Knowledge-Based Systems 88 (2015) 134–143

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Next-song recommendation with temporal dynamics

Ke Ji [a,b], Runyuan Sun [a,b,*], Wenhao Shu [c], Xiang Li [d]

[a] School of Information Science and Engineering, University of Jinan, Jinan, China
[b] Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan, China
[c] School of Information Engineering, East China Jiaotong University, Nanchang, China
[d] Psychology Department of Renmin University, Beijing, China

ARTICLE INFO

Article history:
Received 13 April 2015
Received in revised form 2 July 2015
Accepted 31 July 2015
Available online 6 August 2015

Keywords:
Music recommendation
Music playlist
Markov embedding
Temporal dynamics
Sequence prediction

ABSTRACT

Music recommendation has become an important way to reduce users' burden in discovering songs that meet their interest from a large-scale online music site. Compared with general behavior, user listening behavior has a very strong time dependence in that users frequently change their music interest in different sessions, where the concept of a "session" is that of a single user continuously listening songs over a period of time. However, most existing methods ignore temporal dynamics of both users and songs across sessions. In this paper, we analyze the temporal characters of a real music dataset from Last.fm and propose Time-based Markov Embedding (TME), a next-song recommendation model via Latent Markov Embedding, which boost the recommendation performance by leveraging temporal information. Specifically, we consider a scenario where user music interest is affected by long-term, short-term and session-term effects. By capturing temporal dynamics in the three effects, our model can track the change of user interest over time. We have conducted experiments on Last.fm dataset. Results demonstrate that with our time-based model, the recommendation accuracy is significantly improved compared to other state-of-the-art methods.

1. Introduction

Music plays an important role in our daily lives, having become one of the most popular entertainment forms. The rapid development of the Internet brings great convenience to people's musical life, largely out of expensive price and storage limit of traditional records and CDs. Especially, with growing use of the mobile Internet and applications, people can surf a lot of online music sites (e.g., Last.fm,[1] Pandora[2]) anytime and anywhere for enjoying music. But as time goes on, the music sites have a sharp increase in the number of songs which makes it hard for users to seek out songs that meet their interest from massive amounts of songs [1]. So music recommendation system [2] that can automatically generate the most interesting songs aimed at different users' interest is particularly important for online music providers to improve the quality of the user services and promote the development of music.

Music recommendation is a playlist prediction problem that generates a list of songs which is related to the given songs of a session. Because of the great commercial demand, there have been many studies on music recommendation [3], which can be classified into three categories: *content-based* methods [4–6], *context-based* methods [7–9] and *collaborative filtering* [10–12]. Furthermore, some hybrid approaches [13–20] have been proposed for merging these methods in order to overcome the shortcomings of single model. Most of these methods consider playlist prediction as an information retrieval problem, while ignoring the sequential nature of playlists.

Sequence prediction is a novel emerged research area of music recommendation in recent years. It reduces the playlist prediction to a sequence modeling setting that is more concerned with the order of songs and the transitions between two adjacent songs in the playlists. Since it is similar to Markov chain approach in natural language processing, some Markov chain-based approaches [21–23] are incorporated to solve the sequence prediction problem. Recent works on Markov embedding of multi-dimensional latent space have been done [24,25]. Experiments show that Markov embedding models substantially outperform traditional n-gram Markov model.

Although these methods address playlist prediction to a large extent, they ignore temporal information. Compared with general recommendation scenario, strong time dependence is a distinctive

* Corresponding author at: School of Information Science and Engineering, University of Jinan, Jinan, China.
E-mail address: sunry@ujn.edu.cn (R. Sun).

[1] http://www.last.fm.
[2] http://www.pandora.com.

http://dx.doi.org/10.1016/j.knosys.2015.07.039
0950-7051/

property of music recommendation. For a song, popularity can affect user listening behaviors. For instance, the song "My Heart Will Go On" can earn many more listeners on the 100th anniversary of the most famous accident of Titanic, but cannot get so many mentions on normal days. For a user, interest goes through concept drifting [26,27]. For instance, it is common for users to have rise and decay of music interest instead of the fixed interest. For a session, various reasons lead to its formation, not just in terms of user interest. For instance, some songs in a session do not meet user previous interest, but he still listens to them because of curiosity or influence of other activities. These temporal changes bring unique challenges in music recommendation. So we require a dynamic model, which can track the changes of user interest, and make dynamic adjustments to the prediction model, so as to provide real-time music recommendation.

