GA Based Optimization of Software Development Effort Estimation

Kavita Choudhary
ITM University, Gurgaon, India
E-mail : kavitapunia@gmail.com

Abstract : *Software effort estimation plays a vital role in the development of software. In recent years, software has become the most expensive component of computer system projects. The major part of cost of software development is due to the human-effort, and most cost estimation methods focus on this aspect and give estimates in terms of person-month. In this paper, a new model is proposed that estimate the effort required for the development of software project using genetic algorithm approach. Software systems are becoming complex and they desire for new, effective and optimized technique with limited resources. A solution to this problem lies in nature where complex species have evolved from simple organisms and constantly become able to adapt to changes in the environment. In case of species, it takes hundreds of generations and years which are not considerable in the field of software engineering. With the use of genetic algorithm, it can be done instantly by simulating the results on various tools of genetic algorithm.*

Keywords: *Effort estimation, Effort estimation models, Walston-Felix model. COCOMO model, SEL model.*

I. Introduction

Software effort estimation is one of the steps to be carried out in project planning. The effective and efficient development of software requires accurate estimates. The objective behind estimation is not only restricted to accuracy but also to control the cost and scope of the project from organization's point of view. The aim of this paper is to propose a model that should provide optimum results. Software researchers and practitioners are providing many effort estimation techniques for several decades but the problem persist in software engineering field. Early software estimation models are based on regression analysis or mathematical derivations. Today's models are based on simulation, neural network, genetic algorithm, soft computing, fuzzy logic modeling etc. The organization of this paper is as follows. Section II elaborates some literature reviews on software effort estimation model. Section III elucidates the analysis of proposed model. Section IV describes the experimental results. Conclusion and potential future work are discussed in the Section V.

II. Related Work

This section provides some background information of various software effort estimation models to be used in this paper work.

A. SEL – Model

The Software Engineering Laboratory (SEL)[2][3] of the University of Maryland has established a model i.e. SEL Model for estimation. Estimation of effort according to SEL model is as follows:-

$E_{SEL} = 1.4 * (L)^{0.93}$

Effort (E in Person-Months) and lines of code(L in thousands of lines of code i.e. KLOC) are used as predictors.

B. Walston-Felix Model

The model developed by Walston and Felix[1][2] at IBM provides a relationship between delivered lines of source code (L in thousands of lines) and effort E (E in person-month). This model constitute various aspects of the software development environment such as user participation, customer-oriented changes, memory constraints etc. According to Walston and Felix model, effort is computed by:-

$E_{W-F} = 5.2* (L)^{0.91}$

C. COCOMO Basic Model

COCOMO[4][5][7] model is proposed by B.W.Boehm. COCOMO model have three sub-models i.e. basic, intermediate and detailed model. Basic model takes the form

$E = a_b * (L)^{b_b}$

E is effort applied in person-month. The values of coefficients ab and bb are defined by Boehm. The basic model aims at estimating in a quick and rough fashion, most of the small to medium sized software projects. Three modes of software development are considered in this model: organic, semi-detached and embedded. The calculations of effort according to three different modes are as follows:-

Basic Model – Organic Mode

$E_{B-O} = 2.4* (L)^{1.05}$

Basic Model – Semi-detached Mode

$E_{B-S} = 3.0* (L)^{1.12}$

Basic Model – Embedded Mode

$E_{B-E} = 3.6* (L)^{1.20}$

D. COCOMO Intermediate Model

In this model, Boehm introduced an additional set of 15 predictors called cost drivers in the intermediate model to take account of the software development environment. Cost drivers are used to adjust the nominal cost of a project to the actual project environment, hence increasing the accuracy of the estimate.

The multiplying factors for all 15 cost drivers are multiplied to get the effort adjustment factor (EAF). Typical values for EAF range from 0.9 to 1.4. The intermediate COCOMO considering three modes takes the form: –

E. Intermediate – Organic Mode

$E_{I-O} = 3.2* (L)^{1.05} * EAF$

Intermediate – Semi-detached Mode

$E_{I-S} = 3.0* (L)^{1.12} * EAF$

Intermediate – Embedded Mode

$E_{I-E} = 2.8* (L)^{1.20} * EAF$

III Analysis of Proposed Effort Estimation Model

The procedural analysis of proposed effort estimation model will be carried out in two stages i.e. Conceptual View and Data Analysis. Details on each stage are described below:-

A. Conceptual View

Walston-Felix model computes effort on the basis of lines of codes (L in thousands of lines).In Intermediate COCOMO model, Boehm uses 15 more predictor variables called Cost drivers are needed to calibrate the nominal cost of a project to the actual project environment. The values are assigned to each cost driver according to the characteristics of the specific software project. These numerical values correspond to the 15 cost drivers are multiplied to get effort adjustment factor i.e. EAF. This EAF value can be multiplied to the effort computed by Walston-Felix effort estimation model to get optimum effort value. Value of EAF ranges from 0.9 to 1.4.

