# Control System for Intelligent Web-site : Conceptual Framework and System Design

Seungjoon Lee<sup>1)</sup>, Euiho Suh<sup>2)</sup>

<sup>1)</sup> Pohang University of Science and Technology (joonie@postech.edu)
<sup>2)</sup> Pohang University of Science and Technology (ehsuh@postech.edu)

#### Abstracts

Today's web provides highly dynamic, personalized and interactive services to users and has gained amazing popularity. However, as the size and complexity of information in a web site grows, it may become difficult for users to browse web-site efficiently. Hence, to gain a competitive advantage over the countless available Web sites, Web site should have some intelligent ability. In order to provide intelligence, we must understand the desires and behaviors of individual users in uncertain environment. This typically involves collecting information about the customers in some direct or indirect way and uses this information to develop intelligent Web sites. We define intelligence as the ability of a Web server system to act appropriately on user's uncertain behavior, where appropriate action is that which increase the probability of success, and success is the achievement of behavioral sub-goals that support the system's ultimate goal. The purpose of this paper is to systematically review all uncertain aspects of intelligence and propose a conceptual framework of control system for intelligent web site and finally design control system. To design control system, we use IDEF0 model and NN (neural network).

### 1. Introduction

According to the Gartner group's report, about 85% Asian-Internet Company will become bankrupt or merge in the near future. (Gartner 2000.5) This dramatic competition between Internet companies has brought about profound changes in the way they design a Web site. And, Web sites are becoming critical for companies and organization for their survival.

Therefore, there have been a lot of efforts on providing highly dynamic, personalized and interactive services to users and have gained some popularity. In particular, personalization of web contents has been widely applied in Web site. They allow users to customize the site for themselves. Most "portal" sites (e.g. Yahoo!) allow manual customization such as list of favorite links, stock quotes of interest, and local weather reports. Some sites also allow users to describe interests and will present information-news articles, for example-relevant to those interests [8].

However, as the size and complexity of information in a web site grows, it may become difficult for users to browse web-site efficiently. Hence, to gain a competitive advantage over the countless available Web sites, Web site should have some intelligent ability. In other words, if a Web site wants to provide all users satisfactory service, it should be able to monitor each user and what users find interesting. Only then can a web site provide appropriate contents to its users. This implies that Web site should have some intelligent abilities [4].

There is a lot of literature on user interface and personalization aspect of Web site design but little on intelligent capability of Web sites. Surfing on the WWW shows that a lot of sites are built in a very ad hoc manner and have little intelligent capabilities.

The purpose of this paper is to define the meaning of intelligence in Web-site and then specify design requirements for system and finally design control system for intelligent Web sites. We propose a method which is "user centered" rather than "web-manager centered." In a "web-manager centered," the focus is on maintenance or implementation issue. In our approach, the starting point is that of potential visitors of the Web site. This approach results in Web sites which are better tailored to their users and will therefore have a higher intelligence and give greater satisfaction

The structure of the remainder of the paper is as follows: In Section 2 of this paper, we describe various researches which related to our work. In Section 3, we examine a meaning of intelligence capabilities in Web site and then use it as a foundation for defining intelligent Web site and for specifying conceptual framework of an intelligent Web-site. In Section 4, IDEF0 is used for illustrating functional capabilities of control system and we propose design of control

system of intelligent Web-site. Finally we end this paper with conclusion.

#### 2. Recent Researches

There is a lot of literature on user interface and personalization aspect of Web site design but little on intelligent capability of Web sites. However there are some similarities between personalization issues and intelligence. Therefore, we introduce here some personalization research from design to implementation and then intelligence issues.

#### 2.1 Web site design

A number of researchers have already recognized the importance of design methods for Web sites and have proposed several methods such as HDM, OOHDM, HDM-2. However, design issues related with personalization are just being introduced in Web community. WSDM is a user-centered method for the design of kiosk Web Sites [1]. The WSDM focused on the construction of "Audience-driven" Web application. In this method, users are classified into user classes and the available data is modeled from the viewpoint of the different user classes. This result in Web sites which are better tailored to their users. Another interesting approach for dealing with personalization is presented in WebML [14]. Using the WebML modeling language it is possible to specify the structure of the application's information base, the structure of nodes and the navigation model. Web ML allows representing well known Web patterns and supports data derivation and user modeling. Personalization in WebML is expressed using event-condition-action rules. Rossi presented the OOHDM approach for specifying and designing personalized behaviors in Web applications [3]. They showed which design abstractions are necessary to build customizable applications and allowed modeling personalization features with a design constructs. But these researches have a limitation not to be based on data that actually reflect the customer behavior in the Web sites.

