# DETECTING COLLABORATION PROFILES IN SUCCESS-BASED MUSIC GENRE NETWORKS

**Gabriel P. Oliveira** **Mariana O. Silva** **Danilo B. Seufitelli**
**Anisio Lacerda** **Mirella M. Moro**
Universidade Federal de Minas Gerais, Brazil
{gabrielpoliveira,mariana.santos,daniloboechat,anisio,mirella}@dcc.ufmg.br

## ABSTRACT

We analyze and identify collaboration profiles in success-based music genre networks. Such networks are built upon data recently collected from both global and regional Spotify weekly charts. Overall, our findings reveal an increase in the number of distinct successful genres from high-potential markets, pointing out that local repertoire is more important than ever on building the global music ecosystem. We also detect collaboration patterns mapped into four different profiles: *Solid*, *Regular*, *Bridge* and *Emerging*, wherein the two first depict higher average success. These findings indicate great opportunities for the music industry by revealing the driving power of inter-genre collaborations within regional and global markets.

## 1. INTRODUCTION

Artist collaborations are more popular than ever, as the landscape of the music industry becomes more complex. This widely adopted strategy is a strong force driving music nowadays, maintaining artists' presence and relevance in the market. Such connections usually help artists bridge the gap between styles and genres, overlapping new fan bases and consequently increasing their numbers. Figure 1 illustrates this phenomenon and highlights the growing trend in the number of collaborations within Billboard Hot 100 Charts. Although the general curve increases over time, genres such as *rap* and *hip-hop* present a collaboration rate higher than others (e.g., *pop* and *rock*). This contrast can be explained by the intrinsic nature of each music genre. For instance, *rap* and *hip-hop* artists frequently collaborate with the *pop* community, mainly as featured artists. Moreover, partnerships involving *pop* music may take place not only through intra-genre collaborations but also through inter-genres, bringing an additional dimension to their songs.

Musicians teaming up is nothing new but has risen far beyond the norm. Remaining an industry of creative growth, it is only natural for music (i.e., all musical scene members)

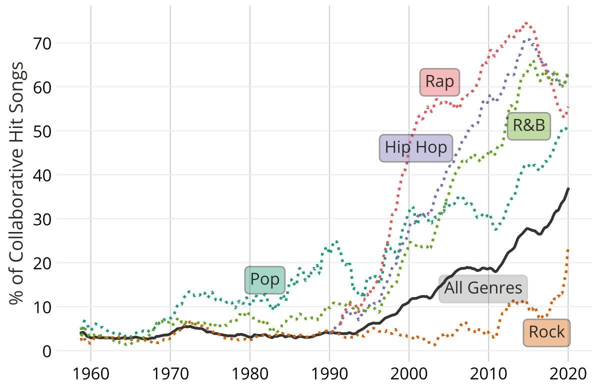


**Figure 1**: Historical frequency of collaborative hit songs for selected genres on Billboard Hot 100 Chart (1958 - 2020).

adapting to new conditions and redefining its layout. Not surprisingly, the Grammy categories were tightened (from 109 to 78, in 2012) as a result of its dynamic nature. [1] That is, the notion of categories and genres are blurred as never before. Through cross-genre collaboration, artists are naturally venturing into new domains and working outside of the category which they had originally been ascribed to. Such a collaboration phenomenon may be drastically reshaping music global environment, by challenging segments of certain genres to come up with something entirely new [1]. Moreover, this gradual revolution is becoming a driving force in creating a more collaborative scenario, making music one of the most innovative art forms.

As this creative market changes, it becomes more unpredictable; and doing both predictive and diagnostic analyses in such a context remains challenging. Still, we believe factors leading to an ideal musical partnership can be understood by exploring collaboration patterns that directly impact its success [1–3]. Hence, we aim to unveil the dynamics of cross-genre connections and collaboration profiles in success-based networks (i.e., connections formed by genres of artists who cooperate and create hit songs). We do so through the following research questions (RQ).

**RQ1:** *Does the regional aspect impact on popular genres and their hit songs?*

**RQ2:** *How has genre collaboration evolved over the past few years?*



[1] Grammy Award: https://en.wikipedia.org/wiki/Grammy_Award

**RQ3:** *Which are the potentially intrinsic factors and indicators that influence the collaboration success?*

In order to answer such questions, we first model genre collaboration in the music scenario as success-based networks (Section 3.1). Then, we build a novel dataset with data from global and regional markets (Section 3.2) and present the network science concepts and metrics required for understanding the paper (Section 3.3). Overall, our analyses and experiments over the networks reveal that:

(1) Individually analyzing regional markets is fundamental, as local genres play a key role on determining hit songs and popular artists (Section 4.1).

