A Similar Product Recommendation Algorithm and Its Application for Global Sales and Local Delivery in Collaborative Commerce

Sang Hyun Choi
Assistant Professor at Department of Industrial and Systems Engineering Engineering Research Institute, Gyeongsang National University.
(chois@gsnu.ac.kr)
Sungmin Kang
Assistant Professor of MIS at College of Business Administration Chung-Ang University.
(smkang@cau.ac.kr)

In this paper, we present a similar product finding algorithm for the cooperative business companies that share the product taxonomy table and have exchangeable products information. The main idea of the proposed algorithm is to compute the aggregated utility ranges over specification values of products in the same product class of the companies and find similar ones between the products. Experimental results from a laboratory application are provided in terms of user satisfaction rates. Their comparative implications with a distance measure-based recommendation algorithm are suggested. The experiment confirmed that our algorithm could be used as a solution for similar product recommendation when a buyer's residence site was different from a receiver's one and the buyer had his/her own incomplete information about the weights of product specifications.

Key words: Product recommendation, Similar product, C-Commerce, Product taxonomy, Utility

Ⅰ. Introduction

Electronic commerce is an Internet application and it depends on key infrastructures such as information technology & telecommunications, social/cultural, commercial, and government/legal. In particular cultural elements such as language, education level, beliefs, and value systems influence technological innovations and entrepreneurial spirit. An online survey conducted by International Data Corporation (IDC) suggests that over 76 percent of Chinese respondents prefer to browse the Internet in their local language, and not in English. When viewed as a global distribution system, international social/cultural barriers remain. Just as many

논문접수일: 2004. 3　　게재확정일: 2004. 11

companies have made serious blunders when marketing in other countries, the lack of boundaries and the complexity of global consumer access magnify these complexities beyond anything previously encountered (Javalgi and Ramsey, 2001). Cultural factors inhibit the diffusion of electronic commerce. They comment that cultural values, including different traditions and habits of trading will impact the speed of e-commerce diffusion. Clearly, cultural issues including beliefs, languages, and value systems seem to present barriers to information sharing (Steinfield and Klein, 1999). In order for companies targeting global consumers to overcome the barriers, they allow the consumers to browse the Internet in their local language. Local companies need to have a cooperative strategy of performing vicarious delivery transaction for ordered goods at another area.

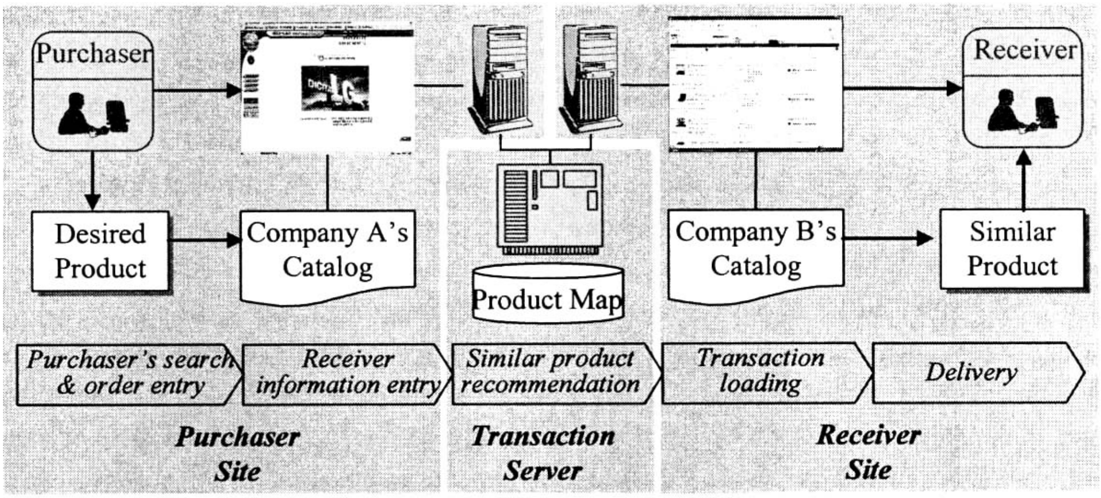
The electronic markets offering full support for all market transactions provide the following services or phases; knowledge exchange, articulation & management of intensions, negotiation & contracting, and settlement in the form of payment and delivery (Grieger, 2003; Schmid, 1997; Schubert, 2000; Slabeva and Schmid, 2000). Usually the online shopping takes a place within a same country. When a consumer who lives in Korea wants to buy a product at Internet shopping mall and present it to a relative who lives in America, there are two ways to do. First way is to access a Korean online shopping website that is written in Korean and then make the company deliver to America. If the consumer who lives in Korea orders the items from the Korean merchant's WWW page and makes the merchant ship the goods to America, he or she must pay for the extra shipping cost (Kalakota and Whinston, 1996). The merchant has to perform such business processes as delivery by airline or ship. Second way is to access an American website at his or her relative's area and make the company deliver it within the same country. The consumer doesn't need to pay for extra delivery from Korea to America. In order to use the web site at another country, there are such limitations as language, shopping culture, and so on (Javalgi and Ramsey, 2001; Steinfield and Klein, 1999).

