2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science

Parameter Estimation of Software Reliability Growth Models by A Modified Whale Optimization Algorithm

Kezhong Lu[1,2]
[1]College of Computer Science and Technology
Nanjing University of Aeronautics and Astronautics
Nanjing 211106, Jiangsu, China;
[2]College of Computer Science, Chizhou University
Chizhou 247100, Anhui, China
e-mail: luke76@163.com

Zongmin Ma[1]
[1]College of Computer Science and Technology
Nanjing University of Aeronautics and Astronautics
Nanjing 211106, Jiangsu, China;
e-mail: zongminma@nuaa.edu.cn

***Abstract*—Software reliability growth models (SRGMs) are non-linear in nature, so they are difficult to estimate the proper parameters. An estimation method based on modified Whale Optimization Algorithm (MWOA) in which parameters are estimated is discussed in this paper. The proposed MWOA shows significant advantages in handling variety of modeling problems such as the exponential model (EXPM), power model (POWM) and delayed S-shaped model (DSSM). However, the fitting error of a single model is relatively large. In view of the different growth characteristics at each stage of the model, a three-stage software reliability growth model is proposed in this paper. MWOA is used to estimate the parameters of three-stage software reliability models. Experimental results show that the fitting accuracy of three-stage model is significantly less than that of a single model. MWOA with the three-stage model can provide a better estimate of the software faults.**

Keywords- Software Reliability Growth Models; Parameter Estimation; Whale Optimization Algorithm;Three-stage Model

I. INTRODUCTION

With the development of computer technology, all kinds of software systems have permeated our modern society and play an increasingly important role [1]. An important issue in developing such software systems is to produce high quality software system that satisfies user requirements. Therefore, the software reliability is becoming more and more important to the researchers. Software reliability modeling is one of the important fields of theoretical research and engineering practice for software reliability. Various software reliability growth models (SRGMs) has been introduced for predicting the reliability of a software system. But every model has some advantages and some disadvantages, so the choice regarding which model to follow for reliability prediction depends upon the requirements of the software [2].

The fault prediction process in SRGMs depends on representing the relationship between time span of software testing and cumulative number of errors detected [3]. Most SRGMs, known in history, have two or three parameters to be estimated. The model parameters are normally in nonlinear relationships. This means that traditional parameter estimation techniques suffer from many problems in finding the best set of parameters to tune the model for better prediction [4].

Whale Optimization Algorithm (WOA) is a relatively new meta-heuristic optimization technique proposed by Mirjalili and Lewis [5], which is inspired by the bubble-net hunting of humpback whales. WOA is easy to implement and has few adjustment parameters, which make it superior than Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO) , and Gravitational Search Algorithm (GSA), etc. [6] As a result, WOA has attracted much attention and has been applied to handle many practical engineering application problems. Fundamentally, WOA is free from restrictive assumptions on the continuity, the existence of derivatives, the unimodality, and etc. It looks very attractive to use WOA for estimating the parameter of software reliability growth models.

In this paper, we analyze the WOA and propose a modified WOA (MWOA). The MWOA is used to predict software reliability by predicting the faults during the software testing process using software faults historical data. Moreover, a three-stage software reliability model is developed for a better estimate of the software faults. The rest of the paper is organized in the following manner. In Section 2, we provide an overview of various SRGMS. In Section 3, the Whale Optimization Algorithm is described briefly. In section 4, a modified Whale Optimization Algorithm that will be applied in this paper is proposed. Detailed experiments results are provided in section 5. Finally, Section 6 concludes the paper.

II. SOFTWARE RELIABILITY GROWTH MODELS

Numerous researchers proposed various SRGMs to predict the quality of software components, three well-known SRGMs and an improved three-stage SRGM used in our study are discussed as follows:

A. Exponential Model (EXPM)

This is a continuous time–independent and identical error behavior model. This model was first introduced by Goel and Okumoto [7] and is also known as the GO model.

$$\mu(t;\beta)=\beta_0\left(1-e^{-\beta_1 t}\right) \quad (1)$$

$$\lambda(t;\beta)=\beta_0\beta_1 e^{-\beta_1 t} \quad (2)$$

Where $\mu(t;\beta)$ and $\lambda(t;\beta)$ represent the mean failure function and the failure intensity function, respectively.

