

文章编号: 1001-0920(2008)03-0258-05

基于完备容差关系的扩充粗糙集模型

盛立, 杨慧中

(江南大学 通信与控制工程学院, 江苏 无锡 214122)

摘要: 针对扩充粗糙集模型所处理的不完备信息系统之间存在的差异, 提出了信息系统完备度的概念; 在此基础上, 提出了基于完备容差关系的扩充粗糙集模型. 与基于容差关系、相似关系、限制容差关系等扩充粗糙集模型相比, 该模型既保留了已有模型的优点, 又在一定程度上克服了已有模型的局限性. 通过实例说明新模型对不完备信息系统的处理更符合实际情况.

关键词: 不完备信息系统; 粗糙集; 容差关系; 相似关系; 限制容差关系

中图分类号: TP18 **文献标识码:** A

Extended rough set model based on completed tolerance relation

SHENGLi, YANG Hui-zhong

(School of Communication and Control Engineering, Jiangnan University, Wuxi 214122, China. Correspondent: SHENGLi, E-mail: victory8209@yahoo.com.cn)

Abstract: The incomplete information systems which are processed by extended rough set models are different from each other. In order to describe these differences, the conception of completed degree is presented. The indiscernibility relation in classical rough set is extended to completed tolerance relation based on completed degree. Compared with the relations in other extensions of rough set, such as tolerance relation, similarity relation and limited tolerance relation, the new relation takes into account the differences between different incomplete information systems. A new extension of rough set based on the completed tolerance relation is put forward to deal with the incomplete systems. Two examples are given to illustrate the validity of the presented model.

Key words: Incomplete information system; Rough set; Tolerance relation; Similarity relation; Limited tolerance relation

1 引言

粗糙集理论是处理不精确、不确定信息的数学工具^[1-4]. 经过 20 余年的发展, 粗糙集理论在智能信息处理、模式识别、数据挖掘等领域得到了广泛的应用^[5-8]. 但经典粗糙集理论只适用于完备的信息系统, 即所有样本对象的属性值都是确切已知的. 由于数据采集能力有限等原因, 现实中绝大多数信息系统都是不完备的, 即某些样本对象的部分属性值是空值.

粗糙集理论适用于处理不完备信息系统. 目前主要有两种方法: 一种是间接处理方法, 即通过数据预处理, 把不完备信息表转化为完备信息表; 另一种是直接处理方法, 即在不完备信息系统下, 对经典粗糙集理论的相关概念进行扩充. 间接方法使用较为

简便, 但容易造成信息丢失, 且在填充信息过程中会引入新的不确定性. 直接方法无需进行数据预处理, 在一定程度上保留了原有系统的特征信息, 因此得到的处理结果更为客观.

常用的直接处理方法有基于容差关系^[9]、相似关系^[10]、限制容差关系^[11]、修正容差关系^[12]等多种扩充粗糙集模型. 这些模型在处理不完备信息时, 只考虑信息系统内部对象间的关系, 没有考虑由于属性值缺失程度不同, 不完备系统之间存在的差异.

本文分析比较了已有的粗糙集扩充模型, 针对不完备系统间的差异提出信息系统完备度的概念, 据此对限制容差关系进行改进; 进一步提出了基于完备容差关系的扩充粗糙集模型, 该模型充分利用不完备系统自身包含的信息, 克服了其他模型的局

收稿日期: 2006-11-24; 修回日期: 2007-04-02.

基金项目: 国家自然科学基金项目(60674092); 江苏省高技术研究项目(BG2006010).

作者简介: 盛立(1982—), 男, 山东莱阳人, 博士生, 从事粗糙集理论、复杂系统控制的研究; 杨慧中(1955—), 女, 江苏无锡人, 教授, 博士生导师, 从事过程建模、优化控制等研究.

限性,更适合对不完备信息系统的处理.

2 粗糙集扩充模型

本节给出对目前常用的几种扩充粗糙集模型的基本定义,并对其性质进行比较分析.

