Novel Approach to Reduce MRE Value for Software Effort Estimation

Anshu Sharma
Research Scholar
Ideal institute of Technology, Ghaziabad,
Uttar Pradesh

Mr. Hariom Sharma
M.Tech Coordinator
Ideal institute of Technology, Ghaziabad,
Uttar Pradesh

Abstract - The efforts are to be estimated before beginning with the development of software. This technique is known as effort estimation. Various models are proposed to measure the efforts accurately. In this work a hybrid formula used that depends upon the values of cost drivers. The generated result will estimate closer efforts than COCOMO. The tool used for the implementation is MATLAB. IVR dataset is being used after taking authentication from the website. This report presented effort estimation models, COCOMO, COCOMO 2, Bailey-Baisly model, Halstead, Bee Colony Optimization Algorithm and the proposed one. In this, there are 47 projects with defined values of cost drivers. The efforts of each model are being calculated and their MMRE is calculated. The result shows that the hybrid model is derived by using TLBO algorithm provides more accurate results than the other estimation models.

Keywords: TLBO,COCOMO,MRE

I. INTRODUCTION

In the area of software development, the main challenging issue is effective software project estimation. If there is no proper, reliable estimation provided in the software development, there will be no proper arrangement as well as control of the project. Even when all the important factors are considered, the software business is not useful in determining the project estimated. It doesn't utilize estimates for improving the development of software [1]. When a project is underestimated the effects such as under-staffing, under-scoping the quality assurance effort, and missing the deadlines resulting in loss of credibility are seen. When the applications try of estimate the estimation, there are equal chances of over-estimation of a project which can also result in providing loss to the applications. If a project is given more number of resources than it actually requires, the resources will be utilized by it. The cost of the product increases due to this reason, which will further result in deploying the estimations stated for the software. This will also result in limiting the resources of the company as they are wasted [2]. The first step towards achieving an effective estimate is to provide an exact estimation for the software size which is be proposed. Along with this, formal descriptions of the needs for project estimation scope, the source of data might start. After receiving the size estimation of a product, it becomes very easy to estimate the effort of it. When the software development lifecycle of a project is defined only then the conversion from software size to total project effort estimation is possible [3]. Further the designing, develop and test of the software are defined for project development. In addition to coding of the software there is much more to the software development project. The third important step in software development project is the determination of project schedule from the effort estimate. The number of people working on a project, the type of work they will do, the starting time and ending time of the project are the factors that are to be involved here. The total estimate of the project is also dependent in a way on the amount of cost an organization allocates to it. In some of the fields the cost allocation is not done at all, and its adjustment is done by increasing the labor costs per hour [4]. The overall costs for the software development is estimated by the software development project manage accordingly. Software cost estimation model is a backhanded measure, which is utilized by software personnel to predict the cost of a project. The development of software product shifts depending upon the earth in which it is being developed. For projects with familiar environment it is anything but difficult to predict the cost of the project. The estimation model is valuable for exchange off between the developer and customer. Organization can understand of what is achievable and deliverable to the customer. Cost drivers are those critical features which affect the project [5]. The cost drivers may vary the cost of building a project. The most important cost driver is size of the project. Size of the project is measured in Kilo lines of code (KLOC). Function points are the empirical measurement to measure size of the project. Constructive Cost model was developed by Barry W Boehm in 1981. It is an algorithmic cost model. Algorithmic cost model is developed taking into account relating the present project to previous projects. It depends on historical information. COCOMO depends on size of the project. The size of the project may vary depending upon the function points. The development time is a very important factor for the Putnam model. The decrement of development time can lead to enormous increment in the person-months that are needed for development [6]. The capacity for estimating accurately the size of software which is to be developed is to

