A Clustering-Based Evidence Reasoning Method

Xinde Li,* Fengyu Wang
Key Laboratory of Measurement and Control of CSE, School of Automation, Southeast University, Ministry of Education, Nanjing 210096, People's Republic of China

Aiming at the counterintuitive phenomena of the Dempster–Shafer method in combining the highly conflictive evidences, a combination method of evidences based on the clustering analysis is proposed in this paper. At first, the cause of conflicts is disclosed from the point of view of the internal and external contradiction. And then, a new similarity measure based on it is proposed by comprehensively considering the Pignistic distance and the sequence according to the size of the basic belief assignments over focal elements. This measure is used to calculate the commonality function of evidences to amend the evidence sources; Meanwhile, the Iterative Self-organizing Data Analysis Techniques Algorithm (ISODATA) method based on the new measure is used for clustering according to the clustering characters of the original evidences. The Dempster rule is applied to combining all the evidences in each clustering into an evidential representative, and the reliability is calculated based on the commonality and the occurrence frequency of the evidences in the clustering. At last, Murphy's method is used to combine these evidential representatives of the different clusterings. The experimental results through a series of numeric examples show that the method proposed in this paper is more effective and superior to others. © 2015 Wiley Periodicals, Inc.

1. INTRODUCTION

Dempster–Shafer theory (DST) is a kind of information processing method first proposed by Dempster in 1967, which can solve the problem of multivalued mapping. Then it was extended by Shafer in 1976.[1] Because it is capable of dealing with uncertainty and has shown excellent performances in the practical engineering, in recent years it has been widely used in many fields, such as uncertainty reasoning, multisensor information fusion, pattern recognition, image processing, fault diagnosis, robot, multiobjective decision, etc. In DST, the basic belief assignments (BBAs) are updated by Dempster's combination rule, which is one of the core foundation of DST. Although its form is relatively simple and suitable for machine implementation, the combinational results might appear the counterintuitive

*Author to whom all correspondence should be addressed; e-mail: xindeli@seu.edu.cn.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 00, 1–24 (2015)

View this article online at wileyonlinelibrary.com. • DOI 10.1002/int.21800

phenomenon in the process of normalization of combination. Since Zadeh[2] noticed this issue, it had attracted widespread attention in this field. From the point of view of philosophy, Haenni[3] gave an analytical result: He used the method Ψ on the model Ξ and obtained an irrational conclusion Σ. So many experts and scholars analyzed and improved DST via the model and method, respectively. And there are also a few people who simultaneously amend the model and method, but most of the studies are carried out on the basis of the complete discernment framework.

Generally speaking, there are two ways of combination in DST: One is the synchronous combination, the other is the sequential combination. For the first one, suppose that evidential sources and constraints are already known, they can be One-time merged in parallel, and the final decision result is obtained. Because of the one-step process in combining all evidences without the ambiguity of the fusion sequence, the final result is optimal and unique; For the second one, suppose evidential sources and their constraints are serially given, and be combined in turn, a decision result is obtained for every time. Because of the degree of similarity (or degree of conflict) between the different evidential sources fed according to the different sequence of combination and the previous combinational result, the final result is usually suboptimal and not unique.

Based on the past researches, by considering the case of the synchronous combination, we analyzed the cause of conflict of evidence under the complex conditions. A synchronous amendment to the model Ξ and the method Ψ was made in this paper, and a distributive evidence reasoning method was proposed. Its novelty lies in considering the drawback of the past description of the distance between the evidences. That is to say, up to now, all the distances were defined according to the corresponding focal elements between two sources of evidence, and the sequence of the size of the assignments of focal elements about the evidence itself was not considered. The sequence might also lead to conflict, which is referred to as "self-conflict or self-contradiction." According to the similarity constraints among evidential sources, the ISODATA (Iterative Self organizing Data Analysis Techniques Algorithm)[4] is used to cluster all the evidences. After that, the commonality and the occurrence frequency of evidences in the inner class are integrated to use for computing the reliability of the representative of evidences. The different methods of combination are used to combine the inner class and interclass evidences.

2. DST AND ITS CONFLICT ANALYSIS

2.1. Under the Same Discernment Framework

Suppose Θ is a limited, mutually exclusive, and complete space, which is called as the discernment framework. Its power set space is 2^Θ. The BBA under the framework is a mapping function: $m : 2^\Theta \mapsto 1$, which satisfies $m(\phi) = 0$, $\sum_{\theta \in 2^\Theta} m(\theta) = 1$. Here ϕ is an empty set. When $1 > m(\theta) > 0$, all θ are called as focal elements (or propositions). The collection of all focal elements θ is called as the core Γ of evidences. Especially, when there is the only focal element Θ with a nonzero assignment in BBA, that is $m(\Theta) = 1$. According to the BBA's definition in Ref. 1, the following belief function (Bel) and plausibility function (PL) are, respectively, defined as

$$Bel\,(A) = \sum_{B \subseteq A} m\,(B) \tag{1}$$

$$Pl\,(A) = \sum_{B \cap A \neq \phi} m\,(B) \tag{2}$$

If $m_i(\cdot)$, $i = 1, 2, \ldots, n$ are n independent BBAs under the same discernment framework, then the combination rule of Dempster is

$$m\,(A) = \frac{\sum\limits_{\cap A_i^j = A} \prod\limits_{i=1 A_i^j \in 2^\Theta}^{n} m_i\left(A_i^j\right)}{1 - \sum\limits_{\cap A_i^j = \phi} \prod\limits_{i=1 A_i^j \in 2^\Theta}^{n} m_i\left(A_i^j\right)} \tag{3}$$

Here $\sum_{\cap A_i^j = \phi} \prod_{i=1 A_i^j \in 2^\Theta}^{n} m_i(A_i^j)$ is called as the conflict factor κ.

2.2. Under the Different Discernment Framework

Suppose there are n different discernment frameworks, i.e., $\Theta_1, \Theta_2, \ldots, \Theta_n$, their power set spaces are $2^{\Theta_1}, 2^{\Theta_2}, \ldots, 2^{\Theta_n}$, respectively, then, the original evidences can be expressed as

$$\begin{aligned}
S_1 &= \left\{A_1^j | A_1^j \in 2^{\Theta 1}; m_{\Theta 1}(A_1^j) \geq 0,\, j = 1, 2, \ldots, |\Gamma_1|\,\right\} \\
S_2 &= \left\{A_2^k | A_2^k \in 2^{\Theta 2}; m_{\Theta 2}(A_2^k) \geq 0,\, k = 1, 2, \ldots, |\Gamma_2|\right\} \\
&\vdots \qquad \vdots \\
S_n &= \left\{A_n^l | A_n^l \in 2^{\Theta_n}; m_{\Theta_n}(A_n^l) \geq 0,\, l = 1, 2, \ldots, |\Gamma_n|\right\}
\end{aligned}$$

Here, the BBAs under the different discernment framework have the character of normalization, that is

$$\begin{cases}
m(\phi) = 0 \\
\sum\limits_{A \in F1} m_{\Theta_1}\left(A_1^j\right) = 1, \quad \forall A_1^j \in 2^{\Theta_1} \\
\sum\limits_{B \in F_2} m_{\Theta_2}\left(A_2^k\right) = 1, \quad \forall A_2^k \in 2^{\Theta_2} \\
\qquad \vdots \\
\sum\limits_{B \in F_n} m_{\Theta_n}\left(A_n^l\right) = 1, \quad \forall A_n^l \in 2^{\Theta_n}
\end{cases}$$

Therefore, the power set space consisting of n original evidences is called $2^{\Theta} = 2^{\Theta_1 \times \Theta_1 \times \cdots \times \Theta_n}$, the BBA function $m(\cdot)$ of the subset of $\Theta_1, \Theta_2, \ldots, \Theta_n$ in the Cartesian product space $\Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$ is $m_{\Theta_n}(\cdot \cap \cdot \cap \cdots \cap \cdot) = m_{\Theta_1}(\cdot) \times m_{\Theta_2}(\cdot) \times \cdots \times m_{\Theta_n}(\cdot)$. Therefore, the combination rule of Dempster under n different discernment frameworks can be expressed as

$$m(A) = \frac{\sum \prod\limits_{i=1, A_i^j \in 2^{\Theta_i}, \cap A_i^j = A}^{n} m_{\Theta_i}(A_i^j)}{1 - \kappa} \tag{4}$$

Here, $\kappa = \sum \prod_{i=1, A_i^j \in 2^{\Theta_i}, \cap A_i^j = \phi}^{n} m_{\Theta_i}(A_i^j)$ refers to the conflict factor.

The following several examples are used to analyze the counterintuitive phenomenon of Dempster's rule of combination.

Example 1. (One example was given by Zadeh in early stage): One patient sees two different doctors, and they think that the patient might get one of the three disease, i.e.,meningitis (M), concussion (C), and cerebral tumor (T). Therefore, the discernment frame Θ is set to be $\{M, C, T\}$. And then, the diagnosis results of the two doctors can be, respectively, given as follows:

Doctor 1: $m_1(M) = 0.99, m_1(T) = 0.01$;

Doctor 2: $m_2(C) = 0.99, m_2(T) = 0.01$.

As shown in the above example, both of them think that the patient has the least possibility of having the cerebral tumor, while he has the higher possibility of having the other two diseases. Therefore, the combinational result is obtained according to Dempster's rule (3): $m_{1,2}(M) = 0, m_{1,2}(C) = 0, m_{1,2}(T) = 1$. This result means that the possibility of having tumor is 100%, which is a counterintuitive phenomenon, because both doctors think the low possibility of having the cerebral tumor. It is noted that the diagnosis of the other two diseases is fully conflictive.

This is an extreme example of conflicting evidences, which reflects the conflictive problems caused by the Dempster's combination rule when dealing with the highly conflictive evidences.

Example 2.

