The Study of Test Evaluation Method Based on Bank Assessment System

JianBin Song
SooChow University,
School of Computer Science and Technology,
Suzhou, Jiangsu, China.

Fei Huang
SooChow University,
School of Computer Science and Technology
Suzhou, Jiangsu, China.

Abstract: In this paper, the personal financial assessment system of Industrial and Commercial Bank of China branch in Suzhou was used for an example, during the designing and testing in parallel, according to test results, we propose a method to calculate the satisfaction of the software system, to judge the performance of the system and to analyze the extent of the various stages of testing changes, then evaluate the quality of the system changes.

Keywords: Bank assessment system, system evaluation and testing, customer satisfaction model

Ⅰ. Introduction

In enterprise systems, the bank operational system has brought a series of difficulties for software testing due to its complexity as well as the high security, accuracy and efficacy level. Especially, in the test appraisal aspect, it still paused in the quite primary stage, the common methods include the defect trend chart, defect severity distribution map and so on. These methods may show the trend and severity of the defects of software very clearly, which certainly helps with the test appraisal. However, due to the defects of insufficient target qualification and lack of vitality of evaluation index, the reproducibility of these methods is low. Besides, the requirement of the training of personals involving in the test in relatively high, which also affects the test appraisal outcome.

In this article, systems engineering methods and analytic hierarchy process were used to establish the target system and determine the weight of various targets, respectively, followed by measuring the degree of satisfaction of each individual target via target satisfaction degree appraisal target system. In the final step, the overall degree of satisfaction were calculated for the evaluation purposes.

Ⅱ. The establishment of appraisal target system

Referring to the demand analysis of the individual financial service inspection system, each module had different priorities, that is, “Red with first priority”, “Green with second priority”, “Yellow with last priority”, respectively. Different modules could result to distinct consequences when encountered problems. For instance, the target with end priority would not affect the system to a great extend, but the one with highest priority could possibly lead to mistakes that can never be corrected. Modules were set as the primary level target, which were given different weights according to their different priorities. Considering the overall degree of satisfaction of the entire system, the degree of satisfaction of each module was calculated first, followed by weighted summation.

Within a module, the ratio of number of problems, the additional proportion and the closure rate during the process of software development could be set as secondary targets, as they were very important data when evaluating the module itself. The ratio of the number of problems equaled to the number of problems discovered in the test over sample size, which reflected the accuracy of the software development procedure. The additional proportion could be calculated from the ratio between the additional number of problems to the total number of mistakes discovered. Closure rate was the percentage of the number of closures over total number of problems. If the problems discovered were found to be solved in another round of testing after modification, then the problems were considered to be closed. In other words, the number of closures directly reflected the speed and quality of solution of problems.

According to the analysis results above, the degree of satisfaction of the individual financial service inspection system was considered as general appraisal goal (O). The module was set as the level target (i.e. criterion level A), followed by further more detailed decomposition, forming an appraisal target objective tree, like Figure 2-1:
Let Q be the overall degree of satisfaction, then the equation for Q is:

$$Q = \sum_{i=1}^{3} Q(A_i) \cdot w_i$$  \hspace{1cm} (1)

where $Q(A_i)$, $i = (1, 2, 3)$ is degree of satisfaction of each level target, $w_i$ is the weight of each level target.

$$Q(A_i) = \sum_{j=1}^{3} Q(B_j) \cdot w_j$$  \hspace{1cm} (2)

where $Q(B_j)$ is one of the degrees of satisfaction of the three secondary level targets, $w_j$ is the weight of each secondary level target.

The equation for the degree of satisfaction of the secondary level target also needed defining. The degree of satisfaction was set as the variable, which the value territory of [0,1], and the target value was set as the independent variable.

The satisfactory range was defined to be the upper and lower limits of the target, based on the historical statistics data as well as the estimation from the development and testing personnel. If the target has degree of satisfaction above the upper limit, then it is considered to be completely satisfied, whereas, the target is completely unsatisfied when its degree of satisfaction is below the lower limit.

