

DO YOU SOUND LIKE YOUR FRIENDS? EXPLORING ARTIST SIMILARITY VIA ARTIST SOCIAL NETWORK RELATIONSHIPS AND AUDIO SIGNAL PROCESSING

BEN FIELDS, MICHAEL CASEY
Goldsmiths Digital Studios
Dept. of Computing
Goldsmiths, University of London

KURT JACOBSON, MARK SANDLER
Centre for Digital Music
Electronic Engineering
Queen Mary, University of London

ABSTRACT

A sample of the Myspace artist network is examined to investigate the relationship between social connectivity and audio-based similarity. Audio data from the Myspace artist pages is analyzed using well-established signal-based music information retrieval techniques. In addition to showing that the Myspace artist network exhibits many of the properties common to social networks, we show there is an ambiguous relationship between audio-based similarity and the social network topology.

1. INTRODUCTION

Myspace¹ has become the de-facto standard for web-based music artist promotion. Although exact figures are not made public, third party estimates suggests there are around 7 million artist pages² on Myspace.

Artists ranging from amateur to the most commercially successful publish Myspace pages. These Myspace artist pages typically include some media – usually streaming audio, video, or both – and a list of “friends” specifying social connections. This combination of media and a user-specified social network provides a unique data set that is unprecedented in both scope and scale. By examining a sample of the Myspace artist network using complex network theory and analyzing the corresponding media using audio-based music information retrieval techniques, we examine the relationship between artists’ social relationships and the audio-based similarity of their respective musical works.

We begin by reviewing relevant literature from complex network theory and signal-based music analysis in Section 2. The methods used to sample the Myspace artist network are discussed in Section 3. The analysis methods are discussed in Section 4. The implications and suggestions for future work are discussed in Section 5.

¹ <http://myspace.com>

² <http://scottelkin.com/archive/2007/05/11/MySpace-Statistics.aspx> reports as of April 2007 ~25 million songs, our estimates approximate 3.5 songs/artist, giving ~7 million artists

2. RELATED WORK

This work uses a combination of complex network theory and signal-based music analysis. Both disciplines apply intuitively to music information retrieval; however, to the best knowledge of the authors, the two have never been applied simultaneously to a single data set.

2.1. Complex Networks

Complex network theory deals with the structure of relationships in complex systems. Using the tools of graph theory and statistical mechanics, physicists have developed models and metrics for describing a diverse set of real-world networks – including social networks, academic citation networks, biological protein networks, and the World-Wide Web. All these networks exhibit several unifying characteristics such as small worldness, scale-free degree distributions, and community structure [5, 12]. Let us briefly discuss some definitions and concepts that will be used in this work.

2.1.1. Network Properties

A given network G is described by a set of *nodes* N connected by a set of *edges* E . Each edge is defined by the pair of nodes it connects (i, j) . If the edges imply directionality, $(i, j) \neq (j, i)$, the network is a *directed network*. Otherwise, it is an *undirected network*. The number of edges incident to the a node i is the *degree* k_i . In a directed network there will be an *indegree* k_i^{in} and an *outdegree* k_i^{out} corresponding to the number of edges pointing into the node and away from the node respectively.

Degree distribution: The *degree distribution* $P(k)$ is the proportion of nodes that have a degree k . The shape of the degree distribution is an important metric for classifying a network – *scale-free* networks have a power-law distribution [12] while *random* networks have a Poisson distribution. The scale-free degree distribution is a property common to many real-world networks. Conceptually, a scale-free distribution indicates the presence of a few very-popular *hubs* that tend to attract more links as the network evolves [5, 12].

Average shortest path: Two nodes i and j are connected if a path exists between them following the edges in the network. The path from i to j may not be unique. The *geodesic path* d_{ij} is the shortest path distance from i to j in number of edges traversed. For the entire network, the average shortest path or mean geodesic distance is l . In a *small-world* network the mean geodesic distance is small relative to the number of nodes in the network [5, 12]. The largest geodesic distance in a network is known as the *diameter*.

2.1.2. Networks and Music

Quite naturally, networks of musicians have been studied in the context of complex network theory – typically viewing the artists as nodes in the network and using either collaboration, influence, or similarity to define network edges. These networks of musicians exhibit many of the properties expected in social networks [4, 6, 14].

