Association for Information Systems

AIS Electronic Library (AISeL)

ICEB 2004 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-5-2004

MOBICORS-Movie: A MOBIle COntents Recommender System for Movie

Jae Kyeong Kim

Yoon Ho Cho

Seungtae Kim

MOBICORS-Movie: A MOBIle COntents Recommender System for Movie

Jaekyeong Kim[1], Yoonho Cho[2], Seungtae Kim[1]

[1] School of Business Administration, Kyung Hee University, Seoul, 130-701, Korea
[2] School of e-Business, Kookmin University, Seoul, 136-702, Korea
jaek@khu.ac.kr, www4u@kookmin.ac.kr

ABSTRACT

In spite of the rapid growth of mobile multimedia contents market, most of the customers experience inconvenience, lengthy search processes and frustration in searching for the specific multimedia contents they want. These difficulties are attributable to the current mobile Internet service method based on inefficient sequential search. To overcome these difficulties, this paper proposes a MOBIle COntents Recommender System for Movie (MOBICORS-Movie), which is designed to reduce customers' search efforts in finding desired movies on the mobile Internet. MOBICORS-Movie consists of three agents: CF (Collaborative Filtering), CBIR (Content-Based Information Retrieval) and RF (Relevance Feedback). These agents collaborate each other to support a customer in finding a desired movie by generating personalized recommendations of movies. To verify the performance of MOBICORS-Movie, the simulation-based experiments were conducted. The experiment results show that MOBICORS-Movie significantly reduces the customer's search effort and can be a realistic solution for movie recommendation in the mobile Internet environment.

Keywords: *mobile commerce, recommender, collaborative filtering, agent, relevance feedback*

1. INTRODUCTION

As mobile Internet technology becomes more increasingly applicable, the mobile multimedia contents market has recorded remarkable growth. In spite of this rapid growth, however, most of the customers experience inconvenience, lengthy search processes and frustration in searching for the specific multimedia contents they want [3]. These difficulties are attributable to the current mobile Internet service method based on inefficient sequential search:

- When customers log on to the mobile Internet site using a mobile phone, they are presented with the best-seller list or a list of multimedia contents in sequence of most recent introduction.
- Customer pages through the list and selects an entry to check out its contents hoping that the contents will be what he or she wants.
- If the customer likes the contents, he or she may make a purchase. Otherwise, the customer repeats the same steps until the customer either has stumbled over the right one or decides to give up.

Using current service method, the expected number of contents the customer views before he or she hits the desired movie far exceeds the acceptable level. Inconvenience, lengthy search processes, and frustration make customers easily annoyed by the amount of irrelevant and uninteresting information they have to filter out. During a search, as more time is consumed, frustration builds with the result that the purchase conversion rate drops. To make search processes more acceptable, a more efficient searching aid that suggests only the movies meeting the individual customer's preference is needed. This paper proposes a MOBIle COntents Recommender System for Movie (MOBICORS-Movie), which is designed to reduce customers' search efforts in finding desired movies on the mobile Internet. The system combines two of the most popular information filtering techniques: Collaborative Filtering (CF) and Content-Based Information Retrieval (CBIR).

MOBICORS-Movie consists of three agents: CF, CBIR and RF (Relevance Feedback). These agents collaborate each other to support a customer in finding a desired movie by generating personalized recommendations of movies.

The simulation-based experiments were conducted to verify the performance of MOBICORS-Movie. The experiment results show that MOBICORS-Movie significantly reduces the customer's search effort, so it can be a realistic solution for multimedia contents recommendation in the mobile Internet environment. We believe that deployment of MOBICORS-Movie offers the following benefits to both consumers and suppliers of mobile multimedia contents:

- Customers can purchase mobile contents with much less search effort and much lower connection time to the mobile Internet, because they can much more easily find desired mobile contents.
- Mobile contents providers can improve the profitability of their business because lower customer frustration in finding desired contents increases revenue through an improved purchase conversion rate.

2. EXISTING TECHNIQUES

2.1 Collaborative Filtering

The recommender system is one of possible solutions to searching for individually preferred contents from a large-contents database. A recommender system is defined as a system that assists customers in finding the items they would like to purchase. One of the most successful recommendation techniques is Collaborative Filtering (CF) [7,8], which has been widely used in a number of different applications.

