World Essays Journal / 4 (1): 30-34, 2016

System Identification by Using Optimized Neural Networks based on Gray Wolf Optimizer Algorithm

Mehdi Ramezani, Mahmood Mohseni Motlagh*

Department of Electrical Engineering, University of Tafresh, Tafresh, Iran.

Corresponding author email: mehdiramezani.ir@gmail.com

Abstract: In this paper, a learning algorithm for Multilayer Perceptron neural networks based on gray wolf optimizer algorithm is proposed. Generally, Multilayer Perceptron (MLP) is known as one of the most popular artificial neural networks model to perform classification task with considerable success. Although, because of the MLP structure complexities and also other problems like local minima trapping, weight interference and over fitting have made neural network training difficult. In this research, the optimized ANN approach is employed to eliminate the considered difficulties and utilize it to estimate the system parameters. GWO is used to train the wavelet parameters and the connections. Results show that the proposed technique can handle the system identification problems.

INTRODUCTION

Recently, Artificial Neural Networks are known as successful tools to classifying a variety of real world like business, science and industry (G. P. Zhang, 2000; Liao and Wen, 2007).

In classification, ANNs require to train properly to be able to generate the desired classifier. In training step, some numbers of examples are trained to the network and then the connection weights of the network are tuned by utilizing the learning algorithm. The main purpose in weights tuning is to learn the network so that network would tuned to the given training data. Multilayer feedforward network (MLP) is the most popular architecture of ANNs. MLP employ a supervised learning approach called Backpropagation (BP) to train the network. Albeit, because of the multi-layered structure, the training speeds are generally very slower as compared to other single layer feed-forward networks (Patra and Pal, 1995).

Local minima trapping, weight interference and overfitting are some problems of the network training in MLP structures which are become to a challenging category (Dehuri and Cho, 2010). Pao and Takefuji (1992) have presented Functional Link Neural Network (FLNN) as an alternative approach to avoid the mentioned problems. In this technique, the hidden layer has been removed from the ANN architecture to reduce the neural architectural complexity and to provide them with an improved representation of the input nodes for the network to be able to apply a non-linear separable classification task (Pao and Takefuji 1992).In the last years, there are introduced several different techniques to identify the linear or non-linear system, and these algorithms employ different structures of knowledge about the system. System Identification is the process of extracting a mathematical model of a physical process using observations. In the other words, it can be considered as the process of mathematical representation or improving the physical system using experimental data. System identification is a method to utilize the output data or both input and output data (Soderstrom and Stoica, 1989). In system modeling three main principles have to be considered such as separation, parsimony and selection (Ljung, 1999).

Applications of system identification in different areas like control, communication, power system and instrumentation is a key factor for achieving a model of the considered physical plant or a new system to be improved. In the system identification, a definite system is identified based on input output data samples. The identification task is to determination of a considered system to estimate it by a mathematical model which has similar characteristics like the main system. A good estimation depends on comparison between the actual output sample and the desired value due to input data up to that instant.

Recently, Artificial Neural Networks are used for modeling complicated systems and are illustrated good results on it (Guyer and Yang, 2000).Multi-layer percepton (MLP) is one of the most popular neural network model in which the connection weight training is generally completed by a back propagation learning algorithm (Rumelhart et al., 1986).

For improving the potency of these network structures and their parameters, some evolutionary algorithms like: Genetic Algorithm (GA) (Qu et al., 2008), Back Propagation (BP) (Tang and Xi, 2008), Pruning Algorithm (Reed, 1993), Simulated Annealing (Souto et al., 2002) and Particle Swarm Optimization (zhang et al., 2002) are analyzed.The considered optimization algorithms can improve the neural network efficiency by different changing in levels like weight training, architecture adaptation and learning rules (zhang et al., 2002).