Another method considered in this paper is based on a shared product taxonomy table between cooperative companies. Business opportunities of this method are in managing the product information (Adomavicius and Tuzhilin, 2001; Brew, 1991; Cho, Kim, and Kim, 2002). It is important to register and recommend similar products among the cooperative companies to the customer. The researches of comparison shopping agent are similar to this research (Doorenbos, Etzioni, and Weld, 1997; Iyengar and Dias, 1998; Yang, Choi, Kim, and Ham, 2000). Bargain-Finder and Jango are the examples of first-stage comparison shoppers that specify the functions that agents must have in order to be applied to Electronic Commerce, and both employ the manual rule extraction method. Shopbot (Doorenbos, Etzioni and Weld, 1997) suggests an automatic rule extraction technique by analyzing and learning the shopping malls. There are lots of researches related to the similar product registration. Most of traditional researches of the clustering algorithm are interested in the traditional recommendation problem which is to find k nearest products considering same and equal weighted product specifications from the large database (Shyu, Haruechaiyasak, and Chen, 2003; Sarwar et al., 2000; Sarwar et al., 2001). The researches are based on similarity measures such as cosine coefficient and distance between products.

The similar product recommendation problem in this paper is different with traditional recommendation problems in two aspects. One is that the cooperative companies share their product information for sales and delivery using the product taxonomy table. The companies share the table that contains exchangeable similar product lists at same classes. The other is that the problem assumes that the products at same class are specified by the same set of product specifications and customers or product managers have their own weight of the features. The weight information is given by them. In some practice, however, it is no simple matter to actually measure the exact values of weights. Rather, product managers or customers give only certain linear relations which express imprecise information about the weights of features.

We suggest an interactive algorithm for finding similar products when the weights of feature are not same and are given by the imprecise information. The main idea of the proposed algorithm is to find the utility range of products in a product class of the companies and register them as exchangeable similar products. The companies then allow consumer to shop and purchase the products at their own residence site and deliver them to another sites. The rest of paper is organized as follows. Section 2 describes the cooperative business process for a similar product recommendation. Section 3 suggests a procedure for the similar product recommendation and computational experiments, and results are included in Section 4. Section 5 concludes the paper.

〈Figure 1〉 Three Components of Collaborative Business Model



Ⅱ. Similar Product Recommendation Problem for a Collaborative Commerce

In the collaborative electronic commerce, three kinds of components exist: purchaser site, delivery site, and transaction server. The purchaser and delivery sites have its web sites and product catalogs and transaction server has a shared product map database between the cooperative sites. The server finds same or similar products with the product that a customer wants to buy and save transaction records such as purchasing information, delivery information, and so on. The records will be used to allocate profit to each related entities according to their activities.

In the collaborative electronic commerce, the overall process consists of four sub processes as shown in 〈Figure 2〉.

