

# Dempster-Shafer Evidence Theory and Study of Some Key Problems

Ying-Jin Lu and Jun He

**Abstract**—As one of the most important mathematical methods, the Dempster-Shafer (D-S) evidence theory has been widely used in date fusion, risk assessment, target identification, knowledge reasoning, and other fields. This paper summarized the development and recent studies of the explanations of D-S model, evidence combination algorithms, and the improvement of the conflict during evidence combination, and also compared all explanation models, algorithms, improvements, and their applicable conditions. We are trying to provide a reference for future research and applications through this summarization.

**Index Terms**—Combination arithmetic, conflict, Dempster-Shafer (D-S) evidence theory, evidence combination.

## 1. Introduction

The Dempster-Shafer (D-S) evidence theory is based on the work of Dempster<sup>[1],[2]</sup> during the 1960s and successfully extended by Shafer<sup>[3]</sup>. D-S evidence theory is an uncertainty reasoning method and it decomposes the entire problem into several subproblems, sub evidences, and then uses the evidence combination rule to get the solution of the problem. Conventional probability theories based on the probability theory and mathematical statistics argue that the probability is only determined by the frequency of the incident completely (evidence), but not related to people's preferences. The probability is purely objective. Bayes subjective probability theory argues that the probability is a measure of people's preferences or subjective intention. But Bayes subjective probability theory only focuses on human's judgment and ignores the objective evidence, so the probability is purely subjective. D-S evidence theory requests emphasize the objectivity of evidence and the

people's preferences both during probability inference.

Generally, D-S evidence theory differs from traditional probability theories which distinguish ignorance and uncertainty explicit in the evidence combination process. Furthermore, D-S theory allows assigning a probability to not only singletons but also a set of multiple alternative elements<sup>[4],[5]</sup>. These unique characteristics make D-S theory particularly suit for designing and implementing complex systems<sup>[6]</sup> and it has been widely used in information fusion, target identification, fault diagnosis, and other fields for this flexibility in evidence polymerization.

The theory of evidence is often interpreted as an extension of the Bayesian theory of probabilities; however, it has also inspired several models of reasoning under uncertainty, which do not require the probabilistic view. In this section, we introduce some basic concepts of D-S evidence theory.

Let  $\Theta$  be a finite nonempty set called the frame of discernment, or simply the frame.  $\Theta$  is composed of a series of mutually exclusive objects, and all the objects to be identified should be included, that is  $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ , where the object  $\theta_i$  is the conclusion the system should make. There are three important functions in D-S theory: basic probability assignment function (BPA), belief function (Bel), and likelihood function (Pl)<sup>[4],[8]</sup>.

Basic probability assignment function: Assuming the discriminate framework  $\Theta$  is known, how to determine the degree of an uncertain element belongs to a subset of  $\Theta$ . For every subset of  $\Theta$ , a probability can be assigned, which is called the basic probability assignment. The definition is as follows:

$$\sum_{A \in \Theta} m(A) = 1, \quad m(\phi) = 0. \quad (1)$$

The set  $\phi$  means a contradiction which cannot be true in any state, so assign  $m(\phi)$  to be 0.

Belief function: The belief function denotes the total degree to which a grade of the information is supported by the obtained evidence. For grades  $A$  and  $B$  satisfying  $B \subseteq A$ ,  $A \subseteq \Theta$ , and  $B \subseteq \Theta$ , define the following function:

$$\text{Bel} : 2^\Theta \rightarrow [0, 1], \quad \text{Bel}(A) = \sum_{B \subseteq A} m(B) \quad (2)$$

where Bel is the belief function of  $\Theta$ .

Likelihood function: The likelihood function denotes the

Manuscript received March 2, 2015; revised February 29, 2016. This work was supported by the Special Project in Humanities and Social Sciences by the Ministry of Education of China (Cultivation of Engineering and Technological Talents) under Grant No. 13JDGC002.

Y.-J. Lu (corresponding author) and J. He are with the School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: luyingjin@uestc.edu.cn; sc\_hejun@139.com).

Digital Object Identifier: 10.11989/JEST.1674-862X.5030211

degree to which the grading cannot be rejected by the obtained evidence. Given a map  $Pl: 2^\Theta \rightarrow [0, 1]$ , it is defined as

$$Pl(A) = 1 - Bel(\bar{A}) = \sum_{A \cap B \neq \phi} m(B). \quad (3)$$

According to (2), it is easy to derive that the quantity of plausibility of  $A$  is equal to the sum of the masses of  $B$ , whose intersection with  $A$  is not empty, as shown in (4). For all  $A \subseteq \Theta$ ,  $Bel(A)$  forms a lower bound for  $A$  that could possibly happen, and  $Pl(A)$  forms an upper bound for  $A$  to be happen, which is given by (5).

$$Pl(A) = \sum_{B|B \cap A \neq \phi} m(B) \quad (4)$$

$$Bel(A) \leq P(A) \leq Pl(A). \quad (5)$$

Given independent belief functions over the same frame of discernment, we can combine the belief into a common agreement concerning a subset of  $2^\Theta$  and quantify the conflicts using Dempster's rule of combination<sup>[3]</sup>. Given two masses  $m_1$  and  $m_2$ , this combination computes a joint mass for the two pieces of evidence under the same frame of discernment. Dempster's rule is calculated as follows:

$$m_{1,2}(A) = \left( \sum_{Y_1 \cap Y_2 = A} m_1(Y_1)m_2(Y_2) \right) / K,$$

where

$$K = 1 - \sum_{Y_1 \cap Y_2 = \phi} m_1(Y_1)m_2(Y_2). \quad (6)$$

The  $K$  represents the conflict measure of the two belief functions. Whenever two or more functions are combined, the combination rule is associative and commutative.

