

DOI:10.3969/j.issn.1673-4785.201312008
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150317.1025.003.html>

采用粒计算的属性权重确定方法

周丹晨

(中国工程物理研究院 机械制造工艺研究所, 四川 绵阳 621900)

摘 要:针对属性值为连续值且无单一决策属性的多属性决策问题,通过宏观和微观的粒计算理论模型的对比分析,提出了一种融合模糊商空间理论和粗糙集理论的属性权重确定方法。首先通过应用模糊商空间理论构建具有分层递阶结构的商空间族,将一系列粒度商空间的样本聚类结果作为相应粒度空间下单一决策属性的分类,然后采用粗糙集理论计算得出所有商空间下各个属性的重要度,综合不同粗细粒度商空间下的属性重要度大小确定了各属性的客观权重。应用实验验证了该方法的合理性、有效性和实用性。

关键词:属性权重;粒计算;模糊商空间;粗糙集;多属性决策

中图分类号:TP18 **文献标志码:**A **文章编号:**1673-4785(2015)02-023-08

中文引用格式:周丹晨. 采用粒计算的属性权重确定方法[J]. 智能系统学报, 2015, 10(2): 273-280.

英文引用格式:ZHOU Danchen. A method for ascertaining the weight of attributes based on granular computing[J]. CAAI Transactions on Intelligent Systems, 2015, 10(2): 273-280.

A method for ascertaining the weight of attributes based on granular computing

ZHOU Danchen

(Institute of Machinery Manufacturing Technology, China Academy of Engineering Physics, Mianyang 621900, China)

Abstract: Aiming at multi-attribute decision making problem without single decision attribute in which the attribute values are in the forms of continuous numbers, a method for ascertaining attributes weight based on fusion of fuzzy quotient spaces theory and rough sets theory is put forward through comparison and analysis of the macrocosmic and microcosmic theoretical models of granular computing. Firstly, the quotient space family with hierarchical structure was established by applying fuzzy quotient spaces theory. The sample clustering resulted in a series of granular quotient spaces were used as classification of single decision attribute in corresponding granular space. Furthermore, the significances of every attribute in all quotient spaces were calculated by applying rough sets theory. Finally, objective attributes weight was ascertained by integrating the significances of every attribute in different granular quotient spaces. The application case showed that the proposed method is reasonable, effective and practical.

Keywords: attributes weight; granular computing; fuzzy quotient space; rough set; multi-attribute decision making

科学合理地确定属性权重关系是多属性决策领域研究的核心问题,关系到决策结果的可靠性与正确性^[1]。属性权重应能从宏观上体现决策目标,有效地反映各属性对目标的影响程度,体现决策者的

评价目的和决策者对不同属性的重视程度,能较好地遵循重要性原则。

确定属性权重的方法有主观赋权法、客观赋权法和主客观相结合的赋权方法。客观赋权法基于决策矩阵信息,从问题数据入手,通过建立一定的数学模型计算出权重系数,从而减少了决策者的主观因素,具有较强的科学性。针对采用连续值、区间数、模糊数、序数等数值形式定量表达的属性值,或者用

收稿日期:2013-12-09. 网络出版日期:2015-03-17.

基金项目:中国工程物理研究院科学技术发展基金资助项目(2013B0203031).

通信作者:周丹晨. E-mail: zdc69@163.com.

语言形式定性描述的属性值,目前主要有主成分分析法^[2]、离差最大化方法^[3-4]、均方差法^[5]、多目标规划法^[6]、信息熵法^[7-8]、聚类分析法^[9]、粗糙集方法^[10]、灰色关联分析方法^[11]等客观赋权法。

对于没有任何先验信息,属性值为连续值且无单一决策属性的多属性决策问题,一些文献提出了基于模糊聚类和粗糙集的客观赋权法^[12-15]。其核心是根据样本集的属性进行模糊聚类,以没有删除任何属性的分类为基准,把没有删除任何一个属性视为一种知识分类,将删除各个属性后视为另一种知识分类,计算删除各属性后的分类相对于总的属性分类的正域,从而求得各个属性的重要程度,最后利用归一化方法来确定各属性的权重。该方法没有客观信息的损失,具有一定的普遍性,但该方法存在以下不足:即如果样本集有 m 个属性,则需要进行 $m+1$ 次模糊聚类计算(其中包括计算量很大的模糊相似度计算以及求出模糊等价矩阵的传递闭包计算),因此当属性数量和样本数量较多时,该方法的计算过程比较繁琐,一定程度上影响了其实用性。

粒计算理论是目前人工智能研究领域的新热点,其理论模型主要分为2类:1)以处理不确定性为主要目标,如以模糊集(也称词计算)理论和粗糙集理论为基础的模型;2)以多粒度计算为目标,如商空间理论^[16-18]。这2类模型的侧重点有所不同,前者侧重于计算对象的不确定性处理,因此重点关注于含糊、不清晰概念的表达和近似推理;而后者核心思想来源于人类问题求解的基本特征之一,即从极不相同的粒度空间上观察和分析同一问题,并很容易地从一个抽象的粒度空间转换到其他的粒度空间,也就是分层次地处理问题的能力,因此更关注于如何从不同粗细粒度的问题空间,即从不同的层次得到对问题不同角度的理解,并最终综合成对问题总的理解。

