

# 不完备邻域多粒度决策理论粗糙集与三支决策

刘丹 徐立新 李敬伟

(河南工学院计算机科学与技术学院 河南 新乡 453003)

**摘要** 多粒度决策理论粗糙集是多粒度视角下三支决策中一种重要的模型。在数值型不完备数据下建立邻域容差关系;在其基础上提出乐观和悲观的邻域多粒度决策理论粗糙集模型。为了弥补这两种模型的局限,提出平均邻域多粒度决策理论粗糙集模型,并分析相关性质以及相互关系。同时为了使所提出的邻域多粒度决策理论粗糙集适用于不完备数据环境,运用区间值的形式表示代价函数,并通过选取不同参数的方式提出一种可变三支决策。实例分析表明,该模型与方法具有一定的合理性与灵活性。

**关键词** 决策理论粗糙集 多粒度 邻域 不完备信息系统 三支决策

中图分类号 TP18

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2019.05.026

## INCOMPLETE NEIGHBORHOOD MULTI-GRANULATION DECISION-THEORETIC ROUGH SET AND THREE-WAY DECISION

Liu Dan Xu Lixin Li Jingwei

(College of Computer Science and Technology, Henan Institute of Technology, Xinxiang 453003, Henan, China)

**Abstract** Multi-granulation decision-theoretic rough set is an important model in three-way decision under multi-granulation perspective. The neighborhood tolerance relation was established under the numerical incomplete data, and we proposed optimistic and pessimistic neighborhood multi-granulation decision-theoretic rough set model. To make up for these two models' limitation, we further proposed the mean neighborhood multi-granulation decision-theoretic rough set model, and analyzed the related properties and relations. In order to make the proposed model suitable for incomplete data environment, we used interval form to express cost functions, and proposed a variable three-way decision by selecting different parameters. The example analysis shows that the model and method have certain rationality and flexibility.

**Keywords** Decision-theoretic rough set Multi-granulation Neighborhood Incomplete information system Three-way decision

## 0 引言

粗糙集理论<sup>[1]</sup>是由波兰学者 Pawlak 提出的一种不确定性数据的分析工具,目前已广泛地运用于分类学习、模式识别以及数据挖掘等领域<sup>[2-4]</sup>。同时粗糙集理论经过几十年的研究,已在不同的领域和不同背景下进行了相关的改进与拓展,诸如概率粗糙集<sup>[5-6]</sup>、变精度粗糙集<sup>[7]</sup>、邻域粗糙集<sup>[8-9]</sup>和模糊粗糙集<sup>[10]</sup>

等。在这些拓展的粗糙集模型中,概率粗糙集是一种常用的粗糙集模型,它通过一对阈值来对经典粗糙集模型进行改进<sup>[5-6]</sup>,因而使其具有容忍噪声数据的特性,目前已引起学者们的广泛研究<sup>[11-12]</sup>。Yao 等<sup>[13]</sup>通过代价的角度来对概率粗糙集模型中的一对阈值进行学习,提出了决策理论粗糙集模型,并且基于该模型诱导出了一种特殊的决策方式,被称之为三支决策<sup>[14]</sup>。三支决策相当于对传统的二支决策进行扩展,它通过最小化代价的视角进行决策,其决策的结果分

为三种情形,分别为接受、延迟和拒绝,这种形式开辟了一种崭新的决策方法,在其他领域获得了大量的应用<sup>[15-17]</sup>。目前决策理论粗糙集和三支决策已成为粗糙集理论研究的热点。

根据粒计算<sup>[18]</sup>的视角,经典粗糙集模型以及各类的扩展模型都可以称之为单粒度下的粗糙集模型。然而在现实的应用环境下,需要通过多层次和多视角对研究对象进行全方位的数据挖掘,因此 Qian 等<sup>[19]</sup>提出了多粒度粗糙集模型,为粗糙集理论的研究提供了全新的思路。为了在多粒度粗糙集模型中融入代价学习, Qian 等<sup>[20]</sup>又提出了基于多粒度的决策理论粗糙集模型,该模型在多粒度数据挖掘和三支决策方面都是一种有效的工具。然而现实的应用环境是复杂多样的,例如数值型数据和含缺失值数据的存在,使得传统的多粒度决策理论粗糙集模型面临着一定的挑战。近年来, Li 等<sup>[21]</sup>提出基于邻域关系的决策理论粗糙集模型,为决策理论粗糙集模型处理数值型数据提供了一种有效的解决方案。而对于不完备数据,基于容差关系的粗糙集理论是一种常用的处理方法<sup>[8,22]</sup>。本文针对不完备的数值型数据,提出一种邻域多粒度的决策理论粗糙集模型。首先在不完备数值型数据中建立邻域容差关系,然后在其基础上提出邻域多粒度决策理论粗糙集模型,包括乐观邻域多粒度决策理论粗糙集和悲观邻域多粒度决策理论粗糙集。由于这两种模型在粗糙逼近刻画方面存在过于宽松和过于苛刻的局限性<sup>[23]</sup>,因此又提出了一种平均邻域多粒度决策理论粗糙集,克服了两种模型的缺点,并分析了三种模型之间的相互关系。

三支决策建立在决策理论粗糙集模型的基础上,其中模型中的阈值直接由给定的代价函数来确定<sup>[14]</sup>,通常代价函数都是具体的数值。然而在不完备信息系统中,由于缺失值的存在,信息系统的不确定性将会进一步增大<sup>[24]</sup>,但是在传统的三支决策模型中,代价函数是一个具体的取值<sup>[25]</sup>。因此本文在不完备邻域多粒度决策理论粗糙集模型的基础上,通过区间值的形式来表示代价函数,然后通过一个参数来确定具体的代价函数值,从而诱导出对应的三支决策结果,本文称之为可变三支决策,理论分析表明了该方法的正确性。最后本文通过一个具体的不完备数值型信息表的实例来分析所提出模型的有效性 with 优越性,同时对模型中的有关参数进行了分析与讨论。

## 1 基本理论

### 1.1 邻域多粒度粗糙集模型

通常,粗糙集理论<sup>[1]</sup>研究的数据对象被称为信息系统,一个信息系统表示为  $IS = (U, At, V)$ ,这里的  $U$  称为论域,是一个非空有限对象集,  $At$  为全体属性集,  $V$  为所有属性的值域。当  $At = C \cup D$ , 且  $C \cap D = \emptyset$ , 其中  $C$  为条件属性集,  $D$  为决策属性集,那么此信息系统又称为决策信息系统。当信息系统的条件属性值均为连续型时,这类信息系统又称为邻域信息系统<sup>[9]</sup>。

经典粗糙集模型是处理不确定数据的一种有效工具,它通过等价关系对论域空间进行知识划分,从而对不确定性的目标对象进行粗糙逼近<sup>[1]</sup>,目前在机器学习和数据挖掘等领域发挥着重要的作用<sup>[4,8-10,15-16]</sup>。在粒计算视角下,经典粗糙集理论相当于以一种单粒度的视角对目标概念进行分析和描述,为了从多个视角来处理和分析问题, Qian 等<sup>[19]</sup>提出了多粒度粗糙集模型。为了使多粒度粗糙集模型能够处理邻域信息系统, Lin 等<sup>[26]</sup>又提出了邻域多粒度粗糙集模型。

**定义 1**<sup>[9]</sup> 考虑一个邻域信息系统  $IS = (U, At, V)$ ,  $B \subseteq At$  在论域  $U$  上诱导的邻域关系定义为:

$$N_B^\delta = \{ (x, y) \in U \times U \mid d_B(x, y) \leq \delta \} \quad (1)$$

式中:  $\delta$  是一非负常数,称为邻域半径,  $d_B(x, y)$  称为对象  $x$  与  $y$  在属性集  $B$  下的距离,常用的距离公式采用闵可夫斯基距离<sup>[9]</sup>。可以看出,邻域关系  $N_B^\delta$  满足自反性和对称性,但不满足传递性。同时,  $\forall x \in U$  在论域上关于邻域关系  $N_B^\delta$  的邻域定义为  $n_B^\delta(x) = \{ y \in U \mid (x, y) \in N_B^\delta \}$ 。

在不引起混淆的情形下,下文适当略去邻域半径  $\delta$  的标记。

**定义 2**<sup>[26]</sup> 考虑一个邻域信息系统  $IS = (U, At, V)$ ,  $B_1, B_2, \dots, B_m \subseteq At$  是一个包含  $m$  个属性子集的属性子集族,它们诱导出的邻域关系分别为  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$ 。对于目标近似对象集  $X \subseteq U$ ,基于这  $m$  个邻域关系的乐观邻域多粒度粗糙集的下近似和上近似分别定义为:

$$\underline{\sum_{i=1}^m N_{B_i}^o(X)} = \{ x \mid n_{B_1}(x) \subseteq X \vee n_{B_2}(x) \subseteq X \vee \dots \vee n_{B_m}(x) \subseteq X \} \quad (2)$$

$$\overline{\sum_{i=1}^m N_{B_i}^o(X)} = \sim \underline{\sum_{i=1}^m N_{B_i}^o(\sim X)} \quad (3)$$

称  $\left\langle \underline{\sum_{i=1}^m N_{B_i}^o(X)}, \overline{\sum_{i=1}^m N_{B_i}^o(X)} \right\rangle$  为  $X$  关于  $m$  个邻域关

系  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$  的乐观邻域多粒度粗糙集。同时,  $X$  关于  $m$  个邻域关系的乐观邻域多粒度正区域, 负区域和边界域分别定义为:

$$POS_{\sum_{i=1}^m N_{B_i}}^O(X) = \overline{\sum_{i=1}^m N_{B_i}^O(X)} \quad (4)$$

$$NEG_{\sum_{i=1}^m N_{B_i}}^O(X) = U - \overline{\sum_{i=1}^m N_{B_i}^O(X)} \quad (5)$$

$$BUN_{\sum_{i=1}^m N_{B_i}}^O(X) = \overline{\sum_{i=1}^m N_{B_i}^O(X)} - \sum_{i=1}^m N_{B_i}^O(X) \quad (6)$$

