

DOI: 10.13382/j.jemi.B2104025

# 基于皮尔逊系数和不确定测度的冲突证据组合方法\*

李姝昊 王晓丹 宋亚飞

(空军工程大学 防空反导学院 西安 710051)

**摘要:** D-S 证据理论在合成冲突较大的证据时会产生直觉相悖问题, 现有的修正证据源的改进方法大多只从单一角度做出改进, 不能全面反映冲突信息特征。针对此问题, 提出了一种基于皮尔逊相关系数和不确定性测度的证据组合方法。首先, 利用皮尔逊相关系数衡量证据之间的相关性, 定义证据的可信度。其次, 引入基于区间概率的不确定度对可信度进行修正, 得到权重。最后, 使用该权重对原始证据进行加权平均, 使用 Dempster 组合规则进行合成。与经典改进方法相比, 该方法能有效处理冲突证据的融合问题, 指认正确命题的准确率达到 0.992 0。与已有皮尔逊系数改进方法相比该方法更具有合理性, 且有较高的准确度。

**关键词:** D-S 证据理论; 证据冲突; 皮尔逊相关系数; 可信度; 不确定测度

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 510.40

## Conflict evidence combination method based on Pearson coefficient and uncertainty measure

Li Shuhao Wang Xiaodan Song Yafei

(Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China)

**Abstract:** D-S evidence theory will produce intuitionistic conflict when synthesizing evidence with large conflict. Since most of the existing improvement methods for correcting evidence sources only make improvements from a single perspective, they cannot fully reflect the characteristics of conflict information. In order to solve this problem, a new evidence combination method based on Pearson correlation coefficient and uncertainty is proposed. Firstly, using the Pearson correlation coefficient to measure the correlation between evidences, to define the credibility of the evidence. Secondly, the uncertainty based on interval probability is introduced to modify the credibility to obtain the weight. Finally, using the weight of the weighted average of the original evidence, synthesized using Dempster combination rule. In the case analysis, compared with the classic improved method, the proposed method can effectively deal with the fusion of conflicting evidence, and the accuracy of identifying the correct proposition reaches 0.992 0. Compared with the existing Pearson coefficient improvement method, the proposed method is more reasonable and has higher accuracy.

**Keywords:** D-S evidence theory; evidence conflict; Pearson correlation coefficient; credibility; uncertainty measure

## 0 引言

D-S (Dempster-Shafer) 证据理论是 Dempster 于 1967 年提出, Shafer 于 1976 年进行进一步完善得到的一种不确定推理方法。由于其具有良好的处理不确定信息的能力, 伴随着信息多源化的发展, 现已被广泛应用到故障诊

断、目标识别、态势感知、可靠性分析等领域<sup>[1-7]</sup>。然而在实际应用时, 在某些特殊的情况下, 传统 Dempster 组合规则在融合高冲突的证据时会产生与直觉相悖的结果。国内外学者对此提出了大量改进方法, 主要分为两类。第 1 类方法认为问题源于证据源本身, 对原始证据源进行修正, 现有方法大多通过证据的可信度大小, 来赋予各证据不同权重, 加权平均后再进行合成<sup>[8-11]</sup>。在这类方

收稿日期: 2021-03-03 Received Date: 2021-03-03

\* 基金项目: 国家自然科学基金(61876189, 61273275, 61703426, 61806219)、陕西省高校科协青年人才托举计划(2019038)、陕西省创新能力支撑计划(2019-065)项目资助

法中,证据权重的确定对证据组合结果有着重要影响。权重确定的核心在于证据距离的选择上,当前涌现出诸多证据距离的度量定义,如 Jusselme 距离<sup>[12]</sup>、基于 Pignistic 概率的距离<sup>[13]</sup>、基于信任区间的距离<sup>[14]</sup>等。研究发现,由于影响证据合成结果的因素较多,只依靠距离函数这一度量规则得到的权重值不能很好地表示证据的重要程度,只是从单一的角度对证据源进行修正,考虑的不够全面<sup>[15]</sup>。第2类是对组合规则进行修改。Fei 等<sup>[16]</sup>从空集入手,将冲突量分配给空集,避免了组合过程中归一化的影响。Deng 等<sup>[17]</sup>修改了识别框架,将识别框架扩展到广义幂集中,给出新的组合规则。李永忠等<sup>[18]</sup>按照证据之间的冲突情况将证据划分为可接受冲突与不可接受冲突,并用不同的规则进行证据组合。

本文从修正证据源的角度出发,提出一种新的证据组合方法。先使用皮尔逊相关系数来衡量证据之间的相关程度,由此确定各证据体的可信度。为避免单一距离函数确定权重时的片面性,引入证据不确定度与可信度结合,分配证据的权重,对原始基本概率分配函数进行修正。最后,对修正后的证据进行加权平均得到平均证据,使用原始 Dempster 组合规则进行证据融合计算。通过与其他改进算法对比,验证了本文提出的新方法的有效性。

## 1 基本理论

### 1.1 D-S 证据理论基础

#### 1) 识别框架

证据理论的推理是建立在识别框架中的,通常以  $\Theta$  表示。识别框架内由有限个元素组成,两两相互独立,设一识别框架  $\Theta = \{A_1, A_2, \dots, A_n\}$ ,  $A_i$  是  $\Theta$  的一个子集,表示第  $i$  个元素。 $\Theta$  中所有元素组成的幂集表示所有可能发生事件的集合,记为  $2^\Theta$ ,可表示为:

$$2^\Theta = \{\emptyset, \{A_1\}, \{A_2\}, \dots, \{A_n\}, \{A_1, A_2\}, \dots, \Theta\} \quad (1)$$

