Problem Chosen	2022	Team Control Number
D	MCM/ICM Summary Sheet	**123456**

How Does Musical Influence Lead to Musical Revolution?

Summary

Music has played an important role in human societies since the beginning of time. And there are many great musicians who have promoted the development of music greatly. As part of an effort to understand the role music has played in the collective human experience,it is supposed to develop a method to quantify musical evolution.To understand more about music and musicians, a lot of research is done.

In TASK 1: First, we take each musician as a node and add edge between musicians by examining if there is influence relationship to build a directed music influence network. Then we use social network analysis (SNA), more specifically, random walking, to analyze the influence network. The Degree Centrality and the Pagerank are calculated. The distributions of both of them are calculated, and it can be seen that they both obey power-law distribution. The person coefficient and spearman coefficient are calculated, revealing that these two indices are not related. So to quantify the music influence, Entropy Weight Methods are used to get the weight of these two indices. Thus, their linear combination is considered to be the music influence. It is found that Stray Cats, The Beatles, Bananarama have the highest music influence. The conclusion is that the network is scale-free, and the average-shortest-path-length is close to 6 (little world characteristic). By detecting the communities, 76 communities are found, with an average of 23 members in each of them.

In TASK 2 and 3: We define music and distance to measure its similarity. According to the literature, distance includes the part based on attribute and the part based on musician. Based on the distance of attribute, a variety of attributes are fitted by entropy weight method to get the calculation results. Based on the Distance of musician, Hausdorff Distance is introduced to get the final distance by combining the two. Considering the similarity within a genre or between genres, we find that musicians in a genre have a higher similarity than between genres.

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.

Keywords: PageRank, Entropy Weight Methods, Random Walking, Weighted Euclidean Distance, Slide Window

Contents

1 Introduction

1.1 Background Research

MUSIC has been part of human societies since the beginning of time as an essential component of cultural heritage. As part of an effort to understand the role music has played in the collective human experience, we have been asked to develop a method to quantify musical evolution. There are many factors that can influence artists when they create a new piece of music, including their innate ingenuity, current social or political events, access to new instruments or tools, or other personal experiences. Our goal is to understand and measure the influence of previously produced music on new music and musical artists.

Some artists can list a dozen or more other artists who they say influenced their own musical work. It has also been suggested that influence can be measured by the degree of similarity between song characteristics, such as structure, rhythm, or lyrics. There are sometimes revolutionary shifts in music, offering new sounds or tempos, such as when a new genre emerges, or there is a reinvention of an existing genre (e.g. classical, pop/rock, jazz, etc.). This can be due to a sequence of small changes, a cooperative effort of artists, a series of influential artists, or a shift within society.

Many songs have similar sounds, and many artists have contributed to major shifts in a musical genre. Sometimes these shifts are due to one artist influencing another. Sometimes it is a change that emerges in response to external events (such as major world events or technological advances). By considering networks of songs and their musical characteristics, we can begin to capture the influence that musical artists have on each other. And, perhaps, we can also gain abetter understanding of how music evolves through societies over time.

1.2 Problem Restatement

Task 1 requires us to build a complex network between influencers and followers based on *influnce_data* data set and develop metrics to capture the music influence in the network. The key to this problem is to define directed influencer network and propose metrics that comprehensively measure the music influence of each influencers in the network.

Task 2 requires us to use various musical characteristics in *full_music_data* data set to measure music similarity and judge whether musicians in the same genre are more similar than those in different genres. It is of great essence to define the distance between music works of artists, from which we can obtain the similarities between artists.

Task 3 requires us to compare similarities and influences between and within genres. It is necessary to define the distance between genres and measure similarities between them. We plan to build a classification tree to distinguish one genre from others. Besides, we can show how genres changing over time and the relationship between genres using visualization.

Task 4 requires us to get a further insight into the similarity of music characters between influencers and followers. We plan to build a Bayesian Network to find the real influencer and use hypothesis test to evaluate whether some music characteristics are more 'contagious' than others.

