

# 基于膜聚类算法的风机振动故障诊断

邹武俊, 田涛, 蒲家蓉, 张宇森

(华北电力大学(北京)控制与计算机工程学院, 北京 102206)

**摘要:** 在工业环境下, 风机振动故障常常需要人工诊断, 诊断效率低, 不易完成实时计算和在线分析判断; 针对上述问题, 提出了一种膜聚类算法可用于风机振动故障的在线智能诊断; 该算法将膜计算的方法引入到聚类中, 并采用概率模型更新种群的方法实现最佳聚类中心的寻优; 算法首先在多个数据集上进行聚类实验, 实验结果显示该算法克服了常规聚类算法聚类结果不稳定, 聚类质量差的缺点; 然后将其应用于风机振动故障在线诊断系统中进行仿真测试, 结果显示所采用的方法能满足风机振动故障在线智能诊断要求, 也可应用于其他各类设备的振动故障在线智能诊断。

**关键词:** 膜计算; 聚类算法; 风机振动; 故障诊断

## Fault Diagnosis of Fan Vibration Based on Membrane Clustering Algorithm

Zou Wujun, Tian Tao, Pu Jiarong, Zhang Yusen

(North China Electric Power University, Beijing 102206, China)

**Abstract:** In industrial environment, fan vibration fault often needs manual diagnosis, which is inefficient and difficult to complete real-time calculation and online analysis and judgment. To solve the above problems, a membrane clustering algorithm is proposed in this paper, which can be used for on-line intelligent diagnosis of fan vibration faults. The algorithm introduces the membrane computing method into clustering, and uses the probability model to update the population method to optimize the best clustering center. The algorithm first carries out clustering experiments on multiple data sets, and the experimental results show that the algorithm overcomes the shortcomings of the unstable clustering results and poor quality of the clustering algorithm. Then it is applied to the on-line diagnosis system of fan vibration fault. The results show that the method can meet the on-line intelligent diagnosis requirement of fan vibration fault and can also be applied to the on-line intelligent diagnosis of vibration fault of other kinds of equipment.

**Keywords:** membrane calculation; clustering algorithm; fan vibration; fault diagnosis

## 0 引言

随着中国制造 2025 计划的提出, 制造行业迎来了新的机遇, 越来越多的机械设备应用于各个行业<sup>[1]</sup>。机械设备的故障会影响生产效率, 造成经济损失, 工业生产中常常采用人工诊断的方法来进行故障诊断, 这种方式具有诊断效率低, 专业性较强, 诊断不及时等缺点。如果能及时的发现故障, 并进行维护将会挽回大量的经济损失。

风机广泛应用于工厂、矿井、隧道、发电等场合的通风、引风、冷却等, 其工作性能的好坏直接或间接的影响生产安全和生产效率, 因此对风机的振动故障的实时在线智能诊断研究具有重要意义。文献 [2-3] 阐述了风机等机械设备常见故障的产生机理和各种故障与振动特征、敏感参数和故障原因之间的对应关系。目前应用于人工智能故障诊断较多的就是基于神经网络的故障诊断<sup>[4]</sup>。也有采用聚类算法进行故障诊断的文献, 如张孝远等<sup>[5]</sup>和李超顺等<sup>[6]</sup>分别利用最小最大核 K 均值聚类算法和引力搜索核聚类算法实现水电机组的振动故障分类, 但此方面的研究相对神

经网络偏少。

聚类是将物理或者抽象对象的集合通过一定的相似系数指标将对象分成多个具有不同相似系数的类。聚类主要应用于模式识别中的语音识别<sup>[7]</sup>、字符识别<sup>[8]</sup>、故障诊断<sup>[9]</sup>等, 机器学习中的聚类算法应用于图像分割<sup>[10]</sup>和机器视觉<sup>[11]</sup>, 图像处理中聚类用于数据压缩和信息检索, 另一个主要应用是数据挖掘<sup>[12]</sup>。在已有的聚类算法中, K-均值聚类算法(K-means)是最常用的一种聚类算法, 由于其聚类效率较高被广泛应用, 但聚类结果不稳定, 聚类质量较差, 很难直接将其运用于实际的工业生产中。为解决上述问题, 本文尝试在常规聚类算法的基础上对聚类算法的性能进行改进, 以满足工业生产的应用。

