JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 29 No. 5, pp. 177-187, May 2024
https://doi.org/10.9708/jksci.2024.29.05.177

Enhancing Music Recommendation Systems Through Emotion Recognition and User Behavior Analysis

Qi Zhang*

*Docotrol couse, Dept. of Global Music, Kyonggi University, Seoul, Korea

[Abstract]

177-Existing music recommendation systems do not sufficiently consider the discrepancy between the intended emotions conveyed by song lyrics and the actual emotions felt by users. In this study, we generate topic vectors for lyrics and user comments using the LDA model, and construct a user preference model by combining user behavior trajectories reflecting time decay effects and playback frequency, along with statistical characteristics. Empirical analysis shows that our proposed model recommends music with higher accuracy compared to existing models that rely solely on lyrics. This research presents a novel methodology for improving personalized music recommendation systems by integrating emotion recognition and user behavior analysis.

▸ **Key words**: Music emotion recognition, User behavior trajectories, Online Music Recommendation System

[요　　약]

요약 배경: 기존 음악 추천 시스템은 가사의 의도된 감정과 사용자가 실제로 느끼는 감정 사이의 불일치를 충분히 고려하지 않았다. 모델: 본 연구에서는 LDA 모델을 활용하여 가사와 사용자 댓글의 주제 벡터를 생성하고, 시간 감쇠 효과와 재생 횟수를 반영한 사용자 행동 궤적과 통계 특성을 결합하여 사용자 선호도 모델을 구축했다. 결과: 실증 분석 결과, 제안 모델이 가사만 활용한 기존 모델보다 높은 정확도로 음악을 추천했다. 시사점: 본 연구는 감정 인식과 사용자 행동 분석을 통합하여 개인화된 음악 추천 시스템을 개선하는 새로운 방법론을 제시한다.

▸ **주제어**: 음악 감정 인식, 사용자 행동 궤적, 온라인 음악 추천 시스템

• First Author: Qi Zhang, Corresponding Author: Qi Zhang
*Qi Zhang (zhangqi1314521@hotmail.com), Dept. of Global Music, Kyonggi University
• Received: 2024. 02. 02, Revised: 2024. 04. 01, Accepted: 2024. 04. 01.
• This paper was submitted to and presented at ICEE-CCA 2023

I. Introduction

The Music Recommender System (MRS) has become a crucial research area, particularly in the vast online music market where competition is intense among platforms like Apple Music, Spotify, Pandora, and Chinese services such as NetEase Cloud Music, QQ Music, and Kugou Music. Attracting users is essential for these platforms to maintain a competitive edge. The solution lies in developing MRS tailored to individual user preferences, enabling platforms to recommend music that resonates with users, thereby increasing their engagement and loyalty [1].

Currently, MRS construction is generally designed around user-object interaction and content-based object description. However, it should be noted that users' taste and demand for music are highly dependent on their internal emotional cognition, music evokes various emotions of users, and users' emotions also affect their music preferences. There is a strong emotional connection between music and users [2]. Therefore, Music Emotion Recognition (MER) has become an active topic in the research field of MRS.

Music Emotion Recognition (MER) involves labeling emotional words in music but integrating it into Music Recommendation Systems (MRS) faces three main challenges. Firstly, MER often overlooks the difference between what is intended to be conveyed (the creator's intent) and what is actually felt (the listener's actual emotions). This difference can be significant[3]. Secondly, listeners' personal emotional states may not align with the emotion of the song, and they may seek to improve or alter their mood through music[4]. Thirdly, emotional responses to songs can vary over time, and the emotional demands of listeners during initial song listening may change during subsequent listens[5]. To address these challenges, music recommendations need to finely adjust individual listeners' preferences and emotional responses to content.

Online music platforms often classify songs into predetermined emotional types for content-based recommendations. For example, NetEase Cloud Music features 12 emotional categories such as nostalgia, romance, and happiness, while QQ Music uses 8 categories such as sentimental and empathetic. Matching users' emotions to these song categories effectively can significantly enhance recommendation accuracy.

