JIEMS
Journal of Industrial Engineering and Management Studies

Vol. 8, No. 1, 2021, pp. 114-134

DOI: 10.22116/jiems.2020.233227.1364

www.jiems.icms.ac.ir

Machine fault diagnosis through entropy weighting using evidential reasoning approach

Mehran Khalaj[1,*], Fereshteh Khalaj[1]

Abstract
This paper presents an approach for the fault diagnosis in the state of fault in a machine by using a combination of the Dempster–Shafer (D-S) theory. At the first, feature extractions in each state have been combined based on evidential reasoning (ER) using kind of sensor information such as vibration, acoustic, pressure, and temperature, to detect and diagnose machine failure. Then, the main fusion will be obtained. In this process, the mass function assignment of any sensors to feature extraction, respectively, in every state of the machine is fused to indicate state quality. Within this framework, we propose a new way for main fusion to derive a consensus decision for fault diagnosis. In this paper, an approach developed to apply the evidential reasoning by defining adaptively weights into the improvement of the D–S evidence theory instead of the probability theory and the D–S evidence theory alone. Instead of using the evidential reasoning approach, this new approach applies entropy weighting in the D-S theory, in which all available data are used for making a decision. Entropy weighting can measure the uncertainty level of the fault decision and assist in obtaining a less uncertain fault decision. It is defined adaptively weights based on ambiguity measures associated with information obtained from each sensor. The ambiguity measure is defined by Shannon's entropy. Many industries use old machines due to cost savings or lack of purchasing power. Maintenance policies in these factories are based on determining their fault experimentally and traditionally. Therefore, the main goal of this paper uses the improved evidence reasoning algorithm using a kind of sensor information to carry out fault diagnosis in these industrials. Then, a numerical example and a case study involving the ball mill machine in fault diagnosis are presented to show the rationality and efficiency of the proposed method.

Keywords: evidential reasoning algorithm (ER algorithm); information fusion; fault diagnosis; multi-sensor fusion; Shannon entropy.

Received: May 2020-30

Revised: December 2020-10

Accepted: December 2020-28

1. Introduction

The performance of structures, the principle of machines, and mechanical systems deteriorate during their service life. Therefore, the ability to fault diagnosis of these systems is becoming

* Corresponding author; mekhalaj1@gmail.com
[1] Department of Industrial Engineering, Robat Karim Branch, Islamic Azad University, Tehran, Iran.

increasingly important from both economic and life-safety viewpoints. A single sensor cannot reliably obtain all the information required for fault diagnosis; thus, machine diagnostics are typically a multi-sensor fusion problem. Data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of greater quality will depend upon the application. In a process of fault diagnosis, according to the different stages of component testing, the method of the fault diagnosis revolves around two main questions (Basir and Yuan 2007); (1) How to acquire precise and reliable information cues about potential faults by incorporating complementary, and possibly redundant, multiple sensors. (2) How to fuse decisions that are derived based on multi-sensor data, which can be imprecise, and conflicting. There has been a substantial amount of research work conducted in the area of decision fusion about making more reasonable inferences based on sensor information. Most of these methods to fault diagnosis are built around of artificial neural network (ANN) (Cheng et al., 2016; Torabi et al., 2016, Yu et al. 2015; Besma et al. 2015; Guedidi et al. 2013), Bayesian inference (Huang et al. 2008; Wang et al. 2016; Cai et al. 2016; Yu 2013), fuzzy logic inference (Laala et al. 2013; Yaghoubi and Amiri, 2015; Rodríguez Ramos et al. 2017; Hang et al. 2016; Zhang et al. 2017) and D-S theory (Oukhellou and Debiolles, 2010; Bhalla et al. 2013, Hoou et al. 2017, Moosavian et al. 2015) and the other methods and application (Mabrouk et al. 2015, Sakhara et al. 2016).

Although there is a lot of fault diagnosis method, these lead to little attention to the consequences of uncertainty. Each traditional technique has been developed based on utilities for fault diagnosis. For example, ANN, in general, has been used to deal with a classification problem in fault diagnosis (Bhalla et al. 2013) and fuzzy inference has been developed to handle perception-based imprecise information (Verbert et al. 2017). Such methods, however, are adopted assumptions regarding the uncertainty, but they are not clearly and adequately express the uncertainty in the process of fault diagnosis (Verbert et al. 2017). Among them, the D-S theory provided a technique to fuse evidence that is obtained from various sources. The D-S theory is an extension of Bayes' theory that is frequently used as a method for dealing with uncertain information (Shafer, 1976). Though Bayesian inference can be employed to determine the probability of the correctness of a decision 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; (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 (Fan et al. 2006, Petersen et al. 2000, Kaftandjian et al. 2003, Parikh et al. 2001) reported the application of the D-S theory in fault diagnosis and defect inspection.

Meanwhile, the evidential reasoning approach is established based on both the D-S theory and decision-theory (Jian-Bo and Singh 1994) and has been extensively developed in all kinds of industries to solve multi-attribute decision-making (MADM) problems under uncertainty (e.g. Guo et al. 2008, Tang and Wu 2011, Singh et al. 2016,). However, most existing ER approach about fault diagnosis has affected significant advancements, but none of the above approaches are considered the relative importance of uncertainty subsequently in a different stage of algorithm deal with incomplete information in the decision-making process (such as; uncertainty about machine feature extractions, total uncertainty in a main fusion of multi-sensor, etc.). Since available data, which are resulted from multi-sensor may be incomplete, imprecise, or incorrect, as well as, human errors of each expert who have evaluated the system's performance, are usually led to the uncertain information and knowledge relating available data, this paper is focused on a method which is considered a different level of uncertainty in

a fault diagnosis decision-making process to minimize the uncertainty and increasing identification accuracy using ER algorithm.
Thus, the objective of this paper is to propose an improved evidential method, which is conceptually simple, and yet can provide much better accuracy and less uncertainty. Therefore, a new method for a fault diagnosis problem is demonstrated on the states of the ball mill machine using multi-sensor measurements under uncertainty. The basic idea is to obtain the appropriate weights for different reports. So, the difference between the two BPAs is measured by belief function. Also, the diversity degree among BPAs can be obtained by combining evidence. According to it, the weight of each BPA can be determined. Finally, we can decide on fault diagnosis by using the Dempster combination rule. An application in fault diagnosis and an example show that our proposed approach can not only increase the accuracy of fault diagnosis but also decrease the uncertain information volume, which is more reasonable. Because of the effectiveness of the D-S theory and ER algorithm to handle uncertainty and to determine the fault states in fault decision-making problems, the proposed method used this theory when the obtained information is incomplete using a multi-sensor. In the proposed method, a new multi-sensor data fusion technique is introduced based on the D-S theory, evidential reasoning, and entropy weighting to diagnose machine failure. Each sensor measurement is considered as a piece of evidence that reveals some information about the state of the machine, also, sensors' relative importance is considered. It seems that the fault identification method based on the weighted and selective information fusion technique can improve fault identification accuracy in comparison with evidential reasoning (ER) algorithm alone and direct information fusion techniques when wide uncertainty is considered to fault decision and the increasing identification accuracy.
A new method is used in a real case study to determine the reliability and confidence extent of the fault diagnosis decision when the available data are insufficient, helping us to make decisions. After calculating the basic probability of machine states with assigned weights and aggregate probability mass of feature extractions in each state, main fusion based on entropy weighting can be calculated, the result shows that the uncertainty of fault diagnosis decision for states of the machine is reduced while the information fusion technique can improve fault identification accuracy in comparison with direct information fusion techniques. Based on the uncertainty quantification (UQ) framework in an engineering system, evidential reasoning (ER) approach has been developed for supporting such decision analysis. This paper is organized as follows. In section 2, we introduce some concepts of the D-S evidence theory. Section 3 describes the ER algorithms. In this algorithm, we are going to fuse the feature extractions from each state of the machine. The process is described as three steps. In section 4, the main fusion technique based on entropy weighting is proposed. Section 5 describes a numerical example in which we demonstrate the effectiveness of the proposed fault detection approach. Conclusions and summary are presented at the end.