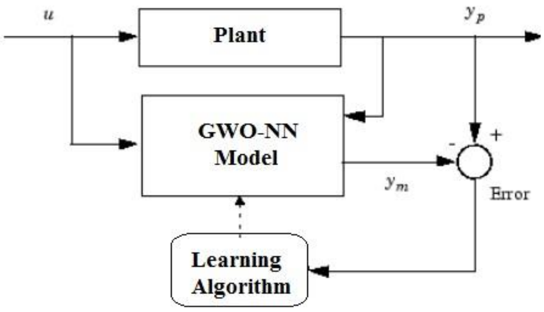

Figure 1. Proposed System Identification Process

In this paper, we describe an overview of neural network and the proposed Gray Wolf Optimizer (GWO) as learning algorithm for achieving better classification capability.

Artificial neural network

The brain can be mathematically simulated by an artificial neural network (ANN) as the programming equivalent. ANN attempts to model the function of a brain by imitating the layout. An ANN comprise of some interconnected neurons.

All of the neurons connections have a particular weight which affects on how much the output from the neuron will change the input to the next neuron. Generally, each neuron has a weight of its own called a bias term which characterizes the considered neuron's influence. ANN stores the information in its weights by training; therefore, it needs to determine the weights. To do this end, several algorithms have been introduced.

One of the common methods for this purpose is back propagation which is used for feed forward networks. In this technique, the error value is evaluated for each of the training pairs and tunes the weights to fit the desired output. This process will be continued in several iterations until the total error on the training set become small enough or when the error value don’t continue to decline. In fact, ANNS benefit supervised learning where the networks are trained by using data for which inputs as well as target outputs are noticed (Hagan et al., 1996).

After training, the network and its weights are ready to use and calculate the output values for new given samples. Back propagation is a gradient descent based technique which is applied on the error space and most likely gets trapped into a local minimum; since, its successful depends on initial (weight) settings. Recently, there are several techniques to compensate this shortcoming. Among these methods, the evolutionary based techniques have been found as a good approach by a proper efficiency. In this paper, a new optimization algorithm, GWO, is utilized for this target.

Gray Wolf Optimizer

Gray Wolf Optimizer (GWO) is a new optimization algorithm proposed by Seyedali et al. (2013). GWO imitates the hunting process of gray wolves in the wildlife. Generally, wolves live in groups which divide into two parts: gray wolves (male and female) managing the other wolves in the pack. By considering the (Mirjalili et al., 2013), the social hierarchy of the pack can be organized as below:

The alphas wolves (α): The leading wolves in the group. They have a duty to make decisions. The alphas orders are forced to the others.

The betas wolves (β): They comprise the second level of wolves after alphas. The main duty of betas wolves is to help and support alphas decisions.

The deltas wolves (δ): these wolves comprise the third level in the wolves social. Deltas wolves used to follow alpha and beta wolves. The delta wolves have 5 categories which can be summarized as below:
Scouts: these wolves monitor and control the boundaries of the zone and alert the pack in case of danger.
Sentinels: the wolves who protect and ensure the safety of the wolves' society.
Elders: these wolves comprise strong wolves which may be used to be alpha or beta wolves in the future.
Hunters: these wolves utilized to back up alpha and beta for hunting and providing food the pack.
Caretakers: these wolves are responsible to take care of the ill, wounded and weak wolves.

The omegas wolves (ω): the lowest levels in the wolves pack which have to follow alpha, beta and delta wolves. The omegas wolves are the last wolves that are allowed to eat.

