多准则分类问题中近似集的增量更新方法

李 艳 靳永飞 吴婷婷 郭娜娜 于 群
(河北大学数学与信息科学学院河北省机器学习与计算智能重点实验室 保定 071002)

摘 要 在优势关系粗糙集方法(DRSA)的框架下,优势关系可用于处理带有序关系属性(准则)的数据,并且已经被广泛用于处理多准则决策问题。然而在实际应用中,当属性集和对象集发生变化时,信息系统会随之不断更新。在这种动态环境下,DRSA中用于属性约简、规则提取以及决策制定的近似集需要得到相应的更新。针对对象集发生变化时(增加或删除一个对象)的多准则分类问题,采用增量方法来更新近似集并提出两种相应的更新算法DRSA1和DRSA2。同时,对不同情况下的更新原则进行了讨论并给出了相关的理论结果与详细的证明。最后给出算例,并在UCI数据集上进行大量的实验,与非增量的方法(传统的DRSA)进行对比,结果充分体现了所提增量方法的有效性与可扩展性。

关键词 优势关系粗糙集,多准则分类,信息系统,近似集,增量更新

中图法分类号 TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.12.012

Incrementally Updating Approximations Approach in Dominance-based Rough Set for Multi-criteria Classification Problems

LI Yan JIN Yong-fei WU Ting-ting GUO Na-na YU Qun
(Key Lab. of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding 071002, China)

Abstract In the framework of dominance-based rough set approach (DRSA), dominance relations are used to handle preference ordered attributes contained in data and these attributes are also called as criteria. DRSA has been widely used in multi-criteria decision-making problems. In real applications, however, due to the variations of attribute set and object set, the information systems are often updated from time to time. Under such dynamic environment, the approximation sets in DRSA are required to be updated correspondingly for their future use in feature reduction, rule extraction, and finally in decision-making. In this paper, focusing on multi-criteria classification problems, we developed incremental methods to update set approximations when an object is inserted or deleted. The updating principles in difference cases were discussed and related theoretical results were given with detailed proofs. Two incremental algorithms, DRSA1 and DRSA2, were proposed to update approximations sets when an object is deleted or inserted respectively. Illustrative examples were also given to support the effectiveness of the proposed incremental methods. The experimental results on UCI data sets demonstrate the obvious improvement for non-incremental method (classic DRSA) in terms of efficiency and scalability by using the incremental approach.

Keywords Dominance relation-based rough set, Multi-criteria classification, Information system, Approximations, Incremental updating

1 引言

Pawlak提出的粗糙集理论(RST)[1]是用于处理数据的不确定性和模糊性的数学工具。基于等价关系,所定义的上、下近似集可用于描述论域中不确定的概念,并进一步用于属性约简以及生成决策规则。RST中的一个核心问题就是知识约简,这也是数据挖掘中最重要的问题之一。在这方面许多学者做了广泛的研究并取得了一定的成果[2-7]。然而,在一些实际问题中,特别是在多准则决策分析中,有一些属性值是有序的,如"得分"属性值分为"高、中、低"。这种类型的属性(也称为准则)经常用于评价论域中的对象,例如根据各科的成绩来对一个学生进行评价,在这种情况下,就不能忽视属性值的偏好性。为了解决这个问题,Greco等人[8-10]首次提出了基于优势关系粗糙集方法(DRSA)的理论框架,它用优势关

到稿日期:2016-05-16 返修日期:2016-09-23 本文受国家自然科学基金(61170040,61473111),河北省自然科学基金(F2014201100,A2014201003)资助。

李 艳(1976—),女,博士,教授,CCF会员,主要研究方向为机器学习、Rough集理论、计算智能,E-mail:ly@hbu.cn;靳永飞(1990—),女,硕士生,主要研究方向为Rough集理论、机器学习,E-mail:742215164@qq.com;吴婷婷(1988—),女,硕士生,主要研究方向为粒计算与知识发现;郭娜娜(1992—),女,硕士生,主要研究方向为粒计算与知识发现;于 群(1993—),女,硕士生,主要研究方向为粒计算与知识发现。

系代替等价关系来描述属性值的偏好顺序，并在这个框架下特别是在多准则决策问题中进行了一些研究[11-13]。到目前为止，DRSA 已被广泛应用于许多实际问题，如水质分析[14]、IT 企业评价[15]、客户行为识别[16]、审计风险判断[17]和群决策制定[18]。

在许多实际应用中，一些收集到的数据信息是动态变化的。动态信息系统中的更新主要包括增加/删除属性（属性集）、对象（对象集）以及属性值。传统 DRSA 不考虑信息系统或决策表的这些变化，每当有属性或对象变化时必须重新计算近似集，这是非常耗时的。

在 RST 及它的扩展文献中，针对属性集和对象集的变化提出了增量更新近似集、属性约简集及决策规则的有效算法。当属性集变化时，李天瑞等人在信息系统删除或增加多个属性的情况下提出了基于特征关系的增量更新近似集的方法[19,20]。文献[21]设计了一种属性添加策略及相应的计算方法来更新基于信息熵的属性约简。考虑到属性的偏好信息，李少勇等人提出了 DRSA 下的增量算法来更新近似集[22]。当属性值随时间变化时，陈红梅等人给出了不完备有序决策系统中关于属性值粗化和细化时的动态更新近似原理[23,24]，并在基于变精度限制容差关系的粗糙集下进一步扩展了这项研究[25]。针对属性值变化，李天瑞在文献[26,27]中讨论了快速更新近似集的方法。在集值的信息系统中，当属性动态变化时，文献[28]提出的增量方法通过更新关系矩阵来更新近似集，文献[29]在属性集值下进一步考虑了偏好顺序并在属性泛化下研究了近似集的动态维护。

信息系统的另一种变化发生在对象集上。对于规则的更新，当增加或者删除一个对象时，Shan 和 Ziarko 基于差别矩阵增量地改进了分类规则[30]。将一个新的样例添加到决策表时，Band 和 Zeungnam 提出了一种增量算法从决策表中找到一组最小的规则[31]。文献[32]提出了更新决策矩阵和规则的方法。为了有效地更新属性约简及近似集，当添加新的对象到决策信息系统时，Hu 等人构建了基于初等集的增量属性约简算法[33]。文献[34-36]也讨论了不同的增量属性约简算法。将一组对象添加到给定的决策表时，Liang 等人提出了一种基于信息熵的增量特征选择算法[37]。文献[38]针对动态数据挖掘，在邻域粗糙集下提出了一种增量更新近似集的方法。基于变精度粗糙集（VPRS）模型，Lui 等人介绍了感兴趣的知识的概念并提出了当对象集随时间变化时诱导感兴趣的知识的增量方法[39,40]。陈红梅等人就增加或删除对象提出了基于 VPRS 模型的增量更新近似集的算法[41]。当增加或删除对象时，文献[42]同时考虑不同类型的属性并提出了复合信息系统中基于复合矩阵的对近似集、正域、负域及边界域进行增量计算的方法。在有序集值信息系统中，文献[43]分析了对象集变化时计算近似集的更新机制并提出了析取/合取信息系统中更新近似集的增量算法。在 DRSA 的框架下，李少勇等人讨论了上下联合的更新算法[44]。

