International Journal of Soft Computing 12 (5-6): 313-317, 2017
ISSN: 1816-9503

The Performance Analysis Comparative Study Depend on Software Reliability Model and Curve Regression Model

Tae-Jin Yang
Department of Academic Cooperation Foundation, Namseoul University,
Cheonan, Republic of Korea

Abstract: Software reliability in the software development process is an important issue. Software process improvement helps in finishing with reliable software product. The proposed process involves evaluation of the parameter of the mean value function and hence, the mean value function of infinite, finite failure model and curve regression model was considered which is based on Non Homogeneous Poisson Process (NHPP). This study, a failure data analysis of applying using time between failures and parameter estimation using maximum likelihood estimation method were made after the efficiency of the data through trend analysis, model selection was performed using the mean square error and coefficient of determination.

Key words: Software reliability model, NHPP, curve regression model, determination, development, efficiency

INTRODUCTION

Software failures were caused by failure of computer systems in our society can lead to huge losses. Thus, software reliability in the software development process is an important issue. These issues of the user requirements meet the cost of testing. Software testing (debugging) in order to reduce costs in terms of changes in the software reliability and testing costs, need to know in advance is more efficient. Thus, the reliability, cost and consideration of release time for software development process are essential. Eventually, the software to predict the contents of a defect in the product development model is needed. Until now, many software reliability models have been proposed. Non-Homogenous Poisson Process (NHPP) models rely on an excellent model (Gokhale and Trivedi, 1999; Goel and Okumoto, 1979) in terms of the error discovery process and if a fault occurs, immediately remove the debugging process and the assumption that no new fault has occurred.

In this field, enhanced non-homogenous poisson process model was presented (Gokhale and Trivedi, 1999), proposed an exponential software reliability. In this model, the total numbers of defects have S-shaped or exponential-shaped with a mean value function was used model (Goel and Okumoto, 1979). The generalized model relies on these models, delayed S-shaped reliability growth model and inflection S-shaped reliability growth model were proposed (Yamada *et al*., 1983). Software reliability problems in change point were proposed software reliability problems in change point (Zhao, 1993) and the generalized reliability growth models proposed (Shyur, 2003). In testing measured coverage, the stability of model with software stability can be evaluated was presented (Pham and Zhang, 2003).

Relatively recently, generalized logistic testing-effort function and the change-point parameter by incorporating efficient techniques to predict software reliability were presented (Huang, 2005). The learning process that software managers to become familiar with the software and test tools for S-type model can be explained (Chiu *et al*., 2008). In addition, the comparative study of NHPP delayed S-shaped and extreme value distribution software reliability model using the perspective of learning effects was studied (Kim, 2013) and the comparative study of software optimal release time based on NHPP software reliability model using exponential and log shaped type for the perspective of learning effect was studied (Shin and Kim, 2013).

The proposed process involves evaluation of the parameter of the mean value function and hence the mean value function of infinite, finite failure model which is based on Non Homogeneous Poisson Process (NHPP) and curve regression models were considered.

Literature review

NHPP model: The mean value function and the intensity function (Gokhale and Trivedi, 1999) for Non-Homogeneous Poisson Process (NHPP) Model are given by:

$$m(t) = \int_0^t \lambda(s)ds, \ \frac{dm(t)}{dt} = \lambda(t) \quad (1)$$

Therefore, N(t) is known Poisson Probability Density Function (PDF) with the parameter m(t). In other words:

$$p[N(t)=n]=\frac{[m(t)]^n}{n!}e^{-m(t)},\ n=0,1,\cdots,\infty \quad (2)$$

MATERIALS AND METHODS

These time domain models for the NHPP process can be described by the probability of failure are possible. This model is the failure intensity function λ(t) expressed differently, also mean value the function m(t) will be expressed differently.

These models are classified into categories, the finite failure NHPP models and infinite failure (Kuo and Yang, 1996). Finite failure NHPP models if they are given sufficient time to test, the expected value of faults has a finite expectation value $(\lim_{t\to\infty} m(t)=\theta<\infty)$ while infinite failure NHPP model assumes that the value is infinite.