Task 5 requires us to identify the characteristics and major artists that signify music evolution. To handle this problem, we propose a Change Point Detection model based on DP algorithm and match the change of music characters with the musical revolution. Then find revolutionaries via Bayesian Network.

Task 6 requires us to develop indicators to analyse dynamic influencers and corresponding influence processes of musical evolution in one genre. In this problem, we will focus on Pop/Rock and then discuss the major contributors to the Pop/Rock evolution through decades.

Task 7 requires us to identify the effects of social, political, or technological changes within the network and find the cultural influence of music. Based on time series analysis, we intend to find the connection between the change in music characteristics and the external events. We also detect several culture-influence of music in different time or circumstances.

2 Preparation of the Model

2.1 Notations

We'd like to begin by defining a list of notations used in the paper as Table 1.

Table 1: Notation

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

2.2 Assumptions and Justifications

To simplify the problem and make it convenient for us to simulate real-life conditions, we make the following basic assumptions, each of which is properly justified.

• 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.

3 Musical Influence Network Analysis

3.1 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 in which we mainly concerned about the degree distribution of the network. To measure the music influence of each musician, we introduce the parameter PageRank based of PageRank algorithm. Finally, we explore the sub-networks of the influence network by the means of GN algorithm.

3.2 Directed Network of Musical Influences

3.2.1 Network Establishment

We consider influencers and followers as nodes and collect all musicians together to make up the set $V = \{v_i\}_{i=1}^{n}$. If artist i has an influence of artist j, then an edge form node i to node j will be generated. All the edges make up the set $E = \{e_i\}_{i=1}^{n}$, and the edges form node i make up set $N(i)$. Each edge has its weight $w\{ij\}$ to measure the influences from node i to node j. Based on the *influence_data* data set, we built a complex network involving $n = 999$ nodes and $m = 9999$ edges in total, as shown in Figure 2.

