IMPROVING SCALABILITY ISSUES USING GIM IN COLLABORATIVE FILTERING BASED ON TAGGING

Shaina Saini[1] and Latha Banda[2]
Department of Computer Science, Lingaya's University, Faridabad, India

ABSTRACT

The paper deals with improving scalability issues in Collaborative filtering through Genre Interestingness measure approach using Tagging. Due to the explosive growth of data and information on web, there is an urgent need for powerful Web Recommender system (RS). RS employ Collaborative filtering that was initially proposed as a framework for filtering information based on the preferences of users. But CF fails seriously to scale up its computation with the growth of both the number of users and items in the database. Apart from that CF encounters two serious limitations with quality evaluation: the sparsity problem and the cold start problem due to the insufficiency of information about the user. To solve these limitations in our research, we combine many information sources as a set of hybrid sources. These hybrid feaures are utilized as the basis for formulating a Genre Interestingness measure (GIM), we propose a unique approach to provide an enhanced recommendation quality from user created tags. This paper is based on the hybrid approach of collaborative filtering, tagging and GIM approach.

KEYWORDS: *Collaborative Filtering, Collaborative Tagging, Genre Interestingness measure, Recommender system.*

I. INTRODUCTION

With the explosive growth of information in the world, the problem of information overload is becoming increasingly acute. The popular use of web as a global information system has flooded us with a tremendous amount of data and information. Due to this explosive growth of data and information on web, there is an urgent need for powerful automated web personalization tools that can assist us in transforming the vast amount of data into useful information. Web Recommender system (RS) is the most successful example of this tool [1]. In other words, these tools ensure that the right information is delivered to the right people at the right time. Web recommender system tailors information access, trim down the information overload, and efficiently guide the user in a personalized manner to interesting items within a very large space of possible options. Typically RS recommend information (URLs, Netnews articles), entertainment (books, movies, restaurants), or individuals (experts). Amazon.com and MovieLens.org are two well-known examples of RS on the web. Recommender systems employ four information filtering techniques [3].

1. Demographic filtering (DMF) categorizes the user based on the user personal attributes and makes recommendations based on demographic classes.
2. Content-based filtering (CBF) suggests items similar to the ones the user preferred in the past.
3. Collaborative filtering (CF) the user will be recommended items people with similar tastes and preferences liked in the past. Group Lens, Movie Lens is some examples of such systems.
4. Hybrid filtering techniques combine more than one filtering technique to enhance the performance like Fab and Amazon.com.

Collaborative filtering (CF) is the most successful and widely used filtering technique for recommender systems. It is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets But CF fails seriously to scale up its computation with the growth of both the number of users and items in the database. Apart from that CF encounters two serious limitations with quality evaluation: the sparsity problem and the cold start problem due to the insufficiency of information about the user. This problem leads to the great scalable challenge for collaborative filtering. A sparse user item matrix causes a Scalability problem for CF. A number of studies have attempted to address problems related to collaborative filtering. To solve these limitations, in our research, we propose a new and unique approach to provide an enhanced recommendation quality derived from user-created tags. Tagging is the process of attaching natural language words as metadata to describe some resource like a movie, photo, book, etc. The proposed approach first determines Similarity between the users created tag. This paper presents the unique approach named as "Genre Interestingness measure". This is a specific contribution toward recommender system.

The rest of this paper is organised as follow: section II describes the problem formulation. Section III describes an overview of related work. Section IV describes the detailed overview of methodology of our proposed work. In section V, the Experiment Performed and results part is described. This presents the effectiveness of our approach. Finally we mention the conclusion and future scope of this paper.

II. PROBLEM FORMULATION

This part mainly contains the Need and Significance of proposed research work. Most recommendation systems employ variations of Collaborative Filtering (CF) for formulating suggestions of items relevant to users' interests. There are various types of problem occurring in CF [2].

1. The Scalability Challenge for Collaborative Filtering- CF requires expensive computations that grow polynomially with the number of users and items in the database.
2. The sparsity problem- It occurs when available data is insufficient for identifying similar users or items (neighbors) due to an immense amount of users and items. In practice, even though users are very active, each individual has only expressed a rating (or purchase) on a very small portion of the items. Likewise, very popular items may have been rated (or purchased) by only a few of the total number of users. Accordingly, it is often the case that there is no intersection at all between two users or two items and hence the similarity is not computable at all.
3. Cold start problem- This problem can be divided into cold-start items and cold-start users. A cold-start user, the focus of the present research, describes a new user that joins a CF-based recommender system and has presented few opinions. With this situation, the system is generally unable to make high quality recommendations.