To address the above challenges, we propose Time-based Markow Embedding (TME), a dynamic next-song recommendation model via Latent Markov Embedding. In addition to the embedded users and songs [25], we also embed sessions into a Euclidean space. Specifically, our whole embedding considers three aspects: distance between users and songs (*long-term effect*) reflects user global interest; distance between songs and songs (*short-term effect*) reflects the sudden song transitions that have low impact on long-term; distance between sessions and song (*session-term effect*) reflects user local interest in each session. By modeling temporal dynamics in the three effects, we extend the effects into the time changing effects. Our model is designed to integrate the time changing effects, and is able to track the change of user interest over time. We have conducted experiments on a real music dataset from Last.fm. The empirical results and analysis demonstrate that our model outperforms other state-of-the-art Markov chain-based methods.

The rest of this paper is organized as follows. We first give an overview of related work in Section 2. Then, we present the preliminaries, including data temporal analysis and problem setting in Section 3. We describe our model in detail in Section 3. We report the experimental results and analysis in Section 4. Finally, we provide a conclusion in Section 5.

2. Related work

Music recommendation has been a fast growing research field of recommender systems. There are several traditional approaches to music recommendation [3].

- ***Content-based*** [4,5] methods extract features directly from the audio signal, compute the similarity among songs based on the features, and recommend songs with high similarity to users. A overview of *content-based* methods is given by [6].
- ***Context-based*** [7–9] methods take context information (e.g., location, emotion and so on) into account. The usage of context information is crucial in music systems because people prefer different songs for different activities such as working, sleeping and studying.
- ***Collaborative filtering*** [12] methods use explicit feedback or implicit feedback to track user listening habits, which need not any additional information. It can be grouped into two general classes: *memory-based*[28] and *model-based* [29].

Content-based methods need audio content analysis that increases computation cost, and cannot achieve rapid prediction in large-scale systems. When context information is not available, *context-based* methods are inapplicable. *Collaborative filtering* methods suffer from some problems, such as cold start and sparsity. It is proved by practice [13,9] that*collaborative filtering* outperforms *content-based* and Context-based on the rating type of data. What's worth mentioning is that *collaborative filtering*methods have been widely used because of their simple practicality, easy to achieve, high accuracy and so on. Some hybrid approaches [13–20] have been proposed to avoid certain limitations of single method. In general, these methods take music recommendation as a similar retrieval problem.

However, unlike traditional recommender system, there are many unique historical playlists existing in user sessions. Every playlist is a sequence of songs where the order of songs imply user's sequential listening behavior. Recently, sequence prediction has become one of the research emphases in music recommendation. In contrast to previous methods, sequence prediction can be seem as a language modeling setting. Markov chain has been incorporated to solve the problem. For example, smoothed n-gram model [21] is the most common method for machine translation and automatic spelling correction. In particular, *Bigram* is a first-order Markov model where transition probabilities are estimated directly for every pair of songs. McFee et al. [22] propose a playlist generation algorithm which characterize playlist as generative models of strings of songs belonging to some unknown language. Chen et al. [24] proposed a new family of learning algorithms *LME* which formulate playlist prediction as a regularized maximum-likelihood embedding of Markov chains in Euclidian space. Wu et al. [25] proposed a personalized next-song recommendation algorithm *PME*, extending *LME* by embedding users into the Euclidian space.

Over the extensive testing in [24,25], the embedding-based models-*LME* and *PME* outperform the statistics-based n-gram models. Often, as time goes on, user interest is evolving. Some interesting activities that happen to users, songs or sessions may lead to a temporary change of user interest. But these methods stay fixed after the model training, and cannot be continuously updated to reflect user present interest. The goal of our work is to track the changes of user interest by leveraging temporal dynamics.

3. Preliminaries

3.1. Data temporal analysis

In this paper, the time drifting data we use is the log of Last.fm and Table 2 shows the basic statistics of the dataset. Based on the temporal information existing in the dataset, we analysis temporal dynamics in user, song and session.

User. Fig. 1(a) shows the statistics of a selected user's changes in listening behavior since his first listening event. The songs that have been listened by the user are randomly divided into four groups. We set a 1-week coarse time resolution, which is sufficient for the rather slow changing listening behavior in practice. For each group of songs, we calculate the number of times the songs listened by the user, and have found that the user pays different attention to them over time based on the four curves that sometimes smooth, sometimes bumpy. This indicates user global interest would change over time.

Song. Fig. 1(b) shows the statistics of a selected song-the number of times the song was listened and the number of the song's listeners since its first listened event. We set a 2-week coarse time resolution, which is sufficient for the rather slow changing song popularity in practice. We have found that the number of times first increases fast, and slowly declines, and then suddenly increases, which also appear for other songs again and again. The number of listeners has a similar trend, but less than the number of times. This indicates song popularity can reflect user current behavioral tendency.