#### 2) 基本概率分配

对于识别框架  $\Theta$  中的任意命题  $A$ ,定义函数  $m: 2^\Theta \rightarrow [0, 1]$ ,满足条件:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) \geq 0 \\ \sum_{A \subseteq \Theta} m(A) = 1 \end{cases} \quad (2)$$

其中  $\emptyset$  为空集,则称函数  $m$  为基本概率分配(basic probability assignment, BPA),  $m(A)$  的含义为证据对命题的支持程度,对于满足  $m(A) > 0$  的  $A$  称为焦元。

#### 3) 信度函数

设  $\Theta$  为一识别框架,  $m(\cdot)$  为基本概率分配函数,则

信度函数的定义为:

$$BEL(A) = \sum_{B \subseteq A} m(B), \forall A \subseteq \Theta \quad (3)$$

对于任意命题  $A \subseteq \Theta$ ,  $BEL(A)$  为  $A$  的每个子集的基本概率分配之和,表示对命题  $A$  为真的信任程度。

#### 4) 似然函数

设  $\Theta$  为一识别框架,  $m(\cdot)$  为基本概率分配函数,则似然函数的定义为:

$$PL(A) = 1 - BEL(\bar{A}) = \sum_{B \cap A \neq \emptyset} m(B) \quad (4)$$

对于任意命题  $A \subseteq \Theta$ ,  $PL(A)$  表示对命题  $A$  为非假的信任程度。

### 1.2 Dempster 组合规则

设  $m_1, m_2, \dots, m_n$  是一识别框架  $\Theta$  内的  $n$  个 BPA 函数,则对这  $n$  个证据的 Dempster 组合规则定义为:

$$m(A) = \begin{cases} \frac{\sum_{\cap A_i = A} \left( \prod_{i=1}^n m_i(A_i) \right)}{1 - k}, \forall A \subseteq \Theta, A \neq \emptyset \\ 0, A = \emptyset \end{cases} \quad (5)$$

其中  $k = \sum_{\cap A_i = \emptyset} \left( \prod_{i=1}^n m_i(A_i) \right)$  称为冲突因子,表示证据之间的冲突程度。

## 2 证据冲突及已有改进方法

### 2.1 D-S 证据理论存在缺陷

在合成冲突较小的证据时, Dempster 组合规则可以获得较好的结果。但是合成冲突较大的证据时,往往得到直觉相悖的结果。Zadeh 第1个提出了 Dempster 组合规则的合成悖论问题,通过一个经典的例子证明了组合规则无法解决证据冲突问题。两位医生 A 和 B 对同一位病人可能患某种病进行诊断,病人的病症可能为脑膜炎,脑震荡,脑肿瘤。在该例子中,识别框架为  $\Theta = \{\text{脑膜炎}, \text{脑震荡}, \text{脑肿瘤}\}$ ,两位医生的诊断结果如下:

医生 A,  $m_A(\text{脑膜炎}) = 0.99, m_A(\text{脑肿瘤}) = 0.01, m_A(\text{脑震荡}) = 0$ ;

医生 B,  $m_B(\text{脑膜炎}) = 0, m_B(\text{脑肿瘤}) = 0.01, m_B(\text{脑震荡}) = 0.99$ 。

使用 Dempster 组合规则进行合成,冲突因子  $k = 0.9999$ ,且合成结果  $m(\text{脑肿瘤}) = 1$ ,明显与直觉结果相反,此时组合规则无法合成正确结果。此外,文献[19-20]表明, Dempster 组合规则还存在信任偏移、证据吸收等问题。

### 2.2 经典的冲突证据改进方法

从修正证据源的角度出发,国内外学者已提出了诸多改进方法,其核心是根据证据相关性对证据体重新分

配权重。具有代表性的相关性测度有 Jousselme 距离、Pignistic 距离、夹角余弦相似度等。

定义 1 Jousselme 距离<sup>[12]</sup>。识别框架  $\Theta$  中 2 个独立 BPA 之间的 Jousselme 距离定义为:

$$d_j = \sqrt{\frac{(\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{D} (\mathbf{m}_1 - \mathbf{m}_2)}{2}} \quad (6)$$

式中:  $\mathbf{m}_i$  是向量形式的 BPA,  $\mathbf{D} = (d_{ij})_{|\Theta| \times |\Theta|}$  为一个  $|\Theta|$  阶方阵, 其组成元素  $d_{ij} = \left| \frac{A_i \cap A_j}{A_i \cup A_j} \right|$  ( $i, j = 1, 2, \dots, 2^{|\Theta|}$ ),  $A_i, A_j$  为识别框架  $\Theta$  的任意非空子集。Jousselme 距离用于衡量证据间的相异性, 通常以  $1 - d_j$  作为证据之间的相似性测度。

定义 2 Pignistic 距离<sup>[13]</sup>。识别框架  $\Theta$  中两个独立 BPA 之间的 Pignistic 距离定义为:

$$d_p = \max |BetP_{m_i}(A) - BetP_{m_j}(A)| \quad (7)$$

其中,  $BetP_{m_i}(A), BetP_{m_j}(A)$  是它们各自对应的 Pignistic 概率函数, 定义为:

$$BetP_m(A) = \sum_{A \subseteq \Theta, B \subseteq \Theta} \frac{|A \cap B|}{|B|} m(B) \quad (8)$$

