USING NETWORK SCIENCES TO RANK MUSICIANS AND COMPOSERS IN BRAZILIAN POPULAR MUSIC

Charith Gunaratna
Computer Sciences
Florida Tech, USA
agunaratna2009@my.fit.edu

Evan Stoner
Computer Sciences
Florida Tech. USA
estoner2010@my.fit.edu

Ronaldo Menezes
Computer Sciences
Florida Tech, USA
rmenezes@cs.fit.edu

ABSTRACT

Music fascinates and touches most people. This fascination leads to opinions about the music pieces that reflects people's exposure and personal experience. This inherent bias of people towards music indicates that *personal opinion* is inappropriate for defining the quality of music and musicians. This paper takes a holistic view of the problem and delves into the understanding of the structure of Brazilian music rooted in Network Sciences. In this paper we work with a large database of albums of Brazilian music and study the structure of collaborations between all the musicians and composers. The collaboration is modelled as a social network of musicians and then analyzed from different perspectives with the goal of describing what we call the structure of that musical genre as well as provide a ranking of musicians and composers.

1. INTRODUCTION

Brazilian Music is admired worldwide due to its diversity and richness of sounds. The music from Brazil is in fact a confluence of many different cultural influences [1, 2]. This process of globalization of the popular music of Brazil has come to a full circle when other genres around the world started to incorporate Brazilian rhythms and refer to Brazilian music as an influence to them. It is known that world greats such as Miles Davis and Frank Sinatra, and more recently the likes of Pat Metheny and Bill Frisell (jazz guitarists), have been influenced by and even worked with many Brazilian musicians.

When it comes to the arts, is hard to define a canon due to subjective opinions. For classical art, the use of networks has improved our ability to understand the importance of many works [13]. In popular art, the definition is a little harder because it could depend on many factors such as sales, and playtime on the radio. However, this paper proposes to use techniques from networks sciences to model the network of collaborations among musicians and derive from the social network a good ranking of musicians and composers in Brazilian music.

Many Brazilian musicians are well-known to people in Brazil and respected for their body of work. In Brazilian popular music (*Música Popular Brasileira* in Portuguese) [11], hereafter referred to as MPB, names such as Tom Jobim, Chico Buarque, and Noel Rosa are likely to be favorites. But does the social network of collaborations in Brazilian support the view of critics about musicians such as the ones mentioned above? What makes a person important to his art? This paper looks initially at collaborations between musicians from a point of view albums recorded. We have build a dataset of Brazilian albums (CDs, LPs, etc) and created a network of musicians in where they are linked if they participated together in at least one album. We then repeat the study with composers who are linked to one another if they wrote a song together. In both instances, the weight of the collaboration is given by how many times the collaboration was repeated. The goal of the study is to improve the understanding of the structure of Brazilian music as well as to use networks for providing a ranking of musicians and composers in MPB.

2. A BRIEF HISTORY OF BRAZILIAN MUSIC

Brazil is a country of continental proportions and, as such, presents a rich variety of sounds and rhythms in its music. Brazil has long been seen as a source of inspiration to many world-class musicians. It is easy to understand that the universality of the music of Brazil is a reflection of the country's history that includes native Brazilians with their rhythms and harmonies, being mixed with European (Portuguese primarily) and African sounds.

Brazilian music was also influenced by sounds from other parts of the world. By the end of the 1950s, one of the most important movements in MPB came to light: the *Bossa Nova*, which introduced to the world names such as Tom Jobim, João Gilberto and Luiz Bonfá. By the end of 1960s, the influence of rock has reached Brazil leading to a movement called *Tropicalismo* led by the likes of Caetano Veloso,

Gilberto Gil and Tom Zé. Other smaller movements: *Jovem Guarda* (driven by a need for songs with simple lyrics) *Pessoal de Minas* (from Minas Gerais State) and *Pessoal do Ceará* (from Ceará State).

What is important to notice is that these movements were influenced different styles: jazz, rock, regional sounds, country music, etc.

