# A Method to Implement Behavior Targeting System Using 3D- Vectors by Customer Preference

*Masayuki Kessoku,*[†] , *Jun-ichi Aoe*[†] *and* *Kazuhiko Tsuda*[††] ,

[†]The University of Tokushima, Tokushima, 7708506, Japan, [††]The University of Tsukuba, Tokyo, 1120012, Japan

## Summary

This paper proposes a method to construct Behavior Targeting, which is an effective advertising method for Web Publishers. One of the challenges on constructing Behavior Targeting is the heavy load on the hardware when extracting customer preference information due to the enormous number of customers and the extremely broad areas that the customers might be interested in. We address this problem with a distributed processing system that extracts customer preference information at the light loaded Access Point Servers, not at the central Marketing Server. To accomplish this, we have constructed a knowledge dictionary that can be dynamically updated and used for extraction of customer preference information at the Access Point Servers. As a result, we have succeeded to reduce the load on the Marketing Server by more than 60%. Another challenge is how to grasp not only the customer preferences but also the trends of the preferences and the markets speedy and accurately. In conventional methods, customer preference and market information are managed with two-dimensional vectors of customer and preference category axes. In this proposed method, we add time axis to make it three-dimensional vectors in order to manage the preference transitions and the market trends. One of the problems of three-dimensional vectors is its huge volume of information. Our method addresses this problem by recording the positions and the values of the points only where the information has changed as time passes. As a result of experiments, we have succeeded to detect groups of customers who show similar preference transitions from the information stored in three-dimensional vectors. Furthermore, we have found a few trend leaders in the groups, which confirm that there is a possibility to make appropriate recommendations to the other group member based on the transitions of the trend leaders' preferences.

***Key words:***
*Customer Preference, Advertisement, Behavior Targeting, Recommendation, 3Dimension.*

## 1. Introduction

In recent years, the needs of customers are more and more diversified. One of the reasons is that the progress of information tools, typified by W.W.W., has enabled customers to obtain much more information than before. The customers with more information tend to purchase goods based on their preferences, rather than coveting for the goods that the people around them posses.

Today is referred as an age of individuality; however it is not necessarily true that everything is diversified. For example, the apparel market has a major trend called "fashion" and the customers tend to express their individualities by selecting only one item that is different from other people.

Information systems are recognized as effective tools to collect preferences of entire market and individual customer [1]. In addition to gathering information, it is reported that the productivity of sales and marketing has been dramatically improved by utilizing Sales Force Automation (SFA) [2], which is a method to analyze sales activities through marketing activities.

The marketing in the Internet businesses requires the way to quickly grasp market trends and preference transitions of individual customer, utilizing interactivity and immediacy of the Internet. It is also required to grasp the effects of advertisements accurately and to propose the effective timing and contents of advertisements to advertisers [3-8].

In these backgrounds, Web Publisher offers a service; such as advertisement and recommendation called "Behavior Targeting" to provide information that may interest a specific customer [9-11]. To improve the accuracy of the Behavior Targeting, they need to extract information such as the market trends and the preference transitions of individual customer, and they also need to manage it accurately. For the purpose of obtaining information about both market trends and customer preference at once, one of the effective methods is to manage customer preference information using two-dimensional vectors with customer and preference category axes. This method is used in many of Customer Relationship Management (CRM) [2] [12] tools. However,

Manuscript received 26th, September, 2005.
Manuscript revised 24th, January, 2006.

this two-dimensional method is not capable of keeping track of market trends and the transitions of customer preferences that change over time.

In this paper, we propose a method that uses three-dimensional vectors with time axis in additions to customer and preference category axes. One of the problems of the three-dimensional method is that the volume of information held in three-dimensional space is enormous so that it is difficult to hold detailed information. The enormous volume of information may also slow down the retrievals of information. Our proposed method addresses these problems by recording the time axis information only on the points where the customer preference changes, while data on customer and preference category axes, which express the current situation of customer preference, is stored in two-dimensional vector space.

Additionally, the system is required 1) to collect the data in real time basis and 2) to analyze the collected data in order to precisely grasp the changing customer preferences and market trends. For this reason, the server to manage three-dimensional vector information requires very high processing power and this makes the system more expensive.

