Article

ANFIS Based Time Series Prediction Method of Bank Cash Flow Optimized by Adaptive Population Activity PSO Algorithm

Jie-Sheng Wang * and Chen-Xu Ning

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114044, Liaoning, China; E-Mail: NCX710754167@163.com

* Author to whom correspondence should be addressed; E-Mail: wang_jiesheng@126.com or wangjiesheng@ustl.edu.cn; Tel.: +86-412-253-8355; Fax: +86-412-253-8244.

Academic Editor: Willy Susilo

Received: 12 May 2015 / Accepted: 19 June 2015 / Published: 24 June 2015

Abstract: In order to improve the accuracy and real-time of all kinds of information in the cash business, and solve the problem which accuracy and stability is not high of the data linkage between cash inventory forecasting and cash management information in the commercial bank, a hybrid learning algorithm is proposed based on adaptive population activity particle swarm optimization (APAPSO) algorithm combined with the least squares method (LMS) to optimize the adaptive network-based fuzzy inference system (ANFIS) model parameters. Through the introduction of metric function of population diversity to ensure the diversity of population and adaptive changes in inertia weight and learning factors, the optimization ability of the particle swarm optimization (PSO) algorithm is improved, which avoids the premature convergence problem of the PSO algorithm. The simulation comparison experiments are carried out with BP-LMS algorithm and standard PSO-LMS by adopting real commercial banks’ cash flow data to verify the effectiveness of the proposed time series prediction of bank cash flow based on improved PSO-ANFIS optimization method. Simulation results show that the optimization speed is faster and the prediction accuracy is higher.

Keywords: time series prediction; bank cash flow; adaptive network-based fuzzy inference system; particle swarm optimization algorithm
1. Introduction

In recent years, time series modeling and prediction are one of the most active research topics in academic research and engineering practice [1–2]. Time series modeling is usually a chronological series of observed data (information) according to the time sequence, whose values are sampled at invariable time intervals. Researchers often predict future changes based on the historical data. For example, according to the situation in the past or the current period of the market sales, changes of stock prices, population growth and the bank’s deposits and withdrawals in the future are predicted. Time series forecasting affects the life of people everywhere, so it has an important practical significance and research prospects in every field of today’s society, which is also an important direction in the computer application field.

The bank cash flow forecasting management information system is designed to create a system management platform for the prediction and analysis of commercial bank cash flow. It will realize the cash flow data statistics summary, the cash flow short-term and long-term predictions, and the management information related to commercial bank cash flow under three levels: secondary branches (Cash Operation Center), branch (Business Library) and Network. Its purpose is to provide effective data of all levels of organization to analyze and assess cash business operation conditions. It will also provide effective system management means for the cash operation managers and decision-making people at all levels.

Artificial neural network (ANN) is a very good approximation method, which has characteristics of adaptive and self-learning [3–4]. However, ANN is easy to fall into local minimum. Combined with the fuzzy inference system, a new kind of nonlinear prediction method was proposed, namely: adaptive neural fuzzy inference system (ANFIS) [5]. This method can use both fuzzy rules and the structure of the neural network to realize adaptive self-learning, thus the prediction accuracy is higher than the single artificial neural network. In order to further improve the prediction precision of the adaptive ANFIS system, the PSO algorithm is also applied to optimize its structure parameters. A new hybrid approach, combining particle swarm optimization and adaptive-network-based fuzzy inference system for short-term wind power prediction in Portugal is proposed, forecasting accuracy is attainable using the proposed approach [4–6]. The radial basis function neural network (RBFNN) with a nonlinear time-varying evolution particle swarm optimization (NTVE-PSO) algorithm is developed, and Simulation results illustrate that the proposed NTVE-PSO-RBFNN has better forecasting accuracy and computational efficiency for different electricity demands [7–10]. An improved PSO-based artificial neural network (ANN) is developed, the results show that the proposed SAPSO-based ANN has a better ability to escape from a local optimum and is more effective than the conventional PSO-based ANN [11–14]. A training algorithm is based on a hybrid of particle swarm optimization (PSO) and evolutionary algorithm (EA) to predict the 100 missing values from a time series of 5000 data points, where experimental results show that PSO-EA algorithm is effective [15].

