Abstract—Software reliability is one of the most important characteristics of software quality. Its measurement and management technologies during the software life-cycle are essential to produce and maintain quality/reliable software systems. In this paper, we discuss software reliability modeling and its applications. As to software reliability modeling, hazard rate and NHPP models are investigated particularly for quantitative software reliability assessment. Further, imperfect debugging and software availability models are also discussed with reference to incorporating practical factors of dynamic software behavior.

Keywords — Product quality/reliability assessment, Software reliability growth modeling, Hazard rate, Nonhomogeneous Poisson process, Imperfect debugging, Software availability, Markov process.

I. INTRODUCTION

A software reliability model (SRM) is a mathematical analysis model for the purpose of measuring and assessing software quality/reliability quantitatively. Many software reliability models have been proposed and applied to practical use because software reliability is considered to be a “must-be quality” characteristic of a software product. The software reliability models can be divided into two classes as shown in Figure 1. One treats the upper software development process, i.e., design and coding phases, and analyzes the reliability factors of the software products and processes, which is categorized in the class of static model. The other deals with testing and operation phases by describing a software failure-occurrence phenomenon or software fault-detection phenomenon, by applying the stochastic/statistics theories and can estimate and predict the software reliability, which is categorized in dynamic model.

In the former class, a software complexity model is well known and can measure the reliability by assessing the complexity based on the structural characteristics of products and the process features to produce the products. In the latter class, a software reliability growth model is especially well known. Further, this model is divided into three categories:

(1) Software failure-occurrence time model

The model which is based on the software failure-occurrence time or the software fault-detection time.

(2) Software fault-detection count model

The model which is based on the number of software failure-occurrence or the number of detected faults.

(3) Software availability model

The model which describes the time-dependent behavior of software system alternating up (operation) and down (restoration or fault correction) states.

The software reliability growth models are utilized for assessing the degree of achievement of software quality, deciding the time to software release for operational use, and evaluating the maintenance cost for faults undetected during the testing phase. We discuss the software reliability growth models.

II. SOFTWARE RELIABILITY GROWTH MODELING

Generally, a mathematical model based on stochastic and statistical theories is useful to describe the software fault-detection phenomena or the software failure-occurrence phenomena and estimate the software reliability quantitatively. During the testing phase in the software development process, software faults are detected and removed with a lot of testing effort expenditures. Then, the number of faults remaining in the software system decreases as the testing goes on. This means that the probability of software failure-occurrence is decreasing, so that the software reliability is increasing and the time interval between software failures becoming longer with the testing time.

A mathematical tool which describes software reliability aspect is a software reliability growth model.

Based on the plausible definitions, we can develop a software reliability growth model based on the assumptions used for the actual environment during the testing phase or the operation phase. Then, we can define the following random variables on the number of detected faults and the software failure-occurrence time (see Figure 2):

\( t = \) the cumulative number of software faults (or the cumulative number of observed software failures) detected up to time \( t \),

\( S_i = \) the \( i \)-th software-failure occurrence time (\( i = 1, 2, \ldots \); \( S_0 = 0 \)),

The model which is based on the number of software failure-occurrence or the number of detected faults.

The model which describes the time-dependent behavior of software system alternating up (operation) and down (restoration or fault correction) states.
\( X_i \): the time interval between (i-1)-st and i-th software failures \((i = 1, 2, \cdots; X_i = 0)\)

Figure 2 shows the occurrence of event \( \{ N t = i \} \) since \( I \) faults have been detected up to time \( t \). From these definitions, we have

\[
S_i = \sum_{k=1}^{i} X_k, \quad X_i = S_i - S_{i-1}.
\]  

\( i \)

\( \lambda x = \phi (\phi > 0), \quad (3) \)

\( \lambda x = \phi x B67 \phi > 0, > 0 \)  

\( \therefore \) then two typical software hazard rate models, respectively called the Jelinski-Moranda model and the Wagoner model can be derived, where \( \phi \) and \( m \) are constant parameters. Usually, it is difficult to assume that a software system is completely fault free or failure free. Then, we have a software hazard rate model called the Moranda model for the case of the infinite number of software failure occurrences as

\[
z_i x = D k; 67 i = 1, 2, \cdots; D > 0; 0 < k < 1,
\]

\( D \) is the initial software hazard rate and \( k \) the decreasing ratio. Eq. (5) describes a software failure-occurrence phenomenon where a software system has high frequency of software failure occurrence during the early stage of the testing or the operation phase and it gradually decreases thereafter.

