Fuzzy Web Ad Selector Based on Web Usage Mining

Sung Min Bae and Sang Chan Park, *Korea Advanced Institute of Science and Technology*

Sung Ho Ha, *Kyungpook National University*

Internet and Web technologies are widely available, making it easier for companies to conduct business and transfer information to customers. Moreover, they speed up financial transactions efficiently, reducing the transaction costs of commercial activities that businesses would normally incur. So, Internet business has created a competitive edge in terms of costs. In such an environment, a successful company wanting to survive and gain a competitive advantage must provide an acceptable bundle of customized services that satisfy customers' needs.

The authors' Web ad selection system divides Web site customers with similar preferences into several segments through Web usage mining. It uses fuzzy rules that express customer segments' surfing patterns on the basis of expert advice, and recommends appropriate ads by fuzzy inference.

Despite the Internet's obvious benefits as a new communication medium—such as targetability, trackability, deliverability, flexibility, and interactivity—Internet advertising gives the same advertising messages to all customers and so has suffered from poor responses.

To raise a Web ad's effectiveness, we propose a Web ad selector that personalizes advertising messages for customers based on their preferences and interests.

Online advertising challenges

As the Internet's capacity expands, so does the range of techniques available to online marketers. The first form of Web ads was the Web site itself. However, as the Web became cluttered with commercial sites, simply building a Web site was not enough to reach Internet customers. The primary advertising models on the Internet today include banner ads, email, chat or communities, sponsorships, and push technology.[1] Banner ads especially are the leading form of online advertising, so the market for them is growing. To raise banner ads' effectiveness, Web sites must deliver the right message to the right customer at the right time. For example, if a person mainly reads an electronic newspaper's sports section, it is more effective to show him or her sports-related ads to maximize a banner ad's effectiveness.

In short, Web sites must give every desirable customer personalized advertising messages. (For more on customized advertising methods, see the "Personalized Web Ad Recommendation" sidebar.) A successful online marketing solution must be able to identify such customers, predict and understand their preferences and interests, identify an appropriate ad, and deliver it in a personalized format directly to customers during their online sessions. As marketing moves to a one-to-one environment, it is important to understand individual customers' needs, differentiate between customers through market segmentation, give customers personalized products and services, and forge closer and deeper relationships with them. From this viewpoint, one-to-one marketing is also called *customer relationship management*, which maximizes a customer's lifetime value to the organization.[2]

The most effective method of gathering Web site visitor and customer information is through a registration form.[3] When customers register for a service, they must often disclose personal data—for example, demographic and psychographic information, including their tastes and interests. Then, advertisers can customize ads based on such data and deliver them to those customers. However, many online customers do not give marketers this information owing to privacy fears. Additionally, customers might not give accurate information about themselves, and Web site administrators could have trouble verifying it. Besides, because online customers' interests change and evolve, asking them to directly express their interests might leave advertisers with obsolete chunks of information.

1094-7167/03/$17.00 © 2003 IEEE
Published by the IEEE Computer Society

Personalized Web Ad Recommendation

Generally, online service providers—including electronic newspapers—analyze their sites' effectiveness from a marketing perspective. *Web marketing* is broadly defined as activities used to draw customers to online stores and retain them. Online marketing techniques include database marketing, one-to-one marketing, and ad targeting (offer targeting).

Database marketing attempts to give customers more personal service. It divides customers into segments based on demographic characteristics such as ZIP code, income, and occupation, then markets to each segment as a group.

One-to-one marketing attempts to overcome marketing's impersonal nature by using technology to assist businesses in treating each customer individually. Recommendation systems help retailers implement one-to-one marketing strategies. Several Web-based personalized recommendation systems have existed. Personalization works by filtering a candidate set of items such as products or Web pages through some representation of a personal profile. Three main filtering approaches exist: *rule-based*, *content-based*, and *collaborative* filtering.[1,2] Decision rule-based filtering lets a Web site administrator manually specify rules on the basis of user demographics or static profiles collected by asking users a series of questions. It delivers content appropriate to a particular user on the basis of those rules. Content-based filtering recommends items on the basis of their similarity to what a given person has liked in the past. Typically, you would represent both items and profiles as vectors in the space of features and compute their similarity using a standard distance metric such as Euclidean distance. Collaborative filtering aims to recommend items that people similar to the target user have liked. It uses information about a group—which can be an entire population of users or a cluster—to produce individual recommendations.

Ad targeting attempts to identify which consumers should receive an ad based on their prior behavior. This approach fits into the current online advertisement spectrum, which includes untargeted, editorial, targeted (filtered), and personalized approaches.

Many characteristics of Internet advertising make it attractive to advertisers:

- *Improved targeting*. In many respects, the Internet offers the ability to better target advertising and so increases advertising expenditure efficiency. Service providers collect customer information on their Internet habits and interests, which can then be used for appropriate advertising.
- *Customization*. As a service provider accumulates customer information, it can customize the virtual store, product, price, and sales process for the customer.
- *Low cost production.* Internet advertising's production costs are less than the production costs of television, radio, or print media advertisements. This Internet advertisement characteristic is extremely attractive.
- *Improved monitoring.* A great advantage of Internet advertising is that a company can monitor its success rate and so apply advertising expenditure more efficiently. It is easy to monitor advertising effectiveness quantitatively. Internet advertisements will likely be paid for on a performance basis such as clicks through or time per view.
- *Global reach.* An audience unrestricted by geographic borders might be able to view Internet advertising. This increased visibility could be used with great effectiveness.

