ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/320184580

HOMMIT: A Sequential Recommendation for Modeling Interest-Transferring via High-Order Markov Model

Conference Paper *in* Lecture Notes in Computer Science · October 2017

DOI: 10.1007/978-3-319-68786-5_30

CITATIONS
0

READS
99

5 authors, including:

Yang Xu
Shandong Normal University
13 PUBLICATIONS **102** CITATIONS

SEE PROFILE

Yupeng Hu
Shandong University
32 PUBLICATIONS **182** CITATIONS

SEE PROFILE

All content following this page was uploaded by Yang Xu on 22 October 2019.

The user has requested enhancement of the downloaded file.

HOMMIT: A Sequential Recommendation for Modeling Interest-Transferring via High-Order Markov Model

Yang Xu, Xiaoguang Hong(✉), Zhaohui Peng, Yupeng Hu, and Guang Yang

School of Computer Science and Technology, Shandong University, Jinan, People's Republic of China
{xuyang0211, hxg, pzh, huyupeng}@sdu.edu.cn, loggyt@yeah.net

Abstract. Capturing user interest accurately is a key task for predicting personalized sequential action in recommender systems. Through preliminary investigation, we find that user interest is stable in short term, while changeable in long term. The user interest changes significantly during the interaction with the system, and the duration of a particular interest and the frequency of transition are also personalized. Based on this finding, a recommendation framework called HOMMIT is proposed, which can identify user interests and adapt an improved high-order Markov chain method to model the dynamic transition process of user interests. It can predict the transition trends of user interest and make personalized sequential recommendation. We evaluate and compare multiple implementations of our framework on two large, real-world datasets. The experiments are conducted to prove the high accuracy of our proposed sequential recommendation framework, which verified the importance of considering interest-transferring in recommendations.

Keywords: Interest modeling · Recommender system · Markov model

1 Introduction

Personalized recommendation systems have been widely used in the case of the rapid expansion of Internet information and become indispensable since they alleviate the information overload, by providing users with personalized information, products or service to satisfy their tastes and preference. Modeling the dynamic evolution process of user interests is the core task of a personalized recommendation system.

In this paper, we are interested in modeling the dynamic evolution process of user interests from sequential behavioral data (e.g. a user's browsing history) to predict user actions such as the next movie to watch, product to purchase, or music to listen. Existing user interest models usually use monotonically decreasing function (e.g. linear function [9], exponential function [2, 7], or forgetting function [6, 8]) to calculate timeliness of user behavior. However, the evolution of user interest is a non-monotonic process which contains rising stage and declining stage [5]. The declining of the last interest will always be accompanied by the growth of new interests. For example, if

© Springer International Publishing AG 2017
A. Bouguettaya et al. (Eds.): WISE 2017, Part II, LNCS 10570, pp. 372–386, 2017.
DOI: 10.1007/978-3-319-68786-5_30

someone has been eating MacDonald's almost every meal recently, but now he is going to have some other food considering the high calorie of hamburgers, which proves that there is no need to make the user's interest of the last action the highest.

Our preliminary investigation shows that user interests have two characteristics. First, user interest changes significantly in the process of user's interaction with the system. There are many reasons for this phenomenon, such as being tired of the old interest, changes of the external environment, and the excitement for new things. Second, the duration of a particular interest and the change frequency are personalized. Some of the users kept a clear interest for a long time before the interest changes, while some others change their interests frequently. Due to these two characteristics, existing interest models cannot effectively describe the dynamic transition process of user interests.

To model interest-transferring and make recommender system be more "human-minded", we propose a new recommendation framework called HOMMIT. Under HOMMIT, we can effectively identify user interests from sequential behavioral data and adapt the high-order Markov chain method to model the dynamic transition process of user interests, and then we can predict the transition trends of user interest and make personalized sequential recommendation. Figure 1 demonstrates an example of how HOMMIT makes recommendations.

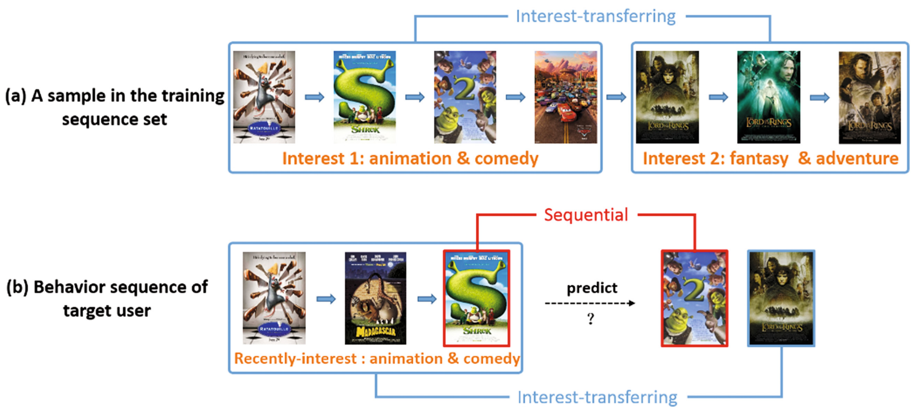

Fig. 1. An example of how HOMMIT makes recommendations

In the example of Fig. 1, (a) is a sample in the training sequence set. HOMMIT learns the sequential patterns and interest-transferring model from the training sequence set. *Shrek 2* and *The Lord of the Rings 1* are recommended to the target user because (1) *Shrek 2* frequently follows the recently-watched movie *Shrek 1*, and (2) the interest '*fantasy & adventure*' frequently follows the recently-interest '*animation & comedy*', and compared with the other movies of the interest '*fantasy & adventure*', *The Lord of the Rings 1* has strong sequential relationship with the recent action of the target user. From this example we can observe that *The Lord of the Rings 1* is still recommended to target user, although it does not seem to match the user's current preferences. This flexibility enables HOMMIT to improve the diversity and accuracy of recommender system and guide users to discover their potential interests. To summarize, the major contributions of this paper are as follows:

- In this paper, we consider the dynamic transition process of user interest which previous studies have rarely mentioned. A novel recommendation framework named HOMMIT for modeling the dynamic transition process of user interest is proposed. HOMMIT can predict the transition trends of user interest and make personalized sequential recommendation.
- In HOMMIT, we design a user interest identification method and an improved high-order Markov chain method to model the dynamic transition process of user interests.
- We systematically compare the proposed HOMMIT approach with other algorithms on real-world datasets. The results confirm that our new method significantly improves the effectiveness of recommendation.

