

# 监督邻域粗糙集

汪琳娜<sup>1,2</sup> 杨 新<sup>3</sup> 杨习贝<sup>4</sup>

(四川工商学院电子信息工程学院 成都 611745)<sup>1</sup>

(里贾纳大学计算机科学学院 萨斯喀彻温 里贾纳 S4S 0A2)<sup>2</sup>

(西南交通大学信息科学与技术学院 成都 611756)<sup>3</sup>

(江苏科技大学计算机科学与工程学院 江苏 镇江 212003)<sup>4</sup>

**摘 要** 传统单一阈值的邻域粗糙集不能有效降低信息的不确定性。考虑对象已有或预测的类别标签信息,通过引入类内和类间两种阈值,提出一种新的邻域粒化方法,并构建了一种基于监督邻域的粗糙集模型。该模型是传统邻域粗糙集的推广形式。通过分析双阈值下的邻域粒子变化规律,给出该模型的粗糙近似质量和条件熵单调性变化定理。最后通过 4 个 UCI 数据集验证了该模型的性能。实验结果显示,可以通过调节监督阈值参数来改善论域的邻域粒化效果,并降低信息的不确定性。

**关键词** 监督邻域,邻域粒化,双阈值,不确定性

**中图法分类号** TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.08.033

## Supervised Neighborhood Rough Set

WANG Lin-na<sup>1,2</sup> YANG Xin<sup>3</sup> YANG Xi-bei<sup>4</sup>

(School of Electronic and Information Engineering, Sichuan Technology and Business University, Chengdu 611745, China)<sup>1</sup>

(Department of Computer Science, University of Regina, Regina, Saskatchewan S4S 0A2, Canada)<sup>2</sup>

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)<sup>3</sup>

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212003, China)<sup>4</sup>

**Abstract** The uncertainty of information can't be efficiently reduced by traditional neighborhood rough set with single threshold. By considering the existing or predicted category label information of the object, this paper introduced two kinds of thresholds, namely, intra-class and inter-class, and proposed a novel neighborhood granulation methods to construct a rough set model based on supervised neighborhood. This model is the generalized form of conventional neighborhood rough set. Moreover, the theorem of monotonic variation with approximate quality and conditional entropy was presented through analyzing the change rules of neighborhood particles under double thresholds. Finally, the performance of the model was demonstrated on four data sets of UCI. The results show that the effect of neighborhood granulation can be improved and the uncertainty of information can be reduced by adjusting supervised threshold parameters.

**Keywords** Supervised neighborhood, Neighborhood granulation, Double thresholds, Uncertainty

## 1 引言

经典 Pawlak 粗糙集<sup>[1]</sup>是一种解决不确定性问题的有效方法,但并不能直接处理数值型数据。针对此问题, Lin<sup>[2-3]</sup>于 1988 年借助拓扑学中内点和闭包的概念,提出了邻域系统的概念;此后, Yao<sup>[4]</sup>、Wu 等<sup>[5]</sup>相继深入研究了邻域系统的性质; Hu 等<sup>[6]</sup>于 2008 年提出了一种基于邻域关系的粗糙集扩展模型。近年来,与邻域粗糙集相关的研究理论和方法被广泛应用于属性约简、特征选择、规则学习和分类器设计等<sup>[7-13]</sup>。

在利用邻域粗糙集模型刻画数值型分类数据的不确定性和不一致性过程中,邻域阈值的选择是一个关键问题。在已

有的研究中,普遍采用单一的邻域阈值来对论域进行粒化。如果阈值过大,会导致邻域粒子过大,即粒化过粗;如果阈值过小,会导致邻域粒子过小,即粒化过细。传统的阈值优化方法通过分析实验效果来给出经验阈值,但通过调整单一阈值只能改变对象的粒化粗细效果,并不能有效降低信息的不确定性。在传统的邻域粒化过程中,通常并不考虑对象的类别标签信息,因此可以称其为一种无监督的粒化方法。根据对象已有或预测的类别标签信息,论域中的任意对象与其他对象间存在类内和类间两种关系,即两个对象在同一类别中时称为类内对象,反之称为类间对象。当运用邻域关系对论域进行粒化时,通常期望邻域粒子中包含较多的类内对象和较

到稿日期:2017-07-21 返修日期:2017-10-29 本文受国家自然科学基金项目(61572242,61573292,71571148),四川省教育厅自然科学基金项目(18ZB0373)资助。

