Research on Information Security Monitoring and Early Warning Mechanism of Internet Application Network Based on Particle Swarm Optimization

Feng Chen[1(✉)], Hong Zou[1], and Xue-sheng Li[2]

[1] Digital Grid Research Institute, CSG, Guangzhou 510507, China
cf9198@126.com
[2] Beifang Minzu University, Yinchuan 750021, China

Abstract. Due to the frequent occurrence of network security incidents, causing unnecessary losses to people, frequent network security incidents are worrying. For the problems of Internet application network information security, attackers use attacks to continuously threaten them. This paper studies the method of information security monitoring and early warning mechanism for Internet application network based on particle swarm optimization. Based on the support vector regression machine, a network security prediction model with multi-group chaotic particle optimization is established. The prediction results are obtained through the network information security monitoring and early warning mechanism, and the prediction results are analyzed and summarized. The results show that the Internet application network information security prediction model based on particle swarm optimization algorithm can provide guidance for the development of network security solutions and strategies, enhance the initiative of network security defense, reduce the losses caused by network attacks, and have better practicality Sex.

Keywords: Particle swarm optimization · Network information security · Monitoring and early warning

1 Introduction

The particle swarm optimization algorithm, also known as the particle swarm optimization algorithm or the flock foraging algorithm, abbreviated as PSO, is a new evolutionary algorithm developed by J. Kennedy and RC Eberhart in recent years. The PSO algorithm is a kind of evolutionary algorithm. It is similar to the simulated annealing algorithm. Starting from the random solution, it finds the optimal solution by iteration, and evaluates the quality of the solution according to the fitness. However, it is simpler than the genetic algorithm rule. It has no genetic algorithm. The "cross" and "mutation" operations seek global optimality by following the current searched optimal values. This kind of algorithm has attracted the attention of the academic community because of its high precision and fast convergence, and it has shown its superiority in

Y.-D. Zhang et al. (Eds.): ICMTEL 2020, LNICST 327, pp. 241–253, 2020.
https://doi.org/10.1007/978-3-030-51103-6_21

solving practical problems. Particle swarm optimization is a parallel algorithm. The information security situation value of Internet application network based on particle swarm optimization is an important indicator to measure network security. The value of its value directly affects the quality of network security [1]. Establishing an effective network information security monitoring and early warning mechanism model can prevent network security incidents and play an important role in network security protection.

Aiming at the shortcomings of particle swarm optimization algorithm, this paper improves it and proposes a multi-group chaotic particle optimization algorithm. The algorithm uses the randomness of chaotic principle, initializes the population particles in the initial stage, and divides the particles into three populations. Different populations adopt different updating strategies, and the convergence speed of the algorithm is accelerated by the synergy and information sharing among the three populations. The variance of the population fitness is used to judge whether the particles fall into local convergence, and chaotic processing is performed on the particles that are partially converged. It escapes from the local convergence point and avoids the phenomenon of "premature maturity" of the population with a certain probability, thus improving the optimization performance of the algorithm. The improved algorithm is tested by four standard test functions and compared with pso and ldw-pso to verify the algorithm has better performance.

2 Improved Particle Swarm Optimization for Network Information Security Prediction

Through the understanding and analysis of the particle swarm optimization algorithm, it can be seen that the exploration ability and development ability of the particle in the algorithm are a contradiction in the optimization process. If the particle swarm algorithm optimizes the simple problem, the global exploration ability of the particle itself and the local development capability is not high, and generally the optimal solution for the optimization problem can be found. However, for complex optimization problems, if the particle happens to be near the global best of the population, the development ability of the particle is weak due to the strong exploration ability of the particle at this time. So that the particles are farther away from the neighborhood where the optimal solution is located; if the particles are in the neighborhood of the most favorable locality, due to the strong development ability of the particles at this time, the exploration ability is weak, and it is easy for the particles to fall into local convergence [2]. Therefore, when the population is initialized, the randomness of the chaotic principle is used to make the particles evenly distributed in the search range, and then the population is divided into multiple populations, so that different populations can exhibit different exploration and development capabilities. Some particle swarms have the ability to explore and optimize. Global search in the process; part of the particles have the ability to develop, local search in the optimization process. Through the information sharing between the particles of the population, mention Optimization of the performance of the algorithm.

