ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on Big Data Assisting the Inheritance of Chinese Ethnic Music in Modern College Art Education

Leqiong Wang[1,*]

[1] Hebei University of Technology, Tianjin 300401, Tianjin, China

SUMMARY: *This article's goal is to discuss the effective passing mechanism of Chinese minority music under the environment of big data. Through the combination of Chinese nationality music components with contemporary nationality music market developing tendencies, it innovates the content of Chinese ethnic music inheritance to meet the personalized needs of modern listeners. This article uses an improved hierarchical clustering algorithm to analyze the attribute data of Chinese ethnic music. Firstly, through data cleaning and patching, a network of citation relationships between high-frequency words is constructed. Next, multiple similarity measurement methods such as Min's distance, Jaccard distance, and Mahala Nobis distance were used to generate a similarity matrix for mixed Chinese ethnic music attribute data, and a hierarchical clustering algorithm based on symmetric regularization was used for clustering solution. Research has found that researchers highly focus on themes such as ethnic music culture education, cultural inheritance and development, and often regard ethnic music teaching as an important way to achieve the inheritance and development of ethnic music culture. Meanwhile, through the analysis of high-frequency keywords and co-occurrence matrices in Chinese ethnic music, the research hotspots and relationships between keywords in the field of Chinese ethnic music were revealed. In the context of big data, the integration of main stream ethnic music components, the exploration of special ethnic music styles and expression methods, and the use of data analysis tools to comprehend audience liking are the keys to push the effective handing down of Chinese ethnic music.The improved hierarchical clustering algorithm provides strong data support for the creation of Chinese ethnic music inheritance content, which helps to create more attractive and unique Chinese ethnic music works, further promoting the inheritance and development of Chinese ethnic music.*

KEYWORDS: *aesthetic education; big data; Chinese ethnic music; Music inheritance; Hierarchical clustering; Similarity matrix*

1 Introduction

In modern aesthetic education in universities, the content of ethnic music inheritance is an important component of the mechanism of ethnic music inheritance, which has a significant impact on the effective inheritance of Chinese ethnic music in the new era [1]. In order to establish a long-term inheritance mechanism that meets the basic development requirements of the big data era, it is necessary to first analyze the characteristics and core elements of ethnic music, in order to create ethnic music inheritance content with distinct ethnic characteristics. Ethnic music has regional and traditional characteristics, reflecting the culture and history of different ethnic groups, showcasing unique ethnic music styles and expressions in specific

*18722523301@163.com
https://doi.org/10.65102/is2026567
Published: 30 April, 2026

regions. In China, ethnic music from different regions has significant differences in modes, rhythms, instruments, and plots, such as the Mongolian long tune and the Uyghur plucked string. The region features of these race music to a certain degree recover the surroundings and culture elements that are merged into race music making, thus they reflect the unique traditions and styles of every race group. At the same time, the folk music of ethnic groups, which is one form of expression that belongs to specific social groups, is generally handed down through oral tradition and handed down from one generation to the next generation. It bears the history, culture, and self-recognition of the country's society, hence it is an important method for people to express their emotions, pass on information, hence push forward the national culture. Therefore it thus performs an important role in the passing of thoughts, feelings, and also cultural values. Its expression forms are diverse, mainly including ballads, folk songs, dance music, etc., all of which reflect the ideology, values, and way of life of specific social groups. In the process of interpreting folk ethnic music, through the melody of the song, the poetic meaning of the lyrics, and the emotional expression of the singer, the audience can intuitively feel a deeper level of national values and spiritual pursuits [2]. From this, it can be seen that these forms of ethnic music expression, the ideological and cultural values conveyed, and the cultural connotations of ethnic regions behind them are the creative basis for the inheritance of ethnic music content.

One of the key elements in the construction of the dissemination mechanism of ethnic and folk music in the context of big data is the target audience positioning of ethnic and folk music. By accurately targeting the target audience, it is possible to better disseminate the cultural connotations of ethnic regions expressed by ethnic folk music, and enhance the influence and social acceptance of ethnic folk music. As a specific form of musical art, ethnic and folk music has strong cultural connotations and a unique music system [3]. Local operas such as Peking Opera, Yu Opera, Cantonese Opera, and Sichuan Opera are characterized by melodious melodies and unique singing styles, accompanied by the performance of traditional instruments such as erhu, pipa, jinghu, and gongs and drums; Mongolian throat singing is a folk music form characterized by throat sounds, which creates unique timbre effects by controlling exhaled breath and throat resonance during singing; The Tibetan suona, mainly played with the traditional wind instrument suona, is famous for its unrestrained and passionate musical expression; Hakka folk songs in southern China are usually sung in a choir style, with a simple, lively, and natural style. From this, it can be seen that the music systems contained in different ethnic folk music also have differences in their target audience. Therefore, before targeting the audience, it is necessary to have a deep understanding of the artistic, ethnic, and regional characteristics of different ethnic music, in order to promote it to audiences interested in these characteristics. Due to the varying preferences and acceptance of music among different age groups, the target audience positioning should also take into account age factors. For example, young people may be more receptive to ethnic music works that blend with popular music, while middle-aged and elderly people are more likely to be attracted to traditional instruments and ethnic tonal characteristics. At the same time, regional attributes and cultural backgrounds can also affect the audience's identification and preference for ethnic and folk music. Therefore, in the process of dissemination, precise positioning can be carried out based on different regional and cultural characteristics to promote ethnic and folk music to appropriate target audiences. In the context of big data, it is necessary to provide more possibilities for target audience positioning through digital technology and big data analysis techniques [4]. Shenren digs music listening records, search preferences, user profiles and other data of intended audience groups on the Internet, so as to accurately grasp the interests of self-labeled audience and provide more possibilities for target audience positioning. By utilizing intelligent recommendation algorithms and personalized customization services, ethnic and folk music

works can be recommended to relevant audiences, increasing their interactive participation and deep experience, thereby consolidating the audience base and providing important guarantees for the effective construction of ethnic and folk music dissemination mechanisms.

