Fault Classification of Hydroelectric Generating Unit Based on Improved Evidence Theory

Jiatang Cheng*, Li Ai, Zhimei Duan and Yan Xiong

The Engineering College of Honghe University, Yunnan Mengzi 661199, China

Abstract: Aiming at the problem of the conventional vibration fault diagnosis technology with inconsistent result of a hydroelectric generating unit, an information fusion method was proposed based on the improved evidence theory. In this algorithm, the original evidence was amended by the credibility factor, and then the synthesis rule of standard evidence theory was utilized to carry out information fusion. The results show that the proposed method can obtain any definitive conclusion even if there is high conflict evidence in the synthesis evidence process, and may avoid the divergent phenomenon when the consistent evidence is fused, and is suitable for the fault classification of hydroelectric generating unit.

Keywords: Classification, Fault diagnosis, Hydroelectric generating unit, Improved evidence theory.

1. INTRODUCTION

Hydroelectric generating unit is the key equipment of hydroelectric system. In the operation of hydroelectric generating, vibration is a common phenomenon. Abnormal vibrations could damage the structure of the unit, thereby reduce operating efficiency and unit output. According to statistics, about 80% of the hydropower units fault will be reflected in the vibration signals [1]. Clearly, through the analysis of the vibration signals of hydropower units, thereby establishing an appropriate diagnostic model has become one of the effective means of diagnosis unit failure [2]. However, due to the reasons of occurrence of hydroelectric generating vibrations, faults are complex and various, they include the factors of electrical, mechanical and hydraulic aspects. Therefore, the failure pattern recognition and classification method has become a hot and difficult research. Obviously, the study of hydropower units' status monitoring and fault classification has become an inevitable trend in the current hydropower operations supporting technology development.

In fact, the current hydropower units fault diagnosis model is often designed with vibration signal from the start in order to extract failure characteristic parameters reflecting the vibration causes, and then artificial intelligence methods are used for fault pattern recognition [3, 4]. However, for vibration fault diagnosis system of hydropower units, there are hundreds of signs involved. Meanwhile, for the same kind of fault in a different domain, it will show signs of different failure characteristics. Given the results of different diagnostic symptoms, domains derived different and sometimes even opposite conclusions, resulting in difficult to locate the fault. Thus, only from different aspects of the unit vibration fault diagnosis and multi-sensor information fusion, it is possible to obtain the consistency of explanation or description.

*Address correspondence to this author at the Engineering College of Honghe University, Yunnan Mengzi 661199, China;
E-mail: chjt@163.com

Evidence theory is a decision-level information fusion method, which can be better integrated with information from different data sources, and has been widely used [5, 6]. However, because the evidence theory emphasizes coordination between the evidences, when the conflict evidence is synthesized, it may draw a perverse decision. Therefore, some scholars have proposed improved methods. Yager assigns the probability of the conflict part to the unknown domain, in order to produce a robust integration conclusion. SUN Quan removes the normalization step, the probability assignment between the conflict evidence is assigned to the corresponding proposition according to a certain percentage, resulting in slow convergence. Lee Pil-way improves the method, the probability of conflict part will be allocated to the various propositions weighted, improving the reliability of conclusions. Although these improved methods are able to solve the problem of synthesis between singular evidences, but will exist divergence in the integration of consistent evidences, thus limiting its practical application.

Therefore, on the basis of the rules of evidence theory synthesis study, the hybrid algorithm of the simulated annealing particle swarm optimized (SAPSO) neural network algorithm combined with the improved evidence theory was proposed for the failure mode identification system. At the first, the mean squared Euclidean distance is calculated to obtain the credibility factors of various evidence, and to correct the original evidence. Secondly, according to the Dempster combination rule of evidence theory, the evidence fusion is achieved and obtains an effective classification of failure modes.

2. IMPROVED EVIDENCE THEORY

2.1. Improved Algorithm

For the identification framework of Θ, evidence m_1, m_2 Dempster synthesis rule is given as follows:

$$m(C)=\frac{1}{1-k}\sum_{A_i\cap B_j=C} m_1(A_i)\times m_2(B_j) \tag{1}$$

where, $d_{ij}=\sqrt{\sum_{k=1}^{n}\left[m_i(A_k)-m_j(A_k)\right]^2}$, it indicates the degree of conflict between the evidence.

