Design of music intelligent recommendation based on user interest preference model

Jianchao Lei[a], Chunqiu Wang*[,a], Xucheng Geng[b]
[a]College of Music and Dance, Huaihua University, Huaihua, China; [b] Finance Office of Huaihua University, Huaihua, China.

ABSTRACT

With the rapid development of the Internet, the number of digital music has increased dramatically. It is very difficult for users to dig out the songs they are interested in from the massive digital music information, and the problem of "information overload" has emerged. A search engine is a tool for retrieving information. When a user submits a query to the search engine, the relevant content is returned to the user after retrieval. The search results of different search engines are different. Through the analysis of users' behaviors, we can understand users' search habits and realize a search engine that is more in line with users' needs. In order to accurately recommend interesting music to users and improve the market share of application objects, this paper designs a music intelligent recommendation system based on user interest preference, which is composed of data acquisition module, offline data processing module and online recommendation module. The data acquisition module collects user registration information, online scoring data of target music, and conducts online survey to collect relevant data. Data analysis and experiments show that the model is superior to the comparison model in terms of training efficiency and accuracy of recommendation, thus verifying the effectiveness of integrating user's long-term and short-term preferences and user influence for recommendation.

Keywords: User interest; Preference model; Smart music recommendation

1. INTRODUCTION

With the rapid development of the Internet, the content that ISP can provide users has also increased dramatically. From the user's point of view, the increase of content has both advantages and disadvantages. The advantage is that users have more choices when obtaining information, and the disadvantage is that it increases the difficulty for users to obtain useful information, that is, it increases the difficulty of information filtering [1]. While users enjoy the convenience of life brought by massive data, they are also troubled by the increased difficulty in processing data. People need to spend more time and energy to find the information they are interested in, and they are easily interfered by other irrelevant information. The increase in information search cost makes users at a loss. There are many solutions to the problem of "information overload", such as navigation websites, search engines and recommendation systems [2]. In personalized recommendation, the analysis of user preferences is a key step, and the next step can be predicted according to user preferences. Most methods calculate the similarity according to the historical scoring data of users, and do not distinguish the importance of historical scoring at different times, which affects the accuracy of recommendation results. In addition, the information of user rating obtained by traditional recommendation methods is very small, and the rating matrix is very sparse. When new users arrive, they have no behavior data, so they can't predict their preferences according to their historical behavior, which will lead to the cold start problem [3].

Information on the Internet exists in many different forms, including text. Other forms of information can usually be expressed in text, such as movies, books and pictures. Without browsing all resources, to find the information you need, you often need to query these text information to achieve the purpose [4]. Generally speaking, most enterprises do not want to handle the internal information of the enterprise externally for the sake of security. In addition, this scheme has a defect that it is not flexible enough. The search results are completely controlled by a third-party search engine, and it is difficult to provide richer information [5]. The implementation of the third scheme is relatively difficult. The main reason is that there is a certain technical threshold for the search engine, and the development requires a relatively large workload and development costs, which are not borne by all enterprises [6].

*Email: 554342852@qq.com

Seventh International Conference on Mechatronics and Intelligent Robotics (ICMIR 2023), edited by Srikanta Patnaik, Tao Shen, Proc. of SPIE Vol. 12779, 127791O · © 2023 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.2689446

Facing the vast digital music, users will encounter the long tail effect in the process of obtaining music for a long time. Most users have heard some of the same popular songs, however, many songs have only been heard by minority people, so many songs have long-term single-digit hits [7]. At present, online music platforms such as Spotify and last.fm abroad, Netease Cloud Music and Douban FM in China have launched the service of filtering and predicting the potential favorite songs of users by analyzing their usage records, and selecting some song candidate sets that users are interested in from a huge music library, which greatly reduces the search time and audition time of users, thus playing a great role in the spread of many previously relatively unpopular songs and improving the freshness of users' favorite songs. At present, there are many methods used in music recommendation research, such as content-based recommendation, model-based recommendation and collaborative filtering method [8].

