

Direct Estimation of Position Bias for Unbiased Learning-to-Rank without Intervention

Grigor Aslanyan

eBay Inc.

2025 Hamilton Avenue

San Jose, CA 95125

gaslanyan@ebay.com

Utkarsh Porwal

eBay Inc.

2025 Hamilton Avenue

San Jose, CA 95125

uporwal@ebay.com

ABSTRACT

The Unbiased Learning-to-Rank framework [17] has been recently introduced as a general approach to systematically remove biases, such as position bias, from learning-to-rank models. The method takes two steps - estimating click propensities and using them to train unbiased models. Most common methods proposed in the literature for estimating propensities involve some degree of intervention in the live search engine. An alternative approach proposed recently uses an Expectation Maximization (EM) algorithm to estimate propensities by using ranking features for estimating relevances [21]. In this work we propose a novel method to estimate propensities which does not use any intervention in live search or rely on any ranking features. Rather, we take advantage of the fact that the same query-document pair may naturally change ranks over time. This typically occurs for eCommerce search because of change of popularity of items¹ over time, existence of time dependent ranking features, or addition or removal of items to the index (an item getting sold or a new item being listed). However, our method is general and can be applied to any search engine for which the rank of the same document may naturally change over time for the same query. We derive a simple likelihood function that depends on propensities only, and by maximizing the likelihood we are able to get estimates of the propensities. We apply this method to eBay search data to estimate click propensities for web and mobile search. We also use simulated data to show that the method gives reliable estimates of the “true” simulated propensities. Finally, we train a simple unbiased learning-to-rank model for eBay search using the estimated propensities and show that it outperforms the baseline model (which does not correct for position bias) on our offline evaluation metrics.

ACM Reference format:

Grigor Aslanyan and Utkarsh Porwal. 2016. Direct Estimation of Position Bias for Unbiased Learning-to-Rank without Intervention. In *Proceedings of Under Review*, , 2018, 10 pages.
DOI: 10.1145/nnnnnnnn.nnnnnnn

¹Here an “item” is the same as a “document”. In this work we use the terms “item” and “document” interchangeably.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Under Review.

© 2018 Copyright held by the owner/author(s). 978-x-xxxx-xxxx-x/YY/MM...\$15.00
DOI: 10.1145/nnnnnnnn.nnnnnnn

1 INTRODUCTION

Modern search engines rely on machine learned methods for ranking the matching results for a given query. Training and evaluation of models for ranking is commonly known as **Learning-to-Rank (LTR)** [12]. There are two common approaches for collecting the data for LTR - **human judgements** and **implicit user feedback**. For human judgements samples of documents are gathered for a sample of queries and sent to human judges who analyze and label each document. The labels can be as simple as *relevant* vs. *not relevant* or can involve more levels of relevance. This labeled data is then used for training and/or evaluation of LTR models. Collecting human judged data can be expensive and time consuming and often infeasible. On the other hand, data from implicit user feedback, such as clicks, is essentially free and abundant. For that reason it is often the preferred method for collecting data for LTR. A major drawback of this method is that the data can be heavily biased. For example, users can only click on documents that have been shown to them (presentation bias) and are more likely to click on higher ranked documents (position bias). A lot of work in the LTR literature has focused on accounting for and removing these biases. In particular, the recent paper by Joachims et. al. [17] has proposed a framework for systematically removing the biases from user feedback data. Following the title of the paper we will refer to this framework as **Unbiased Learning-to-Rank**. In particular, the authors have focused on removing the position bias by first estimating the click propensities and then using the inverse propensities as weights in the loss function. They have shown that this method results in an unbiased loss function and hence an unbiased model.

Unbiased Learning-to-Rank is a very appealing method for removing the inherent biases. However, to be able to successfully apply it one needs to first get a reliable estimate of click propensities. The method proposed in [17] uses result randomization in the live search engine to estimate click propensities. Result randomization can negatively impact the quality of the search results, which will in turn result in poor user experience and potential loss of revenue for the company [21]. It also adds bookkeeping overhead.

Wang et. al. [21] have proposed an alternative method for estimating click propensities without result randomization. The method uses a regression-based Expectation Maximization (EM) algorithm to estimate the propensities. This method uses the ranking features to estimate relevances and has shown promising results for personal search. The only drawback of this method is the use of the ranking features for estimating the relevances. In the absence of good features this algorithm may not be able to produce good estimates of the propensity.

In this paper we propose a novel method for estimating click propensities without any intervention in the live search results page, such as result randomization. We use query-document pairs that appear more than once at different ranks to estimate click propensities. Note that each query-document pair does not have to appear a large number of times - it can be as few as twice. In comparison to the EM based algorithm in [21] our method does not rely on any ranking features to estimate the relevances. In fact, we completely eliminate the relevances from the likelihood function and directly estimate the propensities by maximizing a simple likelihood function. The main novelty of our work is that our method does not use any interventions in the live search engine and does not rely on indirect estimates of relevance through ranking features. Here we focus on eCommerce search as an example since it is natural for ranking to change over time for eCommerce platforms due to change of popularity of items over time, existence of time-dependent ranking features, and addition and removal of items to the index (items being sold or new items being listed). However, the method developed here is general and can be applied to any platform for which ranking can naturally change over time.

Our method can be further generalized to drop the requirement of the ranking to naturally change over time, which will make it applicable to practically any search engine. Our likelihood function can be used in a Bayesian setting for query-document pairs which did not necessarily appear at multiple different ranks. With appropriate prior functions one can use a Bayesian approach (for example, through Gibbs sampling) to estimate posteriors for the propensities. We plan to work on this Bayesian approach in a future study.

We use simulated data to test our method and get good results. We then apply our method on actual data from eBay search logs to estimate click propensities for both web and mobile search. Finally, we use our estimated propensities to train a simple unbiased learning-to-rank model for eBay search and compare it with a baseline “biased” model which does not account for position bias. Our results show that the unbiased model with our estimated propensities significantly outperforms the baseline on our offline evaluation metrics.

The rest of the paper is organized as follows. In Section 2 we discuss some of the related work in the literature. In Section 3 we give a brief summary of the Unbiased Learning-to-Rank framework. In Section 4 we introduce our method for estimating click propensities. In Section 5 we apply our method to eBay search logs and estimate propensities for web and mobile search. In Section 6 we apply our method to simulated data and show that we are able to obtain very good estimates of the “true” simulated propensities. In Section 7 we train and evaluate simple proof-of-concept learning-to-rank models for eBay search using our estimated propensities and show that the unbiased model outperforms the baseline. Finally, we summarize our work in Section 8 and discuss possible future directions for this research.

