

Social Playlists and Bottleneck Measurements:

Exploiting Musician Social Graphs Using Content-Based Dissimilarity and Pairwise Maximum Flow Values

Benjamin Fields¹, Kurt Jacobson², Christophe Rhodes¹ and Michael Casey²

(1) Goldsmiths Digital Studios, Goldsmiths, University of London

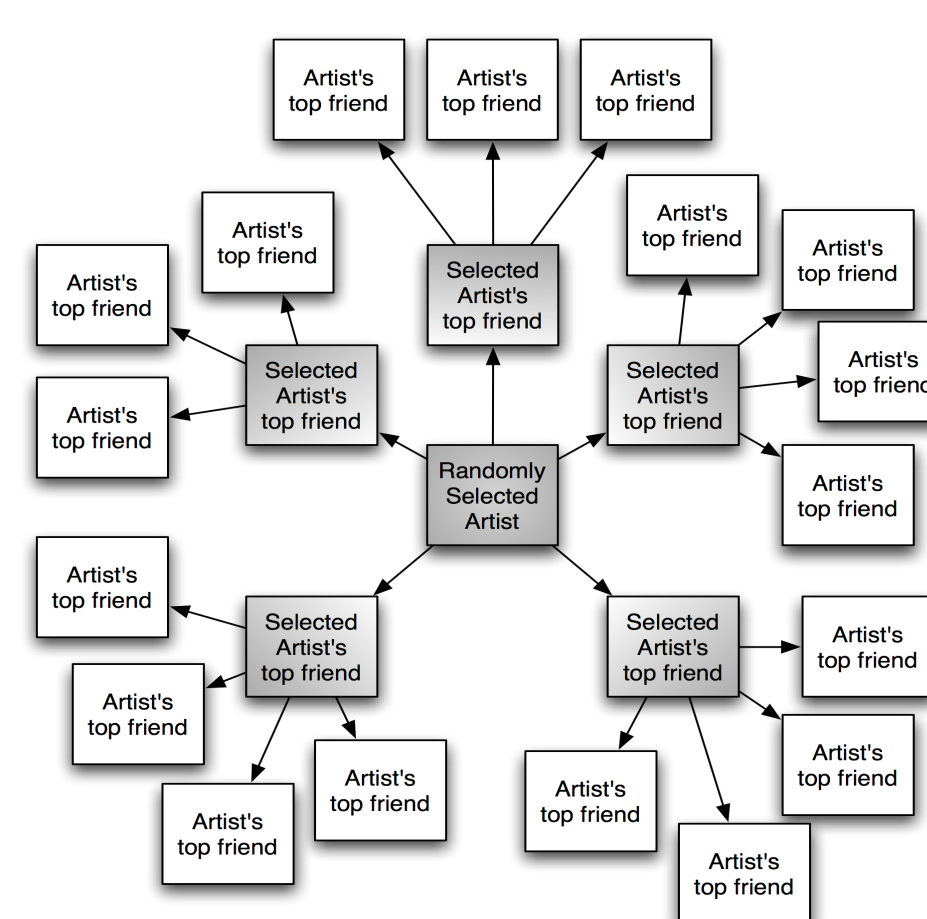
(2) Centre for Digital Music, Queen Mary, University of London

Abstract

We have sampled the artist social network of Myspace and to it applied the pairwise relational connectivity measure Minimum Cut/Maximum Flow. These values are then compared to a pairwise acoustic Earth Mover's Distance measure and the relationship is discussed. Further, a means of constructing playlists using the maximum flow value to exploit both the social and acoustic distances is realized.

