Journal of Electronics and Information Science (2022)
Clausius Scientific Press, Canada
DOI: 10.23977/jeis.2022.070116
ISSN 2371-9524 Vol. 7 Num. 1

Analysis and Research on the influence of Music based on degree-centered European Spatial Distribution

Wenna Wu

College of Science, Chang'an University, Xi'an, Shaanxi, 710064, China

Keywords: Hubs Centrality, tSNE, Music influence

Abstract: Music creation is affected by many factors, especially music and music artists created in the past. First of all, this paper constructs a multi-layer network based on a given data set to visualize the relationship among genres, artists, songs and features. After calculating the centrality score, the sum of the scores of all subsequent nodes is described recursively to describe the authority and influence of music artists. After correlation analysis, we find that there is almost no linear correlation between its features. Then, we use tSNE (random neighbor embedding) to reduce its feature space to three-dimensional space, and use Euclidean distance to measure similarity in this space, indicating that these artists and songs are distributed in this space, even after most people are undersampled.

1. Introduction

Music is the creation of human beings, which is inextricably linked with human beings. In the process of human civilization, music plays a very important role in social development, individual development and educational development in its unique way. Since ancient times, music, as an indispensable part of human development, has been exerting an imperceptible influence on our lives [1]. At the same time, music and music are also influencing each other. Artists in the creation of music, talent, hot current events, the use of musical instruments, personal life experience and other factors will have an impact on their music, artists also have a mutual influence that can not be ignored. Therefore, we hope to develop a method that can quantify the development of music and measure the influence of music more vividly and concretely through the similarity of various features of songs, the creation of other artists and different genres.

2. Initial data and standardized processing

2.1 Normalization

(1) Min-Max Normalization for Tempo and Popularity:

In the process of normalization, use Min-Max Normalization for tempo and popularity, linearly change the original data to map the value to [0-1].

$$x^* = \frac{x - \min}{\max - \min} \tag{1}$$

(2) Z-score Normalization for Loudness and Duration:

Perform data standardization processing on the original data according to the following transformation function [2]. We use this method to normalize loudness and duration.

$$x^{*} = \frac{x - \mu}{\sigma} \tag{2}$$

Where μ represents the mean of all sample data, σ represents the standard deviation of all sample data

2.2 Linear Correlation Analysis

Visualize the correlation between various features through the heat map, which more intuitively reflects the correlation between music features, see Figure 1. It can be found that the correlations are all low, and there is a partial negative correlation.

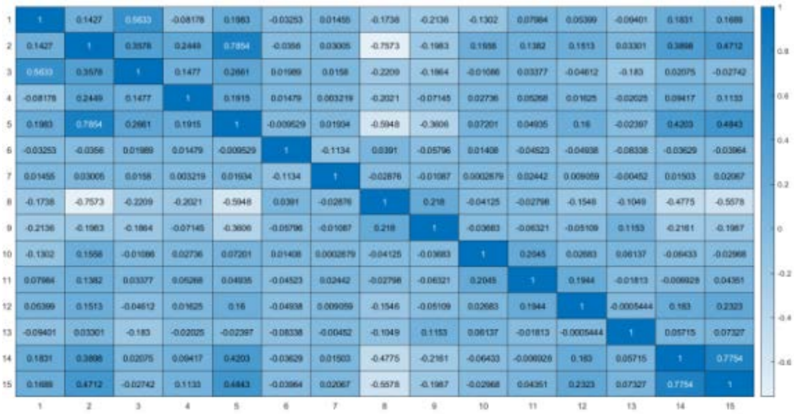

Figure 1: Correlation Coefficient Heat Map

2.3 Data Overview

Visualize the given data sets to establish the multilayer network [3], as shown in Figure 1. All nodes can be divided into four categories, i.e. the four layers drawn in the network.

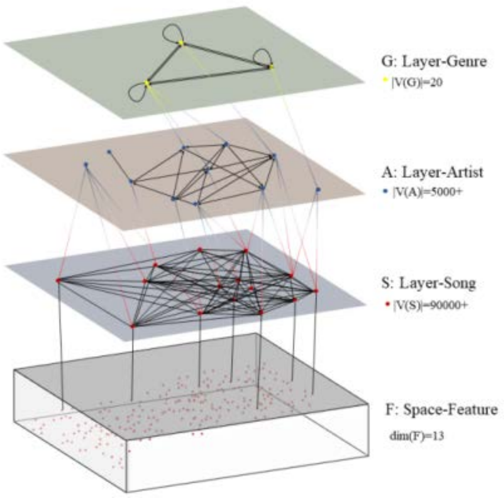

Figure 2: Multilayer network

3. Music Influence Network Model

3.1 Initial Network Construction

Firstly, we drew the non-directional influence relationship network diagram between the influencers and the folloer, see Figure 2. The graph is unit that all edge weighted 1 and the arrow pointed from influencer to the follower [4].