2 RELATED WORK

Implicit feedback such as clicks are commonly used to train user facing machine learned systems such as ranking or recommender

systems. Clicks are preferred over human judged labels as they are available in plenty, are available readily and are collected in a natural environment. However, such user behavior data can only be collected over the items shown to the users. This injects a presentation bias in the collected data. This affects the machine learned systems as they are trained on user feedback data as positives and negatives. It is not feasible to present many choices to the user and it affects the performance of these systems as we can not get an accurate estimate of positives and negatives for training with feedback available only on selective samples. This situation is aggravated by the fact that the feedback of the user not only depends on the presentation, it also depends on where the item was presented. This is a subclass of the presentation bias called position bias. Joachims et al. [17] proved that if the collected user behavior data discounts the position bias accurately then the learned system will be the same as the one learned on true relevance signals.

Several approaches have been proposed to de-bias the collected user behavior data. One of the most common approaches is the use of click models. Click models are used to make hypotheses about the user behavior and then the true relevance is estimated by optimizing the likelihood of the collected clicks. There are several types of click models. One such model is a random click model (RCM) [9] where it is assumed that every document has the same probability of getting clicked and that probability is the model parameter. In a rank based click through rate model (RCTR) it is assumed that the probability of every document being clicked depends on its rank. Therefore, the total number of model parameters is the total number of ranks in the ranking system. Another model is the document based CTR model (DCTR) [8] where the click through rates are estimated for each query-document pair. In this model the total number of model parameters is the total number of query-document pairs. This model is prone to overfitting as the number of parameters grows with the training data size. Most commonly used click models are the position based model (PBM) [8, 16] and the cascade model (CM) [8]. In PBM the hypothesis is that a document is only clicked if it is observed and the user found it attractive or relevant. In CM the hypothesis is that the user sequentially scans the whole document top to bottom and clicks when the document is found to be relevant. In this model the top document is always observed and consecutive documents are only observed if the previous ones were observed and were not deemed relevant. In our proposed method we make a similar hypothesis such as the position based method where the observation probability depends on the rank and the probability of relevance only depends on the query-document pair. However, our approach is to learn the click propensities instead of learning the true relevance by optimizing the likelihood of the collected clicks. More advanced click models, such as the user browsing model (UBM) [9], the dependent click model (DCM) [11], the click chain model (CCM) [10], and the dynamic Bayesian network model (DBN) [6] are also proposed. Chuklin et al. [7] provides a comprehensive overview of click models.

Click models are trained on the collected user behavior data. Interleaving is another option that is deployed at the time of data collection. In interleaving different rank lists can be interleaved together and presented to the user. By comparing the clicks on the swapped results one can learn the unbiased user preference.

Different methods for interleaving have been proposed. In the balanced interleave method [15] a new interleaved ranked list is generated for every query. The document constraint method [13] accounts for the relation between documents. Hofmann et al. [14] proposed a probabilistic interleaving method that addressed some of the drawbacks of the balanced interleave method and the document constraint method. One limitation of the interleaving method is that often the experimentation platform in eCommerce companies is not tied to just search. It supports A/B testing for all teams, such as checkout and advertisements. Therefore, the interleaving ranked list may not be supported as it is pertinent only for search ranking.

A more recent approach to address presentation bias is the unbiased learning-to-rank approach. In this click propensities are estimated and then the inverse propensities are used as weights in the loss function. Click propensities are estimated by presenting the same items at different ranks to account for click biases without explicitly estimating the query-document relevance. Click propensity estimation can either be done randomly or in a more principled manner. Radlinski et al. [18] presented the FairPairs algorithm that randomly flips pairs of results in the ranking presented to the user. They called it randomization with minimal invasion. Carterette et al. [4] also presented a minimally invasive algorithm for offline evaluation. Joachims et al. [17] proposed randomized intervention to estimate the propensity model. Radlinski et al. [19], on the other hand, proposed alteration in ranking in a more informed manner using Multi-Armed Bandits. The main drawback of randomization for propensity estimation is that it can cause bad user experience, book keeping overhead, and a potential loss in revenue. Wang et al. [21] proposed a method to estimate propensities without randomization using the EM algorithm. In most of the existing methods, propensity estimation is done first. Once the propensities are learned, an unbiased ranker is trained using the learned propensities. Recently Ai et al. [1] proposed a dual learning algorithm that learns an unbiased ranker and the propensities together.

3 UNBIASED LEARNING-TO-RANK

In this section we introduce our notation and give a summary of the unbiased learning to rank methodology from [17]. Note that our notation differs from that of [17].

First, let us assume that we have unbiased data. The data consists of a sample \mathbf{Q} of i.i.d. (independent and identically distributed) queries. Each query $q_i \in \mathbf{Q}$ comes with a set of documents \mathbf{Y}_i with their known relevances. We denote the documents in \mathbf{Y}_i by y_{ij} and their corresponding relevances by r_{ij} . We would like to train a LTR model which is a function $F(q, y)$ that computes a score for a given query q and document y . These scores can then be used to rank the documents (higher scored documents will be ranked higher). The model is trained to minimize a *loss function* for the training data. The loss function will have smaller values for ranking functions which give higher ranks to relevant documents and lower ranks to irrelevant documents in the training data. There are multiple approaches to choosing a model F and a loss function.

There are three main classes of LTR models - **pointwise**, **pairwise**, and **listwise** [3, 12]. Pointwise LTR models use document level loss functions, regardless of the query. This means that the

model will try to produce scores for each document that is as close to its relevance as possible. Pairwise models, on the other hand, use pairs of documents with desired orders. For example, pairs of documents could contain one relevant and one irrelevant document for the query. The model tries to produce ranking scores that minimize the number of out-of-order pairs. In other words, it does not matter so much if the scores are close to relevances, as long as the ranking order for each pair of relevant and irrelevant documents is correct. Finally, listwise models take the whole query into account. Instead of using pairs the models will use the full list of documents for each query and will try to rank them as closely as possible to their desired order.

We assume that the loss function takes the form:

$$\Delta(F) = \sum_{y_{ij}} \delta(y_{ij}|F) \quad (1)$$

where the sum is taken over all documents y_{ij} in the training set, and δ denotes a document level loss function. A simple example of a quadratic loss function would be $\delta_{\text{quad}}(y_{ij}|F) = (r_{ij} - F(y_{ij}))^2$, where r_{ij} is the relevance of document y_{ij} . Another popular choice is the cross-entropy loss function if the relevances are binary (0 or 1): $\delta_{CE}(y_{ij}|F) = -r_{ij} \log F(y_{ij}) - (1 - r_{ij}) \log(1 - F(y_{ij}))$.

Note that the form (1) does not restrict the loss function to be pointwise - it can also be pairwise or listwise. In fact, it has been shown in the literature that the unbiased learning-to-rank framework works better for pairwise/listwise models than pointwise ones [17, 21]. We will discuss the reasons behind this later in this section. A popular example of a listwise loss function (and one we will use later for our models) is the DCG (Discounted Cumulative Gain):

$$\text{DCG} = \sum_{i,j} \frac{\text{rel}_{ij}}{\log_2(j+1)} \quad (2)$$

where i is the index of the query, j is the rank of a given document and rel_{ij} is its relevance, and the sum is taken over all queries and all documents for each query.

