A Personalized Collaborative Filtering Recommendation Using Association Rules Mining and Self-Organizing Map

Hongwu Ye
Zhejiang Textile & Fashion College, Ningbo 315211, P. R. China
Email: hongwu_ye@163.com

***Abstract*—With the development of the Internet, the problem of information overload is becoming increasing serious. People all have experienced the feeling of being overwhelmed by the number of new books, articles, and proceedings coming out each year. Many researchers pay more attention on building a proper tool which can help users obtain personalized resources. Personalized recommendation systems are one such software tool used to help users obtain recommendations for unseen items based on their preferences. The commonly used personalized recommendation system methods are content-based filtering, collaborative filtering, and association rules mining. Unfortunately, each method has its drawbacks. This paper presented a personalized collaborative filtering recommendation method combining the association rules mining and self-organizing map. It used the association rules mining to fill the vacant where necessary. Then, it employs clustering function of self-organizing map to form nearest neighbors of the target item and it produces prediction of the target user to the target item using item-based collaborative filtering. The recommendation method combining association rules mining and collaborative filtering can alleviate the data sparsity problem in the recommender systems.**

***Index Terms*—personalized service, recommender systems, association rules mining, collaborative filtering, mean absolute error**

I. Introduction

Currently, the development of the Internet brought us a lot of information, we almost can not handle. To solve the information overload, all kinds of recommendation systems have been established to assist and complement the natural social process. Recommended system has been developed to automate the process recommended. These systems are recommended web resources, online news, and all kinds of movies. Large-scale commercial application of the recommendation system can be found in many e-commerce sites such as Amazon, CDNow, Dangdang. These commercial systems proposed in the past transactions and feedback from potential consumers based products. They are becoming a standard e-commerce technology, which convert browsers to buyers as part of e-commerce sales, increase cross-selling, and build customer loyalty [1].

The most personalized recommendation system uses three type of technology, content-based filtering, collaborative filtering and association rules, recommend products to customers [2,3]. The first content-based filtering, it is similar to the proposed project attempts to a given user in the past very much. It is based on the content between them and the user profile comparison. The second approach to collaborative filtering to identify the user's taste is similar to that to users, and recommended their favorite projects. Given a set of projects, users can express their attempted projects, rating. Recommended by the user can compare the ratings of other users who find the 'most similar to similar to some of the standard, then the user's proposal, similar to the user in the past like the project. See the project results, based on projections from the neighbors with known scores. The third method of association rules, people will by association rule mining proposal.

Traditional collaborative filtering challenge is as follows [4,5,6]:

Sparsity: Even if the user is very active, there are many items in the user database to provide the total number of projects score low ratings. Collaborative filtering algorithm as the main cooperation in the project, rated contains similar measures calculated on the basis of sparse large level may lead to poor accuracy.

Scalability: collaborative filtering algorithm seems to be in the project, interesting user filter efficiency. However, they need the calculation is very expensive and growth of the user and the database is not proportional to the number of items.

Cold start: The project can not recommended, unless it has been a number of user ratings. This problem applies to new projects, and particularly harmful to users with eclectic tastes. Similarly, a new user to rate the project before a sufficient number of collaborative filtering algorithm can provide accurate advice.

In this paper, we presented a personalized collaborative filtering recommendation method combining the association rules mining and self-organizing map. It used the association rules mining to fill the vacant where necessary. Then, it employs clustering function of self-organizing map to form nearest

doi:10.4304/jsw.6.4.732-739

neighbors of the target item and it produces prediction of the target user to the target item using item-based collaborative filtering. The recommendation method combining association rules mining and collaborative filtering can alleviate the data sparsity problem in the recommender systems.

II. Traditional Collaborative Recommendation Algorithm

A. User Item Rating Matrix

Filtering recommendation algorithm involves the traditional to the target user on the target item rating prediction that the user does not give ratings based on observed user ratings project coordination tasks. And user evaluation of the project is in the central database. Each user is represented by the term rating pairs in the user list, which contains the ratings have been expressed by the rij i-j-user items provided, the following table [7,8].

