International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Novel Approach to Enhance Personalized Recommendation

G. Venkatesh[1], M. Sri Devi[2]

[1]M.Tech Student, Department of CSE, Anurag Group of Institutions, Hyderabad, India
[2]Associate Professor, Department of CSE, Anurag Group of Institutions, Hyderabad, India

Abstract: ***Recommendation techniques are important in the fields of E-commerce and other facilities like online shopping. One of the main problems is dynamically providing high-quality recommendation on less data. In this paper, a new dynamic personalized recommendation approach is introduced, in which information contained in both ratings and profile contents are utilized by inventing internal relations between ratings, a set of dynamic features are developed to describe user preferences in different phases , and then a recommendation is made by adaptively weighting the features. This approach performed well on public data sets.***

Keywords: Dynamic Recommendation, Dynamic Features, Multiple Phases of Interest, Time Series Analysis, Hybrid Recommendation

1. Introduction

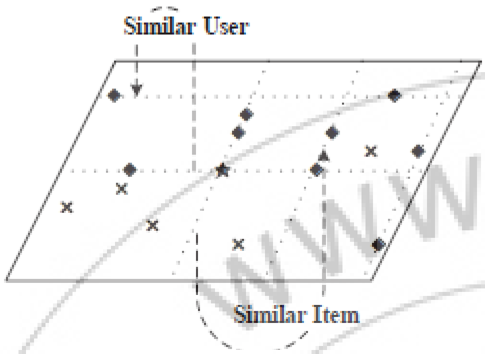
Now a day's the internet has become an inevitable part of our daily lives, and it acts as mediator and provides a platform for enterprises to deliver details about products and services to the customers easily. As the amount of this kind of information is increasing quickly, one great challenge is ensuring that proper content can be delivered quickly to the exact customers. Personalized recommendation is the best way to improve customer satisfaction. [1], [2]. Here we have mainly three methods to recommendation engines based on different data analysis methods. They are rule-based, content-based and collaborative filtering [3], [4]. Among them, collaborative filtering (CF) requires only data about past user behavior like ratings, and its two major approaches are the neighborhood approach and latent factor approach. The neighborhood methods can be user-based or item-based. These are used to find similar-taste users or similar items on the basis of co-ratings, and prediction is done based on the ratings of the nearest neighbors [5], [6], [7]. Latent factor approach try to learn internal relations from the pattern of ratings using techniques like matrix factorization [8] and use the factors to compute the usefulness of items to users. CF has been proved to perform well in situations where user interests are relatively constant. In most dynamic scenarios, there are mainly two issues that prevent accurate prediction of ratings the sparsity [3] and the dynamic nature. Since a user can only rate a very small area of all items, the U ×I rating matrix is quite sparse or less and the amount of data for estimating a user rating is very far. While latent factor models involve most ratings to capture the general interest of users, they still have problems in catching up with the drifting signal in dynamic recommendation because of less data, and it is hard to physically explain the reason of the involving.

2. Problem Definition

The dynamic nature decides that users' preferences may drift over time in dynamic recommendation, resulting in different taste to the items in different phases of interest, but it is not well understood in previous studies [9]. In our observation, the interest cycle differs from user to user, and the pattern how user interest's changes cannot be exactly described by several simple decay functions. Moreover, CF methods normally face the cold-start problem which is increased in the dynamic scenario since the number of new users and new items would be high. Some researchers have already tried to solve the above problems.

3. Proposed System

Hybrid approaches which integrate content-based and collaborative filtering methods in different ways were proposed to alleviate the sparsity or less data problem [3], [10], where more information were extracted than just in each of them. In this paper, we present a novel hybrid dynamic recommendation approach. First, in order to use more information while keeping data consistency, we use user profile and item details to increase the co-rate relation between ratings through each attribute, as shown in Fig. 1.(b). These ratings can reflect similar users' interests and provide useful data for recommendation. Similarly, in order to enable the approach to catch up with the changing of signals quickly and to be updated frequently, a set of dynamic features are developed based on time series analysis (TSA) technique, and similar ratings in each phase of interest are added up by applying TSA to describe users' preferences and items' reputations. Then we propose a personalized recommendation approach by adaptively weighting the features according to the amount of utilized rating data. This approach is effective with dynamic content and also outperforms existing algorithms. The main use of this paper can be explained as follows: (a) Lot of information can be used for recommended systems by finding the similar relation among related user profile and item content. Compared with the previous works such as [4], [12], [10], we use the similarity among data in each profile attribute so that huge content information is used, especially content in those attributes which are very hard to be quantified. (b) A novel set of dynamic features is introduced to describe users' interests, which is more flexible and easy to model the changes of interests in various individual phases of interest compared with dynamic models used in previous works, since the features are developed according to sequential characteristics of users' interest and a linear model of the features can catch up with changes in user preferences. (c) An adaptive weighting approach is designed to integrate the dynamic values for personalized recommendation, and in which time and data factors are used to adapt with dynamic recommendation on small data.



