

免疫粒子群算法及其在矿井提升机故障诊断中的应用

汪楚娇, 夏士雄, 牛 强

(中国矿业大学计算机科学与技术学院, 江苏徐州 221116)

摘 要: 基于人工免疫系统的故障诊断方法是人工智能领域发展起来的一个十分活跃的分支. 为了提高免疫算法在矿井提升机故障诊断系统中的执行效率, 通过对诊断问题进行更精确的建模和分析, 提出了将免疫模型和离散粒子群进化算法相结合的提升机系统的故障诊断方法. 该方法在免疫形态空间中采用核主元形式的相似性度量, 解决了传统距离判别函数法在故障诊断中存在误差较大等问题. 仿真结果表明, 该方法能够适应诊断过程中出现的不确定性, 并实现多故障诊断.

关键词: 提升机; 粒子群算法; 故障诊断; 免疫模型

中图分类号: TP301 **文献标识码:** A **文章编号:** 0372-2112 (2010) 2A-094-05

Artificial Immune Particle Swarm Optimization for Fault Diagnosis of Mine hoist

WANG Chu-Jiao, XIA Shi-Xiong, NIU Qiang

(School of Computer Science and Technology, China University of Mining and Technology, Jiangsu 221116, China)

Abstract: This paper presents an intelligent methodology for diagnosing incipient faults in mine hoist. In this fault diagnosis system, in order to enhance the immune algorithms performance, we propose the improved immune-based symbiotic a new evolutionary learning algorithm. This new evolutionary learning algorithm is based on Discrete Particle Swarm Optimization (DPSO) technique to improve the mutation mechanism. Also to solve the problem that exists in fault diagnosis based on the traditional method using distance discriminant function, an improved method based on immunity strategy with similarity measurement of principle component kernel is presented. The effectiveness of the DPSO based immune algorithms is demonstrated through the classification of the fault signals in mine hoist. Simulation results show that the new scheduling algorithm can deal with the uncertainty situation and be suitable for multi-faults diagnosis, compared to the traditional scheduling algorithms.

Key words: particle swarm optimization; immunity; self; nonself; fault diagnosis

1 引言

受到生物免疫系统的启发, 近年来学术界引发了人工免疫系统(Artificial Immune System, 简称 AIS)的研究热潮, 并逐渐用其解决工程实际问题^[1]. 在故障诊断领域, Dasgupta 利用免疫系统阴性选择机理提出一个高效检测算法用于铣削刀具断裂的检测, 试验结果表明它可有效地用于自动安全临界操作监测; Ishiguro 将免疫网络模型用于在线设备故障诊断; Taniguchi 和 Dote 将人工免疫网络模型用于交流驱动和 UPS 的控制和诊断. 总之, 从文献检索来看, 目前, 国内外专家学者在设备故障诊断、计算机病毒预防等方面做了一些初步的工作^[2]. 但总体来说, 国内外对人工免疫系统在故障诊断领域中的应用研究刚刚起步, 实际应用更少, 而故障诊断是继信息安全之后另一个从免疫系统直接映射而来的人工免

疫系统应用领域, 深入挖掘人工免疫系统仿生机理有助于进一步开展基于人工免疫系统模型的故障诊断研究.

本文针对提升机故障诊断的特点, 引入节约覆盖集理论^[3], 对故障诊断模型问题重新给出了细致详尽的描述. 提出了一种新的二进制量子粒子群优化算法, 并与基于核主元的人工免疫概念相结合实现对矿井提升机故障的检测, 建立了免疫粒子群算法求解故障诊断问题的网络模型. 将这两种方法结合可以加快极小诊断的求解过程. 多个算例仿真表明该方法是有效可行. 较好地解决了提升机故障诊断中遇到的推理速度慢和先验知识获取困难等问题.

2 故障诊断问题描述以及数学模型

下面给出基于概率因果和免疫网络模型的故障诊断技术的基本概念和思想.

