

基于容差关系的不完备可变精度多粒度粗糙集^{*}

许 伟¹, 吴 陈¹, 杨习贝^{1,2}

(1. 江苏科技大学 计算机科学与工程学院, 江苏 镇江 212003; 2. 南京理工大学 计算机科学与技术学院, 南京 210094)

摘 要: 在不完备信息系统中, 为了融合可变精度粗糙集和多粒度粗糙集的各自优点, 提出一种基于容差关系的不完备可变精度多粒度粗糙集模型。研究了基于容差关系的可变精度乐观多粒度粗糙集和悲观多粒度粗糙集的相关性质。通过对可变精度多粒度粗糙集和经典多粒度粗糙集的对比分析, 结果表明, 基于容差关系的不完备可变精度多粒度粗糙集拥有更高的近似精度, 实例分析的结果也验证了该理论的可行性。

关键词: 不完备信息系统; 容差关系; 多粒度粗糙集; 可变精度多粒度粗糙集

中图分类号: TP18

文献标志码: A

文章编号: 1001-3695(2013)06-1712-04

doi:10.3969/j.issn.1001-3695.2013.06.028

Incomplete variable precision multi-granulation rough sets based on tolerance relation

XU Wei¹, WU Chen¹, YANG Xi-bei^{1,2}

(1. School of Computer Science & Technology, Jiangsu University of Science & Technology, Zhenjiang Jiangsu 212003, China; 2. School of Computer Science & Technology, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract: To integrate the good points of variable precision rough set and multi-granulations rough set in the incomplete information system, this paper proposed the incomplete variable precision multi-granulation rough sets based on tolerance relation, and discussed the properties of variable precision multi-granulation optimistic approximation and variable precision multi-granulation pessimistic approximation. Through the analysis of variable precision multi-granulations rough set and the classical multi-granulation rough set, it concluded that incomplete variable precision multi-granulation rough sets based on tolerance relation have a higher approximation. At last, results of the experiment show the effectiveness of the theory.

Key words: incomplete information system (IIS); tolerance relation; multi-granulation rough set; variable precision multi-granulations rough set

在 Pawlak 经典粗糙集^[1]的研究中, 大多数是针对完备信息系统而言, 然而在实际问题中由于记录丢失、获取代价、获取手段限制等原因, 通常都是不完备信息系统。Kryszkiewicz^[2]提出的容差关系就是为了处理不完备信息系统而定义的一种二元关系。目前, 在容差关系的基础上, 许多学者引入了变精度思想, 文献[3]提出了可变精度粗糙集模型, 在这个模型中, 给定一个阈值, 当对象所在的等价类在某种程度上包含于集合中时, 就认为这个对象属于该集合, 即在一个给定错误率条件下将尽可能多的对象划分为一类。这一推广对克服信息系统中噪声数据对知识获取的影响有着重要意义。

粒计算^[4-6]是一门融合了粗糙集、模糊集及人工智能等多种理论的新学科。经典粗糙集和目前很多拓展的粗糙集模型都是建立在单个不可分辨的二元关系上的。文献[7~10]从多个粒结构角度出发, 提出了多粒度粗糙集概念, 给出了乐观多粒度和悲观多粒度粗糙集模型。将经典粗糙集从单个粒结构推广到多个粒结构中, 并证明了 Pawlak 经典粗糙集模型是多粒度粗糙集的特例。本文在不完备信息系统中基于容差关系的多粒度粗糙集模型, 结合可变精度粗糙集和多粒度粗糙集的各自优点, 提出基于容差关系下的可变精度多粒度粗糙集。

1 基本概念

1.1 可变精度粗糙集

Pawlak 经典粗糙集的一个局限性是它所处理的分类必须是完全正确的或肯定的, 因而没有解决某种程度上的“包含”与“属于”。为克服经典粗糙集的这一局限性, Ziarko^[3]于1993年提出了可变精度粗糙集模型。

定义1 设 U 为一个有限的非空论域, $\forall X, Y \subseteq U$, 令

$$e(X, Y) = \begin{cases} 1 - \frac{|X \cap Y|}{|X|} & |X| \neq 0 \\ 0 & |X| = 0 \end{cases} \quad (1)$$

其中: $|X|$ 表示集合 X 的基数, $e(X, Y)$ 表示集合 X 对于集合 Y 的相对错误分类率。令 $0 \leq \beta < 0.5$, 若 $e(X, Y) \leq \beta$, 则称 X 以误差 β 多数包含于 Y , 记为 $X \subseteq_{\beta} Y$ 。

