A Unified Contextual Bandit Framework for Long- and Short-Term Recommendations

Maryam Tavakol and Ulf Brefeld

LEUPHANA UNIVERSITÄT LÜNEBURG

TECHNISCHE UNIVERSITÄT DARMSTADT

Motivation

Personalized recommendation:

Extracting user interests from long-term interactions of user with system

Followed by short-term zeitgeist

Unified model to capture

Long-term part + Short-term component

Framework: Contextual Multi-Armed Bandit e.g., LinUCB

Unified Model

current user

outcome

context

item features

bias term

$$\mathbb{E}[r_{t,a_i}|u_j] = \underbrace{\boldsymbol{\theta}_i^\top \boldsymbol{x}_t}_{Short-term} + \underbrace{\boldsymbol{\beta}_j^\top \boldsymbol{z}_{a_i}}_{Long-term} + b_i$$

parameters of short-term model

parameters of long-term model

Simplified Models:

Focus on the item model

- Short-Term: $\mathbb{E}[r_{t,a_i}] = \boldsymbol{\theta}_i^\top \boldsymbol{x}_t$
- Short-Term+Avg: $\mathbb{E}[r_{t,a_i}] = \boldsymbol{\theta}_i^\top \boldsymbol{x}_t + \boldsymbol{\beta}^\top \boldsymbol{z}_{a_i}$

Focus on the user model

- Long-Term: $\mathbb{E}[r_{t,a_i}|u_j] = \boldsymbol{\beta}_j^\top \boldsymbol{z}_{a_i}$
- Long-Term+Avg: $\mathbb{E}[r_{t,a_i}|u_j] = \boldsymbol{\beta}_j^\top \boldsymbol{z}_{a_i} + \boldsymbol{\theta}^\top \boldsymbol{z}_{a_i}$

General Optimization

Objective function with arbitrary loss, $V(\cdot, r_t)$

$$\sup_{\boldsymbol{\alpha}} \inf_{\substack{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n \\ \boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m \\ \boldsymbol{b},\boldsymbol{y}}} C\sum_{t=1}^{T} V(y_t, r_t) + \frac{1}{2}\sum_i \|\boldsymbol{\theta}_i\|^2 + \frac{\mu}{2}\sum_j \|\boldsymbol{\beta}_j\|^2$$
$$-\sum_{t=1}^{T} \alpha_t(\boldsymbol{\theta}_i^\top \boldsymbol{x}_t + \boldsymbol{\beta}_j^\top \boldsymbol{z}_t + b_t - y_t)$$

Using the **Fenchel-Legendre** conjugate of loss function in the dual space

$$\sup_{\boldsymbol{\alpha}, \mathbf{1}^\top\boldsymbol{\alpha}=0} -C\sum_{t=1}^{T} V^*(-\frac{\alpha_t}{C}, r_t) - \frac{1}{2}\sum_i \boldsymbol{\alpha}^\top (X \circ \boldsymbol{\delta}_i)(X \circ \boldsymbol{\delta}_i)^\top \boldsymbol{\alpha}$$
$$-\frac{1}{2\mu}\sum_j \boldsymbol{\alpha}^\top (Z \circ \boldsymbol{\phi}_j)(Z \circ \boldsymbol{\phi}_j)^\top \boldsymbol{\alpha}$$

Optimization:

- Gradient-based approaches (in dual or primal)
  - Calculating the gradient depends on the loss function
- Model parameters, $(\boldsymbol{\theta}_i, \boldsymbol{\beta}_j)$, are obtained from $\boldsymbol{\alpha}$
- Kernel functions applicable

Algorithm:

for $t = 1, 2, ..., T$ **do**
  Observe user u_t and context $\mathbf{x}_t$
  for all $a \in A_t$ **do**
    Observe arm features $\mathbf{z}_a$
    $v_{t,a}$ = mean reward + confidence bound
  end for
  Choose arm $a_t = \arg\max_a v_{t,a}$, and observe payoff r_t
  Obtain $\boldsymbol{\alpha}$ by optimizing the objective
  Compute $\boldsymbol{\theta}_t$ and $\boldsymbol{\beta}_t$ from $\boldsymbol{\alpha}$
end for

