International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 09, 2020
ISSN: 1475-7192

Equipment life prediction based on genetic algorithm under Weibull distribution

Khalid Talib Othman[1], Qusay Essam Hamid[2]

Abstract---*This paper studies the problem of equipment reliability life prediction under the Weibull distribution. The improved genetic algorithm is mainly used to improve the genetic algorithm coding, objective function and genetic operation to realize the estimation of Weibull parameters. The Weibull distribution model is obtained and the reliability life model of the equipment.*

Type of Paper--- *Review*

Keywords: *Weibull distribution; genetic algorithm; life prediction*

Introduction:

Correctly predicting the remaining life of the equipment [1] is of great significance for ensuring the safe operation of the equipment and improving economic benefits. Pan Dong et al. [2] applied the Achard wear model to fully consider the gear load and speed on the gear tooth surface wear. A gear wear life prediction model was established, but the actual application value of the model was not reflected. Xia et al. [3] established a four-parameter Weibull distribution model for failure research. Dragan Juki et al. [4] implemented a weighted error fitting method. Estimates of Weibull parameters, but lack of comparative explanation of estimation methods. Tan et al. [5] used maximum likelihood estimation to estimate the parameters of the two-parameter Weibull distribution, and used examples to verify the feasibility of the method. But for the commonly used three-parameter The Weibull distribution has not been studied accordingly. Zhou et al. [6] explained and compared several commonly used Weibull distribution parameter estimation methods. In contrast, the maximum likelihood estimation method has superior estimation accuracy. However, the intelligent algorithm is not combined with the estimation method, which makes the calculation difficult.

To sum up, the Weibull distribution is one of the most commonly used probability distributions in the reliability analysis of equipment life. Estimate the parameters of the Weibull distribution to obtain the reliability life of the equipment, so as to predict the life [2]. The parameter estimation of Weibull distribution mostly uses maximum likelihood estimation or least square estimation, etc. The use of these two methods not only has the difficulty of calculation, but also it is difficult to ensure the convergence of the calculation process. Therefore, look for new intelligent methods for Weibull parameter estimation is imperative.

Genetic algorithm is an optimization method that simulates the selection and genetic mechanism of nature to find the optimal solution. Wu Junjie et al. [8] use genetic algorithm to find the optimal route. Li Wei et al. [9] improve the selection operator of genetic algorithm , Realized the optimization of BP neural network. Ren Ziwu et al. [11] used particle swarm optimization algorithm to optimize genetic algorithm, and achieved good results. Therefore, because genetic algorithm does not require continuous function, scalability, and potential parallelism It has been widely used. And, according to the actual situation, the genetic algorithm can be improved accordingly and achieve very effective results. This article is based on the Weibull distribution, using the improved genetic algorithm and the maximum similarity However, the estimation is combined to estimate the parameters of the Weibull distribution, and then predict the life of the equipment. Due to the characteristics of the genetic algorithm itself, the parameters of the Weibull distribution can be easily and simply estimated.

1Iraqi University
2Iraqi University/ College of Education for Women

Selection of equipment life distribution model:

The advantages and characteristics of Weibull distribution:

In the dependableness analysis of apparatus life, the ordinarily used distribution models embrace Weibull distribution, statistical distribution, lognormal distribution, exponential distribution, Lord Rayleigh distribution, etc. Among them, Weibull distribution is that the most typically used distribution in dependableness analysis. it absolutely was introduced by the Swedish scientist Weibull in 1939. it's the theoretical basis for dependableness analysis and life testing. Weibull distribution [3] has several forms, together with one-parameter, two-parameter, three-parameter or mixed Weibull. Three-parameter The Weibull distribution may be a comparatively complete distribution, that is extremely versatile in fitting random knowledge and has sturdy ability. Therefore, the three-parameter Weibull distribution will additional accurately categorical the chance distribution of the instrumentation life, and, within the field of dependableness analysis many ordinarily used distributions, like exponential distribution and Lord Rayleigh distribution, are often thought to be special cases of the three-parameter Weibull distribution. The 3-parameter Weibull distribution is decided by the three parameters of form, scale (range) and position. parameter is that the most a vital parameter determines the essential shape of the distribution density curve; the size parameter functions to enlarge or cut back the curve, however doesn't have an effect on the form of the distribution. The distribution operate of the three-parameter Weibull distribution is (assuming that the instrumentation has no failure before time γ):

$$F(X) = 1 - \exp\left[-\left(\frac{x-\gamma}{a}\right)^{\beta}\right], x \geq \gamma \quad (1)$$

