Research and Practice of Search Engine Technology in Mobile Internet Music Field

Lingdan Wang¹,*

¹Guangzhou College of Technology and Business, Guang Zhou, 510850, Guangdong, China

*Corresponding author e-mail: zouyd6@mail.sysu.edu.cn

Abstract. With the continuous progress of global information construction, massive text information and multimedia information can be retrieved and used on the Internet, and music, as an important part of multimedia, has a huge amount of information. Music is a kind of art, which expresses people’s thoughts and feelings and reflects real life through the artistic image formed by organized music. Music is not only an art, but also a social phenomenon of human beings, which is produced with the appearance of human beings. In order to optimize the music search experience of mobile Internet users, new requirements are put forward for the performance of search engine, search precision and recall, and timeliness of update. The classification of music files in the existing Internet music library is only based on the superficial text information such as style, age, creator, etc., without considering the influence of music on human emotion. In order to improve and solve these two problems, this paper proposes a semantic based music retrieval.

Keywords: Search Engine, Music, Search, Internet

1. Introduction

With the rapid development of the Internet, all kinds of information are flooding the Internet, and the amount of online data is almost exponentially increasing [1]. Many people use the Internet as an important means to acquire knowledge and understand society [2]. With the continuous progress of global information construction, massive text information and multimedia information can be retrieved and used on the Internet, and music, as an important part of multimedia, has a huge amount of information. Information on the Internet exists in many different forms, including words, and other forms of information can usually be expressed in words [3]. Information retrieval technology meets people's needs, but because of its universal nature, it still cannot meet users' requests from different backgrounds, different purposes and different periods. Personalized service technology is put forward
to solve this problem, which provides different services for different users to meet different needs [4].

A search engine is a tool for retrieving information. Developers can prepare information and data in various ways. Search engines can store relevant information in index files. After users submit the query content to the search engine, after retrieval Return the relevant content to the user [5]. Personalized service learns users' interests and behaviors by collecting and analyzing user information, so as to achieve the purpose of active recommendation. Personalized service technology can fully improve the service quality and access efficiency of the site, thereby attracting more visitors.

With the rapid popularity of the Internet, the music entertainment industry has experienced a major transformation from death to life, from physical to virtual, from offline to online, and from traditional media to new media [6]. In order to optimize the music search experience of mobile Internet users, new requirements have been put forward for search engine performance, search accuracy and recall rates, and timeliness of updates. Users have only two concepts of familiarity and unfamiliarity with music. If it is familiar music, the traditional retrieval method can solve the user's needs [7]. The recommendation system is an important part of the personalized service technology. It infers the content that the user may be interested in based on the user's existing behavior. When users use search engines, they need to provide certain information to search engines. This is a very active behavior, and the needs of users are relatively clear [8]. However, the recommendation system is just the opposite. The recommendation system does not require the user to give input information when using it. With the rapid development of digital music, it has also brought tremendous changes to the entire music industry. The traditional record industry is experiencing a violent impact and undergoing a process of nirvana rebirth [9]. The classification of music files in existing music libraries is only based on superficial text information such as style, age, creator, etc., without considering the influence of music on human emotions. In order to improve and solve these two problems, this paper proposes a semantic-based music retrieval.

2. Key issues of music search

Different from general search engines, music search is a vertical field, and users' demands are to find songs, and their demands are relatively concentrated. However, it often happens that songs have the same title and even singers have the same name. In search results, the search relevance of different resources is often the same. Therefore, when sorting search results, other factors should be considered besides relevance, such as the popularity of songs. Before establishing ontology concept set, we can consider reusing existing ontology. Because ontology itself is used for knowledge sharing and reuse, if there is an existing ontology associated with the related concepts of design ontology, then reusing the existing ontology is the best choice. Music is the art of sound, which directly affects human hearing, then enters human thoughts and influences human behavior [10]. In the vast number of music tracks, to find out the valuable tracks and recommend them to related users, we need to explore the internal relationship between users and music tracks and evaluate them. Music can be used to express people's longing for life and future, and also to vent their melancholy, sadness, anger and fear.

Up to now, the quality of music has been greatly improved, and human beings have a wider space to choose music. The importance of music to human life is self-evident. In mobile music applications, search function is a basic function which is widely used. Nowadays, mobile music applications are widely used, and almost all music applications give obvious search entries, which shows the importance of search function. In music search, the high recall rate requires the search engine to return
as many songs as possible, and the search relevance of the results is low. High precision requires that the proportion of songs related to query conditions in the returned songs is as high as possible, and the search relevance of the results is high. High precision shows users results with high correlation with query conditions, but it will affect the number of search results. In the actual implementation process, they are controlled by the search correlation, and cannot have both. Human's choice of music should evolve along the development process from passive to active, from mechanization to humanization, from unpredictable to foreseeable.

