A Hybrid Multimedia Contents Recommendation Procedure for a New Item Problem in M-commerce*

Jae Kyeong Kim
School of Business Administration, Kyung Hee University
(jaek@khu.ac.kr)

Yoon Ho Cho
School of e-Business, Kookmin University
(www4u@kookmin.ac.kr)

Mi Yeon Kang
School of Business Administration, Kyung Hee University
(jaek@khu.ac.kr)

Hyea Kyeong Kim
School of Business Administration, Kyung Hee University
(jaek@khu.ac.kr)

Currently the mobile web service is growing with a tremendous speed and mobile contents are spreading extensively. However, it is hard to search what the user wants because of some limitations of cellular phones. And the music is the most popular content, but many users experience frustrations to search their desired music. To solve these problems, this research proposes a hybrid recommendation system, MOBICORS-music (MOBIle COntents Recommender System for Music). Basically it follows the procedure of Collaborative Filtering (CF) system, but it uses Contents-Based (CB) data representation for neighborhood formation and recommendation of new music. Based on this data representation, MOBICORS-music solves the new item ramp-up problem and results better performance than existing CF systems. The procedure of MOBICORS-music is explained step by step with an illustrative example.

Keywords : Collaborative filtering, recommender system, mobile commerce, multimedia contents

Received: November 2004 Accepted: June 2006 Corresponding Author: Jae Kyeong Kim

1. Introduction

Recently as the wireless devices have been distributed very rapidly, the mobile web service is spreading more extensively. Among this fast growth of mobile web contents, the most popular contents are mainly multimedia contents like music, character image, picture and etc. Especially, the music contents

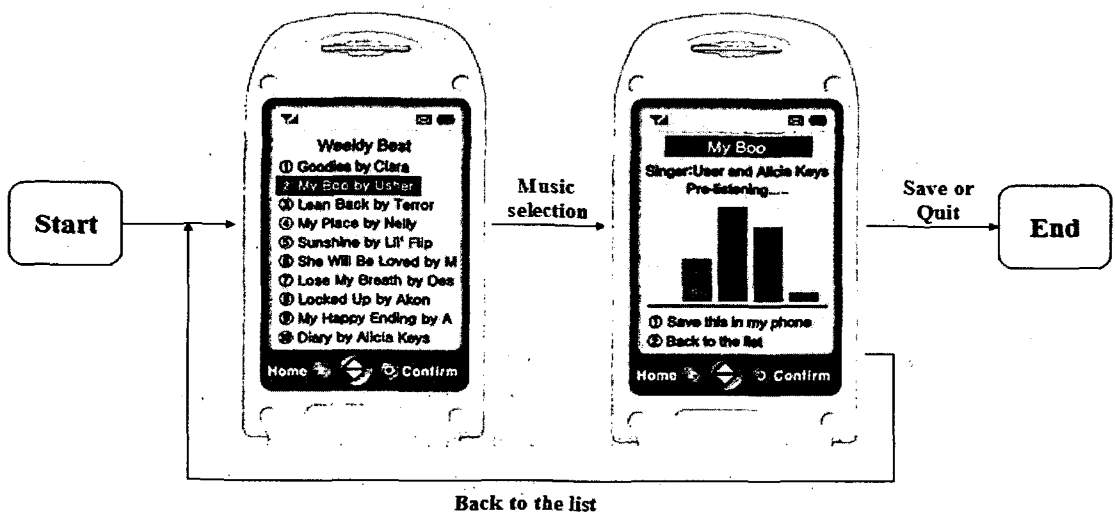
* This research was supported by the Kyung Hee University Research Fund in 2005(KHU-20050410).

take a very large portion in the mobile content market (Korea IT Industry Promotion Agency, 2003).

In spite of the popularity and fast growth of music contents, many users experience frustration in searching for the desired music because of a sequential search process. When a user logs on to a music service site with his/her cellular phone, the user is presented with a weekly or monthly best-selling or newest music list. Then the user searches through the list and he/she might select an entry to inspect. If the user liked the music, he/she might save it in his/her cellular phone, otherwise, the user should repeat the same steps until he/she meets the desired music or decides to quit purchasing. [Fig. 1] shows the current search process in a cellular phone.

However, the current search process takes too much time and effortsearching for the desired music. This difficulty results from limitations of cellular phones such as small LCD, tiny keypad, and sophisticated browser compared to PC. Furthermore, users had to specify their preferences related to the features of music such as title, singer, composer, and genre. It would give much burden to them because the features are hard to be described obviously. As a result, sequential search process, limitations of cellular phones and difficulty of explaining music content make both content providers and mobile users need recommender systems.

The recommender system is designed to assist users in finding the items they would like to purchase. Collaborative Filtering (CF) is known to be the most successful recommender system and has been widely used in a number of different applications. Nevertheless, the existing CF systems have several problems such as sparsity, scalability, and new item ramp-up problems. There are a large number of researches for the sparsity and scalability problems. On the other hand, the new item ramp-up problem has not been



[Fig. 1] The current search process

solved yet. In mobile web environment, new music is frequently supplied, and their purchasing ratio is considerably high, therefore music recommender systems need to solve this new item ramp-up problem.

