

FINDING COMMUNITY STRUCTURE IN MUSIC GENRES NETWORKS

Débora C. Corrêa, Luciano da F. Costa

Instituto de Física de São Carlos

Universidade de São Paulo

debcris.cor@gmail.com, luciano@ifsc.usp.br

Alexandre L. M. Levada

Departamento de Computação

Universidade Federal de São Carlos

alexandre@dc.ufscar.br

ABSTRACT

Complex networks have shown to be promising mechanisms to represent several aspects of nature, since their topological and structural features help in the understanding of relations, properties and intrinsic characteristics of the data. In this context, we propose to build music networks in order to find community structures of music genres. Our main contributions are twofold: 1) Define a totally unsupervised approach for music genres discrimination; 2) Incorporate topological features in music data analysis. We compared different distance metrics and clustering algorithms. Each song is represented by a vector of conditional probabilities for the note values in its percussion track. Initial results indicate the effectiveness of the proposed methodology.

1. INTRODUCTION

Complex networks have received much attention in recent years due to their capability of characterizing and helping in the understanding of many interdisciplinary aspects of the real-world [3]. Regarding music and artistic aspects, music networks have been studied and their topological characteristics shown to be useful for the analysis of dynamics and relations between the involved elements. Examples are the work of Gleiser and Danon [13] concerning a collaboration network of jazz artists and bands; the work of Parket *et al* [8] about a social network of contemporaneous musicians; and the work of Cano *et al* [12] involving an analysis of the similarities between songs and bands.

Community structures have also been studied in music networks. Teitelbaum *et al* [19] analysed two different social networks using similarities and collaborative attributes of music artists. They described some organization patterns and they comment aspects that reflect in the growth of such networks. Lambiotte and Ausloos [17] addressed the diffi-

culty for a general agreement of the genre taxonomy through an empirical analysis of web-downloaded data.

Although there are several works in the literature that provide significant results for more complex case of audio-based analysis [7], in audio files all information is mixed together. Differently, the use of symbolic format like MIDI, may indicate a clearer analysis of what is in fact contributing for the discrimination of the genres [15]. On the other hand, Markov models on high-level rhythm features is an area relatively few explored nowadays. Markov chains in rhythm features and their capability for discriminating music genres has been studied by [3]. The authors investigated that use of Markov chains with memory one and two suggests an evidence that the pattern of note values in the percussion may differ from one genre to another.

Our main goal is to analyse the community structure of music networks, which is a new and promising research area. We believe that mixing temporal features (rhythmic patterns) and global topology information from proper music networks can be effective in understanding the relationship of music genres. We summarize our main contributions as: comparison of different 1) distance metrics, and 2) community detection algorithms in order to find community structures in the music networks, defining a completely unsupervised and low computational cost approach.

The remainder of the paper is organized as follows: section 2 describes the proposed method; and section 3 presents the primarily experiments and provide some discussions. Finally, section 4 shows the conclusions and final remarks.

2. METHOD

2.1 Data Description

The database consists of 280 samples (or songs) in MIDI format equally divided into four genres: blues, *mpb* (Brazilian popular music), reggae and rock. Although it indicates a small database, these songs contain high variability in their rhythmic patterns. Besides, this database allows a qualitative investigation of the music graphs (by visual inspection of their topology). Our motivation for choosing these four genres is the availability of online MIDI samples with con-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2011 International Society for Music Information Retrieval.

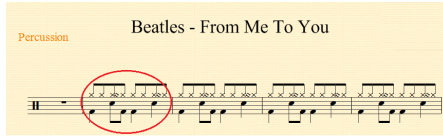


Figure 1. Example of a percussion track.

Beat	4	4	4.5	5	5	5	5.5	5.5	6	6	6.5	7	7	7
Relative Duration	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5	1

Table 1. Matrix representation of second measure of the percussion in Figure 1. First beat starts at 0.

siderable quality and the different tendencies they represent.

