

文章编号: 1000-5862(2020)03-0275-07

一种基于证据焦元距离的权重计算方法

李俊杰, 吴根秀*, 焦黎洋

(江西师范大学数学与信息科学学院, 江西 南昌 330022)

摘要: 针对经典的 D-S 合成方法常会出现与实际相悖的不足, 该文提出了一种考虑焦元交互影响的权重合成方法. 该方法结合香农熵的思想定义了新的焦元距离测度, 基于新距离定义各证据间的权重, 然后根据权重对原始证据函数进行修正, 降低异常证据的重要性, 再利用 D-S 规则进行证据合成. 数值分析说明: 这一方法可排除干扰性信息, 降低冲突, 合成结果优于其他经典的计算方法, 为 D-S 证据理论提供一种有效的计算方法.

关键词: D-S 合成; 焦元交互影响; 焦元距离; 证据理论

中图分类号: TP 391 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2020.03.11

0 引言

证据理论是被 A. P. Dempster^[1] 在 1967 年提出的, 然后由 G. Shafer 在 1976 年出版的《证据的数学理论》专著中发展起来的一种不精确推理理论^[2]. 现在已被广泛运用于网络安全、信息融合等多个领域^[3-5], 但是在实际运用过程中, 当证据之间存在冲突时会出现一些不合理的结论^[6], 例如信任偏移问题、一票否决问题等反直观的现象. 基于此, 近 20 年来国内外许多学者对其进行了进一步地研究, 提出了一些改进的方法^[7-23], 主要通过如下方式对证据融合进行了改进: (i) 修正证据源. 这种方法认为问题来自证据源本身, 即先对证据源进行预处理, 再用 D-S 证据合成规则进行合成. 如 C. K. Murphy^[7] 的算术平均法认为每个证据都是同等重要; 邓勇等^[8] 提出了对每个证据赋予权重的平均值法; 韩德强等^[9] 利用不确定度对证据源进行修正; 毕文豪等^[10] 基于相似度进行修正证据源; 孟光磊等^[11] 提出了新的权重计算合成; 陈云翔等^[12] 考虑时域对证据源修正; 张扬等^[13] 提出了基于数据分聚类的融合方法; 陈元超^[14] 优化了证据冲突的融合算法. 还有基于各种距离赋予权重来修正证据源的方法, 如马氏距离^[15]、Pignistic 概率距离^[16]、Jousselme 距离^[17]. (ii) 修改组合规则. 这种方法认为是对产生的冲突值处理不当, 应对其重新分配. 李弼程等^[18] 按各个

命题的平均支持程度分配冲突值, 但未考虑各证据交叉问题; 孙全等^[19] 采用了加权和平均的形式, 认为证据间的冲突部分可用, 可用程度取决于可信度, 有着比较大主观随意性; R. R. Yager^[20] 则把冲突值全部赋予未知, 实际上是对冲突值的否定; M. C. Florea 等^[21] 将交运算与并运算结合提出了一种新的组合规则. (iii) 对证据源与合成规则同时修正. 陈炜军等^[22] 基于信息容量和证据距离对证据源进行预处理, 再基于可信度修正融合规则; 卢正才等^[23] 提出了基于一般框架下的修正融合方法.

以上方法在不同程度上都对合成结果有所改善, 但总体上来说不够理想, 对证据组合规则修改有所不妥, 因为 Dempster 组合规则本身没什么问题, 并且具有良好的性质, 满足交换律、结合律, 在证据源多的情况下也较好进行处理, 只是在证据源出现异常时会出现反直观的效果, 因此应从修正证据源出发, 排除异常证据的干扰. 本文通过分析 2 种常用的证据冲突的距离函数, 提出了一种关于焦元交互影响的新距离测度, 并给出基于此距离测度的方法来修正证据源.

1 证据理论基本概念

定义 1^[2] 在 D-S 证据理论中, 设 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ 为有限识别框架, 其中元素的个数是有限

收稿日期: 2019-11-17

基金项目: 国家自然科学基金(61462045)和江西省学位与研究生教育教学改革研究(JXYJG-2015-034)资助项目.

通信作者: 吴根秀(1965-), 女, 江西南丰人, 教授, 主要从事不确定性推理与信息融合的研究. E-mail: wgx_nc@sina.com

的,并且元素两两互斥 $m: 2^\Theta \rightarrow [0, 1]$ 满足

$$\begin{cases} m(\emptyset) = 0, \\ \sum_{A_i \subseteq \Theta} m(A_i) = 1, \end{cases}$$

称 m 为 Θ 上的基本概率赋值函数(简称 BPA 函数), $\forall A \subseteq \Theta$ 若 $m(A) > 0$, 则称 A 为 Θ 的一个焦元, $m(A)$ 反映了证据对 A 的信任度.

