

Article

# An Improved Evidential-IOWA Sensor Data Fusion Approach in Fault Diagnosis

Yongchuan Tang \*, Deyun Zhou \*, Miaoyan Zhuang, Xueyi Fang and Chunhe Xie

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China; zhuang-my@mail.nwpu.edu.cn (M.Z.); xyfang@nwpu.edu.cn (X.F.); xiechunhe@mail.nwpu.edu.cn (C.X.)

\* Correspondence: tangyongchuan@mail.nwpu.edu.cn (Y.T.); dyzhou@nwpu.edu.cn (D.Z.)

Received: 30 July 2017; Accepted: 14 September 2017; Published: 18 September 2017

**Abstract:** As an important tool of information fusion, Dempster–Shafer evidence theory is widely applied in handling the uncertain information in fault diagnosis. However, an incorrect result may be obtained if the combined evidence is highly conflicting, which may lead to failure in locating the fault. To deal with the problem, an improved evidential-Induced Ordered Weighted Averaging (IOWA) sensor data fusion approach is proposed in the frame of Dempster–Shafer evidence theory. In the new method, the IOWA operator is used to determine the weight of different sensor data source, while determining the parameter of the IOWA, both the distance of evidence and the belief entropy are taken into consideration. First, based on the global distance of evidence and the global belief entropy, the  $\alpha$  value of IOWA is obtained. Simultaneously, a weight vector is given based on the maximum entropy method model. Then, according to IOWA operator, the evidence are modified before applying the Dempster's combination rule. The proposed method has a better performance in conflict management and fault diagnosis due to the fact that the information volume of each evidence is taken into consideration. A numerical example and a case study in fault diagnosis are presented to show the rationality and efficiency of the proposed method.

**Keywords:** Dempster–Shafer evidence theory; belief entropy; distance of evidence; IOWA operator; fault diagnosis; sensor data fusion

## 1. Introduction

The structure of the modern engineering system is more and more complex [1,2], and how to maintain the safety of these systems is a critical problem. Various types of faults may occur because of long-time continuous operation and the changing environmental factors, which may bring great threats to human life [3–6]. Therefore, fault diagnosis plays an important role in real applications in daily life [7–10]. In practical applications, a multi-sensor system is widely used in fault diagnosis to make a comprehensive judgment [11–13]. For example, fault detection and isolation have been successfully used on the well known Airbus aircraft [14,15], which plays a key role in ensuring the safety of the aircraft [16,17]. However, the information, which may be obtained from a multi-sensor system, is heterogeneous and imprecision [18]. Therefore, it is essential that the uncertain information is pre-processed before data fusion and decision-making [19,20].

Information fusion is a key technology of combining multi-source information [21,22]. To address the uncertain information, some mathematical tools focused on data fusion have been introduced, such as fuzzy sets theory [23,24], Dempster–Shafer evidence theory [25–27], comprehensive optimization algorithm [28,29] and so on [30–32]. As an important tool in information fusion, Dempster–Shafer evidence theory is widely applied in fault diagnosis [33], pattern recognition [34–37], multi-criteria decision-making [38–40], risk analysis [41–44], controller design [45,46] and so on [47–49]. However, an incorrect result may be obtained if the combined evidence is highly conflicting. To handle this problem, many methods have been presented [50–52].

In the frame of Dempster–Shafer evidence theory, while dealing with the conflicting data fusion, one kind of method is to modify the conventional combination rule. Yager modifies Dempster’s combination rule through redistributing the conflicting evidence [53]. However, this method may destroy the good properties of Dempster’s combination rule, such as the commutativity and associativity. In addition, it is unreasonable to blame the combination rule if the incorrect results are caused by sensor failure. Another typical method is to modify the evidence before applying Dempster’s combination rule. Murphy’s method averages the evidence, which does not consider the difference among the evidence [54]. The distance of evidence is used to obtain the weight in Deng et al.’s method [55], which does remedy the disadvantage of Murphy’s method to a certain extent.

In this paper, an improved evidential-Induced Ordered Weighted Averaging (IOWA) sensor data fusion method is proposed in dealing with multi-sensor data fusion in fault diagnosis. Firstly, according to the global distance of evidence  $d_g$  and the global belief entropy  $E_d^g$ ,  $\alpha$  value of the maximum entropy method (MEM) is established. Namely, the  $\alpha$  value is jointly determined by  $d_g$  and  $E_d^g$ . Secondly, a weight vector  $W = (w_1, w_2, \dots, w_n)^T$  is generated based on the MEM model. After that, the evidence are modified by the new IOWA-based weight factor. Finally, the obtained evidence is combined  $(n - 1)$  times with Dempster’s combination rule. A numerical example and a case study on fault diagnosis verify the validity and reasonability of the proposed method.

This rest of this paper is organized as follows. The preliminaries are introduced in Section 2. In Section 3, a new evidential-IOWA sensor data fusion method is proposed. The application of the new method is presented in Section 4. Conclusions are given in Section 5.

## 2. Preliminaries

### 2.1. Dempster–Shafer Evidence Theory

Dempster–Shafer evidence theory was introduced by Dempster and then developed by Shafer, which is usually applied to manage the conflicting evidence [56,57].

Let  $\Theta$  be the frame of discernment, and be defined as  $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ . A basic probability assignment (BPA)  $m : 2^\Theta \rightarrow [0, 1]$ , is defined as follows [25,26]:

$$\begin{cases} m(\emptyset) = 0, \\ \sum_{A \subseteq \Theta} m(A) = 1. \end{cases} \quad (1)$$

when  $m(A) > 0$ ,  $A$  is called a focal element.

Suppose  $m_1$  and  $m_2$  are two BPAs on the frame of discernment  $\Theta$ , Dempster’s combination rule is defined as follows [25]:

$$m(A) = \begin{cases} \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - k}, & A \neq \emptyset, \\ 0, & A = \emptyset, \end{cases} \quad (2)$$

where  $k = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ , is regarded as a measure of conflict between  $m_1$  and  $m_2$ . The larger the  $k$ , the larger the degree of conflict.

### 2.2. Jousselme Distance

Jousselme distance is presented to measure of the difference—or the lack of similarity—between any two BPAs, which is introduced as follows.

Let  $m_1$  and  $m_2$  be two BPAs on the frame of discernment  $\Theta$ , then the distance between  $m_1$  and  $m_2$  is [58]:

$$d(m_1, m_2) = \sqrt{\frac{1}{2}(\vec{m}_1 - \vec{m}_2)^T \underline{D}(\vec{m}_1 - \vec{m}_2)}, \quad (3)$$

where  $\underline{D}$  is an  $2^{|\Theta|} \times 2^{|\Theta|}$  matrix whose elements are

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad A, B \in 2^\Theta. \quad (4)$$

### 2.3. Belief Entropy

Deng entropy is the generalization of Shannon entropy [59], which is defined as follows [60]:

$$E_d = - \sum_i m(B_i) \log_2 \frac{m(B_i)}{2^{|B_i|} - 1}, \quad (5)$$

where  $B_i$  is a proposition in the BPAs, and  $|B_i|$  is the cardinality of  $B_i$ .

