Available online at www.sciencedirect.com

SciVerse ScienceDirect

Physics Procedia 25 (2012) 506 – 512

Physics Procedia

2012 International Conference on Solid State Devices and Materials Science

Prediction of the Chaotic Time Series Based on Chaotic Simulated Annealing and Support Vector Machine

Hu Yuxia [1], Zhang Hongtao[2]

[1]College of Electric Engineering , Zhengzhou University, Zhengzhou, China

[2]Institute of Electric power , North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou, China

Abstract

The regression accuracy and generalization performance of the support vector regression (SVR) model depend on a proper setting of its parameters. An optimal selection approach of SVR parameters was put forward based on chaotic simulated annealing algorithm (CSAA), the key parameters C and ε of SVM and the radial basis kernel parameter g were optimized within the global scope. The support vector regression model was established for chaotic time series prediction by using the optimum parameters. The time series of Lorenz system was used to testify the effectiveness of the model. The root mean square error of prediction reached 8.756×10^{-4} . Simulation results show that the optimal selection approach based on CSAA is available and the CSAA-SVR model can predict the chaotic time series accurately.

Keywords: support vector machine; chaotic simulated annealing algorithm; chaotic time series prediction; phase space reconstruction.

1. Introduction

With the development of chaos theory and its application, the analysis and forecasting of chaotic time series has become a very important research direction in signal processing field in recent years. The prediction of chaotic time series has been used to stock finance, electricity load forecasting, geological environment, weather forecasting [1-4].

Support Vector Machine (SVM) is new machine learning based on statistical theory[5]. It can solve small-sample, non-linear and high dimension problems by using structural risk minimization (SRM) instead of empirical risk minimization (ERM). It has stronger generalization ability and a very good application potentiality has been shown in classification and regression. The parameters selection of support vector machine is very important. The accuracy of classification or regression is determined by selecting a group of appropriate parameters. In recent years many researches on model selection have

1875-3892 © 2012 Published by Elsevier B.V. Selection and/or peer-review under responsibility of Garry Lee
Open access under CC BY-NC-ND license. doi:10.1016/j.phpro.2012.03.118

been done including grid search[6], mutative scale chaos optimization algorithm[7], Genetic Algorithm[8], etc.

Chaotic simulated annealing algorithm (CSAA) is a global optimization algorithm based on of simulated annealing algorithm (SAA) and chaos optimization algorithm (COA), possessing the advantages of efficiency, robustness and flexibility of SAA and the regularity, ergodicity and intrinsic stochastic properties of chaotic motion. The advantages above mentioned make CSAA more superior than the simulated annealing algorithm searching randomly, and CSAA can converge to the global optimal solution quickly [9-10].

In this paper, a chaotic time series forecasting model was established based on SVM, and the key parameters C 、 ε and the radial basis kernel function parameter g were optimized and trained by using CSAA. Then, the time series of Lorenz system was used to testify the effectiveness of the model, and Simulation results show that the optimal selection approach based on CSAA is available and the Convergence rate is faster than standard simulated annealing algorithm.

2. Overview of SVM and Prediction Model for Chaotic Time Series

2.1 SVM regression (SVR) theory

Prediction of chaotic time series can be attributed to support vector machine regression problems. In SVR, the basic idea is to map the data into a higher dimensional feature space via a nonlinear mapping $\Phi(\bullet)$ and then to do linear regression in the space. Therefore, regression approximation addresses the problem of estimating a function based on a given data set $\{\mathbf{x}_i, y_i;\ i=1,2,\cdots,N\}$ (where $\mathbf{x}_i \in R^n$ is the input vector and $y_i \in R$ is the desired value). SVM approximates the function with the form

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}} \bullet \Phi(\mathbf{x}) + b, \quad \Phi: R^n \to F, \mathbf{w} \in F \tag{1}$$

where $\{\Phi_i(\mathbf{x})\}_{i=1}^N$ are the data in features space, $\{\mathbf{w}_i\}_{k=1}^N$ and b are coefficients. They can be estimated by minimizing the regularized risk function:

$$R(C) = C\frac{1}{l}\sum_{i=1}^{N} L_\varepsilon(y_i, f(x_i)) + \frac{1}{2}\|\mathbf{w}\|^2 \tag{2}$$