If we had data for a fair sample of queries and a fair sample of documents for these queries then minimizing the loss function (1) above would result in an unbiased model (this is known in the literature as Empirical Risk Minimization [20]). However, the data is often biased. For example, if we use click logs to determine relevances (a click on the document means the document is relevant, and no click means irrelevant) then we will only have data for documents that the users have actually seen. Using only that data for training will introduce a bias since some documents are more likely to be seen by a user than others. For example, documents that were ranked highly for a given query are more likely to be seen and receive clicks than documents ranked at lower positions (position bias). If we only use the data for documents for which the relevances have been revealed to us (e.g. the user has seen the document and decided to click or not click) we will end up with a biased model. On the other hand, we have no way of including the data for which the relevances have not been revealed. The Unbiased learning-to-rank methodology [17] introduces a modification to the loss function (1) such that it becomes an unbiased estimator for the true loss even if the data is biased. The requirement is that we should know the probabilities of observing the relevances for all of the documents in the data (in the context of this

paper observing the relevance means that the user has examined the document and decided to click on it or not). In other words, for each document y_{ij} in the training data we know the probability $p(y_{ij})$ of the relevance of that document being observed. This probability is commonly referred to as the **propensity**. For example, position bias would imply that $p(y_{ij})$ is larger for documents that had higher ranks when shown to the user compared to the ones at lower ranks. If the propensities are known then an unbiased estimator of the loss function (1) is

$$\hat{\Delta}(F) = \sum_{y_{ij}} \frac{o(y_{ij})\delta(y_{ij}|F)}{p(y_{ij})} = \sum_{y_{ij}: o(y_{ij})=1} \frac{\delta(y_{ij}|F)}{p(y_{ij})}. \quad (3)$$

Here $o(y_{ij})$ denotes if the document y_{ij} has been *observed* (1 if it has been observed, 0 otherwise). The document being observed is equivalent to the relevance being revealed to us. The equation above only includes data that has been observed, so it can be used in practice. To show that (3) is an unbiased estimator of (1) we compute the expected value of (3):

$$\begin{aligned} \mathbb{E}_{o(y_{ij})}[\hat{\Delta}(F)] &= \sum_{y_{ij}} \mathbb{E}_{o(y_{ij})} \left[\frac{o(y_{ij})\delta(y_{ij}|F)}{p(y_{ij})} \right] \\ &= \sum_{y_{ij}} \frac{p(y_{ij})\delta(y_{ij}|F)}{p(y_{ij})} \\ &= \sum_{y_{ij}} \delta(y_{ij}|F) \\ &= \Delta(F). \end{aligned} \quad (4)$$

The unbiased loss function (3) can be used if we know all of the observed documents, as well as their propensities. However, in practice it might be hard to know all of the observed documents. We know for sure that clicked documents have been observed, but we may not have information about documents that have not been clicked. One approach is to assume that all of the documents have been observed up to the lowest ranked clicked document. However, this may be inaccurate since users sometimes skip certain documents and scroll down. Also, the users may observe more documents beyond the last clicked one without clicking. For this reason it is more desirable to have a loss function that includes only “relevant” documents, such as clicked ones. This is why pairwise/listwise models work better than pointwise ones for unbiased learning-to-rank [17, 21]. For example, the DCG loss function (2) only includes relevant documents, assuming $\text{rel}_{ij} = 0$ for irrelevant ones.

In this paper we will focus on position bias. We will assume that the probability of observing a document depends only on the original rank of the document (i.e. the rank at which the document was presented to the user at the time of data collection). We denote the propensity at rank i by p_i for $i \in [1, R_{\max}]$, where R_{\max} is the maximum rank used in the dataset. If these propensities are known then the Unbiased learning-to-rank method described above can be used to remove the position bias from the LTR model.

4 PROPENSITY ESTIMATION METHOD

The method proposed by Joachims et. al. [17] for estimating click propensities is running an experimental intervention in the live search engine, where the documents at two selected ranks are swapped.

By comparing the click through rates at these ranks before and after swapping one can easily estimate the ratios of propensities at these ranks. This process can be repeated for pairs of all different ranks to estimate the ratios of all of the propensities. For the loss function (3) it is sufficient to compute the propensities up to a constant multiplier. Therefore, estimating ratios of all of the propensities is enough for the purpose of training an unbiased LTR model.

Here we propose a novel methodology for estimating click propensities without any intervention. For some search engines, especially in eCommerce, the same query-document pair may naturally appear more than once at different ranks. Using the click data on such documents we can accurately estimate click propensities. It is not required that the same query-document pair should appear at different ranks a large number of times. It is sufficient to have only two data points for the same query-document pair, as long as they are at different ranks.

We model clicks by the following simple model (also used in [17]) - *The probability of a click on a given document is the product of the probability of observing the document and the probability of clicking on the document for the given query assuming that it has been observed*. We assume that the probability of observing a document depends only on its rank and the probability of clicking on the document for a given query if it is observed depends only on the query and the document. Mathematically:

$$\begin{aligned} p(c = 1|q, y) &= p(o = 1|q, y)p(c = 1|q, y, o = 1) \\ &= p(o = 1|\text{rank}(y))p(c = 1|q, y, o = 1) \\ &= p_{\text{rank}(y)}p(c = 1|q, y, o = 1) \end{aligned} \quad (5)$$

where q denotes a query, y denotes a document, c denotes a click (0 or 1), o denotes observation (0 or 1), and p_i denotes the propensity at rank i .

Let us assume that our data D consists of N query-document pairs x_j for $j \in [1, N]$. For a query-document pair x_j we will denote the probability of clicking on the document after observing it by z_j . For each query-document pair x_j we have a set of ranks r_{jk} where the document has appeared for the query, and clicks c_{jk} denoting if the document was clicked or not (1 or 0) when it appeared at rank r_{jk} , for $k \in [1, m_j]$. Here we assume that the query-document pair x_j has appeared m_j separate times. For now we do not assume that m_j must be greater than 1 - it can be any positive integer.

The probability of a click for query-document pair x_j where the document appeared at rank r_{jk} is, according to (5):

$$p(c = 1) = p_{r_{jk}}z_j. \quad (6)$$

It follows that

$$p(c = 0) = 1 - p_{r_{jk}}z_j. \quad (7)$$

We can now introduce the following likelihood function:

$$\mathcal{L}(p_i, z_j|D) = \prod_{j=1}^N \prod_{k=1}^{m_j} \left[c_{jk}p_{r_{jk}}z_j + (1 - c_{jk})(1 - p_{r_{jk}}z_j) \right]. \quad (8)$$

Here the parameters are the propensities p_i and the “relevances” z_j (relevance here means probability of clicking for a given query-document pair assuming that the document has been observed). Theoretically, the parameters can be estimated by maximizing the likelihood function above. However, this can be challenging due

to the large number of parameters z_j . In fact, we are not even interested in estimating the z_j - we only need to estimate the propensities p_i , and the z_j are nuisance parameters.