D-S evidence theory makes off “uncertainty” and “do not know” accurately, it is more accord with our daily behavior, so D-S evidence theory is practicable in engineering. It has been widely used in date fusion, risk assessment, target identification, knowledge reasoning, and other fields. However, several key problems of the evidence theory have not reached consensus, which restricts its further application and development. In recent years, scholars have made a lot of work on the explanation of D-S evidence theory, how to improve the evidence synthesis rules, and how to avoid the paradox during evidence synthesis. Although some scholars have reviewed these studies, their reviews are most about the explanation of D-S evidence theory and how to improve the evidence synthesis rules, not including the conflict comes from both synthesis rules and the source of evidences. And in the process of evidence synthesis, focal elements explosion often brings huge amount of calculation, but previous studies have not paid attention to this point. In this paper, we are trying to make a review on the research of the explanation of D-S evidence theory, the algorithms of evidence combination, and the conflict during evidence combining combined with the recent

research.

## 2. Explanation of D-S Evidence Theory

Ever since Shafer put forwards the framework of D-S theory in [3], many scholars have tried to explain the basic concepts that Shafer ignored, but unfortunately, no one is acknowledged by all scholars. There are four main explanations now: upper and lower probability interpretation, general Bayesian, random decoder model, and transferable belief model.

Upper and lower probability interpretation model<sup>[2]</sup>: Given a probability space  $(\Theta, \ell, p)$ ,  $p$  is a probability measured on  $(\Theta, \ell)$ , and  $\ell$  is a set of  $\Theta$ . If we define  $p_*$  and  $p^*$  for extending  $p$  to  $2^\Theta$ , there are  $p^*(A) \geq p_*(A)$ ,  $p^*(A) = 1 - p_*(\bar{A})$ , then  $A$  is a measurable set if and only if  $p^*(A) = p_*(A) = p(A)$ <sup>[9][11]</sup>. It is not hard to see that the concepts are exactly similar and the belief function and likelihood function are both defined on the decision space. This explanation model can be used even the prior knowledge does not meet the probability of additive. But the shortage is that upper and lower probability interpretation model cannot explain the combinational rule of D-S theory, and the lower probability function does not satisfy the definition of the belief function.

General Bayesian model: When all focal elements meet the independence condition of Bayesian theory, the D-S combination formula is degraded as a Bayesian formula, that is to say, the Bayesian formula is a special case of D-S synthetic formula, all data fusion using Bayesian formula can be used instead of D-S formula. The D-S method satisfies the weaker probability requirement, so the fusion result is often superior to Bayesian method.

Random decoder model: In order to explain the belief function, Shafer and Tversky<sup>[12]</sup> proposed a random decoder model. In this model, all evidence corresponds with a preset  $\mathcal{A}$  and probability  $p$ , if we judge evidence  $B$  is true, we need to preset  $p(c_i|B) = p(c_i)$ ,  $c_i \in \mathcal{A}$ . The unreasonable assumption that the evidence do not change the probability distribution of  $\mathcal{A}$  was criticized by Levi<sup>[13]</sup> and Smets and Kennes<sup>[14][15]</sup>. The random decoder divides all evidence into reliable evidence and shaky evidence accordance with peoples' intuition, but for complex situations, the decoder is not intuitive<sup>[12][17]</sup>. The above three kinds of models are based on the probability theory<sup>[18]</sup>.

Transferable belief model: In order to solve the problem of the preset of the random decoder model, Smet and Kennes<sup>[14][15]</sup> studied the reliability updating of the D-S model, and put forward the transferable belief model. This model presets the evidence is insufficient. The transferable belief model distinguishes two deferent levels: faith level and decision-making level. The faith level is used for acquisition, assignment, and update of belief, belonging to static portions of the model. The decision-making level transfers the belief into decision probability and makes decision, belonging to dynamic portions of the model. To measure the belief, Smets introduced

an inadequate reasoning principle to make the belief distributed on no information. Set the probability function on the faith level as  $\text{betP}$ , then

$$\text{betP}(x, m) = \sum_{x \in A} \frac{m(A)}{|A|}. \quad (7)$$

Then the belief distribution can be gotten from the linear system of equation. The transferable belief model is independent of probability theory, but for the faith that comes from the game, this model inevitably faces the prisoner's dilemma<sup>[19]</sup>.

In addition, Zadeh<sup>[20]</sup>, Dubois and Prade<sup>[21]</sup>, and Pawlak<sup>[22]</sup> are also committed to explain the evidence theory.

