

AP 聚类改进免疫算法用于航空发动机故障诊断

曹愈远, 张博文, 李艳军

(南京航空航天大学 民航学院, 南京 211106)

摘 要: 在免疫算法训练过程中引入近邻传播(AP)聚类与熵权法,对训练样本进行聚类与权值计算,将权值引入免疫算法中样本选择阈值的计算,以解决训练过程采用固定选择阈值所造成的检测器在部分区域过拟合,部分区域欠拟合的问题。结果表明:改进的免疫算法用于典型非线性函数的寻优时,迭代性能均优于传统免疫算法,并在大部分情况下优于粒子群算法与量子遗传算法,在进行某型发动机故障诊断的实例实验时,改进后的算法的诊断准确率达到 98.06%,高于传统免疫算法的 92.60%。

关 键 词: 故障诊断; 近邻传播(AP); 免疫算法; 熵权法; 混沌理论

中图分类号: V263.6; TH17; TK148

文献标志码: A

AP clustering improved immune algorithm for aeroengine fault diagnosis

CAO Yuyuan, ZHANG Bowen, LI Yanjun

(College of Civil Aviation,

Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: In the process of immune algorithm training, the affinity propagation (AP) clustering and entropy weight method were introduced, the training samples were clustered and weighted, and the weights were introduced into the calculation of the sample selection threshold in the immune algorithm to solve the problem of a fixed selection threshold in the training process, which led to over fitting of the detector in a partial area, and under-fitting of the partial area. Result showed that, when the improved immune algorithm was used for the optimization of typical nonlinear functions, the iterative performance was better than the traditional immune algorithm. In most cases it was better than the particle swarm optimization algorithm and the quantum genetic algorithm, in the case of a certain type of engine fault diagnosis. The improved algorithm had a diagnostic accuracy of 98.06%, which was higher than 92.60% of the traditional immune algorithm.

Key words: fault diagnosis; affinity propagation(AP); immune algorithm; entropy weight method; chaos theory

收稿日期: 2019-01-16

基金项目: 航空科学基金(20153352040)

作者简介: 曹愈远(1983—),男,讲师,博士生,主要研究方向为航空器故障诊断、航空器运行安全风险分析与控制技术。

E-mail: caoyuyuan@nuaa.edu.cn

引用格式: 曹愈远,张博文,李艳军. AP 聚类改进免疫算法用于航空发动机故障诊断[J]. 航空动力学报, 2019, 34(8): 1795-1804.
CAO Yuyuan, ZHANG Bowen, LI Yanjun. AP clustering improved immune algorithm for aeroengine fault diagnosis[J]. Journal of Aerospace Power, 2019, 34(8): 1795-1804.

根据人体免疫系统发展起来的人工免疫算法具有学习、记忆、分布式、并行性等特点,较好的解决了故障诊断领域的一些问题. 目前大多数面向免疫算法的研究主要是改进免疫算子策略. 文献[1-11]中已经将多种免疫算子进行了优化改进,成功提出了基于免疫球蛋白、贪婪算法和删除交叉算子、最小平均距离免疫场等改进免疫算法,文献[12-15]考虑了样本权重的差异性,采用聚类等相关算法得到了样本的权重系数并成功的应用于对遗传算法的改进优化.

本文提出基于近邻传播(affinity propagation, AP)优化混沌免疫克隆算法的诊断模型. 通过对样本空间中具有相似分布特性的样本进行聚类,并引入熵权法计算样本权值^[16-18],最后采用混沌免疫算法进行故障诊断.

1 航空发动机磨损故障分析

磨损故障主要来源于发动机在运行过程中,轴承和齿轮等旋转零部件在相互运动过程中产生的摩擦. 随着磨损的加深,零部件间会在不被察觉的情况下发生损坏,进一步造成发动机的故障,得到磨损情况征兆较为常用的方法是油液分析技术,通过人工、仪器等手段判别油液的成分构成、油品品质等信息,主要的技术手段包括铁谱分析、光谱分析和磁塞监控技术等.

利用仪器获取油液中磨粒质量浓度的原理在于,发动机正常磨损工作状态下,油液中的磨粒质量浓度与粒度处于一个动态平衡;而当发生了异

常磨损时,磨粒质量浓度与粒度分布会超过监测指标. 若假设初始状态磨粒质量浓度为零,可以建立如下的特定尺寸磨粒质量浓度的公式:

$$\rho_i = \frac{P_i}{K} \left(1 - e^{-\frac{A}{V}}\right) = \frac{P_i}{EQ} \left(1 - e^{-\frac{EQ}{V}}\right) \quad (1)$$

式中 ρ_i 为特定尺寸粒度 i 的磨粒质量浓度; P_i 为特定尺寸粒度 i 的磨损率; A 为排除率常数; K 排除率常数; V 为油液体积; t 为时间; Q 为流量; E 为油滤出效率. 该式表明磨粒质量浓度正比于磨损率,在正常工作的情况下,磨损率是常数,磨粒质量浓度以指数形式达到平衡,只有当发生异常磨损时,磨粒质量浓度会因为磨损常数的变化而发生改变,可以通过监测油液中特定尺寸磨粒质量浓度来鉴别故障状态.

