How Does Musical Influence Lead to Musical Revolution?

Summary

During the revolution of music, interactions within and between genres can have a critical impact on the direction of musical change. In this paper, we develop models to analyze the influence between musicians, measure the degree of similarity between music, and reveal the changes in genres over time, as well as the influence of social, cultural, and historical aspects.

In TASK 1: First, we take each musician as a node and calculate the directed influence between musicians as weights by considering the factors of **time span** and **genre span** to build a **directed music influence network**. Then we use social network analysis (**SNA**) to analyze the influence network. The Degree Centrality of musicians in the network is calculated, and the **Pagerank-modified eigenvector centrality** is further used to measure the influence of musicians(*MI*), obtaining that Bob Dylan, The Rolling Stones, Chuck Berry, and Elvis Presley have the highest influence. We also find that *MI* obeys **power-law distribution**, implying that the influence network between influencers and followers is a **scale-free network**. Finally, we explore the sub-network to analyze the meaning of the subset of *MI*.

In TASK 2 & 3: First, we use Cosine similarity and Euclidean similarity to measure **global similarity** S_{global} and **local similarity** S_{part} in two aspects of music, and consider the influence of popular trends to build **Multi-dimensional music similarity evaluation model** to measure the similarity of music (*MS*). Music within genres is found to be more similar than between genres. Then, to investigate the distinguishing features of genres, we combine different features and built the **Discrimination evaluation model of feature combination** to select the feature combinations that better distinguish genres from each other. Finally, we analyze changes in genres over time.

In TASK 4: To investigate whether the influencers actually influence the followers' music, we develop a **Long-short term influence model based on sliding window** to calculate the influence degree score (*IDS*), considering the time sequence of release of the followers' music and the influencers' music, and determine the influence threshold. The result shows that followers are indeed influenced by the influencers. Subsequently, to determine the "contagiousness" of specific musical characteristics to musicians, we also consider specific musical characteristics, analyze the degree of closeness (*DC*) of characteristics of followers' and influencers' music, and establish the **Contagious evaluation model of musical characteristics** to find the most contagious musical characteristics for different followers.

In TASK 5 & 6 & 7: First we establish **constraint equations** to identify musical revolution points based on the musical characteristics of the genre and the number of musicians in the genre. And based on the *popularity* before and after the musical revolution point to identify the main influencers in the musical revolution. Then we take 10 years as the examination period to calculate the *IDS* of the analyzed object to everyone within the genre based on the *IDS* in Task 4, and weight the sum to get the dynamic influencer index (*DII*) for identifying the leading figures within the genre in different generations. Ultimately, we analyze the relationship between music and politics, culture, and society.

Finally, we perform a sensitivity analysis of the model and investigate the effect of changes in the variable parameters of the model on the results.

Keywords: SNA, Slide Window, Scale-free Network, Pareto's Principle, PageRank, Combination of Features

Contents

1 Introduction

1.1 Background and Problem Statement

Music is a cultural product that people use to express their emotions and relax their bodies and minds. The influence of musicians on each other has a very important impact on the development and evolution of music. The development of computer technology allows us unprecedented access to data and computing power to analyze the network of influences between musicians, from which we can extract and analyze the key factors that influence the process of musical development, which will also help us understand the meaning of music better. In this problem, we need to accomplish the following objectives:

• Create a music influence network to analyze the influence between musicians and capture the music influence parameters in the influence network.

• Develop similarity evaluation model to explore the degree of influence and similarity between or within genres and between musicians.

• Identify indicators of revolution in genres and analyze the interplay between them in relation to society, politics, and culture.

• Applying a dynamic analysis of the similarities and influences between and within genres of musicians over time.

1.2 Our work

The work we have done in this problem is mainly shown in the following Figure(1).
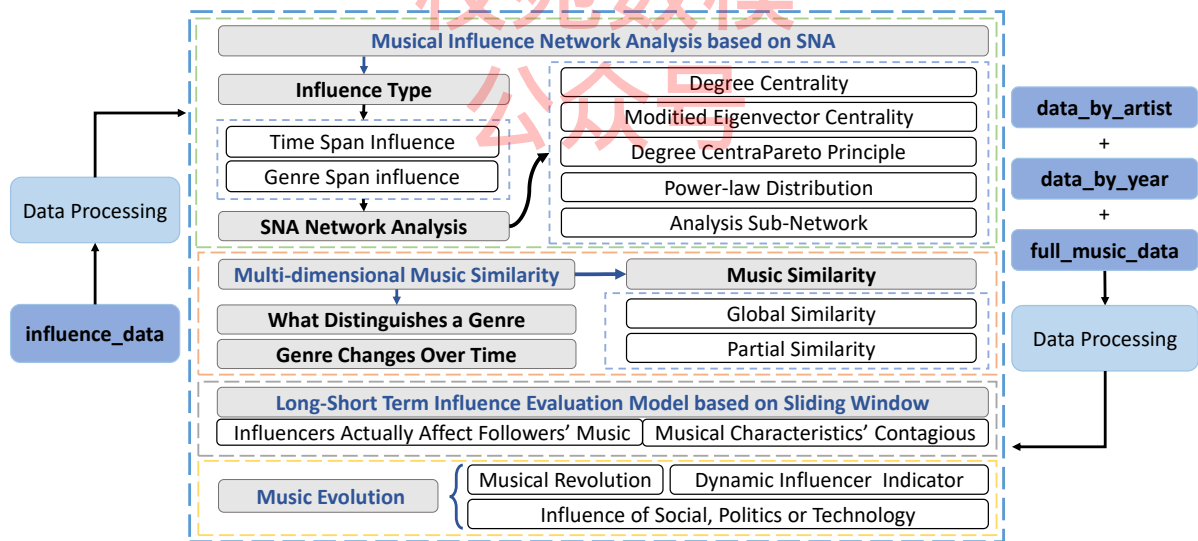
Figure 1: Our Work

2 Preparation of the Models

2.1 Assumptions

We made the following assumptions to help us with our modeling. These assumptions are the promise of our subsequent analysis.

• Assuming that the data is accurate and there is no falsified data: This means our analysis accords with the truth.

• Assuming that followers can only be influenced by songs released before the current year, and that songs after the current year have no effect on the follower's for that year: this is done to measure the true influenceability of the influencers in a sliding window model.

• Assuming that the decade when the musician begin his carrer is the year when they debut, because the accurate years when the musicians debut are not given.

• Assuming that the musician's genre is limited to the genre provided by the data, and have no genre changes, that all his/her songs are in the style of his or her genre

2.2 Notations

The primary notations used in this paper are listed in Table 1.

Table 1: Parameter Settings

parameter	description
T	time-span influence factor
D_{Ci}	degree centrality
E_{Ci}	eigenvector centrality
MI	music influence
S_{global}	global similarity
S_{part}	partial similarity
$V^{(i)}$	music feature vector
MS	multi-similarity
SMS	short-term similarity
LMS	long-term similarity
IDS	influence degree score
DC	degree of closeness
DII	dynamic influencer indicator

3 TASK 1: Musical Influence Network Analysis Based on SNA

3.1 Data Processing

An overview of the given data shows that some songs are sung by more than one person (e.g. *It Can't Happen Here* and *I Wonder If I Take You Home*), and such songs cannot reflect the characteristics of a single musician, so to avoid disturbing the model, we remove the data of a total of 4410 chorus songs (4.4% of the original data). Since there are singers with only choral music, they have no music data after the deletion of choral music, which are invalid data. We remove these people, including those who are not shared between two dataset, to prevent any effect on the model. The data processing is shown in the Figure(2).
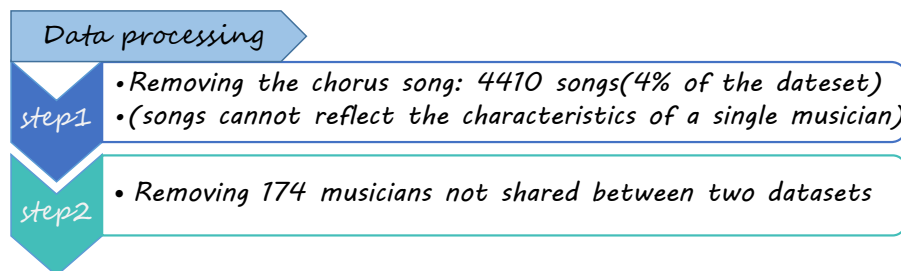

Figure 2: Data Processing

3.2 Model Overview

To analyze the musical influences among musicians, we establish a directed network of musical influences with each musician as a node and directional influences among musicians as

weights. We then further explore the influence among musicians using social network analysis (SNA) to capture the *music influence* in the network to measure the influence of each musician. Finally, we explore the sub-networks of the influence network and explore the significance of *music influence* in the sub-networks.

3.3 Musical Influence Network Analysis Based on SNA

3.3.1 Establishment of Musical Influence Network and Analysis Using SNA

▶ Establishment of Musical Influence Network

The directed graph can be well used to describe the mutual influence relationship among musicians. Based on the relationship between influencers and followers among musicians given by *influence_data*, we build a directed network of musical influence among musicians with each musician as a node and the directed influence among musicians as a weight.

Exploring the data reveals that the types of influence between musicians are divided into two categories: influence within the same genre and influence among different genres. Influence among musicians exists not only that a predecessor influences a junior, but also that a junior influences a predecessor. Therefore, we classify the types of inter-musician influence as follows.

From genre span: Influence occurs within the same genre and between different genres.

From time span: Junior influences predecessor and predecessor influences junior.

Fully considering the difference between the span of genres and the span of time in which the influence occurs, we define the weight of each side of the music influence network:

● **Time span impact factor**

Assuming that existing musician i influences musician j, the time difference Δt is defined as follows:

$$\Delta t = t_j - t_i \tag{1}$$

where t_i and t_j are the debut times of musician i and musician j, respectively.