Collaborative filtering is an information filtering technique that depends on human beings' evaluations of items. It is an attempt to automate the "word of mouth" recommendations that we receive on a daily basis from our family, friends, and colleagues. It identifies customers whose tastes are similar to those of a given customer and it recommends items those customers have liked in the past. In general, CF-based recommender systems make recommendations according to the following steps [4,7]: (1) A customer provides the system with preference ratings on items that may be used to build a customer profile. (2) The system applies statistical or machine learning techniques to find a set of customers, known as neighbors, who had in the past exhibited similar behaviors (i.e., either they had purchased a similar set of items or they had given the set similar ratings). A neighborhood is formed based on the degree of similarity between a target customer and other customers. (3) Once a neighborhood is formed for a target customer, the system generates a set of items that the target customer is most likely to purchase by analyzing the items in which neighbors have shown an interest (*top-n recommendation*).

Although the CF is the most successful recommendation technique, it suffers from two major shortcomings. First, when there is a shortage of ratings, CF suffers from a sparsity problem [1,2,4,7]. Most similarity measures used in CF work properly only when there exists an acceptable level of ratings across customers in common. An increase in the number of customers and items worsens the sparsity problem, because the likelihood of different customers rating common items decreases. Such sparsity in ratings makes the formation of neighborhoods inaccurate, thereby resulting in poor recommendations. Next, CF suffers from a new item problem [1,5]. Since CF recommends an item based on customers' ratings on the item, it does not recommend a newly introduced item until some ratings of the item become available. The new item problem becomes even worse when the turnover rate of items is high. The Third issue is related to *scalability* [4,7]. Recommender systems for mobile customers have to deal with large amount of multimedia contents and millions of customers. Because these systems usually handle very high-dimensional profiles to form the neighborhood, the nearest neighbor algorithm is often very time-consuming and scales poorly in practice.

2.2 Contents-based Information Retrieval

Content-based information retrieval represents a movie as a point in the multi-dimensional feature space and performs similarity-based retrieval using its contents-based keywords. In CBIR, customer describes movies using a query that is a set of example keywords. A query is internally represented as multiple points (i.e. query points) that have keywords of example movies. In general, CBIR systems retrieve movies according to the following steps [3,9]: (1) A system presents a query to the customer via the results of collaborative filtering as a request for desired movies. (2) The system searches for movies similar to the query. The similarity between a movie in the database and a query is calculated using the distance between corresponding points in the feature space. (3) The movies with the highest degree of similarity are retrieved and recommended to the customer.

In spite of the virtues of CBIR in retrieving movies similar to a query, CBIR rarely brings a customer to the desired movies immediately. The reason for this is that any combination of example movies may not precisely represent the movies that a customer desires. For a system to handle this gap properly, it needs the ability to learn about what movie the customer really wants through iterative interactions. The customer's current preference on the presented movies needs to be fed back so that CBIR can learn from this preference to retrieve, in the next iteration, movies more similar to the one customer really wants. This learning process, the *relevance feedback*, is an essential mechanism for a faster search of desired movies. The degree of preference of a movie in MOBICORS-Movie is expressed in three-level (e.g. *preferred, neutral, unpreferred*) weights. We will refer to a set of preferred movies as a *preferred set*. The movies in the preferred set are used for query refinement for the purpose of learning customer's current preference. And relevance feedback is performed using Rocchio's algorithm [6].

3. MOBICORS-Movie

3.1 Recommendation Procedure

MOBICORS-Movie is designed to reduce customers' search efforts in finding desired movies on the mobile Internet. The system consists of three agents: CF, CBIR and RF (Relevance Feedback). These agents collaborate each other to support a customer in finding a desired movie by generating personalized recommendations of movies. Figure 1 shows the recommendation procedure of MOBICORS-Movie.

First, an initial recommended movie list is generated by CF agent and shown up to the mobile customer. A customer selects one of the recommended movies using just a movie title. Then the customer sees the detailed information of selected movie, such as brief introduction of the movie, director, actors, actress, and so on. If he/she buys the movie, then he/she can watch the movie by mobile phone after payment process. If the customer

replied as "unprefered", the recommended movie list generated by CF agent except the selected movie is shown up to the mobile customer.
If the customer doesn't buy the movie but replied as "preferred", CBIR agent and RF agent search for similar movie(s). Then the customer sees the detailed information of selected movie. The procedure is terminated until the customer buys the movie or disconnects the searching process.