Content-based music recommendation generally considers the underlying music features such as the audio and rhythm of music for analysis and recommendation. However, there are significant semantic differences between the underlying music features and the user's understanding of music, and the recommendation effect is not good. For this purpose, an intelligent music recommendation system based on user interest preference is designed. On the basis of designing the overall architecture of the data acquisition module, offline data processing module and online recommendation module, taking into account the user's music interest preference and personalized needs, the collaborative filtering algorithm is introduced to determine the music recommendation preference using the number of nearest neighbors to achieve accurate music recommendation for users.

2. INTELLIGENT MUSIC RECOMMENDATION SYSTEM BASED ON USER INTEREST PREFERENCE

2.1 Overall structure design of the system

Recommendation system is an effective method to solve the problem of information overload. As a new intelligent information service mode, it has been widely used in video, music, social network relations, digital libraries and other fields. The process of recommending users mainly includes two stages: first, the preference acquisition stage, which calculates the user's preference vector according to the user's behavior data; The second is to generate recommendation results, and make recommendations through established user models and historical data of target users [9]. Compared with previous data sets used for recommendation research, YahooMusic data sets are larger in scale and have another important feature: there is hierarchical information between items. In the traditional data set used for recommendation field research, all items have a flat structure, but in YahooMusic data set, there is a hierarchical structure information between items. Specifically, items can be divided into four types: singer, album, song and genre [10]. In addition to music data, this hierarchical structure information can be obtained in many cases, for example, shopping websites like Amazon, goods and categories also have such a relationship. As we will analyze in detail later, making full use of this hierarchical information is of great help to improve the recommendation accuracy.

Generally speaking, users' listening to music is a continuous behavior, that is, they can listen to more than one music at a time, and the music they listen to in a certain period of time is relevant. Therefore, users' interests and preferences in the short term are regular. In addition, users' continuous listening behavior makes it possible to collect more user data in the short term, Therefore, it is feasible and necessary to recommend real-time music to users according to their short-term music listening preferences. Search engines give search results based on search keywords. Generally speaking, the same keywords get the same results, while the recommendation system goes further. It can find users' interests and preferences by studying the patterns or characteristics of users' historical behavior. After personalized calculation, the process of discovering users' interests and preferences will be automated. It can also find users' interests and then guide users, Let them find their own information needs. A good recommendation system should not only meet the user's demand for personalized recommendation, but also explore the user's unknown interest points. Even in some application scenarios, some users rely on the recommendation system. The music intelligent recommendation system designed in this paper is mainly composed of data acquisition module, offline data processing module and online recommendation module. The overall structure design is shown in Figure 1.

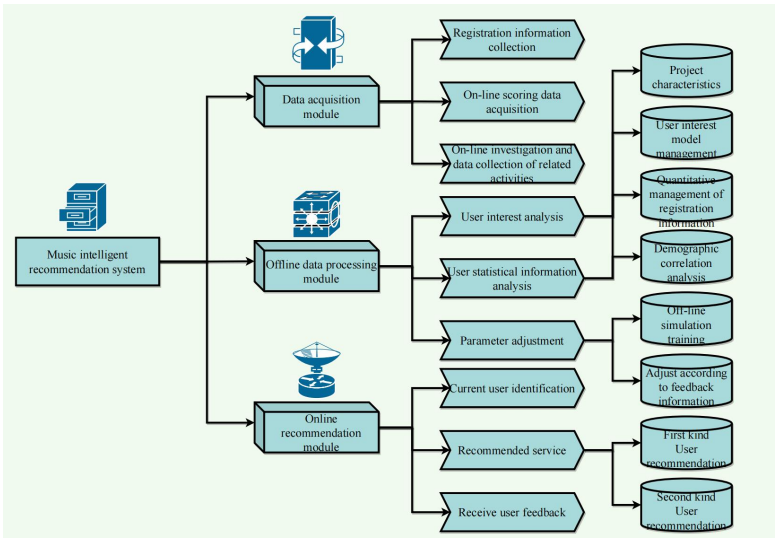

Figure 1. Overall structure of the system

The main function of data acquisition module is to collect user registration information and online rating data of target music, and conduct online investigation. The off-line data processing module is based on the data collected by the data acquisition module, analyzing the user's interest preference and counting the user's information. According to the analysis results of user interest preference, the online recommendation module recommends music that can meet the needs of different users.

