2013 Second IIAI International Conference on Advanced Applied Informatics

# Neighborhood User Estimation from Web Access Log in EC Service

Tomohiro Koketsu
*Graduate School of Engineering*
*Osaka Prefecture University*
*Sakai, Osaka, Japan, 599–8531*
*Email: kouketsu@sig.cs.osakafu-u.ac.jp*

Hidekazu Yanagimoto
*Graduate School of Engineering*
*Osaka Prefecture University*
*Sakai, Osaka, Japan, 599–8531*
*Email: hidekazu@kis.osakafu-u.ac.jp*

Michifumi Yoshioka
*Graduate School of Engineering*
*Osaka Prefecture University*
*Sakai, Osaka, Japan, 599–8531*
*Email: yoshioka@cs.osakafu-u.ac.jp*

***Abstract*—An aim of this paper is to find neighborhood users using customers' access logs in an Electric Commerce site. In general recommendation services neighborhood users usually are defined according to their order histories or their demographic information. Since a neighborhood user estimation algorithm does not define the neighborhood of users that have never bought any products in an EC site, a cold start problem happens in collaborative filtering. To overcome the cold start problem we make user profile from his/her access logs instead of order histories. We have to select access logs that show user's intent clearly since access logs include user's activities that are not related to product purchase. Hence, we focus on Web pages that many users who purchase the same products visited. We assume the Web pages affect user's purchase and use them as characteristic Web pages to predict users' intent. Since we think such the characteristic Web pages are different in product categories we define a set of Web pages in each category. After finding the characteristic Web pages in all categories, we make a feature vector which elements denote the characteristic Web pages and value denotes whether the user visited the Web pages or not. We calculate the similarities among product categories using sets of users buying a product included in a category to check a user profile discriminates categories. Carrying out evaluation experiments, we confirmed that there is high similarity among users who purchase products in the same category and the similarity is a discriminative criterion. And we estimate neighborhood of new users, which have never bought any products, and confirm that the neighborhood includes many users that have bought products in the same category.**

***Keywords*-access log analysis, recommendation system, cold start problem, PageRank, neighborhood estimation.**

## I. Introduction

Users' action histories are stored as access logs in Web servers. Using the access logs, you can extract users' intents and predict their next activities. Especially, it is important to predict to predict users' activities in Electric Commerce (EC) sites, since it improves users' experiment. In this paper, we focus on access logs in EC sites.

In EC sites recommendation systems use collaborative filtering algorithm using customers' order history data or users' demographic information. If we can recommend suitable products for users searching a products that fit their preference, we achieve a high conversion rate. One of the famous recommendation systems is collaborative filtering using customers' order histories or users' demographic information. In these systems it is important to find neighborhood users who have a similar preference since the system predicts products using the neighborhood users' order histories. However, there is a problem that we cannot find neighborhood users for users who visit the site for the first time or have never purchased any products.This problem is called cold start problems. In this paper, we tackle with the cold start problem using users' access logs in an EC site instead of users' order histories. From access logs we can find users' activities and using the activities we can predict users' intents. Since we can get access logs of users that never purchased any products or registered their demographic information completely, we avoid the cold start problem using the access logs.

Many access logs in EC sites' database include information on customers' preferemces of products or manufacturers and so on. For example, we can find what device users use, from where the users come here, how long the users stay in a Web page, and so on[1], [2], [3]. As a approach to solve the cold start problem, we have the goal to define appropriate neighborhood users from access logs. However, the access logs are more ambiguous than the order histories in terms of quality of information to predict users' intents, since the access logs include various kinds of user's intents.To avoid the drawback of access logs we select access logs that show users' intent clearly. Hence, we focus on Web pages that many users who purchase the same products visited. The Web pages affect users' decision of product selection or category selection. Since such the Web pages are different in categories, we have to define a set of Web pages (We call these pages "$Characteristics\ Pages$; $CPs$") in each category. We used our previous research result to select $CPs$[4]. After selecting the $CPs$, we make a feature vector which elements denote the $CPs$ and value denotes whether the user visited the $CP$ or not.