式中:  $|B|$  为集合的势, 表示集合中包含元素的个数。

定义 3 夹角余弦相似度<sup>[21]</sup>。识别框架  $\Theta$  中两个独立 BPA 之间的夹角余弦相似度定义为:

$$cor(\mathbf{m}_1, \mathbf{m}_2) = \frac{\langle \mathbf{m}_1, \mathbf{m}_2 \rangle}{\|\mathbf{m}_1\| \cdot \|\mathbf{m}_2\|} \quad (9)$$

式中:  $\langle \mathbf{m}_1, \mathbf{m}_2 \rangle$  代表两个向量的内积,  $\|\mathbf{m}_2\|$  表示向量的模。

### 3 新的证据合成方法

#### 3.1 基于皮尔逊系数的相似性测度

设  $m_1, m_2$  为识别框架  $\Theta$  的两个证据体, 根据皮尔逊相关系数, 计算两个证据之间的相关性:

$$\rho_{12} = \frac{\text{cov}(m_1, m_2)}{\sigma_{m_1} \sigma_{m_2}} = \frac{E((m_1 - \mu_{m_1})(m_2 - \mu_{m_2}))}{\sigma_{m_1} \sigma_{m_2}} \quad (10)$$

其中,  $\text{cov}(m_1, m_2)$  表示两个证据之间的协方差;  $E$  表示数学期望;  $\mu_{m_i}$  和  $\sigma_{m_i}$  计算式分别如式(11)、(12)。

$$\mu_{m_i} = E(m_i) \quad (11)$$

$$\sigma_{m_i}^2 = E(m_i^2) - E^2(m_i) \quad (12)$$

皮尔逊相关性系数的范围为  $[-1, 1]$ , 负值表示负相关, 负值越小说明负相关程度越高。由于取值为负数时不满足一般的相关性系数的非负条件, 同时为了减小非正相关证据所占权重的计算, 这里引入阶跃函数进行修正:

$$\delta(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases} \quad (13)$$

修正后的相关性系数为:

$$sim(m_i, m_j) = |\rho_{ij}| \times \delta(\rho_{ij}) \quad (14)$$

易证明,  $sim(m_i, m_j)$  满足如下性质:

$$1) 0 \leq sim(m_1, m_2) \leq 1;$$

$$2) sim(m_i, m_j) = sim(m_j, m_i);$$

$$3) sim(m_i, m_j) = 1 \Leftrightarrow m_i = m_j.$$

#### 3.2 基于区间概率的证据不确定性测度

使用 3.1 节的方法只是从单一的角度对证据进行评价, 只考虑了证据之间的相关性, 考虑不够全面。因此, 本文在使用皮尔逊相关系数作为衡量证据相似性的基础上, 引入证据不确定度的定义, 从不同的角度对冲突证据进行修正。

通常情况下, 只靠证据距离无法有效衡量某一证据对于所有证据间全局冲突的影响。在证据理论中, 基于 BPA 可以获得信度区间  $[BEL(A), PL(A)]$ , 所有单元元素子集的信度区间可以看作是区间概率。对于识别框架  $\Theta = \{A_1, A_2, \dots, A_n\}$  上的 BPA, 所有的信度区间构成识别框架  $\Theta$  上的区间概率分布。区间概率的不确定性与证据的不确定性一样, 都是由不一致性和不精确性两部分组成。利用信度区间对证据不确定性进行量化时, 使用所有区间概率的中值  $\frac{PL(A_i) + BEL(A_i)}{2}$  来计算不一致性, 利用区间长度  $PL(A_i) - BEL(A_i)$  来计算不精确性。基于这种思路, 定义一种新的不确定测度<sup>[22]</sup>。

对于每个给定的证据体, 使用式(3)和(4)计算每个命题的信度函数  $BEL$  和似然函数  $PL$ , 设  $m$  为识别框架  $\Theta$  上的 BPA,  $i$  为证据体个数。其不确定度可表示为:

$$SU(m) = \sum_{i=1}^n \left[ \frac{BEL(A_i) + PL(A_i)}{2} \cdot \log_2 \frac{BEL(A_i) + PL(A_i)}{2} + \frac{PL(A_i) - BEL(A_i)}{2} \right] \quad (15)$$

#### 3.3 新的证据权重分配方法

Murphy 提出的证据组合方法是系统收集到  $n$  个证据时, 对  $n$  个证据的基本概率分配函数直接进行加权平均得到加权平均证据, 然后使用 Dempster 组合规则进行  $n-1$  次组合得到最终结果。邓勇等<sup>[23]</sup> 在 Murphy 的方法基础上, 考虑证据之间的关联程度, 建立了证据相似性测度来衡量证据的相关性。在本文方法中, 以皮尔逊相关系数建立证据之间的相关性测度, 确定证据可信度, 又综合考虑了证据的不确定度。通过分析可知, 一个证据的可信度代表着它被其他证据的支持程度, 在证据融合时应加大可信度较高的证据的权重; 同时, 一个证据的不确定度代表着它与其他证据的冲突程度, 在证据融合时应

降低不确定度较大的证据的权重。因此本文结合可信度与不确定度共同来确定证据的权重系数,对原始证据进行修正,将修正后的证据 BPA 进行加权平均,再利用 Dempster 组合规则进行融合,以此来达到解决证据冲突的目的。新方法的流程如图 1 所示。

