

Control-based Bidding for Mobile Livestreaming Ads with Exposure Guarantee

Haoqi Zhang
Shanghai Jiao Tong University
Shanghai, China
zhanghaoqi39@sjtu.edu.cn

Junqi Jin
Alibaba Group
Beijing, China
junqi.jjq@alibaba-inc.com

Zhenzhe Zheng*
Shanghai Jiao Tong University
Shanghai, China
zhengzhenzhe@sjtu.edu.cn

Fan Wu
Shanghai Jiao Tong University
Shanghai, China
fwu@cs.sjtu.edu.cn

Haiyang Xu
Alibaba Group
Beijing, China
shenzhou.xhy@alibaba-inc.com

Jian Xu
Alibaba Group
Beijing, China
xiyu.xj@alibaba-inc.com

ABSTRACT

Mobile livestreaming ads are becoming a popular approach for brand promotion and product marketing. However, a large number of advertisers fail to achieve their desired advertising performance due to the lack of ad exposure guarantee in the dynamic advertising environment. In this work, we propose a bidding-based ad delivery algorithm for mobile livestreaming ads that can provide advertisers with bidding strategies for optimizing diverse marketing objectives under general ad performance guaranteed constraints, such as ad exposure and cost-efficiency constraints. By modeling the problem as an online integer programming and applying primal-dual theory, we can derive the bidding strategy from solving the optimal dual variables. The initialization of the dual variables is realized through a deep neural network that captures the complex relation between dual variables and dynamic advertising environments. We further propose a control-based bidding algorithm to adjust the dual variables in an online manner based on the real-time advertising performance feedback and constraints. Experiments on a real-world industrial dataset demonstrate the effectiveness of our bidding algorithm in terms of optimizing marketing objectives and guaranteeing ad constraints.

CCS CONCEPTS

• Information systems → Display advertising; • Applied computing → Electronic commerce.

*Z. Zheng is the corresponding author.

This work was supported in part by Science and Technology Innovation 2030 – “New Generation Artificial Intelligence” Major Project No. 2018AAA0100905, in part by China NSF grant No. 62132018, 61902248, 62025204, in part by Shanghai Science and Technology fund 20PJ1407900, in part by Alibaba Group through Alibaba Innovation Research Program, and in part by Tencent Rhino Bird Key Research Project. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

CIKM '22, October 17–21, 2022, Atlanta, GA, USA

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9236-5/22/10...\$15.00

<https://doi.org/10.1145/3511808.3557269>

KEYWORDS

E-commerce; Online Advertising; Control-based Bidding

ACM Reference Format:

Haoqi Zhang, Junqi Jin, Zhenzhe Zheng, Fan Wu, Haiyang Xu, and Jian Xu. 2022. Control-based Bidding for Mobile Livestreaming Ads with Exposure Guarantee. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22)*, October 17–21, 2022, Atlanta, GA, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3511808.3557269>

1 INTRODUCTION

With the rapid development of mobile computing, many new formats of advertisements have emerged, such as mobile social media ads [29], location-based ads [27] and mobile livestreaming ads. Among them, mobile livestreaming ads [26] can promote products through real-time product demonstrations and frequent interaction between hosts and users. Such intensive host-user interaction and vivid demonstrations make mobile livestreaming ads more entertaining and attractive than traditional advertising formats. The market scale of livestreaming ads achieved over 170 billion dollars in China in 2021 [20], and will reach 25 billion in US by 2023 [26].

Despite the rapidly growing market scale, mobile livestreaming ads also bring several new challenges to efficient ad delivery within dynamic advertising environments. First, *ad exposure guarantee*: a good ad delivery strategy should guarantee livestreaming ads receive adequate exposure or page views (PVs) within a short period of time, *e.g.*, two to three hours, quite different from the traditional ad campaigns lasting for a long time. A guarantee on ad exposure can enable advertisers to estimate expected advertising performance in advance, and reduce the risk of their ad investments. Second, *dynamic advertising environments*: the advertising context can fluctuate significantly, since the user impression distribution of the ad platform varies greatly in both quality and quantity, and the host can promote various products at different times. Furthermore, the specific products within livestreaming ads are always unknown to the ad platform in advance, and thus each ad is considered as a new ad, resulting in the notorious cold start problem. Thus, a good ad delivery strategy should adapt to the dynamic environment and handle the cold start problem well. Finally, *multi-objective under multi-constraint*: livestreaming ads have diverse marketing objectives, such as increasing gross merchandise volume (GMV) and the popularity of hosts. At the same time, advertisers also set multiple

constraints on performance metrics, such as Return On Investment (ROI) and Pay-Per-Click (PPC), to ensure the cost-efficiency of ad delivery. This leads to a challenging multi-objective optimization problem with general ad delivery constraints.

The existing works are not sufficient to jointly solve the above three challenges. A common solution to satisfy the required ad exposure is guaranteed delivery (GD) [5, 8, 14], in which advertisers sign contracts with the ad platforms in advance for delivering a guaranteed amount of ad exposure at negotiated prices. The ad platform then executes these contracts by matching advertisers with pre-selected user impressions, resulting in inefficient ad allocation in dynamic environments. Furthermore, in most cases, GD only considers the performance of ad exposure, but not other marketing objectives and constraints. A potential solution for overcoming the disadvantages of GD is real-time bidding (RTB) [33, 34]. By participating in online ad auctions, the advertisers could optimize multiple marketing objectives in a flexible manner [22, 32]. However, the bidding environment experiences an erratic fluctuation due to the uncertain quantity and quality of user impressions and the dynamics of bidding strategies from competing advertisers, so the performance obtained by RTB is quite unstable.

In this work, we consider a new format of e-commerce advertisement: mobile livestreaming ads, and investigate the problem of its exposure guarantee in dynamic advertising environments. We jointly leverage the advantages of GD and RTB, and propose a control-based bidding algorithm to maximize advertisers' diverse marketing objectives under general constraints. We formulate the problem as an online integer programming, and show that the optimal bidding strategy can be derived based on the corresponding optimal dual variables. To achieve this, we first design a deep neural network, namely Λ -Network, and build the connection between dual variables and environment features. With this connection, we can establish a good initialization for the dual variables, overcoming the cold-start problem. To further adjust dual variables in dynamic environments, we leverage the idea of Model Predictive Control (MPC) [7], and update dual variables based on the real-time advertising performance feedback. To enhance the control capability of MPC, we jointly consider a proposed historical model and a real-time model, exploiting both historical knowledge and real-time information simultaneously. The historical model is a multi-variable feedback control module, and the real-time model is implemented by a periodic linear programming (LP) over the collected information during the ongoing ad campaign. We conduct comprehensive experiments over an industrial dataset to evaluate the performance of our algorithm. The experiment results demonstrate that the control-based bidding algorithm can simultaneously satisfy the requirement of exposure guarantee and optimize the multiple marketing objectives in dynamic advertising environments.

Our contributions in this work can be summarized as follows:

- We investigate the problem of exposure guarantee for mobile livestreaming ads in dynamic advertising environments. We formulate this problem of bid optimization as an online integer programming and build a connection between the bidding strategy and the corresponding dual variables. We detour the challenge of searching for an appropriate initialization of dual variables by approximating this process through a deep neural network.

- We propose a control-based bidding algorithm that combines a historical model and a real-time model to dynamically adjust dual variables. The historical model overcomes the cold start problem at the early stage, and the real-time model can fully leverage the newly collected information. Our theoretical results show that such a control-based bidding algorithm can well optimize marketing objectives and keep performance constraints at the same time.