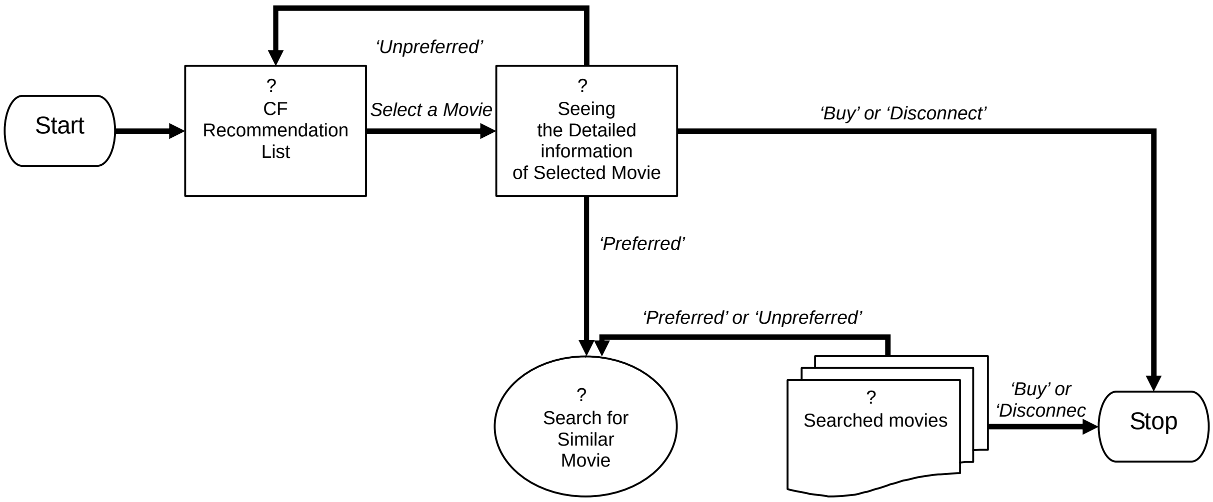

Figure 1. The procedure of MOBICORS-Movie

3.2 System Architecture

The architecture of MOBICORS-Movie is shown at Figure 2.

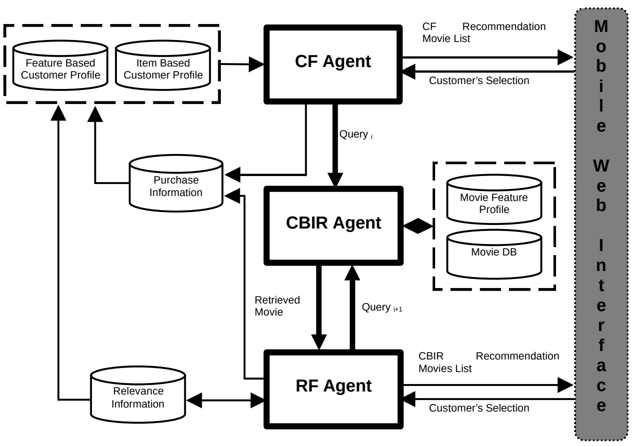

Figure 2. The architecture of MOBICORS-Movie

The CF agent generates a list of recommended movies and provides an initial movie to the CBIR agent. This agent creates the customer profile using purchase and preference information to identify neighbors and generate recommendations. When the CF-generated recommendation list is presented, a customer skims through the list to see if there are any movies of interest. Then, the customer selects an entry to view the movie.

After viewing, the customer may decide to purchase the movie or decide whether to use the movie as a query for content-based search of similar movies or to go back to the CF-generated recommendation list.

When the customer decides to use the viewed movie as a starting query for further search, the viewed movie is passed to the CBIR agent as an initial query, and the agent retrieves movies based on similarity between the query and other movies in the database. For all movies in the database, this agent calculates the similarities from the query and generates a list of the most similar movies as recommendations. It then passes the retrieved movies to the RF agent. RF agent presents the retrieved movies to the customer one by one, and interactively elicits the user's preference judgment on the presented movies.

At any point in this presentation session, the customer may decide to buy a movie or decide to quit. After all of preference judgment, either preferred or unpreferred, are made, the RF agent updates the preference information and purchase databases with all the fed back preference and/or purchase information respectively for later use by the CF agent when the customer revisits the site. If all of presented movies are marked as unpreferred, the search session returns to the CF-generated recommendation list. Otherwise, the RF agent learns the customer's current preference using his/her preference judgment, formulates the new query, and passes the query to the CBIR agent for the next iteration of retrieval.

3.3 Profile Management

Movie Profile: A movie profile includes information about the characteristics of movies. A movie profile is represented by matrix $V=(v_{jk})$, where j and k implies jth movie and kth keyword. If jth movie contains kth keyword, the value of v_{jk} is 1, otherwise it is 0. For example, the number of movies is 7, and that of keywords is 6, an example of movie profiles is represented as Table 1.

