Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 27 (2006) 366–376
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Fault diagnosis of machines based on D–S evidence theory. Part 1: D–S evidence theory and its improvement

Xianfeng Fan, Ming J. Zuo *

Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G8

Received 15 June 2004; received in revised form 30 August 2005
Available online 13 October 2005

Communicated by K. Tumer

Abstract

In this paper, conventional D–S evidence theory is improved through the introduction of fuzzy membership function, importance index, and conflict factor in order to address the issues of evidence sufficiency, evidence importance, and conflicting evidences in the practical application of D–S evidence theory. New decision rules based on the improved D–S evidence theory are proposed. Examples are given to illustrate that the improved D–S evidence theory is better able to perform fault diagnosis through fusing multi-source information than conventional D–S evidence theory.

Keywords: D–S evidence theory; Information fusion; Fuzzy theory; Conflict resolution

1. Introduction

There is no single sensor that can reliably obtain all the information required for fault diagnosis. With the development of sensor technology and signal processing methods, a great deal of information can be obtained. New challenges have arisen with regard to making more reasonable inferences based on multi-source information. Bayesian inference can be used to update the probability of a hypothesis given a piece of observed information. This property makes possible multi-source information fusion. For details on Bayesian inference readers can refer to Gelman et al. (2004). Though Bayesian inference can be employed to determine the probability of a decision's correctness based on prior information, it has some disadvantages: (1) the knowledge required to generate the prior probability distributions may not be available; (2) instabilities may occur when conflicting data is presented and/or the number of unknown propositions is large compared to the known propositions (Roemer et al., 2001); (3) information available to the decision-maker must be characterized by a specific distribution or an exact assertion of the truth of a proposition; and (4) it offers little opportunity to express incomplete information or partial belief (Ducey, 2001).

Based on Dempster's research, Shafer (1976) proposed Dempster–Shafer (D–S) evidence theory as an alternative to Bayesian inference. D–S evidence theory can be considered to be a general extension of Bayesian theory and can robustly deal with incomplete data. It allows the representation of both imprecision and uncertainty (Beynon et al., 2001; Hégarat-Mascle et al., 2003). Rather than computing probabilities of propositions, it computes probabilities that evidence supports the propositions and offers an alternative approach to dealing with uncertainty reasoning based on incomplete information. D–S evidence theory tackles the prior probability issue by keeping track of an explicit probabilistic measure of a possible lack of information. It is suitable for taking into account the disparity between knowledge types (Kaftandjian et al., 2003) because it is able

DOI of original article: 10.1016/j.patrec.2005.08.024
* Corresponding author. Tel.: +1 780 492 4466; fax: +1 780 492 2200.
E-mail address: ming.zuo@ualberta.ca (M.J. Zuo).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2005.08.025

to provide a federative framework (Fabre et al., 2001) and it combines cumulative evidence for changing prior opinions in the light of new evidence (Parikh et al., 2001).

Parikh et al. (2001) proposed a new method based on the fact that the use of predictive accuracy for basic probability assignment can improve performance over that provided by traditional basic probability assignment methods. Wang et al. (1999) proposed the combined fuzzy logic/D–S evidence theory method. Although Jones et al. (2002) noted that it is necessary to ensure the validity of basic probability assignment; they did not give a method to do so. Additional research is needed for application of D–S evidence theory in practical decision-making, because evidence sufficiency, evidence importance, and conflicting evidence were not addressed in conventional D–S evidence theory. This provided the motivation for this research combining fuzzy set theory, weight of evidence, conflict resolution, and D–S evidence theory.

In this paper, we report a new method based on D–S evidence theory that can increase the accuracy of decision-making through fusing multi-source information. The rest of the paper is organized as outlined below. Section 2 describes D–S evidence theory and its disadvantages. Section 3 reports the improvement of conventional D–S evidence theory. The issues of evidence sufficiency, evidence importance, and conflicting evidence are addressed. Decision rules based on the improved D–S evidence theory are given. A flowchart based on the proposed method is provided. In Section 4, an example is given to demonstrate the validity of the improved D–S evidence theory. Conclusions are summarized at the end.

2. D–S evidence theory

Let Θ be a finite nonempty set of mutually exclusive alternatives, and be called the frame of discernment (Shafer, 1976). This frame of discernment contains every possible hypothesis. In fault diagnosis, we have a set of failure hypotheses, $\Theta = \{F_i | i = 1, 2, \ldots, N\}$, where F_i is the hypothesis of "Fault F_i is present". We need to evaluate the strength of belief in each fault hypothesis.

Basic Probability Assignment (BPA) is a function, $m: 2^{\Theta} \to [0, 1]$, such that $m(\emptyset) = 0$ where $\emptyset$ denotes an empty set, and $\sum_{X \subseteq \Theta} m(X) = 1$. The power set 2^{Θ} is the set of all the subsets of Θ including itself (Beynon et al., 2001; Parikh et al., 2001). Given a piece of evidence, a belief level between $[0, 1]$, denoted by $m(\cdot)$, is assigned to each subset of Θ. Each subset contains one or more hypothesis. No belief ought to be committed to an empty set. All the BPAs add up to unity.

If evidence for a hypothesis exists, the corresponding BPA is said to have been fired by the evidence. This BPA will be considered in information fusion. Otherwise, the BPA for the hypothesis has not been fired and will not be employed in information fusion. For example, a piece of evidence is obtained from a series of collected data, and the evidence is for faults $\{F_1\}$, $\{F_2\}$, and $\{F_4, F_5\}$. Then, it can be said that the evidence for faults F_1, F_2, F_4, or, F_5 is present. Assuming $m(\{F_1\}) = 0.6$, $m(\{F_2\}) = 0.3$, and $m(\{F_4, F_5\}) = 0.1$, it can be said that these BPAs are fired by the evidence. This tells us that F_1, F_2, $\{F_4, F_5\}$ are present and have belief levels of 60%, 30%, and 10%, respectively.

The value of 0 indicates zero belief in a hypothesis, while the value of 1 indicates total belief. A value between these two limits indicates partial belief. If the BPAs for only a few hypotheses are known, the complementary BPA value is assigned to the frame of discernment; that is, $m(\Theta)$ is equal to one minus the sum of the known BPAs. This value then represents ignorance. Any subset X of Θ for which $m(X)$ is known which is not equal to zero is called a focal element. The BPA of Θ denoted by $m(\Theta)$ is considered to represent the amount of uncertainty, and is called ignorance (Beynon et al., 2001). For example, $m(\Theta) = m(\{F_1, F_2, \ldots, F_N\}) = 0.2$ means that the chance for fault F_1 or $F_2 \ldots$ or F_N to exist is 20%. In other words, there is a 20% uncertainty regarding fault sources based on available evidence. If existing evidence cannot distinguish between hypotheses F_i and F_j, a BPA value is assigned to the set including both F_i and F_j, denoted as $m(\{F_i, F_j\})$. The BPA of set X, $m(X)$, represents the possibility that the correct hypothesis is in the set X but is in no particular subset of X (Klir and Wierman, 1998). Because evidence theory allows some belief to be uncommitted, it is possible to develop lower and upper bounds of uncertainty for any subset of 2^{Θ}.

The total belief level committed to X, $Bel: 2^{\Theta} \to [0, 1]$, can be obtained by calculating the belief function for X; this adds the BPAs of all subsets of X:

$$Bel(X) = \sum_{Y \subseteq X} m(Y), \text{ for any } X \subseteq \Theta. \tag{1}$$

The difference between $m(X)$ and $Bel(X)$ is that $m(X)$ measures the assignment of belief only to X, not the total assignment of belief to X and all its subsets. $Bel(X)$ measures the total assignment of belief to X and all its subsets. $Bel(X)$ represents the belief level that a proposition lies in X or any subset of X. For example, $Bel(\{F_1, F_2\}) = m(\{F_1\}) + m(\{F_2\}) + m(\{F_1, F_2\})$. Here, $X = \{F_1, F_2\}$ and Y in Eq. (1) takes $\{F_1\}$, $\{F_2\}$, and $\{F_1, F_2\}$.