The rest of this paper is organized as follows. Section 2 reviews the related works on sequential recommendation and user interest modeling. Section 3 presents some notations and the problem formulation, and Sect. 4 introduces our novel algorithm for interest identification. In Sect. 5, we details the personalized interest-transferring model and the recommendation approach. Experiments and discussions are given in Sect. 6. Conclusions are drawn in Sect. 7.

2 Related Work

In this section, we discuss existing research related to our work, including sequential recommendation and interest model based on item recommendation.

Sequential recommendation. Existing works on sequential recommendation mostly utilize the Markov chain to predict a user's next action. Rendle [10] combined the matrix factorization model on modeling personal preference and Markov Chains on modeling sequential patterns. Ruining [11] proposed Fossil which fuses similarity-based models with Markov Chains to predict sequential behavior. Cheng et al. [12] exploited the first-order Markov chain in the check-in sequence to make next-visit recommendation where users' next choice is considered to be only relevant to the last location. Instead of applying Markov Chain on user action sequence, we utilize improved high-order Markov chain to extract interest transition pattern from user interest sequence. On the whole, the main difference between our work and the previous approaches is the utilization of both 'action-level' and 'interest-level' Markov chains which can predict the transition trends of user interest and make personalized sequential recommendation.

Interest model based recommendation. User interest has a strong timeliness in recommender systems. Modeling user interest have been studied by several researchers. Yin [13] applied linear function on modeling the time decay of user interest. Xu [5] proposed a recommendation framework named SimIUC, which can identify multiple user interests and model the dynamic evolution process of user interests with the inverted-U-curves. Chen [6, 8] utilized the memory forgetting curve to model the human interest-forgetting curve, and integrated the interest-forgetting mechanism with Markov model to make item recommendations. Ding [2] proposed TItemKNN which is a time weighted item-based collaborative filtering method by reducing the influence of out-of-date information to allow for high performance when predicting a user's next action. These recommender systems only model the time decay of user interests and do not take interest-transferring into account. In this paper, we focus on discovering user interest transition patterns and predicting the trend of interest evolution by mining sequential behavioral data.

3 Preliminaries

We first introduce some notations in HOMMIT recommendation framework and then give a formal statement of the interest-transferring based sequential prediction problem.

3.1 Notations

Let $\mathcal{U} = \left\{u_1, u_2, \ldots, u_{|\mathcal{U}|}\right\}$ denote the set of users in the system and $\mathcal{I} = \left\{i_1, i_2, \ldots, i_{|\mathcal{I}|}\right\}$ be the set of items (e.g. movies, music, or books) in the given data set. We let $\mathbb{X} = \left\{x_1, x_2, \ldots, x_{|\mathbb{X}|}\right\}$ denote the set of global interests which were identified from user sequential action data S. Each user u is associated with a sequence of actions (e.g. movies watched by u) $S_u = \left\{S_1^u, S_2^u, \ldots, S_{|S_u|}^u\right\}$, where $S_k^u \in \mathcal{I}$, and a sequence of interests $X_u = \left\{x_1^u, x_2^u, \ldots, x_{|X_u|}^u\right\}$, where $x_k^u \in \mathbb{X}$. $S_u^{'} = \left\{S_1^{u'}, S_2^{u'}, \ldots, S_{|S_u^{'}|}^{u'}\right\}$ is the processed action sequence of user u treated by filtering out noise and labeling interest. Each action in $S_u^{'}$ is labeled with a specific global interest. Let $ROI_x = \{i_1, i_2, \ldots, i_{|ROI_x|}\}$ denote the distinct set of actions (or "items") which were labeled with global interest x in the sequential action data S where the sequential signal is ignored, and we call it as the region of interest x. $ROI = \{ROI_{x_1}, ROI_{x_2}, \ldots, ROI_{x_{|\mathbb{X}|}}\}$ represents the set of ROI_x, where $x \in \mathbb{X}$.

Table 1. Notations

Notation	Description
$\mathcal{U}, \mathcal{I}$	User set, item set
$\mathbb{X}$	Global interest set
S	User sequential action data, $S = \{S_{u_1}, S_{u_2}, \ldots, S_{u_{|\mathcal{U}|}}\}$
S_u	Action sequence of user u, $S_u = \{s_1^u, s_2^u, \ldots, s_{|S_u|}^u\}$
$S_u^{'}$	Processed action sequence of user u
X_u	Interest sequence of user u, $X_u = \{x_1^u, x_2^u, \ldots, x_{|X_u|}^u\}$
X	The set of interest sequences of all users, $X = \left\{X_{u_1}, X_{u_2}, \ldots, X_{u_{|\mathcal{U}|}}\right\}$
ITM_u	Personalized interest-transferring model of user u

(*continued*)

Table 1. (*continued*)

