ANOMALY DETECTION BY NEURAL NETWORK MODELS AND STATISTICAL TIME SERIES ANALYSIS

R. KOZMA, M. KITAMURA, M. SAKUMA, Y. YOKOYAMA

Department of Nuclear Engineering, Tohoku University
Aramaki-Aza, Aoba-ku, Sendai 980, Japan

Abstract

The problem of detecting weak anomalies in temporal signals is addressed. The performance of statistical methods utilizing the evaluation of the intensity of time-dependent fluctuations is compared with the results obtained by a layered artificial neural network model. The desired accuracy of the approximation by the neural network at the end of the learning phase has been estimated by analyzing the statistics of the learning data. The application of the obtained results to the analysis of actual anomaly data from a nuclear reactor showed that neural networks can identify the onset of anomalies with a reasonable success, while usual statistical methods were unable to make distinction between normal and abnormal patterns.

Introduction

Artificial neural networks have been successfully applied to the diagnostics and monitoring of dynamic changes in complex systems; see, e.g., Guo and Uhrig (1992). In the framework of the present study, we consider anomalies which appear step-by-step. These anomalies might remain hidden at the first stage of their development. They can be identified only later, depending on the sensitivity of the applied monitoring method. Time series analysis can be performed both in time- and frequency domains. A widely used tool of the time-domain studies is the sequential probability ratio test (SPRT) introduced by Wald (1947). SPRT and its various advanced versions have been applied to a wide range of anomaly detection problems.

In the present paper, however, we perform frequency-domain studies and evaluate the auto power spectral densities (APSDs) of the time series. An important advantage of the frequency-domain analysis is the additional information obtained on the spectral distribution of the anomaly. This information can be used for the characterization of the identified anomaly. In order to improve the statistical accuracy of the frequency-domain studies, average APSDs are calculated over a certain time interval. The need to perform such an averaging is a clear disadvantage compared to time-domain methods which perform sample-by-sample evaluation and, therefore, can respond faster to the onset of anomaly. In order to reduce this disadvantage, very short time intervals are used for the evaluation of the APSDs.

A 3-layer, fully connected, feed-forward neural network has been applied in the present study to the identification of APSDs corresponding to different anomalies. We applied modified backpropagation learning algorithm with structural learning; see Ishikawa (1992). Based on the evaluation of the statistical accuracy of the input learning spectra, the desired final sum of squared errors between the network output and target values has been estimated. Also the error of the statistical discriminators based on the standard deviation and the band-passed root mean square intensity of the temporal fluctuations have been determined. In the last part of the paper, actual neutron detector signals with very weak anomalies are analyzed. These anomalies cannot be identified by the above statistical methods. The neural network, however, can discriminate between these anomaly patterns as well.

Structural Learning Algorithm for Anomaly Detection

The applied neural network has 3 layers: input, hidden and output layers. We use a feed-forward network with fully connected architecture and sigmoidal transfer function. The learning algorithm is modified back-propagation with structural learning. The power spectral densities enter the network through the input layer with 128 nodes. Each of the input nodes corresponds to a frequency point of the power spectral density

0-7803-1901-X/94 $4.00 ©1994 IEEE

function. The number of nodes in the output layer equals to the number of anomaly types to be identified, plus one additional node representing the absence of anomalies. The i-th node assumes the value 1 if the i-th anomaly occurs and the value 0 in all the other cases. The hidden layer has a large number of nodes; typically half of the number of input nodes. As a result of the structural learning, the number of *active* hidden nodes (i.e. nodes with weights significantly deviating from zero) decreases drastically. Therefore, this neural network is not sensitive to the hidden layer size. In the following, the main features of the applied structural learning algortithm with forgetting is summarized. The basic idea is to update the connection weights as follows (Ishikawa, 1992):

$$\Delta w_{ij} = \Delta w'_{ij} - \varepsilon sgn(w_{ij}) \tag{1}$$

Here $\Delta w'_{ij}$ is the change of the ij-th weight using standard backpropagation algorithm, ε is the so-called forgetting rate. The second term on the right hand side of Eq. (1) describes a permanent decreasing tendency for the connection weights. Indeed, the weights decrease continuously, unless they are reinforced by the backpropagation rule. The corresponding cost function is given by:

$$I = \sum_i (y_i - y_i^*)^2 + \varepsilon' \sum_{i,j} |w_{ij}| \tag{2}$$

y_i and y_i^* are the actual and the target values of the network outputs, respectively. $\varepsilon' = \lambda\varepsilon$, where λ is the learning rate. After completing the learning in accordance with Eqs. (1) and (2), usually a significantly reduced network is obtained, in which only a few hidden nodes remain active. This skeletonized network architecture makes it possible to identify certain (quasi-) causal relationships between dominant frequency ranges and anomaly types. Such a knowledge acquisition has been reported earlier by Kitamura (1992).