汪琳娜(1986—),女,硕士,讲师,主要研究方向为粒计算、粗糙集和三支决策等;杨 新(1981—),男,博士生,副教授,主要研究方向为粒计算、粗糙集和三支决策等;杨习贝(1980—),男,博士后,副教授,主要研究方向为粗糙集理论、粒计算与机器学习, E-mail: zhenjiangyangxibei@163.com(通信作者)。

少的类间对象。但是,使用单一阈值并不能很好地实现以上目标。因此,在过程中邻域粒化过程中引入类内阈值和类间阈值的概念,通过有效控制两个阈值来增大或减小邻域粒子,使邻域粒子中类对象内的数量增加而类间对象的数量减少。通过这种有监督的粒化方法,可以大大提升邻域粗糙集刻画不确定性知识的能力。

为了充分利用监督学习下的类别标签信息,本文提出了一种基于类内阈值和类间阈值的邻域粒化方法,构建了相应的监督邻域粗糙集模型;然后给出了该模型下的近似质量和条件熵这两种不确定性信息度量方式,并证明了以上两种度量的单调性变化规律;最后通过 UCI 数据集验证了该模型的效果。

### 2 二元邻域关系

下面简要介绍二元邻域关系的基本概念和性质<sup>[6]</sup>。

给定一个决策信息系统  $S=(U, A, V, f)$ , 其中  $U$  表示非空有限的对象集合;  $A$  是非空有限的属性集合,  $A=C \cup D$ ,  $C \cap D = \emptyset$ , 其中  $C$  表示条件属性集,  $D$  表示决策属性集;  $V = \bigcup_{a \in A} V_a$  是值域,  $V_a$  是对象在属性  $a$  下的所有可能取值;  $f: U \times A \rightarrow V$  是一个映射函数, 表示对象和属性取值的映射关系。

**定义 1**<sup>[6]</sup>(单一阈值邻域) 给定任意的  $x_i \in U$  和  $B \subseteq C$ , 定义  $x_i$  的  $\delta$  邻域集为:

$$\delta_B(x_i) = \{x_j | x_j \in U, \Delta_B(x_i, x_j) \leq \delta\} \quad (1)$$

其中,  $\Delta$  是一个距离函数。对于任意的  $x_1, x_2, x_3 \in U$ ,  $\Delta$  满足以下性质:

- 1)  $\Delta(x_1, x_2) \geq 0$ ;
- 2)  $\Delta(x_1, x_2) = 0$ , 当且仅当  $x_1 = x_2$ ;
- 3)  $\Delta(x_1, x_2) = \Delta(x_2, x_1)$ ;
- 4)  $\Delta(x_1, x_3) \leq \Delta(x_1, x_2) + \Delta(x_2, x_3)$ 。

针对  $N$  维特征空间, 下面给出 3 种常用的距离函数。

$$\Delta_p(x_1, x_2) = \left( \sum_{i=1}^N |f(x_1, a_i) - f(x_2, a_i)|^p \right)^{1/p} \quad (2)$$

其中, 当  $p=1$  时,  $\Delta_1(x_1, x_2)$  为曼哈顿距离; 当  $p=2$  时,  $\Delta_2(x_1, x_2)$  为欧氏距离; 当  $p=\infty$  时,  $\Delta_\infty(x_1, x_2)$  为切比雪夫距离。简便起见, 本文采用欧氏距离。

基于邻域关系可以得到论域  $U$  的一个覆盖, 表示为邻域粒子族  $\{\delta_B(x_i) | i=1, 2, \dots, n\}$ 。用关系矩阵  $M=(r_{ij})_{n \times n}$  来描述二元邻域关系, 其中:

$$r_{ij} = \begin{cases} 1, & \Delta(x_i, x_j) \leq \delta \\ 0, & \text{其他} \end{cases} \quad (3)$$

以上二元邻域关系满足自反性和对称性, 但不一定满足传递性。当  $\delta=0$  时, 二元邻域关系退化为等价关系。

### 3 监督邻域粗糙集

在传统邻域粗糙集中, 单一阈值  $\delta$  的变化直接影响对象的粒化效果和属性约简的质量, 但仅通过调整阈值  $\delta$  来改变数据的粒化结构并不能更好地提升对不确定性信息的刻画能力。为此, 考虑监督学习中已有或预测的对象类别标签, 引入类内阈值和类内阈值的概念来分别刻画类内对象和类间对象的距离差异, 给出一种基于监督邻域的粗糙集扩展模型。

