Intelligent Data Analysis 15 (2011) 399–411
DOI 10.3233/IDA-2010-0474
IOS Press

A novel information fusion method based on Dempster-Shafer evidence theory for conflict resolution

Jianping Yang[a,b], Hong-Zhong Huang[a,*], Qiang Miao[a] and Rui Sun[a]
[a]*School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China*
[b]*Unit 95389 of People's Liberation Army, Guilin, China*

Abstract. Evidence conflict that may cause the counter-intuitive results is one of the most concerns for information fusion by Dempster-Shafer's (D-S) evidence theory. To deal with the issue and manage evidence conflict greatly for the improvement of belief convergence, evidence conflict and belief convergence are investigated based on the analysis of the coherence degree between two sources of evidence. Moreover, the stochastic interpretation for basic probability assignment (BPAs) is illustrated. In addition, a few methods in dealing with evidence conflict are analyzed and compared. Then, a new paradox combination algorithm based on an absolute difference factor of two pieces of evidence and a relative difference factor of two pieces of evidence for a specific hypothesis are proposed with the consideration of local attributions to local conflict. The newly proposed algorithm is verified by the numerical example. The analysis shows the efficiency of the proposed method to improve the performance of belief convergence, in which the comparison studies indicate the advantages of the proposed method as well.

Keywords: D-S evidence theory, belief convergence, evidence conflict, information fusion

1. Introduction

It is well known that the single source information cannot reflect the real properties of the complicated systems. More sensors are employed to acquire information to catch the exact characteristics for the systems. With the development of data acquisition systems, computer science and sensor techniques, the mass information of a system for state identification purpose can be collected, which bring new challenges to deal with the problem of massive information by effective methods. Therefore, a lot of researches have been conducted to handle with this issue. Bomberger [1] proposed an approach to higher-level information fusion based on semantic knowledge networks composed of simulated spiking neurons. The semantic networks had been organized into knowledge network modules, each of which represented a given domain of knowledge. However, when more realistic and complex situations appear, the method needs more options to represent corresponding relationships among evidence. Perlovsky [2] described a mathematical technique on the flexible and adaptive high-level fusion of integrated sensor-communication of cognition-language systems. Yue [3] proposed a collision detection system consisting of two specialized neural networks to extract and fuse different visual cues. Carpenter [4] established an

*Corresponding author. Tel.: +86 28 6183 0248; Fax: +86 28 6183 0229; E-mail: hzhuang@uestc.edu.cn.

information fusion system by using distributed code representation that was able to exploit the capacity of neural networks with the purpose to produce self-organizing expert systems that could discover hierarchical knowledge structures. Kamberova [5] presented an approach of information fusion based on statistical decision theory and application of the method to mobile robot localization. Bayes decision rule based on the probability theory is a classical way of data fusion [6]. With the increase of system complexity and application of more sensors, the disadvantage of Bayes theory occurs. This theory demands a large quantity of known information to construct the priori probability. With the strict constraints of the source axioms of probability, the construction of the priori database becomes very difficult. Moreover, it is unable to easily model the information that is imprecise, incomplete or not totally reliable [7]. Unfortunately, some collected information may be imprecise, incomplete or not totally reliable in practice, especially for systems contain lots of components. This makes it very difficult to construct the priori probability and limits the implementation of Bayes decision rule for this case [7]. Thus, to deal with these disadvantages, the theories of uncertainty have been appeared as alternatives to Bayes theory, such as Dempster-Shafer (D-S) evidence theory [8], fuzzy sets theory and possibility theory [9–11].

D-S evidence theory is proposed based on the research of Dempster [8]. The advantage of the method is in the information fusion with the considerations of both information imprecision and uncertainty in multi-information analysis. This theory employs prior probability called Basic Probability Assignment (BPA) to attain posteriori confidence interval, which is presented by the upper and lower limits called Plausibility Function (*Pl*) and Belief Function (*Bel*), respectively. Thus, Dempster's combination rule of D-S evidence theory can combine the evidence obtained from different evidence sources to update a belief on a specific individual or the set of some individuals. Because of the capability to provide a federative framework, evidence theory is suitable to take into account the disparity of the knowledge types. In addition, it can be considered as the generalization of the Bayesian inference to process the uncertain data associated with no exclusive hypotheses. Because D-S evidence theory is more adaptive to the probabilistic nature of the data, it has been used in machinery fault diagnosis [12–15], medical diagnosis [16], knowledge discovery [17], and so on.

Though D-S evidence theory is attractive, Zadeh [18] has indicated that Dempster's combination rule might produce counter-intuitive results (also named combination paradox) when highly conflicting evidence occur. Many scholars have made some efforts to solve the issue. Yager [19] proposed an algorithm to distribute conflict belief to unknown proposition completely. This algorithm is more reasonable than that of D-S evidence theory in dealing with the combination paradox. However, it is conservative to assign belief that the combination result is undesirable in combining multiple sources of evidence. Dubois and Prade [20] investigated the combining uncertain evidence stemming from several sources, which are nonreliable sources, nonexhaustive sources, inconsistent sources, and dependent sources in the field of artificial intelligence and proposed the disjunctive combination rule. Smets [21] proposed the transferable belief model (TBM) and unnormalized combination rule to manage non-conflict and conflict evidence. Lefevre [22] thought that conflict management was a major problem especially during the fusion of many information sources. Consequently, a formalism to describe a family of combination operators had been defined. Moreover, a generic framework for the fusion of information sources modeled has presented by means of belief mass functions. The average rule of combination proposed by Murphy [23] was just to average all the BPAs of relevant hypothesis to get new belief assignments. However, the method proposed by Murphy cannot generate a desired belief convergence. Chen [24] presented a modified averaging method to combine conflicting evidence based on the distance of evidence and gave the weighted average of the evidence for information fusion.