Despite being simpler to analyse than audio files, MIDI formats have the advantage of being a symbolic representation, which offers a deeply analysis of the involved elements and takes much less space. We used the Sibelius software and the free Midi Toolbox for Matlab computing environment [18]. In this toolbox a MIDI file is represented as a note matrix that provides information like relative duration (in beats), MIDI channel, MIDI pitch, among others. The relative note duration is represented in this matrix through relative numbers (for example, 1 for quarter note, 0.5 for eighth note, 0.25 for sixteenth note and so on). Sibelius software has an option called “Live Playback”. If this option is not marked, the note values in the MIDI file respects their relative proportion (e.g., the eighth note is always 0.5). In this way, we can solve possible fluctuations in tempo.

For each song the track related to the percussion is extracted. We propose that the percussion track of a song is intrinsically suitable to represent the rhythm in terms of note values dynamics. Once we have separated the percussion track, we can obtain a vector that contains the sequence of relative note values present in it. The instrumentation is not been considered. If two or more note events occurs at the same beat, the median duration of them is taken. To illustrate the idea, Figure 1 shows the first measures of the percussion track of the music *From Me To You* (The Beatles).

Part of the percussion matrix corresponding to the second measure is indicated in Table 1. As we can see, different instrument events occur at a same beat. Taking the median value in such cases, the final note duration vector of this measure will be: [0.5 0.5 0.5 0.5 0.5 0.5 0.5]. For each song in the database, we compute the note vector of the whole percussion. All these steps can be automatically performed.

2.2 Markov modeling for note duration dynamics

Markov chains use a conditional probability structure to calculate the probability of future events based on one or more past events [5]. We can analyse different numbers of past

events, which indicates the order of the chain. A first order Markov chain takes into consideration only a predecessor of an event. If instead, the predecessor’s predecessor is considered, then we have a second order Markov chain, and so on. Generally, an n th-order Markov chain is represented by a transition matrix of $n + 1$ dimensions. This is an interesting matrix, since it gives the information about the likelihood of an event’s occurrence, given the previous n states.

In our case, the events are the relative note values of the percussion in the songs, obtained with the steps described in section 2.1. For each song (represented by a vector of note values), we compute the first and second order transition matrices. Therefore, we have the probability that each note value or a pair of note values is followed by other note duration in the song. Higher-order Markov chains tend to incorporate senses of phrasal structure [2], while first-order ones help to identify more often subsequent notes.

In order to reduce data dimensionality, we performed a preliminary analysis of the relative frequency of note values and pairs of note values concerning all the songs, in a way that extremely rare transitions were discarded. For the first order Markov chain we have a matrix of probabilities with 18 rows and 18 columns (we considered 18 different note values in this dataset). Each entry (i, j) of this matrix expresses the probability that a note value i is followed by a note value j in the percussion of the respective song. Then this matrix is treated as a 1×364 feature vector.

For the second order Markov chain, the matrix of probabilities for each song is 167 (rows) \times 18 columns, treated as a 1×3006 ($167 * 18$) feature vector (we considered 167 different pair of note values). Similar, each entry (i, j) of this matrix expresses the probability that a specific pair of note values represented in line i follows a specific note value j . If we concatenate both feature vectors we will have the final feature vector of each song with 3330 elements. It is interesting to mention that, we experimented to built the music networks considering first and second order probabilities separately. However, for both isolated cases, the Clauset-Newman-Moore community detection algorithm clustered 5 different groups, while considering feature vectors composed by the concatenation of first and second order models led to the detection of 4 groups. This fact suggests that a single Markov chain is not sufficiently to model all the dynamics that characterizes the 4 original genres. Another evidence is that when we consider both Markov chains, the accuracy obtained in the classification of these four genres is higher: 70% for first-order Markov chain, 85% for the second-order, and 92% for both chains. (We used the Bayesian classifier under Gaussian hypothesis.)