**定义 2**(监督邻域) 给定一个决策信息系统  $S=(U, A, V, f)$ ,  $B \subseteq C$ 。决策属性  $D$  将  $U$  划分为  $N$  个决策类, 表示为

$D = \{D_k\}, k=1, 2, \dots, N$ 。设类内阈值  $\delta^I$  和类间阈值  $\delta^O$  满足  $0 \leq \delta^I \leq 1, 0 \leq \delta^O \leq 1$ , 给定任意的  $x_i \in U$ , 则定义  $x_i$  的监督邻域集为:

$$\delta_B^I(x_i) = \{x_j | x_j \in D_k, \Delta_B(x_i, x_j) \leq \delta^I\} \cup \{x_j | x_j \in D_k, x_j \notin D_k, \Delta_B(x_i, x_j) \leq \delta^O\} \quad (4)$$

其中, 当两个对象属于同一个决策类时使用类内阈值  $\delta^I$ , 当两个对象不属于同一个决策类时使用类间阈值  $\delta^O$ 。

相对于传统单一阈值  $\delta$ , 可以通过合理改变监督阈值参数  $\delta^I$  和  $\delta^O$  来降低信息的不确定性。主要有以下两种方法:

1)  $0 \leq \delta^O < \delta^I = \delta \leq 1$ , 使类内阈值  $\delta^I$  等于传统单一阈值  $\delta$ , 减小类间阈值  $\delta^O$ , 此时收紧类间对象的距离阈值;

2)  $0 \leq \delta^O = \delta < \delta^I \leq 1$ , 使类间阈值  $\delta^O$  等于传统单一阈值  $\delta$ , 增大类内阈值  $\delta^I$ , 此时放松类内对象的距离阈值。

显然, 如果  $0 \leq \delta^I = \delta = \delta^O \leq 1$ , 那么类内阈值  $\delta^I$  和类间阈值  $\delta^O$  均等于传统单一阈值  $\delta$ , 此时监督邻域关系退化为传统单一阈值的邻域关系。

篇幅所限, 只讨论第一种方法。传统单一阈值的邻域关系并不区分类内对象和类间对象的距离差别, 采用单一阈值调节邻域半径, 容易造成类内对象不在同一邻域粒子中而类间对象却在同一邻域粒子中。因此, 考虑监督学习中的类别标签信息, 采用双阈值来分别处理类内对象和类间对象的二元邻域关系, 以优化邻域粒化方法, 并降低信息的不确定性。

下面以二分类问题为例, 以五角星标记第一类样本, 以菱形标记第二类样本, 空心的五角星和菱形表示与对象  $x$  存在二元邻域关系的对象集, 即对象  $x$  的邻域粒子。在单一阈值的邻域关系下, 以阈值  $\delta$  为邻域半径, 从图 1(a) 中可以观察到对象  $x$  的  $\delta$  邻域内既有第一类样本又有第二类样本, 且两类样本的数量接近, 此时邻域粒化效果不佳。因此, 考虑监督学习下的样本类别标记信息, 把对象间的距离阈值设定为两种类型, 即同一决策类内对象间的类内阈值  $\delta^I$  和不同决策类对象间的类间阈值  $\delta^O$ 。相对于单一阈值  $\delta$ , 可以通过调节双阈值  $\delta^I$  和  $\delta^O$  来进一步改善邻域粒化效果。

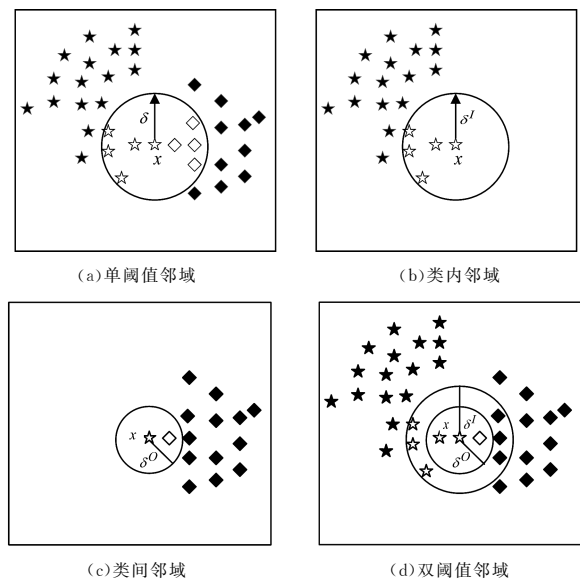


图 1 单一阈值和双阈值下二元邻域关系的对比

Fig. 1 Comparison of binary neighborhood relation between single threshold and dual thresholds