给定信息表 $S = (U, A)$, 其中 $A = C \cup \{d\}$, C 是条件属性集合, d 是决策属性. 记遗漏值为“*”.

Kryszkiewicz^[9] 给出的容差关系定义如下:

定义 1 容差关系 T 定义为

$$T(x, y) = \{(x, y) \in U \times U \mid \forall a \in A (a(x) = * \wedge a(y) = * \wedge a(x) = a(y))\}. \quad (1)$$

对象 $x \in U$ 的容差类, 在容差关系下, 对象集 X 的下近似(\underline{TX}) 和上近似(\overline{TX}) 分别定义为

$$\begin{cases} T(x) = \{y \in U \mid (x, y) \in T\}, \\ \underline{TX} = \{x \in U \mid T(x) \subseteq X\}, \\ \overline{TX} = \{x \in U \mid T(x) \cap X \neq \emptyset\}. \end{cases} \quad (2)$$

容易看出, 容差关系是自反和对称的, 但不一定是传递的. 容差关系认为, 未知值“*”与任何已知属性值相等, 会导致两个个体在没有明确相同已知属性信息的情况下(如 $x = \{1, *, 2, *, 3, *\}$ 和 $y = \{*, 4, *, 5, *, 6\}$), 划分在同一个容差类, 因此其划分粒度过大.

Stefanowski^[10] 等提出了非对称相似关系, 定义如下:

定义 2 相似关系 S 定义为

$$S(x, y) = \{(x, y) \in U \times U \mid \forall a \in A (a(x) = * \wedge a(y) = a(x))\}. \quad (3)$$

显然, 相似关系是自反和传递的, 但不一定对称. 对于任意对象 $x \in U$, 可定义两个相似类

$$\begin{cases} S(x) = \{y \in U \mid (y, x) \in S\}, \\ S^{-1}(x) = \{y \in U \mid (x, y) \in S\}. \end{cases} \quad (4)$$

在相似关系下, 对象集 X 的下近似(\underline{SX}) 和上近似(\overline{SX}) 分别定义为

$$\begin{cases} \underline{SX} = \{x \in U \mid S^{-1}(x) \subseteq X\}, \\ \overline{SX} = \{x \in U \mid S(x) \cap X \neq \emptyset\}. \end{cases} \quad (5)$$

基于相似关系的扩充粗糙集模型存在一定的缺陷: 两个对象大部分已知属性值相同(如 $x = \{1, 2, 3, 4, 5, *\}$ 和 $y = \{1, 2, 3, 4, *, 6\}$), 直观上可判为相似, 却因不满足相似关系而被划分到不同的相似类, 因此其划分粒度过小.

针对容差关系和相似关系的缺点, 王国胤提出了基于限制容差关系的粗糙集扩充模型如下:

定义 3^[11] 限制容差关系 L 定义为

$$L(x, y) = \{(x, y) \in U \times U \mid \forall a \in A (a(x) = * \wedge a(y) = * \wedge a(x) = a(y))\}.$$

$$\begin{aligned} a(y) = * & \wedge [P(x) \cap P(y) \neq \emptyset \\ \forall a \in A ((a(x) = * \wedge a(y) = *) & \\ (a(x) = a(y))) &] \}, \end{aligned} \quad (6)$$

其中 $P(x) = \{a \in A \mid a(x) \neq *\}$.

对象 $x \in U$ 的限制容差类, 在限制容差关系下, 对象集 X 的下近似(\underline{LX}) 和上近似(\overline{LX}) 定义为

$$\begin{cases} L(x) = \{y \in U \mid (x, y) \in L\}, \\ \underline{LX} = \{x \in U \mid L(x) \subseteq X\}, \\ \overline{LX} = \{x \in U \mid L(x) \cap X \neq \emptyset\}. \end{cases} \quad (7)$$

限制容差关系具有自反性和对称性, 但不具有传递性. 它的划分粒度介于容差关系与相似关系之间, 因此基于限制容差关系的扩充粗糙集模型优于基于容差或相似关系的扩充粗糙集模型. 但限制容差关系也存在一定的局限性. 文献[12]指出, 其划分粒度可能不一致, 并提出了基于修正容差关系的粗糙集模型.