TABLE I
USER-ITEM RATINGS TABLE

Item / User	I1	I2	… …	In
U1	r11	r12	… …	r1n
U2	r21	r22	… …	r2n
… …	… …	… …	… …	… …
Um	rm1	rm2	… …	rmn

Where rij denotes the score of item j rated by an active user i. If user i has not rated item j, then rij =0. The symbol m denotes the total number of users, and n denotes the total number of items.

B. Measuring the Rating Similarity

Collaborative filtering method has been generally researchers and has a similar number of publications and the actual implementation of the case proved practitioners. Although there are many algorithms, the basic idea is common to calculate some of the measures used in the proposed project on the basis of the similarity of the similarity of the user. Collaborative filtering algorithm is the similarity between users, use is known as user-based collaborative filtering [9,10].

A similarity is between a set of measures and the correlation between two vectors measures. When these vectors are related to the time value of similarity is called user-based model similar to the user, and when they and when it is called project-based model for similar projects related. Similar measures can be effectively used to balance the prediction algorithm in the meaning of the ratings, therefore, to improve accuracy.

There are several similarities in the algorithm collaborative filtering algorithms [1,3]: Pearson correlation, cosine vector similarity, adjusted cosine vector similarity, mean square deviation and Spearman correlation.

Pearson’s correlation, as following formula, measures the linear correlation between two vectors of ratings.

$$sim(p,q)=\frac{\sum_{t\in Ipq}(R_{pc}-A_p)(R_{qc}-A_q)}{\sqrt{\sum_{t\in Ipq}(R_{pc}-A_p)^2\sum_{t\in Ipq}(R_{qc}-A_q)^2}} \quad (1)$$

Where Rpc is the rating of the item c by user p, Ap is the average rating of user p for all the co-rated items, and Ipq is the items set both rating by user p and user q.

The cosine measure, as following formula, looks at the angle between two vectors of ratings where a smaller angle is regarded as implying greater similarity.

$$sim(p,q)=\frac{\sum_{k=1}^{n}R_{pk}R_{qk}}{\sqrt{\sum_{k=1}^{n}R_{pk}^2\sum_{k=1}^{n}R_{qk}^2}} \quad (2)$$

Where Rpk is the rating of the item k by user p and n is the number of items co-rated by both users. And if the rating is null, it can be set to zero.

The adjusted cosine, as following formula, is used in some collaborative filtering methods for similarity among users where the difference in each user’s use of the rating scale is taken into account.

$$sim(p,q)=\frac{\sum_{t\in I_{pq}}(R_{pc}-A_c)(R_{qc}-A_c)}{\sqrt{\sum_{t\in I_{pq}}(R_{pc}-A_c)^2*\sum_{t\in I_{pq}}(R_{qc}-A_c)^2}} \quad (3)$$

Where Rpc is the rating of the item c by user p, Ac is the average rating of user p for all the co-rated items, and Ipq is the items set both rating by user p and user q.

Literature provides a collaborative filtering method successfully demonstrated abundant evidence. However, there are some ways inadequate. Collaborative filtering method is called sparse data is fragile and has cold start problems. Data sparsity refers to the lack of data, or sparse problems. Reference to cold start issues proposed new projects or where new user recommendations with difficulties and can not provide them with adequate rating in the systems.

C. Selecting Neighbors

Choose a neighbor who will serve as a referee. Both techniques have been employed in the collaborative filtering recommendation system.

Threshold-based selection, according to the similarity of the user exceeds a critical value as the target user's neighbors think.

Top - N technology, the best nitrogen neighbors and N is given first.

D. Producing Prediction

Since we have got the membership of user, we can calculate the weighted average of neighbors’ ratings, weighted by their similarity to the target user.

The rating of the target user u to the target item t is as following:

$$P_{ut} = A_u + \frac{\sum_{i=1}^{c}(R_{it} - A_i) * sim(u,i)}{\sum_{i=1}^{c} sim(u,i)} \tag{4}$$

Where A_u is the average rating of the target user u to the items, Rit is the rating of the neighbour user i to the target item t, Am is the average rating of the neighbour user i to the items, sim(u, i) is the similarity of the target user u and the neighbour user i, and c is the number of the neighbours.