Let θ denote the expected number of faults that would be detected given finite failure NHPP models. Then, the mean value function of the finite failure NHPP models can also be written as:

$$m(t)=\theta F(t) \quad (3)$$

Note that F(t) is a CDF (cumulative distribution function). From Eq. 3, the (instantaneous) failure intensity λ(t) in case of the finite failure NHPP models is given by:

$$\lambda(t)=\theta F'(t)=\theta f(t) \quad (4)$$

In finite model, at the time of each repair, a new defect is assumed not to occur. However, the actual situation at the point of repair new failure may occur.

Add to this situation, the RVS (Record Value Statistics) model could be used NHPP model and mean value function was as follows (Kuo and Yang, 1996):

$$m(t)=-\ln[1\text{-}F(t)] \quad (5)$$

Equation 5 is mean value function of infinite failure NHPP model. Therefore, from Eq. 5 using the related equations of NHPP in Eq. 1, intensity function can be the hazard function (h(t)). In other words:

$$\lambda(t)=m'(t)=f(t)/(1-F(t))=h(t) \quad (6)$$

Note that ∫(t) is a PDF (the Probability Density Function). Let $\{t_n, n = 1, 2, ...\}$ denote the sequence of times between successive software failures. Then, t_n denote the time between (n-1st) and nth failure. Let x_n denotes nth failure time, so that:

$$x_n=\sum_{k=1}^{n} t_k\ (k=1,2,...,n;\ 0\le x_1\le x_2\le\cdots\le x_n) \quad (7)$$

The joint density or the likelihood function of $x_1, x_2, ..., x_n$ can be written as Gokhale and Trivedi (1999):

$$f_{X_1,X_2,...,X_n}(x_1,x_2,...,x_n)=e^{-m(x_n)}\prod_{i=1}^{n}\lambda(x_i) \quad (8)$$

For a given sequence of software failure times $(x_1, x_2, ..., x_n)$ that are realizations of the random variables $(X_1, X_2, ..., X_n)$ the parameters of the software reliability NHPP models are estimated using the Maximum Likelihood method (MLE) (Gokhale and Trivedi, 1999; Kuo and Yang, 1996).

Software reliability model based on finite and infinite NHPP using Rayleigh distribution: In this study, Rayleigh distribution model was applied. The Rayleigh distribution was originally derived in connection with a problem in acoustics and has been used in modeling certain features of electronic waves and as the distance distribution between individuals in a spatial Poisson process. Most frequently, however, it appears as a suitable model in life testing and reliability theory. Rayleigh distribution and the distribution function of the probability density function are known as follows (Shin and Kim, 2014). Note: b>0, t∈(0, ∞):

$$f(t)=2bt\exp(-bt^2),\ F(t)=1-\exp(-bt^2) \quad (10)$$

As a result, finite failure NHPP intensity function and the mean value function can be expressed, using Eq. 3 and 4 as:

$$\begin{aligned}\lambda(t)&=\theta f(t)=2\theta bt\exp(-bt^2),\\ m(t)&=\theta F(t)=\theta\left[1-\exp(-bt^2)\right]\end{aligned} \quad (11)$$

Note that b refers to the shape parameter. In addition, using Eq. 10, the hazard function is derived as follows:

$$h(t)=f(t)/(1-F(t))=2bt \quad (12)$$

The likelihood function, using Eq. 8 is as follows:

$$L_{NHPP}(\Theta|\underline{x})=\left[\prod_{i=1}^{n}2\theta bx_i e^{-bx_i^2}\right]\exp\left[-\theta(1-e^{-bx_n^2})\right] \quad (13)$$

Where:
$\underline{x}$ = $x_1, x_2, x_3, ..., x_n$
Θ = Parameter space

Using Eq. 13, $\hat{\theta}_{MLE}$ and $\hat{b}_{MLE}$ are obtained as the solutions of the following equation:

$$\hat{\theta} = \frac{n}{1-e^{-\hat{b}x_n^2}}, \frac{n}{\hat{b}} = \sum_{i=1}^{n} x_i^2 + \hat{\theta} x_n^2 e^{-\hat{b}x_n^2} \quad (14)$$