#### 2.2 personalized recommendation

Personalized recommendation methods guess where the user wants to go and take her there immediately (or at least providing a link). There have been three major approaches to provide personalized recommendation. In the content based approach, it recommends items that are similar to what the users has liked in the past. It filters content by keywords, string matching and so on. Most Web search engines use these techniques, but they harness only a fraction of the indexable Web and, even then, require users to sift through many results to determine relevant selections. One reason for the low coverage is that the majority of Web pages are dynamically generated and hence not directly accessible via hyperlinks. Another reason is the lack of sophisticated conceptual models for Web information retrieval [12]. In the collaborative filtering approach, it identifies other users that have showed similar preference to the given users and recommends what they have liked. It mines user-access patterns, Web logs, preferences, and profiles to tailor the content provided at specific sites. These days, Data Mining has become an important area of research. Yu (1999) explored how some of data mining techniques can be applied to personalization and introduced various data mining techniques(such as clustering, similarity indexing, Association rule etc.) and their applications to personalization.

Researches in this approach have some intelligent aspect. Dua et al. (2000) presented an automatic discovery method that discovers frequent access routines for unique clients from Web access log files. This method helps understanding user surfing behaviors and is useful for improving Web site domain tree.

Sarukkai (2000) tried to directly capture the sequence of link traversal in a probabilistic manner. They made probabilistic link sequence model using maximum likelihood estimation of Markov models and robust extension.

Ramakrishnan (2000) presented PIPE methodology which is a programmatic framework to design personalization systems. PIPE is based on three concepts: partial evaluation, data mining of semi-structured data, and information integration.

However, each of above researches is only interested in user pattern analysis which is part of intelligence.

#### 2.3 Intelligence

There are little literatures which deal with Intelligence issue of Web site. However, we can find some clue about intelligence of Web-site from some literatures which deals with personalization, user path prediction and recommendation.

Perkowitz et al. (2000) discusses the notion of adaptive web sites which semi-automatically improve their organization by learning from visitor access patterns. The PageGather algorithm uses page co-occurrence frequencies to find clusters of related but unlinked pages. Based on PageGather algorithm, index pages are created for easier navigation. Index page is a page consisting of links to a set of pages that cover a particular topic. They also argue that the concept of adaptive Web sites challenge will both drive AI (Artificial Intelligence) advances and yield valuable technology. However their works have some weakness. They focus on simplifying navigation for individual users and therefore visitor's access pattern is the only input to their system.

Lai et al. (2000) propose system architecture for intelligent browsing on the Web. They firstly provide a classification framework for Web browsing activities using two dimension-Web familiarity and purpose of browsing, and then explore the required intelligent capabilities for the a Web site and finally propose a system architecture. They applied intelligent agent concept to their system architecture. They can't explore fully intelligent aspects of Web site but their works are the first step toward the intelligent Web site. Especially, their works have meaning to our research by providing a useful classification framework for browsing activities, which can serve as a foundation to understand user behavior and laying foundation for intelligent capabilities for a Web site.

### 3. Conceptual Frameworks of Intelligent Web Site.

### 3.1 Definition – Intelligent Web site

Lai et al.(2000) suggested intelligent capabilities for improving browsing activities based on requirements.(See table 1) However, they are not enough to provide an intelligent environment for their users for the meaning of intelligence in Web sites are not fully discovered.

Table 1 Intelligent capability for improving browsing activities.			
User type	Required intelligent capabilities		
For users unfamiliar with the Web	Recommend hot links		
site	Highlight the most valuable contents		
	Provide default search keywords		
	Provide intelligent search		
For users familiar with the Web	Recommend hot links		
site	Highlight the most valuable contents		
	Provide intelligent search		
	Highlight new contents		
	Provide customized contents		
	Monitor personal status		

	1 111 0 1	• • •	
I able I Intelligent cana	ability for impro	ving hrowsing	activities
Table I Intelligent capa	tomey for impro	The browsing	activities.

Much is unknown about intelligence and much will remain beyond human comprehension for a very long time and there are many different perspectives. From the viewpoint of control theory, intelligence is the integration of knowledge and feedback into a sensory-interactive goal-directed control system that can make plans, and generate effective, purposeful action directed toward achieving them. From the viewpoint of psychology, intelligence is the integration of perception, reason, emotion, and behavior in a sensing, perceiving, knowing, caring, planning, acting system that can succeed in achieving its goal in the world [7].

