



Data Construction of Music Culture Communication Based on Data Mining Technology

Sufang Wang^(✉) and Guixiang Lou

Dongying Vocational Institute, Dongying 257091, Shandong, China
wangsuofangyin@163.com

Abstract. In recent years, my country's big data and information technologies such as the Internet have made great progress, and have been widely used in various fields. The application in the field of music communication has greatly impacted the traditional music communication methods and channels. Internet technology has promoted the development of the music industry, and various emerging music styles have emerged one after another, and users' appreciation preferences also vary from person to person. But at the same time, whether it is a folk song or other excellent pieces of traditional music, after being discovered, with the intervention of the media, it will bring a certain amount of hype and commercial purposes, and they will be deleted and modified, and in the end only unrecognizable "boutiques" will be left. The rapid development of society is inseparable from the joint action of many factors. Culture is one of the important factors driving the evolution of social self-organization system. To develop music culture, we should first preserve the original flavor. Only correct preservation can provide a good foundation for all kinds of development in the future, and only better maintaining the tradition can better nationalize. Based on data mining technology, this paper expounds the basic characteristics of music culture in the new media era, and studies the basic strategies of music culture communication and protection in the new media era, in order to promote China's music culture towards sustainable development.

Keywords: Data mining technology · Communication of music culture · Data-based construction

1 Introduction

At present, China has entered the era of big data information, and various new Internet technologies such as big data and cloud computing have revolutionized many aspects of our lives [1]. The inheritance and development of music culture takes place in the process of communication, and communication, as the basis for the survival of music culture, is inseparable from the media at every stage of development [2]. As a classic art form, music has had a profound influence on people's lives. Music can be used as a way of entertainment to relax. It can also be used as a medium to spread culture; Can be

used as a theme to lead the trend of the new era [3]. With the advent of the new media era, the media dominated by digital technology and network technology have played an important role in the inheritance and dissemination of music culture, but the ways and methods of communication have also changed [4]. The inheritance and development of music culture takes place in the process of communication, and communication, as the basis for the survival of music culture, is inseparable from the media at every stage of development [5].

As a music industry that is highly dependent on communication technology, it is one of the first fields to be deeply affected by the communication technology revolution in the era of big data. In addition, music has played an important role in the evolution of human group culture, including the preservation of human knowledge, the promotion of ritual activities and collective behavior, and the promotion of human group cohesion. In particular, digital music, which has gradually developed in recent years, is likely to innovate and change the way and channels of popular music transmission [6]. First of all, the development of digital music brings possibilities for the innovation of communication methods and channels. Digital music stored in digital format is mainly spread through the Internet, and its biggest feature is that no matter how many times the same piece of music is played or downloaded, the quality of the piece will not change [7]. From the perspective of the public, people can no longer enjoy music only through radio and television stations and buying records. They can obtain high-quality music with only a multimedia computer.

The emergence of the new media era has broken the development of traditional music media in China, which is reflected in the time, scope and space of communication. In addition, in the use of new media, media users can send or receive audio, video and text images related to music culture at any time and any place. Music is not only a work of art, but also a cultural consumer goods widely disseminated and consumed by audiences at different levels [8]. This feature of digital music gets rid of the medium carrying music works, and is the necessary basis for it to adapt to the high-speed development of the Internet [9]. In the era of big data, human beings have recorded massive amounts of electronic data through a large number of sensor devices deployed in their daily lives. We use social perception as a research tool to analyze the big data of human social behaviors and interactions, study social culture and its evolution rules, and try to discover human behavior patterns and build explanatory models for them. This feature enables digital music to adapt to the development requirements of the Internet, and completely gets rid of the traditional media that carries music works. In recent years, various music styles emerge one after another, fully demonstrating the personality of music creators [10]. The diversity of music makes music style an important feature of Internet music retrieval, and most people are used to choosing the music they might like according to their favorite style. The openness of new media makes the world a whole, and music cultures from different countries and regions can be displayed on the new media. In the new era when big data and cloud computing become mainstream information processing technologies, the dissemination of digital music needs a new communication model.

2 Outline of Basic Theory

2.1 Overview of Data Mining

In the past century, the Internet has become an important channel for people to store, obtain and transmit information. However, with the development of Internet technology and the improvement of storage media, information explosion is an important problem that can not be ignored. Music platform operators to accurately compared to users recommend music works, you need to pass the identity of the user ID and mobile client download information to the user, often listen to the music, and comment on information behavior related to the real-time collection, on the basis of the filter, filter, after analyzing and mining the data information that exist in the contact and characteristics, In this way, users' behaviors and habits can be fully grasped, and corresponding tracks can be pushed accurately for different end users. Data mining technology is more dependent on the powerful computing power of computers, and the whole process of knowledge discovery is roughly shown in Fig. 1.

