

一种改进的证据合成公式

肖明珠^{1,2}, 陈光¹

(1. 电子科技大学自动化工程系, 四川成都 610054; 2. 中国工程物理研究院电子工程研究所, 四川绵阳 621900)

摘要: 本文讨论了 Dempster 公式、Yager 公式和其他几种证据合成公式存在的不足, 提出了一种改进的证据合成公式. 该公式可合成多路冲突性和不冲突性证据, 冲突性证据按加权平均合成, 不冲突性证据按按与运算合成并反映了相互间交叉融合程度. 实验表明该公式能合成多种包括高度冲突的证据, 具有很好的实用性.

关键词: 证据理论; 合成公式; 数据融合

中图分类号: O21; TP274 **文献标识码:** A **文章编号:** 0372 2112 (2005) 09-1714 03

A Modified Combination Rule of Evidence Theory

XIAO Ming-zhu^{1,2}, CHEN Guang-ju¹

(1. Dept. of Automation Engineering, University of Electronics Science and Technology of China, Chengdu, Sichuan 610054, China;

2. Institute of Electronic Engineering, China Academy of Engineering Physics, Mianyan, 621900, China)

Abstract: This paper discusses the shortcomings of Dempster's rule, Yager's rule and other rules, and then presents a modified combination rule to deal with consistence or inconsistency evidences obtained from multiple sources. The modified rule adapts AND operation to combine consistent evidences and reflects the intersection of focal elements, and allocates the conflict probability to very inconsistency focus element according to its average supported degree. Experiments show that the new combination rule is very reliable and rational for all kinds of evidences including highly conflicting evidences.

Key words: evidence theory; combination rules; data fusion

1 引言

D-S 理论也叫证据理论, 是一种关于不确定性推理的理论, 不仅可以用来处理偶然性不确定性问题, 还可以用来处理认识性不确定性问题. D-S 理论的一个重要内容就是能够从多种证据源获取的证据, 并对证据间的冲突性模型化.

D-S 理论中最早的证据合成公式是 Dempster 公式, Dempster 公式特别强调多种证据的协调性, 抛弃所有冲突的证据. 在采用 Dempster 公式合成高度冲突的证据时, 合成结果常有悖常理. 于是一些学者提出了新的修改公式, 如 Yager 合成公式、加权平均公式, 文[4] 和文[5] 中的公式等. 这些公式的一些重要区别是对证据冲突性的处理上.

本文在分析以上合成公式的基础上提出了一种新的合成公式, 该公式全面考虑了冲突性证据和不冲突性证据的合成, 冲突性证据按加权平均合成, 不冲突性证据按交叉融合程度合成.

2 D-S 理论的证据合成公式

设 X 为论域上的有限全集, $P(X)$ 为 X 上的幂集, 在 D-S 理论下, 有下列函数关系成立:

$$m: P(X) \rightarrow [0, 1] \quad (1)$$

$$m(\phi) = 0 \quad (2)$$

$$\sum_{A \in \mathcal{A}} m(A) = 1 \quad (3)$$

这里 m 称为基本分配, 对于任何一个集合 $A \in P(X)$, 如果 $m(A) \neq 0$, 则称为焦元. 所有焦元的集合记为 F , 一对 (F, m) 便称为一个证据体.

D-S 理论中的证据合成是对来自多个相互独立的证据源的数据进行简化和总结, 是一种信息综合, 类似的综合技术包括算术平均、几何平均、最大值、最小值等.

D-S 理论中最早的证据合成公式是 Dempster 公式, Dempster 证据合成公式如下:

$$m(\phi) = 0 \quad (4)$$

$$m(A) = \frac{1}{1-K} \sum_{A_i \cap A_j \cap A_k \cap \dots = A} m_1(A_i) m_2(A_j) m_3(A_k) \dots \quad (5)$$

$$\text{其中 } K = \sum_{A_i \cap A_j \cap A_k \cap \dots = \phi} m_1(A_i) m_2(A_j) m_3(A_k) \dots$$

Dempster 合成公式中 $1-K$ 是归一化因子, 所起的效果是完全抛弃证据间的冲突性并把所有与冲突相关的概率分配于空集, 这是一种严格的与运算. 因此在合成高度冲突的证据时, 合成结果常有悖常理^[3~5].

为了解决高度冲突证据合成问题, Yager 对 Dempster 公式作了改进, 提出了 Yager 合成公式^[3].

$$m(\phi) = 0 \tag{6}$$

$$m(A) = \sum_{A_i \cap A_j \cap A_k \dots = A} m_1(A_1) m_2(A_j) m_3(A_k) \dots A \neq \phi, X \tag{7}$$

$$m(X) = \sum_{A_i \cap A_j \cap A_k \dots = X} m_1(A_1) m_2(A_j) m_3(A_k) \dots + K \tag{8}$$

其中 K 与 Dempster 公式中定义一样.

