Elsevier Editorial System(tm) for Information Sciences

Manuscript Draft

Manuscript Number:

Title: A new representation of basic probability assignment in Dempster-Shafer theory

Article Type: Full length article

Keywords: Dempster-Shafer theory; basic probability assignment; linear space; data fusion; distance; entropy.

Corresponding Author: Professor Yong Deng, Ph.D

Corresponding Author's Institution: University of Electronic Science and Technology of China

First Author: Ziyuan Luo

Order of Authors: Ziyuan Luo; Yong Deng, Ph.D

Abstract: Because of the superiority in dealing with uncertainty expression, Dempster-Shafer theory (D-S theory) is widely used in decision theory. In D-S theory, the basic probability assignment (BPA) is the basis and core. Recently, some researchers represent BPA on a N-dimension frame of discernment (FOD) as 2^N-dimension vector in Descartes coordinate system. However, the concept of orthogonality in this method is confused and inexplicable. A new representation method of BPA is proposed in this paper. The BPA on a N-dimension FOD is represented as N-dimension vector with parameters in this method. Then BPA is expressed as subset of N-dimension Cartesian space. The essence of this method is to convert BPA to probability distribution (PD) with parameters. Based on this method, problems in D-S theory can be solved, which include the fusion of BPAs, the distance between BPAs, the correspondence between BPA and probability, and the entropy of BPAs. This representation conforms to the definition of orthogonality, and can get satisfactory computing results.

Suggested Reviewers:

A new representation of basic probability assignment in Dempster-Shafer theory

Ziyuan Luo[a], Yong Deng[b,*]

[a]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
[b]Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, China

Abstract

Because of the superiority in dealing with uncertainty expression, Dempster-Shafer theory (D-S theory) is widely used in decision theory. In D-S theory, the basic probability assignment (BPA) is the basis and core. Recently, some researchers represent BPA on a N-dimension frame of discernment (FOD) as 2^N-dimension vector in Descartes coordinate system. However, the concept of orthogonality in this method is confused and inexplicable. A new representation method of BPA is proposed in this paper. The BPA on a N-dimension FOD is represented as N-dimension vector with parameters in this method. Then BPA is expressed as subset of N-dimension Cartesian space. The essence of this method is to convert BPA to probability distribution (PD) with parameters. Based on this method, problems in D-S theory can be solved, which include the fusion of BPAs, the distance between BPAs, the correspondence between BPA and probability, and the entropy of BPAs. This representation conforms to the definition of orthogonality, and can get satisfactory computing results.

*E-mail address: dengentropy@uestc.edu.cn; prof.deng@hotmail.com

Keywords: Dempster-Shafer theory; basic probability assignment; linear space; data fusion; distance; entropy.

---

1. INTRODUCTION

Dempster-Shafer theory (D-S theory) [6, 29], as an important and widely used uncertain reasoning method, has been receiving increasingly attention. D-S theory assigns probabilities to the power set of events, so it can effectively deal with uncertainty and unknown problems. Because of the superiority in pattern recognition [9, 21, 24] and decision making [1, 3], D-S theory has been applied in various fields [26, 28, 27, 5, 41].

In D-S theory, A complete set of incompatible basic hypotheses is called a frame of discernment (FOD), which represents all possible answers to a problem. The degree of trust assigned to each subset of FOD is called the basic probability assignment (BPA). Since the N-dimension FOD has 2^N subsets, some researchers represent the BPA as 2^N-dimension vector in Descartes coordinate system [20, 4]. However, each dimension in Cartesian space is orthogonal to another. In other words, the dimensions separately represented by any two subsets in the set are mutually orthogonal. There are at least two problems. One is that how can two sets with non-empty intersection be mutual orthogonal? The other is that how can empty set represent a equipotent dimension with other non-empty sets? Because of these problems, this method has great limitations.

Based on the idea of converting BPA to probability distribution (PD), a new interpretation of D-S theory is proposed in this paper. The BPA on a N-dimension FOD is represented as N-dimension vector with parameters in this method. Each dimension indicates a hypothesis in FOD. In fact, BPA is represented as a subset of N-dimension space because of the variable parameters. Since the hypotheses in FOD are incompatible, the problem of *orthogonal* is no longer exists. Based on this representation method, several problems in D-S theory have been studied, which include the fusion of BPAs, the distance between BPAs, the correspondence between BPA and probability, and the entropy of BPAs.

The paper is organized as follows. In section 2, we review the basic definitions about D-S theory, the fusion of BPAs, and the traditional vector representation of BPA. In section 3, we propose and discuss the improved representation of BPA. In section 4, we discuss the fusion of BPAs by the new method. In section 5, we define and explain the distance between BPAs. In section 6, we discuss the correspondence between BPA and probability. In section 7, we discuss the entropy of BPAs. In section 8, we have a brief summarization.

2. PRELIMINARIES

2.1. Dempster-Shafer theory

D-S theory can be applied to expert systems, and has the ability to deal with uncertain information. As an uncertain reasoning method, the main feature of the evidence theory is to satisfy the weaker conditions than the Bayesian probability theory, and it has the ability to directly express uncertainty and unknown.