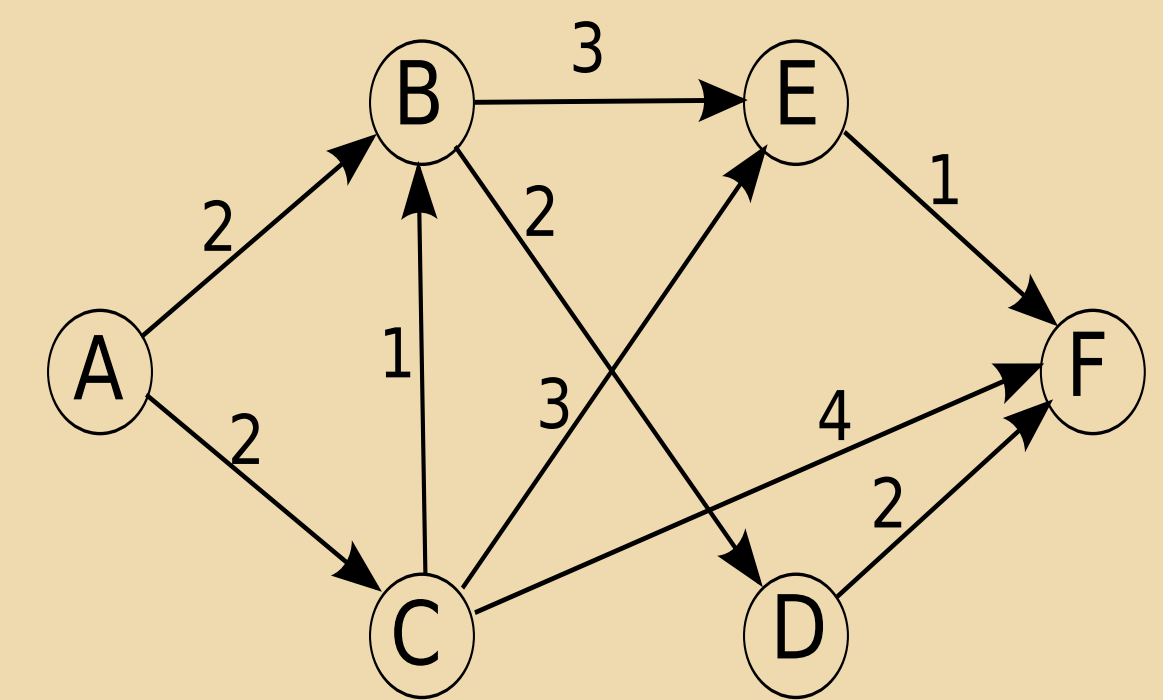
Sampling Myspace

- Only looking at the Artist Network
- A small portion of statistically significant size is used
- Snowball sampling (breadth first) starts at a random entry point and follows only top friends



