International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Recommendation System Based on Behavioral Variability of User

Neethu Raj

Lecturer in Computer Science, University Institute of Technology, Mukhathala, Kollam, Kerala, India

Abstract: ***In real world web applications content recommendations are the most important thing to get better response to a user. Most of all users have less time span to search content in this busy environment. So they always seek improved result with minimum delay. Optimized content recommendation provides good content delivery. User actions and user feedback play a vital role in recommender systems. User feedback may be implicit user feedback or explicit user ratings on the recommended items. Appropriate user action interpretation is critical for a recommender system. This paper builds an online learning framework for personalized recommendation. The main contribution in this paper is an approach of interpreting users' actions for the online learning to achieve better item relevance estimation.***

Keywords: Action interpretation, content optimization, personalization, recommender systems

1. Introduction

The recent development in web technologies helps web users to gather recent information. Due to the explosive growth of social network websites and online user generated content systems personalization is a desirable feature for each website [1]. While browsing on a portal website, time is less and amount of data resulting from search engine is more. To optimize content delivered to a user most of the search engines use content recommendation. The software tools and techniques providing suggestions for items to users are called content recommendation. Recommender systems are a subclass of information filtering system which predict the users rating or preference on an item. The suggestions given by the recommendation system are based on user's decision making process etc.ss like what to buy, what to read, what to watch. The term content item is used to denote what the system recommends to a user. There are multiple content vendors to provide more information to each user, and there are plenty of contents .But in these busy world web users usually has short attention spans while browsing for information. So to deliver right content to each user it is necessary to optimize the content. This content optimization is done by identifying which one is the most attractive content for each person. Normally human editors select a set of content items from a candidate pool to present to a user. This selection can avoid low quality contents, but human effort is expensive. When there is a large pool of candidate items, editorial selection does not guarantee personally relevant contents are recommended to the user. So a proper content optimization scheme is necessary. Due to the explosive growth of social network websites, personalization is a desirable feature for each website to provide right content to each user by identifying their preferences. This property is called personalization. Personalized content recommendation is a process of collecting details about website users and analysing their current and past user behaviour then based on this analysis recommend better content to each user.

In recent years recommendation systems are used everywhere. The main characteristic of a recommendation system is that they draw the interest of user and provide better response to a user. The main application areas are movies, news, books, research articles, search queries and products like dress, fancy, electronic gadgets etc [1]. And also there are recommendation systems for restaurants, life insurance, financial services, jokes, experts, twitter followers etc. Personalized recommendation approaches are of two types, content based filtering and collaborative filtering. In content based filtering a profile is generated for a user based on the content descriptions of the content items that are previously rated by the user. Drawback of this method is its limited capability to recommend content items that are different than those previously rated by the users. So the most widely used technique is Collaborative filtering, in which users ratings are analyzed and recommend items by leveraging preferences from other users with similar tastes.

But this method has a disadvantage called cold start problem. Hybrid recommendation methods avoid some of these weaknesses and still have some problems with it.

2. Problem Statement

The primary objective of recommender system is to avoid information overloading problem. Users prefer the fast retrieval of data while browsing the portals. Content recommendation helps to overcome information overloading problem by choosing the best matching contents. Thus users get better response without wasting their time. Most of the recommendation systems face various difficulties while identifying high quality items and providing this to users.

This problem statement clearly states that,

- An online recommender system is needed with dynamic content uploading and so data should be up-to-date.
- Recommendation should be based on user taste, url rank, topic frequency, similarity with other user etc.
- Personalization should be based on age, gender etc

3. Related Works

This work [1] focuses on building an offline recommendation model and proposes an online learning framework for personalized recommendation. Also leverage

Licensed Under Creative Commons Attribution CC BY

user behavior information to combine the two techniques. In particular, this work applies user action interpretation to model relevance feedback used in content-based filtering. And employ user behavior-based segmentation, which follows the direction of collaborative filtering, to improve the effectiveness of the content recommendations.

3.1 User Segmentation

User segmentation for personalization is due to the fact that the proposed Clustering algorithms actually group users by interests and preferences that are implicitly demonstrated by their behaviors. Once the interest patterns are determined by clustering algorithms, a user will be assigned to a segment by her profile features. Fortunately, user profile features also highly correlate with behaviors and interests. Thus, the user segment assignment is usually reliable except when the user is new to the site so that her profile features are poor.