图 1(b)和图 1(c)分别描述了运用类内阈值  $\delta^l$  和类间阈值  $\delta^o$  对对象  $x$  的类内对象和类间对象进行邻域刻画的过程。在图 1(d)中,假设类内阈值  $\delta^l$  与传统单一阈值邻域  $\delta$  相等,即  $\delta^l = \delta$ ,此时减小类间阈值,使得  $\delta^o < \delta$ 。可以观察到,与对象  $x$  存在邻域关系的对象数量减少了,即邻域粒子变细了。对比图 1(a)和图 1(d)可以进一步发现,减少的对象均与对象  $x$  具有不同的决策类,因此可以得出:监督邻域关系可以得到更好的邻域粒化效果。下面给出相应的监督邻域粗糙集的定义。

**定义 3** 给定一个决策信息系统  $S = (U, A, V, f)$ ,  $B \subseteq C$ ,  $\delta_B^l(x_i)$  为监督邻域集。对于任意的  $X \subseteq U$ , 监督邻域粗糙集的上、下近似集可以分别表示为:

$$\overline{NX} = \{x_i | \delta_B^l(x_i) \cap X \neq \emptyset, x_i \in U\} \quad (5)$$

$$NX = \{x_i | \delta_B^l(x_i) \subseteq X, x_i \in U\}$$

显然,关于  $X$  的边界域可以表示为:

$$BND(X) = \overline{NX} - NX \quad (6)$$

**定义 4** 给定一个决策信息系统  $S = (U, A, V, f)$ , 决策属性  $D$  将  $U$  划分为  $N$  个决策类,表示为  $D = \{D_k\}$ , 其中  $k = 1, 2, \dots, N$ 。决策  $D$  在监督邻域粗糙集中的上、下近似集可以分别表示为:

$$\overline{ND} = \bigcup_{j=1}^N \overline{ND}_j \quad (7)$$

$$\underline{ND} = \bigcup_{j=1}^N \underline{ND}_j$$

其中,正域和边界域分别表示为:

$$POS(D) = \underline{ND} \quad (8)$$

$$BND(D) = \overline{ND} - \underline{ND} \quad (9)$$

且  $POS(D) \cap BND(D) = \emptyset$ ,  $POS(D) \cup BND(D) = U$ 。

由此可见,监督邻域粗糙集把论域  $U$  划分成正域和边界域。正域越大,边界域越小,知识的不确定性就越小。为了深入研究监督邻域粗糙集对信息的不确定刻画效果,引入邻域粗糙集中知识粒度、近似质量和条件熵的定义,其中  $|\cdot|$  表示集合的基数。

**定义 5**<sup>[14]</sup> 给定一个决策信息系统  $S = (U, A, V, f)$ ,  $B \subseteq C$ , 则基于属性集  $B$  的知识粒度度量可以定义为:

$$GK(B) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} |\delta_B(x_i)| \quad (10)$$

**定义 6**<sup>[15]</sup> 给定一个决策信息系统  $S = (U, A, V, f)$ ,  $B \subseteq C$ , 决策属性  $D$  关于条件属性  $B$  的依赖度(也称为近似质量)可以定义为:

$$\gamma_B(D) = \frac{|POS_B(D)|}{|U|} \quad (11)$$

**定义 7**<sup>[15]</sup> 给定一个决策信息系统  $S = (U, A, V, f)$ ,  $B \subseteq C$ , 决策属性  $D$  关于条件属性  $B$  的条件熵可以定义为:

$$H(D|B) = -\frac{1}{|U|} \sum_{i=1}^{|U|} |D(x_i) \cap \delta_B(x_i)| \log(p(D(x_i) | \delta_B(x_i))) \quad (12)$$

其中,  $p(D(x_i) | \delta_B(x_i)) = \frac{|D(x_i) \cap \delta_B(x_i)|}{|\delta_B(x_i)|}$ ,  $D(x_i)$  为包含对象  $x_i$  的决策类。