- We conduct extensive experiments on a real-world industrial dataset. The experiment results demonstrate the superiority of the proposed control-based bidding algorithm in terms of ad performance constraints satisfaction and marketing objective optimization, achieving 95.75% of ad exposure and 93.03% of marketing objective of the optimal offline solution on average.

2 RELATED WORKS

Guaranteed Delivery: The GD allocation problem can be formulated as an online matching over a user-advertiser bipartite graph. The online matching problem is solved in [15, 21] under the assumption that the arriving user impressions are drawn independently from a certain distribution. The primal-dual method is introduced in [10], under which Chen et al. further proposed High Water Mark (HWM) algorithm by assigning an equal fraction of available user impressions for each contract [8]. SHALE algorithm is designed in [5] by converting the optimal dual variables to a good solution for the primal GD allocation problem. Fang et al. modeled the problem at an individual level and developed a real-time pacing strategy for fulfilling contracts smoothly [14]. Different from these GD approaches, we guarantee ad exposure by finding a proper bidding strategy instead of directly matching user impressions and ads.

Bid Optimization with Multiple Constraints: In online advertising, the bid optimization problem is well studied [23, 25, 35]. The classical formulation is budget constraint bidding (BCB) with a single optimization objective [6, 30, 34]. In [34], Zhang et al. proposed a bidding strategy with the form of $\frac{v}{\lambda}$, where v is the value of each impression, and the budget constraint can be satisfied by adjusting the pacing parameter λ . In [30], Wu et al. solved the budget constraint bidding by applying a model-free reinforcement learning method to learn a proper λ . However, these works did not consider other objectives and constraints, such as PPC and PV. In [33], by applying a feedback control method to adjust the bidding strategy based on real-time performance, other KPIs such as PPC and auction-winning-rate could be optimized. In [16], Ghosh et al. proposed an algorithm to acquire the given number of ad exposure within the preset budget. The work [18, 22, 32] attempted to extend BCB to multiple constraints bidding (MCB) with diverse optimization objectives. In [22], Kitts et al. proposed a generic algorithm to address multiple constraints as well as optimize the target advertising value by truncating bid prices. In [32], the problem of GMV maximization under budget and PPC constraints is formulated as an online LP problem. By leveraging the primal-dual method, Yang et al. derived the bidding strategy based on the optimal dual variables. In [18], He et al. further proposed a policy-based reinforcement learning method to adjust the bidding strategy during advertising. However, none of these approaches considered the guaranteed delivery of ad exposure, which is critical to livestreaming ads.

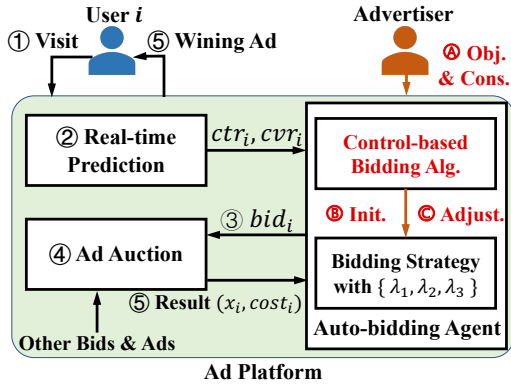


Figure 1: A Guaranteed Ad Delivery System.

Online algorithms with guaranteed competitive ratios have also been proposed to solve the online LP in the literature of operation research [2, 17]. In [2], the optimal dual variables are approximated by solving an offline LP on the observed information until the current time slot. In [17], Gupta et al. solved the online LP problem by converting it to the load balance problem, which can be solved with the expert algorithm. Different from previous works, we consider not only packing constraints about advertisers' limited resources (e.g., budget) but also covering constraints for ensuring the ad exposure. The dynamic advertising environment also brings high exploration costs and intensifies the cold start problem, requiring us to jointly consider the historical knowledge from previous LP instances and real-time information from ongoing environments.

3 PRELIMINARIES AND PROBLEM FORMULATION

In this section, we review ad guaranteed delivery process in dynamic advertising environments. We then give the formal mathematical formulation for multi-objective optimization under general constraints. We finally derive the optimal bidding strategy.

3.1 An Ad Delivery System

We first present the classical ad delivery system as shown in Fig. 1, which usually contains three entities: *users*, *advertisers*, and *ad platform*. A typical workflow of the advertising system contains five steps: 1) Once a user visits the ad platform, an ad display request would be sent to the auto-bidding agents that may be interested in displaying ads to this user. 2) The ad platform provides the predicted advertising performance of the user-advertiser pair, such as Click-Through Rate (CTR) [9, 36] and Conversion Rate (CVR) [31] for the auto-bidding agents. 3) The auto-bidding agents submit bid prices along with the ads back to the ad auction. 4) After receiving bidding prices of all agents, the ad platform launches a Generalized Second Price (GSP) auction [13] to determine the winning ad. 5) The user interacts with the displayed winning ad, and the interaction feedback is used as a reference for payment calculation. In this work, we adopt Cost-Per-Click (CPC) method [19], in which the winning advertiser would be charged only if the user clicks her ad.

Our bidding-based guaranteed ad delivery components can be integrated into the above procedure by introducing three additional steps: A) At the beginning of ad campaigns, advertisers submit

targeting advertising performance to auto-bidding agents, that is, specific marketing objectives and multiple constraints over ad exposure, budget, expected PPC and other KPIs. B) After receiving the ad delivery requests, the auto-bidding agents evaluate the feasibility of these requests, and initialize bidding strategies. C) Within the dynamic advertising environments, agents adaptively adjust the bidding strategies based on real-time feedback to fulfill the ad delivery requests.

3.2 Problem Formulation

We formulate the problem of guaranteed ad delivery in mobile livestreaming ads by specifying the advertiser's marketing objectives and constraints. Instead of simply optimizing a single KPI, we model the advertiser's general optimization objective by a weighted linear combination of multiple ad performance indicators. During the ad campaign, we assume there are a total of N users visiting the ad platform in sequence. For each user impression i , the ad platform can predict the ad performance, such as predicted CTR ctr_i , expected GMV $gmvi$, and also the expected $cost_i$ of displaying the ads to user i . We can formulate the problem of guaranteed ad delivery as the following online integer programming (Online IP):

$$\max_{\mathbf{X}} \sum_{i=1}^N x_i \times (\alpha \times gmvi + \beta \times ctr_i), \quad (1)$$

$$s.t. \sum_{i=1}^N x_i \times cost_i \leq Budget, \quad (2)$$

$$\frac{\sum_{i=1}^N x_i \times cost_i}{\sum_{i=1}^N x_i \times ctr_i} \leq PPC, \quad (3)$$

$$\sum_{i=1}^N x_i \geq PV, \quad (4)$$

$$x_i \in \{0, 1\}, \forall i \in [1, N],$$

where the vector $\mathbf{X} = (x_1, x_2, \dots, x_N)$ indicates whether user i is selected to display the ad, i.e., $x_i = 1$ when user i is selected, and $x_i = 0$, otherwise. We consider three types of representative ad performance constraints in the above formulation. The constraint in (2) is the classical budget constraint [34]. Inequality (3) represents the constraints for the ratio of performance indicators, such as average Pay-Per-Click (PPC). ROI constraint also falls into this type of constraint. Inequality (4) is about constraint over guaranteed ad performance, such as PV constraint for guaranteeing ad exposure.