Where $L_\varepsilon(y_i, f(x_i))$ is loss function measuring the approximation errors between expected output y_i and calculated output $f(\mathbf{x}_i)$, and C is regularization constant determining the tradeoff between the training error and the generalization performance. The second term $\frac{1}{2}\|\mathbf{w}\|^2$ is used as a measurement of function flatness. Introduction of relaxation factor ξ, ξ^* leads Eq. (2) to the following constrained function:

$$\min\ J(\mathbf{w}, \xi, \xi^*) = \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + C\sum_{i=1}^{N}(\xi + \xi^*)$$

$$\text{s.t.} \quad \begin{cases} y_i - \mathbf{w}^T \Phi(\mathbf{x}_i) - b \le \varepsilon + \xi_i \\ \mathbf{w}^T \Phi(\mathbf{x}_i) + b - y_i \le \varepsilon + \xi^*{}_i \\ \xi, \xi^* \ge 0 \end{cases} \tag{3}$$

Finally, by introducing Lagrange multipliers and exploiting to the optimality constraints, the decision Eq. (1) has become the follow form:

$$f(\mathbf{x}) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b \tag{4}$$

Where $\alpha_i, \alpha_i^*, i = 1, \cdots, N$ are Lagrange multipliers, satisfying the equalities $\alpha_i \times \alpha_i^* = 0$, $\alpha_i > 0$ and $\alpha_i^* > 0$. They are obtained by maximizing the dual formula of Eq. (3), which has the following form:

$$\begin{aligned} \max L(\boldsymbol{\alpha}, \boldsymbol{\alpha}^*) = & -\frac{1}{2} \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)(\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_l) \\ & - \varepsilon \sum_{i=1}^{N} y_i (\alpha_i - \alpha_i^*) + \sum_{i=1}^{N} y_i (\alpha_i - \alpha_i^*) \end{aligned} \tag{5}$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i, \alpha_i^* \le C, \ \ i = 1, 2, \cdots, N \end{cases}$$

According to the nature of SVM regression, most of the α_i and α_i^* are zeros. Hence, The final formulation can be arrived by solving optimization problem mentioned above and the support vector regression has the following form:

$$f(\mathbf{x}) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(\mathbf{x}, \mathbf{x}_i) + b \tag{6}$$

Although nonlinear function $\Phi(\bullet)$ is usually unknown, all computations related to $\Phi(\bullet)$ could be reduced to the form $K(\mathbf{x}_i, \mathbf{x}) = \Phi(\mathbf{x}_i)^{\mathrm{T}} \bullet \Phi(\mathbf{x})$. So what we need to do is selecting the appropriate kernel function. The advantages of RBF kernel function make it universally applied to SVM. The RBF kernel function can be applied to any sample by choosing right parameters.

2.2 Prediction model for chaotic time series based on SVM

1) Phase space reconstruction

Chaotic time series prediction is based on the theory of phase space reconstruction under Takens embedding theorem. A reconstructed phase space is a m-dimensional metric space into which a time series is embedded[10].

Given an observation chaotic time series $\{x(t)\}, t = 1, 2, \cdots, N$, selecting the embedding dimension m, the delay time τ, the phase space can be expressed as the following form:

$$\begin{aligned} X(t) &= (x(t), x(t-\tau), \cdots x(t-(m-1)\tau), \\ X(t) &\in R^m, (t = 1, 2, \cdots, M) \end{aligned} \tag{7}$$

where $X(t)$ is a vector or a point in the construction phase space, $M = N - (m-1)\tau$ is the number of points in the reconstructed phase space. If the embedding dimension is large enough, the phase space is homeomorphic to the state space that generated the time series, that is, the phase space contains the same information as the original state space. There is a determinism map $f(\bullet)$ meeting the following equation:

$$x(t+T) = f(X(t)) \tag{8}$$

Local model is usually used to predict the next-step, in which, $T > 0$ is the Step of forward prediction. $f(\bullet)$ is so called the predicting model for chaotic time series.