2 基于 AP 聚类的样本空间约简

AP 聚类算法在待聚类数据对象之间传递相似度和代表度,相似度矩阵用 S 表示, $S(i, j)$ 表示数据点 i 与 j 之间的相似度,通常采用欧氏距离进行度量 $S(i, j) = -\|x_i - x_j\|^2$, $S(k, k)$ 被称为偏向性参数,通常采取对应行的中位数进行表示. 代表度用矩阵 R 表示, $R(i, j)$ 表示从数据点 i 传递到数据的 j 的信息,表示点 j 作为点 i 的代表点的合适程度. 有效性用矩阵 A 表示, $A(i, j)$ 表示点 i 选择点 j 作为代表点的有效程度. AP 算法通过迭代更新每个数据点的代表度与有效度,最终可以得到一个收敛的矩阵,代表度矩阵的更新规则如下:

$$\begin{cases} R(i, j) = S(i, j) - \max\{A(i, j') + S(i, j')\} & i \neq j \\ R(j, j) = S(j, j) - \max\{A(j, j') + S(j, j')\} & i = j \end{cases} \quad (2)$$

有效度矩阵更新规则如下:

$$\begin{cases} A(i, j) = \min\{0, R(j, j) + \sum_{i' \neq \{i, j\}} \max\{0, r(i'j)\}\} & i \neq j \\ A(j, j) = \sum_{i \neq j} \max\{0, r(i'j)\} & i = j \end{cases} \quad (3)$$

决策矩阵为 $A+R$,在矩阵收敛时,将决策矩阵中的对角线元素大于零的数据点作为聚类中心,其余的点按照决策矩阵中最大值所在位置依次分配到各个聚类中心. 为了保证决策矩阵的稳定收敛,引入阻尼因子 λ ,迭代公式变更如下:

$$\begin{cases} R_{i+1} = \lambda R_{i+1} + (1 - \lambda) R_i \\ A_{i+1} = \lambda A_{i+1} + (1 - \lambda) A_i \end{cases} \quad (4)$$

3 基于熵权法的权值计算

通过 AP 聚类算法得到了样本空间的聚类结

果,每一个聚类中心作为免疫克隆算法的训练样本. 每个聚类中心选取 4 个指标作为免疫算法训练检测器时选择阈值的确定依据,分别是类内最大距离 $L_{\max}$ 、类内最小距离 $L_{\min}$ 、类内平均距离 L_{mean} 和类内数据集数 L ,采用熵权法确定权值的具体步骤如下:

1) 归一化处理,将 4 项指标的数值范围都归一化到 $[0, 1]$ 区间.

2) 计算第 q 个聚类中心的第 s 个指标的比重

$$y_{qs} = \frac{x_{qs}}{\sum_{q=1}^m x_{qs}} \quad (5)$$

3) 计算第 q 个聚类中心的第 s 个指标的信息熵

$$e_{qs} = -M \sum_{q=1}^m y_{qs} \ln y_{qs} \quad (6)$$

式中 M 为常数 $\frac{1}{\ln m}$, 如果 y_{qs} 为 0, 则 $y_{qs} \ln y_{qs}$ 的值也取零。

4) 计算第 q 个聚类中心的第 s 个指标的权值

$$W_{qs} = \frac{1 - e_{qs}}{N - \sum_{s=1}^N e_{qs}} \quad (7)$$

式中 N 为指标个数。按照上述步骤可以依次求解出每一个聚类中心各项指标的权值系数 W_q , 该系数可以反映出聚类中心的训练权重。

4 基于混沌理论的免疫克隆算法

引入混沌理论对克隆选择算法优化, 步骤如下文。

1) 采用混沌理论初始化检测器, 混沌理论的主要思想是通过迭代在一区间类进行无固定周期的循环行为, 迭代公式如下:

$$X_{k+1} = \eta X_k (1 - X_k) \quad (8)$$

式中 k 为混沌迭代次数, η 为控制参数, 用来控制序列进入混沌状态的速率, 根据文献[19], 一般取 [3, 4] 间的随机数。对种群中的每一项特征参数按照式(8)进行混沌初始化, 得到抗体的初始种群。

2) 计算种群中抗体的适应度, 计算每一抗体与抗原间的亲和力距离, 常用的亲和力距离度量有欧氏距离、切比雪夫距离、海明距离等, 本文提出一种距离度量方式, 计算公式如下:

$$H_{m,n} = \frac{1}{e^{\|X_m - X_n\|^2}} \quad (9)$$

该式表示数据点 m 与数据点 n 之间的亲和力, 引入指数表示形式来改进传统的欧式亲和力距离计算。

3) 将当前种群中满足亲和力选择阈值要求的抗体筛选出来, 存入记忆库中。该步骤包含两个过程。

第一个过程为选出满足选择阈值要求的抗体, 传统的选择阈值是固定的, 对于分布存在差异的训练样本, 容易造成检测器部分区域过拟合, 部分区域欠拟合的不平衡现象。本文提出使用熵权法来自适应的确定选择阈值, 每一个样本的选择

阈值按照下式进行计算:

$$p = p_1 \alpha + p_2 (1 - \alpha) \quad (10)$$

式中 α 为采用熵权法对每一样本点确定出的权值, p_1 、 p_2 为选择阈值设定范围, $p_1 < p_2$ 。

第二个过程为鉴别筛选出来的抗体与记忆库中已有抗体间的相似度, 采用式(9)计算抗体与抗体间的亲和力, 将大于抑制阈值 ϵ 的抗体从记忆库中删除, 这一过程的引入可以有效地避免部分区域的过拟合, 使检测器可以较为均匀的布满整个样本空间。

4) 对种群进行克隆操作, 计算每一个抗体需要克隆的数量, 按照下式进行计算:

$$n_i = \text{Int} \left\{ \frac{F_i}{\sum_{i=1}^{S_{\text{pop}}} F_i} N_c \right\} \quad (11)$$

