

不确定性研讨信息的结构化与融合方法

向 东 赵 勇 陈 阳

(华中科技大学系统工程研究所 武汉 430074)

摘 要 群体研讨中专家思维的不确定性、研讨信息的非结构化给达成研讨共识带来了困难。针对此问题提出了一个由自然属性与人工属性集成的研讨信息模型,结合论点框架、有效论点组和支持分配函数等概念,研究了不确定性研讨信息提取的机制,探讨了一种基于平均论据的信息融合方法,旨在促进认识的螺旋上升和群体共识的达成。最后结合以汽车为对象的客户需求研讨,用案例说明了模型和方法的可行性与有效性。

关键词 群体研讨,结构化信息,信息融合

中图分类号 TP391 文献标识码 A

Method of Structure and Fusion for Uncertainty Seminar Information

XIANG Dong ZHAO Yong CHEN Yang

(Institute of Systems Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract In group discussion, because the thinking of experts is uncertainty and information is unstructured, group discussion is difficult to reach consensus. To solve this problem, this article gives a information model which is composited of natural and artificial property. Mechanism of how to extract uncertainty information is studied with argument framework, valid arguments group and distribution function. A method is discussed, which is information fusion based on average argument. All is to promote awareness of the spiral and groups to reach consensus. Finally, a case is provided to prove feasibility and effectiveness of model and method, which is discussion about customer's demand of car.

Keywords Group discussion, Structured information, Information fusion

1 引言

综合集成研讨厅是以综合集成方法为指导,解决复杂决策问题为目的,将专家体系、知识体系和机器体系三者综合集成构建的一个智能人机系统。与一般决策支持系统相比,研讨厅更强调专家的智慧与思维过程,用于借助专家群体的主动性和创造思维来解决一些定性与定量相结合的问题。通常,专家在研讨中除希望获取信息并以某种恰当方式表达自己的意见外,还具有与群体保持相容的意愿。但是专家个体因为有限理性、知识背景与经验的不同,对问题的思考势必存在局限性或者差异性,所以他们的研讨发言带有不确定性,并且其内容具有非结构化的特点,使得专家意见的集结非常困难,导致了研讨效率的低下。为了提高研讨效率,加快研讨进程和群体意见的收敛,不确定性研讨信息的处理与融合就显得非常重要。

针对此问题,早期文献^[1]中被称为 IBIS(Issue-based Information System)的模型强调以议题为中心,构建了适用于逻辑推理与辩论的信息模型。虽然重视了研讨信息量化的结构化与自动化的推理过程,但忽视了专家的主动性和灵活性,没有反映对不确定性研讨信息的处理。文献^[2,3]通过在实体数据模型中定义基本属性来结构化研讨信息,借助关联函数实现了研讨信息之间的联系。但过分追求量化的建模对

以提案为中心的研讨不具有一般性,给出的关联函数不能完全刻画出不确定性信息之间的定量关系。文献^[4]通过整理专家发言对方案的偏好信息,构建了发言语义关系模型,采用语义距离的方法定量计算了群体共识度,但所定义的语义关系不够全面且不准确,导致用该方法计算不确定性信息的范围狭窄。文献^[5,6]针对群体研讨中定性信息的模糊性,试图建立一个具有统一标准的信息处理模型。该模型根据专家偏好对专家发言进行分类,然后进行多粒度的量化,最后用计算公式得到群体共识程度。但方法中多粒度量化的过程只限于给定的域值,缺乏更一般性的考虑。文献^[7-9]分别根据研讨专家的偏好建立了各自的定量信息群体研讨支持模型,但是由于专家个体的差异,其偏好呈现出多样性的特点,模型本身只适合于处理特定环境下的问题,缺乏对原始不确定研讨信息进行结构化而得到专家偏好的研究。文献^[10]给出了一种面向群体共识涌现的研讨信息组织模型,将专家发言分为根据、论证、模态限定等要素,其中的共识涌现图是主持人基于定量方法对方案进行筛选的一种图形化表示。虽然模型中的模态限定要素能一定程度上将不确定研讨信息量化,但群体共识达成局限于一般群决策方法,没有方法上的创新。

