Yuichi Washida
Gtraduate School of Commerce and Management, Hitotsubashi University
Innovation Laboratory, Hakuhodo Inc.
1-28-29 Hongo, #201, Bunkyo-Ku, Tokyo, 113-0033, Japan.
kym06001@nifty.com

Hiroaki Tanaka
Department of System Sciences, The University of Tokyo,
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
h.tanaka0528@gmail.com

Kazuhiro Ueda
Interfaculty Initiative in Information Studies, The University of Tokyo
JST CREST
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
ueda@gregorio.c.u-tokyo.ac.jp

Mapping of associations between in-store and online purchases: A proposal for a new product recommendation method

Abstract

Objectives:

Previous research on purchase history has focused solely on in-store or online purchases;there has not been much focus on the association between the two channels. In this study, we used a product mapping method to analyze the relationship between the in-store and online purchases of online shoppers. A new product recommendation method was introduced in light of the product association network created from this data, and was tested for effectiveness.

Methods:

In order to create a product association network, we conducted a consumption diary survey on Yahoo Japan consumer survey panel and calculated the correlations between the purchase frequencies for each possible pair of product categories. The survey targeted ordinary Japanese consumers who frequently shopped online. The survey participants kept a record of all of their consumption activities (including product name, price, purchase channel and place, and situation) on a questionnaire sheet for four days. We received 436 responses. The online purchase history data of those respondents were provided by Yahoo Japan. In all procedures,each respondent's data was strictly masked to maintain privacy and anonymity. We categorized all purchase data into 116 product categories and indicated each as a node on the product association network. We completed the network map by linking all pairs of product nodes which had high correlations with each other.

After creating and analyzing the product association network, we conducted an experiment to persuade consumers who bought online products to buy in-store products that were found to be associated, but also belonged to a different category. For this purpose, product recommendation emails were sent by Yahoo Japan to their online shopping customers. More than one million recommendation emails were sent to the Japanese consumer market in a variety of experimental conditions. click rates and purchase rates were observed or each condition.

Results:

We found that many products purchased online were connected to each other, and that many products purchased in-store were connected to each other. In other words, consumers who tended to buy certain goods online tended to also buy other goods online, whereas consumers who bought goods in-store tended to buy other goods in-store as well, such that there appeared to be a tendency to polarize toward a single purchase channel. However, some items bought online (online products) that were associated to items bought in-store (in-store products) were also identified. Such products could be connecting gateways between the two channels. For instance, in-store snacks was linked to online office supplies. This means that consumers who bought snacks in-store tended to buy office supply products online. If consumers of online office supplies are recommended to buy snacks online, they may change their purchase channel for snacks. This implies a new strategy of online product recommendation that may get consumers across the barrier between the online and in-store markets. The results of the product recommendation experiment suggested that an item pair of medium correlation strength increased the effectiveness of the purchase recommendation.

Conclusions:

The present study introduces a new product recommendation method based on a product association network. Upon comparison of this method with a traditional one used in previous studies, this method appears to have much potential or practical use.

Keywords:

in-store and online shopping, association analysis, item network, diversified recommendation

Introduction and Objectives

In recent years, the e-commerce market has not only increased in scale, but is currently establishing itself as a viable alternative to the traditional storefront (MIC, 2008). However, it is not well understood how consumers differentiate their use between the two channels—that is, would a consumer who buys a specific item online buy other items online, buy them in-store, or not buy them at all? Because the growth of the e-commerce market is supported by the increase in online purchases by each individual consumer (Nomura Research Institute Information Technology Consulting Division, 2010), it is critical to understand how consumers actually shop online. But as far as we know, there has not been previous research that focuses directly on the association between online and in-store purchases. Thus, the first purpose of the present study is to clarify the relationship between internet shoppers' online and in-store purchasing behaviors.

Also of interest are online product recommendation systems that display new products to consumers depending on their preferences and product characteristics. These systems typically utilize online shoppers' activity data (e.g. purchase history, viewing history) to create preference models (Kamishima, 2007), and have been used widely on online shopping sites in recent years. With such a system, however, it seems unlikely that shoppers would be recommended a variety of products from different categories—they would more likely be recommended many similar items. Also, it overlooks preferences about products that online shoppers purchase in-store, and consequently does not include such products in its recommendations. In the event that a consumer is recommended to buy online a product that they typically buy in-store, it would probably be more effective when the relationship between in-store products and online products are strong than when they are weak. As a result, we believe that if the relationship between online and in-store products is clarified, it may be possible to make better online product recommendations. Thus, in addition to determining what various factors affect the strength of the relationship between in-store and online purchases, the present study also seeks to clarify the effectiveness of product recommendation advertisements.