When designing a search engine, it is necessary to define the search relevance, compromise the recall rate and the precision rate, and adjust the algorithm in time according to the user's behavior feedback. When analyzing user behavior, it is necessary to combine functions to see if the functions provided by search engines can effectively help users complete the search. By observing the users' search logs and access logs of music applications, we can find that users tend to input fewer words when using the search function, and only a few users will completely input longer query keywords for longer query parts [11]. With the continuous progress of Internet retrieval technology and the emergence and deepening of semantic technology, many research institutions and researchers base music retrieval on content, and there are some retrieval modes such as humming-based music retrieval and music retrieval based on music score. Humming-based music retrieval realizes that users can retrieve related music by humming a piece of music melody through voice equipment, but the accuracy of humming is high. If the user can't sing or can't sing correctly, it is difficult to search.

3. Construction of music search engine system

3.1. Hierarchical search strategy

The natural language information put forward by the user firstly enters the grammar analysis module for grammar analysis, and then the analyzed results are submitted to the semantic analysis module for semantic analysis, and finally the obtained results are displayed on the user browsing interface. Music is an abstract art that can reflect the real life emotion of human beings. Different users have different interests in music. According to the differences of users' emotions, social conditions, regions and personal styles, the ways and habits of listening to music are also very different. All search applications have the same core code, and data access and data search in different channels use different indexes and search configurations to establish a general search framework. This search framework flattens the source data, and can process any type of data. It defines the correspondence between the fields in the source data and the fields in the index from the configuration level, and processes the fields in the source data in the way specified by the configuration and adds them to the index document. Different from reading books and watching movies, users often have a high probability of listening to familiar music that they have heard in a certain period of time. Because it takes a long time to read a book and watch a movie, users generally do not spend it repeatedly in a short time. By contrast, the time cost of listening to music is much lower, and it is very likely that there will be repeated consumption for music that they like a little.

The basic features of music are extracted, and the complex features are analyzed on this basis. According to the basic and complex features of music, the overall features of music are identified. See fig. 1 for specific structure.
In the face of music whose emotional content can't be seen from its name, how to find the music you want to listen to accurately is the most urgent demand of users for music retrieval system. The semantic similarity calculation of words or sentences is carried out when the mapping between keyword sets and concepts in ontology cannot reach the predetermined matching degree, and the mapping matching is carried out by calculating the semantic similarity between keyword sets and individuals in music domain ontology. Semantic-based music retrieval system is to solve the human demand for humanization of music retrieval, improve the intelligence and predictability of music retrieval, and realize the function of music retrieval based on natural language [12]. The technology of each functional module cited by the system is mature, and it is compatible with the traditional music retrieval mode. It has low requirements on the performance of hardware and software, which is conducive to popularization.

In the process of analyzing the user behavior of the music listening log, it is found that many users have distinct characteristics in the way of using music application, some users like to cut songs frequently, while some users like to enter a music list to play automatically after opening the music application. Music domain ontology describes various semantic relationships between concepts in music domain. Key words are mapped into ontology concepts, and semantic relationships are established between key words. For those users with wider interests, it can help users find the types that they have not touched, find out the new categories that users are most likely interested in through the correlation analysis between categories, and train the breadth interest model by taking the characteristics of music in the new categories as input. Sentence semantic similarity calculation is to assist mapping by calculating the semantic similarity between keyword sets and individuals in music domain ontology when the semantic mapping between keyword sets and concepts in ontology fails to reach the predetermined matching degree.

3.2. Integration and search of song extension information
In the actual retrieval process, it is not necessary to match all words with ontology concepts. Combining with the classification of language parts of speech, we can abandon all the function words at first, and then combine with the retrieval of emotional music, so we only need to leave nouns, verbs and adjectives in real words as keywords to build a decision tree model. We need to start from a single root node, and select one of the many attributes in the training data as the test attribute as the splitting standard of the current node according to the quantitative evaluation standard of the selected feature selection. The calculation method of word similarity based on corpus statistics needs to be based on a large-scale corpus, and the accuracy of the calculation results is directly related to the size and perfection of the corpus. After lexical analysis and keyword extraction of the natural language input by the user, a keyword set that can be understood by the machine is obtained, and semantic mapping between keyword set and ontology concept is needed to realize the retrieval query of ontology [13]. When the machine performs this mapping process, it must consider the complexity of natural language. In the semantic-based music retrieval model, the query language input by users is changeable, and the related words will not be input according to the definition of concepts. Therefore, the semantic mapping between query language and concepts will not reach the predetermined matching degree. This requires the semantic similarity calculation and semantic mapping between words or sentences in query language and ontology concepts.