Let Θ be an exhaustive set of all hypotheses of a random variable, and the elements in Θ are mutually exclusive. The set Θ is called the frame of discernment (FOD) [6, 29]. Let Θ have N elements, which is expressed as follows:

$$\Theta = \{H_1, H_2, H_3, \cdots, H_N\}. \tag{1}$$

The power set of Θ, represented by 2^{Θ} , contains all possible subsets of Θ. Obviously, the set 2^{Θ} have 2^N elements. Such that

$$
\begin{aligned}
&2^{\Theta}=\{\emptyset,\{H_1\},\{H_2\},\cdots,\{H_N\},\{H_1\cup H_2\},\\
&\{H_1\cup H_3\},\cdots,\Theta\}.
\end{aligned}
\tag{2}
$$

A pivotal conception in D-S theory is the BPA. A BPA is a mapping m from 2^{Θ} to $[0,1]$ defined as [6, 29]

$$m:2^{\Theta}\to[0,1], \tag{3}$$

which satisfies the following conditions:

$$\sum_{A\in 2^{\Theta}} m(A)=1, \tag{4}$$

$$m(\emptyset)=0. \tag{5}$$

Based on the BPA m, belief function Bel and plausibility function Pl are defined as follows [6, 29]:

$$Pl(A)=\sum_{B\in 2^{\Theta};B\cap A\neq\emptyset} m(B), \tag{6}$$

$$Bel(A)=\sum_{B\in 2^{\Theta};B\subseteq A} m(B). \tag{7}$$

2.2. Dempster's rule of combination

How to combine BPAs from different information sources is a major problem in D-S theory [6, 38, 31, 25, 32, 36, 12]. Because of the great uncertainty of belief function, BPAs from different sources are different. Dempster is the first one to define a combination method [6]. Dempster's rule is to get a new BPA by calculating the orthogonal sum of the known BPAs.

Given two BPAs, $m^{(1)}$ and $m^{(2)}$, let $m^{(12)}$ denote the BPA resulting from $m^{(1)}$ and $m^{(2)}$. Using $\oplus$ to denote orthogonal sum, the Dempster's rule of combination can be expressed as follows [6]:

$$m^{(12)}(A) = \begin{cases} 0, & A = \emptyset, \\ \mu \cdot \sum_{A_1 \cup A_2 = A} m^{(1)}(A_1) \cdot m^{(2)}(A_2), & A \neq \emptyset, \end{cases} \tag{8}$$

where

$$\mu = \sum_{A_1 \cup A_2 \neq \emptyset} m^{(1)}(A_1) \cdot m^{(2)}(A_2). \tag{9}$$

If we ignore the normalization factor, the above formula can be simplified as

$$m^{(12)}(A) = \sum_{A_1 \cup A_2 = A} m^{(1)}(A_1) \cdot m^{(2)}(A_2). \tag{10}$$

2.3. Vector representation of BPA

From a linear algebraic perspective, BPA set on a N-dimension FOD can be represented as a 2^N-dimension vector in Descartes coordinate system. Given a FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$, the power set is 2^Θ. Suppose that

$$A_i \in 2^\Theta \quad \left(i = 1, 2, 3, \cdots, 2^N\right). \tag{11}$$

Then the BPA m set on Θ can be represented as a vector $\vec{M}$, which is expressed as

$$\vec{M} = \left(m(A_1), m(A_2), m(A_3), \cdots, m(A_{2^N})\right)^\mathsf{T}. \tag{12}$$

Moreover, A_i can be habitually defined as

$$i = 1 + \sum_{j \in B} 2^j, \tag{13}$$

$$A_i = \{H_j | j \in B\}. \tag{14}$$

Clearly, in this vector representation method, the dimensions separately represented by any two subsets in Θ are mutually orthogonal. However, at least two problems have been caused. One is that how can two sets with non-empty intersection be mutual orthogonal? The other is that how can empty set represent a equipotent dimension with other non-empty sets? Because of these problems, this method has great limitations.

3. Improved representation of BPA

3.1. Proposed representation method

Definition 1. *Given a FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$, the BPA m on Θ can be represented as a vector $\vec{M}$.*

$$\vec{M} = (M_1, M_2, M_3, \cdots, M_N)^{\mathsf{T}}, \tag{15}$$

where

$$M_j = \sum_{A_i \subseteq \Theta} m(A_i)\kappa_{(H_j|A_i)} \quad \left(j = 1, 2, 3, \cdots, N; i = 1, 2, 3, \cdots, 2^N\right), \tag{16}$$

and the variable parameters $\kappa_{(H_j|A_i)}\left(j = 1, 2, 3, \cdots, N; i = 1, 2, 3, \cdots, 2^N\right)$ satisfy the following conditions:

$$i. \quad \kappa_{(H_j|A_i)} = 0, \ if\ H_j \notin A_i; \tag{17}$$

$$ii. \quad \kappa_{(H_j|A_i)} \in (0, 1], \ if\ H_j \in A_i; \tag{18}$$

$$iii. \quad \sum_{H_j \in A_i} \kappa_{(H_j|A_i)} = 1, \ for\ a\ fixed\ A_i \subseteq \Theta. \tag{19}$$

$\vec{M}$ is called basic probability assignment vector (BPAV), which is not a a vector in the traditional sense as elements of $\vec{M}$ is not fully determined. M_i is called probability assignment quantity (PAQ). It is clear that

$$\sum_{k=1}^{N} M_k = 1. \tag{20}$$

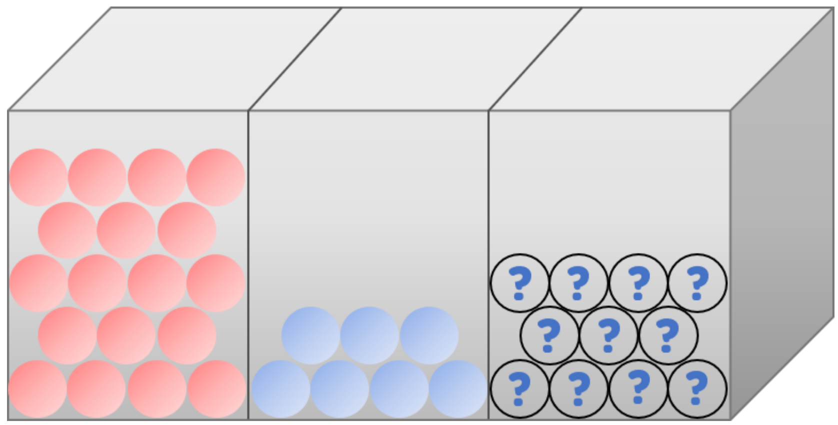

Figure 1: The example of a box containing color balls

$\kappa_{(H_j|A_i)}$ is variable parameter, which cannot be identified only through the BPA m. However, if we get enough information from other sources to identify these parameters, then the BPA is converted to a probability distribution (PD).