III. Related Works

As described in [11], the authors discuss learning a profile of user interests for recommending information sources such as web pages or news articles. They describe the types of information available to determine whether to recommend a particular page to a particular user. This information includes the content of the page, the ratings of the user on other pages and the contents of these pages, the ratings given to that page by other users and the ratings of these other users on other pages and demographic information about users. They describe how each type of information may be used individually and then discuss an approach to combining recommendations from multiple sources. They illustrate each approach and the combined approach in the context of recommending restaurants.

Hybrid approaches use elements of both methods to improve performance and overcome shortcomings. In the reference [12], they propose a hybrid approach based on content-based and collaborative filtering, implemented in MoRe, a movie recommendation system. They also provide empirical comparison of the hybrid approach to the base methods of collaborative and content-based filtering and draw useful conclusions upon their performance.

Reference [13] proposes a novel, unified, and systematic approach to combine collaborative and content-based filtering for ranking and user preference prediction. The framework incorporates all available information by coupling together multiple learning problems and using a suitable kernel or similarity function between user-item pairs.

The huge volume of distributed information that is nowadays available in electronic multimedia documents forces a lot of people to consume a significant percentage of their time looking for documents that contain information useful to them. In previous work [14], they suggest a model for the automation of content-based electronic document filtering, supporting multimedia documents in a wide variety of forms. The model is based on multi-agent technology and utilizes an adaptive knowledge base organized as a set of logical rules. Implementations of the model using the client-server architecture should be able to efficiently access documents distributed over an intranet or the Internet.

With the development of e-commerce and the proliferation of easily accessible information, recommender systems have become a popular technique to prune large information spaces so that users are directed toward those items that best meet their needs and preferences. In reference [15], the authors describe a new filtering approach that combines the content-based filter and collaborative filter to capitalize on their respective strengths, and thereby achieves a good performance. They present a series of recommendations on the selection of the appropriate factors and also look into different techniques for calculating user-user similarities based on the integrated information extracted from user profiles and user ratings.

Collaborative filters are frequently used in e-commerce to provide a heightened user experience and to tempt users into making purchases by recommending items and drawing the user's attention to additional products. Purchasing of digital media over the Internet continues to be popular and e-commerce giants such as Amazon.com, CDNOW.com and Launch.com heavily employ Automated Collaborative Filtering. Reference [16] demonstrates a system for comparing musical compositions and provides an indication of how similar two or more musical pieces are to each other. It is shown that a significant amount of similarity exists between music compositions analyzed from within the same genre. It is proposed that a similarity metric could be incorporated into existing systems to provide a powerful and effective recommendation system that will cater specifically for a user's preferences, and thus encourage purchase.

In order to have an effective command of the relationship between customers and products, as described in [17], the authors have constructed a personalized recommender system which incorporates content-based, collaborative filtering, and data mining techniques. They also introduced a new scoring approach to determine customers' interest scores on products. To demonstrate how the system works, they used it to analyze real cosmetic data and generate a recommender score table for sellers to refer to.

In previous work [18], they conduct a broad and systematic study on different mixture models for collaborative filtering. They discuss general issues related to using a mixture model for collaborative filtering, and propose three properties that a graphical model is expected to satisfy. Using these properties, they thoroughly examine five different mixture models, including Bayesian Clustering , Aspect Model , Flexible Mixture Model , Joint Mixture Model , and the Decoupled Model . They compare these models both analytically and experimentally.

Instead of performing content indexing or content analysis, collaborative filtering systems rely entirely on interest ratings from members of a participating community. Since predictions are based on human ratings, collaborative filtering systems have the potential to provide filtering based on complex attributes, such as quality, taste, or aesthetics. The reference [19] provides a set of recommendations to guide design of neighborhood-based prediction systems, based on the results of an empirical study. They apply an analysis framework that divides the neighborhood-based prediction approach into three components and then examines variants of the key parameters in each component. The three components identified are similarity computation, neighbor selection, and rating combination.

In the reference [20], the authors present an expert software agent, named Traveller, that assists users in the tourism and travel domain. This agent combines collaborative filtering with content-based recommendations and demographic information about customers to suggest package holidays and tours. The combination of techniques in this hybrid approach takes advantage of the positive aspects of each technique and overcomes the difficulties shown by each of them when used in isolation. The results obtained when evaluating the agent demonstrate the benefits of using a combined technique to specify experts' knowledge.

