随机扰动,重新计算准确率 $L_{k,i}^{OOB}$;
- 4)当 $k=2, 3, \dots, K$ 时,重复步骤2、3;
- 5)计算特征 x_i 重要性,即

$$V_{\text{IM}i} = \frac{1}{K} \sum_{k=1}^K (L_k^{\text{OOB}} - L_{k,i}^{\text{OOB}}) \quad (1)$$

6) 对 $V_{\text{IM}i}$ 降序排列, 得到特征重要度排序, VIM 越大特征越重要。

2 MAEPSO 算法

由于 PSO 算法中粒子的运动轨迹是一个有限的并逐渐减小的区域, 无法覆盖整个可行空间解, 因此, 存在早熟收敛、易陷入局部最优等不足^[10]。鉴于此, 引入多尺度协同变异粒子群算法, 该算法粒子速度变异能力由粒子所在子群高斯变异算子决定, 不同子群的变异算子尺度由该子群的适应度决定, 有助于粒子群在搜索空间中分散式搜索, 增强算法全局寻优能力, 同时变异算子的大小随着适应度的提升逐渐减小, 在算法后期有利于提高最优解的精度。

2.1 多尺度高斯变异算子

设粒子变异尺度为 M , 并设置 M 个高斯变异算子标准差的初始值, 其值为优化变量取值范围内的随机值:

$$\epsilon_0 = (\epsilon_1^0, \epsilon_2^0, \dots, \epsilon_M^0) \quad (2)$$

随着算法的迭代, 不同尺度变异算子的标准差会改变, 规律如下: 首先, 将种群中的 N 个粒子依据适应度 F 的大小升序排列; 然后, 将排好序的种群按照适应度大小划分为 M 个子群, 每个子群的粒子个数为 $P = N/M$ 。 t 为算法当前迭代次数, 则第 m 个子群在第 t 次迭代时平均适应度值可定义为

$$F_{X_m^t} = \frac{\sum_{i=1}^P F(x_i^m)}{p} \quad (3)$$

式中 $F(x_i^m)$ 为第 m 个子群中第 i 个粒子的适应度。

因此, 第 m 个子群变异算子标准差 ϵ_m^{t-1} 可以表示为

$$\left\{ \begin{aligned} \epsilon_m^{t-1} &= \exp\left(\frac{M \cdot F_{X_m^t} - \sum_{m=1}^M F_{X_m^t}}{F_{X_{\max}^t} - F_{X_{\min}^t}}\right) \\ F_{X_{\max}^t} &= \max(F_{X_1^t}, F_{X_2^t}, \dots, F_{X_M^t}) \\ F_{X_{\min}^t} &= \min(F_{X_1^t}, F_{X_2^t}, \dots, F_{X_M^t}) \end{aligned} \right. \quad (4)$$

由式(4)可以看出, 变异算子大小取决于所在子群适应能力, 并且排在后面的变异算子可能会很大。

大尺度变异算子在解空间进行粗搜索, 使粒子快速定位至适应度 F 较好时的最优解区域; 小尺度变异算子在进化后期完成局部精细化探索, 从而实现整个解空间覆盖性搜索。为减小变异算子标准差过大对粒子搜索的性能产生不利影响, 规定: 若 $\epsilon_m^t > W/4$, 则

$$\epsilon_m^t = \left| \frac{W}{4} - \epsilon_m^t \right| \quad (5)$$

式中 W 为待优化变量的空间宽度, $W \in [-V_{\max}, V_{\max}]$, V_{\max} 为粒子最大速度。重复式(5), 直到满足 $\epsilon_m^t < W/4$ 。

为最大范围实现空间搜索能力, 粒子速度 V_{ij} 达到一定阈值 T_j 时按伪代码进行一次变异操作, 计算变异后所有粒子适应度及 M 个变异算子, 取适应度最大位置作为下一次粒子新的逃逸点, 代码描述如下:

$$\begin{aligned} &\text{if } (V_{ij} < T_j) \text{ then} \\ &\quad F(x_i + \text{rand} \cdot \epsilon_i^t) = \\ &\quad \min_{0 \leq d \leq M, d \neq m} F(x_i + \text{rand} \cdot \epsilon_d^t) \\ &\quad \text{if } (F(x_i + \text{rand} \cdot \epsilon_i^t) < F(x_i + \text{rand} \cdot V_{\max})) \\ &\quad \quad V_{ij} = \text{rand} \cdot \epsilon_i^t \\ &\quad \text{else } V_{ij} = \text{rand} \cdot V_{\max} \end{aligned} \quad (6)$$

