

2000

Web Ads Selection for One-to-One Advertising Using Neuro-Fuzzy Systems

Dan Jong Kim

SUNY Buffalo, djk5@acsu.buffalo.edu

M. Ramesh

Lucent Technologies, mramesh@lucent.com

H. R. Rao

SUNY Buffalo, mgmtrao@acsu.buffalo.edu

Follow this and additional works at: <http://aisel.aisnet.org/amcis2000>

Recommended Citation

Kim, Dan Jong; Ramesh, M.; and Rao, H. R., "Web Ads Selection for One-to-One Advertising Using Neuro-Fuzzy Systems" (2000).
AMCIS 2000 Proceedings. 389.

<http://aisel.aisnet.org/amcis2000/389>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2000 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Web Ads Selection for One-to-One Advertising Using Neuro-Fuzzy Systems

Dan Jong Kim (djk5@acsu.buffalo.edu)

Department of Management Science and Systems, State University of New York at Buffalo

M. Ramesh (Mramesh@lucent.com) Lucent Technologies

H. Raghav Rao (mgmtrao@acsu.buffalo.edu)

Department of Management Science and Systems, State University of New York at Buffalo

Abstract

The purpose of this research is to develop a web ad selection model for one-to-one advertisement using neuro-fuzzy systems. The objectives of the paper are i) to present a one-to-one web advertising model to develop a "learning relationship" with a customer, ii) to suggest the fuzzy inference approach for web ad selection instead of the classical mathematical programming approach to determine a precise match of a specific ad type for a specific customer, and iii) to describe how to select web advertisements based on customers' profile data using a sample scenario. The paper develops a method to identify web ads on a web site based on the customer's general behavior and demographic information. This targeting is based on preferences and quantifiable demographic data. A key aspect of the approach in this paper is that it uses the neuro-fuzzy (combining neural network and fuzzy systems) perspective borrowed from the engineering literature in soft computing.

1. Introduction

One of the major challenges facing any business-to-customer (B2C) electronic commerce (EC) initiative is to understand its customers. This understanding needs to occur at two levels: products and customers. With information about product requirements, a manufacturer or retailer can ensure that it builds or stocks what its customers want. With information on the demographics of its customers, the business can target sales advertisements directly at the current or prospective customers who are more likely to buy. In this paper we focus primarily on the customer perspectives.

Web advertising has emerged as a dominant business model for generating revenue. However, it is important to develop models for choosing web ads appropriate for targeted customers. In order to select ads more efficiently, it is necessary to have a clear definition of the targeted customers. For a travel package, for example, the targeted customer is defined as "a group of middle-aged customers having over middle income and having interests in travel". This information could be converted into precisely defined selection criteria by an expert's knowledge of this area. These could be: 1) Age: between 35 to 55, 2) Income: greater than \$35,000, and 3) Interest: total spending for

travel items such as airlines, tickets, cruises, lodging and vacation getaways is at least \$1,000. Unfortunately, such criteria are not plausible in some instances: an individual, 34 years old having an income of \$63,000 with \$2,500, total spending for travel items would belong to the target customer group but does not meet the selection criteria. However, another example, a 41 year old man having an income of \$36,000 with \$1,000 total spending will be selected by these selection criteria even though he may not fit the target group as well as the first individual did. To resolve such a conflict, we can apply fuzzy logic in order to implement the target customer selection process in the same way humans would make such decisions. That is they would never use strict and fixed thresholds but would use their intuition and past experience.

Since precisely defined selection criteria conflict in current models, we can apply fuzzy criteria that make possible common-sense relationships between conditions and the actual goal. Significant research has been done in applying fuzzy logic and neural networks for specific problems (Wong, Bodnovich and Selvi, 1997). However, research on web ad selection for one-to-one advertising using neuro-fuzzy systems is almost non-existent. This is one of the contributions of our research. To our knowledge only two articles have discussed this topic. Yager(1997) and Sullivan(1999), both use a general framework and a probabilistic scheme for the competitive selection of advertisements on the Internet without learning capabilities.