The system proposed in this paper reduces the load on server side by executing data management and keyword extraction processes on access point side. Moreover, this configuration enables real time collection of customer preference data. This system executes keyword-learning process on server side using the room created by reducing loads on the server. The keywords learned on server are reflected on the keyword extraction process on access point side.

## 2. Outline of Behavior Targeting Generation System

This chapter illustrates the outline of the Behavior Targeting generation system we propose. The Behavior Targeting service offers a page of summarized information that may interest a specific customer. It naturally offers different information according to the individual customer.

Figure 1 shows the outline of Behavior Targeting generation system. The Template-page is a page used as a template for Behavior Targeting generation. The Template-page contains the information to be provided to all customers, just like regular Web pages, and it also contains empty slots to provide information specific to the customer. Each slot has the tag information to indicate which information should be embedded, such as "Weather", "Advertisement", and "Information" slots.

The Slot Database stores the data to be embedded into the slots in Template-page. The Behavior Targeting generation module generates a Behavior Targeting for a specific customer by retrieving information that may interest the customer from the Slot Database and by embedding it on the Template-page.

For example, if a customer is a golfer who lives in Osaka, the Weather slot may show the weather information of Osaka region, the Information slot may show the result of the Masters, and the Advertisement slot may show the ads for golf shoes and a packaged golf tour.

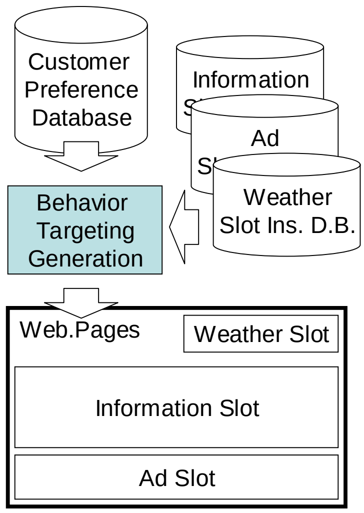



Fig. 1 Outline of Behavior Targeting generation system.

## 3. Collection Method for Preference Information

The most important point to generate a Behavior Targeting that provides useful information for a specific customer is how accurately the customer preferences can be grasped.

When a customer uses W.W.W., he/she connects to the Internet via an access point of an ISP, which means that the Internet usage information of a customer can be collected at the access point. In our proposed method, we allocate Data Collection Modules to each access point as shown in Figure 2. The Data Collection Modules collect preference information at the access points and send it to the Marketing Server. The Marketing Server receives the information and manages it in the Customer Preference Database in one lump.

### 3.1 Collection of Customer Preference Information at Access Points

A Data Collection module at an access point analyzes the Web pages accessed by a customer and extracts the keywords from them. The extracted keywords are sent to Data Management Module. These processes are executed in real time whenever a customer accesses a Web page. Analyzing extracted keywords enables us to determine the preference category of the accessed Web pages and to specify which category the customer is interested in.

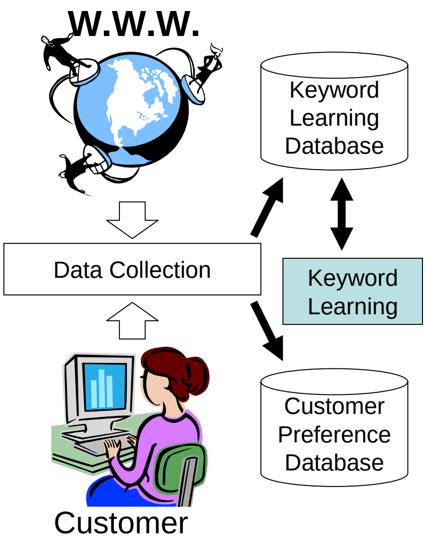



Fig. 2 Outline of Preference Information Collection System

Our system extracts a compound word as well as a simple word as a keyword. For example, a compound word "Steven Spielberg" is an important keyword for a preference category "Movie-SF". To extract this kind of compound words, we store the keywords to be extracted in the Keyword Databases. The extracted keywords are sent to the Data Register Module in the Marketing Server.

## 3.2 Management of Customer Preference data for Marketing Server

Marketing Server consists of the following five blocks:

1. Customer Preference Database
2. Keyword Learning Database
3. Data Register Module
4. Data Update Module
5. Keyword Analysis Module

The Customer Preference Database is composed of Customer Preference vectors, which are the two-dimensional vectors of customer and category axes as shown in Figure 3. Superposing these Customer Preference vectors as illustrated in Figure 4 implements the three-dimensional vectors with customer, category, and time axes.