Taking into account the actual situation, the satisfactory range of various targets in the individual financial service inspection system is listed in Table 2-1:

### Table 2-1 degree of satisfaction appraisal target system

| Target | First indicator | Second indicator | The scope of the satisfaction | Remark |
|--------|----------------|-----------------|------------------------------|--------|
| 0: satisfaction | A1. The Last Priority | B11: The number of problems rate | [0, 0.33] | Decreasing satisfaction |
| | | B12: The additional question rate | [0.05, 0.2] | Decreasing satisfaction |
| | | B13: The closure rate | [0.35, 0.8] | Increasing satisfaction |
| | A2. The second highest priority | B21: The number of problems rate | [0, 0.33] | Decreasing satisfaction |
| | | B22: The additional question rate | [0.05, 0.2] | Decreasing satisfaction |
| | | B23: The closure rate | [0.35, 0.8] | Increasing satisfaction |
| | A3. The highest priority | B31: The number of problems rate | [0, 0.33] | Decreasing satisfaction |
| | | B32: The additional question rate | [0.05, 0.2] | Decreasing satisfaction |
| | | B33: The closure rate | [0.35, 0.8] | Increasing satisfaction |

II. Definition of the weight

In this article, the analytic hierarchy process was used to determine the weight of various targets. This process established the judgment matrix by comparing the importance of each target first, and the weight was then calculated from the extraction of root.

A. Establishment of judgment matrix

The importance of the target at the same level of objective tree was compared with each other by the testing and development personnel. The judgment matrix was the established with the reference of the standards stated below.

### Table 3-1 target importance grading standard

| Indicator and compare | Very unimportant | Unimportant | Slightly important | Equal | Somewhat unimportant | Very important | Important | Vital |
|-----------------------|------------------|------------|-------------------|-------|----------------------|---------------|----------|-------|
| 1/9                   | 3/5              | 5/7        | 7/9               | 1     | 1/3                  | 2             | 3        | 9     |
| 1/8                   | 1/5              | 1/2        | 1/3               | 1/2   | 1/4                  | 1/6           | 1/8      | 1/9   |

Table 3-2 below shows the judgment matrix obtained after comparing the relative importance of various targets listed in Table 2-1:

### Table 3-2 Primary level target judgment matrix A1

| | A1 | A2 | A3 |
|---|----|----|----|
| O | 1  | 1/3 | 1/9 |
| A1 | 1  | 1  | 5/7 |
| A2 | 3  | 1  | 1/5 |
| A3 | 9  | 5  | 1  |

### Table 3-3 Secondary level of target judgment matrix (A1)

| | B11 | B12 | B13 |
|---|-----|-----|-----|
| B11 | 1   | 1/3 | 1/3 |
| B12 | 3   | 1   | 1/2 |
| B13 | 3   | 2   | 1   |

### Table 3-4 Secondary level of target judgment matrix (A2)

| | B21 | B22 | B23 |
|---|-----|-----|-----|
| B21 | 1   | 2   | 1/2 |
| B22 | 1/2 | 1   | 1/3 |
| B23 | 2   | 3   | 1   |

### Table 3-5 Secondary level of target judgment matrix (A3)

| | B31 | B32 | B33 |
|---|-----|-----|-----|
| B31 | 1   | 2   | 1/2 |
| B32 | 1/2 | 1   | 1/3 |
| B33 | 2   | 3   | 1   |

B. Calculation of target weight

In this article, the weight was obtained by multiplication. Take the target in Table 3-2 as an example:

First, according to the row vector, get the judge matrix normalization:

$$b_{ij} = \frac{a_{ij}}{\sum_i a_{ij}}$$
Then multiply and get the cube root line by line:
\[ v_i = \left( \prod_{j=1}^{3} b_{ij} \right)^{1/3} \]

Then get \( v_i \) normalization:

Then, the biggest characteristic root
\[ \lambda = \frac{1}{n} \sum_{j=1}^{n} (Av)_j \]
is:
\[ \lambda = \frac{1}{3} \left( 0.2121 + 0.5366 + 2.2628 \right) \]
\[ = 3.0297 \]

