ISSN No. 0976-5697

Volume 5, No. 1, Jan-Feb 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Merging of two Estimation Techniques will generate better results than single one

Syeda Binish Zahra
Lecturer at Department of Computer Science
Lahore Garrison University Lahore, Pakistan
s.binishzahra@gmail.com

Abstract: In last two decades different models and techniques was invented to solve the problem of estimation. But no one (inventor) gives guarantee that all estimation factors are resolve by his estimation technique. Each estimator tries to solve some factors of estimation. Now accuracy of estimation models depends on how many factors this model solve. My proposed approach of estimation is merger of two estimation techniques (Wideband Delphi and Case-Based Reasoning). My point is that when we merge two or more techniques, it will give us better results from previous ones. The above statement has been proved by my proposed estimation models and the conclusion table verifies my statement.

Keywords: Case-Based Reasoning (CBR), Function Point (FP)

I. INTRODUCTION

Accurate and timely estimation depends on how exactly, one has identified the requirements of the proposed software and understood the available resources as well as the historical data related to software estimation [1].

Software project estimation is very challenging field due to its rapidly changing nature, this made accuracy of software become very difficult in all domains. One of the most important goals for the software estimation developer is to construct a useful estimation model that accurately predict the effect, time, size and cost of software project.

a. Aim is to maximize accuracy in prediction and produce estimates that are as close as possible to the actual values.
b. The technologies and processes are changing rapidly in software development, which is affecting the performance and behavior of software development and their significances. [2]

If we properly manage time, resources, cost, budget, and schedule that software project estimation gives 95% accuracy.

A. Case-Based Reasoning (CBR):

Case-based reasoning also known as Estimation by analogy [3], is an estimation approach that use past cases to make new project reliable. This approach is reliable and beneficial because its performance is much better than algorithmic model and extremely near the approach of expert.

First CBR system stored cases in a database. This database provide extreme base of knowledge that is useful for the developer of projects. Developers use this knowledge base to retrieve the most similar case with their new project. Another advantage is that results of CBR systems are more accurate than any Expert Judgment approach [4].

CBR system depends on four steps (also known as CBR-cycle) [5, 14]

Retrieve*:* get similar cases from database
Reuse: solution recommended with these similar cases.
Revise*:* fit into new case
Retain*:* new complete case put into database

a. ***Advantages:***

a) Requirement of expert is null
b) Provide akin to thinking of human
c) Accurate prediction to handle cases even failure
d) Provide efficient reasoning [6]
e) Allowing faster knowledge acquisition [6]
f) Provide unique explanation capability

b. ***Limitation:***

a) Case data hard to gather
b) Limited prediction about cases

B. Wide-Band Delphi:

Delphi technique was developed at Rand Corporation in 1969 and in 1981 Barry Boehm refines and renews the concepts of Delphi as Wide-band Delphi. This technique estimate schedule, effort and plan. Specifications are used as input and assumption about estimation, detailed task list and effort given as output. The whole process divided into six steps [7]. These steps are as follow:

a) Planning
b) Kickoff meeting
c) Individual Preparation
d) Estimation Meeting
e) Assemble Tasks
f) Review Results

In this technique a team of experts is selected for software development. This team consists of 3 to 7 member with a person who plays the role of moderator [1]. The whole technique is divided into 6 step and these steps are conducted in a sequence. Each team member interacts with each other because Wide-band Delphi incorporates much interaction and communication among participants. Team member's works individually and show their work at meeting that conducted in rounds.

Firstly all work is assigned to team members individually. In first round, results are gathered in tabular form from each participant and then these results returned to each participant for second round. New problem arise from first collected data and participants try to answer all the problems that arise at previous round. The round system will continue until all results are accurate and problem is shoveled [8].

© 2010-14, IJARCS All Rights Reserved

This technique is useful only when participants are experts in their decisions and able to give more "expert's opinion" about the problem. A person is said to be "experienced" in particular field when he spend many of years in that field, but also it gives no guarantee of future knowledge because new needs and requirements.

a. Advantages:

a) Relatively simply process and useful in the absence of historical data.
b) Communication among participants at each stage that ensure estimate are not over works so every point is discuss and final output in much accurate instead to individual estimation [1]
c) Remove politics

b. Limitations:

a) Depends on required management co-operation
b) Team member should be agreed on agreement
c) Team members should be experienced and able to share their ideas clearly
d) Time- consuming as many participate involved [4]
e) Useless when any of participant is absent.
f) Peoples that are not interested, was forcedly because they consider being experienced.
g) Expensive method
h) Sometime difficult to coordinate and motivate a group of experts with diverse interest and busy schedules.

II. RESEARCH APPROACH AND FRAMEWORK

Input:

The problem related to software project management is entered into this system for estimation. That problem is going to become the input of that estimation model.

