

模拟电路故障诊断的信息融合新方法

马 超, 陈西宏, 邓均明, 王光明

(空军工程大学 导弹学院, 陕西 三原 713800)

摘 要: 针对测试信息不足造成模拟电路故障诊断准确率较低的问题, 为充分利用有限测试信息, 提出一种模拟电路故障诊断信息融合新方法. 首先, 将采集的故障样本集变换到不同特征空间, 然后利用所提出的马氏距离分布熵求取各特征空间的相对优势分类集, 在此基础上, 定义相对优势属性约简提取各特征空间的局部最佳可分性信息, 最后, 对基分类器结合所提出的自适应类模糊密度赋值方法进行模糊积分融合. 国际标准电路故障诊断实例表明, 所提方法能有效提高模拟电路的故障诊断率.

关键词: 模拟电路故障诊断; 信息融合; 相对优势属性约简; 模糊密度赋值

中图分类号: TP181; TN707

文献标识码: A

文章编号: 1000-7180(2011)04-0009-04

A New Information Fusion Algorithm for Analog Circuit Fault Diagnosis

MA Chao, CHEN Xi-hong, DENG Jun-ming, WANG Guang-ming

(The Missile Institute of AFEU, Sanyuan 713800, China)

Abstract: To solve the problem that test information in analog circuit diagnosis is always insufficient, a new algorithm of information fusion is presented to improve the diagnosis accuracy in analog circuit. Firstly, fault samples are transformed in different feature spaces, and the relative dominance class set for each feature space is computed by distribution entropy of mahalanobis distance. Subsequently, the relative dominance attribute reduction is defined to extract local prime separability information. And finally, the trained basic classifiers are ensembled by fuzzy integral. The fault instance indicates that the presented method can improve diagnosis accuracy effectively.

Key words: analog circuit fault diagnosis; information fusion; relative dominance attribute reduction; fuzzy density evaluation

1 引言

模拟电路故障诊断是近代电路理论的一个前沿领域, 其基础理论研究始于 1970 年代初的元件可解性问题的研究^[1]. 可测节点的不足以及测试手段的限制使得模拟电路故障诊断所能利用的测试信息十分有限, 不利于故障诊断准确率的提高. 信息融合技术能将不同的诊断信息进行有效融合, 是解决测试信息不足的有效方法, 已经被应用到模拟电路故障诊断领域^[2-5]. 文献[2]通过测试被诊断元件的工作温度和工作电压得出两传感器的信度函数分配,

利用 D-S 联合规则进行融合确定故障元件. 文献[3-4]采集了可测节点电压和不同激励下的输出输入增益两类测试信息, 利用神经网络进行初级诊断, 分别通过 D-S 证据理论和模糊积分进行融合诊断.

上述方法中信息融合技术使用的诊断信息均来自异类测试信息, 但测试手段的限制使这些方法实用性不强, 为进一步充分利用测试信息, 本文将故障样本集变换到不同特征空间, 然后在每个特征空间内提取包含局部最佳可分性诊断信息的特征子集, 用所提取特征子集训练基分类器, 最后对基分类器进行集成以实现不同特征空间局部最佳可分性诊断

收稿日期: 2010-09-13; 修回日期: 2010-12-23

基金项目: 国家自然科学基金项目(60971118)

信息的有效融合。

2 融合诊断模型

设模拟电路故障样本集合为 $S = \{x_1, x_2, \dots, x_n\}$, 故障类别为 $D = \{X_1, X_2, \dots, X_d\}$. 经过 m 种特征提取方法, 得到样本集合 S 的 m 个特征空间, 各特征空间的特征记为 $\{C_1 = (c_1, c_2, \dots, c_{l_1}), C_2 = (c_1, c_2, \dots, c_{l_2}), \dots, C_m = (c_1, c_2, \dots, c_{l_m})\}$. 将 m 个特征空间组成多决策表系统 $T = \{T_1, T_2, \dots, T_m\}$, 其中 $T_j = (U, C_j, D)$, $U \subseteq S$ 为训练样本。

2.1 相对优势属性约简

利用文献[6]中在每个子决策表上建立邻域粗集模型, 提出相对优势属性约简(relative dominance attribute reduction, RDAR)的概念以获得包含各子决策表最佳局部可分性信息的特征子集。邻域粗集模型及其性质见文献[6]。

