Identifying Video Sources by Identifying Audio Compression

Zafar Rafii

04/13/18

Gracenote: Entertainment Data and Tech

Music
Industry standard music data and media recognition powering the world's leading digital music platforms.

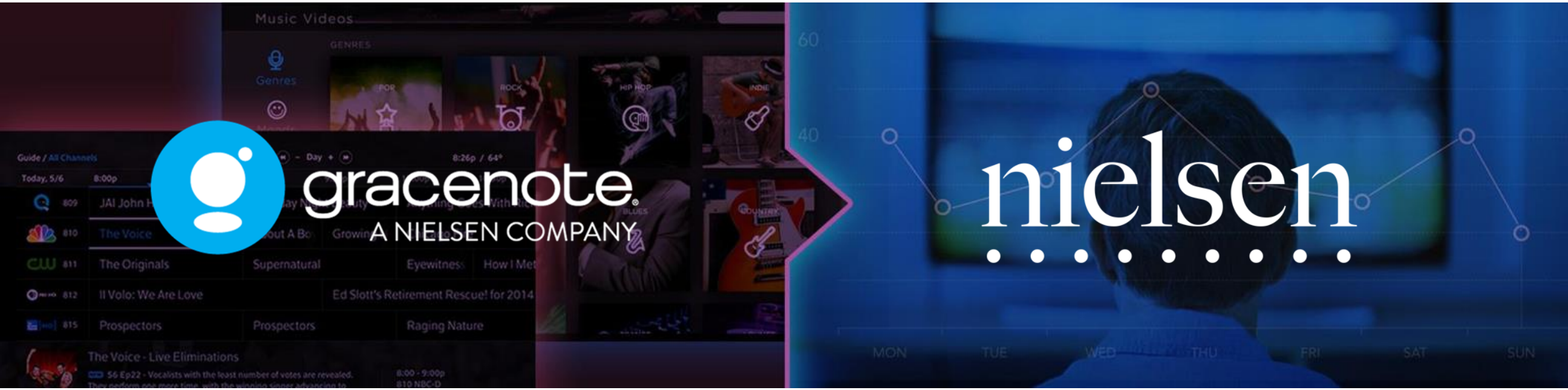
Auto
Music data and media recognition solutions fueling infotainment systems in +100M cars from every major OEM and supplier.

Video
TV listings and descriptive movie and TV data driving user interfaces of the world's top pay-TV providers and OTT services.

Sports
Sports scores and statistics covering the world's top leagues, events, teams and players - powering leading online media.

Personalization
Industry standard ACR technology featured in more than 25M connected TVs from top brands.

Gracenote: A Nielsen Company



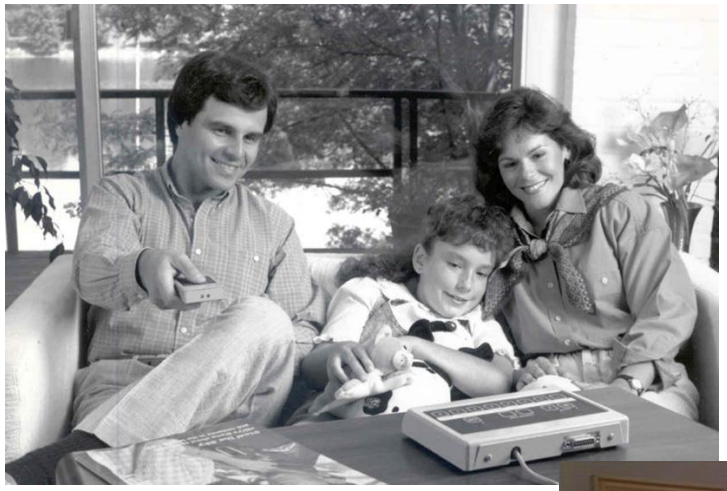
Media Discovery

Gracenote data powers product interfaces and discovery algorithms for top TV, movie and music platforms – helping millions of consumers find what to watch and listen to daily.

Media Consumption

Nielsen's gold standard audience measurement solutions determine what millions of people are watching and listening to across these same devices, platforms and services.

Nielsen: Audience Measurement



Project

- Problem: In some cases, the only data that can be obtained is the audio-video content, which does not help for identifying the source.

- Assumption: Different video sources use different lossy formats for their audio content, introducing distinct compression artifacts in it.