A. Estimation Process:

Step 1:

a. Select a Team:

Project manager select a team [9] including moderator, and 5 to 7 persons that perform job of estimator (not necessary they all are experienced).

Step 2:

a. Meeting for Planning:

This session begins by project manager with defining and scoping the problem. The problem is broken into small manageable parts. The splitting can be either by features or by phases of project [9]. The purpose of breaking problem is to estimate more accurately. It considers being beneficial and thumb of rule to break the project into 10 to 15 parts. This is the responsibility of project manager to break the problem very carefully. If break part is not clear then it became more costly and time-consuming to correct them. The participants included in estimation are moderator (who schedule and plans meeting), project manager and two to five other estimators. Project manager selected estimation team and assign work individually to estimators. The selection of estimation team is depend on estimator's interest as well as who are able to predict. To inform all the participants about schedules of meeting is the responsibility of moderator.

Step 3:

a. Initial Estimation:

Basic process of estimation is started at that step [7]. At that step responsibility of moderator is to explain estimation requirements to participants who are unfamiliar with that. The responsibility of each participant is to develop his / her task independently. The participant takes his / her work and compares their work with the cases that are already placed in case library. This library contains cases that taken from completed projects [10].

b. Retrieve and Reuse:

a) Retrieve: Each participant gather his / her data which is similar to new case from database.
b) Reuse: After retrieving the similar cases, participants reuse that similar case to solve the problem.

Each participant use case-library to find similar projects. By pervious project results, estimator can analyze the proper / actual needs of data. These results give confidence to estimator and they can convey their knowledge more accurately by showing their work results rather than describe their troughs in words. In the case if required data is not available in the database / library then participants should make their assumptions about the project. Each participant must write their assumption to convey the thoughts to other participants and also to stakeholder.

Step 4:

a. Estimation Meeting:

Estimation meeting is conducted by moderator to collect all the individual estimated results by each participant. Thus different estimation results are collected by participants. Each participant discusses their points or knowledge about their individual task and gives their assumptions. Also discuss issues about estimation and highlight questions they have related to problem [11]. The advantage of that point is that each participant can effectively communicate their knowledge and also able to given reasons about their estimation because of database cases usage. At this stage responsibility of moderator is to encourage all the team members and maintaining a friendly and impartial environment [11].

b. Revise:

Estimation meeting collect all the data from each participant and combine it. Then all participants revise the proposed solution by merging all participants' data. If there is any problem in finding solution or inappropriate results, then whole problem is re-estimate through 'Retrieve and Reuse' step. All estimators should define their work in fine form to reduce confusion that may be occurring at merging stage.

Step 5:

a. Assembling Task:

The moderator and the project manager assemble all the project work (that each member conducted in their

© 2010-14, IJARCS All Rights Reserved

individual work) into a single large work. When whole work is merged into single task then all duplicate tasks being eliminate. The purpose of eliminating duplication is to reach at some reasonable solution that gathers from different estimates of individual participants [7].

Step 6:

a. Show Results:

This final step is completed into two phases.

a) Estimation reviews the results (summarized work) and reaches on final outcome. Moderator together all the team members for 30-to-60 minute to review the estimation activity and show results to all involved participants.
b) This is the responsibility of moderator as well as project manager to 'Retain' the new project into database for future need.

B. Proposed Technique Strengths:

a. Relatively simply process and also useful in the absence of historical data.
b. Communication among participants at each stage that ensure estimate are not over works so every point is discuss and final output in much accurate instead to individual estimation
c. Remove politics estimation is not based.
d. Requirement of expert is decreases
e. Provide akin to thinking of human
f. Accurate prediction to handle cases even failure
g. Provide efficient reasoning
h. Wide range for prediction about cases
i. Allowing faster knowledge acquisition
j. Provide unique explanation capability
k. Depends on required management co-operation
l. Team member should be agreed on agreement
m. Not Time- consuming
n. Still useful when any of participant is absent.
o. People that no interested, was not forcedly
p. Cheap method (in cost, in effort)

III. CASE STUDY

A. Case Study: Best Case :

This case study is about "**Ericsson**" which is the biggest distributed telecom system.

The software of that telecom system contains several languages (C, Java, Perl and Erlang). The software size of latest release (2005 release) is more than 450 KSLOC (Kilo SLOC) or about 1000 KSLOC measured in equivalent in C code.

a. Approach:

As we see there are lots of common features in handset of 2001 and 2012. 2001 was the first time when Colour screen, Bluetooth and SMS with T9 featured mobile phone was launched by Ericsson. For that particular time Ericsson might suffer from high cost of research and development but after that particular model almost every handset of Ericsson (later branded as Sony Ericsson) contain Colour screen, Bluetooth and SMS with T9 feature.