be known in the PUTNAM model. Halsted's metrics is used for the purpose of estimating the program volume, complexity of the method as well as the effort of program. The program complexity is defined as the total number of operators and operands used in a program. Teaching-Learning-Based Optimization (TLBO) is a new meta-heuristic optimization algorithm. This optimization technique is used for the continuous non-linear large-scale optimization. Similar to the teaching and learning basic relation, this technology has its methodology. There is an impact on the optimization procedure according to the performances of learners. The teachers are a great influence on the outputs given in a class. It is basically a structure based system [7]. There are various practical optimization problems to which this algorithm has been used for various real world applications which involve the mechanical design, planar steel frames, welding, etc. When in connection with the continuous function optimization, there were various methods executed by the TLBO algorithm. This also has various disadvantages. There are limited benchmarks on which the algorithm is thus suitable to perform. There are limited function to which have dimensions that are also limited.

II. LITERATURE REVIEW

Peyman Khazaei, et.al (2016) proposed that in the day-ahead power systems scheduling, system operators formulated and solved the unit commitment (UC) problem to determine ON/OFF status and power dispatch of the producing units [8]. Reenactment examines depicted the convergence speed and exhaustion of the proposed TLBO algorithm to solve the UC problem. The proposed algorithm was compared with several existing methods. The numerical results showed the effectiveness of the proposed TLBO which was compared with several well known evolutionary algorithms, i.e., DP, PSO, and SLF A. The TLBO not just gave a solution bring down operating costs, additionally had a lower computation time. In addition, adequate spinning reserve was given to alleviate the effect of rapid load/generation changes because of unexpected disturbances. **Yu-Huei Cheng (2016)** proposed that numerous single nucleotide polymorphisms (SNPs) for complex genetic diseases were genotyped by polymerase chain response restriction piece length polymorphism (PCR-RFLP). A feasible PCR-RFLP primer match was to be designed. Also there was a need to discover accessible restriction enzymes which could perceive the target SNP for PCR experimental purposes [9]. In order to enhance the mining work for restricted enzymes which are imposed on the new REBASE the SNP-RELPing method is utilized. For enhancing the TLBOMPD, the renewed PCR-RFLP which is extracted from numerous SNPs is utilized. The experimental results which were achieved were made to compare with the GAMPD results. This would be helpful for enhancing the reliability of this proposed technique. **Chalotra et.al [2015]** introduced that the target of momentum examination was applying Bee Colony Optimization (BCO) meta-heuristic way to deal with enhance the parameters of COCOMO model for enhancing software cost estimation. BCO methodology a "bottom-up" way to deal with modeling where uncommon sorts of artificial agents are made by similarity with bees which are utilized to take care of complex combinatorial optimization issues [10]. The proposed model validation was done utilizing Interactive Voice Response software venture dataset of an organization. The BCO approach creates different partial arrangements and best arrangement is chosen based on Mean Magnitude of Relative Error. The results acquired demonstrate that the proposed BCO based model can enhance the precision of cost estimation furthermore beat different models. **David L. Gonzalez-Alvarez, et.al (2014)** proposed that proteins are molecules that shape the mass of living creatures. The proteins exist in dissociated shapes like amino-acids [11]. Various learning stages were used to enhance the knowledge (quality) of the population-based evolutionary algorithm. This algorithm is characterized as a group of individuals. So, adequate quality solutions could be achieved by providing enhancements to already existing techniques. Six instances composed of various related protein sequences were gathered from the PROSITE database. These were used to evaluate the performance of the proposed technique. The proposed technique helped in making predictions and also helped in improvement of quality of the solutions which were found by biological tools. **M. Ramakrishna Murty, et.al (2014)** proposed a reliable, accurate and robust optimization technique plot for global optimization over constant spaces. A variation of the teaching factor TF in traditional TLBO algorithm is proposed in this paper. Numerical difficulties were observed in the constrained as well as unconstrained optimizations. Problems such as multi-modality, dimensionality and differentiability were found [12]. To solve the global numerical optimization problems, the authors proposed variation of teaching factor in the TLBO algorithm. It was seen through the results evaluated that the proposed algorithm has enhanced the performance in terms of convergence which also includes the some changes in the teaching factor algorithm. For managing the complex numerical optimization issues, the TLBO algorithm involved with new teaching factor has proved to be efficient. **E.B. Elanchezhian, et.al (2014)** proposed in this paper [13], that for the purpose of the accuracy of Economic Dispatch (ED) there are certain algorithms proposed. The algorithms involve the cubic cost models as well as the optimization algorithms. The aims of this paper involve the efficiency and achievability. The TLBO convergence properties are analyzed and variety of results is achieved. According to the results, the efficiency of the algorithm is improved such that the solution has the required characteristics and also is robust in comparison to the already existing various techniques. Other

