Five Approaches to Collecting Tags for Music

**Approach**

**Survey**
- Experts are paid to annotate songs using a standard form

**Social Tags**
- Large community contributes tags using a social network
- Collective wisdom of crowds
- Unlimited vocabulary
- Provides social context

**Game Tags**
- Players produce tags as they play a video game
- Collective wisdom of the crowds
- Fast paced for rapid data collection

**Webtags**
- Analyze a corpus of music reviews, artist bios, blogs, discussion boards
- Large corpus of publicly-available documents
- Provides social context

**Autotags**
- Annotate audio content using signal processing and machine learning
- Not affected by cold-start problem
- No direct human involvement

**Hybrid**
- Combination of approaches

**Weaknesses**
- All Songs
  - Time consuming approach lacks scalability
  - Human-labor intensive
  - Small, predetermined vocabulary
- Weaknesses
  - Sparce/missing data in long tail
  - Produces weak labeling
  - Ad-hoc annotation behavior

**Strengths**
- All Songs
  - Strong labeling
  - High-quality annotations
  - Custom-tailored vocabulary

**Example**
- Songs using a standardized form
- Experts are paid to annotate
- The vocabulary consists of 109 tags that
- natunes.
- There are 87 “long tail” songs from  
  Maim of 3 individuals.
- This data serves as the ground truth.

**Algorithm**

**CAL 500 Data Set**
- Paid 55 undergraduates to annotate 50 songs by 50 artists using a vocabulary of 100 tags
- Each song was annotated by a minimum of 3 individuals.

This data serves as the ground truth. There are 87 “long tail” songs from Mag-

**ListenGame**
- During a two week pilot study of Lis-
  tenGame, we collected 5,773 tags for 260
  of the CAL500 songs from 440 players.
- Each of the 27,250 song-tag pairs were
  presented 1.8 times on average.

**Ad-hoc annotation behavior**
- Produces weak labeling
- Sparse/missing data in long tail

**Rank-Based Interleaving**
- Given a tag, rank songs based on their best rank according to other approaches

**Supervised Multilabel Model**
- MFCC-Delta Feature Space
- One GMM per tag
- Mixture Hierarchies EM to train GMMs
- Produces “Semantic Multinomial” distribution over tag vocabulary for each novel song
- Top performing system in 2008 MIREX Audio Tag Classification Task

**Other hybrid approaches**
- Other hybrid approaches
  1. Kernel Combination
  2. RankBoost
  3. Calibrated Score Averaging

**RankBoost**
- Given a tag, rank songs based on their best rank according to other approaches

**Calibrated Score Averaging**
- Given a tag, rank songs based on their best rank according to other approaches

**Hybrid**
- Combination of approaches
  - Social feedback and audio content
  - Use strengths, remove weaknesses
  - Multi-threaded approach to cold-start problem

**Combining high quality tagging data**
- Based solely on audio content, no context

**Data, Papers, and additional information can be found at:**
http://cosmal.ucsd.edu/cal/