提取局部最佳可分性信息, 首先要定义样本的可分性判别准则。本文提出一种新的样本类别可分性判据—马氏距离分布熵(distribution entropy of mahalanobis distance, DEMD)。

定义 1 样本集中各点相对于空间某点的马氏距离直方图的熵为马氏距离分布熵。

定义 2 某类样本集各点相对于该类样本集均值的马氏距离分布熵为自马氏距离分布熵(ADEMD), 相对于异类样本集均值的马氏距离分布熵为互马氏距离分布熵(CDEMD)。

定义 3 定义某类样本 X_i 的互马氏距离分布熵之和与其自马氏距离分布熵的比值为其类别可分性信息 $H(X_i)$, 即 $H(X_i) = \sum_{i=1}^{d-1} C_{E'} / A_E$ ($C_{E'}$ 为 X_i 的 CDEMD, A_E 为 X_i 的 ADEMD)。

由熵的性质可知, 当样本空间的样本点均匀分布时, 马氏距离分布熵为最大; 当样本空间的样本点分布趋于集中时, 马氏距离分布熵会变小。所以, 当 A_E 较小时, X_i 中的样本集在空间中聚集度高, 而当 A_E 较大时, 聚集度低; 当 C_E 较小时, X_i 中样本与其他类别样本间可分性差, 当 C_E 较大时, 样本间可分性好。由熵的可加性原理可知, $H(X_i)$ 越大, X_i 中样本的可分性越好, 因此 $H(X_i)$ 可作为 X_i 的可分性判别准则。

马氏距离的计算公式为

$$(x_i - \bar{\omega})^T J^{-1} (x_i - \bar{\omega}) = D_i^2 \quad (1)$$

式中, 样本 $x_i \in X_i$, D_i^2 为样本 x_i 的马氏距离; $\bar{\omega}$ 为样本均值; J 为样本协方差矩阵; J^{-1} 为 J 的逆矩阵。

令 $D_{\max}^2 = \max(D_i^2)$, $D_{\min}^2 = \min(D_i^2)$, 则对于样本 x_i , 其 $D_i^2 \in [D_{\max}^2, D_{\min}^2]$. 设常数 $M (M \ll m, m$ 为 X_i 的样本个数), 取 $\Delta D^2 = (D_{\max}^2 - D_{\min}^2) / M$, 得到 M 个连续区间 $\Omega_i, i = 1, 2, \dots, M$. 设马氏距离 D_i^2 中, 属于区间 Ω_i 的数目设为 $p_i (i = 1, 2, \dots, M)$. 马氏距离 D_i^2 属于区间 Ω_i 的概率为 $P_i = p_i / m, i = 1, 2, \dots, M$, 由此可得马氏距离分布直方图。对马氏距离统计直方图求熵, 可得马氏距离分布熵

$$E(P) = - \sum_{i=1}^M P_i \log_2 P_i \quad (2)$$

当 $\bar{\omega}$ 为同类样本均值即可按照上述过程求得自马氏距离分布熵, 当 $\bar{\omega}$ 为异类样本均值时可求得互马氏距离分布熵。

对每个子决策表 $T_i = (U, C_j, D)$, 求可分性判别向量 $H_j = (H_j(X_1), H_j(X_2), \dots, H_j(X_d))$ 及其相对优势分类集:

$$R_j = \{i, H_j(X_i) = \max_{j' \leq d} H_{j'}(X_i)\} \quad (3)$$

若 $R_j \neq \emptyset$, 称 $T_j = (U, C_j, D)$ 为相对优势决策表。

定义 4 设 $T_j = \langle U, C_j, D \rangle$ 为相对优势决策表, Y_j 为决策表 T_j 的相对优势分类集, 称 $B \subseteq C_j$ 为相对优势属性约简, 若 B 满足:

- (1) $\forall a \in B, Pos_{B-a}(R_j) < Pos_B(R_j)$;
- (2) $Pos_B(R_j) = Pos_{C_j}(R_j)$

利用文献[7]提出的基于遗传算法的数值属性约简算法求多决策表系统的所有相对优势属性约简, 记为 $R_{DAD1}, R_{DAD2}, \dots, R_{DAD_{q'}} (q' \geq q)$. 设 $L = \{b_1, b_2, \dots, b_{q'}\}$ 为由相对优势属性约简训练所得基分类器, 则每个基分类器都包含了各特征空间的局部最佳可分性诊断信息。