3 基于完备容差关系的扩展粗糙集模型

3.1 信息系统的完备度

为了对不完备系统进行有效的处理, 上述各模型都对完备系统中的不可分辨关系进行扩充. 容差关系和相似关系是扩充不可分辨关系的两个极端: 容差关系划分粒度太大, 容易将没有相同已知属性信息的个体分到同一个容差类, 相似关系划分粒度太小, 可能将具有很多相同已知属性信息的个体分到不同的相似类. 限制容差关系划分粒度介于容差关系与相似关系之间, 但存在划分粒度可能不一致等不足.

这些扩充的粗糙集模型共同的特点是改变经典模型中的不可分辨关系, 构造新的二元关系, 以便对不完备信息系统中的对象进行有效分类. 但这些扩充模型都只考虑对象之间的关系, 没有考虑由于属性值缺失程度不同而引起的不完备系统之间的差异. 这种差异可能导致不同不完备系统中相同的两组对象具有不同的分类. 例如 $\{3, 2, 1, *\}$ 与 $\{*, 2, *, 1\}$ 这组对象, 在某些不完备系统中可能是可分辨的, 但在另一些不完备系统中可能是不可分辨的.

对不完备信息的理解, 存在遗漏语意和缺席语意两种语意解释^[13]: 遗漏语意认为未知值暂时缺失, 将来总能得到, 可与任意值匹配; 缺席语意认为未知值无法得到, 不能与任一值相比较. 可认为容差关系是在遗漏语意下解释不完备信息, 相似关系是在缺席语意下解释不完备信息. 实际上, 应从遗漏语意和缺席语意两方面来理解不完备信息系统. 为了描述不同不完备系统的差异, 引入信息系统完备度的概念如下:

定义 4 设不完备信息系统 $S = (U, A)$, 其

中: U 是对象集, A 是属性集. 对象 x_i 的完备度 $i = |P(x_i)| / |A|$, 式中: $P(x_i) = \{a \in A \mid a(x_i) = 1\}$, $x_i \in U$, $| \cdot |$ 表示集合“ \cdot ”的基数. 则整个信息系统 S 的完备度为

$$= \frac{1}{n} \sum_{i=1}^n |P(x_i)| / |A| \quad (8)$$

由定义知 $(0, 1]$. 较小说明信息系统的完备度较低, 即已知属性值较少而未知属性值较多; 较大说明信息系统的完备度较高, 即已知属性值较多而未知属性值较少.

由定义 4 知, 当对象数量 $|U|$ 较大时, 能在一定程度上反映信息系统的完备情况. 系统自身属性值的缺失情况, 对 S 也有很大的影响. 信息系统的完备度可辨别不同不完备系统之间的差异, 计算简单且仅依靠不完备信息系统自身的属性, 不需要人为地设定参数, 因此能较客观地描述这种差异.

3.2 完备容差关系

信息系统的完备度是对不完备信息系统特性的一种描述. 下面在此概念的基础上对限制容差关系进行改进, 给出完备容差关系.

定义 5 设不完备信息系统 $S = (U, A)$, 其中: U 是对象集, A 是属性集. 设 S 的完备度为 α , 在 U 上二元关系 M 为完备容差关系定义如下:

$$M(x, y) = \frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)} \geq \alpha \quad (9)$$

其中 $P(x) = \{a \in A \mid a(x) = 1\}$.

不同的不完备信息系统之间存在差异, 而完备度 α 能唯一表征系统的不完备情况, 因此用 α 作为判定完备容差关系的标准之一.