上述研究多是针对某个特定的研讨问题和阶段试图实现不确定研讨信息的量化,大部分强调的是信息处理方对不确定信息的识别与处理,忽视了信息提供方的主动性和智慧

的不可替代性的特点,并且群体共识多以群决策方法为主,忽略了不确定性研讨信息的融合过程。本文在借鉴国内外相关研究工作的基础上,根据专家发言信息不确定性的特点提出了一个由自然属性与人工属性两方面集成的研讨信息模型,并提出了主张确定值、论点框架、支持分配函数等概念,然后给出了发言信息提取的机制,随后研究了研讨信息进一步融合与结论生成的方法,最后结合汽车产品的客户需求研讨,给出了一个不确定性研讨信息结构化、信息提取、融合与结论生成的案例。

2 研讨模型的建立

一般地,从研讨议题的提出到结论的生成要经过议题提出、时间地点的确定、专家邀请、专家发言、发言记录与整理、意见综合和结论生成等多个阶段,其中参与研讨的角色包括管理员、主持人和专家等^[11]。据此,一个一般性的研讨模型可定义为

$$\text{Discussion} ::= \langle \text{type} \rangle \langle \text{time} \rangle \langle \text{place} \rangle \langle \text{master} \rangle \langle \text{dtype} \rangle \\ \langle \text{process} \rangle \langle \text{goal} \rangle \langle \text{restriction} \rangle \langle \text{information} \rangle \langle \text{conclusion} \rangle$$

式中, type、time、place、masters、dtype、process、goal、restriction、information 和 conclusion 分别表示议题类型、研讨时间、研讨地点、参与专家、研讨方式(同步、异步或两者结合)、研讨流程、研讨目标、约束条件、研讨中的信息和研讨结论。

一个典型的研讨流程如图 1 所示。主持人首先提出议题,确定议题类型、研讨地点、参与专家、研讨目标和约束条件等;然后,管理员根据主持人的要求布置研讨环境、研讨规则和会场设备等;再后,专家进入研讨厅,自由发表看法和广泛交流意见;最后,由主持人采用定性、定量方法对整理后的专家发言进行综合,得到研讨结论(conclusion)。此过程中,研讨的信息带有不确定性,不适合用自动化的方式对其进行识别和整理。

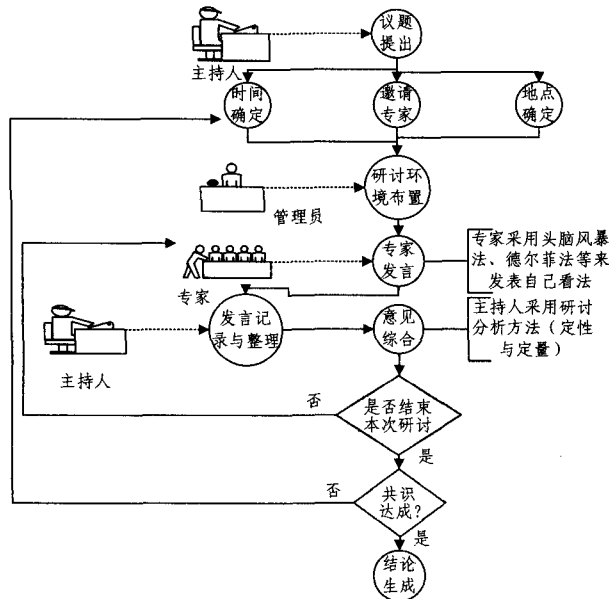


图1 一般研讨流程图

2.1 研讨信息模型

研讨信息的不确定性表现为多个方面:专家观点的主观

性;针对同一个问题、同一个专家给出两个不同的看法,导致专家的观点没有唯一性;针对同一个问题、不同的专家给出不同的看法,导致专家意见没有一致性。这些不确定性因素均给研讨结论的生成带来了困难。为了避免信息间冲突,得到较为稳定的一致性结果,因此研讨信息模型是信息处理与融合的基础。