定义 2^[2] 设 m_1, m_2 都为识别框架 Θ 上的基本概率赋值(BPA)函数, 定义 m 为

$$m(C) = \begin{cases} \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - K}, & C \neq \emptyset, \\ 0, & C = \emptyset, \end{cases}$$

其中 $K = \sum_{A \cap B = \emptyset} m_1(A) m_2(B)$ 称 K 为 m_1 与 m_2 之间的冲突系数, 反映证据间的冲突程度. 定义 m 为 m_1 与 m_2 的 D-S 合成函数, 记为 $m = m_1 \oplus m_2$.

2 D-S 证据合成规则产生的问题

在运用 D-S 合成规则进行证据合成时, 当传感器不稳定并出现异常证据时会出现与直觉不符或无法进行决策的结果. 以如下 2 个问题为例.

2.1 一票否决问题

例 1 设识别框架 $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $m_1, m_2, m_3, \dots, m_n$ 为同一识别框架 Θ 的 n 个基本概率赋值函数, 具体取值为 $m_1(\theta_1) = 0, m_1(\theta_2) = 0.01, m_1(\theta_3) = 0.99; m_2(\theta_1) = 0.99, m_2(\theta_2) = 0.01, m_2(\theta_3) = 0$, 其中 $m_2 = m_3 = m_4 = \dots = m_n$, 经过 D-S 合成后结果 $m = m_1 \oplus m_2 \oplus \dots \oplus m_n$ 为

$$m(\theta_1) = 0, m(\theta_2) = 1, m(\theta_3) = 0.$$

虽然每条证据对 θ_2 的支持度都很低, 但合成结果仍然认为 θ_2 为真, 而且与证据个数 n 无关, 这显然不符合实际情形.

2.2 信任偏移问题

例 2 设识别框架 $\Theta = \{\theta_1, \theta_2, \theta_3\}$, 有 n 条证据分别为 $m_i(\theta_1) = 0.1, m_i(\Theta) = 0.9 (i = 1, 2, \dots, n)$, 运用 D-S 合成规则可得 $m = m_1 \oplus m_2 \oplus \dots \oplus m_n$ 为

$$m(A) = \begin{cases} 1 - 0.9^n, & A = \theta_1, \\ 0.9^n, & A = \Theta, \end{cases}$$

由上述可知, 每个证据对 θ_1 的信任值都很低, 但是由于证据的增多, 从而合成结果使得 θ_1 的信任值趋

近于 1, 这显然也与常理不相符合.

通过上述例子发现经典 D-S 合成方法的不足, 因为在信息融合中对所有的证据都视为同等重要的, 而在实际问题中, 会经常出现传感器系统故障或者干扰性信息, 导致信息不准确, 从而使得最终合成效果不理想, 所以需要原始证据进行处理. 通过在证据之间引入距离测度, 刻画出证据间的差异, 从而得到证据的冲突度, 最终确定权重.

3 证据焦元距离的选取

3.1 Jousselme 距离与 Pignistic 概率距离

证据距离用来描述证据之间的差异性或者相似性, 两者之间距离越大表示它们之间的差异性越大, 相似度越小. 常用的距离函数有 Pignistic 概率距离^[16]、Jousselme 距离^[17]等证据距离. 通过对这 2 种距离的比较分析, 提出了一种新的度量焦元距离的测度.

定义 3^[17] 设 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, m_1, m_2 分别为同一识别框架 Θ 上的基本概率赋值(BPA)函数, 称

$$d_J(m_1, m_2) = \sqrt{(\vec{m}_1 - \vec{m}_2)^T D (\vec{m}_1 - \vec{m}_2) / 2}$$

为证据 m_1, m_2 的 Jousselme 距离. 其中 \vec{m} 是由 BPA 生成的 $2^{|\Theta|} - 1$ 维的列向量, D 是一个 $(2^{|\Theta|} - 1) \times (2^{|\Theta|} - 1)$ 的矩阵, 其中元素定义为 $D(a, b) = |a \cap b| / |a \cup b|$, 这里 a, b 表示 Θ 的子集, $|A|$ 为焦元 A 的基数.

定义 4^[16] 设识别框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, m 为 Θ 上基本概率赋值(BPA)函数, 若 $A \subseteq \Theta$, 则称

$$P_m(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} m(B)$$

为焦元 A 的 Pignistic 概率.

