

RESEARCH ARTICLE

# A new weighting factor in combining belief function

Deyun Zhou<sup>1</sup>, Qian Pan<sup>1\*</sup>, Gyan Chhipi-Shrestha<sup>2</sup>, Xiaoyang Li<sup>1</sup>, Kun Zhang<sup>1</sup>, Kasun Hewage<sup>2</sup>, Rehan Sadiq<sup>2</sup>

**1** School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, **2** School of Engineering, University of British Columbia Okanagan, Kelowna, BC, Canada

\* [panq@mail.nwpu.edu.cn](mailto:panq@mail.nwpu.edu.cn), [paco.pan@foxmail.com](mailto:paco.pan@foxmail.com)



**OPEN ACCESS**

**Citation:** Zhou D, Pan Q, Chhipi-Shrestha G, Li X, Zhang K, Hewage K, et al. (2017) A new weighting factor in combining belief function. PLoS ONE 12 (5): e0177695. <https://doi.org/10.1371/journal.pone.0177695>

**Editor:** Yong Deng, Southwest University, CHINA

**Received:** December 19, 2016

**Accepted:** May 2, 2017

**Published:** May 25, 2017

**Copyright:** © 2017 Zhou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data in Example 5 could be obtained in the published paper of "Combination of sources of evidence with different discounting factors based on a new dissimilarity measure." The data in the application of identification of major land use activities contributing to river pollution are from "River water quality assessment of the Manahara river by using macroinvertebrates as biological indicators, Kathmandu Valley" study are available in Figshare (<https://doi.org/10.6084/m9.figshare.4980803>).

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant No.

## Abstract

Dempster-Shafer evidence theory has been widely used in various applications. However, to solve the problem of counter-intuitive outcomes by using classical Dempster-Shafer combination rule is still an open issue while fusing the conflicting evidences. Many approaches based on discounted evidence and weighted average evidence have been investigated and have made significant improvements. Nevertheless, all of these approaches have inherent flaws. In this paper, a new weighting factor is proposed to address this problem. First, a modified dissimilarity measurement is proposed which is characterized by both distance and conflict between evidences. Second, a measurement of information volume of each evidence based on Deng entropy is introduced. Then two kinds of weight derived from aforementioned measurement are combined to obtain a new weighting factor and a weighted average method based on the new weighting factor is proposed. Numerical examples are used to illustrate the validity and effectiveness of the proposed method. In the end, the new method is applied to a real-life application of river water quality monitoring, which effectively identify the major land use activities contributing to river pollution.

## Introduction

Dempster-Shafer (D-S) evidence theory provides a reasonable and efficient way to deal with the information which is uncertain and discordant. It has been extensively used in various applications related to decision-making such as information fusion [1–3], uncertain reasoning [4], fault diagnosis [5], risk analysis [6–9], cognitive map [10], target recognition and association [11–13]. Unlike the probability theory and Bayesian theory, the D-S evidence theory requires few prior conditions and knowledge when information is processed. For example, the Evidential Reasoning (ER) algorithm is a generalized Bayesian inference process and the ER rule reveals that the combined degree of joint support for a proposition from two pieces of independent evidence constitutes two parts in general [14–16]. When there is no priori information, the ER rule will reduce to the D-S combination rule. Moreover, the combination rule of the D-S evidence theory satisfies some of mathematical properties, such as commutativity and associativity. However, counter-intuitive results may occurred by the normalization step of the classical D-S combination rule when collected sources of evidence highly conflict with

61401363) (K.Z.); Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grants (Grant No. CRDPJ 446638-12) (K.H.); Science and Technology on Avionics Integration Laboratory and Aeronautical Science Foundation (Grant No. 20155153034) (K.Z.); Fundamental Research Funds for the Central Universities (Grant No. 3102016AXX005, 3102015BJJ1JGZ009) (K.Z.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

each other, as pointed out by Zadeh [17]. The effectiveness of the D-S evidence theory will be considerably reduced by this deficiency.

Evidently, it is crucial to handle the evidences with high conflict. In the last few years, many researchers have carried out comprehensive research and have applied a series of modifications to the conventional evidence combination rule [18–24]. In general, the existing methods can be divided into two kinds of solutions: revisal of the combination rule and revisal of the evidences, for the problem of high level conflict evidence fusion. Proponents of the first argument present that illogical results are caused by inappropriate distribution of the conflict information. Therefore, the modified methods based on the revisal of the combination rule mainly focus on altering the assignment of conflict information [2, 18, 21, 25–27]. Among them, solutions of the transferable belief model (TBM) and Dezert-Smarandache theory (DSmT) are more popular. The TBM develops a method to transfer the basic belief assignments (BBAs) to probabilities, but the method only can be used in closed world [25, 28]. DSmT extends the assignment universe of BBAs from a power set to a super power set which is more thorough and complete, and corresponding combination models and rules are developed as well [27]. Nonetheless, the modified combination rules have limitations in some situation. For example, most of the modified combination rules are not commutative and associative and are time consuming when dealing with a large amount of evidences.

The other modification, revisal of the evidences, preprocesses conflict evidences before combination process. The favorable mathematical properties of the D-S combination rule are reserved in the improvement as they do not change the D-S combination rule. Many related work have been proposed to support this modification method [22–24, 29–32]. Murphy [29] generates a new evidence by averaging  $N$  evidences with equal weights and then combine it with  $N-1$  times. Based on this idea, Deng [22] proposes a weighted averaging method to obtain the new evidence. Besides the weighted averaging method, the discounting method also plays an important role in preprocessing conflict evidences. The weighting factor of both the weighted average method and discounting method can be identified by evidence distance, which is usually used to describe the conflict or dissimilarity [33–35]. Liu [35] argues that the conflict coefficient  $k$  in the evidence theory is inadequate to reflect the degree of conflict and dissimilarity between evidences, and he utilizes a two-dimension cell  $\langle \text{evidence conflict, evidence distance} \rangle$  to measure the dissimilarity between evidences. The cell is indeed more comprehensive and adequate than the single coefficient  $k$  when describing the dissimilarity, but it also has intrinsic shortcomings in practical situation. For instance, the conflict tolerance threshold  $\varepsilon$  is largely subjective and depends on the perception of a decision maker. In addition, the cell is not syncretized in the combination rule. A dissimilarity measure is proposed on the basis of Hamacher T-conorm fusion rules given by Liu, who considers not only the evidence conflict and distance but also combines it in the combination rule as a discount [36]. Nevertheless, there are twofold limitations associated with the mathematical modeling. First, since the conflict factors only use the maximal subjective probability of the BBAs, it cannot solve the situation related to the propositions with equal belief values, which are investigated in Section 3. Second, combining the evidence one by one has a low convergence rate.

Besides evidence conflict and distance, the evidence volume is another criterion to measure the importance of an evidence [37]. If an evidence has more information, it should have a greater impact on the final aggregated result. Deng entropy [38], as a generalization of Shannon entropy, can measure the evidence information volume under the framework of D-S evidence theory. In this paper, Deng entropy and modified dissimilarity measure are used to form a new weighting factor. Then the new combination rule of evidence is carried out based on the new weighting factor, which has improved the versatility and has a fast convergence rate.

This paper is organized as follows. Section 2 describes some basic concepts related to the D-S evidence theory and dissimilarity measure. Section 3 presents problems of existing conflict coefficients, especially the limitations of Liu’s method. Section 4 investigates the new weighting factor of modified dissimilarity and Deng entropy, and some examples and analysis are presented to show the superiority and effectiveness of proposed method. In Section 5, the proposed method is used in a real-life application of the identification of water pollution sources. Finally, conclusions are drawn in Section 6.

## Preliminaries

### 2.1 Basics of D-S evidence theory

*Definition 1.* Suppose  $\Theta$  be a nonempty finite set of mutually exclusive alternatives and defined as frame of discernment. Set of all the possible subsets of  $\Theta$ , denoted by  $2^\Theta$ , is called power set. The mapping  $m: 2^\Theta \rightarrow [0,1]$  is defined as the basic belief assignment (BBA) (also known as basic probability assignment, BPA) [39, 40]. The BBA satisfies

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{1}$$

$$m(\emptyset) = 0 \tag{2}$$

where  $m(A)$  reflects the strength of each of evidence support for the proposition  $A$  in the frame of discernment, and  $\emptyset$  denotes the empty set of  $\Theta$ .  $A$  is called the focal element, if  $m(A) > 0$ .

*Definition 2.* The belief function  $Bel(A)$  and plausibility function  $Pl(A)$  from a BBA are defined as

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{3}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{4}$$

where  $Bel(A)$  represents the amount of belief that definitely support  $A$ , and the  $Pl(A)$  could be viewed as the amount of belief that potentially placed in  $A$ .

*Definition 3.* Let  $m_1$  and  $m_2$  be two BBAs defined on the same frame  $\Theta$ . D-S evidence theory combination rule is expressed as

$$m(A) = \begin{cases} \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - k} & A \neq \emptyset \\ 0 & A = \emptyset \end{cases} \tag{5}$$

with

$$k = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \tag{6}$$

where  $k$  is named as conflict coefficient to measure the degree of conflict between two BBAs. The combination rule is out of work when  $k = 1$ .

Zadeh [17] presents a famous example that the D-S combination rule will produce an unexpected result. Suppose a frame is  $\Theta = \{A,B,C\}$  and two BBAs are given as

$$m_1 : m_1(A) = 0.99, m_1(B) = 0.01$$

$$m_2 : m_2(B) = 0.01, m_2(C) = 0.99$$

by the D-S combination rule, the aggregated result is  $k = 0.9999$ ,  $m(A) = m(C) = 0$  and  $m(B) = 1$ , which is obviously counter-intuitive and unreasonable.

