

Research Article

Accuracy Optimization of Synchronization Parameters in Chaotic Time Series Prediction

JianWei Lin* and Zibin Xu

Institute of Information Technology, Zhejiang Shuren University, Hangzhou, 310014, Zhejiang, China.

Received 23 September 2014; Revised 15 October 2014; Accepted 10 November 2014

Abstract

In view of the situation that current chaotic time series prediction model still exists the problem of low accuracy, this paper proposes a synchronous parameter optimization strategy of chaotic time series model based on improved genetic algorithm. It firstly takes the random direction method to construct the initial population of genetic algorithm to speed up the convergence, and then prescribes a limit to the objective function with improved punishment function. Then it optimizes the crossover, mutation and duplication operators of the genetic operator to avoid the premature convergence. Moreover, it adopts the complex method to optimize the local optimization ability of the standard genetic algorithm on the basis of improved fitness function. Finally, this paper applies the improved algorithm into the synchronous parameters optimization of chaotic time series prediction model. Simulation results show that the proposed algorithm has higher accuracy and efficiency in parameter optimization.

Keywords: Chaotic time series, chaotic prediction, synchronization parameters optimization, improved genetic algorithm, accuracy optimization, genetic operator, complex method, punishment function optimization.

1. Introduction

Chaotic prediction, especially chaotic time series prediction is of great importance in chaos theory study [1]. Time series refer to a variable or index in the natural science and social science. It is a set of numerical value ordered according to the appearance time and same time interval [2]. Chaotic time series is a kind of time series with chaotic characteristics, and chaotic time series prediction uses past and present observation data, establishing sequence model by some method to speculate the future of sequence. This is an important branch of statistics developed in recent decades [3]. At present, the prediction of chaotic time series has a wide application in target tracking, run-off prediction, weather forecast, nonlinear market analysis, communication signal recognition etc. [4].

In recent years, people pay more and more attention to the nonlinear system especially the time sequence under the chaotic background. Since Packard et.al proposed two methods of reconstructing phase space from chaotic time series: derivative reconstruction and coordinate delay reconstruction, the theory and methods of chaotic time series analysis have been further studied [5]. Broomhead and King put forward a method of singular value decomposition (SVD), also called as principal component analysis (PCA) that used derivative of time series to reconstruct the phase space [6]. Packard took the delay-coordinate of a variable in original system to reconstruct prediction model network of

phase construction [7]. Takens justified if the embedded dimension and delay time is appropriate with mathematical methods, the reconstructed phase space is in diffeomorphism with original kinetic system, which provided technological foundation for the phase space reconstruction [8]. Hense A firstly introduced chaos theory into hydrology study, deducing the existence of singular attractor according to the relevant dimension value of rainfall time series [9]. Sangoyomi used chaotic method to predict the water volume of Great Salt Lake, superior to traditional prediction methods, triggering people's interest on hydrology chaotic prediction [10]. Ding Jing studied the property of flood with chaotic methods and concluded out a series of hydrology research method [11]. Chen Ping discovered singular attractor in the macroscopic tendency factor, followed by abundant studies in exchange rate, stock index and inflation [12]. Mori and Urano applied chaos theory into the prediction of load series, judged the chaos characteristic with Lyapunov index, and predicted the series after phase space reconstruction with multilayer perceptron [13]. Badel and Dudnai proposed the exponential weighting prediction and distance weighting prediction method respectively, and large amount of numerical experiment and applications showed that these methods were better than local zero-order and mean prediction method in prediction performance and anti-noise property [14]. Mukherjee proved the support vector machine was better to predict the time series than traditional methods, but he only discussed one chaotic time series and some kernel functions [15]. Fraser gave recursive algorithm calculated by mutual information, which considered the nonlinear characteristic of time series but occupied large

* E-mail linjianwei1980@yeah.net

calculated amount [16]. Lin Jiayu put forward complex autocorrelation method without nonlinear characteristic of time series [17]. The C-C method, proposed by Kim, was used to estimate the embedded dimensions and delay time, which efficiently decreased the calculated amount and kept the nonlinear characteristic at the same time. Though there are no strong foundation, it shows unique features in real work [18].

