2009 International Conference on Machine Learning and Computing
IPCSIT vol.3 (2011) © (2011) IACSIT Press, Singapore

D-S Evidence Theory in the Application of Turbine Fault Diagnosis

Xia Fei [1 +], Zhang Hao [1], Xu Longhu [1], Peng Daogang [1]

[1] College of Electric Power and Automation Engineering,
Shanghai University of Electric Power, Shanghai, China

Abstract. According to the complexity of turbine on power plant, a kind of fault diagnosis method based on information fusion is proposed in this paper. Multi-sensor data is adopted and D-S evidence theory is applied, thus the uncertainty of information can be effectively overcome. Considering the data credibility of various information sources, the data fusion method is applied to the simulation experiments of steam turbine fault diagnosis. Experimental results show that multi-sensor information is used in this method which can carry out effective fault diagnosis.

Keywords: information fusion, fault diagnosis, D-S evidence theory, credibility

1. Introduction

Electric industry is the basic industry of the national economy [1], power system operation conditions directly affects the normal operation and the steady development of the national economy. Mechanical equipments in the modern industrial processing treatment are developing with such characters as maximization, complication, high-speed, automation. As the device productivity is more and more higher, the mechanical structure is more and more complex, and the interrelationship and coupling degree of different process parts is more and more close, the whole production will be interrupted if a mechanical fault is appeared. In order to master the equipment operating condition [2], avoid the catastrophic equipment accident, and diagnose various complex equipment vibration in time, work of equipment fault diagnosis technique, especially the online monitoring and diagnosis [3, 4] of critical units in production, has been carried out successively at home and abroad since the 1980s, which has supported the safe operation of equipments preferably, and achieved a remarkable economic benefits. The combination of steam turbine and generator calls turbo-generator unit, the demand of reliable operating in conditions of high-speed, full-loaded, continuous has become higher and higher with the capacity of turbine-generator unit continuously increasing. The monitoring and effective fault diagnosis of the turbo-generator unit operating conditions is important measure to guarantee the safe and normal operation of turbo-generator unit.

The sensors' performance must be better in the turbine fault diagnosis system as the working condition of turbine is poor. Meanwhile, there are various reasons for turbine fault, and the reliability of sensors used for data transmission is affected by a lot of factors. When the sensor has faults, system will lose efficiency because sensor data in the system top-end are inaccuracy. Therefore, how to deal with the uncertainty caused by the required information which can't be fully supplied by single sensor, and how to guarantee system stability and efficiency are the basic problem. Turbo-generator unit is the important equipment in power production, the equipment structure is complex and the operating condition is special. Therefore, when carrying out fault diagnosis in large complex system such as turbo-generator unit, the more reliable and accurate detection and diagnosis on turbo-generator unit can be done only by getting multi-dimensional

+ Corresponding author. Tel.: + 86-21-65430410; fax: +86-21-65430410.
E-mail address: xiafeiblue@163.com.

information from the same object in many aspects and utilizing them comprehensively. In other words, more reliable fault diagnosis results can be obtained only by integrating judgements from various sensors.

Multi-sensor data fusion is the multi-level and aspect processing of the data from many sensors. Then new meaningful information will exist. And this information could not be achieved from any single sensor. Data fusion is completed automatic checking, contacting, evaluation and compound processing of the data from many signal sources. Multi-sensor data fusion could work on various levels in sensor signal process. According to the abstract extent of sensor signal, data fusion could be divided into three levels, and they are pixel-level fusion, feature level fusion and decision-level fusion. Decision-level fusion, which is senior fusion, makes the optimal decision basing on certain criteria and the credibility of each decision. Sensor data will be fused from decision-level in order to carry out fault diagnosis in this paper.

In the coming text, fault types of the steam turbine generator unit are introduced in part 2. The data fusion methods and their achievement in fault diagnosis are discussed in part 3. The application in fault diagnosis being proved by turbine simulation experiment is described in part 4. Finally, part 5 makes a conclusion of the whole paper.

2. Fault Type

There are many manifestations of the faults for steam turbine generator sets [5]. There are Eigen-values differences for units vibration signals the various faults correspond. First, the monitoring signals got by sensors are pre-processed, and then the representative Eigen-values that reflect the units' fault signals are got. Rotors are the very important parts in steam turbine generator sets. They have complex running status. Under the high temperature and high speed conditions, they undertake huge centrifugal force and torsion stress produced by the passing power. And at the same time, they undertake huge thermal stress. And they could even result bending and vibration. The usual steam turbine generator rotor faults are of the next types as follow.