The weight of each primary and secondary level target obtained is shown below:

| Table 3-6 Primary level target weight |
|--------------------------------------|
| O | A1 | A2 | A3 |
|---|----|----|----|
| Weight | 0.0704 | 0.1782 | 0.7514 |

| Table 3-7 Secondary level of target weight (A1) |
|-----------------------------------------------|
| O | A1 | A2 | A3 |
|---|----|----|----|
| Weight | 0.0704 | 0.1782 | 0.7514 |

| Table 3-8 Secondary level of target weight (A2) |
|-----------------------------------------------|
| A2 | B21 | B22 | B23 |
|---|----|----|----|
| Weight | 0.297 | 0.1634 | 0.5396 |

| Table 3-9 Secondary level of target weight (A3) |
|-----------------------------------------------|
| A3 | B31 | B32 | B33 |
|---|----|----|----|
| Weight | 0.297 | 0.1634 | 0.5396 |

C. Uniformity examination

Because the production of judgment matrix was very subjective, the results often appeared to be inconsistent. Therefore, the uniformity of the matrix needed to be examined. Take the Table 3-2 target as an example.

Calculation of consistency index:

Inconsistency:
\[ CI = \frac{\lambda - n}{n-1} = \frac{3.0297 - 3}{3 - 1} = 0.015 \]

| Table 3-10 Random consistency index RI value |
|---------------------------------------------|
| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| RI | 0 | 0 | 0 | 0.58 | 0.9 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 |

Consistency Ratio
\[ CR = \frac{CI}{RI} = \frac{0.015}{0.58} = 0.026 < 0.1 \]

Therefore, the uniformity of this judgment matrix was acceptable.

IV. Calculation of the degree of satisfaction

First, according to the test data collected, the degree of satisfaction of the secondary level target was calculated. For example, for the proportion of the number of problems, the satisfactory range of this target was \([0.1,0.5]\), and its degree of satisfaction decreased progressively, then the equation of the degree of satisfaction of this target was expressed as follow:

\[ Q(r) = \begin{cases} 1, & r < 0.1 \\ 0 - 1.01 \cdot (r - 0.1), & 0.1 \leq r < 0.5 \\ 0, & r \geq 0.5 \end{cases} \]

After testing, the proportion of the number of problems with highest first priority was \(6/26=0.2308\), then the degree of satisfaction was

\[ Q = 0.5 - 0.2308 = 0.673 \]

Table 4-1 listed the degree of satisfaction, the weight of secondary indicators as well as the weighted degree of satisfaction obtained by multiplying: \(Q(B_i)^*w_i\).

| Table 4-1 two level of target degree of satisfaction computation |
|---------------------------------------------------------------|
| The secondary indicator | The satisfaction range | value | satisfaction weight | Weighted degree of satisfaction | remark |
|-------------------------|------------------------|------|----------------------|-------------------------------|--------|
| b11: The number of problems rate (The Last Priority) | [0.1,0.33] | 4/4=1 | 0 | 0.1397 | 0 Increasing satisfaction |
| b12: The additional question rate (The Last Priority) | [0.05,0.2] | 0 | 1 | 0.3235 | 0.3235 Decreasing satisfaction |
| b13: The closure rate (The Last Priority) | [0.5,0.8] | 1 | 1 | 0.5278 | 0.5278 Increasing satisfaction |
| b21: The number of problems rate (The second highest priority) | [0.1,0.33] | 30/94=0.3191 | 0.0474 | 0.297 | 0.0141 Decreasing satisfaction |
| b22: The additional question rate (The second highest priority) | [0.05,0.2] | 18/30=0.6 | 0 | 0.1634 | 0 Decreasing satisfaction |
| b23: The closure rate (The second highest priority) | [0.5,0.8] | 1 | 1 | 0.5396 | 0 Increasing satisfaction |
| b31: The number of problems rate (The highest priority) | [0.1,0.33] | 6/26=0.2308 | 0.4313 | 0.297 | 0.1281 Decreasing satisfaction |
| b32: The additional question rate (The highest priority) | [0.05,0.2] | 0 | 1 | 0.5396 | 0 Increasing satisfaction |
| b33: The closure rate (The highest priority) | [0.5,0.8] | 1 | 1 | 0.5278 | 0.5278 Increasing satisfaction |

