IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 11, 2019, accepted July 26, 2019, date of publication August 1, 2019, date of current version August 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932390

Divergence Measure of Belief Function and Its Application in Data Fusion

YUTONG SONG[1] AND YONG DENG[2]
[1]Glasgow College, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu 610054, China

Corresponding author: Yong Deng (dengentropy@uestc.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61573290 and Grant 61503237.

ABSTRACT Divergence measure is widely used in many applications. To efficiently deal with uncertainty in real applications, basic probability assignment (*BPA*) in Dempster-Shafer evidence theory, instead of probability distribution, is adopted. As a result, an open issue is that how to measure the divergence of *BPA*. In this paper, a new divergence measure of two *BPA*s is proposed. The proposed divergence measure is the generalization of Kullback-Leibler divergence since when the *BPA* is degenerated as probability distribution, the proposed belief divergence is equal to Kullback-Leibler divergence. Furthermore, compared with existing belief divergence measure, the new method has a better performance under the situation with a great degree of uncertainty and ambiguity. Numerical examples are used to illustrate the efficiency of the proposed divergence measure. In addition, based on the proposed belief divergence measure, a combination model is proposed to address data fusion. Finally, an example in target recognition is shown to illustrate the advantage of the new belief divergence in handling not only extreme uncertainty, but also highly conflicting data.

INDEX TERMS Kullback-Leibler divergence, Dempster-Shafer evidence theory, basic probability assignment, target recognition.

I. INTRODUCTION

Kullback-Leibler divergence [1] is a measure of how one probability distribution is different from another probability distribution, which is widely used in uncertain information processing [2]–[4]. Kullback-Leibler divergence can be modelled by Shannon entropy [5], [6]. To deal with fuzzy data, divergence measure between fuzzy sets is presented [7]–[9]. One of the most important applications of divergence measures may be statistical data processing and a short review is given in [10]. Some works aim to extend the classical measure to handle more complex data [11]–[14]. For example, a typical work is to extend Jensen-Shannon divergence in Hilbert space. The discussion of divergence measure does not yet stop [15] and more recent progress can refer [16].

The real world is very complicated with different types of uncertainty [17]–[20]. Not only the probabilistic information, but also the fuzziness exists in complex system [21]. Hence, lots of math tools such as fuzzy sets [22]–[27], Z-numbers [28]–[30], D numbers [31]–[33], belief structures [34]–[36], game theory [37]–[41], entropy function [42]–[44], Bayesian network and other network methods [45]–[47] are presented. Among these tools, the basic probability assignment(*BPA*) in Dempster-Shafer evidence theory [48], [49] has many merits to deal with uncertainty. For example, compared with probability distribution, it can allocate masses not only to the propositions consisting of single objects [50], [51]. Due to its efficiency to model and combine uncertain information, evidence theory is applied to a lots of real engineerings such as target recognition [52], [53], fault diagnosis [54], [55], sensor data fusion [56]–[58], risk and reliability analysis [59], [59]–[61] conflicting management [62]–[67], decision making [68]–[71], pattern recognition [72]–[75], uncertainty modelling [76], [77] and uncertainty measurement [65], [78]–[81]. These real systems often need to calculate the difference between two *BPA*s. As a result, it is necessary to develop the divergence measure of *BPA*, similar to Kullback-Leibler divergence of probability distribution.

Some typical works have been done [82]. Fei and Deng [83], [84] proposed a divergence measure based on Deng entropy [85], defined as relative entropy of Deng entropy. Xiao proposed a new belief divergence, named as belief Jensen-Shannon(*BJS*) divergence [86] to deal with sensor data fusion. However, all of the existing methods fail to reflect the effect of different kinds of subsets so *BPA* is allocated to single subsets and multiple subsets in the same way. In some

The associate editor coordinating the review of this manuscript and approving it for publication was Daniel Benevides Da Costa.

 This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

real applications under high uncertain environment, the combination results given by combination based on it could be unreliable in some degree.

To address this issue, a new belief divergence is proposed and numerical examples are given to illustrate the properties. The proposed divergence measure is the generalization of Kullback-Leibler divergence for probability distribution since when the basic probability assignment (*BPA*) is degenerated as probability distribution and the belief is only assigned to single subsets, the proposed belief divergence is equal to Kullback-Leibler divergence. One of its advantages is to illustrate the effect of the multiple subsets in *BPA*s when the measuring belief divergence among them. In order to indicate the efficiency of handling uncertain data in highly ambiguous environment, a combination model based on the proposed belief divergence measure is proposed to handle data fusion issues. Moreover, a real application in target recognition is given. From the results of the application, the combination modal based on new belief divergence can give a higher supporting degree to the target than existing classical methods.

The remainder of this paper is organized as follows. Section 2 introduces some preliminaries. In Section 3, a new divergence measure for *BPA* is proposed. The real application in target recognition is illustrated to show the advantage of the new method in Section 4. Finally, the conclusion is given in Section 5.

II. PRELIMINARIES

In this section, some basic preliminaries on Kullback-Leibler divergence [1], Dempster-Shafer evidence theory [48], [49], pignistic probability transformation [87] and belief Jensen-Shannon divergence [86] are introduced.

A. KULLBACK-LEIBLER DIVERGENCE

Kullback-Leibler divergence divergence [1] is widely used in information theory to measure the different degree between two probabilities.

Definition 1: Given two probability distribution $A = \{A(x_1), A(x_2), \ldots, A(x_n)\}$ *and* $B = \{B(x_1), B(x_2), \ldots, B(x_n)\}$, *Kullback-Leibler divergence between A and B is defined as [1]*

$$Div_{KL}(A, B) = \sum_{i=1}^{n} A(x_i)\log_2\left(\frac{A(x_i)}{B(x_i)}\right) \quad (1)$$

One of the properties of Kullback-Leibler divergence is not symmetric [1].

$$Div_{KL}(A, B) \neq Div_{KL}(B, A)$$

So in some situations, the symmetric way is used as Jessen-Shannon divergence [88], shown as follows.

$$D_{KL} = \frac{Div_{KL}(A, B) + Div_{KL}(B, A)}{2} \quad (2)$$

In addition, in the calculation of Kullback-Leibler divergence, to avoid the denominator to be zero, a very small value such as 10^{-8} is used to replace zero.