Obviously, when k=1, m(C) is no definition. When k→1, Eq. (1) will come up with counterintuitive results. Therefore, in order to reflect the difference between the evidence, the concept of distance function is introduced and denoted as follows:

$$d_{ij}=\sqrt{\sum_{k=1}^{n}\left[m_i(A_k)-m_j(A_k)\right]^2} \tag{2}$$

At this time, we can get a distance matrix $D_{n\times n}$:

$$D_{n\times n}=\begin{bmatrix} 0 & d_{12} & \cdots & d_{1n} \\ d_{21} & 0 & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots & 0 \end{bmatrix} \tag{3}$$

For $D_{n\times n}$ as normalized, that is $\tilde{d}_{ij}=d_{ij}/\sqrt{2}$, make $\tilde{d}_{ij}\in[0,1]$. Define the mean square Euclidean distance of evidence A_i to evidence set A is s_i, to indicate coherence between them, that is:

$$s_i=\frac{1}{2n}\sum_{j=1}^{n} d_{ij}^2 \tag{4}$$

where, n is the number of evidence. $s_i\in[0,1)$, its size reflects the degree of difference between A_i evidence and other evidences. The large value of s_i indicates that there is a big difference compared with other evidence, then the credibility factor ε_i of evidence A_i should be small. On the contrary, the credibility factor ε_i should be taken a larger value. Let $\varepsilon_i=f(s_i)$, $f(s_i)$ shall satisfy: (1) 0 <$f(s_i)$≤1, so that the evidence cannot be negated; (2) monotonically decreases with increasing s_i. When s_i is small, $f(s_i)$ should slowly decay, otherwise it will quickly decay to zero. Therefore, $f(s_i)$ should be an exponential curve. Confidence factor is defined as the ε_i :

$$\varepsilon_i=(1-s_i)k^{-s_i} \tag{5}$$

Validation analysis [7], when $k=e^{-1}$, $f(s_i)$ curve satisfies the above requirements. In this case, the use of ε_i on the original evidence to be amended. Set the basic probability assignment (BPA) of the original evidence for $m_i(A_j)$, BPA corrected for $m_i'(A_j)$, then:

$$\begin{cases} m_i'(A_j)=\varepsilon_i m_i(A_j) \\ m_i'(\Theta)=1-\sum_{j=1}^{N}\varepsilon_i m_i(A_j) \end{cases} \tag{6}$$

When the BPA of original evidence has been amendment, the evidence fusion has been accomplished using the Dempster synthetic rule.

2.2. Case Analysis

Example 1: Assuming identification framework Θ = {*A*, *B*, *C*}, m_1, m_2 and m_3 corresponding respectively to the basic probability assignment: $m_1(A)$=0.98, $m_1(B)$=0.01, $m_1(C)$=0.01; $m_2(B)$=0.01, $m_2(C)$=0.99; $m_3(A)$=0.95, $m_3(B)$=0.05. Now by using several methods for information integration, the synthesis results are shown in Table **1**.

Obviously, although two of the three evidence support the proposition A with almost probability 1, yet it led to the evidence 2 supports the proposition C conclusion. As can be seen from Table **1**, since Dempster combination rule cannot effectively deal with integration issues between the evidence of strong conflict, resulting in full support of proposition B conclusions. While Yager law can not get a undoubted judgment, the synthesis result has complete uncertainty; Lee Pil-way law uses the conflicting evidence information to some extent, but the credibility of supporting the proposition A is 0.8474, the converge rate is slower to the same conclusion, and it needs more evidence to compensate for the error of the influence of adverse evidence. However, the proposed method not only effectively performs singular evidence synthesis, but also has a faster convergence rate. At this time, the credibility of supporting proposition A is 0.9805, the uncertainty is 0.0041. Therefore, the synthesis result is reasonable.

Although some improved synthesis rules can better fuse the strong conflictive evidences, yet they give poor results in the synthesis of the consistent evidence. Below, another example demonstrates the feasibility of this method.

Table 1. The fusion results comparison of singular evidence.