On the other hand, using Eq. 5 and 6, intensity function and mean value function of the infinite NHPP can be expressed as:

$$\lambda(t) = m'(t) = f(t)/(1-F(t)) = h(t) = 2bt \quad (15)$$

$$m(t) = -\ln\left[1-F(t)\right] = bt^2 \quad (16)$$

Where:
f(t) = A PDF (the Probability Density Function)
F(h) = A CDF (Cumulative Distribution Function)
h(t) = Hazard function

The likelihood function, using Eq. 8 is as follows:

$$L_{NHPP}(\Theta|\underline{x}) = \left[\prod_{i=1}^{n} 2bx_i\right]\exp(-bx_2^n) \quad (17)$$

The log-likelihood function to Maximum Likelihood Estimation (MLE) is derived as follows using Eq. 17:

$$\ln L_{NHPP}(\Theta|\underline{x}) = n\ln 2 + n\ln b + \sum_{i=1}^{n} \ln x_i \text{ -} bx_n^2 = 0 \quad (18)$$

Therefore, the partial differential equation with respect to b using Eq. 18, the maximum likelihood estimation can be calculated as follows:

$$\hat{b}_{MLE} = \frac{n}{x_n^2} \quad (19)$$

Logarithm curve and growth curve regression model Now, we consider the time series regression models. We might use simple linear regression:

$$Y_t = a + bt + w_t \quad (20)$$

Where:
Y_t = Total number of failure observed at time t
t = Not variable but failure time
a and b = Parameters
w_t = Error term

and was considered the logarithm curve model and growth curve model instead of linear regression model. The logarithm model and growth curve models are as follows (Ra and Kim, 2013; Song and Chang, 2014):

$$Y_t = a + b\ln t, \ln Y_t = a + bt \quad (21)$$

In this study, applied models were listed in Table 1 for the sake of model comparison.

Table 1: Mean value function of reliability model and time curves regression

Types	Models	Mean value function
Software reliability	Infinite NHPP	$m(t) = -\ln(1\text{-}F(t)) = bt^2$
	Finite NHPP	$m(t) = \theta\ F(t) = \theta\ [1\text{-}\exp(\text{-}bt^2)]$
Curve regression	Logarithm	$Y_t = a+b\ \ln t$
	Growth	$\ln Y_t = a+bt$

Table 2: Failure time data (Hayakawa and Telfar, 2000)

Failure number	Failure time (h)	Failure number	Failure time (h)
1	0.479	16	10.771
2	0.745	17	10.906
3	1.022	18	11.183
4	1.576	19	11.779
5	2.610	20	12.536
6	3.559	21	12.973
7	4.252	22	15.203
8	4.849	23	15.640
9	4.966	24	15.980
10	5.136	25	16.385
11	5.253	26	16.960
12	6.527	27	17.237
13	6.996	28	17.600
14	8.170	29	18.122
15	8.863	30	18.735

Comparison criterion for the effectiveness: In order to investigate the effectiveness of the proposed model, the comparison criterion used the mean square error and R^2 (Gokhale and Trivedi, 1999; Chiu *et al*., 2008). The Mean Square Error (MSE) measures the deviation between the predicted values with the actual observations. It is defined as:

$$MSE = \frac{\sum_{i=1}^{n}[m(x_i) - \hat{m}(x_i)]^2}{n-k} \quad (22)$$

Where:
m_i = The total cumulated number of errors observed within time $(0, t_i)$
$\hat{m}_i$ = The estimated cumulative number of errors at time x_i obtained from the fitting mean value function
n = The number of observations
k = The number of parameters to be estimated

R-square (R^2) can measure how successful the fit is in explaining the variation of the data. It is defined as Table 2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}[m(x_i) - \hat{m}(x_i)]^2}{\sum_{i=1}^{n}\left(m(x_i) - \sum_{j=1}^{n} m(x_j)/n\right)^2} \quad (23)$$

RESULTS AND DISCUSSION

Illustration: We construct the corresponding software reliability growth model by using the date in Table 1 (Hayakawa and Telfar, 2000). In general, the Laplace trend test analysis is used (Kanoun and Laprie, 1996) for reliability property. As a result of this test in this Fig. 1 as indicated in the Laplace factor is between 2 and -2, reliability growth shows the properties. Thus, using this data, it is possible to estimate the reliability (Kanoun and Laprie, 1996).