In order to enhance the efficiency of Recommendations on the web, it is very necessary to propose the solution of above written problems. A number of studies have been attempted to address problems related to collaborative filtering. For improving the scalability issue, we develop a set of hybrid features that combines one of the user and item properties. These features are based on Genre Interestingness Measure (GIM).This is described in the Section IV of this paper. To solve these limitations, in our research, we propose a new and unique approach to provide an enhanced recommendation quality derived from user-created tags.

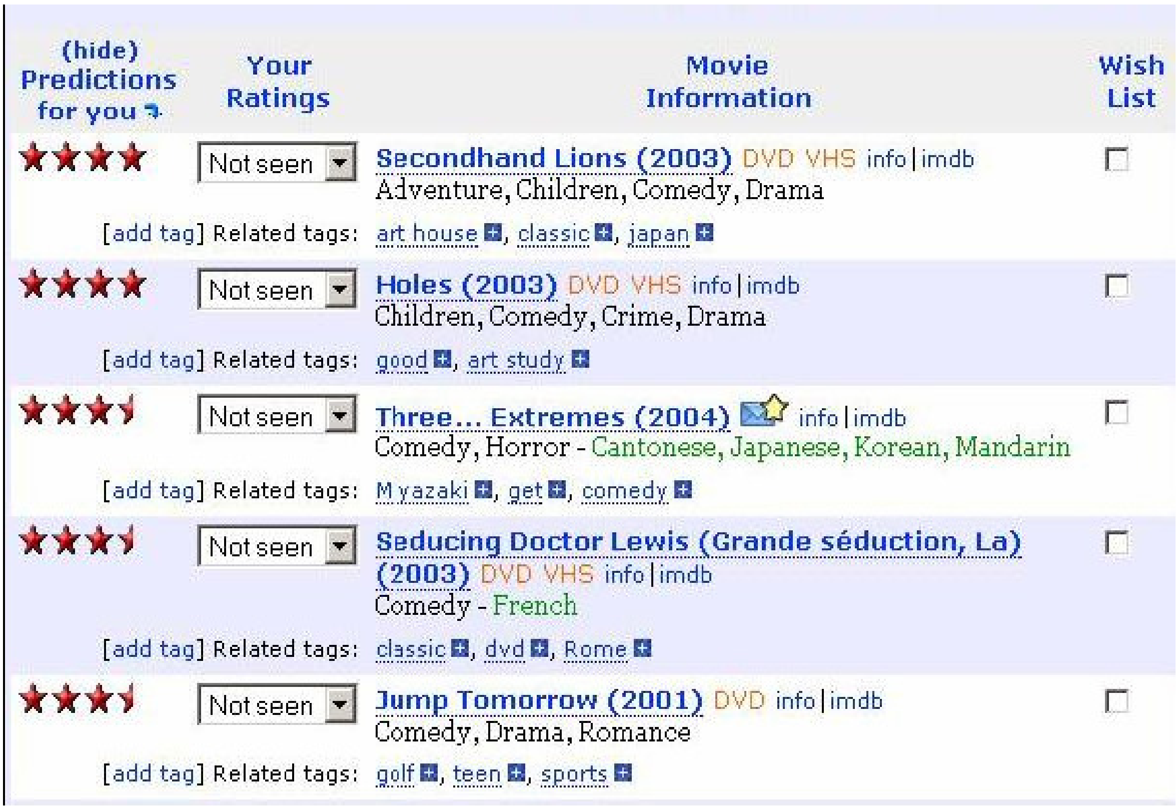
Collaborative tagging, which allows many users to annotate content with descriptive keywords (i.e., tags) is employed as an approach in order to grasp and filter users' preferences for items. Tagging is not new, but has recently become useful and popular as one effective way of classifying items for future search, sharing information, and filtering. In terms of user-created tags, they imply users' preferences and opinions about items as well as Meta data about them. For this purpose we are taking the data set from the site movielens.com. There are Four types of data set are used- User data, Movie data, Rating data, Tag data. Therefore, by using the collaborative filtering based on collaborative tagging and Genre Interestingness Measure (GIM) approach, we can improve the scalability issues.

III. RELATED WORK

In this section background knowledge of collaborative filtering, Collaborative tagging and their similarity measure are introduced.

3.1. Collaborative Filtering

One of the potent personalization technologies powering the adaptive web is collaborative filtering. It is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets. CF technology brings together the opinions of large interconnected communities on the web, supporting filtering of substantial quantities of data. For example Movie Lens is a collaborative filtering system for movies. A user of Movie Lens rates movies using 1 to 5 stars, where 1 is "Awful" and 5 is "Must See". Movie Lens then uses the ratings of the community to recommend other movies that user might be interested in (Fig. 1), predict what that user might rate a movie, or perform other tasks [4].



