

一种新的证据冲突分析方法

邓 勇^{1,2}, 王 栋², 李 齐², 章雅娟¹

(1. 西南大学 计算机与信息科学学院, 重庆 400715; 2. 上海交通大学 电子信息与电气工程学院, 上海 200240)

摘要: 当证据高度冲突时, 使用DS组合规则经常会得到错误的结果, 因此对DS证据理论的合成方法进行改进成为人们研究的重点. 但是如何去度量或确定证据之间冲突的程度却常常被人忽视. 在以往的研究中均使用冲突系数 k 作为表征证据之间冲突的量, 但是研究表明, k 不能很好的描述证据之间的冲突. 对表征证据冲突的变量进行了研究, 提出了一种新的表征证据之间冲突的变量-关联系数, 基于偏熵和混合熵, 定义了关联系数, 该系数可以定量的表示证据之间的冲突. 当关联系数接近1时证据之间冲突很小, 当关联系数接近0时证据之间高度冲突. 算例验证了所提出的基于关联系数的冲突表示方法的有效性.

关键词: D-S证据理论; 组合规则; 证据冲突; 偏熵; 关联系数

中图分类号: TP273 **文献标识码:** A

A new method to analyze evidence conflict

DENG Yong^{1,2}, WANG Dong², LI Qi², ZHANG Ya-juan¹

(1. College of Computer and Information Sciences, Southwest University, Chongqing 400715, China;

2. School of Electronic, Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: Dempster Shafer evidence theory is widely used in data fusion. However, incorrect results may be obtained by classical Dempster combination rule when collected evidence highly conflict with each other. Extensive work has been done to improve the DS combination rule, but how to measure the conflict among two pieces of evidence is an issue which has not been well considered. We define a new factor, the relative coefficient, for measuring the conflict between two pieces of evidence based on the partial entropy and the mixture entropy. When the relative coefficient is close to one, the conflict degree between two pieces of evidence is very low. On the other hand, when the relative coefficient is close to zero, the conflict degree between two pieces of evidence is very high. Some numerical examples are given to illustrate the efficiency of the proposed method.

Key words: D-S evidence theory; combination rule; evidence conflict; partial entropy; relative coefficient

1 引言(Introduction)

当D-S证据理论在信息融合领域应用非常广泛, 但是在证据高度冲突时会出现有悖常理的结果^[1~3]. 研究人员为了处理这个问题, 提出了很多改进的合成方法, 探讨了两个证据之间的冲突存在于什么地方, 冲突应该重新分配给哪些子集, 应该以什么比例分配给这些子集^[4~6]. 由于D-S证据理论在智能信息融合领域得到了广泛应用, D-S的改进方法再一次受到人们的重视^[7~10].

在选择合适的合成方法时确定证据之间是否冲突是个至关重要的问题^[10]. 目前还没有一个统一的方法来度量证据之间的冲突, 人们通常把分配给空集的合成概率指派, 即冲突系数 k 作为度量证据冲突的标准. 但是研究表明, 冲突系数 k 不能很好的度量证据之间的冲突, 比如: 即使两个完全相同的证据之

间也可能得到一个很大的冲突系数 k .

为了更好地描述证据之间的冲突程度, 在引入偏熵和混合熵概念的情况下, 本文提出了关联系数这一种新的表征证据之间冲突的变量. 本文认为如果两个证据之间是一致的, 它们对每个子集的支持度应该是相似的, 那它们的关联性就强, 就可以认为它们是没有冲突的. 如果两个证据之间的关联性差, 那就可以认为它们之间是有矛盾的, 是冲突的. 当关联系数接近1时, 认为证据之间冲突很小, 当关联系数接近0时, 认为证据之间高度冲突. 通过这种方法可以判断证据之间是否存在冲突, 以决定是否选用D-S组合规则或是其他的合成方法.

2 D-S理论的基本概念(The basic concepts of D-S evidence theory)

限于篇幅, 本节仅仅简单地介绍D-S证据理论的

基本原理, 详细的信息可参考文献[11].

