Personalized Music Recommendation Algorithm Based On Hybrid Collaborative Filtering Technology

Wang Wenzhen
Horqin Vocational College of Art;Tongliao 028000, China

***Abstract*—With the continuous growth of music resources, the problem of recommending suitable music for users has become a research hotspot. In this paper, association rules and music genes are added to music collaborative filtering personalized recommendation system to establish a hybrid recommendation model. The structure of the model is described and the recommendation process and recommendation algorithm of personalized recommendation are described in detail. By analyzing users' interests and preferences for different music gene features, the algorithm comprehensively analyses users' behavior, and uses the similarity of interests among different users to construct the neighborhood relationship among them. The recommendation algorithm is validated by combining two factors, and the expected recommendation results are achieved.**

Keywords-music recommendation; collaborative filtering; user interests; gene

I. INTRODUCTION

Traditional music retrieval basically uses metadata to retrieve eligible music. Such retrieval method requires users to remember the relevant information of the target track for retrieval, and it is no longer suitable for users' fast life rhythm. In modern life, what users need is to be able to continuously play music that meet their interests. The emergence of music recommendation system enables users to quickly and continuously obtain music that meets their requirements. However, with the growing scale of business websites, people are increasingly desired for recommendation services. Traditional collaborative filtering technology is also helpless in the face of data sparsity and algorithm complexity, which leads to the decline of recommendation quality. How to provide different users with a list of songs t, or to recommend songs that are of interest to them, has become a problem that needs to be solved in the current music recommendation system.

In this paper, the characteristics of music and the advantages and disadvantages of related recommendation algorithms are analyzed comprehensively and meticulously, and a hybrid recommendation model based on collaborative filtering and music gene is proposed. In the recommendation model based on collaborative filtering, user item matrix is constructed first, and the user similarity analysis is performed to find the nearest neighbor users and get a recommendation list. In the recommendation model based on music gene, by analyzing the structure of music gene, several user preferences are selected and calculated to generate a recommendation list. In the mixed recommendation mode, a final recommendation list is obtained by weight of the above two. Finally, a personalized music recommendation system is designed and implemented, and the system is tested by Web music website based on B/S architecture. By evaluating the recommendation results, we find that the overall performance of the system is good, and it can effectively recommend the results that users may be interested in to a certain extent, which basically achieve the expected goal of the system design.

II. RECOMMENDATION SYSTEM AND RELATED ALGORITHMS

A. Recommendation System

Recommendation system is a kind of information service technology which is mainly to solve the problem of choosing difficult caused by massive information. It can help users get the information and data they want in an effective time. The recommendation system used in business is to imitate the mode that people recommend items in real life. In the actual commodity transaction, the guides can provide certain similar recommendation according to the customer's experience used before. They can also recommend according to the customer's identity background information. Certainly, they can also provide some recommendation according to the customer's neighbors' interests.

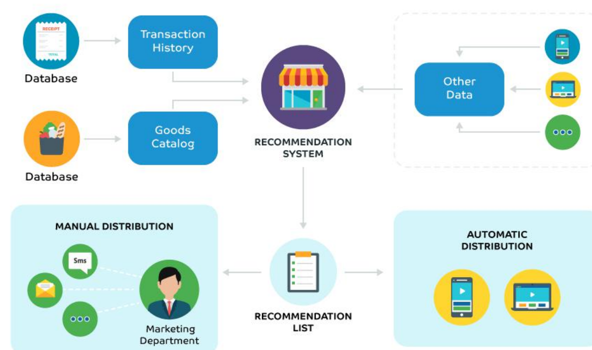


Figure 1. Working principle of recommendation system

Figure 1 depicts how the recommendation system works. As a new intelligent information mode, personalized recommendation technology can recommend the information that customers are interested in very accurately according to the requirements of customers or through the analysis of users' interests. At present, there are many recommendation technologies used, among which collaborative filtering technology is the earliest and most mature technology.