The entropy can definitely degenerate to the Shannon entropy especially when the belief is only assigned to single element. Namely,

$$E_d = - \sum_i m(C_i) \log_2 \frac{m(C_i)}{2^{|C_i|} - 1} = - \sum_i m(C_i) \log_2 m(C_i), \quad (6)$$

and, for  $m_1(A) = \frac{2^{|A|} - 1}{\sum_{B \subseteq X} 2^{|B|} - 1}$ ,  $A, B \subseteq X$ ,  $m_1$  is the mass function having the maximum Deng entropy for the frame of discernment  $X = \{a, b, c\}$ , and its uncertainty can also be calculated by  $\sum_{B \subseteq X} \log_2(2^{|B|} - 1)$ .

### 2.4. IOWA Operator

The Induced Ordered Weighted Averaging (IOWA) operator [61], which is introduced by Yager and Filev, is a more general type of the Ordered Weighted Averaging (OWA) operator. An important feature of this operator is that the ordering of the arguments is induced by another variable called the order inducing variable.

Assume there are  $n$  two-tuple OWA pair  $\langle u_i, a_i \rangle$ ,  $i = 1, \dots, n$  that has an associated weight vector  $W = (w_1, w_2, \dots, w_n)^T$  of dimension  $n$  having the following properties:

$$\begin{aligned} 0 &\leq w_j \leq 1, \\ \sum_{j=1}^n w_j &= 1. \end{aligned} \quad (7)$$

Then, the IOWA operator is defined as follows [61]:

$$F_w(\langle u_1, a_1 \rangle, \dots, \langle u_i, a_i \rangle) = \sum_{j=1}^n w_j b_j, \quad (8)$$

where  $b_j$  is the  $a_i$  of the OWA pair having the  $j$ th largest  $u_i$ .  $u_i$  is referred as the order inducing variable and  $a_i$  is referred as the argument variable.

*orness*, which is associated with the weight vector  $W = (w_1, w_2, \dots, w_n)^T$ , is defined as follows:

$$\alpha = orness(W) = \frac{1}{n-1} \sum_{j=1}^n w_j (n-j), \quad (9)$$

where  $0 \leq orness \leq 1$ .

### 2.5. Maximum Entropy Method

To apply the IOWA operator in fault diagnosis, a very crucial issue is to determine its weight. The weight problem is denoted as a constrained nonlinear optimization model in the MEM model, which is presented by O'Hagan. The weight is gained by the following optimization model [62]:

$$\begin{aligned}
 & \text{Maximize : } \text{Disp}(W) = - \sum_{j=1}^n w_j \ln(w_j), \\
 & \text{S.t } \text{orness}(W) = \alpha = \frac{1}{n-1} \sum_{j=1}^n w_j(n-j), \\
 & \sum_{j=1}^n w_j = 1, \\
 & 0 \leq w_j \leq 1 \text{ for } j = 1, \dots, n.
 \end{aligned} \tag{10}$$

Suppose  $n = 5$  and the weights satisfy different degrees of *orness* :  $\alpha = 0, 0.1, \dots, 1$ , then the weight vector is determined by MEM model, which is shown in Figure 1.

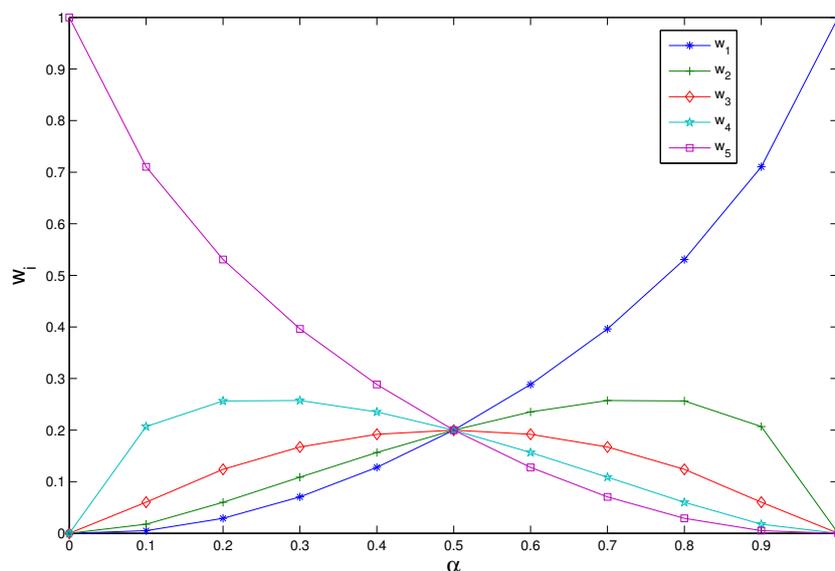
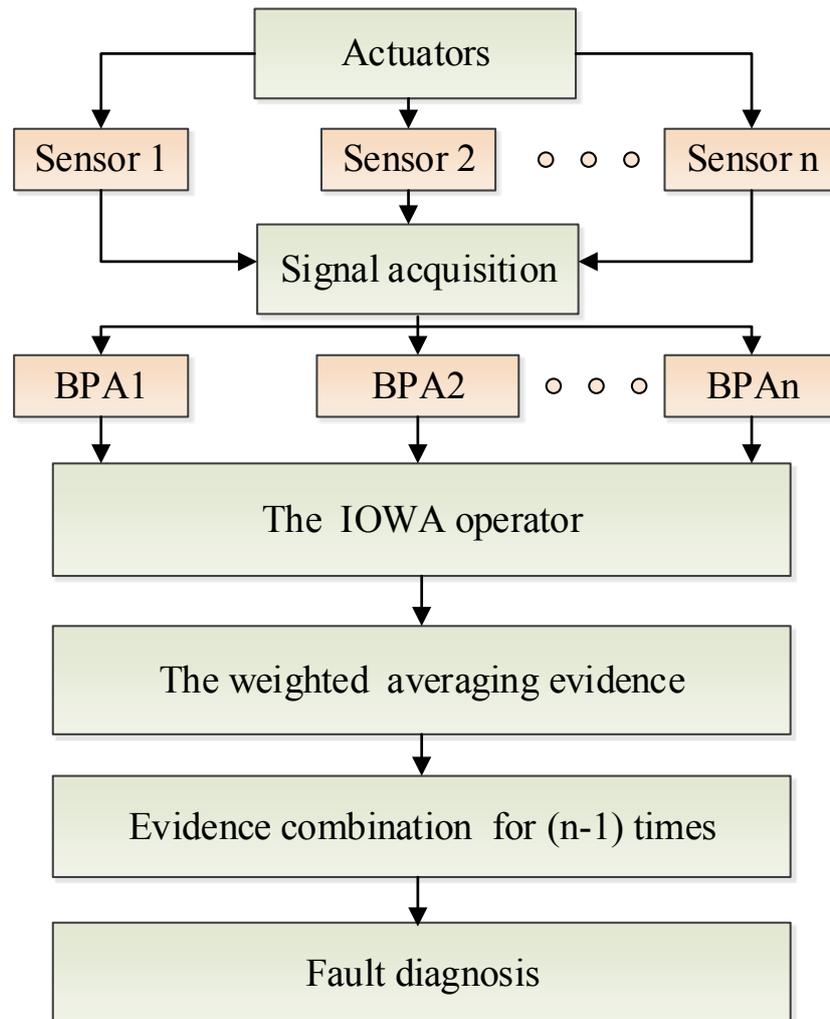


Figure 1. Variation of the weight with *orness* degree.

