Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 3, pp. 1923–1931, DOI 10.12694/scpe.v25i3.2792

CONTEXT-AWARE MUSIC RECOMMENDATION ALGORITHM COMBINING CLASSIFICATION AND COLLABORATIVE FILTERING

XIAOLING WU* AND GUODONG SUN†

Abstract. As an effective solution to the problem of information overload, personalized recommendations have received widespread attention in the music field. A context-aware music recommendation algorithm combining classification and collaborative filtering is proposed based on user context information. Firstly, the similarity analysis of the user situation is carried out. A preliminary list of recommended songs is obtained by collaborative filtering. The machine learning method is used to classify music in different scenes to get the preferences of music types in different situations. Finally, the recommendation list obtained by collaborative filtering is combined with the music type preference obtained by the classification model and personalized music recommendations for users in different situations. This algorithm not only effectively reduces the complexity of the recommendation process Experiments show that the proposed algorithm can effectively improve the accuracy of users' music recommendations.

Key words: Music recommendation; Personalized recommendation; Situational awareness; Improved Collaborative Filtering

1. Introduction. Music is a good pastime in human life. With the progress of science and technology, the development of music resources has reached an unprecedented height. A personalized recommendation is an effective way to alleviate the problem of "information overload." It has received more and more attention and application in the music industry. Major music platforms can now carry out personalized music recommendations, such as Spotify, Pandora, Douban Music, NetEase Music, etc. Many platforms have gained a reputation for recommending songs more accurately. Currently, most music recommendations are based on the user's usage habits to tap into the user's long-term preferences. However, the environment often affects users' short-term song preferences in song recommendations. Scenarios include people's psychological state, behavior, external environment and many other aspects [1]. Focusing only on users and recommended products will impact recommendation results. Take the user's environment into account in personalized music recommendations. It can make individual recommendations to users based on their situation.

By effectively utilizing a large amount of information in the context, personalized recommendations more in line with users can be achieved to improve its accuracy and user experience. This is the focus of scholars at home and abroad [2]. This paper uses multiple class models based on the classical collaborative screening method. Complete a scene-oriented music recommendation method. The traditional collaborative filtering algorithm can create situational awareness by integrating the similarity calculation method of user situational information. The fusion classification model improves the performance of the recommendation system, effectively reduces the complexity of the recommendation process and improves the cold start problem. After implementing this rule, users can be provided with a personalized song recommendation according to the user's actual situation.

2. Recommendation method based on collaborative screening. In essence, the learning recommendation system is realized through machine learning. Through the collection of students' basic information and analysis of students' learning habits, learning behaviors and test results, the external and internal learning needs of students are identified [4]. Students actively seek knowledge in the knowledge base to meet their needs—a dynamic generation of learning paths to facilitate learners to complete learning better.

2.1. Overall system framework. Collaborative screening is one of the most widely used methods at present. The algorithm can be divided into user and item-collaborative filtering [3]. The "nearest neighbor" is

*College of Music, Handan University, Handan, Hebei, 056000, China
†Students' Affairs Division, Hebei University of Engineering, Handan, Hebei, 056000, China (Corresponding author, sunguodong@hebeu.edu.cn)

closest to the target object and recommends the goods with "nearest neighbor preference" to it. The latter is the recommendation of items with similar historical interests to the user to the target user.

2.2. User recommendation methods in collaborative screening. Firstly, a new method based on the user-entry scoring matrix is proposed. Predict the evaluation of an item. Make practical information recommendations to specific users. Standard methods of user similarity calculation include the Pearson correlation coefficient method, cosine similarity method and improved cosine similarity method. Pearson correlation analysis is one of the most widely used:

$$sim(u,u\prime) = \frac{\sqrt{\sum\limits_{i\in I(u,u\prime)} \left|(r_{(u,i)} - \bar{r}_{(u)})(r_{(u\prime,i)} - \bar{r}_{(u\prime)})(r_{(u,i)}+\bar{r}_{(u)})(r_{(u\prime,i)}+\bar{r}_{(u\prime)})\right|}}{\sqrt{\left|\sum\limits_{i\in I(u,u\prime)} (r_{(u,i)} - \bar{r}_{(u)})^2\right|}\sqrt{\left|\sum\limits_{i\in I(u,u\prime)} (r_{(u\prime,i)} - \bar{r}_{(u\prime)})^2\right|}} \tag{2.1}$$

