Proceedings of the 40th Chinese Control Conference
July 26-28, 2021, Shanghai, China

Simulated Annealing Strategy in Chaotic Neural Network with Chebyshev Polynomials

Yaoqun Xu[1,2], Zhenhua Yang[1], Xinxin Zhen[1]

1. School of Computer and Information Engineering, Harbin University of Commerce, Harbin, 150028

2. Institute of System Engineering, Harbin University of Commerce, Harbin, 150028
E-mail: xuyq@hrbcu.edu.cn

Abstract: By combining the Chebyshev polynomials and the Sigmoid function into a new non-monotonic excitation function, a new transient chaotic neural network model is constructed. The dynamic characteristics of the single neuron are analyzed by the time evolution graph and the inverted bifurcation graph of the largest Lyapunov exponent. Verify the rationality of the additional energy term of the constructed model. The effectiveness of the model is verified by nonlinear function optimization and traveling salesman problem. The simulation results show that the newly constructed model can effectively avoid the problem of converging to a local minimum in the optimization process, and can effectively solve the combinatorial optimization problem.

Key Words: Chebyshev polynomials, Chaotic neural network, Lyapunov exponents, Traveling salesman problem

1 Introduction

Inspired by the chaotic dynamic system, Hopfield first studied neural networks from the perspective of dynamic systems. By introducing the energy function, the minimal solution of the energy function was successfully mapped to the equilibrium state of the network, and the optimization problem of the neural network was transformed into the evolution problem of the dynamic system, which pioneered the application of the neural network in computer science. Although Hopfield neural network can successfully solve combinatorial optimization problems, due to its use of gradient descent strategy, in the process of optimization, there will be infeasible solutions and local minimum solutions that converge to the optimization problem. This limits the application of Hopfield neural network. In order to improve the optimization performance of Hopfield neural network, many scholars, inspired by the chaotic characteristics of biological neurons, introduced chaotic dynamics into Hopfield neural network and proposed a chaotic neural network model. This model uses the global ergodicity of chaotic dynamics to effectively avoid the optimization process from falling into a local minimum solution. In recent years, many scholars have discovered that some chaotic neural networks with non-monotonic activation functions are better than Chen's[1] chaotic neural networks in solving optimization problems. Potapov[2] proposed that the network can accelerate into a chaotic state under the action of a non-monotonic increasing function. Shuai[3] et al. proposed that the activation function of neural network can be diversified. He et al. proposed a new chaotic neural network model by introducing transient chaos and time-varying gain into the excitation function. Hu et al. proposed a new chaotic neural network whose activation function is a weighted sum of variable frequency sine[4] function and Sigmoid function. Xiu et al. proposed a network model composed of a linear combination of Sigmoid function and Gauss function as the activation function. Xu[5] et al. proposed an excitation function that is the sum of the Sigmoid function and the continuous wavelet function, and verified the advantages of the continuous wavelet function through comparison and simulation with the Gauss function. Ye et al. proposed to use the linear combination of Legendre[6] function and Sigmoid with higher nonlinearity and better function approximation ability as a new excitation function, and verified the effectiveness of the model through function optimization and traveling salesman problem. In this paper, the combination of Chebyshev polynomial with higher approximation ability and Sigmoid function is selected as the new activation function to construct a new chaotic neuron model. The chaotic dynamics of neurons are analyzed through the inverted bifurcation diagram and the time evolution diagram of the largest Lyapunov exponent. On this basis, a new transient chaotic neural network is constructed, and the effectiveness of the model is verified by function optimization and traveling salesman problem[7].

2 SCF Transient Chaotic Neural Network Model

2.1 Chebyshev Polynomials

The Chebyshev polynomials[8] is derived from the expansion of the cosine function and sine function of multiple angles, and is a special function in computational mathematics. The first type of Chebyshev polynomials is a set of orthogonal polynomials, which are solutions to

This work is supported by Nature Science Foundation of Heilongjiang Provice (LH2019F044)

differential equations. The Chebyshev polynomial[9] of the first kind is described as:

$$T_n(x)=\frac{1}{4\pi i}\oint\frac{(1-t^2)t^{-n-1}}{(1-2tx+t^2)}dt, n=0,1,2\cdots \quad (1)$$

When n=1,2,3,5, the image of Chebyshev polynomials is shown in Fig. 1.

