International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 9
September 2017

Case Based Approach to Collaborative Filtering Recommender System[1]

Apurva Vashist[2], Bhu Dev Sharma[3]

[1]The paper is drawn heavily upon the dissertation on the same topic submitted by the first author to Institute of Electronics and Telecommunication Engineering (IETE) as part of Advanced Level Course in Computer Science.

[2].Academic Dean,Priyadarshini College of Computer Science , Greater Noida.

[3].Former Professor of Mathmatics, Clark Atlanta University, Atlanta, GA, USA

Abstract

Recommender systems provide personal advice to the user for the item that user might be interested in. In the paper, two types of techniques are used for performing recommendation, firstly, Collaborative Filtering (CF) and secondly, Case Based Reasoning (CBR).

Collaborative Filtering recommends the item according to the similar taste of group of people. According to that group of people user select the item. In the case based reasoning, new problems are solved by adopting the solutions given for old ones. When these recommendation techniques are combined together the system is called the Hybrid system. This system improves the performance. In this paper, three types of hybrid systems are used i.eWeighted, Switching and Cascade. Here, the attempts have been made to analyse the performance of these three types of hybrid systems. In order to analyse the performance Movie Lens Data set has been used. Experimental results are presented to demonstrate the performance of the proposed hybrid system.

Indexing Terms: web mining, artificial intelligence, case based reasoning, hybrid system, recommender system.

1. Introduction

Recommender systems are applications that provide personalized advice to users about products or services they might be interested in. Recommender systems are playing a major role in the digital and social networking revolution and becoming a part of everyday life. Recommender system provides advice to user about items they wish to purchase or examine. Recommendations made by such system can help users navigate through large information spaces of product descriptions, new articles or other items. The main task of a recommender system is to locate items, information sources and people related to the interest and preferences of a single person or a group of people. They have become fundamental applications in electronic commerce and information access, providing suggestions that effectively prune large information spaces so that users are directed toward those items that best meet their needs and preferences. A variety of techniques have been proposed for performing recommendation, including content-based, collaborative, knowledge-based and other techniques (Burke, 2000a).

2. Different Types of RecommenderTechniques

2.1 Collaborativefiltering

In collaborative filtering, the user will be recommended items that people with similar tastes and preferences liked in the past. Collaborative filtering recommendation is probably the most familiar, most widely implemented and most mature of the technologies. Collaborative recommender system aggregate ratings or recommendations of objects, recognize commonalities between users on the basis of their rating, and generate new recommendations based on inter-user comparisons. A typical user profile in a collaborative system consists of a vector if items and their ratings, continuously augmented as the user interact with the system over time. Some systems used time-based discounting of ratings to account for drift in user interests (Billsus&Pazzani,2000).

2.2 Content BasedFiltering

This recommendation depends only user not the group of people. If the user preferred in past then the next time he will take the same item. A content-based recommender depends on learning method i.e user select the item according to the past experience.

2.3 Case BasedReasoning

The main idea of CBR is to solve new problems by adapting the solutions given for old ones. When we apply CBR to recommender systems, the importance of the recommendation process lays with the case base representation. We propose, as a representation, a list of experiences (cases) of the user in certain items. Experiences are represented by means of objective attributes describing the item (case definition) and subjective attributes describing implicit or explicit interests of the user in this item (case solution).Assumingthattheuser'sinterestinanewitemissimilartotheuser'sinterestin Similar past experience,when a new item comes up,the recommender system predicts the user's interest in the new item based on interest attributes of similar experience.