(2) In general, genre collaborations are increasing, with emerging local genres hitting global success – despite the differences in the evolution of regional markets (Section 4.2).

(3) Genre collaborations analyses describe three significant factors (*Attractiveness, Affinity* and *Influence*, Section 5.1) to uncover four profiles (*Solid, Regular, Bridge* and *Emerging*, Section 5.3).

## 2. RELATED WORK

Genres are fundamental within the musical scenario by aggregating songs that share common characteristics. Hence, they are frequently used in the field of Music Information Retrieval (MIR), which aims to extract relevant information from music content. In fact, several tasks are genre-dependent or directly related to them, such as automatic genre classification, which has been largely studied by the MIR community over the years [4–8]. Nonetheless, there are also genre-aware studies assessing music source separation [9], genre modeling [10], preferences [11], disambiguation/translation [12, 13], new datasets [14] and ontologies [15]. Network science, the core of our methodology, has also been used to model genres into influence networks [2] and song communities [16].

Hit Song Science (HSS), which tackles the problem of predicting the popularity of a given song, is also an emerging field within MIR. Thus, different studies analyze the impact of acoustic and social features in musical success [17–20], some of them including genre information [21,22]. Moreover, Silva et al. [1] address collaboration as a key factor in success, using topological properties to detect relevant profiles in artist networks. Such an approach is novel and promising in HSS, but it is restricted to the artist and song levels. Therefore, studying collaboration from a genre perspective may reveal important information on how artists from different communities team up to make a new hit song. To the best of our knowledge, we are the first to build a success-based genre network and detect collaboration profiles within it, going deeper into the intrinsic factors that make up a successful collaboration.

## 3. METHODOLOGY

This work aims to detect collaboration profiles over music genre networks. Building a genre network (Section 3.1) requires a proper dataset (Section 3.2), and finding its profiles needs different metrics (Section 3.3), as described next.

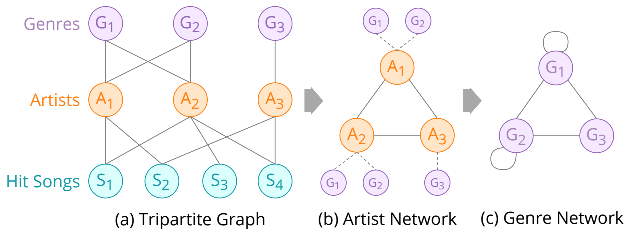


**Figure 2**: Reduction from the tripartite (a) to the one-mode Genre Collaboration Network (c). The intermediate step is an Artist Network with genre information (b). Artists and genres are linked when hit songs involve both nodes.

### 3.1 Genre Collaboration Network

A Collaboration Network is usually modeled as a graph formed by nodes (vertices) that may be connected through edges. For example, nodes represent artists and are connected by an edge if the respective pair of artists has collaborated in a song. Now, to analyze the interactions between genres, we model music collaboration as a tripartite graph, in which nodes are divided into three sets: genres, artists, and hit songs; i.e., the minimum elements to evaluate success. The building process of the genre network from the tripartite model is illustrated in Figure 2. Collaborative hit songs are sung by two or more artists, regardless of their participation (e.g., a typical *feat.* or a duet). We also equally consider all genres linked to an artist because they shape how such an artist is seen by fans and music industry.

For analyzing the interaction between musical genres, we reduce the tripartite model into a one-mode network in which nodes are exclusively genres. However, such a reduction is only possible by executing an intermediate step: building the artist collaboration network, Figure 2(b). In such a network, two artists are connected when both collaborate in one or more hit songs. The genres are not lost, as they are linked directly to the artists.

We may now build the final network by connecting the genres of artists who collaborate in the artist network. The edges are undirected and weighted by the number of hit songs involving artists from both genres, Figure 2(c). Also, self-loop edges are allowed, as there are hit songs from artists of the same genre. For example, the song *Old Town Road* [2] by Lil Nas X and Billy Ray Cyrus generates an edge between these artists in the intermediate network; and each of Lil Nas X's genres (*pop rap, country pop* and *hip hop*) is linked to Cyrus' only genre (*country*) with weight 1.

### 3.2 Dataset Building Process

Over recent years, the world has seen a dramatic change in the way people consume music, moving from physical records to streaming services. Since 2017, such services have become the main source of revenue within the global

[2] #1 Song of 2019 according to Billboard Year-End Hot 100 Chart: https://billboard.com/charts/year-end/2019/hot-100-songs

recorded music market. In fact, streaming revenues increased by 75.4% from then, reaching a total amount of US$ 11.4 billion by the end of 2019.[3] Thus, we build our dataset by using data from Spotify, the most popular global audio streaming service, with more than 286 million users across 79 markets.[4] It provides a weekly chart of the 200 most streamed songs in all its markets, and an aggregated global chart. We collect global and regional charts from January 2017 to December 2019, considering eight of the top 10[5] music markets according to IFPI: United States, Japan[6], United Kingdom, Germany, France, Canada, Australia, and Brazil. We also use Spotify API[7] to gather information about the hit songs and artists present in the charts, such as all collaborating artists within a song (since the charts only provide the main ones) and their respective genres, which is the core of this work. Our final dataset contains 1,370 charts from 156 weeks, comprising 13,880 hit songs and 3,612 artists from 896 different music genres.