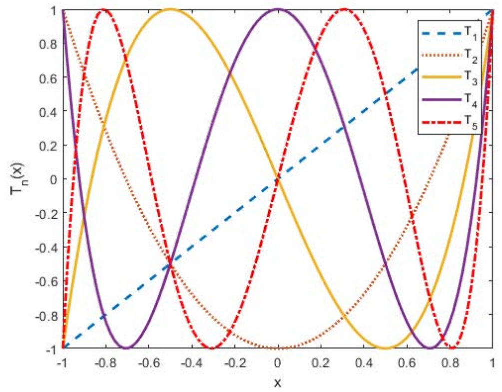


Fig. 1: Chebyshev polynomials image

2.2 Chaotic Dynamics of Single Neuron

Combine the Chebyshev polynomials and the Sigmoid function into a new activation function to construct a new transient chaotic neuron model, call this model the Sigmoid-Chebyshev-Function model, referred to as SCF model, as shown:

$$x(t)=f(y(t)) \quad (2)$$

$$y(t+1)=ky(t)-z(t)(x(t)-I_0) \quad (3)$$

$$z(t+1)=(1-\beta)z(t) \quad (4)$$

$$f(u)=\begin{cases}\sum_{i=1}^{n}\frac{\lambda_i}{2}(T_{2i+1}(u)+1)+(1-\sum_{i=1}^{n}\lambda_i)S(u), -1\le u\le 1\\ S(u), otherwise\end{cases} \quad (5)$$

$$T_n(x)=\frac{1}{4\pi i}\oint\frac{(1-t^2)t^{-n-1}}{(1-2tx+t^2)}dt, n=0,1,2\cdots \quad (6)$$

$$s(u)=1/(1+\exp(-u/\varepsilon_0)) \quad (7)$$

In this model, $x(t)$ is the activation function, that is, the output of the neuron at time t; $y(t)$ is the internal state of the neuron at time t; the value range of k is $0\le k\le 1$, indicating the ability of the neuron to retain its internal state; $f(u)$ is the activation function of the model, which is composed of the Chebyshev polynomials and the Sigmoid function; ε_0 is the steepness coefficient of the activation function; β is the simulated annealing parameter; $z(t)$ is the self-feedback Connection item; I_0 is a positive parameter; λ is a combination parameter, and the value range is $0\le\lambda\le 1$. The activation function in the model is a combination of Chebyshev polynomial and Sigmoid function of $\lambda=6/10$ and $n=3$, as shown in Fig. 2.

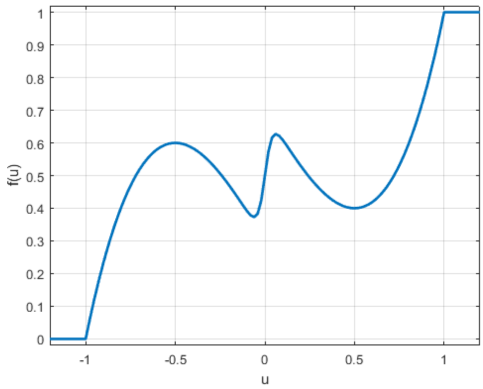


Fig. 2: Image of the activation function

The chaotic dynamic characteristics of neurons can be analyzed by the time evolution diagram of the largest Lyapunov exponent and the inverted bifurcation diagram. Because chaotic motion is extremely sensitive to initial value conditions, two orbits generated by similar initial values are separated exponentially over time. The Lyapunov exponent is an index used to quantitatively describe the degree of orbital separation. The calculation formula of Lyapunov exponent is as follows:

$$\lambda_L=\lim_{n\to\infty}\frac{1}{n}\sum_{n}^{n-1}\ln\left|\frac{dy(t+1)}{dy(t)}\right| \quad (8)$$

For the SCF chaotic neuron model, set n in the Chebyshev polynomial in the activation function to 3, and there are the following equations.

