

一种改进的证据合成规则

蒲书缙, 杨 雷, 杨莘元, 胡炜薇

(哈尔滨工程大学信息与通信工程学院, 哈尔滨 150001)

摘 要: 就 D-S 证据合成规则对于冲突证据合成存在的不足, 不少学者基于证据冲突概率再分配思想, 提出了相应的修正证据合成规则。针对这些修正合成规则的不足, 提出了一种改进的合成规则, 该规则在综合考虑证据源可信度和证据支持度的基础上, 分析证据集的有效性, 合理地实现信息融合。数值实验结果表明, 新的合成规则不仅克服了现有规则的不足, 而且其结果向合理性方向收敛度好, 提高了系统的鲁棒性。

关键词: 证据理论; 合成规则; 证据冲突; 证据有效性; 数据融合

A Modified Combination Rule of Evidence Theory

PU Shujin, YANG Lei, YANG Shenyuan, HU Weiwei

(School of Information and Communication Engineering, Harbin Engineering University, Harbin 150001)

【Abstract】 According to the insufficiencies existing in D-S evidence theory when they deal with the conflicting evidences, many scholars present modified rules base on reallocation of the basic probability assigned to conflict. A new combination rule is presented to mend the disadvantages of the rules. This rule analyzes the validity of evidences according to the extent of conflict with the use of the reliability of evidential sources and support measure of the evidence, and realizes information fusion rationally. Experiments show that the new combination rule overcomes the insufficiencies, and the outputs are constringent on the aspect of rationality, which enhances robustness of systems effectively.

【Key words】 Evidence theory; Combination rule; Evidence conflict; Evidence validity; Data fusion

在D-S证据理论中,证据合成思想是建立在认为证据完全可靠的前提下,在实际问题中并不一定具备这种前提条件,特别是环境复杂多变的情况下,传感器可能提供很不相同、甚至是相互矛盾的数据。对于证据冲突的情况,D-S理论的合成公式存在不足,合成结果有悖常理。尽管修正的方法有很多^[1-7],有些方法也得到了较好的效果,但这些方法大多没有考虑系统实际应用背景,不能保证融合结果向确定和可靠的方向收敛。本文分析了D-S理论及几种针对冲突证据合成的修正证据理论的不足,在此基础上,提出了一种新的合成规则。

1 D-S 证据理论规则及冲突证据处理的局限

在 D-S 证据理论中首先定义的是一个识别框架 Θ , 由一些互斥且穷举的元素组成。对于问题域中的命题都应包含于集函数 $m: 2^\Theta \rightarrow [0,1]$, 满足: (1) $m(\emptyset) = 0$; (2) $\sum_{A \subset \Theta} m(A) = 1$ 。 m 为识别框架 Θ 上的基本概率分配函数: $\forall A \subset \Theta$, $m(A)$ 称为 A 的基本置信度。

D-S 证据合成是将证据集划分为 2 个或多个不相关的部分, 并利用 Dempster 组合规则将它们组合起来。若用 2 个证据进行组合, 则其组合规则为

$$m(C) = \begin{cases} \frac{\sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1-k} & \forall C \subset \Theta, C \neq \emptyset \\ 0 & C = \emptyset \end{cases} \quad (1)$$

式中: $k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) < 1$, 它是一个衡量用于融合的各

个证据之间冲突程度的系数, 也称证据的冲突概率。在证据严重冲突的情况下, Dempster 组合结果往往与实际情况不相符合。主要原因是: 对置信度实行的是“逻辑与”的操作,

不可避免“一票否决权”问题; 将冲突概率 k 舍去并且对置信度进行了归一化, 丢弃了冲突部分的信息量。

2 冲突概率的再分配思想及对应证据合成规则

对于高度冲突和完全冲突证据的融合, 单一地否定冲突命题或者单一地否定冲突命题中的某一个, 所得的融合结果都可能是不正确的。冲突证据组合关键是如何合理地冲突信息, 通常的方法是分配冲突证据所携带的冲突概率, 这种冲突概率再分配思想的合成规则可以统一表示为

