

文章编号: 1000-8152(2009)07-0812-03

基于证据可信度的证据合成新方法

张盛刚, 李巍华, 丁 康

(华南理工大学 机械与汽车工程学院, 广东 广州 510640)

摘要: Dempster-Shafer理论合成高冲突证据时结果不合常理, 许多改进方法能较好地解决这一问题, 但证据比较一致时, 这些方法性能相对于Dempster-Shafer合成规则较差. 鉴于此, 提出了一种基于证据可信度的证据合成方法. 首先建立一个证据距离矩阵, 之后求出系统中各证据到证据集的均方欧式距离, 通过引入可信度函数来获得证据源的可信度因子 ε 并对证据源进行修正, 最后使用Dempster-Shafer合成公式对修正后的证据加以合成. 数值实验结果表明, 改进后的方法不仅适用于证据高冲突情况, 而且也适用于证据比较一致的情况.

关键词: 证据理论; 可信度; 冲突证据; 一致证据; 合成方法

中图分类号: TP274 **文献标识码:** A

A novel approach to evidence combination based on the evidence credibility

ZHANG Sheng-gang, LI Wei-hua, DING Kang

(School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: Dempster-Shafer combination results may be not in agreement with the real condition when the evidences are highly conflict; and this problem has been widely investigated by some improved methods. However, those methods are not as good as Dempster-Shafer combination rule when the evidences are consistent. Therefore, a novel evidence combination method is proposed in this paper based on evidence credibility. First, the mean square Euclidian distance between evidence and evidence subset is obtained after establishing the evidence distance matrix, and the credible factor ε is calculated to modify the evidence resource correspondingly by a confidence function, and then these modified evidences are combined together according to the D-S rule. Simulation experiments demonstrate the effectiveness of this new improved algorithm when dealing with conflicting evidences and consistent ones.

Key words: evidence theory; credibility; conflicting evidences; consistent evidences; combination methods

1 引言(Introduction)

Dempster-Shafer^[1,2]证据理论(D-S理论)为不确定信息的表达和合成提供了自然而强有力的方法. 但在组合高冲突证据时, D-S理论得到的结果却往往有悖常理^[3]. 针对此问题, 已提出了许多改进的证据合成方法^[4~10], 但这些改进方法在组合较一致证据时, 得到的结果却不合常理. 实际中, 各个证据源不是100%可靠的, 当证据间冲突较大时, 说明各证据源存在不同程度的奇异, 可以通过对证据源进行修正或预处理来降低证据间的冲突. 本文以此为出发点, 提出了一种新的证据合成方法, 详细论述了该方法的原理. 最后通过仿真实验, 与D-S方法和几种典型的改进方法进行了比较.

2 基于证据可信度的证据合成方法(A combination method based on evidence credibility)

关于D-S理论的基本概念, 参见文献[1,2], 这里不再赘述. 设识别框架 $\Theta = \{A_1, A_2, \dots, A_N\}$, 证据集 $E = \{E_1, E_2, \dots, E_n\}$ 对应的基本概率赋值函数为 m_1, m_2, \dots, m_n . 定义证据距离矩阵 $D_{n \times n}$:

$$D = \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix}. \quad (1)$$

其中

$$d_{ij} = \sqrt{\sum_{k=1}^N [m_i(A_k) - m_j(A_k)]^2}. \quad (2)$$

N 为 Θ 中命题个数, d_{ij} 的最大值 $\max d_{ij} = \sqrt{2}$. 对 D 做归一化处理, 得正则化距离矩阵 \tilde{D} , \tilde{D} 中元素 $\tilde{d}_{ij} = d_{ij}/\sqrt{2}$, 则 $\tilde{d}_{ij} \in [0, 1]$. 定义证据 E_i 到证据集 E 的均方欧式距离为

$$s_i = \frac{1}{n} \sum_{j=1}^n \tilde{d}_{ij}^2 = \frac{1}{2n} \sum_{j=1}^n d_{ij}^2. \quad (3)$$

