基于加权粒度和优势关系的程度多粒度粗糙集近似集的动态并行更新算法

赵艺琳[1] 姜 麟[1] 米允龙[2] 李金海[1,3]

(昆明理工大学理学院 昆明 650500)[1] (中国科学院大学计算机与控制学院 北京 101408)[2]

(昆明理工大学数据科学研究中心 昆明 650500)[3]

摘 要 随着大数据集的不断更新,经典的多粒度粗糙集理论不再适用。为此,提出加权粒度优势关系程度悲观多粒度粗糙集与加权粒度优势关系程度乐观多粒度粗糙集的相关理论。在此基础上,给出了一种基于加权粒度和优势关系的程度多粒度粗糙集近似集的动态并行更新算法。最后,通过实验验证了所提算法的有效性,其能够应对海量动态更新的数据变化并提升运行效率。

关键词 多粒度粗糙集,加权,优势关系,并行更新算法

中图法分类号 TP182 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.10.003

Dynamic Parallel Updating Algorithm for Approximate Sets of Graded Multi-granulation Rough Set Based on Weighting Granulations and Dominance Relation

ZHAO Yi-lin[1] JIANG Lin[1] MI Yun-long[2] LI Jin-hai[1,3]

(Faculty of Science, Kunming University of Science and Technology, Kunming 650500, China)[1]

(School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing 101408, China)[2]

(Data Science Research Center, Kunming University of Science and Technology, Kunming 650500, China)[3]

Abstract With the continuous updating of large data sets, the classical multi-granulation rough set theory is no longer practical. Therefore, this paper put forward the related theory of graded pessimistic multi-granulation rough set with weighting granulations and dominance relation, graded optimistic multi-granulation rough set with weighting granulations and dominance relation. On the basis of this improved theory, this paper proposed a dynamic parallel updating algorithm for approximate sets of graded multi-granulation rough set based on weighting granulations and dominance relation. Finally, the experiment verifies the effectiveness of the proposed algorithm, which is able to handle data with massive dynamic updates and improve running efficiency.

Keywords Multi-granulation rough set, Weighting, Dominance relation, Parallel updating algorithm

1 引言

粗糙集理论是由 Pawlak 教授于 1982 年提出的一种能够定量分析和处理不精确、不一致、不完整等各种信息的有效工具[1],目前已经在模式识别[2]、知识获取[3]和图像处理[4]等领域得到了广泛应用。

然而,从粒计算角度分析,经典的粗糙集理论是建立在单粒度基础上的,而现实中的问题往往是在多角度和多粒度环境下考虑的,因此经典粗糙集并不能有效求解这类问题[5-11]。为此,一些学者将粗糙集拓展到多粒度粗糙集并获得了较好的成果[12-18]。多粒度粗糙集的核心思想是以多个等价关系(或拓展的一般关系)为基础,利用上近似算子、下近似算子对知识进行组合和有效刻画,从而进行粒度结构约简和规则获取等方面的研究。近期,多粒度粗糙集的探讨慢慢转向交叉融合方向。比如,Li 等[19]从规则提取的角度比较了多粒度粗糙集和概念格的优劣,Lin 等[20]将证据理论与多粒度粗糙集结合起来,Yang 等[21]把代价敏感分析与多粒度粗糙集融合起来。

然而,面对海量数据,虽然已有学者研究了单粒度或多粒度粗糙集的动态更新算法[22-30],但是随着数据量的不断增大,已有的多粒度粗糙集动态更新算法由于均在串行条件下设计实施方案[26-30],因此难以达到较高的计算效率。另一方面,就实际应用背景而言,将加权粒度和优势关系同时融入多粒度粗糙集是非常重要的,具体参见文献[31]描述的患者治

到稿日期:2018-04-17 返修日期:2018-05-18 本文受国家自然科学基金(61562050,KKGD201707071),云南省教育厅基金(KKJB201707008)资助。

赵艺琳(1994—),女,硕士生,主要研究方向为并行数据挖掘和机器学习;**姜 麟**(1969—),男,硕士,副教授,主要研究方向为智能计算和并行计算,E-mail:tojianglin@163.com(通信作者);**米允龙**(1987—),男,博士生,主要研究方向为数据挖掘、机器学习与认知计算;**李金海**(1984—),男,博士,教授,主要研究方向为粗糙集、概念格与粒计算。

疗诊断过程分析。

基于上述两个基本认识，本文针对多粒度粗糙集，通过融合加权粒度和优势关系，提出一种多粒度粗糙集近似集的动态并行更新算法。实验表明，该算法能够适应数据不断增加的情景，并且有着较高的计算效率。

本文第 2 节介绍与本文相关的粗糙集理论的相关知识；第 3 节给出了加权粒度优势关系程度多粒度粗糙集近似集的更新理论；第 4 节对所提出的模型与方法进行实例验证，以说明本文在多粒度粗糙集模型的基础上融入加权粒度和优势关系的必要性；第 5 节给出并行算法，并进行数值实验评估；最后总结全文，并给出需要进一步研究的问题。

2 预备知识

定义 1[12] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集，$\forall X\subseteq U$，则 X 关于属性子集 A 的乐观多粒度粗糙集的下近似集和上近似集分别定义为：

$$\underline{\sum_{i=1}^{m}A_i^O}(X)=\{x\in U:[x]_{A_1}\subseteq X\vee[x]_{A_2}\subseteq X\vee\cdots\vee[x]_{A_m}\subseteq X\} \tag{1}$$

$$\overline{\sum_{i=1}^{m}A_i^O}(X)=\sim\underline{\sum_{i=1}^{m}A_i^O}(\sim X) \tag{2}$$

其中，$\sim X$ 表示 X 的绝对补，称序$\langle\underline{\sum_{i=1}^{m}A_i^O}(X),\overline{\sum_{i=1}^{m}A_i^O}(X)\rangle$为集合 X 关于属性子集 A 的乐观多粒度粗糙集。

定义 2[12] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集，$\forall X\subseteq U$，则 X 关于属性子集 A 的悲观多粒度粗糙集的下近似集和上近似集分别定义为：

$$\underline{\sum_{i=1}^{m}A_i^P}(X)=\{x\in U:[x]_{A_1}\subseteq X\wedge[x]_{A_2}\subseteq X\wedge\cdots\wedge[x]_{A_m}\subseteq X\} \tag{3}$$

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

其中，$\sim X$ 表示 X 的绝对补，称序偶$\langle\underline{\sum_{i=1}^{m}A_i^P}(X),\overline{\sum_{i=1}^{m}A_i^P}(X)\rangle$为集合 X 关于属性子集 A 的悲观多粒度粗糙集。

定义 3[10] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A\subseteq AT$，$k$ 为非负常数，$\forall X\subseteq U$，X 关于 A 依程度 k 的下近似算子 $\underline{A}_k$ 和上近似算子 $\overline{A}_k$ 分别定义为：

$$\underline{A}_k=\{x\in U:|[x]_A|-|[x]_A\cap X|\leqslant k\} \tag{5}$$

$$\overline{A}_k=\{x\in U:|[x]_A\cap X|>k\} \tag{6}$$

称序偶$\langle\underline{A}_k,\overline{A}_k\rangle$为集合 X 关于 A 依程度 k 的程度粗糙集。

定义 4[15] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集，k 为非负常数，$\forall X\subseteq U$，则程度乐观多粒度粗糙集的下近似和上近似分别定义为：

$$\underline{\sum_{i=1}^{m}A_{i_k}^O}(X)=\{x\in U:|[x]_{A_1}|-|[x]_{A_1}\cap X|\leqslant k\vee|[x]_{A_2}|-|[x]_{A_2}\cap X|\leqslant k\vee\cdots\vee|[x]_{A_m}|-|[x]_{A_m}\cap X|\leqslant k\} \tag{7}$$

$$\overline{\sum_{i=1}^{m}A_{i_k}^O}(X)=\sim\underline{\sum_{i=1}^{m}A_{i_k}^O}(\sim X) \tag{8}$$

