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 is crucial in improving advertising performance. Recently, the bid optimization problem has converged to two 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 paper proposes an online receding control methodology to solve million-scaled bid optimization problems in practice. Firstly, we theoretically analyze the structure of the optimal bidding strategy. Further, an iterative 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 million-scaled linear programming problems, we propose a sampling-based efficient solution. Besides, a Recurrent-Neural-Network-based auction scale prediction is used to adapt to the changing market. Finally, a series of carefully designed 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, 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]–[6].

In e-commercial websites like *Taobao*, online advertising is used to facilitate purchases [4], [7]. Fig. 1 illustrates that whenever an online user views and triggers an **impression**, that is, an opportunity to display advertising (AD) in front of users. On the backend of websites, dozens of interested advertisers are recalled to compete in an auction with individual bids, and all of them hope to win the impression to display their ADs. Then, an auction agent uses the Second Price Auction (SPA) to determine the winning ads, in which the highest bid wins and would pay for the second-highest bid [8]. In e-commercial advertising, charging upon click is commonly adopted, i.e., the winning AD will be **displayed for free and charged once clicked** [9]–[11]. Finally, the user sees the winning AD and reacts, e.g., clicking through, purchasing, etc. Typically, the time from triggering an impression to showing the winning AD is within 0.1 seconds to avoid long page load times.

On the other hand, advertisers usually set a campaign for each AD, including the displayed AD, the target groups, **marketing objectives** (e.g., Gross Merchandise Volume (GMV)), **commercial constraints**, etc. [4], [12]. Constraints contain the budget

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.

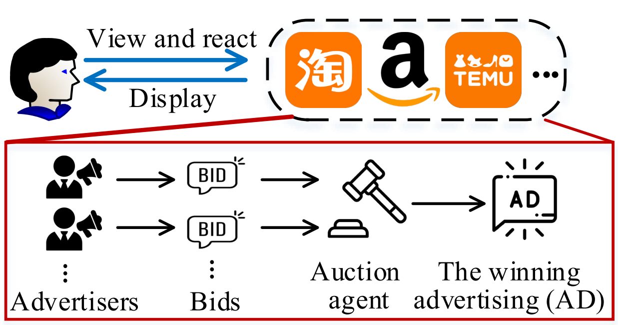

Fig. 1. Brief process of online advertising.

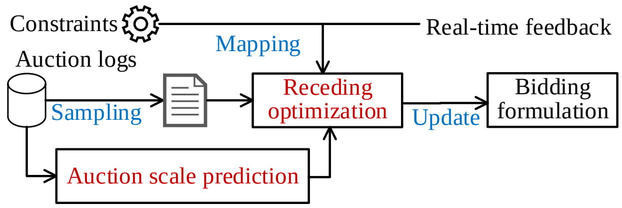

Fig. 2. The proposed methodology.

and cost per action (CPA), e.g., cost per click (CPC). Fig. 1 reveals that only the winning AD can be displayed, implying advertisers must bid carefully for each auction to achieve objectives. Thus, designing an efficient bidding strategy, also called the **bid optimization**, is crucial in achieving better advertising performance [4], [5], [12], [13].

Bid optimization has been thoroughly investigated both in research and industry [5], [10], [13]–[21]. Recently, the problem formulation has converged to **BCB** (Budget-Constrained Bidding, it only has the budget constraint) and **MCB** (Multi-Constraints Bidding, it has the budget and CPA constraints), where BCB is a special case of MCB [10], [16]. Solving BCB/MCB mainly relies on an efficient bidding strategy learned from a bunch of historical auction logs. In *Taobao*, a typical auction log corresponds to an *auction quality tuple*:

$$(timestamp, second\_price, pctr, pcvr, item\_price), \tag{1}$$

which represents the occurrence time of the auction, the second-highest price, the predicted click-through rate, the predicted conversion rate, and the item price, respectively. The atom log can be used to calculate various impression-level values based on marketing objectives, e.g., GMV [6], [13].

Given a bunch of auction logs, the problem is first formulated as a combinatorial optimization problem [13], then relaxed to linear programming [16]. Most existing works try to solve it via mixed-constructed strategies or learning methods [4], [10], [12], [13], [16], [18]. But the problem is still challenging, and the previous works have the following deficiencies worthy of improvement:

- **Large-scale and real-time.** An AD usually participates in millions of relevant auctions each day, implying the corresponding bid optimization may have millions of variables. Solving such a problem seems exhausting, even in a distributed fashion [22]. Nevertheless, e-commerce sites like *Taobao* have millions of such complex problems to solve daily, making it hard to reach an optimal solution for each AD. [7], [10], [16], [23].
- **Curse-of-dimension**. Most existing works, like learning-based and control-based, try to alleviate the above challenge by characterizing the bid optimization problem with a few normalized features, e.g., remnant budget ratio, to train a strategy network (or design a controller) and generalize to all ADs [13], [15]–[17], [24]–[26]. Such generalization seems promising but usually embeds very few features due to the poor convergence of the training process, resulting in unsatisfactory performance.
- **Personalization and adaptability.** In practice, these strategies often yield similar and non-personalized regulation processes. Furthermore, existing works directly optimize or train using the stationary auction log and do not take advantage of today's auction information for strategy adjustment [10], [16], [17], [20], [23]. They might be fragile when the market changes [5], [27].

A. *Preview of our work*

This paper proposes an online receding control methodology to tackle BCB/MCB problems. The methodology is depicted in Fig. 2. Firstly, a receding optimization is scheduled periodically to solve the remaining horizon optimization in the regulation process, which updates a constructed optimal bidding formulation to tackle the **curse-of-dimension**. It is a straightforward way to solve the problem and maintain **personalization**. To tackle **large-scale and real-time**, the optimization process only uses a small part of randomly sampled auction logs rather than the entire logs or fitted models in previous works. Considering the **adaptability**, the optimization process is scheduled periodically based on real-time constraints feedback and the sampled auction logs. Besides, we build an auction scale prediction module to predict the future of the remaining horizon, given the timely occurring auction scales used to modify for the changing market.

The major contributions are summarized as follows:

(a) We propose a self-contained online receding methodology to BCB/MCB. It has been tested and implemented on a real e-commerce platform for over a year.
(b) To tackle the **curse-of-dimension**, we theoretically analyze the structure of the optimal bidding strategy first. Further, using the open-loop feedback control, an iterative approach is designed, making the solved bidding strategy performance non-decreasing after each update.

(c) For the **personalization**, we construct an optimal bid formulation that tries to solve the unique problem of each AD separately. Meanwhile, considering **large-scale** and **real-time**, a sampled-based efficient solution is designed. Empirical experiments show it uses 3% data of the whole platform and achieves more than 90% accuracy. Equipped with an auction scale prediction module, the proposed approach elegantly tackles the **adaptability**.
(d) **A month-long scientific online A/B experiment implemented on Taobao** demonstrates that our work outperforms state-of-the-art and industrial methods. Compared to the online baseline, our method improves marketing objective by 50% and profitability by 30% and decreases cost per transaction by 10%. Compared to state-of-the-art, we still have $5\% \sim 10\%$ uplifts for each metric.

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

II. Related works

In typical online advertising, AD exchange sells impressions via auction. The Supply Side Platform (SSP) integrates the media resource, and the Demand Side Platform (DSP) works as the agency to help advertisers gain impressions. Nevertheless, an e-commercial platform like Taobao incorporates the roles of media, SSP, and DSP due to its scale and proprietary data. Thus, bid optimization can be mainly divided into **social welfare-aware** and **selfish bid optimization**.

A. Social welfare-aware bid optimization

On platforms like Taobao, advertisers compete in their self-contained media. Thanks to its proprietary data and computing resources, the platform even works as a bidding agency for most competing advertisers. To this end, Zhu *et al.* [7] proposed the *Optimized Cost Per Click* bidding algorithm to optimize the global return of an impression and respect the bid limit constraint of the advertisers. Jin *et al.* [17] 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.* [23] tried to model the problem in a *Multi-Agent Cooperation Games* manner but explicitly included the platform revenue as a constraint to strike a balance between the rewards of different advertisers and platform revenue. For super platforms like Taobao, social welfare seems more attractive and directly optimizing social welfare, which apparently achieves a higher global return. However, inappropriate assumptions about advertiser demands usually exist in constructing social welfare. Furthermore, advertisers are not motivated to cooperate with their competitors, therefore introducing the risk of using individual bid authorization for the common good.

B. Selfish bid optimization

Selfish advertisers are more reasonable in practice, which is a common setting in DSP bidding due to its profit form. 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 [14], [28], [29]. Apparently, rigorous bid ranges might miss high-quality but expensive impressions. Nowadays, BCB and MCB have become the mainstream formulation of bid optimization, in which advertisers do not restrict bid limits but only lays a daily budget constraint [13]–[15] and/or CPA constraint [10], [23], [30]. Under SPA, the optimal bidding formula is proved by functional analysis [13] and Lagrangian dual analysis [16]. After that, solving BCB/MCB reduces to finding the optimal biding parameters.

Given a bunch of auction logs, parameters are firstly solved via linear programming solvers [13], [16]. 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. Control strategies are the first attempt to generalize the bidding strategy. It makes the current states, e.g., the consumed budget, follow the referred by adjusting the bidding parameters according to tracking errors. Zhang *et al.* [13] and Karlsson [31] proposed error feedback for BCB, and Yang *et al.* [16] provided a double PID-based method for MCB.