B. DEMPSTER-SHAFER EVIDENCE THEORY

The basic concepts of evidence theory, including *BPA* and Dempster combination rule, are introduced as follows.

Definition 2: Θ *is the set of N elements which represent mutually exclusive and exhaustive hypotheses.* Θ *Let be the frame of discernment [48], [49]:*

$$\Theta = \{H_1, H_2, \cdots, H_i, \cdots, H_N\} \quad (3)$$

The power set of Θ *is denoted by* 2^Θ, *and*

$$2^\Theta = \{\emptyset, \{H_1\}, \cdots, \{H_n\}, \{H_1, H_2\}, \cdots, \{H_1, \cdots, H_N\}\} \quad (4)$$

where $\emptyset$ is an empty set.

Definition 3: A mass function m, also called as BPA, is a mapping of 2^Θ, *defined as follows [48], [49].*

$$m: \quad 2^\Theta \rightarrow [0, 1] \quad (5)$$

which satisfies the following conditions:

$$m(\emptyset) = 0 \quad \sum_{A\in 2^\Theta} m(A) = 1 \quad 0 \leq m(A) \leq 1 \quad A \in 2^\Theta \quad (6)$$

The mass $m(A)$ represents how strongly the evidence supports A.

The Dempster combination rule can be used to obtain the combined evidence.

Definition 4: Given two BPAs m_1 *and* m_2, *Dempster combination rule is shown as follows [48].*

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \dfrac{\sum_{B\bigcap C=A} m_1(B)m_2(C)}{1-K} \end{cases} \quad (7)$$

where $K = \sum_{B\bigcap C=\emptyset} m_1(B)m_2(C)$.

It is markable that K is the coefficient to measure the conflict between evidence in evidence theory, and the combination rules could only be used when $K < 1$.

C. PIGNISTIC PROBABILITY TRANSFORMATION

Pignistic probability transform [87] is defined to transfer a *BPA* into a probability distribution.

Definition 5: Let m be a basic belief assignment on the frame of discernment Θ. *Its associated pignistic probability function BetP on* Θ *is defined as [87]*

$$BetP(A) = \sum_{W\subseteq\Theta, A\subseteq W} \frac{1}{|W|}\frac{m(W)}{(1-m(\phi))}, \quad \forall A \subseteq \Theta \quad (8)$$

where $|W|$ is the number of elements of Θ in W. The transformation between m and $BetP$ is called the pignistic transformation.

D. BELIEF JESSEN-SHANNON DIVERGENCE

BPA is generalization of probability distributions, to some degree. As a result, it is obvious that an open issue is to propose the divergence measure of *BPA*. Recently, a belief Jessen-Shannon divergence is applied to deal with target recognition based on sensor data fusion [86].

Definition 6: Let A_i be a hypothesis of the belief function m, and m_1 and m_2 be two BPAs on the same frame of discernment Θ, containing n. mutually exclusive hypothesis. BJS divergence between m_1 and m_2 is defined as [86]

$$BJS(m_1, m_2) = \frac{1}{2}[Div_{KL}(m_1, \frac{m_1 + m_2}{2}) + Div_{KL}(m_2, \frac{m_1 + m_2}{2}) \quad (9)$$

Belief Jensen-Shannon divergence utilizes the mass functions by taking place of probability functions in the Jessen-Shannon divergence. When all hypothesis are assigned to single elements, the belief Jessen-Shannon divergence will degenerate as Jessen-Shannon divergence [86].

III. A NEW DIVERGENCE MEASURE OF BPA

In Dempster-Shafer evidence theory, how to measure the divergence degree among evidences is still a hot issue. A new divergence measure of belief function is proposed in this section.

A. THE PROPOSED DIVERGENCE

In this section, a new divergence measure is defined as follows.

Definition 7: Given two BPA m_1 and m_2, the divergence between m_1 and m_2 is defined as follows.

$$D(m_1, m_2) = \sum_i \frac{1}{2^{|F_i|}-1} m_1(F_i)\log(\frac{m_1(F_i)}{m_2(F_i)}) \quad (10)$$

where $|F_i|$ is the cardinal number of F_i.

where m is a mass function defined on the frame of discernment Θ, and H_i is the focal element of m, $|H_i|$ is the cardinality of H_i.

As the formula of belief divergence measure shows, the belief for each focal element H_i is divided by a term $(2^{H_i}-1)$ which represents the potential number of states in H_i. Compared with Kullback-Leibler divergence in probability theory, elements in D-S evidence theory is consisting of multiple subsets, so the belief and relevant measures are allocated in multiple subsets.

For the sake of symmetry, a symmetrical divergence based on the proposed method is defined as follows.

Definition 8:

$$D(m_1, m_2) = D(m_2, m_1) = \frac{Div(m_1, m_2) + Div(m_2, m_1)}{2} \quad (11)$$

In the following of this paper, the symmetrical divergence is used as the belief divergence measure.

The divergence measure is the generalization of Kullback-Leibler divergence for probability since when *BPA* is degenerated as probability, divergence measure is equal to Kullback-Leibler divergence. The new divergence measure could allocate masses not only to the propositions consisting of single elements, but also to the unions of such objects so it has a better performance in modelling both of uncertainty and imprecision.

B. NUMERICAL EXAMPLES

Some numerical examples are used to illustrate the properties of the proposed measure.

Example 9: Assuming a frame of discernment is $\Omega = \{A, B, C\}$ which is complete and two BPAs m_1 and m_2 are given as follows.

m_1 :
$m_1(A) = 0.5000, m_1(A, B) = 0.4000, m_1(A, B, C) = 0.1000$
m_2 :
$m_2(A) = 0.6000, m_2(A, B) = 0.2000, m_2(A, B, C) = 0.2000$

The belief divergence between m_1 and m_2 is

$$D(m_1, m_2) = \frac{Div(m_1, m_2) + Div(m_2, m_1)}{2} = 0.0372.$$

From Example 9, we could know that the belief divergence between m_1 and m_2 is equal to that between m_2 and m_1. This property is important for divergence to determine the difference among bodies of evidences.