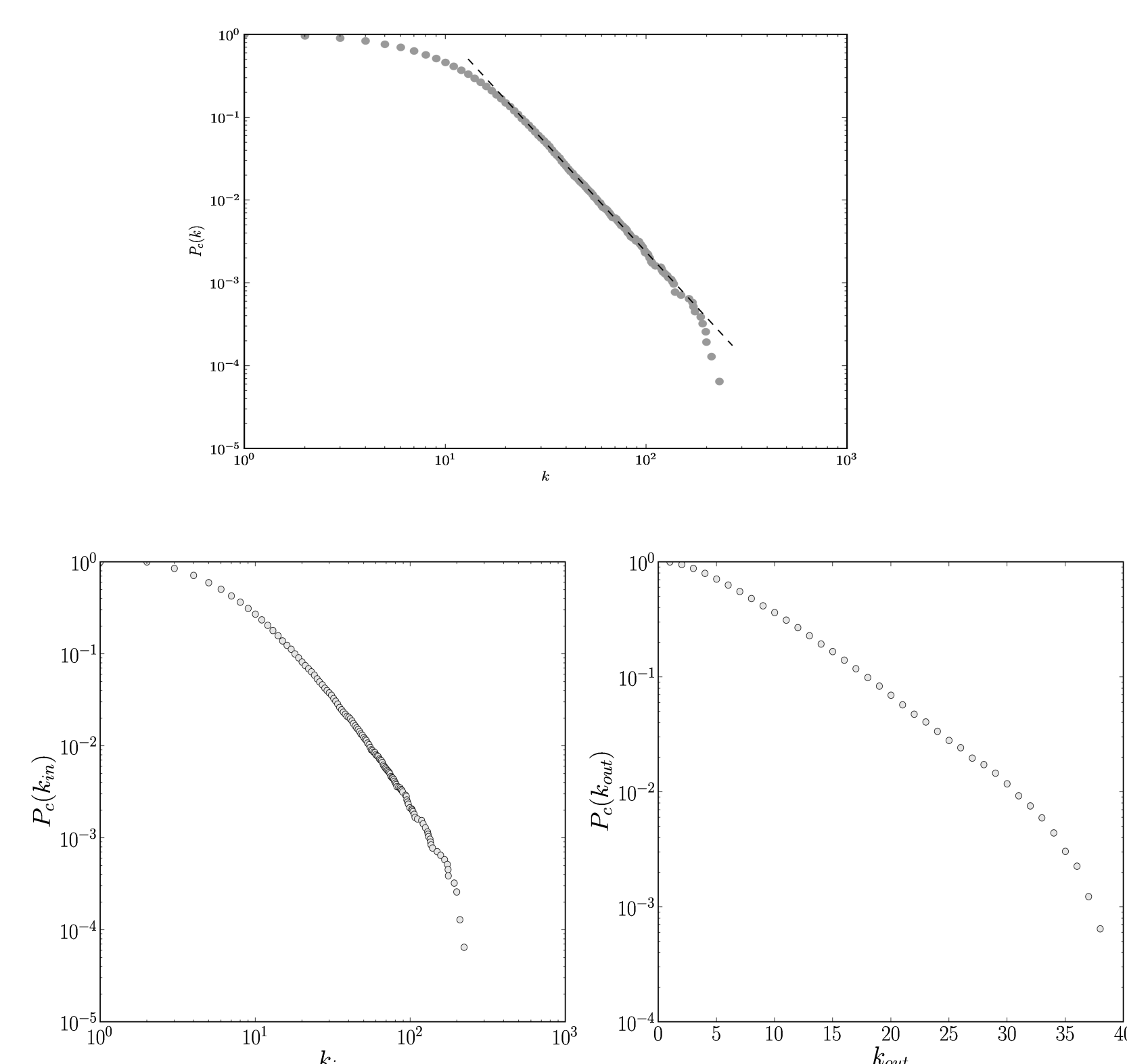
Maximum Flow/Minimum Cut

- nodes as a collection of sources and sinks