In this study, we analyzed purchase history data and experimental data provided by Yahoo Japan Corporation. Due to a nondisclosure agreement with Yahoo Japan, the actual numbers of purchased goods or the effectiveness of advertising is not released in this paper.

Literature Review

Previous studies have touched upon the relationship between online and in-store purchases. Farag et al. (2007) found that internet search frequency positively increased with online purchase frequency and entering stores for non-routine shopping, and vice versa. In another study, Levin et al. (2003) studied the relationship between purchases and the purchase channel used, and presented the important features of purchase channels in each step of the purchasing process. Purchase behavior analyses have also been performed; for example, a market basket analysis in the field of data mining (Berry & Linoff, 1997) and purchase history data from internet shopping sites have been done in a similar manner. Overall, however, it appears that research focusing on the direct association between online and in-store purchases has not yet been done.

There is also some previous research pertaining to product recommendation systems on e-commerce sites. For example, Zieglar et al. (2005) increased the variety in product recommendations by listing products that were already recommended alongside products with low relatedness (different category). However, overall, research on product recommendation systems does not focus much on recommending items of categories different from purchased items, nor the relationship between online and in-store purchases. By determining the effectiveness of product recommendations based on the relationship between in-store and online purchases, we believe that a product recommendation system that has more variety than the traditional model can be realized without greatly reducing the advertising effectiveness.

Association Analysis

In this study, online shoppers' in-store purchase data and online purchase history data were collected for each individual in a data set provided by Yahoo Japan. By computing the number of in-store purchases and online purchases for each individual and correlating the number of purchases for each product category, we created a product network showing the associations between them.

Method

The following describes the methodology used for the association analysis and the creation of the product association network.

Participants

The in-store purchase data and online purchase history data of 436 online shoppers were collected, such that each individual's data could be distinguished. In order to associate in-store purchase data with online purchase history data, respondents were given IDs for the purpose of this study. As such, in all procedures, each respondent's data was strictly masked to maintain privacy and anonymity.

In this study, it was assumed that analyzing people who were very active online shoppers would provide association data that would be easier to work with than consumers who were not as active. For this reason, participants were limited to consumers who had purchased more than 12 times within the year on Yahoo Japan Auctions and Yahoo Japan Shopping.

Data Collection Methods

In order to collect data about in-store purchases, participants were asked to keep an online consumption record (including product name, price, purchase channel and place, and situation) for a four-day period starting on June 18, 2010 and ending on June 21, 2010. This was done on Yahoo Japan consumer survey panel.

In addition, online purchase history data was collected from participants' Yahoo Japan Auctions and Yahoo Japan Shopping log data from June 22, 2009 to June 21, 2010. The last four days of data were confirmed to be identical to those submitted in the participants' self-logged records.

Product Categorization

By referring to product names in the log data, all of the products within the data were sorted into 58 categories (Table 1). For the purpose of this study, online products and in-store products, even those technically in the same category, were treated as separate items (e.g. online clothes vs. in-store clothes). Thus, a total of 116 product categories were created (online and in-store, 2 x 58).

Product categories were created based on purchase history data and existing categories on Yahoo Japan Auctions and Yahoo Japan Shopping. Large product categories in this study (e.g. home electronics, sporting goods) are also used on other online shopping sites such as Rakuten Ichiba and Amazon.co.jp.

Food-related products, however, were divided into 16 smaller, more specific categories for separate analysis (Table 1, categories 1-16). Because the online supermarket business has expanded greatly in the past few years (METI, 2010), there is much attention on further increasing online sales of food-related products. This kind of categorization attempts to cater to these interests. Also, online shopping for some products can be increased more easily than that for other products; a finer categorization would better differentiate between them. Two more categories of "food bought at convenience store" and "food bought at food store" were added in the first data pretreatment, described in detail in the next section.

Data Pretreatment for Analysis

After calculating the number of purchases in each product category for each consumer in the sample, three pretreatments were performed on the data: first, an addition of two product categories; second, a scaling of the purchase volume to a time period of one month; and third, a scaling of the purchase volume on Yahoo Japan Auctions and Yahoo Japan Shopping in order to estimate the purchase volume at other online shopping sites.