In a recent Forrester survey, banner ads ranked first in popularity among online advertising techniques. We can ascribe their popularity to relatively low-cost, high visibility on a Web page, better targeting, ease of tracking, and sheer simplicity. Banners come in different shapes and sizes, and you can place them virtually anywhere on a Web page, but they usually appear toward the top or bottom. Most advertising inventories are banner advertisements sold on a cost per thousand (CPM) impressions basis. CPM is measured by visitor impression, which is generally defined as a site's charge for every 1,000 times visitors show an ad. The site's server records impressions, namely in the Web log file.

In an e-commerce environment, analyzing such information embedded in click-stream data is critical to improve Web marketing's effectiveness for online service providers. Recently, both industry and academia have conducted much research on Web server log analysis. Some of these efforts have shown how to use data mining techniques to characterize and model Web site access patterns in electronic commerce scenarios.[3]

References

1. B. Mobasher, R. Cooley, and J. Srivastava, "Automatic Personalization Based on Web Usage Mining," *Comm. ACM*, vol. 43, no. 8, Aug. 2000, pp. 142–151.
2. G.S. Linoff and M.J.A. Berry, *Mining the Web: Transforming Customer Data into Customer Value*, John Wiley & Sons, 2001.
3. R. Kohavi and F. Provost, Applications of Data Mining To Electronic Commerce: A special issue of *Data Mining and Knowledge Discovery*, vol. 5, nos. 1/2, Jan.-Apr. 2001.

An alternative way to gather customer information is to capture visit information stored on server log files. Mining the Web server log files greatly lessens fears about customers' online privacy, tracks visitors as they navigate a Web site, and helps identify their current tastes and interests with help from machine learning technologies.

Our Web ad selector enables one-to-one Web marketing for e-newspaper providers. It gives customers personalized, real-time ads whenever they visit a Web site, which improves their satisfaction and their responses to the ads. The ad selector mines server log files containing customers' Web navigation patterns and decides on proper Web ads using a fuzzy controller, which performs fuzzy inference.

The fuzzy Web ad selector

We divide our proposed Web ad selector into three parts (see Figure 1): *customer clustering*, *fuzzy inference*, and *Web ad matching*.

Customer clustering uses the self-organizing map (SOM)—a special type of neural network using an unsupervised learning scheme—to mine Web server log files, split customers into different segments on the basis of their similar preferences and interests, and assign each customer to the resulting customer segment. Advertising experts determine the customer segments' viewing patterns and suggest corresponding ads. These patterns and ads form the fuzzy rules later used in fuzzy inference.

Fuzzy inference looks at the number of a

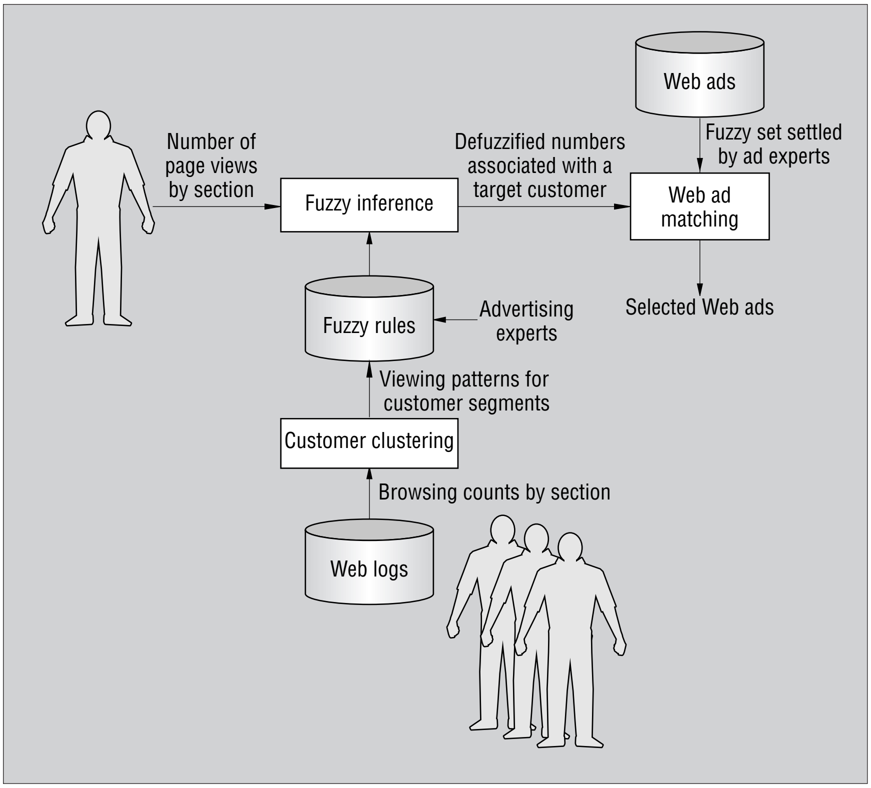

Figure 1. The fuzzy Web ad selector's framework.

target customer's page views (the number of requests for a specific page that a Web server successfully delivered) per news section as an input vector. It produces a fuzzy set and an actual value for the corresponding scalar associated with each solution variable (that is, ad category). The ad selector system then feeds these into the selector's Web ad matching part, which chooses appropriate Web ads by calculating a Hamming distance metric and forwards them to the target customer.