(hide) Predictions for you	Your Ratings	Movie Information	Wish List
★★★★	Not seen ▾	Secondhand Lions (2003) DVD VHS info | imdb Adventure, Children, Comedy, Drama	□
	[add tag] Related tags:	art house ⊞, classic ⊞, japan ⊞	
★★★★	Not seen ▾	Holes (2003) DVD VHS info | imdb Children, Comedy, Crime, Drama	□
	[add tag] Related tags:	good ⊞, art study ⊞	
★★★½	Not seen ▾	Three... Extremes (2004) info | imdb Comedy, Horror - Cantonese, Japanese, Korean, Mandarin	□
	[add tag] Related tags:	M yazaki ⊞, get ⊞, comedy ⊞	
★★★½	Not seen ▾	Seducing Doctor Lewis (Grande séduction, La) (2003) DVD VHS info | imdb Comedy - French	□
	[add tag] Related tags:	classic ⊞, dvd ⊞, Rome ⊞	
★★★½	Not seen ▾	Jump Tomorrow (2001) DVD info | imdb Comedy, Drama, Romance	□
	[add tag] Related tags:	golf ⊞, teen ⊞, sports ⊞	

Figure 1: Movie Lens uses collaborative filtering to predict that this user is likely to rate the movie "Holes" 4 out of 5 stars

3.1.1. Types of Collaborative Filtering

There are two types of collaborative filtering.

1. Memory based collaborative filtering- This mechanism uses user rating data to compute similarity between users or items. This is used for making recommendations. This was the earlier mechanism and is used in many commercial systems. It is easy to implement and is effective. Typical examples of this mechanism are neighborhood based CF and item-based/user-based top-N recommendations. The neighborhood-based algorithm calculates the similarity between two users or items produces a prediction for the user taking the weighted average of all the ratings. Multiple mechanisms such as Pearson correlation and vector cosine based similarity are used for this [5].
2. Model based collaborative filtering- Models are developed using data mining, machine learning algorithms to find patterns based on training data. These are used to make predictions for real data. There are many model based CF algorithms. These include Bayesian Networks, clustering models,

latent semantic models such as singular value decomposition, probabilistic latent semantic analysis. This approach has a more holistic goal to uncover latent factors that explain observed ratings. Most of the models are based on creating a classification or clustering technique to identify the user based on the test set. The number of the parameters can be reduced based on types of principal component analysis.

3.2. Collaborative Tagging and Folksonomy

Collaborative tagging describes the process by which many users add metadata in the form of keywords to shared content. Tagging advocates a grass root approach to form a so called"Folksonomy", which is neither hierarchical nor exclusive. With tagging, a user can enter labels in a free form to tag any object; it therefore relieves users much burden of fitting objects into a universal ontology. Meanwhile, a user can use a certain tag combination to express the interest in objects tagged by other users, e.g., tags (renewable, energy) for objects tagged by both the keywords renewable and energy [7]. Recently, collaborative tagging has grown in popularity on the web, on sites that allow users to tag bookmarks, photographs and other content. The paper analyses the structure of collaborative tagging systems as well as their dynamical aspects. Specifically, we discovered regularities in user activity, tag frequencies, kinds of tags used, bursts of popularity in book marking and a remarkable stability in the relative proportions of tags within a given URL. We also present a dynamical model of collaborative tagging that predicts these stable patterns and relates them to imitation and shared knowledge.

3.3. Neighborhood formation using tagging

The most important task in CF-based recommendations is the similarity measurement because different measurements lead to different neighbor users, in turn, leading to different recommendations. Since the user–item matrix R is usually very sparse, which is one of the limitations of CF, it is often the case that two users do not share a sufficient number of items selected in common for computing similarity. For this reason, in our research, we select the best neighbors, often called k nearest neighbors, with tag frequencies of the corresponding user in the user–tag matrix, A. In order to find the k nearest neighbor (KNN), cosine similarity, which quantifies the similarity of two vectors according to their angle, is employed to measure the similarity values between a target user and every other user.

KNN includes users who have a higher similarity score than the other users and means a set of users who prefer more similar tags with a target user. In cosine similarity between users, two users are treated as two vectors in the m- dimensional space of tags. In addition, we also consider the number of users for tags, namely the inverse user frequency. Consider two tags, t1 and t2, both having been tagged by user u and v; however, just 10 users used tag t1, whereas 100 users used tag t2. In this situation, tag t1, tagged by fewer users, is relatively more reliable for the similarity of user u and v than tag t2 tagged by many users. Likewise with the inverse document frequency, the main idea is that tags used by many users present less contribution with regard to capturing similarity, than tags used by a smaller number of users [2].