- Solution: By analyzing the compression artifacts in the audio content, we can then infer the video source: audio compression identification!

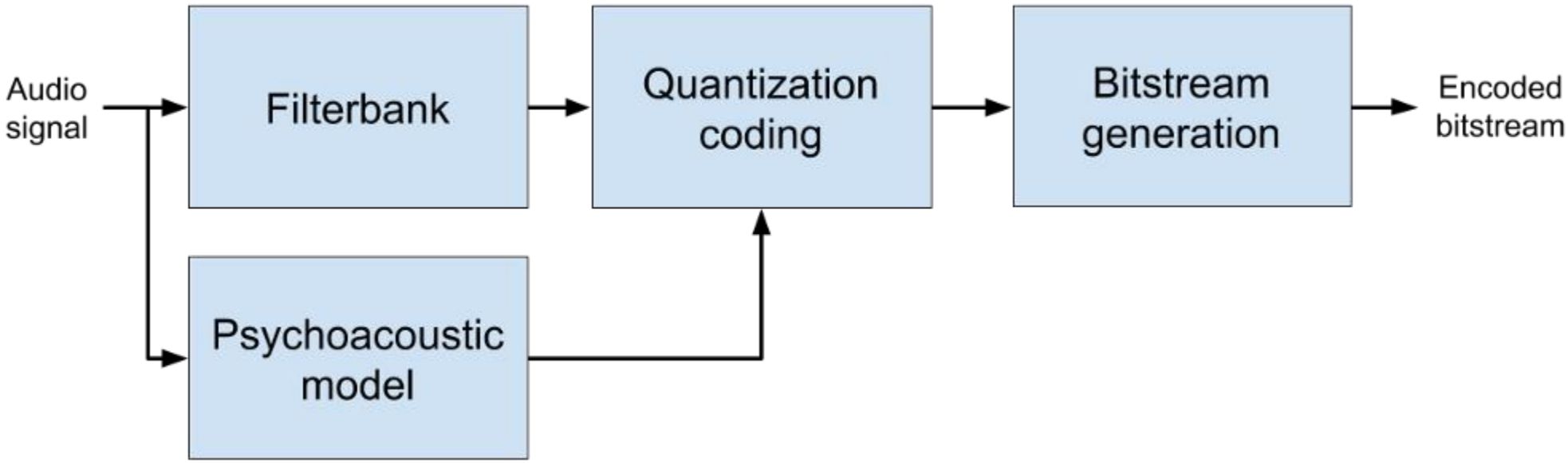
Lossy Compression

- The objective of data compression is to reduce the size of data for more efficient storage or transmission, while preserving the quality.

- While lossless compression encodes data in a reversible manner, lossy compression removes perceptually less significant information.

- Lossy audio coding formats such as MP3, AAC, AC-3, Vorbis, and WMA are widely used in audio/video files and radio/TV broadcasting.

Lossy Compression Structure



https://deezer.io/deezer-at-icassp-2017-347cd296bd45

Filterbank Parameters

- Time-frequency transform:
  - Polyphase Quadrature Filter (PQF)
  - Modified Discrete Cosine Transform (MDCT)
  - Hybrid structure (e.g., PQF+MDCT for MP3)
- Window function:
  - Sine window (e.g., MP3 and AAC)
  - Slope window (e.g., Vorbis)
  - Kaiser-Bessel-derived (KBD) window (e.g., AAC and AC-3)
- Window length:
  - Normal (long) window length (typically, with half-overlapping)
  - Short window length for transients (e.g., long 2048 vs short 384 samples for AAC)
  - Hybrid stop and start windows between long and short windows