分布估计算法(Estimation of Distribution Algorithms, EDA)<sup>[13]</sup>是当前国际进化计算领域的一种新的进化研究热点。分布估计算法与遗传算法不同, 没有传统的交叉、变异等遗传操作, 它采用统计学习的方法建立解空间内个体分布的概率模型, 然后根据概率模型随机采样产生新的群体, 如此不断迭代, 完善概率模型, 实现种群的进化。膜计算在 1998 年第一次被罗马尼亚科学院和欧洲科学院院士提出, 它是从活细胞结构和功能中或从组织和器官等细胞群协作中抽象出来的计算模型, 是一种分布式、并行式的计算模型<sup>[14]</sup>。

收稿日期: 2018-08-01; 修回日期: 2018-08-25。

**作者简介:** 邹武俊(1995-), 男, 云南宣威人, 硕士研究生, 主要从事风力发电系统故障诊断技术方向的研究。

田涛(1968-), 男, 四川人, 博士, 副教授, 主要从事风力发电控制与仿真方向的研究。

为了解决常规聚类算法聚类结果不稳定，聚类质量差的缺点，本文将膜计算和 EDA 算法的思想引入到聚类算法中，将膜计算并行计算和 EDA 全局寻优的优势结合起来，提出了一种膜聚类算法，首先用标准数据集对算法进行聚类性能测试，然后将该算法应用于风机的振动故障诊断系统中，验证其在实际工程应用中的有效性。

## 1 传统聚类算法

### 1.1 数据聚类

数据聚类是通过某个评价指标，对一个数据集中的数据进行分组的过程。一个  $d$  维欧式空间的数据聚类可描述成这样的过程：将一个含  $n$  有个数据点的数据集根据某种相似评价指标划分到几个类别中，记  $X = \{X_1, X_2, \dots, X_n\}$  是一个由  $n$  个待分类的数据点组成的数据集，其中  $X_i = \{X_{i1}, X_{i2}, \dots, X_{id}\}$  是一个  $d$  维向量， $i = 1, 2, \dots, n$ ，且  $X_{ij}$  表示  $X_i$  的第  $j$  个特征数值。对于数据集  $X$ ，聚类算法通过找到  $K$  个簇的一个划分  $\{C_1, C_2, \dots, C_K\}$ ，使相同簇中的数据之间有最大的相似性，而不同簇之间的数据有最大的相异性。数据聚类分类后，各类别数据簇中数据应尽可能的紧凑，而不同类别的数据簇之间尽可能的分开，且满足以下性质：

- 1)  $C_i \neq \phi$ ，对任意  $i \in \{1, 2, \dots, K\}$ ；
- 2)  $C_i \cap C_j = \phi$ ，对任意  $i \neq j$  且  $i, j \in \{1, 2, \dots, K\}$ ；
- 3)  $\bigcup_{i=1}^K C_i = S$ 。

在评价两个类别的相似性时，常采用欧几里得距离作为相似度评价指标，任意两个  $d$  维的数据点  $X_i$  和  $X_j$  的欧几里得距离计算公式如下：

$$d(X_i, X_j) = \sqrt{\sum_{p=1}^d (x_{ip} - x_{jp})^2} = \|X_i - X_j\| \quad (1)$$

数据聚类过程中为了发现  $K$  个数据簇的最优类别划分，聚类问题常常等价简化为一个性能指标评价函数的最优化问题，对于  $K$  个数据簇的聚类质量度量，通常采用整体均方差 (MSE) 作为聚类性能指标评价函数，即：

$$J_m(C_1, C_2, \dots, C_K) = \sum_{i=1}^K \sum_{X_j \in C_i} \|X_j - Z_i\| \quad (2)$$

