

# *Research on music influence based on TOPSIS Model*

**Yanfeng Li**

*College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China*

**Keywords:** entropy weight method, TOPSIS comprehensive weighting method, cluster analysis, time series

**Abstract:** In this paper, based on the analysis of data using data mining technology, based on the evolution of music with events and the influence of artists, the following discussion is carried out. In this paper, a graph model based on TOPSIS comprehensive weight determination method is established. Based on entropy weight method and variation coefficient method, directed weight graph influence is established for the given data set\_data.csv. The music influence parameters in the graph model are mined. On this basis, the music characteristics of artists are clustered within and between schools to measure the similarity of artists, showing the similarity and influence between schools and within schools; in addition, it analyzes how the style of schools changes with the passage of time. By comparing the average musical characteristics of influencers and followers, it is proved that some musical characteristics of influencers play a similar role in influencing the musical color of a particular artist. From the analysis of the changes of the total amount of music genres existing in each era with time, and the genres with better development trend or strong representativeness in each era, we can express the influence of music on time or environment, and understand the value of music through the Internet through the influence of visualization.

## **1. Introduction**

### **1.1 Background**

The influence of music has been a part of human society and an important part of cultural heritage since ancient times. The goal of this paper is to understand and measure the impact of previously produced music on new music and music artists. Music sometimes has revolutionary changes to provide new sounds or rhythms, which can be attributed to a series of small changes, such as the cooperative efforts of artists, a series of influential artists, or changes within society. Sometimes these changes are due to the influence of one artist on another. Sometimes, it is a response to external events, such as major world events or technological progress. By considering the song network and its musical characteristics, and capturing the interaction between musicians, we can better understand how music evolves in society over time.

## 1.2 Problem retelling

Based on the data sets provided by integrated collective music (ICM), this paper studies the evolution and revolutionary trend of artists and genres

Create (multiple) directional music influence network, develop parameters to capture the "music influence" in this network;

Using full\_MUSIC\_ Data set to develop music similarity measurement model;

The similarities and influences between and within schools are compared;

Explore whether the determined influencers affect their respective artists;

Find out whether there is a significant leap in the process of music evolution;

Identify indicators that can reveal dynamic influencers and explain how genres or artists change over time;

Determine the impact of social, political, or technological changes on the network.

## 2. Model hypothesis

1. Assuming that the social background and other factors at that time are ignored, only the data set is modeled.

2. Assuming that the music influence given by the model changes little with time, we can approximately assume that the music influence of music genre influencer remains unchanged for a long time.

## 3. Modeling

### 3.1 Improved TOPSIS scoring model based on network graph

In this paper, a graphic model based on TOPSIS is proposed, and a comprehensive weight determination method based on entropy weight method and variation coefficient method is proposed. The model is more suitable for the data model of this subject, and has the advantages of practical weight determination, fast calculation speed and small error.

Network graph model:  $g = (v, \varepsilon)$ , where  $v$  is the node,  $v = \{1, \dots, V\}$ ,  $\varepsilon$  is the edge connecting the node, and  $\varepsilon = \{(s, t): s, t \in v\}$ . In the directed weighted graph model, if  $(s, t) \in \varepsilon$ , then  $G = (s, t) =$ , where is the weight of the node from  $s$  to  $t$

Coefficient of variation method

The coefficient of variation of each index is calculated as follows:

$$V_i = \frac{\sigma_i}{x_i} (i = 1, 2, 3, 4, \dots, n)$$

Where,  $\sigma_i$  is the variance of index  $I$ ,  $x_i$  is the standard deviation of index  $I$ , and  $\bar{x}_i$  is the average value of index  $I$ .

The weight of each index is as follows:

$$W_i = \frac{v_i}{\sum_{i=1}^n v}$$

### Entropy weight method

The steps of entropy weight method are as follows:

1. Data standardization, the index data standardization.
2. The information entropy of each index was calculated

3. Determine the weight of each indicator

### TOPSIS Model of comprehensive weight determination

Algorithm steps:

1. The polarity consistency matrix is obtained by polarity treatment of evaluation index
2. The normalized data matrix is normalized, and the processing formula is

$$Z_{ij} = \frac{X_{ij}^*}{\sqrt{\sum_{i=1}^n X_{ij}^{*2}}}$$

3. Determine the positive ideal solution  $Z^+$  and the negative ideal solution  $Z^-$

4. The Euclidean distance from the evaluation object to the positive and negative ideal solution is calculated.

$$D_i^+ = \sqrt{\sum_{j=1}^m \omega_j (Z_{ij\max} - Z_{ij})^2}$$

$$D_i^- = \sqrt{\sum_{j=1}^m \omega_j (Z_{ij\min} - Z_{ij})^2}$$

5. The relative closeness between each scheme and the optimal scheme is calculated

6. Comparison scheme

The value of relative closeness is between 0 and 1, and the closer to 1, the closer to the best evaluation level.

### 3.2 K-means clustering model

The ultimate goal of K-means clustering algorithm is to divide data objects into K clusters according to the input parameter k

Firstly, the K value of the number of clusters to be divided is specified, and K initial data object points are randomly selected as the initial clustering centers. Next, the similarity between the remaining data objects and K initial clustering centers is calculated, and the data objects are classified into the nearest clustering centers. Finally, adjust the new class and recalculate the center of the new class. Until the two calculations, the cluster center does not change.

Since the number of K clusters in K-means clustering algorithm depends on the specification of user parameters, the model will become very subjective. In order to overcome this subjectivity, this paper proposes a method to determine the number of K clusters, the elbow method, whose core index is SSE (sum of squares of errors).