2.2. Signal-based Music Analysis

A variety of methods have been developed for signal-based music analysis, characterizing a music signal by its timbre, harmony, rhythm, or structure. One of the most widely used methods is the application of Mel-frequency cepstral coefficients (MFCC) to the modeling of timbre [10]. In combination with various statistical techniques, MFCCs have been successfully applied to music similarity and genre classification tasks [3, 11, 13]. A common approach for computing timbre-based similarity between two songs or collections of songs creates Gaussian mixtures models (GMM) describing the MFCCs and comparing the GMMs using a statistical distance measure. Often the earth mover's distance (EMD)[15], a technique first used in computer vision, is the distance measure used for this purpose. The EMD algorithm finds the minimum work required to transform one distribution into another.

2.3. Bringing It Together

Recently some work has been conducted exploring the interplay between user/artist generated metadata and content based similarity/retrieval. Most of this work [7, 16] focuses on various ways of exploiting the human generated metadata to filter content prior to, or instead of, conducting content-based analysis, similar to the techniques discussed in 2.2, in order to reduce computational load.

3. SAMPLING MYSPACE

The Myspace social network presents a variety of challenges. For one, the massive size prohibits analyzing the graph in its entirety, even when considering only the artist pages. Therefore we sample a small yet sufficiently large portion of the network. Also, the Myspace social network is filled with noisy data – plagued by spammers and orphaned accounts. We limit the scope of our sampling in a way that minimizes this noise. And finally, there currently

is no interface for easily collecting the network data from Myspace. Our data is collected using web crawling and HTML scraping techniques³.

3.1. Artist Pages

It is important to note we are only concerned with a subset of the Myspace social network – the Myspace *artist* network. Myspace artist pages are different from standard Myspace pages in that they include a distinct audio player application. We use the presence or absence of this player to determine whether or not a given page is an artist page.

A Myspace page will most often include a top friends list. This is a hyperlinked list of other Myspace accounts explicitly specified by the user. The top friends list is limited in length with a maximum length of 40 friends (the default length is 16 friends). In constructing our sampled artist network, we use the top friends list to create a set of directed edges between artists. Only top friends who also have artist pages are added to the sampled network; standard Myspace pages are ignored. We also ignore the remainder of the friends list (i.e. friends that are not specified by the user as top friends), assuming these relationships are not as relevant. This reduces the amount of noise in the sampled network but also artificially limits the outdegree of each node. Our sampling method is based on the assumption that artists specified as top friends have some meaningful musical connection for the user – whether through collaboration, stylistic similarity, friendship, or artistic influence.

The audio files associated with each artist page in the sampled network are also collected for feature extraction. Cached versions of the audio files are downloaded and audio features are extracted.

3.2. Snowball Sampling

There are several network sampling methods; however, for the Myspace artist network, snowball sampling is the most appropriate method [1, 9]. In this method, the sample begins with a seed node (artist page), then the seed node's neighbors (top friends), then the neighbors' neighbors, are added to the sample. This breadth-first sampling is continued until a particular sampling ratio is achieved. Here, we randomly select a seed artist⁴ and collect all artist nodes within 6 edges to collect 15,478 nodes. If the size of the Myspace artist network is around 7 million, then this is close to the 0.25% sampling ratio suggested for accurate degree distribution estimation in sampled networks. However, it is insufficient for estimating other topological metrics such as the clustering coefficient and assortativity [8]. Of course, a complete network topology is not our primary concern here.

³ Myspace scraping is done using tools from the MyPySpace project available at <http://mypyspace.sourceforge.net>

⁴ The artist is *Karna Zoo*, Myspace url: <http://www.myspace.com/index.cfm?fuseaction=user.viewProfile&friendID=134901208>

	n	m	$\langle k \rangle$	l	d_{max}
undirected	15478	91326	11.801	4.479	9
directed	15478	120487	15.569	6.426	16

Table 1. The network statistics for the Myspace artist network sample where n is the number of nodes, m is the number of edges, $\langle k \rangle$ is the average degree, l is the mean geodesic distance, and d_{max} is the diameter, as defined in Section 2.1.1.

With snowball sampling there is a tendency to oversample hubs because they have many links and are easily picked up early in the breadth-first sampling. This property would reduce the degree distribution exponent and produce a heavier tail but preserve the power-law nature of the network [9].