**定义 3**<sup>[26]</sup> 考虑邻域信息系统  $IS = (U, At, V)$ ,  $B_1, B_2, \dots, B_m \subseteq At$  是一个包含  $m$  个属性子集的属性子集族, 它们诱导出的邻域关系分别为  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$ 。对于目标近似对象集  $X \subseteq U$ , 基于这  $m$  个邻域关系的悲观邻域多粒度粗糙集的下近似和上近似分别定义为:

$$\underbrace{\sum_{i=1}^m N_{B_i}^P(X)}_{X \wedge \dots \wedge n_{B_m}(x) \subseteq X} = \{x \mid n_{B_1}(x) \subseteq X \wedge n_{B_2}(x) \subseteq X \wedge \dots \wedge n_{B_m}(x) \subseteq X\} \quad (7)$$

$$\overline{\sum_{i=1}^m N_{B_i}^P(X)} = \sim \underbrace{\sum_{i=1}^m N_{B_i}^P(\sim X)} \quad (8)$$

称  $\left\langle \underbrace{\sum_{i=1}^m N_{B_i}^P(X)}, \overline{\sum_{i=1}^m N_{B_i}^P(X)} \right\rangle$  为  $X$  关于  $m$  个邻域关系  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$  的悲观邻域多粒度粗糙集。同时,  $X$  关于  $m$  个邻域关系的悲观邻域多粒度正区域, 负区域和边界域分别定义为:

$$POS_{\sum_{i=1}^m N_{B_i}}^P(X) = \overline{\sum_{i=1}^m N_{B_i}^P(X)} \quad (9)$$

$$NEG_{\sum_{i=1}^m N_{B_i}}^P(X) = U - \overline{\sum_{i=1}^m N_{B_i}^P(X)} \quad (10)$$

$$BUN_{\sum_{i=1}^m N_{B_i}}^P(X) = \overline{\sum_{i=1}^m N_{B_i}^P(X)} - \sum_{i=1}^m N_{B_i}^P(X) \quad (11)$$

目前邻域多粒度粗糙集模型已成为多粒度粗糙集领域的一种重要的模型。

### 1.2 决策理论粗糙集模型

Yao 等<sup>[5,13-14]</sup> 以代价的理论为基础, 通过贝叶斯决策过程来确定概率粗糙集中的两个阈值参数, 推导出具有最小代价化的决策规则, 即三支决策。

对于贝叶斯决策方法, 定义一个非空有限状态集为  $\Omega = \{\omega_1, \omega_2, \dots, \omega_s\}$ , 一个非空有限行为集为  $\Gamma = \{a_1, a_2, \dots, a_t\}$ , 对于  $\forall x \in U$ , 设对象  $x$  处于状态  $\omega_j$  ( $1 \leq j \leq s$ ) 的条件概率为  $P(\omega_j | x)$ , 令  $\lambda(a_i | \omega_j)$  表示  $x$  处于状态  $\omega_j$  采取行为  $a_i$  ( $1 \leq i \leq t$ ) 时的代价函数(损失函数)。对于对象  $x$ , 可以得到它采取行为  $a_i$  的预期代价, 表示为:

$$R(a_i | x) = \sum_{j=1}^s \lambda(a_i | \omega_j) \cdot P(\omega_j | x) \quad (12)$$

在决策理论粗糙集模型中, 所研究的数据对象为信息系统, 那么对象的状态集可表示为  $\Omega = \{X, \sim X\}$ , 行为集可表示为  $\Gamma = \{a_p, a_B, a_N\}$ , 其中  $a_p$  表示对象被分类到  $X$  的正区域,  $a_B$  表示对象被分类到  $X$  的边界域,  $a_N$  表示对象被分类到  $X$  的负区域。由于决策理论粗糙集模型中只包含两种状态和三种行为, 因此可以通过表格的形式去展示出每种状态采取每种行为的代价函数, 具体如表 1 所示。

表 1 所有形式的代价函数

$\Gamma$	$X$	$\sim X$
$a_p$	$\lambda_{pp}$	$\lambda_{pN}$
$a_B$	$\lambda_{BP}$	$\lambda_{BN}$
$a_N$	$\lambda_{NP}$	$\lambda_{NN}$

在表 1 中,  $\lambda_{pp}, \lambda_{BP}$  和  $\lambda_{NP}$  分别表示对象  $x$  隶属于  $X$  时分别采取  $a_p, a_B$  和  $a_N$  时的代价,  $\lambda_{pN}, \lambda_{BN}$  和  $\lambda_{NN}$  分别表示对象  $x$  不隶属于  $X$  时分别采取  $a_p, a_B$  和  $a_N$  时的代价, 通过条件概率  $P(X | [x])$  表示对象  $x$  隶属于  $X$  的程度, 这里的  $[x]$  表示对象  $x$  的等价类。因此可以得到对象  $x$  采取三种行为的预期代价, 分别表示为:

$$R(a_p | [x]) = \lambda_{pp} \cdot P(X | [x]) + \lambda_{pN} \cdot P(\sim X | [x]) \quad (13)$$

$$R(a_B | [x]) = \lambda_{BP} \cdot P(X | [x]) + \lambda_{BN} \cdot P(\sim X | [x]) \quad (14)$$

$$R(a_N | [x]) = \lambda_{NP} \cdot P(X | [x]) + \lambda_{NN} \cdot P(\sim X | [x]) \quad (15)$$

根据贝叶斯决策方法, 可以得到如下三个最小化代价的决策规则:

- (P) 如果  $R(a_p | [x]) < R(a_B | [x])$ , 并且  $R(a_p | [x]) < R(a_N | [x])$ , 那么  $x \in POS(X)$ ;
- (B) 如果  $R(a_B | [x]) < R(a_p | [x])$ , 并且  $R(a_B | [x]) < R(a_N | [x])$ , 那么  $x \in BUN(X)$ ;
- (N) 如果  $R(a_N | [x]) < R(a_p | [x])$ , 并且  $R(a_N | [x]) < R(a_B | [x])$ , 那么  $x \in NEG(X)$ 。

由于  $P(X | [x]) + P(\sim X | [x]) = 1$ , 这里 Yao 等<sup>[14]</sup> 假设  $0 \leq \lambda_{pp} \leq \lambda_{BP} \leq \lambda_{NP}$  和  $0 \leq \lambda_{NN} \leq \lambda_{BN} \leq \lambda_{pN}$ , 那么决策规则 (P)、(B) 和 (N) 可以进一步表示为:

(P) 如果  $P(X | [x]) \geq \alpha$  且  $P(X | [x]) \geq \gamma$ , 那么  $x \in POS(X)$ ;

(B) 如果  $P(X | [x]) < \alpha$  且  $P(X | [x]) < \beta$ , 那么  $x \in BUN(X)$ ;

(C) 如果  $P(X | [x]) < \gamma$  且  $P(X | [x]) \leq \beta$ , 那么

$x \in NEG(X)$ 。

其中:

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} \quad (16)$$

$$\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})} \quad (17)$$

$$\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})} \quad (18)$$

由于  $0 \leq \beta < \gamma < \alpha \leq 1$ , 因此:

(P) 如果  $P(X | [x]) \geq \alpha$ , 那么  $x \in POS(X)$ ;

(B) 如果  $\beta < P(X | [x]) < \alpha$ , 那么  $x \in BUN(X)$ ;

(N) 如果  $P(X | [x]) \leq \beta$ , 那么  $x \in NEG(X)$ 。

根据以上推导, Yao 等给出了决策理论粗糙集模型上下近似的定义。

**定义 4**<sup>[14]</sup> 对于信息系统  $IS = (U, At, V)$ , 设代价函数如表 1 所示, 令  $B \subseteq At$ , 那么  $\forall X \subseteq U$  关于  $B$  的决策理论粗糙集模型的下近似与上近似分别定义为:

$$\underline{B}^{(\alpha, \beta)}(X) = \{x | P(X | [x]) \geq \alpha\} \quad (19)$$

$$\overline{B}^{(\alpha, \beta)}(X) = \{x | P(X | [x]) > \beta\} \quad (20)$$

同时,  $X$  关于  $B$  的决策理论正区域、边界域和负区域分别定义为:

$$POS_B^{(\alpha, \beta)}(X) = \underline{B}^{(\alpha, \beta)}(X) \quad (21)$$

$$BUN_B^{(\alpha, \beta)}(X) = \overline{B}^{(\alpha, \beta)}(X) - \underline{B}^{(\alpha, \beta)}(X) \quad (22)$$

$$NEG_B^{(\alpha, \beta)}(X) = U - \overline{B}^{(\alpha, \beta)}(X) \quad (23)$$

## 2 不完备邻域多粒度决策理论粗糙集

Lin 等<sup>[26]</sup>提出的邻域多粒度粗糙集模型只能适用于完备型的邻域信息系统。现实中包含缺失属性值的邻域信息系统大量存在<sup>[8]</sup>, 为了解决 Lin 的邻域多粒度粗糙集模型的局限性, 可以考虑将该模型在不完备邻域信息系统进行拓展。近年来, Li 等<sup>[21]</sup>提出了基于邻域关系的决策理论粗糙集模型, 因此在此基础上, 本文进一步提出不完备邻域多粒度决策理论粗糙集模型。

**定义 5**<sup>[21]</sup> 考虑邻域信息系统  $IS = (U, At, V)$ , 对于  $\forall x \in U$  和  $X \subseteq U$ ,  $x$  在  $B \subseteq At$  下的邻域为  $n_B^\delta(x)$ , 设条件概率  $P(X | n_B^\delta(x)) = \frac{|n_B^\delta(x) \cap X|}{|n_B^\delta(x)|}$ , 那么  $X$  关于  $B$  的邻域决策理论粗糙集的下近似和上近似分别为:

$$\underline{N}_B^{(\alpha, \beta)}(X) = \{x | P(X | n_B^\delta(x)) \geq \alpha\} \quad (24)$$

$$\overline{N}_B^{(\alpha, \beta)}(X) = \{x | P(X | n_B^\delta(x)) > \beta\} \quad (25)$$

