The Chaos Differential Evolution Optimization Algorithm and its Application to Support Vector Regression Machine

Wei Liang, Laibin Zhang, Mingda Wang *
College of Mechanical and Transportation Engineering, China University of Petroleum, Beijing, China
E-mail: soclassic@sina.com

***Abstract*—The Differential Evolution (DE) population-based algorithm is an optimal algorithm with powerful global searching capability, but it is usually in low convergence speed and presents bad searching capability in the later evolution stage. A new Chaos Differential Evolution algorithm (CDE) based on the cat map is proposed which combines DE and chaotic searching algorithm. Firstly, the chaotic distributed superiority of the cat map is analyzed in this paper. Secondly, the detailed implementation of CDE is introduced. Finally, the effectiveness of CDE is verified in the numerical tests. The Support Vector Regression machine (SVR) is an effective tool to solve the problem of nonlinear prediction, but its prediction accuracy and generalization performances depend on the selection of parameters greatly. So, the CDE is applied to SVR to build an optimized prediction model called CDE-SVR. Then the new prediction model is applied to the short-time regression prediction of the chaotic time series and the boundary extension of the mechanical vibration signals. The results of the two experiments demonstrate the effectiveness of the CDE-SVR.**

***Index Terms*—differential evolution, chaotic cat map, support vector regression machine, parameters optimization, boundary extension**

I. INTRODUCTION

Differential evolution algorithm introduced by Stron and Price for the real parameter optimization problems is a new kind of global optimization algorithm [1]. With the advantage of faster convergence speed, less adjustable parameters, better robustness and simpler algorithm, the DE algorithm has been achieved fine application effects in neural network training, filter design, cluster analysis[2,3].

Although the convergence of DE is fast, some problems of DE need to be solved yet, such as: in the later stage of DE, the convergence speed is slow, even to the extent that falling into the local optimum and presenting the premature. In order to overcome these problems, the chaotic searching that with the property of randomness, ergodicity and initial sensitivity is used to improve the DE algorithm, which is called Chaos Differential Evolution (CDE).

The Support Vector Machine (SVM) [4] that based on the statistical learning theory and structural risk minimization principle has been successfully applied in solving the problems of classification and regression. Since the performances of SVM greatly depend on the kernel function type and especially the nuclear parameters as well as the punishment parameters, the parameter selection has always been a hot issue in the SVM theory and application.

The mainly methods of SVM parameter selection are grid searching and gradient descent algorithm in the early stage and the optimal methods recently. For example: the approach based on the Genetic Algorithm (GA), the Simulated Annealing algorithm and the Partical Swarm Optimization (PSO) algorithm [5,6,7].While these parameters selection approaches based on the optimization algorithm has cut the searching time and reduced the dependence on the initial values, the GA and SA were difficult to be implemented and the PSO was easy to fall into local optimum, which would bring a low optimization efficiency.

This paper proposed a hybrid model based on the CDE algorithm and the Support Vector Regression machine (SVR) model, which is called CDE-SVR. The proposed model is then applied to the short-time regression prediction of the chaotic time series of Chens and the boundary extension of the mechanical vibration signals. The application results verify the effectiveness of CDE-SVR.

II. CONVENTIONAL DE ALGORITHM AND FITNESS VARIANCE OF THE POPULATION

A. Conventional DE algorithm

The differential evolution method based on population algorithm is used to approximate the global optimal solution. Generally speaking, the problem of global optimization can be transformed into solving the following minimization problem: $\min_{x\in D} : f(P)$, where $\boldsymbol{P} = (p_1, p_2, \cdots, p_n) \in D \subset \boldsymbol{R}^n$ is a vector of n-dimension, $f : D \to \boldsymbol{R}$ is the objective function. The generation k could be: $\boldsymbol{P}^k = \{p^k{}_1, \cdots, p^k{}_{NP}\}$, with that the number of individuals is NP . As in other evolutionary algorithms the main operators in DE are: mutation, recombination and selection.

1) Mutation

This operator in DE is rather different than in other evolutionary algorithms. In this step, three individuals, $p^k{}_{r1}$ 、 $p^k{}_{r2}$ and $p^k{}_{r3}$, are randomly chosen from the

doi:10.4304/jsw.6.7.1297-1304

current generation-k. The first individual p^k_{r1} is the base of the mutated vector and we could get the mutant v_i by the following formula:

$$v_i = p^k_{r1} + F(p^k_{r2} - p^k_{r3}) \tag{1}$$

Where, $i = 1, \cdots, NP$; $r1, r2, r3 \in \{1, \cdots, NP\}$; $i \neq r1 \neq r2 \neq r3$. The parameter $F \in [0,2]$ is a scaling factor that controls the amplification of the differential variation.

2) Recombination (crossover)

The crossover operator increases the diversity of the mutated individual by means of the combination of two solutions, mutant (v_i) and target (p^k_i) individuals. And we can get the trial individual u_i by the following formula:

$$u_{i,j} = \begin{cases} v_{i,j} & \text{, if } rand_j(0,1) \le C_r \\ p^k_{i,j} & \text{, else} \end{cases} \tag{2}$$

Where: $u_i = (u_{i,1}, \cdots, u_{i,n})$, $rand_j(0,1) \in [0,1]$ is a random number, and $C_r \in [0,1]$ is a crossover factor, which is used to control the probability of the replacement.

