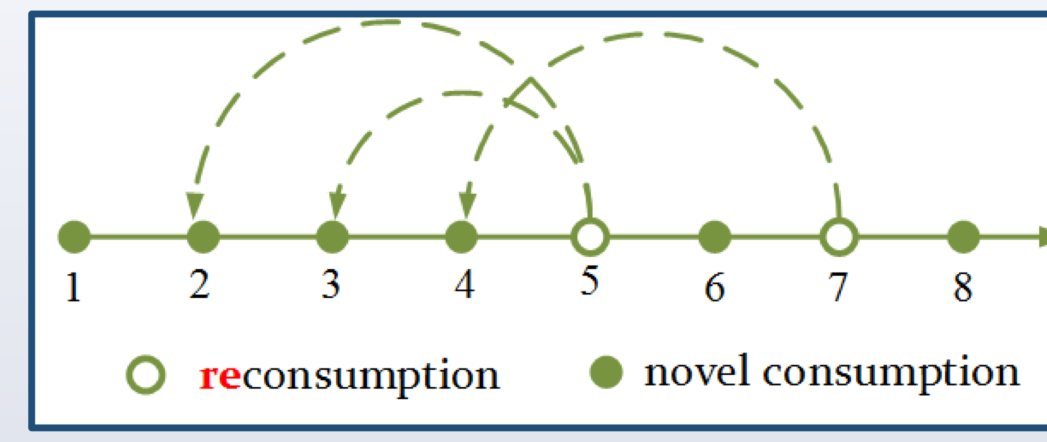
Will You "Reconsume" the Near Past? Fast Prediction on Short-term Reconsumption Behaviors

Jun Chen, Chaokun Wang, Jianmin Wang
School of Software, Tsinghua University, Beijing 100084, P.R. China

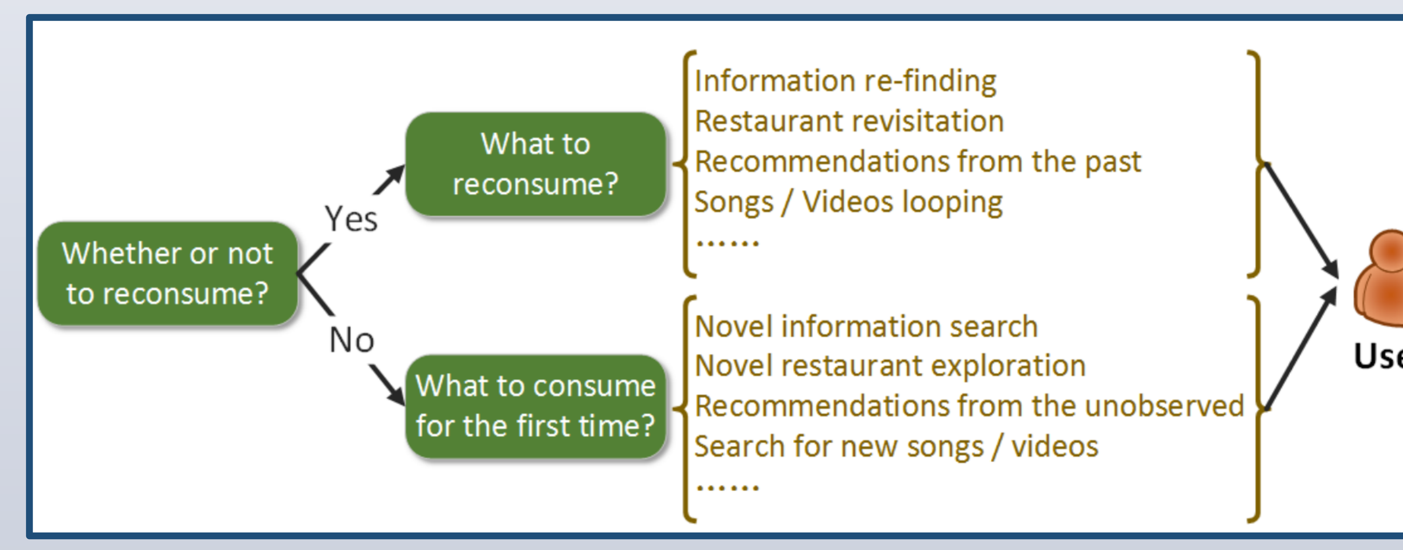
Short-Term REConsumption (STREC) Behaviors

- Reconsume items which have recently been consumed by a same user.
- Alternated now and then with novel consumption behaviors.