定义2 设信息系统 $IS = \langle U, AT \rangle$ 。对于 $X \subseteq U, A \subseteq AT$, $0 \leq \beta < 0.5$, 则 X 关于 A 的 β 下近似和 β 上近似为

$$\underline{A}_{\beta}(X) = \{x \in U; e([x]_A, X) \leq \beta\} \quad (2)$$

$$\overline{A}_{\beta}(X) = \{x \in U; e([x]_A, X) < 1 - \beta\} \quad (3)$$

由 $[\underline{A}_{\beta}(X), \overline{A}_{\beta}(X)]$ 得到的粗糙集模型称为可变精度粗糙集模型。

收稿日期: 2012-09-27; **修回日期:** 2012-10-30 **基金项目:** 国家自然科学基金重点资助项目(61100116); 江苏省自然科学基金重点资助项目(BK2011492); 江苏科技大学研究生创新计划资助项目(CXLX12_0693)

作者简介: 许伟(1987-), 女, 硕士, 主要研究方向为粒计算、数据挖掘(xuweij1987@sina.com); 吴陈(1961-), 男, 教授, 硕士, 主要研究方向为模式识别、智能信息处理等; 杨习贝(1980-), 男, 硕士, 博士(后), 主要研究方向为粒计算、粗糙集等。

1.2 不完备信息系统中的容差关系

在容差关系模型中,最重要的一个思想是将信息表中缺失的属性值都赋予“*”值,*可能是任意值。因此,对于一个决策信息系统 $IS = \langle U, AT \cup \{d\} \rangle$,其中 AT 为条件属性, $\{d\}$ 为决策属性,在未知属性值都是遗漏的情况下, Kryszkiewicz^[2] 提出了容差关系的定义。

定义3 给定不完备决策系统 $IS = \langle U, AT \cup \{D\} \rangle$,对于具有空值的属性子集 $A \subseteq AT$,记空值为“*”,定义 A 上的容差关系 $T(A)$ 为

$$T(A) = \{(x, y) \in U^2 : \forall a \in A \\ a(x) = a(y) \vee a(x) = * \vee a(y) = *\} \quad (4)$$

定义4 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$, $A \subseteq AT, x \subseteq U$,在容差关系 $T(A)$ 下, X 的下、上近似集分别定义为

$$\underline{A}_T(X) = \{x \in U : T_A(x) \subseteq X\} \quad (5)$$

$$\overline{A}_T(X) = \{x \in U : T_A(x) \cap X \neq \emptyset\} \quad (6)$$

其中: $T_A(x) = \{y \in U : (x, y) \in T(A)\}$ 表示对象 x 相对于集合 A 的容差类。

1.3 基于容差关系的多粒度粗糙集模型

基于容差关系的不完备信息系统中,我国学者钱宇华等人已经提出了多粒度粗糙集模型。多粒度粗糙集是从粒计算的角度出发,将单个不可分辨二元关系导出单个粒结构,扩展到由一组二元关系导出的多个粒结构,采用多个粒结构中的知识来近似表示概念。

定义5^[8] 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集,对 $\forall X \subseteq U$,在容差关系族 $T(A_1), T(A_2), \dots, T(A_m)$ 的乐观多粒度粗糙集的下、上近似分别为

$$\underline{\sum_{i=1}^m A_{i\beta}^o}(X) = \{x \in U : \\ T_{A_1}(x) \subseteq X \vee T_{A_2}(x) \subseteq X \vee \dots \vee T_{A_m}(x) \subseteq X\} \quad (7)$$

$$\overline{\sum_{i=1}^m A_{i\beta}^o}(X) = \sim \sim \sum_{i=1}^m A_{i\beta}^o(\sim X) \quad (8)$$

定义6^[10] 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集,对 $\forall X \subseteq U$,在容差关系族 $T(A_1), T(A_2), \dots, T(A_m)$ 的悲观多粒度粗糙集的下、上近似分别为

$$\underline{\sum_{i=1}^m A_{i\beta}^p}(X) = \{x \in U : \\ T_{A_1}(x) \subseteq X \wedge T_{A_2}(x) \subseteq X \wedge \dots \wedge T_{A_m}(x) \subseteq X\} \quad (9)$$

$$\overline{\sum_{i=1}^m A_{i\beta}^p}(X) = \sim \sim \sum_{i=1}^m A_{i\beta}^p(\sim X) \quad (10)$$

2 基于容差关系的可变精度多粒度粗糙集

2.1 基于容差关系的可变精度多粒度粗糙集定义

定义7 设不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$, $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集,对 $\forall X \subseteq U$,在容差关系族 $T(A_1), T(A_2), \dots, T(A_m)$ 的可变精度乐观多粒度下、上近似分别定义为

$$\underline{\sum_{i=1}^m A_{i\beta}^o}(X) = \{x \in U : \\ e(T_{A_1}(x), X) \leq \beta \vee \dots \vee e(T_{A_m}(x), X) \leq \beta\} \quad (11)$$

$$\overline{\sum_{i=1}^m A_{i\beta}^o}(X) = \sim \sim \sum_{i=1}^m A_{i\beta}^o(\sim X) \quad (12)$$