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 上的基本可信度分配(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D-S证据理论中, 对事件 A 的描述采用区间 $[\text{Bel}(A), \text{Pl}(A)]$, Bel 和 Pl 分别称为信任函数(belief function)和似然函数(plausibility function).

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.

3 Dempster组合规则存在的问题(The existing problems of Dempster combination rule)

在Dempster组合规则中, k 是一个用于衡量融合的各个证据之间冲突程度的系数. 如果 $k = 1$, 就不能使用Dempster组合规则进行信息融合; 而当 $k \rightarrow 1$ 时, 即对高度冲突的证据进行正则化处理将会导致与直觉相悖的结果.

例1 在文献[6]中, 假定识别框架 $U = (a, b, c)$,

证据1: $m_1\{a\} = 0.99, m_1\{b\} = 0.01$;

证据2: $m_2\{b\} = 0.01, m_2\{c\} = 0.99$.

可以看出 m_1 与 m_2 之间是高度矛盾的, 根据D-S组合规则对 m_1 和 m_2 进行组合的结果为 $m_1\{a\} = 0, m_1\{b\} = 1, m_1\{c\} = 0$. 该结果违背常理, 一个可信度很低的命题 $\{b\}$ 在组合后具有了最大信任度. 同时, 在证据完全冲突的情况下, 合成公式的分母为零, D-S组合无法进行.

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. 例2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:

$$m_1\{a\} = m_1\{b\} = m_1\{c\} = m_1\{d\} = m_1\{e\} = 0.2;$$

$$m_2\{a\} = m_2\{b\} = m_2\{c\} = m_2\{d\} = m_2\{e\} = 0.2.$$

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 $k = 0.8$, 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, 两者之间根本不存在冲突. 这说明D-S理论中的冲突系数 k 不能来表征证据之间的冲突. 那该用什么变量来表征证据之间的冲突呢?

2006年Liu发表了一篇专门分析证据冲突的文章^[12], 她对这一问题展开了研究. 她认为, 传统的 k 并不能有效的度量证据之间的冲突程度, 她引入了pignistic概率距离, 与冲突系数 k 两者结合起来描述冲突的大小.

她认为 $cf(m_1, m_2) = \langle k, \text{difBet}P \rangle$ 可以描述两个证据之间的关系, 判断它们是否有冲突, 是否高度冲突, D-S证据合成公式是否适用. $\text{difBet}P$ 的定义如下:

$$\text{Bet}P_m(A) = \sum_{\omega \subseteq A} \text{Bet}P_m(\omega), \quad (5)$$

$$\text{difBet}P_{m_1}^{m_2} =$$

$$\max_{A \subseteq \Omega} \sum_{A \subseteq \Omega} (|\text{Bet}P_{m_1}(A) - \text{Bet}P_{m_2}(A)|). \quad (6)$$

她根据两个因素的取值大小, 讨论了4种情况下证据之间的冲突:

A) 当 $k, \text{difBet}P$ 两个值都很小时, 表明证据 m_1, m_2 之间的冲突很小, 可以使用D-S证据合成规则.

B) 当 k 是个较大的值时, $\text{difBet}P$ 是个较小的值时, 表明证据 m_1, m_2 之间没有明显的区别, 使用D-S证据合成规则时要慎重.

C) 当 k 是个较小的值时, $\text{difBet}P$ 是个较大的值时, 表明证据 m_1, m_2 之间有明显的区别, 不建议用D-S证据合成规则.

D) 当 $k, \text{difBet}P$ 两个值都很大时, 表明证据 m_1, m_2 之间有明显的区别, 此时证据之间高度冲突, 不能用D-S证据合成规则.

Liu认为可以采用 $k, \text{difBet}P$ 这两个因子联合来判断证据之间是否有冲突, 是否是高度冲突. 但是她并没有提出一个公式来统一这两个因子, 只是单纯的就这两个因子的大小进行分类, 这样难免会出现

误差. 那么, 能否有一个公式来统一这两个因子, 或单独提出一个因子来表征证据之间的冲突呢?

