2014 年 1 月 郑州大学学报(工学版) Jan. 2014
第 35 卷 第 1 期 Journal of Zhengzhou University (Engineering Science) Vol. 35 No. 1

文章编号:1671 - 6833(2014)01 - 0043 - 03

带有可信度的偏序偏好决策信息的聚合方法

陈树伟, 周 威, 蔡丽娜

(郑州大学 电气工程学院, 河南 郑州 450001)

摘 要: 提出了一种解决带有可信度的偏序形式的偏好信息的决策方法. 对于决策者给出的带有可信度的偏好信息,通过把每个偏好看成一个序列,对于每对方案,利用上下文概率计算一个方案排到另一个方案之前的概率. 然后,根据偏好排序算法对这些概率进行聚合,得到一个近似最优的总方案排序. 并通过实例说明其有效性.

关键词: 决策;偏好;偏序;可信度;聚合

中图分类号: TP182 **文献标志码**: A doi:10.3969/j.issn.1671 - 6833.2014.01.010

0 引言

在实际应用中,决策是至关重要的一步,如组织管理、财务规划和产品评价等. 多数情况下,是根据现有的标准或偏好信息[1],在一组方案中选择出最合适的一个或多个方案. 在许多现实的决策问题中,对于决策者而言,运用方案间的优先顺序形式比提供定量的偏好度更自然和合理. 此外,现实中的决策问题有很多的条件或标准的限制,可是由于决策问题的不确定性和复杂性[2],不是所有的条件都能够同时得到满足,有时,由于时间的限制、知识或数据的缺少、有限的专业知识等等,决策者对于每个属性的判断并不总是可行的. 另外,偏序结构还避免了对一些无关信息的比较.

对于偏好排序目前使用较广泛的是:波达计数法或波达规则[3-4],这种方法存在的问题是:通常,绝对的分数并不容易得到;同样的分数对于不同的决策者可能意义不同[5-6]. 为了克服这些问题,Cohen[5]提出了根据决策者偏好关系生成备选方案的近似最优排序的算法,而无须给各方案分配分数. 在此基础上,wang[7]提出了根据决策者给出的偏好关系计算每对方案偏好度的方法,然后将 Cohen 提出的算法应用于这些偏好度,生成备选方案的近似最优排序.

现有的偏好决策方法很少考虑由不同决策者给出的偏好信息的可信度,以及其他不确定性,比如不同意见的一致性程度. 这些不确定因素总是伴随在个人偏好表达的过程中. 因此有必要研究有关带有可信度的偏序偏好信息的聚合问题.

笔者将决策者给出的带有可信度的偏序形式的偏好信息作为序列,对于每一对方案,利用上下文概率函数计算一种方案被排到另一种方案之前的概率. 然后,运用偏好排序算法[5]得到一个总的方案排序,这个总的排序可以最大程度地体现所有决策者给出的偏好信息[1].

1 带有可信度的偏好信息的聚合

考虑如下决策问题:一组决策者或专家 $E=\{e_1, e_2, \cdots, e_n\}$,其对应的权重:$W=\{\omega_1, \omega_2, \cdots, \omega_n\}$,$\sum\omega_i=1$. 对一组备选方案 $A=\{a_1, a_2, \cdots, a_m\}$进行评估并给出对这组方案的偏好. 记决策者给出的所有偏好的集合为 D,在这里 D 中的每一个元素都是一个偏序,如:$a_1a_2a_5$, a_4a_3, $a_3a_5a_2a_1$. 另外偏好集合 D 伴随一组可信度:$B=\{b_1, b_2, \cdots, b_N\}$,其中 N 是 D 中偏好的个数且 $0\leqslant b_i\leqslant 1$, $i=1$, $i=1,\cdots, N$,每一个 B 中的元素对应的表示 D 中偏好的可信度,如 $a_1a_2a_5$的可信度是 $b_1=0.85$. 我们的任务是找到一个备选方案的排序,使其最大程度地体现所有决策者给出的偏好信息.

采用 wang[7]的观点:我们认为:根据决策者提供的偏好信息计算一个方案在另一个方案之前

收稿日期:2013 - 08 - 20;**修订日期**:2013 - 10 - 20

基金项目:国家自然科学基金资助项目(61175055); 教育部高校博士点专项科研基金(20114101110005)

作者简介:陈树伟(1977 -),男,河南新乡人,郑州大学讲师,博士,主要从事模糊控制与决策等方面的研究,E - mail:swchen@zzu.edu.cn.

的概率,比给每个方案分配绝对的分数更合理. 这些概率是通过把每个偏好当作一个序列后基于上下文概率得到的. 接下来,运用偏好排序算法把得到的概率作为输入来产生一个总的排序结果,而且,这个总的排序结果可以最大程度地体现所有决策者给出的偏好信息.