The plausibility function defined below measures the extent to which we fail to disbelieve the hypothesis of X, $Pl: 2^{\Theta} \to [0, 1]$:

$$Pl(X) = \sum_{Y \cap X \neq \emptyset} m(Y), \text{ for any } X \subseteq \Theta \text{ and all } Y \subseteq \Theta. \tag{2}$$

For example, if $\Theta = \{F_1, F_2, F_3\}$, then $Pl(\{F_1, F_2\}) = m(\{F_1\}) + m(\{F_2\}) + m(\{F_1, F_2\}) + m(\{F_1, F_3\}) + m(\{F_2, F_3\}) + m(\Theta)$. Here $X = \{F_1, F_2\}$, Y in Eq. (2) takes $\{F_1\}$, $\{F_2\}$, $\{F_1, F_2\}$, $\{F_1, F_3\}$, $\{F_2, F_3\}$, and Θ.

Both imprecision and uncertainty can be represented by Bel and Pl (Beynon et al., 2001). In Bayesian theory, the uncertainty of an event is expressed through a single value, the probability that the event will occur, and the imprecision of the measurement of the event is assumed to be null.

In D–S evidence theory, the belief value of hypothesis A is interpreted as the minimum uncertainty value of A, and its plausibility value is interpreted as the maximum uncertainty value of A. The relationship between Bel and Pl is

$$Pl(A) = 1 - Bel(\overline{A}), \tag{3}$$

$$Pl(A) \geqslant Bel(A). \tag{4}$$

$Bel(A)$ and $Pl(A)$ are the lower limit and the upper limit, respectively, of the belief level of hypothesis A. $\overline{A}$ is the negation of hypothesis A. $[Bel(A), Pl(A)]$ is the confidence interval which describes the uncertainty of A. $Pl(A) - Bel(A)$ for $A \subseteq \Theta$ represents the ignorance level in hypothesis A. If the information for fusion is missing or unreliable, the difference between $Bel(A)$ and $Pl(A)$ will increase. This difference provides a measurement of the imprecision and the uncertainty of the belief level in decision-making (Hégarat-Mascle et al., 2003). The above discussion is illustrated in Fig. 1.

Multiple evidences can be fused using Dempster's combination rule, shown in Eq. (5), which is also called the orthogonal sum of evidences (Beynon et al., 2001). Evidences related to any subsets X and Y of Θ can be used to calculate the belief level for some new hypothesis C. $C = X \cap Y$. If $C = \emptyset$, it means evidences conflict with each other totally and the belief level in hypothesis C is then null.

$$m(C) = m_i(X) \oplus m_{i'}(Y) = \begin{cases} 0, & \text{If } X \cap Y = \emptyset, \\ \dfrac{\sum_{X\cap Y = C, \forall X, Y \subseteq \Theta} m_i(X) \times m_{i'}(Y)}{1 - \sum_{X \cap Y = \emptyset, \forall X, Y \subseteq \Theta} m_i(X) \times m_{i'}(Y)}, & \text{If } X \cap Y \neq \emptyset, \end{cases} \tag{5}$$

where i (i') denotes the ith (i'th) evidence, $m_i(X)$ is the BPA of X supported by evidence i, and $m_{i'}(Y)$ is the BPA of Y supported by evidence i'. Let

$$k_{i,i'} = \sum_{X \cap Y = \emptyset} m_i(X) \times m_{i'}(Y). \tag{6}$$

The quantity $k_{i,i'}$ is called the conflict factor between the two pieces of evidence i and i', where $k_{i,i'} \neq 1$. It is obvious that $k_{i,i'} = k_{i',i}$.

Dempster's combination rule can be generalized to more than two hypotheses, as shown in Eq. (7). It is both commutative and associative. When hypotheses are fired by N pieces of evidence, Eq. (7) can be used to fuse these evidences. The final result represents the synthetic effects of all evidences.

$$m = m_1 \oplus m_2 \oplus \cdots \oplus m_N = (((m_1 \oplus m_2) \oplus \cdots) \oplus m_N). \tag{7}$$

According to Eq. (1), there is a one-to-one correspondence between Bel and m. The orthogonal sum of belief functions $Bel = Bel_1 \oplus Bel_2$ can be calculated based on the final BPAs according to Eq. (5).

We have to note that totally contradictory evidences cannot be combined with one another using Dempster's rule. Such conflicts may result in an incorrect fusion result. This is illustrated in Example 1.

Example 1. Assuming a machine's failure modes include F_1, F_2, and F_3. The fault hypothesis set is $\Theta = \{F_1, F_2, F_3\}$. Assume that two evidences, 1 and 2, have been obtained. The BPAs of faults supported by such evidences are $m_1(\{F_1\}) = 0.9$, $m_1(\{F_3\}) = 0.1$, $m_2(\{F_2\}) = 0.9$, and $m_2(\{F_3\}) = 0.1$. These two evidences do not support any other subsets of 2^{Θ}. The given $m_1(\{F_1\})$ value of 0.9 and $m_1(\{F_3\})$ value of 0.1 indicate that evidence 1 provides a belief value of 90% supporting hypothesis F_1 and a belief value of 10% supporting hypothesis F_3. Evidence 2 provides belief values of 90% and 10% supporting hypotheses F_2 and F_3, respectively. If we apply Eq. (5) directly, the BPA of hypothesis F_3 based on these two evidences is

$$m(\{F_3\}) = \frac{0.1 \times 0.1}{1 - 0.9 \times 0.1 - 0.1 \times 0.9 - 0.9 \times 0.9} = 1.$$

This would indicate that hypothesis F_3 is 100% correct. Obviously, the conclusion is wrong because the two evidences do not really support F_3 very well. Each of the two evidences supports hypothesis F_3 with a BPA of 0.1.

Such a mistake arises from the fact that the two evidences have some agreement with each other on fault F_3 only (though these BPAs are very low). They contradict each other greatly on F_1, and F_2. This is one of the limitations of conventional D–S evidence theory; it sometimes generates wrong conclusions for decision-making. Some researchers have tried to solve this problem. Yager improved D–S evidence theory by classifying the conflicting evidences into set Θ (Yager, 1987). Though conflicts are removed, some useful information is lost. Further improvements are needed for effective decision-making.

3. The proposed improvement of D–S evidence theory

Considering the characteristics of practical applications of D–S evidence theory, the following issues should be addressed.

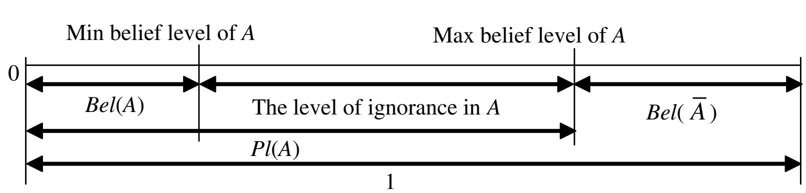

Fig. 1. The relation between *Bel* and *Pl*.