Reinforcement learning (RL) and machine learning (ML) methods have recently been widely adopted due to their power and elasticity in solving complex decision-making problems. For BCB, Cai *et al.* [24], Nedelec *et al.* [25], Zhai *et al.* [32], Grislain *et al.* [15] investigated different types of networks. Directly applying RL to MCB is not straightforward and requires careful reward shaping for CPA constraints [10]. Although applicable, the optimality might not hold since CPA constraints lay an overall constraint of the whole auction process and can not be easily accumulated like the objective value. By incorporating more features in training, RL/ML is a natural extension to PID methods that use only cost speed. However, to deal with all problems and ADs, RL/ML can only embed very few features due to poor convergence and the trade-off between generalization and personalization.

Compared with social welfare-aware methods, self-oriented bid optimization guarantees procedural justice. Bid optimization only answers for better advertising performance. In this way, auction mechanism [8] can be justified with the assumption that each participant will try to achieve their own optimal, which decouples the platform revenue and advertiser performance elegantly. **The method we proposed belongs to the selfish version.**

TABLE I
NOTATIONS

Notation	Description
$\mathbf{x}$	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.

III. BID OPTIMIZATION FORMULATION

We formally model the bid optimization problem, considering both BCB and MCB. Notations refer to Table I. For the bid optimization problem, each auction is characterized by an embedding vector $\mathbf{x}$ that is usually equivalent to (1). Assume there are T auctions and under SPA, the problem is to find a bidding strategy π such that:

$$\begin{aligned} &\max_{\pi_t} \sum_{t=1}^{T} \mathbb{E}\left[v_t \mathbf{1}_t(w_t, \pi_t)\right] \\ &\text{s.t.} \quad \sum_{t=1}^{T} c_t \mathbf{1}_t(w_t, \pi_t) \leq B, \ \frac{\sum_{t=1}^{T} \mathbb{E}\left[c_t \mathbf{1}_t(w_t, \pi_t)\right]}{\sum_{t=1}^{T} \mathbb{E}\left[a_t \mathbf{1}_t(w_t, \pi_t)\right]} \leq C, \end{aligned} \tag{2}$$

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 $b_t \geq w_t$, $\mathbf{1}_t(w_t, \pi_t) = 1$. Otherwise, $\mathbf{1}_t(w_t, \pi_t) = 0$. Both v and a have varied implementations in practice [6], [12]. The objective is to maximize the total expected v. The first constraint is the total cost cannot exceed B. The second constraint limits CPA.

A. Some notes and details

Randomness of $\mathbf{x}$: Firstly, many features can be used to describe an impression, e.g., cookies, time, etc. It is the feature construction problem that exceeds our interests [3]–[5], [13], [27]. But 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 (PDF) $p_{\mathbf{x}}$ [30], [33]. Therefore, both the objective and the CPA constraint of (2) are in the expectation sense about $\mathbf{x}$. However, the budget constraint is 'hard' because advertisers cannot bid if they exhaust the budget.

Feature mappings: Due to $\mathbf{x}$'s construction is not unique, bid optimization usually uses its estimated values to bid, e.g., the predicted click-through rate p^{CTR}, the predicted conversion rate p^{CVR}, etc. It is the feature mapping and estimation problem that is beyond our interests [3], [12], [27], [34]–[36]. But, for modeling and analysis, those estimated values can be regarded as functions of $\mathbf{x}$ [7], [10], [13], [16], [23], [30], [37]. For convenience, we assume they are prior to this paper. Besides, by combining different estimated values, we can personally model varying advertisers' marketing objectives with v and a. For example, if someone wants to maximize the total number of transactions and hold the cost-per-click constraint, we can set $v = p^{\mathrm{CTR}} p^{\mathrm{CVR}}$ and $a = p^{\mathrm{CTR}}$ in (2). 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 the current $\mathbf{x}$.

Cost in click advertising: Nowadays, within e-commercial advertising, advertisers usually only need to **pay once for a winning impression that has been clicked-through** [7], [9], [11], which is the scenario considered in this paper. Therefore, the cost of an auction, c, is a piecewise function: if the winning impression has been clicked-through, $c = w$; otherwise, $c = 0$. Consequently, the cost c also is a function of $\mathbf{x}$. This billing method differs from display ADs with impression settlement considered in the previous works, which pay for the winning impression no matter whether it is clicked-through or not [10], [12], [13], [16], [24], [25], [30]–[32].

Auction logs format: Generally, to solve (2), AD campaigns have access to the historical auction logs. E-commercial platforms like *Taobao* have the full auction logs and can know the winning price of each partitioned auction for each AD. It differs from DSP, which only has the value v of the winning auctions and leads to a censored bid optimization [33], [38]. 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 the proposed method with the auction logs:

$$\{(t, w, p^{\mathrm{CTR}}, p^{\mathrm{CVR}}, i_p)\}_{t=1}^{T}, \tag{3}$$

where $t =$ 'year-month-day-hour-minute-second' is a string that represents the occurrence time of the auction. The predicted click-through rate and conversion rate are p^{CTR} and p^{CVR}, respectively. The item price is denoted as i_p.

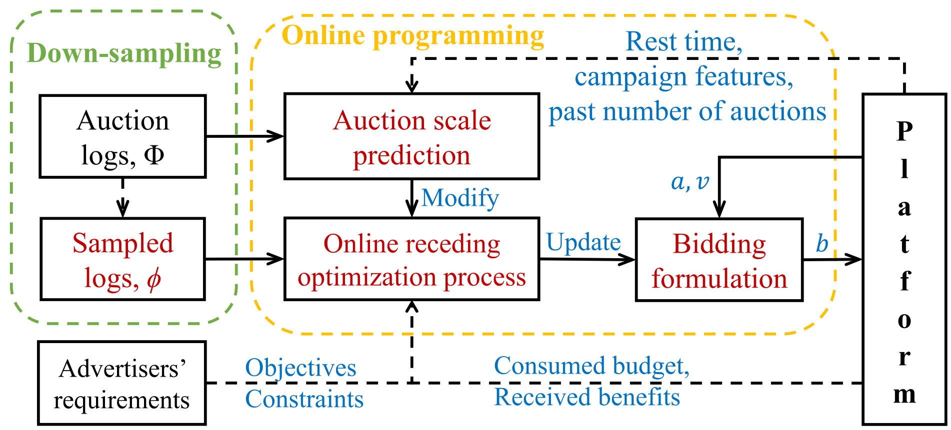

Fig. 3. Overview of our methodology. Solid lines represent real-time interactions, while dashed lines indicate periodic updates. Boxes with red letters are the designed modules. Blue letters denote the exchanged information.

IV. The Proposed Methodology

An overview of our methodology is shown in Fig. 3:

- **Down-sampling.** For each AD campaign with the historical auction logs, we randomly sample a collection of logs with specific sizes and cache it as the sampled logs. This process is carried out once a day.
- **Online programming.** For each auction, the bidding formulation bids a price b based on v and a. An online receding optimization process periodically updates parameters based on the sampled logs, constraints, etc. Meanwhile, the auction scale prediction module is used to modify for market fluctuations.

A. Online Receding Optimization Process

Fundamentally, (2) is a sequential decision-making problem; a common idea is analyzing it in the Markov Decision Process (MDP). Formally, we transform (2) 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:
$$r_t = v_t \mathbf{1}_t(w_t, b_t) - \lambda(c_t - C a_t) \mathbf{1}_t(w_t, b_t), \tag{4}$$
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 (2) is to solve

$$\arg\max_{\pi_t} \sum_{t=1}^{T} \mathbb{E}\,[r_t]. \tag{5}$$

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

(a) Due to the randomness of $\mathbf{x}$, getting the optimal solution may not be trivial. Meanwhile, the exact and adaptive solution for each AD campaign is also impractical and computationally exhausting as the market uncertainty.

(b) The MDP form (5) does not equal to the original problem (2) due to the existence of the CPA constraint. Also, different choices of λ would lead to varying results.

In a DP context, the optimal strategy performance improves when extra information is available, but it may make the intractable calculation. An alternative is called Approximate Dynamic Programming (ADP), which purposely ignores some extra information and pursues more computational-friendly results [39]. Regarding our bid optimization problem, the current state is known, and the future impression features are i.i.d and follow $p_{\mathbf{x}}$. Thus, we introduce the *open-loop feedback control (OLFC)* [39] and construct a receding optimization process to iteratively solve (5). In short, as one of the methods stems from the rollout mechanism, OLFC uses the current information as feedback to determine the probability of the future state. However, it solves the problem as if no further information would be received, using an open-loop receding time horizon optimization over the future evolution of the dynamic [39]. We introduce OLFC because although it is ADP, its result is **non-decreasing**, which means the strategy optimized by OLFC is not worse after each update [39].

