Advertiser-First: A Receding Horizon Bid Optimization Strategy for Online Advertising

Ke Fang, *Student Member, IEEE*, Hao Liu, Chao Li, *Member, IEEE*, Junfeng Wu, *Senior Member, IEEE*, Yang Tan, *Student Member, IEEE*, Qiuqiang Lin, and Qingyu Cao

Abstract

Online advertising has been the mainstream monetization approach for Internet-based companies, in which bid optimization plays a crucial role in enhancing advertising performance. Currently, the bid optimization problem has narrowed down to two specific forms: Budget-Constrained Bidding (BCB) and Multi-Constraint Bidding (MCB). Existing solutions try to solve BCB/MCB via linear programming solvers, learning methods, or feedback control. However, in large-scale complex e-commerce, they still suffer from inefficiency, poor convergence, or slow adaptation to the changing market. This research presents an online receding optimization method as a solution for practical bid optimization problems. We conduct a theoretical analysis of the optimal bidding strategy's structure. Further, an online receding optimization process is designed based on open-loop feedback control, which periodically updates a constructed optimal bid formulation that can be solved by linear programming. Then, considering large-scale linear programming problems, we propose an efficient down sampling scheme. Besides, a neural-network-based auction scale prediction is used to adapt to the changing market. Finally, a series of online A/B experiments on *Taobao Sponsored Search* compare our work to industrial methods and state-of-the-art from several aspects. The proposed method has been implemented on *Taobao*, a billion-scaled online advertising business, for over a year.

Index Terms

Online advertising, Bid optimization, Dynamic programming, Bootstrap

I. Introduction

Online advertising is a marketing channel that advertisers use to touch potential consumers with their promoted items/services via online media, be it the feeds in online social network apps, the items in e-commercial websites, etc. [1], [2].

For e-commercial platforms, online advertising is leveraged to facilitate purchases [3]–[6]. In Fig. 1, the black dashed box depicts the process: when a user views and triggers an *impression*, which refers to an opportunity to display advertisements. On the backend of websites/Apps, dozens of interested advertisers are recalled to compete in an auction with individual bids, each hoping to win the impression to display their advertisements. Then, an auction agent employs the Second Price Auction (SPA), where the highest bid wins and would pay for the second-highest bid [1]. Typically, a pay-per-click model is adopted, meaning the winner's advertisement is *displayed for free and charged once clicked* [7]. Ultimately, the user sees the winner's advertisement and may respond by clicking, purchasing, etc. The time from triggering an impression to displaying the advertisement is within 0.1 seconds to avoid long page loading [1], [8].

On the other hand, advertisers set up an advertising campaign for each advertisement, encompassing the target groups, marketing objectives, constraints, etc. [1]. These constraints include the budget and cost per action (CPA). Fig. 1 shows that only the winner's advertisement can be displayed, which requires advertisers must bid carefully in each auction to meet their objectives. Therefore, designing an efficient bidding strategy, also known as *bid optimization*, is crucial for achieving better advertising performance [1], [8].

Bid optimization has been a focal point of research and industrial application. Bid optimization has been extensively investigated in both research and industry [1], [8]–[12]. Recently, it has converged to *BCB* (Budget-Constrained Bidding) and *MCB* (Multi-Constraints Bidding). BCB only focuses on the budget constraint, whereas MCB takes into both the budget and CPA constraints. And BCB can be seen as a special case of MCB. Addressing BCB/MCB primarily hinges on developing an efficient bidding strategy informed by historical auction logs. Given a bunch of auction logs, the problem is first formulated using functional analysis [9], then relaxed to combinatorial optimization [5]. Most existing works try to solve it via mixed-constructed strategies or learning methods [1]. But the problem remains challenging, and the previous works have the following deficiencies worthy of improvement:

- **(C1) Limited representation**. Most existing works try to characterize the problem with finite normalized features. Then, these features are employed to train a strategy and generalize to all advertising campaigns [1]. Such generalization seems

Manuscript received

Ke Fang, Chao Li, and Yang Tan are with the College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, China. (email: {nruticat, chaoli, tymine}@zju.edu.cn). Hao Liu, Qiuqiang Lin, and Qingyu Cao are with Alibaba Group, Hangzhou, Zhejiang, China. (email: {liubo.lh, linqiuqiang.lqq, qingyu.cqy}@alibaba-inc.com). Junfeng Wu is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, China. (email: junfengwu@cuhk.edu.cn)

Ke Fang and Hao Liu contributed equally to this work. Corresponding author: Junfeng Wu.

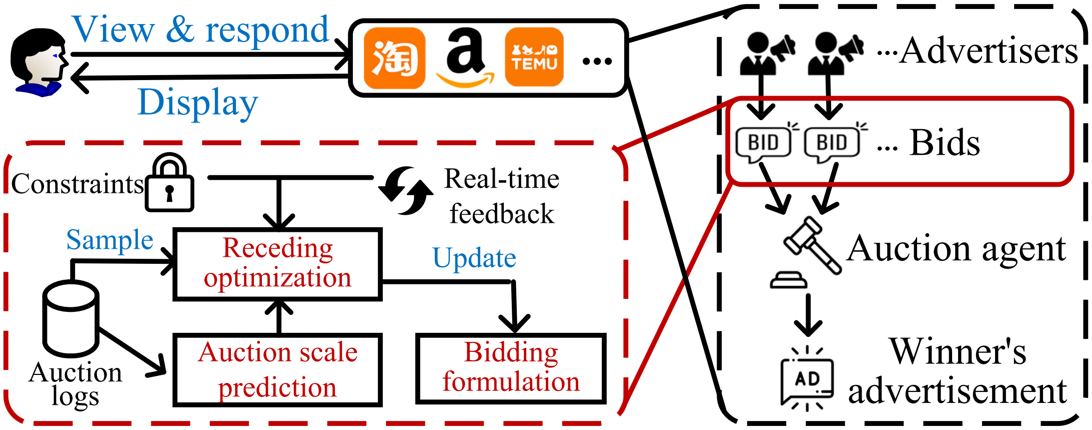


Fig. 1. The process of online advertising (the black dashed box) and the proposed bid optimization method (the red dashed box).

promising but usually embeds very few features due to the poor convergence of the training process, resulting in inadequate performance.

- **(C2) Personalization and adaptability.** In practice, past studies often yield in similar and non-personalized regulation processes. Besides, they typically optimize or train using the stationary auction log and do not fully utilize real time platform feedback for adjustment, making them vulnerable to market volatility [13].
- **(C3) Large-scale operations and real-time processing**. E-commerce platforms typically manage millions of advertising campaigns, each involving millions of auctions daily. Handling such a large scale in real time is daunting and significantly complicates the task of achieving an optimal solution for each advertising campaign [1], [8].

This paper proposes a receding horizon bid optimization methodology to tackle BCB/MCB, outlined by the red dashed box in Fig. 1. Initially, a theoretical analysis is conducted on the structure of the bid optimization solution, which reveals that it is a dynamic, complex strategy closely tied to each advertising campaign's impressions and consumption. Instead of using a few features to model the strategy, we construct an optimal bidding formulation as the strategy, elegantly circumventing (C1). Further, for (C2), each advertising campaign's formulation is solved by linear programming that takes into account its own impressions. In addition, it is periodically updated by a carefully designed receding optimization process, which timely incorporates auction scale prediction and real-time platform feedback on advertising performance. Lastly, an efficient down sampling scheme is introduced to address (C3). Our main contributions are:

(a) We proposed a novel bid optimization method for solving BCB/MCB in online advertising. It has been validated through a scientific month-long online A/B experiments on *Taobao* and implemented for over a year.
(b) A receding optimization process is devised to circumvent (C1) while addressing (C2). Each advertising campaign utilizes linear programming to periodically update a constructed optimal bidding formulation based on its impressions and real-time advertising performance.
(c) Further, an efficient down sampling scheme is introduced to address (C3), resulting in $90\%+$ accuracy with around 3% of the data. In addition, an auction scale prediction is designed to to enhance the capability in (C2).

The remainder is organized as follows. Section II summarises the related works. The bid optimization is formulated in Section III. Section IV introduces our work. The online A/B experiments are discussed in Section V. Section VI discusses limitations and further work. Section VII concludes the paper.

II. Related Works

Typically, the Demand-Side Platform (DSP) facilitates advertisers in securing impressions, serving as an intermediary that consolidates varied advertising requests from multiple online media sources. While DSP primarily focuses on bid optimization to maximize individual gains for a single advertiser, comprehensive platforms, equipped with their proprietary media and extensive advertisers base, alter their focus towards maximizing social welfare. This strategic shift aims to enhance overall returns. In this context, we will review the literature on bid optimization through the perspectives of individual bid strategies and social welfare-aware strategies.

A. Individual bid optimization

In the early stage, advertisers tend to set a fixed bid for a query set or a targeted crowd, and pacing/optimal threshold throttling is utilized to determine whether to bid or not to optimize the performance [10]. Apparently, rigorous bid ranges might miss high-quality but expensive impressions. Besides, the diversification of advertising services has led advertisers to shift from strict bid limits to imposing only daily budget and/or CPA constraints, necessitating a more flexible bid optimization approach [8]. Thus, Zhang *et al.* [9] established the optimal bidding formulation under SPA. This was derived through functional analysis, which models the bidding strategy as a weighted summation of related impression features and assumes prior knowledge of those stochastic models. Yang et al. [5] achieved similar results by formulating the problem as a combinatorial optimization.

After that, solving the bid optimization reduces to finding the optimal weights of the constructed bidding formulation. Nevertheless, once the variable dimension becomes million-sized, and there are millions of such problems, it becomes intractable to apply it in practice. Generalizations seem promising by solving a small set of problems and applying the experience to all the others. *Feedback Control (FC)* strategies are the first attempt, employing a two-step method: initially planning the budget consumption and CPA for the day to establish references, followed by a feedback controller that adjusts weights to ensure current consumption and CPA track these references. Zhang *et al.* [9] and Karlsson [14] proposed error feedback for BCB, and Yang *et al.* [5] provided a double Proportional–Integral–Derivative (PID) for MCB.

More recently, Machine Learning (ML) have recently been widely adopted due to its power and elasticity in solving complex decision-making problems. For BCB, Nedelec *et al.* [15], Zhai *et al.* [16], Grislain *et al.* [11] explored various network architectures. For MCB, He *et al.* [4] provided the *Unified Solution to Constrained Bidding (USCB)* to address CPA constraints, which introduces the ratio of the current CPA and the constraint as an additional state. However, directly applying ML to MCB is not straightforward and requires careful reward shaping for CPA constraints. Although applicable, the optimality might not hold since CPA constraints lay an overall constraint of the whole auction process and can not be accumulated like the budget. Besides, to deal with all advertising campaigns, ML can only embed very few features due to the convergence and training consumption.

B. Social welfare-aware bid optimization

E-commerce platforms can improve overall returns by integrating social welfare into bid optimization, using their exclusive data and computational resources. As a result, social welfare-aware bid optimization emerges alongside selfish bid optimization. Zhu *et al.* [3] proposed the *Optimized Cost Per Click* method to optimize the global return of an impression and respect advertisers' bid limitation. Jin *et al.* [12] proposed a Distributed Coordinated Multi-Agent Bidding framework to optimize the global return but failed to avoid potential collusion, which is unacceptable for the platform. Guan *et al.* [17] tried to model the problem in a *Multi-Agent Cooperation Games (MACG)* manner but explicitly included the platform revenue as a constraint to strike a balance between the rewards of different advertisers and platform revenue. However, advertisers are not motivated to cooperate with their competitors, therefore introducing the risk of using individual bid authorization for the common good.

