

# COPR: Consistency-Oriented Pre-Ranking for Online Advertising

Zhishan Zhao, Jingyue Gao\*  
Alibaba Group  
Beijing, China

Yu Zhang, Shuguang Han  
Alibaba Group  
Beijing, China

Siyuan Lou, Xiang-Rong Sheng  
Alibaba Group  
Beijing, China

Zhe Wang, Han Zhu†  
Alibaba Group  
Beijing, China

Yuning Jiang, Jian Xu  
Alibaba Group  
Beijing, China

Bo Zheng  
Alibaba Group  
Beijing, China

## ABSTRACT

Cascading architecture has been widely adopted in large-scale advertising systems to balance efficiency and effectiveness. In this architecture, the pre-ranking model is expected to be a lightweight approximation of the ranking model, which handles more candidates with strict latency requirements. Due to the gap in model capacity, the pre-ranking and ranking models usually generate inconsistent ranked results, thus hurting the overall system effectiveness. The paradigm of score alignment is proposed to regularize their raw scores to be consistent. However, it suffers from inevitable alignment errors and error amplification by bids when applied in online advertising. To this end, we introduce a consistency-oriented pre-ranking framework for online advertising, which employs a chunk-based sampling module and a plug-and-play rank alignment module to explicitly optimize consistency of ECPM-ranked results. A  $\Delta NDCG$ -based weighting mechanism is adopted to better distinguish the importance of inter-chunk samples in optimization. Both online and offline experiments have validated the superiority of our framework. When deployed in Taobao display advertising system, it achieves an improvement of up to +12.3% CTR and +5.6% RPM.

## KEYWORDS

pre-ranking, cascading architecture, consistency, online advertising

### ACM Reference Format:

Zhishan Zhao, Jingyue Gao, Yu Zhang, Shuguang Han, Siyuan Lou, Xiang-Rong Sheng, Zhe Wang, Han Zhu, Yuning Jiang, Jian Xu, Bo Zheng. 2023. COPR: Consistency-Oriented Pre-Ranking for Online Advertising. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 7 pages. <https://doi.org/XXXXXXX.XXXXXX>

## 1 INTRODUCTION

Online advertising has become a major source of revenue for many web platforms [4, 15]. Advertisers ensure effective promotion of

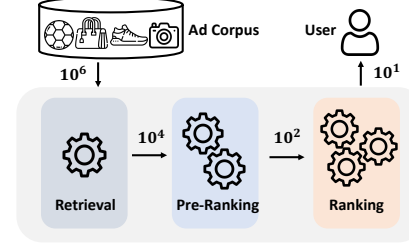


Figure 1: An illustration of the typical three-phase cascading architecture for online advertising systems.

products by bidding and paying for user actions (e.g., click and purchase)<sup>1</sup> on advertisements (i.e., ads). To maximize platform revenue, the advertising system typically ranks ads based on their Expected Cost Per Mille (ECPM) [25] and selects top ones for impression:

$$ECPM = 1000 \times bid \times pCTR, \quad (1)$$

where  $bid$  is the price that the advertiser is willing to pay and  $pCTR$  is the predicted click-through rate (CTR) denoting the probability that the user clicks the ad.

Under strict latency requirements in online deployment, it is infeasible for complex CTR models [5, 12, 18, 26] with high inference cost to handle millions of candidates in the ad corpus. To balance efficiency and effectiveness, a common practice in industrial systems is to adopt a cascading architecture [3, 10, 19, 23], which filters ads through multiple phases with increasingly complex models as illustrated in Fig. 1. Particularly, the retrieval model first retrieves tens of thousands of relevant ads from the corpus. Afterwards, the pre-ranking model outputs  $pCTR$  for retrieved candidates, where top hundreds with highest ECPM are sent to the ranking model for final selection. To handle a larger candidate set, the pre-ranking model is usually designed to be lightweight, which works more efficiently but less accurately compared with the ranking model.

Pre-Ranking has recently received increasing attention due to its importance in the cascading architecture. Huang et al. [8] propose a two-tower model that maps users and candidates into latent vectors and calculates their inner products. To enable high-order feature interactions, Li et al. [10] add fine-grained interactions between two towers and Wang et al. [23] propose to use deep neural network with squeeze-and-excitation block. Despite improvement of accuracy, there is still a non-negligible gap between the pre-ranking and ranking models. They may generate significantly different ranked

\*Zhishan Zhao and Jingyue Gao contribute equally to this work.

†Han Zhu is the corresponding author.

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

Conference acronym 'XX, June 03–05, 2023, Woodstock, NY

© 2023 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXX>

<sup>1</sup>Without loss of generality, we regard click as the action in this paper

results on the same candidate set. Such **inconsistency** hinders the overall system effectiveness. For example, **top ads selected from the pre-ranking phase could be less competitive in the ranking phase**, causing waste of the computational resource. Also, ads which are preferred in the ranking phase could be unfortunately discarded in the pre-ranking phase, leading to sub-optimal results.

