
Goldsmiths
UNIVERSITY OF LONDON

Social Playlists and Bottleneck Measurements:

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

Ben Fields
b.fields@gold.ac.uk

Kurt Jacobson

Christophe Rhodes

Michael Casey

Goldsmiths
UNIVERSITY OF LONDON

Overview

- Motivation
- Background - Networks and Flow
- Dataset
- Experiments
- Going Forward

motivation

motivation

content-based glass ceiling

- affects most content based MIR tasks
- very difficult to get performance above 75 - 85% accuracy
- various theories attempt to explain this phenomenon
- is 100% accuracy really necessary?

motivation

novelty curves

- What makes an interesting playlist/recommender?
- Do you always want to listen to tracks that ‘sound similar’?
- Homogenous > Ideal > Random

motivation

social/cultural awareness

- Are all things that sound similar really related?
- What about cultural context?
- What about relationships?

background

Networks and Flow

background

Complex Networks

- structure of relationships in complex systems
- Graph Theory and Statistical Mechanics
- Properties of a complex network:
  - small world-ness
  - scale-free degree distributions
  - community structure

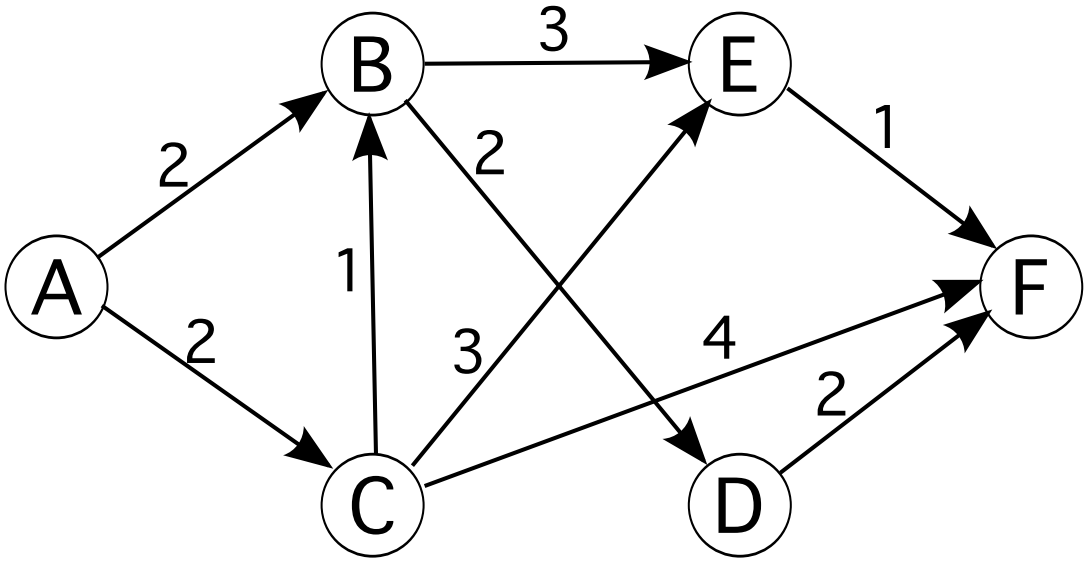
background

Max Flow/Min Cut

- nodes as a collection of sources and sinks for *traffic* or *current*
- weighted edges represent capacity
- looking for maximum capacity

