FINAL REPORT OF THE MINOR RESEARCH PROJECT

ON

“EXPLORING SOFTWARE COST ESTIMATION TECHNIQUES FOR SMALL AND MEDIUM SIZED PROJECTS”

Submitted to

University Grant Commission

Western Regional Office, Pune

By

Principal Investigator

Mrs. Suvarna Ranjeet Jagtap

Department of Computer Science

Bharati Vidyapeeth University

Yashwantrao Mohite College of Arts,Science and commerce

Pune – 411 038.

(From 21-05-2013 To 20-05-2015)

1.INTRODUCTION

Software project development is a tedious and complex activity which should be carried out in a planned manner to achieve success. IT companies face a lot of problems and losses due to software cost overruns during project development process. There are number of issues related to cost overruns such as lack of knowledge related to client's requirements, technology change , incompetent team members , geographical distances etc.

Software cost estimation is an approximate judgment of cost required to complete a software project .It is defined as "A set of techniques and procedures that is used to derive the software cost estimate"

It is often measured in terms of efforts as Person – months/years. The overall cost of the project includes managerial cost , development cost and the cost of resources needed. Here , development cost also includes training cost and quality assurance cost.

Basically software cost estimation will never be exact science because there are too many variables such as human , political , environmental , technical .Software development process includes complex activities that are difficult to judge. It may also include different development environments which takes into account varied number of working hours of the team members such as 40+ hrs in North America ,35+hrs in Europe and again 40+hrs in India.

Software cost estimation is needed for the following purposes:-

1) Initial project bidding ,budgeting and planning
2) Cost control planning
3) Protects integrity of the project
4) Suffers from efforts and cost overruns
5) Faces problems due to unclear requirements, new design , new development tools and development.

Usually a software project is viewed as an amount of lines of source code generated during the development process. Hence , there are three types of software project in terms of the size namely small ,medium and large.

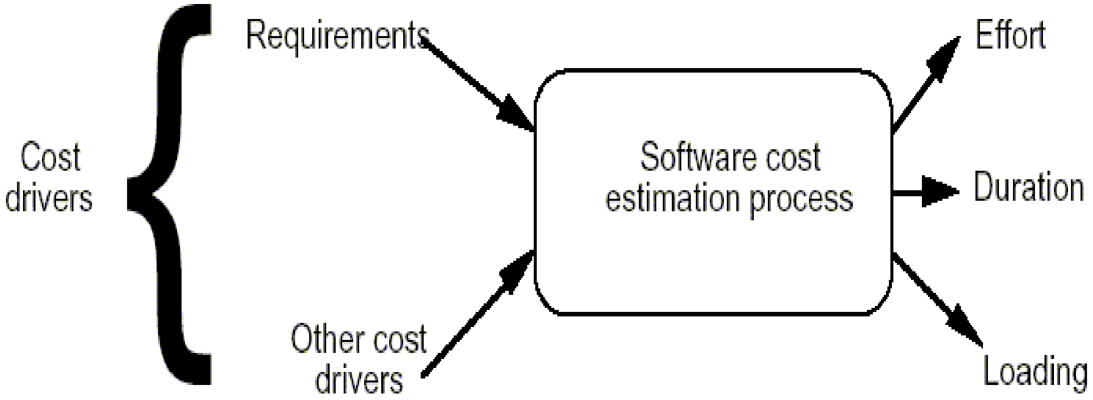


Figure 1.0 : Classical view of software estimation process (Vigder and Kark, 1994)

Most of the software cost estimation models views the estimation process as being a function that is computed from a set of cost drivers. And in most cost estimation techniques the primary cost driver or the most important cost driver is believed to be the software requirements. As illustrated in figure 1, in a classical view of software estimation process, the software requirements are the primary input to the process and also form the basis for the cost estimation. The cost estimate will then be adjusted accordingly to a number of other cost drivers to arrive at the final estimate. So what is cost driver? Cost driver is anything that may or will affect the cost of the software. Cost driver are things such as design methodology, skill-levels, risk assessment, personnel experience, programming language or system complexity.

In a classical view of the estimation process, it will generate three outputs - efforts, duration and loading. The following is a brief description of the outputs:

- Manpower loading - number of personnel (which also includes management personnel) that are allocated to the project as a function of time.
- Project duration - time that is needed to complete the project.
- Effort - amount of effort required to complete the project and is usually measured in units as man-months (MM) or person-months (PM).

The outputs (loading, duration and effort) are usually computed as fixed number with or without tolerance in the classical view. But in reality, the cost estimation process is more complex than what is shown in figure 1. Many of the data that are inputs to the process are modified or refined during the software cost estimation process.

The cost estimation accuracy helps to determine how well or how accurate our estimation is when using a particular model or technique. We can assess the performance of the software estimation technique by:

- Absolute Error (Epred - Eact)
- Percentage or Relative Error (Epred - Eact) / Eact
- Mean Magnitude of Relative Error

Each of the error calculation techniques has advantages and disadvantages. For example, absolute error fails to measure the size of the project, and mean magnitude of relative error will mask any systematic bias. There is a lot of software cost estimation methods or techniques in the software industry. Here are a few techniques that will be discussed in this document:

- Algorithmic (Parametric) model
- Expert Judgment (Expertise Based)
- Top - Down
- Bottom - Up
- Estimation by Analogy
- Price to Win Estimation

2.REVIEW OF RESEARCH AND DEVELOPMENT IN THE SUBJECT

Software development is an expensive process and compared to hardware expenses, it is the major cost for information system budgets. Investigation reports showed that majority of the projects end up with different budgets than estimates, 63% over-running them and 14% under running them [4]. Therefore, in order to manage the software projects effectively and determine economic feasibility and control of project, it is important to have accurate, reliable cost estimates [3].Accurate cost estimation is a challenging task in software management. Inaccurate estimations (over/under) may cause opportunity loss, nonproductive projects, low quality software, less profit products, and even the break-down of the project in the mid-stream. On the contrary, with the knowledge of cost of software project early in development cycle, management can approve or reject project proposal or rearrange resources accordingly [2], [5].

