Design of Fault Detection/Diagnosis Model for Thermal Storage System

Song Pan
Research Laboratory
Sanko Air Conditioning Co., LTD.
Nagoya, Japan
han@sanko-air.co.jp

Mingjie Zheng
Research Laboratory
Sanko Air Conditioning Co., LTD.
Nagoya, Japan
zheng@sanko-air.co.jp

Nobuo Nakahara
Nakahara Laboratory
Environmental Sys-Tech
Nagoya, Japan
pl.nobuo.naka@kfx.biglobe.ne.jp

Abstract: The authors simulated the human ability of pattern recognition mathematically, through finding the state-change characteristics of the objective system from actual measurements using statistical analytical techniques. The Fourier analysis method with the space invariance that can extract the changing characteristics of the system state, such as phase and frequency, is chosen as a typical technique. The distinction function based on the system state such as phase and frequency, is chosen as a typical technique. The Maharanobis’ pan-distance is used to identify and judge the essence of the event. In addition, human learning, recognition, and optimal judgment process of any event can be simulated by optimizing the most effective parameters and their numbers for detection and diagnosis by the use of variable selection method.

In previous papers by authors [2], the two optimization methods for the most effective detection and diagnosis vector, the variable selection method and the differentiation rate increment method, in which linear distinction function based on the Maharanobis’ pan-distance was used, have been reported. In the present paper, a new method for optimal model selection based on AIC (Akaike Information Criteria) is examined. In addition, by examining the influence to the detection and diagnosis rate of the optimal detection and diagnosis vector, the best convergence criteria of AIC was confirmed.

Key words: optimization, AIC, fault detection, fault diagnosis, thermal storage system

1. INTRODUCTION

As the development of applied calculation technology, studies on reproducing human recognizing ability to outside world phenomena with computer are getting more and more popular. In this study, through finding the system state changing characteristics from actual measurement using statistical analytical technique, authors tried mathematically to reproduce human pattern recognition ability. In general, human pattern recognition ability is divided roughly into two processes, a) extracting the necessary information from outside event, and b) identifying and judging the essence of the event. Although the first process could be executed by specialist with abundant experience, in many cases, by the reason of localization of specialists’ knowledge, it is necessary to extract the feature of the object system by statistical analyses of the measurement data. In this case, it is necessary to devise the observation mechanism cleverly. The Fourier analysis method with the space invariance that can extract the changing characteristics of the system state such as phase and frequency, is taken up as a proper technique. However, in general, because the noise with a higher-order frequency is included in the measurement data, it is necessary to delete the noise by acting the low-frequency pass type filter to the Fourier analysis value. As the second process, for parameters made from measured data have mutual correlation generally, it is necessary to use the distinction function based on the Maharanobis pan-distance. In addition, human learning, recognition and optimal judgment process to an event could be simulated by optimizing the most effective parameters and their numbers for detection and diagnosis by the use of variable selection method.

Concerning the subject a) above-mentioned, Fourier analysis was executed measured temperatures in the thermal storage system tank in the past paper [1]. Acting the similar low region street type filter to the Fourier value, twenty-four parameters available for extracting the feature of the object system have been designed using statistical technique. Concerning the subject b) above-mentioned, linear distinction functions based on the Maharanobis’ pan-distance, an optimizing method comprised of the most effective fault detection and diagnosis (to be called FDD, hereafter) vector using variable selection method and detection rate increment method were reported in the past paper [1]. In this paper, an optimum model using Akaike’s Information Criterion (to be called as AIC, hereafter) has newly been applied and the best convergence judgment standard was selected by examining the influence on the detection and diagnosis rate.

2. MODEL CONSTRUCTION WITH AIC

Model construction by amount of AIC change is a
technique by which the system model is constructed and significant parameters are identified based on the statistical feature of measurement data. This technique was applied to optimizing fault detection and diagnosis model in this paper. In general, it can be guessed that the more the number of parameters is, the higher the distinction capability of a system state is, because the square of Mahalanobis distance $D_q^2$ among the distribution centers of various system states broadens. As in a thermal storage system presents discussed there are only two states, normal state and the faulty state of the valve at the inlet side of the chiller, as shown in Fig. 1, $D_q^2$ is shown by equation (1).

$$D_q^2 = (\mu^1 - \mu^2)^T \sum^{-1} (\mu^1 - \mu^2)$$

where $\mu^1 = [\mu_1^1, \mu_2^1, \mu_3^1, \ldots, \mu_q^1]^T$, $\mu^2 = [\mu_1^2, \mu_2^2, \mu_3^2, \ldots, \mu_q^2]^T$, and $\sum = \begin{bmatrix}
\sigma_1^2 & \sigma_1 \sigma_2 & \ldots & \sigma_1 \sigma_q \\
\sigma_2 \sigma_1 & \sigma_2^2 & \ldots & \sigma_2 \sigma_q \\
\ldots & \ldots & \ldots & \ldots \\
\sigma_q \sigma_1 & \sigma_q \sigma_2 & \ldots & \sigma_q^2
\end{bmatrix}$.

