

一种处理冲突证据的 D-S 证据权重计算方法

费 翔, 周 健

(合肥工业大学计算机与信息学院, 合肥 230009)

摘 要: D-S 证据理论是一种有效处理不确定信息的方法, 在数据融合等领域得到了广泛应用。经典 D-S 证据理论在融合冲突较大的信息时通常会得到有悖于常理的结论。为此, 引入证据矛盾系数的概念, 提出一种能有效利用冲突信息的证据权重计算方法, 采用 D-S 证据理论组合规则进行数据融合。算例分析结果表明, 相比同类方法, 提出方法收敛速度更快, 融合结果更准确, 能克服经典 D-S 证据理论的不足, 提高融合结果的可信度。

关键词: D-S 证据理论; 冲突证据; 矛盾系数; 数据融合; 证据可信度

中文引用格式: 费 翔, 周 健. 一种处理冲突证据的 D-S 证据权重计算方法[J]. 计算机工程, 2016, 42(2): 142-145.

英文引用格式: Fei Xiang, Zhou Jian. A D-S Evidence Weight Computing Method for Conflict Evidence[J]. Computer Engineering, 2016, 42(2): 142-145.

A D-S Evidence Weight Computing Method for Conflict Evidence

FEI Xiang, ZHOU Jian

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

【Abstract】 D-S evidence theory is an effective method dealing with the dubious information and is widely used in the field of data fusion. There generates a wrong conclusion when using classical D-S evidence theory to handle the highly conflict evidence information. This paper introduces the concept of contradiction ratio between conflict evidences, uses the conflict information effectively and presents a new determine method for evidence weight. After that, it uses D-S evidence theory combination rules for data processing. The fusion results show the efficiency of the proposed method and provide a new way for D-S evidence theory algorithm improving.

【Key words】 D-S evidence theory; conflict evidence; contradiction ratio; data fusion; evidence credibility ratio

DOI: 10.3969/j.issn.1000-3428.2016.02.026

1 概述

Dempster 在 1967 年提出多值映射^[1]理论, 随后 Shafer 针对已有的成果进行研究, 提出了 Shafer 证据理论, 并将 Dempster 合成规则推广到一般的情况, 提出了经典的 Dempster-Shafer 证据理论, 即 D-S 证据理论^[2]。作为一种不确定推理方法, D-S 证据理论能够准确描述不确定以及不知道信息, 相比于其他推理方法, 可以在不具备先验信息的基础上处理不确定信息。D-S 证据理论自提出以来, 因为其能对不确定事件给出不确定概率, 广泛应用于不确定信息决策^[3-5]、模式识别^[6]、多传感器信息融合^[7-8]等领域。本文基于证据矛盾程度的思想, 引入证据矛盾系数的概念, 提出一种利用冲突信息的证据权重计算方法, 采用 D-S 证据理论组合规则进行数据融合。

2 D-S 证据理论及其存在的问题

2.1 D-S 证据理论

在 D-S 证据理论中, 识别框架 Θ 指所研究对象的全集, Θ 元素之间相互排斥且为离散值。由识别框架 Θ 中所有子集构成的集合称为 Θ 的幂集, 记作 2^Θ 。若存在映射 $m: 2^\Theta \rightarrow [0, 1]$, 且满足:

$$m(\phi) = 0; \sum_{A \subseteq \Theta} m(A) = 1 \quad (1)$$

则称符合上述条件的映射 m 为框架 Θ 上的一个基本概率指派 (Basic Probability Assignment, BPA) 函数。 $\forall A \subseteq \Theta$, 函数 $m(A)$ 反映了证据体对事件 A 的相信程度。对于空集 ϕ , 有 $m(\phi) = 0$, 表示对空集不产生任何信任; 对于 Θ 的全体子集, 有 $\sum_{A \subseteq \Theta} m(A) = 1$, 表示所有事件信任度之和必须为 1。此外, 若 $m(A) > 0$, 则

基金项目: 国家自然科学基金资助项目“基于复杂网络理论的网络安全量化评估模型研究”(60873194)。

作者简介: 费 翔 (1990 -), 男, 硕士研究生, 主研方向为网络安全; 周 健, 教授、博士。

收稿日期: 2015-02-09 **修回日期:** 2015-03-11 **E-mail:** 352448696@qq.com

称 A 为识别框架 Θ 中的一个焦元。

设识别框架 Θ 上有 2 个独立证据体 E_1, E_2 。 m_1, m_2 分别为其对应的 BPA 函数,焦元分别是 $A_1, A_2, \dots, A_n$ 和 $B_1, B_2, \dots, B_n$, D-S 证据理论组合规则为:

$$m(C) = \begin{cases} \frac{\sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1 - K} & C \neq \phi \\ 0 & C = \phi \end{cases} \quad (2)$$