2.3 Music Networks

A complex network is a graph that exhibits a relatively sophisticated structure between its elements when compared

to regular and uniformly random structures. Basically speaking, a network may be composed by vertices, edges (or links) and a mapping that associates a weight in the connection of two vertices. The edges usually has the form $w(i, j)$ indicating a link from vertex i to vertex j with weight $w(i, j)$. Representing music genres as complex networks may be interesting to study relations between the genres characteristics, through a systematic analysis of topological and structural features of the network.

From the first and second-order transition matrixes of the Markov chains we can build a music network. Each vertex represents a song. The links between them represent the distance of the two respective songs, considering their vectors of conditional probabilities of the note values. However, with a full-connected network it may be difficult to obtain intricate structure. There are several forms to define which vertices will be connected and several distance metrics. We propose some possibilities in the following and try to form clusters of vertices that can represent the music genres.

3. EXPERIMENTS AND DISCUSSION

It is worthwhile to mention that the proposed characterization of the music genres is performed in an unsupervised way (community finding algorithms). The obtained groups are based on similarities in the feature set and the classes are not supposed to be known in advance. To illustrate the complexity of the problem, Figure 2 presents the first and second components (new features) obtained by LDA (Linear Discriminant Analysis), which is a supervised technique for feature analysis whose principal aim is to maximize class separability. Even with the LDA new features, reggae and rock classes are still overlapped. This overlapping could be observed in all performed experiments. Considering the rhythms patterns, rock and reggae music are pretty similar.

We know that the use of only four genres with seventy samples each may represent a small dataset. Our purpose is to perform an initial study of rhythmic features and its representation, but with an evidence that the proposed features may be useful and viable for genre characterization.

3.1 Community detection on K -NN graphs

Through the dynamics of the note values in the percussion we built several networks. From the point of view of partitioning the genres into communities, different groups may be obtained, depending on the used criteria. In this section, we used the Clauset, Newman and Moore [1] and Girvan and Newman [10] algorithms for community detection. Such algorithms are widely known in the complex networks literature. The former is based on a hierarchical clustering of the dataset. The latter is based on centrality metrics to determine the community boundaries.

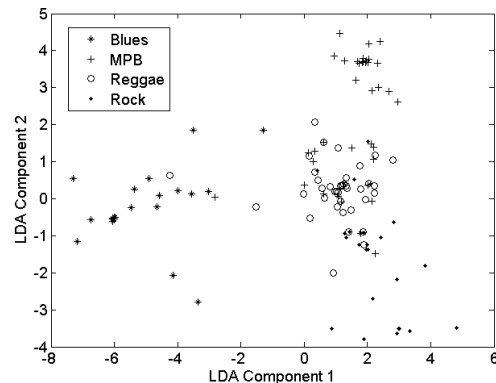


Figure 2. The first and second features obtained by LDA.

	<i>Mpb</i>	Rock	Blues	Reggae
G1	41	29	15	21
G2	4	13	1	11
G3	17	15	50	20
G4	8	13	4	18

Table 2. The groups in network of Figure 3.

For each of the following cases, the networks may be built as follows: 1) From the feature matrix (with 280 lines (the songs) and 3330 columns (the features)), we computed the distance between each pair of feature vector (or each pair of song). This led to a 280×280 symmetric matrix of distances, with zero values in the diagonal. In this case, we have a full network, with all vertices connected to each other; 2) For each song (or vertex), we only link the K nearest songs of it. The weight of each link is the distance between this pair of songs; 3) Consider the obtained K -regular network. Or; 4) For each vertex, take the mean distance, considering the linked vertices. Keep the link between vertices only with their distance is smaller than the mean distance. The main variations of the networks analysed here are consequence of the choice of different distance metrics, different values of K , and the execution or not of step 4.

For the network showed in Figure 3 we used the cosine distance, $K = 10$ and kept the network 10 regular. The songs are spread as indicated in Table 2. Each group has a different dominant class. Blues and *mpb* songs are concentrated in G3 and G1, respectively. Reggae songs are almost equally divided into the groups. Rock songs are almost 50% in G1, overlapping with *mpb* songs. The other 50% divided into the remaining groups. This behavior substantially reflects the projections of LDA in Figure 2. The G3 group reflects the blues songs that are more discriminative. The G1 group reflects mainly the overlapping present in *mpb*, reggae and rock. And G2 and G4 mainly reflect the overlapping between reggae and rock songs.