Example 1. *As shown in Figure 1, There is a box. All the information is known as follow: 1)there are 100 balls in the box; 2)The balls are only red and blue; 3)50 of the balls are red; 20 of the balls are blue; the color of the remaining 30 balls is unknown.. Take a ball out of the box randomly.*

The FOD here is $\Theta = \{Red, Blue\}$. *The BPA is* $m(Red) = 0.5, m(Blue) = 0.2, m(\Theta) = 0.3$. *Using the proposed method, the BPA can be represented as*

$$\vec{M} = \left(0.5 + 0.3\kappa_{(Red|\Theta)}, 0.2 + 0.3\kappa_{(Blue|\Theta)}\right)^{\mathsf{T}}, \tag{21}$$

Clearly, if the color distribution of the remaining 30 balls is got, the parameters $\kappa_{(Red|\Theta)}, \kappa_{(Blue|\Theta)}$ *can be identified. For example, 20 of the remaining balls are red, and 10 of them are blue. The parameters can be identified as follows:*

$$\kappa_{(Red|\Theta)} = \frac{2}{3}; \tag{22}$$

$$\kappa_{(Blue|\Theta)} = \frac{1}{3}. \tag{23}$$

Then the vector $\vec{M} = (0.7, 0.3)^{\mathsf{T}}$. *In fact, the BPA here is converted to a PD, and* $P(Red) = \vec{M}(1), P(Blue) = \vec{M}(2)$.

3.2. *Geometric interpretation of the method*

The essence of the proposed representation method is to convert BPA to PD with parameters. The variability of the parameters indicates the fuzziness of BPA. Vectors with variable parameters can be represented in Cartesian space. A PD is a point in Cartesian space, while the BPA is a subset of the space.

Example 2. *Given a FOD $\Theta = \{H_1, H_2, H_3\}$, the BPA m is $m(H_1) = 0.3$, $m(H_2) = 0.1$, $m(H_1, H_2) = 0.2$, $m(H_1, H_3) = 0.2$, $m(\Theta) = 0.2$.*

The vector representation of m is

$$\begin{aligned}\vec{M} = (&0.3 + 0.2\kappa_{(H_1|H_1,H_2)} + 0.2\kappa_{(H_1|H_1,H_3)} + 0.2\kappa_{(H_1|\Theta)},\\ &0.1 + 0.2\kappa_{(H_2|H_1,H_2)} + 0.2\kappa_{(H_2|\Theta)},\\ &0.2\kappa_{(H_3|H_1,H_3)} + 0.2\kappa_{(H_3|\Theta)})^{\mathsf{T}}.\end{aligned} \tag{24}$$

As shown in Figure 2, $\vec{M}$ can be represented in Cartesian space. The BPA is a subset of the space, not a point. In some cases, the image can be used to represent the BPA m or BPAV $\vec{M}$.

4. Combination of BPAs

Definition 2. *Suppose there are 2 PAQs from different information sources, $M_k^{(1)}$ and $M_k^{(2)}$, which are expressed as follows:*

$$M_k^{(1)} = \sum_{A_i \subseteq \Theta} m^{(1)}(A_i)\kappa_{(H_k|A_i)} \quad \left(i = 1, 2, 3, \cdots, 2^N\right), \tag{25}$$

$$M_k^{(2)} = \sum_{A_j \subseteq \Theta} m^{(2)}(A_j)\kappa_{(H_k|A_j)} \quad \left(j = 1, 2, 3, \cdots, 2^N\right). \tag{26}$$

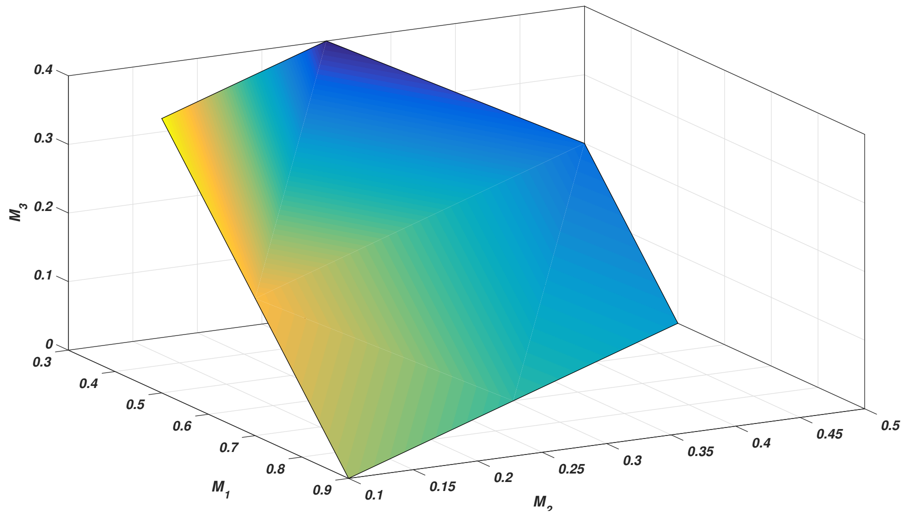

Figure 2: Geometric interpretation of BPAV

The dot product $\cdot$ of $M_k^{(1)}$ and $M_k^{(2)}$ is defined as follows:

$$\begin{aligned} M_k^{(1)} \cdot M_k^{(2)} &= \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m^{(1)}(A_i)\kappa_{(H_k|A_i)} \cdot m^{(2)}(A_j)\kappa_{(H_k|A_j)} \\ &= \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m(A_i)^{(1)} \cdot m^{(2)}(A_j)\left[\kappa_{(H_k|A_i)} \cdot \kappa_{(H_k|A_j)}\right] \\ &= \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m(A_i)^{(1)} \cdot m^{(2)}(A_j)\kappa_{(H_k|A_i \cap A_j)}. \end{aligned} \tag{27}$$

