

# A Review on Cost and Effort Estimation Approach for Software Development

Geetika Batra<sup>1</sup>, Kuntal Barua<sup>2</sup>

<sup>1</sup>M.Tech Scholar, Department of Computer Science, L.N.C.T, Indore

<sup>2</sup>Asst Professor, Department of Computer Science, L.N.C.T, Indore

**Abstract**— *The development of software in a success manner depends on accurate estimation, as numerous factors are responsible for the overall assessment of any project. Estimation of cost and effort is most complicated and challenging task in software industry. Many estimation models are introduced by the time, that concludes estimation is not a precise science and demanding of new methodologies should be proposed day by day. This paper highlights general overview of cost estimation different techniques and metrics including latest trends in this field. Identification of Problems and solutions to those problems is motive of this review. Introducing the new approach is supportive for the reduction of cost and effort.*

**Key Words**—Cost Estimation Models, Effort Estimation, Software Engineering, Metrics.

## I. INTRODUCTION

Software engineering, a layered technology is responsible for the development of software product in software industry. An approach used by software industry to extend any project is a Software Development Life Cycle. Various SDLC models are employed for software development [7]. These models consist of different phases have its own significance. On completion of any projects with the help of SDLC approach there leads some escape factors by programmer. Rework, extension, maintenance on projects results in increase of cost and effort. Level of complexities and size also affects the accuracy of estimation. The most crucial task and still open challenge is estimation; accuracy of any project depends on many factors like business plan, recourses required, resources used, customer expectations, impact of changes and replanning [5]. The field of software is growing rapidly, each cost models has its own pros and cons. Developing software cost continues to increase and causes hindrance for the accurate estimation. In few researches various estimation methods are introduced including algorithmic methods, expert judgment method, analogy method, top-down and bottom-up method [3].

In this paper we summarize different estimation techniques and models like COCOMO, SLIM, Checkpoint, and SEER etc [2]. Study of size estimation techniques like LOC, function point, feature point, and object point [1] that plays an important role in cost estimation. This leads to conclude which type of methods and models use to generate the exact estimation. To decrease the budget of project different approaches, methods and models can be suggested, as increase of cost and effort estimation is biggest problem for

the project manager to deal with. By summarizing literature review problems are identified and solution in the form of a new approach can be developed. As each approach at some or other end had its own pitfalls. The software has one important characteristic that the cost always grows with the time, as time increases the cost always increase. In this case, this paper introduces novel approach, for reducing it.

This section describes an overview of estimation of projects Section 2, Literature review includes Background that gives the description of the existing work and the problem description of methods and models of cost and effort estimation and Literature Extraction which mentions advantages and disadvantages of techniques and sort them to find the best approach. Section 3, Problem identification includes Problem Definition and its solution, Section 4 includes Conclusion.

## II. LITERATURE REVIEW

### A. Background

The several approaches for the cost estimation techniques are developed. It is classified into following: Model Based -SLIM, checkpoint, SEER, COCOMO. Expertise Based-Delphi, Rule-Based, Dynamics-Based-Abdel-HamidM adnick. Learning Oriented Neural, Case based. Regression Based-OLS, Robust. Composite Bayesian-COCOCMO-II. Each technique has their own significance and even its disadvantages are also highlighted. This paper concludes that no one model or single method should be favored over others. The key to achieve the goal i.e. estimation, can be done through variety o tools and methods and then work upon the area that what reasons effects estimation. Project planning also plays an important role in the budget of any software development [2].