上述提到的研究明显地减少了更新近似集、属性约简及规则提取的计算量，但是大部分只能处理符号值属性。文献[22,44]提出的基于优势关系的方法在属性值中能够处理具有偏好顺序的信息，因此这些方法均适用于多准则分类问题。然而，在这些研究中，条件属性与决策属性都要求是单调的，即所有属性值都存在偏好顺序，这不满足多准则分类问题的特点。在实际应用中，一些数据只有条件属性是单调的而决策属性是符号值，这类问题被认为是多准则分类问题。例如，UCI 数据库中的 Iris 数据，对象（花）根据萼片和花瓣的宽度和长度分为 3 种类型。Chen 和 Li 等人就这类问题讨论了属性约简和规则提取方法，并提出了正域约简[45,46]的概念。本文侧重于多准则分类问题，并考虑信息系统或决策系统中对象的变化。在 DRSA 的框架内，当论域中增加或删除一个对象时，提出一种更新近似集的增量更新原则。当对象集有任何变化时，优势类和近似集无需重新计算，可以得到有效的更新。本文的理论结果与文献[41]中的 VPRS 方法有本质的不同，这是由于条件属性集的优势类构成一个覆盖而非论域的划分；此外也与文献[44]的工作相区别，本文更新的是由符号决策属性产生的等价类的上、下近似，而非有序决策类的上、下联合。相应地，我们得到的更新原则形式简单且与文献[44]完全不同，能够用于多准则分类问题。基于这些更新原则，提出了相应的增量算法并给出了说明性的算例及实验结果来支持理论结果。

2 基本概念

作为预备知识，本节介绍了基于优势关系的几个概念，包括目标信息系统，优势/劣势关系/类，上、下近似集，正域和正域约简。

定义 1（目标信息系统） 四元组 $S=(U,A,V,f)$ 是一个目标信息系统，其中 $U=\{x_1,x_2,\cdots,x_n\}$ 是对象的一个非空集合，称为论域；$A=C\cup D, C\cap D=\emptyset$ 是一个有限属性集，其中 C 为条件属性集，D 为决策属性集；V 是属性值的集合；$f: U\times A\rightarrow V$，指在每个属性上对 U 中每个对象给出一个 V 中的值，即 $\forall a\in A, x\in U, f(x,a)\in V$。

在实际应用中，一个信息系统通常用一个二维表来表示，其中行表示论域中的对象，列表示描述对象的属性，每个条目是对应每个对象的属性值。如果决策属性是符号值，那么这些属性基于等价关系构成论域的一个划分（记作 U/D），U/D 中的每个集合称作 U 中的一个目标概念或者决策类。

定义 2（优势/劣势关系） 设 $S=(U,A,V,f)$ 为一个目标信息系统，$B\subseteq A$，记

$$R_B^{\leqslant}=\{(x_i,x_j)\in U\times U: f(x_i,a)\leqslant f(x_j,a), \forall a\in B\}$$

$$R_B^{\geqslant}=\{(x_i,x_j)\in U\times U: f(x_i,a)\geqslant f_l(x_j,a), \forall a\in B\}$$

$R_B^{\leqslant}, R_B^{\geqslant}$ 分别称为信息系统上的优势关系、劣势关系。根据定义 2，$[x_i]_B^{\leqslant}=\{x_j\in U:(x_i,x_j)\in R_B^{\leqslant}\}=\{x_j\in U: f(x_i,a)\leqslant f(x_j,a), \forall a_l\in B\}$ 表示 x_i 关于属性集 B 的优势类，它包含优于 x_i 的所有对象。对应地，可以定义论域中每个对象的劣势类。

定义 3（基于优势关系的上/下近似） 对于每个概念 $X\subseteq U$，定义 X 关于优势关系 R 的上近似与下近似分别为：

$$\overline{R}_B^{\leqslant}(X)=\{x_i\in U:[x_i]_B^{\leqslant}\cap X\neq\emptyset\},$$

$$\underline{R}_B^{\leqslant}(X)=\{x_i\in U:[x_i]_B^{\leqslant}\subseteq X\}$$

类似地，也可以定义 X 关于劣势关系的近似集。

定义 4（正域及基于优势关系的正域约简） 设 $S=(U,A,V,f)$ 为一个目标信息系统，条件属性集 C 关于决策属性集 D 的相对正域定义为 $POS_C(D)=\bigcup_{X\in U/D}\underline{R}_C^{\leqslant}(X)$。对于 $B\subseteq C$，如果 $POS_B(D)=POS_C(D)$，那么 B 被称为一个正域协调集；如果 B 的任何一个子集都不是协调集，那么 B 就是条件

属性集C的一个正域约简。

根据上述定义，近似集用来描述肯定或可能属于给定目标概念X的对象集，正域是指能够精确分类到划分U/D中的任何一个概念中所有对象的集合。很显然，这些概念对于做决策非常重要。在文献[39,40]中，正域约简已被用于多准则分类问题，但是只在条件属性上引入了优势/劣势关系。在动态环境中，正域及正域约简的计算取决于近似集的更新。因此，计算改变后的近似集对于进一步得到新的属性约简是必不可少的。

3 近似集的更新原则

设$S=(U,A,V,f)$，$U=\{x_1,x_2,\cdots,x_n\}$，$A=C\cup D$。在一个多准则分类问题中，在条件属性集C上定义一个优势关系R来刻画属性值之间的偏好关系。U/R构成论域的一个覆盖，而U/D形成论域的一个划分。

考虑对象集随时间变化而属性集保持不变的动态环境，将讨论动态信息系统中计算近似集的增量式更新原则及算法。其基本思想是假定原来的优势类和近似集是已知的，并且在以后的更新过程中可以对它们进行重新利用。当在原先的论域U中删除或插入一个对象$x\in U$时，考虑更新任何给定的决策类或目标概念$X\in U/D$的上、下近似集。为简便，在以后的章节中记一个对象原先的优势类为$[x]_R^{\leqslant}$，上、下近似集分别为$\overline{R}_C^{\leqslant}(X)$、$\underline{R}_C^{\leqslant}(X)$；更新后的论域，决策类，优势类，上、下近似集分别记作U'，X'，$[x]'^{\leqslant}_R$，$\overline{R}_C^{\leqslant}(X')$，$\underline{R}_C^{\leqslant}(X')$。

3.1 当从论域中删除一个对象x时更新X的近似集

本节讨论当从论域U中删除一个对象时，给定决策类的上、下近似集的更新规则。其主要思想是检查删除的对象与给定的决策类之间的关系，并根据不同情况得到不同的更新规则。

3.1.1 更新决策类X的下近似

命题1 如果$x\notin X$，那么$\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)\cup\{x_j\mid x_j\notin \underline{R}_C^{\leqslant}(X),([x_j]_R^{\leqslant}-\{x\})\subseteq X\}$。

证明：$x\notin X$即$X'=X$。考虑删除的对象x，由于$x\notin X$且$x\in[x]_R^{\leqslant}$，因此有$[x]_R^{\leqslant}\not\subset X$。

对于任意的$x_j\ (x_j\neq x)$及$[x_j]_R^{\leqslant}\subseteq X$，有$x\notin X\Rightarrow x\notin [x_j]_R^{\leqslant}$，所以更新后的优势类为$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\subseteq X$。这表明在原先下近似$\underline{R}_C^{\leqslant}(X)$中的对象仍在更新后的下近似$\underline{R}_C^{\leqslant}(X')$中。

对于任意$x_j\ (x_j\neq x)$及$[x_j]_R^{\leqslant}\not\subset X$，如果$x\in[x_j]_R^{\leqslant}$，$[x_j]_R^{\leqslant}-\{x\}\subseteq X$，那么$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}-\{x\}\subseteq X$。在这种情况下，下近似需要加入这样的$x_j$而得到更新。