早期英国哲学家 Toulmin 提出了一个由主张、根据、论证、支援、模态限定、反驳等概念构成的法律学模型 Toulmin^[12],但由于其复杂法律辩论的背景,并且模型本身没有考虑信息的不确定性和不完全性,因此不能直接用于不确定性研讨信息。这里,在 Toulmin 模型的基础上引入专家主观性量度的元素,定义一种由自然属性 Word1 与人工属性 Word2 集成的研讨信息模型,其数据结构如图 2 所示,其 BNF 范式如下:

$$\text{Word} ::= \langle \text{Word1} \rangle \langle \text{Word2} \rangle$$

自然属性是指不需要人工提取的专家原始发言信息,这些信息属于确定性信息。其范式表达如下:

$$\text{Word1} ::= \{ \text{id}, \text{person}, \text{time}, \text{content} \}$$

式中, id 为发言标识,用于标识专家发言的唯一性; person, time, content 分别表示“发言人”、“发言时间”、“发言内容”。

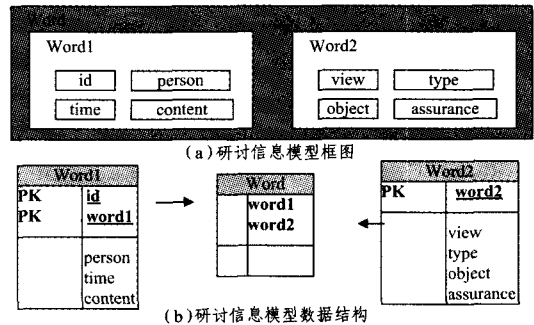


图2 研讨信模型框图与数据结构

人工属性是指需要由主持人从专家原始发言中识别和提取的属性。其范式表达如下

$$\text{Word2} ::= \{ \text{view}, \text{type}, \text{object}, \text{assurance} \}$$

式中, view, type, object, assurance 分别表示“观点”、“观点类型”、“发言对象”、“主张确定值”。

观点(view)指的是从原始发言中提取的中心思想。观点类型(type)分为论点和论据两类,论点指一个疑难问题供其他专家进行辩论,论据指一些事实或者数据等用来支持论点。人工属性中的主张确定值(assurance)是研讨信息的不确定性量度,其取值位于 0 到 1 之间,值越大表明信息的确定程度越大。需要指出的是主张确定值一般由专家在发表看法时给出,当专家没有给定的情况下由主持人根据专家的权威性给定。

以上研讨信息模型对信息进行了结构化。为了更进一步融合研讨信息,还需要对信息进行辨识与处理。

2.2 研讨信息提取机制

研讨信息提取是对专家自由发表信息的进一步归类 and 逻辑化整理,目的在于理清信息的关系,将相关信息纳入同一个共识问题中,以便达成信息融合与群体共识。研讨信息提取的机制包括提取论点框架和构造支持分配函数。提取论点框

架的过程是对议题进行分解的过程。

定义 1 假设一个研讨议题,对于该议题所能认识到的所有可能的论点的完备集合用 θ 来表示,且 θ 中所有元素都相互排斥,则称此集合 θ 为论点框架,可表示为 $\theta = \{a_1, a_2, \dots, a_i, \dots, a_n\}$ 。式中 a_n 称为论点框架 θ 的一个论点或元素; n 是元素的个数; $j=1, 2, \dots, n$ 。由论点框架 θ 的所有子集组成的一个集合称为 θ 的幂集,记为 2^θ 。

从专家发言中整理出的每一条论据用数学的方法可表示为论点框架 θ 的一个子集 A ,称为有效论点组。设论点框架 $\theta = \{a_1, a_2, \dots, a_i, \dots, a_n\}$,专家 i 给出的一个论据可表示为 $A_i = \{a_1, a_2, \dots\}$, k 个专家给出的多个论据对应的有效论点组为 $\{\{A_{11}, A_{12}, \dots\}, \{A_{21}, A_{22}, \dots\}, \dots, \{A_{k1}, A_{k2}, \dots\}\}$ 。

一个论点为真的程度完全由专家决定,是人主观想象的结果。在一组给定的论据与一个给定的论点之间没有客观联系能确定一个精确的支持度。但是对于一个论点,专家可以给出一个数字来表示他本人给出的论据组支持论点的程度。由此定义支持分配函数 m ,用于表示专家给出的论据组对论点的支持程度。