There exist several challenges in solving the above online integer programming. First, advertisers participate in ad auctions by offering bids. Instead of directly deciding \mathbf{X} , we need to design an algorithm to provide an appropriate bidding strategy. Second, the designed algorithm needs to be an online one, which requires deciding a bid price once a user impression arrives, without the knowledge about the subsequent user impressions and future competition. The erratic fluctuation in dynamic environments further complicates the design of online bidding algorithms, making previous approaches mainly leveraging historical data [32] hard to adapt to the dynamic environments. Third, the newly introduced PV constraint in (4) makes it hard to fulfill all the constraints by simply focusing on budget pacing or PPC controlling [22, 33], and

it may make the optimization problem infeasible. That is, the ad delivery request from advertisers may not be fulfilled by the current advertising environments, and we should figure out infeasible instances before solving the optimization problems.

3.3 Optimal Bidding Strategy

We overcome the above challenges by converting the original problem to searching for the optimal dual variables. We first relax the online IP as an online LP problem, in which X are relaxed to be continuous variables, that is, $x_i \in [0, 1]$.

To solve the LP problem, we refer to the classical primal-dual method [12]. Let the dual variables of Budget, PPC, and PV constraints be λ_1 , λ_2 , and λ_3 , respectively, and y_i be the dual variable of $x_i \leq 1$. Then, the dual problem of LP can be defined as:

$$\begin{aligned} \min_{\lambda_1, \lambda_2, \lambda_3, y_i} \quad & \lambda_1 \times \text{Budget} + \sum_{i=1}^N y_i - \lambda_3 \times \text{PV} \\ \text{s.t.} \quad & \alpha \times \text{gmv}_i + \beta \times \text{ctr}_i - (\lambda_1 + \lambda_2) \times \text{cost}_i + \\ & \lambda_2 \times \text{PPC} \times \text{ctr}_i + \lambda_3 \leq y_i, \quad \forall i, \\ & \lambda_1, \lambda_2, \lambda_3, y_i \geq 0, \quad \forall i. \end{aligned} \quad (5)$$

For the optimal solution x^* in the primal problem, we denote the corresponding optimal solution in the dual problem as $\Lambda^* = \{\lambda_1^*, \lambda_2^*, \lambda_3^*\}$, and y^* . With this dual problem formulation, we can build a connection between dual variables and the bidding strategy. It can be shown that solving the user impression selection of the LP is (almost) equivalent to setting the following bidding strategy [2, 32]:

$$\text{bid}_i = \frac{\alpha \times \text{gmv}_i + \beta \times \text{ctr}_i + \lambda_2^* \times \text{PPC} \times \text{ctr}_i + \lambda_3^*}{(\lambda_1^* + \lambda_2^*) \times \text{ctr}_i}. \quad (6)$$

We illustrate the intuition behind the analysis here. On the one hand, according to the rule of GSP auction, the selection decision for user impression i under this bidding strategy is as follows:

$$x'_i = \begin{cases} 1, & \text{if } \text{bid}_i \times \text{ctr}_i > \text{cost}_i, \\ 0, & \text{Otherwise.} \end{cases} \quad (7)$$

On the other hand, according to the complementary slackness and the constraints in (5), we can have

$$\begin{aligned} \text{bid}_i \times \text{ctr}_i - \text{cost}_i - \frac{y_i^*}{\lambda_1^* + \lambda_2^*} &\leq 0, \quad (x_i^* - 1) \times y_i^* = 0, \\ x_i^* \times (\text{bid}_i \times \text{ctr}_i - \text{cost}_i - \frac{y_i^*}{\lambda_1^* + \lambda_2^*}) &= 0, \quad \lambda_1^*, \lambda_2^*, \lambda_3^*, y_i^* \geq 0, \quad \forall i. \end{aligned} \quad (8)$$

With these two aspects in (7) and (8), we can verify that the decision variable x'_i is equal to x_i^* as long as $\text{bid}_i \times \text{ctr}_i \neq \text{cost}_i$. In the dynamic advertising environment with a high volume of ad auctions, the proportion of instances with $\text{bid}_i \times \text{ctr}_i = \text{cost}_i$ is quite small, which implies that the user impression selection result from the bidding strategy in (6) is almost equivalent to the optimal solution of LP. Since the decision variables in the LP are integers as long as $\text{bid}_i \times \text{ctr}_i \neq \text{cost}_i$, we also have demonstrated the relaxation from IP to LP would not cause much distortion. Thus, once we have the optimal dual variables, we can derive the optimal bidding strategy by (6). In the following discussion, we design an efficient online algorithm to search for the optimal dual variables $\Lambda^* = \{\lambda_1^*, \lambda_2^*, \lambda_3^*\}$, from the perspectives of initialization and dynamic adjustment.

4 AD ENVIRONMENT REPRESENTATION

In this section, we introduce a deep neural network to represent the dynamic advertising environment, which is critical to both the initialization and dynamic adjustment of the dual variables in (6).

4.1 Network Design

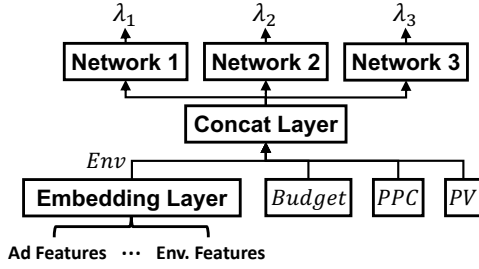
To calculate the dual variables, classical approaches need to collect historical LP instances and then conduct an offline LP. We abstract this process as a mapping from user impression information and constraint parameters to the optimal dual variables, which can be defined as $\Lambda^* = f(\text{Users}, \text{Budget}, \text{PPC}, \text{PV})$, where Users represents the predicted advertising performance for user impressions, including their predicted ctr , gmv , and cost , and Budget , PPC , and PV are the parameters of the performance constraints in (2), (3), and (4). However, there exist three difficulties in finding such a mapping function. First, the information of all user impressions is not available until the end of ad campaigns in the online setting. Second, it is time-consuming to solve the large offline LPs in (1) repeatedly, which involves thousands of decision variables in advertising environments. Third, the mapping function f could be quite complicated, and cannot be expressed as a closed-form expression.

To overcome the above challenges, we design a deep neural network to approximate the mapping function f in a data-driven manner. We introduce a representation layer to model the dynamic advertising environment, which can capture the unavailable user impression information in the future to some extent. This idea comes from the observations that we can regard each arriving user impression as a sample from the overall user impression distribution condition on the current ad's features, which is directly decided by the advertising environment. Compared with user impression information, advertising environment encodes much richer information, such as the information of ad auctions, the ad features like ad category, product price, and etc. Thus, we replace the unavailable user impression information with the environment information in the mapping function as $\Lambda^* = f(\text{Env}, \text{Budget}, \text{PPC}, \text{PV})$, where we transform high dimensional features into dense embedding, *i.e.*, $\text{Env} = \text{Embedding}(\text{ad features}, \text{env features} \dots)$.

