Research Article

Application of D-S Evidence Fusion Method in the Fault Detection of Temperature Sensor

Zheng Dou,[1] **Xiaochun Xu,**[1] **Yun Lin,**[1] **and Ruolin Zhou**[2]

[1] *College of Information and Communication Engineering, Harbin Engineering University, Room 148, Building 21,*
145 Nantong Street, Harbin, Heilongjiang 150001, China
[2] *Department of Electrical and Computer Engineering, Western New England University, Springfield, MA, USA*

Correspondence should be addressed to Yun Lin; linyun@hrbeu.edu.cn

Due to the complexity and dangerousness of drying process, the fault detection of temperature sensor is very difficult and dangerous in actual working practice and the detection effectiveness is not satisfying. For this problem, in this paper, based on the idea of information fusion and the requirements of D-S evidence method, a D-S evidence fusion structure with two layers was introduced to detect the temperature sensor fault in drying process. The first layer was data layer to establish the basic belief assignment function of evidence which could be realized by BP Neural Network. The second layer was decision layer to detect and locate the sensor fault which could be realized by D-S evidence fusion method. According to the numerical simulation results, the working conditions of sensors could be described effectively and accurately by this method, so that it could be used to detect and locate the sensor fault.

1. Introduction

Information fusion is a useful technique to integrate heterogeneous data from different information sources. By increasing comprehensiveness while decreasing uncertainty of information, information fusion can be used to improve the quality of decision using the redundancy and complementariness of different information sources. As one of the most important methods in information fusion, the Dempster-Shafer evidence theory (D-S theory) [1, 2], which is an improvement of the Bayesian theory, has been widely used in information systems [3–11]. A significant improvement of the D-S approach over traditional probabilistic approach is that it allows for the allocation of a probability mass to sets or intervals and can handle both stochastic uncertainty and subjective uncertainty. The D-S evidence theory is a flexible and powerful mathematical tool for handling uncertain, incomplete, and imprecise information for at least the following three reasons. Firstly, by representing the imprecision and uncertainty of a body of knowledge via the notion of evidence, belief can be committed to a singleton or a compound set. Secondly, the evidence combination rule of

the D-S evidence theory can provide an interesting operator to integrate multiple information acquired from different data sources. Finally, the decision on optimal hypothesis choice can be made in a rational and flexible manner.

In drying industry process, the supervision on working conditions of sensors is very important and difficult, and its role is to detect, locate, and isolate the fault sensor as quickly and accurately as possible. But, due to the complexity of sensor and the uncertainty of working environment, the monitoring data is usually uncertain, incomplete, or imprecise, which leads to the reduction of accuracy rate. Therefore, in this paper, a two-layer information fusion structure based on BP Neural Network and D-S evidence fusion method was presented for the supervision on working conditions of sensor in drying industry process. Firstly, according to the monitoring data obtained from different sensor sources, BP Neural Network was used to establish the basic belief assignment function of evidence for every single sensor source. Then, the D-S evidence combination rule was used to fuse those evidences. Finally, according to the fusion result, the working conditions of sensor could be described effectively and accurately. In this fusion process,

on the one hand, the BP Neural Network could provide the ability of self-learning, self-adaptation, and fault tolerance; on the other hand, the D-S evidence method could express and handle the uncertain, incomplete, and imprecise information. Therefore, this method could further improve the accuracy and robustness of sensor monitoring system, which was proved by the numerical simulation results.

2. Preliminaries

2.1. Dempster-Shafer Evidence Theory. The mathematical basis of evidence theory, which was introduced by Dempster [1] and extended by Shafer [2], pays attention to the question of belief in the proposition systems. "Belief" in a proposition is not the same as the "chance" of the proposition being true. When forming propositions, evidence can be considered as a similar way, and the Dempster-Shafer (D-S) theory pays attention to "evidence," "weights of evidence," and "belief in evidence." Obviously, the belief structure in the evidence theory conforms with the Bayesian Probability Model [2], and thus the evidence theory can be viewed as a generalization and improvement of the classic probability theory. Because of its ability in dealing with uncertainty and imprecision problems, the D-S theory can be widely used in many fields [3–11]. Formally, the evidence theory concerns with the following preliminary notations.