However, this method has the following limitations: (a) the reason of belief convergence is not addressed even if belief convergence performance may be accepted, (b) the reason to employ evidence distance for removing evidence conflict is not explained clearly, and (c) the efficiency of the algorithm needs improvement.

Based on the above review, the conflict issue has not yet been solved very well. It should be noted that the real reason disclosure to generate combination paradox might benefit the solution of conflict issues. Unfortunately, this has not been investigated based on our literature review. Therefore, the conflict reason with the desire to modify the traditional D-S evidence theory is concerned and a novel information fusion method that may greatly eliminate conflict effect for information fusion is proposed.

The rest of the paper is organized as below. D-S evidence theory is briefly introduced in Section 2. The stochastic interpretation for BPAs is expressed in Section 3. The reasons of conflict issue existing in D-S evidence theory are investigated in Section 4. A novel information fusion method is proposed in Section 5. Comparison studies are performed in Section 6. Conclusions are presented at last.

2. A brief introduction of D-S evidence theory

D-S evidence theory is represented by a finite nonempty exhaustive set of mutually exclusive possibilities called a frame of discernment, Θ [8]. 2^{Θ} is the power set of Θ, which includes all the possible subsets of Θ. There are 2^n elements in 2^{Θ}, if Θ has n elements. Let q_i be the ith possibility and $i = 1,2, \cdots, n$, then,

$$2^{\Theta}=\{\emptyset, \{q_1\}, \ldots, \{q_n\}, \{q_1, q_2\}, \ldots, \{q_1, q_n\}, \ldots, \{q_{n-1}, q_n\}, \{q_1, q_2, q_3\}, \ldots, \{q_1, q_2, \ldots, q_n\}\} \quad (1)$$

where $\emptyset$ denotes the empty set. The subsets $\emptyset, \{q_1\}, \cdots, \{q_n\}$ including only one element are called singletons.

Definition 1. Basic probability assignment function (BPA): $m(X) : 2^{\Theta} \to [0, 1]$, and satisfies $m(\emptyset) = 0$, $\sum_{X \subseteq \Theta} m(X) = 1$.

where X is the element of 2^{Θ}. $m(X)$ is a measure of the belief attributed exactly to the hypothesis X. $m(\emptyset) = 0$ means that the existing evidence supports no element of the domain. $m(X) = 1$ states that an existing evidence only supports X in the domain. $\sum_{X \subseteq \Theta} m(X) = 1$ guarantees the normalization of evidence. Here it should be noted that Smets proposed the open world assumption (OWA) and considered that it is possible to have $m(\emptyset) \neq 0$ [21]. Contrarily, $m(\emptyset) = 0$ is the close world assumption (CWA) in classical D-S evidence theory. In order to keep consistent with conventional theory, this paper concentrates on CWA.

Definition 2. Belief function (Bel):

$$Bel(X) = \sum_{Y \subseteq X, X \subseteq \Theta} m(Y) \quad (2)$$

where $Bel(X)$ represents the total amount of probability that must be distributed among the elements of X. It reflects inevitability, signifies the total degree of belief of X and constitutes a lower limit function on the probability of X [8].

Table 1
The BPAs for Example 1

	A	B	AB
S_1	$m_1(A) = 0.4$	$m_1(B) = 0.5$	$m_1(AB) = 0.1$
S_2	$m_2(A) = 0.6$	$m_2(B) = 0.2$	$m_2(AB) = 0.2$

Definition 3. Plausibility function (Pl):

$$Pl(X) = \sum_{Y \cap X \neq \Phi, X, Y \subseteq \Theta} m(Y) = 1 - Bel(\overline{X}) \tag{3}$$

where $\overline{X}$ is the negation of a hypothesis X. $Pl(X)$ measures the maximal amount of probability that can be distributed among the elements in X. It describes the total belief degree related to X and constitutes an upper limit function on the probability of X [13,26]. Moreover,

$$Bel(X) \leqslant m(X) \leqslant Pl(X) \tag{4}$$

The fusion of multiple evidence can be performed by the Dempster's combination rule that is defined as follow. Given two Basic probability assignment functions $m_i(X)$ and $m_j(Y)$, the Dempster's combination rule can be defined by

$$m(C) = m_i(X) \oplus m_j(Y) = \begin{cases} 0 & X \cap Y = \emptyset \\ \frac{\sum_{X \cap Y = C, \forall X, Y \subseteq \Theta} m_i(X) \cdot m_j(Y)}{\sum_{X \cap Y = \emptyset, \forall X, Y \subseteq \Theta} m_i(X) \cdot m_j(Y)} & X \cap Y \neq \emptyset \end{cases} \tag{5}$$

where $m_{i(j)}(C)$ denotes the BPA of C that is supported by the $i^{th}(j^{th})$ evidence.