In view of the situation that current chaotic time series prediction model shows bad prediction accuracy, this paper optimizes the prediction model based on genetic model, and improves the initial population, objective function, genetic operator and searching ability of genetic algorithm to improve the accuracy of chaotic time series prediction.

2. The Model Construction of Chaotic Time Series Prediction

In the study of chaotic time series, either the inner calculation of chaotic system or the construction and prediction of chaotic model is based on phase space, therefore, the phase space reconstruction is of great importance in analysis of chaotic time series.

Suppose M is d dimension space, $\varphi: M \rightarrow M$, φ is a smooth diffeomorphism; $y: M \rightarrow R$, y has the two-order continuous derivatives; $\Phi(\varphi, y): M \rightarrow R^{2d+1}$, here,

$$\Phi(\varphi, y) = (y(x), y(\varphi(x)), y(\varphi^2(x)), \dots, y(\varphi^{2d}(x))) \quad (1)$$

Then $\Phi(\varphi, y)$ is an embedding from M to R^{2d+1} .

For the scalar time series $\{x_1, x_2, \dots, x_{n+1}, x_n\}$ of d dimensional attractor, an embedded phase space time with m dimension is always existent in a topology equivalent significance, with the premise of $m \geq 2d + 1$. Namely, if the m and delay time τ is appropriate, then the reconstructed phase space is,

$$Y(t_i) = \{x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m-1)\tau)\} \quad (2)$$

where $i=1, 2, \dots$. It will recover the kinetic properties of chaotic system itself from the aspect of invariant topology.

This paper consider the observed chaotic time series $\{x_t | t=1, 2, \dots, N\}$, and the kinetic equation of chaotic system current time series is,

$$x_{t+1} = f(X_t) \quad (3)$$

X_t is the vector constituted by the observed chaotic time series component x_t .

The best embedded dimension m and delay time τ is used for phase space reconstruction to obtain the vector,

$$X'_t = [x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}]^T \quad (4)$$

In a similar way, X'_{t+1} is obtained.

From (1), the relationship between the reconstructed vector X'_t and X'_{t+1} is,

$$X'_{t+1} = f'(X'_t) \quad (5)$$

Here, f' and f is diffeomorphism.

When constructing chaotic time series prediction model, it is required to determine the embedded dimension m and time delay τ . However, the time series observed in real life is limited with noise. Therefore this paper optimized these two key parameters synchronously.

3. The Synchronization Parameters Optimization Based on Improved Genetic Algorithm

Genetic algorithm possesses a fitness function as a special driving force to guide the algorithm operation with strong global searching ability. It does not require the derivable and differentiable property of objective function. Therefore, this paper takes genetic algorithm to optimize synchronously the prediction model parameters, and firstly improves the genetic algorithm to increase the accuracy.

3.1. Population Initialization Based on Random Direction Method

In order to avoid the omission of some information in searching space, this paper takes the random direction method to construct the initial population.

For chromosome $X = x_1 x_2 \dots x_n$ where $x_i (i=1, 2, \dots, n)$ represents the genetic gene, with value interval $[x_{i\min}, x_{i\max}]$, ($i=1, \dots, n$) suppose the initial population size is P , then this problem is seen as how to select P individuals randomly.

The population initialization based on random direction has following steps.

(1) Determining the value interval,

$$S = [x_{1\min}, x_{1\max}] \cdot [x_{2\min}, x_{2\max}] \cdot \dots \cdot [x_{n\min}, x_{n\max}] = \prod_{i=1}^n [x_{i\min}, x_{i\max}] \quad (6)$$

(2) Data initialization,

Randomly set a starting point X_0 , denoted as base point. Initialization $k=0$, is used for recording the individual number; initialized array $X(P)$, for storing individuals. A random vector, ρ_0 is random vector uniformly distributed on the sphere with radius δ .