This paper focuses on NetEase Cloud Music and establishes user preferences based on platform interactions. To resolve discrepancies between user preferences and emotional expressions in song lyrics, this research utilizes both objective text vectors from user listening records of Chinese song lyrics and subjective text vectors from users' direct song comments. This integrates the intended emotions of songs that resonate with users' feelings. To account for changes in user emotions over time, short-term (song playback trajectories) and long-term (statistical characteristics of listening behaviors) behavioral characteristics are combined to develop a more accurate user preference model. This approach provides a new method for emotion-based online music recommendations and underscores the value of user feedback in the music industry.

II. Literature Review

2.1 Music recommendation

In recent years, Music Recommendation Systems (MRS) have garnered widespread attention in both the academic and industrial sectors, both in China and internationally. A successful MRS can sift through hundreds of millions of songs from various music libraries to select music that users love, thereby preventing choice overload [6]. An effective MRS should consider internal factors (such as a user's personality and emotions), external factors (such as user behavior), and contextual factors (such as the weather conditions, social status, and listening location at the time of listening). However, current MRS research has not sufficiently focused on the element of emotion. In fact, emotion is a crucial psychological construct that influences users' music preferences and their selection of MRS.

Emotion-based MRS have three main objectives: (1) to identify the emotional characteristics of music; (2) to recognize the emotional state of the user; and (3) to understand the interaction between music and the user. Currently, most popular online music platforms use tags to recommend music. These tags reflect the social group's emotional identification with the songs and to some extent influence the listening behavior of individual users. However, the emotional perception of music by individual users is unique and does not completely align with the group's emotions. Research shows that analyzing users' comments on music platforms can better capture their emotional perception of music, leading to more effective recommendations [7].

Furthermore, time is a primary contextual element that affects user interest. It directly impacts the change in user interest preferences and the effectiveness of the recommendation system. Therefore, research should select appropriate factors to examine the temporal changes in users' emotions. Although there is a correlation between music preference and personality, a general theory of music preference can only be developed through investigations in different countries and cultures.

2.2 User preference and user behavior

The concept of user preference, originating from philosophical discussions, has been notably addressed by Aristotle, who described it as an individual's tendency in evaluating multiple options or states. Recent years have witnessed the evolution of user preference modeling as a distinct element in personalized service research. Scholars have categorized user preferences into positive and negative, leading to the development of formal models for personalized recommender systems. Techniques from Multi-Criteria Decision Analysis (MCDA), consumer psychology, and other theoretical frameworks have been integrated with collaborative filtering methods to construct advanced models that outperform basic rating systems [8].

A significant innovation in this field is the introduction of the time factor, acknowledging the dynamic nature of user preferences. Contemporary studies incorporate time to examine the evolution and fading of preferences. Notably, a method has been proposed that combines short-term memory models based on query-related concepts and user interaction data, introducing a forgetting factor to refine personalization [9]. Additionally, a time-aware recommendation system has been designed, leveraging neural collaborative filtering and combining long-term interaction histories with recent user feedback.

However, a comprehensive understanding of the factors driving changes in user preferences and their impact on recommendations remains limited. User behavior analysis, a crucial aspect of consumer psychology, provides valuable insights into human thought processes [10]. This research often involves analyzing users' online behavior, including search engine queries, e-commerce interactions, and reading habits. While studies on music playing behavior trajectories are sparse, findings from various disciplines indicate a strong link between user behavior, emotional expression, and preferences [11]. This suggests the potential value in examining music playing trajectories to understand user preferences [12].

2.3 LDA topic model and user recommendastion

Latent Dirichlet Allocation (LDA), a prominent topic in natural language processing, has attracted significant attention for its effectiveness in dimensionality reduction and has been implemented in major platforms like Microsoft's lightLDA and

Tencent's LDA. Current LDA research predominantly concentrates on extracting the textual content's objective aspects [13]. However, incorporating subjective user preference data, particularly emotional elements, is still a nascent area.

Some studies have started integrating human emotion into LDA. For instance, research has explored user sentiments at various textual levels—word, sentence, and text—focusing on the conveyed meanings or attitudes rather than aligning the model with individual behavioral patterns. The User Sentiment-Topic Model (USTM) stands out by incorporating sentiment information to discern user topics and sentiments, classifying terms as neutral, positive, or negative to gauge sentiment trends. Another approach, the Affective Topic Mode (ATM), aims to detect emotions induced by social media content, introducing a layer that connects content with reader emotions [14].