For the same network, Figure 4 shows the groups ob-

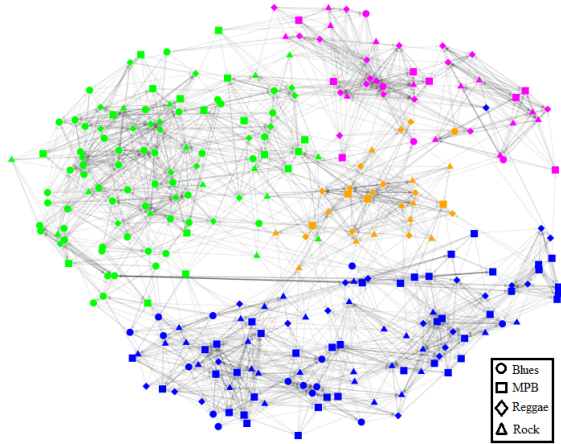


Figure 3. The network of genres. Cosine distance. Groups formed by the Clauset-Newman-Moore algorithm. All colored images available at <http://cyvision.ifsc.usp.br/deboracorreia/MusicandComplexNetworks.html>

tained by the Girvan and Newman algorithm. Since it is an algorithm based on vertex centrality indices, the network was split into nine groups. The result is still interesting since many songs of a same genre are placed together in each group. In addition, this result opens a promising further studies aimed at analysing the presence of sub-genres in these small groups. Are, for example, blues-rock or pop-rock songs more concentrated in a specific group? This is an interesting study that can benefit of this investigative work.

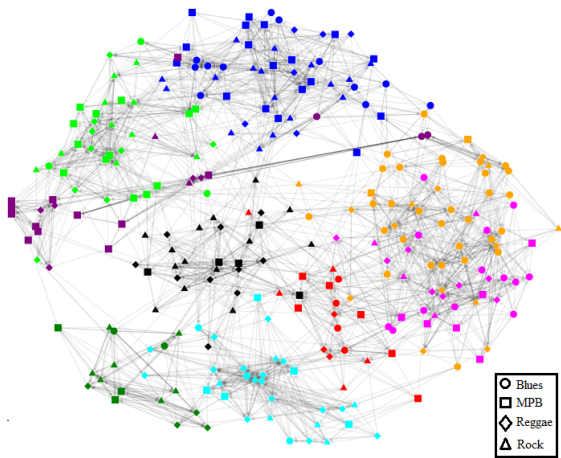


Figure 4. The network of genres. Cosine distance. Groups formed by the Girvan and Newman algorithm.

If instead of cosine distance, we use the Euclidian distance, we will get the network in Figure 5, according to the Clauset, Newman and Moore algorithm. Table 3 shows the groups. Reggae songs are more concentrated (31 in G3); and G4 is smaller than in the first case, with only 12 songs.

Considering all the experiments, including those not pre-

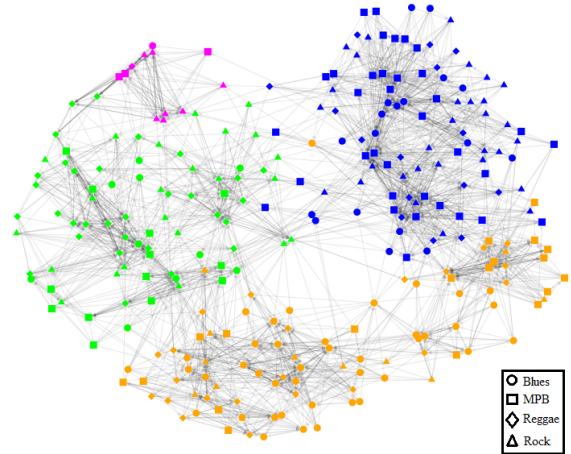


Figure 5. The network genres. Euclidian distance. Groups formed by the Clauset-Newman-Moore algorithm.