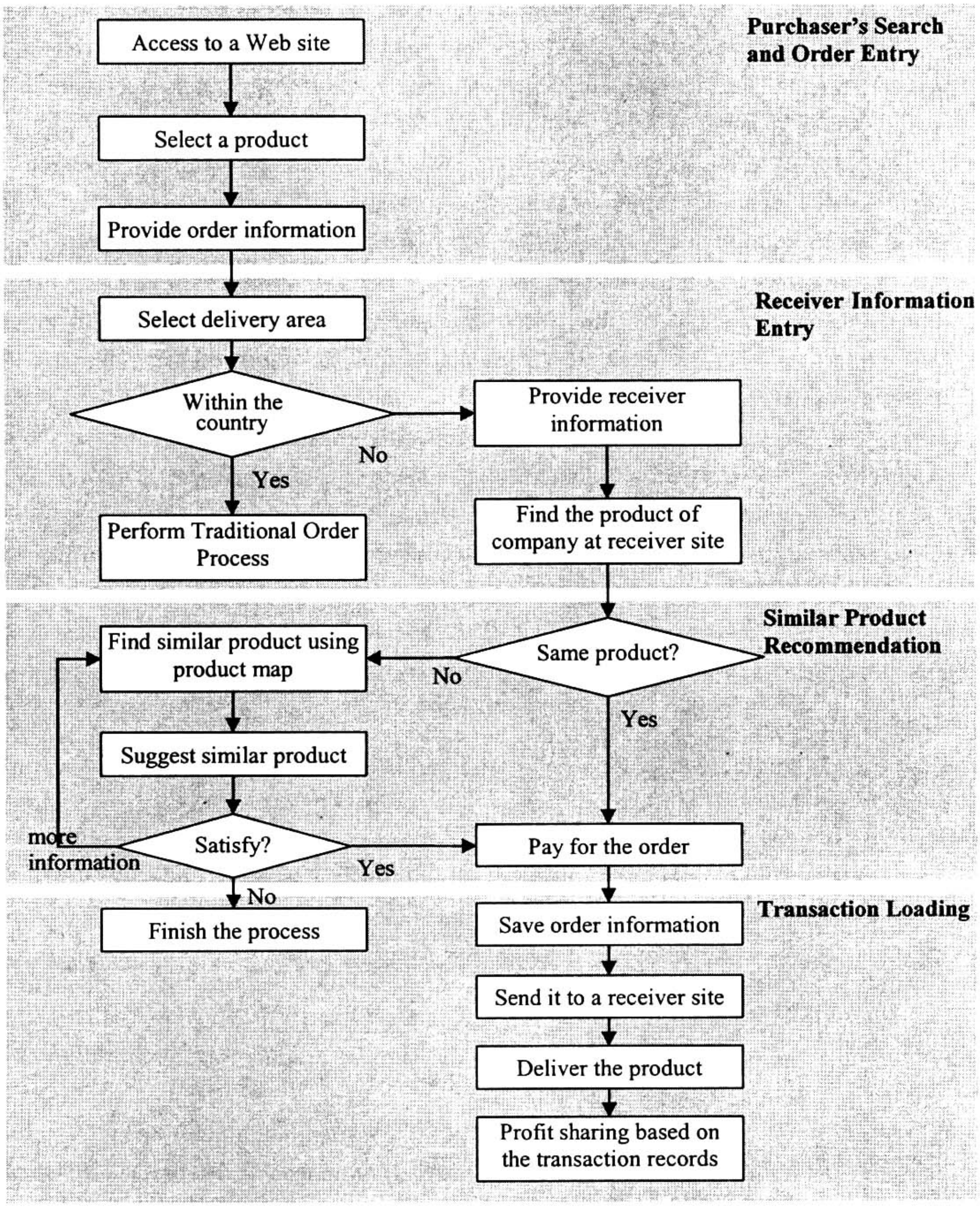
2.1 Purchaser's Search and Order Entry

An user who wants to give a present at Internet shopping mall accesses a website at his or her own residence areas. The user surfs the web and finds a product in a product catalog. If the user searches for the desired product, the user would provide the mall with product order information.

2.2 Receiver Information Entry

The purchaser selects a delivery address type of whether a receipt address is within the residence area. If the receiver address is within the purchaser's country, a traditional

〈Figure 2〉 A Process for the Cooperative Electronic Commerce Model



delivery process is performed by the purchaser site. If the purchaser country isn't same as receiver one, a transaction server finds a cooperative company close to the receiver address and search a same product among the catalog of the company. The server finds companies in the nearest order. If the server finds same product, the delivery site get a delivery order from the server. Otherwise go to next process.

2.3 Similar Product Recommendation

The transaction server searches similar products at the product map of a shared product database. The product map is a product relationship matrix that defines the exchangeable relations between products of the cooperative companies. They share orderable products and their customer orders any product among their shard products. The product map helps the company to drive down delivery cost and reduce prices to its customer. Furthermore, it increases cooperation between the companies as they strive for quick deliveries and low inventories. A detail algorithm for finding the similar product is presented in the next section.

The server shows the user a set of k similar products with the item he/she has just selected. If the user is not satisfied with anything of interest, he or she doesn't purchase it on-line and the transaction is finished. If the user does find alternative product of interest, he or she elects to purchase it on-line, make an order, and provide payment information. Lastly, the consumer selects the means of payment.

2.4 Transaction Loading & Delivery

The consumer sends the merchant a complete order including receiver's address. The merchant requests payment authorization from the consumer's bank and sends the customer a confirmation of the order shipment and payment. The merchant saves the transaction and makes the cooperative company at the receiver area ships the goods to the recorded address. The server saves the transaction. The company at the receiver's area ships the goods in the shipment method as requested by the purchaser.

III. Similar Product Recommendation Procedure

3.1 Mathematical Problem Definition for Similar Product Recommendation

In this section, we formally define the similar product recommendation problem. The goal considered in this problem is to find the most similar product with the product of cooperative company. The selected product is being owned by purchaser site (shortly, A company). Suppose that the company has total of K classes or categories. Each class is characterized by a set of product specifications. Then a new product is assigned to a class having a same set of product specifications. The product at the kth class has a set of M product specifications. We define the following terminologies.

$I=\{i\}_{i=1,N}$: a set of N products at kth class
$J=\{j\}_{j=1,M}$: a set of M product specifications at kth class
w_j : importance of jth specification
x_{ij} : jth specification value of ith product
$x_i = \{x_{i1}, x_{i2}, \ldots, x_{iM}\}$ for i (I: the specification values of ith product at the same class
$x_s^A = \{x_{s1}^A, x_{s2}^A, \ldots, x_{sj}^A, \ldots, x_{sM}^A\}$: the product that the customer has just selected at the A company's internet shopping site

Multi-criteria decision analysis is applied to this problem because the analysis deals with situations in which decision alternatives, such as products, are evaluated on a finite number of attributes, such as specifications. One of the best known and the most widely used ways to evaluate alternative $x=(x_1, \ldots, x_M)$ is to utilize the weighted additive decomposition

$$v(x) = \sum_{j=1}^{M} w_j v_j(x) \tag{1}$$

of a value function v. Here, v_j is the marginal value function of specification j such that v_j : $x_j \to ([0, 1]$ and $w_j \geq 0$ is the weight which represents the relative importance of the jth specification. This evaluation function has been well examined in the literature (e.g., Keeney and Raiffa, 1976; Dyer and Sarin, 1979; Kim and Choi, 2001). Suppose that there is a finite set of all available products, $X=\{x_1, \cdots, x_i, \cdots, x_N\} \subset R^M$ and the cooperative company doesn't have the product x_s^A. The company then searches a similar product x_s, having the same or closest aggregated value.