IV. PROPOSED MODEL

The framework of our proposed model is shown in Figure 2. The detail of each part in the model is illustrated below [6]:

The first phase of this section contains the collaborative filtering based on collaborative tagging. For improving the scalability issue, we develop a set of hybrid features that combines one of the user and item properties. These features are based on Genre Interestingness Measure (GIM). The next phase contains the similarity computattion of the user-item matrix and user-tag matrix. After this prediction and Recommendation is done. The rest of this section contains the Testing phase and this phase is accomplished by MAE analyis. This shows the final result of this proposed work.

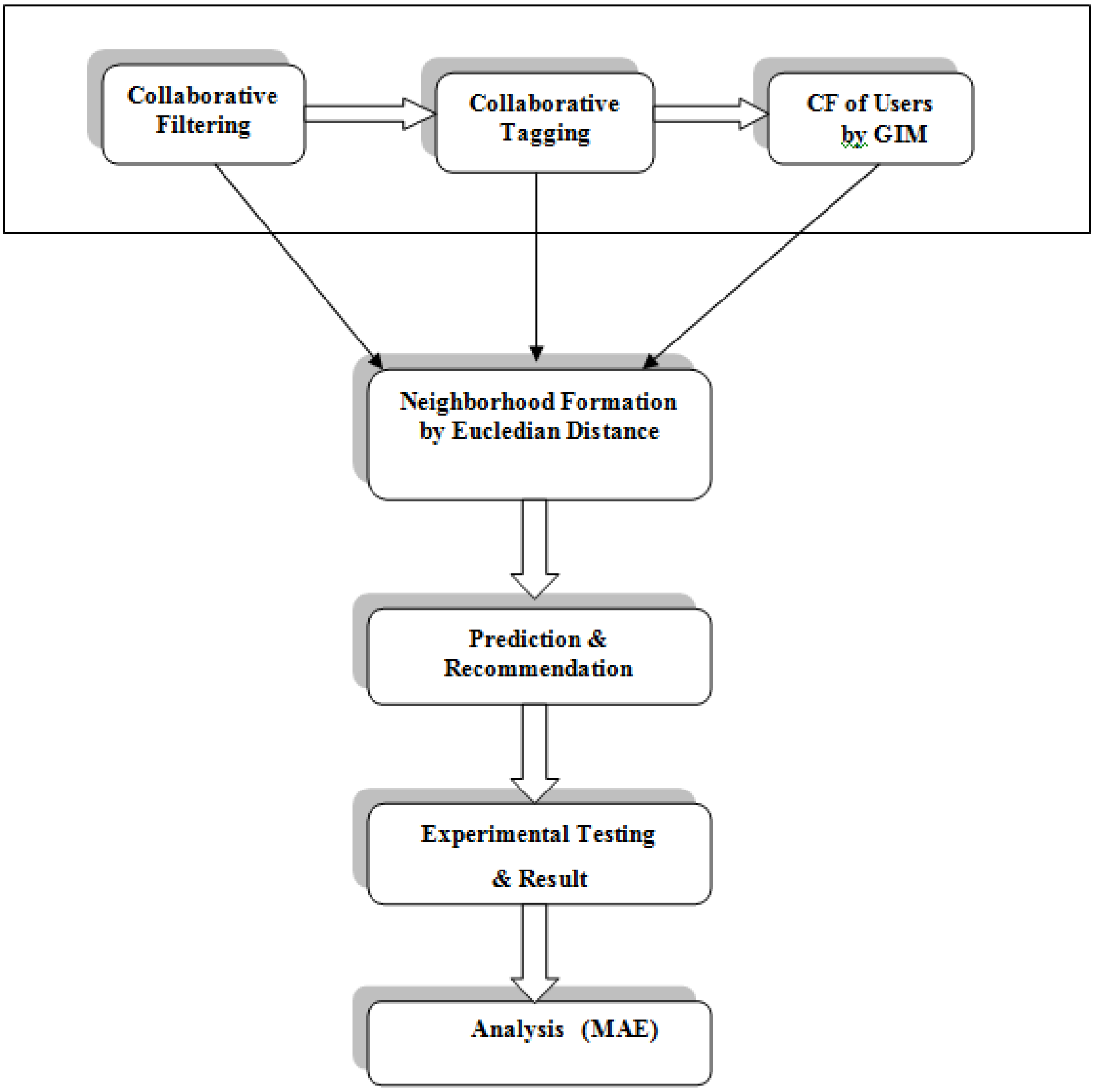

Collaborative Filtering
Collaborative Tagging
CF of Users by GIM
Neighborhood Formation by Eucledian Distance
Prediction & Recommendation
Experimental Testing & Result
Analysis (MAE)

