

Combining belief functions based on distance of evidence

Deng Yong^{a,2}, Shi WenKang^a, Zhu ZhenFu^b, Liu Qi^{c,*,1}

^a*School of Electronics and Information Technology, Shanghai Jiao Tong University, Shanghai 200030, People's Republic of China*

^b*National Defence Key Laboratory of Target and Environment Feature, Beijing 100854, People's Republic of China*

^c*Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Bioinformation Center,
320 YueYang Road, Shanghai 200031, People's Republic of China*

Received 1 May 2003; accepted 1 April 2004

Available online 24 August 2004

Abstract

A modified average method to combine belief function based on distance measures of evidence is proposed. The weight of each body of evidence (BOE) is taken into account. A numerical example is shown to illustrate the use of the proposed method to combine conflicting evidence. Some open issues are discussed in the final section.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Decision-making; Evidence theory; Distance function

1. Introduction

Evidence theory is widely used in many fields such as Information Fusion and decision-making [1,2]. A crucial role in evidence theory is played by Dempster's combination rule that has several interesting mathematical properties such as commutativity and associa-

tivity [1]. However, illogical results may be obtained by classical Dempster combination rule when collected evidence highly conflicts each other. Many methods are proposed to solve this problem [3]. Recently, Murphy [4] presents another problem of classical Dempster's combination rule, the failure to balance multiple bodies of evidence. It is shown that, of the proposed methods, averaging best solves the normalization problems and has much attractive features such as identifying combination problems, showing the distribution of the beliefs and preserving a record of ignorance. However, averaging does not offer convergence toward certainty. To solve this problem, Murphy suggests incorporating average belief into the Dempster's combining rule. However, it can be easily seen that simple average assigns equal weight to each body of evidence and does not consider the

* Corresponding author. Tel.: +86 21 54920080; fax: +86 21 54920143.

E-mail addresses: dengyong@sjtu.edu.cn (D. Yong), liuqi@sibs.ac.cn (L. Qi).

¹ Mail address: 320 YueYang Road, Bioinformation Center, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, People's Republic of China.

² Co-corresponding author. Mail address: Department 8, School of Electronics and Information Technology, Shanghai Jiaotong University, Shanghai, 200030, P.R. China.

association relationship among the evidence collected from multi-sources, which is not reasonable in some real application systems. In this note, based on a distance function of evidence, a weighted average approach to combine conflicting evidence is proposed.

2. Murphy's [4] average approach

In Dempster–Shafer [1,8] evidence theory, there is a fixed set of N mutually exclusive and exhaustive elements, called the *frame of discernment*, which is symbolized by $\Theta = \{H_1, H_2, \dots, H_N\}$. Let us denote $P(\Theta)$ as the power set composed of 2^N elements A of Θ

$$P(\Theta) = \{\emptyset, \{H_1\}, \{H_2\}, \dots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \dots, \Theta\} \quad (1)$$

A basic probability assignment (BPA) is a function from $P(\Theta)$ to $[0, 1]$ defined by:

$$m : \begin{array}{l} P(\Theta) \rightarrow [0, 1] \\ A \mapsto m(A) \end{array} \quad (2)$$

and which satisfies the following conditions:

$$\begin{aligned} \sum_{A \in P(\Theta)} m(A) &= 1 \\ m(\emptyset) &= 0 \end{aligned} \quad (3)$$

The mass $m(A)$ represents how strongly the evidence supports A . The elements of $P(\Theta)$ that have a nonzero mass are called *focal elements*. A body of evidence (BOE) is the set of all the focal elements

$$(\mathcal{R}, m) = \{[A, m(A)] \in AP(\Theta) \text{ and } m(A) > 0\} \quad (4)$$

\mathcal{R} is a subset of $P(\Theta)$, and each of $A \in \mathcal{R}$ has a fixed value $m(A)$. Two bodies of evidence m_1 and m_2 can be combined with Dempster's orthogonal rule as follows