定义 5^[16] 在 D-S 证据理论中, 设识别框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, m_1, m_2 为 Θ 上的基本概率赋值(BPA)函数, 称

$$d_P(m_1, m_2) = \max_{A \subseteq \Theta} (|P_{m_1}(A) - P_{m_2}(A)|)$$

为证据 m_1, m_2 的 Pignistic 概率距离.

定义 3 和定义 5 中的 Jousselme 距离和 Pignistic 概率距离在一定程度上反映了证据之间的差异, 距离越大, 证据 m_1, m_2 之间的差异性越大.

定义 6 m_1, m_2 为 Θ 上的基本概率赋值(BPA)函数, 对 $A_i \subseteq \Theta$ 称

$$a_i = m_1(A_i) - m_2(A_i)$$

为命题 A_i 对 m_1, m_2 之间的距离产生的影响值; 对 $A_i, B_j \subseteq \Theta$ 称 $a_{ij} = a_i - a_j$ 为命题 A_i 和 B_j 对 m_1, m_2 之间的距离产生的交互影响值.

定理1 设识别框架 $\Theta = \{\theta_1, \theta_2\}$ m_1, m_2 为 Θ 上的基本概率赋值函数, 且

$$\begin{aligned} m_1(\theta_1) &= m_{11}, m_1(\theta_2) = m_{12}, m_1(\theta_1\theta_2) = m_{13}; \\ m_2(\theta_1) &= m_{21}, m_2(\theta_2) = m_{22}, m_2(\theta_1\theta_2) = m_{23}; \end{aligned}$$

令 $a_i = m_{1i} - m_{2i} (i = 1, 2, 3)$, 则

$$d_j(m_1, m_2) = \sqrt{(a_1^2 + a_2^2)/2}.$$

$$\begin{aligned} \text{证 } d_j(m_1, m_2) &= \sqrt{(\vec{m}_1 - \vec{m}_2)^T D (\vec{m}_1 - \vec{m}_2)/2} = \\ &= \sqrt{(a_1 \ a_2 \ a_3) \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix} (a_1 \ a_2 \ a_3)^T / 2} = \\ &= \sqrt{(a_1^2 + a_2^2 + a_3^2 + a_1a_3 + a_2a_3)/2} = \\ &= \sqrt{(a_1^2 + a_2^2 + a_3(a_1 + a_2 + a_3))/2}, \end{aligned}$$

因为 $a_1 + a_2 + a_3 = 0$, 所以 $d_j(m_1, m_2) = \sqrt{(a_1^2 + a_2^2)/2}$. 定理1得证.

推论1 识别框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ 上的2个基本概率赋值函数 m_1, m_2 的 $d_j(m_1, m_2)$ 是单个命题影响值 a_i 的函数.

定理2 条件同定理1, 则 $d_p(m_1, m_2) = |a_1 - a_2|/2$.

推论2 识别框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ 上的2个基本概率赋值函数 m_1, m_2 的 $d_p(m_1, m_2)$ 是命题 A_i 与 A_j 的交互影响值 $a_{ij} = a_i - a_j$ 的函数.

由定理1、定理2及推论1、推论2可知 d_j 距离仅与单个命题产生的影响值有关, 而 d_p 距离只与命题的交互影响值有关, 这就导致了2种距离在应用时会产生与实际相悖的结果, 下面例3和例4可说明问题.

例3 设识别框架 $\Theta = \{a, b, c\}$ m_1, m_2, m_3 都为 Θ 上的基本概率赋值函数, 且

$$\begin{aligned} m_1(a) &= 0.5, m_1(b) = 0.2, m_1(c) = 0.2, \\ m_1(ac) &= 0.1; m_2(a) = 0.2, m_2(b) = 0.3, m_2(c) = \\ 0.5; m_3(a) &= 0, m_3(b) = 0.1, m_3(c) = 0.9. \end{aligned}$$

由定义3得 $d_j(m_1, m_2) = 0.3162 < d_j(m_2, m_3) = 0.3464$. 而由直观分析可得: 证据 m_2 和证据 m_3 支持 c , 而证据 m_1 支持的是 a . 直觉上 m_2 与 m_3 之间的距离应该小于 m_1 与 m_2 之间的距离, 但 d_j 与直觉不符.