3.1. Evidence sufficiency

In practice, the data collected from sensors contain errors and the features extracted may not be certain. This affects the sufficiency of evidence. For example, when a sensor returns a value, the error in the sensor regarding this value is a function of the true value and other factors. As a result, we may not have adequate evidence to support a given hypothesis and to fire the corresponding BPA. In addition, if we consider expert opinions in decision-making, different experts may make different decisions given the same evidence. Each expert makes his own assessment of the evidence based on his own criteria (Kumar and Karmarkar, 2001). Suppose that, in the collected vibration data, we observe that the amplitudes at two frequencies f_1 and f_2, are 0.049 mm/s^2 and 0.051 mm/s^2, respectively. Because of the measurement errors of sensors, it is hard to say whether or not 0.049 mm/s^2 and 0.051 mm/s^2 are the real values. Using traditional D–S theory, we treat them as real values. In fault diagnosis, expert A may use an alarm threshold of 0.048, while expert B may use an alarm threshold of 0.050. Both evidences are sufficient according to expert A; while only the second evidence of 0.051 is sufficient according to expert B. We know however, that 0.049 and 0.051 are not that different from each other. Thus, it is necessary to find a way to represent the evidence sufficiency that depends on sensor error and expert's opinions in Dempster's combination rule.

Fuzzy theory can be used to optimize decision-making (Huang, 1997) and can provide a framework for representation of information sufficiency. Fuzzy theory is used in this paper to represent evidence sufficiency in an objective and quantitative manner. For each monitored parameter, a fuzzy membership function is introduced to translate the location of the feature in a plot into interpreted information, similar to that expressed by experts as "the amplitude is high". Though some research results of Fuzzy and D–S evidence theory are reported by Yager (1999) and Kaftandjian et al. (2003), nobody has yet used membership function to represent evidence sufficiency. Fuzzy membership function can describe evidence sufficiency well by expressing the fact that the amplitude is, to a certain degree, either large or small. A trapezoidal fuzzy number can be represented by a quadruplet $\boldsymbol{x} = (a_1, a_2, a_3, a_4)$. Its membership function is given below (Yager, 1999):

$$\mu(x) = \begin{cases} 0, & x < a_1, \\ \dfrac{x - a_1}{a_2 - a_1}, & a_1 \leqslant x < a_2, \\ 1, & a_2 \leqslant x \leqslant a_3, \\ \dfrac{a_4 - x}{a_4 - a_3}, & a_3 < x \leqslant a_4, \\ 0, & x > a_4. \end{cases} \tag{8}$$

Based on Eq. (8), the semi-trapezoidal function given in Eq. (9) is proposed to represent evidence sufficiency. This function is depicted in Fig. 2. Assigning a threshold to an evidence is done not only to ensure that reasonable results are obtained but also to prevent BPAs from being misfired. This threshold can avoid the non-sufficiency of evidence induced by sensor errors and expert opinions. This non-sufficiency level can be measured by $\mu(x)$ effectively.

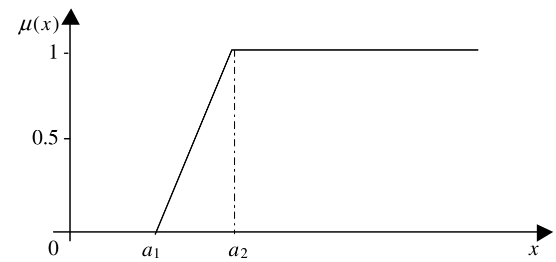

Fig. 2. Semi-trapezoidal fuzzy number $\boldsymbol{x} = (a_1, a_2)$.

$$\mu(x) = \begin{cases} 0, & 0 \leqslant x \leqslant a_1, \\ \dfrac{x - a_1}{a_2 - a_1}, & a_1 < x < a_2, \\ 1, & x \geqslant a_2. \end{cases} \tag{9}$$

Fig. 3 shows the power spectral density of a gearbox vibration signal. Assume that the sensor error is of an amplitude less than 1% of the real value. We can make Eq. (9) represent the evidence sufficiency of the sidebands' amplitude around the meshing frequencies of 266 Hz and 320 Hz. When the gearbox is normal, experts A and B will give the acceptable thresholds of gearbox vibration as, a_1^{A} and a_1^{B}, respectively. When one gear tooth is totally missing, experts A and B will give the alarm thresholds of vibration as, a_2^{A} and a_2^{B}, respectively. This means that expert A believes that the gearbox is normal when the vibration is below the value of a_1^{A}. If the vibration value is in the range of $[a_1^{\mathrm{A}}, a_2^{\mathrm{A}}]$, expert A considers the gear as having a partial fault. If the vibration value is beyond the value of a_2^{A}, expert A considers the gear as having a total fault. The experience of expert B is considered similarly using a_1^{B} and a_2^{B}. The influence of sensor error is considered by modifying these thresholds. Let $a_1 = \min(a_1^{\mathrm{A}}, a_1^{\mathrm{B}})(1 - 1\%) = 0.99\min(a_1^{\mathrm{A}}, a_1^{\mathrm{B}})$ and $a_2 = \max(a_2^{\mathrm{A}}, a_2^{\mathrm{B}})(1 + 1\%) = 1.01 \max(a_2^{\mathrm{A}}, a_2^{\mathrm{B}})$. When the amplitude x is greater than a_2, $\mu(x) = 1$; this means that the evidence is totally sufficient. When the amplitude x is greater than a_1 but less than a_2, $0 < \mu(x) < 1$; this means that the evidence is partially sufficient. When the amplitude x is less than a_1, $\mu(x) = 0$; this means that the evidence is considered as being completely non-existent.

For the evidence of E_i, its sufficiency can be represented by Eq. (9) and μ_i is called the sufficiency index. Evidence sufficiency can be incorporated through affecting attenuation of BPA, given in Eq. (10). The new BPA is denoted as $m_{i,*}(\cdot)$.

$$m_{i,*}(A) = \begin{cases} \mu_i \cdot m_i(A), & A \subset \Theta, \\ 1 - \sum_{B \subset \Theta} \mu_i \cdot m_i(B), & B \subset \Theta, \quad A = \Theta, \end{cases} \tag{10}$$

where $i = 1, 2, \ldots, M$, M is the number of evidences, and $*$ denotes that the BPA has incorporated evidence sufficiency.

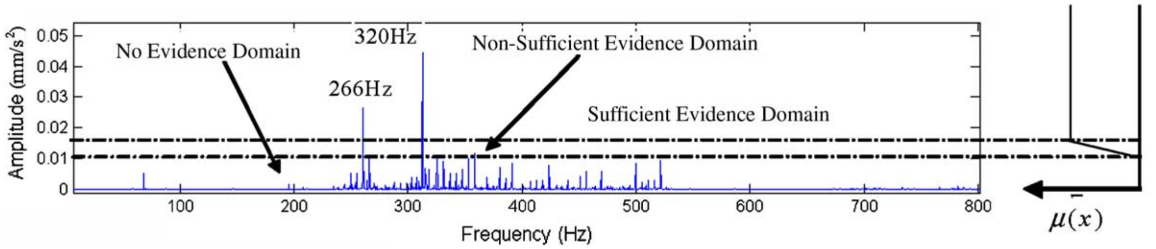

Fig. 3. Calculation of evidence sufficiency.

When a piece of evidence is partially sufficient, the corresponding BPA may be fired, but it has a lower chance of occurrence. The belief level for the hypothesis corresponding to the evidence, i, will be lower. Assume $m_1(\{F_1\}) = 0.6$, $m_1(\{F_2, F_3\}) = 0.2$, and $m_1(\Theta) = 0.2$; if the evidence obtained is not totally sufficient and the sufficiency index is 90%, according to Eq. (10), the new BPAs become $m_{1,*}(\{F_1\}) = 0.6 \times 0.9 = 0.54$, $m_{1,*}(\{F_2, F_3\}) = 0.2 \times 0.9 = 0.18$, and $m_{1,*}(\Theta) = 1 - 0.54 - 0.18 = 0.28$, respectively.