From Figure 1, we can conclude that: the value of the weight vector is closer to the average value  $W = (1/n, 1/n, \dots, 1/n)^T$ ; the value of  $\alpha$  is closer to  $\alpha = 0.5$ ; the value of the weight vector is closer to  $W = (1, 0, \dots, 0)^T$ ; the value of  $\alpha$  is closer to  $\alpha = 1$ . Namely, the smaller the credibility gap among BPAs, the more average for weight distribution.

### 3. The Evidential IOWA-Based Fault Diagnosis Method

As shown in Figure 2, in the fault diagnosis technique, typically, the first step should be information collecting from actuators. Secondly, all hypotheses are modelled (by BPAs in the frame of Dempster–Shafer evidence theory). Thirdly, the evidence is modified according to the IOWA operator. Finally, data fusion is applied for fault diagnosis and decision-making. Here, how to get an appropriate weight to modify the evidence is very important for locating the possible fault accurately. In the proposed method, the MEM model based on the distance of evidence and the belief entropy are used to generate the appropriate weight of evidence.



**Figure 2.** Overall structure of fault diagnosis based on sensor data fusion.

### 3.1. The Evidential-IOWA Parameter

Recently, the IOWA operator has aroused the attention of scholars and is widely used in real applications [63–65]. However, there are some problems while using the IOWA operator. For example, the  $\alpha$  value of a constraint condition usually depends on the experience of the experts, which does not lead to an objective result. In this paper, based on the the distance of evidence and the belief entropy, the  $\alpha$  value is induced as an objective weight.

#### 3.1.1. Definition of $\alpha$ in IOWA

The distance of evidence and the belief entropy are jointly considered to determine the  $\alpha$  value. The value of  $\alpha$  is defined as follows:

$$\alpha = \frac{1}{2}(\alpha_1 + \alpha_2) = \frac{1}{2}(e^{d_g \cdot \ln 0.5} + 0.5E_d^g), \quad (11)$$

where  $d_g$  is the global distance of evidence,  $E_d^g$  is the global belief entropy, and  $0 \leq d_g \leq 1$ ,  $0 \leq E_d^g \leq 1$ ,  $0.5 \leq \alpha \leq 1$ .  $\alpha_1$  is a data-driven value based on the distance of evidence, and  $\alpha_2$  is another data-driven value based on belief entropy.

### 3.1.2. Definition of $\alpha_1$ Based on the Distance of Evidence

Assume that there are many pieces of evidence for fault diagnosis. The Jousselme distances  $d_{ij}$ ,  $i, j = 1, 2, \dots, n$  between two evidence  $m_i$  and  $m_j$  can be calculated according to Equation (3), and the distance matrix (DM) is defined as follows:

$$DM = [d_{ij}] = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}. \quad (12)$$

The average distance of evidence of  $m_i$ ,  $i = 1, 2, \dots, n$ , with respect to the other evidence, denoted as  $\bar{d}_i$ , is defined as follows:

$$\bar{d}_i = \frac{\sum_{j=1, j \neq i}^n d_{ij}}{n-1}, \quad i = 1, 2, \dots, n, \quad (13)$$

then, the global distance of evidence among all the evidence  $d_g$  is defined as follows:

$$d_g = \frac{\sum_{i=1}^n \bar{d}_i}{n}, \quad i = 1, 2, \dots, n. \quad (14)$$

If the global distance of evidence  $d_g$  has a big value, the smaller the global similarity degree among the diagnosed results, the smaller the credibility degree of each sensor. In other words, the smaller the weight gap among the BPAs, the more average the weight distribution is, which means that the value of  $\alpha$  is closer to  $\alpha = 0.5$ . If  $d_g = 1$ , which means that the diagnosed fault type of multi-sensor is entirely different; in this case, the credibility degree of each evidence is the same with each other. Thus, the evidence should be assigned the same weight, namely, the weight vector is  $W = (1/n, 1/n, \dots, 1/n)^T$  and  $\alpha = 0.5$ .

Conversely, the smaller the value of  $d_g$ , the greater the global similarity degree of the diagnosed results, so the BPAs can be represented approximately by less or even one BPA with a high credibility degree. That is to say, the BPA with high credibility degree is given a greater weight and the BPA with a low credibility degree is given a small weight. Thus, the smaller the value of  $d_g$ , the more inequality of the weight distribution, which means the value of  $\alpha$  is closer to  $\alpha = 1$ . If  $d_g = 0$ , which means that the diagnosed results are similar, so the BPA can be represented by any BPAs. Considering the consistency of the algorithm, the initial weight is assigned as  $W = (1, 0, \dots, 0)^T$  and  $\alpha = 1$ .

Based on the above analysis, a relational formula of the degree of orness  $\alpha_1$  is defined as follows:

$$\alpha_1 = e^{d_g \cdot \ln 0.5}, \quad (15)$$

where  $d_g$  is the global distance of evidence, and  $0 \leq d_g \leq 1$ ,  $0.5 \leq \alpha_1 \leq 1$ .

### 3.1.3. Definition of $\alpha_2$ Based on the Belief Entropy

Deng entropy is an efficient tool to measure uncertainty, not only under the situation where the uncertainty is represented by a probability distribution, but also under the situation where the uncertainty is represented by the BPAs. Thus, this entropy is used to determine the  $\alpha$  value.

The global belief entropy  $E_d^g$  is defined as follows:

$$E_d^g = \frac{\sum_{i=1}^n E_{di}}{n \cdot (E_d)_{\max}}, \quad (16)$$

where  $E_{d_i}$  is the belief entropy of the evidence  $m_i$ .  $(E_d)_{\max}$  is the maximum belief entropy on the frame of discernment  $X$ , which is defined as:

$$(E_d)_{\max} = \log \sum_{B \subseteq X} (2^{|B|} - 1). \quad (17)$$

The greater the global belief entropy  $E_d^g$ , the greater the global uncertainty of the diagnosed faults. Therefore, the weight distribution should be more average, and the  $\alpha$  value is more close to 0.5. If  $E_d^g = 1$ , it shows that the diagnosed faults is entirely uncertainty, so they should be assigned to the same weight, that is,  $\alpha = 0.5$ .