## 2.2 Jousselme distance

Jousselme distance [32], considering both the mass and cardinality of focal elements of each BBA, is commonly used as the measure of dissimilarity.

*Definition 4.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$ , containing  $N$  mutually exclusive and exhaustive propositions. The Jousselme distance between  $m_1$  and  $m_2$  are defined as

$$d_{m_2}^{m_1} = \sqrt{0.5 * (\|m_1\|^2 + \|m_2\|^2 - 2\langle m_1, m_2 \rangle)} \tag{7}$$

where  $\|m_1\|^2 = \langle m_1, m_1 \rangle$ ,  $\|m_2\|^2 = \langle m_2, m_2 \rangle$  and  $\langle m_1, m_2 \rangle$  is given by

$$\langle m_1, m_2 \rangle = \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m_1(A_i) m_2(B_j) \frac{|A_i \cap B_j|}{|A_i \cup B_j|} \tag{8}$$

with  $A_i$  and  $B_j$  are the elements of the power set  $2^\Theta$ .  $|A_i \cap B_j|$  and  $|A_i \cup B_j|$  denote the cardinality intersection set and union set of  $A_i$  and  $B_j$ .

## 2.3 Probabilistic-based distance

Since the probabilistic transformation has an ability to convert a BBA from the focal elements into a probability measure of distinct atomic, it provides a probabilistic-based distance to measure the dissimilarity of two evidences [41].

*Definition 5.* Let  $m$  be a BBA on a frame  $\Theta$ , and the probabilistic expression of a singleton element  $B$  in  $\Theta$  could be obtained by pignistic probability function

$$BetP_m(B) = \sum_{A \in 2^\Theta, B \subseteq A} \frac{m(A)}{|A|} \tag{9}$$

where  $|A|$  is the cardinality of proposition  $A$ . If  $|A| = 1$ , then  $B = A$  and  $BetP(B) = BetP(A) = m(A)$ .

*Definition 6.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$  and let  $BetP_{m_1}$  and  $BetP_{m_2}$  be the results of pignistic probability transformation of  $m_1$  and  $m_2$ , the probabilistic-based distance  $difBetP_{m_1}^{m_2}$  is defined as

$$difBetP_{m_1}^{m_2} = \max_{A \in \Theta} (|BetP_{m_1}(A) - BetP_{m_2}(A)|) \tag{10}$$

and the Murkowski distance [36] proposed by Liu is defined as

$$distP_{m_1}^{m_2} = \sum_{A_i \in \Theta} 0.5 * (|BetP_{m_1}(A_i) - BetP_{m_2}(A_i)|) \tag{11}$$

## 2.4 Combinatorial dissimilarity measure

Some compound dissimilarity measures are presented based on the conflict coefficient, evidence distance and probabilistic-based distance.

*Definition 7.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$ , and a combinatorial dissimilarity measure based on the conflict coefficient  $k_{m_1}^{m_2}$  and probabilistic-based distance  $difBetP_{m_1}^{m_2}$  is defined as

$$cf_{m_2}^{m_1} = \langle k_{m_1}^{m_2}, difBetP_{m_1}^{m_2} \rangle \tag{12}$$

$m_1$  and  $m_2$  are in conflict, iff both  $k_{m_1}^{m_2} > \epsilon$  and  $difBetP_{m_1}^{m_2} > \epsilon$ .  $\epsilon \in [0,1]$  denotes the threshold of conflict tolerance, and identified according to different applications [35].

*Definition 8.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$ , and a combinatorial dissimilarity measure [42] based on the conflict coefficient  $k_{m_1}^{m_2}$  and Jousselme distance  $d_{J_{m_1}}^{m_2}$  is defined as

$$k_d = \frac{1}{2} \times (k_{m_1}^{m_2} + d_{J_{m_1}}^{m_2}) \tag{13}$$

*Definition 9.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$  and let  $BetP_{m_1}$  and  $BetP_{m_2}$  be the results of pignistic probability transformation of  $m_1$  and  $m_2$ . Then a combinatorial dissimilarity measure based on the Hamacher T-conorm fusion rules [43] is defined as

$$DismP_{m_1}^{m_2} \triangleq T(distP_{m_1}^{m_2}, ConfP_{m_1}^{m_2}) = \frac{distP_{m_1}^{m_2} + ConfP_{m_1}^{m_2}}{1 + distP_{m_1}^{m_2} \cdot ConfP_{m_1}^{m_2}} \tag{14}$$

where  $ConfP_{m_1}^{m_2}$  denotes the conflict coefficient based on the pignistic probability,

$$ConfP_{m_1}^{m_2} = \begin{cases} 0, & \text{if } BetP(X_{m_1}^{max}) \cap BetP(X_{m_2}^{max}) \neq \emptyset \\ BetP(X_{m_1}^{max}) \cdot BetP(X_{m_2}^{max}), & \text{otherwise} \end{cases} \tag{15}$$

where  $BetP(X_{m_i}^{max}) = arg \max_{x \in \Theta} BetP_{m_i}(x), i = 1, 2$

### 2.5 Deng entropy

Deng entropy, as a generalization of Shannon entropy, provide a solution to measure the information volume of a BBA. It is observed that the Deng entropy and Shannon entropy correspond to an uncertain degree of measurement [38].

*Definition 10.* Let  $m$  be a BBA on the frame  $\Theta$  and the Deng entropy of  $m$  is defined as

$$E_d = -\sum_i m(A_i) \log \frac{m(A_i)}{2^{|A_i|} - 1} \tag{16}$$

where  $A_i$  is a proposition in BBA  $m$ , and  $|A_i|$  is the cardinality of  $A_i$ . The Deng entropy will become identical to Shannon entropy if  $|A_i| = 1$ , that is

$$E_d = -\sum_i m(A_i) \log \frac{m(A_i)}{2^{|A_i|} - 1} = -\sum_i m(A_i) \log m(A_i) \tag{17}$$

### Limitations of exiting dissimilarity measurements between BBAs

*Example 1.* Let  $m_1, m_2$  and  $m_3$  be three BBAs on the same frame  $\Theta$  with four propositions  $\Theta = \{A_1, A_2, A_3, A_4\}$ . The three BBAs are given as

$$m_1 : m_1(A_1) = m_1(A_2) = m_1(A_3) = m_1(A_4) = 0.25$$

$$m_2 : m_2(A_1) = m_2(A_2) = m_2(A_3) = m_2(A_4) = 0.25$$

$$m_3 : m_3(A_1) = m_3(A_2) = m_3(A_3) = 1/3$$

we can get the conflict coefficients  $k_{m_1}^{m_2} = 0.75$  and  $k_{m_1}^{m_3} = 0.67$  by using Eq (6) between the BBAs. The result shows that the degree of conflict between  $m_1$  and  $m_2$  is bigger than the degree of conflict between  $m_1$  and  $m_3$ , and they both are in relative high conflict. In fact, there is no conflict intuitively between  $m_1$  and  $m_2$  because they are the same. By using the Eq (10) and Eq

(12), we can get the  $difBetP_{m_1}^{m_2} = 0$  and  $cf(m_1, m_2) = \langle 0.75, 0 \rangle$ , which illustrates that  $m_1$  and  $m_2$  are consistent and measurement of coefficient  $k$  cannot measure the degree of conflict between the evidences in this situation. Although the combined measurement implies that the D-S combination rule should be used, it cannot conclude that how much the error will be conduct by using the combination rule. Therefore, the combination rule has a limitation in terms of providing an explicit expression and cannot be used directly in the combination rule.

Example 2. Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta = \{A_1, A_2, \dots, A_{2n}\}$ , such that

$$m_1 : m_1(A_1) = m_1(A_2) = \dots = m_1(A_n) = \frac{1}{n}$$

$$m_2 : m_2(A_{n+1}) = m_2(A_{n+2}) = \dots = m_2(A_{2n}) = \frac{1}{n}$$

It is obvious that the  $m_1$  and  $m_2$  are totally contrary to each other as they support the different propositions. The different dissimilarities between  $m_1$  and  $m_2$  are displayed in Fig 1.

From Fig 1, it is evident that the values of the  $d_j$ ,  $k_d$  and  $difBetP$  are 1, when  $n = 1$ , which are intuitive. But when  $n > 1$ , the values of  $d_j$  and  $difBetP$  tend to 0, and  $k_d$  tends to 0.6, indicating that  $m_1$  and  $m_2$  are getting closer and less conflict with the increase of  $n$ , which are counter-intuitive and abnormal. Only the  $DismP$  keeps 1 with the increase with  $n$ , meaning that the  $m_1$  and  $m_2$  are totally in disagreement with each other. Therefore,  $d_j$ ,  $k_d$  and  $difBetP$  cannot be used as measurement of the dissimilarity between BBAs in this example.