background

Max Flow/Min Cut

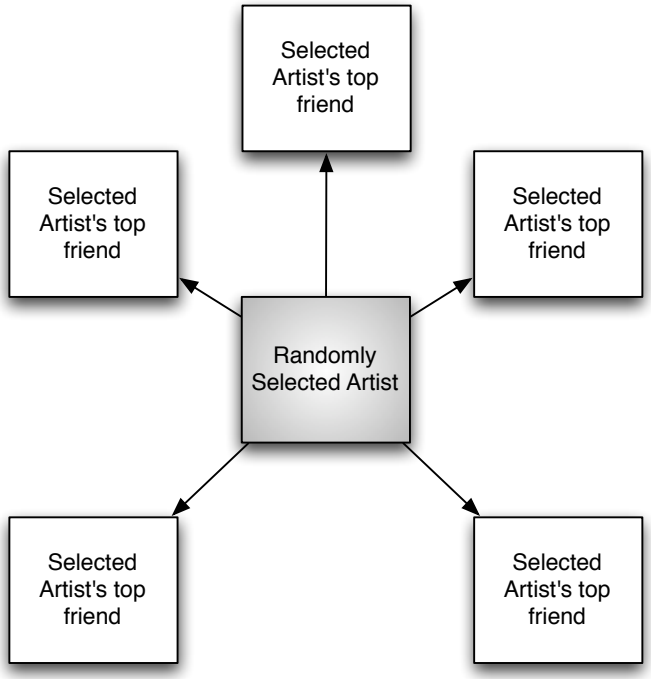


dataset

dataset

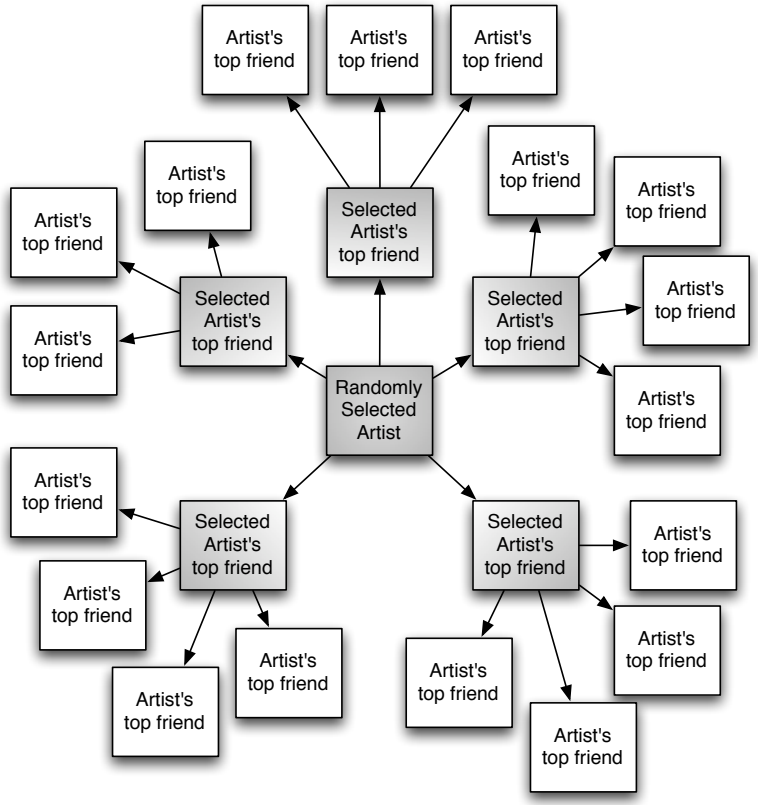
Sampling Myspace

dataset
Sampling Myspace



dataset

Sampling Myspace



Goldsmiths
UNIVERSITY OF LONDON

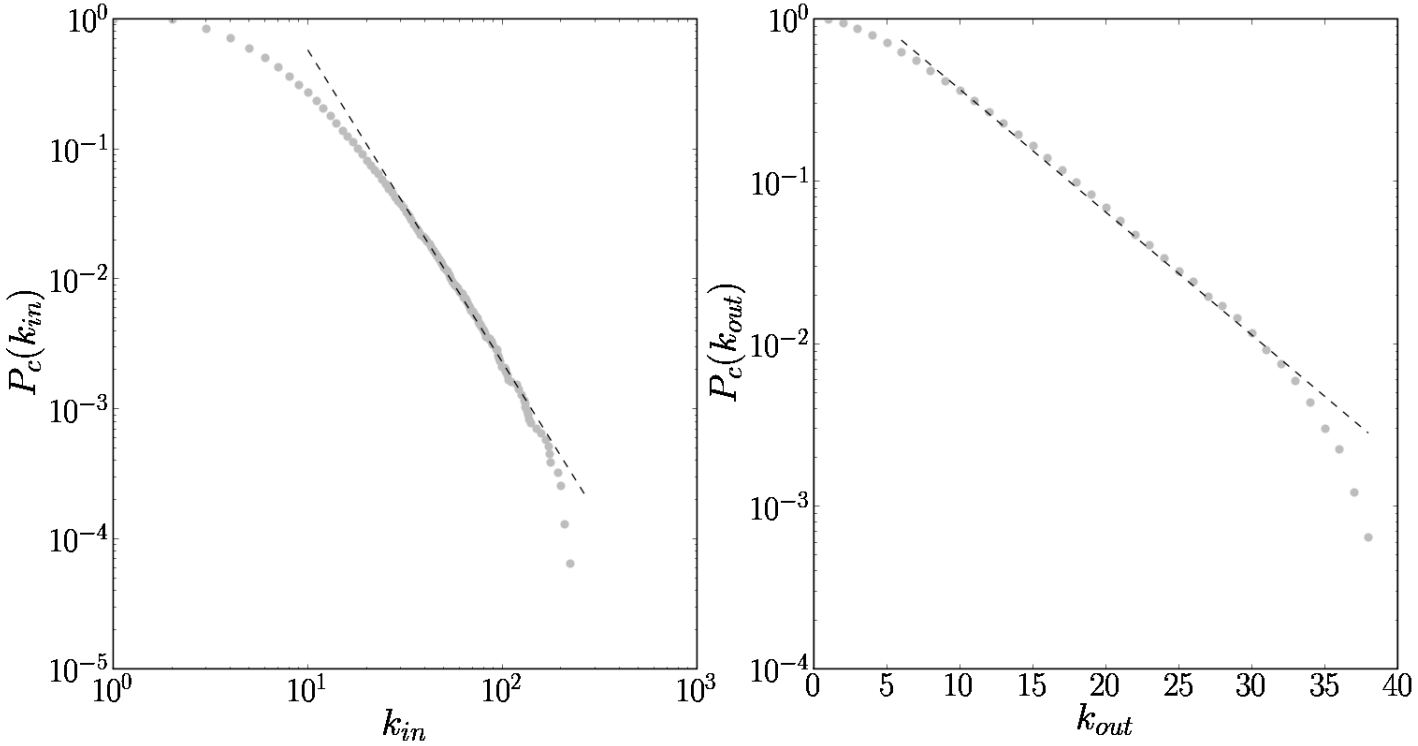
dataset

Sampling Myspace

- scale-free (mostly)
- 15,478 nodes (artist pages)
- 120,487 directed edges
- 91,326 undirected edges
- avg. degree
  - 15.5 as a directed graph
  - 11.8 when undirected