2.1. Rotor imbalance

Rotor imbalance faults are divided into two status of rotor quality eccentric and rotor part defect. Rotor quality eccentric is caused by rotor's manufacturing error, assembly error and material uneven. These are called initial imbalance. Rotor part defect is caused by corrosion, dielectric structure wearing and fatigue stress for rotors in running process. These cause the rotor's parts such as impellers and leaves are partly damaged or there are fragments flying out of them. And they could cause new imbalance of rotors, and cause abnormal vibration. Rotor quality eccentric and rotor part defect are two different fault types, but they have the same mechanism.

2.2. Rotor hot bending

There are many reasons for rotor hot bending such as too huge stress on the shafts, uneven texture in rotors, water influence in the cylinders, rotor partly loses bear edge or has cracks, active and static friction, inadequate warm-up and rotor's central hole has oil or water.

2.3. Rotor cracks

There are many reasons for rotor cracks, and the main two reasons are as follow. At first, the neighbour start of the cooler has huge defects on rotors, when the rated parameter steam washing engine are used, steam would condense and exothermic on the surface of metal. Then the cylinder and rotor's temperature on their surface would sharply rise. Especially the rotor has large surface area, so its speed of temperature rising is much faster, and as a result, hot-pressing stress on the rotor's surface would rise. When the hot-pressing stress is beyond the metal's yield limits, partly plastic deformation could happen here. As the rotor is continuously heated, the thermal stress it undertakes is reduced. But the plastic deformation could not recover automatically as the reduction of the rotor hot-pressing stress. It could result remaining tensile stress in the areas around. Under high temperature condition, the remaining stress could reduce as time passes, and this is called metal relaxation phenomenon. In the front steam seal in the max shaft diameter and regulation of class, this phenomenon become more obviously. If the neighbour start times increases, the extent of damage could be serious, and then the rotor's surface could quickly produce fatigue cracks.

3. Diagnosis Method

In the field of fault diagnosis, because of the complex of the device and the instability of running condition, the fault's feature signals reflected by the device are uncertainty. Data fusion is the most effective method of combined these signals. This method is based on the classic D-S evidence theory which the combined sensors' data in decision level. It could timely master the abnormal status and the early signs in the process of device running. Then it analyzes the causes of faults, and takes the corresponding actions, and eradicates the fault in budding status. This measure could avoid and reduce happen of huge accidents.

3.1. Evidence theory

Evidence theory [6] is a reasoning method of uncertainty. It is promoted by Dempster initially in 1967. He takes the method of many-valued mapping gets the up and down limit of the probability. And then it is promoted and has become evidential reasoning by Shafer in 1976. So it is called D-S theory. D-S evidential theory is used as a reasoning method of uncertainty, has large development these years, and receives more and more concern. It grasps the unknown and uncertain in the problems better than traditional probability theory, and it provides a very useful relative evidence synthesis method. And so it could fuse a number of evidences provided by sources. This method is successfully applied in target recognition and other fields. The main content of D-S theory is as follow.

In the D-S reasoning method, the basic entity is the identification of the framework. It consists of all the incompatible proposition's probable values (called singleton). This collection is defined as follow:

$$\Theta = \{H_n\} \tag{1}$$

But the basic probability distribution function expresses the precise trust for object recognition H given by a signal source S_j, and for all the j, $j \in [1,Q]$, and the Q represents the number of signal sources. We could get the following basic probably divisions function for the signal source Sj.

$$m_j : 2^{\Theta} \to [0,1] \tag{2}$$

We could prove that this function has the following features.

$$m_j(\Phi) = 0 \tag{3}$$

$$\sum_{H \in \Theta} m_j(H) = 1 \tag{4}$$

The D-S evidence theory provides a very useful synthesis of the formula. It makes us could compose many evidences provided by different evidence sources. The formula is as following.

$$m(H) = \frac{1}{1-k} \sum_{H_i \cap H_j \cdots = H} m_1(H_i) m_2(H_j) \cdots \qquad \forall H \in 2^{\Theta} \tag{5}$$

And the k represents the evidence's conflict level. The coefficient *1/(1-k*) is called normalization factor. Its function is to avoid assigning the non-zero probability to the empty set when coalescence.

3.2. Credibility function

Based on complexity of the huge equipment and the uncertainty of the running circumstances, different signal sources provide different signals in many aspects. We could calculate the level of the signal source Sj according to the different sample set of signal sources.

