The Estimation And Prediction Model Of The Software Reliability With The Project Size Index

Dmytro Fedasyuk, Maksym Seniv, Yaroslav Chabanyuk, Ulyana Khimka

Abstract - **The new mathematical model introducing a quality index of the software project size determined at the testing stage for the software reliability estimation and prediction is suggested..**

Keywords - **index of the project, the software reliability, nonhomogeneous Poisson process.**

I. INTRODUCTION

We offer a new mathematical model of software evaluation and reliability prediction with index of the project size, which assumes sizes on interval defining the project size. The introduced index is the function parameter of cumulative number failures and occurrence intensity errors.

II. FORMALIZATION OF THE INDEX OF PROJECT SIZE

Nowadays, there are many works on software reliability models development. In some works the approach to the model building, according to which software behavior in terms of reliability is undefined by nature since each program is unique, is suggested. The review of reliable as well as other methods of software analysis shows that all the models involve the change of errors characteristics in the programme within the limits of their correction, i.e. the time before a regular error occurrence increases [1].

TABLE 1

THE RELATION BETWEEN THE NONHOMOGENEOUS POISSON PROCESS MODELS

MODEL NAME	The failures revealing intensity function $\lambda(t)$
Goel-Okumoto Model	$\lambda(t) = Nb\exp(-bt)$
Schneidewind Model	$\lambda(t) = \alpha_0 \exp(-\beta t)$
Musa Basic Model	$\lambda(t) = \beta_0\beta_1 \exp(-\beta_1 t))$
S-shaped Model	$\lambda(t) = \alpha\beta^2 t exp(-\beta t))$
Generalized model of nonhomogeneous Poisson process	$\lambda(t) = \alpha\beta^{n+1} t^n exp(-\beta t))$

The table presents the failures revealing intensity function $\lambda(t)$ for the basic and proposed models. Their parameters are linked in the following relations:

$N = \alpha,\ \beta_0 = \alpha,\ b = \beta,\ \beta_1 = \beta,\ \alpha_0 = \alpha\beta,$ where α - the total number of failures in the course of observations, β - the change rate of the failures revealing intensity function.

It should be mentioned that the generalized model of nonhomogeneous Poisson process is proposed in the work [2].

To improve the accuracy of this model, the author suggests applying such curve intensity form of the failures detection, which introduces an additional parameter n for the project size estimation, where the choice of parameter n depends on the testing process with the following recommended sizes:

n =0 – for a small project, where the developer is tester at the same time (Musa, Goel-Okumoto and Schneidewind models);

n =1 – for the average project, where testing and designing of software are produced by different persons from one working group (S-shaped model);

n =2 – for a large project, where the software development and testing groups work on the project simultaneously;

n=3 – for a very large project, where testing and development departments are independent.

The researches of this model have been conducted and the obtained graph of the failures revealing intensity curve is presented in Figure 1:

Following that the proposed size of parameter β [2]in case n =3 the curve of failures revealing intensity function $\lambda(t)$ was negative, we have offered to introduce the continuous index s, which defines the project size and generalizes S - shaped model of the software reliability prediction, i.e. the following view of the failures revealing intensity function $\lambda(t)$ is suggested:

$$\lambda(t) = \alpha\beta^{s+1} t^s exp(-\beta t)) \qquad (1)$$

When s =1, the model with a real index of the project size is reduced to the S-shaped model.

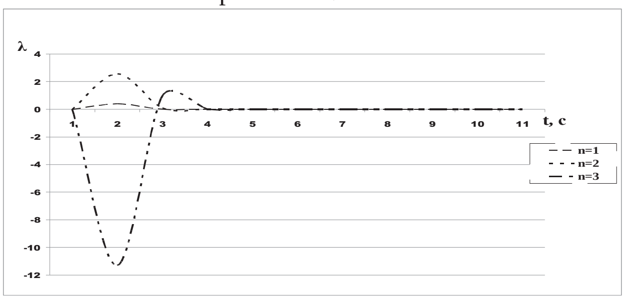

Fig.1 The Graph of failures revealing intensity curve

For the intensity (1) cumulative function is

$$\mu(t) = \int_0^t \lambda(\tau)d\tau = \alpha\left[-\beta^s t^s e^{-\beta t} + s\Gamma_{\beta t}(s)\right], \qquad (2)$$

where $\Gamma_z(p) = \int_0^z t^{p-1}e^{-t}dt$, $\operatorname{Re} p > 0$, - incomplete Γ - function. It should be noted that, when $s = 1$, cumulative function (2) coincides with a view of cumulative function S - shaped model. Thereby, when $t \to \infty$ with (2) we get

$$\mu(\infty) = \alpha s\Gamma(s), \qquad (3)$$

where $\Gamma(s)$ - Γ - function. Since when $s = 1$ $\Gamma(1) = 1$, then $\mu(\infty) = \alpha$,that equal to S - shaped model.

The equation (1) or (2) is to be named the model with the project size index. The method of maximum likelihood is to be applied to determine the point estimations of the parameters α, β, s.