Definition 3. *Suppose there are 2 BPAVs on the same FOD from different information sources, $\vec{M^{(1)}}$ and $\vec{M^{(2)}}$, which are expressed as*

$$\vec{M^{(1)}} = \left(M_1^{(1)}, M_2^{(1)}, M_3^{(1)}, \cdots, M_N^{(1)}\right)^\mathsf{T}, \tag{28}$$

$$\vec{M^{(2)}} = \left(M_1^{(2)}, M_2^{(2)}, M_3^{(2)}, \cdots, M_N^{(2)}\right)^\mathsf{T}. \tag{29}$$

The direct product $\otimes$ of $\vec{M^{(1)}}$ and $\vec{M^{(2)}}$ is defined as follows:

$$\vec{M^{(1)}} \otimes \vec{M^{(2)}} = \left(M_1^{(1)} \cdot M_1^{(2)}, M_2^{(1)} \cdot M_2^{(2)}, M_3^{(1)} \cdot M_3^{(2)}, \cdots, M_N^{(1)} \cdot M_N^{(2)}\right)^\mathsf{T}. \tag{30}$$

Definition 4. *Given 2 BPAVs on the same FOD from different information sources, $\vec{M^{(1)}}$ and $\vec{M^{(2)}}$, $\vec{M^{(12)}}$ denotes the BPA resulting from combining $\vec{M^{(1)}}$ and $\vec{M^{(2)}}$. $\vec{M^{(12)}}$ is defined as*

$$\vec{M^{(12)}} = \vec{M^{(1)}} \otimes \vec{M^{(2)}}. \tag{31}$$

In fact, defining combination in this way has exactly the same meaning as Dempsters rule of combination as equation (10). It is easy to find that direct product $\otimes$ is commutative and associative, which can prove that Dempsters rule of combination is commutative and associative. Equation (31) can be extended to n different information sources. Suppose that $\vec{M^c}$ is the result of combing $\vec{M^{(1)}}, \vec{M^{(2)}}, \vec{M^{(3)}}, \cdots, \vec{M^{(n)}}$. $\vec{M^c}$ can be expressed as

$$\vec{M^c} = \vec{M^{(1)}} \otimes \vec{M^{(2)}} \otimes \vec{M^{(3)}} \otimes \cdots \otimes \vec{M^{(n)}}. \tag{32}$$

Example 3. *Given a FOD $\Theta = \{H_1, H_2\}$, suppose that $m^{(1)}$ and $m^{(2)}$ are two BPAs on Θ from different sources. $m^{(12)}$ is the BPA after combination. The BPAs are known as follows:*

$$\begin{aligned} &m^{(1)}(H_1) = 0.7, \quad m^{(1)}(H_2) = 0.2, \quad m^{(1)}(H_1, H_2) = 0.1; \\ &m^{(2)}(H_1) = 0.6, \quad m^{(2)}(H_1, H_2) = 0.4. \end{aligned} \tag{33}$$

On the one hand, according to equation (10), $m^{(12)}$ can be calculated as

$$m^{(12)}(H_1) = 0.76, \quad m^{(1)}(H_2) = 0.12, \quad m^{(1)}(H_1, H_2) = 0.04. \tag{34}$$

On the other hand, combination can be realized through BPAV. The BPAV of $m^{(1)}$ and $m^{(2)}$ separately are

$$\begin{aligned} \vec{M^{(1)}} &= \left(0.7 + 0.1\kappa_{(H_1|\Theta)}, 0.2 + 0.1\kappa_{(H_2|\Theta)}\right)^{\mathsf{T}}; \\ \vec{M^{(2)}} &= \left(0.6 + 0.4\kappa_{(H_1|\Theta)}, 0.4\kappa_{(H_2|\Theta)}\right)^{\mathsf{T}}. \end{aligned} \tag{35}$$

Therefore, $\vec{M^{(12)}}$ can be calculated as

$$\begin{aligned} \vec{M^{(12)}} &= \vec{M^{(1)}} \otimes \vec{M^{(2)}} \\ &= \left(0.76 + 0.04\kappa_{(H_1|\Theta)}, 0.12 + 0.04\kappa_{(H_2|\Theta)}\right)^{\mathsf{T}}. \end{aligned} \tag{36}$$

By observing equation (34) and equation (36), it can be found that the results are exactly the same.

5. Distance between BPAs

To quantify the similarity between BPAs, researchers have defined different distances between BPAs and apply it to solve problems [8, 14, 22, 18, 23, 19, 17]. Based on the proposed representation method, a new definition of distance D is proposed. $D\left(m^{(1)}, m^{(2)}\right)$ denotes the distance between $m^{(1)}$ and $m^{(2)}$, while $D\left(\vec{M^{(1)}}, \vec{M^{(2)}}\right)$ denotes the distance between two BPAVs, $\vec{M^{(1)}}$ and $\vec{M^{(2)}}$. The proposed definition is based on the definition of distance in linear space.