3) Selection

The $f(p^k_i)$ and $f(u_i)$ are calculated and the new individual p^{k+1}_i is selected by the following formula:

$$p^{k+1}_i = \begin{cases} u_i & \text{, if } f(u_i) \le f(p^k_i) \\ p^k_i & \text{, else} \end{cases} \tag{3}$$

B. Fitness variance of the population

With the evolution of population, the individual differences become smaller, and the individual position determines the individual fitness. Therefore, the fitness of all individuals could be used to determine the state of the populations.

Supposing the number of individuals is NP , f_i is the fitness of the ith individual, f_{avg} is the average fitness, σ^2 is the fitness variance of the population (PFV), σ^2 could be defined as follows:

$$\sigma^2 = \sum_{i=1}^{NP} \left(\frac{f_i - f_{avg}}{f} \right)^2 \tag{4}$$

Where, f is a normalized scaling factor which is used to limit σ^2 , its value could be get by the following formula:

$$f = \begin{cases} \max | f_i - f_{avg} | & \text{if}: \max | f_i - f_{avg} | > 1 \\ 1 & \text{else} \end{cases} \tag{5}$$

As known from the formula above, σ^2 reflects the aggregation of individuals. σ^2 is smaller, the population tends to converge, conversely, population is in the random searching stage. With the increasing of iteration, σ^2 become smaller. So, given a threshold T , if $\sigma^2 < T$, the DE algorithm could be considered in the stage of later search, which means the DE has been fallen into a local optimum.

III. RESEARCH ON CHAOTIC CHARACTERISTICS OF CAT MAP

A. Introduction of the chaotic cat map

The classic Arnold cat map is a two-dimensional invertible chaotic map [8] described by:

$$\begin{cases} x_{n+1} = (x_n + y_n) \pmod 1 \\ y_{n+1} = (x_n + 2y_n) \pmod 1 \end{cases} \tag{6}$$

Its matrix form is:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod } 1) = S \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod } 1)$$

Where, $S = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $x \,(\text{mod } 1)$ is used for the fractional parts of a real number x by subtracting or adding an appropriate integer.

The Lyapunov characteristic exponents of the map are the eigenvalues σ_1 and σ_2 of the matrix S , given by

$$\sigma_1 = \frac{(3+\sqrt{5})}{2} > 1 \quad , \quad \sigma_2 = \frac{(3-\sqrt{5})}{2} < 1$$

So, the map is known to be chaotic.

B. Chaos optimization and chaotic characteristics of the cat map

Chaos optimization method is one global optimization technology by using of the chaotic characteristics. It requires the chaotic sequence which with nice characteristics of ergodicity and distribution. There are many chaos optimizations using logistic map as the chaotic sequence generator [9], but the distribution of this sequence is uneven, which affect the global searching performance and efficiency seriously. Another chaotic sequence generator of tent map is proposed in reference [10]. While the tent map itself is easy to fall into small cycles or fixed points, the distribution of its chaotic sequence is uneven, especially when the length of this sequence is short.

The cat map is applied to DE optimization in this paper and its chaotic characteristics: ergodicity and distribution are examined.

The ergodicity of chaos map $f(x)$ is that: for each integrable function $\psi(x)$ and initial value x_0 , the following formula should be satisfied.

$$\lim_{n\to\infty}\frac{1}{N}\sum_{n=0}^{N}\psi[f^n(x_0)]=\int\rho(x)\psi(x)\mathrm{d}x \qquad (7)$$

Where, $\rho(x)$ is the orbit distribution density with the definition of $\rho(x)=\lim_{n\to\infty}(1/N)\sum_{n=1}^{N}\delta(x-x_n)$. f^n is n-compound of $f(x)$.

In reference [11], the ergodicity of the cat map is supplied. And the chaotic distribution characteristic of cat map is examined below.

The distribution maps of logistic map, tent map and cat map that in the range of $[0,1]$ with two different iteration times are shown below.

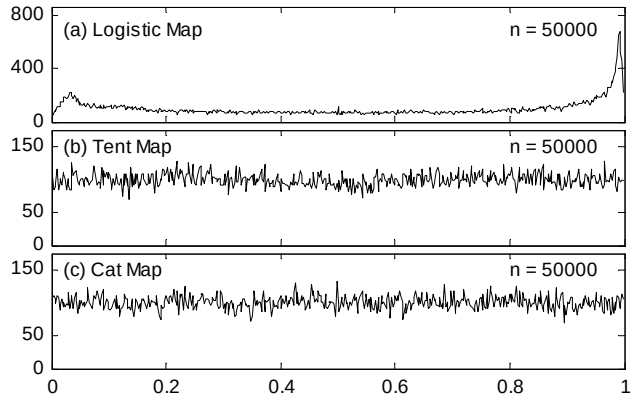

Figure 1. Distribution of maps iterations (iteration 50000 times)

