ICEB 2004 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-5-2004

An Extended Collaborative Filtering-based Recommendation Procedure for Multimedia Contents in M-Commerce

Meeyeon Kang

Yoon Ho Cho

Jaekyeong Kim

An Extended Collaborative Filtering-based Recommendation Procedure for Multimedia Contents in M-Commerce

Meeyeon Kang [1], Yoonho Cho [2], Jaekyeong 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
kkangmee@lycos.co.kr, www4u@kookmin.ac.kr

ABSTRACT

As mobile market grows more and more fast, the mobile contents market, especially music contents for mobile phones have record remarkable growth. In spite of this rapid growth, mobile web users experience high levels of frustration to search the desired music. And new musics are very profitable to the content providers, but the existing CF system can't recommend them. To solve this problem, we propose an extended CF system to reflect the user's real preference by representing users in the feature space. We represent the musics using the music's content based acoustic feature like timbral, MFCCs, rhythmic, and pitch contents in multi-dimensional feature space, and then select a neighborhood with distance based function. And for new music recommendation, we match the new music with other users' preference.
To verify the performance of the proposed system, the simulation imitating the real user's decision-making and context in conducted. Through comparison with the pure CF, we validate our system's performance

Keywords: *collaborative filtering, recommender system, mobile commerce, music information retrieval*

1. INTRODUCTION

With the popularity of the mobile device, technologies and applications, the mobile contents market grows very rapidly. In mobile web environment the multimedia contents like musics, character images, and pictures are the most popular contents, and especially the music contents take a very large portion among them [7]. Although the popularity of the multimedia contents increases very fast, many users frustrate to search the music which they really want because of mobile phone's limitations like the small LCD, tiny keypad, and low memory and music's unique character. Music is hard to explain its feature like tempo, pitch, beat and other spectral features; therefore many users have to search for the desired music by scanning the offered music list one by one or keyword directly. For reducing mobile users' searching efforts and time, contents providers would adopt a recommender system.

Collaborative filtering (CF) is known as the most successful system in recommender systems. But the CF system results in many problems because of its input data representation. The general CF system uses a user-item binary matrix; therefore it is hard to find the real neighbors.

And in mobile environment new musics are very frequently supplied, and their purchasing ratio is considerably high, however existing CF systems can't recommend new musics. A new music doesn't be purchased by anyone, it can't have any ratings.

In this paper we propose the system which can solve these problems. This system extracts the music's various content-based acoustic features like MFCC, tempo, beat, pitch, and then represents these musics in a multi-dimensional feature space. As doing this, we can identify the neighbors who are really similar to the target. And we can recommend new musics by giving a virtual rating to them in the feature space. This paper describes how can solve these problems in mobile web environment.

The rest of the paper is organized as follows: Section 2 reviews the related work on CF system and music information retrieval. Section 3 gives a brief overview of the proposed system. Section 4 presents the detail recommendation procedure. Section 5 describes the experimental design and evaluations for system. Finally, Section 6 provides our conclusions.

2. RELATED WORK

2.1 Collaborative Filtering

Collaborative filtering (CF) is the most familiar, most widely implemented and most mature of the technologies in recommender system [1]. It is defined as one which makes recommendations by finding correlations among users of a recommender system [8]. The goal of CF system is to suggest items for a target user based on the user's previous preference and the opinions of other like-mined users. But the CF system has many problems like sparsity, and scalability problem, but these problems have been explored in [2,6,10,11].

Beyond these problems the CF system has a radical problem related its input data representation. The CF system uses a user profile which is composed with m user's ratings about n items, and is represented by the $m \times n$ user-item binary matrix. When a user has purchased a music, a general CF system gives a rating to the music. This causes that the system is hard to identify user's real neighbors because it uses only common musics for forming the target user's neighborhood. Therefore it is possible that someone whose preference is similar to the target user's preference didn't purchase common musics.