式中 Int 为上取整函数, F_i 为第 i 个抗体的适应度, S_{pop} 为种群容量, N_c 为控制常数, 一般取 10。通过式(11)得到每一抗体的克隆数目, 进行克隆得到高维种群。

5) 对高维种群进行混沌变异操作, 本文结合文献[19], 提出一种变异策略, 具体过程如下:

对父代高维种群 $\{C_1, C_2, \dots, C_n\}$ 中的每一抗体, 以 C_i 为初始值生成一混沌序列, 混沌序列的长度取高维种群的种群容量, 分别进行变异操作。

(1) 混沌正变异。按照式(12)生成新的抗体

$$C_{\text{new}} = C + (Y - C) X(N) e^{\frac{-S}{S_{\text{pop}}}} \cdot \frac{P}{S_{\text{pop}}} \quad (12)$$

式中 Y 为控制参数, 取当前抗原, X 为生成的混沌序列, N 为一小于混沌序列长度的随机整数值, S 为当前记忆库抗体个数, J_{pop} 为记忆库容量, P 为当前抗体的适应度在种群中的排名。

(2) 混沌负变异。按照式(13)生成新的抗体

$$C_{\text{new}} = C - (Y - C) X(N) e^{\frac{-S}{S_{\text{pop}}}} \cdot \frac{P}{S_{\text{pop}}} \quad (13)$$

参数意义与混沌正变异一致。

(3) 混沌混合变异。按照式(14)计算变异概率 q

$$q = e^{\frac{S_{\text{pop}} - 1}{S_{\text{pop}} + P}} \quad (14)$$

式中 S_{pop} 为种群容量, P 为当前抗体的适应度在种群中的排名, 如果产生的随机数大于 q , 进行混沌正变异, 否则进行混沌负变异。

采用 3 种混沌变异策略生成了 3 个不同的变异后个体, 选择 3 个变异后个体与父代个体中适应度最大的值存入高维种群中。

6) 选择适应度最优的前 S_{pop} 个抗体存入新

种群,判断是否满足停止条件,若满足则停止;否则回到第 2)步。

5 实例验证

5.1 混沌免疫算法性能验证^[2]

为了验证本文提出的混沌免疫克隆算法的优劣,对表 1 中的三维空间典型非线性函数进行寻优。表 1 中的 7 个函数 $f_1 \sim f_7$ 的最小值都为 0,函数的维数 D 均取 3。为了适应于对函数进行寻优,对本文算法做出如下修改:

1) 在寻优过程中,变异策略中的抗原 Y 未知,使用 $[0,1]$ 间的随机数 R_{rand} 代替。

2) 寻优过程不需要生成检测器,将变异策略中的当前记忆库容量 S 置为 0。

在种群容量 S_{pop} 分别取 50、100、150、200,对 7 个函数各运行 30 次,以找到种群容量对函数性能的影响,确定最合适的种群容量。分别取运行 30 次中的最优、最差以及平均寻优结果,得到表 2。当函数的维度发生改变时,需要重新进行试错实验选择最合适的种群容量。

表 1 测试函数
Table 1 Test function

函数名称	表达式	最小值
Sphere	$f_1 = \sum_{i=1}^D x_i^2$	0
Quadrio	$f_2 = \sum_{i=1}^D \left(\sum_{j=1}^i x_j \right)^2$	0
Sohwefel 2.22	$f_3 = \sum_{i=1}^D x_i + \prod_{i=1}^D x_i$	0
Step	$f_4 = \sum_{i=1}^D (x_i + 0.5)^2$	0
Rastrigin	$f_5 = \sum_{i=1}^D [x_i^2 - 10\cos(2\pi x_i) + 10]$	0
Griewank	$f_6 = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	0
Ackley	$f_7 = -20\exp\left\{-0.2\sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} - \exp\left[\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right]\right\} + 20 + e$	0

表 2 实验结果
Table 2 Experiment results

函数	$s_{\text{pop}}=50$			$s_{\text{pop}}=100$			$s_{\text{pop}}=150$			$s_{\text{pop}}=200$		
	最优值	最差值	平均值	最优值	最差值	平均值	最优值	最差值	平均值	最优值	最差值	平均值
$f_1/10^{-5}$	4.93	7.23	6.71	0.825	12.2	3.38	2.15	4.24	3.06	1.57	28.8	13.5
$f_2/10^{-5}$	3.26	96.1	15.8	2.09	17.7	8.31	9.56	49.0	21.6	4.31	35.2	9.21
$f_3/10^{-3}$	3.20	9.10	5.20	0	9.80	0.527	0	6.70	0.423	0	9.70	0.343
$f_4/10^{-2}$	4.42	6.99	6.91	5.48	29.7	14.9	5.60	34.0	16.8	12.2	52.0	29.1
$f_5/10^{-2}$	0	20.3	8.77	0	103	4.43	0	101	3.37	0	100	3.33
$f_6/10^{-7}$	0	273	72.4	0	61.8	3.99	31.9	481	147	38.3	697	175
$f_7/10^{-2}$	0.910	5.15	2.27	1.19	3.89	2.26	1.03	3.63	1.84	1.03	6.05	2.18

从表 2 中的数据可以得出如下结论:

1) 种群容量对于算法的寻优性能有影响,随着种群容量的增加,寻优性能有所提高。

2) 在本文进行验证的几个函数上,当种群容

量达到 100 时,算法的寻优性能有较大的改观,容量继续增加,性能提升不大。

3) 文中算法的寻优空间为三维,因此在三维空间进行寻优时,将种群容量设定为 100 是合

适的。

将本文算法的种群容量置为 100,分别用粒子群算法、免疫克隆算法、量子遗传算法和本文算法对 7 个函数进行寻优。免疫克隆算法采用随机变异策略,即在克隆后的种群中进行随机赋值变异,取适应度高的存入下代种群,几种算法的种群容量均取 100。不同的算法在 7 个验证函数上的迭代图如图 1~图 7 所示。