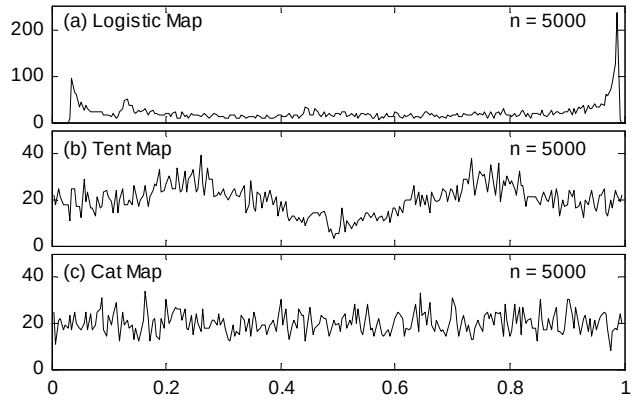

Figure 2. Distribution of maps iterations (iteration 5000 times)

It can be observed form Fig.1 and Fig.2 that while the number of iteration is 50000, the number that in the range of $[0.9,1]$ is over 12000, so the distribution of the logistic map is quite uneven. The distribution of tent map is relatively uniform when the iteration times is high, as shown in Fig.1 (b), but the distribution presents double humps when the iteration times is low, as shown in Fig.2 (b). So the distribution of this map is unstable. It can be observed from Fig.1(c) and Fig.2(c) that the distribution of the cat map is uniform and stable. Furthermore, the initial value of cat map can be 0 and 1, while the other two maps cannot, which is useful while the optimum value is in the boundary. Therefore, the chaotic characteristic of the cat map is better than that of logistic map and tent map.

IV. THE IMPROVED DE ALGORIGHM THAT USING CHAOTIC SEARCHING

A. The searching mechanism that based on chaotic cat map

Chaos is a nonlinear phenomenon that widespread in nature, its movement is ergodicity, randomness and initial sensitive. Chaotic movement could traverse all the status according to its own law. So the optimization search by use of chaotic variable is better than that using random variable. The chaotic searching performs better when in a small searching space and is easy to jump out of the local optimum. Therefore, the chaotic searching is used when DE fall into the local optimum, to improve the DE performance.

It is considered that the DE has fallen into the local optimum, when the PFV is less than a threshold T . At this time, supposing the best population is:

$$P^*=(p^*_1,p^*_2,\cdots,p^*_D)$$

Generating two random vectors $X_0=(x_{0,1},x_{0,2},\cdots,x_{0,D})$ and $Y_0=(y_{0,1},y_{0,2},\cdots,y_{0,D})$, where $x_{0,i},y_{0,i}\in[0,1]$ and taking them into formula (6). The two-dimensional cat map chaotic sequence x_i and y_i are calculated by iteration. Supposing $C_i=X_i=(x_{0,1},x_{0,2},\cdots,x_{0,D})$, $\{C_n,n=1,2,\cdots,N\}$ is the chaotic sequence group to use. Where, N is the limited maximum searching times.

Then the chaotic variable is transformed to the optimization variable space in the way of formula (8) and the optimized variable is $C'_n=(c'_{n,1},c'_{n,2},\cdots,c'_{n,D})$.

$$c_{i,j}{}'=p^*_i+r_i*c_{i,j} \qquad (8)$$

It is known from formula (8) that the chaotic variable is extended to the circular region where the best population P^* is the origin and r_i is the radius. The r_i is an adjust factor. When r_i is relative large, it is benefit to global searching and in a low convergence speed, conversely, it is limited in a small region near to the best population which is useful to improve the searching precision.

B. The improved DE that based on the chaotic searching

The chaos differential evolution algorithm (CDE) is proposed and its main idea is:

Step.1 The conventional DE is being executed until it is fall into the local optimum (the PFV is less than a threshold);

Step.2 The chaotic sequence from cat map is calculated and optimized;

Step.3 The random population is instead by the optimized chaotic sequence and the following operations of recombination and selection is executed.

The flow chart of CDE is shown as below.

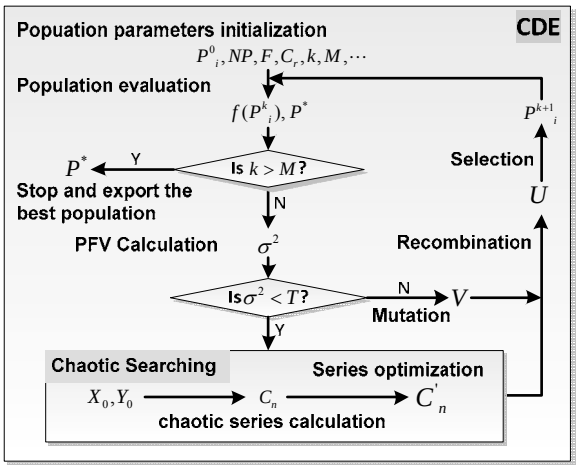

Figure 3. Flow chart of CDE

C. Numerical simulation of CDE

Four Benchmark questions are used to compare DE and CDE. These functions, their optimum status and optimum values are shown as below.