Yager 公式对于不冲突性的证据仍采用了与运算合成方式, 但没有抛弃所有冲突性证据, 而是把支持证据冲突的那部分概率全部赋予了未知领域. 虽然 Yager 公式能合成高度冲突的证据, 但由于对冲突性证据是完全否定的, 因而在证据源多于两个时, 合成的效果有时并不理想^[4, 5]. 于是文[4]、文[5]对冲突性证据分别提出了改进的处理措施.

文[4]引入了一种证据可信度的概念, 对冲突性证据采用加权和形式. 文[4]认为, 即使证据间存在冲突, 它们也是可用的, 并且可用度取决于证据的可信度.

文[5]合成公式基本上与文[4]一样, 只不过是废弃了证据可信度的概念, 把支持证据冲突的概率按各个命题的平均支持程度加权进行分配.

Dempster 公式、Yager 公式、文[4]公式、文[5]公式对不冲突性的证据都采用了与运算的合成方式, 文[4]和文[5]对冲突性证据采用了平均加权技术. 总的来说, 两种技术的综合运用是比较科学的, 即解决了冲突证据的利用处理问题, 又保证了不冲突性证据聚焦趋向合成. 但是在以上几种合成公式中, 都没有考虑证据间的相互交叉和融合程度. 从直觉判断上, 证据间的交叉融合程度不一样, 其合成证据的信任度也应不一样.

文[6]虽然提出了反映证据交叉融合程度的聚焦系数, 但文[6]提出的合成公式只适用 2 路证据合成, 对多路证据合成则无能为力, 而且文[6]的聚焦系数等于证据集合基的比例, 是平均性运算关系, 这违背了不冲突性证据与运算的合成法则, 因而不足取.

3 一种新的证据合成公式

综合以上的讨论, 提出了一种修改的证据合成公式. 设 A_i, B_j, C_k, \dots , 分别是 N 个不同证据源的焦点. 则

$$m(\phi) = 0 \tag{9}$$

$$m(A) = \frac{1-K}{P} \sum_{A_i \cap A_j \cap A_k \dots = A} \frac{|A|}{|A_i| |B_j| |C_k| \dots} m_1(A_i) m_2(A_j) m_3(A_k) \dots A \neq \phi \tag{10}$$

其中

$$K = \sum_{A_i \cap B_j \cap C_k \dots = \phi} m_1(A_i) m_2(B_j) m_3(C_k) \dots \tag{11}$$

$$P = \sum_{A_i \cap B_j \cap C_k \dots = \phi} \frac{|A_i \cap B_j \cap C_k \dots|}{|A_i| |B_j| |C_k| \dots} m_1(A_i) m_2(B_j) m_3(C_k) \dots \tag{12}$$

$$Q(A) = \frac{1}{n} \sum_{1 \leq i \leq n} m_1(A) \tag{13}$$

以上合成公式充分考虑了冲突性和不冲突性证据的合成

要求, 冲突性证据按加权平均合成, 不冲突性证据按与运算合成, 并且反映了证据间的交叉融合程度.

K 表示证据间总的冲突程度; $Q(A)$ 反映证据对 A 的平均支持程度; $\frac{|A_i \cap B_j \cap C_k \dots|}{|A_i| |B_j| |C_k| \dots}$ 反映证据交叉融合程度, 具有与运算关系; P 表示不冲突证据交叉融合后总的概率.

4 实验结果与应用讨论

下面以 2 个例子来说明本文所提出的合成公式的实验效果.

例 1 以文[5]给出的例 1 为例. 设 $\Theta = \{A, B, C\}$, $m_1(A) = 0.99$, $m_1(B) = 0.01$, $m_2(B) = 0.01$, $m_2(C) = 0.99$.

利用 Dempster 公式、Yager 公式、文[4]公式、文[5]公式及本文公式合成的结果如表 1 所示.

表 1 5 种公式对冲突性证据的合成结果

	$m_{12}(A)$	$m_{12}(B)$	$m_{12}(C)$	$m_{12}(\Theta)$
Dempster	0	1	0	0
Yager	0	0.0001	0	0.9999
文[4]	0.182	0.004	0.182	0.632
文[5]	0.495	0.01	0.495	0
本文	0.495	0.01	0.495	0

例 1 是一个高度冲突的证据, 冲突程度 $K = 0.9999$. 由表 1 可以看出, 尽管 m_1 和 m_2 对 B 的支持率很低, 但合成结果却认为 B 为真, 显然有悖常理. Yager 公式虽然把冲突的证据完全否定了, 但把支持证据冲突的那部分概率全部赋予了未知领域, 随意抛弃了信息. 文[4]公式没有完全否定冲突证据, 所得结果只有 0.182, 与常理有一定差距. 按理说, 既然两个证据几乎各支持一个命题, 那么合成结果对每一个命题的支持率都接近 0.5. 本文合成的结果与文[5]的结果很好地接近了这一数值. 通过本例, 可以得知本文的合成公式在处理冲突性证据方面比前三种公式优越, 与文[5]公式具有一样的效果.