There are multiple approaches that one can take to estimate the propensities depending on the data itself. First of all, let us consider the query-document pairs that appeared only at one rank (either the query-document pair appeared only once or multiple times but always at the same rank). The parameters p_i and z_j appear only as a product of each other in the likelihood function above. If a query-document pair x_j appeared only at rank i then we will only get terms like $p_i z_j$. So these query-document pairs could be helpful in estimating the product of the propensity at the rank that they appeared at and the relevance z_j but not each one individually. With z_j unknown, this would not help to estimate the propensity at all. However, we should mention that in the presence of a reliable prior for z_j and/or p_i the likelihood function above can be used even for those query-document pairs that appeared only at one rank. In this case it would be more useful to take a Bayesian approach and estimate the posterior distribution for the propensities, for example using Gibbs sampling [5]. That study is beyond the scope of this work and will be explored by the authors in a future study.

From now on we will assume that the query-document pairs appear at least at two different ranks. Another extreme is the case when each query-document pair appears a large number of times at different ranks. In this case the estimation of the propensities (or more precisely their ratios) could be much simpler than maximizing the likelihood function above. For a pair of fixed ranks i and j one can take all query-document pairs that appeared at both of those ranks, then simply count the total number of clicks received at rank i and rank j and take their ratio. That would give a good estimate for the ratio p_i/p_j . We prove this in subsection 4.1 below. Of course, for that estimate to be reliable one would need a large number of query-document pairs that appeared at both of these ranks. But if our data consists of query-document pairs that appeared a large number of times at different ranks then there should be enough data to estimate the ratios of propensities for the top ranks at least.

Let us now consider the case when the data consists of a large number of query-document pairs that appeared a few times (can be as few as twice) at different ranks, but the query-document pairs do not appear a large enough number of times to be able to estimate the propensity ratios by simply counting clicks as described above. In this case we will actually need to maximize the likelihood above and somehow eliminate the nuisance parameters z_j to get estimates for the p_i . To the best of our knowledge this case has not been studied in the literature before and we will focus the rest of this work on this case. Also, the data we have collected from eBay search logs falls in this category, as will be discussed in Section 5.

If a query-document pair appeared only a few times there is a good chance that it did not receive any clicks. These query-document pairs will not help in estimating the propensities by likelihood maximization because of the unknown parameter z_j . Specifically, for such query-document pairs we will have the terms

$$\prod_{k=1}^{m_j} (1 - p_{r_{jk}} z_j).$$

If we use the maximum likelihood approach for estimating the parameters then the maximum will be reached by $z_j = 0$ for which the terms above will be 1. That means that these query-document pairs without any clicks will not change the maximum likelihood estimate of the propensities. For that reason we will only keep query-document pairs that received at least one click. However, we cannot simply drop the terms from the likelihood function for query-document pairs that did not receive any clicks. Doing so would bias the data towards query-document pairs with a higher likelihood of click. Instead, we will replace the likelihood function above by a conditional probability. Specifically, the likelihood function (8) computes the probability of the click data $\{c_{jk}\}$ obtained for that query-document pair (i.e. the clicks the query-document pair received and did not receive at different ranks). If we are keeping only the query-document pairs that received at least one click then we need to replace that probability by a conditional probability - the probability of the click data $\{c_{jk}\}$ under the condition that there was at least one click received: $\sum_k c_{jk} > 0$. The likelihood function for the query-document pair x_j will take the form:

$$\begin{aligned} \mathcal{L}_j(p_i, z_j | D_j) &= P\left(D_j \mid \sum_k c_{jk} > 0\right) \\ &= \frac{P(D_j \cap \sum_k c_{jk} > 0)}{P(\sum_k c_{jk} > 0)} \\ &= \frac{P(D_j)}{P(\sum_k c_{jk} > 0)} \\ &= \frac{\prod_{k=1}^{m_j} [c_{jk} p_{r_{jk}} z_j + (1 - c_{jk})(1 - p_{r_{jk}} z_j)]}{1 - \prod_{k=1}^{m_j} (1 - p_{r_{jk}} z_j)}. \end{aligned} \quad (9)$$

Here \mathcal{L}_j denotes the likelihood function for the query-document pair x_j , $D_j = \{c_{jk}\}$ denotes the click data for query-document pair j , and P denotes probability. $\sum_k c_{jk} > 0$ simply means that there was at least one click. In the first line above we have replaced the probability of data D_j by a conditional probability. The second line uses the formula for conditional probability. The probability of D_j and at least one click just equals to probability of D_j since we are only keeping query-document pairs that received at least one click. This is how the third line is derived. Finally, in the last line we have explicitly written out $P(D_j)$ in the numerator as above in (8) and the probability of at least one click in the denominator (the probability of no click is $\prod_{k=1}^{m_j} (1 - p_{r_{jk}} z_j)$ so the probability of at least one click is 1 minus that).

The full likelihood is then the product of \mathcal{L}_j for all query-document pairs:

$$\mathcal{L}(p_i, z_j | D) = \prod_{\substack{j=1 \\ \sum_k c_{jk} > 0}}^N \frac{\prod_{k=1}^{m_j} [c_{jk} p_{r_{jk}} z_j + (1 - c_{jk})(1 - p_{r_{jk}} z_j)]}{1 - \prod_{k=1}^{m_j} (1 - p_{r_{jk}} z_j)}. \quad (10)$$

From now on we will assume by default that our dataset contains only query-document pairs that received at least one click and will omit the subscript $\sum_k c_{jk} > 0$.

Our last step will be to simplify the likelihood function (10). Typically the click probabilities $p_i z_j$ are not very large (i.e. not close to 1). This is the probability that the query-document pair j will

get a click when displayed at rank i . To simplify the likelihood for each query-document pair we will only keep terms linear in $p_i z_j$ and drop higher order terms like $p_{i_1} z_{j_1} p_{i_2} z_{j_2}$. As discussed later in Section 5, this assumption is valid for our data. In general, we expect this assumption to be valid for most search engines. It is certainly a valid assumption for lower ranks since click through rates are typically much smaller for lower ranks. One would simply need to verify this assumption for topmost ranks. Since we are dropping product terms the largest ones would be between ranks 1 and 2. For most search engines the click through rates at rank 2 are around 10% or below, which we believe is small enough to be able to safely ignore the product terms mentioned above (they would be at least 10 times smaller than linear terms). We empirically show using simulations in Section 6 that this assumption works very well for data similar to eBay data. If for other search engines the click through rates are much larger for topmost ranks we suggest keeping only those query-document pairs that appeared at least once at a lower enough rank. Also, using the methodology of simulations from Section 6 one can verify how well this assumption works for their particular data.