总之，只需要检查那些不属于$\underline{R}_C^{\leqslant}(X)$的对象及包含删除对象$x$的优势类。

命题2 如果$x\in X$，i. e.，$X'=X-\{x\}$，那么X的下近似的更新公式为：

$$\underline{R}_C^{\leqslant}(X')=\begin{cases}(\underline{R}_C^{\leqslant}(X)-\{x\}), & x\in \underline{R}_C^{\leqslant}(X)\\ \underline{R}_C^{\leqslant}(X), & \text{otherwise}\end{cases}$$

证明：首先考虑删除的对象x，如果$x\in\underline{R}_C^{\leqslant}(X)$，$X$更新后的下近似应计算为$\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)-\{x\}$。反之，如果$[x]_R^{\leqslant}\not\subset X$，则$\underline{R}_C^{\leqslant}(X)$保持不变。

对于$\forall x_j\in U(x_j\neq x)$，如果$[x_j]_R^{\leqslant}\subseteq X$，则有$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}-\{x\}\subseteq X-\{x\}=X'$；如果$[x_j]_R^{\leqslant}\not\subset X$，那么至少存在一个对象$x_k\in[x_j]_R^{\leqslant}$，使得$x_k\notin X$。因$x\in X$，故有$x_k\neq x$，从而有$[x_j]_R^{\leqslant}\not\subset X-\{x\}=X'$。

因此，对于任意对象$x_j\neq x$，$x_j\in\underline{R}_C^{\leqslant}(X')$当且仅当$x_j\in\underline{R}_C^{\leqslant}(X)$。

3.1.2 更新决策类X的上近似

命题3 如果$x\notin X$，i. e.，$X'=X$，那么X的上近似的更新公式为：

$$\overline{R}_C^{\leqslant}(X')=\begin{cases}(\overline{R}_C^{\leqslant}(X)-\{x\}), & x\in \overline{R}_C^{\leqslant}(X)\\ \overline{R}_C^{\leqslant}(X), & x\notin \overline{R}_C^{\leqslant}(X)\end{cases}$$

证明：考虑删除的对象x，如果$x\in\overline{R}_C^{\leqslant}(X)$，即$[x]_R^{\leqslant}\cap\overline{R}_C^{\leqslant}(X)\neq\emptyset$，需要从上近似中移除$x$，即$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)-\{x\}$。

对于任意对象$x_j\in\overline{R}_C^{\leqslant}(X)(x_j\neq x)$，有$[x_j]_R^{\leqslant}\cap\overline{R}_C^{\leqslant}(X)\neq\emptyset$。由于$x\notin X$，一定存在$x_k(\neq x)\in[x_j]_R^{\leqslant}\cap X$，因此$[x_j]'^{\leqslant}_R\cap X=([x_j]_R^{\leqslant}-\{x\})\cap X\neq\emptyset$，即$x_j\in\overline{R}_C^{\leqslant}(X')$。

对于任意对象$x_j\notin\overline{R}_C^{\leqslant}(X)\ (x_j\neq x)$，有$[x_j]_R^{\leqslant}\cap\overline{R}_C^{\leqslant}(X)=\emptyset$。明显地，$[x_j]'^{\leqslant}_R\cap X=([x_j]_R^{\leqslant}-\{x\})\cap X=\emptyset$，即$x_j\notin\overline{R}_C^{\leqslant}(X)\Rightarrow x_j\notin\overline{R}_C^{\leqslant}(X')$。

命题4 如果$x\in X$，i. e.，$X'=X-\{x\}$，那么上近似更新为$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)-\{x\}-\{x_j\mid[x_j]_R^{\leqslant}\cap X=\{x\}\}$。

证明：证明过程与之前的命题的证明类似，因此在这里省略。

3.1.3 增量DRSA1：删除一个对象时的增量算法

根据命题1—命题4的结果，给出基于原始划分U/R、近似集$\underline{R}_C^{\leqslant}(X)$和$\overline{R}_C^{\leqslant}(X)$、条件属性集$C$、对于每个$a\in C$的属性值$V_a$及删除的对象$x\in X$的增量算法。

输入：$U/R=\{[x_1]_R^{\leqslant},[x_2]_R^{\leqslant},\cdots,[x_n]_R^{\leqslant}\}$，$1\leqslant n\leqslant|U|$，$\underline{R}_C^{\leqslant}(X)$，$\overline{R}_C^{\leqslant}(X)$，$V_a$，$C$，$x$

输出：$\underline{R}_C^{\leqslant}(X')$，$\overline{R}_C^{\leqslant}(X')$

Step1 判定$[x_j]_R^{\leqslant}$是否为空，更新优势集和V_a：

若$x\in[x_j]_R^{\leqslant}$，则$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}-\{x\}$；若$[x_j]_R^{\leqslant}=\emptyset$，更新$V_a$；

Step2 判定x是否属于X，分两种情况更新X的近似集：

(1)$x\notin X$

若$[x_j]_R^{\leqslant}-\{x\}\subseteq X$，则$\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X')\cup\{x_j\}$；

若$x\in\overline{R}_C^{\leqslant}(X)$，则$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)-\{x\}$。

(2)$x\in X$，$X'=X-\{x\}$

若$[x]_R^{\leqslant}\subseteq X$，则$\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)-\{x\}$；

若$x\in[x_j]_R^{\leqslant}\cap X$且$[x_j]_R^{\leqslant}\cap X=\{x\}$，则$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)-\{x\}-\{x_j\}$；

若$\exists x_k\neq x\in[x_j]_R^{\leqslant}\cap X$，则$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)-\{x\}$。

Step3 返回更新后的近似集$\underline{R}_C^{\leqslant}(X')$，$\overline{R}_C^{\leqslant}(X')$。

3.2 当插入一个对象x^+时更新X的近似集

当向论域中插入一个新的对象时，更新近似集的计算要比从论域中删除一个当前对象更加复杂。本节将讨论不同情况下的更新规则。

命题5 优势类的更新规则。

(I)如果$x_j\in[x^+]_R^{\leqslant}\Leftrightarrow x^+\in[x_j]'^{\leqslant}_R$，那么$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$；

(II)如果$x_j\notin[x^+]_R^{\leqslant}\Leftrightarrow x^+\notin[x_j]_R^{\leqslant}$，那么$[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}$。

3.2.1 更新决策类 $X'\in U'/D$ 的下近似

命题 6 如果 $x^+\notin X'$，即 $X'=X$，那么 X 的下近似的更新规则为 $\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)-\{x_{j\in(a)}\}$，其中 $\{x_{j\in(a)}\}=\{x_j\mid x_j\in U,x^+\in[x_j]_R^{\leqslant}\}$。

证明：首先计算出新的优势类 $[x^+]_R^{\leqslant}$，其他的优势类根据命题 5 进行更新：1）如果 x^+ 优于 x_j，即 $x_j\in[x^+]_R^{\leqslant}\Leftrightarrow x^+\in[x_j]'^{\leqslant}_R$，那么 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$，因为 $x^+\notin X'$，从而有 $[x_j]'^{\leqslant}_R\not\subset X$；2）若 $x_j\notin[x^+]_R^{\leqslant}\Leftrightarrow x^+\notin[x_j]_R^{\leqslant}$，那么检查 $[x^+]_R^{\leqslant}$ 及其他包含在原先下近似 $\underline{R}_C^{\leqslant}(X)$ 中的优势类。

