

Community Expert based Recommendation for solving First Rater Problem

Akshi Kumar
Dept. of Computer Engineering,
Delhi Technological University, India

MPS Bhatia
Division of Computer Engineering,
Netaji Subhas Institute of Technology, India

ABSTRACT

Information overload on the Web is a well recognized problem [1], where users find it increasingly difficult to locate the right information at the right time. Recommender system [2, 3] comes to the rescue for such a consumer. However, despite all advances, the current generation of recommender systems still requires further improvements to make recommendation methods more effective and applicable to an even broader range of real-life applications. We propose and investigate CER system, a *Community Expert based Recommendation system*. The paradigm is realized by an Interest Mining module which defines a constructing algorithm for Interest Group by uncovering shared interest relationships between people, using their blog document entries and interest similarity relations. Once the interest similarity group is constructed, then we identify an expert from each of the groups so formed. Expert identification from the Collaborative Interest Group is the key to recommendation as it is only the expert's blog whose recommendation is considered compared to systems which require a large set of customer preferences for predicting the new preferences accurately for effective Collaborative filtering-based recommendation, solving the most prominent problem existent in collaborative filtering, the *First-Rater or the cold-start problem*. The initial results show that the CER is a motivating technique.

Keywords

Recommender Systems, First-Rater, Interest Group

1. INTRODUCTION

According to worldwidewebsize.com, the indexed Web contains at least 8.28 billion pages (Wednesday, 14 December, 2011). Given these numbers, it is evident that the complexity of finding relevant information on the Web has become increasingly intricate and crucial. In fact, "information overload" on the Web is a well recognized problem, where users find it increasingly difficult to locate the right information at the right time. As a response to this problem, much research has been done with the goal of providing users with more proactive and personalized information services. These techniques either lie within the fields of *information retrieval* and *information filtering*. While information retrieval systems filter information by letting users specify explicitly what information is needed, information filtering systems strive to adapt the user's long-term interests and filter information based on user profiles. Closely related to information filtering is the idea of having systems that act as personalized decision guides for users. *Recommender systems* have proved to help achieving this goal

by using the opinions of a community of users to help individuals in the community more effectively identify content of interest from a potentially overwhelming set of choices [4]. Recommender systems help in addressing this information overload problem by retrieving the information desired by the user based on his/her similar users' tastes and preferences. Examples of such applications include recommending books, CDs and other products at Amazon.com [5], movies by MovieLens [6].

Two recommendation strategies that have come to dominate are content-based and collaborative filtering. Content-based filtering relies on rich content descriptions of the items that are being recommended [7], while collaborative filtering recommendations are motivated by the observation that we often look to our friends for recommendations [8]. Seminal collaborative filtering systems included GroupLens [9] and Firefly [10]. Even though collaborative filtering often performs better than content-based filtering when lots of user ratings are available, it has limitations [11, 12, 13] namely the First Rater /Cold Start problem, Sparsity problem & Scalability issues. The cold-start problem [13] is based on the fact that collaborative filtering based systems require a large set of customer preferences for predicting the new preferences accurately, i.e., it is where no historical ratings on items or users are available..

Collaborative filtering-based systems require a large set of customer preferences for predicting the new preferences accurately. That is, for a reliable collaborative filtering-based prediction system, we need to have a large set of customer ratings on a variety of items available on the database. As the system can predict accurately only after it has gathered a large set of opinions, customers will not be willing to express their detailed preferences in the initial stages (if it is not going to help them in any way, despite offering preference ratings).

Ongoing increase in wide-area network connectivity promise vastly augmented opportunities for collaboration and resource sharing. Now-a-days, various social networking sites like Facebook, Orkut, MySpace, YouTube have gained so much popularity and we cannot ignore them. They have become one of the most important applications of Web 2.0 [14]. They allow people to build connection networks with other people in an easy and timely way and allow them to share various kinds of information and to use a set of services like picture sharing, blogs, wikis etc.