Each field of Customer Preference vectors contains the score that indicates the strength of preference of a specific customer for a specific category. For example, in Figure 3, the customer A scores “6” on preferences for the category of Rugby, and the customer B scores “4” on preference for the category of tennis.

Keyword Learning Database stores the keywords that indicate various preference categories. The data is stored as a set of the keyword strings, the preference category that the keyword is related, and the score to indicate the strength of preference. For example, data should look like below:

((ball: handball, 4: tennis, 2).

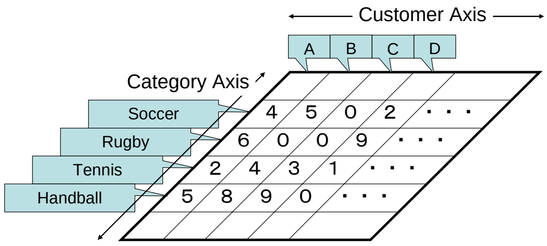



Fig. 3 Composition of Customer Preference vector

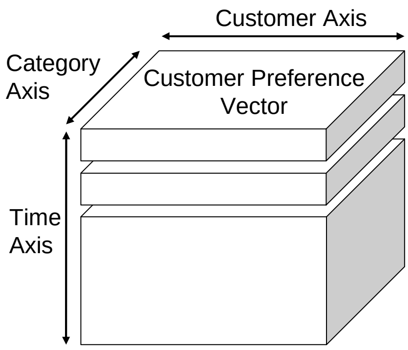



Fig. 4 Conceptual Diagram for Composition of Customer Preference Database

This data indicates that the keyword "ball" scores 4 for the preference category of "football" and scores 2 for the preference category of "tennis". As described above, one keyword can be related to two or more preference categories.

The Data Register Module receives keywords from the Data Collection Module and calculates the score of the Web page accessed by the customer using keywords and scores stored in the Keyword Learning Database. The calculated score is added to the score of the customer stored in the Customer Preference Database. Repeating these processes stores data in the Customer Preference Database that indicates which customer accesses the Web pages of which preference category.

### 3.3 Learning Process in Keyword Learning Database

The information in the Keyword Learning Database is updated by the Keyword Analysis Module in the Marketing Server. Then the Data Update Module reflects the updated information into Keyword Databases at access points. This section describes these processes.

A Data Collection Module at an access point performs morphological analysis of the Web pages accessed by a customer. It extracts keywords (NamedEntity, onomastic words, onomastic word strings, and word strings stored in the Keyword Database) and sends them to the Data Register Module in the Marketing Server as a set of keywords.

When the word strings in the received set of keywords has been registered with the Keyword Learning Database, the Data Register Module retrieves the preference scores from the database and appends them to the word strings. Then the Data Register Module calculates the scores of the word string, registers them with the Customer Preference Database, and sends them to the Keyword Analysis Module.

The Keyword Analysis Module receives the set of keywords and calculates the customer preference score for each category. When the set of keywords contains the word strings that are not registered with the Keyword Learning Database, the Keyword Analysis Module appends the calculated scores to them and registers them with the database.

The Data Update Module updates the Keyword Databases at the access points by registering the word strings data that are newly registered with the Keyword Learning Database. This process should synchronize the Keyword Learning Database and the Keyword Databases

The Keyword Databases is used to retrieve the word strings that can be keywords. This is the same logic as to extract multiple keywords from a text string. To speed up this process, we implement the Keyword Databases based on the structure of String Pattern Matching Machine for multiple keywords, which is proposed by Aho et al [13]. For this reason, it is not easy to update the word strings that are already registered with the Keyword Databases. We address this problem using the algorithm to append keywords to Machine-AC, which is proposed by Tsuda et al [14].

The described keyword learning process enables us to always use the updated Keyword Database for keyword extraction at the access points, which consequently achieves the distributed keyword extraction processes. The whole process of grasping customer preferences is composed of three sub-processes: extracting keywords, retrieving preference scores from the Keyword Learning Database, and registering data with the Customer Preference Database. Among these processes, the keyword extraction requires the largest CPU power, which is 75% of the total. Our proposed method distributes the keyword extraction process to each access points. This requires an additional process to update Keyword Databases at the access points, but it still reduces 60% of CPU usage on the Marketing Server. As a result, we are able to reduce the installation cost of hardware required for Behavior Targeting services by half at the estimation level.

