

文章编号: 1001-0920(2009)06-0957-04

云决策分析方法

柳炳祥, 李海林, 杨丽彬

(景德镇陶瓷学院 信息工程学院, 江西 景德镇 333403)

摘要: 提出一种基于不确定性人工智能的决策分析方法. 该方法利用云理论中的云变换和概念提升方法来刻画属性因素的概念云模型, 并对每个因素都分配相应的云权重; 使用 X -条件云和极大判定法, 找出每个因素的最佳评判语言云模型; 让云权重参与综合评判, 最终通过云计算得到基于云滴分布的综合评价结果. 实验结果表明, 与传统的模糊决策分析相比, 它更具有易理解性、广普适性和高准确性等特点.

关键词: 云理论; 决策分析; 模糊理论; 云计算

中图分类号: TP18

文献标识码: A

Cloud decision analysis method

LIU Bing-xiang, LI Hai-lin, YANG Li-bin

(School of Information Engineering, JDZ Ceramic Institute, Jingdezhen 333403, China. Correspondent: LIU Bing-xiang, E-mail: lbx1966@163.com)

Abstract: The analysis method based on the artificial intelligence with uncertainty is introduced. The ways of the cloud transformation and the climbed-up conception are used to represent the cloud models of different attributes factors and distributes the cloud weight for each factor. Then the best evaluation cloud models for each factor are found by using the relative cloud theory. At the same time, the cloud weight participates in computing with the synthesis evaluation. Finally, the distributing of cloud drops expresses the result of evaluation. The experiment results show that it is easier to understand the meaning of the result and has wider universality, higher veracity than the fuzzy decision analysis.

Key words: Cloud theory; Decision analysis; Fuzzy theory; Cloud computation

1 引言

由于自然语言的不确定性, 人工智能研究主要体现在知识和推理的模糊性和随机性的研究. 传统的方法主要有统计学、模糊数学、粗糙集等理论. 统计学家认为, 事件的发生是非此即彼性的, 不存在亦此亦彼的中间状态; 模糊学家则认为, 事件的发生是相对的, 是亦此亦彼的, 具有很强的模糊性. 但在现实生活中, 事件发生同时具有模糊性与随机性, 并且二者有一定的关联. 因此, 李德毅在统计学家和模糊数学家的基础上, 提出了定性定量不确定性转换模型, 即云模型^[1,2], 实现了定性概念与定量之间的不确定转换.

评价客体的评判结果取决于这个对象在一系列因素上的评判结果, 综合评判就是由分指标上的评判结果得到综合评判结果的过程. 范定国等^[3]使用

云模型理论用定性方法对单因素进行评价, 但存在一些不足, 主要表现在: 权重参与构建云模型数字特征时, 对期望值上界溢出的修正函数不科学, 原因是上溢数据的流失会导致概念定性的误差; 客体的满意度虽然实现了硬划分, 但并没有对因素的各个评语实现客观软划分, 原因是与云模型的定性概念相背离; 在权重参与计算时, 没有构造云权重参与云计算, 使得计算结果依赖于主观权重的设定. 基于上述缺陷, 本文提出一种能对评价客体进行定性评价的云决策分析方法. 该方法不仅能够有效地克服这些缺点, 而且能够使得综合评判结果更贴近实际.

2 模糊决策分析

模糊综合评判决策的数学模型^[4]是对受多种因素影响的事物作出全面评价的一种十分有效的多因素决策方法, 具体如下:

收稿日期: 2008-05-04; 修回日期: 2008-07-07.

基金项目: 国家自然科学基金项目(60661003).

作者简介: 柳炳祥(1966—), 男, 江西九江人, 教授, 博士, 从事数据挖掘、决策支持系统的研究; 李海林(1982—), 男, 福建龙岩人, 硕士生, 从事数据挖掘、智能控制的研究.

设有如下2个论域:

$$U = \{u_1, u_2, \dots, u_n\},$$

$$V = \{v_1, v_2, \dots, v_m\}.$$

其中: U 为 n 种因素或指标组成的集合, 即因素集; V 为 m 种评判组成的集合, 即评判集.