In the estimation process, apart from the models, metrics is also responsible for accuracy. Metrics is a measurement of size of project, and various metrics techniques are introduced in the field if software engineering. Some popular techniques are line of codes (LOC), function points, object points and feature points. The use of good metrics helps in better estimation. The review also mentions that researchers face the problem that as software grows in importance and size, it also grows in complexity, makes the calculation difficult. There are various methodologies for effort estimation, mostly used are: Algorithmic, top-bottom, bottom-up, expert judgment price-to-win and estimation by analogy. The models for estimation can be divided according the metrics like function oriented metrics introduces SEER-SEM ,

Checkpoint Model, ESTIMAC, SLIM etc. As COCOMO used LOC whereas COCOMO-II used function oriented. This paper concludes that function oriented metrics is best technique for the accuracy as it includes many factors for the assessment [1].

The Project completion within a time and budget depend on the accuracy of estimation. In survey, comparison of all methods like expert judgment, estimation by analogy etc and highlights its advantages and disadvantages. The new approach of collapsing the data mining technique with software engineering is introduced for estimation. 2CEE i.e. 21<sup>st</sup> Century effort estimation model is developed which has been encoded in window based tool. By using the concept of data mining it analyze the past project and generates new pattern. The other methods are also effective but as accurate calculation of effort and cost is endless therefore new approaches should developed gradually. To take decision which is better is somewhat difficult, as each models and methods have their own importance and significance [3]. The purpose of the paper is to present the most applicable models and methods for effort estimation. In software engineering effort indicate measure of workforce and defined as total time members take for development to perform a specific task. It is expressed as man-day, man-months, and man-year. Some reasons for vary of effort estimation are: Project approval, project management, defining of project task etc. The number of different effort estimation, techniques, models and attentiveness among researchers is required to improve estimation [10]

Software Cost Estimation is process of predicting the effort required to develop a software system. This paper provides general overview of methods including recent trends in the field of software. The major categories of model are algorithmic and non-algorithmic having its own strength and weakness. Selection is a key factor of accuracy. The estimates depend on effort, project duration, cost etc. This paper concludes that cost estimation will remain complex problem and researches should indulge to approach new technique for this task. Models based on Artificial Intelligence techniques should be used for more accurate estimation [5].

The purpose of this paper is to introduce about the software metrics and models, as the most popular metrics are LOC that is considered as oldest and not good but widely used for estimation. In function point analysis there are some countable measures of information domain and calculation of software complexity like: internal files, external files, internal and external outputs and finally external enquiries. These requirements made it complex but essential for the accurate for the cost and effort estimation. Other metrics known as software science or Halstead equation is also introduced for estimation as it depends on program length, vocabulary, volume, difficulties etc, though it was found better than LOC but it had its own weakness [6].

Primary task of software industry in software developing is estimation; this paper concludes that no single technique is best in all situations. Concept of new approaches with the use

of soft computing and machine learning is also a successful idea. The different techniques must use to reach the destination. Comparison of approaches is not a solution, hybrid techniques can be used as better utilization [9]

Estimation Knowledge is essential in software industries and review shows that this works will continue with development in this field. This paper emphasis that mostly 60-80% projects suffer overruns of schedules or effort. Frequently used techniques in the companies are expert judgment as the other models do not gives confirmation for the correctness and accurateness of estimates. There is a lack of survey in the analysis phase. The remaining portion of this paper discusses some research questions with evidence that leads to some conclusion. Firstly, up to what extend deviation of project from the original plans regarding its cost schedule and effort? The selection of method for estimation in a particular project and will it performs in a systematic manner? Next, importance of accuracy in the estimation how it can create problem in the completion of project? Finally, what can be the reasons that results in the deviation and responsible for overruns? In the solution to these questions some problems are answered and it can be solution in the form of new approach for estimation [8].

The development of Linux operating system, was most booming open project and widely used. In this paper the early costs of the Linux is discussed and mention the new developing cost using the new calculation approach to reduce the cost [4]. The approach for the estimation of Fedora 9 in the year 2008 by using COCOMO model for the assessment and LOC as metrics, discussing that LOC is better and widely used in the present scenario for the calculation of cost.