The purpose of this research is to develop a web ad selection model for one-to-one advertisements using neuro-fuzzy systems. The objectives of the paper are i) to present a one-to-one web advertising model to develop a "learning relationship" with a customer, ii) to suggest the fuzzy inference approach for web ad selection instead of the classical mathematical programming approach to determine a precise match of a specific ad type for a specific customer, and iii) to describe how to select web advertisements based on the customer's profile data using a sample scenario. The focus of this research is to identify web ads on a web site based on the customer's general behavior and demographic information. This targeting is based on preferences and quantifiable demographic data. A key aspect of the approach in this paper is that it uses the neuro-fuzzy (combining neural network and fuzzy

systems) perspective borrowed from the engineering literature in soft computing (Zadeh, 1994).

This paper is organized as follows. Section 2 describes the one-to-one advertising model. Section 3 presents the web ads selection process in terms of a sample scenario and details the simulated example, the structure of the system and the result of the simulation. Section 4 shows extracting rules using neuro-fuzzy systems and section 5 concludes the paper.

2. One-to-one advertising model

One of the most powerful capabilities of the web is enabling one-to-one customization¹. The one-to-one future will be characterized by customized production, individually addressable media, and one-to-one marketing. This could totally change the rules of business competition and growth (Peppers and Rogers, 1994).

In order to build enduring one-to-one relationships, a company must continuously learn from interactions with individual customers. The way to attract the attention of an individual is to develop a "learning relationship" with the customers. Find out what s/he is interested in, and over time, continue to ask preference questions to build a profile on each individual customer and continue to update an individual profile through her or his web behavior history like browsing, searching, clicking web banner, purchasing products and so on. The "learning relationship" can involve collecting information about web site customers through customer tracking technologies. Web customer tracking technologies such as registration/subscription/authenticated access, cookies², log pages, domain (IP) address of linking pages, and collaborative filtering make it possible to collect customer preference information on an individual basis (Dewan, Jing and Seidmann 1999; Raghu, Kennen, Rao and Whinston, 2000).

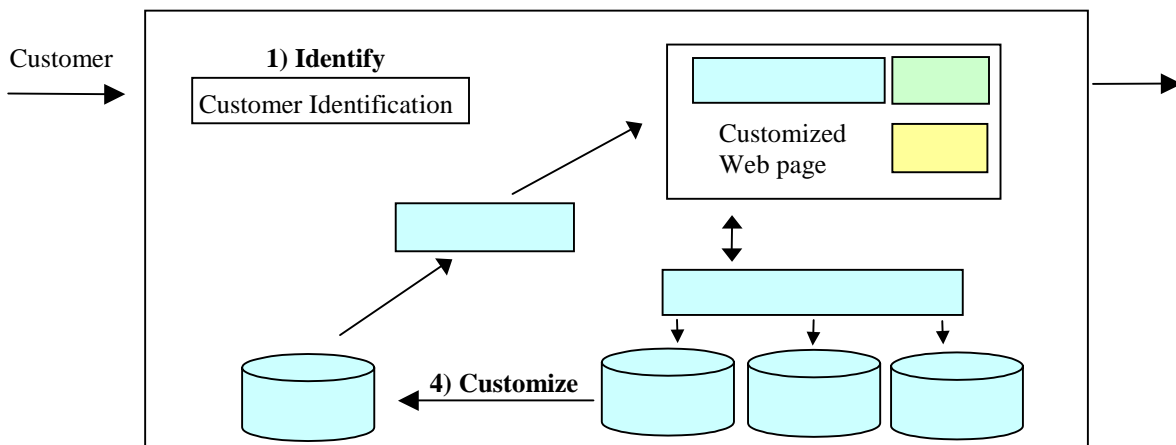
This section will describe a basic process for the selection of web ads for the targeted customer. Peppers,

Rogers and Dorf (1999) have proposed four implementation steps that can be used as a guide for one-to-one marketing. They are 1) identify customers, 2) differentiate customers 3) interact with customers and 4) customize some aspect of an enterprise's behavior toward customer, based on that customer's needs and value. We can use their four step guide as a basic framework to implement one-to-one web advertising (see Figure 1).