In GWO, The first purpose of hunting a prey is to circle it by α, β and ω which can be mathematically modeled as below:

$$X\ (\mathrm{t}+1) = \mathrm{X}_p(\mathrm{t}) + \mathrm{A}.\mathrm{D} \qquad (1)$$

Here ***X*** represent the gray wolf position. the number of iteration is defined by *t*. ***X*** *p* is prey position and ***D*** is:

$$D = \left| C.X_P(\mathrm{t}+1) - \mathrm{X}(\mathrm{t}) \right| \qquad (2)$$

and *C* can be obtained by the equation below:

$$A = 2a.r_1 - a \qquad (3)$$

$$C = 2r_2$$

where *a* is a linearly deduced from 2 to 0 through the number of iterations and is employed for controlling the trade-off between exploitation and exploration. For updating the value of variable:

$$a = 2 - t(2/\mathrm{NumIter}) \qquad (4)$$

$$X\ (\mathrm{t}+1) = (\mathrm{X}_1 + \mathrm{X}_2 + \mathrm{X}_3)/3 \qquad (5)$$

Here *NumIter* defines the number of iterations. r_1 and r_2 are random vectors between [0, 1] which are employed to find the optimal solution. Good knowledge about the potential location of prey can be achieved by Alpha, Beta and Delta where they help the Omega to follow the suitable positions. The values of X_1, X_2 and X_3 are illustrated in the equations below:

$$\mathrm{X}_1 = \left| \mathrm{X}_\alpha - A_1 D_\alpha \right| \qquad (6)$$

$$\mathrm{X}_2 = \left| \mathrm{X}_\beta - A_2 D_\beta \right| \qquad (7)$$

$$\mathrm{X}_3 = \left| \mathrm{X}_\delta - A_3 D_\delta \right| \qquad (8)$$

In iteration *t,* the best 3 solutions are: X_1, X_2 and X_3. Here D_α, D_β and D_δ are:

$$D_\alpha = \left| C_1 X_\alpha - X \right| \qquad (9)$$

$$D_\beta = \left| C_2 X_\beta - X \right| \qquad (10)$$

$$D_\delta = \left| C_3 X_\delta - X \right| \qquad (11)$$

Gray Wolf Optimizer of Neural Networks (ANN-GWO)

Here, GWO is hybridized by ANN for achieving a precise ANN and for removing the drawback of the back-propagation algorithm. Te approach includes two important steps: the first step is to train ANN using GWO. Here, GWO is utilized for finding the optimal initial weights. The second step is to test the results of the proposed GWO-ANN approach. This idea can improve the efficiency of the back-propagation to seek global optima in the search space. The weights are achieved as a vector of variables for the proposed GWO-ANN method.

The cost function in here is based on the Root Mean Square Error (RMSE), which finds the error between the real input and the predicted output. RMSE is presented in the equation below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(T_i - P_i)^2} \qquad (12)$$

where *Ti* and *Pi* are the target output and predicted value from ANN. A lower value of *RMSE* indicates a better model.

SIMULATION RESULTS

For identification of a considered system, we have proposed here a GWO based Neural Network with supervised learning method. Here, we have considered a system as:

$$H\ (\mathrm{s}) = \frac{0.01}{s^2 + 0.5s + 2} \qquad (13)$$

This can be represented in the state space as:

$$A=\begin{bmatrix} 0 & 1 \\ -2 & -0.5 \end{bmatrix}; \qquad (14)$$

$$b=\begin{bmatrix} 0 \\ 0.01 \end{bmatrix};$$

c=[1 0];

d=0;

For identifying the considered system, two hidden layer with 10 neurons are selected. Here, sigmoid function is utilized as the activation function. The weights are initialized and updated in each iteration by GWO algorithm.

In Fig. 1, the estimation error surf of the network using the proposed approach is shown. A better view of the error can be shown in the fig. 2. As it can be seen from figure 2, the error value in 25 iterations is reduced from 100 into 1. The step response of the proposed technique and the original system for 25 iterations are shown in fig.3. As it can be seen, the error value for the step response of the proposed GWO-ANN is small rather than the original system and the proposed algorithm shows a good efficiency to this end.

