Transaction of Azerbaijan National Academy of Sciences, Series of Physical-Technical and Mathematical Sciences: Informatics and Control Problems, Vol. XXXIV, No.6, 2014
www.icp.az/2014/6-20.pdf

P.I. BIDJUK, A.S. GASANOV, S.H. ABDULLAYEV

STRUCTURAL AND PARAMETRIC SYNTHESIS OF PREDICTIVE RBF NEURAL NETWORKS USING ARTIFICIAL IMMUNE SYSTEMS

The purpose of the study is in development of methodology for the structural and parametric syntheses of predictive radial basis function neural networks using the ideas from artificial immune systems. The settings for the RBF neural networks were determined using appropriately constructed immune system, and combined method for forecasting time series with controlled parameters of model is proposed. It was established that increase in the population size slows down the neural network learning process, but on the other hand it resulted in improvement of the models quality. The combined forecasting algorithm and wavelet neural network shows higher accuracy of prediction than the combined algorithm and RBF network, while the latter has a higher rate of training. It was also established that for a higher level of mutation, which implies a high variability of clones of the population, the training is faster, but stability of the process is lower, which decreases the probability of finding the global optimum.

Keywords: artificial immune systems, neural networks, radial basis functions, forecasting

1. Introduction.

Development, formal representation and design of immune and hybrid immune systems suggests the presence of three components (Fig.1):

1) the scheme of the components of the AIS;
2) one or more measures to quantify the states of the system (affinity and measures to evaluate the fitness),
3) immune algorithms that control the behavior of the system.

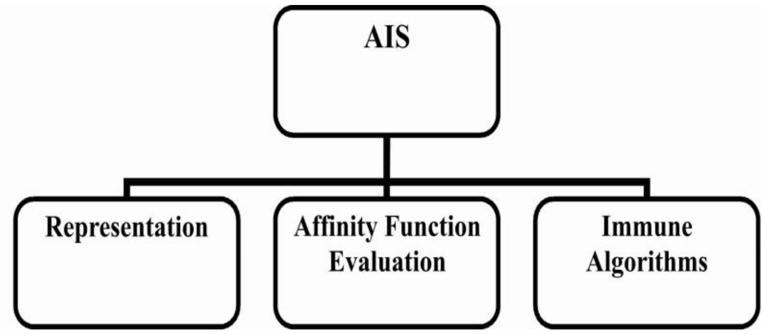


Fig.1 Structural components of AIS

The representation accepted for immune cells and molecules is an expanded version of the shape-space approach. Currently, the most frequently used are the following main types of shape-space: real-valued, integer, and symbolic shape-space [1, 2]. More complex types of shape-space can be used in addition to these forms, such as neural networks, fuzzy neural network, fractal and DNA shape-space. In this paper, the optional choice for shape-space is selected based on radial basis functions (RBF) neural network architecture.

2. Theoretical Part.

As shown in [1], application of gradient methods for local search that are used to construct neuromodels in some cases is not acceptable or impossible. The generalized optimization problem of synthesis of a neural network on the training set can be formulated as follows:

$$Net = Net(\mathrm{M}, \Omega, \mathrm{B}, \Delta, \mathrm{A}),$$

for which

$$\zeta(Net, X, Y) \rightarrow \min,$$

where M is a matrix that determines the presence of synaptic connections between elements of the system (receptors, neurons); $\Omega = \Omega(\mathrm{M})$ is a matrix of weights, which correspond to those present in the network; $\mathrm{B} = \mathrm{B}(\mathrm{M})$ is the bias vector of the neural network; $\Delta = \Delta(\mathrm{M})$ is the vector of discriminant functions for elements of the neural network; $\mathrm{A} = \mathrm{A}(\mathrm{M})$ is the vector of activation function for neural network; $\zeta(Net, X, Y)$ is the criterion for determining the effectiveness of a neural network model to approximate the relationship between inputs X and their corresponding parameter vector of output values Y.

