Recommending Generalized Products in Collaborative Filtering

Aye Kaday Htun, Wint Thida Zaw
University of Computer Studies, Yangon
aye.kady@gmail.com

Abstract

Recommender systems are particularly useful for computer users, here decisions must be normally taken in a short time and the effort required for interacting with the system must be limited as much as possible. The Recommendation systems can help the user to take a decision suggesting those products which best suit his needs and preferences. Recommendation systems have been an important application area and the focus of considerable recent academic and commercial interest. In the classical collaborative filtering recommendation approach, the voting prediction method is based on the computation of the similarity of the active user, to whom a recommendation has to be made, with the other users .Collaborative filtering (CF) describes a variety of processes that automate the interactions of human advisors; a collaborative filter recommends items based upon the opinions of human advisors. In this paper, we implement the recommendation system are the use of Voting by Category method in memory-based collaborative filtering method.

1. Introduction

Nowadays, Recommender System is new types of internet-based software tools, designed to help user find their way through today's complex on-line shops and entertainment website. This leads to a clear denied for automated methods that locate and retrieve information with respect to user's individual interests. More and more people accessing the Internet also provide new possibilities to organize and recommend information.

Recommender systems can be present in all sorts of systems and situations, and thus can be implemented in many different ways. Recommender system provides one way of recommend or suggests items or products to the customer based on his/her preference. These systems are often used by e-commence websites as marketing tools to increase revenue by presenting products that the customer is likely to buy. An Internet site using a recommender system can exploit knowledge of customer's like and dislikes to build an understanding of their individual needs and there by increase customer loyally.

Collaborative filtering techniques have been successful in enabling the prediction of user preferences in the recommendation systems. There are three major processes in the recommendation systems: object data collections and representations, similarity decisions, and recommendation computations. Collaborative filtering aims at finding the relationships among the new individual and the existing data in order to further determine the similarity and provide recommendations. Collaborative filtering is a three stage process of finding similar users, computing predicted voting, and applying the predictions as recommendations to the user.

In this paper, we present an approach to give the prediction result of user's requirements using voting by category algorithm with memory-based collaborative filtering method.

2. Background Theory

This paper is intended to generalize products in the Recommendation system of Collaborative Filtering. So, this paper is based on the voting by category algorithm with the use of Memory-based approach.

2.1. Recommendation System

Recommendation system can provide targeted product suggestions for users either overwhelmed by the large number of alter native options offers nowadays in e-commerce websites, or not having enough knowledge to autonomously select the most suited product. Recommendation systems are designed to allow users to locate the preferable items quickly and to avoid the possible information overloads. This system have gained large acceptance among users for the task of assisting them in searching ,sorting classifying, filtering and sharing the vast amount of information now available on the Web [2].

Recommendation systems are particularly useful for digital device users, here secessions must be

normally taken in a short time and the effort required for interacting with the system must be limits as much as possible. The purpose of the recommender system is to help user to deal with information overload in the internet in which numbers of users and contents are growth every time and provide products recommendations, contents and services to user.

Recommendation system can generally be categorized in five major types. They are

(1) Content-based filtering

(2) Demographic filtering

(3) Collaborative filtering

(4) Knowledge-based Method

(5) Utility-based Method

Current Recommendation system widely used on the content-based filtering and collaborative filtering methods. In this paper, the focus will be on the collaborative filtering.

2.2. Collaborative Filtering Method

Collaborative Filtering is one of the most popular and successful filtering techniques that has been used to date. It is applied in a setting where users have a choice between a number of items (e.g., all laptops in a dg world shop) and provide votes to items that they know about. Collaborative Filtering helps users to make choices based on the opinions of similar users in a system and find relevant items that they may not have explored so far. The basic idea employed is that users who agree with each other on some items based on their voting are likely to agree or disagree on future items. To make predictions for a given user, collaborative filtering algorithms typically find similar users in a system, and assign weights to the level of similarity. The preferences of this set of similar users is combined and weighted with the assigned weights. This technique has its basis in every-day life where people consider the opinions of similar minded people in order to decide what they want to buy next [9].

Collaborative filtering can help e-commerce in converting web surfers into buyers by personalization of the web interface. It can also improved cross-sell by suggesting other products the consumer might be interested in. In a world where an e-commerce site's competitors are only a click or two away, gaining customer loyalties is an essential business strategy [3]. Collaborative filtering can improve the loyalty by creating a value-added relationship between supplier and consumer. Collaborative filtering systems try to predict the utility of items for a particular user based on the items previously voted by other users.