Let

$$K_{ij} = \sum_{X \cap Y = \emptyset, \forall X, Y \subseteq \Theta} m_i(X) \cdot m_j(Y) \tag{6}$$

where K_{ij} is called the conflict factor which expresses the conflict degree between the i^{th} and j^{th} evidence. $0 \leqslant K_{ij} \leqslant 1$. $K_{ij} = 0$ means that evidence i and j have no conflict. While $K_{ij} = 1$ or 0 $< K_{ij} < 1$ represents that two pieces of evidence have complete conflict or partial conflict to support an opinion. An example is employed to illustrate the implementation of Dempster's combination rule. Let S_i be the i^{th} source of evidence.

Example 1. Assume two sources of evidence S_1 and S_2 obtained from a system and the corresponding BPAs are shown in Table 1. A, B, and *AB* are all the propositions discussed in this example.

According to Eq. (5), $m(A) = 0.61$, $m(B) = 0.35$, $m(AB) = 0.03$ are obtained, respectively. The conflict factor $K_{12} = 0.38$ is obtained according to Eq. (6). It is clear that both $m(A) > m_1(A)$ and $m(A) > m_2(A)$. In addition, $m(AB) < m_1(AB)$, $m(AB) < m_2(AB)$ and $m_2(B) < m(B) < m_1(B)$ can be also noticed. Therefore, these may be concluded though it is hard to explain $m(B)$. Moreover, it can be found that the support to A is strengthened and the support to *AB* is weakened with the fusion of multiple sources of evidence. This changing process of brief assignment through combination is also called brief convergence.

3. Stochastic interpretation for BPAs

With the brief review of D-S evidence theory, it can be gained that a body of evidence represents the basic belief assignment information through BPA in a situation at a given time. A BPA can be taken as a discrete random function whose variable is a probability distribution $m(\cdot)$ of 2^Θ. Consequently, A BPA can be easily represented using vector notation whose elements are discrete random $m(\cdot)$of 2^Θ and be dealed with by elementary vector algebra.

Definition 4. Probability vector $(\mathbf{p})$: $\mathbf{p} = (p_1, p_2, \cdots, p_n)$, which satisfies the following conditions:

$$0 \leqslant p_i \leqslant 1, i = 1, 2, \cdots, n \tag{7}$$

$$\sum_{i=1}^{n} p_i = 1 \tag{8}$$

According to the definition of a BPA, it can be received that a BPA is a special case of probability vector which have 2^n elements and can be noted as $\mathbf{m} = (m(\emptyset), m(X_1), m(X_2), \cdots, m(X_{2^n-1}))$. Consequently, the elements of $\mathbf{m}$ satisfy:

$$\sum_{i=1}^{2^n-1} m(X_i) = 1 \text{ and } 0 \leqslant m(X_i) \leqslant 1, i = 1, 2, \cdots, 2^n - 1 \tag{9}$$

$X_i \in 2^\Theta$ where $m(\emptyset) = 0$.

Definition 5. Probability unit vector $(\mathbf{e}_i)$: the probability vector, formed by

$$\mathbf{e}_i = (\underbrace{0, 0, \cdots, 0}_{i-1}, 1, \underbrace{0, \cdots, 0}_{n-i}), \quad i = 1, 2, \cdots, n \tag{10}$$

is called probability unit vector.

In the application of D-S evidence theory to information fusion, the best result is expected to be a singleton. For example, object identification, fault diagnosis and so on. In order to express the ideal combination result, probability unit vector is used. In other words, this means that the combined BPA of absolutely certain of a singleton is equal to 1. Because the frame of discernment Θ includes the empty set $\emptyset$ and belief is distributed in the power set of Θ, the probability unit vector is changed to

$$\mathbf{e}_i^m = (0, \underbrace{0, 0, \cdots, 0}_{i-2}, 1, \underbrace{0, \cdots, 0}_{2^n-i}), \quad 2 \leqslant i \leqslant 2^n \tag{11}$$

Definition 6. Stochastic matrix $(\mathbf{P})$:

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mn} \end{pmatrix} \tag{12}$$

matrix $\mathbf{P}$ which satisfies every row vector is probability vector called Stochastic matrix. D-S evidence theory can fuse multiple sources of evidence and every source information of evidence is expressed by the BPA function. Consequently, every source of evidence can be used as a row of Stochastic matrix. The BPAs of k independent sources of evidence can form $k \times 2^n$Stochastic matrix. This matrix is also called as the mass Stochastic matrix which can be denoted as$\mathbf{M}$.

$$\mathbf{M} = \begin{pmatrix} m_1(\emptyset) & m_1(X_1) & \cdots & m_1(X_{2^n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ m_k(\emptyset) & m_k(X_1) & \cdots & m_k(X_{2^n-1}) \end{pmatrix} \tag{13}$$

$$X_i \in 2^{\Theta}$$

The matrix $\mathbf{M}$ characterizes all information available of evidence source which has to be combined to solve the fusion problem. D-S combination rule can be applied to combine k rows of $\mathbf{M}$ to get fusion result. Simultaneously, it is easy to analyze the relation between different sources of evidence.

4. The effects of BPA difference to both belief convergence and evidence conflict

Based on the analysis in Section 2, it can be noticed that the evidence theory can get more precise characteristics of a system compared with that obtained by a single evidence. The main idea is to obtain a belief convergence through the combination of multiple sources of evidence including suspending evidence. It should be noted that when the opinions of evidence are different, evidence conflict may occur. The combining results obtained by D-S evidence theory might be wrong correspondingly. This phenomenon is called combination paradox in this study. To solve the issue, it is significant to analyze the principles of belief convergence and disclose the reason of combination paradox. When the BPAs of evidence are the same, it is clear that there is no conflict among evidence. However, different BPAs of evidence may generate conflict that results in combination paradox [27]. Therefore, the research is extended as follows.