Figure 2: Proposed Model

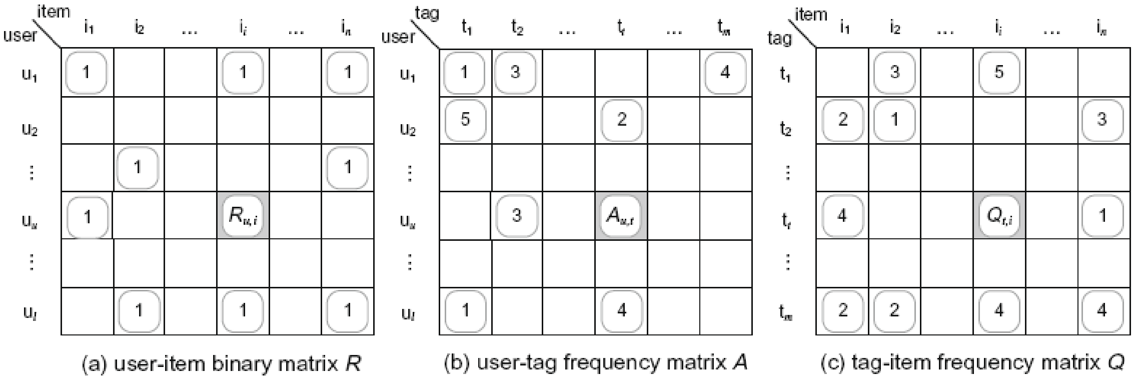
4.1. Collaborative filtering based on collaborative tagging

As mentioned above, it is the starting phase of Figure 2.This phase cotains the three matrices which are described as follow.

1. User–item binary matrix, R- If there is a list of l users U={u1,u2,…,ul}, a list of n items I={i1,i2,…,in}, and a mapping between user–item pairs and the opinions, user–item data can be represented as a l × n binary matrix, R, referred to as a user–item matrix. The matrix rows represent users, the columns represent items, and Ru,i represents the historical preference of a user u on an item i. Each Ru,i gis set to 1 if a user u has selected (or tagged) an item i or 0 otherwise [2].

2. User–tag frequency matrix, A- For a set of *m* tags *T*= {t_1, *t2,...tm*}, tag usages of *l* users can be represented as a *l* × *m user–tag matrix*, *A*. The matrix rows represent users, the columns represent tags, and $A_{u,t}$ represents the number of items that a user *u* has tagged with a tag *t*.

3. Tag–item frequency matrix, Q- This is a *m* × *n* matrix of tags against items that have as elements the frequencies of tags to items. The matrix rows represent tags, the columns represent items; and $Q_{t,i}$ implies the number of users who have tagged an item *i* with a tag *t*.



(a) user-item binary matrix *R* (b) user-tag frequency matrix *A* (c) tag-item frequency matrix *Q*

Figure 3: Three matrices for a tag based collaborative filtering system

4.2. Genre Interestingness Measure

It is a vector representation of active users with their respected genres. This can be explained as per the following chart.It is a new innovative approach under which, mappings of the user data and their particular genre rarings is occur. The Genre Feature Specifies if the movie is an action, adventure, comedy, crime, animation, horror and so on. There are total 18 genres in our data set. For this approach, a user gives the rating to a particular item (movie) and point out the genre means which one genre is present in this movie. For example- Either a movie comedy based or action based and so on. A single movie can belong to more than one genre.In the following chart, a user as u1 specifies the genres present in item 1 say movie1.The '*' symbol is used for the presence of a genre of the movie. Similarly the same user gives the rating to the second movie and it is repeated up to ten movies, at the last, when we squeeze these vectors of ten movies, then it realized that for user u1 the G1 and G16 genres are present. Similary same procedures are used for user u2, u2, up to u10. After suqueezing these vectors of ten users the final result produes as a big matrix as shown in figure 4. This matrix shows the binary mapping in between the user and the genre.