In order to make accurate cost estimations in the literature some models and techniques exists, such as COCOMO, SLIM and FPA. Other than these formal models, some approaches depending on expert intuition or guess have also been used, while determining the project cost, Studies claim that both formal and expert estimations are beneficial for some projects, where as, not useful for some others.

Although both approaches help engineers to optimize development process, improve control over resources and minimize project failure risk [7], it is still difficult to make accurate estimation [7]. The reason behind this is that cost estimation is affected by multiple parameters related to technologies, scheduling, manager - team member skills and experiences, mentality and culture, team cohesion, productivity, project size, complexity, reliability, quality and many more. These parameters are not easy to measure and manage and also have positive or negative impact on estimation accuracy [7].

Many researchers have used different techniques for software cost estimation such as judgment based estimation(expert estimation) which is the most commonly used approach by the software industry[4].The use of regression analysis is to derive predictive equations for software metrics using nontraditional methods like neural networks ,fuzzy logic models etc[9].Analogy based methods are the best performers for software cost estimation[3].There is a research which shows that Fuzzy logic based systems can deal with imprecision which characterizes the early phases of most software development projects[1]. A lot of research has been done using the various machine learning methods[2].

3.SAMPLE PROJECT DESCRIPTION AND SCOPE

The sample project undertaken to find out various software cost estimation techniques used for estimating cost of small sized projects . As it is difficult to incept such projects using typical software development models such as waterfall or spiral.

The sample project has the following preferences :-

1]The data collected through a rigorous survey is transferred to tables and an excel sheet .

2]It has been decided to build a model in JAVA using NETBEANS which applies Cocomo II estimates for small sized projects . Also an add-in of JAVA , JEXCELAPI is introduced so that linear regression can be carried out on the collected data.

3] Since the project is small ,detailed design or elaboration phase is not necessary.

4] The project has time constraint.

4. OBJECTIVES OF THE PROJECT

The set of objectives were proposed during project inception :-

- To identify the estimation processes resulting in systematic lowering of estimation errors
- To identify which Software Cost Estimation model to use
- To identify factors to improve evaluation and training of people responsible for estimation
- To identify the factors that lead to estimation errors which will enable improved risk management.

5. RESEARCH METHODOLOGY

To meet the above stated objectives the following methodology is applied:-

a) A detailed study of recent literature was carried out to find the traditional approaches used to estimate the cost of software projects by critically analysing a variety of research papers published in international as well as national research journals , books on models and techniques used .

b) Structured interviews were also conducted for collecting the facts where IT people who were resistant to give any information in written . Similarly , record reviews were also taken into consideration for collecting factual information.

c) A rigorous survey was conducted where a compact questionnaire was prepared to collect relevant information about small sized projects and then distributed to project managers ,team leaders , senior application developers and few business analysts.

d) Then the data collected through the survey was pre-processed and made ready to use.

e) Then a model is developed using NETBEANS to calculate the cost of software projects using COCOMOII estimate.

f) Also linear regression is carried out on the collected sample data and results are validated.

6.WORK DONE DURING THE PROJECT PERIOD, RESULTS ACHIEVED

For the survey , a questionnaire was distributed personally to more than 300 IT people such as project managers, business analysts ,team leaders and senior developers spread across various IT companies located in Hinjewadi and Magarpatta IT sectors in Pune as well as some IT people outside pune were contacted through e-mail.

- The sample size of the responded questionnaire was 150 ,out of which 90 were fully filled and remaining were incomplete.
- During pre-processing of the data, relevant fields were considered and 58 projects data was finalised for further analysis.
- Adequate efforts were taken to handle the data skilfully and necessary classification is done.
- The outcomes of the project carried out are presented in a research paper and sent to an international journal for publication.
- Details of the questionnaire prepared , screen shots of the model developed and results are attached separately .

PM is calculated by using formula

PM = A * size ^ E * EM

Here A is domain specific constant ,E is a exponent used for scaling factors while EM is used as a product of effort multipliers.

Table 3.0 The difference between actual efforts and estimated efforts

Sr. No.	**Project-ID**	**KSLOC**	**Actual Efforts**	**Estimated Efforts**
1	P01	4.85	119.5	6.49
2	P12	4.73	32.5	10.65
3	P21	4.49	40.3	9.23
4	P24	3.92	146.8	6.82
5	P30	4.87	121.9	5.28
6	P33	3.17	25.2	10.65
7	P37	2.92	8.4	2.43
8	P42	4.50	10.8	3.85
9	P49	4.85	352.8	72.92
10	P51	5.10	72.5	21.91

As shown in the above table, there is a vast difference between actual and estimated efforts. There are The following issues have been observed :-

1)Not much attention given to small sized projects as far as project planning is concerned. Usually price to win strategy is applied to such projects by most of the IT companies.

