Proceedings of the 33rd Chinese Control Conference
July 28-30, 2014, Nanjing, China

The Study of Method about Diagnosis Prediction Based on Adaptive Filtering and HMM

CHU Xiaoyan[1], ZHAO Ling[1], HUANG Darong[1, 2]

1. Institute of Information Science and Engineering, Chongqing Jiaotong University, Chongqing, 400074, China
E-mail: xyanchu@126.com

2. Institute of Automatic Control and Complex Systems, Duisburg-Essen University, Duisburg 47057, Germany
E-mail: dahuangjs@gmail.com

Abstract: A new method of diagnosis prediction based on adaptive filtering and HMM is proposed. Firstly, the feature extraction of time-sequence about collected diagnosis information is conducted, achieving time-sequence status information, and then on the basis of the results the future device status information vectors are obtained by means of adaptive filtering. Secondly, the HMMs for all diagnosis statuses are established and be trained according to the algorithm of Baum-Welch, meanwhile HMMs is regarded as the classifier. Lastly, the obtained time-sequence about future status are inputted into HMMs, getting the prediction results in accordance with the criterion of maximum likelihood. The experimental results prove the feasibility for this method, and its prediction accuracy is improved to some extent.

Key Words: Diagnosis Prediction, HMM, Adaptive Filtering

1 Introduction

The effective fault prediction, which is based on Condition Based Maintenance (CBM), can be realized by means of detecting the status of equipment, providing reasonable suggestions for maintenance strategy, and it is an effective method for inadequate maintenance and excess maintenance. The technology has become a hot research field on equipment maintenance recently, but the effective fault prediction should be come true before the implement of CBM. Generally speaking, the occurrence of equipment fault is a multi-state process varying from a normal state to a failed state, and these states can not be detected directly, but the change of states can be reflected through detecting the signals by means of the sensors and its like. The faults feature vectors can be got by analyzing and extracting the detected signals, and then on the base of the history signals, the vectors are regarded as the input of the established mathematical model, realizing the fault prediction. The reference [1] deals with the problems of fault diagnosis and fault-tolerant control for systems with delayed measurements and states; in the reference [2], the new fault diagnosis method based on fractal theory is proposed and applied to a wind system and the results show the method's feasibility; in reference [3], a new safety performance evaluation of the fault-prediction technology was proposed on misclassification cost, and the further development trend of the fault prediction was discussed. However, the Hidden Markov Model was also used in the field of fault prediction and the fault prediction frame [4] was proposed at the same time. The hypothesis of HMM makes the future states related with the current states, regardless of the past states. In fact, the observation probabilities at any time not only depend on the current states, but also rely on the past states, that is to say, the assuming above is unreasonable. The probability values are obtained by the fault feature vectors being input into the HMM classifier, and under the same conditions, the probability values are greatly effected by the fault feature vectors. In other words, the future states are bound to be effect by not only current state feature vectors but the past.

Based on these, the method about Diagnosis Prediction Based on Adaptive Filtering and HMM is proposed in this paper. After the process for the historic data, the future state feature vectors can be obtained by means of the algorithm of adaptive filtering, and then the future state of equipment can be identified by regarding the vectors got above as the input of HMM classifier.

2 The Fault Prediction Model Based on the Adaptive Filtering and HMM

2.1 Hidden Markov Model

The HMM [5], which is developed on the basis of the Markov chain, is a probability statistics model. Its recognition performance is very good so that it is widely used in the field of fault diagnosis and state detection. However, the complexity of the real problem is higher than the complexity describing by Markov model, and the events observed in HMM is not one to one to the status in device, but the existence or absence of device status is perceived through the stochastic processes. If the status is presented, the characteristics should be perceived in this process and the device statues cannot be observed, so the model is call Hidden Markov Model, which is abbreviated as HMM. The most is that the statistical learning and probabilistic reasoning can be completed by the HMM independently,

[*] This work is supported by National Nature Science Foundation under Grant 61004118, the National Nature Science Foundation under Grant 61304104 and the China Scholarship Conucil.

and the precise knowledge is not required in the process of solving problems.

The Hidden Markov Model is a doubly stochastic process, and one is used to describe the state transition in Markov chain, and the other is a process depicting the statistical correspondence relationship between the device status and the observation. Using the statistical probability model HMM to describe the observation sequence not only possesses rigorous mathematical structure, but also expresses the complete behavioral characteristics of the overall observation sequence.

