Genetic Algorithm Based Optimal Testing Effort Allocation Problem for Modular Software

Anu G. Aggarwal¹, P. K. Kapur², Gurjeet Kaur³ and Ravi Kumar⁴

Abstract - Software reliability growth models (SRGM) are used to assess modular software quantitatively and predict the reliability of each of the modules during module testing phase. In the last few decades various SRGM’s have been proposed in literature. However, it is difficult to select the best model from a plethora of models available. To reduce this difficulty, unified modeling approaches have been proposed by many researchers. In this paper we present a generalized framework for software reliability growth modeling with respect to testing effort expenditure and incorporate the faults of different severity. We have used different standard probability distribution functions for representing failure observation and fault detection/correction times. The faults in the software are labeled as simple, hard and complex faults. Developing reliable modular software is necessary. But, at the same time the testing effort available during the testing time is limited. Consequently, it is important for the project manager to allocate these limited resources among the modules optimally during the testing process. In this paper we have formulated an optimization problem in which the total number of faults removed from modular software is (which include simple, hard and complex faults) maximized subject to budgetary and reliability constraints. To solve the optimization problem we have used genetic algorithm. One numerical example has been discussed to illustrate the solution of the formulated optimal effort allocation problem.

Index Terms - Non-homogenous Poisson process, software reliability growth model, Probability Distribution Functions, Fault Severity, Genetic Algorithm

1. INTRODUCTION

Nowadays large and complex software systems are developed by integrating a number of small and independent modules. Modules can be visualized as independent softwares performing predefined tasks, mostly developed by separate teams of programmers and sometimes at different geographical locations. During the development of modular software, faults can crop in the modules due to human imperfection. These faults manifest themselves in terms of failures when the modules are tested independently during the module testing phase of software development life cycle. However, in today’s computer invaded world these failures can lead to big losses in terms of money, time and life. Thus it is very important to evaluate software reliability of each module during modular testing phase.

To assess modular software quantitatively and predict the reliability of each of the modules during module testing, software reliability growth models (SRGM) are used. Numerous SRGM’s, which relate the number of failures (fault identified) and the Execution time (CPU time/Calendar time) have been discussed in the literature [19,5,3]. All these SRGMs assume that the faults in the software are of the same type. However, this assumption is not truly representative of reality. The software includes different types of faults, and each fault requires different strategies and different amounts of testing effort for removal. Ohba [8] refined the Goel-Okumoto[1] model by assuming that the fault detection/removal rate increases with time and that there are two types of faults in the software. SRGM proposed by Bittanti et al. [22] and Kapur and Garg [13] has similar forms as that of Ohba [8] but they developed under different set of assumptions. These models can describe both exponential and S-shaped growth curves and therefore are termed as flexible models [22, 8, 13]. Kapur et al. [16] developed Flexible software reliability growth model with testing effort dependent learning process in which two types of software faults were taken. Further, they proposed an SRGM with three types of faults [19]. The first type of fault was modeled by an Exponential model of Goel and Okumoto [1]. The second type was modeled by Delayed S-shaped model of Yamada et al. [21]. The third type was modeled by a three-stage Erlang model proposed by Kapur et al. [19]. The total removal phenomenon was modeled by the superposition of the three SRGMs. Shatnawi and Kapur [11] later proposed a generalized model based on classification of the faults in the software system according to their removal complexity.