At first, we could divide every signal source sample set's data into study set and test set according to different fault types. We could set up initial histogram basing on the study condensing sets. The histogram at the same time is the initial evaluation of the study condensing set distribution. We assign that the number of types concluded is K. To the types concluded in the initial histogram, we could optimize it according to the signal guidelines. We could merger the similar types, and then get the most optimistic histogram concluding the type number of $k_{n,j}$. The most optimistic histogram represents the evaluation value of $\hat{\lambda}^{a}_{n,j}$ for study condensing sample distribution. Next, basing on the most optimistic histogram of signal source study set, we could get the test set sample's histogram. Here, we take the same signal source condensing histogram type number of $k_{n,j}$ [7, 8] and the same step to construct the test set's histogram. The histogram represents the evaluation $\hat{\lambda}^{v}_{n,j}$ of sample distribution in test set. At last, from the study condensing sample distribution $\hat{\lambda}^{a}_{n,j}$ and the test condensing sample distribution $\hat{\lambda}^{v}_{n,j}$ from the same signal source, we could calculate the Hellinger distance between the two.

$$Hell(\hat{\lambda}_{n,j}^{a},\hat{\lambda}_{n,j}^{v})=1-\sum_{r=1}^{k_{n,j}}\sqrt{\hat{\lambda}_{n,j}^{a}(B_r)\hat{\lambda}_{n,j}^{v}(B_r)} \tag{7}$$

So we get the signal source's credibility function.

$$q_{n,j}=1-Hell(\hat{\lambda}_{n,j}^{a},\hat{\lambda}_{n,j}^{v}) \tag{8}$$

When $q_{n,j}$=1, it represents that the signal source's credibility is pretty good and it has pretty high quality. Under this condition, when we fuse the signals from many signal sources, we could determine the credibility of the signal source.

3.3. Fusion diagnosis

After considering the reliability of different sensors calculated by credibility functions, we could use the fusion rules to fuse the signals provided by every sensor. The steps are as following.

1) Get the probability distribution to every signal sources, the calculation formula is as following.

$$m_j(H)=\bigoplus_{n=1}^{N} m_{n,j}(H) \tag{9}$$

2） Fuse the different signal sources probability distribution, and then get the final distribution results, the formula is as following.

$$m(H)=\bigoplus_{j=1}^{J} m_j(H) \tag{10}$$

3） According the criteria, we judge the type the fault to be recognized belongs to. The criteria this paper takes is the probability criteria signed as BetP. The BetP finally is got from the division result being fused as the following formula.

$$BetP(H_n)=\sum_{N\subseteq\Omega}\frac{|H_n\cap H|}{|H|}m(H) \tag{11}$$

The max value represents the recognition result of the assumption recognition objects. And the |.| represents the basic number containing the set H.

The fault diagnosis mode's diagram of the technology of multi-sensor data fusion is as follow.

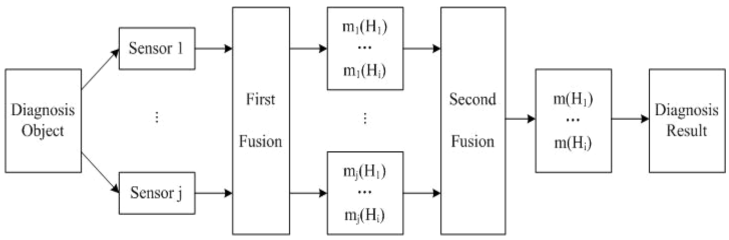

Figure1.The fault diagnosis mode's diagram of the technology of multi-sensor data fusion

4. Stimulation Experiment

There are many types of the usual faults of steam turbine. Next, we shall introduce a practical example basing on the turbine rotor's fault diagnosis to description that how to apply the method above to the fault diagnosis area. Because the usual fault that the turbine rotor has concludes three fault mode of rotor imbalance, rotor hot bending and rotor cracks. That means the recognition framework concludes three types

TABLE I. THE FUNCTION VALUES OF TURBINE ROTOR FAULT RELIABILITY

	$m(u_1)$	**$m(u_2)$**	**$m(u_3)$**	**$m(\Theta)$**
Current Sensor m_1	0.4626	0.2763	0.0919	0.1692
Temperature Sensor m_2	0.3127	0.4729	0.0077	0.2067
Vibration Sensor m_3	0.4139	0.3010	0.0877	0.1974

of rotor imbalance, rotor hot bending and rotor cracks. Every fault mode's characteristic parameter is achieved by three different sensors. In this example, the current sensor S_1, the temperature sensor S_2, and the vibration sensor S3 are used as three signal sources to offer evidence. Table I shows the basic probability division of each signal source as above.