In general, juniors are more likely to be influenced by predecessors as they grow up, while predecessors are less likely to be influenced by juniors. Therefore, we can conclude that the probability of the older generation being influenced by the younger generation is lower. And if this happens, it means the influence of the younger generation is very strong. Compared to predecessors influencing juniors, juniors influencing predecessors should be considered that the juniors have a stronger influence. We therefore define the time-span influence factor T as follows

$$T = \begin{cases} -\beta \Delta t, & \text{if } \Delta t < 0 \\ \Delta t, & \text{if } \Delta t \geq 0 \end{cases} \tag{2}$$

In equation(2), β is a set parameter and $\beta > 1$, indicating the difference in the effect of the two influence categories.

● **Genre span impact factor**

Generally speaking, musicians in the same genre are more likely to influence each other because they share the same direction and research areas. In contrast, musicians from different genres are less likely to influence each other because their work styles rarely overlap. Therefore, we argue that cross-genre influencers have stronger influence compared to within-genre influencers. The genre-spanning influence factor a_{ij} between influencer i and follower j is given as follows:

$$a_{ij} = \begin{cases} 1, & \text{cross-genre} \\ 0, & \text{within-genre} \end{cases} \tag{3}$$

• **Edge weights of a directed graph**

Combining the above analysis on the time span and genre span impact factors, we give the formula for the directed graph weights w_{ij} as follows:

$$w_{ij} = (1 + a_{ij}) \cdot e^{\frac{T}{80}} \tag{4}$$

In equation(4), w_{ij} is the weight of the directed edge between musician nodes i to j. The time interval $|\Delta t|$ between the influencer and the influenced is at most 80, so T is normalized using $\frac{T}{80}$. w_{ij} also represents the intensity of the influence of musician i on musician j. The larger the value of w_{ij}, the stronger the intensity of musician i's influence.

▶ SNA Network Analysis Method

In order to capture the *musical influence* in the established musical influence directed network, we analyze the directed network using social network analysis (SNA) method[1]. The influence of each musician is measured and analyzed by calculating the Degree Centrality of each musician in the musical influence directed network and further calculating the Eigenvector Centrality modified by the PageRank idea. The specific steps are as follows:

Step 1: Calculating Degree Centrality

In an undirected network, the degree of a node can be used to measure the centrality of that point. Analogously, in the music influence directed network we built, we can use the sum of the out-degree weights of a node to measure the importance of a node's influence at the microscopic level. The assumption behind this metric is that nodes with many connections are important nodes, which means that musicians who have an impact on more people have a stronger influence. We therefore give the definition of degree centrality (D_{Ci}) for point i as shown in the equation(5)

$$D_{Ci} = \sum_{j=1}^{n} w_{ij} d_{ij} \tag{5}$$

In equation(5) n represents the total number of nodes in the network and d_{ij} denotes the existence of directed edges between node i to node j. And d_{ij} is defined as follows:

$$d_{ij} = \begin{cases} 1, & \text{no directed edge } from\ i\ to\ j \\ 0, & \text{there is directed edge } from\ i\ to\ j \end{cases} \tag{6}$$

Step 2: Calculating Eigenvector Centrality

The importance of a node depends both on the number of its neighboring nodes and on the importance of its neighboring nodes. This means that the more important the nodes connected to node i, the more important node i will be. We define the eigenvector centrality E_{Ci} to measure the importance of node i at the macro level, as shown in the iterative equation(7)

$$iterative\ equation: \quad E_{Ci} = \lambda \sum_{j=1}^{n} d_{ij} E_{Cj} \tag{7}$$

where *lambda* is a scaling constant and the initial value of the iteration is set in the equation(8)

$$initial\ value: \quad E_{Cj} = D_{Cj}, \quad j = 1 \sim n \tag{8}$$

However, for the directed network we build, the above analysis process only takes the importance of neighboring nodes into account. To increase the influence of the nodes themselves on the outside, we modify the equation based on the PageRank idea.

PageRank[2] is an algorithm proposed by Sergey Brin and Lawrence Page to rank the importance of nodes in a network. The PageRank algorithm assumes that a node exerts influence

on the system is the result of its own influence combining with the results of its followers' influences. Therefore, we modify the model shown in the following equation: Therefore, we modify the model shown in the following equation:

$$\text{iterative equation:} \quad E_{Ci} = \lambda \sum_{j=1}^{n} d_{ij} E_{Cj} + (1 - \lambda) E_{Ci} \tag{9}$$

After several iterations, E_{Ci} converges to a fixed value E_{Ci}'. We define the *musical influence* (MI_i) of musician i as follows:

$$MI_i = \frac{E_{Ci}'}{max\{E_{Cj}', \ j = 1, 2, \cdots, n\}} \tag{10}$$

3.3.2 Model Test on Musical Influence Network

Using all data given in the *influence_data*, the *musical influence* (*MI*) of each musician is calculated, and the *MI* of somes musician are shown in the Table(2).

Table 2: *MI* of Musicians

Musician	MI	Musician	MI
The Beatles	1	Bob Dylan	0.738
The Rolling Stones	0.546	Chuck Berry	0.5
Hank Williams	0.424	Elvis Presley	0.424
Jimi Hendrix	0.397	Little Richard	0.39
James Brown	0.387	The Velvet Underground	0.356

As can be seen from the above table, The Beatles has the highest *MI*, followed by Bob Dylan, The Rolling Stones, Chuck Berry, Elvis Presley and so on. The similarity between the most influential musicians we analyzed and *The 20 Most Representative Musicians in the World Jointly* selected by CNN and Songlines[3] (a world-class music publication) is relatively high. We can conclude that the establishment and analysis process of the model is accurate, reasonable and feasible.

3.4 Analysis of The Result

Most real-world networks are not random networks, and a few nodes tend to have a large number of connections, while most nodes have only a small number of connections, generally obeying the Pareto Principle law[4]. Complex networks with degree distribution conforming to power-law distribution are usually called scale-free networks[5].

In section 3.2.3, we calculated the *musical influence* (*MI*) of each musician. A line graph of the frequency distribution of *MI* is plotted as shown in the blue solid line in the Figure(3).
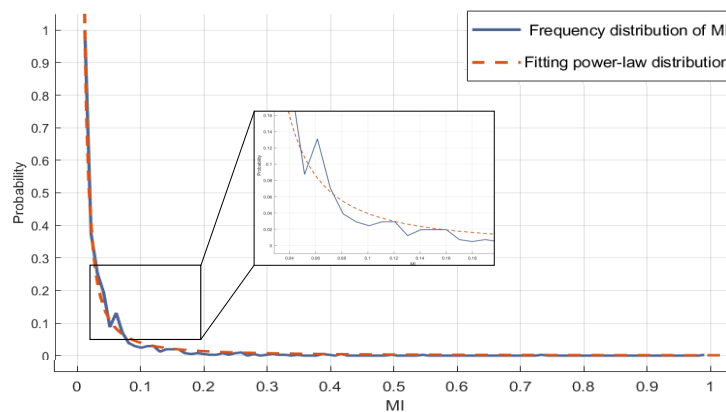

Figure 3: The Frequency Distribution of *MI*

To investigate the nature of the *MI* distribution, we fit the frequency distribution curve of *MI*. Observe that the distribution of *MI* in the figure resembles a power-law distribution, so the complementary cumulative distribution function of the power-law distribution is defined as equation(11)

$$P(X|X > x) = cx^{-a}, \qquad c > 0, a > 0 \tag{11}$$

We solve for the power-law distribution parameters using the maximum likelihood estimation, and the distribution parameters at a confidence level of 95% are shown in the Table(3).

Table 3: Distribution Parameters

distribution parameter	estimated value	95% Confidence interval
c	0.001173	(0.0009809, 0.001366)
α	-1.526	(-1.565, -1.488)

With the goodness of fit $R^2 = 0.9931$, we can consider that the Musical Influence Network is a scale-free network. Musical influence network obeys the law of Pareto Principle, i.e., only a small number of musicians influence many other musicians, while most musicians do not influence other musicians. This also means that a small number of highly influential musicians will have an impact on the music world. And this will be analyzed in detail in the subsequent analysis.

3.5　Analysis of Musical Influence Sub-network

In the model building process, we categorize time span and genre span: influence within genre, influence across genre, influence of predecessors on juniors, and influence of juniors on predecessors. We build independent sub-networks on genre and time span separately and use SNA analysis to explore each musician's influence on predecessors and on juniors EI_{pre} & EI_{jun}, influence within and across genres EI_{wit} & EI_{out}. EI_{pre}, EI_{jun}, EI_{wit}, and EI_{out} are all subsets of the total influence *EI*.

The cross-genre influence data, within-genre influence data, predecessor influence data, and juniors influence data are selected from *influence_data*, respectively, to establish the musical influence sub network. Then the *musical influence*s are calculated separately and normalized. The results of the calculation of each type of influence and the musical influence sub-network are shown in the Figure(4).