B. Collaborative Filtering Algorithm

Collaborative filtering technology is the most widely studied and applied e-commerce recommendation

978-1-7281-4463-4/19/$31.00 ©2019 IEEE
DOI 10.1109/ICSGEA.2019.00071

technology. The idea of collaborative filtering algorithm is that users can be divided into multiple user groups according to their interests. Users in each user group have similar interests, so they can recommend to the target users based on the preferences of other users in the group for a particular item or content. Collaborative filtering algorithm analyses the interest of target users, compares the behavior preference information of target users with other users' preference information, finds out the most similar neighbor user set with target users, and predicts the interest or preference of target users according to the interest or preference of neighbor user set. To achieve collaborative filtering, the following procedures are needed:

(1) Collect the preferences of different users;

(2) Find similar users or user groups;

(3) Recommendation.

The basis of collaborative filtering is composed of the above points. From this we can see that the core problem of collaborative filtering algorithm research is to confirm the similarity of users and to recommend target information according to the similarity.

III. Hybrid Recommendation Algorithm Based on Collaborative Filtering and Music Gene

A. *Hybrid Recommendation Principle*

The application scope of recommendation based on collaborative filtering is wider than other recommendation, and the accuracy of recommendation results is relatively high, but there are also multiple problems such as sparsity and cold start. Content-based recommendation does not rely on user evaluation information, and there are no shortcomings of sparsity and cold start of collaborative filtering, but the application scope is relatively smaller.

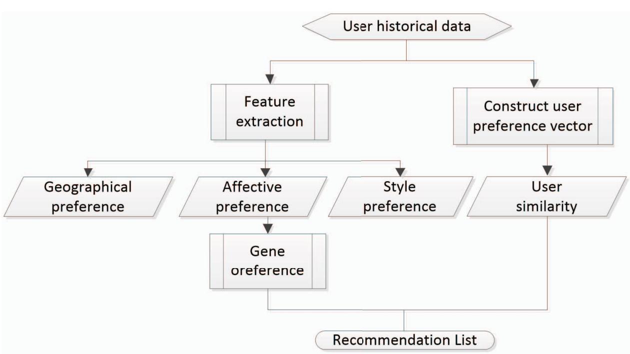


Figure 2. The detailed flow chart of mixed recommendation

For the shortcomings of these recommendations, the recommendation based on hybrid model can adopt different weight strategies in different situations. When the user's evaluation information is less, the weight based on music gene recommendation is increased; when the user's evaluation information is more, the weight based on collaborative filtering recommendation is increased. It makes full use of the advantages of collaborative filtering and music gene recommendation. The overall process of mixed recommendation is shown as figure 2.

B. *Collaborative Filtering Based on SVD*

Singular value decomposition(SVD) is a matrix decomposition technique, which can effectively extract the characteristics of algebra and deeply reveal the internal structure of matrix. It is applied to collaborative filtering algorithm, which can reduce the dimension of item space. The noise of the decomposed low-dimensional orthogonal matrix is lower than that of the original matrix. At the same time, users can score each item in the reduced dimension item space, and reveal the potential association between users and items more effectively. Singular value decomposition decomposes a matrix of $m \times n$ into three matrices T, S, D. T is a $m \times m$ orthogonal matrix , D is a $n \times n$ orthogonal matrix, S is a $m \times n$ diagonal matrix and the elements on the principal diagonal satisfies $\sigma_1 \geq \sigma_1 \geq ... \geq \sigma_n \geq 0$ (if $m \geq n$).

The elements on the S principal diagonal are singular value, which can be expressed as the approximation of a given matrix to a matrix with a lower rank. Generally, for matrices R= T × S × DT, where T, S, D must be full rank. The singular value decomposition has one advantage: It allows the existence of a simplified approximate matrix corresponding to S, preserving the largest K maximum singular values. For a new diagonal matrix S_k , T and D can be converted to T_k and D_k. Then the new matrix $R_k = T_k \times S_k \times D^T_k$ ($R_k \approx R$) can be generated. Singular value decomposition can create the all the ranks which equals to k and are similar to R . The predicted comments of user on project i is:

$$P_{a,j} = \overline{R}_a + T_k \times \sqrt{S^T_k}(a) \times \sqrt{S_k} D^T_k(i) \tag{1}$$

where $\overline{R}_a$ is the evaluation score on all the scoring items.

