

实域粗糙集理论及属性约简

肖迪^{1,2} 胡寿松¹

摘要 本文引入属性和属性子集广义重要度的概念以及空间中的广义近邻关系, 提出广义近邻关系下的实域粗糙集扩展模型. 在实域粗糙集理论中, 利用广义近邻关系在全局中划分相容模块, 构成集合的下、上近似, 避免了 Pawlak 粗糙集必须量化数据的麻烦. 另外, 本文给出了实域粗糙集的属性约简定义和一种贪心算法, 分析了约简属性集合的质量. 最后, 通过实例验证了本文理论和方法的正确性和有效性.

关键词 实域粗糙集理论, 属性约简, 广义重要度, 广义近邻关系, 广义欧氏距离
中图分类号 TP18

Real Rough Set Theory and Attribute Reduction

XIAO Di^{1,2} HU Shou-Song¹

Abstract The notions of general important degree of an attribute and attributes subset are introduced, and thereupon a kind of general Euclidean distance and a general neighborhood relation are set up. And then a real rough set theory based on the general neighborhood relation is proposed. The theory partitions the universe into tolerant modules under general neighborhood relationship the give the lower and upper approximation of the set without discretizing the data. Furthermore, the definition and greedy arithmetic of attribute reduction in real rough set theory are proposed, and the quality of reduction results is analyzed. Finally, an example is given to show the validity and effectiveness of the theory and method of this paper.

Key words Real rough set theory, attribute reduction, general important degree, general neighborhood relation, general Euclidean distance

1 引言

在 Pawlak 粗糙集理论^[1,2] 或其他的扩展粗糙集理论, 如可变精度粗糙集^[3] 和相容粗糙集理论^[4] 的研究中, 属性值的离散化或量化问题是一个重点研究问题. 而离散化过程中断点集的选取是离散化效果优劣的关键. 这是因为断点的存在使得离散化对噪声非常敏感. 如果断点的位置很微妙, 数据被噪声污染前后的大小恰好位于断点位置的前后, 而这类断点又很多的话, 那么离散化数据表前后的差别就很大了. 因此离散化的优劣对于问题的解决影响很大. 为了避免离散化不当造成不良后果, 本文提出了实域粗糙集理论. 对于属性值集为连续实数的

决策系统, 它不需要经过离散化过程, 而根据属性的特征定义了属性的广义重要度, 从而可以重新度量空间中样本之间的相似性距离, 然后以广义欧氏距离为基础构成了空间中的广义近邻关系, 在此关系下定义了实域粗糙集及集合的关系和性质.

另外, 本文给出了实域粗糙集理论的属性约简的定义, 它与 Pawlak 粗糙集理论的属性约简定义不同. 实域属性约简是以全局的一般分类为着眼点, 而 Pawlak 属性约简是以全局的确定性分类为着眼点. 但是前者根据不确定分类的集合大小考虑了实域属性约简的质量问题, 因此它对确定性分类也是有效果的.

2 属性的广义重要度

定理 1. 在 Pawlak 粗糙集理论中, 属性重要度的值为零是该属性可约简的必要而非充分条件.

Pawlak 粗糙集中的属性重要度是针对确定性决策元素的影响进行研究的. 下面介绍一种广义属性重要度, 它是针对所有决策元素而研究的.

定理 2. 给定决策系统 $S = \langle U, C \cup d, V \rangle$, $U/d = \{d_1, d_2, \dots, d_{r(d)}\}$, $\exists a \in C$, 设 $a(d_i) \subset V$, 表示属性 a 对应决策 d_i 的属性值子集区间. 如果有 $a(d_i) \cap a(d_j) = \phi$, ($i \neq j, i, j = 1, 2, \dots, r(d)$), 则

收稿日期 2005-8-2 收修改稿日期 2006-4-24
Received August 2, 2005; in revised form April 24, 2006
国家自然科学基金 (60234010), 航空科学基金 (05E52031) 和国防基础科研项目 (K1603060318) 资助

Supported by National Natural Science Foundation of P. R. China (60234010), Aeronautical Science Foundation of P. R. China (05E52031) and the Basal Science Foundation of National Defense (K1603060318)

1. 南京航空航天大学自动化学院 南京 210016 2. 南京工业大学自动化学院 南京 210009

1. College of Automation Engineering Nanjing University of Aeronautics and Astronautics, Nanjing 210016 2. College of Automation Engineering, Nanjing University of Technology, Nanjing 210009

DOI: 10.1360/aas-007-0253

有 $U/a = U/d$ 成立.

