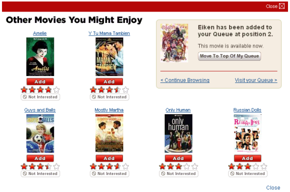
Online Learning to Rank with Features

Authors: Shuai Li, Tor Lattimore, Csaba Szepesvári

The Chinese University of Hong Kong
DeepMind
University of Alberta

Learning to Rank



Amazon, YouTube, Facebook, Netflix, Taobao

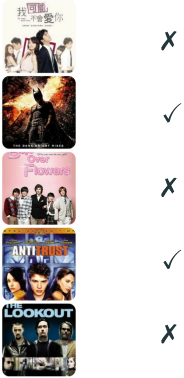
Online Learning to Rank

- There are L items and $K \leq L$ positions
- At each time $t = 1, 2, \ldots,$
  - Choose an ordered list $A_t = (a_1^t, \ldots, a_K^t)$
  - Show the user the list
  - Receive click feedback $C_{t1}, \ldots, C_{tK} \in \{0, 1\}$, per position
- Objective: Maximize the expected number of clicks

$$\mathbb{E}\left[\sum_{t=1}^{T} \sum_{k=1}^{K} C_{tk}\right]$$

Click Models

- Click models describe how users interact with item lists
- Cascade Model (CM)
  - Assumes the user checks the list from position 1 to position K, clicks at the first satisfying item and stops
- Dependent Click Model (DCM)
  - Further assumes there is a satisfaction probability after click
- Position-Based Model (PBM)
  - Assumes the user click probability on an item a of position k can be factored into item attractiveness and position bias
- Generic model
  - Make as few assumptions as possible about the click model

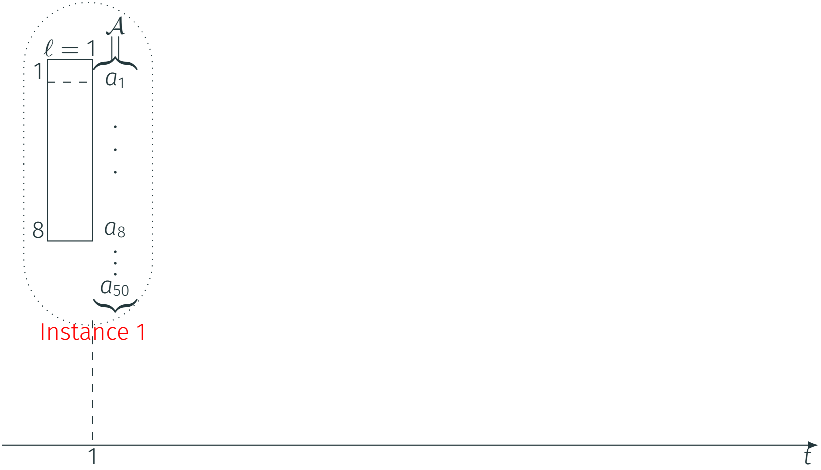


RecurRank

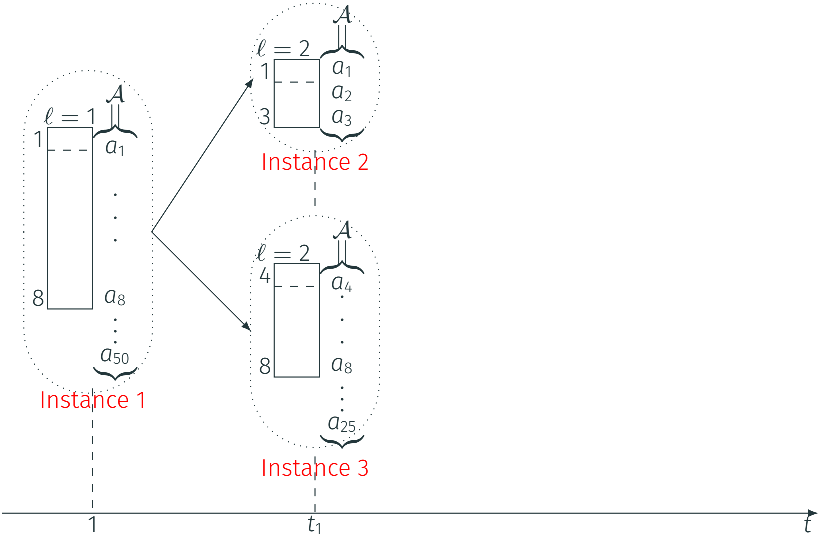
- Each item a is represented by a feature vector $x_a \in \mathbb{R}^d$
- The attractiveness of item a is $\alpha(a) = \theta^\top x_a$
- Click probability factors: $\mathbb{P}_t(C_{ti} = 1) = \alpha(a_i^t)\chi(A_t, i)$ where χ is the examination probability, which satisfies reasonable assumptions