In this study, numerical conversion data (Failure time (hours)$\times$10) in order to facilitate the parameter estimation was used. The result of parameter estimation has been summarized in Table 3. These calculations, solving numerically, the initial values given to 0.001 and 3 and tolerance value for width of interval (10^{-5}) given using C-language checking adequate convergent were performed iteration of 100 times. In this Table 3, coefficients of curve regression ware estimated using Statistics Package SPSS 18 (Version).

The result of Mean Square Error (MSE) and coefficient of determination (R^2) are has been summarized in Table 3. For software reliability in Table 3, MSE (which measures the difference between the actual value and the predicted value) show that finite NHPP Model than infinite NHPP has small value. Therefore, the finite NHPP is appreciably better than the infinite NHPP Model. Also, for curve regression, MSE show that logarithm model than growth model has small value. Therefore, the logarithm model is appreciably better than the growth model. R^2 (which means that the predictive power of the difference between predicted values) show that finite NHPP than other model has high value, finite model is the utility model. Eventually, in terms of MSE and R^2, finite NHPP regard as best model because MSE is the smallest and R^2 is the highest than other models.

The result of mean value functions are has been summarized in Fig. 2. In Fig. 2, patterns of mean value function have the tendency of non-decreasing form. Also, finite NHPP show that finite NHPP model than any models is the utility model because finite NHPP model than any model estimates close to the true value from mean value functions.

Table 3: Parameter estimation and model comparison

Types	Models	MLE	Model comparison MSE	R^2
Software reliability				
Infinite NHPP	-	$\hat{b} = 8.55\times10^{-4}$	31.1951	0.60
Finite NHPP	$\hat{\theta} = 35.306$	$\hat{b} = 5.43\times10^{-5}$	9.4979	0.88
Curve regression				
Logarithm	$\hat{a} = -18.712$	$\hat{b} = 8.086$	10.5051	0.84
Growth	$\hat{a} = 1.253$	$\hat{b} = 0.013$	17.1149	0.79

MLE: Maximum Likelihood Estimation; MSE: Mean Squared Error; R^2: Coefficient of determination

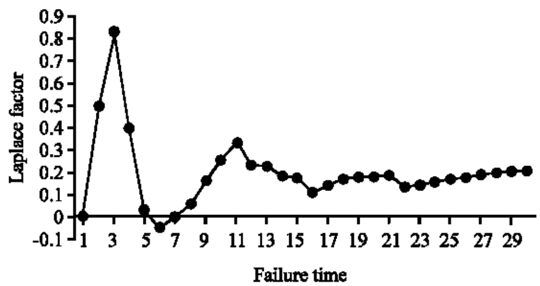

Fig. 1: The results of Laplace trend test

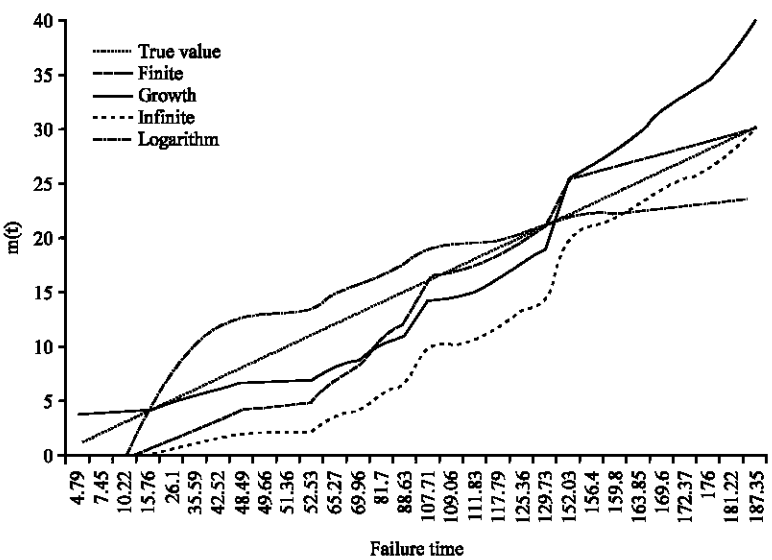

Fig. 2: Mean value function finite, infinite NHPP and time curve regression

CONCLUSION

The mean value functions for NHPP software reliability model and time curve regression models are considered. In this study, the following conclusions were obtained.

