Grenze International Journal of Engineering and Technology, Jul Issue

# Calibrating Intermediate COCOMO Model using Memetic Algorithm

Aditya Verma[1] and Preeti Malik[2]
[1]Graphic Era Hill University/CSE, Dehradun, India
Email: adityaverma.gehu@gmail.com
[2]Graphic Era Deemed to be University, Dehradun, India
Email: Preetishivach2009@gmail.com

***Abstract*—In past few years, numerous analysts and industries related to software given critical consideration on the evaluation of programming effort. The fundamental task in SDLC is software cost estimation. It is also considered as most important part for managing project cost, time, and quality. Numerous researchers provided various cost estimating models, for example, algorithmic models, non-algorithmic models, AI related models, statistical based and nature motivated models. In this work, memetic calculation (MA) is utilized for enhancing the coefficients of intermediate COCOMO model to improve precision for predicting the effort of the project. The procedure is applied on COCOMO NASA programming dataset.**

***Index Terms*— COCOMO, Intermediate COCOMO, Memetic Algorithm, Software Cost Estimation.**

## I. Introduction

The most essential part in software evolution is estimation of cost of the product. Software cost estimation acts as a major part in receiving or refusal of its development. Proper cost estimation ensures programming to complete on financial plan and time. It is most important to understand the cost by proper estimation for project management, improved quality. In software development process for calculating/predicting effort is software cost estimation. Number of estimation models has been developed in last few decades [1, 2, 3]. Expert judgment, Algorithmic strategy, non-Algorithmic strategy, and Analogy based strategy are the techniques proposed by various researchers. LOC and function points are metrics on which most of the software estimation models rely upon for size measure. The precision of the size estimation particularly focuses on the prediction of software costs.

COCOMO [5] is a recursive cost estimation model provided by a product engineer Barry Boehm. The model consists of four parameters a, b, c, and d. To get the ideal result, we have to modify the coefficients, and for this purpose we have to apply meta-heuristic techniques to get enhanced values. The different heuristic techniques utilized to get the enhanced results for effort estimation are Genetic algorithm [6, 7, 8, 9], Differential Evolution [15], Genetic programming [10], PSO [11] and many more.

This research work presents a short investigation of MA which is used when upgrading intermediate COCOMO parameters. The applied technique is implemented using the NASA programming dataset. The paper is organized as, section II. defines knowledge of the models and heuristic algorithm used to optimize parameters, section III. Gives the related work done in context of the method used, section IV. Depicts the procedure

*Grenze ID: 01.GIJET.7.2.5_1*

performed, section V. depicts the dataset used and the outcomes acquired and section VI. finally shows the conclusion of the paper.

## II. BACKGROUND

### *A. The Cost Constructive Model (COCOMO MODEL)*

The most popular algorithmic strategy for effort prediction is COCOMO model. In 1981, Barry Boehm built the model for estimating programming cost [13]. With the help of equations, we can figure out the cost of the projects. The model was established by 63 complete projects from different domain. Basic COCOMO, Intermediate COCOMO and Detailed COCOMO are three types in which COCOMO model was divided. Project size is most significant parameter while estimating effort of the software. It is in the form of source line of code (SLOC) or thousand-line of code (KLOC). The projects are classified into three categories to calculate the cost of the product depending upon size of the project as represented in table I. Equation 1 provides the formula for calculating effort of project.

$$effort = a * (kloc)^{b} \text{ Person month} \qquad (1)$$

### *B. Intermediate COCOMO Model*

This model is a part of COCOMO family which was utilized for optimizing the parameters. According to basic model we can calculate the project effort by the help of line of code (LOC) only, but there are more factors that contribute for calculating effort. Reliability, Experience, potentiality, complexity, performance etc. are other critical factors associated with the project effort. These characteristics are defined as cost drivers [16]. We have to rank these characteristics from 1 to 5 on the basis of project functionality. Besides the project size there are 15 other criteria used to measure project effort [14]. The (2) is used for calculating the effort of intermediate COCOMO [4]. a and b are the coefficients and E is known as effort multiplier which depends on the type of programming and characteristics of a project in equation 2. Emi is constructed by merging procedure, product including improvement qualities. Intermediate model is also partitioned into three projects Organic, Semi-detached, and Embedded. The types of the projects with their coefficients are shown in Table II.