Framework of Discernment. Firstly, evidence theory supposes a set of hypotheses θ as the framework of discernment, which can be defined as follows:

$$\theta = \{H_1, H_2, \ldots, H_N\}, \tag{1}$$

where the set θ is composed of N exclusive and exhaustive hypotheses. In this paper, it represents the temperature sensors. Assume the power set $P(\theta)$ is composed of the 2^N propositions of θ as follows:

$$2^\theta = \{\emptyset, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \{H_1 \cup H_2\}, \\ \{H_1 \cup H_3\}, \ldots, \{H_1 \cup H_2 \cdots \cup H_N\}\}, \tag{2}$$

where $\emptyset$ denotes the empty set. Then, the subset containing only one element is called singleton.

Mass Functions, Focal Elements, and Kernel Elements. When the framework of discernment is determined, the mass function m can be defined as a mapping of the power set $P(\theta)$ to a number between 0 and 1, which is shown as follows:

$$m : P(\theta) \longrightarrow [0, 1], \tag{3}$$

and it also satisfies the following conditions:

$$\sum_{A \subset P(\theta)} m(A) = 1, \\ m(\emptyset) = 0. \tag{4}$$

The mass function m is also called the basic probability assignment (BPA) function, and $m(A)$ represents the proportion of all relevant and available evidences that supports the claim that a particular element of θ belongs to the set A but not to a particular subset of A. Any subset A of $P(\theta)$ satisfying that $m(A) > 0$ is called a focal element, and $C = \bigcup_{m(A) \neq 0} A$ is called a kernel element of mass function m in θ.

Belief and Plausibility Functions. The belief function Bel is defined as

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B). \tag{5}$$

The plausibility function Pl is defined as

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B), \tag{6}$$

$$\text{Pl}(A) = 1 - \text{Bel}(\overline{A}) \text{Pl}(\emptyset) = 0. \tag{7}$$

The belief function Bel(A) measures the total amount of probability that must be distributed among the elements of A. It reflects inevitability and signifies the total degree of belief of A, which constitutes a lower limit function on the probability of A. On the other hand, the plausibility function Pl(A) measures the maximal amount of probability that can be distributed among the elements of A, which describes the total belief degree related to A and constitutes an upper limit function on the probability of A. The relationship between Bel(A) and Pl(A) is shown in Figure 1, and the interval [Bel(A), Pl(A)] is named as belief interval.

Rule of Evidence Combination. Suppose m_1 and m_2 are two mass functions formed based on the information from two different information sources in the same frame of discernment θ and that Dempster's rule of combination (also called orthogonal sum), noted by $m = m_1 \oplus m_2$, is the first one within the framework of evidence theory which can combine two BPA m_1 and m_2 to yield a new BPA:

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) \cdot m_2(C)}{1 - k}, \\ k = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C), \tag{8}$$

where k represents a basic probability mass associated with conflicts among the sources of evidence. Here, k can be determined by summing the products of mass functions of all pairwise sets without intersection and it is often interpreted as a measure of conflict between the data sources. The larger the value of k is, the more conflicting the sources are and the less informative their combination is.

2.2. BP Neural Network Theory. The BP Neural Network [12, 13] is one of the most important and popular techniques in the field of Neural Network, and it is also a kind of supervised learning neural networks, the principle behind which involves using the steepest gradient descent method to get any small approximation. A general model of the BP is shown in Figure 2.