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

$$\begin{aligned} \arg\max_{\bar{\pi}_t} \; & \sum_{t=n}^{T} \mathbb{E}\Bigg[v_t \mathbf{1}_t(w_t, \bar{\pi}_t) - \lambda\Big(B_n + \sum_{t=n}^{T} \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] \\ & \qquad - CA_n - C \sum_{t=n}^{T} \mathbb{E}[a_t \mathbf{1}_t(w_t, \bar{\pi}_t]\Big)\Bigg], \\ \text{s.t.} \; & B_n + \sum_{t=n}^{T} \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] \leq B, \end{aligned} \tag{6}$$

where $A_n = \sum_{t=1}^{n-1} a_t \mathbf{1}_t(w_t, b_t)$ is the received amount of actions, $B_n = \sum_{t=1}^{n-1} c_t \mathbf{1}_t(w_t, b_t)$ is the current consumption. Due to the introduction of OLFC, (6) cannot hold the budget constraint in the same form as (2). In other words, (6) only maintains the budget constraint in the expected sense, which is inevitable because we pursue a more tractable solution. For the n-th auction arrives, $\bar{\pi}_n$ is solved by (6), which only relates to $\mathbf{x}_n$. Then, we use it to bid a price $b_n = \bar{\pi}_n(\mathbf{x}_n)$.

To further analyze the bidding rules, we define

$$z_t := \mathbb{E}[\mathbf{1}(w_t, \bar{\pi}_t)] \in [0, 1], \tag{7}$$

where z_t can be regarded as the winning probability for the t-th auction with $\bar{\pi}_t$. Thus, (6) can be rewritten as the optimization problem concerning z_t in the following:

$$\begin{aligned} \arg\max_{z_t} \; & \sum_{t=m}^{T} \mathbb{E}\left[v_t z_t\right] \\ \text{s.t.} \; & \sum_{t=n}^{T} c_t z_t \leq B - B_n, \; \frac{B_n + \sum_{t=1}^{T} \mathbb{E}\left[c_t z_t\right]}{A_n + \sum_{t=1}^{T} \mathbb{E}\left[a_t z_t\right]} \leq C, \end{aligned} \tag{8}$$

which covers the gap between (5) and (2).

B. Bidding Formulation

So far, we know that bid optimization is to solve the optimal winning probability z in (8), and the bidding rule $\bar{\pi}$ is a way to achieve this probability. However, solving (8) still is not trivial because we do not know the exact PDF $p_{\mathbf{x}}$ of each AD campaign. Fortunately, with the i.i.d assumption, the daily auction logs can be seen as a collection of samples from $\mathcal{X}$ of each AD campaign. Thus, we can solve it in a data-driven manner, that is, by replacing expectations with logs:

$$\begin{aligned} \arg\max_{z_t} \; & \sum_{t=n}^{\hat{T}} \hat{v}_t z_t \\ \text{s.t.} \; & B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \leq B, \; \frac{B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t}{A_n + \sum_{t=n}^{\hat{T}} \hat{a}_t z_t} \leq C, \end{aligned} \tag{9}$$

where the superscript $\wedge$ means the value recorded in auction logs. Due to the click advertising and the i.i.d assumption of $\mathbf{x}$, $\hat{c}_t = \mathbb{E}[c_t \mathbf{1}_t(w_t, \bar{\pi}_t)] = \hat{w}_t \hat{p}_t^{\mathrm{CTR}}$. Clearly, (9) is a linear programming problem concerning z, and we have various methods to use [40]. **However, the solution cannot be used to bid but only tells us which impressions are worth winning.** Then, we introduce

$$q_t = \frac{\hat{v}_t + \beta^* C \hat{a}_t}{\hat{p}_t^{\mathrm{CTR}}(\alpha^* + \beta^*)}, \tag{10}$$

where parameters α^*, β^* are solved by the dual problem of (9) concerning z [10], [16]. Further, we have the following theorem to reveal the relationship between q and (9).

Theorem 1: For a bunch of auction logs with $\hat{T}$ records, suppose (9) and (10) are feasible. We have the following statements under the SPA mechanism:

(a) $\forall n = 1, \ldots, \hat{T}$, the optimal solution of (9), $\{z_n^*, \ldots, z_{\hat{T}}^*\}$, 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 (10) 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 (9), $\{z_n^*, \ldots, z_{\hat{T}}^*\}$, 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 (10) is one of the optimal bidding strategies for (9) in hindsight.

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

$$\begin{aligned} &\arg\max_{z_t} \sum_{t=n}^{\hat{T}} \hat{v}_t z_t \\ &\text{s.t. } \; B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t \leq B, \; \frac{B_n + \sum_{t=n}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} z_t}{A_n + \sum_{t=n}^{\hat{T}} \hat{a}_t z_t} \leq C, \end{aligned} \tag{11}$$

Rewrite it as follows:

$$\begin{aligned} &\arg\max_{z_t, a_t, d, g} \sum_{t=m}^{\hat{T}} \hat{v}_t \hat{q}_t(\hat{\mathbf{x}}_t) \\ &\text{s.t. } g + \sum_{t=m}^{\hat{T}} \hat{w}_t \hat{p}_t^{\mathrm{CTR}} \hat{q}_t(\hat{\mathbf{x}}_t) = B - B_n, \\ &z_t + a_t = 1, \; t = m, \ldots, \hat{T}, \\ &d + B_n - CA_n + \sum_{t=m}^{\hat{T}} z_t \left(\hat{w}_t \hat{p}_t^{\mathrm{CTR}} - C\hat{a}_t\right) = 0, \\ &g \geq 0, d \geq 0, z_t \geq 0, a_t \geq 0, t = m, \ldots, \hat{T}, \end{aligned} \tag{12}$$

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 (12). The BFS of (12) have at least $(\hat{T} - m + 1)$ number of 0. If (11) is feasible, then (12) is feasible, and there exists an optimal solution that is a BFS for (12). 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 (11), it has at least $(\hat{T} - n - 1)$ number of 0 and 1.

For the second statement, since we can write (11) to its dual form as:

$$\begin{aligned} &\arg\min_{\alpha, \beta, \gamma_t} (B - B_n)\alpha + \sum_{t=m}^{\hat{T}} \gamma_t + \beta(B_n - A_n C) \\ &\text{s.t. } \; \hat{w}_t \hat{p}_t^{\mathrm{CTR}} \alpha + \left(\hat{w}_t \hat{p}_t^{\mathrm{CTR}} - C\hat{a}_t\right)\beta - \hat{v}_t + \gamma_t \geq 0, t = m, \ldots, \hat{T}, \\ &\alpha \geq 0, \beta \geq 0, \gamma_t \geq 0, t = n, \ldots, \hat{T}. \end{aligned} \tag{13}$$

Let α^*, β^* be the optimal to (13) and define

$$q_t = \frac{1}{\hat{p}_t^{\mathrm{CTR}}} \frac{\hat{v}_t + \beta^* C\hat{a}_t}{\alpha^* + \beta^*} \tag{14}$$

Based on the dual theory, the optimal solutions of (13) satisfy $\forall t = n, \ldots, \hat{T}$:

$$z_t^* \left\{\hat{w}_t \hat{p}_t^{\mathrm{CTR}} (\alpha^* + \beta^*) - \hat{v}_t - Ca_t\beta^* + \gamma_t^*\right\} = 0, \tag{15}$$

$$(z_t^* - 1)\gamma_t^* = 0, \tag{16}$$

Then,

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

If $\mu_t < \hat{w}_t$, then $\hat{p}_t^{\mathrm{CTR}} (\hat{w}_t - q_t)(\alpha^* + \beta^*) + \gamma_t^* > 0$, which further implies that $z_t^* = 0$ by (15). 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, most AD campaigns have around 20000 auctions daily [5], [10], [11], [16], which satisfies the assumption $T \gg 2$ in Theorem 4.2.

Consequently, we use (10) as the designed bidding rule $\bar{\pi}$. For each AD campaign, we solve its dual problem of (9) based on its auction logs and get the dual solution α^* and β^*. Finally, for each coming impression, we bid a price b_t based on $\bar{\pi}_t$,

$$b_t = \bar{\pi}_t = \frac{v_t + \beta^* C a_t}{p_t^{\mathrm{CTR}} (\alpha^* + \beta^*)}. \tag{17}$$

Algorithm 1 Efficient solution using down-sampling

1: **Input:** Constraint B, C. Auction log Φ. 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: **Bidding parameters initialization:**
10:     Solve α_N and β_N in (10) based on B, C, and ϕ.
11:     $\alpha_0 = \alpha_N$, $\beta_0 = \beta_N$.
12: **For the n-th auction:**
13:     Delete records in ϕ and Φ, whose 'hour-minute-second' is earlier than the current time.
14:     $N_t =$ the number of records in ϕ.
15:     $\hat{T}_t =$ the number of records in Φ.
16:     **If** $N_t > 0$ **then**
17:         Solve α_N and β_N in (10) based on $(B - B_n)N_t/\hat{T}_t$, C, A_n,and ϕ.
18:         $\alpha_n = \alpha_N$, $\beta_n = \beta_N$.
19:     **Else**
20:         $\alpha_n = \alpha_{n-1}$, $\beta_n = \beta_{n-1}$.
21:     **End if**
22:     Update (17) with α_n and β_n, then bid a price.
23:     Wait for the auction result, cost c_n, and benefit a_n.
24:     $B_{n+1} = B_n + c_n$, $A_{n+1} = A_n + a_n$.
25:     $n = n + 1$.
26:     Wait for the next auction, and **go back to Step 14**.
27: **End for**

C. Efficient Solution using Down-sampling

Till now, the problem is to solve α^*, β^* in (10). Many solvers are available, but it is still costly when $\hat{T}$ is large (over 10^5) [22]. Considering millions of campaigns on the platform and the high real-time requirement, it struggles to afford the required computation resources. Note that, with the i.i.d assumption, the bid optimization problem is solved in a data-driven manner, which uses auction logs to solve α^* and β^* in (10). An intuitive idea to improve the efficiency is *Bootstrap* [41], i.e., downsample a smaller collection from the auction logs to solve for α and β instead of using the entire logs. Following this idea, we propose an efficient solution using down-sampling; its pseudo-code is presented in Algorithm 1, where Φ denotes the entire auction logs of an AD campaign. Moreover, for Algorithm 1, we have the following theorem to guarantee its effectiveness.