III. Bid optimization formulation

We formally model the bid optimization problem, considering both BCB and MCB. See Table I for notations. For the bid optimization problem, each auction is characterized by a feature vector $\mathbf{x}$. Assume there are T auctions and under SPA, the problem is to find a bidding strategy π such that:

$$\begin{aligned} \max_{\pi_t} \ & \sum_{t=1}^{T} \mathbb{E}\left[v_t \mathbf{1}_t(w_t, \pi_t)\right] \\ \text{s.t.} \ & \sum_{t=1}^{T} c_t \mathbf{1}_t(w_t, \pi_t) \le B, \ \frac{\sum_{t=1}^{T} \mathbb{E}\left[c_t \mathbf{1}_t(w_t, \pi_t)\right]}{\sum_{t=1}^{T} \mathbb{E}\left[a_t \mathbf{1}_t(w_t, \pi_t)\right]} \le C, \end{aligned} \tag{1}$$

where $b_t = \pi_t(\cdot)$. The indicator function $\mathbf{1}_t(w_t, \pi_t) = \mathbf{1}_t(w_t, b_t) = \{0, 1\}$ denotes the auction result. If $b_t \ge w_t$, $\mathbf{1}_t(w_t, \pi_t) = 1$. Otherwise, $\mathbf{1}_t(w_t, \pi_t) = 0$. Both v and a have varied implementations in practice [1], [8]. The objective is to maximize the total expected v. The first constraint is the total cost cannot exceed B. The second constraint limits CPA.

TABLE I
NOTATIONS

Notation	Description
$\mathbf{x} \in \mathcal{X}$	A high-dimensional impression feature vector.
$v = v(\mathbf{x})$	The estimated value of an impression.
$a = a(\mathbf{x})$	The estimated amount of action of an impression.
b	Bid price.
$w = w(\mathbf{x})$	The second-highest/winning price of an auction.
$c = c(\mathbf{x})$	The cost of an auction.
B	The budget constraint, typically in a day.
C	The cost-per-action (CPA) constraint, typically in a day.

A. Some notes and discussions

Randomness of $\mathbf{x}$: Many features can be used to describe an impression, e.g., cookies, time, etc. It is the feature construction problem that exceeds our interests [13]. Formally, due to the randomness of an impression, we can regard $\mathbf{x}$ as an independent and identically distributed (i.i.d) random variable and follows a probability density function $p_{\mathbf{x}}$ [9]. Consequently, both the objective function and the CPA constraint are defined in terms of the expectation with respect to $\mathbf{x}$.

Feature mappings: Due to the non-unique construction of $\mathbf{x}$, bid optimization commonly employs its evaluated values, e.g., the predicted click-through rate p^{CTR}, the predicted conversion rate p^{CVR}, etc. The problems of value evaluation, however, lie beyond the scope of our interests [8]. For the purposes of modeling and analysis, these values can be treated as functions of $\mathbf{x}$ and are prior to this paper [1], [3]–[5]. Additionally, by combining different evaluated values, we can tailor varying marketing objectives with v and a. For example, if an advertiser aims to maximize the total number of transactions and hold the cost-per-click, we can set $v = p^{\mathrm{CTR}} p^{\mathrm{CVR}}$, $a = p^{\mathrm{CTR}}$. Therefore, v and a are functions of $\mathbf{x}$. Similarly, w is a function of $\mathbf{x}$ because, under SPA, it is the maximum bid among advertisers based on $\mathbf{x}$.

Cost in click advertising: Nowadays, advertisers usually only need to *pay once for a winning impression that has been clicked-through* [7], which is the scenario considered in this paper. Therefore, the cost of an auction, c, is a piecewise function of $\mathbf{x}$: if the winning impression has been clicked, $c = w$; otherwise, $c = 0$.This billing method differs from display advertising with impression settlement considered in the previous works, which pay for the winning impression no matter whether it is clicked or not [1], [4], [5], [8], [14]–[16].

Auction logs: E-commercial platforms like *Taobao* have the full logs, which differs from DSP that only has v of the winning auctions and leads to a censored bid optimization [1]. We claim that the censored version is a special case of bid optimization, and their formulations are universal, only with some pre-treatment of the auction logs. Thus, we elaborate on the details of our work with the following auction logs:

$$\{(d, t, w, p^{\mathrm{CTR}}, p^{\mathrm{CVR}}, i_p)\}_{t=1}^{T}, \tag{2}$$

where $d =$ 'year-month-day' is the date, $t =$ 'hour-minute-second' is a timestamp that represents the occurrence time of the auction. The item price is denoted as i_p.

IV. The Proposed Methodology

The systematic framework of our method, termed *Online receding control with prediction (RCP)*, is outlined in the red dashed box of Fig. 2. It is primarily composed of four parts: (a) Online receding optimization: This core component periodically updates the parameters α and β in the bidding formulation. (b) Bidding formulation module: It uses a constructed formulation to bid a price b based on v and a. (c) Auction logs down-sampling module: It randomly sample a collection of the auction logs and cache it for online usage. (d) Auction scale prediction: It is utilized to assist in adjusting for market Volatility. Additionally, taking into the interaction with the platform and the down-sampling consumption, two more details are discussed in the following,

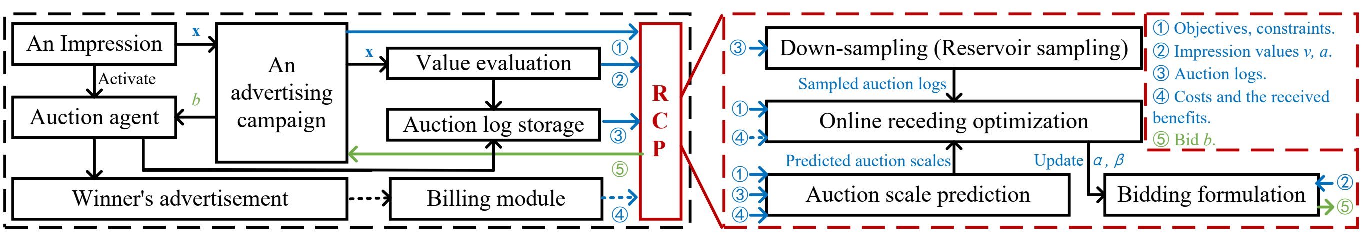


Fig. 2. The systematic framework of our work. The black dashed box is the e-commerce platform, the red dashed box is our work. Solid lines represents real-time interactions, while dashed lines are periodical. For an impression $\mathbf{x}$, it activates an auction agent and recalls interested advertising campaigns. For each advertising campaign, it forwards $\mathbf{x}$ to the value evaluation, which maps $\mathbf{x}$ to v and a. *RCP* then bids a price b and send it back to the agent.

(a) **Billing module.** In click advertising, platforms do not know the cost for each impression until users' responses (click or not) have been observed. However, this time delay is remarkably longer than the interval between two consecutive auctions for an advertising campaign [13], implying it is impractical to update the bidding strategy for each auction. Thus, a billing module is employed to periodically report the received benefits and cost. Consequently, the bidding strategy is updated periodically. After each update, the bidding strategy remains fixed until *RCP* receives new reports.

(b) **Reservoir sampling**. Considering millions of advertising campaigns are on the platform, simultaneously down-sampling is costly. Therefore, we use *reservoir sampling* [18], which runs parallel with the bidding process. The sampled logs are completed at the end of each day, which can be directly and seamlessly used for the next day.

A. Online Receding Optimization Process

Fundamentally, (1) presents a sequential decision-making problem, usually analyzed within the framework of the Markov Decision Process (MDP). Formally, we transform (1) into a finite time horizon MDP $\{\mathcal{S}, \mathcal{B}, p, \mathcal{R}, T\}$:

- **State,** $s_t \in \mathcal{S}$. At the start of the t-th auction, the state $s_t = [B_t, \mathbf{x}_t^\top]^\top$, where B_t is the current consumption.
- **Action (Bid),** $b_t \in \mathcal{B}$. For the t-th auction, the bid price b_t is the action, and $b_t \in [0, B - B_t]$.
- **The state transition probability,** $p_t = \mathbb{P}(s_{t+1} \mid s_t, b_t)$. It is described by two parts: firstly, b_t does not affect $\mathbf{x}_t$ and $\mathbf{x}_{t+1}$ because $\mathbf{x}$ is i.i.d and fully follows $p_{\mathbf{x}}$. Then, given a state B_t and an action b_t, the next state $B_{t+1} = B_t + c_t \mathbf{1}_t(w_t, b_t)$, if $B_t < B$; otherwise, $B_{t+1} = B_t$.
- **Reward,** $r_t = r(s_t, b_t) \in \mathcal{R}$. Please note that we cannot know whether the CPA constraint has been violated until all auctions are completed. Therefore, the reward for the t-th auction is a composite function as follows:

$$r_t = v_t \mathbf{1}_t(w_t, b_t) - \lambda(c_t - Ca_t)\mathbf{1}_t(w_t, b_t), \tag{3}$$

  where $\lambda > 0$ is the Lagrange multiplier. It regards the CPA constraint as a punishment for the objective.
- **Time horizon,** T, which is the total number of auctions.

Therefore, the bid optimization problem (1) is to solve

$$\arg\max_{\pi_t} \sum_{t=1}^{T} \mathbb{E}\left[r_t\right]. \tag{4}$$

We can know that the desired bidding strategy $\pi = \{\pi_t(\mathbf{x}_t, B_t) | t = 1, \ldots, T\}$ is a strategy set that relates to B_t and $\mathbf{x}_t$. Intuitively, Dynamic Programming (DP) can solve it in a data-driven manner. However, there exists two problems:

(a) Given the randomness of $\mathbf{x}$, obtaining the optimal solution is non-trivial and computationally intensive.

(b) The MDP form (4) does not equal to the original problem (1) due to the existence of the CPA constraint. Also, different choices of λ would lead to varying results.

In the context of DP, the optimal strategy performance improves with the availability of extra information; however, it may make the intractable computation [19]. An alternative approach, known as Approximate Dynamic Programming (ADP), deliberately overlooks certain extra information in favor of more computationally efficient solutions. With respect to bid optimization, the current state is known, and future impression features are i.i.d according to $p_{\mathbf{x}}$. Therefore, we introduce *Open-Loop Feedback Control (OLFC)* and construct a receding optimization process to iteratively solve (4). Essentially, as a method derived from the rollout mechanism, OLFC uses current information as feedback to estimate the probability of the future state. Yet, it addresses the problem under the assumption of no further information, employing an open-loop receding time horizon optimization over the system's future evolution. We opt for OLFC because, despite being a form of ADP, it ensures a **non-decreasing** outcome, meaning that the optimized strategy's performance does not deteriorate with each update [19].

Specifically, for the n-th auction, $n = 1, \ldots, T$, we shift to solve the bidding rule $\{\bar{\pi}_n, \ldots, \bar{\pi}_T\}$ as follows:

$$\begin{aligned} &\arg\max_{\bar{\pi}_t} \sum_{t=n}^{T} \mathbb{E}\left[v_t \mathbf{1}_t(w_t, \bar{\pi}_t) - \lambda\Big(B_n + \sum_{t=n}^{T} \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] - CA_n - C\sum_{t=n}^{T} \mathbb{E}[a_t \mathbf{1}_t(w_t, \bar{\pi}_t)]\Big)\right], \\ &\text{s.t.} \;\; B_n + \sum_{t=n}^{T} \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] \le B, \end{aligned} \tag{5}$$

where $A_n = \sum_{t=1}^{n-1} a_t \mathbf{1}_t(w_t, b_t)$ is the received amount of actions. Due to the introduction of OLFC, (5) only maintains the budget constraint in the expected sense, which is inevitable as we pursue a more tractable solution. $\bar{\pi}_n$ is solved by (5), which only relates to $\mathbf{x}_n$. Then, we bid a price $b_n = \bar{\pi}_n(\mathbf{x}_n)$.

To further analyze the bidding rules, we define

$$z_t := \mathbb{E}[\mathbf{1}(w_t, \bar{\pi}_t)] \in [0, 1], \tag{6}$$

where z_t can be regarded as the winning probability for the t-th auction with $\bar{\pi}_t$. Thus, (5) can be rewritten as the optimization problem concerning z_t in the following:

$$\begin{aligned} &\arg\max_{z_t} \sum_{t=m}^{T} \mathbb{E}\left[v_t z_t\right] \\ &\text{s.t.} \;\; \sum_{t=n}^{T} c_t z_t \le B - B_n, \; \frac{B_n + \sum_{t=1}^{T} \mathbb{E}\left[c_t z_t\right]}{A_n + \sum_{t=1}^{T} \mathbb{E}\left[a_t z_t\right]} \le C, \end{aligned} \tag{7}$$

which covers the gap between (4) and (1).

