

文章编号: 1000-6893(2008)

广义证据理论及应用(3): 新的冲突表示模型

邓勇^{1,4}, 蒋雯², 韩德强³

(1. 上海交通大学 电子信息与电气工程学院, 上海 200240)

(2. 西北工业大学 电子信息学院, 陕西 西安 710072)

(3. 西安交通大学 综合自动化研究所, 陕西 西安 710049)

(4. 西南大学 计算机与信息科学学院, 重庆 400715)

Generalized Evidence Theory and Its Application (3): A Novel Model to represent Conflict

Deng Yong^{1,4}, Jiang Wen², Han Deqiang³

(1. School of Electronics Information and Electric Engineering, Shanghai Jiao Tong University, Shanghai, 200240)

(2. School of Electronics and Information, Northwestern Polytechnical University, Xi'an, 710072)

(3. Institute of Integrated Automation, Xi'an Jiao Tong University, Xi'an, 710049)

(4. College of computer and information sciences, Southwest University, Chongqing, 400715)

摘要: 早期关于冲突证据的研究主要集中在修改 Dempster 组合规则, 最新的研究表明证据理论中的经典冲突系数无法合理度量证据之间的冲突程度并引入新的参数来, 但现有冲突表示方法都没有考虑辨识框架不完整这一情况。针对这个问题, 本文提出了一个适用于开放世界的广义冲突表示模型, 基于广义基本概率指派并结合广义证据距离, 使用一个新的二元组综合表示证据之间的冲突。对几种主要的现有方法进行了比较和分析, 新冲突表示方法主要贡献在于: 能够定量给出辨识框架不完整导致冲突的可能性, 并对是否使用 Dempster 组合规则还是使用广义证据理论中的广义组合规则给出了一个判断依据。组合策略是: 如果判断系统处于开放世界, 则不使用 Dempster 组合规则, 而使用广义组合规则。相对于已有的判断依据, 本文所提出的策略物理意义更为清晰。最后用算例说明了本文所提出方法的有效性。

关键词: 广义证据理论; 广义基本概率指派; 冲突表示; 冲突管理; 广义证据距离

中图分类号: (TP391) **文献标识码:** A

Abstract: One of the most important open issues is that the classical Dempster Shafer evidence theory is not efficient to handle highly conflicting evidence combination. The earlier researches were done heavily on the improvement of Dempster combination rule. However, the current research shows that not only the combination rule but also the classical conflict coefficient in evidence theory is not correct to determine the conflict degree between two pieces of evidence. One of the most important of existing methods is that they do not take consideration of the open world situation, whose frame of discernment is incomplete. To solve this problem, a novel method to determine the conflict degree between evidence is proposed in the framework of our proposed generalized evidence theory. The generalized basic probability assignment of the empty set is introduced to make up of a triple model with the other two coefficients, namely the classical conflicts coefficient and the distance of evidence. The main contribution of the proposed method is that the situation that frame of discernment is incomplete can be recognized by the generalized basic probability assignment value of the empty set. As a result, the strategy how to use the Dempster rule can be easily developed. The suggestion of our proposed strategy is that the Dempster rule can **ONLY** be used in the situation that it is confirmed to be under the close world, whose frame of discernment is complete. On the other hand, the proposed generalized combination rule is recommended if the situation is under the open world. The physical significance of the proposed is much more clear and more easily to be understand than exiting methods. Finally, some numerical examples are used to illustrate the efficiency of the proposed conflict representation model.

Key words: generalized evidence theory; generalized basic probability assignment; conflict management; empty set; generalized evidence distance

Zadeh曾给出一个算例表明Dempster Shafer证据理论在证据高度冲突时会产生于直觉矛盾的错误结果,这一问题引起了极大的关注。对众多现有的研究进行归纳分析,可以分为基于修改Dempster组合规则的方法和修改数据的方法。

修改规则的观点可概括为:冲突管理是改进和发展证据推理的关键问题。证据高度冲突下使用Dempster组合规则会产生不合理结论,这是由该规则的归一化步骤所产生的,所以需要修改Dempster组合规则。新的组合规则需要重点研究如何将冲突重新分配这一问题,该问题可进一步细分为:冲突应该重新分配给哪些子集;在确定可以接受冲突的子集后,冲突应该以什么比例分配给这些子集。这一流派的代表人物是Lefevre等人提出的“统一信度函数组合方法”^[1]。这一思路下的其他主要方法还有:Yager^[2,3], Inagaki^[4], 李弼程等的引入命题的平均支持程度的方法^[5], 邓勇等的考虑证据的可信度的方法^[6], 孙全等^[7], Zhang^[8]的方法, 条件Dempster组合规则^[9], 向阳等和陈寅等根据子集基数的大小决定聚焦的权重的方法^[10,11], Li Xiang等的修正组合方法^[12], 杜文吉等提出了加权组合规则^[13], 徐凌宇等基于证据损耗的观点来解释和提升证据组合规则^[14], 王壮等提出一种基于均衡信度分配准则的冲突证据组合规则^[15], Dinael等提出的minC组合规则细化了分配空间和局部冲突^[16], 提出了潜在冲突的概念, Smardndache等提出的冲突比例分配规则PCR3^[17], 李剑峰等提出的类似改进规则都属于这类方法^[18]。