其中， $Z_1, Z_2, \dots, Z_K$  分别是  $K$  个数据簇  $C_1, C_2, \dots, C_K$  的聚类中心，且  $Z_i \in R^d, i = 1, 2, \dots, K$ 。

### 1.2 K-means 算法

K-means 算法是聚类问题中常见的一种非监督学习方法。算法首先从  $n$  个待分类数据中任意选择  $k$  个对象作为初始的聚类中心，然后对其余的数据对象，计算它们与这些聚类中心的相似度，分别将它们分配到与其最相似的数据类别中，然后再计算每个数据类别的新的聚类中心；不断重复以上过程直到聚类性能指标评价函数收敛为止。一般采用均方差作为聚类性能指标评价函数，采用欧几里得距离作为相似度评价指标。算法步骤如下：

1) 确定数据的分类数，从样本数据中任选  $K$  个对象作为初始聚类中心  $c_1, c_2, \dots, c_K$ ；

2) 将数据集  $X$  中的每个样本数据按照相似度评价指标分配到某个聚类中心所代表的数据类别中；

3) 按照以下公式计算每个数据类别新的聚类中心  $c'_i (i = 1, 2, \dots, K)$ ：

$$c'_i = \frac{1}{N_i} \sum_{X \in S_i} X \quad (3)$$

其中： $N_i$  为第  $i$  个数据类别代表的聚类区域  $S_i$  所包含的数据样本个数；

4) 若  $c'_i \neq c_i (i = 1, 2, \dots, K)$ ，则返回步骤 2)；否则算法收敛，计算结束。

K-means 算法为了防止迭代过程中步骤 4) 不满足结束条件而出现死循环的情况，可以在算法初始化时，设置最大迭代步数来避免这种情况发生<sup>[15]</sup>。

## 2 本文膜聚类算法

本文提出了一种在膜计算框架下融入分布估计算法的膜聚类算法，称之为 EDAPS 算法。它将分布估计算法利用概率模型更新种群的方法引入到膜进化机制中，对同一膜中的个体采用概率模型产生子代更新种群，同时结合膜计算中的转运机制，促进不同膜间优秀个体的交流，促进了膜中的种群进化，实现解空间最佳聚类中心的寻优。

### 2.1 膜结构

算法中所采用的膜系统是一种组织型膜系统 (Tissue-like P Systems)。此膜系统以环境作为输出，初始时空，在计算过程中，最优对象输出存储到环境中，并将最优个体传递到每一个基本膜中，促进每个膜中的种群进化，同时保证算法的收敛性，膜结构如图 1 所示。其中  $M_1, M_2, \dots, M_q$  为  $q$  个基本膜，开始时膜  $M_1, \dots, M_{q-1}$  初始化生成一组初始种群；膜  $M_q$  用于接收膜  $M_1, \dots, M_{q-1}$  中的较优个体，同时将膜中较优的个体传递到基本膜中； $W$  为环境，初始时空，用于存储当前出现过的最优个体。

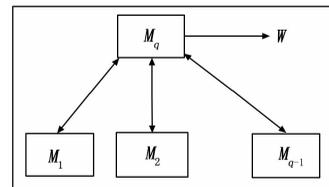


图 1 本文膜聚类算法的膜结构

初始时，系统将为  $q-1$  个膜生成多个初始对象，每个初始对象代表一组假定的最优簇中心。假定数据集  $X$  有  $K$  个簇  $C_1, C_2, \dots, C_K$ ，对应的簇中心分别为  $z_1, z_2, \dots, z_K$ ，且簇中心和数据点都是  $d$  维的向量，则每个基本膜中的对象可设计为如下的  $k \times d$  维向量：

$$Z = (z_{11}, z_{12}, \dots, z_{1d}, \dots, z_{k1}, z_{k2}, \dots, z_{kd}) \quad (4)$$

其中： $z_{i1}, z_{i2}, \dots, z_{id}$  对应第  $i$  个簇中心的  $d$  个分量， $i = 1, 2, \dots, K$ 。