相对于传统单一阈值的邻域粗糙集,监督邻域粗糙集在类内阈值等于传统单一阈值,减小类间阈值的情况下,邻域粒化效果会进一步增强,而且信息的不确定性会明显降低。下面给出并证明决策属性  $D$  关于条件属性  $B$  的近似质量和条件熵的单调性变化定理。

**定理 1** 在监督邻域粗糙集中,假设  $\delta_1^l = \delta_1^o = \delta$ ,  $\delta_2^l = \delta$ ,  $\delta_2^o < \delta$ , 即保持类内阈值不变,减小类间阈值,则决策属性  $D$  关于条件属性  $B$  的近似质量广义单调增加。

证明:假设一个决策信息系统  $S = (U, A, V, f)$ ,  $B \subseteq C$ ,  $\forall x_i \in U, D_j \in D, |U| = n, |D| = N$ 。当双阈值为  $\delta_1^l = \delta_1^o = \delta$  时,  $x_i$  的邻域粒子表示为  $\delta_B(x_i)$ ,  $\gamma_B^1(D)$  表示单一阈值邻域粗糙集的近似质量;当双阈值为  $\delta_2^l = \delta, \delta_2^o < \delta$  时,  $x_i$  的邻域粒子表示为  $\delta_B^l(x_i)$ ,  $\gamma_B^2(D)$  表示双阈值邻域粗糙集的近似质量。

$\forall x_i \in U, D_j \in D$ , 由已知条件显然可得  $\delta_B(x_i) \supseteq \delta_B^l(x_i)$ ,  $|\delta_B(x_i)| \geq |\delta_B^l(x_i)|$ , 且  $|\delta_B(x_i) \cap D_j| = |\delta_B^l(x_i) \cap D_j|$ 。因此有:

$$|\{x_i | \delta_B(x_i) \subseteq D_j, x_i \in U\}| \leq |\{x_i | \delta_B^l(x_i) \subseteq D_j, x_i \in U\}|$$

即  $|\underline{ND}_j^1| \leq |\underline{ND}_j^2|$ 。因此可以得到:

$$|POS_B^1(D)| = |\underline{ND}^1| \leq |POS_B^2(D)| = |\underline{ND}^2|$$

$$\gamma_B^1(D) = \frac{|POS_B^1(D)|}{|U|} \leq \gamma_B^2(D) = \frac{|POS_B^2(D)|}{|U|}$$

定理得证。

**定理 2** 在监督邻域粗糙集中,假设  $\delta_1^l = \delta_1^o = \delta$ ,  $\delta_2^l = \delta$ ,  $\delta_2^o < \delta$ , 即保持类内阈值不变,减小类间阈值,则决策属性  $D$  关于条件属性  $B$  的条件熵广义单调减小。

证明:沿用定理 1 的假设,  $H^1(D|B)$  为单一阈值邻域粗糙集的条件熵,  $H^2(D|B)$  为监督邻域粗糙集的条件熵。由定义 7 可得:

$$H^1(D|B) = -\frac{1}{|U|} \sum_{i=1}^n |D(x_i) \cap \delta_B(x_i)| \log \frac{|D(x_i) \cap \delta_B(x_i)|}{|\delta_B(x_i)|}$$

$$H^2(D|B) = -\frac{1}{|U|} \sum_{i=1}^n |D(x_i) \cap \delta_B^l(x_i)| \log \frac{|D(x_i) \cap \delta_B^l(x_i)|}{|\delta_B^l(x_i)|}$$

因为  $|D(x_i) \cap \delta_B(x_i)| = |D(x_i) \cap \delta_B^l(x_i)|$ , 所以:

$$H^2(D|B) - H^1(D|B)$$

$$= -\frac{1}{|U|} \left( \sum_{i=1}^n |D(x_i) \cap \delta_B^l(x_i)| \log \frac{|D(x_i) \cap \delta_B^l(x_i)|}{|\delta_B^l(x_i)|} - \sum_{i=1}^n |D(x_i) \cap \delta_B(x_i)| \log \frac{|D(x_i) \cap \delta_B(x_i)|}{|\delta_B(x_i)|} \right)$$

$$= -\frac{1}{|U|} \sum_{i=1}^n |D(x_i) \cap \delta_B^l(x_i)| \log \frac{|\delta_B(x_i)|}{|\delta_B^l(x_i)|}$$

因为  $|\delta_B(x_i)| \geq |\delta_B^l(x_i)|$ , 所以有  $H^2(D|B) - H^1(D|B) \leq 0$ , 即有  $H^2(D|B) \leq H^1(D|B)$ 。定理得证。