4.1. The BPAs of evidence are the same

In this case, the BPAs of all sources of evidence of each same subset in the discernment frame are completely the same. In other words, the elements of the row and the column are equal to each other in the mass Stochastic matrix $\mathbf{M}$. It is easy to achieve the rank of $\mathbf{M}$ equals to 1. There is no diversity between two different sources of evidence. Therefore, the BPA of the pre-combination and post-combination should be equal. This does not generate belief convergence. In order to be understood better, Example 2 is used to analyze this case.

Example 2. Let $\Theta = \{A, B, C, D\}$, two sources of evidence S_1 and S_2. The BPAs of the two sources of evidence are given in Table 2, and the mass Stochastic matrix $\mathbf{M}_1$ is expressed by $\mathbf{M}_1 = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$.

Based on Eqs (5) and (6), $m(A) = 0.25$, $m(B) = 0.25$, $m(C) = 0.25$, $m(D) = 0.25$ and $K_{12} = 0.75$ are obtained respectively. Comparing the updated BPAs with the given BPAs, it can be found that the BPAs have no influence even though more sources of evidence are considered when the given BPAs are the same for each subset. This can be concluded that it is not able to obtain the updated BPAs if there is no difference among the given evidence BPAs. The desired belief convergence does not present. In

Table 2
The BPAs for Example 2

	A	B	C	D
S_1	$m_1(A) = 0.25$	$m_1(B) = 0.25$	$m_1(C) = 0.25$	$m_1(D) = 0.25$
S_2	$m_2(A) = 0.5$	$m_2(B) = 0.25$	$m_2(C) = 0.25$	$m_2(D) = 0.25$

addition, the existing evidence has no conflict in this example. However, the K_{12} obtained by Eq. (6) is 0.75. This is inconsistent with the statement that K_{12} should be equal to 0 for the evidence without conflict.

4.2. *For each subset, the BPAs of different sources of evidence are the same; the BPAs for different subsets are different*

Because the BPAs for different subsets supported by multiple sources of evidence are different in this kind of cases, the fusion results of multiple sources of evidence by combination rule may be different. This shows that the difference of the BPAs of evidence may be a reason to obtain the updated BPAs. In this case, the row of mass Stochastic matrix $\mathbf{M}$is equal to each other and the rank of $\mathbf{M}$ is equal to 1. The BPAs of k independent sources of evidence are the same, $m_g(X_i) = m_h(X_i)$, $X_i \in 2^{\Theta}$. $m_g(\cdot)$ and $m_l(\cdot)$ are two random sources of evidence. This can be illustrated by Example 3.

Example 3. Let $\Theta = \{A, B, C\}$. The BPAs supported by two sources of evidence S_1 and S_2 are shown in the mass Stochastic matrix $\mathbf{M}_2$.

$$\mathbf{M}_2 = \begin{pmatrix} m_1(A) & m_1(B) & m_1(C) \\ m_2(A) & m_2(B) & m_2(C) \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.6 & 0.3 & 0.1 \end{pmatrix}$$

Based on Eq. (6), $K_{12} = 0.54$ is attained. Using Eq. (5), $m(A) = 0.7826$, $m(B) = 0.1957$, $m(C) = 0.0217$ are obtained. Comparing the updated BPAs and the given ones, it is noticed that $m(A)$, $m(B)$, and $m(C)$ change from 0.6, 0.3, and 0.1 to 0.7826, 0.1957, and 0.0217, respectively. From this example, the combined BPAs are changed and brief convergence is generated. Moreover, information of two sources of evidence is completely the same. The reason for brief convergence is only the difference of BPA of single evidence. Consequently, it can be received that the difference among evidence BPAs is one of the reasons to update BPAs. Simultaneously K_{12} is not equal to 0. Based on the above analysis, this may be concluded that for non-conflict evidence, K_{12} may not be equal to 0 such as $K_{12} = 0.54$.

4.3. *The evidence BPAs are different for a same subset in a discernment frame*

When the BPAs of two sources of evidence for a subset in a discernment frame are different, the updated BPAs can be obtained. The row of the mass Stochastic matrix M is unequal to each other and the rank of M is unequal to 1. Example 4 is employed to illustrate this situation.

Example 4. Let $\Theta = \{A, B, AB\}$. The BPAs supported by two sources of evidence S_1 and S_2 are shown in the mass Stochastic matrix $\mathbf{M}_3$.

$$\mathbf{M}_3 = \begin{pmatrix} m_1(A) & m_1(B) & m_1(AB) \\ m_2(A) & m_2(B) & m_2(AB) \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{pmatrix}$$

Based on the BPAs in $\mathbf{M}_3$ and Eq. (6), $K_{12} = 0.31$ is obtained. Through Eq. (5), the updated BPAs as $m(A) = 0.7102$, $m(B) = 0.1449$, $m(AB) = 0.1449$ are concluded. It can be gained that $m(A)$

supported by two different sources of evidence are updated from 0.7 and 0.5 to 0.7102 eventually. It means that the belief of A based on sources of evidence S_1 and S_2 has been enhanced. Additionally, the decreasing of both $m(B)$ and $m(AB)$ can be noticed. Therefore, it may be concluded that the difference of the BPAs for a same subset is one of the reasons to update BPAs as well.