This papersuggests a MOBICORS-music, MOBIle COntents Recommender System for Music to recommend multimedia contents for mobile Web users. The basic idea of MOBICORS-music is CF, and it adapts contents-based (CB) filtering algorithm by representing items in feature space. MOBICORS-music finds neighbors using both the web log information of music and their featureinformation. As music is represented in the feature space, individual user is also represented as a cluster in the feature space because the user is related with his/her purchased music or pre-listed music set. Under this situation where music and user are represented in the feature space, MOBICORS-music tries to solve the new item ramp-up problem.

The rest of the paper is organized as follows. In section two, recommendation systems and the music features are explained. In section three, MOBICORS-music and its recommendation procedure are presented. The step-by-step explanation of the procedure of MOBICORS-music with an illustrative example is presented in section four. Finally conclusions and further research area are provided in section five.

2. Related Work

2.1 Collaborative Filtering

The recommender system is one of the possible solutions to searching for individually preferred contents from a large-content database. A recommender system is defined as a system that assists customers in finding the products or contents they would like to purchase.

CF is an information filtering technique that depends on human beings' evaluations of contents. It is an attempt to automate the "word of mouth" recommendations. It identifies customers whose tastes are similar to those of a given customer and it recommends contents those customers have liked in the past. In general, CF-based recommender systems make recommendations according to the following steps [3]: (1) A customer provides the system with preference ratings on contents 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. 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 contents that the target customer is most likely to purchase by analyzing the contents in which neighbors have shown an interest.

Although the CF is the most successful recommendation technique, it suffers from the following shortcomings. First, when there is a shortage of ratings, CF suffers from a sparsity problem [3-6]. Most similarity measures used in CF work properly only when there exists an acceptable level of ratings across customers in common. Next, CF has a scalability problem. This problem could be generated due to the rapidly growing users and items

(Kim et al., 2004). When CF system finds neighbors, it spends too much computation time, and scales poorly in practice. With ten millions of users and items, a typical mobile recommender system running existing algorithms will suffer serious scalability problem (Sarwar et al., 2001). Finally, CF suffers from a new item problem [4,6]. Since CF recommends an item based on customers' ratings on the contents, it does not recommend a newly introduced contents until some ratings of the contents become available. The new contents problem becomes even worse when the turnover rate of contents is high.

In attempts to address these three inherent problems of CF, researchers have proposed many variations of hybrid approaches that combine CF with other recommending techniques [4,5,6]. Despite their success in other applications, little of the previously proposed hybrid approaches can be an adequate solution for new music recommendations.

2.2 Content-based Filtering

Content-based (CB) filtering systems (Lang, 1995; Pazzani, Muramatsu & Billsus, 1996; Mooney & Roy, 1999) identify properties associated with the contents of interest, and recommend those contents that are most similar to contents that the given customer has liked in the past. CB systems build a user profile that is valuable when a customer encounters new content that has not been rated before. CB systems require a source of content information and do not provide much in the way of serendipitous discovery (Balabanović & Shoham, 1997; Good, et al., 1999). Most CB systems are thus ineffective in domains without explicit descriptions.

However, as a technique for music information retrieval (MIR) is progressed, researches for music genre classification have been growing with an amount of attention since digital music has emerged on the web. They assume that the components of particular music genre share certain characteristics typically related to the instrumentation, rhythmic structure, and pitch content of the music (Tzanetakis et al., 2002; Lambrou et al., 1998; Chen et al., 2001; Tao et al., 2003). There are a large number of feature sets used in music genre classification, mainly originating from the area of speech recognition, which has been proposed to represent audio signal. Recently many nonspeech signal analysis techniques are proposed. Especially, Tzanetakis et al.(2002) introduced the content based features, which contained features of timbral texture, rhythmic content, and pitch content. The timbral texture features are mostly defined by the nature of the timbral elements of the texture and the relationship between those elements and the timbral groups they form. The rhythmic content features characterize the movement of music signals over time and contain information about the regularity of the rhythm, beat and tempo. The pitch content features generally describe the melody and harmony information about music signals. The combination of these features is effective to classify the music genre, we thus use universally confirmed music features for CB filtering. And MARSYAS (Tzanetakis, 2000), known as one of the best music feature extractor in MIR, is adapted for the music feature extractor of MOBICORS-music.