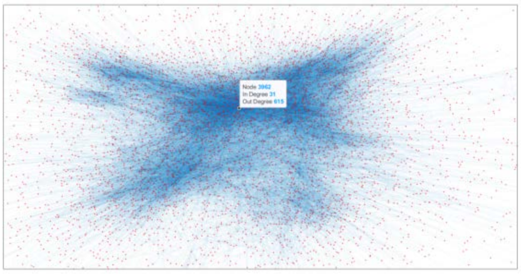

Figure 3: Initial Artist influence Network: A

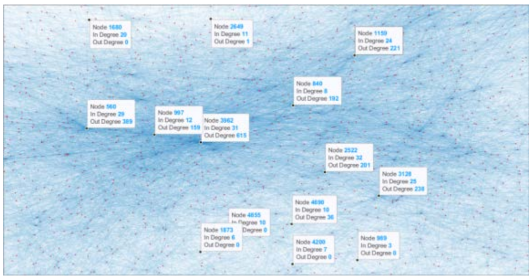

Figure 4: Enlarged View of Initial Artist influence Network

From the figure above, deeper the blue is, the more frequency is the influence among the nodes in corresponding area. We compute the hubs centrality for every nodes in A that was just shown at the figure above, and then we replace the weight of out edges of V (A) with *hcen*(V (A)). Then the artist influence network was built completely.

According to data analysis, it can be found that most of the scores are less than the degree of enrollment, while the highly regarded artists are more influential. From the figure, it can be found that the artists with the nodes numbered 560 and 3962 according to their influence, the out degree is significantly greater than the in degree.

3.2 Extract subnetwork

3.2.1 Age

First, the drawn network diagram is aggregated according to the year, the nodes at the same time are merged and the relationship between the corresponding edges (With weight) is inherited.

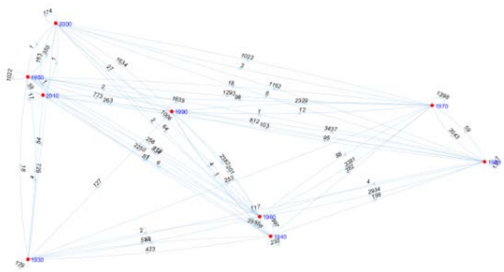

Figure 5: Influence Diagram of Empowerment Year

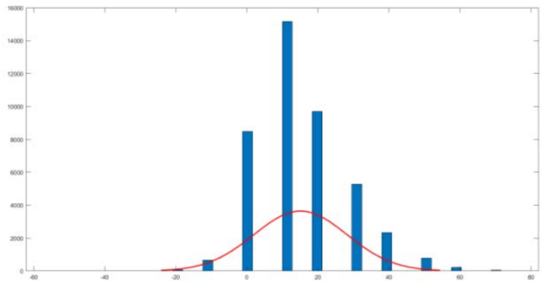

Figure 6: Distribution of Δyears

In the graph A_{new} obtained after aggregation, each node represents a year. After analyzing the data of different years, the following probability distribution diagram can be obtained. The result is shown in figure below. we can see that when the year is ten years apart, the weight ratio is the largest.

3.2.2 Genre

After drawing the network of the influence relationship of 20 music genres, the weight of different genres of music can be obtained through aggregation calculation, and the weight relationship is shown in the figure. The larger the blue circle around the node representing a certain genre, the larger the range of influence and the greater the influence.