Under the simplifying assumption we get for the denominator in (10):

$$1 - \prod_{k=1}^{m_j} (1 - p_{r_{jk}} z_j) \approx 1 - \left(1 - \sum_{k=1}^{m_j} p_{r_{jk}} z_j\right) = z_j \sum_{k=1}^{m_j} p_{r_{jk}}. \quad (11)$$

Let us now simplify the numerator of (10). Firstly, since the click probabilities are not large and each query-document pair appears only a few times we can assume there is only one click per query-document pair². We can assume $c_{jl_j} = 1$ and $c_{jk} = 0$ for $k \neq l_j$. The numerator then simplifies to

$$\begin{aligned} \prod_{k=1}^{m_j} [c_{jk} p_{r_{jk}} z_j + (1 - c_{jk})(1 - p_{r_{jk}} z_j)] &= p_{r_{jl_j}} z_j \prod_{\substack{k=1 \\ k \neq l_j}}^{m_j} (1 - p_{r_{jk}} z_j) \\ &\approx p_{r_{jl_j}} z_j. \end{aligned} \quad (12)$$

Using (11) and (12) the likelihood function (10) simplifies to

$$\mathcal{L}(p_i, z_j | D) = \prod_{j=1}^N \frac{p_{r_{jl_j}} z_j}{z_j \sum_{k=1}^{m_j} p_{r_{jk}}} = \prod_{j=1}^N \frac{p_{r_{jl_j}}}{\sum_{k=1}^{m_j} p_{r_{jk}}}. \quad (13)$$

In the last step z_j cancels out from the numerator and the denominator. Our assumption of small click probabilities, together with keeping only query-document pairs that received at least one click allowed us to simplify the likelihood function to be only a function of propensities. Now we can simply maximize the likelihood (13) to estimate the propensities.

Eq. (13) makes it clear why we need to include the requirement that each query-document pair should appear more than once at different ranks. If we have a query-document pair that appeared only once (or multiple times but always at the same rank) then the numerator and the denominator would cancel each other out in

(13). For that reason we will keep only query-document pairs that appeared at two different ranks at least.

It is numerically better to maximize the log-likelihood, which takes the form:

$$\log \mathcal{L}(p_i | D) = \sum_{j=1}^N \left(\log(p_{r_{jl_j}}) - \log \sum_{k=1}^{m_j} p_{r_{jk}} \right). \quad (14)$$

Let us summarize our method of estimating propensities. Our data consists of query-document for which the document appeared at more than one rank and got at least one click. We then maximize the log-likelihood (14) to estimate the propensities. The log-likelihood is a sum over all query-document pairs in our dataset. For each query-document pair we take the log of the propensity for the rank at which the document got clicked (assuming there is only one) minus the log of the sum of the propensities for all the ranks where the document appeared. The only assumption we made in order to get the simple form of (14) is that the click probabilities are not large. As we will see in Section 5 this is a reasonable assumption for eBay search. Our methodology of simulations in Section 6 can be used to verify the validity of the assumption. We have also discussed alternative approaches for estimating the click propensities for cases when the assumption might not work very well.

4.1 Propensity Ratio Estimation

Here we consider the case when for two fixed ranks i and j a large number of query-document pairs appear at both of these ranks. As mentioned in Section 4 one can simply compute the number of clicks from those query-document pairs for each rank and take the ratio of those numbers to estimate the ratio p_i/p_j . We prove that statement below.

Let us assume that query-document pairs $\{x_k, k = 1, \dots, K\}$ appeared for both ranks i and j . We will first assume that these query-document pairs appeared exactly once for each rank. We will later relax that assumption. The probability of a click at rank i for query-document pair x_k is $p_i z_k$ where z_k is the “relevance” for query-document pair x_k (i.e. the probability of a click under the assumption that it was observed). So the expected number of clicks N_i for rank i for all query-document pairs will be:

$$\mathbb{E}[N_i] = \sum_{k=1}^K p_i z_k = p_i \sum_{k=1}^K z_k.$$

Similarly:

$$\mathbb{E}[N_j] = \sum_{k=1}^K p_j z_k = p_j \sum_{k=1}^K z_k.$$

Taking the ratio³:

$$\frac{\mathbb{E}[N_i]}{\mathbb{E}[N_j]} = \frac{p_i \sum_{k=1}^K z_k}{p_j \sum_{k=1}^K z_k} = \frac{p_i}{p_j}.$$

Replacing the expected values of N_i and N_j above by their actual observed values we can get an estimator for p_i/p_j .

²This is true for our data as discussed later in Section 5. For the cases when most query-document pairs receive multiple clicks we suggest using a different method, such as computing the ratios of propensities by computing the ratios of numbers of clicks as discussed above

³Note that here we are taking the ratio of expected values rather than the expected value of the ratio. While in general they are not the same, to the first order approximation they are the same. We have empirically verified through simulations that taking the ratio of the observed values N_i/N_j asymptotically approaches the ratio p_i/p_j as the total number of query-document pairs is increased.

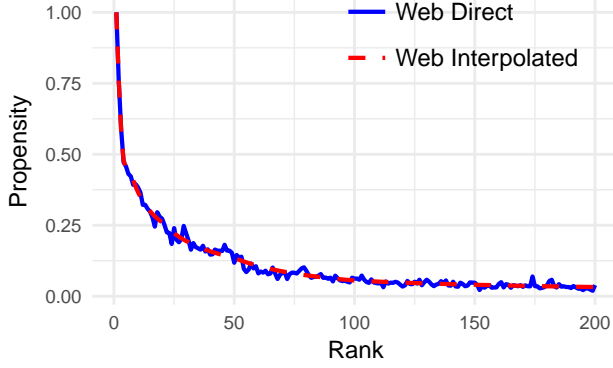


Figure 1: Click propensity estimated for eBay search for web data. The solid blue line is the direct estimation of propensities for each rank, and the red dashed line is the estimation using interpolation.

So far we have assumed that each query-document pair appears exactly once at each rank. The above proof can be easily extended to the case when the same query-document pair appears multiple times at each rank. In this case instead of counting clicks directly we will count clicks per impressions, i.e. for each query-document pair we will count the total number of clicks and divide by the total number of impressions. The expected value for clicks per impressions is still the click probability $p_i z_k$. Then we will take the sum for clicks per impressions for each rank and take the ratio. The derivation will remain exactly the same. The only difference is that N_i and N_j will now denote the sum of clicks per impressions instead of the total number of clicks.