Theorem 2: For BCB, suppose (10) 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.

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

$$\begin{aligned} &\underset{q_t}{\text{maximize}} \ \sum_{t=1}^{\hat{T}} \hat{v}_t \\ &\text{s.t.} \ \ \sum_{t=1}^{\hat{T}} \hat{w}_t \hat{p}_t^{\text{CTR}} z_t \leq B, \end{aligned} \tag{18}$$

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

$$\begin{aligned} &\underset{\alpha,\gamma_t}{\text{minimize}} \ B\alpha + \sum_{t=1}^{\hat{T}} \gamma_t \\ &\text{s.t.} \ \ \hat{w}_t \hat{p}_t^{\text{CTR}} \alpha - \hat{v}_t + \gamma_t \geq 0, \ \forall t = 1, \ldots, \hat{T}, \\ &\qquad \alpha \geq 0, \\ &\qquad \gamma_t \geq 0, \ \forall t = 1, \ldots, \hat{T}, \end{aligned} \tag{19}$$

Let $\alpha^*, \gamma_1^*, \ldots, \gamma_{\hat{T}}^*$ denote the optimal solution to (19). By the complementary conditions for the optimal primal variables z_t^* and dual variables α^* of (18) and (19), we have

$$z_t^* \left[\hat{p}_t^{\mathrm{CTR}} \left(\hat{w}_t - \frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}} \right) \alpha^* + \gamma_t^* \right] = 0, \tag{20}$$

$$(z_t^* - 1)\gamma_t^* = 0, \tag{21}$$

If $\hat{w}_t > \frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}}$, for the optimal condition (20),

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

implying that $z_t^* = 0$ by (20). If $\hat{w}_t < \frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}}$, for the dual constraint in (19), 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 (21), we readily have $z_t^* = 0$. In summary, we have the following relation between $z_t *$ and $\frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}}$,

$$z_t^* = \begin{cases} 1, & \text{if } \frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}} > \hat{w}_t, \\ 0, & \text{if } \frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}} < \hat{w}_t, \end{cases}, \tag{22}$$

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 (18), we have

$$\alpha^* = \sup \left\{ \alpha \mid \sum_{t=1}^{\hat{T}} \hat{w}_t p_t^{\mathrm{CTR}} \mathbf{1}_{\left(\frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}} > \hat{w}_t \right)} \leq B \right\} \tag{23}$$

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,

$$\alpha^* = \sup \alpha, \tag{24}$$

Further, when we use Algorithm 1 to solve (19) with a sampling size $N \leq \hat{T}$, the optimal solution is regarded as α_N^*.

Based on (23),

$$\mathbf{1}_{\frac{\hat{v}_t}{\alpha^* \hat{p}_t^{\mathrm{CTR}}} > \hat{w}_t} = \mathbf{1}_{\alpha < \frac{\hat{v}_t}{\hat{w}_t \hat{p}_t^{\mathrm{CTR}}}}, \tag{25}$$

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.

Finally, we present an example to help understand why α is the quantile. For the feasible condition (23), 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, 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.). $\square$

The sketch is that solving (10) actually is a quantile estimation problem of a complex random variable, stemming from $\mathbf{x}$. And the sampling process does not affect the quantile [42]. We further guess the above also is true for MCB (α_N and β_N converge to α^* and β^*, respectively, almost surely.), but it is tough to elaborate since it relates to multiple random variables, and those are coupled with dual variables. To our knowledge, no work has answered the convergence guarantee for the sampling in solving BCB/MCB until now. On the other hand, mainstream fashion is still BCB, such that it can satisfy most cases, and the extension of MCB is our main work in the future. 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 in practical use (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**.

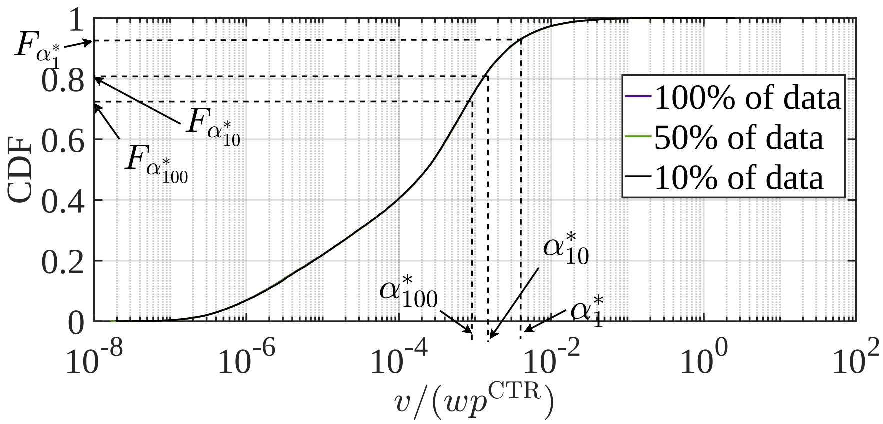

Fig. 4. An example illustrates the idea of using down-sampling to solve parameters efficiently. The horizon axis is the rate of v and wp^{CTR} recorded in the logs; the vertical axis is its CDF. Randomly pick 100%, 50%, and 10% logs and plot their CDFs; they are almost identical, showing the sampling 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.

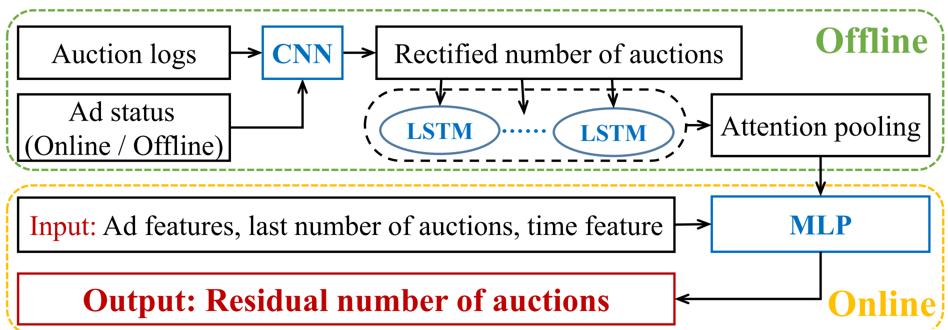

Fig. 5. The framework of auction scale prediction

D. Auction Scale Prediction

To compensate for market fluctuation, predictions are also helpful. Nevertheless, the exacts of each impression are not strictly necessary in our work because, in solving (10), they are regarded as samples from $\mathbf{x}$. Besides, due to the introduction of down-sampling, the factor that directly affects solving (10) are constraints related to the residual number of auctions.

Therefore, we pay more attention to designing an auction scale prediction module to have a better scale of the residual auction. The prediction framework is shown in Fig. 5. Firstly, we divide a day into 96 intervals, each 15 minutes. To do so, one is for the training convergence because the auction scale varies largely with different daily campaigns, from 0 to over a million, which may lead to a scant reward and state construction. On the other hand, the auction scale in a day may be time-related among different intervals. Separating a day into several intervals may reveal this relationship. In the offline part, for many reasons (budget is exhausted, etc.), an ad campaign may have different statuses (go online/offline) at the same moment on different days. Further, we use Convolutional Neural Network (CNN) to rectify them [15], where the kernel size is 3×3. Then, 96 long-short-term-memories (LSTMs) and attention pooling are used to construct relationships among different prediction timestamps. Soft attention with one neuron is used in attention pooling. For the online part, a 128-dimensional vector, which consists of and encodes the predicted ad campaign features, the auction scales of each past interval, and the time feature, inputs the network. A multi-layer perceptron (MLP) uses them with the results from the offline part to predict the residual number of auctions. All networks are fully connected and have two layers with 256 neurons in each layer. The activation function is *tanh*. All networks are trained by the mean squared error loss function. Finally, the network output is the predicted auction scales for each interval in the future. This prediction module is just a prototype for the work, which can become more profound in the future. Also, We conducted experiments to verify it in Section B.

E. Summary and implementation details

To sum up, Algorithm 2 is the pseudocode of our work. In addition, there are some implementation details,

Algorithm 2 Online receding control with prediction (RCP)

Input: Constraint B, C. The sampled logs ϕ. 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.5.
  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, $n = 1, \ldots, n_b$**:**
  Delete records in ϕ, whose 'hour-minute-second' is earlier than the current time.
  $N_n =$ the number of records in ϕ.
  $\hat{T}_n =$ the residual auction scale predicted by Fig.5.
  If $N_n > 0$ **then**
    Solve α^* and β^* in (10) 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 (17) with α_n and β_n.
  Repeat
    Bid prices based on (17) for each coming impression.
  Until receive reports $\{B_n, A_n\}$ from billing modules.
  $n = n + 1$.
End for

(a) **Delayed deduction mechanism.** This paper focuses on click advertising (refer to Section III); platforms do not know the cost of each auctioned impression until users' reactions have been observed (click or not). However, this time delay is remarkably larger than the interval between two consecutive auctions for an AD campaign [5], [11], [12], [27], implying it is impractical to update the bidding strategy for each auction. Thus, e-commerce platforms usually have a billing module that periodically reports forms of cost and received benefits; and the bidding strategy also updates periodically. Assume the billing module reports n_b times in a day, the bidding strategy also updates n_b times. After each update, the bidding strategy is fixed until the new reports arrive.