3.2. Evidence importance

In reality, not all the evidences have the same importance in decision-making. Some evidences are more important than others. Conventional D–S evidence theory does not differentiate the importance of different evidences. In fault diagnosis, because different evidences make different contributions to different faults, evidence importance should be considered for specific fault diagnosis through multi-resource information fusion. For diagnosing the fault of F_1, the importance indexes of all evidences can be written as $\{v_{i,1} | i = 1, 2, \ldots, M\}$. When diagnosing the fault of F_2, we use the importance indexes of all evidences denoted by $\{v_{i,2} | i = 1, 2, \ldots, M\}$.

the disease is dermatosis, then Doctor B's diagnosis should carry more weight than Doctor A's. According to the disease a patient may have, different weights will be employed; therefore, the importance of BPAs should be thought of in terms of a matrix. Based on this idea, we recommend that an importance weight; $w_{i,j}$ ($0 \leqslant w_{i,j} \leqslant 1$), be used to reflect the importance of evidence i for the hypothesis of fault j. The importance matrix, $\boldsymbol{W}$, showing the correlation between evidences and faults, is given in Eq. (11).

$$\boldsymbol{W} = \begin{array}{c} \\ E_1 \\ E_2 \\ \vdots \\ E_M \end{array} \begin{array}{c} \begin{array}{cccc} F_1 & F_2 & \cdots & F_N \end{array} \\ \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,N} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M,1} & w_{M,2} & \cdots & w_{M,N} \end{bmatrix} \end{array}, \tag{11}$$

where $\sum_{i=1}^{M} w_{i,j} = 1$, $j = 1, 2, \ldots, N, F_j$ denotes the jth fault, and E_i denotes the ith evidence. This means that for each fault j, the sum of the weights of evidences $E_1, E_2, \ldots,$ and E_M equals 1.

For all the subsets of Θ, we can calculate the importance matrix, $\boldsymbol{W}^*$, between evidences and subsets of Θ, as given in Eq. (12).

$$\boldsymbol{W}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ \vdots \\ E_M \end{array} \begin{array}{c} \begin{array}{cccccccccccc} \{F_1\} & \{F_2\} & \cdots & \{F_N\} & \{F_1,F_2\} & \{F_1,F_3\} & \{F_1,F_4\} & \cdots & \{F_2,F_3\} & \cdots & \{F_j \mid j \neq N\} & \Theta \end{array} \\ \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,N} & \sum_{j=1,2} w_{1,j} & \sum_{j=1,3} w_{1,j} & \sum_{j=1,4} w_{1,j} & \cdots & \sum_{j=2,3} w_{1,j} & \cdots & \sum_{j \neq N} w_{1,j} & \sum_{j} w_{1,j} \\ w_{2,1} & w_{2,2} & \cdots & w_{1,N} & \sum_{j=1,2} w_{2,j} & \sum_{j=1,3} w_{2,j} & \sum_{j=1,4} w_{2,j} & \cdots & \sum_{j=2,3} w_{2,j} & \cdots & \sum_{j \neq N} w_{2,j} & \sum_{j} w_{2,j} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ w_{M,1} & w_{M,2} & \cdots & w_{M,N} & \sum_{j=1,2} w_{M,j} & \sum_{j=1,3} w_{M,j} & \sum_{j=1,4} w_{M,j} & \cdots & \sum_{j=2,3} w_{M,j} & \cdots & \sum_{j \neq N} w_{M,j} & \sum_{j} w_{M,j} \end{bmatrix} \end{array}. \tag{12}$$

The following example can help us understand the importance of evidence. There are two doctors and three kinds of diseases: hypertension, dermatosis, and tracheitis. Doctor A has more experience in diagnosing hypertension and tracheitis while Doctor B is better at diagnosing dermatosis. In addition, they both have some knowledge of hypertension, dermatosis, and tracheitis. If the disease is tracheitis, then the weight given to Doctor A's diagnosis should be greater than the weight given to Doctor B's. If

Because there are $(2^N - 1)$ subsets in 2^Θ, there are $(2^N - 1)$ columns in Eq. (12). For each column in Eq. (12), the largest element in the column will be taken as the denominator. Thus, the new importance index of BPA, denoted by $v_{i,j'}$ ($i = 1, 2, \ldots, M$; $j' = 1, 2, \ldots, 2^N - 1$) can be expressed as in Eq. (13).

$$v_{i,j'} = \frac{w^*_{i,j'}}{\max\{w^*_{i,j'} | i = 1, 2, \ldots, M\}} \quad \text{for each } j', \tag{13}$$

where $w^*_{i,j'}$ denotes the element of the matrix, $\boldsymbol{W}^*$, and Eq. (14) is the matrix of the importance index obtained according to Eqs. (12) and (13).

$$\boldsymbol{V}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ \vdots \\ E_M \end{array} \begin{array}{c} \begin{array}{ccccccccccc} \{F_1\} & \{F_2\} & \cdots & \{F_N\} & \{F_1,F_2\} & \{F_1,F_3\} & \{F_1,F_4\} & \cdots & \{F_2,F_3\} & \cdots & \{F_j|j\neq N\} & \Theta \end{array} \\ \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,N} & v_{1,N+1} & v_{1,N+2} & v_{1,N+3} & \cdots & v_{1,2N} & \cdots & v_{1,2^N-2} & v_{1,2^N-1} \\ v_{2,1} & v_{2,2} & \cdots & v_{1,N} & v_{2,N+1} & v_{2,N+2} & v_{2,N+3} & \cdots & v_{2,2N} & \cdots & v_{2,2^N-2} & v_{2,2^N-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ v_{M,1} & v_{M,2} & \cdots & v_{M,N} & v_{M,N+1} & v_{M,N+2} & v_{M,N+3} & \cdots & v_{M,2N} & \cdots & v_{M,2^N-2} & v_{M,2^N-1} \end{bmatrix} \end{array}. \tag{14}$$

The importance of evidence is introduced through modifying the BPA as shown in Eq. (15). The new BPA is denoted as $m_{i,**}(\cdot)$.

$$m_{i,**}(A) = \begin{cases} v_{i,j'} \cdot m_i(A), & A \subset \Theta, \\ 1 - \sum_{B\subset\Theta} v_{i,j'} \cdot m_i(B), & B \subset \Theta, \quad A = \Theta, \end{cases} \tag{15}$$

where ** represents that the BPA has been modified by the evidence importance.

Example 2. The evidences in Example 1 are used here. Assuming that the second evidence importance is larger than the first evidence importance for all three fault modes, $v_{1,1} = v_{1,2} = v_{1,3} = 0.3$ and $v_{2,1} = v_{2,2} = v_{2,3} = 0.7$ are used to express the difference. In Table 1, the importance of the first evidence is 42.86% (0.3/0.7) of the importance of the second evidence. The new BPAs are calculated and shown in Table 1.

Based on the new BPAs, $m_{1,**}$ and $m_{2,**}$, the following fusion results can be obtained using Eq. (5):

$$m(\{F_1\}) = \frac{0.3852\times 0 + 0\times 0}{1-0.3852\times 0.9 - 0.3852\times 0.1 - 0.9\times 0.0428} = 0,$$

$$m(\{F_2\}) = \frac{0.9\times 0.572}{1-0.3852\times 0.9 - 0.3852\times 0.1 - 0.9\times 0.0428} = 0.8933,$$

$$m(\{F_3\}) = \frac{0.1\times 0.0428 + 0.1\times 0.572}{1-0.3852\times 0.9 - 0.3852\times 0.1 - 0.9\times 0.0428} = 0.1067,$$

$$m(\Theta) = \frac{0.572\times 0}{1-0.3852\times 0.9 - 0.3852\times 0.1 - 0.9\times 0.0428} = 0.$$

Using Eqs. (1) and (2), we have $Bel(\{F_1\}) = 0$, $Pl(\{F_1\}) = 0$, $Bel(\{F_2\}) = 0.8933$, $Pl(\{F_2\}) = 0.8933$, $Bel(\{F_3\}) = 0.1067$, and $Pl(\{F_3\}) = 0.1067$. The BPA of 0.8933 for hypothesis F_2 is the largest. Obviously, this is consistent with common sense. As a comparison with Example 1, this example shows that the introduction of importance can make fused results more reasonable.