Fusion Method	*m*(*A*)	*m*(*B*)	*m*(*C*)	*m*(Θ)
D-S method	0.0000	1.0000	0.0000	0.0000
Yager method	0.0000	0.0000	0.0000	1.0000
Lee Pil-way method	0.8474	0.0167	0.1359	0.0000
Proposed method	0.9805	0.0044	0.0110	0.0041

Table 2. Comparison of the fusion results of consistent evidences.

Fusion Method	*m*(*A*)	*m*(*B*)	*m*(*C*)	*m*(Θ)
Dempster method	0.8000	0.0800	0.1200	0.0000
Yager method	0.1200	0.0120	0.0180	0.8500
Lee Pil-way method	0.6249	0.1719	0.2032	0.0000
Proposed method	0.7997	0.0801	0.1200	0.0002

Example 2: Assuming identification framework $\Theta = \{A, B, C\}$, m_1, m_2, m_3 corresponding respectively to the basic probability assignment: $m_1(A)=0.5$, $m_1(B)=0.2$, $m_1(C)=0.3$; $m_2(A)=0.4$, $m_2(B)=0.3$, $m_2(C)=0.3$; $m_3(A)=0.6$, $m_3(B)=0.2$, $m_3(C)=0.2$. Obviously, these evidence supports the proposition A. At this time the extent of the conflict between the evidence is weak. Here, still using 4 kinds of evidence synthesis methods in Example 1, the integration results are shown in Table **2**.

Obviously, for integration issues arising between relatively consistent evidences, Dempster method can get satisfactory result. Lee Pil-way synthesis method, in a relatively small credibility, gets only a maximum of 0.6249, which shows some divergence. And for the proposed method, the basic probability assignment is 0.7997, the convergence rate is faster, and the result is almost the same as that the Dempster synthesis method. Therefore, it showed that the method is also applicable to the information fusion for consistent evidences.

3. HYDROPOWER UNITS FAULT CLASSIFICATION

3.1. The Basic Probability Assignment Constructor

As the SAPSO combines the global optimization capabilities of particle swarm algorithm and the ability to jump out local optimal solution of simulated annealing algorithm, so that the algorithm can quickly find the global optimal solution. Therefore, this paper combines simulated annealing particle swarm algorithm and BP neural network to form a SAPSO-BP hybrid algorithm, as the basic probability distribution function of a hydroelectric generating integrated fault diagnosis model. The use of SAPSO algorithm optimizes BP network weights, threshold parameters, and the fitness function to a minimum. SAPSO-BP algorithm's specific implementation process is as follows [8]:

① Determine the topology of BP neural network.

② Initialize the particle swarm.

③ Determine the fitness function, select the MSE of neural network output as a PSO fitness function f.

④ Fitness evaluation. The position of the each current individual particle and fitness store in the p_i; the position of best individual particle of the global optimal solutions p_{best} in the population and fitness store in the p_g.

⑤ Determining the initial temperature, $t_0 = f(p_g)/\ln 5$.

⑥ Determine adaptation value: using eq.(1) to determine the current temperature, the adaptation value of each p_i.

$$TF(p_i) = \frac{e^{-(f(p_i)-f(p_g))/t}}{\sum_{i=1}^{N} e^{-(f(p_i)-f(p_g))/t}} \tag{7}$$

⑦ Adopting roulette strategy determine a substitute value p_g of the global optimum from all p_i, using formula (2) update the velocity and position of each particle.

$$\begin{aligned} v_{id}(k+1) &= \phi\{v_{id}(k) + c_1 r_1 (p_{id}(k) - x_{id}(k)) + c_2 r_2 (p_{gd}(k) - x_{id}(k))\} \\ x_{id}(k+1) &= x_{id}(k) + v_{id}(k+1) \end{aligned} \tag{8}$$

where: v_{id} and x_{id} are the current particles velocity vector and the position vector, respectively. k is the current number of iterations. c_1 and c_2 are learning factors. ϕ is the shrinkage factor, $\phi = 2/\left|2 - c - \sqrt{c^2 - 4c}\right|$, $c = c_1 + c_2$. r_1 and r_2 are the two random numbers.

