

D-S 证据理论中一种新的冲突证据融合方法

蒋黎明 何加浪 张宏

(南京理工大学计算机科学与技术学院 南京 210094)

摘要 针对 D-S 证据理论在融合高度冲突的证据时可能导致与直观结果相悖的问题,通过分析现有的几种冲突证据融合方法,提出了一种有效处理冲突证据的融合方法。该方法综合了模型修正和规则修正这两类方法的优点。首先基于冲突系数矩阵计算证据的可信度,并以此作为折扣因子来修正证据结构,然后基于局部冲突、局部分配原则确定证据合成时的冲突分配空间和分配权重。实验结果表明,该方法对合理的合成方向具有较快的收敛速度,同时降低了决策风险。

关键词 D-S 证据理论,组合规则,冲突,可信度,折扣因子

中图分类号 TP18 **文献标识码** A

New Fusion Approach for Conflicting Evidence in D-S Theory of Evidence

JIANG Li-ming HE Jia-lang ZHANG Hong

(College of Computer, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract A new method was developed for dealing with seriously conflicting evidence when the Dempster combination result can not identify the actual conditions. The method combines the advantages of justification of evidence information and combination rules. The reliability of evidence, computed by a conflict matrix, is taken as discount factor to correct the belief functions. And then the problem of space and weight for conflict redistribution is addressed based on the idea of distributing local conflict to local focal elements. Compared with the existing methods, the new method can applied to both cases of conflicting and coincidence. And the numerical simulation shows that the method improves the reliability and rationality of combination results, and also accelerates convergence and reduces the decision risk.

Keywords D-S evidence theory, Combination rule, Conflict, Reliability, Discount factor

1 引言

D-S 证据理论^[1,2]是 20 世纪 70 年代发展起来的一种数学工具,它是对贝叶斯理论的一种扩展,主要优点是:不需要先验信息,对不确定信息的描述采用“区间”的方法,解决了关于不确定性信息的表示方法,在区分“不知道”与“不确定”方面以及精确反映证据聚合方面显示出很大的灵活性,因而在多源信息融合、目标识别和决策分析等领域得到广泛应用。

Dempster 定义的组合规则^[3]是融合和更新证据信息的有效方法。然而,当证据源高度冲突时,Dempster 组合规则会得出有悖于常理的结论。为了解决这一问题,国内外许多学者针对 D-S 证据理论在冲突证据合成方面的缺陷做了大量修正工作^[5-9]。可概括为两大类:一是基于修正融合模型的方法,首先对冲突证据进行预处理,然后用证据组合规则融合证据,典型代表有折扣系数法和加权平均法;二是基于修正组合规则的方法,主要解决冲突的分配空间和权重问题,典型代表有全局分配方法和局部分配方法。

综合考虑上述两类方法,本文提出一种新的组合方法:首先对证据进行修正和预处理,同时基于局部冲突、局部分配原

则确定一种合理的权重分配方法。

2 Dempster 组合规则及失效问题

设 Bel_1 和 Bel_2 是同一识别框架 Θ 上的信度函数, m_1 和 m_2 分别是其对应的基本概率分配函数, 焦点分别为 A_i 和 B_j ($i=1, 2, \dots, n; j=1, 2, \dots, m$), 则按照 Dempster 组合规则有

$$m(\emptyset) = 0 \tag{1}$$

$$m(A) = \frac{1}{1 - k} \sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j), A \neq \emptyset \tag{2}$$

$$k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) \tag{3}$$

式中, k 称为基本概率分配函数 m_1 和 m_2 之间的冲突系数。如果 $k \rightarrow 0$, m_1 和 m_2 之间没有冲突; 当 $k \rightarrow 1$ 时, m_1 和 m_2 之间严重冲突, 此时应用 Dempster 规则进行证据合成, 会得到与直觉相悖的结果^[4]。

3 D-S 证据理论中证据冲突的改进方法

3.1 典型的修正融合模型方法分析

3.1.1 折扣系数法

当某一证据的似真度为 $t \in (0, 1)$, 则采用折扣操作对的

到稿日期:2010-05-06 返修日期:2010-08-23 本文受国家自然科学基金资助项目(90718021), 自主科研专项计划(2010ZYTS035), 高等学校博士学科点专项科研基金(20093219120024)资助。