(3) Compare $X = X_0 + \rho_0$ with X_0 one by one, if the fitness value of X is better than that of X_0 , then $k = k + 1$, and if $k < P$, then $X(k) = X$, otherwise $k = P$, procedure ends.

(4) If the fitness value of X is worse than that of X_0 , then compare the fitness value of $X = X_0 - \rho_0$ and X_0 , if the fitness value of X is better than that of X_0 then $k = k + 1$, and if $k < P$, then $X(k) = X$, otherwise $k = P$, procedure ends.

(5) In step (3) and (4), if the corresponding fitness value of X is worse than that of X_0 , then go back to step (2).

(6) Prolonging twice the distance of $X_0 X$ to get the reference point R_0 and randomly generating a vector ρ_0 , a random vector uniformly distributed on the sphere with radius δ .

(7) Comparing the fitness value of $R = R_0 + \rho_0$ and R_0 , if the fitness value of R is better than that of R_0 , then $k = k + 1$. And

if $k < P$, then $X(k) = X$; otherwise, $k = P$ and procedure ends.

(8) If the fitness value of R is worse than that of R_0 , then comparing the fitness value of $R = R_0 - \rho_0$ and that of R_0 . If the fitness value of R is better than that of R_0 , then $k = k + 1$. And if $k < P$, then $X(k) = X$; otherwise, $k = P$ and procedure ends.

(9) If both the corresponding fitness value of R in step (7) and (8) is worse than that of R_0 , then go back to step (6).

(10) Comparing the fitness value of $X(k-1)$ and $X(k)$, if the fitness of $X(k)$ is better than that of $X(k-1)$, then $X_0 = X(k-1)$, $X = X(k)$, and then go back to step (6).

(11) If the fitness of $X(k)$ is worse than that of $X(k-1)$, then $X_0 = X(k-2)$, and go back to step (2).

By these steps, we select several better individuals to form the initial population to speed up the convergence.

3.2. Objective Function Optimized By Punishment Function

In standard genetic algorithm, general punishment functions have a common shortcoming that the optimal solution of $G(X)$ is not always the one of $f(X)$. Therefore, according to the degree of individual violating the constraint conditions, this paper optimizes it and constrains the objective function.

The average infeasible degree M of population is a measure for the violation degree of constraints,

$$M = \frac{\sum_{X \in S} E(X)}{P} \quad (7)$$

Here, S is the set of population, P is population size.

On the basis of average infeasibility, the infeasible threshold is introduced,

$$N = T \cdot M \quad (8)$$

Here, T changes with the evolutionary times. At the beginning of evolution, the search scope is expanded, so that T can be valued largely, and in order to ensure the feasibility of solution at the late period of evolution, the value of T should decrease with evolutionary time.

Due to the property of T , we refer it as annealing factor, and define it as,

$$T = \frac{D-d}{D} \quad (9)$$

Here, D is the maximum iteration time the algorithm sets; d is current iteration number.

Because punishment is introduced for dealing with inequality constraints, the improved punishment function method of this paper forms new objective function $G(X)$ which is also a punishment function. This punishment function has the mapping relationship with original objective function $f(X)$ as follows.

For solving the global maximum solution of objective function, the mapping relationship is,

$$G(X) = \begin{cases} f(X), E(X) \leq A \\ f(X) - E(X), E(X) > A \end{cases} \quad (10)$$

For the global minimum solution problem, the mapping relationship is,

$$G(X) = \begin{cases} f(X), E(X) \leq A \\ f(X) + E(X), E(X) > A \end{cases} \quad (11)$$

where $f(X)$ is the original objective function, $E(X)$ is the degree of individual violation of constraints.

