基于改进的可能度优势关系的排序方法研究

张其文 王雪勤 庄新磊
(兰州理工大学计算机与通信学院 兰州 730050)

摘 要 优势关系是解决多属性决策问题的一种重要方法,也是研究不完备区间值信息系统的一种重要方法。在不完备区间值信息系统排序方法的研究中,针对属性过多从而可能引起排序失效的问题,提出了一种改进的可能度优势关系排序方法。这种新的排序方法把区间值优势关系和可能度相结合来改进经典优势关系定义的不足,并给出了平均综合优势度的定义。最后,通过与 ∝-β 优势关系和容差优势关系等排序方法在具体算例中的比较分析,证明了改进的可能度优势关系排序方法不仅能够解决不完备区间值信息系统中的问题,而且能够使排序结果更合理、有效。

关键词 区间值信息系统,优势关系,多属性决策,不完备信息系统,可能度

中图法分类号 TP301 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.11.056

Research of Ranking Method Based on Improved Possible Degree Dominance Relation

ZHANG Qi-wen WANG Xue-qin ZHUANG Xin-lei
(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract The dominance relation is an important method of solving the multi-attribute decision problem and researching the incomplete interval-valued information systems. According to the research of the ranking methods in the incomplete interval valued information system, we proposed a ranking method of improved possible degree dominance relation to solve the problem that too many attributes maybe result in the ranking failure. This new ranking method combines the interval-valued dominance relation with the possible degree, improves the weakness of the definition of classic dominance relation and defines the average comprehensive dominance degree. Finally, compared with other ranking methods such as ∝-β dominance relation and tolerance dominance relation in specific cases, this ranking method of improved possible degree dominance relation can not only solve the problem in the incomplete interval-valued information systems, but also make the ranking results more reasonable and effective.

Keywords Interval-valued information systems, Dominance relation, Multi-attribute decision, Incomplete information systems, Possible degree

1 引言

粗糙集理论是由波兰学者 Pawlak[1] 于 1982 提出的一种数学分析理论。经典的粗糙集理论是基于等价关系建立起来的,并且它只应用于完备系统中。而现实生活中,许多待处理的信息都是不完备且以区间值的形式存在。对此,Greco 等人提出了基于优势关系的粗糙集(DRSA)[2-4],利用优势关系代替等价关系对传统粗糙集理论进行扩展研究,此后,扩展优势关系、相似优势关系和置信优势关系等[5]相继被提出以解决不完备偏好决策问题。目前,优势关系已经广泛用于对决策对象进行排序,如何找到一种区分度高、快速、合理和有效的排序方法是一个值得研究的重要方向。

关于利用优势关系对决策对象进行排序的研究越来越多。如文献[6]针对属性权重未知的多属性决策问题提出用属性值的优势关系确定属性权重,并利用优势关系对决策对象进行排序并择优;文献[7]针对经典优势关系过于严格可能导致排序失效的问题,提出容差优势关系的概念;文献[8]利用熵权和优势关系的结合建立一种综合排序方法;文献[9]针对不完备的多属性决策问题,提出一种基于优势关系的排序方法;文献[10]提出一种基于概率可信度的区间数排序方法;文献[11]针对当前区间序信息系统上 5 种优势关系的不足,提出了∂可能度优势关系;文献[12]针对属性值是区间数的多属性决策问题提出了一种新的 ∝-β 优势关系。

本文在上述研究工作的基础上,针对不完备区间值信息系统中存在的问题,提出一种改进的可能度优势关系排序方法。这种新的方法使排序结果更合理,有更强的区分度,能够更好地解决排序问题。

2 不完备区间值信息系统

定义 1[13] 设 $S=(U,A,V,f)$是区间值信息系统,其中 $U=\{x_1,x_2,\cdots,x_n\}$是一个非空有限对象集,称为论域;$A=\{a_1,a_2,\cdots,a_m\}$ 表示所有属性的集合;$V=\bigcup_{a\in A}V_a$,V_a 表示属性 a 的取值范围,在属性 a 下对象 x 的区间值为 $f(x,a)=$

到稿日期:2014-11-15 返修日期:2015-03-13 本文受甘肃省自然科学基金(2011GS04147),国家自然科学基金(61263047)资助。

张其文(1975-),男,副教授,硕士生导师,主要研究方向为 Ontology、智能信息处理技术、语义 Web、知识工程等;王雪勤(1989-),女,硕士,主要研究方向为数据挖掘;庄新磊(1989-),男,硕士,主要研究方向为人脸识别、稀疏表示。

$[a^L(x), a^U(x)]$，$a^L(x), a^U(x)\in R$ 且 $a^L(x)\leqslant a^U(x)$，$a^L(x)$ 和 $a^U(x)$ 分别为区间的下限和上限。当 $a^L(x)=a^U(x)$ 时，$f(x,a)$ 退化为一个实数。因此，单值是区间值的一个特殊的形式。

在区间值信息系统中，若 $V=\bigcup_{a\in A}V_a\cup\{*\}$，其中 $*$ 表示未知的属性值，那么属性 a 下对象 x 的未知区间值存在这样 3 种形式：$f(x,a)=[a^L(x),*]$，$f(x,a)=[*,a^U(x)]$，$f(x,a)=[*^L,*^U]$。这样的区间值信息系统称为不完备区间值信息系统。