3. MUSICIAN COLLABORATIONS AS SOCIAL NETWORKS

The understanding of musical relationships between styles and cultures, as well as the relation between music and other sciences (particularly Math) have for a long time been of interest to musicologists, independently of the music origin, be it classical, popular, or other genre [3, 7, 14, 17]. More recently we have seen a revival of works on musical relationships due to the demand for recommendation systems in the online world [6, 10, 15]. Companies would like to know more about people's taste based on prior knowledge about their likes and dislikes. There are many approaches for recommendation systems and in one way or the other they require some understanding of musical relationships.

Since the late 1990s we have been seeing the emergence of a new multidisciplinary field, named *Network Sciences*. This field provides a framework for modeling interactions between entities so as to reveal properties at a macro level which may not be noticeable at the individual level.

Techniques from Network Sciences have been successfully applied to music. In general, the works relating music and networks do not attempt to create recommendation systems although that can be seen as a consequence of the understanding of the relationships. Park et al. [12] have described a study in which a social network of contemporary musicians have been created from the allmusic.com (AMG) and compared it with another music network in which musicians are connected based on critics views of their similarities. Gleiser and Danon [9] studied communities in Jazz using the edge-betweenness community detection algorithm from Girvan and Newman [8]. The network was created by linking musicians if they played in the same band. The community analysis found that racial divisions exists within Jazz bands with groups members being mostly black or mostly white. Gleiser and Danon have also created a jazz band network in which bands are linked if they have a musician in common. The jazz band analysis found that communities of bands are divided based on the location they generally record.

Recently, the application of concepts of complex networks have been discussed as very useful to systems dealing with music recommendation [5]. As we move increasingly towards online delivery of music and as the concept of an album is replaced by people picking and choosing individual songs they enjoy, recommendation becomes an important process to the music industry. Music recommendation systems are also crucial in a world where the availability of music can easily overwhelm the listener. In this paper, we move closer towards understanding the structure of the network of collaborations in Brazilian music which in turn may aid the development of recommendation systems for MPB.

4. BUILDING SOCIAL NETWORKS FROM COLLABORATIONS

The first step in our study was to collect a dataset related to Brazilian Music. There are many sites available online with catalogues of records (CDs, LPs) of MPB. The two most famous ones are: Ricardo Cravo Albin's dictionary of Brazilian music [1] and Maria Luiza Kfouri's personal discography [2]. Although the former is more extensive it lacks a information about the songs and the musicians of each album. We opted to go with the later because it is quite complete about musicians who participate on the record, all songs in the album, the composers of each song, and the musicians involved in the recording.

After all was done, we had a dataset with 6,149 albums of which 5,302 feature musicians. There are 506 albums with only one musician, therefore, because of the way we define an edge, these musicians would not appear in the network unless they appear in another album that feature two or more musicians. There are 16,718 musicians that contributed to 85,133 tracks. There are 10,490 composers and 1,913 artists. In order to better understand the structure of Brazilian music we concentrate on musicians (who play the music) and composers (who write the music).

4.1 Metrics

The literature in Network Sciences includes a number of metrics that can be computed to characterize a network which, in turn, may reveal interesting patterns in the relationships of nodes. The analysis of metrics related to the topology of networks have long been used in Social Networks in an area generally referred to as Social Network Analysis (SNA) [16]. In this paper we concentrate on two measures of nodes in the social networks we deal with because they enable us to rank nodes.

Node Degree: The degree of a node is a metric that refers to how many connections the entity represented by the node has in the social network. Higher degree is generally associated with a higher influence in the network because that node can quickly reach many others.

Pagerank: Although the degree looks at the importance of a node, it considers the importance in isolation. However, it is generally the case that the importance of a node depends on the importance of nodes that have a relation with it. In PageRank, important nodes pass on their importance to other nodes they are connected

[1] www.dicionariompb.com.br
[2] www.discosdobrasil.com.br

to. If an important node points to many other nodes, its importance is weighted by the number of connections it has.