其中, T_j 记录种群中第 j 维的当前速度阈值且恒为正数, 其自适应设定原理参考文献[11], x_i 为粒子位置, rand 为均匀分布在 $[0, 1]$ 范围内的随机数。

3 基于随机森林特征优选的 MAEPSO-ELM 算法变压器故障诊断

将随机森林特征优选与 MAEPSO-ELM 算法变压器 DGA 故障诊断方法分为 3 步: 建立候选特征集、确定最优输入特征以及 ELM 参数优化过程。

3.1 建立候选特征集

本文采用广东电科院提供的已确认变压器故障类型样本, 共 417 组。根据相关规程, 诊断结果可划分为 7 种状态: 正常(N)、低能放电(D₁)、高能放电(D₂)、局部放电(PD)、中低温过热(T₁₂)、高温过热(T₃)、放电兼过热(TD), 分别用 {1, 2, 3, 4, 5, 6, 7} 表示, 样本分布如表 1 所示。随机选取 80 组样本为训练集、186 组为验证集、120 组为测试集。

表 1 变压器故障样本

Table 1 Transformer fault samples

故障类型	故障样本	故障类型	故障样本
N	51	T ₁₂	21
D ₁	108	T ₃	44
D ₂	66	TD	31
PD	96		

目前,常用 CH₄、C₂H₂、C₂H₄、C₂H₆、H₂ 的体积分数检测变压器是否发生故障而利用其对故障状态与性质进行更明确判断时依旧存在较大程度偏差。IEC 三比值法、Duval 三角形法、无编码比值法等一系列方法则常用于判定具体的故障类型。此外,文献[12]表明,与 DGA 全部数据相比,DGA 气体中 2 种气体的相对比值和变压器运行状态联系更紧密。为此,以上述 5 种气体为典型故障特征气体,并参照相关比值法构造故障特征的规则,选取关键气体的相关比率以及气体之间的总体占比作为候选特征集,如表 2 所示,其中 CH 为总烃含量。

表 2 候选输入特征

Table 2 Candidate input features

特征编号	故障特征	特征编号	故障特征
1	CH ₄	12	C ₂ H ₄ /CH ₄
2	C ₂ H ₂	13	C ₂ H ₄ /H ₂
3	C ₂ H ₄	14	C ₂ H ₆ /CH ₄
4	C ₂ H ₆	15	C ₂ H ₆ /H ₂
5	H ₂	16	CH ₄ /H ₂
6	CH	17	C ₂ H ₂ /CH
7	C ₂ H ₂ /H ₂	18	H ₂ /CH
8	C ₂ H ₂ /C ₂ H ₄	19	C ₂ H ₄ /CH
9	C ₂ H ₂ /C ₂ H ₆	20	CH ₄ /CH
10	C ₂ H ₂ /CH ₄	21	C ₂ H ₆ /CH
11	C ₂ H ₄ /C ₂ H ₆	22	CH ₄ +C ₂ H ₄ /CH

3.2 确定最优输入特征

首先,在训练集中通过随机森林算法计算各候选输入特征的 VIM 值并降序排列;然后,假设最优输入特征集合为空集,依据特征重要度排序,使用序列前向选择法,每次无放回地选取排序最前的特征加入最优输入特征集合,在验证集上利用 MAEPSO-ELM 模型进行训练,得到最优输入特征集合的分类准确率。遍历完所有特征后,最优输入特征集合即为分类准确率最高对应的特征。最优输入特征选择流程如图 1 所示。