例4 设识别框架 $\Theta = \{\theta_1, \theta_2\}$ m_1, m_2, m_3 为 Θ 上的基本概率赋值函数, 且

$$\begin{aligned} m_1(\theta_1) &= 0.4, m_1(\theta_2) = 0.3, m_1(\theta_1\theta_2) = 0.3; \\ m_2(\theta_1) &= 0.1, m_2(\theta_2) = 0.2, m_2(\theta_1\theta_2) = 0.7; \\ m_3(\theta_1) &= 0.2, m_3(\theta_2) = 0.1, m_3(\theta_1\theta_2) = 0.7. \end{aligned}$$

由定义5得 $d_p(m_1, m_2) = d_p(m_2, m_3) = 0.1$, m_1 与 m_2 之间的距离和 m_2 与 m_3 之间的距离相等, 即2者的差异性相同. 而3个证据之间 m_2 与 m_3 都支持 $\theta_1\theta_2$, 而 m_1 支持 θ_1 要更大些. 由直观分析知 m_2 与 m_3 之间的距离应小于 m_1 与 m_2 之间的距离, 但 d_p 与直觉不符.

3.2 证据的焦元距离

为了克服上述证据距离的不足, 综合考虑单一命题产生的影响值和命题之间的交互影响值, 并借鉴香农熵的思想, 引入新的焦元距离函数.

定义7 设 Θ 为有限识别框架 m_1, m_2 为 Θ 的基本概率赋值函数, 称

$$\begin{aligned} d_w(m_1, m_2) &= - \sum_{A \subseteq \Theta} (\lg(1 - |m_1(A) - m_2(A)|)) / \\ &|A| - \sum_{\substack{A \cap B \neq \emptyset \\ A \neq B, A, B \subseteq \Theta}} (|A \cap B| \lg(1 - |m_1(A) - \\ &m_2(A)| - |m_1(B) - m_2(B)|)) / |A \cup B| \end{aligned}$$

为证据 m_1, m_2 的焦元距离函数.

该距离的第1部分反映的是单个命题的影响, 第2部分反映的是命题的交互影响. 同理, 在推论1的条件下 $d_w(m_1, m_2)$ 是 a_i 和 a_{ij} 的函数.

由定义7可得, 在例3中 $d_w(m_1, m_2) \approx 0.6972 > d_w(m_2, m_3) \approx 0.4156$, 在例4中 $d_w(m_1, m_2) \approx 0.5122 > d_w(m_2, m_3) \approx 0.1830$, 这些都与直观一致, 该距离既刻画了证据之间的差异性, 又考虑了证据函数的聚焦度.

3.3 4种距离函数的比较分析

定义8 设 Θ 为有限识别框架 m_1, m_2 为 Θ 的基本概率赋值函数, 称

$$\begin{aligned} d_{WE}(m_1, m_2) &= - \sum_{A \subseteq \Theta} (\lg(1 - |m_1(A) - \\ &m_2(A)|)) / |A| \end{aligned}$$

为证据 m_1, m_2 的单焦元证据距离函数.

为对比分析 d_j, d_p, d_w, d_{WE} 4个距离, 随机生成了100对基本概率赋值函数, 通过 Matlab 计算出100对BPA的4个距离并绘制出图1.

