Modeling the Interest-Forgetting Curve for Music Recommendation

Jun CHEN, Chaokun WANG, Jianmin WANG

School of Software, Tsinghua University, Beijing 100084, P.R. China

chenjun14@mails.thu.edu.cn, chaokun@mail.tsinghua.edu.cn, jimwang@mail.tsinghua.edu.cn

Abstract

Music recommendation plays a key role in our daily lives as well as in the multimedia industry. This paper adapts the memory forgetting curve to model the human interest-forgetting curve for music recommendations based on the observation of recency effects in people's listening to music. Two music recommendation methods are proposed using this model with respect to the sequence-based and the IFC-based transition probabilities, respectively. We also bring forward a learning method to approximate the global optimal or personalized interest-forgetting speed(s). The experimental results show that our methods can significantly improve the accuracy in music recommendations. Meanwhile, the IFC-based method outperforms the sequence-based method when recommendation list is short at each time.

Music Recommendation

Given: Song set $\mathcal{S} = \{s_1, \dots, s_{|\mathcal{S}|}\}$,

User set $\mathcal{U} = \{u_1, \dots, u_{|\mathcal{U}|}\}$,

Row-normalized song transition probability matrix $M(s_i, s_j) \geq 0$,

Playlist set of user u, $\mathcal{P}^u = \{p_1^u, \dots, p_{l_u}^u\}$

$p_i^u \in \mathcal{S}^*$ (Kleene closure of $\mathcal{S}$),

Current playlist (not expired) of user u, $\widehat{p^u}$.

Recommend Top-K songs by ranking:

$$\mathcal{S}_{rec} = \underset{s_i \in \mathcal{S}}{\operatorname{argmax}} \Pr(s_i | \widehat{p^u}, \mathcal{P}^u, M)$$

$$\Pr(s_i | \widehat{p^u}, \mathcal{P}^u, M) = \sum_{s_j \in \widehat{p^u}} \Pr(s_j | \widehat{p^u}) \, M(s_j, s_i)$$

Candidate set:

$$\mathcal{C}_{\widehat{p^u}} = \{s_x | \exists s_j \in \widehat{p^u} \wedge M(s_j, s_i) > 0 \wedge s_x \notin \widehat{p^u}\}$$

Neighborhood-normalized transition probability:

$$M_{\mathcal{C}_{\widehat{p^u}}}(s_j, s_i) = \frac{M(s_j, s_i)}{\sum_{s_x \in \mathcal{C}_{\widehat{p^u}}} M(s_j, s_x)}$$

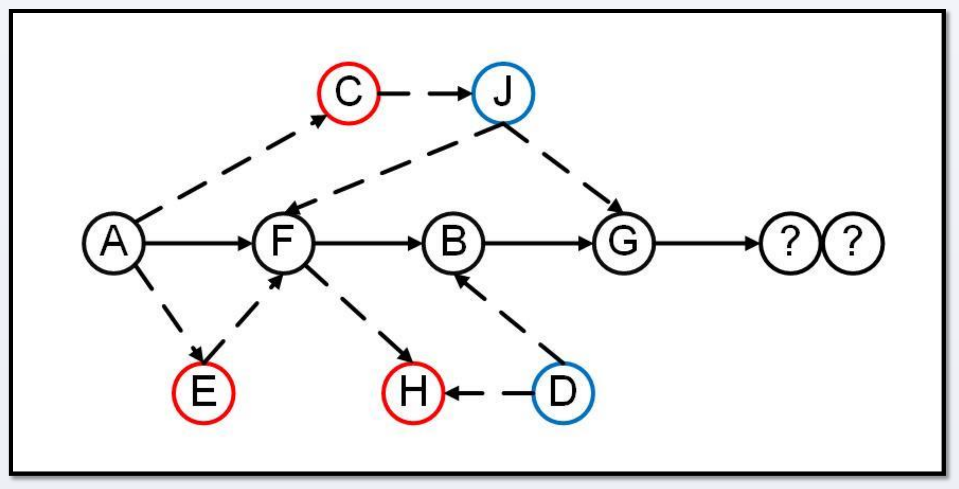

Fig.1 Recommendation illustration. Music A, F, B, G are in the current playlist. Music C, E, H are candidates due to reachability. Music J and D are out of consideration for recommendation at this time.

Interest-Forgetting Curve

Memory retention as time elapses:

$$R = e^{-\frac{t}{S}},$$

S, strength of memory; t, elapsed time.

Interest retention as time step increases:

$$R = e^{-\alpha t},$$

α, (personalized) interest forgetting speed.

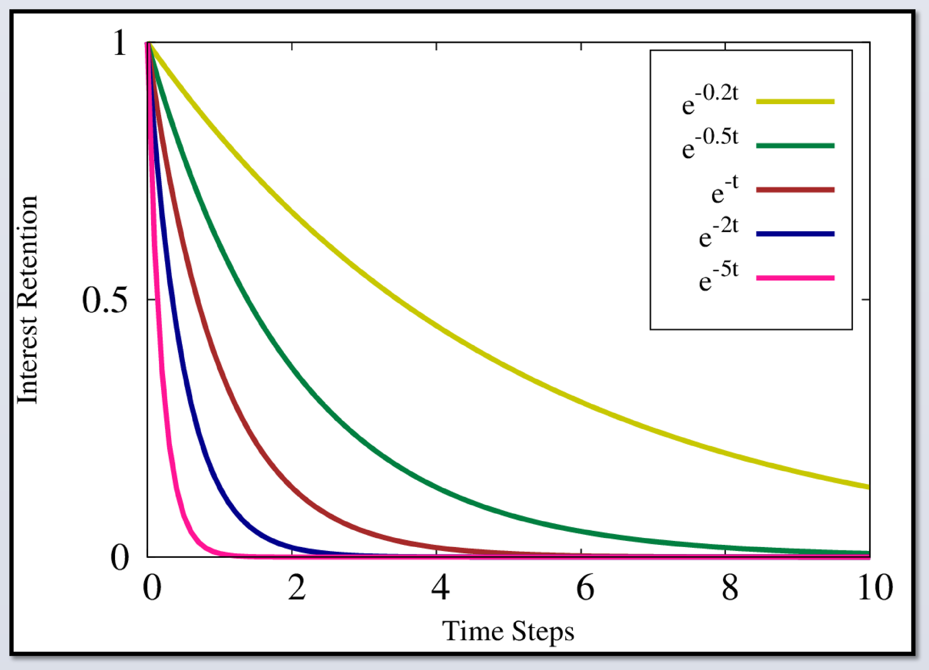

Fig.2 Users's interest on a certain item loses as time elapses.

Modeling IFC in Recommendation

Transition Probabilities

(1) Sequence-based:

$$M_{SEQ}(s_i, s_j) = \frac{\sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}^u} \mathbb{I}_{\{s_i s_j\} seq \subseteq p}}{\sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}^u} \mathbb{I}_{\{s_i\} seq \subseteq p}},$$

$\{s_i, s_j\} seq$ is a subsequence of playlist p. $\mathbb{I}_{cond}$ is the indicator function, which returns 1 when $cond$ is satisfied, othewise, returns 0.

(2) IFC-based:

$$M_{IFC}(s_i, s_j) = \frac{\sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}^u} e^{-\alpha(p(s_j) - p(s_i))} \mathbb{I}_{\{s_i s_j\} seq \subseteq p}}{\sum_{s_x} \sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}^u} e^{-\alpha(p(s_x) - p(s_i))} \mathbb{I}_{\{s_i s_x\} seq \subseteq p}},$$

$p(s_i)$ represents the position index of s_i in p.

Prior Probabilities:

$$\Pr(s_j | \widehat{p^u}) = \frac{e^{-\alpha(|\widehat{p^u}| - \widehat{p^u}(s_j))}}{\sum_{s_x \in \widehat{p^u}} e^{-\alpha(|\widehat{p^u}| - \widehat{p^u}(s_x))}}$$

Learning Interest Forgetting Speed

$$G(\alpha) = \max_{\alpha \in \mathbb{R}_{\geq 0}} \frac{1}{|\mathcal{P}_{train}|} \sum_{p \in \mathcal{P}_{train}} \Pr(\widehat{s^p} | p \backslash \{\widehat{s^p}\}, \mathcal{P}_{train}, M_{\mathcal{C}_p})$$

$$= \max_{\alpha \in \mathbb{R}_{\geq 0}} \frac{1}{|\mathcal{P}_{train}|} \sum_{p \in \mathcal{P}_{train}} \sum_{s_j \in p \backslash \{\widehat{s^p}\}} \Pr(s_j | p \backslash \{\widehat{s^p}\}) \, M_{\mathcal{C}_p}(s_j, \widehat{s^p})$$

Maximize $G(\alpha)$ to obtain the optimized global α, or personalized α_u for each individual user u.

Evaluations

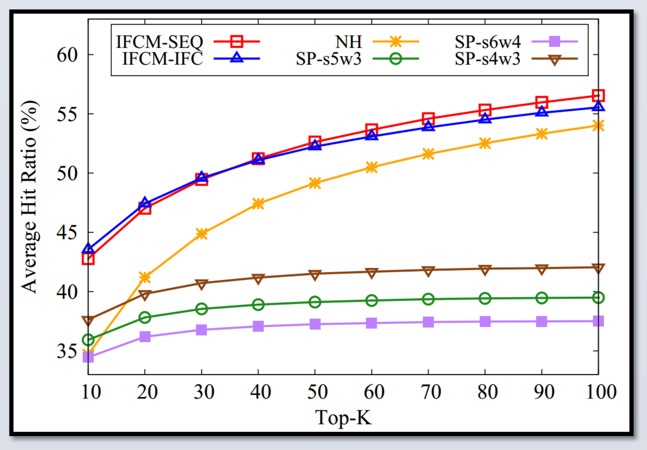

Fig.3 Accuracy comparison with baselines. Neighborhood method (NH), Sequential Pattern method with different supports and window sizes (SP-*).

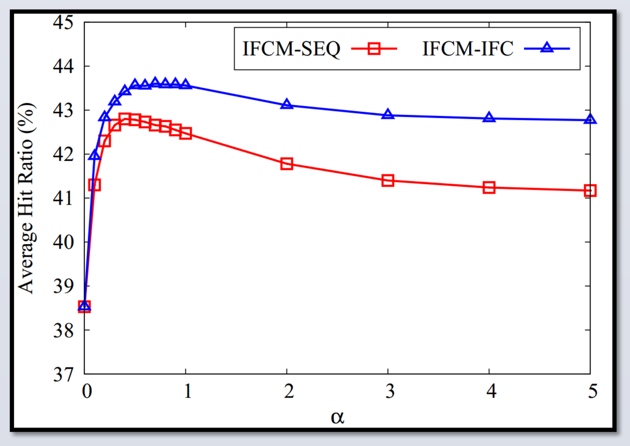

Fig.4 Comparison on the selection of transition probability expression.

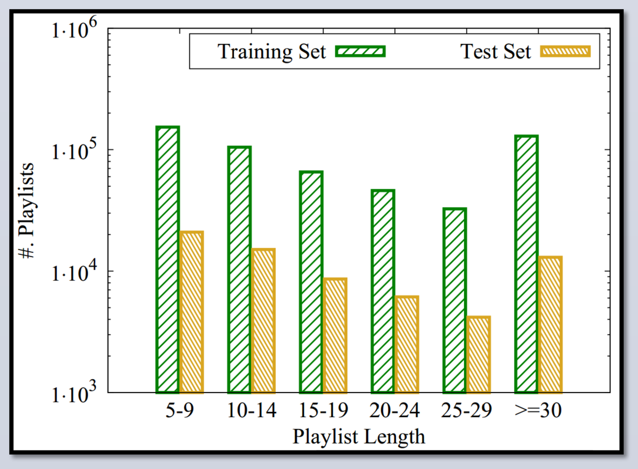

Fig.5 Distribution of playlist length in the training set and the test set.

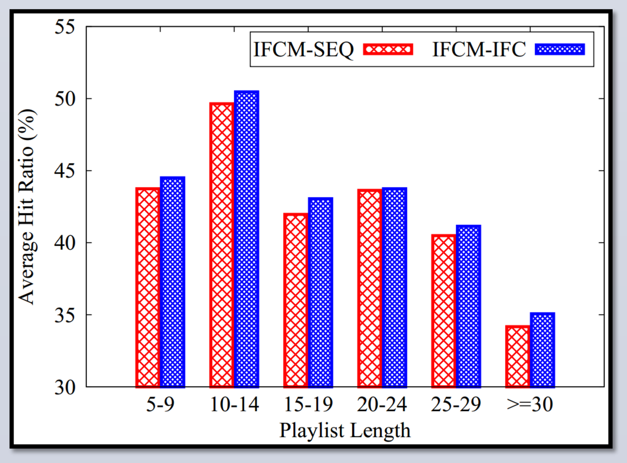

Fig.6 Accuracy under different groups of playlist length.