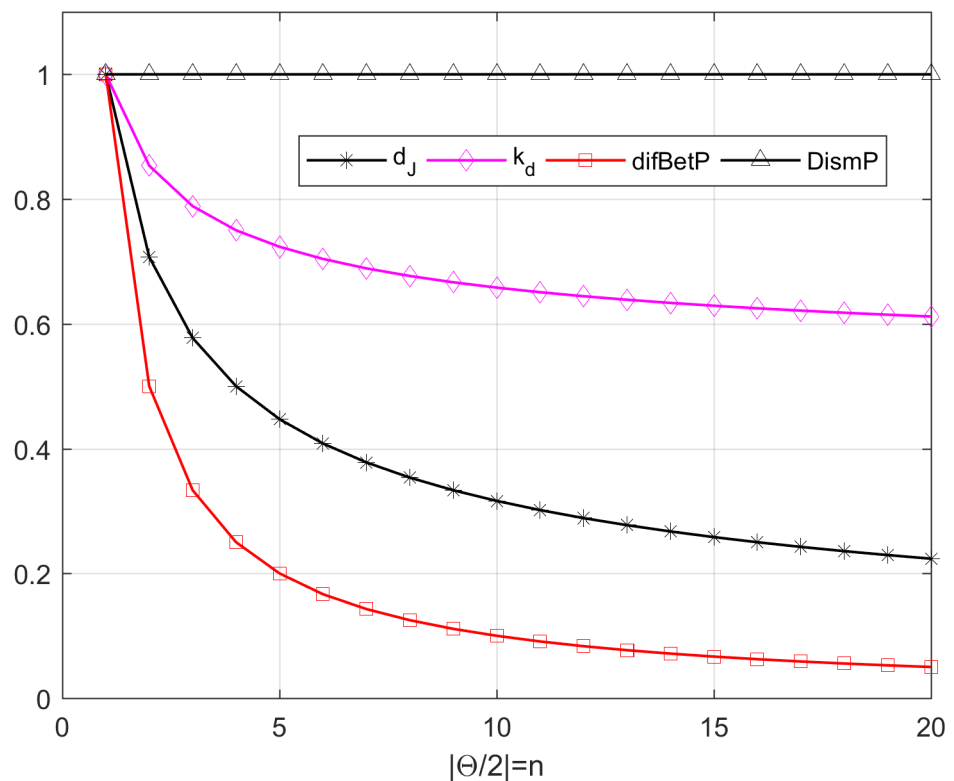


Fig 1. Different dissimilarity measurements.

<https://doi.org/10.1371/journal.pone.0177695.g001>

Since *DismP* considers not only the distance but also the conflict between BBAs, the measurement based on the Hamacher T-conorm fusion rules provides a general method of the dissimilarity. However, it has deficiency as shown in *Example 3*.

*Example 3.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame of discernment  $\Theta = \{A_1, A_2, \dots, A_{20}\}$ . For notation conciseness 1, 2, and so forth have been used to denote  $A_1, A_2$ , and so forth in the frame. The two pairs of BBAs are shown as

1st Pair:

$$m_1 : m_1(2, 3, 4) = 0.05, m_1(7) = 0.05, m_1(\Theta) = 0.1, m_1(\Delta) = 0.8$$

$$m_2 : m_2(1, 2, 3, 4, 5) = 1$$

2nd Pair:

$$m_1 : m_1(2, 3, 4) = 0.05, m_1(7) = 0.05, m_1(\Theta) = 0.1, m_1(\Delta) = 0.8$$

$$m_2 : m_2(1, 2, 3, 4, 5) = 0.5, m_2(6, 7, 8, 9, 10) = 0.5$$

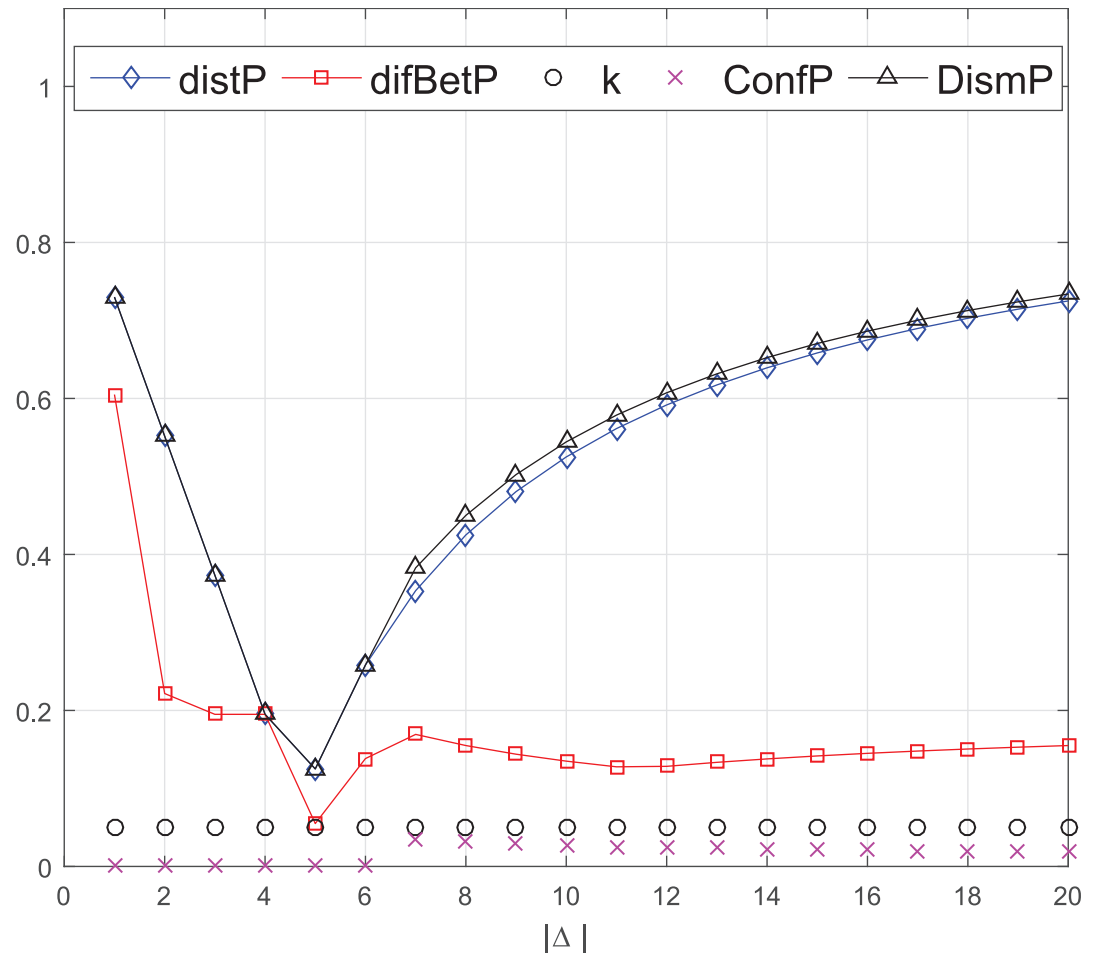
where the  $\Delta$  is a subset of  $\Theta$ . This example considers 20 cases of the subset  $\Delta$ , which increases by adding a new element at each case from  $\Delta = \{1\}$  to  $\Delta = \{1, 2, \dots, 20\}$ . The comparison of the dissimilarity measurements between  $m_1$  and  $m_2$  of the two pairs are shown in Figs 2 and 3 respectively.

From Figs 2 and 3, it can be seen that the *distP* and *DismP* are very close and follow the same trend, since the *DismP* are mainly decided by the *distP* and *ConfP*. As calculated in Eq (15), the maximal pignistic probabilities in both  $m_1$  and  $m_2$  always have intersection. So *ConfP* is small when the  $\Delta$  increases from  $\{1\}$  to  $\{1, 2, \dots, 20\}$ . However, it is quite distinct situation of the two pairs of BBAs. In the 1st pair, both BBAs distribute their major belief to the same elements when the cases from 1 to 6, which cause that the *ConfP* keeps 0. This is reasonable as the  $m_2$  only has one focal element  $m_2(1, 2, 3, 4, 5) = 1$  which corresponds to classical conflict coefficient  $k$ . In the 2nd pair,  $m_2$  has two equal focal elements  $m_2(1, 2, 3, 4, 5) = 0.5$  and  $m_2(6, 7, 8, 9, 10) = 0.5$ . As the case from 1 to 5, there should be a notable dissimilarity between  $m_1$  and  $m_2$ , which is shown as  $k$  in Fig 3. Nevertheless, the pignistic probability transformation divides the belief equally to each single proposition as  $BetP_{m_2}(1) = BetP_{m_2}(2) = \dots = BetP_{m_2}(10) = 0.1$ , indicating that the *ConfP* considers the dissimilarity as 0. Therefore, the dissimilarity measures of *ConfP* and *DismP* are illogical in this situation. Although the classical conflict coefficient  $k$  could depicts the dissimilarity from cases 1 to 5, it cannot reflect the variety of divergence degree as the case increases. Neither does the *difBetP*.

## Combining belief function with a new weighting factor

### 4.1 A modified dissimilarity measure

In this section, a modified dissimilarity measure is proposed which is based on the Hamacher T-conorm fusion rules to describe the dissimilarity between BBAs. The dissimilarity measurement based on Hamacher T-conorm rules satisfy two important properties of commutativity and monotonicity. The commutativity could ensure that the dissimilarity matrix is symmetrical and no matter the fusion order of two evidences is, their dissimilarity is coincident. The monotonicity provides that dissimilarity measurement has single variation trend in a specific interval, which is easy to compare the dissimilarity between evidences.



**Fig 2. Comparison of dissimilarity measures of the 1st pair evidence.**

<https://doi.org/10.1371/journal.pone.0177695.g002>

*Definition 11.* Let  $m_1$  and  $m_2$  be two BBAs on the same frame  $\Theta$ . The modified dissimilarity measure is defined as

$$MDismP_{m_1}^{m_2} \triangleq T\left(distP_{m_1}^{m_2}, k_{m_1}^{m_2}\right) = \frac{distP_{m_1}^{m_2} + k_{m_1}^{m_2}}{1 + distP_{m_1}^{m_2} \cdot k_{m_1}^{m_2}} \quad (18)$$

where  $k_{m_1}^{m_2}$  is the classical conflict coefficient.

$$k_{m_1}^{m_2} = \sum_{A_i \cap A_j = \emptyset} m_1(A_i) m_2(A_j) \quad (19)$$

The modified dissimilarity still satisfies the basic properties of commutativity and monotonicity:

(1) Commutativity:

$$MDismP(x, y) = MDismP(y, x) \quad (20)$$

(2) Monotonicity:

$$0 \leq MDismP(x, y) \leq MDismP(x', y) \leq MDismP(x', y') \leq 1 \quad (21)$$

where  $0 \leq x \leq x' \leq 1$  and  $0 \leq y \leq y' \leq 1$ .