Even if the distinction rate of identified detection and diagnosis model rises with more parameters than necessary, however, it may be only effective for the data used. Therefore, it is important to select an appropriate judgment standard of convergence about optimized calculation to select optimal detection/diagnosis model that includes truly effective and noiseless information for distinction. Parameters calculated from the measured data have formation useful for detection/diagnosis and harmful noise at the same time. Moreover, a part of useful information for detection/diagnosis that is explained with a parameter may have been already explained with other parameters, because parameters are correlated with each other. Therefore, when q pieces of parameters are already selected, it is necessary to judge whether to add newly more r pieces of parameters or not by comparing added advantage and disadvantage, that is, truly useful information for detection/diagnosis and harmful noises. When Mahalanobis’s pan-distances between the distribution centers of normal state and fault state for q parameters and q+r parameters are assumed $D_q^2$ and $D_{q+r}^2$, respectively; it is reported that $AIC$ can judge better whether any increased distance due to added r pieces of parameters is just an error or contributes effective detection/diagnosis, as AIC is thought to be a more convenient index of judgement for model availability. The $AIC$ is calculated with a function consisted of maximum likelihood and the number of free parameters as shown in the equation (2).

$$AIC = -2 \times \log(L^*) + 2 \times q$$

Assuming that parameter vector of each state $X^j$ follows regular distribution $N_p (\mu^j, \Sigma)$ and that each parameter is mutually independent, the joint likelihood $L$ of all the measurement data could be calculated with formula (3).

$$L = 2\pi^{-n/2} |S|^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} (x_j - \mu^j)' \Sigma^{-1} (x_j - \mu^j) \right\}$$

When there are 50 or more parameters, using Least Squares Method, $\mu_1$, $\mu_2$, and $\sum$ in formula (3) can be presumed as vector of the mean values of each state sample data $x^1$, $x^2$ and covariance matrix $S$ respectively. By substituting them into formula (3), the maximum likelihood of the model can be expressed as equation (4).

$$L^* = 2\pi^{-n/2} |S|^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} (x_j - \bar{x})' S^{-1} (x_j - \bar{x}) \right\}$$

When $AIC$ values for q pieces of parameters and for q+1 pieces of parameters are assumed as $AIC_q$ and $AIC_{q+1}$, respectively, at parameter incremental method for optimization of FDD model, changed value of AIC, $\Delta AIC_{q+1}$, is induced as equation (5). 

$$\Delta AIC_{q+1} = AIC_{q+1} - AIC_q = -\log \left\{ 1 + \frac{F_q}{N_q} \right\} + 2$$

$$F_q = \frac{N - q - r - 1}{r} \left( \frac{D_{q+1}^2 - D_q^2}{M(N-2)} + \frac{D_q^2}{N_q N_2} \right) = N_q + N_2$$

If $\Delta AIC_{q+1}$ is negative, it means that added new parameter has brought in more effective information than noise for FDD. Then the parameter newly added shall be taken in the model, because it will advantageously contribute to recognize the system state. If $\Delta AIC_{q+1}$ is positive, it means that added parameter has brought in more noise than effective information for FDD.

When judgment standard $\Delta AIC$ is assumed as almost zero value, it means that the chosen model will have a characteristics too complex to identify. Therefore in general, it was pointed out that -1.0 of $\Delta AIC$ change should be considered as the border of judgment of significance for adding a new parameter.
3. CALCULATION EXAMPLE

