Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004

THE APPLICATION OF RBF NETWORKS BASED ON ARTIFICIAL IMMUNE ALGORITHM IN THE PERFORMANCE PREDICTION OF STEEL BARS

YING ZHOU[1,2], DE-LING ZHENG[2], ZHI-LIANG QIU[1], GUO-YA DONG[1]

[1] School of electrical engineering and automation, Hebei University of technology Tianjin 300130, China
[2] Information Engineering School, University of Science and Technology Beijing 100083, China
E-MAIL: zhouyingq066@sina.com

Abstract:

This paper presents a novel radial basis function (RBF) neural network model based on immune recognition principle. This model can choose the number and location of the hidden layer centers by applying the principles of recognition, memory, learning and self-organized adjustment, and can determine the weights of the output layer by adopting least square algorithm. This novel model is applied to predict the performances of hot-rolled steel bars, and it achieves good effect. Simulation results show that this model proposed in the paper has the advantages of less computation and higher precision, compared with the k-means algorithm.

Keywords:

Artificial immune; immune recognition; RBF neural network; hot-rolled steel bars; performance prediction

1. Introduction

The hot-rolled steel bar is a steel product with high rate of finished products. Its mechanical performance is one of the main indexes reflecting steel bar quality, and have great influence on users' economic benefit. The main factors affecting the mechanical performance are its chemical components and main technical parameters in the hot-rolled process [1]. Because the rolling system is a complex nonlinear system and the production process is disturbed by various random factors, it is extremely difficult to express the influence of technical parameters on the mechanical performance in the form of math. Hence, the fast and exact prediction of product mechanical performances and prompt adjustment of production technology are significant for the improvement of the quality.

Over the last few years, the Artificial Immune System (AIS) has become a new research area and has shown powerful ability of information processing and problem solving in optimization, machine learning, information security, and fault diagnosis [2-4]. The combination of the artificial immune system and the artificial neural network has becomed the research hotspot and has been applied to solve the intelligent technical problems, including the improvement of the network performance and the design of the neural network controller based on the immune feedback principle [5].

The RBF neural network is capable of approximating arbitrary nonlinear mapping, and the algorithm is simple and useful. It has been successfully applied in signal processing, system modeling, process control, and fault diagnosis [6]. In the process of the RBF designing and training, the number and positions of the RBF hidden layer centers directly influence the approximate capability of the network, and centers are required to cover the entire input space. If the number of the RBF hidden layer centers is too large, it will remarkably increase computation, and weaken network generalization capability. In reverse, if the number of centers is too small, it will lead to weaker classification capability and poorer self-adaptive. In order to overcome above disadvantages, a novel RBF neural network model is proposed in this paper based on the principle that the biological immune system can recognize antigens and generate memory antibodies so that the entire input space can be covered with less centers. In this paper, the model is applied for the performance prediction of the hot-roll steel bars in order to improve the mechanical performance of steel.

2. Immune Principle

Immunity is a specific physiological response of organism. When a pathogen invades our body, the immune system can recognize and eliminate it by the lymphocyte cells throughout body to keep the homeostasis of internal environment. The immune cells recognize antigens by complementary match of epitope (attached to the antigen) and paratope (attached to the immune cell). The match is a continuous learning process, and the best-matched antibodies are finally selected to bind antigens for eliminating them.

Immune memory is: the antibodies that can recognize antigens can be reserved by immune system as memory cells. When the same antigen invades again, the corresponding memory cells can be activated to generate a large number of antibodies to recognize antigens, so the recognition time is shortened.

Immune adjustment is: a large number of antibodies were produced in the course of immune response, so it reduces the stimulation of antigens to immune cells, therefore suppresses the differentiation and proliferation of antibodies. At the meantime, the mutual suppression and stimulation also exist between antibodies. The mutual restriction relation between antigens and antibodies, antibodies and antibodies makes the immune system remain a dynamic balance.

