文章编号:1001-506X(2010)12-2603-05

基于邻域粒化的小生境微粒群混合数据约简

赵佰亭,陈希军,曾庆双
(哈尔滨工业大学空间控制与惯性技术研究中心,黑龙江 哈尔滨 150001)

摘 要:混合决策系统中同时包含了符号型属性和数值型属性,经典粗糙集处理数值型属性时需要进行离散化,这样会造成信息的丢失。基于邻域粒化的思想,提出了小生境微粒群约简方法,分析了邻域距离函数的选择和大小对分类精度和约简属性数量的影响。邻域粒化的方法可以直接处理数值型属性,微粒群全局优化的特性可以有效的求解全部约简,小生境技术的采用避免了微粒群算法的早熟收敛。选取 UCI 数据集进行了仿真实验,结果表明该方法可以快速有效地求解混合决策系统的约简,而不影响系统的分类精度。

关键词:人工智能;粗糙集;小生境技术;微粒群

中图分类号:TP 18 **文献标志码**:A **DOI**:10.3969/j.issn.1001-506X.2010.12.23

Hybrid attributes reduction based on neighborhood granulation and niche PSO algorithm

ZHAO Bai-ting, CHEN Xi-jun, ZENG Qing-shuang
(*Space Control and Inertial Technology Research Center, Harbin Inst. of Technology, Harbin 150001, China*)

Abstract: Hybrid decision systems include character attributes and numerical attributes. The lost of information when discretize the numerical attributes by Pawlak rough set is introduced. A reduction algorithm based on the neighborhood rough set model and the niche particle swarm optimization (PSO) algorithm is proposed. The affection of neighborhood operator to the reduction and classification is discussed also. Numerical attributes can be dealt directly by neighborhood relations. The PSO algorithm is a global optimization algorithm and can get all reductions. The use of the niche technology can avoid the premature convergence of the PSO. Experimental results demonstrate the validity and feasibility of the proposed algorithm, in application to four University of California at Irvine (UCI) machine learning databases.

Keywords: artificial intelligence; rough set; niche technology; particle swarm optimization (PSO)

0 引 言

粗糙集理论是在经典集合论基础上发展起来的处理不完备、不确定、不一致信息的数学工具。1995 年文献[1]的发表,引起了计算机和应用数学领域研究人员的广泛关注。目前该理论已经成为机器学习领域的一个十分活跃的研究分支。粗糙集方法被广泛地应用于工业、医疗、商业、制药、金融、化学、材料、地理、社会科学、分子生物等诸多领域,取得了巨大的成功[2]。

经典的 Pawlak 粗糙集理论无法处理数值型属性,包含数值型属性的决策系统在应用粗糙集进行研究时需要进行离散化处理。但是由于属性值的分布通常是不均匀的,并且属性值的噪声点会对这些方法的离散结果产生明显的影响,这些方法在实际应用中通常难以取得满意的效果[3]。另外,实际应用中普遍存在着既有符号型数据又有数值型数据的混合数据库,为了解决这一问题,提出了模糊粗糙集模型[4]、相似关系粗糙集模型[5]和邻域关系模型。邻域粗糙集模型通过点的邻域来粒化论域,将邻域作为基本信息粒子。与经典粗糙集方法相比,该方法省去了离散化的过程,模型可以直接分析数值型数据,与模糊信息熵方法和相似关系粗糙集相比,也表现出更优的性能,拓展了经典粗糙集理论的应用范围。

属性约简是粗糙集理论研究的热点问题之一,目前已经提出了许多属性约简算法,其中基于差别矩阵、近似质量和信息熵的三种方法是最为基本的属性约简方法[6],其他方法都是在这三种方法的基础上对算法的效率、性能等方面的改进研究。为了可靠地解决全局优化问题,人们模拟自然界的一些自然现象,发展起来一系列仿生型智能优化

收稿日期:2009-03-11; **修回日期**:2010-01-07。
基金项目:国防科技预研基金(9140A17030207HT0150);“十一五”总装备部预研基金(51309030102)资助课题
作者简介:赵佰亭(1981-),男,博士,主要研究方向为智能控制、机器学习和故障诊断等。E-mail:zhaobaiting@gmail.com