The optimality criterion for a neural network model can be applied in the form of mean square error:

$$\zeta = \sum_{p=1}^{m} \left(y_p - y(Net, \Psi_p)\right)^2,$$

where Ψ_p is the set of values for the p-th instance; $y(Net, \Psi_p)$ is the value of neural network output obtained for the set of values Ψ_p.

2.1. Synthesis of computational structures using immune algorithms for solving forecasting problem.

It is known that the quality of forecasts can be improved by combining the results obtained by different methods. Block diagram of RBF network is shown in Fig. 2. The RBF network consists of input, single hidden (radial basis) and linear (output) layers. The neurons of hidden layer operate on the principle of centering on the training sample elements. The centers are supported by the weight matrix (W^r). The function “dist” is used for calculating the Euclidean distance between input vector (X) and the corresponding center. Around each center, there is a region called the radius. Radius (sensitivity of the network) is adjusted by means of smoothing coefficients vector: $(\sigma_1, ..., \sigma_m)$.

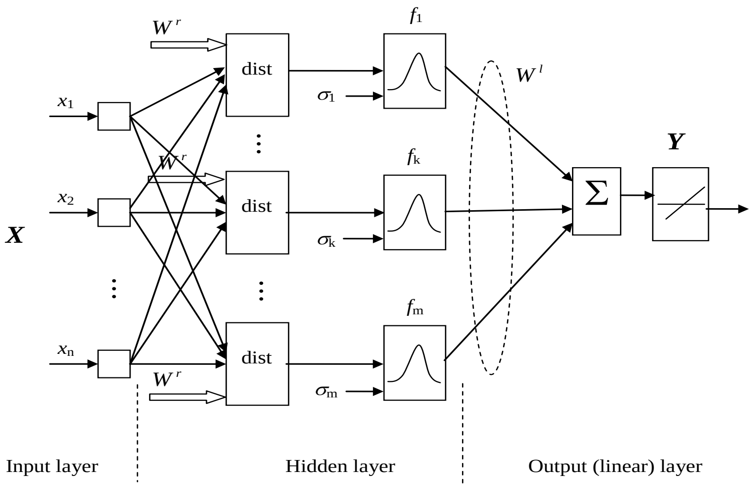


Fig. 2. Architecture of RBF network

Behavior of RBF network depends on the number and position of radial basis functions of the hidden layer. Indeed, for any real *n*-dimensional input vector $x = (x_1, x_2, ..., x_n)$, where $x \in X \subset \Re_n$, the output of the network will be determined as follows:

$$y_i = \sum_{k=1}^{m} w_{ik}^l \, f_k \left(dist\left(x, w_k^r\right), \sigma_k \right),$$

where, $w_{ik}^l \in W^l, i = \overline{1, p}$ is the weight of the linear layer; $w_k^r \in W^r$ are centers of radial basis functions. If Gaussian function is used as a basis one, then:

$$f_k(x) = -\frac{dist(x, w_k^r)^2}{2\sigma_k^2}, \; k = \overline{1, m}.$$

In the context of approximation problem, the network is used to find the function, $y : \Re_n \to \Re$, satisfying equation (5) at $p = 1$. Suppose we have a sample of training data points : $X_1, ..., X_S$, $X_i \in \Re_n$. If the output values for each of these points of $d_1, ..., d_S$, $d_i \in \Re$ are known, then every basis function can be centered on one point of X_i. Asymptotically the number of centers, and therefore the hidden layer neurons, will be equal to the number of data points of the training sample, M. In this case appear at least two problems. First, low ability to generalize as far as the presence of excessive number of neurons in the hidden layer has a negative impact on the approximation of the new data (not participating in the training sample), and second, a large size of training sample will inevitably cause problems of computational nature. To overcome these difficulties, the network should be simplified by reducing the number of basis functions, what in turn poses a new challenge touching upon their optimal centering. The basic idea of the work is to use an immune network for identification of centers of radial basis functions, i.e. solving the problem of recognition and clustering that solves the problem of determining the number of input values. After that a clonal selection algorithm is used for constructing an optimal architecture of radial basis neural network (number and type of RBF-neurons in the hidden layer functions) as well as optimizing the weights and parameters of radial basis functions.