2.2 糊积分融合

常用的多数投票法不能融合基分类器度量层提供的丰富信息, 所以本文采用模糊积分法对基分类器包含的可分性诊断信息进行融合。

定义 5 设 L 为任意集合, $P(L)$ 为 L 的幂集。如果集合函数 g 满足:

- ① $g(\emptyset) = 0, g(L) = 1$;
- ② $g(A) \leq g(B)$, if $A \subset B$ 且 $A, B \in P(L)$;
- ③ $\lim_{i \rightarrow \infty} g(A_i) \leq g(\lim_{i \rightarrow \infty} A_i)$, 若 $A_i \in P(L)$, 且 A_i 是单调的, 称 g 是 $P(L)$ 上的一个模糊测度。

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B) \quad (\lambda > -1) \quad (4)$$

若 g 对所有的 $A, B \in P(L)$, 且 $A \cap B = \emptyset$, 还满足式(4)所示的可加性时, 就称为 Sugeno 测

度 g_λ .

$L = \{b_1, b_2, \dots, b_q\}$ 为基分类器集合, 并记 $g_i = g_\lambda(\{b_i\})$, $g_1, g_2, \dots, g_q$ 的值称为模糊密度, $g_\lambda(L)$ 可表示为

$$g_\lambda(L) = \sum_{i=1}^n g_i + \lambda \sum_{i=1}^{n-1} \sum_{j=i+1}^n g_i g_j + \lambda^{n-1} g_1 g_2 \quad (5)$$

当 $\lambda \neq 0$ 时, 根据模糊测度的性质, λ 值可由下式确定:

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g_i) \quad (6)$$

设 f 为定义在集合 Q 上的可测函数, 它实现映射 $f: Q \rightarrow [0, 1]$, 则 f 在 $A \subset Q$ 上关于模糊测度 g 的 Sugeno 模糊积分定义为

$$\int_A f(\cdot) g(\cdot) = \max_{i=1}^n \{ \min[f(q_i), g(A_i)] \} \quad (7)$$

当 g 为 g_λ 模糊测度时, $g(A_i)$ 可按式递归算得

$$\begin{cases} g(A_i) = g(\{q_1\}) = g \\ g(A_i) = g_i + g(A_{i-1}) + \lambda g_i g(A_{i-1}) \\ 1 < i \leq n \end{cases} \quad (8)$$

模糊密度赋值是实现模糊积分融合的关键, 其赋值方法应根据具体应用进行调整. 本文的模糊密度赋值应考虑两点, 一是包含局部最佳可分析信息的基分类器对样本空间的某一类或某几类具有较高的分类精度, 所以本文采用类模糊密度赋值方法来体现同一分类器判决样本属于不同类别时重要性的不同. 二是实际的模拟电路故障诊断过程中, 难免受到噪声、干扰等因素的影响, 所以模糊密度应体现样本判决结果包含的动态信息. 为此, 本文提出自适应类模糊密度赋值方法.

设 μ_{ij} 为基分类器 b_i 对某类样本 X_j 的分类精度, 则分类器 b_i 对类别 X_j 的类模糊密度为

$$g_i^j = \mu_{ij} / \sum_{i=1}^q \mu_{ij} \quad (9)$$

由式(9)可知, 分类器类模糊密度 g_i^j 只与分类器 b_i 在类别 X_i 上的分类精度有关, 而与在其他类别上的分类精度无关, 保证了高数值分类器类模糊密度在其对应样本的分类过程中能发挥应有作用.

为了使类模糊密度 g_i^j 根据具体待测样本自适应调整, 针对任意样本 x_i , 对每个类模糊密度 g_i^j 定义一个系数 $\gamma_{ij}(x_i)$:

$$\gamma_{ij}(x_i) = \varepsilon_{ij}(x_i) - \frac{1}{d} \sum_{j=1}^d \varepsilon_{ij'}(x_i), i = 1, 2, \dots, q \quad (10)$$

$$\gamma_{ij}^*(x_i) = (\gamma_{ij}(x_i) - \gamma_{ij}(x_i)_{\min}) / (\gamma_{ij}(x_i)_{\max} - \gamma_{ij}(x_i)_{\min}) \quad (11)$$

式中, $\varepsilon_{ij}(x_i)$ 表示分类器 b_i 判断样本 x_i 属于类别 X_j 的概率. $\gamma_{ij}^*(x_i)$ 表示 $\gamma_{ij}(x_i)$ 的归一化结果.

