SURVEY ON SOFTWARE COST BY FUZZY LOGIC

Naveen Malik[1], Sandip Kumar Goyal[2], Vinisha Malik[3]
[1]PhD Research Scholar, [2]Professor, [3]PhD Research Scholar
Department of Computer Science & Engineering[1,2,3] MM University Mullana, Ambala, Haryana, India

[1]Naveenmalik317@gmail.com, [2]skgmmec@gmail.com, [3]Vinishamalik317@gmail.com

Abstract — *The growing application of software and resource constraints in software projects development need a more accurate estimate of the cost and effort because of the importance in program planning, coordinated scheduling and resource management including the number of programming's and software design using tools and modern methods of modeling. Effectively control of investment for software development is achieved by accurate cost estimation..The results show that all three membership functions vary by 1-2 % in the effort estimation where as if compared; COCOMO 2 Effort varies by 5-10 % from Fuzzy output. The traditional method consumes a lot of time and a lot of methods are non mathematical due to which the predicted results may be irrelevant.* ***Application/Improvements:*** *The work has been tested on 5 projects. When the parameters for triangular function are executed, values for embedded 1, semidetached 2 are enhanced as compare to organic, semi detached 1, embedded 2. After the evaluation of trapezoidal function, values of semidetached 1 and embedded 1 are more as compare to others. Same as before, the values for embedded 1 and semidetached 2 are more in Bell membership function. Slight improvement has been seen in organic value. Similarly, for COCOMO II the results have been obtained.*

In the evaluation of software projects is commonly used deterministic method. But software world is totally different from the linear variables and nowadays for performance and estimation should be used nonlinear and non-probabilistic methods. In this paper, we have studied the SCE Using Fuzzy Logic (FL) and we have compared it with COCOMO model. Results of investigations show that FL is a performance model for SCE.

KEYWORDS

Software Cost Estimation, COCOMO, Fuzzy Logic

1. INTRODUCTION

The SCE is an issue that Longley engaged software project manager's mind. In general it can be stated that one of the important tasks for managers in the field of software development is an attempt to find correlations between the impressive resources in software development projects to accurate estimation. Also in some cases software projects development companies due to incorrect estimates lead to the loss of resources and the lack from useful human power. In software development process, failure is inevitable and in the form of the costs imposed on the managers who are directly involved in the cycle related to software projects [1, 2, 3]. This problem is considered as a negative factor in the software production and development. Therefore, it should be taken advantage of advanced techniques in order to avoid the risk of failure of software projects to obtain accurate estimates.

Most of software projects are executed in dynamic and complex environments, so that lack confidence and cost is their intrinsic properties [4]. This uncertainty caused that mostly software projects do not achieve notable success in predetermined goals. This led to problems such as lack of project's development, reduction in performance and dissatisfaction in customers. Currently, there are several factors that lead to the decrease of estimation accuracy which can point to the followings [5, 6]:

- Managers who lack appropriate experience
- Development teams that do not have suitable experience with programming languages
- Customers who are continually demand developments and changes.

To estimate the cost and effort of software projects, different algorithmic models such as COCOMO I [7], COCOMO II [8], SLIM [9, 10] and FP [11] has been used. The purpose of SCE is to increase the probability of project success and this is done through the identification and evaluation of systematic effort and cost. Accurate estimation in the initial phase of the project is vital for project managers and software Development Companies in order to be successful in projects and reasonable in costs. In recent years, artificial intelligence models in combination with algorithmic models had good performances in SCE [12, 13, and 14]. Algorithmic models such as COCOMO are dependent on a number of cost factors that these factors are used to calculate the non-functional characteristics process [7, 8]. These models try to formulation the connection between features of efforts and the size of the project. These models based on criteria such as quantities like number of Line of Code (LOC) or the degree of effective factors in estimation. The size of projects can be gauged by these units and then the amount of required effort and costs are calculated. The general trend of SCE includes the following assumptions [5, 15]:

- The use of previous experience to determine factors in which the cost and effort estimation can be effective for new projects, for example, the number of people on the project. It should be noted that the size of the LOC varies in software projects and always has a strong relationship with effort and cost.
- Evaluation of the accuracy of estimation models

Because, estimation is a complex and uncertain phenomenon, due to uncertainty and ambiguity in determining factors of development and underdevelopment, the use of algorithmic models like COCOMO can be ineffective. FL is formed due to analyze of systems that the dependencies between variables is very complex [16]. The existence of such complexity in engineering and the other different sciences is common. An important link which connects things of this type is imprecision, ambiguous and uncertain nature of reality.