2)Story points/milestones were decided at the initial stages but not much heed given to it afterwards.

3)Usually small sized projects have fixed cost and complexity level is nominal.

4)Bottom approach is most suitable for small sized projects and review of activity based estimation must be considered to get accurate estimation.

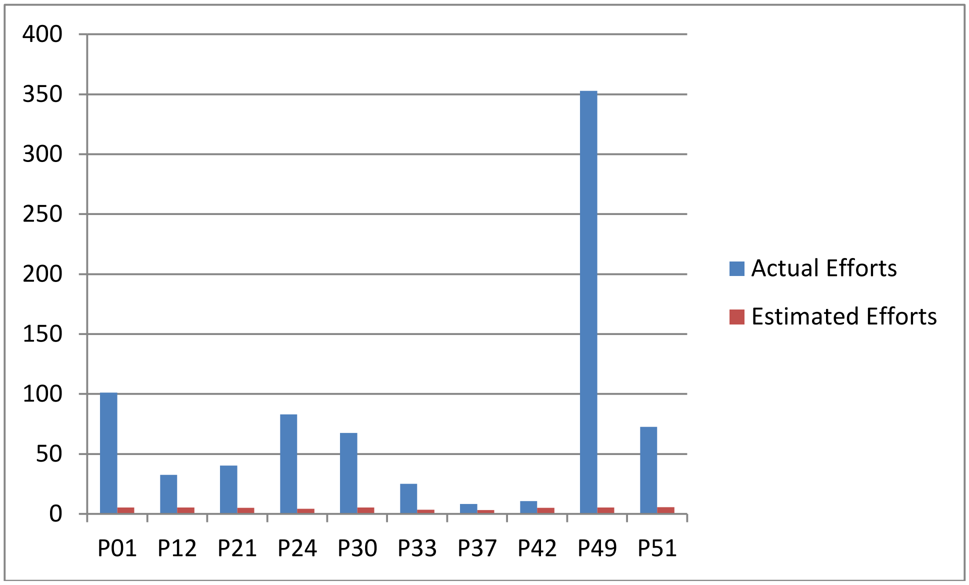


Fig. 2.0 Comparison between Actual and Estimated efforts

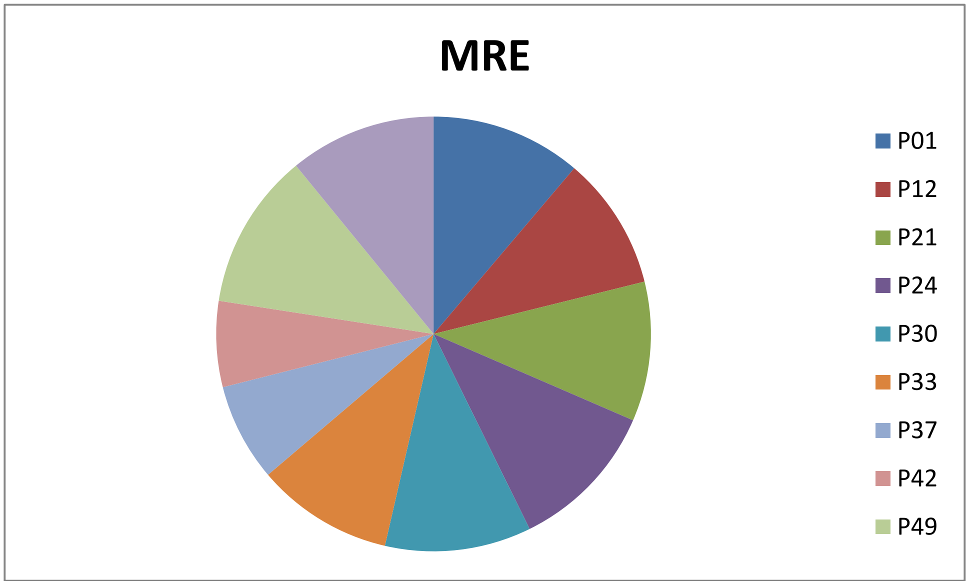


Fig. 3.0 Project wise Mean Relative Error

The above table shows vast difference between actual and estimated efforts as small size projects always suffer lack of planned activity . The following column chart maps the actual efforts and estimated efforts.