Instantiation

Squared loss:

$$\max_{\boldsymbol{\alpha}, \mathbf{1}^\top\boldsymbol{\alpha}=0} -\frac{1}{2C}\boldsymbol{\alpha}^\top\boldsymbol{\alpha} + \boldsymbol{r}^\top\boldsymbol{\alpha}$$
$$-\frac{1}{2}\boldsymbol{\alpha}^\top[(\sum_i \boldsymbol{\delta}_i \otimes \boldsymbol{\delta}_i^\top) \circ XX^\top + \frac{1}{\mu}(\sum_i \boldsymbol{\phi}_i \otimes \boldsymbol{\phi}_i^\top) \circ ZZ^\top]\boldsymbol{\alpha}$$

$$UCB =: c\sigma$$
$$\sigma^2 = \boldsymbol{x}_t^\top (X^\top X)^{-1} \boldsymbol{x}_t + \boldsymbol{z}_t^\top (Z^\top Z)^{-1} \boldsymbol{z}_t$$

Logistic loss:

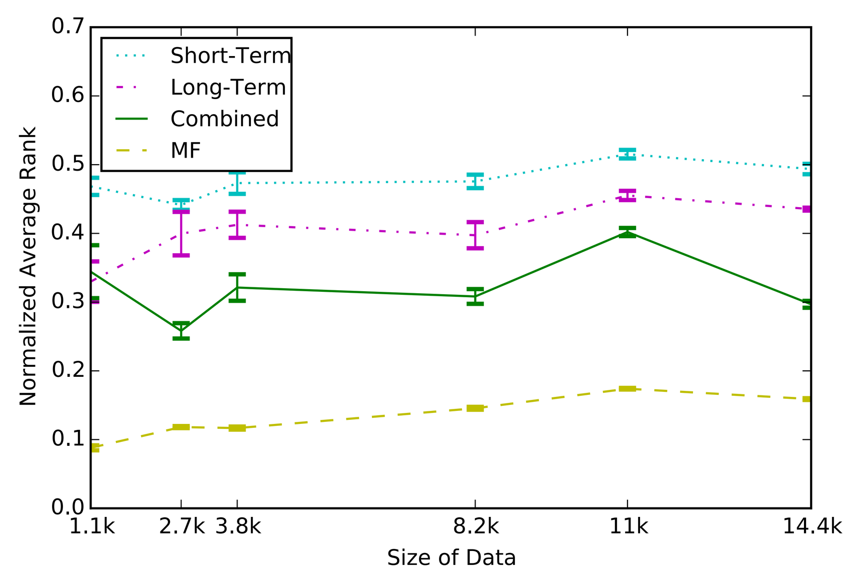
$$\min_{\boldsymbol{\alpha}, \mathbf{1}^\top\boldsymbol{\alpha}=0} C\sum_{t=1}^{T} [(1 - \frac{\alpha_t}{Cr_t})\log(1 - \frac{\alpha_t}{Cr_t}) + \frac{\alpha_t}{Cr_t}\log(\frac{\alpha_t}{Cr_t})]$$
$$+\frac{1}{2}\boldsymbol{\alpha}^\top[(\sum_i \boldsymbol{\delta}_i \otimes \boldsymbol{\delta}_i^\top) \circ XX^\top + \frac{1}{\mu}(\sum_i \boldsymbol{\phi}_i \otimes \boldsymbol{\phi}_i^\top) \circ ZZ^\top]\boldsymbol{\alpha}$$

$$UCB =: c\sigma$$
$$\sigma^2 = \boldsymbol{x}_t^\top (X^\top V_a X)^{-1} \boldsymbol{x}_t + \boldsymbol{z}_t^\top (Z^\top V_u Z)^{-1} \boldsymbol{z}_t$$

