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. 4, Issue 6, June 2016

Hybrid Clustering for Friend Recommender

Jyoti Sharma, Pinki Tanwar

M.Tech Student, Dept. of C.S, JMIT Radaur, India

Assistant Professor, Dept. of C.S, JMIT Radaur, India

ABSTRACT: Friends recommendation is crucial for the growth of social networks. At the early stage of social networks, the network is small with few users, it is easy to browse over other candidates profiles to make a friend request. In contrast with past, the number of social network users reaches an unbelievable level. Social networking sites not only provide a way to keep in contact with friends, but they can also offer opportunities to full fill informational needs of user. Social network is becoming an increasingly popular media for information sharing. These sites are the way of establishing virtual communities and online relationships by making online friends. Most existing methods for recommending friends in social networks only aim at achieving high recommendation success rate. The network grown from such recommendations is not optimized for content spread. In this paper, we will discuss about the similarity based analyses and clustering based analyses and hybrid method along with a proposed method.We also validate our results using performance metrics like accuracy, recall, precision, f-measure.

KEYWORDS: Friend Recommendation, Social Network, Hybrid based recommendation

I. INTRODUCTION

Social network services,e.g, Facebook and Twitter in U.S.A., QQ and Weibo in China, have grown greatly in recent years. Friend recommendation is crucial for the growth of social networks. At the early stage of social networks, the network is small with few users, it is easy to browse over other candidates profiles to make a friend request. In contrast with past, the number of social network users reaches an unbelievable level.

RECOMMENDER SYSTEM :-A recommender system is a software tool that supports users in identifying the most interesting items. With the development of Web 2.0, the study of social-based recommender systems started. One challenge with existing social networking services is how to recommend a better friend to a user. Most of them relies on pre-existing user relationships to select friend candidates. e.g, facebook relies on a social link analysis among those who already share common friends and recommends symmetrical users as potential friends. But, this approach may not be the most appropriate.
Recommender system is generally of two types :- content-based filtering and collaborative filtering.

I. **Collaborative Filtering:-** Collaborative filtering recommender systems operate on the concept that users who have exhibited similar interests in the past will tend to rate items in a similar fashion. The predicted rating of an item d for some user u is calculated by collecting a predetermined number k of u's nearest neighbors and aggregating the ratings that they have provided for d in a way that reacts the degree of similarity between u and those neighbors. The nearest neighbors of u are the users that exhibit the highest degree of similarity to u, and are commonly determined using the k-Nearest Neighbor approach. **C**ollaborative filtering is the process of filtering information or patterns using techniques which involves collaboration among various agents, viewpoints, data sources, etc. Applications of collaborative filtering involves very large data sets. Collaborative filtering methods have been applied to many non-identical kinds of data including: financial data, such as financial service institutions that combine many financial sources; monitoring and sensing data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; or in electronic commerce and web applications where the focus is on user data, etc.

II. **Content-based filtering** :- Pure Collaborative Filtering recommenders only utilize the user ratings matrix, either directly, or to induce a collaborative model. These approaches treat all users and items as atomic units, where predictions are made without regard to the specifics of individual users or items. However, one can

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. 4, Issue 6, June 2016

make a better personalized recommendation by knowing more about a user, such as demographic information , or about an item, such as the director and genre of a movie.Content-based recommenders refer to such approaches, that provide recommendations by comparing representations of content describing an item to representations of content that interests the user.These approaches are sometimes also referred to as *content-based filtering*.Much research in this area has focused on recommending items with associated *textual* content, such as web-pages, books, and movies; where the web-pages themselves or associated content like descriptions and user reviews are available.As such, several approaches have treated this problem as an Information Retrieval(IR) task, where the content associated with the user's preferences is treated as a query, and the unrated documents are scored with relevance/similarity to this query.To classify a new document, it is compared with each prototype vector and given a predicted rating based on the cosine similarity to each category.