$$v(x_s) = \sum_{j=1}^{M} w_j v_j(x_{sj}) \cong v(x_s^A) \tag{2}$$

Then this problem implies finding products having same or similar values as $v(x_s^A)$ among a set of available products, X, at the cooperative company. In this paper, we define the similar products, x_s and x_s^A, as the products that don't have a strict dominance relation and have the most similar utility values. The strict dominance concept is suggested by Park, Kim, and Yoon (1996). The condition is that if the minimal value of $\sum_j w_j[v_j(x_{sj})-v_j(x_{sj}^A)]$ is equal to or more than 0, x_s strictly dominates x_s^A. If the decision parameters w_j and $v_j(\cdot)$ are all exactly or numerically assessed by the decision maker, finding x_s is done by simple calculation using (2). In some practice, however, it is no simple matter to actually measure the exact values of weights. Rather, the decision maker gives only certain linear relations which express imprecise information about the weights. Examples of imprecise weights are in the five forms of bounded descriptions: weak preference ($w_i \geq w_j$), strict preference ($w_i - w_j \geq \varepsilon$), preference with multiple ($w_i \geq a_{ij} w_j$), interval preference ($l_i \leq w_i \leq u_i$), and preference difference ($w_i - w_j \leq w_l - w_m$). Use of linear programming approaches for some special and/or arbitrary forms of imprecise weights has been discussed in the literature: while identifying only non-dominated alternatives, methods for simultaneously dealing with dominance and potential optimality can be found (Eum, Park, and Kim, 2001). In this paper, we suggest the utility range concept which is described in the next section.

3.2 An Interactive Procedure Given Exact Specification Values

The procedure is based on an utility range concept that an utility of a product can be represented by a range. We assume that a product utility value on a specification can be computed by the normalization formula (3). The weights of specification within a same class are in the form of constraints given by product managers. And, the aggregated product utility range of all specifications is computed by solving LP models having the constraints about the importance relationship between product specifications. Our procedure for finding similar products is composed of the following 4 sub-steps.

Step 1: Gather and normalize product specification values

In this step, a product manager gathers information of specification values of the products at the cooperative company. We define the utility of a product specification as the normalized value computed by following formula and the utility is between 0 and 1.

$$v_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad \text{for specification j with better for larger}$$

or

$$v_{ij} = \frac{x_{ij} - \max_i x_{ij}}{\min_i x_{ij} - \max_i x_{ij}} \quad \text{for specification j with better for smaller} \tag{3}$$

Step 2: Compute the utility ranges of current products

This step is used to compute the utility range of products at kth level. The utility of ith product is the weighted sum over all specification utilities, that is, $\sum_j w_j v_{ij}$. It is very difficult to find exact weight values and compute the product utility. The weight information represents the relationship between specification weights with the 5 forms of incomplete information provided by an user. We define the relationship set as Φ_W which is a set derived from the manager's incomplete information regarding the relative importance of specifications. We can get the utility ranges of ith product, $[v_i(\min), v_i(\max)]$, by following formula.

$$v_i(\min) = \min \Sigma_j w_j \cdot v_{ij} \text{ subject to } \Phi_W \text{ and}$$
$$vi(\max) = \max \Sigma_j w_j \cdot v_{ij} \text{ subject to } \Phi_W \quad (4)$$

Finally we can get the expected utility, $E[v_i]$, by computing an average value of $v_i(\min)$ and $v_i(\max)$.

Step 3: Classify the products by the utility ranges

In this step we subdivide the products at the same class into similar product classes using the expected utilities of products. From the expected utilities, we obtain the abstract difference values between products, $DE[v_{mn}] = \left|\overline{v_n} - \overline{v_m}\right|$, where, $\overline{v_n}$ and $\overline{v_m}$ mean respectively average utility value of product v_n and v_m. If two products don't have a strict dominance relation and the value is equal to or less than threshold value, δ, then nth and mth products are registered into product table as a similar product class. This procedure is performed over all products pairs. After having the similar product pairs, the sub-classes at kth class are identified.

Step 4: Determine the sub-class of a new product

Based on the current product taxonomy and catalog, the product manager registers a new product. He/She gathers the specification values for the new product and performs the above classification steps through three steps. If all specification values of the new product are between current minimum and maximum values, it is classified into one among the pre-specified sub-classes. If it has a new minimum or maximum specification value, the utility values of all product specifications are changed and current sub-classes are changed.

3.3 An Interactive Procedure Given Inexact Specification Values

Step 1: Gather and normalize exact specification values

The Step 1 in section 3.2 is performed for the products having exact specification values.

Step 2.0: Input incomplete information and compute value ranges of inexact specification values

If there are unknown specification values, this step is performed. Otherwise go to step 2. The manager enters relationship information of unknown specification values. We define the relationship set as Φ_v. Then we compute value ranges of unknown specification values by following formula.

$$v_{ij}^{min} = \text{minimize s.t. } \Phi_v \text{ and}$$
$$v_{ij}^{max} = \text{maximize s.t. } \Phi_v \qquad (5)$$

Step 2: Compute aggregated value ranges of products

The Step 2 in section 3.2 is performed except unknown specification computation. For minimum or maximum value computation, unknown specification values are replaced by v_{ij}^{min} or v_{ij}^{max} respectively. Finally we can we can get the expected utility, E[v_i], by computing an average value of v_i(min) and v_i(max). The rest stes are performed in a same way as in section 3.2.

