Effort Estimation in Software Cost Using Team Characteristics Based on Fuzzy Analogy Method – A Diverse Approach

S. Malathi[1] and S. Sridhar[2]

[1] Dept of CSE, Sathyabama University,
Chennai, Tamilnadu, India
malathi_raghu@hotmail.com
[2] Dept of CSE & IT, Sathyabama University,
Chennai, Tamilnadu, India
drssridhar@yahoo.com

Abstract. The dramatic increase in the scope of software cost estimation has paved way for the enhanced research to develop different methods for estimating the software effort. Estimation of effort in software cost based on Fuzzy Analogy is one of the most popular existing methods. Usually, only the project characteristics are considered for the effort estimation whereas the team characteristics also play a significant role. This paper presents a diverse approach where the features of team characteristics like joy and skill are considered in addition to the project features. The empirical results are validated with the historical datasets having both categorical and numerical data by considering hypothetical data of team characteristics. The outcome of this paper signifies that the usage of team characteristics improves the performance and accuracy of software effort estimation.

Keywords: cost estimation, fuzzy analogy, team characteristics, effort estimation, hypothetical data.

1 Introduction

A critical factor during the software development process [1] is the ability of estimating the effort accurately at the early stages of development. Moreover, it is noted that none of the existing methods have dealt with the team characteristics effectively. A project team is a social system where personal and co-operative characters play an important role to the success of the organization. The team members in a project have emotions such as joy and skill, which can be either positive or negative, [2] during the duration of the project. The positive feature is related to joy and negative feature is related to fear. However, the degree of joy is one of the salient features affecting the project features.

All estimation methods consider only project features for effort estimation. In this paper, a different approach based on Fuzzy Analogy method has been proposed

V.V. Das and P.M. El-Kafrawy (Eds.): SPIT 2012, LNICST 117, pp. 1–8, 2014.

incorporating the team member characters like joy and skill. The hypothetical data of team member characteristics is considered for this study to estimate the effort.

Section 2 deals with the milieu of study including analogy, fuzzy logic, personality and emotion while Section 3 depicts the proposed method. Section 4 evaluates the performance and Section 5 concludes the paper.

2 Related Work

2.1 Analogy

Analogy based estimation is the process of finding the similar projects among existing ones and assess the estimate for the target project in the incipient stage [3]. Many researchers accept the solution of analogy method since it is derived on the basis of human way of reasoning rather than arcane chains of layers. The drawback of analogy is the degree of similarity between the historic and the target project [4]. Genetic algorithm [5] is used in analogy as one of the search technique for the similarity measure and also overcome the complex optimization issues.

2.2 Fuzzy Method

The term fuzzy came into existence with the fuzzy sets [6] for dealing the linguistic variables. However, fuzzy method cannot overcome the imprecision and uncertainty problem in an effective manner while handling the categorical datasets [7]. Fuzzy integrated with analogy, fuzzy analogy, [8] was developed to overcome this problem. The performance of the effort is improved by using the feature subset algorithm based on fuzzy logic in analogy concept [9].

2.3 Personality and Emotion

Personality [10] and mood plays an important role in the human behaviour. This forms a key characteristic feature for the team members during the project period. The personality behaviour is evaluated based on personnel opinion, attitude and other features related to team operation [11]. Every member in a team will have a topology which dominates the formation of a group based on five personality factors [12] relating the team personality and performance effort. Out of the five, two features, joy and skill are considered. Joy is one of the pairs of emotions used to produce the initial mood whereas skill is the capability of doing things in an innovative way and also being intellectual [13]. These features are also closely linked to an agile software method [14].

3 Propose Work

The proposed method is done in two steps.

3.1 Fuzzy Analogy

Fuzzy analogy [15] is the fuzzification of classical analogy procedure. It is comprised of three steps. 1) Identification of cases, 2) Retrieval of similar cases and 3) Case adaptation.

