小 型 微 型 计 算 机 系 统　　2014 年 11 月 第 11 期
Journal of Chinese Computer Systems　　Vol. 35 No. 11 2014

复合有序信息系统

罗 川，李天瑞，陈红梅，张钧波

（西南交通大学 信息科学与技术学院，成都 610031）
E-mail：luochuan@my.swjtu.edu.cn

摘 要：优势关系粗糙集模型是粗糙集的一种扩展模型，能够处理具有偏好信息的多准则决策问题. 有序信息系统中的数据处理与知识获取是优势关系粗糙集模型的重要应用之一. 现实应用中，一个有序信息系统中可能存在多种类型的数据，在不同属性上对象之间也可能存在各种不同的优势关系. 基于此，在充分考虑各种不同类型的数据上可能具有不同优势关系的基础上，提出一种新的集成不同优势关系的复合优势关系，定义了基于此关系下的复合有序信息系统，并建立了复合优势关系粗糙集模型. 最后，利用矩阵运算能够较直观地体现构造化方法的特点，通过优势关系的矩阵表示和矩阵的相关运算，给出了该模型中上、下近似集的矩阵计算方法，并通过实例对该方法的有效性进行了验证.

关 键 词：粗糙集；复合有序信息系统；优势关系；矩阵运算
中图分类号：TP18　　**文献标识码**：A　　**文 章 编 号**：1000-1220(2014)11-2523-05

Composite Ordered Information Systems

LUO Chuan，LI Tian-rui，CHEN Hong-mei，ZHANG Jun-bo
（School of Information Science and Technology，Southwest Jiaotong University，Chengdu 610031，China）

Abstract：Dominance-based Rough Set Approach (DRSA) is an extended rough set model，which has been introduced to deal with multi-criteria decision problem with preference information. Data processing and knowledge acquisition in the ordered information system is one of the most important applications of DRSA. In real-world applications，there may exist multiple different types of data in an ordered information system and the preference relations between objects may also different under different criterion. Considering the multiple dominance relations w. r. t. different types of data in an ordered information system，the composite dominance relation is proposed in this paper. Then，we present a composite DRSA for the composite ordered information system. Finally，since the character of constructive approach can be reflected more intuitively according to matrix operations，we construct the matrix-based approach for computing the lower and upper approximations in the composite ordered information system. Some numerical examples are used to illustrate the validity of the proposed approach.

Key words：Rough sets；composite ordered information system；dominance relation；matrix operations

1 引 言

粗糙集理论由波兰学者 Pawlak 于 1982 年提出，是一种新型的用于数据表达和分析的数学工具[1]. 该理论能够在不需要任何特定领域的先验知识的前提下，针对不精确、不清晰和不完整等各种不确定信息进行有效的处理. 近年来，随着人们对粗糙集理论的不断深入研究，该理论已经在医疗诊断、过程控制、知识发现、模式识别、人工智能、图像处理等众多领域获得了广泛应用[2-4].

经典粗糙集模型假设论域中每个所研究的对象都附带有一定的信息，如果一些对象由同样的信息来表征，则认为它们是相似的，或者说是不可辨识的，通过利用不可辨识关系（等价关系，满足自反性、对称性和传递性），可以建立论域的一个划分，得到不可辨识的等价类，进而构造集合的上、下近似算子对知识库中不确定的知识进行刻画[5]. 多属性决策问题是决策理论研究的一个重要内容，是实际应用中使用较多的一类决策模式，已被广泛应用于投资决策、质量评估、人才考核、经济效益综合评价等诸多领域[6]. 如何将粗糙集理论引入到多属性决策中已成为目前这一理论应用到实际决策分析与支持中的一个重要研究课题. 由于经典粗糙集基于严格的等价关系，只考虑到对象间的不可辨识性，却未能顾及属性值之间的偏好有序关系，因此这极大的限制了该理论在属性具偏好信息的多属性决策问题上的应用范围. 为此，Greco 等提出了一种基于优势关系的粗糙集方法（Dominance-based Rough Set Approach，DRSA），该方法考虑到了属性值之间存在的偏好关系，将不可辨识关系改进为优势关系，并按照决策者的习惯把决策类以联合的形式表示出来，进而弥补了传统粗糙集在解决该类问题时的

收稿日期：2013-07-22　收修改稿日期：2013-08-13　基金项目：国家自然科学基金项目（60873108，61175047，61100117）资助；NSAF 联合基金项目（U1230117）资助；教育部人文社会科学研究青年基金项目（11YJC630127）资助；中央高校基本科研业务费专项资金专题研究项目（SWJTU11ZT08，SWJTU12CX091，SWJTU12CX117）资助. 作者简介：罗 川，男，1987 年生，博士研究生，主要研究方向为粒计算、粗糙集、数据挖掘和云计算等；李天瑞，男，1969 年生，博士，教授，博士生导师，主要研究方向为智能信息处理、粗糙集、粒计算和云计算等；陈红梅，女，1971 年生，博士，副教授，主要研究方向为数据挖掘、粒计算、粗糙集等；张钧波，男，1986 年生，博士研究生，主要研究方向为粒计算、粗糙集、数据挖掘和云计算等.