### 2.2 膜进化规则

与常规的进化机制不同，本文算法采用概率模型采样

产生子代更新每个细胞膜中的个体 (对象)。

选择规则: 每个膜中的所有个体可视为一个相对独立的种群, 首先对每个膜中的所有个体按照公式 (2) 计算其性能函数值, 并进行排序, 选择其中较好的个体保留。

概率模型: 考虑到数据模型一般都可以用正态分布来表示, 本文就采用正态分布的概率模型。设连续型随机变量  $X$  服从正态分布, 其概率密度为:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x-\mu}{\sigma}} \quad (-\infty < x < \infty) \quad (5)$$

正态分布随机变量  $x$  的均值  $E(x) = \mu$  和方差  $D(x) = \sigma^2$ 。由概率论可知,  $\mu, \sigma$  的估计值分别为:

$$\tilde{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (6)$$

$$\tilde{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^2} \quad (7)$$

由计算机仿真理论可知, 设  $\mu_1$  和  $\mu_2$  是两个独立的 (0,1) 均匀分布随机数, 则正态分布  $X \sim N(\mu, \sigma^2)$  的随机数为<sup>[16]</sup>:

$$x_1 = \mu + \sigma(-2\ln\mu_1)^{\frac{1}{2}} \cos 2\pi\mu_2 \quad (8)$$

$$x_2 = \mu + \sigma(-2\ln\mu_1)^{\frac{1}{2}} \sin 2\pi\mu_2 \quad (9)$$

子代产生: 通过选择规则选择膜中的一部分个体计算其均值和方差得到数据的概率模型 (5), 按照公式 (8) 或公式 (9) 产生子代个体。

转运规则: 细胞膜之间的转运通道用于它们之间对象的交换与共享, 如图 1 所示, 膜  $M_1, \dots$ , 膜  $M_{q-1}$  与膜  $M_q$  之间都为双向通道, 用于对象交换, 即  $q-1$  个膜中的较优对象都发送到膜  $M_q$  中, 且膜  $M_q$  中的最优个体也共享到其他各个膜中, 膜  $M_q$  中对比产生的最优对象作为最优解输出到环境中  $W$  保存。经过一定的进化步数后, 环境中所保存的最有对象  $W$  即为最终输出结果。

### 2.3 算法步骤

本文提出的膜聚类算法是一个度为 5, 采用概率模型进化机制的组织型 P 系统, 主要通过概率模型不断更新 P 系统中各子系统的群体, 为待聚类的数据集搜索其最优的聚类中心, 算法描述如下:

- 1) 随机初始化产生每个膜中的个体;
- 2) 分别计算膜  $M_1, \dots$ , 膜  $M_{q-1}$  中的每个个体的性能函数值并进行排序;
- 3) 将膜  $M_1, \dots$ , 膜  $M_{q-1}$  中的较优对象传递到膜  $M_q$  中, 膜  $M_q$  中对比选择膜中最优的部分对象传递到膜  $M_1, \dots$ , 膜  $M_{q-1}$  中替换原来较差的对象, 最优对象输出到环境  $W$  中;
- 4) 选择膜  $M_1, \dots$ , 膜  $M_{q-1}$  中较优的部分个体计算其均值方差, 建立其对应的概率模型, 并按照公式 (8) 或公式 (9) 更新膜中的个体;
- 5) 重复步骤 2) ~4), 直到满足结束条件;
- 6) 输出环境中的对象  $W$ , 即为最终聚类中心。

## 3 仿真实验与实验结果分析

### 3.1 实验条件与数据集

实验中采用的计算机使用 Intel (R) i5-5200U 2.20

GHz CPU, 4 GB DDR3 RAM; 算法编辑软件为 Matlab R2014a。为了评估本文中的膜聚类算法, 实验中使用了 4 个数据集, 其中 1 个是 4 维的人工数据集 (Man), 另外 3 个是来自于 UCI 的真实数据集 (Iris, Seeds, Glass), 表 1 给出了 4 个数据集的简单描述。