与修改组合规则的第一大类方法不同, 第二大类方法是基于修改原始证据源的方法。该类方法的观点认为: Dempster组合规则本身没有错。当证据高度冲突时, 应该首先对冲突证据进行预处理, 然后再使用Dempster组合规则。这类方法以Haenni等为代表^[19], 其他典型方法有: 何兵等在考虑集合势的基础上, 提出另一种平均算法^[20, 21]。Murphy算术平均方法^[22]; Deng等人的加权平均方法^[23], Jøsang等人的提出一种基于二元框

架的一致性组合算子^[24]。

目前研究人员倾向Haenni的观点, 修改数据模型的论文占主流地位, 以似乎这场争论要以修改数据模型坚持使用Dempster组合规则这一思想的胜利而告终。各方的争论还在继续, 坚持修改组合规则的人员并未放弃自己的观点^[28,29]。但是Liu在权威杂志Artificial Intelligence上发表的论文为这个方向的研究提出了一个新的问题^[25]: 如何有效表示证据之间的冲突? Liu分析了证据理论中经典冲突系数的不足, 提出了一种新的冲突表示方法, 基本思路是引入TBM中的pignistic概率距离与经典证据理论的冲突系数相结合构成一个二元组, 联合表示证据之间的冲突程度。进一步地, Liu对Dempster组合规则的使用条件进行了探讨。

Liu的工作是非常有意义的: 这么多研究冲突处理的工作, 普遍使用的冲突系数都存在着问题, 这些研究并未仔细深入讨论过冲突到底该如何表示, 就从这点来看, 冲突处理远不是“是否坚持使用Dempster组合规则”这么简单。

本文我们将分析指出Liu的冲突表示方法的不足, 在广义证据理论框架下进一步推广Liu的工作, 并给出Dempster组合规则使用的条件。我们的观点是: 如果是辨识框架不完整时, 应该修改组合规则, 使用我们所提出的广义组合规则; 如果辨识框架完整, 即使冲突较大, 仍然应该坚持使用Dempster组合规则, 但使用前应该修改数据模型。这一思想把Lefevre、Haenni和Liu的观点有机结合并统一起来。

本文是系列论文的第3篇, 主要的工作是在广义证据理论的框架下提出了一个新的证据冲突系数表示模型, 新模型是一个广义冲突系数与广义证据距离的二元组, 其中广义冲突系数是一个关于空集GBPA的函数。新冲突系数的这两个参数能够分别表示辨识框架不完整和传感器受干扰两大类导致冲突的因素的影响。结合新冲突表示模型提出了一种合理的Dempster使用策略, 相比已有的方法, 所提出的组合规则使用策略具有更为清晰的物理意义。全文结构如下: 第1节给出了一些基本理论, 简单介绍了冲突问题, 分析了最新冲突表示方式的不足。第2节提出了广义证据理论下的冲突表示方法, 并提出了Dempster组合规则的使用策略。第3节用一些算例比较了新方法和现有方法, 分析了新方法的有效性。第4节小结了论文的工作。

收稿日期: 2010-xx-xx; 修订日期: 2010-xx-xx (留编辑处理)

基金项目: 国家重点基础研究发展规划(973)项目(2007CB311006), 国家自然科学基金(60874105, 60904099), 教育部新世纪优秀人才支持计划(NCET-08-0345), 上海市青年科技启明星计划(09QA1402900), 航空科学基金(20090557004, 20095153022), 上海交通大学“晨星学者计划”资助(T241460612), 西北工业大学学校科技创新基金(2008KJ02022), 部委基金资助。

通讯作者: 邓勇 E-mail: dengyong@sjtu.edu.cn

1. 基本理论和现有冲突系数

1.1 DS证据理论

设 U 是变量 X 的所有可能值的穷举集合, 并且 U 中的元素是互斥的, 称 U 为 X 的一个识别框架。 U 由一完备的互不相容的陈述集合组成, U 的幂集 2^U 构成命题集合 2^U 。当 U 中元素的个数为 n 时, 命题集合所代表的空间大小为 2^n 。

定义1 设 U 为识别框架, U 的幂集 2^U 构成命题集合 2^U , $\forall A \subset \Theta$, 如果集函数 $m: 2^U \rightarrow [0,1]$ 满足:

$$m(\Phi) = 0 \quad (1)$$

$$\sum_{A \subset U} m(A) = 1 \quad (2)$$

则称 m 为框架 U 上的基本概率指派(BPA)。BPA反映了证据对识别框架中的命题 A 的支持程度, 即 $m(A)$ 。若 $\forall A \subset \Theta$, 且满足 $m(A) > 0$, 则称 A 为焦元。所有的焦元集合称为核。

在此基础上提出的D-S 证据理论的证据组合规则表述如下:

$$m(A) = \frac{1}{1-k} \sum_{A_i \cap B_j \cap C_l \cap \dots = A} m_1(A_i) m_2(B_j) m_3(C_l) \dots \quad (3)$$

$$k = \sum_{A_i \cap B_j \cap C_l \cap \dots = \phi} m_1(A_i) m_2(B_j) m_3(C_l) \dots \quad (4)$$

在传统证据理论中系数 k 表示证据间冲突程度, 其值越大说明证据之间的冲突越大。

1.2 广义证据理论

我们在系列论文第1篇中提出了适用于开放世界的广义证据理论 (Generalized evidence theory, GET), 简单介绍如下:

定义2 设 U 为开放世界的辨识框架, U 的幂集 2^U 构成命题集合, $\forall A \subset \Theta$, 如果函数 $m: 2^U \rightarrow [0,1]$ 满足:

$$\sum_{A \subset U} m_G(A) = 1 \quad (5)$$

则称 m 为框架 U 上的广义基本概率指派(Generalized Basic Probability Assignment, GBPA)。GBPA与经典BPA的区别是没有 $m(\Phi)$ 等于0的限制, 空集也可以是焦元, 也可以是核的一部分。如果 $m(\Phi)$ 等于0, 则GBPA就退化为经典