n 为证据个数, 满足 $s_i \in [0, 1]$. s_i 的大小反映了证据 E_i 同其他证据的差异程度, s_i 较大时, 说明证据 E_i 同其他证据存在较大差异, 该证据的可信度因子 ε_i 较小, 即 $s_i \rightarrow 1$ 时, $\varepsilon_i \rightarrow 0$; 反之, 说明该证据的可信度因子 ε_i 较大, 即 $s_i \rightarrow 0$ 时, $\varepsilon_i \rightarrow 1$. 设 $\varepsilon_i = f(s_i)$, 则 $f(s_i)$ 应满足: $f(s_i)$ 应随 s_i 的增大单调递减, 同时, s_i 较小时, 说明证据可信度较高, $f(s_i)$ 应衰减缓慢, 反之, $f(s_i)$ 应迅速衰减到零. 因此, $f(s_i)$ 应该是一条指数关系曲线, 同文献[5]中可信度的定义相似, 定义证据的可信度因子

$$\varepsilon_i = f(s_i) = (1 - s_i)k^{-s_i}. \quad (4)$$

上式中 k 为待确定的未知参数, 对式(4)求导数, 得

$$f'(s_i) = (s_i \ln k - \ln k - 1)k^{-s_i}. \quad (5)$$

根据要求, 要保证 $f(s_i)$ 在 $s_i \in [0, 1]$ 时是减函数, 则需 $f'(s_i) < 0$. 求解不等式 $(s_i \ln k - \ln k - 1)k^{-s_i} < 0$, 得 $k \geq e^{-1}$. 图1所示为 k 取不同值时 ε_i 随 s_i 变化的关系曲线, 只有 $k = e^{-1}$ 时的关系曲线最符合要求. 因此

$$\varepsilon_i = (1 - s_i)\exp s_i. \quad (6)$$

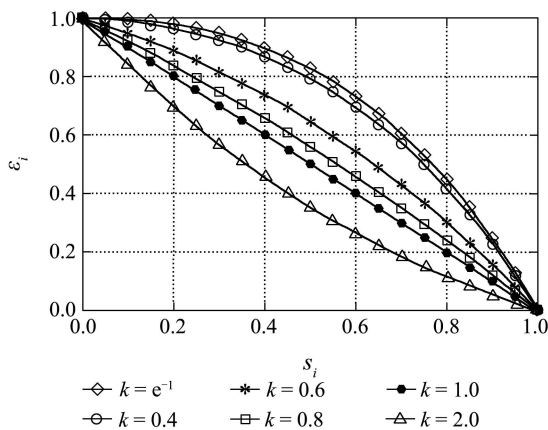


图1 k 取不同值时, 可信度因子与均方欧式距离的关系曲线

Fig. 1 The relationship between credible factor and mean square Euclidian distance when k taking different values

在确定了证据 E_i 的可信度因子 ε_i 后, 就可以对原

始证据进行修正. 设原始证据的BPA为 $m_i(A_j)$, 修正后的BPA为 $m''_i(A_j)$, 则

$$m''_i(A_j) = m_i(A_j) \cdot \varepsilon_i, A_j \neq \Theta, \quad (7a)$$

$$m''_i(\Theta) = 1 - \sum_{j=1}^N m_i(A_j) \cdot \varepsilon_i, A_j \neq \Theta. \quad (7b)$$

对原始证据进行修正后, 用D-S合成公式进行组合就可以得到合理的结果.

3 算例仿真(Simulation examples)

为了验证组合效果, 对本文提出的改进方法进行了数值仿真验算, 并与几种代表性的改进方法进行对比.

例1 在故障诊断中, 设识别框架 Θ 为故障A、故障B、故障C. 系统使用4种传感器对设备进行检测, 证据集 $E = \{E_1, E_2, E_3, E_4\}$, 各故障的BPA值如下:

$$m_1(A) = 0.98, m_1(B) = 0.01, m_1(C) = 0.01,$$

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

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

$$m_4(A) = 0.98, m_4(B) = 0, m_4(C) = 0.02.$$

D-S方法^[2]、Yager方法^[4]、孙全方法^[5]、李弼程方法^[6]及本文方法组合的结果如表1. 从表1可见看出: D-S方法和Yager方法的合成结果中故障A的概率分配值 $m(A)$ 始终为0, 显然不合常理; 孙全方法和李弼程方法虽然能合成高冲突证据, 但其收敛到一致信息的速度太慢, 4个证据合成时故障A的概率分配值 $m(A)$ 只有0.4234和0.7150; 而本文提出的方法不但能合成高冲突证据, 而且收敛速度也很快, 4个证据合成时故障A的概率分配值 $m(A)$ 达到0.9987.