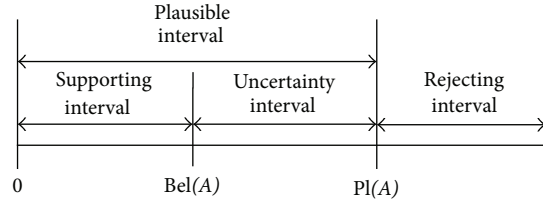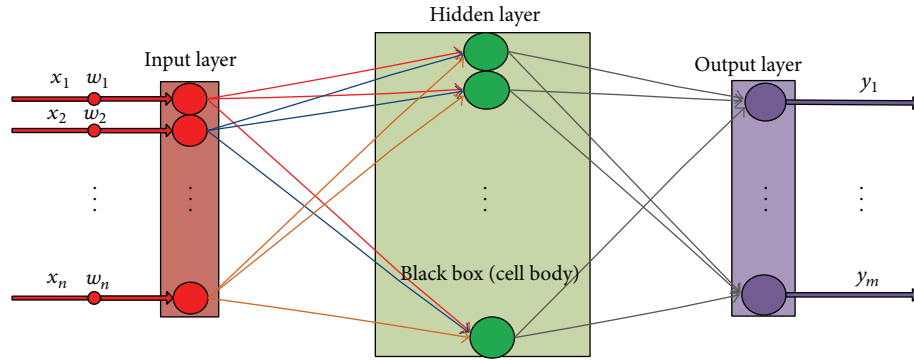In Figure 2, there are three layers in BP Neural Network (BPNN): input layer, hidden layer, and output layer. Two

FIGURE 1: Schematic diagram of Bel(A) and Pl(A).

FIGURE 2: Structure of the BP Neural Network.

nodes of each pair of adjacent layers are directly connected, to form a link. Each link has a weighted value representing the correlation between two nodes. Assuming there are n input neurons, then the weighted values can be updated using a training process described by the following equations in two steps.

(1) *Hidden Layer Stage.* The outputs of all neurons in the hidden layer can be calculated as follows:

$$\text{net}_j = \sum_{i=0}^{n} v_{ij} x_i \quad j = 1, 2, \ldots, m,$$

$$y_j = f_H\left(\text{net}_j\right) \quad j = 1, 2, \ldots, m, \tag{9}$$

where v_{ij} are the weights of neurons, net_j is the activation value of the jth node, y_j is the output of the hidden layer, and f_H is called as the activation function of a node, which is usually a sigmoid function described as follows:

$$y_j = f_H\left(\text{net}_j\right) = \frac{1}{1 + \exp\left(-\text{net}_j\right)}. \tag{10}$$

(2) *Output Layer Stage.* The outputs of all neurons in the output layer can be calculated as follows:

$$\text{Output} = f_O\left(\sum_{j=0}^{m} \omega_{jk} y_j\right), \tag{11}$$

where ω_{jk} are the weights value of output and f_O is the activation function, which is usually a line function. All weights are initially assigned with random values and modified by the delta rule according to the data of learning samples.

3. The Fault Detection Model Based on D-S Evidence Theory

3.1. Detection Model of Sensor Fault. As discussed above, the D-S evidence theory has a strong ability to deal with uncertain, incomplete, and imprecise information. However, there is no general method to calculate BPA in D-S evidence theory. Therefore, in this paper, a structure of three layers is proposed to detect, locate, and isolate the fault sensor, which is shown in Figure 3.

The first layer is data layer, which is used to gather and acquire data. Here, it is supposed that there are N sensors for supervising the drying industry process.

The next two layers are very important, so they are described in detail in the following parts.

3.2. Description of Data Layer. The second layer is data-fusion layer, which is also called data preprocessing. In this part, BP Neural Network is used to get BPA of evidence, because it has many advantages such as robustness for uncertain model, strong matching for nonlinear model, short training period, high accuracy of values, and easily-adjusted network. The data-fusion layer is a two-input and one-output process. Two-input is the supervising data provided by sensors i and j, and one-output is $m_{ij} = (\{OK_i, OK_j\})$, abbreviated as $m_{ij} = (OK)$, which means "the sensors i and j are working well." If the number of sensors is N, and the output number of BP Neural Network is C_N^2.