## 4. Necessity of Three-Dimensional Vectors

The proposed method implements the Customer Preference Database in three-dimensional vector structure. The worst weakness of three-dimensional vector method is its enormous volume of data. This chapter explains the necessity of three-dimensional vector method even though it has the weakness on data volume.

In the progress of IT in recent years, data mining [4-5] is one of the remarkable methods to analyze customer information. For example, data mining tools are used for storing in-house data, analyzing customer information, and working out customer-oriented business strategies.

Data mining includes various methods such as market basket analysis, memory based reasoning, cluster analysis, decision tree, and neural network. By using the clustering method to find similar records, you obtain the initial data to start analyses of customer information to grasp market trends and customer preferences. In other words, the valuable information extracted from lump of data by in-house business analysts can be the starting point for planning business strategies.

The Behavior Targeting service is exactly personalize marketing. It will be a very effective marketing tool for a web publisher if it provides the information that precisely interests a specific customer. However, it can get a bad reputation if it provides meaningless information for the customer. Therefore, the technology to accurately grasp customer preferences is very important [19-22] for providing the effective Behavior Targeting service.

Two-dimensional vector method with customer and preference category axes is one of the effective methods to grasp market trends and customer preferences from analyses of customer information, and is actually used in many marketing tools. This is because the vector space in this method is pure image of markets, and each customer axis represents an individual customer, so that the preferences of individual customer can be kept track by looking at a specific customer axis along with its preference category axis.

However, this method only allows managing accumulated strength of preferences in preference category axis, and it is difficult to keep track of changes in market trends and in customer preferences over time. Therefore, if a preference category aroused customers' interest much in the past but does not arouse much in lately, the preference category still indicate high score. This means that even if a preference category rapidly arouses customers' interest, it scores low at the initial stage and it is difficult to detect it. However, the highest buying inclination is indicated in the very early stage of preference categories, which means that the changes of customer preference must be detected accurately at its early stage to conduct effective sales promotions activities.

Some of the management methods in two-dimensional vector space consider the passage of time. In one of these methods, the value of each field element is decremented by multiplying it in every specific period by a specific modulus of less than 1. These results a high scored preference category in the past to be less scored if the category is losing interests from customers [4-5]. This method considers the preference categories in the past; however it is impossible to keep track of the transitions of preference categories. To keep track of the changes of customer preferences accurately, the preference information must be managed in three-dimensional vector structure with the additional time axis.

## 5. Configuration of Customer Preference Database

The three-dimensional vector method is an effective method to manage transition information of market trends and customer preferences that change over time. However, we need to address the problem of enormous data volume in order to apply it for data mining. In this chapter, we propose a method to compress the Customer Preference Database that is composed of three-dimensional vectors.

### 5.1 Concept of Managing Customer Preference Database

Figure 5 illustrates the configuration of the Customer Preference Database and the outline of storing method for customer preferences. This system is to manage customer preference information V to keep track of the transitions of customer preference categories and market trends that change over time.

Firstly, we calculate Sc = {x1, $x_2$,}, a set of preference strength score x, at the preference strength score calculation module to keep track of the observed behavior of customer c as the customer's preference. The customer c is managed and maintained in a customer management table that has a list C of customers or customer groups as a database. Each element x of the set Sc is the preference strength score that corresponds to the preference categories management table, which maintains a list of preference categories of customers.

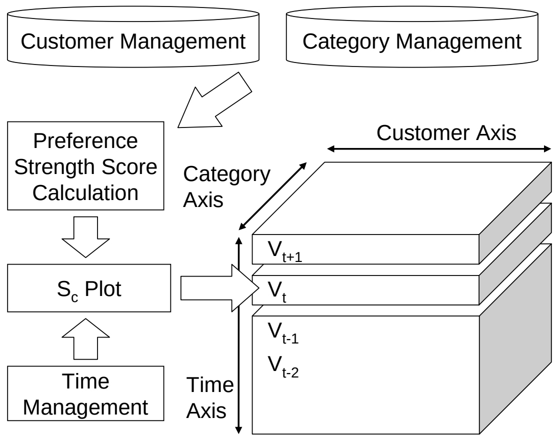



Fig. 5 Configuration of Customer Preference Database

Next, at the preference category plot module, we plot customer preference information Vc = (c, Sc, t), where the customer c behaves at the time t with the preference strength scores Sc, on to the preference field which composed of customer, preference category, and time axes. The customer axis and the preference category axis correspond to the customer management table and the preference category management table respectively, and the time t is plotted on the time axis that is a series of unit time managed at the time management module.

