Expert Systems with Applications 41 (2014) 2608–2618

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Music recommendation using text analysis on song requests to radio stations

Ziwon Hyung, Kibeom Lee, Kyogu Lee *

Music and Audio Research Group, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Music recommendation
Text mining
Song-document association
Latent Semantic Analysis
Probabilistic Latent Semantic Analysis

ABSTRACT

Recommending appropriate music to users has always been a difficult task. In this paper, we propose a novel method in recommending music by analyzing the textual input of users. To this end, we mine a large corpus of documents from a Korean radio station's online bulletin board. Each document, written by the listener, is composed of a song request associated with a brief, personal story. We assume that such stories are closely related with the background of the song requests and thus, our system performs text analysis to recommend songs that were requested from other similar stories. We evaluate our system using conventional metrics along with a user evaluation test. Results show that there is close correlation between document similarity and song similarity, indicating the potential of using text as a source to recommending music.

1. Introduction

Lower costs of hardware and advances in technology has led to a rapid growth in digital music. However, the vast amount of music available has made it more difficult for users to find the music they enjoy, known as the Paradox of Choice (Schwartz, 2009). Now, it has become more important than ever to filter out music that are relevant to each user. Thus, recommender systems have become invaluable due to their ability to offer only the relevant items from the enormous amount of accessible data (Hu, Koren, & Volinsky, 2008). Although there are various approaches for recommender systems, they can largely be generalized into two categories depending on how the items are analyzed: (1) collaborative filtering-based recommenders and (2) content-based recommenders.

Collaborative filtering-based music recommender systems find similar users or items, using information such as purchase histories and item ratings, to recommend music to users. For this approach, it is important for the items to have sufficient information, such as ratings and reviews, provided by the users. The lack of such information is a major problem, known as the 'Cold-Start Problem', in which the recommender system cannot process new items correctly since there is not enough, if any, information to analyze. Another problem with collaborative filtering methods is the poor diversity of recommended items (Zhang & Hurley, 2008). This problem stems from the Long Tail phenomenon (Anderson, 2008), in which the majority of users consume the very few, popular items while the rest of the users consume the less popular items. Celma showed that the music industry follows the Long Tail phenomenon and since collaborative filtering methods utilizes the preferences of users to generate recommendation, the system is only able to provide recommendations from a pool of generally popular items (Celma & Cano, 2008).

Content-based music recommender systems analyze the low-level acoustic features or metadata such as information on genres, artists, and lyrics to find similar items (Bogdanov et al., 2012). Unlike the collaborative filtering methods, this approach does not face the cold-start problem or popularity bias. However, content-based approaches that use acoustic features require a significant amount of computational power to analyze the content. Thus, with digital music libraries of modern sizes, such approaches are less efficient for commercial use compared to collaborative filtering methods.

Since both methods proved to have their own drawbacks, there have been several approaches in combining these two methods. Lu and Tseng proposed a hybrid music recommendation system by combining content-based, collaboration-based, and emotion-based methods (Lu & Tseng, 2009). They differed the weights assigned to each different methods based on the user's listening behavior. Such hybrid methods have shown to overcome some of the problems in each recommendation methods.

However, a common problem found in the methods explained above is the absence of taking into account the user's situation or background when searching for music to listen to. Hyung et al. introduced a recommender system that performs Latent Semantic Analysis (LSA) on textual information conveying the user's current situation to extract the context and recommend the appropriate music (Hyung, Lee, & Lee, 2012). Hyung and Lee also showed that

* Corresponding author. Tel.: +82 31 888 9139.
E-mail addresses: ziotoss@snu.ac.kr (Z. Hyung), kiblee@snu.ac.kr (K. Lee), kglee@snu.ac.kr (K. Lee).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.10.035

Probabilistic Latent Semantic Analysis (pLSA) can be used for text analysis (Hyung & Lee, 2013).

The concept of the proposed music recommender is shown in Fig. 1. Users write short letters to the radio program, where they describe their personal stories followed by song requests. The idea of the proposed system was to perform LSA or pLSA on these letters to search for similar letters. The hypothesis was that letters with common song requests will likely have similar situations or contexts. Given a query document, the proposed system should be able to recommend relevant songs by searching for similar letters and using the associated song requests. From here on, we use the term 'document' to reference such letters written to radio stations.

While our previous works showed possibility in using documents written by the users to extract contextual information when recommending music, they lacked comparison between the two text analysis algorithms and the data set used was small. A more crucial problem is that our previous works did not prove our hypothesis that people in similar situation will prefer similar music. In this paper, we perform a thorough experiment using various controllable parameters of LSA and pLSA algorithms along with several datasets to statistically compare the performance of the two text analysis algorithms. Moreover, we perform a user evaluation test to qualitatively prove our hypothesis mentioned above.