3.3. Description of Decision Layer. The third layer has actually united different frameworks of discernment. The inputs are the outputs of the first layer, and the prior knowledge acquired from different sensors is used to calculate
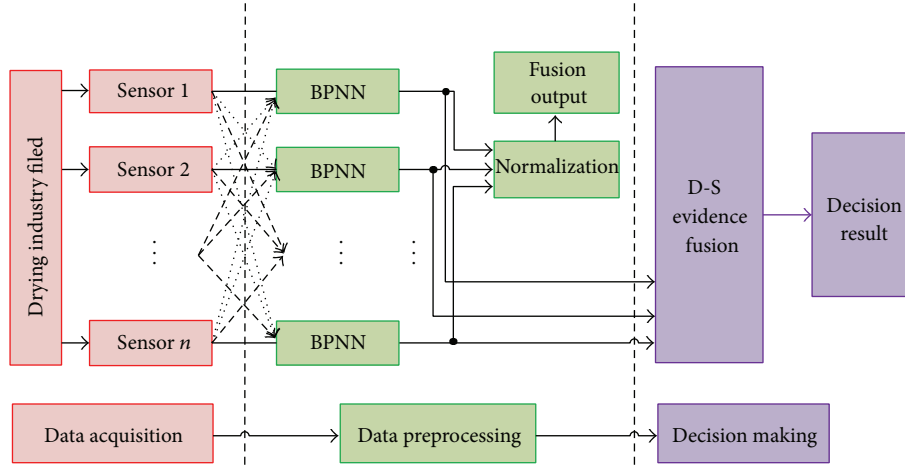
FIGURE 3: Detection model of sensor fault.

the evidence on the different frameworks of discernment. However, according to the requirements of evidence theory, the combination rule is true only under the unified frameworks of discernment. Therefore, in the second layer, the different frameworks of discernment must be united. As well known, it is possible to combine the two evidences within the different frameworks of discernment θ and θ', because they are compatible. In order to combine and merge these evidences, the relationships between θ and θ' must be defined. There are two operations, refinement and coarsening [14–16], which can express the correspondences in the form of compatibility rules. In this paper, the BPAs defined on different framework of discernment are united into a common framework of discernment by the refinement operation, and the BPA of each sensor defined on its own framework of discernment is calculated by the coarsening operation. In fact, a refinement operation unifies compatible elements of θ' to an element of θ, and a coarsening operation is the antagonist relation.

A basic probability assignment of sensor S_i is defined on the set $\theta_i = \{OK_i, KO_i\}$, where OK_i means "sensor i is working well" and KO_i means "sensor j is faulty." Meantime, this framework of discernment is defined as θ_{ij}, which is the Cartesian product of θ_i and θ_j:

$$\theta_{ij} = \left(\left\{OK_i, OK_j\right\}, \left\{OK_i, KO_j\right\}, \left\{KO_i, OK_j\right\}, \left\{KO_i, KO_j\right\}\right). \tag{12}$$

Therefore, m_{ij} is defined on the set of θ_{ij}:

$$m_{ij} \mid 2^{\theta_i \times \theta_j} \longrightarrow [0, 1],$$

$$m_{ij}\left(\left\{OK_i, OK_j\right\}\right) = d\left(V_i, V_j\right),$$

$$m_{ij}\left(\left\{OK_i, KO_j\right\} \cup \left\{KO_i, OK_j\right\} \cup \left\{KO_i, KO_j\right\}\right)$$

$$= 1 - d\left(V_i, V_j\right), \tag{13}$$

where $d(V_i, V_j)$ represents a normalized distance between sensor data V_i and V_j. Of course, the function d, such

as residual generation methods or multivariate statistical methods, can also be used.