In the Figure(4), the size of nodes in each subnetwork represents the *MI* of the corresponding type of singer, while the thickness of the directed curve represents the weight of influence. As can be seen from the Figure(4), the occurrence frequency of influence within genres is higher than that between genres, and the occurrence frequency of influence of predecessors on their successors is higher than that of the latter on their predecessors. It can be concluded that in the interaction of musicians, musicians are more likely to influence musicians in the same genre, and musicians are more likely to influence younger generations rather than predecessors. It can also be seen from The *MI* subseries that The Beatles and Bob Dylan have a huge influence on musicians both in and out of The genre, and Bob Dylan goes one step further and has a strong influence on his predecessors as well.
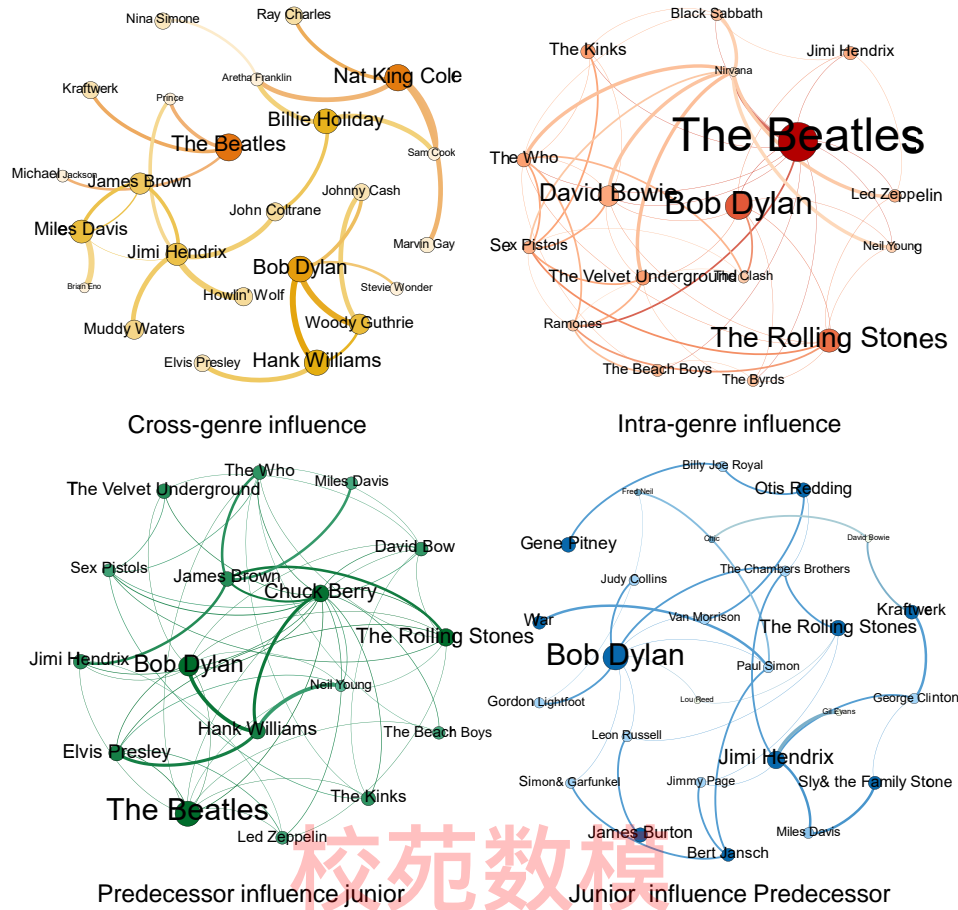
Figure 4: Sub-network of Musical Influence

4 TASK 2 & 3: Multi-dimensional Music Similarity Evaluation Model

4.1 Model Overview

To adequately measure the similarity between the music, we evaluate the similarity between the music using cosine similarity and Euclidean similarity, respectively. The global similarity S_{global} between the music is first calculated. We further explore the partial similarity S_{part} between music. Finally, the trend of the year when the music was released is considered to influence S_{time}. The multidimensional similarity evaluation between music is constructed to obtain the similarity between music. An overview of the model is shown in the Figure(5).
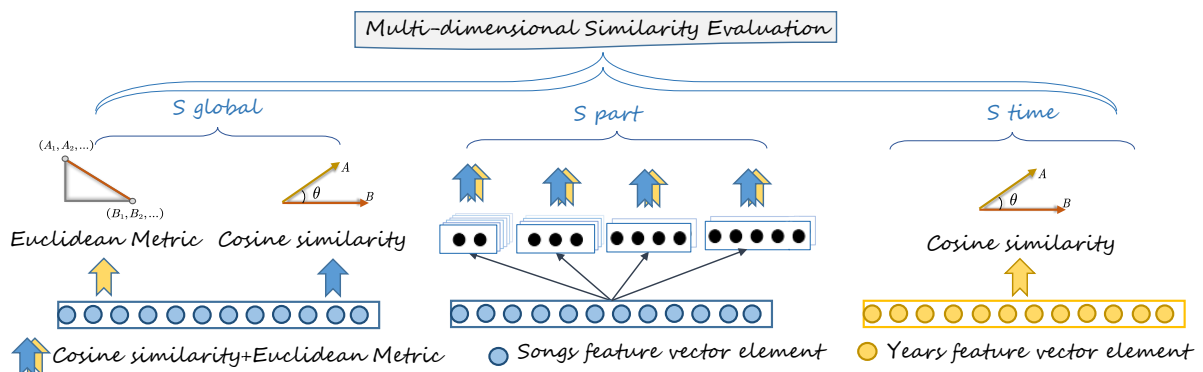

Figure 5: Multi-dimensional Music Similarity Evaluation Model

4.2 Establishment of Multi-dimensional Music Similarity Evaluation Model

4.2.1 Global Similarity of Music

- **Euclidean similarity**

The Euclidean similarity measure is based on the absolute distance between the feature vectors and is used to reflect the absolute difference in the feature values between the two songs. Then the distance between the eigenvectors of music A and music B $distance_{AB}$ is defined as follows:

$$distance_{AB} = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2} \tag{12}$$

Then the Euclidean similarity between music AB $S_E(\mathbf{A},\mathbf{B})$ is defined by the following equation:

$$S_E(\mathbf{A},\mathbf{B}) = \frac{1}{1 + distance_{AB}} \tag{13}$$

- **Cosine similarity**

Different from Euclidean similarity, cosine similarity uses the cosine of the angle between two vectors in vector space to measure the similarity between two vectors. The cosine similarity focuses more on the difference between the two vectors in terms of direction rather than the difference between the values. The cosine similarity between the eigenvectors of music A and music B $S_{cos}(\mathbf{A},\mathbf{B})$ is defined by the following equation(14)

$$S_{cos}(\mathbf{A},\mathbf{B}) = 1 + \frac{cos\theta}{2} = 1 + \frac{\mathbf{A} \cdot \mathbf{B}}{2 \cdot ||\mathbf{A}|| \cdot ||\mathbf{B}||} = 1 + \frac{1}{2} \cdot \frac{\sum_{i=1}^{n} A_i \cdot B_i}{\sqrt{\sum_{i=1}^{n}(A_i^2)} + \sqrt{\sum_{i=1}^{n}(B_i^2)}} \tag{14}$$

The final global similarity S_{global} between music A and B is defined by the following equation(15)

$$S_{global}(\mathbf{A},\mathbf{B}) = S_E(\mathbf{A},\mathbf{B}) + S_{cos}(\mathbf{A},\mathbf{B}) \tag{15}$$

4.2.2 Partial Similarity of Music

The global similarity measure between music is defined above, and the high-latitude feature vectors containing all the music features are considered. But this does not fully reflect the similarities between the music. The similarity measurement of music should be multi-dimensional and separate measurements should be made between different characteristic dimensions of music. In fact, in life, music can be considered similar if it is similar in some aspects such as rhythm, emotional tendencies, and instrumentation, but not in all characteristics. At the same time, songs from the same genre are not similar in all characteristics but may be similar only in certain aspects. To weigh the similarity between the whole dimension and the partial dimension of music, we explore the local similarity of music from the micro point of view, in terms of each feature.

We randomly select two, three, and four features from the key features of music to form different feature vectors **V**. There are a total of N new feature vectors composed of the extracted features.

$$N = C_{12}^2 + C_{12}^3 + C_{12}^4 + C_{12}^5 = 1573$$

The new feature vectors are $V^{(1)}, V^{(2)}, \cdots, V^{(1573)}$, respectively. For music i and music j, we first calculate the S_{global} of each combination of their features and select the highest 50 S_{global}

to calculate the average value as the partial similarity ($S_{part}(i,j)$) between music i and music j. $S_{part}(i,j)$ is defined in the equation(16).

$$S_{part}(i,j) = \frac{\sum_{k=1}^{50} S_{global}(\mathbf{V_i^{(k)}}, \mathbf{V_j^{(k)}})}{50} \tag{16}$$

4.2.3 Fashion Trends Impact Factor

Considering that music is a literary and artistic product with strong social and cultural attributes, we should also consider the influence of the popular trend in the year of music release when analyzing the similarity of two pieces of music. Usually, the style of the song will evolve in the direction of the popular trend of the year. If there is a big difference in the popularity of the two songs in the year of their release, then the difference in melody and vocals between the two songs will change with the pop trends of the two songs in the year of their release.

Therefore, when measuring music similarity, the influence of pop trends on the music itself should be reduced, so as to explore the intrinsic similarity of songs after stripping away the impact of pop trends. For the two songs released in year a and year b, we defined the trend impact factor S_{time} as:

$$S_{time} = S_{global}(Year_a, Year_b) \tag{17}$$

Where $Year_a$ and $Year_b$ are the music style attribute vectors of year a and b, respectively.

In summary, we establish a multidimensional music similarity (*MS*) evaluation model considering the influence of popular trends, defined as follows, β is a parameter defined later ($0<\delta<1$).

$$MS = ((1-\delta)S_{global} + \delta S_{part}) \cdot (1 + \frac{0.1}{S_{time}}) \tag{18}$$

4.3 Evaluation of Music Similarity

We count the number of genres to which all influencers and followers belonged, and plot the following Figure(6) for each type of musician's genre.
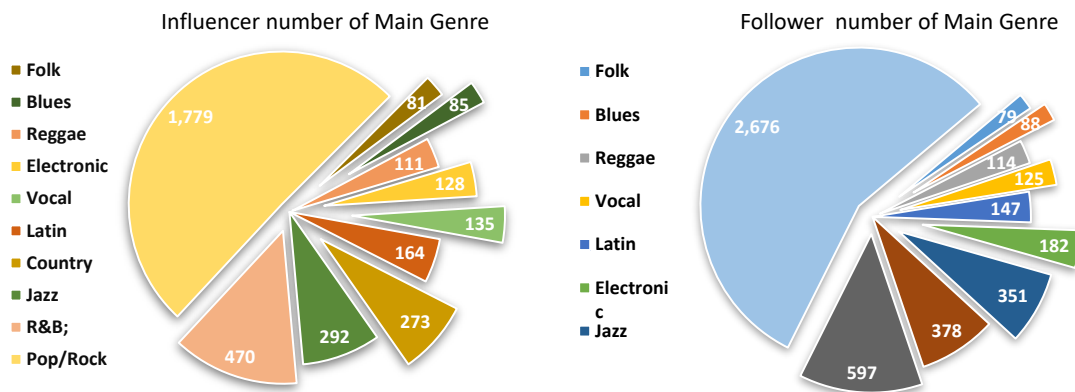

Figure 6: Statistics of Musicians by Genre

As can be seen from the Figure(6), Pop/Rock, R&B, and Jazz have the largest number of influencers, accounting for nearly 75% of all musicians. Therefore the subsequent analysis is based mainly on songs and musicians from these three genres.