这里的  $\alpha, \beta$  诱导方式同 1.2 节。类似地,  $X$  关于  $B$  的邻域决策理论粗糙集正区域、边界域和负区域分别定

义为:

$$POS_{N_B}^{(\alpha, \beta)}(X) = \underline{N}_B^{(\alpha, \beta)}(X) \quad (26)$$

$$BUN_{N_B}^{(\alpha, \beta)}(X) = \overline{N}_B^{(\alpha, \beta)}(X) - \underline{N}_B^{(\alpha, \beta)}(X) \quad (27)$$

$$NEG_{N_B}^{(\alpha, \beta)}(X) = U - \overline{N}_B^{(\alpha, \beta)}(X) \quad (28)$$

Qian 等<sup>[20]</sup>提出了在完备型信息系统中提出了多粒度决策理论粗糙集模型, 在 Qian 的基础上结合定义 5, 提出一种不完备邻域多粒度决策理论粗糙集模型。

在提出不完备邻域多粒度决策理论粗糙集模型前, 需要首先确定不完备邻域信息系统中邻域的定义, 本文这里采用 Zhao 等<sup>[27]</sup>提出的邻域容差关系。

**定义 6**<sup>[27]</sup> 称  $IIS = (U, At, V)$  为不完备邻域信息系统, 利用“\*”表示缺失的属性值, 即  $* \in V$ 。设邻域半径为  $\delta$ , 定义  $B \subseteq At$  在论域  $U$  上的邻域容差关系为:

$$NT_B^\delta = \{(x, y) \in U \times U | a(x) = * \vee a(y) = * \vee d_{|a|}(x, y) \leq \delta, \forall a \in B\} \quad (29)$$

同时,  $\forall x \in U$  在邻域容差关系  $NT_B^\delta$  下的邻域容差类为  $nt_B^\delta(x) = \{y | (x, y) \in NT_B^\delta\}$ 。下文中, 将邻域容差类简称为邻域。

### 2.1 不完备乐观邻域多粒度决策理论粗糙集模型

根据 Qian 等<sup>[20]</sup>学者提出的乐观多粒度决策理论粗糙集模型, 在本节中, 提出不完备信息系统下的乐观邻域多粒度决策理论粗糙集模型。

**定义 7** 对于不完备邻域信息系统  $IIS = (U, At, V)$ ,  $B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 它们诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于目标近似对象集  $X \subseteq U$ ,  $X$  关于  $m$  个邻域容差关系的乐观邻域多粒度决策理论粗糙集的下近似和上近似分别定义为:

$$\underline{\sum_{i=1}^m N_{B_i \alpha}^0(X)} = \{x | P(X | nt_{B_1}(x)) \geq \alpha \vee P(X | nt_{B_2}(x)) \geq \alpha \vee \dots \vee P(X | nt_{B_m}(x)) \geq \alpha\} \quad (30)$$

$$\overline{\sum_{i=1}^m N_{B_i \beta}^0(X)} = U - \{x | P(X | nt_{B_1}(x)) \leq \beta \wedge P(X | nt_{B_2}(x)) \leq \beta \wedge \dots \wedge P(X | nt_{B_m}(x)) \leq \beta\} \quad (31)$$

称  $\left\langle \underline{\sum_{i=1}^m N_{B_i \alpha}^0(X)}, \overline{\sum_{i=1}^m N_{B_i \alpha}^0(X)} \right\rangle$  为  $X$  关于  $m$  个邻域容差关系  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$  的乐观邻域多粒度决策理论粗糙集。这里的  $P(X | nt_{B_2}(x))$  表示条件概率, 满足  $P(X | nt_{B_2}(x)) = \frac{|nt_{B_2}(x) \cap X|}{|nt_{B_2}(x)|}$ 。 $\alpha$  和  $\beta$  是一对阈值。

同时,  $X$  关于  $m$  个邻域容差关系的乐观邻域多粒度决策理论粗糙集的正区域、边界域和负区域分别定义为:

$$POS_{\sum_{i=1}^m N_{B_i}}^O(X) = \sum_{i=1}^m \overline{N_{B_i\alpha}}^O(X) \quad (32)$$

$$BUN_{\sum_{i=1}^m N_{B_i}}^O(X) = \sum_{i=1}^m \overline{N_{B_i\beta}}^O(X) - \sum_{i=1}^m \overline{N_{B_i\alpha}}^O(X) \quad (33)$$

$$NEG_{\sum_{i=1}^m N_{B_i}}^O(X) = U - \sum_{i=1}^m \overline{N_{B_i\beta}}^O(X) \quad (34)$$

根据定义 7 中关于乐观邻域多粒度决策理论粗糙集模型的定义以及文献 [19 - 20], 可以得到如下性质。

**性质 1** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $\forall X \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^O(X) \supseteq \overline{N_{B_i\alpha}}(X) \quad 1 \leq i \leq m \quad (35)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}}^O(X) \subseteq \overline{N_{B_i\beta}}(X) \quad 1 \leq i \leq m \quad (36)$$

这里的  $\overline{N_{B_i\alpha}}(X) = \{x \mid P(X \mid nt_{B_i}(x)) \geq \alpha\}$ , 并且  $\overline{N_{B_i\beta}}(X) = \{x \mid P(X \mid nt_{B_i}(x)) \geq \beta\}$ 。

**性质 2** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $\forall X \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^O(X) = \bigcup_{i=1}^m \overline{N_{B_i\alpha}}(X) \quad (37)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}}^O(X) = \bigcap_{i=1}^m \overline{N_{B_i\beta}}(X) \quad (38)$$

式中:  $\overline{N_{B_i\alpha}}(X)$  和  $\overline{N_{B_i\beta}}(X)$  的定义同性质 1。

**性质 3** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $X_1 \subseteq X_2 \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^O(X_1) \subseteq \overline{\sum_{i=1}^m N_{B_i\alpha}}^O(X_2) \quad (39)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}}^O(X_1) \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}}^O(X_2) \quad (40)$$

## 2.2 不完备悲观邻域多粒度决策理论粗糙集模型

根据 Qian 等<sup>[20]</sup>学者提出的悲观多粒度决策理论粗糙集模型, 这里同样提出不完备信息系统下的悲观邻域多粒度决策理论粗糙集模型。

**定义 8** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于目标近似对象集  $X \subseteq U$ ,  $X$  关于  $m$  个邻域容差关系的悲观邻域多粒度决策理论粗糙集的下近似和上近似分别定义为:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^P(X) = \{x \mid P(X \mid nt_{B_1}(x)) \geq \alpha \wedge P(X \mid nt_{B_2}(x)) \geq \alpha \wedge \dots \wedge P(X \mid nt_{B_m}(x)) \geq \alpha\} \quad (41)$$

$$\underline{\sum_{i=1}^m N_{B_i\beta}}^P(X) = U - \{x \mid P(X \mid nt_{B_1}(x)) \leq \beta \vee P(X \mid nt_{B_2}(x)) \leq \beta \vee \dots \vee P(X \mid nt_{B_m}(x)) \leq \beta\} \quad (42)$$

称  $\left\langle \overline{\sum_{i=1}^m N_{B_i\alpha}}^P(X), \underline{\sum_{i=1}^m N_{B_i\alpha}}^P(X) \right\rangle$  为  $X$  关于  $m$  个邻域容差关系  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$  的悲观邻域多粒度决策理论粗糙集。同理,  $X$  关于  $m$  个邻域容差关系的悲观邻域多粒度决策理论粗糙集的正区域、边界域和负区域分别定义为:

$$POS_{\sum_{i=1}^m N_{B_i}}^P(X) = \sum_{i=1}^m \overline{N_{B_i\alpha}}^P(X) \quad (43)$$

$$BUN_{\sum_{i=1}^m N_{B_i}}^P(X) = \sum_{i=1}^m \overline{N_{B_i\beta}}^P(X) - \sum_{i=1}^m \overline{N_{B_i\alpha}}^P(X) \quad (44)$$

$$NEG_{\sum_{i=1}^m N_{B_i}}^P(X) = U - \sum_{i=1}^m \overline{N_{B_i\beta}}^P(X) \quad (45)$$

同样地, 悲观邻域多粒度决策理论粗糙集也满足如下性质。

**性质 4** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $\forall X \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^P(X) \subseteq \overline{N_{B_i\alpha}}(X) \quad 1 \leq i \leq m \quad (46)$$

$$\underline{\sum_{i=1}^m N_{B_i\beta}}^P(X) \supseteq \underline{N_{B_i\beta}}(X) \quad 1 \leq i \leq m \quad (47)$$

式中:  $\overline{N_{B_i\alpha}}(X)$  和  $\underline{N_{B_i\beta}}(X)$  的定义同性质 1。

**性质 5** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $\forall X \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}}^P(X) = \bigcap_{i=1}^m \overline{N_{B_i\alpha}}(X) \quad (48)$$

$$\underline{\sum_{i=1}^m N_{B_i\beta}}^P(X) = \bigcup_{i=1}^m \underline{N_{B_i\beta}}(X) \quad (49)$$

式中： $\overline{N_{B_i\alpha}}(X)$ 和 $\overline{N_{B_i\beta}}(X)$ 的定义同性质 1。

**性质 6** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $X_1 \subseteq X_2 \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}^P(X_1)} \subseteq \overline{\sum_{i=1}^m N_{B_i\alpha}^P(X_2)} \quad (50)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}^P(X_1)} \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}^P(X_2)} \quad (51)$$

### 2.3 不完备平均邻域多粒度决策理论粗糙集模型

根据定义 7, 在乐观邻域多粒度决策理论粗糙集中, 下近似集中的元素只需任意满足条件概率大于阈值  $\alpha$ , 而对于定义 8 中的悲观邻域多粒度决策理论粗糙集, 其下近似集的元素必须满足所有的条件概率大于阈值  $\alpha$ , 因此, 这两种下近似的定义存在着过于宽松和过于严格的缺陷。近年来, 学者们也提出了一些改进的方法<sup>[20,23]</sup>。本节在文献[20]的基础上, 提出一种平均邻域多粒度决策理论粗糙集。

**定义 9** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于目标近似对象集  $X \subseteq U, X$  关于  $m$  个邻域容差关系的平均邻域多粒度决策理论粗糙集的下近似和上近似分别定义为:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} = \{x \mid (P(X \mid nt_{B_1}(x)) + P(X \mid nt_{B_2}(x)) + \dots + P(X \mid nt_{B_m}(x)))/m \geq \alpha\} \quad (52)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} = U - \{x \mid (P(X \mid nt_{B_1}(x)) + P(X \mid nt_{B_2}(x)) + \dots + P(X \mid nt_{B_m}(x)))/m \leq \beta\} \quad (53)$$