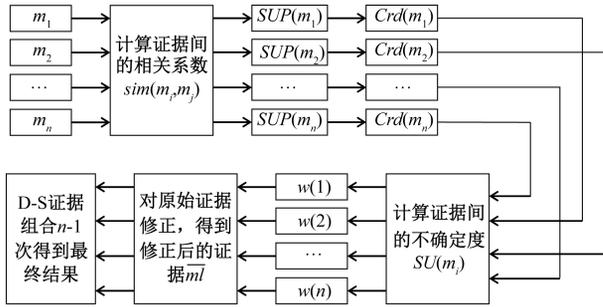


图 1 新的证据组合方法流程

Fig. 1 A flow chart of new evidence combination method

1) 按照式 (10) ~ (14) 计算两两证据之间的相关系数  $sim(m_i, m_j)$ , 并建立相关性矩阵  $SM$ 。

$$SM = \begin{bmatrix} sim(m_1, m_1) & \cdots & sim(m_1, m_n) \\ sim(m_2, m_1) & \cdots & sim(m_2, m_n) \\ \cdots & \cdots & \cdots \\ sim(m_n, m_1) & \cdots & sim(m_n, m_n) \end{bmatrix} \quad (16)$$

2) 计算证据  $m_i$  的支持度  $SUP(m_i)$ 。在两个证据进行融合时,若证据间的冲突较大,会得到两个证据的支持度均为 0 的结果,在步骤 3) 及之后的计算时,出现分母为 0 现象,使得计算公式无意义。为避免支持度计算结果为 0 对后续流程及识别框架的影响,在两个冲突较大的证据融合时,对支持度进行修正,将计算结果赋值为一个较小的数  $\varepsilon$ ,取  $\varepsilon$  为 0.001。

$$SUP(m_i) = \sum_{j=1, j \neq i}^n sim(m_i, m_j) \quad (17)$$

3) 由支持度计算证据  $m_i$  的可信度  $Crd(m_i)$

$$Crd(m_i) = \frac{SUP(m_i)}{\sum_{i=1}^n SUP(m_i)} \quad (18)$$

4) 根据式 (16) 计算各个证据的不确定度  $SU(m_i)$ 。

5) 根据证据可信度和不确定度计算权重,进行归一化处理。记第  $i$  个证据的权重为  $w(i)$ ,则第  $i$  个证据的权重定义为:

$$w(i) = \frac{Crd(m_i) \times e^{SU(m_i)}}{\sum_{j=1}^n (Crd(m_j) \times e^{SU(m_j)})} \quad (19)$$

6) 使用归一化后的权重系数对原始证据加权平均,进行修正得到修正后的证据  $\bar{m}_i$ :

$$\bar{m}_i = \sum_{i=1}^n (w(i) \times m_i) \quad (20)$$

7) 使用 D-S 证据合成规则将加权平均后的  $\bar{m}_i$  组合  $n-1$  次,得到最终结果。

### 4 算例分析

对比几种传统证据组合方法和皮尔逊改进方法两方面验证所提方法的有效性。例 1 主要对比传统证据组合方法,例 2 主要对比已有皮尔逊改进方法。

例 1 设识别框架  $\Theta = \{A, B, C\}$ , 现有 5 个传感器获得的基本概率分配函数如表 1 所示。

表 1 冲突证据源

Table 1 Conflict evidence data

证据体	A	B	C	AC
$m_1$	0.41	0.29	0.3	0
$m_2$	0	0.9	0.1	0
$m_3$	0.58	0.07	0	0.35
$m_4$	0.55	0.1	0	0.35
$m_5$	0.6	0.1	0	0.3

根据第 3 节所述求解过程,确定权重所需的各参数值,对 5 个证据进行融合时各参数如表 2 所示。除证据  $m_2$  外,其他 4 个证据  $m_1, m_3, m_4, m_5$  都认为命题 A 发生的可能性最大,赋予了 A 最大的支持度。而  $m_2$  对命题 A 的支持度为 0,与其他 4 个证据有着截然相反的结论,此时证据  $m_2$  相较于其他 4 条证据来说就是一条高冲突证据。通过对比 Dempster 方法、Murphy 的证据组合方法、邓勇的改进方法说明本文方法的有效性。通过 4 种组合方法对算例中的证据进行组合,得到的结果如表 3 所示,不同方法的对比如图 2 所示。

表 2 各证据参数值

Table 2 Each evidence parameter value

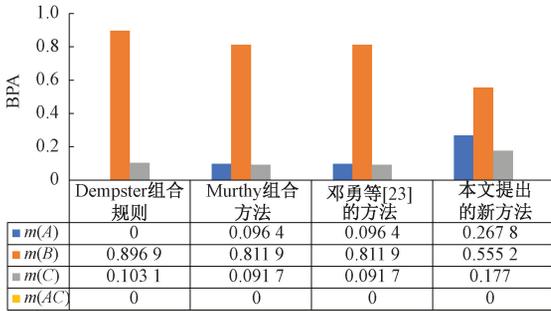
	SUP	Crd	SU	w
$m_1$	0.451	0.066	1.566	0.078
$m_2$	0.179	0.026	0.469	0.010
$m_3$	2.047	0.298	1.365	0.291
$m_4$	2.022	0.295	1.459	0.316
$m_5$	2.160	0.315	1.354	0.304

当证据高度冲突时,使用经典的 Dempster 组合规则将产生反直觉的结果,无法反映实际情况。虽然 Murphy 的组合方法,邓勇等<sup>[23]</sup>的方法都可以给出合理的结果,但它们的融合效果都不如本文提出的方法。此外,本文方法的收敛性能优于对比的其他方法。产生上述结果的主要原因是,通过利用证据距离和不确定性度量,冲突证据的最终权重大大降低,它对最终组合结果的影响大大地削弱。