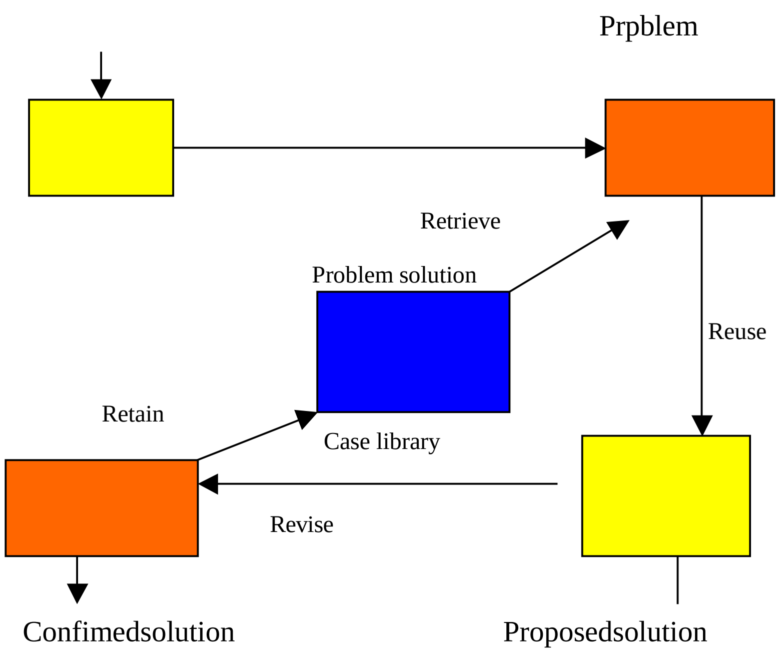

Fig.2 The CBR problem solving cycle

3. Hybrid Recommender Systems

Hybrid recommender system combines two or more recommendation techniques to gain better performance. If it combines both techniques, it is called hybrid system. Such hybrid could offer good performance even with little or no user data. Several recommendation systems use hybrid approach by combining different recommendation techniques, which helps to avoid certain limitations of recommendation techniques (Adomavicius and Alexander, 2005).

Hybridization Methods

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 9
September 2017

Hybridization method Description

Weighted The scores (or votes) of severalrecommendationtechniquesare combined together to produce a singlerecommendation.

Mixed Recommendations from several differentrecommendersarepresented at the sametime.Feature combination Features from different recommendation data sourcesarethrown together into a single recommendation algorithm.Cascade One recommender refines the recommendationsgivenbyanother.

Feature augmentation Output from one technique is used as an input featuretoanother.

Switching The system switches between recommendation techniques depending on the current situation

4.Implementation&Result

Presents proposal design and implementation of hybrid recommender system that integrates case based and collaborative filtering. The work includes empirical comparisons by using weighted, switching and cascade hybrid techniques. This proposed work is implemented by using movie lens dataset.

1 The database contains ratings from 72,916 users on 1,628 movies. The analysis covers 35,527 users who gave at least 20 votes over the total of 1623 movies. We randomly select 100 users out of 35,527 users and divide them into training set (80 users) and test set (20users).

2 From the above selected training data, 15 users and 15 movies items were selected as test data to generate the User × Item matrix with theratings.

3 The Correlation based Similarity algorithm was used to find the similarity between item-item.

4 To calculate the predicted ratings weighted sum technique was used. Then, finally the Mean Absolute Error between the actual ratings predicted ratings wascalculated.

5 Combined CF& CBR Approach by using the age factor. Then, same above Collaborative Filtering algorithm was run and then the results werecompared.

Weighted Kappa Algorithm, to calculate the affinity (correlation coefficient) from the observed and expected frequencies as:

$$WK = \frac{Po(w)-Pe(w)}{1-Pe(w)}$$

where*WK* represents the Weighted Kappa value (affinity);Po(w) represents the observed frequencies and Pe(w) represents the expected frequencies .