Then, a processing phase focuses on the artists' genres, because Spotify assigns a list of very specific genres to each artist. For example, Jay-Z (one of the most popular rappers) is assigned to both *east coast hip hop* and *hip hop* genres, which may be described only by *hip hop*. To simplify our modeling and further analyses, we choose to map all specific genres to more embracing and well-established *super-genres*. Note that the regional aspects are not lost in such a mapping, because our analyses are made separately for each considered market. Hence, the 896 existing genres are now mapped into 162 *super-genres*. The dataset and genre mapping are publicly available on our project page.[8]

**Table 1**: Most popular music genres in each considered market from 2017 to 2019.

| | Genre | Songs | Arts. | | Genre | Songs | Arts. | | Genre | Songs | Arts. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Global* | pop | 1,715 | 424 | *Canada* | pop | 1,790 | 402 | *UK* | pop | 1,772 | 371 |
| | hip hop | 1,192 | 281 | | rap | 1,762 | 209 | | hip hop | 1,138 | 232 |
| | rap | 1,184 | 195 | | hip hop | 1,511 | 232 | | rap | 974 | 167 |
| | pop rap | 845 | 130 | | pop rap | 1,355 | 149 | | dance pop | 922 | 178 |
| | dance pop | 832 | 165 | | trap | 1,139 | 154 | | pop rap | 660 | 120 |
| *Australia* | pop | 1,646 | 411 | *Germany* | hip hop | 2,604 | 352 | *Japan* | j-pop | 797 | 163 |
| | rap | 873 | 165 | | pop | 1,665 | 479 | | pop | 705 | 210 |
| | dance pop | 822 | 171 | | rap | 1,223 | 205 | | dance pop | 438 | 103 |
| | hip hop | 792 | 191 | | dance pop | 650 | 159 | | j-rock | 312 | 72 |
| | pop rap | 657 | 133 | | pop rap | 462 | 109 | | r&b | 276 | 82 |
| *Brazil* | pop | 1,072 | 256 | *France* | pop | 3,138 | 470 | *USA* | rap | 2,057 | 231 |
| | sertanejo | 565 | 82 | | hip hop | 2,660 | 285 | | hip hop | 1,719 | 241 |
| | brazilian funk | 559 | 156 | | rap | 2,526 | 245 | | pop | 1,704 | 340 |
| | dance pop | 415 | 101 | | francoton | 1,097 | 82 | | pop rap | 1,518 | 139 |
| | electro | 307 | 93 | | dance pop | 391 | 119 | | trap | 1,370 | 172 |

### 3.3 Network Science Metrics

In this work, we use well-established network science metrics to analyze musical collaboration. Such metrics consider the network topological features, i.e., they relate to the network structure (nodes and edges) as follows.[9]

***Degree and Weighted Degree.*** These metrics refer to the connectivity of each node in the network. The degree of a node is its amount of incident edges, and the weighted degree is the sum of the edges' weight. In our genre collaboration network, degree stands for the number of genre connections, and weighted degree represents the number of hit songs shared by both genres.

***Clustering Coefficient (CC).*** Measures the tendency of neighbors of a node to be connected themselves. The higher its value, the more interconnected the node neighborhood.

***Common Neighbors (CN).*** The number of neighbors that a given pair of nodes has in common in a network.

***Neighborhood Overlap (NO).*** The ratio between the common neighbors of a given pair of nodes and the union set of their neighbors. Edges with low NO reveal local bridges in the network, i.e., nodes traveling in "social circles", having almost no common connection.

***Preferential Attatchment (PA).*** The probability of a given pair of nodes connecting in the future. The intuition is the more neighbors they have, the more likely they are to connect in the future.

***Edge Betweenness (EB).*** The fraction of shortest paths that go through an edge in the network. Edges with a high score represent a bridge-like connector between two parts of the network, and their removal may affect the communication between others due to the lost common shortest paths.

***Resource Allocation (RA).*** For a pair of nodes, it represents the fraction of a resource (e.g., information) that a node can send to another through its common neighbors.

## 4. EXPLORATORY ANALYSIS

We perform an exploratory analysis of the data collected in two main steps. First, we analyze Spotify charts for each market to detect popular genres (Section 4.1). Then, we characterize the genre collaboration network to understand the evolution of each market (Section 4.2).

