International Journal of Reliability and Applications
Vol. 4, No. 3, pp. 131-139, 2003

Modelling A Relationship Between Reliability and Software Coverage

Joong-Yang Park * and **Young Soon Kim**
Department of Statistics and Information
Gyeongsang National University, Jinju, KOREA

Jae Heung Park
Department of Computer Science
Gyeongsang National University, Jinju, KOREA

Abstract. There is a new trend of incorporating software coverage metrics into software modelling. This paper proposes and empirically evaluates a software reliability growth model, which relates reliability to coverage. The proposed model is derived by modifying the assumptions on which Veevers and Marshall model is based.

Key Words : *coverage metric, software reliability, software reliability growth model*

1. INTRODUCTION

In recent years there is growing use of computer systems. Software is an integral part of most critical computer systems. Since failures of a software system can cause severe consequences in terms of human life, environment impact, or economical losses, an important quality attribute of a software system is the degree to which it can be relied upon to perform its intended function. One of quantifiable measures for this quality attribute is software reliability, which is formally defined as the probability that no failure occurs in a specified environment during a specified exposure period.

Software reliability has attracted considerable attention for the last 2 decades. Many software reliability models have been developed to estimate software reliability measures such as the initial fault content, the time to next failure, the number of remaining faults and the software reliability function. Software reliability growth models (SRGMs) are the software reliability models concerned with the relationship between the cumulative number of faults detected by software testing or the

*Corresponding author.
E-mail address: joongyang@nongae.gsnu.ac.kr

time interval between software failures and the time span of the software testing. Generally the time is measured by the execution time or the number of test cases executed. Refer to Goel (1985), Gokhale, Marinos and Trivedi (1996), Horgan et al. (1995), Lyu (1996), Malaiya and Srimani (1990), Musa et al. (1987), Ramamoorthy and Bastani (1982), Shantikumar (1983) and Wood (1997) for general discussion on software reliability models.

A software system can be regarded as a collection of constructs such as statements, blocks, decisions, *c*-uses and *p*-uses. A block is a portion of code that executes together, that is, consecutive code fragment without a branch, whereas a desicion refers to a branch. A def-use is a pair of statements at which a variable is defined and used, respectively. If the use appears in a computational expression, the def-use is termed as a *c*-use. If the use appears inside a predicate, it is termed as a *p*-use. A construct is said to be covered if this construct is executed. A coverage metric is defined as the fraction of constructs covered by a test. Popular coverage metrics are statement coverage, block coverage, decision coverage, *c*-use coverage, and *p*-use coverage. As mentioned in Varadan (1995) and Wood (1997), a new approach for modelling software reliability is to utilize test coverage measures. This approach is based on the following observations:

1. The more system is covered, the more likely reliable is the software system;
2. Reliability modelling can be improved by taking into account the structural coverage, even though structural testing may or may not be used.

The research results show a strong relation between coverage and software reliability. Chen, Mathur and Rego (1995) have investigated the correlation between coverage and software reliability using experimentation with randomly generated flow graphs. They present an empirical result that there is a need for making use of coverage measures in reliability estimation. They also show that in the absence of coverage measures one is likely to overestimate reliability. Frate et al. (1995) perform a similar investigation in empirical setting. Vouk (1992) investigates the relation between coverage and fault detection rate. Dalal, Horgan and Kettenring (1993) examine the correlation between coverage and the fault removal rate. They give a scatter plot of the number of faults detected during system testing versus the block coverage achieved during unit testing for 28 software systems. The plot shows that modules covered more thoroughly during unit testing are much less likely to contain faults. These studies have shown that there exists a strong relationship between reliability and coverage and provided a basis for the coverage-based SRGMs. Several coverage-based SRGMs have been developed by Chen et al. (1992), Chen, Lyu and Wong (1996, 1997), Gokhale et al. (1996a, 1996b), Grottke (1999, 2000), Malaiya et al.(1994, 1996), Piwowarski et al. (1993), and Veevers and Marshall (1994). Similarly to the traditional time domain SRGMs, most coverage-based SRGMs are usually characterized by their mean value functions relating the expected number of detected faults to coverage. The coverage-based SRGM appeared in Veevers (1990) and Veevers and Marshall (1994) is the only coverage-based SRGM directly relating reliability to coverage.

This paper proposes a new coverage-based SRGM which describes a relationship between reliability and coverage. The SRGM of Veevers (1990) and Veevers and Marshall (1994) is first reviewed in Section 2. Section 3 presents the new coverage-based SRGM and its application to two data sets are illustrated in Section 4. Finally concluding remarks are given in Section 5.

2. REVIEW OF VEEVERS AND MARSHALL SRGM

We represent the software under testing by a set Ω, whose elements are all constructs of the software. Consider $\mathcal{C}$ be a coverage metric defined on Ω such that $\mathcal{C}(\emptyset) = 0$, $\mathcal{C}(\Omega) = 1$ and $\mathcal{C}(\omega_i) < \mathcal{C}(\omega_j)$ where ω_i and ω_j are subsets of Ω and $\omega_i \subset \omega_j$. Constructs may be statements, blocks, branches, p-uses or c-uses depending on the coverage metric employed. Let c be the value of $\mathcal{C}$ at the current stage of testing and U be the set of uncovered constructs. Then $c = |\Omega - U|/|\Omega|$, where $|\cdot|$ is the cardinality of a set. Reliability of the software under testing when $\mathcal{C} = c$ is denoted by $r(c)$. If further testing covers a portion of U, there will be an increment dc in coverage. The coverage increment will produce an increment $dr(c)$ in reliability. Assuming that the total potential contribution of constructs in U to the reliability is $1 - r(c)$ and that all constructs in U are equally likely to be executed, Veevers and Marshall (1994) model reliability increment $dr(c)$ caused by dc to be proportional to $(1 - r(c))/|U|$. Thus they derived the following differential equation.

$$dr(c) = p\,(1 - r(c))\,dc, \tag{2.1}$$

where $0 < p$ is the proportional factor. Solving this differential equation with the initial condition $r(0) = r_{min}$, the relationship between r and c was obtained as

$$r(c) = 1 - (1 - r_{min})\exp(-pc). \tag{2.2}$$

3. A NEW COVERAGE-BASED SRGM

We reconsider the assumptions on which Veevers and Marshall model is based. Veevers and Marshall (1994) assumes that when the current coverage is c, the potential contribution of U to the reliability is $1 - r(c)$. This assumption implies that the coverage can reach 1 as testing proceeds and that reliability attains its theoretical maximum value 1 when the coverage is 1. However, it is generally accepted that 100% coverage is rarely achieved in practice because of, for example, the presence of infeasible blocks, p-uses and c-uses. Thus U can be further partitioned into two sets, UI and UF, which denote the set of uncovered infeasible constructs and the set of uncovered feasible constructs, respectively. It should be also noted that faults may be found in the covered constructs. That is, finding more faults does not necessarily mean that coverage increases. This and the subsume relationship among coverage metrics indicate that 100% coverage does not guarantee 100% reliability. Therefore it is reasonable to assume that there exist upper limits c_{max} and r_{max} for c and $r(c)$ for the given testing strategy. Here $c_{max} = |\Omega - UI|/|\Omega|$ and $r_{max} = r(c_{max})$.

Following the above discussion, the reliability increment $dr(c)$ caused by the coverage increment dc can be assumed to be proportional to $(r_{max} - r(c))/|UF|$. Since

$$\frac{r_{max} - r(c)}{|UF|} = \frac{r_{max} - r(c)}{|UF|/|\Omega|}\frac{1}{|\Omega|} = \frac{r_{max} - r(c)}{c_{max} - c}\frac{1}{|\Omega|}, \tag{3.1}$$

we obtain the following differential equation:

$$dr(c) = p\frac{r_{max} - r(c)}{c_{max} - c}dc, \tag{3.2}$$

Solving the differential equation (3.2) with the initial condition $r(0) = r_{min}$, the relationship between r and c is obtained as

$$r(c) = r_{max} - (r_{max} - r_{min})\left(1 - \frac{c}{c_{max}}\right)^p, \tag{3.3}$$

where $0 \leq c \leq c_{max} \leq 1$, $0 \leq r_{min} \leq r_{max} \leq 1$ and $0 < p$. However, $r(c)$ remains constant for $c_{max} \leq c \leq 1$, i.e., $r(c) = r_{max}$ for $c_{max} \leq c \leq 1$.

4. NUMERICAL EXAMPLES

We perform an empirical evaluation of the coverage-based SRGM proposed in the previous section. Two data sets are used for the empirical evaluation. First data set is the one that appeared in Veevers (1990). It includes failure rates of 13 faults and values of three coverage metrics at the instances of fault detection. The three coverage metrics are statement coverage, block coverage and LCSAJ (linear code sequence and jump) defined in Woodward, Hedley and Hennell (1980). Failure rates of the software system are computed by summing failure rates of the remaining faults at the instances of fault detection. Reliabilities of the software are estimated as $\exp(-\lambda t)$ for some specified value of t, where λ is the failure rate of the software. Values of block coverage metric and corresponding reliabilities for $t = 500$ are given in Table 1. Both Veevers and Marshall model and the newly proposed model are applied to this data set. The least squares estimates are obtained and tabulated in Table 2 and the fitted $r(c)$ curves are depicted in Figure 1. It can be verified that similar results are obtained for statement coverage and LCSAJ.

Table 1. Values of reliabilities and block coverage: Veevers data.

$r(c)$	c
0.5881	0.18
0.7568	0.31
0.8464	0.32
0.8643	0.33
0.9154	0.34
0.9349	0.36
0.9564	0.38
0.9719	0.44
0.9849	0.59
0.9919	0.60
0.9973	0.82

Table 2. Least squares estimates: Veevers data.

Parameter	Veevers and Marshall model	Proposed model
r_{min}	0.0000	0.0000
r_{max}	-	1.0000
c_{max}	-	0.8650
p	5.9699	4.1241

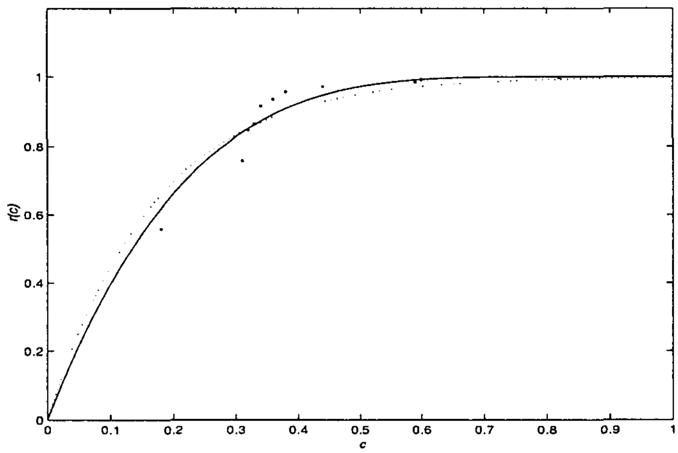

Figure 1. $r(c)$ curves fitted to Veevers data. (dotted line: Veevers and Marshall model, solid line: proposed model.)

Second data set was collected from a simulation experiment conducted by Dr. Mathur in Purdue University. The experiment tests a software system with 22 injected faults. 109 test cases were applied and 17 of the 22 injected faults were dis-

Table 3. Values of reliabilities and four coverage metrics: Mathur data.

	Coverage metric			
$r(c)$	Block	Decision	c-use	p-use
0.4033	0.00	0.00	0.00	0.00
0.5927	0.34	0.20	0.26	0.23
0.6807	0.42	0.28	0.34	0.30
0.7599	0.48	0.33	0.40	0.34
0.7693	0.59	0.44	0.51	0.43
0.7732	0.60	0.45	0.52	0.44
0.8918	0.60	0.46	0.52	0.45
0.9518	0.70	0.53	0.62	0.50
0.9784	0.73	0.60	0.66	0.56
0.9823	0.77	0.65	0.72	0.63
0.9958	0.77	0.65	0.72	0.63
1.0000	0.78	0.66	0.72	0.64

Table 4. Least squares estimates of the proposed model: Mathur data.

	Coverage metric			
Parameter	Block	Decision	c-use	p-use
r_{min}	0.3611	0.4004	0.3938	0.3837
r_{max}	1.0000	1.0000	1.0000	1.0000
c_{max}	0.7954	0.6600	0.7200	0.6400
p	1.0000	1.1323	1.0041	1.1167

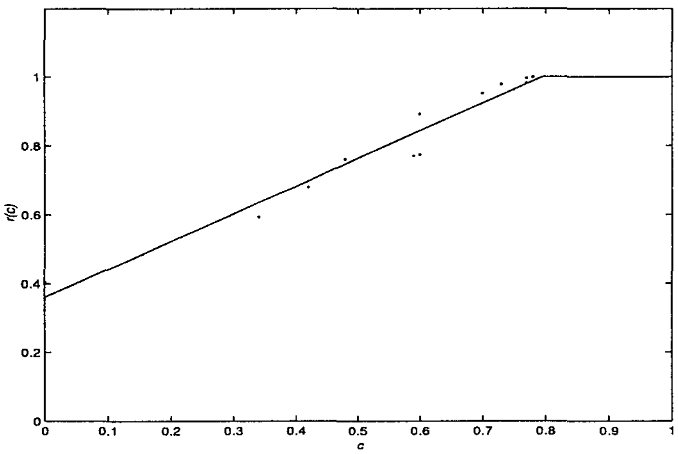

Figure 2. $r(c)$ curve fitted to Mathur data: block coverage.

discovered. Table 3 shows values of four coverage metrics and corresponding reliabilities. The least squares estimates of parameters are given in Table 4 and the estimated $r(c)$ curve for block coverage metric is plotted in Figure 2.

5. CONCLUDING REMARKS

We proposed a new coverage-based SRGM, which is based on more reasonable and practical assumptions than Veevers and Marshall model. Then its applicability was illustrated by means of numerical examples. Empirical evaluation indicates that the proposed model describes reliability growth behavior fairly well. However, the proposed model is inherently of limited use since it is not easy to collect reliability estimates during testing. Further research will be directed to this problem.

REFERENCES

M. H. Chen, J. R. Horgan, A. P. Mathur, and V. J. Rego (1992). A Time/Structure Based Model for Estimating Software Reliability, Technical Report SERC-TR-117-P, Purdue University.

M. H. Chen, M. R. Lyu and W. E. Wong (1996). An Empirical Study of the Correlation Between Code Coverage and Reliability Estimation, *Proceedings of the 3rd IEEE International Symposium on Software Metrics*, Berlin, Germany.

M. H. Chen, M. R. Lyu and W. E. Wong (1997). Incorporating Code Coverage in the Reliability Estimation for Fault-Tolerant Software, *Proceedings of the 16th IEEE Symposium on Reliable Distributed System*, 45-52, Durham, NC.

M. H. Chen, A. P. Mathur, and V. J. Rego (1995). Effect of Testing Techniques on Software Reliability Estimates Obtained Using A Time Domain Model, *IEEE Transactions on Reliability*, 44, 97-103.

S. R. Dalal, J. R. Horgan and J. R. Kettenring (1993). Reliable Software and Communication: Software Quality, Reliability, and Safety, *Proceedings of the 15th IEEE International Conference on Software Engineering*, 425-435, Baltimore, MD.

F. D. Frate, P. Garg, A. P. Mathur, and A. Pasquini (1995). On the Correlation Between Code Coverage and Software Reliability, *Proceedings of the 6th IEEE International Symposium on Software Reliability Engineering*, 124-132, Toulouse, France.

A. L. Goel (1985). Software Reliability Model: Assumptions, Limitations, and Applicability, *IEE Transactions on Software Engineering*, SE-11, 12, 1411-1423.

S. S. Gokhale, P. N. Marinos, and K. S. Trivedi (1996). Important Milestones in Software Reliability Modeling, *Communications in Reliability, Maintainability and Serviceability.*

S. S. Gokhale, T. Philip, P. N. Marinos, and K. S. Trivedi (1996a). Non-Homogeneous Markov Software Reliability Model with Imperfect Repair, Technical Report TR-96/12, CACC Duke University.

S. S. Gokhale, T. Philip, P. N. Marinos, and K. S. Trivedi (1996b). Unification of Finite Failure Non-Homogeneous Poisson Process Models Through Test Coverage, Technical Report TR-96/36, CACC Duke University.

M. Grottke (1999). Software Reliability Model Study, Technical Report IST-1999-55017, Chair of Statistics, University of Erlangen-Nuremberg, Germany.

M. Grottke (2000). A Vector Model for Structural Coverage Growth and the Number of Failure Occurrences, *Proceedings of the 13th International Symposium on Software Reliability Engineering.*

J. R. Horgan, A. P. Mathur, A. Pasquini, and V. J. Rego (1995). Perils of Software Reliability Modeling, Technical Report SERC-TR-160-p, Software Engineering Research Center, Purdue University.

M. R. Lyu (1996). *Handbook of Software Reliability Engineering* , McGraw-Hill.

Y. K. Malaiya, N. Li, J. Bieman, R. Karcich, and R. Skibbe (1994). The Relationship Between Test Coverage and Reliability, *Proceedings of the 5th International Symposium on Software Reliability Engineering*, 186-195, Monterey, CA.

Y. K. Malaiya, N. Li, J. Bieman, R. Karcich, and R. Skibbe (1996). Software Test Coverage and Reliability, Technical Report CS-96-128, Colorado State University.

Y. K. Malaiya and P. K. Srimani (1990). *Software Reliability Models: Theoretical Developments, Evaluation and Applications*, IEEE Computer Society Press, Los Alamos, California.

J. D. Musa, A. Iannino, and K. Okumoto (1987). *Software Reliability: Measurement, Prediction, Application*, McGraw-Hill.

P. Piwowarski, M. Ohba and J. Caruso (1993) Coverage Measurement Experience During Function Test, *Proceedings of the 15th International Conference on Software Engineering*, 287-300, Baltimore, MD.

C. V. Ramamoorthy and F. B. Bastani (1982). Software Reliability - Status and Perspective, *IEEE Tansactions on Software Engineering*, SE-8, 8, 354-371.

J. G. Shanthikumar (1983). Software Reliability Model: A Review, *Microelectronics and Reliability*, 23, 5, 903-943.

G. S. Varadan (1995). Trends in Reliability and Test Strategies, *IEEE Software*, 12, 3, 10.

A. Veevers (1990). Software Coverage Metrics and Operational Reliability, *Proceedings of the IFAC/EWICS/SARS Symposium*, Gatwick, UK, 17, 67-69.

A. Veevers and A. Marshall (1994). A Relationship Between Software Coverage Metrics and Reliability, *Software Testing, Verification and Reliability*, 4, 3-8.

M. A. Vouk (1992). Using Reliability Models During Testing with Non-Operational Profiles, *Proceedings of the 2nd Bellcore/Purdue Workshop on Issues in Software Reliability Estimation*, 103-111.

A. Wood (1997). Software Reliability Growth Models: Assumptions vs. Reality, *Proceedings of the 8th International Symposium on Software Reliability Engineering*, 136-141, Albuquerque, New Mexico.

M. R. Woodward, D. Hedley and M. A. Hennell (1980). Experience with Path Analysis and Testing of Programs, *IEEE Transactions on Software Engineering*, 6, 278-286.