通过 X 的 β 可变精度多粒度粗糙集的下、上近似集,还可以得到其边界域集合:

$$Bn_{\sum_{i=1}^m A_{i\beta}}^o(X) = \underline{\sum_{i=1}^m A_{i\beta}^o}(X) - \underline{\sum_{i=1}^m A_{i\beta}^o}(X) \quad (13)$$

定理1 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, $\forall X \subseteq U$,可知:

$$\underline{\sum_{i=1}^m A_{i\beta}^o}(X) = \{x \in U :$$

$$e(T_{A_1}(x), X) < 1 - \beta \wedge \dots \wedge e(T_{A_m}(x), X) < 1 - \beta\} \quad (14)$$

证明 由定义7知:

$$\begin{aligned} x \in \underline{\sum_{i=1}^m A_{i\beta}^o}(X) &\Leftrightarrow x \notin \overline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) \Leftrightarrow \\ e(T_{A_1}(x), \sim X) &> \beta \wedge \dots \wedge e(T_{A_m}(x), \sim X) > \beta \Leftrightarrow \\ 1 - \frac{|T_{A_1}(X) \cap (\sim X)|}{|T_{A_1}(X)|} &> \beta \wedge \dots \wedge 1 - \frac{|T_{A_m}(X) \cap (\sim X)|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ 1 - \frac{|T_{A_1}(X) - T_{A_1}(X) \cap X|}{|T_{A_1}(X)|} &> \beta \wedge \dots \wedge 1 - \frac{|T_{A_m}(X) - T_{A_m}(X) \cap X|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ 1 - 1 + \frac{|T_{A_1}(X) \cap X|}{|T_{A_1}(X)|} &> \beta \wedge \dots \wedge 1 - 1 + \frac{|T_{A_m}(X) \cap X|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ e(T_{A_1}(x), X) &< 1 - \beta \wedge \dots \wedge e(T_{A_m}(x), X) < 1 - \beta \end{aligned}$$

定义8 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集,对 $\forall X \subseteq U$,在容差关系族 $T(A_1), T(A_2), \dots, T(A_m)$ 的可变精度悲观多粒度下、上近似分别为

$$\underline{\sum_{i=1}^m A_{i\beta}^p}(X) = \{x \in U : \\ e(T_{A_1}(x), X) \leq \beta \wedge \dots \wedge e(T_{A_m}(x), X) \leq \beta\} \quad (15)$$

$$\overline{\sum_{i=1}^m A_{i\beta}^p}(X) = \sim \sim \sum_{i=1}^m A_{i\beta}^p(\sim X) \quad (16)$$

通过 X 的 β 可变精度多粒度粗糙集的下、上近似集,还可以得到其边界域集合:

$$Bn_{\sum_{i=1}^m A_{i\beta}}^p(X) = \underline{\sum_{i=1}^m A_{i\beta}^p}(X) - \underline{\sum_{i=1}^m A_{i\beta}^p}(X) \quad (17)$$

定理2 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$,令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, $\forall X \subseteq U$,可得

$$\underline{\sum_{i=1}^m A_{i\beta}^p}(X) = \{x \in U : \\ e(T_{A_1}(x), X) < 1 - \beta \vee \dots \vee e(T_{A_m}(x), X) < 1 - \beta\} \quad (18)$$

证明 由定义8知:

$$\begin{aligned} x \in \underline{\sum_{i=1}^m A_{i\beta}^p}(X) &\Leftrightarrow x \notin \overline{\sum_{i=1}^m A_{i\beta}^p}(\sim X) \Leftrightarrow \\ e(T_{A_1}(x), \sim X) &> \beta \vee \dots \vee e(T_{A_m}(x), \sim X) > \beta \Leftrightarrow \\ 1 - \frac{|T_{A_1}(X) \cap (\sim X)|}{|T_{A_1}(X)|} &> \beta \vee \dots \vee 1 - \frac{|T_{A_m}(X) \cap (\sim X)|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ 1 - \frac{|T_{A_1}(X) - T_{A_1}(X) \cap X|}{|T_{A_1}(X)|} &> \beta \vee \dots \vee 1 - \frac{|T_{A_m}(X) - T_{A_m}(X) \cap X|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ 1 - 1 + \frac{|T_{A_1}(X) \cap X|}{|T_{A_1}(X)|} &> \beta \vee \dots \vee 1 - 1 + \frac{|T_{A_m}(X) \cap X|}{|T_{A_m}(X)|} > \beta \Leftrightarrow \\ e(T_{A_1}(x), X) &< 1 - \beta \vee \dots \vee e(T_{A_m}(x), X) < 1 - \beta \end{aligned}$$