$r_{(u,i)}, r_{(u\prime,i)}$ indicates the evaluation of users u and $u\prime$ on item i. $\bar{r}_{(u)}$ and $\bar{r}_{(u\prime)}$ represent the average of the scores of users u and $u\prime$ for all entries. The nearest neighbor set D is generated to the target user based on similarity. Then, based on D, the user's score of the item is converted into the K-nearest neighbor score prediction formula (2.2) to predict the user's preference. The result is used as the basis for recommending the target users.

$$p(\varepsilon,i) = \bar{r}_{(\varepsilon)} + \frac{\sum\limits_{\varepsilon_t\in S} [sim(\varepsilon,\varepsilon_t)(r_{(\varepsilon t,j)} - \bar{r}_{(\varepsilon,t)})]}{\sum\limits_{\varepsilon_t\in S} sim(\varepsilon,\varepsilon_t)} \tag{2.2}$$

2.3. Recommended methods for item classification in collaborative screening. First, a similar group of items is found based on the score and solution process in formula (2.3). Record as project set I. This method predicts the target items and the information suitable for its characteristics is recommended.

$$sim(i,j) = \frac{\sum\limits_{u\in S_{(i,j)}} \sqrt{(r_{(u,i)} - \bar{r}_{(i)})(r_{(u,j)} - \bar{r}_{(j)})(r_{(u,i)}+\bar{r}_{(i)})(r_{(u,j)}+\bar{r}_{(j)})}}{\sqrt{\sum\limits_{s\in S_{(i,j)}} (r_{(u,i)} - \bar{r}_{(i)})(r_{(u,i)}+\bar{r}_{(i)})(r_{(u,j)} - \bar{r}_{(j)})(r_{(u,j)}+\bar{r}_{(j)})}} \tag{2.3}$$

$\bar{r}_{(i)}, \bar{r}_{(j)}$ is the average of all users' scores for items i and j. $S_{(i,j)}$ is the set of users. Equation (4) is obtained from the predicted user score ε in the nearest neighbor group I for entry i. Then give the corresponding product recommendation according to the predicted results.

$$p(\varepsilon,i) = \bar{r}_{(i)} + \frac{\sum\limits_{j\in I_{(i,j)}} [sim(i,j)(r_{(\varepsilon,j)} - \bar{r}_{(j)})]}{\sum\limits_{j\in I_{(i,j)}} sim(i,j)} \tag{2.4}$$

$I_{(i,j)}$ represents the set most closely related to item i.

2.4. Some critical issues about collaborative filtering.

Data sparsity problem. The recommended method of collaborative filtering is based on the user-entry score matrix. The higher the score density, the more accurate the similarity calculation and the higher the accuracy of its recommendation. However, due to the extreme sparsity of a large number of user-item evaluation samples, the efficiency of the recommendation algorithm is greatly affected. Literature [4] proposes an algorithm based on matrix and SVD to solve the problem of performance degradation caused by sparse data. The aim is to solve the sparse problem of the score matrix. Literature [5] adopted the clustering method to reduce the dimensionality of the score matrix to solve the problem of sparse data and improve the accuracy of recommendations.

Research on the recommended cold start problem. The cold start problem is when a new user or project is added to the recommendation system; because there is not enough object scoring data, the similarity of the target object will be difficult, failing to recommend the project to the new user. This problem will cause the user to lose confidence in the recommendation mechanism and thus be rejected by the user. At this point, the recommendation mechanism fails. The literature on the cold start phenomenon [6] proposes an algorithm based on n-sequence access analysis logic and most frequent item extraction based on the cold start problem of cooperative filtering. This eliminates the shortcomings of the user in the cold startup process. Literature [7] proposes a collaborative filtering algorithm based on a K-nearest neighbor-based attribute-feature graph. Literature [8] proposes a recommendation algorithm based on association rules, demonstrating its advantages in dealing with sparse data and cold start problems. Literature [9] designed the CUTA Time recommendation algorithm to solve the cold start problem of new projects based on user (project) attributes and rating date.

Real-time research of recommendation. In a recommendation system, the amount of recommendation computation increases with the increase of users and projects. At present, rapid and effective information recommendation is a serious challenge. At present, collaborative screening technology faces an urgent problem: improving its real-time performance while ensuring the accuracy of recommendations. Literature [10] proposes a collaborative screening algorithm that can enhance the real-time performance of recommendations and users' real-time feedback updates. Literature [11] designed a real-time recommendation system for collaborative filtering based on Spark distribution and demonstrated the system's reliability.