We next illustrate this idea of approximation by designing a Λ -network and show the network architecture in Fig. 2. The Λ -network first concatenates the performance constraint parameters and the environment embedding, and then feeds them into three separate regression networks. The output of each regression network is the corresponding dual variables. When training Λ -network, the loss function is the mean square error between the outputted dual variables and the ground truth optimal dual variables Λ^* , and we update the network parameters by minimizing the average loss. It is worth noting that we can integrate a more general neural network to approximate the mapping function by leveraging more advertising features and using more complex network structures to enhance the approximation ability. The detailed design of more complicated neural networks is out of the scope of this work.

4.2 Applications

A direct application of the above Λ -network is to initialize the dual variables using the collected information about the advertising environment and the constraints. Another important application is


 Figure 2: The Λ -network.

that we can quantify the relation between the constraint parameters and the dual variables, which would be exploited for the dynamic adjustment of dual variables. Specifically, we can use the derivatives of the mapping function f , approximated by the gradient of the Λ -network, to represent the linear relation between the change of the constraint parameters and that of the dual variables:

$$\begin{pmatrix} d\lambda_1 \\ d\lambda_2 \\ d\lambda_3 \end{pmatrix} = J \cdot \begin{pmatrix} dBudget \\ dPPC \\ dPV \end{pmatrix},$$

where J is a 3×3 Jacobi matrix and can be derived from the gradient of the Λ -network with inputs Env , $Budget$, PPC , PV :

$$J = \begin{pmatrix} \frac{\partial f_{\lambda_1}}{\partial Budget} & \frac{\partial f_{\lambda_1}}{\partial PPC} & \frac{\partial f_{\lambda_1}}{\partial PV} \\ \frac{\partial f_{\lambda_2}}{\partial Budget} & \frac{\partial f_{\lambda_2}}{\partial PPC} & \frac{\partial f_{\lambda_2}}{\partial PV} \\ \frac{\partial f_{\lambda_3}}{\partial Budget} & \frac{\partial f_{\lambda_3}}{\partial PPC} & \frac{\partial f_{\lambda_3}}{\partial PV} \end{pmatrix}. \quad (9)$$

We discuss how to use such a relation for dynamically adjusting the dual variables in the next section.

We can also apply the Λ -network to evaluate the feasibility of the LP in advance, which is non-trivial due to the newly added PV constraint for guaranteed delivery. Based on duality theory [12], we can derive that the primal problem is infeasible if and only if the dual problem is unbounded. According to the optimization objective in (5), the dual problem is unbounded implies $\lambda_3^* \rightarrow \inf$, i.e., the value of PV reaches infinite. This can be reflected by certain patterns in the predicted dual variables outputted by the Λ -network, and we leverage Support Vector Machine (SVM) [28] to capture such patterns. In the simulation experiment where we generate 181,953 groups of constraints based on real-world ad campaign logs, our approach can achieve an accuracy of 85.9% and AUC of 89.4% for the feasibility validation task.

5 FEEDBACK CONTROL FOR DYNAMIC BIDDING

In this section, we design a feedback control algorithm under Model Predictive Control (MPC) [7] framework to timely adjust the bidding strategy in a dynamic environment.

5.1 Model Predictive Control Framework

As shown in Fig. 3, Model Predictive Control (MPC) is a feedback control framework that can handle the coupling effects of multiple inputs and outputs of a dynamic system. The goal of MPC is to keep the outputs (system feedback) close to the pre-defined references by

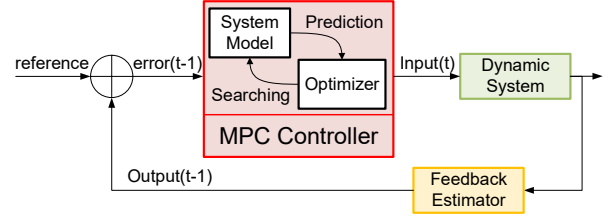


Figure 3: A Model Predictive Control framework.

searching for the best dynamic system inputs. A classic MPC contains two important components: a model for the dynamic system and an optimizer for inputs. By the model, we can predict the possible output for a given input to the dynamic system. The optimizer searches for the best inputs in the next time step by minimizing the error between the predicted output and the desired reference.

We next show the MPC framework for control-based bidding in advertising. We regard the whole advertising environment and bidding strategy as the dynamic system. The inputs of this dynamic system are the dual variables, and the outputs are the current advertising performance, such as the spent cost, average PPC, and acquired PV. The goal of the MPC is to ensure the total cost, the average PPC, and the accumulated PV at the end of the campaign approach to the preset constraints by keeping the outputs close to the references in each time period. Specifically, suppose the duration of the ad campaign is MT with M time periods and each with T time slots. The input of the system in the period n is $\{\lambda_1^n, \lambda_2^n, \lambda_3^n\}$, which determines the bidding strategy and the system outputs, i.e., $Cost_n$, PPC_n and PV_n , indicating the total cost, averaged PPC and PV acquired within time period n . At the end of each time period, the MPC adjusts the dual variables according to the error between the (dynamic) references and performance feedback. A simple definition of the references for time period $[nT, (n+1)T)$ is:

$$Budget_n^{ref} = \frac{w_n}{M} Budget, \quad PPC_n^{ref} = PPC, \quad PV_n^{ref} = \frac{w_n}{M} PV,$$

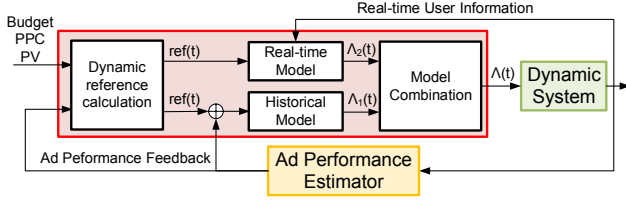
where $\mathbf{W} = \{w_0, w_1, \dots, w_{M-1}\}$ with $\sum_{i=0}^{M-1} w_i = M$ are weight factors describing the environment fluctuation in different time periods. By building the connection of input, output, and the references of the dynamic system in this way, the preset constraints can be satisfied through dynamic adjustment of the dual variables. However, one problem is still unclear: whether this MPC can maximize the advertiser's optimization objective? For this question, we give the following theorem. The detailed proof is displayed in the supplementary material [1].

THEOREM 5.1. For dual variables $\Lambda' = \{\lambda'_1, \lambda'_2, \lambda'_3\}$, suppose the corresponding ad allocation result for (Online IP) is $\mathbf{X}' = \{x'_1, x'_2, \dots, x'_N\}$ and the optimal solution is $\mathbf{X}^* = \{x_1^*, x_2^*, \dots, x_N^*\}$. The upper bound of the gap between the optimization objective achieved by \mathbf{X}' and \mathbf{X}^* is:

$$\sum_{i=1}^N (x_i^* - x'_i) \times (\alpha \times gm v_i + \beta \times ctr_i) \leq \lambda'_1 \times B + \lambda'_2 \times C + \lambda'_3 \times D,$$

$$\text{where } B = Budget - \sum_{i=1}^N x'_i cost_i, \quad C = PPC \sum_{i=1}^N x'_i ctr_i - \sum_{i=1}^N x'_i cost_i,$$

$$D = \sum_{i=1}^N x'_i - PV.$$


Figure 4: The control-based bidding algorithm.

According to Theorem 5.1, when we keep the total cost, the average PPC, and the accumulated PV close to the required constraint parameters, *i.e.*, B , C , and D vanishes to zero, the gap of the optimization objective would also disappear. We can conclude that the MPC feedback control framework can maximize the advertiser's objective and keep constraints in the feasible region simultaneously.