	<i>Mpb</i>	Blues	Reggae	Rock
G1	35	18	15	30
G2	19	41	23	10
G3	13	10	31	23
G4	3	1	1	7

Table 3. The groups in network of Figure 5.

sented here, we can describe some overall characteristics of the clusters found by the Clauset-Newman-Moore algorithm. The most discriminative genre is blues. In most experiments one group was always small. Actually, in some variations the algorithm returned three large groups. This may indicate that, although we have four genres labeled by the usual taxonomy, in terms of the proposed rhythm features there are only three. If we listen to the whole song, we may differ the genres in a successful way. But if we listen to only the percussion track of each song, this discrimination may be harder and one song could be labeled into more than one genre. Therefore, considering that we have a completely unsupervised approach, the proposed investigation indicates that note duration dynamics can be a useful information in characterizing and discriminating music genres.

3.2 Spectral graph partitioning

Topologic-based graph metrics are generally correlated and dependent [16]. For this reason, spectral analysis is a powerful tool that has been widely explored in the characterization of graphs and complex networks. The basic idea can be summarized as follows: in mathematical terms, when we analyze a graph in the spectral domain we have a representation in terms of orthogonal components, which means that information is somehow uncorrelated. Thus, proper analysis of eigenvalues and eigenvectors of adjacency or laplacian matrices identifies aspects that cannot be seen in the topol-

	Reggae	Blues	Rock	<i>Mpb</i>
G1	20	4	13	8
G2	21	46	6	15
G3	17	9	23	20
G4	12	11	18	27

Table 4. The groups in network of Figure 6.

ogy domain. Please, refer to [4, 11] for a good review on the mathematical fundamentals of algebraic graph theory.

In this paper, we use a spectral graph partitioning method based on the analysis of the eigenvalues of the Laplacian matrix. Let \mathbf{A} and \mathbf{B} be the adjacency and incidence matrices of a graph $G = \{V, E\}$, where V is a set of vertices and E is a set of edges. The Laplacian matrix, \mathbf{Q} , is given by:

$$\mathbf{Q} = \mathbf{B}\mathbf{B}^T = \Delta - \mathbf{A} \quad (1)$$

where Δ is a diagonal matrix of the degrees of V .

The second smallest eigenvalue of the Laplacian matrix is known as the algebraic connectivity of a graph and it has many interesting properties. More precisely, the eigenvector associated to this eigenvalue, known as the Fiedler vector [9], has proven to be directly related to graph connectivity. Often, in practice, the signs of the Fiedler vector can be used to partition a graph in two regions. This can be seen as a quantization to binary digits, zero or one.

Here, we propose to do a quantization of the Fiedler vector coefficients in C values, where C represents the number of desirable clusters or groups. By doing so, we are essentially partitioning a graph or network in C subgraphs or communities, which is equivalent to finding $C - 1$ valleys in the histogram that represents the distribution of its coefficient values. In this paper, the thresholds were chosen by visual inspection of the histogram, but several methods for automatic multilevel threshold estimation are available in the image processing literature [14]. A deeper mathematical analysis and discussion about the eigenvectors of the Laplacian matrix and its properties can be found in [16].

For the following experiment, we used the non-regular network generated by first building a K-NN graph with $K = 30$ and then, for each vertex v , cutting the edges whose weights were above a threshold obtained by averaging the weights of every edge incident on v . Thus, the resulting network is not modeled as a k-regular graph anymore. Figure 6 shows the resulting network, with the four detected clusters. The Fiedler vector for this graph and the corresponding histogram for the distribution of its coefficients are plotted in Figures 7 and 8, respectively. The distribution of coefficient values of the second smallest eigenvector of the Laplacian matrix clearly indicates the presence of different clusters or communities in the network. Table 4 shows the groups for the spectral partition. Rock and *mpb* songs are more spread in the four groups than in the former cases.