由于因素集中各个因素所处的地位不同, 作用也不一样, 导致权重不同, 从而评判也不同. 因此, 综合评判应该是 V 上的一个模糊子集 B , 即 $B = (b_1, b_2, \dots, b_m) \in \mathbf{P}(V)$, 其中 b_j 反映了第 j 种评判在综合评判中所占的地位. 综合评判 B 依赖于各个因素的权重, 它应该是 U 上的模糊子集 $A = (a_1, a_2, \dots, a_n) \in \mathbf{P}(U)$ 且 $\sum_{i=1}^n a_i = 1$. 其中: a_i 表示 U 中第 i 个因素的权重, $\mathbf{P}(U)$ 表示 U 的幂集.

如果有一个从 U 到 V 的模糊关系 $R = (r_{ij})_{n \times m}$, 那么模糊关系 R 可诱导出 U 到 V 的模糊变换 T_R . 其中可以将 T_R 看作由权重 A 得到的综合评判 B 的数学模型.

由以上分析, 可构造以下数学模型:

- 1) 因素集 $U = \{u_1, u_2, \dots, u_n\}$;
- 2) 评判集 $V = \{v_1, v_2, \dots, v_m\}$;
- 3) 单因素评判

$$U \xrightarrow{\mathbf{P}(V)} u_i / (u_i) = \{r_{i1}, r_{i2}, \dots, r_{im}\} \in \mathbf{P}(V),$$

$R = (r_{ij})_{n \times m}$ 为从 U 到 V 的模糊关系. 这样便由 $\{U, V, R\}$ 构成了一个模糊综合决策模型.

对于权重 $A = (a_1, a_2, \dots, a_n) \in \mathbf{P}(U)$, 根据实际情况选取合适的模糊算子, 可得到综合评判 $B = A \circ R$; 再按模糊数学的最大隶属原则, 最后得出综合评判结果.

3 云决策分析

由第2节的介绍可知, 在模糊综合决策的数学模型中, 尽管因素集的元素概念不相同, 但对于每个因素集的元素概念的评价元素却都一样, 即对事物的不同属性用同样的语言加以描述. 模糊决策分析之所以要建立相同的评判语言集, 是因为它是一种基于矩阵和向量之间的运算, 而向量与矩阵的运算很大程度上要遵循它们的维数相等. 因此, 模型决策分析是一种利用同种评判因子集来评价事物多种不同属性因素且最终全面考虑属性因素而作出综合评判的方法. 然而, 在生产实践中, 对事物因素属性的评价语言往往是各不相同、丰富多彩的. 即同种事物的不同属性因素可以用不同的评价语言进行单因素的描述, 而最后的综合评价也可以用不同的评价语言作出最后的综合评判, 即用来评价对象的语言个

数或种类都不相同. 在此种情况下, 模糊综合评判模型已失去了原先的效能, 况且在处理不确定性的问题时, 如果过分地强调评判集因子语言值的大小, 会使得整个模糊处理机制趋向于数值的精确化运算, 与不确定分析原理存有一定的矛盾. 因此, 在处理不确定性问题时, 利用不确定性人工智能进行决策分析才是最佳的选择, 故提出一种基于不确定性人工智能的决策分析模型——云决策分析.

云决策分析的主要思想是: 利用云变换和概念提升^[5] 或者根据专家经验给出刻画各属性因素的概念云模型; 对每个因素都分配相应的云权重(用云来描述权重); 根据群体对事物各因素的评判数据, 利用 X -条件云和极大判定法找出每个因素的最佳评判语言云模型; 让云权重参与综合评判, 最后得出该事物的云决策分析结果.