目前，处理不完备区间值信息系统的方法主要有两种思路：一种是把不完备信息转化为完备信息再进行处理，另一种是直接从不完备信息系统中获取规则。而把不完备信息转化为完备信息再进行处理的方法有 3 种：删除法、填充法和扩展法[14]。本文认为未知属性是真实存在的，只是被遗漏了，所以利用填充法把不完备信息转化为完备信息再进行处理，处理办法如下[15]：

若 $f(x,a)=[a^L(x),*]$，取 $*=\max\{a^L(x),\max_{x\in U}\{a^U(x)\}\}$；

若 $f(x,a)=[*,a^U(x)]$，取 $*=\min\{a^U(x),\min_{x\in U}\{a^L(x)\}\}$；

若 $f(x,a)=[*^L,*^U]$，取 $*^L=\min_{x\in U}\{a^L(x)\}$，$*^U=\max_{x\in U}\{a^U(x)\}$。

3 可能度与优势关系

3.1 区间数优势关系

定义 2[13,16,17] 设区间值信息系统 $S=(U,A,V,f)$，$B\subseteq A$，对象的优势关系和优势类定义为：

(1)上优势关系 $R_B^{U\geqslant}$ 和优势类 $[x]_B^{U\geqslant}$：

$R_B^{U\geqslant}=\{(x,y)\in U\times U\mid a^U(x)\geqslant a^U(y),\forall a\in B\}$

$[x]_B^{U\geqslant}=\{y\in U\mid (y,x)\in R_B^{U\geqslant}\}$

(2)下优势关系 $R_B^{L\geqslant}$ 和优势类 $[x]_B^{L\geqslant}$：

$R_B^{L\geqslant}=\{(x,y)\in U\times U\mid a^L(x)\geqslant a^L(y),\forall a\in B\}$

$[x]_B^{L\geqslant}=\{y\in U\mid (y,x)\in R_B^{L\geqslant}\}$

(3)上下优势关系 $R_B^{UL\geqslant}$ 和优势类 $[x]_B^{UL\geqslant}$：

$R_B^{UL\geqslant}=\{(x,y)\in U\times U\mid a^U(x)\geqslant a^L(y),\forall a\in B\}$

$[x]_B^{UL\geqslant}=\{y\in U\mid (y,x)\in R_B^{UL\geqslant}\}$

(4)下上优势关系 $R_B^{LU\geqslant}$ 和优势类 $[x]_B^{LU\geqslant}$：

$R_B^{LU\geqslant}=\{(x,y)\in U\times U\mid a^L(x)\geqslant a^U(y),\forall a\in B\}$

$[x]_B^{LU\geqslant}=\{y\in U\mid (y,x)\in R_B^{LU\geqslant}\}$

(5)上上下下优势关系 $R_B^{UULL\geqslant}$ 和优势类 $[x]_B^{UULL\geqslant}$：

$R_B^{UULL\geqslant}=\{(x,y)\in U\times U\mid a^U(x)\geqslant a^U(y),a^L(x)\geqslant a^L(y),\forall a\in B\}$

$[x]_B^{UULL\geqslant}=\{y\in U\mid (y,x)\in R_B^{UULL\geqslant}\}$

对以上几种区间数优势关系进行分析，发现它们都存在一定的缺陷，具体缺陷如下：

(1)上优势关系定义过于宽松，如[99.99999,100]和[1,100.00001]，[1,100.00001]优于[99.99999,100]，但是事实上，[99.99999,100]优于[1,100.00001]的可能性很大，与事实不符，因此会产生不合理的结果。

(2)下优势关系定义也过于宽松，如[1,1.1]和[0.99999,100]，[1,1.1]优于[0.99999,100]，但事实上，[0.99999,100]优于[1,1.1]的可能性很大，也与事实不符，因此也会产生不合理的结果。

(3)上下优势关系定义也过于宽松，如[2,3]和[2.99,50]，[2,3]优于[2.99,50]，事实上[2.99,50]优于[2,3]可能性比较大，与事实不符。

(4)下上优势类定义太过严格，如[5.9,50]和[1,6]，它们不满足下上优势关系，但是[5.9,50]优于[1,6]的可能性很大，易出现误分。

(5)上上下下优势关系同时满足上优势关系和下优势关系，但很难实现部分区间的比较，从而导致分类的遗漏，如[2,3]和[1.99,100]无法比较，而[1.99,100]优于[2,3]的可能性很大。

通过对区间数优势关系存在的缺陷的分析，徐智明等人提出了∂-可能度优势关系，它通过∂的变化来调节优势关系的优于程度，能够弥补以上区间数优势关系过宽或过窄带来的不足。

3.2 ∂-可能度优势关系

定义 3[11] 设区间值信息系统 $S=(U,A,V,f)$，对任意的 $x_i,x_j\in U$，$a\in A$，$f(x_i,a)=[a^L(x_i),a^U(x_i)]$，$f(x_j,a)=[a^L(x_j),a^U(x_j)]$，令 $l_i^a=a^U(x_i)-a^L(x_i)$，$l_j^a=a^U(x_j)-a^L(x_j)$。

若 $f(x_i,a)$ 和 $f(x_j,a)$ 不同为实数，则

$$\partial_{ji}^a=\frac{\min\{l_i^a+l_j^a,\max\{a^U(x_j)-a^L(x_i),0\}\}}{l_i^a+l_j^a}$$

若 $f(x_i,a)$ 和 $f(x_j,a)$ 同为实数，则

$$\partial_{ji}^a=\begin{cases}0, & f(x_i,a)>f(x_j,a)\\ 0.5, & f(x_i,a)=f(x_j,a)\\ 1, & f(x_i,a)<f(x_j,a)\end{cases}$$

