

Multiple-Fault and Degradation Degree Simultaneous Diagnosis Based on Artificial Neural Network and Hidden Semi-Markov Model

Huang Darong, Chu Xiaoyan, Zhao Ling, Tang Jianping
Chongqing Jiaotong University
Institute of Information Science and Engineering
Chongqing, China

Abstract—Based on the theory of Artificial Neural Networks and Hidden Semi-Markov Model, a hierarchical diagnosis network (ANN-HSMM) is proposed with the respect to multiple-faults and their degradation degree simultaneous diagnosis. ANN-HSMM consists of several sub-networks, and aims at dividing a larger pattern into several smaller subspaces, so that the sub-network can be trained on the subspace, respectively, and the whole network is capable of multiple-faults and their degradation degree simultaneous diagnosis. The experiment results show that ANN-HSMM can not only recognize multiple-faults diagnosis, but also achieve the degradation degree diagnosis for the corresponding faults, and is available for real-time condition monitoring and diagnosis.

Keywords—Multiple-Fault; Degradation Degree; ANN; HSMM

I. INTRODUCTION

It is the fault diagnosis and the identification for corresponding fault degradation degree that is the basis to maintain the critical transmission parts for machinery and equipment to operation smoothly and steady. In [1], aiming at the problem of fault diagnosis and fault-tolerant control for system with delayed measurements and states, an innovative solution is proposed; in [2], a fractal dimension calculation method for discrete signals in the fractal theory was proposed and was applied to extract the fractal dimension feature vectors and classified various fault types; in [3], a new safety performance evaluation of the fault-prediction technology was research based on misclassification cost, and the future development trend of the fault prediction technology was discussed.

Meanwhile, Artificial Neural Network (ANN) and Hidden Markov Model (HMM) are applied in the field of the fault diagnosis, and the studies about Hidden Semi-Markov Model (HSMM) are launched subsequently, and the diagnosis accuracy is improved to some extent. Aiming at the problem that the static multiple fault models in existence can not fit the need of the more and more complicated equipments as the development of electron technique, [4] put forward a dynamic multiple fault model which is based on Hidden Semi-Markov Model (HSMM) and make the process simpler, meanwhile the original dynamic multiple fault problems are decomposed into several independent sub-problems and the Binary Particle

This work is supported by National Nature Science Foundation under Grant (NO. 61004118, 61304104) and the China Scholarship Council.

Swarm Optimization (BPSO) is utilized to solve this problem; respecting to multiple-faults simultaneous diagnosis for the rotating machine, [5] proposed a hierarchical diagnosis network (HDANN) based on the theory of Artificial Neural Networks; in [6], the work status of pump are divided into five categories, and each kind of the four faults is subdivided into three kinds of health statuses respectively, and then Hidden Markov Model (HMM) is used to recognize the real work statuses from the thirteen statuses, realizing the fault and its degradation degree simultaneous diagnosis. However, the model reveals good recognition performance on faults, but bad for faults' degradation degree, which is on account of that the state dwell time probability of various faults' degradation degree obeys exponential distribution on the condition of Markov assumption, causing the dwell time for each degraded state can not be reacted in the evolution of equipment failure. In accordance with the theory of HMM, HSMM is proposed on the condition of introducing the dwell time for each degraded state. HSMM, owning rigorous mathematical structure, can make the behavioral characteristics of the entire observation sequence being expressed completely. However, it is difficult to recognize the multiple-faults by means of HSMM due to the characteristics and the maximum likelihood criterion of HSMM. Artificial Neural Network (ANN) not only having good biology basis and data base, but also owning good ability of generality and fault tolerance, which can deal with the problem of multiple-faults diagnosis, but the accuracy for the identification of the faults' degraded state is not high. Thus, it is critical to research the model of multiple-fault and degradation degree simultaneous diagnosis based on the advantages of Artificial Neural Network and Hidden Semi-Markov Model.

This paper presents a new fault diagnosis method using ANN and HSMM, realizing the multiple-fault and degradation degree simultaneous diagnosis.