从以上两个定理可以看出,相对于传统单一阈值的邻域粗糙集,根据监督信息合理控制和调整双阈值,能够提高决策属性  $D$  关于条件属性  $B$  的近似质量,并减小条件熵,从而降低信息的不确定性。下面选用 UCI 数据集来进一步验证和分析监督邻域粗糙集在研究不确定性信息时的可行性和效果。

## 4 实验分析

为了验证所提模型的效果,在 UCI 数据集中挑选了 4 组数据,具体数据信息如表 1 所列。

表 1 4 组数据(UCI)

Table 1 Four datasets (UCI)

ID	数据集	样本个数	属性个数	类别个数
1	Wine	178	14	3
2	Sonar	208	61	2
3	German	1000	25	2
4	Statlog	2310	19	7

所有实验都在 PC 机上运行,其配置为: Intel (R) Core (TM) i5-4210U CPU @ 1.70 GHz, 12 GB 内存, Windows10 操作系统, Matlab2016a 实验平台。为充分验证模型的效果,采用以下两种方法开展相关实验。方法一:设定监督邻域类内阈值  $\delta^I$  与传统单一邻域阈值  $\delta$  相等,即  $\delta^I = \delta = 0.3$ ,选取 5

组监督邻域类间阈值  $\delta^O$ : 0.25, 0.26, 0.27, 0.28, 0.29, 0.3。把 4 个 UCI 数据集分别按属性数量分成 10 个相等的子数据集,把第一个数据子集作为测试集 1,前两个数据子集累加作为测试集 2,以此类推,得到 10 个测试子集,对 5 组监督邻域阈值在 10 个测试子集上分别计算邻域近似质量和邻域条件熵。方法二:选取 10 组类内阈值  $\delta^I = 0.21, 0.22, \dots, 0.3$ ,在每组中设置类间阈值  $\delta^O$  分别等于类内阈值的 0.5, 0.6, 0.7, 0.8, 0.9, 1 倍,对 4 个 UCI 数据集分别计算不同双阈值设置下的邻域近似质量和邻域条件熵。具体实验结果如图 2 和图 3 所示。

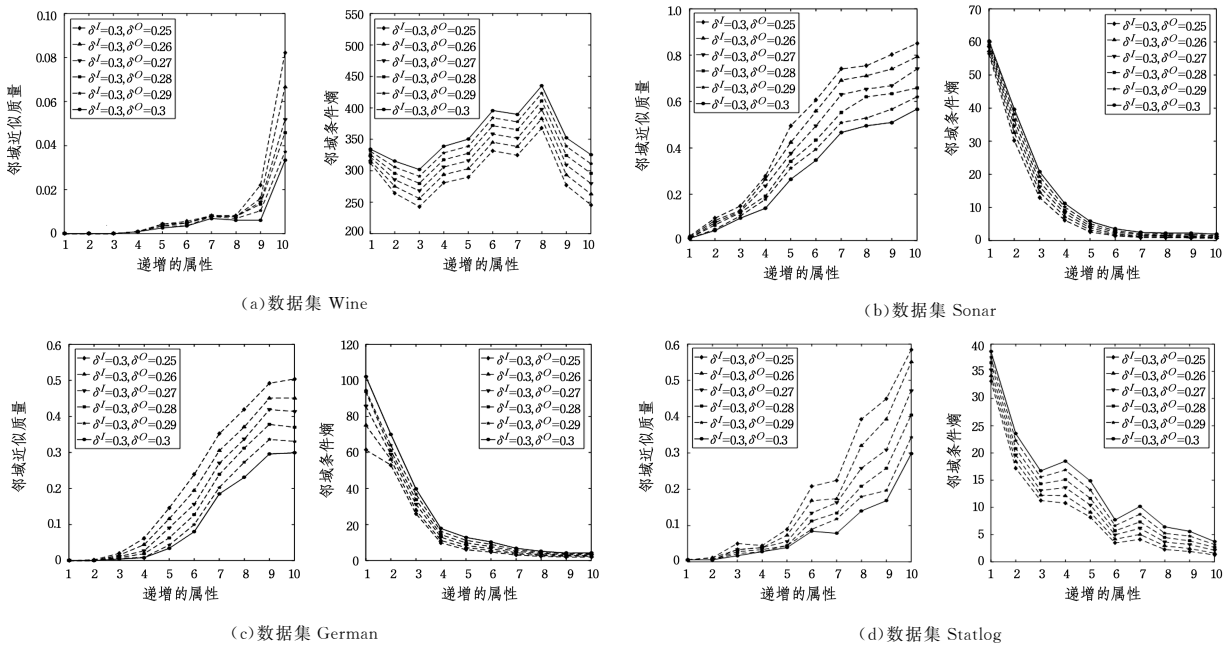


图 2 在属性递增情况下的不确定性刻画对比分析

Fig. 2 Contrastive analysis with description of uncertainty under addition of attributes