2. literature review

In a mechanical machine, which might have many components, some systems are very complex, reflecting with each other (Yang et al. 2018). when something happens unexpectedly in the systems and causes various problems due to different reasons, such as bad condition or environment, human mistake, or a long time of working. Thus, multi-sensors can be reported information that is extremely significant to make a reasonable decision in fault diagnosis (Zhang and Deng. 2018). To make a logical decision when using multi-sensor and data fusion methods, some works have been proposed to handle uncertainty (Rong et al., 2018; Meng et al. 2018), such as fuzzy set theory (Xiao and Ding. 2018; Zhang and Deng. 2018), Z numbers (Kang et al. 2019), D numbers (Xiao, 2018), R numbers (Seiti and Hafezalkotob, 2019) and so on. One of the most used math tools in sensor data fusion is evidence theory (Shafer, 1967). Uncertain information can be modeled with basic probability assignments efficiently rather than many methods (Han and Deng, 2018b). Also, the Dempster rule can efficiently combine the sensor reports from different sources (Zhang et al. 2017). Evidence theory is known as the de facto standard in decision making, due to the desirable properties (Deng et al. 2019) in various research such as; risk and reliability analysis (Meng et al. 2019), system optimization (Xu et al. 2016), and pattern recognition (Han and Deng, 2018). However, an open issue of evidential sensor data fusion is that the illogical results will be obtained when sensor reports conflict with each other to a high degree (Sun and Deng, 2019). Many methods were presented to address this issue (Wang et al. 2018). For example, Smets and Kennes (1994) defined the conjunctive and disjunctive rules, Yager (1987) presented the process of normalizing in D-S combination rule, Murphy (2000) combined the conflicting evidence with the average operation, Fan and Zuo (2006) proposed a combination method to fuse conflicting evidence in fault diagnosis, Dubois and Prades method (1988), Lefevre et al. (2002) and so on. Jiang et al. (2016) applied belief entropy into sensor data fusion and received the best performance.

These methods have some advantages and disadvantages, for example; one of the disadvantages is that some information is not fully used, Jiang et al. (2016) proposed an evidential sensor fusion method in fault diagnosis but they are not considering the distance information and the difference of information. In this paper, we tried to propose an improved evidential method, which is conceptually simple, and yet can provide much better accuracy and less uncertainty.

The basic idea is to obtain the appropriate weights for different reports. The difference between the two BPAs is measured by belief entropy (Deng, 2016). Then the diversity degree among BPAs can be obtained by combining evidence that is the difference (Jousselme et al., 2001). According to it, the weight of each BPA can be determined by entropy. we can decide for fault diagnosis by using the Dempster combination rule. An application in fault diagnosis and an example show that our proposed approach can not only increase the accuracy of fault diagnosis but also decrease the uncertain information volume, which is more reasonable. According to the proposed method, a real application in fault diagnosis and an example are given as follow section. Furthermore, in fault diagnosis and other sensor data fusion systems, the reports of different sensors may be influenced by some complex environments, leading them less reliable. Therefore, how to efficiently determine the reliability of each report, or to say, the weight of each report is very important. To address this issue, we propose an improved method based on the belief of entropy and uncertain evidence. The method considers both the degree of conflict and the difference of information volume among the evidence. An application and an example illustrate the efficiency of the method in evidential sensor fusion. It shows that the proposed method is more efficient for highly conflicting evidence with better performance of convergence and less uncertainty, compared with the existing methods. Some related advantages and disadvantages of different methods are discussed as follows.

Table 1. Some studies conducted by researchers

Research	Reference	Field and subject	Description
Dempster's method	(Dempster, 1968)	Deal with imprecise and uncertain information is widely used in fusing information.	When it comes to highly conflicting evidence, the method will always lead to some illogical results.
Murphy's method	(Murph, 2000)	Uses an average operation to combine the conflicting evidence, which can deal with highly conflicting evidence to some extent.	The difference and relationship of evidence are neglected.
Deng's method	(Deng et al. 2004)	It is different from Jiang's method; evidence distance is used to calculate. The weight rather than the belief entropy.	The similarity between the two pieces of evidence is proposed.
Jiang's method	(Jiang et al. 2016)	Belief entropy is used to calculate the weight of each evidence. It considers the difference in evidence.	makes it more reasonable than Murphy's method
The proposed method	-	Considering both belief entropy and evidence distance, which has a better result. Although it might be not very suitable in some situations	It leverages the advantages of Jiang's method and Deng's method.

3. D-S (evidence) theory

The D-S theory was initially introduced by Dempster (1968) and then Shafer showed the benefits of belief functions for modeling uncertain knowledge (Shafer 1976). The D-S theory allows one to combine evidence from different sources and arrive at a degree of belief that takes into account all the available evidence. It is represented by a belief and plausibility function. The probability in A falls somewhere between $Bel(A)$ and $Pl(A)$. $Bel(A)$ represents the evidence, we have for A directly. Therefore, the probability (A) cannot be less than this value. $Pl(A)$ represents the maximum share of the evidence, which can be assigned for all sets that intersect with A. In this section, some mathematical elements of the D-S Theory are recalled.

3.1. Basic concepts

Let $\Theta = \{A_1, A_2 \ldots, A_N\}$ be a frame of discernment, in which all elements are assumed to be mutually exclusive and exhaustive. The power set of Θ is denoted by $2^\theta = \{A | A \subseteq \theta\}$ Basic Probability Assignment (BPA) is a function that can be mathematically defined by 2^θ in $[0,1]$, such that, $m(\phi)=0$ where ϕ denotes an empty set, and $\sum_{A\subseteq\theta} m(A)=1$. The belief function ($Bel$) and the plausibility function (Pl) are defined as follows:

$$
\begin{aligned}
Bel(A) &= \sum_{\phi \neq B \subseteq A} m(B) \qquad \forall A \subseteq \theta \\
Pl(A) &= \sum_{B \cap A \neq \phi} m(B) \qquad \forall A \subseteq \theta
\end{aligned} \tag{1}
$$

which $Bel(A)$ represents the sum of masses in all subsets of A, whereas, $Pl(A)$ corresponds to the sum of masses committed to those subsets which don't discredit A (F. Khalaj and M. Khalaj, 2020; M. khalaj et al., 2020).

3.2. Combination of belief functions

Multiple evidence can be fused by using Dempster's combination rules, shown in equation (2), which also is known as the orthogonal sum. This sum is both commutative and associative.

$$m(\phi)=0$$
$$m(A)=\frac{1}{1-k}\sum_{B\cap C=A} m_1(B)\cdot m_2(C) \tag{2}$$

with

$$k=\sum_{B\cap C=A} m_1(B)\cdot m_2(C)>0 \tag{3}$$

where the term k is called the conflict factor between two pieces of evidence, it reflects the conflict degree between them.

3.3. Evidence reliability

When the information is provided by sensors, which are not reliable to result in the belief functions, a coefficient α is used to discount the belief. This coefficient will transfer the belief into the set θ. Thus, the discounted belief function m^{α} can be obtained by the following formula (Shafer 1976)

$$\begin{cases} m^{\alpha}(A)=\alpha\times m(A) & A\subset\theta \\ m^{\alpha}(\theta)=(1-\alpha)+\alpha\times m(\theta) \end{cases} \quad \text{where} \quad \alpha\in[0,1] \tag{4}$$

4. Feature extractions fusion using ER algorithm

As (Jian-Bo and Xu 2002, Yang 2001) many multiple attributes decision analysis (MADA) problems are characterized by both quantitative and qualitative attributes with various types of uncertainties. Incompleteness (or ignorance) is among the most common uncertainties in decision analysis. The evidential reasoning (ER) approach has been developed in the 1990s and in recent years to support the solution of MADA problems with ignorance as a kind of probabilistic uncertainty. The ER algorithm that satisfies several common-sense rules governing any approximate reasoning based aggregation procedures. In this paper, the ER approach is developed to aggregate feature extractions in any state of the machine using the sensor information. The process is briefly described as the following steps to generate an overall assessment by aggregating subjective judgments.

Step 1: Definition and representation of a multiple feature extraction

Define a set of L basic feature extractions in every state which is denoted by $E=\{e_1,e_2,\ldots,e_L\}$. Suppose the breakdown is complete; the L basic feature extractions include all the factors influencing the assessment of the general feature extractions. Estimate the relative weights of the feature extractions where w_i is the relative weight for basic feature extraction e_i and is normalized so that $0\le w_i\le 1$ and equation (5) are fulfilled. (The sum of the weights must be one for each state).

$$\sum_{i=1}^{L} w_i = 1 \tag{5}$$

Define N distinctive machine states A_n as a complete set of standards for assessing each option on all feature extractions in each state. $H = \{A_1, A_2, \ldots, A_n\}$. For each feature extractions e_i and machine states A_n, a degree of belief β_n is assigned. The degree of belief denotes the source's level of confidence when assessing the level of fulfillment of a certain property.