2) SVM predicting model for chaotic time series

Setting the $f(\bullet)$ of Eq. (8) as the support vector regression corresponding to Eq. (6) and $T=1$, the one-step prediction model for chaotic time series based on SVM has the following form:

$$\hat{x}(t+1)=\sum_{i=1}^{M}(\alpha_i-\alpha_i^*)\mathrm{K}(X(t),X(i))+b \tag{9}$$

where $\hat{x}(t+1)$ is the value of one-step forecasting, then the next points in phase space can be expressed as $\hat{X}(t+1)=(\hat{x}(t+1),x(t+1-\tau),\cdots x(t-(m-2)\tau))$。

3. Csaa For Parameters Optimization of SVM

3.1 CSAA

One-dimensional Logistic map is usually applied to chaos optimization algorithm whose mathematical model is the following form:

$$\mathrm{Z}_{k+1}=\mu\mathrm{Z}_k(1-\mathrm{Z}_k),\ \ \mathrm{Z}_k\in[0,1],\ \ k=1,2,\cdots, \tag{10}$$

where Z_k is the value of the variable Z at the kth iteration, Z_k in the interval [0, 1], μ is a so-called bifurcation parameter of the system ($0\le\mu\le4$). If $3.5699456\cdots\le\mu\le4$, logistic mapping works in chaotic state, that is, no steady-state solution, the system is a full map interval of [0, 1]. The logistic map has special characters such as the ergodic property, stochastic property and sensitivity dependence on initial conditions of chaos.

Consider the following continuous global optimization problem

$$J=\begin{cases}\min\ \ f(\mathbf{x})\\ \text{s.t.}\quad x_i\in[a_i,b_i],\ \ i=1,2,\cdots,m\end{cases} \tag{11}$$

where $[a_i,b_i]\subset R,\ \ i=1,2,\cdots,m.$ and f is a real-valued continuous function. Based on the chaos system and simulated annealing, the chaos simulated annealing algorithm (CSAA) is proposed for solving the problem J, which combines the features of chaotic system and simulated annealing. CSAA for solving problem J is described as follows.

Step 1: Chaotic initialize. Given an initial value Z_0. Generate the different chaotic variables Z_i^k, $i=1,2,\cdots,m,\ k=1,2,\cdots,N$ by formula (10), where N is the length of chaotic time series, k is a random integer in set $\{1,\cdots,N\}$, the length of chaotic time series should not be much bigger. The initial solution $\mathbf{x}_0=(x_{0,1},x_{0,2},\cdots,x_{0,m})^{\mathrm{T}}$ is produced by the formula

$$x_{0,i}=a_i+(b_i-a_i)Z_i^k,\qquad i=1,2,\cdots,m \tag{12}$$

Step 2: Initialize the initial temperature. Calculate the values of the objective function of the N feasible solutions obtained in step 1, record the maximum value $f_{\max}$ and minimum value $f_{\min}$, and take $f_{\min}$ as the best value, take $\mathbf{x}$ Corresponding to $f_{\min}$ as the current optimal solution X^*, the Corresponding chaotic state is denoted by C^*. If the acceptance probability of the solutions with Poor performance is P_{r}, according to Boltzman distribution, the initial temperature can be obtained by the following formula:

$$T_0=\frac{f_{\max}-f^*}{\ln P_{\mathrm{r}}} \tag{13}$$

Step 3: Simulated annealing process.

(1) set $Z^1=Z^k,k=1,\ n=0$.

(2) while ($T > T_{\min}$) do

(i) Generate a new solution $\mathbf{y}_n = (y_{n,1}, y_{n,2}, \cdots, y_{n,m})^{\mathrm{T}}$ based on the formula $y_{n,i} = x_{n,i} + \alpha \times (b_i - a_i) \times Z_i^k$, i is randomly chosen from the set {1, 2, …, m}. k is a random integer in set {1, …, N}. α is a variable which decreases by the formula $\alpha = \alpha \times e^{-\beta}$ in each iteration. Other components of $\mathbf{y}_n$ are same as that of $\mathbf{x}_n$.