$$effort = \left(a * (kloc)^{b}\right) em_i \text{ Person month} \qquad (2)$$

### *C. Memetic Algorithm*

MA is inspired by meme-documentation from Dawkins. This is a unit of data that repeats itself as the commercial thoughts of individuals [17]. MA attaches the handiness of GA with a couple of heuristic methods like hill climbing, Tabu search etc. MA utilizes the duality of GA and cultural advancements. MA is a meta-heuristic algorithm which combines global as well as local search. On account of neighborhood search capacities, which are improved MA can discover solutions with high quality. Two basic hybridization strategies rely on the "Baldwin impact" and "Lamarckism" thought [17]. Approach of Baldwin stated that local enhancements can interfaces and allow the local searches to change the individual fitness, but genotype remain unaltered [17]. As demonstrated by Lamarckism, characteristics got by individual during its lifetime may wind up heritable attributes. Genotype and fitness of any individual can be modified during neighbourhood enhancement stages that are shown according to this process. The principle goal of MA is it exploits the population dependent on global search methods.

TABLE I. COEFFICIENTS OF BASIC MODEL

| Coefficients | a | b | c | d |
|---|---|---|---|---|
| Organic mode | 2.4 | 1.05 | 2.5 | 0.38 |
| Semidetached mode | 3.0 | 1.12 | 2.5 | 0.35 |
| Embedded Mode | 3.6 | 1.20 | 2.5 | 0.32 |

TABLE II.COEFFICIENTS OF INTERMEDIATE COCOMO MODEL

| Coefficients | a | b | c | d |
|---|---|---|---|---|
| Organic mode | 3.2 | 1.05 | 2.5 | 0.38 |
| Semidetached mode | 3.0 | 1.12 | 2.5 | 0.35 |
| Embedded Mode | 2.8 | 1.20 | 2.5 | 0.32 |

The basic steps followed by MA are defined as:
Algorithm: Memetic Algorithm (MA)
Begin
Initialization (Population)
Evaluation (population)

```
Continue Until (termination criteria does not satisfied) do
    {
        Selection (populations)
       Crossover (population)
      Mutation (population)
      Local Search applied (hill climbing)
  }
Select individuals for next generations;
End do
End
```

*D. Hill Climbing*

A method of mathematical optimization which is applied to move towards [17] the solution is known as Hill climbing technique. This method is a part of local search algorithm. This approach is iterative in nature, begins with initial answer to question which tries to find a better solution by making improvements to original solution. Hill climbing is local search estimation where a loop is run to achieve the best individual to make posterity. In the event that produces posterity is better than its parent, then interchange the values [17].

Pseudo code of local search Hill climbing algorithm is shown below: [17]

1. Start with parent which is the best-known solution at this point.
2. Continue with the loop until end condition does not fulfill.
   2.1. Generate new solution that is present nearby to the parent.
   2.2. Check if new individual is superior to parent, then makes the new solution parent and terminates if loop as well as while loop also.
   2.3. At last step return the best solution or individual.

## III. Related Work

In past, several works have been done to optimize the cost estimation models. GA, PSO, Genetic programming etc. are some methods that were utilized to optimize the coefficients of estimation strategies. This segment is devoted to writing (literature) in a similar exploration zone.

Sachan et al. used GA for enhancing the coefficients of COCOMO. The proposed algorithm uses NASA dataset for the calculation which consists of 18 projects. Manhattan distance is the difference between effort and predicted effort which is used to calculate the fitness value of projects. As indicated by acquired outcome the technique utilized gives better outcome when contrasted with effort generated by NASA. [8]

Dhiman and Diwaker modify the COCOMO II configuration using GA. TURKISH and INDUSTRY dataset. The comparison between two results represents that GA based coefficients produce better outcomes when contrasted with current COCOMO II PA. Mostly the outcomes acquired utilizing GA coefficients give better result than the actual effort. [8]

Galinina et al. applies GA to refine the coefficients of COCOMO model. The dataset on which this technique is applied consist of some records for organic base model and some for semi-detached base model. For training and testing purpose dataset was divided, in which first few records are kept for training and rest are used for testing. During comparison with current COCOMO it was seen that optimized model provides better result for organic mode using GA. No such difference is seen for semi-detached base optimized model [7].