其中，∂_{ji}^a 为在属性 a 下对象 x_j 优于 x_i 的概率，由此可得到以下性质：

(1) $\partial_{ji}^a\in[0,1]$；

(2) $\partial_{ji}^a+\partial_{ij}^a=1$；

(3) $\partial_{ii}^a=0.5$。

定义 4 设区间值信息系统 $S=(U,A,V,f)$，$B\subseteq A$，给定可能度 $\partial_{ji}^a\in[0,1]$，则可定义∂-可能度优势关系 $R_B^{\geqslant\partial}$ 和∂-可能度优势类 $[x_i]_B^{\geqslant\partial}$ 为：

$R_B^{\geqslant\partial}=\{(x_j,x_i)\in U\times U\mid \partial_{ji}^a\geqslant\partial,\forall a\in B\}$

$[x_i]_B^{\geqslant\partial}=\{x_j\in U\mid (x_j,x_i)\in R_B^{\geqslant\partial}\}$

4 改进的可能度优势关系排序方法

虽然徐智明等人提出的∂-可能度优势关系可以通过∂的变化来调节优势关系的优于程度，并且能够在一定程度上弥补区间数优势关系过宽或过窄的不足，但是它仍然存在不足。因为，在∂-可能度优势关系中，不仅∂的确定受主观因素的影响，从而导致排序结果带有较强的主观性，而且经典优势关系定义太过严格，可能导致排序失效(即排序对象区分不开)。

因此，本文针对∂-可能度优势关系存在的不足，并在此基础上，适当放宽可比较属性的个数 N 和适当缩小可能度∂的取值范围，提出了一种改进的可能度优势关系排序方法。改进后的排序方法能够有效地避免因∂的确定带来的主观性和经典优势关系带来的排序失效问题。

4.1 可能度∂取值范围的确定

若在属性 a 下对象 x_j 优于对象 x_i 的可能度为∂_{ji}^a，则在属性 a 下对象 x_j 劣于对象 x_i 的可能度为∂_{ij}^a。若对象 x_j 在真正意义上优于对象 x_i，则有$\partial_{ji}^a \geqslant \partial_{ij}^a$，又因为$\partial_{ji}^a + \partial_{ij}^a = 1$ 且$\partial_{ji}^a \in [0,1]$，所以，$0.5 \leqslant \partial_{ji}^a \leqslant 1$，即$\partial_{ji}^a \in [0.5,1]$。

通过以上分析，$\partial \in [0.5,1]$时分类才更合理。

4.2 可比较属性的个数 N 的确定

在经典优势关系的定义中，一个对象的所有属性必须都优于另一个对象的所有对应的属性时，才能够说明这个对象比另一个对象更优。但在现实生活中，当对象的属性过多，并且对象属性之间互有优劣时，很难符合经典优势关系的定义，这很可能导致排序失效。如对象 $x=\{2,8,4,7,2,1,0,9\}$，$y=\{3,4,5,1,0,0,1,2\}$，对象 x 除了第 1，3，7 个属性比对象 y 劣以外，其他的属性都比对象 y 的属性优。如果利用经典优势关系，那么对象 x 既不优于对象 y 也不劣于对象 y，此时，就会导致排序失效。而事实上，只要对象 x 的属性有 4 个以上都优于对象 y 的属性，那么我们就可以认为对象 x 优于对象 y 是合理的。

若对象 x 和对象 y 的属性个数都为 m，且对象 x 优于对象 y 的属性的个数为 N，对象 x 劣于对象 y 的属性的个数为 M，则 $N+M=m$。若对象 x 优于对象 y，有 $N>M$，则 $N>\frac{m}{2}$，又因为对象属性的个数为整数，所以 N 的取值范围为 $[m/2+1,m]$，其中/为取整。

通过以上分析，可比较属性的个数 $N\in[m/2+1,m]$。

4.3 基于改进的可能度优势关系排序方法

由于可比较属性的个数 N 不同，得到的∂-可能度优势关系 $R_B^{\geqslant\partial}$ 不同，∂-可能度优势类$[x_i]_B^{\geqslant\partial}$ 也不同。因此，可能度优势关系和可能度优势类需要重新定义。

定义 5 设区间值信息系统 $S=(U,A,V,f)$，$B\subseteq A$，当满足可能度$\partial\in[0.5,1]$的可比较属性的个数为 N 时，可定义可能度优势关系 $R_{B_N}^{\geqslant 0.5}$ 和可能度优势类$[x_i]_{B_N}^{\geqslant 0.5}$为：

$$R_{B_N}^{\geqslant 0.5}=\{(x_j,x_i)\in U\times U \mid \partial_{ji}^a\geqslant 0.5,\forall a\in B\}$$

$$[x_i]_{B_N}^{\geqslant 0.5}=\{x_j\in U \mid (x_j,x_i)\in R_{B_N}^{\geqslant 0.5}\}$$

性质 1 若 $B\subseteq A$，可能度优势关系 $R_{B_N}^{\geqslant 0.5}$ 和可能度优势类$[x_i]_{B_N}^{\geqslant 0.5}$的性质如下：

(1)若 $N_1\leqslant N_2$，则$|[x_i]_{B_{N_1}}^{\geqslant 0.5}|\geqslant|[x_i]_{B_{N_2}}^{\geqslant 0.5}|$；

(2)若 $B_{N_1}\subseteq B_{N_2}$，则 $R_{B_{N_1}}^{\geqslant 0.5}\supseteq R_{B_{N_2}}^{\geqslant 0.5}$；