various observations are depicted in this paper. Not only the composite ED model is to be achieved, but also the various objectives are to be fulfilled. With the help of TLBO algorithm, the issues were solved and there are no dependencies of the previous techniques here.

III. RESEARCH METHODOLOGY

The TLBO algorithm is the algorithm which is used for the optimization. In this research TLBO algorithm is applied to reduce MRE value of the COCOMO Model by estimating predicted efforts more accurately. The TLBO algorithm comprises of methods which help each individual to take in something different and to enhance himself. The base of this algorithm has be derived from a normal teacher-earner methodology of a classroom. The TLBO algorithm holds the basics of traditional learning methods that are seen in a teacher and a learner. There are two essential methods of learning involved in it. The first is to learn through the teacher. It is also known as a teacher phase. The second is the learning that is done through interaction with different learners. This is known as the learner phase. TLBO is a population based algorithm. The population comprises of the gatherings of students (learners).

Following are the steps which followed to implement proposed work:

1. In the first step the dataset is loaded which of ivr in which 21 projects are given and correspond to each project their KLOC value is given.
2. In the second step the KLOC value is considered as the initial population which is input to learner phase
3. The learner phase also take input the desired error value and their the loop is executed means until minimum error is achieved by considering the KLOC value
4. In the last phase values of one iteration is compared with the previous iteration and when minimum error is achieved as that phase algorithm will exit the loop and display final result in terms of MMRE value.