2.2 Design of data acquisition module

For music, this recommendation item is greatly influenced by contextual factors, but the traditional music recommendation system seldom considers contextual factors. This chapter will propose a hybrid music recommendation algorithm based on LFM model and the current time situation of users, which will comprehensively analyze and predict the historical interaction between users and music platforms, mainly focusing on users' implicit feedback on music and introducing the stability of users' personal hidden interest preferences based on the current time period. It reduces the dependence of users on other historical behaviors, makes the recommended results more in line with the recommendations of the current mainstream music platforms, and pays attention to the internal relationship between the changes of users' hidden interests and preferences and time. Nowadays, the use of mobile music applications is very common, and almost all music applications will give obvious search entries, which shows the importance of search function. Through investigating several popular mobile music applications on the market, it is found that the basic functional requirements of music search engine are not many, mainly including: display of search hot words, intelligent prompt of search keywords, intelligent matching, song name/singer/album search, lyrics search, etc. By observing the user search log and access log of the music application, it can be found that users tend to enter fewer words when using the search function. For the long query part, only a few users will enter the long query keywords completely; When the search keyword entered by the user is not specific song, the user's search intention is vague, and the number of page turns is more than the purposeful search; The user's input is inaccurate, which is usually a minor error.

Usually, ETL tools extract data (such as relational data and flat data files) from heterogeneous data sources distributed in the business system to the temporary middle layer for cleaning, transformation, integration and finally loading into the data warehouse or data mart, which is the basis of online analytical processing and data mining, and provides a unified data basis for assisting decision-making, analysis and query. In this paper, the collaborative filtering method based on analyzing users' access behavior is used to design an intelligent recommendation system for music websites. The main method is to derive users' evaluation values of music from the data warehouse, so as to find a set of users with similar interests by clustering calculation, so that the recommendation engine can make directional recommendations to users. The running flow of the data acquisition module of this system is shown in Figure 2.

Figure 2. Operation process design of data acquisition module

The data acquisition module can obtain the data resources required for system operation when the system is in the initial startup state or when it is running offline independently. The key role of the data acquisition module is to efficiently obtain massive and effective data, record the data information in the appropriate format, and generate the corresponding files, so as to facilitate the subsequent clustering processing and database storage.

3. INTELLIGENT MUSIC RECOMMENDATION DESIGN BASED ON USER INTEREST PREFERENCE MODEL

3.1 User interest preference calculation model

Based on the user's historical music search information, a time-based user interest preference calculation method is proposed. For the music recommendation system, the key point of the user's needs is the accuracy of the music recommended by the system. For the user, the user's recognition of the recommendation effect of the music recommendation system is the most important indicator. In the traditional LFM algorithm, the selection of recommended items is based on the prediction score of all items. Therefore, for LFM algorithm, the most important goal is to make the predicted score of the project to be predicted equal to or close to the actual score. From the aspects of improving the complexity of the algorithm, solving the sparsity of raw data, and adding bias items, starting from various factors not considered in the recommendation process, we can improve the accuracy of the recommendation algorithm and solve the problems of cold start by introducing new and reasonable calculation factors into the formula, or reconstructing a more reasonable calculation formula. Obviously, in the case that users can't establish labels for their new interests in time, explicit information is not enough to express complete user interests, and implicit information is based on users' current behavior and lacks knowledge of users' long-term interests. Using the combination of the two methods to obtain users' interest characteristics can effectively make up for the defect that the characteristics of interest data obtained by a single method are not comprehensive, and then establish a complete and accurate user interest model. Therefore, this paper introduces collaborative filtering recommendation algorithm. Based on the analysis of the similarity of interests of multiple users, considering the personalization of users, the number of neighbors is used to determine the preference of music recommendation, and the accuracy of music recommendation for users is improved.