对于每个 $x_j\in U'$，如果 $x_j=x^+$，由于 $x^+\notin X'$，即 $X'=X$，则有 $[x_j]_R^{\leqslant}=[x^+]_R^{\leqslant}\not\subset X$；如果 $[x_j]_R^{\leqslant}\subseteq X$ 且满足 1)，那么 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}\not\subset X$，$\underline{R}_C^{\leqslant}(X')$ 应更新为 $\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)-\{x_j\}$；如果 $[x_j]_R^{\leqslant}\subseteq X$ 且满足 2)，那么 $[x_j]'^{\leqslant}_R$ 仍包含在 X' 中且 $\underline{R}_C^{\leqslant}(X')$ 保持不变。

总之，可以得到命题 6 中的更新规则。

命题 7 如果 $x^+\in X'$，即 $X'=X\cup\{x^+\}$，那么对于每个 $x_j\in U'$，X' 的下近似的更新规则为：

$$\underline{R}_C^{\leqslant}(X')=\begin{cases}\underline{R}_C^{\leqslant}(X)\cup\{x^+\}, & [x^+]_R^{\leqslant}\subseteq X'\\ \underline{R}_C^{\leqslant}(X), & \text{otherwise}\end{cases}$$

证明：如果 $[x^+]_R^{\leqslant}\subseteq X'$，更新 $\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)\cup\{x^+\}$。

对于任何其他对象 $x_j\in U$，$x_j\in\underline{R}_C^{\leqslant}(X)$，即 $[x_j]_R^{\leqslant}\subseteq X$。如果 $x^+\in[x_j]'^{\leqslant}_R$，根据命题 5 有 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$，由于 $x^+\in X'$ 且 $[x_j]_R^{\leqslant}\subseteq X$ 及 $[x_j]'^{\leqslant}_R\subseteq X'$，在这种情况下，仍有 $x_j\in\underline{R}_C^{\leqslant}(X')$。很容易证明在其他情况下下近似保持不变。

3.2.2 更新决策类 $X'\in U'/D$ 的上近似

命题 8 如果 $x^+\notin X'$，即 $X'=X$，上近似的更新规则为：

$$\overline{R}_C^{\leqslant}(X')=\begin{cases}\overline{R}_C^{\leqslant}(X)\cup\{x^+\}, & [x^+]_R^{\leqslant}\cap X'\neq\emptyset\\ \overline{R}_C^{\leqslant}(X), & \text{otherwise}\end{cases}$$

证明：如果 $[x^+]_R^{\leqslant}\cap X'\neq\emptyset$，显然 $\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)\cup\{x^+\}$ 成立。

对于任意 $x_j\in U$，存在两种情况：

1）$x_j\in\overline{R}_A^{\leqslant}(X)$，即 $[x_j]_R^{\leqslant}\cap X\neq\emptyset$。由于 $[x_j]'^{\leqslant}_R\supseteq[x_j]_R^{\leqslant}\Rightarrow[x_j]'^{\leqslant}_R\cap X\neq\emptyset$，因此有 $x_j\in\overline{R}_C^{\leqslant}(X')$。

2）$x_j\notin\overline{R}_A^{\leqslant}(X)$，即 $[x_j]_R^{\leqslant}\cap X=\emptyset$。因为 $x^+\notin X'$ 且 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$ 或 $[x_j]_R^{\leqslant}$，$[x_j]'^{\leqslant}_R\cap X=\emptyset$ 成立。此时，$x_j\notin\overline{R}_C^{\leqslant}(X')$。

因此，只需要计算插入对象的优势类 $[x^+]_R^{\leqslant}$ 及检查其与 X 是否有非空交集。

命题 9 如果 $x^+\in X'$，即 $X'=X\cup\{x^+\}$，则上近似的更新规则为 $\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)\cup\{x^+\}\cup\{x_j\mid x_j\notin\overline{R}_C^{\leqslant}(X),x^+\in[x_j]_R^{\leqslant}\}$。

证明：由于 $x^+\in X'$，很明显 $[x^+]_R^{\leqslant}\cap X'\neq\emptyset$。

对于 $x_j\in U$，存在两种情况：

1）与命题 8 的证明类似。

2）$x_j\notin\overline{R}_C^{\leqslant}(X)$，即 $[x_j]_R^{\leqslant}\cap X=\emptyset$。如果 $x^+\in[x_j]_R^{\leqslant}$，那么 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$，由于 $x^+\in X'$，即 $[x_j]_R^{\leqslant}\cap X'\neq\emptyset$，从而有 $x_j\in\overline{R}_C^{\leqslant}(X')$。

总之 $\overline{R}_C^{\leqslant}(X)\cup\{x^+\}\cup\{x_j\mid x_j\notin\overline{R}_C^{\leqslant}(X),x^+\in[x_j]_R^{\leqslant}\}$。

3.2.3 增量 DRSA2：插入一个对象时的增量算法

输入：$U/R=\{[x_1]_R^{\leqslant},[x_2]_R^{\leqslant},\cdots,[x_n]_R^{\leqslant}\}(1\leqslant n\leqslant|U|)\underline{R}_C^{\leqslant}(X),\overline{R}_C^{\leqslant}(X),V_a,C,x$

输出：$\underline{R}_C^{\leqslant}(X'),\overline{R}_C^{\leqslant}(X')$

Step1 更新 x_j 的优势类：

若 $x^+\in[x_j]_R^{\leqslant}$，则 $[x_j]'^{\leqslant}_R=[x_j]_R^{\leqslant}\cup\{x^+\}$；

Step2 更新 X' 的下近似集，分两种情况：

(1) $x^+\notin X',X'=X$

若 $[x_j]_R^{\leqslant}\subseteq X$ 且 $x^+\in[x_j]'^{\leqslant}_R$，则 $\underline{R}_C^{\leqslant}(X')=\underline{R}_C^{\leqslant}(X)-\{x_j\}$；

(2) 若 $[x^+]_R^{\leqslant}\subseteq X'$，则 $\underline{R}_C^{\leqslant}(X)\cup\{x^+\}$；

Step3 更新 X' 的上近似集：

若 $x^+\notin\overline{R}_C^{\leqslant}(X),x^+\in[x_j]'^{\leqslant}_R$，则 $\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)\cup\{x_j\}$；

Step4 返回更新后的近似集 $\underline{R}_C^{\leqslant}(X'),\overline{R}_C^{\leqslant}(X')$。

4 数值算例

本节通过一些简单的例子来说明在决策系统中删除或插入 1 个对象时所提增量更新近似集方法的有效性。基于优势关系的 3 个决策系统如表 1 所列，其中左边是原先具有 6 个对象和 3 个属性的决策系统；中间为插入对象 x_7 的决策系统；右侧是插入对象 x_8 的决策系统，d 是决策属性。

表 1 基于优势关系的决策系统

U	a_1	a_2	a_3	d	U	a_1	a_2	a_3	d	U	a_1	a_2	a_3	d
x_1	1	2	1	3	x_1	1	2	1	3	x_1	1	2	1	3
x_2	3	2	2	2	x_2	3	2	2	2	x_2	3	2	2	2
x_3	1	1	2	1	x_3	1	1	2	1	x_3	1	1	2	1
x_4	2	1	3	2	x_4	2	1	3	2	x_4	2	1	3	2
x_5	3	3	2	3	x_5	3	3	2	3	x_5	3	3	2	3
x_6	3	2	3	1	x_6	3	2	3	1	x_6	3	2	3	1
					x_7	2	1	2	1	x_8	1	3	2	3

属性集 $C=\{a_1,a_2,a_3\}$ 上的优势关系为：

$R_C^{\leqslant}=\{(x_1,x_1),(x_1,x_2),(x_1,x_5),(x_1,x_6),(x_2,x_2),(x_2,x_5),(x_2,x_6),(x_3,x_2),(x_3,x_3),(x_3,x_4),(x_3,x_5),(x_3,x_6),(x_4,x_4),(x_4,x_6),(x_5,x_5),(x_6,x_6)\}$