Under the assistance of large data technology, this paper puts forward a research method for the inheritance mechanism of Chinese national music which adopts an improved hierarchical clustering algorithm. This method can comprehensively put together ethnic music elements, and complement and combine creative inspirations with ethnic area culture features, together with the development tendency of the modern ethnic music market. When we promote the promotion of the quality and influence of ethnic music works, it can more effectively satisfy the individual demands of current audiences, and thus provide an important inheritance content foundation for the building of ethnic music inheritance mechanism under the environment of big data. In the meantime, through making improvement on the hierarchical clustering algorithm for the analysis of massive ethnic music data and connected information, clues and latent topics of creation inspiration are able to be found, and shared elements among different regional and ethnic music can hence be researched. This offers more origins of ethnic music creation inspiration for makers of ethnic music inheritance content, and hence hence creates more attractive special ethnic music works, facilitating the effective inheritance of ethnic music.

2 Related research

Big data technology has promoted the modernization of various industries and fields, and has also innovated the inheritance and development model of traditional culture such as ethnic music, enhancing the mechanism of information inheritance function and helping ethnic music break away from the trend of marginalization. Folk ethnic music is a symbol of the spirit of the nation and the times. Innovative development in the big data environment cannot be subjective and arbitrary. It is necessary to follow its inheritance attributes to develop appropriate connotations and extensions of ethnic music, in order to achieve the expected effect of ethnic music inheritance.

The regional characteristics of ethnic music are prominent, which are not present in other types of ethnic music such as pop music. The origin of ethnic music cannot be separated from human social practice, so local cultural customs, geographical characteristics, and the ways of thinking and behavior of local people are key factors that affect the emergence of ethnic music forms and the development of resources. The generation of ethnic music resources and artistic styles are related to psychological characteristics such as production and labor practices, seeking good fortune and avoiding disasters. Once the form of ethnic music is formed, it has stability. In the process of stable inheritance, ethnic music resources not only deeply reflect the spiritual temperament of the times, ethnic groups, and regions, but also embody unique aesthetic concepts, tastes, and ideals. It can be seen that ethnic music is a symbol with a certain characteristic connotation and special significance in the current environment of big data development.

Ethnic music has iconic artistic connotations and extensions, and the form of folk ethnic music is a core force formed by the integration of form and content of ethnic music, directly influenced by class, ethnicity, and the economic level and aesthetic consciousness of the times. Ethnic music is an important branch of China's excellent cultural system, and there are many factors that influence the formation of ethnic music forms, which are interrelated and permeating. The origin of ethnic music can be traced back to the Tang and Song dynasties. As a product of the fusion of foreign and local cultures, it has formed a form of ethnic music with local characteristics and ancient music heritage [5]. In terms of its mode system, scale form, playing instruments, rhythm and melody, it not only inherits tradition, but also constantly

develops through innovation. Therefore, folk ethnic music resources must have the genetic characteristics of the overall Chinese ethnic culture. In the big data environment, the development of any ethnic culture should be rooted in the context of traditional Chinese culture such as Yanhuang culture, and precisely because of the attribute dependent characteristics of ethnic music, its inheritance dimensions can be expanded in new situations such as big data.

The artistic style of ethnic music is directly influenced by the psychological characteristics, ideological concepts, and aesthetic tastes of the masses. Folk ethnic music with ethnic styles is also accumulated, inherited, and continued through generations in a "collective unconscious" way [6]. The regional, cultural, and psychological characteristics of ethnic music are prominent, reflecting the unique perspectives, creative ideas, and artistic qualities of creators who carefully observe and deeply experience life. The dialect singing style of ethnic music and the performance form of local representative instruments all have a strong local cultural color. In the context of big data, while showcasing ancient music with local cultural elements to the public, Nanyin should also focus on promoting and presenting its role as a spiritual bond for overseas Chinese.

Firstly, ethnic music has diverse forms, which can be classified into folk songs, rap, opera, song and dance, instrumental music, etc. according to the characteristics of Chinese ethnic music forms. The artistic system of Chinese ethnic music has become increasingly perfect in the development of ancient and modern times, and its function of carrying and inheriting information has become stronger, gradually forming forms of ethnic music such as notation, music, and melody [7, 8]. As a legacy of the Jin and Tang dynasties, Nanyin has been influenced by Yuan opera, Kunqu opera, Yiyang opera, Buddhist opera, and local opera in its integrated development, with profound and rich artistic expression. Secondly, there are differences in tonality. The tonality of ethnic music includes Kunshan accent, coconut accent, and leather spring accent, especially in regional ethnic music performances of dialects, which are more unique in terms of lyrics pronunciation and vocal tunes [9]. Finally, there is a diversification of musical instruments. There are various instruments used to perform ethnic music repertoire, including percussion instruments such as drum boards and large and small gongs; The orchestral instruments such as Jinghu, Yueqin, and Xiaosanxian are rich carriers of information inheritance. The rich inheritance of musical instrument information in the big data environment has also deepened the integration of traditional ethnic instruments such as clappers, double bells, and bell ringing with folk music creation, creating more ethnic music works with unique artistic attributes and national temperament characteristics that conform to contemporary public aesthetics.

Ethnic music has a certain sense of time and space, and has a positioning function in the inheritance of big data environment [10, 11]. Ethnic music created by integrating typical modes, dialectal processing, and unique timbres of traditional instruments from different regions highlights the regional texture of the works, enabling the public to quickly recognize the cultural background and style of ethnic music, and associate it with the local people's characteristics such as boldness, unique customs and habits such as worship, and unique lifestyles such as tea cultivation [12].

3 Data sources

At present, the only database in China that can support research on academic citation relationships is the China Social Science Citation Index (CSSCI), also known as the Nanjing University Core Journal (C Journal) database [13]. The footnotes or reference data open to users on its platform lay the foundation for building citation networks among scholars. At the same

time, in terms of the quality of the included literature, CSSCI still holds a leading position in the field of humanities and social sciences, so the conclusions obtained from it as a sample are representative.