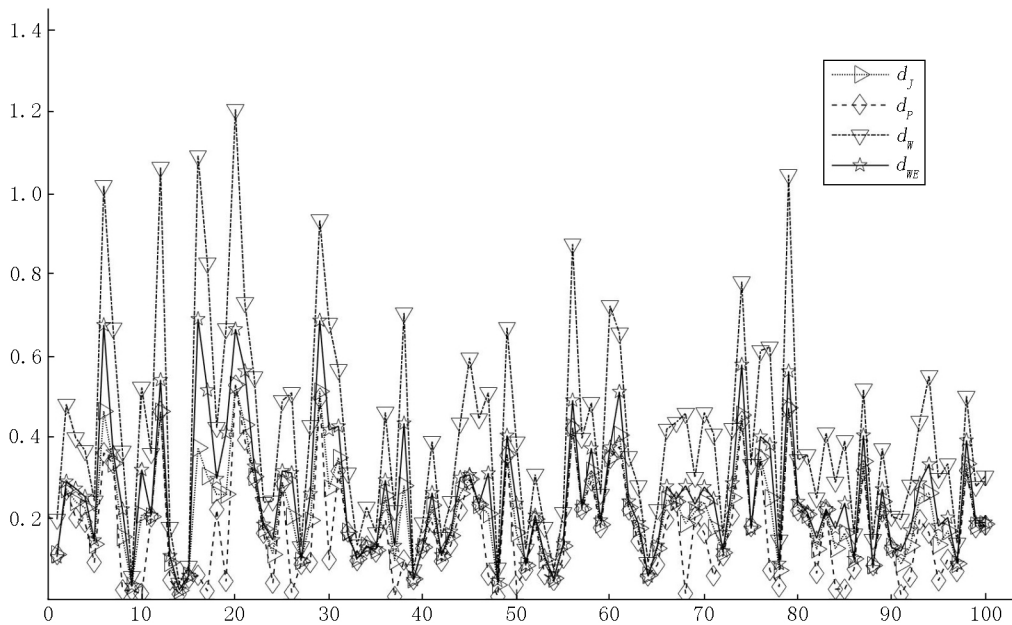


图1 100 对 BPA 的 4 个距离的对比图

由图1可知:在100组证据中4个证据反映的趋势大部分比较吻合,这也说明了定义7的焦元距离 d_w 的合理性.另一方面,在每组随机产生的证据中,焦元距离 d_w 普遍都比其他3个距离更大,且 d_w 值跨度比较大,这说明 d_w 刻画证据之间的差异较明显,对于干扰证据灵敏度更高,故选取 d_w 来刻画证据之间的差异性.

4 证据权重计算及其合成算法

4.1 证据权重的相关定义

定义9 设 Θ 为有限识别框架 m_1, m_2, \dots, m_n 都为 Θ 上的基本概率赋值函数,称

$$C_i = D_i / \sum_{i=1}^n D_i$$

为证据 m_i 与其他证据的冲突度,其中 $D_i = \sum_{j=1, j \neq i}^n d_w(m_i, m_j)$.

定义10 基本概率赋值函数 m_1, m_2, \dots, m_n 的冲突度分别是 C_1, C_2, \dots, C_n ,则称

$$\omega_i = (1 - C_i) / \sum_{i=1}^n (1 - C_i), \quad \omega_i \in [0, 1]$$

为证据 m_i 的权重.

4.2 考虑命题交互影响的权重合成方法

设 Θ 为有限识别框架 m_1, m_2, \dots, m_n 为 Θ 的基本概率赋值函数.改进权重的合成方法步骤如下:

Step 1 由定义7计算2个证据之间的距离 $d_w(m_i, m_j)$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, n$).

Step 2 由定义9得到 D_i 和 C_i ,由定义10得到第 i 个证据权重 ω_i ($i = 1, 2, \dots, n$).

Step 3 根据权重 ω_i ($i = 1, 2, \dots, n$)得到修正

$$\text{证据为 } m(A) = \sum_{i=1}^n \omega_i m_i(A), \quad A \subseteq \Theta.$$

Step 4 对修正证据 m 进行 $n-1$ 次D-S合成,从而得到最终的合成结果.

5 实例分析

为了验证本文方法的有效性和合理性,下面从2个实例中给予说明.从单焦元上分析,例5是单焦元证据函数从2个证据到5个证据进行合成,并和D-S^[2]、R. R. Yager^[20]、陈炜军等^[22]、卢正才等^[23]的方法进行比较.

例5 设识别框架 $\Theta = \{a, b, c\}$,5个证据函数分别如下:

$$\begin{aligned} m_1(a) &= 0.98, m_1(b) = 0.01, m_1(c) = 0.01; \\ m_2(a) &= 0.90, m_2(b) = 0.05, m_2(c) = 0.05; \\ m_3(a) &= 0.90, m_3(b) = 0.05, m_3(c) = 0.05; \\ m_4(a) &= 0, m_4(b) = 0.01, m_4(c) = 0.99; \\ m_5(a) &= 0.99, m_5(b) = 0.01, m_5(c) = 0. \end{aligned}$$

5种合成方法的结果如表1所示.从合成结果可以看出:在没有出现干扰证据之前,5种方法合成结果都比较合理,但当出现干扰证据 m_4 时,合成结果出现了较大差异,这是由于有些方法不能识别干扰证据、缓解冲突而导致的.在自冲突方面,如图2所示,文献[23]与本文方法能较好地缓解冲突,其他自冲突值都为1,这明显不合理,这里体现了 d_w 优越性;在聚焦方面,如图3所示,当出现干扰证据 m_4 时,本文方法 $m(a)$ 的波动较小,具有更强的干扰性,并且在聚焦速率上来说,本文具有更快的聚焦速率.