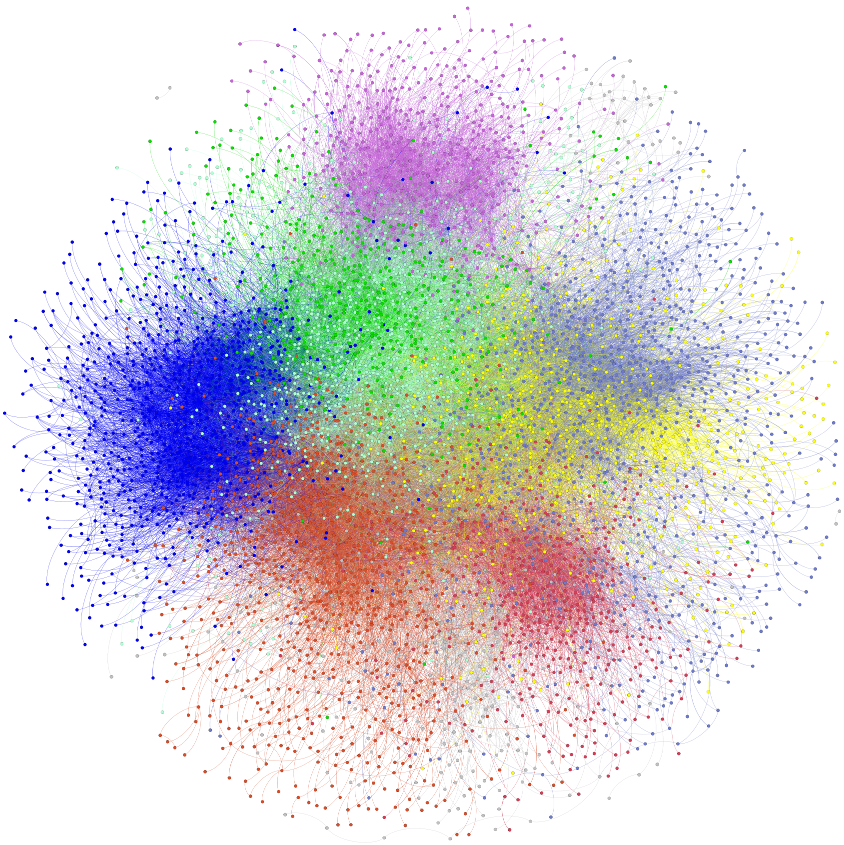


Figure 1: Directed Network of Musical Influences

Considering the quantity of data, we also explore the influences among genres, as shown in

Figure 3. The more the influence is, the thicker the edge will be.

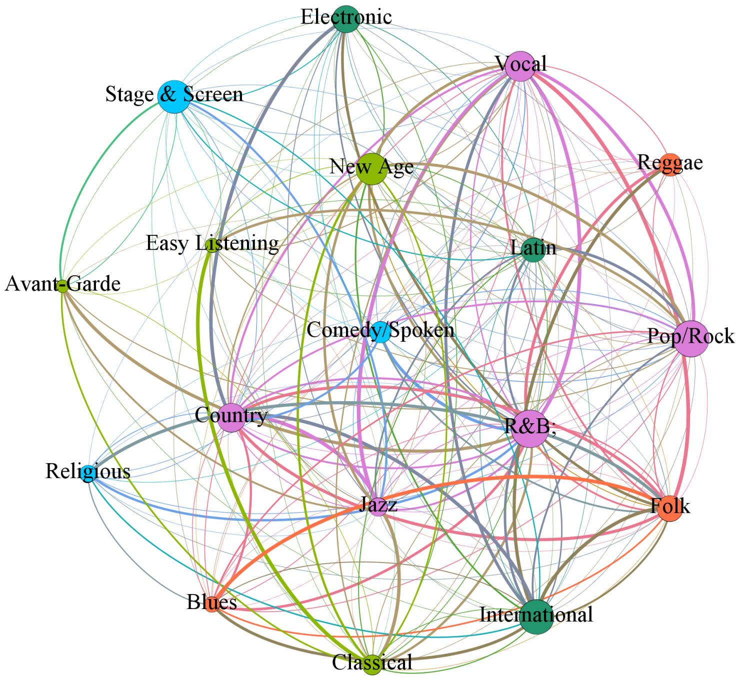


Figure 2: Influence Network among Genres

3.2.2 Degree Distribution Analysis

Degree distribution is an important parameter which describes how much nodes are the main influencer or followers and how much nodes have little influence.

Define the probability that the network medium is k to be

$$P(k) = \frac{n}{N} \tag{1}$$

where the total number of nodes is N and the number of k-degree node is n.

We find the distribution of degree is a good fit for the power law distribution. We use a fitting method to prove that. Define the fitting function

$$y = \alpha_0 x^{-\alpha_1} \tag{2}$$

Respectively fit the In-degree and Out-degree, and the calculation result is shown as Table 2 and Figure 4.

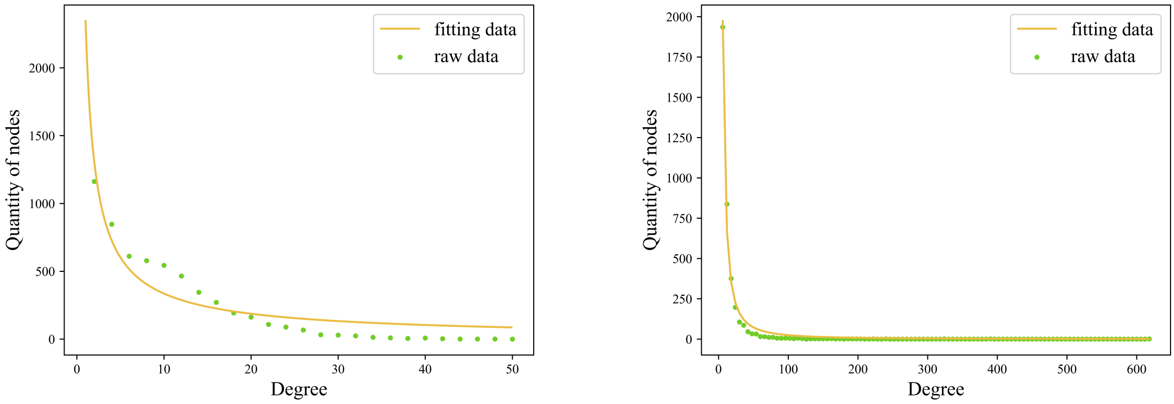
3.3 Evaluation Model of Musical Influence