由定义知, 完备容差关系具有自反性和对称性, 但不具备传递性. 例如 $x_1 = \{3, *, 1, *\}$, $x_2 = \{3, 2, 1, *\}$, $x_3 = \{*, 2, *, *\}$, 由定义 5 容易算出

$$\frac{|P(x_1) \cap P(x_2)|}{\min(|P(x_1)|, |P(x_2)|)} = 1$$

所以 $(x_1, x_2) \in M$. 同理, $(x_2, x_3) \in M$, 但

$$\frac{|P(x_1) \cap P(x_3)|}{\min(|P(x_1)|, |P(x_3)|)} = 0 < \alpha$$

所以 $(x_1, x_3) \notin M$.

定义 6 设不完备信息系统 $S = (U, A)$, $x \in U, X \subseteq U$, M 是完备容差关系. 对象 x 的完备容差类 $M(x)$ 及完备容差关系下 X 的下近似 $\underline{M}X$ 和上近似 $\overline{M}X$ 分别定义为

$$\begin{cases} M(x) = \{y \in U \mid (x, y) \in M\}, \\ \underline{M}X = \{x \in U \mid M(x) \subseteq X\}, \\ \overline{M}X = \{x \in U \mid M(x) \cap X \neq \emptyset\}. \end{cases} \quad (10)$$

将完备容差关系与限制容差关系进行对比和综合, 可得如下定理:

定理 1 设不完备信息系统 $S = (U, A)$, M 是完备容差关系, L 是限制容差关系. 如果 $\forall x \in U (P(x) \neq \emptyset)$, 则对于 $\forall X \subseteq U$ 都有 $\underline{L}X \subseteq \underline{M}X, \overline{M}X \subseteq \overline{L}X$.

证明 对于 U 中任意两个对象 x 和 y , 如果

$$\frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)} \geq \alpha$$

则必有 $(x, y) \in M$, 即 $\forall x, y \in U (M(x, y) \Rightarrow L(x, y))$. 由定义 5 和定义 3 知, $M(x) = \{y \in U \mid (x, y) \in M\}, L(x) = \{y \in U \mid (x, y) \in L\}$, 因此 $\forall y \in U (y \in M(x) \Rightarrow y \in L(x))$; 反之则不成立. 所以 $M(x) \subseteq L(x)$. 由定义 6 和定义 3 易得, $\underline{M}X \subseteq \underline{L}X, \overline{M}X \subseteq \overline{L}X$.

因为 $M(x) \subseteq L(x)$, 所以对于 U 中任意对象 x , 如果 $L(x) \subseteq X$, 则 $M(x) \subseteq X$. 又因为 $x \in \underline{L}X \Leftrightarrow L(x) \subseteq X, x \in \underline{M}X \Leftrightarrow M(x) \subseteq X$, 所以 $\forall x \in U (x \in \underline{L}X \Rightarrow x \in \underline{M}X)$; 反之则不成立. 所以 $\underline{L}X \subseteq \underline{M}X$.

定理 2 设不完备信息系统 $S = (U, A)$, M 是完备容差关系, S 是相似关系. 如果 $\forall x \in U (P(x) \neq \emptyset)$, 则对于 $\forall X \subseteq U$ 都有 $\underline{M}X \subseteq \underline{S}X, \overline{S}X \subseteq \overline{M}X$.

证明 由定义 2 中相似关系的定义知, 如果 (x, y) 满足相似关系 S , 则 $\frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)} = \alpha$, 所以

$$\frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)} = \alpha$$

由定义 6 知, (x, y) 一定满足完备容差关系 M . 同理, 如果 (y, x) 满足相似关系 S , 则也一定满足完备容差关系 M . 即 $\forall x, y \in U (S(x, y) \Rightarrow M(x, y)), \forall x, y \in U (S(y, x) \Rightarrow M(x, y))$.

由定义 2 和定义 6 得, $S(x) = \{y \in U \mid (y, x) \in S\}, M(x) = \{y \in U \mid (x, y) \in M\}$. 而 $\forall x, y \in U (S(y, x) \Rightarrow M(x, y))$, 所以 $\forall y \in U (y \in S(x) \Rightarrow y \in M(x))$; 反之则不能成立. 所以 $S(x) \subseteq M(x)$. 同理可得 $S^{-1}(x) \subseteq M(x)$. 因为 $\overline{S}X = \overline{M}X, \underline{M}X \subseteq \underline{S}X$, 所以 $\underline{M}X \subseteq \underline{S}X, \overline{S}X \subseteq \overline{M}X$.