Notation	Description
ROI_x	The region of interest x, $ROI_x = \{i_1, i_2, \ldots, i_{\lvert ROI_x \rvert}\}$
ROI	The set of ROI_x, where $x \in \mathbb{X}$
$P(x_i \lvert x_j)$	The one-step transition probability from interest $\mathbf{x}_j$ to interest $\boldsymbol{x}_i$
$P(i_a \lvert i_b)$	The one-step transition probability from item $\boldsymbol{i}_b$ to item $\boldsymbol{i}_a$
$\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$	The interest confidence of user $\boldsymbol{u}$ on interest $\boldsymbol{x}$ at time $\boldsymbol{t}$
$\mathcal{T}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$	The timeliness of user interest of user $\boldsymbol{u}$ on interest $\boldsymbol{x}$ at time $\boldsymbol{t}$

We adapt an improved high-order Markov chain method to model the dynamic transition process of user interests. ITM_u is the personalized interest-transferring model of user u. Notations used throughout this paper are summarized in Table 1.

3.2 Problem Formulation

Given users $\mathcal{U} = \{u_1, u_2, \ldots, u_{\lvert \mathcal{U} \rvert}\}$ and their sequential behavioral data S. For the target user $v (v \in \mathcal{U})$ with the action sequence $S_v = \{s_1^v, s_2^v, \ldots, s_{\lvert S_v \rvert}^v\}$, personalized sequential recommendation is to predict k items that target user might be attracted to in the near future. The most common personalized recommendation approaches focus on modeling timeliness of user interest, without considering the characteristics of interest transition. Our main task in this paper is modeling user interest-transferring and recommending items which not only satisfy users' currently-tastes, but also cater to the potential future interests of users.

In order to achieve this purpose, we propose a recommendation framework called HOMMIT and the main components of this framework are shown in Fig. 2. At first, we identify the global interests from user sequential action data. Each action in user action sequence is labeled with a specific global interest index. Then, we learn the personalized interest-transferring model of users on their interest sequences respectively. Last, given an action sequence of target user, HOMMIT will recommend the TOP-N items to the target user according to his/her interest-transferring model and sequential relationship. In the rest of this paper, we will introduce each component of HOMMIT in details.

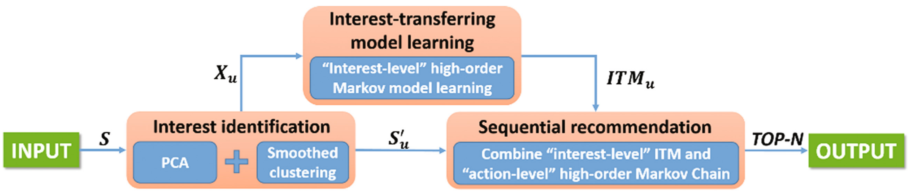

Fig. 2. Overview of the HOMMIT recommendation framework

4 Interest Identification

As discussed in the previous section, in order to model interest-transferring, the first phase of HOMMIT is to identify global interests from user sequential behavioral data. The key challenge is to discover global interests which are applicable to all users and exploit the discovered global interests to generate users' interest sequences. To address this challenge, we propose an interest identification algorithm based on PCA and smoothed clustering.

4.1 Principle Components Analysis

Principal component analysis (PCA) is a widely used statistical technique in unsupervised dimension reduction [14]. Let $S_u = \{s_1^u, s_2^u, \ldots, s_n^u\}$ denote the action sequence of user u, and each action s_i^u is represented with an m-dimensional binary feature vector $(a_{i1}^u, a_{i2}^u, \ldots, a_{ij}^u, \ldots, a_{im}^u)$, whose element a_{ij}^u is 1 if the action i contains feature j, otherwise, a_{ij}^u is 0. In this way, the action sequence of user u can be represented as an $|n| * |m|$ binary matrix A_u. Considering that this feature representation approach is of great dimension, we use Principal Components Analysis (PCA) to extract latent action features. The original data matrix is described as $A = (A_{u_1}^T A_{u_2}^T \cdots A_{u_{|\mathcal{U}|}}^T)^T$, and score matrix $P = (P_{u_1}^T P_{u_2}^T \cdots P_{u_{|\mathcal{U}|}}^T)^T$ is the representation of A in the principal component. Rows of P correspond to user sequential actions and columns correspond to latent action features.

4.2 Smoothed Clustering

Considering that user interest is stable in short term, while changeable in long term, we utilize central moving average method in each column of $P_{u_i} (i = 1, 2, \ldots, |\mathcal{U}|)$ to smooth out short-term fluctuations and highlight longer-term trends or cycles. The central moving average method is formulated as follows. Given a column of P_{u_i}, which is represented as a number series $C = (C_1, C_2, \ldots, C_l)$, and a fixed odd window size $(2n + 1)$, the main task of central moving average method is to generate a smoothed number series $SC = (SC_1, SC_2, \ldots, SC_l)$, where

$$SC_t = \begin{cases} SC_{n+1}, & 1 \le t \le n \\ \frac{1}{2n+1} \sum_{i=-n}^{n} C_{t+i}, & n < t \le l - n \\ SC_{l-n}, & l - n < t \le l \end{cases} \tag{1}$$

We represent the smoothed score matrix as $P' = \left(P_{u_1}^{\prime T} P_{u_2}^{\prime T} \cdots P_{u_{|\mathcal{U}|}}^{\prime T}\right)^T$. A part of the smoothing result of score matrix block P_{u_8} of the 8th user in MovieLens1M dataset [15] in 2-dimensional principle component space is shown in Fig. 3, where each action is represented as a corresponding type of point. There are three fragments of different types of user action in this example which were identified by the interest identification algorithm. In this paper, we consider the "type" as the user interest. Figure 3 shows that the clustering characteristics of user actions have become more obvious in smoothed score matrix and the same type of actions shows the significant sequential continuity. These properties will be propitious to the segmentation of action sequence in the following work.