We describe the complete set of a given model by: $\lambda = (N, M, \pi, A, B)$. Where, π is the initial state probability distribution vector, A stands for state transition probability matrix, B denotes a set of observation probability functions, N represents the number of the state Markov chain model, and M is the number of observations corresponding to every state.

There are three basic algorithms [6, 7] are included in the HMM: firstly, we use the forward (or the backward) procedure to find $p(O \mid \lambda)$ for some $O = (o_1, o_2, \cdots, o_T)$; secondly, given some O and some λ, the Viterbi algorithm is utilized to find the best state sequence $q = (q_1, q_2, \cdots, q_T)$ that explains O ; thirdly, The Baum-Welch (also called forward or EM for HMMs) algorithm is used to find $\lambda^* = \arg\max_{\lambda} p(O \mid \lambda)$. The training and decision-making of HMM can be solved on these three algorithms.

2.2 Adaptive Filtering

2.2.1 The Basic Process of Adaptive Filtering

On acknowledging the adaptive filtering theory [8], the future state of the device can be forecasted by getting the weighted value of the historical observations of the time series. The critical problem is to search a set of best weights, and the specific approach is shown as below: firstly, a predicted value and prediction error can be obtained by means of a couple of given weights; secondly, according to the prediction error, the weights are adjusted to reduce the error, and repeated operation is made until the prediction error is reduced to a minimum through a set of weight, which is called the best weights. Due to its similarity with the process noise transmission in field of communication engineering, this method is called adaptive filtering.

The formula of adaptive filtering can be written as below:

$$\hat{o}_{T+1} = w_1 o_T + w_2 o_{T-1} + \cdots + w_N o_{T-N+1} \tag{1}$$

Where, $\hat{o}_{T+1}$ denotes the eigenvectors at $T+1$; w_i is the state feature weights at $T-i+1$; o_{T-i+1} stands for state eigenvector at $T-i+1$; N indicates the number of the weights. And the adjustment formula for weights can be expressed as following:

$$w_i' = w_i + 2k.e_{i+1} o_{T-i+1} \tag{2}$$

Where, $i = 1, 2, \cdots, N, T = N, N+1, \cdots, n$. n is the number of vectors, w_i is the i-th weight before adjustment, w_i' is the i-th weight after adjustment, k denotes the learning constant, and e_{i+1} stands for the prediction error vectors. The formula (2) indicates that the adjusted weights in i-th times can be expressed by pulsing the old weights and the error adjustments, and the error adjustments includes prediction error vectors, the original status observation vectors and the learning constant. The learning constant k is determined by the adjusted speed of weights.

2.2.2 The Determination of N, k and Initial Weights

The learning constant k and the weights number N should be determined firstly when the weights are adjusted, and the method of autocorrelation coefficient is utilized to determine the value of N . Generally, lets $k = \frac{1}{N}$, and the various value of k can also be adopted to determine a better k , which can help to obtain better results. It is worth to mention that it is critical to determine the initial weights, for simplicity w_i can be got by the following formula:

$$w_i = \frac{1}{N} \tag{3}$$

The method, whose calculation is simple, can not only select the weights and the learning constant based on the predictive purposes to realize control prediction, but also all historic vectors are used during the select of best weights, and the weights are updated constantly associated with the change of track data, realizing the prediction improved.

2.3 The Model About Diagnosis Prediction Based on Adaptive Filtering and HMM

The result got from HMM classifier is determined by the maximum output in HMMs on the basis of Markov hypothesis, but the observation probability at any time not only dependent on the current status, but also rely on the past status over a period of time, that is to say, the suppose above is unreasonable. The probability can be got by putting the status feature vectors into HMM classifier, and the probability can be affected by status feature vectors greatly under the same condition. In other words, the future status feature vectors are bound to be affected by the current status and the past status at the same time. Therefore, the status information can not be made full use of in the original model on suppose above, which lead the forecast risk exist in the process of prediction. The approach based on adaptive filtering can only predict the future status information on the basis of the historic observation, and the classification for status cannot be achieved.

HMM can realize classification of status but does not make full use of the historic status information, on the contrary, the historic status information can be made full use of in the method of adaptive filtering, but the classification can not be achieved. Based on these facts, the adaptive filtering method can be utilized to make up the unreasonable suppose, and the matching degree between various HMM and the predictive signal can be computed by means of the characteristics [9, 12], which comes from HMM processing the continuous dynamic signal, and in all results, the status having maximum input is the possible status in the future. Based on the discussion above, the model about diagnosis prediction based on adaptive filtering and HMM is proposed.