缺陷[7,8].

有序信息系统的知识获取与决策分析是优势关系粗糙集方法的最重要应用之一[9,10]. 在 Greco 等针对单值完备的有序信息系统提出了基于优势关系的粗糙集方法之后，众多研究学者针对各种数据类型的有序信息系统，都提出了相应的基于优势关系的粗糙集扩展模型. Shao 等将优势粗糙集模型引入到不完备有序信息系统中，定义了一种新的优势关系(满足自反性，不一定满足对称性和传递性)，并给出了不完备有序信息系统下基于优势粗糙集的规则提取和知识约简方法[11]. Yang 等在不完备有序信息系统中提出了两种不同的相似优势关系(自反、传递)，并提出了四种近似分布约简. 并将优势关系的概念引入到多粒度空间中，在不完备信息系统中构建了基于优势关系的多粒度粗糙集模型[12]. Qian 等在集值信息系统中引入了两种优势关系，建立了合取、析取型集值有序信息系统模型，并分别讨论了这两类信息系统中基于优势粗糙集模型的属性约简与规则提取方法[13]. 杨习贝等提出了基于可变精度优势关系下的析取集值有序信息系统，并利用对象在信息系统的整体优势度引入了一种对象排序的新方法[14]. Qian 等考虑集值信息系统中对象之间的优势程度，提出了模糊优势关系的概念，建立了集值信息系统中的模糊优势关系粗糙集模型[15]. 由于区间值信息系统是单值信息系统的另外一种扩展模型，他们还研究了区间值有序信息系统中的优势关系，建立了区间有序信息系统并给出了相应的知识获取方法[16]. Yang 等将优势粗糙集模型扩展到不完备区间值信息系统中，并通过六种不同的相对约简，提出了最优决策规则获取方法[17]. 杨青山等在区间值信息系统中分析讨论了现有的四种优势关系各自存在的局限性，利用概率学的基本原理提出了一种新的优势关系，建立了基于 α-优势关系的扩展粗糙集模型[18].

在实际应用中，信息系统中数据的形式是多样的，对象间的关系也是多样化的. 如何构造一个粗糙集模型可以综合考虑各种数据类型的特点、集成不同的二元关系，已成为目前粗糙集理论研究中的一个研究热点. An 和 Tong 考虑到一个信息系统中可能同时存在名义型、定量型、偏序型三种类型的数据，进而提出了一种基于不可分辨关系，相似关系和优势关系的全局二元关系，并讨论了全局二元关系下的上、下近似的定义及其相关性质[19]. Abu-Donia 基于多类型知识库(自反关系、相容关系、优势关系、等价关系)，通过多种二元关系和多种近似算子的交集两种方式，构造了两种不同类型的近似算子，并讨论了其相关性质[20]. Zhang 等考虑到信息系统中存在多种类型的数据形式(数值型、集值、区间值、缺失值等)，给出了复合信息系统的定义，并通过矩阵的方法[21]构造了基于复合二元关系的上、下近似算子[22]. Chen 等进一步提出了基于复合关系的概率粗糙集模型，并研究了该模型下的最大分布约简方法[23]. 本文以有序信息系统为研究对象，考虑信息系统中可能存在多种数据类型的属性值(单值、集值、区间值)以及相对应的偏序关系，建立复合有序信息系统，提出复合优势关系的概念，并给出基于复合有序信息系统下的优势粗糙集扩展模型. 最后，利用关系的矩阵表示和矩阵运算为该模型中上、下近似集的求解问题提供一种直观、有效的计算方法.

2 粗糙集及有序信息系统

2.1 单值有序信息系统

定义 1[2]. 称 $S=(U,C\cup\{d\},V,f)$ 是单值信息系统，其中，$U=\{x_1,x_2,\cdots,x_n\}$ 为非空有限的对象集合，也称为论域；$C=\{a_1,a_2,\cdots,a_m\}$ 为非空有限的属性集合，称为条件属性集，$\{d\}$ 为决策属性，并且 $C\cap\{d\}=\varnothing$；$V=\cup_{a\in C}V_a$ 是属性值的集合，V_a 表示条件属性 $a\in C$ 的值域，V_d 表示决策属性的值域；$f:U\times A\to V$ 为对象属性值映射，满足 $\forall_{x\in U,a\in C}|f(x,a)|=1$，其中，$\forall a\in C,x\in U,f(x,a)$ 指定 U 中的对象 x 关于条件属性 a 的属性值，$|\cdot|$ 表示集合的势.

定义 2[2]. 设 $S=(U,C\cup\{d\},V,f)$ 是一个单值信息系统，若所有的条件属性都是递增或递减偏好有序的，且决策属性 d 的属性值表示对象的全序关系，则称 S 为单值有序信息系统.