Rating of user u1 for item1- * * * * * * * * * * * * *

u1	**sex**	**age**	**occupation**	**1**	**2**	**..…..................................**	**17**	**18**

Rating of user u1 for item2- * *

u1	**sex**	**age**	**occupation**	**1**	**2**	**......................................…**	**17**	**18**

Up to item 10

Rating of user u1 for all item

u1	**sex**	**age**	**occupation**	**2**	**16**	**…........**

Similarly for user u2 * *

u1	**sex**	**age**	**occupation**	**1**	**2**	**…**	**17**	**18**

Similarly for user u3 …………………… & Up to

Similarly for user u10…………………

Figure 4: Vector representation of GIM approach

The above procedure is used for making a matrix which shows the mapping in between the different users and their particular genre. The "1" is used for the presence of a particular genre and "0" is used for the absence.

	G1	G2			.G18
U1	1	0	1	0	1	0
U2	1	1	0	1	0	1
U3						
U4						
U5	1			1		0
U6						
U7					0	
U8		1				0
U9			1			
U10	1			0	0	

Figure 5: GIM Matrix showing Genre Interestingness Measure

4.3. Neighborhood Formation

Neighbors simply mean a group of likeminded users with a target user or a set of similar items with the items that have been already been identified as being preferred by the target user. The most important task in CF-based recommendations is the similarity measurement because different measurements lead to different neighbor users, in turn, leading to different recommendations. Since the user–item matrix R is usually very sparse, which is one of the limitations of CF, it is often the case that two users do not share a sufficient number of items selected in common for computing similarity. For this reason, in our research, we select the best neighbors, often called k nearest neighbors, with tag frequencies of the corresponding user in the user–tag matrix, A. There are various methods for similarity computation [3].

1. The neighborhood formation of user Tag matrix is done by Cosine Similarity- Let l be the total number of users in the system and n_t the number of users tagging with a tag t. Then, the inverse user frequency for a tag t, iuf_t, is computed: $iuf_t = \log(l/n_t)$. If all users have tagged using tag t, then the value of iuf_t is zero, $iuf_t = 0$. When the inverse user frequency is applied to the cosine similarity technique, the similarity between two users, u and v, is measured by the following equation (1).

$$sim(u,v) = \cos(\vec{u},\vec{v}) = \frac{\sum_{t\in T}(A_{u,t}\cdot iuf_t)(A_{v,t}\cdot iuf_t)}{\sqrt{\sum_{t\in T}(A_{u,t}\cdot iuf_t)^2}\sqrt{\sum_{t\in T}(A_{v,t}\cdot iuf_t)^2}} \qquad (1)$$

Users u and v are in user–tag matrix, A. In addition, iuft refers to the inverse user frequency of tag t. The similarity score between two users is in the range of [0, 1]. The higher score a user has, the more similar he/she is to a target user [2].

2. The neighborhood formation of user item matrix is done by the formula of Euclidean distance**.** It is given by the following equation (2)

$$d(\mathbf{x},\mathbf{y}) = \frac{1}{z}\sum_{i=1}^{z}\sqrt{\sum_{j=1}^{N}(x_{i,j}-y_{i,j})^2}. \qquad (2)$$

Here $x_{i,j}$ is the jth feature for the common item s_i, N is the number of features, and $z = |S_{xy}|$, the cardinality of S_{xy} [3].

4.4. Predictions and Recommendations

In this phase, RS assign a predicted rating to all the items seen by the neighborhood set and not by the active user. The predicted rating, $pr_{a,j}$, indicates the expected interestingness of the item s_j to the user u_a, is usually computed as an aggregate of the ratings of user's (u_a) neighborhood set for the same item s_j

$$pr_{a,j} = aggr_{u_c \in C}\, r_{c,j}, \qquad (3)$$

Where C denotes the set of neighbors who have rated item s_j.The most widely used aggregation function is the weighted sum[1] which is called also Resnick's prediction formula.

$$pr_{a,j} = m_a + k \sum_{u_c \in C} d(a,c) \times (r_{c,j} - m_c). \qquad (4)$$

The multiplier k serves as a normalizing factor [3].

4.5. Experimental Testing

For this phase movielens dataset are used. In this phase, "Ten-fold cross validation" scheme is used. Cross-validation, sometimes called rotation estimation, is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. One round of cross-validation involves partitioning a sample of data into complementary subsets, performing the analysis on one subset (called the training set), and validating the analysis on the other subset (called the validation set or testing set). To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the rounds. For e.g.