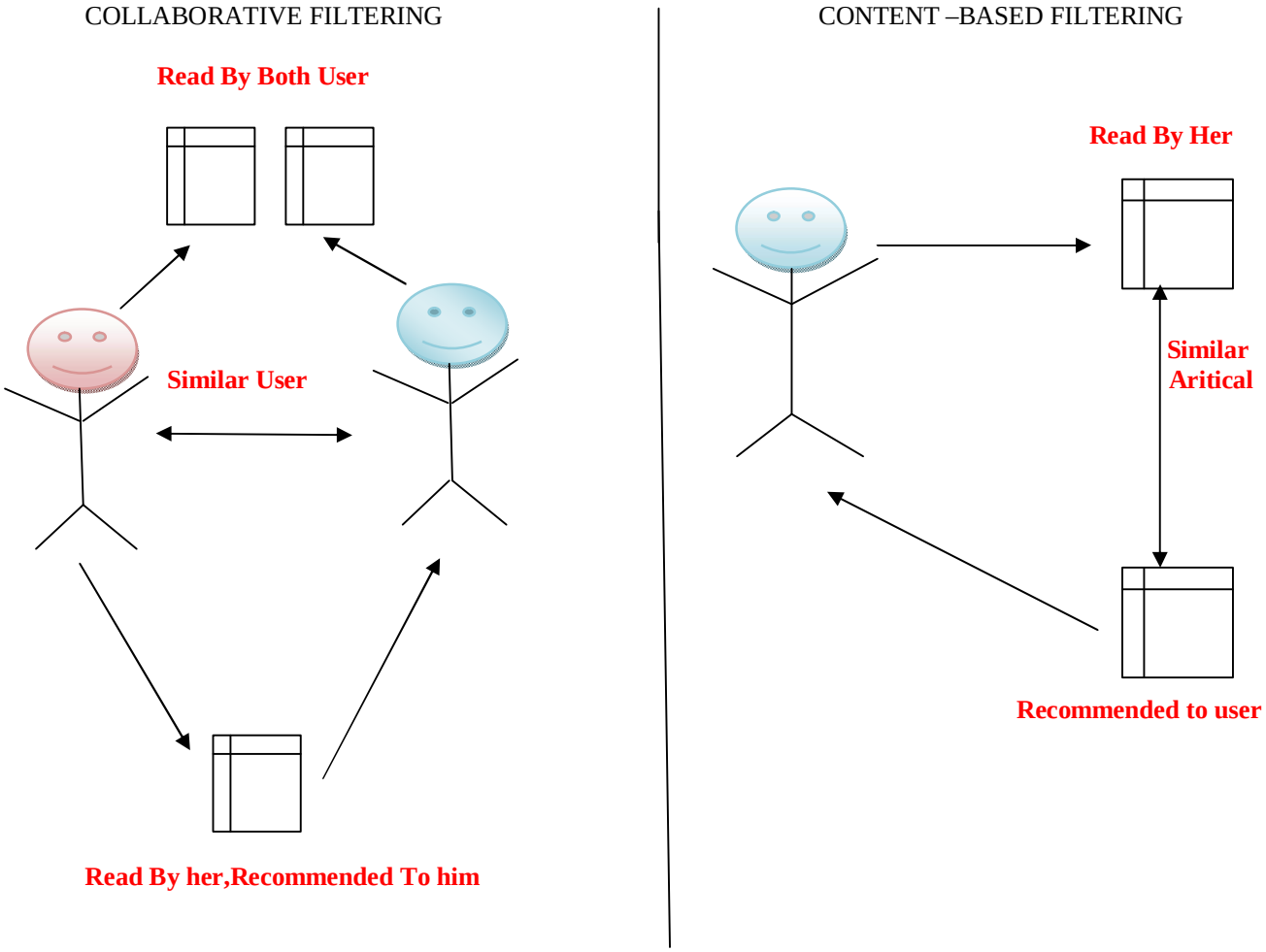


Figure 2: Collaborative and Content-Based Filtering

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. 4, Issue 6, June 2016

II. RELATED WORK

Zhi Yu, Can Wang[*], Jiajun Bu, Xin Wang, Yue Wu, Chun Chen [1], In this paper, we propose a novel friend recommendation method ACR-FoF (algebraic connectivity regularized friends-of-friends) that take both success rate and content spread in the network.

C. Martinez-Cruz[a,*], C.Porcel[a,*], J.Bernabé-Moreno[b],E.Herrera-Viedma[b,*] [2],. In this paper we develop an ontology to characterize the trust between users using the fuzzy linguistic modeling, so that in the recommendation process we do not take into account users with similar ratings history but users in which each user can trust. We provide a method to aggregate the trust information captured in the trust-ontology and to update the user profiles based on the feedback.

Zhoubao Sun[a], Lixin Han[a,*], Wenliang Huang[a], Xueting Wang[a], Xiaoqin Zeng[a],Min Wang[a], Hong Yan[b,c] [3], In this paper, we propose a social regularization approach that incorporates social network information to benefit recommender systems. Both users' friendships and rating records are employed to predict the missing values(tags) in the user-item matrix.Mainly, we use a biclustering algorithm to identify the most suitable group of friends for generating different final recommendations. We employ both friendships among users and rating records (tags) to predict the missing values (tags) in the user-item matrix.