定义 1 设备或系统的状态可以使用一定的数据模式(如特征向量)来描述.从监控对象运行数据中提取描述其工作状态的 m 维归一化状态空间向量称为系统状态特征向量 E .故障症状集 $E = \{e_1, e_2, \dots, e_m\}$ 表示网络潜在的所有故障症状, $E_l \subseteq E$ 是当前观测到的故障症状.

定义 2 候选故障集 $F = \{f_1, f_2, \dots, f_n\}$ 表示所有可能发生的故障, $F_l \subseteq F$ 代表当前发生的故障.

在已知故障症状 E_l 的情况下求解最小诊断集过程就是搜索 F_l 并使 F_l 最小的过程.这个组合优化问题是一个 NP 完全问题.

定义 3 监控对象正常工作状态下,所获得的系统状态特征向量称为自己模式集合,简称为自己集,记作 $V_s = (V_{s_1}, V_{s_2}, \dots, V_{s_n}), 0 \leq V_{s_i} \leq 1, i = 1, 2, \dots, n$ 监控对象异常工作状态下,所获得的系统状态特征向量称为非己模式集合,简称为非己集 $V_N = (V_{N_1}, V_{N_2}, \dots, V_{N_n}), 0 \leq V_{s_i} \leq 1, i = 1, 2, \dots, n$.

定义 4 在 M 维形态空间中,由 M 个 N 维状态检测归一化向量(抗体):

$B = (b_1, b_2, \dots, b_n), 0 \leq b_i \leq n, i = 1, 2, \dots, n$ 所组成的抗体集合记为:

$D = \{B_1, B_2, \dots, B_m\}, 0 \leq B_j \leq m, j = 1, 2, \dots, m$ 称为状态检测器.

系统对外部抗原的免疫应答过程,即是系统中的抗体集合动态进化学学习过程,其中抗原与抗体或抗体与抗体之间的匹配^[4],依据它们之间的亲和力大小进行.

定义 5 设 $X = (x_1, x_2, \dots, x_n)$ 为状态空间上的 N 维归一化向量, x_i, x_j 之间的亲和力记作:

$$A(x_i, x_j) = \frac{1}{1 + d(x_i, x_j)} \quad (1)$$

$d(x_i, x_j)$ 为 x_i, x_j 之间的距离.免疫分子之间的亲和力(Affinity)为免疫分子之间的相似程度,距离越小,其亲和力值越大,反之亦然.

3 粒子群算法简介

粒子群(Particle Swarm Optimization, PSO)算法最初由 Kennedy 提出^[5],其更新公式如下:

$$V_{di} = w \times V_{di} + c_1 \times r_1 \times (p_{di} - x_{di}) + c_2 \times r_2 \times (p_i - x_{di}) \quad (2)$$

$$x_{di} = x_{di} + V_{di} \quad (3)$$

其中: V 代表粒子的速度,每一维粒子的速度有一最大速度 V_{\max} ,若 $V \geq V_{\max}$ 则 $V = V_{\max}$, x 表示粒子的位置,下标 d 表示第 d 个粒子, i 表示当前粒子的第 i 位元素; P_{di} 表示粒子自身的历史最优解; P_i 表示种群全局

最优解, r_1, r_2 为在 $[0, 1]$ 间均匀分布的随机数, c_1, c_2 是学习因子,通常设 $c_1 = c_2 = 2$.其后 Kennedy 于 1997 年提出一种离散二进制 PSO(Discrete PSO)算法^[6],用于求解离散进制优化问题. DPSO 算法保留了原始 PSO 的速度公式,粒子定义为一组由 0, 1 组成的二进制向量,每个粒子元素速度通过预先设定的 S 形限幅转换函数 $S(V_{di})$ 转换为粒子该位元素取 1 的概率.

$$S(V_{di}) = 1 / (1 + e^{-v_{di}}) \quad (4)$$

同时为提高优化效率,文中利用一种改进的离散粒子群算法^[7]生成最小完备测试集,根据故障测试集优化问题的具体特点,构造粒子的表达方式和编码规则,建立粒子群的速度-位置模型.