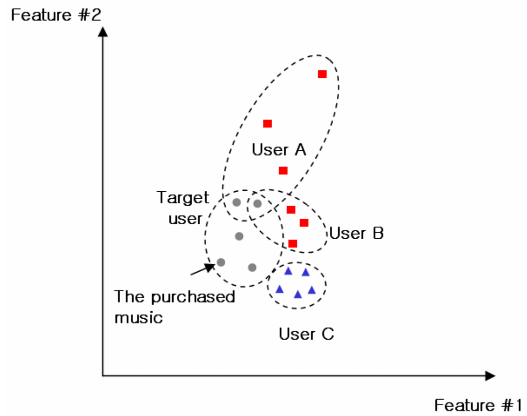

Figure 1. CF system in feature space

Figure 1 shows an existing CF system's problems. Generally the existing CF system selects user A as the target user's neighbor because he purchased the common music. But user A's preference is quite different from target user's preference when their preferred musics represent in feature space. In Figure 1, user B and user C are more similar to the target user rather than user A. Moreover the system can't find hidden neighbors such user C as the target user's neighbor.

And the CF system has the new-item ramp up problem. The new item ramp up problem is the most difficult and unsolved problem in the CF system. In the mobile web environment, new musics are very frequently provided, their purchasing ratio is very high, and many contents providers try to find out this problem. But the existing CF system can't recommend the new music.

In Section 4 we describe how to solve the problems related the input data representation and new item ramp-up problem.

2.2 Music Information Retrieval

As the popularity of digital music has been increasing, efficient and accurate automatic music information processing will be an extremely important issue, and many systems try to classify the music genre, and describe the music contents automatically. Many researches have suggested the automatic music classification methods, especially music feature representation. Many different features can be used for music classification, e.g., content-based acoustic features including tonality, pitch, and beat.

The content-based acoustic features are classified into timbral features, rhythmic content features, and pitch content features. Timbral features are mostly originated from traditional speech recognition techniques. Typical timbral features include Spectral Centroid, Spectral Rolloff, Spectral Flux, Energy, Zero Crossing, Mel-Frequency Cepstral Coefficients (MFCCs) [9]. Rhythmic content features constrain information about the regularity of the rhythm, the beat and tempo information [9]. Pitch content features deals with the frequency information of the music bands and are obtained using various pitch detection techniques.

3. OVERALL PROCESS

We design to deal with the problems of the CF system's input data representation and a new music recommendation on the mobile web. The system follows a general CF process, and is added a new music recommendation module.

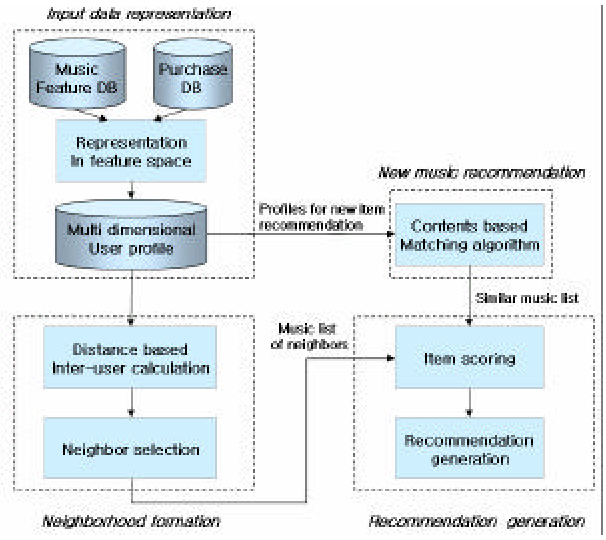

Figure 2. The overall process

The input data representation module creates a user profile using the music feature database and purchase database. The Music feature database is organized with the features like timbral, MFCC, rhythmic, and pitch features, which are extracted by the MARSYAS system [12]. The Purchase database is organized with the individual user's purchasing or pre-listening histories. Using this module, the system reflects the individual user's preference, can represent each user as a point in the multi-dimensional feature space.

The neighborhood formation module selects the target user's neighbors. The existing CF systems calculate the inter-user correlation by the Cosine formulation or Pearson Coefficient; however this way is hard to identify the neighbors who are really close to the target user as described Section 2.1. Therefore this module calculates the distance between the target user and other users to identify his real neighbors, and forms his neighborhood. Finally it gathers neighbors' purchasing or pre-listening musics, and gives the list of these musics to the recommendation generation module.

The recommendation generation module creates a recommendation list for the target user. This module scores the musics which neighbors purchased or pre-listened. As many neighbors purchase or pre-listen, this score increase.

