Five Approaches to Collecting Tags for Music

Douglas Turnbull, Luke Barrington, Gert Lanckriet
Computer Audition Laboratory, UC San Diego

Approach	Survey	Social Tags	Game Tags	Webtags	Autotags	Hybrid
Summary	Experts are paid to annotate songs using a stardard form	Large community contributes tags using a social network	Players produce tags as they play a video game	Analyze a corpus of music reviews, artist bios, blogs, discussion boards	Annotate audio content using signal processing and machine learning	Combination of approaches
Strengths	Custom-tailored vocabulary High-quality annotations Strong labeling	Collective wisdom of crowds Unlimited vocabulary Provides social context	Collective wisdom of the crowds Fast paced for rapid data collection Entertaining incentives produce high-quality tags	Large corpus of publicly-available documents No direct human involvement Provides social context	Not affected by cold-start problem No direct human involvement Produces strong labeling	Combine social context and audio content Use strengths, remove weaknesses Multi-tiered approach to cold-start problem
Weaknesses	Small, predetermined vocabulary Human-labor intensive Time consuming approach lacks scalability	Ad-hoc annotation behavior Produces weak labeling Sparse/missing data in long tail	“Gaming” the system Difficult to create viral gaming experience Based on short clips, rather than songs	Text-mining introduces noise Produces weak labeling Sparce/missing data in long tail	Computationally intensive Limited by training data Based solely on audio content, no context	Increased system complexity Combining data sources can be tricky
Example	PANDORA created by the Music Genome Project	last.fm the social music revolution	listen game	Google	cal Semantic Music Discovery Engine	echotron
Algorithm	**CAL 500 Data Set** [1] Paid 55 undergraduates to annotate 500 songs by 500 artists using a vocabulary of tags. Each song was annoated my a minimum of 3 individuals. This data serves as the ground truth. There are 87 “long tail” songs from Magnatunes. The vocabulary consists of 109 tags that relate to genre, instrumentation, emotion, usage, rhythm, vocals, and other musical characteristics.	**Audioscrobbler** Attempted to collect a list of tags associated with each CAL500 song and each CAL500 artist from Last.fm’s Audioscrobbler website. For each song, the song and artist lists were combined. The combined list was matched to the CAL500 vocabulary. We attempted to use synonyms, alternative spellings, and wildcard matching to improve coverage.	**ListenGame** [2] During a two week pilot study of ListenGame, we collected 16,500 tags for 250 of the CAL500 songs from 440 players. Each of the 27,250 song-tag pairs were present 1.8 times on average.	**Relevance Scoring** [3] 1) Collect Corpus query google with song, artist and album 2) Query Corpus with Tag find most relevant song given tag **Site-Specific** use web documents from sites that are known to have high-quality content (Rolling Stones, AMG AllMusic, etc **Weighted RS** weight pages by tag relevance	**Supervised Multilabel Model** [1] MFCC+Delta Feature Space One GMM per tag Mixture Hiearchies 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	**Rank-Based Interleaving** Given a tag, rank songs based on their best rank according to other approaches **Other hybrid approachs** 1) Kernel Combination [4] 2) RankBoost [5] 3) Calibarated Score Averaging [5]
All Songs: Density	1.00	0.23	0.37	0.67	1.00	1.00
All Songs: AUC-ROC	1.00	0.62	0.65	0.66	0.69	0.74
All Songs: Top 10 Prec	0.97	0.37	0.32	0.32	0.33	0.38
Long Tail: Density	1.00	0.03	Based on our experimental setup, the long tail results are misleading because the selection of songs in ListenGame is independent of popularity.	0.25	1.00	1.00
Long Tail: AUC-ROC	1.00	0.54		0.56	0.70	0.71
Long Tail: Top 10 Prec	0.57	0.19		0.18	0.27	0.28

Please Contact: Douglas Turnbull <turnbull@cs.swarthmore.edu>
Data, Papers, and additional Information can be found at: http://cosmal.ucsd.edu/cal/

[1] Turnbull, Barrington, Torres, Lanckriet *Semantic Annotation and Retrieval Music and Sound Effects.* TASLP 2008
[2] Turnbull, Liu Barrington, Lanckriet *Using Games to Collect Semantic Information About Music.* ISMIR 2007
[3] Knees, Pohle, Schedl, Schnitzer, Seyerlehner *A Document-Centered Approach to a Natural Language Music Search Engine.* ECIR 2008
[4] Barrington, Yazdani, Turnbull, Lanckriet *Combining Feature Kernels for Semantic Music Retrieval.* ISMIR 2008
[5] Turnbull *Design and Development of a Semantic Music Discovery Engine.* Ph.D. Thesis, UC San Diego 2008