

Feedback Control of Real-Time Display Advertising

Weinan Zhang¹, Yifei Rong^{2,1}, Jun Wang¹, Tianchi Zhu³, Xiaofan Wang⁴

¹University College London, ²YOYI Inc., ³Big Tree Times Co., ⁴Shanghai Jiao Tong University

¹{w.zhang, j.wang}@cs.ucl.ac.uk, ²yifei.rong@yozi.com.cn

³tc@bigtree.mobi, ⁴xfwang@sjtu.edu.cn

ABSTRACT

Real-Time Bidding (RTB) is revolutionising display advertising by facilitating per-impression auctions to buy ad impressions as they are being generated. Being able to use impression-level data, such as user cookies, encourages user behaviour targeting, and hence has significantly improved the effectiveness of ad campaigns. However, a fundamental drawback of RTB is its instability because the bid decision is made per impression and there are enormous fluctuations in campaigns' key performance indicators (KPIs). As such, advertisers face great difficulty in controlling their campaign performance against the associated costs. In this paper, we propose a feedback control mechanism for RTB which helps advertisers dynamically adjust the bids to effectively control the KPIs, e.g., the auction winning ratio and the effective cost per click. We further formulate an optimisation framework to show that the proposed feedback control mechanism also has the ability of optimising campaign performance. By settling the effective cost per click at an optimal reference value, the number of campaign's ad clicks can be maximised with the budget constraint. Our empirical study based on real-world data verifies the effectiveness and robustness of our RTB control system in various situations. The proposed feedback control mechanism has also been deployed on a commercial RTB platform and the online test has shown its success in generating controllable advertising performance.

Keywords

Feedback Control, Demand-Side Platform, Real-Time Bidding, Display Advertising

1. INTRODUCTION

Emerged in 2009, Real-Time Bidding (RTB) has become a new paradigm in display advertising [22, 12]. Different from the conventional human negotiation or pre-setting a fixed price for impressions, RTB creates an impression-level auction and enables advertisers to bid for individual impression through computer algorithms served by demand-side platforms (DSPs) [33]. The bid decision could depend on the evaluation of both the utility (e.g., the likelihood and economic value of an impression for generating click or conversion) and the cost (e.g., the actual paid price) of each ad

impression. More importantly, real-time information such as the specific user demographics, interest segments and various context information is leveraged to help the bidding algorithms evaluate each ad impression. With the real-time decision making mechanism, it is reported that RTB yields significantly higher return-on-investment (ROI) than other online advertising forms [31].

Despite the ability of delivering performance-driven advertising, RTB, unfortunately, results in high volatilities, measured by major Key Performance Indicators (KPIs), such as CPM (cost per mille), AWR (auction winning ratio), eCPC (effective cost per click) and CTR (click-through rate). To illustrate this, Figure 1 plots the four major KPIs over time for two example campaigns in a real-world RTB dataset. All four KPIs fluctuate heavily across the time under a widely-used bidding strategy [25]. Such instability causes advertisers ample difficulty in optimising and controlling the KPIs against their cost.

In this paper, we propose to employ feedback control theory [2] to solve the instability problem in RTB. Feedback controllers are widely used in various applications for maintaining dynamically changing variables at the predefined reference values. The application scenarios range from the plane direction control [23] to the robot artificial intelligence [26]. In our RTB scenario, the specific KPI value, depending on the requirements from the advertisers, is regarded as the variable we want to control with a pre-specified reference value. Our study focuses on two use cases. (i) For performance-driven advertising, we concern with the feedback control of the average cost on acquiring a click, measured by effective cost per click (eCPC). (ii) For branding based advertising, to ensure a certain high exposure of a campaign, we focus on the control of the ratio of winning the auctions for the targeted impressions, measured by auction winning ratio (AWR). More specifically, we take each of them as the control input signal and consider the gain (the adjustment value) of bid price as the control output signal for each incoming ad display opportunity (the bid request). We develop two controllers to test: the widely used proportional-integral-derivative (PID) controller [6] and the waterlevel-based (WL) controller [10]. We conduct large-scale experiments to test the feedback control performance with different settings of reference value and reference dynamics. Through the empirical study, we find that the PID and WL controllers are capable of controlling eCPC and AWR, while PID further provides a better control accuracy and robustness than WL.

Furthermore, we investigate whether the proposed feedback control can be employed for controllable bid optimisation. It is common that the performance of an ad campaign (e.g., eCPC) varies from different channels (e.g., ad exchanges, user geographic regions and PC/mobile devices)

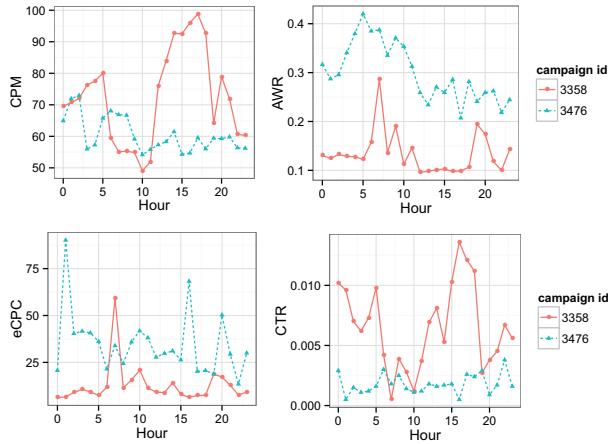


Figure 1: The instability of CPM (cost per mille), AWR (auction winning ratio), eCPC (effective cost per click), and CTR (click-through rate) for two sample campaigns without a controller. Dataset: iPinYou.

[34]. If one can reallocate some budget from less cost-effective channels to more cost-effective ones, the campaign-level performance would improve [35]. In this paper, we formulate the multi-channel bid optimisation problem and propose a model to calculate the optimal reference eCPC for each channel. Our experiments show that the campaign-level click number and eCPC achieve significant improvements with the same budget.

Moreover, the proposed feedback control mechanism has been implemented and integrated in a commercial DSP. The conducted live test shows that in a real and noisy setting the proposed feedback mechanism has the ability to produce controllable advertising performance.

To sum up, the contributions of our work are as follows. (i) We study the instability problem in RTB and investigate its solution by leveraging the feedback control mechanism. (ii) Comprehensive offline and online experiments show that PID controller is better than other alternatives and finds the optimal way to settle the variable in almost all studied cases. (iii) We further discover that feedback controllers are of great potential to perform bid optimisation through settling the eCPC at the reference value calculated by our proposed mathematical click maximisation framework.

The rest of this paper is organised as follows. Section 2 provides preliminaries for RTB and feedback control. Our solution is formally presented in Section 3. The empirical study is reported in Section 4, while the online deployment and live test are given in Section 5. Section 6 discusses the related work and we finally conclude this paper in Section 7.

2. PRELIMINARIES

To make the paper self-contained, in this section, we take a brief review on the RTB eco-system, bidding strategies, and some basics of feedback control theory.

2.1 RTB Flow Steps

The interaction process among the main components of the RTB eco-system is summarised into the following steps: (0) when a user visits an ad-supported site (e.g., web pages, streaming videos and mobile apps), each ad placement will trigger a call for ad (ad request) to the ad exchange. (1) The ad exchange sends the bid requests for this particular ad impression to each advertiser's DSP bidding agent, along with the available information such as the user and context information. (2) With the information of the bid request

and each of its qualified ads, the bidding agent calculates a bid price. Then the bid response (ad, bid price) is sent back to the exchange. (3) Having received the bid responses from the advertisers, the ad exchange hosts an auction and picks the ad with the highest bid as the auction winner. (4) Then the winner is notified of the auction winning from the ad exchange. (5) Finally, the winner's ad will be shown to the visitor along with the regular content of the publisher's site. It is commonly known that a long time page-loading would greatly reduce users' satisfactory [22]. Thus, advertiser bidding agents are usually required to return a bid in a very short time frame (e.g., 100 ms). (6) The user's feedback (e.g., click and conversion) on the displayed ad is tracked and finally sent back to the winner advertiser. For a detailed discussion about RTB eco-systems, we refer to [34, 31]. The above interaction steps have the corresponding positions in Figure 2, as we will discuss later.