- RecurRank (Recursive Ranking)
- For each phase ℓ
  - Use first position for exploration
  - Use remaining positions for exploitation, rank best items first
- Split items and positions when the phase ends
- Recursively call the algorithm with increased phase

Example

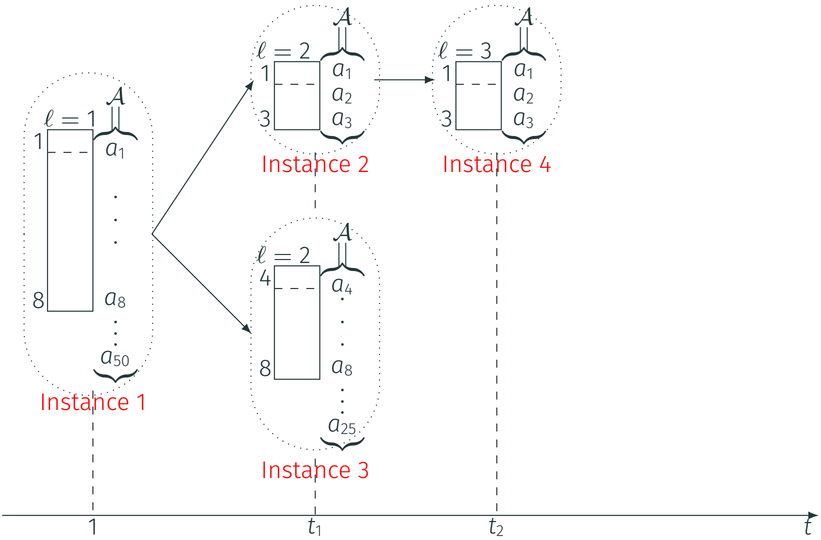


Example



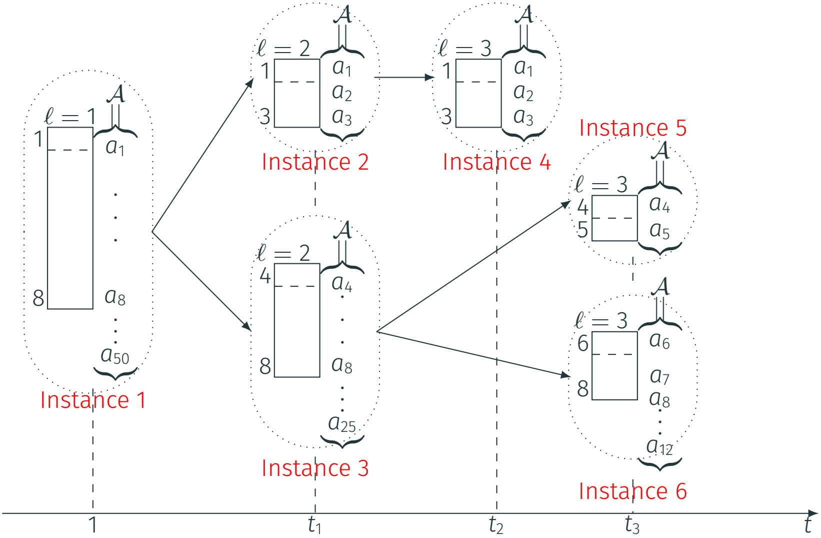
$\mathcal{A}$
$\ell = 1$
1
a_1
8
a_8
a_{50}
Instance 1
$\mathcal{A}$
$\ell = 2$
1
3
a_1
a_2
a_3
Instance 2
$\mathcal{A}$
$\ell = 2$
4
8
a_4
a_8
a_{25}
Instance 3
1
t_1
t

Example



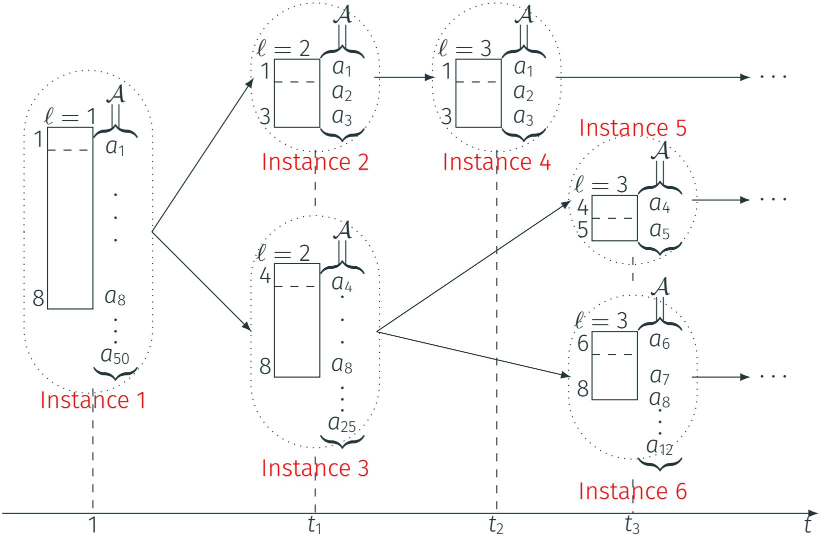
$\mathcal{A}$
$\ell = 1$
1
8
a_1
a_8
a_{50}
Instance 1
$\ell = 2$
1
3
a_1
a_2
a_3
Instance 2
$\ell = 3$
Instance 4
4
a_4
a_{25}
Instance 3
t_1
t_2
t

