User Action Interpretation for Online Content Optimization

Leena Dhangar
ME (Computer)
University of Pune
SITRC, Mahiravani, Nasik, India

A. D. Potgantwar
Assistant Professor
HOD, Computer Department
SITRC, Mahiravani, Nasik, India

ABSTRACT
Search engine advertising has become a significant element of the Web browsing experience. Choosing the right items for the query and the order in which they are displayed greatly affects the probability that a user will see and click on each item. This ranking has a strong impact on the revenue the search engine receives from the clicking on advertisements. Displaying the items to the user that they prefer to click on improves user satisfaction. Therefore, it is important to be able to accurately estimate the click-through rate of ads in the system. So the user's experience depends crucially upon the quality of content recommendations. This paper presents an overview of the content recommendation, namely how to recommend a small set of items to a user from an underlying pool of content items according to user's interest. Therefore, we build an online learning framework for personalized recommendation using recommender system. This paper focuses on an approach of interpreting users' actions for the online learning to achieve better item relevance estimation. So that User is provided with the content in which he is interested. And finally the items are ranked according to the user's interest based on the click through rate (CTR).

Keywords
Content Recommendation, Recommender system, Click Through rate (CTR)

1. INTRODUCTION
THE web has become the central distribution channel for information from traditional sources such as news outlets as well as rapidly growing user-generated content. Developing effective algorithmic approaches to deliver such content when user visits web portals is a fundamental problem. Search engines use ranking algorithms to return the most relevant links in response to a user's keyword query. Whereas, portals that cater to users who browse a site are typically programmed manually. This is because content is harder to assess for relevance, topicality, freshness, and personal preference; there is a wide range in the quality; and there are no reliable quality or trust metrics.

We consider the problem of optimizing content displayed in a module that is the focal point on a major Yahoo! portal; the page also provides several other services such as latest news.

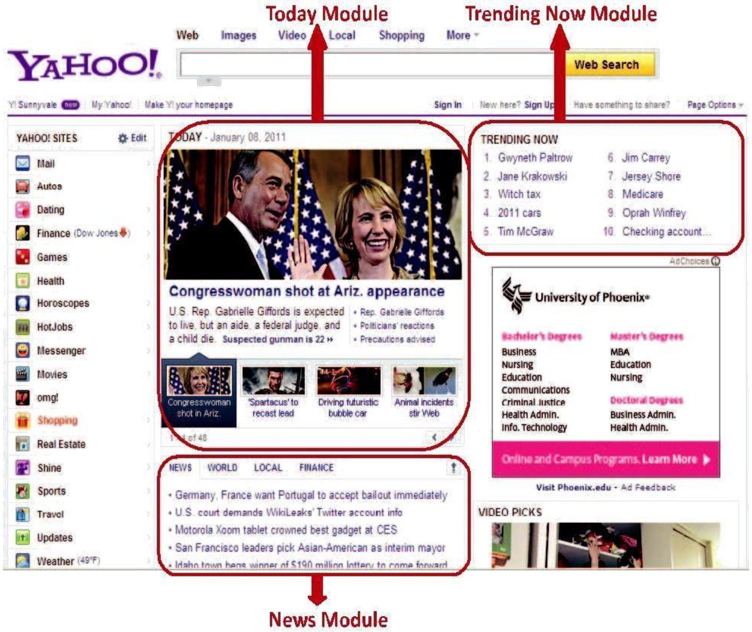

Figure 1: Snapshot of Yahoo! Front Page

Figure (1) shows a panel with slots labelled Today Module (F1), News Module (F2), and Trending Now Module (F3). Slot F1, which accounts for a large fraction of clicks is prominent, and an article is displayed on F1 receives many more clicks than when it is displayed at F2, F3. The available articles are refreshed continually. A few new articles get pushed into the system periodically (every few hours) and replace some old articles. The editors [2] keep up with important new stories (e.g., breaking news) and eliminate irrelevant and fading stories, and ensure that the pool of articles is consistent. The same articles are seen by all users visiting the page. We consider how to choose the best set of articles to display on the module to a given user. Since the mix of content in the available pool already incorporates constraints, we focus on choosing articles to maximize overall click-through [3] rate (CTR), which is the total number of clicks divided by total number of views for a time interval. But each visit generates a "view" event on the Today Module, though the visitor may not pay any attention to that module.