国内外对冲突信息融合的研究主要集中在修改组合规则, 比如文献[13~16], 本文作者课题组也对这一问题进行了研究, 通过考虑传感器可靠性来解决冲突信息融合^[17~20]. 但是, 致证据冲突的原因不局限于此, 两个传感器如果观测的是两个不同的目标, 则也可能导致信息的冲突, 因此, 本文从传感器报告关联的角度来判断证据之间是否存在冲突以及表示的方法, 并提出了关联系数这一新的因子来判断证据之间的冲突.

4 用关联系数表征证据冲突(Characterization of evidential conflict with correlation coefficient)

判断证据是否存在冲突, 冲突是否明显, 可以考虑两个证据之间的关联性, 如果两个证据之间没有冲突或冲突很小, 那么它们的关联性就好, 即对同一命题的支持度接近: 如果两个证据高度冲突, 说明两者之间相似度低, 它们的关联性就不强, 因此可以用关联系数这个概念来描述证据之间的冲突.

作为事物不确定性度量——熵的概念为之拓广, 本文引入随机变量的偏熵和关联熵^[21], 分析其特性并定义关联系数, 这一概念在同一性检验、一致性分析等领域中得到广泛应用^[22].

定义 2 设随机变量 X 与 Y 的分布为

$$X = \left\{ \begin{matrix} a_1 \cdots a_k \\ p_1 \cdots p_k \end{matrix} \right\}, Y = \left\{ \begin{matrix} b_1 \cdots b_k \\ q_1 \cdots q_k \end{matrix} \right\},$$

那么随机变量 X 关于随机变量 Y 的偏熵定义为:

$$H_Y(X) = \sum_{k=1}^n q_k \log p_k, \quad (7)$$

$$H(X) = \sum_{k=1}^n q_k \log q_k. \quad (8)$$

定义 3 随机变量 X 与 Y 之间的关联熵定义为它们的偏熵之和, 即

$$H(X, Y) = H_Y(X) + H_X(Y). \quad (9)$$

定义 4 随机变量的偏关联系数与关联系数分别定义为:

$$r_Y(X) = \frac{H(Y)}{H_Y(X)}, \quad (10)$$

$$r_X(Y) = \frac{H(X)}{H_X(Y)}, \quad (11)$$

$$r(X, Y) = \frac{H(X \otimes Y)}{H(X, Y)} = \frac{H(X) + H(Y)}{H_Y(X) + H_X(Y)}. \quad (12)$$

关联系数具有以下性质: $0 \leq r_X(Y), r_Y(X), r(X, Y) \leq 1$, 当且仅当 X, Y 为同分布时, $r_X(Y) = r_Y(X) = r(X, Y) = 1$. 此性质表明关联系数 r 是随机变量 X 与 Y 的分布之同一性, 一致性的特征度量.

把这个关联系数应用到数据融合中, $X = \left\{ \begin{matrix} a_1 \cdots a_k \\ p_1 \cdots p_k \end{matrix} \right\}$ 中的 a_1, \dots, a_k 表示证据中的子集, 而 p_1, \dots, p_k 表示的是子集所对应的概率指派, 这样就可以把熵的概念引入到数据融合中了.

但是在进行数据计算时, 会产生一些问题, 例如在处理例1时, 由于证据1的 $m(c) = 0$, 而证据2中对应的项 $m(c) = 0.99$, 在计算过程中会出现 $0.99 \log 0 = -\infty$ 的情况, 这样的话, 计算关联系数 r 时, 分母就会变为无穷, 会得到一个接近于0的关联系数. 如果证据中有等于0的项, 而其对应项不为0时, 即使两个证据之间没有冲突, 计算的关联系数也会是趋近于0的数, 这样就不能判断证据之间的冲突了.

为了度量模糊数之间的关联性, 辛小龙等^[23]提出了一个解决方法, 当证据中出现为0的项时, 会有 $\log 0 = -\infty$, 为了避免计算过程中出现无穷大, 用一个很小的正数 1×10^{-12} 代替0, 这不会影响最终的分析结果. 但是这种方法在处理数据时也会遇到问题, 如下:

例 3 假定识别框架 $U = (a, b, c)$,

证据1: $m_1\{a\} = 0.01, m_1\{b\} = 0.99$;

证据2: $m_2\{b\} = 0.99, m_2\{c\} = 0.01$.