定义 2 设 θ 是论点框架, A 表示论点框架 θ 的任一子集,记 $A \subseteq \theta$,且满足

$$\begin{cases} m(\varphi) = 0 \\ \sum_{A \subseteq \theta} m(A) = 1 \end{cases}$$

式中, $m(A)$ 称为对 A 的支持分配函数,其值表示对 A 的支持度。该函数刻画了某一个或者一些论据对 A 的支持程度。为了描述所有论据对 A 的支持程度,定义支持函数 $S(A)$ 的概念。

定义 3 设 θ 是论点框架,如果 A 表示论点框架 θ 的任一子集, B 又是 A 的任一子集,记作 $B \subseteq A \subseteq \theta$,且满足

$$S(A) = \sum_{B \subseteq A} m(B)$$

式中, $S(A)$ 称为所有论据对 A 的支持函数,其值表示所有论据对 A 的支持度。该定义意味着 B 在逻辑上蕴涵于 A ,所以对于 A 总的支持度是对 B 的支持度的总和。

对研讨信息进行辨识,提取论点框架与构造支持分配函数的流程如图 3 所示。

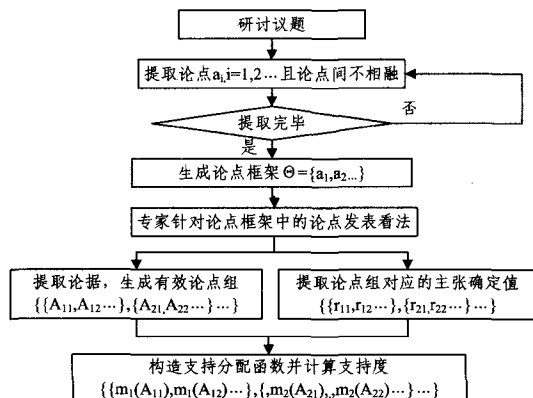


图 3 论点框架与支持分配函数构造流程图

图 3 中的步骤如下

Step1 对研讨议题进行分解,提取论点 a_i 。

Step2 论点是否提取完毕。如果提取完毕,则生成论点框架 $\theta = \{a_1, a_2, a_3, \dots\}$;如果没有提取完毕,则反复提取。

Step3 专家针对论点框架自由发表看法。

Step4 根据专家发言提取论据,生成有效论据组 $\{\{A_{11}, A_{12}, \dots\}, \{A_{21}, A_{22}, \dots\}, \dots\}$ 。提取对应的主张确定值 $\{\{r_{11}, r_{12}, \dots\}, \{r_{21}, r_{22}, \dots\}, \dots\}$ 。

Step5 构造支持分配函数并计算支持度。

3 研讨信息融合与结论生成

研讨信息融合与结论生成包括构造支持度矩阵、求证据的加权平均支持度、构造修正后的证据支持度矩阵、信息融合和生成二元组等步骤。其目的是得到论点框架中所有论点的可信度,由阈值去筛选论点从而产生最终方案。论点的可信度反映的是论点的可信程度。一般来说,一个论点被其他论据支持的程度越高,则论点的可信程度越大。

设某一个研讨议题,其论点框架 $\theta = \{a_1, a_2, a_3, \dots\}$,参与研讨的有 n 个专家 $C = \{c_1, c_2, \dots, c_n\}$,由 n 个专家给出的论据所建立的有效论点组为 $\{A_{11}, A_{12}, \dots, A_{21}, A_{22}, \dots, A_{n1}, \dots, A_{nk}\}$,对应的主张确定值为 $\{r_{11}, r_{12}, \dots, r_{21}, r_{22}, \dots, r_{n1}, \dots, r_{nk}\}$ 。其信息融合与结论生成的过程如下。