算法，如模拟退火方法、遗传算法[7]、微粒群算法[8]等。但是文献[7]提出的遗传算法约简方法，存在收敛速度慢，容易陷入局部最优的问题。文献[8]提出的微粒群算法，模拟遗传算法引入变异操作，形式上强化了微粒的多样性，却打乱了初始群体在可行域上分布的平衡性，从而最终导致单一模式收敛。本文引入小生境技术，提出了基于小生境粒子群的属性约简算法。小生境技术的引入解决了粒子群属性约简早熟收敛的问题，并避免了算法中的参数对具体优化目标的敏感性和单一收敛性，仿真实验结果显示该算法在总体性能上优于文献[4,7-8]。

1 邻域粗糙集模型

粗糙集理论采用决策系统的概念对问题进行描述，并以数据形式表示决策系统，其中行表示对象，列表示对象的属性。

称四元组 $DT=(U,C\cup D,V,f)$ 是一个决策系统，其中 $U=\{x_1,x_2,\cdots x_n\}$ 为对象的非空有限集合，称为论域；$C=\{\alpha|\alpha\in C\}$ 为属性的非空有限集合；$D=\{d|d\in D\}$ 称为决策属性集；$V=\cup V_\alpha(\forall\alpha\in C\cup D)$ 表示信息函数 f 的值域，V_α 为属性 α 的值域；$f=\{f_\alpha|f_\alpha:U\rightarrow V_\alpha,\forall\alpha\in C\cup D\}$ 表示决策表的信息函数。

如果 V 是包含名义型和数值型等不同的属性变量，则称该决策系统为混合决策系统。在混合决策系统中，基于等价类的不可分辨关系不再适用。邻域粗糙集模型拓展了等价关系，用邻域关系来度量不可分辨关系，能够直接处理数值型属性，拓展了经典粗糙集理论的应用范围[9]。

给定一个由 N 个属性描述的分类数据，可以将其形式化为一个决策系统 $\langle U,A,D\rangle$。$U=\{x_1,x_2,\cdots,x_N\}$ 是全部样本构成的集合，$A=\{a_1,a_2,\cdots,a_N\}$ 是描述样本的属性集合，D 是分类决策属性。由属性张成一个 N 维的空间，则样本就是空间的点集。

定义1 设 U 是非空集合，若对 U 中的任意元素 x_i、x_j、x_k，都存在唯一确定的实函数 Δ 与之对应，而且 Δ 满足[9]

(1) $\Delta(x_i,x_j)\geqslant 0$，当且仅当 $x_i=x_j$，$\Delta(x_i,x_j)=0$；

(2) $\Delta(x_i,x_j)=\Delta(x_j,x_i)$；

(3) $\Delta(x_i,x_k)\leqslant\Delta(x_i,x_j)+\Delta(x_j,x_k)$。

称 Δ 是 U 上的距离函数，并且称 $\langle U,\Delta\rangle$ 为距离空间，也称为度量空间。

在 N 维欧式空间中，给定任意两点 x_i 和 x_j，其距离为

$$\Delta(x_i,x_j)=\left(\sum_{i=1}^{N}(x_{li}-x_{lj})^2\right)^{1/2} \tag{1}$$

定义2 设 $\langle U,\Delta\rangle$ 是非空度量空间，称点集 $\delta(x)=\{y|\Delta(x,y)\leqslant\delta,y\in U\}$ 为 x 的 δ 邻域，$\delta\geqslant 0$[9]。

当数值和符号属性共存时，设 $B_1\subseteq A$ 和 $B_2\subseteq A$ 分别是数值属性和符号属性，那么样本 x 的邻域定义为

(1) $\delta_{B_1}(x)=\{x_i|\Delta_{B_1}(x,x_i)\leqslant\delta,x_i\in U\}$；

(2) $\delta_{B_2}(x)=\{x_i|\Delta_{B_2}(x,x_i)=0,x_i\in U\}$；

(3) $\delta_{B_1\cup B_2}(x)=\{x_i|\Delta_{B_1}(x,x_i)\leqslant\delta\wedge\Delta_{B_2}(x,x_i)=0,x_i\in U\}$。