1) Sphere Model

$$f_1(x) = \sum_{i=1}^{n} x_i^2, |x_i| \le 10 (i = 1, 2, \cdots, n)$$

$$\min: f_1(x) = f_1(0, 0, \cdots, 0) = 0$$

2) Rosenbrock Function

$$f_2(x) = \sum_{i=1}^{n} [100(x_{i+1} - x_i)^2 + (x_i - 1)^2], |x_i| \le 50 \ (i = 1, 2, \cdots, n)$$

$$\min: f_2(x) = f_2(1, 1, \cdots, 1) = 0$$

3) Greiwank Function

$$f_3(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos(\frac{x_i}{\sqrt{i}}) + 1, |x_i| \le 600 \ (i = 1, 2, \cdots, n)$$

$$\min: f_3(x) = f_3(0, 0, \cdots, 0) = 0$$

4) Rastrigrin Function

$$f_4(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi x_i) + 10], |x_i| \le 5.12 (i = 1, 2, \cdots, n)$$

$$\min: f_4(x) = f_4(0, 0, \cdots, 0) = 0$$

Where, $f_1(x)$ is a single peak function, $f_2(x)$, $f_3(x)$ and $f_4(x)$ are all multiple peaks function which is difficult to find their global optimum. The initial parameters of the experiments are shown in the following table.

TABLE 1. INITIAL PARAMETERS OF THE EXPERIMENTS

Function	NP	M	N	T	F	C_r
$f_1(x)$	10	1000	100	0.01	0.5	0.1
$f_2(x)$	15	1500	150	0.5	0.6	0.2
$f_3(x)$	20	1000	150	0.01	0.5	0.1
$f_4(x)$	20	1000	150	0.01	0.5	0.1

Since the initial populations and the recombination process contains many random values, the experiment results is not repetitiveness. So twenty experiments are tested for each algorithm, and the statistical results of these experiments are shown as below.

TABLE 2. STATISTICAL RESULTS OF THESE EXPERIMENTS

Function	Algorithm	Maximum optimum	Minimum optimum	Mean optimum	Times *
$f_1(x)$	DE	2.34	1.15e-19	0.13	16
	CDE	1.10e-11	3.72e-20	8.35e-13	20
$f_2(x)$	DE	2.51e-6	1.53e-13	1.30e-7	19
	CDE	9.45e-19	1.38e-19	3.34e-19	20
$f_3(x)$	DE	0.01	0	0.01	17
	CDE	8.58e-14	0	4.29e-15	20
$f_4(x)$	DE	0.99	1.25e-11	0.19	18
	CDE	5.00e-10	6.23e-11	1.72e-10	20

* The times of algorithm accessing the global optimum

It can be observed from Table.2 that most of the maximum optimum, minimum optimum and mean optimum calculated by CDE is better than DE. The times of CDE accessing the global optimum is greater than that of DE. Therefore, both the optimization precision and the global searching ability are improved in CDE.

Furthermore, Fig.4 is the trace of the mean best fitness (MBF) in the optimization iteration. It can be known that because of the chaotic searching in the later stage, the convergence speed and precious of the CDE are enhanced greatly.

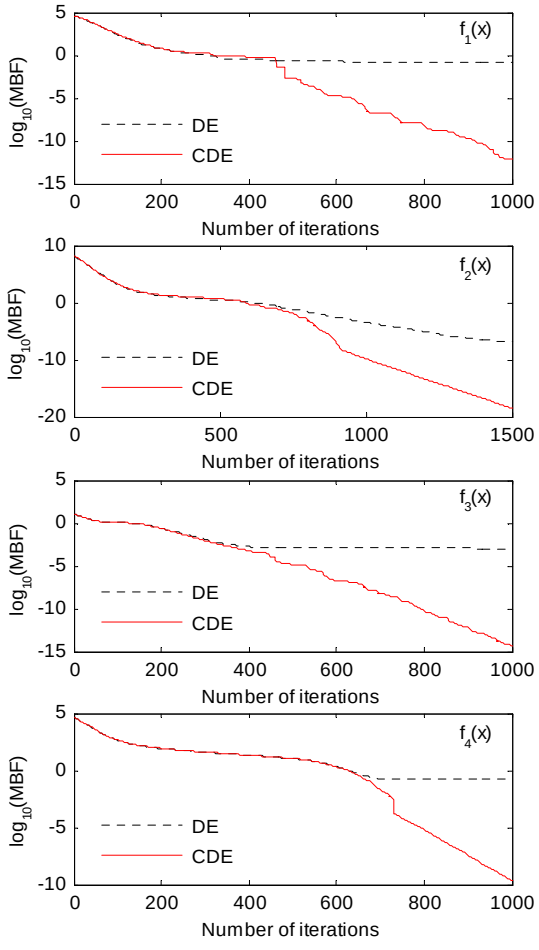

Figure 4. Optimization map of DE and CDE

V. SUPPORT VECTOR REGRESSION AND PARAMETERS ANALYSIS

The basic idea in SVR is to map the dataset $\{(x_1, y_1),\cdots,(x_l, y_l)\} \subset \boldsymbol{R}^n \times \boldsymbol{R}$ into a high dimensional feature space via non-linear mapping, wherein they are correlated linearly with the outputs [12]. The SVR formalism considers the following linear estimation function:

$$f(x, w) = \langle w, \phi(x)\rangle + b \tag{9}$$

Where, $w \in \boldsymbol{R}^n$ is weight vector, $b \in \boldsymbol{R}$ is a constant, $\phi(x)$ denotes a mapping function in the feature space.

