基于联合属性重要度的决策风险最小化属性约简

徐菲菲 毕忠勤 雷景生
(上海电力学院计算机科学与技术学院 上海 200090)

摘 要 经典粗糙集属性约简基本都是保持正域、负域和边界域不变,而决策粗糙集对属性的增减过程不具备单调性,因此不可能同时保持3个区域均不变。在决策粗糙集模型中,作出决策更应该考虑风险最小化原则,因此提出一种改进的风险最小化属性约简方法,在属性的选取过程中同时考虑所选取的属性子集对决策的划分能力,即联合属性重要度以及风险最小化。实验证明所提方法是有效的。

关键词 属性约简,风险最小化,联合属性重要度,决策粗糙集
中图法分类号 TP18 **文献标识码** A

Attribute Reduction Based on Cost Minimization and Significance of Joint Attributes

XU Fei-fei BI Zhong-qin LEI Jing-sheng
(College of Computer Science and Technology,Shanghai University of Electric Power,Shanghai 200090,China)

Abstract The classical rough set attribute reduction is mainly based on maintaining positive region, boundary region and negative region unchanged. In the decision rough set model, the reduction procedure for adding or deleting an attribute is no longer monotonous, so that three regions can not keep all unchanged. In decision theoretic rough set model, decision making should take consideration of minimizing the cost. Therefore, this paper put forward a method for attribute reduction based on minimizing the cost, while considering the classification ability of selected attribute subset to the decision-making, which is named as the significance of joint attributes. Experiments show that our method is effective.

Keywords Attribute reduction, Minimum cost, Significance of joint attributes, Decision rough sets

1 引言

粗糙集理论[1]由波兰数学家Pawlak于1982年提出,是一种处理不确定、不精确数据的数学工具。该理论不需要任何先验知识即可对数据进行处理。近年来,粗糙集理论越来越受到学者的广泛关注,无论在理论上还是应用上都取得了较为重要的成果[2-11]。经典的粗糙集建立在严格的等价关系的基础上,只能处理离散型的数据,不适合处理现实生活中包含噪声或不一致的数据,于是众多学者对粗糙集理论进行了扩展,提出了概率粗糙集[12]、邻域粗糙集[13]等等。

1992年,Yao将贝叶斯风险相关理论引入至概率粗糙集中,提出了决策粗糙集模型(Decision-Theoretic Rough Set Model,DTRS)[14]。该模型将概率粗糙集中的阈值通过贝叶斯决策过程得到,依据一些具体的概念(如风险或开销等),使得阈值的获取具有清晰的语义或解释[15]。在经典的Pawlak粗糙集中,等价类将整个论域分为3个部分:完全属于某个集合的所有等价类构成正域、可能但不完全属于某个集合的所有等价类构成边界域、完全不属于某个集合的所有等价类构成负域。Pawlak粗糙集并没有考虑到决策规则的容错性,完全正确和确定的规则才能进入正域。基于此,在决策粗糙集模型中,Yao等人提出从正域里获取的正规则表示接受某事物(acceptance);从负域里获取的负规则表示拒绝某事物(rejection);落在边界域上的规则表示需要进一步观察,即延迟决策(deferment)。

属性约简是粗糙集理论中的一个重要研究内容。在经典粗糙集中,属性约简是在保持系统整体分类能力不变的情况下,通过删除对决策分类不重要的或无关的属性,使得可以用较少的知识获得与原知识库相同的分类能力,从而大大提高数据处理的效率。而在决策粗糙集模型中,不同于Pawlak粗糙集模型的约简理论,决策粗糙集的约简需要考虑到不同的分类性能[16]。在决策粗糙集理论中,决策区域和决策规则与属性增减之间并不具备单调性,使得在属性约简前后决策表的正域、负域和边界域可能发生不同的变化,如何判断区域变化的好坏就成为评价约简是否合适的重要问题。Li等人[17]提出了一种正域扩大约简方法;Zhang和Miao[20]采用集合区域提出一种新分类区域,分析了3种DTRS分类区域扩大分类正域的机理。这两种方法都源于集合区域。在决策粗糙集约简研究中,另一种则是采用优化决策方案。Jia等人[18]给出了决策粗糙集模型下基于决策风险(损失)最小化的属性约简定义,但该方法并未考虑到条件属性集中各个属性的重要