Table 1. A Movie Profile

	Betrayal	Crime	Cult-favorite	Drama	Police	Twist-in-the-end
Godfather	1	1	0	1	1	1
Pulp Fiction	0	1	1	1	0	0
Reservoir Dogs	0	1	1	0	1	1
Fight Club	0	1	1	1	0	1
Fargo	0	1	1	0	1	0
Psycho	0	0	1	0	1	1
Goodfellas	1	1	0	1	1	0

Feature-based Customer Profile: A feature-based customer profile represents customer's preference on keywords. It is composed of customer-keyword matrix $F = (f_{ik})$, where f_{ik} implies the preference degree of ith customer on kth keyword. For example, the number of customers is 4, and that of keywords is 6, an example of feature-based customer profiles is represented as Table 2.

Table 2. A Feature-based Customer Profile

	Betrayal	Crime	Cult-favorite	Drama	Police	Twist-in-the-end
Kim	5	10	0	5	-10	0
Cho	10	5	-5	5	5	10
Kang	5	15	5	10	-5	0
Lee	0	10	0	10	0	-5

Item-based Customer Profile: An item-based customer profile represents customer's preference on movies. It is composed of customer-movie matrix $P = (p_{ij})$, where p_{ij} implies the preference degree of ith customer on jth movie. For example, the number of customers is 4, and that of movies is 7, an example of item-based customer profile is represented as Table 3..

An item-based customer profile includes information about customer's preferences on movies. Generation of CF recommendations depends totally on item-based customer profile. MOBICORS-Movie builds the item-based customer profile using purchase records and fed-back binary preference information, either preferred or unpreferred.

Table 3. An Item-based Customer Profile

	Godfather	Pulp Fiction	Reservoir Dogs	Fight Club	Fargo	Psycho	Goodfellas
Kim	5	15	0	15	0	-10	5
Cho	15	-5	5	5	-5	0	10
Kang	15	10	15	15	15	0	5
Lee	10	5	5	10	5	-5	5

The profile is updated as the following equation:

$$p_{ij} = \begin{cases} +2 & \text{if customer } i \text{ has purchased movie } j \\ +1 & \text{if customer } i \text{ has marked movie } j \text{ as preferred} \\ -1 & \text{if customer } i \text{ has marked movie } j \text{ as unprefrred} \\ 0 & \text{if customer } i \text{ has not seen movie } j \end{cases}$$

.

Note that the brief introduction of movie that customer viewed from CF recommendation list and decided to use or not to use as an initial query for CBIR is counted as a preferred or unpreferred movie, respectively. As shown in above equation, the cells of the rating matrix have four possible different values. We place the highest ratings on the previously purchased movies, because they should reflect the customer's taste the most strongly. The movies marked as preferred or unpreferred are represented in ratings of half the magnitude of the purchased ones with an opposite sign. Assigning -1 to unpreferred images gives a better chance to unseen movies over unpreferred ones in generating recommendations. Please note that further research work is necessary to justify heuristic choice of values used for preference ratings, since their relative sizes in magnitude could affect recommendations of CF.

The ratings in customer profile are constantly replaced with newly obtained purchase and preference information to dynamically reflect customer's most recent preference. This is significantly different from the customer profiles used in traditional CF techniques [7,8].

4. EXPERIMENTAL EVALUATION

4.1 Experiment

For the purpose of the performance evaluation of MOBICORS-Movie, we developed a Web-based application system running on a PC with exactly the same user interfaces as the mobile phone-based MOBICORS-Movie system. Using the system, we carried out the experiments with the intent of answering three major questions:

(1) How much performance improvement does MOBICORS-Movie deliver as his/her search and purchase information in customer profile is accumulated?

(2) How does the number of movies CBIR agent search affect the performance of MOBICORS-Movie?

(3) How much performance improvement does MOBICORS-Movie deliver compared to other recommender systems in mobile phone environment?

For the experiments, the 250 movies from IMDB(www.imdb.com) were used. MOBICORS-Movie used the top 100 keywords sorted by frequency from

IMDB. Virtually generated 1000 mobile Web customers are participated in the experiment. They are set to purchase 5 movies randomly, and their initial purchase information is stored as an initial feature based customer profile and item based customer profile.