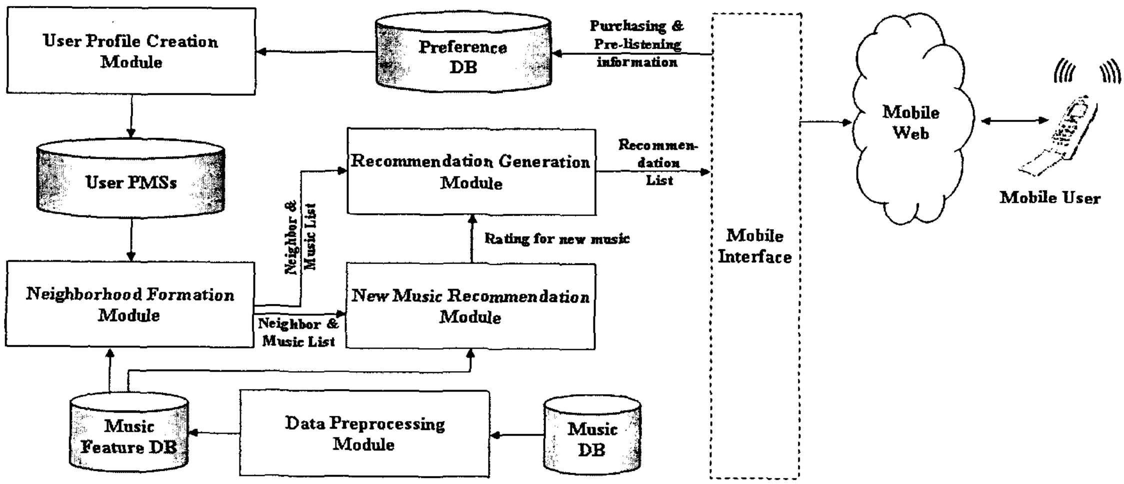
3. MOBICORS-music

MOBICORS-music is developed to recommend music contents for mobile web users. It adapts the CB technology by expressing music contents as a combination of features. The approach of MOBICORS-music would increase the number of ratings by the use of the pre-listened music information and make it possible for new music to be recommended using extracted features describe their perceptual properties. MOBICORS-music is organized with five modules, such as data preprocessing, user profile creation, neighborhood formation, recommendation generation and new music recommendation module. The overview of MOBICORS-music is shown in [Fig. 2]

Data preprocessing module converts the data format of music, extracts features of music, and organizes the music feature database. *User Profile Creation module* constructs the preferred music set (PMS) with the preference database. *Neighborhood Formation module* selects similar users to the target user by a distance-based inter-user similarity. This module calculates the distance between the pre-listened and purchased music set of the target user and those of other users represented in the feature space. Recommendation Generation module creates a recommendation list for the target user. This module computes the Purchase Likeliness Scores (PLS) of the music, and organizes the recommendation list composed of music with higher score. New Music Recommendation module determines whether to recommend new music, or not. This module would define the target user and his/her neighbors' preferred boundaries, and if some new music belonged to them, then the new music is recommended.

3.1 Data Preprocessing

This module extracts features from music in



[Fig. 2] Architecture of MOBICORS-music

music database and then stores the extracted features into music feature database. Our music feature database is represented by a music-feature matrix M = (mik) where i=1 to n, k=1 to l, i and k denote a music content and a music feature respectively. Music-feature matrix M is used in customer profile creation module and new music recommendation module.

3.2 Customer Profile Creation

Customer profile creation module creates the Preferred Music Set (PMS) from the customers' behavior. PMS is composed of purchased and pre-listened music and their ratings. The rating r_i^a, a degree of preference of customer a on the corresponding music m_i, is settled implicitly as the following equation (1):

$$r_i^a = \begin{cases} 1.0 & \text{If customer a has purchased the music } m_i \\ 0.5 & \text{If customer a has pre-listened the music } m_i \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

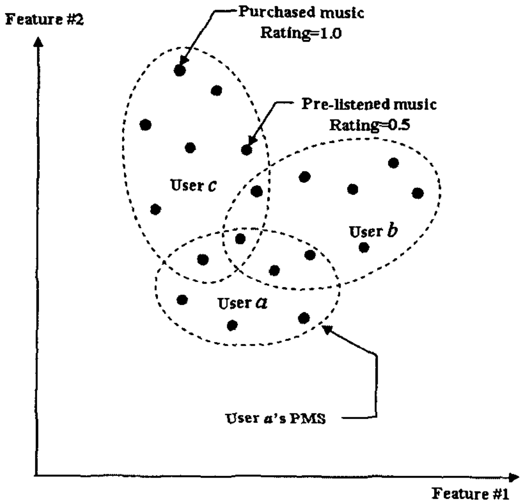
As equation (1) shows, the rating r_i^a has three possible values. We rate the previously purchased music contents highest because they should reflect the customer's taste the most strongly. The pre-listened music has ratings that are one half of the magnitude of the purchased ones. We assign zero to unlisted music. *This definition of preference rating is adapted* from prior researches (Lawrence et al., 2001; Kim et al., 2003).

With the music m_i and corresponding rating r_i^a, the user a's PMS is defined as follows:

$$\text{PMS}^a = \{(m_i, r_i^a) \mid only\ if\ r_i^a \neq 0\} \quad (2)$$

The PMS is constantly updated with newly obtained information and used to search for both similar customers as neighbors and relevant new music contents as recommendations.

The illustration of the PMS is shown in [Fig. 3] to help the understanding of PMS more easily.



[Fig. 3] PMSs in the feature space

User *a, b,* and c are repressented in two dimensional feature space with their preferred musics. Represented by PMS, the system can find out each user's music preference. For example, the preference of user ais more affected by the feature #2 because his/her preferred musics are centered on specific value of feature #2. In the same way, the preference of user c is more intended on the feature #1 than feature #2.