对应地，优势类为：

$[x_1]_C^{\leqslant}=\{x_1,x_2,x_5,x_6\}$；$[x_2]_C^{\leqslant}=\{x_2,x_5,x_6\}$；$[x_3]_C^{\leqslant}=\{x_2,x_3,x_4,x_5,x_6\}$；$[x_4]_C^{\leqslant}=\{x_4,x_6\}$；$[x_5]_C^{\leqslant}=\{x_5\}$；$[x_6]_C^{\leqslant}=\{x_6\}$。

设 $X=\{x_1,x_5\}$ 是一个基于决策属性 d 的决策类，其中 d 是一个不具有偏好顺序的常规属性，则可以基于优势类得到 X 的近似集：$\underline{R}_C^{\leqslant}(X)=\{x_5\}$；$\overline{R}_C^{\leqslant}(X)=\{x_1,x_2,x_3,x_5\}$。

4.1 删除一个对象时更新近似集

例 1 从左侧表中删除对象 x_6，从而有 $U'=\{x_1,x_2,\cdots,x_5\}$，$X'=X=\{x_1,x_5\}$。

(1)传统方法

首先通过成对比较 U' 中的对象，优势类的集合应更新为 $E'=\{\{x_1,x_2,x_5\},\{x_2,x_5\},\{x_2,x_3,x_4,x_5\},\{x_4\},\{x_5\}\}$。这部分的计算复杂度为 $(n-1)(n-2)*m/2$（其中 n 表示行数，m 表示条件属性的个数），即计算复杂度是 $O(mn^2)$；然后检查每个优势类与 X' 的交集，这一步的计算复杂度为 $O(n)$；最后得到更新后的近似集 $\underline{R}_C^{\leqslant}(X')=\{x_5\}$，$\overline{R}_C^{\leqslant}(X')=\overline{R}_C^{\leqslant}(X)=\{x_1,x_2,x_3,x_5\}$。总体而言，计算复杂度为 $O(mn^2)$。

(2)增量更新方法

根据命题 1 的结论，有$\underline{R}_{\bar{C}}^{\leqslant}(X')=\underline{R}_{\bar{C}}^{\leqslant}(X)=\{x_5\}$，只需计算 4 个优势类与 X 的交集，最坏的情况下的计算复杂度为 $O(n^2)$。

另一方面，根据命题 3，因被删除对象的优势类$[x_6]_{\bar{C}}^{\leqslant}=\{x_6\}$不包含在原先的上近似$\overline{R}_{\bar{C}}^{\leqslant}(X)$中，故可直接得到更新后的上近似$\overline{R}_{\bar{C}}^{\leqslant}(X')=\overline{R}_{\bar{C}}^{\leqslant}(X)=\{x_1,x_2,x_3,x_5\}$，即更新上近似无需其他计算。

例 2　删除对象 x_5，有 $U'=\{x_1,x_2,x_3,x_4,x_6\}$，$X'=X-\{x_5\}=\{x_1\}$，满足命题 2 和命题 4 的条件。

(1)传统方法

与例 1 类似，传统方法的计算复杂度为 $O(mn^2)$。优势类的集合更新为 $E'=\{\{x_1,x_2,x_6\},\{x_2,x_6\},\{x_2,x_3,x_4,x_6\},\{x_4,x_6\},\{x_6\}\}$，更新后的近似集为$\underline{R}_{\bar{C}}^{\leqslant}(X')=\emptyset$，$\overline{R}_{\bar{C}}^{\leqslant}(X')=\{x_1\}$。

(2)增量更新方法

根据命题 2，由于被删除的对象 x_5 属于 X 的下近似，则有$\underline{R}_{\bar{C}}^{\leqslant}(X')=(\underline{R}_{\bar{C}}^{\leqslant}(X)-\{x_5\})=\emptyset$，这部分无需其他额外的计算。

根据命题 4，计算每个优势类与 X 的交集，有$[x_2]_{\bar{C}}^{\leqslant}\cap X=\{x_5\}$，$[x_3]_{\bar{C}}^{\leqslant}\cap X=\{x_5\}$，从而有$\overline{R}_{\bar{C}}^{\leqslant}(X_1')=\overline{R}_{\bar{C}}^{\leqslant}(X)-\{x_5\}-\{x_2\}\cup\{x_3\}=\{x_1\}$。

4.2　插入一个对象时更新近似集

例 1　插入对象 $x_7=\{2,1,2,1\}$，有 $U'=\{x_1,x_2,\cdots,x_6,x_7\}$，$X'=X=\{x_1,x_5\}$，满足命题 6 和命题 8 的条件。

(1)传统方法

与前面的章节类似，首先更新后的优势类的集合为 $E'=\{\{x_1,x_2,x_6\},\{x_2,x_5,x_6\},\{x_2,x_3,x_4,x_5,x_6,x_7\},\{x_4,x_6\},\{x_5\},\{x_6\},\{x_2,x_4,x_5,x_6,x_7\}\}$，则可计算出每个优势类与 X' 的交集，最后便可得到 X' 的上、下近似$\underline{R}_{\bar{C}}^{\leqslant}(X')=\underline{R}_{\bar{C}}^{\leqslant}(X)=\{x_5\}$，$\overline{R}_{\bar{C}}^{\leqslant}(X')=\{x_1,x_2,x_3,x_5,x_7\}$。计算复杂度仍然为 $O(mn^2)$。

(2)增量更新方法

根据命题 6，首先插入的对象 $x_7=\{2,1,2,1\}$的优势类与劣势类更新为$[x_7]_{\bar{C}}^{\leqslant}=\{x_2,x_4,x_5,x_6,x_7\}$，$[x_7]_{\bar{C}}^{\geqslant}=\{x_3\}$，且计算复杂度为 $O(m(n-1))$；其次更新劣势类$[x_7]_{\bar{C}}^{\geqslant}$ 中对象的优势类，有$[x_3]_{\bar{C}}'^{\leqslant}=[x_3]_{\bar{C}}^{\leqslant}\cup\{x_7\}$(将 x_7 添加到原来的优势类中)，其他对象的优势类保持不变，这一步的计算复杂度为$|[x_7]_{\bar{C}}^{\geqslant}|$；然后检查这些优势类是否包含在 X 中，有$[x_3]_{\bar{C}}^{\leqslant}\not\subset X$；最后更新后的下近似为$\underline{R}_{\bar{C}}^{\leqslant}(X')=\underline{R}_{\bar{C}}^{\leqslant}(X)=\{x_5\}$，且总体计算复杂度为 $O(mn)$。

上近似可根据命题 8 进行更新，与下近似的更新过程类似，首先需要更新每个对象的优势类，然后只需检查 X' 与$[x_7]_{\bar{C}}^{\leqslant}=\{x_2,x_4,x_5,x_6,x_7\}$的交集。在这个例子中，它们具有非空交集，于是有$\overline{R}_{\bar{C}}^{\leqslant}(X')=\overline{R}_{\bar{C}}^{\leqslant}(X)\cup\{x_7\}=\{x_1,x_2,x_3,x_5,x_7\}$。最坏情况下的计算复杂度为 $O(n^2)$。

例 2　插入对象 $x_8=\{1,3,2,3\}$，有 $U'=\{x_1,x_2,\cdots,x_6,x_8\}$，$X'=X\cup\{x_8\}=\{x_1,x_5,x_8\}$，满足命题 7 和命题 9 的条件。