This study obtained journal literature data from CSSCI from 2004 to 2024, a total of 21 years, and underwent data cleaning and patching steps. The explanation of the repair situation is as follows: Based on the citation data provided by CSSCI, a citation relationship network between journal articles can be directly constructed, and a citation relationship network between authors can be indirectly generated through the authorship of the articles. However, according to the author's statistics, only about 20% of the citation records in journal literature belong to journal literature, while the remaining 80% belong to book literature. To this end, based on the authorship of the cited books and literature, the citation relationships in this section are extracted and merged with the citation relationships of journal literature to jointly generate a citation relationship network between authors, ensuring comprehensive coverage of the data source. The total sample statistics information is as follows: there are a total of 17704 articles on ethnic music (including literature with the middle figure classification number J6 in non-ethnic music journals), with 7465 authors; Among them, there were 5900 authors in the sample population who had mutual citation relationships, with 55445 citation relationships; Generate a mutual citation network diagram and perform cluster analysis based on the screening criterion of citation frequency ≥ 50 within the group.

The so-called co citation relationship network matrix in this article is also known as the "co citation network/matrix", which means that if two authors (two points) are simultaneously cited by others [14], an association can be constructed between them. This association is a representation of "similarity", that is, if they are frequently cited by more people at the same time, the similarity between the two authors will be stronger. In this regard, whether there is and to what extent there is similarity between many authors is the focus of this article's introduction on the "interrelated approach", which is used to characterize the disciplinary structure. Therefore, in order to capture significant similarity, the top 300 authors with the highest citation frequency were selected and a co citation matrix was constructed among them. Draw a simple community dataset of 300 highly cited authors with research domain similarity (the higher the co citation frequency, the more similar they are). Through this data group, we can only roughly see the distribution and clustering relationships in the fields of Chinese ethnic music history, ethnic musicology, ancient Chinese ethnic music history, modern and contemporary ethnic music, and there are still some rough areas. Due to the frequent co citation relationships and high network density among highly cited scholars, the aim is to use clustering lineages and routing networks for data clustering analysis of co citation networks.

4 Improved Hierarchical Clustering Algorithm

4.1 Similarity measurement

This article uses an improved hierarchical clustering algorithm to analyze the research on big data assisting the inheritance of Chinese ethnic music in modern university aesthetic education. Common similarity measures in hierarchical clustering algorithms include Minkowski distance, cosine similarity, Mahalanobis distance, Jaccard distance, Hamming distance, etc. [15]. For the sample dataset $X=\{x_1,x_2,\cdots,x_n\}$, n represents the number of samples, and $x_i=(x_{i1},x_{i2},\cdots,x_{im})$ has a sample dimension of m. For any two points $x_i,x_j\in X$, the common measurement methods are as follows.

(1) Min's distance. The Min's distance can treat all attribute dimensions of Chinese ethnic music equally, without considering the distribution characteristics of various Chinese ethnic

music attributes, which may have a negative impact on the clustering results. The calculation form is as follows [16]:

$$d\left(x_i, x_j\right)=\left(\sum_{q=1}^{m}\left|x_{iq}-x_{jq}\right|^{p}\right)^{\frac{1}{p}} \tag{1}$$

(2) Jaccard distance. The similarity between two sets of Chinese ethnic music attributes A and B is defined as [17]:

$$J(A,B)=\frac{|A\cap B|}{|A|+|B|-|A\cap B|} \tag{2}$$

Due to the fact that this similarity measurement method focuses more on whether the common features of Chinese ethnic music attribute samples are consistent, and cannot measure the specific size of differences, it is mainly used to calculate the similarity of binary attribute objects of Chinese ethnic music and is often used for calculating the similarity of Chinese ethnic music text attributes.

(3) Mahalanobis distance: It is an effective method for calculating the similarity between two unknown Chinese ethnic music attribute sample sets. Unlike Euclidean distance, it takes into account the connections between various characteristics and is scale independent, i.e., independent of the measurement scale. Its definition is as follows [18]:

$$d_m\left(x_i, x_j\right)=\sqrt{\left(x_i-x_j\right)^T \Sigma^{-1}\left(x_i-x_j\right)} \tag{3}$$

where, Σ^{-1} represents the inverse matrix of the covariance matrix of Chinese ethnic music attribute samples. This similarity measure can eliminate the impact of significant differences in scale between different Chinese ethnic music attributes and overcome the disadvantage of multiple linear correlations in Chinese ethnic music attribute data. However, the inverse matrix of the covariance matrix may not exist, and other methods need to be used for processing.

The above measures of similarity in Chinese ethnic music attributes have their own advantages and disadvantages, and are only suitable for measuring the similarity of single type data. It is difficult to achieve good results for mixed type attribute data.

4.2 Hierarchical clustering algorithm

Hierarchical clustering algorithm is a clustering algorithm based on similarity theory [19]. It transforms the clustering problem of the original Chinese ethnic music attribute data $X=\{x_1, x_2, \cdots, x_n\}$ into an optimal similarity partitioning problem. Any Chinese ethnic music attribute data sample x_i is regarded as a vertex $v_i \in V$ in the graph, and the points are connected by weighted edges. That is, for $v_j, v_i \in V$, $e_{ij} \in E$ represents the edge between these two vertices, and $w_{ij} \in W$ represents the weight of edge e_{ij}, thus obtaining an undirected graph $G=(V,E)$. The edge weights between points that are farther apart are low, while the edge weights between points that are closer together are high. According to a certain division criterion, the original graph G formed can be sliced to obtain k subgraphs $g_1, g_2, \cdots, g_k$ with empty intersections. By minimizing the sum of edge weights between different subgraphs g_i and g_j and maximizing the sum of edge weights within each subgraph, the final clustering goal can be achieved.

The objective function of the symmetric regularized hierarchical clustering algorithm for

Chinese ethnic music attributes is [20]:

$$f = \min_F tr\left(F^T L_{sym} F\right),\ s.t.\ F^T F = I \tag{4}$$

where, $F \in R^{n \times k}$ contains k orthogonal vectors; L_{sym} is the Laplacian matrix of symmetric regularization.