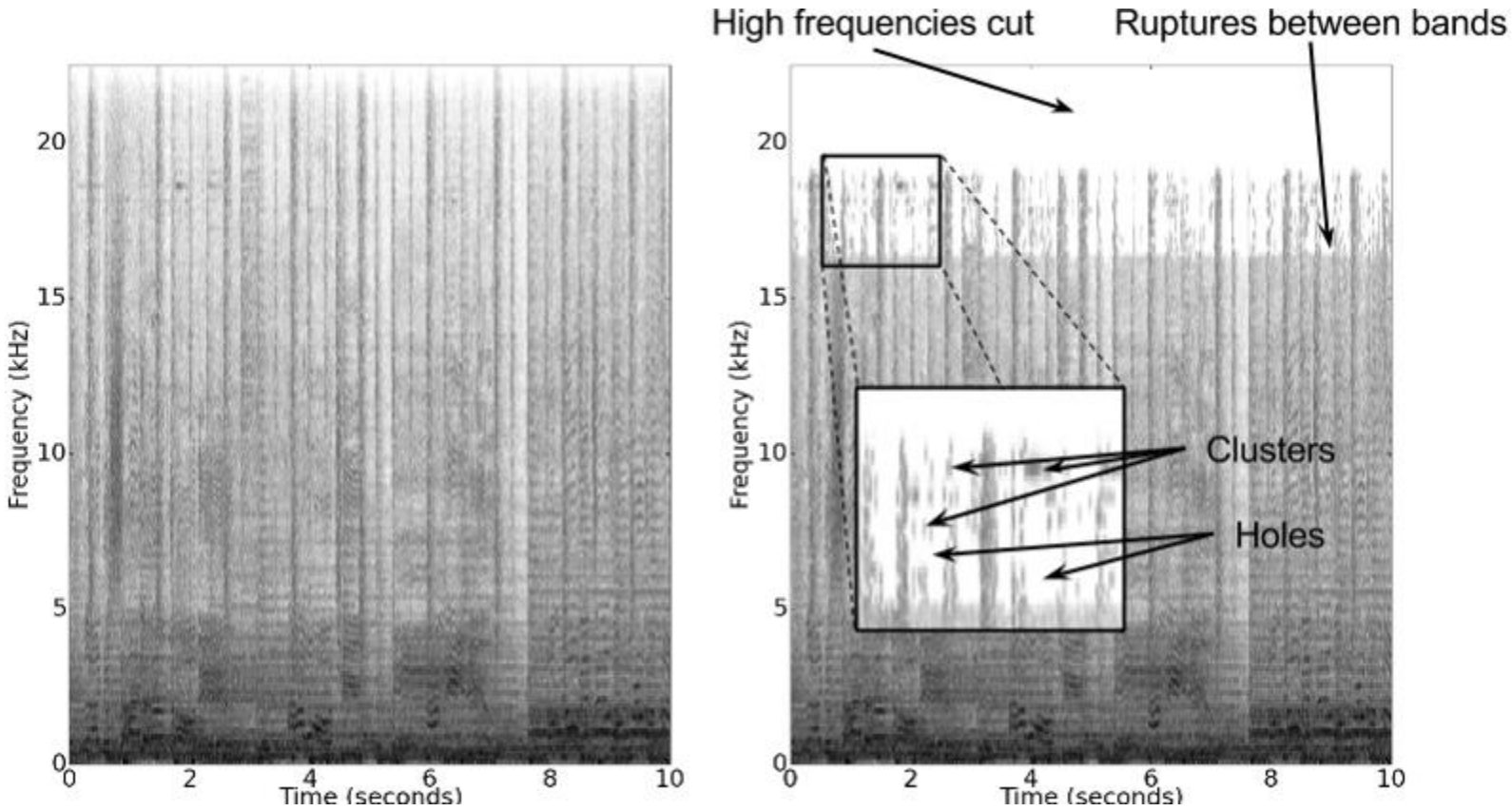
Lossy Compression Identification

- The goal is to identify the compression parameters, the bit rate, the coding format, etc., which were used by the lossy compression.

- Applications include detection of audio alterations, authentication of the audio quality, reverse-engineering of the encoding process, etc.

- The typical approaches in the literature use signal processing (i.e., time-frequency analysis) and/or machine learning (e.g., SVMs).

Lossy Compression Artifacts



https://deezer.io/deezer-at-icassp-2017-347cd296bd45

Observations

- The compression artifacts will be the most apparent when the analysis happens on the same audio samples as in the encoding.
- The compression artifacts will be more apparent if the analysis uses the same compression parameters as in the encoding.
- The compression artifacts can then be measured by looking at the energy in the spectrogram for a given framing and parameters.