Reference [21] contains a sample of the research carried out by us in this important area, focusing the work towards two of its most representative techniques: "Content Based filtering" and "Collaborative filtering." These techniques have been studied from different points of view, allowing to create a solid framework which involves the necessary criteria for designing and creating a tool using the most outstanding characteristics of each technique. They provide a view to facilitate the work of people devoted to the search, depuration and distribution of information.

The two of the most famous techniques in collaborative filtering are the so-called User-Based collaborative filtering and Item-Based collaborative filtering. As described in [22], the authors claim that each of them takes only one-directional information from the user-item ratings matrix to generate recommendations. In other words, the former combines user similarities and the latter tries to make a prediction by utilizing item similarities. They observe the same appearance in the other collaborative filtering area using binary user-item matrix in which transactions, i.e. purchase or non-purchase, are marked. It means that they may use only half of the total information from the given data set. Completing the missing part of usable information they proposed a new prediction method, two-way cooperative collaborative filtering which takes both vertical and horizontal information, in the ensemble respect. The proposed prediction scheme does not fix its collaborative filtering technique but associates two predictions, which come from different collaborative filtering algorithms, by weighted averaging. To decide fair weights the four cases, equivalent case, user-winning case, item-winning case, and prediction-impossible case are categorized by measuring the amount of information which each collaborative filtering utilizes, or the degree of the reliability of a prediction model. They also embedded bagging in the prediction frame to make more accurate predictions.

Reference [23] proposes a novel, unified approach that systematically integrates all available training information such as past user-item ratings as well as attributes of items or users to learn a prediction function. The key ingredient of the method is the design of a suitable kernel or similarity function between user-item pairs that allows simultaneous generalization across the user and item dimensions. They propose an online algorithm that generalizes perceptron learning.

The reference [24] proposes the integrated contextual information as the foundation concept of multidimensional recommendation model, and uses the online analytical processing ability of data warehousing to solve the contradicting problems among hierarchy ratings. The evaluation studies show that by establishing additional customer profiles and using multidimensional analyses to find the key factors affecting customer perceptions, the proposed approach increases the recommendation quality.

In the previous work [25], the authors develop recommendation algorithms with provable performance guarantees in a probabilistic mixture model for collaborative filtering proposed by Hoffman and Puzicha. They identify certain novel parameters of mixture models that are closely connected with the best achievable performance of are commendation algorithm; they show that for any system in which these parameters are bounded, it is possible to give recommendations whose quality converges to optimal as the amount of data grows.

IV. EMPLOYING ASSOCIATION RULES MINING TO SMOOTHING

A. *Association rules mining*

Mining association rules, mining is one of the most studied in data mining. It serves as a useful tool for finding correlations between items in large databases. It will explore the possibility that a specific item, when there are present any other items in the same transaction. Association rules X and Y are of the form are two sets of items X => Y is a condition. The interpretation of the commerce in the context of association rules, the customer, X entries, for anyone wishing to, Y to buy when you purchase the product.

The apriori is the important algorithm in the algorithms of association rules mining. The main idea of the apriori is scanning the database repeatedly. The most important step in mining association is the generation of frequent item sets. In apriori algorithm, most time is consumed for scanning the database repeatedly [4,6].

Let I = { i1 , i2 , …, im } be a set of all items, where an item is an object with some predefined attributes. A transaction T = < tid , It > is a duple, where tid is the identifier of the transaction. A transaction database T consist s of a set of transactions. An itemset is a subset of the set of items.

Definition 1: An association rule takes the form X =>Y where X < I , Y < I , and X ∩ Y = O. The support of the rule X =>Y in the transaction database is :

support (X =>Y) = | { T ：X ∪ Y ∪ T , T ∈ D} | / | D |

Definition 2: The confidence of the rule X =>Y in transaction database is :

confidence (X => Y) = | { T ：X ∪ Y ∪ T , T ∈ D} | / | { T ：X < T , T ∈ D} |.

