
The Quest for Ground Truth in Musical Artist Similarity

Dan Ellis, Brian Whitman, Adam Berenzweig, Steve Lawrence
Columbia, MIT, NEC

Outline

1. Similarity and Artists
2. Sources of Information
3. Evaluation
4. Applications and Conclusions



1. Similarity and music

- Want similarity for recommendation etc.
- Multiple possible information sources
 - Do they agree? Which is best?
- Exercise: define single 'best-fit' pairwise artist distance matrix:

	abba	ace of base	aerosmith	a-ha
abba	0	2.5	4.6	1.2
ace of base	2.5	0	4.4	1.4
aerosmith	4.6	4.4	0	4.1
a-ha	1.2	1.4	4.1	0

- Training data for audio-based metrics?



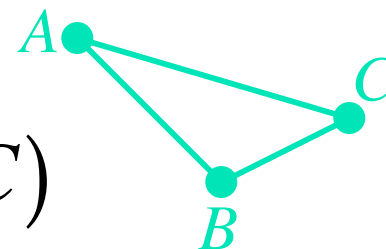
Problems

- Similarity judgment
 - ... is individual
 - ... is deep
 - ... depends on the context/criteria
 - ... can be asymmetric
- Many aspects of similarity
- But .. **try it and see what happens**



Distances

- Similarity (1..0) or distance (0..∞)?
 - How to interpret the numbers:
 $Sim = \Pr(\text{artist A called artist B}) ?$
 - Relate the two: $dist = (-\log[sim])^k$
- Symmetry? “A is like B” \Leftrightarrow “B is like A”
- Triangle inequality:
 $dist(A,C) \leq dist(A,B) + dist(B,C)$
- Distances \rightarrow Geometric embedding



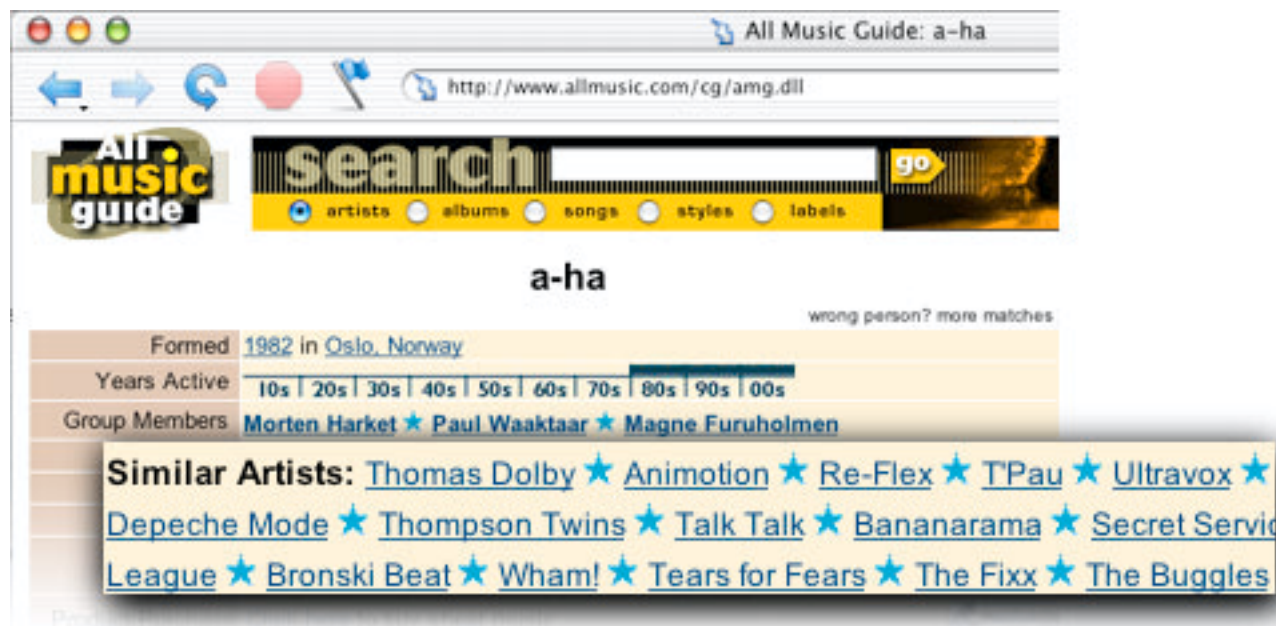
2. Data sources

- Many potential sources of subjective ‘ground-truth’ data:
 - Explicit descriptions (“*A* is similar to *B*, *C*..”)
 - Behavioral i.e. used in similar contexts
 - Described in similar terms
- Issues:
 - Converting to numerical values
 - Extending to cover every pair of artists