$$\frac{dy(t+1)}{dy(t)}=k-z(t)\frac{dx(t)}{dy(t)} \quad (9)$$

$$\frac{dx(t)}{dy(t)}=\frac{\lambda}{2}(12y(t)^2-3)+\frac{1-\lambda}{\varepsilon_0}(s(y(t)))(1-s(y(t))) \quad (10)$$

In one-dimensional mapping, $\lambda_L>0$ shows that the adjacent orbital index is separated, and the motion orbit is unstable in every part, and is in a chaotic state, where the magnitude of the value indicates the strength of the chaotic state. $\lambda_L<0$ shows that the adjacent orbital motions eventually move closer together, the phase volume shrinks and the motion is stable. $\lambda_L=0$ indicates that the model is at the critical value of stability and chaos.

In order to make the neuron show chaotic characteristics, set the parameters as follows:

$\varepsilon_0=0.02$, $y(1)=0.1$, $z(1)=0.98$, $k=1$, $I_0=0.57$, β=0.001, $\lambda=3/10$.

The time evolution diagram and inverted bifurcation diagram of the neuron's largest Lyapunov exponent are shown in Fig. 3.

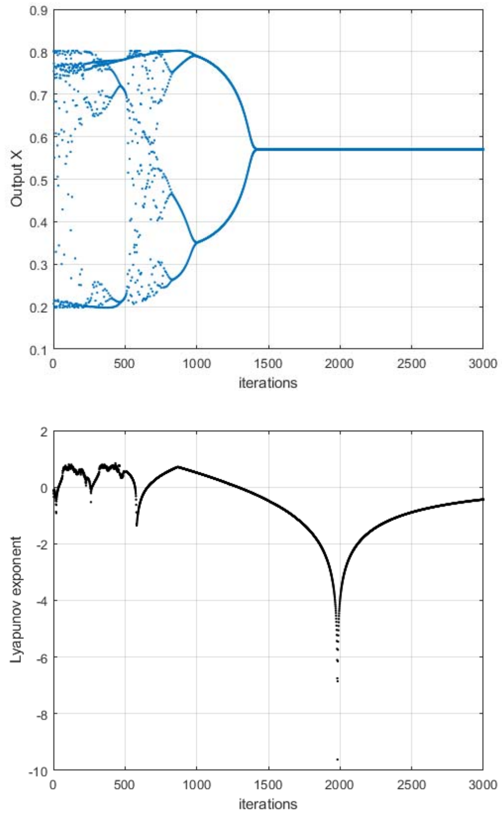


Fig. 3: The inverted bifurcation diagram and the time evolution diagram of the largest Lyapunov exponent of the single neuron for $\lambda = 3/10$

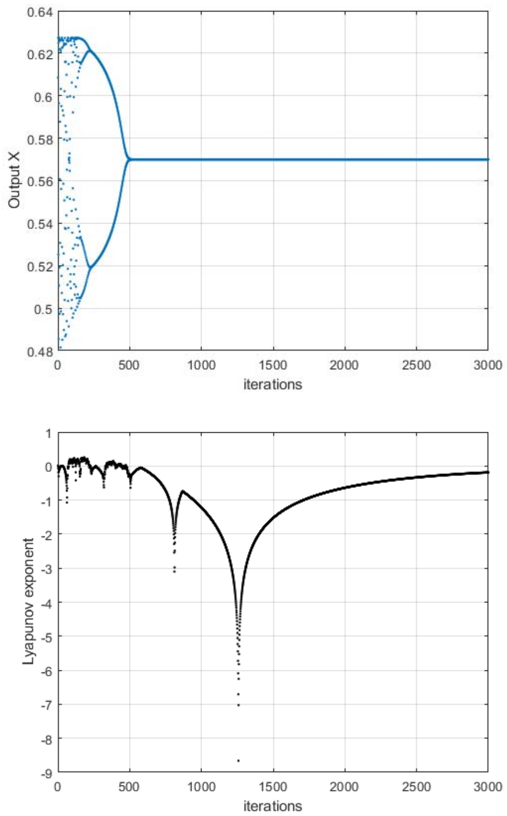


Fig. 4: The inverted bifurcation diagram and the time evolution diagram of the largest Lyapunov exponent of the single neuron for $\lambda = 6/10$

In order to observe the influence of parameter λ on the chaotic characteristics of neurons, set $\lambda = 6/10$. Other parameters are consistent with the above parameters. The time evolution diagram and inverse bifurcation diagram of the maximum Lyapunov exponent are shown in Fig. 4.