The prediction based on the theory of adaptive filtering and HMM can be classified as three steps: data acquisition, prediction model training and the faults prediction. As shown in Fig 1, firstly, the future status feature vectors are obtained by processing the historic data on the theory of adaptive filtering; secondly, the faults characteristic vectors are got by extracting the training samples, and then establishing various HMM corresponding to different fault status, and the Baum-Welch algorithm [10] is utilized to train the established HMM, obtaining the HMMs; thirdly, on the algorithm of Viterbi, the various probability corresponding to every status can be produced by putting the vectors got in the first step into the trained HMMs, and the status, corresponding to the model possessing maximum output probability, is the predictive result.

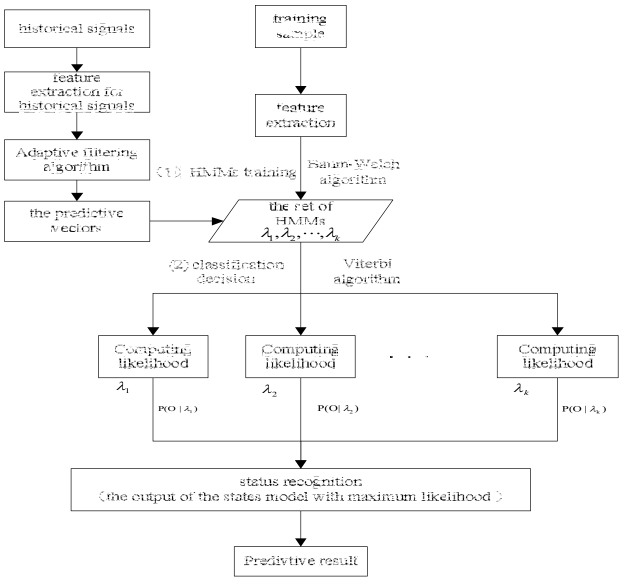

Fig. 1: The predictive diagnosis model based on adaptive filtering and HMM

3 The Forecast Example

3.1 The Prediction of Status Characteristic Vectors

The common fault eigenvectors at various times, as shown in table 1, are obtained on the basis of the field experience and the research on fault mechanism, and then the corresponding characteristic vector at $t+1$ can be got by means of MATLAB.

Table 1: The time sequence characteristic vectors

time	The historical characteristic vector o_t
…	…
t-3	$(0, 0.619, 0.785, 0.005, 0.008, 0.032, 0)^T$
t-2	$(0.106, 0.506, 0.425, 0.649, 0.070, 0.354, 0)^T$
t-1	$(0.360, 0.258, 0.063, 0.844, 0.049, 0.292, 0)^T$
t	$(0.070, 0.279, 0.869, 0.353, 0.001, 0.191, 0)^T$

There are four observations in the table 1. Selecting two weights to calculate the predictive eigenvector at $t+1$, and then the initial weight w_i is got in accordance with the formula (3) on basis of $N=2$. Given $k=0.9$ and the value of t vary from 2. When $T=t-2$ the predictive vector can be got as the formula (1): $\hat{o}_{T+1} = \hat{o}_{t-1} = w_{t-3}o_{t-3} + w_{t-2}o_{t-2}$, and the predictive error can be shown as $e_{T+1} = e_{t-1} \quad = o_{t-1} - \hat{o}_{t-1}$. The adjust weights can be calculated according to the formula (2): $w_1^{'} = w_1 + k.e_{t-1}.o_{t-2}$, and $w_2^{'} = w_2 + k.e_{t-1}.o_{t-2}$. At this time one adjustment for weights are completed, and then the repeated operation above is conducted. At the end of the first adjustment, if the predictive error meets the given accuracy and the changes in weights are very small, a set of best weights are obtained; otherwise the above operation is repeated on basis of $T=t-2$ until that the prediction error has no significant improvement, and the weights is regarded as the best weights at this time, which can be used to predict the status eigenvector at $t+1$. However, due to the large workload of the adjustment for weights, the MATLAB must be utilized in the process of calculation, and the weights $w_1^{'} = 0.534$, $w_2^{'} = -0.595$ are got by means of the MATLAB. Thus, in accordance with the formula (1), the eigenvector at $t+1$ can be got as $o_{t+1} = (0.151, 0.378, 0.248, 0.241, 0.026, 0.042, 0)^T$.