Every upcoming handset model usually contains more features than its previous one. and this is the need for the companies for their survival in this fast moving market because customer demand something new special when there is place for it .and if in production of every new handset Ericsson start to reproduce all those features which are already been exercised then the cost of production (also called cost-of-sales in some companies) go's higher and higher. This is not friendly policy for company because high cost means high sale price. however new model contain many new features but it also contain previous features .so, reproduction of previous features should not performed in fact as they were successive projects so they should be as it merged and that cost should saved or use in research of new features rather than using for sunk purposes.

b. Results [12]:

a) In 2001 Ericsson launches its first handset "Ericsson T68" (simply T68) with the price of 500 € (57500/-Rs)
b) That set have following features:
c) Color screen
d) Passive LCD with 101 x 80 and 256 colors
e) Bluetooth, IrDA port
f) GPRS 3 + 1
g) Tri-band compatibility
h) SMS with T9, EMS, WAP
i) Customizable monophonic ring tones
j) Phone cost is 500 € (57500/-Rs)

"Ericsson" reuse its components about 60 % and 40 % of its components are new. This case study supports my approach, by using my approach (reusing); it will decrease the overall cost as well as time. First set is too costly when it compares to latest mobile sets. First set is 57500/-Rs with limited features and today (2012) set price is 39000/-Rs with lot of features and support all communication linkage. Ericsson official survey show that their quality increase after reusing their components and their cost and time is decreases.

B. Case study: Average Case :

GAO, Customs Service Modernization: Serious Management and Technical Weaknesses Must Be Corrected, GAO/AIMD-99-41 (Washington, D.C, Feb. 2004)

GAO analyzed the U.S. Customs Service approach to deriving its $1.05 billion Automated
Commercial Environment life-cycle cost estimate with Software Engineering Institute (SEI) criteria. SEI had seven questions for decision makers to use in assessing the reliability of a project's cost estimate and detailed criteria to help evaluate how well a project satisfies each question. Among the criteria were several very significant and closely intertwined requirements that are at the core of effective cost estimating. Specifically, embedded in several of the questions were requirements for using (1) formal cost models;
(2) structured and documented processes for determining the software size and reuse inputs to the models; and (3) relevant, measured, and normalized historical cost data (estimated and actual) to calibrate the models.
GAO found that Customs did not satisfy any of these requirements. Instead of using a cost model, it used an unsophisticated spreadsheet to extrapolate the cost of each Automated Commercial Environment increment. Its approach to determining software size and reuse was not documented and was not well supported or convincing. Customs had no historical project cost data when it

© 2010-14, IJARCS All Rights Reserved

developed the $1.05 billion estimate and did not account for relevant, measured, and normalized differences in the increments. Clearly, such fundamental changes can dramatically affect system costs and should have been addressed explicitly in Customs' cost estimates.

C. Case study: Worst Case:

GAO, NASA: Lack of Disciplined Cost-Estimating Processes Hinders Effective Program Management, GAO-04-642 (Washington, D.C, May 28, 2004) Units:

GAO found that the National Aeronautics and Space Administration's (NASA) basic cost estimating processes-an important tool for managing programs-lacked the discipline needed to ensure that program estimates were reasonable. Specially, none of the 10 NASA programs GAO reviewed in detail met all GAO's cost estimating criteria, which are based on criteria Carnegie Mellon University's software Engineering Institute developed. Moreover, none of the 10 programs fully met certain key criteria-including clearly defining the program's life cycle to establish program commitment and manage program costs, as required by NASA.
In addition, only 3 programs provided a breakdown of the work t be performed. Without this knowledge, the programs' estimated costs could be understated and thereby subject to underfunding and cost overruns, putting programs at risk of being reduced in scope or requiring additional funding to meet their objectives. Finally, only 2 programs had a process in place for measuring cost and performance to identify risks.

In study about the problem of cost estimation by NASA, for its projects which are ten in number of. The main issue was that the cost of the projects was not correct, but in my view it was obvious to be; because they even don't know about the life cycle of the projects. When life of the project is uncertain then budget got upset. As we know one of the major budgeting techniques is life cycle costing, which is widely used in budgeting. And if we don't know the project life cycle how can we allocate funds to it. In this situation allocation of funding is also uncertain which will either understating or overstating of funds.

But one think is noticing at this particular point which that in the first project we were might first on this type of project but in second one we were able to made a estimate of cost for the project latest better than the first one but for this propose we should use the statistics of first project. And in the same way in third project; the estimate can be more reliable because now we have two previous experiences. In the mean while at the tenth project we have an experience of nine cases to utilize but if we want to utilize them.

If when the second or third project rises they should mend their direction in the way that they should split component of the project in their definable parts. Estimate for these parts should be made and estimates then should be compare with the estimates of the pervious projects and the variance (favorable our adverse) then should be remove by applying the previous experience about specific component of the present project.