论域中所有对象的邻域形成了论域的粒化，邻域粒子族构成了论域空间中的基本概念系统，通过这些基本概念，我们可以逼近空间中的任意概念。邻域具有如下性质：

(1) $\delta(x_i)\neq\varnothing$，因为 $x_i\in\delta(x_i)$；

(2) $x_j\in\delta(x_i)\Rightarrow x_i\in\delta(x_j)$；

(3) $\bigcup_{i=1}^{n}\delta(x_i)=U$。

定义3 给定邻域近似空间 $NAS=\langle U,N\rangle$ 以及 $X\subseteq U$，X 在邻域近似空间 $NAS=\langle U,N\rangle$ 的下近似与上近似分别定义为[9]

$$\begin{cases}\underline{N}X=\{x_i\mid\delta(x_i)\subseteq X,x_i\in U\}\\ \bar{N}X=\{x_i\mid\delta(x_i)\cap X\neq\varnothing,x_i\in U\}\end{cases} \tag{2}$$

$\forall X\subseteq U$，有 $\bar{N}X\supseteq X\supseteq\underline{N}X$，并定义 X 在近似空间 $NAS=\langle U,N\rangle$ 中的近似边界为

$$BN(X)=\bar{N}X-\underline{N}X \tag{3}$$

$\underline{N}X$ 也称为 X 在近似空间 $NAS=\langle U,N\rangle$ 中的正域，它是那些邻域被 X 所完全包含的对象的集合；相应地，$\bar{N}X$ 是那些邻域与 X 交集非空的对象的集合。那些邻域与 X 完全无关的对象称之为 X 的负域，定义为

$$NEG(X)=U-\bar{N}X=\{x_i\mid\delta(x_i)\cap X=\varnothing\} \tag{4}$$

定义4 给定样本集合 U，A 是描述 U 的条件属性集合，D 是决策属性，如果 A 生成一簇论域上的邻域关系 N，则称 $NDS=\langle U,N,D\rangle$ 为邻域决策系统。

定义5 给定邻域决策系统 $NDS=\langle U,A,D\rangle$，D 将 U 划分为 N 个等价类：$X_1,X_2,\cdots,X_N$，$B\subseteq A$ 生成 U 上的邻域关系 N_B，那么决策 D 关于 B 的邻域下近似和上近似分别为[9]

$$\begin{cases}\underline{N_B}D=\{\underline{N_B}X_1,\underline{N_B}X_2,\cdots,\underline{N_B}X_N\}\\ \overline{N_B}D=\{\overline{N_B}X_1,\overline{N_B}X_2,\cdots,\overline{N_B}X_N\}\end{cases} \tag{5}$$

决策 D 的下近似也称为决策正域，正域的大小反映了分类问题在给定属性空间中的可分离程度。正域越大，表明各类的重叠区域即边界越少。

$\forall X\subseteq U$，有 $\underline{N}X\supseteq X\supseteq\bar{N}X$，决策的边界可定义为

$$BN(D)=\overline{N_B}D-\underline{N_B}D \tag{6}$$

当 $BN(D)=\varnothing$，即 $\overline{N_B}D=\underline{N_B}D$ 时，称 X 在近似空间中 $\langle U,N\rangle$ 是可定义的，否则称 X 是粗糙的。

给定一个邻域决策系统 $NDS=\langle U,A,D\rangle$，决策属性 D 对条件属性 $B\subseteq A$ 的依赖度为[9]

$$\gamma_B(D)=Card(\underline{N_B}D)/Card(U) \tag{7}$$

依赖度 $\gamma_C(D)$ 表示在条件属性 C 下能够确切划入决策在 U/D 的对象占论域中总对象数的比率，表达了决策属性对条件属性的依赖程度。当属性的依赖度为零时，该属性是冗余属性。依赖度反映了决策系统中的条件属性并不是

