STATE FORECASTING FOR ROTARY MACHINE BASED ON NEURAL NETWORK AND GENETIC ALGORITHM

Hongmei Liu
School of Automation Science and Electrical Engineering, Beihang University,
Beijing, 100083, China
liuhongmeigg@tom.com

Shaoping Wang
School of Automation Science and Electrical Engineering, Beihang University,
Beijing, 100083, China,
shaopingwang@vip.sina.com

Pingchao Ouyang
School of Automation Science and Electrical Engineering, Beihang University,
Beijing, 100083, China
Ouy2004@tom.com

ABSTRACT
A state forecasting is a key technology to achieve the advanced predictive maintenance. A Prediction based on neural network is a new approach to realize the state predicting. The present neural networks predicting models are comparatively poor in adaptability to environment and in predicting accuracy, therefore, a new rotary machine online state forecasting method based on the genetic algorithm (GA) and neural network (NN) was presented. GA was used for dynamical optimizing the structure parameters of BP network to obtain the optimal network structure. A training algorithm combining GA with BP was adopted to avoid the local minimum and to heighten the learning precision. The state predicting results for hydraulic pump indicate that the predicting model purposed may dynamically optimize the structure parameters in accordance with different conditions, and gained satisfactory results.

Keywords: Rotary machinery, state forecasting, genetic algorithm, optimization, hydraulic pump

1 INTRODUCTION
A trend predicting is a key technology to achieve the advanced predictive maintenance of rotary machine. The deeper the research on state predicting is, the more serious the nonlinear problem. Consequently, the traditional mathematical prediction method is difficult to achieve the more satisfactory results. Recently, the introduction of artificial intelligent method such as neural network (NN) and genetic algorithm (GA) exploits a new approach to state prediction for rotary machine.

However, neural network structure for state predicting is mainly determined by experiences and experiments, which has following drawbacks:

1) Neural network structures available are so many that it is impossible to search for all structures by experiment, even for a simplest problem.

2) Once the network structure is determined, its structure parameters can not be modified according to working conditions. That is to say, the network is not an open learning system, but a static system without adaptability to the environment.

3) Error back propagation (BP) algorithm is a local optimization algorithm with a slow convergence speed, and easy to get into local minimum. Consequently, in order to improve the online prediction effect, it is necessary to modify neural network prediction model.

Genetic algorithm can simultaneously search for several points in solution space, which overcomes the drawbacks of BP algorithm and is easy to find the global optimization solution. Furthermore, genetic algorithm has higher search efficiency, because it is not an exhaustive search, but a heuristic search. Therefore, genetic algorithm is adopted to determine the approximate range of the optimization solution, and BP algorithm is applied to gain the optimum neural network model.

2 GA AND ITS APPLICATION IN NEURAL NETWORK MODELING
The simple encoding technology and reproduction mechanism is adopted in genetic algorithm to express the complicated phenomenon and to solve quite difficult problem.

Due to the above drawbacks of the neural network, a great number of experiments are carried out before a satisfactory network is obtained. Moreover, it is difficult for the simple network structure determined by experience and the single learning algorithm to guarantee the network precision in different working condition. As a global search approach, genetic algorithm has higher convergence precision and convergence speed. Due to the diversity of network structure parameters, GA is adopted to find an optimum network structure, which makes network parameters more rational than that of determined by experience and experiments. And a training algorithm combining GA with BP is adopted to avoid the local minimum and to heighten the learning precision.

The following problems need to be considered in improving neural network prediction model with GA.

(1) Network topological structure. It is composed of network structure parameters, such as the number of layers and the number of units in each layer etc, and connection mode of the network, for instance, whether adjacent layer nodes is complete connection or not, and whether feedback is allowed or not.

(2) Learning algorithm. What algorithm is chosen to adjust the network weights?

(3) The data input into the network. It includes the selecting and smooth processing for the training sample, eliminating singular point and extracting the trend item etc.