However, these studies mainly focus on emotions as conveyed by texts, treating them as objective data and not integrating direct user feedback. Recent research acknowledges that themes evolve over time, prompting efforts to appropriately incorporate the time element into thematic models.

In music recommendation, studies utilizing LDA for emotion-based suggestions are sparse. There's a need to blend individual emotional states with collective emotional perceptions, along with recognizing preferences that shift over time, including long-term and short-term behavioral trends. Since understanding the emotional theme of a song parallels textual theme cognition and isn't static, employing LDA can offer a more nuanced and accurate foundation for music recommendation decisions [15].

III. Modelling Design

This paper outlines a four-step process to develop our user preference model. First, we gather user behavior characteristics from NetEase Cloud Music through a web crawler. Next, we enhance our understanding of user music preferences by analyzing both user-generated comments and song lyrics on NetEase Cloud Music using the LDA model, categorizing them into thematic groups to create a comprehensive textual profile.

The third step involves constructing a model that integrates these textual profiles with temporal aspects of user behavior. This model assigns time-related attributes to the texts, combines user behavior statistics, and balances long-term behavioral trends to form a cohesive user preference model [16].

Finally, we conduct an empirical study and comparative analysis. Using data from NetEase Cloud Music, we test the model's validity and compare its effectiveness in music recommendation against other methods. The structure of our research is depicted in Figure 1.

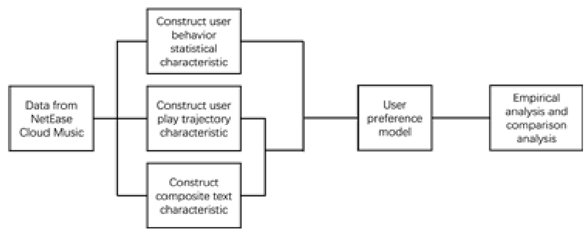

Fig. 1. Overall research framework

3.1 Constructing user behavior statistical characteristic and user play trajectory characteristic

Major online music platforms gauge user music preferences through two primary methods. The first involves user interactions like commenting, replying, and liking songs they enjoy. The second reflects preferences through listening histories, such as frequent replays of favored songs. This study aims to build a model focusing on two aspects: user behavior statistics and playback trajectory characteristics. This model is based on data extracted from music playback histories using Python.

Song characteristics vary by genre. For instance, rap songs typically feature shorter durations with more lyrics, while folk songs are longer with fewer lyrics. Additionally, user emotions vary throughout the day, influencing their music choices—upbeat music is preferred in the morning, whereas softer tunes are favored in the evening. Therefore, constructing a reliable user behavior statistical profile should consider factors like average listening duration, lyric length, the ratio of music to lyric length, and peak listening times. We denote user behavior statistical characteristic vector as F_{STA}:

$$F_{STA} = (F_{ad}, F_{al}, F_{lr}, F_{pp}, F_{pm})$$

F_{ad} represents users' average listening duration; F_{al} represents average length of lyrics; F_{lr} represents the ratio of average length of music to average length of lyrics; F_{pp} represents high-frequency listening period; F_{pm} represents times that the most frequently listened music being played. Every song has its playing trajectory when user listen to it. Those trajectory includes name of the song, the ID of the song, lyrics, period of the song, name of singer, name of the album, the time stamp listened by the user.

Every time user listens to a song, there will be a play track left, including: song name, song ID, lyric information, song duration, singer name, album name, timestamp, user comments, and the number of times the song has been played. Define the user play track characteristic vector F_s:

$$F_s = (F_{sn}, F_{sid}, F_{sl}, F_{sd}, F_{sr}, F_{an}, F_{time}, F_{uc}, F_{count})$$

F_{sn} denotes the name of the song; F_{sid} denotes the ID number of the song; F_{sl} denotes the lyrics information; F_{sd} denotes the duration of the song; F_{sr} denotes the name of the singer; F_{an} denotes the name of the album; F_{time} denotes the timestamp of the song that the user has listened to; F_{uc} denotes the user's music comments. F_{count} denotes the number of times that the user has listened to this song.

