

基于改进 SSA 优化 MDS-SVM 的变压器故障诊断方法

谢国民[†], 蔺晓雨

(辽宁工程技术大学 电气与控制工程学院, 辽宁 葫芦岛 125105)

摘要: 为了提高变压器故障诊断精度, 提出一种基于改进 SSA 优化 MDS-SVM 的变压器故障诊断方法. 首先利用多维尺度缩放法 (Multiple Dimensional Scaling, MDS) 对 20 维变压器故障特征数据进行特征提取, 降低高维数据存在的稀疏性和多重共线性; 然后引入樽海鞘群算法 (Salp Swarm Algorithm, SSA), 并对该算法进行改进, 增置信赖机制和突变, 提高算法的收敛速度和收敛能力. 通过与原始 SSA、PSO、GWO 和 β -GWO 算法进行寻优测试对比, 验证了改进 SSA 算法的优越性. 最后使用改进 SSA 算法对 MDS 降低维度和支持向量机 (Support Vector Machine, SVM) 的参数联合寻优, 构建新的故障诊断模型. 分析比较其与常用算法优化的 SVM 故障诊断模型、BP 神经网络 (Back Propagation Neural Network, BPNN)、K 最近邻 (K-Nearest Neighbor, KNN) 以及随机森林 (Random Forest, RF) 故障诊断模型的故障诊断精确度, 结果证明基于改进 SSA 的 MDS-SVM 变压器故障诊断模型的精确度高于其它算法模型, 且泛化能力较强.

关键词: 变压器; 故障诊断; 多维尺度缩放法; 樽海鞘算法; 支持向量机; 算法改进

中图分类号: TM407; TP18

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1437

开放科学 (资源服务) 标识码 (OSID):



Transformer Fault Diagnosis Method Based on Improved SSA Optimized MDS-SVM

XIE Guo-min[†], LIN Xiao-yu

(Faculty of Electrical and Control Engineering, Liaoning Technical University, Liaoning 125105, China)

Abstract: In order to improve the accuracy of transformer fault diagnosis, a transformer fault diagnosis method based on improved SSA optimized mds-svm is proposed. Firstly, multi dimensional scaling (MDS) is used to extract features from 20 dimensional transformer fault feature data to reduce the sparsity and multicollinearity of high-dimensional data; Then, the paper introduces the salp swarm algorithm (SSA) and improves the algorithm by adding trust mechanism and mutation to improve the convergence speed and ability of the algorithm. By comparing with the original SSA, PSO, GWO and β -GWO algorithm is tested and compared to verify the superiority of the improved SSA algorithm. Finally, the improved SSA algorithm is used to reduce the dimension of MDS and optimize the parameters of support vector machine (SVM) to build a new fault diagnosis model. The fault diagnosis accuracy is analyzed and compared with the SVM fault diagnosis model optimized by common algorithms, BP neural network (BPNN), k-nearest neighbor (KNN) and random forest (RF) fault diagnosis models. The results show that the accuracy of MDS-SVM transformer fault diagnosis model based on improved SSA is higher than that of other algorithm models, and the generalization ability is stronger.

Keywords: transformer; fault diagnosis; multiple dimensional scaling; salp swarm algorithm; support vector machine; algorithm improvement

0 引言

电力变压器作为电力系统输配电、用户配电系统当中使用的关键设备, 承担着不可替代的重要作用^[1]. 当其发生故障时若不能及时诊断, 会使电力系

统正常安全运行得不到保证, 会使供电系统运行产生巨大的安全隐患, 带来不可预计的经济损失, 甚至可能会对人们的生命安全造成严重威胁^[2]. 因此如何确保故障诊断的高精度意义重大.

收稿日期: 2021-08-15; 录用日期: 2021-11-10.

基金项目: 国家自然科学基金项目 (51974151); 辽宁省教育厅重点实验室基金项目 (LJZS003).

[†]通讯作者. E-mail: 804764995@qq.com.

当前油浸式电力变压器在输配电系统中被普遍使用,而且基于变压器油中的溶解气体成分分析(Dissolved Gas Analysis, DGA)法成为了一种主要的变压器故障诊断手段.传统依托此种分析方法进行变压器故障诊断包括改良的三比值法、Rogers 比值法、Nomograph 法和 Duval 三角形法等^[3],但这些诊断方法都有一定局限,大部分需要依赖于经验知识,导致故障诊断的精度不能得到很好的保证.随着人工智能、机器学习和数据挖掘等技术日渐成熟,各种机器学习方法被一一应用到变压器故障诊断当中,如支持向量机、神经网络、专家系统、随机森林等等^[4-6].但人工神经网络需要大量数据样本,学习成本相对较高,构建的系统也相对复杂,专家系统需要制定大量的知识规则,学习能力较差,随机森林算法训练空间和时间受决策树数量的影响,而且当数据集噪声较大时容易陷入过拟合.

相比于上述机器学习方法,支持向量机算法简单,鲁棒性和泛化能力相对出色,本身也属于小样本学习方法,不需要大量数据,正好与变压器故障数据较少的特点相吻合,故本文选用 SVM 作为主体模型进行故障诊断,但支持向量机的分类精确度与数据样本维度 惩罚因子 C 以及核函数自由参量 g 的选择密切相关,因此如何对数据进行降维处理,如何找到最优 C , g 参数成为提高支持向量机的故障诊断精确度的关键^[7-8].而群体智能寻优算法的应用可以有效解决参数寻优问题,文献 [9]、[10] 分别选用 PSO, GWO 算法对 SVM 的参数进行寻优,但这些算法在收敛速度,全局寻优能力等方面都存在不足,容易陷入局部最优,影响支持向量机的分类准确度.

本文提出一种基于改进樽海鞘算法优化 MDS-SVM 的变压器故障诊断方法.通过 MDS 在降低高维特征数据维度的基础上尽可能保持了数据彼此之间的相似性不发生变化,然后对樽海鞘算法进行改进,全面提高算法收敛速度和全局寻优能力,并利用改进后的樽海鞘算法对 MDS 最适降低维数和 SVM 的参数进行联合寻优,构建 MDS-SVM 故障诊断模型,最后和传统 PSO, GWO, β -GWO 算法优化的 SVM 故障诊断模型以及其他传统故障诊断模型进行故障诊断精确度对比实验,验证了提出的变压器故障诊断方法的高精度和可行性.