Based on the principle of structural risk minimization, the SVR learning problem is recast as the optimization problem:

$$\text{Min: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi^*_i) \tag{10}$$

$$\text{Subject to:} \begin{cases} f(x_i) - y_i \le \varepsilon + \xi^*_i \\ y_i - f(x_i) \le \varepsilon + \xi_i \\ \xi_i, \xi^*_i \ge 0;\ i = 1,\cdots,l \end{cases}$$

Where, C is the regularization constant used to specify the trade-off between the empirical risk and regularization term. Two positive slack variables, ξ_i and ξ^*_i, $i = 1,2,\cdots,n$ can be used to measure the deviation from the boundaries of the ε-insensitive zone. That is, they represent the distance from actual values to the corresponding boundary values of ε-insensitive zone. Both C and ε are user-determined parameters. And a schematic representation of the SVR using ε-insensitive loss function is illustrated in Fig.5.

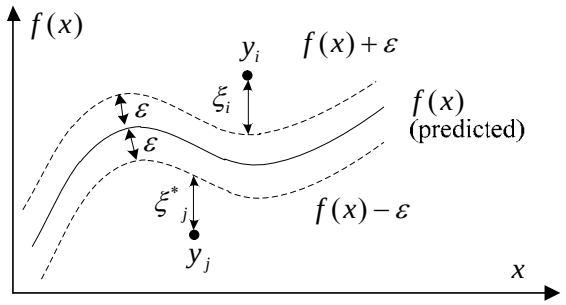

Figure 5. Sketch map of the SVR using ε-insensitive loss function

By using Lagrange function and Karush-Kuhn-Tucker conditions to the (10), it thus yields the following dual optional form:

$$\begin{aligned} \text{Max}: &-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha^*_i - \alpha_i)(\alpha^*_j - \alpha_j)K(x_i, x_j) \\ &+\sum_{i=1}^{l} y_i(\alpha^*_i - \alpha_i) - \varepsilon\sum_{i=1}^{l}(\alpha^*_i + \alpha_i) \end{aligned} \tag{11}$$

Subject to the constraints:

$$\begin{cases} \sum_{i=1}^{k}(\alpha_i - \alpha^*_i) = 0 \\ 0 \le \alpha_i, \alpha^*_i \le C,\quad i = 1,\cdots,l \end{cases}$$

Where, the Lagrange multipliers α_i, α^*_i satisfy the equality $\alpha_i\alpha^*_i = 0$. The Lagrange multipliers are calculated and an optimal desired weight vector of the regression hyper-plan is: $w = \sum_{i=1}^{l}(\alpha^*_i - \alpha_i)\phi(x_i)$. Hence, the general form of the SVR-based regression function can be written as:

$$f(x) = \sum_{i=1}^{l}(\alpha^*_i - \alpha_i)K(x, x_i) + b \tag{12}$$

Where, $K(x_i, x_j)$ is called the kernel function and the values of it equals the inner product of two vectors, x_i and x_j, in the feature space $\phi(x_i)$ and $\phi(x_j)$; that is, $K(x_i, x_j) = \phi(x_i)\phi(x_j)$.

Any function that meets Mercer's condition can be used as the kernel function [13]. The radial basis function (RBF) is applied in this paper as kernel function, which is defined as:

$$K(x_i, x_j) = \exp\left(-\|x_i - x_j\|_2^2 \big/ 2\sigma^2\right)$$

Where, σ^2 is the width of the RBF.

As just mentioned, when to designan effective model, the values of the two essential model parameters (C, ε) and one kernel function parameter (σ^2) in SVR have to be chosen carefully in advance. The regularization parameter C determines the tradeoff cost between minimizing the training error and minimizing model complexity, which will reduce the generalization capability when it was set too small or excessive. The parameter ε defines the non-linear mapping from the input space to some high-dimensional feature space which determines the number of support vectors. And σ^2 reflect correlation of the support vector, which also determines both the generalization capability and the prediction accuracy.

VI. SVR PREDICTION MODEL BASED ON CDE ALGORITHM

The schematic diagram of SVR parameters selection based on the CDE algorithm(CDE-SVR) is presented in Fig.6. One SVR model with kernel function of RBF has three parameters: σ, C and ε. The fitness function is mean square deviation (MSE) which is treated as a direct response to the performance of SVR.