In the first pretreatment, two product categories were added. Convenience stores and stores inside train stations were termed "convenience store," whereas supermarkets, meat shops, and bakeries were termed "food store." Food items bought at such stores were not included in the product categories 1-16 in Table 1, but were instead given new categories: "food item bought at a convenience store" (CVS) or "food item bought at a food store" (SM). Because convenience stores and food stores are where consumers typically purchase food items, consumers were thought to be more resistant to buying such food items online than they would be to buy food items at a department store, for example. Also, because there were far more food items at convenience stores or food stores than at department stores or other stores, there was a concern that data from department stores and other stores would not be reflected in the results. For these reasons, food items were categorized in this way.

The second pretreatment scaled all data to a period of one month. As mentioned earlier, the data collection period for the in-store purchase data was four days, whereas the period for the online purchase history was one year. For this reason, both sets of data were scaled to a period of one month. The in-store purchase quantities for June 18 (Fri) and June 21 (Mon) were multiplied by 10 (20 weekdays in 4 weeks), and the quantities for June 19 (Sat) and June 20 (Sun) were multiplied by 4 (8 weekend days in 4 weeks), and thus the data were scaled to one month (28 days total). The online purchase history data was scaled by first estimating the percentage of the number of items within the year (July 1, 2009 to June 31, 2010) that were bought in a period of one month starting from June 1, 2010. The estimated percentage was 0.087, and data were scaled accordingly.

Finally, in the third pretreatment, the online purchase history data was scaled to estimate purchase data in all internet shopping sites by using an estimate of the share ratio of Yahoo Japan Auctions and Yahoo Japan Shopping. To find this ratio, the 436 participants were asked how frequently they shopped at each site, such that the percentage total equaled 100%. The options were (1) Yahoo Japan Shopping, (2) Yahoo Japan Auctions, (3) Rakuten Ichiba, (4) Amazon, and (5) Other. Options (1) and (2) together were termed the Yahoo Japan share ratio. By multiplying the Yahoo Japan share ratio's reciprocal with purchase quantity, the purchase quantities in all internet shopping sites were estimated. For example, if a participant answered 10% for (1) and 30% for (2), then the Yahoo Japan share ratio would be 40% and all other internet shopping sites would have 60% share. Data pretreatment would then involve multiplying the purchase quantity by 2.5, the reciprocal of 0.4. Because a Yahoo Japan share ratio of close to 0% would cause too much distortion due to the large reciprocal, any participants with Yahoo Japan share ratios below 10% were reassigned 10% as the Yahoo Japan share ratio.

Product Association Network

After the data pretreatments were performed, the correlation coefficient was calculated for all possible pairs of product categories. Using this data and NETDRAW by Analytic Technologies, we created a product association network (Figure 1). Nodes represent product categories as labeled, and lines purchase association links. The node color represents the 436 participants' purchase ratio (PR) in each category, standardized based on the daily necessities category, the largest in-store category after pretreatment.

Links below a threshold of $r = 0.20$ were not displayed; 0.20 was deemed to be sufficient to observe clear relationships between products with less than 200 links. Also, the lowest threshold to detect significance at a $p = 0.01$ level with $n = 436$ was $r = 0.124$, so the threshold of $r = 0.20$ allows for significance. The categories lined up on the left side of Figure 1 are the categories that did not have links with any other category, according to the threshold.

Findings

The product association network (Figure 1) has the following main features: an upper-right cluster of in-store products (in-store product cluster), a bottom-left cluster of online products (online product cluster), a bridge linking the in-store product cluster and the online product cluster, and finally a cluster independent from both the in-store and online product clusters. The in-store product cluster formed mainly around food (in-store) and clothes (in-store), and connected to some home and lifestyle products (online). On the other hand, the online product cluster formed mainly around food (online), home and lifestyle products (online), and clothes (online), with some connections with electronics (online, in-store). There appeared to be no direct connection between the in-store product cluster and the online product cluster. In electronics (online, in-store), pairs that combined the in-store and online channels were seen; for example, AV equipment (online) and home electronics (in-store) were a directly linked pair, as well as computers (online) and computer peripherals (in-store). The independent cluster consisted of hobby-related categories such antiques, comics, CDs and music software, books and magazines, hobby, video games, and undergarments.

