Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.8, No.3, 2017

Qualitative Inquiry of Software Cost and Effort Estimation Techniques

Muhammad Naeem
Department of Computer Science, University of Agriculture Faisalabad, Pakistan

Zain Ul Hassan
Department of Computer Science, University of Agriculture Faisalabad, Pakistan

Abstract
Software cost and effort estimation is one of the supremeimperative asset of software industry. Effort and costprediction plays a vibrantpart in the failure or success of the software project implementation and as well as profited revenue. To survive in the market, a software house need to estimate the development cost and time before starting of project. Cost and effort estimation is most challenging decision making task, because number of developmentnecessities are not well defined and never easy to forecast on starting stage of Software building process. Number of cost and estimation techniques are currently being used, but none of these are given the 100% of accuracy in actual prediction, as these technique fall accurate in one development environment and fail in other one. Organizations want to automate the software effort and cost estimation procedures.This paper unfold the different technique of predicting cost and effort estimation based of literature review.

1. Introduction

Software cost estimation and effort prediction has been always an important research domain, as the accuracy in prediction is inevitable for not only success of the software but also for software industry management to assist in the planning, scheduling budgeting and smooth running of software projects.

Early research in this field has been made by Berry Boehm, Constructive cost model is the one of the famous technique in the software cost and effort prediction, COCOMO-I and COCOMO-II playing their great part in the estimation in software houses. These techniques are discussed here in next section.

Approximately 67% of total software projects cross the software estimated cost, this failure in prediction is due to difference of development environments. Different strategies support different environment properly. So in making the cost estimation domain and area specific, we minimize the failure rate of estimation.

Software cost can also be estimated in a proper way if we develop the standers of the attribute in the software estimation. Different factors are involve in the standardizing the attributes like Product factor, Project factor, Platform factor and Personnel factor. Estimation of factor rather than estimating a comprehensive estimation of integrated system is more favorable in better prediction.

Proceed diagram, like use-case diagram can also be thought as an emerging technique of cost estimation of s/w projects. The proceed diagram depict the use of hierarchy on objects and entities in the system. Unadjusted function count of the functionality can be measure in terms of component by using this strategy.

In system estimation and designing process, it is very significant to manage a relationship among“effort, schedule and quality”, all these factors combined known as “magic triangle”. A Bayesian Belief Network is way of measurement of this triangle that is a form of directed graph in which the node states indicate the probabilities and the edges between these nodes represent the dependencies.

2. Common Estimation Methodologies

The estimation methodologies are:

Analogy method: estimated by comparing with the completed project and the type if it founds. Top-down method: concerned with the generalfeatures of the software that is going to be build. Bottom up method: estimate every component individually and then combines all of these in the form of project

Estimation Techniques are: Parametric include LOC method, FP metrics, IFPUG’s FPA, Feature Points, 3D function point COCOMO I and COCOMO-II. Heuristic Approach includes Expert Judgment Methods, thumb Rule, Delphi Technique.

We can divide these techniques in two categories that are parametric (statistical, numerical and historical analysis) and non-parametric (based on ANN, regression trees, analogy etc.) models. In algorithmic models Boehm’s COCOMO method is used for forecasting the requisiteemployees per month in development, also provide effort estimation needed. Boehm proposed 3 level models 1. Organic (up to 50 KDSI), 2. Semidetached (up to 300 KDSI), 3.Embedded. (Poonam, 2013)

3. Domain Effort Estimation in Software Projects

The domain specific effort estimation in contrary with the overall project cost and effort estimation. The

accuracy of the estimation of overall project is much lower. Even in most commonly used process like "waterfall Model", scatteringdesign of cost and effort and other variable factors are never thoroughly examined. A research use regression model to explain that the domain specific estimation is more accurate than other formal approaches explained in many other researches. Accuracy rate of overall estimation results in 29 % accuracy while domain specific estimation increases this rate up to 74%. So during cost and effort estimation, for better prediction, we divide effort into predefined phases.

ISBSG data repository records were analyzed that contain 4106 S/W project of different fields. The different fields have different phase distribution characteristics. The development type like ND, RD, EN have an impact on phase distribution. FPC is used to measure the magnitude of S/W by ISBSG, so concluded that the "software size and effort are inversely proportional". (Yigit *et al.,* 2009)

4. 4G language software and Estimation

The cost estimation methods like function point analysis, line of code measurement and COCOMO were being uses from three decades, but these techniques fails to provide the accurate estimation while development in the fourth generation environment like component based development, because these environment like C# and VB. Net heavily composed of form, reports, table and screens. So a technique is proposed here to achieve better and accurate estimation. In this regard the struggled is toward the basic three issues. Like which model should be used? How to measure the size of the software (LOC or FP) and what characteristics a good estimate have?

The old methods that are being use are SLOC measuring the line of source code, but this have a shortcoming that we cannot estimate SLOC at the starting phases of SDLC. Function point estimation is the alternative of SLOC that can estimate the cost and effort at the starting stages of SDLC.In thisway for the predict of 4thGeneration Language applications, first we measure the size of application using

Size=Size_C + Size where Size_C is size of the software parts like forms, Reports and tables is transferred Kilo Line of Code (KLOC). Size_C can be drawn as

Size_C = Size_F+Size_R +Size_T.