Some pioneering studies [19, 20] propose to align the pre-ranking and ranking models via distillation on pCTR scores. **The pre-ranking model is encouraged to generate same scores as the ranking model** [19] or generate high scores for top candidates selected by the ranking model [20]. Although exhibiting encouraging performance, the paradigm of **score alignment** suffers from the following issues, especially when applied to the advertising system:

- **Inevitable alignment errors.** Due to simpler architecture and fewer parameters for efficiency concerns, the capacity of the pre-ranking model is limited, making it difficult to well approximate original scores of the complex ranking model. Thus even with explicit optimization, there still exist errors in aligning their scores to be exactly the same.
- **Error amplification in ECPM ranks<sup>2</sup>.** In both pre-ranking and ranking phases, ads are ranked according to their ECPM as Eq. (1), which is jointly determined by the pCTR score and the bid. Thus the influence of alignment errors could be amplified due to existence of bids. As shown in Table 1, when multiplied by corresponding bids, even a tiny difference in pCTR scores of the pre-ranking and ranking models leads to completely different ranked results.

**Table 1: A toy example of error amplification in ECPM ranks. Though pCTR scores of two phases are similar, their ranked results after considering bids are different.**

Candidates	Pre-Ranking				Ranking			
	bid	pCTR	ECPM	rank	bid	pCTR	ECPM	rank
A	21	0.1	2.1	2	21	0.11	2.31	1
B	11	0.2	2.2	1	11	0.19	2.09	2

Above issues call for rethinking the necessity of strictly aligning pCTR scores in the advertising system. Essentially, given a set of candidates, it is **not their absolute pCTR scores but their relative ECPM ranks** that determine the results of each phase. Therefore, to achieve consistent results, the pre-ranking model is not required to output same pCTR scores as the ranking model. Instead, it only needs to output scores which yield same ECPM ranks when multiplied by bids. In this way, the requirement of score alignment can be relaxed to that of **rank alignment**, which is more easier to meet. Moreover, when optimizing pCTR scores for consistent ECPM ranks, the influence of bids can be taken into account beforehand, thus alleviating the issue of error amplification.

To this end, we introduce a **Consistency-Oriented Pre-Ranking (COPR)** framework for online advertising, which explicitly optimize the pre-ranking model towards consistency with the ranking model. Particularly, we collect historical logs of the ranking phase, where each log records a ECPM-ranked list of candidates. COPR segments the list into fixed-sized chunks. Each chunk is endowed

with certain level of priority from the view of the ranking phase. With pairs of ads sampled from different chunks, COPR learns an plug-and-play rank alignment module which aims to consistently distinguish their priority using scores at the pre-ranking phase. Moreover, we adopts a  $\Delta NDCG$ -based weighting mechanism to better distinguish the importance of inter-chunk pairs in optimization.

Our main contributions can be summarized as follows:

- To the best of our knowledge, we are the first to explicitly optimize the pre-ranking model towards consistency with the ranking model in the widely-used cascading architecture for online advertising.
- We propose a novel consistency-oriented pre-ranking framework named COPR, which employs a chunk-based sampling module and a plug-and-play rank alignment module for effective improvement of consistency.
- We conduct extensive experiments on public and industrial datasets. Both offline and online results validate that the proposed COPR framework significantly outperforms state-of-the-art baselines. When deployed in Taobao display advertising system, it achieves an improvement of up to +12.3% CTR and +5.6% RPM.

## 2 RELATED WORK

In this section, we briefly review studies about pre-ranking.

Located in the middle of the cascading architecture, the pre-ranking system has played an indispensable role for many large-scale industrial systems [16, 23]. The development of a pre-ranking model is mainly for balancing the system effectiveness and efficiency, as the downstream ranking model usually cannot deal with tens of thousands of candidates. To this end, techniques such as the dual-tower modeling [8, 24] are commonly adopted. However, this paradigm limits feature interactions between users and items to the form of vector product, which often results in extensive performance degradation.

Another line of work strives to enhance high-order feature interactions, and explores the ways to reduce the online latency. Li et al. [10] add fine-grained and early feature interactions between two towers. Wang et al. [23] propose to use fully-connected layers and employ various techniques from the perspectives of both modeling efficiency and engineering optimization. Specifically, a **Squeeze-and-Excitation** module [7] is utilized to choose the most useful feature set, and meanwhile system parallelism and low-precision computation are exploited whenever possible for latency optimization. Ma et al. [16] propose a feature selection algorithm based on **feature complexity and variational dropout (FSCD)** to search a set of effective and efficient features for pre-ranking. A similar study [11] uses **network architecture searching (NAS)** to determine the optimal set of features and corresponding architectures. These studies mainly focus on improving the accuracy of the pre-ranking model but neglects its interaction with the subsequent ranking model, leading to inconsistent ranked results.

Several studies propose to align the pre-ranking and ranking models in terms of pCTR scores via knowledge distillation. RD [20] encourages the lightweight student model to score higher for candidates selected by the larger teacher model, which is often used

<sup>2</sup>We use **ECPM rank** to denote the order of an ad in the ECPM-ranked list.