Discussion

First, consistent with common sense, basic needs such as food, clothes, and home and lifestyle products were correlated, regardless of whether they were purchased in-store or online. Also, product categories such as antiques and hobby were isolated from other online or in-store products; it appears that certain consumers focus their consumption on these products.

In terms of new findings, it appeared that consumers tended to polarize toward a single channel. According to previous statistical data (METI, 2010; MIC, 2008), ordinary consumers still buy most items in-store, but frequent online shoppers like the participants in the present study buy a portion of products online. Thus, it would be reasonable for products that are bought increasingly more often online to be correlated with in-store products, hence using both purchase channels. However, the results showed that consumers who bought products online (especially food and clothes) would also buy other products online, whereas consumers who bought them in-store also bought other products in-store. In this sense, there appeared to be a polarization into two types of online shoppers. Although online purchases seem to have increased according to previous statistics, the results of the present study suggest that the spread of e-commerce is not progressing as much as previously thought.

In addition, because food is bought in large quantities and because its sales online have not increased much, we would expect further development of the online food market. Yet, according to the results, it is likely that the traditional product recommendation system does not recommend food products to internet shoppers who actively buy food items in-store.

Consequently, in order to recommend similar items online to internet shoppers who buy such items in-store, it may be necessary to implement a system that recommends products beyond category barriers. Thus, by recommending food (online) to consumers who buy items such as office supplies (online), it can be expected that the sales volume of food (online) would increase more efficiently because of the high correlation between food (in-store) and office supplies (online).

Limitations

There are some precautions concerning the data analysis. Only information that was thought to be necessary for analysis was collected from Yahoo Japan's purchase history data. In addition, participants were selected at random from the internet shopping sites on the condition that, not only were they frequent shoppers, but also that they had active communication, among other conditions. Some particularly active participants were chosen arbitrarily for the purpose of this study as well. Also, the purchase history data does not represent all of Yahoo Japan Auction or Yahoo Japan Shopping user trends. Thus, there is some precaution of overgeneralization.

Product Recommendation Experiment

In this experiment, advertisement emails were sent to consumers (recommendation recipients) who bought a certain category of items on an online shopping site. The email recommended them to buy online a product typically bought in-store (recommended item) that was correlated with the purchased item's category. A test of statistical significance and a qualitative comparison was performed on the relationship between the effectiveness of the advertisement and various factors, such as the correlation strength between the recommended item and the purchased item.

Method

In this experiment, email advertisements were sent to consumers who bought a specific category of products on Yahoo Japan Auctions or Yahoo Japan Shopping. These emails recommended an in-store product that was correlated to the purchased online product to be bought on Yahoo Japan Shopping. The emails were sent on December 21, 2010, and the advertisement effectiveness was measured on December, 27, 2010. The effectiveness of the advertisement was measured in two ways: (1) the ratio of consumers who clicked on the product link after receiving the email (click rate) and (2) the ratio of consumers who bought something on Yahoo Japan Shopping within 1 week of the email being sent (purchase rate).

Two Scenarios

There are two scenarios in which an in-store product could be recommended to be bought online after a consumer buys an online product from a specified category: (1) the same item can be recommended to buyers who purchased from different product categories (Figure 2a) and (2) buyers of a single product category can be recommended different kinds of items (Figure 2b). As such, the first scenario (Scenario A) attempts to increase the online sales of a specific in-store product. It investigates which online product categories would have suitable recommendation recipients for a specified in-store product category (1junk) to be purchased online more often. The second scenario (Scenario B) attempts to broaden existing product genres without targeting a specific product; instead, it targeted recommendation recipients who bought from a single product category (2lady) and sent product recommendations from two different categories (1deli, 2toilet). It tests which product categories would be effective to recommend for online product categories that already have a large market. Separate experiments were conducted in each scenario. The structure of the two scenarios is shown in Table 2, alongside the two factors and specific recipient-item pairings.

Two Factors: Market Size and Correlation

Market size and correlation were studied as factors that may affect the effectiveness of the recommendation advertisement (Table 2). Market size refers to the annual volume of online products bought. Two market size categories—“large” and “small”—were created. The average quantity was taken by summing the quantities in all of the product categories in the online purchase history data for the 436 participants and dividing it by 58, the number of online product categories. Products with quantities higher than the average were labeled as having a “large” market size; those lower than average were labeled “small.” Correlations were labeled either “strong” or “medium.” The basis of comparison was the value of $r = 0.094$, which was barely significant with a sample of 436 participants.