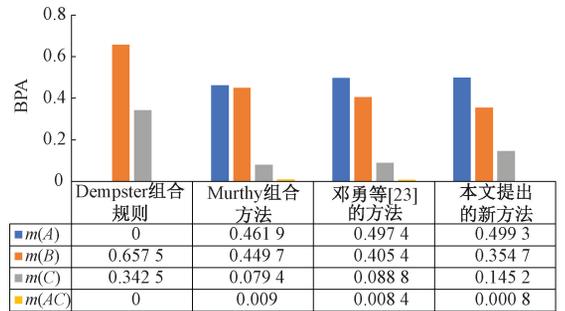
表 3 几种不同组合方法的证据组合结果

Table 3 Evidence combination outcomes based on different combination rules

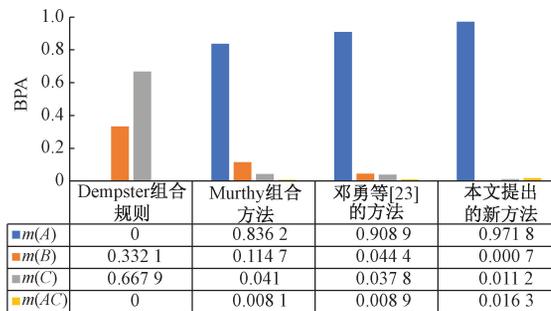
组合方法	$m_1, m_2$	$m_1, m_2, m_3$	$m_1, m_2, m_3, m_4$	$m_1, m_2, m_3, m_4, m_5$
	$m(A) = 0$	$m(A) = 0$	$m(A) = 0$	$m(A) = 0$
Dempster 组合规则	$m(B) = 0.896\ 9$	$m(B) = 0.657\ 5$	$m(B) = 0.332\ 1$	$m(B) = 0.142\ 2$
	$m(C) = 0.103\ 1$	$m(C) = 0.342\ 5$	$m(C) = 0.667\ 9$	$m(C) = 0.857\ 8$
	$m(AC) = 0$	$m(AC) = 0$	$m(AC) = 0$	$m(AC) = 0$
Murthy 组合方法	$m(A) = 0.096\ 4$	$m(A) = 0.461\ 9$	$m(A) = 0.836\ 2$	$m(A) = 0.962\ 0$
	$m(B) = 0.811\ 9$	$m(B) = 0.449\ 7$	$m(B) = 0.114\ 7$	$m(B) = 0.021\ 0$
	$m(C) = 0.091\ 7$	$m(C) = 0.079\ 4$	$m(C) = 0.041\ 0$	$m(C) = 0.013\ 8$
邓勇等 <sup>[23]</sup> 的方法	$m(AC) = 0$	$m(AC) = 0.009\ 0$	$m(AC) = 0.008\ 1$	$m(AC) = 0.003\ 2$
	$m(A) = 0.096\ 4$	$m(A) = 0.497\ 4$	$m(A) = 0.908\ 9$	$m(A) = 0.982\ 0$
	$m(B) = 0.811\ 9$	$m(B) = 0.405\ 4$	$m(B) = 0.044\ 4$	$m(B) = 0.003\ 9$
本文方法	$m(C) = 0.091\ 7$	$m(C) = 0.088\ 8$	$m(C) = 0.037\ 8$	$m(C) = 0.010\ 7$
	$m(AC) = 0$	$m(AC) = 0.008\ 4$	$m(AC) = 0.008\ 9$	$m(AC) = 0.003\ 4$
	$m(A) = 0.267\ 8$	$m(A) = 0.499\ 3$	$m(A) = 0.971\ 8$	$m(A) = 0.992\ 0$
	$m(B) = 0.555\ 2$	$m(B) = 0.354\ 7$	$m(B) = 0.000\ 7$	$m(B) = 0.000\ 1$
	$m(C) = 0.177\ 0$	$m(C) = 0.145\ 2$	$m(C) = 0.011\ 2$	$m(C) = 0.002\ 5$
	$m(AC) = 0$	$m(AC) = 0.000\ 8$	$m(AC) = 0.016\ 3$	$m(AC) = 0.005\ 4$



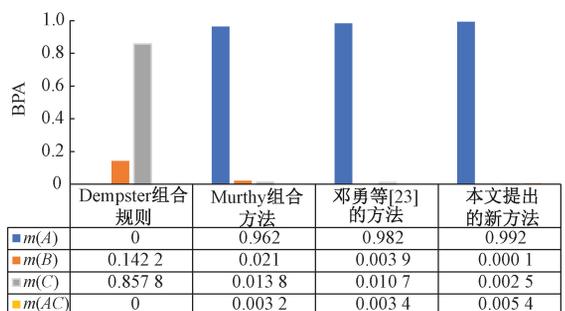
(a)  $m_1, m_2$  融合结果  
(a) Fusion result of  $m_1, m_2$



(b)  $m_1, m_2, m_3$  融合结果  
(b) Fusion result of  $m_1, m_2, m_3$



(c)  $m_1, m_2, m_3, m_4$  融合结果  
(c) Fusion result of  $m_1, m_2, m_3, m_4$



(d)  $m_1, m_2, m_3, m_4, m_5$  融合结果  
(d) Fusion result of  $m_1, m_2, m_3, m_4, m_5$