Example 10: Assuming a frame of discernment is $\Omega = \{A, B, C\}$ which is complete and two BPAs m_1 and m_2 are given as follows.

m_1 :
$m_1(A)=0.6000, m_1(A, B)=0.1000, m_1(A, B, C)=0.3000$
m_2 :
$m_2(A)=0.6000, m_2(A, B)=0.1000, m_2(A, B, C)=0.3000$

The belief divergence between m_1 and m_2 is

$$D(m_1, m_2) = \frac{Div(m_1, m_2) + Div(m_2, m_1)}{2} = 0.0000.$$

From Example 10, for the same *BPA*, its corresponding divergence is zero.

Example 11: Assuming a frame of discernment is $\Omega = \{A, B, C\}$ which is complete and two BPAs m_1 and m_2 are given as follows.

m_1 :
$m_1(A) = 0.5000, m_1(B) = 0.4000, m_1(C) = 0.1000$
m_2 :
$m_2(A) = 0.3000, m_2(B) = 0.5000, m_2(C) = 0.2000$

Kullback-Leibler divergence between m_1 and m_2 is:

$$D_{KL}(m_1, m_2) = 0.0968$$

The belief divergence between m_1 and m_2 is:

$$D(m_1, m_2) = 0.0968$$

It can proved that when *BPA* is degenerated as probability and the belief is only assigned to single subsets, the proposed belief divergence is equal to Kullback-Leibler divergence.

C. PROPERTIES

Significant properties of the proposed divergence are shown as follows. Some remarkable properties of Kullback-Leibler divergence are preserved.

(1) $D(m_1, m_2)$ is symmetric.

(2) $D(m_1, m_2)$ is always nonnegative.

(3) $D(m_1, m_2) = 0$ if and only if *BPA* m_1 and m_2 are identical.

(4) When basic probability assignments m_1 and m_2 are degenerated as probability distributions, $D(m_1, m_2)$ is equal to $D_{KL}(m_1,m_2)$, Kullback-Leibler divergence between m_1 and m_2.

D. JUSTIFICATION OF PROPOSED DIVERGENCE MEASURE

In this section, the rationality of proposed divergence is explained with the use of pignistic probability transformation (*PPT*). After transforming *BPA* into probability distribution, pignistic probability can be seen as probability distribution and then the classical Kullback-Leibler divergence is calculated.

A comparison experiment is given to illustrate the reasonability of new divergence. The experiment is about how Kullback-Leibler divergence between pignistic probability and proposed divergence between *BPA*s varies when the subset sizes of *BPA* changes. The result shows that the varying trend of Kullback-Leibler divergence consistent with our new method and thus explains that our method is reasonable intuitively.

E. A JUSTIFICATION EXPERIMENT

Assume there are two sensor reports in Table 1. There are n possible targets and X_i means ith target.

TABLE 1. The value of two *BPA*s.

	X_1	X_2	$\{X_3, ..., X_n\}$
m_1	0.7000	0.1000	0.1000
m_2	0.7000	0.2000	0.2000

The result for proposed divergence is showed in Figure 2.

As the the value of *Num* increases, the value of the proposed divergence varies as shown in Figure 1. It shows that value of proposed divergence decreases as the number of focal elements increases. This advantage illustrates that the proposed divergence can show the effect of multiple sets, while the classical *BJS* divergence fails to reflect the variance of the multiple sets.

The proposed method shows that divergence measure decreases in Figure 1.

A divergence measure based on *PPT* is used to justify the varying trend is reasonable. *BPA* can be transformed into probability distribution and then the divergence between two probability distributions can be obtained by Kullback-Leibler divergence. Table 2 shows the transformed probability distribution.

The calculation of Kullback-Leibler divergence between the two pignistic probability distributions is shown as follows.

$$Div_{KL}(PPT_1, PPT_2) = \sum_{i=1}^{n} PPT_1(x_i)\log_2(\frac{PPT_1(x_i)}{PPT_2(x_i)})$$

The *PPT* method shows that the divergence measure decreases in Figure 2.

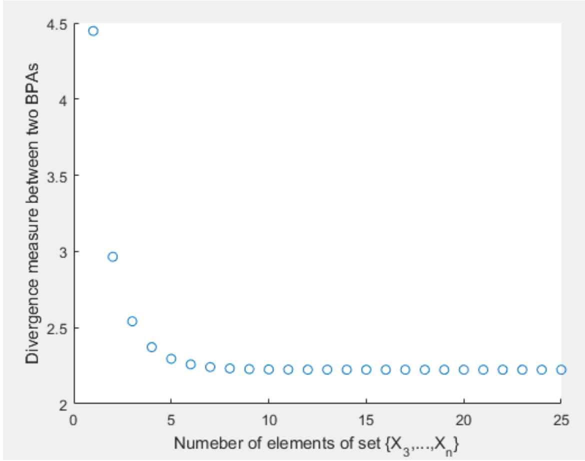

FIGURE 1. The value of new belief divergence for different sizes of BPA.

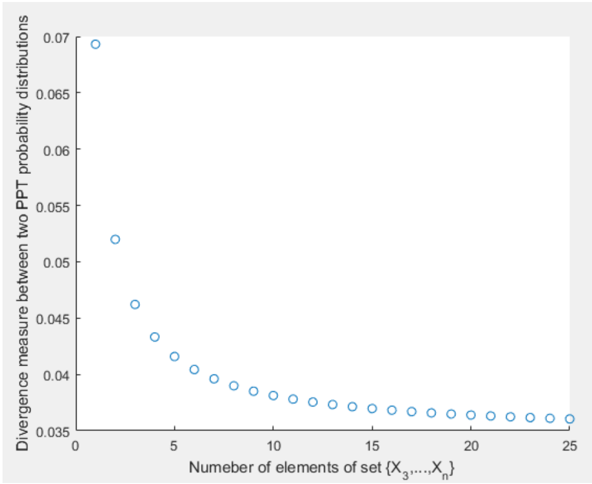

FIGURE 2. The value of Kullback-Leibler divergence for transformed probability distribution.