Collaborative filtering systems are often distinguished by whether they operate over implicit versus explicit votes. Explicit voting refers to a user consciously expressing his or her preference for a title, usually on a discrete numerical scale. Collaborative filtering algorithms are now widely used in Internet applications, with considerable success; use collaborative filtering to provide personalized information filtering for users.

Collaborative recommendations can be grouped into two general classes: memory-based (or heuristic-based) and model-based.

Memory-based algorithms essentially are heuristics that make voting predictions based on the entire collection of previously voted items by the users.

Model-based algorithms use the collection of voting to learn a model, which is then used to make voting predictions [2].

In our system, we use memory-based collaborative recommendation.

2.3. Memory-Based approach

Memory-based algorithms are the more prevalent of the two categories and use all available data in order to make a prediction for the selected user. The system database contains sets of user preferences, recording the transactions that are made by all users of the system. Memory-based CF algorithms retain all relevant data in memory and compute the required prediction on demand in real. The advantage of this approach is that new data provided by a user can immediately be taken into account. Typically, this provides a better usability experience, as the user can see how his/her actions are immediately utilized by the system. However, the scalability of such systems is not arbitrary; using memory based algorithms for real-world systems requires optimizations and some approximations have to be made, which can counter the accuracy of the original method. Nonetheless, several algorithms have been proposed for memory-based CF due to their high accuracy and simplicity of implementation.

Memory-based methods for collaborative filtering predict new voting by averaging (weighted) voting between, respectively, pairs of similar users or items. In practice, a large number of voting from similar users or similar items are not available, due to the sparsity inherent to voting data. Memory-based algorithm is the most popular prediction technique in collaborative filtering applications [8]. The basic idea is to compute the active user’s vote on a target item as a weighted average of the votes given to that

item by other like-minded users. In equation (1), the user database therefore consists of a set of votes $v_{i,j}$ corresponding to the vote for user i on item j. If I_i is the set of items on which user i has voted, then we can done the mean vote for user i as:

$$\bar{v}_i = \frac{1}{|I_i|} \sum_{j \in I_i} v_{i,j} \quad \textbf{(1)}$$

In memory-based collaborative filtering algorithms, we predict the votes of the active user based on some partial information regarding the active user and a set of weights calculated from the user database. The basic idea is to compute the active consumer's predicted vote of a product as a weighted average of the votes given to that product by other consumers. In equation (2), the prediction vote for an active user for item j , $P_{a,j}$ is a weighted sum of the votes of other users given by:

$$p_{a,j} = \bar{v}_a + k \sum_{i=1}^{n} w(a,i)(v_{i,j} - \bar{v}_i) \quad (2)$$

where *n* is the number of the consumers who voted product *j*. v_i and v_a are the average voting given by users i and a. $v_{i,j}$ is the vote cast by *i* on *j*. $w(a,i)$is the weight given to every user i from active user. *k* is a normalizing factor such that the absolute values of the weights sum to unity.

In the following, we distinguish between the various collaborative filtering algorithms in terms of the details of the weight calculation. The weight is calculated by comparing a set of common products, which the active user and all other users in the database have voted. Three major methods to define the weights are Mean Squared Differences, Pearson Correlation and Vector Similarity. The most popular memory-based algorithm uses a similarity measure called Pearson's Correlation. This is a standard measure in statistics, which is applied here with only a small modification: similarity is measured based only on items where votes are available for both users. Pearson's Correlation based Collaborative Filtering: Predicted votes v(i, j) are computed as defined in Equation (3), with similarity weights w(a, i) defined as follows:

$$w(a,i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 \sum_j (v_{i,j} - \bar{v}_i)^2}} \quad (3)$$

The memory-based CF algorithms and address the problem of deciding which instances to use during prediction, in order to reduce time complexity, and to improve the accuracy by avoiding noise and over fitting.

2.3.1. Voting by Category. In some collaborative filtering applications, the dimensions of the users' voting matrix could become unmanageable, preventing the practical calculations over the over matrix. There could be very few common votes to the same products if not using default voting method mentioned before, however, providing default votes may not improve the performance. Basically, they assume the existence of small number of generated clusters or pre- existing categories to which products can be assigned. Then transfer the voting matrix into much lower dimension by transfer users' voting to products into the voting to categories [10].

In equation (4) are calculated the new votes of user to categories ,

$$v_{i,c} = \overline{v_{i,j}}, \qquad j \in c \quad (4)$$

$v_{i,j}$ corresponding to the vote for user i on item j, v_{ic} corresponding to the vote for user i on category c. Now the entry of the new matrix is the average over the votes of the products per each category for a given user.