$$m(\emptyset) = 0 \\ m(A) = \sum_{A_i \cap B_j \cap C_j \dots = A} m_1(A_i) \cdot m_2(B_j) \cdot m_3(C_j) \dots + f(A) \quad (2)$$

$$\forall A \neq \emptyset$$

$f(A)$ 为证据冲突概率的分配函数, 满足: $f(A) \geq 0$, $f(\emptyset) = 0$

且

$$\sum_{A \subset \Theta} f(A) = k \quad (3)$$

$$k = \sum_{A_i \cap B_j \cap C_j \dots = \emptyset} m_1(A_i) \cdot m_2(B_j) \cdot m_3(C_j) \dots \quad (4)$$

基于冲突概率再分配的组合规则的主要代表有:

(1) Yager 合成规则

Yager^[2,3]认为, 尽管有多个证据提供了关于命题的置信度, 但是由冲突证据提供的关于命题的置信度是不可靠的, 因此对整个命题仍然所知甚少。据此, Yager 去掉了 D-S 组合规则中的归一化过程, 把冲突证据携带的那部分置信度赋给了未知命题。所以, 在 Yager 组合规则中取 $f(\emptyset) = k$,

作者简介: 蒲书缙(1980 -), 女, 博士生, 主研方向: 数据融合, 目标识别, 智能信息处理; 杨 雷, 博士生; 杨莘元, 教授、博导; 胡炜薇, 博士生

收稿日期: 2005-12-21 **E-mail:** psj1980@126.com

$$f(A) = 0 \quad (A \neq \Theta).$$

该规则把冲突信息转换成信息融合后的不确定性。其缺点是：1)其结论过于保守，完全否定了冲突信息的可用性，且对证据的扰动非常敏感；2)当冲突的证据数目超过2个时，无法获得合理的结果，组合规则的鲁棒性仍然不好。

(2)修正合成规则 I

为了修正 Yager 组合规则的不足，文献[4]提出了一种证据合成规则，把支持证据冲突的概率按各个命题的平均置信度加权进行分配，即令

$$q(A) = \frac{1}{n} \sum_{1 \leq i \leq n} m_i(A) \quad (5)$$

$$f(A) = k \cdot q(A) \quad (6)$$

该规则忽略了证据冲突对命题合成置信度的影响，没有区分冲突和不冲突证据，对所有的证据都进行平均处理，没有起到压缩和减少错误信息及加强有用和正确信息的作用。

(3)修正合成规则 II

文献[5,6]认为，尽管证据之间存在冲突，但它们是部分可用的，并且这种可用程度取决于证据的可信度。在这种证据合成公式中：

$$f(A) = k \cdot \varepsilon \cdot q(A) \quad \forall A \neq \emptyset, \Theta \quad (7)$$

$$f(\Theta) = k \cdot \varepsilon \cdot q(\Theta) + k(1 - \varepsilon) \quad (8)$$

式中， ε 表示了证据集的有效性系数， $q(A)$ 表示命题 A 平均置信度， $k(1 - \varepsilon)$ 表示冲突概率剩余部分，直接分配给未知命题，它表示了证据的不可靠使得融合结果不确定性的增加。

3 一种改进的冲突概率再分配证据合成规则

事实上，造成证据冲突的原因是多方面的，证据的合成要因事而异，针对实际问题的特点，尤其是对冲突的证据不能全盘否定或者片面地选取。在建立证据合并规则时，要考虑以下几个方面的因素：(1)证据源的可信度；(2)证据的支持度；(3)证据合成向合理性方向收敛取决于证据源的可信度和证据支持度对证据集有效性的影响。根据以上分析，本文在冲突概率再分配的证据合成公式基础上，对冲突分配函数进行了修正。

3.1 证据集有效性分析

3.1.1 证据源的可信度

在实际工作系统中，识别框架的基本概率分配是由不同的传感器(证据源)提供的，由于多种因素的影响、知识掌握的有限性、传感器自身性能的不足，因此证据源本身是存在质量问题的。证据源的可信度表示了证据源的质量。若 crd_i 表示第 i 个证据源的可信度，将其归一化，就可得到各证据源的相对可信度，下文中的证据源可信度都是指相对可信度。

$$acrd_i = crd_i / \sum_{i=1}^n crd_i \quad i = 1, \dots, n \quad (9)$$

3.1.2 证据的支持度

证据支持度是指，对于同一识别框架的 n 个证据，其中一条证据信息被其他 $n-1$ 条证据支持的程度，即表示这条证据被其他证据看来是真实的可能程度。对于证据支持度的估计，采用了证据距离^[7]。