Segmenting Web site customers

"Which Web ads are the most suitable for a target customer?" is the most important question for service providers and advertisers who want to maximize their Web ads' effectiveness. To answer this question, service providers should determine Web site customers' habits and interests, which can be discovered by mining their Web page navigation patterns. Researchers have used Web mining in several distinct ways depending on the type of Web data they wished to mine. The most recognized Web mining categories are content mining, structure mining, and usage mining.[4]

Web usage mining uses data mining algorithms to automatically discover and extract patterns from Web usage data to predict user behavior while the user interacts with the Web. Although Web usage mining has some limitations, such as sparse usage data or regular changes in site content, it has several advantages over the other techniques mentioned. One advantage is its ability to dynamically develop user profiles from user patterns while reducing the need to explicitly obtain subjective user ratings or registration-based personal preferences prone to biases. So, the ad selection system's performance does not degrade over time. Web usage mining can help a Web site determine a customer's lifetime value, cross marketing strategies across products, and a promotional campaign's effectiveness. It can also provide information on how to restructure a Web site to create a more effective Web site presence and can encourage more effective management of customer communication and Web server infrastructure.

Ronald R. Yager suggested fuzzy-set-based intelligent agents,[5] which can autonomously and instantaneously determine which ad to display on a Web site on the basis of the viewer's characteristics. However, he used two characteristics of site visitors—age and income—to formulate their profiles. Generally, e-news is offered free, and an e-news service provider does not have its customers' demographic profiles. This makes it difficult to apply Yager's method to Web ad recommendation in an e-news service. To recommend the right Web ads to the right users, we must therefore try to find clues from a Web log through Web usage mining.

Before extracting customers' access histories, on which we can run the mining algorithms, we must address several data preprocessing issues, such as data cleaning and identifying unique users, sessions, and transactions.[4] We can accomplish data cleaning—cleaning a server log to eliminate irrelevant items—by checking a file name's suffix in the uniform resource locator. We can remove all log entries with filename suffixes that indicate graphic files and leave log entries indicating Web pages containing e-news. We can identify individual customers and their sessions using various preprocessing methods.[3,6]

Once we have identified customer access histories, access pattern mining follows, possibly including path analysis, discovery of association rules and sequential patterns, and clustering and classification. To segment customers, we perform a clustering analysis of Web page traversal, which customers form as they navigate Web pages that contain electronic news by section. Traversal histories for segmentation purposes might display three features: a customer number or customer ID; the time since each customer last visited a section (*recency* view); and the frequency with which each customer has visited the section (*frequency* view).

Figure 2 shows a process for segmenting customers. For analytical convenience, we appropriately classify an electronic newspaper's sections into seven categories—POL (politics), ECO (economy), SOC (social), WOR (world), CUL (culture and life), SCI (science and technology), and SPO (sports). We use these sections later as antecedents of fuzzy rules, which ad experts will generate through knowledge engineering. The fuzzy rules represent customers' viewing electronic news and experts' suggesting ads. Then we count how often each customer has visited each section of the e-newspaper during the analysis period and feed this information to the SOM.

Generally, a Web mining system using SOM has two phases: learning and use. In the learning phase, SOM analyzes Web access data and builds a customer behavior model. This phase is often time consuming and might require assistance from human analysts. After the model is built, the mining system enters a use phase in which you can rapidly and easily apply the model to customer situations.

Using SOM, we derived seven dominant customer segments (see Figure 2). In training SOM, output segments are restricted to 10 or fewer due to managerial convenience. We can represent average access counts per segment with four triangular fuzzy numbers such as S (small), MS (medium small), MB (medium big), and B (big), or three triangular fuzzy numbers such as S (small), M (medium), and B (big). Customers or Web ads generally have multiple characteristics which can't be specified in a crisp, numeric value. In this sense, fuzzy numbers are suitable for expressing customers' or Web ads' fuzzy characteristics because they help us derive the best guess from the gray information having multiple characteristics.

For example, Figure 3 shows a fuzzy partition of the POL section.

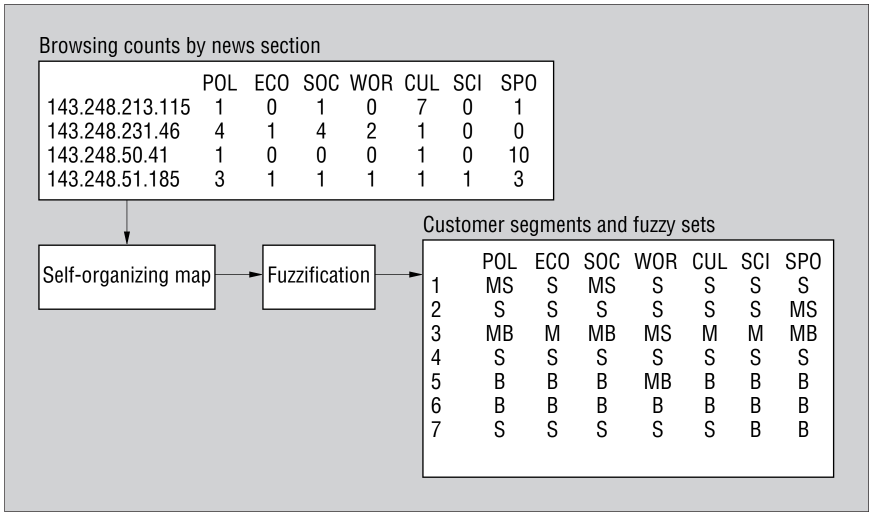

Figure 2. Process for customer segmentation and fuzzification.