### 4. Solution

For the seven problems, this paper gives the following solutions:

#### Question 1:

To create a directed graph from influencer to follower, two parameters of the directed network can be mined: the number of followers, the type of influencer and the proportion of influencer. Music influence shows that the more followers the affected people have, the higher the genre proportion of these followers is, which is consistent with the genre of the affected people, indicating that the greater the music influence of the affected people, the greater the music influence of their genre.

According to the TOPSIS comprehensive weight determination method, the final weight of the

digraph is obtained. According to this, the degree of influencers in the data set, the proportion of the same genre, and some weight scores are listed. It can be concluded that the Beatles, Bob Dylan, and the rolling stones have great influence on music. Combined with the facts, it can be seen that the music influence mined by the model is reliable.

#### **Question 2:**

By integrating the influencer and influence data, 14 music feature attributes, including dance and energy, are clustered, and finally three clustering categories are obtained through the elbow method. The percentage of the highest occupation proportion is used as the comparison index of artists' similarity between schools. In addition, all the data in each genre are clustered, and the similarity index percentage of the highest proportion of digital artists is removed as the genre.

Comparing the similarity indicators of artists between and within genres, we can see that most artists within genres are more similar than those between them. However, we can not rule out the other two cases. The first case is that the sample size of some genres is not enough, and the indicators between genres are the same as those within genres. The second case is that the cultural background of genres leads to the indicators within genres. The number is less than the inter genre index.

#### **Question 3:**

Through the analysis of the given data, it is found that there are 20 genres and their respective 13 musical features. These data are processed as follows:

According to different genres, we analyze the music features of these 20 genres, and find that different music features have different influences on different genres, and different stream parties have different influences on music features.

At the same time, for each genre, by observing the changes of music characteristics of each genre, it is found that no matter which genre, the variance changes of dance ability, energy, value, mode, acoustics, instrumentality, vitality and language are very high, but they are very small (less than one), so no matter which genre, these music characteristics are very similar.

#### **Question 4:**

This paper explains the influence of influencers on followers through the proportion of followers and the schools of influencers. Taking Bob Dylan as an example, it is found that 82.7% of his followers belong to the same genre, which shows that "influential people" do influence the music created by his followers. Then we compare the average value of these followers's musical features with that of the influencers. By calculating the average value of their 13 musical features, we find that the average value of their 13 musical features is basically the same as that of their 13 musical features, which indicates that these musical features have a similar impact on the music of a particular artist. Use.

#### **Question 5:**

According to the solution of problem 3 and problem 1, the inflection point or the point with high or low growth rate of time series line graph may change. According to the number of followers of all the influencers of the same genre in this period, if the number of followers exceeds the number of followers of the influencers of the same genre in the same period, the number of followers can be determined. This feature is the characteristics of the music reformers of this genre in this period, which can be observed by constructing a network diagram: select genre I in the sequence diagram to appear inflection point or non growth rate. In the year of J, all the influencers of genre I are listed. Choose the year of J with the most attention. The beginning node is the name of the influencer, and the end node is the name of the follower. The number of degrees of the influencer can reflect the influence of the music reformer in the genre and era.

#### **Question 6:**

With the development of the times, various genres of music are also developing. In the process of evolution, the music characteristics of this school have gradually changed. Pop / rock music is still an example. With the development of the times, this school of music first developed rapidly, and then

gradually disappeared after reaching the peak. But in the process of this evolution, the music characteristics within the genre are gradually changing. With the development of the times, the music characteristics of the school are mainly manifested in dance ability at the beginning of its rise, music value at its peak and music loudness at its decline.

#### **Question 7:**

Due to the underdevelopment of music technology in the 1930s and 1940s, and the social instability during the world war, country music and Jazz were the most popular in the 1930s and 1940s due to the characteristics of their music schools. Pop / rock got a good development in the 1970s and 1980s. The electronic genre appeared in the 1970s. At this time, the world is in a relatively peaceful era. People have more entertainment space and pop / rock has developed rapidly. Due to the evolution of the times and the development of science and technology, more changes have taken place in music. Therefore, the electronic genre has also developed. It has been well developed.

Therefore, we can express the social situation, technological change and political background at that time from the time series of the development of music schools, the total number of music schools in each period, the development trend in each period.

## **5. Conclusion**

In this paper, data mining is carried out on the data set, and a series of models are established to fit the data in the data set\_music\_data.csv. Such as dance, energy, valence and rhythm, the development trend of music genres over time, the relationship between influencers and followers, etc. the model data has been strictly classified and screened, and the attachment data is complete and effective. However, since data sets are limited to certain types, adding more abundant data may affect the effect of the model. Here, we think that our model can be improved from the following aspects: the value of music can be measured by integrating the background of various music genres in the popular era and the background of various music genres; the influence of a certain music genre on other music genres can be included in the weight, and the weight calculation should take into account the influence of the times and the corresponding social environment, and adjust accordingly to obtain the results. For the music influence of artists, the age of the artist and the corresponding social environment should also be considered when calculating the weight.

The above is one of the solutions to further improve the model, through the above improvements, this model will be more in line with the actual situation, and can better and more accurately give music influence.

## **References**

- [1] Kutlu gundogdu, F., Kahraman, C.: *spherical fuzzy sets and spherical fuzzy TOPSIS method*. *J. Intell. Fuzzy Syst.* 36(1), 337–352 (2019)
- [2] Shih, H.S., shyur, H.J., Lee, E.S.: *the extension of TOPSIS in group decision making*. *Mathematics and computer modeling* 45801 – 813 (2007)
- [3] Lopez, M.M., Heymann, e., senar, M.A.: *dynamic heuristic analysis of workflow scheduling in grid system*. *Proceedings of the 5th IEEE International Symposium on parallel and distributed computing* (2006)
- [4] Kavyasrujana D., Rao B.C. *used cosine similarity measure for hierarchical clustering of sentence extraction*. (2015)