3.2 The Training for HMMs

Compared with the discrete HMM, the continuous HMM [9] has some characteristics, such as obvious classification effects, small predictive error and so on. In order to keep consistencies with the adaptive filtering, the second-order HMM [10] 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. When the size of sample is small, the reliable model can be obtained by selecting the smaller values. The Markov chain with twelve statues is selected in this paper, and the initial probability is $\pi = [1,0,0,0,0,0,0,0,0,0,0,0]$. 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 [11] are utilized to obtain the initial values and the Baum-Welch [7, 13] with several observation sequences is utilized to train the predictive model based on HMM to increase the robustness.

A couple of initial parameters of HMM are estimated by the K-means algorithm and clustering algorithm, and then ten groups of data are utilized to train the various HMM, such as the normal status HMM_1 (S_1), imbalance normal status HMM_2 (S_2), misalignment status HMM_3 (S_3), friction status HMM_4 (S_4), bearing loose status HMM_5 (S_5), cabinet pedestal looseness HMM_6 (S_6), instability status HMM_7 (S_7), subharmonic resonance status HMM_8 (S_8), shaft transverse cracks status HMM_9 (S_9), gap vibration status HMM_{10} (S_{10}), pressure pulsation status HMM_{11} (S_{11}) and oil whirl status HMM_{12} (S_{12}).

3.3 The Results of the Prediction and its Analysis

The MATLAB program is written according to the flowchart of algorithm as Fig 1, and the predictive eigenvector in 3.1 is input into the trained HMMs in 3.2 and the results are shown in table 2.

Table 2: The results of HMM classifier

Status (S)	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
Probability (P)	0.08	0.37	0.77	0.83	0.05	0.87	0.21	0.15	0.40	0.67	0.07	0.13

Associated with the results in table 2 and the principle of maximum like-hood, the HMM of cabinet pedestal looseness will likely occur in the next time, and the friction, misalignment and gap vibration also occur at the same time. And the most possible status in the next time is one or more of the four statuses mentioned above. Cabinet pedestal looseness will mostly occur in the fans, and the occurrence of this fault may lead to the results that the vibration signals are affected, which causes the other faults at the same time; misalignment leads to eccentric, which may give rise to friction to a certain extent, and eccentric may also be aroused by gap vibration. However, the value of the HMM order is selected according to experience and its selected method has no theoretical basis to support. Moreover, due to the maximum likelihood criterion, some risk will exist in this faults prediction method, but the risk is within the acceptable range and the predictive accuracy is improved to some extent.

4 Conclusion

The diagnosis prediction model based on adaptive filtering and HMM in proposed in this paper on basis of analyzing the contradiction between HMM suppose and the reality. Firstly, the future status vectors are obtained through processing the historical signal information by means of adaptive filtering. Secondly, the trained HMMs are regarded as the classifier, and the predicted vectors are put into it so that the fault can be obtained in accordance with the maximum likelihood criterion. This method can not only maintain the advantages of HMM, but also make up its shortcomings. The model is utilized in the faults prediction experiment of fans. The experimental results prove the feasibility and effectiveness for this method, and its prediction accuracy is improved to some extent, meanwhile the predictive error risk is within the acceptable range.

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] Peng, Y., & Dong, M. A hybrid approach of hmm and gray model for age-dependent health prediction of engineering assets [J]. Expert Systems with Application, 2011, 38(10): 12946-12953.

[5] ZHU Ming, GUO Chun-Sheng. Hidden Markov Model and Its Application and Progress [J]. Computer System & Application, 2010, 19(7): 255-259.

[6] YUE Xia, ZHANG Chun-liang, QUAN Yan-ming, ZHU Hou-yao. Fault Diagnosis of Nuclear-Powered Equipment Based on HMM & SVM [J]. Nuclear Power Engineering, 2012, 33(3): 104-108.

[7] 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.

[8] Li Xian-xiao. The Basis of Time Series Analysis [M]. Wu-Han: Huazhong University Press, 1991.

[9] FENG Chang-jian, SHAO Qiang, KANG Jing. Fault Diagnosis for Baring Based on HMMs with Gaussian Mixtures [J]. TURBINE TECHNOLOGUY, 2011, 53(3): 205-208.

[10] 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.

[11] Lawrance R Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition [J]. Proceedings of the IEEE, 1989, 77(2): 257-286.

[12] YUE Xia. HMM-based Fault Diagnosis Technology under Complex Conditions [D]. Guang Zhou: South China University of Technology, 2012.

[13] G. Georgoulas, M. O.Mustafa, I. P. Tsoumas, etc. Principal Component Analysis of the star-up transient and Hidden Markov Modeling for broken rotor bar fault diagnosis in asynchronous machines [J]. Experts Systems with Application, 2013, 40: 7024-7033.