Based on the above analysis of three situations, it can be verified that the difference among BPAs is potentially useful information for fusion information and the analysis of evidence conflict. Therefore, for the combination rule, it is significant to utilize the difference of evidence BPAs effectively. It should be noted that the extreme situation of the third case may appear when the BPAs of relevant unit conflict completely using the classical Dempster-Shafer's (D-S) combination rule. Example 5 can be used to present the situation.

Example 5. Let $\Theta = \{A, B, C\}$. The BPAs supported by two sources of evidence S_1 and S_2 are given in the mass stochastic matrix $\mathbf{M}_4$.

$$\mathbf{M}_4 = \begin{pmatrix} m_1(A) & m_1(B) & m_1(C) \\ m_2(A) & m_2(B) & m_2(C) \end{pmatrix} = \begin{pmatrix} 0.99 & 0.01 & 0 \\ 0 & 0.01 & 0.99 \end{pmatrix}$$

Using D-S combination rule Eq. (5), $K_{12} = 0.9999$, $m(A) = 0$, $m(B) = 1$ can be obtained, and $m(C) = 0$. From these results, it can be seen that both S_1 andS_2 offer little support to B, i.e. $m_1(B) = 0.01$ and $m_2(B) = 0.01$. However, the updated result offers complete support to B, i.e. $m(B) = 1$, using the classical D-S combination rule. This means that the opinion with very low belief level such as $m_1(B) = 0.01$ and $m_2(B) = 0.01$ will come true after information fusion. This is obviously counter-intuitive result that cannot be accepted. Therefore, the modification of D-S combination rule is necessary. In addition, the conflict reason for D-S combination rule obtained in this study may have a great potential to remove the evidence conflict and improve the capability of information fusion.

5. The proposed combination algorithm

5.1. The novel paradox combination algorithm

The difference of BPAs of subset is useful information for belief convergence, which is the essential reason of belief convergence using D-S evidence theory to combine multiple sources of evidence in information fusion. Consequently, it is important to manage the difference of the BPAs in order to obtain belief convergence. Simultaneously, the key question of the solution of convergence paradox is decided with the rational management of the difference of BPAs. Consequently, this can be proposed that a novel paradox combination algorithm with the desire to obtain reasonable information fusion results no matter how the conflict level is. The combination rule, belief update, and common sense are considered in the proposed algorithm. Simultaneously, two factors, i.e. the absolute difference factor of two pieces of evidence and the relative difference factor of two pieces of evidence for a specific hypothesis, are proposed to depict the conflicting degree between two sources of evidence and express the attribution degree of convergence of belief difference of every subset. The absolute difference factor of two pieces of evidence and the relative difference factor of two pieces of evidence for a specific hypothesis are defined as below.

Definition 7. The absolute difference factor of two pieces of evidence is defined by

$$E_{ij} = \sum_k |m_i(X_k) - m_j(X_k)|, \quad k = 1, 2 \cdots, 2^n \tag{14}$$

E_{ij} represents the difference degree between sources of evidence i and j. When E_{ij} becomes larger, the difference between the two sources of evidence will be larger. Because $0 \leqslant m(\cdot) \leqslant 1$, it is easy to know that $0 \leqslant E_{ij} \leqslant 2$. When $E_{ij} = 0$, it corresponds to the second situation discussed in Section 4. When E_{ij} is close or equal to 2, it corresponds to the third situation discussed in Section 4. Evidence conflict will be severer than other situations.

Definition 8. The relative difference factor of two pieces of evidence for a specific hypothesis is given by

$$E^r_{ij}(X_k) = \frac{|m_i(X_k) - m_j(X_k)|}{\sum\limits_k |m_i(X_k) - m_j(X_k)|}, \quad k = 1, 2 \cdots, 2^n \tag{15}$$

If $E^r_{ij}(X_k)$ is small, it means the similarity of two sources of evidence i and j is big. When $E^r_{ij}(X_k) = 0$, it means the BPAs of X_k supported by two sources of evidence i and j are the same.

With the introduction of the relative difference factor of two pieces of evidence for a specific hypothesis, it can be proposed to obtain the updated BPAs for the case $E_{ij} \neq 0$ by

$$m(X_k) = \begin{cases} 0 & X \cap Y = \emptyset \\ E^r_{ij}(X_k) \cdot K_{ij} + \sum_{X \cap Y = X_k, \forall X, Y \subseteq \Theta} m_i(X) \cdot m_j(Y) & X \cap Y \neq \emptyset \text{ and } X \cap Y \neq \Theta \end{cases} \tag{16}$$

and

$$m(\Theta) = 1 - \sum_k (E^r_{ij}(X_k) \cdot K_{ij} + \sum_{X \cap Y = X_k, \forall X, Y \subseteq \Theta} m_i(X) \cdot m_j(Y)) \tag{17}$$

When $E_{ij} = 0$, the updated BPAs by Eq. (5) are obtained. Using Eq. (16) to solve Example 5, $E_{12} = 1.98$, $E^r_{12}(A) = 0.5$, $E^r_{12}(B) = 0$, $E^r_{12}(C) = 0.5$ are concluded respectively. The updated BPAs as $m(A) = 0.4995$, $m(B) = 0.001$, and $m(C) = 0.4995$ are obtained. Because the BPAs of two sources of evidence to support hypothesis B are the same and have very small belief contribution, the BPAs to hypothesis B should be less after combination. From the above results, $m(B)$ has changed from 0.1 to 0.001. This result shows the novel combination algorithm is more reasonable and rational than D-S combination rule. Simultaneously, because of the same weight to various evidence, the BPAs to hypotheses A and C are extremely close to 0.5. Therefore, the combination results clearly indicate that the proposed combination algorithm is effective.