定理 2 说明, 如果一个属性它对应不同的决策的值集是互不相交的, 则这个属性可以完全划分全局, 而且与决策属性划分一致. 当然, 这个属性也就是最重要的属性, 是属性集的核心. 因此可以说, 对一个属性值空间, 各类区间差异越大, 越容易根据此属性值判断分类, 因此该属性也越重要; 若某属性值各类区间完全一致, 则该属性必须联合其他属性值的特征一起进行决策分类, 相应地, 这个属性的重要度也就薄弱了. 这里的重要度就是广义重要度, 它度量的是对全局的决策能力, 而非对确定性决策元素的影响程度.

定义 1. 给定决策系统 $S = \langle U, C \cup d, V \rangle$, $U/d = \{d_1, d_2, \dots, d_{r(d)}\}$, $\forall a \in C$, 定义属性 a 的广义重要度为

$$\sigma_g(a) = \begin{cases} 1 - \frac{1}{C_{r(d)}^2} \sum_{\substack{i \neq j \\ i, j = 1}}^{r(d)} \frac{a(d_i) \cap a(d_j)}{\max \text{cross}(a(d))} & \text{其他} \\ 1 & \forall i, j, a(d_i) \cap a(d_j) = \phi \end{cases} \quad (1)$$

其中 $C_{r(d)}^2$ 表示从 $r(d)$ 个数中取 2 的组合, $a(d_i) \cap a(d_j) \subset V$ 表示属性 a 对应决策值 d_i 的属性值子集与对应决策值 d_j 的属性值子集的交集部分, $\max \text{cross}(a(d)) \subset V$ 表示属性 a 对应全部两两决策值的属性值子集的所有交集所包围的最大区间.

定义 2. 给定决策系统 $S = \langle U, C \cup d \rangle$, $U/d = \{d_1, d_2, \dots, d_{r(d)}\}$, $\forall B \subseteq C$, 则属性子集 B 的广义重要度为

$$\sigma_g(B) = \begin{cases} 1 - \frac{1}{C_{r(d)}^2} \sum_{\substack{i \neq j \\ i, j = 1}}^{r(d)} \frac{B(d_i) \cap B(d_j)}{\max \text{cross}(B(d))} & \text{其他} \\ 1 & \forall i, j, B(d_i) \cap B(d_j) = \phi \end{cases} \quad (2)$$

其中 $C_{r(d)}^2$ 表示从 $r(d)$ 个数中取 2 的组合, $B(d_i) \cap B(d_j)$ 表示属性子集 B 对应的决策 d_i 所包围的属性值区域与它对应的决策 d_j 所包围的属性值区域的交集, $\max \text{cross}(B(d))$ 表示属性子集 B 对应全部两两决策所有相交部分所包围的最大区域.

例. 下面给出一个简单的实数决策表, 分析一下广义属性重要度和广义属性子集重要度的定义和计算. 设实数值决策表如表 1 所示

表 1 实数值决策表

Table 1 Real decision table

U	a_1	a_2	a_3	d
x_1	1.6	0.4	5.7	1
x_2	2.2	0.6	6.8	1
x_3	1.9	0.4	5.3	2
x_4	3.4	0.5	4.8	2
x_5	2.0	0.4	6.9	3
x_6	2.9	0.6	7.5	3

对上面的决策表, $U/d = \{d_1, d_2, d_3\} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}\}$. 属性 a_1 的属性值全域为 $\max \text{cross}(a_1(d)) = [1.9, 2.9]$, 它的广义属性重要度为

$$\sigma_g(a_1) = 1 - \frac{1}{C_3^2} \sum_{i \neq j}^3 \frac{a_1(d_i) \cap a_1(d_j)}{[1.9, 2.9]} = 1 - \frac{1}{3} \left(\frac{[1.9, 2.2] + [2.0, 2.2] + [2.0, 2.9]}{[1.9, 2.9]} \right) = 0.533$$

同样地, 属性 a_2, a_3 的广义重要度分别为 $\sigma_g(a_2) = \frac{1}{3}$, $\sigma_g(a_3) = 1$. 属性 a_2 对应各个决策的值集重合的很大, 因此它的广义重要度就很小, 即对决策分类的帮助很小. 而属性 a_3 对应各个决策的值集均不相交, 因此它的广义重要度为 1, 即可凭借这一个属性进行分类.