Step 2: Basic probability assignments for each basic feature extraction

Let m be a basic probability mass representing the degree to which the ith basic feature extractions e_i support a hypothesis that the general feature extractions are assessed to the nth the machine states A_n.

$$m_{n,i} = w_i \beta_{n,i} \qquad n = 1, \ldots N \tag{6}$$

Let $m_{H,i}$ be the remaining probability mass unassigned to each basic feature extractions e_i, $m_{H,i}$ is calculated as follows:

$$m_{H,i} = 1 - \sum_{n=1}^{N} m_{n,i} = 1 - w_i \sum_{n=1}^{N} \beta_{n,i} \qquad i = 1, \ldots L \tag{7}$$

Decompose $m_{H,i}$ into $\bar{m}_{H,i}$ and $\tilde{m}_{H,i}$ as follows:

$$\bar{m}_{H,i} = 1 - w_i \tag{8}$$

$$\tilde{m}_{H,i} = w_i (1 - \sum_{n=1}^{N} \beta_{n,i}) \tag{9}$$

with $m_{H,i} = \bar{m}_{H,i} + \tilde{m}_{H,i}$ (10)

Step 3: Combined probability assignments for a general feature extraction

In this step, the assessments of the basic feature extractions that constitute the general property are aggregated to form a single assessment of the general property. The probability masses assigned to the various assessment states as well as the probability mass left unassigned are denoted by:

$m_{n,I(L)}$ $(n = 1,2, \ldots, N)$, $\bar{m}_{H,I(L)}$, $\tilde{m}_{H,I(L)}$ and $m_{H,I(L)}$

Let $I(1) = 1$, then we get $m_{n,I(L)} = m_{n,1}$ $(n = 1, \ldots N)$, $\bar{m}_{H,I(1)} = \bar{m}_{H,1}$, $\tilde{m}_{H,I(1)} = \tilde{m}_{H,1}$ and $m_{H,I(1)} = m_{H,1}$.

The combined probability masses can be generated by aggregating all the basic probability assignments using the following recursive ER algorithm:

$$\{A_n\}: \quad \begin{aligned} & m_{n,I(i+1)} = K_{I(i+1)} \left[m_{n,I(i)} m_{n,i+1} + m_{H,I(i)} m_{n,i+1} + m_{n,I(i)} m_{H,i+1} \right] \\ & n = 1, \ldots, N \end{aligned} \tag{11}$$

We continue to let $i = 1$, which leads to that in Eq. (11) the termsmeasures the $m_{n,2}$ and $m_{n,1}$ and degree of both feature extractions e_1 and e_2 supporting the general features extraction to assesses A_n. The terms $m_{n,1}$ and $m_{H,2}$ measures the degree of only e_1 supporting to assesses A_n. The term $m_{H,1}$, $m_{n,2}$ measures the degree of only e_2 supporting to assess A_n.

$$\{H\}: \quad m_{H,I(i)} = \tilde{m}_{H,I(i)} + \bar{m}_{H,I(i)}$$

$$\tilde{m}_{H,I(i+1)} = K_{I(i+1)}\left[\tilde{m}_{H,I(i)}\tilde{m}_{H,i+1} + \bar{m}_{H,I(i)}\tilde{m}_{H,i+1} + \tilde{m}_{H,I(i)}\bar{m}_{H,i+1}\right] \tag{12}$$

$$\bar{m}_{H,I(i+1)} = K_{I(i+1)}\left[\bar{m}_{H,I(i)}\bar{m}_{H,i+1}\right] \tag{13}$$

$$K_{I(i+1)} = \left[1 - \sum_{t=1}^{N}\sum_{\substack{j=1 \\ j\neq t}}^{N} m_{t,I(i)} m_{j,i+1}\right]^{-1} \quad i = \{1,2,\ldots,L-1\} \tag{14}$$

In Eq. (14) the terms $\tilde{m}_{H,1}$ and $\tilde{m}_{H,2}$ measure the degree which cannot be assessed to any individual states due to the incomplete assessments for both e_1 and e_2. The terms $\tilde{m}_{H,1}$ and $\bar{m}_{H,2}$ measure the degree which cannot be assessed due to the incomplete assessments for e_1. The terms $\bar{m}_{H,1}$ and $\tilde{m}_{H,2}$ measure the degree which cannot be assessed due to the incomplete assessments for e_2. In Eq. (13), the terms $\bar{m}_{H,1}$ and $\bar{m}_{H,2}$ measure the degree which has not yet been assessed to individual states due to the relative importance of e_1 and e_2 after e_1 and e_2 have been aggregated. $m_{H,I(2)}$ and $m_{n,I(2)}$ as calculated by Eq. (14) is used to normalize $K_{I(2)}$ so that $\sum_{n=1}^{N} m_{n,I(2)} + m_{H,I(2)} = 1$

5. Main fusion based on entropy weighting

Let us return to the fault diagnosis problem with N state of the machine $H = \{A_1, A_2, \ldots, A_N\}$. Also assume that we test these states using M sensor information ($S_1, S_2, \ldots, S_M$), respectively. We combined the information from each sensor for every state of the machine to indicate the quality of every state. Each probability assignment is considered as the belief quantified for feature extractions from the information source provided by the multi-sensor. However, this information does not by itself provide 100% certainty as complete evidence sufficient for making decisions about the fault in the state of the machine. Therefore, it may be helpful to identify somehow the quality of information regarding the multi-sensor and to take this measure into account the main fusion. Intuitively, if the uncertainty associated with ith sensor is high, it would make us more ambiguous in the decision making. This way of defining weights associated with the main fusion using the measure of Shannon entropy as following (Huynh et al. 2009). Let us denote $m_i(\{A_j\})$ calculated BPA of jth state ($j = 1,2,\ldots,N$) regarding the ith sensor. Then the weight is defined by (Huynh et al. 2009):

$$w_i = 1 - \frac{H(m_i\{A_j\})}{\log(N)} \qquad \begin{array}{l} i = 1,2,\ldots,M \\ j = 1,2,\ldots,N \end{array} \tag{15}$$

Where H is a Shannon entropy expression as following:

$$H(m_i\{A_j\}) = -\sum_{j=1}^{N} m_i\{A_j\}\log(m_i\{A_j\}) \tag{16}$$

Now, our purpose is to combine all pieces of evidence $m_i(\{A_j\})$ from individual sensor ($S_1, S_2, \ldots, S_M$) taking into account their weights' w_i respectively, to overall mass function $m(A)$ for making a final decision. Where $A = (A_1, A_2, \ldots, A_N)$ indicate the different states of the machine. Formally, such an overall mass function $m(A)$ can be formulated in the general form of the following:

$$m(A) = \bigoplus_{i=1}^{M} w_i \otimes m_i(\{A_j\}) \tag{17}$$

where $\otimes$ is the discounting operator and $\oplus$ is a combination operator. Two useful operations that play a central role in the manipulation of belief functions are discounting and Dempster's rule of combination. The discounting operation is used when a source of information provides a BPA m, but one knows that this source has a probability α of reliability (Huynh et al. 2009). Then one may adopt $1-\alpha$ as one's discount rate, which results in a new BPA m^{α} defined by Eq. (6). Accordingly, Eq. (17) represents the proposed weighted coefficient associated with the machine fault diagnosis where $H(m_i\{A_j\}) \leq \log(N)$ and $0 \leq w_i \leq 1$. It is similar to the formula in reference (Huynh et al. 2010). The effectiveness of such a mechanism depends on uncertainty regarding fault sources based on available evidence. This proposed method offers an entropy weighting to calculate the measure of uncertainty level.
As mentioned in Shafer (1976), an obvious way to use discounting with Dempster's rule of combination is to discount all mass functions $m_i(\{A_j\})$ ($i = 1,2,\ldots,M$) at corresponding rates ($1-w_i$) ($i = 1,2,\ldots,M$) before combining them. This discounting–and–orthogonal sum combination strategy is carried out as follows. First, from each mass function $m_i(\{A_j\})$ and its associated weight w_i, we obtain the corresponding discounted mass function, denoted by $m_i^{\omega}(\{A_j\})$, as follows:

$$m_i^{\omega}(\{A_j\}) = w_i \times m_i(\{A_j\}) \quad for\ j = 1,2,\ldots,M \tag{18}$$

$$m_i^{\omega}(H) = 1 - w_i \tag{19}$$

Then, Dempster's rule of combination allows us to combine all $m_i^{\omega}(\{A_j\}$, ($i = 1,2,\ldots,M$), under the independent assumption of information sources for generating the overall mass function $m(A)$. Note that, by definition, focal elements of each $m_i^{\omega}(\{A_j\}$ are either singleton sets or the whole frame of discernment. It is easy to see that $m(A)$ also verifies this property if applicable. Final combines a collection of $m_i^{\omega}(\{A_j\}$ ($i = 1, 2, \ldots, M$) is calculated as follows:

$$m(A) = m_1^{\omega}(\{A_j\}) \oplus m_2^{\omega}(\{A_j\}) \oplus \cdots \oplus m_M^{\omega}(\{A_j\}) \tag{20}$$