3.1 Outline of the system and data measured

Data for analysis is taken from measured temperatures in the thermal storage tank of T Hospital HVAC system shown as in Fig. 1 from May, 1995, to September, 1996. In the early stage of measurements, a fault phenomenon was found occurred from May 1 to July 21, 1995, which was caused by improper operation of the three-way valve V1, where the supply water to the chiller was drawn only from the highest temperature part of the tank in cooling mode. The fault was restored on July 22, and the system kept normal operation, since. From the measured data it was recognized that almost no chilled water was used between October 28, 1995 and April 24, 1996, so that the data during this period are useless. As a result, the 65 days’ data between May 1 and July 21, 1995, and the 91 days’ data between July 22 and October 27, 1995 were assumed as faulty and normal, respectively, to generate FDD model. The 165 days’ data between April 25 and September 30 were assumed as normal data for diagnosis.

\[ F_{kl} = \frac{C_i C_i S_{kl}}{S_{kl}} \]

**Fig. 1 Simplified thermal storage HVAC system diagram of T Hospital**

3.2 Preparing preliminary FDD parameters

First of all, two dimensional Fourier analysis as shown in equation (6) was worked out on data for FDD model making through extracting system state features. Next, 24 preliminary FDD parameters were created after the process for freeing from noise by applying similar low region passage filter on the temperature profile in the thermal storage tank and the Fourier analysis results. In the following expressions, i and j are time and the number of tanks respectively, and k and l are the number of Fourier conversion values along the time and space, respectively.

- \[ P_1 = \text{Min}_{j=1, M} \left( \sum_{i=1}^{N} X_{ij} / N \right) \]
- \[ P_2 = \sum_{j=1}^{M} \left( X_{\text{max}, i} - X_{\text{min}, i} \right) / M \]
- \[ X_{\text{max}, i} = \text{Max}(X_{ij}), X_{\text{min}, i} = \text{Min}(X_{ij}) \]
- \[ P_3 = \left( \sum_{i=1}^{N} \left( \sum_{j=1}^{M} X_{ij} / M \right) \right) / N - \left( \sum_{i=1}^{N} \left( \sum_{j=1}^{M} X_{ij} / MN \right) \right)^{1/2} \]
- \[ P_4 = \text{Max}_{i=1, N} \left( \sum_{j=1}^{M} X_{ij} / M \right) \]
- \[ P_5 = \text{Min}_{i=1, N} \left( \sum_{j=1}^{M} X_{ij} / M \right) \]
- \[ P_6 = \text{Max}_{j=1, M} \left( \sum_{i=1}^{N} X_{ij} / M \right) - \text{Min}_{j=1, M} \left( \sum_{i=1}^{N} X_{ij} / M \right) \]
- \[ P_7 = \left( \sum_{i=1}^{N} \left( X_{\text{max}, i} - \text{Min}(X_{ij}) \right) \right) / N - \left( \sum_{j=1}^{M} X_{\text{max}, i} / N \right)^{1/2} \]
- \[ P_8 = \text{Max}_{j=1, M} \left( \text{Max}(X_{ij}) - \text{Min}_{j=1, M}(X_{ij}) \right) \]
- \[ P_9 = \text{Min}_{i=1, N} \left( \text{Max}(X_{ij}) - \text{Min}_{j=1, M}(X_{ij}) \right) \]
- \[ P_{10} = \text{Max}_{j=1, M} \left[ \text{Max}(X_{ij}) - \text{Min}_{j=1, M}(X_{ij}) \right] - \text{Min}_{i=1, N} \left[ \text{Max}_{j=1, M}(X_{ij}) - \text{Min}_{j=1, M}(X_{ij}) \right] \]
- \[ P_{11} = \sum_{k=1}^{N} \sum_{l=1}^{M} F_{kl} \]
- \[ P_{12} = \sum_{k=1}^{N} \sum_{l=1}^{M} F_{kl}^2 \]
- \[ P_{13} = \text{Max}(F_{kl}) \quad k = 1, M / 2 \quad l = 1, M / 2 \]
- \[ P_{14} = \text{Max}(F_{kl}) \quad k = N / 2, N \quad l = 1, M / 2 \]
assumed as -1.0, if P8 is added to two parameters (P1, P17), \( \Delta AIC_{q+1} \) is calculated as -0.43, that is larger than the standard -1.0, so that adding P8 is judged as useless to increase optimality for explaining state of the system, then the optimal FDD model is decided to be (P1, P17). On the other hand, when the judgment standard is assumed to be 0.0, the optimal FDD model becomes (P1, P2, P10, P17).