表1 5种方法组合高冲突证据的结果

Table 1 Fusion results of highly conflict evidences with five methods mentioned above

	D-S	Yager	孙全	李弼程	本文
$m(A)$	0	0	0.4234	0.7150	0.9987
$m(B)$	0	0	0.0030	0.0050	0
$m(C)$	1	0.0001	0.1658	0.2800	0.0013
$m(\Theta)$	0	0.9999	0.4078	0	0

例1中, 直接用D-S方法组合时证据冲突系数 $k = 1.0000$, 用本文方法对证据源进行修正后证据冲突系数 $k = 0.7848$, 显然降低了奇异证据对组合结果的影响.

D-S理论组合高冲突证据时结果有悖常理, 而现有的一些改进方法虽能组合高冲突证据, 但组合较一致证据时的结果却不太合理. 实际应用中, 当组合的证据在不断变化时, 证据的冲突系数也在不断变化. 因此, 需要一种能对各种冲突强度证据进行有效

组合的方法. 例1已验证了本文所提出方法能有效地组合高冲突证据, 下面通过例2来检验该方法组合低冲突证据的能力.

例2 设识别框架 Θ 为疾病A、疾病B、疾病C. 4位专家 $\{E_1, E_2, E_3, E_4\}$ 认为病人患各种疾病的BPA值如下:

$$\begin{aligned} m_1(A) &= 0.6, m_1(B) = 0.1, m_1(C) = 0.3, \\ m_2(A) &= 0.7, m_2(B) = 0.2, m_2(C) = 0.1, \\ m_3(A) &= 0.5, m_3(B) = 0.2, m_3(C) = 0.3, \\ m_4(A) &= 0.65, m_4(B) = 0.1, m_4(C) = 0.25. \end{aligned}$$

可以看出, 4位专家都认为病人患疾病A的概率较大, 4个证据间冲突较小.

5种方法组合的结果如表2. 从表2可以看出: D-S方法组合低冲突证据时能取得令人满意的效果, 4个证据组合时, 命题A的概率分配值 $m(A)$ 迅速收敛到1; 使用Yager方法、孙全方法和李弼程方法对证据进行组合时, 命题A的概率分配值 $m(A)$ 不但不会收敛到1, 反而发散, 显然不合常理; 本文提出的方法对证据进行组合, 命题A的概率分配值 $m(A)$ 迅速收敛, 4个证据组合时和D-S合成方法得到的结果几乎一致. 可见, 本文提出的改进方法组合低冲突证据同样有效.

表2 5种方法组合低冲突证据的结果
Table 2 Fusion results of slowly conflict evidences with five methods mentioned above

	D-S	Yager	孙全	李弼程	本文
$m(A)$	0.9813	0.1365	0.4405	0.6638	0.9809
$m(B)$	0.0029	0.0004	0.0748	0.1295	0.0029
$m(C)$	0.0158	0.0022	0.1201	0.2067	0.0162
$m(\Theta)$	0	0.8609	0.3646	0	0

4 结束语(Conclusion)

由于人为或环境等因素的影响, 信息融合系统中个别或少数证据存在严重的奇异, 导致合成结果往往有悖常理. 本文提出的新方法通过引入证据可信度因子对原始证据源进行修正, 降低了奇异证据对组合结果的影响. 理论分析和数值实验结果表明, 该方法主要有以下特点: 1) 避免了一些改进方法靠人为给定证据权重系数对证据进行修正的主观因素的影响, 新方法得到的合成结果更可靠; 2) 不仅能有效组合高冲突证据, 同样适用于组合较一致的证据, 是一种较为通用的证据合成方法; 3) 新方法计算的时间复杂度有所增加, 即, 建立证据距离矩阵时增加时间复杂度 $O(n \times (n-1) \times N/2)$, 求均方欧式距离和可信度因子时增加时间复杂度 $O(2n)$, 最坏情况下,