同等重要的，可以简单的这样理解，决策属性对条件属性的依赖程度越高，则该属性的重要程度越高。

在基于邻域关系依赖度的基础上，可以得出如下约简的定义。

定义 6　给定一个邻域决策系统 $NDS=\{U,A,D\}$，设 $B\subseteq A$，称 B 是 A 的一个相对约简，如果 B 满足[9]

(1) 充分条件：$\gamma_B(D)=\gamma_A(D)$；

(2) 必要条件：$\forall a\in B,\gamma_{B-a}(D)<\gamma_B(D)$。

在现阶段针对邻域粗糙集模型的研究中，文献[9]中提出了基于前向贪心搜索的属性约简算法，每次将依赖度最大的属性加入到约简集中，直到添加新的属性依赖度不再变化为止，但是在大多数情况下得到的约简并不是最小约简，仍然包含冗余属性，为此，本文提出了采用小生境微粒群优化算法来求解混合数据的相对约简。

在求解混合数据的相对约简过程中，邻域算子 δ 大小的选择会对约简中包含的属性数量以及约简后决策系统的分类精度产生影响。

如图 1 所示，δ 越大系统的分类边界区域也就越大，不能确切分类的样本也就越多 δ 取值越小则分类的复杂度越高。不同的邻域大小会得到不同的约简和不同的分类精度，如何选择邻域算子是一个重要的研究内容，另外邻域距离函数的选择也可能会对分类精度产生影响，也是一个重要的研究内容。

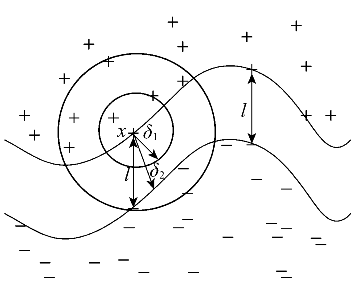

图 1　δ 对分类精度的影响

2　小生境微粒群约简算法

微粒群算法是 Kennedy 和 Eberhart[10] 等首先提出并改进的一种基于群体智能的优化算法，通过对动物社会行为的观察，发现在群体中对信息的社会共享提供了一个演化的优势，并以此作为开发算法的基础。算法中加入了近邻的速度匹配，考虑了多维搜索和根据距离的加速，并引入了惯性权重来更好的控制开发和探索。其主要特点是，群体搜索策略和群体中个体之间的信息交换，搜索不依赖于梯度信息。实践证明，微粒群算法对于组合优化中的 NP 难问题[11]非常有效。例如微粒群算法已经在求解旅行商问题、背包问题、装箱问题、图形划分问题等方面得到成功的应用。

为了改善微粒群优化算法的全局搜索性能，提高搜索的有效性及优化效率，本文将小生境技术引入到微粒群优化算法中，形成小生境微粒群优化算法。小生境微粒群优化算法的最基本思想是将生物学中处于分离的孤立地理小生境中的不同物种之间是不进行竞争等信息交流而独立进化这一概念移植到微粒群算法当中，以保持算法在迭代过程中微粒群的多样性[11]。

2.1　微粒适应度的评价函数

通过编码将样本表示成搜索空间的微粒串结构数据，首先随机产生 m 个长度为 n 的由 $\{0,1\}$ 组成的二进制串 x_{id}，同时 x_{id} 又表示微粒群的初始位置。然后随机产生 $[0,1]$ 范围的随机数 v_{id} 作为微粒子的初始速度。根据混合决策系统属性约简的实际需求，选取了如下形式的适应度函数

$$F(x)=\beta*f(x)*P(x) \tag{8}$$

式中，$f(x)=(n-card(x))/n$，$card(x)$ 是决策系统的正域；n 为微粒位置的长度；$P(x)=(1+e^{\alpha(\gamma_0-\gamma_x(D))})^{-1}$；$\gamma_x(D)$ 为依赖度；γ_0 是阈值。

微粒中单个微粒位置的优劣由微粒适应度函数来确定，适应度函数是对微粒位置的适应性进行评价的指标，所以适应度函数的选取非常重要，决定了混合决策系统最终能否找到最小约简。