The structure of the paper is as follows: in Section 2, we will discuss about SCE and evaluation criteria; in Section 3, we will study the FL membership functions and their application in the SCE; in Section 4, we will discussion the FL membership functions in SCE and finally in Section 5, we will explain the conclusions and future works.

2. SOFTWARE COST ESTIMATION

Naturally, SCE for software projects includes coordination among all developmental activities, design, production monitoring, maintenance, etc. [1]. Accurate estimation of software project causes that internal and external processes and employee activities, efforts and costs to be coordinated. So, before design and implementation of software projects providing the model for them is essential and can be the most difficult tasks in software projects development. In the process of software projects production to reduce cost and schedule and probabilistic risks estimate must be taken to avoid project failure [3, 12]. The importance of SCE is more evident when we know each evaluation in the estimation of cost contains the positive and negative consequences and its balancing at any point of time is one of the most complex management issues. Evaluation of the estimates performance has always been the important and valuable category in software development projects. An issue that could has a significant impact on the future of software companies. So the more accurate the criterion of evaluation of software projects, the better estimate would be achieved. Table (1) shows the most important SCE criteria for software projects.

Table 1. SCE Evaluation Criteria

Evaluation Criteria	Description
$MRE = \frac{\lvert act_i - est_i \rvert}{act_i} \times 100$	Magnitude of Relative Error (MRE). MRE criteria value error of estimated for each of the projects compared to the actual model obtains.
$MMRE = \frac{1}{n} \sum_{i=1}^{n} MRE_i$	Mean Magnitude of Relative Error (MMRE). MMRE as the criteria error of the mean value of the project are considered.
$PRED(x) = \frac{1}{n} \times \sum_{i=1}^{n} \begin{cases} 1, & if MRE \le x \\ 0, & otherwise \end{cases}$	Percentage Relative Error Deviation (PRED). PRED criteria in order to use to estimate the accuracy of the models.
$MARE = \sum_{i=1}^{n} \left(\frac{\lvert act_i - est_i \rvert}{act_i} \times 100 \right) / n$	Mean Absolute Relative Error (MARE). MARE criteria in order to error of the mean values of the applied projects.
$VARE = Var\left[\frac{\lvert act_i - est_i \rvert}{act_i} \right] \times 100$	Variance Absolute Relative Error (VARE). VARE criteria in order to percent of variance to estimate the value of each project can be calculated.
$VAF = \left(1 - \frac{\mathrm{var}(act_i - est_i)}{\mathrm{var}(est_i)}\right) \times 100$	Variance Account For (VAF). VAF is used in the context of statistical models whose main purpose is the prediction of future outcomes on the basis of other related information.
$BRE = \frac{\lvert act_i - est_i \rvert}{\min(act_i, est_i)}$	Balance Relative Error (BRE). BRE criteria in order to accurately estimate the error rate has been used.

Models for SCE should be applied that have necessary performance in the evaluation of estimate criteria. The model which has a lower MRE is better than the model which has higher MRE. The model which has a lower MMRE is better than the model which has higher MMRE. The model which has a higher PRED is better than the model which has lower PRED. The model which has a lower MARE is better than the model which has higher MARE. The model which has a lower VARE is better than the model which has higher VARE. The model has a higher VAF is better than the model which has lower VAF. The model that has a lower BRE is better than the model which has higher BRE.

3. FUZZY LOGIC

FL theory was presented in 1965 because of the uncertainty in data and information and imprecision in the existence of vagueness [16]. FL is not a random or unlikely method and in fact, this method itself introduces a special system to deal with the ambiguous and non-deterministic situations. The essential characteristic of fuzzy theory is displaying uncertain data and also can be the operation and application of mathematical programming. Each fuzzy set can be shown with a membership function which represents the membership grade of element x in the reference set X to fuzzy set A. If the degree of membership of an element is set to be zero, that member is fully withdrawn from the set and if it will be equal to 1, that member is quite in the set. If the degree of membership of a member is between 0 and 1, this number represents the partial membership degree.