The above models explain D-S evidence theory from the sources of evidence theory, the conditions of focal elements, the reliability of evidence, and reliability updating. These models focus on different aspects, making their applicable scopes different. The upper and lower probability interpretation model is suitable for the application that the prior knowledge does not meet the probability of additive, but when all focal elements are independent, the fusion result of D-S evidence theory is more superior than that of Bayesian method. The random decoder method more suits for the environment that the evidence is simple and the reliability can be clearly distinguished. The transferable belief model is not bound by the probability function, but the faith from the game brings prisoner's dilemma. In the actual application, we need to choose the most suitable model according to the practical problem.

### 3. Algorithms of Evidence Combination

The most intuitive shortage of D-S evidence theory is the tremendous calculation from focal elements. In general,  $n$  elements in framework  $\Theta$  often bring  $2^n - 1$  focal elements. If there are 20 elements in framework  $\Theta$ , there are  $1.048576 \times 10^7$  possible focal elements. To solve this problem, there are two main ways: fast algorithm of a special evidence structure and approximation algorithm of decreasing the number of focal elements.

Barnett<sup>[23]</sup> designed a fast algorithm for a simple evidence structure that the evidence supports a hypothesis or not. For evidence reasoning problems that the evidence space can be expressed as tree shape hierarchies (such as medical diagnosis), Gordon and Shortliffe<sup>[24]</sup> designed another fast D-S algorithm. Pearl<sup>[25]</sup> using Bayesian inference in the hypothesis space simplified the calculation process and the amount of calculation. But Shafer<sup>[26]</sup> found in the highly conflict evidence, the result of calculation error is large, so they improved the D-S method and gave a precise algorithm under the hierarchy condition. This kind of algorithm fully embodies the Dempster synthesis rules, and the calculation result is relatively accurate, but the application scope is narrow.

The approximation algorithm is the most efficient method for inducing focal elements. Voorbraak<sup>[27]</sup> found using

Bayesian approximation to replace the reliability function would not affect the result of the synthesis of Dempster's rule, and proved that the reliability function of Bayesian approximate synthesis is equal to the combination of Bayesian reliability function approximation, this method greatly reduces the amount of calculation. Consonant approximation was proposed by Dubois and Prade<sup>[21]</sup>. Elements calculated by this method are nested and less than the assumption of the recognition framework. Consonant approximation is good at evidence expression, but often brings large error, so it is not suitable for practical applications. Tessem<sup>[28]</sup> chose the focal elements of big masses to approximately calculate and put forward the  $(k, l, x)$  approximation method. The  $(k, l, x)$  approximation method is especially suitable for fast rule strength calculation, it not only improves the speed of evidence synthesis but also basically does not affect the decision of the mass functions. Simard *et al.*<sup>[29]</sup> suggested the truncated D-S algorithm. It always keeps the basic probability assignment of "do not know" not zero, namely not depriving existence of the after arrival focal elements, readjusts the  $m(\Theta)$  after each synthetic value, and retains the basic probability of focal elements after the "trim". The algorithm has the advantages of both reducing the computation and ensuring the adaptability of the algorithm; the biggest shortage is that the evidence synthesis order has an impact on the result of the calculation.

In a practical application, the Bayesian approximation method and  $(k, l, x)$  approximation method, in essence, are the conversion of BPAs to Bayesian probability. The difference is how to transfer "not sure" and "do not know" approximate BPAs into "ok" and "know" probabilities. In some sense, the Simard approximate algorithm is closer to the "style" of the conventional D-S method.

Inspired by Pignistic probabilities convert, Burger and Cuzzolin<sup>[30]</sup> put forward two kinds of  $k$ -additive BPAs. The hierarchical clustering method was put forward by Denoeux<sup>[31]</sup> to realize the approximation of inner and outer BPAs. The hierarchical mass distribution method was proposed to achieve the BPA approximation<sup>[32]</sup>. Han *et al.*<sup>[33]</sup> used the distance of evidences and uncertainty measurement to optimize the BPA approximation.

The fast algorithm of a special evidence structure and the method of inducing focal elements have different application environments, advantages, and disadvantages, the principles of choosing a suitable algorithm in a data fusion system include 1) the number of focal elements, 2) the distribution of mass functions, 3) how many mass functions to synthesis, 4) the form of the initial reliability function (a Bayesian reliability function, a belief function, or a simple support function), and 5) the method used for the express of evidence or automatic decision-making.

### 4. Conflict during Evidence Combining

The D-S evidence theory is an important tool for

uncertainty reasoning. In evidence theory, the famous commutative and associative Dempster's rule is used for evidence combinations, when all the sources are considered equally reliable. Although Dempster's rule of combination is well-founded theoretically, its lack of robustness is considered as a limitation by researchers in this field<sup>[34]</sup>. This is because counterintuitive results are obtained in some cases, especially when there is a high conflict among bodies of evidence.