图 2 不同组合方法的融合结果

Fig. 2 Fusion results of different combination methods

例 2 现有 5 种探测器对网络中存在的潜在威胁进行探测, A、B、C 为 3 种不同类型的攻击, 假设某探测器探测到的 BPA 如表 4、5 所示, 表 5 为受干扰情况下探测器存在较大误差的情况。

文献[24-25]也使用了皮尔逊系数对证据冲突进行修正, 与本文不同的是, 文献[24]并未使用传统 Dempster 组合规则进行证据融合计算, 且只使用了皮尔逊系数计算证据体之间的相关程度, 进而对证据权重进行判定。

表 4 数据正常情况下的证据源

Table 4 Evidence source under normal data

证据体	A	B	C
$m_1$	0.9	0	0.1
$m_2$	0.88	0.01	0.11
$m_3$	0.5	0.2	0.3
$m_4$	0.98	0.01	0.01
$m_5$	0.9	0.05	0.05

表 5 受干扰情况下的冲突证据源

Table 5 Conflict evidence source under interference

证据体	A	B	C
$m_1$	0.9	0	0.1
$m_2$	0	0.01	0.99
$m_3$	0.5	0.2	0.3
$m_4$	0.98	0.01	0.01
$m_5$	0.9	0.05	0.05

表 6 数据正常情况下证据合成结果

Table 6 Evidence fusion results under normal data

组合方法	$m_1, m_2$	$m_1, m_2, m_3$	$m_1, m_2, m_3, m_4$	$m_1, m_2, m_3, m_4, m_5$
文献[24]	$m(A) = 0.967\ 33$	$m(A) = 0.852\ 384\ 5$	$m(A) = 0.885\ 389$	$m(A) = 0.889\ 182$
	$m(B) = 0.000\ 985$	$m(B) = 0.041\ 435$	$m(B) = 0.032\ 643$	$m(B) = 0.033\ 95$
	$m(C) = 0.031\ 68$	$m(C) = 0.104\ 332\ 4$	$m(C) = 0.077\ 738$	$m(C) = 0.072\ 12$
	$m(\Theta) = 0$	$m(\Theta) = 0.001\ 847$	$m(\Theta) = 0.004\ 228$	$m(\Theta) = 0.001\ 74$
文献[25]	$m(A) = 0.986\ 28$	$m(A) = 0.991\ 72$	$m(A) = 0.999\ 91$	$m(A) = 0.999\ 995$
	$m(B) = 0.000\ 02$	$m(B) = 0.000\ 02$	$m(B) = 0$	$m(B) = 0$
	$m(C) = 0.013\ 70$	$m(C) = 0.008\ 26$	$m(C) = 0.000\ 08$	$m(C) = 0.000\ 005$
	$m(A) = 0.986\ 15$	$m(A) = 0.963\ 4$	$m(A) = 0.994\ 8$	$m(A) = 0.9994\ 4$
本文	$m(B) = 0.000\ 035$	$m(B) = 0.004\ 9$	$m(B) = 0.000\ 4$	$m(B) = 0.000\ 03$
	$m(C) = 0.013\ 815$	$m(C) = 0.031\ 8$	$m(C) = 0.004\ 8$	$m(C) = 0.000\ 53$

表 7 受干扰情况下证据合成结果

Table 7 Evidence fusion results under interference

组合方法	$m_1, m_2$	$m_1, m_2, m_3$	$m_1, m_2, m_3, m_4$	$m_1, m_2, m_3, m_4, m_5$
文献[24]	$m(A) = 0.405\ 45$	$m(A) = 0.597\ 8$	$m(A) = 0.758\ 134$	$m(A) = 0.801\ 92$
	$m(B) = 0.004\ 505$	$m(B) = 0.057\ 19$	$m(B) = 0.029\ 02$	$m(B) = 0.026\ 97$
	$m(C) = 0.590\ 045$	$m(C) = 0.175\ 69$	$m(C) = 0.052\ 64$	$m(C) = 0.030\ 20$
	$m(\Theta) = 0$	$m(\Theta) = 0.169\ 31$	$m(\Theta) = 0.160\ 2$	$m(\Theta) = 0.140\ 89$
文献[25]	$m(A) = 0.017\ 85$	$m(A) = 0.952\ 79$	$m(A) = 0.999\ 86$	$m(A) = 0.999\ 996$
	$m(B) = 0.000\ 19$	$m(B) = 0.000\ 01$	$m(B) = 0$	$m(B) = 0$
	$m(C) = 0.981\ 95$	$m(C) = 0.047\ 20$	$m(C) = 0.000\ 14$	$m(C) = 0.000\ 004$
	$m(A) = 0.576\ 411$	$m(A) = 0.924\ 5$	$m(A) = 0.989\ 8$	$m(A) = 0.999\ 0$
本文方法	$m(B) = 0.000\ 033$	$m(B) = 0.013\ 1$	$m(B) = 0.001\ 2$	$m(B) = 0.000\ 1$
	$m(C) = 0.423\ 556$	$m(C) = 0.062\ 4$	$m(C) = 0.009\ 7$	$m(C) = 0.000\ 9$

文献[25]使用皮尔逊系数计算证据相关性,分配给证据相应的权重。本文不仅应用了皮尔逊相关系数计算证据之间的相关性,还引入了不确定性来完善单一测度对证据权重分配的不足。且本文使用了原始 Dempster 组合规则进行证据融合计算,保留了算法的完整性。使用文献[23-24]的方法与本文方法进行对比实验,正常情况下和受干扰情况下的证据源的证据合成结果如表 6、7 所示。

