

Study on Reliability Improvement for the Acquisition of Web Application among Government Agencies

Kwoung Hee Choi

R&D Evaluation Center
Korea Institute of S&T Evaluation
and Planning
Seoul City, ROK

Sung Yul Rhew

School of Computing
Soongsil University
Seoul City, ROK

Abstract - Software reliability prediction is a statistical method to put in place a timely software development practice useful for objective assessment of bidders. The current study suggests one research method that enables reliability assessment on such previous projects by studying user satisfaction and the amount of delay in delivery time. If incorporated into the existing acquisition process, the reliability assessment method will further enhance objectivity and accuracy in bidder selection process. GQM(Goal Quality Method) paradigm was used to identify assessment metrics for bidder evaluation and questionnaires were collected from users to create user satisfaction index. In addition, 'weight of evidence', the most appropriate categorical method, was used to isolate attributes of each variable that may contribute to reliability assessment. Also, regression analysis was used to come up with reliability metrics model.

Keywords: Software Reliability, Assessment of Bidder, GQM, Weight of Evidence

1 Introduction

The Software Development Plan 2006 for Government Agencies has earmarked government budget for 250 projects. Out of the 250 projects, 96 projects or 38.4% is worth less than 100 million KRW, and 74 or 29.6% is worth between 100 million to 300 million KRW. Also, 68% of the total is small or mid-sized projects worth less than 300 million KRW while more than 96% of the project developments depend on complete outsourcing or partial outsourcing [1].

Outsourcing partners are screened based on technical proposals submitted which will be assessed against an Evaluation Standard for Software Technology[2]. It is appropriate to employ an evaluation process for large projects only with written proposals as most of bidders are market-proven players. However, bidders for small and mid-sized, short-term projects are mostly small and mid-sized companies and thus, it is required to assess the reliability of the previous projects they have conducted in a practical manner.

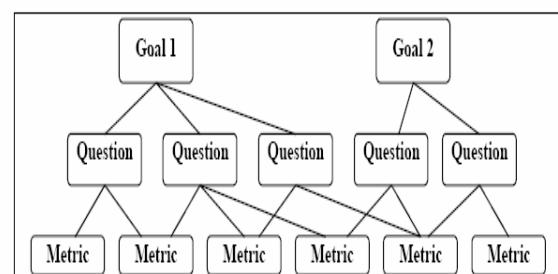
The current study suggests one research method that enables reliability assessment on such previous projects by studying user satisfaction and the amount of delay in delivery time. If incorporated into the existing acquisition process, the reliability assessment method will further enhance objectivity and accuracy in bidder selection process.

GQM(Goal Quality Method) paradigm was used to identify assessment metrics for bidder evaluation and questionnaires were collected from users to create user satisfaction index. In addition, 'weight of evidence', the most appropriate categorical method, was used to isolate attributes of each variable that may contribute to reliability assessment. Also, regression analysis was used to come up with reliability metrics model.

2 Related Research

2.1 GQM Paradigm

The GQM Paradigm is based on the idea that measurement should be goal-oriented; i.e., all data collection should be based on a rationale that is explicitly documented[7].



[Figure 1] GQM paradigm

GQM-based measurement programs should be planned and performed according to the following principles:

1. The analysis task to be performed must be specified precisely and explicitly using a detailed measurement goal;
2. Metrics must be derived in a top-down fashion based on goals and questions;
3. Each metric must have an underlying rationale that is explicitly documented;
4. The data that are gathered for the metrics must be interpreted in a bottom-up fashion using the GQM goal and questions; and
5. The people from whose viewpoint the measurement goal is formulated must be deeply involved in the definition and interpretation of the measurement goal.

Two techniques are useful for developing GQM plans. First, goal templates assist in generating a GQM goal. The template identifies five major aspects, namely the object, purpose, quality focus, viewpoint, and environment of a measurement program.

Second, abstraction sheets are defined to assist in collecting the information necessary to build a detailed GQM plan.

A GQM abstraction sheet is a document, often a single sheet of paper that helps elicit and structure information during an interview and assists in constructing, refining, and reviewing a single GQM plan.

2.2 Metrics on Software Reliability

Software reliability represents a user-oriented view of software quality. Analysis on software reliability is conducted to put in place a timely software development practice via statistical metrics. Reliability analysis consists of prediction and estimation depending on analysis time horizon. Prediction refers to activities taking place from project planning to the phase right before tests and it utilizes empirical data set to find the number of defects during these phases before testing.

Estimation refers to activities from integration to acceptance test and it utilizes operation profile –the number of defects during observation- to calculate estimated errors as well as associated cost at the point of software delivery and afterwards.