Supposing R_k is the refinement operation from θ_{ij} to θ_{ijk} and θ_{ijk} is the Cartesian product of θ_{ij} and θ_{ik}, then the combination rule of m_{ij} and m_{ik} is intersection operation rule described as follows:

$$m_{ij}(A) \oplus m_{ik}(B) = m_{ijk}(C)$$

$$\forall (A, B) \in \left(2^{\theta_i \times \theta_j}, 2^{\theta_i \times \theta_k}\right),$$

$$C \in 2^{\theta_i \times \theta_j \times \theta_k},$$

$$m_{ij}(A) \oplus m_{ik}(B) = m_{ijk}\left[R_k(A) \cap R_j(B)\right],$$

$$R_k(A) \cap R_j(B) = \left(\bigcup_{X \in \theta_k} [\{A, X\}]\right) \cap \left(\bigcup_{X \in \theta_j} [\{B, X\}]\right). \tag{14}$$

Supposing there are M sensors in practice, then there are C_M^2 outputs data preprocessed by BPNN. When it is refinement operation, the two different frameworks of discernment must be compatible. For example, m_{ij} and m_{ik} can be refined, but m_{ij} and m_{kl} cannot be refined. Therefore, there are in total $C_M^2 C_{M-2}^2 / 2 (M > 3)$ kinds of combination modes which cannot be refined.

Supposing S_i is the coarsening operation from θ_{ijk} to θ_i and N_{ij} and N_{ik} are the kernels of m_{ij} and m_{ik}, $\exists (A, B) \in (N_{ij}, N_{ik})$, then

$$S_i\left[R_k(A) \cap R_j(B)\right] \subset P_{\theta_{ijk} \to \theta_i}(D), \tag{15}$$

where $D \in 2^{\theta_i}$ and

$$m_i(D) = \sum_{B \in 2^{\theta_j}, C \in 2^{\theta_k}} m_{ijk}(\{A, B, C\}). \tag{16}$$

Because the combination operation performs all intersections within focal elements of each refined belief assignment,

TABLE 1: Fault preprocessing result of BP Neural Network.

	$m_{12}(A)$	$m_{13}(A)$	$m_{14}(A)$	$m_{23}(A)$	$m_{24}(A)$	$m_{34}(A)$
Fault decision result	0.215	0.213	0.219	0.946	0.281	0.276

TABLE 2: Fault decision result of D-S evidence fusion.

Combination mode	Belief interval of evidence			
	[Bel$_1$ Pl$_1$]	[Bel$_2$ Pl$_2$]	[Bel$_3$ Pl$_3$]	[Bel$_4$ Pl$_4$]
$m_{12} \oplus m_{13}$	[0.306 0.183]	[0.278 0.948]	[0.352 0.753]	
$m_{12} \oplus m_{14}$	[0.382 0.348]	[0.244 0.818]		[0.244 0.303]
$m_{12} \oplus m_{23}$	[0.201 0.156]	[0.193 0.947]	[0.284 0.788]	
$m_{12} \oplus m_{24}$	[0.397 0.282]	[0.910 0.896]		[0.215 0.240]
$m_{13} \oplus m_{14}$	[0.188 0.258]		[0.749 0.828]	[0.278 0.232]
$m_{13} \oplus m_{23}$	[0.489 0.427]	[0.268 0.872]	[0.291 0.920]	
$m_{13} \oplus m_{34}$	[0.293 0.374]		[0.810 0.865]	[0.417 0.348]
$m_{14} \oplus m_{24}$	[0.201 0.338]	[0.881 0.948]		[0.263 0.361]
$m_{14} \oplus m_{34}$	[0.352 0.400]		[0.936 0.979]	[0.338 0.374]
$m_{23} \oplus m_{24}$		[0.834 0.750]	[0.953 0.848]	[0.489 0.427]
$m_{12} \oplus m_{13}$		[0.865 0.876]	[0.809 0.750]	[0.415 0.352]
$m_{12} \oplus m_{13}$		[0.822 0.920]	[0.788 0.850]	[0.395 0.343]
Evidence fusion results	*[0.023 0.041]*	*[1.000 1.000]*	*[1.000 1.000]*	*[0.067 0.079]*
Fault diagnosis results	Sensors 1 and 2 are fault, Sensors 3 and 4 are not working well.			

it must guarantee that all possible intersection operations should have been done. The intersection operations between evidence sources can be expressed only needing to declare the reference set of the corresponding belief assignments, so it is very easy to add new evidence sources (new sensors) without affecting existing functions.