设属性子集 $B = \{a_1, a_2\} \subset C$, 则属性子集 B 对应的空间为以 a_2, a_1 为横纵坐标形成的二维空间, 如图 1 所示. $\max \text{cross}(B(d))$ 为图 1 中的 A、B、C、D、E 和 F 区域的总和.

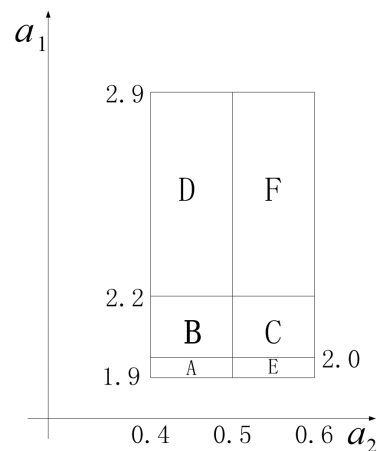


图 1 两属性的属性子集对应决策的相交值域划分

Fig. 1 Decision-making intersect range partition relative to attribute subset

决策 d_1 与决策 d_2 的 B 交集 $B(d_1) \cap B(d_2)$

为 A 加 B 的区域, 决策 d_2 与 d_3 决策的 B 交集 $B(d_2) \cap B(d_3)$ 为图 1 中 B 加上 D 的区域, 决策 d_1 与决策 d_3 的 B 交集 $B(d_1) \cap B(d_3)$ 为 B 加上 C 的区域.

经计算可得,

$$\frac{B(d_1) \cap B(d_2)}{\max cross(B(d))} = \frac{(2.2 - 1.9)(0.5 - 0.4)}{(2.9 - 1.9)(0.6 - 0.4)} = \frac{3}{20}$$

$$\sigma_g(B) = 1 - \frac{1}{C_3^2} \sum_{\substack{i \neq j \\ i, j = 1}}^3 \frac{B(d_i) \cap B(d_j)}{\max cross(B(d))} = 1 - \frac{1}{3} \left(\frac{0.03 + 0.04 + 0.09}{1.0 * 0.2} \right) = \frac{11}{15}$$

另外, 由计算可知, 属性集 C 的广义重要度为 $\sigma_g(C) = 1$.

定理 3. 给定决策系统 $S = \langle U, C \cup d, V \rangle$, 若 $B_1 \subseteq B_2 \subseteq C$, 则 $\sigma_g(B_1) \leq \sigma_g(B_2)$.

定理 4. 对一决策系统, 如果它的属性集的广义重要度的值为 1, 则该系统中的决策类是线性可分的, 否则就是线性不可分的.

实际上, 决策系统属性集的广义重要度若为 1, 表明应用该属性集可以为全局进行确定性的分类. 如果属性集的广义重要度不为 1, 则该决策系统中的各决策类之间存在交集, 这些相交的部分就不能确定分类, 因此可以利用粗糙集理论进行研究.

3 空间的属性广义重要度欧氏距离

很多情况下我们按照样本间的相似性把集合划分为若干个子集, 最后根据某种表示聚类质量的准则函数确定划分的结果. 一些常用的距离度量可以作为这种相似性度量. 最常用的就是欧氏距离, 例如均值聚类, 自组织映射等都是采用这种度量方法. 但是普通的欧氏距离仅仅是测量样本在特征空间中的实际距离, 而没有考虑每个特征向量对样本划分的不同影响, 因此仅从普通距离的角度度量样本间的相似性, 对分类方差不等、或呈非球形分布的样本来说, 一般分类结果不会很好. 因此, 我们提出利用第 2 节中的属性广义重要度, 定义一种空间的广义重要度距离, 它可以考虑到每个属性的特征特点及对分类决策的影响, 因而具有更好的分类相似性, 提高分类的准确率.

定义 3. 设 $\mathbf{x}, \mathbf{y} \in R^n$, a_1, a_2, \dots, a_n 表示 R^n 空间中的 n 个方向, 则 \mathbf{x}, \mathbf{y} 的属性广义重要度欧氏

距离为

$$d_g(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n \sigma_g(a_i)(x_i - y_i)^2}$$

简称为广义欧氏距离. 其中 $x_i, y_i (i = 1, \dots, n)$ 为 \mathbf{x}, \mathbf{y} 在各个方向上的分量.

如果不考虑属性的广义重要度, 也就是不考虑各个向量的特征情况, 每个样本在空间中的各个分量不具有特别的特征, 彼此都相等, 即 $\sigma_g(a_i) = 1, \forall i$, 则广义欧氏距离简化为普通的欧氏距离. 因此可以说欧氏距离是广义欧氏距离的一个特例.