for traffic or current

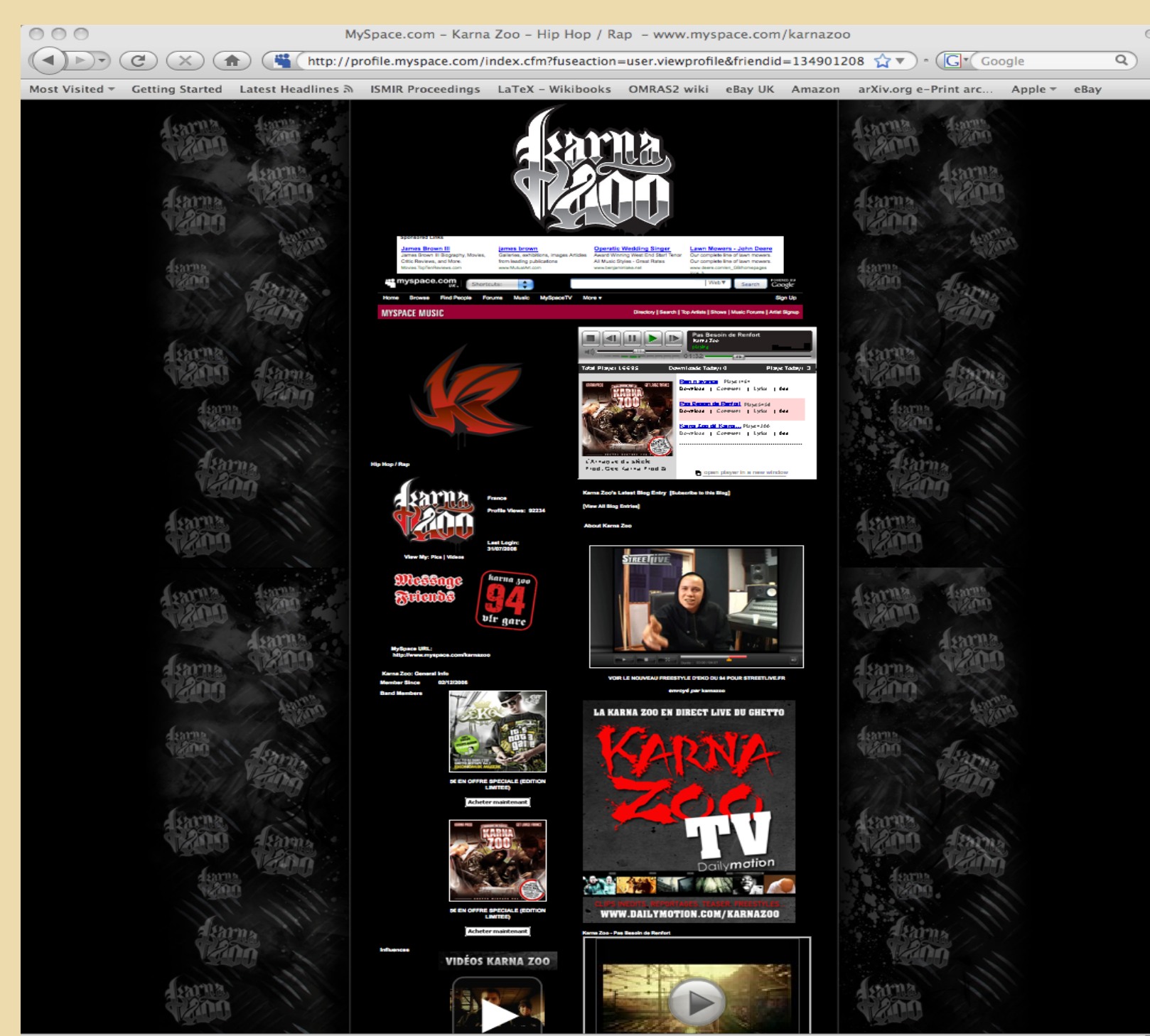


- weighted edges represent capacity