Explicit similarity statements

- *All Music Guide* similar artists lists



The screenshot shows a web browser window with the URL <http://www.allmusic.com/cg/amg.dll>. The page title is "All Music Guide: a-ha". The search bar contains "a-ha" and the "go" button is highlighted. Below the search bar, the artist's name "a-ha" is displayed. The page includes a table with the following information:

Formed	1982 in Oslo, Norway
Years Active	10s 20s 30s 40s 50s 60s 70s 80s 90s 00s
Group Members	Morten Harket ★ Paul Waaktaar ★ Magne Furuholmen

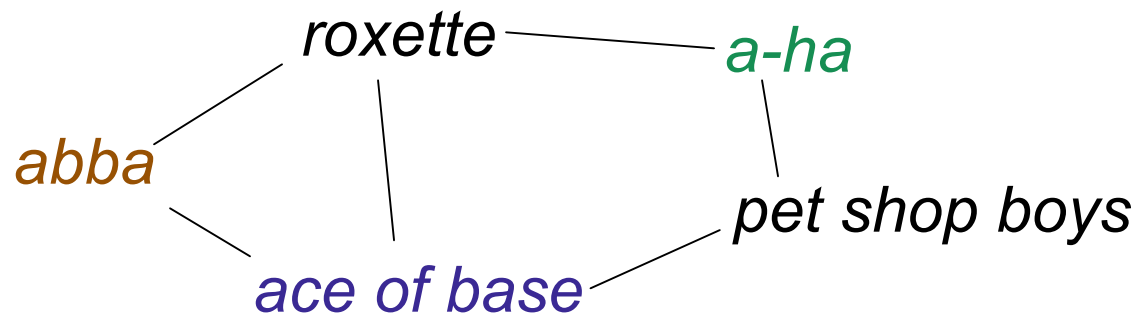
Below the table, a "Similar Artists" section lists the following artists: Thomas Dolby ★ Animation ★ Re-Flex ★ TPau ★ Ultravox ★ Depeche Mode ★ Thompson Twins ★ Talk Talk ★ Bananarama ★ Secret Service League ★ Bronski Beat ★ Wham! ★ Tears for Fears ★ The Fixx ★ The Buggles.

- Binary similarity (in/not in list)



Erdős Distance

- Distance from A to B is minimum # hops



	abba	ace of base	aerosmith	a-ha
abba	0	1	5	2
ace of base	1	0	4	2
aerosmith	5	4	0	4
a-ha	2	2	4	0

- also 'Resistive' Erdős



Collection co-occurrence

- Assumption:
User U possess music by A and B
 $\Rightarrow A$ and B are similar
 - problematic for individuals, averages out
- *OpenNap* data:
Use Napster protocol to record the collections of ~ 3000 users
 - 400,000 identified tracks
- $P(A|B) \times$ popularity mismatch discount



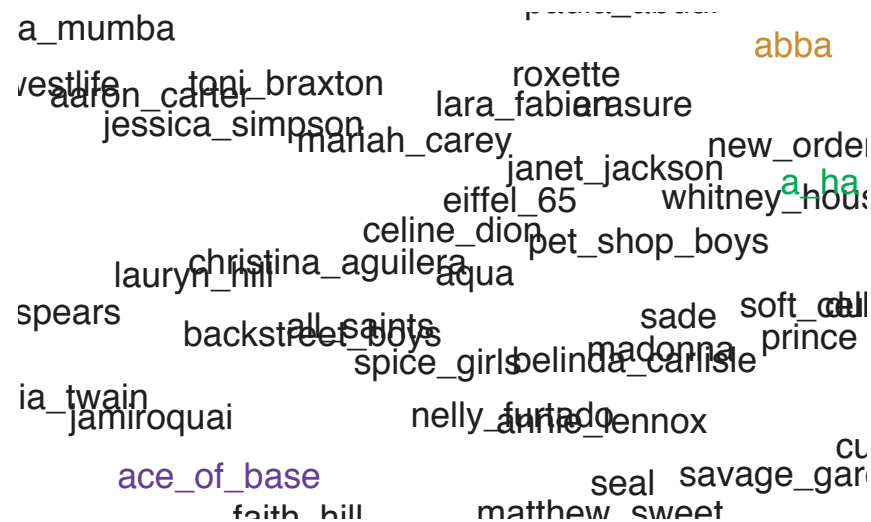
Similarity of associated text

- Find web pages mentioning particular artists (“google abba music”)
- *tf•idf* weighted terms for each artist
- Similarity of artists
= weightings of common terms
- *Klepmit* data



Multidimensional Scaling

- Map distances to spatial arrangement minimizing MSE of Euclidean distances
- Regularizes distance matrices:



- Issues: # dimensions, distance warping



3. Evaluation

- How to compare distance metrics?
- Need *independent* test data
- Collect new evaluation data set
 - Sparse sampling of artist judgments to validate the complete set



Information elicitation

- ~~“What is similarity of A and B?”~~
 - no consistency, arbitrary scale
- **Target + Choices** paradigm:

Which artist is most similar to:
Janet Jackson?