The major problem in the above MPC is to model the dynamic system and design the output optimizer. However, the dynamic advertising environment is too complicated to have a concise model as in traditional dynamic systems [32]: the bidding landscape from competitive advertisers is quite different in various ad auctions; user impressions experience erratic fluctuation in terms of quantity and quality. Without a precise model for the dynamic system, the result from the optimizer can be quite unstable, and the searching process is time-consuming. We next propose an end-to-end control-based bidding algorithm, combining the dynamic system model and the output optimizer together.

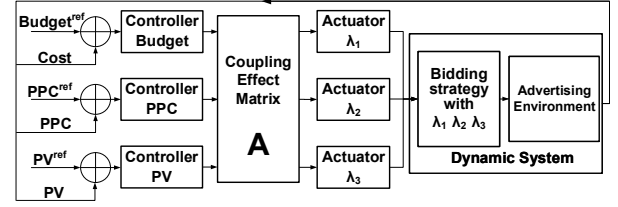
5.2 A Control-based Bidding Algorithm

The complete process of the proposed control-based bidding algorithm is shown in Fig. 4, where the main algorithm components are highlighted in the red box, including dynamic reference calculation, historical model, real-time model, and model combination. Designing this algorithm only using historical experiences can cause huge performance degradation in dynamic environments. While simply relying on real-time information may cause a severe cold-start problem and suffers from incorrect dynamic references due to the incompleteness of data. Thus, we design a *historical model* as well as a *real-time model*, and combine them in an appropriate way.

Historical Model with Multi-variable Feedback Control: As shown in Fig. 5, the historical model is based on the multi-variable feedback control [24], which decouples multi-inputs (error between ad performance feedback and references) and multi-outputs (the adjustments to the dual variables) by using the relation between the change of dual variables and the fluctuation of constraint parameters, captured by the Λ -network. Similar to Section 4.2, when the change of constraint parameters (error) $\Delta Budget$, ΔPPC , ΔPV are small, we can use a linear model to approximate this relation as shown in (10). Since the control signals are generated based on the linear combination of error, the change of constraint parameters can be transformed to the control signals of dual variables in (11).

$$\begin{pmatrix} \Delta \lambda_1 \\ \Delta \lambda_2 \\ \Delta \lambda_3 \end{pmatrix} = A \cdot \begin{pmatrix} \Delta Budget \\ \Delta PPC \\ \Delta PV \end{pmatrix}. \quad (10)$$

$$\begin{pmatrix} u_{\lambda_1}(t) \\ u_{\lambda_2}(t) \\ u_{\lambda_3}(t) \end{pmatrix} = A \cdot \begin{pmatrix} u_{budget}(t) \\ u_{ppc}(t) \\ u_{pv}(t) \end{pmatrix}. \quad (11)$$


Figure 5: The Historical Model.

Instead of fixing the coupling effect matrix A , we use the Jacobi matrix in (9) from Λ -network to represent the coupling effect¹.

We next set the controller, reference, and the actuator to generate the control signals of constraint parameters, *i.e.*, u_{budget} , u_{ppc} , u_{pv} . Different from the classic PID controller [4], we require the controller to be simple and robust to keep the basic performance as we have another real-time model for further adjustment. Thus, we use Water-Level (WL) controller [11] with the control function: $u(t) = u(t-1) + \gamma \times e(t)$, where the error in the n -th time period can be calculated as²:

$$\begin{aligned} e_{budget} &= (Budget_n^{ref} - Cost_n) \times \frac{M}{w_n}, \\ e_{ppc} &= PPC_n^{ref} - PPC_n, \\ e_{pv} &= (PV_n^{ref} - PV_n) \times \frac{M}{w_n}. \end{aligned} \quad (12)$$

It is worth noting that we rescale the error from the single period with T time slots to the entire advertising process with a total of MT time slots. This is because the gradient of Λ -network approximates the relation between dual variables and the constraint parameters over the whole advertising period. Due to the lack of the integral controller, WL controller can not well handle the accumulative error [33]. We further enhance the control capability of WL controller by using dynamic references instead of fixed references:

$$\begin{aligned} Budget_n^{ref} &= \frac{w_{n+1} \times \max(0, Budget - \sum_{k=0}^n Cost_k)}{M - \sum_{k=0}^n w_k} \times \frac{w_n}{w_{n+1}}, \\ PPC_n^{ref} &= \frac{\max(0, M \times PPC - \sum_{k=0}^n w_k PPC_k)}{M - \sum_{k=0}^n w_k}, \\ PV_n^{ref} &= \frac{w_{n+1} \times \max(0, PV - \sum_{k=0}^n PV_k)}{M - \sum_{k=0}^n w_k} \times \frac{w_n}{w_{n+1}}. \end{aligned} \quad (13)$$

The above dynamic references are the cost, PPC, and PV we need to achieve in the next time period if we want the overall ad performance to satisfy the desired ones based on the realized performance so far, and we rescale them from time period $(n+1)$ to n . The actuator we use in WL controller is linear-form: $\lambda((n+1)T) = \lambda(0) + u_{\lambda}(nT)$.

Real-time Model with Periodic Linear Programming: We introduce a real-time model to leverage the fine-grained online information, enabling the bidding algorithm to adapt to the current ongoing environment rapidly. At time slot t , the advertising platform can collect the advertising instances during a period $[t - T_0, t)$ with an observing interval T_0 . We denote the user impressions in period $[t - T_0, t)$ as a set $U(t, T_0)$. We use this information to predict

¹The main advantage is that the Λ -network encodes the knowledge learned from historical instances, and can be generalized to calculate the matrix A .

²We use e_{budget} , e_{ppc} and e_{pv} instead of $\Delta Budget$, ΔPPC , and ΔPV , to represent the error of the constraint parameters as in literature.

the optimal dual variables in the following observation period. The key assumption behind this is that the advertising environment is relatively stable during a short period, meaning that the user impressions arriving in $[t, t + T_0]$ are likely to follow the identical and independent distributions with $U(t, T_0)$. We can then combine the current references with the user impressions in $U(t, T_0)$ to calculate the optimal dual variables through offline linear programming. We use the dynamic references defined in (13) with rescaling as the constraints in this offline linear programming. The offline linear programming is shown as follows:

$$\begin{aligned}
 & \max_{\mathbf{X}} \sum_{i \in U(t, T_0)} x_i \times (\alpha \times gmv_i + \beta \times ctr_i) \\
 \text{s.t.} \quad & \sum_{i \in U(t, T_0)} x_i \times cost_i \leq Budget_n^{ref} \times \frac{\sum_{kT \in [t-T_0, t]} w_k}{w_n}, \\
 & \frac{\sum_{i \in U(t, T_0)} x_i \times cost_i}{\sum_{i \in U(t, T_0)} x_i \times ctr_i} \leq PPC_n^{ref}, \\
 & \sum_{i \in U(t, T_0)} x_i \geq PV_n^{ref} \times \frac{\sum_{kT \in [t-T_0, t]} w_k}{w_n}, \\
 & x_i \in [0, 1], \forall i \in U(t, T_0).
 \end{aligned} \tag{14}$$

Although the real-time adaption model can capture the current advertising environment better than the historical model, it suffers from the cold start problem and the incompleteness of data. We next combine the historical model and the real-time model to improve the overall performance.