3. Discussion on personalized music situational cognition.

3.1. Research on personal style characteristics. Personalized music shows immediacy, situational dependence, mobility, and randomness of user needs. In the personalized music environment, the changes in users' time, space, social relations and other related situations make users' interests and needs change. The lives of music users are constantly changing. When the user's information needs to be timed, the situational factors will directly impact the user's interest tendency [12]. Environmental factors can be weather, seasons, time, space, etc. Mobile phone music's mobility, scene dependence and other characteristics make mobile phone music users will have more personal needs.

3.2. Overview of Situational Perception Knowledge.

Concept 1. $Z, Z = \{E_1, E_2, E_3, \cdots, E_n\}, Z \neq \delta$ scenario can be expressed as various types of non-empty sets a about users. E_j stands for item j in this text. n is the number of features in the scene. Case Z_i can be expressed in terms of $Z_i = \{E_{i1}, E_{i2}, E_{i3}, \cdots, E_{in}\}$. In a specific scenario Z_i, E_{ij} is the value of an attribute E_j. Situations can be divided into material situations and social situations.

Concept 2. Action scenario. Context attribute $Z_j = \{x | x = gZ_i(D)\}, D$ represents the variable of the user's location in the action context.

Concept 3. Situational awareness. Context awareness refers to the operation process in which the applied device can sense the context information, process the device's context information, and use the context data information. In fact, in a generalized computing environment, various applications of situation processing can be called situational cognition. Early scene cognition technology has been widely used in pervasive computing, data mining, information retrieval and other fields [13]. As a universal, flexible and personalized mobile scene, it has a broad application prospect in music, traffic navigation and tourism. Therefore, scene cognition service has become a profit point of music festivals.

3.3. Clustering algorithm. Clustering refers to dividing physical or virtual objects into groups of similar objects. So that objects in the same cluster have a high degree of similarity. The difference in goals is not the same in clustering various categories. An efficient clustering method based on K-means is proposed. Clustering entries and users in different scenarios obtain top-N suggestions.

Basic Model of Mobile Phone Situational Cognition Service. The platform collects many situational information, including current situational information, interest preferences, personal characteristics, etc. A customer-oriented personalized recommendation algorithm with high real-time performance and accuracy is proposed. Figure 3.1 shows the architecture of its standard architecture model [14]. The first layer is the mobile intelligent terminal interaction layer. The second layer is the mobile recommendation service layer. The third level is the data preprocessing layer. The fourth level is the data collection layer.

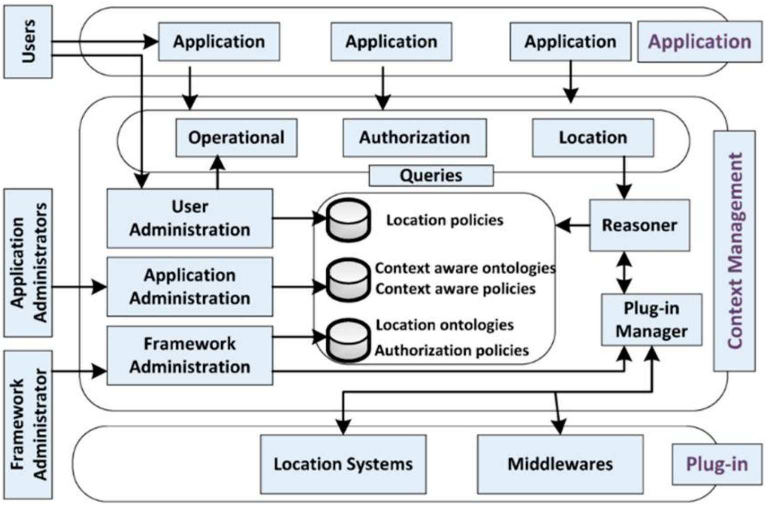

Fig. 3.1: Mobile situational awareness service model.

Context-aware user-item cluster collaborative screening. The traditional collaborative filtering recommendation algorithm mainly considers two dimensions: user and project. However, in the individual music behavior, the user's situation constantly changes, and the user's needs also change. Moreover, the calculation scale of the nearest neighbor also increases linearly with the expansion of the personalized music scale, resulting in a continuous decline in the recommendation accuracy [15]. This project proposes a collaborative screening strategy based on mobile context awareness and clustering. The collaborative filtering method of user-commodity clustering under scene perception is studied. Thus, it can better meet users' personalized requirements for personal music.