3.3 Neighborhood Formation

MOBICORS-music finds the neighbors using Centroid Euclidean distance function as a similarity measure. All the music in PMS are represented by one weighted centroid as its representative, where rating values of each music are applied to weights of own features to incorporate the difference between the purchased music and pre-listened music. And then, the distance between weighted centroids is used as sim(c,a) to denote the similarity between a target customer c and other customer a. It is determined the neighbor set $H = \{h1, h2, ..., hL\}$ such that $c \notin H$ and $sim(c, h_1)$ is the highest, $sim(c, h_2)$ is the next highest, and so on. We calculate the similarity using the weighted Centroid Euclidean distance function as follows:

$$sim(c,a) = \frac{\underset{b \in H}{Max}[\,d(c,b)\,] - d(c,a)}{\underset{b \in H}{Max}[\,d(c,b)\,] - \underset{b \in H}{Min}[\,d(c,b)\,]} \qquad (3)$$

$$d(c,a) = \sqrt{\frac{\sum_{k=1}^{l} (O_k^c - O_k^a)^2}{l}} \qquad (4)$$

,where b implies any customer in the neighbor set H, and $d(c,a)$ is a distance function between the target user c and other customer b, and $Max[d(c,b)]$ and $Min[d(c,b)]$ denote the maximum and minimum distance between two customers c and b, respectively. Let O^a and O^b be the weighted centroids of PMS^a and PMS^b, respectively. Then O_k^a and O_k^b are kth feature value of O^a and O^b, respectively and l is the total number of features.

Equation (3) shows that the neighbor of target customer is constructed as the neighbors with shorter distance than others. As customers updates his/her PMS, PMS also updates the weighted centroid to reflect the customer's current preference. For this purpose, we define O_k^a and O_k^b in Equation (5) as:

$$O_k^a = \frac{\sum_{x=1}^{n^a} r_x^a \cdot m_{xk}^a}{n^a} \text{ and } O_k^b = \frac{\sum_{y=1}^{n^b} r_y^b \cdot m_{yk}^b}{n^b} \qquad (5)$$

,where m_x^a and m_y^b are the xth music in PMS^a and yth music in PMS^b, respectively. And m_{xk}^a and m_{yk}^b are the kth feature value of m_x^a and m_y^b, respectively. r_x^a and r_y^b are the rating value of the music mx rated by customer a and that of music my by customer b, respectively, and na and nb are the total number of music contents in PMS^a and PMS^b, respectively.

3.4 Recommendation Generation

MOBICORS-music generates a recommendation list of Nmusic contents for a target customer c, $RL = \{R_1, R_2, ... R_i, ..., R_N\}$, such that Ri{the music that c has already purchased}. For the RL, PLS(c, mi) is used which denotes the Purchase Likeliness Score of the target customer c on music mi, which is purchased by the members in H, but not purchased by the target customer c. That is, the PLS(c, R1) is the highest, and PLS(c, R2) is the next highest, and so on ; PLS(c, mi) is computed as the following equation (6):

$$PLS(c,m_i) = \frac{\sum_{a \in H} (r_{m_i}^a - \overline{r^a}) \times sim(c,a)}{\sum_{a \in H} sim(c,a)} \qquad (6)$$

,where $\overline{r^a}$ is the customer a's average rating, and $r_{m_i}^a$ is the customer a's rating on the music m_i.

3.5 New Music Recommendation

New Music Recommendation module determines whether to recommend new music to target customer, or not. For the recommendation of new music, the preference area of each customer is assumed to be the feature space composed of the features of preferred music. The basic idea of recommending new music is that if a new music is belonged into the preference area, then it could be preferred by the target customer. Thus, new music recommendation module would define the preferred boundary as preference area based on a customer's PMS. And if new music belongs to the boundaries of a target customer and his/her neighbors, then it searches for the most similar item within the preferred boundary and applies the rating value of the similar item for calculating the PLS of the new music.

3.5.1 Preference boundary

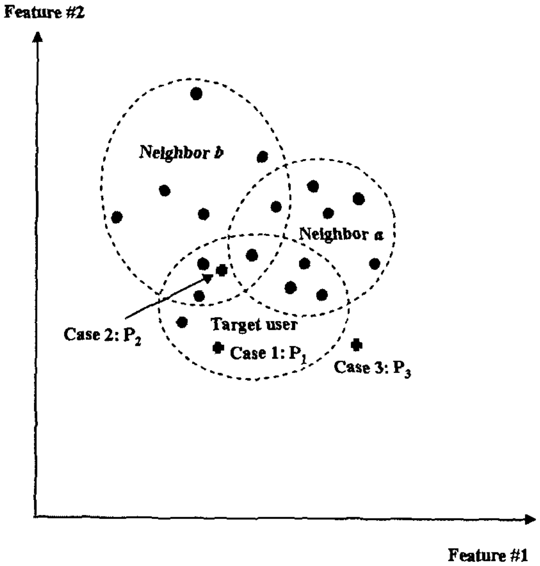
The problem with new music recommendation is that all customers might not prefer all of the new music contents. For example, it is inappropriate to recommend the newly introduced jazzmusic to a customer who likes music having hard core rhythm and strong beat. Therefore the system should find out what customer would prefer the new music. For such a purpose, the preference boundary of each customer is defined in the feature space from the music in each customer's PMS.