B. *Mapping user-item Matrixes to Transactions*

Collaborative filtering user-item ratings data are usually represented as preference matrixes. They will change the transaction database for mining association rules. Each transaction includes a transaction ID and content. TID is the transaction ID of the user's user ID for the transaction to which they belong. The content of the item ID and assessments have been evaluated by the user.

C. *Algorithm*

The Apriori algorithm calculates the frequent item sets in a database using many repeated iterations. All the frequent item sets calculated in the ith iteration are called k item sets. Each iteration consists of two steps: generating the candidate item sets, and calculating and choosing the candidate item sets. Its kernel thought is as follows[5,6]:

(1) L1 = {Large 1-Item };
(2) for (k = 2; kk - 1 ! = 0; k + +)
(3) Ck = Apriori-gen (Lk - 1) ;
(4) for all transaction t∈D do begin
(5) Ct = SubSet (Ck , t) ;
(6) for all candidates c∈Ct do
(7) c. count + +;
(8) end
(9) Lk = { c∈Ck };
(10) end
(11) UkLk
(12) end

The essence of Apriori algorithm is that all the non empty subitems of frequent itemsets must be frequent. It covers two steps: conjunction and pruning.

V. USING SELF-ORGANIZING MAP TO FORM NEAREST NEIGHBORS

A. *Self-organizing map(SOM)*

Self-organizing map was first proposed in 1981 by the Finnish scholar based Kohonen. As an unsupervised learning neural network model, soil organic matter have been widely used in many fields because it is brought forward. SOM network structure consists of two layers, the upper layer and the output is the input layer to the next. The number of neurons in the acquisition layer is responsible for the nod that the data input and the number of variables is the same. One-dimensional or two-dimensional network formed in the output layer, and network to ensure that the domain of the relationship. SOM network is a whole network connection structure, each input layer neurons nod to connect all neurons in output layer nodded. When the input vectors the Euclidean and some weight, which link neurons in the output layer nod from the input layer, at least, nod these neurons are activated the appropriate weight, as the output of the network behavior. At the same time, the connection weights are modified and become more continuous input vector, output neurons, also known as winning neuron, and the corresponding amendments to the connection weights until the termination limit satisfaction [26].

The Self-organizing map training algorithm proposed by Kohonen is summarized as follows [27].

Step 1. Initialization: Choose random values for the initial weights wj (0).

Step 2. Winner Finding: Find the winning neuron j* at time k, using the minimum-distance Euclidean criterion

$$j^* = \arg\min_j \| x(k) - w_j \|,\ j = 1, ..., N^2 \qquad (5)$$

where x(k) = [x1(k), · · · , xn(k)] represents the kth input pattern, N^2 is the total number of neurons, and || · || indicates the Euclidean norm.

Step 3. Weights Updating: Adjust the weights of the winner and its neighbors, using the following rule:

$$w_j(k+1) = w_j(k) + p(k)N_{j^*}(k)(x(k) - w_j(k)) \quad (6)$$

where p(k) is a positive constant and $N_{j^*}(k)$ is the topological neighborhood function of the winner neuron at time k. It should be emphasized that the success of the map formation is critically dependent on how the values of the main parameters (i.e., p(k) and $N_{j^*}(k)$), initial values of weight vectors, and the number of iterations are prespecified.

B. *Using SOM to cluster items*

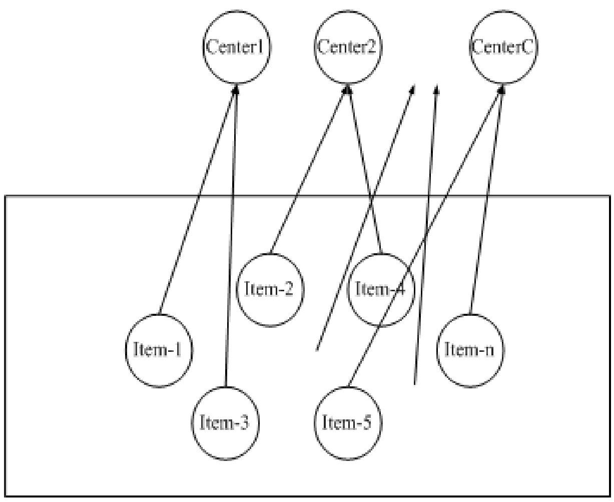

Fig 1. Employing SOM to cluster items

The item-based collaborative filtering recommendation system model, we focus on the SOM's clustering ability. For collaborative filtering algorithms, the formation of neighbor is an important step. We believe that the function of self-organizing map clustering of excellence, first formed in Figure 1 shows a neighbor recently used this method the target project.