Estimation of the Accuracy of the Anomaly Detection

First, we estimate the statistical uncertainty of the input data. In the present study, we can perform this estimation easily by analyzing the statistics of the power spectral density functions both theoretically and experimentally. The relative variance of a spectrum point is given by the relationship (Jenkins, 1968): $\sigma^2_{APSD}/APSD = W_b/2N$. Here N is the number of averaged data records, W_b is the window parameter. In the statistical evaluations, the band-passed root mean square ($RMS(f_1, f_2)$) fluctuation intensity is commonly used. It is defined as the integral of the APSD over the frequency band $[f_1, f_2]$. The relative variance of the RMS can be calculated simply by dividing the relative variance of the spectrum by n, where n is the number of spectral points in the frequency range between f_1 and f_2.

The accuracy of the output values of the neural network can be calculated according to the formula:

$$SSE^* = \sum_{k=1}^{m} x_k N_k B(m) \tag{3}$$

Here SSE^* is the theoretically attainable value of the sum of squared errors of the neural network; $(1 - x_k)$ is the confidence level of the identification in the k-th class of anomalies; N_k is the number of patterns in the k-th class; $B(m)$ is the squared error of random output distribution in the case of m classes. The calculation of x_k and $B(m)$ depend on the actual problem. In the case of 128-point APSDs at the input layer, the corresponding expressions write:

$$B(m) = \left(\frac{1}{m}\right)^2 \frac{m-1}{m} + \left(\frac{m-1}{m}\right)^2 \frac{1}{m} = \frac{m-1}{m^2}; \qquad x_k = \prod_{i=1}^{128} p_k(i) \tag{4}$$

Here $p_k(i)$ is the ambiguity of the i-th frequency point. $p_k(i)$ can be estimated by assuming that the probability distribution of a spectrum point is characterized by two parameters: the mean value and variance. In the case of three patterns, for example, we have three normal distributions. The ambiguity of pattern identification means that that these distributions overlap. Based on the known mean and variance values, the extent of the overlap, i.e. $p_k(i)$, can be estimated. By assuming the statistical independence of the spectral points, the overall uncertainty of the identification of the spectral densities is calculated by multiplying the individual ambiguities.

Results of Anomaly Detection

The above neural network has been applied to the analysis of the signals of ex-core ionization chambers. For experimental details; see Kozma (1992). Very weak anomalies have been introduced artificially in order to test the efficiency of different anomaly detection methods. The signals have been sampled at 32 Hz and auto power spectral densities have been calculated over each 8 s. The APSDs contain 128 frequency points up to 16 Hz. Accordingly, the input layer of the neural net has 128 nodes, while the hidden layer had initially 64 nodes. We have analyzed two types of anomalies in addition to the normal state, thus the output layer contained 3 nodes.

Each of the three training sets of APSDs contained 26 spectra. No changes could be detected in the mean values of the signals as a function of the anomaly state. Also the variances of the APSDs did not show characteristic deviations from the normal values in the case of anomalies. In Fig. 1, the APSDs of a signal are shown at different anomaly conditions. The 26 short-time APSDs are indicated by scattered dots, while the average of the 26 APSDs are shown by solid lines for each of the 3 conditions. The dotted lines are widely spread around the average APSDs.

There is an observable difference between the average APSDs in Fig. 1, especially at higher frequencies. The theoretical value of the standard deviation of the RMS is a few % in this case, so the anomaly is detectable. The dotted curves which correspond to the short-time APSDs, however, significantly overlap and their large standard deviations mask any trend in the APSDs. This conclusion is supported by the data in Table I, where the 7 Hz to 16 Hz band-passed RMS^2 and its standard deviation is given for the 3 anomaly states. It is clearly seen in Table I that σ_{RMS^2} is significantly larger than any change in the average RMS^2, therefore, no anomaly detection is possible on the basis of statistical analysis of RMS^2.

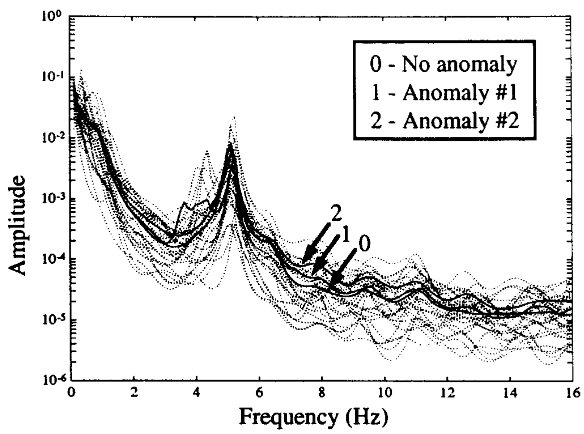

Figure 1 Auto power spectral densities at different anomalies; solid lines : averaged APSDs (208 s); dots : short APSDs (8 s).