定义1 给定论域上的2个云 C_1 和 C_2 , 分别描述为 $C_1(Ex_1, En_1, He_1)$, $C_2(Ex_2, En_2, He_2)$, 则有如下2个云的加运算和乘运算:

$$C_1 + C_2 = (Ex_1 + Ex_2, \sqrt{En_1^2 + En_2^2}, \sqrt{He_1^2 + He_2^2}),$$

$$C_1 * C_2 = \begin{bmatrix} Ex \\ En \\ He \end{bmatrix} = \begin{bmatrix} Ex_1 Ex_2 \\ | Ex_1 Ex_2 | \times \sqrt{\left(\frac{En_1}{Ex_1}\right)^2 + \left(\frac{En_2}{Ex_2}\right)^2} \\ | Ex_1 Ex_2 | \times \sqrt{\left(\frac{He_1}{Ex_1}\right)^2 + \left(\frac{He_2}{Ex_2}\right)^2} \end{bmatrix}.$$

定义2 向量云模型是由云模型作为分量组成的向量. 若2个向量云模型的分被描述为 $X = (XC_1, XC_2, \dots, XC_n)$ 和 $Y = (YC_1, YC_2, \dots, YC_n)$, 则向量云模型的内积为

$$X \circ Y = X * Y = (XC_1, XC_2, \dots, XC_n) \begin{bmatrix} YC_1 \\ YC_2 \\ \dots \\ YC_n \end{bmatrix} = XY(Ex, En, He).$$

下面对云决策分析过程进行详细的阐述:

1) 云的词计算所涉及的概念必须属于同一个论域才有意义, 因而对于云决策分析的后期阶段, 要进行词计算首先就得将所有的概念转化为能够进行云运算的同一论域中. 因此, 不同属性的量纲消除是进行云决策分析的第1步, 即将所有的属性值都映射到区间 $[0, 1]$ 内.

2) 与模糊决策分析一样,列出因素集 $U = \{u_1, u_2, \dots, u_n\}$. 若 $n = 4$, 则假设因素集为服装的花色、样式、质量及价格. 而评价集 $V = \{V_1, V_2, \dots, V_{n+1}\}$, 若 $n = 4$, 则该集合元素分别表示对应每个属性因素的评价语言集合和综合评判该事物的语言集合. 如对于花色的评价语言集合 $V_1 = \{\text{漂亮, 一般, 很丑}\}$, 而对样式的评价语言集合 $V_2 = \{\text{时髦, 过时}\}$ 以及对该服装的综合评价语言集 $V_5 = \{\text{优品, 中品, 次品}\}$. 因此, 评价集中元素的语言集因子的个数和描述可以各不相同, 这是与模糊决策分析的不同点之一.

事实上, 在此过程中, 若先前没有建立好各语言的云模型, 则需要通过云变换和语言概念提升, 将连续性数据离散化成定性描述概念, 最终得到符合实际要求的概念云模型. 这个过程相应于模糊决策分析中的单因素评价过程, 二者的区别主要在于定量数据与定性概念的描述差异上, 即模糊决策是用定量数据对因素进行评价, 而云决策分析则用云模型对其进行评价, 使得属性因素的不确定性得到了不确定的评价, 这符合实际要求. 如花色的评价语言集 V_1 用云模型表示成 $V_1 = \{C_1^1, C_2^1, C_3^1\}$, 其中 C_i^j 表示第 j 个属性的第 i 个评价因子.

3) 对每个因素进行云权重分配, 各个云权重组成的集合为 $W = \{WC_1, WC_2, \dots, WC_n\}$, $WC_i = \{wEx_i, wEn_i, wHe_i\}$. 可以将云权重理解为传统权重的“软化”, 也称为“软权重”, 3 个数字特征的数字期望 wEx 就是传统理念上的权重, 而熵与超熵是用来软化权重的参数, 可根据具体研究的问题来确定 wEn 和 wHe 的值.

4) 根据群体对各个因素的打分求平均值, 即 $X = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$. 利用 X - 条件云发生器和极大判定法找出各个被激活的最佳评价概念云模型 $R = \{RC_1, RC_2, \dots, RC_n\}$.

5) 让云权重参与运算. 利用定义 1 和定义 2, 计算出最终的综合评价云模型 RC . 即

$$RC = R * W = (RC_1, RC_2, \dots, RC_n) \begin{bmatrix} WC_1 \\ WC_2 \\ \dots \\ WC_n \end{bmatrix}$$

6) 根据云模型 RC 在综合评价概念云模型云图的分布情况, 便可发现该群体对事物(服装)的评判结果. 该结果已不是一个定值, 而是一个统观全局的分布图. 这比用一个定量数据来评价事物更直观、更科学、更符合人类的思维方式.

