Optimal Release Policy for Covariate Software Reliability Models

Ebenezer Yawlui[1], Priscila Silva[1], Vidhyashree Nagaraju[2], and Lance Fiondella[1]

[1]University of Massachusetts Dartmouth, [2]Bastion Technologies Inc

UMass | Dartmouth

Abstract

The optimal time to release a software is a common problem of broad concern to software engineers, where the goal is to minimize cost by balancing the cost of fixing defects before or after release as well as the cost of testing. However, the vast majority of these models are based on defect discovery models that are a function of time and can therefore only provide guidance on the amount of additional effort required. To overcome this limitation, this research presents a **software optimal release model based on cost criteria incorporating the covariate software defect detection model** to explicitly links test activities to defect discovery, and determine optimal allocation of multiple test activities to minimize overall cost of producing software.

Covariate Defect Discovery Models

Discrete Cox proportional hazards (DCPH) Non-Homogenous Poisson Process (NHPP) Software Reliability Growth Models (SRGM) correlates r covariates to the number of events in each of n intervals, through the *mean value function*

$$m(\mathbf{x}) = \omega \sum_{i=1}^{n} p_{i,\mathbf{x}_i}$$

where

- $\omega > 0$ - number of defects that would be discovered with infinite testing
- $p_{i,\mathbf{x}_i}$ - probability that a defect is discovered in the i^{th} interval, given that it was not discovered in the first $(i-1)$ intervals described as

$$p_{i,\mathbf{x}_i} = \left(1-(1-h(i))^{g(\mathbf{x}_i;\boldsymbol{\beta})}\right)\prod_{k=1}^{i-1}(1-h(k))^{g(\mathbf{x}_k;\boldsymbol{\beta})}$$

- $h(\cdot)$ - baseline hazard function
- $\boldsymbol{\beta}$ - vector of r parameters contained in DCPH model

To estimate the parameters of the DCPH model, the log-likelihood function is

$$LL(\boldsymbol{\gamma},\boldsymbol{\beta},\omega) = -\omega\sum_{i=1}^{n} p_{i,x_i} + \sum_{i=1}^{n} y_i \ln(\omega) + \sum_{i=1}^{n} y_i \ln(p_{i,x_i}) - \sum_{i=1}^{n} \ln(y_i!)$$

where

- $\boldsymbol{\gamma}$ - vector of model parameters contained in the hazard function
- y_i - number of defects discovered in the i^{th} interval.

Optimal Test Allocation Based On Cost Criterion

Given model $\widehat{m}(\mathbf{x})$ fitted to test activities $\mathbf{x}_{n\times r}$ and vector of observed failures $\mathbf{y}$, we seek to allocate resources to test activities in interval $n+1$ in a manner that minimizes cost

$$argmin\ \boldsymbol{C}(\boldsymbol{X})$$

subject to

$$\sum_{i=1}^{n} \alpha_i \mathbf{X}_{i,n+1} \leq \boldsymbol{B}$$

where

- $\boldsymbol{C}(\boldsymbol{X})$ - Total cost to release the software
- $\boldsymbol{B} > \mathbf{0}$ – Budget Constraint

The cost function for optimal release model based on cost criteria incorporating covariates is

$$\mathbf{C}(\mathbf{X}) = \alpha_{pre}\widehat{m}(\mathbf{X}) + \alpha_{post}\left(\widehat{m}(\mathbf{x}) - \widehat{m}(\mathbf{X})\right) + \sum_{i=1}^{r}\alpha_i\sum_{j=1}^{n+1}\mathrm{X}_{i,j}$$

where

- $\widehat{m}(\mathbf{X})$ - estimated number of defects to be discovered prior to release according to the fitted model
- $\widehat{m}(\mathbf{x})$ - estimated number of defects to be discovered throughout the software lifecycle
- α_{pre} - cost of removing a defect before release
- α_{post} - cost of removing a defect after release

Illustrations: Dataset DS1

DS1 (Shibata et al., 2006): $n = 17$ weeks of observation

Week	Execution Time (hr)	Failure Ident. Work (person hr)	Computer Time-Failure Ident. (hr)	Failure Ident.
1	0.0531	4	1.0	1
2	0.0619	20	0	1
3	0.1580	1	0.5	2
4	0.0810	1	0.5	1
5	1.0460	32	2.0	8
6	1.7500	32	5.0	9
7	2.9600	24	4.5	6
8	4.9700	24	2.5	7
9	0.4200	24	4.0	4
10	4.7000	30	2.0	3
11	0.9000	0	0	0
12	1.5000	8	4.0	4
13	2.0000	8	6.0	1
14	1.2000	12	4.0	0
15	1.2000	20	6.0	2
16	2.2000	32	10	2
17	7.6000	24	8.0	3
Total	32.8000	296	60.0	54

Covariates E (Execution Times) and C (Computer Time-Failure Ident.) are used.