从表 6 可以看出,冲突较小甚至不存在冲突的证据融合时,文献[24]与本文的方法都可以得到最佳结果,在证据数量增加的过程中,文献[24]的方法对合理命题的融合精度没有明显的增加,在两个证据融合时对命题 A 的支持度最高,新增证据体的指认未能对结果有较大影响,有一定不足。本文的方法在 3 个证据进行融合时,所得结果对命题 A 的支持度相比两个证据融合时略有下降,是因为证据  $m_3$  对命题 A 的支持度稍有降低,在融合

时对结果有一定影响,所得结果是合理的。本文方法不仅可以正确的帮助决策者指认合理命题,相较于文献[24]更具有较高收敛性。从表 7 看出,当存在较大冲突证据进行融合时,文献[24]将一部分冲突分配至识别框架中,不利于决策者的指认,本文方法在收敛速度和精

度上都明显高于文献[24]。文献[25]只利用了证据的相关性进行简单的加权,忽略了证据信任程度和焦元支持度不一致的情况,在表 6 和 7 的数据中看到,一些数值变化较小的焦元在最终融合结果中被赋予了 0 值,结论并不严谨。本文方法相较于文献[25],在指认正确命题

(A) 的准确率上略有差距,但差距微小,并不影响对正确命题的指认。与文献[25]的方法相比,本文方法更具合理性。

## 5 结 论

本文提出了一种新的基于皮尔逊相关系数和不确定性度量的加权证据组合方法,该方法能有效地解决冲突证据的组合问题,克服了单一依靠证据距离进行改进方法的不足。通过与现有方法的比较,本文提出的方法在处理高度冲突的证据时有较好的融合效果,最大程度降低了冲突证据带来的组合问题。与其他引入皮尔逊相关系数的改进方法相比,本文提出的方法更具有合理性且收敛性更高,利于决策。

## 参考文献

- [ 1 ] 于晓英,董煜,董昱. 基于多方法证据融合的轨道电路故障诊断[J]. 铁道学报,2021,43(2):86-94.  
YU X Y, DONG Y, DONG Y. Fault diagnosis of track circuit based on multi method evidence fusion [ J ]. Journal of the China Railway Society, 2021, 43 ( 2 ): 86-94.
- [ 2 ] 黄新波,马玉涛,朱永灿. 基于信息融合和 M-RVM 的变压器故障诊断方法[J]. 电力自动化设备, 2020, 40(12): 218-225.  
HUANG X B, MA Y T, ZHU Y C. Transformer fault diagnosis method based on information fusion and M-RVM[J]. Electric Power Automation Equipment, 2020, 40(12): 218-225.
- [ 3 ] 刘娟,胡敏,黄忠. 基于区域 NSBP 特征的加权证据融合表情识别[J]. 电子测量与仪器学报,2020,34(11): 132-139.  
LIU J, HU M, HUANG ZH. Weighted evidence fusion expression recognition based on regional NSBP features[J]. Journal of Electronic Measurement and Instrumentation, 2020,34(11):132-139.
- [ 4 ] 姬晓飞,石宇辰,王昱,等. D-S 理论多分类器融合的光学遥感图像多目标识别[J]. 电子测量与仪器学报, 2020,34(5):127-132.  
JI X F, SHI Y CH, WANG Y, et al. D-S theory based multi-classifier fusion optical remote sensing image target recognition[J]. Journal of Electronic Measurement and Instrumentation,2020,34(5):127-132.
- [ 5 ] 寇广,汤光明,徐梓棕. 网络安全态势感知中 D-S 证据理论冲突证据的修正研究[J]. 计算机科学,2015, 42(7):200-203.  
KOU G, TANG G M, XU Z Q. Research on revising conflict evidence of D-S evidence theory in network security situation awareness [ J ]. Computer Science, 2015,42(7):200-203.
- [ 6 ] 陈宇奇,徐廷学,李志强,等. 基于证据 GO 法的复杂多态系统动态可靠性分析[J]. 系统工程与电子技术, 2020,42(1):230-237.  
CHEN Y Q, XU T X, LI Z Q, et al. Dynamic reliability analysis of complex multi-state system based on evidence GO method [ J ]. Systems Engineering and Electronics, 2020,42(1):230-237.
- [ 7 ] 丁家满,原琦,任东磊,等. 不确定性信息条件下系统可靠性分析[J]. 仪器仪表学报, 2019, 40(4): 153-162.  
DING J M, YUAN Q, REN D L, et al. System reliability analysis under uncertain information [ J ]. Chinese Journal of Scientific Instrument, 2019, 40(4): 153-162.
- [ 8 ] 由东媛,曹梦龙,姜凯. D-S 证据理论中冲突证据的改进方法研究[J]. 电子测量技术,2018,41(23):29-33.  
YOU D Y, CAO M L, JIANG K. Research on improved method of conflict evidence in D-S evidence theory [ J ]. Electronic Measurement Technology,2018,41(23):29-33.
- [ 9 ] LI J, WANG W, GAO X, et al. Strong conflicting evidences fusion based on belief interval distance measurement [ J ]. Cluster Computing, 2019, 22: 6589-6598.
- [ 10 ] 杨晓萍,廉伟健,李孟杰,等. 基于 Pignistic 概率函数和相关系数的证据组合方法[J]. 通信学报,2019, 40(3):73-79.  
YANG X P, LIAN W J, LI M J, et al. Evidence combination method based on Pignistic function transformation and correlation coefficient [ J ]. Journal on Communications, 2019, 40(3): 73-79.
- [ 11 ] 朱京伟,王晓丹,宋亚飞,等. 基于幂 Pignistic 概率距离的加权证据组合方法[J]. 通信学报,2018,39(1): 117-125.  
ZHU J W, WANG X D, SONG Y F, et al. Weighted evidence combination method based on power-Pignistic probability distance [ J ]. Journal on Communications, 2018, 39(1): 117-125.
- [ 12 ] 彭颖,胡增辉,沈怀荣. 一种修正证据距离[J]. 电子与信息学报,2013,35(7):1624-1629.  
PENG Y, HU Z H, SHEN H R. A modified distance of evidence [ J ]. Journal of Electronics & Information Technology, 2013,35(7):1624-1629.
- [ 13 ] 肖建于,童敏明,朱昌杰,等. 基于 pignistic 概率距离的改进证据组合规则[J]. 上海交通大学学报, 2012, 46(4):636-641.  
XIAO J Y, TONG M M, ZHU CH J, et al. Improved