2.2　速度和位置的更新

计算各微粒新的速度和位置。对每一代，其 d 维根据式(9)变化

$$\begin{aligned} v_{id}&=\omega v_{id}+c_1 rand()(p_{id}-x_{id})+\\ &\quad c_2 Rand()(p_{gd}-x_{id}) \end{aligned} \tag{9}$$

$$x_{id}=\begin{cases}1,\ \boldsymbol{\rho}_{id}<sig(v_{id})\\ 0,\ \boldsymbol{\rho}_{id}\geqslant sig(v_{id})\end{cases} \tag{10}$$

式中，ω 是权重函数、c_1 和 c_2 为加速度常数、$rand(\ )$ 和 $Rand(\ )$ 是两个在 $[0,1]$ 范围内变化的随机数、p_{id} 为微粒 i 经过的最好位置(适应度最大)、p_{gd} 为群体所有微粒经过的最好位置、向量 $\boldsymbol{\rho}_{id}$ 的各个分量是在 $[0,1]$ 范围取值的随机数、$sig(\ )$ 为 $sigmoid$ 函数。

式(9)中第一部分为微粒前一代的速度，微粒的速度被限制在最大速度 $v_{\max}$ 以内；第二部分为认知部分，表示微粒的思考；第三部分为社会部分，表示微粒间的信息共享与相互合作。

2.3　交叉运算

按照预先设定的概率 P_c 随机选择 $m\times P_c$ 对微粒子，对每一对微粒子的每一位按相同的概率进行交叉。设交叉后的微粒子位置为 $x'_1=a'_{11}a'_{12}\cdots a'_{1n}$，$x'_2=a'_{21}a'_{22}\cdots a'_{2n}$，则交叉运算为

$$O(P_c,rand):\begin{cases} a'_{1i}=\begin{cases}a_{1i},\ rand>0.5\\ a_{2i},\ rand\leqslant 0.5\end{cases}\\ a'_{2i}=\begin{cases}a_{2i},\ rand>0.5\\ a_{1i},\ rand\leqslant 0.5\end{cases}\end{cases} \tag{11}$$

计算新产生微粒的适应度值，如果其中一个或者两个

的适应度值优于交叉操作之前对应微粒的适应度值，则接受它们为新微粒；否则，回到交叉之前的状态。如果新微粒的适应度增大，则更新个体最好位置和群体最好位置。

2.4 小生境优化

生物学上，小生境是指特定环境下的一种组织结构。小生境技术就是将每一代个体划分为若干类，每个类中选出若干适应度较大的个体作为一个类的优秀代表组成一个群。每个微粒种群独立的搜索最优解，全局最优解在所有微粒群的最优解中产生。各微粒群的独立进化可以避免全部粒子陷入局部最优，可以更好的保持解的多样性，同时具有很高的全局寻优能力和收敛速度。小生境优化的基本思想如下：

(1) 如果微粒速度增量过小，则随机产生[0,1]范围的随机数 v_{id} 作为微粒子的速度，强迫微粒运动，防止陷入局部极值。

(2) 如果两微粒位置接近，将适应度差的微粒重新启动，限制微粒间距离，使其分散于整个搜索空间。

(3) 在整个微粒群中，将适应度差的微粒重新启动，放弃搜索该部分空间，加快搜索速度。

3 约简算法实现

创建微粒群 X，随机产生 m 个微粒子，并对微粒群的位置和速度初始化。对于核中的属性，微粒中的对应位取 1，其他位随机取 0 或者 1，这样可以加快算法的收敛。算法的伪代码如下：