称  $\left\langle \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)}, \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} \right\rangle$  为  $X$  关于  $m$  个邻域容差关系  $N_{B_1}, N_{B_2}, \dots, N_{B_m}$  的平均邻域多粒度决策理论粗糙集。同理,  $X$  关于  $m$  个邻域容差关系的平均邻域多粒度决策理论粗糙集的正区域、边界域和负区域分别定义为:

$$POS_{\sum_{i=1}^m N_{B_i}}^M(X) = \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} \quad (54)$$

$$BUN_{\sum_{i=1}^m N_{B_i}}^M(X) = \overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} - \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} \quad (55)$$

$$NEG_{\sum_{i=1}^m N_{B_i}}^M(X) = U - \overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} \quad (56)$$

平均邻域多粒度决策理论粗糙集满足如下性质。

**性质 7** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $X_1 \subseteq X_2 \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}^M(X_1)} \subseteq \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X_2)} \quad (57)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}^M(X_1)} \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}^M(X_2)} \quad (58)$$

**性质 8** 对于不完备邻域信息系统  $IIS = (U, At, V), B_1, B_2, \dots, B_m \subseteq At$  是一个属性子集族, 诱导出的邻域容差关系分别为  $NT_{B_1}, NT_{B_2}, \dots, NT_{B_m}$ 。对于  $\forall X \subseteq U$ , 满足:

$$\overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i\alpha}^O(X)} \quad (59)$$

$$\overline{\sum_{i=1}^m N_{B_i\alpha}^P(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)} \quad (60)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}^O(X)} \quad (61)$$

$$\overline{\sum_{i=1}^m N_{B_i\beta}^P(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} \quad (62)$$

证明: 对于  $\forall x \in \overline{\sum_{i=1}^m N_{B_i\alpha}^M(X)}$ , 有  $(P(X \mid nt_{B_1}(x)) + \dots + P(X \mid nt_{B_m}(x)))/m \geq \alpha$ , 即存在  $1 \leq k \leq m$  有  $P(X \mid nt_{B_k}(x)) \geq \alpha$ , 因此满足  $x \in \overline{\sum_{i=1}^m N_{B_i\alpha}^O(X)}$ 。所以式(59)成立。

对于  $\forall x \in \overline{\sum_{i=1}^m N_{B_i\alpha}^P(X)}$ , 有:

$$P(X \mid nt_{B_1}(x)) \geq \alpha \wedge \dots \wedge P(X \mid nt_{B_m}(x)) \geq \alpha,$$

那么:

$$(P(X \mid nt_{B_1}(x)) + \dots + P(X \mid nt_{B_m}(x)))/m \geq \alpha,$$

所以式(60)成立。

对于  $\forall x \in NEG_{\sum_{i=1}^m N_{B_i}}^O(X)$ , 有:

$$P(X \mid nt_{B_1}(x)) \leq \beta \wedge \dots \wedge P(X \mid nt_{B_m}(x)) \leq \beta,$$

那么:

$$(P(X \mid nt_{B_1}(x)) + \dots + P(X \mid nt_{B_m}(x)))/m \leq \beta,$$

即  $x \in NEG_{\sum_{i=1}^m N_{B_i}}^M(X)$ , 所以  $NEG_{\sum_{i=1}^m N_{B_i}}^O(X) \subseteq NEG_{\sum_{i=1}^m N_{B_i}}^M(X)$ 。

根据定义 7 和定义 9 有  $\overline{\sum_{i=1}^m N_{B_i\beta}^M(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i\beta}^O(X)}$ ,

因此式(61)成立。

对于  $\forall x \in NEG_{\sum_{i=1}^m N_{B_i}}^M(X)$ , 有:

$$(P(X | nt_{B_1}(x)) + \dots + P(X | nt_{B_m}(x))) / m \leq \beta,$$

那么存在  $1 \leq k \leq m$  有  $P(X | nt_{B_k}(x)) \leq \beta$ , 因此  $x \in NEG_{\sum_{i=1}^m N_{B_i}}^P(X)$ , 所以  $NEG_{\sum_{i=1}^m N_{B_i}}^M(X) \subseteq NEG_{\sum_{i=1}^m N_{B_i}}^P(X)$ 。根

据定义 8 和定义 9 有  $\overline{\sum_{i=1}^m N_{B_i}^P(X)} \subseteq \overline{\sum_{i=1}^m N_{B_i}^M(X)}$ , 因此式(62)成立。

性质 4 揭示了这三种邻域多粒度决策理论粗糙集之间的关系。从中可以看出,乐观情形的下近似刻画得较为宽松,使得下近似集较大,而平均情形的下近似适当限制了乐观情形的定义,平均情形的下近似集小于乐观的下近似集;悲观情形下的下近似刻画得较为严格,使得下近似集较小,而平均情形的下近似集放宽了悲观下近似集的定义,使得平均情形的下近似集大于悲观的下近似集。因此,针对乐观和悲观两种模型的局限性,平均情形下的模型有着很大的改进。

决策偏好是一种重要的多属性决策问题。在第 2 节中提出了三种邻域多粒度决策理论粗糙集,这三种模型均通过阈值  $\alpha$  和  $\beta$  来限制目标概念的粗糙近似,因此这里只需要对阈值的大小进行调整便可以达到决策的偏好程度。对于目标概念  $X$ ,如果决策者是  $X$  厌恶型,那么可以提高阈值  $\alpha$ ,从而提高目标概念的逼近程度,反之则降低阈值  $\alpha$ 。同理对于目标概念  $\sim X$ ,可以通过调整阈值  $\beta$  达到同样的效果。而在所提出的模型中,阈值  $\alpha$  和  $\beta$  直接由代价函数得到,因此对于决策偏好问题,本质上可以通过选择适当的代价函数来实现,即决策厌恶型可以选择高的误分类代价,反之选择低的误分类代价。

### 3 可变三支决策

第 2 节中给出了三种不完备邻域多粒度决策理论粗糙集模型,每种模型都将论域划分为三个互不相交的区域,即正区域、边界域和负区域,基于这三个区域便可以导出邻域多粒度决策理论粗糙集模型下的三支决策。

对于  $X \subseteq U$ , 设  $POS(X)$ 、 $BUN(X)$  和  $NEG(X)$  分别代表了各种不完备邻域多粒度决策理论粗糙集模型下的三个区域。那么可以导出三种决策规则:

(P) 若  $x \in POS(X)$ , 则  $Dec(x) \xrightarrow{+} Dec(X)$ ;

(B) 若  $x \in BUN(X)$ , 则  $Dec(x) \xrightarrow{?} Dec(X)$ ;

(N) 若  $x \in NEG(X)$ , 则  $Dec(x) \xrightarrow{-} Dec(X)$ 。

其中  $Dec(\cdot)$  表示  $\cdot$  的描述<sup>[14]</sup>,  $\xrightarrow{+}$  表示接受决策,  $\xrightarrow{?}$  表示延迟决策,  $\xrightarrow{-}$  表示拒绝决策。

根据 Yao 提出的三支决策理论,代价函数是三支决策诱导的关键,通过代价函数,可以计算出一对阈值  $\alpha$  和  $\beta$ ,从而依据决策理论粗糙集模型对论域划分成三个区域。目前的决策理论粗糙集模型基于完备信息系统构建,代价函数是一个事先给定的数值,但是在不完备信息系统中,由于缺失值的存在,信息系统的不确定性会进一步增大<sup>[24]</sup>,这种不确定性的增大固会产生代价函数也出现一定的不确定性<sup>[25]</sup>,而传统的代价函数是以固定值的形式给出,因而并不适用于不完备信息系统。为此,本文不再将代价函数设定为单一的数值,而是以一种区间值的形式来表现代价函数的这种不确定性,通过区间值的形式可以调控代价函数的大小,进而可以建立不同程度的三支决策。本文将这种决策诱导方式命名为可变三支决策。

类似于表 1,将代价函数扩展成区间值的形式如表 2 所示。定义代价函数  $\tilde{\lambda} = [\lambda^{\leftarrow}, \lambda^{\rightarrow}]$ , 其中  $\lambda^{\leftarrow}$  表示区间值的下界,  $\lambda^{\rightarrow}$  表示区间值的上界。

表 2 区间值代价函数

$\Gamma$	$X$	$\sim X$
$a_P$	$\tilde{\lambda}_{PP} = [\lambda_{PP}^{\leftarrow}, \lambda_{PP}^{\rightarrow}]$	$\tilde{\lambda}_{PN} = [\lambda_{PN}^{\leftarrow}, \lambda_{PN}^{\rightarrow}]$
$a_B$	$\tilde{\lambda}_{BP} = [\lambda_{BP}^{\leftarrow}, \lambda_{BP}^{\rightarrow}]$	$\tilde{\lambda}_{BN} = [\lambda_{BN}^{\leftarrow}, \lambda_{BN}^{\rightarrow}]$
$a_N$	$\tilde{\lambda}_{NP} = [\lambda_{NP}^{\leftarrow}, \lambda_{NP}^{\rightarrow}]$	$\tilde{\lambda}_{NN} = [\lambda_{NN}^{\leftarrow}, \lambda_{NN}^{\rightarrow}]$

根据表 2,可以得到如下关系:

$$\lambda_{PP}^{\leftarrow} \leq \lambda_{BP}^{\leftarrow} < \lambda_{NP}^{\leftarrow}, \lambda_{PP}^{\rightarrow} \leq \lambda_{BP}^{\rightarrow} < \lambda_{NP}^{\rightarrow};$$

$$\lambda_{PN}^{\leftarrow} \leq \lambda_{BN}^{\leftarrow} < \lambda_{NN}^{\leftarrow}, \lambda_{PN}^{\rightarrow} \leq \lambda_{BN}^{\rightarrow} < \lambda_{NN}^{\rightarrow}.$$