### 5.2 Compression Method for Vector Space

The volume of information will be enormous in the Customer Preference Database when it is managed in three-dimensional vectors with customer, preference category, and time axes. This proposed system separately manages time axis from the two-dimensional preference field of customer and preference category axes (see Figure 6). On the two-dimensional preference field, the customer preference information is expressed with preference vectors. For example, the current customer preference information is represented as a current preference vector $V_t$ = (C, $S_t$, t), and the customer preference information at one unit time past is represented as a past preference vector $V_{t-1}$ = (C, $S_{t-1}$, t-1).

Firstly, the preference transition extraction module compares a current preference vector $V_t$ and a past preference vector $V_{t-1}$, and extracts the preference fields with different preference strength. Then the preference transition conversion module converts the extracted information into a coordinate point (c, s) on the customer and the preference category axes and a set preference strength score $x_{t-1}$ of the past preference vector, and store them into the preference transition storage module. In this way, the past preference vectors are stored in smaller volume.

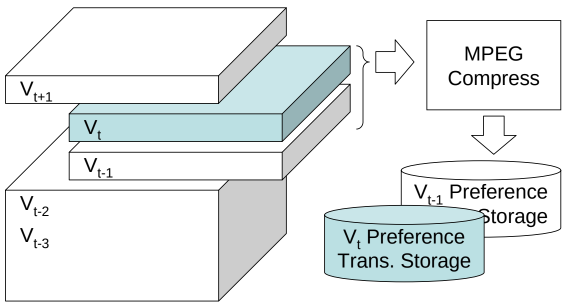



Fig. 6 Storing Method for Preference Field per Unit Time

When a unit time managed in time management module passes, the current preference vector $V_t$ in the one unit time past becomes the past preference vector. Similarly, a new customer preference strength score calculated based on the customer's behavior at the past unit time becomes a current preference vector $V_{t+1}$ = (C, $S_{t+1}$, t+1).

By repeating this procedure, the past preference vectors at each unit time are stored in preference transition module. Consequently, preference information of unit time is stored as sets of current preference vectors and preference transitions in the past. This method enables us to store information for the Customer Preference Database in small volume and without omission.

# 6. Evaluation of Customer Preference Database

This chapter describes the evaluation result of Customer Preference Database constructed in three-dimensional vectors. The first section explains the methods of the experiment, the second section shows the evaluation of data compressibility rate in three-dimensional vectors, and then third section describes the analytical results of customer preference information.

## 6.1 Method of Experiment

We have conducted a small size experiment by applying the proposed system to Behavior Targeting generation of a Web Publisher. This section describes the outline of the experiment and evaluates the proposed system on the experiment.
The experiment is conducted in the following conditions:

- Number of customers: 99 (employees and part-timers of the Web Publisher)
- Number of categories: 538 (amended based on the contents categories defined by contents ID forum [19])
- Time period: 10 weeks (one week is considered as one cycle)

The Data Collection Modules at access points extract keywords from the text data of Web pages accessed by the customers. The extraction is performed only on the <body> part of the HTML files and emphases with tags are not considered.

A Data Collection Module firstly performs morphological analysis to divide the text data into words and to obtain the parts of speech information. Then it extracts keywords as NamedEntity, onomastic words, onomastic word strings, and word strings stored in the Keyword Database.

A set of the extracted keywords and the customer ID ci are sent to the Data Register Module of the Marketing Server. The Data Register Module check each of the received keywords against the Keyword Learning Database to retrieve the score x of the keyword on each preference category St. This process is performed on every keyword to calculate Sc = {x1, x2, ...}, which is a set of the preference strength scores x for the preference category St. The Sc is written to the position where the customer ID ci and the preference category St are crossed in the current preference vector space Vt. When a value is already written in the position, the new score is simply added to the existing value.

## 6.2 Evaluation of Data Compressibility Rate in Three-Dimensional Vector

The vector space V is a two-dimensional vector space that is composed of category and customer axes for a unit time. It is implemented as integer arrays of 4 bytes and simply stores a unit of data in 12 bytes in total (4 bytes each for category axis, customer axis, and preference strength score).