3. RBF neural network

A RBF neural network is composed of input layer, the hidden layer and the output layer, illustrated in Figure 1. It

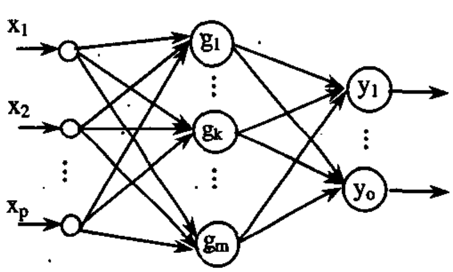

Figure 1. RBF neural network

can be regard as a feedforword network. The activation function of the hidden layer is RBFs.

For a RBF network with p-dimensional input vector $X_i=[X_{i1},X_{i2},\cdots,X_{ip}]^T \in X$ and N input vectors $X=[X_1,X_2,\cdots,X_N]$, the RBF network output is $Y=[y_1,y_2,\cdots,y_o]$, where o is the number of output nodes, and can be computed by the following equation:

$$y_i = W_i^T G = \sum_{j=1}^{m} w_{ij} g_j \ , i=1,2,\cdots,o \tag{1}$$

where $W_i=[w_{i1},w_{i2},\cdots,w_{im}]^T, i=1,2,\cdots,o$, are the connection weights form the hidden layer to the output layer, $G=[g_1,g_2,\cdots,g_m]$ is the vector of radial basis function in hidden layer.

The paper chooses the Gaussian function as the basis function:

$$g_j = \psi_j(\|X_i - c_j\|/\sigma) = e^{-\frac{\|X_i-c_j\|^2}{\sigma^2}} \tag{2}$$
$$j=1,2,\cdots,m \quad i=1,2,\cdots,N$$

where X_i is the i-th input vector, c_j is the j-th center of the Gaussian function, $\|X_i - c_j\|$ is a Euclidean norm and σ is the width of the Gaussian function:

$$\sigma = \frac{d_m}{\sqrt{2M}} \tag{3}$$

where M is the number of centers and d_m is the maximum distance between them [6].

The key factor, influences the performance of the network, is the selection of the basis function centers, rather than the type of the nonlinear function adopted in the RBF network. An inappropriate choice of centers might lead to unsatisfactory performance of RBF network. At present the main methods to determine the RBF network centers are: random selection of centers and the k-means clustering algorithm. The limitation of the random selection of centers is that errors are larger. The limitation of the k-means selection of centers is that the number of clusters should be given in advance and it is easy to converge to the local minimum.

The paper applies the artificial immune principle to choose and optimize the RBF network centers, and a prior determination of the number of the hidden layer centers is unnecessary. The input space is covered as large as possible by lesser centers. Herein, antigens represent the input data, and antibodies represent the hidden layer centers. An antibody memory set, established by applying the immune recognition algorithm, can obviously improve the leaning speed when similar input data invade again.

4. Theory of artificial immune recognition

4.1. Implementation of Affinity

The immune system is an information processing system with high efficiency. Eacing the large number of various antigens, the immune system has to recognize them for their posterior elimination. The process of recognizing antigens is to search for antibodies with the maximum affinity with antigens. In this paper, the affinity between antibody and antigen is defined by the equation (4), while the one between antibody and antibody by equation (5). The details are described as follows.

Definition 1: S^L is an L-dimensional shape-space constructed by antigens and antibodies,
$AB = \{Ab_i | i = 1,2,\cdots,n\}$ is the aggregate of antibodies,
$AG = \{Ag_j | j = 1,2,\cdots,m\}$ is the aggregate of antigens,
Ag_j and Ab_i are vectors of the two aggregates, respectively.