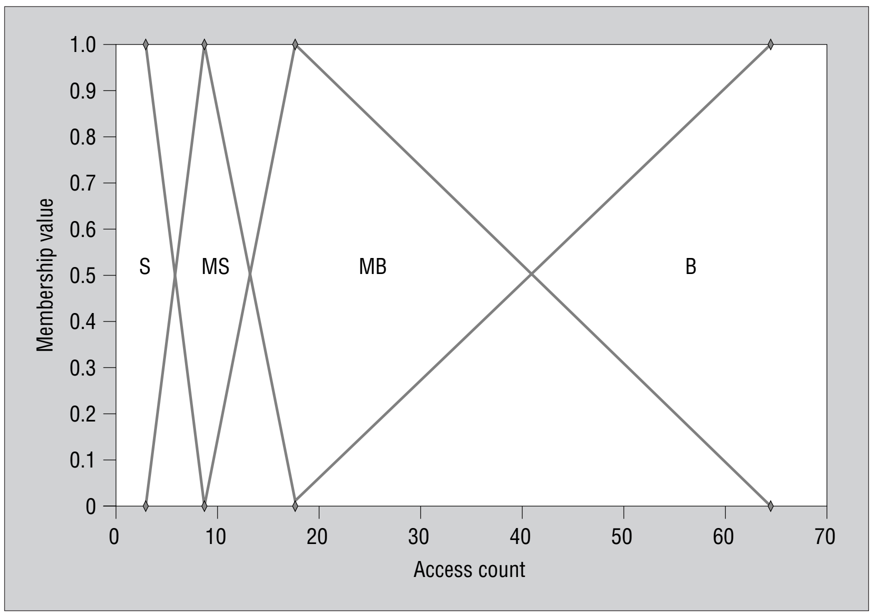

Figure 3. A fuzzy partition of the political section.

Categorizing Web ads

To build an experimental database of Web ads for an ad selector system, ad experts randomly select and analyze ads from a major e-newspaper provider ranked in the Korean daily newspaper market's top three according to "World Press Trend 2002." The ad experts then classify the ads into six main categories according to their characteristics and assign them appropriate fuzzy sets:

- MAG: magazine, newspaper, and news-related ads
- ECO: business, banking, and stock-related ads
- COM: computer, information, and telecommunication-related ads
- SPO: sports-related ads
- ENT: hobby and entertainment-related ads
- RAN: random ads

Eventually, ad experts collected nearly 60 ads for the experimental database. They selected eight to 10 ads—each with either a strong or weak degree of relevance—for each section. You can collect more ads, but the Web ad database will only increase in size. No changes will occur in the derivation of fuzzy rules, the mechanism of inference, and defuzzification, which have nothing to do with the number of ads.

The six ad categories are used as consequents in the right-hand side (RHS), or action part, of a fuzzy rule executed only to the degree that the left-hand side (LHS), or condition part, is true. Each category has membership ranging continuously between 0.0 and 1.0 and has three triangular fuzzy numbers such as S (small), M (medium), and B (big). The "RAN" category is needed when the Web ad selector cannot decide which ads are appropriate to a target customer. If the RAN category has a higher value than others, then a Web ad selector will show the randomly selected ads.

Designing fuzzy rules and the fuzzy inference engine

Our fuzzy inference engine receives a target customer's access counts per e-newspaper section as an input vector value. It processes the information through the fuzzy rule base and produces a fuzzy set for each ad category, which is then converted to a real number through defuzzification.

Fuzzy rule design. Figure 4 shows examples from the 32 fuzzy rules we elicited from ad experts. The fuzzy rule has an if-then form in which the LHS is a conjunction of triangular fuzzy numbers for access counts by section, and the RHS is a conjunction of triangular fuzzy numbers for ad categories. First, ad experts explain which ads are suitable for each customer segment. For example, the first rule

	POL	ECO	SOC	WOR	CUL	SCI	SPO		MAG	ECO	COM	SPO	ENT	RAN
R1	S	S	S	S	S	S	MS	**THEN**	S	S	S	B	S	S
R2	B	B	B	B	B	B	B	**THEN**	B	B	B	B	B	S
R3	B	B	B	MS	B	B	B	**THEN**	M	M	B	B	B	S
R4	B	B	B	B	B	B	B	**THEN**	B	B	B	B	B	S
R5		B						**THEN**	M	M	S	S	S	S
R6				MB				**THEN**	S	S	S	S	S	M
				...								...		
				...								...		

Figure 4. Sample fuzzy rules representing the expertise of ad experts.

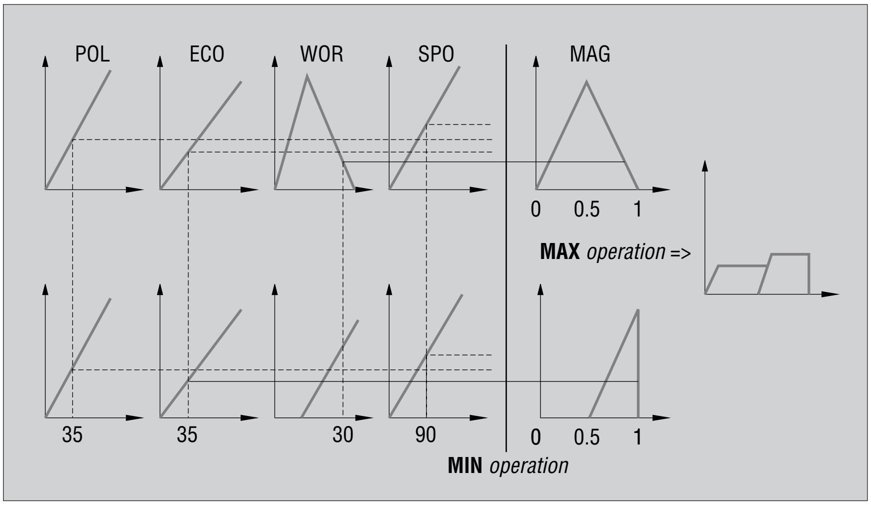

Figure 5. Fuzzy inference for the consequent "MAG."