算法 邻域粒化模型微粒群约简算法

输入：$NDT=\{U,A,D\}$和初始种群 X；

输出：p_{id}，p_{gd}。

步骤 1 $\varnothing \rightarrow p_{id}, p_{gd}$；

步骤 2 For each $x_i \in X$；

步骤 3 根据式(9)计算 v_{id}；

步骤 4 根据式(10)计算 x_{id}；

步骤 5 If 速度增量过小，重启该微粒速度，go to 步骤 2；

步骤 6 根据式(8)计算微粒适应度；

步骤 7 If 微粒间欧氏距离小于给定值，重启适应度差的微粒，go to 步骤 2；

步骤 8 End for；

步骤 9 搜索 p_{id}，p_{gd}；

步骤 10 计算 $P_c=\dfrac{1}{1+\mathrm{e}^{\alpha G}}+\beta$，按式(11)进行交叉操作；

步骤 11 根据式(8)计算微粒适应度；

步骤 12 If $F(x)$增大；

步骤 13 搜索 p_{id}，p_{gd}；

步骤 14 Else 返回交叉操作前的状态；

步骤 15 If 连续 t 代微粒的适应度不再提高，或者达到最大飞行代数，则终止计算，输出 p_{id}，p_{gd}；

步骤 16 Else go to 步骤 2。

4 算例分析

为验证算法的正确性，选用美国加州大学 Irvine 分校的机器学习数据库“Heart1”、“Thyroid”、“Vote”和“Wdbc”进行了仿真实验。采用 SVM 分类器对系统的预测精度进行评价，分别用原始样本数据和约简后的样本数据来训练 SVM 分类器，然后用系统的预测精度来评价微粒群属性约简算法求得的约简质量，仿真结果如表 1 所示。

表 1 约简结果

数据名称	样本数	属性数	原始预测精度	约简预测精度
Heart1	270	13	74.01	81.41
Thyroid	9 172	29	63.68	67.77
Vote	435	16	95.33	93.63
Wdbc	569	31	91.97	90.31

通过实验结果可以看出，本文使用的约简算法有效地降低了属性的数量，而基本不改变预测的精度，有效地减少了数据库的冗余数据，达到了约简的目的。

为了研究领域算子的选择及大小对分类精度和约简后属性数量的影响，选用“Heart1”进行了约简和预测仿真实验。图 2 和图 3 给出了“Heart1”分类精度和约简后属性个数随 δ 大小变化的曲线，邻域算子 δ 选用 2 范数距离函数。图 4 和图 5 给出了邻域算子 δ 选用无穷范数距离函数，“Heart1”分类精度和约简后属性个数随 δ 大小变化的曲线。

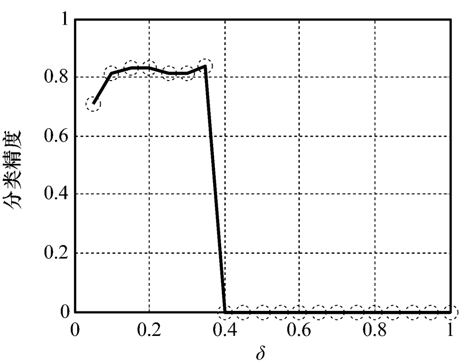

图 2 Heart1 分类精度随 δ 变化曲线($\delta=2$ 范数)

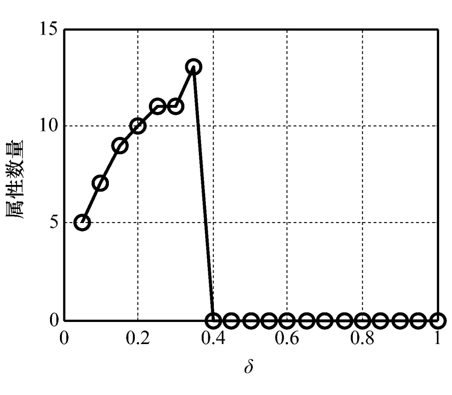

图 3 Heart1 约简属性数量随 δ 变化曲线($\delta=2$ 范数)

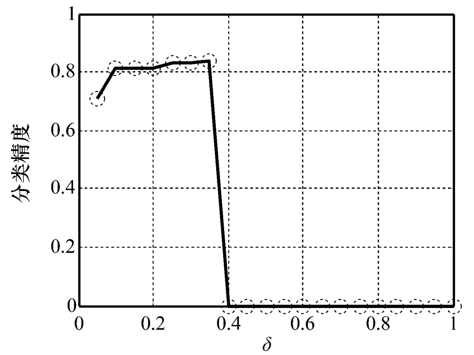

图 4　Heart1 分类精度随 δ 变化曲线(δ=无穷范数)

