Personalized Recommendation on Web Data

Avinash Palve
Computer Engg, TCOER
Savitribai Phule Pune University, India.

Akshay Tamhane
Computer Engg, TCOER
Savitribai Phule Pune University, India.

Akshay Mitkal
Computer Engg, TCOER
Savitribai Phule Pune University, India.

Ashwin Iyer
Computer Engg, TCOER
Savitribai Phule Pune University, India.

Prasad Satam
Computer Engg, TCOER
Savitribai Phule Pune University, India.

Abstract— E-commerce and web-based services are more effective if recommendations are provided. Recommendations are carried out dynamically on web data. In this paper, such recommendation technique is proposed, which has information about rating and profile contents. These two are used to maintain a relation. The user preferences are recorded for this recommendation. Along with the preferences, the features are adaptively weighted for personalized recommendation. Using recommendation technique helps in avoiding the skipping of the end users’ preferred item.

Keywords— Recommendation; collaborative filtering; co-ratings; sparse data; dynamic features.

I. INTRODUCTION

Nowadays the internet has become an important part of human lives. It is a platform for sharing information across the world. With increase in its use, the data is expanding at a faster rate. Here comes the challenge of providing the precise content to the end user. This is where personalized recommendation comes into play for the end users’ satisfaction.

Rule-based, content-based and collaborative filtering is applied for various data analysis. The Collaborative filtering (CF) requires data about past user behaviour like ratings. It has two feature approaches viz; the neighbourhood methods and latent factor model. In the neighbourhood method, firstly it is user-oriented where it tries to find like-minded users. Secondly, it is item-oriented where it tries to find similar items based on co-ratings, and predictions based on ratings of the adjacent neighbours. Latent factor models use matrix factorization which is used compute factors and the usefulness of items to users. Collaborative filtering (CF) is a success where user preferences are static.

The sparsity of data and the dynamic feature are the two hurdles in predicting the ratings. A user can rate only a very small proportion of all items, the UxI rating matrix is quite sparse and the amount for estimating a candidate rating is far from enough. Sparsity at the same time makes latent factor struggle to catch up with slow dynamic recommendation signal which cannot be explained physically. The dynamic nature leads to difference in choices where recommendation are concerned. In our work, the interest cycle differs from user to user. At the same time, more changes cannot be exactly described by several simple decay functions. Collaborative Filtering approach is the cold-start problem which is increased in the dynamic situations as the rate of new users and items would be high. In this paper, we present a hybrid dynamic recommendation method. First we use more information while keeping the data consistent. In order to do this we use user profiles and item contents to extend the co-rate relations between ratings through each element of users, as show in Fig.1. The ratings can reflect similar users’ preferences, which in turn provides useful information for
recommendation. Correspondingly, in order to enable the algorithm to maintain the changing of signals quickly and to be updated conveniently, based on time series analysis (TSA) technique a set of dynamic features are proposed, and relevant ratings in each phase of interest are added up by applying TSA to describe user’s preferences and item’s reputations. Then we proposed a personalized recommendation algorithm by adaptively weighting. The result of the proposed algorithm is effective with dynamic data and per-forms better than the previous algorithms.

![Fig.1](image)

This paper gives us following: (a) large number of data can be used for recommender systems by inspecting the relation among related user profile and items. (b) A novel set of dynamic algorithm is proposed for describing users’ preferences that is highly flexible, effective way to structure the impacts of preferences in various phases of interest in comparison with dynamic methods used in earlier works, since the features are designed considering users’ interest and a change in them. (c) An adaptive weighting algorithm is formulated to combine the dynamic factors for personalized recommendation, with the help of factors.

II. THE PROPOSED METHOD

User profiles will be used to extend the co-rating relation. Next, dynamic features will be proposed to show users’ preferences or items popularity in different stages. Finally, adaptive algorithm for personalized recommendation comes into picture.

A. Relation mining of rating data

The information for monitoring users’ preferences are from three places viz; user profile, item profiles and historical rating records during recommendation. Existing algorithms mainly rely on the co-rate relation. But this will not efficient in calculation while the data is sparse as it limits the amount of data during prediction. So, to overcome this we introduce a semi co-rate relation for finding useful ratings for dynamic personalized recommendation.

B. Dynamic feature extraction

To compute better recommendation algorithm, three kinds of methods were proposed such as instance selection, time-window (usually time decay function) and ensemble learning. This technique considers computation, accuracy and flexibility to describe users’ multi-phase preferences.

C. Adaptive weighting algorithm

The parameters are quantified in the feature extraction as per the previous step, so now it’s easy to organize them for accurate rating estimation by using adaptive weighting. Sizes of all the relevant subsets are also computed in MPD (Multiple Phase Division) and could reflect on data density. The adaptive linear model is described as below.
\[ \hat{R}_{j,k} = \sum_{s} \sum_{d} (\alpha_{s,d} + \beta(\#R_s^d))b_{u_j}(s)b_{i_k}(s)feas_{s,d}, \]

with: \( \alpha_{s,d} \geq 0, \beta > 0. \)

Where,

- \( R_{j,k} \) – Estimated rating
- \( U_j \) – User rating
- \( i_k \) – Item
- \( T_{j,k} \) – Time point
- \( feas_{s,d} (s = 1,2,\ldots, d = 1,2,\ldots) \) got by applying Multiple Phase Division

### III. Evaluation

Root-mean-square error (RMSE), is used to evaluate the proposed recommendation algorithm. In traditional RMSE evaluation, training and testing data are randomly sampled which is not based on time. So, it would result in current prediction based on future data. Hence, Replay-match evaluation has been proposed to address this issue by Li et al whose evaluation results are more stable for dynamic recommendation.

1) To evaluate the accuracies of above mentioned dynamic recommendation algorithms as follows:
2) Sort the complete dataset in natural time order, use a certain training ratio to determine its corresponding splitting.
3) Use the previous part as the training set to adjust all parameters.
4) Run algorithm on this testing set and generate estimated rating for each user-item pair.
5) Compare each estimated ratings and real ratings with in the testing set and calculate RMSE for them.
6) Use variety of ratios and cycle through the last four steps.

### IV. Conclusions

In this paper, we proposed a personalized recommendation algorithm for web data. Here number of rating on sparse data is used in a prediction by involving more neighbouring ratings based on user and also item profiles. Dynamic features are formulated and grouped for describing the preference information based on TSA methodology, and in the end a recommendation is made by adaptively weighting these features using available information in different phases of interest.

### Acknowledgment

We are grateful to people who have helped us in completing our paper, without whom the completion of this document would not have been possible. Firstly we would like to thank our esteemed guide, Prof. Avinash Palve, for his guidance at all the times and that he provided his support and guidance without complaints at any time of the day and also for the bright ideas and inputs that they gave for the document.

We would also like to thank Prof. S. B. Chaudhari (HOD COMPUTER) for his keen interest and Cooperation towards this document. We wish to express our sincere gratitude to TCOER, Pune, for providing facilities like Internet and Library. We would also like to convey thanks to all those who directly and indirectly helped us in completing this document.

### References


Avinash Palve, Akshay Tamhane, Akshay Mitkal, Ashwin Iyer, Prasad Satam