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K} \quad (5)$$

where

$$K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \quad (6)$$

Zadeh [5] has underlined that normalization procedure in the combination rule involves counter-

intuitive behaviors when evidence conflicts. Recently, an averaging method is proposed by Murphy [4]. Murphy suggests that, if all the evidence is available at the same time, one can average the masses, and calculate the combined masses by combining the average values multiple times. As pointed out by Voorbraak [6], in combination with other evidence, the property that probability assigned to a set is not divided among its elements but remains with all the elements can result in an element of the multi-element set receiving a larger belief than seems justified. The result of Murphy's averaging approach seems more reasonable than that of combination without averaging. For more detailed information, please refer to Ref. [4].

3. Modified average approach

In Murphy's approach, all bodies of evidence seem equally important. However, in real system, it is not always the case. The main idea of the proposed approach is that the importance of each body of evidence may be different. How can we determine the importance or the weight of each piece of evidence? In general, if a body of evidence is supported by other collected evidence greatly, this piece of evidence should be more important and has more effect on the final combination results. On the contrary, if a body of evidence is highly conflicting with other bodies of evidence, this piece of evidence should be less important and has little effect on the final combination results. To determine the value of each weight, the distance between two bodies of evidence is used [7]. For more detailed information, please refer to Ref. [7].

Suppose the distance between two bodies of evidence (\mathcal{R}_i, m_i) and (\mathcal{R}_j, m_j) can be calculated by the algorithm in Ref. [7] and is denoted as $d(m_i, m_j)$. The similarity measure Sim_{ij} between the two bodies of evidence and (\mathcal{R}_j, m_j) is defined as:

$$\text{Sim}(m_i, m_j) = 1 - d(m_i, m_j) \quad (7)$$

Suppose the number of bodies of evidence is k . After all the degrees of similarity between the bodies of evidence are obtained, we can construct a *similarity*

measure matrix (SMM), which gives us insight into the agreement between the bodies of evidence.

$$SMM = \begin{bmatrix} 1 & S_{12} & \cdots & S_{1j} & \cdots & S_{1k} \\ \vdots & \vdots & & \vdots & & \vdots \\ S_{i1} & S_{i2} & \cdots & S_{ij} & \cdots & S_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ S_{k1} & S_{n2} & \cdots & S_{nj} & \cdots & 1 \end{bmatrix} \quad (8)$$

The support degree of the body of evidence (\mathcal{R}_i, m_i) ($i=1,2,\dots,k$) is defined as:

$$Sup(m_i) = \sum_{\substack{j=1 \\ j \neq i}}^k Sim(m_i, m_j) \quad (9)$$

The credibility degree Crd_i of the body of evidence (\mathcal{R}_i, m_i) ($i=1,2,\dots,k$) is defined as:

$$Crd_i = \frac{Sup(m_i)}{\sum_{i=1}^k Sup(m_i)} \quad (10)$$

It can be easily seen that $\sum_{i=1}^n Crd_i = 1$, thus, the credibility degree is actually a weight, which shows the relative importance of the collected evidence.

After the definition of the credibility degree, the modified average (or the weight average) of the evidence MAE is given as:

$$MAE(m) = \sum_{i=1}^n (Crd_i \times m_i) \quad (11)$$

If there are n pieces of evidence, one can use the classical Dempster's rule to combine the weighted average of the masses $n-1$ times, which is the same as Murphy's approach [4]. As can be seen from Eqs.

(9)–(11), if a body of evidence is supported by other bodies of evidence greatly, its credibility degree is high and this evidence has more effect on the final combination results. On the contrary, if a piece of evidence is always conflicting with other evidence with high degree, its credibility degree is low and this evidence should have less effect on the final combination results.