1.1 上下文概率

设 V 是一个集合,F 是 V 上的 δ 域或扩展数据空间,如:若 V 是有限的,则在 V 上的 δ 域即是 V 的幂集,对于 V 上的概率函数 P 是从 F 到[0,1]上的映射,上下文概率[5]是一个函数 $G:F\to[0,1]$对于 $X\in F$, 有

$$G(X) = \sum_{E\in F}P(E)f(X\cap E)/K. \quad (1)$$

式中:$K=\sum_{E\in F}P(E)f(E)$是一个标准化因子,且对于 $X\in F$,设非负函数 $f(X)$满足 $f(x_1\cup x_2)=f(x_1)+f(x_2)$,当 $x_1\cap x_2=\phi$.

例如,假设 $f(X)=|X|$,$f(X)$是 X 的基数. ,都称作一个邻域,并且,若 E 与 X 不重叠(即 $f(X\cap E)\neq 0$),那么,E 是 X 的一个邻域. 对于 $t\in V$,如果$t\in E$,我们称:E 涵盖了 t.

基于初始概率 P,我们可以根据公式(1)计算出上下文概率. 但在实际中,P 通常是未知的. 这时我们可以根据样本数据 D,按照公式(2)的方法去估算上下文概率.

定理[1][7]设 N 是样本数据集合 D 中元素的个数,由于 $t\in V$,得

$$\tilde{G}(t) = \frac{1}{NK}\sum_{x\in D}\mathrm{cov}(t,x). \quad (2)$$

式中:$\mathrm{cov}(t,x)$是指 t 和 x 的所有相同邻域的数目.

基于上下文概率函数,wang[8]通过把每个偏好看成一个序列,提出了一种计算一对偏好可以排成一个序列的概率方法. 但是,该方法却没有考虑到可信度的问题,为了解决这个问题,我们可以提出一种新的上下文概率计算公式为

$$\tilde{G}(t) = \frac{1}{NK}\sum_{i=1}^{N}b_i\mathrm{cov}(t,x_i). \quad (3)$$

这里的 $x_i\in D$,且 b_i是 x_i的可信度.

1.2 cov(t,x_i)的计算

要计算公式(3)的结果,首先求出 $\mathrm{cov}(t,x_i)$的大小,即两个序列的所有公共子序列的一个算法.

算法 1 求两个序列的公共子序列

输入:序列 α,β

输出:α 和 β 的公共子序列列表 $L_{\alpha,\beta}$

其流程图如图 1 所示.

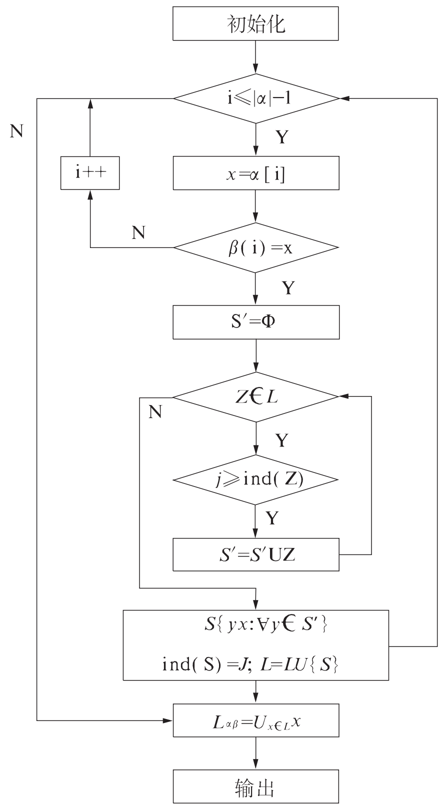

图 1 求公共子序列的算法流程图

Fig. 1 Algorithm flow chart of seeking common subsequences

例 1. 对于序列 $\alpha=qwer,\beta=qewr$ 由算法 1 可得.

初始化:$S=\{\Phi\}$,$\mathrm{ind}(S)=0$

$L=\{\{\Phi\}\}$

$i=0$: $S=\{q\}$,$\mathrm{ind}(S)=0$

$L=\{\{\Phi\},\{q\}\}$

$i=1$: $S=\{w,qw\}$,$\mathrm{ind}(S)=2$

$L=\{\{\Phi\},\{q\},\{w,qw\}\}$

$i=2$: $S=\{e,qe\}$,$\mathrm{ind}(S)=1$

$L=\{\{\Phi\},\{q\},\{e,qe\},\{w,qw\}\}$

$i=3$: $S=\{r,qr,er,qer,wr,qwr\}$,$\mathrm{ind}(S)=3$

$L=\{\{\Phi\},\{q\},\{e,qe\},\{w,qw\},\{r,qr,er,qer,wr,qwr\}\}$

从而,公共子序列为:$L_{\alpha,\beta}=\{\Phi,q,e,qe,w,qw,r,qr,er,qer,wr,qwr\}$,进而可得 $\mathrm{ncm}(\alpha,\beta)=12$.

1.3 计算一个方案排在另一方案之前的概率

对于两个备选方案 α_i 和 α_j,为了计算方案在 α_j 之前的概率,我们设 $\alpha=\alpha_i\alpha_j$,可以利用下面的公式得到这个概率

$$\tilde{G}(t) = \frac{1}{NK}\sum_{i=1}^{N}b_i\mathrm{ncm}(\alpha,x_i). \quad (4)$$

式中:N 是集合 D 中序列的个数;K 是与 α 无关的

标准化因子；$x_i \in D$；b_j 是 x_i 对应的可信度；$\mathrm{ncm}(\alpha, x_i)$是序列 α 和 x_i 的所有公共子序列的数目. 最后根据所有序列成对比较的概率大小，利用偏好排序算法[3]得到对于所有方案的一个近似最优的总体排序结果.