For example, $m(A)$ after using ith and tth sensors are calculated as follows:

$$m(A)=m_i^{\varpi}(\{A_j\})\oplus m_t^{\varpi}(\{A_j\})=$$
$$\frac{m_i^{\varpi}(\{A_j\})\cdot m_t^{\varpi}(\{A_j\})+m_i^{\varpi}(\{A_j\})\cdot m_t^{\varpi}(H)+m_i^{\varpi}(H)\cdot m_t^{\varpi}(\{A_j\})}{1-\sum_{p}\sum_{\substack{q\\q\neq p}} m_i^{\varpi}(\{A_p\})\cdot m_t^{\varpi}(\{A_q\})} \tag{21}$$

Where,

$$K=1-\sum_{p}\sum_{\substack{q\\q\neq p}} m_i^{\varpi}(\{A_p\})\cdot m_t^{\varpi}(\{A_q\}) \tag{22}$$

Eqs. (20) and (21) is discounting-and-orthogonal sum combination (Dempster's rule of combination), which strategy is carried out as follows, First, from each mass function and its associated weight, it is obtained the corresponding discounted mass function, denoted by, second, under the independent assumption of information sources for generating, (based on the formula (18) and (19), after discounting strategy), a combination of all is done based on Eqs. (20) and (21). The effectiveness of such a mechanism depends on uncertainty regarding fault sources based on available evidence. This proposed method offers an entropy weighting to calculate the measure of uncertainty level.

6. Case study

Conditional monitoring is a process of monitoring a parameter of a condition in machinery, such that a significant change is indicative of a developing failure. It is a major component of predictive maintenance. The use of conditional monitoring allows maintenance to be scheduled, or other actions to be taken to avoid the consequences of failure before the failure occurs. Nevertheless, a deviation from a reference value (e.g. acoustic emission or vibration behavior) must occur to identify impending damages. In this work, an attempt has been made to monitor the tool condition (in-process) using an acoustic emission sensor and vibration sensor in a milling machine. In this section, an example is given to explain the proposed method of the sensor fusion system. In the example, we demonstrate how the proposed method can be used to decide the fault state machine.

Figure 1. The ball mill machine

One of the more critical machines in milling is often the ball mill. It is a cylindrical device used in grinding (or mixing) materials like ores, chemicals, ceramic, raw materials, and paints (Fig.1). If the ball mill is not capable of thinning the material, no product leaves the plant.
The major components of the ball mill at one steel facility are the pair of large spools that feed material into and out of the mill. Each spool is operated by a set of 2000kW variable frequency

drive motors, which power a gearbox that in turn drives the spools (Fig.1). The reliability of such a motor/gearbox combination is critical to the product output of the steel mill, by monitoring acoustic emission inside a reservoir and gearbox bearing vibration with two sensors, imminent failure was avoided at this facility. The ball mill is a low-speed rotating cylindrical or conical roller that ends in a steel drum. Work on the cylinder of steel balls constantly hit and extrusion. In a low-speed rotary machine like the ball mill machine, the acoustic sensor has higher importance than other sensors because we know the status of the device directly by using an acoustic sensor in the ball mill. Another vibration sensor that helps to condition monitoring is a vibration sensor. It is installed on the bearing to detect vibrations of the gearbox. In this case study, the weighty importance of the vibration sensor is more about the weight importance of the acoustic sensor.

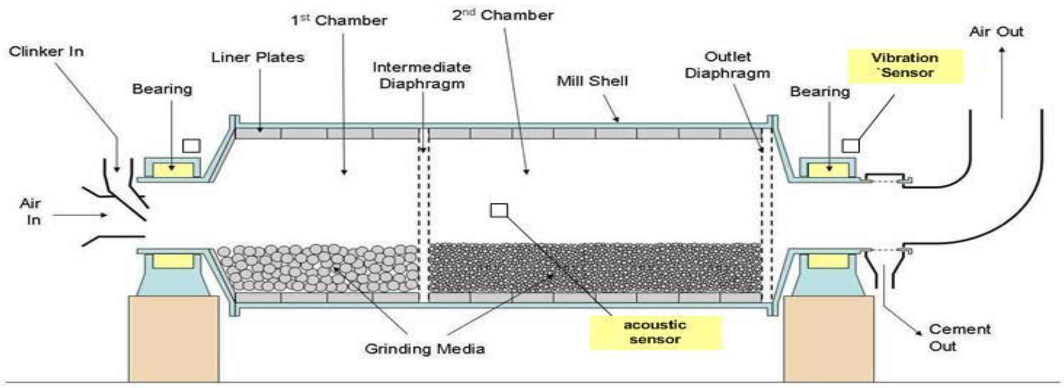

Figure 2. The ball mill machine used in grinding materials

We use two sensors to perform an online inspection of three states $H=\{A_1, A_2, A_3\}$ in the machine (Fig 2). For vibration sensors, we have assessed two features extracted e e_1 and e_2 for each state. The basic feature extracted is assessed to be $\{A_1\}$ by an observer, with a degree of belief $\beta_{1,1}$ of 40%, $\{A_2\}$of 0 %. We $\beta_{3,1}$ to the degree of belief $\{A_3\}$05nd a by % $\beta_{2,1}$ ‘ can calculate basic probability masses with the values of the weight $w_1 = 0.35$ as follows:

$m_{1,1} = w_1\beta_{1,1} = 0.35\times 0.4 = 0.14$

$m_{2,1} = w_1\beta_{2,1} = 0.35\times 0.5 = 0.115$

$m_{3,1} = w_1\beta_{3,1} = 0.35\times 0.0 = 0$

Where

$m_{N,1}$ $(N=1,2,3)$ is The probability mass of feature extraction e_1 for three states $\{A_1\}, \{A_2\}$, $\{A_3\}$.

In a similar way to the acoustic sensor, we calculate the probability mass of feature extracted e_2 using the values of the weight $w_2 = 0.65$ and the degree of belief $\beta_{n,2}$. Probability mass for the state $\{A_1\}$, $\{A_2\}$ and $\{A_3\}$ regarding that feature extracted e_2, respectively is:

$m_{1,2} = w_2\beta_{1,2} = 0.65\times 0.1 = 0.065$

$m_{2,2} = w_2\beta_{2,2} = 0.65\times 0.75 = 0.4875$

$m_{3,2} = w_2\beta_{3,2} = 0.65\times 0.15 = 0.0975$

Unassigned probability masses are calculated for $e_i \quad (i=1,2)$, in the spreadsheet show as follows:

$m_{H,1} = 0.685$ $\qquad$ $m_{H,2} = 0.35$

$\overline{m}_{H,1} = 1 - w_1 = 1 - 0.35 = 0.65$ $\qquad$ $\overline{m}_{H,2} = 1 - w_2 = 1 - 0.65 = 0.35$

$$\tilde{m}_{H,1} = w_1(1 - \sum_{n=1}^{N} \beta_{n,1}) = w_1(1 - (\beta_{1,1} + \beta_{2,1} + \beta_{3,1})) = 0.35(1 - 0.4 - 0.5 - 0) = 0.035$$

$$\tilde{m}_{H,2} = w_2(1 - \sum_{n=1}^{N} \beta_{n,2}) = 0.65(1 - 0.1 - 0075 - 0.15) = 0$$