We use k-means clustering to cluster the actions in groups based on smoothed score matrix P'. Each group is considered as a global interest and each action in user action sequence is labeled with the corresponding global interest index. The subsequence in which each action was labeled with the same interest index, is considered as an interest fragment. We filter out the interest fragments with length less than φ as noise. Then, we obtain the interest sequence X_u and action sequence with interest labeled S'_u for each user u.

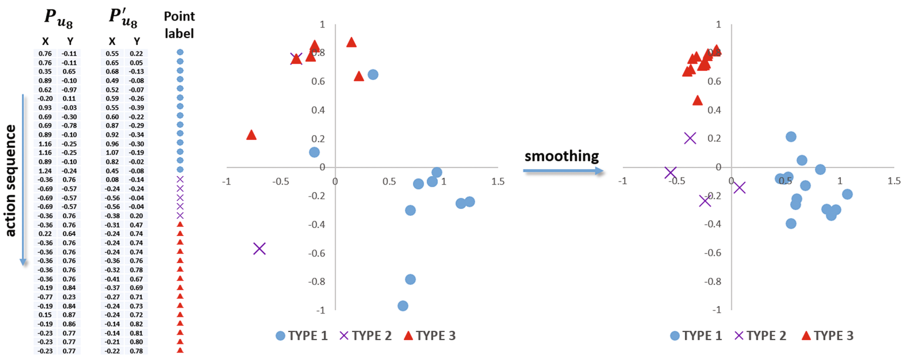

Fig. 3. 2-dimensional smoothing result

5 Interest-Transferring Model and Sequential Recommendation

5.1 Interest-Transferring Model

After interest identification, we improve the high-order Markov method proposed by Raftery [3] to learn the **I**nterest-**T**ransferring **M**odel (**ITM**) of users. We are tackling interest transition prediction tasks which are formulated as follows. Let $\mathbb{X} = \{x_1, x_2, \ldots, x_{|\mathbb{X}|}\}$ denote the set of global interests, and $X^t = \left\{X^t_{u_1}, X^t_{u_2}, \ldots, X^t_{u_{|\mathcal{U}|}}\right\}$ be the set of interest sequences of all users. Given an interest sequence of user u at time t: $X^t_u = \{x^u_1, x^u_2, \ldots, x^u_t\}$, where $x^u_k \in \mathbb{X}$. Our objective is to obtain the transition probability distribution of interest state of user u at time $t+1$. Under the k th-order Markov model setting, the transition probability of interest x at time $t+1$ is defined as:

$$\begin{aligned} P\left(x|x^u_t, x^u_{t-1}, \ldots, x^u_{t+1-k}\right) &= P\left(X^u_{t+1} = x|X^u_t = x^u_t, X^u_{t-1} = x^u_{t-1}, \ldots, X^u_{t+1-k} = x^u_{t+1-k}\right) \\ &= \sum_{i=1}^{k} \lambda^{u,t}_i P(x|x^u_{t+1-i}) \end{aligned} \tag{2}$$

where x_i^u is the interest owned by user u at time i, and X_i^u is a random variable which represents an arbitrary interest in $\mathbb{X}$. $P(x|x_{t+1-i})$ represents the one-step transition probability from interest x_{t+1-i} to interest x, and $\lambda_i^{u,t}$ is a balancing component on each one-step transition probability.

Let $X^t = \left\{X_{u_1}^t, X_{u_2}^t, \ldots, X_{u_{|\mathcal{U}|}}^t\right\}$ be the training set. Given an observing sequence X_u^t, the goal of ITM learning is to maximize the probability of predicting the last interest:

$$\operatorname{argmax}_{\Theta} \prod\nolimits_{X_u^t \in X^t} P\left(x|x_t^u, x_{t-1}^u, \ldots, x_{t+1-k}^u\right) \tag{3}$$

which is equivalent to minimizing the negative log-likelihood:

$$\begin{aligned} \operatorname{argmin}_{\Theta} \mathcal{L} &= -\sum\nolimits_{X_u^t \in X^t} \ln\left(P\left(x|x_t^u, x_{t-1}^u, \ldots, x_{t+1-k}^u\right)\right) \\ &= -\sum\nolimits_{X_u^t \in X^t} \ln\left(\sum_{i=1}^{k} \lambda_i^{u,t} P(x|x_{t+1-i}^u)\right) \\ \text{s.t.}\, 0 \le \lambda_i^{u,t}&, 0 \le P(x|x_{t+1-i}^u) \le 1. \end{aligned} \tag{4}$$

where Θ is the set of parameters in the personalized ITM. We utilize gradient descent method to seek the optimal Θ, then obtain the interest probability distributions at time $t+1$ based on Eq. (2).

One-Step Transition Probability. As described in Sect. 4, the user interests were identified from users' behavioral data globally. Therefore, the user interests under the personalized ITM should be global-related, and the one-step transition probabilities between two interests are supposed to be fixed. We define the one-step transition probability from interest x_j to interest x_i as:

$$P\left(x_i|x_j\right) = \frac{\sum_{X_u^t \in X^t} \mathbb{I}_{\{x_j, x_i\} \subseteq X_u^t}}{\sum_{X_u^t \in X^t} \mathbb{I}_{x_j \subseteq X_u^t}} \tag{5}$$

where $\mathbb{I}_{\{\cdot\} \subseteq X_u^t}$ is a counter function that returns the number of times that the subsequence $\{\cdot\}$ appears in X_u^t. This expression of one-step transition probability defines how often the interest x_i will be observed after the occurrence of interest x_j in all users' interest sequences.