The modification of BPAs considering both evidence sufficiency and evidence importance is shown in Eq. (16).

$$m_{i,\bullet}(A) = \begin{cases} \alpha_{i,j'} \cdot m_i(A), & A \subset \Theta, \\ 1 - \sum_{B\subset\Theta} \alpha_{i,j'} \cdot m_i(B), & B \subset \Theta, \quad A = \Theta, \end{cases} \tag{16}$$

where $m_{i,\bullet}(\cdot)$ is the BPA incorporating both sufficiency and importance of evidences, $\alpha_{i,j'}$ is the combining index of sufficiency and importance, and $\alpha_{i,j'} = v_{i,j'} \cdot \mu_i$.

The term $(1 - \sum_{B\subset\Theta} \alpha_{i,j'} \cdot m_i(B))$ in Eq. (16) indicates the loss of BPA due to $\alpha_{i,j'}$. The value of $(1 - \sum_{B\subset\Theta} \alpha_{i,j'} \cdot m_i(B))$ will be added to the BPA of Θ. As long as at least one evidence's importance is different from that of the others, according to the definition of $\alpha_{i,j'}$, we have $\alpha_{i,j'} = v_{i,j'} \cdot \mu_i < 1$. Using Eq. (16), it is easy to find that $m_{i,\bullet}(A) = \alpha_{i,j'} \cdot m_i(A) = v_{i,j'} \cdot \mu_i \cdot m_i(A) < m_i(A)$. Thus, the BPA for the hypothesis supported by less important and more uncertain evidences will be significantly reduced. According to Eq. (6), conflicts will be weakened through the attenuation of the hypothesis's BPAs and the concurrent increasing of Θ's BPA.

3.3. Conflicts among evidences

Evidences may conflict with one another. For each pair of evidences, a conflict factor denoted by $k_{i,i'}$, is defined in Eq. (6). If two evidences conflict with each other, and both are fully sufficient and equally important, the indexes of

Table 1
BPAs considering weights

BPA	$m\{F_1\}$ (Weight|importance index) (New BPA)	$m\{F_2\}$ (Weight|importance index) (New BPA)	$m\{F_3\}$ (Weight|importance index) (New BPA)	$m(\Theta)$ (New BPA)
m_1 ($m_{1,**}$)	0.9 $(0.3 \mid v_{1,1} = 0.3/0.7 = 0.428)$ $(0.9\times 0.428 = 0.3852)$	0 $(0.3 \mid v_{1,2} = 0.3/0.7 = 0.428)$ $(0\times 0.428 = 0)$	0.1 $(0.3 \mid v_{1,3} = 0.3/0.7 = 0.428)$ $(0.1\times 0.428 = 0.0428)$	0 $(1 - 0.3852 - 0.0428 = 0.572)$
m_2 ($m_{2,**}$)	0 $(0.7 \mid v_{2,1} = 0.7/0.7 = 1)$ $(0\times 1 = 0)$	0.9 $(0.7 \mid v_{2,2} = 0.7/0.7 = 1)$ $(0.9\times 1 = 0.9)$	0.1 $(0.7 \mid v_{2,3} = 0.7/0.7 = 1)$ $(0.1\times 1 = 0.1)$	0 $(1 - 0.9 - 0.1 = 0)$

sufficiency and importance proposed in the earlier section are not adequate for conflict resolution. In this case, we propose to classify evidences into two groups: conflicting and non-conflicting. First, we need to evaluate the conflict threshold, denoted by δ. δ represents the permitted conflict level among evidences. If the conflict factor between two evidences is greater than δ, one of the two evidences will be put into the conflicting group while the other remains in the non-conflicting group. Clearly, δ is the critical factor to be determined. If $k_{i,i'} > \delta$ and $k_{i'',i'} > \delta$, only evidence i' will be placed into the conflicting group. Each of the evidences has to be classified as being either in the conflicting group or in the non-conflicting group. We will apply Dempster's combination rule to fuse all evidences in the non-conflicting group first. The combined BPA for all evidences in the non-conflicting group will then be combined with the BPAs of the evidences in the conflicting group. This procedure avoids considering conflicting evidences at the beginning. Once the non-conflicting evidences have been combined, the weight of the combined BPAs is higher than that of the BPAs of the remaining conflicting evidences. According to Eqs. (10) and (6), the conflict between the evidences combined in the non-conflicting group and the evidences in the conflicting group will be relatively small.

As illustrated in Fig. 4, suppose that there are M evidences. Correspondingly, there are M sets of BPAs. According to the definition of conflict factor in Eq. (6), the conflict factor of each pair of evidences can be calculated and denoted as $k_{i,i'}$. Here, we assume that the conflict factor $k_{2,3}$ of evidences 2 and 3 is the only value that is beyond the threshold, δ. Thus, the conflicting group is comprised of the BPA, $m_{2,\bullet}(\cdot)$. The non-conflicting group includes all $\{m_{i,\bullet}(\cdot)|i \neq 2\}$. All the evidences in the non-conflicting group will be combined first, and then the result will be combined with evidence 2. This method can better avoid information loss than the method proposed by Yager (1987). For more detailed discussion of these issues, readers may refer to Fan (2003).

3.4. The modified Dempster's combination rule

For non-conflicting evidences, we apply Dempster's combination rule to fuse the BPAs modified by both the sufficiency index and the importance index. For combination with conflicting evidences, however, we propose a non-conflict factor to be used to modify Dempster's combination rule. Suppose that the conflict factor $k_{i,i'}$ between evidences i and i' is greater than δ. We define the non-conflict factor between these two evidences as follows:

$$\beta_{i,i''} = \begin{cases} 1 - (k_{i,i'} - \delta), & \text{if } k_{i,i'} > \delta, \\ 1, & \text{if } k_{i,i'} \leqslant \delta. \end{cases} \tag{17}$$

Obviously, $\beta_{i,i'} = \beta_{i',i} \cdot \beta_{i,i'}$ is used to represent the degree of non-conflict. The modified Dempster's combination rule is given in Eqs. (18) and (19). The BPA loss due to $\beta_{i,i'}$ is added to $m(\Theta)$.

$$m(C) = m_{i,\bullet}(X) \oplus m_{i',\bullet}(Y) = \begin{cases} 0, & X \cap Y = \emptyset, \\ \dfrac{\sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m_{i,\bullet}(X)\times m_{i',\bullet}(Y)}{1-\sum_{X\cap Y=\emptyset,\forall X,Y\subseteq\Theta} m_{i,\bullet}(X)\times m_{i',\bullet}(Y)}, & X\cap Y \neq \emptyset,\ k_{i,i'} < \delta, \end{cases} \tag{18}$$

$$m(C) = m'_{i,\bullet}(X) \oplus m'_{i',\bullet}(Y) = \begin{cases} 0, & X \cap Y = \emptyset, \\ \dfrac{\sum_{X\cap Y=C,\forall X,Y\subseteq\Theta} m'_{i,\bullet}(X)\times m'_{i',\bullet}(Y)}{1-\sum_{X\cap Y=\emptyset,\forall X,Y\subseteq\Theta} m'_{i,\bullet}(X)\times m'_{i',\bullet}(Y)}, & X\cap Y \neq \emptyset,\ k_{i,i'} \geqslant \delta, \end{cases} \tag{19}$$

where

$$m'_{i,\bullet}(X) = \begin{cases} \beta_{i,i'} \cdot m_{i,\bullet}(X), & X \subset \Theta, \\ 1 - \sum \beta_{i,i'} \cdot m_{i,\bullet}(A), & X = \Theta, \quad A \subset \Theta, \end{cases}$$

$$m'_{i',\bullet}(Y) = \begin{cases} \beta_{i,i'} \cdot m_{i',\bullet}(Y), & Y \subset \Theta, \\ 1 - \sum \beta_{i,i'} \cdot m_{i',\bullet}(B), & Y = \Theta, \quad B \subset \Theta, \end{cases}$$

where $k_{i,i'}$, $m_{i,\bullet}(m_{i',\bullet})$, and $\beta_{i,i'}$ are calculated with Eqs. (6), (16) and (17), respectively.