### 4.1 Music Genres Overview

To answer *RQ1*, we analyze charts of eight markets over three years (see Methodology) in Table 1. Each country has its own musical inclinations, although the predominant genres are mostly *pop*/*pop rap*, *hip-hop*, and *rap*. Such preference may be due to the increasing number of collaboration songs among artists from different musical genres, as revealed in Figure 1: growing collaborations of *pop*, *rap*, *hip-hop*, and *R&B* in recent years. Also, except for *R&B*, they are the main genres at the top-5 genre lists on most markets; i.e., such genres are among both the most collaborative ones and the most listened on the globe. Moreover, there are three markets with local genres on their top-5 list: Brazil with *sertanejo* and *brazilian funk*; France with *francoton*; and Japan with *j-pop* and *j-rock*. Although local, such genres are potentially good choices for record companies to encourage musical collaborations. Note local engagement shapes the global environment, ensuring that music culture within such countries can develop and progress.

[3] IFPI Global Music Report 2019: https://gmr.ifpi.org/
[4] Spotify Company Info: https://newsroom.spotify.com/company-info/
[5] Data from South Korea and China was not available in Spotify.
[6] The first Japanese weekly chart is from August 31, 2017.
[7] Spotify API: https://developer.spotify.com/
[8] Project Bàde: https://bit.ly/proj-Bade
[9] For more information on such metrics, see references [23–25]

**Table 2**: Network characterization for global and three regional markets, representing the groups of countries with similar network evolution. Underlined values are the highest metric value for a specific market throughout the considered period.

| Metric | Global | | | USA (Group 1) | | | Brazil (Group 2) | | | UK (Group 3) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 |
| Number of genres (nodes) | 72 | 79 | 89 | 76 | 73 | 83 | 58 | 63 | 61 | 74 | 76 | 79 |
| Number of collabs (edges) | 564 | 583 | 709 | 542 | 522 | 670 | 453 | 524 | 392 | 610 | 605 | 627 |
| Average degree | 15.7 | 14.8 | 15.9 | 14.3 | 14.3 | 16.1 | 15.6 | 16.6 | 12.9 | 16.5 | 15.9 | 15.9 |
| Average degree (weighted) | 256.9 | 247.4 | 236.7 | 324.6 | 287.9 | 241.4 | 136.1 | 133.0 | 95.3 | 216.5 | 203.6 | 159.5 |
| Density | 0.221 | 0.189 | 0.181 | 190 | 0.199 | 197 | 0.274 | 268 | 214 | 0.226 | 212 | 204 |
| Average Clustering Coefficient | 0.743 | 0.757 | 0.754 | 0.762 | 760 | 726 | 0.770 | 758 | 677 | 724 | 0.754 | 738 |
| Number of self-loops | 24 | 21 | 28 | 25 | 22 | 27 | 24 | 29 | 27 | 28 | 25 | 30 |

## 4.2 Network Characterization

After analyzing the charts, we build the genre collaboration network for each market and year to find out how genres connect to answer $RQ2$. With nine markets (global and eight countries) during three years, we analyze 27 networks [10]. For each network, we calculate basic statistics on its nodes and edges, as well as structural metrics (Section 3.3). Table 2 shows the results for selected markets.

First, the global genre networks reveal the world is more open to new successful genres (number of nodes/genres growth). Also, the number of genre connections (edges) increased considerably, meaning more collaborative hit songs are coming from artists whose genres are not linked in prior networks, opening up new opportunities for those genres to acquire new listeners. The networks average degree remains stable, while its weighted value decreases over the years. This could reveal a growth in the number of collaborations of well-established genres with emerging ones, represented by edges with low weight values (few hit songs). Still, such low-degree emerging genres may become popular shortly, expanding their collaborations to other unexplored genres. For instance, *k-pop* connections double as it spreads worldwide, approaching genres such as *reggaeton* (e.g., the collaboration between J-Hope from BTS and Becky G in the song *Chicken Noodle Soup*, September 2019).

For regional markets, we classify the countries into three groups, according to the similarities in networks' evolution: *(i)* USA and Canada; *(ii)* Brazil, France, Germany and Japan; *(iii)* UK and Australia. As the global network, countries in the first group have an increasing average degree and a decreasing average clustering coefficient, thus indicating a stronger tendency to diversify the inter-genre collaborations. Then, the second group includes non-English speaking countries with decreasing connectivity metrics in 2019, after a peak in 2018. The last group has countries in which more genres are becoming successful, while the connections are not increasing in the same proportion.