4.2 Networks of Collaborations in Brazilian Music

One pre-condition to perform network analysis is to correctly chose what the nodes in the network represent and what is used for the relationship between these nodes [4]. In this paper we would like to understand the structure of Brazilian music by looking at networks of musicians and composers. These networks will allow us to move a step closer to answering questions like: *who are the seminal individuals in the Brazilian music world?*

In our first network, we look at the structure of people who play the music, what we call the Network of Musicians (NoM). Secondly, we look at who is writing the music being played, what we call Network of Composers (NoC). To create these networks we have to look at the dataset and find appropriate information by projecting the dataset on these two kids of relationships. In the NoM, a musician is linked to another if they have participated together in at least one album. For the NoC we have used composers as nodes and the relationship between them exists if they have composed some music together—Brazilian music is in fact quite unique in this sense since most songs are born out of collaborations. In both network instances, since a person can participate in more than one collaboration, we use a weighted representation of the relationship in which the weight of the edge (i, j), w_{ij}, represents the total number of albums the musicians i and j have played together for the case of NoM, and how many songs they have composed together for the case of NoC. The NoM contains 16,442 nodes and 844,223 edges, while the NoC is a much sparser network with 8,152 nodes and 12,923 edges.

5. ANALYSIS OF THE NETWORKS

The social network we analyzed contains works from more than 60 years of Brazilian music. When discussing the influence of a person in a social network the number of collaborations she has is of prime importance. In social network terms, the number of collaborations is expressed by the degree of the node in the network. For instance, if a node x representing a person collaborated with 4 others his degree, $deg(x) = 4$. Note however that degree does not consider the "size" of the collaboration, so if a person collaborated with the another 5 times, only the weighted degree, $wdeg$ captures this information. In order to have a complete picture we need both the degree (number of different collaborations) and weighted degree (number of total collaborations). We have used the entire dataset and ranked musicians and composers by the number of collaborators. Tables 1 and 2 show the rankings by degree but we also display the weighted degree.

Table 1 shows the list of musicians in Brazilian music. Most of these are probably unknown to the general public because they form what we like to call the "scaffolding of Brazilian music". With a few exceptions, these are the musicians who are respected in their art but generally do not work as leaders in recordings. Some of the numbers presented are quite impressive. Despite the incompleteness of our dataset (see Section 6 for description of our future work), we see many musicians who have collaborated with more than 2,000 others, a feat not so easily achievable. These musicians are able to carry influences from an album to another and are major contributors of cross-fertilization between brazilian styles.

Another interesting observation from Table 1 is that the national instrument from Brazil, the classical guitar (Hornbostel-Sachs number 321.322), is not present. We believe that this is the case because the musicians above belong to this "scaffolding" class which works on albums as supporting members and not as the main personnel. The table shows the importance of classical instruments even for popular music.

Table 2 describes the ranking of composers according to degrees. Here the disparities are more prominent between deg and $wdeg$. This is expected because some composers collaborate with few others but write many compositions with them. For instance, this is the case with Vinicius de Moraes (in bold in Table 2) has deg=59 but $wdeg$=3,392. It is worth noticing that our $wdeg$ is based on the total number of compositions that appears in the dataset (not on unique compositions); this choice is made on purpose for the composers study because we want $wdeg$ to be more than just a count of different compositions but also give a notion of importance of the individual. If a composers has then one collaboration (deg=1) but that composition has been recorded 1,000 times in the dataset, his $wdeg$=1,000. For us that composer is important to the structure of brazilian music although she has not composed many pieces—she would be important because his composition has been frequently recorded.

Table 2. List of top 30 composers by the number of different collaborations (deg). However some of the collaborations are repeated, meaning that the composers may write more than one song with a collaborator. The weighted degree ($wdeg$) column is an indication of repeated collaborations. Names in shown in bold are used as specific examples in the text.