Aiming at the existed problem in the prediction of commercial bank cash flow, a hybrid learning algorithm is proposed based on an improved PSO algorithm combined with LMS to optimize the ANFIS’ configuration parameters, which is adopted to realize the prediction of cash flow time series. The simulation results show the effectiveness of the proposed method. The paper is organized as follows: in Section 2, the technique of the adaptive network-based fuzzy inference system (ANFIS) is introduced.
The optimization of ANFIS parameters based on improved PSO-LMS algorithm is presented in Section 3. The simulation experiments and results analysis are introduced in detail in Section 4. Finally, the conclusion illustrates the last part.

2. Adaptive Network-Based Fuzzy Inference System (ANFIS)

J.-S.R. Jang proposed an adaptive network-based fuzzy inference system (ANFIS) based on the T-S model in the early 1990s [17]. It is a new type of neural network, whose main feature is the organic blend of fuzzy logic and neural network. Sugeno Fuzzy model, also known as TSK fuzzy model, was put forward by Takagi, Sugeno and Kang [4], which is a systematic method to produce fuzzy rules based on a given input-output data set. Because the linearity of rules depends on the system input variables, the Sugeno model is an ideal multivariable controller, which can be applied to nonlinear dynamic systems with variety operating conditions.

The typical structure of ANFIS is shown in Figure 1 [17]. Assume that the considered fuzzy inference system has two inputs x and y, single output f. For the first order Sugeno fuzzy model, the common rule set with two fuzzy if – then rules is described as follows.

Rule 1: if x is A₁ and y is B₁, then:

\[ f_1 = p_1x + q_1y + r_1 \]  \hspace{1cm} (1)

Rule 2: if x is A₂ and y is B₂, then:

\[ f_2 = p_2x + q_2y + r_2 \]  \hspace{1cm} (2)

The first layer contains adaptive rules and membership functions, the second layer is used to calculate the firing strength, the third layer is the product of the firing strength, the fourth layer is the summation of the firing strength, and the fifth layer is the output of the system.

It can be seen from the ANFIS structure that the system has two adaptation layers (layer 1 and layer 4). The first layer has three adjustable premise parameters related to input membership functions. The fourth layer has three adjustable conclusions parameters related to the first-order polynomial. The root mean square error (RMSE) under the current premise parameters and conclusion parameters is calculated by:

\[ RMSE = \sqrt{\frac{1}{t} \sum_{k=1}^{t} (\hat{f}_k - f_k)^2} \]  \hspace{1cm} (3)
3. Optimization of ANFIS Parameters Based on Improved PSO-LMS Algorithm

3.1. Improved PSO Algorithm

In order to avoid the algorithm’s premature fall into local optimum and improve the population diversity and the convergence precision, the metric function of population diversity is defined based on the analysis of the convergence of the PSO algorithm. An adaptive population activity particle swarm optimization (APAPSO) algorithm is proposed according to the relationship between the population diversity and particle velocity. In this proposed improved PSO algorithm, the inertia weight is adjusted adaptively on the basis of the changes of the population diversity and regulation function, so as to improve the particle swarm diversity and the ability of the algorithm jumping out of the local optimal solutions.