4 算法的编码和相应的算子

基于免疫和二进制粒子群算法的故障诊断算法为了去除不必要的变量以减小数据维数,提高故障诊断的实时性能和正确率,引入了免疫系统中的基于核主元的亲和力作为适应度函数(式 7)指导粒子更新,以剔除与故障不相关的变量.首先给出算法的编码和相应的算子.

4.1 基于核主元相似度的亲和力定义

将各种故障模式的聚类中心作为参考模式,把待检样本与参考模式的距离作为判别函数,待检样本应隶属于与其距离最短的故障模式类.这就是常用的基于距离函数的故障诊断方法^[8].

免疫形态空间 S 是一个 m 维的实数空间.免疫算法的核心是采用一种合适的免疫形态空间中免疫分子亲和力的定义方法.原始算法直接采用欧氏距离的倒数作为亲和力,这是很局限的.如果各分量方差不等而呈其它形态,上述算法并不能得到好的分类效果^[9].这里提出采用一种称为主元核相似度的亲和力定义方法.

通常类核可由单位正交矢量 $S_1, S_2, \dots, S_r (0 \leq r \leq n)$ 所张成的空间表示^[9],即 $V_i = \{v, S_1, S_2, \dots, S_r\} (0 \leq r \leq n)$.式中 v 是特征空间中的一点.每一类故障就用 $V_{ij} = \{v, S_{i1}, S_{i2}, \dots, S_{ir}\} (0 \leq r \leq n)$ 代表,一个待分类样本 x 与故障类的距离采用 x 到 V_{ij} 的正交距离,定义一个核 $K_j = P_j^T x_j$ 来表示一个类 V_{ij} ,称为主元核.其中 $P_j = (p_1, p_2, \dots, p_k)$ 是和 X 矩阵的协方差矩阵的 k 个最大特征值相对应的规格化特征向量系统.

因此可以规定一个样本 X_j 与主元核 K_j 之间的正交距离来表示样本与核主元之间的相似程度^[9],即

$$d(X_i, K_j) = \left[\|X_i\|^2 - \sum_{j=1}^k (X_i^T P_j)^2 \right]^{\frac{1}{2}} \quad (5)$$

定义 6 主元子空间上的样本之间的相似性程度

可以用主元子空间上的欧氏距离来度量,即

$$d(\mathbf{X}_i, \mathbf{X}_j) = [(\mathbf{P}_i^T \mathbf{X}_i - \mathbf{P}_j^T \mathbf{X}_j)^T - (\mathbf{P}_i^T \mathbf{X}_i - \mathbf{P}_j^T \mathbf{X}_j)]^{\frac{1}{2}} \quad (6)$$

这样形态空间中的免疫分子之间的亲和力为:

$$A_{ij} = 1/1 + d(\mathbf{X}_i, \mathbf{X}_j) \quad (7)$$

4.2 编码方式

定义 7 $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ 表示一个 $n \times m$ 的与系统关联的 0/1 矩阵.

通过这个定义, A 的每一行对应一个个体, 每一列对应一个故障. 用二值向量代表集合, 则诊断集的个体编码为 (b_1, b_2, \dots, b_m) 的形式, 其中 b_i 取值为 0 或 1.

假设选取典型故障 M 类, 将每一个个体位串的长度定义为 M 位, 随机产生 N 个个体组成一个群体, 个体位串中值为 1 的所对应的故障即构成了诊断假设 F_i , 表示诊断问题的一个可能解.

如上所述, 每个抗体均由一个 $(1 \times S)$ 矩阵表示, 其中 $S = M + L + 1$, 其结构如表 1 所示. 整个群体的数据结构便是一个 $N \times (M + L + 1)$ 的矩阵.

