

# Facilitating Comprehensive Benchmarking Experiments on the Million Song Dataset

Alexander Schindler and Rudolf Mayer and Andreas Rauber



Institute of Software Technology and Interactive Systems  
Information and Software Engineering Group  
Vienna University of Technology



## Motivation

### Advantages of the MSD:

- Test algorithms on large-scale collection
- Real-world scenarios
- Freely available
- Inter-linked to other data ressources

### Shortcommings of the MSD:

- Source audio files are not available
- Sample files are not easily available
- Researchers are limited to features provided
- No possibility to develop new or test different features

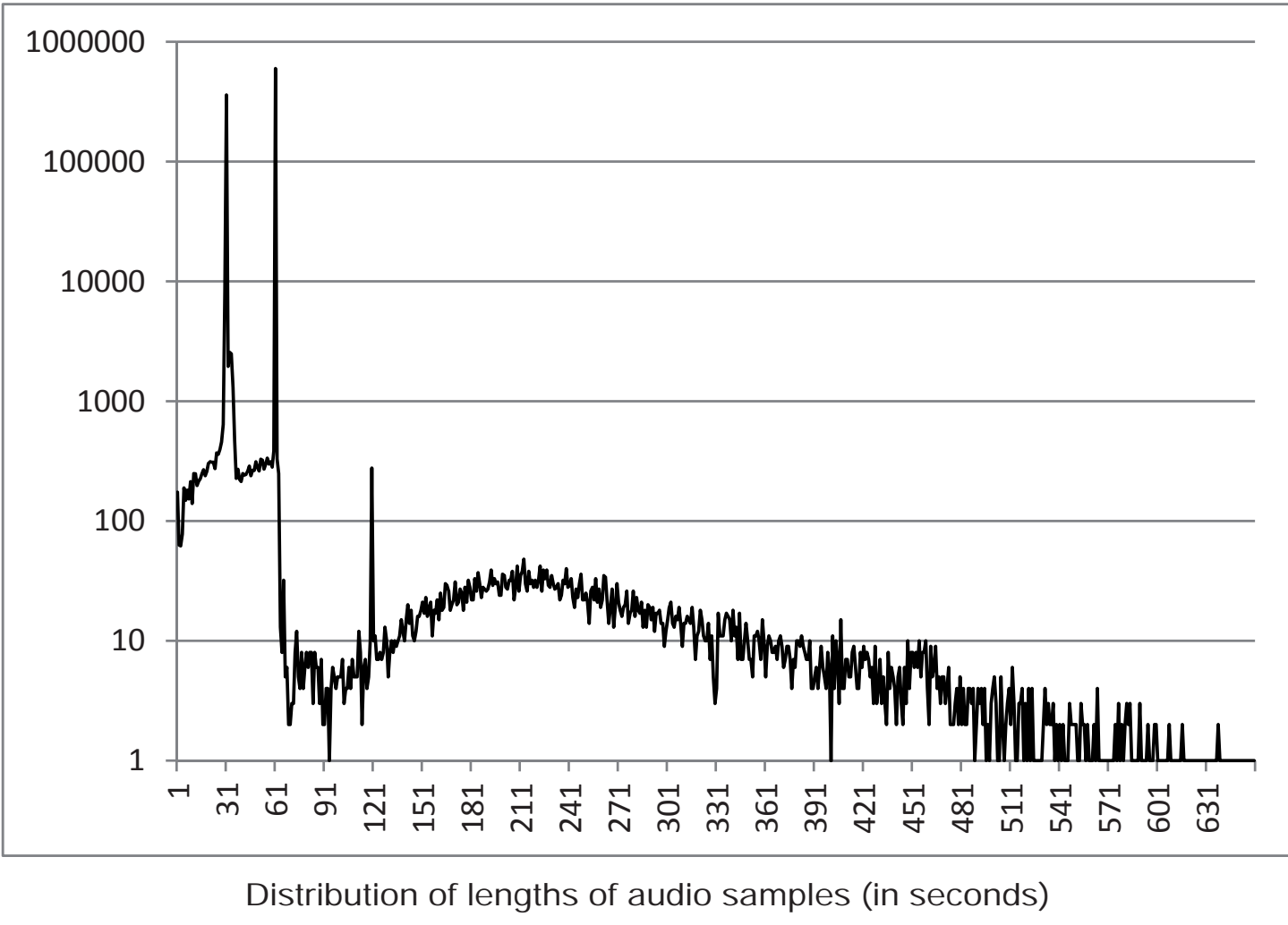
### Contribution:

- A wide range of conventional features
- Three sets of Genre and Style labels extracted from Allmusic.com
- A wide range of partitions and splits for instant use and comparability
- A benchmarking platform for distributing
  - further conventional or newly developed features
  - research results

## Statistics

Downloads		
MP3s	994,960	
Size in Gigabyte	621	
Samplerate		
22.050	768,710	77.26%
44.100	226,169	22.73%
other	81	0.01%
Bitrate		
128	646,120	64.94%
64	343,344	34.51%
other (VBR)	5,494	0.55%
Sample length		
~30 sec (+/- 1)	368,003	36.99%
~60 sec (+/- 1)	597,550	60.06%
other	29,407	2.96%
Channels		
Mono	6,342	0.64%
Stereo	150,779	15.15%
Joint stereo / dual channel	837,839	84.21%