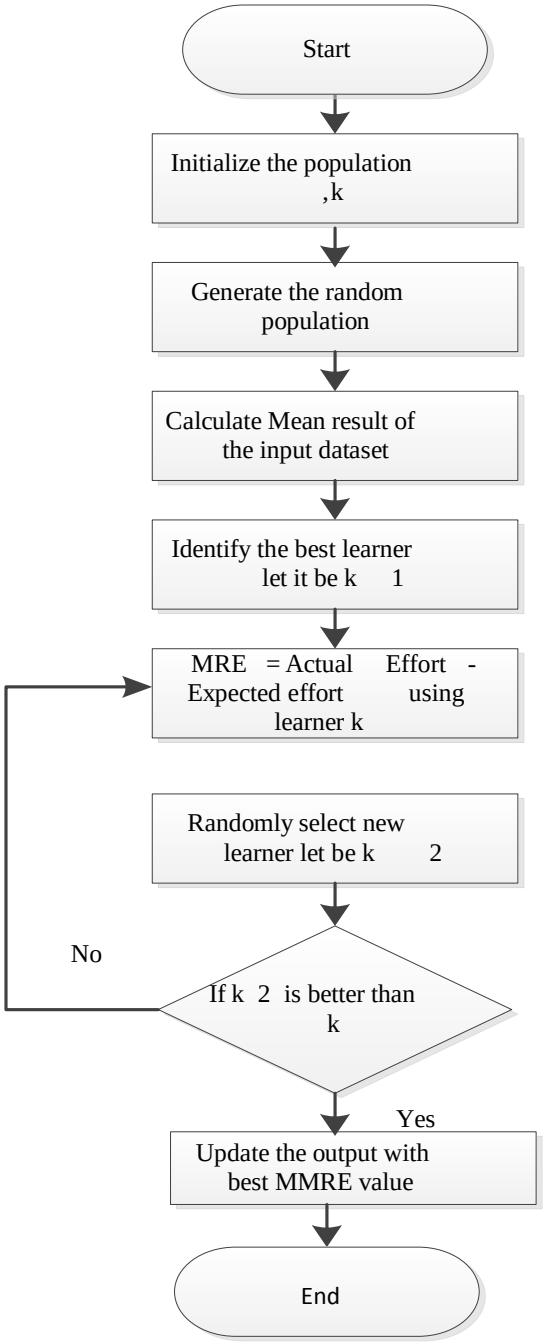

Fig. 1: Flow Chart

IV. EXPERIMENTAL RESULTS

The proposed work has been implemented in MATLAB and its results are evaluated as shown below.

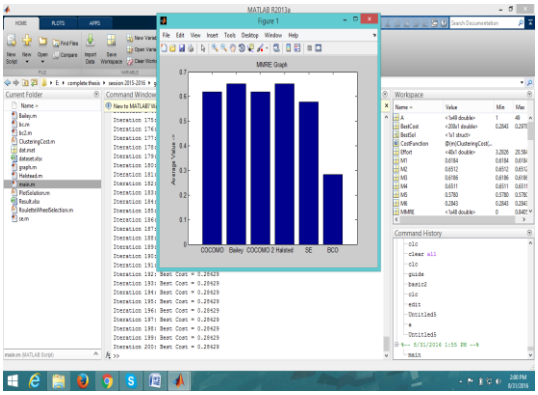

Fig 2: Comparison of all models with bee colony

As shown in figure 2, the various models like COCOMO, Bailey and bee colony optimization algorithms are compared in terms of MRE Value. The BEE colony has least MRE value as compared to other models

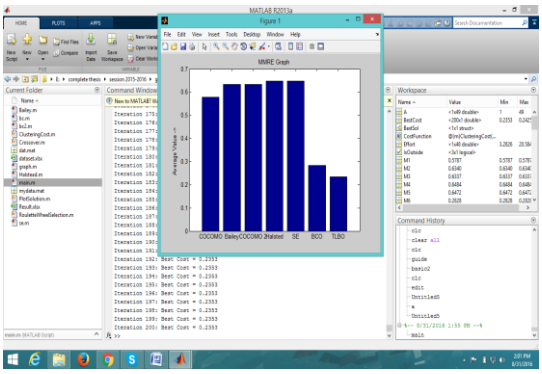

Fig 3: TLBO algorithm performance

As shown in figure 3, the various models like COCOMO, Bailey and bee colony optimization and TLBO algorithms are compared in terms of MRE Value. The TLBO colony has least MRE value as compared to other models

S No	Project Type	COCOMO	COCOMO-II	Bailey	Halstead	SE	DE	TLBO
1	Organic	13.45	21.16	45.34	49.80	34.16	60.15	14.56
2	Semi-Detached	20.12	27.28	51.34	43.56	30.12	65.46	17.19
3	Embedded	34.67	47.78	56.78	62.35	59.87	54.56	45.90

Table 1: Comparison of Variance

S No	Project Type	COCOMO	COCOMO-II	Bailey	Halstead	SE	DE	TLBO
1	Organic	0.79	0.81	0.74	0.83	0.87	0.72	0.68
2	Semi-Detached	0.77	0.74	0.71	0.68	0.70	0.64	0.61
3	Embedded	0.91	0.89	0.88	0.92	0.89	0.74	0.71

Table 2: Comparison of MMRE

S No	Project Type	COCOMO	COCOMO-II	Bailey	Halstead	SE	DE	TLBO
1	Organic	0.34	0.42	0.44	0.52	0.65	0.72	0.78
2	Semi-Detached	0.39	0.41	0.56	0.59	0.62	0.67	0.79
3	Embedded	0.41	0.54	0.59	0.65	0.69	0.74	0.77

Table 3: Comparison of Prediction

V. CONCLUSION

The effort estimation is the technique which will estimate the efforts for the software development. In the recent times various algorithms has been proposed which analyze the efforts for development. These models are COCOMO, COCOMO-II, Bailey, Dotty etc. In this work, we are working on COCOMO model which is based on KLOC values. In this work, we are using IVR dataset in which 47 projects are considered and KLOC of each project is given in the dataset. To reduce MRE value TLBO algorithm is applied which is based on learner and teaching phase. The proposed and existing algorithms are implemented in MATLAB and it is been analyzed that MRE value is reduced with the use of TLBO algorithm.