Table 2.1: User Segmentation Based on Demographic Features

Segment	Age Range and gender
f-u20	10 < age <= 20, gender = female
f-u40	20 < age <= 40, gender = female
f-u80	40 < age <= 80, gender = female
m-u20	10 < age <= 20, gender = male
m-u40	20 < age <= 40, gender = male
m-u80	40 < age <= 80, gender = male
Unk	Unknown age or gender

3.2 Architecture

Fig 1 A random learning bucket is used for exploration purpose. At the end of each time interval, the model for each candidate item is updated based on the users' clicks and views in random learning bucket during this time interval. In the next time interval, the updated model is applied to the corresponding candidate item in the serving bucket. In this way, all the candidate items are displayed by ranking scores (computed by their corresponding updated models) in the serving bucket.

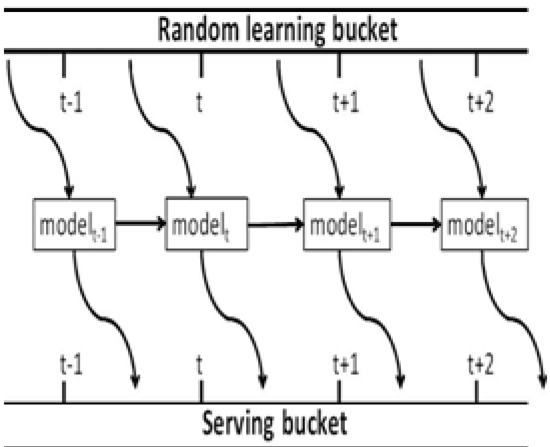

Figure 1: Online learning flowchart

3.3 Disadvantages

This work handles static data which added simultaneously by the admin but it create some problems. The administrator need to be update data like news, videos etc accurately and simultaneously. Hence both the content pool and users' interests change very frequently, and offline models cannot be updated according to such changes very efficiently. This work uses a technique called user segmentation for personalization. The user segment assignment is usually reliable except when the user is new to the site so that her profile features are poor. In this work the features that used for personalization is age of the user. The optimized result may be deviated from the actual result due to single feature. So it is better to enhance the paper by adding some modifications and which will improve the system.

4. Proposed Method

In real world web applications content recommendations are the most important thing to get better response to a user. Most of all users have less time span to search content in this busy environment. So they always seek improved result with minimum delay. Optimized content recommendation provides good content delivery. User actions play a vital role in recommender systems. Hence for a recommender system accurate interpretation of user action is important. This work builds an online learning framework for personalized recommendation. We design a recommender system based on historical user activity. There is a browsing system which records each user activity and estimates their rank. Based on the rank top ranked contents were recommended to the user. The main contribution in this work is an approach of interpreting user's actions for the online learning to achieve better item relevance estimation.

To address all the challenges of previous works this work introduces a personalized recommender system as a web browser. To achieve personalized recommendation this work first categorizes users into diverse groups based on their interests as modeled by user action information, and then serve users in each group with recommendation modeled by user actions of those users in the same group.

4.1 Recommender System

Recommender systems have become extremely common in recent years, and are applied in a variety of applications. [9]The most popular ones are probably movies, music, news, books, research articles, search queries, social tags, and products in general. Recommender systems typically produce a list of recommendations in one of two ways - through collaborative or content-based filtering. Collaborative filtering approaches building a model from a user's past behavior (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users; then use that model to predict items (or ratings for items) that the user may have an interest in. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend additional items with similar properties.

4.2 Rank calculation

Recommendation systems mainly based on ranking system. In this system, ranks were calculated for various url links. Rank calculation is affected by some factors as follows,

Licensed Under Creative Commons Attribution CC BY

4.2.1 URL Rank

URL is an acronym for Uniform Resource Locator and is a reference to a resource on the Internet that specifies the location of the resource on a computer network and a mechanism for retrieving it. A URL has two main components: Protocol identifier: and Resource name. User login to the site and can search any contents on the search page. URL corresponding to the search is stored in the database. When users search the same URL more times url rank increases. Rank calculation is based on the users search frequency.

4.2.2 Domain rank

A domain name is an identification string that defines an administrative authority. When a particular item is searched the same item may include in various domains. Some can have more priority than others so domain should also have importance in ranking. Domain rank increases with increase in url rank.

4.2.3 Topic rank

Rank is calculated for each topic on the search. Topic rank increases with url rank.

4.2.4 User-Topic affinity

For obtaining user taste in each area first we find out user topic relation.

4.2.5 Semantic Relation

Similarity of a user taste with other users can be obtained by calculating semantic relationship. This means relationship between selected words of that user with words on other selected sites.

4.2.6 Normalizing values

Rank calculation is an important process to recommend top ranked content to the user. So the value we calculated must be accurate. Hence calculated ranks must be normalized such that redundancy can be avoided and efficiency can be improved.

Normalized value=Actual value/Maximum value

Normalized value means a percentage of corresponding value related to maximum.