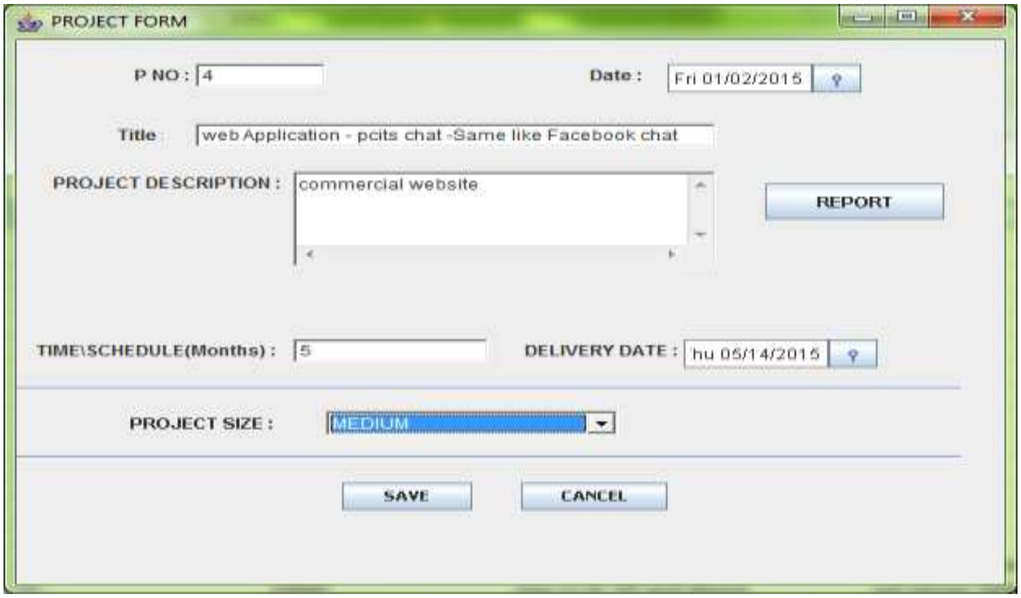
7.CONCLUSIVE REMARKS

In general, the implementation of this minor research project is done in the best possible way within the available time, resources and facilities.

This being the Novel Project for software cost estimation, most probably FIRST of this type conducted using factual data rather than standard data from database repositories. An appropriate experience and background is obtained for more detail and systematic work. This project report will definitely help the researchers as further guidelines.

An attempt will be made to continue the research work undertaken for this project in the form of a more extensive and full fledged major Research Project which will be submitted to UGC.

8.SCREEN SHOTS OF THE MODEL DEVELOPED



P NO : 4

Date : Fri 01/02/2015

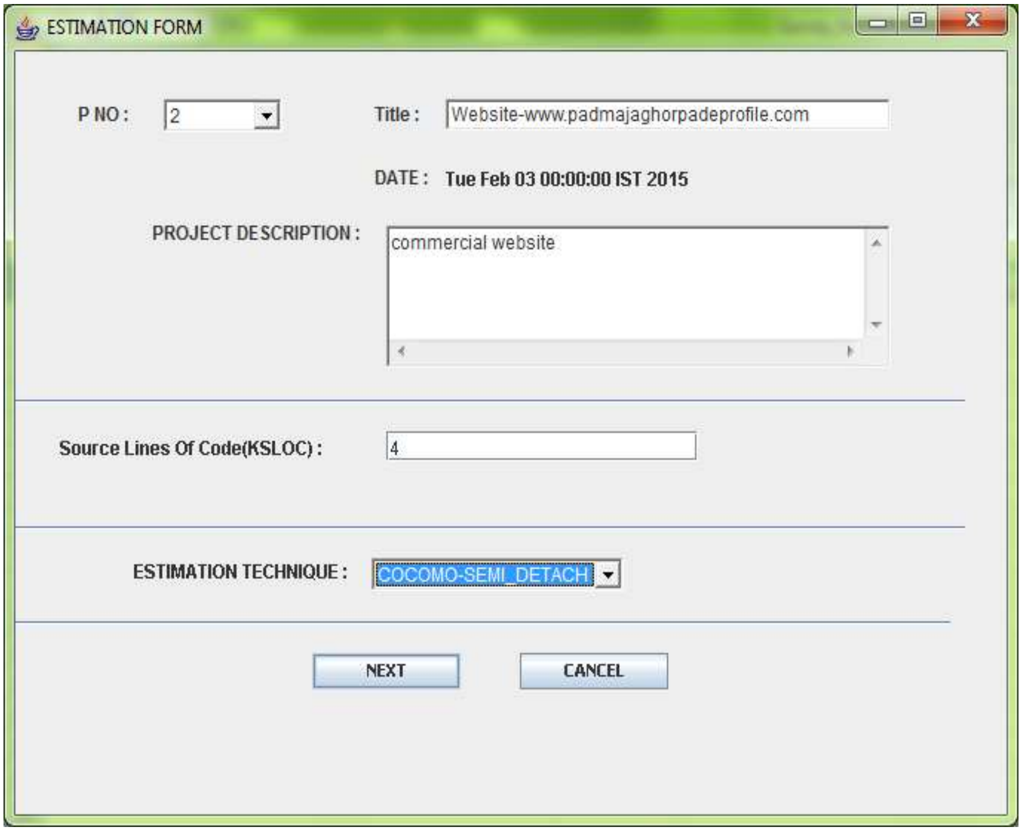
Title web Application - pcits chat -Same like Facebook chat