Pairing Recommendation Recipients and Recommended Items

The product categories that the recommendation recipients bought from and the recommended products are paired in Table 2. Snacks (1junk) was chosen as the recommendation item for Scenario A; consumers of women's clothing (2lady) were targeted to receive a recommendation in Scenario B. 1junk connected in-store products with online products (Figure 1), and 2lady had an especially large market size.

Method of Analysis

The structure of analysis is shown in Table 3. For each Scenario, a chi-square test was performed to determine the effects of each factor on the click rate and the purchase rate. The effects were also compared qualitatively. When performing the chi-square tests on the click rate in Scenario A and the purchase rate in Scenarios A and B, an angular transformation method (Iwahara, 1965) was used because emails were sent to different groups. Also, in order to test for a difference in purchase rates when there was an email advertisement as opposed to when there was not, a control group that did not receive an email advertisement was created. Significance was evaluated at $\alpha = 0.05$.

Comparing the New Product Recommendation Method with a Traditional Method

We compared the product recommendation method used in the present study with a traditional recommendation method, in terms of advertising effectiveness measured by click rates and purchase rates. Tables 4 and 5 compare the click and purchase rates respectively in the present study and two previous studies on Yahoo Japan. In Previous Study 1, items similar to those that online shoppers viewed were recommended, whereas in Previous Study 2, items were recommended without targeting specific consumers.

Because the conditions for calculating for click rate and purchase rate were different in the two studies, the results of the present study were recalculated according to the conditions used in the previous Yahoo Japan studies. For this reason, there are some effects for which the groups in each scenario cannot be compared. The values for click rate (total) and purchase rate (LC) in the two tables were normalized to the Yahoo Japan conditions, such that the Yahoo Japan value was 1.00. The click rate (total) is the percentage of the people who clicked on the email advertisement out of all the people who received the advertisement. The purchase rate (LC) refers to the percentage of people who purchased the recommended product right after clicking on the email advertisement.

Findings

The click rates for each group are shown in Table 4, while the purchase rates are shown in Table 5. Due to the nondisclosure agreement with Yahoo Japan, the actual values of the chi-square and the sum of squares results are not listed. The click rates were normalized such that the medium correlation group in Scenario B had a click rate of 1.00. The purchase rates were normalized such that the group in Scenario B with medium correlation and advertisement emails sent had a purchase rate value of 1.00.

Click Rates and Purchase Rates

In Scenario A, both the correlation coefficient and the market size had a significant effect on click rate ($p < 0.01$). The effect of correlation coefficient had a $p < 0.01$, the effect of market size a $p < 0.01$, and the interaction between these two factors a $p < 0.01$. A medium correlation had a higher click rate for both large market size ($p < 0.01$) and small market size groups ($p < 0.05$). For the groups with strong correlations, the group with the small market size had a greater click rate ($p < 0.01$). However, for the medium correlation groups, market size did not have a significant effect on click rate ($p > 0.05$). In Scenario B, the correlation strength had a significant effect on click rate ($p < 0.01$).

In Scenario A, only the strength of correlation had a significant effect on purchase rate ($p < 0.01$) according to a chi-square test. The email advertisement did not have a significant effect, and there was no significant interaction between the strength of correlation and the email advertisement. In Scenario B, neither correlation strength nor the email advertisement had a significant effect on purchase rate according to a chi-square test; there was also no significant interaction between these two factors.

Overall, it was found that a medium correlation had a higher click rate than a strong correlation, and that a large market size had a higher click rate than a small market size. Also, in both Scenarios A and B, there was no difference in purchase rate between the group that received the email advertisement and the group that did not.

Comparing the New Product Recommendation Method with a Traditional Method

According to the click rate and click rate (total) in Table 4, the overall click rate is lower for the present study's new recommendation method than for the previous studies' recommendation of items in similar categories. However, when comparing only those groups that were recommended items with medium correlation, the new recommendation method was about 70% to 90% as effective as the recommendation method used in the previous study; thus, the click rate for these groups was not greatly reduced. Also, the purchase rates and the purchase rate (LC) for the new recommendation method are not lower than those for the recommendation method used in previous study 2. This suggests that the new recommendation method is not inferior to the older recommendation method in increasing the purchase rate via email advertisements.