Combination of Historical model and Real-time model:

At the time slot t , we denote the dual variables of the historical and real-time model as $\Lambda_1(t)$ and $\Lambda_2(t)$, respectively. The final output of the combined model can be formulated as: $\Lambda(t) = \tau(t) \times \Lambda_1(t) + (1 - \tau(t)) \times \Lambda_2(t)$, where $\tau(t)$ is the weight factor and $\tau(t) = 1 - \min(k(\frac{t}{MT}), p)$, where $p \in [0, 1]$ and $k \geq p$. The weight factor $\tau(t)$ is close to 1 at the beginning and gradually reduces to $1 - p$ as t increases. The intuition behind this $\tau(t)$ is that we pay more attention to the historical model at the beginning due to the cold start problem. As time evolves, the advertising environment becomes relatively stable, and we rely more on the real-time model in the later stage. The parameter k controls the speed of shifting the attention from the historical model to the real-time model and is decided based on the severity of the cold start problem. The parameter p reflects the quality and the credibility of the collected real-time information, which should be small when the incompleteness of data is severe, or the latency to collect the data is large.

6 EXPERIMENT RESULTS

6.1 Dataset and Metrics

The experiments are based on an industry dataset from Mobile Taobao, the largest e-commerce platform in China. The dataset consists of 94 ad campaigns within six sequential days with a total of 7.4 GB ad logs. Each ad campaign contains hundreds to tens of thousands of ad logs, each of which represents an advertising opportunity, including the user impression arrival time, predicted ad KPIs and expected cost. We group these ad logs by their arrival time and the ad campaign they belong to. As livestreaming ads usually last for a short period of time, we set the advertising duration to

be an hour in the following experiments and generate constraint parameters based on the statistic user impression distribution to avoid infeasible or invalid constraints. We divide the ad logs in the first five days as the training data and the last day as the test data.

From this data set, we observe that the ad environment is significantly dynamic in practice. The difference in CTR and CVR between different days reaches over 15%, and the fluctuation in user impression volume can even exceed 450%. The user impression volume and ad KPIs can also change greatly at different time periods within the same day: compared with the early morning between 4 am and 6 am, the user impression volume between 9 pm and 11 pm is 5 times larger and the average CVR increases by 70%. These results demonstrate the necessity of a proper environment representation and adaptive bidding algorithms³.

The evaluation metrics can be divided into three categories:

- 1) Gap between the achieved marketing objective and the optimal offline solution (refer as M.O. for short).
- 2) Gap between the obtained ad KPIs and the preset constraint parameters, measuring the ability to keep the bidding algorithm in line with dynamic environment⁴.
- 3) Constraint satisfaction ratio indicates the proportion of samples with the three constraints satisfied (refer as Ratio for short). Due to the difficulty in satisfying the three constraints rigorously, we allow a 10% drift for these constraints.

6.2 Control Frequency Selection

The frequency of feedback control is critical for the trade-off between the control capability and the efficiency of online systems. We evaluate the impact of feedback control frequency on the performance of the historical model, and show the result in Fig. 6(a). We can observe that a high control frequency reacts quickly against environment fluctuation and indeed improves the performance, increasing the constraint satisfaction ratio from 65% to 78%. Such improvement diminishes with the increase of frequency, and reaches a bottleneck when the frequency exceeds 12 times per hour, which is then chosen for the control frequency of the historical model.

For the real-time model, its control frequency should be smaller than the historical model since the offline LP on large-scale data is much more complicated than the multi-variable feedback control. The real-time model needs detailed information about all the arrived user impressions, which is always incomplete due to data loss and data delay in practice. We model the probability that the user impression's information is not available at time T as a time-varying function $L(T - T_a)$, where T_a is the user impression's arrival time and $L(t) = p_{loss} + p_{delay}(t)$. Here p_{loss} is the probability that the data is lost during transmission, and $p_{delay}(t)$ is the probability that the data is temporarily delayed. We set $p_{loss} = 5\%$ and $p_{delay}(t) = \max(0, 30\% - 3\% \times t)$ in the experiment. The experiment result under different control frequencies is shown in Fig. 6(b). We observe that

³In the following experiments, the dynamic ad environments mainly refer to various user impression volumes and ad performance indicators for different ad campaigns on different days/hours. The user impression volume within a short period, e.g., one hour, is relatively stable.

⁴The gap between performance A and preset constraint B is evaluated as $\max(B - A, 0)/B$, and we calculate the average gap over all the test samples. The advertising process will terminate when the budget is exhausted. Since spending too much money on cheap user impressions can harm the overall ad performance, we consider both insufficient and abundant parts when calculating the gap of PPC, which is $\frac{|ppc' - PPC|}{PPC}$.

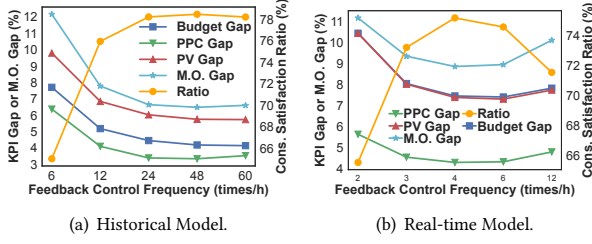


Figure 6: Control frequency analysis.

Table 1: The ablation study of Historical Model.

Exp. Index	Budget	PPC	PV	M.O.	Ratio
1	29.76%	10.50%	28.76%	23.18%	22.94%
2	24.98%	9.09%	24.35%	19.85%	27.41%
3	5.66%	42.66%	8.19%	14.03%	7.22%
4	9.56%	5.05%	10.46%	9.25%	60.49%
5	10.89%	4.85%	11.24%	9.45%	58.75%
6	4.79%	6.25%	6.25%	9.91%	75.88%
Our Alg.	5.17%	4.10%	6.83%	7.78%	75.92%

the low control frequency lacks enough control capability, while the high control frequency can cause severe data delay, increasing the average L from 10% to 27.5%. Considering the control capability and computation resources, we set the frequency of the real-time model to be 4 times per hour with the average L to be 15%.

6.3 Ablation Study

We evaluate the effectiveness of different components in our control-based bidding algorithm through ablation study.

Historical Model. We conduct the ablation experiments of the historical model from high-level components to fine-grained components. We consider the following six ablation experiments:

- 1) Remove the feedback control system and fix the dual variables generated by Λ -network during the whole advertising process.
- 2) Replace the feedback control system with the periodic prediction of Λ -network by applying the updated dynamic references with rescaling as the network input at the end of each time period to re-generate the dual variables.
- 3) Remove the advertising environment features in Λ -network.
- 4) Remove the dynamic references of WL controller.
- 5) Replace the WL controller with the PID controller with hyper-parameters decided through grid-search on training data.
- 6) Replace the feedback control with three independent WL controllers. That is, setting all the elements of the coupling effect matrix A except the main diagonal to 0.

The experiment results are shown in Tab. 1, and we can get the following observations/insights for each of the above experiments:

- 1) It is necessary to adjust the bidding algorithm in line with the dynamic environment. The dual variables initialized by the Λ -network cannot handle the fluctuation of the advertising environment, increasing the PV gap by 21.93%.
- 2) Although periodically adjusting the dual variables through the Λ -network prediction can improve the performance compared with no adjustment, the dual variables are not very precise, leading

Table 2: The ablation study of Real-time Model.