2.2 Bidding Strategies

A basic problem for DSP bidding agents is to figure out how much to bid for an incoming bid request. The bid decision depends on two factors for each ad impression: the utility (e.g., CTR, expected revenue) and cost (i.e., expected charged price) [33]. In a widely adopted bidding strategy [25], the utility is evaluated by CTR estimation while the base bid price is tuned based on the bid landscape [9] for the cost evaluation. The generalised bidding strategy in [25] is

$$b(t) = b_0 \frac{\theta_t}{\theta_0}, \quad (1)$$

where θ_t is the estimated CTR for the bid request at moment t ; θ_0 is the average CTR under a target condition (e.g., a user interest segment); and b_0 is the tuned base bid price for the target condition. In this work, we adopt this widely used bidding strategy and adopt a logistic CTR estimator [27].

2.3 Feedback Control Theory

Feedback control theory deals with the reaction and control of dynamic systems from feedback and outside noise [2]. The usual objective of feedback control theory is to control a dynamic system so that the system output follows a desired control signal, called the reference, which may be a fixed or changing value. To attain this objective, a controller is designed to monitor the output and compares it with the reference. The difference between actual and desired output, called the error factor, is applied as feedback from the dynamic system to the control system. With the specific control function, the controller outputs the control signal, which is then transformed by the actuator into the system input signal sent back to the dynamic system. These processes form a feedback control loop between the dynamic system and the controller. Control techniques are widely used in various engineering applications for maintaining some signals at the predefined or changing reference values, such as plane navigation [23] and water distribution control [10].

3. RTB FEEDBACK CONTROL SYSTEM

Figure 2 presents the diagram of the proposed RTB feedback control system. The traditional bidding strategy is represented as the *bid calculator* module in the DSP bidding agent. The controller plays as a role which adjusts the bid price from the bid calculator.

Specifically, the *monitor* receives the auction win notice from the ad exchange and the user click feedback from the ad tracking system, which as a whole we regard as the dynamic system. Then the current KPI values, such as AWR and

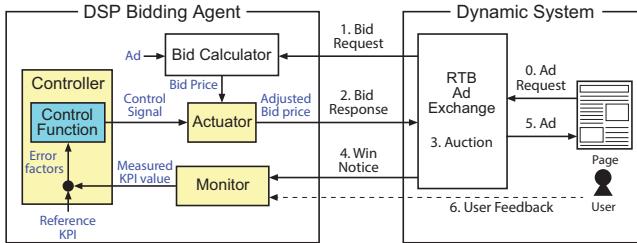


Figure 2: Feedback controller integrated in the RTB system.

eCPC can be calculated. If the task is to control the eCPC with the reference value, the error factor between the reference eCPC and the measured eCPC is calculated then sent into the control function. The output control signal is sent to the *actuator*, which uses the control signal to adjust the original bid price from the bid calculator. The adjusted bid price is packaged with the qualified ad into the bid response and sent back to the ad exchange for auction.

3.1 Actuator

For the bid request at the moment t , the actuator takes into the current control signal $\phi(t)$ to adjust the bid price from $b(t)$ (Eq. (1)) to a new value $b_a(t)$. In our model, the control signal, which will be mathematically defined in the next subsections, is a gain on the bid price. Generally, when the control signal $\phi(t)$ is zero, there should be no bid adjustment. There could be different actuator models, and in our work we choose to use

$$b_a(t) = b(t) \exp\{\phi(t)\}, \quad (2)$$

where the model satisfies $b_a(t) = b(t)$ when $\phi(t) = 0$. Other models such as the linear model $b_a(t) \equiv b(t)(1 + \phi(t))$ are also investigated in our study but it performs poorly in the situations when a big negative control signal is sent to the actuator, where the linear actuator will usually respond a negative or a zero bid, which is meaningless in our scenario. By contrast, the exponential model is a suitable solution to addressing the above drawback because it naturally avoids generating a negative bid. In the later empirical study we mainly report the analysis based on the exponential-form actuator model.

3.2 PID Controller

The first controller we investigate is the classic PID controller [6]. As its name implies, a PID controller produces the control signal from a linear combination of the proportional factor, the integral factor and the derivative factor based on the error factor:

$$e(t_k) = x_r - x(t_k), \quad (3)$$

$$\phi(t_{k+1}) \leftarrow \lambda_P e(t_k) + \lambda_I \sum_{j=1}^k e(t_j) \Delta t_j + \lambda_D \frac{\Delta e(t_k)}{\Delta t_k}, \quad (4)$$

where the error factor $e(t_k)$ is the reference value x_r minus the current controlled variable value $x(t_k)$, the update time interval is given as $\Delta t_j = t_j - t_{j-1}$, the change of error factors is $\Delta e(t_k) = e(t_k) - e(t_{k-1})$, and λ_P , λ_I , λ_D are the weight parameters for each control factor. Note that here the control factors are all in discrete time (t_1, t_2, \dots) because bidding events are discrete and it is practical to periodically update the control factors. All control factors $(\phi(t), e(t_k), \lambda_P, \lambda_I, \lambda_D)$ remain the same between two updates. Thus for all time t between t_k and t_{k+1} , the control signal $\phi(t)$ in Eq. (2) equals $\phi(t_k)$. We see that P factor tends to push the current variable value to the reference value; I

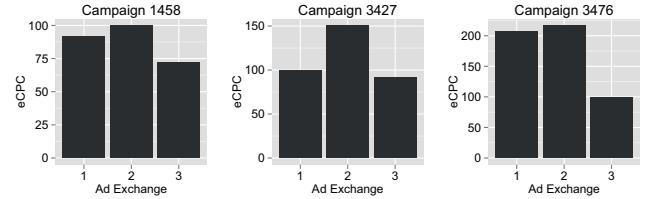


Figure 3: Different eCPCs across different ad exchanges. Dataset: iPinYou.

factor reduces the accumulative error from the beginning to the current time; D factor controls the fluctuation of the variable.

3.3 Waterlevel-based Controller

The Waterlevel-based (WL) controller is another feedback control model which was originally used to switching devices controlled by water level [10]:

$$\phi(t_{k+1}) \leftarrow \phi(t_k) + \gamma(x_r - x(t_k)), \quad (5)$$

where γ is the step size parameter for $\phi(t_k)$ update in exponential scale.

Compared to PID, the WL controller only takes the difference between the variable value and the reference value into consideration. Moreover, it provides a sequential control signal. That is, the next control signal is an adjustment based on the previous one.

3.4 Setting References for Click Maximisation

Given that the feedback controller is an effective tool to deliver advertisers' KPI goal, in this subsection, we demonstrate that the feedback control mechanism can be leveraged as a model-free click maximisation framework embedded with any bidding strategies [25, 33] and performs automatic budget allocation [17] across different channels via setting smart reference values.

When an advertiser specifies the targeted audience (usually also combined with ad impression contextual categories) for their specific campaign, the impressions that fit the target rules may come from separate channels such as different ad exchanges, user regions, users' PC/mobile devices etc. It is common that the DSP integrates with several ad exchanges and delivers the required ad impressions from all those ad exchanges (as long as the impressions fit the target rule), although the market prices [1] may be significantly different. Figure 3 illustrates that, for the same campaign, there is a difference in terms of eCPC across different ad exchanges. As pointed out in [34], the differences are also found in other channels such as user regions and devices.