3.1.1. Metric Function of Population Diversity

The population diversity reflects the distribution of particles in the searching space. The active degree of particles can be expressed by the current velocity of particles. The higher the active degree, the bigger the search scope of the population, therefore the diversity of the population can be reflected through the population activity. The average activity level of the population particles can be quantified as the mean square error of velocities. So it can be very good to understand and grasp the current state of the entire population by studying the changes in the size of population velocity mean square error. The speed of the population particle mean square error is defined as the population activity, which is described as follows:

$$PA(t) = \frac{1}{ND} \sum_{i=1}^{N} \sum_{j=1}^{D} [v_{ij}(t) - v_{avg}]^2$$

where $t$ is the number of iterations; $N$ and $D$ are the number of particles and the spatial dimension respectively; $v_{ij}$ is $j$-th dimensional velocity of the $i$-th particle; $v_{avg}$ represents the average velocity of the particle swarm.

The above definition shows that the group velocity variance $PA(t)$ can reflect the extent of the aggregation of all particles in the solution space, that is to say the diversity of particles in the population. The smaller the $PA(t)$, the smaller the diversity of particles. Conversely, the diversity is bigger. Iterative search of particles is a non-linear optimization process. The current iteration algebra of particles can be understood as a moment for the particles in the search process. Because the inertia weight in the PSO algorithm is between 0 and 1, therefore, the metric function of the population diversity is designed as follows:

$$F(t) = 1 - \frac{2}{\pi} \arctan(PA)$$

where the value of $F(t)$ is the diversity of the population metric function at the $t$ time. According to the Equation (5), when the population activity is smaller, the $F(t)$ value is larger. Conversely, when the population activity is larger, the $F(t)$ value is smaller. When the velocity mean square error of the particle swarm decreases gradually to zero, it indicates that the diversity of the population continues to decrease.
and the particles tend to be consistent gradually, which shows that the size of the population diversity measure function value can represent the difference between the particles in the population.

3.1.2. Adaptive Inertia Weight Adjustment Mechanism

The studies have shown that with the running of PSO algorithm, the individual particles in the population will tend to be consistent eventually. Thus the value of the function $F(t)$ will be bigger and bigger. Therefore, it cannot only be relied to represent the population diversity. At present most of the literature widely adopts linear decreasing strategy of inertia weight, that is to say $w$ will decline in the constant speed with the number of iterations $t$. Because the strategy with weight values is small in the later algorithm, the global search ability is weak. So it will make the algorithm is easy to fall into local optimum and the constant speed decreasing will also reduce the efficiency of searching. As the iteration proceeds, $F(t)$ should be weakened. Therefore, the regulatory function $\phi(t)$ is introduced, which is shown in the Equation (6):

$$
\begin{align*}
\phi(t) &= \exp\left[-\frac{t^2}{2\delta^2}\right] \\
\delta &= \frac{T}{3}
\end{align*}
$$

where $t$ is the number of current iteration; $T$ is the epoch of algorithm termination. When $T$ is 300, the chart of $\phi(t)$ is shown in Figure 2.

![Figure 2. Curve of regulating function.](image)

It can be seen from Figure 2 that if the size of $F(t) \times \phi(t)$ is as a measure of the degree of particles tending to be consistent in the whole population, on the basis of $F(t) \times \phi(t)$, the important parameter of the PSO algorithm, inertial factor is adaptive dynamic adjusted according to the changes of optimal conditions. This can not only help avoiding the defects of PSO algorithm easily falling into local extreme, but also speed up the convergence speed of the algorithm and improve the convergence precision of the algorithm in a great extent. So in the APAPSO algorithm, the strategy of inertia weight
adaptive dynamic adjustment is used. It changes along with the population diversity metric function and regulation function. The inertia weight adjustment mechanism is shown in Equation (7):

$$w(t) = w_{\text{min}} + (w_{\text{max}} - w_{\text{min}}) \times F(t) \times \varphi(t)$$

(7)

where \( t \) is the number of current iteration; \( w_{\text{min}} \) and \( w_{\text{max}} \) are the minimum and maximum values of the inertia weight, respectively.