4. ANALYSIS

We begin by analyzing the structure of the Myspace artist network sample – showing that it conforms, in most respects, to the topology expected from such a social network. We then develop a simple metric for exploring the interaction between signal-based music similarity and the network structure.

4.1. Network Analysis

The Myspace artist network sample exhibits many of the network characteristics common to social networks and other real-world networks. Some of the network statistics are summarized in Table 1.

Although the network is constructed as a directed network, for our purposes we convert to an undirected network to simplify analysis. Each edge is considered bi-directional, that is $(i, j) = (j, i)$, and if a reflexive pair of edges existed in the directed graph, only one bi-directional edge exists in the undirected graph. An examination of the directed graph is reserved for later work.

The degree distribution for the undirected network is plotted in Figure 1 on a log-log scale. As mentioned earlier, it is common to find a power-law degree distribution in social networks [12]. However, exponential degree distributions have been reported previously in some types of music recommendation networks [4]. This is especially true for networks with imposed degree limits. For moderate degree values ($35 < k < 200$), our sample shows a power-law distribution. For lower degree values, the distribution is closer to exponential. This may be related to the fact that our network has an out degree limit imposed by Myspace restricting the maximum number of top friends ($k_{out} \leq 40$). The power-law fit also breaks down for high values of k – most likely due to the limited scope of our sample. Similar “broad-scale” degree distributions have been reported for citation networks and movie actor networks [2].

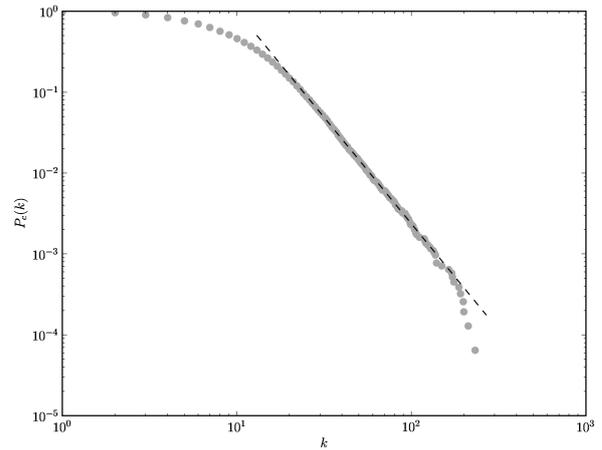


Figure 1. The cumulative degree distributions for the Myspace artist network sample. For moderate values of k , the distribution follows a power-law (indicated by the dotted line), but for low and high values the decay is exponential.

4.2. Signal-based analysis

MFCCs are extracted from each audio signal using a Hamming window on 8192 sample FFT windows with 4096 sample overlap. These FFT windows are gathered into 100ms non-overlapping frames. All MFCCs are created with the `fftExtract` tool⁵. For each artist node a GMM is built from the concatenation of MFCC frames for all songs found on each artist’s Myspace page (generally between 1 and 4 songs although some artists have more). An $n \times n$ matrix is populated with the earth mover’s distance λ_{ij} between the GMMs corresponding to each pair of nodes in the sample.

4.3. Relationship with signal-based measures

We explore a simple relation between audio signal dissimilarity and network structure using a box and whisker plot. The plot is shown in Figure 2. For all pairs of artists i and j , the EMD dissimilarity is found (λ_{ij}). These dissimilarities are grouped according to the geodesic distance in the undirected network between the artist nodes i and j , d_{ij} . There appears to be no clear correlation between these λ values and geodesic distance. The Pearson product-moment correlation coefficient confirms this giving a ρ of -0.0016 , with a p value of 1.50×10^{-20} . This should be viewed in the context of the number of pairwise relationships used, implying it is stable, at least for the community of artists found via this sample of the network.

5. DISCUSSION AND FUTURE WORK

Whatever slight trend may seem evident in Figure 2, it is clear from the Pearson ρ shown in Section 4.3 that no correlation exists in this set. Clearly then an attempt to use