本文基于2类粒计算理论模型的思想,同样针对属性值为连续值且无单一决策属性的多属性决策问题,提出了一种融合模糊商空间理论和粗糙集理论的客观赋权法,以期在现有研究的基础上,在保证权重确定过程合理性的前提下,降低计算复杂度,进一步提高计算结果的准确性和方法的实用性。

1 粒计算理论基础

1.1 模糊商空间理论

模糊商空间理论是将模糊集合论引入经典商空间理论,利用模糊等价关系将精确粒度下的商空间理论推广到模糊粒度计算中,且所有模糊商空间全

体构成一个完备半序格,从而为粒计算提供了一种有力的数学模型和工具^[19-20]。模糊商空间理论能够更好地反映人类处理不确定问题的若干特点,即信息的确定与不确定、概念的清晰与模糊都是相对的,都与问题的粒度粗细有关。

定义1^[19] 设 X 是论域, X 上的一个模糊集 A 是指 $\forall x \in X$,有一个指定的数 $\mu_A \in [0,1]$,称为 x 对 A 的隶属程度,映射: $\mu_A: X \rightarrow [0,1], x \rightarrow \mu_A(x)$ 称为 A 的隶属函数。令 $T(X)$ 表示 X 上一切模糊子集的集合,则 $T(X)$ 实际上是由 $\mu: X \rightarrow [0,1]$ 这个函数组成的函数空间。

定义2^[19] 设 $R \in T(X \times X)$,若满足

- 1) $\forall x \in X, R(x, x) = 1$,
- 2) $\forall x, y \in X, R(x, y) = R(y, x)$,
- 3) $\forall x, y, z \in X, R(x, z) \geq \sup_y (\min(R(x, y), R(y, z)))$,

则称 R 是 X 上的一个模糊等价关系。

命题1^[19] 设 R 是 X 上的一个模糊等价关系,令 $R_\lambda = \{(x, y) \mid R(x, y) \geq \lambda\}, 0 \leq \lambda \leq 1$,则 R_λ 是 X 上的一个普通等价关系,称 R_λ 为 R 的截关系。令等价关系 R_λ 对应的商空间为 $X(\lambda)$,可得到如下性质:若 $0 \leq \lambda_2 \leq \lambda_1 \leq 1 \Leftrightarrow R_{\lambda_1} \supseteq R_{\lambda_2} \Leftrightarrow X(\lambda_2) \subseteq X(\lambda_1)$ 的商集。于是,商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ 按照商集的包含关系构成一个有序链,称 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ 为 X 上的一个分层递阶结构。

命题2^[19] 给定 X 上的一个模糊等价关系,则对应一个 X 上的分层递阶结构。

1.2 粗糙集理论

粗糙集理论是一种处理模糊性和不精确性问题的数学工具^[21-22]。其本质思想是利用等价关系(不可分辨关系)建立论域的一个划分,得到不可区分的等价类,从而构建一个近似空间。在此近似空间上,用精确的上近似集和下近似集来逼近一个边界模糊的集合。

定义3^[21] 给定一个信息系统 $S = \langle U, A, V, f \rangle$,其中 U 为研究对象的非空有限集合,称为论域; A 为属性的非空有限集合, $A = C \cup D$,其中 C 为条件属性集, D 为决策属性集; V 是属性值的集合, $V = \bigcup_{a \in A} V_a$, V_a 是属性 a 的值域; f 为信息函数, $f: U \times A \rightarrow V$,它为每个对象的每个属性赋予信息值。对于每个子集 $X \subseteq U$ 和一个等价关系 $R \in \text{ind}(S)$,称 $R(X) = \bigcup \{Y_i \in U/R \mid Y_i \subseteq X\}$ 为 X 的 R 下近似集。

定义4^[21] 条件属性集 C 与决策属性集 D 之间的依赖程度 $\gamma(C, D)$ 定义为

$$\gamma(C, D) = |\text{POS}_C(D)| / |U| \quad (1)$$

式中: $\text{POS}_C(D)$ 为 C 相对于 D 的正域:
 $\text{POS}_C(D) = \bigcup_{Y \in U/D} C(Y)$, 其表示根据条件属性集 C 能够被准确地分入由决策属性 D 所确定的分类元素的集合。 $|U|$ 表示整个集合对象的个数。