(3)若 $B_{N_1}\subseteq B_{N_2}$，则$[x_i]_{B_{N_1}}^{\geqslant 0.5}\supseteq[x_i]_{B_{N_2}}^{\geqslant 0.5}$。

证明：(1)当可比较属性的个数 N 越小时，满足可能度$\partial\in[0.5,1]$的类就越多，则可能度优势类的个数就越多；相反，当可比较属性的个数 N 越大时，满足可能度$\partial\in[0.5,1]$的类就越少，则可能度优势类的个数就越少。因此，此性质得证。

(2)若 $B_{N_1}\subseteq B_{N_2}$，则可比较属性的个数 $N_1\leqslant N_2$，可比较属性的个数 N 越少，满足可能度$\partial\in[0.5,1]$的类就越多，即可能度优势关系的集合就越多，反之，则可能度优势关系的集合就越少，从而得到 $R_{B_{N_1}}^{\geqslant 0.5}\supseteq R_{B_{N_2}}^{\geqslant 0.5}$。因此，此性质得证。

(3)若 $B_{N_1}\subseteq B_{N_2}$，则可比较属性的个数 $N_1\leqslant N_2$ 且 $R_{B_{N_1}}^{\geqslant 0.5}\supseteq R_{B_{N_2}}^{\geqslant 0.5}$，再由定义 5 得，$[x_i]_{B_{N_1}}^{\geqslant 0.5}\supseteq[x_i]_{B_{N_2}}^{\geqslant 0.5}$。因此，此性质得证。

本文在文献[18]的基础上，结合改进的可能度优势关系和优势类，给出新的优势度、综合优势度的定义。

定义 6 设区间值信息系统 $S=(U,A,V,f)$，$B\subseteq A$，对象 x_i,x_j 关于 $R_{B_N}^{\geqslant 0.5}$ 的优势度为：

$$R_{B_N}(x_i,x_j)=\begin{cases}\dfrac{|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|}{|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|+|[x_i]_{B_N}^{\geqslant 0.5}-[x_j]_{B_N}^{\geqslant 0.5}|}, & [x_j]_{B_N}^{\geqslant 0.5}\neq[x_i]_{B_N}^{\geqslant 0.5}\\ 0.5, & [x_j]_{B_N}^{\geqslant 0.5}=[x_i]_{B_N}^{\geqslant 0.5}\end{cases}$$

其中，$N\in[m/2+1,m]$，$[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}$ 表示从集合$[x_j]_{B_N}^{\geqslant 0.5}$中除去同时存在于$[x_i]_{B_N}^{\geqslant 0.5}$中的元素的集合，表示单独优于 x_j 而不优于 x_i 的对象集，即可表达 x_i 优于 x_j 的部分，同理$[x_i]_{B_N}^{\geqslant 0.5}-[x_j]_{B_N}^{\geqslant 0.5}$表示对象 x_j 优于对象 x_i 的部分。$|\cdot|$表示集合的基数。

性质 2 $x_i\in U$，$n=U$，$B\subseteq A$，$R_{B_N}(x_i,x_j)$的性质如下：

(1)$R_{B_N}(x_i,x_j)\in[0,1]$；

(2)$R_{B_N}(x_i,x_i)=0.5$；

(3)$R_{B_N}(x_i,x_j)+R_{B_N}(x_j,x_i)=1$；

(4)$\sum_{i=1}^{n}\sum_{j=1}^{n}R_{B_N}(x_i,x_j)=\frac{n^2}{2}$。

证明：(1)当$[x_j]_{B_N}^{\geqslant 0.5}\subseteq[x_i]_{B_N}^{\geqslant 0.5}$，则$|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|=0$，即 $R_{B_N}(x_i,x_j)=0$，此时 $R_{B_N}(x_i,x_j)$最小。

当$[x_j]_{B_N}^{\geqslant 0.5}\supseteq[x_i]_{B_N}^{\geqslant 0.5}$，则 $R_{B_N}(x_i,x_j)=1$，此时 $R_{B_N}(x_i,x_j)$最大。

综上所述，$R_{B_N}(x_i,x_j)\in[0,1]$成立。

(2)由定义 6 知，当$[x_j]_{B_N}^{\geqslant 0.5}=[x_i]_{B_N}^{\geqslant 0.5}$时，$R_{B_N}(x_i,x_j)=0.5$。因为$[x_i]_{B_N}^{\geqslant 0.5}=[x_i]_{B_N}^{\geqslant 0.5}$，所以 $R_{B_N}(x_i,x_i)=0.5$。因此，此性质得证。

(3)由定义 6 知，当$[x_j]_{B_N}^{\geqslant 0.5}=[x_i]_{B_N}^{\geqslant 0.5}$时，有 $R_{B_N}(x_i,x_j)=0.5$，$R_{B_N}(x_j,x_i)=0.5$。显然，$R_{B_N}(x_i,x_j)+R_{B_N}(x_j,x_i)=1$。当$[x_j]_{B_N}^{\geqslant 0.5}\neq[x_i]_{B_N}^{\geqslant 0.5}$时，有

$$R_{B_N}(x_i,x_j)=\frac{|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|}{|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|+|[x_i]_{B_N}^{\geqslant 0.5}-[x_j]_{B_N}^{\geqslant 0.5}|}$$

$$R_{B_N}(x_j,x_i)=\frac{|[x_i]_{B_N}^{\geqslant 0.5}-[x_j]_{B_N}^{\geqslant 0.5}|}{|[x_i]_{B_N}^{\geqslant 0.5}-[x_j]_{B_N}^{\geqslant 0.5}|+|[x_j]_{B_N}^{\geqslant 0.5}-[x_i]_{B_N}^{\geqslant 0.5}|}$$

