iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 5 **Issue:** I **Month of publication:** January 2017

**DOI:**

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 5 Issue I, January 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

# An Incremental-And-Static-Combined Scheme for Matrix-Factorization-Based Collaborative Filtering

K. P. Ashvitha[1], B. Jerlyn Flora[2], S. Apoorva[3], D. Lakshmi[4]
[123]Student, Fourth Year, [4]Professor
[1,2,3,4]Department of CSE, Panimalar Institute Of Technology, Chennai-600123.

***Abstract:*** ***The last decade has witnessed a tremendous growth of Web services as a major technology for sharing data, computing resources, and programs on the Web. With increasing adoption and presence of Web services, designing novel approaches for efficient and effective Web service recommendation has become of paramount importance.***
***In existing web services discovery and recommendation approaches focus on keyword-dominant Web service search engines, which possess many limitations such as poor recommendation performance and heavy dependence on correct and complex queries from users. Recent research efforts on Web service recommendation center on two prominent approaches: collaborative filtering and content-based recommendation. Unfortunately, both approaches have some drawbacks, which restrict their applicability in Web service recommendation. In proposed system for recommendation we will be using Agglomerative Hierarchal Clustering or Hierarchal Agglomerative Clustering for effective recommendation in web-services. our approach considers simultaneously both rating data (e.g., QoS) and semantic content data (e.g., functionalities) of Web services using a probabilistic generative model.***

## I. INTRODUCTION

Collaborative Filtering (CF) is a natural choice for designing Recommender Systems (RSs) since it provides recommendations by centrally analyzing the user-item rating matrix alone [1]. CF can generate user specific recommendations based on historical user preferences. Inside a CF, the user interests on involved items (e.g., films, books, purchase records, etc.) are quantized into a user-item rating matrix, where high ratings denote strong preferences.
So the problem of CF can be considered as the problem of missing data estimation, in which the main task is to predict the unknown user-item pairs based on known entries with minimum accumulative error [2]. CF supposes that users sharing the same ratings on past items likely to agree on new items. Research on CF can be grouped into two categories: memory-based and model-baseddel-based [3,4].
Memory-based methods compute similarities between users or between items and apply them to recognize the top most similar neighbors. Then the unknown rating is predicted by combining the known rating of the neighbors. However, there exist three essential challenges for memory-based approaches. The first one is sparsity

## II. LITERATURE SURVEY

*A. SURVEY OF THE STATE-OF-THE-ART AND POSSIBLE EXTENSION :G.ADOMAVICIUS AND A.TUZHILIN - JUNE2005*

Abstract This paper presents an overview of the field of recommender systems and describes the current generation of recommendation methods that are usually classified into the following three main categories: content-based, collaborative, and hybrid recommendation approaches.
This paper also describes various limitations of current recommendation methods and discusses possible extensions that can improve recommendation capabilities and make recommender systems applicable to an even broader range of applications. These extensions include, among others, an improvement of understanding of users and items, incorporation of the contextual information into the recommendation process, support for multcriteria ratings, and a provision of more flexible and less intrusive types of recommendations.

*B. Improvement of aggregate recommendation diversity and accuracy using ranking- g.adomavicius andy.kwon -may 2012*

Abstract Recommendation systems are becoming necessary for individual user and also for providing recommendations at

individual level in various types of businesses.

Recommender system is a personalized information filtering technique used to identify desired number of items based on interest of user. The system uses data on past user ratings by applying various techniques. This techniques concentrate to improve accuracy in recommendations, with recommendation accuracy it is also necessary improve aggregate diversity of recommendation. In this paper, we proposed number of item ranking techniques and different ratings prediction algorithm to improve recommendation accuracy and aggregate diversity by using realworld rating dataset. Keywords: Recommender system, Recommendation diversity, collaborative filtering, Ranking function.

*C. Typicality-based collaborative filtering recommendation system- y.cai,h. Leung, q.li,h.min, j.tang, and j.li,"march 2014*

Abstract Collaborative filtering (CF) is an important and popular technology for recommender systems.