Suppose the discriminate framework  $\Theta=\{A, B, C\}$  and the BPAs of the two evidence are  $m_1(A)=0.99$ ,  $m_1(B)=0.01$ ,  $m_1(C)=0$ ,  $m_2(A)=0$ ,  $m_2(B)=0.01$ , and  $m_2(C)=0.99$ .

From Dempster formula, the conflict measures  $K=m(\phi)=0.0099+0.9801+0.0099=0.9999$ .

The fusion results are  $m_{12}(A)=0$ ,  $m_{12}(B)=1$ , and  $m_{12}(C)=0$ .

Although the support degrees to  $B$  of  $m_1$  and  $m_2$  are comparatively very low, but the fusion results think  $B$  is true. This is obviously perverse. Such results are harmful to decision-making.

There exist two major viewpoints on the so-called counterintuitive combination results. The first is that the counterintuitive results are due to Dempster's rule of combination, especially its normalization step. Thus, a number of researchers have proposed alternative combination rules<sup>[35]-[45]</sup> that use various strategies to redistribute the conflict and provide a fusion tool that produces results that match expectations, such as Yager rules, Lefevre method, DP rules, Quan Sun allocation method, Shanying Zhang allocation method, etc. The second viewpoint is that the counterintuitive results are due to the evidence that is combined, i.e., the data model<sup>[46]-[54]</sup>. According to this viewpoint, there are no counter-intuitive behavior results from the use of Dempster's rule of combination, and the mass functions should be regenerated or modified before combination occurs, such as discount coefficient method, Murphy average method, Jousselme method, etc.

Alternative combination rules are designed for the conflicts assigned on total evidence, and all conflicts will be allocated to all propositions on proportional. Yager<sup>[35]</sup> suggested that the conflicts are the root cause of the failure, all conflicting evidence is unable to provide effective information, so he assigned all conflicts to unknowns  $m(\theta)$ . The improved formula can be used in high conflicting evidence combination, but the irrational distribution will lead to unreasonable results for assigning all conflicting evidence to the unknown. Lefevre *et al.*<sup>[36]</sup> thought conflicting information cannot be completely abandoned. We should extract and analyze the conflicts, then add the combination rules to get the new combination rule, and finally put forward the unify reliability function combination method. Dubois and Prade<sup>[37]</sup> assigned the value of the mass function to all conflicting focal elements, but since there is no distinguish between different focal elements, the composite result is more uncertain. Smets<sup>[38]</sup> believed that the counterintuitive combination is the result of the uncompleted recognition framework, so they treated an empty set as the unknown elements and assigned all conflicts unknown. These

methods change the close of evidence theory and bring more problems. Sun<sup>[39]</sup> thought all evidence credibility is congruent, defined the validity of the evidence coefficient through calculating the average of two conflicts, and gave all conflicts in proportion to each proposition.

In addition, Martin and Osswald<sup>[40]</sup>, Smarandache and Dezert<sup>[41]</sup>, Deng *et al.*<sup>[42]</sup>, and Zhang *et al.*<sup>[43]</sup> proposed improved algorithms of evidence combination, but most of these methods only meet the specific application background. All of them pay too much attention to the allocate space and proportion of the conflicts but neglect the cause for the evidence that is unreliable.

Some methods focusing on correcting the source of evidence have been given. Haenni<sup>[44]</sup> suggested that the combination rule of D-S theory has a solid mathematical basis and is the promotion of Bayesian method. When the evidence is conflicting, the source of evidence should be modified. In order to solve this problem, Shafer<sup>[3]</sup> put forward a general discount coefficient method, however, in practical applications, the reliability of information is different, and the discount factor will also change. Murphy<sup>[45]</sup> calculated the average of all evidence credibility before evidence fusion, but he ignored the credibility of the evidence and the correlation between evidence, the combination results are not ideal. For this phenomenon, Xu *et al.*<sup>[46]</sup> introduced an effective factor to measure the reliability of the evidence sources. Liang *et al.*<sup>[47]</sup> introduced the concept of experts, but these values need to obtain a priori knowledge, so the method is not universal. In addition, Deng *et al.*<sup>[42]</sup> and Ding<sup>[48]</sup> also proposed the method of correcting source. Evidence source revisions speed up the convergence speed of the evidence synthesis and increase the synthesis of reliability, but are easy to cause the losing of information.

As the methods of correcting synthesis and modifying the source of evidence are hard to get general and reasonable applications. Recent years many scholars<sup>[49]-[53]</sup> began to put forward a combination method of these two methods. They tried to take advantage of both so to obtain a more reasonable method. But most of these synthesis methods' theoretical basis is insecure, which only can be applied to specific examples, so it is very difficult to find a truly universal and reasonable fusion method.

Although both types of viewpoints are rational, we prefer the idea that the unreliable source is the cause for the counterintuitive results. One necessary condition for using Dempster's rule of combination is that all the sources are equally reliable. However, in many real applications, all the sources of evidence to be combined may not have equal reliability. Therefore, we think that the correcting of evidence sources to be combined should be modified according to the reliability of their sources, providing a correct assessment of the given problem. The effects of the evidence from more reliable sources should be strengthened, and at the same time, the effects of the evidence from less reliable sources should be weakened.