⁵ source code at <http://omras2.doc.gold.ac.uk/software/fftextract/>

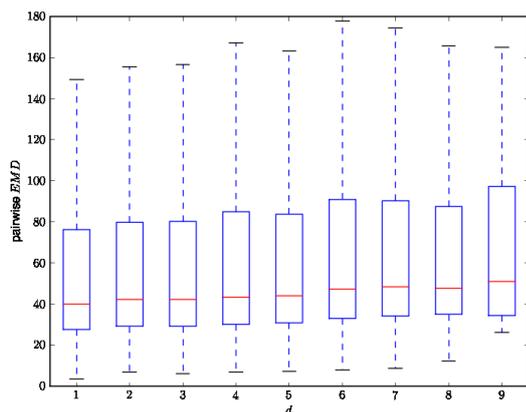


Figure 2. The box and whisker plot showing the spread of pair-wise artist dissimilarity grouped by geodesic distance as found on the artist graph.

geodesic distance to predict acoustic similarity would not success. While artists that are friends may sound similar, the assertion of friendship cannot be taken to imply an acoustic similarity.

The lack of a correlation seen between the artist network and the EMD dissimilarities is certainly enough to consider more exhaustive exploration of how content based retrieval systems can interoperate in the ever-growing space of social networks. The relationship seen in this sample is clear enough, but does it extend into the entirety of the Myspace network? More generally, does this trend occur uniformly across different communities of artists or are there significant differences between communities of artists as to the sonic similarity of friends?

6. ACKNOWLEDGEMENTS

This work is supported as a part of the OMRAS2 project, EPSRC numbers EP/E02274X/1 and EP/E017614/1.

7. REFERENCES

- [1] AHN, Y.-Y., HAN, S., KWAK, H., MOON, S., AND JEONG, H. Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th international conference on World Wide Web* (Alberta, Canada, May 2007), IW3C2.
- [2] AMARAL, L. A. N., SCALA, A., BARTHÉLÉMY, M., AND STANLEY, H. E. Classes of small-world networks. In *Proceeding of the National Academy of Sciences* (2000).
- [3] BERENZWEIG, A., LOGAN, B., ELLIS, D. P. W., AND WHITMAN, B. P. W. A large-scale evaluation of acoustic and subjective music-similarity measures. *Computer Music J.* 28, 2 (2004), 63–76.
- [4] CANO, P., CELMA, O., AND KOPPENBERGER, M. Topology of music recommendation networks. *Chaos* 16 (2006).
- [5] DA F. COSTA, L., RODRIGUES, F. A., TRAVIESO, G., AND BOAS, P. R. V. Characterization of complex networks: a survey of measurements. *Advanced Physics* 56 (2007), 167–242.
- [6] GLEISER, P., AND DANON, L. Community structure in jazz. *Advances in Complex Systems* 6, 4 (2003), 565–573.
- [7] KNEES, P., POHLE, T., SCHEDL, M., AND WIDMER, G. Combining audio-based similarity with web-based data to accelerate automatic music playlist generation. In *Proc. 8th ACM international workshop on Multimedia information retrieval* (2006), pp. 147 – 154.
- [8] KWAK, H., HAN, S., AHN, Y.-Y., MOON, S., AND JEONG, H. Impact of snowball sampling ratios on network characteristics estimation: A case study of cyworld. Tech. Rep. CS/TR-2006-262, KAIST, November 2006.
- [9] LEE, S. H., KIM, P.-J., AND JEONG, H. Statistical properties of sampled networks. *Physical Review E* 73 (January 2006), 102–109.
- [10] LOGAN, B. Mel frequency cepstral coefficients for music modeling. In *Proc. of Int. Symposium on Music Information Retrieval* (2000).
- [11] LOGAN, B., AND SALOMON, A. A music similarity function based on signal analysis. *Multimedia and Expo, 2001. ICME 2001. IEEE International Conference on* (22-25 Aug. 2001), 745–748.
- [12] NEWMAN, M. E. J. The structure and function of complex networks. *SIAM Review* 45 (2003), 167–256.
- [13] PAMPALK, E. *Computational Models of Music Similarity and their Application in Music Information Retrieval*. PhD thesis, Technischen Universität Wien, May 2006.
- [14] PARK, J., CELMA, O., KOPPENBERGER, M., CANO, P., AND BULDU, J. M. The social network of contemporary popular musicians. *Physics and Society* (2006).
- [15] RUBNER, Y., TOMASI, C., AND GUIBAS, L. J. The earth mover’s distance as a metric for image retrieval. *International Journal of Computer Vision* 40, 2 (2000), 99–121.
- [16] SLANEY, M., AND WHITE, W. Similarity based on rating data. In *Proc. of Int. Symposium on Music Information Retrieval* (2007).