

# 基于证据间距离和不确定性度量的证据组合方法

夏文俊<sup>1,2</sup>, 朱林户<sup>1</sup>, 黄邵军<sup>1</sup>, 吕中凯<sup>1</sup>

(1. 空军工程大学 理学院, 陕西 西安 710051; 2. 95907 部队, 甘肃 酒泉 732750)

**摘 要:**针对 Dempster 组合公式无法组合冲突证据的问题,提出了一种证据组合的权重分配方法。该方法充分考虑证据间的关系和证据本身的特性,用证据间距离度量证据间的不一致程度,用证据的不确定度来度量证据本身的不确定性;在此基础上扩展了文献[7]提出的权重确定准则,认为证据组合规则既要考虑使组合后证据与各源证据间的距离和尽量小,也要注重降低组合后证据本身的不确定性。最后根据新的准则给出了权重因子的确定算法和证据组合方法。算例表明,该方法改进了文献[7]权重分配方法的结果,且使权重分配更加灵活。

**关键词:**D-S 证据理论;证据间距离;不确定性度量;信息融合

**DOI:**10.3969/j.issn.1009-3516.2010.02.020

**中图分类号:**TP274;O21 **文献标识码:**A **文章编号:**1009-3516(2010)02-0091-04

证据理论在信息融合领域中应用广泛。1984 年 Zadeh 首次指出了 Dempster 组合规则的缺陷<sup>[1]</sup>,此后对改进证据理论的研究就层出不穷<sup>[2-6]</sup>。证据间的距离被用来度量证据的一致程度,证据的不确定性测度用来度量证据本身的不一致性和不具体性,它们能反映证据间的差异和证据的性质,因此可以被用来研究修改 Dempster 组合规则的方法。修改 Dempster 组合规则可以视作按一定权重把冲突证据分配给各焦元。Lefevre 等人提出了一种权重分配方法的统一表示公式<sup>[2]</sup>,Lefevre、Jsang<sup>[5]</sup>、吴根秀<sup>[7]</sup>分别给出了计算权重的方法。本文在研究了证据间的距离和证据的不确定性度量的基础上给出了一种权重分配方法,改进了文献[7]中的结果,使权重分配更加灵活。

## 1 冲突证据的权重分配和证据组合规则的改进

限于篇幅,证据理论的基础知识在此不再赘述,读者可参见文献[6]。先回顾一下 Zadeh 反例:

例 1 设识别框架  $\Theta = \{A, B, C\}$ , 2 个 BPA 如下:

$$m_1(A) = 0.99, m_1(B) = 0.01, m_1(C) = 0; m_2(A) = 0, m_2(B) = 0.01, m_2(C) = 0.99;$$

利用 Dempster 组合规则有:  $m(A) = 0, m(B) = 1, m(C) = 0$ 。这明显违背了我们的直觉。

这种结果是由于归一化引起的。我们来分析一下归一化的过程,由于  $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ ,  $1 - K =$

$$\sum_{B \cap C \neq \emptyset} m_1(B)m_2(C), \text{ 所以 } m(A) = \frac{1}{1 - K} \sum_{B \cap C = A} m_1(B)m_2(C) = \sum_{B \cap C = A} m_1(B)m_2(C) + \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{\sum_{B \cap C \neq \emptyset} m_1(B)m_2(C)} K。$$

可见,  $\forall A \subseteq \Theta$ , 其 BPA 可以看成 2 部分,一部分是完全支持 A 的部分,另一部分是冲突中分配给 A 的部分。Dempster 组合规则的归一化其实是将冲突的证据按一定的比例分配给各焦元。基于这一认识,很多学者

\* 收稿日期:2008-12-22

基金项目:陕西省自然科学基金资助项目(SJ08F10)

作者简介:夏文俊(1984-),男,湖北黄冈人,硕士生,主要从事信息融合研究;E-mail:xwj1128@126.com

朱林户(1950-),男,陕西武功人,教授,博士生导师,主要从事概率度量空间及其在应用、随机集及其在信息融合中的应用研究。

提出了基于权重分配的改进组合规则。2002年 Lefevre 等人<sup>[3]</sup>提出了一种“统一信度函数组合方法”:

$$m(A) = m_{\cap}(A) + \omega(A, m)K \quad (1)$$

式中,  $m_{\cap}(A) = \sum_{B \cap C = A} m_1(B)m_2(C)$ ,  $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ ,  $\omega(A, m)$  为冲突部分分配给  $A$  的权重系数, 满足  $\sum_{A \in \Theta} \omega(A, m) = 1$ 。