## 3 PROCEDURE OF GA AND THE CHOICE OF FEATURE PARAMETERS

Basic process of GA is as follows:

1. Stochastically produce \( N \) individuals, and form the initial population.
2. Compute the fitness value of the individuals.
3. Arrange the individuals in fitness value sequence.
4. Select parents according to fitness value acting as the selection weights, and produce an offspring through crossover operation and mutation operation.
5. Substitute the newly-produced offspring for individual with the lowest fitness in population and return to (2).

The selection of individuals: GA is influenced by the number of individuals \( N \). If \( N \) is too small, GA will be worse or can not find the solution to the problem at all because too few individuals can not supply enough samplings. If \( N \) is too large, calculation quantity will increase, and convergence speed will become slow. Generally, the number of individuals is selected as \( N_p = 10\text{--}50 \).

The selection of crossover rate: after ranking fitness function, stochastically select two individuals in the former \( N_p / 2 \) individuals and carry out crossover operation at the crossover rate \( P_c = 0.5 \). After circulating \( N_p / 2 \) times, newly-produced individuals replace the late \( N_p / 2 \) individuals in sequence. Thus crossover operation is completed in a generation. Crossover rate controls the frequency of crossover operation. If \( P_c \) is too large, individual with high fitness value will be destroyed; if \( P_c \) is too small, the search will stop. Therefore, generally, \( P_c = 0.25\text{--}0.75 \).

The selection of mutation rate: mutation operation is the second factor increasing the population diversity. If mutation rate \( P_m \) is too small, new gene block will not be produced; and if \( P_m \) is too large, GA will become random search. Generally, \( P_m = 0.01\text{--}0.2 \).

## 4 GANN MODELING

### 4.1 Optimizing the network structure parameters using GA

Optimization for the network structure with GA actually seeks for an optimum network structure parameters to obtain the best prediction effect.

The structure parameters of a feed forward network consist of the number of hidden layers, the number of nodes in each hidden layers, learning rate and momentum factor, in which the number of hidden layers and that of nodes in each hidden layer represent sufficiency of network input information, learning rate determines network learning speed, and momentum factor restrains the occurrence of oscillation. It can be expressed as follows:

\[
\Delta W_{kj}(m+1) = \eta \delta_k O_j + \alpha \Delta W_{kj}(m)
\]

where \( \Delta W_{kj}(m) \) is the increment of the weights, \( \delta_k \) is the gradient of the \( k \)th layer, \( O_j \) is the output of the \( j \)th layer, \( \alpha \) represents the momentum factor, and \( \eta \) is the learning rate.

(1) **Encode the parameters.** Every chromosome string is encoded with 18 bit gene codes shown in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta )</td>
<td>( \alpha )</td>
</tr>
</tbody>
</table>

Where \( H_n \) represents the number of the hidden layers, \( H_1 \) stands for the number of the nodes in the first hidden layer, and \( H_2 \) is the number of the nodes in the second hidden layer.

The calculation formulas are as follows:
\[ \eta = 0.1 \times (2 + 2^0 \beta_3 + 2^1 \beta_2 + 2^2 \beta_4) \]
\[ \alpha = 0.1 \times (2 + 2^0 \beta_5 + 2^1 \beta_6 + 2^2 \beta_4) \]
\[ H_n = \begin{cases} 1 & \beta_1 \beta_8 = 0 \\ 2 & \beta_1 \beta_8 \neq 0 \end{cases} \]
\[ H_1 = 2^0 \beta_{13} + 2^1 \beta_{12} + 2^2 \beta_{11} + 2^3 \beta_{10} + 2^4 \beta_9 \]
\[ H_2 = 2^0 \beta_{18} + 2^1 \beta_{17} + 2^2 \beta_{16} + 2^3 \beta_{15} + 2^4 \beta_{14} \]

(2) Determining Fitness function. The evaluation criterion of network structure generally consists of convergence precision and convergence rapidity which are included in fitness function. Given a maximum iterative times for every individual which is a chromosome string standing for a network structure, if a certain individual can not convergence with prescribed accuracy in given iterative times, the individual will be eliminated from the population and its fitness function is set \( F = 0 \). For those individuals that can convergence with given accuracy, the concept of convergence precision ratio is purposed, which acts as their fitness function. Convergence precision ratio can be expressed as follows:

\[ F = a_{es} = \frac{\log(E_0 - E_1)}{-\log t} \]  \hspace{1cm} (3)

Where \( a_{es} \) represents convergence precision ratio, \( E_0 \) \((<10^{-3})\) is the given convergence precision, \( E_1 \) \((<10^{-5})\) is the actual convergence precision, and \( t \) stands for run time of the procedure.