EAF = 0.9 to 1.4

According to the proposed model, we get optimum effort value by multiplying EAF with effort computed by Walston-Felix model instead of multiplying EAF with effort calculated by COCOMO Intermediate model.

$E_{Proposed} = E_{W-F} * EAF$

B. Data Analysis

Tis section is classified into two phases.

Phase I : – Estimation of standard models using genetic algorithm approach. Here, GA Tool of Matlab is used for the implementation of SEL model, Walston-Felix model, COCOMO basic-organic mode, COCOMO basic semidetached mode, COCOMO basic embedded mode, COCOMO intermediate organic mode, COCOMO intermediate semidetached mode, COCOMO intermediate embedded mode. Here, comparison is made between actual effort and effort calculated using genetic algorithm for all the above mentioned models. Table I shows the data.

Phase II : – Estimation of proposed model using genetic algorithm.

In this section, comparison is made between effort calculated using COCOMO intermediate parameter with effort adjustment effort and effort of proposed model i.e. calculated by Walston-Felix model with effort adjustment factors.

$E_{I-E} = E * EAF$ .... (1)

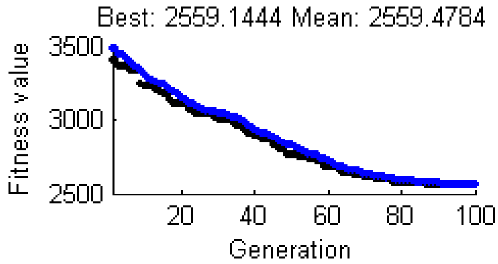
$E_{Proposed}=E_{W-F}*EAF$ ....(2)

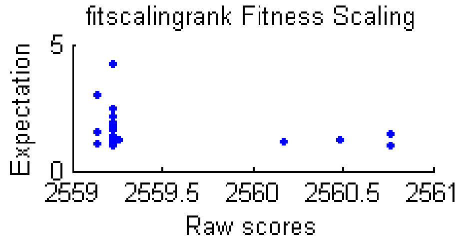
Here, also we compute actual effort for eq. (1) and (2) as well as effort using genetic algorithm on eq. (1) and (2) are also calculated. Table-II represents the variations. It shows that proposed model provides optimum value of effort by simple calculations as well as by applying genetic algorithm approach.

IV. Experimental Results

Simulation Results – SIZE (L)= 400

A. Intermediate-Embedded





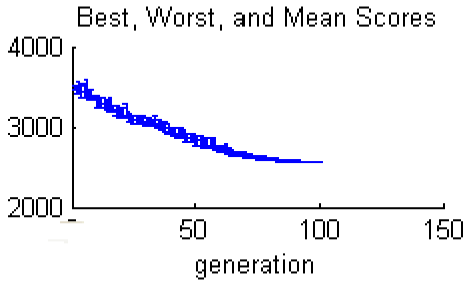


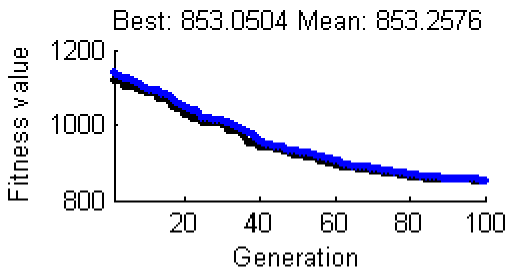
Table 1

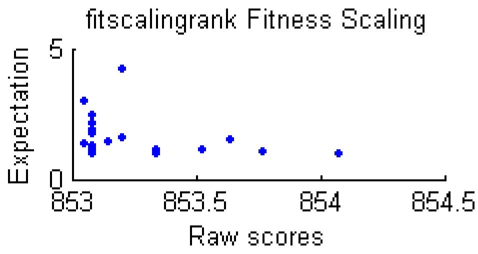
S.No.	Model	L(Size)	Equation	Actual Effort	GA Effort
1	SEL Model	400	$E_{SEL}=1.4* (L)^{0.93}$	368.1661	257.299
2	Walston-Felix	400	$E_{W-F}=5.2* (L)^{0.91}$	1213.0479	899.1549
3	Basic-organic	40	$E_{B-O}=2.4* (L)^{1.05}$	115.4448	51.3607
4	Basic-semidetached	250	$E_{B-S} =3.0* (L)^{1.12}$	1454.8361	876.7732
5	Basic-embedded	400	$E_{B-O} =3.6* (L)^{1.20}$	4772.8138	3279.1143
6	Intermediate-organic	40	$E_{I-O} = 3.2* (L)^{1.05} * EAF$	143.8904	46.0347
7	Inter.- semide-tached	250	$E_{I-S} = 3.0* (L)^{1.12} * EAF$	1359.9808	1094.7614
8	Inter.-embedded	400	$E_{I-E} = 2.8* (L)^{1.20} * EAF$	3470.1538	2559.4784