其中, K 为矛盾因子,其计算公式为: $K = \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j)$, 且满足条件 $0 \leq K \leq 1$ 。矛盾因子 K 与 2 条证据之间的冲突程度有关,当 K 的值越接近 1 时,证据体之间的冲突程度越大,融合结果越不准确;当 K 的值越接近 0,证据体之间的冲突程度越小,融合结果越准确;当 $K = 1$ 时,证据体之间相互矛盾,不能对其进行有效融合。

2.2 D-S 证据理论存在的问题

作为一种不确定知识的推理方法, D-S 证据理论被广泛应用于不确定信息的推理问题。但经典的 D-S 证据理论运用存在一定的局限性,即只有融合的证据体冲突较小时,才能得到较好的融合结果。但是实际运用中经常遇到证据之间存在较大的冲突甚至证据之间完全矛盾时,这种情况下 D-S 证据理论则采取完全丢弃冲突信息的方法,往往得到违背直观的结果,主要可以分为以下 4 种情形:

(1) 一般冲突问题^[9]

例 1 有识别框架 $\Theta = \{A, B, C\}$, 已知 2 个证据体 E_1 和 E_2 , BPA 函数如下: $m_1: m_1(A) = 0.99, m_1(B) = 0.01, m_1(C) = 0, m_2: m_2(A) = 0, m_2(B) = 0.01, m_2(C) = 0.99$ 。

显然 m_1 和 m_2 是高度冲突的,利用 DS 证据理论进行融合后得到。可以发现,尽管证据 $m_1(B) = m_2(B) = 0.01$, 2 个证据对 B 的信任度都很小,融合后 B 的信任度却为 1,相反本来拥有高信任度的 A 和 C ,融合后的信任度却为 0,这是有悖于常理的。

(2) 一票否决问题^[9]

设识别框架 $\Theta = \{A, B, C\}$, 当证据体彻底否定焦元时,该命题的融合结果也被一票否决。对于上文的例子,有 $m_1(A) = 0$, 则不论其余的证据体对 A 的信任度如何,其融合结果仍然为 $m(A) = 0$,与现实不符。

(3) 融合规则失效问题

当冲突系数 $K = 1$ 时,此时 2 个证据体的证据完全冲突, D-S 证据理论组合公式的分母为 0,组合公式失效, D-S 证据理论不能对其进行融合。

(4) 鲁棒性问题^[9]

鲁棒性反映了某个焦元的 BPA 函数发生细微

变动时,其融合结果产生变化的程度。对例 1 的信任函数做微小变动可得到下面的例子。

例 2 有识别框架 $\Theta = \{A, B, C\}$, 已知 2 个证据体 E_1 和 E_2 , BPA 函数如下: $m_1: m_1(A) = 0.99, m_1(B) = 0.01, m_1(C) = 0, m_2: m_2(A) = 0.01, m_2(B) = 0.01, m_2(C) = 0.98$ 。

可以看出,当证据源做微小的变动, $m_2(A)$ 由 0 变为 0.01, $m_2(C)$ 由 0.99 变为 0.98, 融合结果 $m(A)$ 却由 0 变为 0.99, 发生了巨大的变化。

针对经典的 D-S 证据理论存在的问题,国内外的专家学者进行了大量研究,研究成果基本上可以分为 2 大类:第 1 类主要针对组合规则的改进,认为 DS 证据理论存在的问题,只有修改 D-S 证据理论合成规则才能处理高度冲突的问题,这一类的代表文献有文献[10-11];第 2 类则是对概率分配函数的修正,认为 D-S 证据理论不存在任何问题,需要修正的是证据源,这一类的代表文献有文献[12-14]。

3 处理冲突证据的新方法

D-S 证据理论在面对高度冲突证据会产生与直觉相反的融合结果,根本原因在于其丢弃了冲突部分的信息。该规则为了保持基本概率函数的归一性,完全遗弃了矛盾因子,使得 2 个证据的每个公共焦元的基本概率赋值变为原来的 $(1 - K)^{-1}$ 倍,这在高冲突下会产生不合理的结果^[15]。应当意识到,冲突本身也是一种信息,对冲突信息的有效利用,可以提高融合结果的准确性。因此,提出了矛盾系数的概念,用矛盾系数描述对于某个具体的命题,证据之间的冲突程度。冲突系数的定义如下:

定义 有识别框架 $\Theta = \{A_1, A_2, \dots, A_n\}$, 2 个证据体 E_1, E_2 对于命题 A_k 的 BPA 函数分别为 $m_1(A_k)$, $m_2(A_k)$, 则这 2 条证据的冲突系数如式(3)所示。

$$C_{i,j}(A_k) = \left(\frac{m_1(A_k) - m_2(A_k)}{m_1(A_k) + m_2(A_k)} \right)^2 \quad (3)$$

从上式可以看出,当 $m_1(A_k)$ 等于 $m_2(A_k)$ 时,矛盾系数 $C_{1,2}(A_k) = 0$, 说明 2 个证据不存在冲突;当 $m_1(A_k)$ 和 $m_2(A_k)$ 不等时, $C_{1,2}(A_k)$ 随着 $m_1(A_k)$ 和 $m_2(A_k)$ 冲突的增大而增大,且在相同矛盾冲突的情况下,矛盾系数随着 $m_1(A_k)$ 和 $m_2(A_k)$ 值的增大而减小;当 $m_1(A_k)$ 和 $m_2(A_k)$ 中任意一个为 0 时,冲突系数为 1,说明证据间存在高度冲突。

对识别框架中的任意一个命题 A_k , 当存在 n 条证据时,命题 A_k 的矛盾矩阵定义如下:

$$G_{A_k} = \begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,n} \\ \vdots & \vdots & & \vdots \\ C_{n,1} & C_{n,2} & \cdots & C_{n,n} \end{pmatrix}$$

对同一命题 A_k , 矛盾矩阵反映了不同证据相互之间的冲突程度。每个证据与其他证据的相互矛盾程度可以利用矛盾矩阵来计算, 如果该证据对于该命题的判断可信较低, 矛盾程度较高, 最后合成时该证据对应的权重也应相应的较小。

每条证据的相互矛盾系数为:

$$G_i(A_k) = \sum_{j=1}^{j=n} C_{i,j}(A_k) \quad (4)$$

在此用 $G_i(A_k)$ 描述证据的基本概率分配函数 m_i 与其他证据的矛盾程度, $V_i(A_k)$ 用来描述在对命题 A_k 合成时 BPA 函数 m_i 所占的权重。当每条证据都与其他证据高度矛盾, 则期望证据之间的矛盾系数为 1, 这时相互矛盾系数 $G_i(A_k)$ 为 n , 该证据的不可信度即 $\frac{G_i(A_k)}{n}$, 因此该证据的权重 $V_i(A_k)$ 的计算公式如下:

$$V_i(A_k) = 1 - \frac{G_i(A_k)}{n} \quad (5)$$

改进后的融合权重的 DS 证据理论组合规则如下:

$$m(C) = \begin{cases} \frac{\sum_{A_i \cap B_j = C} m_1(A_i) V_1(A_i) \cdot m_2(B_j) V_2(B_j)}{1 - K} & C \neq \phi \\ 0 & C = \phi \end{cases}$$

其中, $K = \sum_{A_i \cap B_j = \phi} m_1(A_i) V_1(A_i) \cdot m_2(B_j) V_2(B_j)$, 反映了改进后证据之间的冲突程度。

4 算例分析

为了验证本文提出方法的有效性, 将经典 D-S 方法、文献[10]方法、文献[11]方法、与本文提出的方法进行分析比较。主要分析各方法对上文所述的 D-S 证据理论不足之处的改进情况。

设 $\Theta = \{A, B, C\}$, 其 3 个证据体的 BPA 函数如下:

$$\begin{aligned} m_1: m_1(A) &= 0.98; m_1(B) = 0.01; m_1(C) = 0.01 \\ m_2: m_2(A) &= 0; m_2(B) = 0.01; m_2(C) = 0.99 \\ m_3: m_3(A) &= 0.99; m_3(B) = 0; m_3(C) = 0.01 \end{aligned}$$