1 变压器故障特征数据选择与降维

油浸式变压器故障时,绝缘油会发生热裂解,产生 H_2 、 CH_4 、 C_2H_2 、 C_2H_4 和 C_2H_6 五种主要气体并溶解于绝缘油中,不同故障类型下溶解的气体

浓度和各气体之间的比例关系各不相同,如发生高能放电和低能放电主要气体均为 H_2 和 C_2H_2 ,但高能放电相对于低能放电来说这两种气体浓度会更高,为了能完整的反映变压器故障类型特征,提高变压器故障诊断精度,本文选用五种气体油中溶解相对浓度、两两气体间比例关系以及五种气体与总烃的比例关系作为特征数据,具体特征数据如表 1.

表 1 变压器故障特征数据

| 20 维特征数据 | | | | |
|-------------------------|--------------------------|----------------------------|----------------------------|----------------------------|
| H_2 | CH_4 | C_2H_2 | C_2H_4 | C_2H_6 |
| $\frac{CH_4}{H_2}$ | $\frac{C_2H_2}{H_2}$ | $\frac{C_2H_4}{H_2}$ | $\frac{C_2H_6}{H_2}$ | $\frac{C_2H_2}{CH_4}$ |
| $\frac{C_2H_4}{CH_4}$ | $\frac{C_2H_2}{C_2H_4}$ | $\frac{CH_4}{C_2H_6}$ | $\frac{C_2H_2}{C_2H_6}$ | $\frac{C_2H_4}{C_2H_6}$ |
| $\frac{H_2}{C_1 + C_2}$ | $\frac{CH_4}{C_1 + C_2}$ | $\frac{C_2H_2}{C_1 + C_2}$ | $\frac{C_2H_4}{C_1 + C_2}$ | $\frac{C_2H_6}{C_1 + C_2}$ |

其中 $C_1 + C_2 = CH_4 + C_2H_2 + C_2H_4 + C_2H_6$.

虽然高维数据可以完整的体现故障特征信息,但在高维空间中数据存在稀疏性和多重共线性,前者不利于 SVM 模型对数据特征进行识别,而后者将会导致求解空间不稳定,从而降低 SVM 的泛化能力.而且特征矩阵维数过大,会导致 SVM 模型计算量过大,进而造成训练时间过长,影响诊断系统的实时性.

传统的降维手段比如主成分分析法(Principal Component Analysis, PCA)为达到降维效果只保留方差大的主成分,导致可能丢失方差小的主成分中对分类结果有影响的特征,局部线性嵌入(Locally Linear Embedding, LLE)虽然可以处理非线性特征数据,但降维效果会受最近邻样本数选取的影响.MDS 算法^[11-12]本质上就是在低维空间用一些数据点与高维空间的数据点一一对应,并且保证数据点彼此之间的相似性尽可能维持不变:

1) 设定数据个数为 n ,则数据空间 $T = \{x_1, x_2, \dots, x_n\}, x_i \in R^d$,其中 d 为数据的维度.

2) 设 $D \in R^{n \times n}$ 为数据之间的距离矩阵,其中 $dist_{ij}$ 代表矩阵 D 中第 i 行第 j 列元素,大小等于数据 x_i 到数据 x_j 之间的距离.

3) 经过降维之后的目标矩阵为 $Z \in R^{d' \times n}, d' < d$, Z 矩阵中任意两个数据的欧氏距离尽可能等于 D 矩阵中的距离,即 $\|z_i - z_j\| \approx dist_{ij}, d'$ 即降维后维度.

越接近原维度, $\|z_i - z_j\| - dist_{ij}$ 的绝对值就越小,因此 d' 的选择会影响特征数据降维效果.

2 改进 SSA 算法

2.1 SSA 算法

通过研究海洋生物樽海鞘群的链式觅食行为, Seyedali Mirjalili 等学者在 2017 年提出的一种全新的群体智能算法^[13], 该算法需要设置的参数较其他群体智能算法来说相对较少, 现已在同步电机多参数识别、短期风电功率预测、行星齿轮箱故障诊断和翼形设计等领域取得了比较好的成果。

首先设定一个 m 维空间, 所有的樽海鞘个体分布其中, 位置矩阵 D 如下所示:

$$D = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_m^1 \\ x_1^2 & x_2^2 & \dots & x_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^N & x_2^N & \dots & x_m^N \end{bmatrix}. \quad (1)$$

N 为樽海鞘种群数量, 前半为领导者, 后半为追随者. x_j^i 和 F_j 分别表示第 i 个个体和食物在 j 维的位置. 则第一个樽海鞘即领导者的位置更新如下:

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j), c_3 \geq 0.5 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j), c_3 < 0.5. \end{cases} \quad (2)$$

式中 ub_j 和 lb_j 用来限定 j 维搜索空间大小, 分别代表上限和下限, c_1 作为 SSA 算法的重要参数, 用来平衡全局搜索和局部开发的能力, 约束公式如下:

$$c_1 = 2e^{-(\frac{4L'}{L})^2} \quad (3)$$

L' 和 L 分别表示目前迭代次数和最大迭代次数. c_2 和 c_3 均为 0 到 1 之间的随机数, c_2 决定领导者个体迭代更新时移动的步长, c_3 决定其移动的方向。

追随者位置的下一迭代更新的位置由自身当前位置和前一个樽海鞘位置决定, 公式如下:

$$x_i^j = \frac{1}{2}(x_j^i + x_j^{i-1}) \quad (4)$$

每一次迭代过程中, 将适应度最优秀的个体位置假定为食物位置, 经过多次迭代, 使整个樽海鞘链整体向食物位置移动, 最终达到最优解^[14-16]。

2.2 SSA 算法的改进

SSA 算法虽然具有设置参数少的优点, 但收敛速度和其他算法相比还具有一定的差距, 而且算法还存在另一个较大缺陷—易陷入局部最优解. 如公式 (4) 所示, 追随者个体进行位置更新时完全取决于它前一个个体的位置, 这将导致如果前面的个体陷入局部最优, 就会连锁反应使整个樽海鞘链陷入局部最优, 从而无法到达全局最优点. 因此为了提高该算法的适用性和泛化能力, 必须对该算法进行改进, 本文从以下两方面对该算法进行优化:

1) 加入信赖机制

对追随者位置更新模式进行改进: 在迭代前期, 追随者个体对前一个个体位置不够信赖, 它自己也会去探寻食物的位置, 但能力有限, 因此在原更新公式中增加食物位置 F_j 以及能力限制因子 ν , 其权重会随着迭代次数增加而减小, 即这种不信赖逐渐降低, 相当于在一个群体当中成员彼此会随着时间越来越信赖对方, 这样可以有效防止陷入局部最优, 并且在这种信赖机制的基础上加入两对彼此互相限制的随机因子, 进一步提高寻优能力, 具体修改公式如下:

$$x_j^i = \frac{1}{2} \{ p \cdot [q \cdot x_j^{i-1} + (1-q) \cdot (\frac{L'}{L})^\omega \cdot x_j^i] + (1-p) \cdot \nu \cdot (1 - \frac{L'}{L}) \cdot F_j \} \quad (5)$$

其中 p 、 $(1-p)$ 和 q 、 $(1-q)$ 为两对彼此限制的 0 到 1 之间的随机数, 用来增加随机性. 引入 $(1-p)$ 和 $(1-q)$ 是为防止出现两项随机数因子同时过小或同时过大从而导致更新位置更加偏离最优位置, 影响算法收敛能力及速度. ω 影响权重因子变化速度, 而能力限制因子 ν 限制追随者寻找食物的能力, 且不宜过高, 否则容易陷入局部最优, 经多次实验比较测试, 二者分别取 8 和 0.2. 在迭代前期, $(\frac{L'}{L})^8$ 非常小, 位置更新主要依据前一个个体位置和食物位置, 迭代后期, $0.2 \cdot (1 - \frac{L'}{L})$ 很小, 位置更新主要依据前一个个体和自身位置. 若迭代次数为 500 次, 设 $\alpha = (\frac{L'}{L})^8$, $\beta = 0.2 \cdot (1 - \frac{L'}{L})$, 则两者变化关系曲线如图 1 所示。

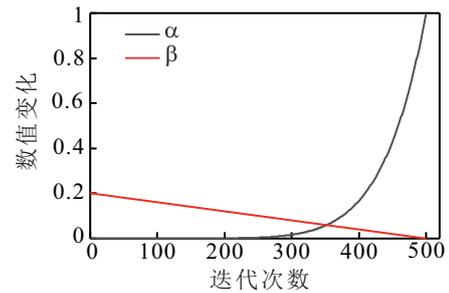


图 1 α 、 β 变化关系曲线

2) 加入突变

为了防止陷入局部最优, 在每一次迭代中, 个体位置更新后都可能发生突变, 并从两个层面对其进行描述, 分别为突变发生概率和突变程度:

1. 突变发生概率遵循卡方分布概率密度函数, 表达式如下:

$$P = \frac{1}{2^{n/2}\Gamma(n/2)} (L'/10)^{\frac{n}{2}-1} e^{-L'/20} \quad (6)$$

式中 n 为自由度, L' 为当前迭代次数。

不同 n 取值下的突变概率曲线图如图 2 所示. 通

过多次测试实验验证,当 n 取 7 时寻优效果最好,故本文 n 取值为 7.

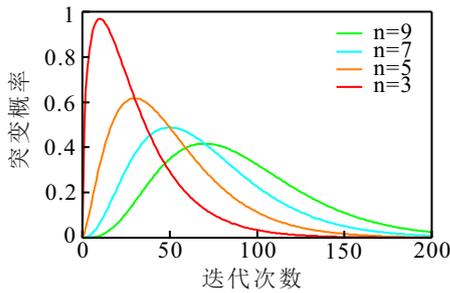


图 2 不同 n 取值下的突变概率测试曲线图

2. 突变程度取决于自身位置以及迭代次数,约束公式如下:

$$x_j^i = x_j^i + (1 - \frac{L^i}{L})x_j^i \quad (7)$$

从图 2 概率变化走势和上述公式可以看出,在迭代初期,突变概率和突变程度逐渐增大,利于增加种群丰富度,有效避免出现陷入局部最优的情况,后期突变概率和突变程度逐渐降低,利于局部开发,加快收敛速度.

2.3 改进 SSA 算法寻优能力测试

为验证改进 SSA 算法的寻优能力,选用表 2 所示的四种函数作为对比不同算法寻优性能的测试函数.

表 2 四种测试函数

| 测试函数 | 维度 | 搜索范围 | 最优解 |
|---|----|-----------------|-----|
| $f_1 = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $ | 10 | $[-100, 100]$ | 0 |
| $f_2 = \max\{ x_i , 1 \leq i \leq n\}$ | 10 | $[-100, 100]$ | 0 |
| $f_3 = \sum_{i=1}^n i \cdot x_i^4 + \text{random}[0, 1]$ | 10 | $[-1.28, 1.28]$ | 0 |
| $f_4 = -20 \exp(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}) - \exp[\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)] + 20 + e$ | 10 | $[-40, 40]$ | 0 |

将改进的 SSA 算法和原始 SSA、PSO、GWO 以及 β -GWO 算法进行寻优性能比较,寻优过程如图 3 至图 6 所示,寻优结果如表 3 所示.

f_1 和 f_2 测试函数表面光滑,仅有全局极小点,改进 SSA 算法较其他算法而言表现出了优异的收敛速度,而 f_3 和 f_4 测试函数均分布若干个局部极小点,可用于测试寻优智能算法跳出局部最优的能力,依据寻优过程曲线图和寻优结果可以看出改进 SSA 算法仍能收敛到全局极小点附近.