- **Similarity of Songs**

Table 4: The Music Chosen from Pop/Rock, R&B, and Jazz

Pop/Rock			
Blind Faith	So Alone	Africa	Long Cool Woman
R&B			
Thinking of You	Long Way Go	Who's Loving You	Don't Stop 'Til You $\cdots$
Jazz			
**** 'N Boogie	What's New	Floatin	Summer Breeze

We select four songs from Pop/Rock, R&B, and Jazz as shown in the Table(4) and then analyzed their similarity.

The established Multi-dimensional Music Similarity Evaluation Model is used to calculate the overall similarity S_{global} between each song and the similarity MS after adding the partial similarity S_{part}. Set $\delta = 0.5$, the calculation results of S_{global} and MS are shown in the Figure(7) and (8).
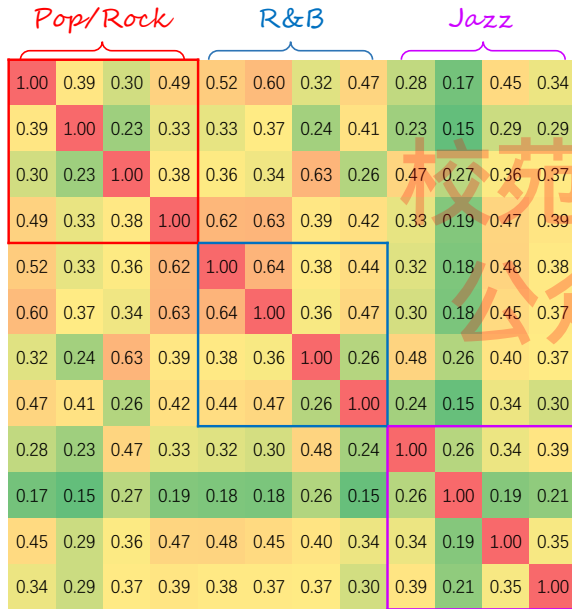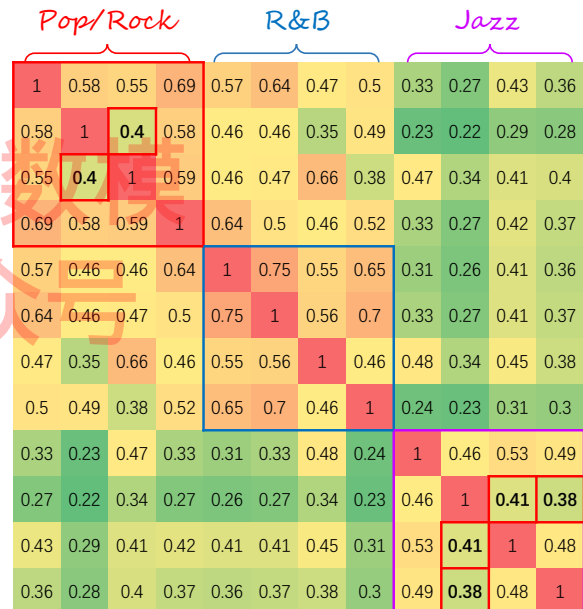

Figure 7: Global Similarity

Figure 8: Multi–Similarity

Figure(7) shows the similarity between songs in each genre without the addition of S_{part}, and Figure(8) shows the similarity between songs in each genre with the addition of S_{part}. As can be seen from the figure that songs within the same genre have higher similarity scores between them. Comparing Figure(7) and Figure(8) with the part circled by the box (i.e., the similarity scores between songs and songs within the same genre), it can be found that: the similarity scores between songs of the same genre have increased after the introduction of the parital similarity of songs (S_{part}), while the similarity between songs of different genres has decreased. The addition of local similarity makes the measurement of musical similarity more accurate and reasonable.

Thus we can consider the discussion about local similarity in 4.2.2 to be accurate: that is, **songs in the same genre may have high similarity in the characteristics of several dimensions**. Songs within the same genre can be distinguished from songs in other genres by a combination of several features.

• **Similarity of Musicians**

We used the established Multi-dimensional Music Similarity Evaluation Model to calculate the similarity of Pop/Rock, R&B, and Jazz musicians within and between genres. The specific values are shown in the Table(5).

Table 5: Similarity of Musicians from Pop/Rock, R&B, and Jazz

Genres	Within genres			Between genres		
	Pop/ROCK	**R&B**	**Jazz**	**Pop and R&B**	**Pop and Jazz**	**R&B and Jazz**
Average	0.659	0.683	0.627	0.647	0.560	0.592
Min	0.399	0.439	0.382	0.309	0.374	0.359
Max	0.917	0.906	0.916	0.949	0.869	0.898

Observing the Table(5), we can learn that the intra-genre artists of each genre are more similar than the similarity with the artists of other genres. Therefore we can assume that the artists within each genre are more similar than the artists between genres. At the same time, we can also see that the average similarity between POP and R&B artists reached 0.647, which is very close to the average similarity within Pop. From this, it can be concluded that there may be more similarities between genres. To visualize the relationship between genres, we used PCA to downscale the characteristics of Pop/Rock, R&B, Jazz and plot the distribution of their musicians in three dimensions, as shown in the Figure(9).
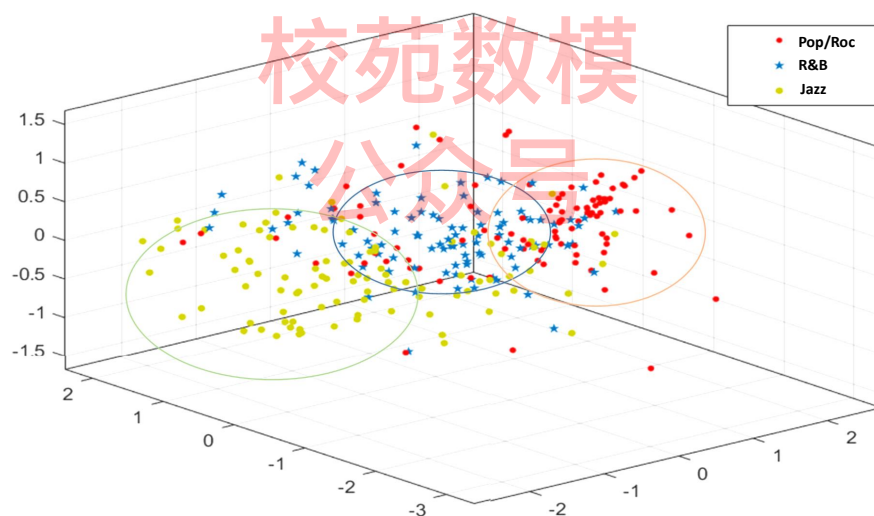

Figure 9: Statistics of Musicians by Genre

As can be seen from the Figure(9), the artists in the genres are more similar, but there is no clear line of distinction between genres. This coincides with the conclusion of our analysis of the Table(5).

4.4 Analysis of the Musical Discrimination and Influences of Genres

4.4.1 Discrimination Evaluation Model of Feature Combination

Combining with the discussion of partial similarity in music in 4.2.2 and 4.3: songs between similar genres may do not share a high degree of similarity in all features, while the combination of several particular features may give music in the same genre a very high degree of similarity and thus distinguish it from music in other genres. Similarly, there may be a high similarity between different genres in the combination of certain characteristics.

To analyze what combination of features distinguishes a genre, we arrange and combine the 12 musical features except *popularity* in two-two, three-three, four-four, and five-five combina-

tions to obtain a total of N' types of new feature combinations: $V^{(1)}, V^{(2)}, \cdots, V^{(1573)}$

$$N' = C_{12}^2 + C_{12}^3 + C_{12}^4 + C_{12}^5 = 1573 \tag{19}$$

Based on the new combination V, the partial similarity S_{part} defined by the equation(16) in 4.2.2 between the genres and the other genres is calculated. The minimum value of S_{part} is $\frac{1}{2}$, which represents the minimum local similarity. For the i_{th} new feature combination $V^{(i)}$, a metric vector $\mathbf{S_{part\ I}} = [S_{part\ 1}, S_{part\ 2}, \cdots, S_{part\ 18}]^T$ is obtained by the partial similarity between the genre to be evaluated and other genres, as shown in the Figure(10).
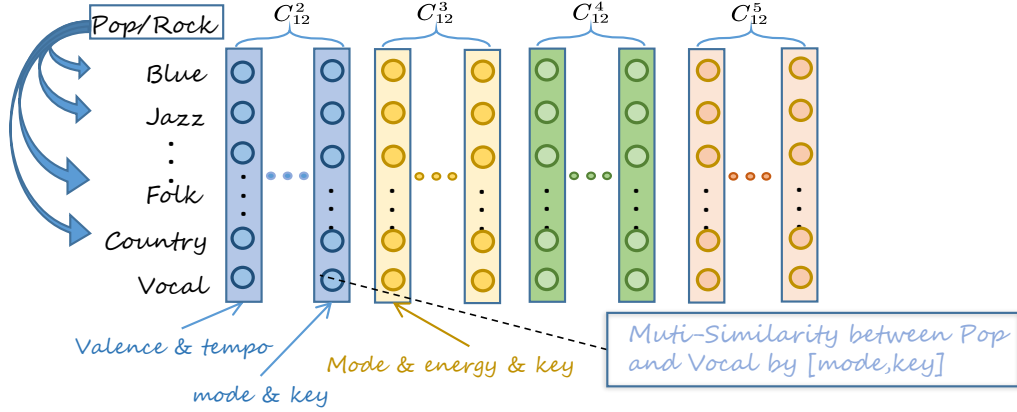

Figure 10: Partial Similarity Evaluation Vector

$\mathbf{S_{part\ i}}$ reflects the similarity of the genre between other genres when only the features contained within $V^{(i)}$ are considered. If the features contained within $V^{(i)}$ can well distinguish the genre from other genres, then these features can be considered as the significant differences between the music of the genre and other genres. Ideally, if the combination of features chosen can completely distinguish the genre from other genres, then the optimal partial similarity vector $\mathbf{S'_{part}} = [\frac{1}{2}, \frac{1}{2}, \cdots, \frac{1}{2}]^T$ is defined.