根据矛盾系数的定义, 对命题 A 进行数据处理, 其矛盾系数为:

$$\begin{aligned} C_{1,1}(A) &= 0; C_{1,2}(A) = 1 \\ C_{1,3}(A) &= \left(\frac{0.98 - 0.99}{0.98 + 0.99} \right)^2 = 0.000\ 026 \end{aligned}$$

可得命题 A 证据矛盾矩阵如下:

$$G_A = \begin{pmatrix} 0 & 1 & 0.000\ 026 \\ 1 & 0 & 1 \\ 0.000\ 026 & 1 & 0 \end{pmatrix}$$

按照相互矛盾系数的定义, 可以求得每条证据

的相互矛盾系数为: $D_1(A) = 0 + 1 + 0.000\ 026 = 1.000\ 026$; $D_2(A) = 1 + 0 + 1 = 2$; $D_3(A) = 0.000\ 026 + 1 + 0 = 1.000\ 026$ 。

再依据权重的计算公式, 可计算得出每条证据的权重为: $V_1(A) = 0.666\ 7$; $V_2(A) = 0$; $V_3(A) = 0.666\ 7$ 。

用求得的权重对基本概率分布函数进行再处理, 可得新的 BPA 函数: $m_1(A) = 0.653\ 3$; $m_2(A) = 0$; $m_3(A) = 0.660\ 0$ 。

同理对命题 B 和 C 进行同样处理, 经过处理后新的 BPA 函数为:

$$\begin{aligned} m_1: m_1(A) &= 0.653\ 3; m_1(B) = 0.006\ 7; m_1(C) = 0.006\ 8; m_1(\Theta) = 0.333\ 2; \\ m_2: m_2(A) &= 0; m_2(B) = 0.006\ 6; m_2(C) = 0.359\ 8; m_2(\Theta) = 0.469\ 8; \\ m_3: m_3(A) &= 0.659\ 9; m_3(B) = 0; m_3(C) = 0.006\ 8; m_3(\Theta) = 0.333\ 2。 \end{aligned}$$

最后对处理后的基本概率指派函数用 D-S 证据理论进行融合, 并与直接运用 D-S 证据理论、文献[10]方法以及文献[11]方法结果进行比较, 比较结果如表 1~表 3 所示。

表 1 $m_1 m_2$ 融合结果比较

合成方法	$m_{12}(A)$	$m_{12}(B)$	$m_{12}(C)$	$m_{12}(\Theta)$
DS 证据理论	0.000 0	0.010 0	0.990 0	0.000 0
文献[10]方法	0.000 0	0.000 1	0.010 0	0.990 0
文献[11]方法	0.180 0	0.003 0	0.189 0	0.625 0
本文方法	0.325 7	0.013 0	0.332 4	0.328 9

表 2 $m_2 m_3$ 融合结果比较

合成方法	$m_{23}(A)$	$m_{23}(B)$	$m_{23}(C)$	$m_{23}(\Theta)$
DS 证据理论	0.000 0	1.000 0	0.000 0	0.000 0
文献[10]方法	0.000 0	0.000 1	0.000 0	0.999 9
文献[11]方法	0.182 1	0.003 8	0.182 1	0.632 0
本文方法	0.328 9	0.013 6	0.328 9	0.328 9

表 3 $m_1 m_2 m_3$ 融合结果比较

合成方法	$m_{123}(A)$	$m_{123}(B)$	$m_{123}(C)$	$m_{123}(\Theta)$
DS 证据理论	0.000 0	0.000	0.000 0	1.000 0
文献[10]方法	0.000 0	0.000	0.000 1	0.999 9
文献[11]方法	0.445 9	0.006	0.148 6	0.399 4
本文方法	0.827 7	0.003	0.063 4	0.106 0

结合表 1、表 2 可以看出, 经典的 DS 证据理论在处理高度冲突的数据时会产生悖论, 融合结果与实际不符。与经典的 D-S 证据理论类似, 文献[10]方法融合的结果也为 0, 且其结果过于保守, 分配给未知领域的信任度过大。除此之外, 经典 D-S 证据理论和文献[10]方法都没有解决鲁棒性的问题, 当证据源发生微小改变时, 其融合结果发生了很大变化。本文提出的方法在鲁棒性上优于上述方法, 而且在