4 实例分析

下面通过实例对上述分析过程作进一步说明,

使用本文提出的云决策分析方法对服装进行综合评价, 并与传统的模糊决策分析过程进行比较.

为了简单介绍云决策分析的方法, 假设对服装的评价属性只有 2 个(质量和价格). 根据云决策分析, 通过云变换和概念提升可以将属性值离散化并得到合适的概念云模型(当然也可以根据专家给出的概念云模型^[6,7]). 假设质量由概念“好”和概念“差”描述, 其对应的云模型分别为 $C_1^1(0.9, 0.15, 0.007)$, $C_2^1(0.3, 0.1, 0.005)$; 价格由概念“便宜”、“适中”和“贵”描述, 其对应的云模型为 $C_2^2(0.9, 0.1, 0.003)$, $C_3^2(0.6, 0.08, 0.006)$, $C_1^2(0.2, 0.13, 0.002)$; 而综合评判的概念云模型由“欢迎”、“一般”和“不欢迎”3 个概念描述, 对应的云模型为 $C_3^3(0.85, 0.12, 0.005)$, $C_2^3(0.63, 0.09, 0.003)$, $C_1^3(0.3, 0.15, 0.009)$. 且对每个属性分配的云权重为 $WC_1(0.6, 0.15, 0.006)$, $WC_2(0.4, 0.1, 0.001)$, 一群客户对 2 个属性因素打分的平均结果为 $X = (x_1, x_2) = (0.8, 0.7)$.

通过上述假设, 根据步骤 4), 可以求得规则前件被激活的概念分别为质量“好”和价格“适中”, 如图 1 和图 2 所示, 箭头所指的灰色圈代表被激活相应概念的点, 故可得到最佳评价概念云模型 $R = (C_1^1, C_2^2)$.

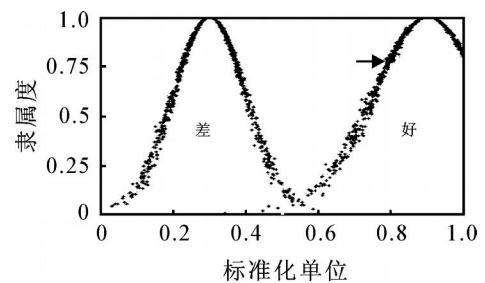


图 1 质量各概念云模型

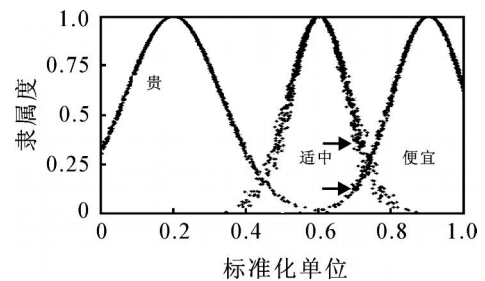


图 2 价格各概念云模型

云权重参与运算过程如下:

$$RC = R * W = (C_1^1, C_2^2) \begin{bmatrix} WC_1 \\ WC_2 \end{bmatrix} = C_1^1 * WC_1 + C_2^2 * WC_2$$

通过上述算法计算得

$$C_1^1 * WC_1 = (0.54, 0.1622, 0.0068),$$

$$\tilde{C}_2 * WC_2 = (0.24, 0.0453, 0.0016).$$

最终得到的云模型描述为

$$RC = (0.78, 0.1684, 0.0070).$$

在未评判之前,对服装的综合描述概念为“欢迎”、“一般”和“不欢迎”,具体云模型如图3所示.而通过上述云决策分析后,评判结果RC在综合评判云图的分布如图4所示,箭头所指的灰色代表评判结果云模型RC.