Exp. Index	Budget	PPC	PV	M.O.	Ratio
1	23.53%	10.61%	22.81%	23.62%	55.11%
2	7.54%	4.33%	6.93%	9.48%	76.12%
Our Alg.	7.44%	4.28%	7.37%	8.84%	75.15%
Dynamic-7.5%	4.79%	3.11%	4.65%	6.70%	85.10%
Preset-7.5%	6.01%	3.20%	5.12%	7.35%	81.44%

to a huge degradation in the final ad performance, satisfying only 27.41% of the advertisers' constraints.

3) The representation of the dynamic environment is critical for the algorithm performance, and the lack of advertising environment information in the Λ -network would cause huge degradation in control capability, making the gap of PPC reach 42.66%.

4) Using WL controller alone would cause accumulative error, and applying dynamic references can greatly improve its control capability, decreasing the gap of PV from 10.46% to 6.83%.

5) The performance of PID controller is far worse than the WL controller especially in PV delivery (11.24%) and constraint satisfaction ratio (58.75%). However, it is worth noting that these two controllers have similar performance in the training set. As PID controller is more powerful with more parameters to decide, it is easier to be misled by the environment in the training set and cause "over-fitting". In contrast, WL controller is simpler and is more robust to the dynamic environment.

6) Although the constraint satisfaction ratio in the algorithm without considering the coupling effect is close to our algorithm in this dataset, we find our algorithm can improve the ability in estimating user impression prices and optimizing marketing objectives, decreasing the gap of PPC and marketing objective by 2%.

From the above discussion, we can summarize that the coupling effect matrix generated based on the Λ -network as well as the WL controller and the dynamic references can improve the control capability of the historical model.

Real-time Model. We next conduct the ablation experiments of the real-time model, and show the results in Tab. 2.

1) Replace the Λ -network-based dual variable initialization with the results of offline LP on historical data in the training set.

2) Replace the dynamic references with the preset constraints when solving periodic offline LP throughout the advertising process.

From the results, we can also find that:

1) The real-time model suffers from the cold start problem, and its performance drops significantly without a proper initialization method, achieving 77.19% of the PVs.

2) Our dynamic references do not achieve significant achievement compared with using the preset constraints. This phenomenon mainly comes from incomplete user impression information. To prove this, we conduct two additional experiments: Dynamic-7.5% and Preset-7.5%, which show the results when the average data incompleteness probability L is improved from 15% to 7.5%. When the incompleteness of data is less severe, the dynamic references can achieve obvious performance improvements, e.g., increasing the constraint satisfaction ratio by 3.66%.

Combined Model Finally, we compare the performance of our combined model with 1) Using the historical model alone. 2) Using the real-time model alone. We set the weight factor $\tau(t)$ as

Table 3: The ablation study of Combined Model.

Model Type	Budget	PPC	PV	M.O.	Ratio
Historical Model	5.17%	4.10%	6.83%	7.78%	75.92%
Real-time Model	7.44%	4.28%	7.37%	8.84%	75.15%
Combined Model	3.46%	3.39%	4.25%	6.97%	85.76%

$1 - \min(2(\frac{t}{MT}), 0.9)$ due to the problems of data delay and data loss. The experiment results in Tab. 3 prove the combined model outperforms only applying the historical model or the real-time model, narrowing the gap of PV by 2.58% and increasing the constraint satisfaction ratio by 9.84%. The results demonstrate the necessity of making full use of historical knowledge as well as adapting to the current environment based on real-time information.

6.4 Comparative study

We compare our algorithm with the approaches from operation research literature (**DLA**, **LPviaLD**) and the bid optimization with multiple constraints (**Cost-min**, **FB-control**, and **M-PID**).

- **DLA** [2]: Dynamic Learning Algorithm (**DLA**) provides bids as in (6) based on periodic LP without any historical knowledge. It uses ad opportunities in the initial period to observe the advertising environment without bidding. The dual variables are initialized through solving offline LP based on the observed information, and would be updated periodically when the new information is collected.

- **LPviaLD** [17]: **LPviaLD** algorithm transforms the online LP into a load balancing problem, which considers balancing three constraints and the objective. This LD problem can be solved through the expert algorithm with Multiplicative Weights Update (MWU) [3], and the bidding strategy can be derived from the weights of experts.

- **Cost-min** [22]: **Cost-min** is a generic algorithm that can address multiple constraints, whose bidding strategy is set as $bid_i = b_0 \times (\alpha \times gmv_i + \beta \times ctr_i)$, and the upper bound of bid_i is set to be PPC . b_0 is initialized by dividing the given PPC by the average value of $(\alpha \times gmv_i + \beta \times ctr_i)$ and would be dynamically adjusted by a feedback control system with three independent PID controllers corresponding to three constraints.

- **FB-control** [33]: **FB-control** is first proposed to control PPC . Its bidding strategy is defined as $bid_i = b_0$, where b_0 is initialized by the preset constraint PPC . We apply a feedback control system with three independent PID controllers to control b_0 .

- **History LP** [32]: This method has the same bidding strategy as in (6), and it initializes the dual variables by conducting offline LP on the historical data that belongs to the same ad campaign. It has no further adjustment for the dual variables during advertising.

- **M-PID** [32]: **M-PID** has the same bidding strategy as in (6), and the initialization is the same as **History LP**. It further applies multi-variable feedback control with PID controllers to adjust the dual variables, and the coupling matrix A is determined based on the training set. To avoid the influence of the dynamic environment, we set the average A of the training set as the coupling effect matrix.

The experiment results are illustrated in Tab. 4. We next summarize the performance of the above bidding algorithms.

For **DLA**, it achieves excellent control in PPC with a gap of 2.03% while its performance in gaining PV is not as good, having a gap of 11.93%. Although the full user impression information in

Table 4: The results of comparative study.

Algorithm	Budget	PPC	PV	M.O.	Ratio
DLA	12.18%	2.03%	11.93%	12.56%	50.77%
LPviaLD	7.18%	8.46%	6.37%	27.13%	44.34%
Cost-min	44.52%	23.08%	47.79%	16.09%	0.03%
FB-control	5.90%	8.70%	11.15%	3.27%	40.21%
History LP	71.02%	21.14%	70.26%	58.05%	3.32%
M-PID	55.55%	17.41%	54.62%	47.65%	9.81%
Our Alg.	3.46%	3.39%	4.25%	6.97%	85.76%

the initial period can reflect the bidding landscape of the ongoing environment, losing these ad opportunities is not affordable in performance optimization, which shows the necessity of leveraging historical knowledge to initialize the bidding strategy.

For **LPviaLD**, it is not as stable as our algorithm and is especially poor in optimizing the marketing objective, reaching the gap of 27.13%, since **LPviaLD** can hardly match Theorem 5.1. After further experiments, we find the low parameter update frequency also weakens its performance.

For **Cost-min**, its PPC is significantly lower than the preset constraint, and thus can hardly gain enough PV, resulting in the constraint satisfaction ratio lower than 1%. Keeping cost-efficiency simply by giving up all expensive user impressions is not applicable after considering PV constraints. **FB-control** can achieve good performance in optimizing marketing objectives while it has the same problem in achieving PV, with the gap of 11.15%. The main reason is that its bidding strategy only considers the value of clicks, without the cost performance in gaining ad exposure.

For **History LP** and **M-PID**, initializing and adjusting the bidding strategy simply based on the historical data without adapting to the current environment would cause huge performance degradation, making the gap of PV exceeds more than 50%.