By using formula (2), the degree of satisfaction \(Q(A1)\) of the primary level appraisal target was obtained from the summation of the weighted degree of satisfaction of each of the secondary level target. Take the last priority level as an
example, where the weighted degree of satisfaction of its secondary indicators were 0.0, 0.3325, 0.5278, respectively. Then the degree of satisfaction of the primary indicator would be 0.8603 which is the summation of the three.

| First indicator | satisfaction | weight | Weighted satisfaction |
|-----------------|--------------|--------|-----------------------|
| A1 - The Last Priority | 0.8603 | 0.0704 | 0.0606 |
| A2 - The second highest priority | 0.0141 | 0.1782 | 0.0025 |
| A3 - The highest priority | 0.5914 | 0.7514 | 0.4444 |

Then by using formula (1), the overall satisfaction was obtained \( Q = 0.5095 \).

V. Evaluation and analysis

Through the calculation of degree of satisfaction, the appraisal target at each level in the evaluation index tree would have their own degree of satisfaction. The degree of satisfaction of each target level could be used to determine whether the result of this test was satisifiable or not. As shown in Table 4-2, the overall degree of satisfaction was 0.7368, i.e. not very satisfied, indicating the need of carrying out a level-by-level inspection from top to bottom to identify the indicator lowering the overall degree of satisfaction.

According to the degree of satisfaction of primary level target, indicator A1 was relatively satisfied with more than 86%, while target A2 and A3 were only partially satisfied with A2 in particular. Therefore, detailed analysis was needed to identify the causes of the insufficient degree of satisfaction of target A2 and A3, that is the highest and second highest priority level.

As shown in Table 4-1, all of the three secondary level indicators affecting target A2 were not very ideal, where were 0.0474, 0 and 0, respectively. Firstly, indicator B21 (i.e. the) was 32%, resulted from the 30 problems encountered in the 94 test cases in A2 unit in this project. The high ratio of number of problems means that there were a lot of misses and errors in the process, meanwhile, due to the incompleteness of the test, there might be a lot of other potential problems that were not found, which meant the number of tests might need increasing. Secondly, indicator B22 (i.e. the additional proportion) had the degree of satisfaction of 0, which requires to pay extra attention to. After re-testing, an extra 18 problems were found, which indicates the inadequate success on fixing problems found at the first place, or the large number of problems that were not found in the first round of testing, reflecting the severity of potential problems. Furthermore, indicator B23 (i.e. the closure rate) was 37%, which revealed the amendment process was relatively slow and the software could not be modified in time, affecting the following tests. On the other hand, the degree of satisfaction of the secondary level indicators affecting target A3 were 0.4313, 1 and 0.5557, respectively. The additional proportion was 0, which means there was not any new problem occurring at all, that is the degree of satisfaction was relatively high. Although there were not a lot of problems, that is only six, it still needed attention due to the high priority of this module. After retesting, there were four problems sorted, which met the acceptable level.

To conclude form the analysis the three secondary level targets above: the biggest problem at the moment was that the development of edition of the module with second highest priority. First of all, the quality of the development was not high enough, with the presence of a lot of problems. Secondly, the modifications were not comprehensive enough, leading to the problems occurring afterwards. Finally, the amendments were not done in time, which therefore lead to failure of finishing to solve a certain amount of problems before the next round of testing.

VI. Conclusion

Now most of the bank business of software development process is still testing after developing, they often get testing after all the codes finished. And many companies, limited to the actual situation, the importance of unit testing is not enough, they often take serious testing until after integration now the cost changes is already quite high . In this paper, the system development and test are in parallel, it establishes evaluation index system and calculates the index weights by AHP, then analysis the results by calculating the satisfaction. It timely detect test and related to the development process in order to adjust the development process, thus to provide timely advices for the development, which is greatly reducing the cost of system changes.

References:
[1] Yongjun Bai. Banking software test evaluation studies. Wuhan: China Water Transport (Academic edition). 2006 (06).
[2] Pressman (U.S.). Software Engineering: Practice's research methods. Beijing: Mechanical Industry Press. 2004.
[3] Qiyuan JIANG . Mathematical model (third edition). Beijing: Higher Education Press. 2003.