3.2. Algorithm Procedure of APAPSO-LMS Algorithm

Each learning process of the hybrid learning algorithm includes the learning stage of the premise parameters and the conclusion parameters. In the premise parameters learning phase, an improved PSO algorithm is adopted for each individual to calculate the excitation intensity and the normalized excitation intensity of all the rules. In the conclusion parameter learning phase, the LMS method is used to identify these linear parameters. After obtaining all parameters, the output error of each input data can be calculated. The specific algorithm procedure is described as follows.

Step 1: Initialize the number of iterations \( n = 1 \). Randomly initialize the particle swarm. The position vector of the \( i \)-th particle is \( X_{id}^n \) and the velocity vector is \( V_{id}^n \) (\( 1 \leq i \leq m, 1 \leq d \leq D, m \) is the size of particle population, \( D \) is the dimension of searching space, that is also the number of prerequisite parameters);

Step 2: Make the position vector of each particle in turn as the premise parameters of ANFIS, and then calculate the incentive intensity \( w_j \) and the normalized incentive intensity (\( 1 \leq j \leq l \), \( l \) is the number of rules). Calculate the coefficient matrix \( A \) by the input data set and the normalized incentive intensity \( \tilde{w}_j \), and then use the least square method to identify the conclusion parameters \( \hat{\theta} \). Finally based on Equation (3) to calculate the root mean square errors (RMSE) of the corresponding particles produced by the ANFIS, which is named as the particle’s fitness value \( \text{RMSE}_i^n \);

Step 3: Compare the current fitness value \( \text{RMSE}_i^n \) of each particle with the best fitness value \( pBest_i \) itself. If \( \text{RMSE}_i^n < pBest_i \), then \( pBest_i = \text{RMSE}_i^n \), \( P_i^n = X_i^n \);

Step 4: Compare the fitness value \( \text{RMSE}_i^n \) of each particle with the best fitness value \( gBest \) of the particle swarm. If \( \text{RMSE}_i^n < gBest \), then \( gBest = \text{RMSE}_i^n \), \( P_g^n = X_i^n \);

Step 5: The velocity \( V_{id}^{n+1} \) and the position vectors \( X_{id}^{n+1} \) of each particle are updated. Based on the Equations (4)–(7), the inertia weight \( w(t) \) of particles is updated.

Step 6: Check whether the termination condition of the PSO algorithm is met. If the preset accuracy or the maximum number of iterations is reached, end the optimizing process and output the optimal solution; otherwise go to Step 2 and continue the next iteration.

4. Time Series Prediction of Bank Cash Flow Based on APAPSO-ANFIS

In order to demonstrate the effectiveness of the proposed APAPSO-ANFIS algorithm, and verify its rationality to predict the bank cash flow time series, based on the collected information and the market data of a commercial bank, the APAPSO-ANFIS is used to realize the time series forecasting of bank cash flow by adopting the MATLAB R2012a simulation platform.

In this paper, the inventory limit data of each day from 2010 to 2012 of a commercial bank are selected as the experimental data (a total of 1095 sample points). The data were carried out the normalization
pre-treatment, where the first 975 data is selected as the training data set, and the remaining 120 data is selected as the testing data set. As used herein, ANFIS contains 16 rules and each input variable is assigned two membership functions. The total numbers of the adjusted parameters are 104, including 24 premise (non-linear) parameters and 80 conclusion (linear) parameters. The number of particles \( N \) is 30; the number of iterations is 100; the learning factors \( c_1 = c_2 = 2 \); the scope of the inertia weight \( w \) is \([0.5, 1.2]\); the membership function of ANFIS is the bell-shaped function. Four input variables are given \([x(t-18), x(t-12), x(t-6), x(t)]\). Thus, the initial membership function and their corresponding termination membership functions after the improved PSO algorithm learning are shown in Figures 3 and 4, respectively.