定义 m_1 和 m_2 是在辨识框架 Θ 上的 2 个基本概率分配，则 m_1 和 m_2 的证据距离可以表示为

$$d_{BPA}(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)D(m_1 - m_2)} \quad (10)$$

D 为一个 $2^N \times 2^N$ 阶矩阵，其中的元素为

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad A, B \in P(\Theta) \quad (11)$$

$P(\Theta)$ 是 Θ 所有子集生成的空间。证据之间的相互支持度是其证据距离的函数，且满足证据距离越小，证据之间的支持度就越高。则证据的相互支持度为

$$sup_{ij} = Sup(m_i, m_j) = 1 - d_{BPA}(m_i, m_j) \quad (12)$$

可以看到， $sup_{i,j}$ 仅是证据 i 被证据 j 支持的程度，而证据 i 的支持度应 $sup_{i,1}, sup_{i,2}, \dots, sup_{i,n}$ ($i \neq j$)，和证据源的可信度共同决定的，因为可信度较大的证据之间的支持程度比可信度较小的证据之间的支持程度对证据结果影响更大。设 sup_i 为证据 i 的加权平均支持程度为

$$sup_i = \sum_{j \neq i}^n acrd_j \cdot sup_{i,j} \quad (13)$$

进行其归一化处理，得到证据的支持度为

$$Sup_i = sup_i / \sum_{i=1}^n sup_i \quad (14)$$

3.1.3 证据集有效性系数

对一个证据集有效性的度量主要取决于证据之间的冲突程度。若假设 m_1, m_2, \dots, m_n 所对应的证据集为 F_1, F_2, \dots, F_n ，证据集 i 和 j 之间的冲突大小为 k_{ij} ，证据集的冲突程度是这 $n(n-1)/2$ 个局部冲突的函数。计算证据集冲突程度时，应考虑个证据源可信度和证据支持度影响，这种影响主要反映在局部冲突有效性上，所以证据集冲突的计算方式可采用线性加权。这里，权重体现的是局部冲突的可用程度和有效性，其满足性质为：(1)证据的综合支持度和证据源的可信度越高，局部冲突的有效性就高，权重也越高；(2)只要任一证据源的可信度为 0，则局部冲突无效，权重为 0。

从证据之间的支持度来看，对于第 s 个证据对证据 i 和证据 j ，其局部冲突被支持度为

$$Sup_{ij}^s = (sup_i + sup_j) / 2 \quad (15)$$

将其进行归一化处理：

$$W_{ij}^s = Sup_{ij}^s / \sum_{q=1}^{n(n-1)/2} Sup_{ij}^q \quad s = 1, \dots, n(n-1)/2 \quad (16)$$

而局部冲突的可信度取决于 2 条证据间的相对可信度：

$$crd_{ij}^s = \min(acrd_i, acrd_j) / \max(acrd_i, acrd_j) \quad (17)$$

则，局部冲突的有效性为

$$\omega_{ij}^s = W_{ij}^s \cdot crd_{ij}^s / \sum_{q=1}^{n(n-1)/2} W_{ij}^q \cdot crd_{ij}^q \quad (18)$$

为了从整个证据集的基础上表现其有效性，将其进行正则化处理。

那么证据集的冲突程度为

$$\tilde{k} = \sum_{q=1}^{n(n-1)/2} \omega_{ij}^q k_{ij}^q \quad i, j \leq n \quad (19)$$

这里 \tilde{k} 反映证据集总体上的冲突程度，与 D-S 理论中的 k 不同，当 k 大时， \tilde{k} 不一定大。那么证据集有效性系数为

$$\varepsilon = e^{-\tilde{k}} \quad (20)$$

3.2 改进的冲突分配函数

对于多个证据，尤其是冲突证据的合成，为了保证其结果向合理性方向收敛，要在全面衡量证据集的基础上，对其冲突概率进行有效的分配。这种改进的证据组合规则，其冲突概率的分配函数形式仍如式(7)、式(8)，而参数 ε 由式(15)~式(20)确定。对于平均置信度 $q(A)$ 的计算，应考虑证据在证

据集中的相对可信度，将对应命题的基本置信度进行加权平均：

$$q(A) = \sum_{1 \leq i \leq n} t_i m_i(A) \quad (21)$$

其中， t_i 为证据的相对可信度

$$t_i = \text{acrd}_i \cdot \text{Sup}_i / \sum_{i=1}^n \text{acrd}_i \cdot \text{Sup}_i \quad (22)$$

4 数值实验分析

若在一个多传感器系统中，系统提供 6 个证据和对应证据源的可信度：

$$m_1 : m_1(A) = 0.011, m_1(B) = 0.98, m_1(C) = 0.009, \text{crd}_1 = 0.81$$

$$m_2 : m_2(A) = 0.005, m_2(B) = 0.005, m_2(C) = 0.99, \text{crd}_2 = 0.9$$

$$m_3 : m_3(A) = 0.05, m_3(B) = 0.97, m_3(C) = 0.025, \text{crd}_3 = 0.8$$

$$m_4 : m_4(A) = 0.01, m_4(B) = 0, m_4(C) = 0.99, \text{crd}_4 = 0.91$$

$$m_5 : m_5(A) = 0.02, m_5(B) = 0.98, m_5(C) = 0, \text{crd}_5 = 0.82$$

$$m_6 : m_6(A) = 0.005, m_6(B) = 0.99, m_6(C) = 0.005, \text{crd}_6 = 0.83$$

利用文中所介绍的几种证据合成公式进行证据融合，其中对于修正规则 II 分为 2 种情况：仅考虑证据的可信度和仅考虑证据的支持度，所得的结果如表 1 所示。

表 1 6 种合成公式的证据组合结果

合成公式	证据	k	\bar{k}	ε	$m(A)$	$m(B)$	$m(C)$	$m(\Theta)$
D-S	m_1, m_2	0.990 5			0.000 6	0.051 7	0.947 8	0.000 0
	m_1, m_2, m_3	0.999 3			0.000 4	0.678 7	0.320 9	0.000 0
	m_1, m_2, m_3, m_4	0.999 8			0.000 0	0.000 0	1.000 0	0.000 0
	m_1, m_2, m_3, m_4, m_5	0.999 9			1.000 0	0.000 0	0.000 0	0.000 0
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9			1.000 0	0.000 0	0.000 0	0.000 0
Yager 规则	m_1, m_2	0.990 5			0.000 0	0.000 5	0.009 0	0.990 5
	m_1, m_2, m_3	0.999 3			0.000 0	0.000 5	0.000 2	0.999 3
	m_1, m_2, m_3, m_4	0.999 8			0.000 0	0.000 0	0.000 2	0.999 8
	m_1, m_2, m_3, m_4, m_5	0.999 9			0.000 0	0.000 0	0.000 1	0.999 9
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9			0.000 0	0.000 0	0.000 1	0.999 9
修正规则 I	m_1, m_2	0.990 5			0.005 7	0.486 1	0.508 2	0.000 0
	m_1, m_2, m_3	0.999 3			0.020 5	0.650 2	0.344 3	0.000 0
	m_1, m_2, m_3, m_4	0.999 8			0.017 9	0.487 5	0.505 9	0.000 0
	m_1, m_2, m_3, m_4, m_5	0.999 9			0.018 3	0.586 1	0.404 6	0.000 0
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9			0.016 1	0.653 4	0.338 0	0.000 0
修正合成规则 II (可信度)	m_1, m_2	0.990 5	0.990 5	0.371 4	0.002 1	0.180 8	0.194 4	0.622 7
	m_1, m_2, m_3	0.999 3	0.712 7	0.490 3	0.010 0	0.319 0	0.168 9	0.509 3
	m_1, m_2, m_3, m_4	0.999 8	0.686 8	0.503 2	0.009 0	0.245 3	0.254 7	0.496 7
	m_1, m_2, m_3, m_4, m_5	0.999 9	0.631 5	0.531 8	0.009 7	0.311 7	0.215 2	0.468 2
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9	0.567 9	0.566 7	0.009 1	0.370 3	0.191 5	0.433 3
修正合成规则 II (支持度)	m_1, m_2	0.990 5	0.990 5	0.371 4	0.002 1	0.180 8	0.194 4	0.622 7
	m_1, m_2, m_3	0.999 3	0.559 0	0.571 8	0.011 7	0.372 0	0.197 0	0.427 9
	m_1, m_2, m_3, m_4	0.999 8	0.687 6	0.502 8	0.009 0	0.245 1	0.254 5	0.497 0
	m_1, m_2, m_3, m_4, m_5	0.999 9	0.595 5	0.551 3	0.010 1	0.323 1	0.223 0	0.448 7
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9	0.492 8	0.610 9	0.009 8	0.399 2	0.206 5	0.389 1
本文的修正规则	m_1, m_2	0.990 5	0.891 4	0.410 1	0.002 3	0.199 6	0.213 7	0.584 3
	m_1, m_2, m_3	0.999 3	0.504 2	0.604 0	0.018 1	0.575 1	0.024 4	0.395 7
	m_1, m_2, m_3, m_4	0.999 8	0.612 5	0.542 0	0.008 9	0.232 1	0.306 5	0.457 9
	m_1, m_2, m_3, m_4, m_5	0.999 9	0.539 4	0.583 1	0.011 9	0.397 0	0.180 3	0.416 9
	$m_1, m_2, m_3, m_4, m_5, m_6$	0.999 9	0.450 5	0.637 3	0.011 8	0.512 7	0.118 6	0.362 7