Overall, considering regional markets individually becomes more important for producers and record labels, as they are delivering more global hits over time. Their distinct behavior emphasizes the strength of cultural aspects on determining how music is consumed and the success of a given genre or artist. In each market, genre connections may reveal distinct profiles, which are an important tool for analyzing successful genre collaborations.

[10] All networks can be visualized in https://bit.ly/proj-Bade

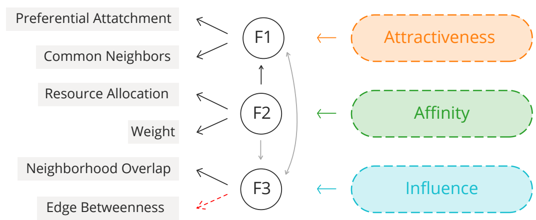


**Figure 3**: EFA diagram. Solid and red dashed lines represent positive and negative correlations, respectively. Dark and lighter lines represent strong $[0.6 - 1.0]$ and weak $[< 0.6]$ correlations, respectively.

## 5. GENRE COLLABORATION PROFILES

This section presents our approach to uncover significant factors that compose a successful music genre collaboration. Inspired by Silva et al. [1], we first extract information from the success-based networks by evaluating six edge-dependent metrics (Section 3.3). We perform an Exploratory Factor Analysis on such metrics to define factors in Section 5.1, and then perform cluster analysis in Section 5.2. Finally, we organize the found clusters into collaboration profiles in Section 5.3, to investigate the key driving factors on successful collaborations and then answer $RQ3$.

## 5.1 Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) [26] is a statistical method designed to underline patterns of correlations among observed variables and extract latent factors. Generally, EFA identifies the number of common factors and the pattern of factor loadings (correlations). It assumes and asserts that manifest (observed) variables are expressed as a linear combination of factors and measurement errors. Each factor explains a particular variance in the variables and may find hidden data patterns. There are two main issues when executing an EFA: *(i)* determining the number of factors to retain for analysis, and *(ii)* selecting the final structure for how the measured variables relate to the factors. For the former, we use the Parallel Analysis criteria [27], which is based on random data simulation. The suggested number of factors to extract is then provided and based on examining the *scree plot* [28] of factors of the observed data with that of a random data matrix of the same size as the original. Finally, the EFA is performed using the well known Ordinary Least Squares (OLS) factoring method and an oblique rotation, allowing factors to correlate with each other.

We use EFA to identify the common factors and the

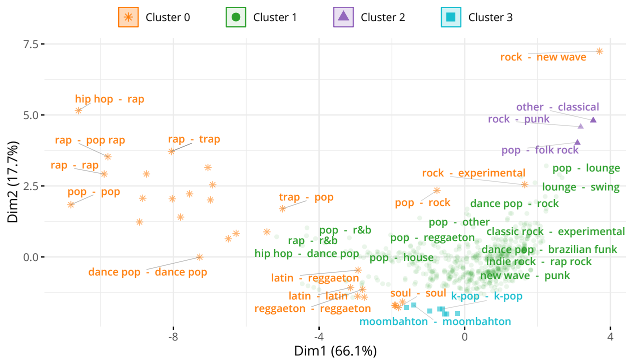


**Figure 4**: Clustering result for the USA network, in 2019, with examples of some genre collaborations in each cluster.

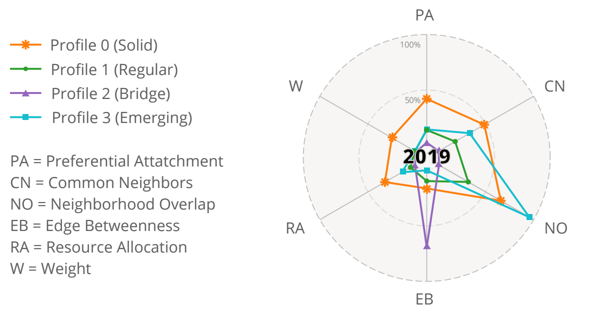


**Figure 5**: Collaboration profiles for all markets (2019). For additional radar plots, see Supplementary Material.

relationships among the edge-dependent metrics of all 27 success-based networks. Overall, the analysis results suggest a three-factor structure within those six metrics. A graphical representation of the emerging structure is in Figure 3. As the three factors are conceptually coherent, we labeled them as follows.

***Attractiveness (F1).*** Factor 1 has high loads for both PA and CN metrics, with a positive correlation between them. Specifically, values close to 0 indicate that two nodes are not close and attracted, while higher values indicate closer nodes. Therefore, this factor corresponds to the predisposition of two nodes to connect in the future.

***Affinity (F2).*** Factor 2 has high loads for both RA and W metrics, with a positive correlation between them. High values indicate strong social ties, and lower ones indicate weak ties. Hence, this factor measures both the frequency of collaboration between two nodes and the social strength.