We carried out some experiments using access logs in a real EC site and discussed the effectiveness as the way to estimate the neighborhood user. We consider that access logs in EC site, but not the other Web site like as portal site, is easy to select $CPs$ since we can understand the result of user behavior as their purchase. From experimental results,

978-0-7695-5071-8/13 $26.00 © 2013 IEEE
DOI 10.1109/IIAI-AAI.2013.67

we confirmed that we get high similarity among users in the same category; using the feature vector as a user profile and the user profile can discriminate categories. And we confirmed that we can define neighborhood of a user using the profile and recommend a product for new users, which have never bought any products.

## II. Related works

As a service using the users' order histories or their demographic information in an online shop, there are online advertising services and product recommendation services. For example, Google Ad and Amazon. These researches have the aim that they find related products using customers' order histories in the online shop[5], [6], [7], [8]. To achieve the aim they search neighborhood users and select products using neighborhood users' preference. In the research field collaborative filtering is used generally. In this method, recommended products have been decided based on order histories that customers have actually purchased. Hence, a cold start problem happens since the method can not recommend appropriate products to users that have never bought any products. Our proposed method solve the cold start problem using access logs to define neighborhood users.

We explain access log analysis method[4] that is one of key ideas in our proposed method. We analyzed customers' access logs and sorted Web pages according to affecting users' purchase. And we could find rival products against an actual order product. First, we constructed a network to express a relationship among Web pages based on transition information that have been recorded in the access logs of the customers who purchased products in the same category. Second, we evaluated the importance of Web pages in the network using PageRank[9], [10], [11]. PageRank evaluate the set of links in the entire Web but we evaluated the set of links based on users' behaviors in an EC site. In this paper, we can regard the Web pages extracted by the access log analysis as the feature of customers' behaviors. We call the Web pages "$Characteristics\ page$" in this paper.

## III. Proposed Method

In this paper, using the $CPs$ we make user profiles. The profiles have some good characteristics. One is to discover neighborhood users efficiently. Moreover, we control computational cost to calculate similarities by selecting the number of $CPs$. Since we can obtain the $CPs$ in advance, when we estimate the neighborhood users, there is no cost.

Then, we explain how to extract $CP$, make user profile and calculate similarities among users as components of proposed method.

### A. Extract Customers' Intents as the Characteristics Page from Access Logs on their Purchase

In this paper, we extract $CPs$ on customers' purchase processes in each category, and use for making user profile to describe user's preference. Based on our previous research[4], we describe how to extract the $CPs$.

First, we construct a network for representing of relationship among Web pages based on their transition information. Specifically, we make an adjacency matrix $H$ using Web page transition information according to the following equation (1).

$$h_{uv} = \begin{cases} n_{uv} & (\text{the number of transition } u \to v) \\ 0 & (\text{no transition, or, } u = v) \end{cases} \quad (1)$$

Second, we transfer the adjacency matrix $H$ to a transition probability matrix $S$ for representing of the users' stochastic behaviors using the following transform. $H$ is a $k \times k$ ($k \geq 2$) matrix. The $k$ denotes the total number of Web pages that have been confirmed transition

$$s_{uv} = \begin{cases} \frac{h_{uv}}{\sum_{i=1}^{k} h_{uv}} & (\text{if } h_{uv} > 0) \\ \frac{1}{k} & (otherwise) \end{cases} \quad (2)$$

After the transformation the following condition holds.

$$\sum_{v=1}^{k} s_{uv} = 1 \quad (3)$$

This model is called random surfer model[9], [12], [13]. Scores in PageRank is defined in an eigenvector corresponding to the maximum eigenvalue of $S$. Strictly speaking, the score represents probabilities under which user on the network walks according to the transition probability matrix S. That is, the probability corresponds to a value of PageRank as the importance of each Web page. Using the method, we obtain a pair of URL and a score of PageRank. Web pages that many customers visited tend to be in top layer. We regard such Web page excluding top pages and order pages and so on, as $CPs$ which affect users' purchase. Hence, we use the $CPs$ to represent customers' behaviors. The $CPs$ are obtained in each category since users take different activities in each category.