PROJECT DESCRIPTION : commercial website

TIME\SCHEDULE(Months) : 5

DELIVERY DATE : hu 05/14/2015

PROJECT SIZE : MEDIUM



P NO : 2

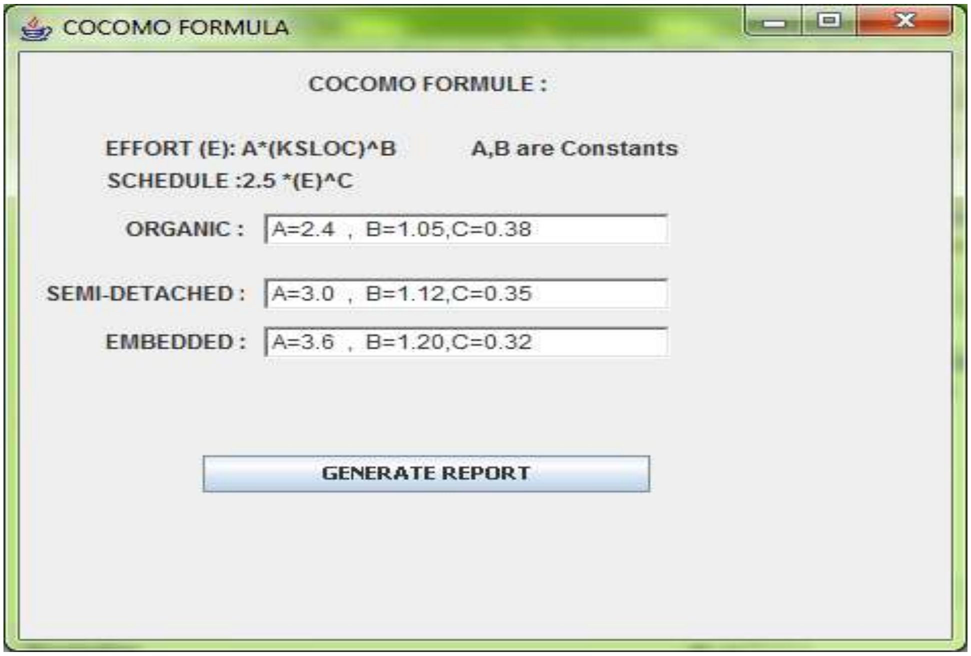
Title : Website-www.padmajaghorpadeprofile.com

DATE : Tue Feb 03 00:00:00 IST 2015

PROJECT DESCRIPTION : commercial website

Source Lines Of Code(KSLOC) : 4

ESTIMATION TECHNIQUE : COCOMO-SEMI_DETACH



COCOMO FORMULE :

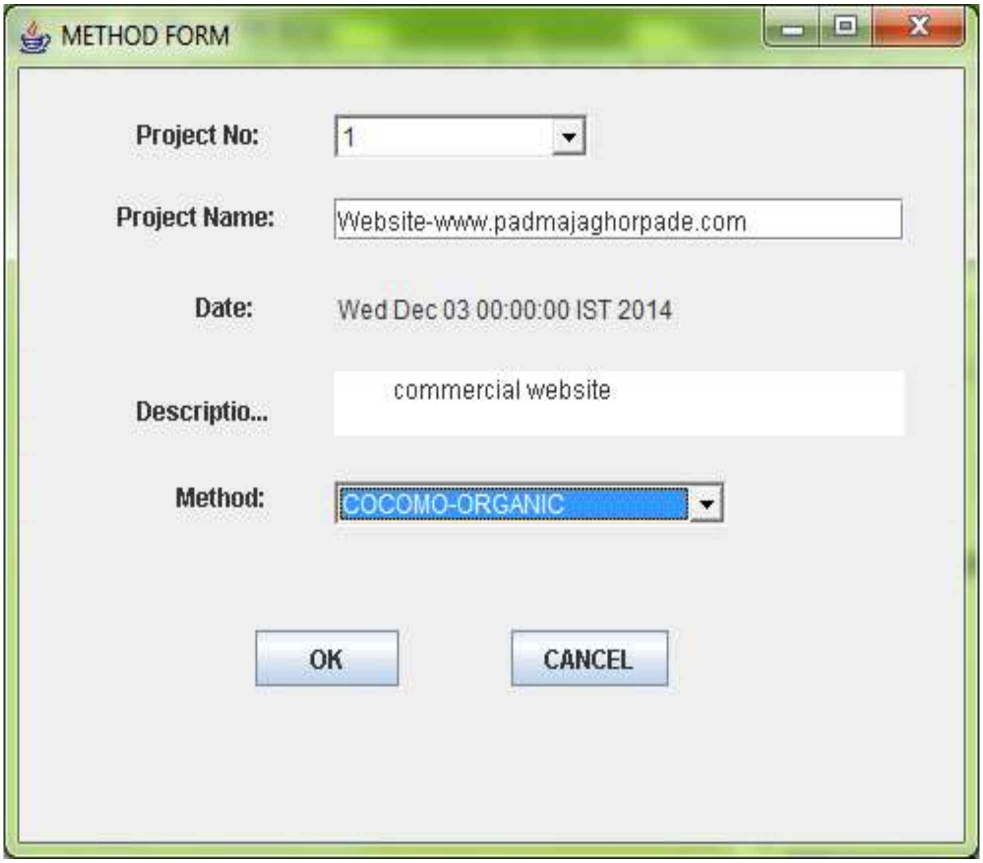
EFFORT (E): A*(KSLOC)^B A,B are Constants