表 1 实验采用的数据集

| 数据集   | 样本数 | 维数 | 分类数 |
|-------|-----|----|-----|
| Man   | 150 | 4  | 4   |
| Iris  | 150 | 4  | 3   |
| Glass | 214 | 9  | 6   |
| Seeds | 210 | 7  | 3   |

### 3.2 输入参数

在实验中, 所提出的膜聚类算法 EDAPS 将与经典的 K-means 聚类算法和 2 种典型的进化算法在上文中的 4 个数据集上进行测试验证, 包括基于遗传算法的聚类算法 (GA-based) 和基于粒子群算法的聚类算法 (PSO-based)。

本文提出的膜聚类算法参数设置如下: 每个细胞中对象个数  $N$  为 100, 最大执行步数 200, 细胞 1-细胞 4 中对象选择比例为 30%, 细胞 5 中的精英保留比例为 10%。

在实现其他两种进化算法 (GA-based 和 PSO-based) 时, 通过几组典型参数的测试结果确定其最优参数设置如下: (1) GA-based 的交叉概率  $P_c = 0.6$ , 变异概率  $P_m = 0.01$ ; (2) PSO-based 参数设置为  $c_1 = c_2 = 2.0$ ,  $\omega = 0.7298$ 。另外, 这些算法和 K-means 算法的进化代数均设置为 200。

### 3.3 实验结果

本文采用典型的整体均方差 (公式 2) 来评价聚类算法的质量, 一般地, 其值越小代表聚类质量越好。考虑到这些算法包含的随机因素, 算法每次运行的结果可能不一样, 所以每个数据集上的每种算法都独立运行 50 次, 统计得到每种算法 50 次运行的平均值、标准差、最小值和最大值, 平均值反应算法的平均性能, 标准差反应算法的鲁棒性, 最小值表示算法运行最优值, 最大值表示算法运行最差值, 实验结果如表 2~表 5 所示。

为了更直观的比较各种算法的中间收敛过程, 实验记录了各种算法在每种数据集上的 50 次独立运行的过程收敛值, 计算得到其平均收敛曲线, 反映各种算法的中间收敛特性, 实验结果如图 1~4 所示。

表 2 人工数据集实验结果

| 算法        | 平均值     | 标准差     | 最小值     | 最大值      |
|-----------|---------|---------|---------|----------|
| K-means   | 88.3516 | 18.7037 | 72.9966 | 125.9297 |
| PSO-based | 82.6126 | 20.3896 | 71.8017 | 121.9654 |
| GA-based  | 72.6985 | 0.1259  | 71.8771 | 72.4871  |
| EDAPS     | 71.8005 | 0.0000  | 71.8005 | 71.8005  |

表 3 Iris 数据集实验结果

| 算法        | 平均值      | 标准差     | 最小值     | 最大值      |
|-----------|----------|---------|---------|----------|
| K-means   | 103.9599 | 11.2825 | 97.3259 | 123.8497 |
| PSO-based | 99.7936  | 9.3860  | 96.6555 | 127.6677 |
| GA-based  | 96.8953  | 0.3181  | 96.6596 | 97.5891  |
| EDAPS     | 96.7350  | 0.1866  | 96.6555 | 97.2414  |

表 4 Seeds 数据集实验结果

| 算法        | 平均值      | 标准差     | 最小值      | 最大值      |
|-----------|----------|---------|----------|----------|
| K-means   | 313.4134 | 0.2537  | 313.2168 | 313.7343 |
| PSO-based | 335.4651 | 11.8804 | 315.5482 | 366.1951 |
| GA-based  | 405.6953 | 0.2439  | 405.2852 | 406.3974 |
| EDAPS     | 311.8283 | 0.1079  | 311.7978 | 312.3060 |