Definition 2: In shape-space S^L, the affinity $a_{ij}(0 \le a_{ij} \le 1)$ between antigen Ag_j and antibody Ab_i is:

$$a_{ij} = 1 - \frac{\|Ab_i - Ag_j\|}{\max\limits_{\substack{1\le i\le n\\ 1\le j\le m}} \|Ab_i - Ag_j\|} \quad \begin{cases} i = 1,2,\cdots,n \\ j = 1,2,\cdots,m \end{cases} \tag{4}$$

where $\|Ab_i - Ag_j\|$ is the norm of Ab_i and Ag_j in S^L. Similarly, the affinity $s_{ij}(0 \le s_{ij} \le 1)$ between the antibody Ab_i and antibody Ab_j is:

$$s_{ij} = 1 - \frac{\|Ab_i - Ab_j\|}{\max\limits_{\substack{1\le i\le n\\ 1\le j\le m}} \|Ab_i - Ab_j\|} \quad \begin{cases} i = 1,2,\cdots,n \\ j = 1,2,\cdots,n \end{cases} \tag{5}$$

4.2. Antibody's clone and variation

In immune system, the antibodies, which have high affinity with the antigens, are selected for proliferation and differentiation (clone), and experiences a hypermutation process at the same time. Clone rate is in direct ratio to affinity. The number of the antibodies clone Nc depends on the affinity given by equation (6) [8]

$$Nc = \sum_{i=1}^{n} round(N - norma\|Ab_i - Ag_j\|N) \tag{6}$$

The antibody's variation Ab^* is given by equation (7),:

$$Ab_i^* = Ab_i - \alpha(Ab_i - Ag_j) \tag{7}$$

where Ab_i is the antibody to be mutated, $\alpha = 1 - e^{-\|Ab_i - Ag_j\|}$ is the ratio of mutation. The higher the affinity, the smaller the α.

Theorem 1: After mutation, the antibody-antigen affinity a_{ij}^* is larger than the previous affinity a_{ij}.

Proof: Set an L-dimensional antibody $Ab_i = \{x_1, x_2, \cdots, x_L\}$, and an L-dimensional antigen $Ag_j = \{y_1, y_2, \cdots, y_L\}$, the variation of the antibody is $Ab_i^* = \{x_1^*, x_2^*, \cdots, x_L^*\}$.

$$\because \quad \|Ab_i - Ag_j\| = \sqrt{\sum_{p=1}^{L}(x_p - y_p)^2} \ge 0,$$

$$\therefore \quad e^{\|Ab_i - Ag_j\|} \ge 1, \quad 0 \le e^{-\|Ab_i - Ag_j\|} \le 1.$$

$$\because \quad \alpha = 1 - e^{-\|Ab_i - Ag_j\|},$$

$$\therefore \quad 0 \le 1 - \alpha \le 1.$$

then:

$$\begin{aligned}
\|Ab_i^* - Ag_j\| &= \sqrt{\sum_{p=1}^{L}[(x_p - \alpha(x_p - y_p)) - y_p]^2} \\
&= \sqrt{\sum_{p=1}^{L}[x_p(1-\alpha) - y_p(1-\alpha)]^2} \\
&= \sqrt{\sum_{p=1}^{L}[(1-\alpha)(x_p - y_p)]^2} = (1-\alpha)\sqrt{\sum_{p=1}^{L}(x_p - y_p)^2} \\
&= (1-\alpha)\|Ab_i - Ag_j\| \le \|Ab_i - Ag_j\|
\end{aligned}$$

therefore: $1 - \dfrac{\|Ab_i^* - Ag_j\|}{\max\limits_{\substack{1\le i\le n\\ 1\le j\le m}} \|Ab_i^* - Ag_j\|} \ge 1 - \dfrac{\|Ab_i - Ag_j\|}{\max\limits_{\substack{1\le i\le n\\ 1\le j\le m}} \|Ab_i - Ag_j\|}$

So that $a_{ij}^* \ge a_{ij}$

4.3. The stimulation and suppression of antibodies

After clone and variation, the antibodies with high antigenic affinity transform into the long living memory cells, others with low antigenic affinity are eliminated. In the immune system, the mutual recognition between antibodies leads to stimulation and suppression. The suppression is achieved by eliminating self-cells with the given suppression threshold ts. The ts is given as below [9]:

$$ts = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} s_{ij}}{N*(N-1)/2} \tag{8}$$

where s_{ij} is the affinity between antibodies Ab_i and Ab_j, N is the total number of antibodies.