In many cases, the drifting of users' interests or items' reputations is not too rapid, which makes it possible to describe temporary state of them with the help of different data value sets. Here first we design a way to make use of profiles to increase the co-rating relation, and after that we introduce a set of dynamic features to show users' interests or items' reputations at each phase of interest, and after that we use an adaptive approach for dynamic personalized recommendation.

3.1 Dynamic Recommendation Module

Personalized recommendation is suitable ways to increase customer satisfaction .Here mainly three approaches are used in recommendation engines based on various data analysis models. They are rule-based, content-based and collaborative filtering models. Among them, collaborative filtering (CF) model requires only data about past user behavior like ratings, and its two main methods are the neighborhood methods and latent factor models. The neighborhood methods can be user-dependent or item-dependent. They try to find like-minded users or similar items on the basis of ratings to a single item by different users or to different items by a single user, and prediction is based on ratings of the nearest users. Latent factor models try to learn internal relations from the pattern of ratings using techniques like matrix factorization and use the factors to compute the usefulness of items to users.

3.2 Relation Mining Of Rating Data

For the sparsity of recommendation information, the actual problem of capturing users' dynamic interests is the lack of useful information, which we gather from three sources – user data, items information and historical rating files. Traditional algorithms heavily depends on the co-rate relation (to the single item by various users or to multiple items by the same user), which is rare when the data is less. Useful ratings are discovered using the similar relation, which is simple, intuitional and physically important when we go one or two steps along, but if we go forward it limits the amount of data used in each prediction. Instead of finding neighboring nodes along co-rate edges in the U ×I plane, we try to find a new way to find useful ratings. We notice that when considering the factors which affect a rating r(u, i), we have to focus more on some attributes of u and i in their profiles, instead of the user himself or the item itself. For example, if the movie "Gone with the Wind" is given good ratings by middle-aged people and normal ratings by teenagers with no doubt, we would firstly check on the age attribute in a user's profile when expecting probable rating the user would give to the movie, instead of other details of the user or how the user has rated remaining movies. As considering this as evident, it may not be necessary to stick only to the co-rate relation, and we introduce the semi-co-rate relation between ratings whose respective user profiles or item contents have same or identical content in one or more attributes. Since semi-co-rate is much less constrained, we enlarge the co-rate relation to it using user profile and item content, and introduce a new way of searching useful ratings for dynamic personalized recommendation.

3.3 Dynamic Feature Extraction Module

User's interests or items' reputations are changing, thus we need to deal with the dynamic nature of data to increase the precision of recommendation approaches, and recent ratings and remote ratings must have various weights in the prediction. These methods help to make report in precision of dynamic recommendation, but they also have their limitations: decay functions mayn't precisely describe the evolution of user interests and only isolating transient noise cannot catch up with the change in data.

3.4 Multiple Phase of Interest Module

A set of dynamic features to describe users' multi-phase preferences in consideration of computation, flexible nature and correctness. It is not possible to learn weights of all ratings for every user, but it is possible to learn the natural weights of ratings in the user's different phases of interest if the phases include ranges of time that are long enough.

3.5 Dynamic Feature Extraction

In order to compute best recommendation algorithm approach, three kinds of methods were introduced such as instance selection, time-window (usually time decay function) and ensemble learning methods. This particular technique contains a set of dynamic features to describe users' multi-phase interests in consideration of computation, correctness and flexibility.

3.6 Adaptive Weighting Algorithm

The parameters are quantified in the feature extraction as per the above step, so now it is very easy to organize them for exact rating estimation by using the method called 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,

$$\hat{R}_{j,k} = \sum_{s}\sum_{d}(\alpha_{s,d} + \beta(\#R_s^d))b_{u_j}(s)b_{i_k}(s)fea_{s,d},$$
$$\text{with}: \alpha_{s,d} \geq 0,\ \beta \geq 0,$$

Where, Rj, k – Estimated rating

Uj – User rating
ik– Item
Tj,k– Time point
feas, d (s =1 ,2,..., d =1 ,2,...) got by applying Multiple Phase Division

4. 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 find the correctness of above mentioned dynamic recommendation approach as follows:
2) Sort the complete dataset in normal time order, and use a certain training ratio to discover its corresponding splitting.
3) Use the existed part as the training set to adjust all parameters.
4) Run algorithm on the testing set and then generate estimated rating for each user-item pair.
5) Compare each and every estimated ratings and real ratings within the testing set and find out RMSE for them.
6) Use different ratios and cycle through the last four steps.