***Influence (F3).*** Factor 3 has high loads for both NO and EB metrics, with a negative correlation between them. Edges with low NO and high EB certainly consist of local bridges in the network. That is, they represent a bridge-like connector between two "social circles". Therefore, this factor corresponds to the importance level of an edge with access to different regions in the network.

### 5.2 Cluster Analysis

The second step of our approach is cluster analysis to group similar music genre connections based on the aforementioned factors. We use DBSCAN [29] as a clustering algorithm, which assigns data points to the same cluster if they are *density-reachable* from each other. Two important parameters are required for DBSCAN: $\epsilon$ defines the radius of neighborhood around a point $x$; and $MinPts$ (minimum points) is the minimum number of neighbors within the $\epsilon$ radius. To choose the optimal $\epsilon$ value, we use a method based on $k$-nearest neighbor distances, which calculates the average of the distances of every data point to its $k$ nearest neighbors. In general, the value of $k$ is specified by the user and corresponds to the $MinPts$ parameter. As a general rule, the $MinPts$ can be derived from the number of dimensions $D$ in the dataset as $MinPts \geq D + 1$. Since we have six topological metrics, we set $MinPts = 7$.

Overall, four distinct clusters were detected in at least one of the 27 collaboration networks. As an example including all clusters, Figure 4 shows the result of the US network in 2019, where Cluster 0 groups the outliers identified by DBSCAN (data points in low-density regions, i.e., not associated with any proper cluster). Clusters 1 and 2 are slightly overlapping, but each covers groups of high-density data points, which is successful information in this analysis. We can also certainly conclude Cluster 3 is separate from the others. Next, we describe each cluster.

### 5.3 Collaboration Profiling

Now that we have detected a set of predominant clusters on all modeled networks, the next step is to look at their characteristics for profiling them and defining proper identities. First, for each network, we calculate the mean of the normalized metrics values grouped by each cluster id. Then, for each year, we plot radar charts for each profile with the mean values of each market present in that profile. Figure 5 shows such radar charts, where each cluster is represented by a polygon that exhibits its identity. To compare the metric values' magnitude of each cluster, we adopt the following scale: *low* is the bottom $30^{th}$ percentile; *medium* is between $30^{th}$ and $80^{th}$ percentile; and *high* is the top $20^{th}$ percentile. Such scale is based on the annual general values, i.e., considering the grouped normalized features of all markets by year.

The differences among the three plots represent minimal changes over the years. However, the distinct shapes show each cluster is *high* or *low* in certain features. Particularly, Cluster 0 presents collaborations with *high* values for *Attractiveness* and *Affinity* factors, but *medium* values for *Influence*. With a similar shape, Cluster 1 presents *medium* values for all four factors. On the other hand, Cluster 2 presents *high* values only for *Influence*, with *low Attractiveness* and *Affinity*. Finally, Cluster 3 is the group with major differences over the years: in general, its collaborations have *medium Attractiveness* and *Affinity*, and *low Influence*. Overall, each curve depicts a distinct collaboration profile, acting as a class descriptor of a cluster.

With the collaboration profiles settled, we can now answer $RQ3$. First, we analyze the distributions of success rate, and then the number of *intra*- and *inter*-genre collaborations for each profile. Here, we define success rate as the average of total streams of songs belonging to the music

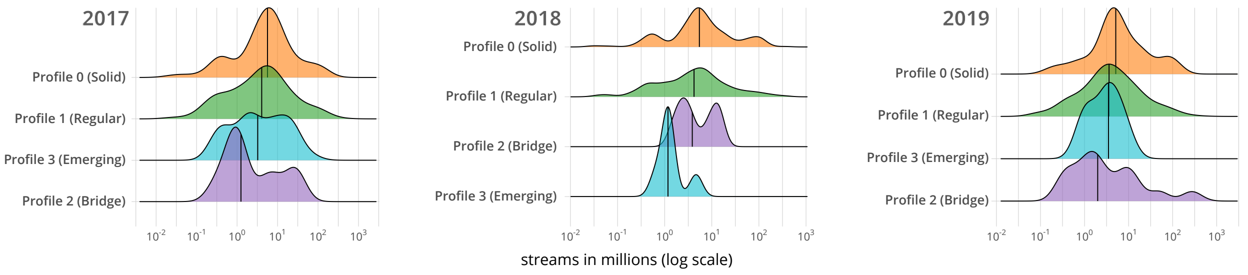


**Figure 6**: Density ridgeline plots of streams in millions for each cluster. Darker vertical lines represent median values.

genres that compose the collaboration (edge) in that year. Figure 6 shows the success density ridgeline plots for each profile, indicating that Profiles 0 and 1 are composed of the most successful music genre collaborations, on average. With results from Table 3, in general, the most successful profiles are those composed of more inter-genre collaborations. Such a result may indicate a strong correlation between musical success and inter-genre collaborations. Indeed, by teaming up with one (or more) person of a different musical style in a song, both artists may draw from one another's fan bases; i.e., they may promote themselves to new public who could increase their fan base and audiences.