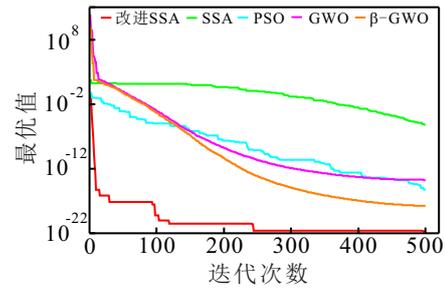


图 3 f_1 函数寻优曲线

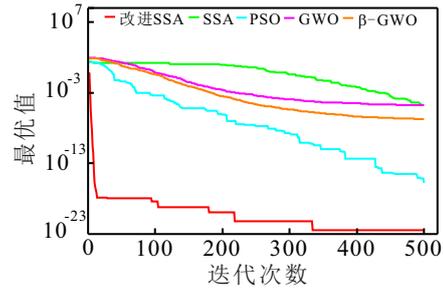


图 4 f_2 函数寻优曲线

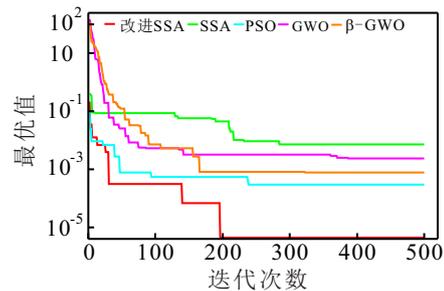


图 5 f_3 函数寻优曲线

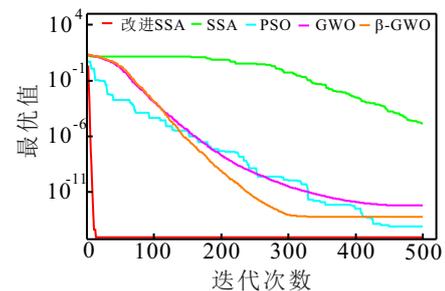


图 6 f_4 函数寻优曲线

通过以上比较实验可以证明:加入信赖机制和概率突变对于改善樽海鞘算法的寻优能力和收敛速度具有十分明显的作用,而且和其他通用算法相比,对于不同测试函数,改进后的樽海鞘算法都具有较强的适应能力.

表 3 各算法对四种测试函数的寻优性能

| 算法类型 | 相同迭代次数 (500 次) 算法寻优结果 | | | |
|--------------|------------------------|------------------------|-----------------------|------------------------|
| | f_1 函数 | f_2 函数 | f_3 函数 | f_4 函数 |
| 改进 SSA | 2.46×10^{-22} | 3.44×10^{-23} | 4.50×10^{-6} | 8.88×10^{-16} |
| SSA | 6.47×10^{-6} | 1.70×10^{-5} | 7.20×10^{-3} | 1.41×10^{-5} |
| PSO | 4.99×10^{-16} | 6.69×10^{-16} | 2.97×10^{-4} | 7.99×10^{-15} |
| GWO | 1.91×10^{-14} | 1.79×10^{-5} | 2.33×10^{-3} | 6.12×10^{-13} |
| β -GWO | 1.86×10^{-18} | 1.94×10^{-7} | 7.70×10^{-4} | 5.77×10^{-14} |

| 算法类型 | 相同收敛精度 (10^{-6}) 算法迭代次数 (限定最大 500 次) | | | |
|--------------|--|----------|----------|----------|
| | f_1 函数 | f_2 函数 | f_3 函数 | f_4 函数 |
| 改进 SSA | 5 | 4 | 196 | 5 |
| SSA | 495 | 未达到 | 未达到 | 未达到 |
| PSO | 119 | 138 | 未达到 | 118 |
| GWO | 130 | 未达到 | 未达到 | 141 |
| β -GWO | 125 | 277 | 未达到 | 132 |

3 改进 SSA 优化的 MDS-SVM 诊断模型

3.1 SVM 分类器

SVM 的分类策略是把分类问题转变为求解一个凸集二次优化问题^[17-19], 即特征空间上的间隔如何才能达到最大, 公式如下:

$$\begin{aligned} & \max \frac{1}{\|\omega\|} \\ & \text{s.t. } y_i(\omega^T x_i + b) \geq 1, i = 1, \dots, n \end{aligned} \quad (8)$$

其中 ω 为所选超平面的法向量, x 为特征向量, y 为结果标签, b 为阈值, n 表示样本个数.

但实际应用中所用到的数据往往并不能够线性可分, 这个时候就需要引入核函数来将数据映射到某个高维空间, 并在高维空间中找到一个最优分离超平面来把在最初数据空间中本不能线性可分的数据进行分类. 由于高斯核函数具有高灵活性、应用广泛特点, 所以本文采用高斯核函数映射变换, 表达式如下:

$$k(x_1, x_2) = \exp(-g\|x_1 - x_2\|^2), g > 0 \quad (9)$$

式中 g 为一个自由参数.

求 $\frac{1}{\|\omega\|}$ 的最大值等价于求 $\frac{1}{2}\|\omega\|^2$ 的最小值, 为了避免某些异常数据点即离群值 (outliers) 对分类效果的影响, 引入松弛变量, 即

$$\begin{aligned} & \min \frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^n \xi_i \\ & \text{s.t. } y_i(\omega^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, n \end{aligned} \quad (10)$$

式中 C 为惩罚因子, 用来平衡函数中两项的权重, ξ_i 为松弛常量.

由式 (9) 和式 (10) 可知, C 和 g 选择是否合适, 将决定 SVM 分类性能好坏.

3.2 MDS-SVM 故障诊断模型

对变压器原始数据采用典型 MDS 降维方法时步骤如下:

- 1) 计算变压器数据的距离矩阵 D .
- 2) 通过 D 矩阵求取 Z 矩阵的内积矩阵 K .
- 3) 对 K 矩阵进行特征分解.
- 4) 对特征值做降序处理得到前 d' 个构成对角矩阵 $\Lambda_{d'}$ 和对应特征向量矩阵 $V_{d'}$.
- 5) 得到目标矩阵 $Z = \Lambda_{d'}^{\frac{1}{2}} V_{d'}^T \in R^{d' \times n}$.

