ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

Active Personalized Recommendation on Large Stream Data

Samuel F lalremruata[1], Sunil G.L[2]

M.Tech Student, Dept of CS&E, Acharya Institute of Technology, Bangalore, Karnataka, India[1]

Assistant Professor, Dept. of CS&E, Acharya Institute of Technology, Bangalore, Karnataka, India[2]

ABSTRACT: With the dramatic increase of modern communication, the internet and multimedia technology comes with large amount of information which make it difficult to provide high quality of recommendation. Information filtering can be more precise if human profiles are includes in the process of filtering. To solve this problem, recommender system help users to make decision that would meet their requirement by predicting their preferences while browsing and searching, based on the user past preferences with other similar group of users.

A recommender system is a useful method that measure similarity between users, in order to increase the quality of recommendation it utilized some of their behaviors features like ratings. Information containing both rating and profile content are making use of it by exploring their co-rate relation between ratings, a set of dynamic features are proposed to describe the user inclination in multiple level and finally recommendation made by adaptively weighting the features. By implementing a webpage which serves as a front-end to a recommender system is use to evaluation results of the overall user experiences.

KEYWORDS: dynamic features, multiple phases of interest, adaptive weighting, recommendation.

I. INTRODUCTION

With the huge quantity of information and popularity of internet, most enterprises are looking for more precise ways of using the current data to give accurate prediction. Based on similarities between users profile, system can provide predictions for a new item to user. Recommender systems develop to obtained some basic features of user, item and rating information to predict how new users will like this particular item. Analyzing large quantity of information is required to predict user preference with the other similar user. Recommender systems play an important role in the field of e-commerce and web services, as data and information keep increasing day by day, one of the main challenging problems which turn out to be interesting in the field of machine learning system is to give an accurate prediction to the recommend user from the large amount of different users. The recommender system need to analysis lots of data from different resources so that user is given the right product or item. The computation power need for recommender system is large and should be efficient to predict user's requirements from the large dataset. Rather than suggesting random products to the user, recommending an item or product based on the user preferences will give fulfillment to user also save more time.

There are many algorithms that could be applied on recommendation system. Among them collaborative filtering, content-based filtering and rule based are the approaches that could be applied on data to predict a user preference. However one of the most well-known predictive system which suggest item for user by collecting their past preferences information from many similar users is known to be collaborative filtering. Collaborative filtering is a method which required history about the user behavior like rating, transaction, activity, etc; its two main techniques are latent factor method and neighboring method. Technique using latent factor model try to learn from the rating pattern from users which are helpful in computing when suggesting items to users. The neighboring method can be the user or item oriented which try to find similar users preferences on the basis of co-rate relation which prediction is based on the nearby neighbor. Collaborative filtering has made great achievement where user preferences are perform in relative static.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

II. RELATED WORK

As the development of e-commerce and mass communication through internet production of data are easily available. Recommender systems have turn out to be a fashionable technique to deal with huge information spaces toward user that best meet their needs and likeness; many systems have been projected for using recommendations such as collaborative method and content-based filtering. A different technique such as content-based filtering selects information based on similar content while collaborative filtering merge the user and other similar users to make a prediction for a target user preferences. In this paper a new filtering approach that come together the content-based filtering and collaborative filtering to take advantage on their individual rule which thereby accomplish a better performance. A sequence of recommendations are choose for having suitable feature which look into different method for analyze user-user resemblance based on the information needed to take out from user profiles and user ratings of item. Finally, it is clear that this approach is compare it with classic filters give the results to demonstrate the helpfulness of our approach.

III. EXISTING SYSTEM

In dynamic data user could only rate small quantity of items from large dataset of information make it hard to predict accurate on rating due its dynamic in nature. These models are hard to seize up with the flow of signal in dynamic data although latent factor model try to contain the wide experience for the users. Simple decay functions cannot describe the change made in the user preferences as the phase of interest period differs from user to user. Recommendation techniques such as content-based, rule-based and collaborative have been proposed to tackle with this problem. This algorithm does not deal well with cold start problem in dynamic situation since rate of new user and item would be high.

IV. PROPOSED METHOD

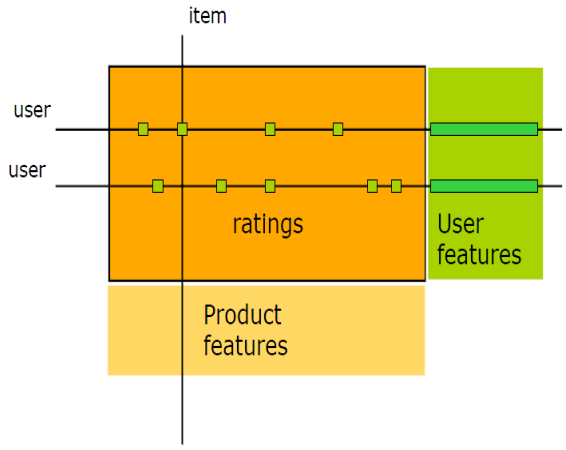

Figure1: The proposed system

The above figure show the proposed system for personalized recommendation, this model is the combination of item-user features matrix and user-item feature matrix. The user-item is constructed from the similar content from the user feature and product features. More precise recommendation can be made by co-rate the relation between both user and item features. In the hybrid dynamic recommendation system, it utilizes different information from historic data for real time prediction on the application. Traditional algorithm rely on the co-rate relation which are not based on time have lower prediction than the proposed system.