Based on the software hazard rate models above, we can derive several software reliability assessment measures. For example, the software reliability function for \( X_i (i = 1, 2, \cdots) \) is given as

\[
R_i(x) = \exp \left[ - \int_{0}^{x} z_i(x) dx \right] \quad (i = 1, 2, \cdots).
\]  

Further, we also discuss NHPP models, which are modeled for random variable \( N(t) \) as typical software reliability growth models. In the NHPP models, a nonhomogeneous Poisson process (NHPP) is assumed for the random variable \( N(t) \), the distribution function of which is given by

\[
\text{Pr}\{ \cdot \} \quad \text{and} \quad E[\cdot] \quad \text{mean the probability and expectation, respectively.} \quad H(t) \quad \text{in Eq. (7) is called a mean value function which indicates the expectation of} \quad N(t), \quad i.e., \quad \text{the expected cumulative number of faults detected (or the expected cumulative number of software failures occurred) in the time interval} \quad (0, t), \quad \text{and} \quad h(t) \quad \text{in Eq. (7) called an intensity function which indicates the instantaneous fault-detection rate at time} \quad t.
\]

\[
\text{From Eq. (7), various software reliability assessment measures can be derived. For examples, the expected number of faults remaining in the system at time} \quad t \quad \text{is given by}
\]

\[
n(t) = a - H(t), \quad (8)
\]

where \( a = H(x) \), i.e., parameter \( a \) denotes the expected initial fault content in the software system. Given that the testing or the operation has been going up to time \( t \), the probability that a software failure does not occur in the time interval } (t, t + x) \} x
\( R(x|t) = \exp[H(t) - H(x + t)] \quad (t \geq 0, x \geq 0) \) (9)

Table 1 A summary of NHPP models.

<table>
<thead>
<tr>
<th>NHPP model</th>
<th>Mean value function ( H(t) )</th>
<th>Intensity function ( h_m(t) )</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential software reliability growth model</td>
<td>( m(t) = a(1 - e^{-bt}) ) ((a &gt; 0, b &gt; 0))</td>
<td>( h_m(t) = a(1 - e^{-bt}) )</td>
<td>A software failure-occurrence phenomenon with a constant fault-detection rate at an arbitrary time is described.</td>
</tr>
<tr>
<td>Modified exponential software reliability growth model</td>
<td>( m_p(t) = a \sum_{i=1}^{2} p_i(1 - e^{-b_i t}) ) ((a &gt; 0, 0 &lt; b_2 &lt; b_1 &lt; 1, \sum_{i=1}^{2} p_i = 1, 0 &lt; p_1 &lt; 1))</td>
<td>( h_p(t) = a \sum_{i=1}^{2} p_i b_i e^{-b_i t} )</td>
<td>A difficulty of software fault-detection during the testing is considered. ( b_1 ) is the fault-detection rate for easily detectable faults; ( b_2 ) is the fault-detection rate for hardly detectable faults</td>
</tr>
<tr>
<td>Delayed S-shaped software reliability growth model</td>
<td>( M(t) = a[1 - (1 + bt)e^{-bt}] ) ((a &gt; 0, b &gt; 0))</td>
<td>( h_M(t) = ab^2te^{-bt} )</td>
<td>A software fault-detection process is described by two successive phenomena, i.e., failure-detection process and fault-isolation process.</td>
</tr>
<tr>
<td>Inflection S-shaped software reliability growth model</td>
<td>( I(t) = \frac{a(1 - e^{-bt})}{1 + e^{-bt}} ) ((a &gt; 0, b &gt; 0, c &gt; 0))</td>
<td>( h_I(t) = \frac{ab(1 + c)e^{-bt}}{(1 + e^{-bt})^2} )</td>
<td>A software failure-occurrence phenomenon with mutual dependency of detected faults is described.</td>
</tr>
<tr>
<td>Testing-effort-dependent software reliability growth model</td>
<td>( T(t) = a[1 - e^{-rW(t)}] ) ((a &gt; 0, 0 &lt; r &lt; 1, \alpha &gt; 0, \beta &gt; 0, m &gt; 0))</td>
<td>( h_T(t) = ar\alpha \beta mt^{m-1}e^{-rW(t)} )</td>
<td>The time-dependent behavior of the amount of testing effort and the cumulative number of detected faults is considered.</td>
</tr>
<tr>
<td>Testing-domain-dependent software reliability growth model</td>
<td>( W(t) = a(1 - e^{-\beta W(t)}) ) ((a &gt; 0, 0 &lt; r &lt; 1, \alpha &gt; 0, \beta &gt; 0, m &gt; 0))</td>
<td>( h_D(t) = \frac{a \alpha \beta}{e-b} ) ((e^{-bt} - e^{-vt}))</td>
<td>The testing domain, which is the set of software functions influenced by executed test cases, is considered.</td>
</tr>
<tr>
<td>Logarithmic Poisson execution time model</td>
<td>( \mu(t) = \frac{1}{2} \ln(\lambda_0 \theta t + 1) ) ((\lambda_0 &gt; 0, \theta &gt; 0))</td>
<td>( \lambda(t) = \frac{\lambda_0}{\lambda_0 \theta t + 1} )</td>
<td>When the testing or operation time is measured on the basis of the number of CPU hours, and exponentially decreasing software failure rate is considered with respect to the cumulative number of software failures.</td>
</tr>
</tbody>
</table>

\( a = \) the expected number of initial fault content in the software system.
\( b, h_i, r = \) the parameter representing the fault-detection rate.
\( c = \) the parameter representing the inflection factor of test personnel.
\( p_i = \) the fault content ratio of Type \( i \) fault \((i = 1, 2)\).  
\( \alpha, \beta, m = \) the parameters which determine the testing-effort function \( W(t) \).
\( v = \) the testing-domain growth rate.
\( \lambda_0 = \) the initial software failure rate.
\( \theta = \) the reduction rate of software failure rate.
In Eq. (9) is a so-called software reliability. Measures of \( MTBF \) (mean time between software failures or fault detections) can be obtained follows:

\[
MTBF_f(t) = \frac{1}{h(t)}, \quad (10)
\]
\[
MTBF_c(t) = \frac{t}{H(t)}. \quad (11)
\]

Faults detected during the testing and operation phases are corrected and removed perfectly. However, debugging actions in real testing and operation environments are not always performed perfectly. For example, typing errors invalidate the fault-correction activity or fault-removal is not carried out precisely due to incorrect analysis of test results. We therefore have an interest in developing a software reliability growth model which assumes an imperfect debugging environment.

Such an imperfect debugging model is expected to estimate reliability assessment measures more accurately.

**A. Imperfect debugging model with perfect correction rate**

To model an imperfect debugging environment, the following assumptions are made:

(1) Each fault which causes a software failures is corrected perfectly with probability \( p \) (0 ≤ \( p \) ≤ 1). It is not corrected with probability \( q = 1-p \). We call \( p \) the perfect debugging rate or the perfect correction rate.

(2) The hazard rate is given by Eq. (5) and decrease geometrically each time a detected fault is corrected (see Figure 3).

(3) The probability that two or more software failures occur simultaneously is negligible.

(4) No new faults are introduced during the debugging. At most one fault is removed when it is corrected, and the correction time is not considered.

Let \( X(t) \) be a random variable representing the cumulative number of faults corrected up to the testing time \( t \). Then, \( X(t) \) forms a Markov process. That is, from assumption (1), when \( I \) faults have been corrected by arbitrary testing time \( t \),

\[
L(t) = \frac{k}{1 + me^{-\alpha t}} \quad (m > 0, \alpha > 0, k > 0), \quad (12)
\]
\[
G(t) = ka(b^t) \quad (0 < a < 1, 0 < b < 1, k > 0). \quad (13)
\]

In Eqs. (12) and (13), assuming that a convergence value of each curve (\( L \propto \) or \( G(\alpha) \)), i.e., parameter \( k \), represents the initial fault content in the software system, it can be estimated by a regression analysis.

**IMPERFECT DEBUGGING MODELING**

Most software reliability growth models proposed so far are based on the assumption of perfect debugging, i.e., that all
Fig. 4 A diagrammatic representation of transitions between states of $X(t)$

$$X(t) = \begin{cases} 
  i, & \text{with probability } q, \\
  i + 1, & \text{with probability } p,
\end{cases} \quad (14)$$

(see Fig. 4). Then, the one-step transition probability for the Markov process that after making a transition into state $i$, the process $\{X(t), t \geq 0\}$ makes a transition into state $j$ by time $t$ is given by

$$Q_{ij}(t) = p_{ij}(1 - \exp[-DK^i t]), \quad (15)$$

where $p_{ij}$ are the transition probabilities from state $i$ to state $j$ and are given by

$$p_{ij} = \begin{cases} 
  q & (i = j) \\
  p & (j = i + 1) \\
  0 & (\text{elsewhere}).
\end{cases} \quad (16)$$

Eq. (15) represents the probability that if $i$ faults have been corrected at time zero, $j$ faults are corrected by time $t$ after the next software failure occurs. Therefore, based on Markov analysis by using the assumptions and stochastic quantities above, we have the software reliability function and the mean time between software failures for $X_i (i = 1, 2, \ldots)$ as

$$R_i(x) = \sum_{s=0}^{i-1} \binom{i-1}{s} p^s q^{i-1-s} \exp[-DK^i x], \quad (17)$$

$$E[X_i] = \int_0^\infty R_i(x) dx = \frac{(p_k + q)^{i-1}}{D}. \quad (18)$$

And if the initial fault content in the system, $N$, is specified the expected cumulative number of faults debugged imperfectly up to time $t$ is given by

$$M(t) = \frac{q}{p} \sum_{n=1}^{N} \sum_{i=0}^{n-1} A_{i,n}(1 - \exp[-pDK^i t]), \quad (19)$$

where $A_{i,n}$ is

$$A_{i,n} = \begin{cases} 
  A_{0,1} = 1 \\
  A_{i,n} = \frac{k \binom{n}{i} \sum_{j=0}^{n-1} (k^j - k^i)}{\prod_{j=0}^{n-1}(k^j - k^i)} \\
  (n = 2, 3, \ldots; i = 0, 1, 2, \ldots, n - 1).
\end{cases} \quad (20)$$

B. Imperfect debugging model for introduced faults

Besides the imperfect debugging factor above in faultcorrection activities, we consider the possibility of introducing new faults in the debugging process. It is assumed that the following two kinds of software failures exist in the dynamic environment, i.e., the testing or user operation phase:

(F1) software failures caused by faults originally latent in the software system prior to the testing (which are called inherent faults),

(F2) software failures caused by faults introduced during the software operation owing to imperfect debugging.