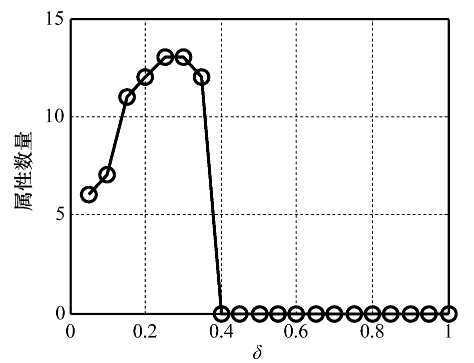

图 5　Heart1 约简属性数量随 δ 变化曲线(δ=无穷范数)

从图 2 和图 4 可以看出,当 δ 较小时,分类问题的粒度很小,较少的特征即可区分决策。当 δ 取值较大,所需特征也越多,当粒度超过一定程度时,任何特征都不足以区分任意样本,δ 的取值和具体分类问题有关。从图 3 和图 5 可以发现,约简中属性的数量随着 δ 的增大而增大,但并不是单调变化。当邻域算子超过某一数值时,任何特征都不足以区分任一样本,得不到任何约简属性,所以合理的选择邻域算子才能取得较好的分类效果。

5　结　论

针对同时包含符号性属性和数值型属性的混合决策系统的属性约简问题,提出了基于邻域粗糙集模型的小生境微粒群约简方法。采用基于邻域等价关系建立的粗糙集模型,用邻域等价关系度量经典粗糙集的不可分辨关系,通过邻域信息粒子逼近论域空间,能够直接处理数值型属性。微粒群算法克服了经典约简算法在求解全部约简时的 NP 难问题。小生境技术的采用避免了微粒群算法的早熟收敛。选用了美国加州大学 Irvine 分校建立的机器学习数据库进行了仿真实验,仿真结果显示,小生境微粒群约简算法可以有效地求解混合数据知识的约简。

参考文献:

[1] Pawkak Z, Grzymala-Busse J, Slowinski R, et al. Rough sets[J]. *Communication of the ACM*, 1995,38(11):89 - 95.

[2] Hu Q H, Zhang C, Chen D G, et al. Gaussian kernel based fuzzy rough sets: model, uncertainty measures and applications[J]. *International Journal of Approximate Reasoning*, 2010, 51(4): 453 -471.

[3] Yu L, Liu H. Efficient feature selection via analysis of relevance and redundancy[J]. *Journal of Machine Learning Research*, 2004,5:1205 - 1224.

[4] Hu Q H, Yu D R, Xie Z X. Information-preserving hybrid data reduction based on fuzzy-rough techniques[J]. *Pattern Recognition Letters*, 2006,27(5):414 - 423.

[5] Slowinski R, Vanderpooten D. A generalized definition of rough approximations based on similarity[J]. *IEEE Trans. on Knowledge and Data Engineering*, 2000,12(2):331 - 336.

[6] 苗夺谦,王国胤,刘清,等. 粒计算:过去、现在与展望[M]. 科学出版社,2007:7 - 47.

[7] 叶玉玲,伞冶. 基于遗传算法的粗糙集混合数据属性约简[J]. 哈尔滨工业大学学报,2008,40(5):683 - 687.

[8] 叶玉玲,伞冶. 一种混合决策系统属性约简算法研究[J]. 系统仿真学报,2007,13(19):2988 - 3004.

[9] 胡清华. 混合数据知识发现的粗糙计算模型和算法[D]. 哈尔滨工业大学,2008.

[10] Kennedy J, Eberhart R C. Particle swarm optimization[C]// *Proc. of the IEEE International Conference on Neural Networks*,1995:1942 - 1948.

[11] Skowron A, Rauszer C. *The discernibility matrices and functions in information systems*[M]// Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Dordrecht: Kluwer Academic Publishers, 1992: 331 - 362.