deg	Name	$wdeg$	deg	Name	$wdeg$
83	Paulo César Pinheiro	1,047	*50*	Chico Buarque	1,186
74	**Arnaldo Antunes**	411	*47*	Francis Hime	488
65	Caetano Veloso	320	*45*	Moraes Moreira	344
61	Aldir Blanc	960	*45*	Ataulfo Alves	257
61	Ivan Lins	668	*44*	Nei Lopes	192
60	Milton Nascimento	917	*44*	Tom Zé	142
60	Gilberto Gil	427	*44*	Martinho da Vila	118
59	**Vinicius de Moraes**	3,392	*43*	Wilson Batista	260
59	Noel Rosa	731	*43*	Itamar Assumpção	115
59	Luiz Gonzaga	706	*42*	Heitor Villa-Lobos	263
57	João Donato	474	*41*	**Carlinhos Brown**	212
56	Nelson Cavaquinho	634	*40*	**Pedro Luís**	72
52	Ronaldo Bastos	364	*39*	Tom Jobim	2,486
52	Hermínio Bello de Carvalho	348	*39*	**Zeca Baleiro**	89
52	Délcio Carvalho	190	*37*	Fausto Nilo	165

Table 1. List of top 30 musicians by the number of different collaborations (deg). However some of the collaborations are repeated, meaning that the musicians may play in many albums with the same musicians. The weighted degree ($wdeg$) column is an indication of repeated collaborations.

deg	Name	$wdeg$	Instrument	deg	Name	$wdeg$	Instrument
3,002	Márcio Eymard Mallard	15,635	Cello	*2,062*	Jorge Helder	7,633	Bass
2,782	José Alves da Silva	15,122	Violin	*2,051*	Wilson das Neves	7,891	Drums
2,659	Jorge Kundert Ranevsky	13,474	Cello	*1,990*	Ricardo Amado	7,878	Violin
2,579	Paschoal Perrota	13,398	Violin	*1,990*	Gordinho	7,477	Percussion
2,563	Jaques Morelenbaum	10,558	Cello	*1,967*	Alfredo Vidal	9,864	Violin
2,470	Walter Hack	13,290	Violin	*1,949*	Jamil Joanes	7,317	Bass
2,445	Alceu de Almeida Reis	12,319	Cello	*1,897*	Jesuína Noronha Passaroto	7,968	Viola
2,405	Robertinho Silva	7,135	Drums	*1,862*	Zé Carlos Bigorna	6,593	Sax, Flute
2,400	João Daltro de Almeida	11,280	Violin	*1,849*	Aizik Meilach Geller	9,263	Violin
2,331	Carlos Eduardo Hack	11,696	Violin	*1,810*	Ovídio Brito	5,376	Percussion
2,314	Frederick Stephany	11,071	Viola	*1,804*	Nailor Proveta	4,438	Sax
2,251	Bernardo Bessler	9,751	Violin	*1,792*	Cristóvão Bastos	8,373	Piano
2,268	Giancarlo Pareschi	12,405	Violin	*1,771*	Márcio Montarroyos	7,243	Trumpet
2,251	Michel Bessler	10,254	Violin	*1,759*	Carlos Malta	4,529	Flute
2,201	Marcos Suzano	5,640	Tambourine	*1,748*	Marie Christine Springuel	7,076	Viola

The list in Table 2 is somewhat surprising at first because of names such as Arnaldo Antunes, Carlinhos Brown, Pedro Luís, and Zeca Baleiro (also in shown in bold). However these names represent the new generation of Brazilian composers who make very good use of social media and collaborate with many other musicians. The rankings in the table considerers all data in the dataset. To better understand the evolution of these rankings we performed a temporal analysis but using pagerank rather than degree ranks.

The first study we have performed using pagerank is shown in Figure 1. These ranks are per decade and follow the position of the top 50 musicians and composers in the most recent decade. It is important to understand that the ranking in decades other than the most recent one is relative to each other. We took the top 50 musicians and composers in the most recent decade and followed their relative ranks in other decades. For instance, Noel Rosa appears as the top ranked composer in Figure 1(right) for the most recent decade but in the 14^{th} position in the 80s; this 14^{th} means relative to the 50 composers listed in the 2000 decade. In absolute terms, Noel Rosa can (it probably is) lower than the 14^{th} position. The connections in the social networks for each decade considers only the collaborations in albums of that decade, which explain sudden changes in the rankings. A musician or composer that was top in a decade may be irrelevant in others because he was not active or because his compositions were not recorded by musicians in that period.