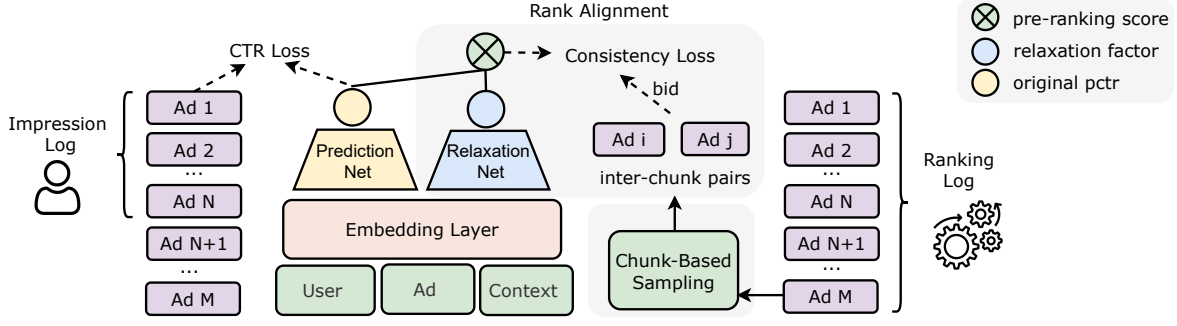


Figure 2: The framework of consistency-oriented pre-ranking.

in training pre-ranking models. RankFlow [19] regularizes the pre-ranking and ranking models to generate same scores for same candidates. Despite encouraging performance, there still exist inevitable errors in score alignment due to discrepancy in model capacity. When applied in online advertising, influence of such errors would be amplified by bids of ads, yielding inconsistent ECPM-ranked results. In this paper, we propose to relax the objective of score alignment to rank alignment, where bids of ads are incorporated and consistency of ranked results between two phases can be explicitly optimized in an effective manner.

### 3 METHODOLOGY

In this section, we first introduce background knowledge about the pre-ranking model, and then describe our proposed COPR framework as illustrated in Fig. 2.

#### 3.1 Background

**Training Data.** When the advertising system serves online traffic as Fig. 1, hundreds of ads are ranked through the ranking phase and recorded to logs, which we refer to as **ranking logs**. Each log contains an ranked list of ads with descending ECPM:

$$\mathbf{R} = [(ad_1, pCTR_1, bid_1), \dots, (ad_M, pCTR_M, bid_M)], \quad (2)$$

where  $pCTR_i$  is the score output by the ranking model for  $i$ -th ad and  $bid_i$  denotes its bid.  $M$  is the number of candidates. Then top  $N$  ads are displayed to the user. User feedback  $y$  (click/non-click) on each displayed ad is recorded to **impression logs**:

$$\mathbf{I} = [(ad_1, y_1), \dots, (ad_N, y_N)]. \quad (3)$$

**Base Model.** The base model for pre-ranking is usually a lightweight CTR model. Here we adopt the architecture of COLD [23]. The input features consist of three parts: user features  $\mathbf{U}$  such as age and gender, ad features  $\mathbf{A}$  such as brand and category, context features  $\mathbf{C}$  such as time and device. After pre-selecting a concise set of features, COLD feeds them into embedding layers and concatenate their embeddings for a compact representation  $\mathbf{x}$ :

$$\mathbf{x} = E(\mathbf{U}) \oplus E(\mathbf{A}) \oplus E(\mathbf{C}). \quad (4)$$

Then it employs a prediction net consists of multiple fully-connected layers to estimate CTR:

$$\hat{y} = \text{Sigmoid}(\text{MLP}(\mathbf{x})) \in [0, 1]. \quad (5)$$

To accurately predict user click  $y$ , the model is optimized with cross entropy loss over impression logs  $\mathbf{I}$ :

$$L_{ctr} = \sum_{\mathbf{I}} [-y \log(\hat{y}) - (1 - y) \log(1 - \hat{y})]. \quad (6)$$

#### 3.2 Consistency-Oriented Pre-Ranking

Though the pre-ranking model is expected to well approximates the ranking model in the cascading system, their gap in model capacity often hinders satisfying approximation. Thus in addition to  $L_{ctr}$ , we aim to explicitly optimize the pre-ranking model towards consistent results with the ranking model over  $\mathbf{R}$ .

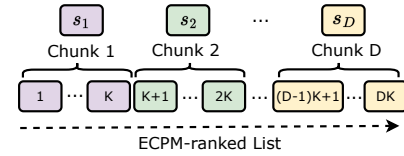


Figure 3: Illustration of chunk-based sampling.