定义 5^[21] 设 a_k 为条件属性集 C 中的一个条件属性(即 $a_k \in C$), 属性 a_k 关于 D 的重要度定义为

$$\text{sig}(a_k, C, D) = \gamma(C, D) - \gamma(C - \{a_k\}, D) \quad (2)$$

式中: $\gamma(C - \{a_k\}, D)$ 表示在 C 中缺少属性 a_k 后, 条件属性集对决策属性集的依赖程度。

2 融合模糊商空间和粗糙集理论的属性权重确定方法

2.1 学术思想

可以看出, 模糊商空间理论和粗糙集理论的共同点是都认为概念粒子可以用子集来表示, 不同粒度下的粒子用不同大小的子集来描述, 所有的粒子都通过等价关系获得划分产生^[17]。但它们之间的区别主要在于: 粗糙集理论的研究对象是由一个多值属性集合描述的对象集合, 但其论域是点集, 各个对象之间没有结构关系或拓扑关系; 而模糊商空间理论是把商集作为描述不同粒度世界的数学模型, 其本质是分层递阶结构, 是一种由粗到细、由表及里来描述样本集的方法, 因此利用模糊商空间理论可以从大量的样本数据中建立一系列具有粒度层次结构的商空间(即对论域进行不同的划分), 然后在各商空间中获取相应的知识, 从而实现从模糊信息粒结构到分层递阶结构再到具有粒度结构的知识的相互转化。从这个意义上来说, 模糊商空间理论可以看成是一种宏观的粒计算模型, 而粗糙集理论可以看成是一种微观的粒计算模型, 即不同粒度均在同一个给定的商空间中进行划分, 是无拓扑结构情况的商空间特例^[17]。

基于以上分析, 本文所提出的融合模糊商空间理论和粗糙集理论的属性权重确定方法的主要思想是: 首先根据样本集的属性值, 通过应用模糊商空间理论得到具有分层递阶结构的商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$, 并把一系列粒度商空间的样本聚类结果作为相应粒度空间下单一决策属性的分类; 然后在各个粒度商空间 $X(\lambda)$ 下, 通过采用粗糙集理论的属性重要度计算方法得出该商空间下各个属性的重要度值, 即从不同粗细粒度的商空间来观察分析各个属性的重要程度, 从而综合得出各属性总的重要度; 最终采用归一化方法来求解各属性的权重。这样通过将宏观和微观的粒计算理论模型

统一起来, 从多粒度、多层次所得到的属性重要度融合信息, 使属性权重的确定过程更为全面客观, 计算结果也相对更为准确。

2.2 流程步骤

该方法的流程如图 1 所示。

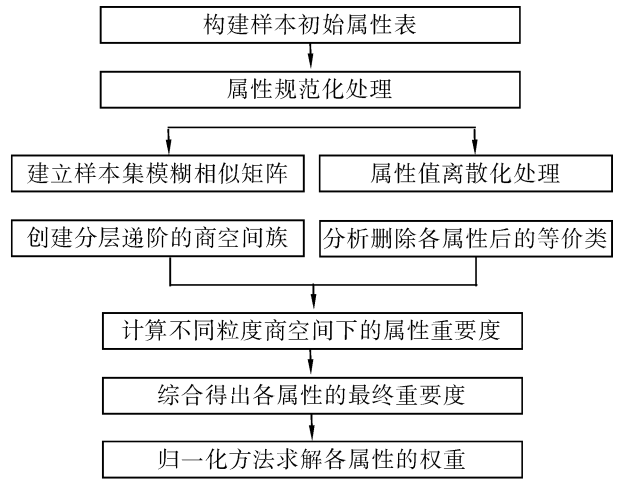


图 1 融合模糊商空间和粗糙集理论的属性权重确定流程
Fig.1 Flow chart of ascertaining attributes weight

其计算步骤如下:

1) 根据需要处理的样本对象属性(指标)项目, 设有待处理的 n 个样本的组成集合

$X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n\}$, 每个样本用 m 个指标特征值向量表示: $\mathbf{x}_i = \{y_{i1}, y_{i2}, \dots, y_{ik}, \dots, y_{im}\}$, 这样就可以构建一个样本初始属性表, 如表 1 所示。

表 1 样本初始属性表
Table 1 Initial attributes of samples

样本	条件属性集					
	a_1	a_2	\cdots	a_k	\cdots	a_m
\mathbf{x}_1	y_{11}	y_{12}	\cdots	y_{1k}	\cdots	y_{1m}
\mathbf{x}_2	y_{21}	y_{22}	\cdots	y_{2k}	\cdots	y_{2m}
\vdots	\vdots	\vdots	\cdots	\vdots	\cdots	\vdots
\mathbf{x}_i	y_{i1}	y_{i2}	\cdots	y_{ik}	\cdots	y_{im}
\vdots	\vdots	\vdots	\cdots	\vdots	\cdots	\vdots
\mathbf{x}_n	y_{n1}	y_{n2}	\cdots	y_{nk}	\cdots	y_{nm}

2) 由于样本初始属性表中各属性的量纲不同、取值范围不同和极性不同, 因此必须首先采用线性比例变换法、极差变换法、比重变换法、向量标准化法等方法, 消除属性量纲和数量级的影响, 并进行极性转换, 将属性值统一规范到固定的区间上, 使得多个属性能够进行比较。

3) 并行的计算步骤, 即一方面基于模糊商空间理论, 为创建具有分层递阶结构的商空间族, 需采用数量积法、相关系数法、夹角余弦法、最大最小法、算

术平均最小法等方法根据规范化处理后的属性值计算样本 x_i 与 x_j 之间的相似关系 $R(x_i, x_j) = \begin{cases} 1, i=j \\ 0 \leq r_{ij} < 1, i \neq j \end{cases}$, 构建所有样本的模糊相似矩阵 $R = [r_{ij}]_{n \times n}$, 由于 $r_{ij} = r_{ji}$, 所以 R 既是自反矩阵也是对称矩阵, 如式 (3) 所示。

$$R = [r_{ij}]_{n \times n} = \begin{bmatrix} 1 & & & & & \\ \vdots & \ddots & & & & \\ r_{i1} & \cdots & 1 & & & \\ \vdots & & \vdots & \ddots & & \\ r_{j1} & \cdots & r_{ji} & \cdots & 1 & \\ \vdots & & \vdots & & \vdots & \ddots \\ r_{n1} & \cdots & r_{ni} & \cdots & r_{nj} & \cdots & 1 \end{bmatrix} \quad (3)$$