However, current CF methods suffer from such problems as data sparsity, recommendation inaccuracy and big-error in predictions. A distinct feature of typicality-based CF is that it finds 'neighbors' of users based on user typicality degrees in user groups (instead of the co-rated items of users, or common users of items, as in traditional CF)[1].

In online social networks people have a dual role as both "users" and "items", e.g., both initiating and receiving contacts. Here the assumption of active users and passive items in traditional collaborative filtering is inapplicable. In this paper we propose a model that fully captures the bilateral role of user interactions within a social network and formulate collaborative filtering methods to enable people to people recommendation.

In this model users can be similar to other users in two ways – either having similar "taste" for the users they contact, or having similar "attractiveness" for the users who contact them. We develop Social Collab, a novel neighbour-based SVD[2] algorithm to predict, for a given user, other users they may like to contact, based on user similarity in terms of both attractiveness and taste. In social networks this goes beyond traditional, merely taste-based, collaborative filtering for item selection [3].

*D. An efficient non-negative matrix-factorization-based approach to collaborative-filtering- X.Luo,M.-C.Zhou,Y.-N.Xia,andQ.-S.Zhu,May 2014*

Abstract: Recommender systems collect various kinds of data to create their recommendations. Collaborative filtering is a common technique in this area. This technique gathers and analyzes information on users preferences, and then estimates what users will like based on their similarity to other users. However, most of current collaborative filtering approaches have faced two problems: sparsity and scalability.

This paper proposes a novel method by applying non-negative matrix factorization, which alleviates these problems via matrix factorization and similarity. Non-negative matrix factorization attempts to find two non-negative matrices whose product can well approximate the original matrix. It also imposes non-negative constraints on the latent factors.

The proposed method presents novel update rules to learn the latent factors for predicting unknown rating. Unlike most of collaborative filtering methods, the proposed method can predict all the unknown ratings. It is easily implemented and its computational complexity is very low. Empirical studies on MovieLens and Book-Crossing datasets display that the proposed method is more tolerant against the problems of sparsity and scalability, and obtains good results.Keywords: recommender system, collaborative filtering, non-negative matrix factorization, latent factors

*E. "Latent Semantic Models for Collaborative Filtering," T. Hofmann.*

Abstract: Collaborative filtering aims at learning predictive models of user preferences, interests or behavior from community data, that is, a database of available user preferences. In this article, we describe a new family of model-based algorithms designed for this task. These algorithms rely on a statistical modelling technique that introduces latent class variables in a mixture model setting to discover user communities and prototypical interest profiles.

We investigate several variations to deal with discrete and continuous response variables as well as with different objective functions. The main advantages of this technique over standard memory-based methods are higher accuracy, constant time prediction, and an explicit and compact model representation.

The latter can also be used to mine for user communitites. The experimental evaluation shows that substantial improvements in accucracy over existing methods and published results can be obtained.Categories and Subject Descriptors: H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—information filtering;I.5.3[Pattern Recognition]: Clustering—algorithmsGeneral Terms: Collaborative filtering, recommender systems, machine learning, mixture models, latent semantic

analysis

*F. "maximum-margin matrix factorization"* n. Srebro, j. D. M. Rennie, and t. S. Jaakola,*

Abstract: We present a novel approach to collaborative prediction, using low-norm instead of low-rank factorizations. The approach is inspired by, and has strong connections to, large-margin linear discrimination. We show how to learn low-norm factorizations by solving a semi-definite program, and discuss generalization error bounds for them.

*G. "Methods for large scale SVD with missing values"- M. Kurucz, A. Benczúr, and K. Csalogány*

ABSTRACT: We compare recommenders based solely on low rank approximations of the rating matrix.

The key difficulty lies in the sparseness of the known ratings within the matrix that cause expactation maximization algorithms converge very slow.

Among the prior publicly known attempts for this problem a gradient boosting approach proved most successful in spite of the fact that the resulting vectors are nonorthogonal and prone to numeric errors.

