Capturing User Interests by Both Exploitation and Exploration

Ka Cheung Sia[1], Shenghuo Zhu[2], Yun Chi[2], Koji Hino[2], and Belle L. Tseng[2]

[1] University of California, Los Angeles, CA 90095, USA
kcsia@cs.ucla.edu
[2] NEC Laboratories America
10080 N. Wolfe Rd, SW3-350, Cupertino, CA 95014 USA
{zsh,ychi,hino,belle}@sv.nec-labs.com

Abstract. Personalization is one of the important research issues in the areas of information retrieval and Web search. Providing personalized services that are tailored toward the specific preferences and interests of a given user can enhance her experience and satisfaction. However, to effectively capture user interests is a challenging research problem. Some challenges include how to quickly capture user interests in an unobtrusive way, how to provide diversified recommendations, and how to track the drifts of user interests in a timely fashion. In this paper, we propose a model for learning user interests and an algorithm that actively captures user interests through an interactive recommendation process. The key advantage of our algorithm is that it takes into account both *exploitation* (recommending items that belong to users' core interest) and *exploration* (discovering potential interests of users). Extensive experiments using synthetic data and a user study show that our algorithm can quickly capture diversified user interests in an unobtrusive way, even when the user interests may drift along time.

1 Introduction

Personalized recommendation systems that provide a user with recommendations on products, news articles, or documents that are tailored toward the his personal interests are being used extensively in e-commerce web sites, news portals, and enterprise documentation portals. As pointed out by the research community recently [1], the five major usability goals for user-adaptive systems are: *privacy*, *controllability*, *unobtrusiveness*, *breadth of experience*, and *predictability and comprehensibility*. We are building a prototype of a *Personal Information Manager* that tries to address the above criteria. Such a system runs on a user's personal computer; it collects the recent important information that matches the user's personal interests from the Web, the blogosphere, and news sites; it then summarizes the collected information and presents to the user in a succinct form.

To solve the above challenging issues, we use a learning framework and propose an algorithm that actively captures user interests through an *unobtrusive* interactive recommendation process. Unlike a greedy algorithm, which only *exploits* the model of users' interests, the proposed algorithm takes into account *exploration*, i.e., it discovers user potential interests through *topic diversification* [2]. In addition, the exploration nature of this algorithm also makes it adapt quickly to *user interest drift* [3] as well. In the following, we will give an overview of the related work and identify our unique contributions as compared to the literature.

Learning user interests has been studied extensively in the area of information retrieval and Web search&mining. In the information retrieval area, relevance feedback [4, 5] has long been used for improving the quality of retrieval. In the Web search and mining area, personalization [6, 7] has been one of the most important research topics. Click history [8–12] is one of the commonly used information to learn a search engine user's interests; some other implicit information such as display time [13] and browse history [14] have also been investigated. While these different approaches have proved effective in various areas, a key point that limits their flexibility is that they are all *passive* in nature. That is, all these approaches *exploit* historic data while ignoring *exploring* additional information from users. In comparison, user interests are *actively explored* in our approach. An active feedback framework [15] is recently proposed for probing user preference by presenting documents that are selected based on a statistical decision theory. It is different from our work in the sense that it requires *explicitly* asking users for feedback and it assumes that the ground truth is available.

2 The Learning Framework

Figure 1 illustrates a system that provides a user with personalized recommendation contents. The system observes the user's activities while she browses the web pages, and shows the user a list of Webpage recommendations. Assuming that *clicking* the link of a recommended Webpage after *reading* a short description of the page indicates that she likes the *topic* of the Web page, the system can learn a user model from this observation and, consequently, provides better recommendations.

