Chapter 13

Building Tag-Aware Groups for Music High-Order Ranking and Topic Discovery

Dimitrios Rafailidis
Aristotle University, Greece

Alexandros Nanopoulos
University of Hildesheim, Germany

Yannis Manolopoulos
Aristotle University, Greece

ABSTRACT

In popular music information retrieval systems, users have the opportunity to tag musical objects to express their personal preferences, thus providing valuable insights about the formulation of user groups/communities. In this article, the authors focus on the analysis of social tagging data to reveal coherent groups characterized by their users, tags and music objects (e.g., songs and artists), which allows for the expression of discovered groups in a multi-aspect way. For each group, this study reveals the most prominent users, tags, and music objects using a generalization of the popular web-ranking concept in the social data domain. Experimenting with real data, the authors' results show that each Tag-Aware group corresponds to a specific music topic, and additionally, a three way ranking analysis is performed inside each group. Building Tag-Aware groups is crucial to offer ways to add structure in the unstructured nature of tags.

INTRODUCTION

Social tagging is the process of adding metadata by users in the form of keywords to annotate information items. In the case of music, the annotated items can be songs, artists, albums and playlists. Social tags are widely used, as high volume sources of descriptive metadata for music. Tags give the opportunity to a user to express his opinion, aiming to capture his personal view of resources he is interested in, including information about genre, mood, instrumentation and quality. A question is arising, why using tags is beneficial

for music information. The paramount motivations are, firstly building playlists by tagging songs and secondly, summarizing and categorizing a user profile using recommendation systems based on tags. This way, social tags have become important for Music Information Retrieval. In particular, tags can provide an insight to user behavior and language usage, e.g., how different is "rap" from "hip-hop". Additional benefits are grouping music items based on tags, finding social groups with shared interests, e.g. people that tag the same items or use the same tags, and generating user profiles from tagging behavior e.g. tag clouds based on tags applied by the user or representing his taste. Therefore, many music discovery and recommendation systems support the social tagging of music. According to Lamere (2008), social tags are used to help searching for items, exploring for new items, finding similar items, and finding other listeners with similar interests.

To assist in exploring and suggesting tags, some systems cluster similar tags together. These tag clusters can assist users in either applying tags or in identifying alternatives that may be useful in their search. Tag clustering has become an important issue for Music Information Retrieval systems. "Tag Radio" and "Tag Cloud" comprise two of the most popular applications of tag clustering. A music search engine can support a user-end tagging or labeling of artists, albums, and songs to create a site-wide folksonomy of music. Users can browse via tags, but the important benefit is the tag radio, permitting users to play music that has been tagged a certain way. The tagging process can be performed by genre "garage rock", mood "chill", artist characteristic "baritone", or any other form of user-defined classification "seen live". Listeners can search for artists or songs that have been tagged with a particular tag, or they can tune into "Tag Radio" where they listen to music that has been tagged with a particular tag.

On the other hand, it has been proved that sometimes tag clustering is not strong enough by itself to satisfy the listeners' requirements. Often users need to obtain results in a ranked list, where the position of each result should be taken into account. A question arising is why ranking is important in Music Information Retrieval. In the more general field of Information Retrieval, ranking of query/ search answers has become mandatory for internet searches. When the answers of a query or search are varying in quality and are large in numbers, it is necessary to rank/order these answers based on some criteria. From a users' viewpoint, ranking is extremely useful especially when associated with the retrieval of a few (top-k) answers. This way, ranking has also become an important issue in music search engines, a basic application of Music Information Retrieval on the internet.

A premature form of ranking in music search engines can be found. For example, people often run upon tag clouds, while visiting music search engines, e.g., Last.fm. A particular song, artist or user can be described by a tag cloud. For example, if a tag cloud describes a certain artist, by clicking on a particular tag in the tag cloud, all artists that have been frequently tagged with that tag will be displayed. In particular, we note that larger fonts in such displays indicate most popular tags, like in Figure 1, where the font size is scaled by the tag's weight (popularity or rank).

Tag clouds consist of collective information from users and provide a visual illustration for these tags. They effectively provide a good overview about a certain music topic like "The Beatles" in Figure 1, and describe it in instant. However, although tag clouds virtually hide the notions of users and music objects (such as songs, artists, albums etc), these two dimensions, e.g., users and music objects are visually missing from a (regular) tag cloud and simultaneously the respective ranking is also omitted. Therefore, from a social music tagging system there is a demand for a more powerful grouping and representation of information based on these three entities (i.e., tags, users, music objects), in which also a ranked list should be provided. Instead of just using tags, here we propose an enriched representation of the three

Related Content

Archive Film Comparison
Maia Zaharieva, Matthias Zeppelzauer, Dalibor Mitrovicand Christian Breiteneder (2010). *International Journal of Multimedia Data Engineering and Management (pp. 41-56).*
www.irma-international.org/article/archive-film-comparison/45754

Multispectral Image Compression, Intelligent Analysis, and Hierarchical Search in Image Databases
Stuart Rubin, Roumen Kountchev, Mariofanna Milanovaand Roumiana Kountcheva (2012). *International Journal of Multimedia Data Engineering and Management (pp. 1-30).*
www.irma-international.org/article/multispectral-image-compression-intelligent-analysis/75454

Electronic Games Improve Adult Learning in Diverse Populations
Robert D. Tennysonand Robert L. Jorczak (2011). *Gaming and Simulations: Concepts, Methodologies, Tools and Applications (pp. 1495-1512).*
www.irma-international.org/chapter/electronic-games-improve-adult-learning/49462

Knowledge-Building through Collaborative Web-Based Learning Community or Ecology in Education
Percy Kwok (2009). *Encyclopedia of Multimedia Technology and Networking, Second Edition (pp. 821-828).*
www.irma-international.org/chapter/knowledge-building-through-collaborative-web/17486

Navigating Through Video Stories Using Clustering Sets
Sheila M. Pinto-Cáceres, Jurandy Almeida, Vânia P. A. Neris, M. Cecília C. Baranauskas, Neucimar J. Leiteand Ricardo da S. Torres (2011). *International Journal of Multimedia Data Engineering and Management (pp. 1-20).*
www.irma-international.org/article/navigating-through-video-stories-using/58048