VI. PRODUCING THE PREDICTION

A. Measuring the item rating similarity

There are several similarity algorithms that have been used [28,29,30]: Pearson correlation, cosine vector similarity, adjusted cosine vector similarity, mean-squared difference and Spearman correlation.

Pearson’s correlation, as following formula, measures the linear correlation between two vectors of ratings as the target item t and the remaining item r.

$$sim(t,r)=\frac{\sum_{i=1}^{m}(R_{it}-A_t)(R_{ir}-A_r)}{\sqrt{\sum_{i=1}^{m}(R_{it}-A_t)^2\sum_{i=1}^{m}(R_{ir}-A_r)^2}} \quad (7)$$

Where R_{it} is the rating of the target item t by user i, R_{ir} is the rating of the remaining item r by user i, A_t is the average rating of the target item t for all the co-rated users, A_r is the average rating of the remaining item r for all the co-rated users, and m is the number of all rating users to the item t and item r.

The cosine measure, as following formula, looks at the angle between two vectors of ratings as the target item t and the remaining item r.

$$sim(t,r)=\frac{\sum_{i=1}^{m}R_{it}R_{ir}}{\sqrt{\sum_{i=1}^{m}R_{it}^2\sum_{i=1}^{m}R_{ir}^2}} \quad (8)$$

Where R_{it} is the rating of the target item t by user i, R_{ir} is the rating of the remaining item r by user i, and m is the number of all rating users to the item t and item r.

The adjusted cosine, as following formula, is used for similarity among items where the difference in each user’s use of the rating scale is taken into account.

$$sim(t,r)=\frac{\sum_{i=1}^{m}(R_{it}-A_i)(R_{ir}-A_i)}{\sqrt{\sum_{i=1}^{m}(R_{it}-A_i)^2\sum_{i=1}^{m}(R_{ir}-A_i)^2}} \quad (9)$$

Where R_{it} is the rating of the target item t by user i, R_{ir} is the rating of the remaining item r by user i, A_i is the average rating of user i for all the co-rated items, and m is the number of all rating users to the item t and item r.

B. Prediction using item-based collaborative filtering

Since we have got the membership of item, we can calculate the weighted average of neighbors’ ratings, weighted by their similarity to the target item.

The rating of the target user u to the target item t is as following:

$$P_{ut}=\frac{\sum_{i=1}^{c}R_{ui}\times sim(t,i)}{\sum_{i=1}^{c}sim(t,i)} \quad (10)$$

Where Rui is the rating of the target user u to the neighbour item i, sim(t, i) is the similarity of the target item t and the neighbour item i, and c is the number of the neighbours.

VII. DATASET AND MEASURENMENT

In this section, we describe the dataset and metrics for the collaborative filtering algorithm.

A. Data Set

We use MovieLens collaborative filtering settings to assess the algorithm's performance data. MovieLens data set collected GroupLens research project, University of Minnesota and MovieLens is a web-based research recommender system, launched in autumn one thousand nine hundred ninety-seven in. MovieLens weekly visits to hundreds of users to rate and get movie recommendations [3 , 31,32]. The site now has more than 45,000 who have expressed different views on the movie 6600 users. We randomly selected enough users to get 100,000 in the 1680 movie rating of at least 20 per user ratings and simple statistical information for the user from the included 1000 users. The five-point rating is a number that negative 1 and grade 2,4 and 5 represents a positive evaluation, said the scale of the contradictions and 3.

B. Performance Measurement

Several indicators have been proposed to evaluate the accuracy of the collaboration filtering recommended method. They fall into two categories: statistical indicators of accuracy and decision-support accuracy metrics [8,9,33,34].