统一信度组合方法是 Smets<sup>[4]</sup>、Yager<sup>[5]</sup>、Dempster 的组合规则的统一表达式<sup>[6]</sup>。

Lefevre 的公式是权重分配组合方法的一般表示形式, 关键问题在于如何确定权重因子  $\omega(A, m)$ 。Lefevre 认为可以采用专家主观确定的方法, 但这种方法很难在实际问题中应用, 因为专家也很难得到较好的权重因子。Lefevre 也给出了利用梯度下降法自适应得到权重因子的方法, 这种可行的方法缺点是需要大量数据训练以得到权重因子。Jsang 提出了权重因子的平均算子法<sup>[5]</sup>, 是一种比较粗略的确定权重因子的方法。吴根秀在引入自冲突概念后从最优化的观点出发提出了一种确定权重因子的方法<sup>[7]</sup>, 得到了较好的结果。

## 2 证据的不确定性度量 and 证据间的距离

一般认为, 不确定性证据分为模糊性(fuzziness)和不明性(ambiguity)2类<sup>[8]</sup>, 本文不考虑模糊证据。证据的不明确性作为一种普通概念是 Klir 等人于 1995 年提出来的<sup>[9]</sup>, 它包括不一致性(discord)和不具体性(nonspecificity)2个方面<sup>[6]</sup>。

$$\text{不具体性: } N(m) = \sum_{A \in \Theta} m(A) \log_2 |A|;$$

$$\text{不一致性: } ST(m) = - \sum_{A \in \Theta} m(A) \log_2 \sum_{B \in \Theta} m(B) \frac{|A \cap B|}{|A|};$$

$$\text{总不确定性为两者之和 } T(m) = N(m) + ST(m) = - \sum_{A \in \Theta} m(A) \log_2 \sum_{B \in \Theta} m(B) \frac{|A \cap B|}{|A|^2};$$

当焦元  $A$  是单点集时, 总不确定性即为不一致性, 且退化为  $T(m) = - \sum_{A \in \Theta} m(A) \log_2 m(A)$ 。

证据的不确定性度量的是证据本身的不具体性和不一致性, 反映的是专家提供的某个证据本身的性质; 而证据之间的距离描述的是证据之间的一致程度。Jousselme 等人研究了 Fixen 和 Mahler 定义的距离<sup>[10]</sup>的非退化性和伪距离性质, 给出了一个性质较好的距离<sup>[11-12]</sup>:

$$d^2(m_1, m_2) = \frac{1}{2} (\langle m_1, m_1 \rangle + \langle m_2, m_2 \rangle - 2 \langle m_1, m_2 \rangle) \quad (2)$$

式中,  $\langle m_1, m_2 \rangle = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} m_1(A_i) m_2(B_j) \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$ ,  $A_i, B_j \in 2^\Theta$ ,  $i, j = 1, 2, \dots, 2^n$ 。

## 3 一种新的权重分配方法

吴根秀<sup>[7]</sup>引入了自冲突的概念, 定义了自冲突强度  $m^*(\emptyset) = \sum_{B \cap C = \emptyset} m(B)m(C)$ , 分析了自冲突强度对组合规则的影响, 并在此基础上提出了权重因子确定的 2 条准则: ①组合后的 BPA 与各证据源的 BPA 的距离和应尽可能小; ②组合后的 BPA 的自冲突强度应尽可能的小。基于此, 文献[7]给出了一种权重因子的线性规划确定方法。

本文认为, 准则②可以一般性地表述为: 组合后的 BPA 的不确定性应尽可能小。事实上, 文献[7]定义的自冲突强度并没有完全反映证据的不确定性程度, 例如:

例 2 识别框架  $\Theta = \{\theta_1, \theta_2, \theta_3\}$ , 事件  $A = \{\theta_1\}$ ,  $B = \{\theta_1, \theta_2\}$ , 已知的 BPA 为  $m(A) = 0.3$ ,  $m(B) = 0.7$ 。按文献[7]的定义, 自冲突强度  $m^*(\emptyset) = 0$ , 然而, 给定的 BPA 显然具有不确定性, 其不确定度为  $T(m) \approx 0.1641$ 。

