Personalized Digital E-library Service Using Users' Profile Information*

Wonik Park[1], Wonil Kim[2], Sanggil Kang[3], Hyunjin Lee[2], and Young-Kuk Kim[1]

[1] Department of Computer Engineering, Chungnam National University, Daejeon, South Korea
{wonik78, ykim}@cnu.ac.kr
[2] College of Electronics and Information, Sejong University, Seoul, South Korea
wikim@sejong.ac.kr, enjoyaje@hotmail.com
[3] Department of Computer Science, University of Suwon, Gyeonggi-do, South Korea
sgkang@suwon.ac.kr

Abstract. We propose a personalized digital E-library system using a collaborative filtering technique, which provides a personalized search list according to users' preference. The proposed system analyzes the registered users' actions such as "clicking" and "borrowing" items. According to the different actions, we provide a weight for calculating the users' preference of each item. However, the list is uniformly provided to the individual users when they search with same keywords. In order to avoid the problem, we customize the order of items in the list according to whether there is any mismatching of profiles among registered users and target users or not.

1 Introduction

Most of the library search systems provide uniformly a list of items, such as books, papers, magazines, etc, to individual users without users' discretion. Usually users are reluctant to spend much time to look up the flood of the unwanted items in the list. To solve the problems, many researchers [1-3] developed the personalized digital library systems.

In this paper, we propose a personalized digital E-library system using a collaborative filtering technique, which provides a personalized search list according to users' preference. The proposed system analyzes the registered users' actions such as "clicking" and "borrowing" items. According to the different actions, we provide a weight for calculating the users' preference of each item. However, the list is uniformly provided to the individual users when they search with same keywords. In order to avoid the problem, we customize the order of items in the list according to whether there is any mismatching of profiles among registered users and target users or not.

The remainder of this paper consists as follows. Chapter 2 explains the proposed system in detail. Chapter 3 briefly explains the over architecture of our system. In Chapter 4, we show the simulated results of our system. Finally chapter 5 will conclude.

* This research was supported by the Ministry of Information and Communication, Korea, under the College Information Technology Research Center Support Program, grant number IITA-2005-C1090-0502-0016.

J. Gonzalo et al. (Eds.): ECDL 2006, LNCS 4172, pp. 528 – 531, 2006.

2 Overall System Architecture

Fig. 1 shows the overall architecture of the proposed E-library service system. The architecture is composed of three modules such as Log Analyzer (LA), Personalization Inference (PI), and Dynamic HTML Generation Machine (DHGM). The LA module collects the registered users' clicking frequency of each item from their log information. If a target user requests a query, the PI module infers the personalized list by our algorithm explained in the following section, which utilizes the obtained users' clicking information, the borrowing information stored in the DB, and users' profile information stored in the Profile DB. The DHGM module shows the inferred personalized list to the target user through Displayed List.

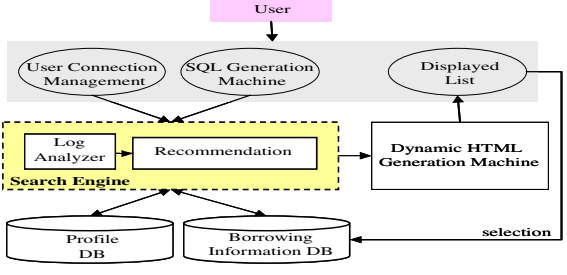

Fig. 1. The overall architecture of the proposed E-library service system

3 Proposed Digital E-library Service

For items retained in Digital E-library such as book, paper, magazine, etc, the preference of each item can be expressed by the actions taken by users. There are two types of actions such as "borrowing and "clicking" items. In conventional method, the preference of an item is expressed with the frequency of clicking for the item by users during a predetermined period, as seen in Equation (1).

$$\Pr ef_i = \sum_{k=1}^{K} c_{k,x_i} \tag{1}$$

where $Pref_i$ is the preference of item x_i, c_{k,x_i} is the frequency of clicking item x_i by user k and K is the total number of users registered in the Digital E-library. Equation (1) is the preference obtained using the action of clicking only. However, the action "borrowing" is usually stronger index for estimating the preference than the action clicking." In order to take into the consideration, we provide a weight w to the action "borrowing" in calculating the frequency of the action of clicking in Equation (1).

$$\Pr ef_i = \sum_{k=1}^{K} (c_{k,x_i} + w \cdot b_{k,x_i}) \tag{2}$$

where b_{k,x_i} is the frequency of borrowing item x_i by user k and $w > 1$. As seen in Equation (2), the accuracy of the preference depends on the value of w. As the

frequency of borrowing an item increases, the preference for the item linearly increases with slop w. Based on the value of the preference, the order of items in the list as the result of a target user's query using one or more than one keyword is determined. However, the list is uniformly provided to the individual users when they search with same keywords. In order to avoid the problem, we customize the order of the list according to users' profile.