3.1.1 Identification of Cases

In this case, the objective is to estimate the software project effort. In the case of numerical value x_0, its fuzzification will be done by the membership function which takes the value of 1 when x is equal to x_0 and 0 otherwise. For categorical value, let us have M attributes and for each attribute M_j, a measure with linguistic values is defined (A_k^j). Each linguistic value, A_k^j, is represented by a fuzzy set with a membership function ($\mu_{A_k^j}$).Rules formulated, based on the fuzzy sets of modes, sizes and efforts .

3.1.2 Retrieval of Similar Cases

This step is based on the choice of a software project similarity measure.These measures evaluate the overall similarity of two projects P_1 and P_2, $d(P_1, P_2)$ by combining all the individual similarities if P_1 and P_2 associated with the various linguistic variables V_j describing P_1 and P_2, $d_{V_j}(P_1, P_2)$. Here, the distance can be measured by employing Euclidean distance method.

$$d_{V_j}(P_1, P_2) = \begin{cases} \max\limits_{k} \min(\mu_{A_k^j}(P_1), \mu_{A_k^j}(P_2)) \\ \max - \min \quad aggregation \\ \sum\limits_{k} \mu_{A_{ik}^j}(P_1) \times \mu_{A_k^j}(P_2) \\ sum - product \quad aggregation \end{cases} \tag{1}$$

Where V_j are the linguistic variable describing the project P_1 and P_2. A_k^j are the fuzzy sets associated with V_j and $\mu_{A_k^j}$ are the membership functions representing fuzzy sets A_k^j .

3.1.3 Case Adaptation

The objective of this step is to derive an estimate for the new project by using the known effort values of similar projects. The objective of this step is to take only the k first projects which are similar to the new project. In the proposed method, all the projects in the data set are used to derive an estimate of the new project. Each historical project will contribute, in the calculation of the effort of the new project, according to its degree of similarity with this project.

3.2 Team Characteristics Evaluation

A personality trait is an enduring pattern of inner behaviour that is extremely inflexible, deviates markedly from the expectation of a person's culture and mood that causes distress.Two of the factors of personality trait are the joy and skill that can increase or decrease the software effort based on their values accordingly. Let Jy and Sl are the joy and skill character value that represent the mood of the members in a software team. Consider them in a class where the value Z can be measured.

If $Jy > thr$ then
Set Sl less than Jy
Else
Set Sl greater than Jy
End if
If $Jy < thr$ then
$Z = value$
End if
If $Sl > thr$ then
$Z = value$
End if

Pseudo code for finding the team characteristics

Since the vagueness and uncertainty of software effort drivers cannot be avoided, a fuzzy model has the advantage of easily verifying the cost drivers by adopting fuzzy sets.

$$Effort = A * (SIZE)^{B+0.01*\sum_{i=1}^{N} d_i} * \prod_{i=1}^{n} EM_i \pm Z \quad (2)$$

where A and B are constants, d is the distance and EM effort multipliers. By using the above formula the effort is estimated.

4 Results and Discussion

The proposed methodology is implemented in JAVA Net Beans. The datasets used in the study is NASA 60[16]. Consider the dataset based on hypothetical assumption having N projects and each project may have different number of workers m. The overall average percentage of the team characters of joy and skill of the employees in a dataset is identified by circulating an analysis form for every individual after the completion of a period or module. The report consists of various objective type questions in two sections related to the behaviour and the performance of the members involved in the project which will be filled by the individuals. Every answer has a credit point and from the report, total weightage is found based on the credit points answered by the individual member. The same procedure is followed for each member in one project and finally the average is taken in terms of percentage. This is repeated for other projects and an overall average percentage is found for the individual characters in a dataset. Therefore the formula for finding the team member characteristics are computed as

$$J_{avg} = \frac{1}{N}\left[\sum_{i=1}^{N}\left(\left(\sum_{i=1}^{m} J_i\right) * f\right)\right] \tag{3}$$

$$S_{avg} = \frac{1}{N}\left[\sum_{i=1}^{N}\left(\left(\sum_{i=1}^{m} S_i\right) * f\right)\right] \tag{4}$$

J_i – represents the credit got for each Joy question.
J_{avg} – represents the value of the average credit for Joy in a team.
m- represents the number of joy questions.
N- Number of persons in a team.
S_i – represents the credit got for each Skill question.
S_{avg} – represents the value of the average credit for Skill in a team.
f – threshold factor which varies for joy and skill.