由表 1 可见，对有冲突的证据集进行融合，D-S 规则的结果是不合理的；Yager 规则完全否定所有的证据，夸大了证据冲突的程度；修正规则 I 部分否定证据冲突，忽略了冲突所造成的不确定性，且平滑了证据之间的冲突。尽管，修正规则 II 克服了以上的不足，既承认了证据冲突给融合结果造成的不确定性，即 $m(\Theta)$ ，同时又消除了证据对命题的“一票否决权”现象，当更多的证据支持命题 B 时， $m(B)$ 的值在升高， $m(\Theta)$ 在减小。然而，修正规则 II 仅从单方面分析了证据冲突，其向合理性方向收敛的幅度小。

本文提出的证据合成规则，考虑了证据源的可信度和证据的支持度，能根据证据集特点，得到合理的融合结果，有效地降低了不可靠证据源对融合结果的影响，提高了向合理性方向的收敛度，减小了融合结果的不确定性，保证了系统的可靠性和鲁棒性。

5 结论

本文在证据冲突概率再分配的基础上提出的修正证据合成规则，结合实际信源可信度和证据的支持度，从证据冲突程度的角度分析了证据集的有效性，合理地分配了冲突概率。实验结果证明，新规则不仅有效地合成了高冲突证据，其效果

果优于 D-S 理论规则及现有的冲突概率再分配组合规则。

参考文献

- 1 Smets P. The Combination of Evidence in the Trans-ferable Belief Model[J]. IEEE Transactions on Patten Analysis and Machine Intelligence, 1990, 12(5): 447- 458.
- 2 Yager R R. On the Dempster-Shafer Framework and New Combination Rules[J]. Information Science, 1989, 41(2): 93-137.
- 3 Yager R R. On the Aggregation of Prioritized Belief STructure. IEEE Trans. on Syst. Man. and Cybern., 1996, 26(6): 708-719.
- 4 李弼程, 王 波. 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.
- 5 孙 全, 叶秀清. 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.
- 6 邓 勇, 施文康. 一种改进的证据推理组合规则[J]. 上海交通大学学报, 2003, 37(8): 1 275-1 278.
- 7 Chen Liangzhou, Shi Wenkang, Deng Yong. A New Fusion Approach Based on Distance of Evidences[J]. Zhejiang University Science, 2005, 6(5): 476-482.