The new music recommendation module solves the new item ramp-up. This module finds the most similar music by the content-based matching algorithm, and generates a similar music list, and gives it to the recommendation generation module.

4. RECOMMENDATION PROCEDURE

4.1 User Profile Creation

A user profile, which is the key component for CF system, includes information about user's preferences on music. This module is organized by two phases, the first phase is the extraction music's features, and second phase is the construction of user's preferred music set.

Phase 1: The extraction music's features

For extracting the music's features, we use the described features in section 2.2. For getting these features, we use the MARSYAS system [10]. This system extracts the timbral, MFCCs, rhythmic, and pitch content features. The timbral features consist of average spectral centroid, average spectral rolloff, average spectral flux, energy, and zerocrossing. MFCC features consist of the mean or each of the first five MFCC coefficients over the frames. The rhythmic features consist of sex features from the rhythm histogram. Pitch features are five features from the pitch histograms.

The system extracts the whole music's features, represents the individual musics in the multi-dimensional feature space as a point.

Phase 2: The construction of user's preferred music set

The general CF system uses the user-item binary matrix, but we use the user's preferred music set as the user profile. In the mobile environment users could do actions like purchasing, pre-listening or giving up buying a music. We set a rating about user's action assuming that these actions reflect user's preference about the music. Therefore we define a rating about user's action.

If user 'a' has purchased music 'x', it could be given a rating 'R_a^x'. We define the 'R_a^x' as the following:

$$R_a^x = \begin{cases} 1.0 & \text{If user 'a' has purchased music x} \\ 0.5 & \text{If user 'a' has pre-listened music x} \\ 0 & \text{If user 'a' has done any action to music x} \end{cases} \quad (1)$$

With the music which user 'a' has done any actions to and its ratings, the system composes the individual user's preferred music set. We defined user a's preferred music set as the following:

$$\textit{Preferred music set} = \{(x_1,\ R_a^{x_1}),\ldots,(x_{Na},\ R_a^{x_{Na}})\} \quad (2)$$

N_a is the number of the music whose R_a^x is non zero, and x_i is the music which user 'a' purchased or pre-listened.

The illustration of the preferred music set is showed in Figure 3.

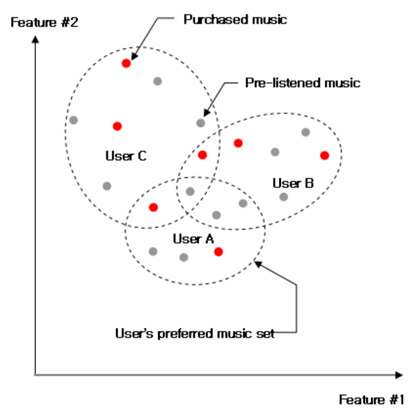

Figure 3. User's preferred music set

In Figure 3, user 'A', 'B', and 'C' are represented in the 2-dimensional feature space. The system represents all of the musics in this space, and binds the purchased and pre-listened musics as user's preferred music set by a dotted line. As shown in Figure 3, the user A's preferred music set includes 7 musics, and consists of 2 purchased musics and 5 pre-listened musics, which reflect his preference. User B and C's preferred music set also are composed by this way.

As doing this way, the user profile is composed by the whole users' preferred music sets.

4.2 Neighborhood Formation

Because whole users and musics are represented in multi-dimensional feature space, the system can find the target user's neighbors with distance based inter-user similarity calculation. If a user is close to the target user in this space, he could have a similar preference the target user's one. To calculate the distance between users, we assume that a user's preferred music set is a cluster.

There are many inter-cluster distance function, and the favorite function is the Centroid Euclidean distance function. This function is very simple and easy to calculate, but it works well in a condition that each cluster is fairly distributed and its form is a circle. However the individual users' preference isn't same and their distributions are difference from each other in shown Figure 3. Therefore we use the Average

Iinter-Cluster distance function [3] which the distance function considering the variance of two clusters, and we modify this function a little as the following:

$$d(a,b) = \frac{\sum_{y=1}^{N_a} \sum_{x=1}^{N_b} \sqrt{\frac{\sum_{i=1}^{N} (R_a^x \cdot F_x^i - R_b^y \cdot F_y^i)^2}{N}}}{N_a \cdot N_b} \quad (3)$$

Where *d(a, b)* is the distance function between user a and b, *N* is the number of features. N_a and N_b is the number of user a's and user b's preferred musics, respectively. R_a^x and R_b^y is the rating about music x by user 'a' and about music y by user 'b', respectively. F_x^i and F_y^i is the value of ith feature of music x and music 'y', respectively.