The above literature review reveals that in the last few decades several SRGM’s have been proposed. This plethora of SRGM’s makes the model selection a tedious task. To reduce this difficulty, unified modeling approaches have been proposed by many researchers. The work in this area started as early as in 1980s with Shantikumar [4] proposing a Generalized birth process model. Gokhale and Trivedi [23] used Testing coverage function to present a unified framework and showed how NHPP based models can be represented by probability...
distribution functions of fault-detection times. Another
unification methodology is based on a systematic study of Fault
detection process (FDP) and Fault correction process (FCP)
where FCPs are described by detection process with time delay.
The idea of modeling FCP as a separate process following the
FDP was first used by Schneidewind [10]. More general
treatment of this concept is due to Xie et al [9] who suggested
modeling of Fault detection process as a NHPP based SRGM
followed by Fault correction process as a delayed detection
process with random time lag. The unification scheme due to
Kapur et al [17] is based on Cumulative Distribution Function
for the detection/correction times and incorporates the concept
of change point in Fault detection rate. These schemes have
proved to be fruitful in obtaining several existing SRGM by
following single methodology and thus present a perceptive
investigation for the study of general models without making
many assumptions. In this paper we made use of such unified
scheme for presenting a generalized framework for software
reliability growth modeling with respect to testing effort
expenditure and incorporate the faults of different severity. We
have used different standard probability distribution functions
for representing failure observation and fault correction times
Also, the total number of faults in the software are labeled as
simple, hard and complex faults. It is assumed that the testing
phase consists of three different processes, namely failure
observation, fault isolation and fault removal. The time delay
between the failure observation and subsequent removal is
assumed to represent the severity of the fault.
Developing reliable modular software is necessary. But, at
the same time the testing effort available during the testing time
is limited. These testing efforts include resources like human
power, CPU hours, and elapsed time, etc. Hence, to develop a
good reliable software system, a project manager must
determine in advance how to effectively allocate these
resources among the various modules. Such optimization
problems are called “Resource Allocation problems”. Many
authors have investigated the problem of resource allocation [2,
7]. Kapur et al [20, 15] studied various resource allocation
problems maximizing the number of faults removed form each
module under constraint on budget and management
aspirations on reliability for exponential and S-shaped SRGMs
[1,19,8]. In this paper we have formulated an optimization
problem in which the total number of faults removed from
modular software is (which include simple, hard and complex
faults) maximized subject to budgetary and reliability
constraints.
To solve the effort allocation problem formulated in this
research paper we use Genetic Algorithm(GA). GA stands up a
powerful tool for solving search & optimization problems. The
complex non linear formulation of the optimal effort allocation
problem is the reason behind choosing genetic algorithm as the
solving tool. GA always considers a population of solutions.
There is no particular requirement on the problem before using
GA’s, as it can be applied to solve any kind of problem.
The paper is organized as follows. Section 2 gives the
generalized framework for developing the software reliability
growth model for faults of different severity. In section 3
parameter estimation and model validation of the proposed
model is done through SPSS. The testing effort allocation
problem is formulated in section 4. In section 5 genetic
algorithm is presented for solving the discussed problem.
Section 6 illustrates the optimization problem solution through
a numerical example. Finally, conclusions are drawn and are
given in section 7.
2.1 Notations
W(t) : Cumulative testing effort in the interval (0,t).
w(t) : Current testing-effort expenditure rate at testing time t.
\frac{d}{dt}W(t) = w(t)
m(W_i) : Expected number of faults removed of type j(j=simple,
Hard, Complex Faults).
m(W_i) : Expected number of total faults removed.
b \quad \text{Constant fault detection rate}.
\beta \quad \text{rate of consumption of testing-effort}
\lambda (W_i) : Intensity function for Fault correction process (FCP)
or Fault correction rate per unit time.
G(W_i), F(W_i), H(W_i) : Testing effort dependent
Probability Distribution Function for Failure observation, Fault
Detection and Fault Correction Times
\begin{align*}
g(W_i), f(W_i), h(W_i) : & \text{Testing effort dependent Probability} \\
\ast & \text{Convolution.} \\
\circ & \text{Steiltjes convolution.}
\end{align*}
2.2 Basic Assumptions
The proposed model is based upon the following basic
assumptions:
1. Failure occurrence, fault detection, or fault removal
phenomenon follows NHPP.
2. Software is subject to failures during execution caused by
faults remaining in the software.
3. The faults existing in the software are of three types: simple,
hard and complex. They are distinguished by the amount of
testing effort needed to remove them
4. Fault removal process is prefect and failure observation/fault
isolation/ fault removal rate is constant.
5. Each time a failure occurs, an immediate effort takes place to
decide the cause of the failure in order to remove it. The time
delay between the failure observation and its subsequent
fault removal is assumed to represent the severity of the
faults. The more severe the fault, more the time delay.
6. The fault isolation/removal rate with respect to testing effort intensity is proportional to the number of observed failures.

2.3 Modeling Testing Effort
The proposed SRGM in this paper takes into account the time dependent variation in testing effort. The testing effort (resources) that govern the pace of testing for almost all the software projects are Manpower and Computer time. To describe the behavior of testing effort, Exponential, Rayleigh, or Weibull function has been used. The testing-effort described by a Weibull-type distribution is given by:

\[ W(t) = \alpha \left[ 1 - \exp\left( - \frac{\beta}{2} \cdot t^2 \right) \right] \]  

In equation (1), if \( g(t) = \beta \cdot t \). Then, there is a Rayleigh curve and the cumulative testing-effort is given by:

\[ W(t) = \alpha \left[ 1 - \exp\left( - \frac{\beta}{2} \cdot t^2 \right) \right] \]  

And if \( g(t) = \gamma \cdot \beta \cdot t^{\gamma - 1} \) in (1), then

\[ W(t) = \alpha \left[ 1 - \exp\left( - \beta \cdot t^\gamma \right) \right] \]  

which is cumulative testing effort of Weibull curve.