In this table, u_1 represents the rotor's imbalance, and u_2 represents the rotor's hot bending, and u_3 represents the rotor cracks. The recognition framework is $\Theta=\{u_1,u_2,u_3\}$. The system at first fuses the evidence of m_1 and m_2, and then calculates the whole confliction coefficient value k_{12}:

$$k_{12}=\sum_{u_i\cap u_j=\Phi} m_1(u_i)m_2(u_j)=m_1(u_1)m_2(u_2)+m_1(u_1)m_2(u_3)+m_1(u_2)m_2(u_1)+m_1(u_2)m_2(u_3)+m_1(u_3)m_2(u_1)+m_1(u_3)m_2(u_2)=0.3830$$

From the following formula we could calculate the assigned reliability function value u_1 of rotor's imbalance fault.

$$m_{12}(u_1)=\frac{\sum_{u_i\cap u_j=u_1} m_1(u_i)m_2(u_j)}{1-k_{12}}=\frac{m_1(u_1)m_2(u_1)+m_1(u_1)m_2(\Theta)+m_1(\Theta)m_2(u_1)}{1-k_{12}}=\frac{0.1447+0.0956+0.0529}{1-0.3830}=0.4752$$

And using the same method we could get that $m_{12}(u_2)$=0.4341, $m_{12}(u_3)$=0.0340, $m_{12}(\Theta)$=0.4341。

According to the data fusion D-S rules, when there are more than one signal source, we could fuse every two signal sources of them by turns. When we fuse the results of the fusion value of current and temperature sensors and the evidence of violation sensors, the table below could be achieved.

TABLE II. THE FAULT DIAGNOSIS RESULTS OF TURBINE ROTOR

	m(u1)	m(u2)	m(u3)	m(Θ)
m12	0.4752	0.4341	0.0340	0.0567
m123	0.5175	0.3847	0.0242	0.0184

From the calculations, we could get the result through fusion. The basic probability value of uncertainty decreases to 0.0184. When the basic probability value decision-making method is taken, the final result is the diagnosis of rotor imbalance with 0.5175(fault reliability value).

We could get the result from this example where the two reliability value got from sensor are approached. The value got from the other sensor conflicts with the previous two types in some extents. At this time, if we only use one function to recognize the fault modes, we would not decide which fault mode it belongs to. That means it'll produce contradictions. In order to get the specious fault diagnosis results, we use data fusion algorithm to fuse the three sensor reliability function values, and at last recognize the faults through the probability criterion of BetP. At last we get the final fused fault recognition mode. The result shows that comparing the fused reliability value with the single-sensor's reliability value, the previous value increases the practical goal's reliability function distribution values. It highly decreases the system's uncertainty and increases its recognition capacity for fault modes.

5. Conclusion

Characteristic parameters of information from multi-sensors can be used in fault diagnosis while adopting data fusion method. The information credibility and fault tolerance ability of diagnosis system are improved via using data fusion method. It has an important application in fault diagnosis of mechanical equipment.

6. References

[1] Y.He, W.Lu and F.Chu, "The research and application of status checkout and fault diagnosis for generator unit ," Journal of Vibration, Measurement & Diagnosis, vol.26(4), Dec. 2006, pp.269-273.

[2] W.Li, P.Wang and Q.Xia, "Application of Steam Turbine On-line Condition Monitoring and Failure Diagnosis Systems," Power Equipment, vol.18(3), Jun. 2004, pp.167-170.

[3] T.Chang, L.Tian, J.Su and Q.Duan, "Design of online evaluation system of running state for electric power system equipment," Journal of North China Electric Power University, vol.32(4), Jul. 2005, pp.44-47.

[4] C.Chen, Q.Li, Y.Liu and Y.Wang, "Intelligent fault diagnosis method for turbo-generator unit," Proceedings of the CSEE, vol.22(5), May 2002, pp.121-124.

[5] J.Guan. Power Source on Power Plant Beijing: China Electric Power Publishing, 2007.

[6] B.Liu, B.Xiao and H.Fu. "Discussion of Data Fusion Algorithm Based on D–S Evidential Theory and Fuzzy Mathematics," Microprocessors, Jun. 2007, pp.70-72.

[7] Azza Ouled Zaid, Chiristian Olivier, Olivier Alata and Francois Marmoiton, "Transrom image coding with global thrsholding: Application to baseline JPEG," Pattern Recognition Letters, vol.24, 2003, pp. 959 – 964.

[8] Olivier C, Courtellemont P and Colot O, "Comparison of Histograms: a Tool for Detection," European Journal of Diagnosis and Safety in Automation, vol.4(3), 1994, pp.335-355.