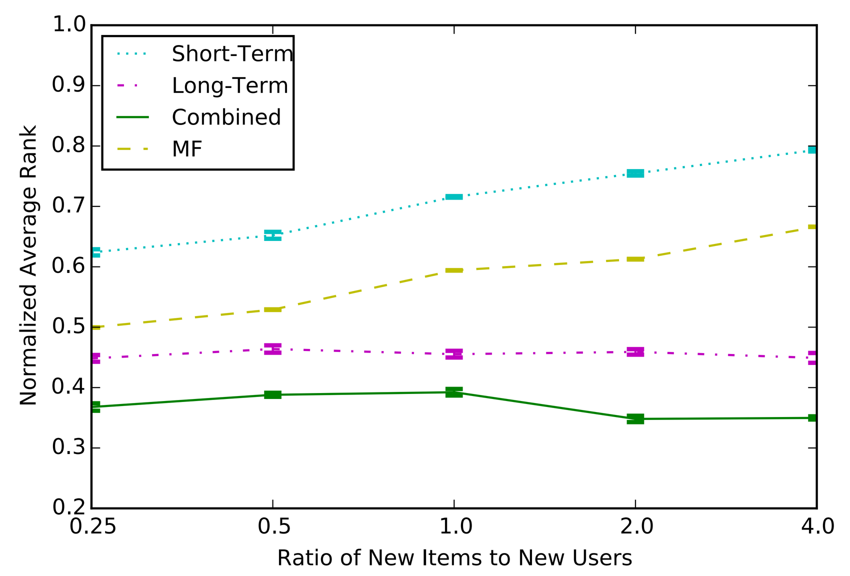
Empirical Study

- Using **squared loss** function
- Dataset: User transactions from Zalando
- Baseline: Matrix Factorization (MF)
- Performance measure: normalized average rank

Performance:

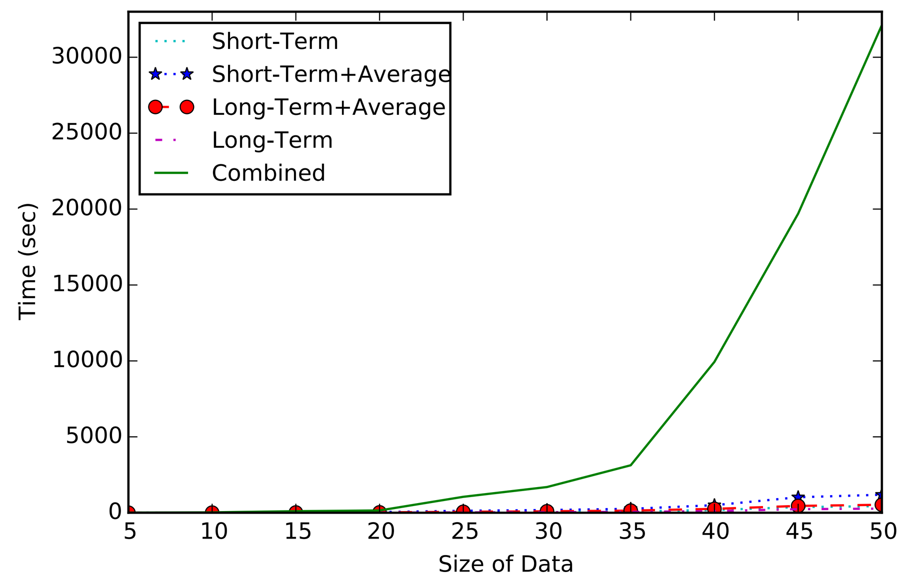


➡ The combined approach outperforms either short- or long-term models —**no new user/item**