For performance evaluation of MOBICORS-Movie, a metric ***views-per-success*** (*vps*), defined as the number of movies viewed by a customer before he/she purchases a movie in a search session, was devised. The *vps* measures the amount of effort a customer takes per successful search [3].

4.2 Results and discussion

The quality of recommendations of CF is known to vary by the size of the neighborhood [4,7]. The neighborhood size yielding optimal performance depends on customer profile, whose contents change all the time. For this reason, comparison between optimal performances of different periods requires increasing number of additional experiments to calculate the optimal neighborhood for every period. However, considering the burden of participants in the experiment, we used the optimal neighborhood size of the first period and fixed the neighborhood size for the rest of the experiment. According to Figure 4, the optimal neighborhood size was determined to be 40. For the rest of the experiment, we fixed the neighborhood size at 40.

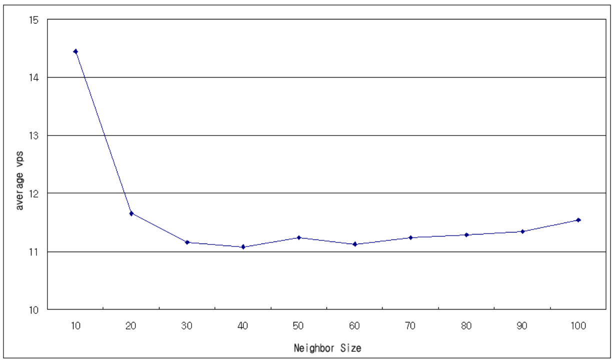

Figure 3. Average *vps* over diverse neighborhood sizes

To answer the question 1 and 2, we observed the variation in *vps* over periods and *n*, number of nearest neighbor movies retrieved in CBIR. This is observed as a decreasing curve in Figure 4, which shows the overall performance of MOBICORS-Movie in terms of *vps* vs. period for five different values of *n*.

As periods move on, more rating information becomes available and the customer profile with more ratings makes neighborhood formation more accurate, thereby improving the quality of CF recommendations. Figure 4 indicates that the variation in *vps* by *n* is significant. Furthermore, it is observed as the difference in *vps* by values of *n* especially in early periods, and as the diminishing trend of difference in *vps* over periods in Figure 4.

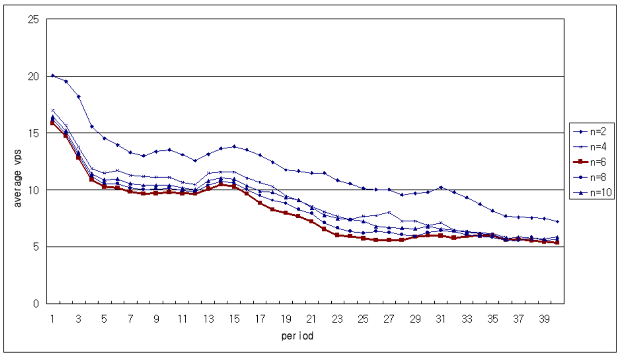

Figure 4. Overall performance of MOBICORS-Movie

Furthermore, as the value of *n* becomes from 2 to 6, the performance becomes better, but the value of *n* goes over 6, the performance becomes worse. In general, sparsity level of customer profile and heterogeneity in customers' buying behaviors are the most important factors influencing the effectiveness of CF [2,3,4]. In cases with high level of either sparsity or heterogeneity, CF's performance is poor, thus we select optimal value of *n*. The value of *n* can be selected considering other factors such as customer's burden of feedback in CBIR or the firm's marketing strategy toward the diversity in recommended movies.

To answer the question 3, the performance of MOBICORS-Movie is compared to those of two other recommender systems, the best-seller-based system (Bestseller) and a typical CF-based recommender system (pure-CF), whose procedure is identical to that of CF-based recommendation of MOBICORS-Movie except that, in building a customer profile, it uses purchase information only.

As compared in Figure 5, the *vps* of MOBICORS-Movie is about 60% and 72% lower, at 40 periods, than that of the pure-CF and best-seller-based systems respectively. In other periods, the performance gab between MOBICORS-Movie and other techniques is higher. Figure 5 also shows that the rates of improvement in *vps* over the 10 periods (i.e. learning speed of the system) of MOBICORS-Movie, pure-CF and best-seller-based systems are 50%, 68% and 55% respectively. These results show that, of the three systems evaluated, MOBICORS- Movie offers not only the lowest *vps* but also the most robust performance over 10 periods. The superior performance of MOBICORS-Movie over the pure-CF results from MOBICORS-Movie's accelerated learning of customer preference from additional preference rating information fed back from CBIR. This evidences that MOBICORS-Movie, as intended, successfully overcomes the sparsity problem of CF.