Fig 1: Web 2.0 Technologies [15]

Summing up, the operative challenges motivating researchers in Web 2.0 setting and recommendation system call for convergence between the two areas of active research. The motivation for the research presented in this research was instigated by the challenges related to the information overload problem being faced by the users in their day-to-day activities and the first-rater problem, the most prominent problem existent in collaborative filtering. It calls for proposing methods that allow users with similar interests and expertise to collaborate among them, and thereby carry out their collaborative work and resource sharing in an efficient manner. Because of this, it is of key interest to set up a collaborative interest group (virtual community) and then identify an expert from each group who will be the centre of the group, and finally providing relevant recommendation to the users in the form of positive, negative or neutral opinion based on the reviews of the expert in his respective blog. Expert identification in online communities is of importance for the following two reasons:

- Online communities can be viewed as knowledge databases where knowledge is accumulated by interactions between the members. That is, we read articles in online communities to get information on specific topics. If we find articles written by experts, we tend to have more confidence in their content.
- In terms of communication dynamics, online communities are spaces where non-experts can communicate with experts. In the real world, communicating with experts is not only difficult but also expensive. However, we can relatively easily communicate with experts in online communities if we know who they are.

The proposed CER system realizes a **Community Expert based Recommendation system**. In the proposed system firstly we build an interest similarity group, an online community which is a virtual space where people who are interested in a specific topic gather and discuss in depth a variety of sub-topics related to the topic using blogs. Although community members can equally discuss with other members, there are members across the expertise spectrum from non-experts to experts. Expert identification involves finding experts on a given topic. Thus, once the group is constructed, as our next step we identify an expert from each of the group. Consequently, in the proposed CER system, it's the opinion of the identified expert within a virtual community built on

shared interest that constitutes the recommendation. Eventually this paradigm helps to overcome the most prominent problem existent in collaborative filtering setting, the *First-Rater or the cold-start problem*, as in our proposed system it is only the expert whose recommendation is considered compared to systems which require a large set of customer preferences for predicting the new preferences accurately for effective Collaborative filtering-based recommendation. These techniques should be easily pluggable into current & future systems for recommendation with a minimal need to make architectural changes or modifications.

2. THE CER SYSTEM

The proposed CER system put forwards a constructing algorithm of Interest Group by uncovering shared interest relationships between people, based on their blog document entries, to let them arrange into groups effectively, to let them share the resources, carry out cooperative work. This framework solves the problem of discovering people who have particular interests or expertise and then providing relevant recommendation to a particular user by retrieving the information as desired by him based on his similar user's tastes and preferences. Eventually this paradigm helps to overcome the most prominent problem existent in collaborative filtering setting, the *First-Rater or the cold-start problem*, as in our proposed system it is only the expert whose recommendation is considered compared to systems which require a large set of customer preferences for predicting the new preferences accurately for effective Collaborative filtering-based recommendation.

The following sub-sections expound the details of the CER system:

2.1. Interest Mining Module

In this module, we focus on the problem of discovering people who have particular interests or expertise. The usual approach is to build interest group lists from explicitly registered data. However, doing so assumes one knows what lists should be built, and who should be included in each list. We present an alternative approach, which can support a finer grained and dynamically adaptive notion of shared interests. Our approach deduces shared interest relationships between people based on interest similarity calculated by the means of entries written on their blog. Using this approach, a user could search for people by requesting a list of people whose interests are similar to several people known to have the interest in question. We propose an Interest Group construction algorithm based on interest similarity, which can cluster researchers with similar interests into the same group and facilitate collaborative work.

The **Collaborative Interest Group** is defined as a group which consists of researchers in the similar area or with shared interests. So, when constructing a group, we try to arrange the researchers with prominent interest similarity into the same group.



Fig 2: Collaborative Interest Group

The following sub-sections expound the details of the Collaborative Interest Group construction [16]:

2.1.1 Interest Vector

Each researcher writes blog entries according to his or her interest. Thus, it can be supposed that terms related to the researcher's interests are present in many entries in his or her blog site. The interest vector of the researcher, V_i , is represented as a bag-of-words with frequently used words being assigned high weights. The interest vector is calculated by the equation described below:

$$V_i = (s_{i1}, s_{i2}, s_{i3}, \dots) \quad 3.1$$

And;

$$s_{ik} = ef_i(w_k) \times \log\left(\frac{N_u}{uf(w_k)}\right) \quad 3.2$$

where s_{ik} means the strength of interest in word w_k ; $ef_i(w_k)$ means the number of entries containing w_k in researcher i 's site; $uf(w_k)$ means the number of researchers who use w_k ; and N_u means the number of researchers. This equation corresponds to the traditional *tf-idf* weighting approach. The entry frequency, $ef_i(w_k)$, corresponds to *tf*, and inverse user frequency, $N_u/uf(w_k)$, corresponds to *idf*. Thus, a word repeatedly used in a small number of blog sites has high weight value.