## 5. Relationship of D-S Evidence Theory and Probability Theory

From Section 1 we know, the D-S evidence theory comes from probability and has a very close relationship with probability theories. From Section 2, we know one of the four main explanations of D-S evidence theory argues that when the BPA is defined on a single subset, the BPA is degraded into probability, as shown in the following example, this is not true.

Suppose  $\Theta = \{\theta_1, \theta_2, \theta_3\}$ , the BPAs are  $m(\{\theta_1\})=0.2$ ,  $m(\{\theta_2\})=0.2$ , and  $m(\{\theta_3\})=0.6$ .

The probabilities are  $p(\{\theta_1\})=0.2$ ,  $p(\{\theta_2\})=0.2$ , and  $p(\{\theta_3\})=0.6$ .

Now we consider whether  $p(\cdot)=m(\cdot)$ , by the additivity of probabilities  $p(\{\theta_1, \theta_2\})=p(\{\theta_1\})+p(\{\theta_2\})=0.4$ .

And for the countable additivity of certain probabilities:  $p(\Theta)=p(\{\theta_1\})+p(\{\theta_2\})+p(\{\theta_3\})=1$ .

All masses of the focal elements except  $\theta_1, \theta_2$ , and  $\theta_3$  are 0. So we have  $m(\{\theta_1, \theta_2\})=0 \neq m(\{\theta_1\})+m(\{\theta_2\})=0.4$ ,  $m(\{\Theta\})=0 \neq m(\{\theta_1\})+m(\{\theta_2\})+m(\{\theta_3\})=1$ .

To obtain the difference more intuitive, the difference between  $p(\cdot)$  and  $m(\cdot)$  is shown in Table 1.

Table 1: Difference between  $p(\cdot)$  and  $m(\cdot)$

$p(\cdot)$	$m(\cdot)$
$p(\theta_1)=0.2$	$m(\{\theta_1\})=0.2$
$p(\theta_2)=0.2$	$m(\{\theta_2\})=0.2$
$p(\theta_3)=0.6$	$m(\{\theta_3\})=0.6$
$p(\theta_1 \cup \theta_2)=p(\theta_1)+p(\theta_2)=0.4$	$m(\{\theta_1 \cup \theta_2\})=0 \neq m(\{\theta_1\})+m(\{\theta_2\})=0.4$
$p(\Theta)=1$	$m(\{\Theta\})=0$

From the Table 1, we know even the BPA is defined on a single subset, the BPA is not satisfying additivity and  $m(\{\Theta\})=1$ , so the BPA is not equivalent to the probability. The BPA is similar to the probability only on formal. On the other hand, the D-S evidence theory can be viewed as an imprecise probability method when the proposition is profiled by upper and lower probabilities, because the probability interval is similar to the belief interval  $[\text{Bel}(A), \text{Pl}(A)]$ .

From Section 2, we know some researchers argue that when the BPA is defined on a single subset, the Dempster's rule is equivalent to Bayes formula. We give an example to compare Dempster's rule and Bayes formula.

Set framework  $\Theta = \{A, B\}$ ,  $P(A)=0.6$ ,  $P(E|A)=0.8$ ,  $P(E|B)=0.2$ , from the Bayes formula, we get:

$$\begin{aligned} P(A|E) &= \frac{P(E|A)P(A)}{P(E|A)P(A) + P(E|B)P(B)} \\ &= \frac{0.8 \times 0.6}{0.8 \times 0.6 + 0.2 \times 0.4} = 0.86 \end{aligned}$$

$$\begin{aligned} P(B|E) &= \frac{P(E|B)P(B)}{P(E|B)P(B) + P(E|A)P(A)} \\ &= \frac{0.2 \times 0.4}{0.2 \times 0.4 + 0.8 \times 0.6} = 0.14. \end{aligned}$$

If we transfer above-mentioned evidence into single-point BPAs, we have  $m_1(A)=0.6$ ,  $m_1(B)=0.4$ ,  $m_2(A)=0.8$ , and  $m_2(B)=0.2$ .

From the Dempster's rule, we have  $m_{12}(A)=0.86$  and  $m_{12}(B)=0.14$ .

The results are same, but it is unable to specify that Dempster's rule is equal to Bayes formula. First, in Dempster's rule, all evidence is equal, our example viewed the prior probability and likelihood function as two independent evidence, it is not reasonable. The second reason is that the example is based on a strong implicit assumption which is  $P(E|A)+P(E|B)=1$ . This assumption is not a necessary condition in Bayes formula, but for BPAs, it is necessary.

In a word, the D-S evidence theory is an inexact promotion of probability.

## 6. Conclusions

Because the D-S evidence theory has the following three requirements, it will not actually achieve expected results: 1) The evidence must be independent, and sometimes it is not easy to meet. 2) There needs a tremendous computing workload during evidence combination. 3) The counterintuitive combination results in evidence combination. In recent years, scholars have made a lot of work on 2) and 3). But for 1), no breakthrough appeared. From the developing of the current D-S evidence theory, the related theories, such as fuzzy set theory<sup>[54][55]</sup>, random set theory<sup>[22]</sup>, analytic hierarchy process<sup>[56]</sup>, and neural network analysis<sup>[57]</sup>, are used to explain and optimize the results of D-S evidence theory.