Self-Transition Probability. In conventional Markov models, most transitions are from one state to another. However, this is not the case with self-transition in interest-transferring Markov model. In a self-transition, the source state and the target state are the same. Under the ITM, we consider the self-transition of user interest as interest-keeping. We use reappearance probability which defines how often the interest will reappear to estimate the self-transition probability approximately. Therefore, the self-transition probability of interest $\boldsymbol{x_i}$ is calculated as:

$$P(x_i|x_i) = \frac{\sum_{u \in \mathcal{U}} \mathbb{I}_{\{x_i, *, x_i\} \subseteq X_u^t}}{\sum_{u \in \mathcal{U}} \mathbb{I}_{x_i \subseteq X_u^t}} \quad (6)$$

where $\mathbb{I}_{\{\cdot\} \subseteq X_u^t}$ is a counting function as described above. '$*$' represents arbitrary interest subsequence without interest x_i. Similar to the one-step transition probabilities, the self-transition probabilities of interest are supposed to be fixed as well.

5.2 Personalized Interest-Transferring Model

So as to improve ITM to be more "human-minded" and personalized, we attempt to incorporate the timeliness of user interest and interest confidence into ITM. Let $\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ be the interest confidence of user $\boldsymbol{u}$ on interest $\boldsymbol{x}$ at time $\boldsymbol{t}$. The interest confidence reflects the strength of user interest and it is proportional to the length of interest fragment. In the interest identification method, the interest fragments with length less than 3 have been filtered out as noise for their low interest confidence. Let $\mathcal{T}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ be the timeliness of user interest of user $\boldsymbol{u}$ on interest $\boldsymbol{x}$ at time $\boldsymbol{t}$. The timeliness of user interest means that the older the interest occurs in the user interest sequence, the larger effectiveness on it will lose. Similar to [8], we combine ITM with timeliness of user interest $\mathcal{T}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ and interest confidence $\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ by defining the λ component in ITM as:

$$\lambda_i^{u,t} = \mathcal{C}_{\boldsymbol{x}_{t+1-i}^{\boldsymbol{u}}}^{\boldsymbol{u},\boldsymbol{t}} \mathcal{T}_{\boldsymbol{x}_{t+1-i}^{\boldsymbol{u}}}^{\boldsymbol{u},\boldsymbol{t}} \quad (7)$$

where $\boldsymbol{x}_{t+1-i}^{\boldsymbol{u}}$ is the interest of $\boldsymbol{u}$ at time $\boldsymbol{t}+1-\boldsymbol{i}$ $(1 \le \boldsymbol{i} \le \boldsymbol{k})$. Next, we will attempt to find proper mathematical expressions of $\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ and $\mathcal{T}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$.

Interest Confidence. The value of interest confidence $\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ reflects the interest strength of user $\boldsymbol{u}$ on interest $\boldsymbol{x}$ at time $\boldsymbol{t}$. The larger the value of $\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}}$ is, the higher effectiveness of $\boldsymbol{x}$ in the interest transition process at time $\boldsymbol{t}$. Intuitively, interest confidence of $\boldsymbol{x}$ should be proportional to the length of corresponding interest fragment. In this paper, we utilize exponential growth (EG) model and rational growth (RG) model to express the interest confidence respectively:

$$\textbf{EG}: \quad \mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}} = \frac{\alpha_u}{1 + \beta_u e^{-\gamma_u l_{u,t}(x)}} \quad (\alpha_u, \beta_u \ge 0, 1 \ge \gamma_u \ge 0) \quad (8)$$

$$\textbf{RG}: \quad \mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{u},\boldsymbol{t}} = 1 + \alpha_u l_{u,t}(x)^{\beta_u} \quad (\alpha_u \ge 0, 1 \ge \beta_u \ge 0) \quad (9)$$

where α_u, β_u, and γ_u are u's bounded personalized parameters. $l_{u,t}(x)$ represents the length of corresponding interest fragment of user $\boldsymbol{u}$'s interest $\boldsymbol{x}$ at time $\boldsymbol{t}$.

Timeliness of User Interest. The timeliness problem of information in recommender systems has been well studied and can be modeled with linear function [9], exponential function [2, 7], forgetting curve [6, 8], or inverted U curve [5]. Similar to interest confidence, in this paper, exponential decay (ED) model and rational decay (RD) model are put forward to express the timeliness of user interest:

$$\textbf{ED}: \quad \mathcal{T}_x^{u,t} = 1 + \alpha_u e^{-\beta_u f_{u,t}(x) + \gamma_u}, f_{u,t}(x) = t - t_{mid}^u(x) \ (\alpha_u, \gamma_u \geq 0, 1 \geq \beta_u \geq 0) \quad (10)$$

$$\textbf{RD}: \quad \mathcal{T}_{\boldsymbol{x}}^{\boldsymbol{u,t}} = \frac{\alpha_u}{\beta_u f_{u,t}(x) - \gamma_u}, f_{u,t}(x) = t - t_{mid}^u(x) \ (\alpha_u, \gamma_u \geq 0, 1 \geq \beta_u \geq 0) \quad (11)$$

where α_u, β_u, and γ_u are u's bounded personalized parameters. $t_{mid}^u(x)$ represents the time index of midpoint of corresponding interest fragment of user $\boldsymbol{u}$'s interest $\boldsymbol{x}$.

5.3 Personalized ITM Based Sequential Recommendation

In this section, we will introduce how to predict user next action based on the combination of action-leveled Markov chain and interest-leveled personalized ITM. Similar to $P(x_i|x_j)$, we define the one-step transition probability from item $\boldsymbol{i_b}$ to item $\boldsymbol{i_a}$ as:

$$P(i_a|i_b) = \frac{\sum_{S_u^t \in S^t} \mathbb{I}_{\{i_b, i_a\} \subseteq S_u^t}}{\sum_{S_u^t \in S^t} \mathbb{I}_{i_b \subseteq S_u^t}} \quad (12)$$

where $\mathbb{I}_{\{\cdot\} \subseteq S_u^t}$ is a counting function that returns the number of times the subsequence $\{\cdot\}$ appears in S_u^t. $P(i_a|i_b)$ expresses how often that item i_a will be observed after the occurrence of item i_b in all users' action sequences.