Definition 5. *$\vec{M^{(1)}}$ and $\vec{M^{(2)}}$ are expressed separately as*

$$\vec{M^{(1)}} = \left(M_1^{(1)}, M_2^{(1)}, M_3^{(1)}, \cdots, M_N^{(1)}\right)^{\mathsf{T}}, \tag{37}$$

$$\vec{M^{(2)}} = \left(M_1^{(2)}, M_2^{(2)}, M_3^{(2)}, \cdots, M_N^{(2)}\right)^{\mathsf{T}}. \tag{38}$$

The distance between two BPAVs are defined as

$$D^2\left(\vec{M^{(1)}}, \vec{M^{(2)}}\right) = \sum_{i=1}^{N} \left(M_i^{(1)} - M_i^{(2)}\right)^2. \tag{39}$$

Clearly, $D\left(\vec{M^{(1)}}, \vec{M^{(2)}}\right)$ has the feature of fuzziness which is not a fixed value. However, this definition is still valuable. As the BPQs satisfy equation (17)-(19), the maximum and minimum of D can be calculated, represented by *maxD* and *minD* respectively.Clearly,

$$maxD \leq \sqrt{2} \quad and \quad minD \geq 0. \tag{40}$$

Example 4. *Given a FOD $\Theta = \{H_1, H_2\}$, the BPA m is as follows:*

$$m(H_1) = 0.5,\ m(H_2) = 0.3,\ m(H_1, H_2) = 0.2. \tag{41}$$

Another BPA m' is

$$m'(H_1) = p,\ m'(H_2) = 0.9 - p,\ m'(H_1, H_2) = 0.1, \tag{42}$$

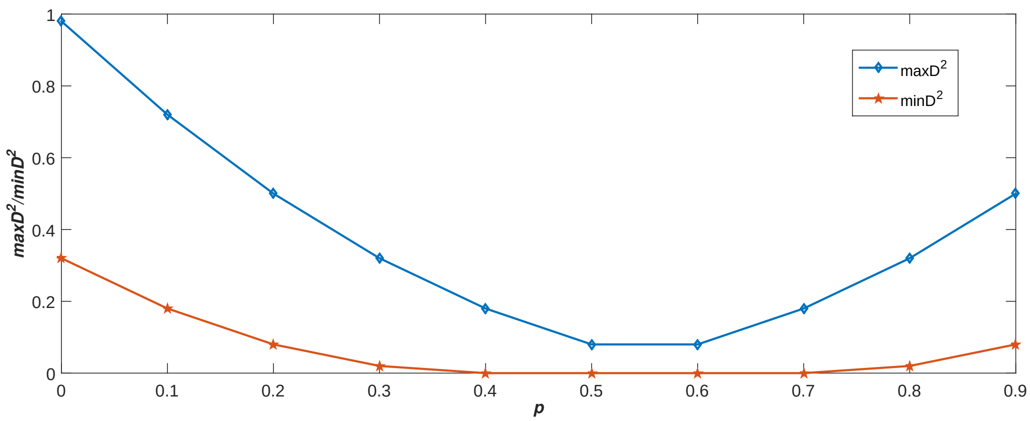

Figure 3: The $maxD^2$ and $minD^2$ between m' and m

where p takes $0,\ 0.1,\ 0.2, \cdots,\ 0.9$ *respectively.*

After calculating the distance D between m and m'*, the results are shown as Table 1 and Figure 3. From the results, when p is between* 0.5 *and* 0.6*, the distance is smallest, which means* m' *is the most similar to m. However, when* $p = 0$*,* m' *and m have the greatest difference, and the distance between them is the biggest.*

p	$maxD^2$	$minD^2$
0	0.98	0.32
0.1	0.72	0.18
0.2	0.50	0.08
0.3	0.32	0.02
0.4	0.18	0
0.5	0.08	0
0.6	0.08	0
0.7	0.18	0
0.8	0.32	0.02
0.9	0.50	0.08

Table 1: The $maxD^2$ and $minD^2$ between m' and m

maxD is also called the *identity distance*, while *minD* is called the *conflict distance*. That means *maxD* is a measure of identity, and *minD* is a measure of conflict. A low *conflict distance* and high *identity distance* illustrate that the BPAs are not in conflict, but also do not support each other.

Example 5. *There is a BPA m on a FOD* $\Theta = \{H_1, H_2\}$ *getting by a known information. Suppose that*

$$m(H_1) = 0.8,\ m(H_2) = 0.1,\ m(H_1, H_2) = 0.1. \tag{43}$$

Now, here comes a new information, by which a BPA m_{new} *is getting. However the new information is useless, which means* $m_{new}(\Theta) = 1$. *There is no additional information. It can be found that minD, or conflict distance, between m and* m_{new} *is 0, so the new information do not conflict with known information. maxD between m and* m_{new} *is 1.27, so the identity distance is high, which means the new information do not support m. Therefore, the identity distance and conflict distance meet the actual situation.*

6. Approximation of BPA with probability

As a BPA assigns probability to each of all the 2^N subsets of the FOD Θ with $|\Theta| = N$. Therefore, the BPA has $2^N - 1$ degrees of freedom, which is large to store and process. Then the problem of approximating BPA with probability arises [35, 34, 2, 13, 11, 10, 40]. A famous example is pignistic transformation [33]. Given a BPA m on FOD Θ. The problem here is to find a transformation function $P : \Theta \to [0, 1]$.

