The Use of Cuckoo Search in Estimating the Parameters of Software Reliability Growth Models

Dr. Najla Akram AL-Saati and Marwa Abd-ALKareem
Software Engineering Dept.
College of Computer Sciences & Mathematics, Mosul, Iraq

International Journal of Computer Science and Information Security (IJCSIS), 2013

2013. 8. 7
Taehyoun Kim
Outlines

- Introduction
- Related Work
- Cuckoo Search (CS)
- Experiments
- Conclusion & Future Work
- Discussion
Introduction

- **Software Reliability**
  - The probability of failure-free software operation for a specified period of time in a specified environment
    

- **Software Reliability Growth Models (SRGMs)**
  - The theoretical models for estimating software reliability growth based on failure detection phenomenon during testing phase
    

*Figure 1. Basic concept of SRGMs*
Parameter Estimation of SRGM

- Traditional SRGMs have some parameters which have physical interpretations.

Ex) Goel-Okumoto Model

Cumulative failure function
\[ m(t) = a(1 - e^{-rt}) \]

- \( a \): the number of total failures
- \( r \): failure detection rate

![Figure 2. cumulative failure function](image)

- Optimal

Failure data
Classical Estimation Techniques

- There are two of the most popular estimation techniques.
  - Maximum Likelihood Estimation (MLE)
  - Least Square Estimation (LSE)

- However, these techniques are not effective to estimate parameters of complicated function like non-linear function.
  - SRGMs are non-linear functions.

- The interest in **heuristic algorithm** began to have effect in solving this problem.
  - There are many nature inspired algorithms.
Related Work

- **Genetic Algorithm (GA) Approach**
  : Survival of the genetically fittest
  

- **Particle Swarm Optimization (PSO) approach**
  : Flock migration
  

- **Ant Colony Optimization (ACO) approach**
  : Shortest path to food source
  
Cuckoo’s Inspiring Behavior (1/2)

**Behavior 1.** They have a strange habit of laying their eggs in public nests.

**Behavior 2.** They remove other’s eggs to boost the hatching probability of their own eggs.

**Behavior 3.** When the first cuckoo baby bird is hatched, he instinctively throw out host eggs to increase his share of food.
Cuckoo’s Inspiring Behavior (2/2)
Cuckoo Search

Basic Steps of Cuckoo Search (1/5)

- Step1. Generate initial population of n host nests.

Ex) Goel-Okumoto model

\[ m(t) = a(1 - e^{-rt}) \]

\[ (a_1, r_1), (a_2, r_2), \ldots, (a_n, r_n) \]: a candidate for optimal parameters
Basic Steps of Cuckoo Search (2/5)

- **Step2. Lay the egg \((a_{k'}, b_{k'})\) in the k nest.**
  - K nest is randomly selected.
  - Cuckoo’s egg is very similar to host egg.
    - \(a_{k'} = a_k + \text{Randomwalk(\text{\it Lévy flight})}_a\)
    - \(r_{k'} = r_k + \text{Randomwalk(\text{\it Lévy flight})}_r\)
Basic Steps of Cuckoo Search (3/5)

- Step 3. Compare the fitness of cuckoo’s egg with the fitness of the host egg.
  - Root Mean Square Error (RMSE)

\[
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (m(t) - m_t)^2}
\]

- \( m(t) \): the number of predicted failures by time \( t \)
- \( m_t \): the number of actual failures by time \( t \)

Which egg is better?
Basic Steps of Cuckoo Search (4/5)

- **Step 4.** If the fitness of cuckoo’s egg is better than host egg, replace the egg in nest $k$ by cuckoo’s egg.

$$fitness(\text{cuckoo's egg}) \geq fitness(\text{host egg})$$

$$fitness(\text{cuckoo's egg}) < fitness(\text{host egg})$$
Basic Steps of Cuckoo Search (5/5)

- Step 5. If host bird notice it, the nest is abandoned and new one is built. ($p < 0.25$) (to avoid local optimization)

- Iterate steps 2 to 5 until termination criterion is met.
### Experimental Set-up (1/2)

#### Comparison with other techniques

<table>
<thead>
<tr>
<th>Reference</th>
<th>Parameter estimation techniques</th>
<th>Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Sheta’s [1]</td>
<td>Particle Swarm Optimization (PSO)</td>
<td>Data Set 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Set 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Set 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Musa Data 3*</td>
</tr>
</tbody>
</table>


**Table1. Estimation techniques and data sets**

Experimental Set-up (2/2)

- **SRGMs for the experiments**

<table>
<thead>
<tr>
<th>SRGMs</th>
<th>Mean value function (cumulative failure function)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goel-Okumoto (GO) model</td>
<td>( m(t) = a(1 - e^{-rt}) )</td>
</tr>
<tr>
<td>Power (POW) model</td>
<td>( m(t) = at^r )</td>
</tr>
<tr>
<td>Yamada Delayed S-Shaped (DSS) model</td>
<td>( m(t) = a(1 - (1 + rt)e^{-rt}) )</td>
</tr>
<tr>
<td>Musa-Okumoto (M-O) model</td>
<td>( m(t) = aln(1 + rt) )</td>
</tr>
</tbody>
</table>

*Table2. SRGMs for the experiments*

- **Parameter settings for the cuckoo search**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower and Upper bound for parameter a</td>
<td>[0.00001 ~ 2000]</td>
</tr>
<tr>
<td>Lower and Upper bound for parameter r</td>
<td>[0.00001 ~ 1]</td>
</tr>
<tr>
<td>Number of Nests</td>
<td>10</td>
</tr>
<tr>
<td>Number of iterations(Generations)</td>
<td>100</td>
</tr>
</tbody>
</table>

*Table3. Parameter settings for cuckoo search*
The impact of training and testing data

- A balance point is training 70% and testing 30%.

