Offline Evaluation of Ranking Policies with Click Models
Shuai Li1 Yasin Abbasi-Yadkori2 Branislav Kveton2 S. Muthukrishnan4 Vishwa Vinay2 Zheng Wen2
1The Chinese University of Hong Kong 2Adobe Research 3Google Research 4Rutgers University

Motivation
- Recommendations happen everywhere, such as Amazon, Facebook, Adobe Stock, Google Play, Netflix
- Suppose the existing policy \( \pi \)

with the expected CTR \( V(\pi) \)
- Can we verify a new policy \( h \)

satisfies \( V(h) \geq V(\pi) \) based on logged data under policy \( \pi \)?

Click Models & Estimators
- List estimator [Strehl2010]

\[
\hat{p}_L(h) = \frac{1}{|S|} \sum_{(A,w) \in S} f(A,w) \min \left\{ \frac{h(a_i|x)}{h(a_j|x)} : M \right\}
\]

- With click-model assumptions, we can build estimators that leverage structures of click feedback

- Document-Based Click Model (DCRM):

\[
\hat{w}(a,k|x) \text{ only depends on item } a
\]

Experiments
- Yandex dataset
- The dataset is recorded over 27 days
- Each record contains
  - a query ID
  - the day when the query occurs
  - 10 displayed items as a response to the query
  - the corresponding click indicators of each displayed items

- Logged dataset \( S \)
  - any records except day \( d \)
  - \( h \) is the empirical distribution over \( S \)
  - Evaluation policy \( h \)
  - the records of day \( d \)
  - \( h \) is the empirical distribution of these records
  - The value \( V(h) \) is the average CTR for these records

- Prediction errors on 100 most frequent queries as a function of clipping parameter \( M \)
  - Records of \( K = 2 \) or 3 positions

- Records of \( K = 10 \) positions with DCG value

- The performance of list estimator deteriorates fast with more positions
- The IP estimator performs best

Setting
- Ground set \( E = \{1, \ldots, K\} \) of \( L \) items
- A list is a \( K \)-permutation of \( E \), which is an element of

\[
\prod \{E \} = \{(a_1, \ldots, a_K) : a_1, a_K \in E; a_1 \neq a_j, i \neq j \}
\]
- Context set \( X \)
- \( \hat{w}(a,k|x) \) : the expected CTR of putting item \( a \) in position \( k \) under context \( x \)
- A policy \( \pi \) is a conditional probability distribution of a list given context \( x \)

- The reward of list \( A \)

\[
f(A,w) = \sum_{k=1}^{K} w(a_k,k)
\]
- The value of a policy \( V(\pi) = E_{E \sim E_{\pi}}[f(A,W(\cdot|x))]\)

- At each time \( t \)
  - the environment draws context \( x_t \) and click realizations \( w_t \)
  - The learner observes \( x_t \) and selects \( A_t \) according to policy \( \pi \)
  - The environment reveals \( (w_t(x_t), a_t))_{t=1}^{K} \)
  - Logged dataset: \( S = \{(x_t,A_t,w_t)^{K}_{t=1}\} \)

Objective
- To design statistically efficient estimators based on logged dataset for any ranking policy

Challenge
- The number of different lists is exponential in \( K \)

Analysis

Proposition 1.[Unbiased in a larger class of policies]
Let \( \mathcal{H}_C \) contains all policies such that \( \hat{p}_L \) is unbiased, for any \( Y \in \{L,IP,1,PBM\} \). Then \( \mathcal{H}_C \subseteq \mathcal{H}_P \subseteq \mathcal{H}_L \cap \mathcal{H}_P \).

Proposition 2.[Lower bias in estimating policy]

\[
E_2[\hat{V}(\pi)] \leq E_2[\hat{V}_P(\pi)] \leq E_2[\hat{V}_L(\pi)] \leq E_2[V(h)]
\]

Proposition 3.[Policy optimization]
Suppose \( \hat{h}_G \) is the best policy under \( V_P \), for any \( Y \in \{L,IP,1,PBM\} \). Then the lower bound on \( \hat{h}_G \) is at least as high as that on \( \hat{h}_L \).

Conclusions
- We propose various estimators for the expected number of clicks on lists generated by ranking policies that leverage the structure of click models
- We prove that our estimators are better than the unstructured list estimators, in the sense that they are less biased and have better guarantees for policy optimization
- Our estimators consistently outperform the list estimator in our experiments

References

Full Paper

Contact
Shuai Li
Email: shuaili@cse.cuhk.edu.hk
Yasin Abbasi-Yadkori
Email: abbasia@adobe.com
Branislav Kveton
Email: bkveton@gmail.com
S. Muthukrishnan
Email: muthu@cs.rutgers.edu
Vishwa Vinay
Email: vinay@adobe.com
Zheng Wen
Email: zwen@adobe.com

References