上面已证明 $S^{-1}(x) \subseteq M(x)$, 由此得出 $\forall x \in U (M(x) \subseteq X \Rightarrow S^{-1}(x) \subseteq X)$. 又因为 $x \in \underline{M}X$ 和 $x \in \underline{S}X$ 分别与 $M(x) \subseteq X$ 和 $S^{-1}(x) \subseteq X$ 等价, 所以 $\forall x \in U (x \in \underline{M}X \Rightarrow x \in \underline{S}X)$; 反之则不成立. 所

以 $MX \subseteq SX$.

由定理 1 和定理 2 及文献[11] 可得到如下结论：

定理 3 设不完备信息系统 $S = (U, A)$, M 是完备容差关系, T 是容差关系, S 是相似关系, L 是限制容差关系. 如果 $\forall_x \cup (P(x) \neq \emptyset)$, 则对于 $\forall X \subseteq U$, 都有 $TX \subseteq LX \subseteq MX \subseteq SX, SX \subseteq MX \subseteq LX \subseteq TX$.

由定理 3 知, 完备容差关系吸收了容差关系、相似关系和限制容差关系的优点, 克服了三者的缺陷. 完备容差关系划分粒度介于限制容差关系与相似关系之间, 并且充分考虑了信息系统完备度的差异, 因此其性能优于上述其他扩充模型.

上述定理在 $\forall_x \cup (P(x) \neq \emptyset)$ 下都成立, 即信息表中任一对象至少包含一个已知的属性值. 如果信息表中存在所有属性值都未知的对象 (如 $x = \{ *, *, *, * \}$), 则上述定理不再成立. 完备容差关系认为, 这样的 x 与其他任意的对象属于不同的完备容差类, 可以分辨. 这是符合逻辑的, 因为 x 所有属性值未知, 与其他对象相同的可能性很小.

4 应用举例

下面通过两个实际的不完备信息系统来分析比较这些扩充粗糙集模型的性能. 表 1 类似于文献[10] 使用的不完备信息表. 为了对比分析, 这里用该表来分析完备容差关系的性质.

表 1 不完备信息系统 1

S	a_1	a_2	a_3	a_4
x_1	3	2	1	0
x_2	2	3	2	0
x_3	2	3	2	0
x_4	*	2	*	1
x_5	*	2	*	1
x_6	2	3	2	1
x_7	3	*	*	3
x_8	*	0	0	*
x_9	3	2	1	3
x_{10}	1	*	*	*
x_{11}	*	2	*	*
x_{12}	3	2	1	*

例 1 设 $S = (U, A)$ 是不完备信息系统 1, 如表 1 所示. 其中: $U = \{ x_1, x_2, \dots, x_{12} \}$ 是对象集, $A = \{ a_1, a_2, a_3, a_4 \}$ 是条件属性集合, 值域为 $\{0, 1, 2, 3\}$. 记遗漏值为“ * ”.

利用容差关系、相似关系和限制容差关系分析此信息表的结果参见文献[11]. 根据定义 4 算出此信息系统的完备度 $= 0.6875$. 由定义 5 给出各个

对象的完备容差类如下：

$$\begin{aligned} M(x_1) &= \{ x_1, x_{11}, x_{12} \}, M(x_2) = \{ x_2, x_3 \}, \\ M(x_3) &= \{ x_2, x_3 \}, M(x_4) = \{ x_4, x_5, x_{11} \}, \\ M(x_5) &= \{ x_4, x_5, x_{11} \}, M(x_6) = \{ x_6 \}, \\ M(x_7) &= \{ x_7, x_9 \}, M(x_8) = \{ x_8 \}, \\ M(x_9) &= \{ x_7, x_9, x_{11}, x_{12} \}, M(x_{10}) = \{ x_{10} \}, \\ M(x_{11}) &= \{ x_1, x_4, x_5, x_9, x_{11}, x_{12} \}, \\ M(x_{12}) &= \{ x_1, x_9, x_{11}, x_{12} \}. \end{aligned}$$