由式(10)可知, $\gamma_{ij}(x)$ 的意义是当分类器 b_i 判断样本 x 属于类别 X_j 的概率大于分类器 b_i 判断样本 x 在各类中的平均概率时, 类模糊密度 g_i^j 应该适当提高, 且 $\gamma_{ij}(x)$ 越大, 提高的幅度越大, 反之亦然.

根据式(11)对类模糊密度 g_i^j 进行修正:

$$g_i^{j*} = \gamma_{ij}^*(x) \cdot g_i^j \quad (12)$$

2.3 融合诊断模型

基于相对优势属性约简和模糊积分的模拟电路故障诊断模型如图1所示.

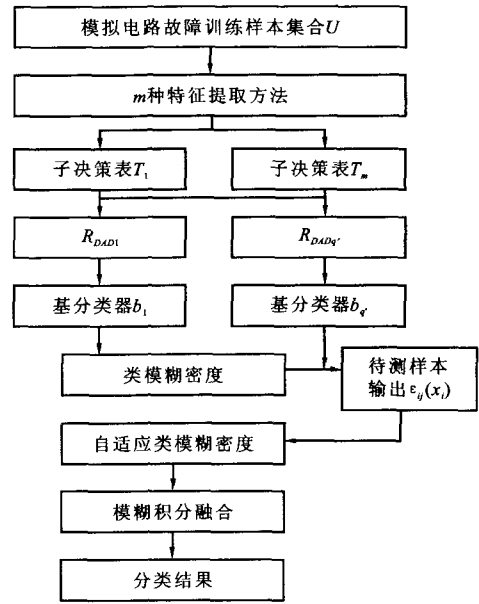


图1 融合诊断模型

3 诊断实例

以双二次滤波电路为诊断实例验证本文方法的优越性. 双二次滤波电路结构如图2所示.

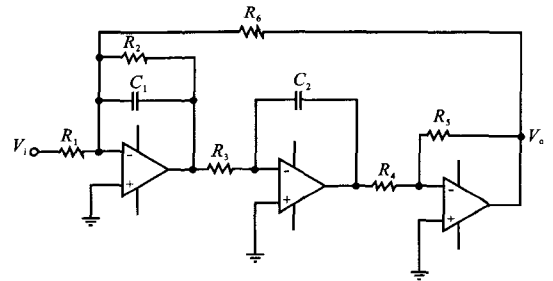


图2 双二次滤波器电路

电路中, 每个元件的标称值为: $R_1 = R_2 = R_3 = R_4 = R_5 = R_6 = 10 \text{ k}\Omega$, $C_1 = C_2 = 20 \text{ nF}$, 电阻和

电容的容差均为 5%, 假设发生软故障时, 元件值偏离其标称值的 $\pm 50\%$. 根据灵敏度分析, R_1 、 C_2 变化时对电路的频率响应影响最大. 假设电路发生单个软故障, 分别是: $R_1 \uparrow$, $R_1 = 15 \text{ k}\Omega$; $C_2 \uparrow$, $C_2 = 30 \text{ nF}$; $R_1 \downarrow$, $R_1 = 5 \text{ k}\Omega$; $C_2 \downarrow$, $C_2 = 10 \text{ nF}$, 其中 $\uparrow$ 代表增故障, $\downarrow$ 代表减故障. 所有电路分析仿真均使用 Multisim7 电路仿真平台. 激励信号为 $(\sin 20\pi kt + \sin 50\pi kt + \sin 80\pi kt + \sin 110\pi kt)$, 对其时域响应按 250 kHz 进行 2.048 ms 的采样, 得到 512 个数据点. 对于电路的正常状态和各种故障状态分别进行 100 次蒙特卡洛仿真获得故障样本集, 共得到分为 5 类的 500 个样本, 每个样本包含 512 个点, 每类故障取 50 个故障向量作为训练样本, 其余 50 个为待测样本. 设为训练样本集, $D = \{X_1, X_2, X_3, X_4, X_5\}$ 为样本分类, 代表 R_1 增故障、 C_2 增故障、 R_1 减故障、 C_2 减故障和无故障五种电路状态.

利用三种特征提取方法对故障样本集合进行变换.

(1) 小波变换. 采用 haar 小波对样本空间进行 6 层小波分解, 提取经归一化处理的第 6 层低频频带和各层高频频带的能量值 $(E'_1, E'_2, \dots, E'_7)$ 作为故障特征, 每个故障向量包含 7 个特征. 建立子决策表 $T_1 = (U, C_1, D)$, 其中 $C_1 = (E'_1, E'_2, \dots, E'_7)$.