SCHEDULE :2.5 *(E)^C

ORGANIC : A=2.4 , B=1.05,C=0.38

SEMI-DETACHED : A=3.0 , B=1.12,C=0.35

EMBEDDED : A=3.6 , B=1.20,C=0.32



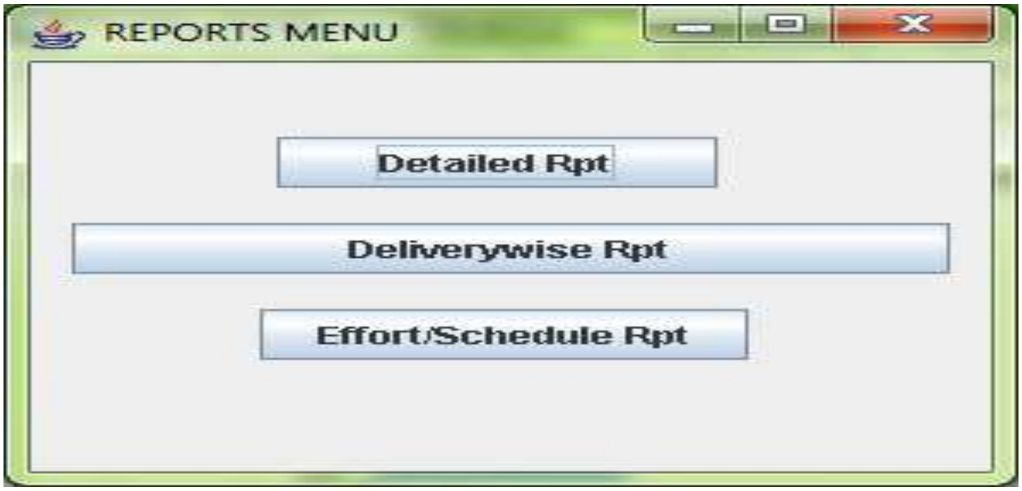
Project No: 1

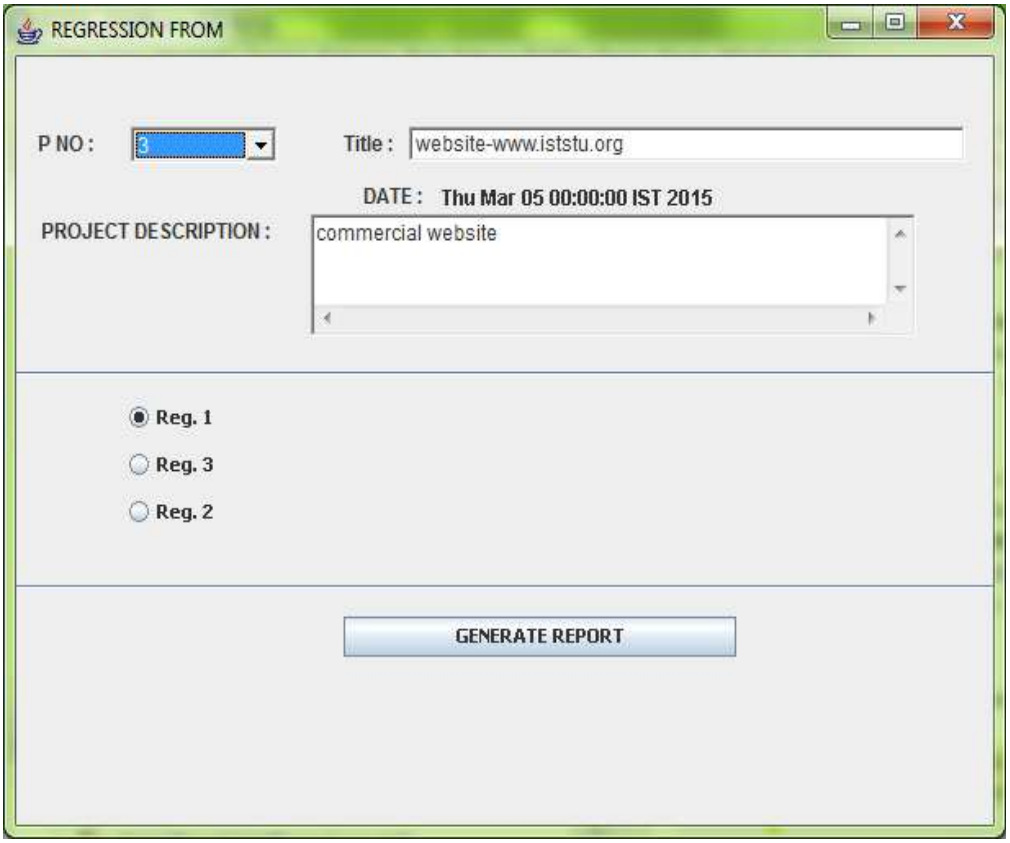
Project Name: Website-www.padmajaghorpade.com