Ⅳ. Computational Experiments

4.1 Experimental Data Sets

Using the procedure discussed in this paper, we carry out the experiments with the intent of answering two major questions, first one is what is different about the procedure in comparison to a recommendation method based on the Euclidean distance measure and second one is how the different relationship sets of specification weights affect the similar product recommendation.

For our experiments, 6 user groups who are interested in buying the Pentium 4 2.0 GHz laptop computer participated in the experiment. We used 6 computer information from Korean L Internet shopping mall at the buyer site and 55 computer information from American B mall at the receiver site. Each group selected a laptop computer in the L mall at the buyer site. We explained the cooperative business model. They couldn't present the product selected at the L mall to the relative in receiver site as the B mall didn't have same product. We were going to suggest similar products based on importance relationships of product specifications. They had to provide the incompletely specified information of specification weights. The computers are specified by RAM size (w_1), hard disk size (w_2), types of display (w_3), CD-ROM (w_4), brand name (w_5), and price (w_6).

We asked each group to enter an agreed set of relationship information between specification weights with the 5 forms of incomplete information. For example, the first group provided the following information (6). The group suggested that the weights of all specifications are equal to or greater than 0.1, and

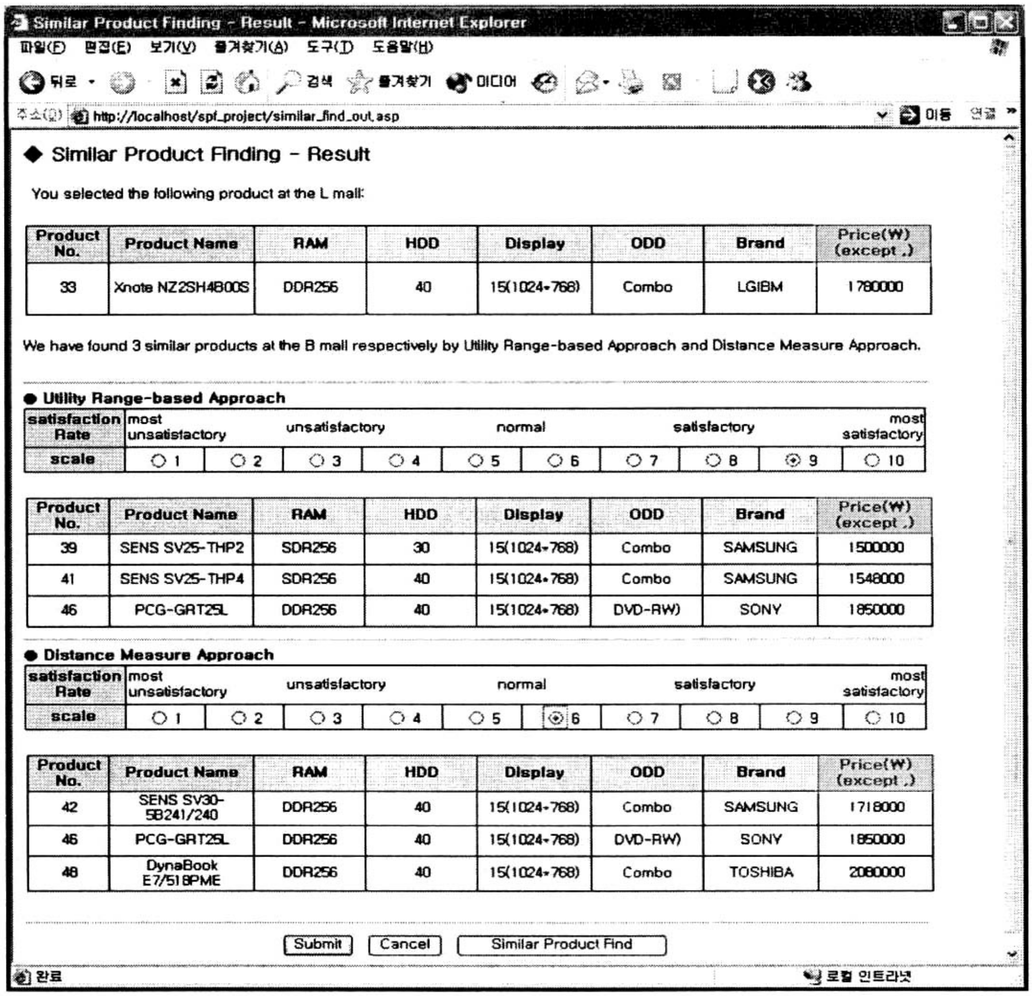
⟨Table 1⟩ Six Results of Similar Product Recommendation Based on Utility Ranges