The Kullback-Leibler divergence between two transformed probability distribution can be obtained as above. As the the value of *Num* increases, the value of Kullback-Leibler divergence varies as shown in Figure 2. It shows that value of proposed divergence decreases as the number of focal elements increases. It can be seen that the trend is similar applied with Kullback-Leibler divergence and proposed divergence. In this way, the rationality of proposed divergence can be proved to some degree.

IV. APPLICATION IN DATA FUSION

In this section, an application of the proposed method in target recognition is used to illustrate its efficiency. First, the data fusion algorithm is briefly introduced. Then, the results and comparisons are shown.

A. DATA FUSION ALGORITHM

Target recognition is a typical sensor data fusion system [89]. Assume that there are n alternatives, denoted as A_i $(i = 1, 2, \ldots, n)$, and k sources of evidence, indicated by $m_j(A_i)$

TABLE 2. Value of two *PPT* probability distributions.

	X_1	X_2	X_3	...	X_i	...	X_n
PPT_1	0.7000	0.1000	$0.1000/(n-2)$	...	$0.1000/(n-2)$	...	$0.1000/(n-2)$
PPT_2	0.7000	0.2000	$0.2000/(n-2)$	...	$0.2000/(n-2)$	...	$0.2000/(n-2)$

is expressed as a *BPA* reported by sensor j ($j = 1, 2, \ldots, k$). The data fusion algorithm is detailed as follows.

Step 1: Calculate the divergence measure matrix (*DMM*) as follows.

$$DMM = \begin{pmatrix} 0\cdots\cdots & D_{1i}\cdots\cdots & D_{1k} \\ D_{i1}\cdots\cdots & 0\cdots\cdots & D_{ik} \\ D_{k1}\cdots\cdots & D_{ki}\cdots\cdots & 0 \end{pmatrix} \quad (12)$$

where D_{ij} is the belief divergence between m_i and m_j.

Step 2: Calculate the supporting degree for m_i based on divergence measure matrix as follows.

$$Sup_i = \frac{1}{\frac{\sum_{j=1, j\neq i}^{k} D_{ij}}{k-1}} \quad (13)$$

Step 3: Normalize the supporting degree to obtain the credibility degree, as the weight of each *BPA*.

$$W_i = \frac{Sup(m_i)}{\sum_{i=1}^{k} Sup(m_i)} \quad (14)$$

The credibility degree is a weight which shows the relative importance of the collected evidence.

Step 4: Obtain the weighted average *BPA* as follows.

$$M(A_j) = \sum_{i=1}^{k} m_i(A_j)W_i \quad (15)$$

where $m_i(A_j)$ is the *BPA* from sensor i about the alternative A_j.

Step 5: Use Dempster rule to combine the averaged weighted evidence $N-1$ times.

B. RESULTS AND COMPARISON

In a target recognition system, possible targets are $\Omega = \{A, B, C\}$. Four sensor reports are collected in Table 2.

First, pieces of evidence are combined by the fusion model proposed in the previous subsection and then make decision about the recognition target. The calculation steps of the proposed method are detailed as follows.

Step 1: Obtain belief divergence measure among pieces of evidences and then built the divergence measure matrix (*DMM*).

$$DMM = \begin{pmatrix} 0.0000 & 2.2984 & 1.0788 & 2.2539 \\ 2.2984 & 0.0000 & 0.0906 & 0.0293 \\ 1.0788 & 0.0906 & 0.0000 & 0.0731 \\ 2.2539 & 0.0293 & 0.0731 & 0.0000 \end{pmatrix}$$

In the calculation of divergence measure matrix, 10^{-8} is used to replace zero to avoid the denominator is zero.

Step 2: Calculate the supporting degree for each m_i from the divergence measure matrix above.

$$Sup_1 = \frac{1}{\frac{\sum_{j=1, j\neq i}^{4} D_{1j}}{4}} = 0.1776$$

$$Sup_2 = \frac{1}{\frac{\sum_{j=1, j\neq i}^{4} D_{2j}}{4}} = 0.4135$$

$$Sup_3 = \frac{1}{\frac{\sum_{j=1, j\neq i}^{4} D_{3j}}{4}} = 0.8048$$

$$Sup_4 = \frac{1}{\frac{\sum_{j=1, j\neq i}^{4} D_{4j}}{4}} = 0.4244$$

If a body of evidence is supported by other bodies evidence greatly, its supporting degree is high and this evidence has more effect on the final combination results.

Step 3: The weight of each evidence is normalized of Sup_i as follows.

$$W_1 = \frac{Sup_1}{\sum_{i=1}^{4} Sup_i} = 0.0976$$

$$W_2 = \frac{Sup_2}{\sum_{i=1}^{4} Sup_i} = 0.2272$$

$$W_3 = \frac{Sup_3}{\sum_{i=1}^{4} Sup_i} = 0.4421$$

$$W_4 = \frac{Sup_4}{\sum_{i=1}^{4} Sup_i} = 0.2331$$

It is obvious that the credibility weight for the first evidence is very low, indicating that it conflicts a lot among evidences.

Step 4: The average weight for each evidence is calculated as follows.

$$M(A) = \sum_{i=1}^{4} m_i(A_1)W_i = 0.2044$$

$$M(B) = \sum_{i=1}^{4} m_i(A_2)W_i = 0.0750$$

$$M(C) = \sum_{i=1}^{4} m_i(A_3)W_i = 0.0906$$

TABLE 3. The sensor reports.

	{A}	{B}	{C}	{A, C}	{A, B, C}
$S_1 : m_1(\bullet)$	0.0000	0.0800	0.1200	0.1000	0.7000
$S_2 : m_2(\bullet)$	0.3000	0.0500	0.0500	0.2000	0.4000
$S_3 : m_3(\bullet)$	0.1500	0.1000	0.1000	0.1500	0.5000
$S_4 : m_4(\bullet)$	0.3000	0.0500	0.1000	0.1000	0.4500

TABLE 4. Supporting degree of possible targets based on different methods.