3. Overview of System Design

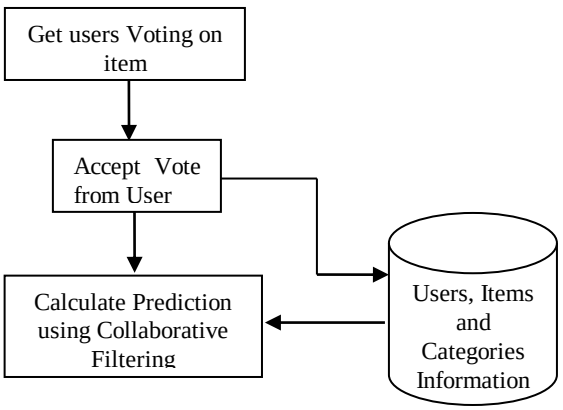
In our system, there are three main steps in Figure(1), to give recommendation for user, (i) get user voting on item (ii) calculate prediction using collaborative filtering (iii) show information predicted value with item and category.

At first, User can present vote to user interest's one product's item in one time. Accept vote from user with value and vote value have a limitation as 1 to 5 votes.

For example, votes value are

(1) Don't like it
(2) Not for me
(3) Normal like
(4) I Like it
(5) I Love it

This voting is used for Collaborative Filtering process. Calculate for each category (Brand) votes by using voting by category in memory-based approach.



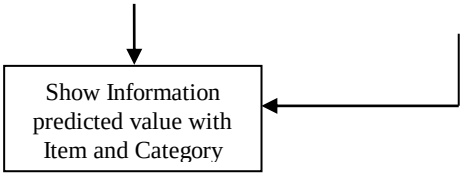


Figure 1. Overview of system design

Filtering system is used to filter the information according to the user's votes. The next process, calculated predictions have four steps.

Step 1, The system accepts the votes on each item from the user. After got the vote value, the system store the data set with item and category for calculate prediction.

Table 1. Voting category by laptop computer

Brand / Name	HP1		Toshiba		Acer	
	DV712	TX210	L305	M200	A150	AS57
AyeAye	5	1				
Ma Ma		3		1		5
Su Su			3		5	4
NgeNge	4		2			

Table (1), Illustrates the user's vote for Laptop Computer respect the category and item. In this table, User Nge Nge is active user and she give vote value 2 in L305 Toshiba laptop , the system start calculate prediction and give recommendation list for Nge Nge .

Brand of Laptop computer find average for one brand, calculate with equation (1) example for HP1 Brand,

$\overline{v1}$(Aye Aye)= $\frac{1}{2}$ (5+1) = 3
$\overline{v2}$(Ma Ma) = 3
$\overline{v3}$(Su Su) = 0
$\overline{v4}$(Nge Nge)=4

Step 2, Average vote of all items in each brand (eg, Table (2), shown in HP1 Brand with column) and average vote of all brand in one product, (eg, Table (2), shown in Average with column).

Table 2. Showing average vote value in brands

Brand / Name	HP1	Toshiba	Acer	Average
Aye Aye	3			3
Ma Ma	3	1	5	3
Su Su		3	4.5	3.75
Nge Nge	4	2	?	3

Step 3, The system compute the weight between the active user and the existing user using Pearson Correlation has been widely and successfully used as a similarity measure between users. In step 3, based on calculate average vote for each category on the user voting using equation (3),
w(Nge Nge, Aye Aye) =0

w(Nge Nge, Ma Ma) =0.71

w(Nge Nge, Su Su) =0.57

Step 4, The weight similarity between active user and existing user are compute equation (3) with the use of equation (2) to find the prediction result of all brands for active user.
p(Nge Nge, HP1) = 3

p(Nge Nge,Toshiba) =3+1[(0.71*(1-3))+0.57*(3-3.75)] = 1.15

p(Nge Nge,Acer)= 3+1[(0.71*(1-3))+0.57*(4.5-3.75)] = 4.85

Then the system computes the prediction vote for the new (active) user and displays the votes list result to the user which is the closet votes given by user. If this user wants to see recommendation result list, the system can present with Brand and item on digital device for nearest of he/she is interested. Collaborative Filtering is to predict the preference of a particular user (active user) based on a database of users' preferences.

At last, Recommendation result show to user's products lists in user gave vote on prediction list to show the user chosen products with Brand or item list.

4. Implementation of the Proposed System

The recommendation system gives predictions which are relevant interested item. The system is applied for digital device information records. Memory-based methods have the advantages of being able to rapidly incorporate the most up-to-date information and relative accurate prediction. Digital devices recommendation system is necessary for buyers who have extensive choice of which brands in product is need for user. The system can support the buyer's requirement with variety of digital devices so that he/she can make choice that meets his/her

design and other supported information. . In the propose system, recommendation system for digital devices are three devices (which includes, Laptop Computer, Digital Camera and Mobile phone) is implemented. In our system, Each products have five Brands respectively as follow.