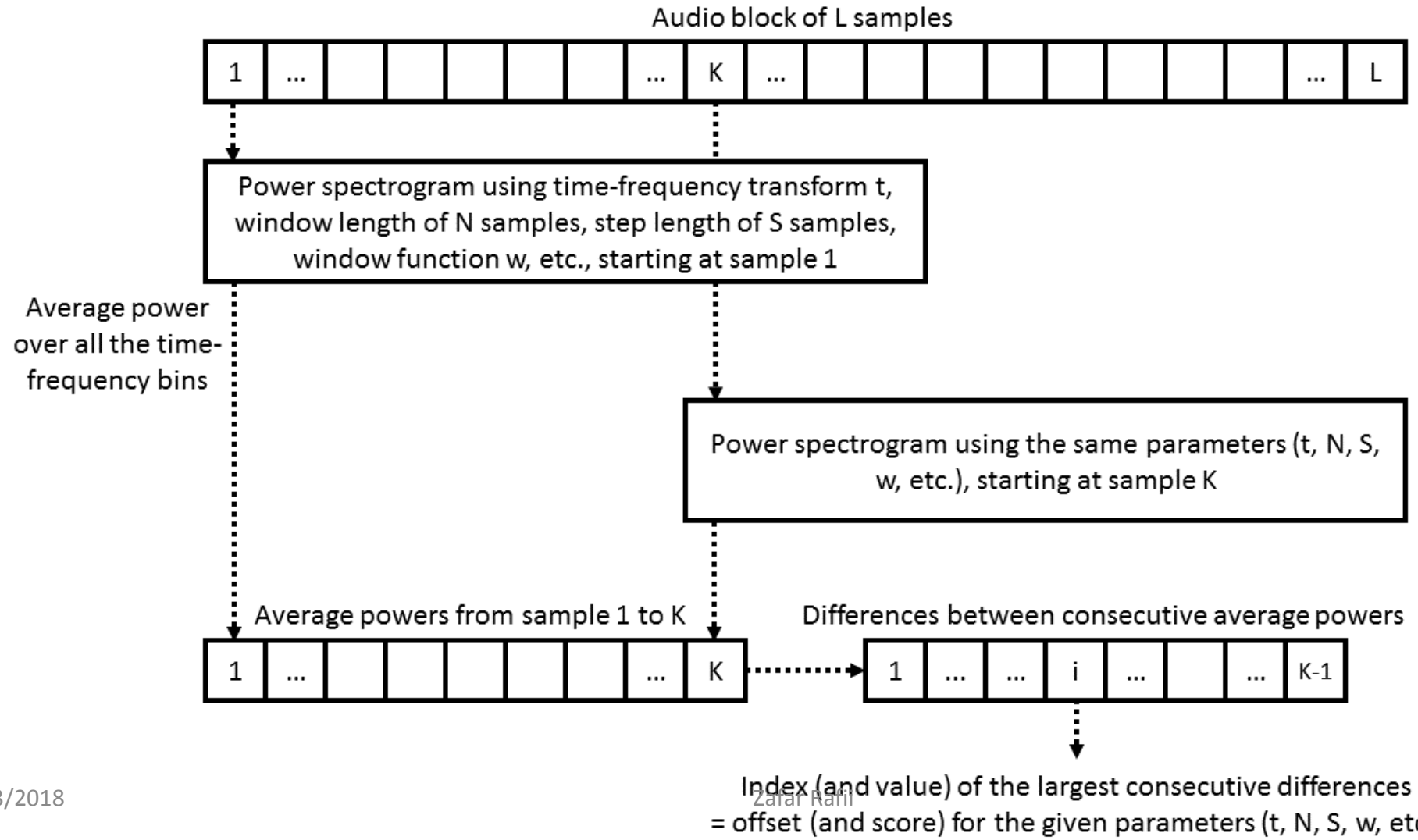
Proposed Approach

1. Measure the differences between the power spectrograms at consecutive offsets and return the largest value and related index.