3.2 Construct composite text characteristics

Composite text characteristic s are characteristic s extracted from user behavioral trajectories, taking into account the objective theme of the song and the personalized interest in the song's theme in the user's comments, so as to reflect the user's own preferences for the song.

Take a song record j in the track set as an example. Using the LDA model to process the lyrics corresponding to this record, we can obtain the subordination degree of the song in terms of each theme. Since the lyrics are constant, each user gets the same subject affiliation vector. Since it is not related to the subjective preference of users, it is defined as the objective text vector F_{LLDA}:

$$F_{LLDA}(j) = \{L^{(0)}, L^{(1)}, \cdots, L^{(j)}, \cdots, L^{(n)}\}$$

in which, $L^{(j)}$ represents the degree of subordination of objective text (lyrics) on the j-th theme. But in the matter of fact, users are more likely to mention in their comments the themes of the song that interest them, which are not necessarily the main meaning of the lyrics. For example, two users, A and B, are both interested in the topic space $\{S1, S2, S3, S4\}$ with the degree of interest $\{0.2, 0.2, 0.2, 0.4\}$. However, user A is accustomed to mentioning only the topic T4 that interests him the most in his comments (his LDA-processed subordination degree is $\{0.05, 0.04, 0.05, 0.86\}$); similarly, when user B comments, he prefers to mention only the topics he is interested in, which is not the main meaning of the song. Similarly, B prefers to focus on the topics he is interested in when commenting (his subordination vector may be $\{0.22, 0.23, 0.25, 0.3\}$). It is obvious that these two users have different preferences. Thus the user comments need to be pre-processed to identify their perceptions of the song content. Define the user comment subject subordination vector as the subjective perception vector F_{SLDA}:

$$F_{SLDA}(j) = \{I^{(0)}, I^{(1)}, \cdots, I^{(j)}, \cdots, I^{(n)}\}$$

Where, $I^{(j)}$ denotes the degree of subordination of the subjective text (the user's music comments) on the ith topic. User comments can be null.

Since subjective perception vectors cannot exist independently from the song contents, the user true preference vector should be the result of combination of objective text vectors and subjective perception vectors, for which, F_{LLDA} and F_{SLDA} need to be integrated into a composite text characteristic $topic\_F_{LDA}$:

$$F_{LDA} = (1-\alpha)F_{LLDA}(j) + \alpha F_{SLDA}(j)$$

where, α is a hyperparameter indicating the importance of the comment.

3.3 Characteristic Fusion and User Preference Modeling

After the establishment of composite text characteristic s, considering the influence of time factor and long-term behavioral characteristics of users on user preferences, these two characteristic s should be combined into user preference model.

(1) combination of user play trajectory characteristic.

A user's play trajectory is a time series. The longer the time elapsed from the current time, the smaller the impact on the user's current preference. Define the vector $time\_opic\_F_{LDA}$ as the preference vector for combined time attenuation effect:

$$time\_topic\_F_{LDA} = topic\_F_{LDA}(j) \times \beta^{\left|\frac{t}{T}\right|}$$

where β is the attenuation step $(0 < \beta < 1)$; T is the attenuation period; t is the time interval between this record and the current time.

In addition, the number of occurrences of a song also has an effect on preference. The first occurrence has the largest effect on user preference, and the effect on preference tends to zero after multiple occurrences. We fit with sigmoid function and define $count\_time\_opic\_F_{LDA}$ as the preference vector for combined number of times:

$$count\_time\_topic\_F_{LDA} = time\_topic\_F_{LDA}(j) \times [S(x) - S(x-1)]$$

where $S(x)$ is a sigmoid function; x denotes the number of times the song appears.

Normalize $count\_time\_topic\_F_{LDA}$ of all m records using equation (7) to obtain the preference vector combined with user play trajectory, denoting as tra_F_{LDA}:

$$tra\_F_{LDA} = \frac{1}{m}\left(\sum_{x=1}^{x=\omega} count\_time\_F_{LDA}(j)^x\right)$$

(2) combination of user behavior statistical characteristic .