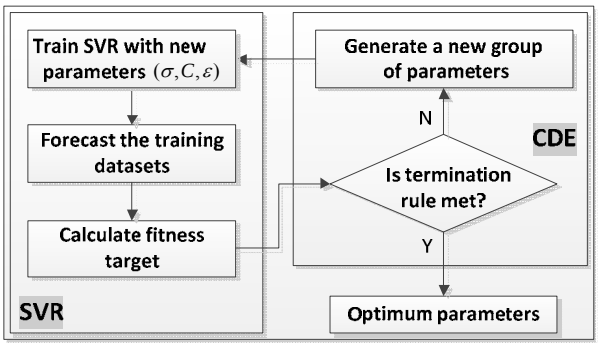

Figure 6. Flowchart of SVR parameters selection using CDE algorithm

The concrete steps of CDE-SVR parameters selection are as follows:

Step 1: Set the number of individuals, crossover factor C_r, and the parameters of individuals. Set the termination conditions: maximum allowable number of iterations and the fitness error limitation (maximum MSE).

Step 2: Generate the mutant v_i with (9).

Step 3: Get the trial individual u_i by the recombination of v_i and $p^k{}_i$ by (10).

Step 4: Get the MSE with the new parameters u_i and execute selection operation using (11).

Step 5: Select the smallest MSE and record this individual.

Step 6: Check whether termination criterion is met. If yes, stop and output the optimum parameters; otherwise, calculate the next generation individual $p^{k+1}{}_i$ and go to step 2.

After that, the SVR model with these optimum parameters is applied to the test datasets, and the common flowchart of one time series prediction based on the DE-SVR is as follows.

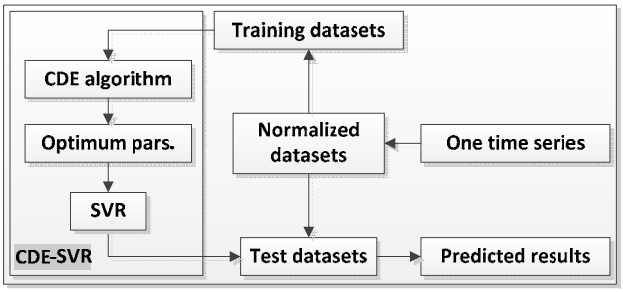

Figure 7. Flowchartof time series prediction based on CDE-SVR

VII. ITS APPLICATION IN CHAOTIC SERIES PREDICTION

A. Chaotic Chens time series and state-space reconstruction

Chaotic systems are deterministic systems which are capable of generating irregular and complex behavior, depending on system parameters. Since chaotic systems are deterministic, the current state of the system is sufficient for determining the entire future evolution of the system. However, as chaotic systems possess positive Lyapunov exponent(s), only short-term predictability is possible, due to exponential divergence [14]. The chaotic signal of Chens dynamical system is researched in this paper which is described by the following differential equations:

$$\begin{cases} \dot{x} = a(y-x) \\ \dot{y} = (c-a)x - xz + cy \\ \dot{z} = xy - bz \end{cases} \tag{13}$$

Where x, y and z are the state variables, the a, b and c are three positive real constants. This system has a chaotic solution when $a=35, b=3, c=28$ .By using the fourth-order Runge-Kutta method with time step 0.01, the time series for X-component (N_{total} =1600) is carried out and shown in Fig.4. Here, the first 200 points are discarded to remove transients, and 1000 points (from 201 to 1200) are selected as training datasets and 200 points (1201-1400) for test datasets.

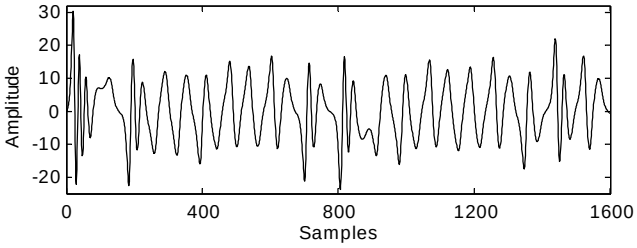

Figure 8. The X-component of Chens chaotic time series

State-space reconstruction is fundamental to both system characterization and forecasting, and the expression of reconstruction matrix $\boldsymbol{X}$ is as follows:

$$\boldsymbol{X} = \begin{bmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ x(2) & x(2+\tau) & \cdots & x(2+(m-1)\tau) \\ \cdots & \cdots & \cdots & \cdots \\ x(n) & x(n+\tau) & \cdots & x(n+(m-1)\tau) \end{bmatrix}$$

Where, m is the embedding dimension and is the embedding delay. The embedding dimension is found using the method of Global False Nearest Neighbors (GFNN) , whilst the embedding time delay is found from the first minimum in the mutual information.

B. The prediction of chaotic sequence using CDE-SVR

In the experiment the parameters of CDE algorithm is set to: $C_r = 0.5$, $\alpha = 0.5$, $NP = 25$, and $n_p = 3$ is the number of SVR parameters, the initial training parameters are: $\sigma^2 = 1.4$, $C = 2$ and $\varepsilon = 0.1$. The optimization termination rule: maximum number of iterations is 200. After training, the prediction of test datasets was carried out by SVR with the optimum parameters.

