Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 10 (3): 221-225, 2010
ISSN 1812-5654

Optimization of the Artificial Neural Networks Using Ant Colony Algorithm to Predict the Variation of Stock Price Index

R. Tehrani and F. Khodayar
Faculty of Management, Tehran University, Jalale Ale Ahmad Avenue, Tehran, Iran

Abstract: This study proposes Ant Colony Optimization (ACO) algorithms approach to determine the connection weights for Artificial Neural Networks (ANNs) and predict the stock price index. Earlier research proposed many varied models of ANN for the method of training the network, feature subset selection and topology optimization. In most of these studies, the optimum weights are not used to improve the learning algorithm. In this study, ACO algorithm is employed not only to improve the learning algorithm, but also to reduce the complexity in feature space. The ACO algorithm optimizes the connection weights between layers in neural network. This method decreases the limitations of the gradient descent algorithm. Experimental results show that ACO algorithm approach to the optimum model in compare to the other conventional models.

Key words: Ant colony algorithm, artificial neural network, optimization, stock price index

INTRODUCTION

Predicting the stock price index is not a simple task. Atienza *et al*. (2000) has been used different techniques in the trading community to predict the tasks. In recent years the concept of neural networks has emerged as one of them.

Among them, there are many studies using data mining techniques including Artificial Neural Networks (ANNs). A neural network is able to work parallel with input variables and consequently handle large sets of data swiftly. The principal strength with the network is its ability to find patterns and irregularities as well as detecting multi-dimensional non-linear connections in data. Zimmermann *et al*. (2000) proved that the latter quality is extremely useful for modeling dynamical systems.

Liu and Setiono (1996) showed that ANN had some limitations in learning the patterns because stock market data has tremendous noise and complex dimensionality. The ANN has preeminent learning ability while it is often confronted with inconsistent and unpredictable performance for noisy data. In addition, sometimes the amount of data is so large that the learning of patterns may not work well.

Many researchers are interested to reduce the data dimension. The reduction and transformation of the irrelevant or redundant data may shorten the running time and yield more generalized results (Zimmermann, 2003). This study proposes a new hybrid model of ANN and ACO algorithm to decrease the above limitations. This study uses ACO to search the optimal or near-optimal thresholds and searches the connection weights between layers in ANN.

Ant Colony Optimization (ACO) is a pattern solution for combinatorial optimization problems. Dorigo *et al*. (1991) and DiCaro and Dorigo (1998) presented the first algorithms which can be classified with in this framework and, since then, many diverse variants of the basic principle have been reported in the literature.

The characteristic of ACO algorithms is their explicit use of elements of previous solutions. In fact, they drive a constructive low-level solution. But including it in a population framework and randomizing the construction in a Monte Carlo way. A Monte Carlo combination of different solution elements is suggested also by Holland (1992), but in the case of ACO, the probability distribution is explicitly defined by previously obtained solution components.

In this study, an ANN model is proposed to predict Stock Price index. The connection weights in ANN are optimized using ACO. The simulation results show that the proposed model is more accurate and reliable.

ARTIFICIAL NEURAL NETWORK MODEL

Many ANN studies relied on the gradient descent algorithm to get the connection weights of the model. Sexton *et al*. (1998) pointed out that the gradient descent algorithm may perform poorly even on simple problems

Corresponding Author: Reza Tehrani, Faculty of Management, Tehran University, Jalale Ale Ahmad Avenue, Tehran, Iran
Tel: +98 21 61117675

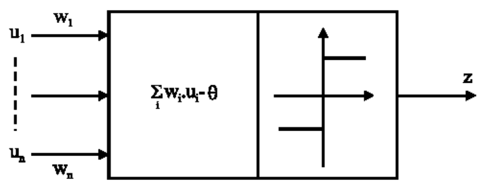

Fig. 1: The artificial neuron with a threshold function

when predicting the holdout data. Their indication stems from the fact that back propagation is a local search algorithm and may tend to fall into a local minimum.