4 实域粗糙集理论

设 U 为非空、有限的集合, 称作论域; A 为属性集合 $A = \{a_1, a_2, \dots, a_m\}$, 并且属性值集都是连续的实数值. 我们考虑广义欧氏距离下的一个二元关系.

定义 4. 设任意 $x, y \in U$, B 为 A 的一非空有限属性子集, 如果下面的不等式成立

$$d_g^B(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{a_i \in B} \sigma_g(a_i)(x_i - y_i)^2} < \delta$$

其中 δ 为正实数, 则称 y 是 x 在 B 下的一个 δ 广义近邻, 也可以说 (x, y) 满足 δ 广义近邻关系.

性质 1. δ 广义近邻关系满足自反性和对称性, 但不满足传递性, 因此是相容关系.

把 x 在 U 中的所有属性子集 B 下的 δ 广义近邻称作 x 的 B - 广义近邻类, 记作 $[x]_B^\delta$, 即

$$[x]_B^\delta = \{y \in U | d_g^B(\mathbf{x}, \mathbf{y}) < \delta\}$$

定义 5. 设 U 为论域, $X \subseteq U$, A 为全体属性集合, $B \subseteq A$, 则 X 在属性子集 B 下的实域下近似集和实域上近似集分别为

$$\underline{BX}^R = \{x \in U | [x]_B^\delta \subseteq X\}$$

$$\overline{BX}^R = \{x \in U | [x]_B^\delta \cap X \neq \emptyset\}$$

X 在属性子集 B 下的实域上、下近似集之差称作 X 在属性子集 B 下的实域边界集, 记作 $BN_B^R(X)$. 即 $BN_B^R(X) = \overline{BX}^R - \underline{BX}^R$.

定义 6. 对 $x \in U, X \subseteq U$, 考虑广义近邻关系下 x 隶属于 X 的程度

$$\mu_X^\delta(x) = \frac{|[x]_B^\delta \cap X|}{|X|}$$

称为广义近邻关系下的实域粗糙隶属度函数. 其中 $|\bullet|$ 表示基数. 显然, $\mu_X^\delta(x) \in [0, 1]$.

Pawlak 不可分辨关系下的粗糙隶属度函数的定义为

$$\mu_X^B(x) = \frac{|[x]_B \cap X|}{|[x]_B|}$$

两种粗糙隶属度函数定义的区别在于它们的分母是不同的. 不可分辨关系是一种等价关系, 它定义的 $[x]_B$ 是一个等价类, 类中的元素之间都具有等价的条件; 而广义近邻关系是一种相容关系, 它定义的 $[x]_B^\delta$ 是一个相容类, 即类中的元素之间不具备等价条件, 只能具备相似的特性.

性质 2. 广义近邻关系下的实域粗糙隶属度函数具有下面的性质:

- 1) 若 $\mu_X^\delta(x) = 1$, 则 $x \in \underline{BX}^R$ 不一定成立;
- 2) 若 $\mu_X^\delta(x) = 0$, 则有 $x \notin \overline{BX}^R$ 成立;
- 3) 若 $0 < \mu_X^\delta(x) < 1$, 则 $x \in \overline{BX}^R$;
- 4) 对任意 $x \in U$, $X, Y \subset U$, 且 $X \cap Y = \phi$,

则

$$\min(\mu_X^\delta(x), \mu_Y^\delta(x)) \leq \mu_{X \cup Y}^\delta(x) \leq \max(\mu_X^\delta(x), \mu_Y^\delta(x)).$$

性质 2 表明: 1) 即使元素相对于一集合的广义近邻粗糙隶属度为 1, 也不能确定这个元素肯定属于这个集合, 也就是说这个元素不是该集合的确定元素, 而只是可能的元素. 这种情况通常在全局中含有不相容决策时会发生. 故广义近邻关系下的粗糙隶属度只能考察元素与集合之间的近似关系. 2) 和 3) 与不可分辨关系下的粗糙隶属度的性质相同. 4) 表明如果全局中有两个不相交的集合, 那么元素对于这两个集合的并集的广义近邻实域粗糙隶属度值的大小介于这个元素分别对这两个集合的广义近邻实域粗糙隶属度的值之间.