3.3. The Optimization of Genetic Operator

To solve the problem of premature convergence of standard genetic algorithm in the reproduction of offspring of the genetic operator, this paper optimizes the crossover operator, mutation operator and duplication operator of genetic operator.

This paper proposed the improved the crossover probability and mutation probability,

$$P_c = \begin{cases} p_{c3} - \frac{p_{c3} - p_{c2}}{1 + \exp\left(A \left(1 - \frac{2(f_{max} - f')}{f_{max} - f_{avg} + r}\right)\right)}, f' \geq f_{avg} \\ p_{c2} - \frac{p_{c2} - p_{c1}}{1 + \exp\left(A \left(1 - \frac{2(f_{avg} - f')}{f_{avg} - f_{min} + r}\right)\right)}, f' < f_{avg} \end{cases} \quad (12)$$

$$P_m = \begin{cases} p_{m3} - \frac{p_{m3} - p_{m1}}{1 + \exp\left(A \left(\frac{2(f_{max} - f')}{f_{max} - f_{avg} + r} - 1\right)\right)}, f' \geq f_{avg} \\ p_{m3} - \frac{p_{m3} - p_{m2}}{1 + \exp\left(A \left(\frac{2(f_{avg} - f')}{f_{avg} - f_{min} + r} - 1\right)\right)}, f' < f_{avg} \end{cases} \quad (13)$$

Here, $p_{c1} > p_{c2} > p_{c3}$, $p_{m1} > p_{m2} > p_{m3}$, $A = 9.9$, r is a small positive number.

But the p_c calculated from equations above focuses on p_{c1} , p_{c2} and p_{c3} , p_m focuses on p_{m1} , p_{m2} and p_{m3} . Its reason is due to the small exponent of the exponential function in the denominator. Therefore, this paper gave the adjusted mutation probability and crossover probability, as follows.

$$P_c = \begin{cases} p_{c3} - \frac{p_{c3} - p_{c2}}{1 + \exp\left(A \left(2 - \frac{3(f_{max} - f')}{f_{max} - f_{avg} + r}\right)\right)}, f' \geq f_{avg} \\ p_{c2} - \frac{p_{c2} - p_{c1}}{1 + \exp\left(A \left(1 - \frac{3(f_{avg} - f')}{f_{avg} - f_{min} + r}\right)\right)}, f' < f_{avg} \end{cases} \quad (14)$$

$$P_m = \begin{cases} P_{m2} - \frac{P_{m2} - P_{m1}}{1 + \exp\left(A \left(\frac{3(f_{max} - f')}{f_{max} - f_{avg} + r} - 2\right)\right)}, f' \geq f_{avg} \\ P_{m3} - \frac{P_{m3} - P_{m2}}{1 + \exp\left(A \left(\frac{3(f_{avg} - f')}{f_{avg} - f_{min} + r} - 1\right)\right)}, f' < f_{avg} \end{cases} \quad (15)$$

For those individuals of poor fitness value, the duplication operator of standard genetic algorithm easily falls into death after operation. This paper combines the probability survival method and optimal retention strategy, so as to improve the variety of population and ensure the convergence of algorithm.

There are several main steps for probability survival method in solving the globally maximum optimization problem of objective function.

Firstly, the mean fitness value \bar{F} of population is calculated, together with the fitness value $F(X)$ of individual X . If

$$F(X) \geq \bar{F} \quad (16)$$

Then, this individual is reserved, otherwise,

$$g(X) = \frac{F(X)}{\sum_{X \in S} F(X)} \quad (17)$$

Randomly generate a number r between 0 and 1, If $g(X) \geq r$, then this individual is reserved, otherwise, it is eliminated. Similarly, the main steps for solving the globally minimum optimization problem are:

Firstly, the mean fitness value \bar{F} is calculated, together with the fitness value $F(X)$ of individual X . If,

$$F(X) \leq \bar{F} \quad (18)$$

Then this individual is reserved, otherwise, do the equation (17).