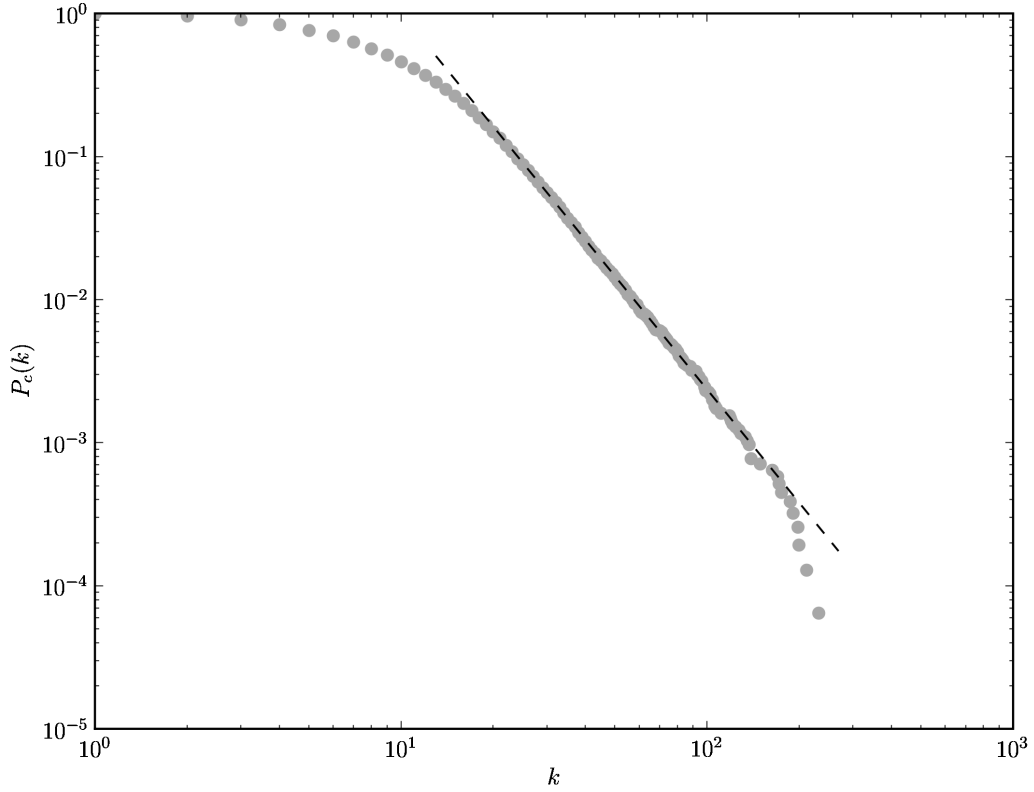
dataset

cumulative degree distribution



dataset

cumulative degree distribution



Goldsmiths
UNIVERSITY OF LONDON

experiment

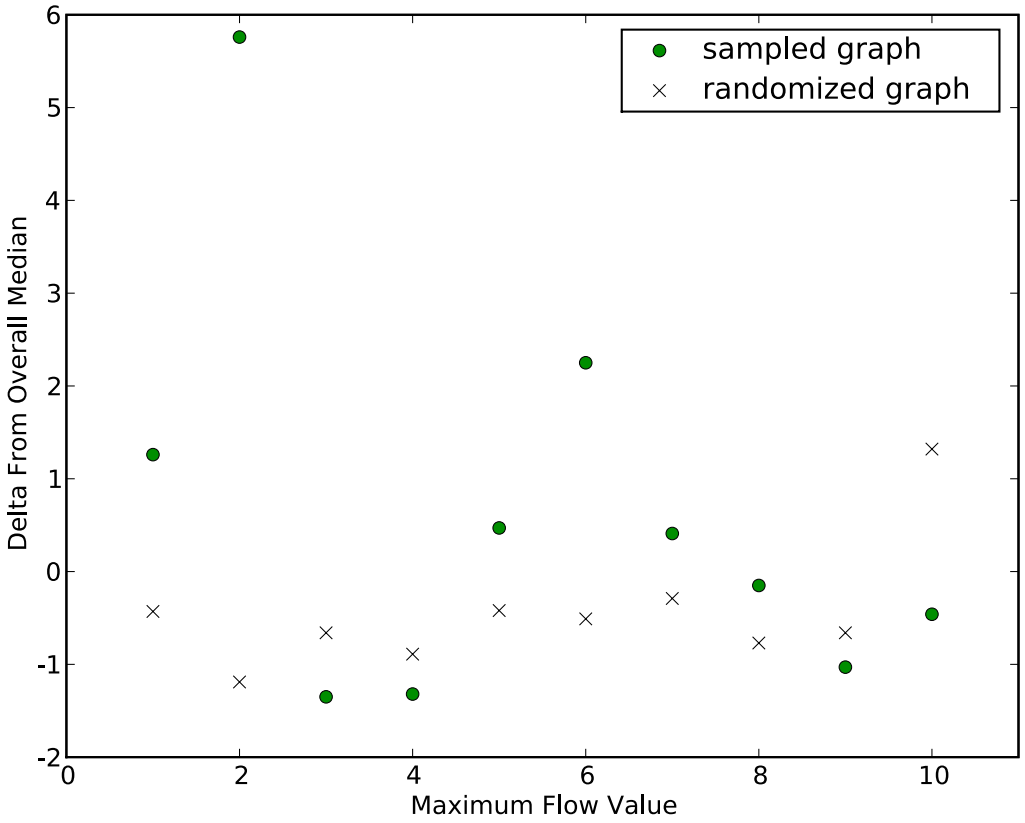
experiments