In addition, it is assumed that one software failure is caused by one fault and that it is impossible to discriminate whether the fault that caused the software failure that has occurred is F1 or F2. As to the software failure-occurrence rate due to F1, the inherent faults are detected with the progress of the operation time. In order to consider two kinds of time dependencies on the decreases of F1, let $a_i(t) (i = 1, 2)$ denote the software failure-occurrence rate for F1. On the other hand, the software failure-occurrence rate due to F2 is denoted as constant $\lambda (\lambda > 0)$, since we assume that F2 occurs randomly throughout the operation. When we consider the software failure-occurrence rate at operation time $t$ is given by

$$h_i(t) = \lambda + a_i(t) \quad (i = 1, 2). \quad (21)$$

From Eq. (21), the expected cumulative number of software failures in the time interval $(0, t]$ (or the expected cumulative number of detected faults) is given by
SOFTWARE AVAILABILITY MODELING

Recently, software performance measures such as the possible utilization factors have begun to be interesting for metrics as well as the hardware products. That is, it is very important to measure and assess software availability, which is defined as the probability that the software system is performing successfully, according to the specification, at a specified time point. Several stochastic models have been proposed so far for software availability measurement and assessment. One group has proposed a software availability model considering a reliability growth process, taking account of the cumulative number of corrected faults. Others have constructed software availability models describing the uncertainty of fault removal. Still others have incorporated the increasing difficulty of fault removal.

The actual operational environment needs to be more clearly reflected in software availability modeling, since software availability is a customer-oriented metrics. In existing models the development of a plausible model is described, which assumes that there exist two types of software failure occurring during the operation phase. Furthermore, an operational software availability model is built up from the viewpoint of restoration scenarios.

The above models have employed Markov processes for describing the stochastic time-dependent behaviors of the systems which alternate between the up state, operating regularly, and the restoration state (down state) when a system is inoperable. Several stochastic metrics for software availability measurement in dynamic environment are derived from the respective models.

We discuss a fundamental software availability model below.

A. Model description

The following assumptions are made for software availability modeling:

(1) The software system is unavailable and starts to be restored as soon as a software failure occurs, and the system cannot operate until the restoration action is complete (see Figure 5).

(2) The restoration action implies debugging activity, which is performed perfectly with probability \( a \) (0 < \( a \) ≤ 1) and imperfectly with probability \( b \) (=1-\( a \)). We call \( a \) the perfect debugging rate. One fault is corrected and removed from the software system when the debugging activity is perfect.
(3) When \( n \) faults have been corrected, the time to the next software failure occurrence and the restoration time follow exponential distributions with means of \( 1/\lambda V \) and \( 1/\mu V \), respectively.

Fig. 8. A state transition diagram for software availability modeling.

(4) The probability that two or more software failures will occur simultaneously is negligible.

Consider a stochastic process \( \{ X_t, t \geq 0 \} \) with the state space \( (W, R) \) where up state vector \( W = [W_n; n = 0,1,2,\cdots] \) and down state vector \( R = [RV; n = 0,1,2,\cdots] \). Then, the events \( \{ X_t = WV \} \) and \( \{ X_t = RV \} \) mean that the system is operating and inoperable, respectively, due to the restoration action at time \( t \), when \( n \) faults have already been corrected.

From assumption (2), when the restoration action has been completed in \( \{ X_t = RV \} \),

\[
X(t) = \begin{cases} W_n & \text{with probability } b_n, \\ W_{n+1} & \text{with probability } a_n, \end{cases}
\]

We use the Moranda model in Eq. (5) to describe the software failure-occurrence phenomenon, i.e., when \( n \) faults have been corrected, the software hazard rate \( \lambda V \) (see Figure 3) is given by

\[
\lambda_n = \lambda (n = 0,1,2,\cdots; D > 0, 0 < k < 1). \tag{24}
\]

The expression of Eq. (24) comes from the point of view that software reliability depends on the debugging efforts, not the residual fault content. We do not note how many faults remain in the software system.