Hsiu-Yu Liao, Kuan-Yu Chen, Duen-Ren Liu[*][4], This paper proposes a novel friend recommendation model, namely a hybrid SVM classifier considering user similarity and virtual contact strengths. In the proposed approach, users' contact activities in virtual worlds are characterized into dynamic features and contact types in order to derive their contact strengths.

Zhou Zhang [a], Yuewen Liu [a,*],Wei Ding [b],Wei (Wayne) Huang [a,c], Qin Su[a], Ping Chen[b] [5],In this study, we introduce a new Friend Recommendation system using a User's Total Attributes Information (FRUTAI) based on the law of total probability.

Parham Moradi[*], Sajad Ahmadian, Fardin Akhlaghian [6],In this paper, we present a model based collaborative filtering method by applying a novel graph clustering algorithm and also considering trust statements. The collaborative filtering approach is a powerful technology for users to find their interesting information. Trust is a concept that has recently attracted much attention and has been considered in online recommendation systems.

NikolaosPolatidis[*],ChristosK.Georgiadis[7], In this paper we propose a multilevel recommendation method with its main purpose being to help users in decision making by providing recommendation so better quality.

III. PROPOSED WORK

While using the similarity based recommender system the time measure factor is not so good that's why we use the cluster based recommender system and hybrid based recommender system in that we will use the both similarity and cluster based recommender system.

- **Clustering based recommendation**:- The main objective is to identify the communities of similar users based on their social relationships and use these communities as a mechanism to make the recommendations. Currently, some recommender systems allow users to build their social networks and use this network as additional information to suggest items to users.

  The aim of cluster analysis is organising a collection of data of uses into clusters based on their similarity.The points within one cluster are more similar to one another than to any other points from the remaining clusters.After that we predict the centre of different cluster.

  In this we also use the k-mean algorithm to improve the accuracy of recommendation.

- **k-mean algorithm :-** K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem.The main idea is to define k centroids, one for each cluster. These centroids shoud be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other.The algorithm is composed of the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2.Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.

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. 4, Issue 6, June 2016

4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

- **Similarity based recommendation :-** It is also the numerical measure of the degree to which the objects are different. For more similar objects the dissimilarity are lower value. Dissimilarity fall in the range of [0,1], with the upper range varying may be from zero to infinity.In this we use the euclidean distance matrix.

  - **Euclidean metrix** :- it is a popular distance calculation algorithm.It is the "ordinary" distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space. The associated norm is called the Euclidean norm**.** Older literature refers to the metric as Pythagorean metric.

In mathematics, a Euclidean distance matrix is an $n \times n$ matrix representing the spacing of a set of n points in Euclidean space. If A is a Euclidean distance matrix and the points x1,x2,…….xn are defined on m-dimensional space, then the elements of A are given by

$$\begin{aligned} A &= (a_{ij}); \\ a_{ij} &= \|x_i - x_j\|_2^2 \end{aligned}$$

By the properties of the 2-norm the matrix A has the following properties.

- All elements on the diagonal of A are zero (i.e. it is a hollow matrix).
- The trace of A is zero (by the above property).
- A is symmetric.
- $\sqrt{a_{ij}} \leq \sqrt{a_{ik}} + \sqrt{a_{kj}}$ (by the triangle inequality)
- $a_{ij} \geq 0$
- The number of unique (distinct) non-zero values within an n-by-n Euclidean distance matrix is bounded above by n(n-1)/2 due to the matrix being symmetric and hollow.
- In dimension m, a Euclidean distance matrix has rank less than or equal to $m+2$. If the points x1,x2,……xn are in general position , the rank is exactly min(n, m + 2).

- **Hybrid based recommendation :-** In this approach we combine the both cluster and similarity based recommendation.

Training set:- In this we take many steps for better recommendation-

STEP 1 : Firstly we take the large dataset
STEP 2 : Then we select appropriate data according to similarity of user data.
STEP 3 : Then we recommend that suggestion to new user based on high similarity.

Testing data :- In this we follow many steps which are explain below –

STEP 1 : *In this firstly predict cluster of large data.*
STEP 2 : Then we divide the cluster according to the similarity of data.
STEP 3 : Then we predict the center of cluster.
STEP 4 : When a new user come then we choose the top k cluster.
STEP 5 : Then we match the data of user to the cluster.
STEP 6 : After that we check the performance matrix.

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. 4, Issue 6, June 2016

❖ **Performance Metrics**

➢ **Precision –:-** Precision is defined as the ratio of relevant items selected to number of items selected. Precision represents the probability that a selected item is relevant is given as .

$$P= N_{rs}/N_s$$

Where Nrs are the relevant items selected Ns is the number of items selected.