Identify customers: the purpose of identifying customers is to recognize a visitor and recall her or his previous visiting behavior or purchases. Cookies have limitations in some situations. For example, a customer can erase his or her cookie file, a customer may use different computer at home and at work, and a computer is used by several different users. Cookies offer an efficient method for identifying customers. **Differentiate customers:** the key idea behind one-to-one web advertising is to differentiate ad offerings based on the customer's previously indicated preferences. **Interact with customers:** at least the retailer's database contains information about three objects: customers, products and transactions. Customer personal profiles are updated by customer actions (searching, browsing, purchasing products, and so on) on this web site. **Customize toward customer:** it is the process for aggregating data from three databases into a customer preference database. In this process, the priority of ads will be stored in the customer preference database. They are generated by an ad selection process which we will suggest.

Once a web customer comes to a website, the customer identification process identifies who s/he is and selects this customer's preference data from the customer preference database (identify). The web advertisement selector then picks up a number of ads for this customer and places them on his/her customized web pages (differentiate). If s/he takes an action such as searching and/or purchasing for some items during web surfing, the

Figure 1: One-to one web advertising process framework



¹ The ability to create customized content for every single web visitor.

² A small piece of information sent by a web server to store on a web browser so it can later be read back from that browser.

profiling software (customer profiler) automatically updates the business databases (interact). New values for customer preference are then calculated and stored in the customer preference database (customize). Figure 1 shows the one-to-one web advertising process.

3. Sample scenario for ad selection

Let us consider the one-to-one advertising for targeted customers within the scope of the following "scenario." We assume that the ABC web retailer site has several different types of ads. All ads are classified by the following categories; Apparel (women, teens and men's clothes, shoes), Book & Videos (bestsellers, movies), Computers (desktops, software, monitors, games), Electronics (TV's, camcorders, cameras), Health & Beauty (vitamins, skin care, cosmetics), Music (CD's, MP3, cassettes), Sport/Sporting goods (boating, baseball, golf), Travel (airlines, tickets, cruises, lodging), Toy & Hobbies (video games, toys, action figures), and so on. Individual types of ads differ from each other in terms of target group characteristics. This site has meticulously collected its customers' demographic and preference data through a registration, survey, purchase and search process. This site has information from three databases: a customer, a product, and a transaction database. The customer profiler (profiling software) in figure 1 has updated these databases over time. The following data is available for the customizing customer preference database.

- Customer: customer name, customer ID, age, sex, marital status, address, income, homeowner, etc.
- Product: product name, product ID, price, cost, product category, etc.
- Transaction: customer ID, product ID, date and time, amount, quantity, etc.

We then process this transaction database with a simple data mining program such that when a record is read with a product that matches one of the categories, we add the amount of the purchase to the corresponding field in the customer preference database. An example customer preference database table may look like this:

- Customer preference: Customer ID, the purchase amounts for the categories, the ranks of preference ads, etc.

The rank of preference ads is the field which could be used by the ad selector to place banner ads on the customer's web pages. As a specific case for simulation, we assume that it is required to select two appropriate ads out of three for each customer who visits this web site.

The target customer groups for the three ad banners are:

Ad_1: a group of middle-aged customers having above middle income and having an interest in travel.