表 5 Glass 数据集实验结果

| 算法        | 平均值      | 标准差     | 最小值      | 最大值      |
|-----------|----------|---------|----------|----------|
| K-means   | 221.9906 | 9.3005  | 215.6775 | 252.5734 |
| PSO-based | 300.0225 | 11.2540 | 272.2721 | 327.9715 |
| GA-based  | 269.4650 | 3.4220  | 261.6796 | 282.2001 |
| EDAPS     | 218.5298 | 3.8702  | 213.9652 | 224.9605 |

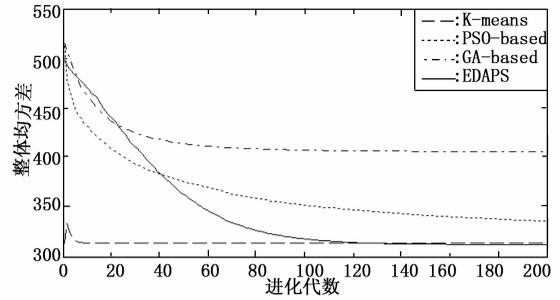


图 4 Seeds 数据集 50 次聚类中心 MSE 值平均变化

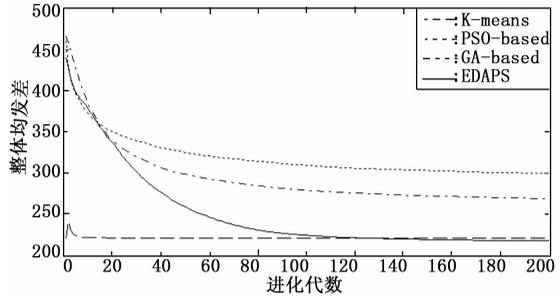


图 5 Glass 数据集 50 次聚类中心 MSE 值平均变化

2) 本文膜聚类算法 EDAPS 在低维和高维聚类问题中都比 K-means 算法、PSO-based 和 GA-based 算法有更好的聚类质量, 且基于该算法的聚类结果标准差相比其他几种算法小很多。实验表明此算法聚类质量好, 聚类结果稳定可靠。

## 4 风机振动故障诊断

### 4.1 风机振动故障诊断系统

风机是工业生产中常用的机械设备, 其故障主要由振动产生, 其故障主要包括不平衡故障。轴承故障、不对中故障以及混合故障。故障诊断时先用本文膜聚类算法对已知的故障数据进行聚类分析, 获得  $M$  个聚类中心  $\{c_1, c_2, \dots, c_M\}$ 。假设  $\{X_1, X_2, \dots, X_n\}$  为待诊断的样本数据, 则样本  $X_n$  的故障类型由公式 (10) 求得。即将待诊断样本数据分配到与各已知聚类中心的欧几里得距离最小的类别中。

$$M^{(i)} = \operatorname{argmin} \|x^{(i)} - c_j\|^2 \quad (10)$$

$$c_j = \frac{\sum_{i=1}^m \{M^{(i)} = j\} x^{(i)}}{\sum_{i=1}^m \{M^{(i)} = j\}} \quad (11)$$

待诊断样本数据分类后加入已知的样本数据集中, 此时样本数据集发生变化, 样本数据集的聚类中心也会发生变化, 可通过公式 (11) 计算其新的聚类中心, 然后进行下一个样本数据的诊断分类。在诊断过程中通过公式 (10) 和公式 (11) 不断迭代, 使诊断系统不断“学习”, 从而不断调整聚类中心完善诊断系统。风机振动故障诊断系统的具体实现步骤如下:

1) 采用 2.3 节算法对已知的故障数据进行分类, 计算得到每一类数据的聚类中心;

2) 选择一个待诊断的样本数据, 通过公式 (10) 将其分配到与其中聚类中心距离最小的类别中;