When we construct networks and extract $CPs$ according to our previous method[4], it is important for us to collect the access log data related to the processes on customers' purchase. In the study, two ranges to collect access logs have been proposed. In this paper, we adopt the range rule considering from first target contact to the order day. Actually 33.5% order in the current dataset have the contacts to the target product before the order day.

### B. Making the User Profile based on CP

Based on the method in III-A, we extract the $CPs$ for each category. Using them we make a user profile as shown in Table 1 from actual access logs for each category. For example, we consider the case that we compare category A and B, the number of customers of category A is $p$ and category B is $q$. Previously, we have to decide the amount

$n$ of the $CP$ used for making a profile. For example, if the number of $n$ is 100, we deal with 200 ($2n$) $CPs$ for the profile since there are two categories.

Table I
THE USER PROFILES OF USERS IN TWO CATEGORIES BEING COMPARED.

| Users$A$+$B$\Feature Quantity(URL) | $w_1$ | $w_2$ | ... | $w_{2n}$ |
|---|---|---|---|---|
| $u_1$ | 1 | 0 | | 0 |
| $u_2$ | 0 | 1 | | 1 |
| ... | | | | |
| $u_{p+q}$ | 1 | 1 | | 1 |

In Table 1 if a user $u_i$ visited a $CPs$, the value of feature vector in his/her profile is 1. If the user did not visited the page, the value of feature vector is 0. Hence, the user profile is a binary vector.

### C. Neighborhood User Estimation

Using the user profile, we define similarity among users to estimate the neighborhood users. We use cosine similarity to define the degree of similarity between users. Cosine similarity is defined below.

$$c_{ij} = \frac{\sum_{l=1}^{2n} x_{il}x_{jl}}{\sqrt{\sum_{l=1}^{2n} x_{il}^2}\sqrt{\sum_{l=1}^{2n} x_{ji}^2}} \quad (4)$$

Where $c_{ij}$ is the cosine similarity between $i$-th user and $j$-th user. $x_{il}$ is the $l$-th element of feature vectors of $i$-th user. $x_{jl}$ is the $l$-th element of feature vectors of $j$-th user. Calculating the similarity among all users according to their profile, we obtain a similarity matrix which scale is $(p+q) \times (p+q)$ in Table 2.

Table II
SIMILARITIES AMONG ALL USERS IN PROFILE INFORMATION.

| Users $A+B$ \ Users $A+B$ | $u_1$ | $u_2$ | ... | $u_{p+q}$ |
|---|---|---|---|---|
| $u_1$ | $c_{11}$ | $c_{12}$ | | $c_{1p+q}$ |
| $u_2$ | $c_{21}$ | $c_{22}$ | | $c_{2p+q}$ |
| ... | | | | |
| $u_{p+q}$ | $c_{p+q1}$ | $c_{p+q2}$ | | $c_{p+qp+q}$ |

From Table 2 we can define the neighborhood user based on the degree of the similarity. We need to confirm an effectiveness of the user profile using $CPs$. Because access log includes various kinds of users' intent and it is more difficult to predict the intents in our proposed method than in order history-based approach.

We examine the results of changing various condition such as the number of the $CPs$. And we discuss the influence of these conditions and confirm the effectiveness of our proposed method.

## IV. EXPERIMENTS

We carry out some experiments using actual access logs in a real EC site to evaluate our proposed method, the neighborhood user estimation using the access log. First, we explain dataset. Second, we explain the experimental procedure. Our experiment consists of 4 components, extraction of $CPs$ that show customers' intents in each category, construction of the users' profile, similarity estimation among users and the neighborhood user estimation. Finally, we discuss the whole experiment.