2.2 基于容差关系的可变精度多粒度粗糙集的性质

定理3 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集,对于 $\forall X, Y \subseteq U$,基于容差关系的可变精度乐观多粒度粗糙集有以下性质:

a) $\underline{\sum_{i=1}^m A_{i\beta}^o}(U) = \underline{\sum_{i=1}^m A_{i\beta}^o}(U) = U, \underline{\sum_{i=1}^m A_{i\beta}^o}(\emptyset) = \underline{\sum_{i=1}^m A_{i\beta}^o}(\emptyset) = \emptyset$;

b) $\underline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) = \sim \underline{\sum_{i=1}^m A_{i\beta}^o}(X), \underline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) = \sim \underline{\sum_{i=1}^m A_{i\beta}^o}(X)$;

c) $X \subseteq Y \Rightarrow \underline{\sum_{i=1}^m A_{i\beta}^o}(X) \subseteq \underline{\sum_{i=1}^m A_{i\beta}^o}(Y) \Rightarrow \underline{\sum_{i=1}^m A_{i\beta}^o}(X) \subseteq \underline{\sum_{i=1}^m A_{i\beta}^o}(Y)$;

d) $\beta_1 \geq \beta_2 \Rightarrow \underline{\sum_{i=1}^m A_{i\beta_1}^o}(X) \supseteq \underline{\sum_{i=1}^m A_{i\beta_2}^o}(X) \Rightarrow \underline{\sum_{i=1}^m A_{i\beta_1}^o}(X) \subseteq \underline{\sum_{i=1}^m A_{i\beta_2}^o}(X)$ 。

证明 a) 由定义7易证。b) 由定义7可知, $\underline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) = \sim \underline{\sum_{i=1}^m A_{i\beta}^o}(X)$, 令 $X = \sim X$, 则有 $\sim \underline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) = \underline{\sum_{i=1}^m A_{i\beta}^o}(X)$, 即 $\underline{\sum_{i=1}^m A_{i\beta}^o}(\sim X) = \sim \underline{\sum_{i=1}^m A_{i\beta}^o}(X)$ 。c) 设 $x \in \underline{\sum_{i=1}^m A_{i\beta}^o}(X)$, 由定义7可知, 必定 $\exists A_i \in \{A_1, A_2, \dots, A_m\}$ 使

$e(T_{A_i}(x), X) \leq \beta$, 又因为 $X \subseteq Y$, 所以 $e(T_{A_i}(x), X) \subseteq e(T_{A_i}(x), Y) \leq \beta$, 因此 $x \in \sum_{i=1}^m A_{i\beta}^o(Y)$, 所以有 $\sum_{i=1}^m A_{i\beta}^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(Y)$ 。类似地, 不难证明 $\sum_{i=1}^m A_{i\beta}^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(Y)$ 。d) 对于 $\forall x \in \sum_{i=1}^m A_{i\beta_2}^o(X)$, 由定义 7 可知, 必定 $\exists A_i \in \{A_1, A_2, \dots, A_m\}$ 使得 $e(T_{A_i}(x), X) \leq \beta_2$, 又因为 $\beta_1 \geq \beta_2$, 所以 $e(T_{A_i}(x), X) \leq \beta_1$, 即 $x \in \sum_{i=1}^m A_{i\beta_1}^o(X)$, 从而可证 $\sum_{i=1}^m A_{i\beta_1}^o(X) \supseteq \sum_{i=1}^m A_{i\beta_2}^o(X)$ 。类似地, 不难证明 $\sum_{i=1}^m A_{i\beta_1}^o(X) \subseteq \sum_{i=1}^m A_{i\beta_2}^o(X)$ 。

定理 4 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, 且 $A_1 \cup A_2 \cup \dots \cup A_m \neq \emptyset$, 对于 $\forall X \subseteq U$, 则有以下性质:

$$\bigcup_{i=1}^m A_{i\beta}^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (19)$$

$$\bigcup_{i=1}^m A_{i\beta}^o(X) \supseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (20)$$

证明 先证明式(20), 对于 $\forall x \in \sum_{i=1}^m A_{i\beta}^o(X)$, 由定理 3 知 $e(T_{A_i}(x), X) < 1 - \beta \wedge \dots \wedge e(T_{A_m}(x), X) < 1 - \beta$, 因此 $x \in \cap_{i=1}^m e(T_{A_i}(x), X) < 1 - \beta$, 又因为 $\cap_{i=1}^m e(T_{A_i}(x), X) \in e(T_{\cup_{i=1}^m A_i}(x), X)$, $\forall x \in U$, 所以有 $x \in e(T_{\cup_{i=1}^m A_i}(x), X) < 1 - \beta$, 即 $x \in \bigcup_{i=1}^m A_{i\beta}^o(X)$, 所以 $\bigcup_{i=1}^m A_{i\beta}^o(X) \supseteq \sum_{i=1}^m A_{i\beta}^o(X)$ 。再证明式(19), 根据相容关系的可变精度乐观多粒度粗糙集模型, 知道 $\bigcup_{i=1}^m A_{i\beta}^o(X) = \sim \bigcup_{i=1}^m A_{i\beta}^o(\sim X)$ 。由式(20)的结果知 $\sim \bigcup_{i=1}^m A_{i\beta}^o(\sim X) \subseteq \sim \sum_{i=1}^m A_{i\beta}^o(\sim X) = \sum_{i=1}^m A_{i\beta}^o(X)$, 即 $\bigcup_{i=1}^m A_{i\beta}^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X)$ 。