2　偏好排序

根据上面计算出的概率大小，我们可以利用一个算法对所有方案进行排序，最后得到一个近似最优总的排序结果. 这个算法的基本内容如下.

算法 2　偏好排序算法

输入：备选方案 X；偏好函数 PREF

输出：一个近似最优排序函数 ρ

```
let V = X
for each v ∈ V do
    φ(v) = ∑_{u∈V} PREF(v,u) − ∑_{u∈V} PREF(u,v)
 while V is not empty do
 Lett = arg max_{u∈V} φ(u)
 Let ρ(t) = |V|
 Let V = V − {t}
for each v ∈ V do
    φ(v) = φ(v) + PREF(t,v) − PREF(v,t)
end
```

算法 2 中的偏好函数 *PREF* 即公式(4)计算得到的 $\tilde{G}$ 概率，而表示具有最大函数值的 V 中的元素.

3　例子

假设有 3 个专家：e_1，e_2，e_3，5 种备选方案：a，b，c，d，e. 每个专家给出的偏好为：e_1：$abcde$，e_2：ab，$aced$，e_3：bc，aed，假设所有专家的权重是相同的，得到一个偏好集合：$D = \{abcde, ab, aced, bc, aed\}$，对于这些偏好的可信度为 $B = \{0.7, 0.9, 0.95, 0.85, 0.95\}$

假设查询序列为 $q = ac$，根据公式(4)来计算概率 $\tilde{G}(q)$，首先计算 $\mathrm{ncm}(q, x)$即序列 q 和 x 的所有公共序列的数目.

对于序列 $x = abcde$. x 的所有子序列为$\{\Phi$，a，b，c，d，e，ab，ac，ad，ae，bc，bd，be，cd，ce，de，abc，abd，abe，acd，ace，ade，bcd，bce，bde，cde，$abcd$，$abce$，$abde$，$acde$，$bcde$，$abcde\}$. ac 序列的所有子集为$\{\Phi, a, c, ac\}$. 因此(由算法 1)可得 $\mathrm{ncm}(ac, abcde) = 4$. 同样地 $\mathrm{ncm}(ac, ab) = 2$，$\mathrm{ncm}(ac, aced) = 4$，$\mathrm{ncm}(ac, bc) = 2$，和 $\mathrm{ncm}(ac, aed) = 2$. 因此，$\tilde{G}(ac) = \frac{1}{NK}(0.7 \times 4 + 0.9 \times 2 + 0.95 \times 4 + 0.85 \times 2 + 0.95 \times 2) = \frac{12}{NK}$. 同样我们可以得到：$\tilde{G}(ab) = \frac{11.9}{NK}$，$\tilde{G}(ad) = \frac{13.05}{NK}$，$\tilde{G}(ae) = \frac{13.05}{NK}$，$\tilde{G}(bc) = \frac{10.85}{NK}$，$\tilde{G}(bd) = \frac{10.1}{NK}$，$\tilde{G}(be) = \frac{10.1}{NK}$，$\tilde{G}(cd) = \frac{11.1}{NK}$，$\tilde{G}(ce) = \frac{11.1}{NK}$和 $\tilde{G}(de) = \frac{10.25}{NK}$. 运用算法 2，可以得到近似最优的总的排序为：$a$，$b$，$c$，$e$，$d$，并且这个排序最大程度地体现了由 3 个专家给出的所有偏好[3]. 决策者可以根据这个结果做出决定. 同时，对于可信度都为 1 的情况，公式(3)将退化为公式(2)，从而运用笔者的方法与未考虑信任度的方法所得结果是一致的.

4　结论

笔者针对带有可信度的偏序形式的偏好信息的决策问题，提出了一种解决方法. 即首先，把决策者给出的带有可信度的每对偏好信息看成一个序列. 其次，对于每对方案，可以利用上下文概率计算出一个方案排在另一个方案之前的概率，再根据偏好排序算法对这些概率进行聚合，得到一个近似最优的总方案排序. 最后，用一个例子说明了这一方法的有效性.

参考文献：

[1] HERRERA F, AAONSO S, CHICLANA F, et al. Computing with words in decision making: foundations, trends and prospects[J]. Fuzzy Optimization and Decision Making, 2009,8(4): 337－364.

[2] CHEN Y-Hsiu, WANG Tien-Chin, WU Chao-Yen. Multi-criteria decision making with fuzzy linguistic preference relations[J]. Applied Mathematical Modelling,2011,35(3): 1322－1330.

[3] HECKELMAN J C. Probabilistic Borda rule voting[J]. Social Choice and Welfare,2003,21(3): 455－468.

[4] BLACK D. Partial justification of the Borda count[J]. Public Choice, 1976,28(1): 1－15.

(下转第 50 页)