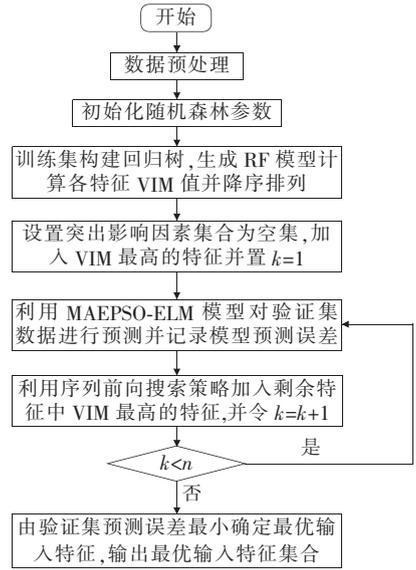


图 1 最优输入特征集合筛选流程

Figure 1 Optimal input feature set screening process

3.3 ELM 参数优化过程

利用 MAEPSO 算法优化 ELM 的输入层权值和隐层阈值,其中 ELM 算法分类准确率作为粒子适应度,算法流程如图 2 所示。

MAEPSO 算法优化步骤如下。

1) 预设网络拓扑结构、PSO 算法的惯性权重、学习因子等固有参数。

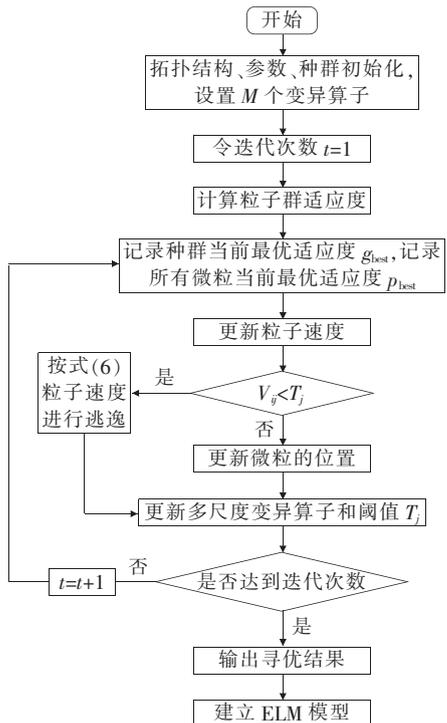


图 2 MAEPSO-ELM 算法流程

Figure 2 Flowchart of MAEPSO-ELM algorithm

2)种群初始化。随机初始化粒子的速度、位置、全局最优解,粒子由 ELM 输入权值和隐层阈值组成,通过训练变压器故障样本数据集计算初始化种群的个体适应度。反复迭代 N 次,共生成 N 个微粒,将其按适应度大小划分为 M 个子群,并设置初代高斯变异算子为 $(\epsilon_1^0, \epsilon_2^0, \dots, \epsilon_M^0)$ 。

3)根据粒子适应度竞争机制,更新所有微粒当前最优适应度 P_{best} 、种群当前最优适应度 g_{best} 。

4)根据 PSO 粒子速度进化公式调整微粒速度。

5)判断粒子速度是否满足 $V_{ij} < T_j$,并且按代码的描述(式(6))对 V_{ij} 进行变异。

6)按 PSO 粒子位置更新公式调整微粒的位置。

7)根据式(3)~(5)更新 M 个高斯变异算子的大小,并更新每一维速度的阈值 T_j 。

8)判断是否满足迭代条件。迭代次数 $t > t_{max}$ 则退出循环,输出最优适应度对应的 ELM 输入权值和隐层阈值,否则转至步骤 3。

9)由训练求出的最优参数建立 ELM 分类模型。

4 算例计算与分析

4.1 基于随机森林的故障特征优选

按本文方法通过随机森林算法计算训练集 22 个候选输入特征 VIM 值,在验证集下结合优化的 ELM,可得到不同输入特征个数时的分类准确率,如图 3 所示。本文随机森林参数设置:回归树数目为 400,节点最小尺寸为 5, mtry 节点值为 8。

