

Intelligent Website Evolution of Public Sector Based on Data Mining Tools

Jang Hee Lee¹ and Gye Hang Hong²

¹ School of Industrial Management, Korea University of Technology and Education,
307 Gajeon-ri, Byeong cheon-myun, Cheonan City,
Choongnam Province 330-708, South Korea
janghlee@kut.ac.kr

² Industrial System and Information Engineering, Korea University
kistduck@nate.com

Abstract. As one of means for the electronic government embodiment, the website construction and its complement of public sector such as government agency and public institution has been importantly considered. The public sector's website is operated for public benefit and consequently needs to be continuously redesigned for the users with lower performance in the satisfaction level and effect of using it based on the served information by evaluating whether the performance is different between the users with the various different backgrounds, areas and etc. In this study we present an intelligent evolution model of public sector's website based on data mining tools in order to improve the whole users' satisfaction and the effects of using it, especially the users with lower performances by continuously redesigning and complementing the current key web pages.

1 Introduction

As one of means for the electronic government embodiment, the website construction of public sectors such as government agencies and public institutions/enterprises have been actively propelled. Electronic government construction aims at providing citizen with service quickly and accurately, effectiveness of government work, innovation by redesigning work process, and raising national competitiveness by improving productivity [1].

Public sector's website ultimately differs with the private sector's website in its seeking the public benefit and equability. It provides beneficial public information to various kinds of users with different types of backgrounds, education levels, various needs and etc. (e.g., citizens, researchers, public servants in a government agency, other government agency and etc.). Since the users of public sector's website are varied and vast, it is difficult to improve the satisfaction of all users on it and the effects of using it based on the served information.

This study presents an intelligent website evolution model of public sector that can evolve its website and consequently satisfy the users of it and give them an improvement of profits after using it. It consists of 4 steps: current website evaluation based on user satisfaction survey, segmentation of the users into groups based on the survey, finding crucial website evaluation factors (i.e., discriminating factors) and key

web pages through pairwise comparison between two groups and finally redesigning the pages in terms of the discriminating factors to suit the targeted group. These steps are continuing and therefore the website evolves.

2 Public Sector's Website Evaluation

Website means the information system of Internet based on web to interactively exchange information and knowledge. The website of public sector provides useful public information to support decision-makings of various people as well as constructs the direct conversation channel between government and people and promotes quick collection of public opinion and citizen's participation for public decision-makings.

The research on the website evaluation of public sector has been recently attempted. The technical aspects of Internet such as the connection speed, error rate, number of URL, convenience of use and user's satisfaction in administration service side were proposed as the major criteria of public sector's website evaluation [2]. The research group of cyberspace policy in the university of Arizona proposed 2 evaluation criteria, transparency and accessibility, which focused on the openness of administration [3]. Alesander and Tate proposed 5 evaluation criteria, the information authority, information accuracy, information objectivity, information currency and information coverage of web page [4]. The web support group in New Zealand proposed the evaluation criteria such as scope of content, layout, accessibility, update frequency, extent of feedback, interactivity, site performance, etc [5].

3 Intelligent Website Evolution Model of Public Sector

We propose a new intelligent evolution model for public sector's website. As stated before, it consists of 4 steps: current website evaluation based on user satisfaction survey, segmentation of the users into groups based on the survey, finding discriminating factors and key web pages through pairwise comparison between two groups and finally redesigning the pages in terms of the discriminating factors to suit the targeted group.

3.1 Current Website Evaluation Based on User Satisfaction Survey

As the first step of intelligent evolution model of public sector's website, we evaluate the level of user's satisfaction about the current website service and the outcomes such as the level of user's reuse intentions, trust on the website and user's profit obtained by using the website information for user's economic behavior according to our new evaluation model of it.

In this study, for the rational design of public sector's website evolution model, we make a new its website evaluation model to use the major criteria for the website evaluation of federal government of the U.S and NCSI (National Customer Satisfaction Index) model for measuring customer satisfaction in Korea. The model describes the cause and effect relationship in the user satisfaction on public sector's website and outcomes such as reuse, trust, and profit/loss (refer to Fig. 1).

