Time of Development and Fitness Analysis of Modified COCOMOII Model for Software Projects

Sivakumar D, Sureshkumar C

Abstract: In the management of software projects, software cost estimate is very important. The absence of precise and premature estimates is a significant source of the inability of many software initiatives. Nevertheless, it remains a challenge for the software industry to correctly estimate the time and the cost of development. It is used to predict how much effort and time the project needs. The precision of the technology proposal usually depends on the estimated time and the strength standard of the template. In this research analysis the time of development and model fitness are analyzed as a major contribution. The optimized coefficient values for A, B, C & D is used to estimate the time of development and model fitness. This experimental analysis proved that the time of development and model fitness analysis shows the great significance in the estimation and their influence in software cost estimation.

Index Terms: Software cost estimation, Time of development, Fitness value, Modified COCOMOII, Coefficient optimization.

I. INTRODUCTION

In real scenario for the software development team as well as customers engaged in the project, precise software cost estimates are important. Software cost estimation is the estimate of the effort, time plan and level of personnel involved in developing software projects [1]. In software project management it is an important task that will reduce the risk. There are many parameters that affect the assessment of the undertaking and thus many methods of assessment [2], [3]. During the past ten years, numerous efforts have been undertaken to develop principles to estimate cost in terms of total efforts in a software unit such as man-months and timing for months [4].

The demands for increasingly larger and more complex software systems have stimulated these efforts, and the realization that large-scale software production projects often come with cost and schedule, which can increasingly take on the software producer. There is growing interest in developing large scale software projects [5], [6]. The ability to characterize an item is measuring the duration of enhancement and the necessary departments are gradually becoming an arduous job. It is because the instruments and methods required by software become more complicated and the contemporary software becomes increasingly expanded in order for software development to be built up and maintained, and the quality objectives are needed to create software systems in a fast-moving evolutionary algorithm situation that leads to a variety of challenging situation [7].

Due to the intangible design of the software itself, estimates of software costs are hard. In addition, during its development the software usually tends to change [8], [9]. This change might necessitate unexpected changes to the estimate. A great deal of research effort was therefore made to get better the precision of the software estimate and a lot of methods for estimating software costs were created [10], [11]. Certain methods are based on algorithms, while others are based on expert opinions.

The software engineer Barry Boehm launched the Constructive Cost Model (COCOMO) in 1981. COCOMO is an algorithmic cost estimation software model which uses a fundamental regression method with parameters derivative beginning past project information and present project properties [12]. The coefficients involved COCOMO have four parameters: A, B, C and D by examining a big amount of software projects [13], [14]. Boehm determined the characteristics of these parameters. However, these parameters could vary from one organization to another. The parameters must be adjusted to achieve the best possible results [15]. The methods of meta-heuristic algorithms are employed to optimize the parameters in order to achieve optimized solutions. Genetic algorithms (GA) are one of the significant and potent optimization algorithms. It is heuristic natural algorithms which are applied for ruling correct and fairly accurate solutions [16], [17]. GA is focused on a number of feasible alternatives to the issue, which are improving iteratively. The creation of forecast systems and wide-ranging data processing [18] is associated with the majority of GA apps. The objective of the present study is to tone COCOMO coefficients, especially COCOMO-II Coefficients of the post Architecture model using the online GA Data Series and some new data collected from the Project.

The reaming part of this paper is prearranged as follows: Segment II literature review that provides insight on previous works. Segment III presents modified COCOMO II model and functionalities. Segment IV describes coefficient optimization with GA. Segment V presents the experiment results and discussions to evaluate the proposed optimization. Segment VI provides the conclusion and areas of the future work.

II. LITERATURE REVIEW

Patil et al.[19] Developed hybrid technique for optimizing NN's weights in combination with Principle Component Analysis (PCA) used for mapping precise input indicators along with NN inputs. The ANN proposal is used to translate input information and to increase the
imprecision of the features of the product development item. Attarzadeh et al.[20]. The proposed model provides an increase of 8.36% prediction of estimation accuracy compared to COCOMO II.