本文受国家自然科学基金(61272437,61305094),上海市教育发展基金会和上海市教育委员会"晨光计划"(13CG58),上海市自然科学基金(13ZR1417500)资助。
徐菲菲(1983-),女,博士,副教授,主要研究方向为粗糙集、数据挖掘、大数据等,E-mail:xufeifei1983@hotmail.com;毕忠勤(1977-),男,博士,副教授,主要研究方向为大数据、云计算、粗糙集等;雷景生(1966-),男,博士,教授,主要研究方向为数据挖掘、大数据等。

性并非同等重要；Yu[19]等人根据不同属性对决策表的决策分类能力大小不同，给出了决策粗糙集模型下属性重要性的定义，并以属性重要性的大小作为属性约简的启发式信息，给出决策风险最小化的属性约简方法。然而，该方法仅仅考虑单个条件属性相对于决策的分类能力，并未考虑到条件属性之间的相关性。假设两个条件属性各自相对于决策都很重要，联合在一起并不一定能增加对决策的分类能力。基于此，本文给出了改进的基于属性重要度的决策风险最小化属性约简方法，通过联合条件属性子集相对于决策分类能力的重要性对属性进行选取。进一步，为了加快约简的速度，本文采取 bottom-up 的方式，大大提升约简获取的效率。实验证明所提方法是有效的。

2 决策风险最小化的属性约简

信息表（决策表）$M=(U,At=C\cup D,V,f)$是一个四元组，其中论域 $U=\{x_1,x,2,\cdots,x_n\}$，条件属性集为 $C=\{a_1,a_2,\cdots,a_m\}$，$D=\{d\}$为决策属性集。设 $R\subseteq C$，则不可分辨关系为：$IND(R)=\{(x,y)\in U\times U\mid \forall a\in R,f(x,a)=f(y,a)\}$。显然，不可分辨关系是等价关系，$U/R$ 是基于等价关系 R 对 U 的一个划分。包含对象 x 的等价类通常记为$[x]_R$，简写为$[x]$。

设条件属性子集 R 和决策属性集 D 在 U 上导出的划分分别为 X,Y，其中 $X=\{X_1,X_2,\cdots,X_i,\cdots,X_s\}$，$Y=\{Y_1,Y_2,\cdots,Y_j,\cdots,Y_v\}$。状态集合 $\Omega=\{Y_j,Y_j^c\}$表示对象是否属于决策类 Y_j，则等价类 X_i 属于 Y_j 和不属于 Y_j 的条件概率分别按如下计算：

$$P(Y_j\mid X_i)=\frac{|X_i\cap Y_j|}{|X_i|}$$

$$P(Y_j^c\mid X_i)=1-P(Y_j\mid X_i)$$

在决策粗糙集模型中，当对象 x 属于对象子集 S 时，λ_{PP}、λ_{BP} 和 λ_{NP} 分别表示将一个对象划分到相应的区域 $POS(S)$、$BND(S)$ 和 $NEG(S)$ 的损失函数。类似地，当对象 x 不属于 S 时，λ_{PN}、λ_{BN}、λ_{NN} 分别表示将一个对象划分到相应的 $POS(S)$、$BND(S)$ 和 $NEG(S)$ 的损失函数。

考虑一种特殊情况，假设损失函数满足：

$$\lambda_{PP}\leqslant\lambda_{BP}<\lambda_{NP} \text{ 且 } \lambda_{NN}\leqslant\lambda_{BN}<\lambda_{PN} \tag{1}$$

其实际意义就是当 x 属于 S 时，将其划分到正域带来的风险要小于或等于将其划分到边界域带来的风险，这两个都小于将其划分到负域的风险；同理，对于不属于 S 的对象 x，将其划分到 S 的负区域所带来的风险要小于或等于将其划分到边界区域的风险，这两者的风险都小于将其划分到 S 的正区域所带来的风险。该假设是符合现实意义的。

令：

$$\alpha=\frac{\lambda_{PN}-\lambda_{BN}}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})}$$

$$\gamma=\frac{\lambda_{PN}-\lambda_{NN}}{(\lambda_{PN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})}$$

$$\beta=\frac{\lambda_{BN}-\lambda_{NN}}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{BP})}$$

如果损失函数满足式（1）的关系，根据 Yao 的三支决策语义规则[15]，则可以推导出 $\alpha\in(0,1]$，$\gamma\in(0,1)$ 和 $\beta\in[0,1)$。因此，有如下符合实际意义的决策规则（P）－（B）：