(a) *Two same uncertain evidences given*:
Doctor 1: $m_1(M) = 0.99, m_1(T) = 0.01$;
Doctor 2: $m_2(M) = 0.99, m_2(T) = 0.01$.
Although the two original evidences are the same, the two doctors are not quite sure about their own diagnosis. According to the definition of the conflict factor κ, here $\kappa = 0.0198 \neq 0$, which means there are still conflicts between them.

(b) *Two same certain evidences given*:
Doctor 1: $m_1(M) = 1, m_1(T) = 0$;
Doctor 2: $m_2(M) = 1, m_2(T) = 0$.
Both doctors are fully sure the patient has the meningitis at the same time. According to the definition of the conflict factor κ, here $\kappa = 0$, which means there is no conflict between them.

(c) *Two opposite evidences given*:
Doctor 1: $m_1(M) = 0, m_1(T) = 1$;
Doctor 2: $m_2(M) = 1, m_2(T) = 0$.

As shown in this example, one doctor is fully sure that the patient has the cerebral tumor, whereas the other is fully sure that the patient has the meningitis. There is not any intersection set between them, so that $\kappa = 1$, which means that the two original evidences are totally conflictive.

For (b) and (c), both of them accord with the intuition, However, for (a), there still exist conflicts between the two same evidences, which does not accord with the intuition.

Here we analyze the cause of conflicts: Three key factors are given by Martin and Osswald in Ref. 5, i.e. the unreliable source of information (like sensor), the unreliable evidences (problems in producing the BBA), and the incomplete discernment framework. However, the cause of conflict in Example 2 has not been revealed quite well from the point of view of Martin and Osswald; the root cause lies in the uncertainty of the evidence itself, which might result in the contradiction and conflict.

Therefore, before combining the highly conflictive evidences by using the classic Dempster combination rule, it is necessary to pay high attention to the cause of the conflict, the degree of the conflict, and the way to reduce the conflict. Since Haenni[3] proposed the guideline of the improvement of the Dempster combination rule, many specialists and scholars have improved the DST through the models and methods, respectively. And there were also a few of people to improve the DST from the two aspects at the same time, but most of them were based on the complete framework of discernment.

(a) *Improvements in the Methods*:
Lefevre et al.[6] believed that the conflict was also a kind of useful information. If the conflict was completely abandoned, this would inevitably lead to the loss of information. So the conflict could be integrated weightedly into the combination rule of unified belief functions. However, Yager's[7] suggestion was contrary to Lefevreet et al.. He believed that all conflicts could not provide any useful information and might be assigned to the unknown item $m(\Theta)$. Martin and Osswald[5]proposed a mixed evidence combination method called discount proportional conflict redistribution by considering the global and partial allocation of conflicts. That is, on the basis of the proportional conflict redistribution, they applied the discounting procedure method to allocate one part of the conflict to some unknown items. Liu[8] considered the causes of conflictive evidence were related not only to the conflict factor k but also to the Pignistic probability distance, which weigh together the conflict degree between the evidence bodies. According to the values of two factors, which results in the conflict, he discussed the classification of conflicts. Wang et al.[9] proposed a new method to combine the conflictive evidences on the basis of Ref. 8. When the conflict is small between two evidences, the combinational result is close to that of DST. When the conflict is greater between the evidences, Wang et al.'s result is more reasonable than that of DST. Deng and co-workers[10] studied the variables characterizing the conflict of evidence and proposed a new correlation coefficient to represent the variables of conflict between evidences based on the partial entropy and mixing entropy, which can quantitatively express the conflict between the evidences. When it is close to 1, the conflict between the evidences is very small, and when close to 0 the conflict is very great. Quan et al.[11] studied from the perspective

of selecting the effective parameter to measure the conflict and proposed a new partial conflict redistribution rule. He and Hu[12] proposed a modified Dempster–Shafer (D–S) evidential combination strategy to solve the problem of the integration of conflictive evidences, which has good convergence and reliability. Yang and Xu[13] established an unified evidential reasoning (ER) rule to handle multiple independent sources of evidence and proved that when each source of evidence is completely reliable, the Dempster rule is a special case of ER rules.

(b) *Improvements in the model*:
Smets[14] analyzed the nature of the evidence combination and proposed the transferable belief model (TBM) based on the nonprobabilistic model. Through a deep comparison with Dempster's model based on probability model, the range of application of the normalized and conjunctive operations was analyzed. Smets pointed out that the existence of a conflict in TBM was not a real problem; a number of fundamental issues of combination of conflictive evidences were further clarified by comparing the different methods to eliminate the conflicts. Han et al.[15] proposed a sequential combination method of weighted evidence based on the evidence variance aiming at the combination of conflictive evidences. First, the definition of evidence variance was given based on the Jousselme distance. For every combinational step, according to the weight produced by the past sequential variance of combination result of all previous evidences and the current sequential variance when the new evidence joins in, the current evidence and last combinational result were amended. After that, they were combined through Dempster's rule. This method must satisfy the premise of sequential combination before its application. Huand colleagues[16] proposed a representation model of the conflict coefficient according to the fact that the conflict factor in the classical theory of evidence cannot reasonably measure the degree of conflict between the evidences and concluded the causes of conflict. Dezert and Smarandache[17] in 2003 proposed Dezert–Smarandache theory based on DS evidence theory and Bayesian probability theory. In its framework, Θ was justly considered n complete propositions containing $\{\theta_1, \theta_2, \ldots, \theta_n\}$. Its superpower set space is built on the lattice model D^{Θ} of Dedekind. There are not any other constraints (mutually exclusive or nonexistence of constraints) between the propositions. Xionget al.[18] on the basis of the distance between the focal elements of the evidence computed the support of every evidence. After a weighted average of all the evidence supports, they obtained a "reference evidence," justifying and amending these initial evidences according to it. And then, combined them with Dempster's rule.

(c) *Simultaneous amendment in the model and method*:
Lu and Qin[19] proposed a method to revise the combinational rules and evidential sources simultaneously and also proposed a general framework for the combination of evidences. Under this framework, first, the two grades of reliability of the evidential sources were processed and then these evidential sources are combined. Finally, the conflict factor is assigned to each focal element according to the weights based on Pignistic probability distance, and a reasonable result of amendment was achieved. Li and Guo[20] proposed a new combinational method based on the credibility of evidence. While the conflict of evidence was assigned, the evidential model was amended. Aiming at the problem of the combination operation "and" for consistent evidences and the proportional assignments of evidence conflict to different proposition, the credibility of every evidence was effectively utilized. Wang et al.[21] proposed a hybrid combination of evidence based on the fuzzy clustering analysis. The method utilizes Pignistic probability distance to construct the fuzzy similar matrix and the validity index of evidence clustering. The transitive closure is used to cluster evidences and compute the credibility. After the amendment of the evidences, they in the same clustering are combined by using the D–S rule, and they in different clustering are combined by using the unified belief function combination method. Quanet al.[22] gave the definition of attribute support of evidence, the classification threshold, and the credibility of the evidence and classified the evidence to ensure the evidence in the same category with good consistency. On the basis of subtriangular norm operator and discount factor analysis, Wangand co-workers[23] used the average distance based on

the evidence triangular norm operator and the conflict factor to classify evidences into three categories, i.e. the credible, nonconflictive, and conflictive evidences and assigned the discount factor 1 to the credible and nonconflictive evidences. And then, they defined the evidential weights according to distance between evidences and got the amended evidence body by combining conflictive evidences according to the weights, to reduce the conflict of evidences and finally complete the combination by using the Dempster rule.

In summary, whatever the amendment of the rules or the model, even if on some special occasions, some effects to combine highly conflictive evidences are achieved. Generally speaking, it is still difficult to solve the problem of conflict combination using the Dempster rule. The strategy of simultaneous amendment of both rules and model to solve this kind of problem has more advantages. However, the current method of classification and clustering of conflictive evidences and amendment of combination rules always has more or less problems, for example, the distance measure between conflictive evidences, clustering analysis and evaluation, the intelligent regulation for weights, and the factors that affect the reliability calculation and so on.

3. EVIDENCE SIMILARITY CHARACTERIZATION

The conflict of evidence may be reflected in the distance between two evidences. It is not a new topic to measure the distance between evidences; there are many experts and scholars engaged in this research field.[24–26] Currently, these distances, i.e., Pignistic distance,[8,27] cosine distance, and Jousselme distance[25] are widely applied. But they justly depend on the corresponding focal elements (or the relevant focal element set) between two sources of evidence to describe and characterize without considering the order of size of the assignment of each focal element in an evidence, which might lead to"self-conflict or self-contradiction." To consider the conflict produced by evidence itself, here a new evidence similarity measure is defined between two evidential sources according to the order of size of the assignment. Prior to this, to give this new similarity measure, first we define the order correlation coefficient between two sets of data.