用上述方法计算关联系数会得到 $r = 0.1956$, 这说明证据之间冲突很大, 关联性不好, 但是按照实际认识, 这两个证据之间关联性应该很好, 它们之间的冲突应该是很小的, 所以说用一个很小的正数来代替0, 这种做法也有它的局限性.

在分析以前的研究结果的基础上, 本文提出了一种类熵的概念, 即随机变量 X 关于随机变量 Y 的类偏熵定义为:

$$H_Y(X) = \sum_{k=1}^n q_k \times^{-5p_k}, \quad (13)$$

$$H(X) = \sum_{k=1}^n q_k \times^{-5q_k}. \quad (14)$$

新的公式采用了和熵类似的思路, 熵的公式中用 \log 来把从0-1的数进行大范围的变化, 提出的类熵的概念也是把0-1的数进行大范围的变化, 因为 e 和 \log 是有反函数的关系, 所以采用 e 这个函数来代替 \log . 但是单纯的 e^{-p_k} 不能完全把0-1的数进行大范围的变化, 这需要在 e^{-p_k} 的基础上增加一个系数, 变成 e^{-5p_k} 的形式, 这个系数是根据验算得出的比较接近 \log 的结果. 在计算关联系数的过程中, 式(9)~(12)是不变的.

对于新定义的类熵和关联系数, 在处理例3时, 会得到关联系数为 $r = 0.9713$, 这与实际认识是一致的, 新的公式很好地处理了0的问题, 无须使用一个较小的正数来代替0, 这样也保证了计算的准确性.

但是这样只是单纯的从数学上讨论了两个证据

之间的关系,没有考虑到证据中基本概率指派对目标的支持关系.比如证据 $m_1\{a\}=0.1, m_1\{b\}=0.8, m_1\{a, c\}=0.1$,如果只是按照给的数据形式来进行计算,就没有考虑到证据对目标 c 的支持度.关联系数反映的是两个不同证据中基本概率指派对同一目标的支持度,如果两个证据对同一目标的支持度高,说明证据之间没有冲突;如果对同一目标的支持度有很大区别,这就说明证据之间存在冲突.因此计算前,要先对证据中的数据进行预处理,得到两个证据中所有出现的目标的支持度,再应用上述的关联系数的公式进行计算,这样得到的结果更加有说服力.

首先,两个证据中出现的所有辨识目标构成一个目标框架 Θ ,然后把证据中的数据按照目标框架的格式写成向量的形式.为了反映证据中不同数据对同一目标的支持度,文献[20]提出了一种BPA之间的距离,该距离的核心思想是将BPA视作向量,并引入一个因子 D 进行修正.基于这一思路,引入因子 D 实现证据中数据对目标框架中每个目标的支持度,形成新的向量.因子 D 是一个矩阵,它反映了目标之间的关系,公式如下:

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}, A, B \in 2^\Theta. \quad (15)$$

把因子 D 乘以目标构成的向量,生成新的向量如下:

$$\vec{v}' = D(i, j) \times \vec{v}. \quad (16)$$

这时就可以用上文中提到的关联系数的公式来进行计算证据之间的冲突.

下面,通过例子来说明关联系数可以用来表征证据之间的冲突.对于上文中提到的例1,应用关联系数的方法进行计算,得到的结果是 $r = 0.0165$.计算过程如下:

$$m_1 = \begin{Bmatrix} m(a) & m(b) & m(c) \\ 0.99 & 0.01 & 0 \end{Bmatrix},$$

$$m_2 = \begin{Bmatrix} m(a) & m(b) & m(c) \\ 0 & 0.01 & 0.99 \end{Bmatrix}.$$

先进行预处理过程,

$$D = \begin{pmatrix} & \{a\} & \{b\} & \{c\} \\ \{a\} & 1 & 0 & 0 \\ \{b\} & 0 & 1 & 0 \\ \{c\} & 0 & 0 & 1 \end{pmatrix},$$

$$\vec{X} = [0.99 \ 0.01 \ 0]^T, \vec{Y} = [0 \ 0.01 \ 0.99]^T,$$

$$\vec{X}' = D\vec{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.99 \\ 0.01 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.99 \\ 0.01 \\ 0 \end{pmatrix},$$

$$\vec{Y}' = D\vec{Y} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0.01 \\ 0.99 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.01 \\ 0.99 \end{pmatrix}.$$