3.3.1 PageRank

According to the literature, we choose PageRank as one of the parameter to describe the musical influence of each node in the network.Given an arbitrary directed graph with n nodes,

Table 2: Result of Fitting

In-degree		**Out-degree**	
α_0	32200.04	α_0	2345.77
α_1	1.5582	α_1	0.8446
Goodness of Fit	0.7460	Goodness of Fit	0.9772



(1)$Indegree$ (2)$Outdegree$

Figure 3: Result of Fitting

a general random walk model is defined on the directed graph, that is, a first-order Markov chain. The transition matrix of the general random walk model is composed of two linear combinations. One is the basic transition matrix M of a directed graph, representing the equal probability of transition from one node to all nodes it connects. The other part is a completely random transition matrix, which represents the transition probability to be $1/n$ from any node to any node. The linear combination coefficient is the damping factor d(here, as usual, we define $d = 0.85$). This general random walk Markov chain has a stationary distribution, denoted by R, and define

$$R = dMR + \frac{1-d}{n}1 \tag{3}$$

to be the PageRank of the directed graph. where 1 is is a n-vector whose components are all 1. Based on random walking, the more the in-degree of a node is, the higher the PageRank will be, representing the higher influence of this node in the network.

As the definition of PageRank is constructive, that is, the definition itself gives the algorithm, here we list the calculation method with power as we use, in Algorithm 1.

3.3.2 Evaluation Model

As PageRank focuses mainly on node in-degree, and the influential information represented by node out-degree is important and indispensable, therefore, we consider the linear combination of PageRank and Out-degre, as our final evaluation index Score

$$Score_{ij} = w_1 PageRank_{ij} + w_2 Outdegree_{ij} \tag{4}$$

where w_1, w_2 is the weight of the two parameters. The higher the Score, the more influential the music. According to the literature, we use entropy weight method to identify them.

Algorithm 1: PageRank

Input: Directed graph with nodes n, transfer matrix M, damping factor d, initial vector x_0, accuracy ϵ.
Output: PageRank matrix R of the directed graph.
Process:
1: $t = 0$, choose the proper initial vector x_0.
2: Calculate the general transition matrix. A

$$A = dM + \frac{1-d}{n}E$$

3: Iterate over and normalize the result vector

$$y_{t+1} = Ax_t$$
$$x_{t+1} = y_{t+1}/\|y_{t+1}\|$$

4: **if** $\|x_{t+1} - x_t\| < \epsilon$, $R = x_t$, **Output**
5: **else** $t = t + 1$, **return Step 3**

3.3.3 Entropy Weight Method

In information theory, entropy is a measure of uncertainty. The greater the uncertainty, the greater the entropy, the more information contained; The less uncertainty there is, the less entropy there is, the less information there is. Usually, the entropy H is defined as

$$H(U) = -\sum_{i=1}^{n} P_i \log P_i \tag{5}$$

where U is the set of all possible events including n values $U_1, U_2, \ldots, U_n$, and the corresponding probability is $P_1, P_2, \ldots, P_n$.