The cost differences provide advertisers a further opportunity to optimise their campaign performance based on eCPCs. To see this, suppose a DSP is integrated to two ad exchanges A and B. For a campaign in this DSP, if its eCPC from exchange A is higher than that from exchange B, which means the inventories from exchange B is more cost effective than those from exchange A, then by reallocating some budget from exchange A to B will potentially reduce the overall eCPC of this campaign. Practically the budget reallocation can be done by reducing the bids for exchange A while increasing the bids for exchange B. Here we formally propose a model of calculating the equilibrium eCPC of each ad exchange, which will be used as the optimal reference eCPC for the feedback control that leads to a maximum number of clicks given the budget constraint.

Mathematically, suppose for a given ad campaign, there are n ad exchanges (could be other channels), i.e., $1, 2, \dots, n$, that have the ad volume for a target rule. In our formula-

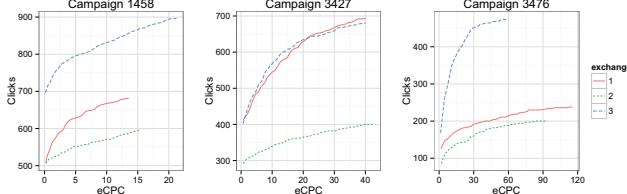


Figure 4: Number of Clicks against eCPC. Clicks and eCPC are calculated across the whole iPinYou training dataset of each campaign by tuning b_0 in Eq. (1).

tion we focus on optimising clicks, while the formulation of conversions can be obtained similarly. Let ξ_i be the eCPC on ad exchange i , and $c_i(\xi_i)$ be the click number that the campaign acquires in the campaign’s lifetime if we tune the bid price to make its eCPC be ξ_i for ad exchange i . For advertisers, they want to maximise the campaign-level click number given the campaign budget B [33]:

$$\max_{\xi_1, \dots, \xi_n} \sum_i c_i(\xi_i) \quad (6)$$

$$\text{s.t. } \sum_i c_i(\xi_i) \xi_i = B. \quad (7)$$

Its Lagrangian is

$$\mathcal{L}(\xi_1, \dots, \xi_n, \alpha) = \sum_i c_i(\xi_i) - \alpha \left(\sum_i c_i(\xi_i) \xi_i - B \right), \quad (8)$$

where α is the Lagrangian multiplier. Then we take its gradient on ξ_i and let it be 0:

$$\frac{\partial \mathcal{L}(\xi_1, \dots, \xi_n, \alpha)}{\partial \xi_i} = c'_i(\xi_i) - \alpha(c'_i(\xi_i) \xi_i + c_i(\xi_i)) = 0, \quad (9)$$

$$\frac{1}{\alpha} = \frac{c'_i(\xi_i) \xi_i + c_i(\xi_i)}{c'_i(\xi_i)} = \xi_i + \frac{c_i(\xi_i)}{c'_i(\xi_i)}, \quad (10)$$

where the equation holds for each ad exchange i . As such, we can use α to bridge the equations for any two ad exchanges i and j :

$$\frac{1}{\alpha} = \xi_i + \frac{c_i(\xi_i)}{c'_i(\xi_i)} = \xi_j + \frac{c_j(\xi_j)}{c'_j(\xi_j)}. \quad (11)$$

So the optimal solution condition is given as follows:

$$\frac{1}{\alpha} = \xi_1 + \frac{c_1(\xi_1)}{c'_1(\xi_1)} = \xi_2 + \frac{c_2(\xi_2)}{c'_2(\xi_2)} = \dots = \xi_n + \frac{c_n(\xi_n)}{c'_n(\xi_n)}, \quad (12)$$

$$\sum_i c_i(\xi_i) \xi_i = B. \quad (13)$$

With sufficient data instances, we find that $c_i(\xi_i)$ is usually a concave and smooth function. Some examples are given in Figure 4. Based on the observation, it is reasonable to define a general polynomial form of the $c_i(\xi_i)$ functions:

$$c_i(\xi_i) = c_i^* a_i \left(\frac{\xi_i}{\xi_i^*} \right)^{b_i}, \quad (14)$$

where ξ_i^* is the campaign’s historic average eCPC on the ad inventories from ad exchange i during the training data period, and c_i^* is the corresponding click number. These two factors are directly obtained from the training data. Parameters a_i and b_i are to be tuned to fit the training data.

Substituting Eq. (14) into Eq. (12) gives

$$\frac{1}{\alpha} = \xi_i + \frac{c_i(\xi_i)}{c'_i(\xi_i)} = \xi_i + \frac{\frac{c_i^* a_i}{\xi_i^*} \xi_i^{b_i}}{\frac{c_i^* a_i}{\xi_i^*} b_i \xi_i^{b_i-1}} = \left(1 + \frac{1}{b_i} \right) \xi_i. \quad (15)$$

We can then rewrite Eq. (12) as

$$\frac{1}{\alpha} = \left(1 + \frac{1}{b_1} \right) \xi_1 = \left(1 + \frac{1}{b_2} \right) \xi_2 = \dots = \left(1 + \frac{1}{b_n} \right) \xi_n. \quad (16)$$

$$\text{Thus } \xi_i = \frac{b_i}{\alpha(b_i + 1)}. \quad (17)$$

Interestingly, from Eq. (17) we find that the equilibrium is *not* in the state that the eCPCs from the exchanges are the same. Instead, it is when any amount of budget reallocated among the exchanges does not make any more total clicks; for instance, in a two-exchange case, the equilibrium reaches when the increase of the clicks from one exchange equals the decrease from the other (Eq. (9)). More specifically, from Eq. (17) we observe that for ad exchange i , if its click function $c_i(\xi_i)$ is quite flat, i.e., the click number increases much slowly as its eCPC increases in a certain area, then its learned b_i should be small. This means the factor $\frac{b_i}{b_i+1}$ is small as well; then from Eq. (17) we can see the optimal eCPC in ad exchange i should be relatively small.

Substituting Eqs. (14) and (17) into Eq. (7) gives

$$\sum_i \frac{c_i^* a_i}{\xi_i^* b_i} \left(\frac{b_i}{b_i + 1} \right)^{b_i+1} \left(\frac{1}{\alpha} \right)^{b_i+1} = B, \quad (18)$$

where for simplicity, we denote for each ad exchange i , its parameter $\frac{c_i^* a_i}{\xi_i^* b_i} \left(\frac{b_i}{b_i + 1} \right)^{b_i+1}$ as δ_i . This give us a simpler form as:

$$\sum_i \delta_i \left(\frac{1}{\alpha} \right)^{b_i+1} = B. \quad (19)$$

There is no closed form to solve Eq. (19) for α . However, as b_i cannot be negative and $\sum_i \delta_i \left(\frac{1}{\alpha} \right)^{b_i+1}$ monotonically increases against $\frac{1}{\alpha}$, one can easily obtain the solution for α by using a numeric solution such as the stochastic gradient decent or the Newton method [5]. Finally, based on the solved α , we can find the optimal eCPC ξ_i for each ad exchange i using Eq. (17). In fact, these eCPCs are the reference value we want the campaign to achieve for the corresponding ad exchanges. We can use PID controllers, by setting x_r in Eq. (3) as ξ_i for each ad exchange i , to achieve these reference eCPCs so as to achieve the maximum number of clicks on the campaign level.

As a special case, if we regard the whole volume of the campaign as one channel, this method can be directly used as a general bid optimisation tool. It makes use of the campaign’s historic data to decide the optimal eCPC and then the click optimisation is performed by control the eCPC to settle at the optimal eCPC as reference. Note that this multi-channel click maximisation framework is flexible to incorporate any bidding strategies.

4. EMPIRICAL STUDY

We conduct comprehensive experiments to study the proposed RTB feedback control mechanism. Our focus in this section is on offline evaluation using a publicly-available real-world dataset. To make our experiment repeatable, we have published the experiment code¹. The online deployment and test on a commercial DSP will be reported in Section 5.