(R1) in Figure 4 represents their expertise that sport-related ads are suitable for customers who belong in segment 2 and are interested in the e-newspaper's sports section. Additionally, ad experts add special rules—which have only one triangular fuzzy number in the LHS—to the fuzzy rule base to make the rules' coverage expand to the universal set. For example, the rule R6 in Figure 4 indicates that the randomly selected ads are suitable to customers who are interested in the WOR section.

Fuzzy inference engine. To infer appropriate Web ads according to the access count by section, we use Mamdani's minimum operation as fuzzy inference. Among fuzzy inference methods, Larsen's product operation is as useful as Mamdani's minimum operation from the computational viewpoint.[7] However, because Mamdani's implementation is often used in various applications, we select his method to infer Web ads and to recommend them online.

When an access count vector is given as (POL, ECO, SOC, WOR, CUL, SCI, SPO) = (x_1, x_2, x_3, x_4, x_5, x_6, x_7), we use the following equation to describe Mamdani's inference:[8]

$$\begin{aligned} W_k &= \mu_{A_1}(x_1) \wedge \mu_{A_2}(x_2) \wedge \mu_{A_3}(x_3) \wedge \mu_{A_4}(x_4) \\ &\quad \wedge \mu_{A_5}(x_5) \wedge \mu_{A_6}(x_6) \wedge \mu_{A_7}(x_7) \\ \mu_C(z) &= \bigvee_{j=1}^{rule\,\#} \left[W_j \wedge \mu_{C_j}(z) \right], \end{aligned} \tag{1}$$

where x_i is the ith input element in the access count vector, $\mu_{A_i}(x_i)$ is the membership grade of the ith antecedent (that is, ith news section), and $\mu_c(z)$ is the fuzzy set of consequent z (that is, the Web ad's z category). We summarize the fuzzy inference process using the following steps:

1. Compute each fuzzy rule's degree of fitness. Because seven antecedents exist, (POL, ECO, SOC, WOR, CUL, SCI, and SPO) in each rule, we obtain seven membership degrees and determine each fuzzy rule's fitness value by a MIN (∧) operation on these membership degrees.
2. Apply the fitness values determined in Step 1 to the consequents of the rules.
3. Combine multiple outputs through a MAX (∨) operation if multiple rules are active for the same input.

For example, when input vector X has (POL, ECO, SOC, WOR, CUL, SCI, SPO) = (35, 35, 35, 30, 35, 35, 90), how can we obtain the fuzzy set of consequent "MAG"? Figure 5 shows the detailed process. Mamdani's inference shows that antecedents "WOR" fired by rule R3 and "ECO" fired by rule R4 have the minimum fitness values. Computation of fuzzy sets for the remaining ad categories follows the same procedures as for "MAG."

Defuzzification. Researchers have studied defuzzification methods for several years and have applied them to fuzzy control and fuzzy expert systems.[9] The main idea behind these methods is to obtain a typical value from a given fuzzy set according to some specified characters, such as max criterion, mean of maximum, and central gravity (that is, centroid) methods. In other words, each defuzzification method provides a correspondence from the set of all fuzzy sets into the set of real numbers.

Defuzzification can transform the fuzzy sets for Web ad categories derived from the fuzzy inference engine into real numbers, which makes fuzzy inference and the resulting fuzzy sets usable. For this article, we use the central gravity method as a defuzzification method. We calculate the central gravity ($Z_{defuzzified}$) of a fuzzy number using

$$Z_{defuzzified} = \frac{\int \mu_C(z) z dz}{\int \mu_C(z) dz} . \tag{2}$$

Matching Web ads. After fuzzy reasoning, we obtain fuzzy sets and defuzzified real numbers associated with a target customer's Web page access patterns. Additionally, ad experts determine the fuzzy set of each ad, especially banner-type ads, in a Web ad database. To choose appropriate Web ads to forward to the target customer, the Web ad selector calculates the Hamming distance between each ad's fuzzy set and the fuzzy set of the target customer's Web page access patterns. The closer the Hamming distance is, the more similar two fuzzy sets are. So, those Web ads are suitable for the target customer. The ad selector calculates the Hamming distance using

$$d(A,B)=\sum_{\substack{i=1\\ x_i\in X}}^{n}\left|\mu_A(x_i)-\mu_B(x_i)\right|, \qquad (3)$$

where $d(A,B)$ is the Hamming distance; μ_A is defuzzified numbers fired by the fuzzy inference engine; μ_B is characteristics of Web ads; and X is categories of Web ads.

Figure 6 shows how to calculate the Hamming distance between a sample ad's fuzzy set and that of a target customer. The sample ad propagates a Web site that allows users to download free software to easily connect to the Internet and falls into the following categories:

- Computer, with a 1.0 degree of membership
- Entertainment, with a 0.5 degree of membership
- Randomness, with a 1.0 degree of membership

The resulting Hamming distance reaches 3.15.