From the purpose of this paper, intelligence will be defined as the ability of a Web server system to act appropriately on user's uncertain behavior, where appropriate action is that which increase the probability of success, and success is the achievement of behavioral sub-goals that support the system's ultimate goal. Also, intelligent Web site will be defined as Web system which has intelligent capability.

Both the criteria of success and the systems ultimate goal are defined external to the intelligent system. For an intelligent Web site the goals and success criteria are typically defined by Web administrator and CIO. Usually, goals of Web site are efficient user's browsing, agreement of company's goal with user goal.

### 3.2 Conceptual frameworks

According to the above discussion, we propose a conceptual framework for intelligent Web-site, as shown in Fig. 1 (This figure is modified version of Albus's work which deals with intelligence theory). There are four elements of intelligent Web site: Monitoring, decision making, world model, and execution. Input to, and output from, intelligent systems are via sensors and actuators (These terms somewhat confusing, but can be applied to this situation).

- 1) *Actuators*: Output from an intelligent Web site is produced by actuators that design dynamic Web contents. Actuators generate directions to appearance of Web pages. An intelligent Web may have tens, hundreds, thousands, even millions of content piece, all of which must be coordinated in order to make a web page.
- 2) *Sensors*: Input to an intelligent Web site is produced by examining log files, which may include client IP, server IP, navigation time, document name, HTTP protocol and file size of any browsing activity occurring at the Web site. Furthermore, CGI parameters which are passed by users have to be caught.



Fig. 1 Conceptual framework for intelligent Web site

- 3) *Monitoring*: perception takes place in a monitoring element that compares sensory observations with expectations generated by a world model. Monitoring algorithms integrate similarities and differences between observations and expectations over time so as to detect and recognize events. In other words, monitoring element catches the user's intention from sensory input data and has a following three functions :
  - A. to capture user status
  - B. to capture user's browsing characteristics
  - C. to update the existing world models
- 4) World model: the world model is the intelligent system's best estimate of knowledge about the world, plus a database management system that stores and retrieves information. The world model also contains a simulation capability that generates expectations and predictions. The world model thus can provide answers to requests for information about the present, past, and probable futures states of the user. The world model provides this information service to the execution system element, so that it can make intelligent plans and choices, to the monitoring system element, in order for it to perform correlation, model matching, and model based recognition of states and events, and to the decision making system element in order for it to compute values such as cost, benefit, uncertainty, importance, attractiveness, etc.[1] The world model is kept up-to-date by the monitoring system element. World model consist of three different DBs, which are for user information, site database and knowledge model database respectively. User information database stores information which is related with user's background and profiles. Knowledge model database stores model which can predict user's behavior and estimate the possible alternative design directions. System database stores thousands of design pieces which can compose whole pages.
- 5) *Decision making*: the decision making system element determines what is good for user's browsing efficiency and accomplishing system's goal. The decision making system observe state of the world (user and system) and the predicted results of hypothesized plans. It computes costs, risks, and benefits both of observed situations and of planned activities. It computes the probability of correctness and assigns believability and uncertainty parameters to state variables. It thus provides the basis for making decisions for choosing one action as opposed to another. Decision making element has a following three major functions :
  - A. to optimize user's browsing activity using the existing world models
  - B. to optimize Web site's goal using the existing world models
  - C. to generate error recovery strategies
- 6) Execution: page design strategy results from an execution system element that selects goals, and plans and executes tasks. Tasks are recursively decomposed into subtasks. And subtasks are sequenced so as to achieve goals. Goals are selected and plans generated by a looping interaction between execution, world modeling and decision making elements. The execution system hypothesized plans, the world model predicts the results of those plans, and the decision making element evaluates those results. The execution system then selects the plans with the highest evaluations for execution.

Each of the system elements of intelligence is reasonably well understood. The phenomenon of intelligence, however, requires more than a set of disconnected elements. Intelligence requires an interconnecting system that enables the various system elements to interact and communicate with each other in intimate and sophisticated ways.

# 4. Functional capabilities of control system

# 4.1 IDEFØ

IDEFØ is a method designed to model the decisions, actions, and activities of an organization or system. IDEFØ was derived from a well-established graphical language, the Structured Analysis and Design Technique (SADT). The United States Air Force commissioned the developers of SADT to develop a function modeling method for analyzing and communicating the functional perspective of a system. Effective IDEFØ models help to organize the analysis of a system and to promote good communication between the analyst and the user.