**3.2.1 Chunk-Based Sampling.** Given candidates  $\{Ad_i\}_1^M$  in ranking logs, an ideal pre-ranking model should output scores that yield same ECPM-ranked list as Eq. (2). Considering its limited capacity, it could be hard to rank hundreds of ad all in correct positions. To reduce the learning difficulty, we partition the ranked list into  $D = \frac{M}{K}$  fixed-sized chunks, each constituting  $K$  adjacent ads, as shown in Fig. 3. We regard ads in the same chunk as candidates with same priority in the ranking phase. The pre-ranking model is not required to distinguish ads in the same chunk. Instead, it only needs to consistently rank candidates in the granularity of chunk. For each chunk, we randomly sample a candidate and endow it with the priority related to this chunk. In this way, for each ranked list, we obtain a concise sub-list:

$$\mathbf{R}_{chunk} = [(ad_{s_d}, pCTR_{s_d}, bid_{s_d}, D - d)]_{d=1}^D, \quad (7)$$

where  $s_d$  is the index of sampled ad in chunk  $d$  and  $D - d$  denotes its priority which the larger the better.

The above chunk-based sampling has two-fold advantages: 1) It provides a flexible way to control the granularity of consistency, which makes the objective reachable for the lightweight

**pre-ranking model.** By increasing the chunk size  $K$ , the objective of consistency gradually shifts from fine-grained to coarse-grained. 2) It effectively reduces the size of ranked list in logs by  $K$  times and still maintains coverage of original lists, which is critical for efficient training in industrial machine learning systems. In our production implementation,  $K$  is set to 10.

**3.2.2 Rank Alignment.** In the following, we introduce how to modify the base model with a plug-and-play rank alignment module.

Instead of regularizing the difference between  $\hat{y}_i$  in Eq. (5) and  $pCTR_i$  in Eq. (7) as score alignment methods [19, 20], we propose to relax the objective to rank alignment on a properly-adjusted pCTR score. Particularly, we employ a relaxation net to learn a factor  $\alpha > 0$ , with which we adjust the original pCTR score:

$$\alpha = \text{ReLU}(\text{MLP}(x)) + 1e^{-6} \in \mathcal{R}^+, \quad (8)$$

$$\tilde{y} = \alpha * \hat{y},$$

where  $\tilde{y}$  denote the adjusted pCTR. Thus ECPM at the pre-ranking phase can be accordingly estimated as  $\tilde{y} * \text{bid}$ , based on which we aim to correctly rank each inter-chunk pair in  $\mathbf{R}_{\text{chunk}}$ . Here we adopt the pairwise logistic loss for its relatively good performance and the simplicity for implementation [1, 17]:

$$L_{\text{rank}} = \sum_{i < j} \log[1 + e^{-\left(\frac{\tilde{y}_{s_i} * \text{bid}_{s_i}}{\tilde{y}_{s_j} * \text{bid}_{s_j}} - 1\right)}]. \quad (9)$$

For each pair of  $ad_{s_i}$  and  $ad_{s_j}$  sampled from different chunks that  $i < j$ , we optimize  $L_{\text{rank}}$  by encouraging  $\tilde{y}_{s_i} * \text{bid}_{s_i} > \tilde{y}_{s_j} * \text{bid}_{s_j}$ , which means  $ad_{s_i}$  would be ranked before  $ad_{s_j}$  by ECPM in the pre-ranking phase. If all inter-chunk pairs can be correctly ranked, we achieve consistent ECPM-ranked results between the pre-ranking and ranking phases over  $\mathbf{R}_{\text{chunk}}$ .

Note that by introducing the relaxation factor  $\alpha$ , we slightly modify the original pCTR score to achieve consistent ranked results if necessary. To maintain original value as much as possible,  $\alpha$  should be around 1. Thus we add a symmetric regularization to penalize the deviation of  $\alpha$  from 1:

$$L_{\text{reg}} = \begin{cases} \alpha - 1 & \alpha > 1 \\ \frac{1}{\alpha} - 1 & \alpha \leq 1 \end{cases}. \quad (10)$$

It is worth mentioning that the proposed rank alignment module does not rely on specific assumption about the architecture of base model. It is an plug-and-play component that can be added to any pre-ranking models for improvement of consistency.

**3.2.3  $\Delta\text{NDCG}$ -Based Pair Weighting.**  $L_{\text{rank}}$  in Eq. (9) fails to consider the relative importance of different pairs in consistency optimization. In practice, consistently ranking ads from chunk 1 and chunk 10 is more important than ranking chunk 11 and chunk 20, since only the top ads will be sent to the ranking phase and displayed to users. It calls for a weighting mechanism that considers chunk-related priorities of candidates.

Intuitively, if pair  $(ad_{s_i}, ad_{s_j})$  in  $L_{\text{rank}}$  are mistakenly ranked, the consistency between the pre-ranking and ranking phase will be hurt. Thus its weight in  $L_{\text{rank}}$  should be determined by the negative impact. As each sampled  $ad_{s_d}$  in  $\mathbf{R}_{\text{chunk}}$  is endowed with priority  $D - d$ , we use NDCG [2, 9] to measure the utility of any ranked list

$p$  of these candidates:

$$\text{DCG} = \sum_{i=1}^D \frac{2^{p_i} - 1}{\log(i + 1)}, \quad (11)$$

$$\text{IDCG} = \sum_{i=1}^D \frac{2^{D-i} - 1}{\log(i + 1)},$$

where  $p_i$  denote the priority of  $i$ -th ad in the permutation and the IDCG is the ideal DCG achieved by  $\mathbf{R}_{\text{chunk}}$ . If we swap the position of  $ad_{s_i}$  and  $ad_{s_j}$  in  $\mathbf{R}_{\text{chunk}}$ , the utility of the list will experience a drop which can be further normalized as:

$$\Delta\text{NDCG}(i, j) = \frac{2^{D-i} - 2^{D-j}}{\text{IDCG}} \left[ \frac{1}{\log(i + 1)} - \frac{1}{\log(j + 1)} \right]. \quad (12)$$

The utility drop is used to re-weight inter-chunk pairs in consistency optimization:

$$L_{\text{rank}} = \sum_{i < j} \Delta\text{NDCG}(i, j) \log[1 + e^{-\left(\frac{\tilde{y}_{s_i} * \text{bid}_{s_i}}{\tilde{y}_{s_j} * \text{bid}_{s_j}} - 1\right)}]. \quad (13)$$

Thus the objective function of COPR can be formulated as:

$$L = \underbrace{L_{\text{ctr}}}_{\text{CTR Loss}} + \underbrace{\lambda_1 L_{\text{rank}} + \lambda_2 L_{\text{reg}}}_{\text{Consistency Loss}}, \quad (14)$$

where  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  are weights for corresponding loss terms. By minimizing  $L$ , we explicitly optimize the pre-ranking model towards consistency with the ranking model via a plug-and-play rank alignment module.

### 3.3 System Deployment

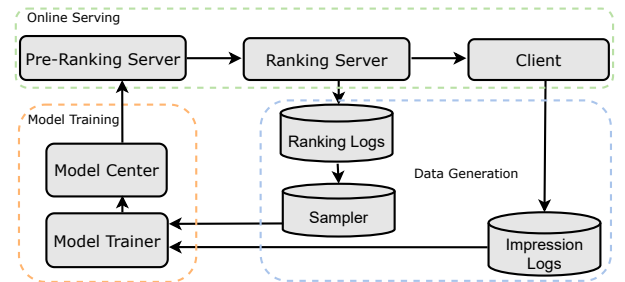


Figure 4: Overview of system pipeline.

We introduce the deployment of COPR in three stages: data generation, model training, and online serving as shown in Fig. 4. **Data Generation.** During online serving, hundreds of ads are ranked through ranking model and recorded to ranking logs, with which we perform chunk-based sampling. The content of each sample includes user index, ad index, chunk index as well as the bid. Note that the bid at the ranking phase could differ from that at the pre-ranking phase [22]. In this case, we record the pre-ranking bid since it influences  $L_{\text{rank}}$  in model training. When ads are displayed to users in the client, we also record user feedback in impression logs, which are used in calculating  $L_{\text{ctr}}$ .



**Model Training.** The training procedure is performed on our ODL (Online Deep Learning) [14] platform, which consumes real-time streaming data to continuously update model parameters. After training with fixed number of steps, the learnt model will be delivered to the Model Center, which manages all online models.

**Online Serving.** Once a new version of pre-ranking model is ready, pre-ranking server will load it from Model Center to replace the online version in service.

## 4 EXPERIMENTS

In this section, we conduct experiments on both public dataset and production dataset to validate the effectiveness of COPR in improving consistency and overall system performance.

### 4.1 Experiment Setup

**Taobao Dataset.** It is a public dataset<sup>3</sup> with 26 million impression logs of 1 million users and 0.8 million items in 8 days. Item price is used as bid. Impressions of first 7 days are used to train DIN [26] as the ranking model. For each impression, we sample 10 candidates and collect ECPM-ranked results by the ranking model to train pre-ranking models. Logs of the last day are used for evaluation. To simulate the cascading process, we sample 100 candidates for each impression, among which the pre-ranking and ranking model sequentially select top 10 and top 1 candidates to display.

**Production Dataset.** It contain 8 days of impression logs and ranking logs collected from our system shown in Fig. 4. These logs are of the magnitude of billions. The first week of logs are used for training and the last day is used for evaluation. According to the scenario that logs come from, it is further divided into two subsets: **Homepage** and **Post-Purchase**.

**Baselines.** COPR is compared with following baselines:

- **Base** adopts the architecture of COLD [23] and is trained on impression logs.
- **Distillation** [6] directly distills predicted scores of the ranking model on impression logs.
- **RankFlow** [19] distills predicted scores of the ranking model on ranking logs and further regularizes the pre-ranking model to generate high scores for candidates selected by the ranking model.
- **COPR w/o  $\Delta NDCG$**  removes the  $\Delta NDCG$ -based weighting mechanism from the COPR framework.