According to the characteristics of entropy, the randomness and disorder degree of an event can be judged by calculating the entropy value, and the dispersion degree of an index can also be judged by using the entropy value. The greater the dispersion degree of an index, the greater the influence (weight) of the index on the comprehensive evaluation.

With entropy weight method, we first do the Z-score normalization for each index x_{ij} by

$$x'_{ij} = \frac{x_{ij} - \overline{x_j}}{S} \tag{6}$$

then calculate the weight of the j index for the i object

$$y_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{m} x'_{ij}} \tag{7}$$

calculate the information entropy of the j index

$$e_j = -K \sum_{i=1}^{m} y_{ij} \ln y_{ij} \tag{8}$$

where $K = 1/\ln m$, and the weight of the j index w_j is

$$w_j = \frac{1 - e_j}{\sum_j 1 - e_j} \tag{9}$$

Table 3: Weight of each index

Index	Weight
$PageRank$	0.6710
$Outdegree$	0.3290

Here, using given data to calculate the weight of PageRank and Out-degree, the result is shown in Table 3.

At this point, we can calculate the score of each musician through PageRank and Out-degree, and measure the influence of each musician according to the score, by

$$Score_{ij} = 0.6710PageRank_{ij} + 0.329Outdegree_{ij} \tag{10}$$

and we will show the detailed result and analysis in Section 3.3.4.

3.3.4 Result in Musical Influence Network

With Formula (10), we calculate the Score for all the musicians to evaluate their musical influence, and the result is shown in Figure 4 and Table 4.

Table 4: Result of Evaluation

Name	Score	Rank
Stray Cats	0.6715	#1
The Beatles	0.3307	#2
Bananarama	0.2310	#3
Johnny Lee	0.2189	#4
Fat Freddy's Drop	0.2180	#5
Bob Dylan	0.2081	#6
The Doors	0.1991	#7
The Rolling Stones	0.1713	#8
Ramones	0.1371	#9
Captain Beyond	0.1366	#10
...	...	...

According to the results, only a few musicians have high influence, and the influence of most musicians is distributed in a lower range. The overall distribution of influence conforms to the power-law distribution, which is consistent with our general understanding.

3.4 Sub-network Exploration

3.4.1 Random Walking Methods

There are many smaller networks in this huge directed network. They are considered to be influential sub-networks. Among them, most nodes of a subnet belong to the same community, or multiple communities. Nodes belonging to the same community are often more closely connected than nodes in other communities. We consider the following basic static geometric characteristics of the network.

○ **Degree distribution**

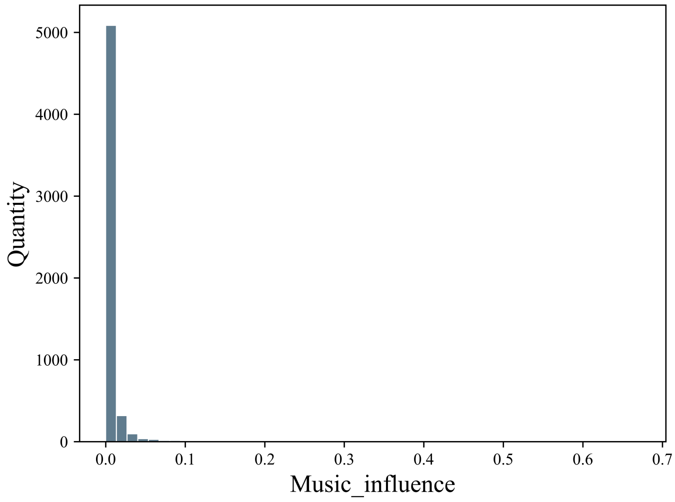


Figure 4: Influence Distribution

The average degree of all nodes in the network is called the average degree of the network.