5. Conclusion

In this paper, we proposed a novel dynamic personalized recommendation algorithm for sparse data. Here numerous rating data is utilized in one prediction by involving more neighboring ratings through each attribute in user and also item profiles. A rich set of dynamic features are designed to describe the preference information based on TSA methodology, and finally a recommendation is made by adaptively weighting the features using available information in different phases of interest. Experimental results on real data also indicate that the proposed algorithm is highly effective, and its computational cost is much acceptable.

References

[1] B. M. Sarwar, G. Karypis, J. A. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: WWW, 2001, pp. 285–295.
[2] P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, Methods and Strategies of Web Personalization, Lecture Notes in Computer Science, Springer, 2007.
[3] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, IEEE Trans. Knowl. Data Eng. 17 (6) (2005) 734–749.
[4] Y. Koren, Collaborative filtering with temporal dynamics, Communications of the ACM 53 (4) (2010) 89–97.
[5] L. Candillier, F. Meyer, M. Boull´e, Comparing state-of-the-art collaborative filtering systems, in: P. Perner (Ed.), MLDM, Vol. 4571 of Lecture Notes in Computer Science, Springer, 2007, pp. 548–562.
[6] K. Yu, A. Schwaighofer, V. Tresp, X. Xu, H. Kriegel, Probabilistic memory-based collaborative filtering, IEEE Transactions on Knowledge and Data Engineering 16 (1) (2004) 56–69.
[7] F. Fouss, A. Pirotte, J. Renders, M. Saerens, Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation, IEEE TKDE 19 (3) (2007) 355–369.
[8] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37.
[9] S. Boutemedjet, D. Ziou, Long-term relevance feedback and feature selection for adaptive content based image suggestion, Pattern Recognition 43 (12) (2010) 3925–3937.
[10] B. M. Kim, Q. Li, C. S. Park, S. G. Kim, J. Y. Kim, A new approach for combining content-based and collaborative filters, J. Intell. Inf. Syst. 27 (1) (2006) 79–91.
[11] G. Prassas, K. C. Pramataris, O. Papaemmanouil, Dynamic recommendations in internet retailing, in: ECIS, 2001.
[12] S. Rendle, C. Freudenthaler, L. Schmidt-Thieme, Factorizing personalized markov chains for next-basket recommendation, in: Proceedings of the 19th WWW, ACM, 2010, pp. 811–820.
[13] C. Xia, X. Jiang, S. Liu, Z. Luo, Z. Yu, Dynamic item-based recommendation algorithm with time decay, in: ICNC, IEEE, 2010, pp. 242–247.
[14] J. Lai, T. Huang, Y. Liaw, A fast k-means clustering algorithm using cluster center displacement, PR 42 (11) (2009) 2551–2556.
[15] A. Tsymbal, The problem of concept drift: definitions and related work, Computer Science Department, Trinity College Dublin.
[16] X. Tang, C. Yang, J. Zhou, Stock price forecasting by combining news mining and time series analysis, in: Web Intelligence, IEEE, 2009, pp. 279–282.
[17] J. Mohammed, Real-time implementation of an efficient rls algorithm based on iir filter for acoustic echo cancellation, in: IEEE/ACS ICCSA, IEEE, 2008, pp. 489–494.
[18] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression,The Annals of statistics 32 (2) (2004) 407–499.
[19] B. Boser, I. Guyon, V. Vapnik, A training algorithm for optimal margin classifiers, in: Proceedings of the fifth annual workshop on Computational learning theory, ACM, 1992, pp. 144–152.
[20] L. Li, W. Chu, J. Langford, X. Wang, Unbiased offline evaluation of contextual-bandit-based news article

recommendation algorithms, in: Proceedings of the fourth ACM international conference on Web search and data mining, ACM, 2011, pp. 297–306.

[21] Y. Koren, Factorization meets the neighborhood: a multifaceted collaborative filtering model, in: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2008, pp. 426–434.

Author Profile

G. Venkatesh received the B.Tech degree in computer science and Engineering from JNTU Hyderabad in 2012 and pursuing M.tech degree in Computer science and Engineering from Anurag Group of Institutions (Formerly CVSR College of Engineering) JNTU Hyderabad.

M. Sri Devi is working as Associate Professor in Computer Science Engineering from ANURAG GROUP OF INSTITUTIONS (Formerly CVSR College of Engineering) JNTU Hyderabad.