4.3 Personalized Recommendation

Personalization is a desirable feature where url links are personalized based on the characteristics of an individual. There are three categories of personalization: Profile / Group based, behavior based and collaboration based. Profile based systems use user features obtained from their profile for recommendation. In behavior based systems user behavior is examined. Collaboration based systems are more advantageous than others. Collaboration based approaches builds a model from a user's past behavior as well as similar decisions made by other users; then use that model to predict items that the user may have an interest in.

When a user login to the site a search history is created based on their search and personalization is from that history. For each user url recommendation is different because url rank is based on each users preference for that content.

4.4 User Segmentation

User segmentation [1] is important for personalization which means, group users by interests and preferences that are implicitly demonstrated by their behaviors. User information includes user id, name, age, sex, age group etc. Users are categorized by their gender or age groups. Based on the age users are placed into four groups as,

Table 3.1: User Segmentation based on age.

Group	Age
1	age below 20
2	Age between 20 and 35
3	Age between 35 and 50
4	age above 50

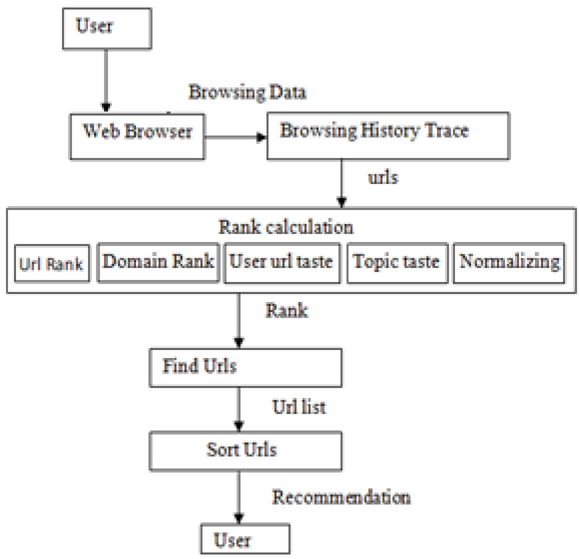

Figure 2 Flow Chart of proposed work

4.5 Algorithm

1) Read all users
   If a new user, then register to the site and then login.
   Else
   User Login
   When a new user registers details were stored into the database. Then read all users from database.
2) Read all urls
   When user enters a search item its url is stored into the database from where it can be read.
3) Calculate url rank.
4) Recommend top ranked urls to the user who login to the site.

4.5.1 Rank calculation Algorithm

1) Initialize 2D matrices ‘rank’ and ‘N-rank’ with total number of rows equal to all url size and total number of columns equal to 8.
2) Get user name uu from user database userinfos.
3) Get urls ui corresponding to the user from url database urlinfo.
4) Select a count v for usage count of url ui by the user uu. It denotes how many times user search that url.
5) Find rank for all users as Ranks[i][0]=Ranks[i][0]+v;
6) For female users set Ranks[i][1]=Ranks[i][1]+v;
7) For male users set Ranks[i][2]=Ranks[i][2]+v;

Licensed Under Creative Commons Attribution CC BY

8) For each age group set Ranks[i][uu.agegroup+3]=Ranks[i][uu.agegroup+3]+v;
9) Find normalized value as, N-Rank[i][j]=rank[i][j]/mx; where mx is the maximum value in that column.
10) Then select maximum rank from each column and place it in top position.

4.5.2 Url Display Process

Admin can categorize urls into url part and domain part.

1) Get all urls from the database.
2) Each links has two parts url part and domain part.
3) Create a table for storing url and its domain.
4) When a url is clicked its details can be viewed as a table with columns userid and coun1t.
5) Which shows how many times the user searches that url.
6) There is a table called word list which stores words which were searched by the user.

4.5.3 Usage Details Form

After categorize urls then there is a view details option and can view the usage details of url. For each user id usage count is calculated. And the words which were searched by user is also displayed as a word list. Usage count is essential for final rank calculation. Then create a table with columns user id, count and add each row when user search any content.

4.5.4 Final Rank calculation Algorithm

1) Set priority values for variables like,
2) UserToLinkRank=2,AgePrefernece=2,SexPrefernece=2, UserTopicTaste=3
   Here higher priority is set for user topic taste than others.
3) Set maximum number of urls that are recommended to the user as max_urls=10
4) For each user who login to the site , find user-word taste, user-url taste, user-age-url taste, user-sex-url taste
5) Then find Best urls as follows,

FindBestUrls ():

i. Initialize array Finalrank [] with size equal to all url size.
ii. For each i from 0 to array length find FinalRank[i]=UserToLinkRank*User_Url_taste[i]+Age_Url_Taste[i]*AgePrefernece+Sex_Url_Taste[i]*SexPrefernece+UserTopicTaste*User_Word_Taste[i];
iii. Set a new vector ‘sorted’ to store sorted urls.
iv. Then select max_urls from sorted vector for recommendation.