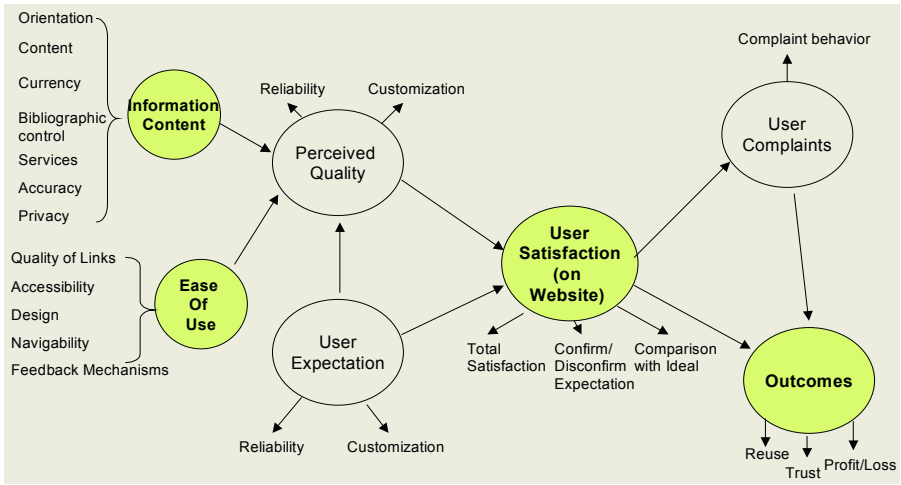


Fig. 1. Evaluation model of public sector’s website in our study

In our model, ‘Information Content’ and ‘Ease of Use’ are major criteria for the website evaluation of federal government of the U.S [6]. ‘Information content’ criteria evaluate the substantive aspects of the website and have 7 sub-criteria, orientation to website, content, currency, bibliographic control, services, accuracy, and privacy. ‘Ease of use’ criteria evaluate physical movement through the website and have 5 sub-criteria, quality of links, feedback mechanisms, accessibility, design and navigability. The ‘Perceived Quality’ after using the public sector’s website affects ‘User Satisfaction’ and accordingly ‘Outcomes’ such as reuse, trust on it and user’s profit/loss.

Based on our model, we conduct the user satisfaction survey on the public sector’s website. The survey asks all the registered users of website to rate 59 questions addressing criteria such as ‘Information Content’, ‘Ease of Use’, the level of user’s satisfaction, reuse intentions, trust on the website and the level of user’s profit to obtain after using the website information for user’s economic behavior, user’s demographic information and so on.

For most questions, a seven-point scale is used with the ratings of very dissatisfied/strongly disagree (1.0), dissatisfied/disagree (2.0), somewhat dissatisfied/weakly disagree (3.0), fair/moderate (4.0), somewhat satisfied/weakly agree (5.0), satisfied/agree (6.0), very satisfied/ strongly agree (7.0). A “don’t know (0)” option is also provided.

3.2 Segmentation of the Users into Groups Based on the Survey

As the second step of intelligent public sector’s website evolution model, we analyze the survey data and segment all the surveyed users into the group with similar patterns in their levels of satisfaction, reuse intentions, trust on the website and profit to obtain after using the website information for their economic behaviors.

To segment all the surveyed users, we use SOM (Self-Organizing Map), a special type of neural network using an unsupervised learning scheme, as a clustering tool

[7]. After SOM using, we can obtain several user groups with similar patterns in the satisfaction, reuse intentions, trust and profits and select an user group among them for the pairwise comparison in the third step, which has the highest level of satisfaction, reuse intentions, trust and profit.

3.3 Identification of Discriminating Factors and Key Web Pages Through Pairwise Comparison Between Two User Groups

As the third step of intelligent public sector's website evolution, we firstly identify the discriminating factors among all factors (e.g., the 'Orientation'-related factors of 'Information Content' criteria, the 'Accessibility'-related factors of 'Ease-of-Use', etc.) that are questioned in the survey, which can discriminate between the two paired user groups, the user group with the highest level of satisfaction, reuse intentions, trust and profit and the other user group among the segmented user groups. The other user group certainly has lower levels than the user group with the highest ones in the satisfaction, reuse intentions, trust and profit.

Pairwise comparison presents constituents such as discriminating factors in the survey and key web pages in web usage patterns in which the other user group with the lower levels is relatively lacking through comparison with the user group with the highest ones. If we obtain ten segmented user groups after SOM using, we conduct 9 pairwise comparisons.

In pairwise comparison, we conduct the analysis of survey data using C4.5 [8], a decision tree learning tool, in order to extract the discriminating factors in the user satisfaction survey. We choose the nodes appeared in the decision tree generated by C4.5 discriminating factors, which can classify the surveyed users into the user with the highest levels of satisfaction, reuse intentions, trust and profit and the other user with lower ones.

In addition to finding discriminating factors, we identify the difference of website usage patterns between the two paired groups by analyzing their web-log data and consequently find the key web pages to be redesigned in the public's sector website for the other users' improvement in the level of satisfaction, reuse intentions, trust and profit (refer to Fig. 2).

For the identification of key web pages, we firstly make the summarization table of user, user's total access frequency and connection time of web page during an analyzing time period and the user group to which the user belongs by analyzing user's web-log data. The table describes the access frequency and connection time of all users belonging to the two paired user groups for all possible web pages of current website.