DEFINITION 1. *Given two sets of data* $\{x_1, x_2, \cdots, x_n\}$,$\{y_1, y_2, \cdots, y_n\}$, *here* $x_1, x_2, \ldots, x_n$ *and* $y_1, y_2, \ldots, y_n$ *are in an ascending order. After sorting, two sets of data are* $x_{p_1}, x_{p_2}, \ldots, x_{p_n}$ *and* $y_{q_1}, y_{q_2}, \ldots, y_{q_n}$, *respectively, meet* $x_{p_1} \leq x_{p_2} \leq \cdots \leq x_{p_n}$ *and* $y_{q_1} \leq y_{q_2} \leq \cdots \leq y_{q_n}$, *for each* p_i, *index its position from* $q_1, q_2, \ldots, q_n$, *assuming it is* q_j, *that is,* $q_j = p_i$. *Note that* $j = f(i)$, *the correlation coefficient is*

$$\mu = \frac{\sum_{i=1}^{n} (i - j)^2}{\sum_{i=1}^{n} [n - (i - 1) - i]^2} \tag{5}$$

It satisfies $0 \le \mu \le 1$. *When* $\mu = 0$, *the convergence of two sets of data is largest; when* $\mu = 1$, *it is reversed.*

3.1. The Consistency Focal Elements between the Two Evidences

DEFINITION 2. *For any two sources of evidence, i.e.,* S_1, S_2, $m_1(\cdot)$, *and* $m_2(\cdot)$ *are the basic belief assignments over n focal elements in the discernment framework* Θ *(note that the assignment is unequal with each other), where* X_i, Y_i *are the serial number in accordance with the order of size of the assignment and the subscript* i *denotes the* ith *proposition subset. The similarity function of evidence to characterize the order of the size of the assignments over subsets as follows:*

$$Sim_{seq}(m_1, m_2) = 1 - \frac{\sum_{i=1}^{n} (X_i - Y_i)^2}{\sum_{i=1}^{n} (n + 1 - 2i)^2} \tag{6}$$

As we know, if there is a similarity function $Sim(m_i, m_j)$, *which is the characterization of distance between any two evidence sources, then the following three basic conditions must be satisfied:*

(1) symmetry: $\forall\, m_i(\cdot), m_j(\cdot)$, $Sim(m_i, m_j) = Sim(m_j, m_i)$,
(2) consistency: $\forall\, m(\cdot)$, $Sim(m, m) = 1$, *and*
(3) nonnegative: $\forall\, m_i(\cdot), m_j(\cdot)$, $1 \ge Sim(m_i, m_j) \ge 0$.
So it may prove that $Sim_{seq}(m_i, m_j)$ *is a similarity measure function, which as follows.*

Proof.

(1) Symmetry: $\forall\, m_i(\cdot), m_j(\cdot)$, $Sim_{seq}(m_i, m_j) = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(n+1-2i)^2} = 1 - \frac{\sum_{i=1}^{n}(Y_i - X_i)^2}{\sum_{i=1}^{n}(n+1-2i)^2} = Sim_{seq}(m_j, m_i)$. So the conclusion is proved.
(2) Consistency: $\forall\, m(\cdot)$, $Sim_{seq}(m, m) = 1 - \frac{\sum_{i=1}^{n}(X_i - X_i)^2}{\sum_{i=1}^{n}(n+1-2i)^2} = 1 - 0 = 1$. So the conclusion is proved.
(3) Nonnegative: $\forall\, m_i(\cdot), m_j(\cdot)$, $Sim_{seq}(m_i, m_j) = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(n+1-2i)^2}$. To prove $0 \le Sim_{seq}(m_i, m_j) \le 1$, so just prove $0 \le \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(n+1-2i)^2} \le 1$. For $\sum_{i=1}^{n}(n + 1 - 2i)^2 = \sum_{i=1}^{n}(X_i - Y_{n-i+1})^2$, just need to prove $\sum_{i=1}^{n}(X_i - Y_i)^2 \le \sum_{i=1}^{n}(X_i - Y_{n-i+1})^2$. Also, since $\sum_{i=1}^{n}(X_i - Y_i)^2 = \sum_{i=1}^{n} X_i^2 - 2X_iY_i + Y_i^2$, $\sum_{i=1}^{n}(X_i - Y_{n-i+1})^2 = \sum_{i=1}^{n} X_i^2 + Y_{n-i+1}^2 - 2X_iY_{n-i+1}$, $\sum_{i=1}^{n} X_i^2 + Y_i^2 = \sum_{i=1}^{n} X_i^2 + Y_{n-i+1}^2$, so just need to prove $\sum_{i=1}^{n} X_iY_i \ge \sum_{i=1}^{n} X_iY_{n-i+1}$, which has been proved in the literature,[28] will be not repeated here. So $1 \ge Sim(m_i, m_j) \ge 0$. □

DEFINITION 3. *For any two sources of evidence, i.e.* S_1, S_2, $m_1(\cdot)$, *and* $m_2(\cdot)$ *are the basic belief assignments over n focal elements in the discernment framework* Θ *(note that the BBA of each subproposition might be same. Assuming that* s_1 *subpropositions' BBAs are same in* $m_1(\cdot)$, *and* s_2 *subpropositions' BBAs are same in* $m_2(\cdot)$.

Wherein, X_i, Y_i are the serial number according to the order of the size of subproposition' BBAs; the subscript i indicates the ith subproposition. Due to the BBAs of some sub-propositions are same. For the evidence S_1, there might be s_1 kinds of sorts. For S_2, there might be s_2 kinds of sorts. Therefore, there are $s_1 \times s_2$ kinds of sorts for S_1 and S_2. The similarity measure function is redefined in this case as follows:

$$Sim'_{seq}(m_1, m_2) = 1 - \frac{\sum_{t=1}^{s_1 s_2} \sum_{i=1}^{n} \left(X^t{}_i - Y^t{}_i\right)^2}{s_1 s_2 \left(\sum_{i=1}^{n} (n + 1 - 2i)^2 \right)} \tag{7}$$

Similarly, it is easy to prove that $Sim'_{seq}(m_X, m_Y)$ is a similarity measure function.

To consider the influence of the distance of evidence, here based on the Pignistic distance (or the similarity measure), the similarity measure is improved further by considering the BBAs' sequence. At first, we will introduce the Pignistic probability distance in the following Definitions 4 and 5.

DEFINITION 4.[27] *Let m be the BBA on the discernment framework Θ, and A, $B \subseteq \Theta$, and satisfies*

$$\underset{A \subseteq \Theta}{Bet}\, P_m(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} m(B) \tag{8}$$

wherein $|B|$ represents the cardinality of the subset B and $Bet_{A \subseteq \Theta}\, P_m(A)$ describes the sum of all the probabilities that m supports the proposition subsets A in power set of 2^{Θ} to be true.

DEFINITION 5.[27] *Suppose that m_i and m_j are BBAs over Θ. $Bet_{A \subseteq \Theta}\, P_{m_i}$ and $Bet_{A \subseteq \Theta}\, P_{m_j}$ are the probability functions of the corresponding subproposition after the Pignistic transformation. So the Pignistic probability distance between m_i and m_j is defined as follows:*

$$DifBetP^{m_i}_{m_j} = \max_{A \subseteq \Theta} \left| Bet\, P_{m_i}(A) - Bet\, P_{m_j}(A) \right| \tag{9}$$

Obviously, $0 \leq DifBetP^{m_i}_{m_j} \leq 1$, it means, the larger Pignistic probability distance between two pieces of evidences is, the greater their difference is, the less their similarity is, and vice versa.

Then, the improved similarity measure is

$$Sim(m_i, m_j) = \left[1 - DifBetP^{m_i}_{m_j}\right] Sim_{seq}(m_i, m_j) \tag{10}$$

Here it is easy to prove that the improved measure function $Sim(m_i, m_j)$ is still a similarity measure function; this is because the product between two similarity measure functions still meets the definition of similarity measure function.

3.2. The Inconsistent Focal Elements between Two Evidences

Suppose that there are two discernment frameworks Θ_1 and Θ_2 and $\Theta_1 \neq \Theta_2$, their power sets are 2^{Θ_1} and 2^{Θ_2}, respectively, and $2^{\Theta_1} \neq 2^{\Theta_2}$, and Γ is called as the kernel of evidences.

3.2.1. *Only Singleton Focal Elements*

Suppose there are two independent evidences $\{\Theta_i, \Gamma_i, m_i\}$, $i = 1, 2$. Herein, m is the basic belief assignment $m(A_i^j)$ on the $\Theta_i, i = 1, 2$, $A_i^j \in \Gamma_i, j = 1, 2, \ldots, |\Gamma_i|$, $\Gamma_1 \neq \Gamma_2$, that is, $\Gamma_1 = \{A_1^1, A_1^2, A_1^3, \ldots, A_1^{|\Gamma_1|}\}$, $\Gamma_2 = \{A_2^1, A_2^2, A_2^3, \ldots A_2^{|\Gamma_2|}\}$; obviously, their kernel spaces are inconsistent. Then, their composite kernel space is $\Gamma = \Gamma_1 \cup \Gamma_2$; the overlap space between the two kernel spaces is $\Gamma_1 \cap \Gamma_2$. Suppose $m_{\bar{i}}(A_i^j) = 0$, $A_i^j \in \Gamma_i - \Gamma_1 \cap \Gamma_2$, $\bar{i}$ is the complementary set of i. According to the formula (10), the similarity measure between the two sources of evidence will be solved.

3.2.2. *Disjunctive Focal Elements*

Suppose there are two independent evidences $\{\Theta_i, \Gamma_i, m_i\}, i = 1, 2$. Herein, m is the basic belief assignment $m(A_i^j)$ on the 2^{Θ_i}, $i = 1, 2$, $A_i^j \in \Gamma_i$, $j = 1, 2, \cdots |\Gamma_i|$, $\Gamma_1 \neq \Gamma_2$.

According to Equation 8, we can obtain the Pignistic probabilities of all subpropositions in the different Θ, i.e.,$Bet_{A_i^j \in \Gamma_i} P_m(A_i^j)$. Suppose that $Bet_{A \in \Theta} P_{\bar{i}}(A_i^j) = 0$, $A_i^j \in \Gamma_i - \Gamma_1 \cap \Gamma_2$, $\bar{i}$ is the complementary set of i, replace $m(\cdot)$ with $Bet\, P_m(\cdot)$. According to Equation 10, the similarity measure between the two evidence sources is easily computed.

4. THE AMENDMENT OF EVIDENTIAL SOURCES

Suppose that there are n independent evidence sources, according to the similarity measure function (10), the similarity measure matrix $C_{n\times n}$ is obtained as follows:

$$C_{n\times n} = \begin{bmatrix} 1 & sim_{12} & sim_{13} & \cdots & sim_{1n} \\ sim_{21} & 1 & sim_{23} & \cdots & sim_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ sim_{n1} & sim_{n2} & sim_{n3} & \cdots & 1 \end{bmatrix} \tag{11}$$

Note that here the elements in matrix (11), i.e., $Sim_{ij} = Sim(m_i, m_j)$ meets $Sim_{ij} = Sim_{ji}$. So the similarity measure matrix $C_{n\times n}$ is a symmetric matrix.