Now the assessments of the basic feature extractions that constitute the general property are aggregated to form a single assessment of the general property for each state. The probability masses assigned to the various assessment states as well as the probability mass left unassigned are denoted by Eqs. (8) and (9). The combined probability masses can be generated by aggregating all the basic probability assignments using the following recursive ER algorithm: (for example) for the state $\{A_1\}$ we have:

$$\{A_1\}: \qquad m_{1,I(2)} = K_{I(2)}\lfloor m_{1,1}m_{1,2} + m_{H,1}m_{1,2} + m_{1,2}m_{H,2} \rfloor = 0.1154$$

$$K_{I(i+1)} = \left[1 - \sum_{t=1}^{N} \sum_{\substack{j=1 \\ j \neq t}}^{N} m_{t,I(i)} m_{j,i+1}\right]^{-1} = 1.12402$$

$$\{H\} \qquad \tilde{m}_{H,I(2)} = 1.12402.\lfloor \tilde{m}_{H,I(1)} \cdot \tilde{m}_{H,2} + \overline{m}_{H,I(1)} \cdot \tilde{m}_{H,2} + \tilde{m}_{H,I(1)} \cdot \overline{m}_{H,2} \rfloor = 0.0138$$

$$\overline{m}_{H,I(2)} = 1.12402.\lfloor \overline{m}_{H,I(1)} \cdot \overline{m}_{H,2} \rfloor = 0.2557$$

$$\Rightarrow m_{H,I(2)} = 0.2557 + 0.0138 = 0.2695$$

Where $m_{1,1}$ is the probability mass of feature extracted e_1 in the state $\{A_1\}$. $m_{1,1}$ Is calculated with the values of the weight $w_1 = 0.35$ and the degree of belief $\beta_{1,1}$. Experimental results for the first stage of the combination are summarized in Table (2). This Table shows the calculated probability of fault diagnosis in each state and the unassigned probability masses in the spreadsheet:

Table 2. Belief degrees and calculated probability masses with assigned weights,

State of machine $H=\{A_1,A_2,A_3\}$	weight	Belief				Probability mass						constant	probability mass (aggregation)					
	ω_i	$\beta_{1,i}$	$\beta_{2,i}$	$\beta_{3,i}$	β_H	$m_{1,i}$	$m_{2,i}$	$m_{3,i}$	$m_{H,i}$	$\overline{m}_{H,i}$	$\tilde{m}_{H,i}$	$K_{I(i+1)}$	$m_{1,I(i+1)}$	$m_{2,I(i+1)}$	$m_{3,I(i+1)}$	$m_{H,I(i+1)}$	$\overline{m}_{H,I(i+1)}$	$\tilde{m}_{H,I(i+1)}$
Feature extraction	0.35	0.4	0.5	0	0.1	0.14	0.175	0	0.685	0.65	.035	1.12402	0.1154	0.5401	0.0751	0.2695	0.2557	0.0138
Feature extraction e_2	0.65	0.1	0.075	0.15	0	0.065	0.487	0.097	0.35	0.35	0							

following the example, Consider the data of Table (3) which depicts the information acquired from the two sensors (the vibration sensor and the acoustic sensor) this information is resulted in three states according to combine information from feature extractions in each state in the first stage. $m_1(\{A_j\})$, ($j=1,2,3$) is the mass function associated with the vibration sensor, and $m_2(\{A_j\})$ is the mass function associated with the acoustic sensor for the three states of the machine. The objective is to diagnose machine failure based on the two sensors.

Table 3. Mass functions of the faults in three states of the machine as assessed based on the two sensors

Sensors	$\{A_1\}$	$\{A_2\}$	$\{A_3\}$	H
(S_1) **Vibration**	0.1154	0.5401	0.0751	0.2695
(S_2) **Acoustic**	0.1942	0.4393	0.0628	0.3038

The BPAs based on the weight associated with the two sensors, according to Table (3) weighted BPAs are listed in Table (4). Furthermore, the main fuse is computed by (21).

Table 4. Mass functions of the faults in three states of the machine and main fusion based on entropy weighting

Sensors	$\{A_1\}$	$\{A_2\}$	$\{A_3\}$	H
(S_1) **Vibration**	0.034	0.159	0.022	0.706
(S_2) **Acoustic**	0.043	0.097	0.014	0.778
Main fusion information	0.0583	0.21	0.028	0.7

Where

$$H(m_1(A))=-(0.1154\times\log 0.1154+0.5401\times\log 0.5401+0.0751\times\log 0.0751)=0.337$$
$$H(m_2(A))=-(0.1942\times\log 0.1942+0.4393\times\log 0.4393+0.0628\times\log 0.0628)=0.371$$

$$w_1 = 1 - \frac{0.337}{\log(3)} = 0.294$$

$$w_2 = 1 - \frac{0.371}{\log(3)} = 0.222$$

$$m_1^{\omega}(H) = 1 - w_1 = 1 - 0.294 = 0.706$$

$$m_2^{\omega}(H) = 1 - w_2 = 1 - 0.222 = 0.778$$

$$K = 1 - \sum_{p} \sum_{\substack{q \\ q \neq p}} m_1^{\omega}(\{A_p\}) \cdot m_2^{\omega}(\{A_q\}) = 1.0162$$

$$m_1^{\omega}(\{A_1\}) \oplus m_2^{\omega}(\{A_1\}) = 0.034 \times 0.043 + 0.034 \times 0.778 + 0.043 \times 0.706 = 0.0583$$

$$j = 1,2,3$$

Similarly, we calculate the probability mass of two states $\{A_2\}$ and $\{A_3\}$ regarding their entropy weighting. We can see clearly from Table (4) that, comparing the mass function value after fusion, by the usage of two kinds of the sensors separately. We can precisely recognize a fault state $\{A_2\}$, it is identical to real diagnosis machine failure. The results of Table (3) are separate from Table (4). Table (3) BPAs or mass functions of failure for ball mill, calculate about weights that reflect the importance of evidence in the diagnosis of fault in each state. The weight of a mass function is determined using a set of defect features and their values are commonly determined via expert opinion. In other word weight of every two pieces of evidence which support fault in each state.

Furthermore, since the information obtained from the sensors is inherently incomplete, uncertain, and imprecise, a fusion mechanism must be devised to minimize such imprecision and uncertainty. The effectiveness of such a mechanism depends on uncertainty regarding fault sources based on available evidence. This proposed method offers an entropy weighting to calculate the measure of uncertainty level, because it is important to decide at what level of abstraction the fusion process is to take place, at the measurement level, and the decision level. We have calculated the new BPAs using entropy weighting that is shown in Table (4) and the fused results are shown in the third row of Table (4). Therefore, the new BPAs (in Table (4)) are determined using the weighting coefficient and the information fusion is calculated to measure the uncertainty level of the decisions. The proposed weighting coefficient is based on Shannon entropies. Therefore, multi-sensor data fusion based on the D-S evident theory effectively raised the fault state recognition ability. Furthermore, this shows the entropy of the collected evidence has improved as a result of information fusion and thus leading to improved diagnostic performance. The proposed method has been widely used in reasoning with uncertainty and information fusion in fault diagnosis systems; it seems that the proposed method can embody both information's uncertainty and expert subjective knowledge in its fault diagnosis, thereby improving the accuracy of the diagnosis.