Adjustable parameters of the neural network are following: a) number of neurons in the hidden layer (*m*); b) centers of radial basis functions (w_k^r); c) coefficients of smoothing (σ_k); d) types of basis functions of the hidden layer; e) the weight of output layer (w_{ik}^l); f) type of activation function of output layer; g) parameters of the activation function of output layer (*a*). Using selected parameters we can obtain the structure of individual AIS as shown in Fig. 3.

f_1	f_2	$\ldots$	f_m	w_{11}^r	$\ldots$	w_{1n}^r	$\ldots$	w_{m1}^r	$\ldots$	w_{mn}^r	σ_1	$\ldots$	σ_m	w_1^l	$\ldots$	w_m^l	a

Fig. 3. The structure of individual (antibody) AIS coding RBF-network

To encode the values by the binary system, the precision (bits per value) is highlighted as a parameter setting AIS. Elements of the string $f_1, ..., f_m$ encode the status of the neurons in the hidden layer [2]. The value of "0" corresponds to the passive or „off“ (the neuron is not involved in calculating of the output value network). The value of "1" shows that a neuron is active (enabled).

This scheme provides automatic search for the optimal number of hidden elements of the RBF network. Consider the example shown in Fig. 4.

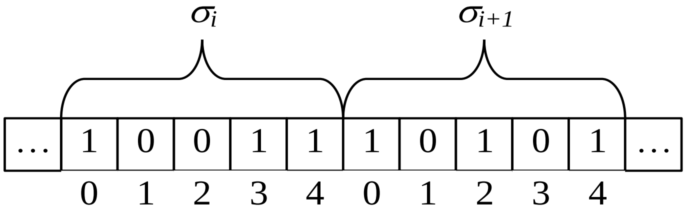


Fig. 4. Land-line antibody with binary encoding

In this example, mutation of the bits with indices 0 or 1 obviously cause more significant changes in the selected parameters than the mutation of the bit 4.

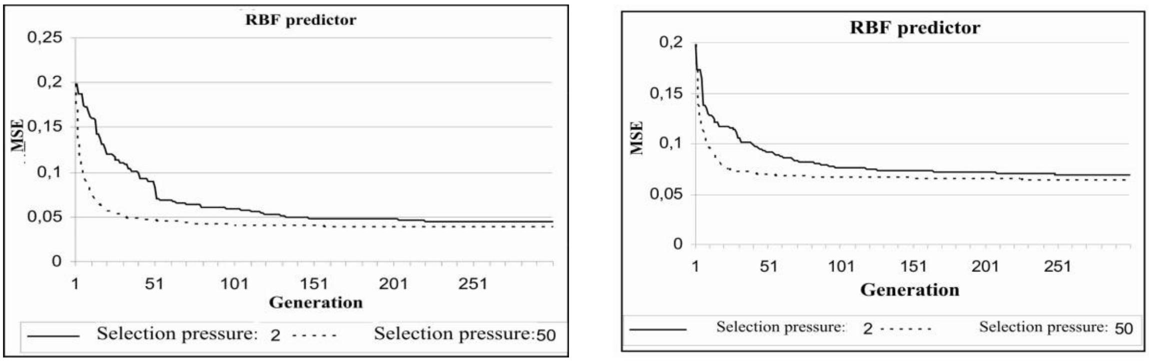
3. Experiments.