2.1.2 Interest Similarity Score

A similarity score represents how similar the interests of a pair of researchers are. If researcher i and j have similar interests, their interest vectors should be similar. Thus, we calculate the similarity score between them, R_{ij} , using the cosine similarity of V_i and V_j as described below.

$$R_{ij} = \frac{V_i \times V_j}{|V_i| |V_j|} \quad 3.3$$

All elements of V_i and V_j are positive and thus the range of R_{ij} is 0 to 1.

2.1.3 Collaborative Interest Group Construction

Construction of an interest group is done to cluster the researchers with similar interests into the same group and facilitate collaborative work. Collaborative Interest Group Construction is done by using the technique of K-means clustering algorithm [17] where K is a user-specified parameter and it refers to the total number of clusters required.

Each point is then assigned to the closest centroid, and each collection of points assigned to a centroid is a cluster. The centroid of each cluster is then updated based on the points assigned to the cluster. We keep repeating this procedure again and again and update steps until no point changes clusters, or equivalently, until the centroids remain the same.

Basic K-means algorithm

- 1: Select K points as initial centroids.
- 2: **repeat**
- 3: Form K clusters by assigning each point to its closest centroid.
- 4: Re-compute the centroid of each cluster.
- 5: **until** centroids do not change.

i. Finding total number of clusters, denoted by K

The value of K is found out by first forming the researcher groups. Total number of researcher groups formed is equal to the total number of researchers and researchers belonging to a particular group can carry out the co-operative work among themselves. Each group will have its respective threshold value which will decide the membership of a particular researcher in that group. T_i denotes the threshold for group i and is found out by averaging all the similarity scores corresponding to researcher i .

Membership criteria: 3.4

If $R_{ij} > T_i$, then researcher j belongs to group i
 else, researcher j belongs to some other group

Now, once all the researcher groups have been formed, then the value of K is equivalent to the minimum number of groups required to cover all the data points.

ii. Assigning Points to the Closest Centroid

To assign a point to the closest centroid, we need a proximity measure that quantifies the notion of 'closest' for the specific data under consideration. In k-means algorithm, there is a value associated with every data point (which is basically indicative of its distance & direction from the origin in some sense) whereas here we only have the similarity values between any two points and not their distance from some 'origin'. Rather, there is no notion of any global origin to begin with. We have resorted to this method to come up with the number of natural clusters because it seemed to be the best way out given that we only had the similarity values to work with. Hence in this case, we use the proximity measure as the distance between any two researchers, denoted by d_{ij} and is given as:

$$d_{ij} = 1 - R_{ij} \quad 3.5$$

where d_{ij} denotes the distance between researchers i and j R_{ij} denotes the similarity score between researchers i and j .

iii. Centroids and Objective Functions

Step 4 of the K-means algorithm is stated as “Re-compute the centroid of each cluster”, since the centroid can vary, depending on the proximity measure for the data and the goal of clustering. The goal of clustering is typically expressed by an objective function that depends on the proximities of the points to one another or to the cluster centroids; e.g., minimize the distance of each point to its closest centroid. The key point is this: once we have specified a proximity measure and an objective function, the centroid that we should choose can often be determined mathematically.

2.2 Accessing Expertise in Collaborative Interest Group

While getting recommendations from somebody with similar tastes is a good start, we might also want something else: that the person making the recommendation is an *expert* on these topics. Various techniques for getting information about expertise may be explored. An open problem thus arises to how can level of expertise be assessed objectively? We propose the solution for this by calculating every researcher’s level of expertise $[e]$ (that is the number of the researchers who have high interest similarity with a specific researcher).

$$e_i = \frac{1}{m} \sum_{j=1}^m ac_{ij} \quad 3.6$$

Suppose there are m researchers, the researcher i ’s level of expertise will be calculated by the following formula:

In this Formula,

$$ac_{ij} = \begin{cases} 1 & \text{if } R_{ij} \geq T_1 \\ 0 & \text{otherwise} \end{cases}$$

T_1 is a pre - determined Threshold Value

Finally, we select the researcher with the highest level of expertise, and take him/her as the center of the group to be constructed. Thus, the expert is identified. Expert identification from the Collaborative Interest Group is the key to recommendation as it is the expert’s blog provides recommendation to the user in the form of positive, negative or neutral, solving the first rater problem.