In order to obtain the universal properties for songs in the same category as well as neutralize the effect of deceasing weights in long-term preference in the previous step due to the time attenuation function, combination of behavior statistical characteristic and preference vector is necessary. In order to combine the user behavior statistical characteristic with tra_F_{LDA}, first of all, we should add parameter ω to process the one-dimension behavior statistical vector F_{STA} , generating a new $1 \times n$ dimensional space vector F^*_{STA} , which yields:

$$F^*_{STA} = F_{STA} \times \omega$$

where parameter ω is a $5 \times n$ dimensional space vector; n is the number of topics. Normalize F^*_{STA} and tra_F_{LDA}, obtaining the final user preference characteristic vector F_{LDA}:

$$F_{LDA} = (1-\gamma)tra\_F_{LDA} + \gamma F^*_{STA}$$

where: γ is a hyperparameter indicating the importance of user behavior statistical characteristic .

3.4 Model training

The key point of model training is to find a combination of parameter α (which indicates the importance of a user's music comments) and parameter γ (which indicates the importance of a user behavior statistical characteristic) for each volunteer. Since the model is designed for each user, the values of these two parameters are personalized.

Considering $\alpha, \gamma \in [0,1]$, we set a step size of 0.1 to compute the user preference vector for different sets of values of parameters α and γ. A good user preference vector should recommend a song set that maximally overlaps with the user's actual favorite songs.

For this purpose, a library of training tunes is built as a recommendation candidate, and the optimal combination of two parameter values can theoretically be obtained by the following steps: (1) Substitute the user information into the user preference model to generate a preference topic vector, compute the JS dispersion with the subject vectors of the candidate songs and rank them, and obtain song ranking as the prediction of user music preference. The proximity of the recommended songs to the user preferences can be measured by the Jensen-Shannon divergence, denoting as D_{JS}:

$$D_{JS}(\theta_F, \theta_S) = \frac{1}{2}\left(\sum_{i=1}^{n}\theta_F^{(i)}\ln\frac{\theta_F^{(i)}}{\omega^{(i)}} + \sum_{i=1}^{n}\theta_S^{(i)}\ln\frac{\theta_S^{(i)}}{\omega^{(i)}}\right)$$

$$\omega^{(i)} = \frac{\theta_F^{(i)} + \theta_S^{(i)}}{2}$$

where θ_F denotes the topic probability distribution in the user preference model; θ_S denotes the topic probability distribution of the recommended song lyrics text; $\theta_F^{(i)}$ and $\theta_S^{(i)}$ denote the values of the same topic in the two texts; n denotes the number of topics. The Jensen-Shannon divergence ranges from $[0,1]$. A smaller Jensen-Shannon divergence value indicates a higher similarity between two topic distributions. In other words, the smaller the Jensen-Shannon divergence value, the better the recommendation effect. (2) Send the trained music library to volunteers, let them identify their favorite songs. (3) Compare ranking of songs predicted in step (1) with the actual favorite songs in step (2), obtain the matching degree between the songs predicted in the model and the actual favorite songs. Choose a pair of parameters that gives the highest matching degree as the value of parameters in the user preference model. In practice, to improve efficiency, in step (1), we choose the top 20 songs with the least Jensen-Shannon divergence value, and in step (2), let user select their top 20 favorite songs, then matching degree (MD) can be calculated as follows:

$$MD = \frac{M_n}{20}$$

M_n is the number of overlapping songs in the set of predicted songs and the set of user's actual favorite songs .

IV. Empirical studies

This study involves training a user preference model using the online music playback data from NetEase Cloud Music users. We aim to determine the optimal parameter values for the model. Once these parameters are established, we will conduct a comparative analysis with other relevant recommendation algorithms. This comparison is designed to validate the effectiveness of our proposed model.

4.1 Data collection and processing

In this research, we focus on Chinese songs and their reviews, using NetEase Cloud Music as our data source due to its extensive collection of high-quality user-generated music reviews. We employed a Python web crawler to gather the data, subsequently storing it in a MySQL database. The process of data collection and preprocessing involved several steps: (1)Initially, the crawler gathered data on 148,326 users; (2)We then filtered out non-Chinese songs, narrowing down to 44,463 users; (3)Further refinement was done by removing users with less than 20 reviews, resulting in 11,698 users; (4) For these remaining users, we collected detailed data including their song playbacks, lyrics, comments, and listening timestamps. The aggregated data on user information, statistics, and song tracks are presented in Tables 1-3.

