Exhausting Meta-heuristic Nature Inspired Approaches for the Parameter Estimation analysis of Software reliability Growth Model

Sangeeta¹, Kapil Sharma² and Manju Bala³

¹²Dept. of Computer Science and Engineering, Delhi Technological University
Shahbad Daulatpur, Main Bawana Road, Delhi, 110042, India
¹sangeeta@dtu.ac.in, ²kapil@ieee.org

³Dept. of Computer Science and Engineering, IP College for Women, Delhi University
31, Sham Nath Marg, Civil Lines, Delhi, 110054, India
Corresponding Author: ²kapil@ieee.org

Abstract. The parameters are the basic building block for any software reliability growth model. The more accurately the parameters are measured more is the accuracy of the Software Reliability Growth Model (SRGM). So there is an intense requirement of a process with which optimum values of the parameters could be measured. Although there are methods like Maximum Likelihood Estimation (MLE), Least Square Estimation (LSE) etc. but are having various constraints and boundaries while measuring the parameters. The authors in this paper take this challenge of identifying the efficient method with which the parameter values can be optimized. The authors have used time domain datasets for estimation process and conducted experiments on the real world datasets and comparison of the results has been done with the actual failure dataset and among the algorithms used. The authors analyzed the capability of meta-heuristic nature inspired algorithms with a new proposed hybrid algorithm for estimating the parameter values of the software reliability growth model. The results obtained are very much satisfactory and might be used on other SRGMs as well and can assess the reliability of any system with more accuracy.

Keywords: We would like to encourage you to list your keywords in this section.

1 Introduction

The computers and software are making intuitive changes to every aspect of human life. In today’s era the quality of the software is a challenging issue but for the software systems scheduled time of detonation is unidentified. It is really too much difficult to have a reliable software due to colossal size of the software, its complexity, changeability, interactive and distributive nature of the software systems. Reliable software offers a grand challenge and it deserves the attention of researchers a lot. Reliability of the software can be used for determining when to release the software
and can be used to govern the future needs of more resources for software testing. Reliability is measured only by using the number of failures in the system. There are two categories of datasets available that can be used for the reliability estimation of any software [1], first one of these are the interval domain based datasets and second one are the time domain based datasets.

In the last thirty five years, number of models have been developed [2-9] for the assessment of reliability of the software and have been studied by the authors in this paper. Specific statistical distribution is used to develop a mathematical expression for the model from the failure dataset. Using the developed model the reliability of the software product can be assessed. These software reliability models are having the parameters which are showing the physiognomies of the model developed and traditional approaches have been used for the parameter estimation. Traditional methods used for parameter estimation includes Least square estimation (LSE), Maximum likelihood estimation (MLE) which involves that at one time only one solution will hold, then locally look the direction in which to move and after taking decision of how far to move select the new current. But these algorithms are not applicable where the objective functions are not differentiable [10]. Also these algorithms do not allow the multiple optima, in this case these will stuck into the local optimum near to where they start and progresses slowly when the number of parameters are large. Each iteration of algorithm will require as many function evaluations as the numbers of parameters are there and the parameter estimation done by the traditional methods are not applicable to the complex software systems. In this paper authors surveyed nature inspired meta-heuristic approaches for the parameter estimation process which have been found to be more satisfactory then the traditional methods and have been applied in various domains by the researchers.

This paper is organized in different sections starting with the Introduction section, followed by Related Researches section. The next section describes the Proposed Work which is followed by the Implementation and Results section. At last the Conclusion section and the References are given.

### 2 Related Researches

Nature inspired optimization algorithms entails only few of the assumptions for the solution that is to be optimized. In [11-13] Particle swarm optimization technique and its variants has been used for parameter estimation and some authors [14-15] also applied Genetic algorithm and its variants for estimation of the parameters for SRGMs. Only three papers have been found [11,14,16] in which evolutionary algorithms have been used for parameter estimation of the SRGM. These are including only the application of PSO and GA.