Discussion

First, we consider why the click rate may have been higher for the medium correlation group. In the present study, participants were limited to online shoppers who were particularly active in their online shopping. The fact that online products and in-store products have a high correlation suggests that, even though in-store products are purchased online increasingly more often, they are still being bought in-store. The groups with medium correlations, on the other hand, may switch purchase channels more easily, and may have had a higher click rate for that reason. On the other hand, the explanation for Scenario A's high click rate for the small market size group may be that the definition of "small" market size in the present study was for products thought to be bought less frequently online. Online shoppers who buy products with a small market size may buy more things online overall compared to ordinary online shoppers; for this reason, the click rate for the snacks category, one thought to be bought less often online, may have been greater than usual in this experiment. The email advertisements may have had little effect on the purchase rate because there are other factors involved in leading up to the ultimate purchase of a product. Many conditions, such as consumer preferences and the willingness of the consumer to purchase at that time, need to be met for the purchase to take place; thus, the effect of the advertisement by itself may have been difficult to observe.

In regards to the new product recommendation method, it can be concluded that the new recommendation method still has potential, even though the email advertisements did not significantly increase the purchase rate of consumers in the present study. Considering how consumers become interested in recommended items, endeavors like this over a longer period may increase consumer interest in products that they usually do not buy online. If this is the case, then the new recommendation method may be effective in increasing online consumption. Therefore, it is suggested that a similar study be done over the long term, rather than over one week as in the present study, in order to understand the trends more fully.

Limitations

We did not focus on the actual purchase activities, while we focused on one of protocol of the actual purchase activities by observing the click rates. The actual purchase tends to be strongly influenced by a price factor besides recommendations. In this study we did not obtain the actual price in the market, therefore, we did not focus on the actual purchase.

Conclusion and Managerial Implications

In the present study's purchase association analysis, a product network correlating in-store products with online products was created using individual in-store purchase data and online purchase history data. According to the results, two opposing trends were discovered. Online shoppers who bought certain items online, especially food and clothes, also bought other items through the online channel; those who bought them in-store also bought other items in-store. In addition, although previous statistical data has suggested that the e-commerce market size has been increasing smoothly, the present study suggests that the association between the two channels is unexpectedly small, even amongst consumers who shop online more frequently than the average consumer. The present study also identified specific categories of in-store products that were correlated with online products; product recommendations with these combinations in mind may potentially increase online purchases.

Based on these results, an experiment was conducted in which email advertisements were sent; online shoppers were recommended to buy in-store products online that were correlated to previous online purchases, but were also in a different product category. The results suggested that, in order to increase interest in products that consumers usually do not buy online, recommending in-store products that have about a medium correlation with the purchased online products would be effective. Recommending products to buyers of online product categories with a smaller market size was also found to be more effective. Thus, these two findings may act as guidelines for effective product recommendation by marketing practitioners based on the association between in-store and online products. Furthermore, upon comparing the new recommendation method to an older method that recommended items in the same category as those that consumers viewed, it was found that the new method was not inferior to the older method. Rather, it may be an effective method in expanding consumer interest in purchasing products online. Marketing practitioners can design a new marketing mix strategy by combining two different product categories.

Future Research

Several suggestions can be made in terms of future directions. First, in the association analysis, only four days of in-store purchase data was collected. However, only four days of data makes it difficult to evaluate the purchase of products such as home electronics and computers; thus, it would be desirable to extend the period of data collection as much as possible.

Also, the type of products that online shoppers buy on Yahoo Japan Auction and Yahoo Japan Shopping may be different from those bought on Rakuten Ichiba, Amazon, or other online shopping sites. Thus, it is suggested that further studies analyze online purchase history data from sites other than Yahoo Japan Auctions or Yahoo Japan Shopping.

Finally, further studies should investigate how long and how many advertisements it would take, using the product recommendation method introduced in this study, for online shoppers to buy products that they usually would not buy online.

References

Berry, M. J., Linoff, G. (1997), *Data Mining Techniques: For Marketing, Sales, and Customer Support*, John Wiley & Sons, New York.

Farag, S., Schwanen, T., Dijst, M., Faber, J. (2007), "Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping", *Transportation Research*, Vol. 41, No. 2, pp. 125-141.