5 实域粗糙集的属性约简定义及属性约简的贪心算法

定义 7. 设 $B \subseteq C$, 若满足下列条件:

- 1) $\sigma_g(B) = \sigma_g(C)$;
- 2) $\forall B' \subset B, \sigma_g(B') \neq \sigma_g(C)$

则称 B 为 C 的约简, 记作 $RED(C)$. 所有约简的交集称作属性的核, 记作 $CORE(C)$,

$$CORE(C) = \cap RED(C)$$

这里采用基于属性的广义重要度的 Beam 搜索^[5]方法进行属性约简.

输入: 决策系统 $S = \langle U, C \cup d \rangle$, $U/d = \{d_1, d_2, \dots, d_{r(d)}\}$, $C = \{a_1, a_2, \dots, a_n\}$ 输出: 属性集 C 的多个约简.

算法步骤如下:

1) 计算所有单个属性的广义重要度 $\sigma_g(a_i)$ ($i = 1, 2, \dots, n$), 形成 n 个 1-元组. 并计算属性集 C 的广义重要度 $\sigma_g(C)$;

2) 从 $\sigma_g(a_i)$ ($i = 1, 2, \dots, n$) 中选择 k 个最大值对应的属性 (k 表示 Beam 的宽度), 并以这些属性作为起点开始属性子集的搜索;

3) 加一个新属性到 k 个属性中的每一个上面, 形成 $k(n-1)$ 个 2-元组, 也就是这个阶段元组的尺寸为 2;

4) 计算每一个 t -元组的属性子集的广义重要度, 并选择广义重要度最大的 k 个属性子集;

5) 再在这 k 个 t -元组的基础上加上一个不在此 k 个 t -元组中的其他属性形成所有可能的 $(t+1)$ -元组;

6) 重复 4) 到 5), 直到某个 t -元组形成的属性子集的广义重要度的值等于 $\sigma_g(C)$, 那么这个 t -元组搜索停止, 而其他的 $(k-1)$ 个 t -元组则继续搜索, 直到也能得到它的属性子集的广义重要度的值等于 $\sigma_g(C)$ 则停止;

7) 最后得到的与 $\sigma_g(C)$ 相等的 t -元组对应的属性子集就是属性约简的结果.

这里的 Beam 搜索方法有一点小的修改, 就是最后的搜索结果中 t -元组中 t 的尺寸可能不都是相同的. 这是因为此处搜索结束的准则与原 Beam 搜索方法不同, 因此结果也不同.

和 Pawlak 粗糙集的属性约简一样, 实值粗糙集的属性约简也是不惟一的. 那么对于多个约简的属性子集来说, 哪一个属性约简结果是最好的呢. 这里根据约简的属性子集对于划分样本决策的不确定的边界集的大小来判断属性约简的质量, 并给出以下定义.

定义 8. 给定决策系统为 $S = \langle U, C \cup d, V \rangle$, $U/d = \{d_1, d_2, \dots, d_{r(d)}\}$. 设 $B_1, B_2 \in RED(C)$, 都是属性 C 的约简, 如果 $\sum_{i=1}^{r(d)} |BN_{B_1}^R(d_i)| < \sum_{i=1}^{r(d)} |BN_{B_2}^R(d_i)|$, 则称约简 B_1 的质量好于约简 B_2 .

这样定义属性约简的质量是因为我们希望得到的约简属性集在相同于整个属性集的分类能力的原则下, 并能使不确定分类的元素个数越少越好.

6 实例验证

以一组某型战斗机的故障飞行数据为例, 建立实数值决策系统如表 2 所示.

表 2 待属性约简的实数值决策系统

Table 2 Real decision system needing attribute reduction

U	a_1	a_2	a_3	a_4	a_5	a_6	d
x_1	1.11562	-0.07397	0.131876	5.0598	0.007014	-0.00046	1
x_2	3.24466	-0.20627	0.408133	5.0666	0.062166	-0.00398	1
...	1
x_{10}	8.58631	-0.60029	1.07413	5.12339	0.516987	-0.03475	1
x_{11}	1.41824	-0.06275	-0.01219	5.05872	0.018001	-0.00082	2
x_{12}	2.67102	-0.12168	-0.00732	5.05847	0.069575	-0.00313	2
...	2
x_{20}	9.23811	-0.56749	0.017876	5.05891	1.63701	-0.08979	2
x_{21}	1.18606	-0.09501	-0.13191	4.89014	0.829897	-0.05499	3
x_{22}	1.02323	-0.06498	-0.23608	4.98404	0.257833	-0.01412	3
...	3
x_{30}	2.02422	-0.1082	-0.22963	5.02577	0.248544	-0.01203	3