The remainder of the paper is organized as follows. In the next section, we discuss several context-based music recommender systems. We also introduce characteristics of the documents written to radio stations in Korea. In Section 3, we go over our proposed system in detail. In Section 4, we provide a statistical evaluation of the proposed system and in Section 5 we present the results of the evaluation. We conclude the paper with a summary and directions for future work in Section 6.

2. Background

Compared to works in collaborative filtering-based algorithms and content-based algorithms, research in context-based algorithms is relatively novel and is starting to attract attention in the research field of music recommendation. In this section we discuss several context-based algorithms and introduce the characteristics of Korean radio broadcasting structure which is important in understanding the proposed system.

2.1. Context-aware recommendation systems

The necessity to use users' contextual information was introduced by Reynolds et al.. Through a survey, they showed that user's activity had a significant impact on the mood of the listeners. Their research showed that the activity in which one was involved in had an impact on the choice of the music one wants to listen to Reynolds, Barry, Burke, and Coyle (2008).

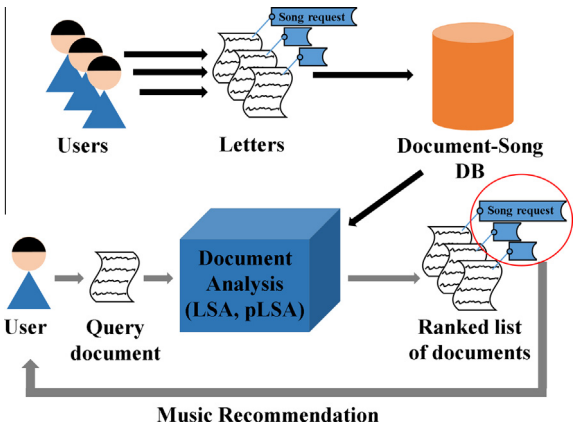

Fig. 1. Concept of the proposed music recommender.

Liu et al. used the time parameter when recommending music (Liu, Hsieh, & Tsai, 2010). They extracted the necessary features from both the symbolic and wave forms of the music files and by using the time parameter as an additional feature, they proposed a context aware music recommender system. Their result proved that the time parameter had a major effect in user's music listening behavior. However, their approach in using context information was limited to time only.

Su et al. proposed a music recommender system that used diverse contextual information, such as heartbeat, body temperature, air temperature, noise volume, humidity, light, motion, time, season, and location (Su, Yeh, Yu, & Tseng, 2010). This information was combined with content analysis on the music data to build a pattern database, which linked music with the users. The recommender system used the links and contextual information to provide more effective recommendation lists. However, generating the pattern database via content analysis confronted scalability issues, which is why it was preprocessed offline. A more critical problem is that the system has to infer the user's activity state using the available contextual information and does not know the activity accurately.

Another important context information is emotion. There have been several attempts to utilize emotion when recommending music. Shan et al. proposed a novel emotion-based music recommendation system where they build a music emotion model by an affinity discovery from film music (Shan, Kuo, Chiang, & Lee, 2009). Once the model is built, they perform Mixed Media Graph, an affinity graph algorithm, to discover the affinities between music features and emotions. However, their system uses 15 types of emotions which we think is not sufficient enough. Not every users will be able to express their context using only these 15 emotional states. Han et al. proposed a context-based music recommendation system that utilized an emotion-state transition model (ESTM), which acted as a bridge between the context of the users and low-level music features (Han, Rho, Jun, & Hwang, 2010). The authors proposed a system where the user selected a situation from a predefined situation list. The selected situation was used together with other features such as genre preference, favorite artist, age, occupation, hobby, etc., to collect the user's overall contextual information. Once the low-level features of music were extracted, their system recommended music according to the user's contextual information using the ESTM. A limitation to this work was that the situations were predefined and it was not possible to insert a new situation when the predefined situations did not fit the user.

Hariri et al. suggested to mine social tags from the most recent sequence of songs liked by the user to infer the contextual information when recommending music. The authors mined social tags from last.fm and instead of using the tags as features, they performed Latent Drichlet Allocation (LDA) topic modeling to determine the latent topics for each song. By performing LDA, the dimensions of the feature space were reduced and noisy data was removed for a better analysis. Using this method, the system inferred the user's current context by using the topics extracted from the user's most recent music playlist. Similar to the work done by Su et al., the system inferred the user's activity state and thus was prone to errors in predicting the actual context.

In order to overcome the problem of inferring the user's context indirectly, we propose a recommender system that uses text documents written explicitly by users that indicate their situation. These documents contain the background for requesting songs that, we believe, can be interpreted as valid contexts.

Since our method explicitly uses the contextual information conveyed in a document provided by the users, we expect that our system, compared to other recommender systems, will more accurately recommend music that is appropriate for the user's