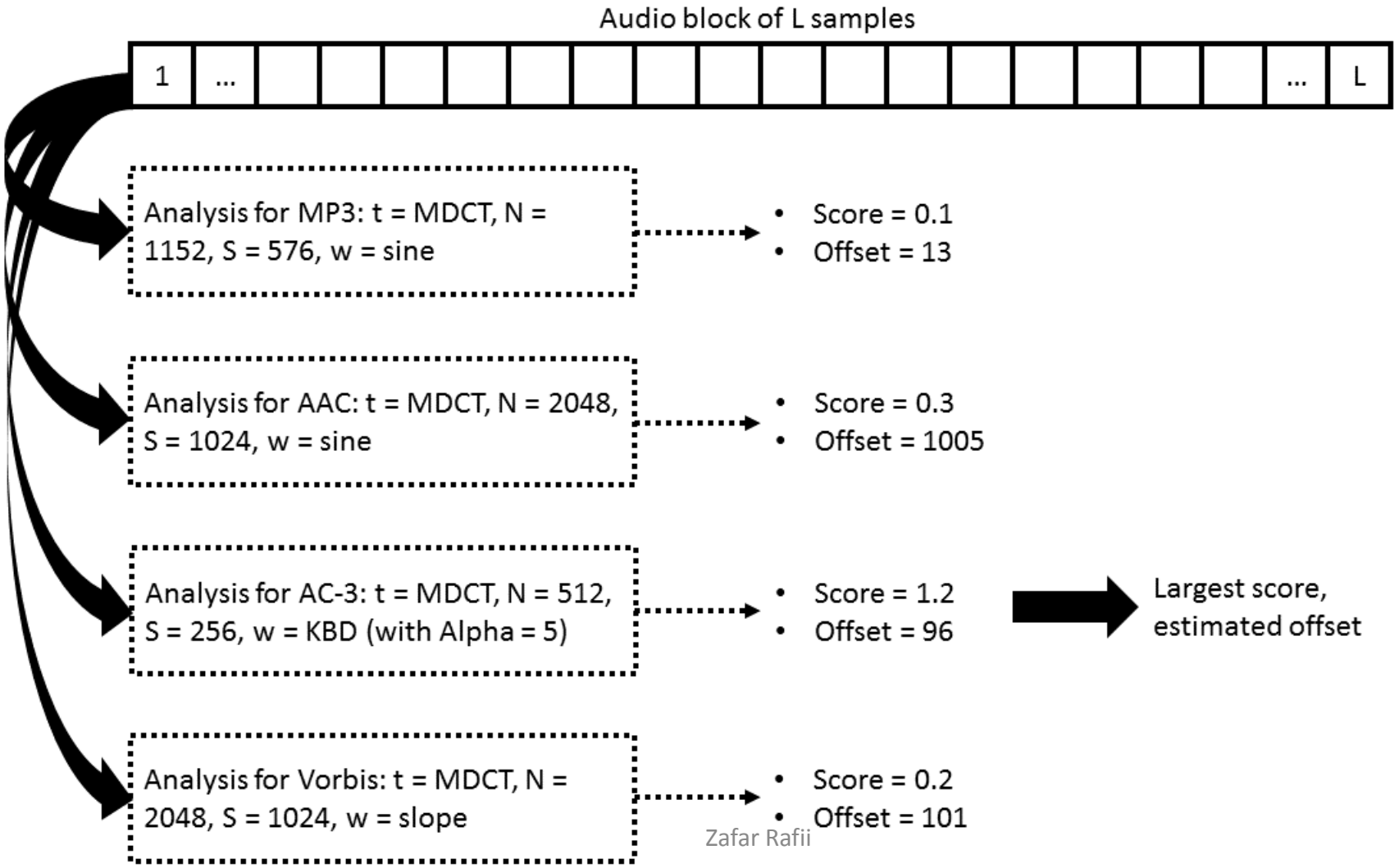
2. Run this process for different sets of parameters corresponding to known coding formats and return the set with the largest score.

3. Run all of this process on successive time blocks in the audio signal and combine the scores (with their indices) to refine the results.

