Cloudspeakers (demo submission for ISMIR 2010)

Chris Bol (Cloudspeakers), Stan van de Burgt (Cloudspeakers), Arjen P. de Vries (CWI)

Cloudspeakers (cloudspeakers.com) is a Dutch music aggregator of links to (legal) audios, videos and reviews. Cloudspeakers' primary objective is to recommend *new* music to the 2000 subscribers, as soon as this music is available - a setting that is notoriously hard for classic collaborative filtering techniques, due to sparseness of data. For, recommending new music implies a cold-start recommendation problem (on the item dimension). Especially when issued by new artists, the lack of usage information that is the basis for collaborative filtering approaches would stop us from making accurate predictions for quite a while after its first online availability.

With this goal and its associated cold-start problem in mind, we have moved away from the common approach to collect clicks or downloads as indicators of musical taste, and focus our efforts on locating music reviewers and their opinions instead. Music reviewers are usually the first to have access to new music. They are also particularly well qualified to assess new music on its merits, and, to compare its style to related but better known music. Cloudspeakers has therefore created an online environment to capture the collective opinion of a selected group of 'taste makers': the music critics of the leading music webzines and blogs. We match people's musical taste using the expert opinion of music critics about new music as a 'taste proxy'.

The first step is to match users to a group of music critics. As soon as users have identified a few favourite artists or albums, to populate their user profile, they are matched to reviewers who have the highest compatibility; based on the critics' reviews and ratings of those artists and albums. Next, if a certain new track or album is significantly tipped by the group of reviewers, this release will be recommended to the specific user. Because these recommendations are triggered by the publication of (early) press, Cloudspeakers recommends new music in a very early stage; sometimes even before the official release.

Cloudspeakers aggregates links to music reviews, legal online audio and video files, and links to artist pages, from sources including MySpace, Twitter and Wikipedia. Music reviews are collected from nearly 200 webzines and blogs in six different languages. Our database lists nearly 7000 reviewers, with all their reviews and ratings; a source that generates over 1000 new reviews per week.

Adding a new webzine or blog is a manual process by creating a regular expression to convert the HTML page to XML (the database contains nearly 400 regular expressions, today). After adding a regular expression, the source will be checked for updates automatically, on a regular basis. For each link that we collect, we identify artist name, album name, the full text of the review and its URL, the name of the source, the review's publication date, the reviewer's name, and, where available, the reviews rating. This information is subsequently matched to the MusicBrainz discography (musicbrainz.org). By matching all our content to MusicBrainz, spelling errors are solved and no duplicate artists are present in DB. MusicBrainz IDs provide a universal and world-unique identifier for music, so different organizations (like BBC, Last.fm) who also use MusicBrainz can easliy exchanged data.

Our proposed demonstration will highlight the particular advantages of recommendation via music critics, based on the data collected from actual operations of the Cloudspeakers service, focusing specifically on making timely recommendations. The accompanying poster will illustrate the system architecture and user experience by example.