[Table 1] Prediction versus Estimation

	Prediction	Estimation
Purpose	To determine in advance What Practice must be employed(How)	To Determine when it is OK to ship(When)
When they are used	Anytime but usually before system testing	Exclusively after the start of system testing
Data Collected	Empirical or historical	Observed failures
Entity measured	Defect density	Current failure rate

3 Reliability Metrics Design

GQM paradigm was used to identify metrics to assess bidders' reliability. By use of Object Template, We can derive detail goal.

[Table 2] Reliability Object Template

	Major aspects
Object	Project Results
Purpose	Prediction
Quality Focus	Reliability
View Point	Acquisitor
Environment	HR, Methodologies

Analysis targets include projects that last under 6 months with less than 300 million KRW. For project reliability variation factors, user satisfaction and the amount of delay before project stabilization were selected.

Also, abstraction sheet was used in order to identify factors in detail that are associated with the quality of objective templates found. Abstraction sheet is mostly a one page document and is useful in clearly delivering information on analysis target during interviews or researches. [Table 3] shows an example of abstraction sheet to articulate research objectives and to review the process.

[Table 3] Reliability Abstraction Sheet

Quality Focus User satisfaction Project delay rate	Variation Factor User satisfaction -Convenience/Functionality (Use IBM Questionnaire) Project delay rate - PM, HR management, Use of methodologies
Baseline Hypothesis User satisfaction -3 or more agree among 5 questions Project delay rate -PM ability, Timely HR input, and use of Risk Management decrease delay time	Impact on Baseline Hypothesis High Satisfaction -Low fault and delay rate Low Delay rate -Competent PM, Stable Human management, Proper risk management

Lastly, it is required to identify a set of questions that can completely define variation factors of reliability abstraction sheet and basic assumptions. Once metrics is designed in line with questions, consistency is ensured for the entire analysis process as each variable will have higher correlation with the research objectives. [Table 4] illustrates a set of questions and metrics based on variation factors of the abstraction sheet and basic assumptions.

[Table 4] Questions and Metrics

Objectives	Metrics on the Reliability of Outsourcing Partners from Acquisitor’s Perspective	
Question	Q1	What are the factors for reliability?
Metrics	M1	User satisfaction
	M2	Project delay rate
Question	Q2	How to measure improvements in reliability (%)?
Metrics	M3	<u>Reduction in Delay as a result of Improvements</u> Initially estimated delay
Question	Q3	How to calculate satisfaction level?
Metrics	M4	1 Overall, I am satisfied with how easy it is to use system
		2 I am able to complete my work quickly using system
		3 It is easy to find the information I needed
		4 I like using the interface of system
		5 system has all the functions and capabilities I expect it to have

To measure user satisfaction which is highly subjective, such methodologies as user experience research and feasibility test on software quality are available. Among these methodologies, user experience research is used in general and the current study has selected 5 items that suit web application from CSUQ(Customer System Usability Questionnaire) developed by IBM[5].

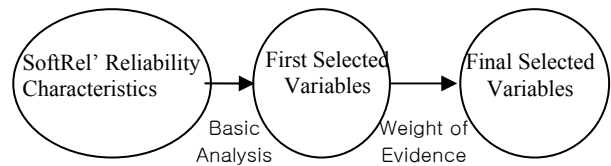
4 Reliability Prediction

Reliability prediction is based on information gathered from questionnaires that are incorporated with identified metrics and distributed to previous and current clients of the bidder.

4.1 Sampling and Variable Selection

37 samples were gathered out of interviews and questionnaires sent to 60 government agencies in order to do research on the results of outsourced projects. Government agencies here refer to a special corporate body that is mandated by relevant law and that is authorized and commissioned by the government to do tasks. Among the 37 samples, 21 were successful while 16 were not successful in terms of user satisfaction and the period required for project stabilization.

Variables for the research are 28 in total: 17 HR-related information and 11 methodologies. To come down with finalized variables for modeling, processes illustrated in [Figure 2] were followed.



[Figure 2] Variable select process

Variables to be used for reliability assessment were drawn from a set of 148 variables developed by SofeRel, a specialized reliability research agency, based on its 15 year study regarding software reliability[3]. The variables can be classified into 5 categories as in [Table 5].

[Table 5] Categories SofeRel' Characteristics

Category	Content
Program Type	Usage of Program
Methodologies	Develop, Test & Verification methods
Work Profile	Fault history management
HR	Education of HR, Knowledge of Work
Product Characteristics	Product and Domain Complexity

As the research targets web application at the stage of acquisition, 28 variables were selected for the first round of study in terms of methodologies and HR.