II. MULTIPLE-FAULT AND DEGRADATION DEGREE SIMULTANEOUS DIAGNOSIS BASED ON ANN-HSMM

A. The Fundamental of ANN

Artificial Neural Network (ANN) [7], a mathematical model which can response the human brain structure and its functions abstractly, is a complex network interconnected by a large number of neurons, and its structure is shown as Fig 1. There are m nodes in the input layer, and the number of output equals to the input; the hidden layer i has q nodes, and f_1 denotes the activation function, ω_{ij} is the connection weights between the input layer j and the hidden layer i ; the output layer k contains l nodes, and f_2 stands for their activation function, ω_{ki} is the connection weights between the hidden layer i and the output layer k . Firstly, the hidden input, which is obtained by weighted sum for the input of the input

$$p_k = \sum_{j=1}^m \omega_{ij} y_j, \quad k = 1, 2, \dots, q.$$

Secondly, in accordance with the activation of the f_1 in hidden layer, the input of the hidden layer can be expressed as $m_k = f_1(\sum_{j=1}^m \omega_{ij} y_j - \theta_k)$, $k = 1, 2, \dots, q$. Lastly, the real input results can be got by the activation function f_2 for the output layer.

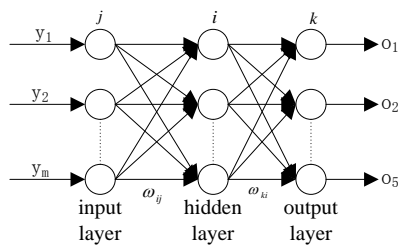


Fig.1. The topology for ANN

The model, which can not only simulate the humans' ability about presentation and storage for knowledge, but also imitate the humans' reasoning behavior for knowledge, is an information system. The objective function, which is established by comparing the actual output with the ideal output, can be used to revise the interconnection weights between the respective neurons by repeated learning, achieving the interconnection weights convergence within the stable range. The trained and learning algorithm of network, which determines the initial weights that connect the related neurons, will adjust the weights automatically with the training patterns' adding. By training the learning algorithm, the network could get the satisfied performance. The common training algorithms include the learning of error correction, competitive learning and so on. And the algorithm of DFP is adopted in this paper. The neural network could acquire the knowledge through the connection structure and steady weight distribution in the related neurons. Meanwhile, the neural network itself could filter out the noise and deal with problems in the presence of noise. Therefore the neural network is suitable for online fault detection and diagnosis.

B. Fundamental of Hidden Semi-Markov

The Hidden Markov Model [8] is a doubly stochastic process: one is a "state" process. Its nature is a Markov chain that can be used to describe the state transfer, and at the same time it is a basic stochastic process; the other is a process called observation sequence value which can describe the corresponding relations of probability and statistics between every state and its observed values. Because the observation sequence value which is due to the implicit state cannot be acquired directly, and the state existence and properties are obtained by perception through a random process. So it is called Hidden Markov model, which is short for HMM denoted by $\lambda = (N, M, \pi, A, B)$.

Where, $A = (a_{ij})_{N \times N}$ denotes state probability transition matrix and $a_{ij} = P(q_{t+1} = \theta_j | q_t = \theta_i), 1 \leq i, j \leq N$; $B = (b_{ik})_{N \times N}$ is on behalf of the observation value probability matrix and $b_{ik} = P(O_t = V_k | q_t = \theta_i), 1 \leq j \leq N, 1 \leq k \leq N$; $\pi = (\pi_1, \pi_2, \dots, \pi_N)$ stands for the initial probability distribution vector and $\pi_i = P(q_1 = \theta_i), 1 \leq i \leq N$; N presents the state number of the model, and N states are denoted by $H = (H_1, H_2, \dots, H_N)$, the state at time t is denoted by s_t ; M is the corresponding number of possible observation values for every state, and M observation values are denoted by $V = \{v_1, v_2, \dots, v_M\}$.