1. [R. Kelly](#)
2. [Paula Abdul](#)
3. [Aaliyah](#)
4. [Milli Vanilli](#)
5. [En Vogue](#)
6. [Kansas](#)
7. [Garbage](#)
8. [Pink](#)
9. [Christina Aguilera](#)

- **Chosen** is *more like target* than (known) others



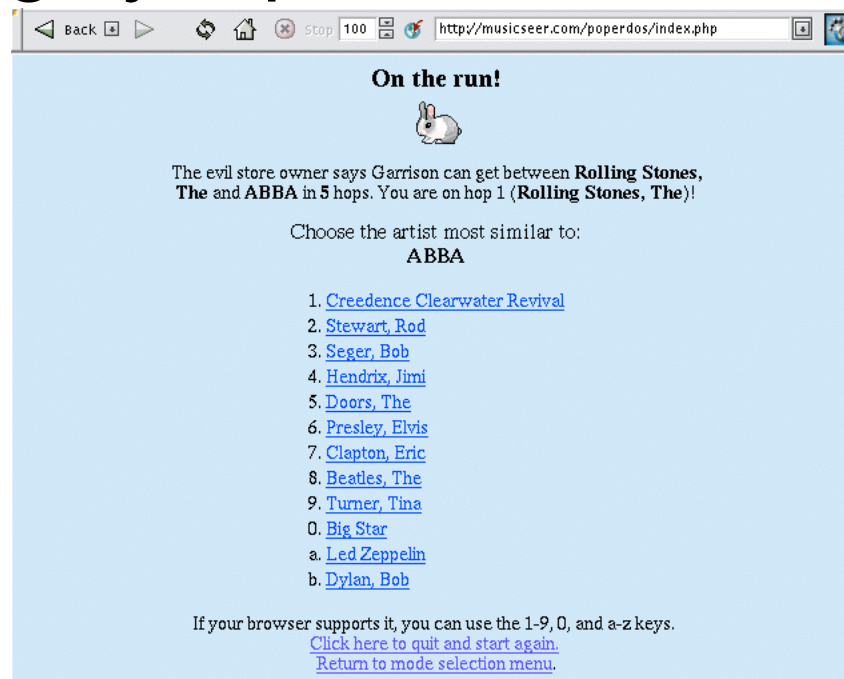
musicseer.com

- Web site to collect artist-similarity judgments from anonymous volunteers
- Two procedures:
 - poperdos game:
Get from artist *A* to *B* in fewest # steps
 - Direct survey:
Choose most similar artist from a list



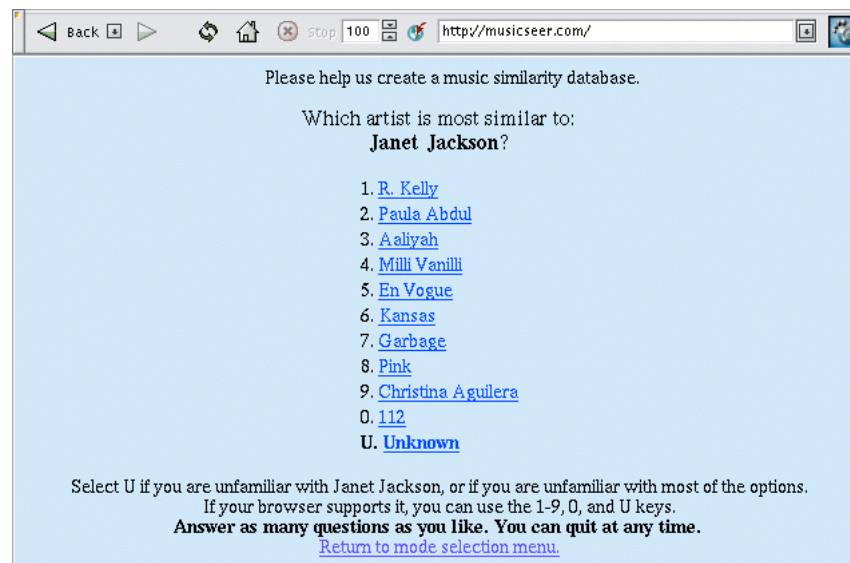
poperdos game

- Competitive motivation of informants?
- Choices from AllMusic lists + OpenNap
→ roughly equidistant choices