IV. CONCLUTION

Table 1.1: Proposed Technique Comparison with Existing Wide-Band Delphi and Case-Based Reasoning

Parameters	Delphi	Case-Base Reasoning	Proposed Technique
Accuracy	Depend on estimator's skills	Good	High
What can estimate	Work break down structure, Effort	Effort	Time, Cost, Effort
Applicable system	Agile development	Data-Base, CBR-DSS, SQUAD	Data-Base System, Agile Development
Available Tools	Delphi freeware	CASPIAN, Remind 1.3 [13]	
Input	Vision and scope documents	Set of Cases	Set of Cases
Applicable Phase	Task Level	CBR-Cycle	
Reliable	Yes	Yes	Yes
Support communication	Yes	Yes	Yes
Reusability	No	Yes	Yes
Support historical Data	No	Yes	Yes
Requirement of expert	High	Null	Low
Provide Reasoning	No	Yes	Yes
Time Consuming	Yes	Yes	No
Cheap method	No	No	Yes

This table and following explanation ensures that the Proposed Estimation Model is more reliable and accurate when it compares with existing approaches. Proposed model support

- ***a. Reusability:*** As it match new problems with previous cases and reuse them.
- ***b. Support communication*:** Each participant can communicate with other fellows, as many meetings conducted. The purpose of meeting is that each participant easily discusses his / her problem and point of view with their co-members.
- ***c. Support historical Data:*** This estimation model use a case library so it supports historical data.
- ***d. Requirement of expert:*** The requirement of expert is low because if one or two expert is not available due to some reason, we can hiear new person. Reason is that we use a library; if a person is not expert then he is able to get guide line from previous cases and able make his own decisions.
- ***e. Provide Reasoning:*** This model enables the estimators to explain their work with proper reasoning. This is natural problem that a person cannot explain his \ her thoughts to other persons, but if they have solid proves then they are explain their point-of-view easy as well as others can understand easily.
- ***f. Time Consuming:*** When we known all or some aspects of problem then we are able to manage our time more easily and eliminate extra time slots.

© 2010-14, IJARCS All Rights Reserved

g. Cheap method: As case-studies show if we reuse the parts of past projects then we are able to decrease our time as well as cost and effort.

V. REFERENCES

[1] Samaresh Mishra, Kabita Hazra, Rajib Mall, "A survey of Metrics for Software Development Effort Estimation" In proceeding of International Journal of Research and Reviews in Computer Science, Vol. 2, No. 5, October 2011, ISSN: 2079-2557.

[2] Anatol Siamionau, "Requirements estimation guideline for IT projects at Scania Networks AB", Master of Science Thesis, Stockholm, Seden 2011, TRITA-ICT-EX-2011:168

[3] Stephen G. MacDonell, Martin J. Shepperd, "Combining techniques to optimize effort predictions in software project management", In proceeding of Journal of System and Software 66 (2003) 91-98

[4] Abdulbasit S. Banga, "Software Estimation Techniques", In proceeding of National Conference (INDIACom-2011), Computing for Nation Development, New Delhi March 10 – 11, 2011.

[5] Pascal Ziegler, "Manual Techniques, Rules of Thumb", seminar on software cost estimation, WS 2002 / 03

[6] Li D. Xu, "Case-Based Reasoning: A major paradigm of artificial intelligence", In proceeding of IEEE, 1994

[7] Karl E. Wiegres, "Stop Promising Miracles", In proceeding of process impact, published in Software Development, Februray 2000.

[8] Dr. Barry Boehm, Sunita Chulani, "Software Development Cost Estimation Approaches" A Survey, 1998, PhD Thesis, Computer Science department at USC, university of Southern California

[9] Payal Jain, Rajneesh Kansal, Er. Zatin Gupta, "Estimation making Software Development More Effective", In proceeding of Grobal Journal of Computere Science and Technology, Vol. 11,Issue 2, Version 1.0, February 2011.

[10] Mel Damondaran, Aqua Netta E. Washington, "Estimation Using Use Case Points", In proceeding of ISECON 2002, San Antonio, 2002

[11] Hareton Leung, Zhang Fan, "Software Cost Estimation", in handbook of software engineering and knowledge engineering, world scientific pub. co, river edge, NJ, 2002

[12] Ericsson reports fourth quarter results January 25, 2011

[13] Mel Damondaran, Aqua Netta E. Washington, "Estimation Using Use Case Points", In proceeding of ISECON 2002, San Antonio, 2002

[14] I. Watson, "Case-Based reasoning is a methodology not a technology", in proceeding of Elsevier Science B.V, Knowledge-Based Systems 12 (1999) 303-308.

© 2010-14, IJARCS All Rights Reserved