- looking for maximum capacity

Topology

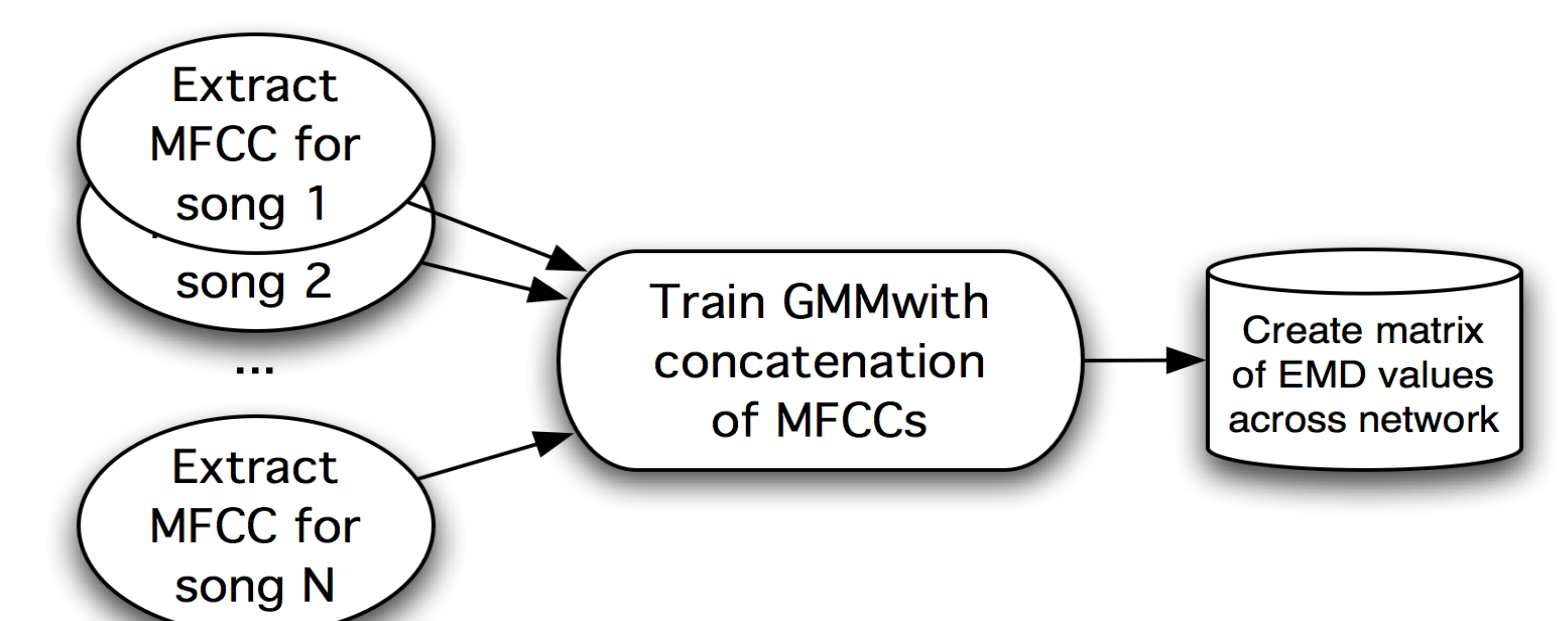


The Artist Page