(2) 小波包变换. 采用 db5 小波对样本空间进行 3 层小波包分解, 提取经归一化处理的第 3 层所有频带的能量值 $(E'_{30}, E'_{31}, \dots, E'_{37})$ 作为故障特征, 每个故障向量包含 8 个特征. 建立子决策表 $T_2 = (U, C_2, D)$, 其中 $C_2 = (E'_{30}, E'_{31}, \dots, E'_{37})$.

(3) 采集响应曲线上等距离 8 个离散点然后进行归一化处理后 $(t'_1, t'_2, \dots, t'_8)$ 作为故障特征, 每个故障向量包含 8 个特征. 建立子决策表 $T_3 = (U, C_3, D)$, 其中 $C_3 = (t'_1, t'_2, \dots, t'_8)$.

设 $M = 30$, 利用马氏距离分布熵求 T_1, T_2, T_3 的可分性判别向量; $H_1 = (1.845 \ 1, 1.557 \ 4, 1.421 \ 3, 1.214 \ 6, 1.789 \ 4)$; $H_2 = (1.664 \ 2, 1.247 \ 8, 1.645 \ 7, 1.243 \ 9, 1.645 \ 8)$; $H_3 = (1.735 \ 2, 1.498 \ 7, 1.542 \ 3, 1.211 \ 3, 1.795 \ 4)$. 可知, 各子决策表的相对优势分类集为 $R_1 = (X_1, X_2)$; $R_2 = (X_3, X_4)$; $R_3 = (X_5)$. 因 R_1, R_2, R_3 均为非空集合, 所以 T_1, T_2, T_3 均为相对优势决策表.

利用文献[7]的遗传算法, 参数设置为: 种群规模 $g = 25$; 变异概率 $P_g = 0.02$, 其余参数设置同文献[7]. 求得相对优势决策表 T_1, T_2, T_3 的相对优势约简分别为 $\{(E'_1, E'_3, E'_6), (E'_2, E'_3, E'_7)\}, \{(E'_{30}, E'_{32},$

$E'_{33}, E'_{37}), (E'_{30}, E'_{34}, E'_{35}, E'_{37}), (E'_{31}, E'_{34}, E'_{36}), \{(t'_1, t'_3, t'_5, t'_7), (t'_2, t'_3, t'_4, t'_6)\}$.

在求得的相对优势属性约简上训练基分类器 $\{b_1, b_2, \dots, b_7\}$, 在基分类器上对训练样本进行分类得到各分类器的类模糊密度 $g_j^1, g_j^2, \dots, g_j^7, 1 \leq j \leq 5$. 根据待测样本的概率输出对类模糊密度进行修正得 $g_j^{1*}, g_j^{2*}, \dots, g_j^{7*}$. 根据式(5)~(8) 计算每一个类别 $X_i (i = 1, 2, \dots, d)$ 的模糊积分 e_k ; 输出最终决策 $k^* = \arg \max_k e_k$.

比较实验针对以下三种方法进行.

A: 神经网络诊断方法^[8];

B: 小波分析和信息融合诊断方法^[5];

C: 本文所提信息融合新方法.

比较实验结果如表 1 所示.

表 1 三种方法比较结果

	$[R_1 \downarrow]$	$[R_1 \uparrow]$	$[C_1 \downarrow]$	$[C_1 \uparrow]$	%
A	90	92	90	88	92
B	94	94	96	92	96
C	98	96	100	98	98

注: 方法 C 中的基分类器采用多分类支持向量机, 其概率输出模型同文献[9].

实验结果分析:

(1) 方法 A 和方法 B 都是采用神经网络作为分类器, 但方法 B 融合了两类测试信息包含的诊断信息, 所以取得了更高的分类精度.

(2) 方法 C 取得了最高的分类精度. 方法 C 的诊断信息提取的是各特征空间的局部最佳可分性信息, 保证了诊断信息间具有很强的互补差异性, 而方法 B 中的诊断信息虽来自异类测试信息, 但是没有在提取过程中保证诊断信息的互补差异性, 故方法 C 取得了更好的分类精度.