4. Numerical example

A fictitious example is illustrated to show the use of the proposed combination rule. In a multisensor-based automatic target recognition system, suppose the real target is A . From five different sensors, the system has collected five bodies of evidence shown as follows:

$$\begin{aligned} (\mathcal{R}_1, m_1) &= ([\{A\}, 0.5], [\{B\}, 0.2], [\{C\}, 0.3]) \\ (\mathcal{R}_2, m_2) &= ([\{A\}, 0], [\{B\}, 0.9], [\{C\}, 0.1]) \\ (\mathcal{R}_3, m_3) &= ([\{A\}, 0.55], [\{B\}, 0.1], [\{A, C\}, 0.35]) \\ (\mathcal{R}_4, m_4) &= ([\{A\}, 0.55], [\{B\}, 0.1], [\{A, C\}, 0.35]) \\ (\mathcal{R}_5, m_5) &= ([\{A\}, 0.60], [\{B\}, 0.1], [\{A, C\}, 0.3]) \end{aligned}$$

The results by different combination rules are shown in Table 1.

As can be seen from Table 1, when conflicting evidence is collected, the classical Dempster's rule for combining beliefs produces illogical results that do not reflect the actual distribution of beliefs. In this case, for the collection of the "bad" evidence m_2 , which may be caused by many factors such as atrocious weather or enemy's jammer or the flaws of the sensor itself, Dempster's combination results show that, though more pieces of evidence collected later support target A , it is impossible that the target is A , which is just

Table 1
Results of different combination rules of evidence

	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
Dempster–Shafer's combination rule	$m(A)=0$ $m(B)=0.8571$ $m(C)=0.1429$	$m(A)=0$ $m(B)=0.6316$ $m(C)=0.3684$	$m(A)=0$ $m(B)=0.3288$ $m(C)=0.6712$	$m(A)=0$ $m(B)=0.1228$ $m(C)=0.8772$
Murphy's average combination rule	$m(A)=0.1543$ $m(B)=0.7469$ $m(C)=0.0988$	$m(A)=0.3500$ $m(B)=0.5224$ $m(C)=0.1276$	$m(A)=0.6027$ $m(B)=0.2627$ $m(C)=0.1346$	$m(A)=0.7958$ $m(B)=0.0932$ $m(C)=0.1110$
Proposed modified average combination rule	$m(A)=0.1543$ $m(B)=0.7469$ $m(C)=0.0988$	$m(A)=0.4861$ $m(B)=0.3481$ $m(C)=0.1657$	$m(A)=0.7773$ $m(B)=0.0628$ $m(C)=0.1600$	$m(A)=0.8909$ $m(B)=0.0086$ $m(C)=0.1005$

against the truth. With incremental evidence, both the simple averaging and weight averaging provide reasonable results. However, when the number of evidence is not adequate to make decision, the proposed method is superior to Murphy's simple average. For example, when the system collects only three pieces of evidence m_1 , m_2 , m_3 , the presented approach draws correct conclusion that the target is A , while the simple average still supports that the target is B . Moreover, as can be seen from the table, the performance of convergence of modified method is better than that of simple average. The main reason for these phenomenon mentioned above is that, by making use of the distance of the evidence, the modified average approach decrease the weight of the "bad" evidence, so the "bad" evidence has less effect on the final combination results.

5. Discussion and conclusions

Dempster's combination operator is a poor solution for the management of the conflict between the various information sources at the normalization step. Of the alternative methods that address the problems, averaging solves the normalization problems to some extent and has much attractive features. The modified average approach based on the distance between the evidence preserves all the desirable properties of the simple average. In addition, compared with simple averaging, the proposed method reflects the association relationship of the evidence and can efficiently handle conflicting evidence with better performance of convergence.

Some problems relative to the weighted average method are discussed as follows.

(1) When is the method useful?

Perhaps the method can be used in decision-making under strong uncertainty situation handled in the framework of Dempster–Shafer theory, such as group decision-making in the autonomous robotics systems and multi-agent systems. In these circumstances, no other information can be obtained except for the collected bodies of evidence while highly conflicting degree makes original Dempster rule produce illogic combination results. It is reasonable to state that an alternative or action can be selected if the alternative or the action is most supported by decision-makers in these data-driven decision-making situations.