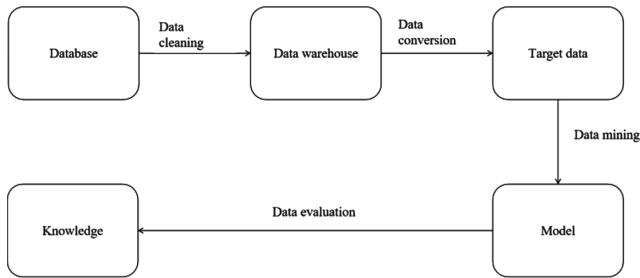


Fig. 1. Process of knowledge discovery

Data on the Internet continues to grow exponentially. New media has various forms of expression and can provide equal opportunities for people to accept music in the development process. To be precise, big data means to collect as many different variables as possible from all aspects to describe various characteristics of people, things, things and phenomena. Each participant can participate in it through different platforms and communicate with each other about music culture, becoming the subject of the new media era. For music perception, we have acquired rich basic cognition in personality influence, social identity, cultural background and cross-cultural influence, and established a relatively complete theoretical framework. In the process of music transmission, music must be connected with human behavior, and the amount of human behavior information data is very large. In the face of the ever-increasing need to record these variables, the number of retrieved data is also increasing rapidly, so the database capacity is getting larger and larger, and finally it is close to infinity. Therefore, a large amount of information will be submerged in the ocean of data. When the traditional media spread music culture, the disseminator is the main center, forming a point-to-point communication mode. In the new media era, the communication of music culture has realized the communication between point and point, point and surface. Processing massive data requires

new technologies and models for storage, calculation and transmission, which brings us many unexpected applications. When the disseminator spreads the basic information of music culture to the network, the receiver will respond in the first time. Besides, it can also forward the basic information of music culture to others, thus fundamentally realizing the basic characteristics of two-way interaction and multiple interactions in the development process of music culture.

2.2 Association Rules Mining Algorithms

Artificial intelligence, statistics and database technology have laid a solid foundation for the birth of data mining technology. Artificial intelligence and statistics are both disciplines that study how to find knowledge from existing data and predict through existing knowledge. The research on the feature mining of cultural products can give full play to the advantages of big data and intelligent analysis means. The process of classification method is to classify the first layer, including melody analysis and harmony analysis. After obtaining the corresponding decision classification, take the results of the first layer classification as the conditions of the second layer classification. Music style classification is a bottom-up process, as shown in Fig. 2.

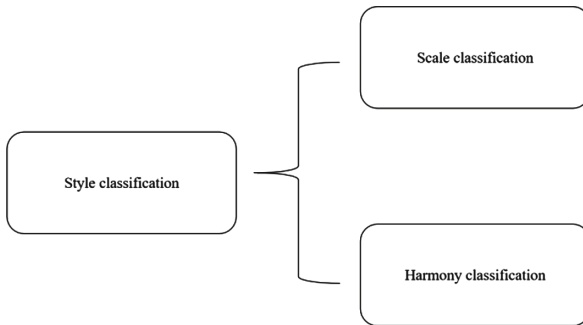


Fig. 2. Music style classification process

Classical algorithm is a classic algorithm for mining frequent itemsets from database. This algorithm connects and prunes the itemsets in database according to the nature. Finally, according to the preset threshold, frequent itemsets and strong association rules can be obtained. Traditional statistics usually adopt the method of sampling statistics. In the era of big data, we analyze the relationship between things by analyzing all existing data related to topics. Now, according to the understanding of the music culture industry, we can know that there is a lack of a unified and standardized cultural system in the process of communication and development of Chinese music culture, which fundamentally restricts the communication of Chinese music culture. Traditional text-based music search has its inherent limitations. Searchers must know the related text content of music works, including the title and introduction of works, while audio-based search and discovery will largely make up for the defects of text search. As an important part of the development of music culture, the music culture market can not

only promote the circulation of music culture products, but also promote the free flow of music culture market. The music culture system has opened up development space for the dissemination of music culture while realizing the optimal allocation of music culture resources.