For each evidential source, there are the degrees supported by other evidences. As we know, the high similarity between the two evidence sources shows that a high degree of mutual support between evidence sources. So, according to $C_{n\times n}$, the commonality of evidence body $Crd(m_i)$ is computed as follows:

$$Crd(m_i) = \frac{Sup(m_i)}{\max_{1\le j\le n}[Sup(m_j)]} \tag{12}$$

Here, $Sup(\mathrm{m_i}) = \sum_{\mathrm{j=1,j\neq i}}^{\mathrm{n}} Sim(m_i, m_j)$, $i = 1, \ldots, n$. According to the commonality of evidence, the initial source of evidence is amended as follows:

$$\begin{cases} m_i'(A) = Crd(m_i)\cdot m_i(A), \forall A \subseteq \Theta \\ m_i{}'(\Theta) = Crd(m_i)\cdot m_i(\Theta) + 1 - Crd(m_i) \end{cases} \tag{13}$$

5. THE CLUSTERING AND COMBINATION OF EVIDENCES

In the synchronous combination, evidences own the characters of clustering. That is to say, evidences with similar opinions are near with each other, their distances between them (conflict) are small, and they are easy to be clustered into one group. On the contrary, evidences with the conflicting opinions are far with each other, their distances (conflicts) are large, and they are not easy to be clustered into one group. If these evidences can be clustered into a class of evidence, the classical Dempster rule can be used directly, because their conflicts between them are small, so that the defects of the Dempster rule are not very prominent. But if the conflict between the clusterings is large, it is not fit to use the classical Dempster rules directly. In this case, the Dempster rule needs to be improved further.

5.1. Evidence Clustering with ISODATA

The ISODATA clustering method is a dynamic and unsupervised-learning clustering algorithm widely used, which is in use of combining and splitting mechanism. The final clustering results are obtained[4] by the unceasing iteration.

When the distances are calculated in the classical ISODATA algorithm, i.e., the distance between the samples and the clustering center and the distance between the clustering centers, all use the Euclidean distance. However, when the evidences are clustered, the correlation of basic belief assignments within the evidence is also considered. So according to (10), the distance between the evidences is defined as follows:

$$D(x, y) = 1 - \left[1 - DifBetP_y^x\right]\cdot Sim_{seq}(x, y) \tag{14}$$

Then, replace the Euclidean distance in the classic ISODATA with the new one in (14), the clustering steps are introduced as follows:

Suppose that each evidence contains υ focal elements, the center of evidences for each category of sources is M_k, which is determined by the average of all evidences N_k (the number of evidential sources) in each clustering. The initial parameters are set as follows: the expected number of categories k, the standard deviation limitation θ_s for the same category of evidential sources, and the minimum distance limitation θ_c for all clustering centers.

Note that if the kernel spaces between two sources are different, or the disjunctive focal elements are involved, we may replace the basic probability $m(\cdot)$ with the Pignistic probability $Bet\ P$.

Step 1: Assign n evidential sources to the nearest clustering center M_k, which can be revised as follows:

$$M_k(f) = \frac{1}{N_k}\sum_{i=1}^{N_k} m_i(f),\ f = f_1, f_2, \cdots, f_\upsilon \tag{15}$$

Step 2: If the iteration is not completed, then determine whether to split in step 3 or merge in step 4 through the number of the existing clustering centers and iterations. If the iteration is completed, go to step 4 directly.

Step 3: Calculating the standard deviation vector of the distance among evidences in each cluster.

$$\sigma_k = (\sigma_{1k}, \sigma_{2k}, \sigma_{3k} \cdots \sigma_{\upsilon k})^T \tag{16}$$

Herein each component of the vector is

$$\sigma_{fk} = \sqrt{\frac{1}{N_k}\sum_{m_i \in M_k} (m_i(f) - M_k(f))^2} \tag{17}$$

where k is the kth clustering of evidence with the center M_k and the number of evidences in clustering k is N_k.

If the largest component of the ith class is greater than the standard deviation upper limitation θ_s of the same category of evidential sources and meets the condition of the splitting, and then, split the ith class and returns step 1. If the condition is not satisfied, then go to step 4.

Step 4: To calculate the distances among all the clustering centers

$$D_{ij} = 1 - \left[1 - DifBetP_{M_j}^{M_i}\right] \cdot Sim_{seq}(M_i, M_j) \tag{18}$$

Let these D_{ij} which is less than the lower limitation θ_c sort in an ascending order, namely

$$D = \{D_{i_1 j_1}, D_{i_2 j_2}, \ldots, D_{i_L j_L}\} \tag{19}$$

and merge any two clustering centers M_{i_k} and M_{j_k}, which are related to $D_{i_k j_k}$ in D_{ij}.

Step 5: Justify whether the iteration is completed. If the iteration is not completed, then return step 1, otherwise the clustering is over. The flow chart is shown in Figure 1.

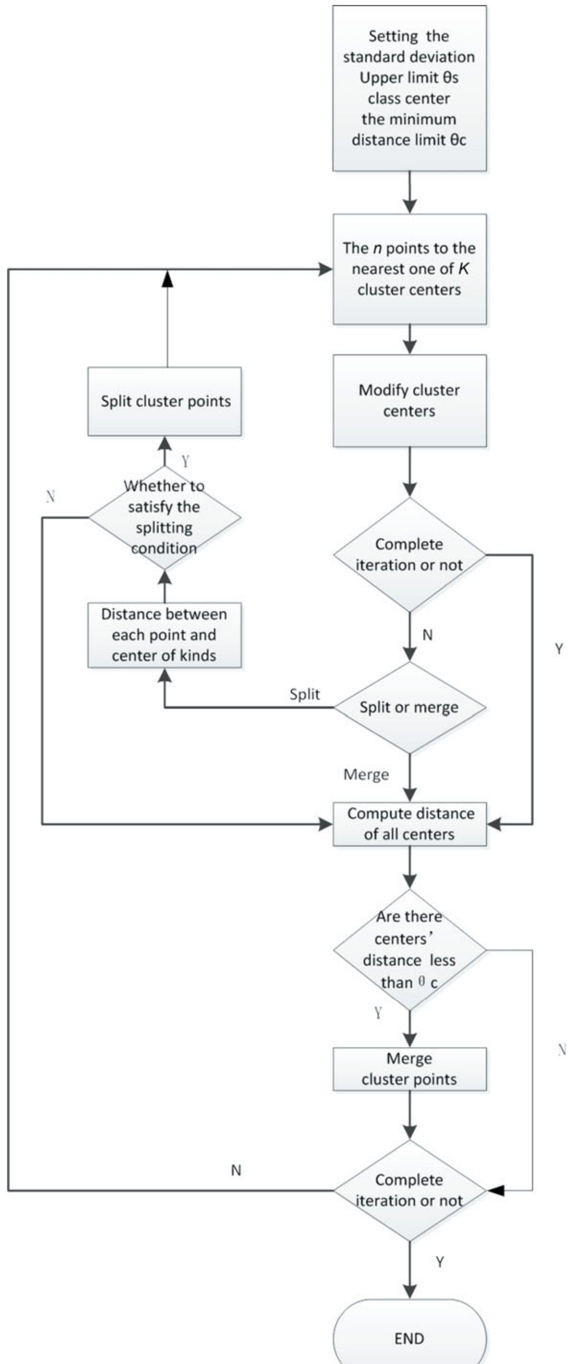

Figure 1. The follow chart of ISODATA.

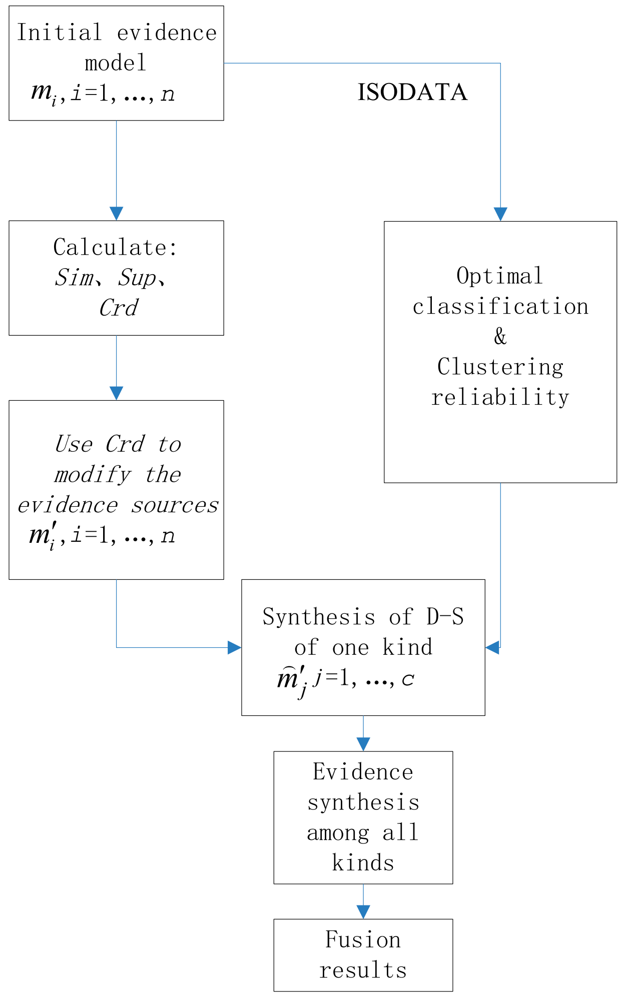

Figure 2. The follow chart of evidence synthesis.