B. Power Model (POWM)

The power model was provided in [8].It is a graphical approach to perform analysis of reliability growth data and is simple and straightforward to understand. The equations which govern the relationship between the time *t* and both $\mu(t;\beta)$ and $\lambda(t;\beta)$ are:

$$\mu(t;\beta)=\beta_0 t^{\beta_1} \quad (3)$$

$$\lambda(t;\beta)=\beta_0\beta_1 t e^{\beta_1-1} \quad (4)$$

978-1-5386-7445-1/18/$31.00 ©2018 IEEE
DOI 10.1109/DCABES.2018.00076

C. Delayed S-Shaped Model (DSSM)

This model describes the software reliability process as a delayed S-shaped model [9]. This model is also a finite failure model. The system equation for $\mu(t;\beta)$ and $\lambda(t;\beta)$ are:

$$\mu(t;\beta) = \beta_0\left(1-(1+\beta_1 t)e^{-\beta_1 t}\right) \tag{5}$$

$$\lambda(t;\beta) = \beta_0\beta_1^2 t^{-\beta_1 t} \tag{6}$$

D. Three-Stage Model (TSM)

There is no universally acceptable model that can be trusted to give accurate results in all circumstances. Every model has some pros and cons so the choice regarding which model to follow for reliability prediction depends upon the requirements of the software [2]. Generally speaking, the software reliability models always go through three stages of generation, development and maturity, and the speed of each stage is different. For S-shaped models, usually at the stage of generation, the speed of change is slow; in the stage of development, the speed of change is quicker; in the mature stage, the speed of change is slow. There is no universally acceptable single model that can be trusted to give accurate results in most or all applications. In view of the different growth characteristics at each stage of the model, a three-stage software reliability growth model is proposed in this paper to better fit the data and improve the accuracy of the model.

For an EXPM example, the corresponding three-stage software reliability growth model is as follows:

$$\mu(t;\beta) = \alpha_1\beta_{01}\left(1-e^{-\beta_{11}t}\right) + \alpha_2\beta_{02}\left(1-e^{-\beta_{12}t}\right) + \alpha_3\beta_{03}\left(1-e^{-\beta_{13}t}\right) \tag{7}$$

$$\lambda(t;\beta) = \alpha_1\beta_{01}\beta_{11}e^{-\beta_{11}t} + \alpha_2\beta_{02}\beta_{12}e^{-\beta_{12}t} + \alpha_3\beta_{03}\beta_{13}e^{-\beta_{13}t} \tag{8}$$

Where, $\begin{cases} \alpha_1 = 1, \alpha_2 = \alpha_3 = 0, when\ t \in [0, t_1] \\ \alpha_2 = 1, \alpha_1 = \alpha_3 = 0, when\ t \in (t_1, t_2] \\ \alpha_3 = 1, \alpha_1 = \alpha_2 = 0, when\ t \in (t_2, T] \end{cases}$,

β_{ij} (i=0, 1;j=1,2,3) are model parameters.

III. Whale Optimization Algorithm

Whale Optimization Algorithm (WOA) is a nature-inspired meta-heuristic optimization algorithm proposed by Mirjalili and Lewis in 2016 [5]. The mathematical model of WOA is described in three sections: encircling prey, bubble net attacking method, and search the prey.

A. Encircling Prey

This behavior is mathematically formulated as:

$$\vec{D} = |\vec{C}\cdot\vec{X^*}(t) - \vec{X}(t)| \tag{9}$$

$$\vec{X}(t+1) = \vec{X^*}(t) - \vec{A}\cdot\vec{D} \tag{10}$$

Where t indicates the current iteration, $\vec{X^*}$ is the position vector of the best solution obtained so far, $\vec{X}$ is the position vector, | | is the absolute value, and · is an element-by-element multiplication. $\vec{A}$ and $\vec{C}$ are coefficient vectors that are calculated as follows:

$$\vec{A} = 2\vec{a}\cdot\vec{r} - \vec{a} \tag{11}$$

$$\vec{C} = 2\vec{r} \tag{12}$$

Where $\vec{a}$ is linearly decreased from 2 to 0 over the course of iterations (in both exploration and exploitation phases, and can be calculated by Eq. (13) and $\vec{r}$ is a random vector in [0,1].

$$\vec{a}=2(1-t/MaxIter) \tag{13}$$

Where t is the iteration number and *MaxIter* is the maximum number of allowed iterations.

B. Bubble-net Attacking Method (Exploitation Phase)

Two strategies are utilized to figure the feeding conduct of humpback whales as follows:

1) Shrinking circling system: This behavior is achieved by decreasing the value of a from 2 to 0 in Eq. (13) over the course of iterations.