Date: Wed Dec 03 00:00:00 IST 2014

Descriptio... commercial website

Method: COCOMO-ORGANIC





P NO : 3

Title : website-www.iststu.org

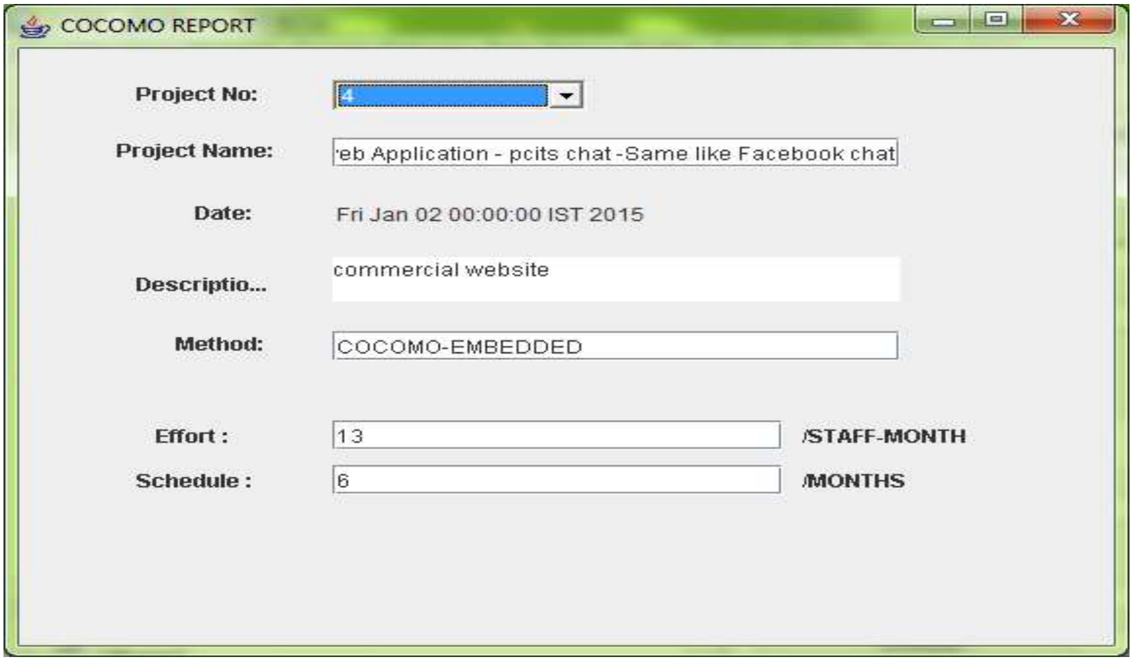
DATE : Thu Mar 05 00:00:00 IST 2015

PROJECT DESCRIPTION : commercial website

● Reg. 1

○ Reg. 3

○ Reg. 2



Project No: 4

Project Name: 'eb Application - pcits chat -Same like Facebook chat

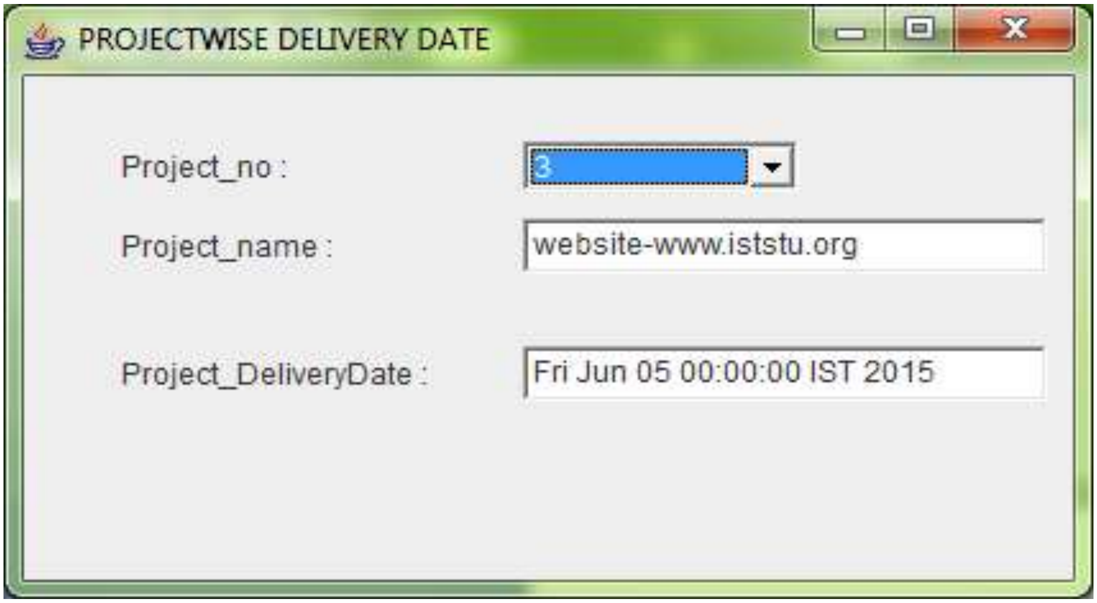
Date: Fri Jan 02 00:00:00 IST 2015

Descriptio... commercial website

Method: COCOMO-EMBEDDED

Effort : 13 /STAFF-MONTH

Schedule : 6 /MONTHS

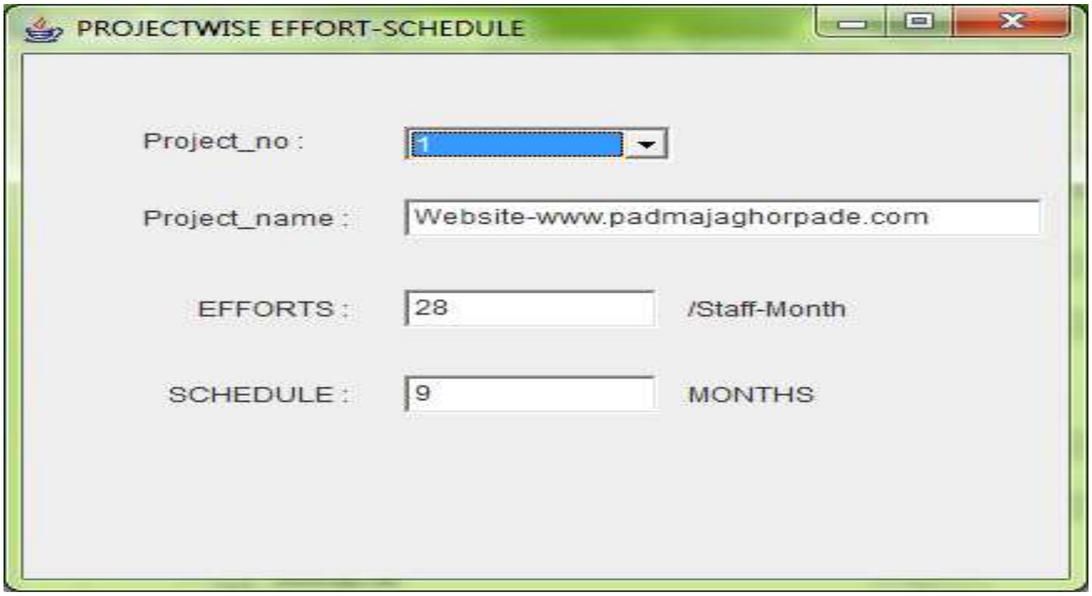


PROJECTWISE DELIVERY DATE

Project_no : 3

Project_name : website-www.iststu.org

Project_DeliveryDate : Fri Jun 05 00:00:00 IST 2015



PROJECTWISE EFFORT-SCHEDULE

Project_no : 1

Project_name : Website-www.padmajaghorpade.com

EFFORTS : 28 /Staff-Month

SCHEDULE : 9 MONTHS

References

1.Agrawal M., Chari K., (2007) “Software Effort, Quality, and Cycle Time: A Study of CMM Level 5 Projects” IEEE Transactions on Software Engineering, Volume:33, Issue: 3, pp. 145 – 156.

2. Heemstra, F.J. "Software Cost Estimation". *http://core.ac.uk/*. N.p., 1992. Web. 21 June 2015.

3. Khatibi, Vahid, and Dayang Jawawi. "Software Cost Estimation Methods: A Review". *http://cisjournal.org/*. N.p., 2010. Web. 21 June 2015.

4.Attarzadeh, I., Hock, O.S.,(2009) “Using Enhancement Method to Improve Earned Value Index to Achieve an Accurate Project Time and Cost Estimation

5. B. W. Boehm, *Software engineering economics*, Englewood Cliffs, NJ: Prentice-Hall, 1981

6. Ashman R. (2004) “Project Estimation: A Simple Use-Case-Based Model” IT Professional Volume: 6, Issue: 4, pp. 40 – 44.

7.Boehm, B.W., *Improving Software Productivity*, Computer , pp 43-57, May 1987.

8. Azath H., Wahidabanu R. S. D., (2011) “Efficient effort estimation system viz. function points and quality assurance coverage” IET Software, Volume: 6, Issue :4, pp. 335–341

9. 30. L. H. Putnam, “A general empirical solution to the macro software sizing and estimating problem”, *IEEE Trans. Soft. Eng*., July 1978, pp. 345-361.

10.Benediktsson O., Dalcher D., (2005) “Estimating size in incremental software development projects” IEE Proceedings online no. 20050019, Volume: 152, Issue:6, pp253-259.