5.2. The Validity of Clustering

To ensure the validity of clustering, it is very important to give an effective evaluation criterion.

Suppose n initial evidences are divided into class c, there are n_j evidences in class j. The clustering criterion function, i.e., the square sum of error J_c is defined as

$$J_c = \sum_{j=1}^{c} \sum_{i=1}^{n_j} \| m_{ij} - \bar{m}_j \|^2 \tag{20}$$

Table I. Given evidences to be clustered

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}
A	0.512	0.588	0.221	0.299	0.488	0.612	0.189	0.301	0.513	0.587	0.212	0.298
B	0.287	0.191	0.198	0.198	0.288	0.192	0.713	0.202	0.288	0.192	0.702	0.199
C	0.201	0.221	0.503	0.503	0.224	0.196	0.098	0.497	0.199	0.221	0.086	0.503

The criterion function of sum of among class distances is defined as

$$J_b = \sum_{j=1}^{c} (\bar{m}_j - \bar{m})^T (\bar{m}_j - \bar{m}) \tag{21}$$

where m_{ij} is the evidence i of the class j, $\bar{m}_j$ is the mean of all evidences in class j, namely

$$\bar{m}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} m_{ij},\ j = 1, 2, \ldots c \tag{22}$$

$\bar{m}$ is the mean of all the samples, namely

$$\bar{m} = \frac{1}{n} \sum_{k=1}^{n} m_k \tag{23}$$

J_c represents the overall error square sum when n evidences are divided into c classes. The smaller it is, the better it is. J_b represents the sum of among class distance. The bigger it is, the better it is.

The two indicators, i.e. J_b and J_c have a direct influence on the clustering results. Here a numerical example (Example 3) is given to compare the clustering effect between the indicators proposed here and mentioned in Ref. 21.

Example 3. By using the transitive closure clustering method and the evaluation indicator,[21] the 12 evidential sources in Table I are clustered into three categories, that is, $\{m_1, m_2, m_5, m_6, m_9, m_{10}\}$, $\{m_3, m_7, m_{11}\}$, and $\{m_4, m_8, m_{12}\}$. According to Equations 20 and 21, we obtained $J_c = 0.0290$ and $J_b = 0.3170$. However, by using the ISODATA method and evaluation indicator, they are clustered into four categories: $\{m_1, m_5, m_9\}$, $\{m_2, m_6, m_{10}\}$, $\{m_3, m_7, m_{11}\}$, and $\{m_4, m_8, m_{12}\}$. According to Equations 20 and (21), we obtained $J_c = 0.0027$ and $J_b = 0.3608$. Since the smaller J_b is the better it is, and the bigger J_b is the better it is. Obviously, the clustering effect of the method presented in this paper is better than that in Ref. 21.

5.3. The Reliability of Clustering

After n initial evidences are divided into c categories, to calculate the reliability of each category of evidences, a new method to compute reliability is proposed here. That is to say, the commonality Crd of each evidence and the occurrence frequency (the ratio between the number of evidences in this category and the number of all evidences) of each category of evidences are considered. Where, the larger the commonality Crd of each evidence in one category is, or the more the evidences are involved, the higher the reliability of the category is. Vice versa. The reliability of clustering weight is defined as follows:

$$\gamma_j = \frac{1}{2} \cdot \frac{\sum_{j=1}^{n_j} Crd_j}{n_j} + \frac{1}{2} \cdot \frac{n_j}{n} \tag{24}$$

According to γ_j, use the following formula (25) to calculate the reliability of the jth category of evidences:

$$\omega_j = \frac{e^{-\beta \cdot \gamma_j}}{\sum_{i=1}^{c} e^{-\beta \cdot \gamma_i}} \tag{25}$$

where $i, j = 1, 2, \ldots, c$, β is the parameter of the negative exponent function. The reliability of clustering will be different according to the different value of β. The value of β is selected according to the experience. Note that $\beta = -5$ in the examples of this paper.

5.4. Combination of Evidences

Suppose there are n discernment frameworks, i.e., $\Theta_1, \Theta_2, \ldots, \Theta_n$, $m_i, i = 1, 2, \ldots, n$ are their basic belief assignments of n evidential sources. First, the evidences are preprocessed and the model $m_i', i = 1, 2, \cdots, n$ are obtained. Second, these evidences are clustered (here suppose the number of clustering is c). As we know, the evidences in one category are relatively similar and have less conflict whereas the evidences in different categories have high conflicts. All evidences in each category are combined according to Dempster rule (3) or (4) (called inner class combination). Therefore, we can obtain c centers of categories, $\widehat{m}'_j$, $j = 1, 2, \ldots, c$, which are called evidential representative. Generally speaking, the evidential representatives have high conflicts, if we directly use the Dempster rule to combine them again; the disadvantage of the Dempster rule in dealing with high conflict becomes very prominent. To overcome the issue, the Murphy method[32] is used to combine these evidential representatives $\widehat{m}'_j$, $j = 1, 2, \ldots, c$, which only amends the evidential model without changing the Dempster combination rule. That is to say, after the basic belief assignments of all evidences are averaged, and the mean evidence is combined with itself for $n - 1$ times according to the Dempster rule.

In this paper, the steps of the combination for the evidential representatives are listed as follows:

(1) The evidential representatives are weightedly averaged according to the reliability of the clustering.

$$\bar{m}(A) = \sum_{i=1}^{c} \omega_i \widehat{m}_i'(A) \tag{26}$$

(2) The evidence $\bar{m}(A)$ after the weighted average is combined for $c - 1$ times according to Dempster rule (3) or (4).

In short, the flow chart of evidential combination is shown in Figure 2, the main steps are given as follows:

Step 1: Calculate the similarity Sim based on the correlation coefficient of n evidences $m_i, i = 1, 2, \ldots, n$ and the commonality Crd of each evidence;
Step 2: amend the original evidences based on the commonality Crd and get the new evidential model $m_i', i = 1, 2, \ldots, n$;
Step 3: cluster the original evidences with the ISODATA method and calculate the reliability ω_j of each category;
Step 4: combine the evidences from the same category by using the Dempster rule, and obtain the new evidential representatives $\widehat{m}_j', j = 1, 2, \ldots, c$;
Step 5: combine the evidential representatives $\widehat{m}_j', j = 1, 2, \ldots, c$ by using Murphy's method.

6. EXPERIMENT

To testify the advantage of our method, three numerical examples (there is one conflictive source, one conflictive and neutral sources, two conflictive sources) are given in this paper to compare with other methods.

6.1. One Conflictive Evidence Involved

Example 4. To validate the effectiveness of the proposed method for a one conflictive evidence involved, two to six evidences are used, respectively, for combination in an ascending order. Our method is compared with other methods, i.e., D-S method, Yager method, Dong Wang method, Lianfeng Wang method, and so on.

Suppose in the discernment framework $\Theta = \{A, B, C\}$, there are six independent evidences $\{\Theta, \Gamma, m_i\}, i = 1, 2, \ldots, 6$ in a system:

$\Gamma_1 = \{\{A\}, \{B\}, \{C\}\},\ m_1 = \{0.60, 0.10, 0.30\}$
$\Gamma_2 = \{\{A\}, \{B\}, \{C\}\},\ m_2 = \{0.55, 0.1, 0.35\}$
$\Gamma_3 = \{\{A\}, \{B\}, \{C\}\},\ m_3 = \{0.00, 0.90, 0.10\}$
$\Gamma_4 = \{\{A\}, \{B\}, \{C\}\},\ m_4 = \{0.55, 0.10, 0.35\}$
$\Gamma_5 = \{\{A\}, \{B\}, \{C\}\},\ m_5 = \{0.55, 0.10, 0.35\}$
$\Gamma_6 = \{\{A\}, \{B\}, \{C\}\},\ m_6 = \{0.55, 0.1, 0.35\}$

Obviously, m_3 has conflict with other sources. When the ISODATA algorithm is used for clustering evidential sources $\{\Theta, \Gamma, m_i\}, i = 1, 2, \ldots, 6$, the parameters are set as follows: The number of initial evidences is used as the expected classification

Table II. The comparison of different methods when one conflict evidence is involved