Products	**Categories**
(1) Laptop Computer (i)	HP1
(ii)	Toshiba
(iii)	Acer
(iv)	Fujitsu
(v)	DELL
(2) Digital Camera (i)	Nikon
(ii)	Canon
(iii)	Sony
(iv)	Olymn
(v)	Panasonic
(3) Mobile Phone (i)	Nokia
(ii)	Sony Errision
(iii)	Samsaung
(iv)	Motorola
(v)	LG

Figure 2. User choose item with laptop brand

Figure (2), show the active user popular list in all brands of laptop computer. The user can give vote to one item in this list and the system save user's vote and calculate prediction.

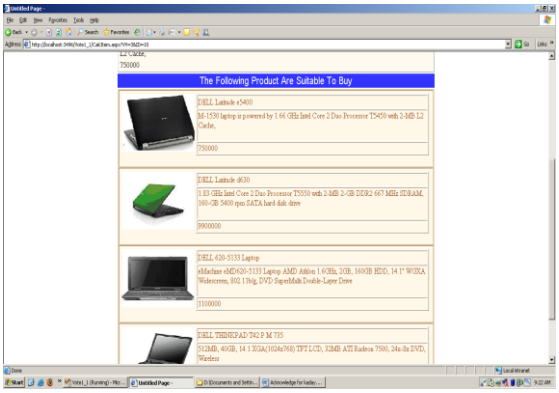


Figure 3. Prediction result list for all items in laptop computer

Figure (3), shows the result of the proposed system in user-based Collaborative Filtering. This system can give prediction results votes for active user; calculate with memory-based voting by category method generalized sample with a high predicted vote recommendation list of laptop computer. In this paper, the system can represent and also calculate that, mobile phone and digital camera give user nearest needs to each product.

All data required for the system is got from dg world shop center. The system is tested by calculate prediction result for 100 people. In this paper, we show the testing result of four people and calculation prediction of laptop computer categories for 4 user. When calculate with brands, the system filtering of all items in each brand.

5. Conclusion

To provide better search results, Recommender systems attempt to reduce information overload and retain customers by selecting a subset of items from a universal set based on user preferences. Collaborative filtering techniques have been proposed to decrease the processing time and the data latency. The results from different recommendation systems indicate that collaborative filtering techniques afford the systems enough ability to provide recommendations to users. The recommendation systems can predict user behavior patterns without any knowledge of the user in advance, and to evaluate the accuracy by the comparing the prediction and the reality. This system uses the voting by category and memory-based method to reduce time and get extensive information. Not only in digital devices but can also be used in other kinds of commodity, such as Movie, Book stores and etc. The goal of this system is to present users with a highly relevant set of items.

References

[1] A.C.Yun-An and D.McLeod, "Collaborative Filtering for Information Recommendation Systems", Department of Computer Science and Integrated Media system Center, University of Southern California, Los Angeles, California, USA.

[2] A.Hinneburg, C.Aggarwal, D.A.Keim, "What is the nearest neighbour in high dimensional spaces?", Processing of the 26th VLDB Conference, Cairo, Egypt, 2000.

[3] B.Sarwar, G.Karpis, J.A.Konstan, J. Riedl, "User-Based Collaborative Filtering Recommendation Algorithm" May 1-5 , 2001.

[4] J. Ben Schafer, Joseph ,A. Konstan, John Riedlm, "E-Commerce Recommendation Applications", GroupLens Research Project, Department of Computer Science and Engineering University of Minnesota, Minneapoils.

[5] J. S. Breese,D. Heckerman and C. Kadie, "Empirical Analysis of Predictive Algorithm for Collaborative Filtering", Proceeding of the 14th Conference on Uncertainty in Arteficial Intelligence,1998 .

[6] K.Yu, Y.Yu, J.Tao, M.Ester and H.P.Kriegel, "Instance Selection Techniques for Memory-based Collaborative Filtering", University of Munich, Tsinghua University,Beijing,China.

[7] Q.Yuan, Y.Xiaoyun, H.Tianping, "Survey of Recommendation Systems and Algorithms", Term Paper for EE 380L: Data Minning Spring, May 5, 2000.

[8] P. Resnick, N. Iacovou, M. Sushak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews" *In Proceedings of* the 1994 Computer Supported Collaborative Work Conference.

[9] R. Jin, J. Y. Chai, and L. Si, "An automatic weighting scheme for collaborative filtering" *In Proc. of* SIGIR,2004.

[10] U. Shardanand, and P. Maes, "Social Information filtering Algorithms for Automating 'Word of Mouth'" *In Proceedings of* CHI'95.