### A. Data

In experiments we use a dataset containing 1,561,205 access logs and 6,656 order histories in an EC site. The customers' behavior data is provided by Joint Association Study Group of Management. The access log data contains 22 attributes as the following Table 3. The order history data contains 12 attributes as the following Table 4.

Table III
THE ATTRIBUTES IN ACCESS LOG DATA

| | |
|---|---|
| ・Serial number | ・Member value |
| ・Session ID | ・Last session ID |
| ・User ID | ・PV(Page View) |
| ・Time and Date of Event | ・Visit classification |
| ・Referrer URL | ・Reservation flag |
| ・Access URL | ・Purchase flag |
| ・Order ID | ・Product ID |
| ・Session starting time | ・Category of products |
| ・Session finishing time | ・Retrieval keyword |
| ・Session time | ・Kinds of browser |
| ・Kinds of OS | |

Table IV
THE ATTRIBUTES IN ORDER HISTORY DATA

| | |
|---|---|
| ・Order ID | ・Contract amount |
| ・User ID | ・Gross amount |
| ・Session ID | ・How point used |
| ・Order date | ・How point obtained |
| ・Product ID | ・Order quantity |
| ・Kinds of manufacturers | ・Category of products |

We use the order histories to decide a target product or a target category. By limitation of product or category we can find rival products, which are products compared with a product frequently. From access logs we select attributes required for experiments as the following Table 5.

Table V
THE USING ATTRIBUTES IN THE DATASET

| Access log data | Order data |
|---|---|
| ・User ID | ・User ID |
| ・Order ID | ・Order ID |
| ・Time and Date of Event | ・Product ID |
| ・Referrer URL | ・Order Date |
| ・Access URL | ・Category of products |
| ・Product ID | ・Kinds of manufacturers |
| ・Session ID | |
| ・Session starting time | |
| ・Session finishing time | |

In the order history data 44 kinds of categories included, for example, shirt or cap, driver and so on. In this paper, as

the target of the experiments we picked up 5 categories. We show the number of customers in the 5 categories in Table 6.

Table VI
ORDER INFORMATION IN EACH TARGET CATEGORY.

| Category | Once | More than once |
|---|---|---|
| $c_1$ | 15 | 174 |
| $c_2$ | 9 | 146 |
| $c_3$ | 30 | 152 |
| $c_4$ | 10 | 158 |
| $c_5$ | 28 | 132 |

In Table 6 $Once$ means that the number of users who purchase a product in the category only once and regarded as new user in our experiments. $More\ than\ once$ mean that the number of users who have purchased products more than once and is candidate of neighborhood users for the new user. We regard users who purchase more than once in the site as training data, users who purchase a product only once as test data. We use training data to extract $CPs$ and define the neighborhood users, test data to evaluate the accuracy of the neighborhood users obtained by our proposed method.

### B. Experimental Procedure

Based on our proposed method explained in the previous section, we conduct experiments according to the following steps. As mentioned above the experiments consists of three components.

Step1. We determine a target category and collect access logs of customers who purchased the same target according to the range rule[4]. Then, using the multiple customers' log data we construct a network and calculate PageRank. We define the $CPs$ in the target category.

Step2. We make user profiles using the $CPs$. The user profile is defined according to actual logs in purchase process.

Step3. Using user profile, we calculate the similarity between two categories.

### C. Experimental Results

In this section, we describe the results for selecting $CPs$ according to our previous method. Then, we show the results of calculating similarity among users.

*1) Extruct important Web pages on customers' purchase:* In this paper, we focus on customers in the 5 categories. The following Table 7 is a part of the result of selecting $CPs$ in the category $c_1$.