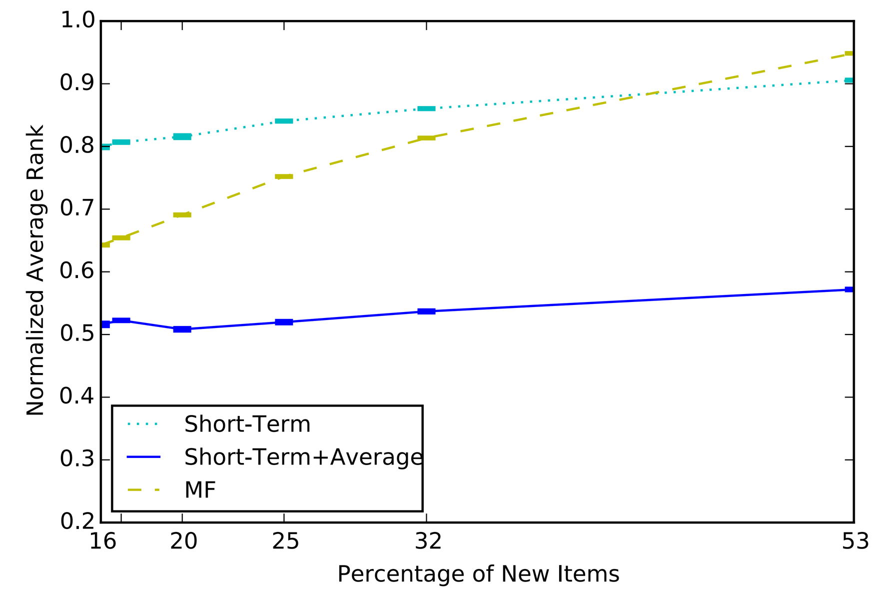


➡ Robustness of combined model in case of **new user/item**

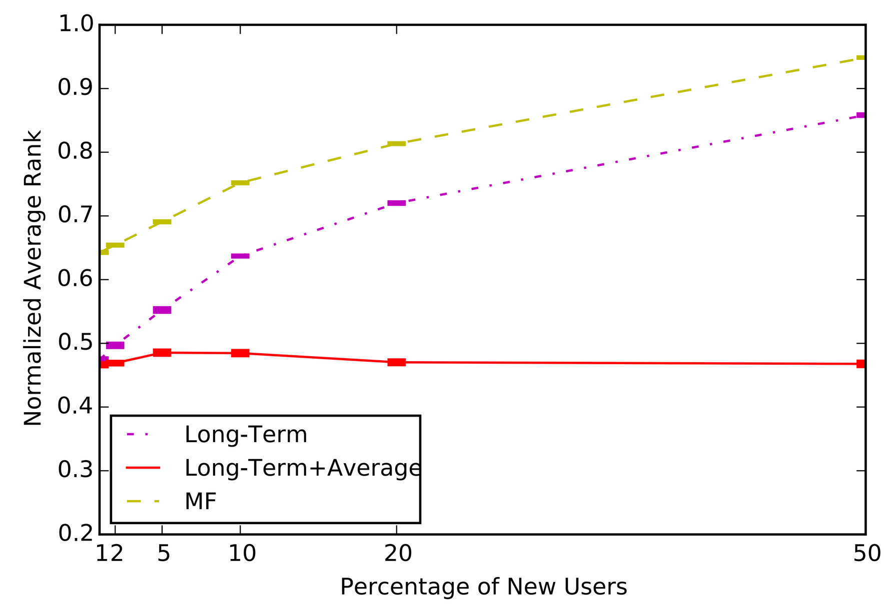
Time Complexity:



➡ The optimization time in combined model is exponential



➡ The average term compensates for the new items



➡ The average term compensates for the new users

Conclusion

- Combining short- and long-term interests of users in one model
- Free choice of loss function and model complexity
- There is no one good model: the choice depends on the application

Maryam Tavakol & Ulf Brefeld,
Machine Learning group,
Leuphana University of Lüneburg,
{tavakol,brefeld}@leuphana.de

Source code available at https://github.com/marytavakol/Bandits