Probability of model staying for some time in a certain state generally obeys the following exponential distribution $P_i(d) = a_{ii}^{d-1} (1 - a_{ii})$ in HMM, so we can get to know that the probability presents exponentially decreasing trend as time increasing. But according to the actual situation, the state staying time does not obey this distribution function in most instances.

HSMM [9] is one of the models used to overcome the above HMM disadvantage. Based on the HMM, HSMM could overcome HMM modeling limitation caused by the Hidden Markov modeling assumption. In HMM, a state only corresponds to one observation value; but in HSMM, a state corresponds to a sectional observation value, it means that the state staying time is added in and the observation value is connected with both the current state and its staying time when transferring from the current state to the next. The HSMM topological structure is shown in Fig.2, in which q_{r-1} represents the start point at time r and the staying time is $d = q_r - q_{r-1}$.

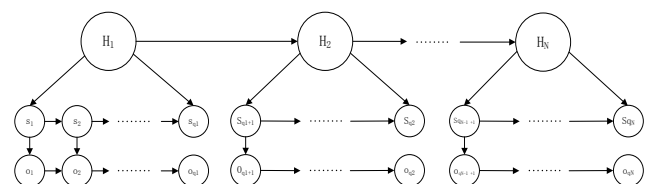


Fig.2. HSMM topological structure

HSMM properties can be described by the initial state distribution vector π , state staying time distribution D , and the probability matrix of observation value; so HSMM can be denoted with $\lambda = (N, M, \pi, A, D, B)$. In HSMM, transition

among macro states meets with the Markov process, but transition among micro states does not, but the model can be trained by according to the algorithm in [10].

C. Fault Diagnosis Model Based on ANN-HSMM Network

Fault diagnosis is divided into before and after two parts in ANN-HSMM network, equals to assign a large-scale diagnosis task to many sub-networks. The first network uses ANN model to diagnosis multiple-faults; based on the above diagnosis results, the second network adopts HSMM to analyze and diagnose the measured signals again to get the degradation degree. This model can identify multiple-faults, diagnose fault degradation degree, and improve the network learning efficiency to meet the practical online diagnosis requirements. Suppose the measured faults can be divided into 5 classes $\{F_0, F_1, F_2, F_3, F_4\}$, in which, exclude from F_0 (normal wok condition), the other 4 faults can further be divided into three different degradation degree faults $\{F_{11}, F_{12}, F_{13}\}$, $\{F_{21}, F_{22}, F_{23}\}$, $\{F_{31}, F_{32}, F_{33}\}$ and $\{F_{41}, F_{42}, F_{43}\}$. So the faults are divided into 13 states and the extracted fault eigenvector equals to $Y = \{y_1, y_2, \dots, y_m\}$.

The first layer is accomplished by the feed-forward single hidden layer neural network; number of the input and output panel points is m and 5 (number of fault kind), every output panel point relates to a certain fault, and if the output value is 1, which means the relative fault exists, otherwise no fault exists. The network training modes are shown in TABLE I.

TABLE I. The first network training mode

Training mode	Input	Output	Relative fault
0	$(y_1^0, y_2^0, \dots, y_m^0)$	0 0 0 0	F_0 (no fault)
1	$(y_1^1, y_2^1, \dots, y_m^1)$	1 1 0 0	F_1
2	$(y_1^2, y_2^2, \dots, y_m^2)$	1 0 1 0	F_2
3	$(y_1^3, y_2^3, \dots, y_m^3)$	1 0 0 1	F_3
4	...	1 0 0 0	F_4
5		1 1 1 0	$F_1 F_2$
6		1 1 0 1	$F_1 F_3$
7		1 1 0 0	$F_1 F_4$
8		1 0 1 1	$F_2 F_3$
9		1 0 1 0	$F_2 F_4$
10		1 0 0 1	$F_3 F_4$
11		1 1 1 1	$F_1 F_2 F_3$
12		1 1 1 0	$F_1 F_2 F_4$
13	...	1 1 0 1	$F_1 F_3 F_4$
14	$(y_1^{14}, y_2^{14}, \dots, y_m^{14})$	1 0 1 1	$F_2 F_3 F_4$
15	$(y_1^{15}, y_2^{15}, \dots, y_m^{15})$	1 1 1 1	$F_1 F_2 F_3 F_4$

The second layer exploit HSMM network, and output values of the first layer can stimulate the 2th layer to function, reclassify the measured signals to identify different order of fault severity, and the network structure is shown in Fig.3.