In the online recommendation module, collaborative filtering algorithm is adopted, which is based on like-minded viewpoints. The core viewpoints of the algorithm are as follows: if users with the same interest preference have searched for a certain music, users with the same interest preference have a greater probability of having obvious preference for a certain music; If users have an interest preference for a certain music, they will also be more inclined to choose music with a higher degree of consistency with the music. User-based recommendation method: given a user, rating data and music items. Some music items have not been scored. Under this condition, the scoring results of other users close to this user can

be used for prediction. The consistency function of interest preference between user A and user B can be described by Formula (1):

$$sim(a,b)=\frac{\sum_{p\in P}(r_{a,p}-\bar{\gamma}_a)(r_{b,p}-\bar{\gamma}_b)}{\sqrt{\sum_{p\in P}(r_{a,p}-\bar{\gamma}_a)}\sqrt{\sum_{p\in P}(r_{b,p}-\bar{\gamma}_b)}} \tag{1}$$

In formula (4), $\bar{\gamma}_a$ and $\bar{\gamma}_b$ represent the score of user a and user b respectively, and $r_{a,p}$ and $r_{b,p}$ represent the score of user a and user b on music p respectively.

Through the statistics of data, it is found that the number of items evaluated by active users and inactive users is not linear, and the number of items evaluated by highly active users can even reach several times of that of moderately active users. Therefore, in order to prevent excessive punishment of the contribution of highly active users, logarithm is used to smooth the punishment of users with different levels of activity. The penalty factor is calculated by Formula (2):

$$a_u^{user-CF}=\frac{1}{\log_{10}(1+|R_u|)} \tag{2}$$

Among them, $|R_u|$ represents the total number of items evaluated by the user u, and the more items evaluated, the more active the user is. When a penalty factor is introduced to suppress the contribution of active users in the prediction scoring stage, the prediction score based on collaborative filtering of users can be expressed by Formula (3):

$$P_{UAIB-User-CF}(u,i)=\sum_{v\in N_u\cap i\in I_v} sim(u,v)*R_{v,i}*a_v^{User-CF} \tag{3}$$

For project-based collaborative filtering, active users u will have a greater possibility of evaluating any two items L and W, so in the calculation of project similarity, the items evaluated by those extremely active users will likely become their neighbors. When the number of extremely active users is very small, the impact will be relatively small, and if there are a little more extremely active users, the impact will increase, so this situation needs to be avoided.

After obtaining the consistency of interest preferences among users, considering the differences of different users' thinking modes and scoring standards, the project-based recommendation method can be analyzed by cosine consistency, and the cosine consistency of music item m and music item n can be determined by Formula (4):

$$sim(m,n)=\frac{\sum_{u\in U}(r_{u,m}-\bar{\gamma}_u)(r_{u,n}-\bar{\gamma}_u)}{\sqrt{\sum_{u\in U}(r_{u,m}-\bar{\gamma}_u)^2}\sqrt{\sum_{p\in P}(r_{u,m}-\bar{\gamma}_u)^2}} \tag{4}$$

User-based recommendation method and project-based recommendation method both involve the process of selecting neighbors. Usually, both methods need to go through a lot of calculation processes, but only a small part of the data is taken, which causes the problem of data sparseness in collaborative recommendation algorithm and reduces the accuracy of recommendation results. Based on the above analysis, it can be seen that the online recommendation module of the system adopts collaborative filtering algorithm, and generates recommendation lists in different ways for different users based on user interest preference, so as to improve the recommendation performance of the system.

