

# **Software Quality Evaluation of Email Management Systems**

**Wonil Kwon**

Software Quality and Evaluation Center

ITTL (Information & Telecommunications Testing Laboratory)

TTA (Telecommunications Technology Association)

267-2 Seo Hyeon-dong, Bundang-gu, SeongNam-si, Gyeonggi-do, Korea

wonil@tta.or.kr / wonil@icu.ac.kr

**Chang-shin Chung**

Software Quality and Evaluation Center

cschung@tta.or.kr

**Sang-duck Lee**

Software Quality and Evaluation Center

sdlee@tta.or.kr

## **Abstract**

The use and the importance of software in various industries including communications area are increasing rapidly while the measurement and the improvement of the software quality are very difficult and unclear. It is because software is intangible and standards are mostly not in place unlike hardware.

An Email Management Systems (EMS) software product is evaluated by testing its functions (behaviors) and checking end-user satisfaction based on international standards. One of the core parts of the evaluation is to find out the weight values of the evaluated items (lists) which are the main focus of this paper. This portion of the overall evaluation has been conducted by the survey of 23 real world users. The weight value is intended to be reusable for the evaluation of the EMS-like software in terms of its usage and functionality.

## **1. Introduction**

For the improvement of software quality, evaluation of the software products is essential to find out where to improve. Various studies have been done in order to develop appropriate and concrete evaluation methodologies and models. One of those approaches is to appraise the target software by testing functions (behaviors) of it or/and checking end-user satisfaction [1, 4, 5, 6] based normally on international standards [3] in the specific categories of software. Those evaluation methodologies have got checklists to be assessed with different weights (priority or importance) [4] for each list which is multiplied by the value after evaluation (of the test cases). This results in the realistic and meaningful evaluation values.

The evaluation model in the form of formula has been generated based on the evaluation

methodology above and the survey of 23 users of a specific software category, EMS(Email Management Systems), so as to evaluate it from the perspective of end-users.

$$E = \sum_{i=1}^n a_i \cdot \frac{U_i}{3} \quad (\text{where, } E = \text{Evaluation value of quality characteristics, } a = \text{weights or}$$

priority (Minimum value, Most likely value, Maximum value),  $U$  = User rated value (0,1,2,3),  $n$  = Evaluation list or item number, 3 for the readability of the  $E$ ) has been used to evaluate the quality characteristics such as functionality, usability, maintainability, etc. of the EMS and all the values of  $a_i$  and  $U_i$  have been derived from the end-user survey.

## **2. Evaluation of an EMS (Email Management Systems)**

The Multilateral Approach for Software Evaluation (MASE) is used during the course of the evaluation. MASE evaluates software in the perspective of user perception, functionality, and user scenarios. The user perception to the evaluated software is measured based on the user survey. The functionality is thoroughly tested by a testing team based on the clearly defined ‘Evaluation Module’ and testing procedure. The user scenario is developed by a group of users, evaluators, and the vendor. The testing experts also test it along the functionality testing. [8]

Of the 3 aspects of MASE, we are only interested in the user perceived evaluation of the software. Other parts of the evaluation methodology will be briefly introduced. However, we will focus on the user perceived factors by analyzing user survey.

### **2.1. Introduction of Tested Software**

The evaluated software is an e-mail management solution, which manages online customer channels in an integrated way in order to answer customer questionnaires conveniently and immediately. It includes the tools for automatic email response, managing simple and repetitive questions, and real-time text and voice chatting messengers.

### **2.2. Evaluation Method**

The user perceived evaluation of the software is one of the 3 important evaluation perspectives in the MASE as mentioned above. In the MASE Software Evaluation Framework (Figure 1), only left-hand side is applicable in this paper.

The evaluation of the user perception needs the user group, survey form, and the analysis of the survey results. The user group is the users who actually use the software in their job, for fun, and in other purposes. The vendor usually has their customer lists that are used for the selection of the user group. The survey form is a modified form of the functionality evaluation module [7].