的BPA。

在GBPA基础上, 我们提出广义组合规则(Generalized combination rule, GCR)。

定义3 设 $\Phi_1 \cap \Phi_2 = \Phi$, 也即空集和空集的交集仍然是空集, 给定两个GBPA, 定义GCR为:

$$m(A) = \frac{(1-m(\Phi)) \sum_{B \cap C = A} m_1(B) m_2(C)}{1-K} \quad (6)$$

$$K = \sum_{B \cap C = \Phi} m_1(B) m_2(C) \quad (7)$$

$$m(\Phi) = m_1(\Phi) m_2(\Phi) \quad (8)$$

$$m(\Phi) = 1 \quad \text{当且仅当} \quad K = 1 \quad (9)$$

公式(7)中的 K 是广义冲突系数, 与经典冲突系数 k 的区别在于: K 是在开放世界下的冲突系数, 既描述了辨识框架中已知焦元命题之间的冲突, 又描述了由于辨识框架本身不完整导致的冲突。如果 $m(\Phi) = 0$ 成立, 则广义冲突系数 K 就退化为经典证据理论中的冲突系数 k 。

1.3 Pignistic概率距离

在介绍冲突问题之间在简单给出两个证据之间的距离函数, 这两个函数将在本文后续研究中使用。

Smets的TBM模型中给出的pignistic概率变换 (Pignistic Probability Transform, PPT) 中, 转换后的概率分布 $BetP$ 为^[26]:

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} \cdot \frac{m(B)}{1-m(\phi)} \quad \forall A \subseteq \Theta. \quad (10)$$

Smets认为在TBM模型的数据传递层次实现信息传递和组合后, 就进入模型的第二个层次——决策层。PPT的实质是将一个BPA函数转化为概率分布。

例1 设 $m(\{\theta_1\}) = 0.5$, $m(\{\theta_2\}) = 0.2$, $m(\{\theta_3\}) = 0.15$, $m(\{\theta_4\}) = 0.1$, $m(\{\theta_1, \theta_2, \theta_3, \theta_4\}) = 0.05$

根据PPT, 则有:

$$BetP(\theta_1) = 0.5 + \frac{0.05}{4} = 0.5125$$

$$BetP(\theta_2) = 0.2 + \frac{0.05}{4} = 0.2125$$

$$BetP(\theta_3) = 0.15 + \frac{0.05}{4} = 0.1625$$

$$BetP(\theta_4) = 0.1 + \frac{0.05}{4} = 0.1125$$

$BetP$ 是一个概率分布, 可以很方便的用于决策问题中。在给定pignistic概率转换后, 可以给出pignistic概率距离^[26]。

假定辨识框架 Θ 上, 性质不同的两个证据的mass函数分别为 m_1 和 m_2 , 它们对应的Pignistic概率函数为 $BetP_{m_1}$ 和 $BetP_{m_2}$, 则有:

$$difBetP_{m_1}^{m_2} = \max_{A \subseteq \Theta} \left(\left| BetP_{m_1}(A) - BetP_{m_2}(A) \right| \right) \quad (11)$$

其中, $\left| BetP_{m_1}(A) - BetP_{m_2}(A) \right|$ 表示两个证据源对于A的赌博信度的差值, 赌博信度距离即所有子集的差值中的最大值。当两个BPA函数比较时, $difBetP_{m_1}^{m_2}$ 可以简写为 $difBetP$, pignistic概率距离也可以简写为 dip 。

1.4 封闭世界中的证据距离^[27]

定义4 设 Θ 为一包含N个两两不同的命题的完备的辨识框架, 但是在证据高度冲突情况下基于经典Dempster组合规则的融合结果存在问题是 Θ 所有子集生成的空间。一个基本概率指派BPA是一个在 $E_p(\Theta)$ 中的坐标系为 $m(A_i)$ 的向量 \bar{m}_i ,

$$\sum_{i=1}^N m(A_i) = 1 \text{ 且 } m(A_i) \geq 1, \dots, 2^N$$

$$A_i \in P(\Theta)$$

定义5 设 Θ 为一个包含N个两两不同的命题的完备的辨识框架, m_1 和 m_2 是在辨识框架 Θ 上的两个BPA, 则 m_1 和 m_2 的距离可以表示为:

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (\bar{m}_1 - \bar{m}_2)^T \bar{D} (\bar{m}_1 - \bar{m}_2)} \quad (12)$$

其中 \bar{D} 为一个 $2^N \times 2^N$ 矩阵, 矩阵中的元素为:

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

1.5 高度冲突的例子

模糊数学的创始人Zadeh^[47]给出一个典型的证据冲突事例:

例2 两个医生对一个病人进行诊断, 病人可能患的是脑膜炎 (M), 脑震荡 (C), 或者是脑肿瘤 (T) 中的一种疾病。因此辨识框架为 $\Theta = \{M, C, T\}$ 。假设两位医生都认为病人患肿瘤的可能性很低, 但是分别认为患其他两种病的可能性很高, 诊断结果如下所示:

$$m_1: m_1(M) = 0.99, m_1(T) = 0.01$$

$$m_2: m_2(C) = 0.99, m_2(T) = 0.01$$

如果用Dempster证据组合公式进行组合, 将会得到意外结论:

$$m(M) = 0, m(C) = 0, m(T) = 1$$

组合结果显示病人患的疾病是脑肿瘤, 这一结果与常理相悖。

2 封闭世界下对Liu冲突表示方法的改进

许多学者在这方面进行了研究, 但是都未能解决什么是证据冲突, 证据冲突的程度如何? 要解决冲突证据的融合, 那么首先就应该判断证据之间是否存在冲突? 冲突的程度有多大? 事实证明D-S证据理论中系数k无法有效度量证据之间的冲突, 最直接的例子就是两个完全一样的证据体, 在这种情况下基本可以认为证据之间是没有冲突的, 但是系数k可以趋于1。下面的例子可以说明这一情况

例2 识别框架 $U = \{a, b, c, d, e\}$: 系统两个证据如下:

$$m_1\{a\} = m_1\{b\} = m_1\{c\} = m_1\{d\} = 0.25$$

$$m_2\{a\} = m_2\{b\} = m_2\{c\} = m_2\{d\} = 0.25$$

在本例中, 如果采用D-S理论中的冲突系数k来表征证据之间的冲突, 可以计算得到 $k = 0.75$, 按照传统的想法, 这两个证据之间是高度冲突的。但实际上, 这是两个完全相同的证据, 两者都表示“对4个目标识别命题的同等支持”两者之间并不存在冲突。这个例子说明D-S理论中的冲突系数k并不能有效表示证据之间的冲突。一个显而易见的问题是: 该用什么变量来表征证据之间的冲突?

Liu对这一问题展开了研究。她指出: 传统的 k 并不能有效地度量证据之间的冲突程度, 解决的途径是引入了pignistic概率距离, 将两者结合起来组成二元数组来描述冲突的大小。通过对 k 和 $difBetP$ 的分析, Liu给出了两个基本信度指派的具体定义。

(基本信度指派的冲突): 假定辨识框架 Θ 上, 性质不同的两个证据的 BPA 函数分别为 m_1 和 m_2 , 定义

$$cf(m_1, m_2) = \langle k, difBetP \rangle \quad (13)$$

是一个二维矩阵, 其中 k 为采用 Dempster 组合规则组合 m_1 和 m_2 时分配给未知集合的基本信度指派, $difBetP$ 为 pignistic 概率距离。当 $difBetP > \varepsilon$ 且 $k > \varepsilon$ (其中 $\varepsilon \in [0, 1]$ 为证据冲突的阈值) 时, 认为 m_1 和 m_2 两个基本信度指派互相冲突。

因为并不存在满足任意两个基本信度指派的冲突要求的“绝对意义上的阈值”, 所以阈值 ε 一般都是根据实际应用情况主观确定的。一般来说, 当 ε 越接近 1.0 时, 对于冲突的容忍越大。

例3 设辨识框架为 $\Theta = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$, 两条传感器的证据报告分别为 m_1 和 m_2 , 其中:

$$m_1(\{\omega_1\}) = 0.8 \quad m_1(\{\omega_2, \omega_3, \omega_4, \omega_5\}) = 0.2 \\ m_2(\Theta) = 1$$

采用 Dempster 组合后分配给未知集合的基本信度指派为 $k = 0$, 传统意义上认为该两条证据报告不存在冲突。根据 Liu 给出的证据冲突的定义, 它们也确实互相之间不存在冲突。然而, m_1 给出的信息更加有效, 而 m_2 对于各个命题的选择比较模糊, 它们之间的差别则是由 $difBetP = 0.6$ 表现出来。也就是说, 这两个证据报告源拥有的信度是不一样的。

综合上述, 当赌博信度距离很大时并不能认为基本信度指派互相冲突, 而 k 很大时也不能认定冲突, 只有把这两个值结合起来才能够准确定义基本信度指派是否冲突。Liu 的方法虽然克服了经典冲突系数 k 的不足, 但是仍然存在问题。

例4 设辨识框架 $\Theta = \{a, b, c\}$, 有两个证据, 其基本概率指派如下:

$$m_1 : m_1(A) = \frac{1}{3}, \quad m_1(b) = \frac{1}{3}, \quad m_1(c) = \frac{1}{3}$$

$$m_2 : m_2(a, b, c) = 1$$

由于 $k = 0$, $difBetP = 0$, 然后得出 $cf_{12} = 0$, 即认为两个证据之间没有冲突。这显然不合常理。因此有必要提出一个更为合理的冲突表示模型。

从物理意义上来看, Liu 的错误在于将“系统状态发生的概率相等”和“对系统一无所知”这两个命题等同起来, 但实际情况不是如此。比

如, 当一个系统告知: 通过大量投掷硬币的实验, 出现正面的概率是 0.5, 出现反面的概率是 0.5。在这种情况下, 并不表示我们对系统一无所知, 而是蕴含着诸多的信息, 比如, 我们可以推断该硬币应该有密度均匀的特性, 如果是在博弈中, 压正面可以赢 10 元, 压负面可以赢 5 元, 则根据平均收益应该决策压正面。反之, 我们对系统一无所知时, 我们并不知道该如何将可能性分配给单子集。也就是 $m(\Omega) = 1$ 表明 $m(a) = m(b) = 0.5$ 是可能的, 也表明 $m(a) = 0.7; m(b) = 0.3$ 和 $m(a) = 1; m(b) = 0$ 等各种情况都是可能的, $m(\Omega) = 1$ 时系统不确定度最大。基于 pignistic 距离无法区分等可能与一无所知两种情况, 并不是一种好的冲突度量方法。基于证据距离我们修正 Liu 方法的不足:

假定辨识框架 Θ 上, 性质不同的两个证据的 BPA 函数分别为 m_1 和 m_2 , 定义

$$cf(m_1, m_2) = \langle k, dis \rangle \quad (14)$$

是一个二维矩阵, 其中 k 为采用 Dempster 组合规则组合 m_1 和 m_2 时分配给未知集合的基本信度指派, 由公式 (4) 确定。dis 为公式 (12) 给出的封闭世界下的证据距离。

2 广义证据理论下的冲突表示模型

本节在 GET 下建立了新的冲突表示模型。首先提出了一个广义证据距离。

1.4 广义证据距离

设 Θ 为一广义证据理论的包含 N 个命题的辨识框架, 一个广义基本概率指派 GBPA 可以是一个在 $E_p(\Theta)$ 中的坐标系为 $m(A_i)$ 的向量 \vec{m}_i ,

$$\sum_{i=1}^N m(A_i) = 1$$

$$\text{且 } m(A_i) \geq 1, \dots, 2^N, A_i \in P(\Theta)$$

m_1 和 m_2 是在辨识框架 Θ 上的两个 GBPA, 则 m_1 和 m_2 的距离可以表示为:

$$d_{GBPA}(m_1, m_2) = \sqrt{\frac{1}{2}(\vec{m}_1 - \vec{m}_2)^T \bar{D}(\vec{m}_1 - \vec{m}_2)} \quad (15)$$

其中 \bar{D} 为一个 $2^N \times 2^N$ 矩阵, 矩阵中的元素为:

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

具体的计算方法是：

$$d_{GBPA}(m_1, m_2) = \sqrt{\frac{1}{2}(\|\vec{m}_1\|^2 + \|\vec{m}_2\|^2 - 2\langle \vec{m}_1, \vec{m}_2 \rangle)}$$

其中 $\|\vec{m}\|^2 = \langle \vec{m}, \vec{m} \rangle$ ， $\langle \vec{m}_1, \vec{m}_2 \rangle$ 为两个向量的内积：

$$\langle \vec{m}_1, \vec{m}_2 \rangle = \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m_1(A_i) m_2(A_j) \frac{|A_i \cap A_j|}{|A_i \cup A_j|}$$

$$A_i, A_j \in P(\Theta)$$

我们在公式(14)中引入了证据距离修正Liu的冲突表示模型具有一定的合理性，但是该冲突表示模型是在封闭世界中，并不适用于辨识框架不完整情况，基于我们所提出的广义证据理论，本文提出一个新的冲突表示模型。

假定辨识框架 Θ 上，性质不同的两个GBPA函数分别为 m_1 和 m_2 ，定义

$$cf_G(m_1, m_2) = \langle K, dis \rangle \quad (16)$$

其中， K 是由公式(7)计算的广义冲突系数；而 dis 则是公式(15)给出的广义证据距离。

相比现有的方法，新冲突表示模型可以有效处理辨识框架不完整情况，当辨识框架完整时，广义冲突系数 K 就退化为经典证据理论中的冲突系数 k ，公式(16)也就退化为公式(14)中所提出的修正冲突表示方法。

从图1给出了现有研究和本文研究的继承与逻辑关系。中可以看出现有工作和本文工作的逻辑关系。在Zadeh提出证据冲突问题后，在Liu之前的研究基本都是用经典冲突系数 k ，Liu引入pignistic概率距离 dip 与 k 结合起来作为冲突表示系数。本文先在封闭世界引入证据距离修正Liu的二元组表示方法，更进一步地，在面向开放世界的广义证据理论下提出了将广义冲突系数与广义证据距离相结合的方法。新方法适用范围更为广泛，表达证据之间的冲突程度更为有效。

2.3 组合规则使用策略

从GCR的物理意义这个角度来看，组合规则的使用策略可以表示如下：

只要确认系统处于封闭世界，此时

$m(\Phi) = 0$ ，就可以使用经典的Dempster组合规则。如果另外一个参数证据距离比较小，表明证据之间冲突程度较小，可以直接使用Dempster组合规则；如果另外一个参数证据距离比较大，表明证据冲突程度较大，此时应该先修正数据模型，再使用Dempster组合规则。

如果确认系统处于开放世界，此时 $m(\Phi) \neq 0$ ，则无论证据距离的大小如何，都不应该使用经典的Dempster组合规则。

综合上述策略可以看出，使用Dempster组合规则与否关键的因素是 $m(\Phi)$ ，如果只是证据间的证据距离较大，这种冲突通常是传感器失效或是受干扰导致，可以在修改数据模型的前提下坚持使用Dempster组合规则。如果生成的GBPA中焦元包括空集，则系统有可能是辨识框架不完整，不应该使用Dempster组合规则。与有冲突就修改规则以及有冲突就修改数据模型而坚持组合规则的两大思路相比，本文所提出的思路更符合逻辑，物理意义更为清晰。

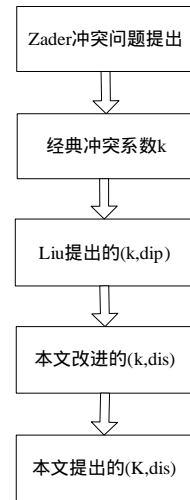


图1 本文研究路线

Fig.1 Research diagram

3. 应用

例5 有两组GBPA，第一组GBPA从 $m_1(a) = 1$ ； $m_1(\Phi) = 0$ 开始， $m_1(a)$ 每次递减0.1， $m_1(\Phi)$ 每次递增0.1；第二组GBPA从 $m_2(a) = 1$ ； $m_2(\Phi) = 0$ 开始， $m_2(a)$ 每次递减0.1， $m_2(\Phi)$ 每次递增0.1，则两组GBPA之间的广义冲突系数可以表示为图(2)。