# 4.2 IDEFØ model for intelligent Web site

The ultimate goal of intelligent Web site is to provide what concerns or interest the users and then get a maximum profit from it. Therefore, we need to have a full understanding of all of the system function, control, input and output. However, in this paper, we limit our context to shopping mall as shown in Fig. 2. For a shopping mall, Web site's goal is to correctly predict user's intended products and show fast and intelligent link to them. Many data - log file, business goal, user profile, product catalog - can be used to predict user's intended product.



## Fig. 2 A-0 model: Predict intended product.

Fig. 2 shows function model of control system for intelligent Web site, which has three activities:

- A. predict item which user want using user's path and pattern and time
- B. Predict item which user want using user's probability and group
- C. Specify item considering business rule.

Fig. 3 shows model of control system, which bears responsibility for updating NN.



Fig. 3 A0 model: IDEF model for control system



Fig.4 A0 model for updating world model

### 4.3 Control system design

Control system for intelligent Web site consists of three activities. Control system outputs expected(intended) target design pieces and instruction which coordinates them and controlled by user profile, business rule and product catalog etc. This activity will be done by mechanisms-NN, Markov chain - as shown in Fig. 3. We show the design of control system in Fig. 5.



Fig. 5 Control system which consists of 2 NN and 1 Markov chains

#### 5. Conclusions

In this paper, we presented a framework for designing intelligent Web sites. The framework is based on intelligent theory and the principles that the Web site should be adapted to and give intelligence to users. This paper not only fully defines intelligent capability in Web site, but also proposes a conceptual framework and control system for intelligent Web site. However, this paper has some limitation. First, we don't build a prototype site. Second, functions of intelligent Web site are not fully identified. Third, feasibility is unclear.

Despite the limitation, this paper is meaningful to suggest the conceptual framework for intelligent web-site and specify functional requirement to control that for the first time.

An interesting problem for future research is that how to apply various AI technologies to intelligent Web site and fully specify the relationship between user behavior and business goal.

### References

- [1] De Troyer, C.J. Leune; WSDM: a user centered design method for Web sites, Vol. 30, pp85-94, 1998.
- [2] Gustavo Rossi, Daniel Schwabe, Fernando Lyardet; Improving Web information systems with navigational patterns, WWW8, 1999.
- [3] Gustavo Rossi, Daniel Schwabe, Robson Mattos Guimaraes; Designing personalized Web Applications, WWW10, 2001.
- [4] Hsiangchu Lai, Tzyy-Ching Yang; A system architecture for intelligent browsing on the Web, Vol. 28, pp219-239, 2000.
- [5] J.A. Harding, K.Popplewell, R.Y.K. fung, A.R. Omar; An intelligent information framework relating customer requirements and product characteristics, Computers in industry, Vol. 44, pp51-65, 2001
- [6] Jae-jun Lee, A methodology for redesign of the web site based on the customer navigation pattern, master's thesis, 2002.
- [7] James S. Albus ; Outline for a theory of intelligence, IEEE transactions on systems, man and cybernetics, Vol. 21, No. 3, pp473-509, 1991.
- [8] Mike Perkowitz, Oren Etzioni; Towards adaptive Web sites: Conceptual framework and case study, Artificial intelligence, Vol. 118, pp245-275, 2000.
- [9] Ming-Syan Chen, Jong Soo Park, Philip S. Yu; Data Mining for Path Traversal Patterns in a Web Environment, Proceedings of the 16<sup>th</sup> ICDCS, 1996.
- [10] Naren Ramakrishnan; PIPE:Web Personalization by Partial Evaluation, IEEE Internet computing, pp21-31, 2000.
- [11] Philip S. Yu; Data mining and personalization technologies.
- [12] Ramakrishnan Srikant, Yinghui Yang; Mining web logs to improve website organization, WWW10, 2001.
- [13] Ramesh R. Sarukkai; Link prediction and path analysis using markov chains, WWW9, 2000.
- [14] Stefano Ceri, Piero Frantemali, Aldo Bongio; Web Modeling Language(WebML): a modeling language for designing Web sites, Computer networks, Vol. 33, pp137-157, 2000.
- [15] Sumeet Dua, Eungchun Cho, S.S. Iyengar; Discovery of Web Frequent Patterns and User Characteristics from Web Access Logs: A Framework for Dyanmic Web Personalization, IEEE, 2000