(1)传统方法

与例 1 类似，有$\underline{R}_{\bar{C}}^{\leqslant}(X')=\underline{R}_{\bar{C}}^{\leqslant}(X)\cup\{x_8\}=\{x_5,x_8\}$，$\overline{R}_{\bar{C}}^{\leqslant}(X')=\{x_1,x_2,x_3,x_5,x_8\}$。

(2)增量更新方法

首先根据命题 5 计算出 x_8 的优势类和劣势类$[x_8]_{\bar{C}}^{\leqslant}=\{x_5,x_8\}$，$[x_8]_{\bar{C}}^{\geqslant}=\{x_1,x_3\}$，那么$[x_8]_{\bar{C}}^{\geqslant}$ 中对象的优势类更新为$[x_1]_{\bar{C}}'^{\leqslant}=[x_1]_{\bar{C}}^{\leqslant}\cup\{x_8\}$，$[x_3]_{\bar{C}}'^{\leqslant}=[x_3]_{\bar{C}}^{\leqslant}\cup\{x_8\}$，且其他对象的优势类保持不变；其次根据命题 7，由于 x_8 的优势类包含在 X' 中，因此下近似集更新为$\underline{R}_{\bar{C}}^{\leqslant}(X')=\underline{R}_{\bar{C}}^{\leqslant}(X)\cup\{x_8\}=\{x_5,x_8\}$；最后根据命题 9，$X'$ 的上近似可直接更新为$\overline{R}_{\bar{C}}^{\leqslant}(X')=\overline{R}_{\bar{C}}^{\leqslant}(X)\cup\{x_8\}\cup_{j\in[x_8]_{\bar{C}}^{\geqslant}}\{x_j\}=\{x_1,x_2,x_3,x_5,x_8\}$，这是因为优势类$[x_8]_{\bar{C}}^{\leqslant}$ 和 X' 都包含 x_8，且$[x_8]_{\bar{C}}^{\geqslant}$ 中那些对象的优势类也包含插入的对象 x_8。

5　实验分析

本节为了验证在对象集发生变化时所提增量方法的有效性，从 UCI 数据集[47]中选取 8 个数据集，这些数据集符合本文研究对象的特征：条件属性是具有偏好关系的，而决策属性是符号值属性，所取数据如表 2 所列。

表 2　数据集

Data sets	N#	M#	C#
Abalone	4177	8	29
Car	1728	6	4
Pima	768	8	2
Balance	625	4	3
Bupa	345	6	2
Sonar	208	60	2
Wine	178	13	3
Iris	150	4	3

注：N#为样例个数，M#为属性个数，C#为类别个数。

所有的算法是在 Windows7，Intel(R) Core(TM) i3-4150 CPU @ 3.50GHz 3.50GHz 4.00GB 环境下通过 Matlab (7.0)实现的。该节就两个方面进行实验对比：1)传统基于优势关系的粗糙集方法 DRSA 与增量算法 DRSA1 和 DRSA2 就运行时间进行对比；2)随着数据集的增大，所有算法在运行时间上进行比较。在 2)中，每个数据被分成大小相同的 5 份，分别记作数据 1、数据 2…数据 5。其中数据 1 被用作第一个数据集，数据 1 与数据 2 的并集用作第二个数据集，依次类推，5 个数据的并集看作第五个数据集。

5.1　非增量方法与增量方法之间运行时间的比较

在表 2 的每个数据集中，从原先的数据集中删除/插入的对象都是随机选择的，且分别利用传统 DRSA、增量算法 DRSA1 和 DRSA2 对所有决策类的上、下近似集进行更新。对上述 3 种算法均运行 100 次并取平均作为最终结果。传统基于优势关系的粗糙集方法(传统 DRSA)与增量 DRSA1 所取数据的运行时间如表 3 所列。传统 DRSA 与增量 DRSA1 运行时间的比值(称为增强率)说明删除一个对象时增量方法比传统方法效率提高的程度。

表 3　传统 DRSA 与增量 DRSA1 的运行时间(单位：s))

Data sets	传统 DRSA	增量 DRSA1	增强率
Abalone	121.7644	1.2267	99.2718
Car	7.5265	0.3185	23.6348
Pima	1.3259	0.0975	13.5928
Balance	1.2632	0.0956	13.2189
Bupa	0.3445	0.0421	8.1825
Sonar	0.1484	0.0091	16.3426
Wine	0.0938	0.0051	18.4580
Iris	0.1249	0.0118	10.6224
Average	16.5740	0.2258	25.4155

对于上述所选取的数据集，传统 DRSA 方法与增量 DR-

SA2 方法运行时间的比较如表 4 所列，其中最后一列是两种算法间的时间比率。与表 3 的结果类似，当在原始数据中插入一个对象时，增量 DRSA2 算法相比传统 DRSA 明显表现出更高的效率。

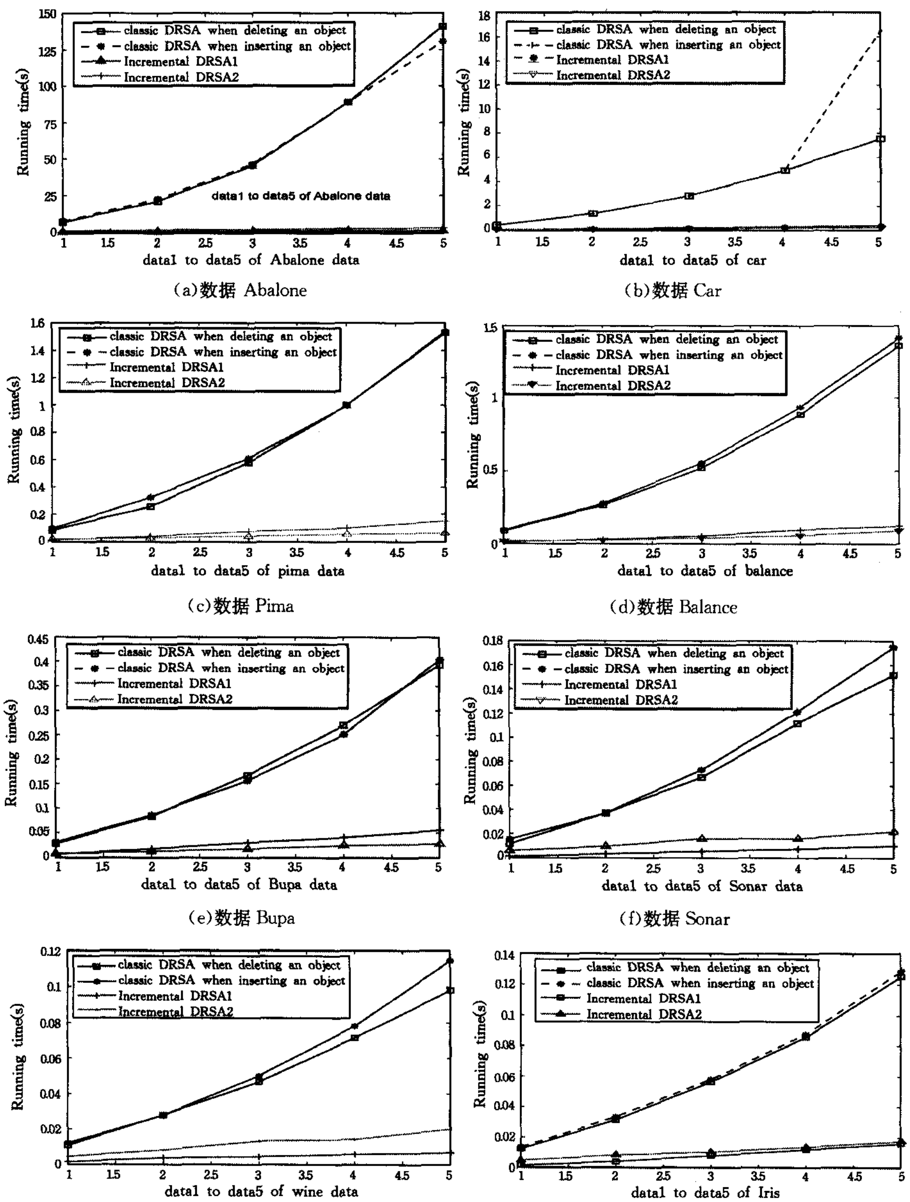
表 4 传统 DRSA 与增量 DRSA2 的运行时间(单位：s)