C. *Recommendation Model Based on Music Gene*

By analyzing and summarizing the various attributes of the song, we can get the music genome composition as shown in figure 3.

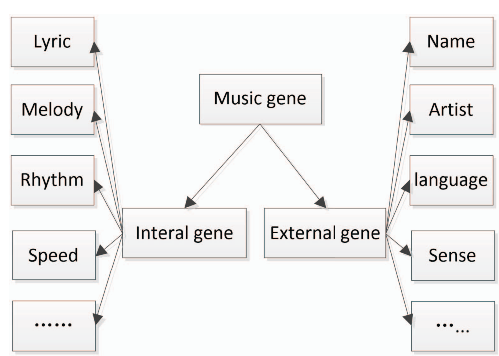


Figure 3. Music gene structure

According to the external gene information of music genome, user preferences for each feature are analyzed separately, and user preference vectors are established. First, define the following concepts:

(1) Singer loyalty degree

Singer loyalty degree is used to describe the user's preference for a singer. There are users who are "fans" of one or some singers who are more interested in their songs. Therefore, first of all, the singer table is established, and for each user, the loyalty $SLD(i,j)$ of users to different singers is calculated according to the information of the song tagging like the following equation

$$SLD(i,j) = \frac{songCount(i,j)}{sumSongCount(i)} \tag{2}$$

where $SLD(i,j)$ denotes the loyalty of user i to singer j; $sumSongCount(i)$ denotes all the music concerned by user; $songCount(i,j)$ denotes the total number of all the music of singer j, to acquire the loyalty list of user.

(2) Song area preference

From the user's attention to music, we can find that there are some users who are fond of music in foreign languages and not interested in Chinese music. Similarly, certain users do not like foreign music, but they are keen on Chinese songs. Therefore, according to the music website, there are four kinds of songs: Putonghua, Cantonese, Japan, Korea and Europe and America. The user preference is calculated for each type of user as

$$SAP(i,j) = \frac{songArea(i,j)}{sumSongCount(i)} \tag{3}$$

where $SAP(i,j)$ denotes the preference on music of user I to area j; $songArea(i,j)$ denotes the number of music belongs to area j, among the music concerned by user.

(3) User preference for free genes

For music, the fixed gene is fixed and almost unique for music. Therefore, users' preferences for different music can be calculated according to their attention behavior. However, for the free gene of music, such aspect of music is uncertain and based on the user's own understanding. Even for the same music, different users will have different understandings. At the same time, for the same understanding, unused users may also use different labels to express. Therefore, for the whole free radical factor, users' preferences for different music features can be obtained according to the user's annotation information. The preference can reflect the characteristics of music that users pay more attention to when looking for music. Therefore, users' preference tables for different characteristics of music free radicals can be established and the calculation equations are shown as follows:

$$TP(i,j) = \frac{songTag(i,j)}{sumTagCount(i)} \tag{4}$$

where $TP(i,j)$ denotes the preference of user i to music feature j; $songTag(i,j)$ denotes the total number of marked lags of user belongs to feature j; $sumTagCount(i)$ denotes the total number of all the tags marked by the user.

IV. EXPERIMENTAL ANALYSIS

The homepage recommendation page of the test system is divided into hot singer and hot song recommendation. When visitors login, the right side shows hot songs recommendation, and the home page shows top 10 songs that play the most times in the week. If logged in, the song recommendation based on collaborative filtering is displayed on the right.