Max Flow v. EMD

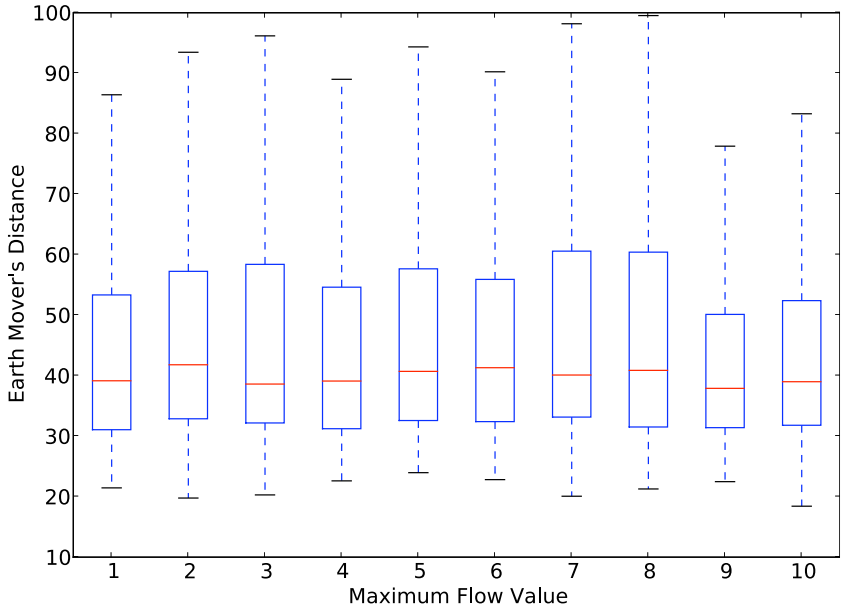
- Pairs of artist nodes grouped based on Maximum Flow
- A randomized network was created as well to compare the relationship
- Results point toward a mostly orthogonal relationship