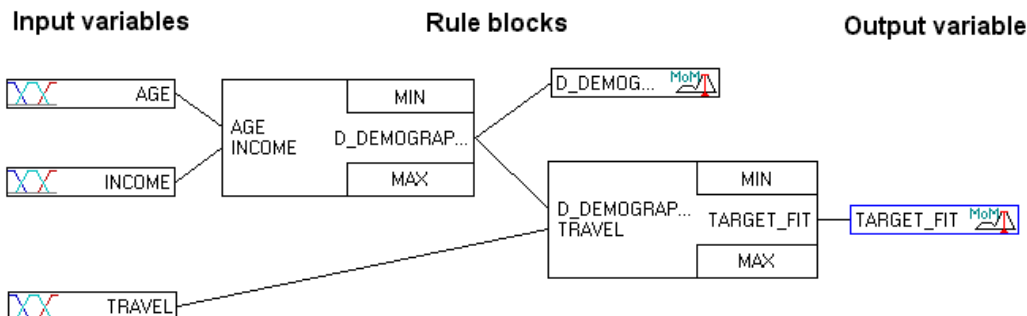
Ad_2: a group of youthful aged customers having low to middle income and having interest in music.

Ad_3: a group of elderly aged customers having high income and having interest in health.

The "classical" approach is to solve this as a mathematical programming problem with a corresponding objective function and a set of constraints. The objective function in this case is to assign advertisements to an individual customer with subject to the given ads target group's constraints. However, there are numerous factors that cannot be built into the formulation of the web ad selection problem as an Integer Programming or an Assignment Problem. It is nearly impossible to determine precisely a specific ad type for a specific customer. Therefore, we can suggest the fuzzy inference approach for each of the ads in this problem. In this simulated example, we show the web ads selection process for "Ad_1" using our sample scenario.

The targeted user group for an "Ad_1" is: a group of middle-aged customers having above middle income and having an interest in travel. For the sake of simplicity, in order to show how to develop a web ad selection model, we use only 3 factors; age, income and total spending for travel items. Age and income are demographic factors and total spending for travel is a customer's preference factor. A customer's preference factor can be collected by a transaction database in the customizing process. Figure 2 shows the structure of this fuzzy inference system including 3 input variables (AGE, INCOME, TRAVEL), 2 rule blocks and 1 output variable (TARGET_FIT). The connecting lines symbolize the data flow. The structure of the system is hierarchical similar to neural networks. At the first rule block (D_DEMOGRAPHIC), two elements (AGE and INCOME) are aggregated into a singular new node which will be used as an input variable for the second rule block along with the customer preference

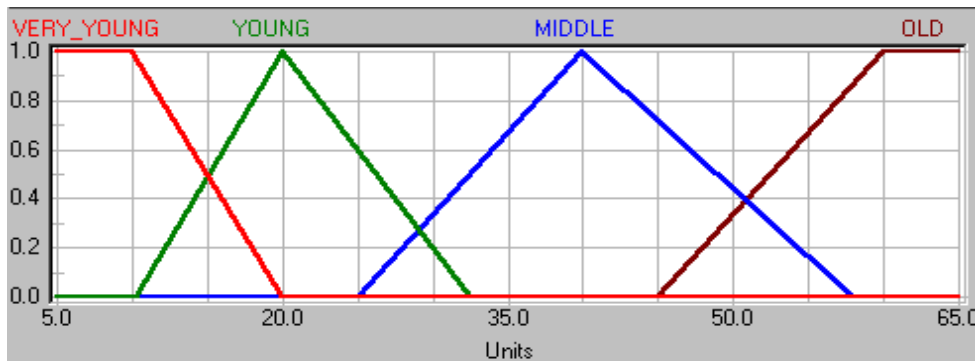
Figure 2: System structure using *FuzzyTech Professional*



variable (TRAVEL). This create two layers of abstraction. The first layer contains the element degree of demographic factors (D_DEMOGRAPHIC) which comprise the information of two input variables. The second layer contains the degree of target fit factors. Because information is condensed at each node, it is called an abstraction. Similar to a human, who takes many input variables into account to come up with one abstract judgment, the aggregation hierarchy proceeds until the output node (TARGET_FIT) is reached at the second layer. The hierarchical decision model squeezes the desired information on target fitness out of the three input variables.