The smaller the global belief entropy  $E_d^g$ , the smaller the global uncertainty of the diagnosed faults. Then, the BPA can be represented approximately by a few or even one BPA of relatively small uncertainty. Therefore, the smaller the  $E_d^g$ , the more inequality the weight distribution, the closer  $\alpha = 1$ . If  $E_d^g = 0$ , the BPA can be represented by any BPAs, that is to say, the weight vector is  $W = (1, 0, \dots, 0)^T$  and  $\alpha = 1$ .

Based on the above analysis, a relational formula of the degree of *orness*  $\alpha_2$  is defined as follows:

$$\alpha_2 = 0.5^{E_d^g}, \quad (18)$$

where  $E_d^g$  is the global belief entropy, and  $0 \leq E_d^g \leq 1$ ,  $0.5 \leq \alpha_2 \leq 1$ .

### 3.1.4. The Weight Vector of IOWA

After obtaining the parameters  $\alpha$ , the weight vector  $W = (w_1, w_2, \dots, w_n)^T$  can be obtained according to the MEM model. Assume that there are  $n$  BPAs  $m_i$ ,  $i = 1, 2, \dots, n$ , the weight vector  $W = (w_1, w_2, \dots, w_n)^T$  can be calculated according to the following steps:

- Step 1** According to Equations (14) and (15), the global distance of evidence  $d_g$  and the  $\alpha_1$  value can be calculated, respectively.
- Step 2** The global belief entropy  $E_d^g$  and the  $\alpha_2$  value are obtained by Equations (16) and (18), respectively.
- Step 3** The  $\alpha$  value and the weight vector  $W$  are calculated based on Equations (11) and (10), respectively.

### 3.2. Multi-Evidential Fusion Model

After getting an appropriate weight vector, the evidence can be modified before using Dempster's combination rule. The evidence are reordered according to the IOWA operator. Assume there are  $n$  BPAs, denoted as  $m_i$ ,  $i = 1, 2, \dots, n$ , the steps of ordering and evidence fusion are defined as follows:

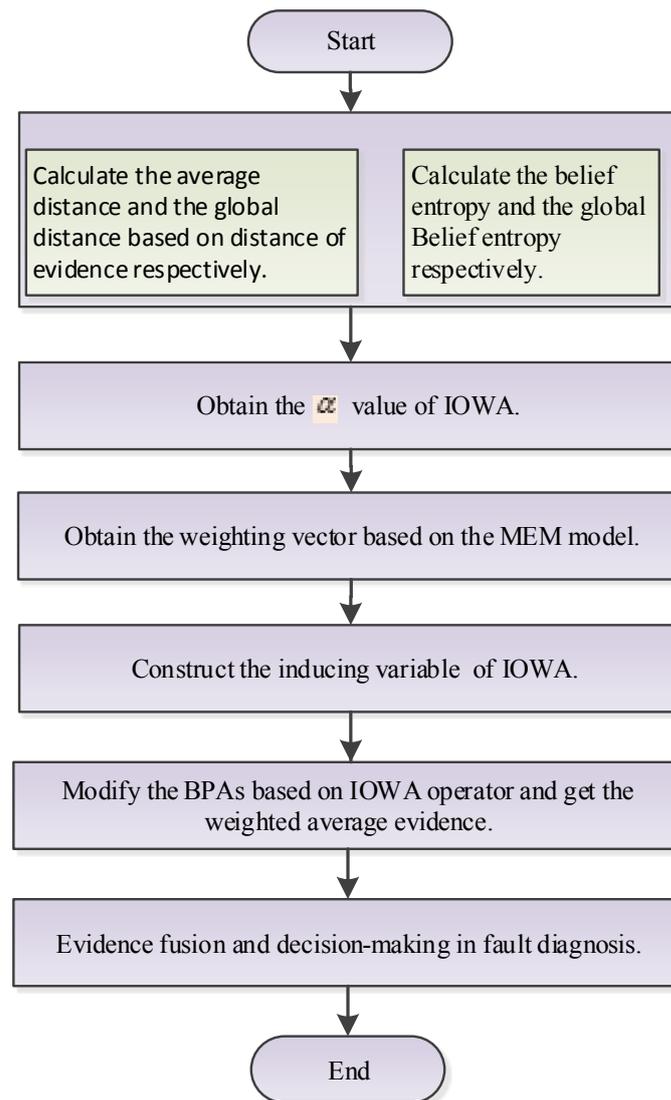
- Step 1** Construct the inducing variable  $S_i$ :

$$S_i = 1 - \bar{d}_i, \quad i = 1, 2, \dots, n, \quad (19)$$

where  $\bar{d}_i$  is the average distance of evidence obtained by Equation (13).

- Step 2** Obtain the OWA pairs  $\langle S_i, M_i \rangle$ ,  $i = 1, 2, \dots, n$ , where  $M_i$  is the argument variable, namely, it is the BPAs of the evidence  $m_i$ .
- Step 3** According to Equation (8), the weighted average evidence can be calculated.
- Step 4** Combine the new evidence with Dempster's combination rule by  $(n - 1)$  times.

With the fusion results, decision-making can be made based on the maximum principle of BPAs. An illustrative explanation of the new method is presented in Figure 3. Firstly, the degree of *orness*  $\alpha$  should be computed based on distance of evidence and belief entropy. Secondly, the weight vector  $W = (w_1, w_2, \dots, w_n)^T$  can be obtained based on the MEM model. Thirdly, a corresponding inducing variable can be constructed. Fourthly, evidence modification and fusion can be achieved. Finally, decision-making in fault diagnosis is based on the fused results.



**Figure 3.** The evidential Induced Ordered Weighted Averaging (IOWA)-based fault diagnosis method.

## 4. Application

### 4.1. Experiment with Artificial Data

This numerical example is used to illustrate how to apply the proposed method in fault diagnosis. Assume that, in the case of motor rotor fault diagnosis, vibration signal is collected by five sensors. There are three faults, denoted as  $A$ ,  $B$  and  $C$ , in motor rotor, which represents the unbalance, misalignment and pedestal looseness fault types, respectively. The BPAs based on these sensors are assumed to be independent and there are abnormal sensor reports, as is shown in Table 1. Intuitively,  $m_2$  comes from abnormal sensor report. Since evidence modelling is another open issue in Dempster–Shafer evidence theory, we do not discuss how to model data with BPAs in this paper. For more detail on how to generate BPAs, please refer to some related work such as [45,46,49].