### ***B. Literature Extraction***

The review from different researchers and surveys clarifies that single approach is not much successful for the estimation as hybrid scheme or combination of more than one techniques can provides the accurate estimation of cost and effort[3] [9].

Software metrics for software size plays significant role as many techniques are introduced for calculation and survey concludes that function point, feature based, object oriented is better than Lines of codes but review concludes that function point is responsible for the overruns of budget and time. The widely used approach used by the companies is LOC which is old fashioned but easier for assessments [4] [8]. The great numbers of models are introduced and different techniques are followed in the reviews but those models are not found reliable in the case of accuracy and still most popular approach expert judgment or estimation by analogy [9].

Each models and methods had its own advantage and disadvantage no such review or survey can mention that the specific model is best in every situations. According to various phases each model is divided such that single model is good or best for each phase [1]. Therefore the suggestions of using combination are required.

### III. PROBLEM IDENTIFICATION

#### A. PROBLEM DEFINITION

The literature review includes background and extraction which describes some problems regarding estimation of cost and effort. Estimation always remains difficult and challenging task. As the review shows that selection of techniques, models and metrics are responsible for the inaccuracy and overruns in the case of budgets and time. This overruns effects the estimations and that directly effects environment. Software possesses characteristics that cost of any product increase with increase in time. Various models are developed but no single is responsible for correct estimation.

The programmers and managers passes through the phases of SDLC where each process affects other. After this there might be some escapes factors and those are rework on same project, extension of project and maintenance.

#### B. SOLUTION DOMAIN

The identification of problem clearly defines the difficulties faced by the researchers. To overcome this problem there are some solutions as reuse [11] [12] can become the key factor for reduction of cost and effort.

1) Reusability concept of OOP can be utilized by using the suitable technique and metric.

2) As for the reusability estimation by analogy is appropriate technique which describes that the past project is compared with new project and if the code is matched then it can be reused.

3) LOC metric can be used to estimate software size as others can cause the overruns within time.

4) Analogy concept [13] needs searching technique in the form of search engine as past projects should be placed in some code repository, and through this method the historical projects can be matched with new projects.

5) If the codes are matched then it can be used in new project reduce the effort which reduces the cost.

6) As tentative cost of new project should be calculated and the old projects cost are already known, it helps to asses reduced cost in case of no. of lines used by new project.

Through this approach the problems can be reduced as estimating by analogy with use of LOC as software size. And the retrieval techniques like cosine similarity Euclidean distance can be used such that reduction of cost and effort can be intended. The steps can be understood by the figure.