对数据进行降维时通常首先人为设定一个可接受精度, 然后分析降低到不同维度数据特征带来的误差, 误差达到设定精度时的维度即为 d' , 这种方法由于人为因素的存在, 通常会带来两种问题, 第一种情况是精度设置过于保守, 从而使特征损失过多, 影响故障诊断精度, 第二种情况是精度过高, 从而使 d' 设置过大, 达不到有效降维的目的, 同样会影响 SVM 模型的故障诊断精确度.

为了探究不同取值对变压器故障诊断精度的影响, 从国家电网南方某电网提供的数据中按故障状态随机挑选出共 270 组变压器数据作为实验样本数据, 然后进行降维实验.

首先求取原始变压器特征数据样本间的欧氏距离, 然后利用 MDS 降维到不同维度, 再求取降维之后低维样本之间的欧氏距离, 以所有样本降维前后距离之差的绝对值的平均值 γ 作为特征保留目标函数, γ 越小, 特征保存越完整, 其表达式如下:

$$\gamma = \frac{\sum_{i=1}^n \sum_{j=1}^n |d1_{ij} - d2_{ij}|}{n^2} \quad (11)$$

式中 n 为样本个数, $d1_{ij}$ 为降维前各样本间欧氏距离, $d2_{ij}$ 为降维后各样本欧氏距离.

通过进行多次降维实验, 做出特征保留目标函数随维度变化曲线如图 7.

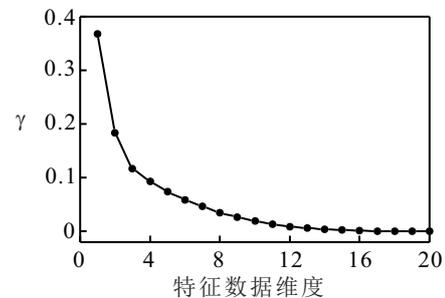


图 7 特征保留目标函数随维度变化曲线

从曲线图中可以看出, 特征保留目标函数 γ 的数值和变化程度随维度增加而减小, 当降到 7 维时已经低于 0.05, 从 14 维之后已经接近于 0. 虽然维数越高, 特征保存越完整, 但维数越高, 也会影响 SVM

的诊断精度,如图8所示.

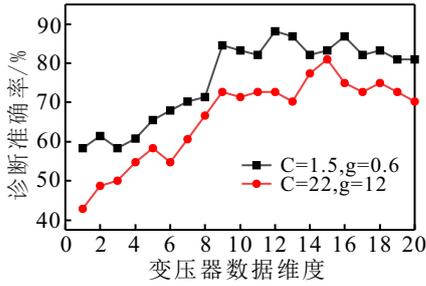


图8 不同维度不同参数 SVM 诊断精度对比

上图随机设定两组 SVM 的 C、g 参数,测试不同维度下模型故障诊断精度,实验结果表明支持向量机参数设置不同,最优降维维度 n 也不同,因此引入改进樽海鞘智能寻优算法,并改进传统故障诊断方法只对支持向量机参数进行寻优,将降维维度也融入寻优算法,对 n、C、g 三个参数同步联合寻优,获取故障诊断最优的 MDS-SVM 模型,寻优步骤如下:

step 1: 设定樽海鞘种群的数量为 30, 维度为 3, 参数 C 的搜索空间范围为 0 ~ 500, 参数 g 的搜索范围为 0 ~ 200, 参数 n 的搜索空间范围为 1 ~ 20, 对种群进行初始化;

step 2: 设置 SVM 相关参数, 设置迭代次数为 500;

step 3: 对训练数据进行 5 折交叉验证, 将平均交叉验证准确率作为适应度, 计算樽海鞘个体的适应度;

step 4: 将每个个体的适应度按照大小进行排列, 适应度最大的个体位置设为食物位置;

step 5: 把经过适应度降序排列的樽海鞘链除去设定为食物位置的樽海鞘之外剩下的平均分成两部分, 前一部分作为领导者, 后一部分作为追随者;

step 6: 按照领导者和追随者位置更新公式对樽海鞘个体的位置进行更新;

step 7: 种群个体发生概率突变;

step 8: 对更新之后的樽海鞘个体再次计算适应度并与食物位置适应度比较, 将大于食物位置适应度的个体位置设定为新的食物位置;

step 9: 判断是否达到最大迭代次数, 若没有达到返回 Step3, 若达到最大迭代次数, 输出当前食物的位置即 C、g 和 n 的最优值, 寻优结束.

故障诊断流程图如图 9.

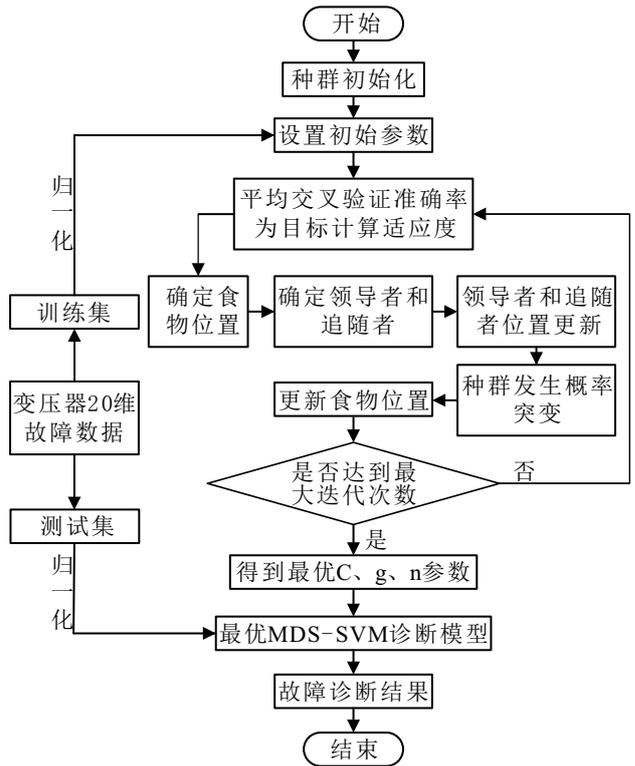


图9 故障诊断流程图

4 故障诊断精度对比实验

将从南方某电网获取的 252 组 20 维变压器特征数据按 2: 1 的比例分成训练集和测试集并对运行状态进行编号如下表.