Important Criteria*	Product Name	RAM		HDD	Display		Brand		Price
		Spec	Util	Util	Spec**	Util	Name	Util	Util
Selected Product	NZ2SH4B	DDR256	0.67	0.67	1024*768	0.6	LGIBM	0.8	0.34
Price Brand	Sens JB242	SDR256	0.33	0.67	1024*768	0.6	Samsung	0.8	0.4
	Sens THP4	SDR256	0.33	0.67	1024*768	0.6	Samsung	0.8	0.42
	PCG-GRT25	DDR256	0.67	0.67	1024*768	0.6	SONY	1	0.31
Price RAM	N7600.4	DDR256	0.67	0.67	1024*768	0.6	TG	0.2	0.47
	N7620.4	DDR256	0.67	0.67	1024*768	0.6	TG	0.2	0.47
	C2220	DDR256	0.67	0.67	1024*768	0.6	Fujitsu	0.4	0.42
Price Display	Sens JB242	SDR256	0.33	0.67	1024*768	0.6	Samsung	0.8	0.4
	N7620.4	DDR256	0.67	0.67	1024*768	0.6	TG	0.2	0.47
	N7600.4	DDR256	0.67	0.67	1024*768	0.6	TG	0.2	0.47
Brand Price	Sens 5B241	DDR256	0.67	0.67	1024*768	0.6	Samsung	0.8	0.36
	NZ3SH4B	DDR256	0.67	0.67	1024*768	0.6	LGIBM	0.8	0.37
	E7/518	DDR256	0.67	0.67	1024*768	0.6	Toshiba	1	0.23
Brand Display	Latitude C840	DDR256	0.67	0.67	1400*1050	0.8	DELL	0.6	0.3
	Sens 5B241	DDR256	0.67	0.67	1024*768	0.6	Samsung	0.8	0.36
	NZ3SH4B00S	DDR256	0.67	0.67	1024*768	0.6	LGIBM	0.8	0.37
Display RAM	S7601.3	DDR256	0.67	0.67	1400*1050	0.8	TG	0.2	0.41
	S7581.3	DDR256	0.67	0.67	1400*1050	0.8	TG	0.2	0.4
	Sens 5B241	DDR256	0.67	0.67	1024*768	0.6	Samsung	0.8	0.36
Distance measure-based approach	Sens 5B241	DDR256	0.67	0.67	1024*768	0.6	Samsung	0.8	0.36
	PCG-GRT25	DDR256	0.67	0.67	1024*768	0.6	SONY	1	0.31
	E7/518	DDR256	0.67	0.67	1024*768	0.6	Toshiba	1	0.23

*: All specification values and utility values of ODD in this table are Combo and 1's. Spec and Util mean respectively Specification value and Utility value.

**: All specification types of Display in this table are 15 inch.

〈Figure 3〉 Results and Satisfaction Rate Survey Sheet of Similar Product Recommendations



Similar Product Finding - Result - Microsoft Internet Explorer

http://localhost/spf_project/similar_find_out.asp

◆ Similar Product Finding - Result

You selected the following product at the L mall:

Product No.	Product Name	RAM	HDD	Display	ODD	Brand	Price(₩) (except.)
33	Xnote NZ2SH4B00S	DDR256	40	15(1024*768)	Combo	LGIBM	1780000

We have found 3 similar products at the B mall respectively by Utility Range-based Approach and Distance Measure Approach.

● Utility Range-based Approach

satisfaction Rate	most unsatisfactory		unsatisfactory		normal		satisfactory			most satisfactory
scale	○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	⊙ 9	○ 10

Product No.	Product Name	RAM	HDD	Display	ODD	Brand	Price(₩) (except.)
39	SENS SV25-THP2	SDR256	30	15(1024*768)	Combo	SAMSUNG	1500000
41	SENS SV25-THP4	SDR256	40	15(1024*768)	Combo	SAMSUNG	1548000
46	PCG-GRT25L	DDR256	40	15(1024*768)	DVD-RW)	SONY	1850000

● Distance Measure Approach

satisfaction Rate	most unsatisfactory		unsatisfactory		normal		satisfactory			most satisfactory
scale	○ 1	○ 2	○ 3	○ 4	○ 5	⊙ 6	○ 7	○ 8	○ 9	○ 10

Product No.	Product Name	RAM	HDD	Display	ODD	Brand	Price(₩) (except.)
42	SENS SV30-5B241/240	DDR256	40	15(1024*768)	Combo	SAMSUNG	1718000
46	PCG-GRT25L	DDR256	40	15(1024*768)	DVD-RW)	SONY	1850000
48	DynaBook E7/518PME	DDR256	40	15(1024*768)	Combo	TOSHIBA	2080000

Submit　Cancel　Similar Product Find

brand & price specifications are more important than RAM size, hard disk size, types of display, and CD-ROM. The following conditions of specification weights are summarized.

$$
\begin{aligned}
& w_5,\ w_6 \geq w_1,\ w_2,\ w_3,\ w_4 \\
& w_1 \geq w_2 \\
& w_4 \geq w_3 \\
& w_3 \geq w_2 \\
& w_5 + w_6 \geq w_1 + w_2 + w_3 + w_4 \\
& w_6 - w_5 \geq w_1 - w_4 \\
& w_1,\ w_2,\ w_3,\ w_4 \geq 0.1 \\
& w_1 + w_2 + w_3 + w_4 + w_5 + w_6 = 1
\end{aligned} \tag{6}
$$

The above group focused on the price and brand specifications. The group reviewed the result of similar products provided by our procedure and didn't satisfy it. The group revised the constraint as following.

$$w_6 \geq 2w_1,\ 2w_2,\ 2w_3,\ 2w_4$$
$$w_1,\ w_2,\ w_3,\ w_4 \geq 0.05 \qquad (7)$$
The rest of constraint is same as (6)

The group is satisfied with the current result and then performed a satisfaction rate survey. For convenience, the rest 5 groups suggested the most and next important specifications. The constraint set are assumed to have same format as formula (7). Second group suggested price and RAM size, third group price and type of display, fourth group brand and price, fifth group brand and type of display, and sixth group type of display and RAM size as shown the important criteria column in Table 1. After gathering the incomplete information of specification weights by each group, we calculated the utility range of 56 computers using formula (4) and obtained the abstract difference values, DE[v_{ij}], between a selected product and the others. We suggested three computers that are closest to the selected product as shown in Table 1.