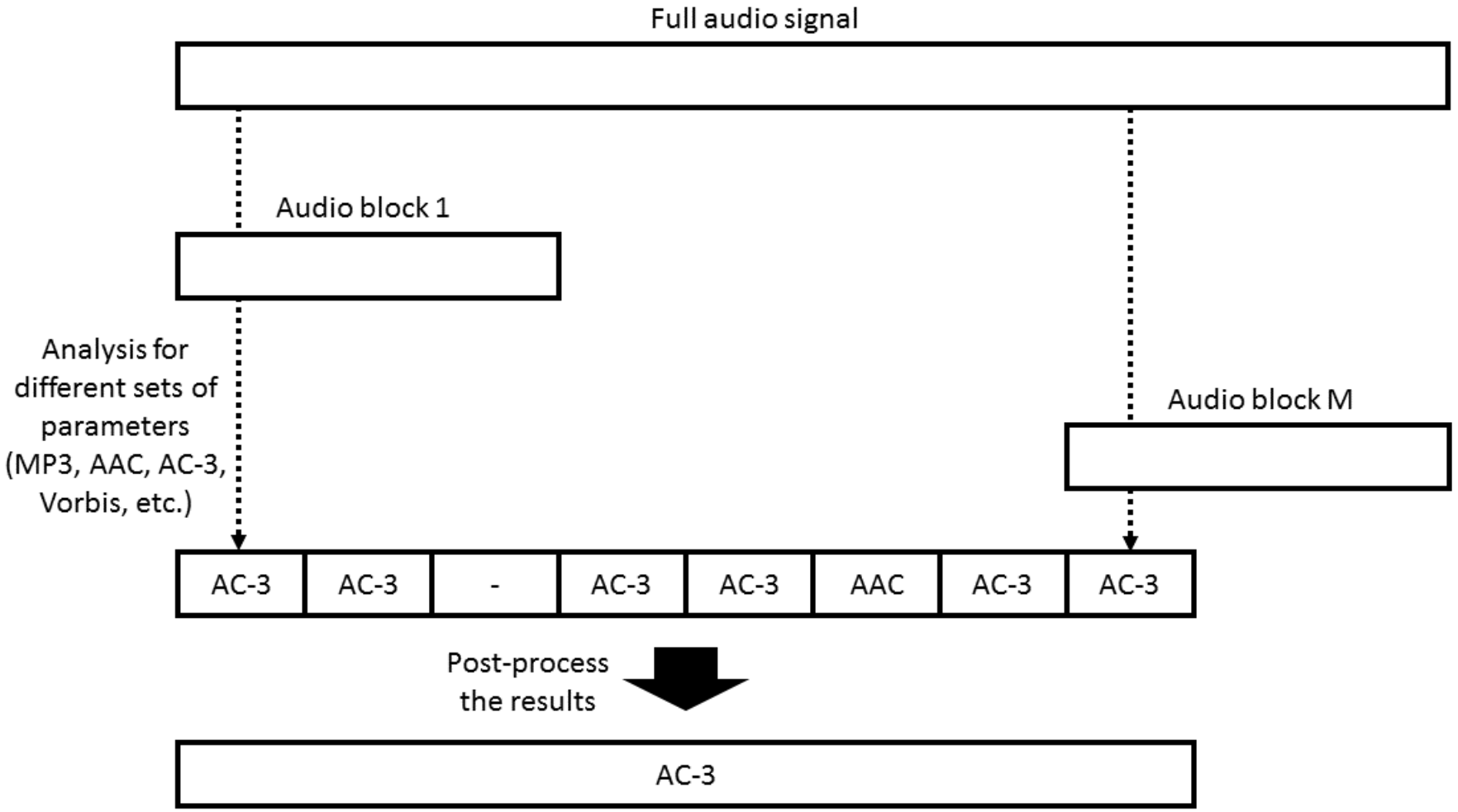
Proposed Approach 1/3



Proposed Approach 2/3

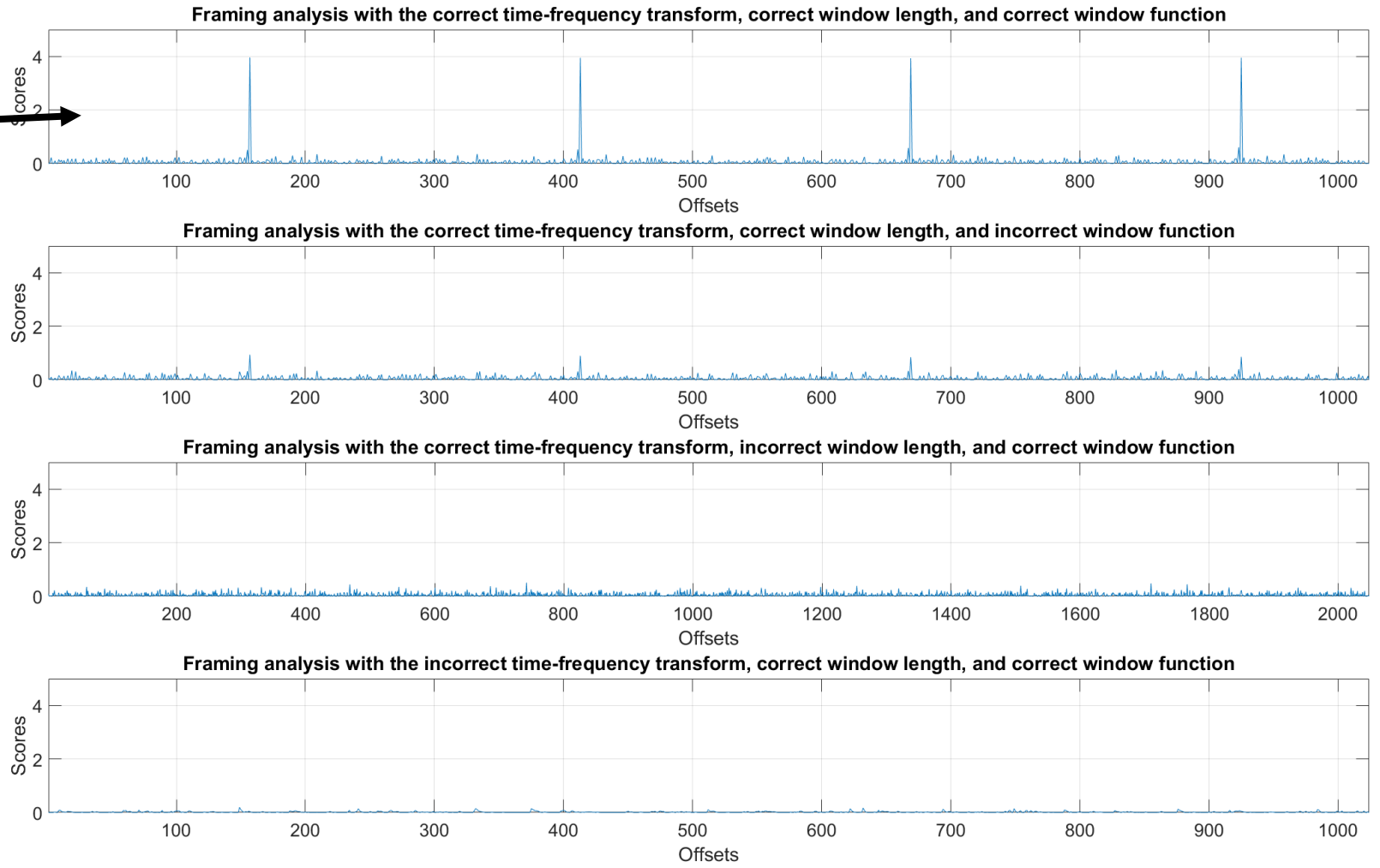


Proposed Approach 3/3

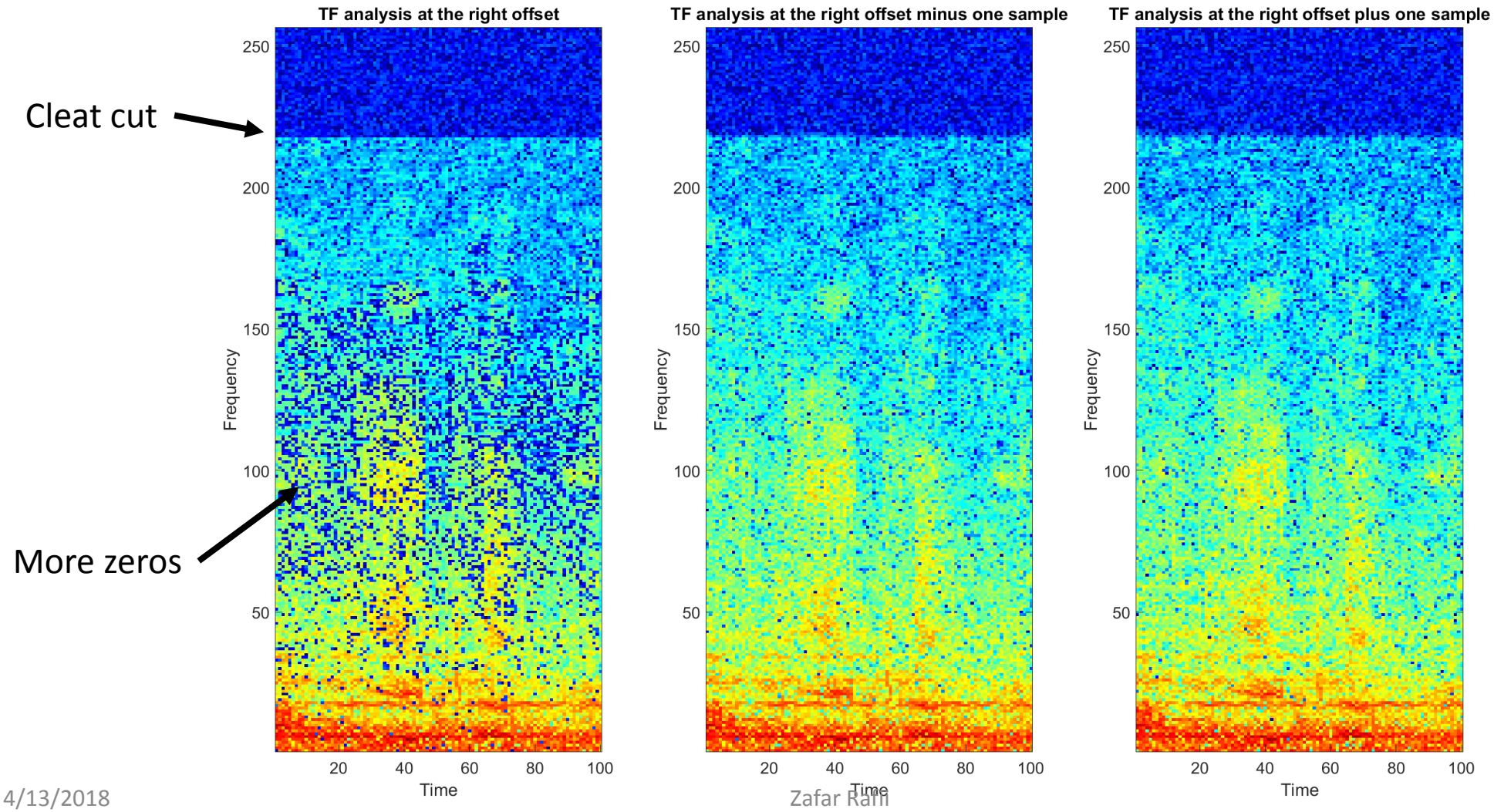


Framing Analysis

Clear peaks



Spectrogram Analysis



Limitations

- Not always discriminative: different sources could use the same compression parameters or an unknown audio encoding.
  - Other cues can also be analyzed (e.g., high-frequency cut), and, in any case, any extra clue is a welcoming additional layer of knowledge for the system.
- Not very robust: external noise, sample desynchronization, and high-bit rates will make the measurements more challenging.