One of our goals is to increase user activities, measured by CTR. We have to rank available articles according to visitors' Interests, and to highlight the most attractive article at the First position. Articles with the highest segmental CTR will be served to user segments respectively. Once we identify users who share similar interests in conjoint analysis, predictive models can be built to classify users (including new visitors) into segments.

2. RELATED WORK

Internet is continuously growing at an exponential rate. So it becomes difficult and time consuming to find the meaningful information in which user is interested. So to remove the unwanted information and find what is really useful for the users is a challenging problem for information filtering. Previous studies focus on offline recommendation model. But it was not good enough to change according to users' interest. Also offline models cannot be updated according to these changes. Therefore we propose an online learning framework for personalized recommendation.

Recommender systems [4] have proven to be an effective method to deal with the problem of information overload in finding relevant products. They not only help to provide relevant items to an individual user but also increase cross-sell by suggesting additional products to the customers and improve consumer loyalty because consumers tend to return to the sites that better serve their needs. Previously human editors were used to manually rank the items according to the users' interest. But it requires expensive human effort. So, various kinds of online recommendation techniques have been proposed, including collaborative filtering (CF) and content based filtering.

A content-based filtering [8] monitor a document stream and pushes documents that match a user profile to the corresponding user. The user may read the delivered documents and provide feedback, which is then used to update the user's profile.

Collaborative filtering (CF) is currently one of the most popular and widely used recommendation techniques.

Collaborative filtering [6] is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). Collaborative filtering [6] approach is that if one person has the same opinion as another person on an issue, then first person is more likely to have opinion like second person on a different issue x than to have the opinion on x of a person chosen randomly.

Collaborative Filtering [7] has many advantages. First, in collaborative filtering, humans determine the relevance, quality, and interest of an item in the information stream. Consequently, filtering can be performed on items that are hard to analyze.

Second, collaborative filtering [5] systems can enhance information-filtering systems by measuring, in dimensions. Humans are capable of analyzing on dimensions such as quality or taste, which are very hard for computer processes Thus collaborative filtering [7] systems can enhance information-filtering systems by measuring, in dimensions beyond that of simple content, how well an item meets a users' need or interests.

Thus online recommendation [6] model is proposed which will rank the items according to the users' interest based on the click through rate (CTR).

2.1 Problem Analysis

Consider the problem of optimizing content displayed in a module that is the focal point on a major Yahoo! portal; the page also provides several other services (e.g., Mail, Weather) and content links. The module is a panel with slots labeled Today Module, Trending Now Module, News Module which accounts for a large fraction of clicks, is prominent. The pool of available articles is refreshed continually. A few new articles programmed by editors get pushed into the system periodically and replace some old articles. So providing the relevant content to the user from large content of pool is a big challenge.

So we focus on users' rating i.e. click-through rate (CTR), which is the total number of clicks divided by total number of views for a time interval. And according to it, content is provided to the user.

3. PROPOSED ARCHITECTURE

A new content publishing system is proposed that selects articles for the users, choosing from the content pool that is frequently refreshed and liked by the users. Otherwise the most likely items are displayed at the bottom of the web page and less likely items are placed at top of the web page.

So, the fundamental problem we must solve is to quickly identify which items are popular. We must also explore the underlying pool constantly to identify alternatives, quickly discarding old stories. Our proposed approach is based on per article performance in near real time through online models. So that users' interest can be estimated using click through rate (CTR).

Therefore, Online Learning Framework [5] for personalized content recommendation is proposed which rank the items by automatically estimating the users' interest. Such recommender system has three characteristics as follows:

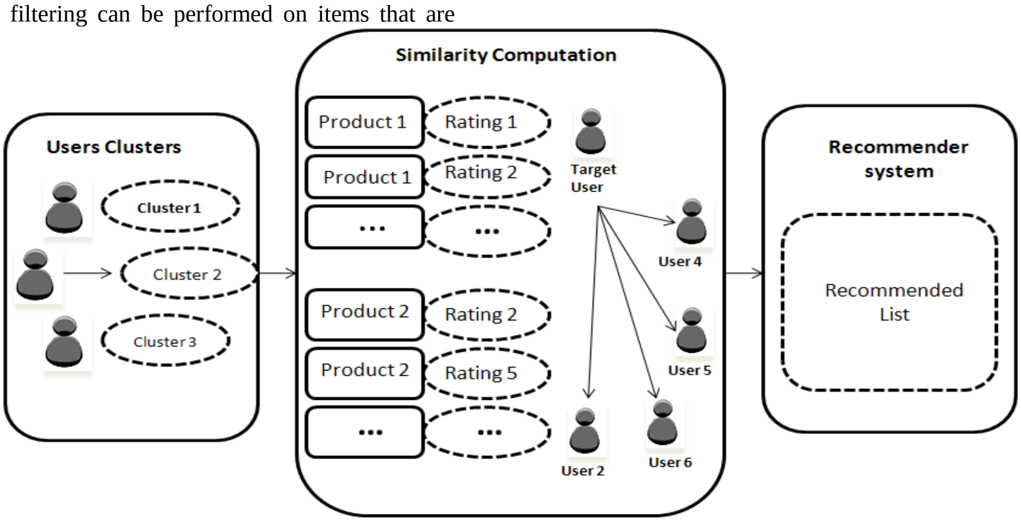

Figure 2: Framework of Recommendation Model

3.1 Online Learning

Online Learning methodology [10] is necessary to provide attractive and relevant content to users. It enables us to model users' behaviors (clicks and views) to update recommendation model in real time. In an online Learning framework random learning bucket is used for estimating CTR. All the models are updated after some time interval. Finally updated models are applied to the bucket and items are displayed to the user in descending order.

3.2 Dedicated Model

It is an effective and dedicated model [10] for each content item which estimates the relevance score. To apply per item model in our online learning framework it is necessary to use particular method for updating per item models so that the items are ranked by the recommendation scores in the descending order.

3.3 Personalization

The users are divided into different groups according to the user profile. And recommender system serves them with the models which are updated. Thus, online recommendation model [11] heavily relies on users' clicks and views which are required for estimating recommendation list. So that content is displayed to the user according to their interest. The top ranked are displayed at the first position and less likely items are displayed at the bottom of the web page

4. FRAMEWORK OF RECOMMENDATION MODEL

Figure (2) shows the framework of recommendation model.

The objective of recommender system is to provide the relevant content to the users by filtering unwanted information. There are three stages of recommender systems.

4.1 Clustering

Clustering [13] is nothing but grouping the items together that are having similar characteristics. So as shown in fig. (2) Users that are having similar interest are grouped together. According to their interest same recommendation list is provided to the users.

4.2 Similarity Computation

When the products or items are displayed to the user, the ratings by the user are taken into consideration. In the process of similarity computation [20] , according to the ratings highly ranked items are first displayed to the user.

4.3 Recommendation

Recommendation is the last stage of recommender system.

In this stage highly ranked items in which user is interested are recommended to the user.

5. PERFORMANCE ANALYSIS

As shown in Fig. 1, there is usually more than one content recommendation module on portal website. Different content modules are likely to compete with each other on a densely packed interface such as the frontpage of the website. Therefore, one user visit on the portal website may not necessarily mean the user is really engaged in all the content modules displayed to the user [1].

Accurate CTR estimation is necessary for recommendation module [14] and it should be based on the events where users were really engaged in this module, instead of all the events where the contents of this module were merely displayed to the users.

The purpose of click through rates [19] is to capture customers' initial response to websites. We can estimate the click through rate as follows:

$$\text{Click through rate (\%)} = \text{Click (\#)} / \text{Impressions (\#)} \qquad (1)$$

Thus, click through rate (CTR) [5] is the number of clicks on particular item to the number of times the particular item is displayed to the user.

Click through rate can be calculated by predicting ranking to each item by the user.

Table 1 Ranking Table for Each Segment

User	Segment	Predicted Ranking
User 1	Sports	5
	Finance	2
	Politics	1
User 2	Sports	2
	Finance	1
	Politics	3

Table 1 shows ranking of each user for every segment. Thus politics segment has first position for user 1 and finance is ranked at first position for second user. In this way ranking is calculated for each user.

6. CONCLUSION

Nowadays the internet is growing exponentially at a fast rate.

When any user searches for information about any topic, large amount of information becomes available to the user which may contain useful as well as unwanted information. So it becomes difficult to extract the meaningful information according to the users' interest from all the available information. To overcome this difficulty recommender systems are proposed. A recommender system satisfies the users' demands by providing relevant content to the users. This paper presents user action interpretation for online personalized content optimization which uses two filtering approaches content based filtering and collaborative filtering. These two filtering approaches filter the data according ti the users' interest.