定理 5 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, 对于 $\forall X, Y \subseteq U$, 基于容差关系的可变精度悲观多粒度粗糙集有以下性质:

a) $\sum_{i=1}^m A_{i\beta}^p(U) = \sum_{i=1}^m A_{i\beta}^p(U) = U$, $\sum_{i=1}^m A_{i\beta}^p(\emptyset) = \sum_{i=1}^m A_{i\beta}^p(\emptyset) = \emptyset$;

b) $\sum_{i=1}^m A_{i\beta}^p(\sim X) = \sim \sum_{i=1}^m A_{i\beta}^p(X)$, $\sum_{i=1}^m A_{i\beta}^p(\sim X) = \sim \sum_{i=1}^m A_{i\beta}^p(X)$;

c) $X \subseteq Y \Rightarrow \sum_{i=1}^m A_{i\beta}^p(X) \subseteq \sum_{i=1}^m A_{i\beta}^p(Y) \Rightarrow \sum_{i=1}^m A_{i\beta}^p(X) \subseteq \sum_{i=1}^m A_{i\beta}^p(Y)$;

c) $\beta_1 \geq \beta_2 \Rightarrow \sum_{i=1}^m A_{i\beta_1}^p(X) \supseteq \sum_{i=1}^m A_{i\beta_2}^p(X) \Rightarrow \sum_{i=1}^m A_{i\beta_1}^p(X) \subseteq \sum_{i=1}^m A_{i\beta_2}^p(X)$ 。

证明 与定理 3 的证明过程类似。

定理 6 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, 且 $A_1 \cup A_2 \cup \dots \cup A_m \neq \emptyset$, 对于 $\forall X \subseteq U$, 则有以下性质:

$$\bigcup_{i=1}^m A_{i\beta}^p(X) \supseteq \sum_{i=1}^m A_{i\beta}^p(X), \bigcup_{i=1}^m A_{i\beta}^p(X) \subseteq \sum_{i=1}^m A_{i\beta}^p(X)$$

证明 与定理 4 的证明过程类似。

2.3 基于相容关系的可变精度多粒度粗糙集之间的关系

定理 7 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, 对于 $\forall X \subseteq U$, 有

$$\sum_{i=1}^m A_{i\beta}^p(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (21)$$

$$\sum_{i=1}^m A_{i\beta}^p(X) \supseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (22)$$

证明 $\forall x \in \sum_{i=1}^m A_{i\beta}^p(X)$, 由定义 8 可知, 必定 $\exists A_k \in \{A_1, A_2, \dots, A_m\}$ 使得 $e(T_{A_k}(x), X) \leq \beta$, 再由定义 7 可知 $\forall x \in \sum_{i=1}^m A_{i\beta}^o(X)$ 显然成立, 从而可证 $\sum_{i=1}^m A_{i\beta}^p(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X)$ 。

类似地, 根据定理 1、2 不难证明 $\sum_{i=1}^m A_{i\beta}^p(X) \supseteq \sum_{i=1}^m A_{i\beta}^o(X)$ 。定理 7 给出了基于容差关系的可变精度乐观多粒度粗糙

集与悲观多粒度粗糙集之间的关系。可以看出, 基于容差关系的可变精度乐观多粒度粗糙集拥有比基于基于容差关系的可变精度悲观多粒度粗糙集更大的下近似集和更小的上近似集。

定理 8 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, 对于 $\forall X \subseteq U$, 有

$$\sum_{i=1}^m A_i^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (23)$$

$$\sum_{i=1}^m A_i^p(X) \supseteq \sum_{i=1}^m A_{i\beta}^p(X) \quad (24)$$

$$\sum_{i=1}^m A_i^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^p(X) \quad (25)$$

$$\sum_{i=1}^m A_i^p(X) \supseteq \sum_{i=1}^m A_{i\beta}^o(X) \quad (26)$$

证明 $\forall x \in \sum_{i=1}^m A_i^o(X)$, 由定义 5 可知, 必定 $\exists A_i \in \{A_1, A_2, \dots, A_m\}$ 使得 $T_{A_i}(x) \subseteq X$, 由定义 1 的相对错误分类率可知, $e(T_{A_i}(x), X) = 0 \leq \beta$, 再根据定义 7 可知, $\forall x \in \sum_{i=1}^m A_{i\beta}^o(X)$ 成立, 从而可证 $\sum_{i=1}^m A_i^o(X) \subseteq \sum_{i=1}^m A_{i\beta}^o(X)$ 。