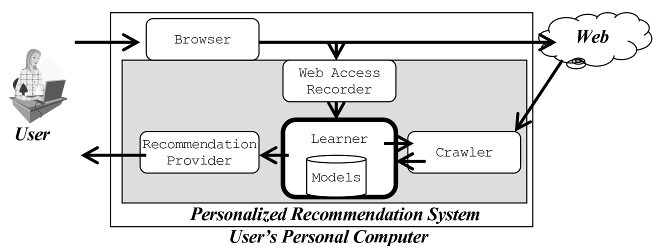

Fig. 1. A diagram of a personalized recommendation system

To model the process of learning user interests, we assume that user interests are represented by a combination of K topics, where K could be a large number. We further assume that each Webpage only belongs to one topic to simplify the model and the analysis. When the description of a recommendation of topic i is read by the user, the user clicks the link of the recommended page with probability $\theta_i = \Pr(\text{click}|\text{read}, \text{topic } i)$. Then the user interest model can be represented by the parameter $\Theta = \{\theta_1, \cdots, \theta_K\}$, which is going to be estimated.

When a recommendation item is shown to the user, it has different chances to attract the user's attention depending on its position on the list. In the Web search engine community, it is observed that the position of an entry on the query-result list heavily affects its chance to be clicked by the user[9]. We capture this phenomenon by a probability model; we denote the probability that the user reads a recommendation at the j-th position of the list as $g(j) = \Pr(\text{read}|j)$, for $1 \leq j \leq K$.

Let variable $R = \{r_1, \cdots, r_K\}$ be the ranking order given to the K topics. Then we can express the utility of this ranking order as

$$U(R;\Theta) = \sum_{i=1}^{K} \Pr(\text{click}|\text{read}, \text{topic } i; \theta_i) \Pr(\text{read}|r_i) = \sum_{i=1}^{K} \theta_i \cdot g(r_i). \quad (1)$$

Given a user model, Θ, the goal of the system is to maximize the utility, $U(R, \Theta)$. As a result, we have two problems to solve: 1) we need to estimate the Θ accurately; 2) we need to choose a ranking order R that maximizes the utility.

Learning Θ We assume the prior of θ_i follows the beta distribution, $B(\theta_i|\alpha_i, \beta_i)$, where α_i and β_i can be initially set to some fixed constants. When the recommendation is ranked at r_i and is not clicked, we have

$$\begin{aligned} d\Pr(\theta_i|\neg\text{click}, r_i) &\propto [1 - \Pr(\text{click}|\theta_i, r_i)]\, d\Pr(\theta_i) = [1 - \theta_i g(r_i)]\, B(\theta_i|\alpha_i, \beta_i) d\theta_i \\ &\approx [1 - \theta_i]^{g(r_i)}\, B(\theta_i|\alpha_i, \beta_i) d\theta_i \propto B(\theta_i|\alpha_i, \beta_i + g(r_i)) d\theta_i, \end{aligned} \quad (2)$$

in which we used the approximation $(1-x)^y \approx 1-xy$. Since $B(\theta_i|\alpha_i, \beta_i+g(r_i))$ is normalized, we have $d\Pr(\theta_i|\neg\text{click}, r_i) \approx B(\theta_i|\alpha_i, \beta_i + g(r_i))d\theta_i$. That is, the posterior distribution of θ_i follows $B(\theta_i|\alpha_i, \beta_i + g(r_i))$ if the user does not click the recommendation of topic i at position r_i of the list. If the recommendation of topic i is clicked, the posterior distribution of θ_i follows $B(\theta_i|\alpha_i + 1, \beta_i)$. Thus, we have the update formula for the distribution of each θ.

Maximizing Utility To maximize the one-step expected utility, we may rank topics according to their expected utilities, $\{\theta_i \cdot g(r_i)\}$ This approach is to *exploit* the best estimation of user interests, Θ, to gain the optimal one-step utility. We call it *greedy* approach. Such an approach puts the best estimated θ's on the list, which may deprive the opportunity of showing the *true* optimal θ's. Without being shown, we are unable to get an accurate estimate of the actual best θ's, which lowers the utility gain in later steps. Showing topics with smaller estimated θ values is known as *exploration*. Therefore, we face a trade-off between exploitation and exploration to gain the optimal overall utility. This was also illustrated in the well-known multi-armed bandit problem [16].

To achieve both goals, we rank topics based on their expected utility plus a term related to their variances instead of solely using the expected utility as in the greedy approach. The term related to the variance is known as the *exploration bonus* [17]. In our case, given $\theta_i \sim B(\cdot|\alpha_i, \beta_i)$, the expected utility is $\alpha_i/(\alpha_i + \beta_i)$, and its variance is $\alpha_i\beta_i/[(\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)]$. We define the exploration bonus as the variance scaled by a weight parameter λ. Hence, the ranking score, a combination of the expected utility and the exploration bonus, is $\alpha_i/(\alpha_i + \beta_i) \times [1 + \lambda\beta_i/[(\alpha_i + \beta_i)(\alpha_i + \beta_i + 1)]]$. More detail derivations and examples are given in another technical report [18].

3 User Study

We carry out a user study experiment to evaluate the performance of our proposed recommendation strategy, *Exploitation and Exploration* (*E&E* in short), and compare it to two other baselines: *random*, which presents each topic, in a random order, the same number of times on average, and *greedy*, which ranks topics and presents topics based on their learned θ_i' respectively (i.e. $\alpha_i/(\alpha_i + \beta_i)$). The two baselines can be considered as two special cases of *E&E* that each focuses on one aspect respectively. Other than the user study, we do simulations to study the properties of each method. Their performance are similar to the findings from the reinforcement learning literature [16], and the details are described in the technical report [18].

In the user study experiment, we randomly select 45 categories (in level two of the hierarchy) from the Open Directory Project (ODP) [19]. Before each experiment begins, we ask the user to indicate, on a scale of one to nine, her interest level on each topic as the ground truth (θ_i) to measure the estimation accuracy. In each iteration of the experiment, the URL of seven Webpages, each coming from a different category, together with their titles and descriptions, are presented to the user. The user is instructed to click on the URLs that she feels interesting until no more is found to proceed to the next iteration. The click records are then used to update the α_i, β_i values of each topic respectively. We interleave three different strategies randomly throughout 75 iterations without informing the user, while each strategy updates its parameters independently for 25 iterations each. Such settings try to minimize any potential bias by comparing the strategies without dividing users into groups or dividing the test into phases. We have recruited ten users from staff members of NEC Labs and students of UCLA Computer Science department to participate in the experiment.

Click utility Figure 2 shows the click utility (as the fractional improvement of the number of cumulative clicks over *random*) averaged over 10 users. The behavior of each strategy is similar to the simulation, where *E&E* performs noticeably better than *greedy*.

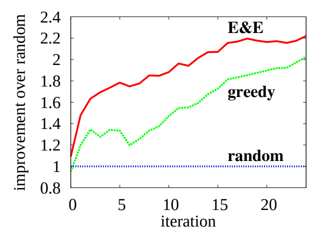

Fig. 2. Comparison of click utility of *E&E*, *greedy*, and *random*.

Estimation error of Θ We map the interest levels indicated by users to click probabilities (θ_i) by using $\frac{x-lb}{ub-lb}$, where x is the level selected, lb is the lowest level selected, and ub is the highest level selected by a user. We compute the normalized mean absolute error of the estimated θ_i's at the 25^{th} iteration of the experiment. The error values of *random*, *E&E*, and *greedy* are 21.6, 23.8, and 24.3, respectively. Their relative performance is similar to the prediction from simulation, however, the difference is less noticeable.

Interest drift We design an experiment in which after a user finishes all iterations of her test, a different user (having different interests) repeats the test using the learned $\alpha_i's$ and $\beta_i's$ as initial values. Such a switch of user simulates the scenario when a user has changed her interests at the end of the 25^{th} iteration. Figure 3 shows the average click utility (as the fractional improvement of the number of cumulative clicks over *random*) of 5 users, from the 26^{th} to 50^{th} iteration. The result clearly shows that *E&E* adapts to changes faster than *greedy* and improves the click utility towards the end.

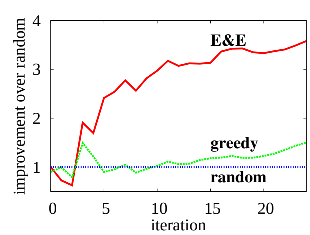

Fig. 3. Comparison of click utility of *E&E*, *greedy*, and *random* strategies under interest drift.

From the result of the user study, we conclude that our *E&E* algorithm outperforms the exploitation-only *greedy* algorithm in terms of click utility, parameter estimation error, and the rate of adaptation to user interest drift.

4 Conclusion and Future Directions

In this paper, we study how to effectively capture user interests in a personalized recommendation system. We propose a learning algorithm that uses both exploitation and exploration to captures user interests, represented as a probabilistic model, through an interactive recommendation process. We demonstrate, through simulations and user studies, that our algorithm achieves higher click utility, lower estimation error, and more agile adaptation to user interest drift against a random algorithm and a greedy algorithm. As suggested by the reviewers, two possible future research directions that make the user modelling more realistic include: a more complicated model that assumes a document belongs to multiple topics; introducing dependency and correlation among topics and recommendation items.

References

1. Jameson, A.: User modeling meets usability goals. In: Proceedings of UM. (2005)
2. Ziegler, C.N., McNee, S.M., Konstan, J.A., Lausen, G.: Improving recommendation lists through topic diversification. In: WWW'05. (2005)
3. Webb, G.I., Pazzani, M.J., Billsus, D.: Machine learning for user modeling. User Modeling and User-Adapted Interaction **11**(1-2) (2001) 19–29
4. Efthimiadis, E.N.: Interactive query expansion: A user-based evaluation in a relevance feedback environment. Journal of the American Society for Information Science **51**(11) (2000)
5. Kelly, D., Dollu, V.D., Fu, X.: The loquacious user: a document-independent source of terms for query expansion. In: SIGIR. (2005)
6. Mobasher, B., Cooley, R., Srivastava, J.: Automatic personalization based on web usage mining. Communications of the ACM **43**(8) (2000)
7. Mobasher, B., Dai, H., Luo, T., Nakagawa, M.: Discovery and evaluation of aggregate usage profiles for web personalization. Data Mining and Knowledge Discovery **6**(1) (2002)
8. Agichtein, E., Brill, E., Dumais, S.: Improving web search ranking by incorporating user behavior information. In: SIGIR. (2006)
9. Agichtein, E., Brill, E., Dumais, S., Ragno, R.: Learning user interaction models for predicting web search result preferences. In: SIGIR. (2006)
10. Qiu, F., Cho, J.: Automatic identification of user interest for personalized search. In: WWW'06. (2006)
11. Jin, X., Zhou, Y., Mobasher, B.: Task-oriented web user modeling for recommendation. In: Proceedings of UM. (2005)
12. Joachims, T.: Optimizing search engines using clickthrough data. In: SIGKDD. (2002)
13. Kelly, D., Belkin, N.J.: Display time as implicit feedback: understanding task effects. In: SIGIR. (2004)
14. Sugiyama, K., Hatano, K., Yoshikawa, M.: Adaptive web search based on user profile constructed without any effort from users. In: WWW'04. (2004)
15. Shen, X., Zhai, C.: Active feedback in ad hoc information retrieval. In: SIGIR. (2005)
16. Gittins, J.C., Jones, D.M.: A dynamic allocation index for the sequential design of experiments. In *et al*, J.G., ed.: Progress in Statistics. Volume I. North-Holland, Amsterdam-London (1974) 241–66
17. Sutton, R.S.: Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: ICML '90, San Mateo CA, Morgan Kaufman (1990) 216–224
18. Sia, K.C., Zhu, S., Chi, Y., Hino, K., Tseng, B.L.: Capturing User Interests by Both Exploitation and Exploration. Technical report, NEC Labs America (2006)
19. Netscape: Dmoz open directory project http://www.dmoz.org.