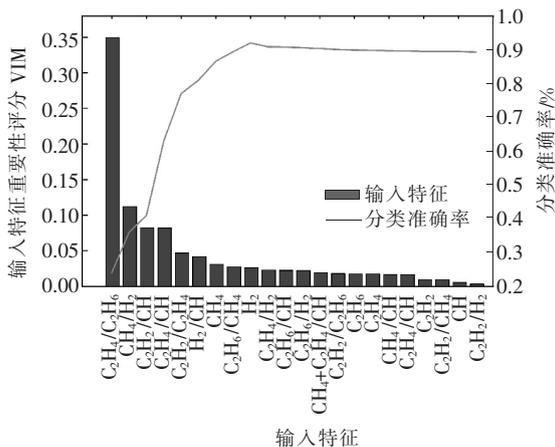


图 3 变量重要性评分及分类准确率曲线

Figure 3 Variable importance score and prediction error curve

由图 3 可知,输入特征维数较低时模型提取表征输出特性的关键信息较困难,分类准确率低;当按照 VIM 值排序逐个增加故障特征量时,进一步丰富了信息,分类准确率有所提高,并且在特征个数为 9 时准确率最大;继续增加特征量,分类准确率又呈下降趋势,最后趋于平稳。这说明在特征增加到一定个数时输入特征混入冗余特征,随之而来的冗余信息对模型的性能有副作用。由此可见特征选择的必要性,选择 VIM 值排序前 9 的特征作为本文方法的最优输入特征,分别为 C_2H_4/C_2H_6 、 CH_4/H_2 、 C_2H_2/CH 、 C_2H_4/CH 、 C_2H_2/C_2H_4 、 H_2/CH 、 $CH_4/C_2H_6/CH_4$ 、 H_2 。

从结果来看,最优输入特征排序的前几个特征包括 CH_4/H_2 、 C_2H_4/C_2H_6 、 C_2H_2/C_2H_4 ,即 IEC 导则推荐的三比值特征量,说明本文方法筛选的特征量与传统变压器故障诊断理论中故障特征比值的选取有较大一致性。

本文方法参数最终设置:种群总规模 $N=50$,尺度个数 $M=5$,惯性权重 ω 采用随进化次数线性递减方式,粒子学习因子 c_1 、 c_2 均取 1.5, V_{max} 为 100,迭代次数 t_{max} 为 160。ELM 输入节点为筛选出的特征维数 9,输出节点个数对应变压器 7 种状态设为 7,采用尝试对比法确定隐层节点数为 50,因此 ELM 拓扑结构为 9-50-7。

4.2 输入量的选择

构建粒子群极限学习机 (PSO-ELM)、量子粒子群极限学习机 (quantum particle swarm extreme learning machine, QPSO-ELM)、MAEPSO-ELM 模型,当输入为最优输入特征时,验证集在相同参数情况下各方法迭代过程的适应度进化曲线如图 4 所示。

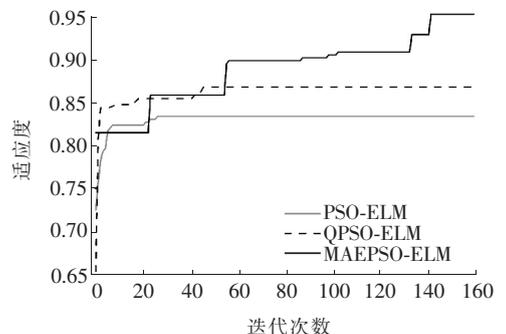


图 4 适应度随迭代次数的变化曲线

Figure 4 Variation curve of fitness with the number of iterations

图 4 表明在 PSO-ELM 模型中,粒子迭代早期就陷入局部最优,导致后续迭代搜索无法优化出适应度更高的曲线,此即为标准 PSO 算法的“早熟”缺陷,在约第 35 次迭代时训练精度收敛于 83.3%;QPSO-ELM 全局寻优性能略好于 PSO-ELM,在第 45 次迭代时进化陷入停滞,并收敛于 86.7%的训练正确率;MAEPSO-ELM 利用多尺度高斯变异算子对粒子速度进行变异,迭代初期粒子训练精度就达到 81%,远大于 PSO-ELM、QPSO-ELM;迭代后期,当粒子适应度陷入局部最优后,通过多尺度协同变异多次跳出局部最优在搜索空间中分散式搜索,在有限的迭代次数内提高搜索精度,最终收敛于 95.2%。综合对比,本文引入的 MAEPSO 算法寻优能力最好。