Method	{A}	{B}	{C}	{A, C}	{A, B, C}
Dempster [48]	0.4879	0.0649	0.1840	0.1666	0.0967
Deng $et al.$ [90]	0.4790	0.0686	0.1868	0.1637	0.1019
Belief Jessen-Shannon divergence [86]	0.4744	0.0666	0.1850	0.1691	0.1049
The proposed method	0.5114	0.0673	0.1752	0.1596	0.0866

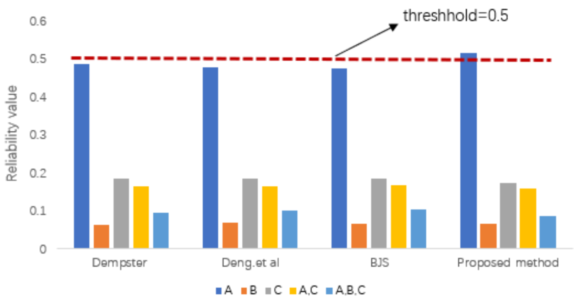

FIGURE 3. Comparison of reliable degrees based on different methods.

$$M(A, C) = \sum_{i=1}^{4} m_i(A_4)W_i = 0.1448$$

$$M(A, B, C) = \sum_{i=1}^{4} m_i(A_5)W_i = 0.4851$$

The averaged weighted evidence shows that the uncertain degree to alternatives is large since the value of *BPA* of {A, B, C} is larger than any single ones.

Step 5: Combine the weighted average *BPA*s by 3 times. Finally, the combining multi-sensor result of the system is obtained as Table 3.

In this experiment, three typical data fusion models, including Dempster rule, Deng *et al*.'s method and belief Jessen-Shannon divergence are selected to compare with the proposed one. Dempster rule is the widest used method based on evidence theory of data fusion and Deng *et al*.'s method is one of the most typical and efficient method to handle conflicting data. Compared with these two methods, the efficiency of fusion model based on new divergence can be convincingly illustrated. The reason to compare with beief Jessen-Shannon divergence is that beief Jessen-Shannon divergence is a typical method to measure divergence between *BPA*s and only difference between the two fusion model is the calculation of divergence measure. So the efficiency of handling highly uncertain and conflicting date can be well illustrated for this comparison.

As can be seen from Table 3, the proposed method has the best performance since our method can get the highest reliability to the correct target (0.5114). Although the difference is not large, only our method can correctly identify the target if we set the threshold at 0.5, which is common in engineering. So it is obvious that in this kind of uncertain environment, target recognition method based on new proposed divergence shows great performance at least it is not worse than any other classical methods.

This application illustrate the efficiency of the new belief divergence measure, not only in conflicting environment, but also in extreme uncertainty. The comparison of values can be intuitively reflected in the Figure 2. The reliability for target A is the most and the value for A, C and A, B, C is the least, which show its great performance in reducing the uncertainty of the final target.

V. CONCLUSION

In order to address the problem of measuring the divergence degree among different evidence, a new belief divergence measure of *BPA* is proposed in this paper. The divergence measure is the generalization of Kullback-Leibler divergence for probability distribution since when the *BPA* is degenerated as probability distribution and the belief is only assigned to single subsets, the proposed belief divergence is equal to Kullback-Leibler divergence. One of the advantages is that the number of elements in subsets is considered. Furthermore, compared with existing belief divergence measures, the proposed divergence can show a better performance under the situation of a greater degree of uncertainty and high conflicts. A real application in target recognition based on sensor data fusion is illustrated the efficiency of the new divergence. Our future work will focus on the exploring efficient data fusion models based on the proposed divergence and its application in engineering to handle uncertain information.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests regarding the publication of this paper.

REFERENCES

[1] S. Kullback, *Information Theory and Statistics*. Chelmsford, MA, USA: Courier, 1997.