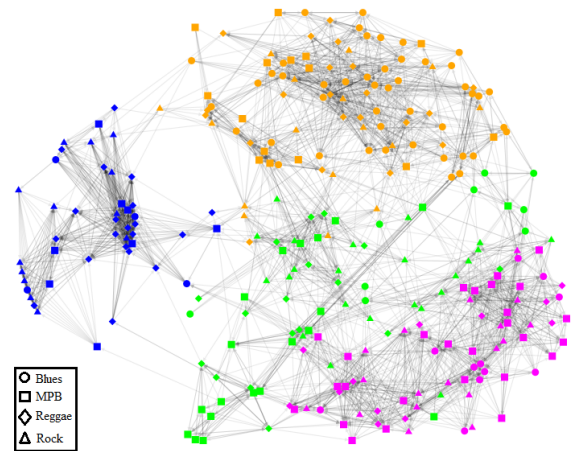


Figure 6. The network of genres by the Fiedler vector.

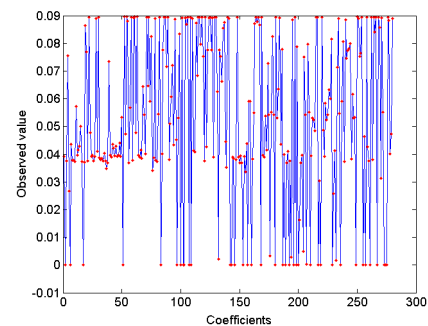


Figure 7. The Fiedler vector for the network in Figure 6.

4. FINAL REMARKS AND ONGOING WORK

In this investigative study we proposed a characterization of music genres by detecting communities in complex music networks. Each vertex represents a song through a feature vector that captures the likelihood of first and second order Markov chains of the note values in the percussion track. The distance between the feature vectors (or between the songs) defines the weight of the links. We tested two different distance metrics (cosine and Euclidian) and two different approaches for finding clusters in the network (traditional algorithms on K-NN graphs and spectral partitioning).

Regarding the formed clusters, we found that the results are promising since in most experiments each cluster is dominated by a different genre. Observing the LDA projections, it is possible to see that many samples from different genres are overlapped (mainly reggae and rock samples). LDA is a supervised technique that maximizes class separability. Therefore, even without any supervised analysis, significant results could be obtained. In addition, most MIDI databases available in the Internet are single-labeled, sometimes with different taxonomies of music genres. In some situations,

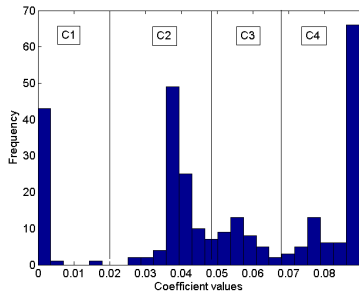


Figure 8. Distribution of coefficient values of the Fiedler vector for the network depicted in Figure 6.

a sample receives different labels in different sites (for example, wikipedia). This introduces noise to the system and reflects in the evaluation of the results.

From the obtained communities and considering the four genres used in this study, we can say that blues is the more discriminative genre. Representing the older genre, and having specific characteristics, blues may have influenced the following genres, which contributed along years for a mixture of some features between genres. Reggae, rock and *mpb* are more similar genres, sharing many overlapped samples. In fact, along years *mpb* music started to include different rhythms like rock and latine music such as reggae and samba. Reggae music, on the other hand, had stylistic origins in jazz, R&B, rocksteady and others. These tendencies are interesting and are somehow reflected in the results. Actually, the use of graph representation (instead of clustering methods in a vector space) is promising, since it combines graph topological features and similarity characteristics in order to infer the data structures.

Music networks is somehow a new research area in the literature. To the best of our knowledge, we could not find a different approach that used partitional network methods for music genres. Comparing with the hierarchical clustering with Euclidian distance metric used in [3], the groups in Table 3 have some differences: the blues songs are significantly more concentrated in one group; the largest group does not concentrate too many samples of all genres, which is not the case in the hierarchical clustering. An advantage of this kind of unsupervised analysis relies on the possibility of the characterization of music sub-genres, which can contribute to the definition of a more unified taxonomy.