⑧ Extreme value update. Update the p_i value of each particle and the p_g value of population.

⑨ Back temperature operation, annealing method to select $t_{k+1} = \lambda t_k$, λ annealed constant.

⑩ Stop searching. If the stop condition is met, the search is stopped, otherwise go to step ⑥.

3.2. Fusion Decision

Through the characteristic parameters of the vibration spectrum and vibration amplitude signs space of hydropower units were extracted, respectively, two independent SAPSO-BP models were used to carry out fault classification for different sign space. Then their outputs were chosen as evidence body of improved evidence theory to obtain the corresponding basic probability assignment, and integration of decision-making in order to identify the failure modes of hydropower units. When the evidence is synthesized, in accordance with decision rules of the biggest credibility factor, that is to say that the diagnostic conclusion is the greatest credible proposition.

4. ALGORITHM VALIDATIONS

4.1. Object Description

Hydroelectric generating unit is a complex nonlinear system that manifested fault mostly in the form of vibration. Fault causes are very complex, and the failure type has a gradation and irregularity. Hydropower unit based on different factors is prone to vibration, the vibration fault types can be divided into mechanical, electromagnetic and hydraulic categories [9]. A detailed analysis of the causes of failure and the characteristics of these three types of vibration was performed. Furthermore, these three failures interact. For example, when hydropower unit is affected by hydraulic factors and led to the rotating part vibration, it will cause changes in the motor air gap magnetic field, and electromagnetic force damping or exacerbate units rotating part vibration. Therefore, since hydroelectric generating vibrations are the results of combined effect of mechanical, electrical and water, and often lead to multiple faults occurring simultaneously, it requires to search for a better fault diagnosis method to achieve effective vibration fault diagnosis of hydropower unit.

On the basis of the vibration fault mechanism of hydropower unit, the fault feature space is divided into vibration spectrum signs subspaces and vibration amplitude signs subspaces, and the fault space is divided into generator failure subspace, rotor bearings failure subspace and overcurrent component fault subspace *etc*. The multi-domain characteristic parameters obtained for each group of fault symptom subspace and fault subspace, were designed respectively through the SAPSO-BP models for modeling analysis, to complete the primary diagnosis of hydropower unit. Then, the results of the symptom domains primary diagnostic are obtained through improved evidence theory to arrive at a final diagnosis conclusion.

Therefore, some parameters, such as the unit vibration spectrum sign domain of six components are as follows:$(0.18\sim0.2)f_0$, $0.5f_0$, f_0, $2f_0$, $3f_o$, $>3f_o$(f_0 for the rotation frequency), and the vibration amplitude of the vibration and speed, load, flow, excitation, oil temperature and other five kinds of fault symptoms parameters associated with the amplitude are chosen as fault identification information. Moreover, the three kinds of common faults found in hydroelectric generating unit, such as rotor unbalance, rotor misalignment and rubbing acted as fault diagnosis domain, and constitute a unit vibration fault diagnosis recognition framework $\Theta=\{y_1,y_2,y_3\}$. Meanwhile, in the simulation, the encoding rules of the unit vibration fault are given as follows: rotor unbalance (100), rotor misalignment (010), Rubbing (001).

4.2. Primary Diagnosis

According to the relational data of hydropower unit's vibration fault types and feature parameters, during the initial diagnosis, the main parameters SAPSO-BP model are set as follows: population size is 30, the maximum number of iterations is 50, learning factor c_1=2.8, c_2=1.3, annealing constant λ=0.5. In the MATLAB programming environment, the SAPSO-BP1 algorithm fitness function curve is shown in Fig. (**1**).

