Large-scale Music Tag Recommendation with Explicit Multiple Attributes

Zhendong Zhao, Xinxi Wang, Qiaoliang Xiang,
Andy M Sarroff, Zhonghua Li, Ye Wang
School of Computing
National University of Singapore
{zhaozhendong, andy660, qiaoliangxiang}@gmail.com
andy.sarroff@nyu.edu, {lizhongh, wangye}@comp.nus.edu.sg

ABSTRACT

Social tagging can provide rich semantic information for large-scale retrieval in music discovery. Such collaborative intelligence, however, also generates a high degree of tags unhelpful to discovery, some of which obfuscate critical information. Towards addressing these shortcomings, tag recommendation for more robust music discovery is an emerging topic of significance for researchers. However, current methods do not consider diversity of music attributes, often using simple heuristics such as tag frequency for filtering out irrelevant tags. Music attributes encompass any number of perceived dimensions, for instance vocalness, genre, and instrumentation. Many of these are underrepresented by current tag recommenders. We propose a scheme for tag recommendation using Explicit Multiple Attributes based on tag semantic similarity and music content. In our approach, the attribute space is explicitly constrained at the outset to a set that minimizes semantic loss and tag noise, while ensuring attribute diversity. Once the user uploads or browses a song, the system recommends a list of relevant tags in each attribute independently. To the best of our knowledge, this is the first method to consider Explicit Multiple Attributes for tag recommendation. Our system is designed for large-scale deployment, on the order of millions of objects. For processing large-scale music data sets, we design parallel algorithms based on the MapReduce framework to perform large-scale music content and social tag analysis, train a model, and compute tag similarity. We evaluate our tag recommendation system on CAL-500 and a large-scale data set ($N = 77,448$ songs) generated by crawling Youtube and Last.fm. Our results indicate that our proposed method is both effective for recommending attribute-diverse relevant tags and efficient at scalable processing.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing; H.5.5 [**Sound and Music Computing**]: Systems

General Terms

Algorithms, Experimentation, Performance

Keywords

Music, tag recommendation, explicit multiple attributes, search

1. INTRODUCTION

In just over a decade, online music distribution services have proliferated, giving music a ubiquitous presence on the Internet. As the availability of online music continues to expand, it becomes imperative to have effective methods that allow humans to satisfactorily explore a large-scale space of mixed content. This is a significant challenge, as there is no predefined universal organization of online multimedia content and because of the well-known semantic gap between human beings and computers, in which computers cannot interpret human meaning with high accuracy. For example, a human may search for a song with the primary keywords, "happy," "Beatles," and "guitar." A human intuitively understands that 'happy' is a common human emotion, "Beatles" is a popular rock band from the 1960's, and "guitar" is a 6-stringed instrument. Yet it is difficult to computationally interpret these words with high semantic accuracy.

Social tagging has gained recent popularity for labeling photos, songs and video clips. Internet users leverage tags found on social websites such as Flickr, Last.fm, and Youtube to help bridge the semantic gap. Because tags are usually generated by humans, they may be semantically robust for describing multimedia items and therefore helpful for discovering new content. However, because they are often generated without constraint, tags can also exhibit significant redundancy, irrelevancy, and noise.

In order to address the deficiencies of socially collaborative tagging, computer based tag recommendation has recently emerged as a significant research topic. Current recommendation systems rely on term frequency metrics to calculate tag importance. However, some attributes of online content are tagged less frequently, leading to attribute sparsity. For instance, music encompasses a high-dimensional space of perceived dimensions, including attributes such as vocalness, genre, and instrumentation. Yet many of these are relatively underrepresented by social tagging. For example, the four most popular tags associated with the musician Kenny G on Last.fm are "saxophone," "smooth jazz," "instrumental jazz," and "easy listening," which are *Instrument* and multiple *Genre* attributes. Thus, three out of the four most popular Kenny G attributes are related to genre. According to [2], *Genre* tags represent 68% of all tags found on Last.fm. Most of the remaining attributes