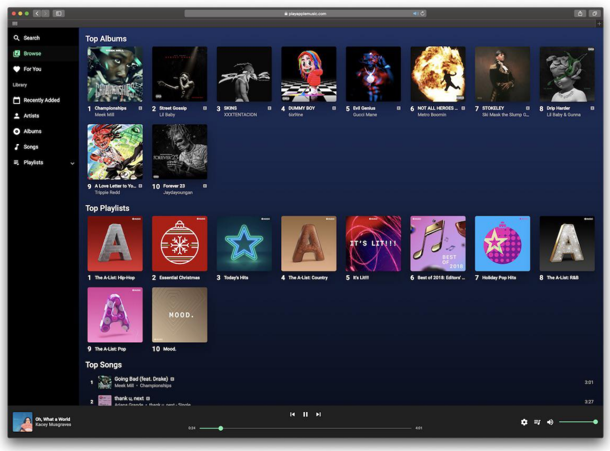


Figure 4. The music searching and recommendation interface

The first experiment set w as all the songs that users pay attention to, and the second set v as the songs that are not paid attention to by users. Then, eight pieces of music are randomly extracted from w as the corpus for user interest analysis. After the tagging information of the user is removed from the eight music pieces, the tagging information is put into w, and the set v is used as the test corpus for the experiment. According to the statistics of the corpus set corresponding to these 10 users, the average number of songs in the set w is 878, and the average number of songs in v is 1974. To test the effectiveness of the system, we invite 20 students to evaluate the recommended results of the system and the recommendation effect of the system is evaluated.

To adapt to the smaller display screen of mobile devices such as mobile phones, 10 music with the highest score are selected as the recommended results in the experiment, and the system is evaluated according to the situation of eight music in the final results. The first N results of the returned results are evaluated and the correct results are tagged. According to the user's annotation, the evaluation system counts the correct results of N results before each recommendation, and takes the proportion of listening as the evaluation index of system performance. The experimental results are shown in figure 5.

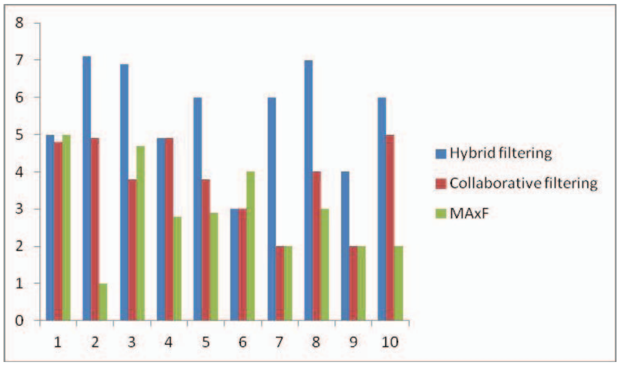
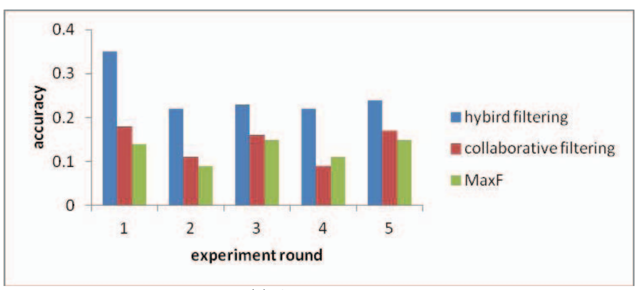
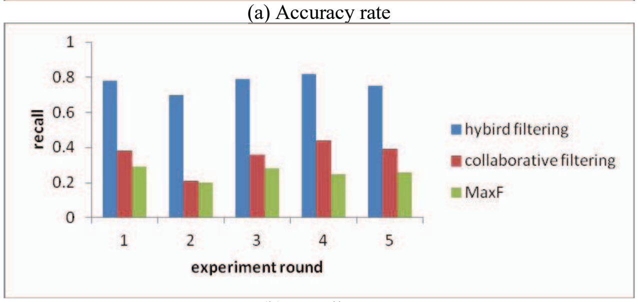