(b) **Reservoir sampling**. Considering millions of AD campaigns are on the platform, simultaneously downsampling is costly. Therefore, we use *reservoir sampling* [43], which runs parallel with the bidding process. The sampled logs are also finished at the end of a day, which can be directly and seamlessly used for the next day.

V. Online A/B experiments

We conduct online A/B experiments on *Taobao Sponsored Search*, one of China's largest online e-commercial platforms with billions of impressions daily. Although both BCB and MCB are considered, we concentrate more on BCB because:

(a) BCB is straightforward for comparison, with only one constraint. But MCB has two, it lacks a unified evaluation criterion to maintain fairness and persuasiveness when comparing them among different methods.

(b) Unlike our method, most previous works (e.g., [7], [10], [16], [16], [17], [20], [23], [31], [37], [44]) solved MCB by introducing the Lagrange multiplier likes (4). However, lacking a unified standard for the Lagrange multiplier selection causes their performance to vary with different multipliers, making comparisons under MCB less convincing than BCB.

A. Evaluation system and pipeline

Firstly, as shown in Fig. 6, we build a budget bucketing system for online A/B experiments. At the start of a day, some AD campaigns are randomly chosen. For each AD campaign, 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 candidate auction is randomly assigned a bucket number, and the bucket bidding strategy 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.

B. Experiment settings and details

- **Business scenarios:** Both BCB and MCB are considered:

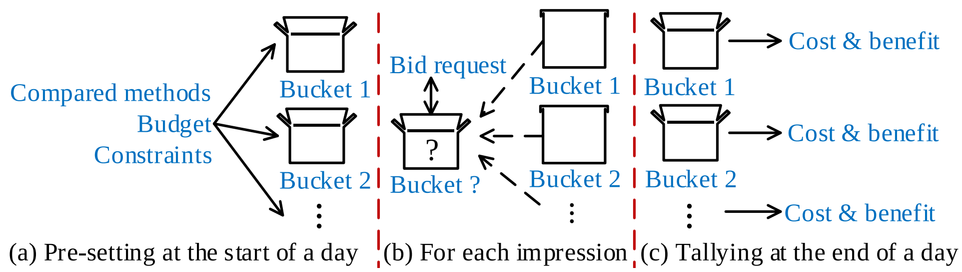

Fig. 6. Budget bucketing system for an AD campaign.

  - **For BCB,** it is the most used advertising service on Taobao. It maximizes the expected Gross Merchandise Volume (GMV) subject to the budget. Thus, in (2) and (17), $v_t = p_t^{\mathrm{CTR}} p_t^{\mathrm{CVR}} i_p$, $b_t = p_t^{\mathrm{CVR}} i_p / \alpha^*$.
  - **For MCB,** it maximizes the expected GMV subject to budget and cost-per-click (CPC) constraints. Thus, in (2) and (17), $v_t = p_t^{\mathrm{CTR}} p_t^{\mathrm{CVR}} i_p$, $a_t = p_t^{\mathrm{CTR}}$, $b_t = (p_t^{\mathrm{CVR}} i_p + \beta^* C)/(\alpha^* + \beta^*)$.
- **Time:** 09/01/2022 - 09/30/2022.
- **Scales:** Unless otherwise stated, we randomly select 2000 campaigns per day for each experiment (totaling around 4 million impressions).
- **Compared methods:** Since our experiments are conducted online, commercial safety and advertising performance cannot be ignored. Therefore, the compared methods must be the mainstream used and have practical applications to ensure the feasibility of experiments.
  (a) Multi-agent cooperation bidding games for bid optimization [23], denoted as **MACG**, which is the basic strategy used on Taobao. Directly comparing it with the other methods may not be fair enough, as it is a social welfare-aware approach. In contrast, our method and the other compared methods are selfish approaches. Thus, we mainly **regard its results as the online baseline for comparison**.
  (b) A unified solution to constrained bidding [10], [20], denoted as **USCB**, which is a state-of-the-art learning-based method. Its inputs contain the current impression values, the remnant budget ratio, the budget spending speed, the left time ratio, and the current CPC ratio for constraint [10]. It uses (17) as the bidding strategy, and output is its parameters.
  (c) Feedback control-based method [16], [31], [44], denoted as **FC**. It is a two-step approach, first planning the budget consumption and CPC within a day and getting the corresponding references, then using the feedback controller to adjust parameters in (17) so that the current consumption and CPC track these references. Since it provides a simple but useful idea to adjust strategies based on real-time feedback, it has been widely studied and used. Its inputs are the current impression values, the tracking errors of the budget consumption, and CPC. It uses (17) as the bidding strategy, and output is its parameters.
  (d) Our proposed method, Algorithm 2, denoted as **RCP**.
- **Methods detail 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 for reference planning. The feedback controller is proportional–integral–derivative, the most used in previous works [16], [31], [44].
  (c) **RCP** uses the previous day's auction logs for down-sampling. The sampling size $N = 10000$. The solver is MindOpt [45]. Auction logs from Aug. 2022 are used for training the auction scale prediction module.
- **Platform settings and others:**
  (a) Billing modules report the consumed budget and the current CPC every 15 minutes.
  (b) Considering commercial safety, each parameter update range is restricted to $\pm 80\%$ around their initial.
  (c) For fairness and persuasion, all methods use the same impression value estimation modules (e.g., $p_t^{\mathrm{CTR}}, p_t^{\mathrm{CVR}}$, etc.) and other supplementary (e.g., auction logs, etc.) provided by the platform.

C. Evaluation method and metrics

First, we count the following metrics for each AD campaign [6], [10]–[12], [16]. Unless otherwise stated, the results from **MACG** are used as the baseline for each campaign to eliminate the effect of different AD scales. Consequently, results from the other methods are normalized by 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 unavoidable error in impression value estimation, we use G to measure the actual performance of each method.

TABLE II
STATISTICS OF THE A/A VALIDATION EXPERIMENT IN BCB.[a]

	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

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

TABLE III
STATISTICS OF THE A/B EXPERIMENT IN BCB BUSINESS SCENARIOS.[a]

	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	0.2823	1.0000	0.2657	1.0000	0.3430	1.0000	0.2457	**1.0000**	–	**1.0000**	0.1552	1.0000	0.1140
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**

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

(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, which is usually also factored into advertisers' cost of sales.

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.

E. A/B experiment and discussion in BCB

In the second experiment, we compare different methods in BCB. The results are shown in Table III. To sum up:

(a) **RCP outperforms the other methods.** Overall, for Avg. of each metric, **RCP** performs best in five metrics; the rest belong to **MACG**. Specifically, for the maximizing objective, $\hat{G}$, **USCB**, **FC**, and **RCP** all perform better than **MACG**. The actual GMV, G, has some deceases compared to $\hat{G}$ because of the estimation error of p^{CTR} and p^{CVR}. Owing to the item price being very small for some ADs, even though we have maximized $\hat{G}$ much better, its improvement is still not significant. Thus, we consider D, ROI, and PPT comprehensively, which eliminates the effect of item prices and directly reflects the advertising capacity. **RCP** increases D and ROI by more than 30% and reduces PPT by nearly 10%, implying the effectiveness and outperforming of our method. Except for **MACG**, **RCP** still has 5%-10% improvements compared to **USCB** and **FC**.
(b) **RCP is relatively more stable.** Consider Std. of each normalized metric in Table III, **RCP** is smaller than the others. However, B_r is an except. In fact, **USCB** tends to exhaust the budget because the budget consumption is one of the states embedded in its input. Consequently, it usually achieves 100% budget consumption, inducing **USCB** achieves the minimal Std. for B_r.
(c) **RCPs are more efficient at winning high-value impressions.** An interesting phenomenon is that **MACG** achieves the lowest B_r and PPC, but its other metrics are remarkably worse than the other method, especially **ROI** and **PPT**. It reveals the low efficiency of **MACG** bidding for high-valued impressions. In other words, it does not tend to win those impressions with higher $\hat{G}$, whereas the impressions whose winning prices are relatively lower. It may not be a wise strategy for maximizing objective $\hat{G}$ and is not satisfactory for advertisers. In contrast, the other methods result in more budget and PPC but achieve significantly greater D and PPT. Remarkably, our method saves nearly 10% of payment for each transaction and results in the best performance (G, $\hat{G}$, D, and ROI), whereas it uses less budget than **USCB** and **FC**. It shows that **RCP** makes every penny count.

F. A/B experiment and discussion in MCB

Then, we conduct the third A/B experiment in MCB and 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 AD campaigns that violate CPC constraints to the total number of AD campaigns in the experiment.