定义 3[2]. 设 $S=(U,C\cup\{d\},V,f)$ 是一个单值有序信息系统，$\forall A\subseteq C$，则定义优势关系为 $R_A^{\geqslant}$：

$$R_A^{\geqslant}=\{(y,x)\in U\times U|f(y,a)\geqslant f(x,a)(\forall a\in A)\}\quad(1)$$

2.2 集值有序信息系统

定义 4[13]. 称 $S=(U,C\cup\{d\},V,f)$ 是集值信息系统，其中，$U=\{x_1,x_2,\cdots,x_n\}$ 为非空有限的对象集合，也称为论域；$C=\{a_1,a_2,\cdots,a_m\}$ 为非空有限的属性集合，称为条件属性集，$\{d\}$ 为决策属性，并且 $C\cap\{d\}=\varnothing$；$V=\cup_{a\in C}V_a$ 是属性值的集合，V_a 表示属性 $a\in C$ 的值域，V_d 表示决策属性的值域；$f:U\times A\to 2^{V_C}$ 为对象属性值映射，是一个集值映射，满足 $\forall_{x\in U,a\in C}|f(x,a)|\geqslant 1$，其中，$\forall a\in C,x\in U,f(x,a)$ 指定 U 中的对象 x 关于条件属性 a 的属性值，$|\cdot|$ 表示集合的势.

定义 5[13]. 设 $S=(U,C\cup\{d\},V,f)$ 是一个集值信息系统，若所有的条件属性都是递增或递减偏好有序的，且决策属性 d 的属性值表示对象的全序关系，则称 S 为集值有序信息系统.

定义 6[13]. 设 $S=((U,C\cup\{d\},V,f)$ 是一个集值有序信息系统，$\forall A\subseteq C$，则定义优势关系为 $R_A^{\supseteq}$：

$$R_A^{\supseteq}=\{(y,x)\in U\times U|f(y,a)\supseteq f(x,a)(\forall a\in A)\}\quad(2)$$

2.3 区间值有序信息系统

定义 7[16]. 称 $S=(U,C\cup\{d\},V,f)$ 是区间值信息系统，其中，$U=\{x_1,x_2,\cdots,x_n\}$ 为非空有限的对象集合，也称为论域；$C=\{a_1,a_2,\cdots,a_m\}$ 为非空有限的属性集合，称为条件属性集，$\{d\}$ 为决策属性，并且 $C\cap\{d\}=\varnothing$；$V=\cup_{a\in A}V_a$ 是属性值的集合，V_a 表示属性 $a\in A$ 的值域，V_d 表示决策属性的值域；$f:U\times A\to V$ 为对象属性值映射，满足 $\forall_{x\in U,a\in C}(f(x,a)\subseteq V_a\wedge f(x,a)=[a^L(x)a^U(x)])$，其中，$\forall a\in C,x\in U,f(x,a)$ 指定 U 中的对象 x 关于条件属性 a 的属性值，$|\cdot|$ 表示集合的势；$a^L(x),a^U(x)\in R$，$a^L(x)$ 和 $a^U(x)$ 分别表示对象 x 在属性 a 下区间取值的下边界和上边界，$a^L(x)\leqslant a^U(x)$.

定义 8[16]. 设 $S=(U,C\cup\{d\},V,f)$ 是一个区间值信息系统，若所有的条件属性都是递增或递减偏好有序的，且决策属性 d 的属性值表示对象的全序关系，则称 S 为区间值有序信息系统.

定义 9[16]. 设 $S=(U,C\cup\{d\},V,f)$ 是一个区间值有序信息系统，$\forall A\subseteq C$，则定义优势关系为 $R_A^{\triangleright}$：

$$R_A^{\triangleright}=\{(y,x)\in U\times U|a^L(y)\geqslant a^L(x),a^U(y)\geqslant a^U(x)$$

$$(\forall a \in A)\}\tag{3}$$

3　复合有序信息系统

定义 10. 称 $S=(U,C\cup\{d\},V,f)$ 是复合信息系统,其中,$U=\{x_1,x_2,\cdots,x_n\}$ 为非空有限的对象集合,也称为论域;$C=\{a_1,a_2,\cdots,a_m\}$ 为非空有限的属性集合,称为条件属性集,$\{d\}$ 为决策属性,并且 $C\cap\{d\}=\emptyset$;$V=\cup_{C_i\subseteq C}V_{C_i}$ 是属性值的集合,其中,C_i 表示同一数据类型的属性集,$C=\cup C_i$,$C_i\cap C_j=\emptyset(i\neq j)$. V_{C_i} 表示条件属性集 C_i 的值域,并且 $V_{C_i}\cap V_{C_j}=\emptyset(i\neq j)$,$V_d$ 表示决策属性的值域,$f:U\times A\to V$ 为对象属性值映射,也称为信息函数,$\forall a\in A,x\in U,f(x,a)$ 指定 U 中的对象 x 关于条件属性 a 的属性值.

定义 11. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合信息系统,若所有的条件属性都是递增或递减偏好有序的,且决策属性 d 的属性值表示对象的全序关系,则称 S 为复合有序信息系统.

例 1. 表 1 给出了一个复合有序信息系统,其中 $U=\{x_1,x_2,x_3,x_4,x_5,x_6\}$ 为论域,$C=\{a_1,a_2,a_3,a_4,a_5\}$ 为条件属性集,d 为决策属性. 根据属性值的数据类型不同,条件属性集 C 可以划分为 $C_1=\{a_1,a_2\}$,$C_2=\{a_3,a_4\}$,$C_3=\{a_5\}$ 分别对应单值、集值、区间值数据类型的属性值.