Given a BPA m on $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$, pignistic function $P_{pignistic}$ is defined as follows [33]:

$$P_{pignistic}(H_i) = \sum_{A \subset \Theta, H_i \in A} \frac{1}{|A|} \frac{m(A)}{1 - m(\emptyset)}, \quad m(\emptyset) \neq 0, \tag{44}$$

where $|A|$ is the cardinality of subset A.

Clearly, belief function Bel and plausibility function Pl are both transformation functions. Belief function and plausibility function can be redefined based on the proposed representation method.

Definition 6. *Given a FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$, the BPAV $\vec{M}$ is*

$$\vec{M} = (M_1, M_2, M_3, \cdots, M_N)^{\mathsf{T}}. \tag{45}$$

The belief function Bel and plausibility function Pl can be defined as

$$Bel(H_i) = min\{M_i\} \quad (i = 1, 2, 3, \cdots, N), \tag{46}$$

$$Pl(H_i) = max\{M_i\} \quad (i = 1, 2, 3, \cdots, N). \tag{47}$$

Some new transformation functions are defined based on the geometric interpretation of the proposed representation method, which has been introduced in section 3.

Definition 7. *Suppose a BPAV $\vec{M}$ on FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$. Use π to denote the area formed by BPAV $\vec{M}$ in N-dimension space. The transformation function P_{COG} is defined as*

$$(P_{COG}(H_1), P_{COG}(H_2), P_{COG}(H_3), \cdots, P_{COG}(H_N)) = COG\{\pi\}, \tag{48}$$

where $COG\{\pi\}$ means the centre of gravity of area π.

Definition 8. *Suppose a BPAV $\vec{M}$ on FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$. Use π to denote the area formed by BPAV $\vec{M}$ in N-dimension space. Suppose the vertexes of π are $\{V_1, V_2, V_k\}$. The transformation function P_{SSD} and P_{SD} are separately defined as*

$$(P_{SSD}(H_1), P_{SSD}(H_2), P_{SSD}(H_3), \cdots, P_{SSD}(H_N)) = \underset{P \in R^N}{\arg\min} \sum_{i=1}^{k} \|P - V_i\|^2, \tag{49}$$

$$(P_{SD}(H_1), P_{SD}(H_2), P_{SD}(H_3), \cdots, P_{SD}(H_N)) = \underset{P \in R^N}{\arg\min} \sum_{i=1}^{k} \|P - V_i\|, \tag{50}$$

where $\|P - V_i\|$ is the distance between P and V_i.

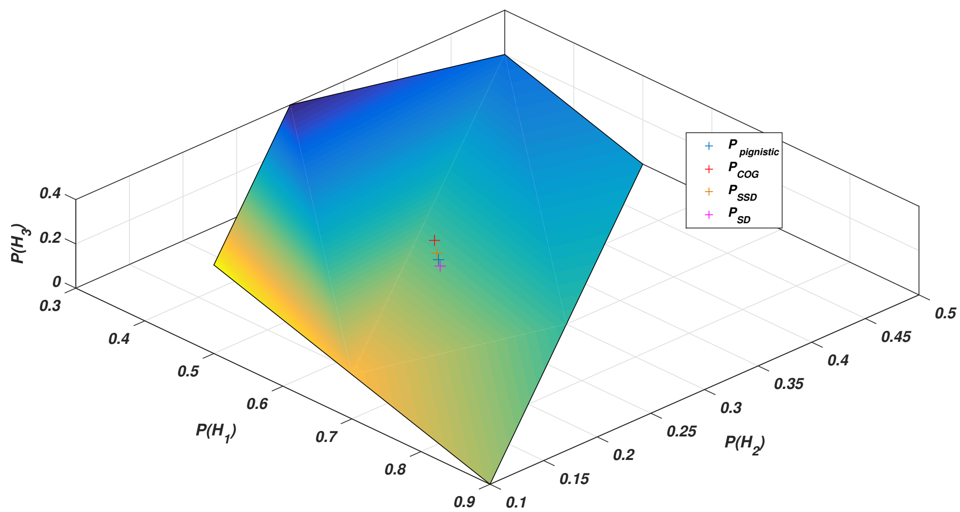

Figure 4: Approximation of BPA with probability

Apparently, function P_{SSD} and P_{SD} are calculated from the square sum of distance and the sum of distance separately.

Example 6. *Use the case in Example 2. The BPA is approximated to probability, and the results are shown as Table 2 and Figure 4. It can be seen that the results are reasonable.*

	$P_{pignistic}$	P_{COG}	P_{SSD}	P_{SD}
$P(H_1)$	0.5667	0.5380	0.5570	0.5760
$P(H_2)$	0.2667	0.2810	0.2710	0.2620
$P(H_3)$	0.1666	0.1810	0.1710	0.1620

Table 2: Approximation of BPA with probability

7. Entropy of BPAs

Shannon has first proposed Shannon entropy in information theory [30]. Entropy for D-S theory indicate the uncertainty of belief function. Many researchers have proposed definitions of entropy of BPA [16, 7, 15, 37, 39]. Based on the proposed method, two types of definition of entropy are given.

Definition 9. *Given a BPA m on FOD $\Theta = \{H_1, H_2, H_3, \cdots, H_N\}$ and a transformation function $P : \Theta \to [0,1]$, the entropy of m, E_P, is defined as*

$$E_P = \sum_{i=1}^{N} P(H_i) \log \left(\frac{1}{P(H_i)} \right). \tag{51}$$

Another type of entropy is defined by regarding the BPAV as PD. Then the entropy is got by applying the Shannon entropy formula.