Data sets	传统 DRSA	增量 DRSA2	增强率
Abalone	119.8284	2.8826	41.5696
Car	7.6262	0.2315	32.9454
Pima	1.3446	0.0624	21.5529
Balance	1.2810	0.0737	17.3733
Bupa	0.3508	0.0252	13.9206
Sonar	0.1529	0.0181	8.4542
Wine	0.0964	0.0228	4.2218
Iris	0.1254	0.0216	5.7956
Average	16.3507	0.4172	18.2292

在表 3 和表 4 中，当数据集增大时(从 Iris 到 Abalone)，通常计算近似集需要的时间会变长，且增量算法与传统 DRSA 相比效率提高得更明显。例如，在两个表中传统 DRSA 与增量算法之间的增强率分别超过 99 和 41。如果比较删除和插入一个对象时的运行时间，传统方法在这两种情况下的运行时间几乎相同，而增量 DRSA1 的运行速度比 DRSA2 更快。这是因为插入一个对象时，在更新优势类(尤其是包含插入的对象)时 DRSA2 需要更多的计算时间。

5.2 数据集不断增大时的运行时间

正如在 5.1 节中所提到的，随着数据集的增大，3 种算法在运行时间上都有所增加，特别是传统 DRSA 最为明显。为了说明所提增量方法的可扩展性，本节探究随着数据集增大时传统 DRSA 与增量算法的计算时间的线性增长趋势。为了实现这一过程，将每个数据随机分成所含对象个数相同的 5 部分，则每个原始数据集通过分别组合第一部分、第二部分…以及最后5 个部分线性增加可形成 5 个数据集(Data1－Data5)。通过所有算法在 Data1－Data5 上的运行，可以明显地观察到计算时间趋势，结果如图 1 所示。图 1(a)－图 1(h)表明传统 DRSA 在每个数据集上从 Data1－Data5 的运行时间明显增加，然而增量算法 DRSA1 和 DRSA2 在数据集变大时运行时间更加稳定，这充分说明了增量方法具有处理大数据集的潜力。



(a)数据 Abalone (b)数据 Car

(c)数据 Pima (d)数据 Balance

(e)数据 Bupa (f)数据 Sonar

(g)数据 Wine (h)数据 Iris

图 1 不同算法在逐渐增大的数据集上运行时间的比较

从 5.1 节及 5.2 节的实验结果可得到如下结论：

(1)在处理动态信息/决策系统时，所提的增量 DRSA1 和 DRSA2 方法比传统 DRSA 方法有效得多。

(2)对于所提的增量方法，就运行时间而言，删除一个对象通常要比插入一个对象成本更低，即 DRSA1 比 DRSA2 运行速度更快，这是因为插入一个对象需要更新所有包含新对象的优势类。

(3)传统 DRSA 的性能会受到数据大小的强烈影响，随数据从 Data1 到 Data5 变化，运行时间呈线性增长，增量算法随数据线性增加时运行时间更加稳定。

结束语 本文主要针对多准则分类问题，提出了增量方法来更新动态信息/决策系统中目标概念的上、下近似集。具体地，考虑了在原数据集中删除或插入一个对象的情况，并在不同情况下提出了相应的近似集更新规则。在粗糙集框架中更新近似集对于动态决策制定至关重要，本文目的在于减少传统 DRSA 在更新近似集过程中的计算时间。给出了更新规则及详细的证明，并基于此提出了两种增量方法 DRSA1 和 DRSA2。通过在 8 个 UCI 数据集上的大量实验，充分证明了当对象集随时间发生变化时所提增量方法在保持知识上的有效性。

需要注意的是，目前主要讨论的是在原信息系统中删除或插入一个对象的情况。当多个对象发生变化时，可将其看作一个对象的多重变化并可重复利用所提的增量方法。进一步研究可能考虑属性集和对象集同时发生变化时的增量方法；当插入一个对象时属性值域的变化也可考虑在内；另一个可能研究的方向是通过结合 DRSA 和变精度粗糙集(VPRS)进一步扩展动态信息系统中的增量算法。

参 考 文 献

[1] Pawlak Z. Rough sets：theoretical aspects of reasoning about data[M]. Kluwer Academic Publishers，Dordrecht，1991

[2] Qian Y H，Liang J Y，Pedrycz W. An efficient accelerator for attribute reduction from incomplete data in rough set framework [J]. Pattern Recognition，2011，44(8)：1658-1670

[3] Liu F，Li T R. Method for Attribute Reduction Based on Rough Sets Boundary Regions[J]. Computer Science，2016，43(3)：242-284(in Chinese)
刘芳，李天瑞. 基于边界域的不完备信息系统属性约简方法[J]. 计算机科学，2016，43(3)：242-284

[4] Blaszczynski J，Greco S，Slowinski R. Inductive discovery of laws using monotonic rules[J]. Engineering Applications of Artificial Intelligence，2012，25(2)：284-294

[5] Abbas A，Liu J. Designing an intelligent recommender system using partial credit model and Bayesian rough set[J]. International Arab Journal of Information Technology，2012，9(2)：179-187

[6] Chang B，Hung H. A study of using RST to create the supplier selection model and decision making rules[J]. Expert Systems with Applications，2010，37(12)：8284-8295

[7] Jiao N. Research on Vertical Segmentation Knowledge Reduction Algorithm Based on Tolerance Rough Set Theory[J]. Computer Science，2016，43(1)：49-52(in Chinese)
焦娜. 相容关系下的分割知识约简算法的研究[J]. 计算机科学，2016，43(1)：49-52

[8] Greco S，Matarazzo B，Slowinski R. Rough sets theory for multicriteria decision analysis [J]. European Journal of Operational Research，2001，129(1)：1-47

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

[10] Greco S，Slowinski R，Matarazzo B，et al. Variable consistency model of dominance-based rough sets approach[C]// Proc. of International Conference on Rough Sets and Current Trends in Computing. Springer-Verlage，2000：170-181

[11] Hu Q H，Yu D R，Guo M Z. Fuzzy preference based rough sets [J]. Information Sciences，2010，180(10)：2003-2022

[12] Hu Q H，Pan W W，Zhang L，et al. Feature selection for monotonic classification[J]. IEEE Transactions on Fuzzy Systems，2012，20(1)：69-81

[13] Hu Q H，Che X J，Zhang L，et al. Rank entropy based decision trees for monotonic classification [J]. IEEE Transactions on Knowledge and Data Engineering，2012，24(11)：2052-2064

[14] Karami J，Alimohammadi A，Seifouri T. Water quality analysis using a variable consistency dominance-based rough set approach[J]. Computers Environment and Urban Systems，2014，43(1)：25-33

[15] Peters G，Poon S. Analyzing IT business values-A dominance based rough sets approach perspective[J]. Expert Systems with Applications，2011，8(9)：11120-11128