显然，$R_{B_N}(x_i,x_j)+R_{B_N}(x_j,x_i)=1$。

综上所述，$R_{B_N}(x_i,x_j)+R_{B_N}(x_j,x_i)=1$ 成立。

(4)由性质 $R_{B_N}(x_i,x_i)=0.5$ 和 $R_{B_N}(x_i,x_j)+R_{B_N}(x_j,x_i)=1$ 可得

$$\sum_{i=1}^{n}\sum_{j=1}^{n}R_{B_N}(x_i,x_j)=0.5\times n+1+2+\cdots+n-1=\frac{n^2}{2}$$

因此，此性质得证。

定义 7 设区间值信息系统 $S=(U,A,V,f)$，$B\subseteq A$，x_i 关于 $R_{B_N}^{\geqslant 0.5}$ 的综合优势度为：

$$R_{B_N}(x_i)=\frac{1}{n}\sum_{j=1}^{n}R_{B_N}(x_i,x_j)$$

性质 3 $x_i\in U$，$n=U$，$B\subseteq A$，$R_{B_N}(x_i)$的性质如下：

(1)$R_{B_N}(x_i)\in[\frac{1}{2n},1-\frac{1}{2n}]$；

(2) $\sum_{i=1}^{n} R_{B_N}(x_i)=\frac{n}{2}$。

证明:(1)对象 x_i 最差的情况是除了 $R_{B_N}(x_i,x_i)=0.5$ 之外其余 $R_{B_N}(x_i,x_j)=0$,有

$$R_{B_N}(x_i)=\frac{0+\cdots+0.5+\cdots+0}{n}=\frac{1}{2n}$$

对象 x_i 最好的情况是除了 $R_{B_N}(x_i,x_i)=0.5$ 之外其余 $R_{B_N}(x_i,x_j)=1$,有

$$R_{B_N}(x_i)=\frac{1+\cdots+0.5+\cdots+1}{n}=\frac{n-1+0.5}{n}=1-\frac{1}{2n}$$

综上所述,$R_{B_N}(x_i)\in[\frac{1}{2n},1-\frac{1}{2n}]$成立。

(2)由定义7得

$$\begin{aligned}\sum_{i=1}^{n} R_{B_N}(x_i)&=\sum_{i=1}^{n}\frac{1}{n}\sum_{j=1}^{n} R_{B_N}(x_i,x_j)\\&=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} R_{B_N}(x_i,x_j)\\&=\frac{1}{n}\times\frac{n^2}{2}\\&=\frac{n}{2}\end{aligned}$$

因此,此性质得证。

因为满足可能度$\partial\in[0.5,1]$的可比较属性的个数 N 不同,得到的综合优势度$R_{B_N}(x_i)$也不同,因此,引入平均综合优势度的定义,它是通过计算所有得到的综合优势度 $R_{B_N}(x_i)$的平均值来求得的。

定义8 设区间值信息系统 $S=(U,A,V,f)$,$B\subseteq A$,x_i 关于 $R_{B_N}^{\geqslant 0.5}$ 的平均综合优势度为:

$$R_B(x_i)=\frac{1}{(m-1)/2+1}\sum_{N=m/2+1}^{m} R_{B_N}(x_i)$$

平均综合优势度 $R_B(x_i)$表示对象 x_i 优于另一个对象的程度,平均综合优势度 $R_B(x_i)$越大,表示对象 x_i 越优。

性质4 $x_i\in U$,$n=U$,$B\subseteq A$,$R_B(x_i)$的性质如下:

(1) $R_B(x_i)\in[\frac{1}{2n},1-\frac{1}{2n}]$;

(2) $\sum_{i=1}^{n} R_B(x_i)=\frac{n}{2}$。

证明:(1)由性质 $R_{B_N}(x_i)\in[\frac{1}{2n},1-\frac{1}{2n}]$得,当每一个 N 值都取 $R_{B_N}(x_i)=\frac{1}{2n}$时,有

$$\begin{aligned}R_B(x_i)_{\min}&=\frac{1}{(m-1)/2+1}\sum_{N=m/2+1}^{m} R_{B_N}(x_i)\\&=\frac{1}{(m-1)/2+1}\times((m-1)/2+1)\times\frac{1}{2n}\\&=\frac{1}{2n}\end{aligned}$$

当每一个 N 值都取 $R_{B_N}(x_i)=1-\frac{1}{2n}$时,有

$$\begin{aligned}R_B(x_i)_{\max}&=\frac{1}{(m-1)/2+1}\sum_{N=m/2+1}^{m} R_{B_N}(x_i)\\&=\frac{1}{(m-1)/2+1}\times((m-1)/2+1)\times(1-\frac{1}{2n})\\&=1-\frac{1}{2n}\end{aligned}$$