Through the time evolution graph and the inverse bifurcation graph of the largest Lyapunov exponent, it can be seen that the SCF chaotic neuron model has transient chaotic behavior, and the neuron model can effectively avoid the model from converging to a local minimum by using the chaotic dynamic characteristics. When the parameter λ of the activation function in the model is 0.3 and 0.6, it can be seen from the simulation results that the former parameter λ_L is greater than 0 more than the latter, indicating that the chaotic state of the former is stronger than the latter. The former is also in the chaotic state for longer than the latter, indicating that the parameter λ in the activation function will affect the time for the model to exit the chaotic state. When the maximum Lyapunov exponent is 0, the inverted bifurcation of the neuron occurs, which is consistent with the inverted bifurcation of the neuron. Through analysis, it is found that the parameter λ affects the strength of the chaotic state of the chaotic neuron and the time for the neuron to exit the chaotic state. The model can control the strength of the chaotic search characteristics of the network by setting the size of the parameter λ to better meet the actual optimization problem.

2.3 SCF Transient Chaotic Neural Network Model

In favor of the above transient chaotic neuron model, the following transient chaotic neural network model is constructed.

$$x_i(t) = f(y_i(t)) \tag{11}$$

$$y_i(t+1) = ky_i(t) + \gamma\left[\sum_{\substack{i=1 \\ j\neq i}}^{n} w_{ij}x_j(t) + I_i\right] - z_i(t)(x_i(t) - I_0) \tag{12}$$

$$z_i(t+1) = (1-\beta)z_i(t) \tag{13}$$

$$f(u) = \begin{cases} \sum_{i=1}^{n} \frac{\lambda_i}{2}(T_{2i+1}(u)+1) + (1-\sum_{i=1}^{n}\lambda_i)S(u), -1 \le u \le 1 \\ S(u), otherwise \end{cases} \tag{14}$$

$$s(u) = 1/(1+\exp(-u/\varepsilon_0)) \tag{15}$$

$$T_n(x) = \frac{1}{4\pi i}\oint \frac{(1-t^2)t^{-n-1}}{(1-2tx+t^2)}dt, n = 0,1,2\cdots \tag{16}$$

Where $y_i(t)$ is the internal state of the i-th neuron at time t; $x_i(t)$ is the activation function, that is, the output of the i-th neuron at time t, and $f(u)$ is the activation function of the network; β are segmented simulated annealing parameter; w_{ij} is the weight of neuron j and i, where $w_{ij} = w_{ji}$, $w_{ii} = 0$; λ is the combined parameter, and the value range is $0 \le \lambda \le 1$; ε_0 is the steepness coefficient of the activation function; I_i is the input deviation of the neuron; k is the damping factor of the neural diaphragm,

$0 \le k \le 1$, which reflects the ability of the network to retain memory or forget the internal state; γ is the input positive scale parameter representing the energy function of the chaotic dynamics the influence of characteristics; I_0 is a positive parameter; T_n is the definition of Chebyshev polynomials.

2.4 Piecewise Simulated Annealing Strategy

The annealing function of the chaotic neural network affects the decay process of the network self-feedback, which in turn affects the strength of the network's chaotic state.

A proper annealing function can effectively use the chaotic dynamics to improve the optimization performance of the network. Choose the following piecewise linear annealing strategy.

$$z(t+1) = \begin{cases} (1-\beta_1)z(t), z(t) > 0.3 \times z(1) \\ (1-\beta_2)z(t), otherwise \end{cases} \tag{17}$$

Set the SCF model parameters as follows:

$\varepsilon_0 = 0.02$, $y(1) = 0.1$, $z(1) = 0.98$, $k = 1$, $I_0 = 0.57$, β=0.001, $\lambda = 3/10$.