Motivations

- A switch problem to narrow down the problem domains.
- Broad applications, e.g. intelligent marketing, recommender systems, web revisitation, information re-finding.
- Challenges: 1) high dynamics of short-term behaviors; 2) multiple influential factors; 3) No representative features.



Feature Extraction

A. Item Popularity

- Average normalized frequencies of items.

B. Item Reconsumption Ratio

- Average normalized reconsumption probability of items.

C. User Reconsumption Ratio

- Reconsumption probability of users.

D. Window Reconsumption Ratio

- Fraction of reconsumptions in the current window.

A

$$h_{IP}(x) = \frac{\log(1+freq(x))}{\max_{y\in\mathbf{X}}\log(1+freq(y))} \qquad h_{IP}(W_k) = \frac{1}{|W_k|}\sum_{x\in W_k} h_{IP}(x)$$

B

$$h_{AIRR}(x) = \log\left(1+\frac{\sum_{u\in\mathbf{U}}\sum_{t\in T_u}\mathbb{I}_{t=x\wedge t\in C_k^{u,t}}}{\sum_{u\in\mathbf{U}}\sum_{t\in T_u}\mathbb{I}_{t=x}}\right)$$

$$h_{IRR}(x) = \frac{h_{AIRR}(x)}{\max_{y\in\mathbf{X}} h_{AIRR}(y)} \qquad h_{IRR}(W_k) = \frac{1}{|W_k|}\sum_{x\in W_k} h_{IRR}(x)$$

C

$$h_{URR}(u) = \frac{\sum_{t\in T_u}\mathbb{I}_{t\in W_k^{u,t}}}{|T_u|}$$

D

$$h_{WRR}(W_k) = 1 - \frac{|DS(W_k)|}{k}$$

Fast Prediction Methods

$$\mathbf{x}_{u,t} = \{h_{IP}(W_k), h_{IRR}(W_k), h_{URR}(u), h_{WRR}(W_k)\}^T$$

1. Linear Method

$$\Pr_{\mathcal{L}}(u,t) = \mathbf{w}^T\mathbf{x}_{u,t}$$

$$\underset{\mathbf{w}_{\mathcal{L}}^*}{\operatorname{argmin}}\ \mathcal{L}(\mathbf{w}) = \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\mathbf{w}^T\mathbf{x}_{u,t} - \mathbb{I}_{t\in W_k^{u,t}})^2 \quad s.t. \sum_i \mathbf{w}_i = \mathbf{1}$$

$$\underset{\mathbf{w}_{\mathcal{L}}^*}{\operatorname{argmin}}\ \mathcal{L}(\mathbf{w}) = \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\mathbf{w}^T\mathbf{x}_{u,t} - \mathbb{I}_{t\in W_k^{u,t}})^2 + \lambda\sum_i \mathbf{w}_i$$