The hierarchical clustering algorithm inherits the traditional clustering algorithm ideas and combines similarity optimization theory, greatly improving the universality of the clustering algorithm. On the one hand, the clustering effect of hierarchical clustering algorithm largely depends on the Laplacian matrix generated by similarity measurement. On the other hand, hierarchical clustering performs well on numerical data, but in practical applications, many datasets are mixed types. Therefore, it is necessary to design more suitable similarity measurement to make hierarchical clustering perform better. This article improves the clustering performance of hierarchical clustering algorithm for mixed Chinese ethnic music attribute data through the design of similarity measurement in the framework of hierarchical clustering algorithm.

4.3 Generation of Similarity Matrix for Mixed Attribute Data

The similarity matrix between the categorical Chinese ethnic music attribute data and the numerical Chinese ethnic music attribute data generated by the formula can be used to construct a similarity matrix for the overall mixed attribute Chinese ethnic music attribute data. Considering that the number of attributes in categorical data is different from that in numerical data, it is unreasonable to simply add the similarity matrices of these two types of data, and Gaussian kernel functions are not used to measure the similarity of categorical data. Therefore, two coefficients α and γ are introduced, and the proportion of the number of attributes contained in categorical data and numerical data in the total number of attributes is taken as α and $(1-\alpha)$, respectively. The calculation formula is:

$$\alpha = M_s / M \tag{5}$$

As a control factor for the similarity of categorical data, the similarity of categorical data can be scaled to a certain extent. The design of this article is as follows:

$$\gamma = e^{-\frac{\max\left(S^*\right)-\min\left(S'\right)}{2\sigma^2}} \tag{6}$$

where, $\max(S'')$ represents the maximum value among all elements in the similarity matrix of the generated numerical Chinese ethnic music attribute data; $\min(S')$ represents the minimum value among all elements in the similarity matrix of the generated categorical data; The difference between the two represents the maximum difference in similarity measurement between the two forms of calculation. As a molecule, the larger the difference between the two, the smaller the parameter value, and the impact of not using Gaussian kernel function on categorical data will also be reduced; On the other hand, in the process of actual code operation experiments, it was found that the difference between the global maximum value of the numerical Chinese ethnic music attribute similarity matrix and the global minimum value of the categorical data similarity matrix was slightly better than the difference between the global minimum value of the numerical Chinese ethnic music attribute similarity matrix and the global maximum value of the categorical Chinese ethnic music attribute data similarity matrix. The

parameter σ has the same value as in formula (2) [21]. Used to mitigate the impact of not using Gaussian kernel function in similarity calculation of categorical data. This article defines the similarity measure between x_i and x_j as $S\left(x_i, x_j\right)$, and the calculation formula is:

$$S\left(x_i, x_j\right) = \alpha \cdot \gamma \cdot \frac{\delta_{i,j}}{\left|x_i^{(s)} \cup x_j^{(s)}\right|}(1-\alpha) \cdot e^{-\frac{\left(\alpha_i^{(t)} - x_j^{(t)}\right)^T vzu\, T^T \left(x_i^{(t)} - x_j^{(t)}\right)}{2\sigma^2}} \tag{7}$$

The hierarchical clustering algorithm represents the attribute data samples of Chinese ethnic music in Figure $G = (V, E)$, where V represents the node set, E represents the edge set, and the weights of edge e_{ij} are represented as w_{ij}. The construction of similar graphs in this article adopts a fully connected approach, represented as:

$$w_{ij} = \begin{cases} w_{ij} = \alpha \cdot \gamma \cdot S'\left(x_i^{(s)}, x_j^{(s)}\right) + (1-\alpha) \cdot S''\left(x_i^{(t)}, x_j^{(t)}\right), (i, j) \in E \\ 0, (i, j) \notin E. \end{cases} \tag{8}$$

The similarity matrix S of the mixed Chinese ethnic music attribute data constructed from this is:

$$S = \left[\alpha \cdot \gamma \cdot S'\left(x_i^{(s)}, x_j^{(s)}\right) + (1-\alpha) \cdot S''\left(x_i^{(t)}, x_j^{(t)}\right)\right]_{n \times n} \tag{9}$$

Based on the obtained similarity matrix S of Chinese ethnic music attributes and the generated degree matrix D, due to the $W = S$ in this article, the symmetric regularized Laplacian matrix obtained is [22]:

$$L_{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = D^{-\frac{1}{2}}(D - S)D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} S D^{-\frac{1}{2}} \tag{10}$$

Based on the Laplacian matrix of Chinese ethnic music attributes formed in this article, symmetric regularized hierarchical clustering is used for clustering solution. The objective function of clustering is:

$$f = \min_F tr\left(F^T L_{sym} F\right),\ s.t.\ F^T F = I \tag{11}$$

where, $F \in R^{n \times k}$ contains k orthogonal vectors $f_1, f_2, \cdots, f_k$, $f_k \in R^n$; L_{sym} is the Laplacian matrix of symmetric regularization. For the solution of the objective function of Chinese ethnic music attributes, the result obtained by calculating the eigenvectors corresponding to the first k smallest eigenvalues of $L_{sym} = D^{-1/2} L D^{-1/2}$ is the optimal solution. Then, the c-means algorithm is used to convert the obtained real valued matrix F into a discrete cluster indicator matrix to obtain the clustering result. For the selection of k value, it is determined based on the traditional intrinsic gap method.

4.4 Algorithm Implementation Process Steps

The algorithm implementation flowchart is shown in Figure 1.