表 1 抗体的数据结构

输入变量二进制编码	输入变量	适应度
M	L	1

4.3 个体选择算子

基于期望值的选择方法可以用比例保留法, 被淘汰的抗体通过粒子群算法生成.

定义 8 望值函数定义:

$$E(\mathbf{X}_i) = \frac{A(\mathbf{X}_i, \mathbf{X}_j)}{S(\mathbf{X}_i)} \quad (8)$$

其中, $A(\mathbf{X}_i, \mathbf{X}_j)$ 为抗体与抗体之间的亲和度, $S(\mathbf{X}_i)$ 为抗体在抗体(解)空间的浓度值.

定义 9 浓度函数定义:

$$S(\mathbf{X}_i) = \frac{1}{N} \sum_{i=1}^N A(\mathbf{X}_i, \mathbf{X}_j) \quad (9)$$

定义 10 所有抗体与抗原的亲和力的平均值为平均亲和力, 可表示为

$$f = \sum_{i,j=1}^N A(\mathbf{X}_i, \mathbf{X}_j) / N \quad (10)$$

式中 N 为抗体的数量.

定义 11 数 q 和平均亲和力的乘积为亲和力阈值, 可表示为:

$$T = q \times f \quad (11)$$

式中 q 的选取可根据具体情况确定.

5 基于免疫原理的故障诊断

故障诊断实质是建立故障征兆空间到故障空间的映射, 实现对故障的识别和诊断. 我们把外来侵犯的抗

原和免疫系统产生的抗体分别与实际求解问题的目的函数及问题的解相对应, 利用人体免疫功能的特点改进和提高粒子群算法执行能力.

5.1 抗原抗体和故障诊断系统的关系

被监测机组设备的状态征兆对应着抗原, 用 E_i 表示, E_i 分别为状态征兆的 P 个规一化的分量. 每种故障模式下的故障检测集合 F_i 对应抗体. 通过训练正常状态下的抗原集合进化得到自我集 S , 通过训练故障模式下的抗原集合进化得到相应的故障检测集合, 即记忆抗体集 NS . 记忆抗体集两两互不相交, 所有的记忆抗体集组成记忆抗体库 Lib , 即:

$N_{s_i} \cap N_{s_j} = \phi$, $Lib = N_{s_1} \cup N_{s_2} \dots N_{s_n}$ 其中 i, j 分别表示不同的故障模式. 当故障诊断系统对应于整个免疫系统时, 若 IM 为全集, 则有:

$$S \cup Lib \subset IM, S \cap Lib = \emptyset$$

5.2 自我集与正常/异常状态识别

采集正常无故障状态下的状态信号(如振动、摆度、压力脉动等), 分析并提取规一化的特征参数, 即初始正常抗原数据. 初始正常抗原集合要进行自我净化. 即两两之间计算亲和力. 之后删除亲和力大于净化阈值的抗原, 剩下的抗原构成自我集. 此过程实质上就是利用反向选择算法确定设备工作是否正常.

5.3 记忆抗体集与故障分类

本文利用动态规模免疫算法确定故障类型. 反向选择算法是受到反向选择机理的启发, 其主要特点在于当利用该算法检测异常时, 并不需要先验知识, 但是该算法也有一些局限性. 使用反向选择算法进行故障诊断时只能发现受检测体是否发生故障, 却无法确定故障类型. 而通过动态规模免疫算法产生每一种故障模式下的记忆抗体集, 即可确定故障类型. 每一种故障模式下的记忆抗体集的形成都要经历一个自学习、自适应的过程. 具体算法步骤如下:

1. 初始化. 将已知故障模式中的样本看作 B 细胞, 对每种故障模式产生训练抗体集合 A . 对于要建立起来的免疫系统, 生成 $N = Nr + M$ 个抗体. 其中 Nr 由 B 细胞分泌而成, M 属于记忆细胞. 设每个抗体的 p_{best} 为初始位置, p_{best} 中的最好值为 g_{best} .