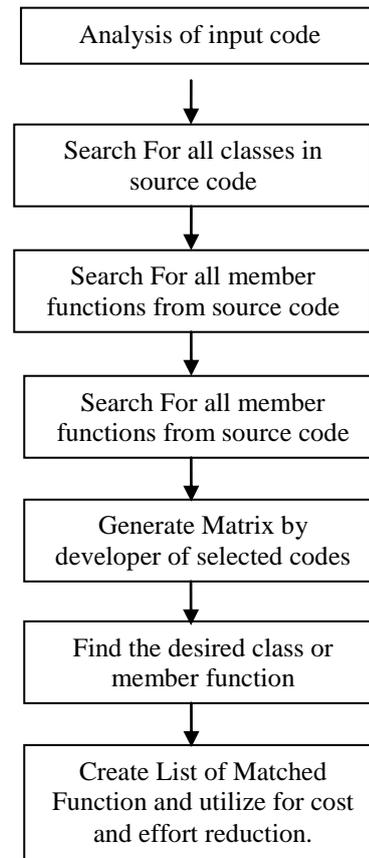


Fig 1 shows the system control to solution domain

### IV. CONCLUSION

Software development in this era is at demanding phase, and estimation of cost and effort in this filed always remains an open challenge and considered to be a complex task. Software engineering and SDLC have their significance presence in the estimation. The review also shows that many reviewers and researchers state that assessment of cost gradually increases or decreases. Though it is an essential task the ignorance is not acceptable. Awareness of project managers and selection of methods are responsible for over budget. Each estimation techniques like COCOMO, SLIM model have it own prospects to be good and at the same time suffered with pitfalls. Estimation by algorithmic, non algorithmic, top-bottom approach or bottom-up approach etc shows their own significance in different manner in the field of software cost estimation. The suggestions of using combination of different techniques and models can be much more efficient as alone model and method are not much effective in estimation Concept of reusability helps in reducing the cost and effort with use of analogy estimation and suitable searching and retrieval techniques. Lastly, gradual increment of new approaches and the hybrid scheme of methods with models can be used.

## REFERENCES

- [1] Samaresh Mishra, Kabita Hazra, and Rajib Mall, "A Survey of Metrics for Software Development Effort Estimation", International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 2, No. 5, October 2011, ISSN: 2079-2557.
- [2] Barry Boehma, Chris Abts and Sunita Chulani, "Software development cost estimation approaches –A survey", Annals of Software Engineering 10 (2000) 177–205.
- [3] Narendra Sharma, Aman Bajpai, Mr. Ratnesh Litoriya "A Comparison of software cost estimation methods: A Survey", The International Journal of Computer Science & Applications (TIJCSA), Volume 1, No. 3, May 2012.
- [4] Amanda McPherson, Brian Proffitt, and Ron Hale-Evans, "Estimating the Total Development Cost of a Linux Distribution", Whitepaper, The Linux Foundation, October 2008.
- [5] Hareton Leung, Zhang Fan "Software Cost Estimation", Department of Computing The Hong Kong Polytechnic University {cshleung, [csfzhang](mailto:csfzhang@comp.polyu.edu.hk)}@comp.polyu.edu.hk.
- [6] Syed Ali Abbas, Xiaofeng Liao, Aqeel Ur Rehman, Afshan Azam, M.I. Abdullah, "Cost Estimation: A Survey of Well-known Historic Cost Estimation Techniques", Journal of Emerging Trends in Computing and Information Sciences, VOL. 3, NO. 4, April 2012.
- [7] Vishwas Massey, Prof. K.J.Satao, "Comparing Various SDLC Models And The New Proposed Model On The Basis Of Available Methodology "International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 4, April 2012.
- [8] Kjetil Moløkken and Magne Jørgensen Simula "A Review of Surveys on Software Effort Estimation", Research Laboratory {kjetilmo,magnej}@simula.n.
- [9] S.K. Mohanty & A.K. Bisoi, "Software Effort Estimation Approaches – A Review", International Journal of Internet Computing ISSN No: 2231 – 6965, VOL- 1, ISS- 3 2012.
- [10] Jovan Živadinović, Ph.D\*, Zorica Medić, Dragan Maksimović, Aleksanda Damnjanović, M.Sc, Slanana Vujčić, "Methods of Effort Estimation in Software Engineering, International Symposium Engineering Management And Competitiveness 2011 (EMC2011) June 24-25, 2011, Zrenjanin, Serbia.
- [11] Daniel Schwabe and Luiselena Esmeraldo, Gustavo Rossi and Fernando "Engineering Web Applications for Reuse", Catholic University of Rio de Janeiro (PUC) Lyardet La Plata University 2001, IEEE.
- [12] Collin McMillan, Member, IEEE, Mark Grechanik, Member, IEEE, Denys Poshyvanyk, Member, IEEE, Chen Fu, Member, IEEE, and Qing Xie, Member, IEEE "Exemplar: A Source Code Search Engine for Finding Highly Relevant Applications", IEEE Transactions on Software Engineering, vol. 38, no. 5, September/October 2012 .
- [13] S.Malathi, Dr.S.Sridhar, "Optimization of Fuzzy Analogy in Software Cost Estimation Using Linguistic Variables", International Conference on Modelling, Optimization and Computing (ICMOC -2012).