Cosine similarity and Euclidean similarity are used to measure the *discrimination*$_i$ of feature combination vector $V^{(i)}$

$$discrimination_i = S_{cos}(\mathbf{S'_{part}}, \mathbf{S_{part\ i}}) + S_E(\mathbf{S'_{part}}, \mathbf{S_{part\ i}}) \tag{20}$$

We selected two genres, Electronic and Latin, for analysis, and obtained the top three discrimination scores as shown in the Table(6).

Table 6: Feature Combination Discrimination

	Feature combination	Discrimination
Electronic	[acousticness, instrumentalness, speechiness]*	1.1588
	[energy, acousticness, instrumentalness, speechiness]*	1.1588
	[energy, acousticness, instrumentalness, liveness, speechiness]	1.1586
	Feature combination	**Discrimination**
Latin	[energy, valence, tempo, danceability]*	1.1474
	[energy, valence, loudness, danceability]	1.1473
	[energy, valence, tempo, danceability, liveness]	1.1470

As can be seen from the Table(6):

● **For Electronic:** Electronic is different from other genres in four aspects, including *acousticness*, *instrumentalness*, *speechiness*, and *energy*. These characteristics are in line with those of electronic-synth music, which has a strong sense of rhythm and is more likely to evoke the desire to dance.

● **For Latin:** Its concentrated characteristics are embodied in *energy*, *valence*, *tempo*, *danceability*, which coincide with the energetic, dynamic, rhythm-centered characteristics of Latin music.

4.4.2 Analysis of Genre Changes Over Time

• Analysis of single genre changes over time

Take Electronic's music characteristics change as an example. The analysis process is shown below:

Step 1: **Data visualization**

We use the four most distinguishing attributes of Electronic: *acousticness*, *instrumentalness*, *speechiness*, and *energy*, which are analyzed in 4.4.1. We add the *number of musicians* within the genre and the *popularity* to obtain six variables. The curves of these six variables over time are plotted as Figure(11).
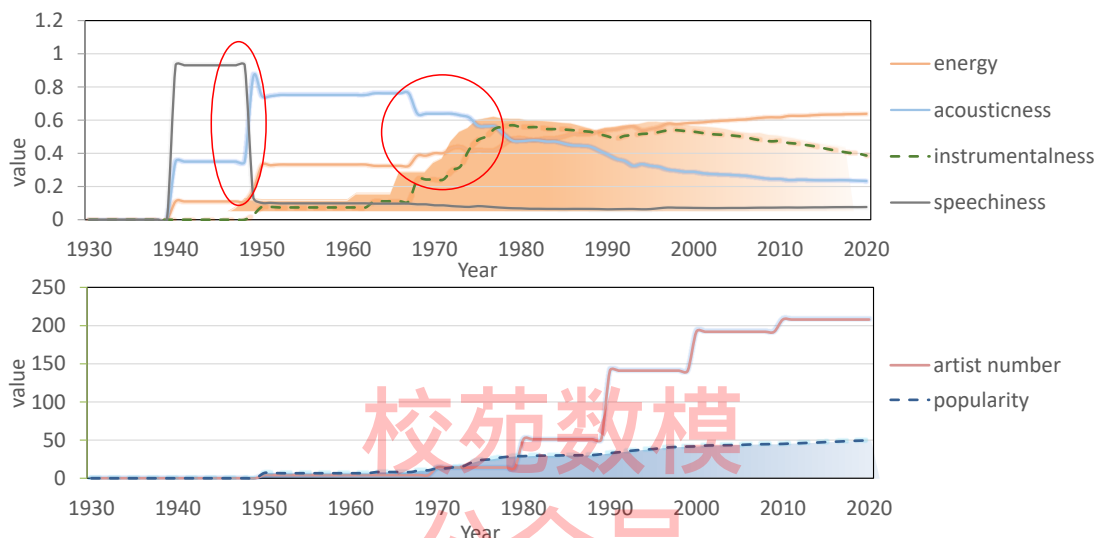

Figure 11: Changes of Electronic Over Time

Step 2: **Identification of mutation points**

From the Figure(11), we can see that electronic music mutated in the 1950s and characteristics changed from 1970 to 1980. In total, it can be divided into three development stages: in the 1950s, 1970~1980, and after the 1980s.

Step 3: **Analysis with historical facts**

1. In the 1950s, there were significant changes in the variety of attributes of electronic music, from 0.93 to 0.10, and the score for acousticness, instrumentalness, iTunes, and energy increased by more than 200%. After the 1950s, electronic music became more intense and electrically enhanced. According to Wikipedia[6], the first electronic music recorded on tape was composed by Louis and Bebe Barron in 1950, signifying the creation of true electronic music.

2. From the 1970s to the 1980s, there were a lot of innovations in electronic music equipment. Digital synthesizers and samplers replaced analog synthesizers, and more musicians could use a synthesizer for music creation.

3. After the 1980s, the Acid House movement in the UK pushed the development of electronic music into the mainstream, and between the late 1990s and the beginning of the 21st century, electronic music using only electronic instruments became increasingly popular, a trend that continues today. Nightclubs around the world are playing electronic music on a regular basis.

• Analysis of correlation of changes between genres over time

We use Easy listening and New age as examples to analyze the interplay of genres during

their change over time. We choose three characteristics of *acousticness*, *tampo*, and *speechiness* of these two genres to analyze how they change over time.
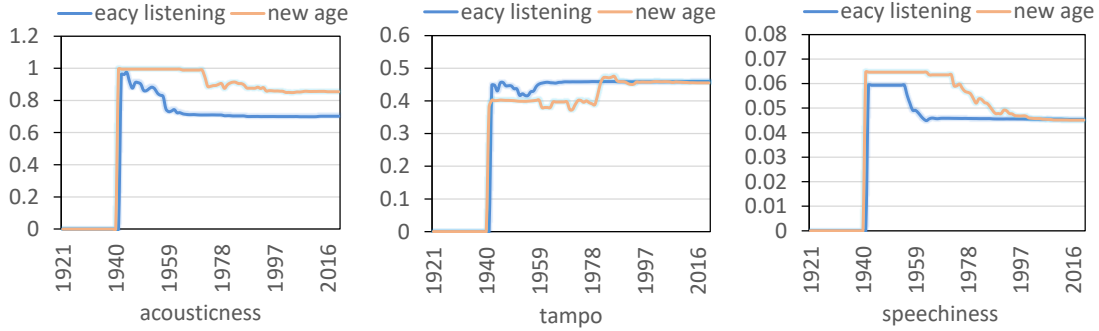

Figure 12: Changes of New age and Easy listening Over Time

As can be seen from the Figure(12), both Easy listening and New age have a similar trend: *acousticness* gradually decreases, *tempo* stabilizes, and *speechiness* continues to decrease on top of the already low values. This is also in line with the development history of the real Easy listening and New age, that is, the rhythm is soothing, there is less spoken word, and gradually introduced a small number of electronic instruments.

Exploring the sequential correlation between the changes of the two curves, we can also learn that the changes of New age always lag behind Easy listening and change in the direction of Easy listening changes. It can be further concluded that New age is influenced by Easy listening as it changes over time, which coincides with the fact that New age started later and is influenced by Easy listening.

5 Task 4 :Long-Short Term Influence Evaluation Model Based on Sliding Window

5.1 Establishment of Sliding Window Model

To explore whether influencers influence the music created by followers, we evaluated the similarity between the music of influencers and followers by *MS*. Considering that the influence of followers by influencers changes continuously over time, we add sliding windows[9] to explore the short-term similarity *MS* of music based on the similarity evaluation model established in TASK 2 & 3, to capture the short-term influence *STS* of influencers on followers' music. Finally, the magnitude of the change in short-term similarity is summed to obtain the long-term influence *LTS* of the influencer on the follower. An overview of the model is shown in the Figure(13).

We arrange the songs of influencers and followers by time, and as the window slides along the time line, new songs are added continuously on the right side, while some outdated songs are dropped on the left. We swipe two windows on the timelines of the works of influencers and followers at different times. Additionally, we ensure that the music in the influencer window is released earlier than the music in the follower window. Then calculate the similarity of influencer's music and follower's music *MS* in both windows.

In order to find the most similar songs (songs with real influence relationships), we define the similarity of the song in the i th follower window to the influencer (MS_i):

$$MS_i = max\{MS(Fl_i, In_j), \quad j = 1 \ to \ n\} \tag{21}$$

Fl_i and In_j are the music within the follower window i and the influencer window j, respectively. $MS(Fl_i, In_j)$ represents the short-term similarity of music within the influencer window j to music within the follower window i. To satisfy that the music within the influencer window is released earlier than the followers, n changes as i changes.
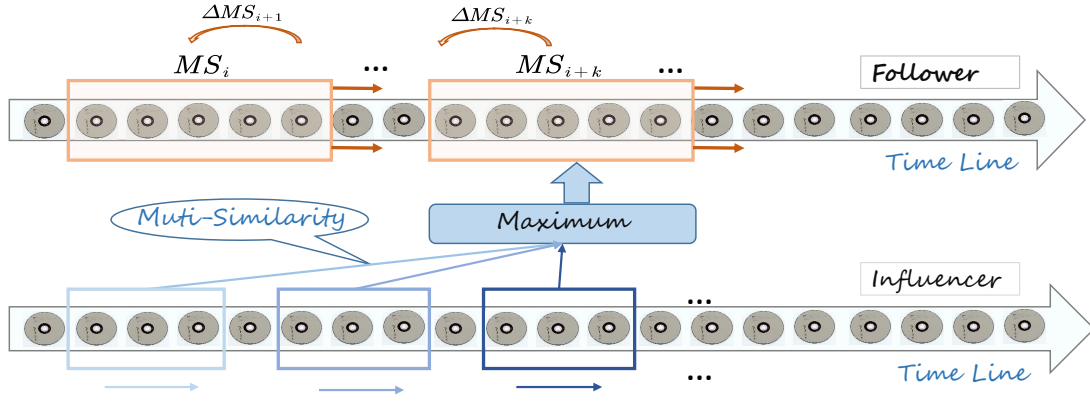
Figure 13: Long-Short Term Influence Evaluation Model Based on Sliding Window

Define the magnitude of similarity change $(\Delta MS)_i$ for the follower's timeline window:

$$(\Delta MS)_i = \frac{MS_i - MS_{i-1}}{MS_{i-1}} \tag{22}$$

Define short-term similarity (STS) and long-term similarity (LTS):

$$\begin{cases} STS = max\{(\Delta MS)_i\} \\ LTS = \sum(\Delta MS)_i \end{cases} \tag{23}$$

We consider that when $STS > 0.01$, i.e., in the short term, there is a period of time when followers' song styles are substantially closer to those of the influencers, and this is when the influencers can be considered to have influenced the followers. When $LTS > 0$, the follower's music style is generally close to the influencer's in the long run, and the influencer can be considered to have a long-term influence on the follower. Defining the degree of influence long-term-short-term score IDS, when $IDS > 0$, the influencer can be considered to have an influence on the follower.

$$IDS = max(0, LTS) + max(0, STS - 0.01) \tag{24}$$

5.2   Test of the Established Model

We set the window size to 5 and the move step to 3 for window sliding detection. The IDS between all influencers and followers are calculated using the model developed above and find that 86.08% of the IDS are greater than 0. It can be concluded that in general influencers do influence the music created by their followers.

We choose the influencer: Black Box, and his followers: Soda Stereo and Dave Matthews Band as an example, the results are shown in the Table(7).

Table 7: Result of Sliding Window Model

Influencer	Follower	LTS	STS	IDS
Black Box	Soda Stereo	0.00033	0.01251	0.000334
Black Box	Dave Matthews Band	0.00553	0.00549	0.00553

It can be seen that Black Box did influence the music making of its followers, Soda Stereo and Dave Matthews Band, to varying degrees.

● **For Soda Stereo:** STS≫LTS, the short-term impact is greater.Inspired by a few songs from an influencer and subsequently less influenced by the influencer.

● **For Dave Matthews Band:** LTS>STS, the long-term impact is greater. Dave Matthews Band has been influenced by influencers for a long time and is a loyal follower.

5.3 Contagious Evaluation of Musical Characteristics

Music has many different characteristics, and in order to explore whether there are more contagious (easy to spread) ones among different characteristics, we established a contagious evaluation system for music characteristics based on the timeline of influencers' and followers' works in the previous section, and the music similarity (*MS*) evaluation model in 4.2.1. For a specific musician (as a follower), the contagious evaluation steps of musical features $v^{(k)}$ are as follows.

Step 1: Build the timelines of the works of this musician (single) and his influencer (possibly multiple) respectively, as shown in Figure(13) (4.4.2).

Step 2: For a musician's song i, it is combined with an influencer's song j, and then calculate the difference $dif(i,j)^{(k)}$ of the musical characteristics $v^{(k)}$ of the two songs.

$$dif(i,j)^{(k)} = v_i^{(k)} - v_j^{(k)} \tag{25}$$

Step 3: Musicians' works may be influenced by only a relatively small number of enlightening works, so there should be the smallest feature difference between them on $v_{(k)}$. We choose the three smallest values of $dif(i,j)^{(k)}$ for weighted summation to obtain the difference $DIF_i^{(k)}$ between the follower's song i on feature $v^{(k)}$ and and its corresponding enlightening work. $dif_1^{(k)}, dif_2^{(k)}, dif_3^{(k)}$ for the first three smallest characteristic differences.

$$DIF_i^{(k)} = \frac{1}{2}dif_1^{(k)} + \frac{1}{3}dif_2^{(k)} + \frac{1}{6}dif_3^{(k)} \tag{26}$$

Step 4: Analyze the contagion of all the features within this musician's works. Define the degree of closeness $DC^{(k)}$ of all songs of a follower to the songs of its influencer on $v^{(k)}$. And $D[d]^{(k)}$ is the variance of all songs of this musician on the feature $v^{(k)}$. The smaller the final value of $DC^{(k)}$ is, the more susceptible the musician is to influencers on $v^{(k)}$, and the more contagious the $v^{(k)}$ is.

$$DC^{(k)} = ((1 + average(DIF_i^{(k)})) \cdot (1 + minimum(DIF_i^{(k)})) \cdot D[d]^{(k)} \tag{27}$$

Using the above analysis process, Soda Stereo and Dave Matthews are selected as followers for the contagious analysis of music characteristics, and the results are obtained as shown in the following Table(8).

Table 8: Contagious of Musical Characteristics

	Characteristics	*speechiness*	*energy*	*instrumentalness*
Soda Stereo	*DC*	0.000249	0.0068	0.007554

	Characteristic	*speechiness*	*instrumentalness*	*danceability*
Dave Matthews	*DC*	8.5279×10^{-5}	9.6721×10^{-5}	0.012702

As can be seen from the Table(8), not all characteristics are highly contagious to a particular musician, and the most contagious characteristics vary from one musician to another. For Soda Stereo, the musical characteristic *speechiness* is the most contagious. For Dave Matthews, the musical characteristics of *speechiness* and *instrumentalness* are the most contagious.

6 Task 5 & 6 & 7: Analysis of Music Revolution Over Time

6.1 The Identification of the Musical revolution

We analyze the given data to identify musical revolutions in terms of two symbol: changes in existing musical genres, the creation of new musical genres.

● **Changes in existing musical genres**

Dramatic changes in musical characteristics within existing musical genres can be recognized as temporal musical revolution. The following are the conditions for determining the point in time when an evolution occurs:

A. The musical characteristics of the genre remained stable before and after the revolution.

B. In the year of the revolution, the musical characteristics of the genre change greatly.

C. There are at least two features that have taken a significant turn at the same time.

The above conditions are quantified using mathematical formulas to explore whether a significant change occurs within a six-year period of analysis. For the characteristic $v^{(k)}$ of the genre, the condition for determining that an evolution occurs in the $3_{th} \sim 4_{th}$ year of the analysis cycle is defined by the equation(28)

$$\begin{cases} |v_3^{(k)} - v_2^{(k)}| < 0.03 \\ |v_2^{(k)} - v_1^{(k)}| < 0.03 \\ |v_4^{(k)} - v_3^{(k)}| > 0.1 \\ |(v_5^{(k)} - v_4^{(k)}) - (v_4^{(k)} - v_3^{(k)})| < 0.02 \ or \ |v_5^{(k)} - v_4^{(k)}| < 0.03 \\ |(v_6^{(k)} - v_5^{(k)}) - (v_5^{(k)} - v_4^{(k)})| < 0.02 \ or \ |v_6^{(k)} - v_5^{(k)}| < 0.03 \end{cases} \qquad (28)$$

Using a six-year time span as the width of the sliding window, a sliding window model is used for analysis to identify the years in which the upper constraint is satisfied, i.e., the time when a revolution occurs. If the trend of change continues in the years following a turning point, we consider a long-term change to have occurred during that period. *Electronic* is selected as the study object to explore the characteristics of significant revolution in *Electronic*. With the above constraint equation, we capture the year 1950 and 1969 as the time points of the revolution. And the characteristics of Electronic in 1950 and 1969 are shown in the Table(9).

Table 9: Identification of Revolution of Electronic

Year	1950			1969	
Characteristic	*speechiness*	*acousticness*	*energy*	*instrumentalness*	*acousticness*
Change	0.796	0.513	0.210	0.131	0.123

From the Table(9), it can be seen that Electronic in 1950 *speechiness*, *energy* and *acousticness* changed a lot. In 1969 *instrumentalness* and *acousticness* changed a lot. This is also similar to the conclusion of our analysis of Electronic in 4.4.2, and it can be considered that the established constraint equations can identify the revolution in the existing music genres.

● **Creation of new musical genres**

To identify the new generation of music genres, we visualize the number of musicians in each genre over time in Figure(14). Each surge in the number of musicians in each genre can be considered as an evolution in the genre. The first surge in the number of musicians in each genre can be identified as the creation of a new genre.

Observe the Figure(14), using the analysis of R&B and Country as an example:

1. The first explosion in the number of musicians in R&B occurred in 1950. There had not been an explosion in the number of musicians in the genre before that, so it can be argued that an evolution happened in 1950, when R&B was created. This also coincides with the start date of the R&B genre, according to Sound and Music from Xinhua News Agency[7].