再对处理过的数据进行关联系数计算,经计算可得:

$$H_Y(X) = 0.9995, H_X(Y) = 0.9995,$$

$$H(X) = 0.0165, H(Y) = 0.0165,$$

$$r(X, Y) = \frac{0.0165 + 0.0165}{0.9995 + 0.9995} = 0.0165.$$

从常识来看,两个证据对同一目标的支持度有明显的区别,证据1中对目标 a 的支持度为0.99,而证据2中对目标 a 的支持度为0,这说明两个证据之间有冲突,并且是高度冲突的.传统的证据冲突系数 k 也证实了这两个证据之间是高度冲突的.计算结果显示这两个证据之间的关联系数很小,这说明证据之间的关联度低,对同一目标的支持度有差别,证据之间是冲突的,而且关联系数趋近于0,说明证据之间是高度冲突的,此时不能使用D-S证据理论.

而对于例2,计算得的关联系数是1,说明证据之间没有冲突,两者是完全一致的,与实际认识也是一致的,通过上面的例子反映可以用关联系数来表征证据之间的冲突.

5 分析证据冲突(The analysis of evidential conflict)

上文已证实冲突系数 k 不能表征证据之间的冲突,下面用几个例子进一步说明关联系数可以用来度量证据之间的冲突.

当两个证据完全相同时,此时应用上述公式得到关联系数 $r = 1$,说明两个证据之间是一致的,对同一目标的支持度是完成相同的,证据之间没有冲突.

当证据是不完全相同时,关联系数同样可以表征证据之间的冲突.通过下面的例子来说明关联系数在判断证据冲突时的有效性.

例4 假定识别框架 $U = (a, b, c)$,

证据1: $m_1\{a\}=0.8, m_1\{b\}=0.1, m_1\{c\}=0.1$;
证据2: $m_2\{a\}=0.1, m_2\{b\}=0.8, m_2\{a, c\}=0.1$.

计算可得关联系数 $r = 0.2811$.

例5 假定识别框架 $U = (a, b, c)$,

证据1: $m_1\{a\} = 0.8, m_1\{b\} = 0.1, m_1\{c\} = 0.1$;
证据2: $m_2\{a\} = 0.8, m_2\{b\} = 0.1, m_2\{a, c\} = 0.1$.

计算可得关联系数 $r = 0.9719$.

由例4和例5可以看出当证据之间存在冲突或高度冲突时,关联系数就较小,而当证据之间没有冲突或冲突较小时,关联系数就较大.

当证据之间都是具有冲突的, 关联系数是否可以反映哪些证据之间的冲突更加严重? 通过下面的例子可以证实关联系数可以实现这个功能, 因为, 冲突更严重的证据对同一目标的支持度差别更大, 两者的关联性就更差, 因此关联系数可以判断那些证据之间冲突更严重.

例 6 假定识别框架 $U = (a, b)$,

证据1: $m_1\{a\} = 0.8, m_1\{b\} = 0.2$;

证据2: $m_2\{a\} = 0.2, m_2\{b\} = 0.8$.

计算可得关联系数 $r = 0.2961$.

例 7 假定识别框架 $U = (a, b)$,

证据1: $m_1\{a\} = 0.9, m_1\{b\} = 0.1$;

证据2: $m_2\{a\} = 0.1, m_2\{b\} = 0.9$.

计算可得关联系数 $r = 0.1292$.

由例6和例7所得的计算结果可以看出, 例6比例7的证据关联系数要大, 说明例6的两个证据之间更一致, 更能反映对同一目标的支持度, 即例7的证据比例6的证据冲突要大, 可见使用关联系数可以来判断两个证据之间是否存在冲突以及冲突的程度.