为验证本文方法高效性,将 MAEPSO-ELM 及对比模型输入特征分为 4 类:①最优输入特征;② DGA 气体比值,即 H_2/CH_4 、 CH_4/CH_4 、 C_2H_6/CH_4 、 C_2H_2/CH_4 、 C_2H_4/CH_4 ;③ IEC 气体比值,即 CH_4/H_2 、 C_2H_4/C_2H_6 、 C_2H_2/C_2H_4 ;④DGA 全数据。

3 种方法在 4 种输入特征下及 IEC 三比值法的测试样本整体诊断结果对比如图 5 所示,得出结论:

- 1)当采用相同诊断方法、最优输入特征作为输入时,诊断准确率比 DGA 全数据高 20%左右;
- 2)当采用相同特征量时,该文方法的识别准确率比 PSO-ELM、QPSO-ELM 故障诊断准确率高 5.6%~13.1%;
- 3)当同时采用最优输入特征与该文方法进行故障诊断时,能获得最高的故障诊断准确率;本文方法识别准确率为 91.3%,高于 IEC 三比值法 32.9%、QPSO-ELM 法 7.9%、PSO-ELM 法 13.1%,诊断效果最好。

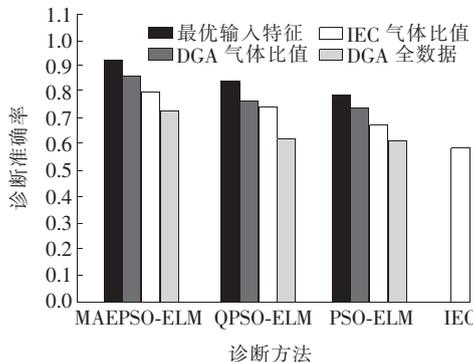


图 5 不同方法诊断准确率

Figure 5 Diagnosis accuracy of different methods

4.3 技术应用

利用本文方法对广东电网某供电局提供的 4 组变压器油色谱在线监测异常数据进行故障诊断,样本如表 3 所示。本文方法和 IEC 三比值法对 4 组样本的诊断结果及经吊芯检查后得出的明确变压器状态结论如表 4 所示。

表 3 故障样本

序号	油中溶解气体含量 $\mu L/L$				
	油中溶解气体含量				
	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
1	124.00	17.50	3.60	33.10	117.10
2	100.11	24.66	7.13	0.00	3.74
3	1582.00	95.50	35.00	48.00	0.00
4	24.21	39.94	84.06	8.78	64.44

表 4 故障样本诊断结果

序号	故障识别结果		吊芯检查结论
	IEC 三比值法	本文方法	
1	低能放电	高能放电	变压器本体 A 相转换开关处有电弧放电,即高能量放电故障发生
	无法识别	中低温过热	
2	无法识别	中低温过热	长期过载电流损耗引起铜过热,导致绝缘导线过热甚至绝缘不良,为低温过热故障
	缺编码组合	局部放电	
3	缺编码组合	局部放电	低压绕组 A 相有 2 根铝线脱焊松动,焊口处有放电痕迹,为局部放电故障
	高能放电	高能放电	
4	高能放电	高能放电	高压侧及低压侧对地绝缘受损,线圈匝间、层间击穿短路,为高能放电故障
	高能放电	高能放电	

由于 IEC 三比值分类边界过于绝对、所选用参数较少不能准确反映故障类型与表现特征之间的关联关系,因而无法正确识别第 1、2、3 组样本。本文方法诊断结果显示 1、4 组样本为高能放电故障,第 2、3 组样本分别为低温过热和局部放电故障,与吊芯检查结论基本一致,验证了该模型的有效性与正确性。

5 结语

本文通过随机森林算法结合 MAEPSO-ELM 进行变压器故障特征量的优选和故障诊断;结合变压器样本数据,筛选出与故障类型最相关的输入特征;并对比本文方法与不同组合的 ELM 在不同输入特征集下的诊断效果;最后,用某供电局的故障数

据进行测试。

1) 随机森林结合智能分类算法进行特征优选,能以较少的优选特征提高分类算法的诊断效果;

2) 通过 3 种不同组合 ELM 模型的参数寻优曲线可知,PSO、QPSO 算法较易陷入局部最优,寻优结果一般;而 MAEPSO 算法多次跳出局部最优,搜索空间中分散式搜索,找到适应值更高的参数,证明了多尺度协同变异机制的有效性;

3) 本文方法测试集的识别准确率为 91.3%,高于 IEC 三比值法 32.9%、QPSO-ELM 法 7.9%、PSO-ELM 法 13.1%,诊断效果最好。

参考文献:

[1] 李亮,范瑾,闫林,等.基于混合采样和支持向量机的变压器故障诊断[J].中国电力,2021,54(12):150-155.
LI Liang, FAN Jin, YAN Lin, et al. Transformer fault diagnosis based on hybrid sampling and support vector machines[J]. Electric Power, 2021, 54(12):150-155.