To answer the first question, we provided users with the results of utility range-based approach and distance measure approach and gathered the response information of users' satisfaction rates as shown in Figure 3. In the distance measure approach, we used same normalized utility values of specification values as our approach. We suggested three computers that are closest to the selected product among 55 computers. For the second question, we summarized the characteristics of the results from 6 utility range-based computations and distance measure-based approach.

4.2 Result and Discussion

A t-Test model was used to analyze users' responses to the post-study questionnaire. The results of the user responses are shown in Table 2. Here, the satisfaction rate is used as a measure to compare the performance of the two approaches. From the table, it can be seen that the average satisfaction rate of utility range-based approach is higher than that of distance measure. Through the post-study questionnaire, we have known that it is because the utility range-based approach allowed users to suggest their opinions about the weight relationships of product specifications. This proves that the utility range-based approach is a viable solution to the problems currently encountered in similar product recommendation on the cooperative business model.

Examining the products obtained from the utility range-based approach and distance

〈Table 2〉 Result of t-Test for Two Approaches

<table>
<tr><th colspan="6">Statistics</th></tr>
<tr><td>TYPE</td><td>Variable</td><td>N</td><td>Mean</td><td>Std Dev</td><td>Std Err</td></tr>
<tr><td>Utility Range</td><td>Satisfaction Rate</td><td>30</td><td>8.9</td><td>0.7589</td><td>0.1385</td></tr>
<tr><td>Distance Measure</td><td>Satisfaction Rate</td><td>30</td><td>6.9</td><td>1.125</td><td>0.2054</td></tr>
<tr><td>RATE</td><td>Diff (1-2)</td><td>30</td><td>2</td><td>0.9595</td><td>0.2477</td></tr>
<tr><th colspan="6">t-Tests</th></tr>
<tr><td>Variable</td><td>Method</td><td>Variances</td><td>DF</td><td>t Value</td><td>Pr > |t|</td></tr>
<tr><td>RATE</td><td>Pooled</td><td>Equal</td><td>58</td><td>8.07</td><td><.0001</td></tr>
<tr><td>RATE</td><td>Satterthwaite</td><td>Unequal</td><td>50.9</td><td>8.07</td><td><.0001</td></tr>
</table>

measure-based approach helps us to show the distinct characteristics of two approaches. The first result was focused on price and brand specifications, and it showed the products that had the same or higher utility values of those specifications. The utility values of RAM and type of display specification were relatively little considered. The second result based on price and RAM-oriented utility showed the products having same utility values of RAM. The utility values of brand specification were relatively little considered and compromised by much higher values of price than the selected product. In the third result focused on price and type of display specifications, the values of display specification are same and ones of price specification were a little higher than the value of selected product. The utility values of brand specification were relatively little considered. In the fourth result, brand and price specifications were focused and had similar values or compromising values. The values of RAM and display specifications were same. The fifth result was focused on brand and display specifications. In this result, the products compromising the values of brand and display specifications were selected. The values of price specification were a little lower than those in the other results. In the sixth result focused on display and RAM specifications, the values of display and RAM were similar to the selected product. The utility values of brand and price specifications were relatively little considered.

In the distance measure approach, the values of RAM, HDD sizes, and display specifications were same and the result was very similar to the fourth result focused on brand and price specifications. The values of brand and price specification were almost same and a little lower than selected product. In the approach, the products having similar values across all the specifications were selected. The users are interested in a set of specific specifications and the approach needs to be revised. For example, they have to provide exact weights of all specification to reflect their opinions. It is not easy to reflect users' opinions about the weights of product specification in a flexible way.

V. Conclusion and Further Study

In this paper, we explained a global electronic business situation that a purchaser residence area is different from a receiver's one. In order to do this kind of business, we suggested an interactive algorithm for finding similar products. This algorithm is based on the user's utility ranges of products. The users can easily enter their preferences of product specifications into incomplete information. Usually the users on the Web have their own preference system. The utility range-based approach suggested in this paper can get users' preference information in the form of incomplete information. It is very easy to reflect users' opinions about the weights of product specification in a flexible way.

To do this, the cooperative companies have to share product taxonomy table and register their own products and exchangeable products. After completing the integration of product categories among cooperative partners, the cooperative business process explained in this paper can be preformed. The goal of this process is to specifically classify products of trading partners into similar product subclasses based on their specification values. This model helps local companies to perform a cooperative strategy of performing vicarious delivery transaction for ordered goods at another area. The business model suggested in this research makes the companies that are difficult to set up physical branches at every corner of the world to cooperate with another companies at the sites in a cost efficient way. The cooperative companies can globally expand their electronic commerce business without a heavy facility investment. The companies are able to give the customer easy access to their own country and save the delivery cost.

The complexity of developing and operating a global electronic commerce model calls for a collaboration platform between the companies that contains rules of similar product registration, profit sharing, and so on. Future research in this model should focus on an in-depth investigation of the required regulations and their economic justification.