2. The first surge in the number of musicians in Country occurred in 1920, and it is also logical to assume that Country had an evolution in 1920, when Country was born. The results
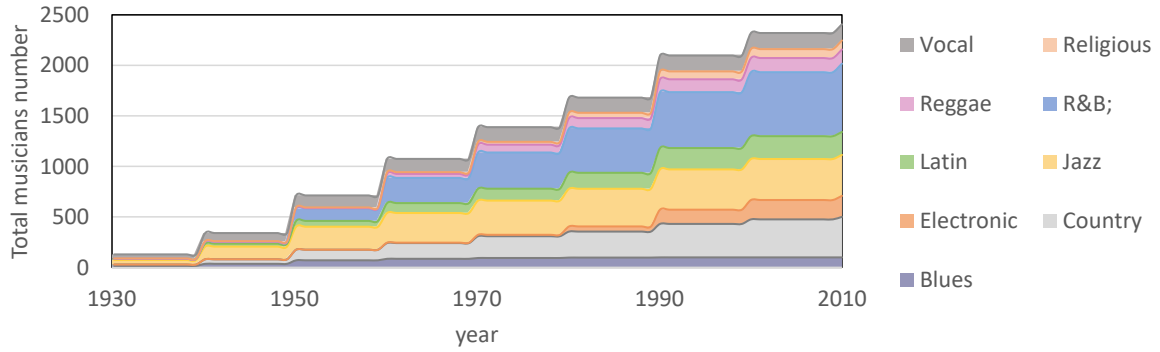
Figure 14: Changes in the Number of Musicians Over Time

of the analysis are also consistent with the facts[7].

6.2 Major Influencers of Revolution

For a genre revolution, we analyze the music published within the genre in the two years before and after the evolution, looking for the most popular artists during that part of the time. We sum the popularity of each musician's works released during that period separately to obtain the popularity of each musician during that period, with the highest being the most influential in the evolution. Electronic is analyzed using the above method.

In 1950, **Freestyle**'s music has the highest sum of popularity: 144, can be considered as a major influencer on the 1950 evolution. In 1969, **Jean-Jacques Perrey**'s music has the highest sum of popularity: 139, can be considered as a major influencer for the 1969 evolution.

The trend of evolution in 1969 was maintained from the beginning of the evolution in 1969 to 1978, and we believe that a long-term change was produced during this period. We also identified the major influencers near 1969 to 1979 as **Kraftwerk** : sum of popularity=822, **Yellow Magic Orchestra**: sum of popularity=495, **Vangelis**: sum of popularity= 473.

6.3 Dynamic Influencer Evaluation Indicator

In the equation(24) in 5.1, we define *IDS* to measure the degree of influence that influencers have on their followers. As genres develop over time, there may be different key influencers within the genre at different times who will influence other musicians within the genre and lead the direction of the genre. To reveal the dynamic influencers within genre *i*, we define the Dynamic Influencer Indicator *DII* as follows.

For musician *i*, first calculate its degree of influence $IDS_{i,j}$ on all musicians within the genre in turn and sum up *IDS*, to obtain the dynamic influencer index DII_i for musician *i*:

$$DII_i = \sum_{j=1}^{n} IDS_{i,j} \tag{29}$$

Within this genre, the musicians are ranked according to the *DII* of each musician. From the analysis of the percentage of major influencers in 3.3, a point threshold of 20% is used based on the law of Pareto's Principle[4]. Musicians with *DII* ranking in the top 20% are then identified as dynamic influencers within this genre. We take Pop/Rock as an example, taking ten years as the analysis period, the *DII* of musicians within the genre of every period is calculated, and we list the two largest *DII* after normalization of musicians every decade shown in the Table(10).

As can be seen from the Table(10), the dynamic influencers within Pop/Rock are not the same in different eras, which is a good indication of the turnover of leading figures within the

Table 10: *DII* of Musicians Over Time

| Year	1950s	1960s	1970s	1980s	1990s
Musician	Elvis Presley	Bob dylan	The Clash	Sex Pistols	Blur
DII	0.9951	0.9926	0.9635	0.8964	0.9131
Musician	Little Richard	Beatles	David Bowie	Nirvana	Beach Boys
DII	0.9814	1	0.9804	0.8413	0.8971

genre. The recognition threshold of *DII* is 0.8127, and musicians above this threshold can be identified as dynamic influencers.

6.4   Influence from Social, Political and Cultural Aspects

Changes in musical genre attributes often reflect the social, political, and cultural aspects of the time to some extent. We explore the social and cultural changes of the time by analyzing the changes in musical attributes such as *valence* related to emotions and *acousticness* related to the use of science and technology in the music influence network across genres.

- **From the general trends in the music industry**

Taking the analysis of the emotional tendency of music as an example, we count the emotional tendency *valence* statistics of several genres of music that account for a relatively large proportion of the genre as shown in the Figure(15).
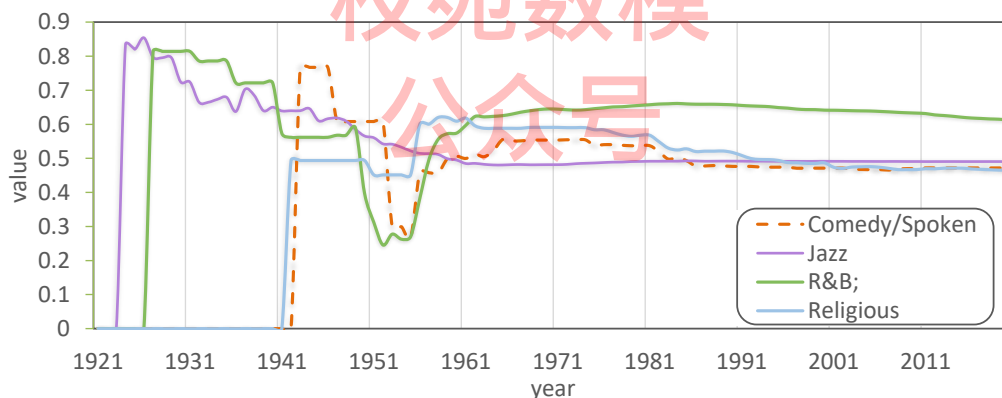
Figure 15: Changes of *valence* Over Time

From the Figure(15), we can see that the emotional tendency of most genre songs started to become negative in 1947 and then became positive after 1965. Combined with the historical fact that the African-American Civil Rights Movement started in 1950 and ended in 1970, African-Americans were oppressed by society prior to the Civil Rights Movement, and the musicians used music to express their long-held negative feelings. After the civil rights movement, the future of African-Americans became clearer and they used music to express their desire for a better life. Thus we can consider that the emotional tendencies of music can, to some extent, reflect the implicit tendencies of society, politics, and culture, and be used to predict social revolution.

- **From trends in specific music genres**

Take the analysis of Electronic characteristics over time in 4.4.2 as an example. As shown in the Figure(11) in 4.4.2, the *acousticness* of Electronic increased significantly in 1950, and so did the *energy*. It can be considered that Electronic has become more energetic and dynamic, which is related to the fact that people were in the ecstasy of achieving victory in the anti-fascist

war after the Second World War in the 1950s. The cultural products of this historical period also fully reflect the optimistic and exciting social trends.

- **From the listener's reaction to the music**

In addition to the music itself, the connection between music and social development can also be glimpsed in the changes in listeners' reactions to music. The trend of *popularity* in each category over time is plotted in Figure(16). It can be seen that there are more and more music listeners, which reflects the unprecedented increase in people's ability to access information through the Internet as a result of the development of science and technology, making music accessible to a wider audience. This will further increase the influence of music on society, politics, and culture.
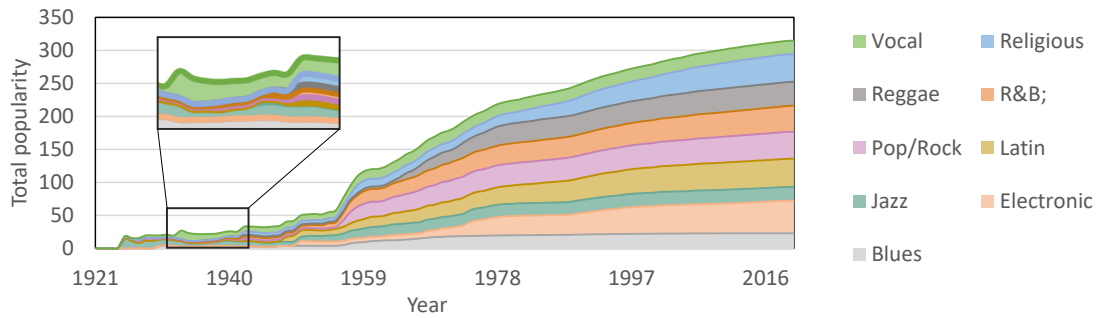

Figure 16: Changes of *popularity* Over Time

7 Sensitive Analysis

7.1 Sensitive Analysis of Model in TASK 2

In Task 2 to build the Multi-dimensional Music Similarity Evaluation Model, we weight the local similarity S_{part} and the global similarity S_{global} to obtain *MS*:

$$MS = ((1-\delta)S_{global} + \delta S_{part}) \cdot (1 + \frac{0.1}{S_{time}}) \tag{30}$$

The different weights δ may affect the evaluation results of similarity. Therefore, we choose 12 pieces of music from 3 genres to study the effect of different δ on the results, and choose $\delta = 0.2, 0.5, 0.8$ to obtain the results of *MS* as Figure(17).

Looking at the above Figure(17) we can see that a change in δ does not change the broad similarity results respectively. However, the number of the parameter can have some effect on some specific categories of music, resulting in changes in musical similarity between or within genres, respectively.

- When δ is small, we can see higher similarity between genres. And S_{global} has a large weight and pays more attention to the differences of songs in all characteristics, which leads to lower similarity within the genre.