3.5. The decision rules

A decision is made according to the results obtained from multi-source information using the modified Dempster's combination rule. There are three popular decision rules: the maximum belief, the maximum plausibility, and the maximum belief without overlapping belief intervals (Hégarat-Mascle et al., 2003). These decision rules, however, are applicable only to the BPA values calculated using conventional D–S evidence theory. In this paper, sufficiency, importance, and the conflict factor of evidences are incorporated to modify the BPAs. The decision rules listed above are no longer applicable. It is necessary to develop new decision rules for decision-making using these modified BPAs for different hypotheses.

Given the development reported thus far in this paper, the following decision rules are proposed. These rules can be applied to BPAs obtained uisng either the method proposed in this paper or the conventional D–S theory.

(a) According to the proposed method for incorporating sufficiency, importance, and the non-conflict factor in calculating new BPAs, the sum of all BPAs for each

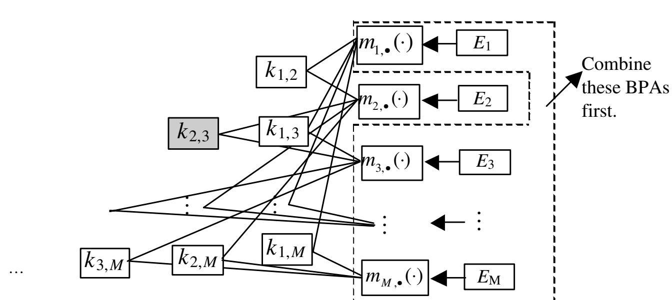

Fig. 4. Classification of evidences.

hypothesis is equal to 1. The BPA of Θ becomes larger in this process. Since the BPA of Θ represents uncertainty, we will not use it in our decision- making.

(b) The BPA for decision-making must be greater than a threshold denoted by Δ. We will use $\Delta = 0.6$ in this paper to illustrate the procedure since at present there are no better ways to determine the value of Δ. It is interpreted as having a belief level of at least 60% for any given fault.

(c) The proposed method calculates different sets of BPAs for each hypothesis. We check if the BPA for the fault that has been hypothesized is beyond the threshold, Δ, in fault diagnosis.

(d) Because the belief and plausibility of evidence can give us the certainty level of results, both a maximum of belief and a maximum of plausibility will be considered in decision-making.

3.6. Flowchart of the proposed method

The flowchart of the proposed procedure is shown in Fig. 5. The proposed method provides more reasonable results than does conventional D–S theory, because sufficiency and importance of evidence are considered. It also provides a mechanism to reduce or eliminate conflict among evidences.

4. Determination of parameters

In the improved D–S evidence theory, there are parameters to be determined such as BPA, sufficiency, importance, and conflict threshold. The derivation of BPA is the most crucial step, since BPA represents knowledge about the actual application, as well as the uncertainty incorporated in the selected information source: $0 \leqslant m(\cdot) \leqslant 1$ (Kaftandjian et al., 2003). BPAs can be determined in one of the following ways (Kurkowska et al., 2000; Salzenstein and Boudraa, 2004): flat BPA method in which all the BPAs are identical, expert knowledge method, occurrence frequency method, and Gaussian distribution modeling method. The second method, expert knowledge, is the one used in this paper.

Because weights reflect the importance of evidence in the diagnosis of fault F_j, weights of every two evidences which support fault F_j should be evaluated first. The procedure is given below.

(1) Evidences are first assumed to be totally sufficient. Choose the first two evidences, and calculate BPAs using Eq. (5). Here, $m_i(X) = v_i m_i(X)$, $m_{i'}(Y) = v_{i'} m_{i'}(Y)$, $m(C) = m(\{F_1\})$, $v_i = \frac{w_i}{\max(w_i, w_i')}$, and $v_i' = \frac{w_i'}{\max(w_i, w_i')}$. We allow the weight w_i to change from 0.1 to 0.9 with a step size of 0.1 and $w_{i'} = 1 - w_i$. The

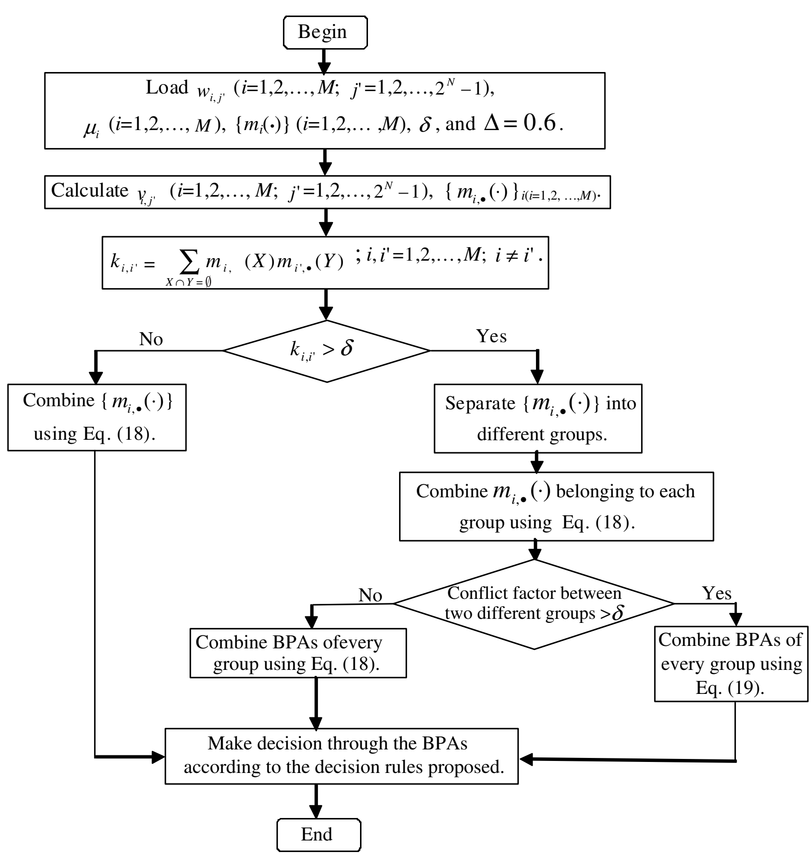

Fig. 5. Flowchart of the decision-making using the improved D–S evidence theory.

specific w_i value that maximizes the $m(\{F_j\})$ will be kept. The optimal weights for other pairs of evidences can be obtained in a similar manner.

(2) Based on the weights of every pair of the evidences, weights among all evidences can be obtained. This gives us the jth column of the matrix in Eq. (11).

Because the conflict factor $k_{i,i'}$ reflects the inconsistent level of evidences, i and i', it is reasonable to use the average conflict factor as a criterion in evidence fusion. That is why, $\delta = \frac{1}{n}\sum_{i'\neq i} k_{i,i'}$ is employed in this paper, where $n = \binom{M}{2}$. Considering Eqs. (17)–(19), this means that if $k_{i,i'}$ is less than the average conflict factor, δ, the total BPA will be kept; otherwise, only partial BPA value will be kept in fusion.