Next, we describe the time-dependent behavior of the restoration action. The restoration action for software systems includes not only the data recovery and the program reload, but also the debugging activities for manifested faults. From the viewpoint of the complexity, there are cases where the faults detected during the early stage of the testing or operation phase have low complexity and are easy to correct/remove, and as the testing is in progress, detected faults have higher complexity and are more difficult to correct/remove. In the above case, it is appropriate that the mean restoration time becomes longer with the increasing number of corrected faults. Accordingly, we express \( \mu_n \) as follows (see Figure 6).

\[
\mu_n = E r^n (n = 0,1,2,\cdots; E > 0, 0 < r \leq 1), \tag{25}
\]

where \( E \) and \( r \) are the initial restoration rate and the decreasing ratio of the restoration rate, respectively. In Eq. (25) the case of \( r=1 \), i.e., \( \mu V = E \), means that the complexity of each fault is random.

Let \( T_n \) and \( U_n (n = 0,1,2,\cdots) \) be the random variables representing the next software failure occurrence and the next restoration time intervals when \( n \) faults have been corrected, in other words the sojourn times in states \( W_n \) and \( R_n \), respectively. Furthermore, let \( Y(t) \) be the random variable representing the cumulative number of faults corrected up to time \( t \). The sample behavior of \( Y(t) \) is illustrated in Figure 7. It is noted that the cumulative number of corrected faults is not always coincident with that of software failures or restoration actions. The sample state transition diagram of \( X(t) \) is illustrated in Figure 8.

**B. Software availability measures**

We can obtain the state occupancy probabilities that the system is in state \( W_n \) and \( R_n \) at time point \( t \) as

\[
P_{W_n}(t) = \Pr\{ X(t) = W_n \} = \frac{g_{n+1}(t)}{a\lambda_n} + \frac{g'_{n+1}(t)}{a\lambda_n\mu_n (n = 0,1,2,\cdots)}, \tag{26}
\]

\[
P_{R_n}(t) = \Pr\{ X(t) = R_n \} = \frac{g_{n+1}(t)}{a\mu_n} (n = 0,1,2,\cdots), \tag{27}
\]

respectively, where \( g_n(t) \) is the probability density function of random variable \( S_n \), which denotes the first passage time to state \( W_n \), and \( g'_n(t) \equiv dg_n(t)/dt \). \( g_n(t) \) and \( g'_n(t) \) can be given analytically.
Fig. 9 Dependence of perfect debugging rate $a$ on $(t)$.

\[ \sum_{n=0}^{\infty} [P_{W_n}(t) + P_{R_n}(t)] = 1. \]  
(28)

The instantaneous availability is defined as

\[ A(t) \equiv \sum_{n=0}^{\infty} P_{W_n}(t), \]  
(29)

which represents the probability that the software system is operating at specified time point $t$. Furthermore, the average software availability over $(0, t]$ is defined as

\[ A_{av}(t) = \frac{1}{t} \int_0^t A(x) \, dx, \]  
(30)

which represents the ratio of system’s operating time to the time interval $(0, t]$. Using Eqs. (26) and (27), we can express Eqs. (29) and (30) as

\[ A(t) = \sum_{n=0}^{\infty} \left[ \frac{g_{n+1}(t)}{a\lambda_n} + \frac{g'_{n+1}(t)}{a\lambda_n\mu_n} \right] \]  
\[ = 1 - \sum_{n=0}^{\infty} \frac{g_{n+1}(t)}{a\mu_n}, \]  
(31)

Fig. 10. Dependence of perfect debugging rate $a$ on $A_{av}(t)$.

\[ A_{av}(t) = \frac{1}{t} \sum_{n=0}^{\infty} \left[ \frac{G_{n+1}(t)}{a\lambda_n} + \frac{g_{n+1}(t)}{a\lambda_n\mu_n} \right] \]  
\[ = 1 - \frac{1}{t} \sum_{n=0}^{\infty} \frac{G_{n+1}(t)}{a\mu_n}, \]  
(32)
respectively, where $G_n(t)$ is the distribution function of $S_n$.

Figures 9 and 10 show numerical illustrations of $t_1$ and $A_{av}(t)$ in Eqs. (31) and (32), respectively.

**ACKNOWLEDGMENT**

This research was partially supported by the JSPS KAKENHI (C), Grant No. 16K01242.

**REFERENCES**