Survey

- Sequence of single judgments
 - More uniform sampling of targets
 - Infer which bands are known to informant
 - Dummy entries ‘catch’ random guessing



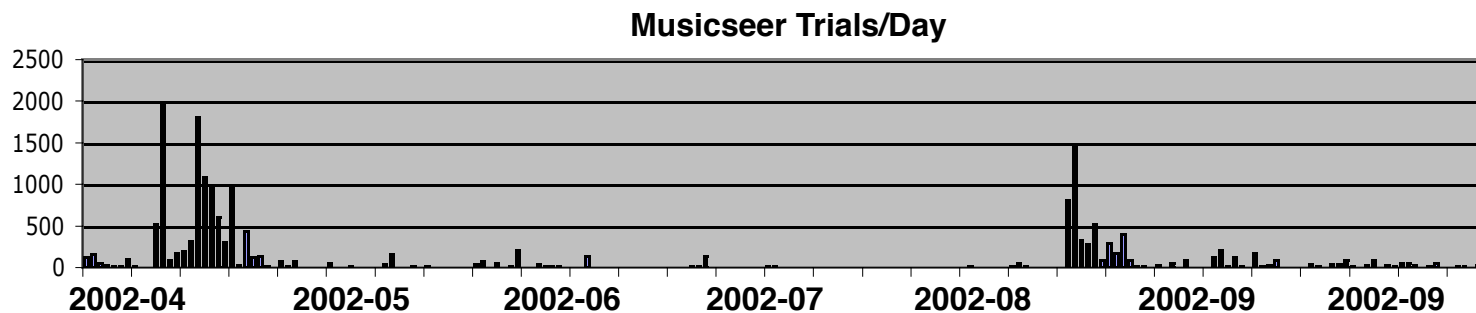
“10,000 random band names”

- Generative model of band names
 - trained on real band name ‘grammars’
+ vocabulary of music-related web pages
- Most popular decoy choices:
 - *Neither Palindrome Taker*
 - *Archbishop Riot*
 - *Skylight Cuisine*
 - *Blond and Bipolar*



musicseer data

- 1,782 users
 - best user: 748 judgments, ten above 200
- 22,470 judgments
 - ~140k triplets <target chosen unchosen>
 - Spike after “10,000 names” blogging



- <http://musicseer.com/results/>



Scoring the similarity metrics

- Avg. ranking of user choice by metric
- Agreement percentages for...
 - judgments
 - triplets

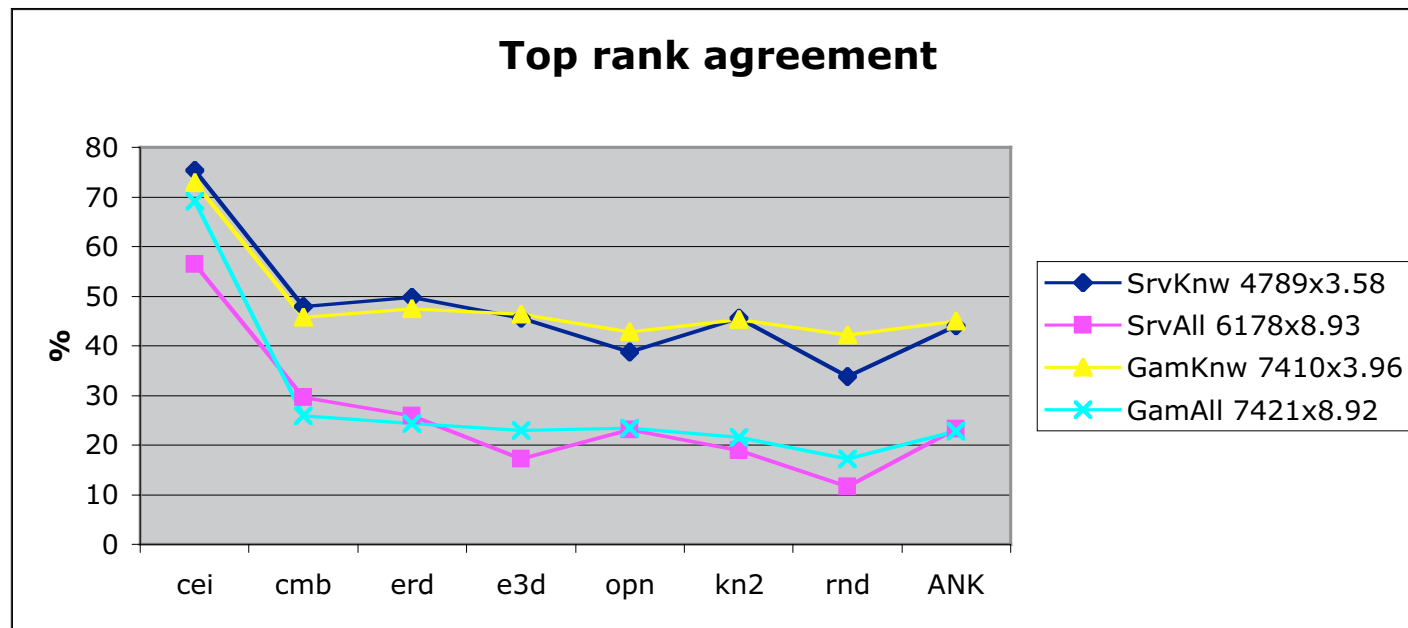
<i>Target:</i>	Police
<i>Choices:</i> <i>(ranked by metric)</i>	U2 Pretenders Sting Queen INXS Phil Collins