图 1~图 7 中横坐标表示迭代次数,纵坐标表示适应度,从图 1~图 7 中可以得到如下结论:

- 1) 在大多数函数(f_1 、 f_2 、 f_3 、 f_5 、 f_6 、 f_7)上,本文提出的算法表现均优于粒子群算法与量子遗传算法。
- 2) 相比于随机变异策略,本文提出的算法在 7 个验证函数上均表现更为优秀。
- 3) 算法的收敛曲线非常平滑,没有突变,对于免疫算法生成均匀检测器是有益的。
- 4) 对于寻优范围在 $[0,1]$ 内的情况,本文提

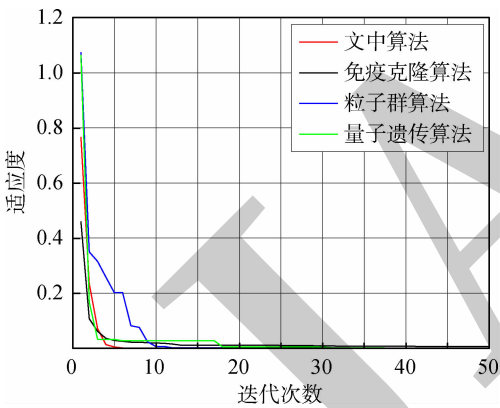


图 1 f_1 函数迭代图
Fig. 1 f_1 function iteration graph

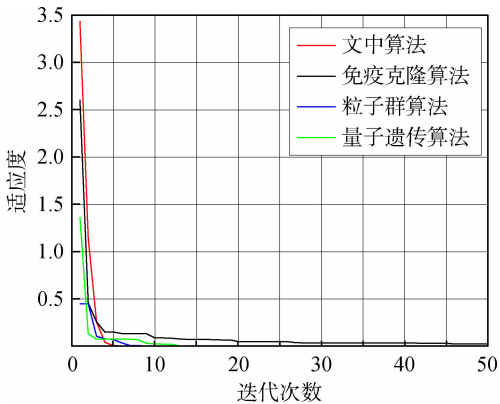


图 2 f_2 函数迭代图
Fig. 2 f_2 function iteration graph

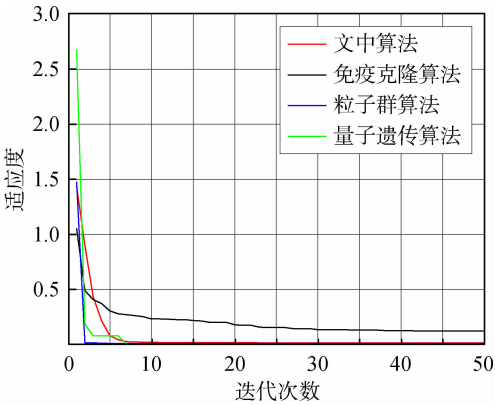


图 3 f_3 函数迭代图
Fig. 3 f_3 function iteration graph

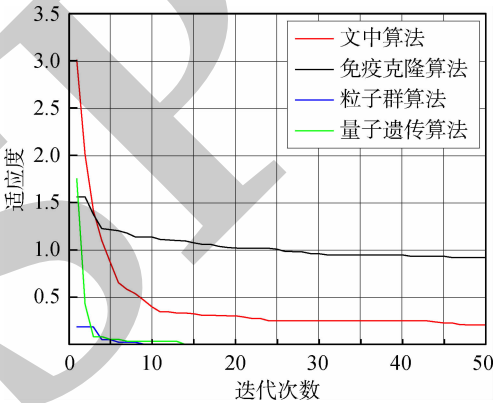


图 4 f_4 函数迭代图
Fig. 4 f_4 function iteration graph

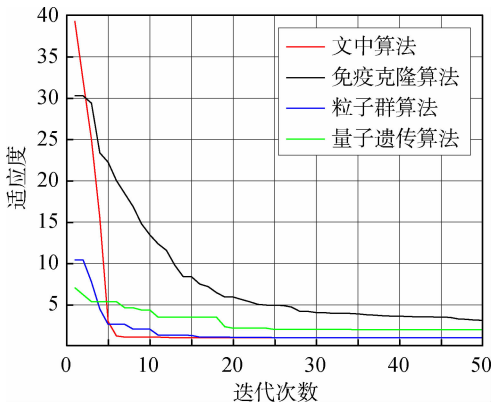


图 5 f_5 函数迭代图
Fig. 5 f_5 function iteration graph

出的算法具有较好的收敛性能。

5.2 航空发动机故障故障诊断实例验证

将算法用于发动机磨损故障诊断,具体步骤如下:

- 1) 引入 K-CV(K-fold cross-validation)方法

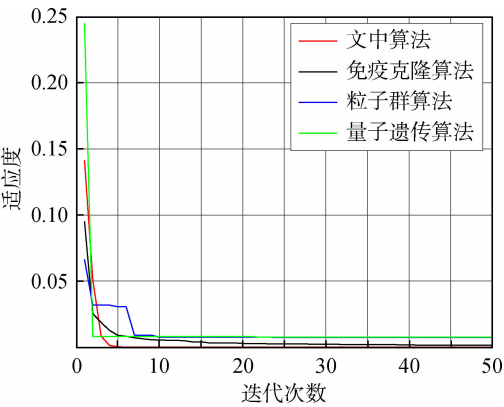


图 6 f_6 函数迭代图
Fig. 6 f_6 function iteration graph

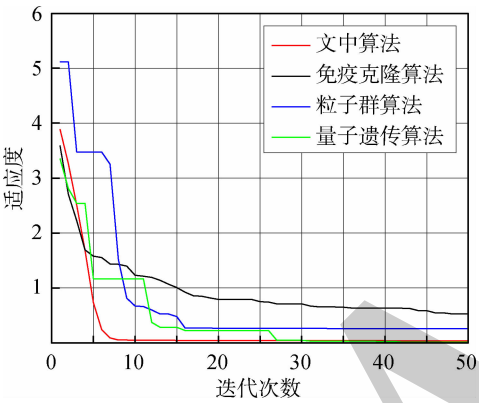


图 7 f_7 函数迭代图
Fig. 7 f_7 function iteration graph

将样本分为 5 组,依次选取其中的 4 组作为训练样本,剩下的一组作为测试样本。

2) 引入 AP 聚类对训练样本进行聚类,得到聚类后的聚类中心,以及类内最小距离、最大距离和平均距离,以及类内数据点个数。

3) 引入熵权法确定每一个聚类中心的权值系数。

4) 对每一个聚类中心求解出自适应的选择阈值,并采用改进的混沌免疫算法确定每一个聚类中心对应的成熟抗体,存入记忆库中。

5) 训练完所有的聚类中心后得到成熟的抗体检测器,采用测试样本对诊断模型进行测试。整个过程的流程图如图 8 所示。

采集到某型发动机处于正常状态、警告状态、异常状态的试车滑油共 260 组油液磨粒质量浓度数据样本,如表 3。

5.2.1 故障与非故障诊断

对数据进行[0,1]归一化,将样本数据分为故障样本与正常样本,其中故障样本包含多种故障

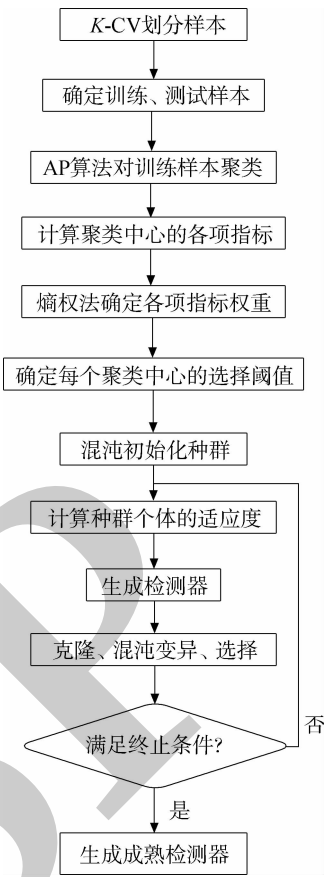


图 8 算法流程图
Fig. 8 Flow chart of algorithm

类型,采用 K-CV 准则将正常样本与故障样本都随机均分为 5 组,依次选取其中的 4 组作为训练样本,剩余的一组作为测试样本,首先对样本是否为故障样本进行诊断。

对正常样本训练集进行 AP 聚类,阻尼因子 $\lambda=0.3$,最大迭代次数取 50,在 $K=5$ 时得到的正常样本训练集聚类结果如图 9。

对故障样本训练集进行 AP 聚类,参数设置如上,在 $K=5$ 时得到的故障样本训练集聚类结果如图 10。

图 10 中三角形为聚类中心,圆形为样本训练集。分别对正常样本聚类中心与故障样本聚类中心进行混沌免疫算法训练。为了保证检测器空间的成熟样本数足够,同时考虑训练时间,多次实验算后记忆库容量取 500,因为是三维空间,根据前文,种群容量可以取 100,对选择阈值范围在[0,1]间进行遍历,得到阈值范围为[0.8,0.9]时较为合适,克隆参数 N_c 取 10,混沌控制参数 η 取[3,4]之间随机数,多次实验得到最优抑制阈值 ϵ 约为 1, $K=5$ 时正常状态检测器与故障状态训练成熟检测器在样本空间中的分布如图 11 所示。

表 3 样本数据
Table 3 Sample data

状态	编号	磨粒个数/(个/mL)		
		切削磨粒	滑动磨粒	疲劳磨粒
正常状态	1	162	1144	3160
	2	29	179	1351
	⋮	⋮	⋮	⋮
	139	125	280	1790
	140	102	420	1958
警告状态	1	828	895	550
	2	611	2634	3465
	⋮	⋮	⋮	⋮
	69	681	580	8253
	70	653	2967	4362
异常状态	1	667	471	8996
	2	375	3042	3339
	⋮	⋮	⋮	⋮
	49	817	3024	12423
	50	231	3730	4248

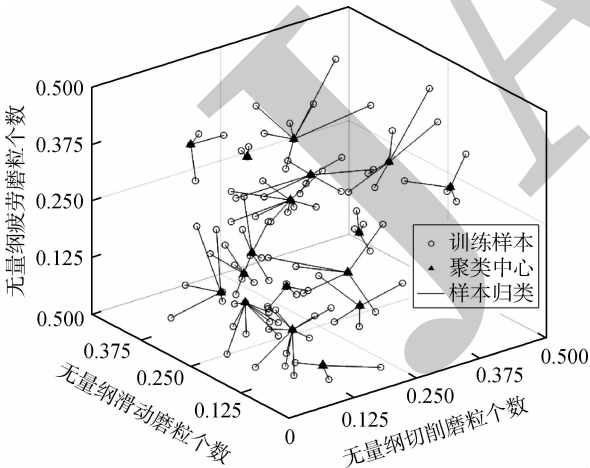


图 9 正常样本聚类结果

Fig. 9 Normal sample clustering results

图 11 中圆形为正常样本检测器,三角形为故障样本检测器。采用训练完毕得到的成熟检测器对训练样本和测试样本进行诊断,诊断阈值取 0.7,进行诊断时,计算待检样本与记忆库中抗体的亲和力,选出与该样本亲和力大于诊断阈值的抗体,其中占最大比例的那一类抗体所对应的抗原即为该待检样本的状态。若有多类抗体所占比