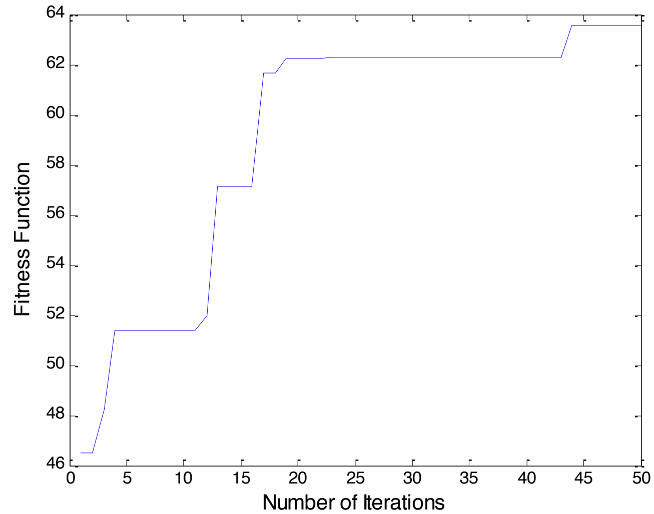

Fig. (1). Fitness change curve of SAPSO-BP1 algorithm.

4.3. Fusion Analysis

Using SAPSO-BP algorithm for the initial fault diagnosis, and its outputs were normalized (to retain four significant figures), the improved evidence theory algorithm in this paper performed evidence fusion and comparing with several common improvement methods synthesis results, the comparison result of the part of the test samples shown in Table **3**.

From Table **3**, it can be seen that the degree of conflict between the evidence is not great. Dempster combination rule can effectively strengthen the credibility of high proposition, weakening the credibility of the low proposition. Yager method assigns conflict evidence to unknown areas, resulting in sample two and five diagnostic results uncertainty. SUN Quan law, to some extent, uses the information of conflict evidence, and thus reduces uncertainty of the results. But the proposition credibility of the diagnostic conclusion is not very high, and the maximum credibility is 0.8145, and the maximum uncertainty is 0.2181. Lee Pil-way method of the average allocation of conflict evidence to each proposition, fully considers the impact of the conflict evidence on the proposition credibility of improved diagnostic results. The synthesis of weak conflict evidence show that the convergence rate is still slow. Owing to use the concept of distance function, the proposed method reflects the difference between the evidence. When the evidence is consistent, the proposed method can draw almost the same fusion results as with Dempster combination rule. Meanwhile, for the No.5, there is a certain conflict within the diagnosis results of SAPSO-BP1 and SAPSO-BP2, where the extent of the conflict k=0.5939. However, when using this method for fault pattern recognition, it can still correctly identify the failure mode of the unit, and the confidence is higher than the corresponding values obtained through Lee Pil-way method and Dempster law method. Hence, the results verify the effectiveness of the improved method of fault classification.

Table 3. Comparison of diagnostic results.

No.	Diagnostic Methods	$m(y_1)$	$m(y_2)$	$m(y_3)$	$m(\Theta)$	Diagnostic Conclusions	Results Evaluation
1	SAPSO-BP1	0.7381	0.1496	0.1123	0	y_1	yes
	SAPSO-BP2	0.6752	0.0945	0.2303	0	y_1	yes
	Dempster method	0.9256	0.0263	0.0481	0	y_1	yes
	Yager method	0.4984	0.0141	0.0259	0.4616	y_1	yes
	SUN Quan method	0.7463	0.0570	0.0860	0.1107	y_1	yes
	Lee Pil-way method	0.8246	0.0705	0.1049	0	y_1	yes
	Proposed method	0.9256	0.0263	0.0481	0	y_1	yes
2	SAPSO-BP1	0.5896	0.3012	0.1092	0	y_1	yes
	SAPSO-BP2	0.6487	0.2615	0.0898	0	y_1	yes
	Dempster method	0.8120	0.1672	0.0208	0	y_1	yes
	Yager method	0.3825	0.0788	0.0098	0.5289	Uncertain	no
	SUN Quan method	0.6416	0.1965	0.0515	0.1104	y_1	yes
	Lee Pil-way method	0.7100	0.2276	0.0624	0	y_1	yes
	Proposed method	0.8120	0.1672	0.0208	0	y_1	Yes
3	SAPSO-BP1	0.1067	0.7920	0.1013	0	y_2	yes
	SAPSO-BP2	0.2251	0.6894	0.0855	0	y_2	yes
	Dempster method	0.0415	0.9435	0.0150	0	y_2	yes
	Yager method	0.0240	0.5460	0.0087	0.4213	y_2	yes
	SUN Quan method	0.0842	0.8145	0.0425	0.0588	y_2	yes
	Lee Pil-way method	0.0939	0.8581	0.0480	0	y_2	yes
	Proposed method	0.0415	0.9435	0.0150	0	y_2	Yes
4	SAPSO-BP1	0.1764	0.8178	0.0058	0	y_2	yes
	SAPSO-BP2	0.0911	0.6974	0.2115	0	y_2	yes
	Dempster method	0.0273	0.9706	0.0021	0	y_2	yes
	Yager method	0.0161	0.5703	0.0012	0.4124	y_2	yes
	SUN Quan method	0.0556	0.7942	0.0333	0.1169	y_2	yes
	Lee Pil-way method	0.0712	0.8828	0.0460	0	y_2	yes
	Proposed method	0.0273	0.9706	0.0021	0	y_2	Yes
5	SAPSO-BP1	0.1005	0.5472	0.3523	0	y_2	no
	SAPSO-BP2	0.0287	0.3131	0.6582	0	y_3	yes
	Dempster method	0.0071	0.4219	0.5710	0	y_3	yes
	Yager method	0.0029	0.1713	0.2319	0.5939	Uncertain	no
	SUN Quan method	0.0272	0.3330	0.4217	0.2181	y_3	yes
	Lee Pil-way method	0.0413	0.4268	0.5319	0	y_3	yes
	Proposed method	0.0070	0.3948	0.5981	0	y_3	Yes