The inverted bifurcation diagram of the neuron is shown in Fig. 5.

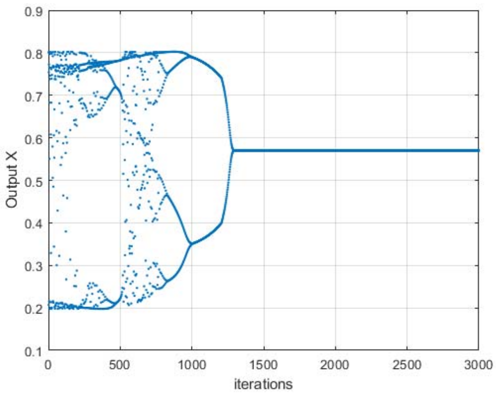


Fig. 5: The inverted bifurcation diagram of the single neuron for $\lambda = 3/10$

Fig. 4 uses an exponential annealing strategy for neurons, and Fig. 5 uses a piecewise linear annealing strategy. Through comparison and simulation, we find that different annealing strategies show different dynamic characteristics. Choosing a suitable annealing strategy can mention the optimal performance of the network.

2.5 Additional Energy Function of SCF

Chaotic neural network is produced to overcome the tendency of Hopfield network to converge to a local minimum in the optimization process. As the types of chaotic neural networks increase, a set of theory is needed to guide the research of chaotic neural networks. The emergence of the unified framework theory[10] provides theoretical support for the research and construction of chaotic neural networks. The unified framework theory[11] studies the optimization principle of chaotic neural networks from the direction of the additional energy term of the energy function. This theory takes the sum of the energy function of the Hopfield network and the additional energy function as the energy function of the network, and combines it with the dynamic equation of the hybrid neural network to introduce the specific expression form of the additional energy term, which is a reasonable construction for the new chaotic neural network.

$$\frac{dy_i(t)}{dt} = -\frac{\partial E}{\partial x_i(t)} \tag{18}$$

$$E = E_{HOP} + H \tag{19}$$

$$E_{HOP} = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_i(t)x_j(t) - \sum_{i=1}^{N} I_i x_i(t) + \frac{1}{\tau}\sum_{i=1}^{N}\int_0^{x_i(t)} f^{-1}(\xi)d\xi \tag{20}$$

$$\begin{aligned} \frac{dy_i(t)}{dt} &= -\left(\frac{\partial E_{HOP}}{\partial x_i(t)} + \frac{\partial H}{\partial x_i(t)}\right) \\ &= -\frac{y_i(t)}{\tau} + \left[\sum_{j=1}^{n} w_{ij} + I_i\right] - \frac{\partial H}{\partial x_i} \end{aligned} \tag{21}$$

Unify equation 21 and equation 2 to obtain the additional energy term of the SCF model, as shown in Equation 21.

$$H_{scf} = \lambda(t)\sum_{i}^{n}\int_0^{x_i} (x_i(t) - I_0)dx_i \tag{22}$$

Substituting equation 22 into equation 21 yields equation 23, and performs Euler discretization of equation 23 to yield equation 24.

$$\frac{dy_i(t)}{dt} = -\frac{y_i(t)}{\tau} + \sum_{j=1}^{n} w_{ij}x_j(t) + I_i - \lambda(t)(x_i(t) - I_0) \tag{23}$$

$$\begin{aligned} y_i(t+1) &= (1 - \frac{\Delta t}{\tau})y_i(t) + \\ &\Delta t\left[\sum_{j=1}^{n} w_{ij}x_j(t) + I_i\right] - \Delta t\lambda(t)(x_i(t) - I_0) \end{aligned} \tag{24}$$

If the following relationship are satisfied:

$k = 1 - \Delta t / \tau$, $\alpha = \Delta t$, $z(t) = \lambda(t)$, then equation 24 and equation 12 are equivalent. Therefore equation 22 is a reasonable additional energy term for the SCF model.