步骤 1 设 n 个专家给出了 k 个有效论点组,则存在 n 个支持分配函数和 $n \times k$ 个支持度。根据专家给出的主张确定值,用下面的公式对每个专家的主张确定值进行再分配。

$$m_i(A_{ij}) = r_{ij} = \frac{r_{ij}}{\sum_{j=1}^k r_{ij}}$$

式中, $i=1, 2, \dots, n$ 。由此构造支持度矩阵:

$$m = \begin{pmatrix} m_1(A_{11}) & m_1(A_{12}) & \dots & m_1(A_{1k}) \\ m_2(A_{21}) & m_2(A_{22}) & \dots & m_2(A_{2k}) \\ \vdots & \vdots & \dots & \vdots \\ m_n(A_{n1}) & m_n(A_{n2}) & \dots & m_n(A_{nk}) \end{pmatrix}$$

步骤 2 n 个专家由于具有不同的权威性,赋予权重系数 x_i 且满足 $\sum_{i=1}^n x_i = 1$ 。论据的加权平均支持度 $\bar{h}_i = \sum_{j=1}^k x_j m_i(A_{ij})$,式中 $j=1, 2, \dots, k$ 。论据的加权平均支持度 $\bar{h} = (\bar{h}_1, \bar{h}_2, \dots, \bar{h}_k)$ 客观上描述了论据支持的平均程度,可以用来修订各专家的支持度,使之向平均支持程度靠近。

步骤 3 用步骤 2 的结果求与加权平均支持度的偏差 $\bar{m}_i(A_{ij}) = m_i(A_{ij}) - \bar{h}_j$,式中 $i=1, 2, \dots, n; j=1, 2, \dots, k$ 。与加权平均支持度偏差的矩阵为

$$\bar{m} = \begin{pmatrix} \bar{m}_1(A_{11}) & \bar{m}_1(A_{12}) & \dots & \bar{m}_1(A_{1k}) \\ \bar{m}_2(A_{21}) & \bar{m}_2(A_{22}) & \dots & \bar{m}_2(A_{2k}) \\ \vdots & \vdots & \dots & \vdots \\ \bar{m}_n(A_{n1}) & \bar{m}_n(A_{n2}) & \dots & \bar{m}_n(A_{nk}) \end{pmatrix}$$

用公式 $e_i = \sum_{j=1}^k \frac{\bar{m}_i^2(A_{ij})}{k}$,式中 $j=1, 2, \dots, k$,求得偏差 $E = (e_1, e_2, \dots, e_n)$ 。

步骤 4 用偏差 E 修正原始的支持度,得到修正后的证据支持度 $m_i^c(A_{ij}) = m_i(A_{ij}) - x_i e_i$,式中 x_i 为专家权重; $i=1, 2, \dots, n; j=1, 2, \dots, k$ 。修正后的论据支持度矩阵为

$$m^c = \begin{pmatrix} m_1^c(A_{11}) & m_1^c(A_{12}) & \dots & m_1^c(A_{1k}) \\ m_2^c(A_{21}) & m_2^c(A_{22}) & \dots & m_2^c(A_{2k}) \\ \vdots & \vdots & \dots & \vdots \\ m_n^c(A_{n1}) & m_n^c(A_{n2}) & \dots & m_n^c(A_{nk}) \end{pmatrix}$$

从以上计算的过程来看,专家权重越大,该专家的支持度与平均证据支持度之间的距离越小,其调整量越小;当专家权重越小时,该专家的支持度与平均证据支持度之间的距离越大,其调整量越大。通过与平均证据比较后的修正最终减小了证据之间的冲突程度。

步骤5 将调整后的 $n \times k$ 个证据进行信息融合,得到综合支持度 $p(B_j)$ 。

$$p(B_j) = \begin{cases} \frac{\sum_{\cap A_{ij}=B_j, 1 \leq j \leq k} \prod m_i^c(A_{ij})}{1-e}, & B_j \neq \varphi \\ 0, & B_j = \varphi \end{cases}$$

式中, $i=1, 2, \dots, n, e = \sum_{\cap A_{ij}=\varphi, 1 \leq j \leq k} \prod m_i^c(A_{ij})$ 反映的是论据的总体冲突程度,其值越大则表明证据间的冲突越大。