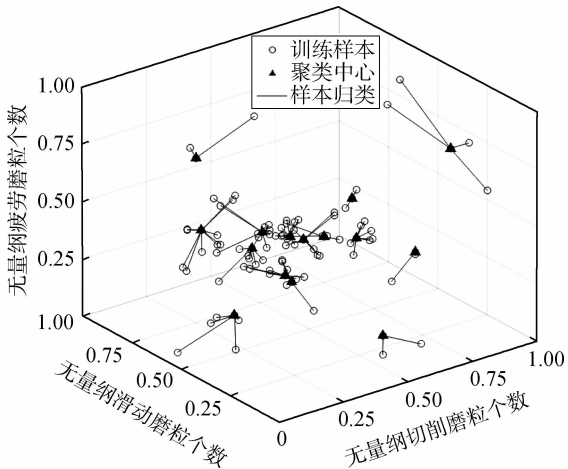


图 10 故障样本聚类结果

Fig. 10 Fault sample clustering results

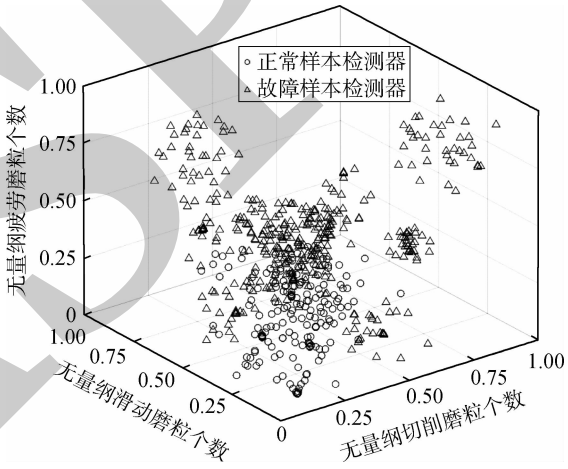


图 11 故障与非故障检测器分布

Fig. 11 Fault and non-fault detector distribution

例相同,则选择平均亲和力最大的那类抗体对应的抗原为待检样本状态。评价该算法有效性的指标有故障检测率 t 和误报率 f ,计算方法分别如下:

故障检测率 t 为

$$t = \frac{F_t}{F_t + F_f} \tag{15}$$

式中 F_t 表示表示故障样本被正确检测为故障的个数; F_f 表示故障样本被检测为正常样本的个数。

误报率 f 为

$$f = \frac{T_t}{T_t + T_f} \tag{16}$$

式中 T_t 表示表示正常样本被检测为正常样本的个数; T_f 表示正常样本被检测为故障样本的个数,采用 K -CV 验证方法得到最终结果如表 4,表中 F 代表误报率,计算如式(16), T 代表故障检测率,计算如式(15), K 代表测试样本组序号。

表 4 故障与非故障诊断准确率

Table 4 Fault and non-fault diagnosis accuracy

K	训练集准确率/%		测试集准确率/%		训练 时间/s
	F	T	F	T	
1	100.000	98.958	100.000	100.000	18.957
2	100.000	100.000	100.000	100.000	16.922
3	98.214	100.000	100.000	100.000	18.222
4	99.107	100.000	96.429	100.000	18.394
5	99.107	100.000	96.429	95.833	16.616
平均值	99.286	99.792	98.572	99.167	17.822

5.2.2 故障类型诊断

完成了故障与非故障诊断后,对于诊断出的故障样本进行故障类型诊断,与故障非故障样本诊断不同,在进行故障类型诊断时,只选取故障样本进行训练。

对 II 类故障样本训练集进行 AP 聚类,相关设置不变在 $K=5$ 时得到的正常样本训练集聚类结果如图 12。

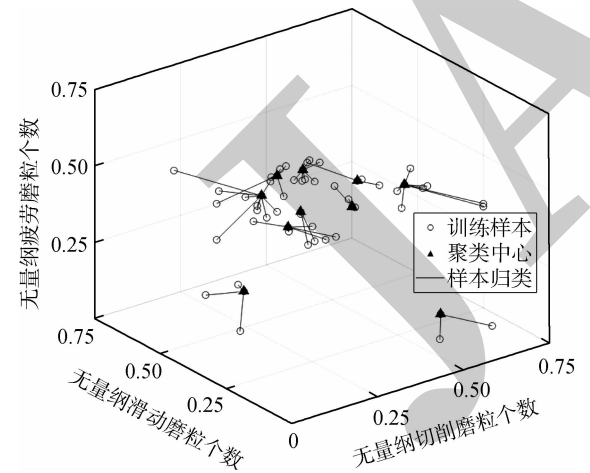


图 12 II 类故障样本聚类结果

Fig. 12 Class II fault sample clustering results

对 III 类故障样本训练集进行 AP 聚类,在 $K=5$ 时得到的故障样本训练集聚类结果如图 13。

与前文类似同样可以得到两类故障样本的成熟检测器在样本空间中的分布如图 14。

图 14 中圆形为 II 类故障样本检测器的分布,三角形为 III 类故障样本检测器的分布。采用训练完毕得到的成熟检测器对训练样本和测试样本进行诊断,诊断阈值取 0.8,评价该算法有效性的指