We systematically explore expectation maximization methods based both on the Lanczos algorithm and power iteration; novel in this paper is the efficient handling of the dense estimate matrix used as input to a next iteration. We also compare sequence transformation methods to speed up convergence

*H. Improving regularized singular value decomposition for collaborative filtering"- A. Paterek,*

ABSTRACT : A key part of a recommender system is a collaborative filtering algorithm predicting users'preferences for items. In this paper we describe different efficient collaborative filtering techniques and a framework for combining them to obtain a good prediction.

The methods described in this paper are the most important parts of a solution predicting users' preferences for movies with error rate 7.04% better on the Netflix Prize dataset than the reference algorithm Netflix Cinematch. The set of predictors used includes algorithms suggested by Netflix Prize contestants: regularized singular value decomposition of data with missing values, K-means, postprocessing SVD with KNN.

We propose extending the set of predictors with the following methods: addition of biases to the regularized SVD, postprocessing SVD with kernel ridge regression, using a separate linear model for each movie, and using methods similar to the regularized SVD, but with fewer parameters.

All predictors and selected 2-way interactions between them are combined using linear regression on a holdout set

## III. EXISTING SYSTEM

In the existing system, The most fundamental challenge for the Big Data applications is to explore the large volumes of data and extract useful information or knowledge for future actions.

Major disadvantage of Web service discovery and recommendation approaches focus on either perishing Universal Description Discovery and Integration registries, or keyword-dominant Web service search engines.

In this approach which possess many limitations such as poor recommendation performance and heavy dependence on correct and complex queries from users.In traditional CF algorithms, to compute similarity between every pair of users or services may take too much time, even exceed the processing capability of current RSs.Consequently, service recommendation based on the similar users or similar services would either lose its timeliness or couldn't be done at all.