表 1　一个复合有序信息系统

Table 1　An composite ordered information system

U	a_1	a_2	a_3	a_4	a_5	d
x_1	1	0.12	$\{E,F,G\}$	$\{F,G\}$	$[75,84]$	1
x_2	3	0.35	$\{E,F\}$	$\{F,G\}$	$[65,80]$	3
x_3	2	0.26	$\{E,F,G\}$	$\{E,F,G\}$	$[85,90]$	3
x_4	5	0.25	$\{E,F\}$	$\{F\}$	$[74,85]$	2
x_5	1	0.05	$\{E\}$	$\{F\}$	$[75,70]$	1
x_6	4	0.50	$\{E,F,G\}$	$\{F,G\}$	$[65,93]$	2

定义 12. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,$\forall x,y\in U,A=\cup A_k\subseteq C,A\subseteq C_k$,则定义复合优势关系为 $CR_A^{\succ}$:

$$CR_A^{\succ}=\{(y,x)\mid(y,x)\in\bigcap_{A_k\in A}R_{A_k}^{*}\}\tag{4}$$

其中,$R_{A_k}^{*}$ 表示同一数据类型的属性集 A_k 上的优势关系.

由以上定义可以看出,复合有序关系 $CR_A^{\succ}$ 满足自反性和传递性,而不一定满足对称性.

根据复合优势关系,对于 $\forall x\in U$,可以构建复合优势类和复合劣势类两种集合如下所示:

1) $CR_A^{+}(x)=\{y\in U\mid(y,x)\in CR_A^{\succ}\}$ 表示根据属性集合 A 中的各种优势关系,所有优于对象 x 的对象集合;

2) $CR_A^{-}(x)=\{y\in U\mid(x,y)\in CR_A^{\succ}\}$ 表示根据属性集合 A 中的各种优势关系,所有劣于对象 x 的对象集合.

定理 1. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,$\forall x\in U,A=\cup A_k\subseteq C,A_k\subseteq C_k$,则有:

1) $CR_A^{\succ}$ 满足自反性和传递性,而不一定满足对称性;

2) $CR_A^{+}(x)=\bigcap_{A_k\in A}CR_{A_k}^{+}(x)$,$CR_A^{-}(x)=\bigcap_{A_k\in A}CR_{A_k}^{-}(x)$.

例 2. 对于表 1 所列的复合有序信息系统 $S=(U,C\cup\{d\},V,f)$,$C=\bigcup_{1\leqslant i\leqslant 3}C_i$,通过在复合有序信息系统中两个对象之间复合优势关系的定义,可得出表 1 在复合优势关系 $CR_C^{\succ}$ 下各对象的优势类和劣势类如下:

优势类:$CR_C^{+}(x_1)=\{x_1,x_3\}$,$CR_C^{+}(x_2)=\{x_2,x_6\}$,$CR_C^{+}(x_3)=\{x_3\}$,$CR_C^{+}(x_4)=\{x_4\}$,$CR_C^{+}(x_5)=\{x_1,x_3,x_5\}$,$CR_C^{+}(x_6)=\{x_6\}$;

劣势类:$CR_C^{-}(x_1)=\{x_1,x_5\}$,$CR_C^{-}(x_2)=\{x_2\}$,$CR_C^{-}(x_3)=\{x_1,x_3,x_5\}$,$CR_C^{-}(x_4)=\{x_4\}$,$CR_C^{-}(x_5)=\{x_5\}$,$CR_C^{-}(x_6)=\{x_2,x_6\}$.

在复合有序信息系统 $S=(U,C\cup\{d\},V,f)$ 中,根据决策属性 d 可将论域 U 划分为有穷个数的决策类集合:$Cl=\{Cl_t,t\in T\}$,$T=\{1,\cdots,n\}$,对于任意 $x\in U$ 属于且只属于一个决策类 $Cl_t\in Cl$. 在基于优势关系的粗糙集方法中,我们考虑的是决策类的向上联集和向下联集的集合近似.

定义 13. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,决策类 Cl_t 的向上联集和向下联集定义为:

$$Cl_t^{\geqslant}=\cup_{s\geqslant t}Cl_s,\ Cl_t^{\leqslant}=\cup_{s\leqslant t}Cl_s,\ t=1,\cdots,n\tag{5}$$

其中,对于任意的 $r,s\in t$,若 $r>s$,则认为 Cl_r 中的对象要优于 Cl_s 中的对象. 对于任意的 $x\in U$,如果 $x\in Cl_t^{\geqslant}$,则表示 x 至少属于决策类 Cl_t,如果 $x\in Cl_t^{\leqslant}$,则表示 x 至多属于决策类 Cl_t.