**Table 1.** The basic probability assignment (BPA) as an example.

$m_i$	{A}	{B}	{C}	{AC}
$m_1$	0.41	0.29	0.30	0.00
$m_2$	0.00	0.90	0.10	0.00
$m_3$	0.58	0.07	0.00	0.35
$m_4$	0.55	0.10	0.00	0.35
$m_5$	0.6	0.10	0.00	0.30

According to the new method shown in Figure 3, firstly, with Equations (13) and (14), the average distance of evidence  $\bar{d}_i$ ,  $i = 1, 2, \dots, 5$  and the global distance of evidence  $d_g$  can be calculated, respectively, and the results are:  $\bar{d}_1 = 0.3456$ ,  $\bar{d}_2 = 0.6647$ ,  $\bar{d}_3 = 0.2661$ ,  $\bar{d}_4 = 0.2564$ ,  $\bar{d}_5 = 0.2641$  and  $d_g = 0.3594$ . With Equations (5) and (16), the belief entropy  $E_{di}$ ,  $i = 1, 2, \dots, 5$  and the global belief entropy  $E_d^g$  can be calculated, respectively, and the results are:  $E_{d1} = 1.5664$ ,  $E_{d2} = 0.4690$ ,  $E_{d3} = 1.8092$ ,  $E_{d4} = 1.8914$ ,  $E_{d5} = 1.7710$  and  $E_d^g = 0.3534$ .

Secondly, the degree of *orness*  $\alpha$  can be calculated by Equation (11):

$$\alpha = \frac{1}{2}(e^{0.3594 \cdot \ln 0.5} + 0.5^{0.3534}) = 0.7811. \quad (20)$$

The weight vector  $W = (w_1, w_2, w_3, w_4, w_5)^T$  is calculated according to Equation (10), and the result is:

$$W = (0.5026, 0.2592, 0.1337, 0.0689, 0.0356)^T. \quad (21)$$

In addition, the inducing variable  $S_i$ ,  $i = 1, 2, \dots, 5$  are calculated according to Equation (19):

$$\begin{aligned} S_1 &= 1 - 0.3456 = 0.6544, \\ S_2 &= 1 - 0.6647 = 0.3353, \\ S_3 &= 1 - 0.2661 = 0.7339, \\ S_4 &= 1 - 0.2564 = 0.7436, \\ S_5 &= 1 - 0.2641 = 0.7359. \end{aligned} \quad (22)$$

Thirdly, according to the ordering variable  $\bar{d}_i$ , the OWA pair  $\langle S_i, M_i \rangle$ ,  $i = 1, 2, \dots, 5$  are ordered as follows:

$$\begin{aligned} &\langle 0.7436, M_4 \rangle, \\ &\langle 0.7359, M_5 \rangle, \\ &\langle 0.7339, M_3 \rangle, \\ &\langle 0.6544, M_1 \rangle, \\ &\langle 0.3353, M_2 \rangle. \end{aligned} \quad (23)$$

Then, the BPAs on each fault are modified according to Equation (8), and the weighted average evidence is:

$$m(A) = 0.5411, \quad m(B) = 0.1325, \quad m(C) = 0.0217, \quad m(AC) = 0.3047. \quad (24)$$

Finally, combining the weighted average evidence with Dempster's combination rule by four times, the final results are shown as follows:

$$m(A) = 0.9914, \quad m(B) = 0.0001, \quad m(C) = 0.0025, \quad m(AC) = 0.0061. \quad (25)$$

In Table 2, we compare the results among several existing methods. It also shows the process of locating the fault type. With the new method, the belief in the fault diagnosis results that  $A$  is the fault type is 99.14%, which is not lower than the other methods.

**Table 2.** Comparison of several existing methods.

BPAs	Methods	$m(A)$	$m(B)$	$m(C)$	$m(AC)$	Faults
$m_1, m_2$	Dempster's method [25]	0	0.8969	0.1031	0	$B$
	Murphy's method [54]	0.0964	0.8119	0.0917	0	$B$
	Deng et al.'s method [55]	0.0964	0.8119	0.0917	0	$B$
	The proposed method	0.0964	0.8119	0.0917	0	$B$
$m_1, m_2, m_3$	Dempster's method [25]	0	0.6350	0.3650	0	$B$
	Murphy's method [54]	0.4939	0.4180	0.0792	0.0090	$A$
	Deng et al.'s method [55]	0.4974	0.4054	0.0888	0.0084	$A$
	The proposed method	0.6960	0.1744	0.1253	0.0056	$A$
$m_1, m_2, m_3, m_4$	Dempster's method [25]	0	0.3321	0.6679	0	$C$
	Murphy's method [54]	0.8362	0.1147	0.0410	0.0081	$A$
	Deng et al.'s method [55]	0.9089	0.0444	0.0379	0.0089	$A$
	The proposed method	0.9683	0.0020	0.0133	0.0163	$A$
$m_1, m_2, m_3, m_4, m_5$	Dempster's method [25]	0	0.1422	0.8578	0	$C$
	Murphy's method [54]	0.9620	0.0210	0.0138	0.0032	$A$
	Deng et al.'s method [55]	0.9820	0.0039	0.0107	0.0034	$A$
	The proposed method	0.9914	0.0001	0.0025	0.0061	$A$

In addition, if Dempster's combination rule is used directly, due to the conflicting evidence  $m_2$ , incorrect results are obtained. The same diagnosis results can be obtained according to the Murphy's method [54], Deng et al.'s method [55] and the proposed method. However, Murphy's method is only a simple arithmetic mean which does not consider the difference among the evidence, while Deng et al.'s method ignores the influence of evidence itself in generating the weight factor. The proposed method takes into consideration more available information before making data fusion and fault diagnosis, e.g., the distance of evidence and the belief entropy.

#### 4.2. A Case Study

In order to verify the effectiveness and success of the proposed evidential-IOWA sensor data fusion approach, the new method is applied to a case study adopted from [66].

Recall the fault diagnosis problem in [66]. Three potential fault types are denoted as  $F_1$ ,  $F_2$  and  $F_3$ ; thus, the fault hypothesis set is  $\Theta = \{F_1, F_2, F_3\}$ . Three sensors report the diagnosis results independently, the diagnosis results are modelled as three bodies of evidence, denoted as  $E_1$ ,  $E_2$  and  $E_3$ , and the BPAs of the diagnosis results are shown in Table 3. Intuitively,  $F_1$  is the fault type because both  $E_1$  and  $E_3$  have a belief of more than 60% on the fault type  $F_1$ , while the  $E_2$  may come from an abnormal sensor in comparison with the other two bodies of evidence. This is a challenge for data fusion, especially for some conventional combination rules, such as Dempster's rule of combination. The proposed method is applied to solve this problem.

**Table 3.** BPAs for fault diagnosis of the case study [66].

Sensor Report	$\{F_1\}$	$\{F_2\}$	$\{F_2, F_3\}$	$\Theta$
$E_1 : m_1(\cdot)$	0.60	0.10	0.10	0.20
$E_2 : m_2(\cdot)$	0.05	0.80	0.05	0.10
$E_3 : m_3(\cdot)$	0.70	0.10	0.10	0.10

According to the proposed method shown in Figure 3, the first step is to calculate the average distance and global distance of the evidence  $E_1$ ,  $E_2$  and  $E_3$ . Based on Equations (13) and (14), the calculation results of the average distance of each piece of evidence, denoted as  $\bar{d}_i(E_i)$ ,  $i = 1, 2, 3$ , and the global distance, denoted as  $d_g(E_i)$ ,  $i = 1, 2, 3$ , are shown in Table 4.