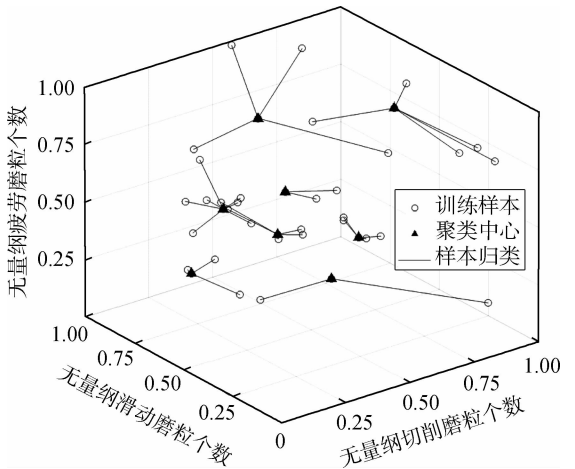


图 13 III 类故障样本聚类结果

Fig. 13 Class III fault sample clustering results

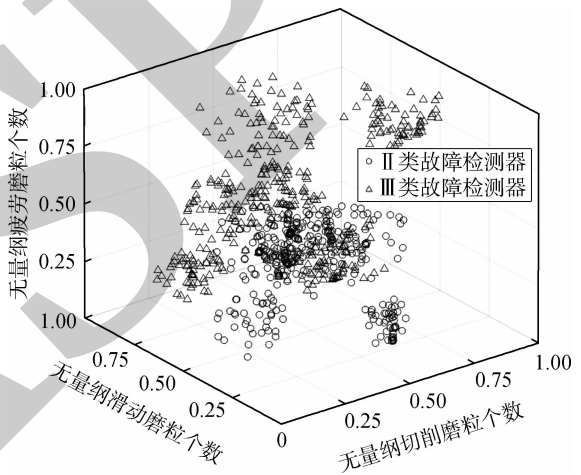


图 14 故障类型检测器分布

Fig. 14 Fault type detector distribution

标为故障检测率 t ,计算方法分别如下:

故障检测率 t

$$t = \frac{F_r}{F_r + F_e} \tag{17}$$

式中 F_r 表示故障样本被正确检测为对应故障类型的个数; F_e 表示故障样本被检测为错误故障类型的个数,采用 K -CV 验证方法得到最终结果如表 5, T_1 代表 II 类故障检测率, T_2 代表 III 类故障检测率,计算如式(17)。

5.2.3 不同诊断模型性能对比

分别采用免疫算法 1(限制记忆库容量)、免疫算法 2(不限制记忆库容量)、支持向量机 SVM (support vector machines)、免疫算法优化支持向量机以及本文提出的算法进行故障诊断,不同诊断模型在 K -CV 验证方法下得到的测试集平均准确率如表 6。

表 5 故障类型诊断准确率

Table 5 Fault type diagnosis accuracy

K	训练集准确率/%		测试集准确率/%		训练 时间/s
	T ₁	T ₂	T ₁	T ₂	
1	100.000	100.000	100.000	100.000	13.597
2	100.000	92.500	100.000	100.000	14.136
3	100.000	97.500	100.000	100.000	14.155
4	100.000	100.000	92.500	100.000	13.510
5	100.000	97.500	100.000	80.000	13.486
平均值	100.000	97.500	98.500	96.000	13.777

表 6 不同模型诊断准确率

Table 6 Diagnostic accuracy of different models

诊断模型	测试集 准确率/%	训练 时间/s	记忆库 容量/个
免疫算法 1	92.60	33.66	500
免疫算法 2	94.50	49.40	>4000
支持向量机	92.90	9.83	
免疫 SVM	97.80	12.45	
文中模型	98.06	31.60	500

与传统的免疫算法相比,在记忆库容量受限时,本文提出的算法在诊断准确率上有一定的提升;在不限制记忆库容量时,传统的免疫算法的诊断准确率有提升,但最终得到的成熟检测器的样本数量极大,耗费时间也较长。与其他几种故障诊断算法相比,本文提出的算法也有较为不错的表现。

6 结 论

本文主要完成了以下工作:

- 1) 引入混沌理论改进传统免疫克隆算法,在不同种群下使用 7 个典型的非线性函数进行测试,确定了种群容量对函数寻优性能的影响。
- 2) 对比了改进的免疫算法与传统免疫算法、粒子群算法和量子遗传算法在性能上的差异,发现改进的免疫算法在多数情况下具有较优的表现。
- 3) 引入 AP 聚类算法来进行训练样本约简,引入熵权法来确定约简后样本的权值,将权值引入免疫算法进行检测器生成。
- 4) 使用本文的诊断模型与几种常用的诊断

模型进行某型航空发动机的故障诊断,本文的模型表现出了较优的诊断性能,可以在记忆库容量一定的情况下有效地提高发动机故障诊断的准确率与计算速度,为其他故障诊断方法的优化提供了参考。

参考文献:

[1] 程超,汪久根. 双列角接触球轴承的免疫算法优化设计[J]. 航空动力学报,2015,30(11):2810-2816.
CHENG Chao,WANG Jiugen. Optimization design of immune algorithm for double row angular contact ball bearings[J]. Journal of Aerospace Power,2015,30(11):2810-2816. (in Chinese)

[2] 宋丹,樊晓平,文中华. 模糊非基因信息记忆的双克隆选择算法[J]. 电子与信息学报,2017,39(2):255-262.
SONG Dan,FAN Xiaoping,WEN Zhonghua. Double clonal selection algorithm for fuzzy non-genetic information memory[J]. Journal of Electronics and Information Technology,2017,39(2):255-262. (in Chinese)

[3] 张弛,贾丽媛,王加阳. 改进的混合免疫算法在约束函数优化中的应用[J]. 中南大学学报,2016,47(6):1940-1946.
ZHANG Chi,JIA Liyuan,WANG Jiayang. Application of improved hybrid immune algorithm in constraint function optimization[J]. Journal of Central South University,2016,47(6):1940-1946. (in Chinese)

[4] PAN Guo,LI Kenli,OUYANG Aijia. Hybrid immune algorithm based on greedy algorithm and delete-cross operator for solving TSP[J]. Soft Computing,2016,20(2):555-566.