定义 14. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,若有 $A\subseteq C$,则对于任意的 $Cl_t^{\geqslant}(l\leqslant t\leqslant n)$,$Cl_t^{\geqslant}$ 的下近似、上近似定义为:

$$\underline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x\in U\mid CR_A^{+}(x)\subseteq Cl_t^{\geqslant}\},$$
$$\overline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x\in U\mid CR_A^{-}(x)\cap Cl_t^{\geqslant}\neq\emptyset\}\tag{6}$$

定义 15. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,若有 $A\subseteq C$,则对于任意的 $Cl_t^{\leqslant}(1\leqslant t\leqslant n)$,$Cl_t^{\leqslant}$ 的下近似、上近似定义为:

$$\underline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x\in U\mid CR_A^{-}(x)\subseteq Cl_t^{\leqslant}\},$$
$$\overline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x\in U\mid CR_A^{+}(x)\cap Cl_t^{\leqslant}\neq\emptyset\}\tag{7}$$

其中,$CR_A^{+}(x)$ 和 $CR_A^{-}(x)$ 分别表示对象 x 关于条件属性集合 A 的复合优势类和复合劣势类.

例 3. 对于表 1 所列的复合有序信息系统 $S=(U,C\cup\{d\},V,f)$,根据决策属性 d 可得到论域上的划分如下:$Cl=\{Cl_1,Cl_2,Cl_3\}=\{\{x_1,x_5\},\{x_4,x_6\},\{x_2,x_3\}\}$. 这里我们考虑决策类 Cl_2 的上、下联集的近似集合,根据定义 2 可得 $Cl_2^{\geqslant}=\{x_2,x_3,x_4,x_6\}$,$Cl_2^{\leqslant}=\{x_1,x_4,x_5,x_6\}$. 则通过定义 2 和定义 3,我们有:$\underline{CR_A^{\succ}}(Cl_2^{\geqslant})=\{x_3,x_4,x_6\}$,$\overline{CR_A^{\succ}}(Cl_2^{\geqslant})=\{x_2,x_3,x_4,x_6\}$;$\underline{CR_A^{\succ}}(Cl_2^{\leqslant})=\{x_1,x_4,x_5\}$,$\overline{CR_A^{\succ}}(Cl_2^{\leqslant})=\{x_1,x_2,x_4,x_5,x_6\}$.

4　复合有序信息系统中基于矩阵的近似集计算方法

定义 16. 设 $U=\{x_1,x_2,\cdots,x_n\}$ 是非空有限论域的集合,X 是论域 U 的任意子集. 则 X 可用布尔列矩阵 $C(X)$ 表示如下:

$$C(X)=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix},c_i=\begin{cases}1,x_i\in X\\0,x_i\notin X\end{cases}\tag{8}$$

简记作: $C(X)=[c_1,c_2,\cdots,c_n]^T$,其中"$T$"表示矩阵的转置.

例4. 对于表1所列的复合有序信息系统 $S=(U,C\cup\{d\},V,f)$, $U=\{x_1,x_2,x_3,x_4,x_5,x_6\}$, $Cl_2^{\geqslant}=\{x_2,x_3,x_4,x_6\}$. 则论域 U 的子集 $Cl_2^{\geqslant}$ 可表示为 $C(Cl_2^{\geqslant})=[0,1,1,1,0,1]^T$.

定义17. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,若有 $A\subseteq C$, $CR_A^{\succ}$ 为由 A 决定的复合优势关系,则可定义关于属性集 A 的优势矩阵和劣势矩阵分别为 $M_A^{\succ}=[m_{ij}]_{n\times n}$ 和 $M_A^{\prec}=[m'_{ij}]_{n\times n}$,其中

$$m_{ij}=\begin{cases}1, x_j\in CR_A^+(x_i)\\ 0, x_j\notin CR_A^+(x_i)\end{cases}, m'_{ij}=\begin{cases}1, x_j\in CR_A^-(x_i)\\ 0, x_j\notin CR_A^-(x_i)\end{cases} \quad (9)$$

定义18. 设 $S=(U,C\cup\{d\},V,f)$ 是 s 一个复合有序信息系统,若有 $A,B\subseteq C$, $CR_A^{\succ}$ 和 $CR_B^{\succ}$ 分别为 A 和 B 的复合优势关系, $M_A^{\Delta}=[\alpha_{ij}]_{n\times n}$ 和 $M_B^{\Delta}=[\alpha_{ij}]_{n\times n}$ $(\Delta=>,<)$ 是 $CR_A^{\succ}$ 和 $CR_B^{\succ}$ 的优势矩阵(劣势矩阵),则矩阵的与运算"$\wedge$"可定义为

$$M_A^{\Delta}\wedge M_B^{\Delta}=[\gamma_{ij}]_{n\times n}, \text{其中 } \gamma_{ij}=\alpha_{ij}\wedge\beta_{ij}\ (i,j=1,2,\cdots,n)$$

定理2. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统, $A=\cup A_k\subseteq C$, $A_k\subseteq C_k$, $CR_A^{\succ}$ 为由 A 决定的复合优势关系, $M_A^{\succ}=[m_{ij}]_{n\times n}$ 和 $M_A^{\prec}=[m'_{ij}]_{n\times n}$ 是 $CR_A^{\succ}$ 的优势矩阵和劣势矩阵,则:

$$M_A^{\succ}=\bigwedge_{A_k\in A}M_{A_k}^{\succ}, M_A^{\prec}=\bigwedge_{A_k\in A}M_{A_k}^{\prec}$$