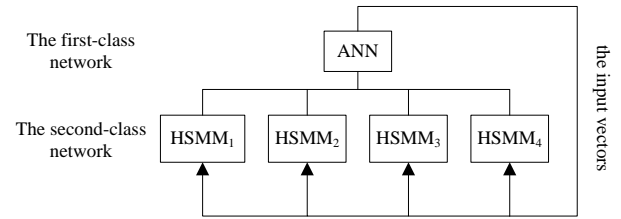


Fig.3. ANN-HSMM network structure

Every network is trained individually according to its grade level. To the trained networks, the inputs are the extracted fault characteristics vectors from the measured objects and the diagnosis should be executed step by step. First ANN network takes effect, and the second corresponding HSMM is stimulated by the results of first-class network, and the diagnostic rules are as follows:

- If the first output value of ANN is 0, it indicates no faults exist in the equipment, and it's over; otherwise turn to 2th step.
- Judge the other output of ANN in order excluded from the first output, then input all the relative serial number that is connected with the state whose output value is greater than 0.5 into empty matrix A; if all the output values of ANN are smaller than 0.5, input the serial number that is connected with the panel point whose output value is maximum into matrix A.
- Stimulate the second network in order to function according to the elements in matrix A, in the relative HSMM library which connects with the stimulated faults, the state connects with the maximum output probability is concerned to be the degree to this equipment and its corresponding fault.
- The ultimate diagnosis result is due to the output of both the first and second network. The output values of the first layer network those are greater than 0.5 or the maximum output value possibly is the fault class to the measured object. Stimulated by the output results of the first layer network, the state connected with the model that has the maximum output probability in HSMM library is the corresponding degradation degree to the related faults.

As the training is independent to the sub-networks of the related grades, so in actual, it only needs to use the practical measured fault samples to retrain the related sub-networks for different grades, and the suitable specific units can be acquired which is easier to be trained and has stronger adaptability.

III. APPLICATION OF SIMULTANEOUS DIAGNOSIS OF MULTIPLE-FAULTS AND SEVERITY OF ANN-HSMM

A. Acquirment of Training Samples

The training eigenvector is structured by the energy ratio in vibration signal band of measured equipment and used to divide the fault category to get the fault training model. For the frequency domain features of the vibration signals of the rotating machinery are acquired through the Fourier transform, and time domain signals can be represented by the superposition of multi sinusoidal signals, if a fault exists, it

means that the original time domain signals superpose one or several different sinusoidal signals which have different frequency. $f(t)$ is supposed to be the time domain signal when the unit has two single-frequency faults, as:

$$f(t) = f_1(t) + f_2(t) = \sin(2\pi f_1 t + \varphi_1) + \sin(2\pi f_2 t + \varphi_2) \quad (1)$$

According to the linear characteristic of Fourier transform, there can be:

$$F(\omega) = F(f_1(t)) + F(f_2(t)) = F_1(\omega) + F_2(\omega) \quad (2)$$

Multi-fault eigenvector can be acquired through the superposition of single-fault eigenvector, in order to protect the fault characteristic, doubly fault eigenvector is acquired by two single-fault eigenvector weighting. Suppose y_i^1, y_i^2 is the i -th relative component of single-fault eigenvector, so the doubly-fault eigenvector can be defined as follows:

$$y_i^{1,2} = \left[\frac{y_i^1}{y_i^1 + y_i^2} \right] y_i^1 + \left[\frac{y_i^2}{y_i^1 + y_i^2} \right] y_i^2 \quad (3)$$

It is the formula (3) that is a filtering process, which can not only maintain the fault characteristics, but also restrain the non-fault features. The response component of the multiple-fault characteristics' vectors can be established similarly to (3), obtaining the standard training samples of the network. Similarly, the statuses' feature vectors of the various faults' degradation degree can be got under different faults' status.