Now, it is possible to get the belief interval of each sensor by (5) and (6), which can be described as belief interval [Bel(OK$_i$), Pl(OK$_i$)]. And then, the new belief interval [Bel(OK$_i$), Pl(OK$_i$)]$_{combine}$ can be calculated by using (8). Thus, the working state of every sensor can be known.

4. Numerical Simulation Analysis in Drying Industry Process

Supposing there are four temperature sensors $\theta = \{H_1, H_2, H_3, H_4\}$ to supervise the drying process and the sampling data of every single sensor is the input of BP Neural Network, then there are $C_4^2 = 6$ outputs, which are $m_{12} = (OK)$, $m_{13} = (OK)$, $m_{14} = (OK)$, $m_{23} = (OK)$, $m_{24} = (OK)$, and $m_{34} = (OK)$. Because of the limitation of space, the preprocessing result of BP Neural Network and part of intermediate result will be given directly in Table 1 without detailed description.

According to the discussion in Sections 3.2 and 3.3, there are $C_6^2 = 15$ kinds of combination modes, but $C_4^2 C_{4-2}^2 / 2 = 3$ kinds of combination modes cannot be refined. Therefore, there are $15 - 3 = 12$ kinds of combination modes that can be refined. Furthermore, according to the method presented in Section 3.3, each two preprocessing results are firstly operated by refinement, then the refinement results are operated by coarsening to get the BPA of each sensor, and finally those

BPAs are fused by using evidence theory combination rule to get the new intervals [Bel(OK), Pl(OK)] shown in Table 2.

According to Table 2, through analyzing every pairwise combination, it can be concluded that the difference of upper and lower limits of belief interval [Bel(OK), Pl(OK)] is very significant, which shows the great uncertainty of sensor state. Therefore, the results of those pairwise combinations cannot be used to decide the sensor state.

In order to fully exploit the maximum of available information and reduce the uncertainty of sensor state, the results of those pairwise combinations should be further fused by evidence combination rule. The new interval [Bel(OK), Pl(OK)] is shown as the font of bold italic in Table 2.

For example, the fusion result of sensor one is [0.026, 0.041]; that is to say, the sum of belief degree of all pieces of evidence that precisely support the proposition "sensor one is working well" is 0.026 and precisely support the proposition "sensor one is fault" is $1 - 0.026 = 0.974$, so the fusion result indicates that "sensor one is fault." Similarly, the sum of belief degree of the proposition "sensor two is working well" is 1.0000 and that of the proposition "sensor two is fault" is $1 - 1.0000 = 0.0000$, so the fusion result indicates that "sensor two is working well." To sum up, sensors one and four are fault, and sensors two and three are working well.

5. Conclusions

In this paper, a modular and generic framework for multiple fault detection and isolation of sensors was presented with a two-layer structure. In data layer, through fully exploiting the sensor data, the data preprocessing could be realized by BP Neural Network to overcome the disadvantages caused by the

change of input sensor data and calculate the BPA of evidence. In decision layer, a modular and generic framework of sensor network for multiple fault detection was presented, and the different but compatible frameworks of discernment were united without affecting the existing relationship, according to the refinement and coarsening operations. Furthermore, new evidence (sensor) could be added to the process very easily and effectively. Therefore, our method had much better expandability, modularity, and flexibility.

After combining all sensors with combination rule, the total uncertainty of sensor state was greatly reduced and the fault sensors could be further exploited by the final fusion results. Numerical simulation results also proved that our new method could be used in practice.