Target customer c's preference boundary is determined using the weighted centroid and standard deviation of PMSc. Based on customer c's preference boundary, new music Pjis recommended to target customer c if $O_k^c - s_k^c \le P_{jk} \le O_k^c + s_k^c$ for all k, where s_k^c is weighted standard deviation of musics in PMS^c and it can be calculated as $\sqrt{\frac{\sum_{x=1}^{n^c}(m_{xk}^c - O_k^c)^2}{n^c}}$. P_{jk} is the kth feature value of new music P_j, O_k^c is the kth feature value of O^c, the weighted centroid PMS^c, and m_{xk}^c is the kth feature value of m_x^c.

However the standard deviation function assumes that the number of music in PMS^c is over 30 and PMS^c follows the normal distribution. But in mobile web environment, the number of musics in their PMS might be smaller than 30. So t-preference boundary is defined by t-distribution with 95% confidence interval, and the confidence level is set up considering the recommending environment. New music P_j is recommended to target customer c with 95% confidence interval, if $O_k^c - t(0.025, n_c - 1) \times \frac{s_k^c}{\sqrt{n_c}} \le P_{jk}$ $\le P_{jk} \le O_k^c + t(0.025, n_c - 1) \times \frac{s_k^c}{\sqrt{n_c}}$, for all k, where $s_k^c = \sqrt{\frac{\sum_{x=1}^{n^c}(m_{xk}^c - O_k^c)^2}{n^c}}$.

In Case 1, and Case 2, P_1 and P_2 are recommended to the target user, but P_3 can not be recommended.

In this way, the system tells whether the user will prefer the new music, or not, but it does not still know how much preferences he/she has on the new music. The new music does not have any rating information so the system can not calculate the PLS of new music. Therefore, the system gives the PLS of the most similar music to the new music.



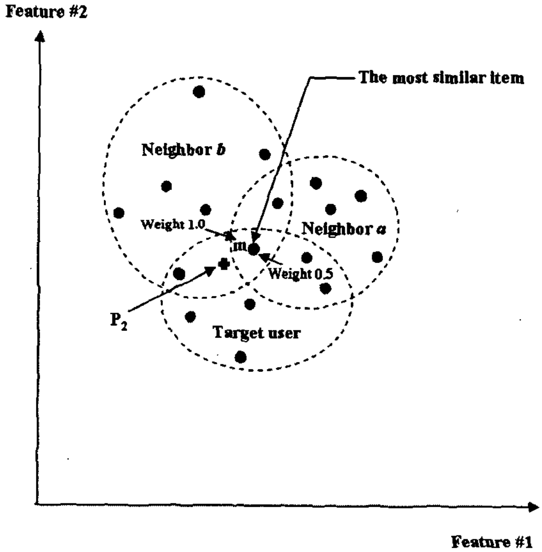
[Fig. 4] User preference boundaries

The procedure for the new music recommendation is organized in two phases: defining the extended neighbors' preference boundaries, and finding the most possible rating of new music.

3.5.2 Finding the Most Similar Music

As new music P_j does not have any rating, the system can not calculate its PLS. For this reason, MOBICORS-music finds the most similar music (MSM), which is defined as the music having the shortest distance to P_j in preference boundary. And then the rating value of new music P_j is substituted by that of MSM.

The illustration of this procedure is shown in [Fig. 5], where P_2 is included in both preference boundaries of target customer c and neighbor b. In this illustration, both $MSM^c_{P_2}$ and $MSM^b_{P_2}$ is music m, therefore P_2 can get the rating value of both r^c_m 0.5 and r^b_m 1.0. With these rating values, the system calculates the new music's PLS. Therefore the recommendation list RL is generated from music rated by neighbors and new music may be included in preference boundaries at the same time.



[Fig. 5] The MSM for new music

With these rating values, the system calculates the new music's PLS value. Therefore the recommendation list RL is composed of existing music and new music.

4. An Illustrative Example

In order to help the understanding of the procedure of MOBICORS-music, this section presents two illustrative examples including a CF based-recommendation procedure and a new music recommendation procedure.

4.1 The CF Recommendation Procedure

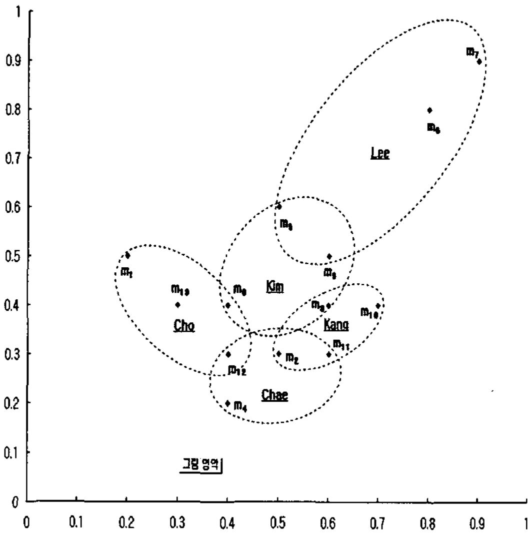
The CF recommendation procedure is explained with five users' PMSs. Five users are named as Kim, Lee, Cho, Kang, and Chaewhere, Kim is the target user. For the simplicity of this illustration, it is assumed that the number of music features is only two, the number of neighbors L is three, and the number of recommendation musics N is five.