In the historical behavior, users mainly show obvious preference for the music of Hong Kong and Taiwan singers. Besides, there are other kinds of music listening behaviors such as European and American singers or mainland singers. Interest will be used to measure the user's preference for Hong Kong and Taiwan singers and their acceptance of music of other types of singers. The calculation of interest degree is expressed by the following formula:

\[ \text{interestDegree} = \log(1 + \#\text{Cats}) \times \ln\left(1 + \frac{\#\text{Cats}}{\#\text{Songs}}\right) \]  

(1)

Among them, \( \text{interestDegree} \) represents the quantitative value of interest, \( \text{Cats} \) represents all the categories that the user listens to, and \( \text{Songs} \) represents all the music that the user listens to.

The main purpose of music recommendation is to quickly and accurately extract music tracks with relevance from massive data, rate their relevance, and recommend them after matching with music tracks collected in the user's music box. The most intuitive way for music to be displayed in front of human beings is the name of music. Music creators try their best to make the name of music contain the emotion expressed by music. There is a certain semantic relationship between concepts in the ontology of music field. Mapping keywords with ontology concepts makes the semantic relationship between keywords established. By observing and analyzing the user behavior in the data set, it can be found that in the real environment, different users have different activity levels, and the popularity of projects often varies greatly. The recommendation basis of project-based collaborative filtering is that if users are interested in a project, they will be interested in similar projects, and whether similar projects are measured by similarity. This similarity is based on user behavior, which can effectively improve the user experience.

4. Conclusion
In the past ten years, the rapid development of the Internet has brought a lot of convenience to human life. However, at present, the data on the Internet is not only massive, but also growing at a high speed. These information resources have different kinds, different qualities and different structures, which makes it difficult for users to accurately locate the small part of information data they need in such massive information data. Traditional retrieval based on text information is to match the user's retrieval request with the text information marked in music, but the premise of user retrieval is that the user knows nothing about the text information of music. This leads to the contradiction between the user's needs and the existing music retrieval mode. In this paper, according to the needs of most users searching for music, the music search technology is studied, and the music search engine will continue to be optimized and improved in combination with natural language recognition technology in the future to meet the more personalized needs in the mobile Internet era. Users only have two concepts of familiarity and strangeness to music. If it is familiar music, the traditional retrieval method can solve the user's needs. With the rapid development of mobile Internet, the ways for users to find music are increasingly diversified and personalized, and search engines will still serve users as infrastructure.

References

[1] Yang Yiping, Zhang Yuzhong, Yang Haojie, et al. Research and Practice of Search Engine Technology in the Field of Mobile Internet Music[J]. Telecommunications Science, 2015, 31(2):152-157.

[2] Zhang Yuanming, Zhang Yuzhong, Liu Di. The demand growth model of electronic dance music information retrieval based on grounded theory[J]. Library and Information Service, 2018, 062(011):35-42.

[3] Gao Xuan. Research on the effectiveness of search engine-based online marketing and false traffic detection [J]. Business Times, 2020, 000(006):76-78.

[4] Jiang Enbo, Qin Lin. Search engine based on structured data [J]. Modern Information, 2019, 039(002):66-72.

[5] Xu Jicang. Review of Research on Optimizing Music Information Resource Organization and Retrieval[J]. Yunnan Science and Technology Management, 2016, 29(001):55-57.

[6] Xu Yunzhi, Shao Xi. Music label subject retrieval based on MT-LDA[J]. Computer Technology and Development, 2016(7):200-204.

[7] Li Wei, Li Zijin, Gao Yongwei. Understanding Digital Music-A Review of Music Information Retrieval Technology[J]. Journal of Fudan University (Natural Science Edition), 2018, 057(003):271-313.

[8] Wang Hanxi, Chen Chong, Zhang Yanping. The key technologies facing "Internet + digital music"[J]. Journal of Wuhan University of Technology, 2017, 315(03):97-102.

[9] Liu Mingxing. Music classification model based on BP neural network [J]. Modern Electronic Technology, 2018, 41(05):144-147.
[10] Zheng Dan. Multi-feature fusion music classification algorithm based on deep belief network[J]. Electronic Design Engineering, 2020, 28(04):132-136.

[11] Liu Yang. Research on music analysis and auxiliary composition based on complex network [J]. Northern Music, 2016, 36(013):161-162.

[12] Qin Jing, Lin Hongfei, Xu Bo. Music retrieval model based on example semantics[J]. Journal of Shandong University (Science Edition), 2017(06):40-48.

[13] Wang Jianjian. Construction and implementation of content-based ethnic music retrieval model[J]. Electronic Design Engineering, 2016, 24(019):66-68,72.