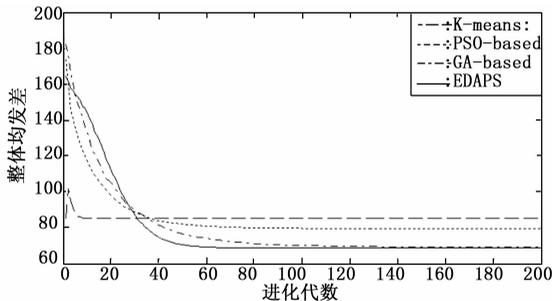


图 2 Main 数据集 50 次聚类中心 MSE 值平均变化

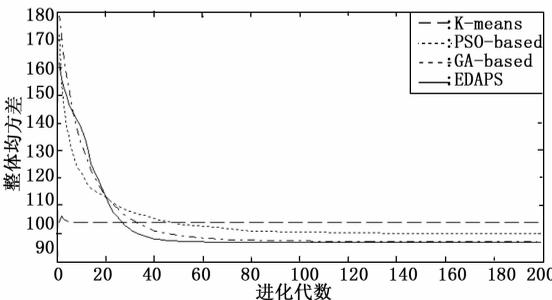


图 3 Iris 数据集 50 次聚类中心 MSE 值平均变化

不同算法在 4 个数据集上测试的实验结果显示:

1) K-means 算法在 4 个数据集上平均迭代次数 7 次左右开始稳定, 聚类效果在 Man 数据集、Iris 数据集、Glass 数据集上都比较差, 聚类结果不稳定; 但在 Seeds 数据集上聚类效果比 PSO-based 和 GA-based 好, 比本文 EDAPS 算法聚类质量差, 这是由于 K-means 算法本身的特性导致算法的性能受数据集内在关系的影响较大;

3) 将被诊断的样本数据加入到已经聚类后的数据集中, 通过公式 (11) 计算其新的聚类中心;

4) 选择下一个待诊断样本数据重复步骤 2) 和步骤 3) 直至所有待诊断样本数据诊断结束。

#### 4.2 故障诊断系统的应用验证

将 4.1 节中设计的风机振动故障诊断系统应用于实际工程中的风机振动故障数据, 进行故障诊断分析。实验所采用的数据为文献 17 中提供的风机振动数据, 数据分为正常状态、不平衡故障状态、轴承故障状态、不对中故障状态和轴承+不平衡故障状态共 5 种数据, 由于不对中故障实验对整体试验造成的破坏较大, 文中仅提供了 1 组实验数据, 故实验以其余 4 种状态下的共 240 组数据作为测试数据进行风机振动故障诊断系统模拟实验。

实验中以每种状态的前 40 组数据作为已知状态进行诊断系统的学习, 用每种状态的后 20 组数据作为待分类测试数据对诊断系统的有效性进行评估。

为了验证本文算法在风机振动故障诊断系统中的可行性, 实验还分别采用了上文中 K-means 算法、PSO-based、GA-based 和本文提出的膜聚类算法 EDAPS 进行故障诊断结果对比, 输入参数与 3.2 节相同。考虑到算法本身具有随机性的特点; 每种算法独立重复诊断 50 次, 其诊断结果正确率的平均值、标准差、最小值和最大值记录于表 6 中。

表 6 风机振动故障诊断结果

| 算法        | 平均值/% | 标准差    | 最小值/% | 最大值/% |
|-----------|-------|--------|-------|-------|
| K-means   | 76.43 | 0.0713 | 66.67 | 86.67 |
| PSO-based | 88.73 | 0.0808 | 75.00 | 96.67 |
| GA-based  | 90.07 | 0.0513 | 86.67 | 95.00 |
| EDAPS     | 94.30 | 0.0431 | 90.00 | 98.33 |

从表 6 中风机振动故障诊断结果可以看出 K-means 算法和 PSO-based 算法诊断结果平均准确率都较低, 且标准差较大, 诊断结果不稳定。GA-based 算法平均诊断准确率达到 90.07%, 而基于本文提出的膜聚类算法的平均故障诊断准确率达到 94.30%, 比其他三种聚类算法都更准确; 此外, 从诊断结果标准差可以看出基于本文 EDAPS 算法的故障诊断系统诊断结果更加稳定可靠。

## 5 结束语

本文提出了一种膜聚类算法, 结合了膜计算和分布估计算法的优势, 首先在 4 个数据集上进行聚类测试, 验证了算法的可行性; 然后将其应用于风机振动故障诊断系统中进行故障诊断测试, 实验结果显示基于该算法的风机振动故障诊断系统能够较准确的进行故障诊断, 诊断结果稳定可靠。可应用于工业环境中各类设备的振动故障在线智能诊断系统中。

#### 参考文献:

[1] 姚立国, 黄海松. 改进 K 均值模拟退火聚类算法的滚动轴承

故障诊断 [J]. 组合机床与自动化加工技术, 2017 (4): 114-117.