In addition, the D-S evidence theory is a form of random sets theory, but the random sets theory lacks statistical techniques. The essence of BPAs is the distribution of random variables, and the Dempster's rule is the compute rule of random sets. Both of these are dependent on the study of random sets theory. So in order to expand the application of D-S theory, the best way is enriching the study of random sets theory.

In terms of applications, the D-S evidence theory has been used in intelligent identification systems<sup>[57]</sup>, fault diagnosis<sup>[58][61]</sup>, human resource management<sup>[62]</sup>, risk assessment<sup>[63][64]</sup>, decision-making evaluation<sup>[65]</sup>, etc. With the research deepening and some key problems' solving, its applications will be more widely.

## References

- [1] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Studies in Fuzziness and Soft Computing*, vol. 219, no. 2, pp. 57-72, 2008.
- [2] A. P. Dempster, "A generalization of Bayesian inference," *Journal of the Royal Statistical Society*, vol. 30, no. 30, pp. 73-104, 2008.
- [3] G. Shafer, *A Mathematical Theory of Evidence*, Princeton: Princeton University Press, 1976.
- [4] J. Kohlas and P. A. Monney, *A Mathematical Theory of*

- Hints: An Approach to the Dempster-Shafer Theory of Evidence*, Springer-Verlag, 1995.
- [5] T. Denoeux, "A neural network classifier based on Dempster-Shafer theory," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 30, no. 2, pp. 131-150, 2000.
  - [6] F. Dong, S. M. Shartz, and H.-P. Xu, "Reasoning under uncertainty for skill detection in online auctions using Dempster-Shafer theory," *Intl. Journal of Software Engineering and Knowledge Engineering*, vol. 20, no. 20, pp. 943-973, 2012.
  - [7] R. P. Srivastava, T. J. Mock, and L. Gao, "The Dempster-Shafer theory: An introduction and fraud risk assessment illustration Australian accounting review," *Australian Accounting Review*, vol. 21, no. 3, pp. 282-291, 2011.
  - [8] E. Nusrat and K. Yamada, "A descriptive decision-making model under uncertainty: Combination of Dempster-Shafer theory and prospect theory," *Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 21, no. 1, pp. 79-102, 2012.
  - [9] R. Fagin and J. Y. Halpern, "A new approach to updating beliefs," in *Proc. of Conf. on Uncertainty in Artificial Intelligence*, 1990, pp. 347-374.
  - [10] J. Y. Halpern and R. Fagin, "Two views of belief: Belief as generalized probability and belief as evidence," *Artificial Intelligence*, vol. 54, no. 3, pp. 275-317, 1992.
  - [11] W. Xiong, "Dempster-Shafer evidence theory and explanation," *Journal of South China Normal University*, vol. 3, pp. 15-21, Jun. 2000.
  - [12] G. Shafer and A. Tversky, "Language and designs for probability judgment," *Cognitive Science*, vol. 9, no. 3, pp. 309-339, 1985.
  - [13] I. Levi. Consonance, dissonance and evidentiary methanisms. *Philosophy*. 1993. [Online]. Available: <https://philpapers.org/rec/LEVCD>
  - [14] P. Smets, "The transferable belief model and other interoretations of Dempster-Shafer's model," in *Uncertainty in Artificial Intelligence*, D.-Y. Li and Y. Du, Ed. Beijing: National Defence Industry Press, 1991, pp. 375-383.
  - [15] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, no. 2, pp. 91-96, 1994.
  - [16] G. Shafer and J. Pearl, *Readings in Uncertain Reasoning Morgan Kaufmann*, Bartington: Morgan Kaufmann Publishers Inc., 1990, pp. 298-312.
  - [17] G. Shafer and R. Srivastava, "The Bayesian and belief-function foronalisms: A general perspective for audifing," in *Reading in Uncertain Reasoning*, G. Shafer and J. Pearl, Ed. Burlington: Morgan kaufmann Publishers Inc., 1990, pp. 482-521.