  - The deviations can be taken care of, to some extent, through some post-processing (as described earlier) and using longer or multiple audio segments.
- Computationally expensive: time-frequency analysis is performed for every sample and for every set of parameters to be tested.
  - There are ways to speed up the computation (e.g., using previously analyzed segments) and the time-frequency analysis can be optimized as well.

Some References

1. ten Kate – Maintaining Audio Quality in Cascaded Psychoacoustics Coding – 1996.
2. Herre and Shlug – Analysis of Decompressed Audio - "The Inverse Decoder“ – 2000.
3. Moehrs et al. – Analysing Decompressed Audio with the "Inverse Decoder" - Towards an Operative Algorithm – 2002.
4. Bosi and Goldberg – Introduction to Digital Audio Coding and Standards – 2003.
5. Yang et al. – Detecting Digital Audio Forgeries by Checking Frame Offsets – 2008.
6. Yang et al. – Defeating Fake-Quality MP3 – 2009.
7. Gärtner et al. – Efficient Cross-Codec Framing Grid Analysis for Audio Tampering Detection – 2016.
8. Seitchter et al. – AAC Encoding Detection and Bitrate Estimation using a Convolutional Neural Network – 2016.
9. Hennequin et al. – Codec Independent Lossy Audio Compression Detection – 2017.
10. Kim and Rafii – Lossy Audio Compression Identification – 2018 (submitted).

Some Links

- http://zafarrafii.com/
- https://github.com/zafarrafii
- http://www.gracenote.com/
- https://sisec.inria.fr/
- https://www.meetup.com/bishbash/