3.1 Description of experimental data.

For the pilot study, two time series representing real processes were selected. First is observations on the volume of monthly sales of tickets for American Airlines data for 12 years; the series contains 144 observations. The first 100 observations (70%) were used as a training sample, the remaining 44 (30%) were used as a test sample. The minimum time lag (immersion depth d) was calculated on the basis of the partial autocorrelation function (PACF). From the PACF values $d \geq 13$, we have chosen the value $d = 14$. The second data sample was daily consumption of electricity for one of the regions in Ukraine. To study the process we used the observations No. 260-330. The last 35 observations were used as a test sample. In this case the minimum time lag, $d \geq 15$, we have chosen the value of $d = 15$. Further on, based on initial data, we demonstrated experimentally the convergence of the developed methods and investigated influence of the main parameters on the AIS learning algorithms.

3.2 Influence of parameters on the convergence of AIS algorithms.

The experiments were set to investigate the influence of three main parameters of AIS: selection pressure, population size of clones, and the level of mutation. First, we set a high selection pressure equal to 50, the size of the population of clones equal to 300, and the level of mutation equal to 0.8. In this case, the selection of the best antibodies will occur as follows. After assessing the entire population, i.e. 50, the antibodies are selected by a certain percentage of the best of them for subsequent cloning. This percentage is set by parameter "factor of selection of the best antibodies", which in this case is 0.7. Thus, for cloning we must select 50 * 0,7 = 35, the best antibodies. The best antibodies (i.e., those which give the smallest error of approximation on the training data) are chosen by tournament selection. The tournament selection involves a random selection from the population, the number of antibodies, specified by the parameter "selection pressure", and the choice of one of the best out of this amount. The tournament is repeated as many times as antibodies should be selected (in this case 35 times). It follows that the larger the tournament is the less likely is a penetration of "weak" antibodies into the population of clones, and the faster the immune algorithm should converge. Then we set a minimal selection pressure equal to 2 and compared the results obtained. Fig. 5 (a) shows the graphs of convergence of the developed combined algorithm for the problem of forecasting the time series of ticket sales and Fig. 5 (b) shows the similar experiments with a number of observations of daily electricity consumption. The results show that: (1) developed algorithms converge to the minimum learning error, which proves the possibility of their use for solving the approximation problems; (2) parameter selection pressure can control the rate of convergence of algorithms that can be effectively used to prevent premature convergence to local optima.



(a) (b)

Fig. 5. (A) Convergence of combined algorithms at different levels of selection pressure: (a) for the monthly ticket sales; (b) for the energy consumption data

More studies have been performed touching on dependence of the time spent on learning and loss errors on the test sample from the population size of clones AIS. In these experiments, the rigidity of selection for all algorithms was set equal to 20; the experimental results are shown in Table 1.

Table 1.
A comparative study of the effect size of the population of clones

Algorithm type	Population size of clones	Data (time series)	Training time	PMSE on test sample
RBF-predictor	100	Sales tickets	35 c	12,65 %
RBF-predictor	300	Sales tickets	120 c	5,34 %
RBF-predictor	100	Electricity consumption	28 c	5,17 %
RBF-predictor	300	Electricity consumption	93 c	4,26 %

Here PMSE is the standard percentage error in percentage.

According to the results shown in Table 1, we can conclude that increase in population size slows down the learning process, on the other hand in improvement of the generated models quality. Furthermore, as seen from the table, combined algorithm IMS and wavelet neural network shows a higher accuracy of prediction than the combined algorithm and RBF network, while the latter has a higher rate of training.

The experimental results show that for a high level of mutation, which implies a high variability of clones of the population, in most cases the training is faster, but the step character of the curve indicates a low stability of the process, thus decreasing the probability of finding the global optimum. The figure shows that the value of errors corresponding to the high level of mutation decreases rapidly at the beginning of training. But it falls at a certain time in one of the local optima, and cannot leave it further on due to the high variability of antibodies, which leads to deterioration in the quality of training. In fact the mutation is the main driving force behind the evolution of the immune system and therefore requires more careful adjustment in accordance with the objectives set in the solution of the forecasting problem.