Combination rules	m_1, m_2	m_1, m_2, m_3	m_1, m_2, m_3, m_4	m_1, m_2, m_3, m_4, m_5	$m_1, m_2, m_3, m_4, m_5, m_6$
D-S[1]	m(A) = 0.7416 m(B) = 0.0224 m(C) = 0.2360	m(A) = 0.0000 m(B) = 0.4615 m(C) = 0.5385	m(A) = 0.0000 m(B) = 0.1967 m(C) = 0.8033	m(A) = 0.0000 m(B) = 0.0654 m(C) = 0.9346	m(A) = 0.0000 m(B) = 0.0196 m(C) = 0.9804
Yager[7]	m(A) = 0.3300 m(B) = 0.0100 m(C) = 0.1050 m(Θ) = 0.5550	m(A) = 0.0000 m(B) = 0.0009 m(C) = 0.0105 m(Θ) = 0.9805	m(A) = 0.0000 m(B) = 0.0009 m(C) = 0.0037 m(Θ) = 0.9954	m(A) = 0.0000 m(B) = 0.0001 m(C) = 0.0013 m(Θ) = 0.9986	m(A) = 0.0000 m(B) = 0.0000 m(C) = 0.0005 m(Θ) = 0.9995
Quan Sun[29]	m(A) = 0.5132 m(B) = 0.0419 m(C) = 0.2085 m(Θ) = 0.2364	m(A) = 0.1740 m(B) = 0.1755 m(C) = 0.1240 m(Θ) = 0.5265	m(A) = 0.2064 m(B) = 0.1466 m(C) = 0.1373 m(Θ) = 0.5097	m(A) = 0.2262 m(B) = 0.1308 m(C) = 0.1471 m(Θ) = 0.4960	m(A) = 0.2396 m(B) = 0.1198 m(C) = 0.1545 m(Θ) = 0.4860
Zhengcai Lu[19]	m(A) = 0.6491 m(B) = 0.0655 m(C) = 0.2854 m(Θ) = 0.0000	m(A) = 0.5255 m(B) = 0.0886 m(C) = 0.2376 m(Θ) = 0.1483	m(A) = 0.5177 m(B) = 0.0966 m(C) = 0.2528 m(Θ) = 0.1329	m(A) = 0.5123 m(B) = 0.1023 m(C) = 0.2731 m(Θ) = 0.1123	m(A) = 0.5104 m(B) = 0.1048 m(C) = 0.2896 m(Θ) = 0.0952
Yanming Xiong[18]	m(A) = 0.5036 m(B) = 0.1384 m(C) = 0.3140 m(Θ) = 0.0440	m(A) = 0.4928 m(B) = 0.1531 m(C) = 0.3110 m(Θ) = 0.0430	m(A) = 0.6511 m(B) = 0.0576 m(C) = 0.2861 m(Θ) = 0.0051	m(A) = 0.7561 m(B) = 0.0183 m(C) = 0.2250 m(Θ) = 0.0006	m(A) = 0.8287 m(B) = 0.0053 m(C) = 0.1659 m(Θ) = 0.0001
Dong Wang[9]	m(A) = 0.4225 m(B) = 0.0285 m(C) = 0.1614 m(Θ) = 0.3876	m(A) = 0.1953 m(B) = 0.4773 m(C) = 0.1279 m(Θ) = 0.1995	m(A) = 0.3427 m(B) = 0.2029 m(C) = 0.2044 m(Θ) = 0.2500	m(A) = 0.4772 m(B) = 0.1129 m(C) = 0.2638 m(Θ) = 0.1460	m(A) = 0.4936 m(B) = 0.0638 m(C) = 0.2425 m(Θ) = 0.2001
Wenli Li[20]	m(A) = 0.6391 m(B) = 0.0658 m(C) = 0.2851 m(Θ) = 0.0100	m(A) = 0.4980 m(B) = 0.1456 m(C) = 0.2961 m(Θ) = 0.0703	m(A) = 0.6383 m(B) = 0.0704 m(C) = 0.2755 m(Θ) = 0.0158	m(A) = 0.6944 m(B) = 0.0528 m(C) = 0.2476 m(Θ) = 0.0042	m(A) = 0.7352 m(B) = 0.0362 m(C) = 0.2275 m(Θ) = 0.0011
Deqiang Han[15]	m(A) = 0.7409 m(B) = 0.0224 m(C) = 0.2367	m(A) = 0.6129 m(B) = 0.2925 m(C) = 0.0946	m(A) = 0.7936 m(B) = 0.0903 m(C) = 0.1161	m(A) = 0.8776 m(B) = 0.0176 m(C) = 0.1048	m(A) = 0.9034 m(B) = 0.0060 m(C) = 0.0905
Lianfeng Wang[21]	m(A) = 0.6428 m(B) = 0.0690 m(C) = 0.2882	m(A) = 0.6021 m(B) = 0.1095 m(C) = 0.2884	m(A) = 0.6576 m(B) = 0.0650 m(C) = 0.2773	m(A) = 0.6846 m(B) = 0.0506 m(C) = 0.2648	Unclassified Classified m(A) = 0.71 m(A) = 0.81 m(B) = 0.03 m(B) = 0.01 m(C) = 0.26 m(C) = 0.18
Our method (pignistic distance)	m(A) = 0.7416 m(B) = 0.0224 m(C) = 0.2360 m(Θ) = 0.0000	m(A) = 0.8699 m(B) = 0.0108 m(C) = 0.1150 m(Θ) = 0.0042	m(A) = 0.9420 m(B) = 0.0020 m(C) = 0.0545 m(Θ) = 0.0015	m(A) = 0.9739 m(B) = 0.0007 m(C) = 0.0246 m(Θ) = 0.0008	m(A) = 0.9879 m(B) = 0.0004 m(C) = 0.0112 m(Θ) = 0.0005
Our method	m(A) = 0.7416 m(B) = 0.0224 m(C) = 0.2360 m(Θ) = 0.0000	m(A) = 0.8882 m(B) = 0.0031 m(C) = 0.1068 m(Θ) = 0.0019	m(A) = 0.9479 m(B) = 0.0004 m(C) = 0.0509 m(Θ) = 0.0008	m(A) = 0.9768 m(B) = 0.0001 m(C) = 0.0227 m(Θ) = 0.0004	m(A) = 0.9896 m(B) = 0.0000 m(C) = 0.0101 m(Θ) = 0.0003

number, the upper limitation of standard deviation is 0.05, and the lower limitation of minimum distance is 0.05. All the evidential sources are divided into two categories, i.e., m_3 is regarded separately as a category, the others are regarded as a category.

As shown in Table II, the results from D–S and Yager methods suddenly become 0, when the evidence m_3 appears, and then the result always keeps 0. Obviously, it is an excessive reaction with the "one ticket veto" phenomenon. Therefore, D–S and Yager methods are not suitable for combining high conflictive evidences. The Sun Quan method also overreacts on conflicts, which has the same disadvantage as D–S and Yager methods. The combinational result of six evidential sources is unsatisfactory even if the rest evidential sources are true (positive), and the convergence speed is very slow. The Lu Zhengcai method has a relatively normal reaction on the conflict, but the result cannot converge with the further combination of positive evidences; and it is obviously abnormal. The Xiong Yanming method gets a reasonable result when the conflict occurs, but the mass value in combining six evidential sources is low. The Wang Dong method reacts violently with conflicts. Although the combinational result converges to the true proposition A, but it converges slowly. For the methods, i.e. Li Wenli, Han Deqiang, and Wang Lianfeng, they all preprocess the evidences, and the reaction on the conflict is normal. The combinational result converges to the true proposition A, and it is close to the final result. Among them, the method of Li Wenli calculates the weight of evidence according to the distance function between focal elements, and the convergence speed is not high. The Han Deqiang method effectively controls the convergence rate of the D–S rule based on the evidential variance correction method and makes the result more credible. But the accuracy of this method is lower than ours; moreover, it is more suitable for sequential combination because it is vulnerable to the order of combination. The Wang Lianfeng method based on the transitive closure clustering effectively avoids the interference impact of evidence m_3. The combinational result converges to true proposition A. But its convergence speed and the accuracy of the final result are lower than ours. The method of calculating the reliability proposed in this paper fully considers the impact of internal and external contradictions of evidence. This greatly reduces the impact of negative evidences on combination result, which shows strong anti-interference ability. As shown in Table II, the final combinational result with our method, i.e., $m(A) = 0.9896$, is significantly better than that of other methods, which has the high accuracy, strong anti-interference ability, and fast convergence speed.

6.2. One Conflictive and Neutral Evidences Involved

Example 5. For one conflictive and neutral evidences involved in a combination, compare the performance of the method proposed in this paper with other methods. Suppose there is a discernment framework $\Theta = \{A, B, C\}$, and there are five independent evidences $\{\Theta, \Gamma, m_i\}$, $i = 1, 2, \ldots, 5$.

$$\Gamma_1 = \{\{A\}, \{B\}, \{C\}\}, \; m_1 = \{0.90, 0.00, 0.10\}$$
$$\Gamma_2 = \{\{A\}, \{B\}, \{C\}\}, \; m_2 = \{0.00, 0.01, 0.99\}$$
$$\Gamma_3 = \{\{A\}, \{B\}, \{C\}\}, \; m_3 = \{0.3334, 0.3333, 0.3333\}$$
$$\Gamma_4 = \{\{A\}, \{B\}, \{C\}\}, \; m_4 = \{0.98, 0.01, 0.01\}$$
$$\Gamma_5 = \{\{A\}, \{B\}, \{C\}\}, \; m_5 = \{0.90, 0.05, 0.05\}$$

Among them, evidence m_2 obviously has conflicts with other evidences, and evidence m_3 is seen as a neutral evidence. But m_3 is lightly inclined to support A for $m_3(A) = 0.3334 > m_3(B) = m_3(C) = 0.3333$.

There might be three different classifications according to the different parameters set. That is, (1) m_2 and m_3, respectively, as one class, the other as one class; (2) m_2 separately as one class, the other as one class; (3) m_2 and m_3 as one class, the other as one class.

As observed from Table III, the Dempster rule does not have a reasonable result with high conflict, even if all the rest evidential sources support the true proposition, and this is unreasonable. The Yager method has the same problem, that is, $m(A)$ and $m(B)$ are always 0, $m(C)$ is decreasing with the increasing number of evidences combined, which support the true proposition, and the value of $m(\Theta)$ is increasing. Obviously, the decision result is more uncertain with the phenomenon of entropy increase, which illustrates this combination result is invalid. The combination result of the methods of Sun Quan, Deng Yong, and Xiong Yanming assign a large part of assignment to $m(\Theta)$, which makes the value of $m(A)$ small. For instance, $m(A) = 0.164679$, $m(C) = 0.298444$ in Sun Quan and Deng Yong methods; $m(A) = 0.246347$, $m(C) = 0.278064$ in the Xiong Yanming method; $m(A) = 0.405365$, $m(C) = 0.594585$ in our method. Our result is more consistent with intuitive response. In the methods of Li Bicheng, Wang Dong, and Han Deqiang, the combinational result is obviously unreasonable that $m(A)$ has a decreasing phenomenon when a neutral evidence source is combined after conflictive evidence because of $m_3(A) = 0.3334 > m_3(B) = m_3(C) = 0.3333$. That is to say, the combinational result decreases while the neutral evidence is involved. For Lu Zhengcai, Li Wenli, and our methods, the value of $m(A)$ slightly increases when the neutral evidence source is combined after the conflictive evidence. But the accuracy of Lu Zhengcai and Li Wenli methods is lower than ours, when the number of evidences combined is five. The final combinational result with our method is $m(A) = 0.999196$ when the evidences are divided into two categories $\{m_2\}\{m_1, m_3, m_4, m_5\}$; $m(A) = 0.993324$ when they are divided into $\{m_2, m_3\}\{m_1, m_4, m_5\}$ and the result $m(A) = 0.998266$ when they are divided into three categories $\{m_2\}\{m_3\}\{m_1, m_4, m_5\}$. This experiment shows that the final combinational results have a little difference when the clustering feature is fuzzy so that the clustering suffers from the impact of the clustering parameter setting, but our method has an obvious advantage over other methods.