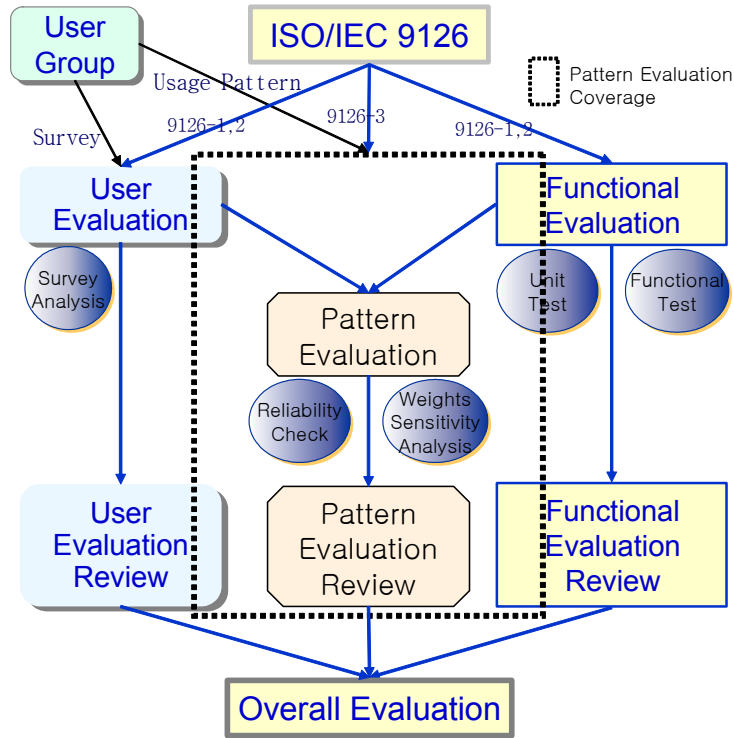


Figure 1. MASE Software Evaluation Framework

### 2.3. User Evaluation by Survey

The survey form is modified from the functionality evaluation module used in this overall evaluation in order to be understandable for the general software users. Functionality evaluation module is based on the ISO/IEC 9126-1, ISO/IEC 9126-2, and ISO/IEC 12119 [1], [3]. The format of the results of the survey analysis will reflect the software quality characteristics defined in the ISO/IEC 9126-1. The survey form asks the importance of the sub-characteristics and the evaluation marks based on their everyday use.

Survey responders are the actual users of the tested software working in Lycos Korea, Asiana Airline, Yahoo! Korea, etc. Those companies use the email heavily for the customer management. The vendor recommended them and 23 of the responds were collected and analyzed.

The evaluation can be expressed in the form of formula below.

$$E = \sum_{i=1}^n a_i \cdot \frac{U_i}{3} \dots\dots\dots (1)$$

where,

$E$  = Evaluation value of quality characteristics

$a$  = Weights or priority (Minimum value, Most likely value, Maximum value),

$U$  = User rated value (0,1,2,3)

$n$  = Evaluation list or item number

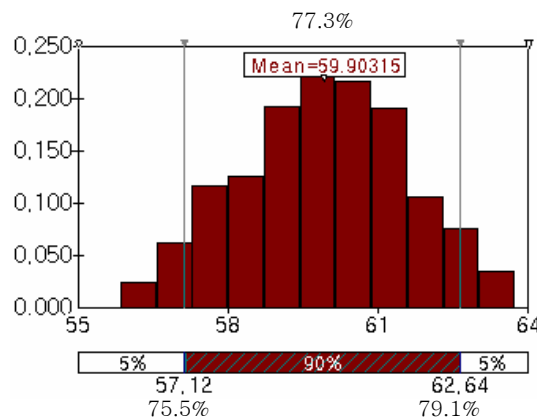
3 : It is used for the readability of the  $E$

The values of  $a_i$  and  $U_i$  have been derived from the end-user survey to evaluate the quality characteristics such as functionality, usability, maintainability, etc. of the EMS. Here the value of  $a_i$  can be reused for the next evaluation by deriving the value of  $U_i$  from the black-box testers with the assumption that they are almost end users through the course of testing.

The weight value is derived by averaging the weigh values the users have given to the evaluation items (quality sub-characteristics). One each maximum and minimum value has been excluded to minimize the abnormality of the survey results.

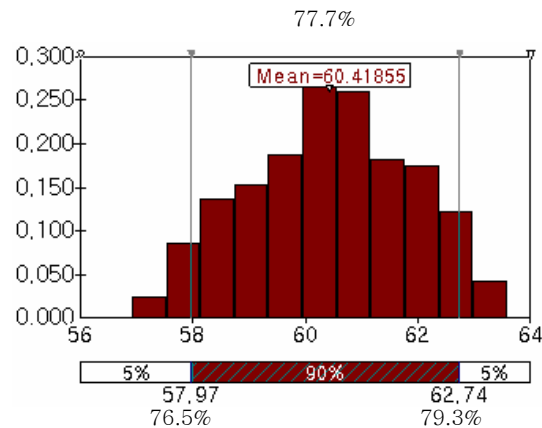
The user rated value is 4 scaled (0, 1, 2, 3 – from very bad to very good). The maximum evaluation value is given when the weight value is 100 and the user rated value is 3. It is actually impossible to get the maximum evaluation value, 100 since the weight value is given by ranged values. The overall evaluation value can be derived by applying these evaluation values to the formula (1) above.

Another overall evaluation value is derived using the Monte Carlo simulation. The result is shown in the Figure 2 below. The averaged overall evaluation value is 57.12~62.64 with the average of 59.90 within 90% of confidence [5].