(2) Why not just discard the differing evidence?

Generally speaking, if a piece of evidence is absolutely nonsense, deleting it may be the best way since it not only achieves credible combination results but also decreases computational complexity to some extent. However, the problem is how can we determine which one among the collected evidence is absolutely nonsense. Hence, to our opinions, it is feasible to assign the weight to each piece of evidence, which can be regarded as its "confidence" to some extent. The proposed algorithm can guarantee that the differing evidence has little reflection on the final combination results. In the extreme situation, the weight of differing evidence can be approximated as zero, which can be seen as discarding the piece of evidence.

(3) When is the method impractical or misleading?

It is well known that the original Dempster combination rule is already computationally expensive. It can be easily seen that, however, the proposed weighted average is more complex than the classical Dempster rule. So, the proposed method cannot efficiently deal with Information Fusion problems with a large amount of data in real time application systems. Another problem of the proposed method is that, if the piece of evidence is a "bad" one and is repeated many times, the credibility degree will be high and the combination results may be illogical. Thus, the frequency of reports of the sensors also has some reflection on final results. Also, some desirable properties of Dempster combination rule such as commutativity and associativity are not preserved in the proposed method. As a result, the order of the evidence has some reflection on final results. To solve these problems, a more simple and intelligent method that takes the reliability and confidence of each sensor into consideration is needed in the future research.

Acknowledgements

Many thanks for anonymous referees' constructive comments for improving the paper. The work is partially supported by Shanghai Nature Foundation under Contact Number 03ZR14065 and the *National Defence Key Laboratory of Target and Environment Feature* (NDKLTEF) under Contact Number 51476040103JW13.

References

- [1] M. Beynon, B. Curry, P. Morgan, The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modeling, *Omega* 28 (2000) 37–50.
- [2] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, *Annual Mathematics and Statistics* 38 (4) (1967) 325–339.
- [3] A.L. Jousselme, D. Grenier, E. Bosse, A new distance between two bodies of evidence, *Information Fusion* 2 (2001) 91–101.
- [4] E. Lefevre, Belief function combination and conflict management, *Information Fusion* 3 (2002) 149–162.
- [5] C.K. Murphy, Combining belief functions when evidence conflicts, *Decision Support Systems* 29 (2000) 1–9.
- [6] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.
- [7] F. Voorbraak, On the justification of Dempster’s rule of combination, *Artificial Intelligence* 48 (1991) 171–197.
- [8] L. Zadeh, A simple view of the Dempster–Shafer theory of evidence and its implication for the rule of combination, *AI Magazine* 7 (1986) 85–90.



Deng Yong was born in 1975, in Hunan, P. R. China. He received the Bachelor degree in physics from Shanxi Normal University, Xi’an, China in 1997, the Master degree in Instrument engineering from Hunan University, Changsha, China in 2000, the PhD degree in precise instrument and mechanic engineering from Shanghai Jiao Tong University, Shanghai, China in 2003. He is currently a lecturer in school of electronics and information technology, Shanghai Jiao

Tong University. His research interests are in the areas of Information Fusion, fuzzy logic and decision analysis.

Shi Wenkang is a professor in Shanghai Jiaotong University. His research interests are in the areas of data fusion.

Zhu Zhenfu is a director of the National Defence Key Laboratory of Target and Environment Feature. His research interests are in the areas of Information Fusion and pattern recognition.



Liu Qi was born in 1977, in Hunan, P. R. China. She received the Bachelor degree and the Master degree in Instrument engineering from Hunan University, Changsha, China in 1997 and 2000, respectively, the PhD degree in biomedical engineering from Shanghai Jiao Tong University, Shanghai, China in 2003. She is currently an assistant researcher. Her research interests are in the areas of intelligent information processing and bioinformatics.