6.3. Two Conflictive Evidences Involved

Example 6. When two conflictive evidences are involved in the combination, compare our method with other methods. Suppose that there are five independent evidences $\{\Theta, \Gamma, m_i\}$, $i = 1, 2, \ldots, 5$ in a system:

$$\Gamma_1 = \{\{A\},\ \{B\},\ \{C\}\},\ m_1 = \{0.60, 0.10, 0.30\}$$
$$\Gamma_2 = \{\{A\},\ \{B\},\ \{C\}\},\ m_2 = \{0.10, 0.10, 0.80\}$$
$$\Gamma_3 = \{\{A\},\ \{B\},\ \{C\}\},\ m_3 = \{0.30, 0.10, 0.60\}$$
$$\Gamma_4 = \{\{A\},\ \{B\},\ \{C\}\},\ m_4 = \{0.70, 0.20, 0.10\}$$
$$\Gamma_5 = \{\{A\},\ \{B\},\ \{C\}\},\ m_5 = \{0.55, 0.10, 0.35\}$$

Among them, m_2 and m_3 are obviously conflictive with m_1, m_4, m_5. Set the clustering parameters as follows: set the number of initial evidences as the expected number of categories; set the upper limitation of standard deviation and the lower limitation of minimum distance as 0.1 and 0.3, respectively. Cluster the evidences into two categories, i.e. $\{m_2, m_3\}$ and $\{m_1, m_4, m_5\}$. As observed

Table III. The comparison of different methods when one conflict and neutral evidences are involved

Combination rules	m_1, m_2	m_1, m_2, m_3	m_1, m_2, m_3, m_4	m_1, m_2, m_3, m_4, m_5
D-S[1]	m(A) = 0 m(B) = 0 m(C) = 1 m(Θ) = 0	m(A) = 0 m(B) = 0 m(C) = 1 m(Θ) = 0	m(A) = 0 m(B) = 0 m(C) = 1 m(Θ) = 0	m(A) = 0 m(B) = 0 m(C) = 1 m(Θ) = 0
Yager[7]	m(A) = 0 m(B) = 0 m(C) = 0.099000 m(Θ) = 0.901000	m(A) = 0 m(B) = 0 m(C) = 0.032997 m(Θ) = 0.967003	m(A) = 0 m(B) = 0 m(C) = 0.00330 m(Θ) = 0.999670	m(A) = 0 m(B) = 0 m(C) = 0.000016 m(Θ) = 0.999984
Quan Sun[29]	m(A) = 0.164679 m(B) = 0.001830 m(C) = 0.298444 m(Θ) = 0.535047	m(A) = 0.188782 m(B) = 0.052545 m(C) = 0.250845 m(Θ) = 0.507828	m(A) = 0.283631 m(B) = 0.045273 m(C) = 0.183997 m(Θ) = 0.487099	m(A) = 0.345473 m(B) = 0.044291 m(C) = 0.163894 m(Θ) = 0.446341
Yong Deng[30]	m(A) = 0.164679 m(B) = 0.001830 m(C) = 0.298444 m(Θ) = 0.535047	m(A) = 0.191374 m(B) = 0.053266 m(C) = 0.253835 m(Θ) = 0.501525	m(A) = 0.293611 m(B) = 0.046866 m(C) = 0.190460 m(Θ) = 0.469063	m(A) = 0.364009 m(B) = 0.047153 m(C) = 0.173439 m(Θ) = 0.415399
Bicheng Li[31]	m(A) = 0.405450 m(B) = 0.004505 m(C) = 0.590045	m(A) = 0.397567 m(B) = 0.110657 m(C) = 0.491775	m(A) = 0.553167 m(B) = 0.088296 m(C) = 0.358537	m(A) = 0.622670 m(B) = 0.080659 m(C) = 0.296671
Dong Wang[9]	m(A) = 0.409247 m(B) = 0.004547 m(C) = 0.505543 m(Θ) = 0.080663	m(A) = 0.371308 m(B) = 0.103349 m(C) = 0.431777 m(Θ) = 0.093566	m(A) = 0.542287 m(B) = 0.086559 m(C) = 0.351167 m(Θ) = 0.019987	m(A) = 0.610227 m(B) = 0.079047 m(C) = 0.290727 m(Θ) = 0.019999
Yanming Xiong[18]	m(A) = 0.246347 m(B) = 0.002394 m(C) = 0.278064 m(Θ) = 0.473196	m(A) = 0.304055 m(B) = 0.153369 m(C) = 0.298275 m(Θ) = 0.244301	m(A) = 0.730667 m(B) = 0.038400 m(C) = 0.101213 m(Θ) = 0.129719	m(A) = 0.967549 m(B) = 0.005317 m(C) = 0.015495 m(Θ) = 0.011638
Deqiang Han[15]	m(A) = 0.405365 m(B) = 0.000050 m(C) = 0.594585	m(A) = 0.359450 m(B) = 0.072132 m(C) = 0.568418	m(A) = 0.825805 m(B) = 0.003393 m(C) = 0.170802	m(A) = 0.982867 m(B) = 0.000931 m(C) = 0.016202
Zhengcai Lu[19]	m(A) = 0.405450 m(B) = 0.004505 m(C) = 0.590045 m(Θ) = 0.000000	m(A) = 0.407566 m(B) = 0.145214 m(C) = 0.447219 m(Θ) = 0.000000	m(A) = 0.663919 m(B) = 0.067720 m(C) = 0.123781 m(Θ) = 0.144580	m(A) = 0.741408 m(B) = 0.057437 m(C) = 0.100032 m(Θ) = 0.101123
Wenli Li[20]	m(A) = 0.405450 m(B) = 0.004505 m(C) = 0.590045	m(A) = 0.414626 m(B) = 0.153884 m(C) = 0.431490	m(A) = 0.731444 m(B) = 0.076423 m(C) = 0.192132	m(A) = 0.823203 m(B) = 0.056724 m(C) = 0.120073
Our method (neutral, conflict and the other each as a category)	m(A) = 0.405365 m(B) = 0.000050 m(C) = 0.594585 m(Θ) = 0.000000	m(A) = 0.727101 m(B) = 0.104511 m(C) = 0.163856 m(Θ) = 0.004532	m(A) = 0.992293 m(B) = 0.002519 m(C) = 0.003522 m(Θ) = 0.001666	m(A) = 0.998266 m(B) = 0.000561 m(C) = 0.000632 m(Θ) = 0.000541
Our method (neutral and conflict as a category, the other as a category)	m(A) = 0.405365 m(B) = 0.000050 m(C) = 0.594585 m(Θ) = 0.000000	m(A) = 0.714291 m(B) = 0.093673 m(C) = 0.185852 m(Θ) = 0.006184	m(A) = 0.977078 m(B) = 0.007041 m(C) = 0.011255 m(Θ) = 0.004626	m(A) = 0.993324 m(B) = 0.002076 m(C) = 0.002618 m(Θ) = 0.001982
Our method (neutral and the other as a category, conflict as a category)	m(A) = 0.405365 m(B) = 0.000050 m(C) = 0.594585 m(Θ) = 0.000000	m(A) = 0.892516 m(B) = 0.027820 m(C) = 0.075301 m(Θ) = 0.004363	m(A) = 0.997546 m(B) = 0.000003 m(C) = 0.001072 m(Θ) = 0.001379	m(A) = 0.999196 m(B) = 0.000001 m(C) = 0.000107 m(Θ) = 0.000696

Table IV. The comparison of different methods when two conflict evidences are involved