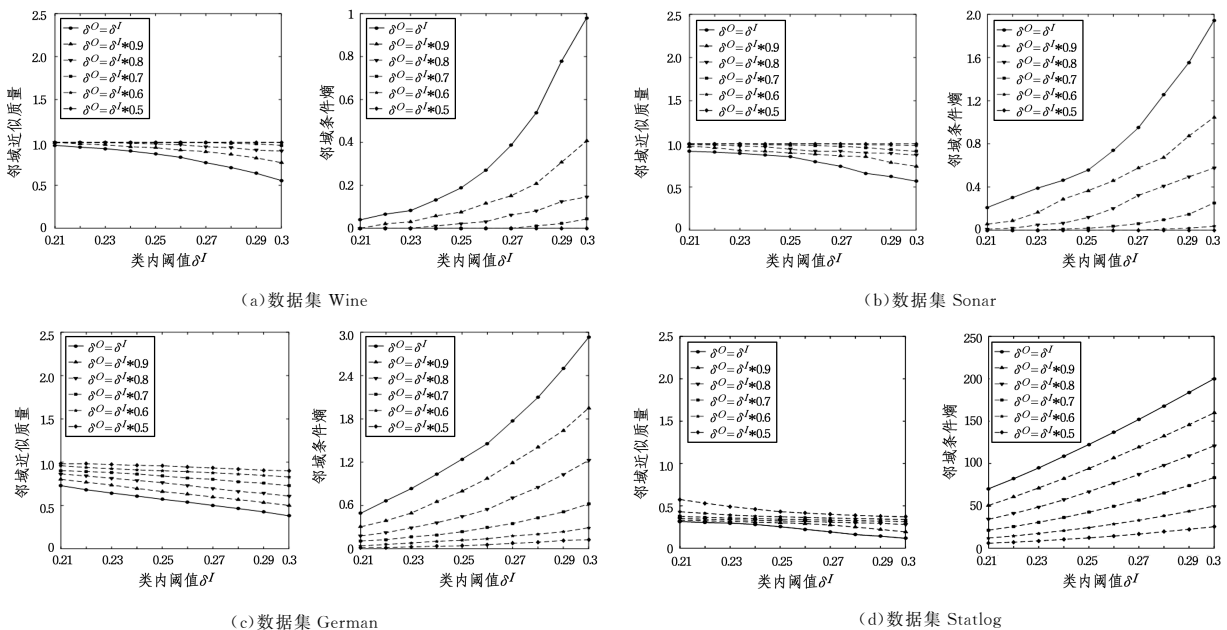


图 3 在阈值变化情况下的不确定性刻画对比分析

Fig. 3 Contrastive analysis with description of uncertainty under change of thresholds

在图2中,保持类内阈值与传统单一阈值相等,逐渐减小类间阈值。可以发现,4个数据集的邻域近似质量单调增加,而邻域条件熵单调减小,即知识的不确定性下降。另外,随着属性数量的逐渐增多,4个数据集的邻域近似质量和邻域条件熵并没有呈现出单调性变化,这是由于在属性变化下采用了多次归一化的实验方法。通过观察图2发现,监督邻域粗糙集在粒度变化情况下可以通过保持类内阈值不变而调整类间阈值来实现对邻域粒化效果和不确定性的有效控制。

由图3可知,在不同的类内阈值下,逐渐减小类间阈值。可以发现,随着类间阈值的减小,4个数据集的邻域近似质量呈现单调增加的趋势,而邻域条件熵呈现单调减小的趋势,知识不确定性也有所下降。另外,随着类内阈值的增大,4个数据集的邻域近似质量单调减小,而邻域条件熵单调增加。通过图3可以发现,在数据集不变的情况下,通过保持类内阈值不变而调整类间阈值也能实现对邻域粒化效果和不确定性的有效控制。