类似于传统的决策理论粗糙集模型,这里定义  $\forall x \in U$  采取三种行为的预期代价为:

$$\begin{aligned} \bar{R}(a_P | nt(x)) &= \tilde{\lambda}_{PP} P(X | nt(x)) + \tilde{\lambda}_{PN} P(\sim X | nt(x)) = \\ &= [\lambda_{PP}^{\leftarrow} P(X | nt(x)), \lambda_{PP}^{\rightarrow} P(X | nt(x))] + \\ &= [\lambda_{PN}^{\leftarrow} P(\sim X | nt(x)), \lambda_{PN}^{\rightarrow} P(\sim X | nt(x))] = \\ &= [\lambda_{PP}^{\leftarrow} P(X | nt(x)) + \lambda_{PN}^{\leftarrow} (1 - P(X | nt(x))), \\ &= \lambda_{PP}^{\rightarrow} P(X | nt(x)) + \lambda_{PN}^{\rightarrow} (1 - P(X | nt(x)))] \end{aligned} \tag{63}$$

$$\begin{aligned} \bar{R}(a_B | nt(x)) &= \tilde{\lambda}_{BP} P(X | nt(x)) + \tilde{\lambda}_{BN} P(\sim X | nt(x)) = \\ &= [\lambda_{BP}^{\leftarrow} P(X | nt(x)), \lambda_{BP}^{\rightarrow} P(X | nt(x))] + \\ &= [\lambda_{BN}^{\leftarrow} P(\sim X | nt(x)), \lambda_{BN}^{\rightarrow} P(\sim X | nt(x))] = \\ &= [\lambda_{BP}^{\leftarrow} P(X | nt(x)) + \lambda_{BN}^{\leftarrow} (1 - P(X | nt(x))), \\ &= \lambda_{BP}^{\rightarrow} P(X | nt(x)) + \lambda_{BN}^{\rightarrow} (1 - P(X | nt(x)))] \end{aligned}$$

$$\lambda_{BP}^{\rightarrow}P(X | nt(x)) + \lambda_{BN}^{\leftarrow}(1 - P(X | nt(x))) \quad (64)$$

$$\begin{aligned} \bar{R}(a_N | nt(x)) &= \bar{\lambda}_{NP}P(X | nt(x)) + \bar{\lambda}_{NN}P(\sim X | nt(x)) = \\ &= [\lambda_{NP}^{\leftarrow}P(X | nt(x)), \lambda_{NP}^{\rightarrow}P(X | nt(x))] + \\ &= [\lambda_{NN}^{\leftarrow}P(\sim X | nt(x)), \lambda_{NN}^{\rightarrow}P(\sim X | nt(x))] = \\ &= [\lambda_{NP}^{\leftarrow}P(X | nt(x)) + \lambda_{NN}^{\leftarrow}(1 - P(X | nt(x))), \\ &= \lambda_{NP}^{\rightarrow}P(X | nt(x)) + \lambda_{NN}^{\rightarrow}(1 - P(X | nt(x)))] \quad (65) \end{aligned}$$

由于代价函数是以区间值的形式表示,因此得到预期代价结果也是一个区间值,即预期的代价结果在一个范围内并不是确定的值。但是在实际应用中,需要得到一个具体的预期代价结果,因此就需要对预期代价区间值中取一个具体的值,为此,这里定义一种区间值的取值函数。

**定义 10** 设区间值  $\omega = [a, b]$ , 定义函数  $\xi_k(\omega) = a + k \cdot (b - a)$ , 这里的  $0 \leq k \leq 1$ 。

在定义 10 中, 函数  $\xi_k(\omega)$  通过  $k$  在区间值  $[a, b]$  中进行取值, 当  $k=0$  时,  $\xi_k(\omega) = a$ , 当  $k=1$  时,  $\xi_k(\omega) = b$ , 当  $0 < k < 1$  时,  $a < \xi_k(\omega) < b$ 。因此可以根据实际需求, 选取不同的  $k$  值, 这样就可以得到区间值中的每一个值。

根据定义 10, 可以进一步得到:

$$\begin{aligned} \xi_k(\bar{R}(a_P | nt(x))) &= \\ &= \lambda_{PP}^{\rightarrow}P(X | nt(x)) + \lambda_{PN}^{\leftarrow}(1 - P(X | nt(x))) + \\ &= k \cdot (\lambda_{PP}^{\rightarrow}P(X | nt(x)) + \lambda_{PN}^{\leftarrow}(1 - P(X | nt(x)))) - \\ &= \lambda_{PP}^{\leftarrow}P(X | nt(x)) - \lambda_{PN}^{\rightarrow}(1 - P(X | nt(x))) = \\ &= (\lambda_{PP}^{\leftarrow} + k \cdot (\lambda_{PP}^{\rightarrow} - \lambda_{PP}^{\leftarrow}))P(X | nt(x)) + \\ &= (\lambda_{PN}^{\leftarrow} + k \cdot (\lambda_{PN}^{\rightarrow} - \lambda_{PN}^{\leftarrow}))(1 - P(X | nt(x))) = \\ &= \xi_k(\bar{\lambda}_{PP})P(X | nt(x)) + \xi_k(\bar{\lambda}_{PN})(1 - P(X | nt(x))); \end{aligned}$$

$$\begin{aligned} \xi_k(\bar{R}(a_B | nt(x))) &= \\ &= \lambda_{BP}^{\leftarrow}P(X | nt(x)) + \lambda_{BN}^{\rightarrow}(1 - P(X | nt(x))) + \\ &= k \cdot (\lambda_{BP}^{\leftarrow}P(X | nt(x)) + \lambda_{BN}^{\rightarrow}(1 - P(X | nt(x)))) - \\ &= \lambda_{BP}^{\rightarrow}P(X | nt(x)) - \lambda_{BN}^{\leftarrow}(1 - P(X | nt(x))) = \\ &= (\lambda_{BP}^{\rightarrow} + k \cdot (\lambda_{BP}^{\leftarrow} - \lambda_{BP}^{\rightarrow}))P(X | nt(x)) + \\ &= (\lambda_{BN}^{\rightarrow} + k \cdot (\lambda_{BN}^{\leftarrow} - \lambda_{BN}^{\rightarrow}))(1 - P(X | nt(x))) = \\ &= \xi_k(\bar{\lambda}_{BP})P(X | nt(x)) + \xi_k(\bar{\lambda}_{BN})(1 - P(X | nt(x))); \end{aligned}$$

$$\begin{aligned} \xi_k(\bar{R}(a_N | nt(x))) &= \\ &= \lambda_{NP}^{\leftarrow}P(X | nt(x)) + \lambda_{NN}^{\rightarrow}(1 - P(X | nt(x))) + \\ &= k \cdot (\lambda_{NP}^{\leftarrow}P(X | nt(x)) + \lambda_{NN}^{\rightarrow}(1 - P(X | nt(x)))) - \\ &= \lambda_{NP}^{\rightarrow}P(X | nt(x)) - \lambda_{NN}^{\leftarrow}(1 - P(X | nt(x))) = \\ &= (\lambda_{NP}^{\rightarrow} + k \cdot (\lambda_{NP}^{\leftarrow} - \lambda_{NP}^{\rightarrow}))P(X | nt(x)) + \\ &= (\lambda_{NN}^{\leftarrow} + k \cdot (\lambda_{NN}^{\rightarrow} - \lambda_{NN}^{\leftarrow}))(1 - P(X | nt(x))) = \\ &= \xi_k(\bar{\lambda}_{NP})P(X | nt(x)) + \xi_k(\bar{\lambda}_{NN})(1 - P(X | nt(x))). \end{aligned}$$

类似于经典决策理论粗糙集模型, 有:

$$\begin{aligned} \xi_k(\bar{\lambda}_{PP}) &\leq \xi_k(\bar{\lambda}_{BP}) < \xi_k(\bar{\lambda}_{NP}) \\ \xi_k(\bar{\lambda}_{NN}) &\leq \xi_k(\bar{\lambda}_{BN}) < \xi_k(\bar{\lambda}_{PN}) \end{aligned}$$

所以:

$$\begin{aligned} \xi_k(\bar{R}(a_P | nt(x))) &\leq \xi_k(\bar{R}(a_B | nt(x))) \Leftrightarrow \\ \xi_k(\bar{\lambda}_{PP})P(X | nt(x)) &+ \xi_k(\bar{\lambda}_{PN})(1 - P(X | nt(x))) \leq \\ \xi_k(\bar{\lambda}_{BP})P(X | nt(x)) &+ \xi_k(\bar{\lambda}_{BN})(1 - P(X | nt(x))) \Leftrightarrow \\ P(X | nt(x)) &\geq \frac{\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{BN})}{(\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{BN})) + (\xi_k(\bar{\lambda}_{BP}) - \xi_k(\bar{\lambda}_{PP}))} \end{aligned}$$

同理:

$$\begin{aligned} \xi_k(\bar{R}(a_B | nt(x))) &\leq \xi_k(\bar{R}(a_N | nt(x))) \Leftrightarrow \\ \xi_k(\bar{\lambda}_{BP})P(X | nt(x)) &+ \xi_k(\bar{\lambda}_{BN})(1 - P(X | nt(x))) \leq \\ \xi_k(\bar{\lambda}_{NP})P(X | nt(x)) &+ \xi_k(\bar{\lambda}_{NN})(1 - P(X | nt(x))) \Leftrightarrow \\ P(X | nt(x)) &\geq \frac{\xi_k(\bar{\lambda}_{BN}) - \xi_k(\bar{\lambda}_{NN})}{(\xi_k(\bar{\lambda}_{BN}) - \xi_k(\bar{\lambda}_{NN})) + (\xi_k(\bar{\lambda}_{NP}) - \xi_k(\bar{\lambda}_{BP}))} \end{aligned}$$

这里令:

$$\bar{\alpha}_k = \frac{\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{BN})}{(\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{BN})) + (\xi_k(\bar{\lambda}_{BP}) - \xi_k(\bar{\lambda}_{PP}))} \quad (66)$$

$$\bar{\beta}_k = \frac{\xi_k(\bar{\lambda}_{BN}) - \xi_k(\bar{\lambda}_{NN})}{(\xi_k(\bar{\lambda}_{BN}) - \xi_k(\bar{\lambda}_{NN})) + (\xi_k(\bar{\lambda}_{NP}) - \xi_k(\bar{\lambda}_{BP}))} \quad (67)$$

$$\bar{\gamma}_k = \frac{\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{NN})}{(\xi_k(\bar{\lambda}_{PN}) - \xi_k(\bar{\lambda}_{NN})) + (\xi_k(\bar{\lambda}_{NP}) - \xi_k(\bar{\lambda}_{PP}))} \quad (68)$$