At this stage, many technologies can carry out audio recognition, but one of the most widely used music retrieval technologies based on fingerprint can effectively extract the fingerprint in the original waveform music, then analyze the extracted fingerprint information and automatically match the music works with the highest similarity. Apriori property is that all non empty subsets of frequent itemsets must be frequent itemsets; The supersets of all non frequent itemsets must be non frequent itemsets. In modern music culture, various social thoughts appear and there are many ideological schools. The advantage of the algorithm is that the algorithm is simple, stable and easy to implement. The disadvantage is that it scans the transaction database too many times and costs a lot.

3 Introduction and Preprocessing of MIDI Music Files

3.1 Introduction to MIDI Files

In this era of Digital music, music file formats are very diverse, and we are familiar with MP3, WAV—like, MID is also a common music file format, its full name is Musical Instrument Digital Interface (Musical Instrument Digital Interface). Music culture, from tradition to modernity, from classics to network, presents the important characteristics of rich variety and long history. It is extremely difficult to spread all music culture. MIDI music files are used as the source data for music style classification analysis, and the reasons are shown in Fig. 3:

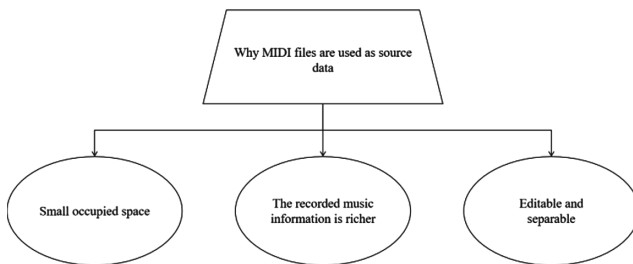


Fig. 3. Why MIDI is the source data

Therefore, when disseminating music culture, it is necessary to take the essence, remove the dross, scientifically sort out, select and refine, spread the ideological value and widely influential music culture in Chinese music culture, and fundamentally occupy the new media The development space, actively create music culture, let people understand music, love music culture. When the total amount of data reaches a certain level, other characteristics of things hidden behind the data will be revealed, and more connections between things can be found. Cloud storage and cloud computing are simple manifestations of cloud computing technology. Users can use the cloud storage service

of cloud music providers to create their own private cloud, then upload all their collected music works to their own private cloud, and then use relevant music playback software to play music on computers and share the music works uploaded to the cloud on mobile phones and other devices, so that users don't need to spend time and effort to store and copy the collected music works, and can give full play to the characteristics of network interaction to the greatest extent. When applied to music style differentiation, each itemset needs to have a decision attribute D . When connecting to generate candidate $K + 1$ itemsets, two frequent K itemsets that can be connected must meet the following requirements:

$$(l_a, l_b) \in \{(l_1, l_2) | \text{diff}(l_1, l_2) = 1, d_1 = d_2\} \quad (1)$$

It is worth noting that the distribution of regional diversity of individual users in Tibet and Xinjiang shows a certain degree of non-monotony in regions with low regional diversity. In the music industry, the data into the hands of the many music company executive, they through a variety of data collection and analysis of comprehensive, not only can clearly understand the grades in each song on flow media service web site, the popularity of the audience, and, more importantly, we collected data through to the large audience feedback information, statistics, analysis and mining, I can clearly grasp the situation of the main audience and the music types they like, and even the listening habits of the audience can be well informed. Considering the correlation between these diversities and the economic development level of each region, it can be found that users in regions with higher economic development level are more likely to prefer minority music with low regional diversity, revealing the relationship between economic development level and the personalized music preference of users.

The MID music file format was put forward in 1980s, which is different from WAV in that MID music file is not a sound waveform signal but a series of instruction signals, such as notes, time values and playing skills, etc. These instructions can directly control MID equipment to perform music.

3.2 Classification Rule Extraction

The extraction of classification rules is to complete the data preprocessing first, and then use the association rules in data mining to classify music styles. In the era of big data, digital processing of music content alone is far from meeting the needs of effective music dissemination, and a more convenient way of management and calculation is needed, that is, dataization. For the active search type Apriori algorithm, its advantages are that it is simple, reliable, and easy to implement. It is a very mature mining algorithm for association rules. In the environment of large transaction databases, the active lookup Apriori algorithm has absolute advantages compared with the matrix-based Apriori algorithm, and the specific advantages are shown in Fig. 4.