The threshold factor is predicted based on the analysis of various reviews conducted during the process. Consider the overall average percentage for the characters like joy and skill hypothetically for Nasa60 dataset are 75% and 45.5%. Figure 1 shows the effort values obtained in using the team characters for Nasa60 dataset.

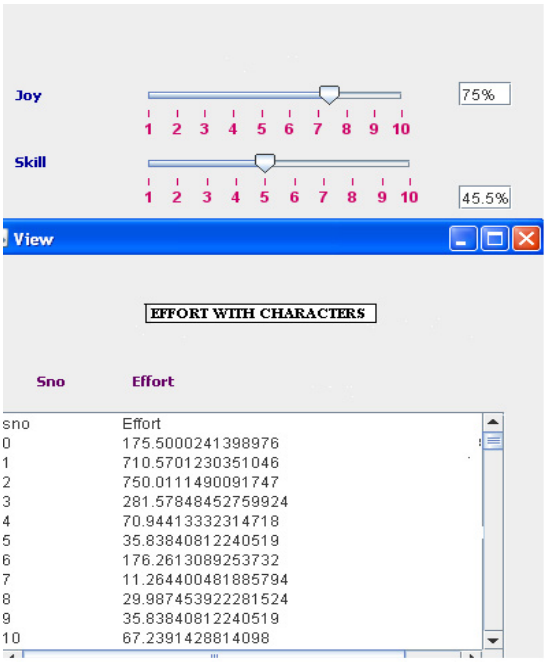

Fig. 1. Estimated Effort with Team Characters for Nasa60 dataset

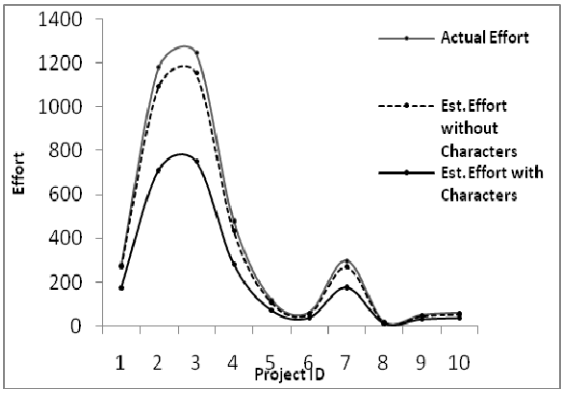

Fig. 2. Comparative Analysis of Effort for Nasa60 dataset

Figure 2 depicts the comparative analysis of the actual and estimated effort for the dataset with and without the team characteristic features. It is clearly seen that better results are obtained when the team characteristic features are included along with the project features. The overall average effort for the dataset in the proposed method is tabulated in table 1 and its comparison is given in figure 3.

Table 1. Avg. Effort comparison with and without team characters

Dataset	No. of Projects	Actual Avg.Effort	Estimated Avg.Effort without Characters	Estimated Avg.Effort with Characters
Nasa60	60	406.413	359.324	234.534