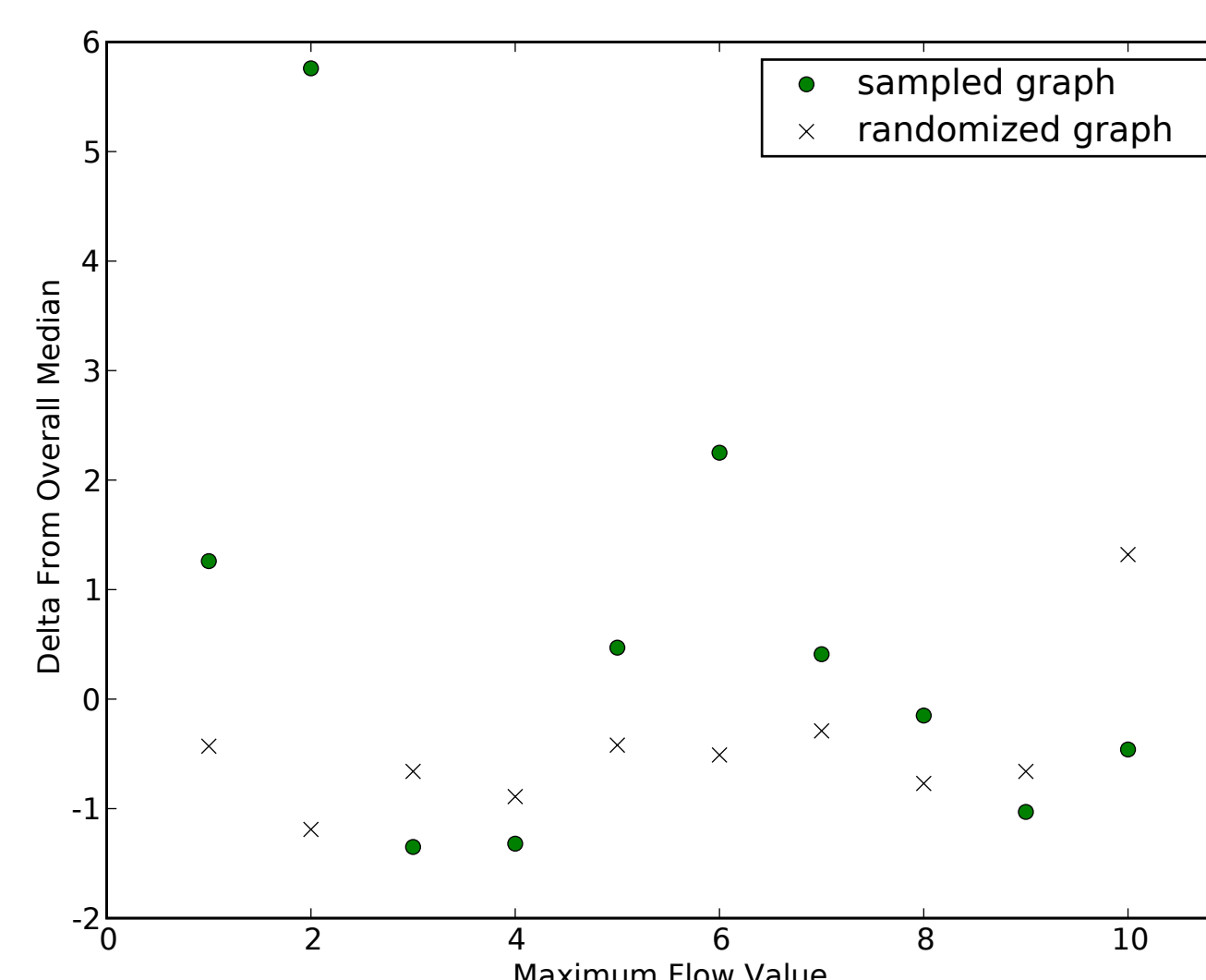
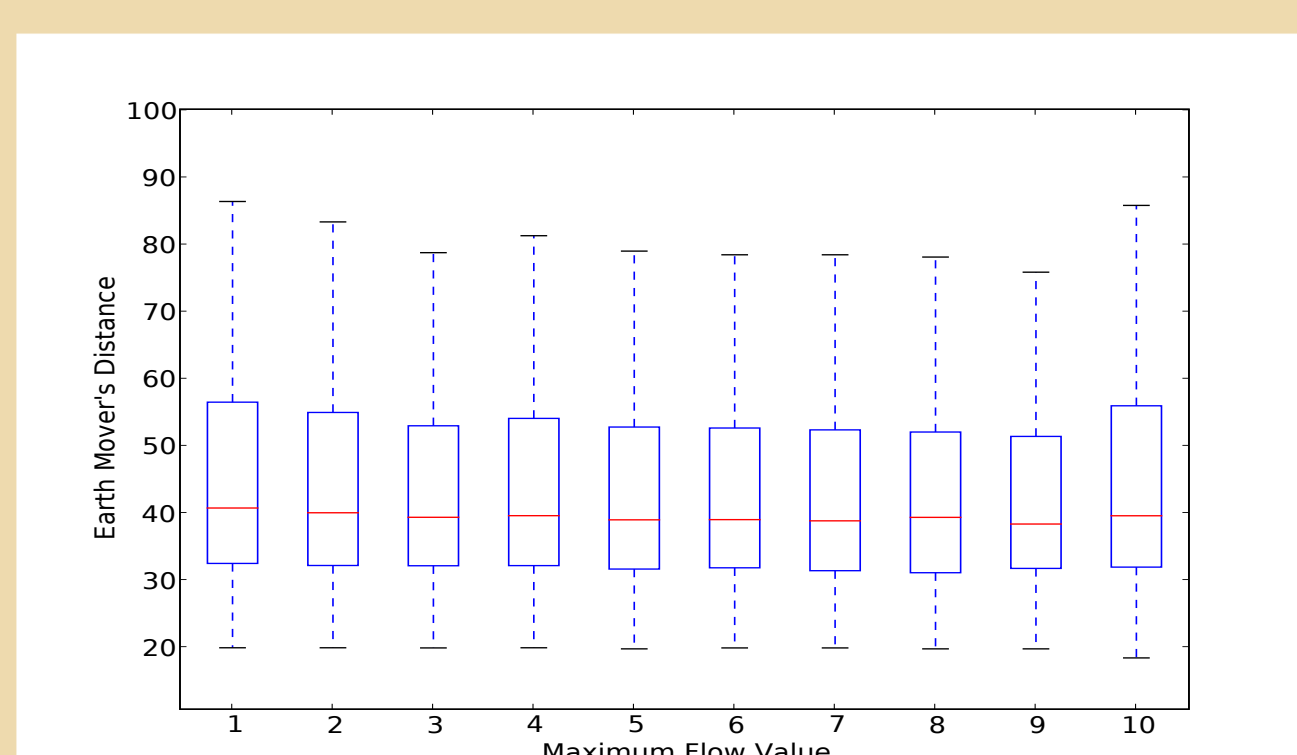
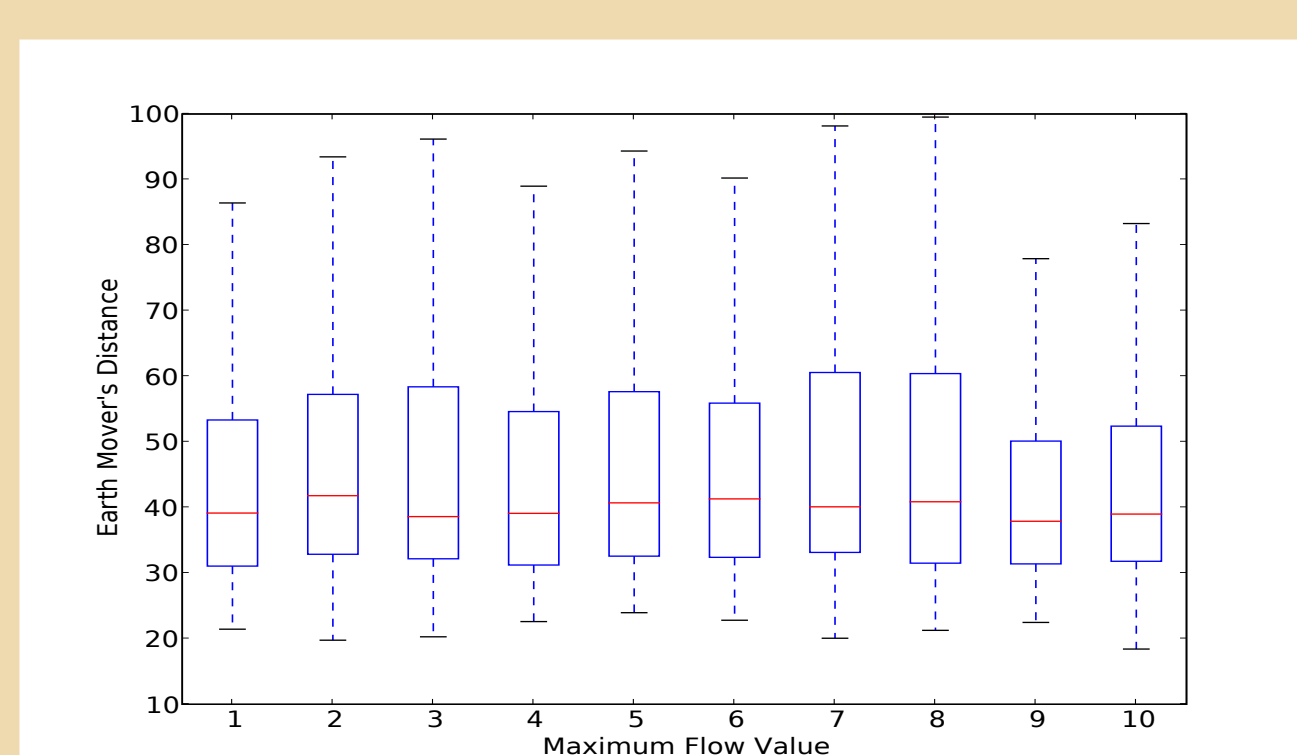