5. The coupled algorithm by the artificial immune and RBF neural network

(1) The RBF network hidden layer centers can be determined by applying the immune principles presented above.

In the algorithm, the input data was regarded as antigens and the hidden layer centers as antibodies. The concrete algorithm is presented as follows:

Step 1: Normalize the input data, regarded as antigens, and then the RBF centers $C_i, i=1,2,\cdots,n$, corresponding to antibodies, are initialized using a new set of data created randomly over the interval 0-1.

Step 2: Process each of the data as follows:

Step2.1: Substitute C_i into the equation (4) instead of Ab_i to determine the antigenic affinities to all centers and sort order;

Step2.2: Apply equations (6), (7) for the clone and variation of centers (antibodies);

Step2.3: After variation, recalculate the affinities between antigens and centers;

Step2.4: Reselect centers with higher affinity to create a memory matrix M_p of centers, and eliminate those centers whose affinity is inferior to threshold (the death of the antibody);

Step2.5: In memory matrix M_p, calculate the affinity between center and center by the equation (5), the higher the affinity, the stronger the suppression. Eliminate the centers whose affinity s_{ij} is lower than suppression threshold ts given by equation (8).

Step 3: Further suppress centers in all center memory matrix C_l obtained from step 2, where $C_l = C_l \cup M_p$. It embodies the clone suppression in immune system, and being used to eliminate self-recognized antibodies.

Step 4: Centers generated randomly are added to memory matrix $C = C_l \cup rand(C)$, then go to step 2 until the number of centers is achieved.

Because the hidden layer centers are the compression clustering mapping of the input data, these centers obtained by above immune algorithm can cover the entire input space with less number. The hidden layer centers compose a memory set, so it can remarkably improve the convergence speed when the RBF network recognizes similar data again. The algorithm is an adaptive clustering learning algorithm that does not necessarily to determine the number of the hidden layer nodes a priori.

(2) The output calculation of the hidden layer

The equation (2) is used to calculate the output of the hidden layer, after centers being determined.

(3) The estimation of the weight matrix W for the output layer of the RBF network.

The least square method is adopted for the estimation of W, as following:

$$GW = Y_d \tag{9}$$

where G is the vector of radial basis functions in the hidden layer, $W = [W_1, W_2, \cdots, W_o]$ is the weight matrix of the output layer, Y_d is the desired output vector of the training data.

W can be obtained by the following equation (10) as,

$$W = G^{+}Y_d = (G^T G)^{-1} G^T Y_d \tag{10}$$

where G^{+} is the Moore-Penrose pseudo-inverse of G [10].

6. Applications

The factors affecting the performance of the hot-rolled steel bar (e.g. yield point, strength in tension, extension percentage.) include the bar diameter, components (e.g. carbon, silicon, manganese, phosphorus, sulfur, vanadium, carbon equivalent, listed in table 1).

Table 1 The pattern character of steel bar

No	Dia (mm)	C (%)	Mn (%)	…	Averaged yield points (MPa)	Strength in tension (MPa)
1	20	0.22	1.5	…	530	647.5
2	20	0.2	1.53	…	515	640
…	…	…	…	…	…	…

It is important to rapidly and accurately predict the mechanical performance based on above technological parameters and primary chemical components. Five variables, chosen from bars with same diameter, were normalized, and are used as the input vectors of RBF network, i.e., antigens. The network centers are regard as antibodies, and the yield strength as the output of network.

150 data sets are chosen for one quarter, 100 sets for training and 50 sets for testing. The suppression threshold is 0.15 given by equation (8). After 15 iterations, the number of centers is 33, the hit rate of prediction is 90%. Results depicted in Figure 2.