B. The Constructed Optimal Bidding Formulation

Thus far, the core problem is to solve the optimal winning probability z in (7), and the bidding rule $\bar{\pi}$ serves as a mechanism to achieve this probability. However, directly solving (7) is complex due to the absence of the knowledge of $p_{\mathbf{x}}$. Nonetheless, under the i.i.d. assumption, daily auction logs can be treated as a collection of samples for each advertising campaign. This allows us to solve it in a data-driven manner, that is, by replacing expectations with logs:

$$\begin{aligned}&\underset{z_n,\ldots,z_{\hat{T}}}{\arg\max} \sum_{t=n}^{\hat{T}} \hat{v}_t z_t \\ &\text{s.t.} \;\; B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \le B, \;\; \frac{B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t}{A_n + \sum_{t=n}^{\hat{T}} \hat{a}_t z_t} \le C,\end{aligned} \tag{8}$$

where the superscript $\wedge$ means the value recorded in auction logs, $\hat{c}_t = \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] = \hat{w}_t \hat{p}_t^{\mathrm{CTR}}$ because of the click advertising. Clearly, (8) is a linear programming concerning z, and we have various methods to use. *However, the solution cannot be used to bid but only tells us which impressions are worth winning.* Then, we introduce

$$q_t = \frac{\hat{v}_t + \beta^* C \hat{a}_t}{\hat{p}_t^{\mathrm{CTR}}(\alpha^* + \beta^*)}, \tag{9}$$

where parameters α^*, β^* are solved by the dual problem of (8) concerning z [5]. Further, we have the following theorem to reveal the relationship between q and (8).

Theorem 1: For a bunch of auction logs with $\hat{T}$ records, suppose (8) and (9) are feasible. We have the following statements under SPA:

(a) $\forall n = 1, \ldots, \hat{T}$, the optimal solution of (8), $\{z_n^*, \ldots, z_{\hat{T}}^*\}$, has at least $(\hat{T} - n - 1)$ number of 0 and 1.
(b) $\forall n = 1, \ldots, \hat{T}$ and $t = n, \ldots, \hat{T}$, if $q_t > \hat{w}_t$, then $z_t^* = 1$. If $q_t < \hat{w}$, then $z_t^* = 0$. If $q_t = \hat{w}_t$, then $z_t^* = [0, 1]$.
(c) $\forall n = 1, \ldots, \hat{T}$, we use (9) as the bidding strategy and get the auction result, $\{\bar{z}_n^*, \ldots, \bar{z}_{\hat{T}}^*\}$. Then comparing it to the optimal solution of (8), $\{z_n^*, \ldots, z_{\hat{T}}^*\}$, the number of different auction result ($\bar{z}_t^* \neq z_t^*$) would not exceed 2. Meanwhile, this gap is the best that we can have.
(d) If $\hat{T} \gg 2$, the formulation (9) is *one of the optimal bidding strategies* for (8).

Proof 1: For the first statement, We hope to solve

$$\begin{aligned}&\underset{z_t}{\arg\max} \sum_{t=n}^{\hat{T}} \hat{v}_t z_t \\ &\text{s.t.} \;\; B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \le B, \;\; \frac{B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t}{A_n + \sum_{t=n}^{\hat{T}} \hat{a}_t z_t} \le C,\end{aligned} \tag{10}$$

Rewrite it as follows:

$$\begin{aligned}&\underset{z_t, a_t, d, g}{\arg\max} \sum_{t=m}^{\hat{T}} \hat{v}_t \hat{q}_t(\hat{\mathbf{x}}_t) \\ &\text{s.t.} \; g + \sum_{t=m}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} \hat{q}_t(\hat{\mathbf{x}}_t) = B - B_n, \\ &\quad z_t + a_t = 1, \; t = m, \ldots, \hat{T}, \\ &\quad d + B_n - C A_n + \sum_{t=m}^{\hat{T}} z_t \left(\hat{w}_t \hat{p}_t^{\mathrm{CTR}} - C \hat{a}_t\right) = 0, \\ &\quad g \ge 0, d \ge 0, z_t \ge 0, a_t \ge 0, t = m, \ldots, \hat{T},\end{aligned} \tag{11}$$

which contains $(2\hat{T} - 2n + 4)$ variables and $(\hat{T} - n + 3)$ equality constraints. We can define the so-called basic feasible solutions (BFS) for the form of LP problems like (11). The BFS of (11) have at least $(\hat{T} - m + 1)$ number of 0. If (10) is feasible, then (11) is feasible, and there exists an optimal solution that is a BFS for (11). Furthermore, due to the constraint $z_t + a_t = 1$, $\forall t = n, \ldots, \hat{T}$, implying if $a_t = 0$ then $z_t = 1$. Thus, when we only consider $\{z_n, \ldots, z_{\hat{T}}\}$ of a BFS as an optimal solution to (10), it has at least $(\hat{T} - n - 1)$ number of 0 and 1.

For the second statement, since we can write (10) to its dual form as:

$$\begin{aligned}
&\arg\min_{\alpha,\beta,\gamma_t} (B-B_n)\alpha + \sum_{t=m}^{\hat{T}} \gamma_t + \beta(B_n - A_n C) \\
&\quad \text{s.t. } \ \hat{w}_t \hat{p}_t^{\text{CTR}} \alpha + \left(\hat{w}_t \hat{p}_t^{\text{CTR}} - C\hat{a}_t\right)\beta - \hat{v}_t + \gamma_t \geq 0, t = m, \ldots, \hat{T}, \\
&\qquad \alpha \geq 0, \beta \geq 0, \gamma_t \geq 0, t = n, \ldots, \hat{T}.
\end{aligned} \tag{12}$$

Let α^*, β^* be the optimal to (12) and define

$$q_t = \frac{1}{\hat{p}_t^{\text{CTR}}} \frac{\hat{v}_t + \beta^* C\hat{a}_t}{\alpha^* + \beta^*} \tag{13}$$

Based on the dual theory, the optimal solutions of (12) satisfy $\forall t = n, \ldots, \hat{T}$:

$$z_t^* \left\{\hat{w}_t \hat{p}_t^{\text{CTR}} (\alpha^* + \beta^*) - \hat{v}_t - C a_t \beta^* + \gamma_t^*\right\} = 0, \tag{14}$$

$$(z_t^* - 1)\gamma_t^* = 0, \tag{15}$$

Then,

$$z_t^* \left\{\hat{p}_t^{\text{CTR}} (\hat{w}_t - q_t)(\alpha^* + \beta^*) + \gamma_t^*\right\} = 0.$$

If $\mu_t < \hat{w}_t$, then $\hat{p}_t^{\text{CTR}} (\hat{w}_t - q_t)(\alpha^* + \beta^*) + \gamma_t^* > 0$, which further implies that $z_t^* = 0$ by (14). In contrast, $q_t > \hat{w}_t$ implies $z_t^* = 1 > 0$.

For the third and the fourth statements. Acoording to the GSP mechanism, if $q_t > \hat{w}_t$, then the bidder will win the impression, $z_t^* > 0$. If $q_t < \hat{w}_t$, then the bidder will lose the impression. In this case $z_t^* = 0$. When $z_t^* > 0$, there are at most 2 z_t^* that is not 1. Noticing that $\hat{T} \gg 2$, it concludes the proof. □

In fact, the assumption $T \gg 2$ is feasible, as most advertising campaigns have around 20,000 auctions daily [13]. To sum up, given an advertising campaign with the previous day's auction log sorted by timestamps 'hour-minute-second', for the n-th auctioned impression $\mathbf{x}_n$, platforms disclose its timestamp 'hour-minute-second', a_n, v_n, B_n, A_n, and p_n^{CTR}. We then insert $\mathbf{x}_n$ into the log based on its timestamp and remove records before $\mathbf{x}_n$. Further, we solve the dual problem of (8) with the log and bid a price b_n:

$$b_n = \bar{\pi}_n = \frac{v_n + \beta^* C a_n}{p_n^{\text{CTR}}(\alpha^* + \beta^*)}. \tag{16}$$

For Theorem 1, we have the following intuitive example to facilitate understanding,

Example 1: We randomly selected an AD campaign and its auction log on Sep. 25, 2022. The AD campaign participated in 60,690 auctions, hoping to maximize the expected GMV. Its budget constraint is CNY 212.43, and CPA constraint is CNY 1.00 (here, the interested action is set to click), which means $B = 212.43$, $k = 1$, $C_1 = 1.00$. The optimal solution of equ.(9) is $\{h_1^*, \ldots, h_{60690}^*\}$. Then we solve the optimal dual solution α^*, β^* and

$$\bar{\pi}_t = \frac{\hat{v}_t + \beta^* \hat{a}_t}{\hat{p}_t^{\text{CTR}}(\alpha^* + \beta^*)},$$

Further, the auction log is replayed by $\bar{\pi}$ and get auction results $\{h_1, \ldots, h_{60690}\}$. As shown in Fig. 3, we compare $\{h_1^*, \ldots, h_{60690}^*\}$ to $\{h_1, \ldots, h_{60690}\}$, which shows that only two auctions' results are different. For these two auctions, the optimal results solved by equ.(9) are $h_{t_1}^* = 0.6220$ and $h_{t_2}^* = 0.5639$. The bid price provided by $\bar{\pi}$ are 0.6168 and 0.3192, while the winning prices recorded in the log are 4.34 and 0.32, which means that $h_{t_1} = h_{t_2} = 0$.

C. Efficient Solution using Down-sampling

Till now, the primary challenge resides in solving the dual problem of (8). Despite the availability of numerous solvers, computational intensity becomes an issue in the bid optimization. Considering millions of advertising campaigns and the stringent real-time requirement, it is burdensome to provide the required resources. Since (9) is formulated in a data-driven fashion, an intuitive idea to improving efficiency is through *Bootstrap* sampling [20], which solves α and β using a subset of the auction logs rather than the entire logs. Capitalizing on this idea, we introduce an efficient solution employing down-sampling. The pseudo-code for this method is detailed in Algorithm 1. Additionally, we provide a theorem to validate the effectiveness of our approach.

Theorem 2: For BCB, suppose (9) is feasible, there exists an optimal solution α^*. Then as the down-sampling size N increases, the solution α_N converges to α^*, almost surely.

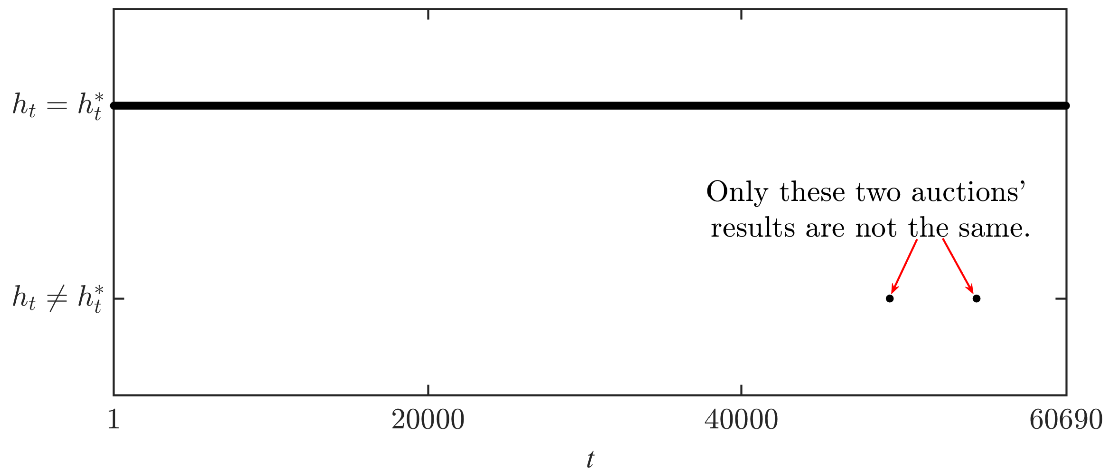


Fig. 3. Comparison of two methods' auction results. The horizontal axis represents the t-th auction. The vertical axis means whether or not h_t^* and h_t are the same.