4 结束语

信息融合技术已经越来越多的应用于模拟电路故障诊断, 现有的方法都是针对不同类测试信息提取诊断信息, 为进一步挖掘有限测试信息, 利用不同特征变换方法从不同侧面提取样本集合包含的诊断信息, 最后通过合适的集成方法对提取的诊断信息进行融合, 实验结果表明, 该方法能有效提高模拟电路故障诊断准确率. 下一步的研究方向是新的模拟电路特征提取方法, 以及如何选择特征提取方法才能使获得的诊断信息具有更大的互补差异性.

(下转第 17 页)

4 结束语

本文通过对模型参数的自动获取技术进行深入研究,实现了主机间可达性、安全防护系统、主机系统和网络服务四个层面的目标网络自动建模。同时,由于状态枚举攻击图存在随着主机数增加而呈指数增长的缺点,因此本文根据预定义的规则进行聚合,生成结点和边数与网络中主机的增加成线性关系的主机攻击图。以上两方面使得本文提出的算法具有较高的效率,能够适用于大规模复杂的目标网络系统。

参考文献:

- [1] 程玲. MANET入侵检测技术的研究[J]. 微电子学与计算机, 2010, 27(6): 57-59.
- [2] 孙强. 基于定量安全风险评估模型的网络安全管理平台[J]. 微电子学与计算机, 2010, 27(5): 70-73.
- [3] Swiler L P, Phillips C, Ellis D, et al. Computer-attack graph generation tool[C]//Proceedings DARPA Information Survivability Conference and Exposition (DISCEX II'01). Anaheim, California: IEEE Computer Society, 2001: 307-321.

- [4] Swiler L P, Phillips C, Gaylor T. A graph-based network-vulnerability analysis system, SAND97-3010/1[R]. New Mexico and Livermore, California: Sandia National Laboratories, Albuquerque, 1998.
- [5] Sushil Jajodia, Steven Noel. Topological vulnerability analysis, center for secure information systems[J]. Computer Science, 2010, 46(4): 139-154.
- [6] Xinming Ou, Wayne F Boyer. A scalable approach to attack graph generation[C]// Proceedings of the 13th ACM conference on Computer and communications security, USA: ACM, 2006: 336-345.
- [7] Xinming Ou. A logic-programming approach to network security analysis[D]. Princeton: Princeton University, 2005.

作者简介:

马俊春 女, (1983-), 博士生. 研究方向为网络安全。

王勇军 男, (1971-), 博导, 教授. 研究方向为高性能网络隔离技术、网络脆弱性分析技术、网络安全态势分析评估技术、网络生存性技术等。

孙继银 男, (1952-), 博导, 教授, CCF 会员. 研究方向为指挥自动化工程, 计算机应用技术。

(上接第12页)

参考文献:

- [1] 禹旺兵, 彭良玉, 禹恒州. 基于小波分析和神经网络的模拟电路故障诊断方法[J]. 微电子学与计算机, 2007, 24(7): 43-46.
- [2] 朱大奇, 于盛林. 基于D-S理论的数据融合算法及其在电路故障诊断用的应用[J]. 电子学报, 2002, 30(2): 221-223.
- [3] 彭敏放, 何怡刚, 王耀南. 基于神经网络与证据理论的模拟电路故障诊断[J]. 电路与系统学报, 2005, 10(1): 35-39.
- [4] 彭敏放, 何怡刚, 王耀南, 等. 模拟电路的融合智能故障诊断[J]. 中国电机工程学报, 2006, 26(3): 19-24.
- [5] 刘美华, 彭良玉. 基于小波分析和信息融合技术的故障诊断[J]. 微电子学与计算机, 2009, 26(1): 166-168.
- [6] 胡清华, 于达仁, 谢宗霞. 基于邻域粒化和粗糙逼近的数

值属性约简[J]. 软件学报, 2008, 19(3): 640-649.

- [7] 曾庆双. 基于邻域粒化和遗传算法的数值属性约简方法[J]. 自动化博览, 2009(8): 69-74.

- [8] 廖薇, 许春冬, 刘锦高. 基于神经网络的模拟电路故障诊断研究[J]. 微电子学与计算机, 2010, 27(5): 125-128.

- [9] 颜根廷, 李传江, 马广富. 支持向量分类器的模糊积分集成方法[J]. 哈尔滨工业大学学报, 2008, 40(7): 1017-1020.

作者简介:

马超 男, (1983-), 博士研究生. 研究方向为防空导弹模拟电路智能故障诊断方法。

陈西宏 男, (1961-), 博士, 教授, 博士生导师. 研究方向为防空导弹武器系统信息技术与智能故障诊断方法等。