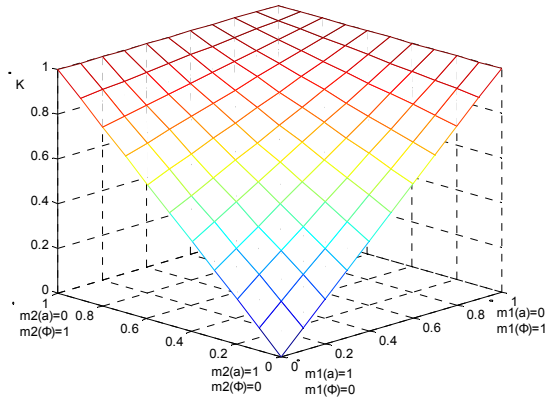


图2 两组GBPA之间的广义冲突系数

Fig.2 Generalized conflict coefficient between two GBPAs

从图(2)可以看出,在 $m_1(a)=1$; $m_1(\Phi)=0$; 且 $m_2(a)=1$; $m_2(\Phi)=0$ 时,广义冲突系数最小,表明系统完全支持命题 a ,且此时系统处于辨识框架完整状态。随着两组GBPA的 $m(a)$ 逐渐递减0.1, $m(\Phi)$ 逐渐递增0.1,广义冲突系数开始增大,这表明系统支持辨识框架不完整的程度越来越大。根据GCR,当两个GBPA中有一个GBPA的 $m(\Phi)=1$ 时,广义冲突系数达到最大为1。

例6 有两组GBPA,第一组GBPA从 $m_1(a)=0$; $m_1(\Phi)=1$ 开始, $m_1(a)$ 每次递增0.1, $m_1(\Phi)$ 每次递减0.1;第二组GBPA从 $m_2(a)=0$; $m_2(\Phi)=1$ 开始, $m_2(a)$ 每次递增0.1, $m_2(\Phi)$ 每次递减0.1,则两组GBPA之间的广义证据距离可以表示为图(3)。

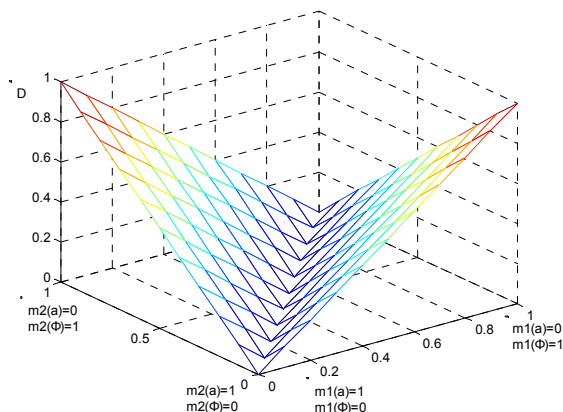


图3 两组GBPA之间的广义证据距离

Fig. 3 Generalized evidence distance between two GBPAs

图3可以看出,在对轴线上广义证据距离为0,因为此时两个证据完全相同。

例7 两个GBPA分别为:

$$m_{x1}(b) = 0.1; m_{x1}(\Phi) = 0.9$$

$$m_{x2}(b) = 0.1; m_{x2}(\Phi) = 0.9$$

如果根据所提出的冲突系数模型计算结果为:

$$cf_G(m_1, m_2) = \langle 0.82, 0 \rangle$$

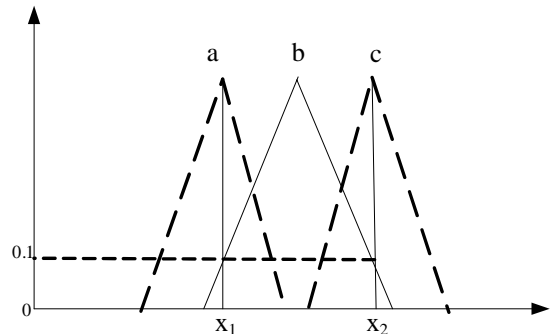


图4 两个属于不同类别的样本

Fig.4 Two samples with different classes

图4表示系统辨识框架中只有类别b,两个虚线三角形分别表示辨识框架中没有的类别a和类别c。样本 x_1 和 x_2 本应该分别属于类别a和类别c,如果仅仅根据广义证据距离进行参考,则会得出错误的结论。上面的结果说明广义证据距离在辨识框架不完整情况虽然能部分反映证据冲突,但是其效果没有广义冲突系数好。

下面两个例子讨论辨识框架完整 $m(\Phi)=0$ 时的冲突表示问题。

例8 设辨识框架为 $\{a, b\}$,有一个BPA表示为: $m_1(a, b)=1$;另外一个BPA初始状态为: $m_2(a)=0$; $m_2(b)=1$ 。以 $m_2(a)$ 每次递增0.1, $m_2(b)$ 每次递减0.1变化,则第2组BPA与 $m_1(a, b)=1$ 之间的证据距离和pignistic概率距离可以用图5表示。

从图5可以看出,当第二组BPA变化为 $m_2(a)=0.5$; $m_2(b)=0.5$ 时, pignistic概率距离的值为0,系统认为两个BPA之间没有冲突;而证据距离的值为0.5,系统认为两个BPA之间是