Algorithm 1 Efficient solution using down-sampling

1: **Input:** Constraints B, C. The previous day's auction log Φ sorted by timestamp 'hour-minute-second'. Down-sampling size N.
2: **Initialization:** $n = 1$, $B_n = 0$, A_n =0, $\hat{T}$ = the number of records in Φ.
3: **Down-sampling:**
4:     **If** $\hat{T} > N$ **then**
5:         $\phi =$ Uniformly sample N records from Φ.
6:     **Else**
7:         $\phi = \Phi$.
8:     **End if**
9: Solve α, β in (9) with B, C, and ϕ. $\alpha_0 = \alpha_N$, $\beta_0 = \beta_N$.
10: **For the n-th auctioned impression $\mathbf{x}_n$:**
11:     Insert $\mathbf{x}_n$ into ϕ and Φ based on its timestamp 'hour-minute-second'. Remove records before $\mathbf{x}_n$.
12:     $N_n =$ the number of records in ϕ.
13:     $\hat{T}_n =$ the number of records in Φ.
14:     **If** $N_n > 0$ **then**
15:         Solve α and β in (9) based on $(B - B_n)N_n/\hat{T}_n$, C, A_n,and ϕ. And $\alpha_n = \alpha_N$, $\beta_n = \beta_N$.
16:     **Else**
17:         $\alpha_n = \alpha_{n-1}$, $\beta_n = \beta_{n-1}$.
18:     **End if**
19:     Update (16) with α_n and β_n, then bid a price.
20:     Wait for the auction result, cost c_n, and benefit a_n.
21:     $B_{n+1} = B_n + c_n$, $A_{n+1} = A_n + a_n$.
22:     $n = n + 1$.
23:     Wait for the next auction, and **go back to Step 10**.
24: **End for**

Proof 2: Consider the following budget-constraint bidding (BCB) problem,

$$\begin{aligned} &\underset{q_t}{\text{maximize}} \sum_{t=1}^{\hat{T}} \hat{v}_t \\ &\text{s.t.} \sum_{t=1}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \leq B, \end{aligned} \tag{17}$$

where the superscript $\wedge$ denotes the value recorded in logs. Then we write its dual form as follows

$$\begin{aligned}
&\underset{\alpha,\gamma_t}{\text{minimize}}\ B\alpha+\sum_{t=1}^{\hat{T}}\gamma_t \\
&\text{s.t.}\ \ \hat{w}_t\hat{p}_t^{\mathrm{CTR}}\alpha-\hat{v}_t+\gamma_t\geq 0,\ \forall t=1,\ldots,\hat{T}, \\
&\qquad \alpha\geq 0, \\
&\qquad \gamma_t\geq 0,\ \forall t=1,\ldots,\hat{T},
\end{aligned} \tag{18}$$

Let $\alpha^*,\gamma_1^*,\ldots,\gamma_{\hat{T}}^*$ denote the optimal solution to (18). By the complementary conditions for the optimal primal variables z_t^* and dual variables α^* of (17) and (18), we have

$$z_t^*\left[\hat{p}_t^{\mathrm{CTR}}\left(\hat{w}_t-\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}\right)\alpha^*+\gamma_t^*\right]=0, \tag{19}$$

$$(z_t^*-1)\gamma_t^*=0, \tag{20}$$

If $\hat{w}_t>\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}$, for the optimal condition (19),

$$\hat{p}_t^{\mathrm{CTR}}\left(\hat{w}_t-\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}\right)\alpha^*+\gamma_t^*>0,$$

implying that $z_t^*=0$ by (19). If $\hat{w}_t<\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}$, for the dual constraint in (18), there exists

$$\hat{w}_t\hat{p}_t^{\mathrm{CTR}}\alpha^*-\hat{v}_t+\gamma_t=(\hat{w}_t-\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}})\alpha^*\hat{p}_t^{\mathrm{CTR}}+\gamma_t\geq 0,$$

implying that $\gamma>0$, and for the condition (20), we readily have $z_t^*=0$. In summary, we have the following relation between z_t* and $\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}$,

$$z_t^*=\begin{cases}1, & \text{if } \frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}>\hat{w}_t,\\ 0, & \text{if } \frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}<\hat{w}_t,\end{cases}, \tag{21}$$

where $z_t^*\in[0,1]$ if $\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}=\hat{w}_t$.

On the other hand, from the feasibility condition of the primal problem (17), we have

$$\alpha^*=\sup\left\{\alpha\mid\sum_{t=1}^{\hat{T}}\hat{w}_tp_t^{\mathrm{CTR}}\mathbf{1}_{(\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}>\hat{w}_t)}\leq B\right\} \tag{22}$$

where $\mathbf{1}(\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}>\hat{w}_t)$ means $\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}>w_t$. $\hat{v}_t$, $\hat{p}_t^{\mathrm{CTR}}$, and $\hat{w}_t$ are functions of the random variable $\mathbf{x}$, which follows a probability density function $p_{\mathbf{x}}$. Then, the optimal solution α^* is the supremum of the above set,

$$\alpha^*=\sup\alpha, \tag{23}$$

Further, when we use Algorithm 1 to solve (18) with a sampling size $N\leq\hat{T}$, the optimal solution is regarded as α_N^*.

Based on (22),

$$\mathbf{1}_{\frac{\hat{v}_t}{\alpha^*\hat{p}_t^{\mathrm{CTR}}}>\hat{w}_t}=\mathbf{1}_{\alpha<\frac{\hat{v}_t}{\hat{w}_t\hat{p}_t^{\mathrm{CTR}}}}, \tag{24}$$

which shows that α is the quantile of $\frac{\hat{v}_t}{\hat{w}_t\hat{p}_t^{\mathrm{CTR}}}$. And α^* is the supremum quantile. Because the sampling process does not affect quantiles, one can use fewer data to estimate quantiles by balancing the accuracy and computation. We have α_N converges to α^* almost surely. Similarly, β_N converges to β^* in MCB almost surely.

□

The essence of solving (9) is a quantile estimation problem for a complex random variable originating from $\mathbf{x}$. The sampling process does not affect the quantile [21]. We further conjecture the above also holds ture for MCB, where α_N and β_N converge to α^* and β^*, respectively, almost surely. However, elaborating on this is challenging due to the involvement of multiple random variables that are coupled with dual variables. To the best of our knowledge, no work has yet addressed the convergence for sampling in solving BCB/MCB. Meanwhile, the mainstream approach remains BCB, which suffices for most scenarios, and the extension to MCB is a primary focus for our future work.

Due to the distribution of $\mathbf{x}$ being non-parametric, Section A proposes a series of experiments for empirically setting a feasible N that balances the accuracy and efficiency (also for verifying the guess of MCB). Finally, we found that $N=10000$ is feasible, *only using 3% data of the whole logs and achieving more than 90% accuracy.*

Finally, we present an example to help understand why α is the quantile. For the feasible condition (22), it is similar to pacing, which selects impressions with high rates of v to wp^{CTR} until the budget is exhausted. Then the last chosen impression's

$v/(wp^{\mathrm{CTR}})$ is regarded as α^* used in its bid formulation. Due to the randomness of $\mathbf{x}$, $v/(wp^{\mathrm{CTR}})$ also is a random variable. Limited by the total budget, α^* relates to a quantile of the distribution of $v/(wp^{\mathrm{CTR}})$, denoted by F_{α^*}. After that, we can win the impression whose v/w is the top $(1 - F_{\alpha^*})$. If we have more budget, F_{α^*} becomes smaller, and we can win more impressions. It is verified in Fig. 4 and Fig. 5, where the budget is multiplied by 1, 10, and 100 times, the corresponding quantiles are $F_{\alpha_1^*}$, $F_{\alpha_{10}^*}$, and $F_{\alpha_{100}^*}$, respectively (The auction log is provided by Taobao on Sep. 15th, 2022.).

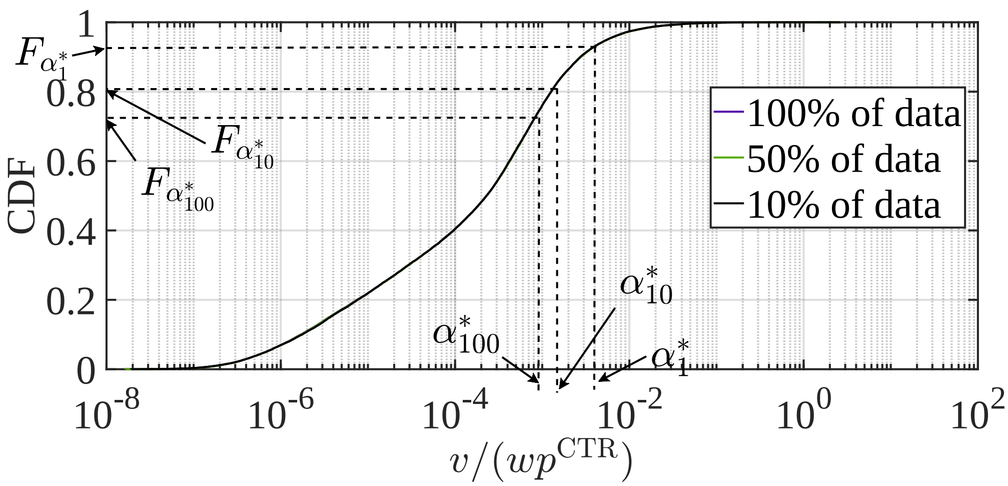


Fig. 4. An example illustrates the idea of using down-sampling to solve parameters efficiently. The horizon axis is the rate of v and wp^{CTR} recorded in the logs; the vertical axis is its CDF. Randomly pick 100%, 50%, and 10% logs and plot their CDFs; they are almost identical, showing the sampling process does not affect the random variable's distribution. Variable α^* is the optimal solution with the whole logs and the subscript means the multiple by which the budget is enlarged. Variable F means the quantile.

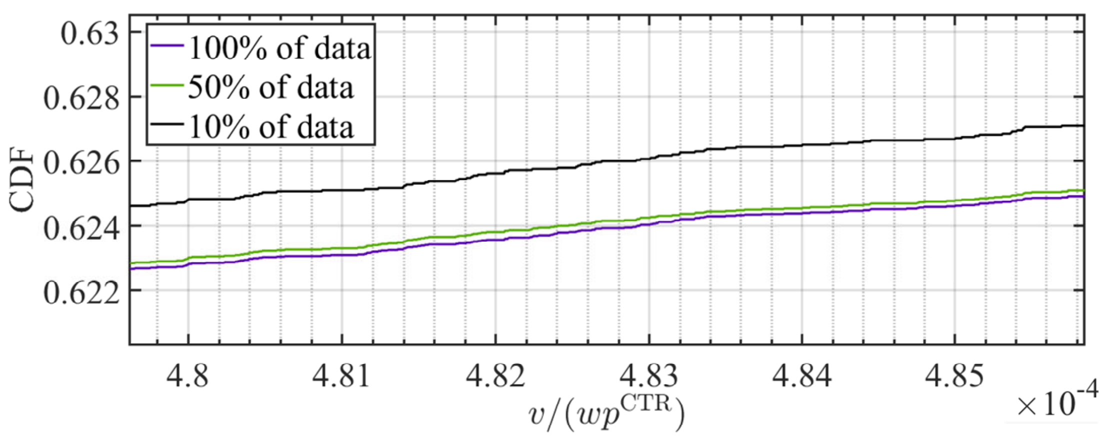


Fig. 5. A zoom-in of Fig. 4 around $\frac{v}{wp^{\mathrm{CTR}}} = 4.83$.

D. Auction Scale Prediction

Predictions are instrumental in compensating for market Volatility, yet it an open problem [1], [11]. In this work, impressions are treated as samples drawn from $\mathcal{X}$, and the introduction of bootstrap allows us to focus the factors that directly affect the solution of (9), which are primarily associated with the constraints related to the residual auction scale. Additionally, an advertising campaign is set to run for a one-day duration [4], [7], [17]. Consequently, our predictive focus is primarily on the residual auction scale for that specific day. However, this presents a challenge as typical advertising campaigns exhibit continuity and temporal relationships within a day. A comparison of two consecutive days reveals periodicity, while a weekly analysis unveils slow time-varying features [11], [13]. Furthermore, practical factors such as advertisers specifying advertising times or prematurely exhausting budgets can lead to incomplete auction logs, potentially causing the advertising campaign to unexpectedly go offline, and thereby misguiding predictions.