表 4 实验数据集分类

| 运行状态 | 正常 | 高能 放电 | 低能 放电 | 局部 放电 | 高温 过热 | 中低温 过热 |
|-------|----|----------|----------|----------|----------|-----------|
| 状态编号 | 1 | 2 | 3 | 4 | 5 | 6 |
| 训练集个数 | 30 | 30 | 20 | 24 | 40 | 24 |
| 测试集个数 | 15 | 15 | 10 | 12 | 20 | 12 |

4.1 不同算法优化 MDS-SVM 及 SVM 的性能比较

1) 首先不对数据进行 MDS 降维, 直接利用改进 SSA, SSA, PSO, GWO 和 β -GWO 算法对支持向量机进行优化, 得到最优 C、g 参数, 然后将测试集数据集归一化后输入不同算法优化的 SVM 模型中进行故障诊断, 结果如图 10 至图 14 所示.

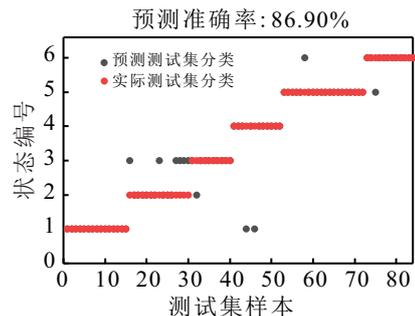


图10 改进 SSA 优化 SVM 故障诊断结果

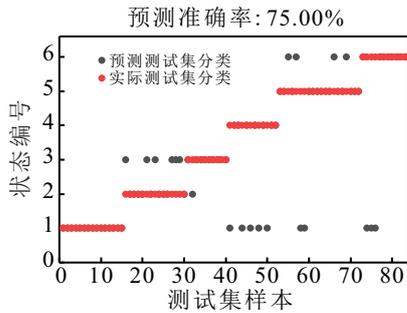


图 11 SSA 优化 SVM 故障诊断结果

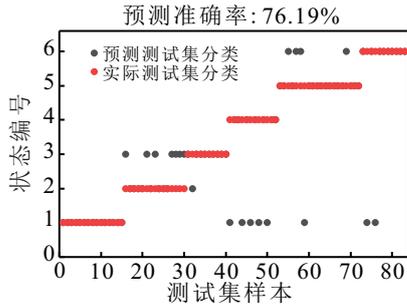


图 12 PSO 优化 SVM 故障诊断结果

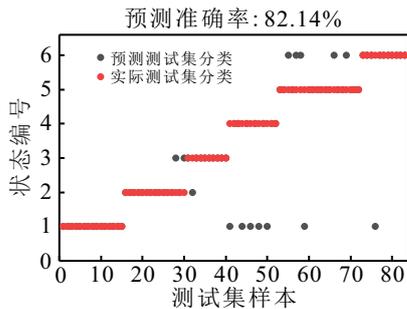


图 13 GWO 优化 SVM 故障诊断结果

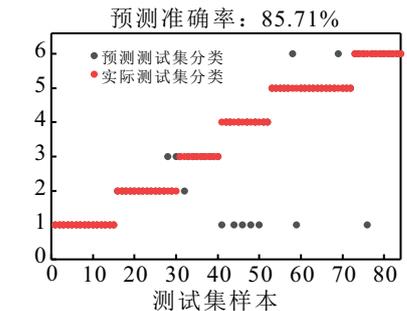


图 14 β -GWO 优化 SVM 故障诊断结果

通过图 8 可以看出:

(1) 未改进的 SSA 算法和 PSO 算法优化的 SVM 对于高能放电和低能放电之间、正常运行和局部放电之间以及高温过热和中低温过热之间的区分存在明显不足, 识别准确率仅为 75% 和 76.19%.

(2) GWO 算法优化的 SVM 提高了高能放电和低能放电的区分能力, 加入 β 序列改进后进一步提高了高温过热和中低温过热的区分能力, 但仍然没有解决正常运行和局部放电识别准确率过低的问题.

(3) 改进 SSA 算法优化的 SVM 对高能放电和低能放电、正常运行和局部放电以及高温过热和中低温过热的区分准确度都有了明显的提高.

2) 应用以上五种算法优化 MDS-SVM 模型, 对降维维度 n 和支持向量机 C 、 g 参数联合寻优, 然后进行故障诊断, 诊断结果如图 15 至图 19 所示, 诊断精度对比如表 5 所示.