构造二元组 $\{(B_1, p(B_1)), (B_2, p(B_2)), \dots, (B_k, p(B_k))\}$ 表示各论据对应的支持度。

步骤6 计算论点框架 $\theta = \{a_1, a_2, \dots, a_j\}$ 中单个论点的支持度 $l(a_j) = \sum_{a_j \subset A} p(B_j)$, 式中 $j=1, 2, \dots, k$ 。由论点支持度得到论点的可信度为

$$u(a_j) = \frac{l(a_j)}{\sum_{j=1}^k a_j}$$

综合论点框架中所有论点的可信度,得到 $U = (u(a_1), u(a_2), \dots, u(a_k))$ 。取阈值筛选论点,得到满足要求的集合 $\{a_1, a_2, \dots, a_z\}$, 其中 $z < k$, 该论点集合则为研讨的最终结果。

4 应用实例

某品牌汽车公司为了满足客户需求,进一步占领市场,获取更多的利润,特需要拟订未来产品的发展方向。专家首先通过研讨拟订可能满足客户需求的所有技术指标,如表1所列。该研讨过程的实质是将问题进行分解,用分解后的多个论点去描述研讨议题,表1中所有的工程技术指标组成论点框架。专家在下一轮研讨中针对这些论点进行讨论,根据个人的知识经验自由地选择所支持的论点并给出论据。

表1 工程技术指标(论点框架)

代码	工程技术指标	代码	工程技术指标
a ₁	耗油量	a ₅	牵引力
a ₂	最大速度	a ₆	舒适度等级
a ₃	发动机功率	a ₇	质量等级
a ₄	价格	a ₈	安全等级

3位专家针对表1中的论点框架(工程技术指标)提出个人看法如下(研讨信息的自然属性已经整理):

专家1(2010 Mon Sep 28 15:01:05):观点1:我认为耗油量、最大速度、发动机功率、价格是公司最需要解决的工程技术指标,该提法的把握为0.5。观点2:耗油量、最大速度、牵引力、舒适度、质量等级是客户最想要的,我们公司应该着重关注这方面的问题,该提法的准确性为0.7。

专家2(2010 Mon Sep 28 15:03:05):观点1:我非常同意专家1所提出的第2个观点,我们应该就这几个方面进行技术改进,以满足客户的需求,提高公司的销售额,我相信这个观点的准确性为0.7。观点2:除了油耗和价格外,其他因素都必须首先考虑。汽车的价格由我们的成本决定,质量好的汽车价格肯定会高,我们不会去降低成本生产质量差的汽

车。我公司的汽车的耗油量已经是同类型产品最小的了,该观点的准确性为0.6。

专家3(2010 Mon Sep 28 15:07:05):观点1:专家2第2个观点说得对,我们公司汽车的耗油量已经非常小了,准确性为0.8。观点2:我针对专家2的第2个观点谈谈我个人看法,我认为汽车价格还是有调整的空间,毕竟现在市场上的汽车价格已经在逐年降低了,我这个观点的准确性为0.5。

由以上专家发言可整理出的有效论点组为 $\{A_1, A_2, A_3, A_4\}$, 对应的主张确定值为 $\{0.5, 0.7, 0.7, 0.6, 0.8, 0.5\}$, 表1中的论点框架为 $\theta = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ 。根据专家知识背景分别设各专家的权重为 $\{0.2, 0.3, 0.5\}$, m_1, m_2, m_3 分别表示3位专家的支持分配函数,详细信息如表2所列。

表2 提取后的研讨信息表

专家	权重	工程技术特点组合(论据)	主张确定值
专家1	0.2	A ₁ = {a ₁ , a ₂ , a ₃ , a ₄ }	0.5
		A ₂ = {a ₁ , a ₂ , a ₅ , a ₆ , a ₇ }	0.7
专家2	0.3	A ₂ = {a ₁ , a ₂ , a ₅ , a ₆ , a ₇ }	0.7
		A ₃ = {a ₂ , a ₃ , a ₅ , a ₆ , a ₇ , a ₈ }	0.6
专家3	0.5	A ₃ = {a ₂ , a ₃ , a ₅ , a ₆ , a ₇ , a ₈ }	0.8
		A ₄ = {a ₁ , a ₂ , a ₃ , a ₅ , a ₆ , a ₇ , a ₈ }	0.5