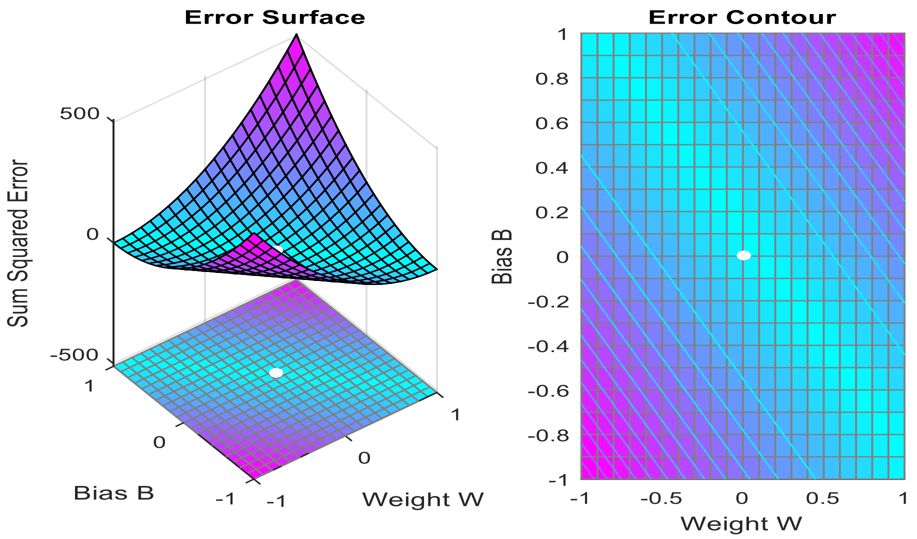

Figure 2. Error surf of the trained GWO-NN

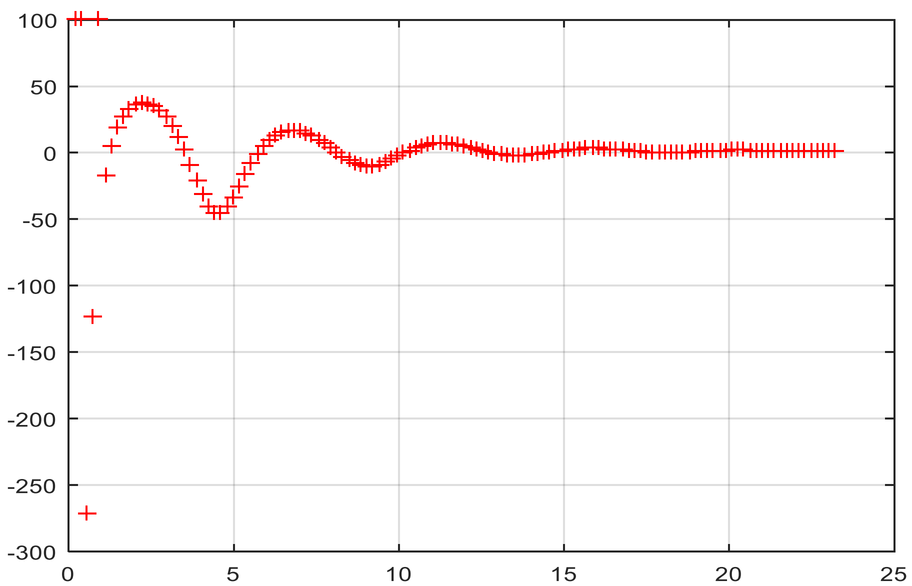

Figure 3. Error value of the proposed method for 25 iterations

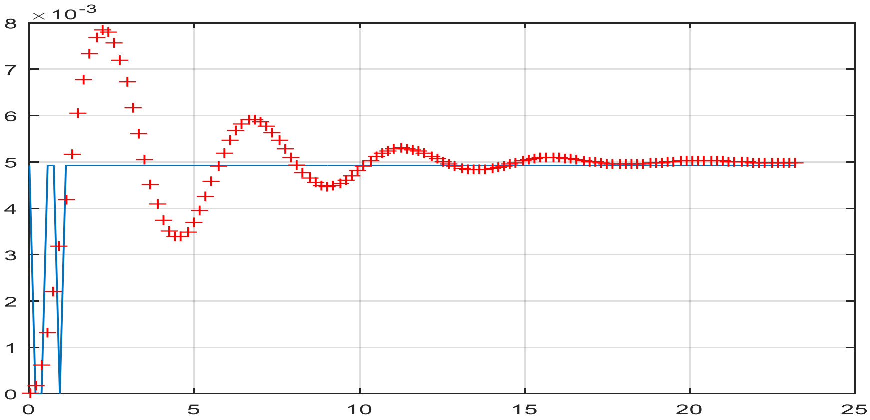

Figure 4 . Step response of the original system (blue-line) and the identified model (red-plus points)

CONCLUSION

In this paper, gray wolf optimizer as a new evolutionary algorithm has been proposed and utilized to design an optimized neural network for system identification purposes. In the proposed algorithm, the network weights are selected by GWO to achieve the minimum value to the root mean square error. Indeed, GWO is used to train the neural network's parameters and the connection switches. The proposed GWO-ANN can handle identification problems.

REFERENCES

C.zhang, H.Shao, Y.Li, "Particle swarm optimization for evolving artificial neural network", 2000 IEEE Intl. Conf. on Systems, Man, and Cybernetics, Vol. 4, pp.2487 – 2490, Oct. 2000.

D. E. Rumelhart, G. E .Hinton, and R.J. Williams, "Learning representations by backpropagating errors", Nature, 323: 533-536, 1986.

D. Guyer, X. Yang, "Use of genetic artificial neural networks and spectral imaging for defect detection on cherries", Computers and Electronics in Agriculture, vol. 29, pp. 179-194, 2000.

G. P. Zhang, "Neural networks for classification: a survey," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 30, pp. 451-462, 2000.

Hagan, M.T., Demuth, H.B. and Beale, M. (1996) Neural Network Design. PWS Publishing Co., Boston.

J. C. Patra and R. N. Pal, "A functional link artificial neural network for adaptive channel equalization," Signal Processing, vol. 43, pp. 181- 195, 1995.