*2) Spiral redesigning position:*A spiral equation is then created between the position of whale and prey to mimic the helix-shaped movement of humpback whales as follows:

$$\vec{X}(t+1) = \vec{D'}\cdot e^{bl}\cdot\cos 2\pi l + \vec{X^*}(t) \tag{14}$$

Where $\vec{D'} = |\vec{X^*}(t) - \vec{X}(t)|$ indicates the distance of the i^{th} whale to the prey, b is a constant for defining the shape of the logarithmic spiral, l is a random number in [-1,1], and · is an element- by-element multiplication.

In hunting whales swim around the prey in above two paths simultaneously. To update whales' positions, 50% probability is taken for above two methods.

C. Search for Prey (Exploration Phase)

In order to have a global optimizer, a randomly chosen solution is used to update the position instead of the best search agent found so far.

$$\vec{D} = |\vec{C}\cdot\overrightarrow{X_{rand}}(t) - \vec{X}(t)| \tag{15}$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}}(t) - \vec{A}\cdot\vec{D} \tag{16}$$

Where $\overrightarrow{X_{rand}}(t)$ is a random position vector (a random whale) chosen from the current population. For further details, the reader may refer to [5].

IV. A Modified Whale Optimization Algorithm

Although WOA has several character, however, it exists still its limitation for some complex problems, especially existed the phenomenon of premature and the deficiency of global search ability, which affect the performance of WOA. A modified WOA (MWOA) is proposed in this paper. The details are as follows:

A. Choice on Dimension

The choice probability p between the shrinking encircling mechanism and the spiral model is same for a whale during optimization. For some complex problems, this choice is difficult to well balance between exploitation and exploration. The choice probability p changes in each dimension of a whale in this paper. So there is a probability of 50% to choose shrinking encircling update or spiral update in each dimension of a whale for enhancing the search performance of WOA (see step 10 in Fig. 2).

B. Exploration Control

The exploration is controlled by the random number A in WOA. A is in the interval $[-a,a]$ where a is decreased from 2 to 0 over the course of iterations according to the Eq.(13). The range of A is shown in Fig. 1(a). According to Eq.(11), when t>*MaxIter*/2, a<1, and |A|<1. So, when t>*MaxIter*/2 there is only encircling prey without search for prey. However, search for prey is an exploration phase, it helps to avoid premature convergence. To improve exploration ability of the later iteration, we define a new control parameter B. The parameter B is calculated as follows:

$$B = (a + 2) \cdot r - 2 \tag{17}$$

Where a is linearly decreased from 2 to 0 by Eq.(13) and r is a random number in [0,1]. So B is a random value in the interval $[-2, a]$. The range of B is shown in Fig. 1(b). A random search agent is chosen when $B \geq 0$, while the best solution is selected when $B<0$ for updating the position of the search agents. According to Fig.1(b), the search agents still have the opportunity to explore in the later iteration (see steps 8 and 12 in Fig. 2).

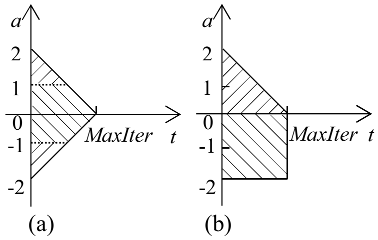

Figure 1. The control paramters: (a) the range of A, (b) the range of B

```
1 Initialize the whales population Xi(i=1,2,…,n)
2 Initialize a, A, C, and t=0
3 Evaluate f(Xi (t)), i=1,2,…,n
4 X*=the best search agent
5 While (t<MaxIter)
6   a=2(1-t/MaxIter)
7   for 1 i=1:n
8     Update A,C, l and B
9     for 2 j=1:d
10      p=rand()
11      if 1(p<0.5)
12        if 2(B<0)
13          Update X'ij by the Eq.(18)
14        else
15          Search a random search agent (Xrand)
16          Update X'ij by the Eq.(20)
17        end if 2
18      else
19        Update X'ij by the Eq.(19)
20      end if 1
21    end for 2
22    Check X'i and amend it
23    if 3 f(X'i)≤f(Xi (t))
24      Xi (t+1))= X'i
25    else
26      Xi (t+1))= Xi (t)
27    end if 3
28  end for 1
29  Update X* if there is a better solution
30  t=t+1
31end while
32return X*
```

Figure 2. Pseudo-code of the MWOA

C. Candidate Solution Selection

In the WOA, the newly generated candidate solutions are accepted without selection. This may cause the rebound of new search agents, which will reduce the convergence speed and the convergence accuracy. To decide whether or not it should become a member of generation t+1, the trial vector $X_i'(t+1)$ is introduced and the Eqs.(10,14,16) are altered to the Eqs.(18,19,20). Then $X_i'(t+1)$ is compared to the target vector $X_i(t)$ using the greedy criterion. If vector $X_i'(t+1)$ yields a smaller cost function value than $X_i(t)$, then $X_i(t+1)$ is set to $X_i'(t+1)$; otherwise, the old value $X_i(t)$ is retained(see steps 23~27 in Fig. 2).