据步骤1中的公式对各专家给出的论据的支持度进行再分配,得到支持分配函数与其值如下

$$m_1(A_1) = 0.42, m_1(A_2) = 0.59, m_2(A_2) = 0.54$$

$$m_2(A_3) = 0.46, m_3(A_3) = 0.62, m_3(A_4) = 0.38$$

由此构造支持分配矩阵

$$m = \begin{bmatrix} 0.42 & 0.59 & 0 & 0 \\ 0 & 0.54 & 0.46 & 0 \\ 0 & 0 & 0.64 & 0.38 \end{bmatrix}$$

由步骤2计算加权平均支持度

$$\bar{h} = [0.084 \quad 0.28 \quad 0.448 \quad 0.19]$$

由步骤3计算与加权平均支持度的偏差矩阵

$$\bar{m} = \begin{bmatrix} 0.336 & 0.31 & -0.448 & -0.19 \\ -0.084 & 0.26 & -0.012 & -0.19 \\ -0.084 & -0.28 & 0.172 & 0.19 \end{bmatrix}$$

从而得到偏差 $E = [0.111 \quad 0.0277 \quad 0.0379]$ 。

由步骤4得到修正后的支持度矩阵

$$m^c = \begin{bmatrix} 0.386 & 0.581 & 0.05 & 0.007 \\ 0.002 & 0.533 & 0.46 & 0.005 \\ 0.003 & 0.008 & 0.614 & 0.373 \end{bmatrix}$$

如果对原始支持度矩阵进行融合,得到的单个指标的可信度将会出现冲突。而经过修正后的支持度已经向平均论据的支持度靠近,一定程度上消除了论据间的冲突。

由步骤5中的公式进行信息融合,得到综合支持度 $P(A) = [0.0005, 0.2051, 0.7543, 0.0312]$, 从而构建二元组 $\{(A_1, 0.0005)(A_2, 0.2051)(A_3, 0.7543)(A_4, 0.0412)\}$ 。

由步骤6中的公式求得单个指标的支持度 $L = [0.2513, 0.9956, 0.2213, 0.0005, 0.9951, 0.9951, 0.9951, 0.79]$, 归一化得单个指标的可信度:

$$U = [0.048, 0.19, 0.042, 0, 0.19, 0.19, 0.19, 0.15]$$

取可信度的阈值为0.1, 筛选得到满足条件的论点集合为 $\{a_2, a_5, a_6, a_7, a_8\}$, 相应的工程技术指标为最大速度、牵引

力、舒适度等级、质量等级、安全等级。

如果单是结合专家权重直接对专家的主张值进行加权求和可以得到这样的结论:汽车的“耗油量”是客户关心的第一工程技术指标。产生这个结论的原因主要是专家3的权重和主张确定值相对较大,由此可以看出此方法给结论带来了不稳定和不准确性。

通过案例的计算、比较与分析可以看出,本文提出的研讨信息模型将不确定性研讨信息进行了结构化与量化,从中提取到了论点框架、有效论点组并构造了支持分配函数,应用信息融合的技术对信息进行综合,得到了论点的可信度,从而生成了结论。为了消除论据冲突,采用论据支持度重新分配与平均证据修正的方法,一定程度上增强了信息融合的可靠性。

结束语 群体研讨中,专家个体因为有限理性、知识背景与经验的不同,对问题的思考会存在局限性或者差异性,因而研讨信息呈现出不确定性的特点。本文针对此特点提出了自然属性与人工属性集成的研讨信息模型。该模型中的主张支持值将专家的主观性进行了量化,通过构造支持分配函数来表示专家对论点的支持程度。从不确定性研讨信息中提取论点框架和有效论点组,为信息的融合提供了数据结构的基础。在信息融合方面,采用论据支持度重新分配方法,并结合平均证据去修正原始支持度,一定程度上缩小了论据间的冲突,增强了信息融合的稳定性和可靠性,促进了共识结论的生成。后续将进一步研究引起信息不确定性的其他因素,扩展模型结构,提高模型的处理能力。