2. 亲和力计算. 分别计算该模式下训练集合 A 与非记忆抗体 Nr 和记忆抗体集 M 中同类别抗体之间的主元核形式的亲和力.

3. 位置与速度更新. 选择亲和力最高的 n_1 个抗体根据式(1)、(2)进行位置与速度的更新, 更新全局最优和个体历史最优; 从而产生选择集 C .

4. 亲和力计算. 重新计算抗原 A 与 C 的亲和力, 删除亲和力小于自然死亡阈值的记忆抗体, 选择 n_2 个亲

和力高的抗体构成临时记忆抗体集 M . 对种群中的每一个个体进行适应度的评价并更新粒子位置与速度.

5. 选择抑制. 计算临时记忆抗体集 M 中记忆抗体之间的亲和力, 删除亲和力大于免疫抑制阈值的抗体.

6. 抗原提呈. 选择训练集中下一个抗原转至步骤 2, 直到该类所有的抗原提呈完毕.

7. 记忆抗体抑制. 对部分记忆抗体集进行抑制, 直至记忆抗体集规模维持在 n_3 , 否则继续步骤.

8. 随机产生 n 个抗体, 转至步骤 2 继续迭代.

经免疫预处理后的记忆细胞集可以表征故障样本特征, 故以记忆细胞集为初始解候选集, 由粒子群算法寻优动态获得原样本集中最佳解空间.

在算法中, 产生初始抗体识别抗原. 根据应答的情况, 初次应答则初始抗体全部随机产生; 再次应答则初始抗体部分随机产生, 部分通过激活免疫记忆细胞, 在包含最优抗体的免疫记忆细胞库中产生.

5.4 免疫粒子群算法求解过程

在免疫粒子群算法中, 将待求解的故障征兆对应为抗原, 搜索空间内的候选解对应为抗体, 根据抗原与抗体的亲和力以及抗体之间的结合力对候选解进行评价和选择, 通过记忆细胞保留局部最优解. 免疫粒子群算法的主要模块如图 1 所示.

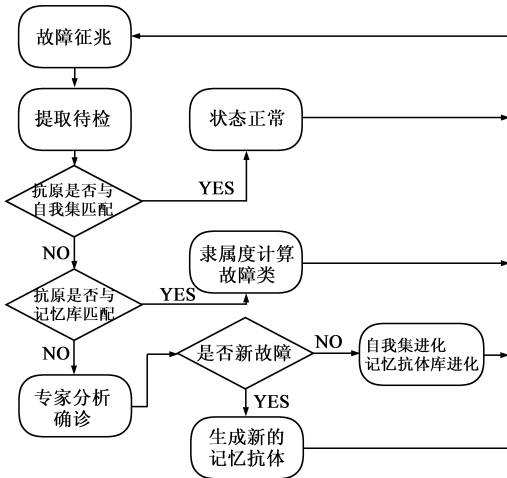


图1 基于免疫原理的故障诊断流程图

首先判断待检抗原是否自我集元素匹配, 根据反向选择算法对正常/异常状态进行判别. 对异常状态的抗原, 再根据各记忆抗体集的阈值判断其是否记忆抗体集匹配. 如待检抗原与已知记忆抗体集都不匹配, 诊断系统就会自动记录该抗原, 并提请专家人工分析. 如专家确诊为新抗原, 则将其加入训练抗体集 A , 实现记忆抗体库的整体进化、更新.