VI. REFERENCES

[1]. Hitesh Kumar Sharma, Ravi Tomar, Ankur Dumka and M. S. Aswal," OpenECOCOMO: The Algorithms and Implementation of Extended COst COnstructive MOdel (E-COCOMO)", 2015, NGCT, 978-1-4673-6809

[2]. Iman Attarzadeh, Siew Hock Ow," Improving Estimation Accuracy of the COCOMO II Using an Adaptive Fuzzy Logic Model", 2011, IEEE, 978-1-4244-7317

[3]. Jun Liu, Zheng Xu, Jianzhong Qiao, Shukuan Lin," A Defect Prediction Model for Software Based on Service Oriented Architecture using EXPERT COCOMO", 2009, IEEE, 978-1-4244-2723

[4]. Petr Musilek, Witold Pedrycz, Nan Sun, Giancarlo Succi," On the Sensitivity of COCOMO II Software Cost Estimation Model", 2002, IEEE, 0-7695-1339
[5]. Mahima Mishra, Saru Dhir," Appraisement of different Software Estimation Models: A Rumination and Contradistinction", 2015, ICSCTI, 978-1-4673-6792-9
[6]. Sumeet Kaur Sehra, Jasneet Kaur, Yadwinder Singh Brar, Navdeep Kaur," Analysis of Data Mining Techniques for Software Effort Estimation", 2014, IEEE, 978-1-4799-3187-3
[7]. Manoj Kumar Debnath, Nimai Charan Patel, Ranjan Kumar Mallick," Automatic Generation Control of a two area multi-unit interconnected power system with Proportional-Integral-Derivative controller with Filter(PIDF) optimized by TLBO algorithm", 2016, ICCPCT 978-1-5090-1277
[8]. Peyman Khazaei, Morteza Dabbaghjamanesh, Ali Kalantarzadeh, Hasan Mousavi," Applying the Modified TLBO Algorithm to Solve the Unit Commitment Problem", 2016, Springer International Publishing Switzerland, 7583026-34-54, Pp-45
[9]. Yu-Huei Cheng," A Novel Teaching-Learning-Based Optimization for Improved Mutagenic Primer Design in Mismatch PCR-RFLP SNP Genotyping", 2016, IEEE/ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, Vol. 13, No. 1
[10]. S. Chalotra, S. K. Sehra, Y. S. Brar, and N. Kaur, "Tuning of COCOMO Model Parameters by using Bee Colony Optimization," Indian J. Sci. Technol., vol. 8, no. 14, Jul. 2015.
[11]. David L. Gonzalez-Alvarez, Miguel A. Vega-Rodriguez and Alvaro Rubio-Largo," Finding Patterns in Protein Sequences by Using a Hybrid Multi-objective Teaching Learning Based Optimization Algorithm", 2014, IEEE/ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, Vol. X, No. Y
[12]. M. Ramakrishna Murty, J.V.R. Murthy, P.V.G.D. Prasad Reddy, Anima Naik, and Suresh Chandra Satapathy," Performance of Teaching Learning Based Optimization Algorithm with Various Teaching Factor Values for Solving Optimization Problems", 2014, Springer International Publishing Switzerland, 78-3-319-02931-3
[13]. E.B. Elanchezhian, S. Subramanian, S. Ganesan," Economic power dispatch with cubic cost models using teaching learning algorithm", 2014, IET Gener. Transm. Distrib., Vol. 8, Iss. 7, pp. 1187–1202