3.4 Effects of judgment standard of significance on distinction and diagnosis rate

In general, if discriminating function composed by FDD vector follows regular distribution, square of Maharanobis' pan-distance from measured data of a state to the center of distribution follows the distribution \( \chi^2(q) \). Accordingly, significance of discriminating analysis using Maharanobis' pan-distance can be tested with \( \chi^2(q) \). In a word, if Maharanobis' pan-distance \( D_q^2 \) is larger than \( \chi^2_{(q)} \) as shown in equation (31), the judgement result due to Maharanobis' pan-distance is not so reliable. The severer the judgment standard of \( \chi^2(q) \) is, the larger the range of discrimination becomes. In the present case, in order to enable to discriminate 99% of data, the judgment standard was assumed as \( \chi^2_{0.01} \).

\[
D_q^2 > \chi^2_{0.01}(q) \tag{31}
\]

Tab2 shows how the FDD model derived from 1995 data as for learning could discriminate the state for 1996 test data as well as for 1995 learning data, according to the significance judgment standard of two AIC variations. As a result, the following are understood.

1)When the judgment standard of the significance of the AIC variation is assumed as -1.0, the optimal FDD model is chosen as (P1,P17). Compared with the case using the model selected by the judgment standard of zero, the diagnosis rate became the best result, though the distinction rate becomes a little lower to the contrary. In a word, -1.0 of the judgement standard is the best for optimization of FDD model using AIC change and avoid excessive-learning of the 1995 learning data.

2)The data far enough from the distribution centers remains un-distinguished, even if the best setting of judgment standard mentioned above is used. It is reasonable to suppose that some other faults might have taken place during the data period in this case.

### Tab.1 Selection process of optimal model with AIC

<table>
<thead>
<tr>
<th>Parameters</th>
<th>-1.0</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_q )</td>
<td>82.6</td>
<td>6.3</td>
</tr>
<tr>
<td>( AIC_{-1} )</td>
<td>6.5</td>
<td>-4.3</td>
</tr>
<tr>
<td>( AIC_{0} )</td>
<td>-0.43</td>
<td>1.95</td>
</tr>
<tr>
<td>( AIC_{0.01} )</td>
<td>-0.36</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Proceedings of the Sixth International Conference for Enhanced Building Operations, Shenzhen, China, November 6 - 9, 2006
### 3.5 Physical meaning of composing parameters of the optimal FDD model

Therefore, parameter combination \((P_1, P_17)\) was selected as the optimal FDD model. Fig.2 shows the FDD distribution using them. The physical meaning of the two parameters is considered as follows.

1. **Physical meaning of \(P_1\)**

   Fig.3 and Fig.4 are temperature profile in the tank and Fourier transform chart in normal and fault state respectively. Fig.5 shows how to read the Fourier transform chart for the temperature profile of the multi-connected complete mixing tanks. From these figures, it is understood the temperature fluctuation in the low temperature side tank is small when three-way valve V1 operates normally. On the other hand, when three-way valve V1 is abnormal, the temperature in the low temperature side tank rise largely and the temperature slopes gently along space axis. According to equation (7), \(P_1\) is the average temperature in the low temperature tank. It is understood the variability of the low temperature tank of normal and fault state above-mentioned can be extracted with \(P_1\).

2. **Physical meaning of \(P_{17}\)**

   According to expression (23), \(P_{17}\) is the average value of maximums in each component of Fourier transformation and can express effective element size of low frequency of temperature variation along time and space axis. The maximum value of the \(S_S\) element is investigated as an example. In the normal state as shown in Fig.3, as the temperature variation frequency along time axis is large, high frequency element in \(S_S\) is large and the maximum value of \(S_S\) becomes small. On the other hand, when three-way valve V1 is abnormal as shown in Fig.4, as the temperature variation frequency along time axis is small, high frequency element in \(S_S\) is small and the maximum value of \(S_S\) becomes large. As the same tendency exists in the other elements of Fourier transform, \(P_{17}\) of fault state is larger than that of normal state as shown in Fig.2. Therefore, \(P_{17}\) is able to identify the difference of temperature variation along time and space axis between the normal and fault states.

### 4. CONCLUSION

In this report, the FDD model was optimized by introducing the computational algorithm of the AIC technique. The knowledge of the FDD of the object system is automatically acquired, and this technique can be said to emulate human process of learning, recognizing and judging optimality by automatically choosing effective parameters for FDD. Some findings obtained from the present research are shown as follows.

1. When the judgment standard of the significance of the \(\text{AIC}^\ast\) variation is assumed as zero, identified FDD model become higher dimensions, and have the tendency of excessive-learning of the learning data. Therefore, the judgment standard is recommended to be -1.0.