**Table 4.** The average distance and global distance of  $E_i$  ( $i = 1, 2, 3$ ).

Evidence Distance-Based Parameter	$\bar{d}_1(E_1)$	$\bar{d}_2(E_2)$	$\bar{d}_3(E_3)$	$d_g(E_i)$
Calculation Result	0.1916	0.3477	0.2033	0.3712

Then, based on Equations (5) and (16), the corresponding belief entropy, denoted as  $E_{di}(E_i)$ ,  $i = 1, 2, 3$ , and the global belief entropy, denoted as  $E_d^g(E_i)$ ,  $i = 1, 2, 3$ , are calculated in Table 5.

**Table 5.** The belief entropy and global belief entropy of  $E_i$  ( $i = 1, 2, 3$ ).

Belief Entropy-Based Parameter	$E_{d1}(E_1)$	$E_{d2}(E_2)$	$E_{d3}(E_3)$	$E_d^g(E_i)$
Calculation Result	2.2909	1.3819	1.7960	0.5884

With Equation (11), the degree of *orness*  $\alpha$  of the case study, denoted as  $\alpha(E_i)$ , is calculated as follows:

$$\alpha(E_i) = \frac{1}{2}(e^{0.3712 \cdot \ln 0.5} + 0.5^{0.5884}) = 0.7189. \quad (26)$$

According to the Maximum Entropy Method defined in Equation (10), the weight vector of the evidence, denoted as  $W(E_i) = (w_1, w_2, w_3)^T$ , can be calculated, and the result is

$$W(E_i) = (0.5771, 0.2836, 0.1393)^T. \quad (27)$$

The inducing variable, denoted as  $S_i(E_i)$  ( $i = 1, 2, 3$ ), can be calculated based on Equation (19) and the parameters in Table 4, and the results are shown as follows:

$$\begin{aligned} S_1(E_1) &= 1 - 0.1916 = 0.8084, \\ S_2(E_2) &= 1 - 0.3477 = 0.6523, \\ S_3(E_3) &= 1 - 0.2033 = 0.7967. \end{aligned} \quad (28)$$

Combining the inducing variables with the parameters in Table 5, the OWA pairs  $\langle S_i(E_i), E_i \rangle$ ,  $i = 1, 2, 3$ , are ordered as follows:

$$\begin{aligned} &\langle 0.8084, E_1 \rangle, \\ &\langle 0.7967, E_3 \rangle, \\ &\langle 0.6523, E_2 \rangle. \end{aligned} \quad (29)$$

Now, the BPAs in Table 3 can be modified according to Equation (8), and the weighted average evidence are as follows:

$$m(F_1) = 0.5517, \quad m(F_2) = 0.1975, \quad m(F_2, F_3) = 0.0930, \quad m(\Theta) = 0.1577. \quad (30)$$

Finally, combining the weighted average evidence with Dempster's combination rule by four times, the fusion results are as follows:

$$m(F_1) = 0.9123, \quad m(F_2) = 0.0810, \quad m(F_2, F_3) = 0.0027, \quad m(\Theta) = 0.0039. \quad (31)$$

The fused results with the proposed method are compared with the method in [66] where this case study comes from, and the comparison result is shown in Table 6.

**Table 6.** Fusion results with different methods.

Fault Types	$\{F_1\}$	$\{F_2\}$	$\{F_2, F_3\}$	$\Theta$
<b>Only Dempster's Rule of Combination</b>	0.4519	0.5048	0.0336	0.0096
<b>Fan et al's Method [66]</b>	0.8119	0.1096	0.0526	0.0259
<b>The Proposed Method</b>	0.9123	0.0810	0.0027	0.0039

It can be concluded from Table 6 that the proposed method has the most distinguishable fusion results on sensor reports, which means a clear indicator on the most possible fault type. The highest belief degree on fault type  $F_1$  is 91.23%, which is higher than the method with Fan et al's method with more than 10%. This is helpful for decision-making in real applications. While the fusion results of fault type  $F_1$  and  $F_2$  with the conventional Dempster's rule of combination are close to each other, it is hard to judge which fault has occurred. The case study verifies the effectiveness of the proposed method. In addition, the case study indicates a better performance of the proposed method in comparison with some of the existing methods.

#### 4.3. Discussion

The effectiveness of the proposed method is verified according to the applications based on both artificial data and the experiment adopted from the literature.

A few reasons contribute to the success of the new method. Firstly, not only the distance of evidence, but also the belief entropy and the IOWA operator are taken into consideration, which means more available information are used while doing information processing. Thus, information loss is decreased. Secondly, the way of getting the degree of *orness*  $a$  of IOWA (based on belief entropy and evidence distance) is data-driven, which is more reliable compared with some subjective methods. Finally, the final fused rule is based on Dempster's rule of combination. The merits of Dempster's rule of combination, such as satisfying the commutativity and associativity, contribute to the effectiveness of the proposed method.

In the fault diagnosis (FD) research area, an FD technique is good if a new method can guarantee that there is no false alarm, no missed detection and a full detection for all considered faulty scenarios [67,68]. The ongoing work of the proposed method should try to focus on this case. In future work, the following situations should be well addressed:

- FD without fault to be sure that the proposed solution doesn't give false alarm,
- FD with a misalignment fault to highlight that we detect this fault well,
- FD with pedestal fault.

## 5. Conclusions

In this paper, in the frame of Dempster–Shafer evidence theory, an improved evidential-IOWA sensor data fusion approach is proposed in dealing with a multi-source data-based fault diagnosis problem. Before applying sensor data fusion for final decision-making, the sensor data comes from different independent sources modelled, as BPA is pre-processed to avoid unreasonable fusion results that may be caused by conflicting evidence. In the new method, the IOWA operator is used to determine the weight of different sensor data sources, and the parameter of the IOWA is based on the distance of evidence and the belief entropy. The proposed method has a better performance in conflict management and fault diagnosis due to the fact that the information volume of each piece of evidence is taken into consideration. The proposed method outperforms the other methods according to the applications.

The ongoing work of the proposed method will be focused on some basic rules of fault diagnosis in industrial environmental scenarios, e.g., no missed detection and a full detection for all considered faulty scenarios should be strictly obeyed while applying the fault diagnosis technique.

**Acknowledgments:** The authors greatly appreciate the editor's encouragement and the two anonymous reviewers' valuable comments and suggestions to improve this paper. The work is partially supported by the National Natural Science Foundation of China (Grant No. 61671384), the Natural Science Basic Research Plan in Shaanxi Province of China (Grant No. 2016JM6018) and the Innovation Foundation for the Doctor Dissertation of Northwestern Polytechnical University (Grant No. CX201705).