同时, 由于粗糙集理论只能分析离散型属性值, 因此需要对连续量属性值进行离散化处理。

4) 该步骤同样为并行的计算步骤, 即一方面需根据得到的样本集模糊相似矩阵创建具有分层递阶结构的商空间族, 另一方面需根据离散化后的样本属性表分析删除各个属性后的等价类 $U/\text{ind}(C - \{a_k\})$ 。

传统的模糊聚类方法在建立模糊相似矩阵 R 后, 需要采用平方法求出 R 的传递闭包: $R \rightarrow R^2 \rightarrow (R^2)^2 \rightarrow \cdots \rightarrow R^{2^k} = \hat{R}$, 将 R 改造成模糊等价矩阵 \hat{R} 后才能得到聚类结果^[23]。文献[24]证明了一个模糊相似矩阵 R 和其模糊等价矩阵 \hat{R} 所分别对应的分层递阶结构是相同的, 从而省去了比较繁琐的传递闭包计算, 同时给出了基于模糊相似矩阵的分层递阶结构聚类算法, 由此就可以得到一个具有商集包含关系的样本商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ 。不同的 λ 取值将分别对应一个不同粒度的商空间, 且 λ 值越大, 商空间的粒度就越细, 样本聚类数也就越多。

5) 以商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ 所形成的一系列粒度商空间的样本聚类结果作为相应粒度空间下单一决策属性的分类, 按照式 (2) 分别计算各粒度商空间下各个属性的重要度。

6) 根据式 (4) 计算得出各属性最终重要度:

$$\text{sig}(a_k, C, D) = \frac{1}{q} \sum_{p=1}^q \text{sig}^{X(\lambda_p)}(a_k, C, D) \quad (4)$$

式中: q 为商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ 的个数, $\text{sig}^{X(\lambda_p)}(a_k, C, D)$ 为商空间 $X(\lambda_p)$ 下属性 a_k 的重要度。

7) 根据各个属性最终重要度的大小, 采用归一

化方法确定各属性的权重, 如式 (5) 所示。

$$w_{a_k} = \frac{\text{sig}(a_k, C, D)}{\sum_{k=1}^m \text{sig}(a_k, C, D)} \quad (5)$$

3 应用算例

将上述提出的属性权重确定方法应用于某多品种小批量制造企业加工中心操作工 2012 年工作绩效综合评价, 共选取 15 个操作人员样本, 评价指标及其属性值如表 2 所示, 其中除加班时间为成本型属性外, 其余属性均为效益型属性。

表 2 某企业加工中心操作工工作绩效综合评价指标
Table 2 The performance evaluation index of machining center operators in an enterprise

操作工 编号	总工时 a_1	生产 工时率 $a_2/\%$	产品 合格率 $a_3/\%$	按期 完成率 $a_4/\%$	加工 难度 a_5	加班 时间 a_6
x_1	3 681.0	99.6	100	98.70	高	294.5
x_2	3 135.6	96.5	99.65	100	较高	509.0
x_3	1 916.5	97.0	99.39	97.01	一般	277.1
x_4	4 098.0	98.7	99.05	96.23	高	601.8
x_5	2 076.9	98.1	99.96	97.92	较高	532.2
x_6	3 937.6	99.0	95.67	95.58	高	462.6
x_7	4 194.3	98.3	98.92	94.67	较高	711.9
x_8	3 456.4	98.4	99.78	98.31	较高	346.7
x_9	4 515.1	99.2	100	100	较高	543.8
x_{10}	2 365.6	100	97.53	100	一般	132.2
x_{11}	2 654.3	96.9	99.74	87.00	一般	259.7
x_{12}	2 975.2	87.0	100	99.61	高	387.3
x_{13}	3 777.2	99.0	100	99.87	较高	474.2
x_{14}	4 354.7	99.5	99.91	94.02	高	636.5
x_{15}	1 306.9	95.3	99.91	98.44	一般	375.7

1) 由于“加工难度”属性为定性指标, 直接按照“高=100, 较高=50, 一般=0”的规则进行处理。对于其余连续型属性值, 采用式 (6) 所示的极差变换法进行属性规范化处理, 将每一个属性值统一于同一数值范围 $[0, 100]$, 规范化后的属性数据表如表 3 所示。

$$\begin{cases} y'_{ik} = \frac{y_{ik} - \min_{i=1}^n(y_{ik})}{\max_{i=1}^n(y_{ik}) - \min_{i=1}^n(y_{ik})} \times 100, a_k \text{ 为效益型属性} \\ y'_{ik} = \frac{\max_{i=1}^n(y_{ik}) - y_{ik}}{\max_{i=1}^n(y_{ik}) - \min_{i=1}^n(y_{ik})} \times 100, a_k \text{ 为成本型属性} \end{cases}$$