In mathematical terms the following equations gives a dense description of the neuron:

$$y = \sum_{i=1}^{n} \omega_i u_i - \theta \tag{1}$$

And

$$z = \psi(y) \tag{2}$$

where, y and Ψ are the net input and activation function respectively.

A simple diagram of a neural network is shown in Fig. 1. In neural computing almost exclusively three different types of activation functions are being used:

- The threshold function

$$\psi(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{3}$$

- The piecewise linear function

$$\psi(x) = \begin{cases} 1 & x \geq \frac{1}{2} \\ x & \frac{-1}{2} < x < \frac{1}{2} \\ 0 & x \leq \frac{-1}{2} \end{cases} \tag{4}$$

- Sigmoid function

$$\psi(x) = \frac{1}{1 + e^{-\alpha x}} \tag{5}$$

where, α controls the slope.

ANT COLONY OPTIMIZATION (ACO)

Dorigo *et al.* (1996) simulate the behavior of real ants using ACO algorithms. They are based on the principle

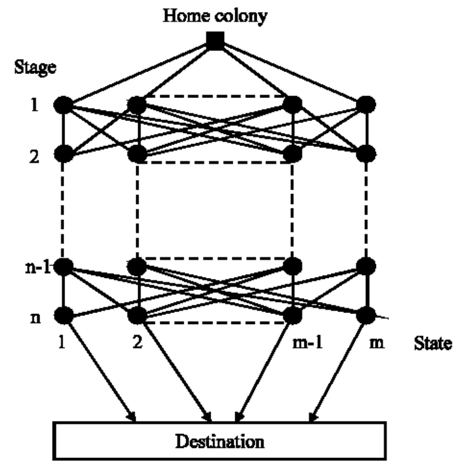

Fig. 2: Search space for an optimization problem

that using simple communication mechanisms, an ant group is able to find the shortest path between any two points. During their trips, a chemical trail (pheromone) is left on the ground. The pheromone guides other ants toward the target point. For one ant, the path is chosen according to the quantity of pheromone. The pheromone evaporates over time. If many ants choose a certain path and lay down pheromones, the quantity of the trail increases and thus, this trail attracts more and more ants. The artificial ants simulate the transitions from one point to another point, according to the improved version of ACO, namely the Max-Min AS (MMAS) algorithm presented by Stutzle and Hoos (1998), as follows:

The search space of the ant colony has been shown in Fig. 2. The ant k maintains a black list (N_r^k) in memory that defines the set of points still to be visited when it is at point r. The ant chooses to go from point r to point s during a tour with a probability given by Stutzle and Hoos (1998).

$$p(r,s) = \frac{\gamma(r,s)}{\sum_{l} \gamma(r,l)} \quad s, l \in N_l^k \tag{6}$$

where, matrix represents the amount of the pheromone trail (pheromone intensity) between points r and s.

Then, the pheromone trail on coupling is updated according to:

$$\gamma(r,s) = \alpha.\gamma(r,s) + \Delta\gamma^k(r,s) \tag{7}$$

where, α with $0<\alpha<1$ is the persistence of the pheromone trail, so that $(1-\alpha)$ to represent the evaporation and is the amount of pheromone that ant k puts on the trail. The pheromone update $\Delta\gamma^k(r, s)$ reflects the desirability of the trail (r, s), such as shorter distance, better performance, etc., depending on the application problem. Since, the best tour is unknown initially, an ant needs to select a trail randomly and deposits pheromone in the trail, where the amount of pheromone will depend upon the pheromone update rule. The randomness implies that pheromone is deposited in all possible trails, not just in the best trail. The trail with favorable update, however, increases the pheromone intensity more than other trails. After all ants have completed their tours, global pheromone is updated in the trails of the ant with the best tour executed. In the next section, the MMAS algorithm is extended and modified to find optimal or near optimal connection weights and thresholds in the ANN.