The users' PMSs are represented as follows: Kim's PMS = $\{(m_3, 0.5), (m_5, 1.0), (m_8, 1.0), (m_9, 0.5)\}$, Lee's PMS = $\{(m_5, 0.5), (m_6, 1.0), (m_7, 1.0), (m_9, 1.0)\}$, Cho's PMS = $\{(m_1, 0.5), (m_8, 1.0), (m_{12}, 1.0), (m_{13}, 1.0)\}$, Kang's PMS = $\{(m_2, 1.0), (m_3, 1.0), (m_{10}, 1.0), (m_{11}, 0.5)\}$, and Chae's PMS = $\{(m_2, 0.5), (m_4, 1.0), (m_{11}, 1.0), (m12, 0.5)\}$. And the music feature database is shown at <Table 1>.

<Table 1> The music feature database

Music \ Feature	1	2
m_1	0.2	0.5
m_2	0.5	0.3
m_3	0.6	0.4
m_4	0.4	0.2
m_5	0.5	0.6
m_6	0.8	0.8
m_7	0.9	0.9
m_8	0.4	0.4
m_9	0.6	0.5
m_{10}	0.7	0.4
m_{11}	0.6	0.3
m_{12}	0.4	0.3
m_{13}	0.3	0.4

[Fig. 6] illustrates the preference boundaries of each user in two-dimensional feature space.



[Fig. 6] User Preference Area

To calculate the distance between users, all of the users' weighted centroids are calculated using equation (5), and the results are shown in <Table 2>.

<Table 2> The weighted centroid of five users

User \ Feature	1	2
Kim	0.375	0.363
Lee	0.638	0.625
Cho	0.3	0.338
Kang	0.525	0.313
Chae	0.363	0.2

With the weighted centroid of each user, the system calculates the distance between target user and other users. For instance, the d(Kim, Lee) is calculated as equation (7).

$$d(\text{Kim,Lee}) = \sqrt{\frac{(0.375-0.638)^2+(0.363-0.625)^2}{2}} = 0.321 \quad (7)$$

In the same way, other user's distances are also calculated and the results are d(Kim,Cho) = 0.077, d(Kim,Kang) = 0.154, and d(Kim,Chae) = 0.116. Thus Kim's neighborhood set H is {Cho, Kang, Chae}.

To compare the result with existing CF systems, the similarity between users is also computed with the Cosine formulation method, and the results are Cos(Kim, Lee) = 0.5, Cos(Kim, Cho) =0.25, Cos(Kim, Kang) = 0.25, Cos(Kim, Chae) = 0. As shown in [Fig. 6], Lee's preference is distributed to a very large area in the space, and it seems to be inappropriate to select Lee as Kim's neighbor. But, the existing CF systems cannot notice this, and then would select Lee as Kim's neighbor. Based on Lee's purchased music, the CF system would recommend incorrect musics such as m_6 and m_7 which are not satisfied by Kim. The feature-based data representation can be escaped from such a risk.

Before calculating the PLS, neighbors' similarities can be computed using the distances. For instance, Sim(Kim,Cho) is calculated as equation (15),

$$\text{sim(Kim,Cho)} = \frac{0.5-0.077}{0.5-0.01} = 0.863 \quad (15)$$

In the same way, other user's similarities are also calculated as follows: sim(Kim, Kang) = 0.706, sim(Kim, Chae) = 0.785 The music set preferred by neighbors are m_1, m_2, m_3, m_4, m_8, m_{10}, m_{11}, m_{12}, and m_{13}. As Kim has already purchased m_3 and m_8, they are excluded in the RL. Therefore, the PLSs of other musics are PLS(Kim, m_1) = 0.231, PLS(Kim, m_2) = 0.488, PLS(Kim, m_4) = 0.769, PLS(Kim, m_{10}) = 0.731, PLS(Kim, m_{11}) = 0.514, PLS(Kim, m_{12}) = 0.511, and PLS(Kim, m_{13}) = 0.731. These results are combined with the results of new music recommendation module for generating RL.

4.2 New Music Recommendation

Suppose that the new music such as m_{14}, m_{15} and m_{16} are provided. <Table 3> shows the feature profiles of them.

<Table 3> New music's profile

Feature / Music	1	2
m_{14}	0.36	0.33
m_{15}	0.45	0.55
m_{16}	0.74	0.12

In this illustration, the number of music in each user's PMS is four, so the t-preference boundary is used. The weighted centroids are presented in the <Table 3>, and the weighted standard deviations are computed, and the results are in <Table 4>.