(ii) Evaluate the change in energy level $\Delta E^* = f(\mathbf{y}_n) - f^*$ and $\Delta E = f(\mathbf{y}_n) - f(\mathbf{x}_n)$.

(iii) If $\Delta E^* \le 0$, update the best solution $\mathbf{x}^* = \mathbf{y}_n$ and the best value $f^* = f(\mathbf{x}^*)$.

(iv) If $\Delta E \le 0$, update current state with new state, $\mathbf{x}_{n+1} = \mathbf{y}_n$.

(v) If $\Delta E > 0$, update current state with new state with probability $\exp(\frac{-\Delta E}{T})$.

(vi) $n = n + 1$.

(3) Decrease temperature T according to annealing schedule by formula $T = \lambda T$.

*Step*4: Output the best solution $\mathbf{x}^*$ and the best value f^*.

3.2 the parameters optimization of SVM

IN SVR, the values of parameter C, σ and ε affect the model complexity in different ways. Parameter C determines the trade-off between model complexity and the degree to which deviations larger than ε are tolerated. Parameter ε controls the width of the $\varepsilon-$ insensitive zone and can affect the number of SV in the optimization problem. Kernel parameter σ determines the kernel width and relates to the input range of the training data set.

In Matlab environment, the LIBSVM developed by scholars Chih-Jen Lin was applied to realize chaotic time series prediction. Setting the ranges of the parameters: $C \in (2^{-5}, 2^{15})$, $\varepsilon \in (2^{-12}, 2^{0})$, $g \in (2^{-15}, 2^{5})$, It was evident that ranges of the parameters were y large enough to meet most systems. The length of Logistic time series was $N = 400$, annealing temperature decreasing coefficient was $\lambda = 0.95$.

4. Chaotic Time Series Prediction Experiments

Lorenz attractor was taken to examine the SVM prediction model optimized by CSAA whose formula is following form:

$$\begin{cases} \dot{x} = a(y - x) \\ \dot{y} = -xz + rz - cy \\ \dot{z} = xy - bz \end{cases} \tag{14}$$

The system is chaotic when $a = 10$, $b = 8/3$, $c = 28$. And the differential equations were solved numerically using 4th order Runge-Kutta integration with a step size $\Delta t = 0.01$ and initial value $x(0) = 0$, $y(0) = 1$, $z(0) = 0$. In order to eliminate the influence of the initial value, the initial 1000 points were discarded, the phase space reconstruction of x chaotic time series was done with the embedding dimension m, the delay time τ. Then 1500 points were taken as training data, 600 points were taken as test data.

The absolute error was taken to measure the accuracy of the prediction model.

$$e(i) = y(i) - \hat{y}(i) \tag{15}$$

Where $y(i)$ and $\hat{y}(i)$ were the actual values and predicted values of chaotic time series.

The root mean square error(RMSE) was used as integral performance index to evaluate the model .

$$RMSE = \sqrt{\sum_{1=1}^{N} (y(i) - \hat{y}(i))^2 / N} \tag{16}$$

Where N was the number of test points.

Normalizing dates to [0 1] and training SVR with CSAA, the optimal values of C, ε, g were obtained, $C = 2333.4$, $\varepsilon = 0.002925$, $g = 0.01603$ .Then the prediction experiment was done with the SVR model obtained, $RMSE = 8.756 \times 10^{-4}$, and the prediction and absolute error curves was shown in Figure 1 and Figure 2.