B. The Training for the Network

The feed-forward signal hidden layer network is adopted in the first-class network, and the number of input nodes is m , the output nodes is 5, the hidden layer nodes can be commutated by the formulation "the number of hidden layer nodes=2*the number of input nodes+1", that is $2m+1$. Then DPF Approach [11] is adopted to train the first network, and the learning objective function can be defined as below:

$$J(W) = \frac{1}{2} \sum_{i=1}^M E_i^2(W) = \frac{1}{2} \sum_{i=1}^M (O_j^M - d_j)^2 \quad (4)$$

Where, W denotes the weights vector of the first-class network; E_i is on behalf of the error between the real input and the ideal input. In accordance with the formula (4), the second-order polynomial Taylor around the minimum point can be regarded as the objective function approximately, obtaining the estimated value of the minimum point. After obtaining W_{k+1} , the connection weights for the neural network can be modified according to the following formula (5-8):

$$W_{k+1} = W_k - H_k E_i(k) \bullet \partial E_i(W_k) / \beta_k \quad (5)$$

$$H_k = \lambda^{-1} \left[H_k - \frac{H_k * \partial E(W_k) * \partial E^T(W_k) * H_k}{\beta_k} \right] \quad (6)$$

$$\beta_k = \lambda + \partial E^T(W_k) * H_k * \partial E(W_k) \quad (7)$$

$$H_1 = I \text{ (Identity Matrix)} \quad (8)$$

Where, $E(W_k)$ is the gradient for E plays on W_k .

As for the training about HSMMs, the second-order HSMM with continuous Gaussian density are chosen, and the Markov chain with its status varying from left side to right, which has better effects and high training speed, is selected. The Markov chain with three statuses is selected in this paper, and the initial probability is $\pi = [1, 0, 0]$. The initial value of A can be obtained by means of uniform choice and B is got randomly. The Baum-Welch algorithm is developed on the basis of climbing algorithm so that the initial value has great effects on achieving the best solutions. Therefore, the K-means algorithm and clustering algorithm are utilized to obtain the initial values and the Baum-Welch [12, 13] with several observation sequences is utilized to train the model to increase the robustness. A couple of initial parameters of HSMM are estimated by the K-means algorithm and clustering algorithm, and then the data obtained above is utilized to train the various HSMM.

C. The Test Results and the Analysis for the Network

The network test is obtained on the trained network ANN-HSMM, achieving multiple-fault and degradation degree simultaneous diagnosis. The test results are shown as below:

- The diagnosis and the classification for the large fault can be realized by means of the first-class network, and the diagnostic accuracy reaches 95%.
- Stimulated by the results of the first-class network, the second corresponding HSMM starts to work, achieving the faults' degradation degree diagnosis. Expected for the two states $\{S_{10}, S_{11}\}$, the other states can be identified correctly by the hierarchical diagnosis network (ANN-HSMM), and the identification accuracy is 90%.

For the same diagnosis problem, the diagnostic performed for the HMM and the ANN-HSMM mentioned above is compared, and the comparison results are shown as below:

TABLE II. The comparison for the recognition rate

	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
HMM	85%	75%	75%	90%	75%	70%	70%	100%	100%	85%	65%	60%	75%
ANN-HSMM	100%	95%	85%	90%	80%	85%	80%	100%	100%	90%	75%	70%	85%

As can be seen from the above table, the diagnosis and the classification for the large faults can be achieved by HMM and ANN-HSMM, but for the problem of faults' degradation degree diagnosis, the diagnosis accuracy for ANN-HSMM are higher than HMM, in other words, the multiple-fault and degradation degree simultaneous diagnosis can be obtained through the hierarchical diagnosis network (ANN-HSMM). The test results reveal that for the same diagnosis problem, regardless of the diagnosis and the classification for the large fault or the faults' degradation degree diagnosis, ANN-HSMM is superior to the HMM in conference [6], therefore the established hierarchical diagnosis network (ANN-HSMM) in this paper is reasonably practicable.