5.2. The novel combination rule

In practical application, multiple sources of evidence can be classified into either conflicting sources or non-conflicting sources, which is proposed by Fan [13]. Simultaneously, the classical D-S combination rule can effectively deal with the evidence of non-conflicting sources. To handle the evidence of conflicting sources, the novel paradox combination algorithm can be applied to obtain a rational combination consequence. Consequently, in practice, these two rules could be appropriately synthesized or combined depending on the conflicting relationship of different sources of evidence. For each pair of

Table 3
Some current combination rules

	Combination algorithm
Yager rule	$\begin{cases} m(C) = \sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m_i(X)\cdot m_j(Y) \\ m(\Theta) = \sum_{X\cap Y=\Phi,\forall X,Y\subseteq\Theta} m_i(X)\cdot m_j(Y) \\ m(\Phi) = 0 \end{cases}$
Murphy rule	$\begin{cases} m(C) = \frac{\sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m_i(X)\cdot m_j(Y)}{1-\sum_{X\cap Y=\Phi,\forall X,Y\subseteq\Theta} m_i(X)\cdot m_j(Y)} \\ m(\Phi) = 0 \\ \sum\limits_{C\in 2^\Theta} m(C) = 1 \end{cases}$
Chen rule	$\begin{cases} m(C) = \frac{\sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m_i*(X)\cdot m_j*(Y)}{1-\sum_{X\cap Y=\Phi,\forall X,Y\subseteq\Theta} m_i*(X)\cdot m_j*(Y)} \\ m*(\Phi) = 0 \\ \sum\limits_{C\in 2^\Theta} m*(C) = 1 \end{cases}$

sources of evidence, the conflict degree is characterized by the conflict factor K_{ij} which is defined in Eq. (6). Therefore, the process of evidence classification can be expressed in the following. Firstly, the conflict factor for each pair of sources of evidence should be calculated and ranked from the largest to the smallest. Secondly, it is necessary to evaluate the conflict threshold denoted by ε for ε∈[0, 1]. The threshold aims to represent the permitted conflict level between two sources of evidence. If K_{il} is greater than ε, one of the two sources of evidence will be put into the conflicting group while the other remains in the non-conflicting group [13]. The criterion to put into the conflicting group is determined based on the conflict degree of this evidence relative to the others. For instance, if $K_{ij} > \varepsilon$ and $K_{il} > \varepsilon$, only evidence i will be put into the conflicting group. Equation (16) can be used as the criterion to consolidate all conflicting sources of evidence as one group. Every source of evidence has to be classified into either the conflicting group or the non-conflicting group. The aggregated BPA for all sources of evidence in the conflicting group will then be combined with the BPAs in the non-conflicting group using Eq. (6). It should be noted that there is no such an "absolute meaningful conflict threshold" which is applicable to all applications. The choice of ε also depends on the specific application. In this paper, $\varepsilon = \frac{1}{n}\sum\limits_{i\neq j}^{K_{ij}}$ is adopted from Fan [13,14], and n represents the total number of the conflict factors. For more details about evidence classification and the choice of ε, readers may refer to Fan [13,14] and Ayoun and Smets [25].

6. Comparison studies

To further verify the proposed method, the comparison studies can be conducted among the proposed method and some current methods such as the traditional D-S combination rule, Yager [19] rule, Murphy [23] rule, Chen [24] rule, and others. These algorithms are presented briefly in Table 3. For Murphy rule, the combination incorporating average belief is included. For Chen rule, $m*(\cdot)$ is the modified average BPA. More details can be referred to [19,23,24].

The example in Ref. [24] is used to analyze and address the comparison studies.

Example 6. Assume that there are three fault modes A, B, and C. Moreover, Assume A is authentic fault mode. The BPAs supported by evidence S_1,S_2,S_3,S_4, and S_5 are shown in the mass stochastic matrix

Table 4
The BPAs obtained from two pieces of evidence S_1 and S_2

	m(A)	m(B)	m(C)	m(Θ)
D-S rule	0	0.8571	0.1429	0
Yager rule	0	0.18	0.03	0.79
Murphy rule	0.1543	0.7496	0.0988	0
Chen rule	0.1543	0.7496	0.0988	0
The proposed rule	0.33	0.44	0.14	0.09

$\mathbf{M}_5$.

$$\mathbf{M}_5 = \begin{pmatrix} m_1(A) & m_1(B) & m_1(C) \\ m_2(A) & m_2(B) & m_2(C) \\ m_3(A) & m_3(B) & m_3(C) \\ m_4(A) & m_4(B) & m_4(C) \\ m_5(A) & m_5(B) & m_5(C) \end{pmatrix} = \begin{pmatrix} 0.6 & 0.2 & 0.3 \\ 0 & 0.9 & 0.1 \\ 0.55 & 0.1 & 0.35 \\ 0.55 & 0.1 & 0.35 \\ 0.55 & 0.1 & 0.35 \end{pmatrix}$$

To investigate the details, the information fusion is performed step by step in this study.