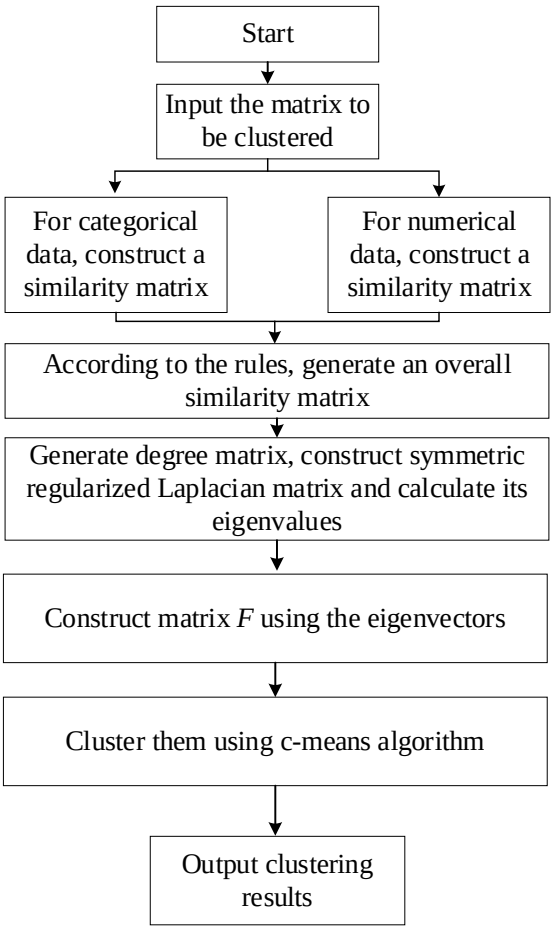

Figure 1: Flow Chart of Improved Hierarchical Clustering Algorithm

The specific implementation process of improving hierarchical clustering algorithm is as follows:

Step 1: Enter the Chinese ethnic music attribute dataset $X = \{x_1, x_2, \cdots, x_n\}$ to be clustered.

Step 2: For categorical Chinese ethnic music attribute data, use the improved Jaccard similarity to construct the similarity matrix S' for categorical data. For numerical Chinese ethnic music attribute data, use the Mahalanobis distance Gaussian kernel similarity calculated using pseudo inverse to construct the numerical data similarity matrix S''.

Step 3: Based on the similarity matrix of Chinese ethnic music attributes and numerical similarity matrix, obtain the similarity matrix S of the overall mixed attribute data according to the rules given in this article.

Step 4: Based on the similarity matrix S of Chinese ethnic music attributes, a degree matrix D is generated, and a Laplacian matrix $L_{sym} = D^{-1/2} L D^{-1/2}$ is constructed to calculate the eigenvalues of L_{sym}. Using the intrinsic gap method, the position where the first extreme intrinsic gap appears is taken as the good value of the number of clusters.

Step 5: Construct $F \in R^{n \times k}$ using the feature vector $f_1, f_2, \cdots, f_k$ corresponding to the first k smallest Chinese ethnic music attribute feature values of L_{sym}.

Step 6: Regularize each row of matrix F so that the sum of squares of the Chinese ethnic music attribute elements in each row is 1.

Step 7: Treat each row of matrix F as a data point and use the c-means algorithm to cluster F.

Step 8: Output the clustering results of Chinese ethnic music attributes.

4.5 Analysis of Algorithm Time Complexity

For the Chinese ethnic music attribute dataset with n sample sizes and k cluster numbers, the time complexity of the improved hierarchical clustering algorithm consists of 7 parts:

1) Construct a similarity matrix S' for categorizing Chinese ethnic music attribute data;

2) Construct a similarity matrix S'' for numerical Chinese ethnic music attribute data;

3) Construct a similarity matrix S for the overall mixed Chinese ethnic music attribute data;

4) Calculate the eigenvalues of the Laplacian matrix L_{sym} of Chinese ethnic music attributes;

5) Construct the feature vector matrix F of Chinese ethnic music attributes;

6) Regularized Chinese ethnic music attribute matrix F;

7) Use the c-means algorithm for clustering Chinese ethnic music attributes.

5 Research Results and Analysis

High frequency keywords are currently the most popular research method in the field of scientific research [23]. From a statistical perspective, high-frequency keywords of Chinese ethnic music attributes can reflect the core content of this research field. It is a collection of a large number of research results in a certain research field during a certain period of time, which helps to determine the development background, hot topics, and trends of this field.

This study obtained 2132 keywords (including repeated words) from 463 papers, and after classifying and merging the keywords, sorted by word frequency, a total of 982 categories of keywords were obtained. The basis for determining the high-frequency words of Chinese ethnic music attributes in this study is the Price formula [24, 25] ($N_{\min} = 0.749\sqrt{N_{\max}}$, $N_{\min}$ represent the minimum threshold frequency of high-frequency words, $N_{\max}$ represents the maximum frequency of keyword words). Using the Price formula, the frequency threshold of high-frequency words of Chinese ethnic music attributes in this study was calculated to be 8.3 (0.731/157). Therefore, this study used the keyword frequency "11" as the minimum threshold and generated a total of 23 high-frequency keywords. The high-frequency words in the research field of "Chinese ethnic music education" are shown in Table 1 for statistical data.

As shown in Table 1, excluding literature search terms, the top 10 high-frequency keywords of Chinese ethnic music attributes are distributed as follows: ethnic music culture education (105 times), ethnic music culture (45 times), inheritance and development (62 times), ethnic music education (54 times), ethnic music (32 times), ethnic (24 times), rights holders (21 times), teaching (18 times), Chinese nation (16 times), and Chinese ethnic minority music (15 times). From the current high-frequency words, it can be preliminarily inferred that researchers highly identify with Chinese ethnic music as a cultural phenomenon, focusing on themes such as Chinese ethnic music culture education, cultural inheritance and development, and emphasizing the exploration of cultural rights subjects in ethnic music, with a greater emphasis on themes related to Chinese ethnic music.

Table 1: High frequency Vocabulary in the Research Field of Chinese Ethnic Music Education

Serial Number	Key fields	Frequency of occurrence	Percentage%	Cumulative percentage%
1	Chinese folk music	159	7.715	7.718
2	Ethnic music and cultural education	104	4.004	12.717
3	Ethnic music culture	62	2.945	14.678
4	inheritance and development	45	2.627	18.295
4	national music education	45	2.464	20.861
6	folk music	32	1.427	22.382
6	nationality	24	1.198	23.460
8	Subject of Rights	21	1.004	24.464
9	teaching	18	0.865	24.437
10	Chinese nation	16	0.663	26.191
11	Chinese Ethnic Minority Music	16	0.661	26.944
12	Traditional Mongolian Ethnic Music Culture	15	0.612	26.662
13	Ethnic music teaching	15	0.613	28.381
14	ethnomusicology	15	0.662	29.045
14	Primary and secondary schools	13	0.621	29.662
16	Ethnic Music Education in Normal Universities	12	0.622	30.294
16	ethnic minority	16	0.623	30.905
18	Multicultural	12	0.462	31.487
19	Sight singing and ear training	17	0.461	32.049
20	Ethnic minority music education	12	0.462	32.627
21	Ethnic music education in universities	10	0.424	33.148

In order to better explore the relationship between high-frequency words of Chinese ethnic music attributes, a "matrix" function was used in Bicomb software to construct a "high-frequency word ≥ 10" Chinese ethnic music attribute discourse matrix (see Table 2) and a co word matrix (see Table 3).