7. Managerial insights

Companies are involved in high competition for reducing the cost of production to maintain their market shares. Because maintenance costs are a significant part of production costs, companies need to fund maintenance effectively (Farugh and Mostafayi, 2019; Behnamian and Rahami, 2020). As explained if we can detect the failure of complex and heavy equipment such as ball mills, we can avoid huge costs. Heavy and complex equipment such as ball mills in case of failure, huge costs, including machine sleep costs, reconstruction costs, and damage due to line stops will impose on the organization. Specialist visual inspection of heavy equipment is impossible, and special sensors are used to aid in diagnosis. Sensors can also give inaccurate and unreliable information to the machine operator, making it difficult to make the final decision to stop and repair the production line. In the obtained method, combining the output of the sensors and their fusion minimizes the uncertainty, which improves the performance of the detector. Failure time detection will be an important factor in preventing the imposition of heavy material and human costs in the factory. Accurate detection of equipment failure will increase the safety and reliability of the devices. Unreasonable stopping of the device in a state that is not broken or failure to identify the location of the defect can increase maintenance and repair costs, while the method provided for combining the data will detect the failure in the event of certainty. Based on the mentioned reasons, however, there are many methods to evaluate the fault diagnosis in the process of ER using multisensory data fusion, which has a great influence on reasoning results. Because of the traditional method based on the expertise of experts to index values between different faults, it is more effective to use the D-S method. An important matter in all kinds of industries is fault detection identification and diagnosis which are evaluated of ambiguity measured information by several sensors and expert's opinions, especially ball mills in ceramic tile, gypsum, minerals such as gold, copper, silver, glass, … industries. However, different methods are used in the mentioned factories; the main purpose of this paper is to provide a practical method for minimizing uncertainty in ball mill fault detection. The existence method will be affected by uncertain value in multisensory information and different indicators of specialist's indicators. To solve the above problems, we propose a method to regularize the reliability value of fault diagnosis. We use the entropy weighted presents as a measure of the uncertainty level of information and mass functions based on experts' information to increase identification accuracy in the decision-making process. Because of considering different approaches to ambiguity in the information obtained from sensors and different opinions of experts as a calculation formula, the reliability of fault diagnosis will be improved in the decision-making process. Therefore, this paper uses the improved evidence reasoning algorithm using the kind of sensor information to carry out fault diagnosis. The main contributions of this paper are as follows; (1) In traditional ER, the reliability only considers set interval in fault data, neglecting the impact of the relationship between the inner data. In the newly proposed method, we used two-stage of considering uncertainty and fused data which is obtained from multiple sensors as accuracy and reliability regularization. In practice, the new reliability is more reliable. (2) To achieve better fault diagnosis results, we combine improved ER and entropy weight in the ball mill fault diagnosis. The experimental result shows that the new method is better than the traditional one.

8. Conclusions and summary

Multisensory data fusion is a broad issue due to the wide range of scenarios that can be applied to fault diagnosis of a system. Thus, finding a model or an algorithm scheme, which can be explicated and implemented in a real system, is an important matter in all kinds of industries. It is important to say that, existing models cannot be used for any kind of application at the same time and it is an unfeasible task. Hence, a view of the different approaches, theories, and implementations in the issue of sensor fusion can be presented intending to be useful as a collection of different ideas that should be combined with the implementation of a real fusion system. The D-S theory is a well-known method for its usefulness to express uncertain judgments and is frequently used as a method for dealing with uncertain information. This paper used this theory and the ER approach as the framework for representing the uncertainty of fault detection identification and diagnosis. Furthermore, entropy weighted presents as a measure of the uncertainty level of information, which is employed to decrease the effect of uncertainty on damage identification and increasing identification accuracy. Accordingly, a multi-sensor implementation of an evidence theory based on a diagnostic system is described. To detect and identify machine faults, integration of multi-sensor (vibration and temperature information) is presented. The proposed method for fault diagnostic problems is established using the various states of the machine and the calculus of safety and reliability system states with uncertain information in the decision-making process. This is formulated in the context of the evidence, reasoning in terms of fault frame of discernment (N state of the machine), mass functions, and evidence combination. The proposed method has considered two stages of calculating mass functions (feature extraction fusion and main fusion). According to the effectiveness of factors, the fault in each state of the machine is determined using multiple information sources (or multi-sensor), then for evaluation of ambiguity measured information by sensors, this study has proposed a new method for defining adaptively weights of each sensor information by entropy measures. This is considered as ambiguity associated with their measurement information in the main fusion and evaluates the performance of information fusion.

The presented example at the end of this paper has explained a case of insufficient available data. The case study is contented of the ball mill with low speed rotating cylindrical and conical roller, where two sensors are installed on the ball mill to sense acoustic and vibrations of the machine. The results using the information acquired from the two sensors separately, (considering entropy weight, mass function value, etc.) are precisely recognized as a fault state $\{A_2\}$ in the real diagnosis machine failure. The results of Table (3) are separate from Table (4). Table (3) is resulted in using BPAs or mass functions and calculated weights that reflected the importance of evidence in the diagnosis of fault in each state. The weight of a mass function is determined using a set of defect features, and their values are commonly determined via expert opinion. In other words, the weight of every two pieces of evidence is supported by a fault in each state. Since information obtained from the sensors is inherently incomplete, uncertain, and imprecise, a fusion mechanism must be devised to minimize such imprecision and uncertainty. The effectiveness of such a mechanism depends on uncertainty regarding fault sources based on available evidence. Because it is important to decide at what level of abstraction the fusion process is to take place, at the measurement level, and at the decision level, the entropy weight has calculated the measure of uncertainty level. Then the new BPAs using entropy weighting that is shown in Table (4) have been calculated and the fused results are shown in the third row of Table (4). Therefore, the new BPAs (in Table (4)) are determined using the weighting coefficient and the information fusion is calculated to measure the uncertainty level of the decisions.

Since its inception, the proposed method can be reasonably and effectively used in uncertain conditions and information fusion in fault diagnosis systems; because it makes inferences with maximum specificity in the presence of uncertain and missing data, dealing with minimal information. It can embody both information's uncertainty and expert subjective knowledge in the fault diagnosis process, thereby improving the accuracy of the fault diagnosis. The proposed algorithm and results concluded that the method is easy to use. Also, the development of the D-S theory for fault diagnosis can not only resolve the decision conflict based on conflict among evidence but also improved the accuracy of fault diagnosis by combining expert knowledge and fusing information from multi-sensor. Thus, it can improve the accuracy of decision making for fault diagnosis by fusing information from multi-sensor. This method as well as can analyze and preparing reports to diagnose a fault with a reduction of overall uncertainty and increase of accuracy and can be followed to implement a real system. Since the ability to machine fault diagnosis of these systems is becoming increasingly important from both economic and life-safety viewpoints, the ER approach with defining adaptively weights, importance index and conflict factors to consider the issues of evidence sufficiency, differing importance of various evidence, and the conflict of different evidence, using the final combination to detect faults in machines, where used and developed the D–S evidence theory to integrate multi-sensor information including vibration, and temperature to detect and identify machine faults. An application and an example illustrate the efficiency of the method in evidential sensor fusion. It shows that the proposed method is more efficient for highly conflicting evidence with better performance of convergence and less uncertainty, compared with the existing methods Some related advantages and disadvantages of different methods are discussed as follows.

Dempster's method, which can well deal with imprecise and uncertain information, is widely used in fusing information. However, when it comes to highly conflicting evidence the method will always lead to some illogical results. Other methods like Murphy's method (Murphy, 2000), which uses an average operation to combine the conflicting evidence, can deal with highly conflicting evidence to some extent. However, the difference and relationship of evidence are neglected. In Jiang's method (Jiang et al. 2016), belief entropy is used to calculate the weight of each evidence. It considers the difference of evidence, which makes it more reasonable than Murphy's method. In Deng's method (Deng et al. 2004), which is different from Jiang's method, evidence distance is used to calculate the weight rather than the belief entropy. And the similarity between the two pieces of evidence is proposed. The proposed method, considering both belief entropy and evidence theory to obtain a better result. Although it might be not very suitable in some situations, it leverages the advantages of Jiang's method and Deng's method.

Acknowledgements

The authors would like to appreciate constructive comments and helpful suggestions from the associate editor and the referees, which have helped to improve the presentation of the paper.

References

Basir, O., and Yuan, X., (2007). "Engine fault diagnosis based on multi-sensor information fusion using Dempster–Shafer evidence theory", *Information Fusion*, Vol. 8, No. 4, pp. 379–386. https://doi.org/10.1016/j.inffus.2005.07.003.

Behnamian, J., Rahami, Z., (2020). "Multi-objective scheduling and assembly line balancing with resource constraint and cost uncertainty: A “box” set robust optimization", *Journal of Industrial Engineering and Management Studies*, Vol. 7, No. 1, pp. 220-232. https://doi.org 10.22116/JIEMS.2020.110250.

Bessam, B., Menacer, A., Boumehraz, M., and Cherif, H., (2015). "Wavelet transform and neural network techniques for inter-turn short circuit diagnosis and location in induction motor", *International Journal of System Assurance Engineering and Management*, Vol. 8(S1), pp. 478–488. https://doi.org/10.1007/s13198-015-0400-4.