The probability density function and the failure rate function are

$$f(x) = \frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{a}\right)^{\beta}\right], x \geq \gamma \quad (2)$$

$$\lambda(x) = \left(\frac{\beta}{\alpha}\right)\left[\frac{x-\gamma}{\alpha}\right]^{\beta-1}, x \geq \gamma \quad (3)$$

The reliability function is

$$R(x) = \exp\left\{-\left[\frac{x-\gamma}{\alpha}\right]^{\beta}\right\}, x \geq \gamma \quad (4)$$

Among them, α, β> 0, γ ≥ 0, α is the scale parameter, β is the shape parameter, and γ is the position parameter [4].

The Weibull distribution must meet the following two conditions: First, it satisfies the randomness of objective things; second, the variable must be greater than zero. The life of the equipment is a continuous random variable with a value greater than zero. Therefore, the equipment life distribution conforms to the Weibull distribution Under the conditions of, the Weibull distribution model can be established to predict the life of the equipment. That is, the Weibull distribution model of the equipment life can be obtained, and the equipment life loss distribution function, probability density function and failure function are respectively (1)-(3) shown [5].

The relationship between equipment failure rate and Weibull parameters:

In the failure rate function (as shown in formula (3)), since α is a scale parameter, its size is reflected by the enlargement or reduction of the horizontal and vertical coordinates of the graph. Its change does not affect the shape of the graph, and γ is the initial loss of the equipment Therefore, in the failure rate formula, the value of β plays a very important role. Let α = 1, γ = 0 can simulate the bathtub curve of the equipment life, as shown in Figure 2.1. It can be seen from the figure that when β <1, λ (x) decreases with the increase of time, the equipment runs in the early failure period I, and the life loss of the equipment is low. When β = 1, λ (x) basically does not increase with the increase of time When β> 1, the λ (x) curve shows an increasing distribution. When β = 3∼4, the λ (x) curve is similar to the normal distribution shape., The life loss of the equipment increases, and the equipment runs in the loss and failure period III [7].

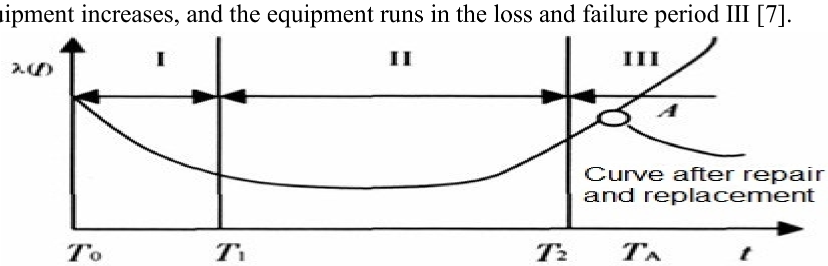

Figure 2.1: Equipment life curve

Using genetic algorithm to estimate Weibull parameters

The parameter estimation methods of Weibull distribution include: maximum likelihood estimation, moment estimation, least square estimation, linear regression estimation, etc. For the two-parameter Weibulll distribution model, the above methods can achieve relatively satisfactory results, and the calculation difficulty is not Very large. But for the three-parameter Weibull distribution model, due to the complexity of the model itself, when using the above method, there are problems such as difficult calculations and insufficient estimation accuracy [6]. Therefore, it is hoped to improve it to reduce calculations The dual effect of improving accuracy and accuracy. Genetic algorithm has the advantages of fast solving speed and high accuracy. The use of genetic algorithm combined with maximum likelihood estimation can achieve satisfactory results.

According to the historical failure data of the same equipment and the parameter characteristics of Weibull distribution, the use of genetic algorithm combined with maximum likelihood estimation can determine the specific values of α, β, and γ, thereby obtaining the Weibull distribution model of the equipment, and The reliability life of the equipment and various related indicators can be calculated. According to the specific Weibull distribution model, the operating conditions of the equipment can be evaluated and the reliability parameters that are practical for design, work and maintenance can be obtained.

Basic genetic algorithm:

Genetic algorithm is a method that uses biological evolution theory to select the approximate optimal solution of the problem[8]. It has the advantages of not requiring function continuity, scalability, potential parallelism, etc., and has been widely used[9] Its implementation process includes operations such as encoding, generating initial population, calculating fitness, selection, duplication, exchange, mutation, repeated iteration, termination, etc. [10].