6 算例诊断与验证

由上述分析, 根据提升机系统模型和算法诊断模型, 首先要确定故障诊断问题的症状集、故障集. 所有

故障信号由故障提升系统产生并通过串口通信协议 RS232 传递给诊断模型, 通过知识库对传入诊断模型

6.1 诊断数据处理

为了得到足够多的输入样本数据, 分别在标准故障征兆模式的基础上, 按(12)式对变量进行叠加.

$$X = X_S + K\delta Random(1, m) \quad (12)$$

式中 X_S 表示标准故障征兆模式样本, K 表示控制变量幅值, δ 为标准差, 函数 $Random(1, m)$ 产生 $[0, 1]$ 之间的随机数. 对矩阵 X 按式(13)进行标准化处理,

$$X'_{mn} = (X_{mn} - \alpha_m) / \delta_m \quad (13)$$

式中 $(X'_{mn})_{1 \leq m \leq 10, 1 \leq n \leq q}$ 为数据矩阵, α_m 为 X 中的元素每列数据的平均值, δ_m 为 X 中的元素每列数据的标准差.

经过上述处理后, 得到标准化的数据矩阵 X' , 求出矩阵 X' 的协方差矩阵的规格化特征向量系统, 得到每个模式类的主元核.

6.2 诊断示例

采用免疫选择算法和标准遗传算法分别对 10 维原始特征进行类别数为 8 的聚类分析. 算法中的其他参数设置如下: 抗原个数为 15, 抗体个数为 8, 抗原抗体净化阈值 0.15.

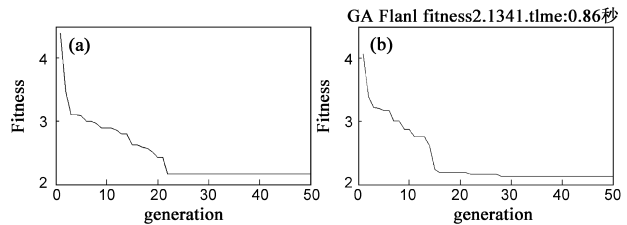


图2 种群规模为10时的适应值变化趋势比较

对于 8 种故障模式我们每种用到 200 个样本数据. 为了测试本文提出的 AI-DPSO 的识别故障效率, 在相同参数设置情况下将结果和 GA 进行了比较. 图 2 分别是对 60 个发生瓦闸系统卡缸故障的样本数据用 AI-DPSO 和 GA 测试的平均结果, 在此将种群规模设置为 10, 对于 GA 的交叉率设为 0.85, 变异率设为 0.09. 从图 2 中可以看出两种算法最终结果差别不是很明显. 因此我们将种群规模设为 50, 从图 3 可以看到 AI-DPSO 的收敛速度提高了, 而 GA 在种群规模扩大的情况下执行效率

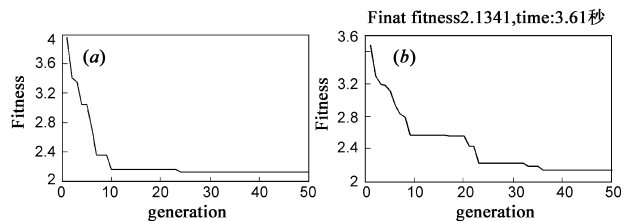


图3 种群规模为50时的适应值变化趋势比较

却降低了。

图 4 分别显示了对于油压站过压、瓦闸系统卡缸以及油泵故障三种故障模式下的标准训练数据进行实验后的平均识别结果。结果表面通过文中提出的方法我们可以很好的对故障进行分类识别。

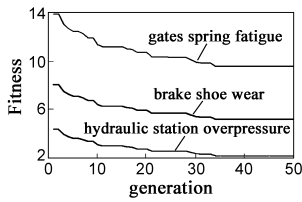


图 4 不同故障模式下的适应值变化趋势

表 2 典型训练样本集

	E ₁	E ₂	E ₃	E ₄	E ₅	E ₆	E ₇	E ₈	E ₉	E ₁₀
F ₁	2.3	-0.4	-0.4	2.3	-0.4	-0.4	2.3	-0.4	-0.4	0
F ₂	0	0	0	0	0	0	-2.2	2.2	2.2	-2
F ₃	-1.5	-1.5	0.38	0.38	0.38	-3.4	0.38	0.38	0.38	0.4
F ₄	-1.9	0.48	-1.9	0.48	0.48	-1.9	-1.9	0.48	0.49	0.5
F ₅	-2.9	0.32	0.32	-2.9	0.32	0.32	0.32	0.32	0.32	0.3
F ₆	-3.1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0
F ₇	-0.3	-0.3	2.9	2.9	-0.3	-0.3	-0.3	-0.3	-0.3	0
F ₈	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	-2.9	-3