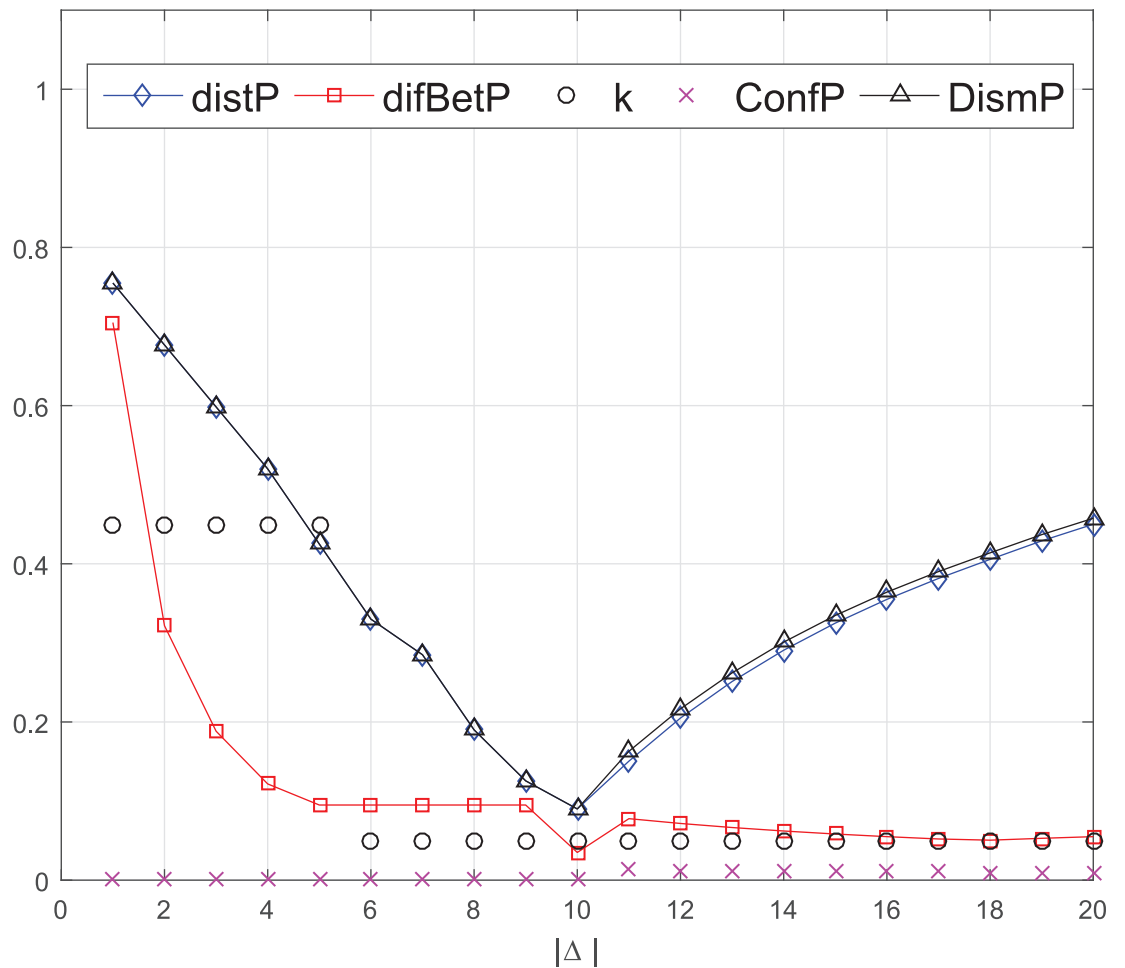


Fig 3. Comparison of dissimilarity measures of the 2nd pair evidence.

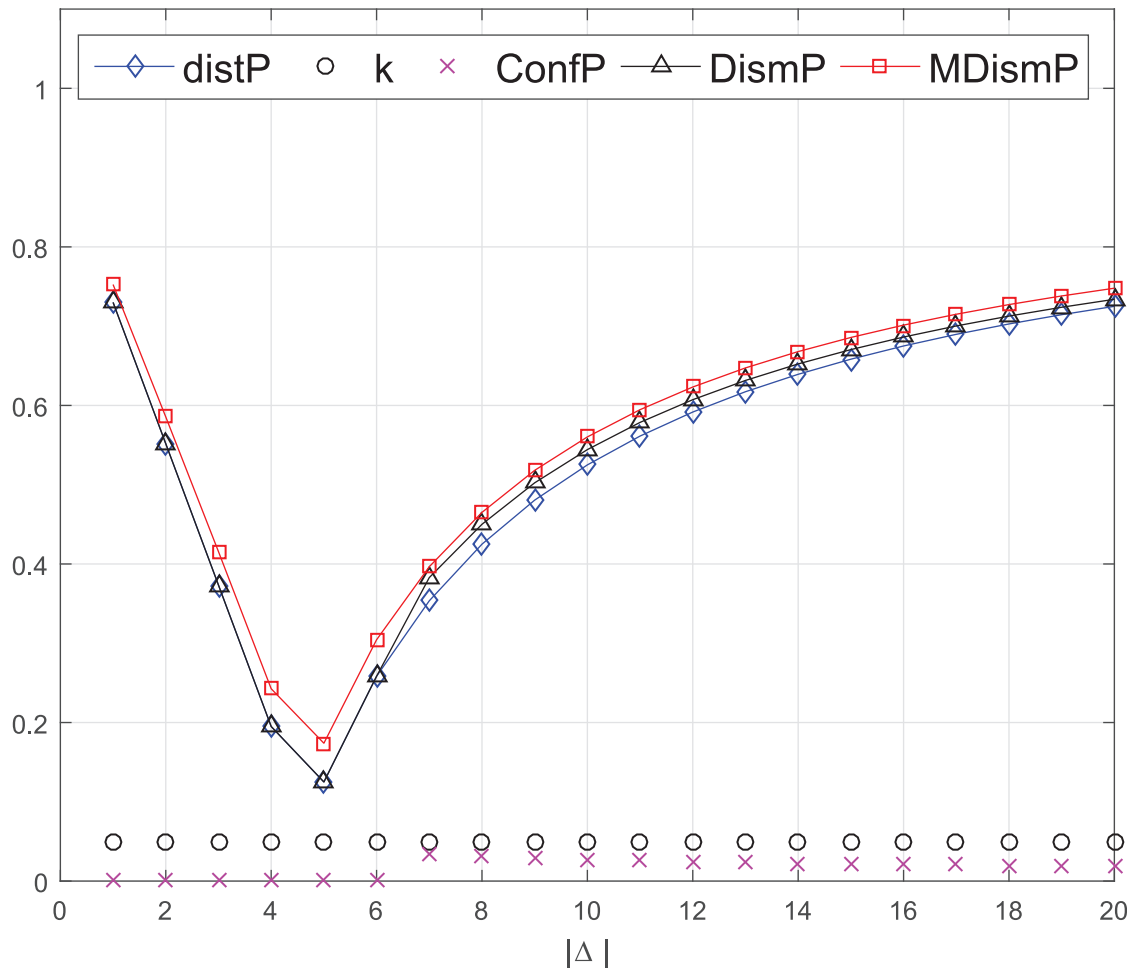
<https://doi.org/10.1371/journal.pone.0177695.g003>

The modified dissimilarity measurement consists of a distance coefficient and a conflict coefficient between two evidences. As both the distance coefficient and the conflict coefficient lie in  $[0,1]$ , the modified measurement is larger than its either components. The evidences have large distance and high conflict with the majority of other evidences would have a larger dissimilarity measurement and vice versa.

It is obvious that the modified dissimilarity measure replaces the conflict coefficient  $ConfP_{m_1}^{m_2}$  with  $k_{m_1}^{m_2}$ . The  $ConfP_{m_1}^{m_2}$  implies that the main conflict results from discordant propositions which are strongly supported by two BBAs respectively. However, the  $ConfP_{m_1}^{m_2}$  cannot handle the situation that one BBA has several propositions with equal belief. As we see it, the conflict coefficient should involve all conflicts existed between BBAs no matter how small the extent of conflicts is. Furthermore, the  $MDismP$  not only maintains good features but also makes up for shortcomings of  $DismP$ .

*Example 4.* Considering two pairs of BBAs from *Example 3* with the proposed dissimilarity measure of  $MDismP$ , the results are plotted in Figs 4 and 5.

For the results of the 1st pair illustrated in Fig 4, the  $distP$ ,  $DismP$  and  $MDismP$  are identical. The lines show a variation tendency from a high dissimilarity when  $\Delta = \{1\}$  to the minimum dissimilarity when  $\Delta = \{1,2,3,4,5\}$  and increase again as  $\Delta$  includes more elements. This is because the  $m_2$  only has one proposition and when  $\Delta = \{1,2,3,4,5\}$ , the propositions with the



**Fig 4. Comparison of dissimilarity of the 1st pair evidence.**

<https://doi.org/10.1371/journal.pone.0177695.g004>

maximum belief of two BBAs are accordant. In Fig 5, when  $m_2$  has two propositions with equal belief value, the results are different. The *MDismP* has a bigger value than the *distP* and *DismP* when  $\Delta$  from  $\{1\}$  to  $\{1,2,3,4,5\}$ . The dissimilarity value of *MDismP* is near 0.8 before cases 6, which means the two BBAs are incompatible with each other. But the dissimilarity values of *distP* and *DismP* are less than 0.6, which seems unreasonable. Based on the analysis of the above examples, a conclusion can be drawn that the modified dissimilarity measure *MDismP* can efficiently reflects the degree of dissimilarity between BBAs.