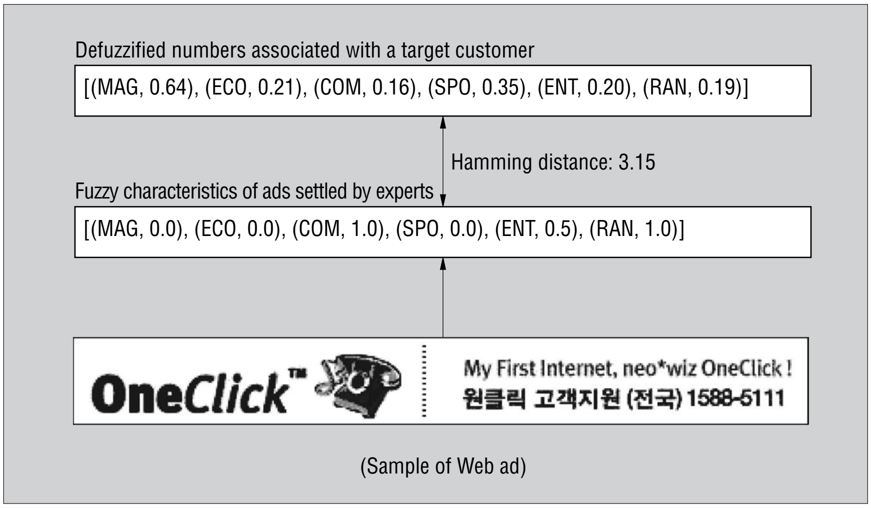

Figure 6. The Hamming distance between the two fuzzy sets. (Web ad courtesy of Neowiz)

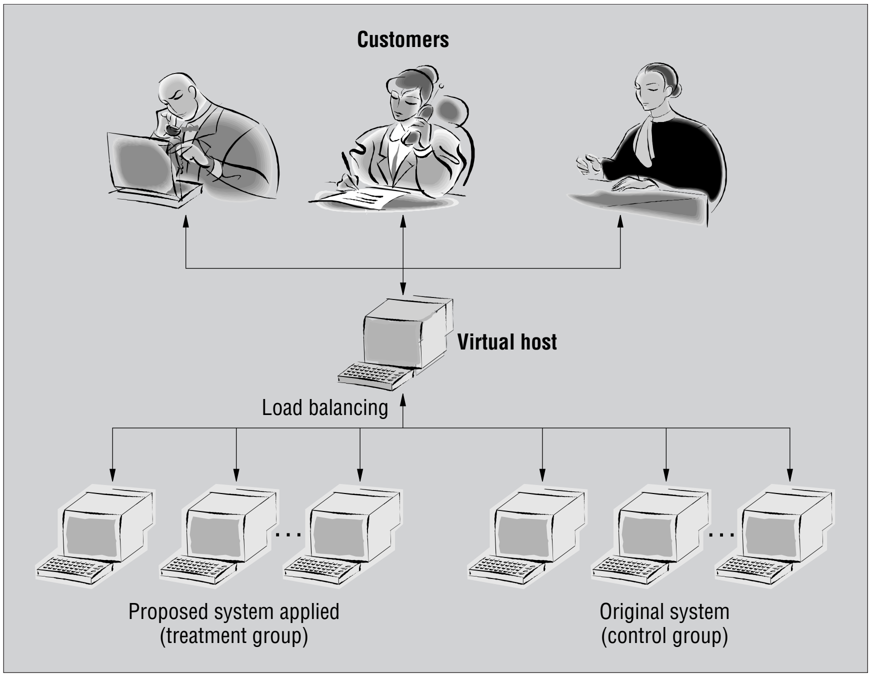

Figure 7. General architecture for experimenting on the fuzzy Web ad selector's performance.

Experimental design and evaluation

To analyze Web ads' effectiveness, we can use several metrics, such as *impression*, *click rate*, *reach*, and *click-through rate* (CTR). Impression measures how many times an advertisement appears on an accessed Web page. For example, if the current accessed page shows three ads, the page has three impressions. Advertisers use impressions to measure the number of views their ads receive. The click rate is the percentage of ad views that resulted in click-through. In Internet marketing, reach is how many different people visit a Web site to see the ad and also what percentage of these people fall into the ad's target audience. A common measure of reach for a Web site is its *unique visitors per month*. In Web site banner advertising, the CTR is the percentage of times that a Web page's viewers clicked on a given banner ad, transmitting the advertiser's Web site to the viewer. For example, if two out of every 100 visitors to a Web page clicked on a given ad, that ad would have a 2 percent click-through rate. In most cases, advertisers would consider a 2 percent click-through rate very successful. In many campaigns, especially where ads become familiar to users, the click-through rate is well below 1 percent.

An advertisement's characteristics determine which metrics advertisers use to measure its performance. For instance, the impression metric is appropriate for a corporate identity promotion ad, whereas CTR is meaningful for the membership invitation ads. In the majority of cases, CTR is the most general criterion of a Web ad's contract; so, we use CTR as a metric for Web ads' performance.

Our experiment has two major goals. The first goal is to verify the fuzzy Web ad selector's effect on the increase of CTR values per section of the e-newspaper. The second is to examine the accuracy of the fuzzy rules the ad experts extracted. So, we must compare the CTR before the ad selector's adoption with the CTR after the adoption to verify its effect and examine the CTR measures by section according to the advertisements' characteristics.

An increase in CTR values can result from various factors. However, disregarding those factors with the characteristics of a one-time event, we only consider the main factors that can discern the cause and effect. We can categorize these into the period factor, age factor, proposed system factor, and so on.