（P）：如果 $P(Y_j\mid X_i)\geqslant\alpha$，则 $x\in POS(X_i)$；

（N）：如果 $P(Y_j\mid X_i)\leqslant\beta$，则 $x\in NEG(X_i)$；

（B）：如果 $\beta<P(Y_j\mid X_i)<\alpha$，则 $x\in BND(X_i)$。

其中 X_i,Y_j 分别是条件属性子集和决策属性集在 U 上导出的等价类，$X_i\in X,Y_j\in Y$。根据上述决策规则就可以把对象 x 划分到相应的区域。

定义 1 在信息表 $M=(U,At=C\cup D,V,f)$ 中，属性集合 $R\subseteq C$ 的决策风险定义为[18]：

$$Cost_R=\sum_{x_i\in POS_{(\alpha,\beta)}(Y|X)}(\lambda_{PP}P(x_i)+\lambda_{PN}(1-P(x_i)))+\sum_{x_j\in BND_{(\alpha,\beta)}(Y|X)}(\lambda_{BP}P(x_j)+\lambda_{BN}(1-P(x_j)))+\sum_{x_k\in NEG_{(\alpha,\beta)}(Y|X)}(\lambda_{NP}P(x_k)+\lambda_{NN}(1-P(x_k)))$$

考虑到正确分类的风险为 0，也就是说 $\lambda_{PP}=\lambda_{NN}=0$，则有

$$Cost_R=\sum_{x_i\in POS_{(\alpha,\beta)}(Y|X)}\lambda_{PN}(1-P(x_i))+\sum_{x_j\in BND_{(\alpha,\beta)}(Y|X)}(\lambda_{BP}P(x_j)+\lambda_{BN}(1-P(x_j)))+\sum_{x_k\in NEG_{(\alpha,\beta)}(Y|X)}\lambda_{NP}P(x_k) \tag{2}$$

其中，$P(x_i)=\max(\frac{|X_i\cap Y_j|}{|X_i|})$，$x_i\in X_i$，$R\subseteq C$。定义 1 表示在决策表中对每个对象做出相应决策时所带来的风险总和。

定义 2 在信息表 $M=(U,At=C\cup U,V,f)$ 中，属性集合 $R\subseteq C$ 是 C 的一个决策属性约简[18]，当且仅当：

（1）$R=\arg(\min_{R\subseteq C}(Cost_R)$；

（2）$\forall R'\subseteq R,Cost_{R'}>Cost_R$。

定义 2 说明在计算约简时将不再考虑约简前后区域的变化，只关注区域变化后所带来的决策风险是否减小。在经典粗糙集下，大部分属性约简的定义都是基于保持整个决策表正域不变的。因为对整个决策表而言，负域为空集，所以保持整个决策表正域不变的同时，实际也就保证了整个决策表的边界域不变，即保持整个决策表的 3 个区域都不变。在决策表中，属性的增加和减少会在一定程度上影响着决策表各个区域的变化，而决策表区域的变化却不是由属性的增减直接决定的。回顾决策粗糙集模型，每一个对象所属的区域是由风险最小化原则决定的，划分到哪一个区域所带来的风险最小，就将该对象划到相应的区域。区域的变化实际上是由风险的变化所决定的，因此每次决策应该依据风险最小化原则进行。所以在决策粗糙集模型中我们不再关注于区域的变化，只关注属性增减后决策风险的变化，由此基于风险最小化的属性约简定义是合理的。也就是说，如果添加一个属性后，使得整个决策表决策风险减少，则该属性是可增加的。

3 基于联合属性重要度的决策风险最小化的约简

3.1 联合属性重要度定义

在经典粗糙集理论中，属性重要性反映的是条件属性子集对决策属性的分类能力，即决策属性对条件属性子集的依赖程度。而在文献[18]中，定义的决策风险最小化的属性约简只考虑了风险代价对对象所处区域划分的影响，并没有考虑条件属性子集对决策属性的分类能力（即属性重要性）。文献[19]将属性重要性引入到风险最小化属性约简中，但算法中仅仅按照单个属性对决策的分类能力进行重要度排序逐一删除。在某些现实数据中，条件属性之间的相关性很强，有可能两个属性对决策的分类能力都很强，但联合在一起，并不能增加对决策的分类能力。因此，给出决策粗糙集模型下属性