最后对证据进行组合时增加时间复杂度 $O(n)$. 所以, 增加时间复杂度上限为 $O(n \times (n-1) \times N/2 + 3n)$, n 和 N 分别为识别框架中证据个数和命题个数, 但在证据数量不是很大时, 增加的计算量是可以容忍的.

参考文献(References):

- [1] DEMPSTER A P. Upper and low probabilities induced by a multi-valued mapping[J]. *Annals of Mathematical Statistics*, 1967, 38(6): 325 - 339.
- [2] SHAFER G. *A Mathematical Theory of Evidence*[M]. Princeton: Princeton University Press, 1976.
- [3] 张所地, 王拉娣. Dempster-Shafer合成法则的悖论[J]. 系统工程理论与实践, 1997, 17(5): 82 - 85. (ZHANG Suodi, WANG Ladi. Absurdities about Dempster-shafer's rule of combination of evidence[J]. *Systems Engineering-Theory & Practice*, 1997, 17(5): 82 - 85.)
- [4] YAGER R R. On the Dempster-Shafer frame work and new combination rules[J]. *Information Sciences*, 1987, 41(2): 93 - 138.
- [5] 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. (SUN Quan, YE Xiuqing, GU Weikang. A new combination rules of evidence theory[J]. *Acta Electronica Sinica*, 2000, 28(8): 117 - 119.)
- [6] 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. (LI Bicheng, WANG Bo, WEI Jun, et al. An efficient combination rule of evidence theory[J]. *Journal of Data Acquisition & Processing*, 2002, 17(1): 33 - 36.)
- [7] 陈增明, 梁昌勇, 蒋翠清, 等. 证据理论悖论消除的一种新方法[J]. 计算机应用研究, 2006, 23(11): 34 - 37. (CHEN Zengming, LIANG Changyong, JIANG Cuiqing, et al. New method of dispelling absurdities of Dempster-Shafer's rule of combination[J]. *Application Research of Computers*, 2006, 23(11): 34 - 37.)
- [8] 叶清, 吴晓平, 宋业新. 引入权重因子的证据合成方法[J]. 火力与指挥控制, 2007, 32(6): 21 - 24. (YE Qing, WU Xiaoping, SONG Yexin. An evidence combination method of introducing weight factors[J]. *Fire Control and Command Control*, 2007, 32(6): 21 - 24.)
- [9] 徐凌宇, 尹国成, 宫义山, 等. 基于不同置信度的证据组合规则及应用[J]. 东北大学学报(自然科学版), 2002, 23(2): 123 - 125. (XU Lingyu, YIN Guocheng, GONG Yishan, et al. Combination rules of various credibility evidences and application[J]. *Journal of Northeastern University(Natural Science)*, 2002, 23(2): 123 - 125.)
- [10] 梁昌勇, 陈增明, 黄永青, 等. Dempster-Shafer合成法则悖论的一种消除方法[J]. 系统工程理论与实践, 2005, 27(3): 7 - 12. (LIANG Changyong, CHEN Zengming, HUANG Yongqing, et al. A method of dispelling the absurdities of Dempster-Shafer's rule of combination[J]. *Systems Engineering-Theory & Practice*, 2005, 27(3): 7 - 12.)

作者简介:

张盛刚 (1983—), 男, 硕士研究生, 主要研究方向为车辆测试技术与故障诊断, E-mail: zhangshg1983@gmail.com;

李巍华 (1973—), 男, 副教授, 主要研究方向为设备状态监测、智能诊断与振动信号分析等, E-mail: whlee@scut.edu.cn;

丁康 (1957—), 男, 教授, 博士生导师, 主要研究方向为车辆性能测试、故障诊断、信号处理等, E-mail: kding@scut.edu.cn.