The architecture for experimentation with the fuzzy Web ad selector's performance comprises multiple Web servers, which operate together and deal with the e-newspaper

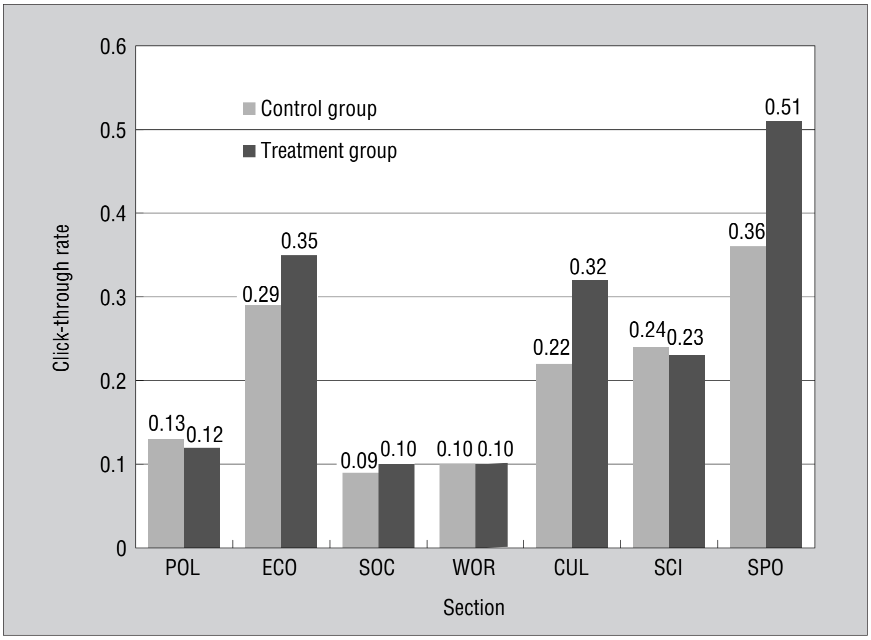

Figure 8. The click-through rate scores per section for the treatment group and control group.

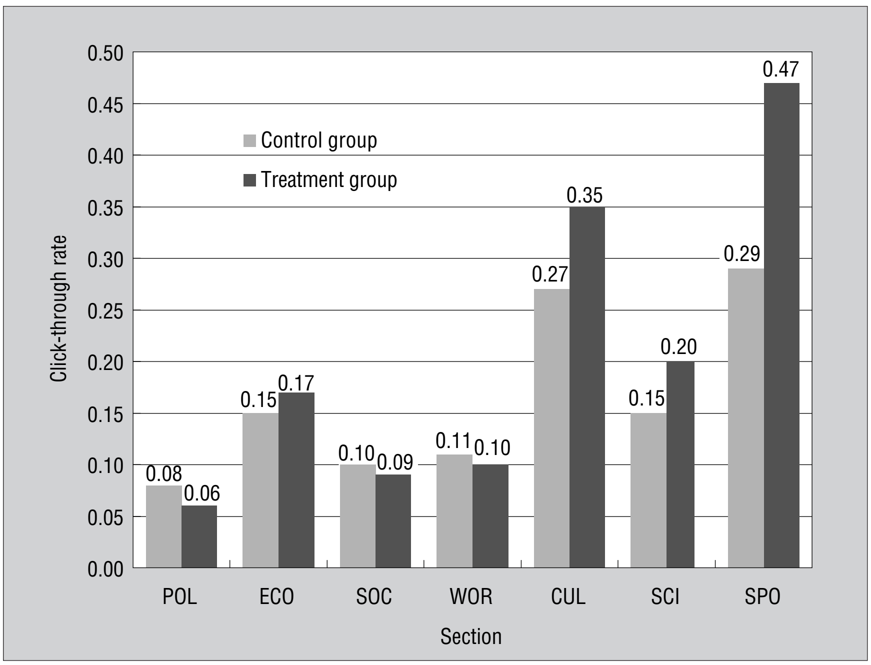

Figure 9. Distribution of the CTR by section for the sports-related ads.

provider's actual customers. Figure 7 shows this architecture.

Our experiment might also use direct user feedback. However, because the system obtains a user's actual click stream as the result of his or her Web interactions, it does not collect or retain any explicit or subjective user ratings. This lets the system track user ratings without annoying the user.

We apply the Web ad selector to the e-news servers using a simple load-balancing scheme. When a user requests e-news services, a virtual server representing all the actual servers receives the request and redirects it to any free-traffic Web server (that is, users do not know which server they connect to). If every Web server is busy, the virtual host redirects the request to a randomly selected Web server. Web servers are divided into two groups: one with the Web ad selector (the treatment group) and the other without it (the control group). So, the treatment group's and control group's Web servers have the same e-news articles, but the treatment group's shows Web ads that the selector recommends.

Under such circumstances, the virtual server might not accomplish random redirection. However, because the amount of access to those two groups is likely to be nearly the same in the long run, we need not be concerned about random sampling. Users will access the e-newspaper without realizing that the experiment exists, which guarantees that the same customers get the same services during the study (two months).

The difference in both groups' CTR values after the two-month experiment indicated the fuzzy Web ad selector's effects because the period and age factors were eliminated through grouping Web servers. Only the system factor remained. To acquire more precise results in the experiment, we made some additional treatments. The control group showed Web ads only once and the same customer could not see the same ads.

Figure 8 illustrates the average CTR scores per section for the two different groups after the two-month experiment.