According to the results, the "dissimilarity matrix" of Chinese ethnic music attributes is represented by a number from 0 to 1, and the size of the number indicates the difference between phrases (the distance difference between two high-frequency words of Chinese ethnic music attributes). The closer the keywords are to 1, the greater the dissimilarity (the farther the relationship), and vice versa, the greater the similarity of the keywords (the closer the relationship). According to the matrix, the order of distance between each keyword and "Chinese ethnic music" from near to far is: ethnic music education (0.749) teaching (0-787) inheritance and development (0.834) ethnic music culture education (0.888) ethnic music culture (0.913) Chinese nation (0.924) ethnic (0.956) rights holder (0.974) ethnic music (0.976). From this, it can be seen that researchers often regard the teaching of Chinese ethnic music as an important way to achieve the inheritance and development of ethnic music culture, and often discuss the relationship between ethnic music, culture, and rights holders.

Table 2: Word Matrix (Partial)

Entry	L1	L2	L3	L4	L4	L6	L7	L8	L9	L10	L11
Chinese folk music	1	1	1	1	1	1	1	1	1	1	1
Ethnic music and cultural education	0	0	0	1	0	0	0	0	0	0	0
Ethnic music culture	0	0	0	1	0	0	0	0	0	1	0
inheritance and development	0	0	0	1	0	0	0	0	0	1	0
national music education	0	0	0	1	0	0	0	0	0	1	0
folk music	0	0	0	1	0	0	0	1	0	1	0
nationality	0	1	0	1	0	0	0	1	0	0	0
Subject of Rights	0	1	0	0	0	0	0	1	0	0	0
teaching	0	1	0	0	0	0	0	1	0	0	0
Chinese nation	0	1	0	0	0	0	0	1	0	0	0
Chinese Ethnic Minority Music	0		0	0	0	0	0	0	0	0	0

Table 3: Common Word Matrix (Partial)

Entry	Chinese folk music	Ethnic music and cultural education	Ethnic music culture	Inheritance and development	National music education	Folk music	Nationality	Subject of Rights	Teaching	Chinese nation	Chinese Ethnic Music
Chinese folk music	149	16	12	14	28	4	4	44	44	3	10
Ethnic music and cultural education	16	112	13	18	8	11	2	68	4	12	2
Ethnic music culture	11	16	62	12	18	16	6	8	6	9	2
Inheritance and development	18	12	12	48	13	3	1	1	9	3	
National music education	23	21	13	13	44	9	1	1	1	11	16
Folk music	12	11	1	2	2	31	9	11	11	3	
Nationality	42	1	2	1	4	8	28	22	3	9	3
Subject of Rights	14	2	3	4	2	4	22	22	11	11	13
Teaching	12	0	3	1	8	8	8	14	16	8	18
Chinese nation	1	12	4	4	1	8	2	2	14	19	14
Chinese Ethnic Music	1	1	1	4	8	1	6	1	14	1	19

The theme includes 11 high-frequency words of Chinese ethnic music attributes, including ethnicity (9), rights holders (10), folk songs (25), ethnic music culture education (4), ethnic music (8), schools (26), ethnic music culture (8), primary and secondary schools (14), Chinese nation (18), higher education ethnic music education (15), and ethnic music teaching (13). The subject of rights is a legal term that refers to the parties who enjoy rights and assume obligations in tax legal relationships. In the protection of intangible cultural heritage of ethnic music culture, it is used to explore the identity of the owner and subject of ethnic culture, aiming to strengthen the subject consciousness and cultural consciousness of ethnic music culture. In the process of creating and inheriting ethnic music culture, the rights holders of culture belong not only to the state, but also to groups and individuals. As the rights holders of ethnic culture, it is necessary to strengthen their rights attributes and construct subject consciousness in order to achieve self-

awareness and self-consciousness of their own ethnic culture. Research has pointed out that traditional ethnic music is deeply rooted in the lives and history of our ancestors, reflecting various aspects of ethnic customs, characteristics, ideology, aesthetic psychology, and moral ethics, collectively forming traditional ethnic culture. In the context of the development of Chinese ethnic music culture education, more emphasis should be placed on the subject consciousness of ethnic music culture (creative subject and inheritance subject), and the identity recognition of ethnic music culture should be strengthened.

School education is an important place for realizing the recognition and return of the subjectivity of the inheritance of Chinese ethnic music [26, 27]. As an excellent traditional culture of the Chinese nation, ethnic music (especially folk ethnic music) originates from the national land and belongs to various ethnic groups. Through school education, ethnic music culture can be continued. In the system of ethnic music culture education, basic education (primary and secondary schools) is the main battlefield, and normal universities are the key areas. Normal universities serve as the "working machine" and "cradle" for cultivating primary and secondary school education teachers, and are an important base for reconstructing the "main value" of Chinese ethnic music culture. Currently, there are still some problems in school education (especially normal universities) in China, such as insufficient curriculum design and lack of a unified and comprehensive ethnic music curriculum system [28]; low comprehensive quality and ability of ethnic music education teachers; insufficient student interest; research mainly focuses on theoretical research, lacking practical research on specific teaching links. In the proposed countermeasures, it is pointed out that ethnic music should be regarded as a cultural phenomenon and cultural heritage, and should be viewed from the perspective of basic ethnic music. Starting to take education seriously (connecting normal universities and primary and secondary schools) Conduct in-depth and detailed research and discussion on the content and aesthetic taste of ethnic music textbooks, in order to gradually establish and improve the school (ethnic) ethnic music education system.