3 Application

3.1 Continuous Function Optimization Problems

Choose the following optimization function[12].

$$\begin{aligned} f(x_1, x_2) = (x_1 - 0.7)^2\left[(x_2 + 0.6)^2 + 0.1\right] + \\ (x_2 - 0.5)^2\left[(x_1 + 0.4)^2 + 0.15\right] \end{aligned} \tag{25}$$

The minimum value of the above optimization function is 0; the minimum point is (0.7,0.5); the local minimum point is (0.6,0.4) and (0.6,0.5). We set parameters as follows:

$\varepsilon_0 = 0.02$, $k = 1$, $\gamma = 0.1$, $y(1) = 0.383$, $y(2) = 0.283$, $z(1) = 0.98$, $z(2) = 0.98$, $I_0 = 0.57$, β=0.002, $\lambda = 3/10$ and $\lambda = 4/10$.

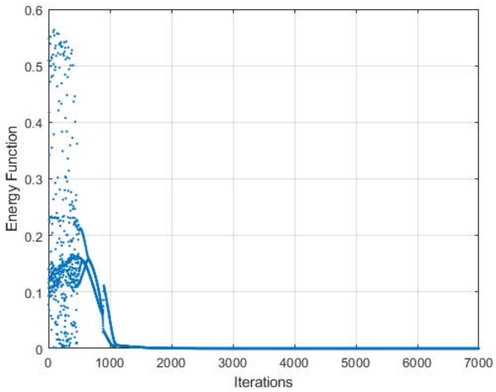


Fig .6: Time evolution diagram of the energy function for $\lambda = 3/10$

When $\lambda = 3/10$, the energy time evolution diagram of the optimization function f is shown in Fig. 6

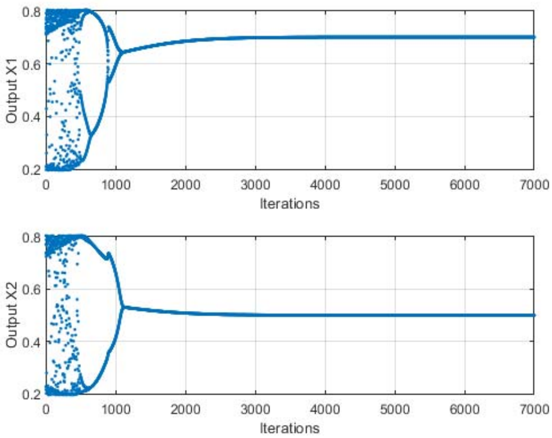


Fig. 7: The optimal solution of x_1, x_2 for $\lambda = 3/10$

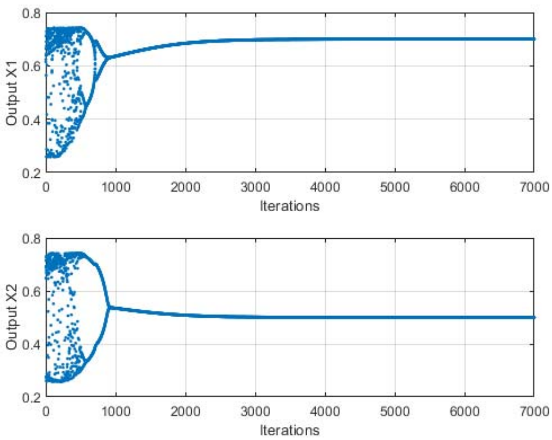


Fig .8: The optimal solution of x_1, x_2 for $\lambda = 4/10$

When $\lambda = 3/10$ and $\lambda = 4/10$ the optimal solution of x_1, x_2 is shown in Fig. 7 and Fig. 8.

Through the above simulation experiments, it is found that as the parameter λ increases, the time for neurons to exit the chaotic search is faster, and the parameter λ affects the chaotic search process of the neuron. By setting a suitable λ ,the optimization ability of the network can be improved.

3.2 Application to 10-city and 20-city TSP

In the field of combinatorial optimization, TSP[13] problem is a very typical problem. Due to the difficulty of its solution, many scholars have been looking for an efficient and practical method to solve this problem The TSP problem can be simply described as "finding a shortest loop containing all n cities, where each city must be visited only once". $(n-1)!/2$ is the number of possible paths for TSP problems in n cities.