表 3 规范化处理后的属性表

样本	a_1	a_2	a_3	a_4	a_5	a_6
x_1	74	97	100	90	100	72
x_2	57	73	92	100	50	35
x_3	19	77	86	77	0	75
x_4	87	90	78	71	100	19
x_5	24	85	99	84	50	31
x_6	82	92	0	66	100	43
x_7	90	87	75	59	50	0
x_8	67	88	95	87	50	63
x_9	100	94	100	100	50	29
x_{10}	33	100	43	100	0	100
x_{11}	42	76	94	0	0	78
x_{12}	52	0	100	97	100	56
x_{13}	77	92	100	99	50	41
x_{14}	95	96	98	54	100	13
x_{15}	0	64	98	88	0	58

2) 采用夹角余弦法计算每一个样本之间的相似关系,其计算公式如式(7)所示,从而建立所有样本的模糊相似矩阵,如式(8)所示。

$$r_{ij} = \cos(x_i, x_j) = \frac{\sum_{k=1}^m y'_{ik} \times y'_{jk}}{\sqrt{\sum_{k=1}^m (y'_{ik})^2 \times \sum_{k=1}^m (y'_{jk})^2}}$$

(7)

3) 根据文献[24]给出的基于模糊相似矩阵的分层递阶结构聚类算法,得到一个有序的样本商空间族 $\{X(\lambda) \mid 0 \leq \lambda \leq 1\}$,不同粒度的商空间分别对应不同的样本聚类结果,如表 4 所示。图 2 所示为与表 4 相对应的根据分层递阶结构绘制的样本聚类结构图。

$$R = [r_{ij}]_{15 \times 15} =$$

1														
0.962	1													
0.855	0.895	1												
0.963	0.931	0.724	1											
0.943	0.975	0.912	0.898	1										
0.875	0.788	0.602	0.905	0.729	1									
0.908	0.924	0.739	0.964	0.882	0.830	1								
0.979	0.983	0.926	0.926	0.960	0.812	0.917	1							
0.945	0.979	0.846	0.950	0.932	0.816	0.975	0.972	1						
0.819	0.837	0.944	0.692	0.823	0.709	0.694	0.883	0.802	1					
0.775	0.735	0.865	0.668	0.764	0.519	0.705	0.829	0.747	0.766	1				
0.888	0.875	0.713	0.837	0.827	0.694	0.739	0.849	0.818	0.626	0.567	1			
0.965	0.994	0.895	0.942	0.964	0.809	0.951	0.990	0.993	0.845	0.777	0.840	1		
0.945	0.910	0.708	0.991	0.887	0.858	0.971	0.911	0.941	0.644	0.710	0.808	0.929	1	
0.813	0.887	0.978	0.684	0.921	0.508	0.696	0.887	0.813	0.882	0.788	0.730	0.868	0.667	1

(8)

表 4 具有分层递阶结构的样本商空间族

Table 4 The family of quotient spaces with hierarchical structure

商空间族	样本聚类结果	聚类数
$X(1)$	$\{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{13}\}, \{x_{14}\}, \{x_{15}\}\}$	15
$X(0.994)$	$\{\{x_1\}, \{x_2, x_{13}\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{14}\}, \{x_{15}\}\}$	14
$X(0.993)$	$\{\{x_1\}, \{x_2, x_9, x_{13}\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{14}\}, \{x_{15}\}\}$	13
$X(0.991)$	$\{\{x_1\}, \{x_2, x_9, x_{13}\}, \{x_3\}, \{x_4, x_{14}\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{15}\}\}$	12
$X(0.990)$	$\{\{x_1\}, \{x_2, x_8, x_9, x_{13}\}, \{x_3\}, \{x_4, x_{14}\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{15}\}\}$	11
$X(0.979)$	$\{\{x_1, x_2, x_8, x_9, x_{13}\}, \{x_3\}, \{x_4, x_{14}\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{15}\}\}$	10
$X(0.978)$	$\{\{x_1, x_2, x_8, x_9, x_{13}\}, \{x_3, x_{15}\}, \{x_4, x_{14}\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}\}$	9
$X(0.975)$	$\{\{x_1, x_2, x_5, x_7, x_8, x_9, x_{13}\}, \{x_3, x_{15}\}, \{x_4, x_{14}\}, \{x_6\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}\}$	7
$X(0.971)$	$\{\{x_1, x_2, x_4, x_5, x_7, x_8, x_9, x_{13}, x_{14}\}, \{x_3, x_{15}\}, \{x_6\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}\}$	6
$X(0.944)$	$\{\{x_1, x_2, x_4, x_5, x_7, x_8, x_9, x_{13}, x_{14}\}, \{x_3, x_{10}, x_{15}\}, \{x_6\}, \{x_{11}\}, \{x_{12}\}\}$	5
$X(0.926)$	$\{\{x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{13}, x_{14}, x_{15}\}, \{x_6\}, \{x_{11}\}, \{x_{12}\}\}$	4
$X(0.905)$	$\{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{13}, x_{14}, x_{15}\}, \{x_{11}\}, \{x_{12}\}\}$	3
$X(0.888)$	$\{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{12}, x_{13}, x_{14}, x_{15}\}, \{x_{11}\}\}$	2
$X(0.865)$	$\{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}\}$	1

