Computationally Efficient Optimization of Plackett-Luce Ranking Models for Relevance and Fairness

Harrie Oosterhuis

July 12, 2021

Radboud University
harrie.oosterhuis@ru.nl
https://twitter.com/HarrieOos

Goal of this work:

- Optimize a **Plackett-Luce** (PL) model for **relevance** or **fairness** ranking metrics,
- with an unbiased method (no heuristic or bounding),
- in a **computationally efficient** way (avoid combinatorial problems).

Contribution: PL-Rank

- A novel **sampling-based** method for quickly estimating PL gradients.
- Derivation to prove the estimation is unbiased.

Background

Background: Problem Setting

For any ranking y, an arbitrary ranking metric uses the weights per rank θ_k, the relevance of the items $P(R = 1 \mid q, d) = \rho_d$, and the policy π with the probability of a ranking $\pi(y \mid q)$:

$$\mathcal{R}(q) = \sum_{y \in \pi} \pi(y \mid q) \sum_{k=1}^{K} \theta_k P(R = 1 \mid q, y_k) = \sum_{y \in \pi} \pi(y) \sum_{k=1}^{K} \theta_k \rho_{y_k} = \mathbb{E}_y \left[\sum_{k=1}^{K} \theta_k \rho_{y_k} \right].$$

This is taken in expectation over a query distribution:

$$\mathcal{R} = \mathbb{E}_q[\mathcal{R}(q)] = \sum_{q \in \mathcal{Q}} P(q) \mathcal{R}(q).$$

This description applies to well-known metrics: precision@k, recall@k, DCG, ARP.

A Plackett-Luce model applies a **SoftMax** to the scores of **unplaced items**:

$$\pi(d \mid y_{1:k}, D) = \frac{\overbrace{\mathbb{1}[d \notin y_{1:k}]e^{m(d)}}^{\text{item score if not placed}}}{\underbrace{\sum_{d' \in D \setminus y_{1:k}} e^{m(d')}}_{\text{sum of all unplaced item scores}}}.$$

The probability of a ranking is the product over each item placement:

$$\pi(y) = \prod_{k=1}^{K} \pi(y_k \mid y_{1:k-1}, \mathcal{D}).$$

We can sample from a Plackett-Luce model by sampling Gumbell Noise: $\zeta_d \sim$ *Gumbell*, and sorting according to $m(d) + \zeta_d$.

The prevalent approach in existing work uses **policy-gradients** with the log-trick:

$$\frac{\delta}{\delta m}\pi(y) = \pi(y)\left[\frac{\delta}{\delta m}\log(\pi(y))\right].$$

Given N samples from π: $y^{(i)} \sim \pi$, the gradient can be unbiasedly estimated:

$$\frac{\delta}{\delta m}\mathcal{R}(q) \approx \frac{1}{N}\sum_{i=1}^{N} \underbrace{\left[\frac{\delta}{\delta m}\log(\pi(y^{(i)}))\right]}_{\text{gradient w.r.t. log prob. of full ranking}} \underbrace{\left(\sum_{k=1}^{K}\theta_k \rho_{y_k^{(i)}}\right)}_{\text{observed reward}}.$$

Method: PL-Rank

A **reward before** rank k should **not influence** the probabilities of the ranking **after** k:

$$\mathcal{R}(q) = \sum_{y \in \pi} \pi(y) \sum_{k=1}^{K} \theta_k \rho_{y_k} = \sum_{k=1}^{K} \theta_k \sum_{y \in \pi} \pi(y) \rho_{y_k} = \sum_{k=1}^{K} \theta_k \sum_{y_{1:k} \in \pi} \pi(y_{1:k}) \rho_{y_k}.$$

Given N samples from π: $y^{(i)} \sim \pi$, the gradient can be unbiasedly estimated:

$$\frac{\delta}{\delta m} \mathcal{R}(q) \approx \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \underbrace{\left[\frac{\delta}{\delta m} \log(\pi(y_k^{(i)} | y_{1:k-1}^{(i)})) \right]}_{\text{prob. of partial ranking up to } k} \underbrace{\sum_{x=k}^{K} \theta_x \rho_{y_x^{(i)}}}_{\text{reward received after } k}.$$