When predicting next action of the target user u at time t, our basic idea is to consider the items in user u's currently-interest fragment as the sources to be injected with user preference. Then we predict the directions of user u's currently-interest transfer via his/her personalized ITM, and user preference will be propagated to candidate items in the target ROIs. Let $\mathcal{P}^{u,t} = \{\wp_{x_1}^{u,t}, \wp_{x_2}^{u,t}, \ldots, \wp_{x_{|\mathbb{X}|}}^{u,t}\}$ represents the transition probability distribution of interest state of user u at time t. The computation method of the preference score of each candidate item i_c is described as formula (13).

$$P_u^t(i_c) = \sum_{i_k \in \mathbb{I}_u^t} \Big(\sum_{x_j \in \mathbb{X} - x_c} \mathrm{I}_{\{i_c \in ROI_{x_j}\}} (1-\omega) \wp_{x_j}^{u,t} P(i_c|i_k) + \mathrm{I}_{\{i_c \in ROI_{x_c}\}} \omega \wp_{x_c}^{u,t} P(i_c|i_k) \Big) \quad (13)$$

where $\mathrm{I}_{\{cond\}}$ is the indicator function, and it will return 1 if *cond* is satisfied, otherwise, it will return 0. x_c is the currently-interest of target user u at time t, and $\mathbb{I}_u^t$ is the set of items in the currently-interest fragment of target user u at time t. As described in Sect. 5.2, we expressed the interest-keeping as the self-transition in personalized ITM, and utilized reappearance probability to estimate the self-transition probability approximately. Therefore, we lead the weight parameter ω into (13) to control the weight of interest-transferring and interest-keeping, as well as to offset the inaccuracy of approximate method.

Finally, we sort the candidates by their preference scores in descending order and recommend the top-N items to the target user u.

6 Experiments

We have conducted a set of experiments to examine the performance of our recommendation method compared with the baselines. We first introduce the experimental settings, and then analyze the evaluation results.

6.1 Experimental Settings

Data Description. There are two real world datasets used in our experiments: MovieLens1M [15] and Hetrec2011-MovieLens-2k [16]. The two datasets are the most widely used stable benchmark datasets in recommendation research projects. The MovieLens1M dataset contains 6,040 users who have issued 999,209 explicit ratings on a 5-point likert scale, referring to 3,883 movies.
The Hetrec2011-MovieLens-2k dataset is an extension of MovieLens10M dataset, and it is published by GroupLeans research group [17]. It has 2113 users, 10,197 movies and 855,598 ratings provided by these users. In the dataset, the detailed information about each movie, i.e. genre, cast or location, is also provided.

Evaluation Metric. We adopt the All-But-One evaluation method and use Hit-Rate (HR) and the Average Reciprocal Hit-Rank (ARHR) [1] as quality measures for sequential recommendation. Our datasets were split into two subsets, the training set and the test set. For every user, the latest item in his/her action sequence is selected as test data and the rest of action sequence is used as training data.

When making recommendation, we use HOMMIT to generate a recommendation list of N items named $R(u,t)$ for each user u at time t. If the test item of the user u appears in $R(u,t)$, we call it a hit. The Hit Ratio is calculated in the following way:

$$HR = \frac{\sum_{u} \mathrm{I}_{\{(T_u \in R(u,t))\}}}{|\mathcal{U}|} \tag{14}$$

where $\mathrm{I}_{\{cond\}}$ is the indicator function, T_u is the ground-truth item from the test set. One limitation of the Hit-Rate measure is that it treats all hits equally regardless of where they appear in the list of the top-N recommended items. That is, a hit that occurs in the first position is treated equally with a hit that occurs in the N-th position. We address this limitation by the average reciprocal hit-rank (ARHR) which is measured as:

$$ARHR = \frac{1}{n}\sum_{i=1}^{h}\frac{1}{p_i} \tag{15}$$

where h is the number of hits that occurred at positions $p_1, p_2, \ldots, p_h$ within the top-N recommendation lists. That is, hits that occur earlier in the top-N lists are weighted higher than those occurring later.

Compared Methods. The most closely related method to ours are (1) user interest model based recommend algorithm, (2) Markov model based sequential recommendation algorithm. Therefore, we examine the performance of the proposed HOMMIT approach by comparing it with ItemKNN [4], TWItemKNN [2], NPMC-MLE [10], and multiple implementations of IFMM [8].

Item-based collaborative filtering method (ItemKNN) is famous recommendation algorithm which uses the most similar items to a user's already-rated items to generate a list of recommendations. Time weight item-based collaborative filtering method (TWItemKNN) utilizes exponential decay function to model information timeliness where the more recent the information, the higher the value of the time function is. NPMC-MLE is a Markov chain based recommend algorithm mentioned in [10] which estimating transition probability by maximum likelihood estimation (MLE) method. NPMC-MLE was designed for next-basket prediction task based on sequential basket data but it can be easily extended to item recommendation for sequential action data. IFMM is a music recommendation framework integrating interest-forgetting property with variable-order Markov model. Multiple implementations of the IFMM (IFMM-LL, IFMM-EX, and IFMM-HY) were compared with our method in the experiments.

We evaluated our methods with two interest confidence ($\mathcal{C}_x^{u,t}$) implementations and two timeliness of user interest ($\mathcal{T}_x^{u,t}$) implementations, i.e., ED_EG, ED_RG, RD_EG, and RD_RG where EG, RG, ED, and RD represent exponential growth model, rational growth model (interest confidence), exponential decay model, and rational decay model (timeliness of user interest), respectively.

6.2 Evaluations

Impact of Parameter $\boldsymbol{\omega}$. We first focus on analyzing the parameter $\boldsymbol{\omega}$, which governs the influence of interest-transferring and interest-keeping. Since TOP-1 and TOP-10 recommendations have the most demanding requirements of recommendation algorithm and also have the most realistic application values, we evaluate the parameter $\boldsymbol{\omega}$ based on the Hit-Rates in TOP-1 and TOP-10 recommendation results. In the first experiment, we vary the parameter $\boldsymbol{\omega}$ from 0, 0.1 to 1. The results of multiple implementations of HOMMIT of using different constant $\boldsymbol{\omega}$ on both datasets are demonstrated in Fig. 4.

The results have shown that ω is important in determining the Hit-Rate, and ignoring either interest-transferring ($\omega = 1$) or interest-keeping ($\omega = 0$) cannot generate good results. Optimal results can be gotten by combining interest-transferring and interest-keeping together. The optimal ω of each implementation of HOMMIT in both datasets are shown in Table 2. In the following experiments, ω are set to the corresponding optimal values.

6.3 Overall Accuracy Performance

In this section, we first evaluated the overall accuracy performance of our proposed methods. Tables 3 and 4 show the hit-rates and average reciprocal hit-ranks of the

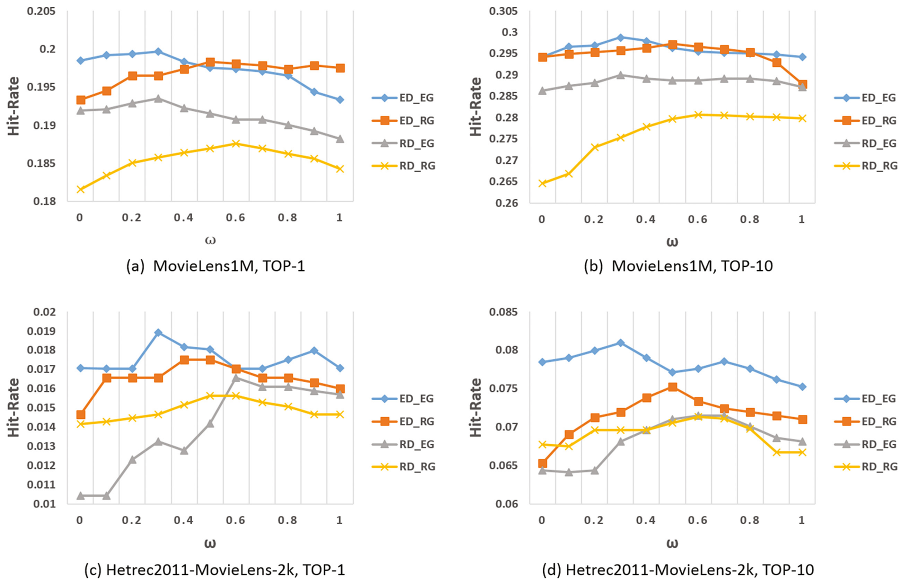

Fig. 4. The impact of ω on Hit-Rate

Table 2. Optimal value of ω

Dataset	ED_EG	ED_RG	RD_EG	RD_RG
MovieLens1M	0.3	0.5	0.3	0.6
Hetrec2011-MovieLens-2k	0.3	0.5	0.6	0.6

proposed methods under different length of recommendation list on both datasets. In general, all the proposed methods have shown promising performance in the experiments. As for the choice of interest confidence expressions, we observed that EG methods outperforms the RG methods in this evaluation. As for the choice of timeliness of user interest expressions, ED methods outperformed the RD methods on both datasets.

Table 3. HR and ARHR of proposed methods on the MovieLens1M

Metric method	Hit-Rate (HR)				Average reciprocal Hit-Rank (ARHR)			
	ED_EG	ED_RG	RD_EG	RD_RG	ED_EG	ED_RG	RD_EG	RD_RG
TOP-1	0.1997	0.1933	0.1912	0.1876	1	1	1	1
TOP-10	0.2969	0.2911	0.2891	0.2803	0.7752	0.7722	0.7719	0.7698
TOP-20	0.3555	0.3501	0.3461	0.3389	0.6761	0.6738	0.6732	0.6689
TOP-50	0.4578	0.4492	0.4442	0.4332	0.5276	0.5229	0.5221	0.5164
TOP-100	0.5509	0.5408	0.5301	0.5205	0.4309	0.4247	0.4239	0.4179

Table 4. HR and ARHR of proposed methods on the Hetrec2011-MovieLens-2k

Metric method	Hit-Rate (HR)				Average reciprocal Hit-Rank (ARHR)			
	ED_EG	ED_RG	RD_EG	RD_RG	ED_EG	ED_RG	RD_EG	RD_RG
TOP-1	0.0189	0.0175	0.0165	0.0156	1	1	1	1
TOP-10	0.0809	0.0752	0.0714	0.0713	0.4459	0.4334	0.4322	0.4281
TOP-20	0.1160	0.1074	0.1041	0.1027	0.3341	0.3267	0.3266	0.3219
TOP-50	0.1822	0.1746	0.1708	0.1680	0.2197	0.2087	0.2055	0.2003
TOP-100	0.2375	0.2297	0.2234	0.2144	0.1706	0.1655	0.1634	0.1618

Next, we compared the accuracy performance between our methods and the baselines. ED_EG is selected as the representative of our methods since it shows the best performance among the 4 proposed methods. The comparison results have been shown in Fig. 5, and we can observe that ED_EG outperforms all the baselines under different length of recommendation list. In particular, ED_EG exhibits excellent accuracy performance in the experiments of shorter recommendation list ($N \leq 10$). The absolute improvement of ED_EG is about 5% to 15% compared with the best of the baseline IFMM-LL on the MovieLens1M dataset, and ED_EG improves IFMM-LL up to 10% to 40% on the Hetrec2011-MovieLens1M-2k dataset. The experiment proves that the personalized interest-transferring model plays an important role in improving recommendation accuracy, and HOMMIT can get better accuracy in sequential recommendation.