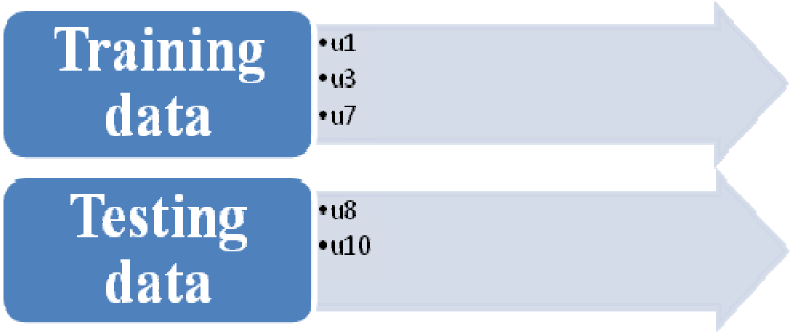


Figure 6: Shows Testing Pattern

The set of training users is used to find a set of neighbors for the active user while the set of active users (50 users) is used to test the performance of the system. During the testing phase, each active user's ratings are divided randomly into two disjoint sets, training ratings (34%) and test rating (66%). The training ratings are used for overall implementation.

4.6. MAE (Mean Absolute Error)

The MAE measures the deviation of predictions generated by the RS from the true ratings specified by the user. The MAE for active user ui [3] is given by the following formula:

$$\mathrm{MAE}(i) = \frac{1}{n_i} \sum_{j=1}^{n_i} |pr_{i,j} - r_{i,j}|,$$

Final result = [MAE $_{(CF)}$ + MAE $_{(CT)}$ + MAE $_{(GIM)}$] /3.

Lower the MAE corresponds to correct predictions of a given RS. This leads to improvement of Scalability.

V. RESULTS AND DISCUSSIONS

This section contains the experiment conducted, their final outcomes and analysis of the result. We conduct several experiments to examine the effetiveness of our new scheme for Collaborative Filtering based on Collaborative Tagging using Genre Interestingness measure in terms of scalability and recommendation quality.

5.1. Data Set

As we know that the experimental data comes from the movielens website. Based on MovieLens dataset we considered 500 users who have rated at least 40 movies, for each movie dataset, we extracted subset of 10,000 users with more than 40 ratings. To compare these algorithms, we experimented with several configurations. For MovieLens dataset the training set to be the first 100, 200 and 300 users. Such a random separation was intended for the execution of ten folds cross validation where all the experiments are repeated ten times for 100 users, 200 users and 300 users. For movie Lens we the testing set 30% of all users.

5.2. Experiment Performed

I. Find out the MAE of collaborative filtering, Tagging and GIM denoted as MAE $_{(CF),}$ MAE $_{(CT)}$ and MAE $_{(GIM).}$

II. We take average value of MAE $_{(CF),}$ and MAE $_{(CT),}$ it is denoted as MAE $_{(CFT).}$

$$CFT = (CF+CT)/2$$

III. We take average of MAE $_{(CF),}$ MAE $_{(CFT),}$ MAE $_{(GIM),}$ It is denoted as MAE $_{(CFTGIM),}$

$$\text{Final value}= (CF+CFT+ GIM)/3 = CFTGIM$$

5.3. Performance

As we mentioned above, our algorithm could solve the problem of scalability. In order to show the performance of our approach, we compare the MAE of Collaborative Filtering, collaborative filtering based on collaborative tagging and collaborative tagging with Genre Interestingness Measure.

Table 1: MAE of CF, CFT, and CFTGIM for 100 users

	MAE		
No. of users	**CF**	**CFT**	**CFTGIM**
10	0.974	0.873	0.841
20	0.961	0.848	0.832
30	0.948	0.828	0.818
40	0.935	0.819	0.801
50	0.898	0.812	0.792
60	0.896	0.808	0.784
70	0.892	0.806	0.781
80	0.886	0.804	0.778
90	0.883	0.802	0.775
100	0.881	0.801	0.772

The results of these three methods are shown in Table 1. It has been clearly shown in the table that the CFTGIM has lower range of MAE as compare to other two i.e.CF and CFT. Collaborative tagging with genre interestingness measure outperforms other two methods in respective MAE and prediction accuracy.

5.4. Analysis of the Results

In this experiment we run the proposed collaborative tagging using genre interestingness measure and compare its results with classical Collaborative filtering and collaborative filtering based on collaborative tagging. After implementation of the proposed approach, we analyzed that MAE for collaborative filtering based on tagging with genre interestingness measure (CFTGIM) is lower than other two methods. The results summerized in the table are plotted as shown in figure 6. From this graph it has been clearly shown that the third one approach i.e CFTGIM always has lower MAE values as compare to traditional CF and CFT approaches. Lower MAE corresponds to more accurate predictions of a given RS.