有冲突的,显然在辨识框架完整时度量证据之间的冲突程度应该选用证据距离,而不是pignistic概率距离。

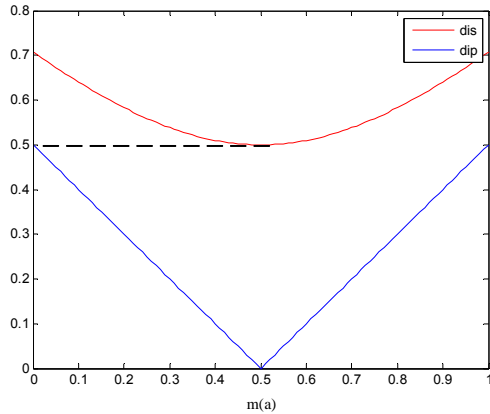


图5 证据距离与pignistic概率距离的比较

Fig. 5 Comparisons with evidence distance and pignistic probability distance

例9 设辨识框架 $\Omega = \{1, 2, 3, \dots, 20\}$, 有两个基本概率指派如下:

$$m_1(7) = 0.6 \quad m_1(A) = 0.4$$

$$m_2(1, 2, 3) = 1$$

其中 A 按照 $\{1\}$, $\{1, 2\}$, $\{1, 2, 3\}$, ... $\{1, 2, 3, \dots, 20\}$ 变化

表1分别给出了证据距离 d_{BPA} 与经典证据理论中的冲突系数 k 的结果。从图中可以看出:不论子集 A 的变化如何,经典的冲突系数始终为0.6,这不符合常理。广义证据距离的大小是随着子集 A 的变化而改变的。当子集 $A = \{1, 2\}$ 时,就比当子集 $A = \{1\}$ 时的冲突要小;类似地,当子集 $A = \{1, 2, 3, 4\}$ 时,就比当子集 $A = \{1, 2, 3, 4, 5\}$ 时的冲突要小;当子集 $A = \{1, 2, 3\}$ 时,新的冲突系数最小。

以上例子的结果说明广义冲突系数在辨识框架完整情况虽然能部分反映证据冲突,但是其效果没有广义证据距离数好。因此,在辨识框架不完整 $m(\Phi) \neq 0$ 时,冲突参数二元组应该以广义冲突系数为主对冲突进行度量和判断;在辨识框架完整 $m(\Phi) = 0$ 时,冲突参数二元组应该以广义证据距离为主对冲突进行度量和判断;

表 1 新冲突系数与经典冲突系数的比较

Table 1 Comparison of new proposed conflict coefficient with classical e conflict coefficient

Case	d_{BPA}	k
$A=\{1\}$	0.7916	0.6
$A=\{1,2\}$	0.7024	0.6
$A=\{1,2,3\}$	0.6	0.6
$A=\{1,\dots,4\}$	0.6782	0.6
$A=\{1,\dots,5\}$	0.7211	0.6
$A=\{1,\dots,6\}$	0.7483	0.6
$A=\{1,\dots,7\}$	0.7982	0.6
$A=\{1,\dots,8\}$	0.8	0.6
$A=\{1,\dots,9\}$	0.8083	0.6
$A=\{1,\dots,10\}$	0.8149	0.6
$A=\{1,\dots,11\}$	0.8202	0.6
$A=\{1,\dots,12\}$	0.8246	0.6
$A=\{1,\dots,13\}$	0.8283	0.6
$A=\{1,\dots,14\}$	0.8315	0.6
$A=\{1,\dots,15\}$	0.8343	0.6
$A=\{1,\dots,16\}$	0.8367	0.6
$A=\{1,\dots,17\}$	0.8388	0.6
$A=\{1,\dots,18\}$	0.8406	0.6
$A=\{1,\dots,19\}$	0.8423	0.6
$A=\{1,\dots,20\}$	0.8438	0.6

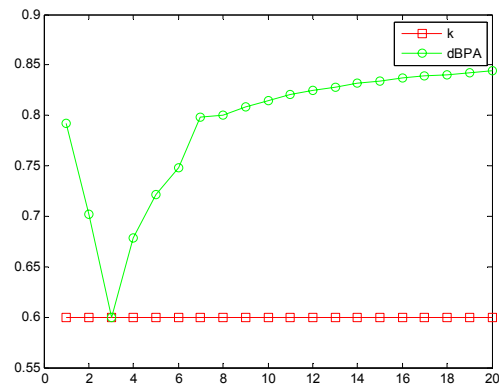


图 6 证据冲突系数的比较

Fig. 6 Comparison of new proposed conflict coefficient with classical e conflict coefficient English

4 结论

本文的主要工作是对证据之间的冲突进行度量。首先分析了现有冲突表示方法的不足。针对将经典冲突系数与pignistic概率距离结合的二元组的问题,引入证据距离进行修正。进一步地,

考虑到现有的冲突表示方法都没有考虑辨识框架不完整这一情况, 本文提出了一个适用于开放世界的冲突表示模型, 将经典冲突系数修正为广义冲突系数, 该系数是空集的广义基本概率指派的函数, 将广义冲突系数与广义证据距离相结合, 建立了广义证据理论中的冲突度量模型。

关于冲突表示模型方面的主要研究结论如下:

在辨识框架不完整情况下, 应该主要参考广义冲突系数的大小作为判断GBPA之间冲突程度的参数。

在辨识框架完整情况下, 所提出的冲突表示模型退化为经典冲突系数和证据距离构成的二元组, 此时应该主要参考证据距离的大小作为判断证据之间冲突程度的参数。

在所提出的冲突表示方法上, 提出了一种Dempster组合规则使用准则: 如果判断系统处于开放世界, 则不能使用Dempster组合规则, 而应该使用广义组合规则。如果判断系统处于封闭世界, 则可以使用Dempster组合规则, 在冲突较大时应该调整数据模型后再使用Dempster组合规则。相对于Liu的判断依据, 本文所提出的策略物理意义更为清晰。本文的冲突分析将应用到后续部分关于冲突证据融合的论文中, 对其他信息融合系统也有一定的参考价值。

参 考 文 献

- [1] Lefevre E, Colot O, Vannoorenberghe P. Belief functions combination and conflict management [J]. Information Fusion, 2002, 3(2): 149-162.
- [2] Yager R R. On the Dempster-Shafer framework and new combination rules[J], Information Sciences, 1989, 41(2): 93-138.
- [3] Yager R R. Quasi-Associative Operations in the Combination of Evidence[J], Kybernetes, 1987, 16(1): 7-41.
- [4] Inagaki T. Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory[J]. IEEE Transactions on Reliability, 1991, 40(2): 182-188.
- [5] 李弼程, 王波, 魏俊. 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.
- [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): 1275-1278.
- [7] 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.
- [8] Zhang L. Representation, independence, and combination of evidence in the Dempster-Shafer theory[C]. In Yager R R, Kacprzyk J, Fedrizzi M. Advances in the Dempster-Shafer Theory of Evidence, New York: John Wiley & Sons Inc., 1994, 51-69.
- [9] Fixsen D, Mahler R. Modified Dempster-Shafer approach to classification[J]. IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans, 1997, 27(1): 96-104.
- [10] 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.
- [11] 陈寅, 林良明, 颜国正. D-S 证据推理在信息融合应用中的存在问题及改进[J]. 系统工程与电子技术, 2000, 22(11): 69-71.
- [12] Li X, Liu Y X, Fu Y W. Decision fusion recognition based on modified evidence rule[J]. Progress in Natural Science, 2001, 11(4): 316-320.
- [13] 杜文吉, 谢维信. D-S 证据理论中的证据组合[J]. 系统工程与电子技术, 1999, 21(12): 92-94.
- [14] 徐凌宇, 张博锋, 徐炜民. D-S 理论中证据损耗分析及改进方法[J]. 软件学报, 2004, 15(1): 69-75.
- [15] 王 壮, 胡卫东, 郁文贤. 基于均衡信度分配准则的冲突证据组合方法[J]. 电子学报, 2001, 29(12A): 1852-1855.
- [16] Daniel M. Associativity in combination of belief functions: a derivation of minC combination[J]. Soft Computing, 2003, 7(4): 288-296.
- [17] Smarandache F, Dezert J. Four versions of the proportional conflict redistribution rules of combination in information fusion[J]. Eprint ar Xiv, cs/0408064, 2004(8): 1-14.
- [18] 李剑峰, 乐光新, 尚勇. 基于改进型 D-S 证据理论的决策融合滤波算法[J]. 电子学报, 2004, 32(7): 1160-1164.
- [19] Haenni R. Are alternative to Dempster's rule of combination real alternative? Comments on "About the belief function combination and the conflict management problem" ——Lefevre et al. [J]. Information Fusion, 2002, 3(3): 237-239.
- [20] 何兵, 胡文丽. 一种修正的 DS 证据融合策略[J]. 航空学报, 2003, 24(6): 559-562.
- [21] 何兵, 胡文丽. 一种分级的 DS 证据合成策略[J]. 计算机工程与应用, 2004, 10(1): 88-90.
- [22] Murphy K. Combining belief functions when evidence

- conflicts[J]. Decision Support Systems, 2000, 29(1): 1-9.
- [23] Deng Y, Shi W K, Zhu Z F et al . Combining belief functions based on distance of evidence[J]. Decision Support Systems, 2004, 38(4): 489-493.
- [24] Jøsang A. The consensus operator for combining beliefs[J]. Artificial Intelligence, 2002, 142(1-2): 157-170.
- [25] Liu W R. Analyzing the degree of conflict among belief functions[J]. Artificial Intelligence, 2006, 170(11): 909-924.
- [26] Smets P, Kennes R. The transferable belief model[J]. Artificial Intelligence. 1994, 66(3): 191-243.
- [27] Jousselme A L, Grenier D and Bosse E. A new distance between two bodies of evidence[J]. Information fusion, 2001, 2(1): 91-101.
- [28] Smets P. Analyzing the combination of conflicting belief functions[J]. Information fusion, 2007, 8(4): 387-412.
- [29] Florea M C, Jousselme A L, Bosse E, et al. Robust combination rules for evidence theory, [J]. Information fusion, 2009, 10(2): 387-412.

作者简介：

邓勇（1975 - ） 男，博士（后），教授，博士生导师。上海交通大学晨星奖励计划SMC优秀学者，入选教育部新世纪优秀人才支持计划和上海市青年科技启明星计划。主要研究方向：信息融合、智能信息处理。

Tel: 021-34204492 13916297024

E-mail: dengyong@sjtu.edu.cn

doctordengyong@yahoo.com.cn

蒋雯（1974 - ） 女，博士，副教授，硕士生导师。主要研究方向：信息融合。

E-mail: jiangwen@nwpu.edu.cn

韩德强（1980 - ） 男，博士，讲师。主要研究方向：信息融合、模式识别、图像处理。

E-mail: dehan@mail1.xjtu.edu.cn