The designed auction scale prediction module is shown in Fig. 6. For the auction logs, we divide a day into 96 intervals, each 15 minutes. One reason is to facilitate training convergence since the auction scale varies from zero to over a million for different advertising campaigns. On the other hand, it helps to reveal temporal relationship throughout a day.

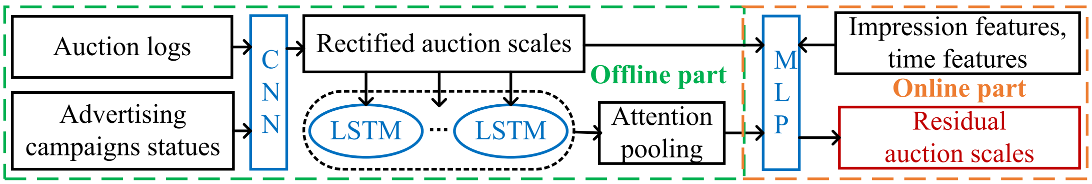


Fig. 6. The auction scale prediction module.

In the offline part, the first is to check whether there exists missing data by using advertising campaigns status. Then, we rectify auction logs by imputing missing data with a Convolutional Neural Network (CNN), inspired by image inpainting [22]. Each advertising campaign's auction logs are rearranged into an image-like form, with days as rows and the chronological auction scales of each interval as pixels. Thus, we obtain the rectified auction scales to establish temporal relationships. Further, 96 Long Short-Term Memories (LSTMs) are employed, known for its superior ability to capture long-term dependencies. However, two new challenges arise: overfitting in advertising campaigns with few impressions and dynamic prediction sequence lengths induced by the online update. Generally, the second issue can be mitigated using pooling, but its standard form is static, which cannot address the first issue. Therefore, we introduce soft attention pooling [16].

In the online part, predicting each interval's auction scale is challenging due to significant magnitude variation. Considering the periodicity, we can use the previous day's rectified auction scales as the baseline and predict its uplifts. Thus, the prediction is decomposed into a classification (predicting increase/decrease) and a regression (predicting the uplift ratio), both of which can be effectively solved by Multi-Layer Perceptron (MLP). Finally, MLP inputs incorporate impression features and time features to account for slow time-varying features of different advertising campaigns.

Honestly, the prediction exceeds our primary scope of interest. This module is a prototype designed to consummate the proposed method, and it has significant potential for future enhancement. Detailed settings and verification experiments are provided in Section B.

V. Online A/B experiments

In summaries, the pseudocode of our method is presented in Algorithm 2. In this section, we conduct online A/B experiments on *Taobao Sponsored Search*, one of China's largest online e-commercial platforms.

A. Evaluation system

Fig. 7, illustrates our budget bucketing system for experiments. At the start of a day, some advertising campaigns are randomly chosen. For each of them, we build some buckets to implement each compared method. The constraints are the same for all buckets, where the budget comes from dividing the original budget into several equal parts. After that, each bucket is isolated from the others. Each auction is randomly assigned a bucket number, and the bucket method is applied. The corresponding cost and benefit are only revealed to the selected bucket. When a bucket exhausts its budget, it can still be chosen but will not participate in the auction. At the end of a day, we tally each bucket's costs and benefits.

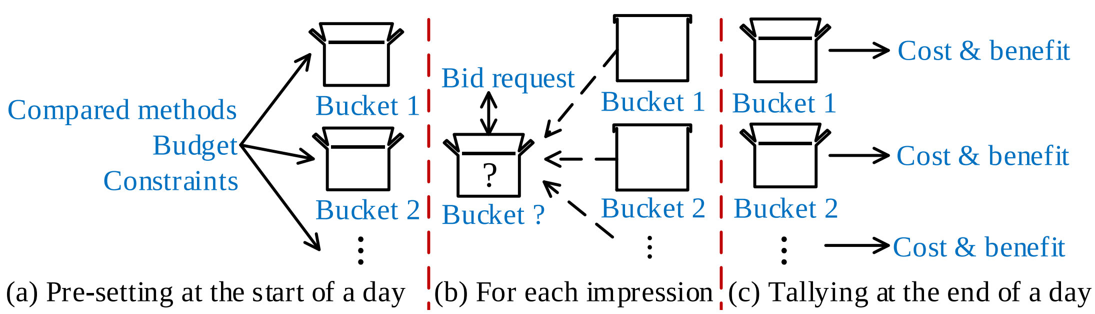


Fig. 7. Budget bucketing system for an advertising campaign.

Algorithm 2 Online receding control with prediction (RCP)

Input: Constraints B, C. The sampled logs ϕ, sorted by timestamp 'hour-minute-second'. Billing module update times n_b.
Initialization: $n=0$, $B_n=0$, $A_n=0$.
  $N_n =$ the number of records in ϕ.
  $\hat{T}_n =$ the residual auction scale predicted by Fig.6.
  Solve α and β based on $BN_n/\hat{T}_n$, C, and ϕ.
  $\alpha_n = \alpha$, $\beta_n = \beta$.
For the n-the update, $n = 1, \ldots, n_b$**:**
  Remove records before the current timestamp 'hour-minute-second' in ϕ.
  $N_n =$ the number of records in ϕ.
  $\hat{T}_n =$ the residual auction scale predicted by Fig.6.
  If $N_n > 0$ **then**
    Solve α and β in (9) based on $(B - B_n)N_n/\hat{T}_n$, A_n, C, and ϕ.
    $\alpha_n = \alpha$, $\beta_n = \beta$.
  Else
    $\alpha_n = \alpha_{n-1}$, $\beta_n = \beta_{n-1}$.
  End if
  Update (16) with α_n and β_n.
  Repeat
    Bid prices based on (16) for each coming impression.
  Until receive reports $\{B_n, A_n\}$ from billing modules.
  $n = n + 1$.
End for

B. Experiment settings

- **Business scenarios:** Both BCB and MCB are considered:
  - **BCB** maximizes the expected Gross Merchandise Volume (eGMV) subject to the budget. Thus, in (1) and (16), $v_t = p_t^{\mathrm{CTR}} p_t^{\mathrm{CVR}} i_p$, $b_t = p_t^{\mathrm{CVR}} i_p / \alpha^*$.
  - **MCB** maximizes eGMV subject to the budget and cost-per-click (CPC) constraints. Thus, $a_t = p_t^{\mathrm{CTR}}$, $v_t = p_t^{\mathrm{CTR}} p_t^{\mathrm{CVR}} i_p$, $b_t = (p_t^{\mathrm{CVR}} i_p + \beta^* C)/(\alpha^* + \beta^*)$.
- **Time:** 09/01/2022 - 09/30/2022.
- **Scales:** Unless otherwise stated, in each day, we randomly selected 2,000 campaigns with the previous day's auction logs for each experiment, totaling around 4 million impressions.
- **Compared methods:** Since our experiments are conducted online, it is imperative to consider both commercial safety and advertising performance. These necessitate that the compared methods must be widely adopted and have practical applications. Consequently, the following methods, which were regarded as the most effective in the latest survey [1] prior to our work, are considered.
  (a) *Multi-agent Cooperation Bidding Games (MACG)* [17], which is the basic strategy used on *Taobao*. We mainly regard its results as the baseline for comparison.
  (b) The *Unified Solution to Constrained Bidding (USCB)* [4], which is a state-of-the-art learning-based method. Its inputs encompass the current impression values, the tracking error in budget consumption, and the tracking error in CPC, etc. It uses (16) to bid and output is its parameters.
  (c) *Feedback Control (FC)* method [5], [14], [23]. This straightforward yet effective strategy for real-time feedback-based adjustments has garnered significant attention and application. Its inputs include the current impression values, the budget consumption tracking error, and the CPC tracking error. It uses (16) to bid and output is its parameters.
  (d) Our proposed method, Algorithm 2, denoted as *RCP*.
- **Methods' settings:**
  (a) *USCB* uses auction logs from Aug. 2022 for training and is re-trained at the end of each week.
  (b) *FC* uses the previous day's auction logs to generate references. The controller is PID [5], [14], [23].
  (c) *RCP* uses the previous day's auction logs for down-sampling. The sampling size $N = 10000$. The solver is MindOpt [24]. Auction logs from Aug. 2022 are used for training the auction scale prediction module.
- **Platform settings and others:**
  (a) Billing modules report the received benefits, costs and the current CPC every 15 minutes.
  (b) Considering commercial safety, each parameter update range is restricted to $\pm 80\%$ around their initial.
  (c) All methods use the same impression value evaluation module and others provided by the platform.

C. Evaluation metrics

We count the following metrics for each advertising campaign [1], [4], [5], [7], [8]. Unless otherwise specified, the results from *MACG* are employed as the baseline.

(a) D, the actual number of transactions.
(b) $\hat{G}$, the expected GMV. It is the maximized objective and measures each method's performance theoretically.
(c) G, the actual GMV. Due to the randomness of users' reactions in reality, we use G to measure the practical performance of each method.
(d) ROI, the quotient of the actual GMV and the consumed budget, which measures method profitability.
(e) B_r, the budget utilization ratio.
(f) PPC, the payment per click, related to CPC constraint.
(g) PPT, the payment per transaction. It measures methods' advertising capability and reveals the average advertising payment of a transaction.

D. Validation of budget bucketing system

To verify the fairness, we conduct the A/A experiment with *MACG* first. The results are shown in Table II. The two buckets have high similarity, proving the system is reliable for the rest A/B experiments. The results of MCB are similar to Table II, and we do not present them due to the limited space.

TABLE II
THE A/A VALIDATION EXPERIMENT RESULTS IN BCB.[1]

	D		$\hat{G}$		G		ROI		B_r		PPC		PPT	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
Bucket 1	1.0000	–	1.0000	–	1.0000	–	1.0000	–	1.0000	–	1.0000	–	1.0000	–
Bucket 2	0.9933	0.0152	1.0038	0.0096	0.9959	0.0219	0.9884	0.0121	1.0134	0.0172	0.9981	0.0051	0.9944	0.0085

[1] **Bucket 1** is the baseline, and all results are normalized by it. 'Avg.' means the average. 'Std.' means the standard deviation.

E. A/B experiment and discussion for BCB

The daily comparison is shown in Fig. 8, and Table III is the statistical results. In conclusion:

TABLE III
THE A/B EXPERIMENT RESULTS IN BCB.

	D		$\hat{G}$		G		ROI		B_r		PPC		PPT	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
MACG	1.0000	–	1.0000	–	1.0000	–	1.0000	–	**1.0000**	–	**1.0000**	–	1.0000	–
USCB	1.2461	0.2823	1.4752	0.2657	1.5123	0.3430	1.3143	0.2457	1.2697	**0.1240**	1.8267	0.1552	1.0120	0.1140
FC	1.1994	0.3147	1.4523	0.3079	1.4423	0.4052	1.2692	0.2708	1.2433	0.1332	1.8301	0.1808	1.0189	0.1509
RCP	**1.3249**	**0.2010**	**1.5025**	**0.2191**	**1.5521**	**0.2917**	**1.3617**	**0.1938**	1.2307	0.1404	1.7328	**0.1392**	**0.9081**	**0.0841**

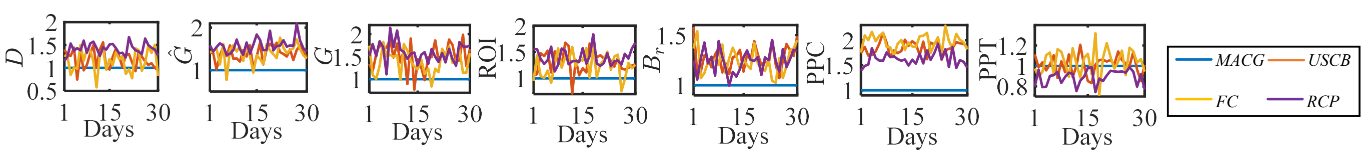


Fig. 8. Daily comparison of *MACG/USCB/FC/RCP* for each metric in BCB.