Bhalla, D., Bansal, R.K., and Gupta, H.O., (2013). "Integrating AI based DGA fault diagnosis using Dempster–Shafer Theory", *International Journal of Electrical Power & Energy Systems*, Vol. 48, pp. 31–38. https://doi.org/10.1016/j.ijepes.2012.11.018.

Cai, B., Liu, H., and Xie, M., (2016). "A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks", *Mechanical Systems and Signal Processing*, Vol. 80, pp. 31–44. https://doi.org/10.1016/j.ymssp.2016.04.019.

Cheng, G., Chen, X., Li, H., Li, P., and Liu, H., (2016). "Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition", *Measurement*, Vol. 91, pp. 140–154. https://doi.org/10.1016/j.measurement.2016.05.059.

Dempster, A.P., (1968). "A Generalization of Bayesian Inference", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 30, No. 2, pp. 205–232. https://doi.org/10.1111/j.2517-6161.1968.tb00722.x.

Deng, Y., (2016). "Deng entropy", *Chaos, Solitons & Fractals*, Vol. 91, pp. 549–553. https://doi.org/10.1016/j.chaos.2016.07.014.

Deng, X., Jiang, W., and Wang, Z., (2019). "Zero-sum polymatrix games with link uncertainty: A Dempster-Shafer theory solution", *Applied Mathematics and Computation*, Vol. 340, pp. 101–112. https://doi.org/10.1016/j.amc.2018.08.032.

Deng, Y., Shi, W.K., Zhu, Z.F., and Liu, Q., (2004). "Combining belief functions based on distance of evidence", *Decision Support Systems*, Vol. 38, No. 3, pp. 489–493. https://doi.org/10.1016/j.dss.2004.04.015.

Fan, X., and Zuo, M.J., (2006). "Fault diagnosis of machines based on D–S evidence theory. Part 2: Application of the improved D–S evidence theory in gearbox fault diagnosis", *Pattern Recognition Letters*, Vol. 27, No. 5, pp. 377–385. https://doi.org/10.1016/j.patrec.2005.08.024.

Farugh, H., Mostafayi, S., Afrasiabi, A., (2019). "Bi-objective robust optimization model for configuring cellular manufacturing system with variable machine reliability and parts demand: A real case study", *Journal of Industrial Engineering and Management Studie*s, Vol. 6, No. 2, pp. 120-146. https://doi.org/120-146. 10.22116/ JIEMS.2019.93028.

Guedidi, S., Zouzou, S. E., Laala, W., Yahia, K., and Sahraoui, M., (2013). "Induction motors broken rotor bars detection using MCSA and neural network: Experimental research", *International Journal of System Assurance Engineering and Management*, Vol. 4, No. 2, pp. 173–181. https://doi.org/10.1007/s13198-013-0149-6.

Guo, M., Yang, J.B., Chin, K.S., and Wang, H., (2008). *The evidential reasoning approach for multi-attribute decision analysis under both fuzzy and interval uncertainty*, *Interval/probabilistic uncertainty and non-classical logics*, part of the advances in soft computing book series (AINSC) 46:129-140.

Han, Y., and Deng, Y., (2018). "An evidential fractal analytic hierarchy process target recognition method, *Defence Science Journal*, Vol. 68, No. 4, p. 367. https://doi.org/10.14429/dsj.68.11737.

Han, Y., and Deng, Y., (2018b). "A novel matrix game with payoffs of Maxitive Belief Structure", *International Journal of Intelligent Systems*, Vol. 34, No. 4, pp. 690–706. https://doi.org/10.1002/int.22072.

Hang, J., Zhang, J., and Cheng, M., (2016). "Application of multi-class fuzzy support vector machine classifier for fault diagnosis of wind turbine", *Fuzzy Sets and Systems*, Vol. 297, pp. 128–140. https://doi.org/10.1016/j.fss.2015.07.005.

Huang, Y., McMurran, R., Dhadyalla, G., and Peter Jones, R., (2008). "Probability based vehicle fault diagnosis: Bayesian network method", *Journal of Intelligent Manufacturing*, Vol. 19, No. 3, pp. 301–311. https://doi.org/10.1007/s10845-008-0083-7.

Hui, K. H., Lim, M. H., Leong, M. S., and Al-Obaidi, S.M., (2017). "Dempster-Shafer evidence theory for multi-bearing faults diagnosis", *Engineering Applications of Artificial Intelligence*, Vol. 57, pp. 160–170. https://doi.org/10.1016/j.engappai.2016.10.017.

Huynh, V.-N., Nguyen, T.T., and Le, C.A., (2010). "Adaptively entropy-based weighting classifiers in combination using Dempster–Shafer theory for word sense disambiguation", *Computer Speech & Language*, Vol. 24, No. 3, pp. 461–473. https://doi.org/10.1016/j.csl.2009.06.003.

Jian-Bo Yang, and Dong-Ling Xu., (2002). "Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty", *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 32, No. 3, pp. 376–393. https://doi.org/10.1109/tsmca.2002.802809.

Jian-Bo Yang, and Singh, M. G., (1994). "An evidential reasoning approach for multiple-attribute decision making with uncertainty", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 1, pp. 1–18. https://doi.org/10.1109/21.259681.

Jiang, W., Wei, B., Xie, C., and Zhou, D., (2016). "An evidential sensor fusion method in fault diagnosis", *Advances in Mechanical Engineering*, Vol. 8, No. 3, 168781401664182. https://doi.org/10.1177/1687814016641820.

Jousselme, A.-L., Grenier, D., and Bossé, É., (2001). "A new distance between two bodies of evidence", *Information Fusion*, Vol. 2, No. 2, pp. 91–101. https://doi.org/10.1016/s1566-2535(01)00026-4.

Kaftandjian, V., Dupuis, O., Babot, D., and Min Zhu, Y., (2003). "Uncertainty modelling using Dempster–Shafer theory for improving detection of weld defects", *Pattern Recognition Letters*, Vol. 24(1–3), pp. 547–564. https://doi.org/10.1016/s0167-8655(02)00276-3.

Kang, B., Deng, Y., Hewage, K., and Sadiq, R. (2019). "A Method of Measuring Uncertainty for Z-Number", *IEEE Transactions on Fuzzy Systems*, Vol. 27, No. 4, pp. 731–738. https://doi.org/10.1109/tfuzz.2018.2868496.

Khalaj, F., Khalaj, M., (2020). "Developed cosine similarity measure on belief function theory: An application in medical diagnosis", *Communications in Statistics - Theory and Methods*, pp. 1-12. https://doi.org/10.1080/03610926.2020.1782935.

Khalaj, M., Tavakkoli-Moghaddam, R., Khalaj, F., Siadat, A., (2020). "New definition of the cross entropy based on the Dempster-Shafer theory and its application in a decision-making process", *Communications in Statistics - Theory and Methods*, Vol. 49, No. 4, pp. 909-923. https://doi.org/10.1080/03610926.2018.1554123.

Laala, W., Zouzou, S.-E., and Guedidi, S., (2013). "Induction motor broken rotor bars detection using fuzzy logic: experimental research", *International Journal of System Assurance Engineering and Management*, Vol. 5, No. 3, pp. 329–336. https://doi.org/10.1007/s13198-013-0171-8.

Mabrouk, A. E., Zouzou, S. E., Khelif, S., and Ghoggal, A., (2015). "On-line fault diagnostics in operating three-phase induction motors by the active and reactive currents", *International Journal of System Assurance Engineering and Management*, Vol. 8(S1), pp. 160–168. https://doi.org/10.1007/s13198-015-0364-4.

Meng, D., Liu, M., Yang, S., Zhang, H., and Ding, R. (2018). A fluid–structure analysis approach and its application in the uncertainty-based multidisciplinary design and optimization for blades. *Advances in Mechanical Engineering*, 10(6), 168781401878341. https://doi.org/10.1177/1687814018783410.

Meng, D., Li, Y., Zhu, S.-P., Lv, G., Correia, J., and de Jesus, A., (2019). "An enhanced reliability index method and its application in reliability-based collaborative design and optimization", *Mathematical Problems in Engineering*, pp. 1–10. https://doi.org/10.1155/2019/4536906.

Moosavian, A., Khazaee, M., Najafi, G., Kettner, M., and Mamat, R., (2015). "Spark plug fault recognition based on sensor fusion and classifier combination using Dempster–Shafer evidence theory", *Applied Acoustics*, Vol. 93, pp. 120–129. https://doi.org/10.1016/j.apacoust.2015.01.008.