➢ **Recall -** It is defined as the ratio of relevant items selected to total number of relevant items available. Recall represents the probability that a relevant item will be selected is given as

$$R=N_{rs}/N_r$$

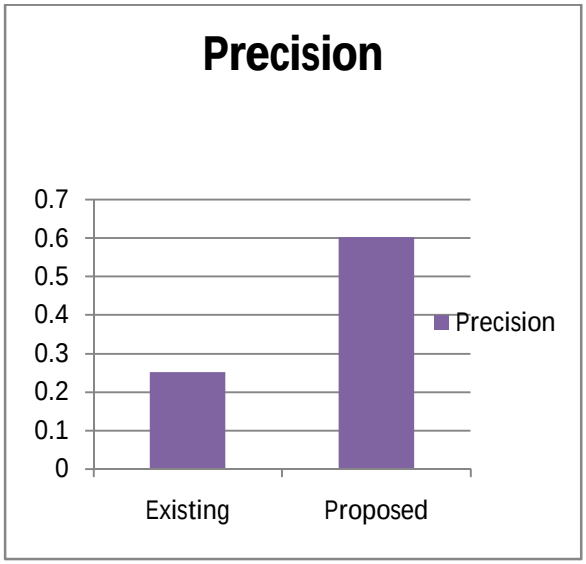
➢ **F-measure :-** A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score:

F = 2.(precision.recall)/(precision+recall)

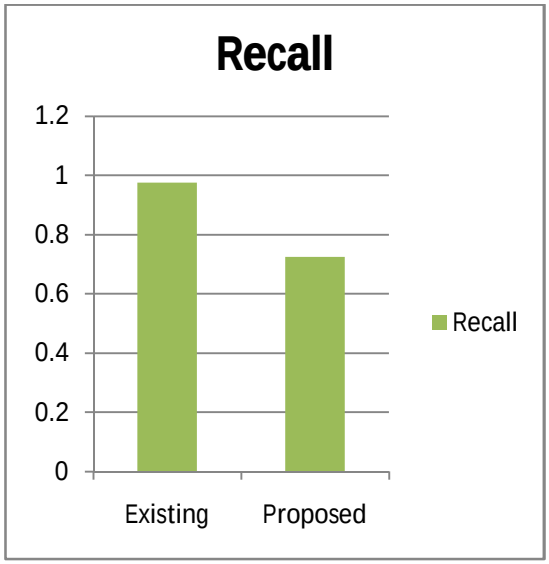
IV. EXPERIMENTS AND RESULTS

Dataset :- To examine the performance of our model, we conducted an experimental study on real-world datasets. We randomly selected 20 users through Social Network. For each user, we collected their profile data that includes author name, location, college, school etc.

Evaluation Metric :- Since users are usually more concerned with top-ranked recommended users, top-ranked users in the results should be rewarded more heavily than those ranked lower.It consists of many factor like precision ,recall, f-measure.



Graph 1. (Precision)



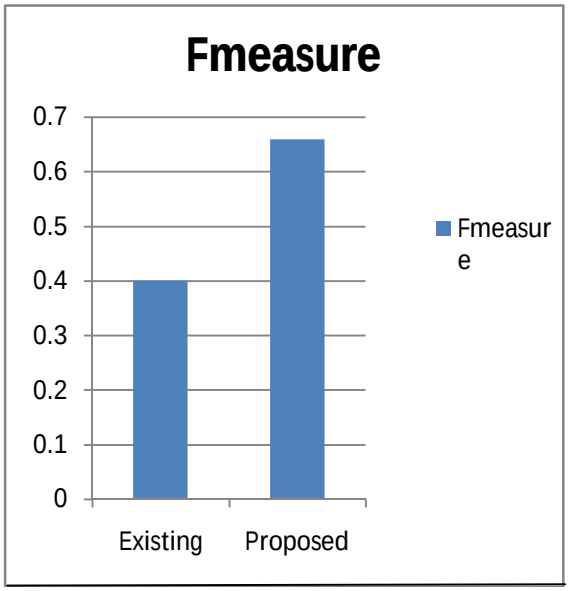
Graph 2. (Recall)