通过大量的计算可知, 关联系数的范围从0到1变化, 当关联系数越接近1时, 说明证据之间关联性越好, 一致性越高, 证据之间的冲突越小; 当关联系数越接近0时, 说明证据之间的冲突越大, 证据甚至是高度冲突的. 利用本文提出的关联系数判断证据之间是否有冲突, 解决了传统证据系数 k 不能很好的判断证据之间冲突的问题, 可以为下一步选择合适的证据合成方法打好基础.

从上文的例子已经看出关联系数这个因子可以表征证据之间的冲突, 本文将采用文献[24]中的例子来说明本文提出的关联系数可以像Liu等提出的pignistic概率距离一样来表征证据之间的冲突.

例 8 在文献[24]中, 设辨识框架 $\Omega = \{1, 2, 3, \dots, 20\}$, 有两个基本概率指派如下:

证据1: $m_1(2, 3, 4) = 0.05, m_1(5) = 0.05, m_1(\Omega) = 0.1, m_1(A) = 0.8$;

证据2: $m_2(1, 2, 3, 4, 5) = 1$.

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

表1给出了当 A 变化时这20个例子中 m_1 和 m_2 证据之间的关联系数与pignistic概率距离的比较. 图1显示了这两者随着子集 A 变化而变化的情况.

如图1所示, Liu等提出的pignistic概率距离 difBetP 所生成的曲线与本文提出的关联系数的曲线正好相反, 即 difBetP 大时, 关联系数小, 这是因为 difBetP 大时反映证据冲突大, 而关联系数则是小时反映两个证据之间关联性小, 两者冲突大. 从两条曲线的比较可以看出, 关联系数也可以表征证据之间的冲突大小, 并且是一个因子来表征的, 这样反映更直观.

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 conflict coefficient

例子	difBetP	关联系数 r	k
$A = \{1\}$	0.605	0.3313	0.05
$A = \{1, 2\}$	0.426	0.6422	0.05
$A = \{1, 2, 3\}$	0.248	0.8579	0.05
$A = \{1, \dots, 4\}$	0.125	0.9290	0.05
$A = \{1, \dots, 5\}$	0.125	0.9308	0.05
$A = \{1, \dots, 6\}$	0.258	0.8913	0.05
$A = \{1, \dots, 7\}$	0.355	0.7014	0.05
$A = \{1, \dots, 8\}$	0.425	0.6773	0.05
$A = \{1, \dots, 9\}$	0.480	0.6436	0.05
$A = \{1, \dots, 10\}$	0.525	0.6066	0.05
$A = \{1, \dots, 11\}$	0.560	0.5698	0.05
$A = \{1, \dots, 12\}$	0.591	0.5351	0.05
$A = \{1, \dots, 13\}$	0.617	0.5031	0.05
$A = \{1, \dots, 14\}$	0.639	0.4740	0.05
$A = \{1, \dots, 15\}$	0.658	0.4477	0.05
$A = \{1, \dots, 16\}$	0.675	0.4241	0.05
$A = \{1, \dots, 17\}$	0.689	0.4027	0.05
$A = \{1, \dots, 18\}$	0.702	0.3835	0.05
$A = \{1, \dots, 19\}$	0.714	0.3661	0.05
$A = \{1, \dots, 20\}$	0.725	0.3627	0.05

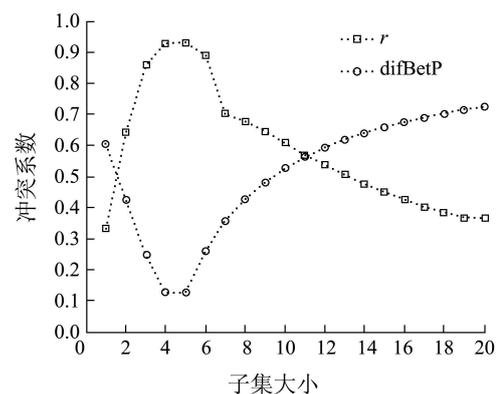


图 1 证据冲突系数的比较
Fig. 1 Comparison of new proposed conflict coefficient with classical conflict coefficient