In Table 7 we can find that customers in the category often visited a specific product page like the rank 1, 3, 4, 8, 9 or 10 and so on. Also we can find the customer has the tendency in visiting a specific manufacturer like the rank 5 to 7. In our previous research[4], it is confirmed that these pages are different from ones evaluated by using page view

Table VII
THE CHARACTERISTICS OF CUSTOMERS'BEHAVIORS IN CATEGORY $c_1$

| Rank | PageRank | Web page(contents) |
|---|---|---|
| 1 | 0.0595317 | A Product |
| 2 | 0.0595317 | A Customer Review |
| 3 | 0.0595317 | A Product |
| 4 | 0.0592781 | A Product |
| 5 | 0.0551616 | A Manufacturer |
| 6 | 0.0548055 | A Manufacturer |
| 7 | 0.0547560 | A Manufacturer |
| 8 | 0.0595317 | A Product |
| 9 | 0.0592781 | A Product |
| 10 | 0.0595317 | A Product |
| | | |
| 12,405 | 8.18607e$^-$07 | An image of Products |

(PV) or staying time (ST) that sre the evaluation criteria on general access log analysis[1]. Since the method evaluate Web pages based on the transitions of pages by customers. Hence, we can consider that these pages have influence to customers' purchase as $CPs$.

In the same way we selected $CPs$ in other categories in advance.

*2) Estimation neighborhood user:* In general, the more features are, the more time cost you need. Therefore, we need to reduce features as much as possible. We set 7 types of user profiles using $CPs$ as top 100, 500, 1,000, 1,500, 2,000, 3,000, and non threshold, for making the user profiles in order to discuss the relation between the amount of $CP$ and the similarity among users.

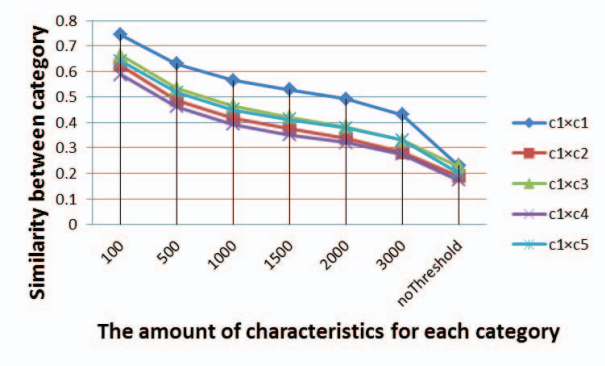


Figure 1. Similarity in customers' behaviors for each category.

In Figure 1 we calculate similarities between category $c_1$ and other categories. We can find that there is high similarity among users who purchase the products in the same category. Hence, the user profile made from access log is effectiveness, since the user profile can gather users that buy products in the same category. And from Figure 1 we find that filtering $CPs$ is superior to noThreshould. In the result of noThreshould it is difficult for us to make clear the difference between category $c_1$ and another category. However, in the result of filtering $CPs$ it is easy to separate category $c_1$. We think that the user profile using filtered

$CPs$ can capture intent of users which bought products in the same category.

From Figure 1 we find that there is no difference in the results of feature vector of user profile in the case of from 100 to 3,000 in terms of neighborhood user estimation. Hence, we can consider it is sufficient to use at least 100 $CPs$ for making user's profile. Therefore, we can expect to reduce the time cost. In addition, we need to set other types of user profiles using less the number of $CPs$, and confirm that an optimal number of $CPs$ enough to make the user profile.

*3) Category estimation of purchases by new users:* In order to the effectiveness of neighborhood user estimation we conducted experiments of category estimation of products purchased by new users. As mentioned above, we regarded the users who purchase only once as test data. When we make the user profile about new user, we delete some Web pages that seem to be related to their purchase directly, for example, Web pages that contain the character of "purchase" or "order".

In this experiments, we make user profiles using 100 $CPs$ in each category. Therefore, the size of the feature vector per user is 500. When we calculate neighborhood users for a new user who is every one of the 87 users in the same way as above, we can define the similarity of 762 users. When we focus on a user in the category $c_1$, we show an example of result in Table 8.