### 4.2 Weighting factors

In this section, we propose a novel method to determine the weighting factors among BBAs based on the modified dissimilarity measure and Deng entropy.

The weight determinations are based on the principle that if an evidence is supported by greater number of evidences, this piece of evidence should be more important and have large effect on the final combination results. Moreover, if an evidence has considerable information, it also should be weighted more [37].

Suppose  $N$  evidences  $\{m_1, m_2, \dots, m_N\}$  are in the same frame of discernment  $\Theta$  and the weight of each evidence is made up of the degree of similarity and information volume. The

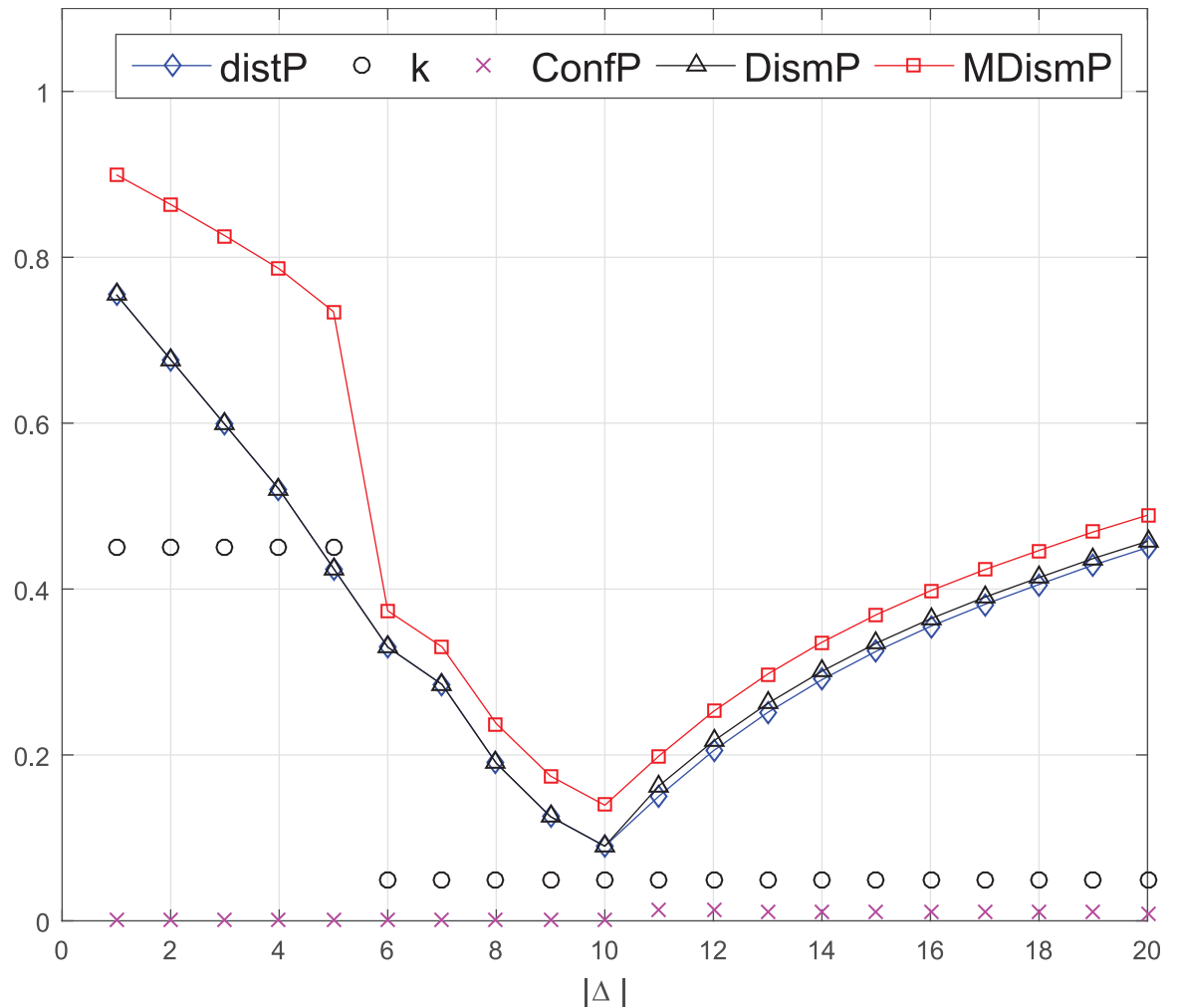


Fig 5. Comparison of dissimilarity of the 2nd pair evidence.

<https://doi.org/10.1371/journal.pone.0177695.g005>

similarity degree  $Sim_{i,j}$  of  $m_i$  and  $m_j$  is defined as

$$Sim_{i,j} = 1 - MDismP(m_i, m_j) \tag{22}$$

the mutual similarity degree matrix  $S_{N \times N}$  is then defined as

$$S_{N \times N} = \begin{bmatrix} 1 & Sim_{1,2} & \cdots & Sim_{1,N} \\ Sim_{2,1} & 1 & \cdots & Sim_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ Sim_{N,1} & Sim_{N,2} & \cdots & 1 \end{bmatrix} \tag{23}$$

The  $S_{N \times N}$  is symmetrical and  $Sim_{ij} = Sim_{ji}$  means that the similarity between two evidences fulfills the commutativity. The diagonal element is 1 means that an evidence is totally similar with itself. The similarity degree matrix helps give an insight into the agreement between evidences, and the weighting factor of dissimilarity could be obtained based on the similarity degree matrix.

The weighting vector  $W^{dis}$  of each evidence is associated with the eigen vector of the maximal positive eigen value  $\lambda_{max}$ , that is  $\lambda_{max} \cdot W^{dis} = S_{N \times N} \cdot W^{dis}$ . The evidence with the largest weight is deemed to be the most important evidence and the weight of each evidence is revised as

$$\omega_i^{dis} = \frac{W_i^{dis}}{\max(W_i^{dis})} \quad i = (1, 2, \dots, N) \tag{24}$$

the information volume of each evidence is measured by Deng entropy and can be calculated by Eq (16). After Deng entropy of each evidence is processed, the weights of information volume of evidence are obtained by

$$\omega_i^{Deng} = \frac{E_d(m_i)}{\max E_d(m_i)} \quad i = (1, 2, \dots, N) \tag{25}$$

then the weight of each evidence based on the proposed method is defined as

$$\omega_i = \frac{\omega_i^{dis} + \omega_i^{Deng}}{\sum_{i=1}^N (\omega_i^{dis} + \omega_i^{Deng})} \quad i = (1, 2, \dots, N) \tag{26}$$

The final weighting factor is measured by the weighting factor of dissimilarity measurement and the weighting factor of evidence information. The two weighting factors describe the final weight from two aspects: the factor of dissimilarity measurement depicts the mutual degree of deviation of an evidence with other evidences; the factor of evidence information illustrates the relative amount of information of an evidence compared to others. The final weighting factor is more thorough than the other weighting (discounting) factors which just depend on the dissimilarity measurement. With the weighting factor of information volume, the new weighting factor still hold the capability to avoid fusion error caused by the single information source failure.

### 4.3 Combination of evidences

The weighted average [29] and discounted [36] methods are two kinds of approaches that have been proposed for combining the evidences. The discounted method distributes the remaining mass value of the discounted mass to the universal set  $\Theta$ , but this would result in more uncertainty. As the discounted methods need to combine the evidences one by one, the calculation will be time consuming and have a low convergence when a large amount of evidences need to be combined. The weighted average method would reinforce each other if the evidences are concordant and would weaken each other if the evidences are in conflict. The belief of propositions after combination will remain distinct from each other. In addition, the weighted average method is easily computational and more reliable and rational. Hence, the weighted average method is applied in this article.

The evidence generated by weighted average method is

$$EWA(m) = \sum_{i=1}^N \omega_i m_i, \quad i = (1, 2, \dots, N) \tag{27}$$

For  $N$  original evidences, we should combine the new evidence  $EWA(m)$  for  $N - 1$  times. In this section, an example of target recognition [36] is presented to show some behaviors of the existing method as well as the proposed method.

*Example 5.* Let  $m_1, m_2, m_3, m_4$  and  $m_5$  be five BBAs on the same frame of discernment  $\Theta = \{A_1, A_2, A_3\}$  as shown in Table 1. The combination results obtained with the proposed methods are shown in Table 2. The convergence is shown in Table 3.  $m_i^i$  means fusion of evidences  $m_1, m_2, \dots, m_i$ , that is  $m_i^i = m_1 \oplus m_2 \oplus \dots \oplus m_i$ .