表 3 多种粗糙集方法比较

Table 3 Compare of manifold rough set method

项目	类型		
	实域粗糙集	Pawlak 粗糙集	可变精度粗糙集
属性约简集	$\{a_1, a_3, a_4, a_5\}$	$\{a_1, a_3, a_4, a_5\}$	$\{a_1, a_3, a_4, a_5\}$
约简运算时间	783 (ms)	1476 (ms)	2997 (ms)
噪声样本的分类能力 (正确率)	93.3%	63.3%	80%
新样本的分类能力 (正确率)	80%	30%	60%
方法比较	较优	较差	良

要求对表 2 的决策系统进行实域属性约简, 并用带有噪声的样本和新样本作为验证集, 与 Pawlak 粗糙集和可变精度粗糙集的属性约简进行分类能力的比较, 其中 Pawlak 粗糙集采用不可分辨矩阵方法^[6]进行约简, 可变精度粗糙集采用文^[7]中的近似增益约简方法, 结果如表 3 所示.

这里实域属性约简中, Beam 的宽度 $k = 2$, 但是因为最后只有一个属性子集的广义重要度与属性全集的广义重要度相同, 因此属性约简的结果只有一个集合, 为 $\{a_1, a_3, a_4, a_5\}$. 这种约简方法只需要计算属性和属性子集的广义重要度, 故它的运算时间相比于其他的两种约简方法要少很多. 不可分辨矩阵的属性约简方法要比较每两个元素的不可分辨属性子集, 还有经过析取、合取运算, 所以运算时间略长. 而可变精度粗糙集的属性约简更繁琐, 它要根据集合的正域、负域计算出局部相对增益和近似增益, 因此运算时间最长.

另外, 从表 3 的比较可知, 实域粗糙集属性约简的结果对 30 个加噪声 (噪声为高斯白噪声, 方差

为 0.00001) 的原样本和 10 个新样本的分类能力都好于其他两种方法. 因此实域粗糙集理论对这个决策系统的分类能力最优, 可变精度粗糙集方法其次, 而 Pawlak 粗糙集方法最差.

7 结束语

本文研究了实数域中粗糙集理论的一些问题, 如实域粗糙下、上近似集, 实域粗糙隶属度函数, 实域属性约简等. 并且提出了实域属性约简的一种基于 Beam 搜索的贪心算法, 对多个属性约简结果分析了约简的质量. 从理论上和实例验证上说明了实域粗糙集理论是正确和有效的. 但是, 对于含有语义属性的决策系统, 实域粗糙集理论是不适用的, 它仅对实数领域效果较好, 因此比较适合工程上的应用.

References

- 1 Zdzislaw Pawlak. Rough set theory. *Künstliche Intelligenz*, 2001, 15(3): 38~39

- 2 Zdzislaw Pawlak, Jerzy W. Grzymala-Busse, Roman Slowinski, Wojciech Ziarko. Rough sets. *Comm of ACM*, 1995, **38**(11): 88~95
- 3 Dominik Slezak, Wojciech Ziarko. Variable precision bayesian rough set model. In: Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. Chongqing, China, Springer Press, 2003. 312~315
- 4 Kim D. Data classification based on tolerant rough set. *Pattern Recognition*, 2001, **34**(8): 1613~1624
- 5 Gupta P, Doermann D, DeMenthon D. Beam search for feature selection in automatic SVM defect classification. *Pattern Recognition*, 2002, **19**(2): 212~215
- 6 Zhou P L, Salah M, Mingins Christine. An effective parallel attribute reduct algorithm based on discernibility matrix. In: Proceedings of the International Conference on Information and Knowledge Engineering. Las Vegas, USA, CSREA Press, 2003. 693~701
- 7 Dominik Slezak, Wojciech Ziarko. Attribute reduction in the bayesian version of variable precision rough set model. *Electronic Notes in Theoretical Computer Science*, 2003, **82**(4): 1~11

tronic Notes in Theoretical Computer Science, 2003, **82**(4): 1~11



肖迪 博士, 讲师, 研究方向为粗糙集理论, 神经网络等. 本文通信作者. E-mail: xiaodi_12@sina.com
(**XIAO Di** Ph.D., lecturer. Her research interest covers rough set theory and neural network. Corresponding author of this paper.)



胡寿松 教授. 主要研究领域为故障诊断、鲁棒控制、非线性系统的自修复等.
(**HU Shou-Song** Professor. His research interest covers fault diagnosis, robust control, and self-repairing of nonlinear systems.)