2.4 Model Development
Let \( a_1, a_2, \) and \( a_3 \) be the simple, hard and complex faults respectively at the beginning of testing. Also ‘a’ is the total fault content i.e. \( a = a_1, a_2, a_3 \).

2.4.1 Modeling Simple Faults
Simple faults are the faults which can be removed instantly as soon as they are observed. The mean value function for the simple faults of the software reliability growth model with respect to testing effort expenditure can be written as [18]:

\[ m_1(W_t) = a_1 \cdot F(W_t) \]  

where, \( F(W_t) \) is testing effort dependent distribution function.

From Equation (5), the instantaneous failure intensity function \( \lambda(W_t) \) is given by:

\[ \lambda(W_t) = a_1 \cdot F'(W_t) \]  

Or we can write

\[ \lambda(W_t) = \frac{dm}{dW_t} = \left[ a_1 - m(W_t) \right] \frac{F'(W_t)}{1 - F(W_t)} \]  

2.4.2 Modeling Hard Faults
The hard faults consume more testing time for the removal. This means that the testing team will have to spend more time to analyze the cause of the failure and therefore requires greater time to remove them. Hence the removal process for hard faults is modeled as a two-stage process and is given by[18]:

\[ m_2(W_t) = a_2 \left( F \otimes G \right)(W_t) \]  

\[ \lambda(W_t) = \frac{\left( f \ast g \right)(W_t)}{1 - (F \otimes G)(W_t)} \left[ a_2 - m(W_t) \right] \]  

2.4.3 Modeling Complex Faults
These faults require more testing time for removal after isolation as compared to hard fault removal. Hence they need to be modeled with greater time lag between failure observation and removal. Thus, the removal process for complex faults is modeled as a three-stage process:

\[ m_3(W_t) = a_3 \left( F \otimes G \otimes H \right)(W_t) \]  

And the instantaneous failure intensity function \( \lambda(W_t) \) is:

\[ \lambda(W_t) = \frac{\left( f \ast g \ast h \right)(W_t)}{1 - (F \otimes G \otimes H)(W_t)} \left[ a_3 - m(W_t) \right] \]  

2.4.4 Modeling Total Faults
The total fault removal phenomenon is the superimposition of the simple, hard and complex faults, and is therefore given as:

\[ m(W_t) = m_1(W_t) + m_2(W_t) + m_3(W_t) \]  

A particular case of the proposed model is tabulated in Table 2.1.
Genetic Algorithm Based Optimal Testing Effort Allocation Problem for Modular Software

2.5 Reliability Evaluation
Using the SRGM we can evaluate the reliability of the software during the progress of testing and predict the reliability at the release time. Reliability of software is defined as “given that the testing has continued up to time t, the probability that a software failure does not occur in time interval \((t, t + \Delta t)\) \((\Delta t \geq 0)\).”. Hence the reliability of software is represented mathematically as

\[
R(t) = R(t + \Delta t | t) = \exp\left(-\left(m(t + \Delta t) - m(t)\right)\right)
\]

Another measure of software reliability at time t is defined as “the ratio of the cumulative number of detected faults at time t to the expected number initial fault content of the software” given by[4]:

\[
R(W_t) = \frac{m(W_t)}{a}
\]

3. PARAMETER ESTIMATION AND MODEL VALIDATION
To measure the performance of the proposed model we have carried out the parameter estimation on the data set cited in M.Ohba [8](DS-I). The software was tested for 19 weeks during which 47.65 computer hours were used and 328 faults were removed. The estimation results for Exponential, Rayleigh, and Weibull function are given in table 3.1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>19029.3</td>
<td>353</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.0001</td>
<td>0.05218</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>-</td>
<td>6.53029</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.992</td>
<td>0.996</td>
</tr>
<tr>
<td>MSE</td>
<td>38.79684</td>
<td>75.31579</td>
</tr>
</tbody>
</table>

Table 3.1: Testing Effort Function Parameter Estimates

4. TESTING RESOURCE ALLOCATION PROBLEM
4.1 Notations:
- \(j\); 1,2,3; Simple faults-1;Hard Faults-2, Complex Faults-3
- \(i\); Module, 1,2..N
- \(N\); Total number of modules
- \(m_i(W_i)\); Mean value function for \(i^{th}\) module
- \(b_{ji}\); Constant fault detection rate for \(j^{th}\) fault type in \(i^{th}\) module
- \(a_{ji}\); Constant, representing the number of \(j^{th}\) fault type lying dormant in \(i^{th}\) module at the beginning of testing.
- \(c_{ji}\); Cost of removing \(j^{th}\) fault from \(i^{th}\) module
- \(W_i\); Testing effort for \(i^{th}\) module
- \(R_i\); Reliability of \(i^{th}\) module
- \(B\); Total cost of removing different types of faults