Combination rules	m_1, m_2	m_1, m_2, m_3	m_1, m_2, m_3, m_4	m_1, m_2, m_3, m_4, m_5
D-S[1]	m(A) = 0.193548 m(B) = 0.032258 m(C) = 0.774194 m(Θ) = 0	m(A) = 0.110429 m(B) = 0.006135 m(C) = 0.883436 m(Θ) = 0	m(A) = 0.463235 m(B) = 0.007353 m(C) = 0.529412 m(Θ) = 0	m(A) = 0.577982 m(B) = 0.001668 m(C) = 0.420350 m(Θ) = 0
Yager[7]	m(A) = 0.060000 m(B) = 0.010000 m(C) = 0.240000 m(Θ) = 0.690000	m(A) = 0.018000 m(B) = 0.001000 m(C) = 0.144000 m(Θ) = 0.837000	m(A) = 0.012600 m(B) = 0.000200 m(C) = 0.014400 m(Θ) = 0.972800	m(A) = 0.006930 m(B) = 0.000020 m(C) = 0.005040 m(Θ) = 0.988010
Quan Sun[29]	m(A) = 0.181130 m(B) = 0.044609 m(C) = 0.430348 m(Θ) = 0.343913	m(A) = 0.171118 m(B) = 0.046936 m(C) = 0.404301 m(Θ) = 0.377645	m(A) = 0.229516 m(B) = 0.063999 m(C) = 0.244076 m(Θ) = 0.462408	m(A) = 0.244791 m(B) = 0.063450 m(C) = 0.232329 m(Θ) = 0.459431
Yong Deng[30]	m(A) = 0.181131 m(B) = 0.044609 m(C) = 0.430348 m(Θ) = 0.343912	m(A) = 0.171694 m(B) = 0.047108 m(C) = 0.405279 m(Θ) = 0.375919	m(A) = 0.230189 m(B) = 0.064197 m(C) = 0.244789 m(Θ) = 0.460825	m(A) = 0.245566 m(B) = 0.063656 m(C) = 0.233069 m(Θ) = 0.457709
Bicheng Li[31]	m(A) = 0.301500 m(B) = 0.079000 m(C) = 0.619500 m(Θ) = 0	m(A) = 0.297000 m(B) = 0.084700 m(C) = 0.618300 m(Θ) = 0	m(A) = 0.426040 m(B) = 0.121800 m(C) = 0.452160 m(Θ) = 0	m(A) = 0.451535 m(B) = 0.118581 m(C) = 0.429884 m(Θ) = 0
Dong Wang[9]	m(A) = 0.246949 m(B) = 0.066985 m(C) = 0.460920 m(Θ) = 0.225146	m(A) = 0.195602 m(B) = 0.056481 m(C) = 0.389224 m(Θ) = 0.358694	m(A) = 0.304180 m(B) = 0.088413 m(C) = 0.322391 m(Θ) = 0.285016	m(A) = 0.318429 m(B) = 0.084366 m(C) = 0.303802 m(Θ) = 0.293403
Yanming Xiong[18]	m(A) = 0.281988 m(B) = 0.068915 m(C) = 0.436950 m(Θ) = 0.212147	m(A) = 0.218618 m(B) = 0.047878 m(C) = 0.720752 m(Θ) = 0.012752	m(A) = 0.438538 m(B) = 0.062424 m(C) = 0.442207 m(Θ) = 0.056831	m(A) = 0.622042 m(B) = 0.032843 m(C) = 0.333968 m(Θ) = 0.011147
Deqiang Han[15]	m(A) = 0.281609 m(B) = 0.022989 m(C) = 0.695402	m(A) = 0.167116 m(B) = 0.007668 m(C) = 0.825216	m(A) = 0.417912 m(B) = 0.023082 m(C) = 0.559006	m(A) = 0.526216 m(B) = 0.008480 m(C) = 0.465304
Zhengcai Lu[19]	m(A) = 0.301500 m(B) = 0.079000 m(C) = 0.619500	m(A) = 0.290025 m(B) = 0.084700 m(C) = 0.625275	m(A) = 0.428742 m(B) = 0.119098 m(C) = 0.452160	m(A) = 0.463885 m(B) = 0.116411 m(C) = 0.419705
Wenli Li[20]	m(A) = 0.301500 m(B) = 0.079000 m(C) = 0.619500	m(A) = 0.282364 m(B) = 0.079232 m(C) = 0.638404	m(A) = 0.434511 m(B) = 0.114697 m(C) = 0.450792	m(A) = 0.473329 m(B) = 0.113191 m(C) = 0.413480
Our method	m(A) = 0.281609 m(B) = 0.022989 m(C) = 0.695402	m(A) = 0.129168 m(B) = 0.021489 m(C) = 0.838878	m(A) = 0.479296 m(B) = 0.011829 m(C) = 0.508875	m(A) = 0.848019 m(B) = 0.004798 m(C) = 0.145478

from Table IV, from the point of view of intuition, the proposition C is true when $m_1 = \{0.60, 0.10, 0.30\}$ and $m_2 = \{0.10, 0.10, 0.80\}$ are combined. Our method get a result $m(C) = 0.695402$, and the method of Han Deqiang has the same conclusion as ours. But the method of Xiong Yanming has a result $m(C) = 0.436950$. The proposition C should be supported when m_1, m_2, and m_3 are combined.

However, only the result with Han Deqiang $m(C) = 0.825216$ is close to ours $m(C) = 0.838878$. The value of $m(C)$ with our method in combining the previous three evidences is the highest among all methods. And then, when another two evidences which support A are involved in the combination, the proposition A should be true from the point of view of intuition. In the final combinational result of five evidential sources, the method of Xiong Yanming has the highest $m(A) = 0.62204$ among all the other methods. But it is far less than that of ours $m(A) = 0.848019$, which shows that our method has a stronger anti-interference ability and a higher accuracy.

7. CONCLUSION

This paper focuses on the counterintuitive phenomenon of D–S combination rule in dealing with highly conflictive information. With the deep analysis of the cause of the evidential conflict and the disadvantage of the similarity measure to weigh evidential conflict, this paper proposes the new similarity measure by considering internal and external contradictions and amends the evidential sources according to the commonality and, then, clusters these evidences using the ISODATA method and get the evidential representatives. After that, our method combines evidential representatives. Our method makes a thorough comparison with other methods through a lot of numeric examples and presents a better solution to the counterintuitive phenomenon than that of other methods. Certainly, our method can still be further improved. For example, more factors can be considered to calculate the reliability of the evidences. Furthermore, the final combinational result might be different according to the different parameters set when the clustering feature is fuzzy, and a better clustering approach is expected to reduce this impact in the future.

Acknowledgments

This work has been supported by NNSF of China (nos. 61175091 and 61573097), Aeronautical Science Foundation of China (no. 20140169002), and supported by Qing Lan Project and Six Major Top-talent Plan.

References

1. Shafer G. A mathematical theory of evidence. Princeton, NJ: Princeton University Press; 1976.
2. Zadeh LA. Review of books: a mathematical theory of evidence. AI Mag 1984;5(3):81–83
3. Haenni R. Are alternatives to Dempster's rule of combination real alternatives. Inform Fusion 2002;3(4):237–239.
4. Bian Z, Zhang X. Pattern recognition. Beijing, People's Republic of China: Tsinghua University Press; 2000.
5. Martin A, Osswald C. Toward a combination rule to deal with partial conflict an d specificity in belief functions theory. In: Int Conf on Information Fusion. Quebec, Canada; 9–12 July 2007.

6. Lefevre E, Colot O, Vannoorenberghe P. Belief functions combination and conflict management. Inform Fusion 2002;3(2):149–162.
7. Yager RR. On the D-S framework and new combination rules. Inform Sci 1987;41(2):93–138.
8. Liu W. Analyzing the degree of conflict among belief functions. Artif Intell 2006;5:909–924.
9. Wang D, Li Q, Jiang W, Xu XB, Deng Y. New method to combine conflict evidence based on pignistic probability distance. Infrared Laser Eng 2009;1:149–154.
10. Deng Y, Wang D, Li Q, Zhang YJ. A new method to analyze evidence conflict. Control Theory Appl 2011;28(6):839–844.
11. Quan W, Wang XD, Wang J, Zhang YX. New combination rule of DST based on local conflict distribution strategy. Dianzi Xuebao (Acta Electron Sin 2012;40(9):1880–1884.
12. He B, Hu HL. A modified DS evidence combination strategy. Acta Aeronaut Astron Sin 2003;24(6):559–562.
13. Yang JB, Xu£¬Evidential DL. Reasoning rule for evidence combination. Artif Intell 2013:205;1–29
14. Smets P. Analyzing the combination of conflicting belief functions. Inform Fusion 2007;8(4):387–412.
15. Han D, Han C, Deng Y, Yang Y. Weighted combination of conflicting evidence based on evidence variance. Acta Electron Sin 2011;39(3A):153–157.
16. Hu LF, Guan X, Deng Y, He Y. Cause-analysis for conflicting evidences in the generalized power space. Control Theory Appl 2011;28(12):1717–1722.
17. Smarandache F, Dezert J. Advances and applications of DSmT for information fusion. Rehoboth, NM: American Research Press. Vols. 1–3. 2004–2009.
18. Xiong YM, Yang ZP, Qu XF. Novel combination method of conflict evidence based on evidential model modification. Control Decis 2011;26(6):883–887.
19. Lu Z, Qin Z. General framework for evidence combination and its approach to highly conflicting evidence fusion. J Tsinghua Univ (Sci Technol) 2011;11:016.
20. Li WL, Guo KH. Combination rules of DS evidence theory and conflict problem. Syst Eng Theory Pract 2010;30(8):1422–1432.
21. Wang LF, Song JS, Zhu Y, Cao JP. Evidence combination based on fuzzy clustering analysis. Syst Eng Electron 2013;35(1):113–119.
22. Quan W, Wang X, Shi Z. Conflict evidence combination strategy based on evidence classification. J Air Force Eng Univ (Nat Sci Ed) 2012;13(3):80–84.
23. Wang J, Wu D, Cao J, Li J. Weighted combination of conflicting evidence based on evidence classification. Comput Sci 2013;40(1):247–250.
24. Jousselme AL, Maupin P. Distances in evidence theory: comprehensive survey and generalizations. Int J Approx Reason 2012;53:118–145.
25. Jousselme AL, Grenier D, Bosse E. A new distance between two bodies of evidence. Inform Fusion 2001;2(2):91–101.
26. Li XD, Dezert J, Smarandache F, Huang XH. Evidence supporting measure of similarity for reducing the complexity in information fusion. Inform Sci 2011;181(10):1818–1835.
27. Smets P. Data fusion in the transferable belief model. In: Proc Third Int Conf. on Information Fusion, Paris, France; 10–13 July 2000. pp. 21–33.
28. Gu Z. The research and application of compositive mathematical methods to water quality analysis, simulation and forecast in reservoirs and lakes. Jilin University, Jilin, China, doctoral thesis, 2006.
29. Sun Q, Ye X, Gu W. A new combination rules of evidence theory. Acta Electron Sin 2000;28(8):117–119.
30. Deng Y, Shi W. A modified combination rule of evidence theory. J Shanghai Jiaotong Univ (Chin ed) 2003;37(8):1275–1278.
31. Li B, Wang B, Wei J, Huang YQ, Guo ZG. Efficient combination rule of evidence theory. In: Multispectral image processing and pattern recognition. International Society for Optics and Photonics; SPIE Proceedings 2001;4554. pp. 237–240.
32. Murphy CK. Combining belief functions when evidence conflicts. Decis Support Syst 2000;29(1):1–9.