4.1 Evaluation Setup

Dataset. We test our system on a publicly available dataset collected from iPinYou DSP [19]. It contains the ad log data from 9 campaigns during 10 days in 2013, which consists of 64.75M bid records, 19.50M impressions, 14.79K

¹<https://github.com/wnzhang/rtbcontrol>

clicks and 16K Chinese Yuan (CNY) expense. According to the data publisher [19], the last three-day data of each campaign is split as the test data and the rest as the training data. The dataset disk size is 35GB. More statistics and analysis of the dataset is available in [34]. The dataset is in a record-per-row format, where each row consists of three parts: (i) The features for this auction, e.g., the time, location, IP address, the URL/domain of the publisher, ad slot size, user interest segments etc. The features of each record are indexed as a 0.7M-dimension sparse binary vector which is fed into a logistic regression CTR estimator of the bidding strategy in Eq. (1); (ii) The auction winning price, which is the threshold of the bid to win this auction; (iii) The user feedback on the ad impression, i.e., click or not.

Evaluation Protocol. We follow the evaluation protocol from previous studies on bid optimisation [33, 34] and an RTB contest [19] to run our experiment. Specifically, for each data record, we pass the feature information to our bidding agent. Our bidding agent generates a new bid based on the CTR prediction and other parameters in Eq. (1). We then compare the generated bid with the logged actual auction winning price. If the bid is higher than the auction winning price, we know the bidding agent has won this auction, paid the winning price, and obtained the ad impression. If from the ad impression record there is a click, then the placement has generated a positive outcome (one click) with a cost equal to the winning price. If there is no click, the placement has resulted in a negative outcome and wasted the money. The control parameters are updated every 2 hours (as one round).

It is worth mentioning that historical user feedback has been widely used for evaluating information retrieval systems [29] and recommender systems [13]. All of them used historic clicks as a proxy for relevancy to train the prediction model as well as to form the ground truth. Similarly, our evaluation protocol keeps the user contexts, displayed ads (creatives etc.), bid requests, and auction environment unchanged. We intend to answer that under the same context if the advertiser were given a different or better bidding strategy or employed a feedback loop, whether they would be able to get more clicks with the budget limitation. The click would stay the same as nothing has been changed for the users. This methodology works well for evaluating bid optimisation [1, 33] and has been adopted in the display advertising industry [19].

Evaluation Measures. We adopt several commonly used measures in feedback control systems [3]. We define the *error band* as the $\pm 10\%$ interval around the reference value. If the controlled variable settles within this area, we consider that the variable is successfully controlled. The **speed** of convergence (to the reference value) is also important. Specifically, we evaluate the *rise time* to check how fast the controlled variable will get into the error band. We also use the *settling time* to evaluate how fast the controlled variable will be successfully restricted into the error band. However, fast convergence may bring the problem of inaccurate control. Thus, two control **accuracy** measures are introduced. We use the *overshoot* to measure the percentage of value that the controlled variable passes over the reference value. After the settling (called the steady state), we use the *RMSE-SS* to evaluate the root mean square error between the controlled variable value and the reference value. At last, we measure the control **stability** by calculating the standard deviation of the variable value after the settling, named as *SD-SS*.

For bid optimisation performance, we use the campaign's total achieved click number and eCPC as the prime evaluation measures. We also monitor the impression related performance such as impression number, AWR and CPM.

Table 1: Overall control performance on eCPC.

Cpg.	Cntr	Rise	Settling	Overshoot	RMSE-SS	SD-SS
1458	PID	1	5	7.73	0.0325	0.0313
	WL	6	36	0	0.0845	0.0103
2259	PID	7	7	8.03	0.0449	0.0411
	WL	6	-	0	-	-
2261	PID	3	23	17.66	0.0299	0.0294
	WL	5	-	0	-	-
2821	PID	17	22	14.47	0.0242	0.0216
	WL	-	-	0	-	-
2997	PID	17	17	0.75	0.0361	0.026
	WL	-	-	0	-	-
3358	PID	3	7	23.89	0.0337	0.0287
	WL	-	-	0	-	-
3386	PID	9	13	7.90	0.0341	0.0341
	WL	-	-	0	-	-
3427	PID	1	12	29.03	0.0396	0.0332
	WL	-	-	0	-	-
3476	PID	1	5	7.64	0.0327	0.031
	WL	1	-	17.11	-	-

Table 2: Overall control performance on AWR.

Cpg.	Cntr	Rise	Settling	Overshoot	RMSE-SS	SD-SS
1458	PID	4	10	16.86	0.0153	0.0093
	WL	3	7	0.00	0.0448	0.0231
2259	PID	4	6	17.08	0.0076	0.0072
	WL	1	13	3.91	0.0833	0.0113
2261	PID	1	4	16.39	0.0205	0.0203
	WL	1	-	2.02	-	-
2821	PID	6	8	16.44	0.0086	0.0086
	WL	1	3	5.77	0.0501	0.0332
2997	PID	1	8	13.68	0.0151	0.0151
	WL	1	-	0.00	-	-
3358	PID	2	8	22.08	0.0250	0.0213
	WL	1	7	0.13	0.0332	0.0211
3386	PID	4	8	18.85	0.0133	0.0118
	WL	1	5	2.95	0.0300	0.0291
3427	PID	2	6	26.63	0.0200	0.0179
	WL	3	13	0.24	0.0482	0.0257
3476	PID	2	6	27.15	0.0175	0.0161
	WL	1	7	1.49	0.0308	0.0271

Empirical Study Organisation. Our empirical study consists of five parts with the focus on controlling two KPIs: eCPC and AWR. (i) In Section 4.2, we answer whether the proposed feedback control systems are practically capable of controlling the KPIs. (ii) In Section 4.3, we study the control difficulty with different reference value settings. (iii) In Section 4.4, we focus on the PID controller and investigate its attributes on settling the target variable. (iv) In Section 4.5, we leverage the PID controllers as a bid optimisation tool and study their performance on optimising the campaign's clicks and eCPC across multiple ad exchanges. (v) Finally, more discussions about PID parameter tuning and online update will be given in Section 4.6.

4.2 Control Capability

For each campaign, we check the performance of the two controllers on two KPIs. We first tune the control parameters on the training data to minimise the settling time. Then we adopt the controllers over the test data and observe the performance. The detailed control performance on each campaign is provided in Table 1 for eCPC² and Table 2 for AWR. Figure 5 shows the controlled KPI curves against the timesteps (i.e., round). The dashed horizontal line means the reference.

We see from the results that (i) all the PID controllers can settle both KPIs within the error band (with the settling time less than 40 rounds), which indicates that the PID control is capable of settling both KPIs at the given reference value. (ii) The WL controller on eCPC does not work that well on test data, even though we could find good parameters on training data. This is due to the fact that WL controller tries to affect the average system behaviour through transient performance feedbacks while facing the

²“-” cells mean invalid because of the failure to rise or settle.

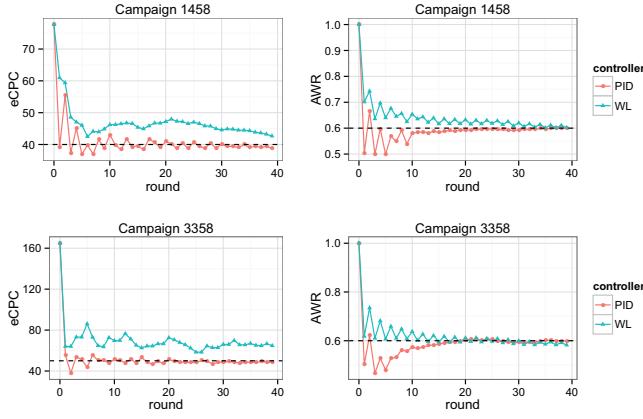


Figure 5: Control performance on eCPC and AWR.