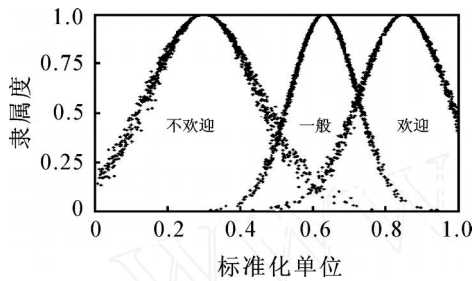


图3 服装综合评价概念云模型

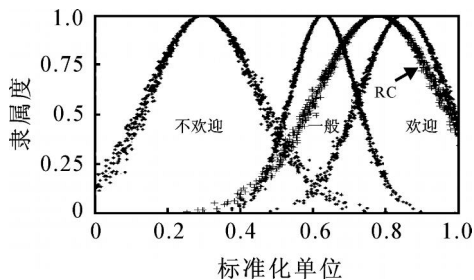


图4 评价结果在原综合概念云图中的分布情况

从图4可以发现,评判结果云模型RC覆盖了概念“一般”和概念“欢迎”,这说明该群体对服装的评价是介于“一般”与“欢迎”之间,而整个评价的具体情况可以从云图分布中获知.这与模糊决策有着本质的区别,模糊决策分析只能通过最大隶属度原则,将决策分析的结果强硬地归于评价集中的某个概念,却不能像云决策分析那样,给出具体的云图分布情况.如果决策用户须知道评判结果到底归属于3个评价概念中的哪个概念,可通过分析箭头所指的灰色云图的云滴被这3个评价概念的云滴占有率来获知.从图4容易发现,评价云模型的云滴绝大部分被“欢迎”概念所占有.当然,也可以通过结果云模型的期望值激活与之相交的概念云模型的最大确定度来决定其归属于哪一类.由图4易获知,结果期望值激活概念“一般”的隶属度明显比激活概念“欢

迎”的隶属度小,因此与模糊决策的综合评价结果一样,服装的云综合评判结果为“欢迎”.

5 结论

在工程实践中,由于事物不确定性的存在,图4的评判结果分布图才是用户最想要的,它不仅能够表达最终的评判结果,而且能够知道其具体的分布情况以及被评价概念的占有比,甚至还可以通过这种分布图对未来的服装销售作进一步的预测和规划.因此,云决策分析不仅使云理论在决策分析理论中发挥了优势,更为重要的是为生产实践提供了可靠的可视化评价结果,其表达的效果更具说服力.

参考文献(References)

- [1] 李德毅. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005.
(Li D Y. Intelligence with uncertainty [M]. Beijing: Nation Defence Industry Press, 2005.)
- [2] 邱凯昌. 空间数据挖掘和知识发现[M]. 武汉: 武汉大学出版社, 2003.
(Di K C. Spatial data mining and knowledge discovery [M]. Wuhan: Wuhan University Press, 2003.)
- [3] 范定国. 一种基于云模型的综合评判模型[J]. 科技情报开发与经济, 2003, 13(12): 157-159.
(Fan D G. Synthesis judge model based on cloud model [J]. Sci-Tech Information Development & Economy, 2003, 13(12): 157-159.)
- [4] 谢季坚. 模糊数学方法及其应用[M]. 武汉: 华中科技大学出版社, 2003, 7: 190-203.
(Xie J J. Method and application of fuzzy mathematics [M]. Wuhan: Publishing House of Huazhong University of Science & Technology, 2003, 7: 190-203.)
- [5] 蒋嵘, 李德毅, 范建华. 数值型数据的泛概念树的自动生成方法[J]. 计算机学报, 2000, 23(5): 470-476.
(Jiang R, Li D Y, Fan J H. Automatic generation of par-concept-tree on numerical data [J]. Chinese J of Computers, 2000, 23(5): 470-476.)
- [6] Li Deyi, Du Yi, Yin Guoding, et al. Commonsense knowledge modeling [C]. 16th World Computer Congress. Beijing, 2000: 245-252.
- [7] Li Deyi, Han J, Shi X M. Knowledge representation and discovery based on linguistic models [M]. Singapore: World Scientific Press, 1977: 3-20.