To sum up, our control-based bidding algorithm outperforms the other methods in both optimizing marketing objectives and satisfying constraints. Our approach can achieve 95.75% of the ad exposure and 93.03% of the marketing objective of the optimal offline solution on average. The overall constraint satisfaction ratio can exceed 85.76% with the average Budget and PPC Gap limited to 3.46% and 3.39%, respectively.

7 CONCLUSION

In this work, we have studied the problem of generalized ad guaranteed delivery for mobile livestreaming ads in dynamic environments, maximizing various marketing objectives under multiple constraints over budget, PPC , and PV at the same time. We have designed a deep neural network, namely Λ -network, to approximate the process of solving the dual variables. To further timely update the bidding strategy, we have proposed a control-based bidding algorithm under the MPC framework, which consists of a historical model and a real-time model. By conducting experiments on the industrial dataset, we have found our algorithm can handle the fluctuation of advertising environments, optimizing marketing objectives and keeping performance constraints.

REFERENCES

- [1] 2022. Supplementary Material. <https://drive.google.com/drive/folders/1Jk7EhrzHPiOgyii3ArfbmuGryKxfZHIE?usp=sharing>.
- [2] Shipra Agrawal, Zizhuo Wang, and Yinyu Ye. 2014. A dynamic near-optimal algorithm for online linear programming. *Operations Research* 62, 4 (2014), 876–890.
- [3] Sanjeev Arora, Elad Hazan, and Satyen Kale. 2012. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing* 8, 1 (2012), 121–164.
- [4] Stuart Bennett. 1993. Development of the PID controller. *IEEE Control Systems Magazine* 13, 6 (1993), 58–62.
- [5] Vijay Bharadwaj, Peiji Chen, Wenjing Ma, Chandrashekar Nagarajan, John Tomlin, Sergei Vassilvitskii, Erik Vee, and Jian Yang. 2012. Shale: an efficient algorithm for allocation of guaranteed display advertising. In *SIGKDD*. 1195–1203.
- [6] Han Cai, Kan Ren, Weinan Zhang, Kleantlis Malialis, Jun Wang, Yong Yu, and Defeng Guo. 2017. Real-time bidding by reinforcement learning in display advertising. In *WSDM*. 661–670.
- [7] Eduardo F Camacho and Carlos Bordons Alba. 2013. *Model predictive control*. Springer Science & Business Media.
- [8] Peiji Chen, Wenjing Ma, Srinath Mandalapu, Chandrashekar Nagarjan, Jayavel Shammugasundaram, Sergei Vassilvitskii, Erik Vee, Manfai Yu, and Jason Zien. 2012. Ad serving using a compact allocation plan. In *EC*. 319–336.
- [9] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishii Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In *Workshop on DLRS*. 7–10.
- [10] Nikhil R Devanur and Thomas P Hayes. 2009. The adwords problem: online keyword matching with budgeted bidders under random permutations. In *EC*. 71–78.
- [11] Bernard W Dezotell. 1936. Water level controller. US Patent 2,043,530.
- [12] W Erwin Diewert. 1974. Applications of duality theory. (1974).
- [13] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review* 97, 1 (2007), 242–259.
- [14] Zhen Fang, Yang Li, Chuanren Liu, Wenxiang Zhu, Yu Zheng, and Wenjun Zhou. 2019. Large-Scale Personalized Delivery for Guaranteed Display Advertising with Real-Time Pacing. In *ICDM*. 190–199.
- [15] Jon Feldman, Aranyak Mehta, Vahab Mirrokni, and Shan Muthukrishnan. 2009. Online stochastic matching: Beating $1-1/e$. In *FOCS*. IEEE, 117–126.
- [16] Arpita Ghosh, Benjamin IP Rubinfeld, Sergei Vassilvitskii, and Martin Zinkevich. 2009. Adaptive bidding for display advertising. In *WWW*. 251–260.
- [17] Anupam Gupta and Marco Molinaro. 2016. How the experts algorithm can help solve LPs online. *Mathematics of Operations Research* 41, 4 (2016), 1404–1431.
- [18] Yue He, Xiujun Chen, Di Wu, Junwei Pan, Qing Tan, Chuan Yu, Jian Xu, and Xiaoqiang Zhu. 2021. A Unified Solution to Constrained Bidding in Online Display Advertising. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2993–3001.
- [19] Yu Hu, Jiwoong Shin, and Zhulei Tang. 2016. Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action. *Management Science* 62, 7 (2016), 2022–2038.
- [20] iiMedia Research. 2022. Annual Research Report on Chinese Online Live Streaming Industry in 2021.
- [21] Chinmay Karande, Aranyak Mehta, and Pushkar Tripathi. 2011. Online bipartite matching with unknown distributions. In *STOC*. 587–596.
- [22] Brendan Kitts, Michael Krishnan, Ishadutta Yadav, Yongbo Zeng, Garrett Badeau, Andrew Potter, Sergey Tolkachov, Ethan Thornburg, and Satyanarayana Reddy Janga. 2017. Ad Serving with Multiple KPIs. In *SIGKDD*. 1853–1861.
- [23] Takanori Maehara, Atsuhiko Narita, Jun Baba, and Takayuki Kawabata. 2018. Optimal Bidding Strategy for Brand Advertising. In *IJCAI*. 424–432.
- [24] James Blake Rawlings and David Q Mayne. 2009. *Model predictive control: Theory and design*. Nob Hill Pub.
- [25] Kan Ren, Weinan Zhang, Ke Chang, Yifei Rong, Yong Yu, and Jun Wang. 2017. Bidding machine: Learning to bid for directly optimizing profits in display advertising. *IEEE Transactions on Knowledge and Data Engineering* 30, 4 (2017), 645–659.
- [26] Coresight Research. 2021. Playbook: Livestreaming E-Commerce in the US.
- [27] Oliver J Rutz and Randolph E Bucklin. 2011. From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research* 48, 1 (2011), 87–102.
- [28] Johan AK Suykens and Joos Vandewalle. 1999. Least squares support vector machine classifiers. *Neural processing letters* 9, 3 (1999), 293–300.
- [29] Hilde AM Voorveld, Guda van Noort, Daniël G Muntinga, and Fred Bronner. 2018. Engagement with social media and social media advertising: The differentiating role of platform type. *Journal of advertising* 47, 1 (2018), 38–54.
- [30] Di Wu, Xiujun Chen, Xun Yang, Hao Wang, Qing Tan, Xiaoxun Zhang, Jian Xu, and Kun Gai. 2018. Budget constrained bidding by model-free reinforcement learning in display advertising. In *CIKM*. 1443–1451.
- [31] Hongxia Yang. 2017. Bayesian heteroscedastic matrix factorization for conversion rate prediction. In *CIKM*. 2407–2410.
- [32] Xun Yang, Yasong Li, Hao Wang, Di Wu, Qing Tan, Jian Xu, and Kun Gai. 2019. Bid optimization by multivariable control in display advertising. In *SIGKDD*. 1966–1974.
- [33] Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. 2016. Feedback control of real-time display advertising. In *WSDM*. 407–416.
- [34] Weinan Zhang, Shuai Yuan, and Jun Wang. 2014. Optimal real-time bidding for display advertising. In *SIGKDD*. 1077–1086.
- [35] Weinan Zhang, Ying Zhang, Bin Gao, Yong Yu, Xiaojie Yuan, and Tie-Yan Liu. 2012. Joint optimization of bid and budget allocation in sponsored search. In *SIGKDD*. 1177–1185.
- [36] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *SIGKDD*. 1059–1068.