Study On Unification of Software Reliability Growth Models Srgms)

Jyoti, Former Assistant Professor, Government College Of Women, Bawani Khera, Bhiwani (Haryana)
Email- Neer4ualways@Gmail.Com

Abstract

In the past 35 years numerous software reliability growth models (SRGMs) have been proposed under diverse testing and debugging (T&D) environments and applied successfully in many real life software projects but no SRGM can claim to be the best in general as the physical interpretation of the T&D is not general. Unified modeling approach proves to be very successful in this regard and provides an excellent platform for obtaining several existing SRGM following single methodology. It forms the main focus of this paper; here we propose a unification modeling approach applying the infinite server queuing theory based on the basic assumptions of an SRGM defining three level of complexity of faults. Our unification methodology can be used to obtain several existing and new SRGMs which consider testing a one stage process with no fault categorization, two/three stage process with random delay functions and hence categorize faults in two/three level complexity.

Keywords: Unification, Software Reliability Growth Models (SRGMs)

Introduction:

Software reliability growth models (SRGMs) are widely used to estimate quantitative software reliability measures. The SRGMs are generally defined as stochastic counting processes regarding the number of faults detected or failures experienced in testing phase. Based on the counting process, we compute some of software reliability measures such as the quantitative software reliability which is defined as a probability that no failure occurs during a certain time interval. Since the seminal works of Jelinski and Moranda and Goel and Okumoto, a number of SRGMs have been competitively proposed to improve their fitting abilities and accuracy of the future failure prediction. In particular, non-homogeneous Poisson processes (NHPPs) have much popularity for the software reliability modeling due to their mathematical tractability.

The quantitative software reliability assessment generally consists of the following steps: (i) Collect software failure data, (ii) Determine a set of candidates of models, (iii) Estimate parameters for all the candidates, (iv) Choose the best model under a certain criterion, and (v) Calculate the quantitative measures using the best model. In the step (i), we collect the failure time data in software testing or the number of failures per working day. The steps (iii) and (iv) can be solved by statistical arguments such as the maximum likelihood (ML) estimation and the theory of information criterion. In the step (v), the software reliability measures can be computed according to their formulas. However, in the step (ii), there is no definitive method to determine a suitable set of candidates. If the number of candidates is large, it causes a large effort in the parameter estimation procedure of the step (iii). At the same time, since there have been over 200 SRGMs proposed in many papers, it is infeasible to estimate the model parameters for all the existing SRGMs. On the other hand, if the set did not include the sufficient number of candidates, it causes the degradation of fitting ability and prediction accuracy. That is, it is important to determine a suitable set of candidates.

The modeling framework is helpful to determine a set of candidates, which involves some of the existing NHPP-based SRGMs. The most well-known and well-defined modeling framework has been proposed by Langberg and Singpurwalla. They revealed the statistical structure of a class of NHPP-based SRGMs. According to their approach, NHPP-based SRGMs consists of only two factors; the number of total software faults before testing (the number of total failures experienced in the software lifetime) and failure time caused by each fault in testing. In particular, when the number of total software faults is given by a Poisson random variable, the framework covers the NHPP-based SRGMs whose mean value function is bounded by a finite value. Then the concrete formulas of NHPP-based SRGMs are dominated by only types of the failure time distribution. In their framework, the candidate selection of NHPP-based SRGMs can be reduced to selecting a set of suitable failure time distributions. In fact, typical NHPP-based SRGMs consist of well-known statistical distributions such as exponential, gamma Pareto logistic and extreme-value (Weibull, Frechet and Gompertz) distributions. Generally, in the area of reliability engineering, the exponential, gamma, Pareto, logistic, extreme-value distributions are also applied to representing the hardware failure time distribution.

Unification of SRGMs

A SRGM is said to exhibit relationships between quantitative variables under some set of assumptions. These assumptions are used so as to portray the system. The emerging research direction for SRGMs that is usually considered is the unification of SRGM. Unification is a more general formulation comprising some typical reliability growth patterns. It is an insightful investigation wherein many models are studied without the need of making many different assumptions. The main aim behind unification of SRGMs is robust assessment of software reliability.

Reliability Models Issues

A number of reliability growth models have been proposed from the basic exponential model developed by Goel and Okumoto (1979). Unfortunately, there is no such model that can be accepted as "the" model for use in various different situations. A model is supposed to be an abstraction for a specific aspect of study and cannot include everything. For that reason we have to specify the factors on which our model is build. This is the reason that the existing reliability growth models are based on assumptions (their limitations). Some of the issues that are faced by SRGMs are as follows: SRGMs are not tested among different situations that can be present when the software is being executed in the operational profile. SRGMs are restricted to certain limitations.

- It is often difficult for the average software engineers to comprehend it.
- Different types of model are required for different application areas e.g. realtime, control intensive, interactive etc.
- Data gathering is expensive and is a difficult task
- The restriction of SRGM to certain assumptions can somehow be solved by the unification methodology, an emerging research area in the software reliability field.

Unification Methodology

The NHPP based SRGMs are formulated considering different T&D environments. The existing models have been successfully applied to some particular reliability growth i.e. observed during testing. Since the T&D environment cannot be generalized, a particular SRGM cannot be applied for any situation. This problem was solved by the unification approach (Shantikumar, 1981; Shantikumar, 1983; Langberg and Singpurwalla, 1985; Miller, 1986; Thompson, 1988; Gokhale et al., 1996; Chen and Singpurwalla, 1997; Gokhale and Trivedi, 1999; Huang et al., 2003). The SRGMs proposed earlier considered only some factors involved in testing but none of them was able to describe a testing environment in general. Therefore, several models were studied for reliability analysis in some particular application and then the best ones were selected. With the unification methodology, several different SRGMs are obtained without the need of making different assumptions. This makes the selection of SRGMs much easier than the earlier methods. Unified approach has also been used by many researchers to make reliability evaluation of the software robust.

Shantikumar (1981) and Shantikumar (1983) initially started the research in unified modelling and proposed a unified birth process model. This model described Jelinski and Moranda (1972) and GO model (1979). Langberg and Singpurwalla (1985) used the Bayesian approach to show that several Software Reliability Models (SRMs) can be comprehensively described. Langberg and Singpurwalla's idea was extended to a wide range of SRGMs by Miller (1986) and Thompson Jr. (1988). The framework of generalized order statistics (GOS) and record value was used by them. All SRMs and NHPP based models proposed in the past can be generalized by selfexciting point processes (Chen and Singpurwalla, 1997). A unified approach for discrete SRGMs using the concept of means was studied by Huang et al. (2003). In order to model the fault detection and correction process, Xie et al. (2007) proposed a generalized scheme. Two unified approaches were proposed by Kapur et al. (2010, 2011a) viz. infinite server queuing theory approach and fault-detection time distribution approach. With the unified modelling approaches (Xie et al., 2007; Kapur et al., 2010; Kapur et al., 2011b) several existing SRGMs that have been developed with diverse testing and debugging environment can be derived.

Unified approach for OSS

A vast literature is available for fault prediction modelling in OSS based on varied set of assumptions. Najeeb ullah and Morisio (2012) discussed the reliability growth for both OSS and CSS and carried out comparison between the two. In their study, it was analysed that Inflection S-shaped model and Gompertz model were best suited for OSS and Musa Okumoto and Inflection S-shaped models performed well for CSS. Further, the reliability growth of OSS was analysed by Najeeb ullah et al. (2012) by making use of eight SRGMs. Singh et al. (2010a) incorporated a change point concept in OSS for modelling reliability growth. In addition to this, various SRGMs were presented by Singh et al. (2010b) in order to predict the failure pattern in OSS.

For this stochastic differential equations were used. OSS reliability models for distributed environment was proposed by Johri et al. (2016) and was based on component specific testing effort. Many researchers (Tamura and Yamada, 2007; Yang et al., 2016; Aggarwal et al., 2017) have put forth the idea of multi-release modelling in OSS. Based on the maintenance effort of the software system an optimal upgrade time for OSS was found by Tamura and Yamada (2007). Multi-release modelling for OSS under the assumptions of perfect debugging was studied by Garmabaki et al. (2015). A multi-release framework for OSS modelling with separate fault detection and correction process was given by Yang et al. (2016). Aggarwal et al. (2017) studied the multi-release OSS reliability by integrating the factors of change point and FRF. They used exponentiated weibull FRF in order to model the real behaviour of the testing environment. Thus, the research work done till now on OSS focusses on formulating SRGMs applicable to a particular reliability growth. Due to the diverse nature of T&D environment, a particular SRGM cannot be applied in general.