ANN OPTIMIZATION USING ACO ALGORITHM

Many fund managers and investors in the stock market generally accept and use certain criteria for technical indicators as the signal of future market trends. Even if a feature represents a continuous measure, the experts usually interpret the values in qualitative terms such as low, medium and high. The reasoning process of ANN may be like that of the human experts. The decreased data can simplify the process of learning and may improve the learned results because it may effectively reduce the noisy and redundant data. The block diagram of the back propagation neural network (BPLT) is shown in Fig. 3.

This study proposes the optimized thresholds based on ACO algorithm may find optimal or near-optimal thresholds for maximum predictive performance because ACO searches the optimal or near-optimal parameters to maximize the fitness function.

The block diagram of ant colony optimization of ANN is shown in Fig. 4. The optimization algorithm consists of three phases. In the first phase, ACO searches optimal or near optimal connection weights and thresholds. The connection weights and the thresholds are initialized into random values before the search process. This study needs three sets of parameters. The first set is the set of connection weights between the input layer and the hidden layer of the network. The second set is the set of connection weights between the hidden layer and the output layer. As mentioned earlier, the above two sets may mitigate the limitation of the gradient descent algorithm. The third set represents the thresholds. A fully connected feed-forward network with one hidden layer and one output shown in Fig. 5.

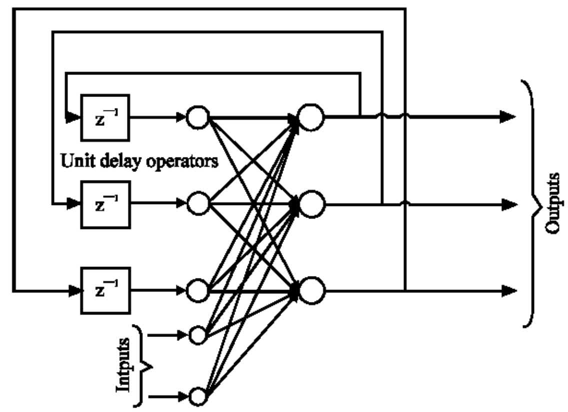

Fig. 3: Block diagram of the back propagation neural network (BPLT)

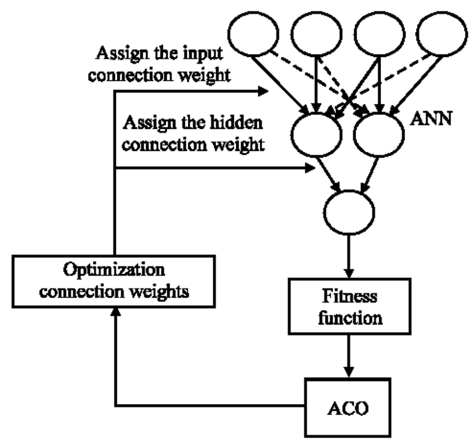

Fig. 4: Block diagram of ant colony optimization of ANN

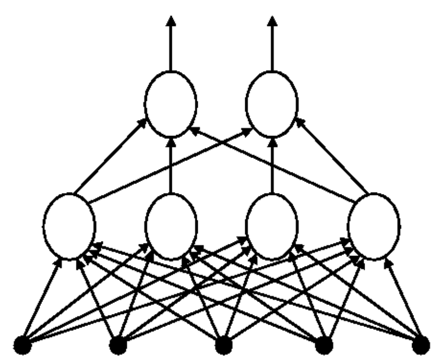

Fig. 5: A fully connected feed-forward network with one hidden layer and one output

RESEARCH DATA AND EXPERIMENTS

The research data used in this study is technical indicators and the direction of change in the daily Iran