![Initial membership function diagram of four input variables.](image)

**Figure 3.** Initial membership function diagram of four input variables.

![Final membership functions diagrams.](image)

**Figure 4.** Final membership functions diagrams.

Clearly seen from the comparison of the charts above, there are changes in the four membership functions after learning, which proves that the improved algorithm is feasible and effective. The prediction results and the error curves are obtained based on the improved PSO algorithm, which are shown in Figures 5 and 6. It can be seen from Figures 5 and 6 that the predicted results are very ideal.
In order to show the advantage of APAPSO-LMS algorithm to optimize the parameters of ANFIS, three different algorithms (the BP-LMS algorithm, PSO-LMS algorithm with inertia weight linear decreasing strategy and the APAPSO-LMS algorithm) are used to optimize ANFIS. The experimental comparison results of RMSE evolution curve is shown in Figure 7. The simulation prediction experimental results are shown in Figures 8 and 9. Respectively, and the comparison results of prediction errors are shown in Figure 10.

**Figure 5.** Prediction results based on APAPSO-LMS algorithm.

**Figure 6.** Error curve based on APAPSO-LMS algorithm.
**Figure 7.** Comparison results of RMSE.

**Figure 8.** Prediction results based on BP-LMS algorithm.
Figure 9. Prediction results based on PSO-LMS algorithm.

Figure 10. Comparison results of prediction errors.

Seen from the Figure 7 that when the BP-LMS method is used to optimize ANFIS and the number of iterations runs to about 55 generation, the optimization error almost reaches a stable value. When using PSO-LMS algorithm of inertia weight decreasing linearly strategy, although the optimization effect is
better than BP-LMS algorithm and the convergence rate is increased, but the result is not very ideal. That is to say that when the number of iterations runs to about 60 generations, the optimization error no longer reduces, the algorithm is probably falling into a local optimum.

When the APAPSO-LMS algorithm is used to optimize ANFIS, the error decreases all the time until 100 generations in the entire optimization process, and the rate of convergence and optimization results are much better than the previous two algorithms. Because the inertia weight of particles can be adjusted adaptively, the global search capability and the local exploitation ability of the algorithm can be balanced well, which can effectively avoid the premature convergence and improve the algorithm comprehensive optimization performance. By comparing the simulation results (Figures 8–10), it can be seen more intuitively that the proposed APAPSO-LMS prediction algorithm has much smaller error than the BP-LMS and PSO-LMS algorithm, and the prediction accuracy is higher. So, it can be concluded that the APAPSO-LMS algorithm is effective for the improvement of PSO algorithm and it has a certain practical significance.

In order to more clearly evaluate the predictive performance of the APAPSO-LMS algorithm, the results analysis is carried out based on the following five performance indexes. The prediction error is the deviation between the predicted results and the actual results, which determines the prediction accuracy. \(y_1, y_2, \ldots, y_n\) are the actual observations of predicted object and \(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n\) are the predicted values.

(1) Absolute error of predicted points
\[
a_t = y_t - \hat{y}_t, \quad t = 1, 2, \ldots, n
\]
where \(a_t\) is the absolute error at the point \(t\). Obviously, \(a_t\) is the most direct measure index of the prediction error, but it is affected by the measurement unit of the predicted objects. So it is unsuitable as the final measure indicator of prediction accuracy.

(2) Relative error of predicted points
\[
\hat{a}_t = \frac{a_t}{y_t} = \frac{y_t - \hat{y}_t}{y_t}, \quad t = 1, 2, \ldots, n
\]
where \(\hat{a}_t\) is the relative error at the point \(t\), which is usually expressed as a percentage and to measure the accuracy of the predicted values relative to the observed values at the predicted point \(t\).

(3) Prediction accuracy of the prediction points
\[
A_t = 1 - \left| y_t - \hat{y}_t \right| / y_t, \quad 0 \leq \left| y_t - \hat{y}_t \right| / y_t \leq 1
\]
\[
A_t = 0 \quad \left| y_t - \hat{y}_t \right| / y_t > 1
\]
where \(A_t\) is the prediction accuracy at the prediction point \(t\).