There are many possibilities for future works. First, many other rhythm attributes can be analysed (like the intensity of the beat), as well as other open music databases [15]. Another interesting work that has been started is the investigation of sub-genres present in sub-clusters of the main groups. It would be promising if a system could be sensitive to various styles inside a genre. Contextual analysis through Markov Random Field models may also bring ben-

efits, since with this kind of modeling we can measure how individual elements are influenced by their neighbors, analyzing spatial configuration patterns of vertices.

5. ACKNOWLEDGMENTS

Debora Correa thanks Fapesp financial support (2009/50142-0) and Luciano da F. Costa thanks CNPq (301303/06-1 and 573583/2008-0) and Fapesp (05/00587-5) financial support.

6. REFERENCES

- [1] A. Clauset, M. E. J. Newman, C. Moore: "Finding Community Structure in Very Large Networks," *Phys. Rev. E* Vol. 70, No. 066111, 2004.
- [2] C. Roads : *The Computer Music Tutorial*, MIT Press, 1996.
- [3] D. C. Correa, J. H. Saito, L. da F. Costa: "Musical Genres: Beating to the Rhythms of Different Drums", *New Journal of Physics* Vol. 12, N. 053030, 2010.
- [4] D. M. Cvetkovic, M. Doob and H. Sachs: *Spectra of Graphs, Theory and Applications*, Johann Ambrosius Barth (Heidelberg), 3 ed., 1995.
- [5] E. Miranda: *Composing with computers*, Focal Press, Oxford, 2001.
- [6] J. Clark and D. A. Holton: *A First Look at Graph Theory*, World Scientific, 1991.
- [7] J-J. Aucouturier and F. Pachet "Representing Musical Genre: A State of the Art", *J. of New Music Research* Vol. 32, No. 1, pp.8393, 2003.
- [8] J. Park, O. Celma, M. Koppenberger, P. Cano, and J. M. Buld: "The social network of contemporary popular musicians", *International Journal of Bifurcation and Chaos* Vol.17. N. 7, pp. 2281-2288, 2007.
- [9] M. Fiedler: "Algebraic Connectivity of graphs", *Czechoslovak Mathematical Journal*, Vol. 23, No. 98, pp. 298-305, 1973.
- [10] M. Girvan and M. E. Newman: "Community structure in social and biological networks", *Statistical Mechanics - Proc. Natl. Acad. Sci. USA* Vol.99, pp. 7821-7826, 2002.
- [11] N. Biggs: *Algebraic Graph Theory*, Cambridge Univ. Press, 1994.
- [12] P. Cano, O. Celma, M. Koppenberger, e J. M. Buld: "Topology of music recommendation networks.", *Chaos* Vol.16, N.013107, 2006.
- [13] P. M. Gleiser and L. Danon: "Community structure in jazz", *Advances in Complex Systems* Vol 6 N. 4, 2003.
- [14] P. S. Liao, T. S. Chen and P. C. Chung: "A Fast Algorithm for Multi-level Thresholding", *Journal of Information Science and Engineering*, Vol. 17, pp. 713-727, 2001.
- [15] C. McKay, I. Fujinaga "Automatic Genre Classification Using Large High-Level Musical Feature Sets", *Proc. of the International Conference on Music Information Retrieval*, pp. 525-530, 2004.
- [16] P. V. Mieghem: *Graph Spectra for Complex Networks*, Cambridge Univ. Press, 2011.
- [17] R. Lambiotte and M. Ausloos: "On the genre-fication of music: a percolation approach", *The European Physical Journal B - Condensed Matter and Complex Systems* Vol. 50, N. 1-2, pp.183-188, 2006.
- [18] T. Eerola and P. Toiviainen: *MIDI Toolbox: MATLAB Tools for Music Research*, University of Jyväskylä, 2004
- [19] T. Teitelbaum, P. Balenzuela, P. Cano, and J. M. Buld: "Community structures and role detection in music networks", *Chaos* Vol. 18, N. 043105, 2008.