Build personalized music, user-project-context model. We adopt personalized music 3D $S - I - Z$ (User, Account, Events) mode. This study aims to more clearly present users' preferences for music items in various contexts (Figure 3.2). The vertical axis is the user dimension [16]. The horizontal axis represents the dimensional dimension of the music entry. Z represents a specific characteristic component of the situation (season, location, weather, occupation, gender, interests, etc.). Coordinate refers to the user S_i score value of I_j in the specific scenario Z_t is $g(S_i, I_j, Z_t)$. A context-based user-entry score matrix $H : S \times I \times Z$ is established.

Scenario-based User-Project clustering. Under the premise of thoroughly combining the scene elements, the user-project clustering operation is realized—offline clustering groups users and items with highly similar situations into the same category.

Determination of scenario similarity. Contextual information is diverse, such as the user's name, gender, online time, hobbies, etc. Specific scene information includes season, weather, time, temperature, location, etc. All the assumed data are quantified as numerical to facilitate the calculation. Due to the complexity and diversity of scenes, the conventional method of scene similarity has been unable to adapt, so the calculation of scene similarity can be carried out by formula (3.1).

$$sim(Z_i, Z_j) = \frac{\sum_{v=1}^{m} \lambda_{ij}^v \times \kappa_{ij}^v}{\sum_{v=1}^{m} \lambda_{ij}^v} \tag{3.1}$$

$\lambda_{ij}^v, \kappa_{ij}^v$ is the indicator function. If the $v$variable in Z_i, Z_j does not appear, it is $\lambda_{ij}^v = 0$; otherwise, it is $\lambda_{ij}^v = 1$. If v variables in Z_i, Z_j are $\kappa_{ij}^v = 1$, then $\kappa_{ij}^v = 1$. The closer the $sim(Z_i, Z_j)$ value is to 1, the higher the similarity is, and the lower the similarity is.

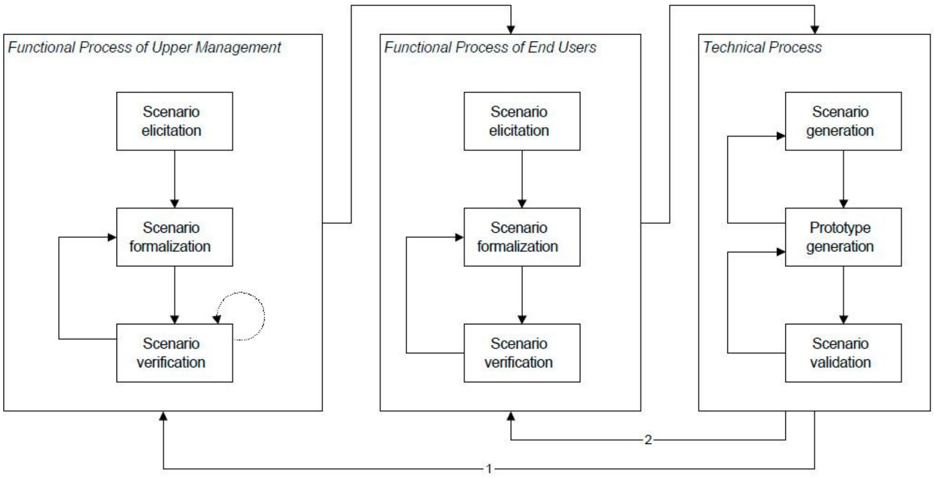

Fig. 3.2: User-project-scenario model.

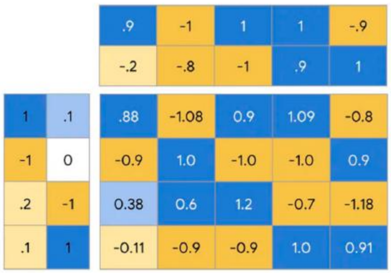

Fig. 3.3: Steps to create a user preferences table.