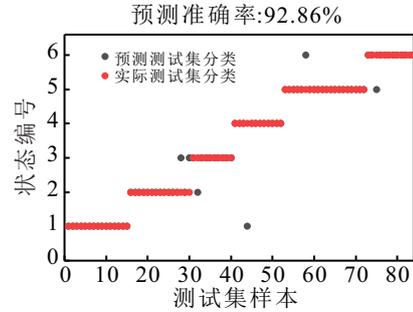


图 15 改进 SSA 优化 MDS-SVM 故障诊断结果

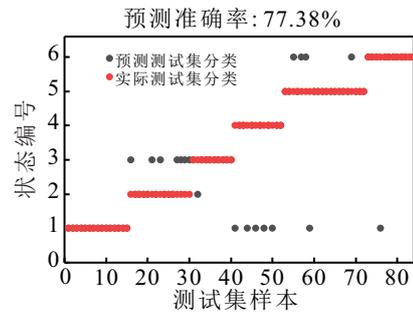


图 16 SSA 优化 MDS-SVM 故障诊断结果

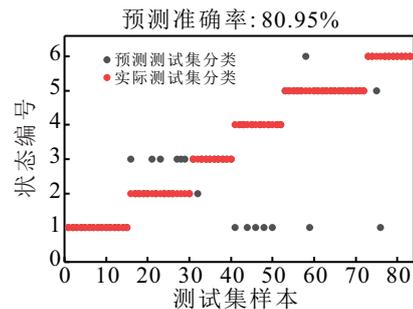


图 17 PSO 优化 MDS-SVM 故障诊断结果

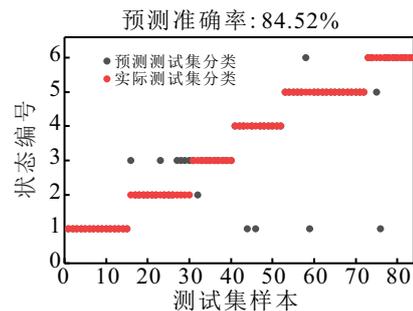


图 18 GWO 优化 MDS-SVM 故障诊断结果

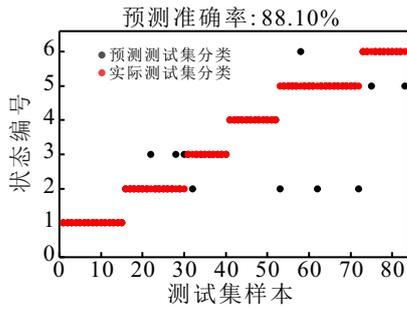


图 19 β -GWO 优化 MDS-SVM 故障诊断结果

通过图 15 至图 19 和表 5 可以得出: 不同算法优化后的 MDS-SVM 故障诊断精度较单一 SVM 均有显著提升, 可见融入 MDS 降维确实能有效的提高故障诊断的精度。

表 5 故障诊断精度对比

| 诊断模型 | 诊断正确率/% | | | | |
|---------|---------|-------|-------|-------|--------------|
| | 改进 SSA | SSA | PSO | GWO | β -GWO |
| SVM | 86.90 | 75.00 | 76.19 | 82.14 | 85.71 |
| MDS-SVM | 92.86 | 77.38 | 80.95 | 84.52 | 88.10 |

4.2 不同故障诊断算法的性能比较

为了进一步验证改进 SSA 优化 MDS-SVM 故障诊断模型的优越性, 本文采用 BP 神经网络 (Back Propagation Neural Network, BPNN)、K 最近邻 (K-Nearest Neighbor, KNN) 以及随机森林 (Random Forest, RF) 故障诊断模型与其进行故障诊断精度对比, 其中, BPNN 算法隐含层神经元个数设为 337, 输出层为 6, 传递函数均采用 logsig 函数, 学习训练算法采用 traingdx 算法, 性能函数最小梯度为 10^{-25} , 学习目标为 10^{-7} , 学习率为 0.1, 训练次数 500 次; KNN 邻点个数设为 1, 距离量度为 cosine, 距离权重为等距离; RF 决策树数量设置成 100, 四种模型故障诊断精度对比如图 20, 可以看出改进 SSA 优化 MDS-SVM 故障诊断模型仍优于其他模型。

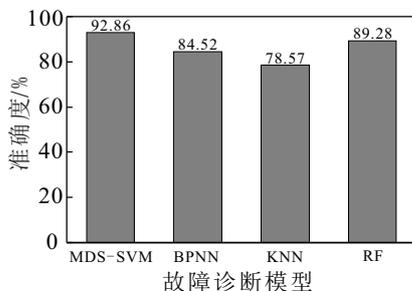


图 20 不同故障诊断模型精度对比图

5 结论

1) 针对变压器故障诊断精确度不高的问题, 利用 MDS 对高维数据进行降维处理, 保证诊断特征信息贡献度的同时, 有效避免了高维数据存在的稀疏

性和多重共线性, 便于 SVM 模型识别数据特征, 提高了识别精度。

2) 鉴于 SSA 算法存在收敛速度慢, 易陷入局部最优的不足, 加入了信赖机制和概率突变, 显著提高了算法的收敛速度和收敛能力, 并且不易陷入局部最优, 最后通过和传统算法进行对比论证了该算法的优越性。

3) 通过系列对比实验, 本文提出的基于改进 SSA 优化 MDS-SVM 的变压器故障诊断方法精确率达到 92.86%, 较传统 PSO-SVM, GWO-SVM, β -GWO-SVM 诊断方法分别提升了 16.67%, 10.72% 和 7.15%, 同时也高于典型 BPNN、KNN 及 RF 故障诊断模型, 具有一定理论意义和工程实用价值。

参考文献 (References)

- [1] 李刚, 于长海, 刘云鹏, 范辉, 文福拴, 宋雨. 电力变压器故障预测与健康管: 挑战与展望 [J]. 电力系统自动化, 2017, 41(23): 156-167.
(Li G, Yu C H, Liu Y P, Fan H, Wen F S, Song Y. Challenges and prospects of fault prognostic and health management for power transformer[J]. Automation of Electric Power Systems, 2017, 41 (23): 156-167.)
- [2] 方涛, 钱晔, 郭灿杰, 宋闯, 王志华, 罗建平, 巴全科. 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, Song C, Wang Z H, Luo J P, Ba Q K. 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.)
- [3] 梁永亮, 郭汉琮, 薛永端. 基于特征气体关联特征的变压器故障诊断方法 [J]. 高电压技术, 2019, 45(02): 386-392.
(Liang Y L, Guo H C, Xue Y D. Transformer fault diagnosis method based on association characteristics of characteristic gases[J]. High Voltage Engineering, 2019, 45 (02): 386-392.)
- [4] 曲岳晗, 赵洪山, 马利波, 赵仕策, 米增强. 多深度神经网络综合的电力变压器故障识别方法 [J/OL]. (2021-01-15)[2021-04-29]. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=ZGDC20210114000&DbName=DKFX2021>.
- [5] 徐阳, 谢天喜, 周志成, 陆云才, 李建生, 曹健, 蒋超. 基于多维度信息融合的实用型变压器故障诊断专家系统 [J]. 中国电力, 2017, 50(01): 85-91.
(Xu Y, Xie T X, Zhou Z C, Lu Y C, Li J S, Cao J, Jiang C. Practical expert system for transformer fault diagnosis based on multi-dimensional information fusion technology[J]. Electric Power, 2017, 50 (01): 85-91.)
- [6] 李鹤健, 徐肖伟, 王科, 赵勇军, 吴世浙, 刘可真. 基于粒子群优化随机森林的变压器故障诊断模型 [J]. 昆明理工大学学报 (自然科学版), 2021, 46(03): 94-101.
(Li H J, Xu X W, Wang K, Zhao Y J, Wu S Z, Liu K