The calculated total size of the application is used to conclude person per month (nominal) as in "COCOMO-II" model. Using$PM_{nominal}$=A× (Size) ^{B}.

The above described method is applied on 19 different projects. The proposed method proved strong with the help of cross-validation and judgments on 4thGeneration Languageprojects. It gives a substantialaccuracy level. (Zia *et al., 2009)*

5. COCOMO-II and ANN(artificial Neural network)

Software cost and effort estimation means we have to estimate the amount of work, time and staff needed to develop software. But the estimation at starting of development is very difficult to analyze. The author also narrated that software cost and effort estimation is still a complex problem that has not overcome up to significant level. There for we need such a model of estimation that increases the accuracy rate and results in better prediction.

This model is a combination of COCOMO-II and artificial neural networks. In studying these model separate we come to know that there are two types of models, that are algorithmic that are constructed on numerical analysis of data from history, like SLIM and COCOMO model and non-algorithmic model that are created on the basis of new methods like "Parkinson, expert judgment and price to win" etc.

COCOMO model was introduced by Berry Boehm, having three further categories. Application composition model best suited for GUI based application. Early design model to get the rough estimation. Post architecture model use to estimate in detail on the base of LOC or FP.

The artificial neural network based on neurons that are similar to human biological neural system. These neurons can be train by using learning algorithm in software estimation process. (Iman and Ow, 2011)

6. Bayesian Belief Networks

In software prediction and designing process, it is imperative to establish a balanced relationship between "effort, schedule and quality"as these factors combined known as "magic triangle".

"A BNN (Bayesian Belief Network) is form of directed graph in which the node states indicate the probabilities and the edges between these nodes represent the dependencies". A BBN model is based on 4 different Sub models, modules are explain below.

A variable related to activities for developing a component uses a variable for estimation. Test activities related variables are involved to check theefficiencyof test related estimation model.

"Residual Defect Estimation" Model contains variables which are used to estimate the number of residual defects after defect removal activities that have taken place Test Estimation Model comprises some variables related to test effort, schedule estimation and resource allocation. After the completing these four sub modules the overall outputs are integrated to fine combined effect of estimation at project level. So using this

way project level estimation and managerial decision making is highly supported. (Hao *et al., 2006)*

7. Standard Set of Attributes for Estimation

In software development the major emphasize of the developers and the customer is to estimate the financial cost, time, resources and techniques, to develop new software's or up gradation of old developed software. In software engineering, the researcher's field has in detail analyzed that how to estimate the cost of the software project as this is very vital for both, the developing organization as well as project manager who is managing the project. Estimation of Cost and effort is actually estimated the overall success of the software product development.

For the high ranking organization have the significant impact of each cost means finance, time and manpower etc. there exist numerous cost and effort estimation methods, but the efficiency of each method basically based on its integralcharacteristics, those are highly recommended by the software engineers and most flexible to the defined projects.

So concluding on that we need a framework to that comprise on characteristics that can be divided into four types: Product factor, Project factor, Platform factor and employees factor. (Khalidi *at al.,* 2012)

8. Comparison of Different Techniques

The accurate prediction of cost and effort has large economic impact on software development. Since last three decades number of methods like COCOMO, expert judgment, analogy method, Putnam Method etc are being implemented. None of these provides the 100% accuracy. In other words it can be said that each method have its own merits and demerits. So this phenomenon creates confusion for the software industry to adopt which one for better predictions.

A research explores the statistical analyses of different model using Chi-Square test, and interprets the following results.

1. There is no substantialdissimilarity between actual value and estimated value thus null hypothesis is accepted for COCOMO –II, Delphi and Analogy method.
2. Putnam method fails in Chi-Square test at H0. So rejected. (Sharma *et al., 2012)*

Reference

Poonam P. (2013). Analysis of the Techniques for Software Cost Estimation. *Third International Conference on Advanced Computing & Communication Technologies.* 978-0-7695-4941-5/12.

Yigit K., E. Kocaguneli and A. B. Bener. (2009). Domain Specific Phase by Phase Effort Estimation in Software Projects. *International Journal of Computer, Information, and Systems Sciences, and Engineering*. 978-1-4244-5023 IEEE.

Zia Z., A. Rashid and K. U. Zaman. (2009). Software cost estimation for component based fourth generation language software applications. *The Institution of Engineering and Technology*. IET Software. Vol. 5, Iss. 1, pp. 103–110.

Iman A. and S.H. Ow. (2011). A Novel Soft Computing Model to Increase the Accuracy of Software Development Cost Estimation. *International Electrical and Electronics Engineering.* 978-1-4244-5586-7/10.

Sharma T.N., A. Bhardwaj, G. R. Kherwa. (2012). Statistical Analysis of various models of Software Cost Estimation. *International Journal of Engineering Research and Applications (IJERA)*. ISSN: 2248-9622.Vol. 2, Issue 3, May-Jun 2012, pp.683-685.

Khalidi N. A., A. Saifan and I. M. Alsmadi. (2012). Selecting a Standard Set of Attributes for Cost Estimation of Software Projects. *International Electrical and Electronics Engineering.* 978-1-4673-1550-0/12.

Hao W., F. Peng and C. Zhang. (2006). Software Project Level Estimation Model Framework based on Bayesian Belief Networks. *Proceedings of the Sixth International Conference on Quality Software.* 0-7695-2718-3/06.