2. If there are data located far enough from the distribution center, the reliability of discrimination for such data by Maharanovis’ pan-distance is less. It is necessary in that case to detect such data with the \(\chi^2\) test of significance.

3. It can be said that the distinction of normal state and faulty state of the objective thermal storage system is best performed by the FDD model using \((P_1, P_{17})\) parameters, because the diagnosis rate is the highest, though the distinction rate is a little lower than using \((P_1, P_2, P_{10}, P_{17})\) parameters.
Fig. 3 Temperature profile and Fourier transformation in normal state (1995/7/26)

Fig. 4 Temperature profile and Fourier transformation in fault state (95/7/15)

Fig. 5 View method of Fourier Transformation of temperature profile

ACKNOWLEDGEMENT

The authors acknowledge Mr. Miyasaka, then Yamatake Building Systems Co. Ltd., for his efforts in preparing measurement data of thermal storage tank, and Prof. Sagara, then Mie University, for his variable advice and FDD Working Group members of the then Thermal Storage Optimization Committee in SHASE, Society of Heating, Air-conditioning and Sanitary Engineers of Japan.

NOMENCLATURE

AIC: Akaike’s Information Criterion
\( D_q^2 \): square of Mahalanobis distance between the distribution centers of normal and fault states
\( \Sigma \): covariance matrix of parameters
\( X'(q) \): test statistic, \( q \): number of free parameter
\( \mu_q \): vector of average value of \( q \) parameters in fault state
\( \mu' \): vector of average value of \( q \) parameters in normal state
\( \sigma \): dispersion of average value of \( q \) parameters in fault and in normal state
\( L \): likelihood, \( \hat{L} \): maximum likelihood
\( \log(\hat{L}) \): maximum logarithmic likelihood
\( \frac{1}{2} \chi^2 \): vector of the average of fault state sample data
\( \frac{1}{2} \chi^2 \): vector of the average of normal state sample data
\( S \): covariance matrix of \( \frac{1}{2} \chi^2 \) and \( \frac{1}{2} \chi^2 \)
\( N_1 \): number of normal data, \( N_2 \): number of fault data
\( N \): number of measure time
\( M \): number of thermal storage tanks
\( X_{rs} \): temperature of tanks in two-dimensional real cyclical function of \( s \) and \( r \)
\( s \): the number of tank, \( r \): time
\( c,s \): cosine and sine part of \( X_{rs} \) respectively
\( F_{kl} \): Fourier transform of \( X_{rs} \) with \( k \) and \( l \) for the number of Fourier transform along time and space
\( k,l \): number of Fourier transform along time and space respectively
\( F_1 \): Minimum value of daily average temperature in each tank
\( F_2 \): Average value of daily maximum temperature difference in each tank
\( F_3 \): Variance of average temperature of all tanks over all time
\( F_4 \): Maximum value of average temperature of all tanks over all time
\( F_5 \): Minimum value of average temperature of all tanks over all time
\( F_6 \): Difference between maximum and minimum value of average temperature of all tanks over all time
\( F_7 \): Variance of maximum temperature differences of all tanks at each time
\( F_8 \): Maximum value of maximum temperature differences of all tanks at each time
\( F_9 \): Minimum value of maximum temperature differences of all tanks at each time
\( F_{10} \): Difference between maximum and minimum values of maximum temperature differences of all tanks at each time
\( F_{11} \): Sum of Fourier transformation
\( F_{12} \): Norm \( \| F \|_2 \) of Fourier transformation
\( F_{13} \): Maximum value in \( C_rC_l \) components of Fourier transformation
\( F_{14} \): Maximum value in \( C_rS_l \) components of Fourier transformation
\( F_{15} \): Maximum value in \( S_rC_l \) components of Fourier transformation
\( F_{16} \): Maximum value in \( S_rS_l \) components of Fourier transformation
\( F_{17} \): Average value of maximums in each component of Fourier transformation
\( F_{18} \): Sum of average value of cycles which exceed threshold value \( b \)
$P_9$: Norm $\|P\|^2$ of average value of cycles which exceed threshold value $b$

$P_{20}$: Frequency for the magnitude of cycles to exceed threshold value $b$

$P_{21}$: Maximum value of cycles varied along time axis which exceed threshold value $b$

$P_{22}$: Maximum value of cycles varied along space axis which exceed threshold value $b$

$P_{23}$: Thermal quantity of storage tank

$P_{24}$: Thermal quantity of heat dissipated

REFERENCES