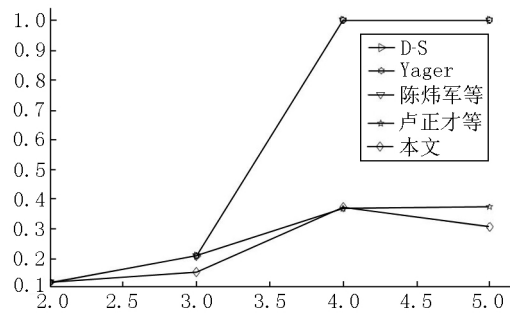


图 2 自冲突与证据个数

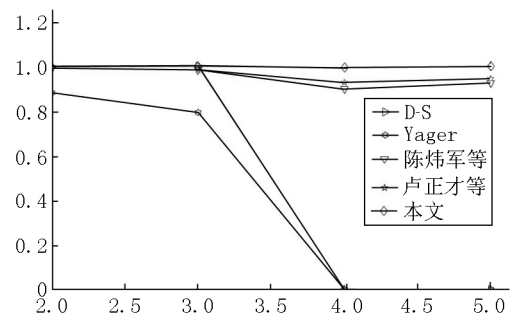


图 3 $m(a)$ 与证据个数的关系

表 1 例 5 中 5 种合成方法的多证据合成结果

合成证据	方法	k	$m(a)$	$m(b)$	$m(c)$	$m(\Theta)$
m_{12}	D-S	0.117 000	0.998 800	0.000 600	0.000 600	0
	R. R. Yager	0.117 000	0.882 000	0.000 500	0.000 500	0.117 000
	陈炜军等	0.117 000	0.992 000	0.004 000	0.004 000	0
	卢正才等	0.117 000	0.992 000	0.004 000	0.004 000	0
	本文	0.114 600	0.998 000	0.001 000	0.001 000	0
m_{123}	D-S	0.206 150	0.999 940	0.000 030	0.000 030	0
	R. R. Yager	0.206 150	0.793 800	0.000 025	0.000 025	0.206 150
	陈炜军等	0.206 150	0.984 694	0.007 653	0.007 653	0
	卢正才等	0.206 150	0.984 678	0.007 661	0.007 661	0
	本文	0.150 400	0.999 835	0.000 082	0.000 082	0
m_{1234}	D-S	0.999 975	0	0.010 000	0.990 000	0
	R. R. Yager	0.999 975	0	0	0.990 000	0.999 975
	陈炜军等	0.999 975	0.898 266	0.036 471	0.065 263	0
	卢正才等	0.364 915	0.928 110	0.011 931	0.021 387	0.038 572
	本文	0.364 915	0.995 504	0.000 003	0.004 492	0
m_{12345}	D-S	1.000 000	0	1.000 000	0	0
	R. R. Yager	1.000 000	0	0	0	1.000 000
	陈炜军等	1.000 000	0.925 886	0.029 992	0.044 122	0
	卢正才等	0.371 310	0.945 765	0.010 271	0.016 238	0.038 572
	本文	0.302 673	0.999 769	0	0.000 231	0

在多焦元的合成问题上,本文分别给出了 5 个不同证据源合成进行了实验,与文献[2]、文献[8]、文献[9]等 3 种合成进行了对比分析。

例 6 设识别框架 $\Theta = \{a, b, c\}$, 5 个证据函数分别如下:

$$\begin{aligned} m_1(a) &= 0.50, m_1(b) = 0.20, m_1(c) = 0.30; \\ m_2(a) &= 0.65, m_2(b) = 0.10, m_2(ac) = 0.25; \\ m_3(a) &= 0, m_3(b) = 0.80, m_3(c) = 0.20; \\ m_4(a) &= 0.75, m_4(b) = 0.10, m_4(c) = 0.15; \\ m_5(a) &= 0.55, m_5(b) = 0.10, m_5(ac) = 0.35. \end{aligned}$$

从上述 5 条证据中可以看出: 证据 m_1, m_2, m_4, m_5 都支持识别目标 a , 证据 m_3 则支持识别目标 b ; 从前面分析可知, 证据 m_3 应与其他证据有较大的差距, 即冲突度较大, 从而其权重应较小, 表 2 是对这 5 条证据的融合结果。在证据 m_1, m_2 融合时, 2 个证据冲突较小, 4 种方法的合成结果都符合直观。但在

后续的合成中, 当证据 m_3 出现时, 文献[2]的方法合成发生了与例 1 指出的一票否决问题, 而本文方法及文献[9]、文献[7]较好地克服了这一缺点, 这是因为这几种方法都对证据源进行修正, 有效地降低了异常证据 m_3 出现造成的干扰影响。在本文方法与文献[9]、文献[7]比较时, 3 种方法合成的结果都比较合理, 最终都是大概率的支持目标 a 。从收敛速率来看, 本文方法明显比文献[9]方法收敛速率更快, 这是因为文献[9]方法对每个证据赋予同样的权重进行了简单平均合成, 不能有效地识别干扰证据, 由于在实际问题中, 难免会遇到传感器异常, 导致证据源失真的情况存在, 应该对每个证据赋予不同的权重, 所以文献[9]方法有其不合理之处。相比文献[7]的合成方法, 相同点都是对证据源进行了分析, 赋予了不同的权重, 得到修正的证据体, 再依据 D-S 证据组合合成, 不同点在于本文在赋予权重的方法上从一种新的思路出发, 分析对比了经典