**Metrics.** We adopt two groups of metrics in evaluation.

- The first group measures the consistency between ECPM-ranked results of the pre-ranking and ranking phases, including HitRatio(**HR@K**), normalized discounted cumulative gain (**NDCG@K**), and mean average precision (**MAP@K**). In HR@K and MAP@K, top 10 candidates selected by the ranking model are treated as relative ones. In NDCG@K, order in ranking logs is used as a proxy of relevance. The standard calculation of these metrics can be found in [13].
- The second group measures the overall system performance. We use Click-Through-Rate (**CTR**) and Revenue Per Mille (**RPM**) similar to [19, 21], which corresponds to user experience and platform revenue, respectively. On public dataset, CTR is simulated as the portion of clicked ads in displayed

ads, and RPM is simulated as the product of CTR and average bid of clicked ads. In production experiment, we perform online A/B test to obtain CTR and RPM on real traffic.

**Hyper-parameters.** The chunk size is set to 2 and 10 on the public dataset and the production dataset, respectively. The number of MLP layers in the prediction net and the relaxation net is 3. The embedding size of raw input features is set to 16.  $\lambda_1$  and  $\lambda_2$  in Eq. (14) are fixed to 1 and 0.2.

### 4.2 Results on Public Dataset

**Table 2: Comparison among COPR and baselines in terms of consistency and overall system performance. Best results are highlighted in bold.**

Method	HR@10	NDCG@10	MAP@10	CTR	RPM
Base	0.544	0.402	0.198	0.0170	112.66
Distillation	0.593	0.461	0.244	0.0176	117.36
RankFlow	0.722	0.467	0.270	0.0182	121.36
COPR w/o $\Delta NDCG$	0.741	0.514	0.327	0.0188	129.20
COPR	<b>0.759</b>	<b>0.530</b>	<b>0.359</b>	0.0191	132.93

Table 2 compares COPR and baselines in terms of consistency and system performance. We only show  $K = 10$  in HR@K, NDCG@K, and MAP@K due to limited space. Results under other settings of  $K$  are similar. From Table 2, we draw the following conclusions.

First, system performance (CTR and RPM) is highly associated with the consistency between the pre-ranking and ranking phases. For COPR and baselines, the higher consistency generally yields the better system performance. It validates our motivation to explicitly optimize consistency between phases in order to improve the overall effectiveness of the cascading system.

Second, COPR achieves best consistent results of all methods, outperforming the state-of-the-art RankFlow by 5.1%, 13.5%, and 33.0% in terms of HR@10, NDCG@10, and MAP@10. We attribute the improvement to our shift of objective from score alignment to rank alignment. By such relaxation, COPR can directly optimize towards consistent ECPM-ranked results and meanwhile reduce the learning difficulty for the lightweight model. Moreover, the influence of bids is considered in training COPR, thus alleviating the issue of error amplification that RankFlow suffers from. We also find that RankFlow is better than Distillation. We think it is because Rankflow aligns scores over ranking logs while the latter is on impression logs which is too sparse.

Third, COPR w/o  $\Delta NDCG$  experiences performance drop compared with COPR. This ablation study verifies the effectiveness of the pair weighting mechanism based on  $\Delta NDCG$ . By emphasizing more on important inter-chunk pairs in consistency optimization, COPR ensures top candidates are more likely to be consistently ranked, which helps improve the overall utility of pre-ranking results.

### 4.3 Results on Production Dataset

We also perform similar evaluation on the production dataset composed of samples from two scenarios. Most conclusions are consistent with those on the public dataset.

<sup>3</sup><https://tianchi.aliyun.com/dataset/dataDetail?dataId=56>

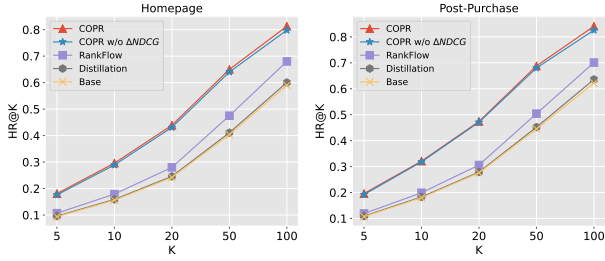


Figure 5: HR@K of different methods in the scenario of Homepage (Left) and Post-Purchase (Right).

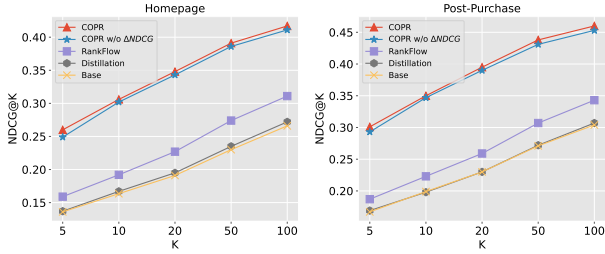


Figure 6: NDCG@K of different methods in the scenario of Homepage (Left) and Post-Purchase (Right).