Table I High-frequency RMS^2 and Standard Deviation$^+$

Anomaly Type	μ_{RMS^2} $(\times 10^{-4})$	σ_{RMS^2} $(\times 10^{-4})$
0	1.8029	1.0704
1	2.0176	1.1057
2	2.9867	2.1603

$^+$ over : [7 Hz, 16 Hz]

On the other hand, the performance of the neural network is 100 % in the case of the learning data set. Moreover, the neural net has a reasonable capability of identifying anomalies in the actual process data. This is illutrated in Fig. 2, where the evolution of an anomaly is shown. From 0 s to 570 s, anomaly # 0 (i.e. normal system state) took place. Between 570 s and 1070 s, transients occurred with various anomalies. After 1070 s, anomaly #2 has been maintained continuously till the end of the experiment. Although, the network is far from 100 % successful anomaly identification, but it certainly could identify the major trends during the experiment. Its performance can be significantly improved by applying advanced methods of logical inference, e.g., based on fuzzy logics.

Based on Eqs. (3) and (4), the desired value of the SSE of the output values of the NN has been estimated. In our case $m = 3$, $B(3) = 2/9$, and $N_1=N_2=N_3= 26$. x_k has been estimated by using the experimentally determined values of the means and variances of the APSD points. For example, the empirical ambiguity function of the 2nd output node of the network (anomaly #1) is shown in Fig. 3.

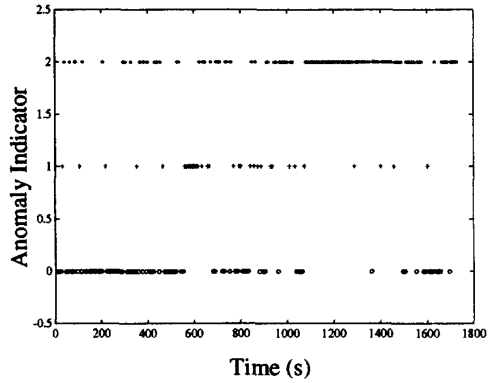

Figure 2 Demonstration of the anomaly detection method. Notations o, +, and * mark anomalies of type 0, 1, and 2, respectively.

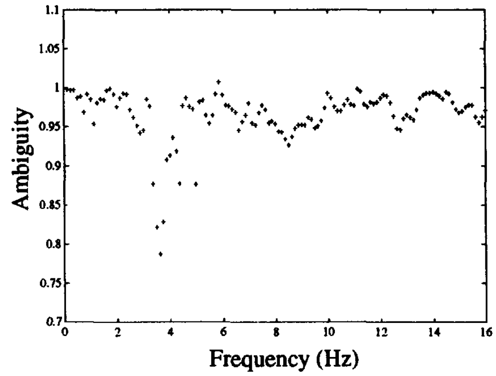

Figure 3 Empirical ambiguity of the individual frequency points for the anomaly represented by the 2nd node of the output layer of the network.

The confidence level of this particular node is 0.985 according to Eq. (4). The theoretical limit of the output error has been evaluated by Eq. (3) to yield $SSE^* = 0.16$. During actual learning, we did not reach this limit. Usually, we stopped learning at SSE values of 0.3 to 0.4.

Concluding Remarks

The desired accuracy attainable during the learning phase in a hetero-associative, layered neural network have been studied. An attempt has been made to estimate this accuracy on the basis of the statistics of the input learning data and the properties of the actual neural network architecture. Neural networks, which have been trained to an error level corresponding to the theoretically determined sum of squared errors, could detect anomalies with a resonable accuracy.

In the analysis, the modified backpropagation algorithm with forgetting has been applied. This algorithm provides the possibility to extract useful information by analyzing the weights of the trained network. In our anomaly identification problem, only 5 nodes remain active from the 64 hidden nodes by the end of the training. It was possible to identify those nodes in the the input layer, which carry the dominant information on each anomaly.

There is an important relationship between the attainable SSE at the input layer of the neural network and the rate of forgetting, ε. We have assigned a rather small value to ε, so that the actual forgetting becomes important if the learning has advanced already; i.e., when the two terms on the right-hand side of Eq. (2) are of the same order. According to our experience, this can lead to a situation when the SSE decreases less intensively. In fact, ε can be used as a control parameter to adjust the final SSE value of the trained network. If we estimate the desired SSE* in advance, e.g. by using Eq. (3), we can set the value of ε accordingly. In this way, the learning by the neural network can become more efficient. The implementation of this strategy is in progress.

References

Guo Z., and R.E. Uhrig (1992) *Proc. SPIE - Int. Soc. Opt. Eng.*, 1709, pp.505-516
Hecht-Nielsen, R. (1990) "Neurocomputing", Addison-Wesley Publ. Co.
Ishikawa, M. (1992) Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks,*Iizuka'92*, pp.855-858
Jenkins, G.M., D.G. Watts (1968) "Spectral Analysis and Its Applications", Holden Day, S.F.
Kitamura, M., et al. (1992) Proc. 2nd Int. Conf. on Fuzzy Logic & Neural Nets, *Iizuka'92*, pp.827-831
Kozma, R., H. van Dam, J.E. Hoogenboom (1992) *Nuclear Technology*, 100, pp.97-110
Wald, A. (1947) "Sequential Analysis", Wiley, New York