3.4. Local Optimization Based on Complex Method

In order to improve the accuracy of improved genetic algorithm in chaotic time series prediction, this paper optimizes the local optimization problem of original algorithm, and firstly adjusts the fitness function.

$$Fitness = \begin{cases} \frac{1}{1 + \exp((f - f_{avg})/c)}, & g \geq 0.3n \\ \frac{1}{1 + \exp(f - f_{avg})}, & g < 0.3n \end{cases} \quad (19)$$

where f is original fitness, if the problem itself is a minimization problem, then f is the objective function, but if the problem itself is a maximization problem, then the problem should be switched to minimization problem; f_{avg} is the mean value of fitness of current population; making difference between the f value of each individual and f_{avg} ,

and g represents the individual number of difference out of $(-10, 10)$; c is the order of magnitude of the absolute maximum value of the difference between f and f_{avg} .

Then, the complex method is used for the local optimization problem.

(1) To set the fitness value from (19) as the objective function.

(2) In each evolutionary step of genetic algorithm, the complex method as a basic operator of genetic algorithm, improves the property of current population. Therefore, in this algorithm, complex method sets each individual as the peak of recombination.

(3) To calculate the objective function of each peak and centroid \bar{X} of the complex, find the best point X_g and test whether it meet the convergence criterion.

$$\|F(X_g) - \bar{F}\| \leq \varepsilon \quad (20)$$

$$\|X_g - \bar{X}\| \leq \delta \quad (21)$$

where,

$$\bar{F} = \frac{1}{P} \sum_{i=1}^P F(X_i) \quad (22)$$

$$\bar{X} = \frac{1}{P} \sum_{i=1}^P X_i \quad (23)$$

Here, P is population size, ε and δ is a preset small positive number, adjustable according to the value of objective function and the requirements of problem. If it meets the termination rule, then iteration ends, otherwise, go into step (4).

(4) To search out the worst point X_b , calculate the centroid X_c of each peak except the X_b . If the centroid meets the constraint conditions, then turn step (5); otherwise, detect the individual number m that doesn't meet the constraint conditions, and form m new individuals in the feasible region determined by constraints to replace the original m individuals, turn step (3).

(5) Firstly, to select reflection coefficient α and calculate reflection point X_m ,

$$X_m = (1 + \alpha) \cdot X_c - \alpha \cdot X_b \quad (24)$$

Then, check whether X_m meets the constraints, if it meets, then turn step (6), otherwise, the constriction operation is executed,

$$X_m = \frac{1}{2}(X_m + \bar{X}) \quad (25)$$

Until it meets the constraint conditions, and turn step (6).

(6) To calculate $F(X_m)$, if $F(X_m) < F(X_b)$, then replacing X_b with X_m to form new complex, turn step (3); otherwise, do the constriction operation and turn the second operation of step (5).

3.5. The Optimization of Synchronization Parameters of Prediction Model Based on Improved Algorithm

Firstly, the initial population generated by improved genetic algorithm is set as the parameters of prediction model, specifically, γ , σ , τ and m , with which the prediction model is constructed to get the error; then the fitness value is calculated, through duplication, crossover, mutation etc. operations to generate population of next generation, and repeat this procedure until the fitness meets the termination rule of improved genetic algorithm; finally, the optimized parameters are used for model prediction.

- (1) To set the iteration number $t = 0$, initialize the population R_{gen} generated by random direction method with size N_{pop} .
- (2) To reconstruct the chaotic time series according to τ and m , learn and establish model with γ and σ , predict and record the accuracy, and calculate the fitness value of each individual.
- (3) If the fitness value of the optimal individual meets the requirement or reaches the preset iteration number, then turn step (6).
- (4) $t = t + 1$.
- (5) To generate new population with improved duplication, crossover and mutation operators, and turn step (2).
- (6) To get current optimal γ , σ , τ and m .
- (7) To reconstruct the chaotic time series with optimal τ and m , train and predict with optimal γ and σ , and finally get the optimal prediction results.