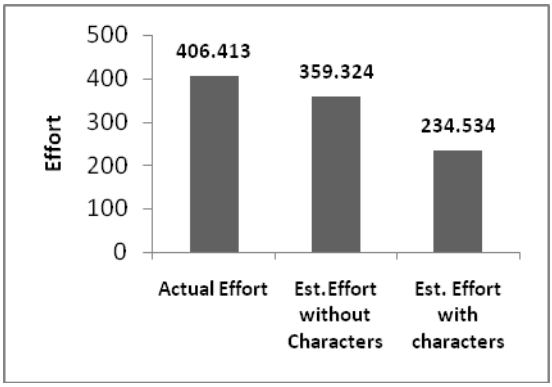

Fig. 3. Comparative Analysis of Average Effort for Nasa60 dataset

5 Conclusion

The complexity of new software has made the estimation of software effort very difficult for the project management. The existing techniques use only the project features for software effort estimation and not the team characteristics in an effective manner. This research paper proposes a different approach using Fuzzy Analogy for the team characteristic features like joy and skill of the employees during the project schedule for estimating the software effort resourcefully. This method is based on reasoning by analogy, fuzzy logic and linguistic quantifiers with team characteristics and can be used for the software projects. From the analysis of results, it is deduced that the proposed method effectively estimates the effort for historical datasets based on the hypothetical models. In addition, the studies have showed a good performance in overcoming the imprecision and uncertainty problem. The future research can be embarked on with real time data and considering other characteristics like neuroticism, agreeableness and culture to predict the effort efficiently.

References

1. Chiu, N.H., Huang, S.-J.: The adjusted analogy-based software effort estimation based on similarity distances. The Journal of Systems and Software 80, 628–640 (2007)
2. Kazemifard, M., Zaeri, A., Ghasem-Aghaee, N., Nematbakhsh, M.A., Mardukhi, F.: Fuzzy Emotional COCOMO II Software Cost Estimation (FECSCE) using Multi-Agent Systems. Applied Soft Computing 11, 2260–2270 (2011)

3. Li, J., Ruhe, G., Al-Emran, A., Richter, M.M.: A Flexible Method for Software Effort Estimation by Analogy. Empirical Software Engineering 12, 65–106 (2006)
4. Azzeh, M., Neagu, D., Cowling, P.I.: Analogy-based software effort estimation using Fuzzy numbers. The Journal of Systems and Software 84, 270–284 (2011)
5. Fedotova, O., Teixeira, L., Alvelos, H.: Software Effort Estimation with Multiple Linear Regression: review and practical application. Journal of Information Science and Engineering (2011)
6. Satyananda Reddy, C., Raju, K.: Improving the Accuracy of Effort Estimation through Fuzzy Set Representation of Size. Journal of Computer Science 5(6), 451–455 (2009)
7. Prasad Reddy, P.V.G.D., Sudha, K.R., Rama Sree, P.: Application of Fuzzy Logic Approach to Software Effort Estimation. International Journal of Advanced Computer Science and Applications 2(5) (2011)
8. Malathi, S., Sridhar, S.: A Classical Fuzzy Approach for Software Effort Estimation on Machine Learning Technique. International Journal of Computer Science Issues 8(6)(1) (November 2011)
9. Azzeh, M., Neagu, D., Cowling, P.I.: Improving Analogy Software Effort Estimation using Fuzzy Feature Subset Selection Algorithm. In: 4th International Conference on Predictive Models in Software Engineering, pp. 71–78 (2008)
10. Kazemifard, M., Ghasem-Aghaee, N., Oren, T.I.: An event based implementation of emotional agents. In: Summer Simulation Conference, Calgary, Canada, pp. 63–67 (2006)
11. Molleman, E., Nauta, A., Jehn, K.A.: Person-Job applied to teamwork: a multilevel approach. Small Group Res. 35, 515–539 (2004)
12. Costa Jr., P.T., McCrae, R.R.: NEO Personality Inventry, Revised. Psychological Assessment Resources, Odessa (1992); (Spanish version, TEA Ediciones, Madrid) (2002)
13. Ghasem-Aghaee, N., Oren, T.I.: Towards fuzzy agents with dynamic personality for human behavior simulation. In: Summer Computer Simulation Conference, Montreal, Canada, pp. 3–10 (2003)
14. Beck, K.: Extreme Programming Explained: Embrace Change. Addison-Wesley, Reading (1999)
15. Malathi, S., Sridhar, S.: Optimization of Fuzzy Analogy in Software cost estimation using linguistic Variables. In: International Conference on Modeling, Optimization and Computing, ICMOC (2012)
16. Sayyad Shirabad, J., Menzies, T.J.: The PROMISE Repository of Software Engineering Databases. School of Information Technology and Engineering, University of Ottawa, Canada (2005), `http://promise.site.uottawa.ca/SERepository`