Table 1: Technical parameter use for stock price prediction

Parameter No.	Definition	Note
1	$DOA_{1,30}[y(t)]/MA_{30}[y(t)]$	Difference of averages in 30 day observational window and normalized with 30 days moving average
2	$MA_3[ROC_{15}[y(t)]]$	3 week rate of change, averaged with days moving average
3	$y(t)/MA_5[y(t)]$	5 days disparity
4	$RSI_{15}[y(t)]$	3 week relative strength index
5	$\sigma_{30}[y(t)]$	Monthly standard deviation
6	$y(t)/MA_{10}[y(t)]$	10 days disparity
7	William $\%R_{15}[y(t)]$	William indicator with 15 days window
8	$DL_5[y(t)]$	Lagged historical value
9	$DL_5[y(t-1)]$	
10	$DL_5[y(t-2)]$	
11	$DL_5[y(t-3)]$	
12	$DL_5[y(t-4)]$	
13	$RDP_{-5}[y(t)]$	Stock index excess return
14	$RDP_{-10}[y(t)]$	
15	$RDP_{-20}[y(t)]$	

Stock Price Index (IRSPI). The total number of samples is 2428 trading days, from July 1995 to September 2009.

We select 15 technical parameters as feature subsets by the review of domain experts and previous researches. Table 1 shows the summary statistics for each feature which are presented by Abdullah and Ganapathy (2000) and Ramirez *et al*. (2004). This study compares linear transformation with the back propagation neural network (BPLT) to the linear transformation with ANN trained by ACO algorithm. In this study, linear transformation means the linear scaling of data into the range of 0.0-1.0.

Linear transformation is generally used to enhance the performance of ANN because most ANN models accept numeric data only in the range from 0-1 or 21-11.

Back propagation neural network (BPLT) uses the gradient descent algorithm to train the network. This is the conventional approach of earlier studies. The number of processing elements in the hidden layer is fixed at 12. This is like the number of feature subsets.

About 20% of the data is used for holdout and 80% for training. The training data is used to search the optimal or near-optimal parameters and is employed to evaluate the fitness function. The holdout data is used to test the results with the data that is not utilized to develop the model.

SIMULATION RESULTS

Linear transformation with the back propagation neural network (BPLT) and the linear transformation with ANN trained by ACO algorithm is simulated using Matlab software. Four auxiliary parameters have been defined to calculate the performance of each model. Table 2 shows these auxiliary parameters which are presented by Ramirez *et al*. (2004).

To compare the performance of the models, we calculate the below parameters:

$$A = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

Table 2: Definition of the auxiliary parameters

R	Auxiliary parameters	Definition
1	TP	No. of positive class predicted correctly as positive class
2	FP	No. of negative class predicted wrongly as positive class
3	FN	No. of positive class predicted wrongly as negative class
4	TN	No. of negative class predicted correctly as negative class

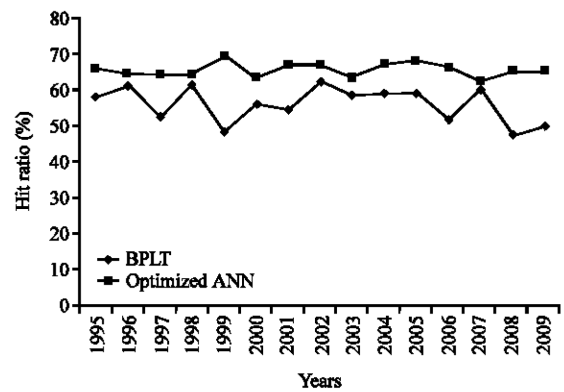

Fig. 6: The average predictive accuracy for BPLT and optimized ANN models

$$A = \frac{TP}{TP + FP} \tag{9}$$

$$A = \frac{TP}{TP + FN} \tag{10}$$

where, A is accuracy, P is precision and R is recall rate or sensitivity.

Two models are compared according to the methods of determining the connection weights and feature transformation. Figure 6 shows the average prediction accuracy of each model.