2.1 Building a Support Vector Regression Machine

Support vector regression is based on some complex relationship between sample data, to determine the functional relationship f (x) between the independent and dependent variables in the sample data for regression, the functional relationship f (x) can be divided into linear and non-linear Linear.

(a) Linear regression

Let the linear regression data be:

$$K = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}, x \in Rn, y \in R \quad (1)$$

Where x is the input of the regression function, the dimension is n, and the general case is n > l. y is the output of the regression function, and the dimension is 1.

Let svr's regression function be:

$$f(x) = (w \cdot x) + b, w \in Rn, b \in R \quad (2)$$

Where w and b are regression functions and the unknown parameters of f (x) are real numbers.$(w \cdot x)$ The dot product of the vector, and the input vector x.

The goal of the support vector regression algorithm is to determine the regression function f (x) so that all sample data are near the regression function f (x), allowing some sample data to deviate from f (x) by a certain distance e, which has a good fit [3]. In the two-dimensional space, the regression function of the sample data is shown in Fig. 1.

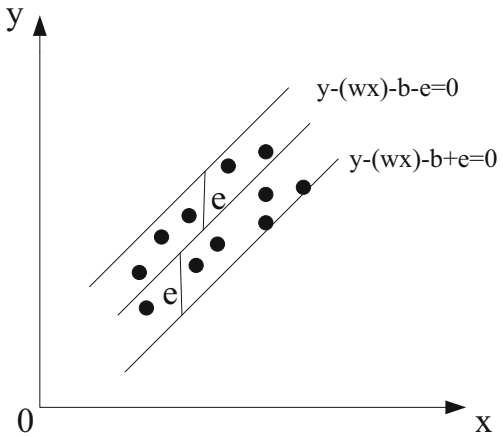

Fig. 1. Schematic diagram of the regression function of the two-dimensional space sample data

As you can see from Fig. 1, all the sample data is in

Y − (w · x) – be = 0 and y − (w · x) – b + e = 0 in the area of two parallel lines, that is, all sample data satisfy the formula (3).

$$-e < f(x_i) - (w \cdot x_i) - b < e,\ i = 1, 2, \ldots, n \tag{3}$$

2.2 Network Information Security Monitoring and Early Warning Mechanism

The three parameters of the deviation parameter e, the penalty parameter C and the kernel function parameter g are determined to determine the network situation prediction model of the SVR. The deviation parameter e is the distance from which the sample data deviates from the regression line, and the number of sample data affecting the support vector. If the value of e is relatively large, the number of sample data of the support vector is relatively small, and the error of the regression estimation is increased, and the precision is lower. If the value of e is relatively small, the number of sample data of the support vector is relatively large, and the number of samples is reduced. The accuracy of the regression estimate is higher. Penalizing the size of the parameter C not only affects the training error of the sample, but also affects the complexity of the regression model [4]. If the value of C is relatively large, the phenomenon of "over-learning" will occur. The larger C value imposes a heavier penalty on the sample data outside the e-band, increasing the number of support vectors, thereby reducing the training error of the sample and increasing The generalization ability of the model; if the value of C is relatively small, it will produce the phenomenon of "under-learning". The smaller C value will lighten the penalty of the sample data outside the e-band, reducing the number of support vectors and increasing the training. Error, reducing the generalization ability of the model. The kernel function parameter g is the width of the RBF, and its magnitude affects the corresponding width of the inner product kernel function on the input variable. If the value of g is too large, the regression function or the discriminant function is too gentle; if the value of g is too small, the training data may have a certain memory ability or an "over-fitting" phenomenon.

In summary, setting reasonable parameters plays an important role in the generalization ability of the support vector regression model. By judging the error rate of the model algorithm, it can be judged whether the value of the model parameters is optimal [5].

