

Personalized News Recommendation System Based on Consumers' Click Behavior

Aarti Gaikwad¹, Geeta Narle², Mohini Malge³, Namrata Mulay⁴, Noorjahan Sayyed⁵

Department of Information Technology
1, 2, 3, 4, 5 D. Y. Patil College Of Engineering Akurdi, Pune

Abstract-The development of the Internet, more and more people read news on the Internet. However, facing the massive data, consumers need to spend much time on choosing news they prefer. Therefore, providing an effective personalized news recommendation for consumers will promote consumer's experience of browsing news greatly. The news browsing sequence of a consumer can be obtained from the consumer's click behaviour on the Internet. Here, some potential associations between news using the news browsing sequence of a consumer will be found. Then, personalized news recommendation for different consumers can be provided according to these potential associations. In this paper, an improved personalized news recommendation algorithm based on consumers' click behaviour is proposed. Through doing experiments on real news browsing data, the recommendation result is better and the new algorithm is proved to be feasible.

Keywords-Recommendation , Browsing , Personalised, Consumers

I. INTRODUCTION

In this project we are going to create one web application which will show news in portal where users will come and will be using it. Behind the scene all user's behavior will be tracked by using logging mechanism (server logging). i.e. if user is clicking some news which he was not interested earlier, we will write some algorithm to conclude user's current interest Modules in this project .

- 1) Users module this will store user personalized profile which will in showing some recommended news on portal when he log in
- 2) Newz module this module will have all information related to newz management (admin module can create different newz here). the newz will be related to each other by parent-child hierarchy
- 3) Analyzer module this module will be core module which is responsible for retrieving useful & analyzed information from raw information. With the development of the Internet, more and more people read news on the Internet.

However, facing the massive data, consumers need to spend much time on choosing news they prefer. Therefore, providing an effective personalized news recommendation for

consumers will promote consumers' experience of browsing news greatly. The research on the personalized recommendation system can be roughly divided into four kinds, namely recommendation based on content, recommendation based on collaborative filtering, recommendation based on knowledge and hybrid recommendation. The recommendation based on content is a method which is widely used in practice. This method finds the relation between different news according to the content. Then, the preference of consumers can be obtained and applied to news recommendation. As text processing technology gets more sophisticated, the research on the recommendation based on content is becoming more and more mature. The recommendation based on collaborative filtering is a method which is widely studied in recommendation system. This method has two branches, namely collaborative filtering based on items and collaborative filtering based on consumers. The collaborative filtering based on items obtains similarity of different items from their scores provided by same consumers and make recommendations according to the similarity of items.

II. IDENTIFY, RESEARCH AND COLLECT

IDEA

- 1) Why Recommendation is important?

Recommender engines (REs) also known as recommender systems are software tools and techniques providing suggestions to a user. The suggestions provided are aimed at supporting their users in various decision making processes such as what items to buy, what music to listen, what profiles to browse, or what news to read. This thesis studies the feasibility of the integration of a recommender engine as a module in a News portal, and shows the process of its design and implementation using the Apache Mahout library. As such our work tackles two major problems which are:

- (1) the implementation of the recommender engine using the Apache Mahout library, and
- (2) the integration of the recommender in News portal.

Prior to the design of the application and the decisions made at the design stage, are shown the analysis of the requirements for the recommender system in the context of the business. The decisions made at design time are explained and the risks involved are analyzed and mitigated. Finally conclusions regarding the integration of the recommender engine in News portal are made.

With the development of network information technology, a lot of news comes into the view of Internet users. We have entered the "information overload" era. So how to find useful information becomes more and more important. Personalized news recommendation is a kind of technology to find the news that users want to get urgently. It is based on the browsing history of many users. Through analysis of their interests, we can realize the personalized news recommendation for different people.

III. STUDIES AND FINDINGS

(1) Data Filtering

There are many computer cache data, and we can find some messy code that don't meet our requirements or are useless for us. So we need to filter out this part of data. It can reduce a lot of unnecessary troubles and make our following work more effective.

(2) Data Preprocessing

Any original data in the computer must adopt a specific mathematical model to represent, and Internet user's cached data is also no exception. After analyzing the data, we know that each record includes User ID, News ID, Vis time, Title, Content and Re time. We need to process the title and content when analyzing the news in Chinese, such as segment words, remove the stop words. If we analyze news in other languages like English, we don't need to process it because it can separate from each other by space. Of course, the original data contains unqualified data, such as noisy data.

(3) Data Analysis

After analyzing the data, we know that each record includes User_ID, News_ID, Title, Content. At the same time, we generally know the users' browsing habits, such as the time of users typically browse news, the total number of news each user browses in this month, the total number of each piece of news browsed by users in a month and so on.

Project Consist Of Three Modules-

A. User Interface Module

B. News Module

C. Analyzer Module

A. User Interface Module

This will store user's personalized profile which will in showing some recommended news on portal when user login.

B. News Module

This module will all information related to news management (Admin module can create different news here). The news will be related to each other by parent child hierarchy.

C. Analyzer Module

This module will be core module which is responsible for retrieving useful and analyse information from raw information

IV. CONCLUSION

In this paper, an improved recommendation algorithm is proposed. The factor of browsing time is added into the construction of association rules. The experiments show that the proposed recommendation algorithm is better than the traditional recommendation algorithm. However, the proposed algorithm pays more attention to the efficiency but ignores the accuracy. And the algorithm just takes advantage of one attribute of consumers' behavior, namely browsing time. So the recommendation result of proposed algorithm gets a little higher than the recommendation result of traditional algorithm. In the following research, more attributes will be added into experiment, such as duration of behavior, and more works will be focused on the improvement of accuracy. Now, there are some ways worth getting a try. The change of consumers' behavior can reflect the change of consumers' interest. So recommendation can be obtained from the change. In addition, accidental behavior may contain some potential consumers' information. Such behavior is abundant in the Internet, so it is useful for recommendation to mine the potential information.

REFERENCES

- [1] J. Liu, P. Dolan, and E. Pedersen. "Personalized news recommendation based on click behavior". In Proc. Of 15th Int. Conf. on IUI, 2010, pp.
- [2] J. G. Liu, T. Zhou, and B. H. Wang, "Research progress of personalized recommendation system," Progress in natural science, vol. 19, pp. 1-15, Jan. 2009.

- [3] C. S. Cui and Q. Z. Wu, “Research on content-based recommendation based on Vague sets”, *Application Research of Computer*, vol. 27, pp. 2109-2111, Jun. 2010