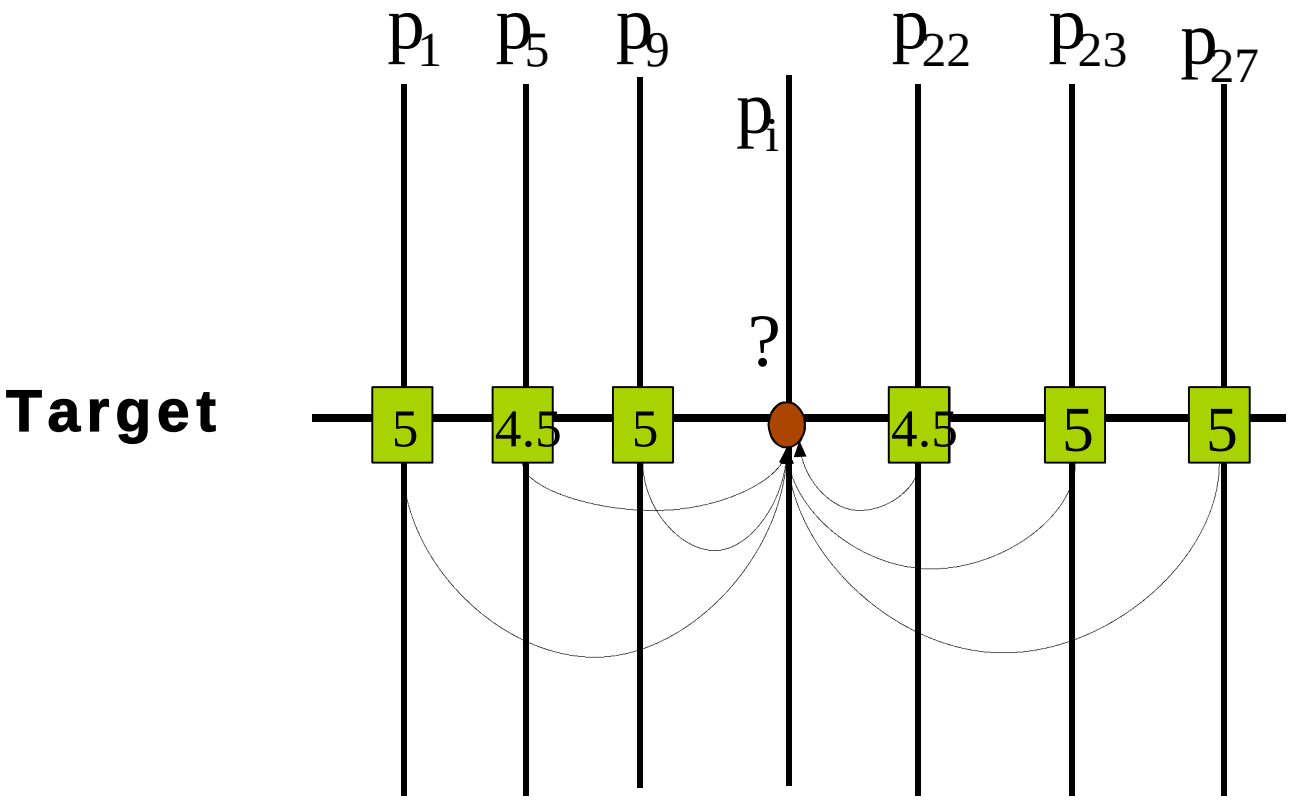
## IV. PROPOSED SYSTEM

In the proposed system We proposed an Agglomerative Hierarchal Clustering or Hierarchal Agglomerative Clustering.Clustering are such techniques that can reduce the data size by a large factor by grouping similar services together. A cluster contains some similar services just like a club contains some like-minded users. This is another reason besides abbreviation that we call this approach Club CF. This approach is enacted around two stages.

In the first stage, the available services are divided into small-scale clusters, in logic, for further processing. At the second stage, a collaborative filtering algorithm is imposed on one of the clusters.This similarity metric computes the Euclidean distance d between two such user points this value alone doesn't constitute a valid similarity metric, because larger values would mean more-distant, and therefore less similar, users. The value should be smaller when users are more similar