Considering training samples and initial weights value will affect the convergence precision of the network, the same training samples and the same initial weights value are used in training.

(3) Arrange the individuals in fitness value sequence, and substitute sequence number for individual rank.

(4) Select parents according to the selection weights, and produce an offspring through crossover operation and mutation operation.

(5) Substitute the newly-produced offspring for the lowest fitness individual in population and return to (2).

Note: the procedure sets a clock function to record the learning time when every individual starts learning.

4.2 Optimizing network weights with GA

After structure parameters of the network are determined and before network formally starts to train, GA is firstly adopted to optimize initial weights of the network, which specifies the convergence direction of network precision. Then BP algorithm is adopted to train the network for several times. Lastly, GA is adopted to optimize weights of the network. This method can effectively inhibit network from plunging into local minimum, Shorten the training time and heighten the convergence precision.

(1) encoding

Neural network has \( p = \sum_{i=1}^{l} (p_i + 1) \) parameters which are connection weights, where \( p_i \) represents the number of nodes in the \( i \) th layer. Every nodes consists of a threshold value \( \theta \), and every parameter can be expressed by a gene string composed of 8 binary bits. Consequently, every parameter can be random encoded as follows:

\[ w_{i1} \quad w_{i2} \quad w_{i3} \quad w_{i4} \]

(2) fitness function

The error between tutorial signal and actual signal acts as fitness function in genetic operation, and can be expressed as follows:

\[ F_j = M - m_j \]  \hspace{1cm} (4)

\[ m_j = \frac{1}{N} \sum (t_i - o_{ij})^2 \]  \hspace{1cm} (5)

where \( F_j \) represents the fitness value of the \( j \) th individual, \( m_j \) is the error of the \( j \) th individual, \( t_i \) stands for the tutorial signal of the \( i \) th sample, \( o_{ij} \) represents the \( i \) th sample actual output of the \( j \) th individual, \( N \) is the number of the training samples, and \( M \) stands for the maximum of given \( m_j \).

\[ F_j \] is a function of \( w \) and \( \theta \), that is \( F_j = f(w, \theta) \). Optimizes the objective function with GA and makes \( F_j \rightarrow M \).

5. APPLICATION OF GANN MODEL IN HYDRAULIC PUMP STATE PREDICTION

In fact, the state forecasting for hydraulic pump mainly predicts the evaluation parameters of states. Vibration intensity and vibration feature spectrum act as state evaluation parameters. It is necessary to choose different network structure according to different data source and different prediction requirement. The requirements for network structure are also different even for the same data source because of the influence of random factors on the spot or in different history time section. Due to GANN model is an online optimization model based on GA, optimal network can dynamically be searched for to realize data prediction according to different kind of data source.
5.1 GANN modeling in state prediction for hydraulic pump

The prediction effect of two methods is compared according to time serials of vibration intensity of hydraulic pump in specified working condition. For the convenience of comparison, selects data in two time sections. The unit of time serials is day. The former 60 days’ data act as training samples to predict the latter 40 days’ data. 20 input nodes and 1 output node are chosen in network prediction model. The aim of choosing one output node is to make full use of the samples. And the aim of choosing several input nodes is to make use of more data information. Input time serials are as follows:

\[
\{X_{t-1}, X_{t-2}, \ldots, X_{t-20}\} = \{X_1, X_2, \ldots, X_{20}\}
\]

Above data will be normalized in actual prediction.