We can see in Figure 1 that composers ranks are more stable than the musicians meaning that the relative ranks are better maintained for composers (the lines not cross as often and as radically). We can also argue that musicians do not have as high longevity as composers. A musician who is very active today may not have been very active a few years back. A clear example of this is the musician Michel Bessler who does not even appear in albums prior to 1980 although he is 10^{th} most important musician of today. Bessler is the spalla of the Brazilian Symphony Orchestra and has participated in many popular albums (see Table 1). Figure 1 confirms this longevity observation, which is expected since the NoM requires active participation of the musician in the recording while the NoC includes people who may have even be deceased but continue to have their music recorded (e.g. Noel Rosa).

Last we look at the evolution of rankings using accumulative networks. While Figure 1 looks at collaborations in isolation, Figure 2 shows the ranking (also according to pagerank) of musicians and composers using an accumulative approach. Here we want to see how the ranks evolve if we consider the collaborations until a particular year but including all information since the first date we have information on the dataset. For instance, the ranking in 2010 considers all the works available in the dataset, that is, the full collaboration network. Antecedent years (2007, 2004, etc.) consider collaborations from the first data available in the dataset until the given year. Hence, the change from one year to another (3 years appart) is due to the work produced in the last 3 years.

The use of accumulated networks allows us to see a little better how the structure changes as new musicians and composers become active. An excellent example of this is Arnaldo Antunes who appears in Figure 2(right) in 4^{th} position but decrease his relative rank quite rapidly until disappearing completely in 1986. Arnaldo Antunes appeared to in Brazilian music scene as a member of a rock band called Titãs in the mid-80s. After leaving the band, he emerged as one of the main composers in Brazil with many collaborators (which influences his pagerank). Most recently, his compositions have been part of recordings of many respected brazilian singers such as Marisa Monte and Cássia Eller.

Another interesting class of composers that we can see in

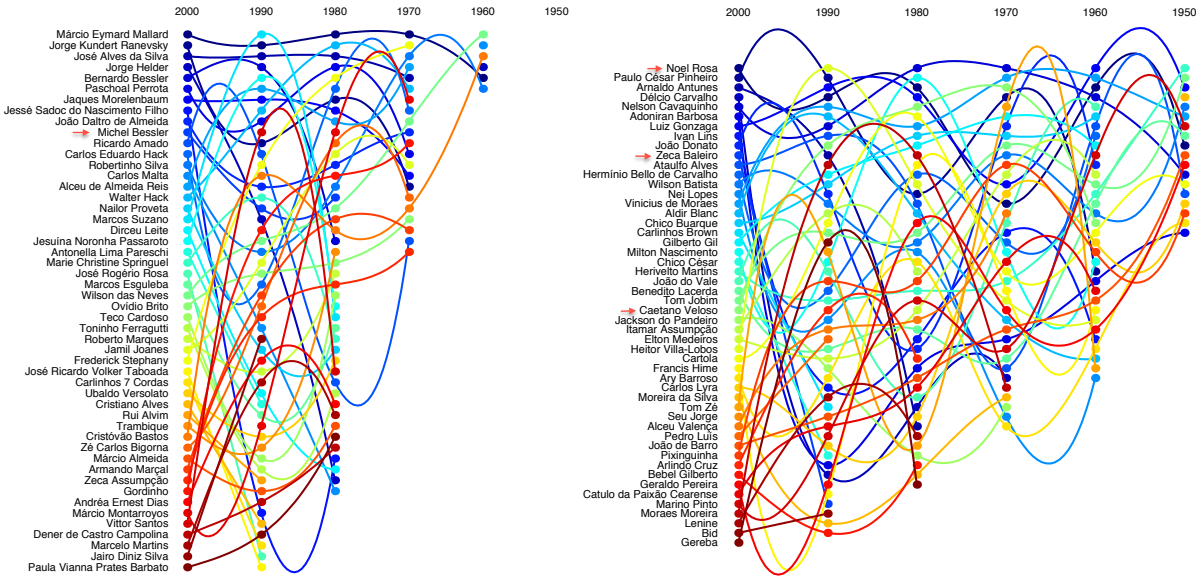

Figure 1. Rank of the top 50 musicians (left) of the last decade and how these ranks evolve per decade. The pagerank of 2000 is absolute but for the other decades it represents how these musicians rank against each other. For instance, none of the musicians ranked today were present in recordings from 1950s (1950-1959). On the right picture, we have the rank of the top 50 composers of the last decade and how these ranks evolve per decade. Individuals marked with an → are examples discussed in the text.