根据改进后的确定权值的准则, 确定权重  $\omega(A, m)$  的问题变为:

$$\min \begin{pmatrix} D \\ T \end{pmatrix}$$

$$\text{s. t.} \begin{cases} \sum_{A \subseteq \Theta} \omega(A, m) = 1 \\ \omega(A, m) \geq 0, \forall A \subseteq \Theta, \end{cases}$$

式中: $D = \sum_{i=1}^k d^2(m, m_i)$  为组合后 BPA 与各证据源 BPA 的距离平方和,表示证据间的不一致程度,称之为不一致度, $k$  为证据数; $T = - \sum_{A \subseteq \Theta} m(A) \log_2 \sum_{B \subseteq \Theta} m(B) \frac{|A \cap B|}{|A|^2}$  为组合后 BPA 的总不确定度。

由上述多目标优化模型得到权重  $\omega(A, m)$  后,就可以用统一信度组合方法来组合证据了。

### 4 算例

用上述方法计算 Zadeh 及例子(例 1) 算得  $m_{\cap}(A) = 0; m_{\cap}(B) = 0.0001; m_{\cap}(C) = 0; m_{\cap}(\emptyset) = 0.9999$ 。又有: $D = d^2(m, m_1) + d^2(m, m_2), T = - \sum_{A \subseteq \Theta} m(A) \log_2 m(A)$ ;且  $m(A) = m_{\cap}(A) + \omega(A, m)m_{\cap}(\emptyset)$ ,式中: $\omega(A, m) + \omega(B, m) + \omega(C, m) = 1; \omega(A, m), \omega(B, m), \omega(C, m) \geq 0$ 。

如果假设目标  $D$  和  $T$  同等重要,解这个多目标优化问题,得到  $\omega(A, m) = 0.5, \omega(B, m) = 0, \omega(C, m) = 0.5$ 。于是组合后的结果为  $m(A) = 0.49995, m(B) = 0.0001, m_2(C) = 0.49995$ 。这是符合人们的推理常识的结果。

本文给出的方法,其优势在于同时考虑了证据间的距离和证据本身的不确定度。考虑证据间的距离使组合后的证据与源证据距离尽量小,可以消解冲突证据权重对组合结果的影响;考虑证据本身的不确定度使组合证据时注重降低证据的不确定性,可以加快组合规则的收敛速度,较快得出识别结果。解多目标优化问题时可以灵活选择解法以得到更加符合实际情况的融合结果。具体来说,当证据间冲突较大时,可以将  $D$  作为主要目标求解,使组合后证据与组合前各证据更为接近,组合的结果更符合常理;当证据间冲突较小时,可以将  $T$  作为主要目标求解权值,注重消除不确定性,加快收敛速度。

例 3 识别框架  $\Theta = \{A, B, C\}$ ,4 个 BPA 分别如下:

$$m_1(A) = 0.5, m_1(B) = 0.2, m_1(C) = 0.3; m_2(A) = 0, m_2(B) = 0.9, m_2(C) = 0.1;$$

$$m_3(A) = 0.55, m_3(B) = 0.1, m_3(C) = 0.35; m_4(A) = 0.55, m_4(B) = 0.1, m_4(C) = 0.35;$$

算得: $m_{\cap}(A) = 0; m_{\cap}(B) = 0.0018; m_{\cap}(C) = 0.0037; m_{\cap}(\emptyset) = 0.9945$ 。

用 Dempster 组合规则、文献[7]的组合规则和本文的规则可得结果见表 1。

表 1 比较结果

Tab. 1 The comparison results

焦元	Dempster	文献[7]	本文				
			主要考虑消解冲突		适当考虑降低不确定度		
	$m$	$\omega$	$m$	$\omega$	$m$	$\omega$	$m$
A	0	0.448 2	0.445 7	0.504 4	0.501 6	0.559 1	0.556 0
B	0.33	0.319 3	0.319 3	0.312 3	0.312 4	0.324 2	0.324 2
C	0.67	0.232 5	0.234 9	0.183 3	0.186 0	0.116 7	0.119 8