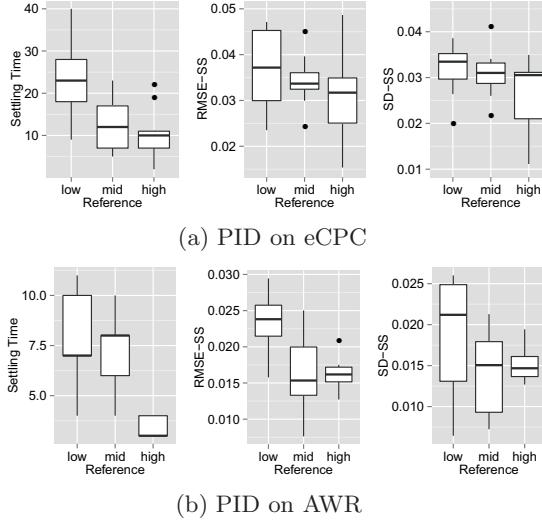


Figure 6: Control difficulty comparison with PID.

huge dynamics of RTB. (iii) For WL on AWR, most campaigns are controllable while there are still two campaigns that fail to settle at the reference value. (iv) Compared to PID on AWR, WL always results in higher RMSE-SS and SD-SS values but lower overshoot percentage. Those control settings with a fairly short rise time usually face a higher overshoot. (v) In addition, we observe that the campaigns with higher CTR estimator AUC performance (referring [34]) normally get shorter settling time.

According to above results, PID controller outperforms the WL controller in the tested RTB cases. We believe this is due to the fact that the integral factor in PID controller helps reduce the accumulative error (i.e., RMSE-SS) and the derivative factor helps reduce the variable fluctuation (i.e., SD-SS). And it is easier to settle the AWR than the eCPC. This is mainly because AWR only depends on the market price distribution while eCPC additionally involves the user feedback, i.e., CTR, where the prediction is associated with significant uncertainty.

4.3 Control Difficulty

In this section, we extend our control capability experiments further by adding higher and lower reference values in comparison. Our goal is to investigate the impact of different levels of reference values on control difficulty. We follow the same scheme to train and test the controllers as Section 4.2. However, instead of showing the exact performance value, our focus here is on the performance comparison with different reference settings.

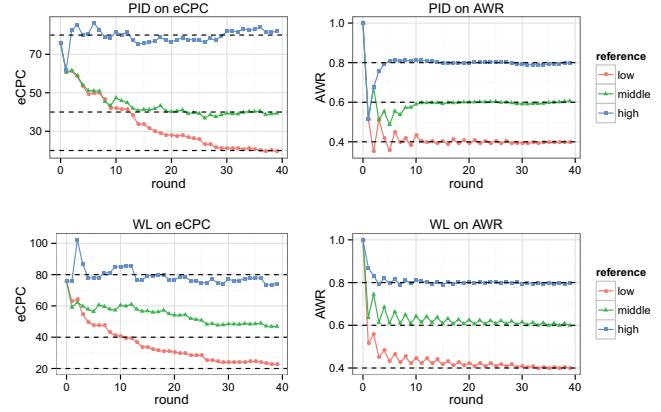


Figure 7: Control performance for campaign 3386 on eCPC and AWR with different reference values.

The distribution of achieved settling time, RMSE-SS and SD-SS, with the setting of three reference levels, i.e., low, middle and high, are shown in the form of box plot [20] in the Figure 6(a) and 6(b) for the eCPC and AWR control with PID. We observe that the average settling time, RMSE-SS and SD-SS, are reduced as the reference values get higher. This shows that generally the control tasks with higher reference eCPC and AWR are easier to achieve because one can simply bid higher to win more and spend more. Also as the higher reference is closer to the initial performance value, the control signal does not bring serious bias or volatility, which leads to the lower RMSE-SS and SD-SS. For the page limit, the control performance with WL is not presented here. The results are similar with PID.

Figure 7 gives the specific control curves of the two controllers with three reference levels on a sample campaign 3386. We find that the reference value which is farthest away from the initial value of the controlled variable brings the largest difficulty for settling, both on eCPC and AWR. This suggests that advertisers setting an ambitious control target will introduce the risk of unsettling or large volatility. The advertisers should try to find a best trade-off between the target value and the practical control performance.

4.4 PID Settling: Static vs. Dynamic References

The combination of proportional, integral and derivative factors enables the PID feedback to automatically adjust the settling progress during the control lifetime with high efficiency [7]. Alternatively, one can empirically adjust the reference value in order to achieve the desired reference value. For example of eCPC control, if the campaign's achieved eCPC is higher than the initial reference value right after exhausting the first half budget, the advertiser might want to lower the reference value in order to accelerate the downward adjustment and finally reach its initial eCPC target before running out of the budget. PID feedback controller implicitly handles such problem via its integration factor [7, 28]. In this section, we investigate with our RTB feedback control mechanism whether it is still necessary for advertisers to intentionally adjust the reference value according to the campaign's real-time performance.

Dynamic Reference Adjustment Model. To simulate the advertisers' strategies to adaptively change the reference value of eCPC and AWR under the budget constraint, we propose a dynamic reference adjustment model to calculate the new reference $x_r(t_{k+1})$ after t_k :

$$x_r(t_{k+1}) = \frac{(B - s(t_k))x_r(t_k)}{Bx(t_k) - s(t_k)x_r}, \quad (20)$$

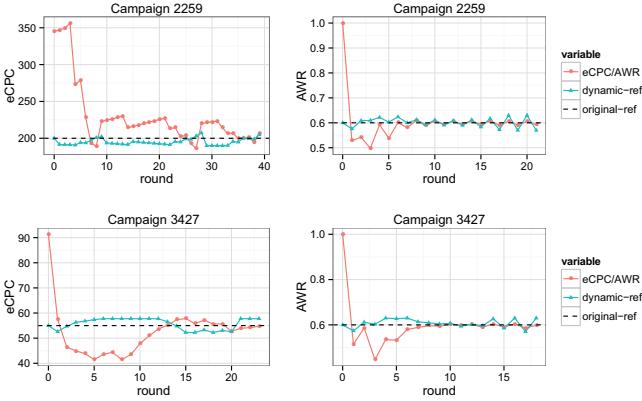


Figure 8: Dynamic reference control with PID.

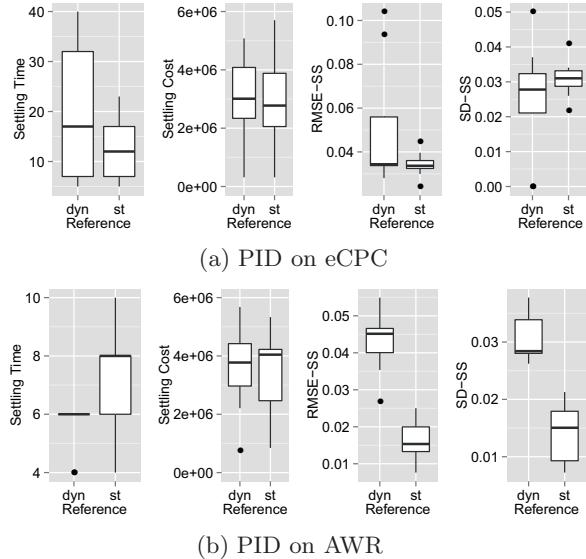


Figure 9: Dynamic v.s. static reference with PID.

where x_r is the initial reference value, $x(t_k)$ is the achieved KPI (eCPC or AWR) at timestep t_k , B is the campaign budget, $s(t_k)$ is the cost so far. We can see from Eq. (20) that when $x_r(t_k) = x_r$, $x_r(t_{k+1})$ will be set the same as x_r ; when $x_r(t_k) > x_r$, $x_r(t_{k+1})$ will be set lower than x_r and vice versa. For readability, we leave the detailed derivation in appendix. Using Eq. (20) we calculate the new reference eCPC/AWR $x_r(t_{k+1})$ and use it to substitute x_r in Eq. (3) to calculated the error factor so as to make the dynamic-reference control.