Informant choice ranked **#2** by metric
⇒ **no agreement** for this *judgment*
... but 4 of 5 *triplets* do agree



Results

- Judgment-level agreement %:
i.e. user choice ranks #1 under metric



- “Ceiling” reflects inconsistency limit



4. Applications

- Original motivation:
 - Comparing ground-truth data sources
 - Targets for training music similarity classifier
 - What happens when you try?
- Actual outcome
 - `musicseer` data
 - Evaluation metrics



Using evaluation data

- Significant information in `musicseer` evaluation data
 - basis for a metric in itself?
- Use these evaluation procedures for acoustic-based similarity measures
- ‘Inconsistency’ may point to *aspects...*



Playola

- Acoustic similarity browser

The screenshot shows the Playola website interface. At the top, there is a search bar with the text 'Search:' and a dropdown menu for 'Artist'. Below the search bar, there are several utility links: [About], [Help], [Turn Samples Off], [Turn Debug On], [Turn Popups Off], and [Logout dpwe].

The main navigation area includes 'Get Playola Selections: 20 songs', a dropdown for 'you recently heard', and a 'Go!' button. There are also links for 'Browse: Artists Albums Playlists' and a 'Range: 0-C' dropdown.

The current artist is 'The Woodbury Muffin Outbreak', with a link to their 'band web page' and a '[Play!]' button. The playlist is currently '-New Playlist-' with '[Add to]' and '[View]' options.

The main content area is divided into two columns. The left column is a table of songs:

	Song Title	Artist	Time	Rating
<input type="checkbox"/>	The Ballad of Tabitha	The Woodbury Muffin Outbreak	4:00	
<input type="checkbox"/>	Monkey Dreams	The Woodbury Muffin Outbreak	2:57	
<input type="checkbox"/>	A Cold Dark Night (Live)	The Woodbury Muffin Outbreak	3:13	
<input type="checkbox"/>	Leo, The Ballad of	The Woodbury Muffin Outbreak	1:48	
<input type="checkbox"/>	Baby I Forgot To Tell You	The Woodbury Muffin Outbreak	4:04	

The right column is the 'Music-Space Browser' with a '[What's This?]' link. It features a 'Feature' bar and a 'Less More' slider. The genres listed are: AltNGrunge, CollegeRock, Country, DanceRock, Electronica, MetalNPunk, NewWave, Rap, RnBSoul, SingerSongwriter, SoftRock, TradRock, Female, and HiFi.

Below the Music-Space Browser is the 'Similar Songs' section with a '[Play this list]' link and a '[What's This?]' link. It contains a table of similar songs:

	Song Title	Artist	Distance	Good Match?
<input type="checkbox"/>	Baby I Forgot To Tell You	The Woodbury Muffin Outbreak	0.00	
<input type="checkbox"/>	Number five	Bizi Chyld	0.07	
<input type="checkbox"/>	Waiting for Your Love	Toto	0.08	
<input type="checkbox"/>	Excerpt from 'CD'	Weirdomusic	0.08	



Conclusions

- Musical artist similarity - a tricky concept
 - but *how* tricky?
 - best metric agreed with 50% of judgments
- Evaluation procedure + data has broader applications
 - Most valuable outcome?
- Web-based user data collection is viable