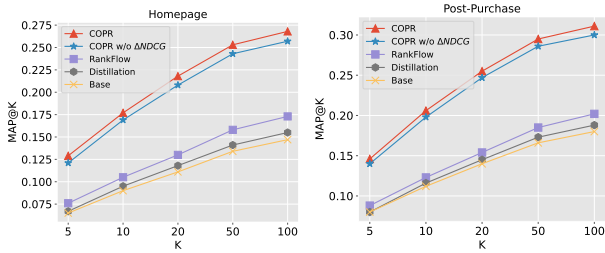


Figure 7: MAP@K of different methods in the scenario of Homepage (Left) and Post-Purchase (Right).

As shown in more details from Fig. 5 to Fig. 7, COPR significantly outperforms other methods in term of HR@K, NDCG@K, and MAP@K with varying  $K$  from 5 to 100 on two scenarios, which demonstrates the stable improvement of consistency achieved by our proposed framework. Moreover, we still observe the gap between COPR and COPR w/o  $\Delta NDCG$ , which shows that the weighting mechanism also works in the large-scale production dataset.

To evaluate system performance in production environment, we perform online A/B test on two scenarios, where these methods are used to serve real users and advertisers. From Table 3 we find that Distillation, RankFlow, and COPR all perform better than the production baseline, among which COPR achieves the largest improvement, with a lift of up to +12.3% CTR and +5.6% RPM. With impressive performance, **COPR has been successfully deployed to serve the main traffic of Taobao display advertising system in the pre-ranking phase since October of 2022.**

Table 3: Relative improvement over the production baseline in online A/B Test. Best results are highlighted in bold.

Method	Homepage		Post-Purchase	
	CTR	RPM	CTR	RPM
Base	-	-	-	-
Distillation	+2.2%	+0.1%	+3.6%	+0.6%
RankFlow	+8.3%	+2.9%	+6.8%	+2.3%
COPR w/o $\Delta NDCG$	+11.5%	+5.0%	+9.6%	+3.7%
COPR	<b>+12.3%</b>	<b>+5.6%</b>	<b>+10.8%</b>	<b>+4.4%</b>

#### 4.4 Qualitative Analysis

Given ranked results from the pre-ranking and ranking phases, we calculate the average pre-ranking position for candidates at each ranking position, based on which we draw the Ranking-PreRanking Curve (RPC). The ideal RPC happens when results are exactly same.

**4.4.1 Error Amplification in ECPM Rank.** As shown in Fig. 8 (Left), RPC by pCTR of RankFlow is close to the ideal curve, showing well alignment of raw pCTR in two phases. However, after ranking by ECPM, RPC of RankFlow largely deviates from the ideal one. It verifies that the involvement of bid in ECPM will amplify the influence of errors in score alignment, leading to more inconsistent ECPM-ranked results. This analysis is consistent with the example in Table 1. Hence we confirm that merely score alignment is not enough for the cascading architecture in online advertising.

**4.4.2 More Consistent ECPM Rank.** Fig. 8 (Right) shows RPC by ECPM of different methods. We observe that compared with Base and RankFlow, RPC of COPR is more close to the ideal curve in almost each ranking position. It qualitatively shows that ECPM-ranked results given by COPR are more consistent with results of the ranking phase. It can be attributed to the design of our consistency-oriented framework, where the rank alignment module directly optimizes towards this objective. The incorporation of bid also helps alleviate the above mentioned error amplification.

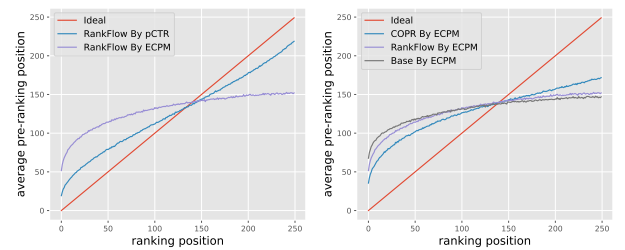


Figure 8: Left: RPC by pCTR and ECPM of RankFlow. Right: RPC by ECPM of different methods.

## 5 CONCLUSION

In this paper, we introduce a consistency-oriented pre-ranking framework for online advertising, which employs a chunk-based sampling module and a plug-and-play rank alignment module to explicitly optimize consistency of ECPM-ranked results. A  $\Delta NDCG$ -based weighting mechanism is also adopted to better distinguish the

importance of inter-chunk samples in optimization. Both online and offline experiments have validated the superiority of our framework. When deployed in Taobao display advertising system, it achieves an improvement of up to +12.3% CTR and +5.6% RPM.