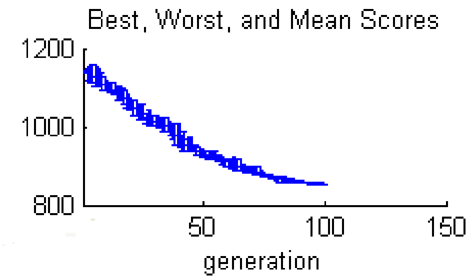
Table 2

S.No.	Model	L(Size)	Equation	Actual Effort	GA Effort
1	Inter-Embedded	400	$E_{I-E} = 2.8* (L)^{1.20} * EAF$	3470.1538	2559.4784
2	Proposed Model	400	$E_{Proposed} = 5.2* (L)^{0.91} * EAF$	1133.9571	853.0504

B. Proposed Model







V. Conclusion

This work explores the inter-relationship among different dimensions of software projects namely models, project size and effort. In this work we have proposed the calculation of optimum effort. The future scope of this work is to calculate magnitude of relative error (i.e. MRE) or absolute percentage error.

References

[1] Yeong-Seok Seo, Kyung-A Yoon, Doo-Hwan Bae, *"An Empirical analysis of software effort estimation with outlier elimination"*, Proceedings of the 4th International workshop on Predictor models in Software Engineering © ACM 2008 (ISBN: 978-1-60558-036-4/08).

[2] Pichai Jodpimai, Peraphon, Advanced Virtual and Intelligent Computing Centre, *"Analysis of Effort Estimation based on Software Project Models"*, IEEE 2009.

[3] Parvinder S. Sandhu, Manisha Prashar, Pourush Bassi and Atul Bisht, *"A model for estimation of efforts in development of software systems"*,World Academy of Science and Technology 56-2009.

[4] Alaa F Sheta, *"Estimation of the COCOMO model parameters using genetic algorithm for NASA Software Projects"*, Journal of Computer Science 2006 ISSN: 1549-3636, 2006 Science Publications.

[5] Andreea Fanea and Laura Diosan, *"Components execution order using Genetic Algorithm"*, STUDIA UNIV. BABES BOLYAI, INFORMATICA, 2005.

[6] Westley Weimer, ThanhVu Nguyen, Claire Le Goues and Stephanie Forrest, *"Automatically Finding Patches Using Genetic Programming"*, ICSE'09, IEEE.

[7] Mark Harman, *"The Current State and Future of Search Based Software Engineering"*, Future of Software Engineering (FOSE'07) 2007, IEEE.

[8] Indrajit Bhattacharyya, *"Vector GA - A Novel Enhancement of Genetic Algorithms for Efficient Multi-variable or Multi-dimensional Search"*, November 2009 Volume 34 Number 6, ACM.

[9] Anne Martens, *"Optimising Multiple Quality Criteria of Service-Oriented Software Architectures"*, QUASOSS'09, August 25, 2009, ACM Publications.

[10] Yujie Yao and Haopeng Chen, *"QoS-aware Service Composition Using NSGA"*, ICIS 2009, November 24-26, 2009 Seoul, Korea 2009, ACM 978-1-60558-710.

[11] Hsin-yi Jiang, Carl Chang, *"Deriving Evaluation Metrics for Applicability of Genetic Algorithms to Optimization Problems"*, GECCO'08, ACM 978-1-60558-130.

Kavita Choudhary completed her M.Tech (IT) from Guru Gobind Singh Indraprastha University, Delhi, India in 2010 and B.E. in Computer Science and Engineering from JECRC, Jaipur, India with Honours in 2005. Presently, she is working as Assistant Professor in ITM University, Gurgaon, India. She has five years of experience in teaching. She has published and presented many papers in national and international conferences. Her research area is Analysis of Effort Estimation using Genetic Algorithm.