[2] 谢国民,倪乐水.基于 IABC 优化 SVM 的变压器故障诊断[J].电力系统保护与控制,2020,48(15):156-163.
XIE Guomin, XIE Guomin. Transformer fault diagnosis based on an artificial bee colony-support vector machine optimization algorithm[J]. Power System Protection and Control, 2020, 48(15):156-163.

[3] 王阳,乐健,周谦,等.基于小波包分析与神经网络的变压器区内外故障判断方法[J].电测与仪表,2020,57(7):1-7.
WANG Yang, LE Jian, ZHOU Qian, et al. Inner and outer zone fault diagnosis method of transformer based on wavelet packet analysis and neural network[J]. Electrical Measurement & Instrumentation, 2020, 57(7):1-7.

[4] 张丞鸣,谢菊芳,胡东,等.基于 QPSO-SVM 与 DGA 五边形解释工具的变压器故障诊断方法[J].高压电器,2021,57(12):117-124.
ZHANG Chengming, XIE Jufang, HU Dong, et al. Fault diagnosis method of transformer based on QPSO-SVM and DGA pentagonal interpretation tool[J]. High Voltage Apparatus, 2021, 57(12):117-124.

[5] 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 Jinxiao, ZHOU Buxiang, TANG Hao, et al. Transformer fault diagnosis with the combination of RapidMiner-modified particle swarm optimization-extreme learning machine algorithm[J]. Proceedings of the CSU-

EPSSA, 2019, 31(3):133-138.

- [6] 周光宇,马松龄.基于机器学习与 DGA 的变压器故障诊断及定位研究[J].高压电器,2020,56(6):262-268.
ZHOU Guangyu, MA Songling. Study of transformer fault diagnosis and location based on machine learning and DGA[J]. High Voltage Apparatus, 2020, 56(6):262-268.
- [7] 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.
WANG Ke, LI Jinzhong, ZHANG Shuqi, et al. New features derived from dissolved gas analysis for fault diagnosis of power transformers[J]. Proceedings of the CSEE, 2016, 36(23):6570-6578.
- [8] 王干军,李锦舒,吴毅江,等.基于随机森林的高压电缆局部放电特征寻优[J].电网技术,2019,43(4):1329-1336.
WANG Ganjun, LI Jinshu, WU Yijiang, et al. Random forest based feature selection for partial discharge recognition of HV cables[J]. Power System Technology, 2019, 43(4):1329-1336.
- [9] 易淑智,刘云凯,杨帆,等.基于改进门控循环单元分位数回归的短期负荷概率预测[J].智慧电力,2021,49(8):84-90.
YI Shuzhi, LIU Yunkai, YANG Fan, et al. Short-term load probability forecasting based on improved quantile regression of gated recurrent unit[J]. Smart Power, 2021, 49(8):84-90.
- [10] 党东升,张树永,葛鹏江,等.基于改进量子粒子群优化支持向量机的变压器故障诊断方法[J].电力科学与技术学报,2019,34(3):108-113.
DANG Dongsheng, ZHANG Shuyong, GE Pengjiang, et al. Transformer fault diagnosis method based on support vector machine optimized by improved quantum-behaved particle swarm optimization[J]. Journal of Electric Power Science and Technology, 2019, 34(3):108-113.
- [11] 班国邦,徐玉韬,蔡欢,等.基于改进量子粒子群算法的直流配电中心削峰填谷策略研究[J].电网与清洁能源,2021,37(12):86-95.
BAN Guobang, XU Yutao, CAI Huan. A study on peak load shaving strategies for DC distribution center based on improved QPSO algorithm[J]. Power System and Clean Energy, 2021, 37(12):86-95.
- [12] 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.