证据冲突冲突的情况下,融合结果的分配精度也比较理想。通过表 3 可以看出,当有新证据源加入时,经典 D-S 证据理论和 Yager 方法对命题 A 的融合结果仍然为 0,存在一票否决的情况。文献[11]方法有效解决了一票否决的问题,但是收敛速度慢,分配精度不够,未知领域 θ 没有明显下降。相比文献[11]方法,本文提出的方法随着支持命题 A 证据的增加,融合结果收敛速度更快,未知领域 θ 的范围更小,较好地反映了实际情况下的融合结果,可靠性较高。

5 结束语

针对 D-S 证据理论在处理高度冲突证据时面临的规则失效问题,本文基于矛盾系数,提出了一种新的改进方法,在一定程度上弥补了 D-S 证据理论的不足。新的方法改进了经典 D-S 证据理论丢弃冲突信息的做法,充分利用了各证据体之间的冲突信息,采用冲突信息求得各证据间的矛盾系数,再计算各证据对于命题的权值,应用权值信息对加权基本概率分布函数进行改进,最后对证据进行融合。算例分析结果证明,相比其他方法,本文方法更符合人们的客观预期,融合结果的收敛速度更快。

参考文献

- [1] Dempster A P. Upper and Lower Probabilities Induced by a Multiplied Mapping[J]. Annals of Mathematical Statistics, 1967, 38(2): 325-339.
- [2] Shafer G. A Mathematical Theory of Evidence[M]. New Jersey, USA: Princeton University Press, 1976.
- [3] Wu Zhiwei. Attribute Reduction Based on Evidence Theory in Incomplete Decision Systems[J]. Information

Sciences, 2008, 178(5): 1355-1371.

- [4] Hua Zhongsheng, Gong Bengang, Xu Xiaoyan. A DS-AHP Approach for Multi-attribute Decision Making Problem with Incomplete Information[J]. Expert System with Application, 2008, 34(3): 2221-2227.
- [5] 陈赠明. 群决策环境下证据理论决策方法应用研究[D]. 合肥: 合肥工业大学, 2006.
- [6] Lin Tzu-chao. Partition Belief Median Filter Based on Dempster-Shafer Theory for Image Processing[J]. Pattern Recognition, 2008, 41(1): 139-151.
- [7] Otman B, Yuan Xiaohong. Engine Fault Diagnosis Based on Multisensor Information Fusion Using Dempster-Shafer Evidence Theory[J]. Information Fusion, 2007, 8(4): 379-386.
- [8] Zeng Dehuai, Xu Jianmin, Xu Gang. Data Fusion for Traffic Incident Detection Using D-S Evidence Theory with Probabilistic SVMs[J]. Journal of Computers, 2008, 3(10): 36-43.
- [9] 丁迎迎, 李洪瑞. 一种简单有效的处理冲突证据的 D-S 改进方法[J]. 指挥控制与仿真, 2011, 32(2): 22-25.
- [10] Yager R R. On the Dempster-Shafer Framework and New Combination Rules[J]. Information Sciences, 1987, 41(2): 93-137.
- [11] 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.
- [12] Murphy C K. Combining Belief Functions When Evidence Conflicts[J]. Decision Support Systems, 2000, 29(1): 1-9.
- [13] 熊彦铭, 杨战平, 屈新芬. 基于模型修正的冲突证据组合新方法[J]. 控制与决策, 2011, 26(6): 883-887.
- [14] 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.
- [15] 许红波, 丁建江, 胡伟稿. DS 规则推广及其在飞机目标识别中的应用研究[J]. 雷达与对抗, 2006, 1(1): 34-38.

编辑 顾逸斐

(上接第 141 页)

- [8] Nguyen M H, Ong S J, Vadhan S. Statistical Zero-knowledge Arguments for NP from Any One-way Function[C]//Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science. Washington D. C., USA: IEEE Press, 2006: 3-14.
- [9] Stinson D R. Cryptography: Theory and Practice[M]. Boca Raton, USA: CRC Press, 2005.
- [10] Rivest R L, Shamir A, Adleman L. A Method for Obtaining Digital Signatures and Public-key Cryptosystems[J]. Communications of the ACM, 1978, 21(2): 120-126.

- [11] ElGamal T. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms[J]. IEEE Transactions on Information Theory, 1985, 31(4): 10-18.
- [12] Itakura K, Nakamura K. A Public-key Cryptosystem Suitable for Digital Multisignatures[J]. NEC Research & Development, 1983, (71): 1-8.
- [13] 王晓峰, 张璟, 王尚平. 多重数字签名方案及其安全性证明[J]. 计算机学报, 2008, 31(1): 176-183.

编辑 陆燕菲