We sort the neighbors ascendingly according to the distance value, and then select the *L* neighbors. Finally we determine the neighborhood set *H* for target user *c*, and $H = \{h_1, h_2, \ldots, h_L\}$ such that $c \notin H$.

4.3 Recommendation Generation

The proposed system generates a list of n musics, $R = \{r_1, r_2, ..., r_n\}$ such that $r_i \notin$ {the music that the target user c has already purchased}. *PLS(c,x)* denotes the Purchase Likeliness Score of the target customer *c* on music *x*, and is computed as the following:

$$PLS(c,x) = \frac{\sum_{a \in H} (R_a^x - \overline{R_a}) \times sim(c,a)}{\sum_{a \in H} sim(c,a)} \quad (4)$$

Where $\overline{R_a}$ is the user a's average rating, *u*, *w* are the other users included the neighborhood set H. Because we use the distance function, we have to covert these distance values into similarity values, and normalize them. $sim(c,a)$ function denotes the similarity between target user c and user a, and it is computed as the following:

$$sim(c,a) = \frac{\underset{u,w \in H}{Max}\,[\,d(u,w)\,] - d(c,a)}{\underset{u,w \in H}{Max}\,[\,d(u,w)\,] - \underset{u,w \in H}{Min}\,[\,d(u,w)\,]} \quad (5)$$

4.4 New Music Recommendation

In mobile web environment, new musics are frequently provided and their purchasing ratio is very high, but the existing CF system can't recommend new musics. The general CF systems use the common items to find neighbors. Therefore they can't recommend musics that anyone didn't have been purchased. New musics can't be purchased by anyone, therefore they don't have ratings. If a new music has a rating, it could be recommended. To solve this problem we give a virtual rating to the new music.

Phase 1: Determination of new musics recommendation

New music also has the above mentioned features, which can be extracted. The proposed system extract new musics' features by the same way described in Section 4.1, and represent the musics in multi-dimensional feature space. Although new music can't have any ratings, it could be included in any user's preferred music set. Therefore we define a condition to determine new music's recommendation, as the following:

$$\begin{array}{c} \textit{If } x_{new} \in \textit{ a's preferred music set} \\ \textit{then } x_{new} \textit{ is recommended} \end{array} \quad (6)$$

Where x_{new} is a new music, $a \in H$, $H = \{c, h_1, h_2, \ldots, h_L\}$, and preferred music set is defined in (2). Why target user c is included in the neighborhood set H is that the new item doesn't purchased by target user, and the system have to reflect the target user's preference as much as possible.

The target user's preferred music set reflects his preference, therefore a new music comes in the target user's preferred music set or his neighbor's one, it would be preferred by them. The illustration of this is showed in Figure 4.

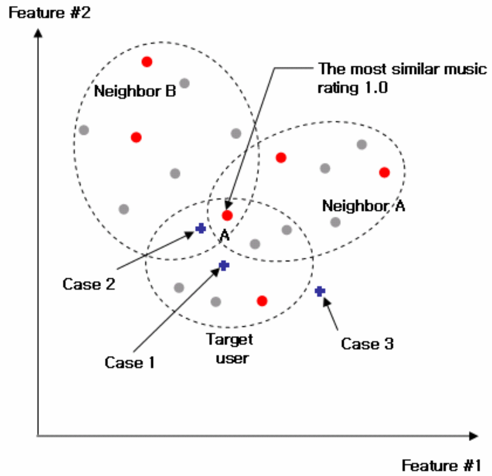

Figure 4. New music recommendation

Case 1, 2, and 3 are new musics and represented in the feature space. As shown Figure 4, case 1 and 2 can be recommended because they are satisfied the condition, but case 3 isn't. Therefore the system determines to recommend the case 1 and 2.