对比利用限制容差关系分析此信息表的结果, 可发现 x_4, x_5, x_7, x_{12} 4 个对象的完备容差类和限制容差类不同. 即 $L(x_4) = \{ x_4, x_5, x_{11}, x_{12} \}, L(x_5) = \{ x_4, x_5, x_{11}, x_{12} \}, L(x_7) = \{ x_7, x_9, x_{12} \}, L(x_{12}) = \{ x_1, x_4, x_5, x_7, x_9, x_{11}, x_{12} \}$. 下面分析出现这种差异的原因.

以对象 x_4 为例, (x_4, x_{12}) 满足限制容差关系, 因此二者分在同一限制容差类, 认为是不可分辨的. 但在完备容差关系下

$$\frac{|P(x_4) \cap P(x_{12})|}{\min(|P(x_4)|, |P(x_{12})|)} = 0.5 < 0.6875,$$

因此 x_4 和 x_{12} 分在不同的完备容差类, 在这个不完备信息表中认为是可分辨的. 此信息表的完备度较大, 认为此信息表较为完备. 在缺席语意下理解不完备信息, 认为它们是不可得到的, 不可能与其他值相等, $\frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)}$ 的可能性较小, 所以划分的粒度相对较小. 同样是 $\{ *, 2, *, 1 \}$ 和 $\{ 3, 2, 1, * \}$, 在其他不完备信息表中则可能是不可分辨的.

例 2 设 $S = (U, A)$ 是不完备信息系统 2, 如表 2 所示. 其中: $U = \{ x_1, x_2, \dots, x_6 \}$ 是对象集, $A = \{ a_1, a_2, a_3, a_4 \}$ 是条件属性集合, 值域为 $\{0, 1, 2, 3\}$.

表 2 不完备信息系统 2

S	a_1	a_2	a_3	a_4
x_1	*	2	*	0
x_2	*	2	*	1
x_3	3	*	1	*
x_4	1	*	*	*
x_5	*	2	*	*
x_6	3	2	1	*

由定义 4 可算出此信息系统的完备度 $= 0.4583$. 根据定义 5 可得各个对象的完备容差类如下：

$$\begin{aligned} M(x_1) &= \{ x_1, x_5, x_6 \}, M(x_2) = \{ x_2, x_5, x_6 \}, \\ M(x_3) &= \{ x_3, x_6 \}, M(x_4) = \{ x_4 \}, \\ M(x_5) &= \{ x_1, x_2, x_5, x_6 \}, \\ M(x_6) &= \{ x_1, x_2, x_5, x_6 \}. \end{aligned}$$

由定义3得到各个对象的限制容差类如下:

$$L(x_1) = \{x_1, x_5, x_6\}, L(x_2) = \{x_2, x_5, x_6\},$$

$$L(x_3) = \{x_3, x_6\}, L(x_4) = \{x_4\},$$

$$L(x_5) = \{x_1, x_2, x_5, x_6\},$$

$$L(x_6) = \{x_1, x_2, x_5, x_6\}.$$

可以看出,完备容差关系和限制容差关系对对象集 U 的分类结果是相同的.

例1中 (x_4, x_{12}) 不满足完备容差关系,但本例中 $x_2 = \{*, 2, *, 1\}$ 和 $x_6 = \{3, 2, 1, *\}$ 与例1中 x_4 和 x_{12} 对应相等, (x_2, x_6) 满足完备容差关系. 因为

$$\frac{|P(x_2) \cap P(x_6)|}{\min(|P(x_2)|, |P(x_6)|)} = 0.5 > 0.4583,$$

所以二者属于同一完备容差类,在不完备信息系统2中认为是不可分辨的. 分析其原因,此信息表缺失值较多,因此由定义4算出的完备度较小,认为此信息表完备度较低. 在遗漏语意下理解表中的不完备信息,认为它们现在缺失仅是因为被遗漏但又确实存在,而且可能与任意值匹配, $\frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)}$ 的可能性较大,划分的粒度相对较大.