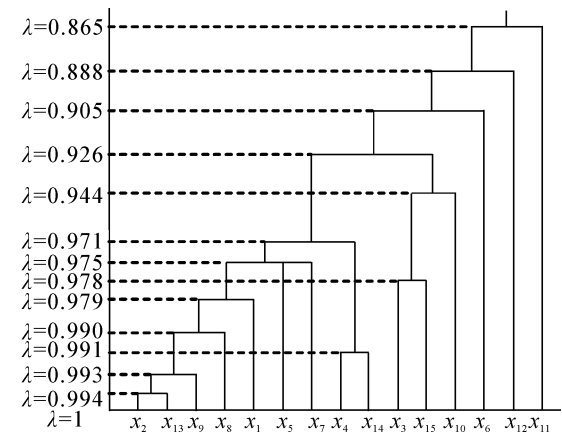


图 2 样本聚类结构图

Fig.2 The structural clustering graph

4) 按照表 5 中的离散区间对表 3 中的属性值进行离散化处理,得到具有离散化属性值的属性表,如表 6 所示。根据表 6,可以分析计算删除各个属性后的等价类。

表 5 连续属性值的离散区间

Table 5 Discrete interval of continuous attribute values

y'_{ik}	$[0,50)$	$[50,80]$	$(80,100]$
离散值	1	2	3
含义	低	中等	高

表 6 离散化处理后的属性表

Table 6 Attributes after discretization

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	2	3	3	3	3	2
x_2	2	2	3	3	2	1
x_3	1	2	3	2	1	2
x_4	3	3	2	2	3	1
x_5	1	3	3	3	2	1
x_6	3	3	1	2	3	1
x_7	3	3	2	2	2	1
x_8	2	3	3	3	2	2
x_9	3	3	3	3	2	1
x_{10}	1	3	1	3	1	3
x_{11}	1	2	3	1	1	2
x_{12}	2	1	3	3	3	2
x_{13}	2	3	3	3	2	1
x_{14}	3	3	3	2	3	1
x_{15}	1	2	3	3	1	2

$$U/\text{ind}(C) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{13}\}, \{x_{14}\}, \{x_{15}\}\}$$
$$U/\text{ind}(C - \{a_1\}) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_9, x_{13}\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{14}\}, \{x_{15}\}\}$$
$$U/\text{ind}(C - \{a_2\}) = \{\{x_1, x_{12}\}, \{x_2, x_{13}\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{14}\}, \{x_{15}\}\}$$

$$U/\text{ind}(C - \{a_3\}) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_6, x_{14}\}, \{x_5\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{13}\}, \{x_{15}\}\}$$
$$U/\text{ind}(C - \{a_4\}) = \{\{x_1\}, \{x_2\}, \{x_3, x_{11}, x_{15}\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{12}\}, \{x_{13}\}, \{x_{14}\}\}$$
$$U/\text{ind}(C - \{a_5\}) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_7\}, \{x_5\}, \{x_6\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{13}\}, \{x_{14}\}, \{x_{15}\}\}$$
$$U/\text{ind}(C - \{a_6\}) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8, x_{13}\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\}, \{x_{14}\}, \{x_{15}\}\}$$

5) 以表 4 中商空间族所形成的 14 个不同粒度商空间的样本聚类结果作为相应粒度空间下单一决策属性的分类,按照式(2)分别计算各粒度空间下各个属性的重要度,再根据式(4)综合计算得出各属性的最终重要度,计算结果如表 7 所示。

表 7 属性重要度的计算结果

Table 7 Computing result of attributes significance

序号	商空间族	属性重要度					
		a_1	a_2	a_3	a_4	a_5	a_6
1	$X(1)$	3/15	4/15	3/15	3/15	2/15	2/15
2	$X(0.994)$	3/15	2/15	3/15	3/15	2/15	2/15
3	$X(0.993)$	3/15	2/15	3/15	3/15	2/15	2/15
4	$X(0.991)$	3/15	2/15	3/15	3/15	2/15	2/15
5	$X(0.990)$	3/15	2/15	3/15	3/15	2/15	0
6	$X(0.979)$	3/15	2/15	3/15	3/15	2/15	0
7	$X(0.978)$	3/15	2/15	3/15	3/15	2/15	0
8	$X(0.975)$	0	2/15	3/15	3/15	2/15	0
9	$X(0.971)$	0	2/15	3/15	3/15	0	0
10	$X(0.944)$	0	2/15	3/15	3/15	0	0
11	$X(0.926)$	0	2/15	3/15	3/15	0	0
12	$X(0.905)$	0	2/15	0	3/15	0	0
13	$X(0.888)$	0	0	0	3/15	0	0
14	$X(0.865)$	0	0	0	0	0	0
综合值		0.1	0.124	0.157	0.186	0.076	0.038