In this paper, the SCF transient chaotic neural network model is applied to solve the TSP[14] problem of 10 cities and 20 cities. The simulation results show that this model has a good ability to solve the TSP problem. Take the energy function as:

$$E = \frac{A}{2}\sum_{i=1}^{n}(\sum_{i=1}^{n} x_{ij} - 1)^2 + \frac{B}{2}\sum_{j=1}^{n}(\sum_{i=1}^{n} x_{ij} - 1)^2 + \frac{D}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} d_{ik} x_{ij}(x_{k,j+1} + x_{k,j-1}) \tag{26}$$

Where, x_{ij} refers to city i visited in order j, parameter $A = B$, d_{ik} represents the distance between city i and city k, and the shortest effective path is equivalent to a global minimum energy value. We select the normalized 10 city coordinates as follows:

(0.4,0.4439);(0.2439,0.1463);(0.1707,0.2293);(0.2293,0.716);(0.5171,0.9414);(0.8732,0.6536);(0.6878,0.5219);(0.8488,0.3609);(0.6683,0.2536);(0.6195,0.2634).

We initialize the parameters as follows:

$A = B = 1$, $D = 2$, $\gamma = 0.27$, $k = 1$, $I_0 = 0.5$, $\varepsilon = 0.02$, $z(1) = 0.8$, β=0.0015, $\lambda = 6/10$.

The shortest path that satisfies the TSP problem of 10 cities under this coordinate is 2.6776. The program runs a total of 5 times, and 200 different initial conditions of x_{ij} are generated randomly in the interval [- 0.1, 0.1]. The results are summarized in Table 1.

Table 1. Results of SCF Model Solving TSP Problem in 10 Cities

Legal path	Optimal path	Legal ratio(%)	Optimal ratio(%)
200	196	100	98
200	198	100	99
199	196	99.5	98
198	195	99	97.5
200	198	100	99

As see from Table 1, this model has a higher performance in solving the 10-city TSP problem under the above-mentioned parameter settings.

In order to study the ability of the SCF model to solve more urban traveling salesman problems, we use the SCF model to solve the 20-city TSP[15] problem, and select the normalized 20-city coordinates as follows:

(0.41,0.94);(0.37,0.84),(0.54,0.67);(0.25,0.62);(0.07,0.64);(0.02,0.99);(0.68,0.58);0.71,0.44);(0.54,0.62);(0.83,0.69

);(0.64,0.60);(0.18,0.54);(0.22,0.60);(0.83,0.46);(0.91,0.38);(0.25,0.38);(0.24,0.42);(0.58,0.69);(0.71 ,0.71);(0.74, 0.78).

We initialize the parameters as follows:

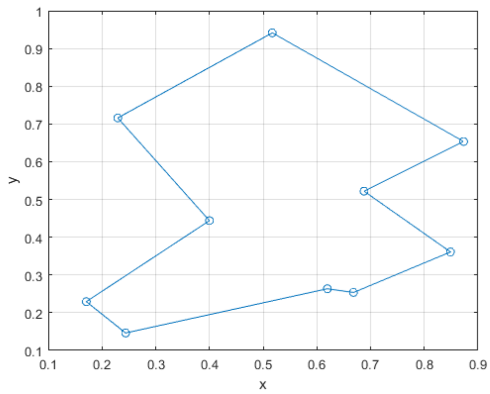
$A = B = 1$, $D = 2$, $\gamma = 1$, $k = 1$, $I_0 = 0.5$, $\varepsilon = 0.02$, $z(1) = 0.98$, $\beta_1 = 0.001$, $\beta_2 = 0.003$, $\lambda = 6/10$.

The shortest path that satisfies the TSP problem of 20 cities under this coordinate is 3.1934. The program runs a total of 5 times, and 100 different initial conditions of x_{ij} are generated randomly in the interval [- 0.1, 0.1]. The results are summarized in Table 2.