从计算结果可以看出,例 3 中的情况,Dempster 组合规则无法得到符合常理的结果,文献[7]的方法等得到较好的组合结果,而本文提出的组合方法能得到最优的结果并且有较好的灵活性,可以根据实际情况对不一致度和不确定度的要求灵活调整。

### 5 结束语

Dempster 组合规则由于在处理高度冲突的证据时得到违背常理的组合结果一直倍受批评,Lefevre 等人

给出了冲突证据的权重分配算法,能较好地解决这一问题,如何确定冲突证据的权重是这类改进组合规则的关键问题。本文在研究证据间距离和证据的不确定性度量的基础上改进了文献[7]提出的权重分配原则,并给出了一种灵活性更大的权重分配算法,在实际计算中得到了较好的结果。

### 参考文献:

- [1] Zadeh L A. Review of Shafer's A Mathematical Theory of Evidence [J]. AI Mag, 1984, 5:81 - 83.
- [2] Lefevre E, Colot O, Vannoorenberghe P. Belief Functions Combination and Conflict Management [J]. Information Fusion, 2002, 3(2): 149 - 162.
- [3] Smets P, Kennes R. The Transferable Belief Model [J]. Artificial Intelligence, 1994, 66(2): 191 - 234.
- [4] Yager R R. On the Dempster - shafer Framework and New Combination Rules [J]. Information Sciences, 1987, 41: 93 - 138.
- [5] Sang A. The Consensus Operator for Combining Beliefs [J]. Artif Intell J, 2002, 142(1 - 2): 157 - 170.
- [6] 郭华伟,施文康,邓勇,等. 证据冲突: 丢弃,发现或化解[J]. 系统工程与电子技术, 2007, 29(6): 890 - 898.  
GUO Huawei, SHI Wenkang, DENG Yong. Evidential Conflict and Its 3D Strategy: Discard, Discover and Disassemble [J]. Systems Engineering and Electronics, 2007, 29(6): 890 - 898. (In Chinese)
- [7] 吴根秀. 冲突证据组合方法[J]. 计算机工程, 2005, 31(9): 151 - 154.  
WU Genxiu. An Approach for Combining Conflict Evidence [J]. Computer Engineering, 2005, 31(9): 151 - 154. (In Chinese)
- [8] Anne Laure Josselme, Chunsheng Liu, Dominic Grenier. Measuring Ambiguity in the Evidence Theory [J]. IEEE Transactions on Systems Man and Cyberneticspart A: Systems and Humans, 2006, 36(5): 890 - 903.
- [9] Klir G J, Yuan B. Fuzzy Sets and Fuzzy Logic: Theory and Applications [M]. Upper Saddle River NJ: Prentice - Hall, 1995.
- [10] Harmance D, Klir G J. Measuring Total Uncertainty in Dempster - Shafer Theory: A Novel Approach [J]. International Journal of General Systems, 1997, 22(4): 405 - 419.
- [11] Dale Fixen, Ronald P, Mahler S. The Modified Dempster - Shafer Approach to Classification [J]. IEEE, Transactions on Systems Man and Cybernetics - Part A: Systems and Humans, 1997, 27(1): 96 - 104.
- [12] Anne - Laure Josselme, Dominic Grenier, Eloi Bossé. A New Distance Between Two Bodies of Evidence [J]. Information Fusion, 2001, 2(3): 91 - 101.

(编辑:田新华)

## Evidence Combination Approach Based on Distance between Evidences and Ambiguity Measure

XIA Wen - jun<sup>1,2</sup>, ZHU Lin - hu<sup>1</sup>, HUANG Shao - jun<sup>1</sup>, LÜ Zhong - kai<sup>1</sup>

(1. Science Institute, Air Force Engineering University, Xi'an 710051, China; 2. Unit 95907, Jiuquan 732750, Gansu, China)

**Abstract:** In order to solve the problem that conflict evidences can not be combined with Dempster's combination rule, a weights redistribution approach to evidence combination is proposed. In this approach the disaccord of evidences and the ambiguity of evidence are measured by distance and ambiguity measure respectively. Based on this, this paper expands the weights determination principles proposed in [7]. Then, new weights determination algorithm and evidence combination approach are suggested, and the result in [7] is improved.

**Key words:** D - S evidence theory; distance between evidences; ambiguity measure; information fusion