In order to verify the effectiveness and superiority of the SVR parameters selection based on the CDE algorithm, the Grid-search SVR and PSO-SVR were also implemented. The optimal parameters and other

implementation results obtained are presented in Table.3. And the prediction results with CDE-SVR are shown in Fig.9. Where, the RMSE is the mean of MSE.

TABLE 3. IMPLEMENTATION RESULTS USING DIFFERENT METHODS

Model	***Grid-SVR***	***PSO-SVR***	***CDE-SVR***
RMSE	0.112	0.082	0.074

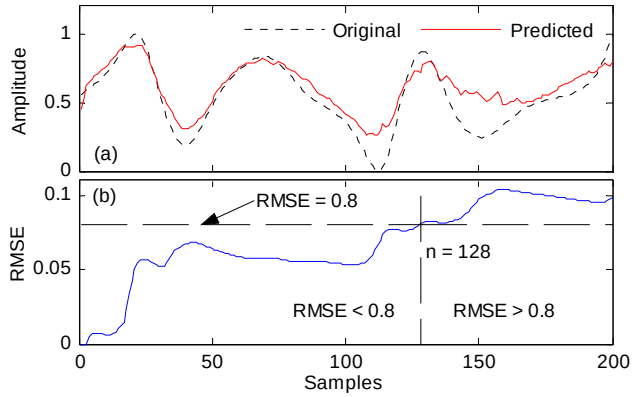

Figure 9. Predicted and actual series using proposed CDE-SVR

Table 3 clearly demonstrated the effectiveness of the proposed CDE-SVR for the prediction of the Chens time series. Compared with the conventional grid searching approach and PSO-SVR, the prediction accuracy of CDE-SVR is higher.

As well known, the chaotic series is difficult to predict. If we set a threshold of RMSE=0.8, which is the maximum error could be accepted; only 128 points ahead is meaningful.

VIII. ITS APPLICATION IN BOUNDARY EXTENTION OF MECHANICAL VIBRATION SIGNALS

Oil pump is the key equipment in the process of oil transportation, its structure are shown in Fig.10. Most of its faults could be reflected in its body vibration signals. At present, the approach of EMD could be used to extract its fault features effectively [15]. But there exists strong end effects in EMD [16], so the boundary of signal should be extended firstly, and the SVR is an appropriate approach to achieve this work.

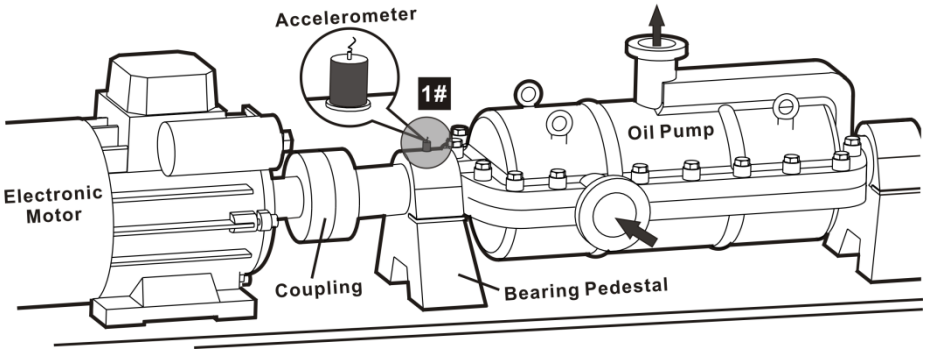

Figure 10. Device structure and the measurement points

The vibration signal shown in Fig.11 was measured in position 1#. The sampling rate is 4096Hz and the instantaneous rotational speed is about 2310rpm (38.5Hz).

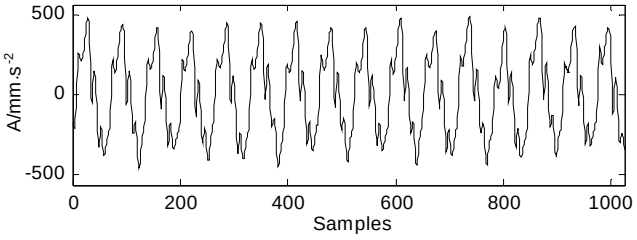

Figure 11. Oil pump vibration signal in 1#

The total length of the signal is 1024 points. In order to examine the extension results, the points of 1~700 are selected as train dataset, and points of 701~900 are set as test dataset. The extension results (prediction results) and the RMSE are shown in Fig.12.