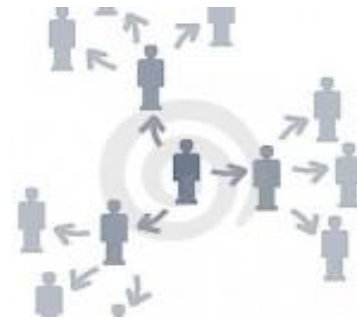


Fig 3: Expert to Solve First-Rater

3. ILLUSTRATION

To clearly illustrate the use and effectiveness of the proposed system, a case study is presented to describe a typical scenario, where

- There are 5 researchers viz. i, j, k, n & m . Therefore, $N_u = 5$
- There are 5 entries in each of the researcher’s blog site.

The following table 1 shows the blog entries of each of the Researcher i, j, k, n & m .

Table 1. Sample blog entries of 5 researchers

Researcher	i	j	k	n	m
Entry					
1.	$W_1, W_{16}, W_3, W_2, W_{17}, W_9, W_{24}, W_{25}$	$W_{14}, W_8, W_6, W_7, W_{17}, W_{21}, W_{25}$	$W_{11}, W_7, W_2, W_9, W_{19}, W_{21}, W_{25}$	$W_{13}, W_{13}, W_{10}, W_{14}, W_{21}, W_{22}$	$W_{10}, W_{15}, W_2, W_{21}, W_{23}, W_{24}$
2.	$W_4, W_2, W_3, W_{14}, W_{11}, W_{18}, W_{17}, W_{21}, W_{23}$	$W_1, W_{16}, W_{11}, W_7, W_{18}, W_{17}, W_6, W_{23}$	$W_{14}, W_{10}, W_4, W_9, W_1, W_9, W_{20}$	$W_{11}, W_{13}, W_6, W_5, W_{20}, W_{21}, W_{22}, W_{25}$	$W_{14}, W_{16}, W_9, W_8, W_{18}, W_{23}, W_{24}$
3.	$W_1, W_2, W_6, W_{13}, W_{20}$	$W_7, W_3, W_{18}, W_8, W_{17}, W_{24}$	$W_9, W_{19}, W_{11}, W_{10}, W_{17}, W_{23}$	$W_{13}, W_{14}, W_{18}, W_{12}, W_{20}, W_{22}$	$W_{15}, W_{19}, W_{11}, W_{16}, W_{20}, W_{23}, W_{24}$
4.	$W_1, W_2, W_4, W_9, W_{15}, W_{10}$	$W_6, W_6, W_7, W_{17}, W_{22}$	$W_{12}, W_9, W_{19}, W_{16}, W_{24}$	$W_{17}, W_{13}, W_2, W_{20}, W_{21}, W_{22}$	$W_{11}, W_{17}, W_6, W_{15}, W_{24}, W_{25}$
5.	$W_1, W_2, W_5, W_3, W_{19}$	$W_7, W_{18}, W_{15}, W_2, W_{18}, W_6, W_{17}, W_1$	$W_{19}, W_9, W_{17}, W_{10}, W_{10}$	$W_{18}, W_7, W_{13}, W_{13}, W_{20}, W_{23}, W_{24}$	$W_3, W_{13}, W_{22}, W_{23}, W_{24}, W_{25}$

3.1 Interest Mining Module

This module is concerned with the construction of Collaborative Interest Group by uncovering shared interest relationships between people, based on their blog document entries. The key point of constructing this Collaborative Interest Group is the calculations of interest similarity relations and application of the K-means clustering technique to cluster researchers with similar interests into the same group.

3.1.1 Interest Vector calculations

We have the interest vector corresponding to each of the researcher i, j, k, n & m represented as V_i, V_j, V_k, V_n, V_m . The calculation for these vectors using equation 3.2 is shown below:

For Researcher i: The Interest Vector is: $V_i = (S_{i1}, S_{i2}, S_{i3}, S_{i4}, S_{i5})$ where ;

$$S_{i1} = ef(w_1) \times \log [5 / uf(w_1)]$$

$$S_{i2} = ef(w_2) \times \log [5 / uf(w_2)]$$

$$S_{i3} = ef(w_3) \times \log [5 / uf(w_3)]$$

$$S_{i4} = ef(w_4) \times \log [5 / uf(w_4)]$$

$$S_{i5} = ef(w_5) \times \log [5 / uf(w_5)]$$

Now, from table 1, we find the values for ef 's and uf 's for the corresponding words:

$$ef(w_1) = 4 ; uf(w_1) = 3 \Rightarrow S_{i1} = 4 * \log (5/3) = 0.8874$$

$$ef(w_2) = 5 ; uf(w_2) = 4 \Rightarrow S_{i2} = 5 * \log (5/4) = 0.4846$$

$$ef(w_3) = 3 ; uf(w_3) = 2 \Rightarrow S_{i3} = 3 * \log (5/2) = 1.1938$$

$$ef(w_4) = 2 ; uf(w_4) = 1 \Rightarrow S_{i4} = 2 * \log (5/1) = 1.3979$$

$$ef(w_5) = 1 ; uf(w_5) = 1 \Rightarrow S_{i5} = 1 * \log (5/1) = 0.6989$$

Thus,

$$V_i = (0.8874, 0.4846, 1.1938, 1.3979, 0.6989)$$

Similarly,

For Researcher j:

$$V_j = (0.8874, 1.9897, 0.7959, 0.4845, 0.6655)$$

For Researcher k:

$$V_k = (1.9897, 0.6655, 0.1938, 0.6988, 1.9897)$$

For Researcher n:

$$V_n = (1.9897, 0.1938, 0.8874, 0.2907, 0.8874)$$

For Researcher m:

$$V_m = (1.1938, 0.4436, 0.3876, 0.4845, 0.1938)$$

3.1.2 Interest Similarity Score calculations

Using the formula defined in equation 3.3, we calculate the values of Similarity Score between each of the 2 researchers:

$$R_{ij} = 0.7063; R_{ik} = 0.7110; R_{in} = 0.7502; R_{im} = 0.8064; R_{jk} = 0.6688; R_{jn} = 0.6132$$

$$R_{jm} = 0.7424; R_{kn} = 0.8786; R_{km} = 0.8140; R_{nm} = 0.9169$$

As all the elements of both the vectors taken at a time to calculate the similarity score are positive, thus the range of similarity score is between 0 to 1.

This indicates that:

- The value of 1 means that the 2 researchers have exactly similar interests and;
- The value of 0 means that the 2 researchers do not have any similar interests at all.

Therefore, we can say that:

- The researchers n & m have almost similar interests (as $R_{nm} = 0.9169$, approx 1)
- The researchers k & n have similar interests to a very great extent (as $R_{kn} = 0.8786$)
- The researchers “ k & m ” and “ i & m ” have quite a lot similar interests (as $R_{km} = 0.8140$ and $R_{im} = 0.8064$)
- The researchers “ j & k ” and “ j & n ” have quite less similar interests (as $R_{jk} = 0.6688$ and $R_{jn} = 0.6132$)

3.1.3 Collaborative Interest Group Construction

We construct the collaborative interest group by using the technique of K-means clustering algorithm with the help of two basic steps. We first construct the researcher groups by finding the membership of each of the researcher using the formula defined in equation 3.4. This step would give us the total number of clusters required, denoted by K . And then we assign points to the closest centroid by taking the proximity measure as the distance between two researchers using the formula defined in equation 3.5.

A. Construction of Researcher Groups

(a) Membership for group i

Step 1: Calculate the threshold for this group, T_i

$$T_i = \frac{1}{5} [R_{ii} + R_{ij} + R_{ik} + R_{in} + R_{im}]$$

$$= \frac{1}{5} [1 + 0.7063 + 0.7110 + 0.7502 + 0.8064]$$

$$= 0.79478$$

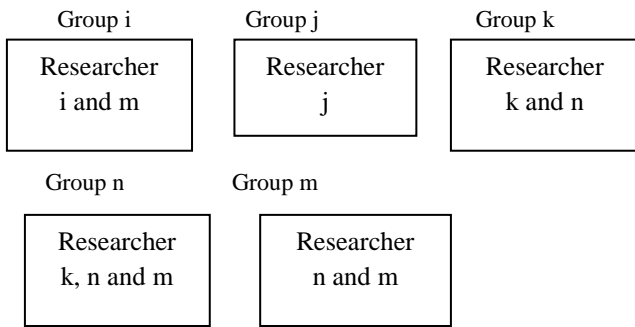
Step 2: Deciding the members for group i

As we can see,

$R_{ii} > T_i$ and $R_{im} > T_i$, therefore Researcher i and

Researcher m belong to group i.