<Table 4> The weighted standard deviation

Feature / User	1	2
Kim	0.096	0.18
Cho	0.141	0.075
Kang	0.171	0.118
Chae	0.18	0.071

In condition that confidence interval is 95%

and degree of freedom is three, t-value is 3.182. Therefore, each user's t-preference boundary is determined as shown in <Table 5>.

As you can see, m_{14} is included in Kim's, Cho's and Kang's preference boundary, and m_{15} is included in Kim's and Cho's preference boundary. But m_{16} cannot be included in any user's one. In next step, the system finds the individual MSM of selected new music set. In the case of m_{14}, the distance between m_{14} and other music in Kim's PMS are computed as follows: $d(m_{14},m_3)=0.101$, $d(m_{14},m_5)=0.215$, $d(m_{14},m_8)=0.057$, $d(m_{14},m_9)=0.071$. As a result, $MSM^{Kim}_{m_{14}}$ is m_8, and the rating value of m_8, 1.0 is given to that of m_{14}. In the same way, the $MSM^{Cho}_{m_{14}}$ is m_{12}, and $MSM^{Kang}_{m_{14}}$ is m_2. And m_{14} is given the rating values of each MSM's value, i.e., 1.0, 1.0, and 1.0, respectively.

<Table 5> t - Preference boundary

User \ Feature	1	2
Kim	(0.22, 0.53)	(0.08, 0.65)
Cho	(0.07, 0.53)	(0.3, 0.38)
Kang	(0.25, 0.8)	(0.12, 0.5)
Chae	(0.08, 0.65)	(0.09, 0.31)

In the case of m_{15}, as the same way, $MSM^{Cho}_{m_{15}}$ is m_{13}, and it is given 1.0 as the rating value of m_{13}. With the rating values of new music, the system would compute their PLSs as follows: PLS(Kim, m_{14}) = 0.751, PLS(Kim, m_{15}) = 0.731. Finally, the system can generate the last RL for Kim composed of both new music, i.e., RL = $\{m_4, m_{14}, m_{15}, m_{10}, m_{13}\}$.

5. Conclusion

Currently the mobile web service is growing with a tremendous speed and mobile contents are spreading extensively. It forces existing recommender systems to deal with the newest items and to ensure high quality of recommendations. In this paper, we focused on these challenging issues of the recommender systems in mobile environment and proposed MOBICORS-music, combining collaborative filtering with content-based filtering. And we made an example, which can show step by step how the system do find target user's neighbors, and how can recommend new music. In this example, we verify that MOBICORS-music is more accurate than the general CF, and generate the recommendation list including both new music and other existing music.

Use of MOBICORS-music is expected to offer the following benefits to both consumers and suppliers of mobile contents: (1) Customers can purchase contents with much less 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 business because lower customer frustration in finding desired contents increases revenue through an improved purchase conversion rate.

While MOBICORS-music is effective and efficient for content recommendations in mobile web environment, it also suggests a real-world application. And with the rapid growth of mobile commerce, the mobile Web based recommender system for other types of multimedia contents, such as video on demand (VOD), will continue to be an area of research interest in the future.

References

[1] Balabanovic, M. and Y. Shoham, "Fab: Content-based, Collaborative Recommendation", *Communications of the ACM*, Vol.40, No.3(1997), 66-72.

[2] Burke, R., "Hybrid Recommender Systems: Survey and Experiments", *User Modeling and User-Adapted Interaction*, Vol.12, No.4(2002), 331-370.

[3] Cho, Y. H. and J. K. Kim, "Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendation in e-Commerce", *Expert systems with application*, Vol.26, No.2(2004), 233-246.

[4] Everitt, B. S., S. Landau, and M. Leese, *Cluster Analysis*, Oxford University Press, 2001.

[5] Goto, M. and Y. Muraoka, "A Beat Tracking System for Acoustic Signals of Music", In Proceedings of the 2th ACM International Conference on Multimedia, San Francisco, USA, (1994), 365-372.

[6] Hung-Chen and L. P. Arbee, "A Music Recommendation System based on Music Data Grouping and User Interests", In Proceedings of the 10th international conference on Information and Knowledge management, Atlanta, USA, (2003), 231-238.

[7] Kim, C. Y., J. K. Lee, Y. H. Cho, and D. H. Kim, "VISCORS: a Visual Contents Recommender System on the Mobile web", *IEEE Intelligent Systems*, Vol.19, No.6(2004), 32-39.

[8] Kim, D. H. and C. W. Chung, "Qcluster: Relevance Feedback Using Adaptive Clustering for Content-Based Image Retrieval", In Proceedings of the 2003 ACM SIGMOD international conference on Management of data, San Diego, USA, (2003), 599-610.

[9] Kim, J. K. and Y. H. Cho, "Recommendation Systems for eCRM, Web Engineering: Principles and Techniques", by Idea Group, Inc., (2004), 1-29.

[10] Kim, J. K. and Y. H. Cho, "A Personalized Recommendation Procedure for Internet Shopping Support", *Electronic Commerce Research and Applications*, Vol.1, No.3-4(2002), 301-313.