4. Simulation Experiment

To verify the effectiveness of this improved algorithm, this paper conducted the simulation experiment. Firstly, the classical OliverTSP problem is used to test the efficiency of genetic algorithm. In this paper, population size is 40, crossover probability, mutation probability. The comparison result of standard genetic algorithm and improved genetic algorithm is shown in table1.

The number of cities	Coordinate error/%		Iterations	
	GA	IM-GA	GA	IM-GA
10	8.23	3.13	21	11
20	10.53	3.46	26	13
30	15.71	3.87	32	16
40	18.37	3.95	45	19
50	20.47	4.15	49	21
60	22.51	4.94	52	22
70	25.74	5.48	59	25
80	28.90	5.81	68	27
90	31.26	6.03	79	30
100	34.42	6.33	93	31

Tab. 1. The results for the test.

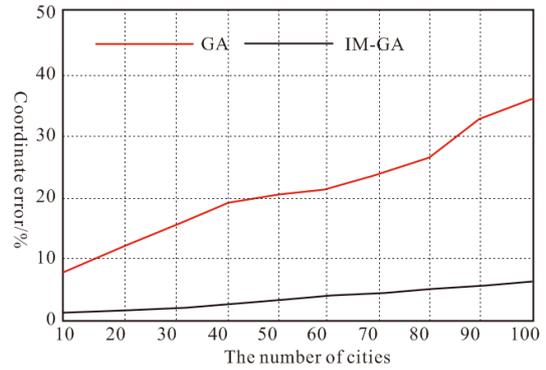


Fig.1. Accuracy analysis of two algorithms for solving.

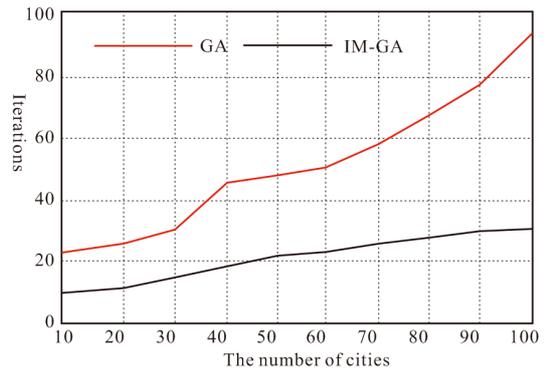


Fig.2. Efficiency analysis of two algorithms for solving.

It is apparent from Figs.1 & 2 that the improved algorithm has higher efficiency and accuracy compared with original genetic algorithm.

Then, we use improved genetic algorithm to optimize the parameters of chaotic time series prediction model synchronously, taking Lorenz chaotic times series as example, with results shown in Figs.3 & 4.

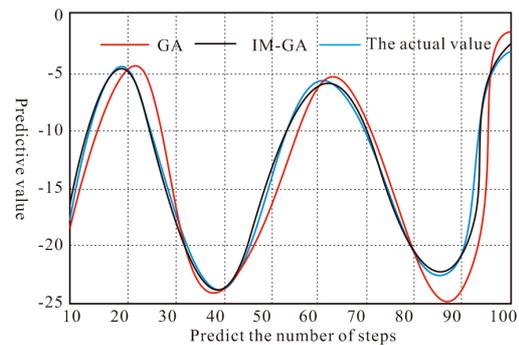


Fig.3. Predictive value comparison of the two algorithms.