First, the conflict factor is calculated between each pair of sources of evidence and rank from large to small, $K_{12} = 0.88$, $K_{23} = K_{24} = K_{25} = 0.875$, $K_{13} = K_{14} = K_{15} = 0.555$, $K_{34} = K_{35} = K_{45} = 0.565$. The conflict threshold, $\varepsilon = \frac{1}{10}(K_{12}+K_{13}+K_{14}+K_{15}+K_{23}+K_{24}+K_{25}+K_{34}+K_{35}+K_{45}) = 0.6865$ is attained. According to Section 5.2, the second source of evidence, S_2, is conflict with the other four sources of evidence. Simultaneously, it can be obtained that S_1 and S_2 are most highly conflicting evidence because the conflict factor K_{12} is maximal. Therefore, Eq. (16) is used to combine S_1 and S_2. In order to compare, the combination results of S_1 and S_2 using five methods are expressed in Table 4.

In Table 4, the results obtained by the classical D-S combination rule do not reflect the actual BPAs. The BPA to A is equal to 0 and the BPA to B is the majority of belief assignment. This result is unacceptable because the true fault mode is A. The BPA to A is also ignored by Yager rule based on Table 4. Additionally, it should be noted that Yager rule distributes the completely conflict BPAs to unknown proposition. The conclusion of Murphy and Chen has been reasonably compared with that of D-S and Yager with the consideration of truths. However, it is not always acceptable for the BPA to A, when there are only two sources of evidence existed, or there are not adequate sources of evidence to make decision that can be obtained with system. Through above combination, the new BPAs is given in Matrix $\mathbf{M}_6$.

$$\mathbf{M}_6 = \begin{pmatrix} m_{1'}(A) & m_{1'}(B) & m_{1'}(C) & m_{1'}(\Theta) \\ m_3(A) & m_3(B) & m_3(C) & m_3(\Theta) \\ m_4(A) & m_4(B) & m_4(C) & m_4(\Theta) \\ m_5(A) & m_5(B) & m_5(C) & m_5(\Theta) \end{pmatrix} = \begin{pmatrix} 0.33 & 0.44 & 0.14 & 0.09 \\ 0.55 & 0.1 & 0.35 & 0 \\ 0.55 & 0.1 & 0.35 & 0 \\ 0.55 & 0.1 & 0.35 & 0 \end{pmatrix}$$

where $m_{1'}(\cdot)$ is expressed for the combined BPAs through evidence S_1 and S_2. The conflict factors between each pair of sources of evidence can be calculated in $\mathbf{M}_6$ and rank from large to small, $K_{1'3} = K_{1'4} = K_{1'5} = 0.642$, $K_{34} = K_{35} = K_{45} = 0.565$. It is obvious that those conflict factors are all less than the conflict threshold $\varepsilon = 0.6865$. Equation (6) should be used to accomplish the combination according to Section 5.2. When the evidence S_3 is analyzed, the results are shown in Table 5.

Because evidence S_3 supports fault mode A, it is obvious that Mode A can be affirmed using the novel algorithm. This conclusion is consistent with the truth. Moreover, the belief assignment to Mode A with the proposed novel algorithm is the largest one among all the results obtained by all the algorithms

Table 5
The BPAs obtained from evidence S_1, S_2 and S_3

	m(A)	m(B)	m(C)	m(Θ)
D-S rule	0	0.6316	0.3468	0
Yager rule	0	0.018	0.0105	0.9715
Murphy rule	0.3504	0.5231	0.1265	0
Chen rule	0.4626	0.3845	0.1529	0
Novel rule	0.6453	0.1480	0.2067	0

Table 6
The BPAs obtained from four and five pieces of evidence

	S_1, S_2, S_3, S_4	S_1, S_2, S_3, S_4, S_5
D-S rule	$m(A) = 0$ $m(B) = 0.3288$ $m(C) = 0.6712$	$m(A) = 0$ $m(B) = 0.1288$ $m(C) = 0.8722$
Yager rule	$m(A) = 0$ $m(B) = 0.0018$ $m(C) = 0.0037$ $m(\Theta) = 0.9945$	$m(A) = 0$ $m(B) = 0.0002$ $m(C) = 0.0013$ $m(\Theta) = 0.9985$
Murphy rule	$m(A) = 0.6027$ $m(B) = 0.2627$ $m(C) = 0.1346$	$m(A) = 0.7958$ $m(B) = 0.0932$ $m(C) = 0.1110$
Chen rule	$m(A) = 0.7419$ $m(B) = 0.1120$ $m(C) = 0.1461$	$m(A) = 0.8827$ $m(B) = 0.0142$ $m(C) = 0.1031$
The proposed method	$m(A) = 0.8029$ $m(B) = 0.0335$ $m(C) = 0.1636$	$m(A) = 0.8972$ $m(B) = 0.0067$ $m(C) = 0.1140$

mentioned in this study and convergence performance is the best. Therefore, the correct conclusion is obtained when three sources of evidence are presented and Mode A will be recognized. Simultaneously, from Table 5 Yager and D-S combination rule cannot obtain the correct mode. Chen combination rule can get the correct mode, but its belief assignment to Mode A is much smaller than the proposed algorithm. Then, subset A can be identified, which means Mode A exists based on the fusion of evidence S_1 to S_3.