Figure 2 is well represented by Caetano Veloso. The accumulated ranking shows that Caetano Veloso has maintained himself active through several decades (by composing and having his songs recorded by other artists) and he is today still the 5^{th} most important composer in Brazil. Compare this to his position in Figure 1; since that analysis considers only recordings per decade in isolation we see that Caetano Veloso is not so well positioned in more recent years. The fact is that Figures 2 and 1 taken together give us a good idea of the ranking of a musician and composer and how it evolves.

Lastly, our results allow us to observe scenarios like what happens to Zeca Baleiro in Figure 1. Because the study takes decades in isolation we can see that he appears high in the rankings but not at all in the accumulative ranking in Figure 2. This is a case where we have a upcoming composer who has been active only very recently and is part of the ranking of the last decade but not yet part of the entire history.

6. CONCLUSION AND FUTURE WORK

In this paper we demonstrated that the construction of social networks and the use of metrics rooted in network sciences may help us understand the structure of Brazilian music. Rankings related to music are always controversial because of the attachment people feel to music. However we believe our approach is less biased and provides a good understanding of the structure of Brazilian music. Our work shows that the network of musicians is less stable than the network of composers. This result is expected because musicians actively participate in the recording while composers participate by having their songs recorded—a composer may even be deceased when his composition is recorded. Our hope was to have a social network of musicians based on them playing together on specific tracks rather than on an album as we believe this is a more accurate representation of the collaboration. However to our knowledge, no dataset of MPB includes the information per track.

The two kinds of rankings provided (and their visualization) also allows us to understand how the rankings change with time. An analysis not included in this paper (due to space restriction) seem to indicate that a composer needs to be well ranked for about 30 years to appear in the accumulative rankings. This appears to indicate that 30 years for Brazilian music a measure of "success" for a composers—what differentiates them from one-hit composers.

We continue to work on the current dataset on many fronts. We are currently collecting more data to make the dataset more complete since it is still incomplete particularly with regards to older recordings. Next, we intend to consider the date of the composition in our analysis although this data is appearing to be very hard to gather. With this information we believe we can have another dimension of the structure of composers. Last, our ultimate goal is to be able to add the concept of reputation to the study and for that we may have to consider a third category of individuals. A composer may become part of the rankings by having one of his compositions recorded by major singers (e.g. Elis Regina). We