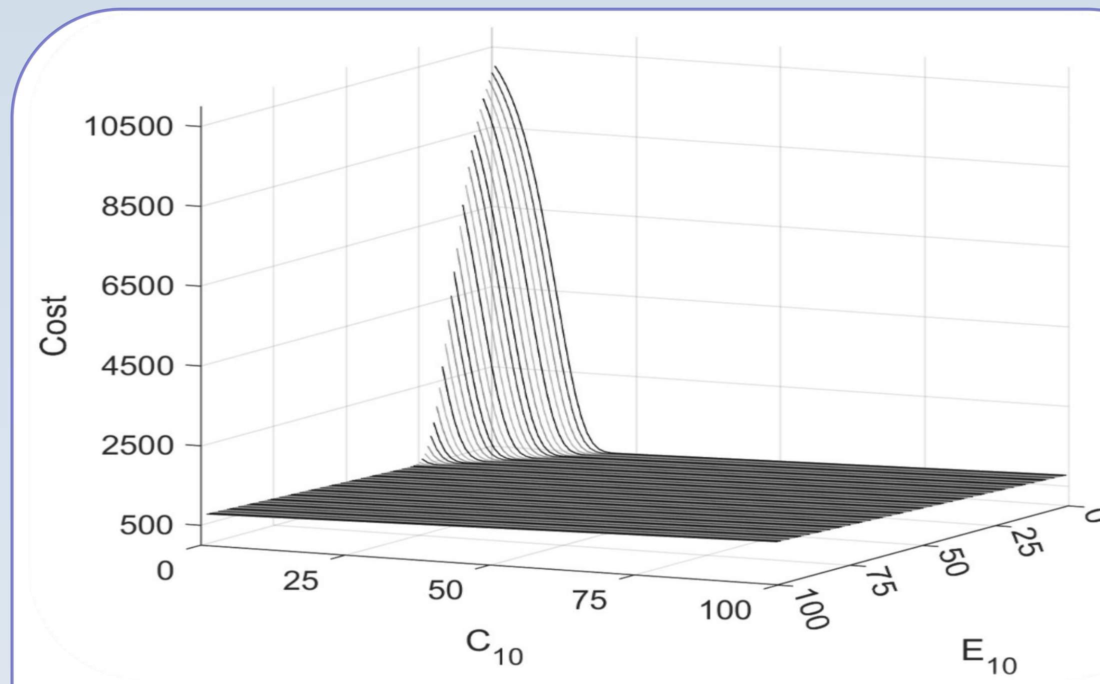

Figure 1 - Impact of test allocation on cost under negative binomial hazard function fitted to E and C covariates of DS1

The minimum of $\mathbf{C}(\mathbf{X})$ according to the fitted model is $\mathbf{C}^* = 561.12$ obtained when $E_{10}^* = 0$ and $C_{10}^* = 22.51$. Cost increases when $C_{10} > C_{10}^*$, and $E_{10} > E_{10}^*$.