Gupta and Sharma applied Bat Algorithm for optimizing Intermediate COCOMO model coefficients. The parameters of all three modes were optimized using NASA 63 dataset. According to the obtained result it was seen that in most of the cases the optimized results are better than the previous results. [14]

Mittal et al. proposed new model based on fuzzy logic. As the size metric of the software project triangular fuzzy number is used. Estimated effort can be optimized by making some modifications to the constants of the current model. While making comparison between the models it was observed that proposed model is better than the other models. [12]

Sabbagh Jafari and Ziaaddini applied harmony search algorithm which is a meta-heuristic algorithm for optimizing COCOMO model. The purpose of this method is to decrease MMRE due which coefficients of the model are optimized. Optimizing the coefficients will increase accuracy in estimation process. NASA dataset was used for testing the results obtained after modifying coefficients for all three modes of COCOMO. Accuracy in effort estimation is increased by the help of Harmony search algorithm. [16]

A new hybrid mutation operator was provided by Manju Sharma after studying various types of mutation operator. In this mutation operator she utilized local search to control the behaviour of the operator [18]. She has contrasted conventional GA and MA comprising of hybrid mutation operator. During this experiment, the researcher has used Roulette wheel and Rank selection operators. The investigations have been coordinated using TSP oliva30 and Eil51 benchmark issues and execution is passed on using MATLAB. Results demonstrate the improvement of MA with hybrid mutation over existing GA with basic mutation operators.

Poonam Garg performed analysis on MA dependent on hill climbing. In this work GA has been used for the cryptanalysis on simple information encryption [19]. For discovering number of keys present the algorithm MA is better than GA which is the conclusion obtained by the researcher at the end.

Ajay Jaiswal and Meena Sharma introduced a new estimation tool known as expert estimator which is a web-based estimation tool [20]. The knowledge of java programming and the eclipse framework was used by the author to develop expert estimator. Risk and cost of the software was estimated by this tool. At every stage of the project development, model can estimate risk from one step to another step during project progress.

Moscato [21] depicted evolutionary calculation utilizing local search algorithm by pursuit that presented MA. Local search and GA are the required techniques to perform MA. GA is utilized for exploration and for exploitation purpose local search is performed. In this way, the balance is maintained between intensification and diversification.

A way to deal with Open Shop Scheduling Problem utilizing Memetic Algorithm is proposed in [22]. Memetic Algorithm contains the blend of GA and to limit the makespan, three nearby pursuit calculations is applied to the Open Shop Problems. Outcome is created by taking care of 20 Open Shop Scheduling Problem with the proposed MA and contrasted with PSO outcomes founded by Sha and Hsu [23].

## IV. Proposed Method

MA was used to optimize intermediate model coefficients. By applying MA with intermediate COCOMO we have received values of a and b that are optimized. While using these new coefficients a and b it was observed that the effort, we estimate is more accurate relative to the effort achieved using traditional values of a and b. The procedure that is followed while applying MA with COCOMO is:

1. Start with initial population.
2. Effort was calculated along with the coefficients of all individuals for all the projects.
3. Fitness function was calculated. Manhattan distance was used for fitness function calculation for individuals, i fitness will be for the project j. Equation 3. represents the formula for MD [7], other measurement parameters are Root Mean Square, MMRE etc.

$$MD = \sum_{i=1 to\_n} \left( | effort_i - estimated\_effort_i \right) \qquad (3)$$

i = coefficient of individual
j = project number
effort = Actual NASA effort
estimated effort = effort calculated

4. Calculate average fitness function for each chromosome with the aid of (4).

$$fitness_i = \frac{1}{n} * \left( \sum_{j=1}^{n} fitness_{ij} \right) \qquad (4)$$

Where,
n= total count of project

5. On chromosomes perform selection step based upon minimum fitness function.
6. In this step as a local search algorithm we applied hill climbing technique.
7. To generate new population one point crossover was performed at this stage.
8. Check termination state if reach then end the cycle, otherwise repeat the steps with newly created population