称序偶$\langle\underline{\sum_{i=1}^{m}A_{i_k}^O}(X),\overline{\sum_{i=1}^{m}A_{i_k}^O}(X)\rangle$为集合 X 依程度 k 的程度乐观多粒度粗糙集。

定义 5[15] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集，k 为非负常数，$\forall X\subseteq U$，则程度悲观多粒度粗糙集的下近似和上近似分别定义为：

$$\underline{\sum_{i=1}^{m}A_{i_k}^P}(X)=\{x\in U:|[x]_{A_1}|-|[x]_{A_1}\cap X|\leqslant k\wedge|[x]_{A_2}|-|[x]_{A_2}\cap X|\leqslant k\wedge\cdots\wedge|[x]_{A_m}|-|[x]_{A_m}\cap X|\leqslant k\} \tag{9}$$

$$\overline{\sum_{i=1}^{m}A_{i_k}^P}(X)=\sim\underline{\sum_{i=1}^{m}A_{i_k}^P}(\sim X) \tag{10}$$

称序偶$\langle\underline{\sum_{i=1}^{m}A_{i_k}^P}(X),\overline{\sum_{i=1}^{m}A_{i_k}^P}(X)\rangle$为集合 X 依程度 k 的程度悲观多粒度粗糙集。

定义 6[11] 设 $S=\langle U,A,V,f\rangle$ 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 的值域且具有偏序关系；$f:U\times A\to V$，其中，对于 $\forall a\in A,x\in U$，有 $f(x,a)\in V_a$，对于 $\forall B\subseteq A$，记 $R_B^{\leqslant}=\{(x_i,x_j)\in U\times U|f(x_i,a)\leqslant f(x_j,a),\forall a\in B\}$，称 $R_B^{\leqslant}$ 为 $S=\langle U,A,V,f\rangle$上的优势关系，优势类$[x_i]_B^{\leqslant}=\{x_j|(x_i,x_j)\in R_B^{\leqslant}\}$为 x_i 的优势类。

定义 7[11] 设 $S=\langle U,A,V,f\rangle$为信息系统，$\forall X\subseteq U$，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：

$$\underline{R_B^{\leqslant}}(X)=\{x_i|[x_i]_B^{\leqslant}\subseteq X\} \tag{11}$$

$$\overline{R_B^{\leqslant}}(X)=\{x_i|[x_i]_B^{\leqslant}\cap X\neq\varnothing\} \tag{12}$$

其中，$\underline{R_B^{\leqslant}}(X)$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。$bnr_B^{\leqslant}=\overline{R_B^{\leqslant}}(X)-\underline{R_B^{\leqslant}}(X)$，表示集合 X 的边界集。

定义 8[18] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$\forall X\subseteq U$，$A\subseteq AT$，$k$ 为非负整数，X 的优势关系程度粗糙集的下近似和上近似分别为：

$$\underline{A}_k^{\leqslant}=\{x\in U:|R_A^{\leqslant}(x)|-|R_A^{\leqslant}(x)\cap X|\leqslant k\wedge R_A^{\leqslant}(x)\cap X\neq\varnothing\} \tag{13}$$

$$\overline{A}_k^{\leqslant}=\{x\in U:|R_A^{\leqslant}(x)\cap X|>k\} \tag{14}$$

程度粗糙集可能存在一个问题，即 $[x]_A\cap X=\varnothing$ 时，$|[x]_A|-|[x]_A\cap X|\leqslant k$ 有可能成立，则 $x\in\underline{A}_k^{\leqslant}(X)$，这是不符合实际的。定义 8 在 X 的优势关系程度粗糙集的下近似中增加了条件 $R_A^{\leqslant}(x)\cap X\neq\varnothing$，避免了这种情况。

定义 9[17] 设信息系统 $IS=\langle U,AT,V,f\rangle$，$\forall X\subseteq U$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集，$0<\beta\leqslant 1$，粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则 X 关于 A 的 β 加权多粒度粗糙集的下近似和上近似为：

$$\underline{\sum_{i=1}^{\beta}A_i}_{\alpha}(X)=\{x\in U:\forall A_i\in T,[x]_{A_i}\subseteq X,T\subseteq A\wedge\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta\} \tag{15}$$

$$\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_i^{\beta}}_{\alpha}(X)=\sim\underline{\sum_{i=1}^{m}}^{\leqslant}A_i^{\beta}(\sim X) \tag{16}$$

定义 10[18]　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,$0<\beta\leqslant 1$,对于 $\forall X\subseteq U$,k 为非负整数,粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$,且 $\sum_{i=1}^{m}\alpha_{A_i}=1$,则 X 关于 A 的加权粒度优势关系程度多粒度粗糙集的下近似集和上近似集为:

$$\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)=\{x\in U:\forall A_i\in T,x\in\underline{A_{i_k}^{\leqslant}}(X),T\subseteq A,\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta\} \tag{17}$$

$$\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\sim\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(\sim X) \tag{18}$$

定义 11　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,对于 $\forall X\subseteq U$,k 为非负整数,粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$,且 $\sum_{i=1}^{m}\alpha_{A_i}=1$,则 X 关于 A 的加权粒度优势关系程度乐观多粒度粗糙集的下近似集和上近似集分别为:

$$\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)=\{x\in U:\forall A_i\in T,x\in\underline{A_{i_k}^{\leqslant}}(X),T\subseteq A,\sum_{i=1}^{|T|}\alpha_{A_i}>0\} \tag{19}$$

$$\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}}_{\alpha}(X)=\sim\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(\sim X) \tag{20}$$

定义 12　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,对于 $\forall X\subseteq U$,k 为非负整数,粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$,且 $\sum_{i=1}^{m}\alpha_{A_i}=1$,则 X 关于 A 的加权粒度优势关系程度悲观多粒度粗糙集的下近似集和上近似集分别为:

$$\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(X)=\{x\in U:\forall A_i\in T,x\in\underline{A_{i_k}^{\leqslant}}(X),T\subseteq A,\sum_{i=1}^{|T|}\alpha_{A_i}=1\} \tag{21}$$

$$\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X)=\sim\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(\sim X) \tag{22}$$

根据多粒度粗糙集的方法,依然可以构建乐观和悲观两种不同加权粒度优势关系程度多粒度粗糙集。本文构建的加权粒度优势关系程度乐观多粒度粗糙集的下近似集中至少有一个具有粒度权重的粒结构满足 $|R_A^{\leqslant}(x)|-|R_A^{\leqslant}(x)\cap X|\leqslant k\wedge R_A^{\leqslant}(x)\cap X\neq\varnothing$ 时,x 就属于 X 的下近似;加权粒度优势关系程度悲观多粒度粗糙集的下近似集中所有具有粒度权重的粒结构都满足 $|R_A^{\leqslant}(x)|-|R_A^{\leqslant}(x)\cap X|\leqslant k\wedge R_A^{\leqslant}(x)\cap X\neq\varnothing$ 时,x 才属于 X 的下近似。加权粒度优势关系程度多粒度粗糙集认为,只要有一定数目的粒结构满足程度条件,即可成为下近似。

3　加权粒度优势关系程度多粒度粗糙集近似集更新理论

加权粒度优势关系程度多粒度粗糙集是在程度多粒度粗糙集的基础上加入了粒度权重与优势关系的概念,使得模型在具有粒度权重和优势关系的基础上允许有一定程度的误差。下面介绍当粒度结构增加时,近似集动态更新的理论。

假设 $IS=\langle U,AT,V,f\rangle$ 为信息系统,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,对于 $\forall X\subseteq U$,当信息系统增加粒度结构时,加权粒度优势关系程度多粒度粗糙集的上近似和下近似分别表示为 $\overline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{\beta}{}_{\alpha}(X)$,$\underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{\beta}(X)$。