**Figure 2. Averaged Overall Evaluation Value**

The functionality evaluation shown in Figure 3 below as an example of evaluating the quality characteristics individually results in 57.95~62.74 with the average value of 60.42 within 90% of confidence [3].



**Figure 3. Evaluation Value of Functionality**

The reliability evaluation results in 59.86~64.85 with the average value of 62.40 within 90% of confidence. The efficiency, usability, maintainability, and portability evaluations result in the ranged value of 58.2~63.55 with the average value of 60.87, 56.9~62.74 with the average value of 59.84, 50.7~55.89 with the average value of 53.32, and 60.09~66.65 with the average value of 63.48 respectively within 90% of confidence. The evaluation value of maintainability is relatively low and that of portability relatively high.

## 2.4. Implication from the Evaluation

The weight (or priority) values shown below in Table 1 will be beneficiary to the software evaluation organization and companies when evaluating similar software products by reusing and benchmarking them. The software companies that are to develop EMS related products would have the chance to reflect what the end-users think important upon their quality assurance or control, customer supports, and the development of new products.

**Table 1. Weight Values of an EMS**

Quality Characteristics	Sub-characteristics	Minimum	Most Likely	Maximum
Functionality	Suitability	72.33	76.89	78.33
	Accuracy	75.00	81.28	86.00
	Security	78.82	83.53	88.41
Reliability	Fault tolerance	79.61	84.72	89.33
	Recoverability	78.89	83.78	88.50
Efficiency	Time behavior	73.89	80.17	85.11
	Resource utilization	74.17	78.89	83.56
Usability	Understandability	75.56	82.39	86.89
	Learnability	73.61	80.83	86.06
	Operability	73.22	78.33	83.83
	Attractiveness	68.33	74.11	80.67
Maintainability	Analyzability	73.82	78.94	83.94
	Changeability	76.00	81.67	86.67
	Stability	69.83	76.06	82.11
Portability	Adaptability	70.64	80.36	86.86
	Installability	78.24	83.59	89.53
	Replaceability	72.33	77.67	83.93
	Co-existence	78.39	84.00	90.22

In the formula (1), software product evaluation will be straightforward and objective once there are weight values and ratings for each evaluation items. The values in the Table 1 and tester rated values will lead to the straightforward evaluation of EMS applications with the assumption that testers are almost end users through the course of testing. Evaluators in the testing or certification organizations actually experience that testers with black-box testing becomes more than the real user of the software through the testing process by the help and training from the vendors, the user interview, etc.

### 3. Conclusion

A way of evaluating software based on the general testing methodology has been introduced and an Email Management Systems (EMS) has been evaluated in a real situation in order to prove the firmness of the evaluation approach. In the evaluation process, the weight values of the evaluated items (lists) have been derived by the user survey. The weight values are important here since they can be reused in the real-world EMS-like software evaluation

without intensive survey. The weight values will be beneficiary to the software evaluation organization and companies when evaluating similar software products by reusing and benchmarking them.

### **3.1. Future Direction**

We are trying to derive the weight values for various softwares while we evaluate them in our software products quality certification program. The evaluation of software products will become easier to deal with once the weight values are derived for various software categories.

### **3.2. Limitations**

Survey was conducted to the 23 users that are insufficient sample sizes for the proper survey analysis. We are now trying to outsource the surveys in order to survey more users and expand the survey coverage in terms of software categories and the number of users. Joint efforts between the survey outsourcing and the software testing expertise will lead to the more objective and precise weight values for software product evaluation.

## **References**

- [1] Hyun Kyung Kim, M.K. Lee, H.R. Kim, Developing a Measurement Scale for User Evaluations of Internet Portal Sites, The Journal of Korean Institute of CALS/EC, Vol. 6, No. 3, p. 127~148, Dec. 2001
- [2] ISO/IEC 12119 – Information Technology – Software Packages – Quality Requirements and Testing, 1994.11
- [3] ISO/IEC 9126 - Information Technology - Software quality characteristics and metrics, 1991
- [4] Jong Moo Lee et al., A Certification Method for the Quality of Package Software Considering the Measurement Scale, Institute of Korea Information Systems 2001 Fall Conference, 2001
- [5] Khaled El Emam, A methodology for Scenario-Based Software Certification, ETRI report, 2001
- [6] Sang-Wook Park et al., Software Evaluation Framework for Package Software Quality Certification, Korea Information Processing Society 2001 Fall Conference, 2001.10
- [7] SQEC (Software Quality Evaluation Center) at ITTL (IT Testing Laboratory) of TTA (Telecommunications Technology Associations), The Software Evaluation Module for Software Testing version 1.1, Internal document of SQEC, 2001.11
- [8] Wonil KWON et al., Multilateral Approach for Software Evaluation, Korea Information Processing Society 2002 Spring Conference, 2002. 04