定义19. 设 $M_A^{\succ}=[m_{ij}]_{n\times n}$ 和 $M_A^{\prec}=[m_{ij}]_{n\times n}$ 是关于复合优势关系 $CR_A^{\succ}$ 的优势矩阵和劣势矩阵,由 $M_A^{\succ}$ 和 $M_A^{\prec}$ 所诱导的对角矩阵 $\Lambda_A^{\succ}$, $\Lambda_A^{\prec}$ 分别定义为

$$\Lambda_A^{\succ}=\begin{bmatrix}\frac{1}{\sum_{k=1}^n m_{1k}} & 0 & \cdots & 0\\ 0 & \frac{1}{\sum_{k=1}^n m_{2k}} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sum_{k=1}^n m_{nk}}\end{bmatrix},$$

$$\Lambda_A^{\prec}=\begin{bmatrix}\frac{1}{\sum_{k=1}^n m'_{1k}} & 0 & \cdots & 0\\ 0 & \frac{1}{\sum_{k=1}^n m'_{2k}} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sum_{k=1}^n m'_{nk}}\end{bmatrix}.$$

简记作: $\Lambda_A^{\succ}=diag\left[\frac{1}{\lambda_1},\frac{1}{\lambda_2},\cdots,\frac{1}{\lambda_n}\right]$ 和 $\Lambda_A^{\succ}=diag\left[\frac{1}{\lambda'_1},\frac{1}{\lambda'_2},\cdots,\frac{1}{\lambda'_n}\right]$,

其中, $\lambda_i=\sum_{j=1}^n m_{ij}$, $\lambda'_i=\sum_{j=1}^n m'_{ij}$ $(1\leqslant i\leqslant n)$.

定理3. $\Lambda_A^{\succ}=diag\left[\frac{1}{|CR_A^+(x_1)|},\frac{1}{|CR_A^+(x_2)|},\cdots,\frac{1}{|CR_A^+(x_n)|}\right]$, $\Lambda_A^{\prec}=diag\left[\frac{1}{|CR_A^-(x_1)|},\frac{1}{|CR_A^-(x_2)|},\cdots,\frac{1}{|CR_A^-(x_n)|}\right]$,其中 $1\leqslant|CR_A^+(x_i)|$, $|CR_A^-(x_i)|\leqslant n$ $(1\leqslant i\leqslant n)$, $|CR_A^+(x_i)|$ 和 $|CR_A^-(x_i)|$ 分别表示根据复合优势关系 $CR_A^{\succ}$,论域 U 中所有优于对象 x_i 的对象个数.

例5. 根据定义2和引理2,可计算例2中优势矩阵 $M_A^{\succ}$ 的诱导对角矩阵如 $\Lambda_A^{\succ}=diag\left[\frac{1}{2},\frac{1}{2},1,1,\frac{1}{3},1\right]$.

定义20. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,若有 $A\subseteq C$, $CR_A^{\succ}$ 为由 A 决定的复合优势关系, $M_A^{\succ}=[m_{ij}]_{n\times n}$ 和 $M_A^{\prec}=[m'_{ij}]_{n\times n}$ 是 $CR_A^{\succ}$ 的优势矩阵和劣势矩阵, $\Lambda_A^{\succ}$ 和 $\Lambda_A^{\prec}$ 是由 $M_A^{\succ}$ 和 $M_A^{\prec}$ 所诱导的对角矩阵,则对于任意的 $Cl_t^{\geqslant}$ $(1\leqslant t\leqslant n)$ 有

1) $Cl_t^{\geqslant}$ 关于 $CR_A^{\succ}$ 的下近似集的列矩阵可定义为:

$$L_A^{\succ}(Cl_t^{\geqslant})=\Lambda_A^{\succ}\bullet(M_A^{\succ}\bullet C(Cl_t^{\geqslant}));$$

2) $Cl_t^{\geqslant}$ 关于 $CR_A^{\succ}$ 的上近似集的列矩阵可定义为:

$$U_A^{\succ}(Cl_t^{\geqslant})=M_A^{\prec}\bullet C(Cl_t^{\geqslant}).$$

其中,"$\bullet$"表示矩阵的数量积.

定义21. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,若有 $A\subseteq C$, $CR_A^{\succ}$ 为由 A 决定的复合优势关系, $M_A^{\succ}=[m_{ij}]_{n\times n}$ 和 $M_A^{\prec}=[m'_{ij}]_{n\times n}$ 是 $CR_A^{\succ}$ 的优势矩阵和劣势矩阵, $\Lambda_A^{\succ}$ 和 $\Lambda_A^{\prec}$ 是由 $M_A^{\succ}$ 和 $M_A^{\prec}$ 所诱导的对角矩阵,则对于任意的 $Cl_t^{\leqslant}$ $(1\leqslant t\leqslant n)$ 有

1) $Cl_t^{\leqslant}$ 关于 $CR_A^{\succ}$ 的下近似集的列矩阵可定义为:

$$L_A^{\succ}(Cl_t^{\leqslant})=\Lambda_A^{\prec}\bullet(M_A^{\prec}\bullet C(Cl_t^{\leqslant}));$$