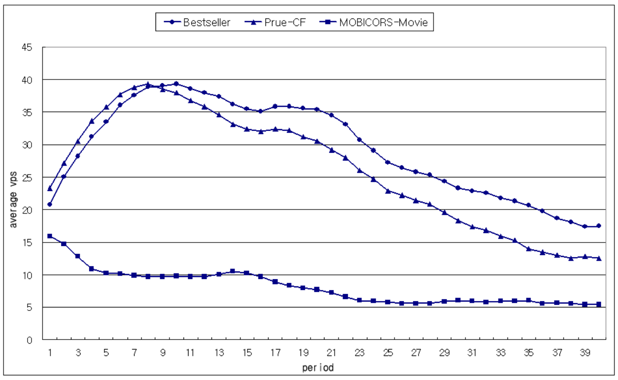

Figure 5. Comparision of MOBICORS-Movie and Other Recommendation Techniques

5. CONCLUSION

It is devised a recommender system, MOBICORS-Movie, to address the real problems encountered in movie watching services in the mobile Web environment. The system combines two techniques from different research domains: recommender system and information retrieval. By employing two techniques, MOBICORS-Movie generates recommendations with the following characteristics: (1) MOBICORS-Movie recommends movies by reflecting the opinions of other customers with similar tastes. (2) MOBICORS-Movie recommends movies with similar keyword features to the ones the customer currently prefers. (3) MOBICORS-Movie recommends movies by learning the customer's preference adaptively via feedback on recommended movies.

To verify the performance of MOBICORS-Movie, the simulation experiments were conducted. From the experiment result, we can conclude that MOBICORS-Movie is a viable solution to the problems currently encountered in movie watching on the mobile Web, and that it can be expected to reduce the search effort, thereby increasing the purchase conversion rate. Use of MOBICORS-Movie offers the following benefits to both consumers and suppliers of mobile contents: (1) Customers can purchase contents with much less search effort and much lower connection time to the mobile Web, because they can much more easily find desired mobile contents. (2) Mobile contents providers can improve the profitability of their businesses because lower customer frustration in finding desired contents increases revenue through an improved purchase conversion rate.

With the rapid growth of the mobile Web service, the mobile Web-based recommender system for other types of multimedia contents, such as music on demand (MOD) will continue to be an area of research interest in the future.

REFERENCES

[1] Balabanovic, M., Y. Shoham, "Fab: Content-based Collaborative Recommendation", *Communication of the ACM*, Vol. 40, No. 3, pp 66-72, 1997.
[2] Cho, Y.H., J.K. Kim, S.H. Kim, "A Personalized Recommender System Based on Web Usage Mining and Decision Tree Induction", *Expert Systems with Applications*, Vol. 23, No. 3, pp 329-342, 2002.
[3] Kim, C.Y., J.K. Lee, Y.H. Cho, D.H. Kim, "VISCORS: a Visual Contents Recommender System on the Mobile Web", *IEEE Intelligent Systems, Special issue on Mining the Web for Actionable Knowledge*, Accepted for publication, 2004.
[4] Kim, J.K., Y.H. Cho, "Using Web Usage Mining and SVD to improve E-Commerce Recommendation Quality", *LNAI 2931: Intelligent Agents and Multi-Agent Systems*, pp 86-97, 2003.
[5] Melville, P., R.J. Mooney, R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations", *Proc. the 18th National Conference on Artificial Intelligence*, pp187-192, 2002.
[6] Rocchio, J.J., "Relevance Feedback in Information Retrieval", In Salton. G., editor, *The SMART Retrieval System - Experiments in Automatic Document Processing*, pp.313-323. Prentice Hall Inc., Englewood Cliffs, NJ., 1971.
[7] Sarwar, B., G. Karypis, J.A. Konstan, J. Riedl, "Analysis of recommendation algorithms for e-commerce", *Proceedings of the ACM Conference on Electronic Commerce*, pp 158-167, 2000.
[8] Schafer, J.B., J.A. Konstan, J. Riedl, "E-commerce recommendation applications", *Data Mining and Knowledge Discovery*, Vol. 5, No. 1-2, pp 115-153, 2001.
[9] Zhou, X.S., T.S. Huang, "Relevance feedback for image retrieval: a comprehensive review," *ACM Multimedia Systems Journal*, Vol. 8, No. 6, pp 536-544, 2003.