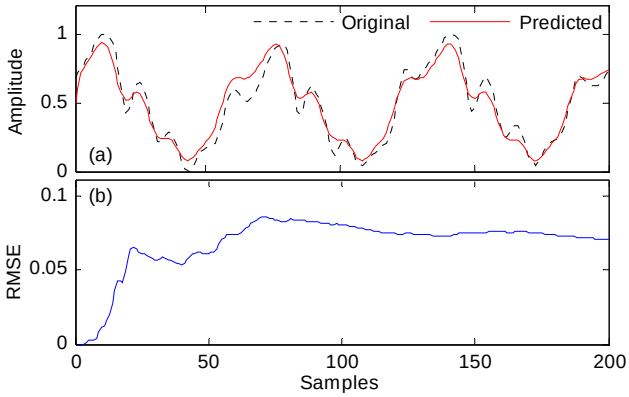

Figure 12. Predicted and actual series using proposed CDE-SVR

As can be seen from Fig.12 that the RMSE between the original signal and predicted signal is small and the period features of the signal is well preserved, which could meet the requirements of the boundary extension. Therefore, the effectiveness of the CDE-SVR is verified and this approach could be applied to the boundary extension of the mechanical vibration signals.

IX. CONCLUSION

A chaos differential evolution optimization algorithm (CDE) that based on the chaotic characteristics analysis of the cat map is proposed in this paper. When the DE fall into the local optimum, the chaotic sequence generated by cat map is used to update the population. The numerical simulation verifies that the optimization speed and precision of the CDE algorithm is better than that of convectional DE algorithm.

Then the CDE is used to SVR and a new optimal prediction model of CDE-SVR is proposed in this paper. The model is applied to the short time prediction of chaotic time series. The results verify that, compared with the prediction model optimized by other method, the prediction accuracy of CDE-SVR is higher than others. The model is also applied to the boundary extension of the mechanical vibration signals. The application results also verify the effectiveness of this new prediction model.

ACKNOWLEDGMENT

This project was supported by National Natural Science Foundation of China (Grant No. 51005247), National High-tech R&D Program of China (Grant No. 2008AA06Z209) and Special Items Fund of Beijing Municipal Commission of Education of China.

REFERENCES

[1] R.M. Storn and K.V. Price. Differential evolution-a simple evolution strategy for global optimization over continuous space. J.Global Opt. 1997(11):341-359.

[2] Leandro dos Santos Coelho and Marianni V C. Improved differential evolution algorithms for handing economic dispatch optimization with generator constraints. Energy Conversion and Management, 2007, 48:1631-1639.

[3] Y.C. Lin, K.S. Hwang and F.S. Wang. Co-evolutionary hybrid differential evolution for mixed-integer optimization problems. Engineering Optimization, 2001,33(6):663-682.

[4] V.N. Vapnik, E. Levin and C.Y. Le. Measuring the VC Dimension of a Learning Machine. Neural Computation, 1994(6):851-876.

[5] P.W. Chen, J.Y. Wang and H.M. Lee. Model selection of SVMs using GA approach. IEEE International Joint Conference on Neural Networks 2004, Piscataway, NJ, IEEE Press, 2004:2035-2040.

[6] C.L. Huang and C.J. Wang. A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications, 2006,31(2):231-240.

[7] X.Y. Peng and H.X. Wu. Parameter selection method for SVM with PSO. Chinese Journal of Electronics, 2006,15(4):638-642.

[8] G.R. Chen, Y.B. Mao and C.K. Cui. A symmetric image encryption scheme based on 3D chaotic cat maps. Chaos, Solutions and Fractals, 2004,21(3):749-761.

[9] C. Choi and J.J. Lee. Chaotic local search algorithm. Artificial Life and Robotics, 1998,2(1):41-47.

[10] H.L. Qin and X.B. Li. A chaotic search method for global optimization on tent map. Electric Machines and Control, 2004, 8(1):67-70.

[11] F. Wang, Y.S. Dai and S.S. Wang. Modified chaos-genetic algorithm. Computer Engineering and Applications, 2010, 46(6):29-32.

[12] N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Publishing House of Electronics Industry, Beijing, 2004.

[13] V.N. Vapnik. The Nature of Statistical Learning Theory. Springer, New York, 2000.

[14] L.T. Por and S. Puthusserypady. Chaotic time series prediction and additive white Gaussian noise. Physics Letters A. 2007,365:309-314.

[15] J.S. Cheng, D.J. Yu and Y. Yang. Fault diagnosis for rotor system based on EMD and fractal dimension. China Mechanical Engineering, 2005,16(12):1088-1091.

[16] J.S. Cheng, D.J. Yu and Y. Yang. Discussion of the end effects in Hilbert-Huang transform. Journal of Vibration and Shock, 2005,24(6):40-47.

Wei Liang was born in Shanxi, China, 1978. He received the Ph.D. in Mechanical Engineering from the China University of Petroleum, in 2005. Currently he serves as an associate professor in the China University of Petroleum. His research interests include dynamic safety assessment of pipeline and machine set.

Laibin Zhang was born in Anhui, China, 1961. He received the Ph.D. in Mechanical Engineering from the Petroleum University, China, in 1991. Currently he serves as a professor in the China University of Petroleum. His research interests include electromechanical integration and machinery faults diagnosis.

Mingda Wang was born in Shandong, China, 1984. He is currently a Ph.D student in China University of Petroleum. His research interests include system optimization and machinery faults diagnosis.