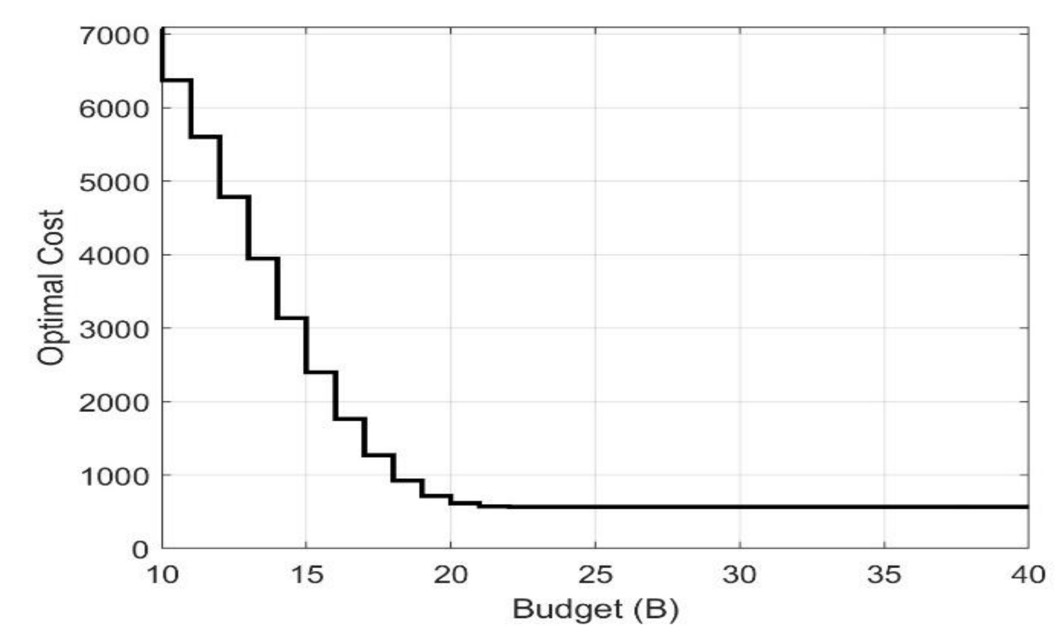

Figure 2 – Minimum cost attainable under optimal test activity allocation given budget B

Cost decreases to the optimal allocation $B = 22.51$ is reached. Increasing the budget for values greater than $B = 23$ does not reduce the cost further.

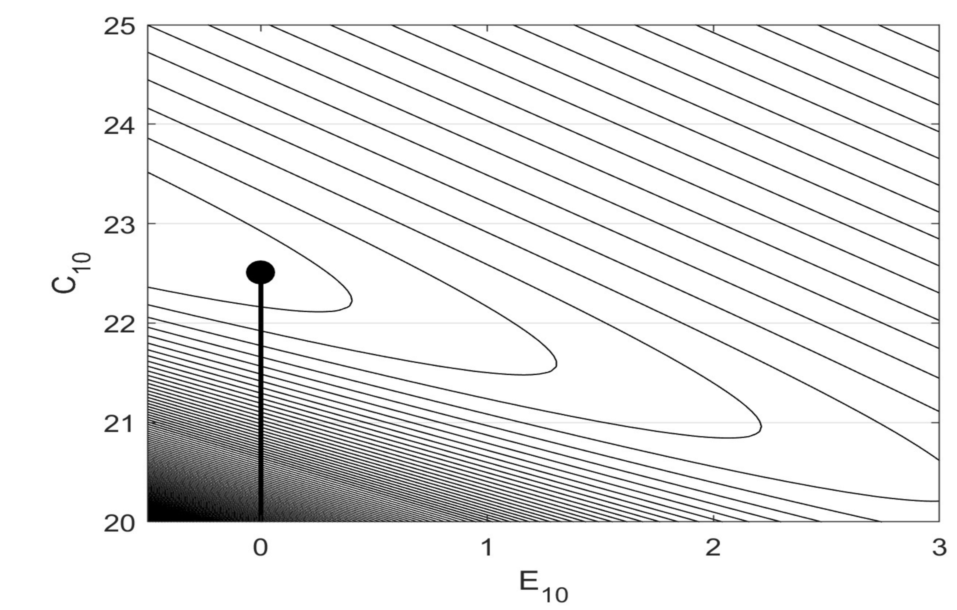

Figure 3 - Contour plot of cost function with optimal test activity allocation given budget B superimposed.

All values of $E_{10}^* = 0$, indicates that allocating additional time to this activity will not compensate the cost of performing the testing.

Summary & Conclusion

This research presented a software optimal release model based on cost criterion

- Incorporating covariate software defect detection model
- Allocating multiple alternative test activities based on a budget constraint in order to maximize defect discovery

Results suggest that optimal release policy based on covariates model provides direct guidance to software engineers on how to allocate efforts across multiple test activities in order to

- minimize cost
- maximize defect discovery

Next Steps and Future Work

Future research will explore

- methods considering parametric uncertainty to identify the optimal test activity allocation when more covariates are considered and the dimensionality of the problem increases
- Develop stable and efficient algorithms in order to establish confidence in release decisions
- Online procedures to iteratively allocate test activities, so that the approach aligns with application during the software testing process.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant Number 1749635. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author Contact Information

Ebenezer Yawlui
Department of Electrical and Computer Engineering
University of Massachusetts Dartmouth
285 Old Westport Road North Dartmouth, MA USA
Email: eyawlui@umassd.edu

Lance Fiondella, PhD
Department of Electrical and Computer Engineering
University of Massachusetts Dartmouth
285 Old Westport Road North Dartmouth, MA USA
Email: lfiondella@umassd.edu