In Fig. 6, the optimized model has higher predictive accuracy than BPLT by 8≈11% for the training data. It is a mistake to compare the prediction accuracy between the training data and holdout data. There is a wide difference between the training data and the holdout data for the two

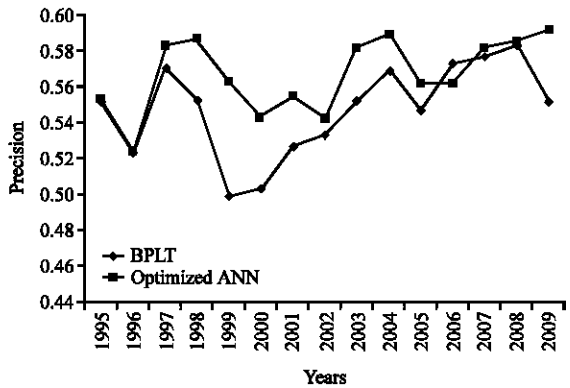

Fig. 7: The average predictive precision for BPLT and optimized ANN models

models. Figure 7 shows the average predictive precision for BPLT and optimized ANN models. The optimized model has higher predictive precision than BPLT by 5≈8% for the training data.

The simulation results show that optimized model with ant colony algorithm is more accurate and reliable. The performance of the optimized model in compare to conventional neural network models is higher by 9≈13% for the training data.

CONCLUSION

In this research, a novel ANN model is proposed and simulated. The connection weights and thresholds are optimized with ant colony algorithm. The ACO searches for the optimal or near-optimal solutions of connection weights in the learning algorithm. The research data used in this study is technical indicators and the direction of change in the daily Iran Stock Price Index (IRSPI). Four auxiliary parameters were defined to calculate the performance of models. The performance of the optimized model is higher than BPLT and conventional models. The optimized model has higher prediction accuracy than BPLT by 8≈11% for the training data.

ACKNOWLEDGMENT

The authors would like to acknowledge the active participation and financial support of the Management Faculty of Tehran University.

REFERENCES

Abdullah, M.H.L.B. and V. Ganapathy, 2000. Neural network ensemble for financial trend prediction. Proc. TENCON, 3: 157-161.

Atienza, F., N. Martinez-Alzamora, J.A. De Velasco, S. Dreiseitl and L. Ohno-Machado, 2000. Risk stratification in heart failure using artificial neural networks. Proc. AMIA Symp., 2000: 32-36.

DiCaro, G. and M. Dorigo, 1998. AntNet: Distributed stigmergetic control for communications networks. J. Artif. Intell. Res., 9: 317-365.

Dorigo, M., V. Maniezzo and A. Colorni, 1991. The ant system: an autocatalytic optimizing process. Technical Report TR91-016, Politecnico di Milano.

Dorigo, M., V. Maniezzo and A. Colorni, 1996. The ant system: Optimization by a colony of cooperating agents. IEEE Trans. Syst. Man Cybernet., 26: 29-41.

Holland, J.H., 1992. Adaptation in Natural and Artificial System. 1st Edn., The MIT Press, USA., ISBN-10: 0262581116.

Liu, H. and R. Setiono, 1996. Dimensionality reduction via discretization. Knowledge-Based Syst., 9: 67-72.

Ramirez, D.A., E.I. Truijillo, M. Juan and G. Gomez, 2004. Choosing variables with a genetic algorithm for econometric models based on neural networks learning and adaptation. Comput. Econ. Finance.

Sexton, R.S., B. Alidaee, R.E. Dorsey and J.D. Johnson, 1998. Global optimization for artificial neural networks: A tabu search application. Eur. J. Operat. Res., 106: 570-584.

Stutzle, T. and H.H. Hoos, 1998. Improvements on the Ant System: Introducing the MAX-MIN Ant System. Springer Verlag, New York, pp: 245-249.

Zimmermann, H.G., R. Grothmann and R. Neuneier, 2000. Forecasting high-dimensional dynamical systems by error correction neural networks with a focus on the german yield curve. Research Paper, Siemens AG, Munich.

Zimmermann, H.G., 2003. System identification and forecasting with advanced neural networks principles, techniques, applications. http://www.idsia.ch/~juergen/zimmermann.pdf.