Example



Instance 1
Instance 2
Instance 3
Instance 4
Instance 5
Instance 6
$\ell = 1$
$\ell = 2$
$\ell = 3$
$\mathcal{A}$
a_1
a_2
a_3
a_4
a_5
a_6
a_7
a_8
a_{12}
a_{25}
a_{50}
1
t_1
t_2
t_3
t

Example


$\mathcal{A}$
$\ell = 1$
1
8
a_1
a_8
a_{50}
Instance 1
$\ell = 2$
1
3
a_1
a_2
a_3
Instance 2
$\ell = 2$
4
8
a_4
a_8
a_{25}
Instance 3
$\ell = 3$
1
3
a_1
a_2
a_3
Instance 4
Instance 5
$\ell = 3$
4
5
a_4
a_5
$\ell = 3$
6
8
a_6
a_7
a_8
a_{12}
Instance 6
1
t_1
t_2
t_3
t

Results

- Regret bound

$$R(T) = O(K\sqrt{dT\log(LT)})$$

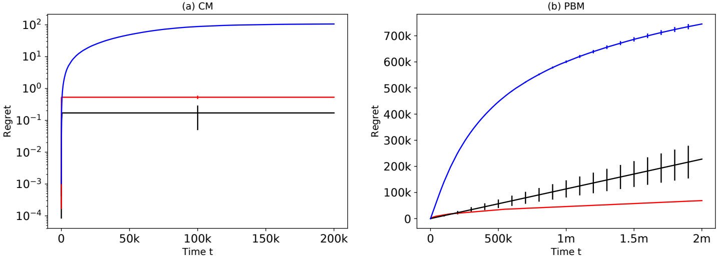
- Improves over existing bound $O\left(\sqrt{K^3LT\log(T)}\right)$

Results

- Regret bound

$$R(T) = O(K\sqrt{dT\log(LT)})$$

- Improves over existing bound $O\left(\sqrt{K^3LT\log(T)}\right)$



—RecurRank —CascadeLinUCB —TopRank

Thank you!

Sumeet Katariya, Branislav Kveton, Csaba Szepesvari, and Zheng Wen.
Dcm bandits: Learning to rank with multiple clicks.
In *International Conference on Machine Learning*, pages 1215–1224, 2016.

Branislav Kveton, Csaba Szepesvari, Zheng Wen, and Azin Ashkan.
Cascading bandits: Learning to rank in the cascade model.
In *International Conference on Machine Learning*, pages 767–776, 2015.

Paul Lagrée, Claire Vernade, and Olivier Cappe.
Multiple-play bandits in the position-based model.
In *Advances in Neural Information Processing Systems*, pages 1597–1605, 2016.

Tor Lattimore, Branislav Kveton, Shuai Li, and Csaba Szepesvari.
Toprank: A practical algorithm for online stochastic ranking.
In *The Conference on Neural Information Processing Systems*, 2018.

Shuai Li, Tor Lattimore, and Csaba Szepesvári.
Online learning to rank with features.
arXiv preprint arXiv:1810.02567, 2018.

Shuai Li, Baoxiang Wang, Shengyu Zhang, and Wei Chen.
Contextual combinatorial cascading bandits.
In *International Conference on Machine Learning*, pages 1245–1253, 2016.

Shuai Li and Shengyu Zhang.
Online clustering of contextual cascading bandits.
In *The AAAI Conference on Artificial Intelligence*, 2018.

Weiwen Liu, Shuai Li, and Shengyu Zhang.
Contextual dependent click bandit algorithm for web recommendation.
In *International Computing and Combinatorics Conference*, pages 39–50. Springer, 2018.

Masrour Zoghi, Tomas Tunys, Mohammad Ghavamzadeh, Branislav Kveton, Csaba Szepesvari, and Zheng Wen.
Online learning to rank in stochastic click models.
In *International Conference on Machine Learning*, pages 4199–4208, 2017.

Shi Zong, Hao Ni, Kenny Sung, Nan Rosemary Ke, Zheng Wen, and Branislav Kveton.
Cascading bandits for large-scale recommendation problems.
In *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*, pages 835–844. AUAI Press, 2016.