综上所述,$R_B(x_i)\in[\frac{1}{2n},1-\frac{1}{2n}]$成立。

(2)由性质 $\sum_{i=1}^{n} R_{B_N}(x_i)=\frac{n}{2}$得,

$$\begin{aligned}\sum_{i=1}^{n} R_B(x_i)&=\sum_{i=1}^{n}\frac{1}{(m-1)/2+1}\sum_{N=m/2+1}^{m} R_{B_N}(x_i)\\&=\frac{1}{(m-1)/2+1}\sum_{i=1}^{n}\sum_{N=m/2+1}^{m} R_{B_N}(x_i)\\&=\frac{1}{(m-1)/2+1}\sum_{N=m/2+1}^{m}\sum_{i=1}^{n} R_{B_N}(x_i)\\&=\frac{1}{(m-1)/2+1}\times((m-1)/2+1)\times\frac{n}{2}\\&=\frac{n}{2}\end{aligned}$$

因此,$\sum_{i=1}^{n} R_B(x_i)=\frac{n}{2}$成立。

5 算例分析

某学院在2011年学期末由全体学生对教师的教学水平进行了评估,共有6个评价指标,即 a_1 为讲课流利;a_2 为分析到位;a_3 为重点突出;a_4 为启发思维;a_5 为课堂氛围;a_6 为课件水平。表1为全院学生对其中7个老师给出的一个评估表 $S=(U,A,V,f)$,其中 $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7\}$为论域,$x_1,x_2,x_3,x_4,x_5,x_6,x_7$ 分别表示被评估的7个教师,$A=\{a_1,a_2,a_3,a_4,a_5,a_6\}$为属性集即评价指标。最终给出教师在各项指标下的得分区间。

表1 不完备的教师评估表

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	[90,95]	[89,*]	[85,95]	[80,*]	[94,97]	[80,94]
x_2	[80,91]	[*,86]	[81,84]	[81,84]	[86,89]	[82,92]
x_3	[86,95]	[91,*]	[91,95]	[90,94]	[92,*]	[91,*]
x_4	[89,95]	[92,*]	[91,95]	[87,90]	[91,95]	[88,92]
x_5	[90,97]	[85,*]	[87,97]	[83,93]	[94,95]	[87,92]
x_6	[86,89]	[90,93]	[85,89]	[83,89]	[90,95]	[85,91]
x_7	[84,89]	[84,85]	[82,87]	[86,89]	[88,92]	[90,91]

根据表1数据,具体排序步骤如下:

首先,将不完备的教师评估表转化为完备的教师评估表,得到表2。

表2 完备的教师评估表

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	[90,95]	[89,93]	[85,95]	[80,94]	[94,97]	[80,94]
x_2	[80,91]	[84,86]	[81,84]	[81,84]	[86,89]	[82,92]
x_3	[86,95]	[91,93]	[91,95]	[90,94]	[92,97]	[91,94]
x_4	[89,95]	[92,93]	[91,95]	[87,90]	[91,95]	[88,92]
x_5	[90,97]	[85,93]	[87,97]	[83,93]	[94,95]	[87,92]
x_6	[86,89]	[90,93]	[85,89]	[83,89]	[90,95]	[85,91]
x_7	[84,89]	[84,85]	[82,87]	[86,89]	[88,92]	[90,91]

其次,根据完备的教师评估表,算出所有对象的可能度 ∂_{ji}^{a}。

再次,根据各个对象都有6个属性,即 $m=6$,得到可比较属性的个数 $N\in[4,6]$,即 $N=4,5,6$。

然后,根据不同 N 值,给出不同的可能度优势类。

当 $N=4$ 时,各个对象的可能度优势类为:

$[x_1]_{B_4}^{\geqslant 0.5}=\{x_1,x_3,x_4,x_5\}$

$[x_2]_{B_4}^{\geqslant 0.5}=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7\}$

$[x_3]_{B_4}^{\geqslant 0.5}=\{x_3\}$

$[x_4]_{B_4}^{\geqslant 0.5}=\{x_3,x_4\}$

$[x_5]_{B_4}^{\geqslant 0.5}=\{x_3,x_4,x_5\}$

$[x_6]_{B_4}^{\geqslant 0.5}=\{x_1,x_3,x_4,x_5,x_6\}$

$[x_7]_{B_4}^{\geqslant 0.5}=\{x_1,x_3,x_4,x_5,x_6,x_7\}$

根据各对象的可能度优势类，得到关于 $R_{B_4}^{\geqslant 0.5}$ 的优势度矩阵为：

$$R_{B_4}(x_i,x_j)=\begin{bmatrix} \frac{1}{2} & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & \frac{1}{2} & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & \frac{1}{2} & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

从而，得到各对象的综合优势度为：$R_{B_4}(x_1)=\frac{1}{2}$，$R_{B_4}(x_2)=\frac{1}{14}$，$R_{B_4}(x_3)=\frac{13}{14}$，$R_{B_4}(x_4)=\frac{11}{14}$，$R_{B_4}(x_5)=\frac{9}{14}$，$R_{B_4}(x_6)=\frac{5}{14}$，$R_{B_4}(x_7)=\frac{3}{14}$。

同理可得，当 $N=5$ 时，各对象的综合优势度为：$R_{B_5}(x_1)=\frac{143}{210}$，$R_{B_5}(x_2)=\frac{1}{14}$，$R_{B_5}(x_3)=\frac{11}{14}$，$R_{B_5}(x_4)=\frac{31}{42}$，$R_{B_5}(x_5)=\frac{25}{42}$，$R_{B_5}(x_6)=\frac{11}{35}$，$R_{B_5}(x_7)=\frac{11}{35}$。

当 $N=6$ 时，各对象的综合优势度为：$R_{B_6}(x_1)=\frac{53}{84}$，$R_{B_6}(x_2)=\frac{2}{21}$，$R_{B_6}(x_3)=\frac{5}{7}$，$R_{B_6}(x_4)=\frac{2}{3}$，$R_{B_6}(x_5)=\frac{53}{84}$，$R_{B_6}(x_6)=\frac{1}{3}$，$R_{B_6}(x_7)=\frac{3}{7}$。