Among the selected variables for the first round, Information Value was used to see if the attributes of the variables are appropriate to show its impact on reliability.

$$(p_{ij} - q_{ij})w_{ij} \quad (1)$$

note, p_{ij} = the number of good quality with attribute j from a variable with i attribute /total number of good quality

q_{ij} = the number of defective quality with attribute j from a variable with i attribute/total number of good quality

$$w_{ij} = \ln(p_{ij}/q_{ij})$$

In general, the relevant variable is selected for the model if information value that exceeds 0.1. In information theory, evidence weight is an index that represents the importance of specific information and it also shows how much information is provided in determining “reliable” or “non-reliable” for reliability assessment. The evidence weight that is greater than 0, means that reliability rate is higher than non-reliability rate.

The selected variables for the first round are sorted out depending on their attributes and on evidence weight. [Table 6] show selected variables that have information value exceeding 0.1.

[Table 6] Summary on Finalized Variables by Category

Variable	Attribute	Defect Rate	Evidence Weight	Information Value
Bidder's Risk Management	Good	5	0.64	1.65
	Not Good	50	-2.35	
Timely HR Input	Timely	29	1.05	0.90
	Delayed	25	-0.97	
Agreement on IA and Design	In Advance	24	0.71	0.74
	Some Time	38	-0.97	
	Later			

4.2 Reliability Model

To test the relationship between finalized variables and project outcomes, a regression analysis was conducted with the success of the project as dependent variable. [Table 7] shows result of regression analysis.

[Table 7] Summary on Finalized Variables by Category

	B	Std Dev.	Beta	t	p
Constant	-0.094	0.144		-0.651	0.519
Risk Management	0.422	0.159	0.365	2.650	0.01
Timely HR Input	0.313	0.135	0.316	2.319	0.02
Analyzed Scope	0.305	0.131	0.302	2.317	0.02
Model	R	R ²	Modified R ²	SE	p
Reliability	0.674	0.454	0.405	0.387	0.0001

As the impact of finalized variables on project deliverables was tested, the following regression model can be derived to explain project reliability.

$$Y = -0.094 + 0.422 V1(\text{bidder's risk management}) + 0.313 V2(\text{timely input of HR}) + 0.305 V3(\text{analysis on development scope}) \quad (2)$$

The model shows that the most important variable in assessing reliability of small/mid-sized, short term projects is the timely input of human resources if other conditions are the same. It was also found that development scope analysis at the beginning of the project, should be accurate and that bidders should cope with risk situations in an appropriate manner.

The score of bidder's reliability can be classified into 3 levels as in [Table 8].

[Table 8] Level of Bidder

Level	1	2	3
Score	-0.094~0.328	0.328~0.641	0.641~0.946

4.3 Model Evaluation

To measure the fitness of model, following [table 9] can be used.

[Table 9] Classified Table

		Classified Group		
		Bad	Good	Sum
Actual Group	Bad	12	4	16
	Good	2	19	21
	Sum	14	23	37

Sensitivity: The percentage of classifying the actual Good Group as good.

$$(19 \div 21) \times 100 = 90\%$$

Specification : The percentage of classifying the actual Bad Group as bad.

$$(12 \div 16) \times 100 = 75\%$$

Accuracy : Probability of classifying the Good Group and Bad Group correctly.

$$((12 + 19) \div 37) \times 100 = 84\%$$

5 Conclusions

Statistical assessment can serve as an objective tool to assess bidders whose reliability has yet to be objectively measured for acquisition projects of government agencies. Based on the most optimal categorical method of 'evidence weight', variables were selected to draw a line between 'reliable' and 'non-reliable' and furthermore, regression analysis was employed to construct a simple model.

In future, it is required to find additional variables that have higher applicability by doing more detailed case studies for diverse projects.

6 References

[1] "2006 Software Demand Survey," Korea Software Industry Association, 2005.

[2] "Evaluation Standard for Software Technology," Ministry of Information and Communication, 2004.

[3] Ann Marie Neufelder, "Ensuring Software Reliability," Marcel Dekker Incorporated, ISBN 0-8247-8762-5

[4] Michael R. Lyu, "The Hand Book of Software Reliability Engineering," McGraw-Hill, ISBN 0-07-039400-8, 1995.

[5] Lewis, J.R., IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use, *International Journal of Human-Computer Interaction*, Vol. 7(1), pp. 57-78, 1995

[6] Standish Group International, Inc., Extreme Chaos, 2001.

[7] Christian Differding, Christopher M. Lott, Technology Package for the Goal Question Metric paradigm, Internal Report, 1996.

[8] Sun Hum, Paek, "Reliability Engineering," 2001.

[9] J. Y. Song, K. H. Choi, J. S. Jang., Aging test and software reliability analysis method for pc-based control,

Proceedings of the 2005 International Conference on Simulation and Modeling, 2005.

[10] Zhang, X. and Pham, H. Comparison of NHPP Software Reliability Models, *Proceedings of the Forth ISSAT*, 1998.