Table VIII
A RESULT OF NEIGHBORHOOD USER ESTIMATION FOR A USER IN CATEGORY 1

| Rank | Similarity | Category of neighborhood user |
|---|---|---|
| 1 | 0.80403 | $c_4$ |
| 2 | 0.76980 | $c_1$ |
| 3 | 0.76980 | $c_1$ |
| 4 | 0.73786 | $c_1$ |
| ... | | |
| 762 | 0 | $c_2$ |

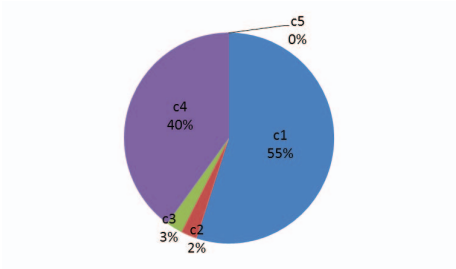


Figure 2. Category occupancy rate of $N = 40$ for a user in the category $c_1$

From Table 8 a nearest neighbor user of target user is in category $c_4$, but the category is different from a category where the target user bought a product. In this paper, we determine an estimated category as a category that is occupied highest percentage in the top $N$ users. If the estimated category matched the category of the new user, we can predict a correct category. From Figure 2 when focusing on the top 40 neighborhood users ($N = 40$), the estimated category is $c_1$. In the case we conducted a correct estimation.

We show the result of 5 categories in Table 9 and the number of the correct estimation in the result in Figure 3.

Table IX
THE RESULT OF CATEGORY ESTIMATION EXPERIMENTS

| Category | $N$=10 | $N$=20 | $N$=30 | $N$=40 | $N$=50 |
|---|---|---|---|---|---|
| $c_1$ | 10 | 9 | 8 | 9 | 9 |
| $c_2$ | 3 | 2 | 1 | 0 | 0 |
| $c_3$ | 23 | 16 | 13 | 14 | 13 |
| $c_4$ | 5 | 7 | 4 | 5 | 6 |
| $c_5$ | 13 | 11 | 16 | 0 | 12 |

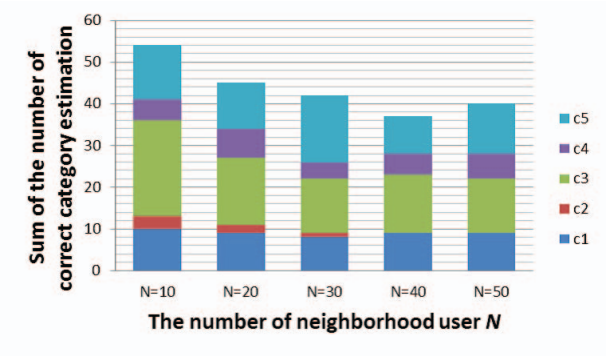


Figure 3. The sum of correct category estimation

Table 9 and Figure 3 show that in the result of $N = 10$ the number of correct answers is most. The rate of accuracy is 54/87 in $N = 10$. The result means that we can carry out the correct approach to 54 new users on their purchase processes, shows that our proposed method is effective.

On the other hand, when we focus on category $c_2$, the accuracy rate decreases as the number of neighborhood users. We consider that users in other categories affect category estimation as noise or the user profile of users in the category 2 is not able to work well.

## V. DISCUSSION

We confirmed the effectiveness of making user profile from actual access logs and neighborhood user estimation. And furthermore, using results of the estimation we conducted experiments for predict the category of products that new users purchase.

However, in category estimation experiments we have to improve the accuracy rate in order to develop the real recommendation service. As improvements, we discuss our experimental results, the comparison with different condition about the amount of $CPs$ when we make the user profile and the quality of $CPs$ from two viewpoints.

### A. Dependency of log data size and variation

Since provided dataset is the information for a certain period of time, the dataset does not include all access log enough to understand users' intents. In this regard, we may have not been made user profile well. For example, in category estimation of category $c_2$, the characteristics of user profiles in category $c_2$ might just not enough in this data. Hence, we need to collect more data and try to apply the proposed method to them.

In addition, according to the type and amount of data, we have to improve the way to make the user profile, since an EC site has been diversifying in recent years in terms of place, time or device for purchase. As an example, we can make user profile reflected the kinds of device or the number of order in the site.