Weibull effort function is chosen to represent the testing effort as it provided the best fit on the testing effort data (based on the highest value of \(R^2\)). Based upon these estimated parameters, parameters of proposed SRGM were estimated. The goodness of fit measures used are Mean Square Error (MSE) and Coefficient of multiple determination (\(R^2\)). The results are compared with SRGM proposed by Kapur et al. [19] with three types of fault. The results are tabulated in table 3.2 (Letting \(b_1=b_2=b_3=b\))

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>353</td>
<td>378</td>
</tr>
<tr>
<td>(b)</td>
<td>0.05218</td>
<td>0.09722</td>
</tr>
<tr>
<td>(\mu)</td>
<td>26.71107</td>
<td>-</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>6.53029</td>
<td>-</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.992</td>
<td>0.992</td>
</tr>
<tr>
<td>MSE</td>
<td>38.79684</td>
<td>75.31579</td>
</tr>
</tbody>
</table>

Table 3.2: Parameter Estimates for DS-I

Figure 3.1: Goodness of Fit Curve for DS-I

Copy Right © BIJIT – 2012; January - June, 2012; Vol. 4 No. 1; ISSN 0973 – 5658
4.2 Mathematical Formulation
Consider software with ‘N’ modules where each module is different in size, complexity, the functions they perform etc. In each module there are three types of faults; simple, hard and complex. The software has to be released in the market at a predefined software release time with limited availability of testing resources expenditure. Further the cost of removing the fault from each module is dependent on its severity. Therefore, the problem of maximizing the faults of each of N independent modules such that reliability of each module is at least $R_0$ is formulated as:

Maximize

$$m(W_i) = \sum_{i=1}^{N} m_i(W_i)$$

$$= \sum_{i=1}^{N} \left( a_1 \left( e^{-b_1 W_i} \right) + \sum_{j=1}^{N} \left( a_2 \left( 1 - (1 + b_2 W_i) e^{-b_3 W_i} \right) \right) + \sum_{j=1}^{N} \left( a_3 \left( \Phi \left( W_i, \mu_i, \sigma_i^2 \right) \right) \right) \right)$$

Subject to:

$$\sum_{i=1}^{N} \left( C_1 m_1(W_i) + C_2 m_2(W_i) + C_3 m_3(W_i) \right) \leq B \quad i = 1, 2...N$$

$$\sum_{i=1}^{N} W_i \leq W \quad i = 1, 2...N$$

$$R_i \geq R_0 \quad i = 1, 2...N \quad (P1)$$

$$W_i \geq 0 \quad i = 1, 2...N$$

5. GENETIC ALGORITHM FOR TESTING RESOURCE ALLOCATION
The above optimization problem is solved by a powerful computerized heuristic search and optimization method, viz. genetic algorithm (GA) that is based on the mechanics of natural selection and natural genetics. In each iteration (called generation), three basic genetic operations i.e., selection, reproduction, crossover and mutation are executed. For implementing the GA in solving the allocation problem, the following basic elements are to be considered.

5.1 Chromosome Representation
Genetic Algorithm starts with the initial population of solutions represented as chromosomes. A chromosome comprises genes where each gene represents a specific attribute of the solution. Here the solution of the testing-effort allocation problem in modular software system includes the effort resources consumed by individual modules. Therefore, a chromosome is a set of modular testing effort consumed as part of the total testing effort availability.

5.2 Initial Population
For a given total testing time $W$, GA generates the initial population randomly. It initialize to random values within the limits of each variable.

5.3 Fitness Of A Chromosome
The fitness is a measure of the quality of the solution it represents in terms of various optimization parameters of the solution. A fit chromosome suggests a better solution. In the effort allocation problem, the fitness function is the objective of testing effort optimization problem along with the penalties of the constraints that are not met.

5.4 Selection
Selection is the process of choosing two parents from the population for crossover. The higher the fitness function, the more chance an individual has to be selected. The selection pressure drives the GA to improve the population fitness over the successive generations. Selection has to be balanced with variation form crossover and mutation. Too strong selection means sub optimal highly fit individuals, will take over the population, reducing the diversity needed for change and progress; too weak selection will result in too slow evolution. We use “Tournament selection” here.

5.5 Crossover
Crossover is the process of taking two parent solutions and producing two similar chromosomes by swapping sets of genes, hoping that at least one child will have genes that improve its fitness. In the testing resource allocation problem, crossover diversifies the population by swapping modules with distinct time consuming, particularly when the population size is small.