Iwahara, Shinkurō (1965), *Estimation techniques for education and psychology: new edition* (in Japanese), Nihon Bunka Kagakusha Co., Tokyo.

Kamishima, Toshihiro (2007), "Algorithms of recommendation systems" (in Japanese), *Transactions of the Japanese Society for Artificial Intelligence*, Vol. 22, No. 6, pp. 826-837.

Levin, A. M., Levin, I. P., Heath, C. E. (2003), "Product category dependent consumer preferences for online and offline shopping features and their influence on multi-channel retail alliances", *Journal of Electronic Commerce Research*, Vol. 4, No. 3, pp. 85-93.

Ministry of Economy, Trade and Industry of Japan (2010), "Report for 2010 electronic commerce research" (in Japanese), available at: http://www.meti.go.jp/policy/it_policy/statistics/outlook/h21houkoku.pdf (accessed 21 June 2011).

Ministry of Internal Affairs and Communications of Japan (2008), "2008 white paper for telecommunications" (in Japanese), available at: http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h20/pdf/20honpen.pdf (accessed 21 June 2011).

Nomura Research Institute Information Technology Consulting Division (2010), *What will happen in the information/telecommunications market from here onwards—IT market navigator 2010* (in Japanese), Toyo Keizai, Inc., Tokyo.

Zieglar, C.N., McNee, S. M., Konstan, J. A., Lausen, G. (2005), “Improving Recommendation Lists Through Topic Diversification”, in *proc. of WWW’05*, pp. 22-32.

Appendix

Table 1. Categories of items and corresponding labels

No	Product Category	Label	No	Product Category	Label
1	Meals	eat	29	Women's clothing	lady
2	Prepared food	deli	30	Men's clothing	men
3	Staple food	rice	31	Inner wear	inner
4	Ingredients	ham	32	Watches, Accessories	acce
5	Fresh food	fresh	33	Shoes, Footwear	shoe
6	Instant food	insta	34	Bags	bag
7	Luxury goods	favo	35	Baby supplies, children's clothes	baby
8	Flavorings	oil	36	Home audio, Home video	av
9	Desserts	sweet	37	Home appliances	ele
10	Snacks	junk	38	Mobile phones	mobi
11	Soft drinks	dri	39	Computers	pc
12	Water	wat	40	Computer peripherals	pc+
13	Milk	milk	41	Office supplies	supp
14	Alcohol	beer	42	Sporting goods	sport
15	Local products	local	43	Luggage and travel supplies	trav
16	Take-out food, food deliveries	pac	44	Bicycles, Bicycle supplies	bicy
17	Cosmetics	cosme	45	Automobiles, Motorcycles, Car and motorcycle goods	car
18	Daily necessities	daily	46	Video games, Gaming software	game
19	Toiletries	toilet	47	Toys	toy
20	Household goods	life	48	Hobby, Culture	hobb
21	Gardening supplies	flow	49	Antiques, Collections	anti
22	Office supplies, stationary	offi	50	Newspapers	news
23	Home improvement tools, DIY kits	diy	51	Books, Magazines	book
24	Dietary supplements, Health food	diet	52	Comics	comi
25	Medicine	medi	53	CDs, Music software	cd
26	Glasses, Contacts, Eye care products	glas	54	DVD, Movie software	dvd
27	Tobacco	smok	55	Digital contents, Applications, Computer software	soft
28	Furniture	funi	56	Pets, Pet supplies	pet
			57	Tickets, Gift certificates	tick
			58	Other	ex
				Food bought at convenience store	CVS
				Food bought at food store	SM