CONCLUSION

SAPSO neural network algorithm is used to construct the probability distribution function based on evidence theory, a combination of currently troubleshooting techniques part of the excellent methods. The study intends to build a fault diagnosis model of hydropower units based on improved evidence theory information fusion, which can effectively prevent the emergence of a single model misdiagnosis phenomenon, with a strong fault tolerance.

Improved theoretical evidence are used in keeping the Dempster synthesis rules remain unchanged, and by introducing credibility factor, corrected the original evidence, reduced the impact of singular evidence on diagnostic results. Meanwhile, for the synthesis of consistent evidence can still get better failures classification results, provide a reference idea for fault classification of other devices.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Bide, Z.; Yuan, T.; Ping, Z.J. Vibration fault diagnosis of hydropower units based on Choquet fuzzy integral. *J. Vibr. Shock* **2013**, *32*, 61-66.

[2] Yun, Z.; Ping, M.X.; Jin, W.; Hui, L.C. Fault diagnosis of hydroelectric sets based on hierarchical fuzzy neural network. *J. Hohai Uni. (Nat. Sci.),* **2009**, *37,* 335-340.

[3] Rong J.; Gang, H.; Hua, W.; Hui, X.J. Vibration fault diagnosis of hydroelectric generating unit by least squares support vector machine based on improved particle swarm optimization. *J. Hydraul. Eng.,* **2011**, *42,*373-377.

[4] Chaoshun, L.; Jianzhong, Z. Parameters identification of hydraulic turbine governing system using improved gravitational search algorithm. *Energ. Conv. Manage*., **2011,** *52,* 374-381.

[5] Liu, J.; Xu, G.; Gao, Y.; Yang, Y. The algorithm of power system security detection comprehensive decision based on dempster-shafer theory improved combination rule. *Trans. China Electrotech. Soc*., **2011**, *26*, 247-255.

[6] Jiatang, C.; Li, A.; Zhimei, D. Transformer fault diagnosis based on improved evidence theory and neural network integrated method. *Power. Sys. Protec. Cont.,* **2013**, *41*, 92-96.

[7] Weihua L.; Shenggang, Z. Fault classification based on improved evidence theory and multiple neural network fusion. *J. Mech. Eng.*, **2010**, *46*, 93-99.

[8] Yong, L.; Jianjun, W.; Lihua, C. Simulated annealing particle swarm optimization algorithm of optimal load dispatch in power plant. *Proc. CSU-EPSA*, **2011**, *23*, 40-44.

[9] Zhang, X.Y. ; Zhou, J.Z.; Wang, C.Q.; Li, C.S. Vibrant fault diagnosis for hydroelectric generator unit considering overlapping fault patterns. *Power. Sys. Protec. Cont.,* **2012**, *40*, 8-14.

Received: November 10, 2014 Revised: January 6, 2015 Accepted: January 28, 2015