Using the fact that π is a Plackett-Luce model, we can estimate the gradient using:

$$\frac{\delta}{\delta m}\mathcal{R}(q) \approx \frac{1}{N}\sum_{d\in\mathcal{D}} \overbrace{\left[\frac{\delta}{\delta m}m(d)\right]}^{\text{grad. w.r.t. score}} \sum_{i=1}^{N} \Bigg(\overbrace{\left(\sum_{k=\text{rank}(d,y^{(i)})}^{K} \theta_k \rho_{y_k^{(i)}}\right)}^{\text{reward following placement}} - \underbrace{\sum_{k=1}^{\text{rank}(d,y^{(i)})} \pi(d \mid y_{1:k-1}^{(i)}) \left(\sum_{x=k}^{K} \theta_x \rho_{y_x^{(i)}}\right)}_{\text{risk imposed by placement probability}} \Bigg).$$

Given N samples, this can be computed in $\mathcal{O}(N \cdot K \cdot D)$.

Flaw: items that are **not** in the top-K of any of the N **sampled** rankings will **always** have a **negative** gradient.

Method: PL-Rank-2

We can avoid the flaw while maintaining the $\mathcal{O}(N \cdot K \cdot D)$ complexity:

$$\frac{\delta}{\delta m}\mathcal{R}(q) \approx \frac{1}{N}\sum_{d\in\mathcal{D}} \overbrace{\left[\frac{\delta}{\delta m}m(d)\right]}^{\text{grad. w.r.t. score}} \sum_{i=1}^{N} \overbrace{\left(\sum_{k=\text{rank}(d,y^{(i)})+1}^{K} \theta_k \rho_{y_k^{(i)}}\right)}^{\text{future reward after placement}}$$
$$+ \underbrace{\sum_{k=1}^{\text{rank}(d,y^{(i)})} \pi(d \mid y_{1:k-1}^{(i)})\left(\theta_k \rho_d - \sum_{x=k}^{K}\theta_x \rho_{y_x^{(i)}}\right)}_{\text{expected direct reward minus the risk of placement}}.$$

Fairness in exposure generally use rank-based exposure:

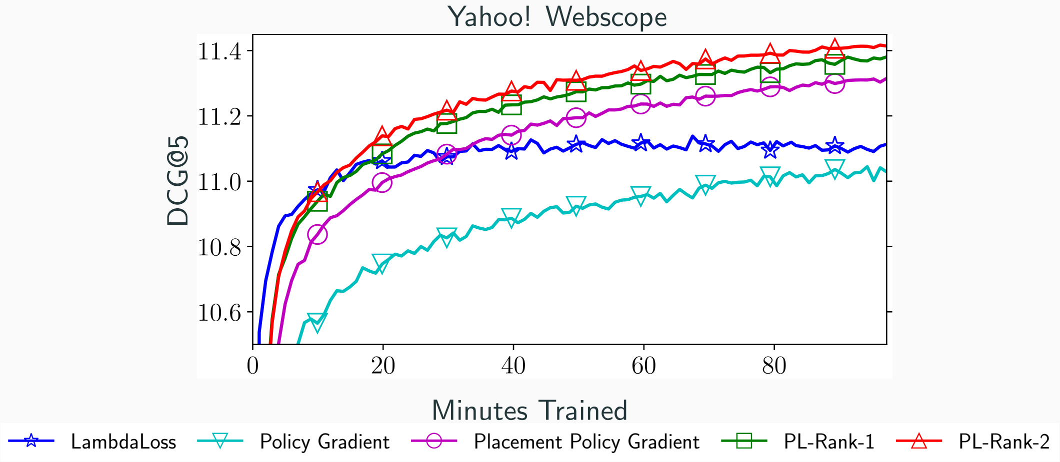
$$\mathcal{E}(q,d) = \mathbb{E}_y\left[\sum_{k=1}^{K} \theta_k \mathbb{1}[y_k = d]\right] = \sum_{y\in\pi} \pi(y) \sum_{k=1}^{K} \theta_k \mathbb{1}[y_k = d].$$

PL-Rank can be used to optimize a fairness based metric $\mathcal{F}$:

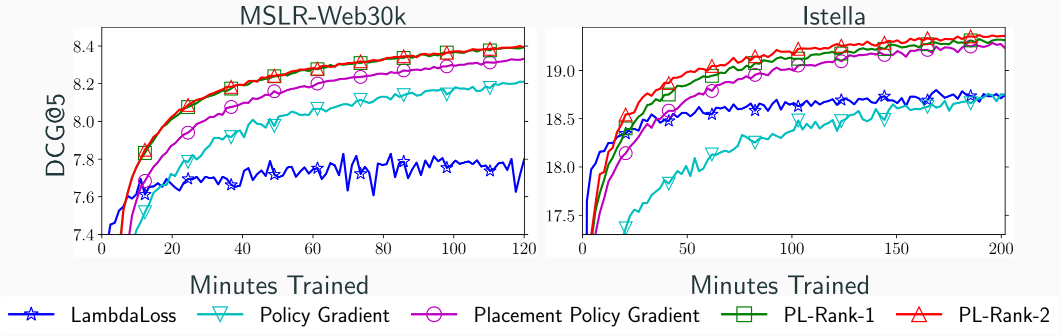
$$\begin{aligned}\frac{\delta}{\delta m}\mathcal{F}(q) = \sum_{d\in\mathcal{D}}\left[\frac{\delta}{\delta m} m(d)\right]\mathbb{E}_y\Bigg[\Bigg(&\sum_{k=\text{rank}(d,y)+1}^{K} \theta_k \left[\frac{\delta \mathcal{F}(q)}{\delta \mathcal{E}_{y_k}}\right]\Bigg) \\ &+ \sum_{k=1}^{\text{rank}(d,y)} \pi(d \,|\, y_{1:k-1})\left(\theta_k \left[\frac{\delta \mathcal{F}(q)}{\delta \mathcal{E}_d}\right] - \sum_{x=k}^{K} \theta_x \left[\frac{\delta \mathcal{F}(q)}{\delta \mathcal{E}_{y_x}}\right]\right)\Bigg].\end{aligned}$$

Experimental Results

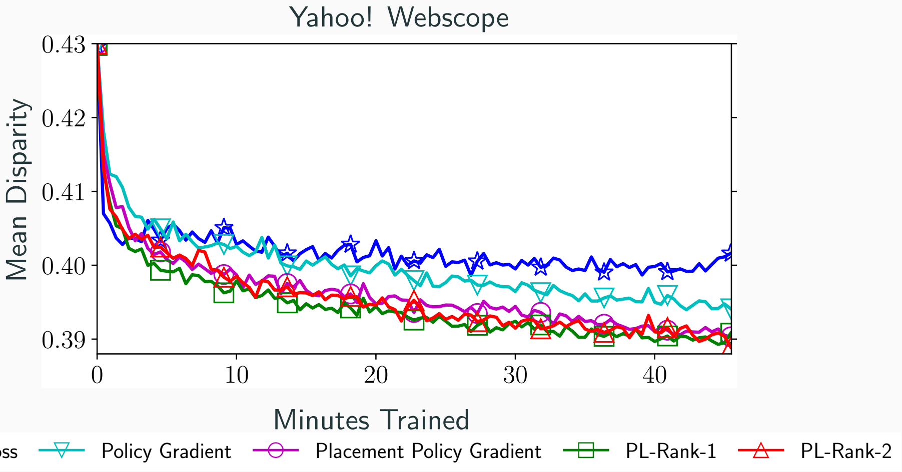
Results on Yahoo! Webscope



Results on All Datasets



Results: Exposure Fairness



Conclusion

PL-Rank: a novel LTR method for Plackett-Luce models:

- **unbiased sample-based** gradient estimation (no heuristic or bounding),
- **computationally efficient** (avoids combinatorial problems).
- applicable to **relevance** and **fairness** ranking metrics.

Continue our work: `https://github.com/HarrieO/2021-SIGIR-plackett-luce`

Future Work: StochasticRank

The **StochasticRank** algorithm (Ustimenko and Prokhorenkova, 2020) uses sampled noise to **stochastically smooth** a ranking function.

This algorithm has strong theoretical properties and could also be applied to Plackett-Luce models with comparable computational complexity.

Very promising direction for finding computationally efficient, effective and broadly applicable LTR.

A. Ustimenko and L. Prokhorenkova. Stochasticrank: Global optimization of scale-free discrete functions. In International Conference on Machine Learning, pages 9669–9679. PMLR, 2020.