Let on interval $(t_i, t_{i+1}], i=\overline{0,n}$ has been revealed m_i mistakes ($\sum_{i=1}^{n} m_i = k$). The function of maximum likelihood $L(\alpha,\beta,s)$ in assumption of Poisson distribution of errors quantity on the interval $(t_i, t_{i+1}], i=\overline{0,n}$ should be plotted:

$$L(\alpha,\beta,s)=\prod_{i=1}^{n}\frac{[\mu(t_i)-\mu(t_{i-1})]^{m_i}}{m_i!}\exp(\mu(t_{i-1})-\mu(t_i))$$

Taking into consideration $\mu(t)$ from the equitation (2), we get:

$$L(\alpha,\beta,s)=\prod_{i=1}^{n}\frac{\alpha^{m_i}}{m_i!}\times$$
$$\times[s(\Gamma_{\beta t_i}(s)-\Gamma_{\beta t_{i-1}}(s))+$$
$$\beta^s(t_{i-1}^s\exp(-\beta t_{i-1})-t_i^s\exp(-\beta t_i)]^{m_i}\times$$
$$\times\exp(\mu(t_{i-1})-\mu(t_i)).$$

The function of likelihood L satisfies the conditions:

1) is differentiated with the arbitrary sizes of the sample;
2) reaches the maximum in the interval of possible sizes.

To get the estimations α, β, s we should solve the system of equations:

$$\begin{cases}\dfrac{\partial L(\alpha,\beta,s)}{\partial\alpha}=0,\\ \dfrac{\partial L(\alpha,\beta,s)}{\partial\beta}=0,\\ \dfrac{\partial L(\alpha,\beta,s)}{\partial s}=0;\end{cases}\qquad(4)$$

The function $\ln L$ is considered instead of function L in the system (4) without reducing the generality and after elementary transformations the set of equations is obtained:

$$\alpha=\frac{k}{s\Gamma_{\beta t_n}(s)-\beta^s t_n^s\exp(-\beta t_n)};$$

$$\sum_{i=1}^{n}\frac{m_i(t_i^{s+1}\exp(-\beta t_i)-t_{i-1}^{s+1}\exp(-\beta t_{i-1}))}{s\Phi_{\beta t_i t_{i-1}}(s)+\beta^s(t_{i-1}^s\exp(-\beta t_{i-1})-t_i^s\exp(-\beta t_i))}-$$
$$-\frac{kt_n^{s+1}\exp(-\beta t_n)}{s\Gamma_{\beta t_n}(s)-\beta^s t_n^s\exp(-\beta t_n)}=0;\quad(5)$$

$$\sum_{i=1}^{n}\frac{m_i}{s\Phi_{\beta t_i t_{i-1}}(s)+\beta^s(t_{i-1}^s\exp(-\beta t_{i-1})-t_i^s\exp(-\beta t_i))}\times$$
$$\times[\Phi_{\beta t_i t_{i-1}}(s)+s\mathrm{F}_{\beta t_i t_{i-1}}(s)+$$
$$\beta^s(t_{i-1}^s\exp(-\beta t_{i-1})\ln(\beta t_{i-1})-t_i^s\exp(-\beta t_i)\ln(\beta t_i))]+$$
$$+\frac{k}{s\Gamma_{\beta t_n}(s)-\beta^s t_n^s\exp(-\beta t_n)}\times$$
$$[\beta^s t_n^s\exp(-\beta t_n)\ln(\beta t_n)-\Gamma_{\beta t_n}(s)-$$
$$-s\int_0^{\beta t_n}\tau^{s-1}\ln\tau\exp(-\tau)d\tau]=0;$$

where $\Phi_{\beta t_i t_{i-1}}(s)=\int_{\beta t_{i-1}}^{\beta t_i}\tau^{s-1}\exp(-\tau)d\tau$,

$$\mathrm{F}_{\beta t_i t_{i-1}}(s)=\int_{\beta t_{i-1}}^{\beta t_i}\tau^{s-1}\ln\tau\exp(-\tau)d\tau.$$

We suggest solving the system of transcendental equations (5) by the approximate method and applying the obtained approximate sizes $\hat{\alpha}, \hat{\beta}, \hat{s}$ for the analysis of the testing termination, the project size, the software product reliability. The representation (2) of cumulative function $\mu(t)$ with obtained parameter $\hat{s}$ form the system (5) gives an opportunity to estimate the parameter $\hat{\alpha}$.

Remark 1. The first stage of the problem algorithm of likelihood function L creation is the intervals introduction $(t_i, t_{i+1}]$, $i=\overline{0,n}$ where the errors quantity m_i has the Poisson distribution.

Remark 2. The determination of point estimations $\hat{\alpha}, \hat{\beta}, \hat{s}$ characteristics, namely validity, efficiency, unbiasedness is the subject of further researches.

Applying the test examples of work [3], we get the following interval criteria of the project size index:

when $s\in[0;0{,}7)$ the project can be classified as small,

when $s\in[0{,}7;1{,}5)$ the project can be classified as medium,

when $s\in[1{,}5;2{,}2)$ the project can be classified as big,

when $s\in[2{,}2;2{,}7)$ - huge.

III. CONCLUSION

Introduced estimation and prediction model of software reliability can be used on enterprise for estimation the total error quantity of software and for prediction the frequency of their appearance. On the basis of given model the criterion of testing adequacy can be determined, which allows to determine the test termination on the enterprise and the precise time of putting the software product into operation. The index of project size allows giving practical recommendations for project managers on production means distribution between the phases of project life cycle.

REFERENCES

[1] Половко А.М., Гуров С.В. Основи теорії надійності. Спб., Bhv. 702с.

[2] Дидковская М.В. Інтеграційне тестування компонентно-базованого програмного забезпечення: критерії, оцінки, метод вибору // НТУУ КПІ, 2005, VII МНПК "Системний аналіз та інформаційні технології", Київ, 2005. С.189.

[3] Kai-Yuan Cai, De-Bin Hu, Cheng-Gang Bai, Hai Hu, Tao Jing *Does software reliability growth behavior follow a non-homogeneous Poisson process* // Information and Software Technology, Vol. 50 (2008), pp. 1232–1247.