Pandora and the music genome project: Song structure analysis tools facilitate new music discovery

Article · September 2006

1 author:

John R Joyce
Arcana Informatica
63 PUBLICATIONS **24** CITATIONS

SEE PROFILE

Published on *Scientific Computing* (http://www.scientificcomputing.com)

Pandora and the Music Genome Project: Song structure analysis tools facilitate new music discovery

John R. Joyce, Ph.D.

Most music sites allow some form of new music discovery. With some, the system is based on a form of rating or recommendation from other users. This could show up like 'Other users also bought XXX,' implying that if you liked this song, there is a good chance that you might like the other too. However, there are a few music services living on the Web that take a different approach. One of these is the Music Genome Project.

Conceived by Tim Westergren, the Music Genome Project attempts to take a more analytical approach. This project was launched in January of 2000 with an attempt to analyze the structure of a song so that they could identify similar songs that a searcher might like. Originally a project of his company Savage Beast, it took five years and 30 experts in music theory to build their database to the point of usefulness. Five years because each of the songs had to be individually listened to and up to 400 musical attributes manually assessed, which required 20 to 30 minutes for each four minutes of song. As of May 2006, this library contains in excess of 400,000 analyzed songs from 20,000+ contemporary artists.

[1]

click to enlarge

Binary Pandora generally requires no installation, as it works its magic through the use of Adobe's Flash plug-in and some judicious Java Script.

Unlike other projects, such as musicologist Alan Lomax's Global Jukebox, the impetus of the Music Genome Project is to be a successful commercial venture. To do that, they are not making the mistake of attempting to compete head-on with ventures such as iTunes and Rhapsody. Instead, they are attempting to take advantage of an economic concept called The Long Tail. In this concept, which posits an environment with limitless selection and cheap/easy distribution, you make your money by going beyond the blockbuster hits to identify the obscure gems that might otherwise be lost. Once a song is identified, listeners are then directed to Amazon.com and other sites for purchase. Until August of 2005, the only ones experiencing the power of this project were those who accessed it via in-store music recommendation kiosks, such as those in Barnes & Noble, Best Buy and Tower Records.

Last year saw the release of its online interface, called Pandora Radio, and a company name change from Savage Beast to Pandora Media. Pandora, from Greek mythology, received many gifts from the gods, including both the gifts of music and curiosity which, in this case, are celebrated and rewarded. This binary Pandora generally requires no installation, as it works its magic through the use of Adobe's Flash plug-in (version 7 or 8, 8 being preferred) and some judicious Java Script. It will work on machines running Windows 2000 or Windows XP with either MS Internet Explorer 7 or 8, or with Firefox. Unlike many services, it will also run on the MacOS X 10.3+ with either Safari or Firefox. The processor should be running at 1 GHz or more, with a

minimum of 256 MB of RAM. A broadband Internet connection is also required, as dial-up connections are not supported.

Unlike its initial indexing, using Pandora is actually very simple. Simply go to the Pandora site and enter an artist name or song title into the main player field and click Create to create your station. Typically, entering an artist's name will result in a more generalized selection than entering a more specific song title. A good example of this would be The Beatles, who had an extremely diverse repertoire. The station's selections can be adjusted by giving a thumbs-up or thumbs-down for additional songs and/or artists. According to their FAQ, giving a song a thumbs-up can add around a hundred songs to a station. Adding an artist on the other hand, might add several hundred songs to a station.

You can listen to up to 10 selections in a session without registering. You can register in one of two ways, both of which supply the same feature set. The most popular is to register for the free version, which could potentially subject you to a barrage of supporting advertisements, although none seemed intrusive when I checked it out. Alternately, you can purchase an annual subscription for $36 or a quarterly subscription for $12, which provides ad-free access. Registering is simple, you just supply an e-mail address as a User ID (they do promise not to distribute it) and a password. They also require you to enter your ZIP code to ensure that you reside in the United States, as that is the only area in which their license will allow them to currently distribute music. There doesn't appear to be any provision to verify the accuracy of the value entered, so you could probably enter any value you wanted, allowing you to download from outside the U.S., but as they said in *Kentucky Fried Movie*, '...that would be wrong!'

The remaining restrictions are all required for compliance with the beloved DMCA (Digital Millennium Copyright Act). Basically, these boil down to not allowing you to directly play a selected song. You can specify the channel as tightly as you want, but the system will always randomize what it's going to play. While the program's structure could be changed by the time you read this, there have been a number of threads on the Net about how to capture the temp files that Pandora downloads. There has been some question regarding whether capturing these temp files violates the DMCA or would fall under the fair use clause, such as recording a TV show for later viewing. It may well depend on the use to which you put any of these captured files. Trying to follow the arcana of some of these bills can be a bit too mind numbing. Many a time that I wish our dear congressmen had to actually write, or at least read, the bills they propose and pass, but I don't know if that would actually help or hurt (having on occasion been exposed to the statements of 'fact' that some of them espouse).

Some reviewers seem to feel that Pandora is best for increasing your depth of exposure to a particular genome, as opposed to broadening your musical exposure. While this may be true to some extent, at the very least, it is a good way to discover additional artists working in a style you like. The songs it tends to pick definitely have a different sound to them than say Yahoo!'s Launchcast radio. As most of the Music Genome selectors have not been revealed, an attempt to create an Open Source version of the Music Genome is being made.

While it appears to violate Pandora's terms of service, another application of potential interest is Pyrrha from Pyrcast Online. Requiring Windows XP, this application purports to allow you to create podcasts of songs to play offline. It has been highly reviewed by some publications, but it appears that the originators have gone to some lengths to hide their identity, which is rather suspicious. If you decide to try it out, be cautious and understand that it might result in your

Pandora account being dropped or malicious software being installed on your machine. At the very least, make sure that all of your security patches are up to date before playing with it. Pandomax is another program that appears to perform a similar function, along with similar reticence regarding who they actually are.

In any event, happy listening!

Music Services Resources

Digg — Pandora Radio	www.digg.com/music/Pandora_streaming_radio_(dugg_earlier)_caches_to_your_TEMP_folder_ [2]
Global Jukebox	www.alanlomax.com/style_globaljukebox_Naimark.html [3]
HowStuffWorks — Pandora Radio	entertainment.howstuffworks.com/pandora.htm [4]
Open Source Music Genome Project	osmgp.pbwiki.com [5]
Pandomax	www.pandomax.com/index.html [6]
Pandora Media	www.pandora.com [7]
Pyrcast Online	www.pyrcast.com [8]

John Joyce is the LIMS manager for Virginia's State Division of Consolidated Laboratory Services. He may be contacted at editor@ScientificComputing.com [9]

Links:
[1] http://www.scientificcomputing.com/sites/scientificcomputing.com/files/legacyimages/0609/sc69OLS_lrg.gif
[2] http://www.digg.com/music/Pandora_streaming_radio_(dugg_earlier)_caches_to_your_TEMP_folder_
[3] http://www.alanlomax.com/style_globaljukebox_Naimark.html
[4] http://entertainment.howstuffworks.com/pandora.htm
[5] http://osmgp.pbwiki.com/
[6] http://www.pandomax.com/index.html
[7] http://www.pandora.com/
[8] http://www.pyrcast.com/
[9] mailto:editor@ScientificComputing.com