As a result of the experiment, even the simple method like this enables us to store the past preference vector information for a unit time, which is managed in the time management module, in the size less than 5% of the preference vector space V in the preference transition conversion module.

- Each customer c is interested in only less than 10% of categories among the ones managed in the preference category table,
- Only 20% of customers in maximum react to new categories that are recognized through advertisements or social phenomena, and
- Customers react to only 5 or 6 new categories in maximum (only 1% of total number of categories) that

are recognized through advertisements or social phenomena.

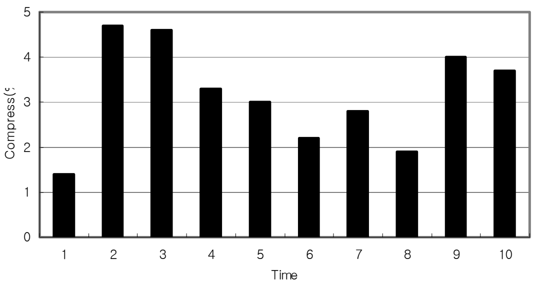



Fig. 7 Compressibility Ratio of Two-dimensional Vector Space V

Implementing a preference field in a three-dimensional array requires enormous size of storage. For example, when you have one million customers and a thousand preference categories, and if a unit time is one week and one year consist of fifty unit times, four bytes of a preference strength score leads the total volume of a preference filed to approximately 200 GB.

The handling is not easy if Customer Preference Database becomes large-scale. However, if it becomes about 5% of data about capacity, i.e., 10GB, like an experiment result, it will become possible to deal with it enough also with a personal computer.

The behavior information of customers enables variety of analyses. For example, transitions of customer preferences and market trends can be observed in various time passages by adjusting the length and the cycle of the unit time. It is also possible to link this system to existing statistics tools to make various analyses such as regression analysis and factor analysis.

## 6.3 Analytical Results of Customer Preference Information

Then we have analyzed the results of clustering customer preference information stored in three-dimensional vectors.

We firstly retrieve the customers who showed preferences on the category "Sub notebook computer" on 6th. and 7th. weeks, and twelve customers are extracted. In this period, the major Japanese PC manufacturers introduced portable sub notebook computers to the market. We investigate the preference transition of these customers from 1st. to 5th. weeks.

Figure 8 shows the preference transitions of these twelve customers from 1st. to 7th. weeks. The figure indicates that some of these customers are clustered into the following two groups:

A) three customers who transited to "Sub notebook computer" from the area "Portable devices" such as "Portable AV devices" or "PDA" (indicated by triangle markers)
B) five customers who transited to "Sub notebook computer" from the area "Personal computer" such as "Desktop computer" or "Notebook computer" via the area "Economics" (indicated by square markers).

As a result of further investigation, we find that the customer $\boldsymbol{A_1}$ (indicated by black triangle markers) and $\boldsymbol{B_1}$ (indicated by black square markers) use the Internet far frequently than the others. Although the customer $\boldsymbol{A_1}$ and $\boldsymbol{B_1}$ show their preferences to the same areas as the others, they always collect information actively and therefore their preferences transit one step ahead of the others. This indicates the possibility to make appropriate recommendations to the other members of each group by predicting the areas they will show interests in the future more accurately based on the preference transitions of the customer $\boldsymbol{A_1}$ and $\boldsymbol{B_1}$.

Instead of clustering customers simply by their preferences, clustering them by the transitions of their preferences achieves higher clustering precision. Furthermore, it enables us to trace the preference transitions of trend leaders in the clustered customer groups, which serves as base information to make recommendations to the other group members for providing more effective Behavior Targeting.

Since this experiment is small sized and is conducted against the customers who work in a single industry, we need to examine in future if the similar results can be achieved in larger experiments with broader ranges of customer types. However, Web Publisher has demographic information (basic customer information) such as address, age, and job, so that the various types of experiments should be available with it.