2.3 Network Security Prediction Model for Multi-group Chaotic Particle Optimization

Set the number of hosts infected with network viruses to x_i. Number of websites that have been tampered with $x_{\min}$. Total number of sites that were implanted in the back door $x_{\max}$, the number of new information security vulnerabilities y_i As the independent variable of the network situation value, the network situation value is taken as the dependent variable. In order to reduce the different dimensions of the sample data and

cause unnecessary errors in the results, the sample data needs to be normalized. Its formula is as shown in (4):

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

The network information security monitoring and early warning mechanism improves the shortcomings of particle swarm optimization algorithm. An MSCPO algorithm is proposed. The performance of the algorithm is verified by the test function, and the algorithm has better performance. Therefore, the parameters of the support vector regression machine model are optimized by MSCPO algorithm, and the model is trained for the sample data set used in this paper. Three parameters of deviating parameter e, penalty parameter C and kernel function parameter g are selected to establish MSCPO-SVR. The prediction model is expected to produce better prediction results [6].

The fitness function of the mscpo algorithm is the mean square error of the support vector regression model:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - x_i^{'})^2 \tag{5}$$

among them, x_i For the true value, $x_i^{'}$ For the predicted value, n is the number of samples.

The establishment process of the network information security monitoring and early warning mechanism:

(1) Observe the original value of the input data and analyze it.
(2) Construct effective predictors, select independent and dependent variables.
(3) Preprocessing of the sample data, that is, normalization.
(4) The first 180 data of the sample data are used as the training set, and the last 10 data are used as the test set.
(5) Initialize the mscpo algorithm.
(6) Optimize the parameters of the support vector regression machine model according to the mscpo algorithm, and train the training set.

Practice, determine the MSCPO-SVR prediction model.

(7) Test the test sample data and output the test results of the model.

The flow chart for establishing the MSCPO-SVR prediction model is shown in Fig. 2.

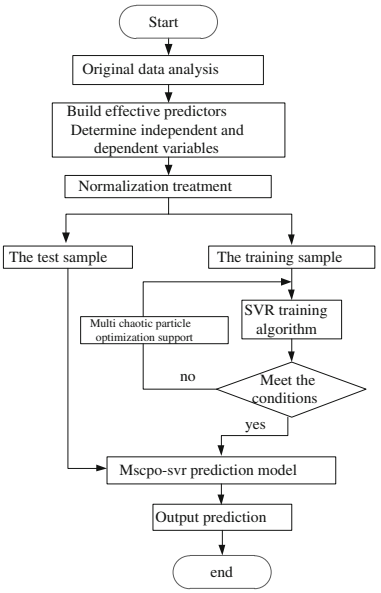

Fig. 2. MSCPO-SVR network situation prediction model flow

2.4 Prediction Results

Since the MSCPO-SVR prediction model has a small complexity and does not require prior knowledge, it can perform better generalization performance on randomly distributed sample data, and thus is applicable to any distributed sample data, so the MSCPO-SVR network situation prediction. The prediction results of the model [7]. MSCPO-SVR prediction model are shown in Table 1.

Table 1. Prediction results of the MSCPO-SVR prediction model

Serial number	Actual value	Predictive value	Absolute error
1	4	3.9150	0.0850
2	4	3.9442	0.1858
3	3	3.1303	0.2303
4	3	4.1372	0.1037
5	3	3.8613	0.2597
6	4	4.2018	0.2307
7	4	2.9263	0.1387
8	4	3.1597	0.3081
9	4	3.8867	0.2133
10	4	4.0932	0.1932

The absolute error of the prediction result of the MSCPO-SVR prediction model is shown in Fig. 3.

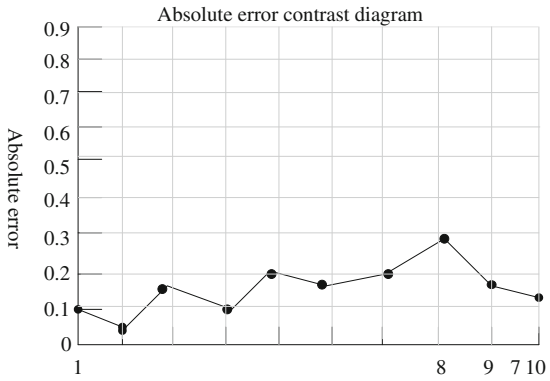

Fig. 3. Absolute error of the prediction result of the MSCPO-SVR prediction model

In order to visually represent the predictive performance of the model, compare its predicted value with the true value, such as