To enhance the precision of recommendation system which deal with dynamic data, recent rating and remote rating should have dissimilar weight in prediction means that data in different phases of interest have different ratio. The proposed system is quite efficient and found that a recent rating does not have higher weight every time. This show the proposed method features are flexibility and highly accurate than the previous system. In this approach user profile and

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

item content to extend the co-rate relation between ratings through each attribute, the involved rating can provide useful information for similar user preferences for recommendation. Due to dynamic in nature changes its signal from time to time, so a novel proposed of time series analysis (TSA) algorithm is enable that can catch up and update regularly, the relevant ratings are added up to describe user and item preferences. Finally a personalized recommendation is proposed by using adaptively weighting the features that utilize some of the rating data.

4.1 Relation mining

In large dynamic information it is difficult to capture the interest users as the pattern of user preferences change from time to time. Existing method mainly rely on co-rate relation which predict its nearest neighbor which are not efficient in dynamic data as it limits prediction on items or products. To overcome this, functional ratings are revealed using the semi co-rate relation between ratings whose corresponding user profiles or item contents have identical content in one or more attributes and propose a new way for dynamic recommendation.

4.2 Dynamic feature extraction

To enhance accuracy on recommendation algorithms in dynamic nature of data recent weight and remote weight should have different weight in prediction. Instance selection, time-window (usually time decay function) and ensemble learning cannot precisely describe the user preferences as it cannot catch up with the changes in data. Learning all the rating from users is impossible but it's possible to learn the general weight of ratings in dynamic data that describe user preferences involved in ratings. Time series analysis algorithm (TSA) is enables to match up with the change in data that are update periodically on the users' inclination. An array of ratings is arranged by time order where time series analysis approach is efficient to dealing with such data to prefer on the basic feature extraction method.

The basic TSA algorithm is:

$$\begin{cases} R_s^d = \{R_{j',k'} | R_{j',k'} \in R_s \ and \ T_{j,k} - T_{j',k'} \geq T_d\}, \\ fea_{s,d} = \sum_{l=1}^{o} \mu(1-\mu)^{l-1} R_{s,l}^d, \end{cases}$$

where,
R^d_s (d = 1, 2, ...) are the secondary subsets.
T_d (d = 1, 2, ...) are a sequence of time differences manually set,
$R^d_{s,l}$ (l = 1, 2, ...)are the rating values listed in reversed order in the subset
μ - is the forgetting element for index smoothing.

A novel approach is proposed to TSA that have a flexible way of feature extraction where weight can be different for different rating can be learned from data which have give accuracy and flexibly than the previous step. This proposed method is known as multiple phase division (MPD).

4.3 Adaptive weighting algorithm

The parameter is measure in the feature extraction of user and item profile from previous step, so now it is simple to organize them for particular rating since they are consistent and it is efficient to learn their weight by using adaptive weighting. Sizes of the weight on relevant associate sets are also calculated in MPD (Multiple Phase Division) that could reflect on data density. The adaptive linear model is described as below:

$$R_{j,k} = \sum_s \sum_d (\alpha_{s,d} + \beta(\#R_s^d)) b_{ik}(s) \, fea_{s,d}$$

with: $\alpha_{s,d} >= 0$, $\beta >= 0$

where,
$R_{j,k}$ – indicate the estimated rating

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

U_j – user rating
i_k– item rating
$T_{j,k}$– time point
feas,d (s =1 ,2,..., d =1 ,2,...) – features calculation
b_{ui} and b_{ik} - binary function of candidate rating
$\alpha_{s,d}$ and β – weighting parameter.

V. CONCLUSION

In this paper, a webpage is build that extracts the basic user attributes from a large amount of data by using distributed processing platform then construct the user preference models by machine learning method. This claim that proposed recommendation system is more efficient than the traditional way of presenting recommendations as one-dimensional lists, in the sense of predict based on rating of the nearest neighbor. These features are designed to describe interest phases on user preferences based on time series analysis (TSA) technique and finally recommendation is made using adaptively weighting the features using information on the interest rating data. The hybrid approaches handle more information that possibly achieves better accuracies if the dynamic nature of data is consistent and well minded.

REFERENCES

[1] 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.
[2] 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 (2007).
[3] 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.
[4] B. M. Sarwar, G. Karypis, J. A. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: WWW, 2001, pp. 285–295.
[5] P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, Methods and Strategies of Web Personalization, Lecture Notes in Computer Science, Springer, 2007.
[6] 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.
[7] Y. Koren, Collaborative filtering with temporal dynamics, Communications of the ACM 53 (4) (2010) 89–97.
[8] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37.
[9] J. Lai, T. Huang, Y. Liaw, A fast k-means clustering algorithm using cluster center displacement, PR 42 (11) (2009) 2551–2556.
[10] A. Tsymbal, The problem of concept drift: definitions and related work, Computer Science Department, Trinity College Dublin.
[11] 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.
[12] 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.
[13] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, The Annals of statistics 32 (2) (2004) 407–499.
[14] 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.

BIOGRAPHY

Samuel F Lalremruata received his B.E degree in Computer Science and Engineering from Bharati Vidyapeeth College of Engg, Navi Mumbai. Now perceiving M.tech degree in Computer Science and Engineering at Acharya Institute of Technology, Bangaluru.