In the year 2014 M. Madheswaran, et al.[21] have focused on the development of a model of software effort estimation centered on artificial neural networks. The main intend of this project was to frame to improve the perfection of the COCOMO model. They used neural networks with multiple layers to provide accommodation for the representation and its parameters to calculate the development effort of software. The complex system network is sophisticated by the use of background spreading algorithm through the processing and comparison of the prediction of the network with the actual effort of a number of training models. The networks have been trained and tested using the COCOMO data set and the outcome shows that the proposed model generates more precise inference than that of the COCOMO model.

The expansion and changes made in the COCOMO II model is provided by Vu. Nguyen (2014) [22] following evaluation of different popular cost estimated software maintenance models. He has shown that there are weaknesses in the different existing models regarding the limited choice of data and the limited range of maintenance activities. In order to evaluate and assess software maintenance projects, the author has evaluated and presented an extended version of the COCOMO II models. To develop the assessment model, the author has employed regression method. In the case of organizations, when data are not adequate to calibrate various evaluation models, the proposed model can be used. This expanded model of COCOMO II also takes the deleted code SLOC metric into its size metric.

In the year 2012, Syed Ali Abbas, et.al. [23] provides baselines on the development of the COCOMO suite model improvements and distinctions between the derivatives to scientists, software organizations and professionals. The authors have stated that over the years, different software cost estimation processes have been supported through a variety of models and that COCOMO (Constructive Cost Model) is considered to be one of the best schedule estimates. This paper discusses an overview of the COCOMO suite model and provides the appropriate environment for their use, concept behind their development, and procedures for their development.

Parida et al. [24] Analyzed designs which are normally used for estimating the magnitude and effort estimation of the back-end portion. In addition to the development of the front end, development of the back end is also crucial. The price of the back-end portion is determined on both sides by the volume of each data item and by the organizational dimension of the database. The experimental results showed how the effort estimate for the background part has to be carried out by means of ER diagrams and how the productivity of any business software can be increased in calculating the actual cost of the background part. M. Pauline et., al [25] has developed an improved system which is designed to minimize effort by improving adjustments to functional sizing techniques. This paper presents fuzzy methods for the classification of quality models as a basis. Empirical validation of multipliers of COCOMO II software development work is analyzed and ratings are defined for cost drivers.

N. Shivakumar et., al [26] examined the attempt of the algorithmic procedure and non-algorithmic procedure and demonstrated that the neuroscience-based integrated estimate is more effective for the assessment procedure than the algorithmic technique. In different procedures the achievement of the estimate is dependent on the precision and stabilization of the technique. In comparison to the methods Doty, Baied and Halstead, they implemented the fuzzy method into existing databases. Finally, the minimum rate of error by Fuzzy was shown in comparison with other methods by experimental results.

Thamarai et., al [27] used the Differential Evolution technique, a minimum mistake percentage is determined in comparison with current techniques in the suggested technique, in order to pick the appropriate design from a number of historical activities which are comparable to a fresh venture.

The proposal of four effort metrics for the software database or background, which can be derived in line with the complexity level of the ER diagnosis, was made by Samaresh Mishra et al. [28].

A model for estimating activity was suggested with the assistance of an ER diagram by Samaresh Mishra et al. [29]. The model takes into account the complexity of the ER graph of the software background in order to evaluate the effort. The metrics used in the model are the relation sizes, the size of the entity and semantic integrity limitations.

### III. MODIFIED COCOMO II MODEL

The software project development exertion is premeditated by means of the subsequent mathematical equation in the post architecture modelling method COCOMO II

\[
\text{Effort} = A S^E \prod_{j=1}^{n} EM_j
\]

(1)

Where,

‘A’ is the constant multiplication

‘S’ is the software project size

‘E’ is an exponential variable for factors of scale and economy related to the magnitude of the software project

‘EM’ is the multiplier of efforts

The exponential variable ‘E’ is dictated through measuring of predefined scale factors (SF) and summing them by using the associated mathematical equation

\[
E = B + 0.01 \sum_{j=1}^{5} SF_j
\]

(2)

The time of development (TDEV) gets from the effort, as indicated by the accompanying equation

\[
TDEV = C(\text{Effort})^F
\]

(3)
The latest calibration approach shows that the ‘C’ multiplier is equal to 3.67 and that the ‘F’ ratio is similarly measured in the scale form of the associated mathematical statement

\[ F = D + 0.02 \sum_{j=1}^{S} SF_j \] (4)

\[ = D + 0.2(E - B) \] (5)

When each variable and multiplier is provided with their nominal values, mathematical statement is provided for the effort and the time schedule as

\[ Effort = 2.94S^{1.1} \] (6)

Time schedule,

\[ TDEV = 3.67Effort^{3.18} \] (7)

The COCOMO II method with algorithmic assessment techniques is evident and strong.