**结束语** 传统的邻域粗糙集使用单一邻域阈值,大大限制了对论域粒化以及降低信息不确定性的能力。本文借助监督学习中对象已有或预测的标记信息,引入类内阈值和类间阈值的概念,提出了一种监督邻域粗糙集。该模型是传统邻域粗糙集的推广形式,可以通过保持类内邻域不变和减小类间邻域来优化邻域粒子群。实验结果显示,该模型在粗糙近似质量和条件信息熵两种度量下均能降低信息的不确定性。下一步工作将在监督邻域系统下对属性约简开展深入探讨和研究。

### 参 考 文 献

- [1] PAWLAK Z. Rough sets[J]. International Journal of Computer & Information Sciences, 1982, 11(5): 341-356.
- [2] LIN T Y. Granular computing: practices, theories, and future directions [M]. New York: Springer, 2009: 4339-4355.
- [3] LIN T Y. Granular computing on binary relation I: Date mining and neighborhood systems [J]. Rough Sets in Knowledge Discovery, 1998(2): 165-166.
- [4] YAO Y Y. Relational interpretation of neighborhood operators and rough set approximation operators [J]. Information Sciences, 1998, 111(198): 239-259.
- [5] WU W Z, ZHANG W X. Neighborhood operator systems and approximations [J]. Information Sciences, 2002, 144(1-4): 201-217.
- [6] HU Q H, YU D R, XIE Z X. Neighborhood classifiers [J]. Expert Systems with Applications, 2008, 34(2): 866-876.
- [7] HU Q H, PEDRYCZ W, YU D R, et al. Selecting discrete and continuous features based on neighborhood decision error minimization [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2010, 40(1): 137-150.
- [8] DUAN J, HU Q H, ZHANG L J, et al. Feature selection for multi-label classification based on neighborhood rough sets [J]. Journal of Computer Research and Development, 2015, 52(1): 56-65.
- [9] CHEN D G, LI W L, ZHANG X, et al. Evidence-theory-based numerical algorithms of attribute reduction with neighborhood-covering rough sets [J]. International Journal of Approximate Reasoning, 2014, 55(3): 908-923.
- [10] YANG X B, ZHANG M, DOU H L, et al. Neighborhood Systems-Based Rough Sets in Incomplete Information System [J]. Knowledge-Based Systems, 2011, 24(6): 858-867.
- [11] YANG X B, CHEN Z H, DOU H L, et al. Neighborhood system based rough set: models and attribute reductions [J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2012, 20(3): 399-419.
- [12] 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.
- [13] CHEN H M, LI T R, CAI Y, et al. Parallel attribute reduction in dominance-based neighborhood rough set [J]. Information Sciences, 2016, 373: 351-368.
- [14] LIANG J Y, SHI Z, LI D, et al. Information entropy, rough entropy and knowledge granulation in incomplete information systems [J]. International Journal of General Systems, 2006, 35(6): 641-654.
- [15] HU Q H, GUO M Z, YU D R, et al. Information entropy for ordinal classification [J]. Science China Information Sciences, 2010, 53(6): 1188-1200.
- [16] DO H, ELBAUM S, ROTHERMEL G. Supporting Controlled Experimentation with Testing Techniques: An Infrastructure and its Potential Impact [J]. Empirical Software Engineering, 2005, 10(4): 405-435.
- [17] ALETI A, GRUNSKIE L. Test data generation with a Kalman filter-based adaptive genetic algorithm [J]. Journal of Systems & Software, 2015, 103(C): 343-352.
- [18] SCHAFFER J D, CARUANA R A, ESHELMAN L J, et al. A study of control parameters affecting online performance of genetic algorithms for function optimization [C] // International Conference on Genetic Algorithms, George Mason University, Fairfax, Virginia, USA. DBLP, 1989: 51-60.
- [15] NIRPAL P B, KALE K V. Using Genetic Algorithm for Automated Efficient Software Test Case Generation for Path Testing [J]. International Journal of Advanced Networking & Applications, 2011, 2(6): 911-915.
- [16] MCMINN P. Evolutionary Search for Test Data in the Presence of State Behaviour [J]. University of Sheffield, 2005, 16(12): 41-46.
- [17] PACHAURI A, SRIVASTAVA G. Automated test data generation for branch testing using genetic algorithm: An improved approach using branch ordering, memory and elitism [J]. Journal of Systems & Software, 2013, 86(5): 1191-1208.

(上接第178页)