Phase 2: Finding the most similar music

We determine whether to recommend the new music or not, but it still doesn't have a rating. To give a virtual rating to the new music, the system searches for the most similar music where it is included. If music x and y in the same user's preferred music set are very similar, they would be given the very similar rating by that user. Most similar music is founded as the follows:

$$\min_{y \in a}\{d(x,y)\} = \sqrt{\frac{\sum_{i=1}^{N} (x_{new}^i - y^i)^2}{N}} \quad (7)$$

Where x_{new}^i and y^i is the ith feature value of new music x_{new} and music y included user a's preferred music set, respectively.

The illustration of this is showed in Figure 4. As shown in phase 1, the case 1 and 2 can be recommended, but we see the case 2. The case 2's most similar music is the music A, and its rating value is 1.0. Therefore the case 2 is given the rating 1.0.

Through this way new music can have a rating, with this rating the system calculates new item's PLS value. Therefore we generate a recommendation list with combination of old musics and new musics.

5. EXPERIMENT EVALUATION

5.1 Experiment Design

For the purpose of the performance evaluation of the proposed system, we developed a simulation program imitating the real user's decision-making and context in conducted. And we use two databases for our experiment. The first dataset contains features of 500 songs which are the most popular songs in mobile environment. The second dataset contains 1000 users' utility. We assume that the individual user's preference is different from each other, and they have their own utilities. Therefore we create utility functions for 1000 users.

Because the CF system has to start with a few transactions, we make the initial transactions. This is conducted by user's utility, and transactions organized by three purchase and five pre-listened transactions. With these transactions, we conduct the initial experiment.

For the new music recommendation, we offer the 100 new musics to the system, conduct the second experiment. We give the new music with no ratings to the system, evaluate the system's performance.

5.2 Evaluation

To evaluate our system, we use two matrix widely used in the information retrieval community namely recall and precision. Each matrix is defined as the following:

$$\text{Recall} = \frac{\text{Purchased musics}}{\text{all recommended musics}}$$

$$\text{Precision} = \frac{\text{Recommended musics}}{\text{all purchased musics}}$$

However these tow measures is often conflicting in nature. Then we user the standard F1 metric that gives equal weight to them both and is computed as the following:

$$\text{F1} = \frac{2 \cdot \text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}$$

We compute F1 for each individual user and calculate the average value to user as our matrix.

6. CONCLUSION

Because of an information overflow for multimedia contents like music, character images, and movie, mobile users need recommendation systems. Especially, music is the most popular contents in the mobile contents market, but there are few music recommender systems to aid mobile users to search musics which they really want. The CF system is known to be useful, but it is inadequate for the multimedia contents. Moreover CF system can't recommend the new item which is very frequently supplied and most contents providers want to sell out. We proposed the extended CF system which can solve these problems in the feature space and verify our system's performance.

REFERENCES

[1] Burke R., "Hybrid recommender systems: Survey and experiments", *User Modeling and User-Adapted Interaction*, 2002.
[2] Cho Y.H., J.K. Kim, "Application of web usage mining and product taxonomy to collaborative recommendation in e-commerce", *Expert systems with application*, 2004.
[3] Everitt B.S., S. Landau, M. Leese, "*Cluster Analysis", Oxford University Press*, 2001
[4] Goto M., Y. Muraoka, "A beat tracking system for acoustic signals of music", *ACM mulmimedia*, pages 365-372, 1994.
[5] 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.
[6] Kim J.K, Y.H. Cho, "Using Web Usage Mining and SVD to Improve E-commerce Recommendation Quality ", *LNCS (Lecture Notes Computer Science): Intelligent Agents and Multi-Agent Systems*, pp 86-97, 2003.
[7] Korea IT industry Promotion Agency, "*The Digital Contents Market Report*", 2003.
[8] Pazzani M.J., "A framework for collaborative, Content-Based and Demographic Filtering", *Artificial Intelligence Review*, 1999.
[9] Rabiner L., B.H. Juang, " *Fundamentals of Speech recognition*", *Prentice-Hall*, NJ, 1993.
[10] Sarwar B.M., G. Karypis, J.A. Konstan, J.T. Riedl, "Application of dimensionality reduction in recommender system - a case study", *ACM*, 2001.
[11] 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* , 2000.
[12] Tzanetakis G., "MARSYAS: A framework for audio analysis", *Organized sound*, Vol.4(3), 2000.