5. Example

The following example is used to demonstrate the effectiveness of the proposed method. Assume the following fault hypothesis set, $\Theta = \{F_1, F_2, F_3\}$, and the following evidence set, $E = \{E_1, E_2, E_3\}$. The BPAs based on these evidences are given in Table 2.

The conflict factors between each pair of evidences are $k_{1,2} = 0.52$, $k_{1,3} = 0.26$, and $k_{2,3} = 0.605$. The conflict threshold, $\delta = \frac{1}{3}(k_{1,2} + k_{1,3} + k_{2,3}) = 0.4617$, is employed. It is clear that the second evidence greatly conflicts with the other two evidences. We assume that the sufficiency indexes for the three evidences are 1, 0.6, and 1, respectively.

Next, the importance index matrix, $\boldsymbol{W}^*$, is obtained based on Table 2. It is given as Eq. (20).

$$\boldsymbol{W}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{cccc} \{F_1\} & \{F_2\} & \{F_2,F_3\} & \Theta \end{array} \\ \begin{bmatrix} 0.47 & 0.09 & 1 & 1 \\ 0.16 & 0.82 & 0.8 & 1 \\ 0.47 & 0.09 & 1 & 1 \end{bmatrix} \end{array}. \tag{20}$$

This importance matrix indicates that for F_1, the BPAs supported by E_1 and E_3 have the same importance and that the importance weights of E_1, E_2, and E_3 are 0.47, 0.16, and 0.47, respectively. Thus, the revised importance index matrix can be obtained according to Eq. (13). It is shown below.

$$\boldsymbol{V}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{cccc} \{F_1\} & \{F_2\} & \{F_2,F_3\} & \Theta \end{array} \\ \begin{bmatrix} 1 & 0.11 & 1 & 1 \\ 0.34 & 1 & 0.8 & 1 \\ 1 & 0.11 & 1 & 1 \end{bmatrix} \end{array}. \tag{21}$$

If we want to determine whether the fault is F_1, we use the first column in the importance matrix in Eq. (21), $[1, 0.34, 1]$, for evidence combination. Thus, the matrix of importance index for F_1 is as follows:

$$\boldsymbol{V}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{cccc} \{F_1\} & \{F_2\} & \{F_2,F_3\} & \Theta \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0.34 & 0.34 & 0.34 & 0.34 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}. \tag{22}$$

The new BPAs adjusted by the sufficiency index and the importance index are given by

$$\begin{aligned}
E_1\colon m_{1,\bullet}(\{F_1\}) &= m_1(\{F_1\}) \times v_{1,1} \times \mu_1 \\
&= 0.6 \times 1 \times 1 = 0.6, \\
m_{1,\bullet}(\{F_2\}) &= m_1(\{F_2\}) \times v_{1,1} \times \mu_1 \\
&= 0.1 \times 1 \times 1 = 0.1, \\
m_{1,\bullet}(\{F_2, F_3\}) &= m_1(\{F_2, F_3\}) \times v_{1,1} \times \mu_1 \\
&= 0.1 \times 1 \times 1 = 0.1, \\
m_{1,\bullet}(\Theta) &= 1 - 0.6 - 0.1 - 0.1 = 0.2; \\
E_2\colon m_{2,\bullet}(\{F_1\}) &= m_2(\{F_1\}) \times v_{2,1} \times \mu_2 \\
&= 0.05 \times 0.34 \times 0.6 = 0.0102, \\
m_{2,\bullet}(\{F_2\}) &= m_2(\{F_2\}) \times v_{2,1} \times \mu_2 \\
&= 0.8 \times 0.34 \times 0.6 = 0.1632, \\
m_{2,\bullet}(\{F_2, F_3\}) &= m_2(\{F_2, F_3\}) \times v_{2,1} \times \mu_2 \\
&= 0.05 \times 0.34 \times 0.6 = 0.0102, \\
m_{2,\bullet}(\Theta) &= 1 - 0.0102 - 0.1632 - 0.0102 = 0.8164; \\
E_3\colon m_{3,\bullet}(\{F_1\}) &= m_3(\{F_1\}) \times v_{3,1} \times \mu_3 \\
&= 0.7 \times 1 \times 1 = 0.7, \\
m_{3,\bullet}(\{F_2\}) &= m_3(\{F_2\}) \times v_{3,1} \times \mu_3 \\
&= 0.1 \times 1 \times 1 = 0.1, \\
m_{3,\bullet}(\{F_2, F_3\}) &= m_3(\{F_2, F_3\}) \times v_{3,1} \times \mu_3 \\
&= 0.1 \times 1 \times 1 = 0.1, \\
m_{3,\bullet}(\Theta) &= 1 - 0.7 - 0.1 - 0.1 = 0.1.
\end{aligned}$$

These new BPAs for fault F_1 are summarized in Table 3.

The new BPAs given in Table 3 are used to calculate the conflict factors for each pair of evidences. The conflict factors are $k_{1,2} = 0.10608$, $k_{1,3} = 0.26$, and $k_{2,3} = 0.12342$. Obviously, these conflict factors are all less than the conflict threshold $\delta = 0.4617$. The conflicts among the evidences are not significant any more. According to Eq. (18), new combined results can be obtained. If some of these conflict factors are bigger than the threshold, Eq.

Table 2
BPAs for the example

	$\{F_1\}$	$\{F_2\}$	$\{F_2, F_3\}$	Θ
E_1: $m_1(\cdot)$	0.6	0.1	0.1	0.2
E_2: $m_2(\cdot)$	0.05	0.8	0.05	0.1
E_3: $m_3(\cdot)$	0.7	0.1	0.1	0.1

Table 3
New BPAs for fault F_1

	$\{F_1\}$	$\{F_2\}$	$\{F_2, F_3\}$	Θ
E_1: $m_{1,\bullet}(\cdot)$	0.6	0.1	0.1	0.2
E_2: $m_{2,\bullet}(\cdot)$	0.0102	0.1632	0.0102	0.8164
E_3: $m_{3,\bullet}(\cdot)$	0.7	0.1	0.1	0.1

(19) will be used. The fused results are shown in the third row of Table 5.

In the same way, fused results based on the assumption that the fault is F_2 can be obtained. The second column in the matrix of importance index in Eq. (21) is used as follows:

$$\boldsymbol{V}^* = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{cccc} \{F_1\} & \{F_2\} & \{F_2,F_3\} & \Theta \end{array} \\ \begin{bmatrix} 0.11 & 0.11 & 0.11 & 0.11 \\ 1 & 1 & 1 & 1 \\ 0.11 & 0.11 & 0.11 & 0.11 \end{bmatrix} \end{array}. \tag{23}$$

The new BPAs incorporating the importance and sufficiency indexes are shown in Table 4. The combined results of all evidences are shown in the fourth row of Table 5.

Table 4
New BPAs for fault F_2

	$\{F_1\}$	$\{F_2\}$	$\{F_2,F_3\}$	Θ
E_1: $m_{1,\bullet}(\cdot)$	0.066	0.011	0.011	0.912
E_2: $m_{2,\bullet}(\cdot)$	0.03	0.480	0.03	0.46
E_3: $m_{3,\bullet}(\cdot)$	0.077	0.011	0.011	0.901

Table 5
Comparison between the results of D–S theory and the proposed method

Methods	No.	$m(\{F_1\})$	$m(\{F_2\})$	$m(\{F_2,F_3\})$	$m(\Theta)$
D–S theory	1	0.4519	0.5048	0.0336	0.0096
The proposed method (weights)					
[0.47 0.16 0.47]/$\{F_1\}$	2	**0.8119**	0.1096	0.0526	0.0259
[0.09 0.82 0.09]/$\{F_2\}$	3	0.0978	**0.4572**	0.0372	0.4078
[1 0.8 1]/$\{F_2,F_3\}$	4	0.8324	0.0871	**0.0543**	0.0261

Similarly, fused results are obtained for fault $\{F_2,F_3\}$ and shown in the fifth row of Table 5. For comparison purposes, results obtained by conventional D–S evidence theory are given in the second row of Table 5.