### B. Selecting condition of Characteristics pages

When we select the $CPs$ in each category, we need to identify users' purchase processes in order to find Web pages that many users who purchase the same products visited. Since the quality of $CPs$ directly affects making user profile and neighborhood user estimation. For example the result of $N = noThreshold$ in Figure 1 shows that the user profile using all access logs cannot discriminate categories since the profiles have unwillingly various users' intents. There is need to continue improvement of how to identify the access logs correctly reflected intents on users' purchase processes.

## VI. Conclusion

In this paper, we have performed some experiments with our proposed method and conducted the neighborhood user estimation from customers' access logs on their purchase processes. From the results we discovered neighborhood users among customers who purchased products in the same category. And we confirmed that we can define neighborhood of a user using the profile and recommend a product for new users, which have never bought any products.

As future works, we have to conduct experiments with various experimental conditions and confirm the validity of $CPs$ extracted from access logs as users' intents.


## Acknowledgment

We thank Joint Association Study Group of Management for help with obtaining the customers' behavior data.


## References


[1] Taku Ogawa, Introduction to Web Analytics, Softbank Creative (2010)

[2] Kira Radinsky, Krysta Svore, Susan Dumais, Jaime Teevan, Alex Bocharov, Eric Horvitz, Modeling and Predicting Behavioral Dynamics on the Web, In $Proceedings\ of\ the\ 21st\ international\ conference\ on\ World\ Wide\ Web$ ($WWW2012$), Pages 599-608, (2012)

[3] Ravi Kumar, Andrew Tomkins, A Characterization of Online Search Behavior, In $Proceedings\ of\ the\ 19th\ international\ conference\ on\ World\ Wide\ Web$ ($WWW2010$), Pages 561-570, (2010)

[4] Tomohiro Koketsu, Hidekazu Yanagimoto, Michifumi Yoshioka, Access Log Analysis with PageRank for EC Service, In $The\ First\ Asian\ Conference\ on\ Information\ Systems$ ($ACIS2012$), (2012)

[5] G. Linden, B. Smith, and J. York: Amazon.com Recommendations: Item-to-Item Collaborative Filtering, IEEE Internet Computing, Vol.7, No. 1, pp.76-80, 2003

[6] J. Gittins, K. Glazebrook, and R. Weber : Multi-armed Bandit Allocation Indices 2nd Edition, John Wiley & Sons Ltd, 2011

[7] P. Auer, N. Cesa-Bianchi, and P. Fischer : Finite-time Analysis of the Multiarmed Bandit Problem, Machine Learning, Vol. 47, pp.235-256, 2002.

[8] R. S. Sutton and A. G. Barto, S. Mikami, M. Minagawa :Reinforcement Learning, Morikita (2000)

[9] L. Page, S. Brin, R. Motwani and T. Winograd, The PageRank citation ranking : bringing order to the Web, Technical report, Stanford University(1998)

[10] Amy N. Langville, Carl D. Meyer, Google's PageRank and Beyond: The Science of Search Engine Rankings, Kyoritsu (2009)

[11] Jon M. Kleinberg, Authoritative source in a hyperlinked environment, In $Journal\ of\ the\ ACM\ Volume\ 46\ Issue\ 5$, Pages 604-632, (1999)

[12] Steven Bethard, Dan Jurafsky, Who Should I Cite? Learning Literature Search Models from Citation Behavior, In $Proceedings\ of\ the\ 19th\ ACM\ international\ conference\ on\ Information\ and\ knowledge\ management$ ($CIKM2010$), Pages 609-618, (2010)

[13] David F. Gleich, Abraham D. Flaxman, Paul G. Constantine, Asela Gunawardana, Tracking the Random Surfer: Empirically Measured Teleportation Parameters in PageRank, In $Proceedings\ of\ the\ 19th\ international\ conference\ on\ World\ wide\ web$ ($WWW2010$), Pages 381-390, (2010)