A linguistic variable translates a numerical value into a linguistic value. The possible values of a linguistic variable are not numbers but so called “linguistic terms”. For instance, to translate the real variable “age” into a linguistic variable, four terms (“VERY_YOUNG”, “YOUNG”, “MIDDLE” and “OLD”) are defined. Each term is defined by a membership function (MBF). Each MBF defines for any value for the input variable the associated degree of membership of the linguistic term. The MBF of all terms of one linguistic variable is displayed in one graph. Figure 3 plots the MBF of the four terms for “AGE.” According to the MBF of “AGE” in Figure 3, an age of 18 is a member of MBFs for the terms: VERY_YOUNG to the degree of 0.2 YOUNG to the degree of 0.8 MIDDLE to the degree of 0.0 OLD to the degree of 0.0

Figure 3: MBF of AGE



In this system, the numerical input variables AGE, INCOME, and TRAVEL need to be translated into linguistic values. This step is called “fuzzification” since it uses fuzzy sets for this translation. Table 1 shows all linguistic variables of this system and their term names.

Figure 4 shows the conceptual structure of fuzzy logic system for web ads selection. Once all numeric input values have been converted to linguistic values, the fuzzy inference step can identify the rules that apply to the current situation and can compute the values of the output linguistic variables. The result of this is again a linguistic value for the linguistic variable TARGET_FIT. For

Table 1: Linguistic variables

Variable Name	Term Names
AGE	VERY_YOUNG, YOUNG, MIDDLE, OLD
INCOME	LOW, MEDIUM, HIGH
TRAVEL	LOW, MEDIUM, HIGH
D_DEMOGRAPHIC	LOW, MEDIUM, HIGH, VERY_HIGH
TARGET_FIT	VERY_LOW, LOW, MEDIUM, HIGH, VERY_HIGH

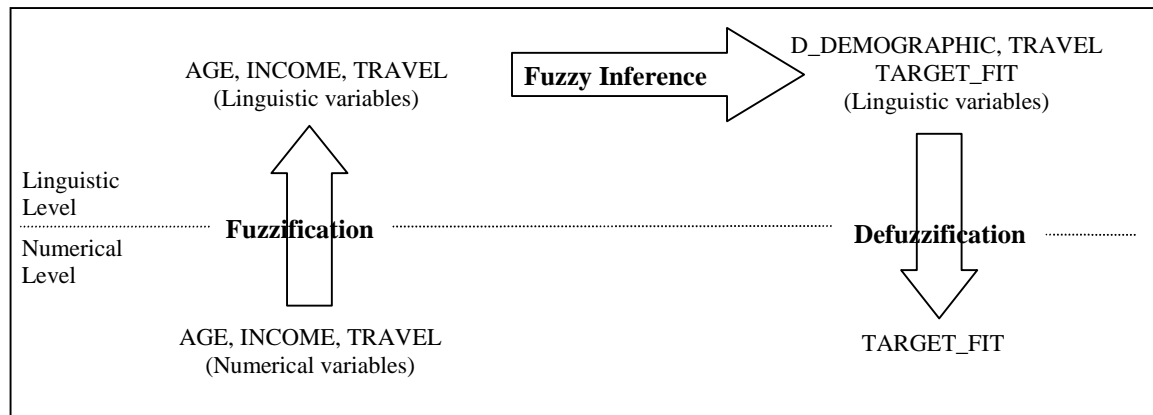
example, this linguistic result could be “VERY_HIGH.” The defuzzification step translates this linguistic result into a numerical value that results in the TARGET_FIT. Linguistic variables have to be defined for all variables used in the if-then rules. The rules are empirical knowledge concerning the operation of a particular process under consideration. The rule block is the control strategy of a fuzzy logic system. Each rule block is confined by the same input and output variables of the rules. The fuzzy inference step can identify the rules that apply to the current situation and can compute the values of the output linguistic variables. Below shows a subset of four rules.

Rule 1: **IF** AGE = VERY_YOUNG **AND** INCOME = LOW **THEN** D_DEMOGRAPHIC = LOW

Rule 2: **IF** AGE = YOUNG **AND** INCOME = MEDIUM **THEN** D_DEMOGRAPHIC = MEDIUM
 Rule 3: **IF** AGE = MIDDLE **AND** INCOME = MEDIUM **THEN** D_DEMOGRAPHIC = HIGH
 Rule 4: **IF** AGE = OLD **AND** INCOME = HIGH **THEN** D_DEMOGRAPHIC = VERY_HIGH

At the end of the fuzzy logic inference, the result for target fitness (TARGET_FIT) is output as a linguistic variable value. Target fitness is the degree to which the customer fits into the web ad. It means that if the TARGET_FIT is VERY_HIGH, the customer will have a very high tendency to click this ad or very high attention for this ad. To use this value for comparisons of other web

Figure 4: Conceptual structure of a fuzzy logic system for web ad selection



ads, it has to be translated into a numerical value. This step is called defuzzification. If the value is close to 1, it is likely that the web banner in question will be selected when this customer is visiting a site.

A part of the final result of the simulation is shown in Table 2. This table shows the degree of demographic factor (D_DEMOGRAPHIC) induced by age and income and the TARGET_FIT is derived from D_DEMOGRAPHIC and TRVEL. Likewise, TARGET_FIT of Ad_2 and Ad_3 could be determined by their own models. The TARGET_FITs of those are listed in the table 2. For our sample scenario, if a customer with id113197 visits this particular site, Ad_2 and Ad_3 will be selected because they have higher TARGET_FIT than Ad_1.

Table 2: Simulation Result

CUSTOMER	AGE	INCOME (\$10,000)	TRAVEL (\$)	Degree of DEMOGRAPHIC	TARGET_FIT Ad_1	TARGET_FIT Ad_2	TARGET_FIT Ad_3
id113197	31	5	50	0.5378	0.4147	0.5208*	0.5552*
id113205	50	7	250	0.9000	0.9367	0.4379	0.8003
id113310	31	7	50	0.5908	0.4561	0.4240	0.5790
id113430	41	8	250	0.8309	0.8992	0.5439	0.9367
id113438	50	7	250	0.9000	0.9367	0.4379	0.8003
id113599	36	8	250	0.7183	0.8334	0.6462	0.9367
id113973	50	9	250	0.9000	0.9367	0.4329	0.8003
id114140	26	7	100	0.4766	0.5400	0.6105	0.6822
id114145	20	2	50	0.1906	0.1823	0.5797	0.2846
id114174	21	2	30	0.1999	0.1466	0.5000	0.2463
id114211	23	2	50	0.2077	0.2016	0.5600	0.3275
id114484	41	6	200	0.8305	0.8766	0.5351	0.9367

*These were two ads selected for customer id113197.

4. Extracting the rules using neuro-fuzzy system

In a real situation, the greatest difficulty in developing and implementing the rule based fuzzy logic model is to formalize the expert's knowledge. To formulate the expert's knowledge more effectively, an attempt could be

made to develop a model based on neural networks. To formulate the expert's knowledge with learning capabilities, we can integrate neural networks technologies. Neuro-fuzzy systems, the combination of fuzzy logic and neural net technology, reaps the advantages from both technologies.

Neural networks can be mapped to a fuzzy system (Jang and Sun, 1996). This enables the use of powerful neural net learning algorithms with fuzzy logic. Fuzzy systems and neural networks are both soft computing approaches to modeling expert behavior. The neural networks and fuzzy systems solve problems by performing function approximation. If we want to use a neural network for solving a given problem, we must describe the problem sufficiently by means of sample data. We do not need a mathematical model for the problem of interest, and we do not need any form of prior knowledge.