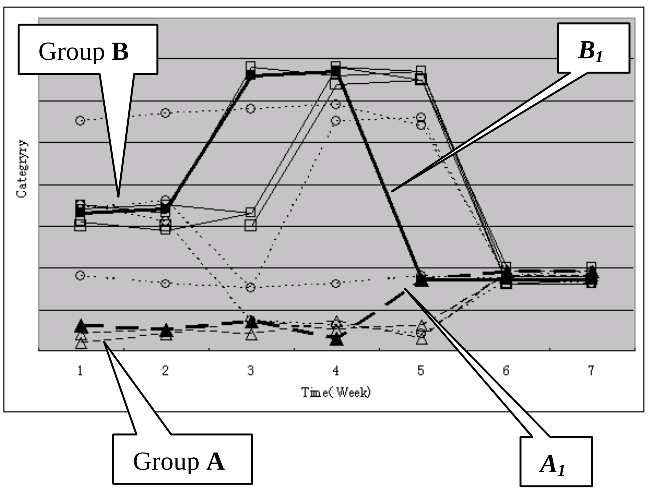



Fig. 8 Preference Transitions of Customers

In the larger sized experiments, the fractionization of customer clustering and the increased number of customer groups may cause processing capacity problems. It is true that the heavy load is imposed on a database when managing customer preference information using multidimensional structured databases such as On-line Analytical Processing (OLAP) [20-22]. In our proposed method, however, the transitions of customer preferences are stored in three-dimensional vectors so that the customer clustering based on the preference transitions is accomplished only by matching vectors and can be executed at considerably higher speed.

## 7. Conclusion

In this paper, we have proposed an information management system to keep track of the transitions of market trends and customer preferences in real time. The extracted information is used for data mining that enables Web Publisher to provide the effective Behavior Targeting service. Using this system, we are able to precisely manage the information of customer preferences and market trends that change over time, which accordingly makes the result of analyses using various analysis tools highly reliable.

This system enables Web Publisher, who provides the Behavior Targeting service, to determine which advertisements or recommendation should be provided to a specific customer. Especially on advertisements, not only the high hit ratio wins the advertisers' esteem but also the valuable information is obtained for determining which category of new advertisers they should target for.

As future work, we need to improve the compression method of the Customer Preference Database. The structure of Customer Preference Database we propose is a set of two-dimensional vectors with customer and preference category axes that are managed per unit time. Considering these two-dimensional vectors as image data, the method to record only the changed points is very similar to the method for compressing image data. In future, we consider applying ISO/IEC IS15938 (MPEG-7) [23], an international standard for image data compression, to our Customer Preference Database. If this is accomplished, we can use low-priced hardware that processes MPEG data at higher speed. This should enable us to use shorter unit time for managing two-dimensional vectors and to grasp the transitions of customer preferences more accurately.

The Keyword Learning Database is also play an important role for grasping customer preferences. We have proposed the method to update the Keyword Learning Database, however, we have not proposed a method to learn the scores for specific categories that each keyword has. In future, we also consider the method to learn the scores automatically.

## References

[1] Moriarity,R.T. and Swartz,G.S., “Automation to Boost Sales and Marketing”, Harvard Business Review, January-February, pp.100-108, 1989.

[2] Morio Nagata, “SFA: Software for Customer Satisfaction”, IPSJ Magazine, Vol.40, No.5, pp.528-531, 1999.

[3] Youji Kohda and Susumu Endo, A New Advertising Business Framework in the 1:1 Marketing, IPSJ Journal, Vol.37, No.6, pp1235-1236, 1996.

[4] Michael J.A. Berry and Gordon Linoff, “Data Mining Techniques: For Marketing”, Sales, and Customer Suppport, John Wiley & Sons, Inc., 1997.

[5] Pieter Adriaans, Dolf Zantinge, “Data Mining”, Addison-Wesley, 1996

[6] Langley,P., “The Computer-aided Discovery of Scientific Knowledge”, Proc. 1st International Conference on Discovery Science, pp.25-39, 1998

[7] Valder-Perez,R.E., “Principles of Human Computer Collaboration for Knowledge Discovery”, Artificial Intelligence 107, pp.335-346, 1999.

[8] M. Balabanovic and Y. Shoham. “Fab: Content-based, collaborative recommendation”. *Communications of the ACM*, 40(3):66–72,1997.

[9] http://www.yahoo.co.jp

[10] http://www.dowjones.com

[11] http://www.revenuescience.com

[12] K.Yamanishi, Data and Text Mining (in Japanese) Journal of Computational Engineering, Oct. 2001.