6 Conclusion

In the current era of booming big data development, in order to promote the high-quality and sustainable inheritance of ethnic music, we should actively draw on mainstream ethnic music elements favored by the public, deeply explore and shape unique ethnic music styles and expression techniques, innovate and expand the essence of ethnic music inheritance, and enhance its appeal to the audience. In this course, under the assistance of data analysis methods, we can accurately know the audience's hobbies and inclinations to ethnic music, and flexibly modify the direction and thoughts of ethnic music inheritance content making according to this feedback message. As an artistic treasure that bears a long history and deep cultural deposit, ethnic music frequently displays its artistic charm through many different music forms and expressing skills. The era is continuously developing, and the masses' aesthetic requirements for mainstream ethnic music are also in a dynamic changing situation. According to this circumstance, when we make a unique style of ethnic music, we have to keep looking for new methods of explanation and manifestation, hence deeply put innovative ideas into the creation system of traditional ethnic music. By integrating ethnic music elements from different regions and standing at the height of international cultural exchange, promoting deep cross-cultural dialogue and integration, ethnic music can radiate a distinct contemporary atmosphere. In this way, it can not only conform to the mainstream values of modern audiences, but also effectively promote the inheritance and innovative development of ethnic music. At the level of expression techniques, creators can cleverly integrate diverse cultural emotional elements into ethnic music creation through modern lyric creation, melody arrangement, and sound effect processing,

helping audiences to better understand and truly feel the core essence of ethnic music inheritance.

Marxism has long pointed out that "human beings are the sum of social relations." In this sense, humans are both the subject and an important component of culture. China has abundant ethnic music and cultural resources. According to incomplete statistics, the ethnic music heritage compiled from 1979 to 2009 alone is shocking, including over 400000 folk songs, 17636 ethnic songs and dances, 341 types of Quyi music, 317 types of opera music, more than 500 types of ethnic musical instruments (75000 instrumental pieces), and over 3000 pieces of Guqin music. These heritage creations are jointly completed by different ethnic groups, groups, and individuals in China through long-term production and life. Therefore, the cultural identity of ethnic music should also be reflected in the entire scope of the country, groups, and individuals. In the fields of music anthropology and ethnomusicology, the viewpoint of "music as culture" or "music within culture" has been proposed. Ethnic music, as an important cultural carrier of our nation, has highlighted its rich historical, morphological, and cultural connotations. As a cultural educational concept, ethnic music should be an important focus of ethnic music teaching, which has been responded to by the music education community in China. Professor Guan Jianhua, an advocate of music education cultural philosophy, once pointed out that "contemporary music education focuses on music communication and mutual understanding between people and between subjects, rather than a fixed learning mode of subject to object aesthetics. Understanding music is about understanding the existence of music (people and culture, people and the world, people and society) rather than just understanding the" works "in existence. Contemporary music education focuses on the real-life world and cultural representations of humanity (including the spiritual world)." Chinese ethnic music has dual values of aesthetics and culture, and the process of ethnic music education needs to highlight them. The main identity of the "self" in ethnic music culture, Both educators and learners should participate in the experience and practice of ethnic music as the main cultural identity, so as to truly feel and express it. Ethnic music teaching should also experience the charm of ethnic music beauty in culture, feel the profound connotation of ethnic culture through a unique aesthetic process, and ultimately achieve the "cultural education" function of ethnic music education.

According to this background, for the sake of effectively push forward the high-efficiency and lasting handing down of nationality music, we have the necessity to deeply carry out exploration and fully make use of the technical dividends which are given by the era of big data. To speak specifically, advanced data analysis apparatuses can be utilized to accurately carry out analysis on the preference tendencies of audiences, and on the basis of these viewpoints, creative strategies can be adjusted in a timely way for the making of ethnic music excellent works that combine modern style and unique charm. At the practical level, we can use big data analysis tools to extensively collect and deeply analyze the ethnic music access behavior data of massive audience groups. By carefully mining these data, we can clearly grasp the differences in audience preferences for different ethnic music genres, styles, and themes, and then make more refined divisions of the audience group. Such data support undoubtedly provides a solid and powerful basis for ethnic music creation, enable our work to more accurately satisfy the requirements of various audience groups and grasp the chance in violent market competition. In addition, we must closely carry out monitoring on audience's feedback and preferences, and therefore make timely adjustment and optimization for ethnic music works on the basis of this. By this method, we are able to guarantee that ethnic music creations forever keep high consistency with the main current aesthetic and individual demands of the nowadays audiences, thereby promoting their wider and more effective inheritance in the era of big data.

About the Author

Leqiong Wang EMBA master, associate professor of vocal music of Hebei University of Technology, member of China Musicians Association, member of Tianjin Musicians Association, director of the Chorus Special Committee of Tianjin Music Association, vice president of the Chinese Minority Vocal Society. Graduated from Tianjin Conservatory of Music, she studied opera and opera in the Department of Drama and Opera Performance of the College of Music of DIT University of Ireland.

Achievements and contributions

1. In 2002, he won the top ten in the Tianjin District and the National American Voice Group of the Second "Hope Star" Competition.

2. During her study in Ireland, she won the first prize in the Mozart, Verdi Vocal Competition and the "Art Ship" Cup Art Song Competition.

3. On June 9, 2005, she successfully held a solo concert at the Dublin City Hall in Ireland, becoming the first Chinese to hold a concert in the hall, which was strongly supported by the Dublin City Government, the Chinese Embassy in Ireland and other departments and news media.

4. In 2016, she went to Astrakhan, Russia to participate in the China-Europe Art Festival and won the silver prize in the vocal competition.

5 In 2019, she won the gold prize in the vocal competition at the China-Europe Art Festival in Frankfurt, Germany, and was invited to serve as a judge of the art festival chorus competition.

Concerts and performances

1. After returning to China, she held many solo concerts in Tianjin.

2. Invited to participate in the 20th anniversary concert of the Tianjin Symphony Orchestra.

3. Recording of the theme song of the Tianjin Spring Festival Gala and various large-scale parties.

4. In 2019, she was invited to perform at the Golden Hall in Vienna, Austria.

Education and contribution

1. Served as the vocal art director and permanent conductor of the "Smile" choir of Hebei University of Technology. The choir has won more than 50 first and second prizes in various vocal competitions held nationwide, Tianjin and Hebei provinces.