5 CLICK PROPENSITIES FOR EBAY SEARCH

We now apply the method developed above on eBay search data to estimate propensities. We collected a small sample (0.2%) of queries for four months of eBay search traffic. For each query we keep the top 500 items (as mentioned before, we use the terms “item” and “document” interchangeably). There are multiple sort types on eBay (such as Best Match, Price Low to High, Time Ending Soonest) and click propensities may differ for different sort types. In this paper we present our results on Best Match sort, and hence we keep only queries for that sort type. Furthermore, there are multiple different platforms for search (such as a web browser or a mobile app) which can have different propensities. We separate our dataset into two platforms - web and mobile, and estimate click propensities for each platform separately. For web queries the users can choose the number of items shown per page. The most common choice (and the default one) is 50 items per page. Click propensities might differ depending on the number of items per page. For web search we estimate the propensities for 50 items per page with list view, and filter our dataset accordingly.

Next, we must identify same query-document pairs and find cases where the document appeared at multiple different ranks. However, we would like to avoid the cases where some detail about the item (such as price) has changed between the queries, which could result in a different probability of a click on that item. eBay has inventory of both auction and fixed price items. Since the

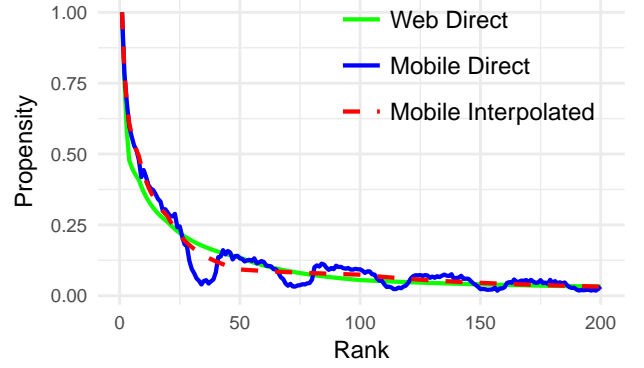


Figure 2: Click propensities for eBay search for mobile data. The solid blue line is the direct estimation at each rank and the red dashed line is the estimation using interpolation. For comparison, we also plot the propensities for web data using interpolation in solid green, which is the same as the red dashed line from Fig. 1.

auction item properties (like current bid amount) change very frequently we exclude auction items from our dataset and keep only fixed price items. Furthermore, we check to make sure that the price of the item has not changed from query to query. We also keep the same query-item pairs from the same day only. This is a further caution to make sure that seasonality effects do not affect the popularity of the item. For the query side we identify two queries to be the same if they have the same keywords, as well as the same category and aspect (such as color, size) constraints. We then keep only those query-document pairs that appeared at two different ranks and got one click in one rank and no click in the other (in our data it is rare for the same query-document pair to appear at more than one rank and get more than one click).⁴

We first check our assumption of small click probabilities. We compute the click through rates at ranks 1-500. As expected, rank 1 has the highest click through rate, but it is still under 10%. For lower ranks it is even smaller. This does justify our assumption of small click probabilities and we can apply our method above to estimate propensities.

We first estimate propensities for web queries. Our dataset consists of about 40,000 query-item pairs, each of which appeared at two different ranks and received a click at one of the ranks. We use two methods for estimating propensities - direct and interpolation. In the direct method we treat the propensity at each rank as a separate parameter. We therefore get 500 different parameters to estimate. In the interpolation method we fix a few different ranks and use the propensities at those ranks as our parameters to estimate. The propensities for all the other ranks are computed as a linear interpolation in the log-log space, i.e. we approximate the log of the propensity as a linear function of the log of the rank. This results in the propensity being a power law of the rank. For the interpolation method our fixed ranks are 1, 2, 4, 8, 20, 50, 100, 200,

⁴Note that keeping only query-document pairs that appeared at two ranks exactly is in no way a requirement of our method. The method is general and can be used for query-document pairs that appeared more than twice. This is just intended to simplify our analysis without a significant loss in data, since it is rare for the same query-document pair to appear at more than two ranks.

300, and 500. We choose a denser grid for higher ranks since there is more data and less noise for higher ranks, and the propensities can be estimated more accurately.

Our resulting propensity for web search is shown in Fig. 1. The solid blue line shows the propensities estimated through the direct method, and the red dashed curve shows the propensities estimated through interpolation. Even though we estimate propensities up to rank 500, we plot them only up to rank 200 so that the higher ranks can be seen more clearly. The red dashed curve passes smoothly through the blue solid curve, which is reassuring. Note that the red dashed curve is not a fit to the blue one. The two are estimated directly from the data. For the blue curve the parameters are all of the propensities at each rank, whereas for the red dashed curve we only parametrize the propensities at select ranks and interpolate in between. We then maximize the likelihood for each case to estimate the parameters. The fact that the red dashed line appears to be a smooth fit to the solid blue shows that the interpolation method is useful in obtaining a smooth and less noisy propensity curve which is still very close to the direct estimation.

The propensities estimated from eBay mobile search data are shown in Fig. 2. As in Fig. 1, the blue solid curve shows direct estimation, and the red dashed curve is estimation using interpolation. For comparison, we plot the propensities from web using interpolation in solid green. The blue solid curve shows a certain periodicity - the propensities seem to drop sharply near rank 25, then go back up at rank 40, drop again around rank 65, then back at rank 80, and so on. In fact, this reflects the way results are loaded in mobile search - 40 at a time. The blue curve seems to indicate that users observe the results at higher ranks with the usual decrease in interest, then they tend to scroll faster to the bottom skipping the results towards the bottom, then as the new batch is loaded they again gain more interest and it continues decreasing as they scroll down and the pattern repeats itself every 40 ranks. The red dashed curve matches the blue one reasonable well, but it fails to capture the periodic dips. This is due to our choice of knots for the linear spline. One can use the blue curve to choose new locations of the knots to be able to get a better interpolation for the propensities. The green solid curve matches fairly well with the blue one except for the dips. This means that the propensities for web and mobile are very similar, except for the periodic dips for mobile. The web results are shown 50 items per page, but we have not found any periodic dips for web search. Perhaps this indicates that for web search users do not tend to scroll quickly towards the end of the page and then regain interest as a new page is loaded. The smooth decline in propensities indicates that for web search users steadily lose interest as they scroll down, but the number of items per page does not affect their behavior. If they make it to the end of the page then they load the next page and continue.