2) $Cl_t^{\leqslant}$ 关于 $CR_A^{\succ}$ 的上近似集的列矩阵可定义为:

$$U_A^{\succ}(Cl_t^{\leqslant})=M_A^{\succ}\bullet C(Cl_t^{\leqslant}).$$

例6. 对于表1所列的复合有序信息系统 $S=(U,C\cup\{d\},V,f)$,我们考虑 $Cl_2^{\geqslant}$ 关于复合优势关系 $CR_C^{\succ}$ 的上、下近似集的布尔列矩阵,根据例1-3的结果可得

$$L_A^{\succ}(Cl_2^{\geqslant})=\Lambda_A^{\succ}\bullet(M_A^{\succ}\bullet C(Cl_2^{\geqslant}))=\begin{bmatrix}1/2&0&0&0&0&0\\0&1/2&0&0&0&0\\0&0&1&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1/3&0\\0&0&0&0&0&1\end{bmatrix}\bullet$$

$$\left(\begin{bmatrix}1&0&1&0&0&0\\0&1&0&0&0&1\\0&0&1&0&0&0\\0&0&0&1&0&0\\1&0&1&0&1&0\\0&0&0&0&0&1\end{bmatrix}\bullet\begin{bmatrix}0\\1\\1\\1\\0\\1\end{bmatrix}\right)=\left[\frac{1}{2},1,1,1,\frac{1}{3},1\right]^T$$

$$U_A^{\succ}(Cl_2^{\geqslant})=M_A^{\prec}\bullet C(Cl_2^{\geqslant})=\begin{bmatrix}1&0&0&0&1&0\\0&1&0&0&0&0\\1&0&1&0&1&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&1&0&0&0&1\end{bmatrix}\bullet\begin{bmatrix}0\\1\\1\\1\\0\\1\end{bmatrix}=$$

$$[0,1,1,1,0,2]^T$$

同理,对于 $Cl_2^{\leqslant}$,根据定义2我们有: $L_A^{\succ}(Cl_2^{\leqslant})=[1,0,2/3,1,1,1/2]^T$, $U_A^{\succ}(Cl_t^{\leqslant})=[1,1,0,1,2,1]^T$.

定理 4. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,$\forall A\subseteq C$,对于任意的 $Cl_t^{\geqslant}(1\leqslant t\leqslant n)$,$Cl_t^{\geqslant}$ 的下、上近似集可分别由下式得出

$$\underline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x_i\in U\,|\,l_i=1,l_i\in L_A^{\succ}(Cl_A^{\geqslant})(1\leqslant i\leqslant n)\}$$

$$\overline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x_i\in U\,|\,\mu_i\geqslant 1,\mu_i\in U_A^{\succ}(Cl_A^{\geqslant})(1\leqslant i\leqslant n)\}$$

其中,$L_A^{\succ}(Cl_A^{\geqslant})$ 和 $U_A^{\succ}(Cl_A^{\geqslant})$ 分别表示 $Cl_t^{\geqslant}$ 的下、上近似集的布尔列矩阵.

定理 5. 设 $S=(U,C\cup\{d\},V,f)$ 是一个复合有序信息系统,$\forall A\subseteq C$,对于任意的 $Cl_t^{\leqslant}(1\leqslant t\leqslant n)$,$Cl_t^{\leqslant}$ 的下、上近似集可分别由下式得出

$$\underline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x_i\in U\,|\,l_i=1,l_i\in L_A^{\succ}(Cl_A^{\leqslant})(1\leqslant i\leqslant n)\}$$

$$\overline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x_i\in U\,|\,\mu_i\geqslant 1,\mu_i\in U_A^{\succ}(Cl_A^{\leqslant})(1\leqslant i\leqslant n)\}$$

其中,$L_A^{\succ}(Cl_A^{\leqslant})$ 和 $U_A^{\succ}(Cl_A^{\leqslant})$ 分别表示 $Cl_t^{\leqslant}$ 的下、上近似集的布尔列矩阵.

例 7. 对于表 1 所列的复合有序信息系统 $S=(U,C\cup\{d\},V,f)$,我们考虑 $Cl_2^{\geqslant}$ 和 $Cl_2^{\leqslant}$ 关于复合优势关系 $CR_C^{\succ}$ 的上、下近似集,根据例 1-3 的结果,$Cl_2^{\geqslant}$ 和 $Cl_2^{\leqslant}$ 的上、下近似集的列矩阵如下所示.

$\underline{CR_A^{\succ}}(Cl_t^{\geqslant})$	$\overline{CR_A^{\succ}}(Cl_t^{\geqslant})$	$\underline{CR_A^{\succ}}(Cl_t^{\leqslant})$	$\overline{CR_A^{\succ}}(Cl_t^{\leqslant})$
$\begin{bmatrix}1/2\\1\\1\\1\\1/3\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\\1\\1\\0\\2\end{bmatrix}$	$\begin{bmatrix}1\\0\\2/3\\1\\1\\1/2\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\\1\\2\\1\end{bmatrix}$

$Cl_2^{\geqslant}$ 和 $Cl_2^{\leqslant}$ 的上、下近似集可直接由定理 1 和 2 得到,如:

$\underline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x_2,x_3,x_4,x_6\}$,$\overline{CR_A^{\succ}}(Cl_t^{\geqslant})=\{x_2,x_3,x_4,x_6\}$;

$\underline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x_1,x_4,x_5\}$,$\overline{CR_A^{\succ}}(Cl_t^{\leqslant})=\{x_1,x_2,x_4,x_5,x_6\}$.