2. Students who have been guided have been admitted to excellent institutions at home and abroad in important examinations such as art examinations and vocal postgraduate examinations.

3. Hold many open vocal classes to share your singing experience and teaching experience.

4. Served as a judge of various vocal competitions for children, primary and secondary schools and colleges for a long time.

References

[1] Li, J., Onlamul, K., You, L., et al. (2024). Transmission guidelines for literacy studies in Si Zhu Yue traditional Chinese folk music. International Journal of Education and Literacy Studies, 12(2), 80-86.

[2] Li, Y., & Chaiyason, N. (2024). Chinese folk music literacy: A study on the Liu Sanjie Ge Yao folk music in Guangxi Zhuang Autonomous Region. International Journal of Education and Literacy Studies, 12(4), 143-150.

[3] Qiu, L., & Wan, X. (2024). Nature's beauty versus urban bustle: Chinese folk music influences food choices by inducing mental imagery of different scenes. Appetite, 199, 107507.

[4] Yan, C., Pattananon, N., & Teeravut, P. (2024). Chinese folk music in higher education in China. Journal of Modern Learning Development, 9(4), 568-576.

[5] Zhang, J. (2024). The impact of Chinese folk music on the mental health and academic motivation of college students. Current Psychology, 43(15), 13339-13348.

[6] Guo, H., Sensai, P., & Karin, K. (2025). Educational perspectives on the literacy of Chinese folk music recomposed for saxophone. International Journal of Education and Literacy Studies, 13(1), 96-103.

[7] Liu, S., & Chuangprakhon, S. (2024). Reviewing the current situation of Huadengxi Chinese folk songs in Yunnan Province. Multidisciplinary Reviews, 7(5), 2024109.

[8] Zhang, X., & Sornyai, P. (2024). Chinese folk songs literacy in the context of Meihua Dagu in Tianjin, China. International Journal of Education and Literacy Studies, 12(4), 116-124.

[9] Yao, Z., Li, R., & Hartanto, Y. (2024). Chinese folk songs can facilitate Chinese language learning–A pilot study. Journal of Psycholinguistic Research, 53(6), 72-83.

[10] Beibei, H. (2024). Research on Jiangnan folk music elements in contemporary ethnic instrumental music. Development, 6(4), 62-70.

[11] Wu, Y., & Chuangprakhon, S. (2024). Preservation and transmission guidelines of Salar vocal folk music. International Journal of Education and Literacy Studies, 12(1), 186-192.

[12] Wang, Y., & Thotham, A. (2024). Literacy transmission of Chinese folk songs in Southern Shaanxi. International Journal of Education and Literacy Studies, 12(3), 143-149.

[13] Yongjie, X., & Meesorn, P. (2024). The constructing the Mongolian short songs folk music for saxophone quartet guidebook for teaching first-year students at Wuhan Conservatory of Music in the People’s Republic of China. Journal of Roi Kaensarn Academi, 9(7), 171-181.

[14] Jialun, L., & Jierong, W. (2024). The literatures on Erhu education and Han folk song in China. Journal of Roi Kaensarn Academi, 9(1), 689-696.

[15] Gao, L. L., Bien, J., & Witten, D. (2024). Selective inference for hierarchical clustering. Journal of the American Statistical Association, 119(545), 332-342.

[16] Chen, Y., Shi, J., Lu, C., et al. (2024). Hierarchical clustering-based collapse mode identification and design optimization of energy-dissipation braces inspired by the triangular resch pattern. Journal of Structural Engineering, 150(5), 04024037.

[17] Lopez-Rodulfo, I. M., Tsochatzis, E. D., Stentoft, E. W., et al. (2024). Partitioning and in vitro bioaccessibility of apple polyphenols during mechanical and physiological extraction: A hierarchical clustering analysis with LC-ESI-QTOF-MS/MS. Food Chemistry, 441, 138320.

[18] Benevento, A., & Durante, F. (2024). Correlation-based hierarchical clustering of time series with spatial constraints. Spatial Statistics, 59, 100797.

[19] Li, S., Lu, L., Liu, Q., et al. (2025). Analysis of deep non-smooth symmetric nonnegative matrix factorization on hierarchical clustering. Applied Intelligence, 55(6), 470.

[20] Matos, J. P., Guilherme, L. Q., da Encarnação, S. G. A., et al. (2025). Influence of sleep quality on recovery and performance in endurance and ultra-endurance runners: Sex differences identified through hierarchical clustering. Healthcare, 13(7), 812-829.

[21] Song, Z., Li, X., & Wang, Y. (2025). Hierarchical clustering-based optimal selection of ground motion records for nonparametric estimating method of structural seismic fragility. Structures, 77, 109130.

[22] Sveshnikov, S., Bocharnikov, V., & Mudrak, Y. (2024). Evaluation and choice of an anti-aircraft missile system under uncertain conditions based on fuzzy-integral calculus and hierarchical cluster analysis. Operations Research and Decisions, 34(2), 135-161.

[23] Zhang, R., & Pu, S. (2024). On the religious philosophy and mysticism elements in Chinese folk dance. Trans/Form/Ação, 47, e02400252.

[24] Pan, X., Maneewattana, C., & Ren, X. (2024). The influence of computer music on Guangxi Zhuang folk songs. Journal of Dhamma for Life, 30(4), 141-154.

[25] Fan, C., & Chuangprakhon, S. (2024). Literacy preservation and transmission of Youyang Chinese folk songs. International Journal of Education and Literacy Studies, 12(2), 118-124.

[26] Zou, W., Hin-on, T., & Sapaso, P. (2024). Implementing practices for the Chinese nationalized piano teaching model in undergraduate music education. Cogent Education, 11(1), 2374173.

[27] Lei, L. (2024). The latest technological developments in Chinese music education: Motifs of national musical culture and folklore in modern electronic music. Education and Information Technologies, 29(9), 10595-10610.

[28] Long, R., Rattachaiwong, N., & Liu, L. (2024). Research on the adaptation and performance communication platform for Chinese Maonan folk songs. Journal of Dhamma for Life, 30(3), 102-118.