In general, the action information of users with the same profile as a target user is more useful for predicting the target user's usage behavior than that of users with different profile. For example, let a target user's profile information be u_1= (Student, Engineering, Computer Engineering, Junior). Also, there is two users' (User 1 and User 2) action information for the book C++ stored in the library database. User 1 with profile information u_1= (Student, Engineering, Computer Engineering, Junior) clicked the book C++ once. User 2 with profile information u_2= (Student, College of Art, Painting, Freshman) borrowed the book C++ once. In this case, even though the action of borrowing is more effective on computing the preference of the book C++ than the action of clicking, the action of clicking is more reliable because the profile information of User 1 is identical with the target user. In order to consider this problem for computing the preference, we modify the frequency of the actions by providing a penalty according to the degree of mismatching profiles between the users in the database and the target user.

$$\Pr ef_i = \sum_{k=1}^{K} p_k (c_{k,x_i} + w \cdot b_{k,x_i}) \tag{3}$$

where p_k is the penalty for the mismatching and $p_k \leq 1$. If there is no mismatching between user k and the target user then $p_k = 1$.

By using Equation (3) for computing the preference of each item, the personalized search list can be provided by according to different target users. Also, as seen in Equation (3), the accuracy of estimating the preference of each item depends on the values of the variables w and p_k. In the experimental section, we show the optimal values of those variables from the empirical experience.

4 Experiment

We implemented our system using the JAVA webserver in the Window NT environment. In the server, we used the JSDK which is Java servlet developer kit 1.4 to run our personalized E-library service program. MS SQL server 2000 was used as the relational database. Also, JDBC (Java Database Connectivity) was used in order to connect database with servlet.

Also, we made a sample E-library website with 1,000 book lists and collected the profile information and the actions of clicking and borrowing of 100 students registered in the Chungnam National University (CNU) in Korea and 50 faculties during one month from October 2005 to November 2005. Their actions were stored in the databases. For 10 target students and 10 target faculties whose actions are not stored in the database, we evaluated the performance of our system by comparing the degree of the satisfaction of our system with that of the E-library system provided by the

CNU. For the weight of borrowing action in Equation (3), we chose $w = 4$. Also, the penalties of mismatching between the k^{th} registered user and a target user are chosen as follows; $p_k = 0.6$. We evaluated the performances of the CNU E-library system and our personalized E-library system by surveying the satisfaction of both systems for each target user.

From Table. 1, 16 target users (3 for "very satisfactory", 13 for "satisfactory") out of 20 for our system expressed their satisfaction at our system, while only 2 target users at the CNU E-library system.

Table. 1. The result of the survey of the satisfaction of both systems for the 20 target users

Evaluation	Target users	
	CNU E-library	**Our system**
very satisfactory	0	**3**
satisfactory	2	**13**
dissatisfied	17	**3**
very dissatisfied	1	**1**

5 Conclusion

In this paper, we proposed the personalized E-library system by considering E-library users' actions such as "clicking" and "borrowing" and their profile information. From the experimental section, it is shown that our system can give satisfaction to E-library users, compared to the existing E-library system.

However, the weight of borrowing action and the penalties of mismatching used in the experiment were obtained from the exhaustive empirical experience. We need to do further study for developing an automatic algorithm in determining the weight and the penalties for various situations.

References

1. Bollacker, K.D., Lawrence S., Giles, C.L.: A System for Automatic Personalized Tracking of Scientific Literature on the Web. Proc. ACM Conference on Digital Libraries (1999) 105-113
2. Lee, W.P., Yang, T.H.: Personalizing Information Appliances: A Multi-agent Framework for TV Program Recommendations. Expert Systems with Applications, vol. 25, no. 3 (2003) 331-341
3. Kamba, T., Bharat, K., Albers, M.C.: An Interactive Personalized Newspaper on the Web. Proc. International World Wide Web Conference (1995) 159-170