6) 按照式(5)确定各属性的权重: $w_{a_1} = 0.147$, $w_{a_2} = 0.182$, $w_{a_3} = 0.230$, $w_{a_4} = 0.273$, $w_{a_5} = 0.112$, $w_{a_6} = 0.056$ 。由计算结果可以看出,各属性的权重由大到小的排列顺序为:按期完成率>产品合格率>生产工时率>总工时>加工难度>加班时间,这与该企业管理者正从以前单纯强调工作量和加班时间,逐步转变为更加重视任务进度的控制、产品质量的稳定、工作效率的提高这一技能人员工作绩效考核评价新思路是完全吻合的,实现了基于数据驱动的客观权重计算与决策者主观偏好的有机统一,由此初步验证了本文所提出方法的合理性和有效性。

7) 为进一步验证该方法的准确性和实用性,采用文献[12-15]所提出方法进行了属性权重计算(由于篇幅所限,本文不给出详细的计算过程)。其计算结果为: $w_{a_1} = 0.141, w_{a_2} = 0.172, w_{a_3} = 0.176, w_{a_4} = 0.191, w_{a_5} = 0.190, w_{a_6} = 0.130$, 即各属性的权重由大到小的排列顺序为:按期完成率>加工难度>产品合格率>生产工时率>总工时>加班时间。通过对比可知,2种方法属性权重计算结果的总体分布是基本一致的,但本文方法的准确性和实用性主要体现在:①以往方法的属性权重计算结果相对比较平均,差距不大,其标准差仅为0.023,因此对企业管理者新的考核评价思路的体现作用不显著;而本文方法的权重标准差为0.072,为以往方法的3.1倍,相比较而言,更能充分发挥管理者所重视指标的导向作用。②2种方法对于“加工难度”指标权重的计算结果差异较大。考虑到“加工难度”仅仅是个定性指标,目前还难以量化,区分度不强,如果权重过大,在具体实施时将很难操作,因此本文方法相对较低的权重是比较合适的。③本文方法的模糊相似度计算次数仅为以往方法的1/7,且无需进行以往方法求出模糊等价矩阵的7×4次传递闭包计算,因此计算量要小得多。

4 结束语

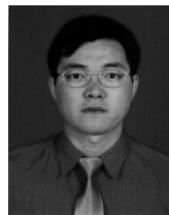
针对属性值为连续值且无单一决策属性的多属性决策问题,提出了一种基于粒计算的属性权重确定方法。与目前的研究成果相比,该方法通过从不同粗细粒度的商空间多角度、多层次地综合分析各个属性的重要程度,使属性权重的确定过程更为全面,计算结果更为准确,应用实例也说明了其合理性和有效性。同时,该方法只需进行一次模糊相似度计算,无需进行传递闭包计算,计算复杂度大幅降低,因此也更具实用性。方法的不足之处主要在于连续属性值的离散化处理过程会造成一定的客观信息的损失,所以如何应用合适的高效离散化算法来减少这种损失需要做进一步的研究。

参考文献:

- [1] 张荣.交互式不确定多属性决策研究[M]. 郑州:河南大学出版社, 2011: 10-23.
- [2] 杨立才,张朋飞,王德伟.基于主成分分析的代谢综合征模糊综合评价方法[J]. 生物医学工程学杂志, 2013, 30(1): 67-70.
YANG Licai, ZHANG Pengfei, WANG Dewei. Fuzzy comprehensive evaluation method of metabolic syndrome based on PCA[J]. Journal of Biomedical Engineering, 2013, 30(1): 67-70.
- [3] XU Zeshui, ZHANG Xiaolu. Hesitant fuzzy multi-attribute decision making based on TOPSIS with incomplete weight information[J]. Knowledge-Based Systems, 2013, 52: 53-64.
- [4] 黄宗盛,胡培,聂佳佳.基于离差最大化的交叉效率评价方法[J]. 运筹与管理, 2012, 21(6): 177-181.
HUANG Zongsheng, HU Pei, NIE Jiajia. Cross efficiency evaluation method based on maximizing deviations[J]. Operations Research and Management Science, 2012, 21(6): 177-181.
- [5] 江文奇.无量纲化方法对属性权重影响的敏感性和方案保序[J]. 系统工程与电子技术, 2012, 34(12): 2520-2523.
JIANG Wenqi. Sensibility and alternative COP analysis of dimensionless methods on effect of attribute weight[J]. Systems Engineering and Electronics, 2012, 34(12): 2520-2523.
- [6] 刘勇,JEFFREY F,刘思峰,等.基于区间直觉模糊的动态多属性灰色关联决策方法[J]. 控制与决策, 2013, 28(9): 1303-1308.
LIU Yong, JEFFREY F, LIU Sifeng, et al. Dynamic multiple attribute grey incidence decision making method based on interval valued intuitionistic fuzzy number[J]. Control and Decision, 2013, 28(9): 1303-1308.
- [7] CHEN Yuan, LI Bing. Dynamic multi-attribute decision making model based on triangular intuitionistic fuzzy numbers[J]. Scientia Iranica, 2011, 18(2): 268-274.
- [8] 陈晓红,戴子敬,刘翔.基于熵和关联系数的区间直觉模糊决策方法[J]. 系统工程与电子技术, 2013, 35(4): 791-795.
CHEN Xiaohong, DAI Zijing, LIU Xiang. Approach to interval-valued intuitionistic fuzzy decision making based on entropy and correlation coefficient[J]. Systems Engineering and Electronics, 2013, 35(4): 791-795.
- [9] 迟国泰,程砚秋,曹勇.基于聚类赋权的科学发展评价模型及实证[J]. 运筹与管理, 2011, 20(5): 94-102.
CHI Guotai, CHENG Yanqiu, CAO Yong. The scientific development evaluation model based on clustering and its empirical study[J]. Operations Research and Management Science, 2011, 20(5): 94-102.
- [10] 苏永华,刘科伟,张进华.基于粗糙集重心理论的公路隧道塌方风险分析[J]. 湖南大学学报:自然科学版, 2013, 40(1): 21-26.
SU Yonghua, LIU Kewei, ZHANG Jinhua. Fuzzy evaluation of collapse incidents in highway tunnel construction based on rough set and barycenter theory[J]. Journal of Hunan University: Natural Sciences, 2013, 40(1): 21-26.
- [11] WEI Guiwu. Gray relational analysis method for intuitionistic fuzzy multiple attribute decision making[J]. Expert Sys-