Finally, in order to keep the integrity of the example, the combination results of different algorithm are presented in Table 6, when evidence S_4 and S_5 are considered as well.

7. Conclusions

In this study, the reason to belief convergence has been analyzed and the origin of combination conflict has been investigated. Based on the discovery and the traditional combination rule, a new algorithm has been proposed through the introduction of two factors, i.e. the absolute difference factor and the relative difference factor of two pieces of evidence for a specific hypothesis. This method is able to manage evidence effectively, take advantage of the useful information of evidence conflict and improve the reliability and rationality of combination results. The advantages of the proposed method are demonstrated through the comparison study with other information fusion algorithms. It may be recognized that some related topics should be concentrated for information fusion in the future, such as to develop an effective indicator for evidence conflict level quantification.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China under the contract number 50775026. The authors wish to express their thanks to Dr. Xianfeng Fan for his helpful discussions, useful advice and correcting the English manuscript of this paper.

References

[1] N.A. Bomberger, A.M. Waxman and B.J. Rhodes, A new approach to higher-lever information fusion using associative learning in semantic networks of spiking neurons, *Information Fusion* **8**(2007), 227–251.
[2] L.I. Perlovsky, Cognitive high level information fusion, *Information Science* **177** (2007), 2099–2188.
[3] S. Yue and F.C. Rind, Visual motion pattern extraction and fusion for collision detection in complex dynamic scenes, *Computer Vision and Image Understanding* **104** (2006), 48–60.
[4] G.A. Carpenter, S. Marens and O.J. Ogas, Self-organizing information fusion and hierarchical knowledge discovery: a new framework using ARTMAP neural networks, *Neural Network* **18** (2005), 287–295.
[5] G. Kamberova, R. Mandelbaum and M. Mintz, Decision-theoretic approach to robust fusion of location data, *Journal of the Franklin Institute-Engineering and Applied Mathematics* **366** (1999), 269–284.
[6] A. Gelman, J.B. Carlin et al., *Bayesian Data Analysis*, 2nd ed, Chapman and Hall/CRC, 2004.
[7] S. Fabre, A. Appriou and X. Briottet, Presentation and description of two classification methods using data fusion based on sensor management, *Information Fusion* **2** (2001), 49–71.
[8] G. Shafer, *A mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
[9] B.N. Nelson and P.D. Gader, Fuzzy set information fusion in landmine detection, *Proceedings of SPIE-The International Society for Optical Engineering* **3710**(11) (1999), 1168–1178.
[10] M. Oussalah, H. Maaref and C. Barret, New fusion methodology approach and application to mobile robotics: investigation in the framework of possibility theory, *Information Fusion* **2** (2001), 31–48.
[11] C. Sossai, P. Bison and G. Chemello, Sensor fusion for localization using possibility theory, *Control Engineering Practice* **7** (1999), 773–782.
[12] O. Basir and X.H. Yuan, Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory, *Information Fusion* **8** (2007), 379–386.
[13] X. Fan and M.J. Zuo, Fault diagnosis of machines based on D-S evidence theory. Part 1: D-S evidence theory and its improvement, *Pattern Recognition Letters* **27**(5) (2006), 366–376.
[14] X. Fan and M.J. Zuo, Fault diagnosis of machines based on D-S evidence theory. Part 2: application of the improved D-S evidence theory in gearbox fault diagnosis, *Pattern Recognition Letters* **27**(5) (2006), 377–385.
[15] C.R. Parikh, M.J. Pont and N.B. Jones, Application of Dempster–Shafer theory in condition monitoring applications: a case study, *Pattern Recognition Letters* **22**(6–7) (2001), 777–785.
[16] R.W. Jones, A. Lowe and M.J. Harrison, A framework for intelligent medical diagnosis using the theory of evidence, *Knowledge-based Systems* **15** (2002), 77–84.
[17] D. Cai, M.F. Mctear and S.I. Mcclean, Knowledge discovery in distributed databases using evidence theory, *International Journal of Intelligent Systems* **15** (2000), 745–761.
[18] L. Zadeh, A simple view of the Dempster-shafer framework and new combination rules, *AI Magazine* **7** (1986), 85–90.
[19] R.R. Yager, On the Dempster-Shafer framework and new combination rules, *Information Sciences* **41** (1987), 93–138.
[20] D. Dubois and H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, *Computational Intelligence* **3** (1988), 244–264.
[21] P. Smets, The combination of evidence in the transferable belief model, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(5) (1990), 447–458.
[22] E. Lefevre, O. Colot and P. Vannoorenberghe, Belief function combination and conflict management, *Information Fusion* **3** (2002), 149–162.
[23] C.K. Murphy, Combining belief functions when evidence conflicts, *Decisions Support Systems* **29** (2000), 1–9.
[24] L.Z. Chen, W.K. Shi et al., A new fusion approach based on distance of evidence, *Journal of Zhejiang University* (*Science*) **6A**(5) (2005), 476–482.
[25] A. Ayoun and P. Smets, Data association in multi-target detection using the transferable belief model, *International Journal of Intelligent Systems* **16**(10) (2001), 1167–1182.
[26] A. Rakar, D.X. Juricic and P. Balle, Transferable belief model in fault diagnosis, *Engineering Application of Artificial Intelligence* **12** (1999), 555–567.
[27] W. Liu, Analyzing the degree of conflict among belief functions, *Artificial Intelligence* **170** (2006), 909–924.