5.6 Mutation
Mutation prevents the algorithm to be trapped in a local minimum. Mutation plays the role of recovering the lost genetic materials as well as for randomly disturbing genetic information. The important parameter in the mutation technique is the mutation probability. The mutation probability decides how often parts of chromosome will be mutated. If there is no mutation, offspring are generated immediately after crossover (or directly copied) without any change. In our problem of testing resource allocation, we have used a mutation probability of 10%.

With the basic modules of genetic algorithm described above, the procedure for solving the optimal effort allocation problem is as follows [6]:

Step 1: Start
Step 2: Generate random population of chromosomes
Step 3: Evaluate the fitness of each chromosome in the population
Step 4: Create a new population by repeating following steps until the new population is complete:

[Selection] Select two parent chromosomes from a population according to their fitness
[Crossover] With a crossover probability, cross over the parents to form new offspring (children). If no crossover is performed, offspring is the exact copy of parents.
[Mutation] With a mutation probability, mutate offspring at each locus (position in chromosome)
[Accepting] Place new offspring in the new population
[Replace] Use new generated population for further sum of the algorithm.
[Test] If the end condition is satisfied, stop and return the best solution in the current population
[Loop] Go to step 3 for fitness evaluation

6. NUMERICAL EXAMPLE
The Effort Allocation Problem described in section 4 is illustrated numerically in this section. Consider a software system consisting of three modules, whose parameters have already been estimated using software failure data. These parameter estimates for each module is shown in Table 6.1. The total testing resources available is assumed to be 5000 units. Total cost for removing the different types of faults is 10000 units. Also, it is desired that the reliability of each module is at least 0.9.

<table>
<thead>
<tr>
<th>Module</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$b$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>357</td>
<td>107</td>
<td>81</td>
<td>0.00368</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>16.2925</td>
<td>5.586</td>
</tr>
<tr>
<td>2</td>
<td>332</td>
<td>97</td>
<td>76</td>
<td>0.00234</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>14.9874</td>
<td>4.123</td>
</tr>
<tr>
<td>3</td>
<td>298</td>
<td>64</td>
<td>32</td>
<td>0.0018</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>12.4567</td>
<td>7.654</td>
</tr>
</tbody>
</table>

Table 6.1: Parameter Estimates for effort allocation problem

Based on the above information, the problem (P1) is solved using genetic algorithm. The parameters used in GA evaluation are given in table 6.2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Size</td>
<td>106</td>
</tr>
<tr>
<td>Number of Generations</td>
<td>26</td>
</tr>
<tr>
<td>Selection Method</td>
<td>Tournament</td>
</tr>
<tr>
<td>Crossover Probability</td>
<td>0.9</td>
</tr>
<tr>
<td>Mutation Probability</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 6.2: Parameter of the GA

The optimal testing time allocation to each type of fault in module and hence total fault removed from each module and their corresponding cost of removing is shown in table 6.3.

<table>
<thead>
<tr>
<th>Module</th>
<th>$W$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>Reliability</th>
<th>Cost of removing faults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1192.22</td>
<td>309</td>
<td>100</td>
<td>81</td>
<td>490</td>
<td>0.978</td>
</tr>
<tr>
<td>2</td>
<td>1602.294</td>
<td>324</td>
<td>86</td>
<td>76</td>
<td>486</td>
<td>0.962</td>
</tr>
<tr>
<td>3</td>
<td>2202.934</td>
<td>292</td>
<td>58</td>
<td>32</td>
<td>382</td>
<td>0.969</td>
</tr>
<tr>
<td>Total</td>
<td>4997.448</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1358</td>
</tr>
</tbody>
</table>

Table 6.3: The optimal testing effort expenditure with the corresponding cost of each module

7. CONCLUSION
In this paper we have discussed the problem for modular software at the unit testing stage. We have made use of unified scheme for presenting a generalized framework for Software reliability growth modeling with respect to testing effort expenditure and incorporated the faults of different severity. The faults in each module are of three types-simple, hard and complex. Further we have optimally allocated the testing effort to each type of fault and the modules and have found out the different types of faults removed in the modules with a fixed budget and a prerequisite level of reliability. Genetic Algorithm is developed to solve the problem of resource allocation. Numerical example is discussed to illustrate the solving of the discussed optimization problem through GA.

FUTURE SCOPE
The present study is done under the assumption of independence of the failures of different modules. In future, dependence of the failures from different modules as well as the architecture styles and connectors reliability can also be studied.

REFERENCES