Extracted Features		
Rhythm Patterns	994,175	99.42%
Marsyas	995,000	99.50%
jMir	993,668	99.67%
Temporal Echonest Features	1.000.000	100.00%
Intersection	992,865	99.29%



## Additional Features extracted from Audio Samples

Temporal Echonest Features	Rhythm Patterns	jMir	Marsyas
Empirically selected combination of features provided by the MSD. Aggregated by calculating statistical measures from the audio descriptorr.	A set of features by Rauber, Lidy et al. based on psychoacoustical models, capturing fluctuations on frequency bands critical to the human auditory system.	Features from the jMir Framework developed by Cory McKay.	Features from the Marsyas Framework de-veloped by Georege Tzanetakis et al.
<a href="http://www.ifs.tuwien.ac.at/mir/msd/temporalEN.html">http://www.ifs.tuwien.ac.at/mir/msd/temporalEN.html</a>	<a href="http://www.ifs.tuwien.ac.at/mir/audiofeatureextraction.html">http://www.ifs.tuwien.ac.at/mir/audiofeatureextraction.html</a>	<a href="http://jmir.sourceforge.net/">http://jmir.sourceforge.net/</a>	<a href="http://marsyas.info/">http://marsyas.info/</a>
<b>Feature: Dim:</b>	<b>Feature: Dim:</b>	<b>Feature: Dim:</b>	<b>Feature: Dim:</b>
TEN 224 Further description of Temporal Echonest Features: A. Schindler, A. Rauber, <i>Capturing the temporal domain in Echonest Features for improved classification effectiveness</i> , Adaptive Multimedia Retrieval (AMR 2012), Copenhagen, Denmark, Oct. 24-25 2012.	RP 1440	MFCC 26	MFCC 52
EN0 12	RH 60	Spectral 16	Chroma 48
EN1 24	SSD 168	Methods o.M. 10	Timbral 124
EN2 24	MVD 420	Area M.o.M. 20	
EN3 90	TRH 420	Lin. Pred. Cod. 20	
EN4 96	TSSD 1176		
EN5 192			
These features are based on the original MSD audio descript-ors and are not extracted from the downloaded audio samples.			

## Annotations

Three different expert annotated gorund truth assignments consisting of genre and style labels are provided. Two partitions of Top level genres with predominating classes and one style/sub-genre partition with balanced distribution of classes.

Genre Labels		Style Labels	
Pop/Rock	238,786	Big Band	3,115
Electronic	41,075	Blues Contemporary	6,874
Rap	20,939	Country Traditional	11,164
Jazz	17,836	Dance	15,114
Latin	17,590	Electronica	10,987
R&B	14,335	Experimental	12,139
International	14,242	Folk International	9,849
Country	11,772	Gospel	6,974
Reggae	6,946	Grunge Emo	6,256
Blues	6,836	Hip Hop Rap	16,100
Vocal	6,195	Jazz Classic	10,024
Folk	5,865	Metal Alternative	14,009
New Age	4,010	Metal Death	9,851
Religious	8814	Metal Heavy	10,784
Comedy/Spoken	2067	Pop Contemporary	13,624
Stage	1614	Pop Indie	18,138
Easy Listening	1545	Pop Latin	7,699
Avant-Garde	1014	Punk	9,610
Classical	556	Reggae	5,232
Childrens	477	RnB Soul	6,238
Holiday	200	Rock Alternative	12,717
		Rock College	16,575
		Rock Contemporary	16,530
		Rock Hard	13,276
		Rock Neo Psychedelia	11,057

## Benchmark Sets

We provide a number of benchmark partitions that researchers can use in their future studies, in order to facilitate repeatability of experiments with the MSD beyond x-fold cross validation. We also encourage and provide a platform for exchange of results obtained and new partitions created via our website:

<http://www.ifs.tuwien.ac.at/mir/msd/>

- Splits with all the ground truth assignments into genre and style classes
- Splits considering the sample rate of the files, i.e. only the 22khz samples, only the 44khz samples, and a set with all audio files.
- A split with a fixed number of training samples, equally sized for each class, with 2,000 and 1,000 samples per class for the genre and style data sets, respectively. This excludes minority classes with less than the required number of samples.
- "Traditional" splits into training and test sets, with 90%, 80%, 66% and 50% size of the training set, applying stratification of the sampling to ensure having the same percentage of training data per class, which is important for minority classes.
- Splits into training and test sets with an album, time or artist filter, i.e. avoiding to have the same artist in both the training and test set; both stratified and non-stratified sets are provided

## Your Features / Your Results

- Copyright restrictions prevent the redistribution of the downloaded audio samples.
- We provide the possibility to extract newly developed features and to publish them on the Web page
- We provide a benchmarking platform where you can publish your results and compare them against others

To further enhance large scale benchmarking based on the Million Song Dataset we offer to extract further features from the downloaded audio samples. If you are interested in having your audio descriptors extracted, please contact us

[{schindler,mayer,rauber}@ifs.tuwien.ac.at](mailto:{schindler,mayer,rauber}@ifs.tuwien.ac.at)