Figure 5. The emergence of 10 music tracks in the first experiment

We randomly grabbed 800 pieces of music from the experiment by web crawler, and divided the data set into training set and testing set according to the ratio of 7:3. We used 40 users to conduct a 30-day cycle recommendation training, and randomly selected 10 users as the target users of recommendation. The recommendation effects of the hybrid recommendation model and the single collaborative recommendation technology are evaluated respectively. Precision and recall are used to evaluate the recommendation effect of the algorithm. Accuracy is defined as the ratio of the user's favorite music items to all the recommended music items. The range of accuracy and recall are in [0,1]. The average accuracy and recall of 10 test users are obtained and the test results are shown in figure 6. The experimental results show that the hybrid recommendation algorithm has higher accuracy and recall rate, which achieves good recommendation results.



(a) Accuracy rate



(b) Recall rate

Figure 6. The result comparison of three algorithms in the second experiment

V. CONCLUSIONS

This paper studies the commonly used algorithms in music recommendation system: collaborative filtering recommendation algorithm based on users and items, and studies the problem of improving the quality of music recommendation. Since music itself aims to express human emotions, music recommendation should also reflect emotional factors. We also study the establishment of music emotion model as an implicit score of recommendation algorithm. The research is helpful to improve the practicability of music recommendation system, and the relevant results also play a reference role in the application of music recommendation system in the field of emotional regulation and psychotherapy.

REFERENCES

[1] Jiao C B. Hybrid Collaborative Filtering Recommendation Algorithm Based on Model Filling. Journal of Software, 2007, 18(7):1685-1694

[2] Zhang C, Ming Y, Jing L, et al. An Improved Hybrid Collaborative Filtering Algorithm Based on Tags and Time Factor. Big Data Mining and Analytics, 2018, 1(2):128-136

[3] Pessemier T D, Dhondt J, Martens L. Hybrid group recommendations for a travel service. Multimedia Tools & Applications, 2016, 76(2):1-25

[4] Su J H, Yeh H H, Yu P S, et al. Music Recommendation Using Content and Context Information Mining. IEEE Intelligent Systems, 2010, 25(1):16-26

[5] Chen Y, Min Y. A Hybrid Collaborative Filtering Algorithm Based on User-Item. Wuhan University Journal of Natural Sciences, 2016, 16(1):16-20

[6] Huang S, Ma J, Cheng P, et al. A Hybrid Multigroup Coclustering Recommendation Framework Based on Information Fusion. Acm Transactions on Intelligent Systems & Technology, 2015, 6(2):1-22

[7] Mohbey K K, Thakur G S. Interesting User Behaviour Prediction in Mobile E-commerce Environment using Constraints. Iete Technical Review, 2015, 32(1):16-28

[8] Pedro J S, Siersdorfer S, Sanderson M. Content redundancy in YouTube and its application to video tagging. Acm Transactions on Information Systems, 2011, 29(3):1-31

[9] Ha I, Oh K J, Jo G S. Personalized advertisement system using social relationship based user modeling. Multimedia Tools & Applications, 2015, 74(20):8801-8819

[10] Wachs J P, Kölsch M, Stern H, et al. Vision-based hand-gesture applications. Communications of the Acm, 2011, 54(2):60-71

[11] Hu Y, Shi W, Li H, et al. Mitigating Data Sparsity Using Similarity Reinforcement-Enhanced Collaborative Filtering. Acm Transactions on Internet Technology, 2017, 17(3):1-20

[12] Liu Z, Qu W, Li H, et al. A hybrid collaborative filtering recommendation mechanism for P2P networks. Future Generation Computer Systems, 2010, 26(8):1409-1417

[13] Panigrahi S, Lenka R K, Stitipragyan A. A Hybrid Distributed Collaborative Filtering Recommender Engine Using Apache Spark. Procedia Computer Science, 2016, 83:1000-1006

[14] Cao J, Wu B, Peng J, et al. Copper-copper iodide hybrid nanostructure as hole transport material for efficient and stable inverted perovskite solar cells. Science China, 2019, 62(3):83-89