[16] Liou J H，Tzeng G H. A dominance-based rough set approach to customer behavior in the airline market[J]. Information Sciences，2010，180(11)：2230-2238

[17] Huang B，Zhuang Y L，Li H X，et al. A Dominance Intuitionistic Fuzzy-Rough Set Approach and its Applications [J]. Applied Mathematical Modeling，2013，37(12/13)：7128-7141

[18] Chakhar S，Saad I. Dominance-based rough set approach for groups in multicriteria classification problems[J]. Decision Support Systems，2012，54(1)：372-380

[19] 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

[20] Liu D，Li T R，Zhang J B. Incremental updating approximations in probabilistic rough sets under the variation of attributes[J]. Knowledge-based Systems，2015，73(81)：81-96

[21] Wang F，Liang J Y，Qian Y H. Attribute reduction：A dimension incremental strategy [J]. Knowledge-Based Systems，2003，39 (6)：95-108

[22] 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

[23] Chen H M，Li T R，Qiao S，et al. A rough set based dynamic maintenance approach for approximations in coarsening and refining attribute values [J]. International Journal of Intelligent Systems，2010，25(10)：1005-1026

[24] Chen H M，Li T R，Ruan D. Maintenance of approximations in incomplete ordered decision systems while attribute values coarsening or refining[J]. Knowledge-Based Systems，2012，31 (7)：140-161

[25] 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，2012，18(5/6)：577-598

[26] Li S Y，Li T R. Incremental update of approximations in dominance-based rough sets approach under the variation of attribute

values[J]. Information Sciences,2015,294(c):348-361

[27] Luo C,Li T R,Chen H M,et al. Fast algorithms for computing rough approximations inset-valued decision systems while updating criteria values[J]. Information Sciences,2015,299(c):221-242

[28] Zhang J B,Li T R,Ruan D,et al. Rough sets based matrix approaches with dynamic attribute variation in set-valued information systems[J]. International Journal of Approximate Reasoning,2012,53(4):620-635

[29] Luo C,Li T R,Chen H M. Dynamic maintenance of approximations in set-valued ordered decision systems under the attribute generalization[J]. Information Sciences,2014,257(2):210-228

[30] Shan N,Ziarko W. Data-based acquisition and incremental modification of classification rules[J]. Computational Intelligence, 1995,11(2):357-370

[31] Bang W C,Zeungnam B. New incremental learning algorithm in the framework of rough set theory[J]. International Journal of Fuzzy Systems,1999,1(1) :25-36

[32] Tong L,An L. Incremental learning of decision rules based on rough set theory[J]. Proc. of World Congress on Intelligent Control and Automation,2002,1(11):420-425

[33] Hu F,Wang G Y,Huang H,et al. Incremental attribute reduction based on elementary sets[M]//Rough Sets,Fuzzy Sets,Data Mining, and Granular Computing. Springer Berlin Heidelberg,2005:185-193

[34] Jing Y G,Li T R,Huang J F,et al. An incremental attribute reduction approach based on knowledge granularity under the attribute generalization[J]. International Journal of Approximate Reasoning,2016,76(6):80-95

[35] Jing Y G,Li T R,Luo C,et al. An incremental approach for attribute reduction based on knowledge granularity[J]. Knowledge-Based Systems,2016,104(1):24-38

[36] Shu W H,Shen H. Incremental feature selection based on rough set in dynamic incomplete data[J]. Pattern Recognition,2014,47(12):3890-3906

[37] Liang J Y,Wang F,Dang C Y,et al. A group incremental approach to feature selection applying rough set technique[J]. IEEE Transactions on Knowledge and Data Engineering,2014, 26(2):294-308

[38] 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

[39] 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

[40] Liu D,Li T R,Ruan D,et al. Incremental learning optimization on knowledge discovery in dynamic business intelligent systems [J]. Journal of Global Optimization,2011,51(2):325-344

[41] 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,2013,25(2):274-284

[42] Zhang J B, Li T R, Chen H M. Composite rough sets for dynamic data mining[J]. Information Sciences,2014,257(2):81-100

[43] Luo C,Li T R,Chen H M,et al. Incremental approaches for updating approximations in set-valued ordered information systems [J]. Knowledge-Based Systems,2013,50(50):218-233

[44] 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

[45] Chen J,Wang G Y,Hu J. Positive domain reduction based on dominance relation in inconstant system[J]. Computer Science, 2008,35(3):216-218(in Chinese)
陈娟,王国胤,胡军. 优势关系下不协调信息系统的正域约简[J]. 计算机科学,2008,35(3):216-218

[46] Li Y,Sun N X,Zhao J,et al. Reductions based on dominance-equivalence relations and rules extraction methods[J]. Computer Science,2011,38(11):220-224(in Chinese)
李艳,孙娜欣,赵津,等. 基于优势-等价关系的几种约简及规则抽取方法[J]. 计算机科学,2011,38(11):220-224

[47] Bache K,Lichma M. UCI Macine Learning Repository[OL]. http://archive. ics. uci. edu/ml

(上接第 49 页)

[12] Goldberg D E. Genetic algorithms in search, optimization and machine leaning[M]. New York: Addison-Wesley Publishing Company,Inc,1989

[13] Yang Ning,Wang Qian. Negative Selection Algorithm Based on Niche Strategy[J]. Computer Science, 2011, 38(10A): 181-184 (in Chinese)
杨宁,王茜. 一种基于小生境策略的阴性选择算法[J]. 计算机科学,2011,38(10A):181-184

[14] Zhu Yu,Han Chang-pei. Improved Genetic Algorithm with Adaptive Convergence Populations[J]. Computer Science, 2012, 39(10):214-217(in Chinese)
朱钰,韩昌配. 一种种群自适应收敛的快速遗传算法[J]. 计算机科学,2012,39(10):214-217

[15] Gu D Q,Liu G F,Yuan J,et al. New algorithm for solving TSP based on dots and lines loop optimization[J]. Journal of PLA University of Science and Technology,2010,11(4):423-427(in Chinese)
顾大权,刘高飞,袁珏,等. 基于点线回路优化求解 TSP 的一个新算法[J]. 解放军理工大学学报,2010,11(4):423-427

[16] Feng X, Ma M Y, Yu H Q. Lake-Energy Optimization Algorithm for Travelling Salesman Problem[J]. Journal of Computer Research and Development,2013,50(9):2015-2027(in Chinese)
冯翔,马美怡,虞慧群. TSP 湖水能量优化算法[J]. 计算机研究与发展,2013,50(9):2015-2027

[17] Zhou Y Q,Huang Z X. Artificial glowworm swarm optimization algorithm for TSP[J]. Control and Decision, 2012, 27(12): 1816-1821(in Chinese)
周永权,黄正新. 求解 TSP 的人工萤火虫群优化算法[J]. 控制与决策,2012,27(12):1816-1821

[18] Zhou Y Q, Huang Z X, Liu H X. Discrete Glowworm Swarm Optimization Algorithm for TSP Problem[J]. Acta Electronic Sinica,2012,40(6):1164-1170(in Chinese)
周永权,黄正新,刘洪霞. 求解 TSP 问题的离散型萤火虫群优化算法[J]. 电子学报,2012,40(6):1164-1170

[19] Rao W Z,Jin C,Huang T Y. Hybrid algorithm of the nearest neighbor and insertion for the traveling salesman problem[J]. Systems Engineering-Theory & Practice, 2011, 31(8): 1419-1428(in Chinese)
饶卫振,金淳,黄英艺. 求解 TSP 问题的最近邻域与插入混合算法[J]. 系统工程理论与实践,2011,31(8):1419-1428