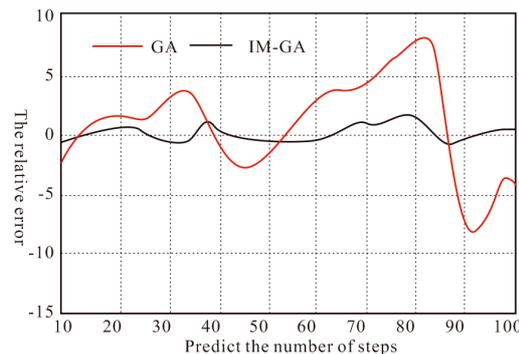


Fig.4. Comparison of two algorithms prediction error.

From the simulation results, it is seen that the proposed algorithm has higher accuracy in synchronous parameters optimization of chaotic time series prediction model.

3. Conclusion

Chaotic time series prediction is a very important and challenging work in chaotic analysis. This paper puts forward a synchronous optimization strategy of chaotic time series prediction model based on improved genetic algorithm. The simulation results show that the proposed algorithm has higher accuracy and efficiency in synchronous

parameters optimization of chaotic time series prediction model.

Acknowledgement

This work was supported by the Department of Science and Technology of Zhejiang Province (Grant No.2014C31065), and the National Natural Science Foundation of China (Grant No.51175473).

References

1. Z. Guoyong, A simple model for probabilistic interval forecasts of wind power chaotic time series, *Acta Phys. Sin.*, vol. 2014, pp.422-430 (2014).
2. H.G.Yu, Risk analysis on long inclined-shaft construction in coalmine by TBM techniques based on multiple variables chaotic time series, *Acta Phys. Sin.*, vol. 2014, pp. 103-110 (2014).
3. H. Ming, Research on multivariate chaotic time series prediction using mRSM model, *Acta Automatica Sinica*, vol. 2014, pp. 822-829 (2014).
4. L. Miao, Novel hidden feature extraction method for chaotic time series prediction, *Chinese Journal of Scientific Instrument*, pp.1-7 (2014).
5. T. Zhoujin, A least square support vector machine prediction algorithm for chaotic time series based on the iterative error correction, *Acta Physica Sinica*, pp.70-79 (2014).
6. Y. Can, Prediction about Chaotic Times Series of Natural Circulation Flow under Rolling Motion, *Nuclear Power Engineering*, pp.60-63 (2014).
7. C. Qing, Application of LSSVM in chaotic time series based on ADE algorithm", *Computer Applications and Software*, pp. 275-277 (2014).
8. Z Yongping, Chaotic time series prediction using add-delete mechanism based regularized extreme learning machine, *Acta Phys. Sin.*, pp.78-85 (2013).
9. H. Ming, Multivariate chaotic time series prediction based on extreme learning machine, *Control Theory & Applications*, pp.1467-1472 (2013).
10. G Junjie, Chaotic time series local prediction based on SVM, *Computer simulation*, pp. 303-306 (2013).
11. W Fuyong, Adaptive forecasting model of dam safety monitoring based on chaotic time series, *Water Power*, pp.72-76 (2013).
12. W. Xiangxi, Chaotic time series prediction based on optimal training subset online fuzzy LSSVM, *Journal of Applied Physics*, pp.411-417 (2013).
13. P. Yuming, Chaotic time series prediction based on QPSO-FNN, *Computer Application and Software*, pp.91-94 (2013).
14. W. Yang, Comparative study on five defuzzification methods in chaotic time series prediction, *Computer simulation*, pp.133-137 (2013).
15. M. Wentao, MCC based adaptive prediction algorithm simulation for chaotic series, *Computer Simulation*, pp. 247-250 (2013).
16. H. Tengfei, Noise estimation and reduction for chaotic time series by wavelet variance decomposition, *Journal of Computer Applications*, pp.890-895 (2013).
17. L. Wen, Parameters determination based on combinative evolutionary algorithm for reconstructing phase-space in chaos time series, *Journal of Central South University: Science and Technology*, pp. 3246-3253 (2013).
18. H. Ming and X Meiling, A hybrid prediction model of multivariable chaotic time series based on error correction, *Acat Phys. Sin.*, pp.106-112 (2013).