The accuracy of statistical data evaluation, a more realistic user ratings from their ratings of the forecast error of the numerical prediction accuracy. Some of them used the average absolute error (MAE), root mean square error (RMSE) between the ratings and forecasts, and relevant. The results of these indicators are calculated data, and generally provided the same conclusion. The accuracy of the statistical measures, the average absolute error (MAE) is employed.

Formally, if n is the number of actual ratings in an item set, then MAE is defined as the average absolute difference between the n pairs. Assume that r1, r2, r3, ..., rn is the prediction of users' ratings, and the corresponding real ratings data set of users is s1, s2, s3, ..., sn. See the MAE definition as following:

$$MAE=\frac{\sum_{i=1}^{n}|r_i-s_i|}{n} \quad (11)$$

The lower the MAE, the more accurate the predictions would be, allowing for better recommendations to be formulated. MAE has been computed for different prediction algorithms and for different levels of sparsity.

VIII. CONCLUSIONS

With the development of the Internet, the problem of information overload is becoming increasing serious. People all have experienced the feeling of being overwhelmed by the number of new books, articles, and proceedings coming out each year. Many researchers pay more attention on building a proper tool which can help users obtain personalized resources. Recommender systems can help people to find interesting things and they are widely used in our life with the development of electronic commerce. The commonly used personalized recommendation system methods are content-based filtering, collaborative filtering, and association rules mining. Unfortunately, each method has its drawbacks.

In this paper, we presented a personalized collaborative filtering recommendation method combining the association rules mining and self-organizing map. It used the association rules mining to fill the vacant where necessary. Then, it employs clustering function of self-organizing map to form nearest neighbors of the target item and it produces prediction of the target user to the target item using item-based collaborative filtering. The recommendation method combining association rules mining and collaborative filtering can alleviate the data sparsity problem in the recommender systems.

REFERENCES

[1] Songjie Gong, Employing User Attribute and Item Attribute to Enhance the Collaborative Filtering Recommendation, Journal of Software, Volume 4, Number 8, October 2009, pp: 883-890.

[2] Yi-Fan Wang, Yu-Liang Chuang, Mei-Hua Hsu, Huan-Chao Keh. A personalized recommender system for the cosmetic business. Expert Systems with Applications 26 (2004) 427–434.

[3] Songjie Gong, A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering, Journal of Software, Volume 5, Number 7, July 2010, pp: 745-752.

[4] LI Pingxiang, CHEN Jiangping, BIAN Fuling, A Developed Algorithm of Apriori Based on Association Analysis, Geo-spatial Information Science, Volume 7, Issue 2, 2004

[5] TAN Ying, YIN Guofu, LI Guibing, CHEN Jianying, Mining Compatibility Rules from Irregular Chinese Traditional Medicine Database by Apriori Agorithm. Journal of Southwest Jiaotong University (English Edition) Vol.15, No.4, 2007

[6] Songjie Gong, Personalized Recommendation System Based on Association Rules Mining and Collaborative Filtering, Applied Mechanics and Materials, Volume 39, pp:540-544.

[7] Yu Li, Liu Lu, Li Xuefeng, A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce, Expert Systems with Applications 28 (2005) 67–77.

[8] George Lekakos, George M. Giaglis, A hybrid approach for improving predictive accuracy of collaborative filtering algorithms, User Model User-Adap Inter (2007) 17:5–40.

[9] Breese J, Hecherman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI'98). 1998. 43~52.

[10] Goldberg D, Nichols D, Oki BM, Terry D. Using collaborative filtering to weave an information tapestry. Communications of the ACM, 1992,35(12):61~70.

[11] MICHAEL J. PAZZANI, A Framework for Collaborative, Content-Based and Demographic Filtering, Artificial Intelligence Review 13: 393–408, 1999.

[12] George Lekakos & Petros Caravelas, A hybrid approach for movie recommendation, Multimed Tools Appl (2008) 36:55–70

[13] Justin Basilico,Thomas Hofmann, A Joint Framework for Collaborative and Content Filtering, SIGIR'04, July 25–29, 2004

[14] N. S. PAPASPYROU, C. E. SGOUROPOULOU and E. S. SKORDALAKIS, A Model of Collaborating Agents for Content-Based Electronic Document Filtering, Journal of Intelligent and Robotic Systems 26: 199–213, 1999.

