GenAuction: A Generative Auction for Online Advertising

Yuchao Ma[1*], Ruohan Qian[2*], Bingzhe Wang[1*], Qi Qi[1†], Wenqiang Liu[2], Qian Tang[2], Zhao Shen[2], Wei Zhong[2], Bo Shen[2], Yixin Su[2], Bin Zou[2], Wen Yi[2], Zhi Guo[2], Shuanglong Li[2], Lin Liu[2]

[1]Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China
[2]Baidu Inc., Beijing, China
{yc.ma, wbz2022, qi.qi}@ruc.edu.cn, {qianruohan, liuwenqiang01, tangqian01, shenzhao, zhongwei04, shenbo04, suyixin, zoubin03, yiwen01, guozhi, lishuanglong, liulin03}@baidu.com

Abstract

Previous ad auctions predominantly relied on rule-based mechanisms, which selected winning advertisements (ads) at the ad-level and subsequently combined them into page views (PVs), leading to suboptimal allocations in multi-round auctions. This limitation stems from the significant computational burden required to design ranking score rules and select winning ad sets, as well as the inability to fully capture contextual information within PVs during ad-level selection. In this paper, we propose a key-performance-indicator (KPI) based auction mechanism that selects winning PVs at the PV-level, modeling the ad allocation as a constrained optimization problem. This approach enables us to address both short-term and long-term KPIs while leveraging the comprehensive contextual information available within PVs. Based on this framework, we design GenAuction, a generative auction mechanism utilizing a Generator-Evaluator architecture powered by Transformer algorithms. The Generator swiftly generates multiple candidate PVs, while the Evaluator selects the optimal PVs based on contextual information, adhering to the objectives and KPIs of multi-round auctions. We conduct extensive experiments using real-world data and online A/B tests to validate that GenAuction efficiently handles multi-objective allocation tasks, demonstrating its efficacy and potential for real-world application.

Datasets — https://drive.google.com/drive/folders/1xHsLdHJRPWXCF2s2kdat7xUnQ6cFIILN?usp=drive_link

Introduction

Online advertising (ad) plays a pivotal role in search platforms, with ad auction constituting a primary source of revenue for platforms. An auction mechanism comprise an allocation mechanism and a payment mechanism. In Ad auctions, advertisers (bidders) submit bids representing their valuations for an ad impression (or click, conversion), and the platform runs the allocation mechanism to display the winning bidders while charging them based on the payment mechanism. Traditional research on auction mechanism primarily focused on designing single-objective auctions. Vickrey, Clarke, and Groves designed the Vickrey-Clarke-Groves (VCG) mechanism, which aims to maximize social welfare (SW) (Vickrey 1961; Clarke 1971; Groves 1973). Subsequently, Myerson designed the Myerson auction to maximize the auctioneer's revenue (Myerson 1981). Building upon these foundations, Overture and Google proposed the Generalized First-Price Auction (GFP) (Despotakis, Ravi, and Sayedi 2021; Han and Liu 2015) and Generalized Second-Price Auction (GSP) (Edelman, Ostrovsky, and Schwarz 2007; Varian 2007; Chawla and Hartline 2013), respectively, for selling ads and implemented them in their ad systems.

As ad systems evolved, a single objective became insufficient to satisfy the needs of platforms, advertisers, and users, as their key performance indicators (KPIs) often conflict. For instance, a revenue-maximizing mechanism might compromise user experience by displaying numerous high-bid but low-quality ads. Consequently, research shifted towards multi-objective auctions (Lahaie and Pennock 2007; Thompson and Leyton-Brown 2013), aiming to optimize a combination of multiple KPIs (e.g., SW, revenue, conversion volume) (Geyik et al. 2016; Bachrach et al. 2014). Multi-objective auctions often employ rule-based mechanisms, which pre-design an aggregation function. This function serves to integrate the scores across various dimensions of a bidder, resulting in a rank score. By designing aggregation functions, various multi-objective auctions can be achieved, such as simultaneously maximizing SW and revenue (Zhang et al. 2021; Liu et al. 2021) or balancing short-term and long-term revenue (e.g., user retention) (Zhang et al. 2019; Lahaie et al. 2018; Deng et al. 2020; Deng and Lahaie 2019). Once rank scores are determined, mechanisms like GSP can be employed to maximize the sum of rank scores. Designing an aggregation function that caters to platform's objectives poses a significant challenge and is a focal point of related research. In multi-slot scenarios, rule-based mechanisms necessitate computing rank scores for all bidders and selecting an optimal combination of bidders, leading to complex computations. Moreover, rule-based mechanisms are static, using fixed aggregation function across all auctions (Liu et al. 2021; Zhang et al. 2021). However, with billions of rounds occurring daily in ad systems, applying

[*]These authors contributed equally.
[†]Corresponding Author.

the same rule with unchanged aggregation function results in suboptimal solutions for multi-round auctions.

To obtain the optimal allocation of multi-round auctions, we proposes a novel approach distinct from rule-based mechanisms, termed KPI-based mechanisms. Specifically, we model the multi-objective allocation problem as a constrained optimization problem, with the primary KPI (objective) serving as the objective function and other KPIs as constraints. KPI-based mechanisms incorporate the KPIs of various stakeholders in the advertising system directly as constraints, whereas rule-based mechanisms necessitate an intermediary step where an aggregation function is devised based on these KPIs, followed by the computation of rank scores for all bidders. Consequently, KPI-based mechanisms is a more straightforward approach. Furthermore, KPI-based mechanisms can address both short-term (single-round) and long-term (multi-round) KPIs, thereby mitigating the suboptimal performance issues encountered by rule-based mechanisms in the context of multi-round auctions.

In a page view (PV), multiple ads are typically displayed. A prevalent assumption in existing research is that ads are mutually independent, implying that their click-through rates (CTRs) and conversion rates (CVRs) are independent, disregarding any inter-ad effects (Bachrach et al. 2014; Huang et al. 2021). However, this assumption contrasts with the realities of ad systems, where users instantaneously assess the quality of ads in a PV, thereby influencing CTRs (Gatti, Lazaric, and Trovò 2012; Gomes, Immorlica, and Markakis 2009; Hummel and McAfee 2014). While recent studies have begun to account for inter-ad influences, most of them have focused solely on the interactions within a given ad set (Liu et al. 2021; Wang et al. 2022; Goel et al. 2023). Nevertheless, in a PV-level, variations in the composition and even the arrangement of ads can alter user engagement, necessitating the integration of contextual information into CTR estimation. Only a handful of studies have considered the contextual interplay within a PV, yet these efforts remain inadequate, primarily capturing the influence of preceding ads on subsequent ones while neglecting the reverse effect (Fotakis, Krysta, and Telelis 2011; Dupret and Piwowarski 2008; Farina and Gatti 2016; Kempe and Mahdian 2008), thus rendering the contextual influence incomplete. The inherent complexity of rule-based mechanisms, particularly in selecting ad combination, escalates significantly when contextual factors are introduced. While deep learning techniques have been used in mechanism design, substantially reducing computational load (Dütting et al. 2019; Feng, Narasimhan, and Parkes 2018; Rahme, Jelassi, and Weinberg 2020; Shen, Tang, and Zuo 2018), no existing work concurrently addresses PV-level information.

In this paper, we devise a Transformer-based generative auction mechanism, GenAuction, which not only handle short-term and long-term KPIs but also incorporates complete contextual information in PV-level, achieving optimal allocation for multi-round auctions. Specifically, GenAuction's allocation mechanism comprises two primary modules: Generator and Evaluator. The Generator generates multiple candidate PVs, while the Evaluator selects the optimal PV based on contextual information, adhering to the objectives and constraints across multiple rounds. Notably, the optimal PV is chosen to maximize performance across multiple rounds, rather than merely optimizing for single round. Furthermore, to ensure GenAuction satisfies Incentive Compatibility (IC), we employ the VCG payment mechanism.

Our contributions can be summarized as follows:

KPI-based mechanism Design. We formulate the allocation problem as an constrained optimization problem and first propose a KPI-based mechanism. This approach circumvents the rank score calculation procedures in rule-based mechanisms and directly aligns with the KPIs of all stakeholders within the ad system.

Complete Contextual Information. To avoid the issue of incomplete contextual information being considered during ad-level selection, we directly select PVs at the PV-level. This approach enables us to evaluate the comprehensive contextual information encapsulated within each PV.

Generative Auction Mechanism. Employing a generator-evaluator architecture and grounded in the VCG payment mechanism, we introduce GenAuction, a novel generative auction mechanism. GenAuction is a KPI-based mechanism where the Generator produces multiple candidate PVs, while the Evaluator selects the optimal PV, guided by contextual information and adhering to objectives and constraints across multiple rounds.

Rigorous Experimental Evaluation. We conducted extensive experiments using real-world data, complemented by rigorous online A/B testing, to demonstrate that GenAuction can efficiently achieve multi-objective allocation.

Further Related Work

The Generator-Evaluator architecture is a computationally efficient framework that has emerged in recent years in the context of recommendation system (RS) (Ren et al. 2024; Feng et al. 2021; Lin et al. 2024; Shi et al. 2023; Xi et al. 2021). However, unlike RS, ad systems necessitate the consideration of bidders' strategic behaviors, rendering the direct application of recommendation algorithms insufficient to satisfy the game-theoretic properties crucial in auction mechanisms, such as IC and Individual Rationality (IR). Similar to our approach, EdgeNet utilizes Transformers to devise its allocation mechanism (Shen et al. 2023). Nevertheless, EdgeNet is unable to handle long-term KPIs and fails to use complete contextual information. Similar to DeepGSP (Zhang et al. 2021), DNA (Liu et al. 2021), and NMA (Liao et al. 2022), GenAuction adopts a payment mechanism based on GSP. Nevertheless, DeepGSP, DNA, and NMA are rule-based mechanisms, primarily focusing on single-round auctions. GenAuction is a KPI-based mechanism for multi-round auctions. By doing so, it effectively addresses more complex short-term and long-term KPIs, enabling better allocation outcomes.

Online Ad System

In the realm of ad systems, we consider a setting where N bidders participate in a T-round auction. When a user searches a query in t, there are K bidders occupying K ad slots in a PV. In round t, the CTR and CVR of bidder n in slot k are denoted as α_{ntk} and β_{ntk}, respectively. γ_{nt} denote the relevance of ad n to PV t. In this paper, we delve into the OCPC (Optimized Cost per Click) pricing model, a prevalent pricing model within multi-round auction frameworks. In OCPC pricing model, bidder n submits a bid b_n, claiming that b_n represents the value v_n of a conversion. Upon receiving the bid profile $\boldsymbol{b} = (b_1, \cdots, b_N)$ from all bidders, the platform runs the auction mechanism $\mathcal{M}(\boldsymbol{x}, \boldsymbol{p})$ in each round, yielding an allocation outcome $\boldsymbol{x}$ and a payment outcome $\boldsymbol{p}$. If bidder n wins the impression and then being clicked, she pays p_{nt} to the platform. The utility of bidder n across all T rounds is $u_n = \sum_{t,k} (v_n x_{ntk} \alpha_{ntk} \beta_{ntk} - p_{nt})$, where x_{ntk} is an indicator variable for whether bidder n is displayed at slot k in round t.

In this paper, we delve into the realm of multi-objective mechanism design, where the primary objective of the platform is to maximize SW, defined as the aggregate values across all bidders, while taking into account both user experience and advertiser business requirements. Specifically, to ensure user experience, the ads displayed must maintain a certain level of relevance to the current PV and refrain from an excessive number of ads. Consequently, the platform imposes three KPIs pertaining to user experience: Γ_t, X_t, and X. Here, Γ_t governs the relevance threshold for ads within each PV, X_t restricts the number of ads per PV, and X imposes a cap on the total number of ads across all PVs. Furthermore, the platform should also consider the advertisers's business demands that, albeit prevalent in real-world advertising systems, have been overlooked in pertinent research. Notably, some advertisers may refrain from their ads appearing alongside certain others due to potential brand image damage. For instance, advertisers promoting luxury products may avoid being placed alongside those selling low-cost items, as this could undermine their respective brand perceptions. Additionally, ads espousing conflicting political or religious viewpoints must never be displayed together. To our knowledge, our work is the first to consider such business demands. We denote the set of bidders incompatible with bidder n as $\mathcal{N}_n$. If bidder n can coexist with any other bidder, then $\mathcal{N}_n = \emptyset$. Moreover, advertisers aim to avoid ad redundancy within the same PV, as users click at most one ad per PV, rendering repeated exposures futile for advertisers. Moreover, the slot is indivisible, i.e., a slot is either occupied by an ad or remains empty. Consequently, the platform confronts the optimization problem $\boldsymbol{Allo}$ when allocating ads.

$$\max \quad \sum_{n,t,k} b_n x_{ntk} \alpha_{ntk} \beta_{ntk} \qquad (\boldsymbol{Allo})$$

$$\text{s.t.} \quad \sum_{n,k} \gamma_{nt} x_{ntk} \geq \Gamma_t, \qquad \forall t \qquad (1)$$

$$\sum_{n,k} x_{ntk} \leq X_t, \qquad \forall t \qquad (2)$$

$$\sum_{n,t,k} x_{ntk} \leq X \qquad (3)$$

$$x_{ntk} \cdot \sum_{i \in [\mathcal{N}_n], k} (x_{ntk} + x_{itk}) \leq 1, \qquad \forall n, t \qquad (4)$$

$$\sum_{k} x_{ntk} \leq 1, \qquad \forall n, t \qquad (5)$$

$$x_{ntk} \in \{0, 1\}, \qquad \forall n, t, k. \qquad (6)$$

In auction $\mathcal{M}(\boldsymbol{x}, \boldsymbol{p})$, bidders strategically submit bids b_n with the objective of maximizing their individual utilities, implying that b_n does not necessarily equate to v_n. Untruthful bidding undermines the stability of the auction. To encourage advertisers to bid truthfully, the auction mechanism implemented by the platform ought to satisfy the property of Incentive Compatibility (IC).

Definition 1 (Incentive Compatibility) *An auction mechanism satisfies the property of Incentive Compatibility (IC) iff the best strategy for each bidder is truthful bidding, i.e.,* $b_n = v_n, \forall n$.

To ensure that the mechanism satisfies IC, we have devised an auction mechanism named GenAution, building upon the foundation of the VCG mechanism.

Definition 2 (GenAution) *The GenAution mechanism comprises an allocation mechanism $\boldsymbol{x}$ and a payment mechanism $\boldsymbol{p}$, defined as follows:*

Allocation Mechanism $\boldsymbol{x}$: The allocation is determined by the optimal solution of problem $\boldsymbol{Allo}$.

Payment Mechanism $\boldsymbol{p}$: For each bidder n, the payment p_n is calculated as the difference between the total value accruing to all other bidders in the optimal allocation when bidder n is present, and the total value accruing to all other bidders in the optimal allocation when bidder n is absent.

$$\begin{aligned} p_{nt} = &\max_{x} (\sum_{i \neq n, k} v_i x_{itk} \alpha_{itk} \beta_{itk}) \\ &- \max_{x^{-n}} (\sum_{i \neq n, k} v_i x_{itk}^{-n} \alpha_{itk} \beta_{itk}), \end{aligned} \qquad (7)$$

where x indicates that bidder n is present, and x^{-n} indicates that bidder n is absent.

It is noteworthy that while we employ the VCG payment mechanism within GenAution, alternative payment schemes could also be utilized to ensure IC, contingent upon the nature of the bidders. For instance, in scenarios where advertisers are value-maximizers, the first-price payment mechanism could potentially be adopted within GenAution. However, the focus of this paper is not on the payment mechanism itself but rather on efficiency allocation.

Methodology

In this section, we present the overall structure and detailed design of GenAuction. As illustrated in Figure 1, GenAuction consists of three primary components: the Generator, which is to create a set of candidate PVs, the Evaluator, which assesses the candidate PVs to select a winning PV, and a VCG Payment module.

The Generator begins by receiving the feature of Ads and user information as input, mapping them into embeddings. These embeddings are then processed by the transformer-based encoder, which integrates information from selected

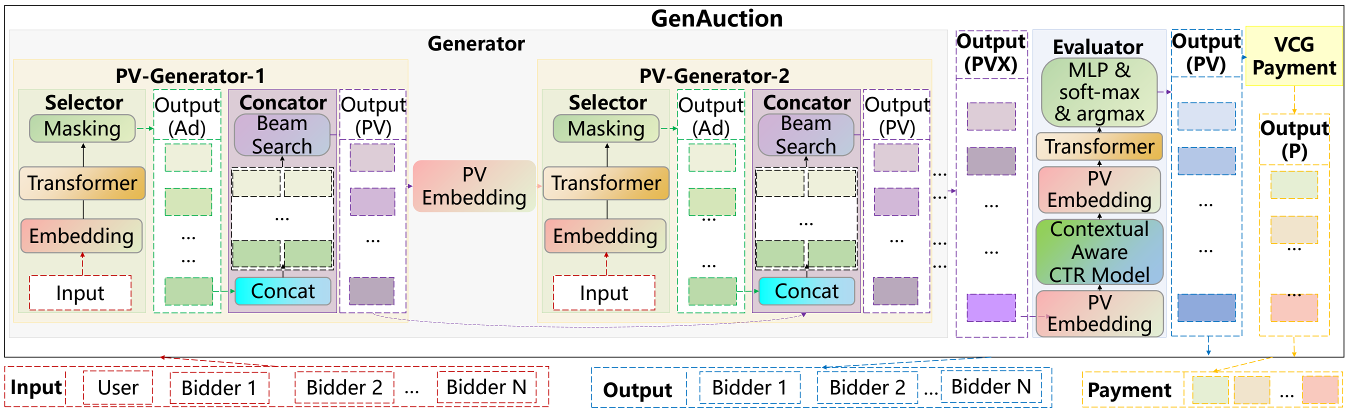

Figure 1: The overall structure of GenAuction, including the Generator responsible for producing candidate PVs and the Evaluator used for assessing candidate PVs and selecting the optimal one.

ads and generates embeddings for subsequent ads based on the previously generated ad sequences. This process is capable of capturing representation of ads. Ads that fail to meet the predefined constraints are subsequently filtered out through a masking strategy. The final set of candidate PVs is produced by the beam search module.

Following the Generator, the Evaluator takes the generated candidate PVs as input, re-embeds these PVs, and applies the attention and final optimization modules to identify the winning PV, which is then displayed to the user. Additionally, the payment module calculates and charges the winning ads in the PV shown to the user. Next, we will provide a detailed description of each module.

Structure of the Generator

In each round $t \in T$, the Generator is composed of X_T cascading PV-Generators, ultimately generating M candidate PVs, which are then passed to the Evaluator. Each PV-Generator primarily consists of two components: a Selector, which chooses M ads from the set of candidate ads, and a Concator, which concatenates these M ads into candidate PVs.

Selector To start with, the first PV-Generator adopts a parameter-sharing embedding layer for all the candidate ads and user information. As shown in Figure 1, we use embedding layers to extract the embeddings from raw sparse features and then concatenate dense features with the embeddings to form d_g-dimensional feature vector $\mathbf{z}_n \in \mathbb{R}^{d_g}$ for the n-th ad and $\mathbf{z}_u \in \mathbb{R}^{d_g}$ for the user's information. Then we stack the features of all candidate ads into the feature matrix $\mathbf{Z}^{\mathbf{A}} \in \mathbb{R}^{N \times d_g}$.

Afterward, the first PV-Generator utilize a self-attention unit (Vaswani 2017) to model the interaction among candidate ads:

$$\mathbf{H}^{\mathbf{A}} = \text{SelfAtt}\left(\mathbf{Q}^{\mathbf{A}}, \mathbf{K}^{\mathbf{A}}, \mathbf{V}^{\mathbf{A}}\right) = \text{softmax}\left(\frac{\mathbf{Q}^{\mathbf{A}}(\mathbf{K}^{\mathbf{A}})^{\top}}{\sqrt{d_g}}\right)\mathbf{V}^{\mathbf{A}},$$

where $\mathbf{Q}^{\mathbf{A}}$, $\mathbf{K}^{\mathbf{A}}$, $\mathbf{V}^{\mathbf{A}}$ represent query, key, and value, respectively. Here query, key, and value are transformed linearly from feature information of all candidate ads:

$$\mathbf{Q}^{\mathbf{A}} = \mathbf{Z}^{\mathbf{A}}\mathbf{W}^{Q}, \mathbf{K}^{\mathbf{A}} = \mathbf{Z}^{\mathbf{A}}\mathbf{W}^{K}, \mathbf{V}^{\mathbf{A}} = \mathbf{Z}^{\mathbf{A}}\mathbf{W}^{V}.$$

For a multi-head version of self-attetion unit, the input is linearly projected into $\mathbf{Q}^{\mathbf{A}}$, $\mathbf{K}^{\mathbf{A}}$ and $\mathbf{V}^{\mathbf{A}}$ with h times using individual linear projections to small dimensions (e.g. $d'_g = \frac{d_g}{h}$). Finally, the output of self-attention (SA) is

$$\begin{aligned} \text{SA} &= [\text{head}_1, \cdots, \text{head}_h]\mathbf{W}^{O}, \\ \text{head}_i &= \text{SelfAtt}(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i). \end{aligned} \tag{8}$$

Afterward, we utilize a multi-layer perceptron to generate the allocation probabilities ($\mathbf{WP}^{\mathbf{A}}$) of all ads.

$$\mathbf{WP}^{\mathbf{A}} = \text{softmax}(\text{FC}(\text{MLP}(\mathbf{H}^{\mathbf{A}}\mathbf{W}^{in})))\mathbf{W}^{out}, \tag{9}$$

where $\mathbf{W}^{in}$ and $\mathbf{W}^{out}$ represent the weight matrices for linear projection.

Then the Transformer layer passes the winning probability matrix $\mathbf{WP}^{\mathbf{A}}$ to the Masking layer, which is responsible for filtering out ads that violate two pre-determined criteria. The first principle is that ads that have already appeared in the candidate list PV_m cannot be added to PV_m again (formula (5)). The other principle is a few of constraints in the industrial ad system, such as diversity constraints, ad quality constraints, etc (formula (4)). Upon passing through the Masking layer, the allocation probability of ads that violate the constraints will be set to zero. Notably, as the initial Selector, the output corresponds to the first ad in the candidate PV. In this case, the constraints do not apply, and the Masking layer does not modify the allocation probability.

Finally, the first Selector can output M optimal ads, $Ad_1^1, \cdots, Ad_M^1$, as the input of the Concator. Let $Ad^1 = (Ad_1^1, \cdots, Ad_M^1)$ represent the set of ads generated by the Selector of PV-Generator-1.

Concator The first Concator takes Ad^1 as input and concatenates these ads onto the M empty PVs in a hierarchical manner, thereby generating M candidate PVs, $PV_1^1, \cdots, PV_M^1$. Let $PV^1 = (PV_1^1, \cdots, PV_M^1)$ represent the set of candidate PVs generated by PV-Generator-1.

Next, the input to PV-Generator-x includes the input to PV-Generator-1, with $x \in X_T$. However, within the Transformer layer of the Selector, the feature embeddings of PV^{x-1} are incorporated. The Selector then outputs M optimal ads, denoted as $Ad^x = (Ad_1^x, \cdots, Ad_M^x)$. The Concator takes Ad^x as input and concatenates these ads onto PV^{x-1} in a hierarchical manner, generating M^2 candidate PVs. Afterward, the Concator utilize beam search module to select M PVs, $PV_1^x, \cdots, PV_M^x$, as the output of the PV-Generator-x. Let $PV^x = (PV_1^x, \cdots, PV_M^x)$ represent the set of candidate PVs generated by the Selector of PV-Generator-x.

Loss Function We use the Lagrangian method to solve the constrained training problem $\boldsymbol{Allo}$:

$$\mathcal{L}_g = -\sum_{n,t,k}[b_n x_{ntk}\alpha_{ntk}\beta_{ntk} + \lambda_i \sum_i (f_i(x_{ntk}) - thr_i)]$$

where $f_i(x_{ntk})$ is the i-th performance metric, and λ_i is a Lagrangian multiplier with thr_i serving as the corresponding threshold.

Structure of the Evaluator

In round t, after receiving M candidate PVs from the Generator, the Evaluator assesses these PVs and computes the rewards for the winning PV. Unlike the Selector in the Generator, which selects the winning ads at the ad level, the Evaluator performs evaluations at the PV level. Intuitively, since PVs are filled incrementally in the Generator, the ads placed below will take the information of the ads above into consideration. However, once Ad^x is concatenated to PV^{x-1} in the Concator, Ad^x can exert an influence on the ads above it. Furthermore, if all X_t ad slots within a PV are filled with ads, the contextual information throughout the entire PV may change. The Evaluator evaluates the candidate PVs based on the contextual information within the entire PV and then selects the optimal PV^t.

In each round $t \in T$, For each candidate PV $PV_m^{X_t} \in PV^{X_t}$, containing X_t ads, the embedding layers of the Evaluator will map the PV level features into embeddings. Then the contextual aware CTR model will re-evaluate and updates the CTR information of these ads in candidate PV $PV_m^{X_t}$. We adopt the position-aware context aggregation (Hou et al. 2023) as the core technique of the contextual aware CTR model. Notably, in order to model the display context for the k-th in each PV, we use the attention mechanism to capture the absolute position and relative position of the k-th ad. With the utilization of contextual aware CTR model, we are capable of enhancing the representation ability of the Evaluator, thereby leading to better performance.

The attention layers of the Evaluator is similar to these of the Generator. We initialize an embedding vector for the k-th ad in the candidate PV as $\mathbf{z}_k \in \mathbb{R}^{d_e}$. The feature matrix of the candidate PV $PV_m^{X_t}$ is represented as $\mathbf{Z}^M \in \mathbb{R}^{X_t \times d_e}$. Then we apply projection to the feature matrix:

$$\mathbf{Q^M} = \mathbf{Z^M}\mathbf{W}_e^Q, \mathbf{K^M} = \mathbf{Z^M}\mathbf{W}_e^K, \mathbf{V^M} = \mathbf{Z^M}\mathbf{W}_e^V.$$

Afterward, we apply the transformation in Eq. (8) to $\mathbf{Q^M}$, $\mathbf{K^M}$, $\mathbf{V^M}$ to obtain the output representation. Finally, we apply Eq. (9) to generate the winning probability WP^m for each PV and screen out the optimal one PV^t. Similar to the generator, the loss function is also a Lagrangian version of the constrained optimization problem $\boldsymbol{Allo}$:

$$\mathcal{L}_e = -\sum_t[\max_{n,k}\sum_{n,k}(b_n x_{ntk}\alpha_{ntk}\beta_{ntk} + \lambda_i \sum_i (f_i(x_{ntk}) - thr_i))]$$

VCG Payment

Upon determining the optimal PV PV^t, the subsequent task is to calculate the payments for winning ads within PV^t. GenAuction adopts VCG mechanism (in Eq. 7) for this goal, uitlizing its capability of ensuring IC and maximizing SW.

Training Procedure

We model our problem as a constrained optimization problem. Although we can integrate the constraints into the objective function using the Lagrangian method, in industrial practice, the process of optimizing the loss function may lead to some constraints being violated, which could affect the efficiency of the platform or harm the user experience. This phenomenon is similar to reinforcement learning, where, after taking certain actions, the agent receives positive or negative feedback and needs to dynamically adjust her actions to maximize the reward. Similarly, the platform needs to ensure that the performance is optimized while obeying certain constraints. With this connection, we model the multi-round optimal PV selection problem as a Constrained Markov Decision Process (CMDP) and solve it through a model-free method, Deep Deterministic Policy Gradient (DDPG). Now we introduce the core elements:

State Space $\mathcal{S}$. A state $s \in \mathcal{S}$ contains the information of ads. Specifically, the following information are essential to represent state: 1) continuous features, such as bid, tCPA, CTR and CVR. 2) discrete information, like the account id, trade id, title, etc. Specifically, for the n-th ad, s_n equals to $\mathbf{z}_n$, which is defined in methodology.

Action Space $\mathcal{A}$. Formally, we denote

$$a^m = \begin{cases} x_{11} & \cdots & x_{1X_t} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NX_t} \end{cases},$$

by decision whether to select the n-th ad to form the candidate PV $PV_m^{X_t}$, where

$$x_{nk} = \begin{cases} 1, & \text{the } n\text{-th ad displayed at slot } k \\ 0, & \text{otherwise} \end{cases}$$

Afterward, the action $a \in \mathcal{A}$ is defined as $a = (a^1, \cdots, a^M)$, representing the decision matrix for PV^{X_t}.

Transition Function $\mathcal{T}$. Concretely, $\mathcal{T}: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ represent the transition from state s_t to state s_{t+1} after adopting action $a \in \mathcal{A}$.

Reward Function $\mathcal{R}$. Specifically, $\mathcal{S} \times \mathcal{A} \to \mathbb{R}$. The reward $\mathbf{r} \in \mathcal{R}$ contains multiple metrics, e.g., social welfare, query landing page relevance, show ratio. Formally,

$$r_t = \sum_{n,k} \{ b_{nt} x_{ntk} \alpha_{ntk} \beta_{ntk} - \sum_i \max\left[0, \left|\lambda_i f_i\left(x_{ntk}\right) - thr_i\right|\right] \},$$

where $f_i(x_{ntk})$ is the i-th performance metric, and λ_i is a Lagrangian multiplier with thr_i serving as the corresponding threshold.

Discount Factor γ. The factor γ determines how much the agent cares about rewards in the future.

Then we define the allocation policy $\phi(a|s) : \mathcal{S} \rightarrow \mathcal{A}$ mapping from state space to action space. With the allocation policy, we can directly form a Q-function $Q^{\phi}(s, a)$ to obtain reward on taking action a at state s:

$$Q^{\phi}(s_t, a_t) = \mathbb{E}_{(s_{t'}, a_{t'}) \sim \phi}[r(s_t, a_t) + \sum_{t'=t+1}^{\infty} \gamma^{(t'-t)} \cdot r(s_{t'}, a_{t'})]$$

Our goal is to optimize the policy $\phi(a|s)$, with which the agents can maximize their cumulative reward:

$$\max_{\phi} \mathcal{J}(\phi) = \mathbb{E}_{s_t \sim \rho_t^{\phi}(\cdot), a_t \sim \pi(\cdot|s_t)}[Q^{\phi}(s_t, a_t)],$$

where $\rho_t^{\phi}(\cdot)$ represents the discounted state distribution under the policy ϕ. The whole training algorithm process is presented in Algorithm 1.

Algorithm 1: Transformer-based DDPG

Initialize a $M+1$ dimensional random process $\mathcal{P}$;
Initialize replay memory $\mathcal{D}$ with capacity $\mathcal{C}$;
Initialize actor ϕ_θ with weight θ and critic Q_ζ with weight ζ;
while *not convergence* **do**
  for $t = 1$ **to** T **do**
    Set $cnt = 0$;
    while *be able to add new ad* **do**
      Observe state s_t^{cnt};
      Take action a_t^{cnt} and get reward r_t^{cnt};
    end
    Store $(s_t^0, a_t^0, \mathbf{r}_t^0, s_t^1, a_t^1, \mathbf{r}_t^1, s_t^2, a_t^2, \mathbf{r}_t^2)$ in $\mathcal{D}$;
    Sample a random batch of $\mathcal{S}$ samples from $\mathcal{D}$;
    Update critic network Q_ζ by minimizing the loss $\mathcal{L} = \frac{1}{T|\mathcal{S}|} \sum^{|\mathcal{S}|} \sum_t^T (Q_\phi(s_t, a_t) - y(s_t, a_t, d))^2$, where $y\left(s_t, a_t, d\right) = r_t^{\text{ecpm}} + \gamma(1-d) Q_\zeta^{\text{targ}}\left(s_{t+1}, \phi_\theta^{\text{targ}}\left(s_{t+1}\right)\right)$;
    Update actor network ϕ_θ by policy gradient $\nabla_\theta J = \frac{1}{T|\mathcal{S}|} \sum^{|\mathcal{S}|} \sum_t^T \nabla_\theta \phi(s^k) \nabla_a Q^{\phi}(s^k, a)|_{a=\phi(s^k)}$;
    Update target network parameters $\theta^{\text{targ}} \leftarrow \rho\theta^{\text{targ}} + (1-\rho)\theta, \phi^{\text{targ}} \leftarrow \rho\phi^{\text{targ}} + (1-\rho)\phi$
  end
end

Experiments

In this section, we conduct both offline and online experiments using real-world datasets to evaluate the performance of GenAuction in multi-objective allocation tasks.

Experimental Settings

Dataset The dataset we use for experiments comes from Baidu, a prominent search engine in China. We randomly sample 500 advertisers' participation in 5.5 million auction rounds from a 30-day online log. This comprehensive dataset is then partitioned into a training set and a test set at a ratio of 9:1 to ensure the rigor and validity of our experiments. For each auction round, our dataset contains a comprehensive suite of information, encompassing the timestamp of the query, user-specific details pertinent to the query (such as user features and search keywords), and advertiser-related attributes (including bid amounts, pCTRs pCVRs, advertisement content, and sets of mutually exclusive ads).

Additionally, the dataset incorporates various constraints reflecting real-world business demands. These constraints can be grouped into three categories: the first pertains to style restrictions for individual ads, the second to exposure constraints within a PV, and the third to exposure limitations across different PVs.

To facilitate a direct comparison with the optimization problem $\boldsymbol{Allo}$, we setup our experiments with a fixed number of advertisers $N = 500$, engaging in a total of $T = 5 \times 10^6$ auction rounds. For each round t, we imposed a limit of slots for display up to $X_t = 3$ ads and a maximum of $X = 2T$ ads being displayed across all T PVs. Furthermore, we introduced a tunable threshold Γ_t to quantify the relevance of ads to individual PVs.

Baselines In our experimental setup, we utilize four auction mechanisms prevalent in industrial ad systems as the baselines for comprehensive evaluation and comparison.

- **uGSP.** Utility-based Generalized Second Price Auction is an extensive version of traditional GSP, adopting the rank score r_{ntk} as a linear combination of bid and other metrics: $r_{ntk}(b_n) = \omega_1 b_n \alpha_{ntk} \beta_{ntk} + o_{ntk}$, where o_{ntk} is the weighted score of other metrics, such as user experience: $o_{ntk} = \omega_2 \gamma_{nt}$. The payment of the n-th bidder is $p_{nt} = x_{ntk} \cdot r_{n,t,k+1}^{-1}(b^{k+1})$, where $r_{n,t,k+1}$ is the rank score of the advertiser in slot $k+1$, and b^{k+1} is the bid of the advertiser in slot $k+1$. Notably, uGSP is capable of selecting ads in greedy manner within single PV.
- **VCG.** VCG is the mechanism aiming to maximize SW (Vickrey 1961; Clarke 1971; Groves 1973). The payment of VCG is the same with GenAuction.
- **tVCG.** The VCG mechanism necessitates the enumeration of all permutations of bidders, followed by the selection of the PV that maximizes the SW. However, this approach becomes computationally prohibitive as the number of bidders grows. To mitigate this issue, the truncated

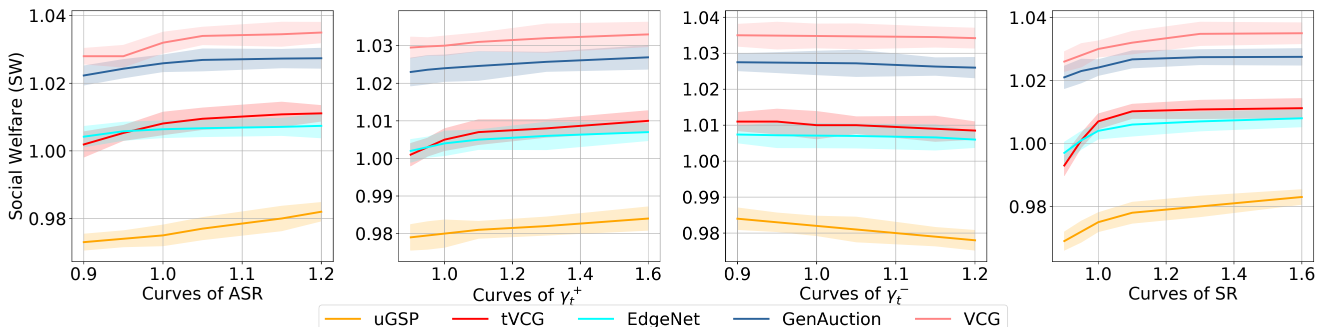

Figure 2: The relationship between cumulative SW and other performance metrics under different mechanisms.

VCG (tVCG) mechanism initially filters the set of bidders by selecting those with high bids, effectively reducing the pool of candidates. Subsequently, VCG is applied solely to this reduced set of high-bidding bidders.

- **EdgeNet.** EdgeNet employs a Transformer-based algorithm to sequentially select ads for constructing PVs (Shen et al. 2023).

Evaluation Metrics To verify GenAuction's superiority over baselines, we adopt several metrics for evaluations:

- **Social Welfare**. SW $= \sum_{n,t,k} b_n x_{ntk} \alpha_{ntk} \beta_{ntk}$.
- **PV Relevance**. We denote γ_t^- as the upper bound on the total quantity of ads for which the relevance metric γ_{nt} is negative, indicating a lack of correlation between ads and the PV. Conversely, γ_t^+ represents the upper limit on the aggregate number of ads whose γ_{nt} is positive, signifying a positive correlation between ads and the PV.
- **Show Ratio (SR)**. We denote $SR = \frac{\sum_{n,t,k} x_{ntk}}{\sum_t X_t}$ by the show ratio of displayed ads across (at most) K slots within T rounds.
- **Ad Show Ratio (ASR)**. We define $\mathbb{1}_t$ as the indicator of whether the PV containing ad or not in round t. Hence, $ASR = \frac{\sum_t \mathbb{1}_t}{T}$.

Offline Evaluations

We compare GenAuction with other baselines through offline experiments. For ease of presentation, we only compare two performance metrics simultaneously in the form of SW $+ \lambda \cdot C$, where C is a performance metric selected from $\{\gamma_t^+, \gamma_t^-, SR, ASR\}$. We plot the curve of the SW as it changes with other performance metrics for different mechanisms, as shown in Fig. 2, where the range of λ variation is marked on the x-axis of each subplot. We begin by selecting a set of metrics $\{\gamma_t^+, \gamma_t^-, SR, ASR\}$, conducting the experiment under the tVCG mechanism, and obtaining the corresponding SW for this set of metrics. Subsequently, we run the experiment using other mechanisms to obtain their respective SW values, which are then normalized by dividing by the SW value from the tVCG mechanism to yield a relative SW index. Finally, we adjust metrics and observe the variations in the calibrated SW. Notably, each offline experiment is repeated 5 times over different random seeds.

Metric	Social Welfare
Improved Ratio	$+8.1\%$

Table 1: The result of online A/B test.

We have found that the curves of GenAuction are above the curves of uGSP, tVCG and EdgeNet, while just below these of VCG, which is the optimal solution to our problem. However, due to the vast scale of candidate ads in the industry, using VCG to compute the scores of all potential ad pages formed by combinations of candidate ads and selecting the optimal ad page not only consumes significant computational resources but also leads to high latency, which can negatively impact user experience. Therefore, VCG can only serve as a theoretical optimal solution. In real-world scenarios, our GenAuction is more practical and valuable for use.

Online A/B test

We present the online experiments by deploying the proposed GenAuction on Baidu through online A/B test. For each PV, the platform returns an ordered list with up to three ads at the top. We choose tVCG as our baseline, a rule-based method, which balance the efficiency and performance. In detail, we calculate hundreds candidate PVs' social welfare under tVCG. To evaluate the performance of tVCG and GenAuction, we conduct online A/B tests using 10% of production traffic over a span of 30 consecutive days.

We mainly focus on the performance of social welfare in the experiments. The results in online A/B tests are shown in Table 1. From the results, we can find that GenAuction achieves obvious promotion for social welfare while these constraints maintaining obeyed compared with tVCG.

Conclusion

In this paper, we propose a KPI-based auction mechanism for ad allocation at the PV-level. To this end, we design a generative auction mechanism, GenAuction, which incorporates a Generator-Evaluator structure based on Transformer. Through extensive evaluations with real-world data, we validate that GenAuction can efficiently handle multi-objective allocation tasks, demonstrating its effectiveness and suitability for addressing complex advertising scenarios.

Acknowledgments

This work was supported by National Natural Science Foundation of China No. 62472428, the Fundamental Research Funds for the Central Universities, the Research Funds of Renmin University of China (No. 22XNKJ07, No. 23XNH028) and Major Innovation & Planning Interdisciplinary Platform for the "Double-First Class" Initiative, Renmin University of China.

References

Bachrach, Y.; Ceppi, S.; Kash, I. A.; Key, P.; and Kurokawa, D. 2014. Optimising trade-offs among stakeholders in ad auctions. In *Proceedings of the fifteenth ACM conference on Economics and computation*, 75–92.

Chawla, S.; and Hartline, J. D. 2013. Auctions with unique equilibria. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, 181–196.

Clarke, E. H. 1971. Multipart pricing of public goods. *Public choice*, 17–33.

Deng, Y.; and Lahaie, S. 2019. Testing dynamic incentive compatibility in display ad auctions. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1616–1624.

Deng, Y.; Lahaie, S.; Mirrokni, V.; and Zuo, S. 2020. A data-driven metric of incentive compatibility. In *Proceedings of The Web Conference 2020*, 1796–1806.

Despotakis, S.; Ravi, R.; and Sayedi, A. 2021. First-price auctions in online display advertising. *Journal of Marketing Research*, 58(5): 888–907.

Dupret, G. E.; and Piwowarski, B. 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*, 331–338.

Dütting, P.; Feng, Z.; Narasimhan, H.; Parkes, D.; and Ravindranath, S. S. 2019. Optimal auctions through deep learning. In *International Conference on Machine Learning*, 1706–1715. PMLR.

Edelman, B.; Ostrovsky, M.; and Schwarz, M. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review*, 97(1): 242–259.

Farina, G.; and Gatti, N. 2016. Ad auctions and cascade model: GSP inefficiency and algorithms. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Feng, Y.; Hu, B.; Gong, Y.; Sun, F.; Liu, Q.; and Ou, W. 2021. GRN: Generative Rerank Network for Context-wise Recommendation. *arXiv preprint arXiv:2104.00860*.

Feng, Z.; Narasimhan, H.; and Parkes, D. C. 2018. Deep learning for revenue-optimal auctions with budgets. In *Proceedings of the 17th international conference on autonomous agents and multiagent systems*, 354–362.

Fotakis, D.; Krysta, P.; and Telelis, O. 2011. Externalities among advertisers in sponsored search. In *International Symposium on Algorithmic Game Theory*, 105–116. Springer.

Gatti, N.; Lazaric, A.; and Trovò, F. 2012. A truthful learning mechanism for contextual multi-slot sponsored search auctions with externalities. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 605–622.

Geyik, S. C.; Faleev, S.; Shen, J.; O'Donnell, S.; and Kolay, S. 2016. Joint optimization of multiple performance metrics in online video advertising. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 471–480.

Goel, G.; Paes Leme, R.; Schneider, J.; Thompson, D.; and Zhang, H. 2023. Eligibility mechanisms: Auctions meet information retrieval. In *Proceedings of the ACM Web Conference 2023*, 3541–3549.

Gomes, R.; Immorlica, N.; and Markakis, E. 2009. Externalities in keyword auctions: An empirical and theoretical assessment. In *International workshop on internet and network economics*, 172–183. Springer.

Groves, T. 1973. Incentives in teams. *Econometrica: Journal of the Econometric Society*, 617–631.

Han, X.; and Liu, S. 2015. Bayes-Nash Equilibrium of the Generalized First-Price Auction. *Mathematical Problems in Engineering*, 2015(1): 301734.

Hou, X.; Wang, Z.; Liu, Q.; Qu, T.; Cheng, J.; and Lei, J. 2023. Deep Context Interest Network for Click-Through Rate Prediction. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 3948–3952.

Huang, J.; Hu, K.; Tang, Q.; Chen, M.; Qi, Y.; Cheng, J.; and Lei, J. 2021. Deep position-wise interaction network for ctr prediction. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1885–1889.

Hummel, P.; and McAfee, R. P. 2014. Position auctions with externalities. In *Web and Internet Economics: 10th International Conference, WINE 2014, Beijing, China, December 14-17, 2014. Proceedings 10*, 417–422. Springer.

Kempe, D.; and Mahdian, M. 2008. A cascade model for externalities in sponsored search. In *International Workshop on Internet and Network Economics*, 585–596. Springer.

Lahaie, S.; Munoz Medina, A.; Sivan, B.; and Vassilvitskii, S. 2018. Testing incentive compatibility in display ad auctions. In *Proceedings of the 2018 World Wide Web Conference*, 1419–1428.

Lahaie, S.; and Pennock, D. M. 2007. Revenue analysis of a family of ranking rules for keyword auctions. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, 50–56.

Liao, G.; Li, X.; Wang, Z.; Yang, F.; Guan, M.; Zhu, B.; Wang, Y.; Wang, X.; and Wang, D. 2022. NMA: neural multi-slot auctions with externalities for online advertising. *arXiv preprint arXiv:2205.10018*.

Lin, X.; Chen, X.; Wang, C.; Shu, H.; Song, L.; Li, B.; and Jiang, P. 2024. Discrete conditional diffusion for reranking in recommendation. In *Companion Proceedings of the ACM on Web Conference 2024*, 161–169.

Liu, X.; Yu, C.; Zhang, Z.; Zheng, Z.; Rong, Y.; Lv, H.; Huo, D.; Wang, Y.; Chen, D.; Xu, J.; et al. 2021. Neural auction: End-to-end learning of auction mechanisms for e-commerce advertising. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 3354–3364.

Myerson, R. B. 1981. Optimal auction design. *Mathematics of operations research*, 6(1): 58–73.

Rahme, J.; Jelassi, S.; and Weinberg, S. M. 2020. Auction learning as a two-player game. *arXiv preprint arXiv:2006.05684*.

Ren, Y.; Yang, Q.; Wu, Y.; Xu, W.; Wang, Y.; and Zhang, Z. 2024. Non-autoregressive Generative Models for Reranking Recommendation. *arXiv preprint arXiv:2402.06871*.

Shen, G.; Sun, S.; Gao, D.; Song, D.; Yang, L.; Wang, Z.; Shi, Y.; and Ning, W. 2023. EdgeNet: Encoder-decoder generative Network for Auction Design in E-commerce Online Advertising. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 4274–4278.

Shen, W.; Tang, P.; and Zuo, S. 2018. Automated mechanism design via neural networks. *arXiv preprint arXiv:1805.03382*.

Shi, X.; Yang, F.; Wang, Z.; Wu, X.; Guan, M.; Liao, G.; Yongkang, W.; Wang, X.; and Wang, D. 2023. PIER: Permutation-Level Interest-Based End-to-End Re-ranking Framework in E-commerce. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 4823–4831.

Thompson, D. R.; and Leyton-Brown, K. 2013. Revenue optimization in the generalized second-price auction. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, 837–852.

Varian, H. R. 2007. Position auctions. *international Journal of industrial Organization*, 25(6): 1163–1178.

Vaswani, A. 2017. Attention is all you need. *arXiv preprint arXiv:1706.03762*.

Vickrey, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1): 8–37.

Wang, Y.; Liu, X.; Zheng, Z.; Zhang, Z.; Xu, M.; Yu, C.; and Wu, F. 2022. On designing a two-stage auction for online advertising. In *Proceedings of the ACM Web Conference 2022*, 90–99.

Xi, Y.; Liu, W.; Dai, X.; Tang, R.; Zhang, W.; Liu, Q.; He, X.; and Yu, Y. 2021. Context-aware reranking with utility maximization for recommendation. *arXiv preprint arXiv:2110.09059*.

Zhang, Y.; Yang, Z.; Wang, L.; and He, L. 2019. Autor3: Au tomated R eal-time R anking with R einforcement Learning in E-commerce Sponsored Search Advertising. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2499–2507.

Zhang, Z.; Liu, X.; Zheng, Z.; Zhang, C.; Xu, M.; Pan, J.; Yu, C.; Wu, F.; Xu, J.; and Gai, K. 2021. Optimizing multiple performance metrics with deep GSP auctions for e-commerce advertising. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, 993–1001.