Constructing User Item Type Preference Scoring Matrix (UPM). The score of an item can directly reflect the user's preference for the same category item. Here, the initial score matrix $H : S \times I \times Z$ can be reconstructed by formula (6). The evaluation matrix of various items by users containing contextual information is established, which is called the "user item category preference scoring matrix".

$$F(S_i)_{Ox} = \sum_{i \in IS_i, x, Z_t} r_{ui}, Z_t / |IS_i, Ox, Z_t| \tag{3.2}$$

I_u, x, Z_t represents the collection of items with characteristic x that are evaluated by the user in ScenarioZ_t. r_{ui}, Z_t is user u real evaluation of item i in the specific situation Z_t. $|I_u, x, Z_t|$ represents how many elements there are in the group I_u, x, Z_t. Start by reading the original score matrix and the relevant information about the categories of entries. Then, the UPM matrix is formed by extracting the user's interest tendency toward the commodity category from (3.2). Figure 3.3 illustrates a simple UPM generation process. For items that are not scored, the data is marked as 0 in the scoring matrix H_1. In the entry type matrix H_2, when the entry has a particular property $O\chi$, its corresponding entry represents 1, and vice versa. The value of the User Presence Matrix (UPM) H_3 is calculated by (3.2).

K-class clusters are obtained by the IC-KM method. K The appropriate value should be chosen so that the collection of items can be better classified rationally. Ensure all items entering the same cluster have similar scores [17]. Filling within the same cluster can effectively eliminate the influence of external indicators on the clustering results and minimize the error rate of the clustering results. This facilitates the filling of the sparse matrix for subsequent scoring.

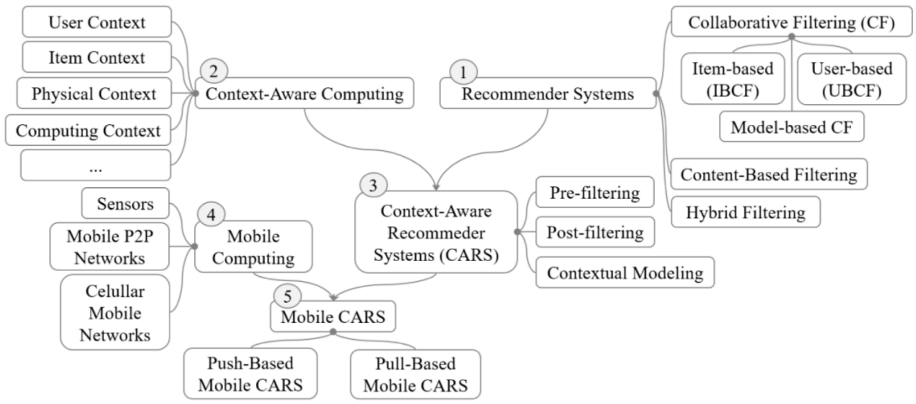

Fig. 3.4: Implementation flow of UIC-collaborative screening algorithm.

Complete user-entry score matrix. A weighted Slope1 algorithm is developed to fill the score matrix according to the similarity between entries. First, the mean of each item is calculated according to the mean in formula (3.3), and the similarity between the items is calculated according to the mean in formula (2.3). The value is then calculated by the weight of the formula (3.4). All the score values missing from the score matrix are obtained in order. Produces A new scoring matrix $H\prime : S \times I \times Z$ with no vacancies. Its scene composition has not changed.

$$dev_{ij} = \sum_{u \in S_{ij}} \frac{r_{ui} - r_{uj}}{sum(S_{ij})} \tag{3.3}$$

$$F(u)_i = \frac{\sum\limits_{j \in I_i} [(dev_{ij} + r_{ui}) \times sum(S_{ij}) \times sim(i,j)]}{\sum\limits_{j \in I_i} [sum(S_{ij}) \times sim(i,j)]} \tag{3.4}$$

S_{ij} is the set of users that evaluate i and j. r_{uj} is the score for item j given by user u. dev_{ij} is the mean of each index i, j. $sum(S_{ij})$ is used to evaluate the number of users for items i and j.

3.4. Context-based user-item cluster recommendation method. This paper presents a user-commodity cluster recommendation method in the context. The volume flow is shown in Figure 3.4 (image cited in Knowledge-Based Systems, 2021, 215:106740).

4. Experimental results and analysis.

4.1. Experimental data. In this paper, 80 volunteers collected music background information daily to test. Experiments and theoretical analysis verify the scenario-based personalized recommendation method. These subjects were wearing special sensors. Collect the user's heart rate, exercise status, and other information, and record the user's situation information in various scenarios. Try to ensure that all categories of background information are covered. The goal is to take full advantage of the value of these features. There are 40,000 records in the original data collected. The complete, clean, preprocessed dataset contains 38,600 records. Finally, the effectiveness of the proposed algorithm is verified by ten cycles of experiments.