最后，得到各对象的平均综合优势度为：$R_B(x_1)=0.604$，$R_B(x_2)=0.079$，$R_B(x_3)=0.810$，$R_B(x_4)=0.730$，$R_B(x_5)=0.623$，$R_B(x_6)=0.335$，$R_B(x_7)=0.319$。

最后得出排序的结果为：

$x_3\succ x_4\succ x_5\succ x_1\succ x_6\succ x_7\succ x_2$

而在文献[12]中，利用基于 ∝-β 优势关系的区间数多属性决策模型得到的排序结果为：

$x_3\succ x_4\succ x_5\succ x_1\sim x_6\succ x_7\sim x_2$

通过排序结果可以看出，本文提出的基于改进的可能度优势关系排序方法能将对象 x_1 和 x_6、对象 x_7 和 x_2 区分开，而基于 ∝-β 优势关系的区间数多属性决策模型却不能。因此，本文提出的排序方法具有较高的区分度，排序更高效，结果更合理。

表 3　与其他排序方法的结果比较

算例	其他排序方法及排序结果	本文排序方法的排序结果
文献[8]中的算例（前 5 个对象）	基于熵权和优势关系相结合的方法 $x_1\succ x_3\succ x_5\succ x_4\succ x_2$	$x_1\succ x_4\succ x_3\succ x_2\succ x_5$
文献[19]中的算例（前 5 个对象）	改进的熵权-层次分析法 $x_1\succ x_3\succ x_5\succ x_4\succ x_2$	$x_4\succ x_3\succ x_1\succ x_2\succ x_5$
文献[7]中的算例（前 6 个对象）	基于容差优势关系的排序方法 $x_3\succ x_1\sim x_6\succ x_4\succ x_5\succ x_2$	$x_3\succ x_6\succ x_4\succ x_5\succ x_1\succ x_2$

考虑到仅将其他不同的排序方法应用到本文算例中不具有较强的说服力，为了验证本文提出的基于改进的可能度优势关系排序方法具有更高效、更客观、更合理、更高的区分度和应用更广泛的特点，我们将此排序方法与其他排序方法在不同算例中进行比较，比较结果如表 3 所列。

结果分析：

本文提出的基于改进的可能度优势关系排序方法与文献[8]中的基于熵权和优势关系相结合的方法相比较，虽然后者能够将主观和客观有效地相结合，也能够在一定程度上有效地避免属性权重带来的主观性，但是它仍然存在一定的主观因素和优势关系太过严格的问题。然而，本文提出的排序方法能够解决其缺陷，从而使排序结果更客观合理。

与文献[19]中的基于改进的熵权-层次分析法相比较，虽然改进的熵权-层次分析法在一定程度上也做到了主观和客观的有效结合，但是它仍然存在一定的主观性。然而，本文提出的排序方法能够有效地避免主观性的存在，从而使排序结果具有较强的可信度和合理性。

与文献[7]中的基于容差优势关系的排序方法相比较，虽然容差优势关系避免了优势关系太过严格的不足，但是从结果来看，它并没有有效地把对象 x_1 和 x_6 区分开，并且容差率 β 的确定也存在一定的主观性。然而，本文提出的排序方法不仅能够有效地避免主观性的存在，而且能把对象 x_1 和 x_6 区分开，具有更高的区分度。

通过与以上不同的排序方法在不同算例中的对比分析，可以发现本文提出的排序方法具有以下显著特点：

从排序结果上看，本文提出的基于改进的可能度优势关系的排序方法能够更好地将不同对象区分开。因此，本文提出的排序方法具有较高的区分度。

从排序方法上看，本文方法不仅能有效地避免主观性的存在，而且能解决排序失效问题；从排序合理性上看，本文方法能使排序结果更客观合理；从应用上看，本文方法不仅能够解决不完备区间信息系统的问题，而且也能解决完备区间和单值信息系统的问题，因此，此排序方法具有更广泛的应用；从操作上看，本文方法操作简单、计算量小。

结束语　本文针对不完备区间值信息系统中存在的问题，在优势关系和可能度的基础上，给出了平均综合优势度的定义，从而提出基于改进的可能度优势关系的排序方法，并通过具体算例验证了该方法的可行性，又通过与其他排序方法进行比较，证明了本文提出的排序方法更高效、更客观、区分度更高、应用更广泛。

本文着重讨论了不同对象的排序问题，对于不完备信息系统将不完备信息转化为完备的信息进行计算，如何直接应用不完备信息系统将是下一步研究的重点问题。

参 考 文 献

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

[2] Greco S，Matarazzo B，Slowinski R. A new rough set approach to multicriteria and multiattribute classification[C]// Proceedings of the First International Conference on Rough Sets and Current Trends in Computing，1998. Berlin：Springer Berlin Heidelberg，1998：60-67