On the other hand we cannot interpret the solution obtained from the learning process. A fuzzy system can be used to solve a problem if we have knowledge about the solution in the form of linguistic if-then rules. The main benefit of a fuzzy system is that it lets you define the desired system behavior with simple if-then relations. In addition, we can use all available engineering know-how to optimize the performance directly. But, we are lost

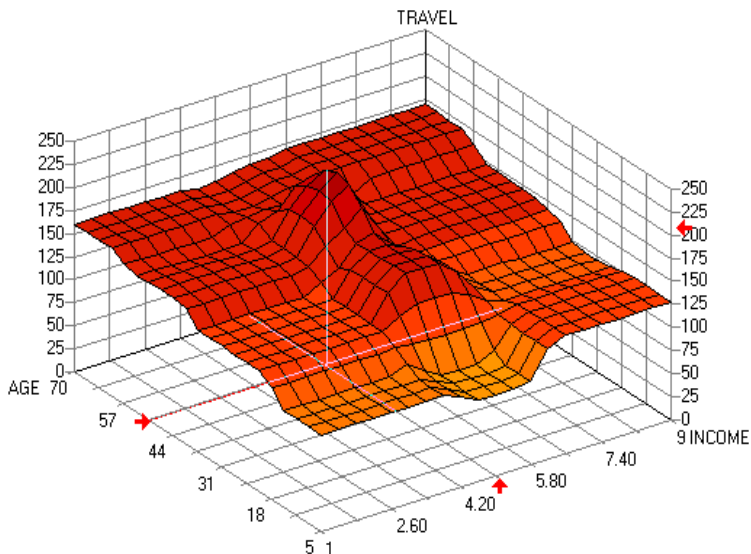
without "if-then" rules, to make the fuzzy system work we may need a long-term tuning process. The main reason for combining fuzzy systems with neural networks is their learning capability. Such a combination should be able to "learn" linguistic rules and/or membership functions, or to optimize existing ones. Both neural networks and fuzzy systems are powerful design techniques that have their strengths and weaknesses. Neural networks can learn from data sets while fuzzy system solutions are easy to verify and optimize. A combination of the explicit knowledge representation of fuzzy systems with the learning power of neural networks results in the neuro-fuzzy system.

For neuro-fuzzy training, we can use sample customer data. The data set contains 100 examples of a customer's age, income and total spending on travel within a certain time period. Next we can use this data for neuro-fuzzy training to extract the knowledge rules that the experts followed to come up with decisions.

Figure 5 shows a 3D plot for a training result of a neuro-fuzzy data file. Likewise we can apply this step for training whole rule bases. Figure 6 shows trained rules sorted by the terms of DoS³ (Degree of Support). Note that the first two rules in the rule block have 1.00 DoS. It means that those two rules perfectly represent the expert's knowledge based on these data.

IF Age IS young AND Income IS high THEN
Interest_travel IS medium
IF Age IS middle AND Income IS high THEN
Interest_travel IS high

Figure 5: 3D plot showing the training result



³ This degree of support (DoS) is also known as the weight of the rule or truth function since it established a one-to-one correspondence between an element in the domain and a truth value indicating its degree of membership in the set.

Figure 6: Trained rules sorted by the terms of DoS

#	IF		THEN	
	AGE	INCOME	DoS	TRAVEL
1	Young	high	1.00	medium
2	Middle	high	1.00	high
3	Very_Young	medium	0.90	low
4	Young	medium	0.77	high
5	Very_Young	low	0.65	low
6	Middle	medium	0.64	high
7	Middle	high	0.60	low
8	Young	low	0.55	high
9	Very_Young	high	0.50	high
10	Very_Young	high	0.48	low
11	Young	low	0.48	low
12	Young	low	0.41	medium
13	Middle	low	0.38	high
14	Old	high	0.37	high

5. Conclusion and discussion

The basic aim of this research is to develop a web ad selection model for one-to-one advertising using a neuro-fuzzy system. This model is derived from multiple sources of factors using approximate knowledge. In this paper we presented a one-to-one web advertising model, a web ads selection process, a simulated example and the structure of the system using a sample scenario. From sets of real world behavior data, expert knowledge or rule can be discovered and trained using a neuro-fuzzy system. We also showed neuro-fuzzy training to extract knowledge rules using sample customer data. To develop a learning relationship with a customer, we suggest the one-to-one web advertising model. However, the most important factor for the success of this one-to-one web advertising model is to maintain the trust of customers.