- tems with Applications, 2011, 38(9): 11671-11677.
- [12] 黄定轩, 武振业, 宗蕴璋. 基于属性重要性的多属性客观权重分配方法[J]. 系统工程理论方法应用, 2004, 13(3): 203-207.
HUANG Dingxuan, WU Zhenye, ZONG Yunzhang. An im-personal multi-attribute weight allocation method based on attribute importance [J]. Systems Engineering – Theory Methodology Applications, 2004, 13(3): 203-207.
- [13] 柳炳祥, 李海林. 基于模糊粗糙集的因素权重分配方法[J]. 控制与决策, 2007, 22(12): 1437-1440.
LIU Bingxiang, LI Hailin. Method of factor weights allocation based on combination of fuzzy and rough set[J]. Control and Decision, 2007, 22(12): 1437-1440.
- [14] 刘文军. 连续值域决策表的一种属性权重确定方法[J]. 模糊系统与数学, 2008, 22(3): 160-166.
LIU Wenjun. An algorithm of obtain weight of continuous attributes in continuous domain decision table [J]. Fuzzy Systems and Mathematics, 2008, 22(3): 160-166.
- [15] 吴静, 吴晓燕, 高忠长. 基于模糊聚类 and 粗糙集的仿真可信性模糊综合评估[J]. 系统工程与电子技术, 2010, 32(4): 770-773.
WU Jing, WU Xiaoyan, GAO Zhongchang. Fuzzy comprehensive evaluation of simulation credibility based on fuzzy clustering analysis and rough sets theory[J]. Systems Engineering and Electronics, 2010, 32(4): 770-773.
- [16] 张钹, 张铃. 粒计算未来发展方向探讨[J]. 重庆邮电大学学报: 自然科学版, 2010, 22(5): 538-540.
ZHANG Bo, ZHANG Ling. Discussion on future development of granular computing[J]. Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2010, 22(5): 538-540.
- [17] 王国胤, 张清华, 胡军. 粒计算研究综述[J]. 智能系统学报, 2007, 2(6): 8-26.
WANG Guoyin, ZHANG Qinghua, HU Jun. An overview of granular computing [J]. CAAI Transactions on Intelligent Systems, 2007, 2(6): 8-26.
- [18] SANKAR K P, SAROJ K M. Natural computing: a problem solving paradigm with granular information processing [J]. Applied Soft Computing, 2013, 13(9): 3944-3955.
- [19] 张铃, 张钹. 模糊商空间理论(模糊粒度计算方法)[J]. 软件学报, 2003, 14(4): 770-776.
ZHANG Ling, ZHANG Bo. Theory of fuzzy quotient space (methods of fuzzy granular computing) [J]. Journal of Software, 2003, 14(4): 770-776.
- [20] ZHANG Ling, ZHANG Bo. The structure analysis of fuzzy sets [J]. International Journal of Approximate Reasoning, 2005, 40(1/2): 92-108.
- [21] 苗夺谦, 李道国. 粗糙集理论、算法与应用[M]. 北京: 清华大学出版社, 2008: 24-81.
MIAO Duoqian, LI Daoguo. Rough sets theory, algorithms and applications [M]. Beijing: Tsinghua University Press, 2008: 24-81.
- [22] 王国胤, 姚一豫, 于洪. 粗糙集理论与应用研究综述[J]. 计算机学报, 2009, 32(7): 1229-1246.
WANG Guoyin, YAO Yiyu, YU Hong. A survey on rough set theory and applications [J]. Chinese Journal of Computer, 2009, 32(7): 1229-1246.
- [23] 杨纶标, 高英仪. 模糊数学原理及应用[M]. 4 版. 广州: 华南理工大学出版社, 2008: 66-77.
- [24] TANG Xuqing, ZHU Ping, CHENG Jiaxing. Cluster analysis based on fuzzy quotient space [J]. Journal of Software, 2008, 19(4): 861-868.

作者简介:



周丹晨, 男, 1969 年生, 高级工程师, 博士, 主要研究方向为人工智能、数字化设计与制造, 发表学术论文 50 余篇, 其中被 SCI、EI 检索 20 余篇。