IV. COEFFICIENT OPTIMIZATION WITH GA

In this research analysis the technical genetic algorithm is used as an optimisation method and the detailed step-by-step process is given below.

There are four model variables, verbalized by means of four genes, and the whole number is null to nine. The first gene is regarded null in this strategy and the three remaining genes are seen as dividing elements [30].

Algorithm: The Pseudo code for Genetic Algorithm

1: Set the required Parameters
2: Select the proper encoding methods
3: Create the initial population
4: WHILE i < IterationMax and FitnessBest < FitnessMax then
5: Complete the Fitness estimate
6: Carry out the Selection procedure
7: Carry out the Crossover procedure
8: Carry out the Mutation process
9: End while
10: Based on maximum fitness value Decode the individual
11: revisit the finest solution

1. Create the individual arbitrary
2. For each undertaking ‘j’ in the dataset, the project effort will be shown using optimized models coefficients of an individual ‘i’.
3. Estimate the project fitness ‘j’ for every individual fitness ‘i’ through one of the subsequent equations

\[ FP_{ij} = \frac{|TDEV_e - TDEV_a|}{TDEV_a} \] (8)

\[ FP_{ij} = \frac{|ASPE_e - ESPE_a|}{ASPE_a} \] (9)

‘i’ - is the quantity of individual
‘j’ - is the quantity of project

\[ TDEV_a^j \] - is an definite project ‘j’ development duration

\[ TDEV_e^j \] - is the estimated project development ‘i’

\[ ASPE_e^j \] - is definite effort of software project

\[ ESPE_a^j \] - is the anticipated or predicted effort of software project, for every entity ‘i’ in jth project

By understanding the legitimate effort for software projects is greater than the estimated or anticipated effort in the project, then fitness has an optimistic importance and if the actual effort in the project is less than that predicted or estimated, fitness has a negative value [31].

4. Every candidate fitness value is ascertained as the regular assessment of all software projects; meticulous fitness of a candidate is calculated from through the step 2 and step 3. The fitness convenience importance anticipated to be minimized [32].

\[ F_i = \frac{1}{n} \sum_{j=1}^{n} F_{ij} \] (10)

‘F’ – represents the fitness
‘i’ – is the candidate numeric representation
‘j’ – is the project number
‘n’ – represents the number of projects taken

5. The situation to stop the iteration specifies when the algorithm should be completed. An alternative option is the finest (minimal) and median survival of a population.

6. This experiment uses the tournament selection process as a method for selection.

7. A uniform crossover approach is used to generate fresh people through people chosen in a preceding phase.

8. The mutant mode of intervention is the voluntary decision of none or tiny amount of people. The likelihood of choosing the person should be small is obviously around 10%. Uniform crossover approach is used to produce fresh people through people chosen in the earlier stage for each person, random scheme of the mutating genes.

9. The fresh population is established with prominent people chosen by step 6 between parents and children.

V. RESULTS AND DISCUSSION

All exertion was conducted in all the way through the methodology chosen for this research study. The updated genetic algorithm parameters, developmental technology, are used to improve the results in the COCOMO II after architecture model which include a consistent crossover.
approach, a choice of competition and a speed of transition of 0.15 estimated at 250, population age 50 and mutation rate 0.01. The highest outcome is achieved in countless phases and an outcome range is obtained which gives the overwhelming person the answer with the transcendent fitness importance.

Best fittest value among all available individual value is, 2 7 9 8 | 0 8 1 3 | 3 5 1 8 | 0 2 9 7

On this basis, the resulting optimized COCOMO II coefficients for the post-architectural model are evaluated and have accompanying qualities, a=2.798, b=0.813, c=3.518, d=0.297. Elderly COCOMO II post architectural model coefficients have the accompanying qualities a=2.94; b=0.91; c=3.67; d=0.28. With the assistance of the optimized coefficients of COCOMO II, the mean relative estimation value or best fitness estimation of the dataset record is 3.79. It is 1.42 times below the average comparative value of the present coefficient COCOMO II post-architecture model, which is equivalent to 5.99.