[13] A.V. Aho and M.J. Corasich, “Efficient String Pattern Matching: An aid Bibliographic Search”, Common. ACM, 18(6) (1975), 333-340

[14] Kazuhiko Tsuda, Masao Fuketa and Jun-ichi Aoe, “An Incremental Algorithm for String Pattern Matching Machines”, Intern. J. Math., Vol.58, pp.33-42

[15] Hisatoshi Mochizuki, Masayuki Kessoku, Masafumi Koyama, Kazuhiko Tsuda : "How to Implement Efficient System for Managing Market Trends and Customer Orientation Transitions″, 5th International Conference on Knowledge-Based Intelligent information Engineering Systems.(KES2001), Osaka-Kyoiku Univ. JP, pp.142-146 (2001.09)

[16] Yoshinori Fukue, Masayuki Kessoku and Kazuhiko Tsuda : "Extracting Purchase Patterns in Convenience Store E-Commerce Market Using Customer Cube Analysis″, 8th International Conference on Knowledge-Based Intelligent

information Engineering Systems.(KES2004), Wellington New Zealand. , pp.501-508, (20-24 September 2004)

[17] Masayuki Kessoku, Masakazu Takahashi and Kazuhiko Tsuda : "A Method of Customer Intention Management for a My Page System" , 8th International Conference on Knowledge-Based Intelligent information Engineering Systems.(KES2004), Wellington New Zealand. , pp.523-529, (20-24 September 2004)

[18] Masayuki Kessoku, Masakazu Takahashi, Kazuhiko Tsuda : "My Page Service Realizing Method by Using Market Expectation Engine", 9th International Conference on Knowledge-Based Intelligent information Engineering Systems.(KES2005), Melbourne Australia., pp.Ⅳ-765-771, (14-16 September 2005)

[19] http://www.cidf.org/english/index.html

[20] S. Chaudhuri and U. Dayal. "An overview of data warehousing and OLAP technology". *ACM SIGMOD Record*, 26(1):65–74, 1997.

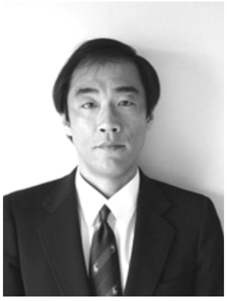
[21] R. Kimball. "*The Data Warehouse Toolkit*". JohnWiley & Sons, Inc., 1996.

[22] J. Srivastava, R. Cooley, M. Deshpande, and P.-N. "Tan. Web usage mining: Discovery and applications of usage patterns from web data". *SIGKDD Explorations*, 1(2):12–23, 2000.

[23] http://www.cselt.stet.it/mpeg.

**Masayuki Kessoku** received the B.S. degrees in Sacred Musicology from in Sacred Music from Elisabeth University of Music in 1984, respectively. He was with Matsushita Electric Industrial Co., Ltd. during 1986-2001, and with NTT DATA CORPORATION during 2001-2005.

**Jun-ichi Aoe** received the B.E. and M.E. degrees in electronic engineering from the University of Tokushima, Tokushima, Japan, in 1974, and 1976, respectively, and the Ph.D. degree in communication engineering from the University of Osaka, Japan, in 1980. Since 1976 he has been with the University of Tokushima. He is currently an Professor in the department of Information Science and Intelligent systems at the University of Tokushima. He is the author of about 50 scientific papers. His rearch interests include computer algorithms in software engineering and natural language processing. Dr. Aoe is editing Computer Algorithms Series of the IEEE Computer Society Press, and two books is published: Computer Algorithms -Key Search Strategies- in 1991, and -String Pattern Matching Strategies- in 1994. He received a Best Author Award from the Information Processing Society of Japan in 1993. He is a member of the Information Processing Society of Japan, the Insutitute of Electronics, Information and Communication Engineers of Japan, the Association for Natural Language Processing of Japan and the IEEE.



**Kazuhiko Tsuda** received the B.S. and Ph.D. degrees in Engineering from in engineering from Tokushima University in 1986 and 1994, respectively. He was with Mitsubishi Electric Corporation during 1986-1990, and with Sumitomo Metal Industries Ltd. during 1991-1998. He is an associate professor in Graduate School of Business Management, University of Tsukuba, Tokyo, Japan during 1998-2005, professor since 2005. His research interests include natural language processing, database, and human-computer interaction. He is a member of The Information Processing Society of Japan and The Institute of Electronics, Information and Communication Engineers.