Results and Discussions. Figure 8 shows the PID control performance with dynamic reference calculated based on Eq. (20). The campaign performance gets stopped at the point where the budget is exhausted. From the figure, we see that for both eCPC and AWR control, the dynamic reference takes an aggressive approach and pushes the eCPC or AWR across the original reference value (dashed line). This actually simulates some advertisers' strategy: when the performance is lower than the reference, then higher the dynamic reference to push the total performance to the initial reference more quickly. Furthermore, for AWR control, we can see the dynamic reference fluctuates seriously when the budget is to be exhausted soon. This is because when there is insufficient budget left, the reference value will be set much high or low by Eq. (20) in order to push the performance back to the initial target. Apparently this is an ineffective solution.

Table 3: Control performance on multi-exchanges with the reference eCPC set for click maximisation.

Cpg.	AdEx	Rise	Settling	Cpg.	AdEx	Rise	Settling
1458	1	13	26	3358	1	9	20
	2	15	18		2	14	39
	3	13	13		3	26	26
2259	1	10	38	3386	1	6	18
	2	3	14		2	12	12
	3	3	29		3	1	1
2261	1	3	30	3427	1	16	16
	2	7	38		2	35	35
	3	0	35		3	23	23
2821	1	6	17	3476	1	18	29
	2	3	10		2	22	28
	3	15	15		3	19	22
	4	4	38				

Furthermore, we directly compare the quantitative control performance between dynamic-reference controllers (dyn) with the standard static-reference ones (st) using PID. Besides the settling time, we also compare the settling cost, which is the spent budget before settling. The overall performance across all the campaigns is shown in Figure 9(a) for eCPC control and Figure 9(b) for AWR control. The results show that (i) for eCPC control, the dynamic-reference controllers do *not* perform better than the static-reference ones; (ii) for AWR control, the dynamic-reference controllers could reduce the settling time and cost, but the accuracy (RMSE-SS) and stability (SD-SS) is much worse than the static-reference controllers. This is because the dynamic reference itself brings volatility (see Figure 8). These results demonstrate that PID controller does perform a good enough way to settling the variable towards the pre-specified reference without the need of dynamically adjusting the reference to accelerate using our methods. Other dynamic reference models might be somewhat effective but this is not the focus of this paper.

4.5 Reference Setting for Click Maximisation

We now study how the proposed feedback control could be used for click optimisation purpose. As we have discussed in Section 3.4, bid requests usually come from different ad exchanges where the market power and thus the CPM prices are disparate. We have shown that given a budget constraint, the number of clicks is maximised if one can control the eCPC in each ad exchange by settling it at an optimal eCPC reference for each of them, respectively.

In this experiment, we build a PID feedback controller for each of its integrated ad exchanges, where their reference eCPCs are calculated via Eqs. (17) and (19). We train the PID parameters on the training data of each campaign, and then test the bidding performance on the test data. As shown in Table 3, the eCPC on all the ad exchanges for all tested campaigns get settled at the reference values³ (settling time less than 40). We denote our multi-exchange eCPC feedback control method as **multiple**. Besides **multiple**, we also test a baseline method which assigns a single optimal uniform eCPC reference across all the ad exchanges, denoted as **uniform**. We also use the linear bidding strategy without feedback control [25] as a baseline, denoted as **none**⁴.

The comparisons over various evaluation measures are reported in Figure 10. We observe that (i) the feedback-control-enabled bidding strategies **uniform** and **multiple** significantly outperform the non-controlled bidding strategy **none** in terms of the number of achieved clicks and eCPC. This suggests that properly controlling eCPCs would lead to an optimal solution for maximising clicks. (ii) By reallo-

³Campaign 2997 is only integrated with one ad exchange, thus not compared here.

⁴Other bidding strategies [33, 17] are also investigated. Producing similar results, they are omitted here for clarity.

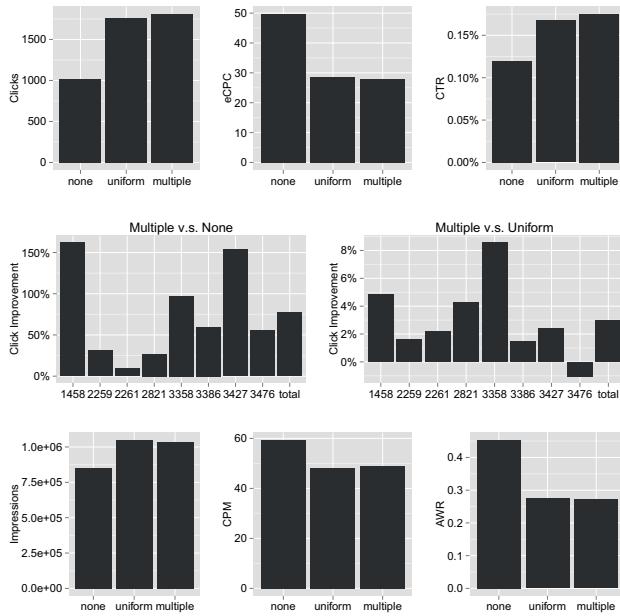


Figure 10: Bid optimisation performance.

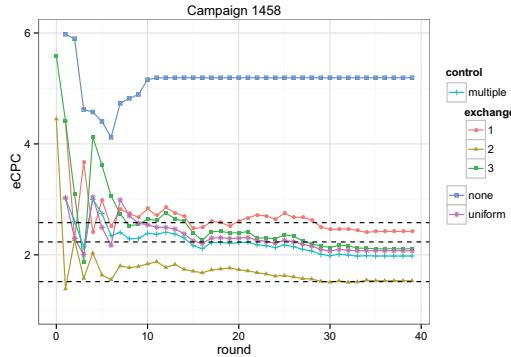


Figure 11: Settlement of multi-exchange feedback control.

cating the budget via setting different reference eCPCs on different ad exchanges, **multiple** further outperforms **uniform** on 7 out of 8 tested campaigns. (iii) On the impression related measures, the feedback-control-enabled bidding strategies earn more impressions than the non-controlled bidding strategy by actively lowering their bids (CPM) and thus AWR, but achieving more bid volumes. This suggests that by allocating more budget to the lower valued impressions, one could potentially generate more clicks. As a by-product, this confirms the theoretical finding reported in [33].

As a case study, Figure 11 plots the settling performance of the three methods on campaign 1458. The three dashed horizontal lines are the reference eCPCs on three ad exchanges. We see that the eCPCs on the three ad exchanges successfully settle at the reference eCPCs. At the same time, the campaign-level eCPC (**multiple**) settles at a lower value than **uniform** and **none**.

4.6 PID Parameter Tuning

In this subsection, we share some lessons learned about PID controller parameter tuning and online update.

Parameter Search. Empirically, λ_D does not change the control performance significantly. Just a small valued λ_D , e.g., 1×10^{-5} , will reduce the overshoot and slightly shorten the settling time. Thus the parameter search is focused on

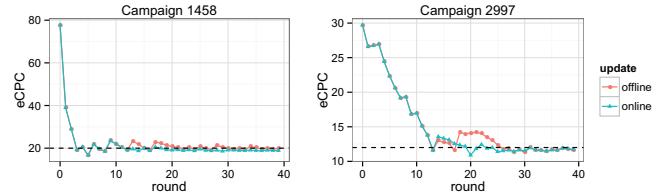


Figure 12: Control with online/offline parameter updating.

λ_P and λ_I . Instead of using the computationally expensive grid search, we perform an adaptive coordinate search. For every update, we fix one parameter and shoot another one to seek for the optimal value leading shortest settling time, and the line searching step length shrinks exponentially for each shooting. Normally after 3 or 4 iterations, the local optima is reached and we find such solution is highly comparable with the expensive grid search.

Setting $\phi(t)$ Bounds. We also find that setting up upper/lower bounds of control signal $\phi(t)$ is important to make KPIs controllable. Due to the dynamics in RTB, it is common that user CTR drops during a period, which makes eCPC much higher. The corresponding feedback would probably result in a large negative gain on the bids, leading extremely low bid price and thus no win, no click and no additional cost at all for remaining rounds. In such case, a proper lower bound (-2) of $\phi(t)$ aims to eliminate above extreme effects by preventing from a seriously negative control signal. In addition, an upper bound (5) is used in order to avoid excessive variable growth beyond the reference value.