However, if the number of sensors is larger and the combination modes are more, there will be a heavy computation burden. Therefore, this method should be further simplified and optimized in the further study.
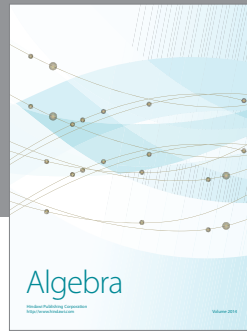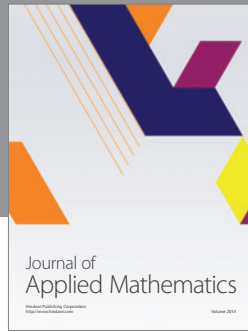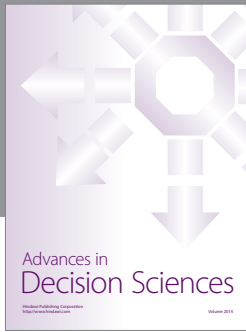
Conflict of Interests

All the authors declared that there was no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals of Mathematical Statistics*, vol. 38, pp. 325–339, 1967.

[2] G. Shafer, *A mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA, 1976.

[3] L. Dymova and P. Sevastjanov, "An interpretation of intuitionistic fuzzy sets in terms of evidence theory: decision making aspect," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 772–782, 2010.

[4] Y. Deng and F. T. S. Chan, "A new fuzzy dempster MCDM method and its application in supplier selection," *Expert Systems with Applications*, vol. 38, no. 8, pp. 9854–9861, 2011.

[5] X. Y. Deng, Q. Liu, Y. Hu, and Y. Deng, "Topper: topology prediction of transmembrane protein based on evidential reasoning," *The Scientific World Journal*, vol. 2013, Article ID 123731, 8 pages, 2013.

[6] X. Y. Su, Y. Deng, S. Mahadevan, and Q. L. Bao, "An improved method for risk evaluation in failure modes and effects analysis of aircraft engine rotor blades," *Engineering Failure Analysis*, vol. 26, pp. 164–174, 2012.

[7] D. Zhu and W. Gu, "Sensor fusion in integrated circuit fault diagnosis using a belief function model," *International Journal of Distributed Sensor Networks*, vol. 4, no. 3, pp. 247–261, 2008.

[8] R. Feng, S. Che, X. Wang, and N. Yu, "A credible routing based on a novel trust mechanism in ad hoc networks," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 652051, 12 pages, 2013.

[9] F. Browne, N. Rooney, W. Liu et al., "Integrating textual analysis and evidential reasoning for decision making in engineering design," *Knowledge-Based Systems*, vol. 52, pp. 165–175, 2013.

[10] D. Niu, Y. Wei, Y. Shi, and H. R. Karimi, "A novel evaluation model for hybrid power system based on vague set and Dempster-Shafer evidence theory," *Mathematical Problems in Engineering*, vol. 2012, Article ID 784389, 12 pages, 2012.

[11] Y. Zhao, J. Li, L. Li, M. Zhang, and L. Guo, "Environmental perception and sensor data fusion for unmanned ground vehicle," *Mathematical Problems in Engineering*, vol. 2013, Article ID 903951, 12 pages, 2013.

[12] Z. Yudong and W. Lenan, "Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network," *Expert Systems with Applications*, vol. 36, no. 5, pp. 8849–8854, 2009.

[13] H. Azami, M.-R. Mosavi, and S. Sanei, "Classification of GPS satellites using improved back propagation training algorithms," *Wireless Personal Communications*, vol. 71, no. 2, pp. 789–803, 2013.

[14] F. Janez and A. Appriou, "Theory of evidence and non-exhaustive frames of discernment: plausibilities correction methods," *International Journal of Approximate Reasoning*, vol. 18, no. 1-2, pp. 1–19, 1998.

[15] J.-P. Steyer, L. Lardon, and O. Bernard, "Sensors network diagnosis in anaerobic digestion processes using evidence theory," *Water Science and Technology*, vol. 50, no. 11, pp. 21–29, 2004.

[16] W. Haitao, L. Qun, and Z. Qiji, "Information fusion of neural networks and evidence theory in fault diagnosis of equipments," *Computer Engineering and Applications*, vol. 22, pp. 13–219, 2004.