Fig. 6 is the bar plot of the BPAs in Table 5. In the first subplot in Fig. 6, the BPAs of $\{F_1\}$ and $\{F_2\}$, 0.4519 and 0.5048, obtained by the conventional D–S combination rule, are very close. Thus, it is hard to say which fault has occurred. The second, third, and fourth subplots in Fig. 6, concern only the BPAs of $\{F_1\}$, $\{F_2\}$, and $\{F_2,F_3\}$, respectively. The BPAs are 0.8119, 0.4572, and 0.0543, respectively. According to the proposed decision rules and the fused results, we compare only those BPAs, that are highlighted in Table 5. Although $m(\{F_2\}) = 0.5048$ is the largest BPA in the second row of Table 5, because $\max(\{m(\cdot)\} = m(\{F_2\}) = 0.5048 < \varDelta = 0.6$, we cannot conclude that the fault is F_2. As for the third–fourth rows of Table 5 or the seconnd–fourth subplots in Fig. 6, it is obvious that we can conclude that the fault is F_1 because $m(\{F_1\}) = 0.8119 > 0.6$, $Bel(\{F_1\}) = 0.8119$, and $Pl(\{F_1\}) = 0.8378$. The ignorance factor of F_1 is 0.0259, which is very small. It is clear that the proposed method can give more explicit diagnostic results than conventional D–S evidence theory.

6. Conclusions and summary

Membership functions and weights of features are introduced into the improvement of D–S evidence theory. The proposed method embodies information's uncertainty and expert's subjective knowledge in decision-making. An example shows that the method is valid and easy to use. From the results obtained in this paper, we conclude that

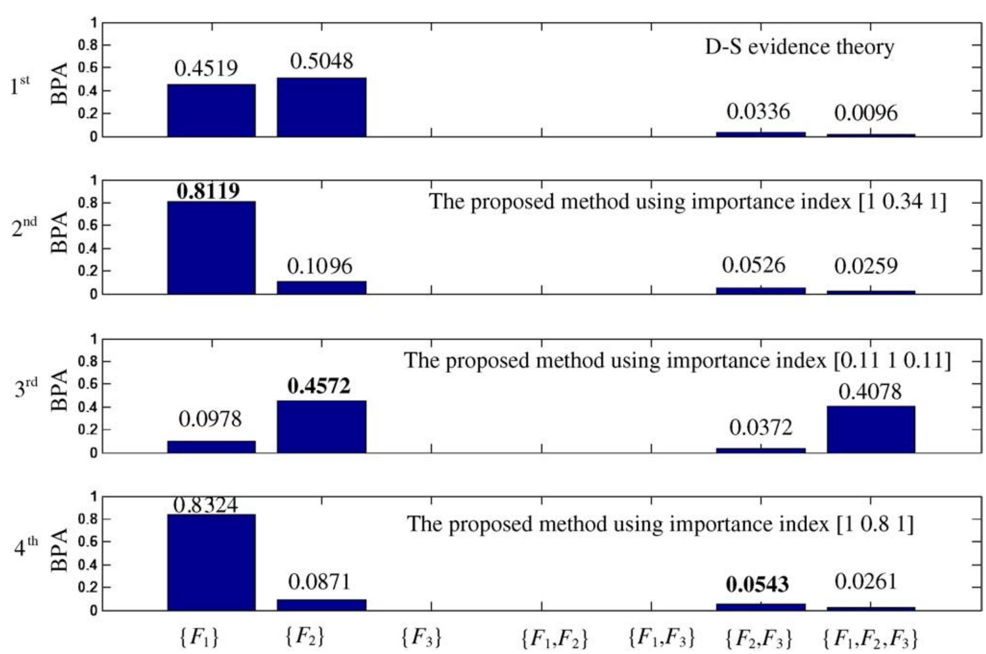

Fig. 6. The bar plot of the analyzed results using different methods.

the improved D–S evidence theory can not only resolve conflicts among evidences, but also combine expert knowledge and multi-source information. It can improve the accuracy of decision-making using multi-source information.

Acknowledgements

This research is supported by the Natural Science and Engineering Research Council of Canada (NSERC). Comments from the associate editor and the referees are appreciated very much.

References

Beynon, M., Cosker, D., Marshall, D., 2001. An expert system for multi-criteria decision making using Dempster–Shafer theory. Expert Syst. Appl. 20, 357–367.

Ducey, M.J., 2001. Representing uncertainty in silvicultural decisions: an application of the Dempster–Shafer theory of evidence. Forest Ecol. Manage. 150, 199–211.

Fabre, S., Appriou, A., Briottet, X., 2001. Presentation and description of two classification methods using data fusion based on sensor management. Inf. Fus. 2, 49–71.

Fan, X.F., 2003. Research of satellite fault diagnostic technology based on multi-agent system. Doctoral Dissertation. Harbin Institute of Technology. March.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian data analysis, second ed. Chapman & Hall/CRC.

Hégarat-Mascle, S.L., Richard, D., Ottlé, C., 2003. Multi-scale data fusion using Dempster–Shafer evidence theory. Integrated Comput. Aided Eng. 10, 9–22.

Huang, H.Z., 1997. Fuzzy multi-objective optimization decision-making of reliability of series system. Microelectron Reliab. 37 (3), 447–449.

Jones, R.W., Lowe, A., Harrison, M.J.A., 2002. Framework for intelligent medical diagnosis using the theory of evidence. Knowledge-Based Syst. 15, 77–84.

Kaftandjian, V., Dupuis, O., Babot, D., Zhu, Y.M., 2003. Uncertainty modeling using Dempster–Shafer theory for improving detection of weld defects. Pattern Recognition Lett. 24 (1–3), 547–564.

Klir, G.J., Wierman, M.J., 1998. Uncertainty-based information: elements of generalized information theory. Heidelberg, Physica-Verlag.

Kumar, A., Karmarkar, G., 2001. Aggregation of opinions using fuzzy numbers. Int. J. Syst. Sci. 32 (12), 1399–1411.

Kurkowska, S.P., Kostka, P., Gertych, A., Straszecka, E., Straszecka, J., 2000. Basic probability assignment calculation in diagnosis support. ESIT September, 14–15, Aachen, Germany: 92–94.

Parikh, C.R., Pont, M.J., Jones, N.B., 2001. Application of Dempster–Shafer theory in condition monitoring applications: a case study. Pattern Recognition Lett. 22 (6–7), 777–785.

Roemer, M.J., Kacprzynski, G.J., Orsagh, R.F., 2001. Assessment of data and knowledge fusion strategies for prognostics and health management. IEEE Aerospace Conf. Proc. 6 (10–17), 2979–2988.

Shafer, G., 1976. A mathematical theory of evidence. Princeton University Press, Princeton, NJ.

Salzenstein, F., Boudraa, A.O., 2004. Iterative estimation of Dempster–Shafer's basic probability assignment: application to multisensor image segmentation. Opt. Eng. 43 (6), 1293–1299.

Wang, P., Propes, N., Khiripet, N., Li, Y., Vachtsevanos, G., 1999. An integrated approach to machine fault diagnosis. IEEE Annual Textile Fiber and Film Industry Technical Conference, Atlanta, May 4–6.

Yager, R.R., 1987. On the Dempster–Shafer framework and new combination rules. Inf. Sci. 41 (2), 93–137.

Yager, R.R., 1999. A class of fuzzy measures generated from a Dempster–Shafer belief structure. Int. J. Intell. Syst. 14, 1239–1247.