To summarize the characteristics of the collaboration profiles, we name each as follows.

– Profile 0 is *Solid Collaboration (Solid)*, composed of well-established collaborations between most popular genres (super-genres), which have been going on for decades. Examples include: *rap* and *hip-hop*, whose collaborative albums are hugely popular; and *hip-hop* and *pop*, whose separating line (between both genres) has become completely blurred in the last decade, mainly in the USA;

– Profile 1 is *Regular Collaboration (Regular)*, composed of the most common collaborations in all markets, which are very similar to solid collaborations but not as engaged. For instance, collaborations between *hip-hop*/*rap*/*pop* and *jazz*/*blues*/*soul*, which can be typical in many markets, but not as consolidated when compared to *Solid* ones;

– Profile 2 is *Bridge Collaboration (Bridge)*, composed of collaborations with high influence, representing bridge-like connectors between two regions of a network (mostly between divergent music styles). Such collaborations may be possible sources of investment to increase connectivity and strengthen ties among different audiences. One example is collaborations between *gospel* and others, such as *rap* and *MPB* (*Brazilian Popular Music*); and

– Profile 3 is *Emerging Collaboration (Emerging)*, formed mainly of collaborations between regional genres. Such partnerships generally occur within the same genre; possibly between one (or more) unknown artist and one (or more) established artist; or maybe in order to easily reach that genre audience. We propose the term *emerging* because such a profile can be seen as a transition phase for beginners, until they establish their fan bases. Examples of regional genres here include *k-pop* (popular music from South Korea), *moombahton* (fusion genre of *house* music and *reggaeton* (from Washington, D.C.), and *forró* (a popular musical genre from Brazilian Northeastern Region).

**Table 3**: Total number of *intra-* and *inter*-genre collaborations in each profile, from 2017 to 2019.

| Collab | *Solid* 2017 | 2018 | 2019 | *Regular* 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|
| Inter-genre | 140 (49%) | 125 (42%) | 103 (51%) | 1,828 (99%) | 1,916 (98%) | 2,165 (94%) |
| Intra-genre | 145 (51%) | 174 (58%) | 99 (49%) | 23 (1%) | 34 (2%) | 128 (6%) |
| **Collab** | ***Bridge* 2017** | **2018** | **2019** | ***Emerging* 2017** | **2018** | **2019** |
| Inter-genre | 10 (100%) | 7 (100%) | 16 (100%) | 3 (7%) | 1 (17%) | 0 (0%) |
| Intra-genre | 0 (0%) | 0 (0%) | 0 (0%) | 40 (93%) | 5 (83%) | 7 (100%) |

## 6. CONCLUSIONS

In this paper, we analyze and identify collaboration profiles in success-based music genre networks. Our results suggest that analyzing regional markets individually is fundamental, as local genres play a key role in determining hit songs and popular artists. Besides the differences in the evolution of regional markets, genre collaborations are also increasing, with emerging local genres achieving global success. Moreover, the networks' structures reveal three main factors that describe a genre collaboration: *Attractiveness*, *Affinity* and *Influence*. Analyzing such factors uncovers four different collaboration profiles: *Solid*, *Regular*, *Bridge* and *Emerging*, which act as class descriptors of successful partnerships. Overall, our results contribute to the understanding of the relation between cross-genre collaboration and hit songs.

Indeed, detecting genre collaboration profiles is a powerful way to assess musical success by describing similar behaviors within collaborative songs from multiple angles. Our findings may act as base material for further research tasks, e.g., prediction and recommendation. The former enables predicting the success of a given song/artist/album, while the latter can be used to point out potentially successful genre/artist collaborations. This not only benefits the MIR community, but also the music industry as a whole. In fact, music industry CEOs may maximize expected success by properly investing in potential artist/genre collaborations. Finally, artists may also profit by identifying the most suitable partnerships to lead the album to early stardom. In conclusion, this work sheds light on the science behind the collaboration phenomenon, providing potential impact to the music industry.

**Future Work.** We plan to consider other data sources and to expand the time period in order to better understand the markets' behavior, enhancing further analyses.

**Acknowledgements.** This work is supported by CAPES, CNPq and FAPEMIG, Brazil.

## 7. REFERENCES


[1] M. O. Silva, L. M. Rocha, and M. M. Moro, "Collaboration Profiles and Their Impact on Musical Success," in *Procs. of ACM/SIGAPP SAC*, Limassol, Cyprus, 2019, pp. 2070–2077.