的重要度定义。

在信息表 $M=(U,At=C\cup D,V,f)$ 中，属性子集 $R\subseteq C$。设条件属性子集 R 和决策属性集 D 在 U 上导出的划分分别为 X^R,Y，其中 $X^R=\{X_1^R,X_2^R,\cdots,X_i^R,\cdots,X_s^R\}$，$Y=\{Y_1,Y_2,\cdots,Y_j,\cdots,Y_v\}$。如果 $x\in X_i^R$，那么在条件属性子集下划分的等价类对决策属性集 D 下划分的等价类 Y_j 的条件概率大小用 $P(x)$ 表示，定义如下：

$$P(x)=\max_j(P(Y_j\mid X_i^R))=\max_j(\frac{|Y_j\cap X_i^R|}{|X_i^R|}),j=1,2,\cdots,v \tag{3}$$

$P(x)$ 反映的是在条件属性子集 R 下划分的等价类对决策属性分类能力的大小。式(3)中取最大值反映的是使得确定性程度最大，这样取值也符合概率统计的实际意义。

定义 3　在信息表 $M=(U,At=C\cup D,V,f)$ 中，n 为对象个数，设条件属性子集 $R\subseteq C$ 与决策属性集 D 在 U 上导出的划分分别为 X^R,Y，其中 $X^R=\{X_1^R,X_2^R,\cdots,X_i^R,\cdots,X_s^R\}$，$Y=\{Y_1,Y_2,\cdots,Y_j,\cdots,Y_v\}$。如果 $x_l\in X_i^R$，则条件属性子集 R 下的重要度 $SGF(R)$ 定义为：

$$SGF(R)=\sum_{l=1}^{n}P(x_l)=\sum_{l=1}^{n}\max_j(P(Y_j\mid X_i^R))=\sum_{l=1}^{n}\max_j(\frac{|Y_j\cap X_i^R|}{|X_i^R|}),l=1,2,\cdots,n,j=1,2,\cdots,v \tag{4}$$

$SGF(R)$ 是对条件属性子集 R 相对于决策分类能力大小的度量。由于考察的对象是整个论域 U，因此对每个对象的条件概率 $P(x_l)$ 求和。

3.2　基于联合属性重要度的决策风险最小化启发式算法

很多时候不需要得到所有的约简结果，一个约简结果已经足够帮助我们得到决策规则并做出决策。因此，采用启发式属性约简算法可以大大地提高属性约简的效率。启发式约简方法通常以重要度作为启发式信息，包含 3 种方法：前向添加、后向删除以及两者结合法。文献[19]采取后向删除法，根据式(2)首先计算约简前整个信息表的决策风险代价，在保证原决策表风险代价不变大的基础上逐个删除属性做进一步约简。然而，当属性较多时，采用后向删除法将耗费大量的时间。因此，本文提出一种基于联合属性重要度的决策风险最小化属性约简算法 JAS&MDC(Attribute Reduction Based on Joint Attribute Significance and Minimum Decision Cost)。首先根据定义 3 选择重要度最大的属性进行添加，如果添加后整个决策表的风险代价小于未添加前的决策表的风险代价，说明此属性可以帮助减小风险代价，同时该属性对决策有较强的分类能力；反之，则算法结束，输出约简结果。

JAS&MDC 算法描述如下：

Input：信息表 $M=(U,At=C\cup D,V,f)$，阈值 α,β

Output：属性约简集合 R

Step1　置 $R\neq\emptyset$；

Step2　根据定义 3，计算每个属性 $a_k\in C$ 的重要度 $SGF(\{a_k\})$，$k=1,2,\cdots,m$，选择使得 $SGF(\{a_k\})$ 最大的属性 a_k 添加到约简集合 R 中(若同时有多个属性达到最大值，则从中选取一个等价类个数最少的属性作为 a_k)；

Step3　计算 $Cost_R$：

1. 求论域 U 中每个对象 x_l 的 $P(x_l)$ 值，$l=1,2,\cdots,n$；
2. 根据决策规则(P)－(B)以及预先给定的 α,β 的值将对象 x_l 加入到正域、边界域、负域中；
3. 根据式(2)计算 $Cost_R$；

Step4　对条件属性集 C－R 重复：

1. 对每个属性 $a_k\in C-R$，计算联合重要度 $SGF(R\cup\{a_k\})$；
2. 选择使得 $SGF(R\cup\{a_k\})$ 最大的属性 a_k(若同时有多个属性达到最大值，则从中选取一个等价类个数最少的属性作为 a_k)；
3. 令 $R'=R\cup\{a_k\}$，计算 $Cost_{R'}$；
4. 如果 $Cost_{R'}\leqslant Cost_R$，则 $R=R'$；否则终止；