1) FindUserUrlTaste()
   i. Initialize array User_Url_taste[] with size equal to all url size.
   ii. for(int i=0;i<Algorithm.User_Url_Rank.Ranks.length;i++)
   iii. {
   iv. User_Url_taste[i]=Algorithm.User_Url_Rank.N_Ranks[i][0];
   v. }
   Normalize value for each url and set the value in the array.
2) FindUserSexUrlTaste()
   i. Initialize array Sex_Url_taste[] with size equal to all url size.
   ii. For male user se=1 and for female user se=2 set value as,
   iii. for(int i=0;i<Sex_Url_Taste.length;i++)
   iv. {
   v. Sex_Url_Taste[i]=Algorithm.User_Url_Rank.N_Ranks[i][se];
   vi. }
3) FindUserAgeUrlTaste()
   i. Initialize array Age_Url_taste[] with size equal to all url size.
   ii. for(int i=0;i<Age_Url_Taste.length;i++)
   iii. {
   iv. Age_Url_Taste[i]=Algorithm.User_Url_Rank.N_Ranks[i][se+2];
   v. }
   Where se denotes the age group of user.
4) FindUserWordTaste()
   i. Initialize array User_Word_taste[] with size equal to all url size.
   ii. Set two variables uw and urlw for storing user word rank and url word rank respectively.
   iii. for(int i=0;i<Algorithm.all_urls.size();i++)
   iv. {
   v. for(int j=0;j<uw.AllWords.size();j++)
   vi. {
   vii. User_Word_Taste[i]=User_Word_Taste[i]+(uw.N_Ranks[j][Me_Position]+1)*urlw.N_Ranks[j][i];
   viii. }
   ix. }
   Here priority is more for user word rank than url word rank.
5) Then send url list in sorted order.

5. Conclusion

This work propose an approach for interpreting users actions in terms of click or browsing then recommend contents based on the user preference. For each user a user profile is created then user action history is developed. Recommendation is based on user preferences obtained from user profile. In previous systems content is static that is news added by the administrator. It will take more time and want to update periodically so this work introduce dynamic content uploading. In this system user get recommendation from dynamic content which is collected when user search contents on the web browser. Rank is calculated by examining the factors such as url rank, domain rank, topic rank, user taste, semantic relations, user topic affinity etc. First we set some priority values and then calculate the rank. Then store the urls in to a sorted array based on the rank and recommend top ranked content. For each url, its relationship with user is calculated and affinity of topic with other users also calculated. User’s age group, gender and click behaviors also used for proper recommendation. For each recommendation user can view the popularity of the content through the chart view. In the future this work can be improved by adding more features on rank calculation like geo location. And hence ranking will be more accurate.

Licensed Under Creative Commons Attribution CC BY

References

[1] Jiang Bian, Anlei Dong, Xiaofeng He, Srihari Reddy, and Yi Chang, "User Action Interpretation for Online Content Optimization", IEEE transactions on knowledge and data engineering, vol. 25, no. 9, year 2013
[2] D. Agarwal, B.-C. Chen, P. Elango, N. Motgi, S.-T. Park, R.Ramakrishnan, S. Roy, and J. Zachariah, "Online Models for Content Optimization," Proc. Neural Information Processing Systems Foundation Conf. (NIPS), 2008.
[3] D. Agarwal, B.-C. Chen, and P. Elango, "Spatio-Temporal Models for Estimating Click-Through Rate," Proc. 18th Int'l World Wide Web Conf. (WWW), 2009.
[4] E. Agichtein, E. Brill, and S. Dumais, "Improving Web Search Ranking by Incorporating User Behavior Information," Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.
[5] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay,"Accurately Interpreting Clickthrough Data as Implicit Feedback," Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2005.
[6] P. Melville, R.J. Mooney, and R. Nagarajan, "Content-Boosted Collaborative Filtering for Improved Recommendations," Proc. 18th Nat'l Conf. Artificial Intelligence (AAAI), 2002.
[7] 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.
[8] J. Liu, P. Dolan, and E.R. Pedersen, "Personalized News Recommendation Based on Click Behavior," Proc. 15th Int'l Conf. Intelligent User Interfaces (IUI), 2010.
[9] Neethu Raj, Suja Rani M S, "An Overview of Content Recommendation Methods," Vol.3 Issue 1 on IJIRCCE Jan.2015
[10] http://en.wikipedia.org/wiki/Web_browser
[11] http://en.wikipedia.org/wiki/Domain_name
[12] http://en.wikipedia.org/wiki/Uniform_resource_locator

Licensed Under Creative Commons Attribution CC BY