In the process of extracting music classification rules using active search Apriori algorithm, it is necessary to set a confidence threshold, extract rules from a frequent itemset, and calculate the confidence of the rules:

$$\text{Confidence} = \frac{\text{support}}{|\text{matched}|} \quad (2)$$

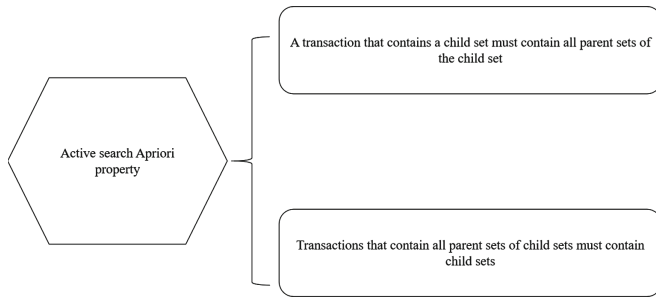


Fig. 4. Properties of active search Apriori

Cloud computing is a kind of pay by usage pattern, this pattern provides available, convenient, on-demand network access, enter the configurable computing resources Shared pool (resources including network, servers, storage, applications and services), these resources can be quick to provide, just in the management of the very few and or little interaction with service providers. At the same time, users can also easily and quickly retrieve music works that meet their requirements through cloud computing, thus effectively saving users' time and energy spent on music retrieval. Kullback-leibler divergence, or relative entropy, used in statistics to calculate the match between two probability distributions, and in machine learning to measure the approximation between two functions, is calculated by the following formula:

$$KL(p\|q) = \sum p(x) \log \frac{p(x)}{q(x)} \quad (3)$$

The active search Apriori algorithm only needs to scan the transaction database once, and when calculating the support of frequent itemsets, it uses the active search strategy to replace the passive scan, which effectively reduces the time complexity. Therefore, we should fully grasp the advantages brought by new media, constantly refine and spread our music culture, and actively push our excellent music culture to the world. In the previous research, the knowledge ontology in the music field has been constructed, in which four sub-ontologies are defined to describe the superior conceptual structure, namely, music events, music works, music figures, music places and music technologies. For each track, we add a unit value to the elements of the track vector corresponding to tags a, b, and c if the playlist that contains tag information (such as tags a, b, and c) contains the track. Overall, the future trend is that cloud computing, as the bottom layer of computing resources, supports the upper layer of big data processing; and the development trend of big data is real-time interactive query efficiency and analysis capabilities. The time complexity of this algorithm and the number of times of scanning the transaction database are lower than those of the algorithm based on transaction compression technology; compared with the matrix-based algorithm, the time complexity is the same, but the space complexity is lower than that of the matrix-based algorithm.

4 Conclusions

The wide application of big data, Internet and other information technologies in music communication, on the one hand, enables us to enjoy music works more conveniently and quickly, on the other hand, it realizes sharing, from music communication methods such as cloud music, online concert and show live broadcast. The rapid development of new technology and its increasingly wide application provide unparalleled imagination for music communication. It brings us not only convenience and quickness, but also “sharing”. This paper combines data mining technology with music theoretical knowledge, takes various elements constituting a complex cultural system as the analysis object, analyzes the regional diversity and Preference Diversity of music, in order to study the cultural structure of China’s online music society, and puts forward a research on the data construction of music culture communication based on data mining technology.

References

1. Ziyue, Z.: Research on music style classification method based on data mining technology. *Nanchang University* **23**(1), 4 (2013)
2. Linlin, Q.: Research on the application of data mining technology in music pop trend prediction. *Southwest University* **25**(2), 8 (2017)
3. Ziyue, Z.: Research on music style classification method based on data mining technology. *Nanchang University* **43**(1), 5 (2014)
4. Ziyue, Z.: Research on music style classification method based on data mining technology. *Nanchang University* **35**(2), 4 (2016)
5. Tingting, Z.: Application of data mining technology based on association rules in music classification. *Modern Electron. Technol.* **46**(7), 6 (2020)
6. Xiao, C., Shuo, C., Min, C.: Research on music curriculum model under cloud computing and data mining technology—comment on algorithmic composition theory and practice. *Chinese Sci. Papers* **14**(7), 1 (2019)
7. Yan, D.: Research on music curriculum model under cloud computing and data mining technology. *Autom. Instrum.* (9), 3 (2018)
8. Ziru, Z., Muxuan, M., Huan, X., et al.: Research and analysis of Chinese folk song industry based on data mining. *Inf. Technol. Inf.* (10), 4 (2019)
9. Xingyu, Y.: Research on music playing data mining based on FP growth algorithm and ARIMA model. *South China Agricultural University* **33**(5), 6 (2017)
10. Yan, L.: Research on fine marketing of music brand value-added services based on data mining. *Pr World* (22), 2 (2022)