Step5　最后得到的 R 就是条件属性 C 相对于 D 的一个决策风险最小化约简。

该算法以空集为起点，在已有子集条件下每次添加最重要的属性至约简子集中，如果添加后整个决策表的风险代价小于未添加前的决策表的风险代价，说明此属性有利于减小风险代价，将其加入到约简集合中；反之，则算法结束，输出该约简集合作为最终的约简结果。

4　实验与分析

为进一步说明算法的有效性，本文选用了 UCI 上的 5 个数据集进行实验。对于每一个数据集分别采用本文算法和文献[18,19]中的算法以及未进行约简的数据集来进行实验，通过约简过程所需要的运行时间以及约简子集的长度、约简子集对决策的分类能力来评价所提算法的有效性。

4.1　实验数据

本实验选用 UCI 上的 5 组数据集进行实验，分别为 soybean-small，MushRoom，Contraceptive Mothod Choice，Chess (King-Rook Vs. King-Pawn)，Spect. Test。$|U|$ 表示数据集中对象的个数，$|C|$ 表示整个数据集的条件属性个数，t 表示算法约简过程的运行时间(单位为秒)，约简子集表示算法所选取的属性列号，3NN 和最近邻中的数据分别表示采用该分类器的十折交叉验证平均分类准确率。

4.2　评价指标

本文将所提算法与文献[18,19]中的进行对比。对约简算法而言，除了考虑其运行时间外，还应考虑算法所得约简结果的好坏。因此本实验根据约简子集的长度以及约简子集相对于决策的分类能力评价算法的优劣，分别采用最近邻和 3NN 作为分类器。为了更好地度量分类准确率，采用十折交叉验证求平均分类准确率。

4.3　实验结果与分析

实验结果如表 1 所列。实验过程中，参数 λ_{BP}、λ_{NP}、λ_{PN}、λ_{BN} 均为随机选取，但满足 $\lambda_{BP}<\lambda_{NP}$ 且 $\lambda_{BN}<\lambda_{PN}$。本实验在 $\alpha=0.75,\beta=0.5$ 下进行。从表 1 可以看出，本文所提基于联合属性重要度的决策风险最小化的属性约简算法能大大减少约简子集中属性的个数，使得约简长度更短。同时，与文献[18]、文献[19]中提出的算法相比，本文算法由于采用前向添加法，也有效地降低了时间复杂度，能快速地获得约简，本算法整体运行时间明显低于其他两种算法。特别当数据集较大时，这种时间复杂度的降低更能快速地提高算法的效率。本文算法在计算约简时，在保证风险代价不增加的情况下，同时保证联合属性子集的属性重要度最大，即所获得的约简对整个论域的划分能力更强，从实验结果中也能反映出本文算法所选取的属性具有较强的分类能力。特别在 Chess 和 Spec. Test 数据集中，本文所提算法比未进行约简的数据集平均分类准确率更高，可以充分说明本文所提约简算法可以过滤部分噪声数据，因此所获得的约简子集是有效的。虽然在 soybean-small 和 Contraceptive Mothod Choice 数据集中，本文所提算法的分类准确率略低于文献[19]算法，但由于本文所选取的属性个数明显较少，因此总体来说本文所提算法所选取的属性子集是较优的。

表 1 算法结果与时间比对

数据集	|U|	|C|	算法	t(s)	约简子集	3NN(%)	最近邻(%)
soybean-small	47	35	未约简			100	100
			文献[19]算法	1.777	21,22	100	100
			文献[18]算法	5.264	1,2,3,4,5,6	76.6	80.85
			本文算法	0.511	22	87.2	87.23
MushRoom	8124	22	未约简			100	100
			文献[19]算法	6.783	5	100	100
			文献[18]算法	30.32	1,2,3,4,5	100	100
			本文算法	0.566	5	100	100
Contraceptive Mothod Choice	1473	8	未约简			49.08	47.86
			文献[19]算法	2.046	1,2,3,4,5,6,7,8	49.08	47.86
			文献[18]算法	18.15	3,7,1,2,6,5,4,8	49.08	47.86
			本文算法	0.579	3	48	48
Chess (King-Rook Vs. King-Pawn)	3196	36	未约简			96.37	98.28
			文献[19]算法	10.352	1,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36	96.5	98.19
			文献[18]算法	122.3	6,7,10,12,16,17,20,23,25,27,30,31,33,1,2,3,4,5,8,9,11,13,14,15,18,19,21,22,24,26,28,29,32	95.15	97.31
			本文算法	34.87	4,6,7,9,10,16,17,21,26,27,30,33,36	98.06	99.16
Spect. Test	187	22	未约简			88.24	87.17
			文献[19]算法	0.444	所有属性	88.24	87.17
			文献[18]算法	7.511	22,1,3,19,13,16,20,21,2,4,5,6,7,8,9	90.37	90.37
			本文算法	0.233	1	92	91.98