We secondly identify key web pages that can distinguish the users (group) with the highest levels from the other users (groups) on the basis of website usage patterns by using C 4.5. Like the same way in the identification of discriminating factors, we choose the nodes appeared in the decision tree generated by C4.5 key web pages. The generated nodes are the connection time or access frequency of key web pages (e.g., the connection time of web page 38, the access frequency of web page 36) among all possible web pages of current website.

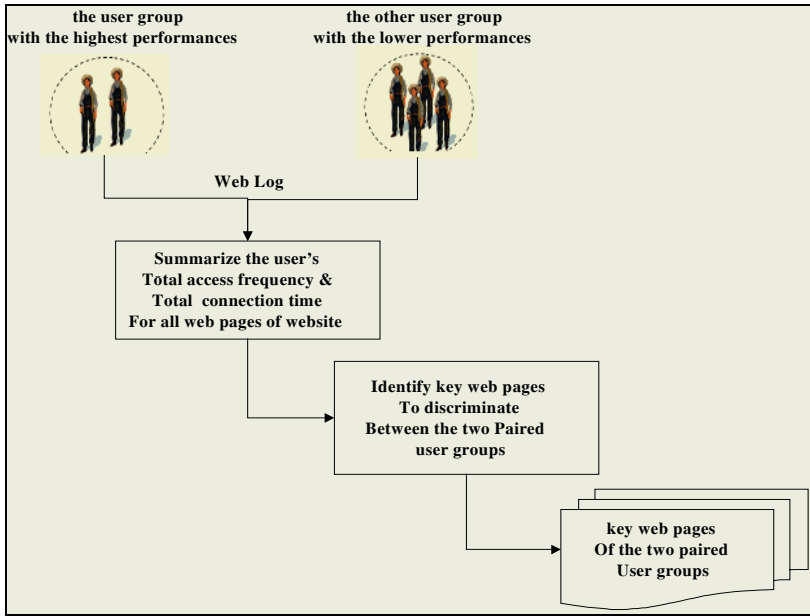


Fig. 2. Key web pages identification steps

3.4 Redesigning Current Key Web Pages in Terms of the Discriminating Factors

After finding the key web pages for all paired user groups, we differently redesign current those web pages (e.g., web page 38 and 36) in terms of the identified discriminating factors (e.g., content-related factors, design-related factors) for the improvement in the level of satisfaction, reuse intentions, trust and profit of the user group with lower levels.

Since the discriminating factors can discriminate between the user group with the highest levels and the other user group with lower ones in the satisfaction, reuse intentions, trust and profit, they influence the level of user's satisfaction, reuse intentions, trust and profit. If we provide the users with lower levels with the redesigned key web pages on the basis of discriminating factors, their level of satisfaction, reuse intentions, trust and profit can be improved.

For the intelligent evolution of public sector's website, we propose that the above described 4 steps are regularly and continuously executed. Through the website evolution, all the users with different backgrounds, education and experience levels, etc. are continuously satisfied with the website, trust and reuse it and can obtain more profits by easy and correct using the website information for their economic behaviors.

4 Conclusion

Incorporating intelligence into website is increasingly important, especially in public sector, as the cross-section of user communities is various, vast and increasingly

growing. Our study presents an intelligent evolution model of public sector's website based on data mining tools, which can continuously provide all the users with the lower levels of satisfaction, trust, reuse and benefits with the beneficial information service to improve their levels by continuously and suitably redesigning key web pages in the website.

Our evolution model can decide what to change next in the website for the performance improvement of users through the analysis of user satisfaction survey data using data mining tools. We will progress further work about the application of our model to other areas and the supplement of our evaluation model for public sector's website in Usability engineering, requirements engineering, prototyping, and information design.

References

1. Lee J. A. and Jung J. W.: Strategy for Implementing High Level e-Government based on Customer Relationship Management. Korean National Computerization Agency. (2004)
2. Korean National Computerization Agency.: Guide of construction, operation and information resource management of public institution's website. Korean National Computerization Agency. (1999)
3. Research group of cyberspace policy of the university of Arizona.: <http://w3.arizona.edu>
4. Alesander and Tate.: <http://www.widener.edu/wolfgram-memorial-library/inform.htm>
5. Web support group of New Zealand.: <http://www.theweb.co.nz/govtweb/0115.html>
6. Kristin R. E., John C. B., Charles R. M., Steven K. W.: Accessing U.S. Federal Government Websites. Government Information Quarterly. Vol. 14. (1997) 173-189
7. Kohonen, T.: Self-Organization & Associative Memory. 3rd ed. Springer-Verlag, Berlin (1989)
8. Quinlan, J. R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers. San Mateo California (1993)