第47卷第1期 中国科学技术大学学报 Vol. 47, No. 1
2017年1月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Jan. 2017

文章编号:0253-2778(2017)1-0040-08

优势关系多粒度粗糙集中近似集动态更新方法

胡成祥,赵国柱

(滁州学院计算机与信息工程学院,安徽滁州 239000)

摘要:随着数据的不断变化,从信息系统中获取有用的信息,可有效地为决策提供依据.为此在多粒度环境下,优势关系多粒度粗糙集中粒度增加时,分析了优势关系乐观多粒度粗糙集和悲观多粒度粗糙集近似集动态更新的定理和相关性质,提出了一种优势关系多粒度粗糙集模型中,当粒度结构动态增加时,近似集更新的算法.该算法的基本思想是不需要重新计算粒度结构变化时信息系统的优势类、下近似集和上近似集,只需根据新增粒度结构的相关信息计算所有对象的优势类;然后根据优势关系乐观多粒度粗糙集和悲观多粒度粗糙集中动态更新近似集的相关定理计算近似集,提高了更新效率.通过与传统的静态算法做比较,验证了本算法的有效性.

关键词:优势关系;多粒度粗糙集;动态更新;近似集

中图分类号:TP 181 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2017.01.006

引用格式:胡成祥,赵国柱.优势关系多粒度粗糙集中近似集动态更新方法[J].中国科学技术大学学报,2017,47(1):40-47.

HU Chengxiang, ZHAO Guozhu. A dominance-based multigranulation rough sets approach for dynamic updating approximations[J]. Journal of University of Science and Technology of China, 2017, 47(1):40-47.

A dominance-based multigranulation rough sets approach for dynamic updating approximations

HU Chengxiang, ZHAO Guozhu

(*School of Computer and Information Engineering, Chuzhou University, Chuzhou* 239000, *China*)

Abstract: With the variation of the collected data, useful information obtained dynamically from the information system plays an important role in decision making. The properties of updating approximations in dominance-based optimistic and pessimistic multigranulation rough sets were discussed. An approach to dynamically updating approximations in dominance-based optimistic and pessimistic multigranulation rough sets while adding a granulation structure in multigranulation environment was presented. The approach does not need to recalculate the dominance classes and approximations of each granulation structure in the universe. The dominance classes of each object were calculated with respect to the added granulation structure, and then the approximations can be obtained by the properties of updating approximations in dominance-based optimistic and pessimistic multigranulation rough sets which can improve the efficiency of

收稿日期:2016-03-01;**修回日期**:2016-09-17

基金项目:安徽省教育厅自然科学基金重点项目(KJ2016A525),安徽省高校自然科学基金(KJ2015B19),滁州学院科研启动基金(2014qd018)资助

作者简介:胡成祥,男,1984年生,硕士/讲师.研究方向:智能信息处理.E-mail: chengxiang0550@163.com

通讯作者:赵国柱,硕士/实验师.E-mail: hhgzju@163.com

updating approximations. The experimental results demonstrate the validity of the proposed approach while comparing with the static algorithm.

Key words: dominance-based; multigranulation rough sets; dynamic updating; approximations

0 引言

粗糙集理论是由波兰学者 Pawlak 提出的一种研究模糊、不精确数据、不完整数据的表达、学习和归纳等的数学理论[1]，目前已经在模式识别、人工智能、知识获取、图像处理、预测分析等领域得到广泛应用[2-4]. Greco 等在完备信息系统中考虑了属性的顺序特征，采用优势关系下粗糙集处理论域中的分类问题，并根据所提出的方法获取粗糙近似集[5]，鲍忠奎等提出了一种优势关系下的直觉模糊粗糙集，给出了在直觉模糊信息系统中进行正域约简的方法[6].

Qian 等指出经典粗糙集及其扩展模型都是从单一的二元关系划分论域空间获取近似集，从粒度的角度出发，采用多粒度的观点，提出了针对完备信息系统的乐观多粒度粗糙集模型和悲观多粒度粗糙集模型[7]. 目前，多粒度粗糙集的研究是热点. Xu 等提出了容差关系下的多粒度粗糙集模型，给出了容差关系下乐观多粒度粗糙集和悲观多粒度粗糙集，提出了两种模型的近似集计算方法和相关性质[8]. Qian 等将完备系统中的多粒度粗糙集模型推广到不完备信息系统中，并提出了不完备系统中的容差关系多粒度粗糙集模型[9]. Huang 等提出了一种模糊环境下模糊多粒度粗糙集模型[10]. 庄颖等将多粒度优势关系粗糙集模型作为一种信息融合技术应用于目标威胁估计，使其适应于多源信息系统[11]. 翟永健等对基于扩展优势关系和限制优势关系的单粒度粗糙集模型进行了扩展，分别提出了不完备信息系统中的乐观和悲观多粒度粗糙集模型[12]. Yang 等提出了多粒度粗糙集中粒度结构增加时知识的获得方法和属性约简方法[13]. Ju 等提出了模糊多粒度粗糙集模型中粒度结构变化时基于粒度结构选择的属性约简方法和粗糙近似集的获取方法[14].