REFERENCES

Adomavicius, G. and A. Tuzhilin (2001), "Expert-Driven Validation of Rule-Based User Models in Personalization Applications," *Data Mining and Knowledge Discovery*, 5 (1-2), 33-58.

Brew, C. (1991), "Systemic Classification and Its Efficiency," *Computational Linguistics*, 17(4), 375-408.

Cho, Y. H., J. K. Kim, and S. H. Kim (2002), "A Personalized Recommender System Based on Web Usage Mining and Decision Tree Induction," *Expert Systems with Applications*, 23(3), 329-342.

Doorenbos, R. B., O. Etzioni, and D. S. Weld (1997), "A Scalable Comparison Shopping Agent for the World Wide Web," *Proceedings of the First International Conference on Autonomous Agent*, 39-48

Dyer J. S. and R. K. Sarin (1979), "Measurable Multiattribute Value Functions," *Operational Research*, 27(4), 810-822.

Eum, Y. S., K. S. Park, and S. H. Kim (2001), "Establishing Dominance and Potential Optimality in Multi-Criteria Analysis with Imprecise Weight and Value," *Computers & Operations Research*, 28, 397-409.

Grieger, M. (2003), "Electronic Marketplace: A Literature Review and a Call for Supply Chain Management Research," *European Journal of Operational Research*, 144, 280- 294

Iyengar, A. and D. Dias (1998), "Distributed Virtual Malls on the World Wide Web," *International Conference on Distributed Computing Systems*, 58-65.

Javalgi, R. and R. Ramsey (2001), "Strategic Issues of E-Commerce as an Alternative Global Distribution System," *International Marketing Review*, 18(4), 376-391.

Kalakota, R. and A. B. Whinston (1996), *Electronic Commerce: A Manager's Guide*, Addison Wesley.

Keeney, R. L. and H. Raiffa (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoff*, New York, Wiley.

Kim, J. K. and S. H. Choi (2001), "An Utility Range Based Interactive Group Support System for Multiattribute Decision Making," *Computers and Operations Research*, 28(5), 485-503.

Park, K. S., S. H. Kim, and W. C. Yoon (1996), "Establishing Strict Dominance between Alternatives with Special Type of Incomplete Information," *European Journal of Operational Research*, 96, 398-406.

Sarwar, B., G. Karypis, J. Konstan, and J. Riedl (2000), "Analysis of Recommendation Algorithms for E-Commerce," *Proceedings of ACM E-Commerce 2000 Conference*, 158-167.

Sarwar, B., G. Karypis, J. Konstan, and J. Riedl (2001), "Item-Based Collaborative Filtering Recommendation Algorithm," *Proceedings of the Tenth International World Wide Web Conference*, 285-295.

Schmid, B. F. (1997), "Requirements for Electronic Markets Architecture," *Electronic Markets*, 7(1), 3-6.

Schubert, P. (2000), "The Pivotal Role of Community Building in Electronic Commerce," *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 1.

Shyu, M. M., C. Haruechaiyasak, and S. C. Chen (2003), "Category Cluster Discovery from Distributed WWW Directories," *Information Sciences*, 155, 181-197.

Slabeva, K. S. and B. F. Schmid (2000), "Internal Electronic Product Catalogs: An Approach Beyond Simple Keywords and Multimedia," *Computer Networks*, 32, 701-715.

Steinfield, C. and S. Klein (1999), "Local vs. Global Issues in Electronic Commerce," *Electronic Markets*, 9(1/2), 1-6.

Yang, J., J. Choi, J. Kim, and H. Ham (2000), "A More Scalable Comparison Shopping Agent," *Engineering of Intelligent Systems*, 766-772.

BargainFinder URL: http://www.bf.cstar.ac.com/bf.

Jango URL: http://www.jango.com

글로벌 판매 및 물류 협업 전자상거래를 위한 유사상품 추천 알고리즘 및 활용

최상현* · 강성민**

요 약

본 연구에서는 판매 및 배송 업무 제휴 기업들을 위한 유사상품 발견 알고리즘을 제시하고자 한다. 제휴 기업들 간에는 제품 분류표를 공유해야 하며 해당 기업들 간 상호 교환 판매가 가능한 제품 정보를 정의해야 한다. 본 연구에서 제시되는 알고리즘의 주요한 아이디어는 제휴 기업들이 보유한 제품들에 대한 동일한 분류 수준에 속해 있는 제품들의 사양 값들에 대해 취합된 효용 가치의 범위를 계산하고 그 제품들의 효용 범위에 근거하여 유사한 상품들을 발견해 내는 것이다. 본 알고리즘의 활용 가능성 검증을 위해 사용자 만족도 실험을 실행하고 그 결과를 제시하였다. 전통적인 유사성 발견 방법론인 거리 척도에 기반을 둔 추천 알고리즘의 결과와 사용자 만족도의 차이를 비교하였다. 본 실험을 통해 제시된 알고리즘이 구매자의 거주지와 수령자의 거주지가 상이하고 구매자가 제품 사양의 중요도에 대해 자신만의 불완전 정보를 보유하고 있을 때 유사상품 추천을 위한 하나의 솔루션으로 활용될 수 있다는 결과를 확인하였다.

한글색인어: 상품 추천, 유사 상품 발견, 협업전자상거래, 제품 분류표, 효용 기반, 다요소 의사결정

* 국립경상대학교 공학연구원 산업시스템공학부
** 중앙대학교 경영학과