## REFERENCES

- [1] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. 2005. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*. 89–96.
- [2] Christopher JC Burges. 2010. From ranknet to lambdarank to lambdamart: An overview. *Learning* 11, 23-581 (2010), 81.
- [3] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.
- [4] Avi Goldfarb and Catherine Tucker. 2011. Online display advertising: Targeting and obtrusiveness. *Marketing Science* 30, 3 (2011), 389–404.
- [5] Huifeng Guo, Bo Chen, Ruiming Tang, Weinan Zhang, Zhenguo Li, and Xiuqiang He. 2021. An embedding learning framework for numerical features in ctr prediction. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2910–2918.
- [6] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. 2015. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531* (2015).
- [7] Jie Hu, Li Shen, and Gang Sun. 2018. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7132–7141.
- [8] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. 2013. Learning deep structured semantic models for web search using clickthrough data. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*. 2333–2338.
- [9] Kalervo Järvelin and Jaana Kekäläinen. 2017. IR evaluation methods for retrieving highly relevant documents. In *ACM SIGIR Forum*, Vol. 51. ACM New York, NY, USA, 243–250.
- [10] Xiangyang Li, Bo Chen, HuiFeng Guo, Jingjie Li, Chenxu Zhu, Xiang Long, Sujian Li, Yichao Wang, Wei Guo, Longxia Mao, et al. 2022. IntTower: the Next Generation of Two-Tower Model for Pre-Ranking System. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 3292–3301.
- [11] Xiang Li, Xiaojiang Zhou, Yao Xiao, Peihao Huang, Dayao Chen, Sheng Chen, and Yunsen Xian. 2022. AutoFAS: Automatic Feature and Architecture Selection for Pre-Ranking System. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3241–3249.
- [12] Bin Liu, Chenxu Zhu, Guilin Li, Weinan Zhang, Jincai Lai, Ruiming Tang, Xiuqiang He, Zhenguo Li, and Yong Yu. 2020. Autofis: Automatic feature interaction selection in factorization models for click-through rate prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2636–2645.
- [13] Tie-Yan Liu et al. 2009. Learning to rank for information retrieval. *Foundations and Trends® in Information Retrieval* 3, 3 (2009), 225–331.
- [14] Qiang Luo, Xiang-Rong Sheng, Jun Jiang, Chi Ma, Shuguang Wang, Haiyang He, Pengtao Yi, Guowang Zhang, Yue Song, and Guorui Zhou. 2020. Bernoulli: A Streaming System with Structured Designs. (2020).
- [15] Ning Ma, Mustafa Ispir, Yuan Li, Yongpeng Yang, Zhe Chen, Derek Zhiyuan Cheng, Lan Nie, and Kishor Barman. 2022. An Online Multi-Task Learning Framework for Google Feed Ads Auction Models. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3477–3485.
- [16] Xu Ma, Pengjie Wang, Hui Zhao, Shaoguo Liu, Chuhan Zhao, Wei Lin, Kuang-Chih Lee, Jian Xu, and Bo Zheng. 2021. Towards a Better Tradeoff between Effectiveness and Efficiency in Pre-Ranking: A Learnable Feature Selection based Approach. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2036–2040.
- [17] Rama Kumar Pasumarthi, Sebastian Bruch, Xuanhui Wang, Cheng Li, Michael Bendersky, Marc Najork, Jan Pfeifer, Nadav Golbandi, Rohan Anil, and Stephan Wolf. 2019. TF-ranking: Scalable tensorflow library for learning-to-rank. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2970–2978.
- [18] Qi Pi, Guorui Zhou, Yujing Zhang, Zhe Wang, Lejian Ren, Ying Fan, Xiaoqiang Zhu, and Kun Gai. 2020. Search-based user interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2685–2692.
- [19] Jiarui Qin, Jiachen Zhu, Bo Chen, Zhirong Liu, Weiwen Liu, Ruiming Tang, Rui Zhang, Yong Yu, and Weinan Zhang. 2022. RankFlow: Joint Optimization of Multi-Stage Cascade Ranking Systems as Flows. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 814–824.
- [20] Jiayi Tang and Ke Wang. 2018. Ranking distillation: Learning compact ranking models with high performance for recommender system. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2289–2298.
- [21] Jun Wang and Shuai Yuan. 2015. Real-time bidding: A new frontier of computational advertising research. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. 415–416.
- [22] Yiqing Wang, Xiangyu Liu, Zhenzhe Zheng, Zhilin Zhang, Miao Xu, Chuan Yu, and Fan Wu. 2022. On Designing a Two-stage Auction for Online Advertising. In *Proceedings of the ACM Web Conference 2022*. 90–99.
- [23] Zhe Wang, Liqin Zhao, Biye Jiang, Guorui Zhou, Xiaoqiang Zhu, and Kun Gai. 2020. Cold: Towards the next generation of pre-ranking system. *arXiv preprint arXiv:2007.16122* (2020).
- [24] Wenjin Wu, Guojun Liu, Hui Ye, Chenshuang Zhang, Tianshu Wu, Daorui Xiao, Wei Lin, and Xiaoyu Zhu. 2018. Eenmf: An end-to-end neural matching framework for e-commerce sponsored search. *arXiv preprint arXiv:1812.01190* (2018).
- [25] Bowen Yuan, Jui-Yang Hsia, Meng-Yuan Yang, Hong Zhu, Chih-Yao Chang, Zhenhua Dong, and Chih-Jen Lin. 2019. Improving ad click prediction by considering non-displayed events. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 329–338.
- [26] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 1059–1068.