那么有:

(P) 如果  $P(X | nt(x)) \geq \bar{\alpha}_k$  且  $P(X | nt(x)) \geq \bar{\gamma}_k$ , 那么  $x \in POS(X)$ ;

(B) 如果  $P(X | nt(x)) < \bar{\alpha}_k$  且  $P(X | nt(x)) > \bar{\beta}_k$ , 那么  $x \in BUN(X)$ ;

(N) 如果  $P(X | nt(x)) \leq \bar{\beta}_k$  且  $P(X | nt(x)) \leq \bar{\gamma}_k$ , 那么  $x \in NEG(X)$ 。

若  $0 \leq \bar{\beta}_k < \bar{\gamma}_k < \bar{\alpha}_k \leq 1$ , 因此上述规则可以简化为:

(P) 如果  $P(X | nt(x)) \geq \bar{\alpha}_k$ , 那么  $x \in POS(X)$ ;

(B) 如果  $\bar{\beta}_k < P(X | nt(x)) < \bar{\alpha}_k$ , 那么  $x \in BUN(X)$ ;

(N) 如果  $P(X | nt(x)) \leq \bar{\beta}_k$ , 那么  $x \in NEG(X)$ 。

因此, 这里只需要根据实际需求选取不同的  $k$  值, 便可以得到相对应的一对阈值  $\bar{\alpha}_k$  和  $\bar{\beta}_k$ , 从而导出对应  $k$  值下的三支决策规则。理论分析表明本节所提出的可变三支决策具有更好的灵活性。

### 4 实例分析

本节将通过一个具体的实例来分析本文所提出的模型。实例所选取的是一个大学生三好学生评估表,众所周知,大学生的三好学生的评估将从思想品德、课程成绩以及社会实践等不同方面对学生进行综合评估,因此这符合了多粒度粗糙集模型实际应用的基本视角,同时加入三支决策的思想,可以使得问题的处理更加科学合理。

表 3 为一个某专业 12 名同学的思想品德、课程成绩以及社会实践三方面测评信息表  $IIS = (U, At, V)$ , 其中的成绩值均为归一化至  $[0, 1]$  区间的结果。该表中包含 6 个属性,即  $\{a_1, a_2, b_1, b_2, c_1, c_2\}$ , 并被分为三大类,其中  $A = \{a_1, a_2\}$  表示两位老师对每位同学的思想品德测评结果,由于有的老师对个别同学不熟悉,因此很难给出具体的评分,所以成绩出现了空缺,表中用“\*”表示。 $B = \{b_1, b_2\}$  表示每位同学中期末考试的综合成绩,由于个别同学未参加,因此也出现了部分空缺。 $C = \{c_1, c_2\}$  表示两位老师对每位同学的社会实践评分,同样存在个别的空缺。对于表 3 中的决策属性 class, 其中决策值为 1 的学生表示最有可能获得三好学生的同学。

表 3 三好学生测评信息表

U	A		B		C		class
	$a_1$	$a_2$	$b_1$	$b_2$	$c_1$	$c_2$	
$u_1$	0.47	0.66	0.76	0.61	0.43	0.31	0
$u_2$	0.73	0.62	0.54	0.76	0.56	0.44	1
$u_3$	0.21	*	0.35	0.26	0.46	0.59	0
$u_4$	0.76	0.61	0.87	0.69	0.26	0.19	1
$u_5$	0.45	0.54	0.67	*	0.75	0.66	0
$u_6$	0.88	0.78	0.34	0.47	0.56	0.62	1
$u_7$	0.17	0.23	0.45	0.36	0.45	0.52	0
$u_8$	0.95	0.86	0.76	0.64	0.67	*	1
$u_9$	0.65	0.57	*	0.57	0.34	0.27	0
$u_{10}$	0.87	0.81	0.88	0.76	0.47	0.58	1
$u_{11}$	*	0.64	0.62	0.59	0.54	0.62	0
$u_{12}$	0.31	0.46	0.68	0.71	0.74	0.65	1

将决策值为 1 的集合设为  $X = \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}$ , 本实例分析的状态集为  $\Omega = \{X, \sim X\}$ , 行为集为  $\Gamma = \{a_p, a_B, a_N\}$ 。该高校的教务部门给出的三好学生测评代价函数如表 4 所示, 由于有缺失值的存在, 因此所示的代价函数是一个区间形式, 便于各专业选取和

调整。

表 4 三好学生测评代价函数

$\Gamma$	$X$	$\sim X$
$a_p$	$\bar{\lambda}_{pp} = [0, 0.2]$	$\bar{\lambda}_{pN} = [0.8, 1.2]$
$a_B$	$\bar{\lambda}_{BP} = [0.2, 0.6]$	$\bar{\lambda}_{BN} = [0.3, 0.6]$
$a_N$	$\bar{\lambda}_{NP} = [0.8, 1.2]$	$\bar{\lambda}_{NN} = [0.2, 0.4]$

采用本文所提出的不完备邻域多粒度决策理论粗糙集对其进行建模, 取属性集  $A, B$  和  $C$  作为该模型的两个粒度层次, 它们诱导出的邻域容差关系分为  $NT_A, NT_B$  和  $NT_C$ , 然后根据  $NT_A + NT_B + NT_C$  三个粒度针对  $X$  进行多粒度决策理论粗糙逼近, 得到  $X$  的三个区域, 即正区域、边界域和负区域。由于本文提出了多粒度决策理论粗糙集的三种具体模型, 分别为乐观模型、悲观模型和平均模型, 因此对于乐观模型, 三个区域记为  $POS^0(X), BUN^0(X)$  和  $NEG^0(X)$ ; 对于悲观模型, 三个区域记为  $POS^p(X), BUN^p(X)$  和  $NEG^p(X)$ ; 对于平均模型, 三个区域记为  $POS^m(X), BUN^m(X)$  和  $NEG^m(X)$ 。由于该模型运用的代价函数是区间形式, 针对定义 10 中的  $\xi_k$  函数, 本实例分析将其中的  $k$  值在  $[0, 1]$  区间以步长 0.2 依次取值, 然后根据对应的代价函数分别进行实验, 依次观察  $k$  对最终结果的影响。另外, 对于该模型的邻域容差关系, 本实验设定邻域半径为 0.1。表 5、表 6 和表 7 分别为对象集  $X$  关于  $NT_A + NT_B + NT_C$  的乐观、悲观和平均模型取不同  $k$  值的三个区域结果。

表 5 乐观邻域多粒度决策理论粗糙集三个区域

$k$ 值	$POS^0(X); BUN^0(X); NEG^0(X)$
0.0	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$
0.2	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$
0.4	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$
0.6	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$
0.8	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$
1.0	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_9, u_{11}\}; \{u_7\}$

表 6 悲观邻域多粒度决策理论粗糙集三个区域

$k$ 值	$POS^p(X); BUN^p(X); NEG^p(X)$
0.0	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$
0.2	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$
0.4	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$
0.6	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$
0.8	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$
1.0	$\{\}; \{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_1, u_3, u_5, u_7, u_9, u_{11}\}$

表 7 平均邻域多粒度决策理论粗糙集三个区域

$k$ 值	$POS^M(X); BUN^M(X); NEG^M(X)$
0.0	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_3, u_5, u_9\}; \{u_1, u_7, u_{11}\}$
0.2	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_3, u_5, u_9\}; \{u_1, u_7, u_{11}\}$
0.4	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_5\}; \{u_1, u_3, u_7, u_9, u_{11}\}$
0.6	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_5\}; \{u_1, u_3, u_7, u_9, u_{11}\}$
0.8	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_5\}; \{u_1, u_3, u_7, u_9, u_{11}\}$
1.0	$\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}; \{u_5\}; \{u_1, u_3, u_7, u_9, u_{11}\}$

表 5 为  $X$  在乐观邻域多粒度决策理论粗糙集的三个区域,根据三支决策模型,即  $POS^O(X)$  中的学生直接评定为三好学生,  $BUN^O(X)$  中的学生有待进一步进行考察,而  $NEG^O(X)$  中的学生直接不评定为三好学生。观察表 5 可以发现,无论  $k$  取何值时,三个区域始终保持不变,即在此评测表中,最终的评测结果不随着代价的变化而变化。同时有  $POS^O(X) = X$ ,表明原先可能评为三好学生的同学都评定为三好学生,  $BUN^O(X) = \{u_1, u_3, u_5, u_9, u_{11}\}$  表明这些同学处于待定状态,而只有  $NEG^O(X) = \{u_7\}$  是直接不评定为三好学生。这样的评定结果刚好符合乐观邻域多粒度决策理论粗糙集的“乐观”态度,在表 3 中,  $\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}$  中的同学都至少有一个成绩满足  $>0.7$ ,该模型直接判定为三好学生,其余的大部分都进入待定状态,而只有一个是直接拒绝。这主要是由于乐观模型中,只要有一个粒度下对象的邻域容差类的包含度满足大于对应阈值,那么该对象便属于正区域中,而所有粒度下对象的邻域容差类的包含度满足小于对应阈值,该对象才隶属于负区域,这样使得正区域的定义比较松弛,而负区域的定义比较苛刻,因而出现了表 5 中的评定结果,并且受代价函数的变化影响较小。所以对于不严格的三好学生评定,那么该模型是一个较好的选择。

表 6 为  $X$  在悲观邻域多粒度决策理论粗糙集的三个区域,观察发现,随着  $k$  取值的变化,三个区域始终保持不变,即在此评测表中,最终的评测结果不随着代价的变化而变化。对于评测结果,  $POS^P(X)$  为空集,即没有任何一个同学直接评定为三好学生。  $BUN^P(X) = X$ ,即原先有较大可能评定为三好学生的同学都进入了待定状态,其余的同学直接拒绝评定为三好学生,这与表 5 中的结果相反,体现出了悲观邻域多粒度决策理论粗糙集的“悲观”态度,  $\{u_2, u_4, u_6, u_8, u_{10}, u_{12}\}$  中的同学虽然都有一个成绩满足  $>0.7$ ,但是也都存在成绩  $<0.7$ ,从而进入待定状态。这主要是由于悲观模型与乐观模型的定义相反,正区域的定义比较苛刻,负区域的定义比较松弛,因而出现了表 6 中的评定