TABLE IV
STATISTICS OF R_{cpc} AND V_c IN MCB.[a]

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

[a] 'Avg.' means the average. 'Max.' means the maximum.
[b] '$\mathrm{P}(V_c > 1.2)$' means the ratio of the number of AD campaigns whose $V_c > 1.2$ to the total number of AD campaigns.

TABLE V
STATISTICS OF THE A/B EXPERIMENT IN MCB BUSINESS SCENARIOS.[a]

	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**

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

(b) V_c, the quotient of the actual CPC to its constraint for an AD campaign. It is only used for AD campaigns that violate the constraint. Although the platform has recommended CPC constraints, some advertisers may still make unreasonable settings. This metric measures the robustness when facing unconscionable circumstances. Compared to the budget constraint, the CPC constraint is "soft" in that it is an average metric over the day and is not readily known precisely in advance. Thus, for advertisers on Taobao, it is also acceptable if the actual CPC does not exceed 20% of the set value [7], [10], [16].

The results are shown in Table IV and Table V. To sum up:

(a) Similar to BCB, Table V shows that **RCP performs better and is more stable** than the others.
(b) Table IV shows that **USCB, FC, and RCP potentially violate the CPC constraint in practice, but MACG does not.** There are three reasons: *Firstly*, the bid optimization problem (2) models the CPC constraint in the expected sense due to the randomness of impressions. *Secondly*, for MCB, **MACG** uses 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, the bid limitation does not exist in **USCB, FC, and RCP**. They have a broader range of bid impressions, especially high-value impressions, thus achieving better performance. *Lastly*, setting a CPC constraint is not as intuitive as budget, which requires advertisers to have a comprehensive overview of the market and balance different benefits and costs. Improper settings can lead to narrow, fragile solutions, significantly increasing CPC violation risks in practice. This prompts us to improve guidance on the rationality of CPC settings in the future.
(c) **Compared to USCB and FC, RCP achieves the lowest violation risk.** Though our method, **RCP**, has around 15% of AD campaigns that violate the CPC constraint, the maximum (Max.) and the average (Avg.) of V_c are remarkably lower than **USCB** and **FC**. Meanwhile, **RCP** has the fewest AD campaigns whose actual CPCs exceed 20% of their constraints, meaning it is likely to be more popular with advertisers.
(d) **RCP is better at winning high-value impressions.** As shown in Table V, four methods have similar PPC, but **RCP**'s PPT is much lower than the other methods. This means that impressions that **RCP** bids have a higher probability of transaction, resulting in higher objectives.

G. Ablation experiment

As mentioned in Fig.3 and Algorithm 2, in **RCP**, online receding optimization and auction scale prediction together help to update the bidding formulation. Therefore, we conduct the ablation experiment to analyze their effect. Specifically,

(a) **RCP**, the entire version of our method.
(b) **SS**, *the static strategy*, which is the version of **RCP** that does not have the online receding optimization process and auction scale prediction. We only solve the bidding formulation at the start of a day.
(c) **RCW**, *online receding control without prediction*, which is the version of **RCP** without auction scale prediction.

The results are shown in Table VI. Overall, three versions perform better than the baseline, **MACG**. Comparing **RCW** to **SS**, almost all metrics are significantly improved by 10%-20% on average, proving that the online receding optimization process is effective and coinciding with the non-decreasing properties of OLFC we used. On the other hand, the average improvement between **RCP** and **RCW** is much smaller. However, the stability of all metrics has been improved remarkably. It shows that introducing auction scale prediction mainly benefits stabilization when facing market fluctuations. Finally, there is an interesting

TABLE VI
STATISTICS OF THE ABLATION EXPERIMENT IN BCB.[a]

	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**

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

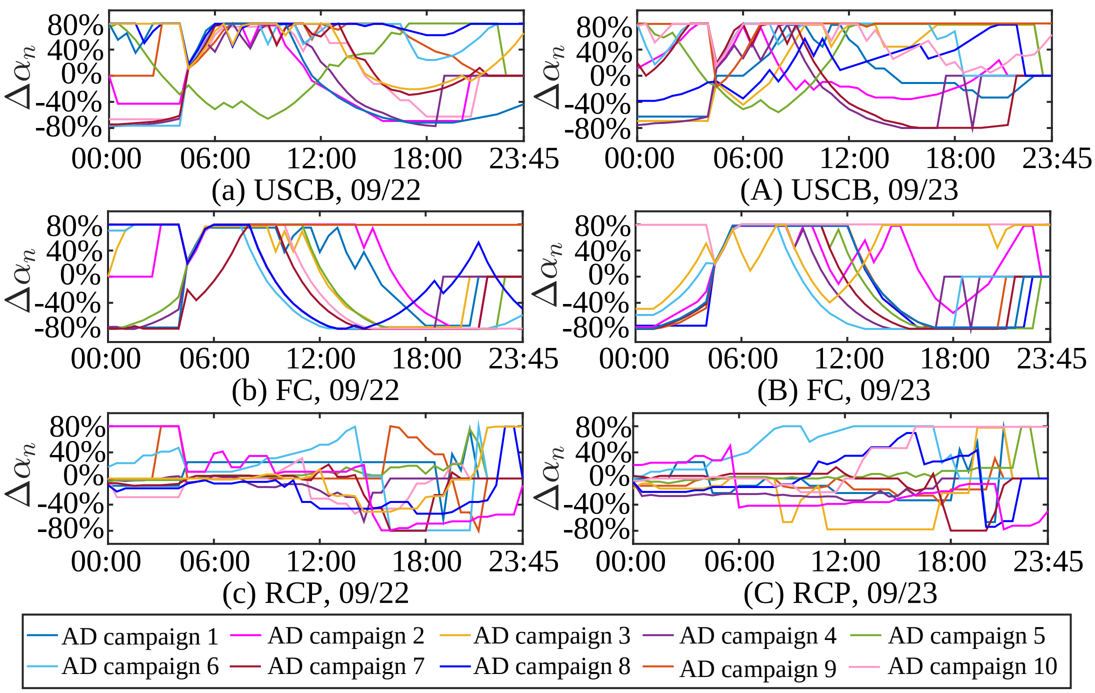

Fig. 7. Regulation processes of **USCB/FC/RCP**. The left column is from 09/22, and the right is from 09/23.

observation that **RCW** leads to the highest budget utilization B_r, while **RCP** is lower than it. In contrast, **RCP** achieves a higher ROI than **RCW**. This partly suggests that introducing predictions is crucial for designing effective and economical bid optimization methods.

H. Adaptability and personalization

Since different AD campaigns sell different items, they have varying target groups, and their impressions received on the same day may differ; even for the same AD, it is too arbitrary to say that the market and its impressions are the same in two days [5], [6]. An ideal strategy can self-adjust based on advertisers' objectives and market-changing. Thanks to (10), objectives are modeled in the optimization problem. Therefore, we focus on analyzing the regulation process of each method to reveal the capability to market changes. Particularly for **USCB/FC/RCP**, as they all use the bid formula (10), we hope to reveal the fundamental difference in our approach compared to previous works. For clarity, we use BCB as an example and provide a case study to discuss (MCB has coupled parameters to bid a price, making it complicated to present). Define

$$\Delta\alpha_n = \alpha_n/\bar{\alpha} - 1, \ n = 1, \dots, 96, \tag{26}$$

where α_n is the n-th update α in (10), $\bar{\alpha}$ is the initial value of each method. Further, 10 AD campaigns from different categories (clothing, electronics, etc.) on two adjacent days are randomly selected; and we plot their regulation process of **RCP**, **USCB**, and **FC** in Fig.7. In summaries:

(a) For the same day, there exists an interesting phenomenon that in **USCB** and **FC**, regulation processes seem to have **a uniform trend: rising before noon and falling at night.** By contrast, **RCP** does not have this trend.
(b) For the same AD campaign, **USCB** and **FC** still use the same strategy as mentioned before, but **RCP** does not.

We preliminarily speculate that **USCB** and **FC** focus too much on generalization to neglect personalization. In contrast, **RCP** solves a unique optimization problem of each AD and makes rolling updates based on real-time feedback and logs. However, comparing which method is better is not straightforward, as the analysis of impressions and markets is complicated and may

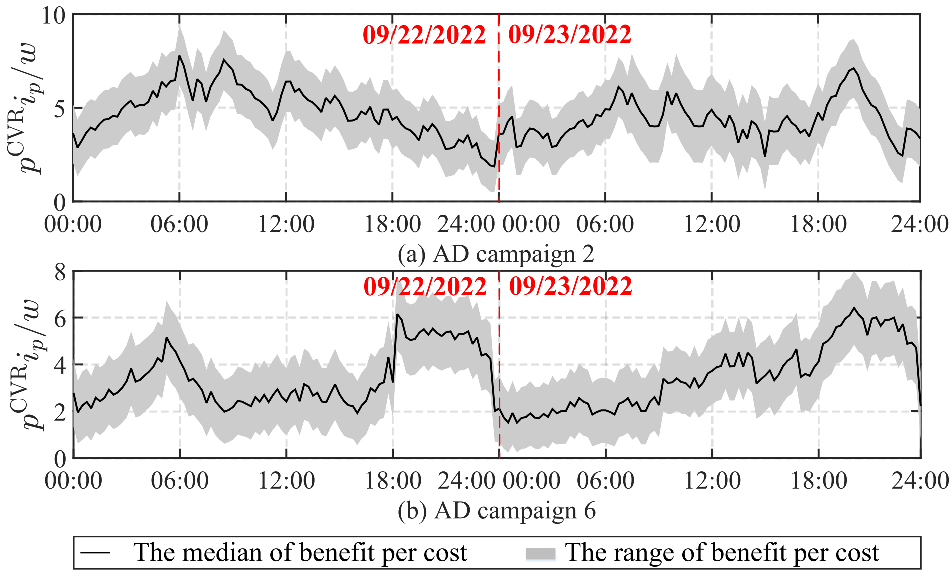

Fig. 8. Comparison of different AD campaigns' benefit per cost, $p^{\mathrm{CVR}} i_p/w$. The black line is the median, and the grey part is its range.