性质 1　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,$0<\beta\leqslant 1$,对于 $\forall X\subseteq U$,k 为非负整数,粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$,且 $\sum_{i=1}^{m}\alpha_{A_i}=1$,则 X 关于 A 的加权粒度优势关系程度悲观多粒度粗糙集、加权粒度优势关系程度乐观多粒度粗糙集和加权粒度优势关系程度多粒度粗糙集具有如下性质:

$$1)\ \underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X) \tag{23}$$

$$2)\ \overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}}_{\alpha}(X)\subseteq\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)\subseteq\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X) \tag{24}$$

证明:1)对于 $\forall x\in U$,若 $x\in\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(X)$,根据定义 12,对于 $\forall A_i\in T,i\in\{1,2,\cdots,m\}$,$x\in\underline{A_{i_k}^{\leqslant}}(X)$,有 $\sum_{j=1}^{|T|}\alpha_{A_j}=1$。对于 $\forall x\in U$,若 $x\in\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)$,根据定义 10,$\forall A_i\in T,i\in\{1,2,\cdots,m\}$,$x\in\underline{A_{i_k}^{\leqslant}}(X)$,有 $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$,因为 $0<\beta\leqslant 1$,即证得 $\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)$。同理可证 $\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)$,从而证得 $\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{P}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{\beta}(X)\subseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)$。

2)由定义 10、定义 11 和定义 12 类似可证得。

性质 2　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,对于 $\forall X\subseteq U$,k 为非负整数,粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$,且 $\sum_{i=1}^{m}\alpha_{A_i}=1$,则对于 X 关于 A 的加权粒度优势关系程度乐观多粒度粗糙集而言,当增加单个粒度结构,粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$,且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时,有如下性质:

$$1)\ \underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}(X)\supseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X) \tag{25}$$

$$2)\ \overline{\underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}}_{\alpha}(X)\subseteq\overline{\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}}_{\alpha}(X) \tag{26}$$

证明:1)对于 $\forall x\in U$,若 $x\in\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)$,根据定义 11,对于 $\forall A_i\in T,i\in\{1,2,\cdots,m\}$,$x\in\underline{A_{i_k}^{\leqslant}}(X)$,有 $\sum_{j=1}^{|T|}\alpha_{A_j}>0$,因此对于 $\forall A_i\in T,i\in\{1,2,\cdots,m,m+1\}$,$x\in\underline{A_{i_k}^{\leqslant}}(X)$,有 $\sum_{i=1}^{|T|}\alpha_{A_i}>0$,即 $x\in\underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}(X)$,从而证得 $\underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}(X)\supseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)$。

2)同理可证 $\underline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}(\sim X)\supseteq\underline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(\sim X)$,即 $\sim\overline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}(X)\supseteq\sim\overline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}(X)$,从而证得 $\overline{\sum_{i=1}^{m+1}}^{\leqslant}A_{i_k}^{O}{}_{\alpha}(X)\subseteq\overline{\sum_{i=1}^{m}}^{\leqslant}A_{i_k}^{O}{}_{\alpha}(X)$。

性质 3　设信息系统 $IS=\langle U,AT,V,f\rangle$,$A=\{A_1,A_2,\cdots,A_m\}$是 AT 的 m 个属性子集,对于 $\forall X\subseteq U$,k 为非负整数,粒

度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则对于 X 关于 A 的基于加权粒度和优势关系的程度悲观多粒度粗糙集而言，当增加单个粒度结构，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时，有如下性质：

$$1)\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X)\subseteq\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X) \tag{27}$$

$$2)\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X)\supseteq\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{P}}_{\alpha}(X) \tag{28}$$

证明：性质 3 的证明与性质 2 的证明类似。

定理 1 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$ 是 AT 的 m 个属性子集，$0<\beta\leqslant1$，对于 $\forall X\subseteq U$，k 为非负整数，粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则对于 X 关于 A 的基于加权粒度和优势关系的程度多粒度粗糙集而言，当增加单个粒度结构，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时，若 $x\in\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，则对于 $\forall A_i\in T$，$x\in\underline{A_{i_k}^{\leqslant}}(X)$，$i\in\{1,2,\cdots,m,m+1\}$，有：

1)若 $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，则 $\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$；

2)否则 $\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)-\{x\}$。

证明：1)若 $x\in\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，根据定义 10，对于 $\forall A_i\in T$，$i\in\{1,2,\cdots,m\}$，$x\in\underline{A_{i_k}^{\leqslant}}(X)$，其中 $T\subseteq A$，$\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，当增加单个粒度结构时，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$，若 x 满足 $\forall A_i\in T$，$i\in\{1,2,\cdots,m,m+1\}$，$x\in\underline{A_{i_k}^{\leqslant}}(X)$，其中 $T\subseteq A$，$\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，根据定义 10，$x\in\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，因此下近似不变，从而证得 $\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$。

2)由定义 10 类似可证。

定理 2 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$ 是 AT 的 m 个属性子集，$0<\beta\leqslant1$，对于 $\forall X\subseteq U$，k 为非负整数，粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则对于 X 关于 A 的基于加权粒度和优势关系的程度多粒度粗糙集而言，当增加单个粒度结构，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时，若 $x\notin\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，则对于 $\forall A_i\in T$，$x\in\underline{A_{i_k}^{\leqslant}}(X)$，$i\in\{1,2,\cdots,m,m+1\}$，有：

1)若 $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，则 $\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)\cup\{x\}$；

2)否则，$\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$。

证明：定理 2 的证明与定理 1 的证明类似。

定理 3 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$ 是 AT 的 m 个属性子集，$0<\beta\leqslant1$，对于 $\forall X\subseteq U$，k 为非负整数，粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则对于 X 关于 A 的基于加权粒度和优势关系的程度多粒度粗糙集而言，当增加单个粒度结构，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时，若 $x\in\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，则对于 $\forall A_i\in T$，$x\notin\overline{A_{i_k}^{\leqslant}}(X)$，$i\in\{1,2,\cdots,m,m+1\}$，有：

1)若 $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，则 $\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)-\{x\}$；

2)否则，$\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$。

证明：定理 3 的证明与定理 1 的证明类似。

定理 4 设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$ 是 AT 的 m 个属性子集，$0<\beta\leqslant1$，对于 $\forall X\subseteq U$，k 为非负整数，粒度空间 A 的粒度权重为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m}\}$，且 $\sum_{i=1}^{m}\alpha_{A_i}=1$，则对于 X 关于 A 的基于加权粒度和优势关系的程度多粒度粗糙集而言，当增加单个粒度结构，粒度空间 A 的粒度权重变为 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$，且 $\sum_{i=1}^{m+1}\alpha_{A_i}=1$ 时，若 $x\notin\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$，则对于 $\forall A_i\in T$，$x\notin\overline{A_{i_k}^{\leqslant}}(X)$，$i\in\{1,2,\cdots,m,m+1\}$：

1)如果 $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$，则 $\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)$；

2)否则，$\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}{}^{\leqslant}A_{i_k}^{\beta}}_{\alpha}(X)\cup\{x\}$。

证明：定理 4 的证明与定理 1 的证明类似。

4 实例验证

以下使用诊断病例对本文所提理论与方法进行验证。设对象集 $U=\{x_1,x_2,\cdots,x_{11}\}$ 为待决策的病例；$C=\{a_1,a_2,a_3,a_4,a_5\}$ 为条件属性集，$a_1,a_2,\cdots,a_5$ 分别表示病人的体温、是否出现恶心、尿频、尿急、尿道肿胀；d 为决策变量，$d=1$ 表示该病例患有肾盂肾炎，$d=0$ 表示该病例未患肾盂肾炎。