Online Parameter Updating. As the DSP running with feedback control, the collected data can be immediately utilised for training a new PID controller and updating the older one. We investigate the possibility of the online updating of PID parameters with the recent data. Specifically, after initialising the PID parameters using training data, we re-train the controller for every 10 rounds (i.e., before round 10, 20 and 30) in the test stage using all previous data with the same parameter searching method as in the training stage. The parameter searching in re-training takes about 10 minutes for each controller, which is far shorter than the round period (2 hours). Figure 12 shows the control performance with PID parameters tuned online and offline respectively. As we can see after the 10th round (i.e., the first online tuning point), the online-tuned PIDs manage to control the eCPC around the reference value more effectively than the offline-tuned one, resulting shorter settling time and lower overshoot. In addition, no obvious disturbance or instability occurs when we switch parameters. With the online parameter updating, we can start to train the controllers based on several-hour training data and adaptively update the parameters from the new data to improve the control performance.

5. ONLINE DEPLOYMENT AND TEST

The proposed RTB feedback control system has been deployed and tested in live on BigTree DSP⁵, a performance-driven mobile advertising DSP in China. BigTree DSP focuses on the programmatic buying for optimal advertising performance on mobile devices, which makes it an ideal place to test our proposed solution.

The deployment environment is based on Aliyun elastic cloud computing servers. A three-node cluster is deployed for the DSP bidding agent, where each node is in Ubuntu 12.04, with 8 core Intel Xeon CPU E5-2630 (2.30GHz) and

⁵<http://www.bigtree.mobi/>

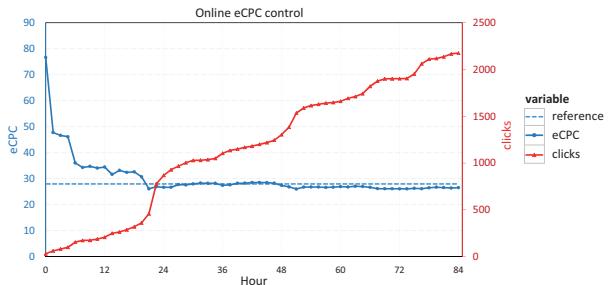


Figure 13: The online eCPC control performance and the accumulative click numbers of a mobile game campaign on BigTree DSP.

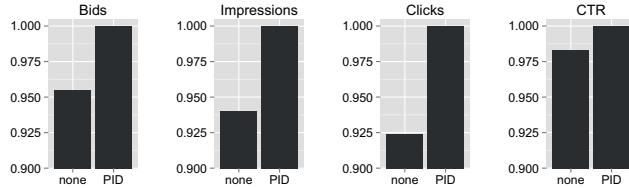


Figure 14: Relative performance for online test.

8GB RAM. The controller module is implemented in Python with uWSGI and Nginx.

For BigTree DSP controller module, we deploy the PID control function and tune its parameters. Specifically, we use the last 6-week bidding log data in 2014 as the training data for tuning PID parameters. A three-fold validation process is performed to evaluate the generalisation of the PID control performance, where the previous week data is used as the training data while the later week data is used for validation. The control factors ($\phi(t), e(t_k)$ in Eq. (4)) are updated for every 90 minutes. After acquiring a set of robust and effective PID parameters, we launch the controller module, including the monitor and actuator submodules, on BigTree DSP.

Figure 13 shows the online eCPC control performance on one of the iOS mobile game campaigns during 84 hours from 7 Jan. 2015 to 10 Jan. 2015. The reference eCPC is set as 28 RMB cent by the advertiser, which is about 0.8 times the average eCPC value of the previous week where there was no control. Following the same training process described in the previous section, we update the online control factors for every 90 minutes. From the result we can see the eCPC value dropped from the beginning 79 cent to 30 during the first day and then settled closed to the reference afterwards.

In the meantime, A/B testing is used to compare with the non-controlled bidding agent (with the same sampling rate but disjoint bid requests). Figure 14 shows the corresponding advertising performance comparison between a non-controlled bidding agent and the PID-control bidding agent during the test period with the same budget. As we can see, by settling the eCPC value around the lower reference eCPC, the PID-control bidding agent acquires more bid volume and win more (higher-CTR) impressions and clicks, which demonstrates its ability of optimising the performance.

Compared with the offline empirical study, the online running is more challenging: (i) all pipeline steps including the update of the CTR estimator, the KPI monitor linked to the database and the PID controller should operate smoothly against the market turbulence; (ii) the real market competition is highly dynamic during the new year period when we launched our test; (iii) other competitors might tune their bidding strategies independently or according to any changes of their performance after we employed the controlled bid-

ding strategy. In sum, the successful eCPC control on an online commercial DSP verifies the effectiveness of our proposed feedback control RTB system.

6. RELATED WORK

Enabling the impression-level evaluation and bidding, much research work has been done on RTB display advertising, including bidding strategy optimisation [25, 33], reserve price optimisation [30], ad exchange auction design [4], and ad tracking [11].

In order to perform the optimal bidding, the DSP bidding agent should estimate both utility and cost of a given ad impression. The impression-level utility evaluation, including CTR and conversion rate (CVR) estimation, is the essential part for each bidding agent in DSPs. In [18] the sparsity problem of CVR estimation is handled by modelling the conversions at different hierarchical levels. The user click behaviour on mobile RTB ads is studied in [24]. On the cost evaluation side, bid landscape modelling and forecasting is much important to inform the bidding agent about the competitiveness of the market. The authors in [9] break down the campaign-level bid landscape forecasting problem into “samples” by targeting rules and then employ a mixture model of log-normal distributions to build the campaign-level bid landscape. The authors in [16] try to reduce the bid landscape forecasting error through frequently re-building the landscape models. Based on the utility and cost evaluation of the ad inventory, bid optimisation is performed to improve the advertising performance under the campaign budget constraint. Given the estimated CTR/CVR, the authors in [25, 18] employ linear bidding functions based on truth-telling attributes of second price auctions. However, given the budget constraint, the advertisers’ bidding behaviour is not truth-telling. The authors in [33] propose a general bid optimisation framework to maximise the desired advertising KPI (e.g., total click number) under the budget constraint. Besides the general bid optimisation, the explicit bidding rules such as frequency and recency capping are studied in [31]. Moreover, the budget pacing [17] which refers to smoothly delivering the campaign budget is another important problem for DSPs.

There are a few research papers on recommender systems leveraging feedback controllers for performance maintenance and improvement. In [21], a rating updating algorithm based on the PID controller is developed to exclude unfair ratings in order to build a robust reputation system. The authors in [32] apply a self-monitoring and self-adaptive approach to perform a dynamic update of the training data fed into the recommender system to automatically balance the computational cost and the prediction accuracy. Furthermore, the authors in [14] adopt the more effective and well-studied PID controller to the data-feeding scheme of recommender systems, which is proved to be practically effective in their studied training task.

Compared to the work of controlling the recommender system performance by changing the number of training cases, our control task in RTB is more challenging, with much various dynamics from advertising environment such as the fluctuation of market price, auction volume and user behaviour. In [8], the authors discuss multiple aspects in a performance-driven RTB system, where the impression volume control is one of discussed aspects. Specifically, WL and a model-based controller are implemented to control the impression volume during each time interval. In [15], feedback control is used to perform budget pacing in order to stabilise the conversion volume. Compared to [8, 15], our work is a more comprehensive study focused on the feedback control techniques to address the instability problem in RTB. Besides

WL, we intensively investigate the PID controller, which takes more factors into consideration than WL. For the controlled KPIs, we look into the control tasks on both eCPC and AWR, which are crucial KPIs for performance-driven campaigns and branding-based campaigns, respectively. In addition, we proposed an effective model to calculate the optimal eCPC reference to maximise the campaign's clicks using feedback controllers.