IV. CONCLUSION

Based on the theory of Artificial Neural Networks and Hidden Semi-Markov Model, a hierarchical diagnosis network (ANN-HSMM) is proposed with the respect to multiple-faults and their degradation degree simultaneous diagnosis. The experiment results show that comparing with the fault diagnosis method in [6], ANN-HSMM can not only recognize multiple-faults diagnosis, but also achieve the degradation degree diagnosis for the corresponding faults, and the diagnostic accuracy is higher, available for real-time condition monitoring and diagnosis.

REFERENCES

- [1] Juan Li, Hong-wei Gao, Peng Zhang and Da-rong Huang. Fault Diagnosis and Optimal Fault-Tolerant Control for Systems with Delayed Measurements and States [J]. International Journal of Control, Automation, and Systems. 2012, 10(1): 150-157.
- [2] Zhao Ling, Huang Darong, Song Jun. Fault diagnosis method based on fractal theory and application in wind power systems [J]. Journal of China Ordnance, 2012, 8(3): 167-173.
- [3] HUANG Da-rong, SONG Jun, ZHAO Gang. Research on Safety Performance Evaluation Method of Fault-prediction Technology Based on Misclassification Cost [J]. Acta Armamentarii, 2011, 32(10): 1292-1297.
- [4] Ge Pengyue, Huang Kaoli, Liu Xiaoqin, Lian Guangyao. Research on Dynamic Multiple Fault Diagnosis Model (DMFD) Based on Hidden semi-Markov Model [J]. Computer Measurement & Control, 2010, 18(6): 1280-1286.
- [5] He Yongyong, Zhong Binglin, Huang Ren. Multiple-Fault Simultaneous Diagnosis Based on Artificial Neural Networks for Rotating Machine [J]. JOURNAL OF SOUTHEAST UNIVERSITY, 1996, 26(5): 39-43.
- [6] YUE Xia. HMM-based Fault Diagnosis Technology under Complex Conditions [D]. Guang Zhou: South China University of Technology, 2012.
- [7] LIU Jing-yan, LI Yu-dong, YANG Xiao-bang. Application of Genetic Neural Network to Gear Fault Diagnosis [J]. 2012, 33(3): 36-39.
- [8] Hu Wei, Gao Lei, Fu Li. Research on motor fault detection method based on optimal order hidden Markov model [J]. Chinese Journal of Scientific Instrument, 2013, 34(3): 524-530.
- [9] WANG Ning, SUN Shu-dong, CAI Zhi-qiang, LI Shu-min. Method of limitation state recognition for two stage equipment based on HSMM [J]. Application Research of Computers, 2011, 12(28): 4560-4563.
- [10] YANG Zhi-bo, DONG Ming. Equipment Fault Diagnosis Using Auto-regression Hidden Markov Model [J]. JOURNAL OF SHANGHAI JIAOTONG UNIVERSITY, 2008, 42(3): 471-479.
- [11] Wang Yaonan. Research of Fast Learning Algorithm for Neural Networks Based on DPF Approach [J]. ACTA SIMULATA SYSTEMATICAL SINICA, 1997, 3: 34-39.
- [12] G. Georgoulas, M. O.Mustafa, I. P. Tsoumas, etc. Principal Component Analysis of the star-up transient and Hidden Semi-Markov Modeling for broken rotor bar fault diagnosis in asynchronous machines [J]. Experts Systems with Application, 2013, 40: 7024-7033.
- [13] Boutros, T. & Liang, M. Detection and diagnosis of bearing and cutting tool faults using hidden Markov models. Mechanical Systems and Signal Processing [J]. 2011, 25: 2102-2124.

IJERT