[5] LIU Xufei,CHUNG T. A modified immunoglobulin-based artificial immune system algorithm for solving the permutation flow shop scheduling problem[J]. Journal of Industrial and Production Engineering,2017,34(7):542-550.

[6] YU Xiao,ZHOU Zijie,RIHA K. Blurred infrared image segmentation using new immune algorithm with minimum mean distanceimmune field[J]. Spectroscopy and Spectral Analysis,2018,38(11):3645-3652.

[7] 魏明军,王月月,金建国. 一种改进免疫算法的入侵检测设计[J]. 西安电子科技大学学报,2016,43(2):126-131.
WEI Mingjun,WANG Yueyue,JIN Jianguo. An intrusion detection design based on improved immune algorithm[J]. Journal of Xidian University,2016,43(2):126-131. (in Chinese)

[8] ZHANG Weiwei,GAO Kui,ZHANG Weizheng. A hybrid clonal selection algorithm with modified combinatorial recombination and success-history based adaptive mutation for numerical optimization[J]. Applied Intelligence,2019,49(2):819-836.

[9] ZHANG Hanye. An improved immune algorithm for simple assembly line balancing problem of type 1[J]. Journal of Algorithms and Computational Technology,2017,11(4):317-326.

[10] GOSCINIAK I. Discussion on semi-immune algorithm behaviour based on fractal analysis[J]. Soft Computing, 2017,21(14):3945-3956.

[11] 李海林,魏苗. 基于近邻传播的限定簇数聚类方法研究[J]. 电子科技大学学报,2018,47(5):733-739.
LI Hailin, WEI Miao. Research on clustering method of limited cluster number based on neighbor propagation[J]. Journal of University of Electronic Science and Technology of China,2018,47(5):733-739. (in Chinese)

[12] 徐青山,娄藕蝶,郑爱霞,等. 基于近邻传播聚类和遗传优化的非侵入式负荷分解方法[J]. 电工技术学报,2018,33(16):3868-3878.
XU Qingshan, LOU Oudie, ZHENG Aixia, et al. Non-intrusive load decomposition method based on neighbor propagation clustering and genetic optimization[J]. Transactions of China Electrotechnical Society, 2018, 33(16): 3868-3878. (in Chinese)

[13] 张伟,许爱强,平殿发,等. 基于近邻传播聚类的航空电子部件 LMK 诊断模型[J]. 北京航空航天大学学报,2018,44(8):1693-1704.
ZHANG Wei, XU Aiqiang, PING Dianfa, et al. Diagnostic model of avionics components based on near-neighbor propagation clustering[J]. Journal of Beijing University of Aeronautics and Astronautics,2018,44(8):1693-1704. (in Chinese)

[14] 李海林,王成,邓晓懿. 基于分量属性近邻传播的多元时间序列数据聚类方法[J]. 控制与决策,2018,33(4):649-656.
LI Hailin, WANG Cheng, DENG Xiaowei. Multivariate time series data clustering method based on component attribute neighbor propagation[J]. Control and Decision, 2018,33(4):649-656. (in Chinese)

[15] 徐明亮,王士同,杭文龙. 一种基于同类约束的半监督近邻反射传播聚类方法[J]. 自动化学报,2016,42(2):255-269.
XU Mingliang, WANG Shitong, HANG Wenlong. A semi-supervised nearest neighbor reflection propagation clustering method based on similar constraints[J]. Acta Automatica Sinica,2016,42(2):255-269. (in Chinese)

[16] 孙文舟,殷晓东,李树军. 基于熵权重的水下载体导航信息融合方法[J]. 武汉大学学报,2018,43(10):1465-1471.
SUN Wenzhou, YIN Xiaodong, LI Shujun. Underwater carrier navigation information fusion method based on entropy weight[J]. Journal of Wuhan University, 2018, 43(10):1465-1471. (in Chinese)

[17] 李文璟,李梦,邢宇哲,等. 基于熵权-灰色模型的电力数据网风险预测[J]. 北京邮电大学学报,2018,41(3):39-45.
LI Wenbiao, LI Meng, XING Ningzhe, et al. Risk prediction of power data network based on entropy weight-gray model[J]. Journal of Beijing University of Posts and Telecommunications,2018,41(3):39-45. (in Chinese)

[18] 陈秀明,刘业政. 基于熵权的多粒度犹豫模糊语言 VIKOR 群推荐方法[J]. 控制与决策,2018,33(1):111-118.
CHEN Xiuming, LIU Yezheng. Multi-granular hesitant fuzzy language VIKOR group recommendation method based on entropy weight[J]. Control and Decision, 2018, 33(1):111-118. (in Chinese)

[19] 焦李成,杜海峰,刘芳,等. 免疫优化计算、学习与识别[M]. 北京科学出版社,2006.

(编辑:王碧琚)