As can be seen from Table 2, the D-S combination rule (without the discounting or weighted average process) concludes that the proposition  $A_2$  is almost be regarded as the

**Table 1. BBAs of five evidences.**

	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$
$A_1$	0.8	0.4	0	0.3	0.45
$A_2$	0.1	0.2	0.95	0.2	0.1
$A_3$	0	0.1	0.05	0.25	0
$\{A_1, A_2\}$	0	0.3	0	0.2	0
$\{A_2, A_3\}$	0	0	0	0	0.15
$\emptyset$	0.1	0	0	0.05	0.3

<https://doi.org/10.1371/journal.pone.0177695.t001>

target. The result is unreasonable and counter-intuitive since majority of evidences distribute the major belief to proposition  $A_1$  and just one evidence  $m_3$  give its major belief to  $A_2$ . Such unexpected behavior is solved by using the discounting and weighted average method to lessen the influence of the evidences which is dissimilar with the other evidences. The larger the dissimilarity of an evidence, the larger discount it will have. However, one also sees that the process of the proposed method as the dissimilarity measure to determine the weight of each evidence generates a more specific and reliable result than the process of discounting factors based on  $d_j$  and *DismP*. In addition, the  $m_1^3$  is larger than the  $m_1^2$  of  $m(A_1)$  of the proposed method, which is completely opposite of  $d_j$  and *DismP* when the  $m_3$  assign its major belief to  $A_2$ . What caused this is that the proposed new weighting factor weakens  $m_3$  twice by the weighting factor of dissimilarity and Deng entropy. For  $d_j$  and *DismP*, the discount factor weakens  $m_3$  once, which overcomes the defect raise by the classical combination rule. Besides, the evidence  $m_3$  has less information when compared to other evidences. So, weighting factor of Deng entropy further weaken the evidence  $m_3$  on the basis of the weighting factor of the dissimilarity factor. The larger  $m_1^3$  of  $m(A_1)$  of the proposed method reflects the proposed new weighting factor has a great capability of anti-interference. Furthermore, the proposed method has a better performance of convergence than the other methods. As can be seen from [Table 3](#), the series results of proposed method achieve the belief level of other methods before the

**Table 2. Comparison results of different methods.**

	$m_1^2$	$m_1^3$	$m_1^4$	$m_1^5$
<b>D – S [36]</b>	$m(A_1) = 0.8451$ $m(A_2) = 0.0986$ $m(A_3) = 0.0140$ $m(A_{1,2}) = 0.0423$	$m(A_2) = 0.9948$ $m(A_3) = 0.0052$	$m(A_2) = 0.9965$ $m(A_3) = 0.0035$	$m(A_2) = 0.9971$ $m(A_3) = 0.0029$
<b><math>d_j</math> [36]</b>	$m(A_1) = 0.7659$ $m(A_2) = 0.1166$ $m(A_3) = 0.0294$ $m(A_{1,2}) = 0.0881$	$m(A_1) = 0.6239$ $m(A_2) = 0.2791$ $m(A_3) = 0.0252$ $m(A_{1,2}) = 0.0718$	$m(A_1) = 0.6858$ $m(A_2) = 0.2645$ $m(A_3) = 0.0146$ $m(A_{1,2}) = 0.0315$	$m(A_1) = 0.7528$ $m(A_2) = 0.2217$ $m(A_3) = 0.0096$ $m(A_{1,2}) = 0.0159$
<b><i>DismP</i> [36]</b>	$m(A_1) = 0.7503$ $m(A_2) = 0.1196$ $m(A_3) = 0.0319$ $m(A_{1,2}) = 0.0957$ $m(\emptyset) = 0.0025$	$m(A_1) = 0.7157$ $m(A_2) = 0.1598$ $m(A_3) = 0.0308$ $m(A_{1,2}) = 0.0913$ $m(\emptyset) = 0.0024$	$m(A_1) = 0.7670$ $m(A_2) = 0.1655$ $m(A_3) = 0.0194$ $m(A_{1,2}) = 0.0477$ $m(\emptyset) = 0.0004$	$m(A_1) = 0.8254$ $m(A_2) = 0.1424$ $m(A_3) = 0.0120$ $m(A_{1,2}) = 0.0198$ $m(A_{2,3}) = 0.0002$ $m(\emptyset) = 0.0002$
<b><i>Proposed Method</i></b>	$m(A_1) = 0.7974$ $m(A_2) = 0.1276$ $m(A_3) = 0.0114$ $m(A_{1,2}) = 0.0610$ $m(\emptyset) = 0.0026$	$m(A_1) = 0.8227$ $m(A_2) = 0.1584$ $m(A_3) = 0.0019$ $m(A_{1,2}) = 0.0169$ $m(\emptyset) = 0.0001$	$m(A_1) = 0.8277$ $m(A_2) = 0.1588$ $m(A_3) = 0.0036$ $m(A_{1,2}) = 0.0099$	$m(A_1) = 0.8651$ $m(A_2) = 0.1249$ $m(A_3) = 0.0048$ $m(A_{1,2}) = 0.0046$ $m(A_{2,3}) = 0.0004$ $m(\emptyset) = 0.0001$

<https://doi.org/10.1371/journal.pone.0177695.t002>

Table 3. The convergence of the proposed method.

Fusion time $n$	$m_1^2$	$m_1^3$	$m_1^4$	$m_1^5$
$n = 1$	$m(A_1) = 0.7974$ $m(A_2) = 0.1276$ $m(A_3) = 0.0114$ $m(A_{1,2}) = 0.0610$ $m(\Theta) = 0.0026$	$m(A_1) = 0.7160$ $m(A_2) = 0.2149$ $m(A_3) = 0.0115$ $m(A_{1,2}) = 0.0553$ $m(\Theta) = 0.0023$	$m(A_1) = 0.6356$ $m(A_2) = 0.2424$ $m(A_3) = 0.0442$ $m(A_{1,2}) = 0.0748$ $m(\Theta) = 0.0029$	$m(A_1) = 0.6158$ $m(A_2) = 0.2247$ $m(A_3) = 0.0555$ $m(A_{1,2}) = 0.0676$ $m(A_{2,3}) = 0.0169$ $m(\Theta) = 0.0195$
$n = 2$		$m(A_1) = 0.8227$ $m(A_2) = 0.1584$ $m(A_3) = 0.0019$ $m(A_{1,2}) = 0.0169$ $m(\Theta) = 0.0001$	<b><math>m(A_1) = 0.7523</math></b> $m(A_2) = 0.2066$ $m(A_3) = 0.0131$ $m(A_{1,2}) = 0.0279$ $m(\Theta) = 0.0002$	$m(A_1) = 0.7292$ $m(A_2) = 0.2051$ $m(A_3) = 0.0270$ $m(A_{1,2}) = 0.0297$ $m(A_{2,3}) = 0.0055$ $m(\Theta) = 0.0035$
$n = 3$			$m(A_1) = 0.8277$ $m(A_2) = 0.1588$ $m(A_3) = 0.0036$ $m(A_{1,2}) = 0.0099$	<b><math>m(A_1) = 0.8086</math></b> $m(A_2) = 0.1654$ $m(A_3) = 0.0118$ $m(A_{1,2}) = 0.0120$ $m(A_{2,3}) = 0.0016$ $m(\Theta) = 0.0006$
$n = 4$				<b><math>m(A_1) = 0.8651</math></b> $m(A_2) = 0.1249$ $m(A_3) = 0.0048$ $m(A_{1,2}) = 0.0046$ $m(A_{2,3}) = 0.0004$ $m(\Theta) = 0.0001$

<https://doi.org/10.1371/journal.pone.0177695.t003>

completion of the fusion process. For example, the three times fusion results of  $m_1^5$  by the proposed method achieves 0.8 which achieves belief level to final result of *DismP* 0.8254, indicating that the proposed method provides faster convergence and is less computationally intensive. In the end, the great difference in  $m_3$ , where almost all belief is given to  $A_2$ , is caused by that evidence  $m_3$  may be especially sensitive to the typical characteristics of the kind of  $A_2$ . The results in Table 2 show that the  $A_2$  with the second largest belief that is much larger than other propositions except  $A_1$ . Therefore,  $A_2$  could be seen as a potential identification results as well.

This work provides a modified dissimilarity measure and evidence information volume measure to determine the weighting factors of evidences involved in the fusion process. The modified dissimilarity measure includes both the distances and all conflicts between evidences and the evidence information volume of each evidence is measured by Deng entropy. Compared with other methods, the proposed method gives a more specific and faster aggregated result. In addition, it is an efficient method in some decision-making applications.

### Identification of major land use activities contributing to river pollution: An application

In this section, the proposed method is applied to a case study in identifying the major land use activities contributing to river pollution and the required data are collected by water quality monitors.

The Manahara River lies in the Kathmandu Valley, Nepal, with a watershed of 256 km<sup>2</sup>. The river originates from Manichud Lekh (ridge) at an elevation of 2352 m and has a length of 30 km. The river originates from a pristine forested region and flows through forest, rural, semi-urban, and urban areas. The water quality was monitored monthly in seven different sites (Sites 1–7) of the river by sensors. The data of Sites 1 to 7, respectively Salinadi, Sankhu, Brahmakhel, Bode, Sinamangal, Imadol, and Chyasar are obtained from [44]. The land use and anthropogenic activities in transition of different sites are given in Table 4.

**Table 4. Land use and river activities in site transition.**

Site	Land use and river activities
1 to 2	A1 <sup>a</sup> , A2
2 to 3	A2, A3
3 to 4	A2, A3, A4
4 to 5	A2, A3, A6
5 to 6	A2, A3, A5, A6
6 to 7	A6

<sup>a</sup>Note: A1: Forest, A2: Agriculture, A3: Bathing, washing & cleansing, A4: Rural settlement (sparse), A5: Industries, A6: Urban settlement (dense).