结果,同样受代价函数的变化影响较小。所以对于严格的三好学生评定,应选择该模型进行评定。

表 7 为  $X$  在平均邻域多粒度决策理论粗糙集三个区域,观察发现所示的结果与表 5 和表 6 不同,在  $k$  的取值较小时,即当  $k$  为  $0 \sim 0.2$  时,  $POS^M(X) = X$ ,即原先较大可能为三好学生的同学直接评定为三好学生,  $BUN^M(X) = \{u_3, u_5, u_9\}$ ,即这些同学进入待定状态,而  $NEG^M(X) = \{u_1, u_7, u_{11}\}$  直接拒绝评定为三好学生,观察表 3 的成绩可以看出,  $\{u_3, u_5, u_9\}$  的评测成绩整体上要优于  $\{u_1, u_7, u_{11}\}$  的评测成绩。当  $k$  的值逐渐增大,即当  $k$  为  $0.4 \sim 1.0$  时,  $POS^M(X)$  保持不变,  $BUN^M(X)$  中只剩下  $u_5$ ,而  $NEG^M(X) = \{u_1, u_3, u_7, u_9, u_{11}\}$ ,这表明,原先待定的同学  $u_3$  和  $u_9$  转变为拒绝评定为三好学生,观察表 3 的评测成绩,可以发现  $u_5$  的成绩要优于  $u_3$  和  $u_9$ 。表 7 中出现这样的结果,主要是由于本文所提出的平均邻域多粒度决策理论粗糙集三个区域是在乐观模型和悲观模型两者之间取了一个折中,分别对两个模型的局限性进行了改善,因此表 7 中的评测结果比较中规中矩。同时对于区间代价函数,  $\bar{\lambda}_{PP} = [0, 0.2]$ ,  $\bar{\lambda}_{PN} = [0.8, 1.2]$ ,  $\bar{\lambda}_{NP} = [0.8, 1.2]$ ,  $\bar{\lambda}_{NN} = [0.2, 0.4]$ ,当  $k$  值逐渐增大,  $\xi_k(\bar{\lambda}_{PP})$  与  $\xi_k(\bar{\lambda}_{PN})$  之间的差距逐渐拉大,  $\xi_k(\bar{\lambda}_{NP})$  与  $\xi_k(\bar{\lambda}_{NN})$  之间的差距也逐渐拉大,错误的分类将引起更大的代价,使得相应的  $\bar{\alpha}_k$  和  $\bar{\beta}_k$  的值会提高,那么评测的结果将会更加严格,原先待定的同学可能会有一部分产生拒绝决策,其结果正如表 7 所示的那样。

综合分析表 5、表 6 和表 7 的结果,可以得出,所提出的乐观邻域多粒度决策理论粗糙集和悲观邻域多粒度决策理论粗糙集在三个区域划分的严格程度方面,都存在一定的缺陷。而所提出的平均邻域多粒度决策理论粗糙集对如上两个模型的缺陷都进行了一定的改善,同时针对不完备数据提出的区间代价函数,并通过取不同  $k$  值的方式得到了不同程度的三支决策规则,为本文所提出的模型又增添了一份灵活性。

在本文所提出的不完备邻域多粒度决策理论粗糙集模型中,其中邻域容差关系是处理数值型不完备数据的关键,其中包含一个邻域半径  $\delta$  的参数,它的取值不同将会对模型的处理结果产生一定的影响,因此接下来将探究邻域半径  $\delta$  对实例结果的影响以及如何选取最佳的取值问题。

表 3 中的评测成绩值已归一化入  $0 \sim 1$  之间,因此将邻域半径在  $[0.05, 0.55]$  之间以  $0.05$  为步长进行取值,  $k$  值在  $[0, 1]$  区间以  $0.1$  为步长进行取值,然后将每一组邻域半径和  $k$  值的参数分别赋入模型中,计算

对象集  $X$  的三个区域。其结果如图 1、图 2 和图 3 所示,图 1 表示的是乐观模型下不同邻域半径和  $k$  值的三个区域结果,图 2 表示的是悲观模型下不同邻域半径和  $k$  值的三个区域结果,图 3 表示的是平均模型下不同邻域半径和  $k$  值的三个区域结果。

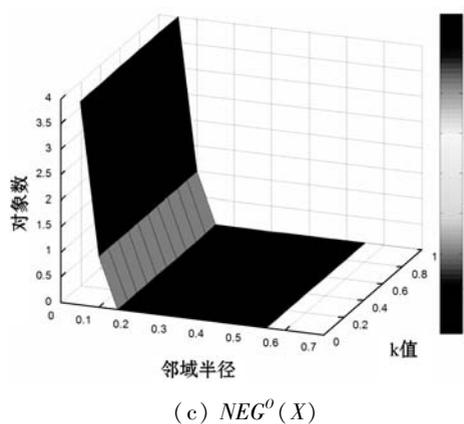
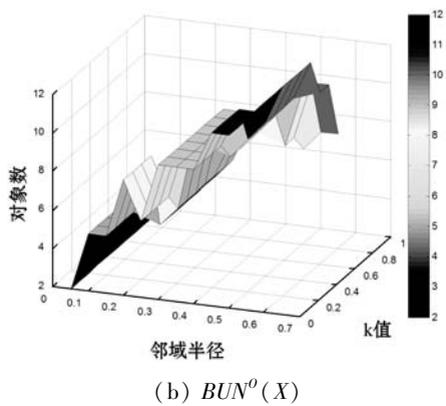
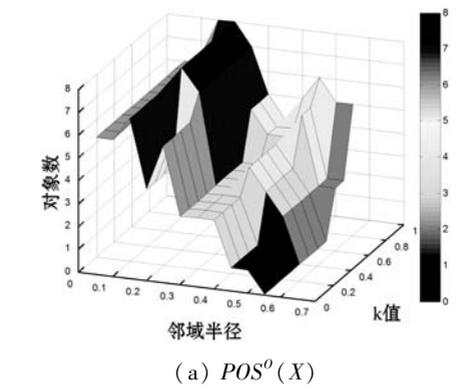


图 1 乐观邻域多粒度决策理论粗糙集模型三个区域大小变化

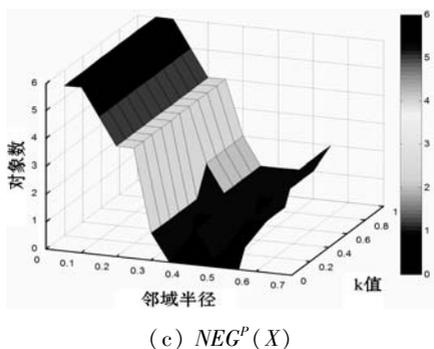
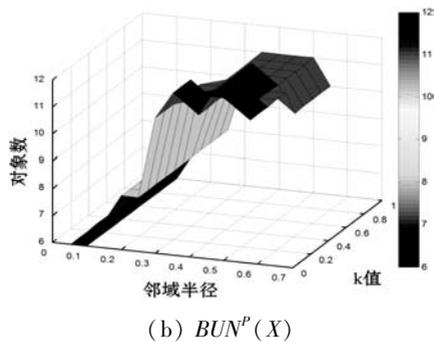
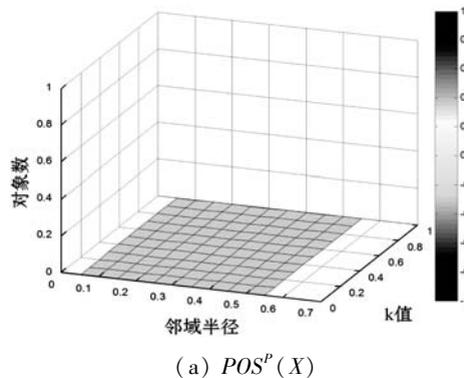


图 2 悲观邻域多粒度决策理论粗糙集模型三个区域大小变化

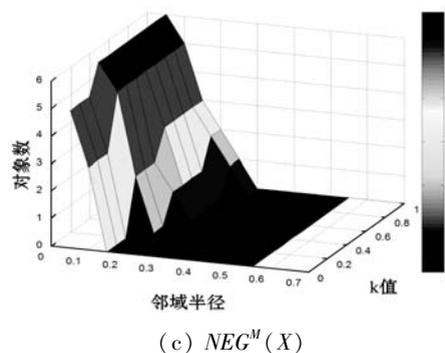
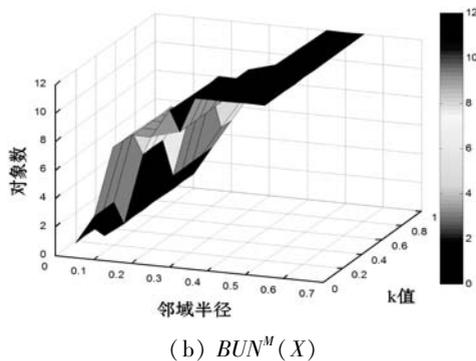
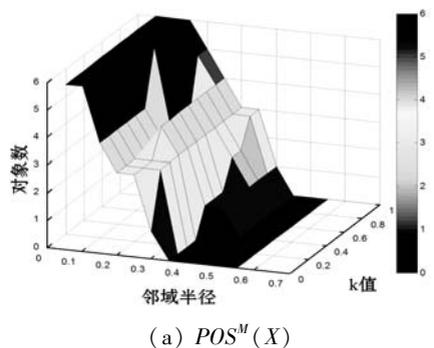