6 RESULTS ON SIMULATIONS

In this section we use simulated data to verify that the method of estimating propensities developed in Section 4 works well. For our simulations we choose the following propensity function as truth:

$$p_i^{\text{sim}} = \min\left(\frac{1}{\log i}, 1\right) \quad (15)$$

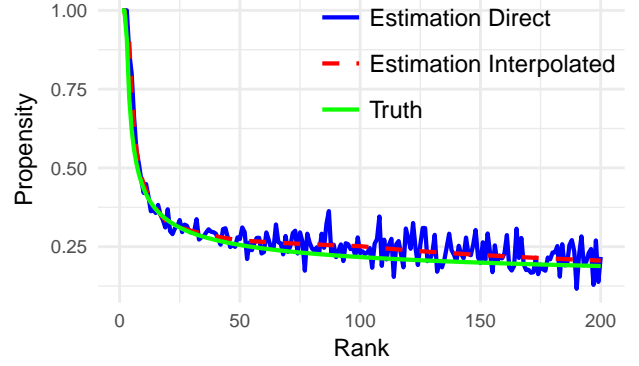


Figure 3: Propensity estimated from simulated data. The green solid curve shows the “true” propensity (15). The blue solid curve is the estimated propensity using the direct estimation method. The red dashed curve is the estimation using interpolation.

which assigns propensity of 1 for ranks 1 and 2, and then decreases as the inverse of the log of the rank.

Other than choosing our own version of propensities we simulate the data to be as similar to the eBay dataset as possible. We generate a large number of query-document pairs and randomly choose a mean rank $rank_{mean}$ for each query-document pair uniformly between 1 and 500. We randomly generate a click probability z for that query-document pair depending on the mean rank $rank_{mean}$. We choose the distribution from which the click probabilities are drawn such that the click through rates at each rank match closely with the click through rates for real data, taking into account the “true” propensities (15). We then generate two different ranks drawn from $\mathcal{N}(rank_{mean}, (rank_{mean}/5)^2)$. For each rank i we compute the probability of a click as zp_i^{sim} . Then we keep only those query-document pairs which got at least one click and appeared at exactly two different ranks, in agreement with our method used for real eBay data. Finally we keep about 40,000 query-document pairs so that the simulated data is similar to the eBay web search data in size. This becomes the simulated data. Note that the real data contains many query-document pairs that appeared only once and some that appeared more than twice. We do not simulate such query-document pairs since they would be eventually all filtered out. In our simple method we keep only those query document pairs which appear twice exactly at two different ranks and receive at least one click. Note that we could have also kept query-document pairs that appeared more than twice. However, as mentioned in the previous Section, in actual eBay data we have much fewer pairs like that compared to the ones that appeared twice. So for the sake of simplicity we choose to only keep pairs that appeared exactly twice, without losing much of the data.

The estimated propensities on the simulated dataset are shown in Fig. 3. The green solid curve shows the true propensity (15), the blue solid curve shows the estimated propensity using the direct estimation method, and the red dashed curve is the estimated propensity using interpolation. As we can see, the estimations closely match with the truth. Furthermore, we can see that the interpolation method gives a better result by reducing the noise

in the estimate. These results show that the propensity estimation method developed in this paper works well.

7 UNBIASED LEARNING-TO-RANK MODELS

In this section we study the improvement in ranking models by using the estimated click propensities for eBay search data. Previous studies have consistently shown that unbiased learning-to-rank models significantly improve ranking metrics compared to their biased counterparts. Specifically, Joachims et. al. [17] have shown that an unbiased learning-to-rank model significantly improves the average rank of relevant results for simulated data. Furthermore, they have performed an online interleaving experiment on a live search engine for scientific articles, which resulted in a significant improvement for the unbiased model. Wang et. al. [21] have shown an improvement in MRR (Mean Reciprocal Rank) for the unbiased learning-to-rank models for personal search.

Here we train simple proof-of concept models to check if unbiased ranking models show improvements over their biased counterparts. The training and evaluation of full production scale models is beyond the scope of this work and will be performed in the future. For our training data we collect a sample of about 40,000 queries which have received at least one click. The sample is collected from four days of search logs. We train listwise ranking models using the LambdaMART algorithm [2]. We use the DCG metric (2) as our loss function. We define rel_{ij} to be 1 if document j was clicked, and 0 otherwise. We train two models - baseline and unbiased. The baseline model uses (2) as its loss function without any correction for position bias and serves as our baseline, while the unbiased model uses inverse-propensity weighted relevances as in (3). We use the propensities estimated for eBay web search as shown in Fig. 1 red dashed curve. Our training and test data are also from web search (i.e. browser) only. We use 25 features for both models, selected from our top ranking features. We also fix the model parameters: the number of trees is 100 and the shrinkage is 0.1. Since these are simple proof-of-concept models we have not done extensive feature selection and parameter tuning.

Our test data contains a sample of about 10,000 queries from four days of eBay search logs. Since the test data also has the same position bias as the training data we cannot rely on standard ranking metrics such as DCG, NDCG (Normalized Discounted Cumulative Gain), or MRR (Mean Reciprocal Rank). Another option would be to use inverse-propensity-weighted versions of these metrics to remove the presentation bias. However, the true propensities are unknown to us and we obviously cannot use estimated propensities for evaluation since part of the evaluation is checking if our estimate of propensities is a good one. For that reason we choose a different approach for evaluation. Namely, we fix the rank of items in the test data, i.e. we select items from different queries that appeared at a given fixed rank. By selecting the items from a fixed rank in the evaluation set we effectively eliminate position bias since all of the items will be affected by position bias the same way (the observation probability is the same for all the items since the rank is the same). Then we compare the two ranking models as classifiers for those items, which means that we evaluate how well the models can distinguish items that were clicked from ones

Table 1: Performance of the unbiased ranking model compared to the baseline model. The unbiased model uses inverse propensity weighting with the propensities estimated in Section 5, while the baseline model does not correct for position bias. The validation set contains documents from a fixed rank. The first column shows that fixed rank. The second and third columns show the AUC metric for the baseline and unbiased models, respectively. The last column shows the improvement in AUC for the unbiased model compared to the baseline.

Rank	AUC Baseline	AUC Unbiased	AUC Improvement
1	0.484	0.499	3.1%
2	0.520	0.535	2.9%
4	0.527	0.549	4.2%
8	0.551	0.570	3.4%
16	0.603	0.644	6.8%
32	0.584	0.588	0.7%

that were not. We use AUC (Area Under the Receiver Operating Characteristic Curve) as our evaluation metric.

The results are presented in Table 1, where we show results for fixed ranks 1, 2, 4, 8, 16, and 32. As we can see, for all ranks the unbiased model outperforms the baseline. The best improvement is for items at fixed rank 16, with an AUC improvement of 6.8%. We also notice that for low ranks the AUC is close to 0.5 for both models, meaning that the models are unable to distinguish between items that are likely to get clicked from ones that are not. This is due to the fact that the items have been ranked with the current production model and low ranked items are the ones that have a high likelihood to get clicked. Clearly our simple models are strongly correlated with the production model. This result simply means that after the production model selects the “best” items the new models do not have a strong distinguishing power. Nevertheless, the unbiased model has a higher distinguishing power than the baseline.