Murphy, C. K., (2000). "Combining belief functions when evidence conflicts", *Decision Support Systems*, Vol. 29, No. 1, pp. 1–9. https://doi.org/10.1016/s0167-9236(99)00084-6.

Oukhellou, L., Debiolles, A., Denœux, T., and Aknin, P. (2010). "Fault diagnosis in railway track circuits using Dempster–Shafer classifier fusion", *Engineering Applications of Artificial Intelligence*, Vol. 23, No. 1, pp. 117–128. https://doi.org/10.1016/j.engappai.2009.06.005.

Parikh, C. R., Pont, M. J., and Barrie Jones, N., (2001). "Application of Dempster–Shafer theory in condition monitoring applications: a case study", *Pattern Recognition Letters*, Vol. 22(6–7), pp. 777–785. https://doi.org/10.1016/s0167-8655(01)00014-9.

Petersen, D. R., Link, R. E., Wang, H. F., and Wang, J. P., (2000). "Fault Diagnosis Theory: Method and Application Based on Multisensor Data Fusion", *Journal of Testing and Evaluation*, Vol. 28, No. 6, pp. 513-518. https://doi.org/10.1520/jte12143j.

Rodríguez Ramos, A., Llanes-Santiago, O., Bernal de Lázaro, J. M., Cruz Corona, C., Silva Neto, A. J., and Verdegay Galdeano, J. L., (2017). "A novel fault diagnosis scheme applying fuzzy clustering algorithms", *Applied Soft Computing*, Vol. 58, pp. 605–619. https://doi.org/10.1016/j.asoc.2017.04.071.

Rong, H., Ge, M., Zhang, G., and Zhu, M. (2018). "An Approach for Detecting Fault Lines in a Small Current Grounding System using Fuzzy Reasoning Spiking Neural P Systems", *International Journal of Computers Communications & Control*, Vol. 13, No. 4, pp. 521–536. https://doi.org/10.15837/ijccc.2018.4.3220.

Sakhara, S., Saad, S., and Nacib, L., (2016). "Diagnosis and detection of short circuit in asynchronous motor using three-phase model", *International Journal of System Assurance Engineering and Management*, Vol. 8, No. 2, pp. 308–317. https://doi.org/10.1007/s13198-016-0435-1.

Seiti, H., and Hafezalkotob, A., (2019). "Developing the R-TOPSIS methodology for risk-based preventive maintenance planning: A case study in rolling mill company", *Computers & Industrial Engineering*, Vol. 128, pp. 622–636. https://doi.org/10.1016/j.cie.2019.01.012.

Shafer, G., (1976). *A mathematical theory of evidence*, Princeton University Press, New Jersey.

Smets, P., and Kennes, R., (1994). "The transferable belief model", *Artificial Intelligence*, Vol. 66, No. 2, pp. 191–234. https://doi.org/10.1016/0004-3702(94)90026-4.

Smith, A. F. M., and Shafer, G., (1976). "A Mathematical theory of evidence", *Biometrics*, Vol. 32, No. 3, p. 703. https://doi.org/10.2307/2529769.

Singh, B., Grover, S., and Singh, V., (2016). "Evaluation of benchmarking attribute for service quality using multi attitude decision making approach", *International Journal of System Assurance Engineering and Management*, Vol. 8(S2), pp. 617–630. https://doi.org/10.1007/s13198-016-0485-4.

Sun, R., and Deng, Y., (2019). "A new method to identify incomplete frame of discernment in evidence theory", *IEEE Access*, Vol. 7, pp. 15547–15555. https://doi.org/10.1109/access.2019.2893884.

Tang, W.H., Wu, Q.H., (2011). *Dealing with uncertainty for dissolved gas analysis, condition monitoring and assessment of power transformers using computational intelligence*, Part of the power systems book series (POWSYS), pp. 125-162.

Torabi Jahromi, A., Er, M.J., Li, X., and Lim, B.S., (2016). "Sequential fuzzy clustering based dynamic fuzzy neural network for fault diagnosis and prognosis", *Neurocomputing*, Vol. 196, pp. 31–41. https://doi.org/10.1016/j.neucom.2016.02.036.

Verbert, K., Babuška, R., and De Schutter, B., (2017). "Bayesian and Dempster–Shafer reasoning for knowledge-based fault diagnosis–A comparative study", *Engineering Applications of Artificial Intelligence*, Vol. 60, pp. 136–150. https://doi.org/10.1016/j.engappai.2017.01.011.

Wang, D., Tsui, KL., Zhou, Q., (2016). "Novel Gauss–Hermite integration based Bayesian inference on optimal wavelet parameters for bearing fault diagnosis", *Mech. Syst. Signal. Process*, Vol 72–73, pp.80-91.

Wang, Y., Zhang, K., and Deng, Y., (2018). "Base belief function: an efficient method of conflict management", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 10, No. 9, pp. 3427–3437. https://doi.org/10.1007/s12652-018-1099-2.

Yang, J.B., (2001). "Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties", *European Journal of Operational Research*, Vol. 131, No. 1, pp. 31–61. https://doi.org/10.1016/s0377-2217(99)00441-5.

Xiao, F., (2019). "A multiple-criteria decision-making method based on D numbers and belief entropy", *International Journal of Fuzzy Systems*, Vol. 21, No. 4, pp. 1144–1153. https://doi.org/10.1007/s40815-019-00620-2.

Xiao, F., (2018). "A novel multi-criteria decision-making method for assessing health-care waste treatment technologies based on D numbers", *Engineering Applications of Artificial Intelligence*, Vol. 71, pp. 216–225. https://doi.org/10.1016/j.engappai.2018.03.002.

Yager, R.R., (1987). "On the Dempster-Shafer framework and new combination rules", *Information Sciences*, Vol. 41, No. 2, pp. 93–137. https://doi.org/10.1016/0020-0255(87)90007-7.

Yaghoubi, A., Amiri, M., (2015). "Designing a new multi-objective fuzzy stochastic DEA model in a dynamic environment to estimate efficiency of decision-making units (Case Study: An Iranian petroleum company), *Journal of Industrial Engineering and Management Studies*, Vol. 2, No. 2, pp. 26-42.

Yang, H.C., Deng, Y., Jones, J., (2018). "Network division method based on cellular growth and physarum-inspired network adaptation", *International Journal of Unconventional Computing*, Vol. 13, No. 6, pp. 477–491.

Yang, J.B., Xu, D.L., (2002). "Nonlinear information aggregation via evidential reasoning in multi attribute decision analysis under uncertainty", *IEEE Trans. Syst., Man, Cybern Part A: Systems and Humans*, Vol. 32, No. 3, pp.1–18.

Yu, J., (2013). "A new fault diagnosis method of multimode processes using Bayesian inference based Gaussian mixture contribution decomposition", *Engineering Applications of Artificial Intelligence*, Vol. 26, No. 1, pp. 456–466. https://doi.org/10.1016/j.engappai.2012.09.003.

Yu, Y., Li, W., Sheng, D., and Chen, J., (2015). "A novel sensor fault diagnosis method based on modified ensemble empirical mode decomposition and probabilistic", *Neural Network. Measurement*, Vol. 68, pp. 328–336. https://doi.org/10.1016/j.measurement.2015.03.003.

Zhang, C., Li, D., Mu, Y., and Song, D., (2017). "An interval-valued hesitant fuzzy multigranulation rough set over two universes model for steam turbine fault diagnosis", *Applied Mathematical Modelling*, Vol. 42, pp. 693–704. https://doi.org/10.1016/j.apm.2016.10.048.

Zhang, H., and Deng, Y., (2018). "Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory", *Advances in Mechanical Engineering*, Vol. 10, No. 11, 168781401880918. https://doi.org/10.1177/1687814018809184.

Zhang, L., Wu, X., Zhu, H., and Abou Rizk, S.M., (2017). "Perceiving safety risk of buildings adjacent to tunneling excavation: An information fusion approach", *Automation in Construction*, Vol. 73, pp. 88–101. https://doi.org/10.1016/j.autcon.2016.09.003.

This article can be cited: Khalaj, M., Khalaj, F., (2021). "Machine fault diagnosis through entropy weighting using evidential reasoning approach", *Journal of Industrial Engineering and Management Studies*, Vol. 8, No. 1, pp. 114-134.