Measuring Acoustic Distance

- Extract sequence of 0.1s 20-bin MFCCs for each song
- Train GMM for artist based on concatenation of song MFCCs
- Create $n \times n$ matrix of Earth Mover's Distance λ_{ij} between the GMMs corresponding to each pair of nodes in the sample



Results



Max Flow	median	deviation	randomized	deviation
1	40.80	1.26	39.10	-0.43
2	45.30	5.76	38.34	-1.19
3	38.18	-1.35	38.97	-0.66
4	38.21	-1.32	38.64	-0.89
5	40.00	0.47	39.11	-0.42
6	41.77	2.25	39.02	-0.51
7	39.94	0.41	39.24	-0.29
8	39.38	-0.15	38.76	-0.77
9	38.50	-1.03	38.87	-0.66
10	39.07	-0.46	40.85	1.32

Table 1: Node pairs average EMD values grouped by actual minimum cut values and randomized minimum cut values, shown with deviations from the global median of 39.53.

	H-value	P-value
From sample	12.46	0.19
Random permutations	9.11	0.43

Table 1: The Kruskal-Wallis one-way ANOVA test results of EMD against maximum flow for both the sampled graph and its random permutations. The H-values are drawn from a chi-square distribution with 10 degrees of freedom.

Conclusions and Future Work

- Expected inverse relationship does not appear to exist between Max Flow and EMD
- Since these spaces are not correlated, can we use them together to make assertions?
- Maximum Flow playlist shows promise in preliminary prototypes

software available:

<http://mypspace.sourceforge.net>

<http://omras2.doc.gold.ac.uk/software/fftextact/>