近似集是粗糙集理论中的重要概念，可以为决策规则获取、特征选择等提供依据. 张清华等利用条件属性对论域形成的知识粒直接构建不确定概念的近似集，相对于粗糙集中的不确定概念的下近似集或上近似集可能有更好的近似度[15]，张清华等在研究粗糙集近似集的基础上，利用条件属性对论域形成的划分空间，直接构建目标模糊集的近似集，提出了粗糙模糊集的近似表示[16]. 当信息系统发生变化时，如何动态地更新近似集对知识获取、规则提取等有重要意义. Liu 等采用一种新的矩阵方法用来描述集值信息系统的动态变化，给出了对象集变化时知识的动态更新算法[17]；张清华等针对信息系统中动态增加的数据，创建了一种知识粒树，采用动态方式有效地获取了信息系统的知识[18]；Chen 等调查了对象集粒度变化时近似集获取方法，设计了数据量增加时知识的动态获取方法[19]；Zhang 等针对邻域粗糙集模型对象集发生变化，给出了近似集动态更新增量方法，可以处理多个对象的增加和删除[20]；Li 等针对优势粗糙集模型中对象集的粗化和细化，给出了近似集动态更新的增量式学习方法，并设计了算法、验证了其有效性[21]；Chan 提出了属性增删时上、下近似和规则的更新方法[22]；Li 等调查了多属性增删时上、下近似的更新方法，将其应用到粗糙集扩展模型，进而获取知识[23-24]；Li 等还提出了一种在优势关系下属性集改变时知识获取算法[25]；Wang 等研究了当属性动态变化时，提出了一种多维增量式的属性约简算法[26]；胡成祥等根据属性变化前后等价类的变化，提出了属性增删时，粗糙集模型和变精度粗糙集中近似集的增量更新方法[27-28]；Chen 等提出了信息系统中关于属性值粗化和细化的定义，讨论了经典粗糙集模型中属性值粗化和细化时近似集的动态更新方法，分析了多粒度多层次情况下属性值粗化和细化的原理[29]；Chen 等提出了不完备信息系统中的限制容差关系变精度粗糙集模型，分析了属性值粗化、细化时粒度变化情况，并给出在此模型中属性值变化时近似集的动态更新算法[30]；Luo 等提出了集值决策信息系统中属性值粗化、细化的概念，讨论了属性值随时间变化时上、下近似的获取算法[31].

1 优势关系多粒度粗糙集相关知识

1.1 优势关系粗糙集

定义 1.1[5]　令 $S=(U,A,V,f)$ 为信息系统，其中 $U=\{x_1,x_2,\cdots,x_n\}$ 是非空的有限对象集，$A=\{a_1,a_2,\cdots,a_m\}$ 是非空的有限属性集，$V=\bigcup_{a\in A}V_a$

是所有属性值的集合，其中 V_a 是属性 a 的值域，V_a 具有偏序关系，$f:U\times A\rightarrow V$，其中，对 $\forall a\in A$，$x\in U$，$f(x,a)\in V_a$. 对 $\forall B\subseteq A$ 记 $R_B^{\leqslant}=\{(x_i,x_j)\in U\times U \mid f(x_i,a)\leqslant f(x_j,a),\forall a\in B\}$，称 $R_B^{\leqslant}$ 为 $S=\{U,A,V,f\}$ 上的优势关系，优势类 $[x_i]_B^{\leqslant}=\{x_j \mid (x_i,x_j)\in R_B^{\leqslant}\}$ 为 x_i 的优势类.

定义 1.2[5]　假设 $S=(U,A,V,f)$ 是一个信息系统，$\forall X\subseteq U$，则集合 X 的下近似和上近似在优势关系 $R_B^{\leqslant}$ 下可以表示为

$$\underline{R_B^{\leqslant}}(X)=\{x_i \mid [x_i]_B^{\leqslant}\subseteq X\} \tag{1}$$

$$\overline{R_B^{\leqslant}}(X)=\{x_i \mid [x_i]_B^{\leqslant}\cap X\neq\varnothing\} \tag{2}$$

式中，$\underline{R_B^{\leqslant}}(X)$ 表示论域 U 中确定属于 X 的元素集合，$\overline{R_B^{\leqslant}}(X)$ 表示论域 U 中可能属于 X 的元素集合，$bnr_B^{\leqslant}(X)=\overline{R_B^{\leqslant}}(X)-\underline{R_B^{\leqslant}}(X)$ 表示集合 X 的边界集.