[2] A. M. Alaa and M. van der Schaar, "Bayesian nonparametric causal inference: Information rates and learning algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 5, pp. 1031–1046, Oct. 2018.
[3] L. Pardo, *Statistical Inference Based on Divergence Measures*. Boca Raton, FL, USA: CRC Press, 2005.
[4] T. Minka, "Divergence measures and message passing," Microsoft Res., Tech. Rep., 2005.
[5] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.
[6] I. S. Dhillon, S. Mallela, and R. Kumar, "A divisive information-theoretic feature clustering algorithm for text classification," *J. Mach. Learn. Res.*, vol. 3, pp. 1265–1287, Mar. 2003.
[7] S. Montes, I. Couso, P. Gil, and C. Bertoluzza, "Divergence measure between fuzzy sets," *Int. J. Approx. Reasoning*, vol. 30, no. 2, pp. 91–105, Jun. 2002.
[8] M. D. Ansari, A. R. Mishra, and F. T. Ansari, "New divergence and entropy measures for intuitionistic fuzzy sets on edge detection," *Int. J. Fuzzy Syst.*, vol. 20, no. 2, pp. 474–487, Feb. 2018.
[9] I. K. Vlachos and G. D. Sergiadis, "Intuitionistic fuzzy information—Applications to pattern recognition," *Pattern Recognit. Lett.*, vol. 28, no. 2, pp. 197–206, Jan. 2007.
[10] M. Basseville, "Divergence measures for statistical data processing—An annotated bibliography," *Signal Process.*, vol. 93, no. 4, pp. 621–633, Apr. 2013.
[11] Y. He, A. B. Hamza, and H. Krim, "A generalized divergence measure for robust image registration," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1211–1220, May 2003.
[12] H. Chen, B. Jiang, and N. Lu, "An improved incipient fault detection method based on Kullback–Leibler divergence," *ISA Trans.*, vol. 79, pp. 127–136, Aug. 2018.
[13] H. Garg, "A novel divergence measure and its based TOPSIS method for multi criteria decision-making under single-valued neutrosophic environment," *J. Intell. Fuzzy Syst.*, vol. 36, no. 1, pp. 101–115, Feb. 2019.
[14] R. Kompass, "A generalized divergence measure for nonnegative matrix factorization," *Neural Comput.*, vol. 19, no. 3, pp. 780–791, Mar. 2007.
[15] F. Topsøe, "Some inequalities for information divergence and related measures of discrimination," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1602–1609, Jul. 2000.
[16] L. Pardo, *Statistical Inference Based on Divergence Measures*. London, U.K.: Chapman & Hall, 2018.
[17] X. Li, W. Pu, and X. Zhao, "Agent action diagram: Toward a model for emergency management system," *Simul. Model. Pract. Theory*, vol. 94, pp. 66–99, Jul. 2019.
[18] Z. Wang, C. T. Bauch, S. Bhattacharyya, A. D'Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathé, and D. Zhao, "Statistical physics of vaccination," *Phys. Rep.*, vol. 664, pp. 1–113, Dec. 2016.
[19] P. Zhu, Q. Zhi, Z. Wang, and Y. Guo, "Stochastic analysis and optimal design of majority systems," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 66, no. 1, pp. 131–135, Jan. 2019.
[20] L. Chen and H. Yu, "Emergency alternative selection based on an E-IFWA approach," *IEEE Access*, vol. 7, pp. 44431–44440, 2019. doi: 10.1109/ACCESS.2019.2908671.
[21] L. A. Zadeh, "Preliminary draft notes on a similarity-based analysis of time-series with applications to prediction, decision and diagnostics," *Int. J. Intell. Syst.*, vol. 34, no. 1, pp. 107–113, 2019.
[22] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.
[23] A. Mardani, M. Nilashi, E. K. Zavadskas, S. R. Awang, H. Zare, and N. M. Jamal, "Decision making methods based on fuzzy aggregation operators: Three decades review from 1986 to 2017," *Int. J. Inf. Technol. Decis. Making*, vol. 17, no. 2, pp. 391–466, 2018.
[24] J. H. Dahooie, E. K. Zavadskas, M. Abolhasani, A. Vanaki, and Z. Turskis, "A novel approach for evaluation of projects using an interval–valued fuzzy additive ratio assessment (ARAS) Method: A case study of oil and gas well drilling projects," *Symmetry*, vol. 10, no. 2, p. 45, 2018.
[25] F. Xiao, "A hybrid fuzzy soft sets decision making method in medical diagnosis," *IEEE Access*, vol. 6, pp. 25300–25312, 2018.
[26] K. Chatterjee, E. K. Zavadskas, J. Tamošaitienė, K. Adhikary, and S. Kar, "A hybrid MCDM technique for risk management in construction projects," *Symmetry*, vol. 10, no. 2, p. 46, 2018.
[27] M. K. Ghorabaee, M. Amiri, E. K. Zavadskas, and J. Antuchevičienė, "Supplier evaluation and selection in fuzzy environments: A review of MADM approaches," *Econ. Res.-Ekonomska Istraživanja*, vol. 30, no. 1, pp. 1073–1118, 2017.
[28] B. Kang, P. Zhang, Z. Gao, G. Chhipi-Shrestha, K. Hewage, and R. Sadiq, "Environmental assessment under uncertainty using Dempster-Shafer theory and Z-numbers," *J. Ambient Intell. Humanized Comput.*, pp. 1–20, Feb. 2019. doi: 10.1007/s12652-019-01228-y.
[29] L. A. Zadeh, "A note on Z-numbers," *Inf. Sci.*, vol. 181, no. 14, pp. 2923–2932, 2011.
[30] R. R. Yager, "On Z-valuations using Zadeh's Z-numbers," *Int. J. Intell. Syst.*, vol. 27, no. 3, pp. 259–278, Mar. 2012.
[31] F. Xiao, "A multiple-criteria decision-making method based on D numbers and belief entropy," *Int. J. Fuzzy Syst.*, vol. 21, no. 4, pp. 1144–1153, Jun. 2019. doi: 10.1007/s40815-019-00620-2.
[32] J. Zhang, D. Zhong, M. Zhao, J. Yu, and F. Lv, "An optimization model for construction stage and zone plans of rockfill dams based on the enhanced whale optimization algorithm," *Energies*, vol. 12, no. 3, p. 466, Feb. 2019.
[33] X. Deng and Y. Deng, "D-AHP method with different credibility of information," *Soft Comput.*, vol. 23, no. 2, pp. 683–691, 2019.
[34] R. R. Yager and N. Alajlan, "Maxitive belief structures and imprecise possibility distributions," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 4, pp. 768–774, Aug. 2017.
[35] Y. Han and Y. Deng, "A novel matrix game with payoffs of maxitive belief structure," *Int. J. Intell. Syst.*, vol. 34, no. 4, pp. 690–706, 2019.
[36] R. R. Yager, "Fuzzy rule bases with generalized belief structure inputs," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 93–98, Jun. 2018.
[37] Z. Li, D. Jia, H. Guo, Y. Geng, C. Shen, Z. Wang, and X. Li, "The effect of multigame on cooperation in spatial network," *Appl. Math. Comput.*, vol. 351, pp. 162–167, Jun. 2019.
[38] Z. Wang, M. Jusup, L. Shi, J.-H. Lee, Y. Iwasa, and S. Boccaletti, "Exploiting a cognitive bias promotes cooperation in social dilemma experiments," *Nature Commun.*, vol. 9, no. 1, p. 2954, Jul. 2018.
[39] C. Liu, H. Guo, Z. Li, X. Gao, and S. Li, "Coevolution of multi-game resolves social dilemma in network population," *Appl. Math. Comput.*, vol. 341, pp. 402–407, Jan. 2019.
[40] C. Liu, J. Shi, T. Li, and J. Liu, "Aspiration driven coevolution resolves social dilemmas in networks," *Appl. Math. Comput.*, vol. 342, pp. 247–254, Feb. 2019.
[41] K. M. A. Kabir, J. Tanimoto, and Z. Wang, "Influence of bolstering network reciprocity in the evolutionary spatial Prisoner's Dilemma game: A perspective," *Eur. Phys. J. B*, vol. 91, no. 12, p. 312, Dec. 2018.
[42] J. Abellán and J. G. Castellano, "Improving the naive Bayes classifier via a quick variable selection method using maximum of entropy," *Entropy*, vol. 19, no. 6, p. 247, Jun. 2017.
[43] A. L. Kuzemsky, "Temporal evolution, directionality of time and irreversibility," *Rivista Del Nuovo Cimento*, vol. 41, no. 10, pp. 513–574, Oct. 2018.
[44] K. Özkan, "Comparing shannon entropy with deng entropy and improved deng entropy for measuring biodiversity when *a priori* data is not clear," *Forestist*, vol. 68, no. 2, pp. 136–140, Jan. 2018.
[45] Z. Huang, L. Yang, and W. Jiang, "Uncertainty measurement with belief entropy on the interference effect in the quantum-like Bayesian networks," *Appl. Math. Comput.*, vol. 347, pp. 417–428, Apr. 2019.
[46] W. Zhu, H. Yang, Y. Jin, and B. Liu, "A method for recognizing fatigue driving based on Dempster–Shafer theory and fuzzy neural network," *Math. Problems Eng.*, vol. 2017, Dec. 2017, Art. no. 6191035.
[47] C. Liu, C. Shen, Y. Geng, S. Li, C. Xia, Z. Tian, L. Shi, R. Wang, S. Boccaletti, and Z. Wang, "Popularity enhances the interdependent network reciprocity," *New J. Phys.*, vol. 20, no. 12, Dec. 2018, Art. no. 123012.
[48] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Statist.*, vol. 38, no. 2, pp. 325–339, 1967.
[49] G. Shafer, *A Mathematical Theory of Evidence*, vol. 42. Princeton, NJ, USA: Princeton Univ. Press, 1976.
[50] A. D. Jaunzemis, M. J. Holzinger, M. W. Chan, and P. P. Shenoy, "Evidence gathering for hypothesis resolution using judicial evidential reasoning," *Inf. Fusion*, vol. 49, pp. 26–45, Sep. 2019.
[51] J. Vandoni, E. Aldea, and S. Le Hégarat-Mascle, "Evidential query-by-committee active learning for pedestrian detection in high-density crowds," *Int. J. Approx. Reasoning*, vol. 104, pp. 166–184, Jan. 2019.
[52] Y. Han and Y. Deng, "An evidential fractal ahp target recognition method," *Defence Sci. J.*, vol. 68, no. 4, pp. 367–373, 2018.
[53] Y. Zhang, Y. Liu, Z. Zhang, and N. Zhao, "Collaborative fusion for distributed target classification using evidence theory in IoT environment," *IEEE Access*, vol. 6, pp. 62314–62323, 2018.
[54] H. Seiti, A. Hafezalkotob, S. Najafi, and M. Khalaj, "A risk-based fuzzy evidential framework for FMEA analysis under uncertainty: An interval-valued DS approach," *J. Intell. Fuzzy Syst.*, vol. 35, no. 2, pp. 1419–1430, Aug. 2018.