结束语 经典粗糙集的属性约简定义基本都是保持正域不变。而在决策粗糙集模型里，由于正域、负域和边界域都是不确定的，属性约简最多只能定义为保持一个区域不变，而并不能同时保持 3 个区域均不变。对于变化的 3 个区域，保持哪个区域不变是依据问题需要而主观定义的，不具备泛化性。在决策粗糙集模型中，作出决策应依据风险最小化原则，使得决策代价最小。因此本文提出了一种基于决策风险最小化的属性约简，在属性的选取上考虑到整个属性子集对决策的划分能力而非单个属性的划分能力。实验证明所提约简算法得到的约简子集较优。

参 考 文 献

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

[2] 王国胤，姚一豫，于洪. 粗糙集理论与应用研究综述[J]. 计算机学报, 2009, 32(7): 1229-1246

[3] Mac P N, Shen Q J R. Rough and fuzzy-rough methods for mammographic data analysis[J]. Intelligent Data Analysis-An International Journal, 2010, 14(2): 225-244

[4] Zhu W. Topological approaches to covering rough sets[J]. Information Sciences, 2007, 177(6): 1499-1508

[5] Qian Y, Liang J, Yao Y, et al. MGRS: A multi-granulation rough set[J]. Information Sciences, 2010, 180(6): 949-970

[6] Suyun Z, Tsang E, Degang C. The Model of Fuzzy Variable Precision Rough Sets[J]. IEEE Transactions on Fuzzy Systems, 2009, 17(2): 451-467

[7] 黄兵，胡作进，周献中. 优势模糊粗糙模型及其在审计风险评估中的应用[J]. 控制与决策, 2009, 24(6): 899-902

[8] Xu F F, Miao D Q, Wei L. Fuzzy-rough attribute reduction via mutual information with an application to cancer classification [J]. Computers & Mathematics with Applications, 2009, 57(6): 1010-1017

[9] 胡清华，赵辉，于达仁. 基于邻域粗糙集的符号与数值属性快速约简算法[J]. 模式识别与人工智能, 2008, 21(6): 732-738

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

[11] 钱进，苗夺谦，张志华，等. MapReduce 框架下并行知识约简算法模型研究[J]. 计算机科学与探索, 2013, 7(1): 35-45

[12] Pawlak Z, Wong S K M, ZiarkoW. Roughsets: probabilistic versus deterministic approach[J]. Inter. Journal of Man-MachineStudies, 1988, 29(1): 81-95

[13] Hu Qing-hua, Yu Da-ren, Liu Jin-fu, et al. Neighborhood rough set based heterogeneous feature subset selection[J]. Information Sciences, 2008, 178(18): 3577-3594

[14] Yao Y Y, Wong S K M. A decision theoretic framework for approximating concepts [J]. Inter. Journal of Man-machine Studies, 1992, 37(6): 793-809

[15] Yao Y Y. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181(6): 1080-1096

[16] Yao Y Y, Zhao Y. Attribute reductions in decision-theoretic rough set models[J]. Information Sciences, 2008, 178(17): 3356-3373

[17] Li H X, Zhou X Z, Zhao J B, et al. Attribute reduction in decision-theoretic rough set model: a further investigation[C]// Proceedings of RSKT2011. LNCS, vol. 6954, 2011: 466-475

[18] Jia Xiu-yi, Liao Wen-he, Tang Zhen-min, et al. Minimum cost attribute reduction in decision-theoretic rough set models[J]. Information Sciences, 2013, 219(10): 151-167

[19] 于洪，姚园，赵军. 一种有效的基于风险最小化的属性约简算法[J]. 南京大学学报(自然科学), 2013, 49(2): 210-216

[20] Zhang Xian-yong, Miao Duo-qian. Region-based quantitative and hierarchical attribute reduction in the two-category decision theoretic rough set model[J]. Knowledge-Based Systems, 2014, 71(11): 146-161