4. Conclusion.

The main results of this study are as follows:

– experimental analysis of the problem of finding the settings for RBF neural networks;

– the combined methods of forecasting time series with controlled parameters is proposed, based on the synthesis of RBF networks using artificial immune systems;

– the results of computational experiments showed higher effectiveness of the proposed combined methods.

References

1. Subbotin, S.O, Oleinik S.A, Oleinik O.O. Not iterative, evolutionary and multi-agent methods of synthesis of fuzzy models and neural networks: Monograph / Edited by S.O. Subbotina. – Zaporizhzhya: ZNTU, 2009. – 375 p. (UKR).
2. De Castro L. N., Timmis J. Artificial Immune Systems as a Novel Soft Computing Paradigm, Soft Computing Journal, vol. 7, Issue 7, July 2003.

UOT 004.042

P.İ. Bidyuk, A.S. Həsənov, S.H. Abdullayev

Süni imun sistemlərdən istifadə etməklə proqnozlaşdırılmış RBF neyron şəbəkələrin struktur və parametrik sintezi

Süni sistemlərin inkişafının nəticələrindən istifadə etməklə radial bazis funksiyalarının əsasında proqnozlaşdırılmış neyron şəbəkələrin struktur və parametrik sintez metodlarının işlənib hazırlanması tədqiqatın əsasını təşkil edir. Süni imun sistemlərin modifikasiyasına uyğun olaraq RBF əsasında neyron sistemlərin qurulması təyin olunmuşdur. Nəzarət olunan parametrlər modeli vasitəsi ilə zaman sıralarının proqnozlaşdırılması üçün kombinasiya edilmiş metod təklif olunmuşdur.Təyin edilmişdir ki, əhalinin artımı neyron şəbəkələrin öyrənmə prosesinin sürətinin azaldılması ilə yanaşı həmin modelin keyfiyyətinin yaxşılaşmasına gətirib çıxarır.Proqnozlaşdırılmanın kombinasiya olunmuş alqoritmi RBF şəbəkə ilə proqnozlaşdırılmadan fərqli olaraq Wavelet əsasında neyron şəbəkələr ilə proqnozlaşdırılma daha yaxşı keyfiyyətə malikdir. Müəyyən olunmuşdur ki, daha yüksək səviyyəli mutasiya zamanı öyrənmə prosesinin daha sürətlə həyata keçirilir, lakin təlimin müqavimətinin azaldılması isə qlobal optimumun müəyyən edilmə ehtimalının azaldılmasina gətirib çıxarır.

Açar sözlər: süni imun sistemlər, neyron şəbəkələr, radial bazis funksiyaları, proqnozlaşdırılmış

УДК 004.042

П.И. Биджук, А.С. Гасанов, С.Г. Абдуллаев

Структурный и параметрический синтез прогнозирующих нейронных сетей RBF, используя искусственные иммунные системы

Целью исследования является разработка метода структурного и параметрического синтеза прогнозирующей нейронной сети на основе радиальных базисных функций с использованием результатов развития искусственных систем. Определены настройки для нейронной сети на основе РБФ с помощью соответственно модифицированной искусственной иммунной системы. Предложен комбинированный метод для прогнозирования временных рядов с контролируемыми параметрами модели. Установлено, что увеличение размера популяции снижает быстродействие процесса обучения нейронной сети, но приводит к улучшению качества модели. Комбинированный алгоритм прогнозирования с нейросетью на основе вейвлетов показал лучшее качество прогнозирования чем с сетью РБФ. Определено, что при более высоких уровнях мутации, обучение происходит быстрее, а устойчивость обучения снижается, что уменьшает вероятность определения глобального оптимума.

Ключевые слова: искусственные иммунные системы, нейронные сети, радиальные базисные функции, прогнозирование

Institute of Information Technology, Azerbaijan NAS Presented 08.10.2014