[15] Byeong Man Kim & Qing Li & Chang Seok Park & Si Gwan Kim & Ju Yeon Kim, A new approach for combining content-based and collaborative filters, J Intell Inf Syst (2006) 27: 79–91

[16] Cunningham, S., Bergen, H., & Grout, V., A Note onContent-Based Collaborative Filtering of Music, IADIS 5th-8th October (2006).

[17] Yi-Fan Wang, Yu-Liang Chuang, Mei-Hua Hsu, Huan-Chao Keh, A personalized recommender system for the cosmetic business, Expert Systems with Applications 26 (2004) 427–434

[18] Rong Jin • Luo Si • Chengxiang Zhai, A study of mixture models for collaborative filtering, Inf Retrieval (2006) 9:357–382

[19] JON HERLOCKER, JOSEPH A. KONSTAN, JOHN RIEDL, An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms, Information Retrieval, 5, 287–310, 2002

[20] Schiaffino, S., & Amandi, A., Building an expert travel agent as a software agent, Expert Systems with Applications (2008), doi:10.1016/j.eswa.2007.11.032

[21] José de J. Pérez-Alcázar , Maritza L. Calderón-Benavides , Cristina N. González-Caro, Towards an Information Filtering System in the Web Integrating Collaborative and Content Based Techniques, Proceedings of the First Conference on Latin American Web Congress, p.222, November 10-12, 2003

[22] Jong-Seok Lee, Sigurdur Olafsson, Two-way cooperative prediction for collaborative filtering recommendations, Expert Systems with Applications: An International Journal Volume 36 , Issue 3 April 2009

[23] J. Basilico and T. Hofmann. Unifying collaborative and content-based filtering. the21st International Conference on Machine Learning (ICML), 2004.

[24] S.S. Weng, B.S. Lin, and W.T. Chen, "Using ContextualInformation and Multidimensional Approach forRecommendation," Expert Systems with Applications, 36(2),2009, pp. 1268-1279

[25] Jon Kleinberg , Mark Sandler, Using mixture models for collaborative filtering, Proceedings of the thirty-sixth

annual ACM symposium on Theory of computing, June 13-16, 2004, Chicago, IL, USA
[26] ZHOU ShaoHua, FU Lue, LIANG BaoLiu, Clustering analysis of ancient celadon based on SOM neural network, Science in China Series E: Technological Sciences, 2008, 51(7):999-1007.
[27] WANG Ling, MU Zhi-Chun,GUO Hui, Combining Self-organizing Feature Map with Support Vector Regression Based on Expert System,ACTA AUTOMATICA SINICA, 2005, 31(4):612-619.
[28] Jong-Seok Lee, Chi-Hyuck Jun, Jaewook Lee, Sooyoung Kim, Classification-based collaborative filtering using market basket data, Expert Systems with Applications 29 (2005) 700–704.
[29] Hyung Jun Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Information Sciences 178 (2008) 37-51.
[30] Songjie Gong, An Efficient Collaborative Recommendation Algorithm Based on Item Clustering, Lecture notes in electrical engineering, Volume 72, pp:381-387.
[31] Gao Fengrong, Xing Chunxiao, Du Xiaoyong, Wang Shan, Personalized Service System Based on Hybrid Filtering for Digital Library, Tsinghua Science and Technology, Volume 12, Number 1, February 2007,1-8.
[32] M.G. Vozalis, K.G. Margaritis, Using SVD and demographic data for the enhancement of generalized Collaborative Filtering, Information Sciences 177 (2007) 3017–3037.
[33] Songjie Gong, An Enhanced Similarity Measure Used in Personalized Recommendation Algorithms, Advanced Materials Research, Volume 159, pp:671-675.
[34] George Lekakos, George M. Giaglis, Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors, Interacting with Computers 18 (2006) 410–431.

Hongwu Ye was born in Ningbo, Zhejiang Province, P.R.China, in 1976. He received B. Sc degree from Zhejiang University and M. Sc degree from Zhejiang University of technology, P.R. China in 2000 and 2006 respectively. He is currently a teacher in Zhejiang Textile & Fashion College, Ningbo, P.R.China.