（下转第 292 页）

法[J]. 控制理论与应用，2004，21(2)：165-168
Fang Min，Wang Bao-shu. Multi-sensor radar radiation source target recognition based on evolution strategy method[J]. Control Theory and Application，2004，21(2)：165-168
[14] 李鸿儒，王晓楠，高仝. 基于免疫进化策略的神经网络优化方法[J]. 东北大学学报(自然科学版)，2008，29(6)：794-797
Li Hong-ru，Wang Xiao-nan，Gao Tong. Optimization Algorithm Based on Immune Evolutionary Strategy of Neural Network[J]. Journal of Northeast University(Natural Science)，2008，29(6)：794-797
[15] Bengio Y. Learning Deep Architectures for AI[J]. Foundations and Trends in Machine Learning，2009，2(1)：1-127
[16] Deep Learning 和 Knowledge Graph 引爆大数据革命[OL]. http://blog.sina.com.cn/s/blog_46d0a3930101fswl.html
[17] Boltzmann 神经网络模型与学习算法 [OL]. http://wenku.baidu.com/view/490dcf748e9951e79b891785.html
[18] 浅谈 Deep Learning 的基本思想和方法[OL]. http://blog.csdn.net.xianlingmao/article/details
[19] Deep Learning[OL]. http://www.cs.nyu.edu/yann/ research/deep/

(上接第 278 页)

[3] Greco S，Matarazzo B，Slowinski R. Rough sets theory for multicriteria decision analysis[J]. European Journal of Operational Research，2001，129(1)：1-47
[4] 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
[5] 苟光磊，王国胤，利节，等. 基于置信优势关系的粗糙集近似模型[J]. 控制与决策，2014，29(7)：1325-1329
Gou Guang-lei，Wang Guo-yin，Li jie，et al. Confidential dominance relation based rough approximation model[J]. Control and Decision，2014，29(7)：1325-1329
[6] Liu Jian，Xue Li，Liu Si-feng，et al. Research on multiple-attribute decision making problems based on the superiority index[J]. Control and Decision，2010，25(7)：1079-1087
[7] 陈万翠，吕跃进，翁世洲. 基于容差优势关系的排序方法及其应用[J]. 计算机应用，2014，34(8)：2170-2174
Chen Wan-cui，Lv Yue-jin，Weng Shi-zhou. Sorting method and its application based on tolerance dominance relation[J]. Journal of Computer Applications，2014，34(8)：2170-2174
[8] 王利东，田晓娟，杨艳冰. 基于熵权与优势关系的教学效果评价方案[J]. 数学的实践与认识，2014，44(10)：8-12
Wang Li-dong，Tian Xiao-juan，Yang Yan-bing. The comprehensive evaluation of teaching based on entropy weight and dominance relation[J]. Journal of Mathematics in Practice and Theory，2014，44(10)：8-12
[9] 李金鹏，岳超源，李武. 一类基于优势关系的不完全信息多属性决策方法[J]. 控制与决策，2013，28(2)：229-234
Li Jin-peng，Yue Chao-yuan，Li Wu. A dominance relation-based decision making approach for multi-attribute decision making problems with incomplete information[J]. Control and Decision，2013，28(2)：229-234
[10] 邱涤珊，贺川，朱晓敏. 基于概率可信度的区间数排序方法[J]. 控制与决策，2012，27(12)：1894-1898
Qiu Di-shan，He Chuan，Zhu Xiao-min. Ranking method research of interval numbers based on probability[J]. Control and Decision，2012，27(12)：1894-1898
[11] 徐智明，刘宏，马琳，等. 基于可能度优势关系的区间序粗糙集模型[J]. 舰船电子工程，2012，32(9)：40-42
Xu Zhi-ming，Liu Hong，Ma Lin，et al. Interval ordered rough set model based on possible degree dominance relation[J]. Ship Electronic Engineering，2012，32(9)：40-42
[12] Hu Ming-li，Li Lin-li. A novel dominance relation and application in interval grey number decision model[J]. Journal of Grey System，2014，26(1)：91-98
[13] Qian Yu-hua，Liang Ji-ye，Dang Chuang-yin. Interval ordered information systems[J]. Computers and Mathematics with Applications，2008，56(8)：1994-2009
[14] 王晓妍. 不完备信息系统中优势关系粗糙集决策方法研究[D]. 哈尔滨：哈尔滨理工大学，2014
Wang Xiao-yan. Study on decision-making approaches based on dominance relation rough sets in incomplete information system[D]. Harbin：Harbin University of Science and Technology，2014
[15] 王斌，邵明文，王金鹤，等. 基于改进的优势关系下的不完备区间值信息系统评估模型[J]. 计算机科学，2014，41(2)：253-256
Wang Bin，Shao Ming-wen，Wang Jin-he，et al. New evaluation model for incomplete interval-valued information system based on improved dominance relations[J]. Computer Science，2014，41(2)：253-256
[16] 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(理学版)，2010，45(9)：7-13
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
[17] 刘学生. 基于粗集的不确定多属性决策排序法的研究[D]. 大连：大连理工大学，2009
Liu Xue-sheng. Research on uncertain multiple attribute decision ranking based on rough sets[D]. Dalian：Dalian University of Technology，2009
[18] 翁世洲，吕跃进，莫京兰. 基于优势关系的排序模型及其保序性约简理论[J]. 广西师范大学学报(自然科学版)，2013，31(3)：37-44
Weng Shi-zhou，Lv Yue-jin，Mo Jing-lan. Ranking model and preserving reduction based on dominance relation[J]. Journal of Guangxi Normal University(Natural Science Edition)，2013，31(3)：37-44
[19] 郭金维，蒲绪强，高祥，等. 一种改进的多目标决策指标权重计算方法[J]. 西安电子科技大学学报，2014，41(6)：118-125
Guo Jin-wei，Pu Xu-qiang，Gao Xiang，et al. Improved method on weights determination of indexes in multi-objective decision[J]. Journal of Xidian University，2014，41(6)：118-125