- Z. Transformer fault diagnosis model based on particle swarm optimization and random forest[J]. Journal of Kunming University of Science and Technology(natural sciences),2021,46(03):94-101.)
- [7] 蔡赛男, 宋卫星, 班利明, 齐小刚, 汤润之. 基于鲸鱼算法优化 LSSVM 的滚动轴承故障诊断 [J/OL].(2020-12-02)[2021-11-06].<https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KZYC2020120200Z&DbName=CAPJ2020>.
- [8] 黄大荣, 柯兰艳, 林梦婷, 孙国玺. 一种参数优化 VMD 多尺度熵的轴承故障诊断新方法 [J]. 控制与决策,2020,35(07):1631-1638.
(Huang D R, Ke L Y, Lin M T, Sun G X. A new fault diagnosis approach for bearing based on multi-scale entropy of the optimized VMD[J]. Control and Decision, 2020,35 (07): 1631-1638.)
- [9] 董红召, 王乐恒, 唐伟, 杨强, 余翊妮. 融合时空特征的 PCA-PSO-SVM 臭氧 (O_3) 预测方法研究 [J]. 中国环境科学,2021,41(02):596-605.
(Dong H Z, Wang L H, Tang W, Yang Q, She Y N. Research on PCA-PSO-SVM ozone prediction considering spatial-temporal features[J]. China Environmental Science, 2021,41(02): 596-605.)
- [10] 高新智, 刘作军, 张燕, 陈玲玲. 基于 GWO-SVM 的下肢假肢穿戴者骑行相位识别 [J]. 浙江大学学报(工学版),2021,55(04):648-657.
(Gao X Z, Liu Z J, Zhang Y, Chen L L. Bicycle riding phase recognition of lower limb amputees based on GWO-SVM[J]. Journal of Zhejiang University (Engineering Science), 2021,55 (04): 648-657.)
- [11] 马朝永, 黄攀, 胥永刚, 付胜. 基于 MDS 和神经网络的滚动轴承故障诊断方法 [J]. 噪声与振动控制,2017,37(04):171-174.
(Ma C Y, Huang P, Xu Y G, Fu S. Rolling bearing fault diagnosis method based on MDS and neural network[J]. Noise and Vibration Control, 2017, 37(04): 171-174.)
- [12] 张安安, 杨林, 何嘉辉, 高春林, 李茜. 基于 MDS 的电缆附件局部放电模式识别 [J]. 电子科技大学学报,2019,48(02):202-207.
(Zhang A A, Yang L, He J H, Gao C L, Li X. Pattern recognition for partial discharge of cable accessories based on multidimensional scaling[J]. Journal of university of electronic science and technology of China, 2019, 48 (02): 202-207.)
- [13] S Mirjalili, A H Gandomi, S Z Mirjalili. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems[J]. Advances in Engineering Software , 2017, 114(6):163-191.
- [14] 张达敏, 陈忠云, 辛梓芸, 张绘娟, 闫威. 基于疯狂自适应的樽海鞘群算法 [J]. 控制与决策,2020,35(09):2112-2120.
(Zhang D M, Chen Z Y, Xin Z Y, Zhang H J, Yan W. Salp swarm algorithm based on craziness and adaptive[J]. Control and Decision , 2020,35 (09): 2112-2120.)
- [15] 俞家珊, 吴雷. 双领导者樽海鞘群算法 [J]. 计算机科学,2021,48(04): 254 - 260.
(Yu J S, Wu L. Two types of leaders salp swarm algorithm[J]. Computer Science, 2021,48 (04): 254-260.)
- [16] 刘景森, 袁蒙蒙, 左方. 面向全局搜索的自适应领导者樽海鞘群算法 [J]. 控制与决策,2021,36(09):2152-2160.
(Liu J S, Yuan M M, Zuo F. Global search-oriented adaptive leader salp swarm algorithm[J]. Control and Decision , 2021,36 (09): 2152-2160.)
- [17] 张轶议, 焦健, 汪可, 郑含博, 房加珂, 周浩. 基于帝国殖民竞争算法优化支持向量机的电力变压器故障诊断模型 [J]. 电力自动化设备,2018,38(01):99-104.
(Zhang Y Y, Jiao J, Wang K, Zheng H B, Fang J K, Zhou H. Power transformer fault diagnosis model based on support vector machine optimized by imperial colonial competitive algorithm[J]. Electric Power Automation Equipment, 2018, 38 (01): 99-104.)
- [18] 田书欣, 周全, 程浩忠, 柳璐, 路亮, 江粟. 基于鸽群优化算法的支持向量机在电力需求总量预测中的应用 [J]. 电力自动化设备,2020,40(05):173-181.
(Tian S X, Zhou Q, Cheng H Z, Liu L, Lu L, Jiang L. Application of pigeon-inspired optimization algorithm based SVM in total power demand forecasting[J]. Electric Power Automation Equipment, 2020, 40(05): 173-181.)
- [19] 李赢, 舒乃秋. 基于模糊聚类和完全二叉树支持向量机的变压器故障诊断 [J]. 电工技术学报,2016,31(04):64-70.
(Li Y, Shu N Q. Transformer fault diagnosis based on fuzzy clustering and complete binary tree support vector machine[J]. Transactions of China Electrotechnical Society, 2016, 31 (04): 64-70.)

作者简介

谢国民(1969—), 男, 副教授, 博士, 从事电力系统智能检测与预测控制等研究, E-mail: 804764995@qq.com;

蔺晓雨(1995—), 男, 硕士生, 从事电力系统智能检测与预测控制的研究, E-mail: 2691651925@qq.com;