In this case fuzzy set A is shown according to equation (1).

$$A = \{(x, \mu_A(x)); \quad x \in X, \mu_A(x) \in [0,1]\} \qquad (1)$$

If a fuzzy set contains of discrete elements x_i, the fuzzy set A is shown according to equation (2).

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \ldots + \mu_A(x_n)/x_n = \sum_{x_i \in X} \mu_A(x_i)/x_i$$

And if the reference set will be continuous, the fuzzy set A is shown according to equation (3).

$$A = \int_{x_i \in X} \mu_A(x_n)/x_n \qquad (3)$$

Fuzzy sets can be defined and maintained its character based on the membership function. The most importance membership functions for fuzzy set includes [16, 17]:

Triangular Membership Function: This membership function with three parameters a, b and c, where a<b<c is defined according to equation (4). In the triangular membership function if the value of the property is greater than the center of the membership function, the center of the membership function must be transferred to the left in order to go further away of these characteristics. The upper limit of the membership functions should be moved to the left. Also, for the case that the value of the property is less than the center of the membership function, the membership function of the mean and the lower limit of membership function must be transferred to the right.

$$triangle(x \mid a,b,c) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ \dfrac{c-x}{c-b} & b \le x \le c \\ 0 & x \ge c \end{cases} = \max\left[\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right] \qquad (4)$$

Trapezoidal Membership Function: This function with the four parameters a, b, c, d where a<b<c<d is defined according to equation (5).

$$\begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x \le b \end{cases} \qquad x-a \quad d-x$$

$$trapezoid(x\,|\,a,b,c,d)=\begin{cases} 1 & b\le x\le c \\ \dfrac{d-x}{d-c} & c\le x\le d \\ 0 & x\ge d \end{cases} = \max[\min(\frac{}{b-a},1,\frac{}{d-c}),0] \qquad (5)$$

Gaussian Membership Function: This membership function with two parameters $\square$, c where $\square$ represents width and c represents the center of the membership function is defined According to equation (6).

$$gauss(x\,|\,\square,c)=e^{-0.5\left(\frac{x-c}{\square}\right)^2} \qquad (6)$$

Generalized Bell Membership Function: This membership functions with three parameters a, b, c where a is the width, b is the slope and c is the center of membership function. Bell membership function is shown according to equation (7).

$$gbell = (x \mid a, b, c) = \frac{1}{1 + \left|\frac{x - c}{a}\right|^{2b}} \qquad (7)$$

Sigmoidal Membership Function: Sigmoidal membership function is defined according to equation (8).

$$sig(x \mid, a, c) = \frac{1}{1 + e^{[-a(x-c)]}} \qquad (8)$$

In equation (8), a value the amount of slope at the point x = c handles. Sigmoidal membership function will open from left to right based on parameter's a. Therefore, this type of membership function is good to represent concepts such as “very good” or “very bad”.

3.1. Fuzzy Logic for Software Cost Estimation

In many engineering sciences lots of quality and quantity factors such as quality, price, flexibility, scalability and performance must be considered for decision-making. To do so would be to determine the factors and weights of fuzzy functions use and fuzzy numbers can be expressed in them. So FL tries to obtain convenient option for issues with estimates and decision making in environments with ambiguous and vague criteria. In recent years, FL has many applications in SCE due to the flexibility and high precision in the estimates. In this section, we review FL models and its applications and also the modeling results of proposed models that have been done by the researchers on the project software dataset.

The SCE is analyzed using fuzzy functions. Method of Triangular Membership Function, Trapezoidal Membership Function and Gaussian Membership Function are used for evaluating on the NASA93 dataset. The proposed method is combined with COCOMO II model. COCOMO II model includes 17 Effort Multipliers (EMs) and 5 Scale Factors (SFs). In Figure (1) is shown a hybrid model.