Table 1. Example of Fundamental User Information

User ID	372915791
Nick Name	Waterfall
Sex	Male
Age	19
City Name	Shandong-Yantai

Table 2. Example of user's statistics

index	average	Std.	Max	Min
average length of songs	241.09	14.10	378	206
average length of lyrics	213.53	42.32	297	126
ratio of length of songs to length of lyrics	1.76	0.38	2.52	1.14
high-frequency listening period	183	219	1092	7

Table 3. Examples of user song playing track

Song ID	108299	376216
Song Name	Do you remember the dream in the youth?	Guard a windoe for you
Lyrics	Like a flower that never adjusts...	Whether it's dusk or dawn...
Durations/s	269	211
Singer	Zongsheg LI	Shuimunianhua
Album	the theory of love	Rhapsody on life
Timestamp		
Create time	2023/8/9 23 : 22	2023/7/13 20 : 04
Count type	10	5
Info	song	song

4.2 Model Training and Parameter Selection

Online music platforms typically categorize songs into specific lyric and melody categories. In this study, we adapted these classifications to the context of Chinese music, especially considering NetEase Cloud Music's categorization system. We organized song texts into 12 sub-themes: nostalgia, romance, sadness, healing, tranquility, cheerfulness, inspiration, relaxation, loneliness, sensuality, touching, and excitement.

For our recommendation model's training library, we selected ten songs from each sub-theme, resulting in a total of 120 songs. A corresponding NetEase Cloud Music playlist was compiled from this selection to aid in model training. We gathered data from 262 volunteers, including their account information and playback history of Chinese songs. Using the approach described in Chapter 2, Section 4, we determined two key parameters of the user preference model. The match between the parameter values and the actual preferences of a user is shown in Table 4. It can be seen from the table that for this user, the recommended list matches his/her real preference best when his/her personalization parameters $\alpha = 0.6$ and $\gamma = 0.3$.

Table 4. Selection and matching degree of parameters α and γ

γ/α	matching degree										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.0	0.20	0.30	0.35	0.31	0.30	0.45	0.65	0.35	0.25	0.30	0.20
0.1	0.30	0.35	0.35	0.35	0.35	0.45	0.65	0.35	0.25	0.25	0.30
0.2	0.30	0.25	0.55	0.20	0.50	0.70	0.70	0.25	0.45	0.20	0.15
0.3	0.30	0.25	0.25	0.45	0.45	0.65	0.80	0.60	0.45	0.25	0.25
0.4	0.20	0.20	0.55	0.55	0.65	0.70	0.70	0.65	0.30	0.20	0.20
0.5	0.20	0.20	0.35	0.35	0.50	0.55	0.70	0.60	0.45	0.30	0.25
0.6	0.35	0.35	0.25	0.35	0.50	0.50	0.50	0.50	0.40	0.40	0.35
0.7	0.25	0.25	0.35	0.25	0.45	0.45	0.35	0.50	0.40	0.30	0.35
0.8	0.20	0.20	0.30	0.25	0.30	0.35	0.20	0.40	0.45	0.30	0.30
0.9	0.15	0.15	0.25	0.30	0.20	0.35	0.30	0.10	0.25	0.25	0.25
1.0	0.10	0.15	0.25	0.25	0.10	0.30	0.25	0.10	0.30	0.25	0.25

From this, the constructed preference model can be applied to calculate the probability of a user's preferred topic , and some of the results are shown in Table 5.

Table 5. Examples of user LDA topic distribution probability

theme	probalitity distribution				
	39012414	40886925	48257292	50497671	52475077
nostalgia	0.052585	0.080628	0.03476	0.063419	0.468985
romance	0.071848	0.03925	0.059518	0.063711	0.620343
sadness,	0.030001	0.0080649	0.071256	0.007088	0.610872
healing	0.049834	0.063865	0.018196	0.012355	0.602106
tranquility	0.073048	0.007449	0.010499	0.032679	0.160873

4.3 Comparative analysis

In text mining, music preference and personalized recommendation research primarily fall into two categories: (1) Simple lyric text theme mining, which analyzes the semantic content of song lyrics to recommend tracks semantically similar to those in a user's listening history; and (2) Text theme mining based on user music reviews, focusing on the integration of review text and song list text in recommendation algorithms. In our study, we used the objective theme vectors and combined text theme vectors of volunteers' Chinese song playback histories as a control group. We then compared the average matching degree of this group with the recommendations generated by our user preference model, using 20 songs from our song library for each group. The results of this comparison are presented in Table 6.