[2] Thomson W T. Theory of Vibration With Application [M]. Prentice, 1972.

[3] Walker K, Shirkhodaie A. Development of a virtual environment for fault diagnosis in rotary machinery [A] Proceedings of the 33rd Southeastern Symposium on System Theory (Cat. No. 01EX460) [C]. IEEE, Piscataway, NJ, USA, 2001, 99-103.

[4] 汪光阳, 周义莲. 风机振动故障诊断综述 [J]. 安徽工业大学学报 (自然科学版), 2006 (1): 64-68.

[5] 张孝远, 张新萍, 苏保平. 基于最小最大核 K 均值聚类算法的水电机组振动故障诊断 [J]. 电力系统保护与控制, 2015, 43 (5): 27-34.

[6] 李超顺, 周建中, 肖 剑. 基于引力搜索核聚类算法的水电机组振动故障诊断 [J]. 中国电机工程学报, 2013, 33 (2): 98-104, 18.

[7] 姚敏锋, 李心广, 杨佳能. 基于语音特征聚类的 HMM 语音识别系统研究 [J]. 微计算机信息, 2012, 28 (10): 458-460.

[8] 王海涛, 黄文杰, 朱永凯, 等. 基于聚类分析与神经网络的车牌字符识别 [J]. 数据采集与处理, 2008 (2): 238-242.

[9] 曹苏群, 侯志伟, 刘 磊. 一种基于聚类的滚动轴承故障诊断方法 [J]. 机械设计与制造, 2012 (5): 239-241.

[10] 张向荣, 蹇晓雪, 焦李成. 基于免疫谱聚类的图像分割 [J]. 软件学报, 2010, 21 (9): 2196-2205.

[11] 杨立志, 周方颖. 基于机器视觉的毛纱管聚类方法 [J]. 毛纺科技, 2017, 45 (8): 85-88.

[12] 周 涛, 陆惠玲. 数据挖掘中聚类算法研究进展 [J]. 计算机工程与应用, 2012, 48 (12): 100-111.

[13] Zhang Q, Mühlhain H. On the convergence of a class of estimation of distribution algorithms [J]. IEEE Transactions on Evolutionary Computation, 2004, 8 (2): 127-136.

[14] 张葛祥, 潘林强. 自然计算的新分支——膜计算 [J]. 计算机学报, 2010, 33 (2): 208-214.

[15] 温凤文, 王洪春. 基于邻域影响的改进粒子群算法的聚类算法 [J]. 重庆师范大学学报 (自然科学版), 2014, 31 (2): 59-62.

[16] 邱 玲, 高 尚, 曹存根. 改进的正态分布的分布估计算法 [J]. 计算机科学, 2015, 42 (8): 32-35.

[17] 范佳卿. 基于实时特征值的风机振动状态监测与数据挖掘的故障诊断研究 [D]. 上海: 上海发电设备成套设计研究院, 2017.

[18] 拓守恒, 邓方安, 周 涛. 一种利用膜计算求解高维函数的全局优化算法 [J]. 计算机工程与应用, 2011, 47 (19): 27-30.

[19] 易 辉, 梅 磊, 李丽娟, 等. 基于多分类相关向量机的水电机组振动故障诊断 [J]. 中国电机工程学报, 2014, 34 (17): 2843-2850.

[20] 蒋恩超, 傅 攀, 张思聪. 小波包和 GA-SVM 在轴承故障诊断中的应用 [J]. 计算机测量与控制, 2017, 25 (10): 7-10.