be coupled. Fortunately, all auction logs are accessible, so we can solve the optimal parameter α in (10) of each update in hindsight, denoted by α^h. Further, for each AD campaign, we get the *optimal update sequence* $\{\alpha_n^h \mid n = 1, \ldots, 96\}$ and calculate the Pearson correlation coefficient, ρ, between it and each method's update sequence $\{\alpha_n \mid n = 1, \ldots, 96\}$. Following this idea, we statistic ρ of all AD campaigns from the BCB A/B experiment, the average ρ of **USCB/FC/RCP**'s result in 0.3133/0.1523/0.6589, respectively. It verifies that **RCP** is closer to the optimal strategy in hindsight than **USCB** and **FC**. On the other hand, 10000 optimal update sequences are randomly selected to check the otherness among different AD campaigns and days, which results in the average of ρ is 0.0972. It means that each AD campaign needs a corresponding personalized strategy each day to adapt to market changes in time.

Thus, compared to **RCP**, we can say that **USCB** and **FC** do not perform well enough in providing personalized strategies and lack adaptability to market changes. In practice, we notice an interesting phenomenon: **USCB/FC**'s regulation processes mainly coincide with the trend of the entire platform: the quality and quantity of total impressions peak at noon and evening. While this may be a generalization for most AD campaigns, many high-value impressions may still missed. This phenomenon may be due to the curse-of-dimension, **USCB/FC** cannot use too many individual features to describe an AD campaign. In contrast, RCP makes the benefit of auction logs to facilitate online updates, implicitly considering each AD campaign's characteristics. Due to the complexity of online advertising, no existing methods can be used to measure a bidding strategy comprehensively. This case study provides a perspective from the optimal update sequence in hindsight and reveals a potential problem in previous works, as well as the balance between generalization and personalization. We hope it can be helpful for future work.

For example, in BCB and for an impression $\mathbf{x}_t$ with w_t, if we want to win it under SPA,

$$b_t \geq w_t \Rightarrow p_t^{\mathrm{CVR}} i_p/\alpha \geq w_t \Rightarrow p_t^{\mathrm{CVR}} i_p/w_t \geq \alpha, \tag{27}$$

where from the perspective of hindsight, the parameter α can be regarded as a threshold that determines whether the impression $\mathbf{x}_t$ is worth winning. Thus, $p_t^{\mathrm{CVR}} i_p/w_t$ can be approximately seen as the impression quality, we called it *benefit per cost*. Limited by the space, we randomly choose two AD campaigns in Fig.8, then plot their $p^{\mathrm{CVR}} i_p/w$ with time in Fig.8, which is aggregated in 15 minutes. To sum up:

(a) **For the same AD, its impression quality distribution varies from day to day.** On 09/22, AD campaign 2 peaked around 06:00-08:00; on 09/23, the peak occurred after 18:00. AD campaign 6 even had two peaks on 09/22.
(b) **Different AD campaigns have varying impression quality distributions.** Clearly, AD campaign 6 had two peaks on 09/22/2022, but AD campaign 2 did not.

Note that **USCB/FC/RCP** use the same bid formulation (10); different impression quality distribution may lead to different optimal dual solution α^* in hindsight. Consequently, in hindsight, they have different optimal budget consumption and the received benefits (i.e., the expected GMV, $\hat{G}$).

Following this idea, we solve the optimal α^* in hindsight with the entire auction logs (Limited by space, we use AD campaign 6 on 09/22/2022 as a case study; the results also exist for other AD campaigns.). Then we replay the auction logs to get the *optimal budget consumption* and the *optimal expected GMV* sequences, denoted as $\{B_1^*, \ldots, B_{96}^*\}$ and $\{\hat{G}_1^*, \ldots, \hat{G}_{96}^*\}$, respectively. Finally, we compare the actual budget consumption and received benefits of each method to optimal sequences to

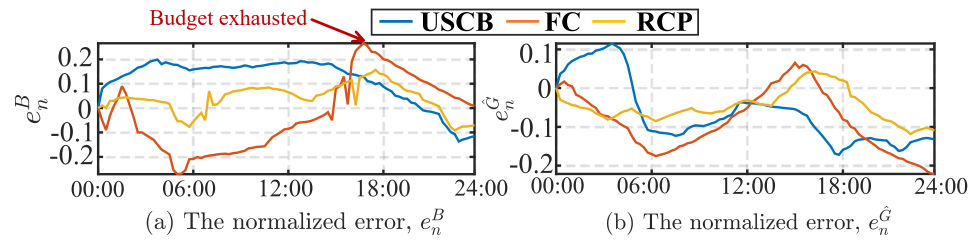

Fig. 9. Comparison of the normalized errors e_n^B and $e_n^{\hat{G}}$.

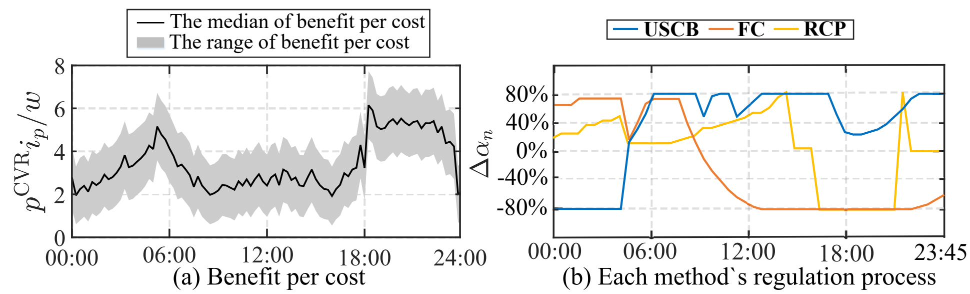

Fig. 10. Benefit per cost and each method's regulation process for AD campaign 6 on 09/22/2022.

compare their adaptability. For a clear comparison, we define the following normalized error with the corresponding optimal sequences

$$e_n^B = B_n/B_n^* - 1,\ e_n^{\hat{G}} = \hat{G}_n/\hat{G}_n^* - 1,\ n = 1, \dots, 96. \tag{28}$$

The result is shown in Fig. 9. For analysis, the benefit per cost and each method's regulation process of AD campaign 6 on 09/22/2022 are plotted in Fig. 10. To sum up:

(a) **Compared to USCB and FC, RCP's trajectories are closer to the optimal ones in hindsight.**

(b) **RCP is relatively more adaptive than USCB and FC.** As shown in Fig.9-(a), three methods' e_n^B have significant difference. Especially, **FC** even exhausted its budget before 18:00, resulting in the lowest $\hat{G}$ and $e_n^{\hat{G}}$. Combining Fig.10, we can know that **FC** bid relatively higher prices during 06:00 - 18:00, which won a batch of low-valued impressions. As a result, it exhausted the budget. **USCB** performed better than **FC**, while it still won too many low-valued impressions before 06:00. Besides, for the peak around 18:00, **USCB** decreased α such that increasing bids, but the degree was too small, which makes it miss part of high-valued impressions. By contrast, **RCP** is more adaptive: it bids the lowest price it can before the first peak. Besides, for the second peak, it forecasts its arrival and decreases α such that bidding a higher price. Consequently, **RCP** achieves the lowest e_n^B and $e_n^{\hat{G}}$ in the end.

(c) **RCP still exists a large room for future improvements.** Although **RCP** outperforms **USCB** and **FC**, it still suffers from some prediction errors, leading to unreasonable operations such as increasing α in the last two hours.

VI. Conclusion

This paper proposes a new methodology for solving bid optimization in e-commerce advertising. The solution consists of down-sampling, receding optimization, and auction scale prediction. A series of online A/B experiments are made on *Taobao Sponsored Search*, and the results show the superiority of our work over different perspectives.

Appendix A
Empirical experiments for sampling size

We randomly select 1000 AD campaigns from Taobao (500 for BCB, 500 for MCB). 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 AD campaign. Each AD campaign is sampled and solved 100 times for each N. **For BCB,** it maximizes the expected Gross Merchandise Volume (GMV) subject to budget constraint. **For MCB,** it maximizes the expected GMV subject to budget and cost-per-click (CPC) constraints. Evaluation metrics are:

(a) **The relative errors** of α and β, defined

$$\tilde{\alpha}_N = \alpha_N/\alpha_{+\infty} - 1, \quad \tilde{b}_N = b_N/b_{+\infty} - 1, \tag{29}$$

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 $\beta_{+\infty}$.