[11] Kim, J. K and Y. H. Cho, "Using Web Usage Mining and SVD to Improve E-commerce Recommendation Quality", *Lecture Notes Computer Science: Intelligent Agents and Multi-Agent Systems*, (2003), 86-97.

[12] Korea IT industry Promotion Agency, *The Digital Contents Market Report*, 2003.

[13] Lambrou, P. Kudmakis, R. Speller, M. Sandler, and A. Linney, Classification of Audio Signals Using Statistical Features on Time and Wavelet Transorm Domains, In Proceeding International Conference Acoustic, Speech and Signal processing, Vol.6(1998), 3621-3621.

[14] Lawrence, R. D., G. S. Almasi, M. S. Kotlyar, and Viveros, S. S. Duri, "Personalization of Supermarket Product Recommendations", *Data Mining and Knowledge Discovery*, Vol. 5, No.1-2(2001), 11-32.

[15] Melville, P., et al., "Content-boosted Collaborative Filtering for Improved Recommendations", In Proceedings of the 18th National Conference on Artificial Intelligence, Edmonton, Canada, (2002), 187-192.

[16] Pazzani, M. J., "A Framework for Collaborative, Content-Based and Demographic Filtering", *Artificial Intelligence Review*, Vol. 13, No.5-6(1999), 393-408.

[17] Rabiner L. and B. H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall, NJ., 1993.

[18] Sarwar B. M., G. Karypis, J. A. Konstan, and J. T. Riedl, "Application of Dimensionality Reduction in Recommender System - a Case Study", In Proceedings of the 13th International Conference on World Wide Web, New York, USA, (2001), 393-402.

[19] Sarwar, B. M., G. Karypis, J. A. Konstan, J. T. Riedl, "Analysis of Recommendation Algorithms for E-Commerce", In Proceedings of ACM Conference on Electronic Commerce, Minneapolis, USA, (2000), 158-167.

[20] Sarwar, B. M., G. Karypis, J. A. Konstan, J. T. Riedl, "Item based Collaborative Filtering Recommendation Algorithms, 10th International World Wide Web Conference, ACM Press, Hong Kong, China, (2001), 285-295.

[21] Tao Li, Mitsunori O., Qi Li, "A Comparative Study on Content-based Music Genre Classification", In Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, Toronto, Canada, (2003), 282-289.

[22] Tzanetakis, G., "MARSYAS: A Framework for Audio Analysis", *Organized sound*, Vol.4, No.3(2000).

[23] Tzanetakis, G., "Automatic Musical Genre Classification of Audio Signals", *IEEE Transactions on Speechand Audio Processing*, Bloomington, Indiana, 2001.

[24] Uitdenbogerd, A. L., Justin Zobel, "Manipulation of Music for Melody Matching", Proceeding of the 6 th ACM International Conference on Multimedia, Bristol, United Kingdom, (1998), 235-240.

[25] Zhang, T., R. Ramakrishnan, M. Livny, "BIRCH: An Efficient Data Clustering Method for Very Large Databases", *ACM SIGMOD Record*, Vol.25, No.2(1996), 03-114.

요약

하이브리드 기법을 이용한 신상품 추천문제 해결방안에 관한 연구 : 모바일 멀티미디어 컨텐츠를 중심으로

김재경* · 조윤호** · 강미연* · 김혜경*

휴대폰, PDA 등 모바일 단말기의 급속한 진화와 광범위한 보급으로 인하여 모바일 웹 서비스가 빠르게 확산되고 있으며 모바일 컨텐츠 시장 또한 급성장하고 있다. 이에 따른 새로운 멀티미디어 컨텐츠의 활발한 공급은 모바일 웹사용자들에게 많은 멀티미디어를 획득할 수 있는 기회를 제공하는 동시에 정보과부하로 인한 컨텐츠 검색의 어려움을 겪게 하고 있다. 본 연구는 신상품에 대한 니즈가 높은 모바일 멀티미디어 컨텐츠의 특성과 기존 유선 웹 환경에 비해 열악한 모바일 웹 환경의 제약 사항을 고려하여, 모바일 웹 서비스 이용 고객이 보다 적은 노력과 비용으로 원하는 멀티미디어 컨텐츠를 신속하게 찾을 수 있도록 지원하는 개인화된 멀티미디어 컨텐츠 추천 방법론을 개발하는 것이다. 이를 위하여 기존 추천시스템에서 대표적으로 사용되는 협업필터링(Collaborative Filtering)기법의 한계를 보완하기 위하여 내용기반 필터링 기법(Content-based Filtering)을 결합한 하이브리드 추천 기법을 개발하였다. 제안한 하이브리드 기법은 모바일 환경에서 적은 계산으로도 높은 추천 성능과 함께 신상품추천이 가능한 방법이며, 이를 구현하기 위하여 멀티미디어 컨텐츠 추천시스템, MOBICORS-music(MOBIle COntents Recommender System for Music)을 개발하였다.

* 경희대학교 경영대학
** 국민대학교 e-비즈니스학부