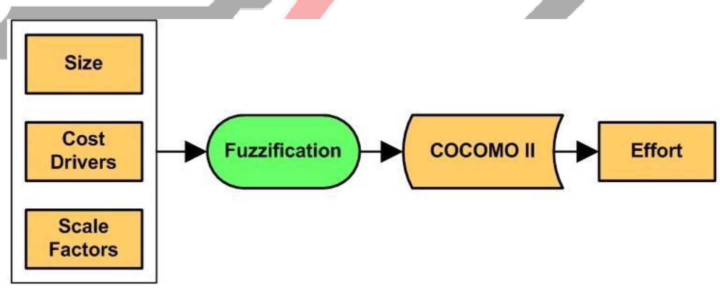

Figure 1. FL hybrid model and COCOMO II

Evaluation and results have been done on the 10 projects of NASA93 dataset. Their results show that fuzzy methods are more accurate in SCE and have less MRE error than the COCOMO II model. And also Gaussian Membership Function among Fuzzy functions has better performance than other models, and in most cases it has reduced MRE error's rate.

A new approach for the SCE has been proposed with a combination of Triangular Membership Function and COCOMO II model [19]. Evaluation has been conducted on 30 projects from NASA software projects. The hybrid model factors of EM, SF and KSLOC (Kilo Source Line of Code) are evaluated using fuzzification. In Figure (2) is shown the hybrid model.

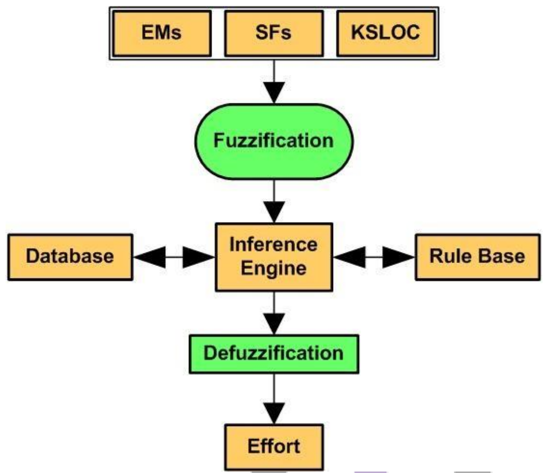

Figure 2. Hybrid Model COCOMO II and the Fuzzy Inference System

Hybrid model inputs consist of 17 EMs, 5 SFs and KSLOC. Inference Engine part is the process of formulating the mapping from a given input to an output using the FL. Rule Base part is the selection of fuzzy rules. In the Database, used membership functions are defined in the fuzzy rules. And the Inference Engine, which includes inference try to give a reasonable output with the help of rules. Defuzzification part converts a fuzzy set to a number. Experimental results show that the accuracy of MMRE and MRE errors in the hybrid model is high when compared with other models. Such that the value of MMRE error in a hybrid model is equal to 7.512, and the MRE error on average for the 30 projects is equal to %63.33. Also in Table (2) is shown the PRED error for the projects.

Table 2. Evaluation of MMRE, VAF and PRED Criteria

Criterias	Models [19]	
	COCOMO II	Hybrid
MMRE	11.003%	7.512%
VAF	95.86%	98.77%
PRED(25)	93.33%	96.33%
PRED(15)	63.33%	93.33%
PRED(10)	50%	80%
PRED(8)	40%	63.33%

FL-COCOMO II hybrid model has proposed based on the membership functions of Triangular, Trapezoidal, Gaussian, Generalized Bell and Sigmoidal then

than COCOMO model. In Table , fuzzy membership functions in terms of Accuracy, Transient Response and Oscillations are evaluated. As you can see the performance of each of the functions are different from each other. Based on investigations, we concluded that it should be considered the performance of the functions for estimate.

Evaluation of the Performance of Fuzzy Functions

Membership Function	Accuracy	Transient Response	Oscillations
Triangular	High	Normal	Normal
Trapezoidal	Medium	Normal	High
Gaussian	High	Normal	Normal
Generalized Bell	Medium	Normal	Normal
Sigmoidal	High	High	Normal

4. CONCLUSIONS AND FUTURE WORKS

In this paper, SCE was investigated using the FL and the capabilities of COCOMO model. Nowadays, the use of FL as a new approach in analyzing software engineering issues such as SCE has improved. One of the issues that helps the correct use of software and prevent the software projects from failure is the accurate estimate of cost. Therefore, this paper analyzes the FL in SCE using membership functions of Triangular, Trapezoidal, Gaussian, Generalized Bell and Sigmoidal. Functions tested and evaluated on a set of software projects dataset and showed that they have better performance than COCOMO model better. We hope in the future with presenting this paper give performed models for SCE using the fuzzy systems.