5　结束语

在实际应用中,一个信息系统中所含有的数据类型可能是多样化的,如:单值、集值、区间值等. 而针对不同类型的数据,对象之间可能具有不同的二元关系. 本文针对有序信息系统,在综合考虑系统中可能存在各种不同类型数据的基础上,建立了复合有序信息系统. 并在此基础上,研究了复合有序信息系统中不同属性下对象间所具有的不同偏序关系,提出了一种新的优势关系—复合优势关系,进而将单一关系的优势关系粗糙集模型扩展到了多种关系相结合的优势关系粗糙集模型. 最后,基于复合优势关系粗糙集模型,提出了一种基于矩阵运算的近似集计算方法,并通过实例验证了所提方法 S 的有效性和可行性. 在未来的工作中,如何在复合优势关系粗糙集模型下进行有效地知识获取和规则提取将是我们研究的重点.

References:

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

[2] Duntsch I,Gediga G. Uncertainty measures of rough set prediction [J]. Artificial Intelligence,1998,106(1):109-137.

[3] Yao Y Y. Three-way decisions with probabilistic rough sets [J]. Information Sciences,2010,180(3):341-353.

[4] Yao Y Y. Two semantic issues in a probabilistic rough set model [J]. Fundamenta Informaticae,2011,108(3-4):249-265.

[5] Pawlak Z,Skowron A. Rudiments of rough sets [J]. Information Science,2007,177(1):28-40.

[6] Iyer N S. A family of dominance rules for multi-attribute decision making under uncertainty [J]. IEEE Transactions on Systems,Man,and Cybernetics Part A,2003,14(2):191-201

[7] Greco S,Matarazzo B,Slowinski R. Rough approximation of preference relation by dominance relation [J]. European Journal of Operational Research,1999,117(1):63-68.

[8] Greco S,Matarazzo B,Slowinski R. Rough sets methodology for sorting problems in presence of multiple attributes and criteria [J]. European Journal of Operational Research,2002,138(2):247-259.

[9] Hu Q H,Gao M Z,Yu D R,et al. Information entropy for ordinal classification [J]. Science in China F:Information Sciences,2010,53(6):1188-1200.

[10] Xu W H,Zhang X Y,Zhang W X. Knowledge granulation [J]. Knowledge Entropy and Knowledge Uncertainty Measure in Ordered Information Systems,2009,9(4):1244-1251.

[11] Shao M W,Zhang W X. Dominance Relation and Rules in an incomplete ordered information system [J]. International Journal of Intelligence Systems,2005,20(1):13-27.

[12] Yang X B,Yang J Y,Wu C,et al. Dominance-based rough set approach and knowledge reductions in incomplete ordered information system [J]. Information Sciences,2008,178(4):1219-1234.

[13] Qian Y H,Dang C Y,Liang J Y,et al. Set-valued ordered information systems [J]. Information Sciences,2009,179(16):2809-2832.

[14] Qian Y H,Liang J Y. On dominance relations in disjunctive set-valued ordered information systems [J]. International Journal of Information Technology & Decision Making,2010,9(1):9-33.

[15] Yang Xi-bei,Zhang Zai-yue,Zhang Ming. Fuzzy dominance-based rough set in set-valued information system [J]. Computer Science,2011,38(2):234-237.

[16] Qian Y H,Liang J Y,Dang C Y. Interval ordered information systems [J]. Computers and Mathematics with Applications,2008,56(8):1994-2009.

[17] Yang X B,Yu D J,Yang J Y,et al. Dominance-based rough set approach to incomplete interval-valued information system [J]. Data & Knowledge Engineering,2009,68(11):1331-1347.

[18] Yang Qing-shan,Wang Guo-yin,Zhang Qing-hua,et al. The interval-valued rough set extended model based on the dominance relation [J]. Journal of Shandong University(Natural Science),2010,45(9):7-13.

[19] An L P,Tong L Y. Rough approximations based on intersection of indiscernibility,similarity and outranking relations [J]. Knowledge-Based Systems,2010,23(6):555-562.

[20] Abu-Donia H M. Multi knowledge based rough approximations and applications [J]. Knowledge-Based Systems,2012,26(1):20-29.

[21] 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.

[22] Zhang J B,Li T R,Chen H M. Composite rough sets [C]. Proceedings of the 4th International Conference on Artificial Intelligence and Computational Intelligence,2012:150-159.

[23] Chen H M,Li T R,Zhang J B,et al. Probabilistic composite rough set and attribute reduction [C]. Proceedings of the 7th International Conference on Intelligent Systems and Knowledge Engineering,2012.

附中文参考文献:

[15] 杨习贝,张再跃,张　明. 集值信息系统中的模糊优势关系粗糙集 [J]. 计算机科学,2011,38(2):234-237.

[18] 杨青山,王国胤,张清华,等. 基于优势关系的区间值粗糙集扩充模型 [J]. 山东大学学报,2010,45(9):7-13.