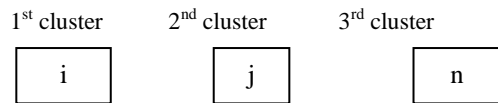
We find Membership for group j, group k, group n, group m in a similar way and the following Researcher Groups are formed with their respective members:



B. Construction of Clusters

(a) Total number of clusters

Now as we know total number of clusters i.e. K is equivalent to the minimum number of groups required to cover all the data points. Therefore, $K=3$. In other words, we can say that there are total three number of clusters required with the centroid as i, j, and n respectively.



(b) Assigning points to the closest Centroid

In this step we assign points (researcher m and k) to the closest centroid by taking the proximity measure as the distance between two researchers. Therefore using the formula defined in equation 3.5, we calculate the distance of these two researchers with each of the above researchers:

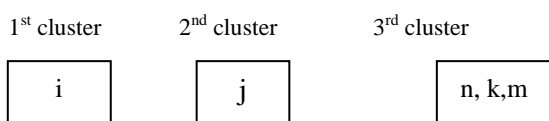
$$d_{ki} = 0.289; \quad d_{kj} = 0.3312; \quad d_{kn} = 0.1214$$

Since d_{kn} is minimum, therefore researcher k belongs to the 3rd cluster with centroid as n.

$$d_{mi} = 0.1936; \quad d_{mj} = 0.2576; \quad d_{mn} = 0.0831$$

Similarly, Since d_{mn} is minimum, therefore researcher m also belongs to the 3rd cluster with centroid as n.

So, after the first iteration we have the following clusters:



Now, the 2nd iteration begins. We will recompute the centroid of the 3rd cluster.

Distance between each of the two researchers is as follows:

$$d_{ij} = 0.2937; \quad d_{in} = 0.2498; \quad d_{jn} = 0.3868;$$

$$d_{km} = 0.186; \quad d_{ki} = 0.289; \quad d_{kj} = 0.3312;$$

$$d_{kn} = 0.1214; \quad d_{mi} = 0.1936; \quad d_{mj} = 0.2576;$$

$$d_{mn} = 0.0831$$

Assuming n to be the centroid:

$$S1 = d_{nm} + d_{nk} = 0.1214 + 0.0831 = 0.2045$$

Assuming m to be the centroid:

$$S2 = d_{mk} + d_{mn} = 0.186 + 0.0831 = 0.2691$$

Assuming k to be the centroid:

$$S3 = d_{km} + d_{kn} = 0.186 + 0.1214 = 0.3074$$

Since S1 is minimum, therefore n remains the centroid.

3.2 Accessing Expertise in Collaborative Interest Group

Expert from the 1st and 2nd cluster are i and j respectively since there are no more researchers in that cluster.

Expert from the 3rd cluster

Using the formula defined in equation 3.6, we find the Researcher n's level of expertise as:-

$$e_n = \frac{1}{3} [ac_{n1} + ac_{n2} + ac_{n3}]$$

$$= \frac{1}{3} [ac_{nn} + ac_{nm} + ac_{nk}]$$

In our experiment, Let $T_1 =$ average of R_{nn} , R_{nm} and R_{nk}

$$= \frac{1}{3} [R_{nn} + R_{nm} + R_{nk}] = \frac{1}{3} [1 + 0.9169 + 0.8786] = 0.9318$$

Now, since

$$R_{nn} = 1 > 0.9318, \text{ therefore } ac_{nn} = 1$$

$$R_{nm} = 0.9169 < 0.9318, \text{ therefore } ac_{nm} = 0$$

$$R_{nk} = 0.8786 < 0.9318, \text{ therefore } ac_{nk} = 0$$

$$\text{Therefore, } e_n = \frac{1}{3} [1 + 0 + 0] = 0.333$$

Similarly;

$$\text{➤ Researcher k's level of expertise} = \frac{1}{3} [ac_{kk} + ac_{kn} + ac_{km}] = \frac{1}{3} [1 + 0 + 0] = 0.3333$$

- Researcher m's level of expertise = $\frac{1}{3} [ac_{mk} + ac_{mn} + ac_{mm}] = \frac{1}{3} [0+0+1] = 0.3333$

Since the expertise value of all the three Researchers viz. n, k and m belonging to the 3rd cluster is same, therefore we conclude by saying that all the three researchers are expert in this particular field and that no one is better than the other. Thus, we can consider the Researcher n to be an expert of the 3rd cluster. Thus finally analyzing expert n's blog provides recommendation to the user in the form of positive, negative or neutral, solving the first rater problem.