For software reliability, mean square error which measures the difference between the actual value and the predicted value show that finite NHPP Model than infinite NHPP has small value. Therefore, the finite NHPP is appreciably better than the infinite NHPP Model. Also, for curve regression, mean square error shows that logarithm model than growth model has small value. Therefore, the logarithm model is appreciably better than the growth model. The coefficient of determination which means that the predictive power of the difference between predicted values show that finite NHPP than other model has high value, finite model is the utility model. Eventually, in terms of mean square error and coefficient of determination, finite NHPP regard as best model because mean square error is the smallest and coefficient of determination is the highest than other models. The patterns of mean value function have the tendency of non-decreasing form. Also, finite NHPP show that finite NHPP Model than any models is the utility model because finite NHPP Model than any model estimates close to the true value from mean value functions. Therefore, in this study, the proposed software reliability model and time curve regression models can be used as reasonable model in this field. As an alternative to this area judge that the content is a valuable research.

ACKNOWLEDGEMENT

Funding for this study was provided by Namseoul University.

REFERENCES

Chiu, K.C., Y.S. Huang and T.Z. Lee, 2008. A study of software reliability growth from the perspective of learning effects. Reliab. Eng. Syst. Safety, 93: 1410-1421.

Goel, A.L. and K. Okumoto, 1979. Time-dependent error-detection rate model for software reliability and other performance measures. IEEE. Trans. Reliab., 28: 206-211.

Gokhale, S.S. and K.S. Trivedi, 1999. A time/structure based software reliability model. Ann. Software Eng., 8: 85-121.

Hayakawa, Y. and G. Telfar, 2000. Mixed poisson-type processes with application in software reliability. Math. Comput. Model., 31: 151-156.

Huang, C.Y., 2005. Performance analysis of software reliability growth models with testing-effort and change-point. J. Syst. Software, 76: 181-194.

Kanoun, K. and J.C. Laprie, 1996. Handbook of Software Reliability Engineering. In: Chapter Trend Analysis, Lyu, M.R. (Ed.). McGrawHill, New York, USA., pp: 401-437.

Kim, H.C., 2013. The comparative study of NHPP delayed S-shaped and extreme value distribution software reliability model using the perspective of learning effects. Int. J. Adv. Comput. Technol., 5: 1210-1218.

Kuo, L. and T.Y. Yang, 1996. Bayesian computation for nonhomogeneous poisson processes in software reliability. J. Am. Stat. Assoc., 91: 763-773.

Pham, H. and X. Zhang, 2003. NHPP software reliability and cost models with testing coverage. Eur. J. Opert. Res., 145: 443-454.

Ra, Y.S. and H.C. Kim, 2013. Study for Predict of the Future Software Failure Time using Nonlinear Regression. In: IT Convergence and Security 2012, Kim, K. and K.Y. Chung (Eds.). Springer, Dordrecht, Netherlands,ISBN:978-94-007-5859-9,pp:1117-1124.

Shin, H. D. and H.C. Kim, 2014. Assessing rayleigh software reliability model based on NHPP using SPC. J. Next Gen. Inform. Technol., 5: 77-84.

Shin, H.D. and H.C. Kim, 2013. The comparative study of software optimal release time based on NHPP software reliability model using exponential and log shaped type for the perspective of learning effects. Int. J. Adv. Comput. Technol., 5: 120-129.

Shyur, H.J., 2003. A stochastic software reliability model with imperfect debugging and change-point. J. Syst. Software, 66: 135-141.

Song, K.Y. and I.H. Chang, 2014. Parameter estimation and prediction for NHPP software reliability model and time series regression in software failure data. J. Chosun Nat. Sci., 7: 67-73.

Yamada, S., M. Obha and S. Osaki, 1983. S-Shaped reliability growth modeling for software error detection. IEEE Trans. Reliab., 12: 475-478.

Zhao, M., 1993. Change-point problems in software and hardware reliability. Commun. Stat. Theory Meth., 22: 757-768.