(4) Mean square error (MSE)
Mean square error (MSE) is a kind of convenient method to measure the average error to evaluate the degree of data change, which is described as follows.
\[
MSE = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2
\]
(5) Computational loading

The method of computational load is used in this paper to add “tic” direction as a started timer, and then put “toc” direction at the end of program as a terminate timer, and returning the total time since the “tic” direction is started.

The above mentioned three prediction algorithms are used to realize the time series prediction of bank cash flow. The simulation results summarized based on the training data and testing data are shown in Table 1. It can be seen from Table 1 that training data obtained accuracy is higher than the testing data, but the program running time is much longer, because the fitting degree of using training data is better and testing data need not spend more time to train the parameters. Although PSO-LMS algorithm may fall into local optimum lead to the computational time of the proposed APAPSO-LMS prediction algorithm is relatively longer, but for the time series prediction of bank cash flow it has the highest accuracy and the effectiveness of the proposed method is verified once again. In conclusion, by comparing the experiments’ results, it can be seen that the proposed prediction method is more suitable for bank cash flow time series forecasting and analysis.

<table>
<thead>
<tr>
<th>Performance indicators</th>
<th>BP-LMS</th>
<th>PSO-LMS</th>
<th>APAPSO-LMS</th>
<th>BP-LMS</th>
<th>PSO-LMS</th>
<th>APAPSO-LMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>$7.54 \times 10^{-5}$</td>
<td>$5.29 \times 10^{-5}$</td>
<td>$2.25 \times 10^{-6}$</td>
<td>$8.39 \times 10^{-5}$</td>
<td>$5.78 \times 10^{-5}$</td>
<td>$2.81 \times 10^{-6}$</td>
</tr>
<tr>
<td>Absolute error</td>
<td>$5.46 \times 10^5$</td>
<td>$4.37 \times 10^5$</td>
<td>$1.27 \times 10^5$</td>
<td>$5.95 \times 10^5$</td>
<td>$4.48 \times 10^5$</td>
<td>$1.62 \times 10^5$</td>
</tr>
<tr>
<td>Relative error of (%)</td>
<td>6.35%</td>
<td>3.19%</td>
<td>1.20%</td>
<td>6.68%</td>
<td>3.76%</td>
<td>1.43%</td>
</tr>
<tr>
<td>Prediction accuracy (%)</td>
<td>93.65%</td>
<td>96.81%</td>
<td>98.80%</td>
<td>93.32%</td>
<td>96.24%</td>
<td>98.57%</td>
</tr>
<tr>
<td>Computational time (s)</td>
<td>18.96 s</td>
<td>150.37 s</td>
<td>168.64 s</td>
<td>3.24 s</td>
<td>3.68 s</td>
<td>3.71 s</td>
</tr>
</tbody>
</table>

5. Conclusions

A hybrid learning algorithm based on adaptive population activity particle swarm optimization (APAPSO) algorithm and the least squares method (LMS) is proposed to optimize the parameters of ANFIS model, which is used to realize the time series forecasting of commercial bank cash flow. By the introduction of measure function of species diversity and adaptive inertia weight adjustment mechanisms, the algorithm optimization capability and convergence accuracy are improved and the convergence rate is accelerated to a great extent. The simulation results verify the proposed algorithm has better applicability for bank cash flow time series forecasting.

Acknowledgments

This work is partially supported by the Program for China Postdoctoral Science Foundation (Grant No. 20110491510), the Program for Liaoning Excellent Talents in University (Grant No.LR2014008), and the Project by Liaoning Provincial Natural Science Foundation of China (Grant No. 2014020177).
Author Contributions

Jie-Sheng Wang participated in the conception, design, interpretation, and commented on the manuscript. A substantial amount of Chen-Xu Ning’s contribution to the draft writing, critical revision data collection, and analysis and algorithm simulation of this paper was undertaken. Both authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).