类似地, 根据上述论证不难证明式(24)~(26)。

定理 8 给出了基于容差关系的可变精度多粒度粗糙集与文献的传统多粒度粗糙集之间的关系。可以看出, 采用错误分类率方法的可变精度多粒度粗糙集比传统的多粒度粗糙集拥有更大的下近似集和更小的上近似集。

3 多粒度粗糙集的近似精度

粗糙集的不确定性是由于边界的存在而引起的, 集合边界域越大则不确定性也越大, 确定性就越小, 从而精确性也就越低, 粗糙程度就越大。为了更好地描述这一点, 下面从粗糙集的近似精度来度量不完备信息系统的多粒度粗糙集。

定义 9 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, $A = \{A_1, A_2, \dots, A_m\}$, $0 \leq \beta < 0.5$, $\forall X \subseteq U$ (其中 $X \neq \emptyset$), 则 X 关于 A 的乐观多粒度、悲观多粒度、可变精度乐观多粒度和可变精度悲观多粒度粗糙集的近似精度分别定义为 $\alpha_o, \alpha_p, \alpha_o^\beta$ 和 α_p^β 。

$$\alpha_o(\sum_{i=1}^m A_i, X) = \frac{|\sum_{i=1}^m A_i^o(X)|}{|\sum_{i=1}^m A_i^o(X)|} \quad (27)$$

$$\alpha_p(\sum_{i=1}^m A_i, X) = \frac{|\sum_{i=1}^m A_i^p(X)|}{|\sum_{i=1}^m A_i^p(X)|}$$

$$\alpha_o^\beta(\sum_{i=1}^m A_i, X) = \frac{|\sum_{i=1}^m A_{i\beta}^o(X)|}{|\sum_{i=1}^m A_{i\beta}^o(X)|}$$

$$\alpha_p^\beta(\sum_{i=1}^m A_i, X) = \frac{|\sum_{i=1}^m A_{i\beta}^p(X)|}{|\sum_{i=1}^m A_{i\beta}^p(X)|} \quad (28)$$

定理 9 不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 。令 $A_1, A_2, \dots, A_m$ 是 AT 的 m 个属性子集, $A = \{A_1, A_2, \dots, A_m\}$, $\forall X \subseteq U$ (其中 $X \neq \emptyset$), 则各近似精度之间有如下关系成立:

$$\alpha_o(\sum_{i=1}^m A_i, X) \leq \alpha_o^\beta(\sum_{i=1}^m A_i, X) \quad (29)$$

$$\alpha_p(\sum_{i=1}^m A_i, X) \leq \alpha_p^\beta(\sum_{i=1}^m A_i, X) \quad (30)$$

根据定理 8 知 $|\sum_{i=1}^m A_i^o(X)| \leq |\sum_{i=1}^m A_{i\beta}^o(X)|$ 且 $|\sum_{i=1}^m A_i^o(X)| \geq |\sum_{i=1}^m A_{i\beta}^o(X)|$ 。因此, $\alpha_o(\sum_{i=1}^m A_i, X) = \frac{|\sum_{i=1}^m A_i^o(X)|}{|\sum_{i=1}^m A_i^o(X)|} \leq \frac{|\sum_{i=1}^m A_{i\beta}^o(X)|}{|\sum_{i=1}^m A_{i\beta}^o(X)|} = \alpha_o^\beta(\sum_{i=1}^m A_i, X)$ 。

同理不难证明 $\alpha_p(\sum_{i=1}^m A_i, X) \leq \alpha_p^\beta(\sum_{i=1}^m A_i, X)$ 。

定理 9 说明了基于容差关系的不完备可变精度乐观和悲观多粒度粗糙集比传统的乐观和悲观多粒度粗糙集拥有更高

的近似精度。

4 实例分析

不完备决策系统 $IS = \langle U, AT \cup \{d\} \rangle$ 如表 1 所示,其中对象集 $U = \{x_1, x_2, \dots, x_{10}\}$, 条件属性集 $AT = \{c_1, c_2, c_3, c_4, c_5\}$, 取 $\beta = 0.35, d$ 为决策属性。

表 1 不完备决策系统

U	c_1	c_2	c_3	c_4	c_5	d	U	c_1	c_2	c_3	c_4	c_5	d
x_1	1	2	2	1	1	yes	x_6	*	1	0	1	2	yes
x_2	0	2	2	1	1	yes	x_7	1	1	1	1	1	no
x_3	0	0	0	0	0	no	x_8	2	1	1	2	2	yes
x_4	0	0	0	0	0	no	x_9	1	2	1	2	0	yes
x_5	0	0	0	0	0	no	x_{10}	0	0	*	0	*	no