Definition 10. *Given a BPAV $\vec{M} = (M_1, M_2, M_3, \cdots, M_N)^{\mathsf{T}}$ on FOD Θ, the entropy of $\vec{M}$, E_Q, is defined as*

$$E_Q = \sum_{i=1}^{N} M_i \log \left(\frac{1}{M_i} \right). \tag{52}$$

Clearly, E_Q is not a fixed value. The maximum and minimum value of E_Q are respectively written as $maxE_Q$ and $minE_Q$. The $meanE_Q$ is defined as

$$meanE_Q = \frac{1}{2} \left(maxE_Q + minE_Q \right). \tag{53}$$

In fact, the entropy of BPA comes from two parts, one is measure of conflict, the other is measure of non-specificity. Shannon entropy can measure the first part, while $|maxE_Q - minE_Q|$ can measure the second. Highly non-specific BPA has a huge gap between $maxE_Q$ and $minE_Q$. Therefore, $meanE_Q$ is a good measure of uncertainty, which considers both sources that cause uncertainty.

Example 7. *Given a FOD $\Theta = \{H_1, H_2\}$, the BPA m is*

$$m(H_1) = 0.3,\ m(H_2) = t,\ m(H_3) = 0.7 - t, \tag{54}$$

where t takes $0,\ 0.1,\ 0.2, \cdots,\ 0.7$ *respectively.*

The entropy $E_{P_{COG}}$, $maxE_Q$, $minE_Q$ and $meanE_Q$ of m are calculated, and the results as shown as Table 3 and Figure 5. It can be found from the figure the $|maxE_Q - minE_Q|$ increase with $m(H_1, H_2)$. Compared to $E_{P_{COG}}$, $meanE_Q$ is a better definition of entropy, which more fully expresses uncertainty.

t	$E_{P_{COG}}$	$maxE_Q$	$minE_Q$	$meanE_Q$
0	0.93	1	0	0.50
0.1	0.97	1	0.47	0.73
0.2	0.99	1	0.72	0.86
0.3	1	1	0.88	0.94
0.4	0.99	1	0.88	0.94
0.5	0.97	1	0.88	0.94
0.6	0.93	0.97	0.88	0.93
0.7	0.88	0.88	0.88	0.88

Table 3: The entropy of BPA

8. Conclusion

In this paper, a new representation method of BPA is proposed. The BPA on a N-dimension FOD is represented as N-dimension vector with variable parameters in this method. Then BPA is expressed as subset of N-dimension Cartesian space and has a clear geometric interpretation. Then the methods in Bayesian theory can be applied. With this representation, problems in D-S theory can be solved, which include the fusion of BPAs, the distance between BPAs, the approximation of BPA with probability, and the entropy of BPAs. This representation conforms to

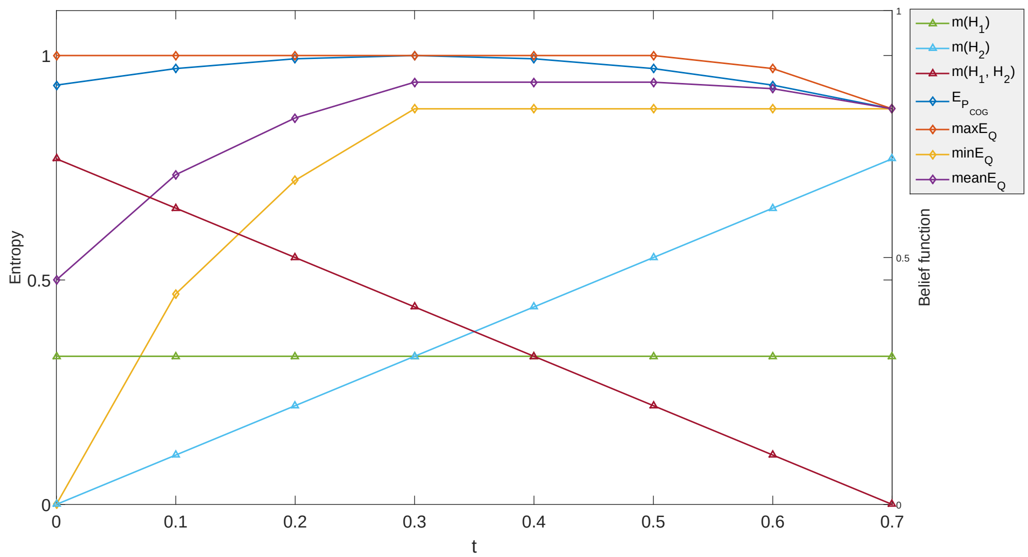

Figure 5: The entropy of BPA

the definition of orthogonality, and can get satisfactory computing results. As the proposed method is basic work in D-S theory, it can be applied to more problems in this field.

Acknowledgment

The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61573290, 61503237).

References

[1] N. O. Attoh-Okine and J. Gibbons. Use of belief function in brownfield infrastructure redevelopment decision making. *Journal of Urban Planning & Development*, 127(3):126–143, 2001.

[2] M. Bauer. Approximation algorithms and decision making in the dempster-shafer theory of evidence ł an empirical study. *International Journal of Approximate Reasoning*, 17(2-3):217–237, 1997.

[3] M. Beynon, B. Curry, and P. Morgan. The dempstercshafer theory of evidence: an alternative approach to multicriteria decision modelling. *Omega*, 28(1):37–50, 2009.

[4] F. Cuzzolin. A geometric approach to the theory of evidence. *IEEE Transactions on Systems Man & Cybernetics Part C*, 38(4):522–534, 2008.

[5] S. Demotier, W. Schon, and T. Denœux. Risk assessment based on weak information using belief functions: a case study in water treatment. *IEEE Transactions on Systems Man & Cybernetics Part C*, 36(3):382–396, 2006.

[6] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.*, 38(2):325–339, 04 1967.