3.2 Analysis of experimental results

In the experiment of this chapter, firstly, the influence of experimental parameters in the proposed personalized music recommendation algorithm on each evaluation index is compared. Then, other recommendation algorithms are compared with other music recommendation algorithms to test the performance of RNDM method. The realization of system functions is the basis for the effective operation of the system, so the black box test method is used to test the functions of this system. The main test cases are: (1) User login successfully: enter the correct user name and password, and expect the output result to be successful login; (2) Unsuccessful user login: wrong user name and password are entered, and the output result is expected to be wrong user name or password. Please re-enter; (3) Generate recommendation results: the input is to click the recommendation list for different users under different conditions, and the expected output result is to correctly

predict the user's interest preference and recommend the corresponding music. In this paper, the user's interest model is established based on tags and user behavior, which only involves the acquisition of explicit information and implicit information. According to the modeling idea, the acquisition of explicit information is mainly reflected in the user's tags, while the acquisition of implicit information is mainly reflected in the user's browsing behavior. In addition, the role of user interest model in improving recommendation quality needs to be combined with recommendation algorithm to be well reflected.

In order to verify the advantages of RNDM method, we compare RNDM with the following three methods, including population-based method, matrixfactorization method and MEM method. As shown in Figure 3 and Figure 4.

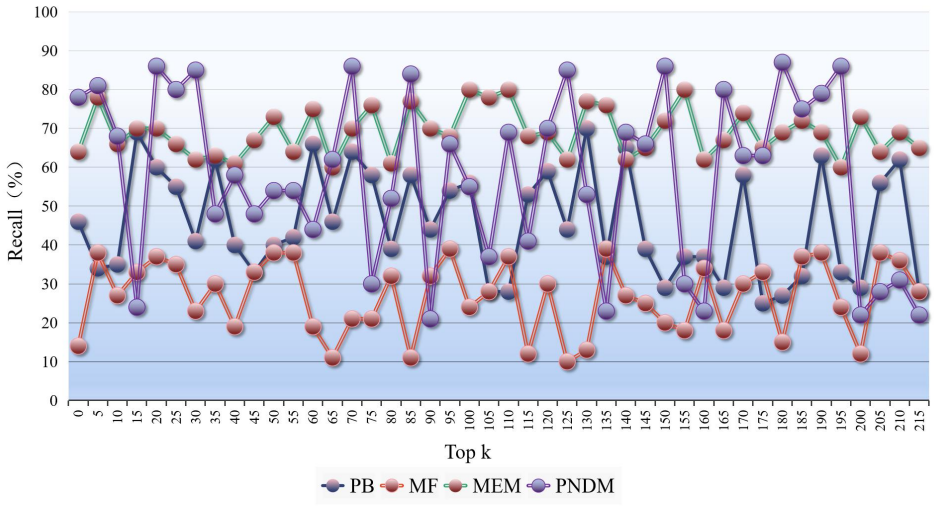

Figure 3. Comparison of four methods in recall rate

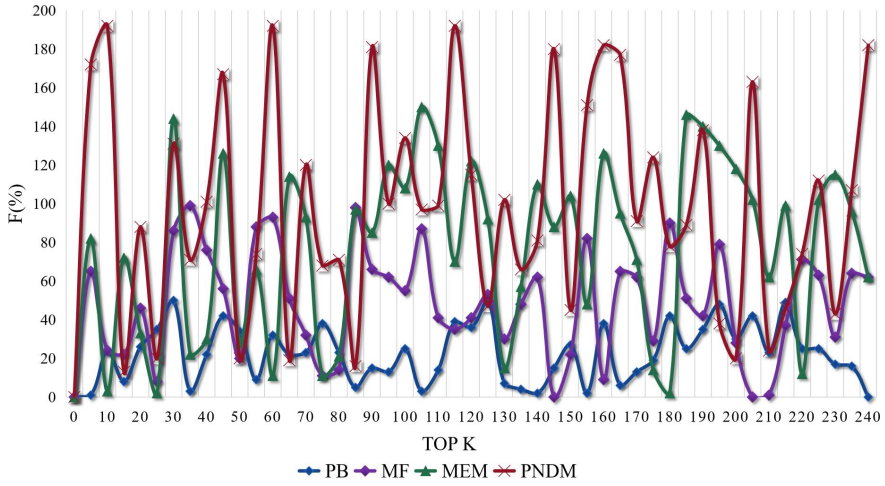

Figure 4. Comparison of four methods on F1

From the analysis of the above figure, it is found that the promotion of the number of neighbors has a direct impact on the recommendation results of this system. Accuracy and recall: The accuracy and recall of this system are positively correlated with the number of neighbors, that is, with the increase of the number of neighbors in the online recommendation module of this system, the accuracy and recall of the system recommendation results are on the rise as a whole. Compared with PB, MF and MEM, RNDM improves the performance by 67%,45.3% and 21.5% respectively. When the value of k is small, the most popular songs are recommended to the target users, and it can be observed that the performance of the method based on popularity is better. However, as the value of k increases gradually, the popularity index of songs decreases, which reduces the effect of music recommendation. In the context and context awareness music recommendation, MEM is not as good as RNDM, but it has better performance than other basic methods, because PB and

