Google Summer of Code 2008: Amarok Mass-Tagging

Synopsis

I will add a feature to Amarok to make music collection management easier. This would entail integrating track data available from FreeDB or Musicbrainz in a very user friendly manner. This could entail interesting ways to automatically identify poorly tagged tracks and update/fix them.

Project

The contribution of mass tagging is fairly straightforward. A system which can effectively perform mass tagging generally possess the ability to identify poorly tagged files, accurately modify tags, and have control through an intuitive user interface.

Poorly tagged songs include inaccurate or misspelled fields, incomplete fields, and missing fields. The system should be able to use any existing tag information to identify a song even in the event of a misspelling or when characters such as an underscore are used in place of a space. In the case of complete absence of useful tag information, it may be possible to identify the track using file name, location in file system (e.g. track is named “Track 1” with no tag data but resides in a folder named “The Beatles – White Album” folder) and/or song signature without prompting the user. This is likely to be the biggest challenge but existing programs such as “EasyTag” and “PicardQt” can be studied for existing solutions.

The second part to this project would entail being able to accurately modify tags. This is trivial and means modification of tags needs to be correct. For an example, given only an artist name and song, it may be possible populate the album field with any album the song appears in or one of the user’s choosing, however, the user may not know the correct album, and the context of the track may need to be inspected.

Finally the user interface is the last challenge presented in implementing this feature. The task of mass tagging songs is something most users will use relatively infrequently. The user interface should be designed to be extremely intuitive to allow users to tag a mixed set of tracks, an entire album, or an entire collection in a way with no learning curve. At the moment right clicking on

an album and selecting “Edit Information for 16 Tracks” allows you apply uniform values (such as an album year). This is an example of where certain user interface for mass tagging should appear, however it must be properly integrated and work well with single track editing as well.

Conditions for Success

- Poorly tagged songs are correctly identified and updated. This also means properly tagged songs or songs the user explicitly tagged in a certain way are not modified.
- User Interface for mass tag editing is cleanly integrated and user friendly
- Mass tagging is fast as possible. This means it probably does not scan every song every single time. It should also use well optimized features which already exist when possible (e.g. MySQL integration)

Benefits to Amarok

1. Quickly and accurately tagged music which will leader better collection management, search, and service synchronization
2. More convenient Amarok user experience – users will not need to install and update another software package which would be used occasionally at best.

Road Map

1. *May to mid June*
   Familiarize myself with KDE/Amarok practices, community members, relevant technologies. Discuss with community members and work with my mentor to ensure proper planning
2. *Mid June to early August*
   Finish design and implement the mass tagging feature. Work on submitting midterm evaluation
3. *Early August to end*
   Scrub code, perform final tests, present to community members for feedback, and work on final evaluation

Biography

I am currently an undergraduate student at the University of Southern California (USC) majoring in Computer Science. Upon completion of my undergraduate degree in May 2008, I am continuing my education to acquire a M.S. in Computer Science from the same institution. It is at USC where I have studied with the likes of Leonard Adleman before moving to a specialization in Software Engineering. I am currently conducting directed research on real time architecturally aware middleware systems. I serve as an officer for the UPE Computer Science Honors Society and am a member of ACM.

I begun using Linux for fun in 2006 and instantly fell in love. As soon as I moved to full time use of Ubuntu, I became familiar with Amarok. To those who know me well consider me an Amarok evangelist. I have encouraged many to become familiar with Amarok and use it as soon as it is available on Windows (Amarok 2.0?).