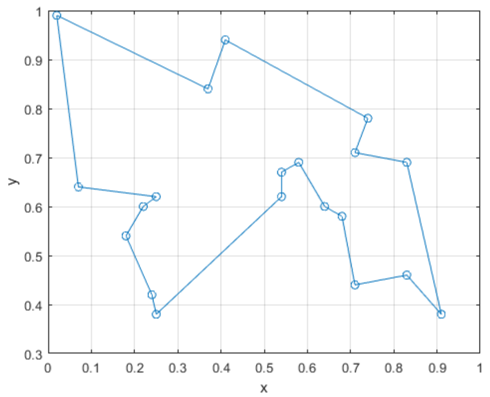
Table 2. Results of SCF Model Solving TSP Problem in 20 Cities

Legal path	Optimal path	Legal ratio(%)	Optimal ratio(%)
94	91	94	91
92	90	92	90
90	89	90	89
93	92	93	92
91	90	91	90

As see from Table 2, this model has an efficiently performance in solving the 20-city TSP problem under the above-mentioned parameter settings.



10-city



20-city

Fig. 9: The shortest path of 10 cities and 20 cities TSP normalized coordinates

4 Conclusion

In this paper, Chebyshev polynomials are introduced into Chen's transient chaotic neural network, and a new transient chaotic neural network is constructed. The unified framework theory is used to verify the rationality of the additional energy term of the SCF chaotic neural network, and this model is applied to solve the optimization function, 10 cities and 20 traveling salesman problems, the influence of parameters on the network is also discussed. The simulation results show that this model has better optimization capabilities. In the future, we hope to make better use of chaotic features to further improve the optimization of network models.

References

[1] L.N. Chen, K. Aihara, Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos, Neural Networks, 1995, 8(6): 915-930.

[2] Potapov A, Ali M K. Robust chaos in neural networks[J]. Physics Letters A, 2000, 277(6): 310-322.

[3] Shuai J, Chen Z, Liu R, et al. Self-evolution neural model[J]. Physics Letters A, 1996, 221(5): 311-316.

[4] Hu Z, Li W, Qiao J. Frequency conversion sinusoidal chaotic neural network based on adaptive simulated annealing. Journal of Electronic, 2019,47(3):613-622.

[5] Xu Y, Sun M, Duan G. Wavelet chaotic neural networks and their application to optimization problems. Lecture Notes in Computer Science. 2006, 3791: 379-384

[6] Ye Y, Xu Y. SLF chaotic neural network model and its application. Computer Engineering and Applications, 2015 (21): 74-79.

[7] Lin S. Computer solutions of the traveling salesman problem[J]. Bell System Technical Journal, 1965, 44(10): 2245-2269.

[8] Imamoglu E, Kaltofen E L. On computing the degree of a Chebyshev polynomial from its value[J]. Journal of Symbolic Computation, 2021, 104: 159-167.

[9] Sukhorukova N, Ugon J, Yost D. Chebyshev multivariate polynomial approximation and point reduction procedure[J]. Constructive Approximation, 2019: 1-16.

[10] Kwok T, Smith K. A unified framework for chaotic neural-network approaches to combinatorial optimization. IEEE Transactions on Neural Networks. 1999, 10(4): 978–981

[11] Kwok T, Smith K. Experimental analysis of chaotic neural network models for combinatorial optimization under a unifying framework. Neural Networks. 2000, 13(7): 731–744

[12] Xu Y, Xu N, Qiu Z. Chaotic Neural Network with Trigonometric Function Self-feedback[C]//2020 39th Chinese Control Conference (CCC). IEEE, 2020: 7396-7401.

[13] Yamada T, Aihara K, Kotani M. Chaotic neural networks and the traveling salesman problem[C]//Proceedings of 1993 International Conference on Neural Networks (IJCNN-93 - Nagoya, Japan). IEEE, 1993, 2: 1549-1552.

[14] Y. Xu, Q. Yu, Optimizing Logistic Distribution Routing Problem Based on Wavelet Chaotic Neural Network, Proceedings of the 2017 Chinese Control and Decision Conference, CCDC 2017, 3460-3464.

[15] Fogel D B. Applying evolutionary programming to selected traveling salesman problems[J]. Cybernetics and systems, 1993, 24(1): 27-36.