Table I. Time of development (TDEV) estimated for the project with modified COCOMO II with the optimized coefficient model for new project data

<table>
<thead>
<tr>
<th>S.No</th>
<th>Project ID</th>
<th>KLOC</th>
<th>Actual TDEV</th>
<th>Modified COCOMO II</th>
<th>Modified COCOMO II With Optimized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P1</td>
<td>23</td>
<td>19.2</td>
<td>12.1</td>
<td>13.5</td>
</tr>
<tr>
<td>2</td>
<td>P2</td>
<td>98</td>
<td>57</td>
<td>43.4</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>P3</td>
<td>10</td>
<td>3.4</td>
<td>1.4</td>
<td>2.1</td>
</tr>
<tr>
<td>4</td>
<td>P4</td>
<td>130</td>
<td>96.4</td>
<td>87</td>
<td>89</td>
</tr>
<tr>
<td>5</td>
<td>P5</td>
<td>79</td>
<td>20.7</td>
<td>12</td>
<td>15.3</td>
</tr>
<tr>
<td>6</td>
<td>P6</td>
<td>86</td>
<td>23</td>
<td>13.4</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>P7</td>
<td>45</td>
<td>14.3</td>
<td>9</td>
<td>11.2</td>
</tr>
</tbody>
</table>

In this regard the TDEV estimated for the project with project ID P3 is 2.8 and it is closer to actual TDEV 3.4 of the same project. The TDEV estimated for the project with project ID P6 is 21.8 whereas the actual TDEV is 23. These type data are considered as the proven evident in the enhanced estimation and it helps the project administrator to maintain the price of the software project in the specified limit.

Table II. Time of development (TDEV) in modified COCOMO II with optimized coefficient model for new project data

<table>
<thead>
<tr>
<th>S.No</th>
<th>Project ID</th>
<th>KLOC</th>
<th>Actual TDEV</th>
<th>COCOMO II</th>
<th>Modified COCOMO II</th>
<th>Modified COCOMO II With Optimized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P1</td>
<td>23</td>
<td>19.2</td>
<td>12.1</td>
<td>13.5</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>P2</td>
<td>98</td>
<td>57</td>
<td>43.4</td>
<td>46</td>
<td>82.4</td>
</tr>
<tr>
<td>3</td>
<td>P3</td>
<td>10</td>
<td>3.4</td>
<td>1.4</td>
<td>2.1</td>
<td>2.8</td>
</tr>
<tr>
<td>4</td>
<td>P4</td>
<td>130</td>
<td>96.4</td>
<td>87</td>
<td>89</td>
<td>93.2</td>
</tr>
<tr>
<td>5</td>
<td>P5</td>
<td>79</td>
<td>20.7</td>
<td>12</td>
<td>15.3</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>P6</td>
<td>86</td>
<td>23</td>
<td>13.4</td>
<td>16</td>
<td>21.8</td>
</tr>
<tr>
<td>7</td>
<td>P7</td>
<td>45</td>
<td>14.3</td>
<td>9</td>
<td>11.2</td>
<td>13</td>
</tr>
</tbody>
</table>

Individual project fitness is evaluated and is listed in Table III. The statistical data output of this experiment clearly indicates that the modified COCOMO II with optimized coefficient model produced very less fitness value when compared to the other models considered for the analysis of this experiment.

Table III. Individual project fitness based on TDEV for COCOMO II dataset

<table>
<thead>
<tr>
<th>S.No</th>
<th>Project ID</th>
<th>COCOMO II</th>
<th>Modified COCOMO II</th>
<th>Modified COCOMO II With Optimized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>23</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>44</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>26</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>33</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>55</td>
<td>33</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>58</td>
<td>44</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>42</td>
<td>37</td>
<td>29</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>47</td>
<td>49</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>41</td>
<td>45</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>51</td>
<td>44</td>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>54</td>
<td>53</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>56</td>
<td>37</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>61</td>
<td>39</td>
<td>34</td>
<td>16</td>
</tr>
</tbody>
</table>

Model Fitness 41.46 31.07 10.76

If the fitness value is less the model prediction is closer to the actual. Very less model fitness value 10.76 is achieved in modified COCOMO II with the optimized coefficient approach, whereas the same value in the COCOMO II model was 41.46. This result shows the clear improvement in the estimation of TDEV.
Table IV. Individual project fitness based on TDEV for new project data