由表 1 得 $U/\text{IND}(\{d\}) = \{D_1, D_2\}$, 其中 $D_1 = \{x_1, x_2, x_6, x_8, x_9\}$, $D_2 = \{x_3, x_4, x_5, x_7, x_{10}\}$ 。根据定义 5 和 6 得

$$\frac{\sum_{i=1}^5 A_i^o(D_1)}{\sum_{i=1}^5 A_i^o(D_2)} = \{x_1, x_2, x_8, x_9\}$$

$$\frac{\sum_{i=1}^5 A_i^o(D_2)}{\sum_{i=1}^5 A_i^o(D_1)} = \{x_3, x_4, x_5, x_{10}\}$$

$$\frac{\sum_{i=1}^5 A_i^p(D_1)}{\sum_{i=1}^5 A_i^p(D_2)} = \{x_1, x_2, x_6, x_7, x_8, x_9\}$$

$$\frac{\sum_{i=1}^5 A_i^p(D_2)}{\sum_{i=1}^5 A_i^p(D_1)} = \{x_3, x_4, x_5, x_{10}\}, \frac{\sum_{i=1}^5 A_i^p(D_1)}{\sum_{i=1}^5 A_i^p(D_2)} = \emptyset,$$

$$\frac{\sum_{i=1}^5 A_i^p(D_2)}{\sum_{i=1}^5 A_i^p(D_1)} = \emptyset, \frac{\sum_{i=1}^5 A_i^p(D_1)}{\sum_{i=1}^5 A_i^p(D_2)} = U, \frac{\sum_{i=1}^5 A_i^p(D_2)}{\sum_{i=1}^5 A_i^p(D_1)} = U$$

根据定义 7 和 8 得

$$\frac{\sum_{i=1}^5 A_{i\beta}^o(D_1)}{\sum_{i=1}^5 A_{i\beta}^o(D_2)} = \{x_1, x_2, x_6, x_7, x_8, x_9\}$$

$$\frac{\sum_{i=1}^5 A_{i\beta}^o(D_2)}{\sum_{i=1}^5 A_{i\beta}^o(D_1)} = \{x_2, x_3, x_4, x_5, x_6, x_9, x_{10}\}$$

$$\frac{\sum_{i=1}^5 A_{i\beta}^p(D_1)}{\sum_{i=1}^5 A_{i\beta}^p(D_2)} = \{x_1, x_2, x_6, x_7, x_8, x_9\}$$

$$\frac{\sum_{i=1}^5 A_{i\beta}^p(D_2)}{\sum_{i=1}^5 A_{i\beta}^p(D_1)} = \{x_3, x_4, x_5, x_{10}\}, \frac{\sum_{i=1}^5 A_{i\beta}^p(D_1)}{\sum_{i=1}^5 A_{i\beta}^p(D_2)} = \emptyset,$$

$$\frac{\sum_{i=1}^5 A_{i\beta}^p(D_2)}{\sum_{i=1}^5 A_{i\beta}^p(D_1)} = \{x_3, x_4, x_5\}, \frac{\sum_{i=1}^5 A_{i\beta}^p(D_1)}{\sum_{i=1}^5 A_{i\beta}^p(D_2)} = U, \frac{\sum_{i=1}^5 A_{i\beta}^p(D_2)}{\sum_{i=1}^5 A_{i\beta}^p(D_1)} = U$$

根据定理 9 得

$$\alpha_o(\sum_{i=1}^5 A_i, D_1) = \frac{2}{3}, \alpha_o(\sum_{i=1}^5 A_i, D_2) = 1$$

$$\alpha_p(\sum_{i=1}^5 A_i, D_1) = 0, \alpha_p(\sum_{i=1}^5 A_i, D_2) = 0$$

$$\alpha_o^\beta(\sum_{i=1}^5 A_i, D_1) = 1, \alpha_o^\beta(\sum_{i=1}^5 A_i, D_2) = \frac{7}{4}$$

(上接第 1677 页)规划;b)在对客户点进行服务优化时,着重考虑客户点之间的路径选择。对客户点的优化通过两方面进行:a)客户点顺序的优化;b)对客户点之间路线选择的优化。本文采用粒子群算法对客户点服务顺序进行规划,并采用动态规划方法对客户点之间的路线进行规划。实验表明,利用交通流变化进行路径规划可减少运输时间,提高货物配送的时效,减少运输成本。

参考文献:

- [1] DANTZIG G, RAMSER J. The truck dispatching problem[J]. *Management Science*, 1959, 6(1): 80-91.
- [2] BOWMAN E H. Production scheduling by the transportation method of linear programming[J]. *Operations Research*, 1956, 4(1): 100-103.
- [3] MALANDRAK I C, DASKIN M S. Time dependent vehicle routing problem: formulation, properties and heuristic algorithms[J]. *Transportation Science*, 1992, 26(3): 185-200.
- [4] ICHOUA S, GENDREAU M, POTVIN J Y. Vehicle dispatching with time dependent travel times[J]. *European Journal of Operations Research*, 2003, 144(2): 379-396.
- [5] HAGHANI A A, JUNG S. A dynamic vehicle routing problem with time-dependent travel times[J]. *Operations Research*, 2005, 32(11): 2959-2986.
- [6] DONATI A V, MONTEMANNI R, CASAGRANDE N, et al. Time dependent vehicle routing problem with a multi ant colony system[J]. *European Journal of Operational Research*, 2008, 185(3): 1174-1191.
- [7] LI Jing-quan, MIRCHANDANI P B, BORENDETEIN D. Real-time

$$\alpha_p^\beta(\sum_{i=1}^5 A_i, D_1) = 0, \alpha_p^\beta(\sum_{i=1}^5 A_i, D_2) = \frac{3}{10}$$

以上计算结果验证了定理 7~9 的正确性。

5 结束语

本文采用错误分类率方法将多粒度粗糙集拓展到容差关系的不完备系统中,提出了基于容差关系的可变精度多粒度粗糙集模型,包括可变精度乐观多粒度粗糙集和可变精度悲观多粒度粗糙集。因为采用了可变精度的方式,从而比传统的多粒度粗糙集拥有更大的下近似集和更小的上近似集,提高了近似精度。接下来的工作将在本文的基础上进一步研究可变精度多粒度粗糙集的约简和决策规则获取问题。

参考文献:

- [1] PAWLAK Z. Rough sets[J]. *International Journal of Computer and Information Sciences*, 1984, 11: 341-356.
- [2] KRYSZKIEWICZ M. Rough set approach to incomplete information systems[J]. *Information Sciences*, 1998, 112(1-4): 39-49.
- [3] ZIARKO W. Variable precision rough set model[J]. *Journal of Computer and System Sciences*, 1993, 46(1): 39-59.
- [4] 苗夺谦, 王国胤, 刘清, 等. 粒计算: 过去、未来和展望[M]. 北京: 科学出版社, 2007.
- [5] 王国胤, 张清华. 不同知识粒度下粗糙集的不确定性研究[J]. *计算机学报*, 2008, 31(9): 1588-1598.
- [6] 王国胤, 姚一豫, 于洪. 粗糙集理论与应用研究综述[J]. *计算机学报*, 2009, 32(7): 1229-1246.
- [7] YAG Y Y. Information granulation and rough set approximation[J]. *International Journal of Intelligent Systems*, 2001, 16(1): 87-104.
- [8] QIAN Yu-hua, LIANG Ji-ye, YAO Yi-yu, et al. MGRS: a multi-granulation rough set[J]. *Information Sciences*, 2010, 180(6): 949-970.
- [9] QIAN Yu-hua, LIANG Ji-ye, DANG Chuang-yin. Incomplete multi-granulation rough set[J]. *IEEE Trans on Systems, Man and Cybernetics, Part A*, 2010, 40(2): 420-431.
- [10] QIAN Yu-hua, LIANG Ji-ye, WEI Wei. Pessimistic rough decision [C]//Proc of the 2nd International Workshop on Rough Sets Theory. 2010: 440-449.

vehicle rerouting problems with time windows[J]. *European Journal of Operational Research*, 2009, 194(3): 711-727.

- [8] KRITZINGER S, DOERNER K F, HARTL R F, et al. Using traffic information for time-dependent vehicle routing: social and behavioral sciences[C]//Proc of the 7th International Conference on City Logistics. 2012: 217-229.
- [9] KOK A L, HANS E W, SCHUTTEN J M J. Vehicle routing under time-dependent travel times: the impact of congestion avoidance[J]. *Computers & Operations Research*, 2012, 39(5): 910-918.
- [10] 邹开俊, 鲁怀伟, 杜三山. 二进制粒子群优化算法在车辆路径问题中的应用[J]. *西北民族大学学报: 自然科学版*, 2010, 31(4): 12-15.
- [11] 张丽艳. 粒子群算法在车辆路径问题中的应用[D]. 上海: 上海交通大学, 2006.
- [12] 张文静. 协同粒子群算法及其在停车场路径优化问题中的应用[D]. 上海: 华东师范大学, 2010.
- [13] 王伟. 公路交通流车速——流量实用关系模型[J]. *东南大学学报*, 2003, 33(4): 486-490.
- [14] 曹翔宇. 城市路段行程时间计算研究[D]. 广州: 中山大学, 2008.
- [15] SALMAN A, AHMAD I, AL-MADANI S. Particle swarm optimization for task assignment problem[J]. *Microprocessors and Microsystem*, 2002, 26(8): 363-371.
- [16] BAISEIRO S R, LOISEAU I, RAMONET J. An ant colony algorithm hybridized with insertion heuristics for the time dependent vehicle routing problem with time windows[J]. *Computers & Operations Research*, 2011, 38(6): 954-966.