TABLE III: SETUP FOR MA PARAMETERS

| **Operators** | **Types that were used** |
|---|---|
| Selection | Rank Selection |
| Crossover | 1-point Crossover |
| Mutation | Not Applied |
| Size of population | 20 |
| Maximum number of Generations | 100 |

## V. Experimental Evaluation And Result

Existing Intermediate model coefficients are a=3.20 and B=1.05. The coefficients received after applying meta-heuristic algorithm are a=3.91 and b=1.05. Using enhanced coefficients, the estimated value of efforts for applied technique is shown in table 4. Fig 1 provides the comparison among effort evaluated by NASA, traditional COCOMO model and Optimized model. The MD of strategy related to MA is 11.8 which is better than intermediate COCOMO as appeared (represented) in Fig 2.

$$effort = \left(3.91 * (kloc)^{1.05}\right) em_i \tag{5}$$

TABLE IV: Evaluated values of effort for Intermediate COCOMO and MA

| S.No. | Size | Intermediate COCOMO Effort | NASA Real Effort | MA |
|---|---|---|---|---|
| 1. | 16.3 | 58.136 | 82 | 71.05 |
| 2. | 35.5 | 131.64 | 192 | 155.9 |
| 3. | 5.5 | 16.83 | 18 | 20.566 |
| 4. | 19.7 | 64 | 60 | 78.51 |
| 5. | 14 | 44.909 | 60 | 54.85 |
| 6. | 10.4 | 32.868 | 50 | 40.148 |
| 7. | 8 | 32.665 | 42 | 39.92 |
| 8. | 48.5 | 182.682 | 239 | 223.26 |
| 9. | 29.5 | 98 | 120 | 119.97 |
| 10. | 13 | 54.386 | 60 | 66.47 |
| 11. | 32.6 | 120.37 | 170 | 147.11 |

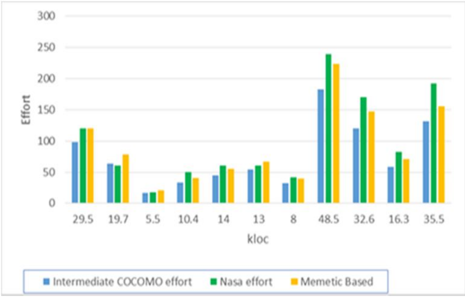


Figure. 1 Graph representing comparison among NASA Intermediate COCOMO and MA effort

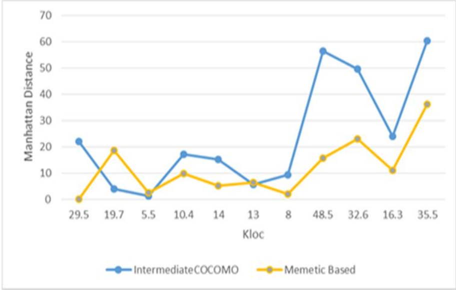


Figure. 2 MD Graph for Proposed calculation and Intermediate COCOMO

## VI. Conclusions

The focused area of the research is to apply MA for upgrading the coefficients of intermediate COCOMO. Organic type projects are used for this research work. To get the required outcomes well known COCOMO NASA dataset helps a lot. The proposed algorithm was tested, and the outcomes acquired were contrasted with the old coefficients of intermediate COCOMO model. It was observed while contrasting that new parameter give preferable results over the old parameters.

## References

[1] Tannu, Yogesh kumar, "Comparative Analysis of Different Software Cost Estimation Methods", International Journal of Computer Science and Mobile Computing vol. 3, no. 6, pp. 547-557, June 2014.

[2] R.Tripathi, Dr. P.K Rai, "Comparative Study of Software Cost estimation Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no.1, January 2016.

[3] S.Shekhar, U. Kumar, "Review of Various Software Cost Estimation Techniques", International Journal of ComputerApplications, vol. 141, no.11, pp.0975-8887, May 2016.

[4] P. Pandey, "Analysis of the Techniques for Software Cost Estimation", Third International Conference on Advanced Computing & Communication Technologies, pp. 16 - 19, 2012.