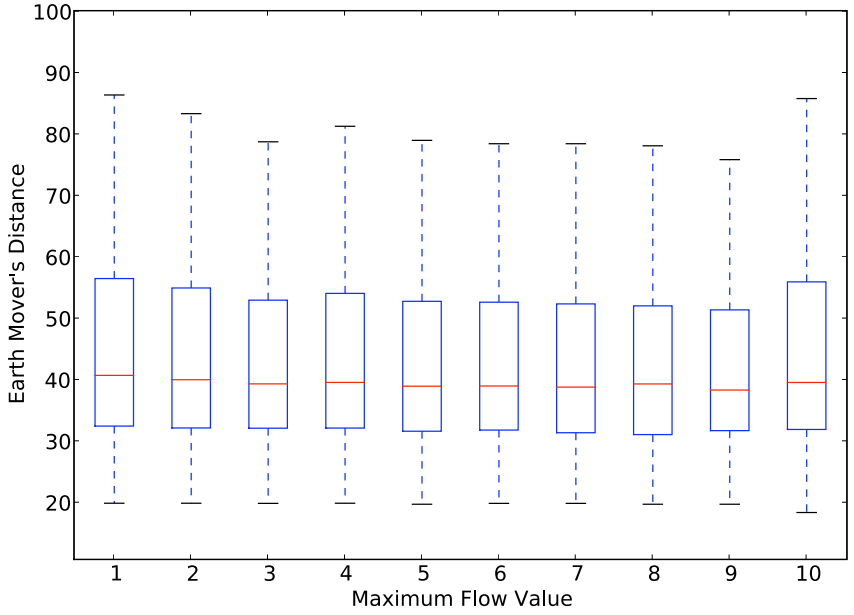
experiments

Max Flow v. EMD



experiments
Max Flow v. EMD





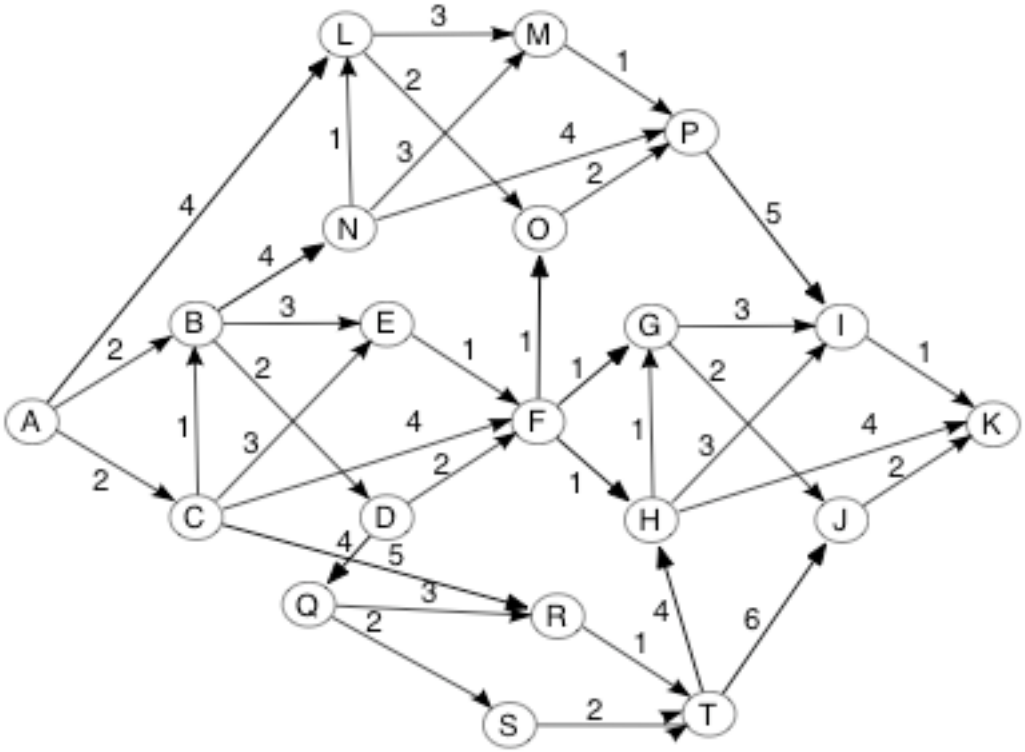
going forward

going forward

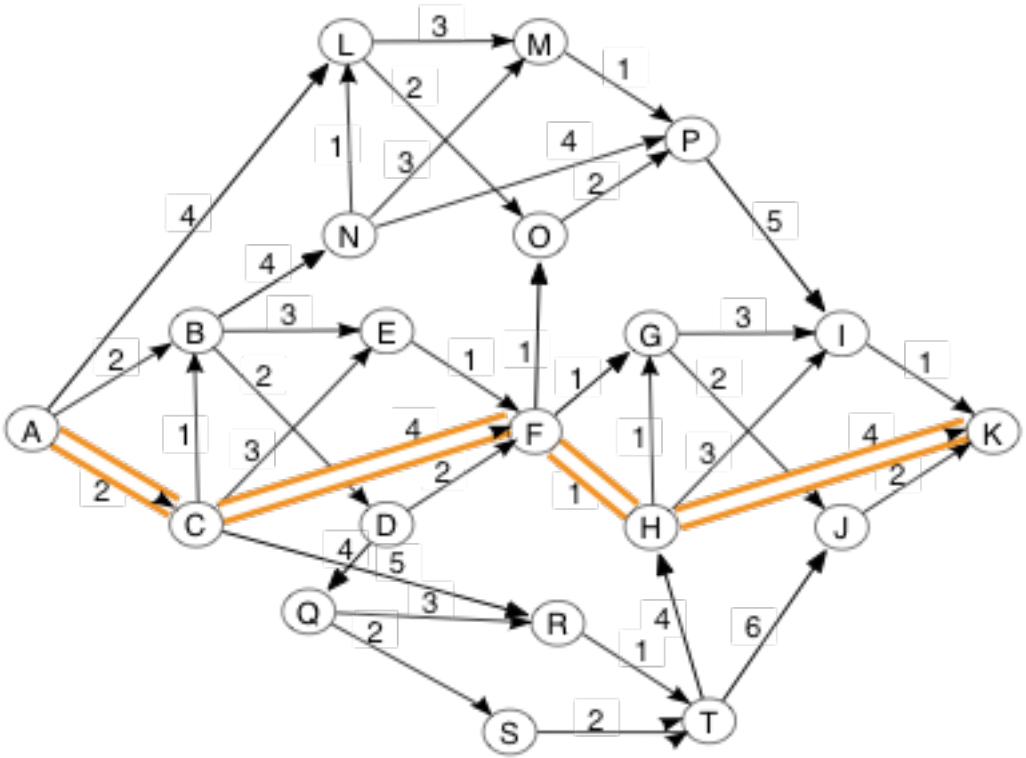
from analysis to synthesis

- How can the data seen in these experiment best be exploited?
- It seems they are mostly non-overlapping
- Recommenders and playlist generators should use both

going forward
playlist generator



going forward
playlist generator



going forward

Weighted Max Flow Playlists

- Max flow presents an interesting opportunity to create playlists using least resistant paths
- Preliminary testing shows promise
- Needs more exhaustive testing

resources

- http://mypyspace.sourceforge.net/
- http://dbtune.org/myspace/
- http://omras2.doc.gold.ac.uk/software/fftExtract/
- http://doc.gold.ac.uk/~map01bf/ismir2008/
- contact: b.fields@gold.ac.uk

resources

- http://mypyspace.sourceforge.net/
- http://dbtune.org/myspace/
- http://omras2.doc.gold.ac.uk/software/fftExtract/
- http://doc.gold.ac.uk/~map01bf/ismir2008/
- contact: b.fields@gold.ac.uk

Questions?


Goldsmiths
UNIVERSITY OF LONDON