- When δ is large, it can be seen that the similarity between genres is lower. S_{part} has a larger weight and pays more attention to the differences in important features of songs, which makes the the similarity within the genre become higher.

Therefore, we choose $\delta = 0.5$ in the model of the main text to make our similarity evaluation more balanced between local and global features.

7.2 Sensitive Analysis of Model in TASK 4

In the sliding window model in Task 4, the size of the window and step size may lead to different results. We use a variance-based sensitivity analysis method (Sobol method)[8] to
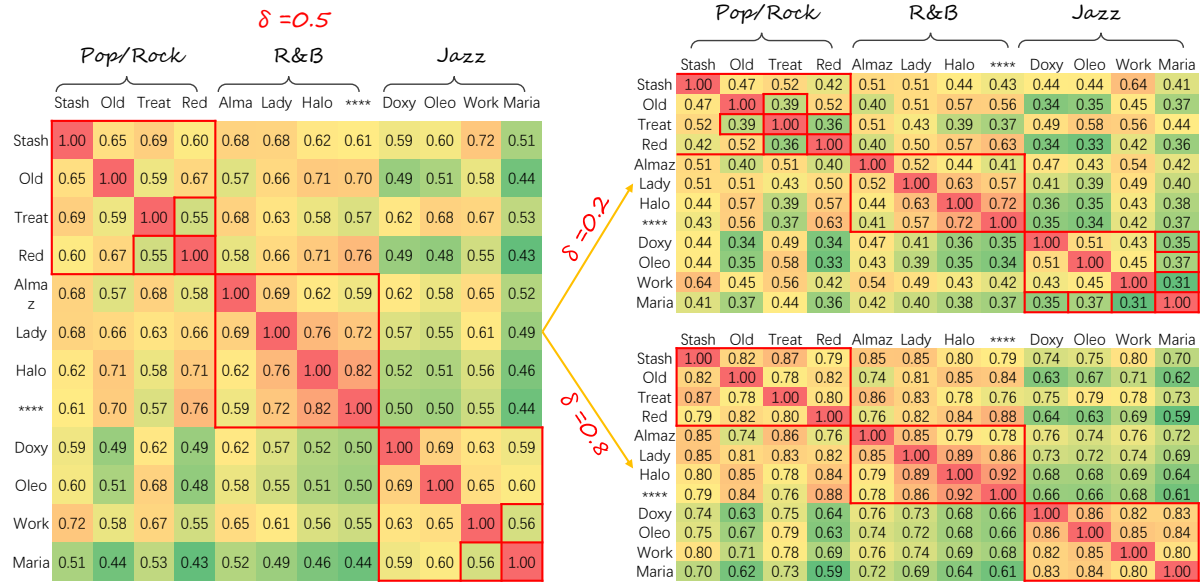
Figure 17: Sensitive Analysis of the Model in Task 2

investigate the effect on the results when multiple hyperparameters are varied simultaneously. The Sobol method simulates Monte Carlo sampling of the parameters in the feasible space and obtains a large number of outputs. The variance of the model output is the sum of all order variances of a single parameter. The equation for the contribution of the computed variance of the Sobol method is as follows:

$$V(Y) = \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} V_{ij} + \cdots + V_{1,\cdots,p}$$

$$S_{T_i} = \frac{E_{X_{\sim i}}(Var_{X_i}(Y|X_{\sim i}))}{Var(Y)} = 1 - \frac{Var_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i})}{Var(Y)}$$

(31)

The above equation measures the proportion of variance change caused by the interaction between one variable and other variables that change simultaneously. We then use the Sobol method on the sliding window model to study the sensitivity of window size and step size and show the sensitivity (S_T) as Figure(18).
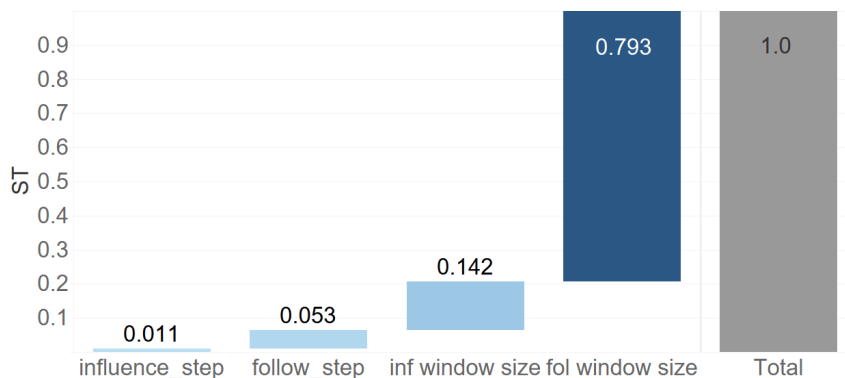

Figure 18: Sensitive Analysis of the Model in Task 4

It can be seen that the change in the size of the window of followers has the greatest impact on the results, which can indicate that the number of music selected in the calculation of short-term influences has the greatest impact on the final influence results. This represents that for different musicians, when influenced by an influencer, different musicians may reflect the

influence of the influencer on several consecutive different songs. Some musicians may be inspired only in the short term with only a small time period of songs reflecting this influence. While others may be inspired in the long term, with multiple songs over a long period of time reflecting the influence of an influencer on them.

8 Strengths and Weaknesses

8.1 Strengths

• We used the theory of SNA to construct a network of influences between musicians and to measure the internal parameters of the network with more well-considered metrics.

• In analyzing whether influencers actually have an impact on followers, we used a sliding window time series approach to synthesize the short- and long-term impact of influencers on followers.

• We have done sufficient visualization work on the algorithms and results of the model to facilitate understanding.

• Most of our models are built with innovative ideas or improved based on original theories and specific problems, making our models both theoretical and innovative.

8.2 Weaknesses

• We could also go deeper into the analysis of certain issues, but due to space limitations, we can only stop here.

• Due to the lack of actual numerical reference, we can only observe the accuracy of our model from the actual situation and cannot numerically measure the accuracy of our model.

References

[1] Zhang Cungang, Li Ming, Lu Demei. *Social Network Analysis: An Important Sociological Research Method* [J]. Gansu Social Sciences, 2004.

[2] Haveliwala T H . *Topic-sensitive PageRank: a context-sensitive ranking algorithm for Web search*[J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(4):p.784-796.

[3] *The 20 Most Representative Musicians in the World Jointly* selected by CNN and Songlines from `http://edition.cnn.com/2010/SHOWBIZ/Music/08/04/your.music.icon/index.html`

[4] Chen Dao-lan. *The Discriminatory Relationship Between the Law of Two and Eight and the Long Tail Theory[J]*. Talent, 2010(15):280.

[5] LIU Yan-xia, XIAO Wen-jun, XI Jian-qing, et al. *Properties of Scale-free Networks with Power Index Less than 2* [J]. Journal of Nanchang University (Science Edition), 2013, 37(004):350-354.

[6] *Electronic Music*.Wikipedia `https://en.wikipedia.org/wiki/Electronic_music`

[7] *Sound and Music – 1890-2010*, World Pop Music Echoes (news.cn) from the Xinhua News Agency `http://fms.news.cn/swf/2016_sjxw/yqfm_2016/index.html`

[8] Zadeh, Farkhondeh Khorashadi,Nossent, Jiri,Sarrazin, Fanny, et al.*Comparison of Variance-based and Moment-independent Global Sensitivity Analysis Approaches by Application to the SWAT Model* [J].Environmental modelling & software,2017,91(May):210-222.

[9] Tan H Q, Niu Q Q. *A Time Series Symbolization Method Based on Sliding Window and Local Features*[J]. Computer Application Research, 2013, 30(003):796-798.

MEMO

To: Experts in The ICM Society
From: Team 2107091
Date: February 8th, 2021
Subject: Analysis of Music Influence and Similarity, Suggestions for Futher Study

Dear experts in the ICM Society:

We are honored to inform you our achievement after performing data analysis and establishing the music influence and similarity evaluation model.

Our team conducted a modeling analysis of the interactions between and within musicians across genres from 1921 to 2020. We develop a directed network for analyzing the interactions between musicians to understand the interactions that exist between musicians. An evaluation model to measure the similarity between the music is developed to provide an evaluation of the similarity between and within genres and to give the distinctive features of music in each genre. An evaluation model based on a sliding window is developed to select the leaders of each genre in different time periods. With the model we developed, we obtained the following valuable conclusions:

• The influence of musicians on each other is consistent with Pareto's Principle, which states that a very small number of musicians have influenced the vast majority of musicians. We can consider that a very small number of outstanding musicians led the entire music world, and in-depth exploration of the works of these musicians may deepen our understanding of music.

• Although there are differences between the genres, they are not completely independent but have no clear boundaries. As a carrier of human emotions, music has similarities among genres. The study of the evolutionary process of musical alternation and fusion among genres is conducive to enhancing our perception of the evolution of music as a whole.

• Changes in music can clearly reflect social change. Music is a tool for people to express their feelings, and by monitoring changes in musical emotions, we can probe into the pent-up emotions in society and thus make predictions about social revolution.

The model we developed still has limitations due to incomplete data and space limitations. We also need more data to achieve the following objectives

• We found that some of the artists are not common to both datasets, and some of the artists' debut dates are unknown, which causes some inaccuracy in our analysis. We hope to obtain a more complete dataset of artists and music.

• If we know the specific year in which the followers were influenced by the influencer, then we can refine the model for determining the contagiousness of music and whether it actually produces change by the change in the music of the follower before and after the year of influence. Also, the dynamic impact indicators of influencers can be refined based on the change in the number of influenced people.

• If we know the album information where the song is located, then we can modify our sliding window model to capture the short-term impact by directly comparing the similarity between albums.

We expect to build a dynamic network in the future in the dimension of time and genre.We also hope to make a more detailed division in genre and to develop a set of visualization system in the future, to more describe the internal development. The above are the results of our team's research, and we sincerely hope they can provide you with useful information and data.