<https://doi.org/10.1371/journal.pone.0177695.t004>

The water quality of the river during the low flow months in 2006 is given in Table 5. Water qualities are good in upstream and bad in downstream regions and varied gradually from Site 1 to 7. Each monitor records the amount of each chemical element in river when the water

**Table 5. Water quality of Manahara river in different sites.**

Month	Site	DO <sup>a</sup> , mg/L	BOD mg/L	Free CO2, mg/L	TA mg/L	Cl, mg/L	NO3-N, mg/L	PO4 -P, mg/L	NH3 -N, mg/L	EC, uS/cm	TDS, mg/L
February	1	9.2	1.2	7.1	26	6.3	0.18	0.1	0.07	60.33	39.9
	2	9.2	1.3	9.2	39	6.7	0.32	0.2	0.2	71	47.1
	3	9.1	6.8	12.5	39	7.2	0.64	0.3	0.26	82	54.4
	4	8.5	14.9	17.8	67	9.0	0.72	0.4	0.61	94	62.4
	5	6.2	27.5	48.3	103	22.2	0.83	0.9	1.16	281.33	186.4
	6	2.2	154.9	106.8	352	57.6	2	4.2	4.42	925	613.0
	7	0	155.0	107.5	379	71.6	3.37	5.3	4.62	957.73	667.0
March	1	9.6	1.2	6.2	25	7.5	0.19	0.1	0.08	58.33	38.6
	2	8.9	1.8	6.9	31	8.4	0.35	0.2	0.22	70	46.4
	3	7.9	8.7	13.4	40	9.0	0.7	0.3	0.28	81.67	54.1
	4	7.3	16.7	14.1	73	13.2	0.75	0.4	0.66	116.33	76.7
	5	5.8	36.9	32.6	151	38.7	0.88	1.1	1.21	405.33	267.5
	6	5.7	83.8	55.9	276	68.7	1.82	3.3	3.25	719.33	474.7
	7	2.5	85.3	59.5	277	78.5	3.26	3.4	8.84	796	474.7
April	1	8.6	1.8	8.5	32	6.2	0.18	0.1	0.07	49.67	32.8
	2	7.2	2.8	8.7	33	7.4	0.25	0.2	0.22	75.33	49.9
	3	6.8	9.8	9.2	37	8.6	0.6	0.2	0.24	78	51.8
	4	6.5	17.8	9.3	41	11.8	0.63	0.4	0.39	91	60.3
	5	4.3	39.3	17.3	63	26.5	0.79	1.3	1.32	192.67	127.3
	6	3.3	113.6	37.1	167	43.2	1.79	2.9	2.86	473	278.3
	7	2.0	149.2	77.0	279	48.4	3.09	3.2	3.22	544.67	359.0
May	1	7.2	2.16	8.6	32	9.3	0.2	0.1	0.07	58.33	38.4
	2	6.7	3.8	13.4	33	9.7	0.26	0.1	0.22	97.33	64.1
	3	6.2	10.5	13.9	41	10.9	0.62	0.3	0.25	114.33	75.3
	4	6.0	18.4	13.9	41	13.3	0.67	0.4	0.55	122.33	80.6
	5	2.8	56.2	13.9	62	17.4	0.75	1.2	1.15	181.33	119.5
	6	2.3	119.0	24.9	130	25.0	2.24	1.8	3.04	296	195.0
	7	1.8	167.3	25.2	139	29.3	3.83	1.8	3.1	389.33	256.6

<sup>a</sup>Note: DO: Dissolved oxygen, BOD: Biochemical Oxygen Demand, TA: Total Alkalinity, Cl: Chloride, EC: Electrical conductivity, TDS: Total Dissolved Solids.

<https://doi.org/10.1371/journal.pone.0177695.t005>

stream flow past each site. The actual variation amount of chemical element in each site should only consider the difference with the previous site. The chemical elements criteria in Table 5 could be divided into benefit criteria and cost criteria.

The change in water quality from one site to next site is normalized by Eqs (28) and (29). For a benefit criterion,

$$N_i = \frac{X_i - X_{i+1}}{X^{max} - X^{min}} \tag{28}$$

where  $\sum_{i=1}^n N_i = 1$  and  $X^{max}$  is the maximum value and  $X^{min}$  is the minimum value of the criterion. For a cost criterion,

$$N_j = \frac{X_{j+1} - X_j}{X^{max} - X^{min}} \tag{29}$$

where  $\sum_{j=1}^n N_j = 1$  and  $X^{max}$  is the maximum value and  $X^{min}$  is the minimum value of the criterion.

For example, the amount of DO (mg/L) in February is 9.2, 2.2 and 0 record in site 1, 6 and 7 respectively. DO (mg/L) is a benefit criterion and the normalization of site 6–7 is as Eq (30) shows.

$$N_6 = \frac{X_6 - X_7}{X_1 - X_7} = \frac{2.2 - 0}{9.2 - 0} = 0.239 \tag{30}$$

The criteria, except the DO, are all cost criteria. The normalized chemical criteria of February are shown in Table 6.

In Table 6, each chemical criterion is seen as an evidence. Therefore, there are ten evidences need to be fused to identify which site causes the major pollution. According to the proposed method, the results of the river water quality monitoring and identification of February are given in Table 7.

From the Table 7, it shows that the A6 urban settlement (dense) is the major cause for the river pollution. All the fusion results are listed in Table 8.

Table 8 shows that the major source of water pollution in all months are related to A6, i.e. urban settlement (dense). This result is true and reasonable as the adjacent area of Site 6 and 7 is A6 that discharged untreated sewage directly to those sites [45, 46], that have the worst water quality with very high BOD, nutrients (NO3-N and PO4-P), NH3-N, EC, and TDS with very low dissolved oxygen. Untreated sewage has high BOD, nutrients (NO3-N and PO4-P), NH3-N, EC, and TDS with very low or almost zero dissolved oxygen [47]. However, the combination (A2, A3, & A6) in Site 4 to 5 and the combination (A2, A3, A5, & A6) in Site 5 to 6 also have urban settlement (A6), which contributed insignificantly because urban settlement

**Table 6. Normalized chemical criteria in different sites of February.**

Month	Site	DO, mg/L	BOD mg/L	Free CO2, mg/L	TA mg/L	Cl, mg/L	NO3-N, mg/L	PO4 -P, mg/L	NH3 -N, mg/L	EC, uS/cm	TDS, mg/L
February	1–2	0.002	0.000	0.021	0.037	0.006	0.044	0.019	0.029	0.012	0.011
	2–3	0.009	0.036	0.033	0.000	0.008	0.100	0.019	0.013	0.012	0.012
	3–4	0.066	0.053	0.053	0.079	0.027	0.025	0.019	0.077	0.013	0.013
	4–5	0.248	0.082	0.303	0.102	0.201	0.034	0.096	0.121	0.209	0.198
	5–6	0.436	0.828	0.583	0.705	0.543	0.367	0.635	0.716	0.717	0.680
	6–7	0.239	0.000	0.007	0.076	0.215	0.429	0.212	0.044	0.036	0.086

<https://doi.org/10.1371/journal.pone.0177695.t006>



**Table 7. Fusion results of February by using the proposed method.**

Land use and river activities	A2	A6	A1 A2	A2 A3	A2, A3, A4	A2, A3, A6	A2, A3, A5, A6
Fusion Result of February	0.057	<b>0.615</b>	0.000	0.151	0.000	0.134	0.041

<https://doi.org/10.1371/journal.pone.0177695.t007>

covered only a smaller land cover, is far from the river, and sewage outfall is absent in Site 4 and lower in Site 5.

The combination (A2, A3, A5, & A6) also has an industry which contributed insignificantly because there is only one beverage industry and its wastewater is diluted by the river flow and/or the sampling time would have been different than the discharge of concentrated industrial wastewater. In addition, the contribution of organic load by agriculture is much lower than that of sewage outfall [48]. Due to this, agriculture (A2) contribution has resulted in an insignificant pollution. Furthermore, the organic load contribution by forest (A1) is lower [49, 50], due to which the combination (A2 & A3) has insignificant contribution too.

In the four months from February to May, the results are similar. This is because the rivers have low flow in these months with February being winter and March to May being the pre-monsoon season [47]. The low flow in a river results in a high concentration of pollutants due to the lack of dilution. Therefore, the results of the application of the proposed method in identifying water polluting land use activities match the actual situation.

## Conclusion

Solving the problem of conflicting information fusion is still under continuous discussion. In this paper, a combination approach of evidences with a new weighting factor has been proposed based on a modified dissimilarity measure and Deng entropy between BBAs. The new weighting factor contains two part: weight of similarity and weight of entropy. The weight of similarity describes the degree of mutual support between BBAs. The larger the dissimilarity with others, the less the weight of similarity of a BBA. The weight of entropy describes the information volume of each BBA. If a BBA has a larger entropy value, meaning it contains more information, the weight of entropy of that BBA would be larger.

After analyzing the features and limitations of the existed dissimilarity measures, a modified dissimilarity measure based on the Hamacher T-conorm fusion rules mixing the probabilistic-based distances with the all conflict between BBAs has been developed. Also, the weight of similarity is determined by dissimilarity measures. The information volume of a BBA is obtained by Deng entropy and is used in the determination of the weight of entropy.

**Table 8. Results of the application of the proposed method.**

Land use and river activities	Fusion Result			
	Feb	Mar	Apr	May
A2	0.057	0.074	0.171	0.179
A6	<b>0.615</b>	<b>0.770</b>	<b>0.619</b>	<b>0.606</b>
A1 & A2	0.000	0.000	0.000	0.000
A2 & A3	0.151	0.114	0.112	0.142
A2, A3, & A4	0.000	0.000	0.000	0.000
A2, A3, & A6	0.134	0.038	0.087	0.069
A2, A3, A5, & A6	0.041	0.001	0.005	0.002

<https://doi.org/10.1371/journal.pone.0177695.t008>

The new weighting factor based on the modified dissimilarity and Deng entropy is presented and applied in the weighted average combining belief function. Several numerical examples analysis show that the proposed method not only obtains a more reasonable and specific result but also has a faster convergence rate. A real application of determining the major land use activities contributing to river pollution is implemented by the proposed method. Result shows that the urban settlement (dense) is the major source of water pollution. This new approach can be of great interest for decision makers in devising strategies to control water pollution and environmental management.