**Author Contributions:** Y.T. and D.Z. designed and performed the research. Y.T. and M.Z. wrote the manuscript. Y.T., M.Z. and C.X. performed the computation. X.F. analyzed the data. All authors discussed the results and commented on the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

## References

1. Patecornell, M. Organizational aspects of engineering system safety: The case of offshore platforms. *Science* **1990**, *250*, 1210–1217.
2. Oosterom, M.; Babuska, R.; Verbruggen, H.B. Soft computing applications in aircraft sensor management and flight control law reconfiguration. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2002**, *32*, 125–139.
3. Picazo-Rodenas, M.J.; Royo, R.; Antonino-Daviu, J.; Roger-Folch, J. Use of the infrared data for heating curve computation in induction motors: Application to fault diagnosis. *Eng. Fail. Anal.* **2013**, *35*, 178–192.
4. Zhang, B.; Tan, A.C.C.; Lin, J.H. Gearbox fault diagnosis of high-speed railway train. *Eng. Fail. Anal.* **2016**, *66*, 407–420.
5. Du, Y.; Lu, X.; Su, X.; Hu, Y.; Deng, Y. New failure mode and effects analysis: An evidential downscaling method. *Qual. Reliab. Eng. Int.* **2016**, *32*, 737–746.
6. Chen, C.; Brown, D.; Sconyers, C.; Zhang, B.; Vachtsevanos, G.; Orchard, M.E. An integrated architecture for fault diagnosis and failure prognosis of complex engineering systems. *Expert Syst. Appl.* **2012**, *39*, 9031–9040.
7. Ferreira de Loza, A.; Cieslak, J.; Henry, D.; Zolghadri, A.; Fridman, L.M. Output tracking of systems subjected to perturbations and a class of actuator faults based on HOSM observation and identification. *Automatica* **2015**, *59*, 200–205.
8. Jiang, W.; Xie, C.; Zhuang, M.; Shou, Y.; Tang, Y. Sensor Data Fusion with Z-Numbers and Its Application in Fault Diagnosis. *Sensors* **2016**, *16*, 1509.
9. Su, X.; Deng, Y.; Mahadevan, S.; Bao, Q. An improved method for risk evaluation in failure modes and effects analysis of aircraft engine rotor blades. *Eng. Fail. Anal.* **2012**, *26*, 164–174.
10. Oukhellou, L.; Debiolles, A.; Denoeux, T.; Aknin, P. Fault diagnosis in railway track circuits using Dempster-Shafer classifier fusion. *Eng. Appl. Artif. Intell.* **2010**, *23*, 117–128.
11. Sun, S.; Lin, H.; Ma, J.; Li, X. Multi-sensor distributed fusion estimation with applications in networked systems: A review paper. *Inf. Fusion* **2017**, *38*, 122–134.
12. Pang, C.; Sun, S. Fusion Predictors for Multisensor Stochastic Uncertain Systems With Missing Measurements and Unknown Measurement Disturbances. *IEEE Sens. J.* **2015**, *15*, 4346–4354.
13. De Loza, A.F.; Cieslak, J.; Henry, D.; Davila, J.; Zolghadri, A. Sensor fault diagnosis using a non-homogeneous high-order sliding mode observer with application to a transport aircraft. *IET Control Theory Appl.* **2015**, *9*, 598–607.
14. Berdjag, D.; Cieslak, J.; Zolghadri, A. Fault diagnosis and monitoring of oscillatory failure case in aircraft inertial system. *Control Eng. Pract.* **2012**, *20*, 1410–1425.
15. Berdjag, D.; Zolghadri, A.; Cieslak, J.; Goupil, P. Fault detection and isolation for redundant aircraft sensors. In proceedings of 2010 Conference on Control and Fault-Tolerant Systems (SysTol), Nice, France, 6–8 October 2010; pp. 137–142.
16. Cieslak, J.; Henry, D.; Zolghadri, A.; Goupil, P. Development of an active fault-tolerant flight control strategy. *J. Guid. Control Dyn.* **2008**, *31*, 135–147.

17. Efimov, D.; Cieslak, J.; Zolghadri, A.; Henry, D. Actuator fault detection in aircraft systems: Oscillatory failure case study. *Annu. Rev. Control* **2013**, *37*, 180–190.
18. Li, C.; Mahadevan, S.; Ling, Y.; Choze, S.; Wang, L. Dynamic Bayesian Network for Aircraft Wing Health Monitoring Digital Twin. *AIAA J.* **2017**, *55*, 930–941.
19. He, Y.; Hu, L.F.; Guan, X.; Deng, Y.; Han, D. New method for measuring the degree of conflict among general basic probability assignments. *Sci. Chin.-Inf. Sci.* **2012**, *55*, 312–321.
20. Tang, Y.; Zhou, D.; Xu, S.; He, Z. A Weighted Belief Entropy-Based Uncertainty Measure for Multi-Sensor Data Fusion. *Sensors* **2017**, *17*, 928.
21. Zhang, L.; Yang, F.B.; Ji, L.; Yuan, H.; Dong, A. A categorization method of infrared polarization and intensity image fusion algorithm based on the transfer ability of difference features. *Infrared Phys. Technol.* **2016**, *79*, 91–100.
22. Wang, Z.; Yang, F.; Peng, Z.; Chen, L.; Ji, L. Multi-sensor image enhanced fusion algorithm based on NSST and top-hat transformation. *Optik Int. J. Light Electron Opt.* **2015**, *126*, 4184–4190.
23. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353.
24. Li, D.W.; Yang, F.; Wang, X.X. Crop region extraction of remote sensing images based on fuzzy ARTMAP and adaptive boost. *J. Intell. Fuzzy Syst.* **2015**, *29*, 2787–2794.
25. Dempster, A.P. Upper and Lower Probabilities Induced by a Multi-valued Mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339.
26. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976.
27. Deng, Y. Generalized evidence theory. *Appl. Intell.* **2015**, *43*, 530–543.
28. Zhang, X.; Mahadevan, S. A Bio-Inspired Approach to Traffic Network Equilibrium Assignment Problem. *IEEE Trans. Cybern.* **2017**, *PP*, 1–12.
29. Guo, H.; Huang, H.; Liang, L. Fuzzy-robust design optimization with multi-quality characteristics. *Eng. Optim.* **2010**, *42*, 87–99.
30. Sun, S.; Tian, T.; Lin, H. State estimators for systems with random parameter matrices, stochastic nonlinearities, fading measurements and correlated noises. *Inf. Sci.* **2017**, *397*, 118–136.
31. Miao, B.; Dou, C.; Jin, X. Main Trend Extraction Based on Irregular Sampling Estimation and Its Application in Storage Volume of Internet Data Center. *Comput. Intell. Neurosci.* **2016**, *2016*, 9328062.
32. Jin, X.B.; Du, J.J.; Bao, J. Target Tracking of a Linear Time Invariant System Under Irregular Sampling. *Int. J. Adv. Rob. Syst.* **2012**, *9*, 219.
33. Yuan, K.; Xiao, F.; Fei, L.; Kang, B.; Deng, Y. Modeling Sensor Reliability in Fault Diagnosis Based on Evidence Theory. *Sensors* **2016**, *16*, 113.
34. Liu, Z.G.; Pan, Q.; Dezert, J.; Mercier, G. Credal classification rule for uncertain data based on belief functions. *Pattern Recognit.* **2014**, *47*, 2532–2541.
35. Han, D.; Liu, W.; Dezert, J.; Yang, Y. A novel approach to pre-extracting support vectors based on the theory of belief functions. *Knowl.-Based Syst.* **2016**, *110*, 210–223.
36. Liu, Z.G.; Pan, Q.; Dezert, J.; Martin, A. Adaptive imputation of missing values for incomplete pattern classification. *Pattern Recognit.* **2016**, *52*, 85–95.
37. Zhou, K.; Martin, A.; Pan, Q.; Liu, Z. ECMdd: Evidential c-medoids clustering with multiple prototypes. *Pattern Recognit.* **2016**, *60*, 239–257.
38. Ghorabae, M.K.; Zavadskas, E.K.; Amiri, M.; Antucheviciene, J. A New Method of Assessment Based on Fuzzy Ranking and Aggregated Weights (AFRAW) for MCDM Problems under Type-2 Fuzzy Environment. *Econ. Comput. Econ. Cybern. Stud. Res.* **2016**, *50*, 39–68.
39. Chin, K.S.; Fu, C.; Wang, Y. A method of determining attribute weights in evidential reasoning approach based on incompatibility among attributes. *Computer. Ind. Eng.* **2015**, *87*, 150–162.
40. Fu, C.; Wang, Y. An interval difference based evidential reasoning approach with unknown attribute weights and utilities of assessment grades. *Computer. Ind. Eng.* **2015**, *81*, 109–117.
41. Fu, C.; Yang, J.B.; Yang, S.L. A group evidential reasoning approach based on expert reliability. *Eur. J. Oper. Res.* **2015**, *246*, 886–893.
42. Zhang, X.; Mahadevan, S.; Deng, X. Reliability analysis with linguistic data: An evidential network approach. *Reliab. Eng. Syst. Saf.* **2017**, *162*, 111–121.
43. Su, X.; Mahadevan, S.; Xu, P.; Deng, Y. Dependence Assessment in Human Reliability Analysis Using Evidence Theory and AHP. *Risk Anal.* **2015**, *35*, 1296–1316.