蒋黎明(1983-), 博士生, 主要研究方向为动态信任建模和人工智能, E-mail: njustilm@163.com; 何加浪(1984-), 博士生, 主要研究方向为软件修复与进化计算; 张宏(1956-), 教授, 博士生导师, 主要研究方向为网络安全与人工智能。

BPA 函数进行处理。Shafer 在文献[2]中首次提出一种通用的用于 BPA 函数的折扣操作,折扣处理后的 BPA 函数为:

$$\begin{cases} m^*(A) = tm(A) & \forall A \subseteq \Theta \\ m^*(\Theta) = 1 - t + tm(\Theta) \end{cases} \quad (4)$$

在实际应用中,不同证据的可靠程度也不一样,并随时间及上下文环境的改变而变化。因而对不同的证据信息而言,其折扣因子 t 也不是固定不变的。而文献[2]中并没有给出如何确定参数 t 值的方法。

3.1.2 Murphy 平均法

文献[5]中, Murphy 提出一种收敛速度比较快的方法。当系统中有 N 条证据时, Murphy 规则首先将计算所有证据对识别框架中的命题支持度的平均值,然后用 D-S 合并规则自身迭代 $N-1$ 次。

$$\bar{m}(X) = \frac{1}{N} \sum_{i=1}^N m_i(X) \quad \forall X \subseteq \Theta \quad (5)$$

$$\tilde{m}(A) = \frac{\sum_{B \cap C = A} \bar{m}(B) \bar{m}(C)}{1 - \bar{k}} \quad (6)$$

$$\bar{k} = \sum_{B \cap C = \emptyset} \bar{m}(B) \bar{m}(C) \quad (7)$$

但 Murphy 方法只是将各证据源的数据作简单的算术平均,并未考虑各证据源的数据可靠性程度的不同,因而 Han^[6] 等人提出了基于粗糙集的证据权值计算方法。

3.2 典型的修正组合规则方法分析

3.2.1 Yager 规则

Yager^[7] 分析了 D-S 理论的失效问题,并提出这样一个观点:既然对冲突证据无法给出合理的抉择,就应该将冲突证据全部赋给未知项。改进后的合成公式如下:

$$m(\emptyset) = 0 \quad (8)$$

$$m(A) = \sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j), A \neq \emptyset \text{ 且 } A \neq \Theta \quad (9)$$

$$m(\Theta) = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) + k \quad (10)$$

尽管此公式可以合成高冲突的证据,但由于 Yager 方法将冲突信息全部分配给全集,造成了冲突信息分配不公平。当合成多个高冲突的证据时,合成结果往往并不理想。

3.2.2 Dubois 规则 (DP)

Dubois 和 Parade 等人^[8] 提出一种新的局部冲突重分配方法,即将 BBA 值赋给冲突焦元的并集:

$$m(\emptyset) = 0 \quad (11)$$

$$m(A) = \sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j) + \sum_{\substack{C_i \cup D_j = A \\ C_i \cap D_j = \emptyset}} m_1(C_i) m_2(D_j) \quad (12)$$

$$\forall A \subseteq \Theta, A \neq \emptyset$$

DP 规则将冲突分配给冲突焦元的并集,增加了证据推理的不确定性,但也没有区分对待 $m_1(C_i) m_2(D_j)$ 和 $m_2(C_i) m_1(D_j)$, 使得冲突分配不合理。

3.2.3 概率冲突分配规则^[9,10]

Smarandache 等人^[9] 提出了一种新的冲突证据组合规则 PCR5, 该规则把支持冲突证据的概率在引发证据冲突的焦元之间进行加权分配。在对两个证据源进行融合时,其规则表达如下:

$$m_{PCR}(\emptyset) = 0 \quad (13)$$

$$m_{PCR}(A) = \sum_{\substack{A_i \cap B_j = A \\ A_i, B_j \subseteq \Theta}} m_1(A_i) m_2(B_j) + \sum_{\substack{A \cap B_j = \emptyset \\ B_j \subseteq \Theta}} \frac{m_1(A)^2 m_2(B_j)}{m_1(A) + m_2(B_j)} + \sum_{\substack{A \cap B_j = \emptyset \\ B_j \subseteq \Theta}} \frac{m_2(A)^2 m_1(B_j)}{m_2(A) + m_1(B_j)}$$

$$A \neq \emptyset \quad (14)$$

文献[10]提出了相类似的方法。这类方法虽然在计算冲突分配权重上比前几种方法更合理,但它也忽视了可靠性低的证据对合成结果的影响。

上述冲突证据分配方法中,都存在一定的不足,它们仅仅关注冲突重分配空间和比例问题,忽视了造成证据冲突的一个重要问题——证据源不可靠。所以,在证据组合时,需要考虑证据的可信度信息,尽可能降低可信度低的证据源对融合结果的影响,降低决策风险。但是,不同的证据源的可靠性程度不同,还存在上下文和时变效应,所以不同的证据源应具有不同的折扣因子。本文提出一种新的冲突证据组合规则,首先基于冲突系数矩阵计算证据的可信度,并以此作为各证据源的折扣因子来修正证据结构,进一步确定了冲突分配的权值。本文方法进一步提高了冲突分配的准确度并降低了决策风险。

3.3 本文方法

设 Θ 为一个包含 N 个互斥命题的识别框架, $m_i(\cdot)$ 和 $m_j(\cdot)$ 分别是证据 E_i 和 E_j 作用于 Θ 所产生的基本概率分配,焦元分别为 A_1, \dots, A_k 和 B_1, \dots, B_l , 则证据 E_i 和 E_j 的冲突系数为:

$$k_{ij} = \begin{cases} \sum_{\substack{A_i \cap B_j = \emptyset \\ A_i, B_j \subseteq \Theta}} m_i(A_i) m_j(B_j), & i \neq j \\ 0, & i = j \end{cases} \quad (15)$$

若针对同一识别框架 Θ 的证据数目为 n , 则可利用式 (15) 计算出任意两证据 E_i 和 E_j 的冲突系数 k_{ij} , 并可进一步表示为一个冲突矩阵 K 。

$$K = \begin{matrix} & \begin{matrix} E_1 & E_2 & \cdots & E_n \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{matrix} & \begin{bmatrix} 0 & k_{12} & \cdots & k_{1n} \\ k_{21} & 0 & \cdots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \cdots & 0 \end{bmatrix} \end{matrix} \quad (16)$$

本文认为,一个证据和其它证据间的冲突越小,则其它证据对它的支持度就越大。所以,我们首先给出局部支持度和全局支持度的概念。

如果证据 E_i 和 E_j 之间的冲突为 k_{ij} , 则可以得到证据 E_i 对证据 E_j 的局部支持度 δ_i^j , 即:

$$\delta_i^j = \begin{cases} \frac{1 - k_{ij}}{1 - k_{ij} + k_{ij} \times k_{ij}}, & i \neq j \\ 1, & i = j \end{cases} \quad (17)$$

通过累加其它证据对 E_i 的局部支持度,得到 E_i 的全局支持度 δ_i , 证据的全局支持度越高,该证据越可信。反之,该证据的可信度就低。在求出 δ_i 之后,可以得到证据 E_i 的相对可信度 ϕ_i :

$$\phi_i = \frac{\delta_i}{\sum_j \delta_j} \quad (18)$$

相对可信度最高的证据是绝对可信的,即绝对可信度为 1, 所以任意证据的绝对可信度可以表示为:

$$\omega_i = \frac{\phi_i}{\max_{j=1,2,\dots,n} \{\phi_j\}}, i=1,2,\dots,n \quad (19)$$

基于上述分析与结论,本文提出了一种新的组合方法,其步骤归纳如下:

(1) 对于 n 条组合证据,采用式 (15) 一式 (19) 计算每条证

据的绝对可信度 $\omega_i (i=1, 2, \dots, n)$;

(2) 基于可信度对原始证据进行修正, 如式(20):

$$m_i(A) = \begin{cases} \omega_i m_i(A), & A \subseteq \Theta \\ 1 - \sum_{B \subseteq \Theta} \omega_i m_i(B), & B \subseteq \Theta, A = \Theta \end{cases} \quad (20)$$

(3) 新的组合规则为:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \sum_{\substack{B \cap C = A \\ B, C \subseteq \Theta}} m_1(B) m_2(C) + \sum q(A) \end{cases} \quad (21)$$

其中,

$$q(A) = \frac{m_1(A)^3}{m_1(A)^3 + m_2(X)^3} m_1(A) m_2(X) + \frac{m_2(A)^3}{m_2(A)^3 + m_1(X)^3} m_2(A) m_1(X), \quad \forall A \subseteq \Theta, A \neq \emptyset \quad (22)$$

$q(A)$ 为各局部冲突中分配给 A 的 BPA 值。新的合成规则区分了证据间的一致性部分和冲突部分。对于一致性部分, 我们采用合取规则; 对于冲突部分, 局部冲突在引发冲突的焦点间进行按权值分配, 确定权值的依据是修正后证据的 BPA 值。

4 数值算例分析

我们讨论一个多传感器系统对敌方飞机的类型 (A, B 或 C) 的识别结果仿真。并通过与 Dempster, Yager, Dubois, Murphy, PCR5 等证据合成方法进行对比, 以说明本文方法的有效性。

表 1 不同证据合成方法之间的比较

方法	E_1, E_2, E_3	E_1, E_2, E_3, E_4	E_1, E_2, E_3, E_4, E_5
Dempster ^[3]	$m(A)=0.2167$	$m(A)=0.3978$	$m(A)=0.4829$
	$m(B)=0.7330$	$m(B)=0.5758$	$m(B)=0.5083$
	$m(C)=0.0498$	$m(C)=0.0263$	$m(C)=0.0087$
	$m(\Theta)=0.0005$	$m(\Theta)=0.0001$	$m(\Theta)=0.0001$
Yager ^[7]	$m(A)=0.0031$	$m(A)=0.0019$	$m(A)=0.0009$
	$m(B)=0.0448$	$m(B)=0.0090$	$m(B)=0.0031$
	$m(C)=0.0011$	$m(C)=0.009$	$m(A)=0.0001$
	$m(\Theta)=0.9510$	$m(\Theta)=0.973$	$m(\Theta)=0.9959$
Dubois ^[8]	$m(A)=0.0031$	$m(A)=0.0019$	$m(A)=0.0009$
	$m(AB)=0.2568$	$m(AB)=0.2329$	$m(AB)=0.2031$
	$m(B)=0.0448$	$m(B)=0.0090$	$m(B)=0.0031$
	$m(AC)=0.2172$	$m(AC)=0.1530$	$m(AC)=0.0921$
	$m(C)=0.0011$	$m(C)=0.0001$	$m(C)=0.0001$
PCR5 ^[9]	$m(BC)=0.1094$	$m(BC)=0.0375$	$m(BC)=0.0178$
	$m(\Theta)=0.3676$	$m(\Theta)=0.5656$	$m(\Theta)=0.6829$
	$m(A)=0.4775$	$m(A)=0.6543$	$m(A)=0.6887$
	$m(B)=0.3478$	$m(B)=0.2547$	$m(B)=0.2683$
	$m(C)=0.1747$	$m(C)=0.0909$	$m(C)=0.0430$
Murphy ^[5]	$m(\Theta)=0$	$m(\Theta)=0$	$m(\Theta)=0$
	$m(A)=0.4694$	$m(A)=0.7096$	$m(A)=0.7897$
	$m(B)=0.4045$	$m(B)=0.2403$	$m(B)=0.1937$
	$m(C)=0.1258$	$m(C)=0.0001$	$m(C)=0.0166$
本文	$m(\Theta)=0.0003$	$m(\Theta)=0.9890$	$m(\Theta)=0$
	$m(A)=0.6371$	$m(A)=0.8459$	$m(A)=0.9028$
	$m(B)=0.2838$	$m(B)=0.1419$	$m(B)=0.0919$
	$m(C)=0.0787$	$m(C)=0.0121$	$m(C)=0.0053$
	$m(\Theta)=0.0004$	$m(\Theta)=0.0001$	$m(\Theta)=0$