## Acknowledgments

The work is partially supported by the National Natural Science Foundation of China (Grant No. 61401363), the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grants (Grant No. CRDPJ 446638–12), the Science and Technology on Avionics Integration Laboratory and Aeronautical Science Foundation (Grant No. 20155153034), the Fundamental Research Funds for the Central Universities (Grant No. 3102016AXXX005, 3102015BJIIGZ009).

## Author Contributions

**Conceptualization:** QP.

**Data curation:** GC-S.

**Formal analysis:** QP GC-S.

**Funding acquisition:** KZ KH.

**Investigation:** QP GC-S.

**Methodology:** QP.

**Project administration:** DZ RS.

**Resources:** DZ RS.

**Software:** QP XL.

**Supervision:** RS.

**Validation:** DZ RS.

**Visualization:** QP GC-S.

**Writing – original draft:** QP.

**Writing – review & editing:** GC-S RS.

## References

1. Jiang W, Yang Y, Luo Y, Qin X. Determining basic probability assignment based on the improved similarity measures of generalized fuzzy numbers. *International Journal of Computers Communications & Control*. 2015; 10(3):333–47.
2. Smets P, Kennes R. The transferable belief model. *Artificial intelligence*. 1994; 66(2):191–234.
3. Rao SS, Annamdas K. A comparative study of evidence theories in the modeling, analysis, and design of engineering systems. *Journal of Mechanical Design*. 2013; 135(6):061006.
4. Fu C, Chin K-S. Robust evidential reasoning approach with unknown attribute weights. *Knowledge-Based Systems*. 2014; 59:9–20.

5. Wang Y, Dai Y, Chen Y-w, Meng F. The evidential reasoning approach to medical diagnosis using intuitionistic fuzzy Dempster-Shafer theory. *International Journal of Computational Intelligence Systems*. 2015; 8(1):75–94.
6. Deng Y, Mahadevan S, Zhou D. Vulnerability Assessment of Physical Protection Systems: A Bio-Inspired Approach. *International Journal of Unconventional Computing*. 2015; 11.
7. Garcia ML. *Vulnerability assessment of physical protection systems*: Butterworth-Heinemann; 2005.
8. Liu H-C, Liu L, Lin Q-L. Fuzzy failure mode and effects analysis using fuzzy evidential reasoning and belief rule-based methodology. *IEEE Transactions on Reliability*. 2013; 62(1):23–36.
9. Su X, Mahadevan S, Xu P, Deng Y. Dependence assessment in human reliability analysis using evidence theory and AHP. *Risk Analysis*. 2015; 35(7):1296–316. <https://doi.org/10.1111/risa.12347> PMID: [25847228](https://pubmed.ncbi.nlm.nih.gov/25847228/)
10. Kang B, Deng Y, Sadiq R, Mahadevan S. Evidential cognitive maps. *Knowledge-Based Systems*. 2012; 35:77–86.
11. Tianlu C, Peiwen Q. Target recognition based on modified combination rule. *Journal of Systems Engineering and Electronics*. 2006; 17(2):279–83.
12. Deng Y, Su X, Wang D, Li Q. Target recognition based on fuzzy dempster data fusion method. *Defence Science Journal*. 2010; 60(5).
13. McKleroy VS, Galbraith JS, Cummings B, Jones P. Adapting evidence-based behavioral interventions for new settings and target populations. *AIDS Education and Prevention*. 2006; 18:59. <https://doi.org/10.1521/aeap.2006.18.suppl.59> PMID: [16987089](https://pubmed.ncbi.nlm.nih.gov/16987089/)
14. Yang J-B, Xu D-L. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*. 2002; 32(3):289–304.
15. Yang J-B, Xu D-L. Evidential reasoning rule for evidence combination. *Artificial Intelligence*. 2013; 205:1–29.
16. Zhang M-J, Wang Y-M, Li L-H, Chen S-Q. A general evidential reasoning algorithm for multi-attribute decision analysis under interval uncertainty. *European Journal of Operational Research*. 2017; 257(3):1005–15.
17. Zadeh LA. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI magazine*. 1986; 7(2):85.
18. Yager RR. On the Dempster-Shafer framework and new combination rules. *Information sciences*. 1987; 41(2):93–137.
19. Smets P. Belief functions: the disjunctive rule of combination and the generalized bayesian theorem. *Classic Works of the Dempster-Shafer Theory of Belief Functions*: Springer; 2008. p. 633–64.
20. Smets P. The combination of evidence in the transferable belief model. *IEEE Transactions on pattern analysis and machine intelligence*. 1990; 12(5):447–58.
21. Dubois D, Prade H. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*. 1988; 4(3):244–64.
22. Yong D, WenKang S, ZhenFu Z, Qi L. Combining belief functions based on distance of evidence. *Decision support systems*. 2004; 38(3):489–93.
23. Liu Z-g, Pan Q, Cheng Y-m, Dezert J. Sequential adaptive combination of unreliable sources of evidence. *Advances and Applications of DSMT for Information Fusion, Vol IV: Collected Works*. 2015:23.
24. Martin A, Jousselme A-L, Osswald C, editors. Conflict measure for the discounting operation on belief functions. *Information Fusion, 2008 11th International Conference on*; 2008: IEEE.
25. Deng Y. Generalized evidence theory. *Applied Intelligence*. 2015; 43(3):530–43.
26. Smarandache F, Dezert J. *Advances and Applications of DSMT for Information Fusion (Collected works)*, second volume: *Collected Works: Infinite Study*; 2006.
27. Dezert J, Smarandache F. *DSMT: A new paradigm shift for information fusion*. arXiv preprint [cs/0610175](https://arxiv.org/abs/0610175). 2006.
28. Dempster AP, Chiu WF. Dempster–Shafer models for object recognition and classification. *International Journal of Intelligent Systems*. 2006; 21(3):283–97.
29. Murphy CK. Combining belief functions when evidence conflicts. *Decision support systems*. 2000; 29(1):1–9.
30. Han D, Deng Y, Han C, Yang Y. Some notes on betting commitment distance in evidence theory. *Science China Information Sciences*. 2012; 55(3):558–65.
31. He Y, Hu L, Guan X, Deng Y, Han D. New method for measuring the degree of conflict among general basic probability assignments. *Science China Information Sciences*. 2012; 55(2):312–21.

32. Jusselme A-L, Grenier D, Bossé É. A new distance between two bodies of evidence. *Information fusion*. 2001; 2(2):91–101.
33. Deng X, Hu Y, Deng Y, Mahadevan S. Environmental impact assessment based on D numbers. *Expert Systems with Applications*. 2014; 41(2):635–43.
34. Ristic B, Smets P. The TBM global distance measure for the association of uncertain combat ID declarations. *Information fusion*. 2006; 7(3):276–84.
35. Liu W. Analyzing the degree of conflict among belief functions. *Artificial Intelligence*. 2006; 170(11):909–24.
36. Liu Z-g, Dezert J, Pan Q, Mercier G. Combination of sources of evidence with different discounting factors based on a new dissimilarity measure. *Decision Support Systems*. 2011; 52(1):133–41.
37. Jiang W, Wei B, Xie C, Zhou D. An evidential sensor fusion method in fault diagnosis. *Advances in Mechanical Engineering*. 2016; 8(3):1687814016641820.
38. Deng Y. *Deng entropy: a generalized Shannon entropy to measure uncertainty*. 2015.
39. Shafer G. *A mathematical theory of evidence*: Princeton university press Princeton; 1976.
40. Dempster AP. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*. 1967:325–39.
41. Smets P. Decision making in the TBM: the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*. 2005; 38(2):133–47.
42. Jiang W, Peng J-y, Deng Y. New representation method of evidential conflict. *Systems Engineering and Electronics*. 2010; 32(3):562–5.
43. Hamacher H. *Über logische Aggregationen nicht-binär explizierter Entscheidungskriterien: Ein axiomatischer Beitrag zur normativen Entscheidungstheorie*: Rita G Fischer Verlag; 1978.
44. Chhipi-Shrestha GK. River water quality assessment of the Manahara river by using macroinvertebrates as biological indicators, Kathmandu Valley. MSc thesis Central Department of Environmental Science, Tribhuvan University. 2007.
45. Chhipi-Shrestha GK. Ecological status of the Manahara river and community initiatives in wastewater management for preservation of the river. *Proceeding of International symposium on Community-led Management of River Environment* August. 2007:163–78.
46. Chhipi-Shrestha GKaP A.M.S. Water quality monitoring of Manahara river: An application of chemical water quality indices. *Journal of SAN Stratigraphical Association of Nepal, Kathmandu*. 2007; 6:55–62.
47. Kumar AY, Reddy MV. Assessment of seasonal effects of municipal sewage pollution on the water quality of an urban canal—a case study of the Buckingham canal at Kalpakkam (India): NO<sub>3</sub>, PO<sub>4</sub>, SO<sub>4</sub>, BOD, COD and DO. *Environmental monitoring and assessment*. 2009; 157(1–4):223–34. <https://doi.org/10.1007/s10661-008-0530-0> PMID: 18949568
48. Zia H, Harris NR, Merrett GV, Rivers M, Coles N. The impact of agricultural activities on water quality: A case for collaborative catchment-scale management using integrated wireless sensor networks. *Computers and electronics in agriculture*. 2013; 96:126–38.
49. Ernst C, Gullick R, Nixon K. Conserving forests to protect water. *Am Water W Assoc*. 2004; 30:1–7.
50. Cunha DGF, Sabogal-Paz LP, Dodds WK. Land use influence on raw surface water quality and treatment costs for drinking supply in São Paulo State (Brazil). *Ecological Engineering*. 2016; 94:516–24.