Also, a matter to be considered in future research would be the efficiency and effectiveness of the proposed one-to-one advertising model as compared empirically with a conventional model in a real situation.

Issues that are important in this area but have not been discussed here are customer privacy and ethical concerns on B2C web based EC models. They are the greatest threats to the development of electronic commerce (Kalin 1998; Kovacich 1998; Monahan 1998; Wagner 1996). To overcome this obstacle, the Federal Trade Commission's proposed legislative model that is built around four basic fair information practices: 1. Notice/awareness (Give consumers notice of what information is collected and how it is used.) 2. Choice/consent (Let consumers choose whether to allow secondary uses of information.) 3. Access/participation (Give consumers reasonable access to their information and let them correct any errors.) 4. Security/integrity (Ensure the information's security.) Information collected by this model can be associated with a customer's personally identifiable information only if that user has agreed to receive personally customized

ads. In addition, web retailers and web advertisers have to maintain internal practices that help to protect the security and confidentiality of customer information by limiting employee access to and use of this information.

References

Altrock, C. v. (1997) *Fuzzy Logic and NeuroFuzzy Applications in Business and Finance*: Prentice Hall PTR.

Anonymous (1998) "FTC sets year-end deadline for industry self-regulation," *ABA Bank Compliance* (19) 9 (Regulatory & Legislative Advisory), pp. 2.

Dewan, R., B. Jing, and A. Seidmann. (1999) "One-to-one Marketing on the Internet." *International Conference on Information Systems, Charlotte, North Carolina, 1999*, pp. 93-102.

Jang, J.-S. R. and C.-T. Sun (1996) *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*: Prentice Hall.

Kalin, S. (1998) "Saving private data," *Cio* (12) 1 (Section 2), pp. 32-42.

Kovacich, G. (1998) "Electronic-Internet Business and Security," *Computers & Security* (17) 2, .

Monahan, P. (1998) "Deconstructing Information Walls: The Impact of the European Data Directive on U.S. Businesses," *Law and Policy in International Business* (29) 2, .

Peppers, D. and M. Rogers (1994) *The one-to-one future.*: Judy Piatus Ltd.

Peppers, D., M. Rogers, and B. Dorf (1999) *The One to One Field Book*: Doubleday.

Raghu, T. S., P. K. Kannan, H. R. Rao, and A. B. Whinston (2000) "Dynamic Profiling of Consumers for Customized Offerings over the Internet: A Model and Analysis," *Working paper*, SUNY Buffalo

Sullivan, C. (1999) "The Times they are a-targeting," *Editor & Publisher* (132) 16, pp. 44-45.

Wagner, M. (1996) "Web on privacy alert," *Computerworld* (30) 18, pp. 1,16.

Wong, B. K., T. A. Bodnovich, and Y. Selvi (1997) "Neural network applications in business: A review and analysis of the literature (1988-95)," *Decision Support Systems* (19) 4, pp. 301-320.

Yager, R. R. (1997) "Intelligent Agents for World Wide Web Advertising Decisions," *International Journal of Intelligent Systems* (12) 5, pp. 379-390.

Zadeh, L. A. (1984) "Making Computers Think Like People," *IEEE Spectrum* (21) 8, pp. 26-32.

Zadeh, L. A. (1994) "Fuzzy logic, neural networks, and soft computing," *Communications of the ACM* (37) 3, pp. 77-84.

Zadeh, L. A. (1996) "Fuzzy control: A personal perspective by Lotfi A. Zadeh," *Control Engineering* (43) 10, pp. 51-52.