Table 6. Matching degree comparison of the recommended results

theme vector	matching degree
lyrics	0.475
composite text	0.535
user play trajectory	0.645
user preference	0.705

The comparison table of matching degrees clearly demonstrates that the text theme vector, which integrates both subjective and objective theme vectors, outperforms the recommendation algorithm that solely uses objective theme vectors for constructing the user preference model. Additionally, the preference model utilizing theme vectors from users' playback trajectories yields better results than the model using text theme vectors alone. Furthermore, the user preference theme vectors developed in this study are shown to be the most effective for recommendation purposes.

Beyond just matching degrees, this study also employed the Average Rank Score metric to assess the accuracy of different recommendation methods. For a user U, the rank score of their favored song S is defined as follows:

$$S_{US} = \frac{K_{US}}{N_U}$$

Where: N_U denotes the length of the recommendation list, i.e. the total number of recommended songs, which is set to 20 here; K_{US} denotes the ranking of song S in the recommendation list. The accuracy of user's recommendation can be measured by the ranking score: the lower the ranking score is, the more the recommendation system tends to rank the songs preferred by the user first, and the recommendation effect is good; on the contrary, the algorithm accuracy is low and the recommendation effect is not good. The results are shown in Table 7.

Table 7. Accuracy comparison of recommended results

theme vector	accuracy
lyrics	0.3988
composite text	0.236
user play trajectory	0.1528
user preference	0.0712

From the comparison of the data in Table 7, it can be seen that the recommendation effect of the user preference model is very similar to the real preference of the users and is significantly better than other topic recommendation methods.

V. Conclusions

In this paper, we employed the LDA model to create both objective subject vectors for songs and subjective subject vectors from user comments, addressing the discrepancy between users' personal feelings and the emotions conveyed by songs. We integrated user behavioral trajectories, accounting for the influence of playback patterns like timing and frequency, and the impact of users' behavioral statistics on preferences. This culminated in the establishment of a user preference model, which we validated using data from NetEase Cloud Music. The empirical findings indicate that the model, focusing on user behavioral trajectories, is effective in terms of both match quality and recommendation accuracy. This research contributes a novel approach to online music personalized recommendations, reinforcing the idea that amalgamating diverse elements in the big data era is key to enhancing Music Recommendation Systems (MRS).

However, it's important to note that user behavior on online music platforms encompasses more than just song listening activities. Online music platforms, like NetEase Cloud Music, are increasingly evolving into social platforms where users engage in activities such as commenting and liking. These social behaviors are not yet incorporated in this study. Future research could expand to include these social behaviors in user behavioral trajectories and explore the broader impact of these comprehensive user activities on song selection and listening experiences.

The recommendation algorithm utilizing only lyric theme vectors, which applies the song's lyrics for personalized music recommendations, shows inferior performance compared to other music recommendation methods. The effectiveness of the comprehensive text theme vector, integrating both lyrics theme vectors and user comment theme vectors, is significantly enhanced. This underscores the strong personal nuance in users' perceptions of songs. While users' interpretations are based on the song lyrics' themes, their subjective perceptions differ from the objective textual expression of the lyrics, with varying depth and intensity among different users for the same song.

The algorithm based on users' playback trajectory theme vectors, considering factors like playback frequency and timing, further improves the recommendation quality. It's noted that users often play their favorite songs repeatedly without necessarily commenting each time, and the frequency of playback is much higher than the number of comments on songs.

The user preference vectors developed in this study outperform those based on playback trajectory theme vectors. This superiority is attributed to the inclusion of user behavioral statistics in the user preference vectors. These statistics mainly reflect long-term user preferences. Experimental results show that although short-term preferences significantly influence current user choices, long-term preferences remain crucial in shaping current preferences.