7. CONCLUSIONS

In this paper, we have proposed a feedback control mechanism for RTB display advertising, with the aim of improving its robustness of achieving the advertiser's KPI goal. We mainly studied PID and WL controllers for controlling the eCPC and AWR KPIs. Through our comprehensive empirical study, we have the following discoveries. (i) Despite of the high dynamics in RTB, the KPI variables are controllable using our feedback control mechanism. (ii) Different reference values bring different control difficulties, which are reflected in the control speed, accuracy and stability. (iii) PID controller naturally finds its best way to settle the variable, and there is no necessity to adjust the reference value for accelerating the PID settling. (iv) By settling the eCPCs to the optimised reference values, the feedback controller is capable of making bid optimisation. Deployed on a commercial DSP, the online test demonstrates the effectiveness of the feedback control mechanism in generating controllable advertising performance. In the future work, we will further study the applications based on feedback controllers in RTB, such as budget pacing and retargeting frequency capping.

8. REFERENCES

- [1] K. Amin, M. Kearns, P. Key, and A. Schwaighofer. Budget optimization for sponsored search: Censored learning in mdps. *UAI*, 2012.
- [2] K. J. Åström and P. Kumar. Control: A perspective. *Automatica*, 50(1):3–43, 2014.
- [3] K. J. Åström and R. M. Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.
- [4] S. Balseiro, O. Besbes, and G. Y. Weintraub. Repeated auctions with budgets in ad exchanges: Approximations and design. *Columbia Business School Research Paper*, (12/55), 2014.
- [5] R. Battiti. First-and second-order methods for learning: between steepest descent and newton's method. *Neural computation*, 4(2):141–166, 1992.
- [6] S. Bennett. Development of the pid controller. *Control Systems*, 13(6):58–62, 1993.
- [7] S. P. Bhattacharyya, H. Chapellat, and L. H. Keel. Robust control. *The Parametric Approach*, by Prentice Hall PTR, 1995.
- [8] Y. Chen, P. Berkhin, B. Anderson, and N. R. Devanur. Real-time bidding algorithms for performance-based display ad allocation. *KDD*, 2011.
- [9] Y. Cui, R. Zhang, W. Li, and J. Mao. Bid landscape forecasting in online ad exchange marketplace. In *KDD*, pages 265–273. ACM, 2011.
- [10] B. W. Dezotell. Water level controller, June 9 1936. US Patent 2,043,530.
- [11] R. Gomer, E. M. Rodrigues, N. Milic-Frayling, and M. Schraefel. Network analysis of third party tracking: User exposure to tracking cookies through search. In *WI*, 2013.
- [12] Google. The arrival of real-time bidding. Technical report, 2011.
- [13] Y. Hu, Y. Koren, and C. Volinsky. Collaborative filtering for implicit feedback datasets. In *ICDM*, 2008.
- [14] T. Jambor, J. Wang, and N. Lathia. Using control theory for stable and efficient recommender systems. In *WWW*, 2012.
- [15] N. Karlsson and J. Zhang. Applications of feedback control in online advertising. In *American Control Conference*, 2013.
- [16] K. J. Lang, B. Moseley, and S. Vassilvitskii. Handling forecast errors while bidding for display advertising. In *WWW*, 2012.
- [17] K.-C. Lee, A. Jalali, and A. Dasdan. Real time bid optimization with smooth budget delivery in online advertising. In *ADKDD*, 2013.
- [18] K.-C. Lee, B. Orten, A. Dasdan, and W. Li. Estimating conversion rate in display advertising from past performance data. In *KDD*, 2012.
- [19] H. Liao, L. Peng, Z. Liu, and X. Shen. ipinyou global rtb bidding algorithm competition dataset. In *ADKDD*, 2014.
- [20] R. McGill, J. W. Tukey, and W. A. Larsen. Variations of box plots. *The American Statistician*, 32(1):12–16, 1978.
- [21] K. Meng, Y. Wang, X. Zhang, X.-c. Xiao, et al. Control theory based rating recommendation for reputation systems. In *ICNSC*, 2006.
- [22] S. Muthukrishnan. Ad exchanges: Research issues. In *Internet and network economics*, pages 1–12. Springer, 2009.
- [23] R. C. Nelson. *Flight stability and automatic control*, volume 2. WCB/McGraw Hill, 1998.
- [24] R. J. Oentaryo, E. P. Lim, D. J. W. Low, D. Lo, and M. Finegold. Predicting response in mobile advertising with hierarchical importance-aware factorization machine. In *WSDM*, 2014.
- [25] C. Perlich, B. Dalessandro, R. Hook, O. Stitelman, T. Raeder, and F. Provost. Bid optimizing and inventory scoring in targeted online advertising. In *KDD*, pages 804–812, 2012.
- [26] G. Peterson and D. J. Cook. Decision-theoretic layered robotic control architecture. In *AAAI/IAAI*, page 976, 1999.
- [27] M. Richardson, E. Dominowska, and R. Ragno. Predicting clicks: estimating the click-through rate for new ads. In *WWW*, pages 521–530. ACM, 2007.
- [28] D. E. Rivera, M. Morari, and S. Skogestad. Internal model control: Pid controller design. *Industrial & engineering chemistry process design and development*, 25(1), 1986.
- [29] J. Xu, C. Chen, G. Xu, H. Li, and E. R. T. Abib. Improving quality of training data for learning to rank using click-through data. In *WSDM*, 2010.
- [30] S. Yuan, B. Chen, J. Wang, P. Mason, and S. Seljan. An Empirical Study of Reserve Price Optimisation in Real-Time Bidding. In *KDD*, 2014.
- [31] S. Yuan, J. Wang, and X. Zhao. Real-time bidding for online advertising: measurement and analysis. In *ADKDD*, 2013.
- [32] V. Zanardi and L. Capra. Dynamic updating of online recommender systems via feed-forward controllers. In *SEAMS*, 2011.
- [33] W. Zhang, S. Yuan, and J. Wang. Optimal real-time bidding for display advertising. In *KDD*, 2014.
- [34] W. Zhang, S. Yuan, J. Wang, and X. Shen. Real-time bidding benchmarking with ipinyou dataset. *arXiv*, 2014.
- [35] W. Zhang, Y. Zhang, B. Gao, Y. Yu, X. Yuan, and T.-Y. Liu. Joint optimization of bid and budget allocation in sponsored search. In *KDD*, 2012.

APPENDIX

Reference Adjust Models. Here we provide the detailed derivation of the proposed dynamic-reference model Eq. (20) in Section 4.4. We mainly introduce the derivation of reference eCPC adjustment, while the derivation of reference AWR adjustment can be obtained similarly.

Let ξ_r be the initial eCPC target, $\xi(t_k)$ be the achieved eCPC before the moment t_k , $s(t_k)$ be the total cost so far, and B be the campaign budget. In such setting, the current achieved click number is $s(t_k)/\xi(t_k)$ and the target click number is B/ξ_r . In order to push the overall eCPC to ξ_r , i.e., push the total click number B/ξ_r with the budget B , the reference eCPC for the remaining time $\xi_r(t_{k+1})$ should satisfy

$$\frac{s(t_k)}{\xi(t_k)} + \frac{B - s(t_k)}{\xi_r(t_{k+1})} = \frac{B}{\xi_r}. \quad (21)$$

Solving the equation we have

$$\xi_r(t_{k+1}) = \frac{(B - s(t_k))\xi_r\xi(t_k)}{B\xi(t_k) - s(t_k)\xi_r}. \quad (22)$$

The derivation of reference AWR adjustment is much similar but with an extra winning function which links between the bid price and winning probability [33]. The result formula is just the same as Eq. (22). Using x as a general notation for eCPC and AWR variables results in Eq. (20) in Section 4.4.