最终我们根据标准训练数据实验的结果,得到每种故障模式下的最优故障征兆数据,经过测试振动信号,特征向量提取后,获得样本数据如表 2 所示。

7 结束语

就免疫系统机理本身来讲,也存在缺憾。为了构造防御系统(最优可行解集合),获得初始抗体(解特征样本),需要进行大量计算^[10]。本文将人工免疫模型与粒子群算法相结合,提出一种新颖的求解极小诊断的算法,并引入到提升机故障诊断中,为提升机系统建立了一个故障诊断系统。实验证明,该系统能够有效识别出系统各种典型故障,方法切实可行。

参考文献:

- [1] Dasgupta D, KrishnaKumar K, Wong D, Berry M, et al. Negative selection algorithm for aircraft fault detection[oL]. <http://issrl.cs.memphis.edu/papers/ais/2004/ICARIS04.pdf>.
- [2] GC Luh, WC Cheng. Immune model-based fault diagnosis[J]. *Mathematics and Computers in Simulation*, 2005, 67(6): 515 - 539.
- [3] Thomas Lukasiewicz. Probabilistic description logic programs [J]. *International Journal of Approximate Reasoning*, 2007, 45 (2): 288 - 307.

- [4] Henry Y. K. Lau, Vicky W. K. Wong. An immunity approach to strategic behavioral control[J]. *Engineering Applications of Artificial Intelligence*, 2007, 20(3): 289 - 306.
- [5] B Liu, L Wang, YH Jin, F Tang, DX Huang. Improved particle swarm optimization combined with chaos[J]. *Chaos, Solitons and Fractals*, 2005, 25(5): 1261 - 1271.
- [6] J Pugh, A Martinoli. Discrete Multi-Valued Particle Swarm Optimization[A]. In *Proc. IEEE Swarm Intelligence Symposium [C]*. Switzerland: 2006.
- [7] 张长胜, 孙吉贵, 欧阳丹彤. 一种自适应离散粒子群算法及其应用研究[J]. *电子学报*, 2009, 37(2): 299 - 304. ZHANG Chang-sheng, SUN Ji-gui, OUYANG Dan-tong. A Self-Adaptive Discrete Particle Swarm Optimization Algorithm [J]. *Acta Electronica Sinica*, 2009, 37(2): 299 - 304. (in Chinese)
- [8] F Caccavale, F Pierri, L Villani. Adaptive Observer for Fault Diagnosis in Nonlinear Discrete-Time Systems [J]. *Journal of Dynamic Systems, Measurement, and Control*, 2008, 130(2): 961 - 970.
- [9] M Filippone, F Camastra, F Masulli, S Rovetta. A survey of kernel and spectral methods for clustering[J]. *Pattern Recognition*, 2008, 41(1): 176 - 190.
- [10] 焦李成, 杜海峰. 人工免疫系统进展与展望[J]. *电子学报*, 2003, 35(10): 1540 - 1548. JIAO Li-cheng, DU Hai-feng. Development and Prospect of the Artificial Immune System [J]. *Acta Electronica Sinica*, 2003, 35(10): 1540 - 1548. (in Chinese)

作者简介:



汪楚娇 女. 1976 年出生于湖北麻城. 博士生, 从事智能优化方法 煤矿数字化的研究.
E-mail: wej@cumt.edu.cn



夏士雄 男. 1961 年 10 月出生. 教授, 博士生导师, 从事智能优化方法 煤矿数字化的研究.
E-mail: xiasx@cumt.edu.cn