REFERENCES

[1] Karn, A. L., Karna, R. K. “Customer centric hybrid recommendation system for E-Commerce applications by integrating hybrid sentiment analysis”, Electronic Commerce Research, Vol. 23, Issue 1, No 13, pp. 279-314, 2022. DOI: 10.1007/s10660-022-09630-z

[2] Liu, Z., Xu, W. “An emotion-based personalized music recommendation framework for emotion improvement”, Information Processing & Management, Vol. 60, Issue 3, 2007. DOI: 10.1016/j.ipm.2022.103256

[3] Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the experience sampling methodology. Musicae Scientiae, 5(1), 9-32. https://doi.org/10.1177/102986490100500102

[4] Witvliet, C. V., & Vrana, S. R. (1996). The emotional impact of instrumental music on affective states and psychophysiological measures. Journal of the American Academy of Audiology, 7(5), 341-351.

[5] Krause, A. E., & North, A. C. (2016). Music listening in everyday life: Devices, selection methods, and digital technology. Psychology of Music, 44(1), 129-147. https://doi.org/10.1177/0305735614559065

[6] Zhao, T., Li, C., Ding, Q., Li, L. (2012). “User-sentiment topic model: refining user's topics with sentiment information”,

Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics.

[7] Rao, Y., Li, Q., Wenyin, L., Wu, Q., Quan, X. (2014). "Affective topic model for social emotion detection", Neural Networks, Vol. 58, pp. 29-37. DOI: 10.1016/j.neunet.2014.05.003

[8] Zhao, F., Ren, X., Yang, S., Han, Q., Zhao, P., Yang, X. "Latent dirichlet allocation model training with differential privacy", IEEE Transactions on Information Forensics and Security, Vol. 16, pp. 1290-1305, 2020. DOI: 10.1109/TIFS.2020.3032021

[9] Zhu, K., Xiao, Y., Zheng, W., Jiao, X., Hsu, C.-H. "A novel context-aware mobile application recommendation approach based on users behavior trajectories", IEEE Access, Vol. 9, pp. 1362-1375, 2020. DOI: 10.1109/ACCESS.2020.3046654

[10] Huang, Z., Stakhiyevich, P. "A time-aware hybrid approach for intelligent recommendation systems for individual and group users", Complexity, Vol. 2021, pp. 1-19, 2021. DOI: 10.1155/2021/8826833

[11] Chkhartishvili, A. G. "The problem of finding the median preference of individuals in a stochastic model", Automation and Remote Control, Vol. 82, pp. 853-862, 2021. DOI: 10.1134/S000511792105009X

[12] Kang, D., Seo, S. "Personalized smart home audio system with automatic music selection based on emotion", Multimedia Tools and Applications, Vol. 78, pp. 3267-3276, 2019. DOI: 10.1007/s11042-018-6733-7

[13] Murciego, Á. L., Jiménez-Bravo, D. M., Román, A. V., Santana, J. F. D. P., Moreno-García, M. N. "Context-aware recommender systems in the music domain: A systematic literature review", Electronics, Vol. 10, No. 13, pp. 1555, 2021. DOI: 10.3390/electronics10131555

[14] Schedl, M., Zamani, H., Chen, C.-W., Deldjoo, Y., Elahi, M. "Current challenges and visions in music recommender systems research", International Journal of Multimedia Information Retrieval, Vol. 7, pp. 95-116, 2018. DOI: 10.48550/arXiv.1710.03208

[15] Schedl, M., Gomez, E., Trent, E. S., Tkalčič, M., Eghbal-Zadeh, H., Martorell, A. "On the interrelation between listener characteristics and the perception of emotions in classical orchestra music", IEEE Transactions on Affective Computing, Vol. 9, No. 4, pp. 507-525, 2017. DOI: 10.1109/TAFFC.2017.2663421

Authors

Qi Zhang graduated from Anshan Normal University in China with a bachelor's degree and a master's degree from Kyonggi University. She received his master's degree in 2021.

In 2021, she started studying for a doctorate in global music at Kyonggi University. Currently dedicated to international cultural research and educational research, electronic music, etc.