4. CONCLUSION

Collaborative filtering-based systems require a large set of customer preferences for predicting the new preferences accurately. That is, for a reliable collaborative filtering-based prediction system, we need to have a large set of customer ratings on a variety of items available on the database. As the system can predict accurately only after it has gathered a large set of opinions, customers will not be willing to express their detailed preferences in the initial stages (if it is not going to help them in any way, despite offering preference ratings). We proposed a paradigm that allows users with similar interests to team up, setting a collaborative interest group (virtual community), subsequently, identifying an expert from each group and eventually based on the reviews in the expert's respective blog, providing relevant recommendation to the users in the form of positive, negative or neutral opinion. This paradigm effectively surmounts the first-rater problem, and provides users with the suitable recommendation.

5. REFERENCES

- [1] Carlson, C. N. (2003). Information overload, retrieval strategies and Internet user empowerment. Proceedings of The Good, the Bad and the Irrelevant, ed. L. Haddon *et al.*, University of Art and Design, Helsinki, Finland, pp.169–173.
- [2] Hill, W., L. Stead, M. Rosenstein, and G. Furnas. (1995). Recommending and evaluating choices in a virtual community of use. Proceedings of the ACM Conference on Human Factors in Computing System, pp.194-201.
- [3] Resnick, P. and Varian, H.R.. (1997). Recommender Systems. Guest Editor's Introduction to the Special Section, Communications of the ACM, 40(3), pp.56-58.
- [4] Mortensen, M. (2007). Design and evaluation of a recommender system, Master Thesis.
- [5] Linden, G., B. Smith, and J. York. (2003). Amazon.com Recommendations: Item-to-Item Collaborative Filtering. IEEE Internet Computing, Vol.7, No.1, pp. 76-80.
- [6] Miller, B. N., I. Albert, S. K. Lam, J. A. Konstan, and J. Riedl. (2003). MovieLens Unplugged: Experiences with an Occasionally Connected Recommender System. Proceedings of the International Conference on Intelligent User Interfaces, pp. 263-268.
- [7] Melville, P., Mooney, R., and Nagarajan, R. (2001). Content-boosted collaborative filtering, Proceedings of the SIGIR-2001 Workshop on Recommender Systems, New Orleans, LA, September 2001
- [8] Badrul, M.S., Karypis, G., Konstan, J. A., and Reidl, J. (2001). Item-based collaborative filtering recommendation algorithms. Proceedings of the 10th international conference on World Wide Web, pp.285-295.
- [9] Resnick, P., N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl. (1994). GroupLens: An open architecture for collaborative filtering of netnews. Proceedings of the Computer Supported Cooperative Work Conference.
- [10] Shardanand, U. and P. Maes. (1995). Social information filtering: Algorithms for automating 'word of mouth'. In Proc. of the Conf. on Human Factors in Computing Systems.
- [11] Balabanovic, M. and Shoham, Y. Fab: Content-Based, Collaborative Recommendation, in Resnick and Varian (eds.), pp. 66-72.
- [12] Lee, W. S. (2001). Collaborative learning for recommender systems. Proceedings of the International Conference on Machine Learning,
- [13] E. Vozalis and K.G.Margaritis. (2003). Analysis of Recommender Systems' Algorithms. presented at HERCMA. pp.732-745.
- [14] Colazzo, L., A. Molinari and N. Villa. (2009). Collaboration vs. Participation: the Role of Virtual Communities in a Web 2.0 world. International Conference on Education Technology and Computer. pp.321-325.
- [15] Barrett, H. (2008). Web 2.0 Tools for Lifelong & Life Wide Learning. Online Course.
- [16] Kumar, A. & Jain, A. (2010). An Algorithmic Framework for Collaborative Interest Group Construction. Recent Trends in Networks and Communications, LNCS-CCIS, Springer, Volume 90, Part 3, 500-508
- [17] Pang-Ning Tan, Michael Steinbach and Vipin Kumar. *Introduction to Data Mining*. Pearson Education.