### Table 4. The impact of training and testing percentage
(a) data set 1, (b) data set 2, (c) data set 3
## Comparison with other techniques (1/4)

- **PSO vs. Cuckoo Search (Data set 1)**


<table>
<thead>
<tr>
<th></th>
<th>Training 70%</th>
<th>Testing 30%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSO</td>
<td>CS</td>
</tr>
<tr>
<td>G-O</td>
<td>20.2565</td>
<td>34.0933</td>
</tr>
<tr>
<td>POW</td>
<td>22.2166</td>
<td>44.8663</td>
</tr>
<tr>
<td>DSS</td>
<td>15.9237</td>
<td>32.6376</td>
</tr>
</tbody>
</table>

**Table 5. Comparison with PSO using Data set 1**

**Figure 3. Comparison with PSO using Data set 1**
Comparison with other techniques (2/4)

- PSO vs. Cuckoo Search (Data set 2)

<table>
<thead>
<tr>
<th></th>
<th>Training 70%</th>
<th>Testing 30%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSO</td>
<td>CS</td>
</tr>
<tr>
<td>G-O</td>
<td>24.9899</td>
<td>33.2311</td>
</tr>
<tr>
<td>POW</td>
<td>32.355</td>
<td>47.0571</td>
</tr>
<tr>
<td>DSS</td>
<td>20.8325</td>
<td>27.9159</td>
</tr>
</tbody>
</table>

Table 6. Comparison with PSO using Data set 2

Figure 4. Comparison with PSO using Data set 2
Comparison with other techniques (3/4)

- PSO vs. Cuckoo Search (Data set 3)

<table>
<thead>
<tr>
<th></th>
<th>Training 70%</th>
<th>Testing 30%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSO</td>
<td>CS</td>
</tr>
<tr>
<td>G-O</td>
<td>12.8925</td>
<td>13.5404</td>
</tr>
<tr>
<td>DSS</td>
<td>18.5807</td>
<td>13.6634</td>
</tr>
</tbody>
</table>

Table 7. Comparison with PSO using Data set 3

Figure 5. Comparison with PSO using Data set 3
Comparison with other techniques (4/4)

- **ACO vs. Ex-ACO vs. Cuckoo Search**


<table>
<thead>
<tr>
<th></th>
<th>Musa Data Set 1 - Training 100%</th>
<th>Musa Data Set 2 - Training 100%</th>
<th>Musa Data Set 3 - Training 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACO</td>
<td>EX-ACO</td>
<td>CS</td>
</tr>
<tr>
<td>G-O</td>
<td>60.0371</td>
<td><strong>28.5891</strong></td>
<td>41.7971</td>
</tr>
<tr>
<td>POW</td>
<td>52.8854</td>
<td><strong>34.0521</strong></td>
<td>45.9783</td>
</tr>
<tr>
<td>DSS</td>
<td>52.8854</td>
<td><strong>33.0461</strong></td>
<td>42.2256</td>
</tr>
<tr>
<td>M-O</td>
<td>26.0385</td>
<td><strong>17.359</strong></td>
<td>41.7732</td>
</tr>
</tbody>
</table>

Table 8. Comparison with ACO and Extended ACO
(a) musa data set 1, (b) musa data set 2, (c) musa data set 3
Conclusion & Discussion

- **Conclusion**
  - Cuckoo Search can be successfully applied to find acceptable solutions to problem of parameter estimation of SRGM.

- **Future work**
  - Other swarm intelligent methods can be applied and compared.
  
    Ex) Harmony Search, Firefly algorithm, Accelerated Particle Swarm Optimization(APSO), etc..
Discussion

- **Cons.**
  - New heuristic algorithm was applied to estimate the parameters of SRGM.
  - The results are obtained very quickly and accurately.

- **Pros.**
  - Experiments were not proper to compare cuckoo search with other heuristics.
Q&A

Thank you
Least Square Estimation (LSE) (1/2)

- Minimize the sum of square of the deviations between actual data and expected data

Ex) parameter estimation of function \( f(x) = ax + b \)

\[
E(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2
\]

When \( \frac{d(E(a,b))}{da} = 0, \frac{d(E(a,b))}{db} = 0 \),

The values of \( a \) and \( b \) are the best estimates of \( a \) and \( b \).

Figure 2. Basic concept of LSE
Least Square Estimation (LSE) (2/2)

Ex) Goel-Okumoto Model

\[ m(t) = a(1 - e^{-rt}) \]

\[ E(a, b) = \sum_{i=1}^{n} (y_i - a(1 - e^{-rt_i}))^2 \]

When \( \frac{d(E(a,b))}{da} = 0, \frac{d(E(a,b))}{db} = 0, \)

The values of \( a \) and \( b \) are the best estimates of \( a \) and \( b \).
Lévy flight (1/2)

- A random walk in which the step-lengths have a probability distribution that is heavy-tailed (to avoid local optimization)

*Figure 6. The basic concept of random walk (Lévy flight)*
Lévy flight (2/2)

- Normal distribution VS. Heavy tail distribution