A clear difference between before and after the adoption of the fuzzy selector does not exist for the Political, Social, World, and Science and Technology sections. However, the Economy, Culture and Life, and Sports sections have different results. Banner ads with solid classification characteristics result in a significant increase in CTR when the system recommends them to the right customers. Recommending ads to customers who mostly look at such sections as Political, Social, or World proves less effective. Those sections do not have specific classification characteristics for selecting the Web ads, so the random rule is more likely to be fired.

Figure 9 depicts the distribution of the average CTR by section especially for the sports-related banner ads. In general, sports-related banner ads are shown in the Sports, and Culture and Life sections. Even though a user's Web surfing characteristics might cause sports-related banner ads to appear in other sections, sections such as Political, Social, or World generally have random ads and have few sports-related banners due to the sections' randomness.

Figure 9 explicitly shows that for the Political, Social, or World sections, little to no difference exists between before and after adoption for the sports-related ads. However, the Sports, Culture and Life, and Science and Technology sections benefit from them. These results imply that the ad experts successfully designed fuzzy rules and classified the Web ads.

Users' profiles remain static and unchanged until the next segmentation. However, customers' interests and needs continuously change. The Web ad selector system must capture this change. Therefore, the average CTR goes down a set threshold, and the selector system starts deriving new profiles through customer segmentation and creating new fuzzy rules for the customers.

The work we present can be extended in several ways for future research. First, we can extend clustering analysis used to segment customers' access patterns to include richer information. The current clustering analysis uses recency, frequency, and monetary (RFM) values as base variables for customer segmentation. Although they performed well, we should consider an alternative set of features with better predictive performance than the RFM values. If we could significantly improve predictive accuracy using the alternative, we would have a basis for future research in this kind of analysis. If the fuzzy Web ad selector acquires customers' demographics and uses them for classifying Web ads, it could improve the CTR scores in the Political, Social, and World sections.

Second, deriving and maintaining fuzzy rule sets are also issues for fuzzy systems. Fuzzy reasoning can produce accurate results as long as we can define a sufficient number of fuzzy sets and rules. The larger the number of sets, the more accurately a fuzzy system reflects the true relationship among variables. On the other hand, it also increases complexity, decreases system robustness, and makes it more difficult to modify the system.

Finally, we must perform more empirical studies covering a longer time range over different types of online service providers. Such research will help us rigorously validate various speculations about Web marketing strategies. Additionally, it will help us understand an online service provider's effectiveness with minimal effort.

The Authors

Sung Min Bae is a postdoctoral fellow in the Department of Industrial Engineering at the Korea Advanced Institute of Science and Technology (KAIST). His research interests include customer relationship management, electronic commerce, data mining, service quality, and total quality management. He received his PhD in industrial engineering from KAIST. Contact him at the Dept. Industrial Eng., Korea Advanced Inst. of Science and Technology, 373-1 Kusung-Dong, Yusong-Gu, Taejon, Korea, 305-701; loveiris@major.kaist.ac.kr.

Sung Ho Ha is a professor of business administration at Kyungpook National University. His research interests include machine learning, data mining, e-commerce, agent systems, and intelligent information systems. He received his PhD in industrial engineering from the Korea Advanced Institute of Science and Technology. He is a member of the ACM and the IEEE Computer Society. Contact him at the School of Business Administration, College of Economics & Commerce, KyungPook National Univ., Sangyeok-dong, Buk-gu, Daegu, Korea, 702-701; hsh@bh.knu.ac.kr.

Sang Chan Park is a professor of industrial engineering at the Korea Advanced Institute of Science and Technology. His teaching and research specialties include artificial intelligence, expert systems, machine learning, total quality management, e-commerce, and supply chain management. He received his PhD in management information systems from the University of Illinois. He is a member of the IEEE. Contact him at the Dept. Industrial Eng., Korea Advanced Inst. of Science and Technology, 373-1 Kusung-Dong, Yusong-Gu, Taejon, Korea, 305-701; sangpark@mail.kaist.ac.kr.

References

1. C. Allen, D. Kania, and B. Yaeckel, *Internet World Guide to One-To-One Web Marketing*, John Wiley & Sons, 1998.
2. J. Peppard, "Customer Relationship Management (CRM) in Financial Services," *European Management J.*, vol. 18, no. 3, June 2000, pp. 312–327.
3. J. Srivastava et al., "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data," *SIGKDD Explorations*, vol. 1, no. 2, Jan. 2000, pp. 12–23.
4. R. Kosala and H. Blockeel, "Web Mining Research: A Survey," *SIGKDD Explorations*, vol. 2, no. 1, July 2000, pp. 1–15.
5. R.R. Yager, "Targeted E-Commerce Marketing Using Fuzzy Intelligent Agents," *IEEE Intelligent Systems*, vol. 15, no. 6, Nov./Dec. 2000, pp. 42–45.
6. R. Cooley, B. Mobasher, and J. Srivastava, "Data Preparation for Mining World Wide Web Browsing Patterns," *Int'l J. Knowledge and Information Systems*, vol.1, no.1, Feb. 1999, pp. 5–32.
7. C.C. Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controller—Part I, Part II," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 20, no. 2, Mar./Apr. 1990, pp. 404–435.
8. E. Cox, *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems*, AP Professional, 1994.
9. M. Ma, A. Kandel, and M. Friedman, "A New Approach for Defuzzification," *Fuzzy Sets and Systems*, vol. 111, no. 3, May 2000, pp. 351–356.