Using Particle Swarm Optimization method, the problem is solved by having a candidate solution in terms of particles. Each particle moves in the search space by using mathematical formulas for the position and velocity updates to make the final global solution. Only few papers have been found by the authors [20] that are having use of PSO [21] in different applications for parameter optimization of software relia-
bility models. In the experimental analysis Trelea-vectorized PSO has been used by the authors. Genetic algorithm was developed by the John Holland and his team members at the University of Michigan. This technique is based on the genetics and natural selection process where survival of fittest is the aim [14], [22]. The genetic algorithm uses three steps processes as selection of the parent from the candidate solutions then cross over them with some probability and then perform mutation operation in order to find the best candidate solution. For the implementation process real valued genetic algorithm has been used by the author’s for the parameter estimation of the software reliability growth models.

After reviewing the meta-heuristic algorithms in various domains, authors applied the discussed algorithms and also tried to analyze the parameter estimation capability of a hybrid algorithm [23] using the Gravitational search algorithm and particle swarm optimization [24]. Using gravitational affects PSOGSA helps in finding the best solution for guiding the heavy masses towards the global optimum positions. This process also increases the speed and overall movement of particles/masses as well and will enhance the exploitation capability in terms of the convergence rate and optimal parameter estimation competency of PSOGSA algorithm. Authors also proposed a new hybrid approach using the PSOGSA and Genetic algorithm and analyze the results on time domain dataset. Previously all these algorithms have not been applied altogether for reliability assessment.

3 Proposed Work

Authors in this paper used time domain datasets for the parameter estimation analysis of the Goel Okumotto [2] software reliability growth models (SRGM) which is based on the non-homogenous Poisson process using different meta-heuristic optimization approaches for the parameter estimation[17]. Goel Okumooto model is one of the basic models for reliability assessment. This is Continuous time The mean value function $m(t)$ and failure intensity $\lambda(t)$ of the model is given by [2]-

\[ m(t) = a(1 - e^{-bt}) \quad \text{and} \quad \lambda(t) = m(t) = abe^{-bt} \]

where $m(t)$ gives the expected number of errors in the software at t time,

- $a$ denotes total number of initial errors in software and $b$ denotes the fault detection rate, $t$ is the time interval.

There is the requirement of estimating the parameters $b$ and $a$ in such a way that these values maps exactly or approximately close to the fault detection rate and actual number of faults values in the real datasets. If the estimated values of the parameters are not closely mapping the actual values of the parameters then the software developers will not be able to make the decisions accurately as what more number of resources will be required in future to achieve the required level of reliability for the software and what will be the exact time of software release. So the intense requirement is to estimate the parameters exactly. Inspiring from the nature or biological world, population based optimization process is becoming more popular from the last two decades [19] comparative to the traditional methods. These optimization methods
goes to a class of stochastic search strategies, modeled after evolutionary mechanism and instead of dealing with a single candidate solution in population based optimization methods these are dealing with a population of candidates that can contribute jointly in reconnoitering the new solution. The benefits of using these methods are not just of speedup but these methods are ample robust and do rationally well. There is a great success of these algorithms in finding the optimum solution. Having seen the benefits of using these algorithms for optimization [16], authors in this paper applied a set of evolutionary algorithms (GA, PSO, PSOGSA) with a proposed hybrid PSOGSAGA algorithm altogether in parameter estimation and reliability assessment capability of software reliability growth model.

Objective function need to be defined precisely for each type of dataset and for each optimization algorithm. In this paper authors have used MLE based fitness function rather than LSE based function[10]. The fitness function used for type 2 data has been developed using the likelihood function and is given in the equation [1] as-

\[
\text{fitness}_{type2} = \sum_{i} \log(\Lambda(s)) - m(s_i)
\]

where  \(s_i\) is successive time of observed failures and \(m(s_i)\) is the estimated number of detected errors

As the value of this fitness increases, the solution will become better towards the global optimum.

### 4 Implementation and Results

For the implementation of the proposed work the authors used GO Software Reliability Growth Model and time domain type-2 data set taken from [1] (page number 141 Data set #4) and [2]. The author implemented the proposed work in Mat Lab using the nature inspired meta-heuristic algorithms Particle Swarm Optimization, Genetic Algorithm, hybrid PSOGSA and PSOGSA-GA algorithm.