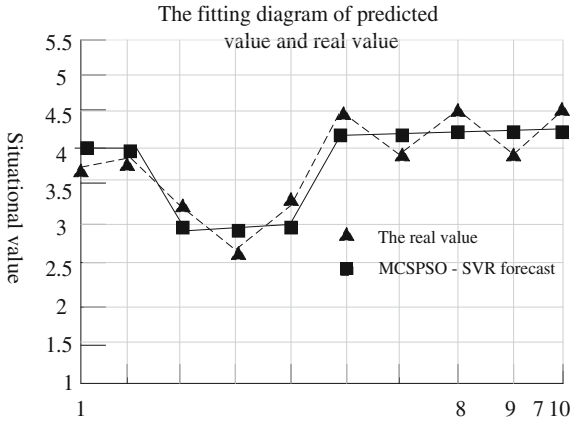

Fig. 4. Fitting the prediction results of the MSCPO-SVR prediction model with the real values

It can be seen from Fig. 3 and Fig. 4 that the absolute error of the prediction result obtained by the MSCPO-SVR prediction model proposed in this paper is relatively small, and the predicted value is close to the real value, and the fitting is better.

2.5 Comparative Analysis of Prediction Results

In order to verify the prediction performance of the MSCPO-SVR prediction model, it is compared with the prediction results of the bp neural network prediction model and the svr prediction model. This experiment uses the following commonly used evaluation indicators to evaluate the experimental prediction results, if the evaluation index is smaller, It indicates that the prediction result is more fitting to the true value.

(1) Absolute error e:

$$e = \left|x_i - x_i^{'}\right| \tag{6}$$

(2) Average absolute error mae:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - x_i^{'}\right| \tag{7}$$

(3) Average relative error mape:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - x_i^{'}\right|\frac{\left|x_i - x_i^{'}\right|}{x_i} \tag{8}$$

(4) Average square root error rmse:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - x_i^{'}\right)^2} \tag{9}$$

among them, x_i For the true value, $x_i^{'}$ For the predicted value, n is the total number of samples.

The bp neural network prediction model and the svr prediction model sample data are used for training, and the test results are obtained. The prediction results of the prediction model are shown in Table 2.

According to formulas (7), (8), (9), the evaluation index values of the prediction results of the bp neural network prediction model and the svr prediction model are shown in Table 3.

The comparison index between the bp neural network prediction model and the svr prediction model is shown in Fig. 5.

Table 2. Comparison prediction results of prediction models

Serial number	Actual value	MSCPO-SVR		BP	
		Predictive value	Absolute error	Predictive value	Absolute error
1	4	3.9150	0.0850	3.2150	0.7850
2	4	3.9442	0.1858	4.7442	0.7442
3	3	3.1303	0.2303	2.2303	0.7697
4	3	4.1372	0.1037	3.7459	0.7459
5	3	3.8613	0.2597	2.3524	0.6476
6	4	4.2018	0.2307	4.8372	0.8372
7	4	2.9263	0.1387	3.3613	0.6387
8	4	3.1597	0.3081	4.2965	0.2965
9	4	3.8867	0.2133	3.4867	0.5133
10	4	4.0932	0.1932	3.6932	0.3068

Table 3. Evaluation index value of the prediction model

Prediction model	MSCPO-SVR	BP
MAE	0.1189	0.6285
MAPE	0.0327	0.1751
RMSE	0.1167	0.6552

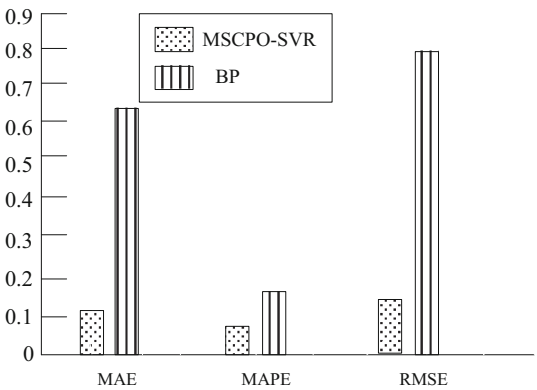

Fig. 5. Comparison chart of predictive model evaluation indicators

In summary, the predicted value of the MSCPO-SVR prediction model proposed in this paper is more fitting to the true value. Determine the three parameters of deviation parameter e, penalty parameter c and kernel function parameter g, establish the MSCPO-SVR prediction model, and test the 10 data after the sample data, and obtain the fitting comparison between the prediction result and the real value Strong.