综合上面两个例子可以看出,限制容差关系在对不完备信息系统中的对象分类时,只考虑对象之间的关系. 如果两个对象的已知属性值都相同,则它们在任意不完备信息系统中的可分辨性是一致的. 完备容差关系分类时,不仅考虑对象之间的关系,也考虑不同不完备信息系统之间的差异. 如果两个对象的已知属性值都相同,则它们在信息系统完备度较低的情况下不可分辨的可能性较大,但在完备度较高的信息系统中可分辨的可能性较大,这样更符合实际情况.

5 结 语

在处理不完备信息系统时,一般需要对经典粗糙集理论进行必要的扩充,目前已有基于容差关系、相似关系、限制容差关系、修正容差关系等的扩充粗糙集模型. 本文分析比较了这些扩充模型的优缺点,给出了信息系统完备度的概念;在此基础上,提出一种基于完备容差关系的扩充粗糙集模型,该模型考虑了不同信息系统完备度之间的差异,既保留了限制容差关系扩充模型的优点,又克服了容差关系、相似关系和限制容差关系等扩充模型的局限性,对不完备信息系统的处理更符合现实情况. 在完备容差关系下,如何进行信息系统的属性约简,将是下一步的工作任务.

参考文献(References)

- [1] Pawlak Z. Rough sets [J]. Int J of computer and information sciences, 1982, 11(5): 341-356.
- [2] Pawlak Z, Skowron A. Rudiments of rough sets [J]. Information Sciences, 2007, 177(1): 3-27.
- [3] Pawlak Z, Skowron A. Rough sets: Some extensions [J]. Information Sciences, 2007, 177(1): 28-40.
- [4] Pawlak Z, Skowron A. Rough sets and boolean reasoning [J]. Information Sciences, 2007, 177(1): 41-73.
- [5] Gora G, Wojna A. RIONA: A new classification system combining rule induction and instance-based learning [J]. Fundamenta Informaticae, 2002, 51(4): 369-390.
- [6] Greco S, Inuiguchi M, Slowinski R. Fuzzy rough sets and multiple-premise gradual decision rules [J]. Int J of Approximate Reasoning, 2006, 41(2): 179-211.
- [7] Grzymala Busse J W, Ziarko W. Data mining and rough set theory [J]. Communications of the ACM, 2000, 43(1): 108-109.
- [8] Hu X, Cercone N. Data mining via discretization, generalization and rough set feature selection [J]. Knowledge and Information Systems, 1999, 1(1): 33-60.
- [9] Kryszkiewicz M. Rough set approach to incomplete information system [J]. Information Sciences, 1998, 11(2): 39-49.
- [10] Stefanowski J, Tsoukias A. On the extension of rough sets under incomplete information [C]. Proc of the 7th Int Workshop on New Directions in Rough Sets, Data Mining and Granular Soft Computing. Berlin: Springer-Verlag, 1999: 73-81.
- [11] 王国胤. 粗糙集理论在不完备信息系统中的扩充 [J]. 计算机研究与发展, 2002, 39(10): 1238-1243. (Wang Guo-yin. Extension of rough set under incomplete information systems [J]. J of Computer Research and Development, 2002, 39(10): 1238-1243.)
- [12] 刘富春. 基于修正容差关系的扩充粗糙集模型 [J]. 计算机工程, 2005, 31(24): 145-147. (Liu Fu-chun. Extended rough set model based on modified tolerance relation [J]. Computer Engineering, 2005, 31(24): 145-147.)
- [13] 王国胤. 粗糙集理论与知识获取 [M]. 西安: 西安交通大学出版社, 2001. (Wang Guo-yin. The theory of rough set and knowledge acquisition [M]. Xi'an: Xi'an Jiaotong University Press, 2001.)