44. Zhou, D.; Tang, Y.; Jiang, W. A modified model of failure mode and effects analysis based on generalized evidence theory. *Math. Probl. Eng.* **2016**, *2016*, 4512383.
45. Yager, R.R.; Filev, D.P. Including probabilistic uncertainty in fuzzy logic controller modeling using Dempster–Shafer theory. *IEEE Trans. Syst. Man Cybern.* **1995**, *25*, 1221–1230.
46. Tang, Y.; Zhou, D.; Jiang, W. A New Fuzzy-Evidential Controller for Stabilization of the Planar Inverted Pendulum System. *PLoS ONE* **2016**, *11*, e0160416.
47. Zhou, K.; Martin, A.; Pan, Q.; Liu, Z.G. Median evidential c-means algorithm and its application to community detection. *Knowl.-Based Syst.* **2015**, *74*, 69–88.
48. Ma, J.; Liu, W.; Miller, P.; Zhou, H. An Evidential Fusion Approach for Gender Profiling. *Inf. Sci.* **2015**, *333*, 10–20.
49. Xu, P.; Su, X.; Mahadevan, S.; Li, C.; Deng, Y. A non-parametric method to determine basic probability assignment for classification problems. *Appl. Intell.* **2014**, *41*, 681–693.
50. Smets, P. Data fusion in the transferable belief model. *Inf. Fusion* **2000**, *1*, 21–33.
51. Lefevre, E.; Colot, O.; Vannoorenberghe, P. Belief functions combination and conflict management. *Inf. Fusion* **2002**, *3*, 149–162.
52. Tang, Y.; Zhou, D.; He, Z.; Xu, S. An improved belief entropy-based uncertainty management approach for sensor data fusion. *Int. J. Distrib. Sens. Netw.* **2017**, *13*, doi:10.1177/1550147717718497.
53. Yager, R.R. On the Dempster–Shafer framework and new combination rules. *Inf. Sci.* **1987**, *41*, 93–137.
54. Murphy, C.K. Combining belief functions when evidence conflicts. *Decis. Support Syst.* **2000**, *29*, 1–9.
55. Deng, Y.; Shi, W.; Zhu, Z.; Liu, Q. Combining belief functions based on distance of evidence. *Decis. Support Syst.* **2004**, *38*, 489–493.
56. Chin, K.S.; Fu, C. Weighted cautious conjunctive rule for belief functions combination. *Inf. Sci.* **2015**, *325*, 70–86.
57. Jiang, W.; Zhuang, M.; Qin, X.; Tang, Y. Conflicting evidence combination based on uncertainty measure and distance of evidence. *SpringerPlus* **2016**, *5*, 1217.
58. Jousselme, A.L.; Grenier, D.; Bosse, E. A new distance between two bodies of evidence. *Inf. Fusion* **2001**, *2*, 91–101.
59. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
60. Deng, Y. Deng entropy. *Chaos Solitons Fractals* **2016**, *91*, 549–553.
61. Yager, R.R.; Filev, D.P. Induced ordered weighted averaging operators. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **1999**, *29*, 141–150.
62. O’Hagan, M. Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. *Sign. Syst. Comput.* **1988**, *29*, 681–689.
63. Zeng, S.; Merigo, J.M.; Su, W. The uncertain probabilistic owa distance operator and its application in group decision making. *Appl. Math. Modell.* **2013**, *37*, 6266–6275.
64. Cho, S.B. Fuzzy aggregation of modular neural networks with ordered weighted averaging operators. *Int. J. Approx. Reason.* **1995**, *13*, 359–375.
65. Peng, Y.; Zhang, Y.; Tang, Y.; Li, S. An incident information management framework based on data integration, data mining, and multi-criteria decision making. *Decis. Support Syst.* **2011**, *51*, 316–327.
66. Fan, X.; Zuo, M.J. Fault diagnosis of machines based on D-S evidence theory. Part 1: D–S evidence theory and its improvement. *Pattern Recognit. Lett.* **2006**, *27*, 366–376.
67. Henry, D.; Cieslak, J.; Zolghadri, A.; Efimov, D.  $H_\infty/H$  - LPV solutions for fault detection of aircraft actuator faults: Bridging the gap between theory and practice. *Int. J. Robust Nonlinear Control* **2015**, *25*, 649–672.
68. Cieslak, J.; Efimov, D.; Zolghadri, A.; Henry, D.; Goupil, P. Design of a non-homogeneous differentiator for actuator oscillatory failure case reconstruction in noisy environment. *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.* **2015**, *229*, 266–275.