(a) *RCP* **outperforms the other methods.** As shown in Fig. 8, for daily metrics D, $\hat{G}$, G, and ROI, *USCB/FC/RCP* consistently surpass *MACG*, with *RCP* (the purple lines) demonstrating the most significant improvements. In terms of B_r, *USCB/FC/RCP* all spend more budget than *MACG*. Regarding PPC, while *USCB/FC/RCP* incurs higher costs per click than *MACG*, *RCP* is marginally lower than *USCB/FC*. Notably, for PPT, *RCP* significantly outperforms *USCB/FC/MACG*, suggesting superior advertising capability. In summary, as detailed in Table III, *RCP* leads in five out of seven metrics, with *MACG* claiming the remaining two. Compared to *MACG*, *RCP* increases D and ROI by over 30% and reduces PPT by nearly 10%. It also shows $5\% - 10\%$ improvements over *USCB* and *FC*.
(b) *RCP* **demonstrates greater stability.** In the daily comparison from Fig. 8 for metrics D, $\hat{G}$, G, and ROI, *RCP*'s performance is consistently superior to MACG, whereas the lines for *FC* and *USCB* occasionally dip below *MACG*. This suggests that *RCP* maintains a more reliable advertising performance and improvements compared to *USCB/FC*. Additionally, *RCP* exhibits less volatility in its curves than *USCB/FC*. Examining Std. of each metric in Table III, *RCP*

is smaller than the others, except for B_r. In fact, *USCB* tends to exhaust the budget because the budget consumption integral part of its input, resulting in the lowest Std. for B_r.

(c) ***RCP* is more efficient at securing high-valued impressions.** Interestingly, *MACG*, despite achieves the lowest B_r and PPC, significantly underperforms in other metrics, particularly ROI and PPT. This indicates *MACG*'s inefficiency in securing high-valued impressions, opting instead for lower-priced ones, which is unsatisfactory for advertisers. In contrast, other methods, while result in higher B_r and PPC, achieve substantially higher D and PPT. Notably, *RCP* stands out by not only reducing PPT by nearly 10% but also delivering superior performance in G, $\hat{G}$, D, and ROI, all with a more conservative budget consumption compared to *USCB* and *FC*. This shows that *RCP* makes every penny count.

F. A/B experiment and discussion for MCB

We firstly define the following two metrics to compare different methods' performance on the CPC constraint.

(a) R_{cpc}, the CPC constraint violation risk. It is the ratio of the number of advertising campaigns that violate CPC constraints to the total number of advertising campaigns.

(b) V_c, the quotient of the actual CPC to its constraint for an advertising campaign. It is only used for advertising campaigns that violate the constraint. This metric measures the robustness when facing unconscionable circumstances. Although the platform has recommended CPC constraints, some advertisers may still make unreasonable settings. Compared to the budget, the CPC constraint is 'soft' because it is an average over the day. Thus, for advertisers on *Taobao*, it is also acceptable if the actual CPC does not exceed 20% of the set value [3]–[5].

Limited by the space, we mainly presented the statistical results in Table IV and Table V. In conclusion:

TABLE IV
THE A/B EXPERIMENT RESULTS IN MCB.

	D		$\hat{G}$		G		ROI		B_r		PPC		PPT	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
MACG	1.0000	–	1.0000	–	1.0000	–	1.0000	–	**1.0000**	–	**1.0000**	–	1.0000	–
USCB	1.2962	0.2829	1.2923	0.3127	1.4098	0.3812	1.3532	0.2598	1.1010	0.1525	1.0529	0.0523	0.8012	0.1923
FC	1.2105	0.3229	1.1522	0.4031	1.2088	0.4532	1.3398	0.4931	1.0434	0.0812	1.0792	0.1009	0.9021	0.1692
RCP	**1.4635**	**0.2352**	**1.3733**	**0.1919**	**1.5126**	**0.2721**	**1.4944**	**0.1532**	1.0290	**0.0312**	1.0208	**0.0103**	**0.6916**	**0.1052**

TABLE V
THE RESULTS OF R_{cpc} AND V_c IN MCB.

	R_{cpc}	Max. of V_c	Max. of V_c	P($V_c > 1.2$)[1]
MACG	**0.00%**	–	–	–
USCB	21.21%	1.3723	1.1323	9.81%
FC	39.23%	1.5231	1.3106	19.25%
RCP	**13.29%**	**1.3123**	**1.0925**	**7.75%**

[1] It means the ratio of the number of advertising campaigns whose $V_c > 1.2$ to the total number of advertising campaigns.

(a) As with BCB, *RCP* is superior as shown in Table IV; further details are omitted for brevity.

(b) **Table V reveals that *USCB*, *FC*, and *RCP* might occasionally violate the CPC constraint, whereas *MACG* does not.** There are three reasons: Firstly, the bid optimization problem (1) models the CPC constraint in the expected sense. Secondly, for MCB, *MACG* employs's a 'hard' upper bound to limit its bids such that it never breaks the CPC constraint. It is practical but can not hold the optimality. By contrast, *USCB*, *FC*, and *RCP* does not, offering a broader bidding range, especially for high-valued impressions, which leads to better performance. Lastly, setting CPC constraints requires agile market acumen and a strategic balance of costs and benefits. Poorly chosen CPC constraint can compromise advertising campaigns' effectiveness and increase the risk of violations. This highlights the need for better CPC constraint setting guidance.

(c) **Compared to *USCB* and *FC*, *RCP* has the lowest violation risk.** Despite a 15% violation rate, *RCP* shows lower Max. and Avg. of V_c. Besides, *RCP* achieves the lowest likelihood of $V_c > 1.2$, indicating greater robustness in extreme scenarios.

G. Ablation experiment

As mentioned in Fig.2 and Algorithm 2, in *RCP*, online receding optimization and auction scale prediction collaboratively refine the bidding formulation. To evaluate their impact, we performed an ablation study:

(a) *RCP*: our complete method incorporating both processes.

(b) *SS*: A static version of *RCP*, lacking online receding optimization and auction scale prediction.

(c) *RCW*: *RCP* without auction scale prediction, retaining online receding optimization.

Table VI presents the finds. Collectively, all *RCP* variants outperform *MACG*. *RCW* shows a marked enhancement over *SS* in nearly all metrics, averaging $10\% - 20\%$ improvements, validating the efficacy of online receding optimization and aligning with the non-decreasing property of OLFC that we used. The gap between *RCP* and *RCW* is less pronounced, yet *RCP* demonstrates superior stability across metrics. It shows that introducing auction scale prediction mainly benefits stability when facing market Volatility.

TABLE VI
THE ABLATION EXPERIMENT RESULTS IN BCB.

	D		$\hat{G}$		G		ROI		B_r		PPC		PPT	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
MACG	1.0000	–	1.0000	–	1.0000	–	1.0000	–	**1.0000**	–	**1.0000**	–	1.0000	–
SS	1.1794	0.3521	1.3207	0.3122	1.3763	0.3921	1.2233	0.2531	1.2168	0.1523	1.7587	0.2332	0.9465	0.1923
RCW	1.2609	0.2813	1.4809	0.2533	1.5322	0.3218	1.3523	0.2191	1.3145	0.1498	1.8562	0.2033	0.9270	0.1822
RCP	**1.3027**	**0.1821**	**1.4833**	**0.1899**	**1.5423**	**0.2839**	**1.3881**	**0.1678**	1.2502	**0.1112**	1.7425	**0.1588**	**0.9101**	**0.0997**

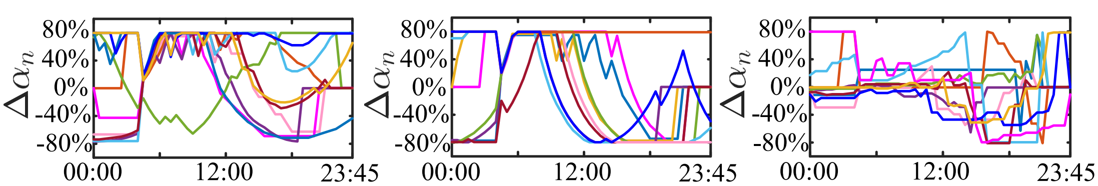
H. A case study of adaptability and personalization

Since different advertising campaigns promote varied items, the impressions they receive on the same day may vary. Moreover, even for the same advertisement, it is unrealistic to assume consistent impressions across days [13]. An ideal strategy should be personalized and adapted to both advertisers' objectives and market shifts. Thanks to (9), objectives are already integrated into the optimization problem. Therefore, we focus on analyzing the regulation process of each method to assess its adaptability and personalization, particularly for *USCB/FC/RCP*, which all use (9). We hope to reveal the fundamental difference between our approach and the previous works. For clarity, we use BCB as an example and provide a case study (MCB has coupled parameters in (9), making it complicated to present). Define

$$\Delta\alpha_n = \alpha_n/\bar{\alpha} - 1,\ n = 1, \ldots, 96, \tag{25}$$

where α_n is the n-th update α in (9), $\bar{\alpha}$ is the initial value of each method. We can graph the variation of $\Delta\alpha_n$ over time to visualize the parameter regulation process of each method. These curves would naturally differ among advertising campaigns and even for the same advertising campaign across days, reflecting personalized requirements and market changes. To quantify the similarity of two curves, we can calculate their Pearson correlation coefficient ρ. The closer ρ is to 0, the more similar the two curves are.

We randomly selected ten advertising campaigns from various categories (clothing, electronics, etc.) and plotted their $\Delta\alpha_n$ on 09/22 and 09/23. To eliminate the effect of different α_0 on $\Delta\alpha_n$, for the same advertising campaign, we set the same initial value a_0 for *USCB*, *FC*, and *RCP*. The results are depicted in Fig. 9 and Fig. 10. In summary,



(a) *USCB* (b) *FC* (c) *RCP*

Fig. 9. The parameter regulation processes for the bidding formulation of *USCB/FC/RCP* on 09/22. Each color line represents an advertising campaign.

(a) Fig. 9 depicts the regulation processes of ten advertising campaigns on 09/22, the average ρ of *USCB/FC/RCP* are 0.1828/0.3631/0.0293, respectively. Interestingly, most of *USCB*/*FC*'s curves show similar patterns, with an initial increase followed by a decline post-noon. This trend is also observed on other days for both *USCB* and *FC*, which does not repeat here. In contrast, *RCP* does not.

(b) Fig. 10 examines the regulation processes for the same advertising campaign on different dates, the average ρ of *USCB/FC/RCP* are 0.2453/0.6143/0.1003, respectively. *USCB/FC* maintain a high similarity in their curves, often following the trend observed in (a). In contrast, *RCP*' curves demonstrate relatively less similarity.

We preliminarily speculate that *USCB* and *FC* may overemphasize generalization, potentially at the expense of personalization and adaptability. In contrast, *RCP* solves a unique optimization problem of each advertising campaign and updates based on real-time feedback. However, determining which is better is not trivial, as the analysis of impressions and the market is intricate and they may be interdependent. Fortunately, with all auction logs being accessible, we can retrospectively solve the optimal α in (9) of each update in hindsight, denoted as α^h. Subsequently, for each advertising campaign, we get the *optimal*

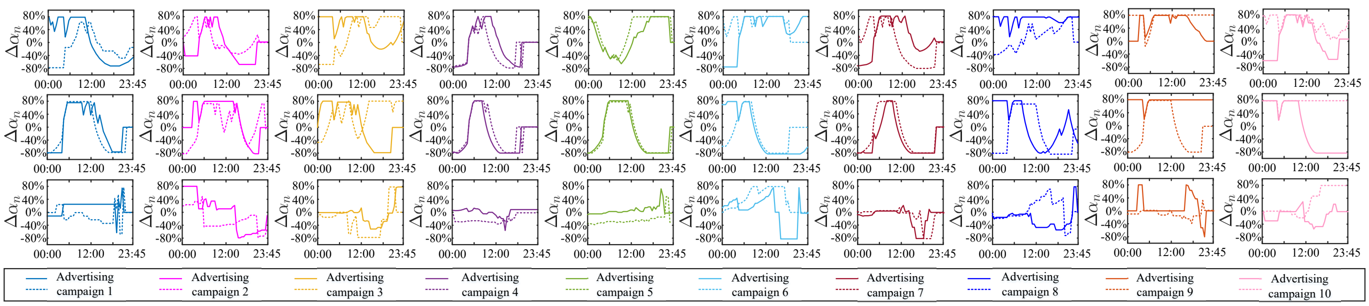


Fig. 10. Regulation processes of *USCB/FC/RCP* on 09/22 and 09/23. The first row comes from *USCB* , the second row comes from *FC* , the third row comes from *RCP*. The solid lines represent the results from 09/22, while the dashed lines represent the results from 09/23.

update sequence $\{\alpha_n^h \mid n = 1, \ldots, 96\}$ and calculate ρ. Firstly, for all advertising campaigns in the BCB A/B experiment, the average ρ of their optimal update sequences is 0.0072 on the same day, while the average ρ is 0.0831 on distinct dates. This indicates that each advertising campaign requires a personalized strategy every day to adapt to market Volatility. Then, for each advertising campaigns in the BCB A/B experiment, we compare its optimal update sequence with each method's update sequence $\{\alpha_n \mid n = 1, \ldots, 96\}$ to calculate ρ. The average ρ of *USCB/FC/RCP* are 0.3133/0.1523/0.6589, respectively. These suggests that *RCP* achieves the higher level of personalization and adaptability.