诊断病例的决策信息系统如表 1 所列。

表 1 决策信息系统

Table 1 Decision information system

U	a_1	a_2	a_3	a_4	a_5	d
x_1	35.9	0	1	1	1	0
x_2	36.0	0	1	1	1	0
x_3	38.0	0	1	0	1	1
x_4	40.0	1	1	1	0	1
x_5	37.9	0	0	0	0	0
x_6	41.0	1	0	1	0	1
x_7	37.6	0	1	1	0	0
x_8	41.5	0	1	0	1	1
x_9	41.2	1	1	1	1	1
x_{10}	40.2	0	0	0	0	0
x_{11}	40.9	1	1	1	0	1

由表 1 可知 $U/IND(D)=\{D_1,D_2\}$，其中 $D_1=\{x_3,x_4,x_6,x_8,x_9,x_{11}\}$，$D_2=\{x_1,x_2,x_5,x_7,x_{10}\}$。设决策信息系统中的属性集 $A=\{A_1,A_2,A_3,A_4\}=\{\{a_1\},\{a_2\},\{a_3,a_4\},\{a_5\}\}$，根据专家的建议[31]对 A 的粒度权值的分配为 $\alpha=$

{0.43,0.1,0.3,0.17},阈值 $\beta=0.7$,$k=3$。

$R^{\leqslant}_{A_1}(x_1)=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_2)=\{x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_3)=\{x_3,x_4,x_6,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_4)=\{x_4,x_6,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_5)=\{x_3,x_4,x_5,x_6,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_6)=\{x_6,x_8,x_9\}$

$R^{\leqslant}_{A_1}(x_7)=\{x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11}\}$

$R^{\leqslant}_{A_1}(x_8)=\{x_8\}$

$R^{\leqslant}_{A_1}(x_9)=\{x_8,x_9\}$

$R^{\leqslant}_{A_1}(x_{10})=R^{\leqslant}_{A_1}(x_{11})=\{x_6,x_8,x_9,x_{10},x_{11}\}$

$\underline{A^{\leqslant}_{1_k}}(D_1)=\{x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11}\}$

$\underline{A^{\leqslant}_{1_k}}(D_2)=\varnothing$。

同理可得:

$\underline{A^{\leqslant}_{2_k}}(D_1)=\{x_4,x_6,x_9,x_{11}\}$

$\underline{A^{\leqslant}_{3_k}}(D_1)=\{x_1,x_2,x_3,x_4,x_6,x_7,x_8,x_9,x_{11}\}$

$\underline{A^{\leqslant}_{4_k}}(D_1)=\underline{A^{\leqslant}_{4_k}}(D_2)=\{x_1,x_2,x_3,x_8,x_9\}$

$\underline{A^{\leqslant}_{2_k}}(D_2)=\varnothing$

$\underline{A^{\leqslant}_{3_k}}(D_2)=\{x_1,x_2,x_4,x_7,x_9,x_{11}\}$

$\underline{\sum_{i=1}^{4}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_1)=\{x_3,x_4,x_6,x_7,x_8,x_9,x_{11}\}$;$\overline{\sum_{i=1}^{4}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_1)=U$

$\underline{\sum_{i=1}^{4}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_2)=\varnothing$;$\overline{\sum_{i=1}^{4}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_2)=\{x_1,x_2,x_5,x_{10}\}$

表 2 在表 1 的信息系统中增加一个粒度结构 a_6,条件属性集变为 $C=\{a_1,a_2,a_3,a_4,a_5,a_6\}$,$a_1,a_2,\cdots,a_6$ 分别表示病人的体温、是否出现恶心、尿频、尿急、尿道肿胀、肾区是否有叩击痛。设决策信息系统中的属性集 $A=\{A_1,A_2,A_3,A_4,A_5\}=\{\{a_1\},\{a_2\},\{a_3,a_4\},\{a_5\},\{a_6\}\}$,根据专家的建议[31]将粒度结构 A 的粒度权值分配变为 $\alpha=\{0.3,0.1,0.2,0.1,0.3\}$,阈值 $\beta=0.7$,$k=3$。

表 2　更新后的决策信息系统

Table 2　Updated decision information system

U	a_1	a_2	a_3	a_4	a_5	a_6	d
x_1	35.9	0	1	1	1	0	0
x_2	36.0	0	1	1	1	0	0
x_3	38.0	0	1	0	1	1	1
x_4	40.0	1	1	1	0	1	1
x_5	37.9	0	0	0	0	1	0
x_6	41.0	1	0	1	0	1	1
x_7	37.6	0	1	1	0	0	0
x_8	41.5	0	1	0	1	1	1
x_9	41.2	1	1	1	1	1	1
x_{10}	40.2	0	0	0	0	0	0
x_{11}	40.9	1	1	1	0	1	1

基于加权粒度和优势关系的程度多粒度粗糙集的近似集为:

$\underline{\sum_{i=1}^{5}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_1)=\{x_3,x_4,x_6,x_8,x_9,x_{11}\}$;$\overline{\sum_{i=1}^{5}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_1)=U$;

$\underline{\sum_{i=1}^{5}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_2)=\varnothing$;$\overline{\sum_{i=1}^{5}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(D_2)=\{x_1,x_2,x_5,x_7,x_{10}\}$

根据上述结果分析,对于对象 x_7,在原始粒度知识空间中有 $x_7\in A^{\leqslant}_{1_k}(D_1)$,$x_7\in A^{\leqslant}_{3_k}(D_1)$,对象 x_7 满足一定数目的加权粒度空间(0.43+0.3=0.73>β)。当增加粒度 A_5 之后,对象 x_7 不满足一定数目的加权粒度空间(0.3+0.2=0.5<β)。因此,将 x_7 从 D_1 的加权粒度优势关系程度多粒度粗糙集下近似中剔除。

5　基于加权粒度和优势关系的程度多粒度粗糙集近似集更新算法

5.1　近似集的动态更新算法

当粒度结构增加时,原始算法因为存在每增加一个粒度结构则需要遍历整个信息系统、计算时间过长等不足,所以不适用于数据量较大的信息系统。为此,本文提出了动态更新算法。首先,计算出新增粒度结构的优势关系;其次,计算其优势关系程度粗糙集的上近似和下近似;最后,根据定理判断加权粒度优势关系程度多粒度粗糙集的上近似、下近似的变化情况,从而实现不需要计算整个信息系统就可以进行更新的效果。动态更新算法的具体步骤如算法 1 所示。

算法 1　Update approximation

输入:原始信息系统 IS=〈U,AT,V,f〉,属性集 $A=\{A_1,A_2,\cdots,A_m\}$;增加的粒度结构 IS_add 与新的粒度权重 $\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$;程度 k 与阈值 β;原信息系统中优势关系程度粗糙集的近似集 $\underline{A^{\leqslant}_{i_k}}(X)$ 和 $\overline{A^{\leqslant}_{i_k}}(X)$;原信息系统中的近似集 $\underline{\sum_{i=1}^{m}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(X)$ 和 $\overline{\sum_{i=1}^{m}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(X)$