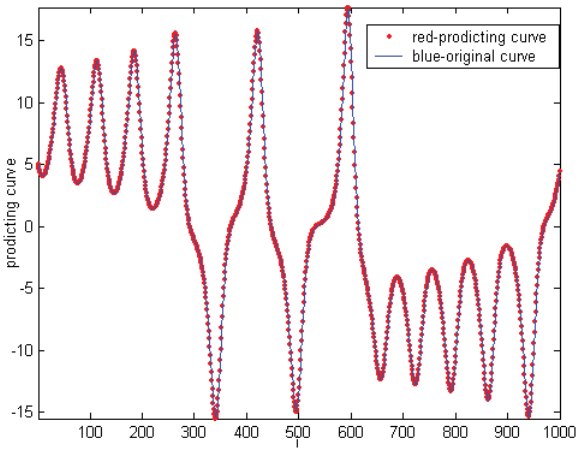

Fig 1. Predicted results and the original curve of Lorenz system

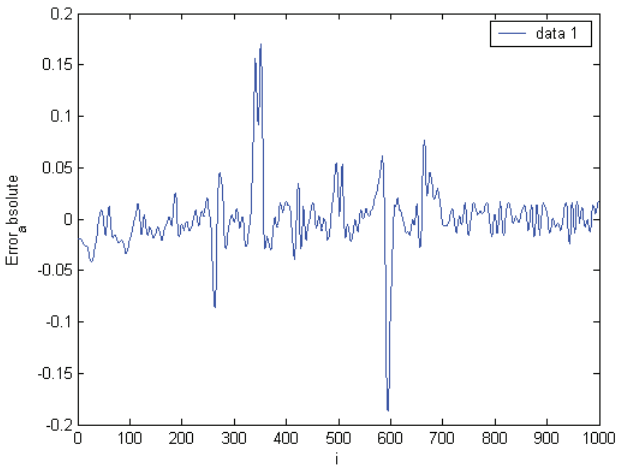

Fig 2. Absolute prediction error curve of Lorenz system

5. Conclusion

IN this paper, the chaotic simulated annealing algorithm was used for selecting the key parameters of support vector machine regression model, and the time series of Lorenz system was employed to examine the SVR model. Simulation results show the effectiveness of the algorithm, that is the SVM prediction model optimized with CSAA has a fast convergence rate and strong generalization capability, and can more accurately predict chaotic time series.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No.30871449) and Education Department of Henan Province Natural Science Research projects (No .2008B510014).

References

[1] Q. Huang，W. Xie, "The prediction of short-term exchange rate based on the phase space reconstruction and kalman," *Computer Applications and Software*, vol.25, no.8, pp79-80, August 2008.

[2] Z. Zhang, Y. Sun, "A new approach of STLF based on combination of phase space reconstruction theory and optimal recursive neural networks," *Electric Power*, vol.37, no. 1, pp19-23, January 2004.

[3] Z. Shan, C. Lin, "Chaos Prediction of Wave Hydrodynamic Pressure Signals Based on Local Support Vectors Machine," *Journal of System Simulation*, 2008, vol.20, no.23, pp6470-6472, December 2008.

[4] G. Wu, P. Zhou, "Prediction Based on Phrase Construction and Its Application in Weather Forecast," *Chinese journal of nature*, vol. 21, no. 2, pp107-110, April 999.

[5] Cortes C, Vapnik V N, "Support Vector Networks," Machine Learning, Vol.20, no.3, pp273-295, September 1995.

[6] C. Hsu, C. Chang , C. Lin, "A practical guide to support vector classification," http://www.csie.ntu.edu.twcjlin/papem/guide/guide.pdf, 2003

[7] Q. Liu, P. Que, C. Fei, "Model selection for SVM using mutative scale chaos optimization algorithm," *Journal of shanghai university*, vol. 10, no. 6, pp531-534, December 2006. 36(4), P101-103.

[8] C. Huang C. Wang, "A GA-based feature selection and parameters optimization," Expert Systems with Applications, no. 31, pp231–240, November 2006.

[9] M. Ji, H. Tang, "Application of chaos in simulated annealing," Chaos, *Solutions and Fractals*, no.24, pp933 – 941, August 2004.

[10] L. Wang, D. Zheng, "Simulated annealing with the state generator based on Cauchy and Gaussian distributions", *Tsinghua Science and Technology*, vol. 40, no.9, pp 109-112, September 2000.