4.2. Evaluation Indicators. The evaluation index of a personalized recommendation algorithm cannot be completely equivalent to that of the classification algorithm. This is because similar criteria, such as accuracy, can evaluate the method. In the evaluation system, its evaluation indicators are more diverse. The central performance is correct rate, recall rate, average absolute difference, diversity, surprise degree, etc. Too much accuracy will result in a smaller number of selected items. Increasing the diversity and surprise of the

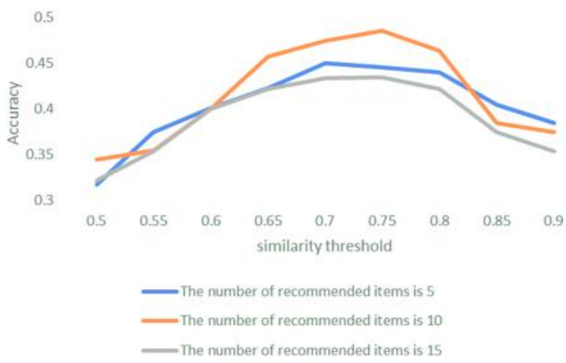

Fig. 4.1: Accuracy of the cooperative screening model.

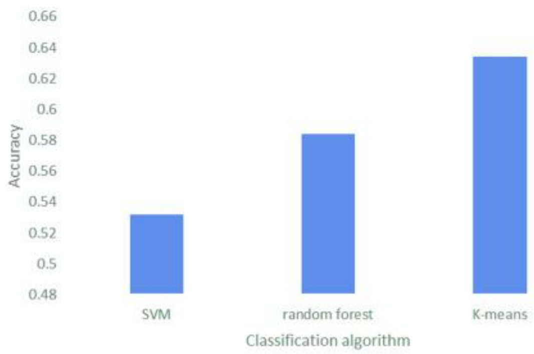

Fig. 4.2: Accuracy of classification model.

recommendation system will inevitably reduce its accuracy. Through the analysis of user satisfaction, a more comprehensive conclusion is drawn.

Accuracy refers to the percentage of recommended lists marked as favorites by the user. $T(u)$ represents the list of associations marked as preferences by the user. $R(u)$ represents the list of suggestions given, then the method of calculating accuracy is

$$\Pr ecision = \frac{\sum\limits_{u \in U} (R(u) \bigcap T(u))}{\sum\limits_{u \in U} |R(u)|} \tag{4.1}$$

In addition, 50 volunteers were tested under different recommendation algorithms to get the evaluation of the system. The user comments on the list of suggestions. The average of these ratings indicates the user's satisfaction with the suggestion system. In addition, based on the feedback information from users with clear goals, the whole recommendation system can be further improved.

4.3. Experimental results and analysis. In the process of collaborative screening, it is necessary to understand the influence of the threshold of user story similarity on the screening results. When the accuracy of the collaborative filtering model is tested according to the size of the threshold, the accuracy results are obtained (Figure 4.1).

Experiments verify the correctness of the algorithm. The average accuracy of the forecast for the classification model is shown in Figure 4.2. Through the analysis of experimental data, it is concluded that both random forest and K-means methods have greatly improved the prediction accuracy.

Fig. 4.3: Accuracy of the fusion model.

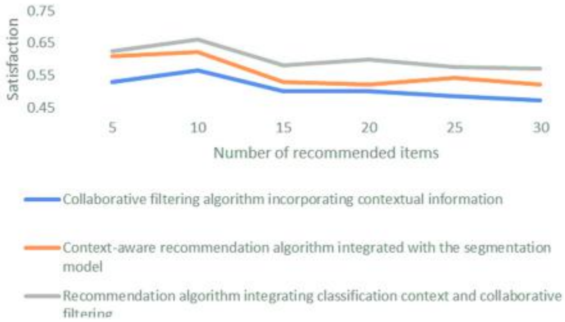

Fig. 4.4: Comparison of user satisfaction.

The precision of this fusion pattern is shown in Figure 4.3. Figure 4.4 shows a comparison of user satisfaction levels. The results show that the accuracy of the method has improved significantly. When K-mean is selected, and the number of recommended items is 10, the accuracy of this method is the best. Compared with the support vector machine method, the accuracy of this method is significantly improved. At the same time, the result obtained by this method is the most satisfactory. The feasibility and effectiveness of this method are demonstrated.

