User Oriented Recommendation System for Web Content

CS533
Information Retrieval Systems

Eren Golge
Tunc Gultekin
Ilker Sarac
Ahmet Iscen

A) Description of Problem

Recommending users new web content and services based on their previous web history, activity and its relation to other users.

B)Motivation/ Importance

The world-wide-web has grown enormously over the years, and there are billions of web pages today. Although, this means that the users have many options to choose from when they need to access an information, or simply just surf the web, they usually cannot find exactly what they are looking for. Therefore, it is important to guide the users to websites that they are likely to enjoy, and our system tries to seek a solution for this problem.

On the other side, knowledge of users' preferences is informative for commercial activities like user specific advertisements. In that respect, our proposed system supplies a underpinning system for their customer relations.

C) Methodology

In the past, recommendation systems for movie sites such as IMDB has been designed, such as **[1]**. This method uses movie ratings given by each user , and find the similarity between users to recommend new movies. This is done by using Pearson correlation to weight all users for similarity, finding *n* closest neighbors for a user, and finally calculating the weighted combination of selected neighbors' ratings to make a recommendation. In **[2]**, a new method is proposed that is similar to **[1]**, however it avoids some of its problems, such as very sparse matrices.

As a trendy approach NetFlix's current recommendation methodology is a successful system **[3].** They are using Collaborative Filtering method for film recommendations by using the information obtained from user's past experience and the interactions to other users. In that way, they are proposing a state of art engine performance.

For our system, we plan to inspired by the method used in **[3]**. In order to do this, instead of collecting ratings, we will collect how many times a user has visited a website, and then find the similarities between users to recommend a new website. Another scenario, that we can try is to cluster users into k clusters first, and then find the similarity between users in the same clusters. Further, we can try to improve the user activity collection by measuring the time user spends on each website, and weighing the website visit frequency according to that.

For collection of user web history, a http port interface will be developed. User need to run that interface at his system and the listener will send the his activities to our web service to process the data accordingly. Content recommendation will be presented to user with the current knowledge of out proposed system.

D) Expected Results

Since Collaborative Filtering approach has been useful in the past for other recommender system, we expect it to work for ours as well. We are hoping that the system will match similar users with each other, group them in same clusters, so that it can predict new websites according to their taste.

E) References

[1] Herlocker, J.; Konstan, J.; Borchers, A.; and Riedl, J. 1999. aAn algorithmic framework for performing collaborative filtering. In SIGIR '99: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 230–237.

[2] Prem Melville, Raymond J. Mooney, and Ramadass Nagarajan. Content-boosted collaborative filtering for improved recommendations. In Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02), July 2002

[3] The BellKor 2008 Solution to the Netflix Prize Bell, R. M., Park, F., Volinsky, C., & Park, F. (2008).