[55] H. Zhang and Y. Deng, "Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory," *Adv. Mech. Eng.*, vol. 10, no. 11, Nov. 2018, Art. no. 1687814018809184. doi: 10.1177/1687814018809184.
[56] X. Su, L. Li, F. Shi, and H. Qian, "Research on the fusion of dependent evidence based on mutual information," *IEEE Access*, vol. 6, pp. 71839–71845, 2018.
[57] X. Su, L. Li, H. Qian, M. Sankaran, and Y. Deng, "A new rule to combine dependent bodies of evidence," *Soft Comput.*, pp. 1–7, Feb. 2019. doi: 10.1007/s00500-019-03804-y.
[58] Y. Song and Y. Deng, "A new method to measure the divergence in evidential sensor data fusion," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 4, Apr. 2019, Art. no. 1550147719841295. doi: 10.1177/1550147719841295.
[59] Z. Wang, J.-M. Gao, R.-X. Wang, K. Chen, Z.-Y. Gao, and Y. Jiang, "Failure mode and effects analysis using Dempster-Shafer theory and TOPSIS method: Application to the gas insulated metal enclosed transmission line (GIL)," *Appl. Soft Comput.*, vol. 70, pp. 633–647, Sep. 2018.
[60] S. Lin, C. Li, F. Xu, D. Liu, and J. Liu, "Risk identification and analysis for new energy power system in China based on D numbers and decision-making trial and evaluation laboratory (DEMATEL)," *J. Cleaner Prod.*, vol. 180, pp. 81–96, Apr. 2018.
[61] Z. Li and L. Chen, "A novel evidential FMEA method by integrating fuzzy belief structure and grey relational projection method," *Eng. Appl. Artif. Intell.*, vol. 77, pp. 136–147, Jan. 2019.
[62] J. Wang, K. Qiao, and Z. Zhang, "An improvement for combination rule in evidence theory," *Future Gener. Comput. Syst.*, vol. 91, pp. 1–9, Feb. 2019.
[63] W. Zhang and Y. Deng, "Combining conflicting evidence using the DEMATEL method," *Soft Comput.*, vol. 23, no. 17, pp. 8207–8216, Sep. 2019. doi: 10.1007/s00500-018-3455-8.
[64] G. Sun, X. Guan, X. Yi, and J. Zhao, "Conflict evidence measurement based on the weighted separate union kernel correlation coefficient," *IEEE Access*, vol. 6, pp. 30458–30472, 2018.
[65] M. D. Mambe, S. Oumtanaga, and G. N. Anoh, "A belief entropy-based approach for conflict resolution in IoT applications," in *Proc. 1st Int. Conf. Smart Cities Communities (SCCIC)*, Jul. 2018, pp. 1–5.
[66] Y. Wang, K. Zhang, and Y. Deng, "Base belief function: An efficient method of conflict management," *J. Ambient Intell. Humanized Comput.*, to be published. doi: 10.1007/s12652-018-1099-2.
[67] J. An, M. Hu, L. Fu, and J. Zhan, "A novel fuzzy approach for combining uncertain conflict evidences in the Dempster–Shafer theory," *IEEE Access*, vol. 7, pp. 7481–7501, 2019.
[68] Z. He and W. Jiang, "A new belief Markov chain model and its application in inventory prediction," *Int. J. Prod. Res.*, vol. 56, no. 8, pp. 2800–2817, 2018.
[69] J.-B. Yang and D.-L. Xu, "Evidential reasoning rule for evidence combination," *Artif. Intell.*, vol. 205, pp. 1–29, Dec. 2013.
[70] J.-B. Yang and D.-L. Xu, "On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 32, no. 3, pp. 289–304, May 2002.
[71] Z. He and W. Jiang, "An evidential Markov decision making model," *Inf. Sci.*, vol. 467, pp. 357–372, Oct. 2018.
[72] Z. Liu, Q. Pan, J. Dezert, and A. Martin, "Combination of classifiers with optimal weight based on evidential reasoning," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1217–1230, Jun. 2017.
[73] Z. Zhang, D. Han, J. Dezert, and Y. Yang, "A new adaptive switching median filter for impulse noise reduction with pre-detection based on evidential reasoning," *Signal Process.*, vol. 147, pp. 173–189, Jun. 2018.
[74] M. Li, Q. Zhang, and Y. Deng, "Evidential identification of influential nodes in network of networks," *Chaos, Solitons Fractals*, vol. 117, pp. 283–296, Dec. 2018.
[75] Z. Liu, Q. Pan, J. Dezert, J.-W. Han, and Y. He, "Classifier fusion with contextual reliability evaluation," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1605–1618, May 2017.
[76] H. Cui, Q. Liu, J. Zhang, and B. Kang, "An improved Deng entropy and its application in pattern recognition," *IEEE Access*, vol. 7, pp. 18284–18292, 2019.
[77] X. Deng, W. Jiang, and Z. Wang, "Zero-sum polymatrix games with link uncertainty: A Dempster–Shafer theory solution," *Appl. Math. Comput.*, vol. 340, pp. 101–112, Jan. 2019.
[78] J. Abellán, "Analyzing properties of Deng entropy in the theory of evidence," *Chaos, Solitons Fractals*, vol. 95, pp. 195–199, Feb. 2017.
[79] M. D. Mambe and T. N'Takpe, N. G. Anoh, and S. Oumtanaga, "A new uncertainty measure in belief entropy framework," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 11, pp. 600–606, Nov. 2018.
[80] Y. Yang and D. Han, "A new distance-based total uncertainty measure in the theory of belief functions," *Knowl.-Based Syst.*, vol. 94, pp. 114–123, Feb. 2016.
[81] L. Pan and Y. Deng, "A new belief entropy to measure uncertainty of basic probability assignments based on belief function and plausibility function," *Entropy*, vol. 20, no. 11, p. 842, 2018.
[82] W. L. Perry and H. E. Stephanou, "Belief function divergence as a classifier," in *Proc. IEEE Int. Symp. Intell. Control*, Aug. 1991, pp. 280–285.
[83] L. Fei and Y. Deng, "Meausre divergence degree of basic probability assignment based on Deng relative entropy," in *Proc. Chin. Control Decis. Conf. (CCDC)*, May 2016, pp. 3857–3859.
[84] L. Fei and Y. Deng, "A new divergence measure for basic probability assignment and its applications in extremely uncertain environments," *Int. J. Intell. Syst.*, vol. 34, no. 4, pp. 584–600, 2018.
[85] Y. Deng, "Deng entropy," *Chaos, Solitons Fractals*, vol. 91, pp. 549–553, Oct. 2016.
[86] F. Xiao, "Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy," *Inf. Fusion*, vol. 46, pp. 23–32, Mar. 2019.
[87] P. Smets, "Decision making in the TBM: The necessity of the pignistic transformation," *Int. J. Approx. Reasoning*, vol. 38, no. 2, pp. 133–147, Feb. 2005.
[88] C. D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999.
[89] X. Guan, H. Liu, X. Yi, and J. Zhao, "The improved combination rule of D numbers and its application in radiation source identification," *Math. Problems Eng.*, vol. 2018, Jul. 2018, Art. no. 6025680.
[90] Y. Deng, W. Shi, Z. Zhu, and Q. Liu, "Combining belief functions based on distance of evidence," *Decis. Support Syst.*, vol. 38, pp. 489–493, Dec. 2004.