11. Berger, J., O., (1993) “Statistical Decision Theory and Bayesian Analysis”.Springer; 2nd edition Verlag, New York.

12. N. A. Parr, “An alternative to the Raleigh Curve Model for Software development effort”, *IEEE on Software Eng.* May 1980.
.

13.Boehm B, Abts C., Brown W., Chulani S., Clark B.k., Horowitz E , Madachy R.,Reifer D. j. and Steece B., (2000) “Software Cost Estimation with COCOMO II” Prentice Hall, Edition 1

14.Jones, T.C., Programming Productivity, McGraw-Hill, NY 1987.

15.Kansala, K., Integrating Risk Assessment with Cost Estimation, IEEE Software, May/June, 1997, pp. 61-68.

16. G. Cantone, A. Cimitile and U. De Carlini, “A comparison of models for software cost estimation and management of software projects”, in *Computer Systems*: *Performance and Simulation*, Elisevier Science Publishers B.V., 1986.

17 University of Calgary, California ,http://www.cpsc.ucalgary.ca/cpsc_research/areas

18.Softstar Systems (2002), Answers to Frequently Asked Questions . http://www.softstarsystems.com/faq.htm.

19. NASA JSC, Basic COCOMO Software Cost Model. http://www.jsc.nasa.gov/bu2/COCOMO.html.

20. Boehm, B., Clark, B., Horowitz, E., Madachy, R., Shelby, R., and Westland C. (1995). Cost Models for Future Software Life Cycle Processes: COCOMO 2.0, Annals of Software Engineering. http://sunset.usc.edu/research/COCOMOII/Docs/stc.pdf.

21.Boehm, B., Chulani, S., and Clark, B. (1997). Calibration Results of COCOMO II.1997.http://sunset.usc.edu/publications/TECHRPTS/1998/usccse98-502/CalPostArch.pdf.

22.Boehm, B., Chulani, S., Clark, B. (1997). Calibrating the COCOMO II Post Architecture Model.http://sunset.usc.edu/Research_Group/Sunita/down/calpap.pdf.

23. Londeix, B. (1987). Cost Estimation for Software Development. Addison-Wesley Publishing Company, New York, NY

24. Pressman, R.S. (2001). Software Engineering: A Practitioner's Approach. McGraw-Hill Higher Education, New York, NY.

25.Vigder, M. R. and Kark, A. W. (1994). Software Cost Estimation and Control. Software Engineering Institute for Information Technology.http://wwwsel.iit.nrc.ca/seldocs/cpdocs/NRC37116.pdf.

26 .Wu, L. (1997). The comparison of the Software Cost Estimating Methods, University of Calgary, Calgary, Canada.http://sern.ucalgary.ca/courses/seng/621/W97/wul/seng621_11.html.