The working steps of the algorithm are [11]:

- Determine the composition and length of the individual string.
- Randomly establish an initial population.
- Calculate the fitness of each individual and select excellent individuals.
- According to the genetic probability, use the following operations to generate new individuals:

a) Copy. Copy existing good individuals and add them to the new group, delete low-quality individuals.
b) Crossover. The two selected individuals will be exchanged, and the new individuals will enter the new group.
c) Mutation. After randomly changing a character of a body, add a new individual to a new group.

(5) Repeat (3) and (4) until the termination condition is reached, and select the best individual as the result of the genetic algorithm.

The coding of the population is the primary problem to be solved by the genetic algorithm, which directly determines the number of iterations of the algorithm and whether it converges. The traditional binary code string is too long to reduce the efficiency of the algorithm, and the binary code itself is not intuitive and accurate. Not high. The determination of the fitness function of the population is a key issue. Generally, the objective function to be solved is used as the fitness function of the genetic algorithm. Genetic operations mainly include selection, crossover and mutation. Traditional selection is based on probability values To be sure, this method may destroy the generation of new individuals in the group. Traditional crossover and mutation are set in advance, and they do not have the ability to automatically adjust the probability of individual crossover and mutation as the value of the fitness function changes.

Improved genetic algorithm

Taking into account the shortcomings of traditional genetic algorithms, the following improvements are made on its basis:

Improvement of coding scheme

The chromosome is designed into a two-layer hierarchical structure. In the field of biomedicine, the chromosome is composed of control genes and sequence genes. The control genes indicate the role and function of this chromosome, and the sequence genes are used to realize its role and function. Control genes Using binary coding, 1 means that the lower-level genes are activated, and the sequence genes are involved in genetic operations. 0 means the lower-level genes are not activated, and the sequence genes are not involved in any genetic operations. The sequence genes are encoded in real numbers. The structure of the chromosome is shown in Figure 3.1 The chromosome shown in Figure 3.1 is composed of 5 control genes and 5 sequence genes represented by integers, which is expressed as X = (4, 5, 7, 8). Such a chromosome structure can shorten the length of the chromosome structure, It is also possible to determine the probability of mutation by controlling genes and protect excellent genes.

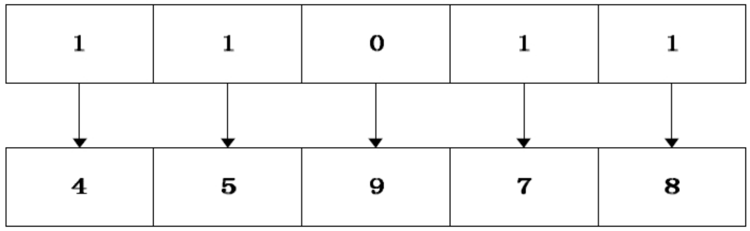

Figure 3.1: Double-layer chromosome structure diagram

The establishment of genetic algorithm objective function

In order to realize parameter estimation with genetic algorithm, the combination of maximum likelihood estimation and genetic algorithm is used to construct the objective function of genetic algorithm as follows:

$$maxF(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i;\theta) \qquad (3.1)$$

$$Equivalent\ to$$

$$minF(\theta) =- \ln L(\theta) =- \sum_{i=1}^{n} \ln f(x_i;\theta) \qquad (3.2)$$

Where $\theta = (\theta 1, \theta 2, ..., \theta m)$ is the parameter to be estimated.

Improvement of genetic manipulation

- Adopt a selection method that retains the best individual [12]. Because the genetic algorithm itself has random selection, even the best individual may be eliminated in the selection algorithm, and the worst individual may also be selected. Therefore, in The genetic algorithm adopts the selection method of retaining the best individuals [13], that is, each generation retains a few of the best individuals not to participate in selection and mutation operations, but to participate in crossover operations, which can ensure that excellent genes are passed to the next generation, so that the population The overall performance has been improved.
- The crossover probability PC and mutation probability PM automatically change with the fitness function value. When the individual fitness value is lower than the overall average fitness value, it indicates that the individual's performance is relatively poor, and a larger PC and PM should be used. Adaptation The greater the degree value, the smaller the PC and PM. The specific values are as follows:

$$\begin{cases} P_C = P_{C1} - (P_{C1} - P_{C2})(f - \bar{f})/(f_{max} - \bar{f}) \cdots f \geq \bar{f} \\ P_C = P_{C1} \cdots f \prec \bar{f} \\ P_M = P_{M1} - (P_{M1} - P_{M2})(f_{max} - \bar{f})/(f_{max} - \bar{f}) \cdots f \geq \bar{f} \\ P_M = P_{M1} \cdots f \prec \bar{f} \end{cases} \qquad (3.3)$$

Where f is the average fitness value of the group, fmax is the maximum fitness value of the group, f is the fitness value of the individual, PC1 = 0.9, PC2 = 0.6, PM1 = 0.1, PM2 = 0.001. In this way, the values of PC and PM vary with The fitness function changes and changes, coupled with the optimal individual preservation strategy, can help the genetic algorithm to converge quickly, and can avoid falling into the local optimal solution.

Improvement of termination conditions. Traditional genetic algorithms generally take the number of genetic iterations to terminate the algorithm. In this way, it is possible to fall into a local optimum and the size of the error cannot be ensured, so the genetic termination conditions are improved-that is, when two rounds The relative error of the optimal fitness of the genetic operation is less than 5 ‰, and the genetic operation ends. In order to avoid falling into the local optimum as much as possible, the restriction of the number of genetics is also added, that is, the genetic algorithm is terminated when both the number of genetics and the error conditions are met [14,15].

The specific process of the genetic algorithm to estimate Weibull parameters

The specific process of Weibull parameter estimation using genetic algorithm is shown in Figure 3.2.

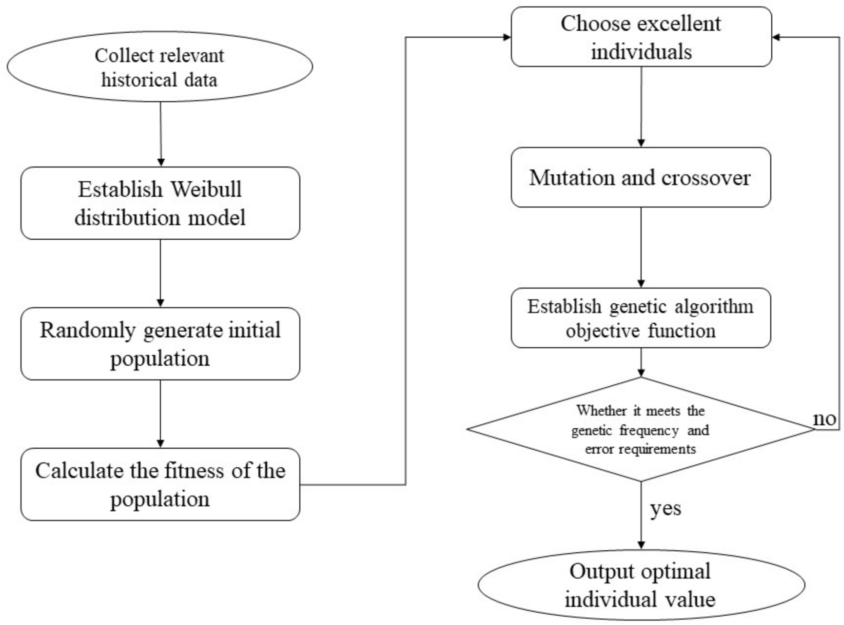

Figure 3.2: The specific flow chart of the genetic algorithm to estimate the weibull parameter

Examples of equipment life prediction

The specific steps of the equipment life prediction based on genetic algorithm under Weibull distribution are as follows:

- Establish a Weibull distribution life model for equipment conforming to three parameters, and determine its probability density number.
- Collect historical failure data of the same equipment.
- Estimate the Weibull parameters using the improved genetic algorithm and historical failure data.
- According to the estimated values of the parameters, determine the life loss degree and reliability of such equipment.

Table 4.1 is a statistical table of pipeline fatigue life failure data (data from literature [16]), pipeline failure data (from the appearance of pipeline fatigue cracks to the time when the requirements for use are not met) are tested by a buried oil and gas magnetic flux leakage detection device developed by a certain institute The data shall prevail. Literature [16] adopts the method of maximum likelihood estimation, through certain calculations, established a relatively satisfactory model, and obtained relatively accurate results. In order to further reduce the amount of calculation and improve the accuracy, this article uses the improved The genetic algorithm establishes the Weibull distribution model. The specific process is as follows:

Table 4.1: Statistical Table of Failure Data of Pipeline Fatigue Life (Unit : Day)

No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Time	160	178	196	218	231	246	255	286	326
No.	10	11	12	13	14	15	16	17	18
Time	358	408	446	512	643	690	700	712	740

Using the concept of maximum likelihood estimation, through Matlab language programming, the traditional genetic algorithm and the improved genetic algorithm are used to obtain the parameter values respectively. The optimization iterations are shown in Figure 4.1 and Figure 4.2:

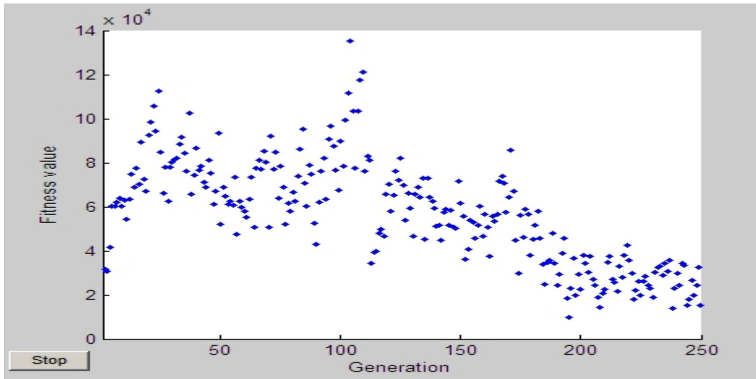

Figure 4.1: Traditional genetic algorithm optimization iteration diagram

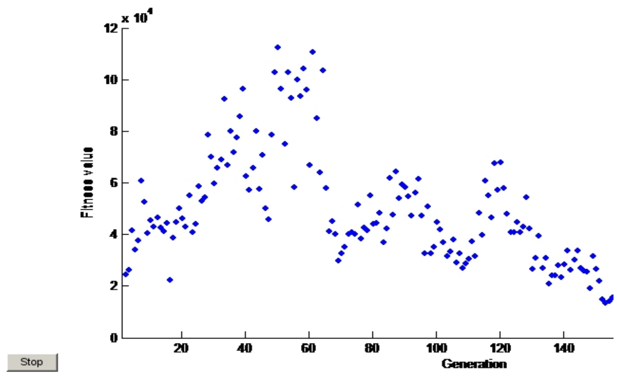

Figure 4.2: Improved genetic algorithm optimization iteration diagram

The results obtained by both and the results of the literature [16] are shown in Table 4.2:

Table 4.2: Results

	Genetic algorithm	Improved Genetic algorithm	MLE
Parameter specific value	$\alpha = 215,\ \beta = 3.12,\ \gamma = 149$	$\alpha = 295,\ \beta = 3.41,\ \gamma = 139$	$\alpha = 274,\ \beta = 3.08,\ \gamma = 159$
number of iterations	500	500	-
Maximum objective function	-192.23	-147.22	-152.83

It can be seen from Table 4.2 that the traditional genetic algorithm requires a large number of iterations, and the accuracy of the results obtained is not high. However, using the improved genetic algorithm not only reduces the fluctuation range of the target value, but also reduces the number of iterations of the algorithm. Moreover, a better optimal individual is obtained, and the obtained accuracy is higher than that in the literature [16], and the amount of calculation is also reduced.

Substituting the parameter values obtained by the improved genetic algorithm into formulas (2.1)-(2.4), the Weibull distribution model can be obtained as:

$$F(t) = 1 - \exp\left[-\left(\frac{t-139}{295}\right)^{3.41}\right], t \geq 139$$

The probability density function and the failure rate function are

$$f(t) = \frac{3.41}{295}\left(\frac{t-139}{295}\right)^{2.41} \exp\left[-\left(\frac{t-139}{295}\right)^{3.41}\right], t \geq 139$$

$$\lambda(t) = \left(\frac{3.41}{295}\right)\left[\frac{t-139}{295}\right]^{2.41}, t \geq 139$$

The reliability function is

$$R(t) = \exp\left\{-\left[\frac{t-139}{295}\right]^{3.41}\right\}, t \geq 139$$

At this point, the Weibull distribution model has been successfully established. According to this model, the failure probability and reliability life of the pipeline operation can be easily calculated. Assuming that the pipeline has been working for 300 days after the occurrence of fatigue cracks, the failure probability is λ(300) = 3 × 10−3, the reliability function R(300) = 85.9%. To ensure that the reliability of fatigue failure caused by fatigue cracks in the pipeline reaches 85.9%, after the occurrence of fatigue cracks, the running time of the defective pipeline is about 300 days That is, after the fatigue cracks appear, after working for about 300 days, the pipeline needs to be excavated, repaired or updated to ensure the safety of oil and gas transportation. With a quantitative analysis, it is greatly convenient for the staff to repair or replace the equipment, saving costs, Improved efficiency and brought considerable profits to the enterprise.

Conclusion:

In this paper, the improved genetic algorithm is used to estimate the relevant parameters of the Weibull distribution, so as to obtain the relevant indicators of the reliability and life of the equipment. According to the concept of the maximum likelihood function, the objective function of the genetic algorithm is constructed, and the characteristics of the genetic algorithm are rationally used to solve The maximum likelihood estimation is difficult to calculate, and it is not easy to converge. On this basis, the improvement of the selection and termination conditions of the genetic algorithm itself can solve the random selection of the genetic algorithm itself and fall into the local optimum to a certain extent. However, this paper only conducts corresponding research on a single Weibull distribution, and does not conduct corresponding parameter estimation and other related studies on multiple Weibull mixed models. This is the deficiency of this paper.

References:

1. Zhang Xiaoli, Chen Xuefeng, etc. Overview of life prediction of major machinery equipment [J]. Journal of Mechanical Engineering, 2011, 47(9): 100–116.
2. Pan Dong, Zhao Yang, etc. Gear wear life prediction method [J]. Journal of Harbin Institute of Technology, 2012, 44(9): 29–33.
3. Xia Zhaopeng, Yu Jianyong. Studying the mechanical properties of jute/cotton blended yarns using the Weibull mode[J]. J. Donghua Univ., 2009, 4: 393–396.
4. Dragan Juki, Darija Markovi. On nonlinear weighted errors-in-variables parameter estimation prob- lem in the three-parameter Weibull model[J]. Appl. Math. Comput., 2010, 215: 3599–3609.
5. Tan Zhibin. A new approach to MLE of Weibull distribution with interval data[J]. Reliab. Engin. Sys. Saf., 2009, 94: 394–403.
6. Zhou Dan. Comparison of parameter estimation methods for transformer Weibull lifetime mod- elling[J]. High Voltage Technology, 2013, 39(5): 1170–1177.
7. Liu Shuxin, Liu Changwu, et al. Study on Weibull parameters of rock strength based on damage multifractal characteristics [J]. Chinese Journal of Geotechnical Engineering, 2011, 33(11): 1786–1791.
8. Wu Junjie. Research on genetic algorithm for optimization of hull assembly line drawing [J]. Journal of Dalian University of Technology, 2012, 52(3): 381–386.
9. Li Wei. Research on optimization of BP neural network based on rough set and improved genetic algorithm [J]. Journal of Northwestern Polytechnical University, 2012, 30(4): 601–606.
10. Liu Dalian, Xu Shangwen. Internal and external crossover genetic algorithm for solving constrained optimization problems [J]. Systems Engineering Theory and Practice, 2012, 32(1): 189–195.
11. Ren Ziwu, Umbrella Zhi. Improvement and performance research of real genetic algorithm [J]. Chinese Journal of Electronics, 2007, 35(2): 269–274.
12. Zhao Liang, Lu Jianhong. Multi-objective reactive power optimization of wind farms based on improved genetic algorithm [J]. Electric Power Automation Equipment, 2010, 30(10): 84–88.

13. Song Yisheng. The gradual evolution strategy of genetic algorithm based on shrinkage accuracy [J]. Journal of PLA University of Science and Technology, 2010, 11(3): 343–347.
14. Fan Qingwu, Wang Pu. The essential analysis of genetic algorithm crossover operator [J]. Journal of Beijing University of Technology, 2010, 36(10): 1328–1336.
15. Cui Jianguo, Zhao Yunlong. Life prediction of aviation generator based on genetic algorithm and ARMA model [J]. Acta Aeronautica Sinica, 2011, 32(8): 1506–1511.
16. Song Zhiqiang, Li Zhuxin. Analysis and assessment of fatigue life of buried oil and gas pipelines based on Weibull distribution [J]. Machine Tool & Hydraulics, 2011, 39(7): 130–133.