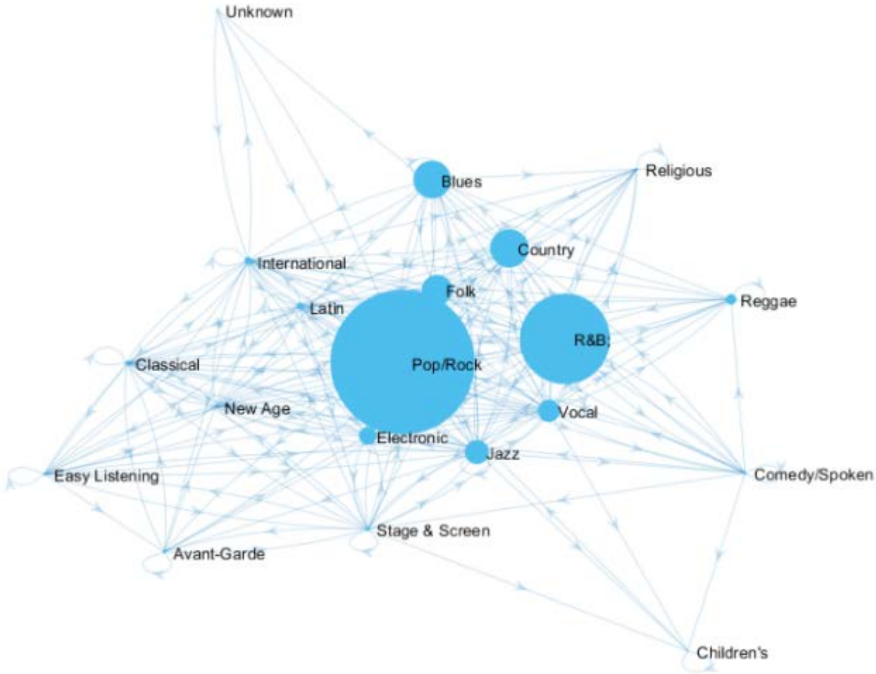

Figure 7: Genre influence Network

It can be seen from figure above that Pop-Rock has the greatest impact. Because of its great influence, in order to draw the picture more beautifully, the scope of influence shown in the picture is reduced by about forty times.

4. Music Similarity Measure

4.1 Music Feature Data Visualization

We first divide all artists in the known data into genres, and in this step, we give up some sample that cannot be labeled for not participating in influence network. Artists without genres (discrete nodes) will be discarded. Then a relationship diagram between S network and A network will be constructed. In the influence network diagram, if a node in the S network has two edges connected to the A network, then the node in the S network is split into two nodes, and the two connected edges are assigned to two new nodes.

4.2 Dimension Reduction by tSNE

Use t-SNE for dimensionality reduction and visualization of high-dimensional data. Since t-SNE mainly focuses on the local structure of the data, it can be used to more accurately visualize the similarity relationship between samples. The Kullback-Leibler (KL) divergence of the joint probability of the original space and the embedding space is used to evaluate the visualization effect, and then the loss function is minimized by gradient descent, and the convergence result is finally obtained. After dimensionality reduction through tSNE, we will get a new network F_{new}.

4.3 Similarity Measured with Euclidean Distance

In the newly constructed three-dimensional network space F_{new}, the Euclidean distance is used to calculate the similar distance between nodes,

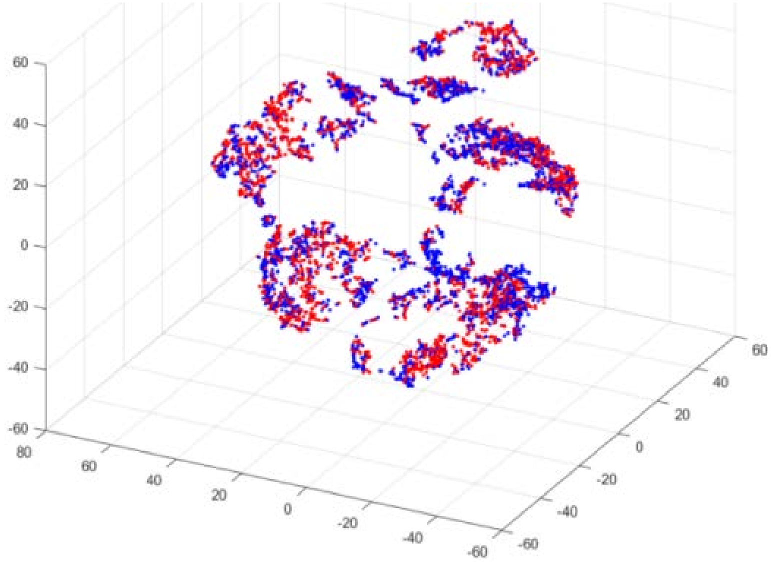

Figure 8: Distribution of Artists Genres, where red refers to Pop/Rock

We can find that the red dots representing Pop/Rock music are not gathered in the same place, but scattered throughout the space, so there is no significant gathering feature within the genre. Artists within genre are no more similar than artists between genres.

5. Conclusion

In order to explore the influence of music and art creation, this paper first visually presents the relationship among schools, artists, songs and features and calculates the centrality score, that is, recursively calculates the sum of the scores of all subsequent nodes. it is found that there is almost no linear correlation between the features. Then, we use tSNE (Random neighbor embedding), use Euclidean distance to measure similarity in this space, and draw scatter plot to show the distribution, which shows that these artists and songs are distributed in this space.

References

[1] Han H, Wang W Y, Mao B H. Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning [C] // Proceedings of the 2005 international conference on Advances in Intelligent Computing - Volume Part I. 2005.

[2] Liu X Y, Wu J, Zhou Z H. Exploratory Undersampling for Class-Imbalance Learning [J]. IEEE Transactions on Systems Man & Cybernetics Part B, 2009, 39(2): 539-550.

[3] Matthias, Mauch, Robert, et al. The evolution of popular music: USA 1960–2010 [J]. Open Science, 2015.

[4] Fitch W T. The evolution of music in comparative perspective. [J]. Annals of the New York Academy of Sciences, 2005, 1060 (Dec): 29.