Volume 5 Issue I, January 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

## V. SYSTEM MODULES

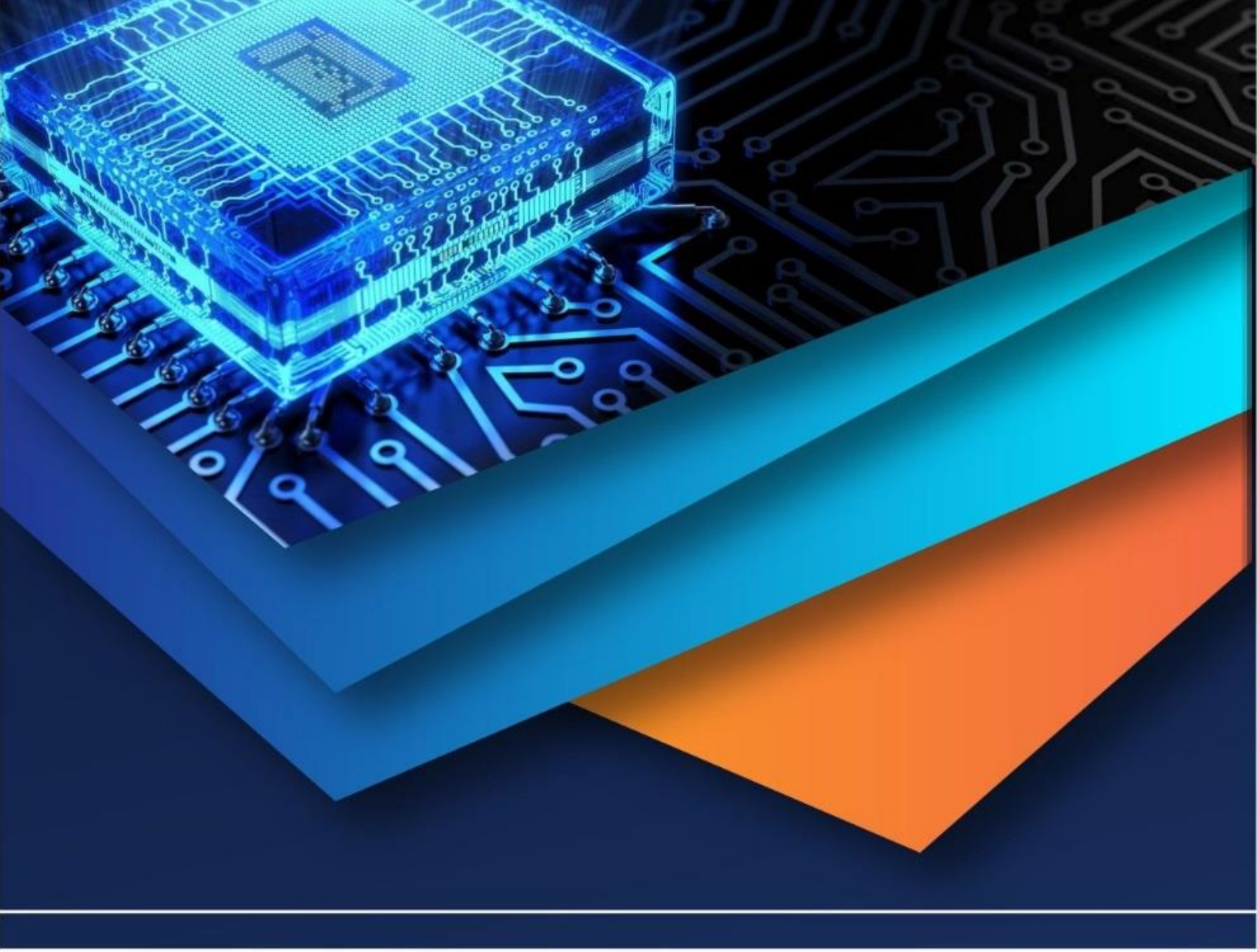



## VI. CONCLUSION

This work focuses on developing an Incremental-and-staticcombined scheme for matrix factorization-based collaborative filtering. The proposed scheme enables designers to build up a recommender with an incremental component to trace the rating-variations, and static component to maintain the context information in the static ratings. The highly accurate and timely recommendations are thus generated through combining the prediction results from them. As an example, this scheme is applied to RMF, thereby resulting in an incremental-and-static-combined RMF-based recommender (IR2). From the experimental results on three industrial-site datasets, we conclude that IR2 can obtain high prediction accuracy close to that of the totally retrained model. Moreover, since our scheme is realized by manipulating rating-variations, it can easily be adapted to the other MF-based recommenders to achieve high prediction accuracy with rating-variations while well meeting the computational requirements.

## REFERENCES

[1] G. Adomavicius and A.Tuzhilin,"Toward thenextgeneration of recommender systems: A survey of the state-of-the-Art and possible extensions," IEEE Trans. Knowl. Data Eng., vol. 17, pp. 734–749, Jun. 2005.
[2] G.AdomaviciusandY.Kwon,"Improvingaggregaterecommendation diversity using ranking-based techniques," IEEE Trans. Knowl. Data Eng., vol. 24, pp. 896–911, May 201
[3] Y.KorenandR.Bell,"Advancesincollaborativefiltering,"inRecommender System Handbook, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor,Eds. NewYork,NY,USA:Springer,2011,pp.145–186.
[4] X.Luo,M.-C.Zhou,Y.-N.Xia,andQ.-S.Zhu,"Anefficientnon-negative matrix-factorization-based approach to collaborative-filtering," IEEETrans.Ind.Informat.,vol.10,pp.1273–1284,May201
[5] T. Hofmann, "Latent Semantic Models for Collaborative Filtering," ACM Trans. Inf. Syst.,vol.22,pp.89–115,2004.
[6] N. Srebro, J. D. M. Rennie, and T. S. Jaakola, "Maximum-margin matrix factorization," Adv. Neural Inform. Process. Syst., vol. 17, pp. 1329–1336, 2005.
[7] M. Kurucz, A. Benczúr, and K. Csalogány, "Methods for large scale SVD with missing values,",CA,USA,2007,pp.7–14.
[8] APaterek,"Improving regularized singular value decomposition for collaborativefiltering,"inProc.13thACMSIGKDDInt.Conf.Knowl. DiscoveryDataMining,SanJose,CA,USA,2007,pp.39–42.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY