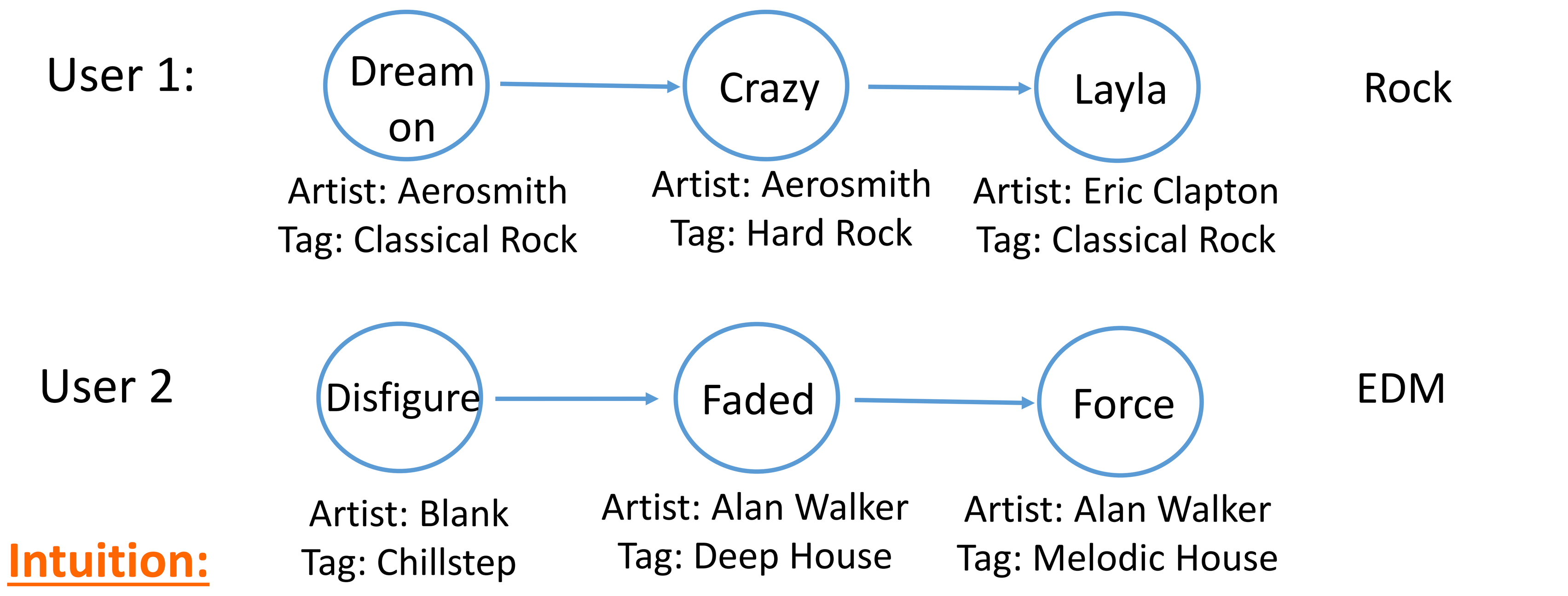
SMU SINGAPORE MANAGEMENT UNIVERSITY
School of Information Systems

Modeling Sequential Preferences with Dynamic User and Context Factors

Duc-Trong Le, Yuan Fang, Hady W. Lauw

Problem

Examples: Given musical playlists from different users

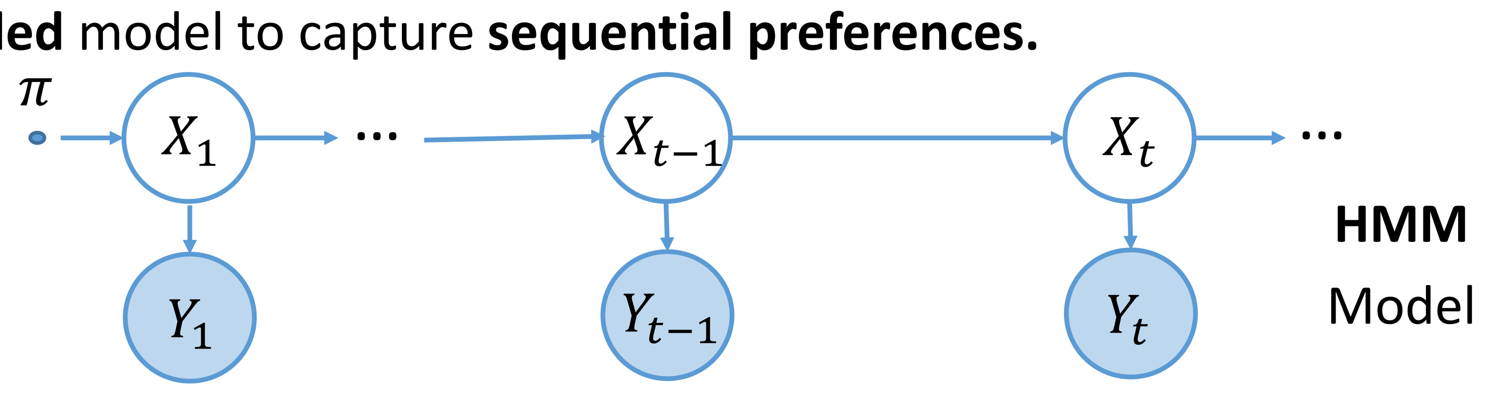


Intuition:

- The sequences are influenced by **latent user factors**.
- The adoption of the **current song** affects the **next selection**, which we refer to as **sequential preference**.
- The sequential preference is triggered by **latent context factors** through **multiple context features** such as: tag, artist, etc.

Framework

Hidden Markov Model (HMM): It is a **well-accepted** and **easy to be extended** model to capture **sequential preferences.**



HMM-Formulation: Given a set of observable items $\mathcal{Y}$ and a set of latent states $\mathcal{X}$, it can be defined by a triplet of parameters $\theta = (\pi, A, B)$

- π is the initial state distribution: $\pi_x \triangleq P(X_1 = x)$;
- A is the transition matrix: $A_{xu} \triangleq P(X_t = u \mid X_{t-1} = x)$;
- B is the emission matrix: $B_{xy} = P(Y_t = y \mid X_t = x)$;

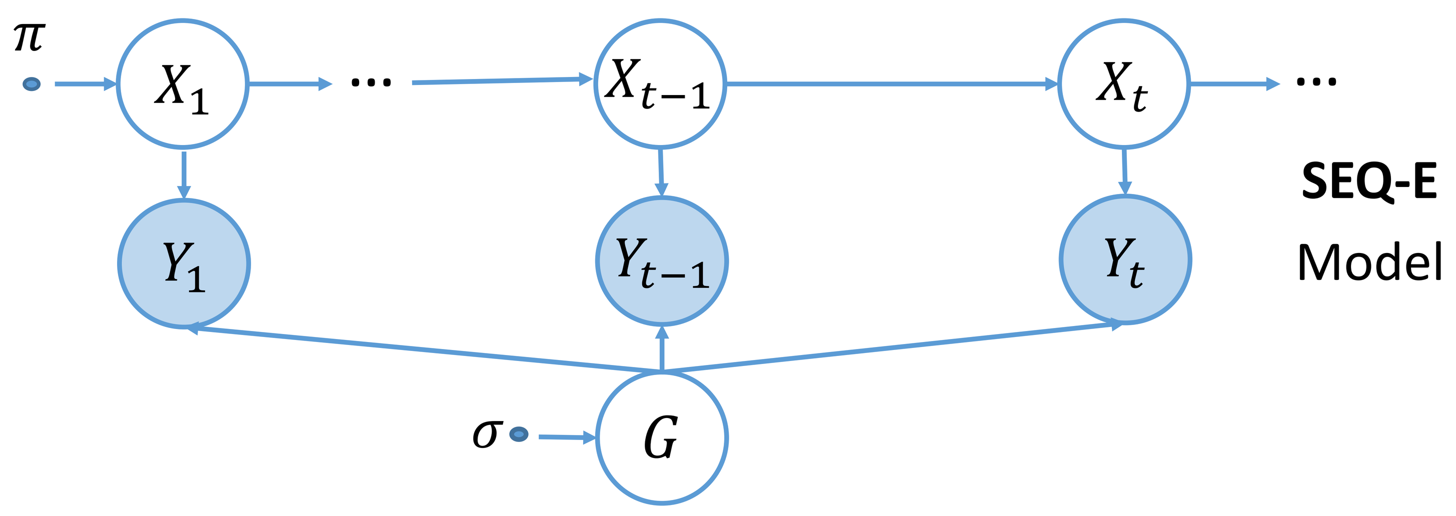
$$\forall\, x, u \in \mathcal{X}; y \in \mathcal{Y};\ t \in \{1, 2, \dots\}$$

Our approach: It is to build a sequential model on top of HMM, which dynamically captures **user and context factors**.

Modeling Dynamic User Factors

Main idea: There exists **different groups** of users

- Users in the **same group** share the **same emission probabilities**
- Users **across groups** may have **different emission probabilities**.



Parameters: $\theta = (\pi, \sigma, A, B)$ with **a set of groups** $\mathcal{G}$

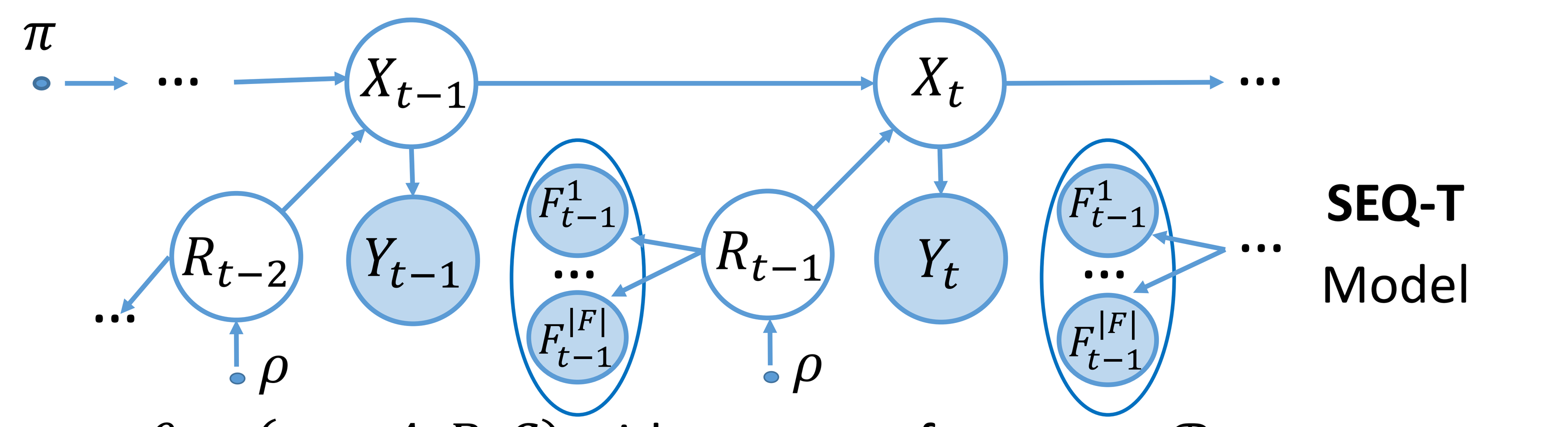
- π and A are the **same** as in a standard HMM
- σ is the **group distribution**: $\sigma_g \triangleq P(G = g)$
- B is the **new emission tensor**: $B_{gxy} \triangleq P(Y_t = y \mid X_t = x, G = g)$

$$\forall\, x, u \in \mathcal{X}; y \in \mathcal{Y}; g \in \mathcal{G};\ t \in \{1, 2, \dots\}$$

Modeling Dynamic Context Factors

Main idea: There exists **context features** and **factors**

- **Transitions** are **affected** by **latent context factors**.
- **Latent context factors** manifest through **observable context features.**

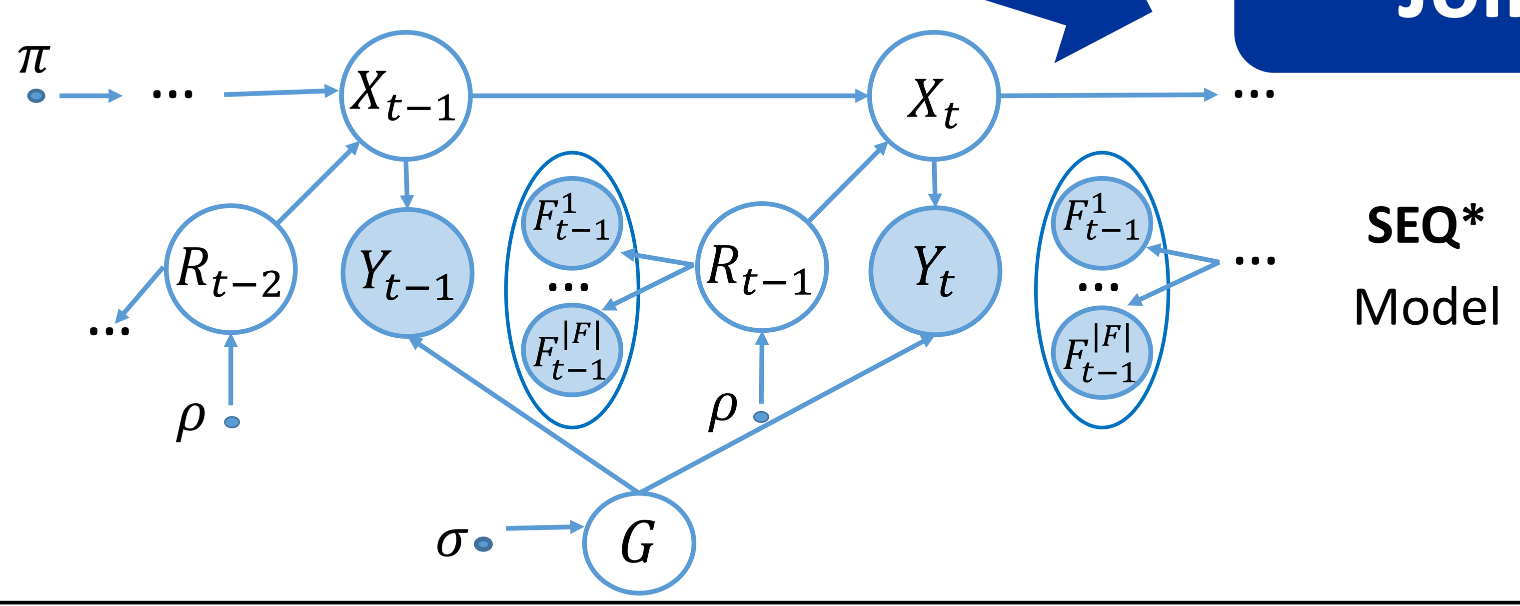


Parameters: $\theta = (\pi, \rho, A, B, C)$ with a context factors set $\mathcal{R}$; a context features set $F = \{F^1, F^2, \dots\}$; each feature F^i takes a values set $\mathcal{F}^i$

- π and B are the **same** as in a standard HMM;
- ρ is the distribution of the **latent context factor**: $\rho_r \triangleq P(R_t = r)$;
- C is the **feature probability matrix**: $C_{rif} \triangleq P\left(F_t^i = f \mid R_t = r\right)$;
- A is the **new transition tensor**: $A_{rxu} \triangleq P(X_t = u \mid X_{t-1} = x, R_{t-1} = r)$

$$\forall\, x, u \in \mathcal{X}; i \in \{1, \dots, |F|\}; f \in \mathcal{F}^i;\ t \in \{1, 2, \dots\}$$

Joint Model



Main idea: Jointly capture both **user** and **context factors** in a single model

Parameters: The **six-tuple** $\theta = (\pi, \sigma, \rho, A, B, C)$ as above

Algorithm: Forward-backward algorithm. For a given sequence length T

$$\theta^* = \operatorname{argmax}_\theta \log P(Y_1, \dots, Y_T, F_1, \dots, F_T; \theta)$$

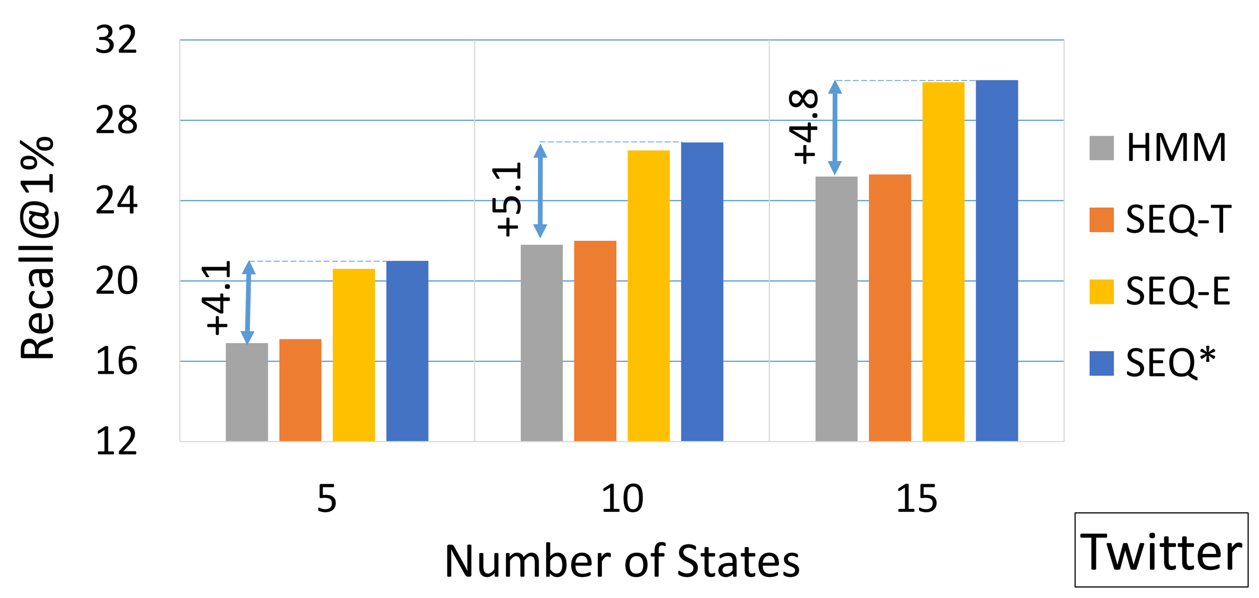
Complexity: $\mathcal{O}(T|\mathcal{R}|(|\mathcal{G}||\mathcal{X}|^2|\mathcal{R}| + |F|))$

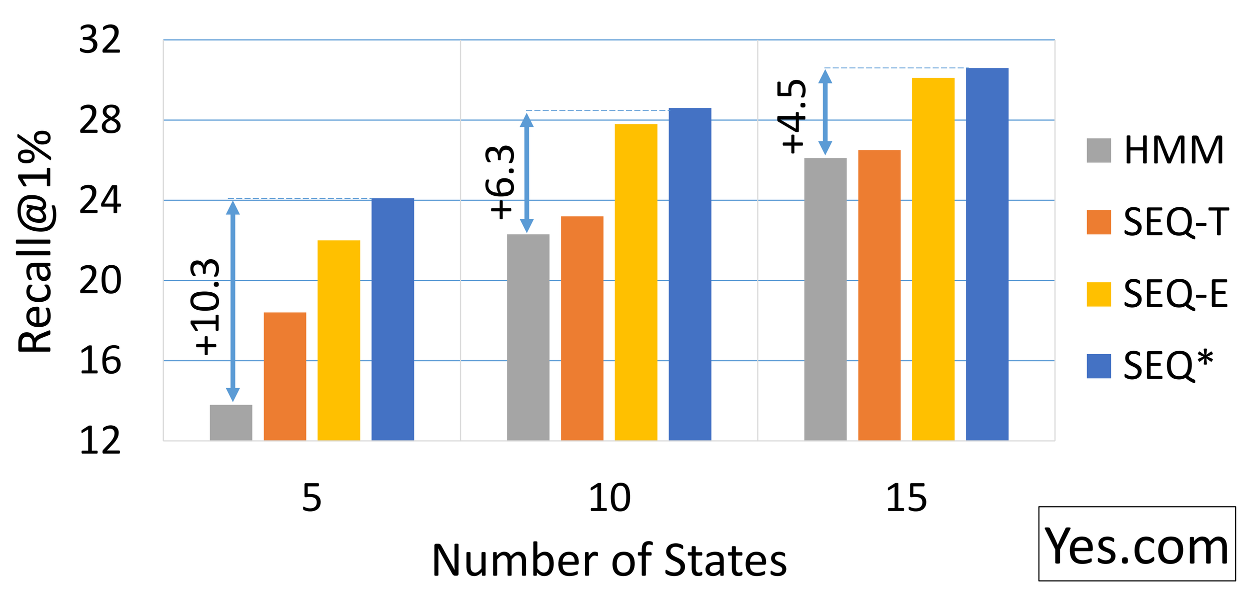
Prediction: $y^* = \operatorname{argmax}_y P(Y_{T+1} = y \mid Y_1, \dots, Y_T, F_1, \dots, F_T; \theta^*)$

Experiments

Datasets: Hashtag sequences (**Twitter**) and song playlists (**Yes.com**).

Methodology: For a given testing sequence length T, consider the **top-K predictions** for the **last item** using: $P(Y_T \mid Y_1, \dots, Y_{T-1}, F_1, \dots, F_{T-1}; \theta)$





Conclusion: Experiments on the two datasets show that **dynamic user and context factors** (of the joint model SEQ*) contribute **statistically significant** improvement as compared to the baseline HMM in term of top-K recommendations for sequences.