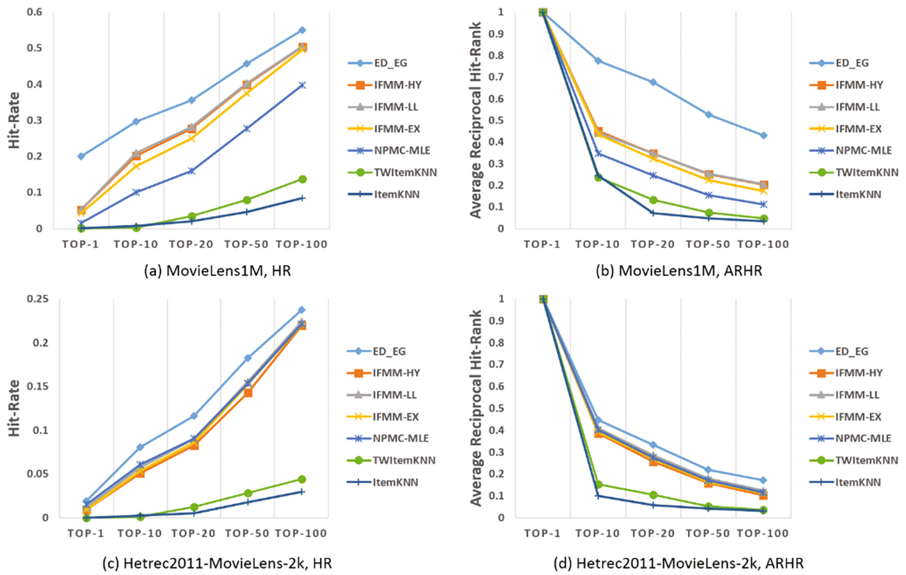

Fig. 5. Comparisons on HR and ARHR of recommendation

7 Conclusion and Future Work

A user's interests have a dynamic transition process during the interaction with the system. Modeling and leveraging this dynamic transition process for sequential recommendation are new great challenges. In this paper, we proposed a novel recommendation framework named HOMMIT, which can identify user interests and adapt an improved high-order Markov chain method to model the dynamic transition process of user interests. Based on that, HOMMIT can predict the transition trends of user interest and make personalized sequential recommendation. The experimental results have shown a significant improvement in the accuracy of our proposed sequential recommendation methods compared with the baselines. In our future works, we will try to make this framework more "human-minded", and the user's interest pattern should be incorporated into this framework.

Acknowledgements. This work is supported by NSF of China (No. 61602237), the Science and Technology Development Plan of Shandong, China (Nos. 2014GGX101047, 2014GGX101019).

References

1. Karypis, G.: Evaluation of item-based top-N recommendation algorithms. In: CIKM 2001, pp. 247–254 (2001)
2. Ding, Y., Li, X.: Time weight collaborative filtering. In: CIKM 2005, pp. 485–492 (2005)
3. Raftery, A.E.: A model for high-order markov chains. J. Roy. Stat. Soc. **47**(3), 528–539 (1985)
4. Deshpande, M., Karypis, G.: Item-based top-N recommendation algorithms. ACM TOIS **22**(1), 143–177 (2004)
5. Xu, Y., Hong, X., Peng, Z., Yang, G., Yu, P. S.: Temporal recommendation via modeling dynamic interests with inverted-U-Curves. In: DASFAA 2016, pp. 313–329 (2016)
6. Chen, J., Wang, C., Wang, J.: Modeling the interest-forgetting curve for music recommendation. In: MM 2014, pp. 921–924 (2014)
7. Toscher, A., Jahrer, M., Bell, R. M.: The BigChaos solution to the Netflix Grand prize (2008)
8. Chen, J., Wang, C., Wang, J.: A personalized interest-forgetting markov model for recommendations. In: AAAI 2015, pp. 16–22 (2015)
9. Koychev, I., Schwab, I.: Adaptation to drifting user's interests. In: ECML 2000 Workshop: Machine Learning in New Information Age (2000)
10. Rendle, S., Freudenthaler, C., Schmidt-Thieme, L.: Factorizing personalized Markov chains for next-basket recommendation. In: WWW 2010, pp. 811–820 (2010)
11. He, R., Mcauley, J.: Fusing similarity models with markov chains for sparse sequential recommendation. In: ICDM 2016, pp. 191–200 (2016)
12. Cheng, C., Yang, H., Lyu, M.R., King, I.: Where you like to go next: successive point-of-interest recommendation. In: IJCAI, pp. 2605–2611 (2013)
13. Yin, B., Yang, Y., Liu, W.: Exploring social activeness and dynamic interest in community-based recommender system. In: WWW 2014, pp. 771–776 (2014)
14. Jolliffe, I.T.: Pincipal component analysis. J. Mark. Res. **25**, 513 (2002)
15. MovieLens: http://grouplens.org/datasets/movielens
16. Hetrec2011-MovieLens-2k: https://grouplens.org/datasets/hetrec-2011
17. GroupLeans: https://grouplens.org

View publication stats