<table>
<thead>
<tr>
<th>S.No</th>
<th>Project ID</th>
<th>Estimated Fitness Value</th>
<th>COCOMO II</th>
<th>Modified COCOMO II</th>
<th>Modified COCOMO II With Optimized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P1</td>
<td>37</td>
<td>30</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P2</td>
<td>24</td>
<td>19</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>P3</td>
<td>59</td>
<td>38</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>P4</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>P5</td>
<td>42</td>
<td>26</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>P6</td>
<td>42</td>
<td>30</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>P7</td>
<td>37</td>
<td>22</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Model Fitness: 35.86 24.71 14.14

The new project data collected for testing the method is used for evaluating the individual project fitness and the model fitness. Both the fitness has very less value when compared the same with another model. The model fitness value 14.14 is measured using modified COCOMO II with optimized coefficient model and is the least value when compared to the remaining two fitness value 35.86 and 24.71. While considering the individual project fitness also the project with project id P7 has the fitness value of 9 in the modified COCOMO II model with optimized coefficients where as the COCOMO II model fitness value is 37 and the modified COCOMO II model fitness value is 22.

In view of better mechanism in representing results the pictorial representation is a more powerful tool when compared to tabular form with statistical values. The Fig.2 clearly indicates that the modified COCOMO II with optimized coefficient model produced enhanced fitness with less difference in actual and estimated TDEV values. The project id 9 has very less value of fitness this indicates that the project id 9 provides very high accuracy in its prediction of effort and TDEV. Almost all testing data shows that the fitness value is very less in coefficient optimized modified COCOMO II model when compared to all other model fitness values.

This modified COCOMO II with optimized coefficient model again proved with the new project data considered, that is this model provides a schedule of estimation which is very closer to the actual schedule. Even though it is not the exact schedule of estimation, it is a better result when compared to the COCOMO II based schedule estimation.
Fig. 4 Fitness representation of modified COCOMO II with optimized coefficient model for new project data

Fig. 4 provides the graphical representation of individual project fitness in estimating the effort and time of development of software project. The fitness value of all the projects has very less value and hence this modified COCOMO II with optimized coefficient model produced a better fitness in estimating the effort and the schedule.

VI. CONCLUSION

This experimental assessment is based on the COCOMO II dataset and current design information, the updated COCOMO II with an optimized coefficient model. In comparison with the consequence calculated using present COCOMO II PA model equations, the assessed growth and fitness moment is capable to produce high assessment ability. The findings show that the energy and moment measured using the optimized coefficient with the genetic algorithm method in the vast bulk of instances gives superior outcomes. This practice is further extended to optimize mobile software forecast precision.

REFERENCES

31. F. Leung, H. Lam, S. Ling, “Tuning of the Structure and Parameters of a Neural Network

Retrieval Number: B3304078219/19/BEIESP
DOI: 10.35940/ijrte.B3304.078219
Published By: Blue Eyes Intelligence Engineering & Sciences Publication

3737


AUTHORS PROFILE

Dr Sivakumar D received his Ph.D (CSE) in 2016, ME (CSE) in 2007, M.Tech (Power Systems) in 2005, and BE (EEE) in 1999. He is currently working as an Associate Professor in the Department of Computer Science and Engineering, RajaRajeswari College of Engineering, Bengaluru. He also worked as a Software Trainer. He is having more than 15 years of Experience in the field of engineering education. He is a reviewer for the International Journal of Technological Research & Innovative Solutions (IJTRIS). His areas of interest in Teaching, Learning & Research are: Software Engineering | Computer Networks | Big Data Analytics | Soft Computing: Neural Network, Genetic Algorithm | Data Base Management | Programming and Data Structures | Java | C | C++ | MATLAB. He is Life Time Member of ISTE.

Mr. Sureshkumar C received his Master of Engineering in Computer science and engineering from Annamalai University, Chidambaram. He is presently working as an assistant professor in the department of Information Technology, KGiSL Institute of Technology, Coimbatore. He has 13 years of industry, research and teaching experience. He has published papers in International conferences (Paper indexed in IEEE Xplore digital library) and Scopus indexed journal. His research interests includes | Data analytics | Software Engineering | Software Testing | Distributed Computing | Cloud Computing.