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i \tag{11}$$

○ **Average path length**

The average path length L of the network is defined as the average value of the distance between any two nodes.

$$L = \frac{1}{C_N^2} \sum_{1 \le i \le j \le N} d_{ij} \tag{12}$$

○ **Clustering coefficient**

The ratio of the actual number of edges and the total number of possible edges between the k neighbor nodes of a node is defined as the clustering coefficient of the node.

$$C_i = \frac{E_i}{C_2^{k_i}} \tag{13}$$

The clustering coefficient of the whole network is the average of the clustering coefficients of all nodes.

3.4.2 Result of Sub-network

So far, we choose pagerank and out-degrees as the parameters of "music influence". As shown in the pictures, these communities with high connectivity tend to have higher degree distribution, lower average path length and higher clustering coefficient, which is consistent with the small world network model. We note that some famous musicians play an important role in their communities. Besides, we found some typical topological structures of some communities in large networks, such as star coupling networks.

Table 5: Result of Sub-network

	Flow Centrality	Average Path Length	Cluster Coefficient
1	4.0	1.558	0.031
2	4.125	1.132	0.017
3	2.211	1.16	0.008
4	3.3	1.623	0.262
5	3.5	1.176	0.042
6	2.0	1.542	0.0

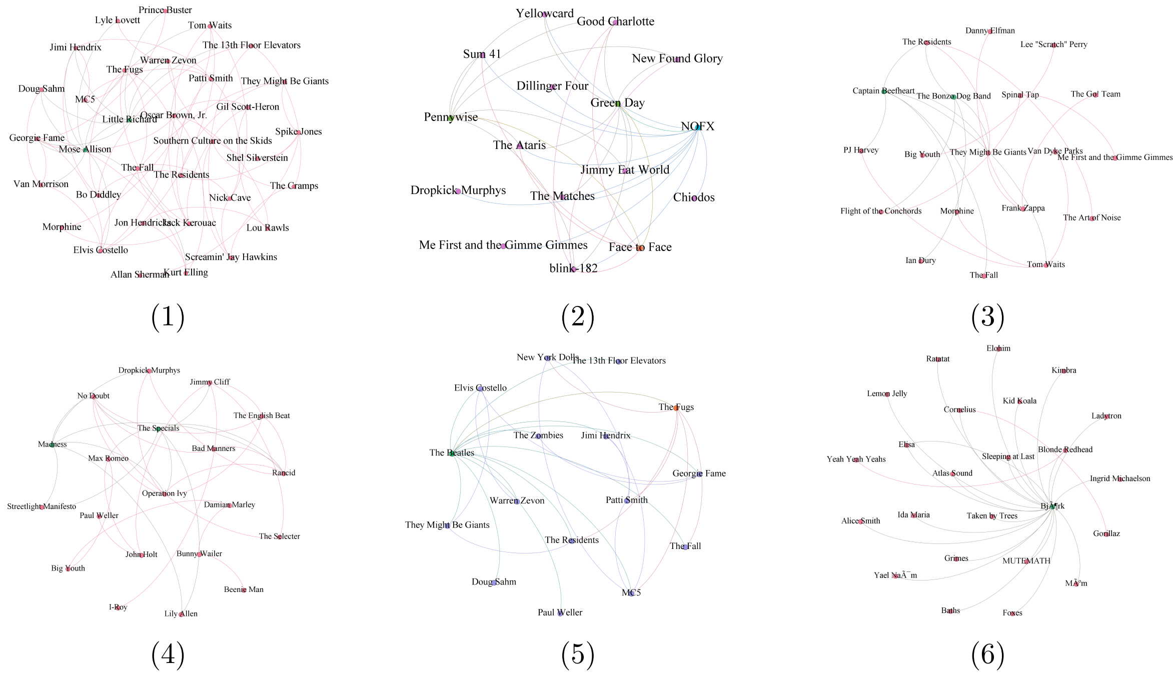


Figure 5: Result of Sub-network

4 Measures of Music Similarity

4.1 Overview

We define music and distance to measure its similarity. According to the literature, distance includes the part based on attribute and the part based on musician. Based on the Distance of attribute, a variety of attributes are fitted by entropy weight method to get the calculation results. Based on the Distance of musician, Hausdorff Distance is introduced to get the final Distance by combining the two.