In practice, we also noted an intriguing phenomenon: the regulation processes of *USCB/FC* mainly align with the whole platform trend, where the quality and quantity of total impressions are at their peak during midday and evening hours. While this may be suitable for most advertising campaigns, many high-value impressions may still missed. This shortcoming may be due to the curse-of-dimension, as *USCB/FC* cannot use too many individual features to characterize each advertising campaign. Conversely, *RCP* makes the benefit of auction logs, implicitly considering each advertising campaign's characteristics. Due to the complexity of online advertising, no existing methods can measure a bidding strategy comprehensively. This case study provides a perspective to highlight a potential issue in previous works as well as the delicate balance that must be struck between generalization and personalization.

I. Computation consumption

Finally, for the practical usage, computation comparison also is desired. We divide it into two parts: *offline preparation* and *online bidding*. Offline preparation is essential for all methods, involving strategy network training for *MACG/USCB*, parameters searching for *FC*, and prediction network training for *RCP*.

The results are shown in Table VII. To sum up: *FC* is the most resource-saving, as it treats the problem as a pre-set reference tracking issue. The majority of its computational resources are used for generating references and parameters searching. *MACG* and *USCB* have similar consumption levels, primarily due to network training. *RCP* has lower offline preparation consumption,as its prediction module is significantly simpler, but it requires additional online computational resources to solve (9). Nevertheless, with AliCloud support, these added computation is cost-effective and does not significantly impact practicality. Besides, *RCP*'s inexpensive and quickly updatable offline preparation may be key to enhancing performance and stability.

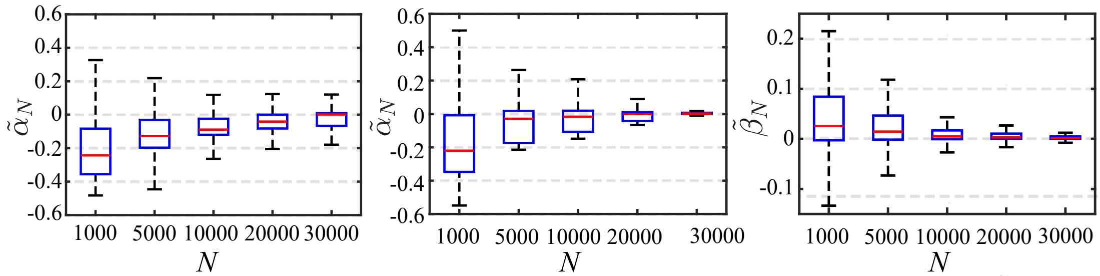
TABLE VII
THE COMPARISON OF COMPUTATION CONSUMPTION.

	Offline preparation		Online bidding	
	Time[1]	CPU / Memory[1]	Time[2]	CPU / Memory[2]
MACG	$\sim$ 3 hours	1.0000 / 1.0000	$\leq 10ms$	1.0000 / 1.0000
USCB	$\sim$ 3 hours	0.9513 / 1.0328	$\leq 10ms$	0.9896 / 0.9551
FC	$\sim$ 20 mins	0.1812 / 0.2691	$\leq 10ms$	0.0268 / 0.0312
RCP	$\sim$ 30 mins	0.3315 / 0.4198	$\leq 10ms$	1.0553 / 1.0674

[1] The total time, total CPU/memory consumption for all campaigns.
[2] The maximal computation time, average CPU/memory consumption for an advertising campaign during an auction.

VI. Limitation and Future Work

It is crucial to note that RCP requires approximately 5% more online computational resources than the state-of-the-art, which, despite being cost-effective, could strain infrastructures, particularly for smaller platforms. Additionally, *RCP* highly relies on the auction logs, which would occur the cold-start problem for new established advertising campaigns. Future improvements include:



(a) BCB, $\tilde{\alpha}_N$ (b) MCB, $\tilde{\alpha}_N$ (c) MCB, $\tilde{\beta}_N$

Fig. 11. Boxplots of the relative errors with different N. (a) is $\tilde{\alpha}_N$ of BCB, (b) is $\tilde{\alpha}_N$ of MCB, (c) is $\tilde{\beta}_N$ of MCB.

(a) Reducing online computation. Essentially, the online computation mainly involves solving a dual linear programming problem, and the latest methods still exceed $O(n^{2.5})$ complexity. Nonetheless, for BCB, we have recently reformulated it as a sorting problem in the latest work [25], which reduces the complexity to $O(nlogn)$. This significantly saves online resources, and we are actively seeking similar solutions for MCB.
(b) Enhancing the prediction module, possibly through better temporal data analysis and leveraging new models.
(c) The cold-start is inevitable for newly established advertising campaigns. It is essential to explore methods to reduce reliance on auction logs and shorten the time for cold-starting.

Additionally, *RCP* is not only utilized in e-commerce but also has the potential to be extended to similar real-time bidding/competitive scenarios, such as optimizing transactions in the microgrid/smartgrid electricity market [26], influence maximization in social networks [27], crowdsourcing [28], etc.

VII. Conclusion

This paper proposes a novel bid optimization method for solving BCB/MCB in online advertising, including receding optimization, bidding formulation, downsampling, and auction scale prediction. We begin with a theoretical analysis of the optimal bidding strategy and develop an online receding optimization process that updates the bid formulation via linear programming. We propose an efficient down-sampling scheme to facilitate large-scale practical application. Additionally, auction scale prediction is integrated to respond to market volatility. A month-long A/B experiment on *Taobao* confirms our method's superiority. Compared to the industrial methods, we improve marketing objectives by 50%, profitability by 30%, and reduces cost per transaction by 10%, while still achieving $5\% - 10\%$ improvements over the state-of-the-art.

Appendix A
Empirical experiments for sampling size

We randomly select 10000 advertising campaigns from *Taobao*. The sampling size is denoted as N, and we set six groups $N = \{1000, 5000, 10000, 20000, 30000, +\infty\}$, where $+\infty$ means we use all records of an advertising campaign. Each advertising campaign is sampled and solved 100 times for each N. **For BCB,** it maximizes the expected Gross Merchandise Volume (eGMV) subject to budget constraint. **For MCB,** it maximizes eGMV subject to budget and cost-per-click (CPC) constraints. Evaluation metrics are:

(a) **The relative errors** of α and β, we define $\tilde{\alpha}_N = \alpha_N/\alpha_{+\infty} - 1$ and $\tilde{b}_N = b_N/b_{+\infty} - 1$, where α_N and β_N denote solutions come from different N. Here, the results solved by the whole records are regarded as the truth, that is, $\alpha_{+\infty}$ and $\beta_{+\infty}$.
(b) **Average resource consumption**, including the solving time T_N, CPU usage C_N, and memory consumption M_N. For clarity, we use average results from $N = 1000$ of each advertising campaign as itself baseline and define

$$\tilde{T}_N = T_N/\bar{T}_{1000},\ \tilde{C}_N = C_N/\bar{C}_{1000},\ \tilde{M}_N = M_N/\bar{M}_{1000}.$$

The solver is MindOpt [24] and the results are shown in Fig. 11 and Table VIII.

(a) The accuracy and stability of solutions are related to N. Increasing N can improve its accuracy and stability but exponentially increases computational resources.
(b) Solution improvements are not infinite as increasing N. When $N > 10000$, improvements are less significant, but the required resources are overgrowing.

In practice, by balancing accuracy and consumption, we choose $N = 10000$ in our work. Meanwhile, *the solving time of an advertising campaign is 10 ms at most* in our practical usage. For the whole platform, the sampled log is only 3% of the entire logs, which is generally preferred.

TABLE VIII
STATISTICS OF THE AVERAGE RESOURCE CONSUMPTION.

N	1000	5000	10000	20000	30000	$+\infty$
$\tilde{T}_N$	1.0000	1.6214	4.7123	13.2363	29.5136	591.3262
$\tilde{C}_N$	1.0000	1.4236	3.2319	8.9000	19.1215	377.4231
$\tilde{M}_N$	1.0000	1.3824	3.7981	10.1252	21.9363	633.7121

Besides, We have empirical examples for visually proving the sub datasets obtained from random sampling and auction logs are identical distributed. For example, given a BCB advertising campaign with 122361 auction records in its auction log, its objective is to maximize the total number of transactions, the constraint is budget. Its parameters are solved the dual form of the following problem,

$$\arg\max_{z_1,\ldots,z_{\hat{T}}} \sum_{t=1}^{\hat{T}} \hat{v}_t z_t \qquad \text{s.t.} \sum_{t=1}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \leq B,$$

where random variables that effect the solution are $\hat{v}$, $\hat{w}$, and $\hat{p}^{\mathrm{CTR}}$. Then for this advertising campaign, we randomly select 100%, 50%, and 10% of data in its record log and plot their distributions in Fig. 12.

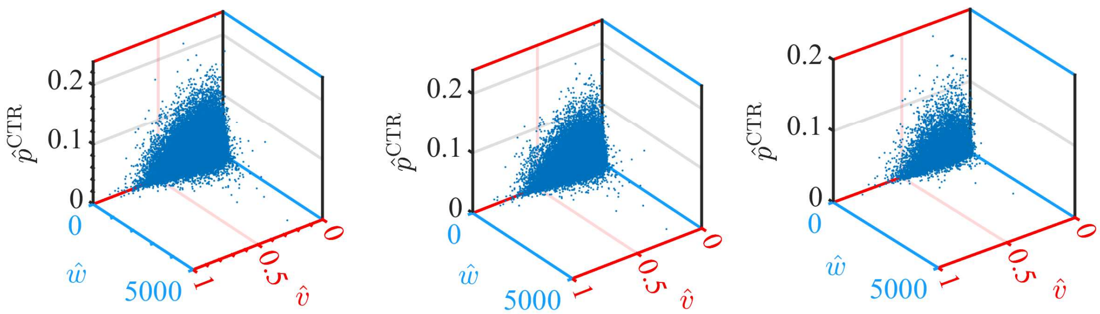


Fig. 12. The distributions with different sampling sizes. The left subfigure use 100% of data, the middle uses 50%, the right uses 10%. The red axis represents $\hat{v}$, the blue axis represents $\hat{w}$, and the black axis represents $\hat{p}^{\mathrm{CTR}}$.

In Fig. 12, three distributions seems to be similar. For further comparison, we plot their marginal distributions in Fig. 13, Fig. 14, and Fig. 15. Thus, we can say that the sub datasets obtained from random sampling and auction logs are identically distributed.