2. Quadratic Method

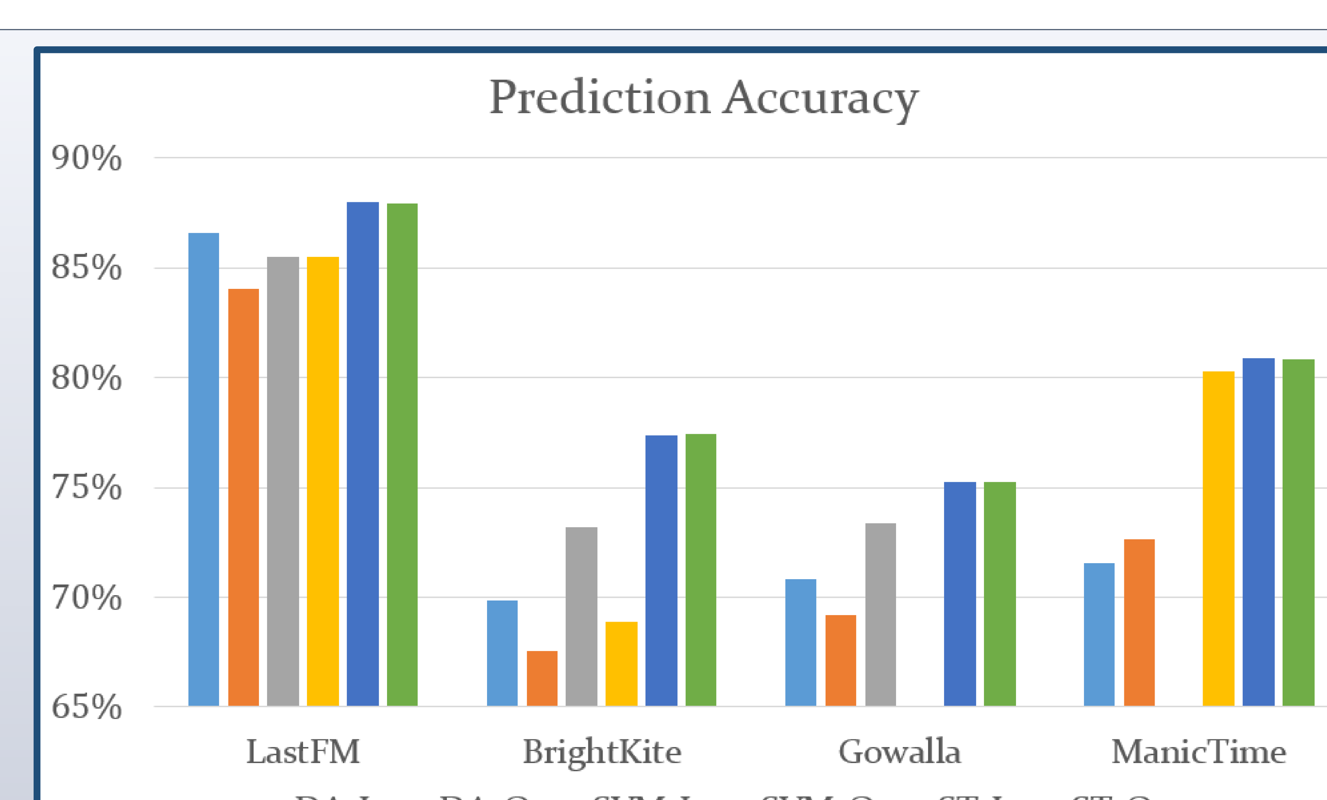
$$\Pr_{\mathcal{Q}}(u,t) = \sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}}$$

$$\underset{\mathbf{w}_{\mathcal{Q}}^*}{\operatorname{argmin}}\ \mathcal{Q}(\mathbf{w}) = \sum_{u\in\mathbf{U}}\sum_{t\in T_u}\left(\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}} - \mathbb{I}_{t\in W_k^{u,t}}\right)^2 \quad s.t.\ \mathbf{w}^T\mathbf{w} = \mathbf{1}$$

$$\underset{\mathbf{w}_{\mathcal{Q}}^*}{\operatorname{argmin}}\ \mathcal{Q}(\mathbf{w}) = \sum_{u\in\mathbf{U}}\sum_{t\in T_u}\left(\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}} - \mathbb{I}_{t\in W_k^{u,t}}\right)^2 + \lambda\mathbf{w}^T\mathbf{w}$$

Experiments

- Collected a new App using data set, ManicTime.
- LastFM, BrightKite, Gowalla data sets.
- On average, 80% prediction accuracy on 4 real-world datasets.
- SVM method is prone to be overfitted.
- The linear and the quadratic methods do not have much difference in prediction accuracy.



The Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15, Austin, USA).
Contact: chenjun14@mails.thu.edu.cn, chaokun@mail.tsinghua.edu.cn.