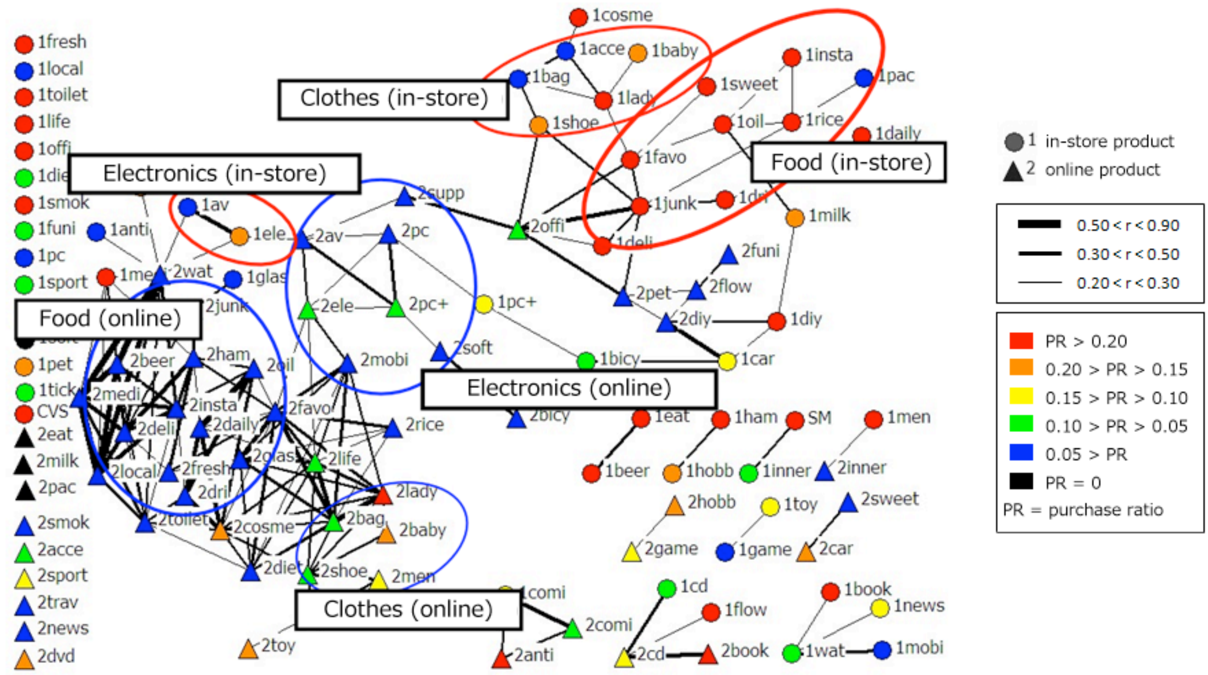
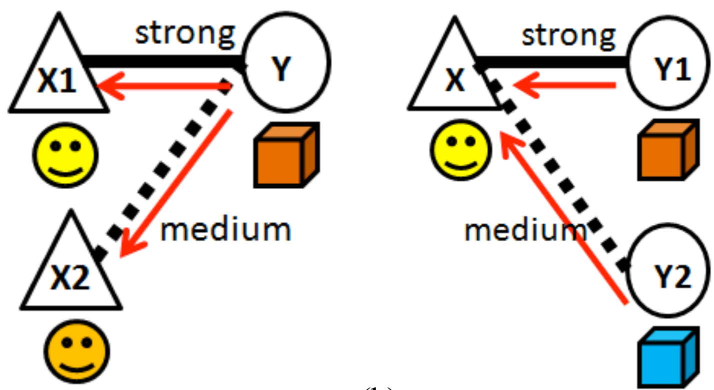


Figure 1. **A product association network**



(a) (b)

Figure 2. **Scenario A (a) and Scenario B (b)**

Table 2. Chart pairing recommendation recipients and recommended items

No	Scenario	Recommendation recipient	Recommended item	Market size	Correlation
1	A	2pet	1junk	small	strong
2		2offi		small	strong
3		2supp		small	strong
4		2pc+		large	strong
5		2trav		small	strong
6		2ele		large	medium
7		2life		small	medium
8		2funi		small	medium
9		2acce		large	medium
10	B	2lady	1deli	large	strong
11			1toilet	large	medium

Table 3. Factors affecting recommendation effectiveness

Scenario	Factor	Click rate	Purchase rate
A	1	correlation	correlation
	2	market size	email
B	1	correlation	correlation
	2		email

Table 4. Click rate comparison between the new recommendation method and a traditionally used method

Scenario	A				B		Previous Study 1
Correlation	strong		medium		strong	Medium	
Market size	large	small	large	small			
Click rate	0.55	0.82	0.84	0.94	0.48	**1.00**	
Click rate (total)	0.78				0.41	0.92	**1.00**

Table 5. Purchase rate comparison between the new recommendation method and a traditionally used method

Scenario	A				B				Previous Study 2
Correlation	strong		medium		strong		medium		
Email	sent	not sent	sent	not sent	sent	not sent	sent	not sent	
Purchase rate	1.23	1.23	1.18	1.17	1.04	1.02	**1.00**	1.00	
Purchase rate (LC)	1.43								**1.00**