5. Conclusion. Context information is an essential factor affecting the effect of personalized music recommendations. A context-aware music recommendation method is designed by combining situational information and collaborative filtering. Experiments show that combining contextual information and collaborative filtering can obtain higher accuracy and user satisfaction. However, the research on this subject needs more expansion and deepening due to the limitation of the data used and the background data included. Only the music type label is used in the classification model, but no more profound analysis of its speech characteristics is carried out, which has a particular impact on the performance of the recommendation algorithm.

Acknowledgments. This work was supported by Research project of Humanities and Social Sciences in Colleges and universities of Hebei Provincial Department of Education: Research on the inheritance of local music in local universities (NO.SY2022099).

REFERENCES

[1] Zhou, K., Yang, C., Li, L., Miao, C., Song, L., Jiang, P., & Su, J. A folksonomy-based collaborative filtering method for crowdsourcing knowledge-sharing communities. Kybernetes, 2023;52(1): 328-343.

[2] Lv, Z., & Song, H. Trust mechanism of feedback trust weight in multimedia network. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2021; 17(4): 1-26.
[3] Liu, K., Xue, F., He, X., Guo, D., & Hong, R. Joint multi-grained popularity-aware graph convolution collaborative filtering for recommendation. IEEE Transactions on Computational Social Systems, 2022; 10(1): 72-83.
[4] Li, L., Wang, Z., Li, C., Chen, L., & Wang, Y. Collaborative filtering recommendation using fusing criteria against shilling attacks. Connection Science, 2022; 34(1): 1678-1696.
[5] Vahidnia, M. H. Point-of-interest recommendation in location-based social networks based on collaborative filtering and spatial kernel weighting. Geocarto international, 2022; 37(26): 13949-13972.
[6] Ilangovan, P. Support Vector Machine based a New Recommendation System for Selecting Movies and Music. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021;12(10): 1425-1429.
[7] Afchar, D., Melchiorre, A., Schedl, M., Hennequin, R., Epure, E., & Moussallam, M. Explainability in music recommender systems. AI Magazine, 2022;43(2): 190-208.
[8] Papadakis, H., Papagrigoriou, A., Panagiotakis, C., Kosmas, E., & Fragopoulou, P. Collaborative filtering recommender systems taxonomy. Knowledge and Information Systems, 2022;64(1): 35-74.
[9] Wu, L., He, X., Wang, X., Zhang, K., & Wang, M. A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation. IEEE Transactions on Knowledge and Data Engineering, 2022; 35(5): 4425-4445.
[10] Yannam, V. R., Kumar, J., Babu, K. S., & Sahoo, B. Improving group recommendation using deep collaborative filtering approach. International Journal of Information Technology, 2023; 15(3): 1489-1497.
[11] Liu, L. The artistic design of user interaction experience for mobile systems based on context-awareness and machine learning. Neural Computing and Applications, 2022;34(9): 6721-6731.
[12] Borges, R., & Stefanidis, K. Feature-blind fairness in collaborative filtering recommender systems. Knowledge and Information Systems, 2022; 64(4): 943-962.
[13] Zitouni, H., Meshoul, S., & Mezioud, C. New contextual collaborative filtering system with application to personalized healthy nutrition education. Journal of King Saud University-Computer and Information Sciences, 2022;34(4): 1124-1137.
[14] Assuncao, W. G., Piccolo, L. S., & Zaina, L. A. Considering emotions and contextual factors in music recommendation: a systematic literature review. Multimedia Tools and Applications, 2022; 81(6): 8367-8407.
[15] Yao, K., Wang, H., Li, Y., Rodrigues, J. J., & de Albuquerque, V. H. C. A group discovery method based on collaborative filtering and knowledge graph for IoT scenarios. IEEE Transactions on Computational Social Systems, 2021; 9(1): 279-290.
[16] Horasan, F. Latent Semantic Indexing-Based Hybrid Collaborative Filtering for Recommender Systems. Arabian Journal for Science and Engineering, 2022; 47(8): 10639-10653.
[17] Ghosh, S., Tyagi, D., Vashisht, D., Yadav, A., & Rajput, D. A comprehensive survey of personalized music identifier system. The International Journal of Recent Technology and Engineering (IJRTE), 2021; 9(6): 90-96.

Edited by: Zhigao Zheng
Special issue on: Graph Powered Big Aerospace Data Processing
Received: Nov 3, 2023
Accepted: Nov 20, 2023