图 3 平均邻域多粒度决策理论粗糙集模型三个区域大小变化

在图1所示的乐观邻域多粒度决策理论粗糙集模型三个区域大小变化中,对于同一个 $k$ 值,随着邻域半径的逐渐增大,正区域 $POS^o(X)$ 是逐渐减小的,当 $\delta = 0.55$ 时, $POS^o(X)$ 几乎为空集;负区域 $NEG^o(X)$ 也是逐渐减小的,并且当 $\delta \geq 0.15$ 时, $NEG^o(X)$ 都为空集;而边界域 $BUN^o(X)$ 满足逐渐增大的趋势,并且最终 $BUN^o(X) = U$ ,这表明过大的邻域半径会使得所有的同学都处于进一步考察状态,因此这不符合预期的评测目的。在图2所示的悲观邻域多粒度决策理论粗糙集模型三个区域大小变化中,随着邻域半径的逐渐增大,正区域 $POS^p(X)$ 都为空集, $NEG^p(X)$ 也是逐渐减小的,边界域 $BUN^p(X)$ 满足逐渐增大并且最终满足 $BUN^p(X) = U$ ,与乐观模型中出现了同样的结果。图3所示的平均模型也是满足同样的规律。因此综合三类模型的实验结果可以发现,选取过大的邻域半径将会对最终的判定结果产生很大的误差,观察各图结果可以看出最佳的邻域半径可选取为 $0.1 \sim 0.2$ 。由于本实例中信息表的属性值都归一化至 $[0, 1]$ 区间,因此实验中得到的最佳邻域半径具有一定的一般性,并且文献[8-9,27]通过大量实验也得出了最佳邻域半径选取为 $0.1 \sim 0.2$ ,因此本文所提出模型中邻域半径应选取为 $0.1 \sim 0.2$ 。

## 5 结语

针对数值型数据和包含缺失值数据等复杂的数据环境,本文将多粒度决策理论粗糙集模型在不完备数值型数据下进行推广,提出了乐观邻域多粒度决策理论粗糙集、悲观邻域多粒度决策理论粗糙集以及平均邻域多粒度决策理论粗糙集,并分析了三种模型的相关性质以及相互关系。同时针对不完备信息系统下决策理论粗糙集模型中代价函数面临的局限性,本文提出了一种可变三支决策,它主要通过区间形式来表示代价函数,然后通过选取不同参数诱导出不同的三支决策结果,使得这种三支决策诱导方法具备了更好的灵活性,最后的实例分析表明了所提出模型与方法的实用性与合理性。本文所提出的模型适用于较为复杂数据环境下的多粒度分析与决策分析,是针对静态的数据建立的,因此接下来将对动态数据环境下作进一步研究。

## 参 考 文 献

- [1] Pawlak Z. Rough Sets[J]. International Journal of Information and Computer Sciences. 1982,11;341-356.
- [2] 张清华, 薛玉斌, 胡峰, 等. 粗糙集近似集不确定性研究[J]. 电子学报, 2016, 44(7):1574-1580.
- [3] Li F C, Yang J N, Jin C X, et al. A new effect-based roughness measure for attribute reduction in information system[J]. Information Sciences, 2017, 378:348-362.
- [4] Zhang X, Mei C L, Chen D G, et al. A fuzzy rough set-based feature selection method using representative instances[J]. Knowledge-Based Systems, 2018,151(1):216-229.
- [5] Yao Y Y. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181(6):1080-1096.
- [6] Yao Y Y. Two semantic issues in a probabilistic rough set model[J]. Fundamenta Informaticae, 2011,108:249-265.
- [7] Kang X P, Miao D Q. A variable precision rough set model based on the granularity of tolerance relation[J]. Knowledge-Based Systems, 2016, 102(15):103-115.
- [8] 姚晟, 徐风, 赵鹏, 等. 基于邻域量化容差关系粗糙集模型的特征选择算法[J]. 模式识别与人工智能, 2017, 30(5):416-428.
- [9] Hu Q H, Yu D R, Liu J F, et al. Neighborhood rough set based heterogeneous feature subset selection[J]. Information Sciences, 2008, 178(18):3577-3594.
- [10] An S, Hu Q H, Pedrycz W, et al. Data-Distribution-Aware fuzzy rough set model and its application to robust classification[J]. IEEE Transactions on Cybernetics, 2015, 46(12):3073-3085.
- [11] Ma J M, Zou C J, Pan X C. Structured probabilistic rough set approximations[J]. International Journal of Approximate Reasoning, 2017, 90:319-332.
- [12] Zhang Q H, Zhang Q, Wang G Y. The uncertainty of probabilistic rough sets in multi-granulation spaces[J]. International Journal of Approximate Reasoning, 2016, 77:38-54.
- [13] Yao Y Y, Wong S K. A decision theoretic framework for approximating concepts[J]. International Journal of Man-Machine Studies, 1992, 37(6):793-809.
- [14] Yao Y Y. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180(3):341-353.
- [15] 张越兵, 苗夺谦, 张志飞. 基于三支决策的多粒度文本情感分类模型[J]. 计算机科学, 2017, 44(12):188-215.
- [16] Zhang H R, Min F. Three-way recommender systems based on random forests[J]. Knowledge-Based Systems, 2016, 91:275-286.
- [17] Xu J F, Miao D Q, Zhang Y J, et al. A three-way decisions model with probabilistic rough sets for stream computing[J]. International Journal of Approximate Reasoning, 2017, 88:1-22.
- [18] Zhong C F, Pedrycz W, Wang D, et al. Granular data imputation: A framework of Granular Computing[J]. Applied Soft Computing, 2016, 46:307-316.
- [19] 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.

- [20] Qian Y H, Zhang H, Sang Y L, et al. Multigranulation decision-theoretic rough sets[J]. International Journal of Approximate Reasoning, 2014, 55(1):225-237.
- [21] Li W W, Huang Z Q, Jia X Y, et al. Neighborhood based decision-theoretic rough set models[J]. International Journal of Approximate Reasoning, 2016, 69:1-17.
- [22] Yuan J L, Chen M C, Jiang T, et al. Complete tolerance relation based parallel filling for incomplete energy big data[J]. Knowledge-Based Systems, 2017, 132(15):215-225.
- [23] 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.
- [24] Dai J H, Xu Q. Approximations and uncertainty measures in incomplete information systems[J]. Information Sciences, 2012, 98(1):62-80.
- [25] Liu D, Liang D C, Wang C C. A novel three-way decision model based on incomplete information system[J]. Knowledge-Based Systems, 2016, 91:32-45.
- [26] Lin G P, Qian Y H, Li J J. NMGRS: Neighborhood-based multigranulation rough sets[J]. International Journal of Approximate Reasoning, 2012, 53(7):1080-1093.
- [27] Zhao H, Qin K Y. Mixed feature selection in incomplete decision table[J]. Knowledge-Based Systems, 2014, 57:181-190.

(上接第42页)

(5) 将质量评价等级集  $V$  与  $B^T$  相乘得到综合质量评价结果分值为:

$$G_1 = V \cdot B_1^T = 0.8680$$

$$G_2 = V \cdot B_2^T = 0.7134$$

根据个性化定制产品质量模糊综合评价结果可以看出通过手机云制造平台定制的产品质量要比普通手机企业评分高。说明利用手机云制造平台使得个性化定制的生产能力和产品质量得到了明显提升,同时,也解决手机企业满足用户需求难度大的问题,将用户需求满意度达到最大化。

## 5 结 语

针对用户对手机个性化需求的增加,手机定制难度大以及手机企业对关键零部件存在研发能力不足等问题显得更加突出。结合云制造理念提出手机制造云模式;建立个性化手机定制的云制造平台系统架构;设计个性化手机定制的云制造平台工作流程和软件框架。本文具有拓展性地将个性化手机定制生产融入云制造技术领域,通过手机制造云平台的统一管控及

优化配置,实现了合理的个性化定制业务并为企业提供最优的制造资源。根据上述分析,说明手机制造云平台完美地解决个性化定制生产能力不足和用户需求难以满足的问题。手机制造云模式不仅促进了手机制造业的发展,同时也带动制造业新模式的创新。

## 参 考 文 献

- [1] 华尔天, 孙琦宗, 刘肖健, 等. 一种面向汽车个性化定制的智能交互方法[J]. 中国机械工程, 2017, 28(19):2275-2281.
- [2] Werner M, Jahn A, Lutz E, et al. Analysis of system parameters for LEO/ICO-satellite communication networks[J]. IEEE Journal on Selected Areas in Communications, 1995, 13(2):371-381.
- [3] 李伯虎, 张霖, 王时龙, 等. 云制造——面向服务的网络化制造新模式[J]. 计算机集成制造系统, 2010, 16(1):1-7.
- [4] Liu Y K, Wang L H, Wang X V. Cloud Manufacturing: Latest Advancements and Future Trends[J]. Procedia Manufacturing, 2018, 25:62-73.
- [5] 航天云网-INDICS 平台[EB/OL]. [2018-04-16]. <http://www.casicloud.com/>.
- [6] 张倩. 华为云强势登陆广州,助力建设“智”造广东——“如果未来可以预见——华为云中国行”系列活动广州站成功举办[J]. 机电工程技术, 2017(9).
- [7] 李伯虎, 张霖, 任磊, 等. 再论云制造[J]. 计算机集成制造系统, 2011, 17(3):449-457.
- [8] Ning F, Zhou W, Zhang F, et al. The architecture of cloud manufacturing and its key technologies research[C]//IEEE International Conference on Cloud Computing and Intelligence Systems. IEEE, 2011:259-263.
- [9] 王庆国, 蔡淑琴, 陈宏峰, 等. 面向客户定制的动态BOM模型及算法[J]. 计算机工程与应用, 2002, 38(24):94-95.
- [10] 朱辉. 需求不确定条件下的主生产计划的制定[J]. 经济师, 2004(10):14-15.
- [11] 宋杰, 王科社, 陈婉, 等. 基于MVVM模式的滚珠丝杠副工艺卡输出设计[J]. 组合机床与自动化加工技术, 2017(3):148-150.
- [12] 张春红. P2P技术全面解析[M]. 人民邮电出版社, 2010.
- [13] Werner M. A Dynamic Routing Concept for ATM-Based Satellite Personal Communication Networks[J]. IEEE Journal on Selected Areas in Communications, 1997, 15(8):1636-1648.
- [14] 顾名字. 基于P2P及B/S搭建教育云平台与流量管理分析[J]. 西南师范大学学报(自然科学版), 2015, 40(7):124-129.
- [15] 姜兴宇, 于天彪, 梁爽, 等. 基于顾客满意的个性化定制质量保证体系研究[J]. 计算机集成制造系统, 2008, 14(8):1608-1616.