的 Pignistic 概率距离、Jousselme 距离,考虑引入关于焦元交互影响,结合香农熵的思想考虑证据的聚焦程度,引入一种新的距离测度,从而赋予证据权重.韩德强等^[9]的方法合成结果符合直观,收敛速率也比前面几种更好,但本文相比文献[9]的方法聚焦

速率更快.通过上述例子表明,本文充分考虑了证据之间的差异性,能有效识别干扰信息,有效地降低了证据间冲突而导致的不良合成结果,且本方法使得收敛速率加快.

表 2 例 6 中 4 种合成方法的多证据合成结果

方法	基本概率赋值函数			
	$m_1 \ m_2$	$m_1 \ m_2 \ m_3$	$m_1 \ m_2 \ m_3 \ m_4$	$m_1 \ m_2 \ m_3 \ m_4 \ m_5$
D-S	$m(a) = 0.825 \ 7$	$m(a) = 0$	$m(a) = 0$	$m(a) = 0$
	$m(b) = 0.036 \ 7$	$m(b) = 0.516 \ 1$	$m(b) = 0.415 \ 6$	$m(b) = 0.168 \ 9$
	$m(c) = 0.137 \ 6$	$m(c) = 0.483 \ 9$	$m(c) = 0.584 \ 4$	$m(c) = 0.831 \ 1$
C. K. Murphy	$m(a) = 0.828 \ 6$	$m(a) = 0.608 \ 8$	$m(a) = 0.886 \ 7$	$m(a) = 0.975 \ 0$
	$m(b) = 0.039 \ 3$	$m(b) = 0.297 \ 0$	$m(b) = 0.086 \ 1$	$m(b) = 0.013 \ 7$
	$m(c) = 0.104 \ 8$	$m(c) = 0.090 \ 7$	$m(c) = 0.027 \ 1$	$m(c) = 0.011 \ 0$
	$m(ac) = 0.027 \ 3$	$m(ac) = 0.003 \ 5$	$m(ac) = 0.000 \ 2$	$m(ac) = 0.000 \ 3$
韩德强等	$m(a) = 0.825 \ 7$	$m(a) = 0.644 \ 4$	$m(a) = 0.919 \ 5$	$m(a) = 0.976 \ 7$
	$m(b) = 0.036 \ 7$	$m(b) = 0.153 \ 0$	$m(b) = 0.024 \ 0$	$m(b) = 0.002 \ 3$
	$m(c) = 0.137 \ 6$	$m(c) = 0.202 \ 6$	$m(c) = 0.056 \ 5$	$m(c) = 0.020 \ 6$
	$m(ac) = 0$	$m(ac) = 0$	$m(ac) = 0$	$m(ac) = 0.000 \ 4$
本文	$m(a) = 0.828 \ 6$	$m(a) = 0.657 \ 9$	$m(a) = 0.925 \ 5$	$m(a) = 0.983 \ 4$
	$m(b) = 0.039 \ 3$	$m(b) = 0.243 \ 3$	$m(b) = 0.050 \ 2$	$m(b) = 0.007 \ 0$
	$m(c) = 0.104 \ 8$	$m(c) = 0.094 \ 7$	$m(c) = 0.024 \ 1$	$m(c) = 0.009 \ 4$
	$m(ac) = 0.027 \ 3$	$m(ac) = 0.003 \ 1$	$m(ac) = 0.000 \ 1$	$m(ac) = 0.000 \ 3$

6 结论

在信息融合中,由于信息干扰、机械故障等各种不确定因素使得到的证据函数有偏差,从而出现目标识别错误的情况,因此为了克服这一困难关键在于如何识别干扰信息和改进 D-S 合成方法,本文提出了一种新的刻画证据之间差异的焦元距离,通过随机产生的 100 组证据的实验对比,说明了焦元距离灵敏度更高,并通过和其他方法比较,验证了本文方法的可行性和合理性.但在应用方面仍还有许多工作需要进一步地研究.

7 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] Shafer G. A mathematical theory of evidence [M]. Princeton: Princeton University Press, 1976.