4.2 Distance among Music

4.2.1 Distance Based on Attributes

There are three different categories of attributes shown in Figure 7, but here, for simplicity's sake, they are considered equally.

The distance is calculated by weighted euclidean distance, and the weights are obtained by entropy weight methods

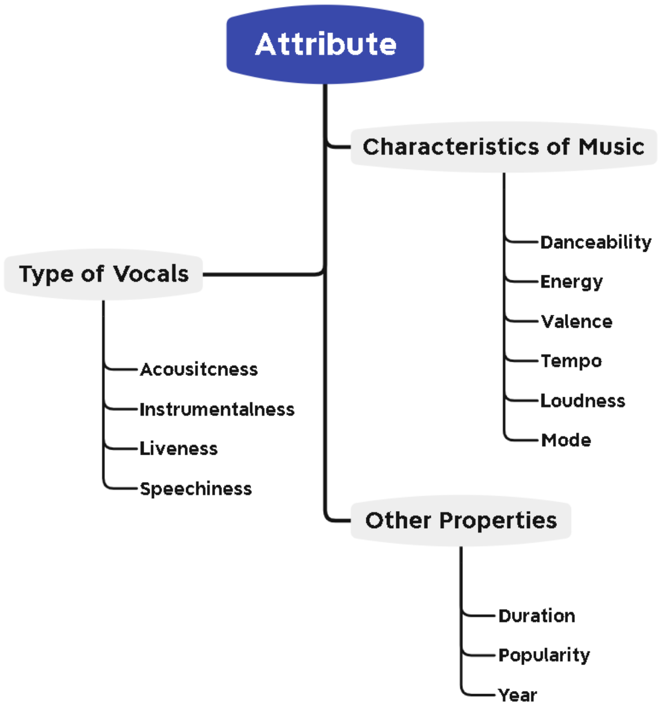


Figure 6: Three Kinds of Attributes

4.2.2 Distance Based on Musician

We introduce Hausdorff distance to measure the distance among musicians. The Hausdorff distance is a kind of distance between two sets. Here the set refers to the musician set. The items in the set, the musicians, co-composed the music.

As Hausdorff's theory, the distance is defined as

$$dist_H(X, Z) = \max(dist_h(X, Z), dist_h(Z, X)) \tag{14}$$

where

$$dist_h(X, Z) = \max_{x \in X} \min_{z \in Z} \|x - z\|_2 \tag{15}$$

4.3 Measures of Music Similarity

Together with the two kinds of distance mentioned above, we define

$$Similarity = \frac{1}{1 + distance} \tag{16}$$

where distance is the combination with the two kinds. Up here, based on the above metrics and the given data set, we can calculate the similarity of music and give detailed results in the next Section.

4.4 Result

We selected 6 representative genres and calculated the average distance between musicians within and between genres respectively, as shown in Table 6.

Table 6: Result of Music Similarity

	Folk	**Blues**	**Vocal**	**New Age**	**International**	**Electronic**
Folk	0.803	0.834	0.834	0.834	0.834	0.834
Blues	0.834	0.809	0.841	0.834	0.841	0.841
Vocal	0.834	0.841	0.816	0.845	0.845	0.845
New Age	0.834	0.841	0.845	0.798	0.831	0.831
International	0.834	0.841	0.845	0.831	0.793	0.823
Electronic	0.834	0.841	0.845	0.831	0.823	0.808

5 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(**??**).

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):

$$MS_i = max\{MS(Fl_i, In_j), \quad j = 1 \; to \; n\} \tag{17}$$

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.