6 结论(Conclusion)

通常情况下认为当冲突系数 k 大时就不能采用 D-S 证据合成理论, 冲突系数 k 被广泛的认为是反映证据之间冲突的变量, 但研究显示冲突系数 k 不能准确的表征证据之间的冲突. 本文采用关联系数这一新的变量来度量证据之间的冲突, 它根据证据之间的关联程度来反映证据之间的冲突, 判断证据之间冲突的程度, 从而进一步确定是否可以采用 D-S 证据理论及选择其他证据合成方法. 本文还有一些问题值得进一步研究: 比如, 公式(13)(14)中本文选用数字5, 基本思路是希望用新的曲线与log曲线尽可能相近, 工程上可以用于估算证据之间的冲突程度. 但是具体选用什么数值, 还缺乏物理意义上的解释.

此外,导致冲突的原因有很多种,比如传感器收到杂波干扰、传感器关联错误等等,本文所提出的冲突表示方法还无法有效区分是何种情况导致冲突.对于本文中提出的分析证据之间冲突的方法与导致冲突原因的关系,还有待于进一步研究.

本文所提出的方法并未考虑辨识框架不完整情况,因此本方法只适用于辨识框架完整,也就是封闭世界下的冲突度量.开放世界下如何度量证据之间的冲突仍然是一个有待解决的问题.

最后应该指出:证据冲突的表示是证据冲突解决的前提,Liu的二元组为这个问题提供了新的思路,但是,冲突的表示至今没有一个统一标准,目前的研究工作只是从定性(趋势上)和定量(冲突表征参数的数据)两个方面综合来判断冲突,给出一个让人感觉合理的方法,在将来的研究中提出一个统一的标准全面地比较各种新冲突表示方法的有效性也是一个有意义的课题.

致谢 感谢同济大学计算机科学与技术系的苗夺谦、魏莱对本文工作的一些意见和改进.感谢审稿专家对本文的意见和修改建议.

参考文献(References):