The NTDS dataset is having 34 failures observed during the production phase, test phase and user phases. Implementation of the proposed work has been done using the processor Intel(R) Core (TM) i5-62000 CPU @ 2.40 GHz having installed memory (RAM) 4.00 GB with system type 64 bit Windows 10 Operating system, x64 based-processor.

#### 4.1 Parameter estimation analysis using Type 2 data

Authors have used the different approaches (GA, PSO, PSOGSA, and PSOGSA-GA) for parameter estimation on time domain dataset for the GO model. For this Type-2(time domain) NTDS dataset has been used and the dataset has been extracted from the failures in the software of the real time multicomputer programming center of the NTDS.
4.2 Implementation Result Analysis

Figure 1 shows the comparison of the actual number of errors in the software and the number of errors estimated over time $t$ by using different nature inspired meta-heuristic algorithms like GA, PSO, PSOGSA and PSOGSA-GA implemented by the authors in this paper.

Fig. 1. Comparison of mean value function among different approaches used

<table>
<thead>
<tr>
<th>Sr. no</th>
<th>Optimization method used</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Genetic algorithm (GA)</td>
<td>34.3050</td>
<td>0.0068</td>
</tr>
<tr>
<td>2</td>
<td>Particle swarm optimization(PSO)</td>
<td>35.9554</td>
<td>0.0068</td>
</tr>
<tr>
<td>3</td>
<td>Hybrid PSOGSA</td>
<td>30.0675</td>
<td>0.0051</td>
</tr>
<tr>
<td>4</td>
<td>Hybrid PSOGSA-GA</td>
<td>30.0585</td>
<td>0.0053</td>
</tr>
</tbody>
</table>

As shown in Table 1 using PSOGSA and PSOGSA-GA algorithms, the values of the parameters ‘$a$’ and ‘$b$’ are better optimized than other algorithms (GA,PSO).

Figure 2 shows the estimated reliability of the software by using the GA, PSO and PSOGSA and proposed Hybrid PSOGSA-GA algorithm for the perfect debugging process based GO Model and the reliability estimation using PSOGSA-GA (70.69%)
algorithm is better than the other algorithms (68.8% GA), (66.9% PSO), (68.6% PSOGSA).

![Comparison of reliability estimation among different approaches used](image)

**Fig. 2.** Comparison of reliability estimation among different approaches used

<table>
<thead>
<tr>
<th>Sr. no</th>
<th>Approach Used</th>
<th>Elapsed time in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Best</td>
</tr>
<tr>
<td>1</td>
<td>GA</td>
<td>51.12580</td>
</tr>
<tr>
<td>2</td>
<td>PSO</td>
<td>1.607097</td>
</tr>
<tr>
<td>3</td>
<td>PSOGSA</td>
<td>1.595033</td>
</tr>
<tr>
<td>4</td>
<td>PSOGSA-GA</td>
<td>50.43287</td>
</tr>
</tbody>
</table>

Table 2. Comparison of Elapsed time for all the used optimization approaches

Table 2 shows the convergence rate of various algorithms with best, worst and average elapsed time and PSOGSA shows the better convergence rate than the other algorithms.

4.3 Summary results

The optimization capability of the evolutionary algorithms have been analyzed by the authors in terms of their parameter estimation capability, reliability assessed, deviation from the actual and estimated values, convergence rate in terms of elapsed time for each algorithm.

The results above shows that the PSOGSA-GA algorithm outperforms than other algorithms in the parameter estimation, error detection and the reliability estimation. The convergence rate of the PSOGSA algorithm is much better than PSO, GA and PSOGSA algorithms.
So the error detection, parameter estimation and reliability estimation of the GO SRGM model using hybrid algorithm gives better results than the GA and PSO algorithms independently.

5 Conclusions

This is the first attempt where various population based optimization algorithm has been applied altogether for the parameter estimation of software reliability growth model. The experimental results have been performed on real datasets and the results estimation is found to be very much acceptable then the earlier parameter estimation approaches. The reliability estimation has been done at each step of analysis and can be used for quality analysis in order to identify the release time, maintenance cost, resource requirements in the software development, and will be helpful in finding the total expenditure involved during the software development process.

References