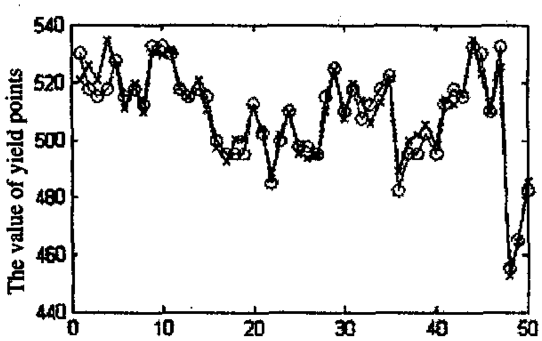

Figure 2. The comparison of yield point between the predicted and real values, where"o"is the actual value, "x" is the predicted value

Figure 3. shows the comparison between the proposed method and k-means clustering algorithms for the selection of centers with error $e < 0.1$. The computation time is 7.5s by the proposed algorithm, whereas 40.3s by k-means clustering algorithm. It leads to the conclusion that the proposed algorithm is feasible.

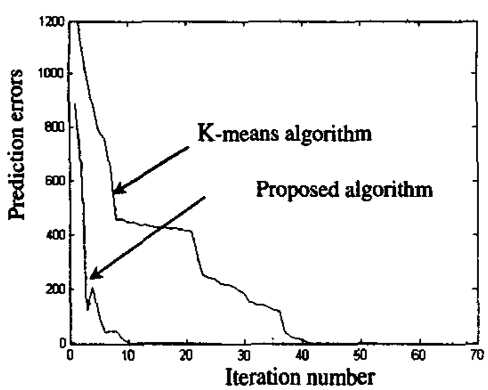

Figure 3. The comparison of convergence between these two Algorithms

7. Conclusions

The paper presents an artificial immune recognition algorithm based on the principle that the immune system can recognizes various antigens and generates memory antibodies. By combining the algorithm with the RBF neural network, an immune neural network model is established and used to predict the yielding points (the mechanical performance of the hot-rolled steel bar). The simulation results show that the algorithm possesses higher prediction hit ratio and needs less calculations. In addition, the algorithm has remarkable effect on the object with a great amount of data.

Acknowledgements

This paper is supported by the Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP), No.20020008004.

References

[1] Dianhui Wang, Zhenyu Liu, and Guodong Wang et al, The mechanical performance prediction of the hot-rolled steel bar by Neural Networks, Iron and Steel, Vol 30, No. 1, pp. 28, 1995.

[2] Deling Zheng, Ruixin Liang, and Dongmei Fu et al, Application of Artificial Immune System and Artificial Immune Genetic Algorithm to optimization, Journal of University of Science and Technology Beijing, Vol 10, No. 3, pp. 57-59,2003

[3] Kesheng Liu, Xianlin Cao, and Haoran Zheng et al, The method of solving TSP problem based on the immune algorithm, Computer Engineering, Vol 26, No. 1, pp. 1-2,2000.

[4] Hunt. J. E, Cooke D. E, Learning Using an artificial immune system. In Journal of Network and Computer Applications: Special Issue on Intelligent Systems: Design and Application, No. 19, pp. 189-212,1996.

[5] Hoffmann. G. W, A Neural network model based on the analogy with the immune system, Tounal of Theoretical Biology, No. 122, pp, 33, 1986

[6] Hongyuan Wang, Guodong Shi, Artificial neural network technique and application, China Petrochemical Press, Beijin, 2002.

[7] Anshen Qi, Chanying Du, Nonlinear models in immunity, Shanghai Science Education Press, Shanghai, 1998.

[8] Castro. De, Fernando. J. An Evolutionary Immune Network for Data Clustering. In Proceeding of the IEEE SBRN, pp. 84-89, November 2000.

[9] Timmis. J, Neal. M, and Hunt. J, An artificial immune system for data analysis, BioSystems, No.55, pp. 143-150,2000.

[10] Gomm. Barry. J, Liyu Ding, Selecting radial basis function network centers with recursive orthogonal least squares training, IEEE Trans. Neural Networks, Vol 11, No. 2, pp. 306, 2000.