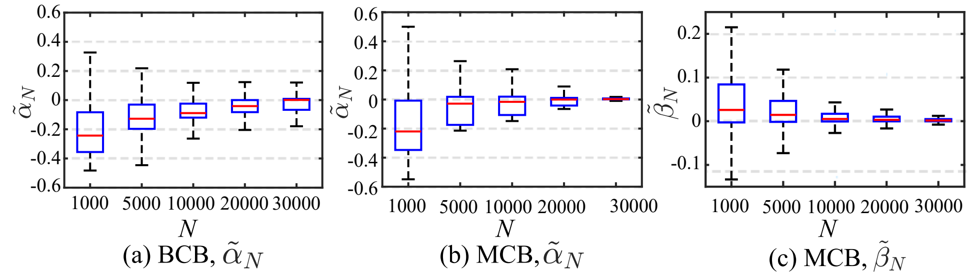

Fig. 11. Boxplots of the relative errors with different N. (a) is $\tilde{\alpha}_N$ of BCB, (b) is $\tilde{\alpha}_N$ of MCB, (c) is $\tilde{\beta}_N$ of MCB.

TABLE VII
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

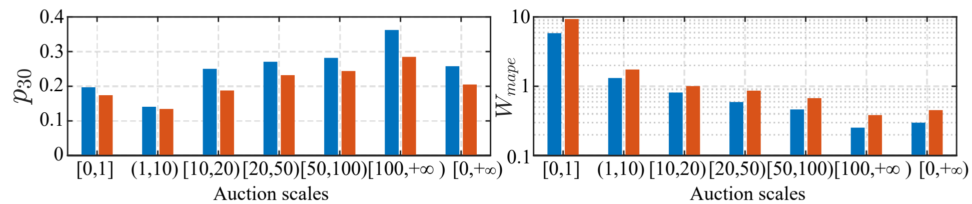

Fig. 12. The comparison of p_{30} and W_{mape}. The blue bar denotes our method. The red bar denotes the approach that directly uses yesterday's data.

(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 AD campaign as itself baseline and define

$$\tilde{T}_N = \frac{T_N}{\bar{T}_{1000}},\ \tilde{C}_N = \frac{C_N}{\bar{C}_{1000}},\ \tilde{M}_N = \frac{M_N}{\bar{M}_{1000}}. \tag{30}$$

Experiments are implemented on Dataworks [46], and the solver is MindOpt [45]. Results are shown in Fig. 11-Table VII.

(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.

Finally, by balancing accuracy and consumption, we choose $N = 10000$ in our work. Meanwhile, **the solving time of an AD 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.

APPENDIX B
EXPERIMENTS OF AUCTION SCALE PREDICTION

We use the auction logs from 05/11/2022-05/17/2022 provided by Taobao for network training, whose set is over 500 GB. Then the network is used to predict the auction scale of each campaign on 05/18/2022-05/19/2022. Several groups are set based on the auction scales to compare our prediction and the one that directly used yesterday's data. There are six groups, which are $[0, 1]$, $(1, 10)$, $[10, 20)$, $[20, 50)$, $[50, 100)$, and $[100, +\infty)$. For example, $[0, 1]$ means that the auction scale is less than/equals to 1 time. The compared metrics are:

- **The probability of the absolute error less than** 30%, denoted by p_{30}, the higher, the better [47]. It means the probability of the prediction is around the truth $\pm 30\%$.
- **The weighted mean absolute percentage error**, denoted by W_{mape}, the lower, the better. In this experiment, all AD campaigns has the same weight.

The results are shown in Fig. 12. When the auction scale increases, the accuracy p_{30} increases, and the absolute error W_{mape} decreases. Compared to the approach that directly uses yesterday's data, the designed prediction module performs better by at least 20%. Though it still has a significant gap to the truth, it is enough for usage [47].

References

[1] J. Aguilar and G. Garcia, "An adaptive intelligent management system of advertising for social networks: A case study of facebook," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 20–32, 2017.

[2] J. Li, X. Ni, and Y. Yuan, "The reserve price of ad impressions in multi-channel real-time bidding markets," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 583–592, 2018.

[3] F. Lyu, X. Tang, H. Guo, R. Tang, X. He, R. Zhang, and X. Liu, "Memorize, factorize, or be naïve: Learning optimal feature interaction methods for ctr prediction," in *2022 IEEE 38th International Conference on Data Engineering (ICDE)*. IEEE, 2022, pp. 1450–1462.

[4] Y. Yuan, F. Wang, J. Li, and R. Qin, "A survey on real time bidding advertising," in *Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics*. IEEE, 2014, pp. 418–423.

[5] S. Yuan, J. Wang, B. Chen, P. Mason, and S. Seljan, "An empirical study of reserve price optimisation in real-time bidding," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1897–1906.

[6] 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.

[7] 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.

[8] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," *The Journal of finance*, vol. 16, no. 1, pp. 8–37, 1961.

[9] D. Bergemann, P. Dütting, R. Paes Leme, and S. Zuo, "Calibrated click-through auctions," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 47–57.

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] K.-C. Lee, A. Jalali, and A. Dasdan, "Real time bid optimization with smooth budget delivery in online advertising," in *Proceedings of the seventh international workshop on data mining for online advertising*, 2013, pp. 1–9.

[19] Y. Chen, P. Berkhin, B. Anderson, and N. R. Devanur, "Real-time bidding algorithms for performance-based display ad allocation," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2011, pp. 1307–1315.

[20] J. Zhao, G. Qiu, Z. Guan, W. Zhao, and X. He, "Deep reinforcement learning for sponsored search real-time bidding," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 1021–1030.

[21] W. C.-H. Wu, M.-Y. Yeh, and M.-S. Chen, "Predicting winning price in real time bidding with censored data," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1305–1314.

[22] G. Dantzig, "Linear programming and extensions," in *Linear programming and extensions*. Princeton university press, 2016.

[23] 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.

[24] H. Cai, K. Ren, W. Zhang, K. Malialis, J. Wang, Y. Yu, and D. Guo, "Real-time bidding by reinforcement learning in display advertising," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 2017, pp. 661–670.

[25] 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.

[26] S. R. Balseiro and Y. Gur, "Learning in repeated auctions with budgets: Regret minimization and equilibrium," *Management Science*, vol. 65, no. 9, pp. 3952–3968, 2019.

[27] B. Chandramouli, J. Goldstein, and S. Duan, "Temporal analytics on big data for web advertising," in *2012 IEEE 28th international conference on data engineering*. IEEE, 2012, pp. 90–101.

[28] J. Fernandez-Tapia, "Optimal budget-pacing for real-time bidding," *Social Science Electronic Publishing*, 2015.

[29] S. C. Geyik, L. Chowdhury, F. Raudies, W. Pu, and J. Shen, "Impression pacing for jobs marketplace at linkedin," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 2020, pp. 2445–2452.

[30] W. Zhang, "Optimal real-time bidding for display advertising," Ph.D. dissertation, UCL (University College London), 2016.

[31] 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.

[32] 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.

[33] W. Zhang, T. Zhou, J. Wang, and J. Xu, "Bid-aware gradient descent for unbiased learning with censored data in display advertising," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 2016, pp. 665–674.

[34] H. Wen, J. Zhang, F. Lv, W. Bao, T. Wang, and Z. Chen, "Hierarchically modeling micro and macro behaviors via multi-task learning for conversion rate prediction," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 2187–2191.

[35] G. Zhou, X. Zhu, C. Song, Y. Fan, H. Zhu, X. Ma, Y. Yan, J. Jin, H. Li, and K. Gai, "Deep interest network for click-through rate prediction," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 1059–1068.

[36] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, "Deepfm: a factorization-machine based neural network for ctr prediction," *arXiv preprint arXiv:1703.04247*, 2017.

[37] K. Ren, W. Zhang, K. Chang, Y. Rong, Y. Yu, and J. Wang, "Bidding machine: Learning to bid for directly optimizing profits in display advertising," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 4, pp. 645–659, 2017.

[38] K. Amin, M. Kearns, P. Key, and A. Schwaighofer, "Budget optimization for sponsored search: Censored learning in mdps," *arXiv preprint arXiv:1210.4847*, 2012.

[39] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.

[40] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*. Athena Scientific Belmont, MA, 1997, vol. 6.
[41] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.
[42] P. Roy, R. Laprise, and P. Gachon, "Sampling errors of quantile estimations from finite samples of data," *arXiv preprint arXiv:1610.03458*, 2016.
[43] 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.
[44] 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.
[45] A. D. Academy, "Mindopt optimization suite." [Online]. Available: https://solver.damo.alibaba.com/htmlpages/page#/en
[46] A. Cloud, "Dataworks." [Online]. Available: https://www.alibabacloud.com/help/en/dataworks/latest/what-is-dataworks
[47] 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.

TABLE VIII
ALL FIGURE CAPTIONS

Fig. 1	Brief process of online advertising.
Fig. 2	The proposed methodology.
Fig. 3	Overview of our methodology. Solid lines represent real-time interactions, while dashed lines indicate periodic updates. Boxes with red letters are the designed modules. Blue letters denote the exchanged information.
Fig. 5	The framework of auction scale prediction.
Fig. 6	Budget bucketing system for an AD campaign.
Fig. 7	Regulation processes of **USCB/FC/RCP**. The left column is from 09/22, and the right is from 09/23.
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.
Fig. 12	The comparison of p_{30} and W_{mape}. The blue bar denotes our method. The red bar denotes the approach that directly uses yesterday's data.