- [1] SMETS P, MAMDANI E H, DUBOIS D, et al. *Non-standard Logics for Automated Reasoning*[M]. New York, American: Academic Press, 1988.
- [2] ZADEH L. Review of a mathematical theory of evidence[J]. *AI Magazine*, 1984, 5(3): 81 – 83.
- [3] ZADEH L. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination[J]. *AI Magazine*, 1986, 7(2): 85 – 90.
- [4] DUBOIS D, PARDE H. Representation and combination of uncertainty with belief functions and possibility measures[J]. *Computational Intelligence*, 1988, 4(3): 244 – 264.
- [5] SMETS P. The combination of evidence in the transferable belief model[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 11(5): 447 – 458.
- [6] YAGER R R. On the Dempster-Shafer framework and new combination rules[J]. *Information Sciences*, 1987, 41(2): 93 – 137.
- [7] JOSANG A. The consensus operator for combining beliefs[J]. *Artificial Intelligence*, 2002, 141(1): 157 – 170.
- [8] LEFEVRE E, COLOT O, VANNOORENBERGHE P. Belief function combination and conflict management[J]. *Information Fusion*, 2002, 3(2): 149 – 162.
- [9] MURPHY C. Combining belief functions when evidence conflicts[J]. *Decision Support Systems*, 2000, 29(1): 1 – 9.
- [10] SMETS P. Analyzing the combination of conflicting belief functions[J]. *Information Fusion*, 2007, 8(4): 387 – 412.
- [11] 康耀红. 数据融合理论[M]. 西安: 西安电子科技大学出版社, 1997.
(KANG Yaohong. *The Theory of Data Fusion*[M]. Xi'an: Xidian University Press, 1997.)
- [12] LIU W R. Analyzing the degree of conflict among belief functions[J]. *Artificial Intelligence*, 2006, 170(11): 909 – 924.
- [13] 何兵, 郝爱民, 赵沁平. 一种基于不确定信息的决策方法[J]. 计算机学报, 2004, 27(2): 281 – 286.
(HE Bing, HAO Aiming, ZHAO Q P. A decision making method based on uncertain information[J]. *Chinese Journal of Computers*, 2004, 27(2): 281 – 286.)
- [14] 邢清华, 雷英杰, 刘付显. 一种按比例分配冲突度的证据推理组合规则[J]. 控制与决策, 2004, 19(12): 1388 – 1392.
(XING Qinghua, LEI Yingjie, LIU Fuxian. One combination rule of evidence theory based on distributing conflict in proportion[J]. *Control and Decision*, 2004, 19(12): 1388 – 1392.)
- [15] 何兵, 胡红丽. 一种修正的DS证据融合策略[J]. 航空学报, 2003, 24(6): 559 – 563.
(HE Bing, HU Hongli. A modified DS evidence combination strategy[J]. *Acta Aeronautica et Astronautica Sinica*, 2003, 24(6): 559 – 563.)
- [16] 潘泉, 张山鹰, 程咏梅, 等. 证据推理的鲁棒性研究[J]. 自动化学报, 2001, 27(6): 798 – 805.
(PAN Quan, ZHANG Shanying, CHENG Yongmei, et al. Some research on robustness of evidence theory[J]. *Acta Automatica Sinica*, 2001, 27(6): 798 – 805.)
- [17] 郭华伟, 施文康, 邓勇. 证据冲突: 丢弃, 发现或化解?[J]. 系统工程与电子技术, 2007, 29(6): 890 – 898.
(GUO Huawei, SHI Wenkang, DENG Yong. Evidential conflict and its 3D strategy: discard, discover and disassemble?[J]. *Systems Engineering and Electronics*, 2007, 29(6): 890 – 898.)
- [18] GUO H W, SHI W K, DENG Y. Evaluating sensor reliability in classification problems based on evidence theory[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2006, 36(5): 970 – 981.
- [19] 邓勇, 朱振福, 钟山. 基于证据理论的模糊信息融合及其在目标识别中的应用[J]. 航空学报, 2005, 26(6): 754 – 758.
(DENG Yong, ZHU Zhenfu, ZHONG Shan. Fuzzy information fusion based on evidence theory and its application in target recognition[J]. *Acta Aeronautica et Astronautica Sinica*, 2005, 26(6): 754 – 758.)
- [20] DENG Y, SHI W K, LIU Q. Combining belief functions based on distance of evidence[J]. *Decision Support Systems*, 2004, 38(3): 489 – 493.
- [21] 王维琼, 辛小龙. 混合偏熵与关联熵[J]. 模糊系统与数学, 2005, 19(4): 60 – 65.
(WANG Weiqiong, XIN Xiaolong. Hybrid partial entropy and relative entropy[J]. *Fuzzy Systems and Mathematics*, 2005, 19(4): 60 – 65.)
- [22] 吴敏金, 白治江. 关联熵及其应用[J]. 华东师范大学学报(自然科学版), 1998(2): 28 – 35.
(WU Minjin, BAI Zhijiang. Relative Entropy and its application[J]. *Journal of Eastchina Normal University(Natural Science)*, 1998(2): 28 – 35.)
- [23] 郭效芝, 辛小龙. Fuzzy集的偏熵与关联熵[J]. 模糊系统与数学, 2005, 19(2): 97 – 103.
(GUO Xiaozhi, XIN Xiaolong. Partial entropy and relative entropy of fuzzy sets[J]. *Fuzzy Systems and Mathematics*, 2005, 19(2): 97 – 103.)
- [24] JOUSSELME A L, GRENIER D, BOSSE E. A new distance between two bodies of evidence[J]. *Information Fusion*, 2001, 2(2): 91 – 101.

作者简介:

- 邓勇** (1975—), 男, 博士, 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, E-mail: ydeng@swu.edu.cn professorsdeng@hotmail.com;
- 王栋** (1980—), 男, 硕士研究生, 主要研究方向为信息融合;
- 李齐** (1976—), 男, 博士研究生, 主要研究方向不确定信息处理;
- 章雅娟** (1989—), 女, 硕士研究生, 主要研究方向风险分析、进化计算.