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

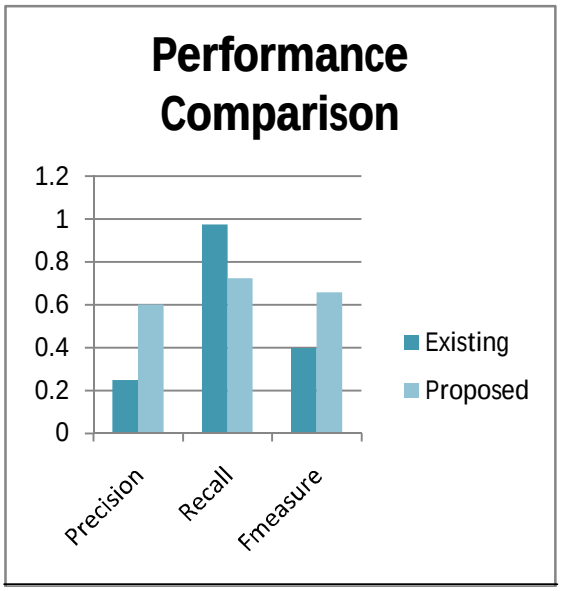
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016



Graph 3.(Fmeasure)



Graph 4. (Performance comparison)

V. CONCLUSION

Friend recommendation is a very essential application in social networking sites.In this paper, We suryed the latest literature review on friend recommender system.We gave brief discussion of friend recommender system and its different techniques that we use in our general life .It mainly focus on the use of recommender system in social network sites for that we used different technologies.In future we will work on how can we recommend friends to another user by using accurate technique,so that we can make friends online.

REFERENCES

[1] Zhi Yu, Can Wang*, Jiajun Bu, Xin Wang, Yue Wu, Chun Chen," Friend recommendation with content spread enhancement in social networks", Information Sciences 309, pp:102-118, 2015.
[2] C. Martinez-Cruz[a,*], C.Porcel[a,*], J.Bernabé-Moreno[b],E.Herrera-Viedma[b,*],"A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling", Information Sciences 311, pp:102-118, 2015.
[3] Zhoubao Sun[a], Lixin Han[a,*], Wenliang Huang[a], Xueting Wang[a], Xiaoqin Zeng[a],Min Wang[a], Hong Yan[b,c]," Recommender systems based on social networks", The Journal of Systems and Software 99 , pp:109-119, 2015.
[4] Hsiu-Yu Liao, Kuan-Yu Chen, Duen-Ren Liu*[4]," Virtual friend recommendations in virtual worlds", Decision Support Systems 69 (2015) , pp:59-69,2015
[5] Zhou Zhang [a], Yuewen Liu [a,*],Wei Ding [b],Wei (Wayne) Huang [a,c], Qin Su[a], Ping Chen[b] ," Proposing a new friend recommendation method, FRUTAI, to enhance social media providers' performance", Decision Support Systems 79 , pp:46-54,2015.
[6] Parham Moradi*, Sajad Ahmadian, Fardin Akhlaghian," An effective trust-based recommendation method using a novel graph clustering algorithm", Physica A 436, pp:462-481, 2015.
[7] NikolaosPolatidis*,ChristosK.Georgiadis," A multi-level collaborative filtering method that improves recommendations", Expert Systems With Applications 48 , pp:100-110, 2016.
[8] Ying Liu[a,b,*] Jiajun Yang[a][8],"Improving Ranking-based Recommendation by Social Information and Negative Similarity", Procedia Computer Science 55 , pp:732-740, 2015.
[9] Parham Moradi *, Sajad Ahmadian," A reliability-based recommendation method to improve trust-aware recommender systems", Expert Systems with Applications 42 , pp:7386-7398, 2015.
[10] Melike Yigit [a], Bilal E. Bilgin [a], Adem Karahoca [b,*]," Extended topology based recommendation system for Unidirectional social networks", Expert Systems with Applications 42 , pp:3653-3661, 2015.
[11] Zhibo Wang, Student Member, IEEE, Jilong Liao, Qing Cao, Member, IEEE et.al," Friendbook: A Semantic Based Friend Recommendation System for Social Networks", IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 14, NO. 3, pp:538-551,MARCH 2015.
[12] Arlina D'cunha, Vandana Patil," Friend Recommendation Techniques in Social Network", 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, India, pp:1-4,Jan. 16-17.

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. 4, Issue 6, June 2016

[13] Sundong Kim supervised by Jae-Gil Lee," Friend Recommendation with a Target User in Social Networking Services", *Korea Advanced Institute of Science and Technology, Daejeon, Korea. pp:* 235-239,Feb, 2018.
[14] Yan Xu, Meilin Zhou, and Siyao Han," Feature Representation for Microblog Followee Recommendation in Classification Framework", 7th International Conference on Advanced Computational Intelligence Mount Wuyi, Fujian, China, pp:318-322, March 27-29, 2015.
[15] Bailing Wang , Junheng Huang et.al," A Collaborative Filtering Algorithm Fusing User-based, Item-based and Social Networks", IEEE International Conference on Big Data (Big Data), pp:2337-2343,2015.