In the future, more interest is in exploring more information about the users' geographic location as well as clicks behaviour and makes the use of that information to recommend relevant content to the user.

7. ACKNOWLEDGMENT

I would like to express my special appreciation and thanks to my advisor Prof. A. D. Potgantwar, you have been a tremendous mentor for me.

8. REFERENCES

[1] Jiang Bian, Anlie Dong, Xiaofeng He, Srihari Reddy, Yi Chang, "User Action Interpretation for Online Content Optimization, VOL.25, No. 9,2013.

[2] L. M. de Campos, J. M. Ferndez-Luna, J. F. Huete, and M. A. Rueda Morales, "Combining Content Based filtering and Collaborative Recommendations: A Hybrid Approach Based on Bayesian Networks ",Int'l

Approximate Reasoning vol. 51, no.7, pp. 785-799, 2010.

[3] D. Agarwal, B- C. Chen, and P. Elango, "Fast Online Learning through Offline Initialization for Time Sensitive Recommendations" Proc. 16th ACM SIGKDD Int'l Cont. Knowledge Discovery and Data Mining (KDD), 2010.

[4] D. Agarwal, B- C. Chen, and P. Elango, "Spatio-Temporal Models for Estimating Click- Through Rate" Proc. 18th Int'l World Wide Web Cont. (WWW), 2009.

[5] Agarwal, B- C. Chen, P. Elango, N. Motgi, S-T. Park, R Ramakrishnan, S. Roy, and J. Zachariah," Online Models for Content Optimization." 2005.

[6] G. Buscher, L. vainest and Dengel, "Segmentation Level Display Time as Implicit Feedback: A Comparison to Eye Tracking" , Proc. 32nd Ann. Int'l ACM SIGIR Cont. Research and Development in Information Retrieval (SIGIR), 2009.

[7] W. Chu and Z. Ghahramani, "Probabilistic Models for Incomplete Multi- Dimensional Arrays", Proc. 12th Int'l Conf. Artificial Intelligence and Statistics, 2009.

[8] W. Chu, S. T. Park, "Personalized Recommendation on Dynamic Content Using Predictive Bilinear Models ", Proc. 18th Int'l World Wide Web Conf. (WWW), 2009.

[9] W. Chu, S. T. Park, T. Beaupre, N. Motgi and A. Phadke, "A Case Study Of Behaviour- Driven Co joint Analysis on Yahoo! Front page Today Module", Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD), 2009.

[10] D. Downey, S. Dumais, D. Liebling, and E. Horvitz, "Under-standing The Relationship between Searchers' Queries and Information Goals," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM), 2008.

[11] J. Ahn, P. Brusilovsky, J. Grady, D. He, and S.Y. Syn, "Open User Profiles for Adaptive News Systems: Help or Harm?" Proc. 16th Int'l World Wide Web Conf. (WWW), 2007.

[12] Z. Guan and E. Cutrell, "An Eye Tracking Study of the Effect of Target Rank on Web Search," Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), 2007.

[13] A. Broder, "Computational Advertising and Recommender Systems", Proc. Second ACM Int'l Conf. Recommender System (RecSys), 2007.

[14] D. Billsus and M. Pazzani, "Adaptive News Access- the Adaptive web Methods and Strategies of web Personalization", Springer-Verlag 2007.

[15] E. Gabrilovich, S. Dumais, and E. Horvitz, "Newsjunkie: Providing Personalized Newsfeeds via Analysis of Information Novelty," Proc.13th Int'l World Wide Web Conf. (WWW), 2004.

[16] L.A. Granka, T. Joachims, and G. Gay, "Eye-Tracking Analysis of User Behaviour in WWW Search," Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2004.

[17] R. Jin, J.Y. Chai, and L. Si, "An Automatic Weighting Scheme for Collaborative Filtering," Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2004.

[18] A. Das, M. Datar, A. Garg, S. Rajaram, "Google News Personalization Scalable Online Collaborative Filtering", Proc. 16th Int'l World Wide Web Cont. (WWW), 2004.

[19] J.L. Herlocker, J.A. Konstan, A. Borchers, and J. Riedl, "An Algorithmic Framework for Performing Collaborative Filtering," Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 1999.

[20] T. Hofmann and J. Puzicha, "Latent Class Models for Collaborative Filtering," Proc. 16th Int'l Joint Conf. Artificial Intelligence (IJCAI), 1999.