  - [18] C.-F. Xu, W.-D. Geng, and Y.-H. Pan, "Review of theory and applications of Dempster-Shafer evidential reasoning method," *Pattern Recognition and Artificial Intelligence*, vol. 12, no. 4, pp. 424-430, 1999.
  - [19] V. Frans, "On the justification of Dempster's rule of combination," *Artificial Intelligence*, vol. 48, no. 2, pp. 171-197, 1991.
  - [20] L. A. Zadeh, "On the validity of Dempster's rule of combination of evidence," in *Proc. of Intl. Symposium on Memory Management*, 1979, pp. 479-524.
  - [21] D. Dubois and H. Prade, "Default reasoning and possibility theory," *Artificial Intelligence*, vol. 35, no. 2, pp. 243-257, 1988.
  - [22] Z. Pawlak, "Rough sets," *Intl. Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341-356, 1982.
  - [23] J. A. Barnett, "Computational methods for a mathematical theory of evidence," *Studies in Fuzziness and Soft Computing*, vol. 219, pp. 868-875, 2008.
  - [24] J. Gordon and E. H. Shortliffe, "A method for managing evidential reasoning in a hierarchical hypothesis space," *Artificial Intelligence*, vol. 26, no. 3, pp. 323-357, 1985.
  - [25] J. Pearl, "On evidential reasoning in a hierarchy of hypothesis," *Artificial Intelligence*, vol. 28, no. 1, pp. 9-15, 1986.
  - [26] G. Shafer and R. Logan, "Implementing Dempster's rule for hierarchical evidence," *Springer Berlin Heidelberg*, vol. 33, no. 3, pp. 271-298, 2008.
  - [27] F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory," *Intl. Man-Machine Studies*, vol. 30, no. 5, pp. 525-536, 1989.
  - [28] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61, no. 2, pp. 315-329, 1993.
  - [29] M. A. Simard, J. Couture, and E. Bosse, "Data fusion of multiple sensors attribute information for target identity estimation using a Dempster-Shafer evidential combination algorithm," in *Proc. of SPIE—Intl. Society for Optical Engineering*, 1996, pp. 577-588.
  - [30] T. Burger and F. Cuzzolin, "The barycenters of the  $k$ -additive dominating belief functions and the Pignistic  $k$ -additive belief functions," in *Proc. of Workshop on the Theory of Belief Functions*, 2010, pp. 1-6.
  - [31] T. Denoeux, "Inner and outer approximation of belief structures using a hierarchical clustering approach," *Intl. Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, vol. 9, no. 4, pp. 437-460, 2001.
  - [32] J. Dezert, D.-Q. Han, and Z.-G. Liu, "Hierarchical DSMP transformation for decision-making under uncertainty," in *Proc. of the 15th Intl. Conf. on Information Fusion*, 2012, pp. 294-301.
  - [33] D.-Q. Han, J. Dezert, and C.-Z. Han, "New basic belief assignment approximations based on optimization," in *Proc. of the 15th Intl. Conf. on Information Fusion*, 2012, pp. 286-293.
  - [34] Y. Yang, D.-Q. Han, and C.-Z. Han, "Discounted combination of unreliable evidence using degree of disagreement," *Intl. Journal of Approximate Reasoning*, vol. 54, no. 8, pp. 1197-1216, 2013.
  - [35] R. R. Yager, "On the Dempster-Shafer framework and new combination rules," *Information Sciences*, vol. 41, no. 2, pp. 93-137, 1987.
  - [36] E. Lefevre, O. Colot, and P. Vannoorenberghe, "Belief functions combination and conflict management," *Information Fusion*, vol. 3, no. 2, pp. 149-162, 2002.
  - [37] D. Dubois and H. Prade, "Representation and combination of uncertainty with belief functions and possibility measures," *Computational Intelligence*, vol. 4, pp. 244-264, Sept. 1998.
  - [38] P. Smets, "Data fusion in the transferable belief model," in *Proc. of Intl. Conf. on Information Fusion*, 2000, pp. 359-364.
  - [39] Q. Sun, X.-Q. Ye, and W.-K. Gu, "A new combination rule of evidence theory," *Acta Electronica Sinica*, vol. 28, no. 8, pp. 117-119, 2000.
  - [40] A. Martin and C. Osswald, "Toward a combination rule to deal with partial conflict and specificity in belief functions theory," *Computer Science*, vol. 9, no. 8, pp. 1-8, 2008.