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

$$(\Delta MS)_i = \frac{MS_i - MS_{i-1}}{MS_{i-1}} \tag{18}$$

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

$$\begin{cases} STS = max\{(\Delta MS)_i\} \\ LTS = \sum (\Delta MS)_i \end{cases} \tag{19}$$

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.

$$IDS = max(0, LTS) + max(0, STS - 0.01) \tag{20}$$

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(**??**) (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.

$$dif(i,j)^{(k)} = v_i^{(k)} - v_j^{(k)} \tag{21}$$

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.

$$DIF_i^{(k)} = \frac{1}{2}dif_1^{(k)} + \frac{1}{3}dif_2^{(k)} + \frac{1}{6}dif_3^{(k)} \tag{22}$$

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.

$$DC^{(k)} = ((1 + average(DIF_i^{(k)})) \cdot (1 + minimum(DIF_i^{(k)})) \cdot D[d]^{(k)} \tag{23}$$

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

Soda Stereo	**Characteristics**	***speechiness***	***energy***	***instrumentalness***
	DC	0.000249	0.0068	0.007554
Dave Matthews	**Characteristic**	***speechiness***	***instrumentalness***	***danceability***
	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 Final Remark

In this section, we'll make a general summary of our work, including the model we built, the strategies we found, the pros & cons of our methods, and the future development of our model.

6.1 Strengths and Weaknesses

6.1.1 Strengths

- **Accurate and intuitive.** The result of the solution has been proved to be accurate. We also made pairs of figures to help the result comprehensible.
- **Contractor-friendly.** Methods could be used just with some sets of contest data, which is not difficult for a musician to provide.
- **Extendable.** The parameter of the model shares high interpretability, which, therefore, makes the solution easy to extend.

6.1.2 Weaknesses

- **Data-based only.** The musician network is an extremely complex system which is impossible for any model to do mechanism analysis completely. Our work is supposed to be carried out based on historical data, as detailed as possible.

- **Simplified temporal networks.** Considering the complexity of model, we simplify the long-time network by average methods. Some kinds of temporal information could be ignored.

- **Hard to be exhaustive.** For comprehensive evaluation methods, it's hard to reach every aspect of a matter and impossible to be totally objective, our methods are no exception.

6.2 Future Model Development

- **Further exploration of mechanism analysis.** More work could be done to find the kernel of the musician network, which could give more explicable ideas for strategies in the future.

- **Wider exploration of Sub-network.** Limited by the variety of the data sets and the complexity of the model, characteristics used in our solution remain restricted, which could be developed to make the model have stronger robustness in the future.

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.

Appendices

Appendix A: Document to ICM society

Dear ICM Association,

We are a research team dedicated to the study of music influence networks. Through the data set provided, we adopt a more reasonable complex network analysis method to study the influence of different musicians and different schools.

Firstly, we use PageRank and out degree to simply describe the influence network. There are many smaller networks in this huge directed network. They are considered to be influential sub-networks. Among them, most nodes of a subnet belong to the same community, or multiple communities. Nodes belonging to the same community are often more closely connected than nodes in other communities. We consider the following basic static geometric characteristics of the network. We note that some famous musicians play an important role in their communities. Besides, we found some typical topological structures of some communities in large networks, such as star coupling networks.

Then, we measure the distance and similarity between different musicians and different music genres. The distance between two musicians is based on attributes includes: characteristics of the music, type of vocals, other properties.As for similarity, we define that it is inversely related to distance.

Finally, we measured the influence of different music genres. We find that Blues has a great influence on Folk from the picture. Similarly, Electronic and R&B; are mainly influenced by Country and Vocal separately. These mean that different music genres are closely related.The larger the point, the greater the impact. At the same time, they form an obvious interactive structure. The thicker the line, the greater the impact.

In the future, with the gradual enrichment of data, we will build a more complex, accurate and effective music influence network.This will be of great help for us to reveal the impact of music on culture. If we can get more data on the impact of society, politics and the Internet on music, our conclusion will be more reliable.

Yours sincerely,

Mathematical modeling team