8 SUMMARY AND FUTURE WORK

In this work we have introduced a new method for estimating click propensities for eCommerce search without randomizing the results during live search. Our method uses query-document pairs that appear more than once and at different ranks. Although we have used eCommerce search as our main example, the method is general and can be applied to any search engine for which ranking naturally changes over time. The clear advantage of our method over result randomization is that it does not affect live search results, which can have a negative impact on the engine as has been shown in the literature [21]. To the best of our knowledge there has been proposed only one other approach for estimating click propensities without result randomization by Wang et. al. [21]. Their method uses an EM (Expectation Maximization) algorithm to estimate the propensities and relies on ranking features for estimating and eliminating relevances. The advantage of our method over the EM method is that we estimate propensities directly and do not have to rely on ranking features. The results of the EM method strongly depend on the goodness of the features and their ability to predict the relevance of the document. For this reason

we believe that our novel approach will find widespread use for unbiased learning-to-rank modeling.

We have used simulated data to show that our method can give accurate estimates of the true propensities. We have applied our method to eBay search results to separately estimate propensities for web and mobile search. We have also trained simple proof-of-concept ranking models and compared the performance of the unbiased model using the estimated propensities to the baseline model, which does not correct for position bias. Using a validation dataset of documents from a fixed rank we have shown that the unbiased model significantly outperforms the baseline in terms of the AUC metric.

This work focuses on estimating propensities from query-document pairs that appear at multiple different ranks. We have derived a simple method for estimating propensity ratios when each query-document pair appears a large number of times at different ranks. More importantly, we have addressed the case when the same query-document pair appears only a few number of times at different ranks (can be as few as twice), which is the case for our data. This method can be generalized to use query-document pairs that appeared at a single rank only by incorporating appropriate priors and using Gibbs sampling to estimate the posterior distribution for propensities. We plan to study this approach in a future work.

Previous studies have shown that the unbiased ranking models significantly outperform their “biased” counterparts both in terms of offline ranking metrics and A/B test results [17, 21]. This work has mostly focussed on propensity estimation. Because of space limitations we have only studied simple proof-of-concept learning-to-rank models. In the future we would like to further evaluate our estimated propensities and their use in ranking models. This can include training production level unbiased learning-to-rank models with appropriate feature selection and parameter tuning. The models can be evaluated offline on human judged data (which will remove the position bias from the validation set) and online through A/B tests.

Click propensities have been estimated in this work under the assumption that they are the same for all queries and all users. In the future it would be desirable to estimate and compare propensities for different classes of queries and/or users. For example, one could compare more frequent and less frequent queries, queries from different categories (such as electronics versus fashion), as well as different user demographics. Also, in this work we have only focused on the “Best Match” sort for eBay. Propensities for other sort types, such as “Price Low to High”, would also be of interest in a future study.

REFERENCES

- [1] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W Bruce Croft. 2018. Unbiased Learning to Rank with Unbiased Propensity Estimation. *arXiv preprint arXiv:1804.05938* (2018).
- [2] Chris J.C. Burges. 2010. *From RankNet to LambdaRank to LambdaMART: An Overview*. Technical Report. <https://www.microsoft.com/en-us/research/publication/from-ranknet-to-lambdarank-to-lambdamart-an-overview/>
- [3] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to Rank: From Pairwise Approach to Listwise Approach. In *Proceedings of the 24th International Conference on Machine Learning (ICML '07)*. ACM, New York, NY, USA, 129–136. <https://doi.org/10.1145/1273496.1273513>
- [4] Ben Carterette and Praveen Chandar. 2018. Offline Comparative Evaluation with Incremental, Minimally-Invasive Online Feedback. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18)*. ACM, New York, NY, USA, 705–714. <https://doi.org/10.1145/3209978.3210050>
- [5] George Casella and Edward I. George. 1992. Explaining the Gibbs Sampler. *The American Statistician* 46, 3 (1992), 167–174. <http://www.jstor.org/stable/2685208>
- [6] Olivier Chapelle and Ya Zhang. 2009. A dynamic bayesian network click model for web search ranking. In *Proceedings of the 18th international conference on World wide web*. ACM, 1–10.
- [7] Aleksandr Chuklin, Ilya Markov, and Maarten de Rijke. 2015. Click models for web search. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 7, 3 (2015), 1–115.
- [8] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. 2008. An experimental comparison of click position-bias models. In *Proceedings of the 2008 international conference on web search and data mining*. ACM, 87–94.
- [9] Georges E Dupret and Benjamin Piwowarski. 2008. A user browsing model to predict search engine click data from past observations.. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 331–338.
- [10] Fan Guo, Chao Liu, Anitha Kannan, Tom Minka, Michael Taylor, Yi-Min Wang, and Christos Faloutsos. 2009. Click chain model in web search. In *Proceedings of the 18th international conference on World wide web*. ACM, 11–20.
- [11] Fan Guo, Chao Liu, and Yi Min Wang. 2009. Efficient multiple-click models in web search. In *Proceedings of the second acm international conference on web search and data mining*. ACM, 124–131.
- [12] Li Hang. 2011. A Short Introduction to Learning to Rank.
- [13] Jing He, Chengxiang Zhai, and Xiaoming Li. 2009. Evaluation of methods for relative comparison of retrieval systems based on clickthroughs. In *Proceedings of the 18th ACM conference on Information and knowledge management*. ACM, 2029–2032.
- [14] Katja Hofmann, Shimon Whiteson, and Maarten De Rijke. 2011. A probabilistic method for inferring preferences from clicks. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 249–258.
- [15] Thorsten Joachims. 2003. Evaluating retrieval performance using clickthrough data.
- [16] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. 2005. Accurately Interpreting Clickthrough Data As Implicit Feedback. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '05)*. ACM, New York, NY, USA, 154–161. <https://doi.org/10.1145/1076034.1076063>
- [17] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. 2017. Unbiased Learning-to-Rank with Biased Feedback. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM '17)*. ACM, New York, NY, USA, 781–789. <https://doi.org/10.1145/3018661.3018699>
- [18] Filip Radlinski and Thorsten Joachims. 2006. Minimally invasive randomization for collecting unbiased preferences from clickthrough logs.
- [19] Filip Radlinski, Robert Kleinberg, and Thorsten Joachims. 2008. Learning diverse rankings with multi-armed bandits. In *Proceedings of the 25th international conference on Machine learning*. ACM, 784–791.
- [20] Vladimir N. Vapnik. 1998. *Statistical Learning Theory*. Wiley-Interscience.
- [21] Xuanhui Wang, Nadav Golbandi, Michael Bendersky, Donald Metzler, and Marc Najork. 2018. Position Bias Estimation for Unbiased Learning to Rank in Personal Search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*. ACM, New York, NY, USA, 610–618. <https://doi.org/10.1145/3159652.3159732>