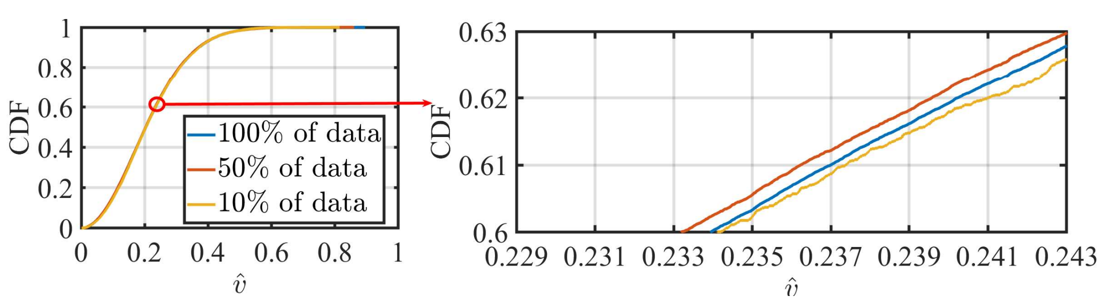


Fig. 13. The cumulative density function (CDF) of $\hat{v}$ with different sampling sizes. The left subfigure provides an overview, while the right offers a zoomed-in perspective around $\hat{v} = 0.237$.

APPENDIX B
DETAILED SETTINGS AND VERIFICATION OF THE AUCTION SCALE PREDICTION

The configuration details are as follows:

- CNN: Kernel size of 3×3.
- LSTM: 256 hidden variables per LSTM, determined through grid searching.
- Soft attention: Utilizes one neuron for attention pooling, with *Softmax* as the activation function.
- MLP: Two layers with 256 neurons each, optimized through grid searching. *tanh* is the activation function.
- Training: *Adam* with cross-entropy loss function.

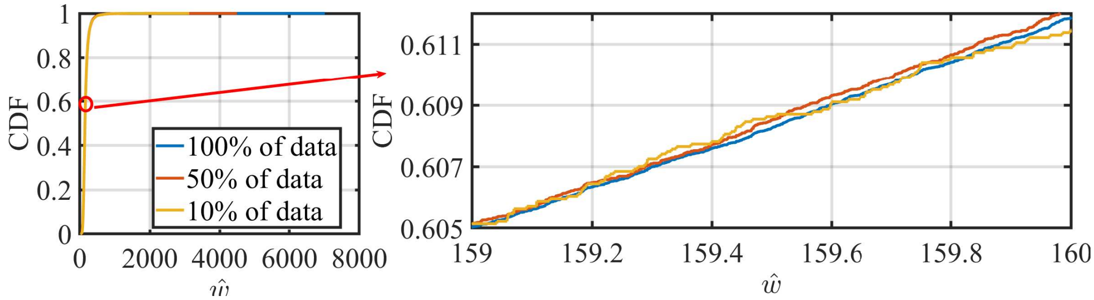


Fig. 14. The cumulative density function (CDF) of $\hat{w}$ with different sampling sizes. The left subfigure provides an overview, while the right offers a zoomed-in perspective around $\hat{w} = 159.4$.

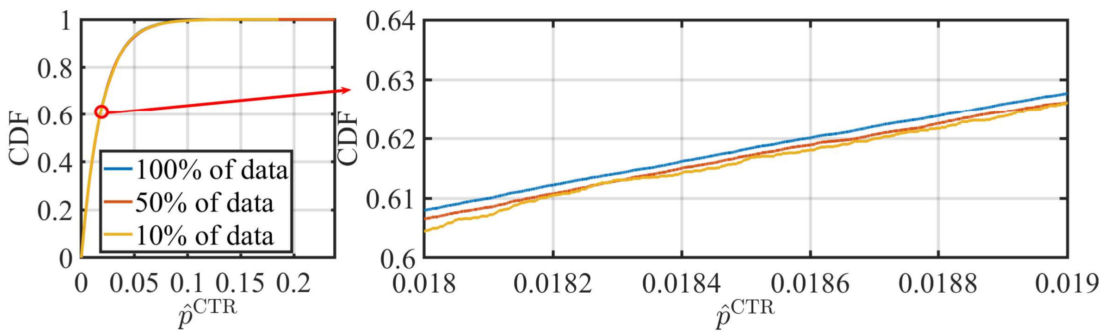


Fig. 15. The cumulative density function (CDF) of $\hat{p}^{\mathrm{CTR}}$ with different sampling sizes. The left subfigure provides an overview, while the right offers a zoomed-in perspective around $\hat{p}^{\mathrm{CTR}} = 0.0185$.

- Features: Impression features include the number of impressions, clicks, etc. Time features consist of current time, residual time, and current time interval.

We use the auction logs from 05/11/2022-05/17/2022 provided by *Taobao* for training. The module predicts the auction scales of each campaign for each interval from 05/18/2022-05/31/2022. To account for the diverse auction scales across various advertising campaigns, we classify them into six distinct groups for comparative analysis: $[0, 1)$, $[1, 10)$, $[10, 20)$, $[20, 50)$, $[50, 100)$, and $[100, +\infty)$. For example, the range $[0, 1)$ denotes advertising campaigns with an average auction scale falling within the interval $[0, 1)$. Comparison metrics are

- **Performance of** 10%, P_{10}. It represents the proportion of samples where the absolute error of the predicted is less than 10% [11], [29]. It reflects the stability and confidence coefficient of the prediction; the higher it is, the better.
- **Weighted Mean Absolute Percentage Error**, W. It reflects the anti-interference ability and convergence of the prediction; the lower it is, the better.

The results are shown in Fig. 16. For advertising campaigns with fewer auction scales, predictions are challenging, as indicated by a lower Avg. of P_{10} and a higher Avg. of W. As the auction scale increases, P_{10} predominantly exhibits an upward trend, whereas W decreases. Despite the periodicity between two adjacent days, the results for 'Yesterday' do not meet expectations, indicating the need for prediction techniques. Compared to previous work, our method outperforms in all aspects except for P_{10} in the group $[1, 10)$. Specifically, when compared to DA-RNN, which has the second-best performance, our method improves P_{10} by around 7% and reduces W by about 5%.

References

[1] W. Ou, B. Chen, X. Dai, W. Zhang, W. Liu, and Y. Yu, "A survey on bid optimization in real-time bidding display advertising," *ACM Trans. Knowl. Discov. Data*, 2023.

[2] J. P. Johnson, T. Jungbauer, and M. Preuss, "Online advertising, data sharing, and consumer control," *Management Science*, 2024.

[3] H. Zhu, J. Jin, C. Tan, F. Pan, Y. Zeng, H. Li, and K. Gai, "Optimized cost per click in taobao display advertising," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 2191–2200.

[4] Y. He, X. Chen, D. Wu, J. Pan, Q. Tan, C. Yu, J. Xu, and X. Zhu, "A unified solution to constrained bidding in online display advertising," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021, pp. 2993–3001.

[5] X. Yang, Y. Li, H. Wang, D. Wu, Q. Tan, J. Xu, and K. Gai, "Bid optimization by multivariable control in display advertising," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 1966–1974.

[6] H. Wang, C. Du, P. Fang, L. He, L. Wang, and B. Zheng, "Adversarial constrained bidding via minimax regret optimization with causality-aware reinforcement learning," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023.

[7] Google.com, "Google ads help," 2022. [Online]. Available: https://support.google.com/google-ads

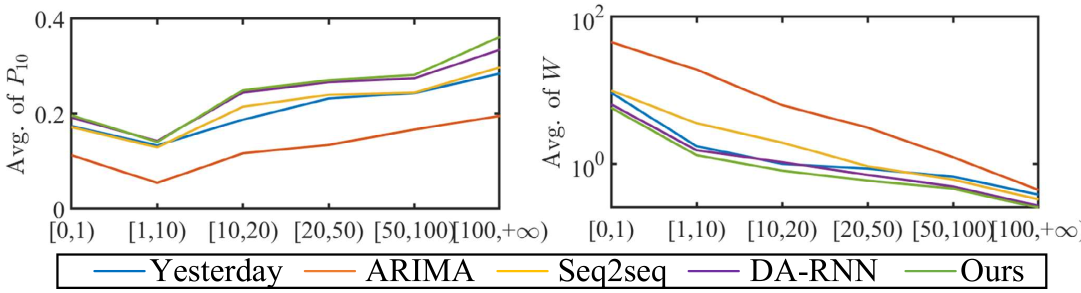


Fig. 16. Average (Avg.) of P_{10} and W in 05/18/2022-05/31/2022. 'Yesterday' means the previous day's auction scale is directly used as the prediction. 'ARIMA' is [30], 'Seq2seq' is [31], 'DA-RNN' is [32].

[8] W. Zhang and J. Xu, "Learning, prediction and optimisation in rtb display advertising," in *Proceedings of the 25th Information and Knowledge Management*, 2014, pp. 1077–1086.
[9] W. Zhang, S. Yuan, and J. Wang, "Optimal real-time bidding for display advertising," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1077–1086.
[10] J. Xu, K.-c. Lee, W. Li, H. Qi, and Q. Lu, "Smart pacing for effective online ad campaign optimization," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 2217–2226.
[11] N. Grislain, N. Perrin, and A. Thabault, "Recurrent neural networks for stochastic control in real-time bidding," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2801–2809.
[12] J. Jin, C. Song, H. Li, K. Gai, J. Wang, and W. Zhang, "Real-time bidding with multi-agent reinforcement learning in display advertising," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. New York, NY, USA: Association for Computing Machinery, 2018, p. 2193–2201.
[13] S. Yuan, J. Wang, and X. Zhao, "Real-time bidding for online advertising: measurement and analysis," in *Proceedings of the seventh international workshop on data mining for online advertising*, 2013, pp. 1–8.
[14] N. Karlsson, "Feedback control in programmatic advertising: The frontier of optimization in real-time bidding," *IEEE Control Systems Magazine*, vol. 40, no. 5, pp. 40–77, 2020.
[15] T. Nedelec, N. El Karoui, and V. Perchet, "Learning to bid in revenue-maximizing auctions," in *International Conference on Machine Learning*. PMLR, 2019, pp. 4781–4789.
[16] S. Zhai, K.-h. Chang, R. Zhang, and Z. Zhang, "Attention based recurrent neural networks for online advertising," in *Proceedings of the 25th International Conference Companion on World Wide Web*, 2016, pp. 141–142.
[17] Z. Guan, H. Wu, Q. Cao, H. Liu, W. Zhao, S. Li, C. Xu, G. Qiu, J. Xu, and B. Zheng, "Multi-agent cooperative bidding games for multi-objective optimization in e-commercial sponsored search," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021, pp. 2899–2909.
[18] K.-H. Li, "Reservoir-sampling algorithms of time complexity o (n (1+ log (n/n)))," *ACM Transactions on Mathematical Software (TOMS)*, vol. 20, no. 4, pp. 481–493, 1994.
[19] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.
[20] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.
[21] P. Roy, R. Laprise, and P. Gachon, "Sampling errors of quantile estimations from finite samples of data," *arXiv preprint arXiv:1610.03458*, 2016.
[22] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.
[23] W. Zhang, Y. Rong, J. Wang, T. Zhu, and X. Wang, "Feedback control of real-time display advertising," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, 2016, pp. 407–416.
[24] A. D. Academy, "Mindopt optimization suite." [Online]. Available: https://solver.damo.alibaba.com/htmlpages/page#/en
[25] K. Fang, J. Wu, H. Liu, and Q. Cao, "Efficient bid optimization method for budget constraint bidding in online advertising," in *The 42nd Chinese Control Conference (CCC)*, 2024.
[26] G. Liu, Y. Xu, and K. Tomsovic, "Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization," *IEEE Transactions on Smart Grid*, pp. 227–237, 2015.
[27] C. Zhang, J. Zhou, J. Wang, J. Fan, and Y. Shi, "Fairness-aware competitive bidding influence maximization in social networks," *IEEE Transactions on Computational Social Systems*, pp. 2147–2159, 2024.
[28] Y. Li, Y. Li, Y. Peng, X. Fu, J. Xu, and M. Xu, "Auction-based crowdsourced first and last mile logistics," *IEEE Transactions on Mobile Computing*, pp. 180–193, 2024.
[29] A. Nath, S. Mukherjee, P. Jain, N. Goyal, and S. Laxman, "Ad impression forecasting for sponsored search," in *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 943–952.
[30] K.-c. Lee, B. Orten, A. Dasdan, and W. Li, "Estimating conversion rate in display advertising from past erformance data," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012, p. 768–776.
[31] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 27, 2014.
[32] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, and G. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," *arXiv preprint arXiv:1704.02971*, 2017.