$$\overrightarrow{X'}(t+1) = \overrightarrow{X^*}(t) - \vec{A} \cdot \vec{D} \tag{18}$$

$$\overrightarrow{X'}(\mathrm{t}+1) = \overrightarrow{D'} \cdot e^{bl} \cdot \cos 2\pi l + \overrightarrow{X^*}(t) \tag{19}$$

$$\overrightarrow{X'}(t+1) = \overrightarrow{X_{rand}}(t) - \vec{A} \cdot \vec{D} \tag{20}$$

V. EXPERIMENTAL EVALUATION

To develop our new technique for solving the parameter estimation problem of software reliability growth model we used the WOA and the proposed MWOA. Our objective is to estimate the parameters such that the error difference, between the actual fault and the estimated faults based the parameter estimated using WOA or MWOA, is minimal. RMSE criterion is used to measure the performance. It is just the square root of the mean square error as shown in Eq. (21):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{21}$$

Where N represents the number of measurements used for estimating the model parameters, y_i represents the i^{th} value of the effort, and $\hat{y}_i$ is the estimated effort. The performances of WOA and MWOA are evaluated by computing the objective evaluation criterion in terms of the fitness.

A. Test/Debug Data Sets

Three real measured test/debug data sets are applied in our proposed approaches. The first data set of 109 measurements and the second data set of 111 measurements presented in [10] are used for our experiments. The third data set has 46 measurements was presented in [11].

B. Parameter Estimation Based on Single Model

In our case, the ranges of the parameters for EXPM, POWM and DSSM models are set as follows: $\beta_0 \epsilon [0,2000]$, $\beta_1 \epsilon [0,1]$. For both the WOA and MWOA, a population size and maximum iteration number equal to 40 and 1000 have been utilized. The result reported in Table I is obtained based on 30 independent runs.

In Table I, it is found that the average RMSEs of MWOA are less than those of WOA in all instances. That shows the average optimization ability of the MWOA is better than that of the WOA. The standard deviations of the MWOA are small in all instances. That shows the MWOA is robust.

For data1 and data 2, we find that the delayed S-shaped model is able to provide the best results using the MWOA tuned parameters. The model error is the minimum compared to other proposed models. The exponential model, for data 3, obtains a best fitting result.

TABLE I. THE RESULTS OF SINGLE MODEL

Data	Model	RMSE of WOA		RMSE of MWOA	
		Avg	Std	Avg	Std
1	EXPM	28.749	0.0575	**28.702**	**5.33e-5**
1	POWM	41.670	0.6719	**41.336**	**3.90e-3**
1	DSSM	14.935	0.0306	**14.902**	**1.31e-13**
2	EXPM	28.124	0.0345	**28.102**	**6.33e-10**
2	POWM	47.943	0.2760	**47.687**	**3.07e-3**

2	DSSM	18.055	0.0034	**18.052**	**4.14e-14**
3	EXPM	12.281	0.3949	**12.161**	**8.48e-4**
3	POWM	12.451	0.2762	**12.278**	**1.15e-3**
3	DSSM	18.604	0.0339	**18.582**	**2.13e-10**

C. Parameter Estimation Based on Three-stage Model

According to section II, there are eight parameters to be estimated, six model parameters, β_{ij} (i=0,1;j=1,2,3), and two subsection parameters, t_i (i=1,2). The ranges of these parameters are given: $\beta_{0j}\in[0,2000]$, $\beta_{1j}\in[0,1]$ (j=1,2,3),t_i (i=1,2) must be integers, and $0\leqslant t_1 \leqslant t_2 \leqslant N_{data}$, N_{data} is the data number of a data set. For the MWOA, a population size and maximum iteration number equal to 40 and 1000 have been utilized. The estimated parameters based on three-stage model for the developed SRGM are given in Table II. From Table II, it is found that the results of three-stage model parameter estimation using MWOA are obviously better than those of single model. For three data sets, we found that the delayed S-shaped model is able to provide the best results using the MWOA tuned parameters. The model error is the minimum compared to other proposed models. The actual and accumulated faults (failures) curves for single models and three-stage models on 46 Measurements are shown in Figs. 3 and 4, respectively. The estimated failures using three-stage DSSM are in good agreement with the actual ones in Fig. 4.