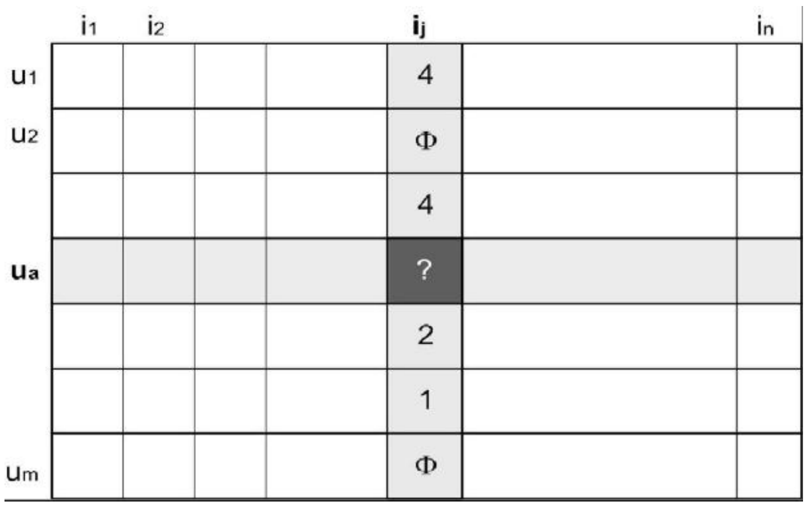

fig.4.1 User × Item matrix for the Collaborative Filtering

4.1 Correlation Coefficient

In collaborative filtering process as computing the expected value of a user prediction, given his/her ratings on other items. The item-based approach looks into the set of items the target user has rated and computes how similar they are to the target item i and then selects k most similar items $\{i_1, i_2, \ldots, i_k\}$. At the same time their corresponding similarities

$\{s_{i1}, s_{i2}, \ldots, s_{ik}\}$are also computed. Once the most similar items are found, the prediction is then computed by taking a weighted average of the target user's ratings on these similar items.(Sarwar,Kargis 2001)

4.2 Adjusted CosineSimilarity

One fundamental difference between the similarity computation in user based CF and item- based CF is that, in user –based CF the similarity is computed along the rows of the matrix but in case of the item –based CF the similarity is computed along the columns(Sarwar,Karkis 2001). The similarity between items i and j is given by

$$sim(i,j) = \frac{\sum_{u \in U}(R_{u,i} - \bar{R_u})(R_{u,j} - \bar{R_u})}{\sqrt{\sum_{u \in U}(R_{u,i} - \bar{R_u})^2}\sqrt{\sum_{u \in U}(R_{u,j} - \bar{R_u})^2}}.$$

4.3 PredictionComputation

In a collaborative filtering system is to generate the output interface in terms of prediction. We isolate the set of most similar items based on the similarity measures and the target ratings and use a technique to obtain predictions.

4.4 Weighted Sum

This method computes the prediction on an item i for a user u by computing the sum of the ratings given by the user on the items similar to i. Each ratings is weighted by thr corresponding similarity $s_{i,j}$between items i and j.The Prediction Pu,i as(Sarwar ,Kargis 2001)

$$P_{u,i} = \frac{\sum_{\text{all similar items, N}}(s_{i,N} * R_{u,N})}{\sum_{\text{all similar items, N}}(|s_{i,N}|)}$$

4.5 Mean Absolute Error

Mean Absolute Error (MAE) is the most common predictive accuracy metric used in recommender systems. MAE is calculated by determining the mean of the errors between a prediction and the associated actual rating .MAE is a measure of the deviation of recommendation from their true user –specified values. For each ratings-prediction pair<pi,qi>this matrix treats the absolute error between them. There are two variations on MAE: calculating it per user, averaging over all users, or per item, averaging over all items. These two ways report different views into the error of an algorithm(Sarwar,Kargis 2001). We choose to calculate MAE per user, as it is more closely reports a user's view of the algorithm. MAE per user is calculated as follows:

$$MAE = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N}$$

4.6 Algorithm For CBR

Let S = A session: { s0, .., sn}, n>=20 where each si consists of a pair <restaurant, rating>

r = a positive rating from S

T = test data for the session, initially { }
P(S, t), a prediction function that predicts the
score of test item t, given ratings from training data in S.
move r from S to T move 6 random s from S to T for i<- 4, 6, 8
p <- t such that P({s0,..si}, t) is maximized if p equals r
correct prediction (i) (Burke 2000b)

The comparison of MAE Hybrid Systems are:

Hybrid System	MAE
Weighted	0.35600
Switching	0.022427
Cascade	0.014951

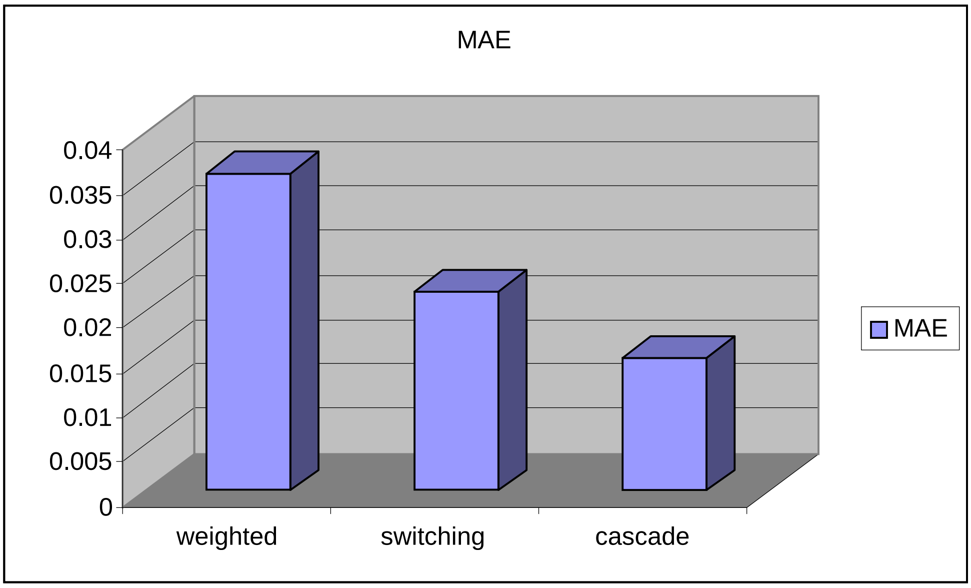

Fig.4.2 Comparing Hybrid Methods

The above chart clearly shows that MAE is decreasing when different methods of hybrid are compared i.e weighted, switching and cascade. It shows that cascade hybrid system outperform weighted and switching systems. From the experimental evaluation of the item- item collaborative filtering scheme we make some important observations. These are given below.

First, the item-item scheme provides better quality of predictions than the use-user (*k*- nearest neighbor) scheme. The improvement inquality is consistent over different neighborhood size and train/test ratio. However, the improvement is not significantly large.

The second observation is that the item neighborhood is fairly static, which can be potentially pre-computed, which results in very high online performance.

5.Conclusions and Future Work

In hybrid recommender systems, two recommender techniques, *i.e*collaborative filtering and case based reasoning are combined for improved performance. Three hybridization techniques are discussed in this project, *i.e*. weighted, switching, and cascade. Experiments indicate that collaborative filtering does improve the performance over the CBR component acting alone.The experiments conclude that the cascade hybrid system gives the best performance compared to weighted and switching.

Future work

In the present work three hybrid techniques have been used. However, in future other hybrid techniques, *e.g.* meta level, feature combination and feature augmentation techniques can also be used. Further extensions may include hybridization using content-based filtering and demographicfiltering.

References

1. Adomavicius, Gediminas and Alexander Tuzhilin2005,'Toward the Next Generation of Recommended System': A Survey of the State-of –the –Art and Possible Extensions. 'IEEE Transactions on Knowledge and Data Engineering', vol. 17, no 6,June
2. BadrulSarwar, George Karypis, Joseph Konstan, and John Riedl,2001,'Item- BasedCollaborative Filtering Recommendation Algorithms'WWW10, pp. 285 - 295.
3. Billsus, D and. Pazzani., M. J 1999, 'A Hybrid User Model for News Classification'. In Kay J. (ed.), UM99 User Modeling – Proceedings of the Seventh International Conference, pages 99–108. Wien, New York: Springer- Verlag
4. Burke, Robin 1999, 'The Wasabi Personal Shopper: A Case Based Recommended System'. In Proceedings of the 11thNational Conference on Innovative Applications of Artificial Intelligence, pp. 844-84, AAAI .
5. Burke, Robin 2000a, 'Knowledge Based Recommender Systems'. A. Kent (ed) in Encyclopedia of Library and Information Systems, vol.69, supplement 32, New York: MarcelDekker.
6. Burke, Robin 2000b,' A case Based Reasoning Approach to Collaborative filtering'. In E.blanzieri and L.Portinale (eds.),Advances in Case Based Reasoning (5thEuropean workshop EWCBR 2000), pages 370-379. New York: Springer Verlag.
7. Burke,Robin 2002,'Hybrid Recommender Systems: Survey and Experiments'. User Modeling And User-AdaptedInteraction'12(4),pp331-370.
8. Eui-Hong (sam) Han and George Karypis.2005,'Feature-Based Recommender System' Proceedings of the 14th CIKM,pp.446-452
9. Mukund Deshpande and George Karypis 2004,'Item-Based Top-N Recommendation Algorithms' ACM Transactions on Information Systems. Volume 22, Issue 1, pp. 143 - 177,.