1.2　优势关系多粒度粗糙集

定义 1.3[6]　令 $S=(U,A,V,f)$ 为信息系统，其中属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，乐观多粒度粗糙集关于集合 X 的下近似集表示为 $\underline{\sum_{i=1}^{m}a_i^O}(X)$，上近似集表示为 $\overline{\sum_{i=1}^{m}a_i^O}(X)$，那么

$$\underline{\sum_{i=1}^{m}a_i}^{O}(X)=\{x\in U:\exists i\in\{1,2,\cdots,m\},[x]_{a_i}\subseteq X\} \tag{3}$$

$$\overline{\sum_{i=1}^{m}a_i}^{O}(X)=\sim\underline{\sum_{i=1}^{m}a_i}^{O}(\sim X) \tag{4}$$

式中，$[x]_{a_i}(1\leqslant i\leqslant m)$ 是元素 x 在属性 a_i 下的等价类，$\sim X$ 是集合 X 的补集.

定义 1.4[6]　令 $S=\{U,A,V,f\}$ 为信息系统，其中 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，悲观多粒度粗糙集关于集合 X 的下近似集表示为 $\underline{\sum_{i=1}^{m}a_i}^{P}(X)$，上近似集表示分别为 $\overline{\sum_{i=1}^{m}a_i}^{P}(X)$，那么

$$\underline{\sum_{i=1}^{m}a_i}^{P}(X)=\{x\in U:\forall i\in\{1,2,\cdots,m\},[x]_{a_i}\subseteq X\} \tag{5}$$

$$\overline{\sum_{i=1}^{m}a_i}^{P}(X)=\sim\underline{\sum_{i=1}^{m}a_i}^{P}(\sim X) \tag{6}$$

式中，$[x]_{a_i}(1\leqslant i\leqslant m)$ 是元素 x 在属性 a_i 下的等价类，$\sim X$ 表示集合 X 的补集.

下面介绍优势关系多粒度粗糙集的相关定义和定理.

定义 1.5[10]　令 $S=(U,A,V,f)$ 为信息系统，其中属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，优势关系乐观多粒度粗糙集关于集合 X 的下近似集和上近似集分别表示为 $\underline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)$ 和 $\overline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)$，其中

$$\underline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)=\{x\in U:\exists i\in\{1,2,\cdots,m\},[x]_{a_i}^{\leqslant}\subseteq X\} \tag{7}$$

$$\overline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)=\sim\underline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(\sim X) \tag{8}$$

式中，$[x]_{a_i}^{\leqslant}(i\in 1,2,\cdots,m)$ 是元素 x 在属性 a_i 下的优势类，$\sim X$ 表示集合 X 的补集.

定义 1.6[10]　令 $S=(U,A,V,f)$ 为信息系统，其中 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，优势关系悲观多粒度粗糙集关于集合 X 的下近似集和上近似集分别表示为 $\underline{\sum_{i=1}^{m}a_i}{}^{P}_{\leqslant}(X)$ 和 $\overline{\sum_{i=1}^{m}a_i}{}^{P}_{\leqslant}(X)$，那么

$$\underline{\sum_{i=1}^{m}a_i}{}^{P}_{\leqslant}(X)=\{x\in U:\forall i\in\{1,2,\cdots,m\},[x]_{a_i}^{\leqslant}\subseteq X\} \tag{9}$$

$$\overline{\sum_{i=1}^{m}a_i}{}^{P}_{\leqslant}(X)=\sim\underline{\sum_{i=1}^{m}a_i}{}^{P}_{\leqslant}(\sim X) \tag{10}$$

式中，$[x]_{a_i}^{\leqslant}(1\leqslant i\leqslant m)$ 是元素 x 在属性 a_i 下的优势类，$\sim X$ 表示集合 X 的补集.

定理 1.1　令 $S=(U,A,V,f)$ 为信息系统，其中属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，可得

$$\overline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)=\{x\in U:\forall i\in\{1,2,\cdots,m\},[x]_{a_i}^{\leqslant}\cap X\neq\varnothing\} \tag{11}$$

证明　假设 $x\in\overline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)$，那么根据式(8)可得，$x\notin\underline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(\sim X)$，所以，根据定义 1.5 可以得到，$x\in\{[x]_{a_1}^{\leqslant}\not\subset\sim X\wedge[x]_{a_2}^{\leqslant}\not\subset\sim X\wedge\cdots\wedge[x]_m^{\leqslant}\not\subset\sim X\}$，因此 $x\in\{[x]_{a_1}^{\leqslant}\cap X\neq\varnothing\wedge[x]_{a_2}^{\leqslant}\cap X\neq\varnothing\wedge\cdots\wedge[x]_m^{\leqslant}\cap X\neq\varnothing\}$；证得 $\overline{\sum_{i=1}^{m}a_i}{}^{O}_{\leqslant}(X)=\{x\in U:\forall i\in\{1,2,\cdots,m\},[x]_{a_i}^{\leqslant}\cap X\neq\varnothing\}$.