参 考 文 献

[1] Kunz W, Rittel W J H. Issues as elements of information systems[R]. Berkeley, CA: University of California, 1970
[2] Hemant K B, Daniel J P, Daewon S. Progress in Web-based decision support technologies[J]. Decision Support Systems, 2007 (43):1083-1095

(上接第 156 页)

从表 2 中可得

$$MRR_{IRD} > MRR_{Borda}$$

由 MRR 查准率的特性可以得出,IRD 排序算法明显要优于 Borda 算法。

结束语 元搜索引擎的核心研究在于对从成员搜索引擎中获得的结果进行重新排序。以往的排序算法在计算上有一定的缺陷,在研究以往算法并集合多种算法后提出修正后的 IRD 算法,经实验结果表明,其较以前的排序算法更有优越性,能有效提高排序结果的效率;将用户期望的信息排在前面。下一步将结合历史值调整计算结果的相关度^[7]。

参 考 文 献

[1] Selberg E, Etzioni O. Multi-service search and comparison using the MetaCrawler [C] // Proceedings of the 4th International World Wide Web Conference. December 1995
[2] Hsu D F, Taksa I. Comparing Rank and Score Combination

[3] Ohbyung K, Kyoung Y K, Kun C L. MM-DSS: Integrating multimedia and decision-making knowledge in decision support systems[J]. Expert Systems with Applications, 2007(32):441-457
[4] 李嘉,张朋柱. 群体研讨支持系统中研讨主题的自动可视化聚类研究[J]. 系统管理报, 2009, 18(3):325-331
[5] Alonso S, Cabrerizo F J, Chiclana F, et al. A Web based consensus support system for group decision making problems and incomplete preferences [J]. Information Sciences, 2010(180): 4477-4495
[6] Alonso S, Viedma E H, Chiclana F, et al. Individual and social strategies to deal with ignorance situations in multi-person decision making[J]. International Journal of Information Technology and Decision Making, 2009, 8(2):313-333
[7] Cabrerizo F J, Alonso S, Viedma E H. A consensus model for group decision making problems with unbalanced fuzzy linguistic information[J]. International Journal of Information Technology and Decision Making, 2009, 8(1):109-131
[8] Chiclana F, Mata F, Martnez L, et al. Integration of a consistency control module within a consensus decision making model [J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2008, 16(1):35-53
[9] Mata F, Mart L, Viedma E H. An adaptative consensus support model for group decision making problems in a multi-granular fuzzylinguistic context [J]. IEEE Transactions on Fuzzy Systems, 2009, 17(2):279-290
[10] Chen Y J. Development of a method for ontology-based empirical knowledge representation and reasoning [J]. Decision Support Systems, 2010, 50(1):1-20
[11] 李德华,熊才权. 一种研讨信息组织模型及其在研讨厅中的应用 [J]. 计算机应用研究, 2008, 25(9):2730-2733
[12] Toulmin S. The uses of argument[M]. New York: Cambridge Univemity, 1958:1-10

Methods for Data Fusion in Information Retrieval [J]. Information Retrieval, 2005, 8(3):449-480

[3] Shimotsuruma, Yamato S, Kanagawa K. Information Retrieval on the Web [J]. ACM Computing Surveys, 2000, 32(2):144-171
[4] 曹林. 元搜索引擎排序技术研究 [D]. 南京: 河海大学, 2009
[5] 樊昌信, 曹丽娜. 通信原理(第六版)[M]. 北京: 国防工业出版社, 2008
[6] Mizzaro S. Relevance: The whole history [J]. Journal of the American Society for Information Science, 1997, 48(9):810-832
[7] Wu Sheng-li, Bi Ya-xin, Zeng Xiao-qin. Retrieval Result Presentation and Evaluation [J]. Lecture Notes in Computer Science, 2010, 6291:125-136
[8] Martl'nez-Santiago, Ureña-López L, Martl'n-Valdivia M. A merging strategy proposal: The 2-step retrieval status value method [J]. Information Retrieval, 2006, 9(1):71-93
[9] Wu Sheng-li. Assigning appropriate weights for the linear combination data fusion method in information retrieval [J]. Expert Systems with Applications, 2009, 36(2):2997-3006