[3] Guan Jiwen, Bell D A. Evidence theory and its applications [M]. North-Holland, Amsterdam, New York: Elsevier Science Publisher, 1992.

[4] 汤永利,李伟杰,于金霞,等. 基于改进 D-S 证据理论的

网络安全态势评估方法 [J]. 南京理工大学学报, 2015, 39(4) : 405-411.

[5] 韩德强,杨艺,韩崇昭. DS 证据理论研究进展及相关问题探讨 [J]. 控制与决策, 2014, 29(1) : 1-11.

[6] 吴根秀. 信任函数组合与局部冲突处理 [J]. 计算机工程与应用, 2004, 40(34) : 81-84, 97.

[7] Murphy C K. Combining belief functions when evidence conflicts [J]. Decision Support Systems, 2000, 29(1) : 1-9.

[8] 邓勇,施文康,朱振福. 一种有效处理冲突证据的组合方法 [J]. 红外与毫米波学报, 2004, 23(1) : 27-32.

[9] 韩德强,邓勇,韩崇昭,等. 利用不确定度的冲突证据组合 [J]. 控制理论与应用, 2011, 28(6) : 788-792.

[10] 毕文豪,张安,李冲. 基于新的证据冲突衡量的加权证据融合方法 [J]. 控制与决策, 2016, 31(1) : 73-78.

[11] 孟光磊,龚光红. 证据源权重的计算及其在证据融合中的应用 [J]. 北京航空航天大学学报, 2010, 36(11) : 1365-1368.

[12] 陈云翔,罗承昆,王攀,等. 考虑可靠性的时域证据组合方法 [J]. 控制与决策, 2018, 33(3) : 463-470.

[13] 张扬,杨建华,侯宏. 一种基于证据理论的数据分聚类融合算法 [J]. 科学技术与工程, 2018, 18(1) : 54-58.

[14] 陈元超. 优化的证据冲突融合方法 [J]. 探测与控制学报, 2018, 40(5) : 116-120.

[15] 鲁睿,徐启建,张杰,等. 一种基于马氏距离的冲突证据组合方法 [J]. 通信技术, 2015, 48(3) : 278-282.

[16] Tessem B. Approximations for efficient computation in the

- theory of evidence [J]. Artificial Intelligence ,1993 ,61 (2) : 315-329.
- [17] Jousselme A L ,Grenier D ,Bossé É. A new distance between two bodies of evidence [J]. Information Fusion , 2001 ,2(2) : 91-101.
- [18] 李弼程 ,王波 ,魏俊 等. 一种有效的证据理论合成公式 [J]. 数据采集与处理 ,2002 ,17(1) : 33-36.
- [19] 孙全 ,叶秀清 ,顾伟康. 一种新的基于证据理论的合成公式 [J]. 电子学报 ,2000 ,28(8) : 117-119.
- [20] Yager R R. On the Dempster-Shafer framework and new combination rules [J]. Information Sciences ,1987 ,41 (2) : 93-137.
- [21] Florea M C ,Jousselme A L ,Bossé É ,et al. Robust combination rules for evidence theory [J]. Information Fusion , 2009 ,10(2) : 183-197.
- [22] 陈炜军 ,景占荣 ,袁芳菲 等. D-S 证据理论的不足及其数学修正 [J]. 中北大学学报: 自然科学版 ,2010 ,31 (2) : 161-168.
- [23] 卢正才 ,覃征. 证据合成的一般框架及高度冲突证据合成方法 [J]. 清华大学学报: 自然科学版 ,2011 ,51 (11) : 1611-1615 ,1626.

The Weight Calculation Method Based on Focal Distance of Evidence

LI Junjie ,WU Genxiu* ,JIAO Liyang

(College of Mathematics and Information Science ,Jiangxi Normal University ,Nanchang Jiangxi 330022 ,China)

Abstract: Because the classical D-S synthesis method often has shortcomings contrary to the reality ,a weight synthesis method considering the interaction of focal elements is proposed in this paper. By finding the shortcomings of several classical distances ,a new focal distance measure is defined in combination with evidence of entropy and the weight of each evidence is generated based on the new distance. Then ,the original evidence function is modified according to the weight ,reducing the importance of abnormal evidence ,and the D-S rule is used for evidence synthesis. Numerical analysis shows that this method can eliminate the interference information and reduce the conflict ,and the synthetic result is better than other classical methods ,which provides an effective calculation method for D-S evidence theory.

Key words: D-S synthesis; focal element interaction; focal distance; evidence theory

(责任编辑: 曾剑锋)