GANN modeling consists of two steps:

Step 1: seek for optimal neural network structure. As described above, the number of hidden layers, the number of the nodes in every hidden layer, learning rate and momentum factor are encoded with an 18-bit gene string. Each gene string is an individual standing for a neural network structure. Population size is 30, iterative generation is 80, 30 individuals (neural network) are trained in each generation. Rank them according to the convergence precision ratio, and stochastically select two individual, then carry out crossover operation at crossover rate of 0.5. After looping for 15 times, newly-generated 15 individuals replace the latter 15 individuals in order. Thus the crossover operation is completed in a generation. And then mutation operation carries out at mutation rate of 0.15. Then genetic operation starts again in next generation. After 80 generation, the optimum network structure is found and its parameter is as follows:

Time section 1: \( H_n = 1 \) \( H_i = 17 \) \( \alpha = 0.4 \) \( \eta = 0.8 \)

Time section 2: \( H_n = 1 \) \( H_i = 23 \) \( \alpha = 0.2 \) \( \eta = 0.8 \)

The network parameters determined by experience are as follows:

\( H_n = 1 \) \( H_i = 26 \) \( \alpha = 0.4 \) \( \eta = 0.8 \)

Step 2: after optimum network structure is determined, initial weights of the network will be optimized. As described in 4.2 section, the number of connection weights is \( p \), every connection weights is expressed by 8-bit binary genes, and all connect weights can be combined together to form a gene string with 8p-bit (a individual). Suppose the population size is 20, and iterative number is 50. Every individual respectively stands for the different initial weights. Then in every generation's genetic operation, 20 neural networks with different initial weights are trained, and mean square deviation acts as fitness function to carry out genetic operation. After 50 generation, optimum network initial weights are obtained, which are used as initial weights to train optimum neural network.

5.2 Adaptive dynamic adjustment of GANN prediction model

GANN model has the self-adaptive ability, which varies with different data source and different historic time section. In trend prediction of vibration intensity for hydraulic pump, in order to obtain optimal prediction effect, GANN prediction model is modified to obtain the new neural network structure every three months to adapt to variation of data source in the different time section.

5.3 Prediction results analysis

Figure 1 and figure 2 are predicted results of neural network determined by experience. Network structure parameters are not changed anymore once it is determined firstly, that is to say, vibration intensity in two continuous time section is predicted with the same network. Figure3 and Fig. 4 are predicted results of dynamic adaptive network optimized by GA. Unsmooth curve 1 is the actual measurement and smooth curve 2 is the predicted value.

![Figure 1](image1)

**Figure 1** Forecasting model with experience in time section 1

![Figure 2](image2)

**Figure 2** Forecasting model with experience in time section 2
Combining Tab.2 and comparing Fig.1 with Fig.2, we can see that GANN model effectively carries out parallel search in the large range and find the optimum neural network. Meanwhile the network weights are also optimized to avoid the local minimum. Therefore, all indexes are better than network model established by experience. Comparing the Fig.1 with Fig.2, we can see that the prediction precision of empirical network model decreases in the second time section than in the first time section. However, from Fig.3 and Fig.4, it can be seen that GANN model has higher prediction precision whether in the first time section or in the second time section. This is mainly due to GANN model carries out the network optimization for the different learning samples in different time section and dynamic modifies the network structure parameters.

### 6 CONCLUSIONS

A trend prediction is a key technology to realize advance maintenance for the rotary machine, and neural network is a new approach to realize the trend prediction. Considering structure parameters of neural network used for predicting rotary machine is determined by experience and experiments, adaptability to environment is worse and prediction precision is lower. Combined with actual working conditions, an online adaptive trend prediction method is purposed. GA is adopted to optimize the BP network structure parameters. A learning algorithm of combining the GA with BP is adopted to train the neural network.

Modified prediction model can adaptively learn according to the different historical data, therefore, optimal structure parameters of the network are dynamical determined. Experiments results demonstrated that prediction precision of improved model is heightened and real time performance is effectively heightened.

### REFERENCES