定理 1.2　令 $S=(U,A,V,f)$ 为信息系统，其中属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对 $\forall X\subseteq U$，可得

$$\overline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)=$$
$$\{x\in U:\exists i\in\{1,2,\cdots,m\},[x]_{a_i}^{\leqslant}\cap X\neq\varnothing\}\tag{12}$$

证明　证明过程类似于定理 1.1 的证明过程.

2　粒度结构增加时优势关系多粒度粗糙集中 近似集动态更新方法

2.1　粒度结构增加时近似集动态获取方法

下面介绍当粒度结构增加时动态更新的定理和性质. 假设 $S=(U,A,V,f)$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，$\forall X\subseteq U$，信息系统中粒度增加时，增加的粒度为 a_{m+1}，优势关系乐观多粒度粗糙集，上、下近似集分别表示为 $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)$，$\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)$，优势关系悲观多粒度粗糙集上、下近似集分别表示为 $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)$，$\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)$.

性质 2.1　令 $S=(U,A,V,f)$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对象集 $U=\{x_1,x_2,\cdots,x_n\}$，$\forall X\subseteq U$，对优势关系乐观多粒度粗糙集而言，当增加单个粒度结构时，可知

① $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)\supseteq\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$；

② $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)\subseteq\overline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$.

证明　① $\forall x\in U$，$x\in\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$，根据优势关系乐观多粒度粗糙集下近似集定义可知，$\exists i\in\{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant}\subseteq X$，因此，$\exists i\in\{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant}\subseteq X$，即 $x\in\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)$，证得 $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)\supseteq\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$.

②同理可证 $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)\subseteq\overline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$.

性质 2.2　令 $S=(U,A,V,f)$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对象集 $U=\{x_1,x_2,\cdots,x_n\}$，$\forall X\subseteq U$，对优势关系悲观多粒度粗糙集而言，当增加单个粒度结构时

① $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)\subseteq\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)$；

② $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)\supseteq\overline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)$.

证明　① $\forall x\in U$，$x\in\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)$，根据悲观多粒度粗糙集定义，可知 $\forall i\in\{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant}\subseteq X$，因此 $\forall i\in\{1,2,\cdots m\}$，$[x]_{a_i}^{\leqslant}\subseteq X$，所以 $x\in\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)$，证得 $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)\subseteq\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)$.

②同理可证 $\overline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{P}(X)\supseteq\overline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{P}(X)$.

定理 2.1　令 $S=(U,A,V,f)$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对象集 $U=\{x_1,x_2,\cdots,x_n\}$，$\forall X\subseteq U$，当增加粒度结构 a_{m+1} 时，如果 $x\notin\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$，那么

① 若 $[x]_{a_{m+1}}^{\leqslant}\subseteq X$，则 $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)=\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)\cup\{x\}$；

②否则，$\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)=\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$.

证明　① $x\notin\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$，根据定义可知 $\forall i\in\{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant}\not\subset X$，因为 $[x]_{a_{m+1}}^{\leqslant}\subseteq X$，所以 $\exists i\in\{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant}\subseteq X$. 根据优势关系乐观多粒度下近似集的定义可知，$x\in\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)$，所以下近似集增加，证得 $\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)=\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)\cup\{x\}$.

② $x\notin\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$ 根据定义可知，$\forall i\in\{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant}\not\subset X$. $\because[x]_{a_{m+1}}^{\leqslant}\not\subset X$，对 $\forall i\in\{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant}\not\subset X$，$x\notin\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)$，当粒结构增加后，下近似集不变，即证得，$\underline{\sum_{i=1}^{m+1}a_i}{}_{\leqslant}^{O}(X)=\underline{\sum_{i=1}^{m}a_i}{}_{\leqslant}^{O}(X)$

定理 2.2　令 $S=(U,A,V,f)$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，对象集 $U=\{x_1,x_2,\cdots,x_n\}$，$\forall X\subseteq U$，当增加粒结构 a_{m+1} 时，

如果 $x \in \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X)$，那么

①若 $[x]_{a_{m+1}}^{\leqslant} \cap X = \varnothing$，则 $\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X) - \{x\}$；

②否则，$\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X)$.

证明 ① $x \in \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X)$ 根据优势关系乐观多粒度粗糙集上近似集定义，$\forall i \in \{1,\cdots m\}$，$[x]_{a_i}^{\leqslant} \cap X \neq \varnothing$，因为 $[x]_{a_{m+1}}^{\leqslant} \cap X = \varnothing$，所以 $\exists i \in \{1,2,\cdots,m+1\}$，若 x 关于所增加的粒度结构的优势关系所形成的优势类满足 $[x]_{a_i}^{\leqslant} \cap X = \varnothing$，根据定义 $x \notin \overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X)$，所以，上近似集减少，证得 $\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X) - \{x\}$.

②因为 $x \in \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X)$，根据优势关系乐观多粒度粗糙集的上近似集定义可知，对 $\forall i \in \{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant} \cap X \neq \varnothing$. 当粒度结构增加时，若 x 关于所增加的粒度结构的优势关系所形成的优势类满足条件 $[x]_{a_{m+1}}^{\leqslant} \cap X \neq \varnothing$，则 $\forall i \in \{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant} \cap X \neq \varnothing$. 根据定义可知，$x \in \overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X)$，所以上近似集不变，证得，$\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{O}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{O}}(X)$.

定理 2.3 令 $S = \{U,A,V,f\}$ 为信息系统，属性集 $A = \{a_1,a_2,\cdots,a_m\}$，对象集 $U = \{x_1,x_2,\cdots,x_n\}$，$\forall X \subseteq U$，当增加粒结构 a_{m+1} 时，如果 $x \in \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，那么

① 若 $[x]_{a_{m+1}}^{\leqslant} \not\subset X$，则 $\underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X) - \{x\}$；

②否则，$\underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$.

证明 ① $x \in \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，根据优势关系悲观多粒度粗糙集下近似集定义，$\forall i \in \{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant} \subseteq X$，因为 $[x]_{a_{m+1}}^{\leqslant} \not\subset X$，所以 x 不满足 $\forall i \in \{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant} \subseteq X$. 根据定义可知，$x \notin \underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X)$，所以证得 $\underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X) - \{x\}$.

② $x \in \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，根据优势关系悲观多粒度粗糙集下近似集定义，$\forall i \in \{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant} \subseteq X$，因为 $[x]_{a_{m+1}}^{\leqslant} \subseteq X$，所以 x 满足 $\forall i \in \{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant} \subseteq X$. 根据定义可知，$x \in \underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X)$，所以下近似集不变. 即证得 $\underline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \underline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$.

定理 2.4 令 $S = \{U,A,V,f\}$ 为信息系统，属性集 $A = \{a_1,a_2,\cdots,a_m\}$，对象集 $U = \{x_1,x_2,\cdots,x_n\}$，$\forall X \subseteq U$，当增加粒结构 a_{m+1} 时，如果 $x \notin \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，那么

①若 $[x]_{a_{m+1}}^{\leqslant} \cap X \neq \varnothing$，则 $\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X) \cup \{x\}$；

②否则，$\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$.

证明 ①因为 $x \notin \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，根据优势关系悲观多粒度粗糙集上近似集定义，$\exists i \in \{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant} \cap X \neq \varnothing$. 因为 $[x]_{a_{m+1}}^{\leqslant} \cap X \neq \varnothing$，所以 x 满足 $\exists i \in \{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant} \cap X \neq \varnothing$，根据定义可知，$x \in \overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X)$，证得 $\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X) \cup \{x\}$.

②因为 $x \notin \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$，根据优势关系悲观多粒度粗糙集上近似集定义，$\forall i \in \{1,2,\cdots,m\}$，$[x]_{a_i}^{\leqslant} \cap X = \varnothing$. 因为 $[x]_{a_{m+1}}^{\leqslant} \cap X = \varnothing$，所以 x 满足 $\forall i \in \{1,2,\cdots,m+1\}$，$[x]_{a_i}^{\leqslant} \cap X = \varnothing$. 根据定义可知，$x \notin \overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X)$，即证得 $\overline{\sum_{i=1}^{m+1} a_i {}_{\leqslant}^{P}}(X) = \overline{\sum_{i=1}^{m} a_i {}_{\leqslant}^{P}}(X)$.

2.2 算法设计与分析

前一部分给出了优势关系多粒度粗糙集中，当

单个粒度结构增加时，信息系统中的近似集变化的相关定理. 当信息系统发生变化时，如果采用原始的静态更新算法，需要重新扫描整个信息系统，求出每个粒度结构下所有对象之间的优势类；然后再根据优势关系乐观和悲观多粒度粗糙集中下近似集和上近似集的定义求某个给定集合的下近似集和上近似集，此过程需要重复扫描信息系统求优势类、下近似集和上近似集，花费的时间代价较多，采用本文提出的动态更新算法，不需要重新扫描所有的信息系统中的优势类，只需要根据增加的粒度结构，计算所有对象相对于增加粒度结构的优势类，根据相关定理判断优势关系下乐观多粒度粗糙集和悲观多粒度粗糙集的上、下近似集的变化情况. 下面根据提出的定理，给出动态更新近似集算法. 其中，当多个粒度结构增加时，可以看成多个单个粒度结构依次增加.

算法 2.1　当粒度结构增加时，优势关系多粒度粗糙集中近似集更新算法

输入：①令 $S=\{U,A,V,f\}$ 为信息系统，属性集 $A=\{a_1,a_2,\cdots,a_m\}$，$\forall X\subseteq U$；

②原信息系统中的优势类$[x]_{a_i}^{\leqslant}$；

③增加的粒度结构 a_{m+1}；

④原信息系统中的近似集 $\underline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$，$\overline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$，$\underline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$，$\overline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$.

输出：信息系统中，增加粒度结构之后的近似集 $\underline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)$，$\underline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)$，$\overline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)$，$\overline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)$.

第一步：对象集中所有对象 $x\in U$

计算每个对象相对 a_{m+1} 的优势类$[x]_{a_{m+1}}^{\leqslant}$

第二步：//根据定理 3 计算优势关系乐观多粒度下近似集

扫描所有的对象判断 $x\notin\underline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$

如果$[x]_{a_{m+1}}^{\leqslant}\subseteq X$，那么

$\underline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)=\underline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)\cup\{x\}$

否则，$\underline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)=\underline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$

第三步：//根据定理 2.2 计算优势关系乐观多粒度上近似集

扫描所有的对象判断 $x\in\overline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$

如果$[x]_{a_{m+1}}^{\leqslant}\cap X=\varnothing$，那么

$\overline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)=\overline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)-\{x\}$

否则，$\overline{\sum_{i=1}^{m+1}a_i\overset{O}{\leqslant}}(X)=\overline{\sum_{i=1}^{m}a_i\overset{O}{\leqslant}}(X)$

第四步：//根据定理 2.3 计算优势关系悲观多粒度下近似集

对所有的对象判断 $x\in\underline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$

如果$[x]_{a_{m+1}}^{\leqslant}\not\subset X$，那么

$\underline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)=\underline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)-\{x\}$

否则，$\underline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)=\underline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$

第五步：//根据定理 2.4 计算优势关系悲观多粒度上近似集

对所有的对象判断 $x\notin\overline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$

如果$[x]_{a_{m+1}}^{\leqslant}\cap X\neq\varnothing$，那么

$\overline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)=\overline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)\cup\{x\}$

否则，$\overline{\sum_{i=1}^{m+1}a_i\overset{P}{\leqslant}}(X)=\overline{\sum_{i=1}^{m}a_i\overset{P}{\leqslant}}(X)$

通过分析可知，设 S 表示原信息系统，$|U|$ 表示对象个数，$|C|$ 表示原信息系统中粒度结构数，对静态算法而言，当增加单个粒度结构时，计算整个论域对象相对于所有粒度结构的优势类的时间复杂度为 $O(|U|^2\cdot(|C|+1))$，扫描所有数据对象计算近似集的时间复杂度为 $O(|U|\cdot|C|)$，故更新近似集总的时间复杂度为 $O(|U|^2\cdot(|C|+1)+|U|\cdot|C|)$；采用动态更新算法，粒度结构增加时，计算增加粒度结构的优势类花费时间为 $O(|U|^2)$，计算近似集的花费时间为 $O(|U|)$，此时，采用所提出算法总的花费时间为 $O(|U|^2+|U|)$，因此动态更新算法的时间复杂度优于静态算法.

3　实验结果与分析

动态更新算法的实验平台为：Windows 7 操作系统，Intel（R） Core（TM） i3-4010U Duo @ 1.70GHz，4G 内存，在 Visual Studio 2012 开发平台中进行仿真实验，算法采用 C++实现. 算法测试采用的 6 个数据集均来自于 UCI 数据库，如表 1 所示. 实验过程中，先对数据进行预处理，选取信息系统中单个粒度结构作为表示增加的粒度结构，信息系统中其余的粒度结构作为原始的信息系统，将每个数据集划分成 10 份，增加粒度结构时采用第 1 份作为初始数据采用静态算法和动态更新算法分别计算近似集；再将数据集中的第 2 份合并到第 1 份数据集中. 在此基础上增加单个粒度结构时，分别计算获取近似集，依次将剩余的数据集增加到先前的数据集中，然后进行近似集的动态更新计算，进行 10 次实验，得到在数据集上静态算法和动态更新算法获取近似集的时间，实验结果如图 1 所示. 每个子图

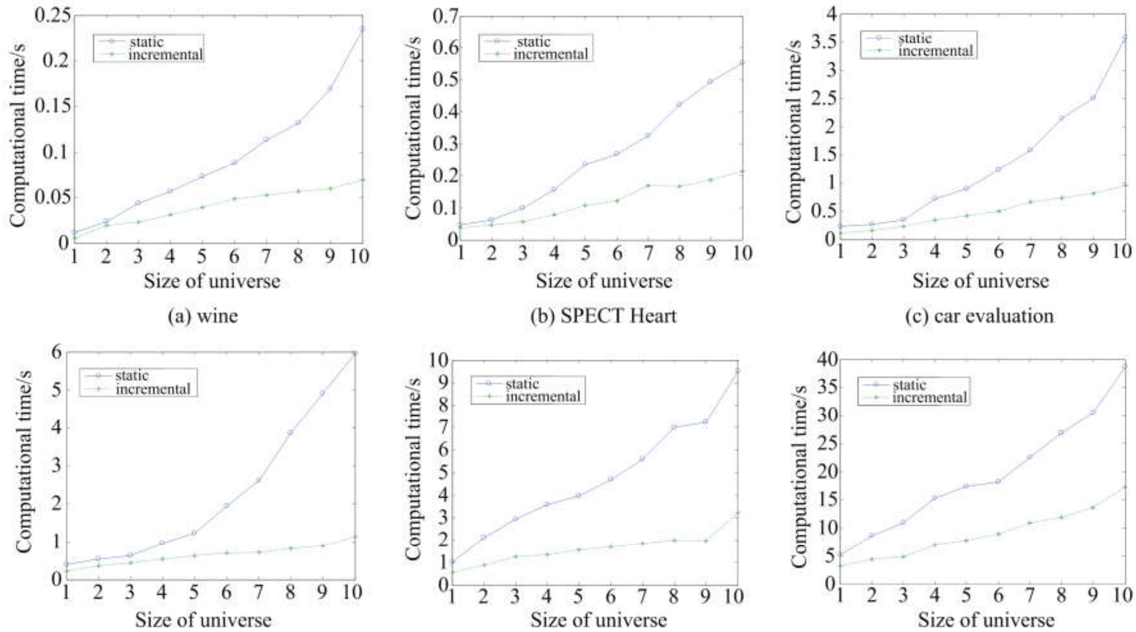
表示一个数据集的静态算法和动态更新算法的计算近似集的时间，x 轴表示数据集的大小，y 轴表示计算近似集的时间.

表 1 实验数据集的描述

Tab. 1 Description of experiment dataset

Data sets	Samples	Attribute sets
wine	178	13
SPECT Heart	267	22
car evaluation	1 728	6
chess(king-rook vs. king-pawn)	3 196	36
statlog	4 435	37
mushroom	8 124	22

从图 1 的每个子图可以看出，静态算法和动态更新算法计算近似集的时间都随着数据集的增大而增加. 在开始阶段，静态算法和动态更新算法的效率相差不多，但随着数据的增加，动态算法获取知识的效率有一定的优势. 分析原因可知，在信息系统粒度结构增加时，采用静态算法求解近似集需要重复计算所有数据对象相对于粒度结构的优势类，在重新扫描时，花费的时间代价较高，而本文提出的动态算法，在计算近似集的过程中，可以使用原有信息系统中的近似集，只需对增加的粒度结构计算每个对象的优势类，然后根据优势关系乐观多粒度粗糙集和悲观多粒度粗糙集的相关知识计算近似集，避免了优势类和近似集的重新计算，花费的时间比静态算法少，提高了效率.



(a) wine (b) SPECT Heart (c) car evaluation
(d) chess (e) statlog (f) mushroom

图 1 粒度结构增加时静态算法与动态更新算法计算时间比较图

Fig. 1 Comparison time of static algorithm and incremental algorithm while adding granulation structure

4 结论

由于信息系统随着数据量的增加而不断变化，本文从多粒度的角度考虑，提出一种优势关系多粒度粗糙集模型中，当粒度结构动态增加时上、下近似集的获取方法，得到了近似集动态更新的相关定理，通过实验验证了该方法在计算近似集时效率比静态算法有所提高；然而，在多粒度环境下，当信息系统中的对象和属性值发生变化时，如何有效地动态获取有用的知识以及多粒度环境下的动态属性约简等都是需要进一步研究的内容.

参考文献(References)

[1] PAWLAK Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5):341-356.

[2] HASHEMI R R, CHOOBINEH F F. A fuzzy rough sets classifier for database mining[J]. International Journal of Smart Engineering System Design, 2002, 4

(2):107-114.

[3] 梁吉业，李德玉. 信息系统中的不确定性与知识获取[M]. 北京：科学出版社，2005.

[4] 王国胤. Rough 集理论与知识获取[M]. 西安：西安交通大学出版社，2001.

[5] Greco S., Matarazzo B., Slowinski R. Rough approximation by dominance relation[J]. International Journal of Intelligent Systems, 2002, 17(2):153-171.

[6] 鲍忠奎，杨善林. 直觉模糊目标信息系统的正域约简[J]. 中国科学技术大学学报，2015，45(4):329-336.
BAO Zhongkui, YANG Shanlin. Positive domain reduction in intuitionistic fuzzy objective information systems [J]. Journal of University of Science and Technology of China, 2015, 45(4):329-336.

[7] QIAN Y H, LIANG J Y, YAO Y Y, et al. MGRS: A multi-granulation rough set[J]. Information Sciences, 2010, 180: 949-970.

[8] XU W H, WANG Q R, ZHANG X T. Multi-granulation rough sets based on tolerance relations[J]. Soft Computing, 2013, 17(7): 1241-1252.

[9] QIAN Y H, LIANG J Y, DANG C Y. Incomplete multigranulation rough set[J]. IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans, 2010, 40(2):420-431.

[10] HUANG B, GUO C X, ZHANG Y L, et al. Intuitionistic fuzzy multigranulation rough sets [J]. Information Sciences, 2014, 277: 299-320.

[11] 庄颖，刘文奇，范敏，等. 集值信息系统上的多粒度优势关系与信息融合[J]. 模式识别与人工智能，2015，28(8):741-749.

[12] 翟永健，张宏. 不完备信息系统中的优势关系多粒度粗糙集[J]. 南京理工大学学报，2012，36(1):66-72.

[13] YANG XIBEI, QI YONG, YU HUALONG, et al. Updating multigranulation rough approximations with increasing of granular structures[J]. Knowledge-Based Systems, 2014, 64:59-69.

[14] JU H G, YANG XB, SONG X N, et al. Dynamic updating multigranulation fuzzy rough set: approximations and reducts[J]. International Journal of Machine Learning & Cybernetics, 2014, 5(6): 981-990.

[15] 张清华，王国胤，肖雨. 粗糙集的近似集[J]. 软件学报，2012，23(7):1745-1759.

[16] 张清华，王进，王国胤. 粗糙模糊集的近似表示[J]. 计算机学报，2015，38(7):1457-1470.

[17] LIU D, LI T R, RUAN D, et al. An incremental approach for inducing knowledge from dynamic information systems [J]. Fundamenta Informaticae, 2009, 94(2):245-260.

[18] 张清华，幸禹可，周玉兰. 基于粒计算的增量式知识获取方法[J]. 电子与信息学报，2011，33(2):435-441.

[19] CHEN H M, LI T R, RUAN D, et al. A rough-set based incremental approach for updating approximations under dynamic maintenance environments[J]. IEEE Transactions on Knowledge & Data Engineering, 2013, 25(2):274-284.

[20] ZHANG J B, LI T R, RUAN D, et al. Neighborhood rough sets for dynamic data mining[J]. International Journal of Intelligent Systems, 2012, 27(4): 317-342.

[21] LI S Y, LI T R, LIU D. Dynamic maintenance of approximations in dominance-based rough set approach under the variation of the object set[J]. International Journal of Intelligent Systems, 2013, 28(8): 729-751.

[22] CHAN C C. A rough set approach to attribute generalization in data mining[J]. Information Sciences, 1998, 107: 177-194.

[23] LI T R, MA J, XU Y, et al. An approach to attribute generalization in incomplete information system[C]// Proceedings of the International Conference on Machine Learning and Cybernetics. Xi′an, China: IEEE Press, 2003: 1678-1691.

[24] LI T R, RUAN D, GEERT W, et al. A rough sets based characteristic relation approach for dynamic attribute generalization in data mining[J]. Knowledge-based Systems, 2007, 20(5): 485-494.

[25] LI S Y, LI T R, LIU D. Incremental updating approximations in dominance-based rough sets approach under the variation of the attribute set[J]. Knowledge-based Systems, 2013, 40(1):17-26.

[26] WANG F, LIANG J Y, QIAN Y H. Attribute reduction: A dimension incremental strategy [J]. Knowledge-Based Systems, 2013, 39(2): 95-108.

[27] 胡成祥，赵瑞斌. 基于经典粗糙集的近似集动态获取方法研究[J]. 数据采集与处理，2015，30(6):1332-1340.

[28] 胡成祥. 变精度粗糙集中属性变化时近似集获取方法[J]. 计算机科学与探索，2015，9(11):1398-1408.

[29] CHEN H M, LI T R, QIAO S J, et, al. A rough set based dynamic maintenance approach for approximations in coarsening and refining attribute values[J]. International Journal of Intelligent System, 2010, 25: 1005-1026.

[30] CHEN H M, LI T R, RUAN D. Dynamic maintenance of approximations under a rough-set based variable precision limited tolerance relation[J]. Journal of Multiple-valued Logic and Soft Computing, 2012, 8(5): 577-598.

[31] LUO C, LI T R, CHEN H M, et al. Fast algorithms for computing rough approximations in set-valued decision systems while updating criteria values [J]. Information Sciences, 2015, 299: 221-242.