TABLE II. THE RMSES OF MWOA BASED ON THREE-STAGE MODEL, RMSE1 FOR SINGLE MODEL, RMSE2 FOR THREE-STAGE MODEL

Data	Model	RMSE1	RMSE2
1	EXPM	28.701	**5.5842**
1	POWM	41.334	**8.2670**
1	DSSM	14.902	**5.1660**
2	EXPM	28.102	**9.2153**
2	POWM	47.686	**12.004**
2	DSSM	18.052	**8.7049**
3	EXPM	12.161	**6.4699**
3	POWM	12.277	**8.6090**
3	DSSM	18.582	**4.5121**

VI. CONCLUSION

WOA is a new swarm intelligence optimization algorithm. This algorithm is not perfect enough. Based on the analysis of WOA, we point out the disadvantages of WOA, and propose a modified WOA (MWOA). We use MWOA to estimate the parameters of single SRGMs. Experimental results show that MWOA outperforms WOA, and gets less RMSEs.

A three-stage software reliability growth model is proposed in this paper. MWOA is used to estimate the parameters of three-stage software reliability models. Experimental results show that the fitting accuracy of three-stage model is significantly better than that of a single model. The three-stage model can provide a better estimate of the software faults.

REFERENCES

[1] M. R. Lyu, "Handbook of software reliability engineering," IEEE Computer Society Press,1996.

[2] K.Bidhan, A. Awasthi, "A Review on Parameter Estimation Techniques of Software Reliability Growth Models," International Journal of Computer Applications Technology and Research, vol. 3, no. 4, pp. 267-272, 2014.

[3] Z. ALRahamneh, M. Reyalat, A. F. Sheta, et al., "A New Software Reliability Growth Model: Genetic-Programming-Based Approach," Journal of Software Engineering and Applications, vol. 4, no. 8, pp. 476-481, 2011.

[4] M. Alweshah, W. Ahmed, H. Aldabbas, "Evolution of Software Reliability Growth Models: A Comparison of Auto-Regression and Genetic Programming Models," International Journal of Computer Applications, vol. 125, no. 3, pp. 20-25, 2015.

[5] S. Mirjalili, A. Lewis, "The Whale Optimization Algorithm," Advances in Engineering Software, vol. 95, pp. 51-67, 2016.

[6] M. Zhong, W. Long, "Whale optimization algorithm with nonlinear control parameter," MATEC Web Conf. vol. 139, pp. 1-5, 2017.

[7] C. Zheng, X. Liu, S. Huang, "Estimating parameters of software reliability models by ant colony algorithm," Journal of Computer Applications, vol.32, no.4, pp.1147-1151, 2012.

[8] X. Xiao, T. Dohi, "Estimating software reliability using extreme value distribution," Software Engineering, Business Continuity, and Education. Springer Berlin Heidelberg, pp. 399-406, 2011.

[9] A. Geol, K. Okumoto, "Time-dependent error-detection rate model for software reliability and other performance measures," IEEE Trans. Reliability, vol. 28, pp. 206–211, 1979.

[10] L. H. Crow, "Reliability for complex repairable systems," Reliability and Biometry, SIAM, pp. 379–410, 1974.

[11] S. Yamada, M. Ohba, and O. S., "S-Shaped relaibility growth modeling for software error detection," IEEE Trans. Reliability, vol. 32, no. 5, pp. 475–484, 2009.

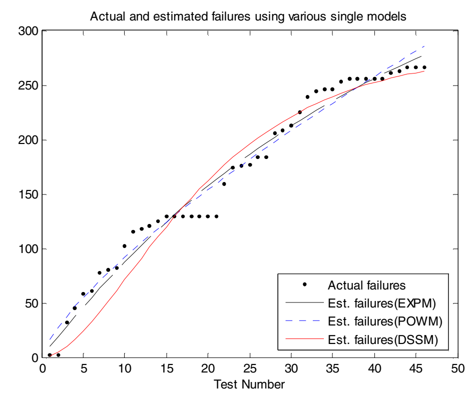

Figure 3. Actual and estimated failures for MWOA using single models on 46 Measurements

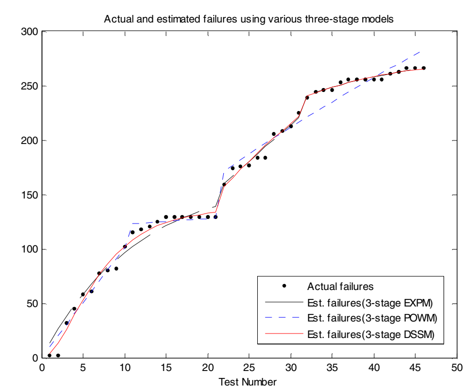

Figure 4. Actual and estimated failures for MWOA using three-stage models on 46 Measurements