YUTONG SONG was appointed as a Research Assistant with the Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, China. She is currently a Junior Student with the Glasgow College of the University of Electronic Science and Technology of China, Chengdu. She does scientific research on information fusion, under the supervision of Prof. Deng.

YONG DENG received the Ph.D. degree in precise instrumentation from Shanghai Jiao Tong University, Shanghai, China, in 2003. From 2005 to 2011, he was an Associate Professor with the Department of Instrument Science and Technology, Shanghai Jiao Tong University. Since 2010, he has been a Professor with the School of Computer and Information Science, Southwest University, Chongqing, China. Since 2012, he has been a Visiting Professor with Vanderbilt University, Nashville, TN, USA. Since 2016, he has been a Professor with the School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China. Since 2017, he has been a Full Professor with the Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, China. He has published more than 200 papers in referred journals, such as the IEEE TRANSACTIONS ON FUZZY SYSTEMS, the IEEE TRANSACTIONS ON CYBERNETICS, the *Decision Support Systems*, and the *European Journal of Operational Research and Scientific Reports*. His research interests include evidence theory, decision-making, information fusion, and complex system modeling. He has served as a program member for many conferences, such as the International Conference on Belief Functions. He has served as many editorial board members such as the Academic Editor of the *PLOS ONE* and the *Applied Intelligence*. He has received numerous honors and awards, including the Elsevier Highly Cited Scientist in China, from 2014 to 2018. He has served as a Reviewer for more than 30 journals, such as the IEEE TRANSACTION ON FUZZY SYSTEMS.