[7] Y. Deng. Deng entropy. *Chaos Solitons & Fractals the Interdisciplinary Journal of Nonlinear Science & Nonequilibrium & Complex Phenomena*, 91:549–553, 2016.

[8] Y. Deng, W. K. Shi, Z. F. Zhu, and Q. Liu. Combining belief functions based on distance of evidence. *Decision Support Systems*, 38(3):489–493, 2005.

[9] T. Denœux. A k-nearest neighbor classification rule based on dempster-shafer theory. *Systems Man & Cybernetics IEEE Transactions on*, 25(5):804–813, 1995.

[10] T. Denœux. *Inner and Outer Approximation of Belief Structures Using a Hierarchical Clustering Approach*. World Scientific Publishing Co., Inc., 2001.

[11] T. Denœux and A. B. Yaghlane. Approximating the combination of belief functions using the fast möbius transform in a coarsened frame. *International Journal of Approximate Reasoning*, 31(1):77–101, 2002.

[12] D. Dubois and H. Prade. On the unicity of dempster rule of combination. *International Journal of Intelligent Systems*, 1(2):133–142, 2010.

[13] R. Haenni and N. Lehmann. Resource bounded and anytime approximation of belief function computations. *International Journal of Approximate Reasoning*, 31(1):103–154, 2002.

[14] D. Q. Han, Y. Deng, C. Z. Han, and Y. Yang. Some notes on betting commitment distance in evidence theory. *Science China*, 55(3):558–565, 2012.

[15] D. HARMANEC and G. J. KLIR. Measuring total uncertainty in dempster-shafer theory: A novel approach. *International Journal of General Systems*, 22(4):405–419, 1994.

[16] R. Jiroušek and P. P. Shenoy. A new definition of entropy of belief functions in the dempster-shafer theory. *International Journal of Approximate Reasoning*, 92:49 – 65, 2018.

[17] A.-L. Jousselme, D. Grenier, and Éloi Bossé. A new distance between two bodies of evidence. *Information Fusion*, 2(2):91 – 101, 2001.

[18] A. L. Jousselme and P. Maupin. *Distances in evidence theory: Comprehensive survey and generalizations*. Elsevier Science Inc., 2012.

[19] A.-L. Jousselme and P. Maupin. Distances in evidence theory: Comprehensive survey and generalizations. *International Journal of Approximate Reasoning*, 53(2):118 – 145, 2012. Theory of Belief Functions (BELIEF 2010).

[20] N. Kami. Integrated formulation of the theory of belief functions from a linear algebraic perspective. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 641–647, 2017.

[21] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin. Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition*, 34(2):299–314, 2001.

[22] M. Li, Y. Hu, Q. Zhang, and Y. Deng. A novel distance function of d numbers and its application in product engineering. *Engineering Applications of Artificial Intelligence*, 47(C):61–67, 2016.

[23] W. Liu. Analyzing the degree of conflict among belief functions. *Artificial Intelligence*, 170(11):909 – 924, 2006.

[24] J. Ma, W. Liu, P. Miller, and H. Zhou. An evidential fusion approach for gender profiling. *Information Sciences*, 333:10 – 20, 2016.

[25] C. K. Murphy. Combining belief functions when evidence conflicts. *Decision Support Systems*, 29(1):1–9, 2000.

[26] R. R. Murphy. Dempster-shafer theory for sensor fusion in autonomous mobile robots. *Robotics & Automation IEEE Transactions on*, 14(2):197–206, 1998.

[27] B. Ristic, M. C. Florea, and Éloi BossÉ. Addendum for the tbm global distance measure for the association of uncertain combat id declarations. *Information Fusion*, 7(3):276–284, 2014.

[28] H. Seraji and N. Serrano. A multisensor decision fusion system for terrain safety assessment. *IEEE Transactions on Robotics*, 25(1):99–108, Feb. 2009.

[29] G. Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.

[30] C. E. Shannon. Ieee xplore abstract - a mathematical theory of communication. *Bell System Technical Journal*, 1948.

[31] P. Smets. The combination of evidence in the transferable belief model. *IEEE Trans*, 12(5):447–458, 1990.

[32] P. Smets. Belief functions: The disjunctive rule of combination and the generalized bayesian theorem. *International Journal of Approximate Reasoning*, 9(1):633–664, 1993.

[33] P. Smets. *Decision making in the TBM: the necessity of the pignistic transformation*. Elsevier Science Inc., 2005.

[34] B. Tessem. *Approximations for efficient computation in the theory of evidence*. Elsevier Science Publishers Ltd., 1993.

[35] F. Voorbraak. *A computationally efficient approximation of Dempster-Shafer theory*. Academic Press Ltd., 1989.

[36] F. Voorbraak. On the justification of dempster's rule of combination. *Artificial Intelligence*, 48(2):171 – 197, 1991.

[37] W. H. Xu, X. Y. Zhang, and W. X. Zhang. Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems. *Applied Soft Computing Journal*, 9(4):1244–1251, 2009.

[38] R. R. Yager. On the dempster-shafer framework and new combination rules. *Information Sciences*, 41(2):93–137, 1987.

[39] R. R. Yager. Uncertainty representation using fuzzy measures. *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society*, 32(1):13–20, 2002.

[40] A. B. Yaghlane, T. Denœux, and K. Mellouli. Coarsening approximations of belief functions. In *Symbolic and Quantitative Approaches To Reasoning with Uncertainty, European Conference, Ecsqaru 2001, Toulouse, France, September 19-21, 2001, Proceedings*, pages 362–373, 2001.

[41] X. Zhou, X. Deng, Y. Deng, and S. Mahadevan. Dependence assessment in human reliability analysis based on d numbers and ahp. *Nuclear Engineering and Design*, 313:243 – 252, 2017.