Thus, there is a dire need of developing a unified framework for predicting faults in OSS based applications. In this thesis, a unified framework for OSS incorporating imperfect debugging, error generation and change point has been formulated. Further, eight different distribution functions have been used to obtain several different SRGMs from this framework. Thus, with the help of unified framework, several different SRGMs can be obtained without the need of making distinct assumptions making the model selection process much simpler.

Unified approach for CSS

A lot of research has been carried out in CSS and many unified frameworks have been proposed each serving different purpose. The studies conducted by researchers incorporating the concept of FRF were as follows: Hamilton and Musa (1978), Musa (1975) studied that the behaviour of FRF for different projects remains stable. Further, Hsu et al. (2011) studied the effect of time-dependent FRF on the reliability of the software. In their study, the behaviour of constant, increasing and decreasing FRF was discussed. A SRGM with change point, constant FRF under the assumptions of imperfect debugging was proposed by Jain et al. (2014a).

In addition, the effect of multiple change points incorporating weibull TEF, constant FRF under imperfect debugging was further studied by Jain et al. (2014b). A single release SRGM with inflection S-shaped FRF was proposed by Pachauri et al. (2015). This SRGM was studied under both perfect and imperfect debugging environment. The reason behind using inflection S-shaped curve was to address the changing behaviour of FRF. Moreover, they also studied a multi-release model incorporating constant FRF taking the debugging environment as imperfect. A multi-release unified model with time-variant FRF and TEF was proposed by Kapur et al. (2016) under perfect debugging environment.

Further, Aggarwal et al. (2017) developed a SRGM with exponentiated FRF as it is considered quite successful for describing diverse environmental conditions. The literature survey shows that many researchers have proposed SRGMs taking FRF and other environmental factors that have an impact on it into account. No research has been carried where the factors of testing effort, change point and error generation have been incorporated all together with FRF and a unified framework has been developed. In this study, we have developed a unified SRGM for CSS incorporating the concept of FRF after going through intense literature survey. With this unified framework, it is possible to derive different SRGMs taking different distribution functions into account. For example, it is possible to derive SRGMs that contains only a combination of testing effort and change point by setting the value of error generation to 0.

Conclusion

the need of developing unified frameworks and its significance has been discussed. The unified approach is the emerging research area in the field of NHPPbased software reliability modelling. Further, the literature regarding various proposed SRGMs has been discussed and the research gap wherein the need for developing a unified framework arises has been identified.

References:

Ahmad, N., Khan, M.G.M., & Rafi, L.S. (2011). ‘Analysis of an inflection S-shaped software reliability model considering log-logistic testing-effort and imperfect debugging’, International Journal of Computer Science and Network Security, vol. 11, no. 1, pp. 161–171.

Bittanti, S., Bolzern, P., Pedrotti, E., Pozzi, M., & Scattolini, R. (1988). ‘A flexible modeling approach for software reliability growth’, In: Goos G, Harmanis J (eds) Software reliability modelling and identification, Springer, Berlin, pp. 101–140.

Bokhari, M. U., & Ahmad, N. (2006). ‘Analysis of a Software Reliability Growth Models: the Case of Log-logistic Test-effort Function’, in: Proceedings of the 17th IASTED International Conference on Modeling and Simulation, Montreal, Canada, pp. 540-545.

Chang, Y. (2001). ‘Estimation of parameters for non-homogeneous Poisson process: software reliability with change point model’, Commun Stat Simulation Comput, vol. 30, no. 3, pp. 623–635.

Chen, Y., & Singpurwalla, N. D. (1997). ‘Unification of software reliability models by self-exciting point processes’, Advances in Applied Probability, vol. 29, issue. 2, pp. 337–352.

Hwang, S., & Pham, H. (2009). ‘Quasi-renewal time-delay fault-removal consideration in software reliability modelling’, IEEE Trans. Systems, Man and CyberneticsPart A: Systems and Humans, vol. 39, no. 1, pp. 200 – 209.

Kapur, P. K., Mishra, P., Shrivastava, A.K., & Khatri, S. K. (2016). ‘Multi-release modelling of a software with testing effort and Fault Reduction Factor’, International conference on Innovation and Challenges in Cyber Security (ICICCS).