

一种新的证据权重系数确定方法

许将军

(中国民航飞行学院航空工程学院 广汉 618307)

摘 要 分析现有的证据权重确定方式存在的缺陷,在此基础上提出了一种新的改进算法。利用各证据与平均证据的相似系数求出证据的初始权重,把权重低于阈值的证据从证据组中剔除,计算剩余证据的权重,经 K 次迭代,直到所有证据的权重都大于阈值,最后将 K 次权重算术平均得到各证据的最新权重系数,对证据加权平均后再利用 D-S 组合规则进行组合。通过算例比较表明该方法的合成结果更加有效。

关键词 冲突证据,平均证据,证据权重,组合规则

中图法分类号 TP18 文献标识码 A

New Method of Determining Evidences Weight Coefficient

XU Jiang-jun

(Aviation Engineering Institute, Civil Aviation Flight University of China, Guanghan 618307, China)

Abstract We analyzed the defects of the existing way to determine the weight of evidence, and gave a new modified method based on the present method. It calculated the evidence primary weight via similarity between evidence and average evidence. When the weight value is less than the threshold value, the evidence is removed from evidence group, and the weight of remaining evidence is calculated. By K iterations, until all the evidence weight is higher than the threshold value, arithmetic average value of K evidence weights is calculated. Finally, D-S combination rules is used to combine the weighted average evidence. An example indicates that the new method can give the more reasonable combination results compared with D-S combination rule and other modified methods.

Keywords Conflict evidence, Average evidence, Evidence weight, Combination rules

D-S 证据理论是由美国数学家 Dempster 提出的一套数学方法^[1-3],主要用来处理不确定性推理问题。由于在证据理论中需要的先验数据比概率推理理论中的更为直观、更容易获得,再加上 Dempster 合成公式可以综合不同专家或数据源的知识或数据,使得证据理论在专家系统、信息融合等领域得到广泛应用。但是,当证据冲突严重和完全冲突时,利用 D-S 证据理论合成会得到与直觉相悖的结果。针对这一问题,不同学者提出了不同的改进方法,本文在分析现有改进方法的基础上,提出了一种新的改进算法,实验证明该方法更加合理有效。

1 D-S 理论组合规则及存在的问题

定义 1 设 Θ 是一个识别框架,在识别框架 Θ 上的 BPA 是一个 $2^\Theta \rightarrow [0, 1]$ 的函数称为 mass 函数,并且满足:

$$m(\Phi) = 0 \text{ 且 } \sum_{A \subseteq \Theta} m(A) = 1$$

假定识别框架 Θ 下的两个证据 E_1 和 E_2 ,其相应的基本信任分配函数为 m_1 和 m_2 ,焦点分别为 A_i 和 B_j ,则 D-S 合成规则为:

$$m(A) = \begin{cases} \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - K}, & \forall A \subseteq U, A \neq \emptyset \\ 0, & A = \emptyset \end{cases} \quad (1)$$

其中,

$$K = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) \quad (2)$$

证据组合规则中, K 是冲突因子,反映了证据的冲突程度。当 $K=1$ 时,证据完全冲突,就不能使用 D-S 合成规则进行融合;当 K 接近 1 时,合成结果又往往与实际相悖。文献^[4]给出了具体实例。

现有两条证据如下:

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

$$m_2(A) = 0, m_2(B) = 0.01, m_2(C) = 0.09$$

根据式(2)得 $K=0.9999$,属于证据高度冲突,采用式(1)合成之后, $m(A)=0, m(B)=1, m(C)=0$,可见支持度最低的 B 经 D-S 合成规则合成后变为了 1,该结果与常理相悖。因此 D-S 证据理论不能正确处理证据高冲突问题。

2 现有的改进方法及其存在的不足

对冲突证据的处理方法有两种,1)修改合成规则;2)修改证据模型。文献^[5]通过比较分析认为改进模型的方法优于组合规则的改进。目前,大多研究人员对关于证据模型修正的研究都是以文献^[6]为基础,并在此基础上做了进一步扩展,引入证据权的概念,考虑了证据之间的相关性,根据证据间的支持度算出证据的权重,然后利用证据权重对各焦点的基本概率进行修正,最后对修正后的证据利用 D-S 组合规则进行合成,如文献^[7-9]。该方法存在两个缺点,

(1) 假设有 n 个证据, 要计算各个证据的两两相似度, 至少要做 $\frac{n(n+1)}{2}$ 次运算, 当证据个数、焦元个数较多时, 证据推理计算量过大, 不满足计算的要求。

(2) 利用证据间的相似系数, 确定其它证据对该证据的支持度, 并以该支持度作为证据的权重并不合理, 如: 假设有 3 个证据, 焦元分别为 A, B, C , 其基本概率分配为:

$$m_1(A)=0.5, m_1(B)=0.5, m_1(C)=0$$

$$m_2(A)=0.5, m_2(B)=0.5, m_2(C)=0$$

$$m_3(A)=0.9, m_3(B)=0, m_3(C)=0.1$$

从原证据模型可以看出 m_1, m_2 相似度高, 而 m_3 与其余两个证据的相似度低, 因此利用文献[7-9]计算, m_1, m_2 获得较高的权重, 而 m_3 因为支持度低, 所以获得较低的权重。但是由于 m_1, m_2 对焦元 A, B 的基本概率分配数相同, 因此具有较高的自冲突。如果对这两个证据赋予较高的权重, 而给 m_3 较小的权重, 则其合成结果使命题 A 的信任度降低。因此本文提出了一种新的计算权重法, 利用各证据的算术平均证据作为基础证据, 然后计算各证据与基础证据的相似度来获得证据的权重。

3 新的证据权重确定方法

步骤 1:

定义 2 设识别框架 Θ 下的证据为 $E_1, E_2, \dots, E_n, m_1, m_2, \dots, m_n$ 是识别框架 Θ 上的基本信任分配函数, 则各证据的平均信任分配函数为:

$$m_a = \frac{1}{n} \sum_{i=1}^n m_i \quad (3)$$

定义 3 假设证据的焦元分别是 A_k 和 B_k , 在定义 2 的基础上, 定义各证据与平均证据的相似系数为:

$$d_{ia} = \frac{\sum_{A_k \cap B_k \neq \Phi} m_i(A_k) m_a(B_k)}{\sqrt{\sum_{i=1}^n m_i^2(A_k)} \sqrt{\sum_{a=1}^n m_a^2(B_k)}} \quad (4)$$

将其归一化处理后得到各证据的相对支持度为:

$$resup(m_i) = \frac{d_{ia}}{\sum_{i=1}^n d_{ia}}, i, j=1, 2, \dots, n \quad (5)$$

从式(5)可以看出, 若一个证据与平均证据的相似度越高, 那么它的相对支持度也越高, 因此可以把相对支持度当作证据 m_i 的初始权重 w_i 。

步骤 2:

在证据权重 w_i 小于 $1/n$ (n 为证据个数) 的证据中, 将证据权重最小的证据定性为冲突证据, 并将该证据从证据组中剔除, 被剔除的证据的权重赋值为 0, 按步骤 1 中的式(3) — 式(5)重新计算各剩余证据的权重。

步骤 3:

重复步骤 1、步骤 2, 直到所有证据的权重都大于 $1/m$ (m 为剩余证据的个数)。

步骤 4: 假设迭代次数为 K , 最后将各证据的 K 次权重算术平均得到各证据的权重。

4 算例分析

假设有 3 组证据 E_1, E_2, E_3, E_4, E_5 , 焦元分别为 A, B, C , 其基本信任分配函数为:

$$m_1(B)=0.5, m_1(A)=0.5, m_1(C)=0;$$

$$m_2(A)=0.5, m_2(B)=0.5, m_2(C)=0;$$

$$m_3(A)=0.7, m_3(B)=0.1, m_3(C)=0.2;$$

$$m_4(A)=0.6, m_4(B)=0.1, m_4(C)=0.3;$$

$$m_5(A)=0, m_5(B)=0.9, m_5(C)=0.1。$$

为验证本方法的有效性, 把 D-S 合成法、文献[7]方法、文献[8]方法和本文方法进行比较, 各种方法的合成结果如表 1 所列。

表 1 几种组合方法的比较

	$m(A)$	$m(B)$	$m(C)$
D-S 组合规则	0	1	0
文献[7]方法	0.718	0.232	0.05
文献[8]方法	0.5	0.5	0
本文方法	0.878	0.123	0.009

结束语 对表 1 中的组合结果进行分析, D-S 组合规则不能正确处理证据高冲突问题, 出现了一票否决的现象; 文献[7]虽然能有效地处理证据冲突问题, 但是合成结果并不理想, 主要有如下原因: 当证据中出现自冲突证据时, 利用各证据之间的相似度来确定证据的权重并不合理; 该方法并没有对冲突证据进行处理。文献[8]在文献[7]的基础上进行了改进, 设立了阈值来处理冲突证据, 通过该算例表明这种方法是错误的, 因为利用文献[8]的方法, 证据 3、证据 4、证据 5 被定性为冲突证据, 最后合成结果出现了无法确定是 A 还是 B 的现象。本文在文献[8]的基础上做了 3 点修改: 1) 以平均证据作为基础证据, 计算各证据与平均证据的相似度来确定该证据的权重, 降低了计算的复杂程度; 2) 对冲突证据判定做了修正, 文献[8]将 $1/n$ (n 为证据个数) 作为阈值, 并将权重小于 $1/n$ 的证据定性为冲突证据, 而本文将权重小于 $1/n$ 的证据中权重最小的那一个证据定位冲突证据, 避免了误将有效证据判定为冲突证据; 3) 将 K 次权重算术平均得到各证据的最终权重。最后利用该方法确定的权重对证据加权平均, 再利用 D-S 组合规则进行组合。表 1 的组合结果表明该方法不仅有效, 而且合成结果更为理想。

参 考 文 献

- [1] Dempster A P. Upper and lower probabilities induced by a multivalued mapping[J]. Annals of Mathematical Statistics, 1967, 38(2):325-339
- [2] Simard M A, et al. Data fusion of multiple sensors attribute information for target identity estimation using a Dempster-Shafer evidential combination algorithm[C]// Proceedings of SPIE International Society for Optical Engineering. 1996, 2759:577-588
- [3] Barnett J A. Computational methods for a mathematical theory of evidence[M]// Classic Works of the Dempster-Shafer Theory of Belief Functions. Springer Berlin Heidelberg, 2008:197-216
- [4] Zadeh L A. Review of books: A mathematical theory of evidence[J]. AI Magazine, 1984, 5(3):81-83
- [5] 韩峰, 杨万海, 袁晓光. 一种有效处理证据冲突的组合方法[J]. 电光与控制, 2010, 17(4):5-9
- [6] Murphy C K. Combining belief functions when evidence conflicts[J]. Decision Support Systems, 2000, 29(1):1-9
- [7] 王肖霞, 杨风暴. 一种处理冲突证据的合成方法[J]. 弹箭与制导学报, 2007, 27(5):255-257
- [8] 黄青, 陈以. 一种改进的基于权重系数的证据合成方法[J]. 传感器与微系统, 2012, 31(7):14-16
- [9] 白剑林, 王煜. 一种解决 D-S 理论冲突的有效方法[J]. 系统与电子技术, 2009, 31(9):2106-2109