MF do not consider the user's preferences; Compared with MEM, we can see that this method is more effective for music recommendation, because it mines the characteristics of users in many aspects and combines them to make recommendations for users according to their comprehensive music preferences.

4. CONCLUSION

With the development of a large number of music streaming services, online music platforms such as YouTube and Netease Cloud Music are gradually emerging. Retrieving music that meets user preferences has become a very challenging task. Traditional search engine tools cannot accurately locate the music that users are interested in. Music plays an important role in people's leisure and entertainment. Internet music users now account for 71.1% of the total users. The scale of the music market is also growing rapidly. The music recommendation field also needs to apply personalized music recommendation technology. It can be predicted that music recommendation will produce greater value in the future. In this paper, an intelligent music recommendation system based on user interest preference is designed. Based on user data and music data, the user's interest preference for the target music type is determined, and the music that the user is interested in is recommended to the user by collaborative filtering recommendation algorithm. The application test results show that the system can accurately recommend interested music to users and improve the market share of application objects.

REFERENCES

[1] Ge, J., Shi, L. L., Liu, L., et al., Intelligent Link Prediction Management Based on Community Discovery and User Behavior Preference in Online Social Networks. Wireless Communications and Mobile Computing, vol.44, no.8, pp.2 (2021).

[2] Zhang, L., Optimization of an Intelligent Music-Playing System Based on Network Communication. Complexity, vol.47, no.9, p.10 (2021).

[3] Dai, D. D., Artificial Intelligence Technology Assisted Music Teaching Design. Scientific programming, vol.13, no.11, pp.24 (2021).

[4] Wu, Y., Lin, J., Ma, Y., A hybrid music recommendation model based on personalized measurement and game theory. Chinese Journal of Electronics, 32, 1-10 (2022).

[5] Wei, L., Design of a Music Recommendation Model on the Basis of Multilayer Attention Representation. Scientific Programming, vol.11, no.8, pp.22 (2022).

[6] Andjelkovic, I., Parra, D., O'Donovan, J., Moodplay: Interactive Music Recommendation based on Artists' Mood Similarity. International Journal of Human-Computer Studies, vol.12, no.1, pp.37 (2018).

[7] Elbir, A., Aydin, N., Music Genre Classification and Music Recommendation by Using Deep Learning. Electronics Letters, vol.56, no.10, pp.46 (2020).

[8] Xiang, D., Zhang, Z., Cross-Border E-Commerce Personalized Recommendation Based on Fuzzy Association Specifications Combined with Complex Preference Model. Mathematical Problems in Engineering, vol.4, no.1, pp.9 (2020).

[9] Zhang, L., Tian, Z., Research on the recommendation of aerobics music adaptation based on computer aided design software. Journal of Intelligent and Fuzzy Systems, vol.12, no.1, pp.2 (2021).

[10] Liang, Y., Fan, Y., Peng, Y., et al., Smart Grid Project Benefit Evaluation Based on a Hybrid Intelligent Model. Sustainability, vol.14, no.8, pp.45 (2022).