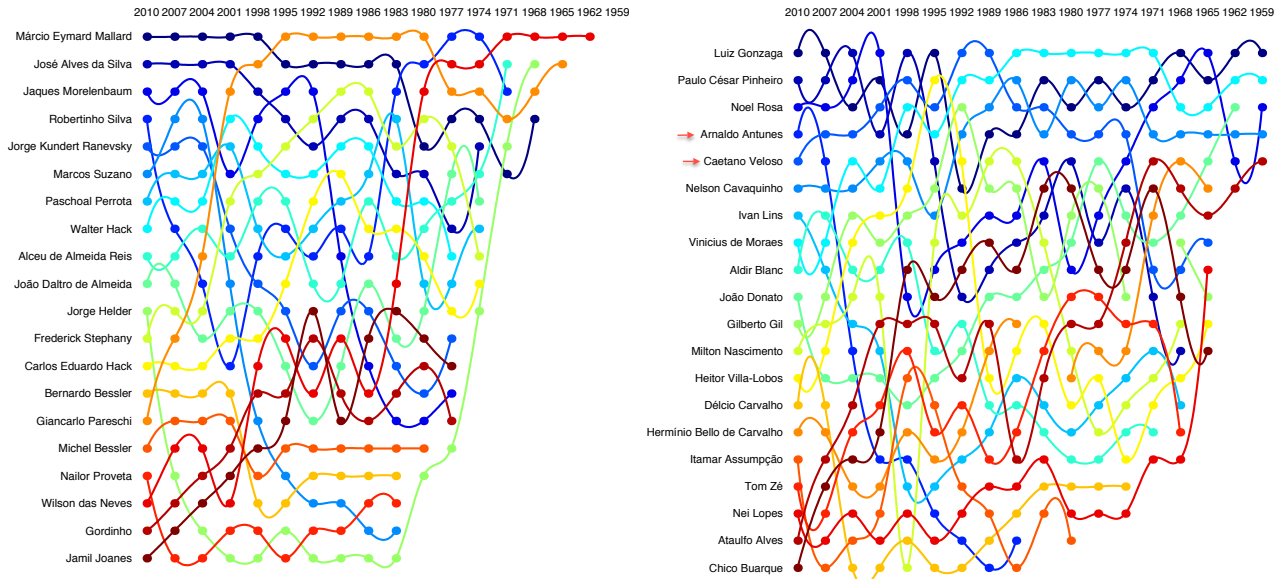

Figure 2. This picture shows the rank of the top 20 musicians (left) and composers (right) for the year 2010 and how they rank against each other in the antecedent years. Note that the ranks for all other years are not absolute. This means that if a person is listed in the 1^{st} position it only means that the person is in the 1^{st} position relative to the other people listed in the year 2010. In this case, the network of collaborations is not taken in isolation, so the data for the year 2007 includes all collaborations until 2007. Individuals marked with an → are examples discussed in the text.

are currently considering how this reputation can be added to the study given that some of these singers and have never composed songs are not musicians either.

7. REFERENCES

[1] Rafael Bastos. The origin of samba as the invention of brazil (why do songs have music?). *British Journal of Ethnomusicology*, 8 IS -:67–96, Jan 1999.

[2] Gerard Béhague. Rap, reggae, rock, or samba: The local and the global in brazilian popular music (1985-95). *Latin American Music Review / Revista de Música Latinoamericana*, 27(1):79–90, Apr 2006.

[3] Wendy S. Boettcher, Sabrina S. Hahn, and Gordon L. Shaw. Mathematics and music: A search for insight into higher brain function. *Leonardo Music Journal*, 4:53–58, 1994.

[4] Carter T. Butts. Revisiting the Foundations of Network Analysis. *Science*, 325(5939):414–416, 2009.

[5] Òscar Celma. *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer Verlag, 2010.

[6] Hung-Chen Chen and Arbee L. P. Chen. A music recommendation system based on music and user grouping. *Journal of Intelligent Information Systems*, 24:113–132, 2005.

[7] Janet M. Cliff. On relationships between folk music and folk games. *Western Folklore*, 51(2):129–151, 1992.

[8] M Girvan and Mark E Newman. Community structure in social and biological networks. *PNAS*, 99(12):7821–7826, Jun 2002.

[9] Pablo M. Gleiser and Leon Danon. Community structure in jazz. *Advances in Complex Systems: A Multidisciplinary Journal*, 6(4):565–573, 2003.

[10] Beth Logan. Music recommendation from song sets. In *Proceedings of the International Society for Music Information Retrieval Conference*, pages 425–428, 2004.

[11] Chris McGowan and Ricardo Pessanha. *The Brazilian Sound*. Temple University Press, 1998.

[12] Juyong Park, Òscar Celma, Markus Koppenberger, Pedro Cano, and Javier M. Buldú. The social network of contemporary popular musicians. *International Journal of Bifurcation and Chaos*, 17(7):2281–2288, 2007.

[13] Maximilian Schich, Sune Lehmann, and Juyong Park. Dissecting the canon: Visual subject co-popularity networks in art research. In *International Workshop on Challenges and Visions in the Social Sciences*, 2008.

[14] John Shepherd. Music, culture and interdisciplinarity: Reflections on relationships. *Popular Music*, 13(2):127–141, 1994.

[15] Ja-Hwung Su, Hsin-Ho Yeh, P.S. Yu, and V.S. Tseng. Music recommendation using content and context information mining. *IEEE Intelligent Systems*, 25(1):16–26, 2010.

[16] Stanley Wasserman. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.

[17] Olly Wilson. The significance of the relationship between afro-american music and west african music. *The Black Perspective in Music*, 2(1):3–22, 1974.