例 2 设有 2 个一致性证据(Consonant Evidence), $\Theta = \{a, b, c, d, e, f, g, h\}$, $m_1(a) = 0.5$, $m_1(a, b, c) = 0.4$, $m_1(a, b, c, d, e, f, g, h) = 0.1$, $m_2(a, b) = 0.3$, $m_2(a, b, c, d, e) = 0.5$, $m_2(a, b, c, d, e, f, g, h) = 0.2$

利用 Dempster 公式、Yager 公式、文[4]公式、文[5]公式及本文公式合成的结果如表 2 所示.

表 2 5 种公式对一致性证据的合成结果

	m_{12} (a)	m_{12} (a, b)	m_{12} (a, b, c)	$m_{12}(a, b, c, d, e)$	$m_{12}(a, b, c, d, e, f, g, h)$
Dempster	0.5	0.15	0.28	0.05	0.02
Yager	0.5	0.15	0.28	0.05	0.02
文[4]	0.5	0.15	0.028	0.05	0.02
文[5]	0.5	0.15	0.028	0.05	0.02
本文	0.3427	0.3364	0.3364	0.0249	0.0779

从表 2 的数据来看, 前四种公式的合成结果是一样的, 因为它们都没有考虑证据之间的交叉重合程度, 对一致性证据失去了细辨能力. 本文合成的 $m_{12}(a)$ 比其他 4 种的要小, 而剩下焦点的分配概率比其他 4 种相应要大. 这是由于焦点

$\{a\}$ 与其他焦元交叉融合的程度相对要小的原因,因此本文提出的合成公式能反映了证据交叉融合的程度.通过证据 1 的 $m_1(a) = 0.5$ 和证据 2 的 $m_2(a, b) = 0.3$, 凭直觉判断合成后的 $m_{12}(a)$ 应小于 0.5 而 $m_{12}(a, b)$ 应大于 0.3, 只有本文合成的结果很好地符合这一判断.

本文的合成公式可广泛应用于数据融合、目标识别、人工智能推理、故障诊断、多传感器信号测量、风险评估等领域.本文的合成公式最佳的应用环境为处理冲突性和非冲突性的数据,协调不同专家的意见,综合反映大多数证据广泛一致的意见.当然,寻找一个万能的合成公式是不现实的,本文的合成公式与其他公式是相互补充完善的关系,需要在实际中灵活运用.

5 结论

本文提出了一种修改的证据合成公式,该公式没有抛弃和否定冲突性证据,而是同时兼顾了冲突性和非冲突性证据,冲突性证据按加权平均合成,非冲突性证据按与运算合成,并且反映了证据间的交叉融合程度.通过本文的 2 个实验表面该公式具有很好的实用效果.

参考文献:

- [1] Shafer G. A Mathematical Theory of Evidence [M]. Princeton U P, Princeton, 1976.
- [2] 朱大奇, 于盛林. 基于 D S 证据理论的数据融合算法及其在电路故障中的应用 [J]. 电子学报, 2002, 30(2): 221- 223.
ZHU Da qi, YU Sheng lin. Data fusion algorithm based on DS theory

and its application for circuit fault diagnosis [J]. Acta Electronica Sinica, 2002, 30(2): 221- 223. (Chinese Source)

- [3] Yager R R. On the D-S framework and new combination rules [J]. Information Sciences, 1987, 41(2): 93- 138.
- [4] 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 Q, Ye X Q, Gu W K. A new combination rules of evidence theory [J]. Acta Electronica Sinica, 2000, 28(8): 117- 119. (Chinese Source)
- [5] 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): 34- 36.
- [6] 向阳, 史习智. 证据理论合成规则的一点修正 [J]. 上海交通大学学报, 1999, 33(3): 357- 360.

作者简介:



肖明珠 男, 1970 年出生于江西省吉水县, 现为电子科技大学博士生, 从事测量技术研究.

陈光福 男, 1939 年出生于四川, 现为电子科技大学教授, 博士生导师, 中国电子学会会士, 英国 IEE 会员, 多年来, 在现代测试理论及方法, 计算机辅助测试, 故障诊断技术与可测性设计、专家系统以及测试系统结构体系等进行了卓有成效的研究工作, 在国内外核心期刊物发表学术论文 100 多篇.