例 1 (目标识别分析) 辨识框架为 $\Theta = \{A = \text{轰炸机}, B = \text{民航客机}, C = \text{战斗机}\}$, 在某一时刻利用传感器的观测信息构造的 5 个证据如下:

$$\begin{cases} E_1: m_1(A)=0.6, m_1(B)=0.32, m_1(C)=0.06, m_1(\Theta)=0.02 \\ E_2: m_2(A)=0.01, m_2(B)=0.35, m_2(C)=0.6, m_2(\Theta)=0.04 \\ E_3: m_3(A)=0.52, m_3(B)=0.4, m_3(C)=0.03, m_3(\Theta)=0.05 \\ E_4: m_4(A)=0.6, m_4(B)=0.2, m_4(C)=0.1, m_4(\Theta)=0.1 \\ E_5: m_5(A)=0.5, m_5(B)=0.35, m_5(C)=0.1, m_5(\Theta)=0.05 \end{cases}$$

各种方法的识别结果的对比如表 1 所列。由表 1 可见, Dempster^[3] 组合公式无法处理冲突情况; Yager^[6] 规则过于保守, 不利于决策; Dubois^[7] 的办法一定程度上可以处理冲突, 通过向上“散焦”暂缓了决策, 显得过于保守, 有时无法处理冲突; 只有在系统收集到 4 个证据时, PCR5^[8], Murphy^[5] 中的方法才可以分辨出目标。而本文的方法在收到第 3 个证据时, 就可以分辨出目标了。在本文的方法中, 通过参考证据的可信度等全局信息来修正证据源, 起到了信息“滤波”的作用。同时在将局部冲突进行分配的过程中, 选择了合理的分配空间和权重, 在最大程度地减少干扰证据对融合结果和决策的影响, 提高收敛速度的同时, 也降低了决策风险, 在证据比较少少的情况下就可以收敛为正确的目标。

结束语 本文研究了 D-S 证据理论的冲突证据融合问题, 通过比较现有的各种改进方法和思路, 综合考虑了融合模型修正和组合规则修正两类方法的优势, 提出了一种新的冲突证据融合方法。通过数值算例分析, 与几种典型方法相比, 结果表明本文提出的融合方法可以有效处理证据冲突问题, 具有较快的收敛速度, 同时降低了决策风险, 提高了在证据冲突时融合结果的可靠性和合理性。

参考文献

- [1] Dempster A. Upper and lower probabilities induced by multivalued mapping[J]. Annals of Mathematical Statistics, 1967, 38: 325-339
- [2] Shafer G. A Mathematical Theory of Evidence [M]. Princeton University Press, 1976
- [3] Dempster A P. A generalization of Bayesian inference [J]. Journal of the Royal Statistical Society, 1968, 30: 205-245
- [4] Zadeh L A. Review of Shafer's mathematical theory of evidence [J]. AI Magazine, 1984, 5: 205-245
- [5] Murphy C. Combining of Belief Functions When Evidence Conflicts [J]. Decision Support Systems, 2000, 29(1): 1-9
- [6] Han Li, Shi Li-ping. Approach to evidence combination based on rough set [C] // 2009 International Conference on Electronic Computer Technology. Macau, China, 2009: 693-697
- [7] Yager R. On the Dempster Shafer Framework and New Combination Rules [J]. Information Sciences, 1987, 41(2): 93-137
- [8] Dubois D, Prade H. Representation and Combination of Uncertainty with Belief Functions and Possibility Measures [J]. Computational Intelligence, 1998(4): 244-264
- [9] Smarandache F, Dezert J. Proportional conflict redistribution rules for information fusion[M]. Smarandache F, Dezert J, eds. Vol. 2. Rehoboth: American Research Press, 2006: 3-68
- [10] Heeyoul C, Seungjin C, Yoonsuck C. Probabilistic Combination of Multiple Evidence[C] // Proceedings of the 16th International Conference on Neural Information Processing. Bangkok, Thailand, 2009: 302-311