Compared with BP neural network and MSCPO-SVR prediction model, the evaluation indexes of MSCPO-SVR prediction model are relatively small, which is better than the other two prediction models [8]. Through the BP neural network and MSCPO-SVR prediction model, the predicted value and the true value fit map show that the MSCPO-SVR prediction model has strong fitting and can maximize the optimization of network information security prediction [9, 10].

3 Simulation Test Experiment

3.1 Experimental Software and Hardware Environment

The experimental hardware configuration is computer configuration processor: Intel i5-3210 M CPU@2.SOGH; memory 10.0 GB; hard disk 1 TB; operating system 64-bit Windows 10 Professional.

The experimental software configuration is MATLAB environment: MATLABR 2016aVersion9.0.0.34136064-bit.

3.2 Comparison of Experimental Results

In order to verify the optimization performance of the MSCPO-SVR network information security monitoring and early warning algorithm proposed in this section, the algorithm of this section is used to test its algorithm and compare it with the pso and ldw-pso algorithms. The maximum number of iterations of the three optimization algorithms. Both are 1000, and the initialization parameters are set as follows.

PSO: w = 0.75, $c_1 = c_2 = 2$;
LDW-PSO: $w_{\max} = 0.9$, $w_{\min} = 0.4$, $c_1 = c_2 = 2$;

The first population in the MSCPO algorithm is w = 0.9, and the second population is w = 0.4, in the third population.$w_{\max} = 0.9$, $w_{\min} = 0.4$.

The three algorithms are independently operated 100 times in the 3 and 10 dimensions of the four standard test functions, and the average optimal solution, the average number of iterations, and the optimization success rate of the three algorithms are recorded. By finding the average number of iterations of the optimal solution, it is used to measure the convergence speed of the algorithm; by finding the average optimal solution size, it is used to measure the convergence accuracy of the algorithm; by the optimization success rate, it is used to measure the algorithm. Overall performance. The specific parameters of the standard test function are shown in Table 4.

Table 4. Standard test function parameter setting table

Test function name	Population size	Ranges	Convergence accuracy
Sphere	100	[−100, 100]	1E–10
Ackley	100	[−10, 10]	1E–6
Rastrigin	100	[−10, 10]	1E–6
Griewank	100	[−10, 10]	1E–6

The experimental results of the standard test functions of the three algorithms pso, ldw-pso and mscpo show that the mscpo algorithm finds the optimal solution with fewer iterations, and the optimization performance is better than the other two algorithms. For the peak function, the pso and ldw-pso algorithms also have better performance, and the average number of iterations is not much different. Because the unimodal function itself does not have a local optimal solution, the algorithm does not fall into local convergence. However, for the 3-D multi-peak test function, when the pso and ldw-pso algorithms are initialized, some particles may be in the vicinity of the optimal solution, so the optimal solution can be found with a certain probability, and the number of iterations is relatively small, but the population is found. The success rate of the optimal solution is relatively low. The average iteration number of the mscpo algorithm is lower than the other two algorithms, and the optimization success rate is significantly higher than the other two algorithms. Because the mscpo algorithm divides the particles into three populations during initialization. In the iterative process, the three populations cooperate and share the optimal value of each population. The chaotic processing of the particles that are trapped in the local optimum helps them escape the local best advantage, so that the algorithm finds the optimal solution with fewer iterations.

This paper is devoted to optimizing the shortcomings of particle swarm optimization algorithm. In order to better study the information security monitoring and early warning mechanism of Internet application network and improve its algorithm, a multi-group chaotic particle optimization algorithm is proposed. The 10D dimension of the MQC algorithm is tested and compared with the pso and ldw-pso algorithms. The experimental results show that:

(1) For the 3-dimensional test function, the pso and ldw-pso algorithms perform better for the single-peak test function because there is no local best advantage in the 3-dimensional single-peak test function. But for the 3-dimensional multi-peak test function Some particles may be in the vicinity of the optimal solution when they are initialized, so the optimal solution can be found, and the success rate of finding the optimal solution is relatively low. However, the optimization performance of the mscpo algorithm proposed in this paper is better than other in the single-peak test function. The two algorithms are not very obvious, but the optimization performance of the multi-peak test function is better than the pso and ldw-pso algorithms, showing good optimization performance.
(2) For the 10-dimensional test function, the pso and ldw-pso algorithms are only valid for the single-peak test function. For the multi-peak test function, the search success rate of the algorithm is very low, almost zero, and it is impossible to avoid the “precocity” of the population. Phenomenon. However, the mscpo algorithm proposed in this paper is better than the pso and ldw-pso algorithms for both single-peak and multi-peak functions. The population is initialized by chaotic mutation, and the population particles are divided into three different populations. Excellent particles adopt chaotic processing, which has certain efficiency to avoid the phenomenon of “early maturity” in the population. However, compared with the 3-dimensional multi-peak test function, although the average number of iterations has increased, the success rate of optimization has decreased, but still the optimal solution can be found with a certain probability.

The improved algorithm is used to optimize the parameters of the support vector regression machine model, the sample data set collected in this paper is trained, the MSCPO-SVR prediction model is established, and the sample data is predicted. The error between the predicted value and the real value is smaller. Compared with the bp neural network and the svr prediction model, the evaluation index of the model is lower than the other two prediction models. It can be seen that the prediction effect of the model is better, and the predicted value is the true value is more fitting.

In summary, the MSCPO-SVR prediction model can provide guidance for the development of network security solutions and strategies, enhance the initiative of network security defense, reduce the losses caused by network attacks, and have better practicability and enhance network security. The initiative of defense.

4 Conclusion

Through the analysis and summary of the research status of network information security monitoring and early warning model technology and particle swarm optimization algorithm, the regression principle of support vector machine is briefly explained. Analyze the situation report published on the website of the National Internet Emergency Center and establish a sample data set. The parameters of the MSCPO optimized support vector regression model are proposed. The first 100 data of the sample data are used as the training set to determine the three parameters of the deviation parameter e, the penalty parameter C and the kernel function parameter g. The MSCPO-SVR prediction model was established, and the 10 data after the sample data were tested. The prediction results obtained were more fitting with the real values. By comparing with BP neural network and MSCPO-SVR prediction model, it is verified that MSCPO-SVR prediction model has strong fitting and practicability for network information security monitoring.

Acknowledgment. China Southern Power Grid technology project "Information Operation Security System V1.0 product development" (2018030102dx00697).

References

1. Qi, Y., Tan, R.: Application of improved particle swarm optimization algorithm-based BP neural network to dam deformation analysis. Water Resour. Hydropower Eng. **48**(2), 118–124 (2017)
2. He, R., Luo, D.: Application of improved particle-swarm-optimization neural network in coalmine safety evaluation. Ind. Saf. Environ. Prot. **44**(11), 33–35 (2018)
3. Luo, S., Liu, C.: A detection method based on particle swarm optimization algorithm and SVM dealing with network intrusion. Mod. Electron. Tech. **40**(10), 31–34 (2017)
4. Xiao, Z.: Mobile Internet application platform of information security research of situation assessment. Comput. Simul. **34**(3), 423–426 (2017)
5. Baoren, Ch., Gu, W., Han Kuai, K., et al.: Research on ad hoc network optimization based on chaotic particle swarm optimization. Trans. Beijing Inst. Technol. **37**(4), 381–385 (2017)

6. Ren, P.-F., Gu, L.-K.: WSN node localization algorithm for power transmission networks based on particle swarm optimization. J. Shenyang Univ. Technol. **40**(5), 63–68 (2018)
7. Tao, C.P.Y.J.Y.J.Y.: Respiratory signals prediction based on particle swarm optimization and back propagation neural networks. Chin. J. Biomed. Eng. **37**(6), 714–719 (2018)
8. Wu, C., Liu, J.-M., Guo, Z.-D.: Use of hybrid fuzzy c-means and probabilistic neural network based on improved particle swarm optimization in the prediction of financial distress. Oper. Res. Manag. Sci. **27**(2), 106–114 (2018)
9. Yan, B., Wang, C.: Application of RBF neural network based on particle swarm optimization algorithm in crack width prediction of sluice pier. Water Power **44**(3), 33–36 (2018)
10. Liu, S., Cheng, X., Fu, W., et al.: Numeric characteristics of generalized M-set with its asymptote. Appl. Math. Comput. **243**, 767–774 (2014)