- combination rule of evidence based on pignistic probability distance [J]. Journal of Shanghai Jiaotong University, 2012, 46(4): 636-641.
- [14] HAN D Q, DEZERT J, YANG Y. Belief interval-based distance measures in the theory of belief functions [J]. IEEE Transactions on Systems Man & Cybernetics Systems, 2016, 99: 1-18.
- [15] 薛大为, 王永, 高康凯. 利用奇异值和虚假度的证据组合方法[J]. 北京邮电大学学报, 2018, 41(1): 95-102.  
XUE D W, WANG Y, GAO K K. Evidence combination method based on singular value and falsity [J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(1): 95-102.
- [16] FEI L G, HU Y, XIAO F Y, et al. A modified TOPSIS method based on D numbers and its applications in human resources selection [J]. Mathematical Problems in Engineering, 2016, 2016: 6145196.
- [17] DENG X Y, LU X, CHAN F T, et al. D-CFPR: D numbers extended consistent fuzzy preference relations [J]. Knowledge based Systems, 2015, 73: 61-68.
- [18] 李永忠, 陈兴亮, 于化龙. 基于改进 DS 证据融合与 ELM 的入侵检测算法 [J]. 计算机应用研究, 2016, 33(10): 3049-3051, 3082.  
LI Y Z, CHEN X L, YU H L. Intrusion detection technologies based on improved evidence fusion of DS and ELM [J]. Application Research of Computers, 2016, 33(10): 3049-3051, 3082.
- [19] 杨艺, 韩德强, 韩崇昭. 基于多准则排序融合的证据组合方法 [J]. 自动化学报, 2012, 38(5): 823-831.  
YANG Y, HAN D Q, HAN CH ZH. Evidence combination based on multi-criteria rank-level fusion [J]. Acta Automatica Sinica, 2012, 38(5): 823-831.
- [20] 柯小路. 证据理论中信任函数的合成方法研究与应用 [D]. 合肥: 中国科学技术大学, 2016.  
KE X L. A Research on combination of belief functions with applications in evidence theory [D]. Hefei: University of Science and Technology of China, 2016.
- [21] 宋亚飞, 王晓丹, 雷蕾, 等. 基于相关系数的证据冲突度量方法 [J]. 通信学报, 2014, 35(5): 95-100.  
SONG Y F, WANG X D, LEI L, et al. Measurement of evidence conflict based on correlation coefficient [J]. Journal on Communications, 2014, 35(5): 95-100.
- [22] WANG X D, SONG Y F. Uncertainty measure in evidence theory with its applications [J]. Applied Intelligence, 2018, 48(7): 1672-1688.
- [23] 邓勇, 施文康, 朱振福. 一种有效处理冲突证据的组合方法 [J]. 红外与毫米波学报, 2004, 23(1): 27-32.  
DENG Y, SHI W K, ZHU ZH F. Efficient combination approach of conflict evidence [J]. Journal of Infrared and Millimeter Waves, 2004, 23(1): 27-32.
- [24] 魏永超. 基于皮尔逊系数的冲突证据合成新方法 [J]. 电讯技术, 2012, 52(4): 466-471.  
WEI Y CH. A novel D-S combination method of conflicting evidences based on pearson correlation coefficient [J]. Telecommunication Engineering, 2012, 52(4): 466-471.
- [25] 张欢, 陆见光, 唐向红. 面向冲突证据的改进 DS 证据理论算法 [J]. 北京航空航天大学学报, 2020, 46(3): 616-623.  
ZHANG H, LU J G, TANG X H. An improved DS evidence theory algorithm for conflict evidence [J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(3): 616-623.

### 作者简介



**李姝昊**, 2018 年于青岛农业大学获得学士学位, 现为空军工程大学硕士研究生, 主要研究方向为态势感知。

E-mail: 529127098@qq.com

**Li Shuhao** received her B. Sc. degree from Qingdao Agricultural University in 2018.

Now she is a M. Sc candidate at Air Force Engineering University. Her main research interest includes situation awareness.



**王晓丹** (通信作者), 2004 年于西北工业大学获得博士学位, 现为空军工程大学教授, 主要研究方向为模式识别、智能信息处理。

E-mail: afeu\_wang@163.com

**Wang Xiaodan** (Corresponding author) received her Ph. D. degree from Northwestern Polytechnical University in 2004. Now she is a professor at Air Force Engineering University. Her main research interests include pattern recognition and intelligent information processing.