REFERENCES

[1] P.C. Pendharkar, "Probabilistic Estimation of Software Size and Effort", Expert Systems with Applications, Vol. 37, pp. 4435-4440, 2010.

[2] T.R. Benala, R. Mall, P. Srikavya, M.V. HariPriya, "Software Effort Estimation Using Data Mining Techniques", Advances in Intelligent Systems and Computing, Vol. 248, pp. 85-92, Springer, 2014.

[3] G. Sivanageswara Rao, Ch. V. Phani Krishna, K. Rajasekhara Rao, "Multi Objective Particle Swarm Optimization for Software Cost Estimation", Advances in Intelligent Systems and Computing, Vol. 248, pp. 125-132, Springer, 2014.

[4] A. Trendowicz, "Why Software Effort Estimation?", The Fraunhofer IESE Series on Software and Systems Engineering, pp. 3-7, Springer, 2013.

[5] W. Zhang, Y. Yang, Q. Wang, "A Study on Software Effort Prediction Using Machine Learning Techniques", Communications in Computer and Information Science, Vol. 275, pp. 1-15, Springer, 2013.

[6] V.P. Cot, B. Oligny, N. Rivard, "Software Metrics: an Overview of Recent Results", the Journal of Systems and Software, Vol. 8, pp. 121-131, 1988.

[7] A.J. Albrecht, J. Gaffney, “Software Function, Source Lines of Code, and Development Effort Prediction: a Software Science Validation”, IEEE Transactions on Software Engineering SE, Vol. 9, No. 6, pp. 639-648, 1983.

[8] F.S. Gharehchopogh, “Neural Networks Application in Software Cost Estimation: A Case Study”, 2011 International Symposium on Innovations in Intelligent Systems and Applications (INISTA 2011), pp. 69-73, IEEE, Istanbul, Turkey, 15-18 June 2011.

[9] Z.A. Khalifelu, F.S. Gharehchopogh, “Comparison and Evaluation Data Mining Techniques with Algorithmic Models in Software Cost Estimation”, Elsevier, Procedia-Technology Journal, ISSN: 2212-0173, Vol. 1, pp. 65-71, 2012.

[10] P. Pandey, “Analysis of the Techniques for Software Cost Estimation”, Rohtak, pp. 16-19, IEEE, 2013.

[11] L. A. Zadeh, “Fuzzy sets”, Information and Control, Vol. 8, pp. 338-353, 1965.

[12] G. J. Klir, B. Yuan, “Fuzzy Sets and Fuzzy Logic”, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[13] A. Malik, V. Pandey, A. Kaushik, “An Analysis of Fuzzy Approaches for COCOMO II”, International Journal Intelligent Systems and Applications, Vol.5, pp. 68-75, 2013.

[14] P. Jha, K.S. Patnaik, “Comparative Analysis of COCOMO81 using Various Fuzzy Membership Functions”, International Journal of Computer Applications (IJCA), Vol. 58, No. 14, pp. 220-27, 2012.

[15] R.R. Yager, D.P. Filev, “SLIDE: A Simple Adaptive Defuzzification Method”, IEEE transaction on Fuzzy Systems, Vol. 1, No. 1, pp. 69-78, February 1992.

[16] R. Agarwal, I. Awasthi, S. Sarwar, “An Empirical Validation of Software Cost Estimation Model Using Fuzzy Technique”, International Journal of Advances in Electrical and Electronics Engineering, Vol. 1, No. 2, pp.188-194, 2012.

[17] J.N.V.R Swarup Kumar, A. Mandala, M.V. Chaitanya, G.V.S.N.R.V. Prasad, “Fuzzy logic for Software Effort Estimation Using Polynomial Regression as Firing Interval”, International Journal Computer Technology Application, Vol. 2, No. 6, pp. 1843-1847, 2011.

[18] M. Kazemifard, A. Zaeri, N. Ghasem-Aghaee, M.A. Nematbakhsh, F. Mardukhi, “Fuzzy Emotional COCOMO II Software Cost Estimation (FECSCE) using Multi-Agent Systems”, Applied Soft Computing, Vol. 11, pp. 2260-2270, Elsevier B.V, 2011.