- [41] F. Smarandache and J. Dezert, "Applications and advances of DSMT for information fusion," *Instrument Standardization & Metrology*, vol. 368, no.2, pp. 137-160, 2006.
- [42] Y. Deng, W.-K. Shi, and Z.-F. Zhu, "Efficient combination approach of conflict evidence," *Journal of Infrared Millimeter Wave*, vol. 23, no. 1, pp. 27-32, 2004.
- [43] S.-Y. Zhang, Q. Pan, and H.-C. Zhang, "Conflict problem of Dempster-Shafer evidence theory," *Acta Aeronautica et Astronautica Sinica*, vol. 22, no. 4, pp. 369-372, 2001.
- [44] R. Haenni, "Are alternatives to Dempster's rule of combination real alternatives: Comments on 'about the belief function combination and the conflict management problem'," *Information Fusion*, vol. 3, no. 3, pp. 237-239, 2002.
- [45] C. K. Murphy, "Combining belief functions when evidence conflicts," *Decision Support Systems*, vol. 29, no. 1, pp. 1-9, 2000.
- [46] L.-Y. Xu, G.-C. Yin, Y.-S. Gong, and H. Zhao, "Combination rules of various credibility evidences and application," *Journal of North Eastern University*, vol. 23, no. 2, pp. 123-125, 2002.
- [47] C.-Y. Liang, Z.-M. Chen, Y.-Q. Huang, and J.-J. Tong, "A method of dispelling the absurdities of Dempster-Shafer's rule of combination," *Systems Engineering-Theory and Practice*, vol. 25, pp. 7-12, Aug. 2005.
- [48] Y.-J. Ding, S.-L. Wang, and X.-D. Zhao, "An improved method of D-S evidence reasoning," in *Proc. of IEEE Intl. Conf. on Communications Technology and Applications*, 2009, pp. 761-765.
- [49] W.-L. Li and K.-H. Guo, "Combination rules of D-S evidence theory and conflict problem," *System Engineering Theory and Practice*, vol. 30, no. 8, pp. 1423-1432, 2010.
- [50] L.-M. Jiang, J.-L. He, and H. Zhang, "New fusion approach for conflicting evidence in D-S theory of evidence," *Computer Science*, vol. 38, no. 4, pp. 236-238, 2011.
- [51] J.-L. Bai and Y. Wang, "Efficient combination approach to conflict evidence for D-S theory," *System Engineering Theory and Practice*, vol. 31, no. 9, pp. 2106-2109, 2009.
- [52] L. Wang, J. Bai, and Y.-R. Zhang, "A new method for conflict evidence processing in D-S theory," *Science Technology and Engineering*, vol. 13, no. 28, pp. 8497-8501, 2013.
- [53] J. Cao and X. Meng, "Efficient combination of conflict evidence for D-S theory," *Application Research of Computers*, vol. 29, no. 5, pp. 1815-1817, 2012.
- [54] N.-X. Xie, G.-Q. Wen, and Z.-W. Li, "A method for fuzzy soft sets in decision making based on grey relational analysis and D-S theory of evidence: Application to medical diagnosis," *Artificial Intelligence in Medicine*, vol. 64, no. 3, pp. 161-171, 2005.
- [55] T.-M. Liu, "Surface target recognition based on evidence theory, fuzzy inference and multisensory information fusion," *Pattern Recognition & Artificial Intelligence*, vol. 12, no. 1, pp. 25-31, 1999.
- [56] B.-G. Gong, "Methods based on Dempster-Shafer theory for multi-attribute decision making with incomplete information," M.S. thesis, School of Management, University of Science and Technology of China, Chengdu, China, 2007.
- [57] M. Khazaei, H. Ahmadi, M. Omid, and M. Khazaei, "Vibration condition monitoring of planetary gears based on decision level data fusion using Dempster-Shafer theory of evidence," *Journal of Vibroengineering*, vol. 14, no. 2, pp. 838-851, 2012.
- [58] A. Moosavian, M. Khazaei, G. Najafi, M. Kettner, and R. Mamat, "Spark plug fault recognition based on sensor fusion and classifier combination using Dempster-Shafer evidence theory," *Applied Acoustics*, vol. 93, no. 2, pp. 120-129, 2015.
- [59] X.-F. Fan and M.-J. Zuo, "Fault diagnosis of machines based on D-S evidence theory. Part 2: Application of the improved D-S evidence theory in gearbox fault diagnosis," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 377-385, 2006.
- [60] X.-P. Wang, B. Qu, and Y.-N. Wen, "Fault diagnosis for complex fault based on the parallel algorithm of evidence theory synthesizer," *Fluid Machinery*, vol. 42, no. 4, pp. 32-36, 2014.
- [61] Y.-J. Zhang, "The motivation model of post-80s knowledge-worker for private enterprise of Kun Ming," M.S. thesis, School of Management, Kun Ming University of Science and Engineering, Kunming, China, 2013.
- [62] C.-B. Li, G.-F. Qing, P. Li, and T.-T. Yin, "Operational risk assessment of distribution network equipment based on rough set and D-S evidence theory," *Journal of Applied Mathematics*, vol. 2013, pp. 1-7, Nov. 2013.
- [63] J.-P. Yang, H.-Z. Huang, L.-P. He, S.-P. Zhua, and D.-W. Wen, "Risk evaluation in failure mode and effects analysis of aircraft turbine rotor blades using Dempster-Shafer evidence theory under uncertainty," *Engineering Failure Analysis*, vol. 18, no. 8, pp. 2084-2092, 2011.
- [64] Y.-J. Guo, C. Wu, and T.-J. Yu, "An assessment algorithm of credit risk based on BP neural network group and D-S evidential theory," *Journal of Hefei University of Technology*, vol. 23, no. 5, pp. 42-45, 2009.
- [65] V. N. Huynh, Y. Nakamori, T. B. Ho, and T. Murai, "Multiple attribute decision making under uncertainty: The evidential reasoning approach revisited," *IEEE Trans. on Systems, Man & Cybernetics: Part A*, vol. 36, no. 4, pp. 804-822, 2006.



**Ying-Jin Lu** was born in Ningxia, China in 1963. He received the B.S. degree from Southwest Petroleum University, Nanchong in 1988, the M.S. degree from University of Electronic Science and Technology of China (UESTC), Chengdu in 2000, and the Ph.D. degree from UESTC in 2004. Now, he is an associate professor with UESTC. His research interests include logistics and supply chain management, system complexity, and business network privacy protection.



**Jun He** was born in Sichuan, China in 1988. She received the B.S. degree from Shandong University of Science and Technology, Qingdao in 2012. She is currently pursuing the M.S. degree with UESTC. Her research interests include logistics and supply chain management, system complexity, and business network privacy protection.