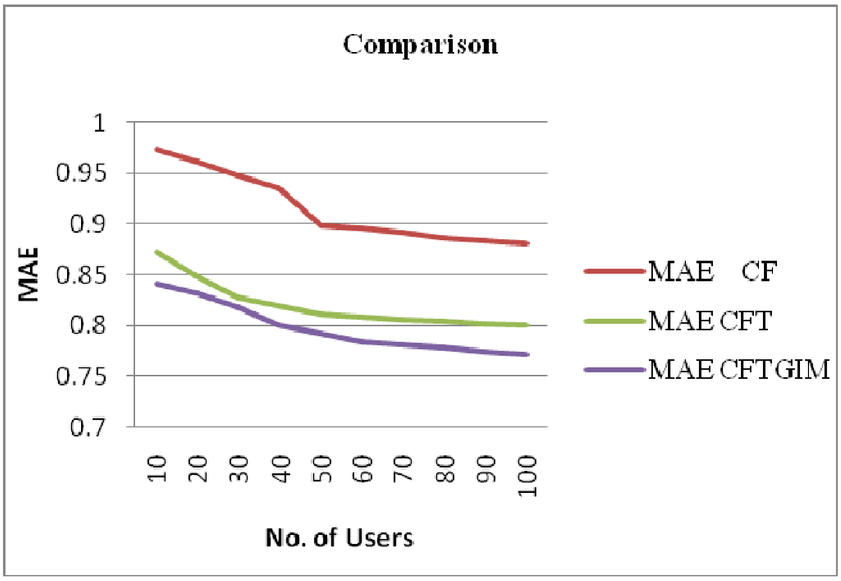


Figure 7: Comparison between MAE variations of three techniques

VI. CONCLUSION AND FUTURE SCOPE

This work has a considerable reduction in the complexity of recommender system (RS). As we know that these complexity is caused by various problem occurring in Collaborative filtering (CF). In order to solve these problems, this paper represents the integration of collaborative filtering, Collaborative tagging and Genre interestingness measure approach. In this paper we analyse the potential of Collaborative tagging to overcome the problem of data sparseness and cold start user. By finding out the MAE of these techniques one by one respectively, we merge up their final outcome. This produces the less error as compared to already present model. This approach makes the system more scalable by reducing the error and thus enhancing the recommendation quality. In the future work, we would like to prform this experiment with more accuracy and consideration according to user's interest. We will work on trust reputation for addressing the Collaborative Tagging (CT) with GIM in the future.

ACKNOWLEDGEMENTS

I would like to express my most sincere appreciation to Ms. Latha Banda, Associate Prof., CSE Dept., Lingaya's University, for their valuable guidance and support for completion of this project.

REFERENCES

[1]. Adomavicius, Tuzhilin, (2005) "Toward the next generation of recommender systems: A survey of the state-of –the-art and poaaible extensions", *IEEE Transaction on Knowledge and Data Engineering, 17(6), 734-749.*

[2]. Heung-Nam Kim, Ae-Ttie Ji, Inay Ha, Geun-Sik Jo, (2010) “Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation”, *ELSEVIER Electronic Commerce Research and Applications 9, 73–83.*

[3]. Mohammad Yahya H. Al-Shamri, Kamal K. Bharadwaj, (2008) “Fuzzy-genetic approach to recommender systems based on a on hybrid user model, ” *ELSEVIER Expert Systems with Applications 35, 1386–1399.*

[4]. Zheng Wen, (2008) “Recommendation System Based on Collaborative Filtering”.

[5]. Buhwan Jeong & Jaewook Lee, (2010) “Improving memory-based collaborative filtering via similarity updating and prediction modulation, ”, *in ELESVIER.*

[6]. Shaina Saini, Latha Banda, (2012) “Enhancing Recommendation Quality by using GIM in Tag based Collaborative Filtering,” *In Proceedings of the National Technical Symposium on “Advancement in Computing Technologies (NTSACT)”, Published by Bonfring ISBN 978-1-4675-1444-6.*

[7]. Zhichen Xu, Yun Fu, Jianchang Mao, “Towards the Semantic Web: Collaborative Tag Suggestions,” *Inc2821 Mission College Blvd., Santa Clara, CA 95054.*

AUTHORS PROFILE

Shaina Saini received her bachelor’s degree in Computer Science from M.D University, Haryana and master’s degree in Computer Science from Lingaya’s university Faridabad. Her areas of interests include Web Mining, Multimedia Technology etc.

Latha Banda received her bachelor’s degree in CSE from J.N.T University, Hyderabad, master’s degree in CSE from I.E.T.E University, Delhi and currently pursuing her Doctoral Degree. She has 9 years of experience in teaching. Currently, she is working as an Associate Professor in the Dept. of Computer Sc. & Engg. at Lingaya’s University, Faridabad. Her areas of interests include Data Mining, Web Personalization, and Recommender System.