输出:信息系统中增加粒度结构之后的近似集 $\underline{\sum_{i=1}^{m+1}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(X)$ 和 $\overline{\sum_{i=1}^{m+1}{}^{\leqslant}A^{\beta}_{i_k}}_{\alpha}(X)$

```
1. BEGIN
2. R^≤_{A_{m+1}}(X)=[ ]/*储存增加的粒度结构优势关系*/
3. R^≤_{A_{m+1}}=advantage(IS_add)/*判断新增粒度结构的优势关系*/
4. FOR each x_i∈U DO/*新增粒度结构的优势关系程度粗糙集近似集*/
5.    IF |R^≤_{A_{m+1}}(x_i)|-|R^≤_{A_{m+1}}(x_i)∩X|≤k∧R^≤_{A_{m+1}}(x_i)∩X≠∅ THEN
6.       x_i∈A^≤_{m+1_k}(X)
7.    ENDIF
8.    IF |R^≤_{A_{m+1}}(x_i)∩X|>k THEN
9.       x_i∈Ā^≤_{m+1_k}(X)
10.   ENDIF
11. ENDFOR
12. FOR each x_i∈U DO/*计算加权粒度优势关系程度多粒度粗糙集近似集*/
13.   IF x∈Σ^{m,≤}_{i=1}A^β_{i_k}(X) AND Σ^{|T|}_{j=1}α_{A_j}<β THEN
14.      Σ^{m+1,≤}_{i=1}A^β_{i_k}(X)=Σ^{m,≤}_{i=1}A^β_{i_k}(X)-{x}
15.   ENDIF
16.   IF x∉Σ^{m,≤}_{i=1}A^β_{i_k}(X) AND Σ^{|T|}_{j=1}α_{A_j}≥β THEN
17.      Σ^{m+1,≤}_{i=1}A^β_{i_k}(X)=Σ^{m,≤}_{i=1}A^β_{i_k}(X)∪{x}
18.   ENDIF
```

19. IF $x\in\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha'_{A_j}\geqslant\beta$ THEN
20. $\overline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)-\{x\}$
21. ENDIF
22. IF $x\notin\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha'_{A_j}<\beta$ THEN
23. $\overline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)\cup\{x\}$
24. ENDIF
25. ENDFOR
26. END

5.2 近似集的动态更新并行算法

MATLAB已经成为数值计算领域的主流工具，其中的并行计算工具箱(Parallel Computing Toolbox，PCT)和并行计算服务(Distributed Computing Server，DCS)为常用的并行计算问题提供了相对简单、快捷的方法。SPMD(Single Program，Multiple Data)是MATLAB支持的一种并行结构，可以对单个程序多个数据的情况进行并行处理。SPMD可以与串行程序结合起来，串行程序在客户端MATLAB上执行。SPMD并行程序在多个lab上执行，每个lab在物理上对应单个CPU核或单个CPU，各计算单元之间通过共享内存结构或网络传输数据[32]。因此，根据上述动态更新算法，利用SPMD并行结构创建分布式阵列，提出动态更新算法的并行构造。算法2是基于算法1进行的改进。

算法2 Para-Update approximation

输入：原始信息系统 $IS=\langle U,AT,V,f\rangle$，属性集 $A=\{A_1,A_2,\cdots,A_m\}$；增加的粒度结构 IS_add 与新的粒度权重；$\alpha=\{\alpha_{A_1},\alpha_{A_2},\cdots,\alpha_{A_m},\alpha_{A_{m+1}}\}$；程度 k 与阈值 β；原信息系统中的优势关系程度粗糙集近似集 $\underline{A_{i_k}^{\leqslant}}(X)$ 和 $\overline{A_{i_k}^{\leqslant}}(X)$；原信息系统中的近似集 $\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ 和 $\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$

输出：信息系统中增加粒度结构之后的近似集 $\underline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ 和 $\overline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$

1. BEGIN
2. $R_{A_{m+1}}^{\leqslant}(X)=[\ ]$ /* 储存新增粒度结构优势关系 */
3. $R_{A_{m+1}}^{\leqslant}(X)=$ advantage(IS_add)/* 得出新增粒度结构的优势关系 */
4. SPMD/* 开启并行结构 */
5. codist=codistributor1d(1) /* 创建沿一维对阵列进行分割的分布式阵列 */
6. $R_{A_{m+1}}^{\leqslant}(X)=$codistributed($R_{A_{m+1}}^{\leqslant}(X)$,codist)/* 将新增优势关系的数据按对象进行分割 */
7. FOR each $x_i\in U$ DO /* 将分割后的优势关系分组执行 */
8. IF $|R_{A_{m+1}}^{\leqslant}(x_i)|-|R_{A_{m+1}}^{\leqslant}(x_i)\cap X|\leqslant k$ AND $R_{A_{m+1}}^{\leqslant}(x_i)\cap X\neq\varnothing$ THEN
9. $x_i\in\underline{A_{m+1_k}^{\leqslant}}(X)$
10. ENDIF
11. IF $|R_{A_{m+1}}^{\leqslant}(x_i)\cap X|>k$ THEN
12. $x_i\in\overline{A_{m+1_k}^{\leqslant}}(X)$
13. ENDIF
14. ENDFOR
15. ENDSPMD/* 关闭并行结构 */
16. PARFOR each $x_i\in U$ DO/* 将数据按对象分组执行 */
17. IF $x\in\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha_{A_j}<\beta$ THEN
18. $\underline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)-\{x\}$
19. ENDIF
20. IF $x\notin\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha_{A_j}\geqslant\beta$ THEN
21. $\underline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)\cup\{x\}$
22. ENDIF
23. IF $x\in\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha'_{A_j}\geqslant\beta$ THEN
24. $\overline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)-\{x\}$
25. ENDIF
26. IF $x\notin\underline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)$ AND $\sum_{j=1}^{|T|}\alpha'_{A_j}<\beta$ THEN
27. $\overline{\sum_{i=1}^{m+1}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)=\overline{\sum_{i=1}^{m}\leqslant A_{i_k}^{\beta}}_{\alpha}(X)\cup\{x\}$
28. ENDIF
29. ENDPARFOR
30. END

下面分析算法1和算法2的时间复杂度。设$|U|$和$|C|$分别代表信息系统中的对象数与粒度结构数，根据算法1，计算新增粒度结构的优势关系时，其最坏时间复杂度为$O(|U|^2)$，计算优势关系程度粗糙集近似集与加权粒度优势关系程度多粒度粗糙集近似集的时间复杂度均为$O(|U|)$，因此算法1的总时间复杂度为$O(|U|^2+2|U|)$。计算算法2的时间复杂度时，设程序运行时有N个节点，$O(|U|/N)$为耗时最长的节点的时间，则计算优势关系程度粗糙集的近似集与加权粒度优势关系程度多粒度粗糙集的近似集的时间复杂度均为$O(|U|/N)$，因此算法2的总时间复杂度为$O(|U|^2+2|U|/N)$。

5.3 动态更新算法的并行调度步骤

MATLAB作为并行计算的支持环境，提供了PCT(Parallel Computing Toolbox)和MDCS(Matlab Distributed Computing Server)。PCT只能支持单个集群节点的并行计算问题，如果要在多个节点上运行并行程序，需要同时配置PCT和MDCS，因此在集群中启动较大规模的并行程序时，首先需要启动和配置MDCS[32]。

该系统分为3个部分：提交job的客户端，管理job的Job Manager，执行任务的worker。算法2的并行过程中客户端首先将并行程序提交给job，由job将任务分配给各个worker，worker接收到这些任务后进行运算，运算完成后将结果返回给Job Manager，再由Job Manager将结果传送给客户端输出。分布计算要求若干个小任务相互独立，任务与任务之间不存在联系时，才能将相同的程序、不同的数据分配到相应的worker中执行。具体执行步骤如下：

Step 1 寻找Job Manager，即寻找适合计算任务的JM；

Step 2 创建job；

Step 3 分配task，即对创建的job分配task，task定义了worker的数量，大数据在此分解成小任务；

Step 4　提交 job，即将任务提交给 Job Manager，Job Manager 将任务分配给 task 中的 worker 进行计算；

Step 5　回收结果，当结果计算完成后，将结果返回给客户端；

Step 6　删除任务，释放内存。

算法 2 的流程如图 1 所示。首先，寻找合适的 Job Manager 创建 job，将并行程序递交给客户端，客户端将需要并行的部分递交给 Job Manager，由 Job Manager 将任务分配给各个 worker 执行，worker 执行完任务后将结果返回给 Job Manager，Job Manager 归并结果，并将结果返回给客户端输出。

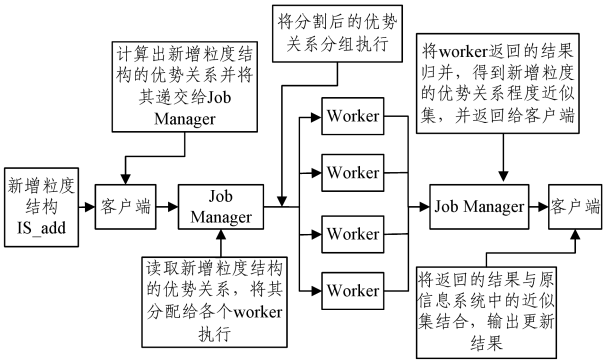

图 1　并行算法的流程

Fig. 1　Flow of parallel algorithm

5.4　数值实验与分析

本文实验环境如下：MDCS 集群由一台主控制节点和两台从节点构成，每一个节点的配置如下：CPU 为 Inter(R) Core(TM)2 Quad Q6600 2.40GHz，4.00 GB 内存和 500 GB 硬盘；操作系统为 Windows 7，MATLAB 2015b Win64 版本，编译环境为 MATLAB 2015b Win64，并行计算平台为 MATLAB Distributed Computing Server 6.7。如表 3 所列，实验采用 UCI 数据库的 6 个数据集，并对数据进行预处理。实验时，从信息系统中选取单个粒度结构作为增加对象，余下数据集作为原始的信息系统。将数据集等量划分成 10 份，第一份数据集作为初始数据进行第一次实验，再将第二份数据集与第一份数据集合并进行第二次实验，依次类推，对每个数据集进行 10 次实验，分别得到静态算法、静态并行算法、动态更新算法和动态并行更新算法在数据集上获取近似集的时间，实验结果如表 4、表 5 所列，其中的数据结果均为执行 10 次数据的均值。本文并行实验将集群的节点数分别控制为 2 和 4，并与串行算法进行对比实验，测试动态并行更新算法的效率，sta-P-2 表示静态并行算法在 2 个 worker 下运行的时间，inc-P-2 和 inc-P-4 分别表示动态并行更新算法在 2 个 worker 和 4 个 worker 下的运行时间。随着数据集的增大，该算法的变化趋势分别如图 2—图 4 所示。图 2 中的每个子图表示一个数据集的静态算法、静态并行算法和动态更新算法计算近似集的时间。图 3 中的每个子图表示一个数据集的静态并行算法和动态并行更新算法计算近似集的时间。图 4 中的每个子图表示一个数据集的动态更新算法和动态更新并行算法计算近似集的时间。

表 3　实验数据集

Table 3　Experimental date sets

Date sets	Samples	Attribute sets
Statlog(German Credit Data)	1000	20
Wilt	4889	6
Thyroid Disease	7200	21
EEG Eye State	14980	15
HTRU2	17898	9
Occupancy Detection	20560	7

表 4　实验结果(1)

Table 4　Experimental results(1)

No	Statlog(German Credit Data)					Wilt					Thyroid Disease				
	static	sta-P-2	incre	inc-P-2	inc-P-4	static	sta-P-2	incre	inc-P-2	inc-P-4	static	sta-P-2	incre	inc-P-2	inc-P-4
1	0.04	0.52	<0.01	0.23	0.24	0.15	0.56	0.03	0.27	0.28	1.33	1.20	0.08	0.30	0.31
2	0.11	0.55	<0.01	0.23	0.28	0.72	0.86	0.16	0.36	0.37	6.58	4.37	0.42	0.60	0.51
3	0.21	0.61	0.01	0.22	0.27	1.46	1.21	0.38	0.59	0.49	15.87	9.76	0.88	1.14	0.91
4	0.41	0.74	0.02	0.24	0.28	2.64	1.97	0.64	0.90	0.73	26.50	16.54	1.33	1.83	1.40
5	0.59	0.81	0.03	0.26	0.28	4.36	3.07	1.03	1.37	1.03	52.60	28.44	2.49	2.60	2.10
6	0.82	0.90	0.05	0.28	0.29	6.81	4.55	1.51	1.75	1.39	74.25	43.34	3.40	3.79	2.76
7	1.03	1.02	0.06	0.30	0.31	9.81	6.29	2.06	2.31	1.80	102.01	60.08	4.66	4.66	3.47
8	1.41	1.23	0.09	0.32	0.33	13.05	8.21	2.73	2.93	2.30	133.84	78.43	6.15	5.83	4.50
9	1.68	1.50	0.11	0.35	0.35	16.49	10.32	3.44	3.62	2.98	166.35	99.69	7.34	7.200	5.90
10	2.29	1.87	0.15	0.37	0.38	19.17	11.37	3.94	4.49	3.38	207.94	124.87	9.48	8.55	7.09

表 5　实验结果(2)

Table 5　Experimental results(2)

No	EGG Eye State					HTRU2					Occupancy Detection				
	static	sta-P-2	incre	inc-P-2	inc-P-4	static	sta-P-2	incre	inc-P-2	inc-P-4	static	sta-P-2	incre	inc-P-2	inc-P-4
1	1.37	1.26	0.35	0.55	0.48	3.49	2.76	0.57	0.81	0.65	2.54	1.94	0.68	0.93	0.74
2	6.37	4.23	1.41	1.79	1.38	17.40	11.04	2.29	2.60	2.01	12.81	7.86	2.66	2.93	2.27
3	15.36	9.56	3.15	3.69	2.53	40.21	24.66	5.04	5.24	3.96	29.22	17.37	5.90	5.78	4.73
4	27.64	16.78	5.47	6.32	4.29	72.60	44.23	9.09	10.02	6.95	56.76	33.17	11.33	11.18	7.83
5	44.36	27.00	8.78	9.43	6.69	124.84	76.22	15.46	15.15	10.65	97.24	56.54	19.26	17.19	12.48
6	69.45	42.40	14.20	12.67	9.31	193.41	116.14	24.03	19.64	15.09	155.72	90.45	30.81	25.27	17.01
7	99.83	60.78	19.79	17.17	12.88	280.46	168.25	34.94	26.63	19.88	213.28	123.54	41.87	35.04	23.76
8	140.46	85.10	28.25	23.53	16.02	369.90	221.35	45.99	36.90	25.88	277.45	160.63	54.42	43.14	29.60
9	180.51	109.79	36.17	29.85	21.07	467.33	280.53	57.85	44.47	31.60	352.55	204.40	69.24	48.65	37.50
10	222.75	135.91	44.19	37.36	26.25	577.42	346.70	71.45	52.16	38.62	434.26	252.71	85.20	67.65	44.67

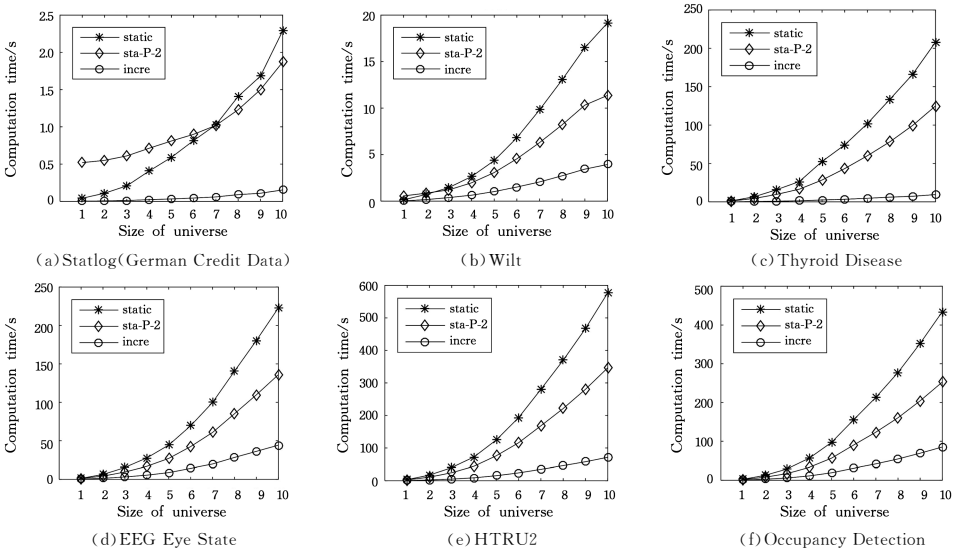

图 2 增加粒度结构时静态算法、静态并行算法与动态更新算法的耗时比较

Fig. 2 Comsuming time comparison of static algorithm, static parallel algorithm and incremental algorithm while adding granulation structure

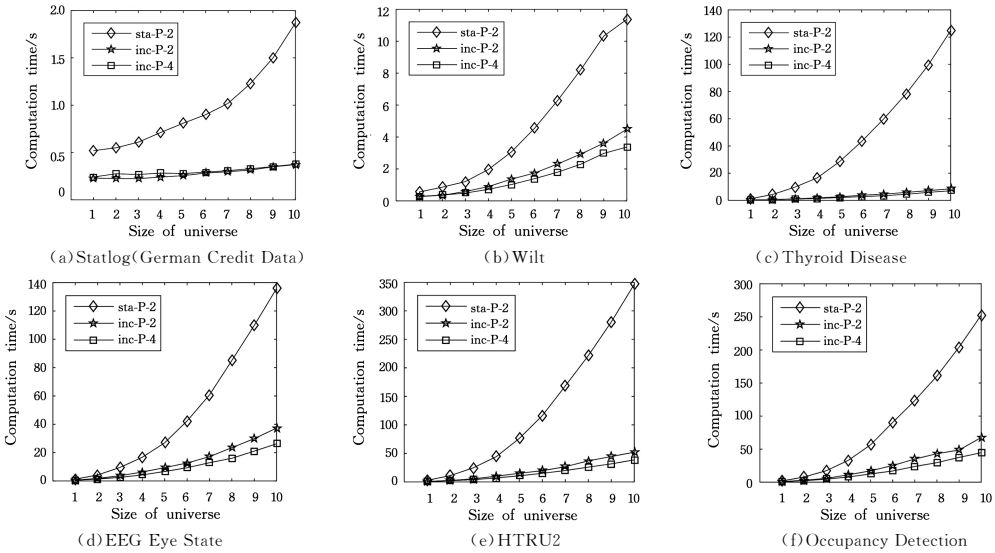

图 3 增加粒度结构时静态并行算法与动态更新并行算法的耗时比较

Fig. 3 Comsuming time comparison of static parallel algorithm and incremental parallel algorithm while adding granulation structure

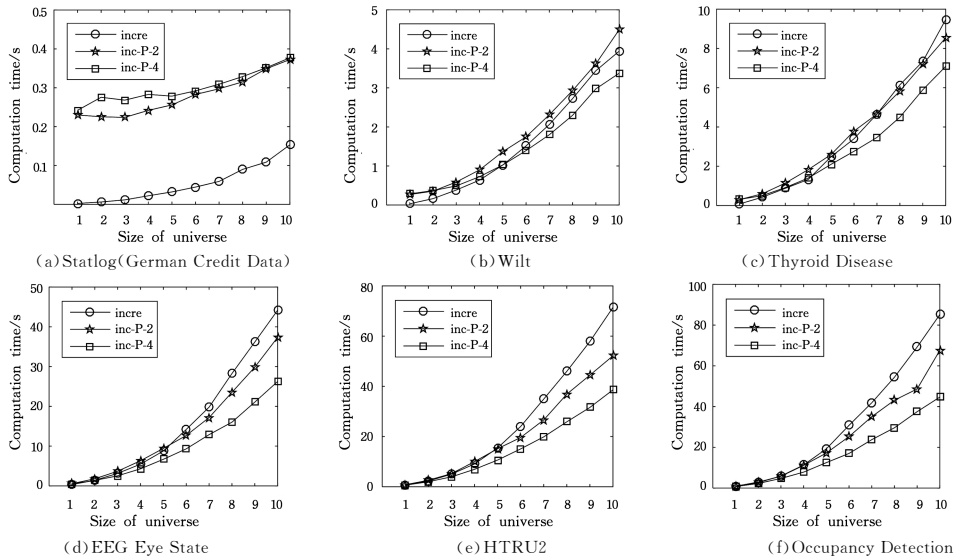

图 4 增加粒度结构时动态更新算法与动态更新并行算法的耗时比较

Fig. 4 Comsuming time comparison of incremental algorithm and incremental parallel algorithm while adding granulation structure

从图 2 的每个子图中可以看出，随着数据的增加，各算法之间的时间差逐渐增大。原始静态算法在计算近似集时需要重新计算每个属性，因此耗时较长；将其并行化后，可以提升算法的效率，若实验集群节点的数目增加，则该算法的效率将进一步得到提升。本文提出的动态算法只需重新计算新增的粒度结构，使用原信息系统中的近似集与新增粒度的近似集进行判断，避免了对信息系统进行重复计算，其耗时比静态算法与静态并行算法少，算法的效率得到提升。从图 3 的每个子图可以看出，在开始阶段，静态并行算法与动态并行更新算法的效率相差不大，但随着数据的增加，动态并行更新算法的运行效率逐渐提升。

因此，将动态更新算法与动态并行更新算法进行比较，如图 4 所示，从每个子图可以看出，在数据量较小的情况下，并行程序的运行效率低于串行程序的运行效率；但随着数据量的增大，并行程序的运行效率表现出一定的优势，并且随着集群节点数目的增加，算法的效率得到进一步提升。这是由于并行运算各个核心之间需要进行数据的通信，当数据集中的数据量较小时，核心之间的数据通信占整体运行时间的比例较高。由此可以看出，当数据量较大时，本文提出的动态并行更新算法可以进一步提升算法的效率，减少计算时间。

结束语　面对大数据集，数据的动态更新变得越来越困难。因此，本文提出一种基于加权粒度与优势关系的程度多粒度粗糙集模型，并对新增粒度结构进行学习，得到上近似空间和下近似空间。实验结果表明，串行算法在数据量较小时效率较高；但随着数据量的增加，并行算法的效率更高，并且随着节点数的增加，并行算法的效率得到进一步提升。

由于本文算法仅对一部分算法进行并行，因此需要进一步探索更好的并行方法，以提升算法效率；同时，当信息系统发生变化时，如何有效地获取决策规则等都是需要研究的问题。

最近，Hao 等在多粒度标记决策信息系统中研究了动态粒度下的最优粒度选择问题[33]，有关实现技术在多粒度粗糙集中有助于考虑保持近似集不变的粒度结构约简动态更新问题。

未来将就以上若干研究问题开展进一步的讨论，以推动多粒度粗糙集理论的发展。

参 考 文 献

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

[2] BAZAN J, PETERS J F, SKOWRON A, et al. Rough set approach to pattern extraction from classifiers[J]. Electronic Notes in Theoretical Computer Science, 2003, 82(4): 20-29.

[3] 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.

[4] PAL S K, MITRA P. Multispectral image segmentation using the rough-set-initialized EM algorithm [J]. IEEE Transactions on Geoscience and Remote Sensing, 2002, 40(11): 2495-2501.

[5] 张文修，仇国芳. 基于粗糙集的不确定性决策[M]. 北京：清华大学出版社，2005.

[6] MIAO D Q, ZHANG Q H, QIAN Y H, et al. From human intelligence to machine implementation model: theories and applications based on granular computing [J]. CAAI Transactions on Intelligent Systems, 2016, 11(6): 743-757. (in Chinese)
苗夺谦，张清华，钱宇华，等. 从人类智能到机器实现模型——粒计算理论与方法[J]. 智能系统学报，2016，11(6)：743-757.

[7] WANG G Y, ZHANG Q H, MA X A, et al. Granular computing models for knowledge uncertainty [J]. Journal of Software, 2011, 24(4): 676-694. (in Chinese)
王国胤，张清华，马希骜，等. 知识不确定性问题的粒计算模型[J]. 软件学报，2011，24(4)：676-694.

[8] ZHANG Y P, ZHANG L, WU T. The representation of different granular worlds: A quotient space [J]. Chinese Journal of Computers, 2004, 27(3): 238-333. (in Chinese)
张燕平，张铃，吴涛. 不同粒度世界的描述法——商空间法[J]. 计算机学报，2004，27(3)：238-333.

[9] LIANG J Y, QIAN Y H, LI D Y, et al. Theory and method of granular computing for big data mining [J]. Science China: Information Sciences, 2015, 45(11): 1355-1369. (in Chinese)
梁吉业，钱宇华，李德玉，等. 大数据挖掘的粒计算理论与方法[J]. 中国科学：信息科学，2015，45(11)：1355-1369.

[10] YAO Y Y, LIN T Y. Generalization of rough sets using modal logics [J]. Intelligent Automation and Soft Computing, 1996, 2(2): 103-120.

[11] GRECO S, MATARAZZO B, SLOWINSKI R. Rough approximation by dominance relations[J]. International Journal of Intelligent Systems, 2002, 17(2): 153-171.

[12] QIAN Y H, LIANG J Y, YAO Y Y, et al. MGRS: a multi-granulation rough set[J]. Information Sciences, 2010, 180(6): 949-970.

[13] QIAN Y H, LIANG J Y, WANG F. A positive approximation based accelerated algorithm to feature selection from incomplete decision tables [J]. Chinese Journal of Computers, 2011, 34(3): 435-442. (in Chinese)
钱宇华，梁吉业，王锋. 面向非完备决策表的正向近似特征选择加速算法[J]. 计算机学报，2011，34(3)：435-442.

[14] XU W H, WANG Q R, LUO S Q. Multi-granulation fuzzy rough sets [J]. Journal of Intelligent and Fuzzy Systems, 2014, 26(3): 1323-1340.

[15] WU Z Y, ZHONG P H, HU J G. Graded multi-granulation rough sets [J]. Fuzzy Systems and Mathematics, 2014, 28(3): 165-172. (in Chinese)
吴志远，钟培华，胡建根. 程度多粒度粗糙集[J]. 模糊系统与数学，2014，28(3)：165-172.

[16] ZHANG M, TANG Z M, XU W Y, et al. Variable multigranulation rough set model [J]. Pattern Recognition and Artificial Intelligence, 2012, 25(4): 709-720. (in Chinese)
张明，唐振民，徐维艳，等. 可变多粒度粗糙集模型[J]. 模式识别与人工智能，2012，25(4)：709-720.

[17] ZHANG M, CHENG K, YANG X B, et al. Multigranulation rough set based on weighted granulations [J]. Control and Decision, 2015, 30(2): 222-228. (in Chinese)
张明，程科，杨习贝，等. 基于加权粒度的多粒度粗糙集[J]. 控制与决策，2015，30(2)：222-228.

[18] WANG X Y, SHEN J L, SHEN Y X. Graded multi-granulation rough set based on weighting granulations and dominance relation[J]. Journal of Shandong University (Natural Science), 2017, 52(3): 97-104. (in Chinese)

汪小燕,沈家兰,申元霞.基于加权粒度和优势关系的程度多粒度粗糙集[J].山东大学学报(理学版),2017,52(3):97-104.

[19] LI J H,REN Y,MEI C L,et al. A comparative study of multigranulation rough sets and concept lattices via rule acquisition [J]. Knowledge-Based Systems,2016,91:152-164.

[20] LIN G P,LIANG J Y,QIAN Y H. An information fusion approach by combining multigranulation rough sets and evidence theory [J]. Information Sciences,2015,314:184-199.

[21] YANG X B,QI Y S,SONG X N,et al. Test cost sensitive multigranulation rough set:Model and minimal cost selection [J]. Information Sciences,2013,250:184-199.

[22] 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.

[23] 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.

[24] 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 and Data Engineering,2012,25(2):274-284.

[25] LIU W B,LI T R,ZOU W L,et al. Approaches for Incrementally Updating Approximations under Characteristic Relation-based Rough Sets While Attribute Values Coarsening and Refining[J]. Computer Science,2010,37(6):248-251. (in Chinese)
刘伟斌,李天瑞,邹维丽,等.特性关系粗糙集下属性值粗化细化时近似集增量更新方法研究[J].计算机科学,2010,37(6):248-251.

[26] YANG X B,QI Y,YU H L,et al. Updating multigranulation rough approximations with increasing of granular structures [J]. Knowledge-Based Systems,2014,64(1):59-69.

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

[28] HU C X,LIU S X,HUANG X L. Dynamic updating approximations in multigranulation rough sets while refining or coarsening attribute values [J]. Knowledge-Based Systems,2017,130:62-73.

[29] HU C X,LIU S X,LIU G X. Matrix-based approaches for dynamic updating approximations in multigranulation rough sets [J]. Knowledge-Based Systems,2017,122:51-63.

[30] HU C X,ZHAO G Z. A dominance-based multigranulation rough sets approach for dynamic updating approximations [J]. Journal of University of Science and Technology of China,2017(1):40-47. (in Chinese)
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(1):40-47.

[31] SUN A W. Experience in the diagnosis and treatment of acute pyelonephritis [J]. Chinese Journal of Medicine,1966,15(1):32-33. (in Chinese)
孙爱文.诊治急性肾盂肾炎的体会[J].中国医刊,1966,15(1):32-33.

[32] 刘维.实战 MATLAB之并行程序设计[M].北京:北京航空航天大学出版社,2012.

[33] HAO C,LI J H,FAN M,et al. Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions [J]. Information Sciences,2017,415:213-232.

(上接第 5 页)

[11] YAO Y Y. Granular computing and sequential three-way decisions [M]// Rough Sets and Knowledge Technology. Berlin:Springer,2013:16-27.

[12] LI H X,ZHANG L B,HUANG B,et al. Sequential three-way decision and granulation for cost-sensitive face recognition [J]. Knowledge-Based Systems,2016,91(C):241-251.

[13] LI H X,ZHANG L B,ZHOU X Z,et al. Cost-sensitive sequential three-way decision modeling using a deep neural network [J]. International Journal of Approximate Reasoning,2017,85(C):68-78.

[14] SAVCHENKO A V. Fast multi-class recognition of piecewise regular objects based on sequential three-way decisions and granular computing [J]. Knowledge-Based Systems,2016,91:252-262.

[15] QIAN J,DANG C Y,YUE X D,et al. Attribute reduction for sequential three-way decisions under dynamic granulation [J]. International Journal of Approximate Reasoning,2017,85:196-216.

[16] LI J H,HUANG C C,QI J J,et al. Three-way cognitive concept learning via multi-granularity [J]. Information Sciences,2017,378(1):244-263.

[17] HAO C,LI J H,FAN M,et al. Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions [J]. Information Sciences,2017,415:213-232.

[18] FANG Y,MIN F,LIU Z H,et al. Sequential three-way decisions based cost-sensitive approach to classification [J]. Journal of Nanjing University (Natural Science),2018,54(1):148-156. (in Chinese)
方宇,闵帆,刘忠慧,等.序贯三支决策的代价敏感分类方法[J].南京大学学报(自然科学),2018,54(1):148-156.

[19] YAO Y Y. Decision-theoretic rough set models [M]// Rough Sets and Knowledge Technology. Berlin:Springer,2007:1-12.

[20] YAO Y Y. Three-way decisions with probabilistic rough sets [J]. Information Sciences,2010,180(3):341-353.

[21] WU W Z,LEUNG Y. Theory and applications of granular labelled partitions in multi-scale decision tables [J]. Information Sciences,2011,181(18):3878-3897.