[5] Boehm, "B.W. Software Engineering Economics", Prentice-Hall, Engle Wood Cliffs, N.J. 1981.

[6] R. k. Sachan, A. Nigam, A. Singh, S. Singh, M. Choudhary, A. Tiwari and D. S. Kushwaha, "Optimizing Basic COCOMO Model using Simplified Genetic Algorithm", Twelfth International Multi-Conference on Information Processing Procedia ,pp.492 – 498,2016.

[7] A. Galinina, O. Burceva, and S. Parshutin, "The Optimization of COCOMO Model Coefficients Using Genetic Algorithms," Information Technology and Management Science, vol. 15, no. 1, pp. 45–51, 2012.

[8] A.Dhiman and C.Diwaker, "Optimization of COCOMO II Effort Estimation using Genetic Algorithm," American International Journal of Research in Science, Technology, Engineering & Mathematics, pp-208-2012, 2013.

[9] M. Alajlan and N. Tagoug, "Optimization of COCOMO-II Model for Effort and Development Time Estimation using Genetic Algorithms," Proc. Of The International Conference on Communications, Computer Science and Information Technology.

[10] A. F. Sheta and A. Al-Afeef, "Software Effort Estimation for Nasa Projects Using Genetic Programming", Journal of Intelligent Computing, vol. 1(3), p. 147, (2010).

[11] C. Hari and P. P. Reddy, "A Fine Parameter Tuning For Cocomo 81 Software Effort Estimation Using Particle Swarm Optimization", Journal of Software Engineering, vol. 5(1), pp. 38–48, (2011).

[12] A. Mittal, K. Parkash, H. Mittal, "Software Cost Estimation Using Fuzzy Logic", ACM SIGSOFT Software Engineering Notes, Vol. 35. No. 1, pp. 1, November 2010.

[13] P. Suri and P. Ranjan, "Comparative Analysis of Software Effort Estimation Techniques", International Journal of Computer Applications, vol. 48(21), pp. 12–19, (2012).

[14] N. Gupta and K. Sharma, "Optimizing intermediate COCOMO model using BAT algorithm," in 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 1649–1653.

[15] S. Aljahdali and A. F. Sheta, "Software effort estimation by tuning COOCMO model parameters using differential evolution," in Computer Systems and Applications (AICCSA), 2010 IEEE/ACS International Conference on, 2010, pp. 1-6.

[16] S. M. Sabbagh Jafari, F. Ziaaddini, "Optimization of Software Cost Estimation using Harmony Search Algorithm" 1st Conference on Swarm Intelligence and Evolutionary Computation (CSIEC2016), Higher Education Complex of Bam, Iran,pp. 131 – 135, 2016.

[17] R. Kumar, S. Tyagi and M. Sharma, "Memetic Algorithm Hybridization of Hill Climbing with Selection Operator", International Journal of Soft Computing and Engineering, Vol.3, No. 2, 2013.

[18] M.Sharma, "Memetic Algorithm with Hybrid mutation Operator", International Journal of Computer Science and Mobile Computing, Vol. 3, No. 1, 2014.

[19] P. Garg, "A Comparison between Memetic algorithm and Genetic algorithm for the cryptanalysis of Simplified Data Encryption Standard algorithm", International Journal of Network Security & Its Application, Vol.1, No.1, 2009.

[20] A. Jaiswal, M.Sharma, "An Expert Estimator Tool to Estimate Project Cost and Risk with early stage of function points", International Journal of Software Engineering and Applications, Vol. 3, No. 5, 2012.

[21] P. Moscato, M.G. Norman, "A Memetic approach for the travelling salesman problem Implementation of a computational ecology for combinatorial optimization on message-passing system", Parallel Computing and Transputer Applications, pp. 177-186, 1992.

[22] B.E. Engin, M.O. sumbul, "A Memetic Algorithm to solve the Open Shop Scheduling Problem", IEEE, 2015.

[23] D.Y. Sha, Cheng-Yu Hsu, "A new particle swarm optimization for the open shop scheduling problem", Computers & Operations Research 35 (2008) 3243 – 3261.