L. Ljung, “System Identification - Theory for the User”, Prentice-Hall, N.J. 2nd edition, 1999.

M.C.P.de Souto, A.Yamazaki, T.B.Ludernir, "Optimization of neural network weights and architecture for odor recognition using simulated annealing", Proc. 2002. Intl. Joint Conf. on Neural Networks, vol.1, 2002, p.547– 552.

Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014) Grey Wolf Optimizer. Advances in Engineering Software, 69, 46-61. http://dx.doi.org/10.1016/j.advengsoft.2013.12.007

P.Tang and Z. Xi, "The Research on BP Neural Network Model Based on Guaranteed Convergence Particle Swarm Optimization", Second Intl. Symp. on Intelligent Information Technology Application, IITA '08, Vol.2, pp.13 – 16, Dec. 2008.

R.Reed, "Pruning algorithms—a survey", IEEE Trans. Neural Networks Vol.4, pp.740–747, 1993.

S. Dehuri and S.-B. Cho, "A comprehensive survey on functional link neural networks and an adaptive PSO–BP learning for CFLNN," Neural Computing & Applications vol. 19, pp. 187-205, 2010.

S.-H. Liao and C.-H. Wen, "Artificial neural networks classification and clustering of methodologies and applications – literature analysis from 1995 to 2005," Expert Systems with Applications, vol. 32, pp. 1-11, 2007.

Soderstrom, T. and Stoica, P.”System Identification”, Prentice Hall, 1989.

X. Qu, J. Feng, W. Sun, "Parallel Genetic Algorithm Model Based on AHP and Neural Networks for Enterprise Comprehensive Business", IEEE Intl. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, pp.897-900, 2008.

Y. H. Pao and Y. Takefuji, "Functional-link net computing: theory, system architecture, and functionalities," Computer, vol. 25, pp. 76-79, 1992.