[2] N. J. Bryan and G. Wang, "Musical influence network analysis and rank of sample-based music," in *ISMIR*, Miami, USA, 2011, pp. 329–334.

[3] C. Baccigalupo *et al.*, "Uncovering affinity of artists to multiple genres from social behaviour data," in *ISMIR*, Philadelphia, USA, 2008, pp. 275–280.

[4] T. Arjannikov and J. Z. Zhang, "An association-based approach to genre classification in music," in *ISMIR*, Taipei, Taiwan, 2014, pp. 95–100.

[5] I. Vatolkin, G. Rudolph, and C. Weihs, "Evaluation of album effect for feature selection in music genre recognition," in *ISMIR*, Malaga, Spain, 2015, pp. 169–175.

[6] S. Oramas *et al.*, "Multi-label music genre classification from audio, text and images using deep features," in *ISMIR*, Suzhou, China, 2017, pp. 23–30.

[7] A. Tsaptsinos, "Lyrics-based music genre classification using a hierarchical attention network," in *ISMIR*, Suzhou, China, 2017, pp. 694–701.

[8] S. S. Ghosal and I. Sarkar, "Novel approach to music genre classification using clustering augmented learning method (CALM)," in *AAAI MAKE*, ser. CEUR Workshop Proceedings, vol. 2600, 2020.

[9] C. Laroche *et al.*, "Genre specific dictionaries for harmonic/percussive source separation," in *ISMIR*, NYC, USA, 2016, pp. 407–413.

[10] M. Prockup *et al.*, "Modeling genre with the music genome project: Comparing human-labeled attributes and audio features," in *ISMIR*, Malaga, Spain, 2015, pp. 31–37.

[11] J. Bansal and M. Woolhouse, "Predictive power of personality on music-genre exclusivity," in *ISMIR*, Malaga, Spain, 2015, pp. 652–658.

[12] R. Hennequin, J. Royo-Letelier, and M. Moussallam, "Audio based disambiguation of music genre tags," in *ISMIR*, Paris, France, 2018, pp. 645–652.

[13] E. V. Epure, A. Khlif, and R. Hennequin, "Leveraging knowledge bases and parallel annotations for music genre translation," in *ISMIR*, Delft, the Netherlands, 2019, pp. 839–846.

[14] D. Bogdanov *et al.*, "The acousticbrainz genre dataset: Multi-source, multi-level, multi-label, and large-scale," in *ISMIR*, NYC, USA, 2019, pp. 360–367.

[15] H. Schreiber, "Genre ontology learning: Comparing curated with crowd-sourced ontologies," in *ISMIR*, NYC, USA, 2016, pp. 400–406.

[16] D. C. Corrêa, A. L. M. Levada, and L. da F. Costa, "Finding community structure in music genres networks," in *ISMIR*, Miami, UAS, 2011, pp. 447–452.

[17] L. Yang *et al.*, "Revisiting the problem of audio-based hit song prediction using convolutional neural networks," in *ICASSP*. IEEE, 2017, pp. 621–625.

[18] C. V. Araujo *et al.*, "Predicting music success based on users' comments on online social networks," in *WebMedia*, Brazil, 2017, pp. 149–156.

[19] F. Calefato *et al.*, "Collaboration success factors in an online music community," in *ACM GROUP*, Sanibel Island, USA, 2018.

[20] A. Cosimato *et al.*, "The conundrum of success in music: Playing it or talking about it?" *IEEE Access*, vol. 7, pp. 123 289–123 298, 2019.

[21] F. Abel *et al.*, "Analyzing the blogosphere for predicting the success of music and movie products," in *ASONAM*, Odense, Denmark, 2010, pp. 276–280.

[22] E. Zangerle *et al.*, "Hit song prediction: Leveraging low- and high-level audio features," in *ISMIR*, Delft, the Netherlands, 2019, pp. 319–326.

[23] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.

[24] D. Liben-Nowell and J. M. Kleinberg, "The link-prediction problem for social networks," *J. Assoc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2007.

[25] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: statistical mechanics and its applications*, vol. 390, no. 6, pp. 1150–1170, 2011.

[26] A. B. Costello and J. Osborne, "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis," *Practical assessment, research, and evaluation*, vol. 10, no. 1, p. 7, 2005.

[27] L. G. Humphreys and R. G. Montanelli Jr, "An investigation of the parallel analysis criterion for determining the number of common factors," *Multivariate Behavioral Research*, vol. 10, no. 2, pp. 193–205, 1975.

[28] R. B. Cattell, "The scree test for the number of factors," *Multivariate behavioral research*, vol. 1, no. 2, pp. 245–276, 1966.

[29] M. Ester *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Procs. of KDD*, Portland, USA, 1996, pp. 226–231.