2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Optimization in Online Advertising via Simultaneous Adaptive Rate and Price Feedback Control

Niklas Karlsson[1]
Amazon Ads

***Abstract*— Online advertising is typically implemented via real-time bidding, and advertising campaigns are then defined as extremely high-dimensional optimization problems. To solve these problems in light of large scale and significant uncertainties, the optimization problems are modularized in a way that makes feedback control a critical component of the solution. The control problem, however, is challenging due to plant uncertainties, nonlinearities, time-variance, and noise. An Oracle would define the control signal in terms of bid price adjustments only; however, we propose the introduction of a companion throttling control signal that creates a useful plant linearity. In this paper, the control problem is redefined in such a way that the linearity is exploited for improved feedback control. A dual lever control algorithm is designed and evaluated in simulations, with promising results.**

***Index Terms* — Adaptive Control, Optimization, Real-time Bidding, Programmatic Advertising**

I. Introduction

Programmatic advertising is an important aspect of the business model for companies such as Amazon, Google, and Facebook. A *Demand Side Platform* (DSP) is an example of such a business model. A DSP provides the service of spending online advertisement budgets optimally. It represents advertisers and is situated between an advertiser and one or more open exchange trading so called ad *impressions*, which are opportunities to show an ad creative to Internet users. The DSP implements advanced algorithms to compute and submit bids on impressions in real time.

The optimization problems are extremely high-dimensional, but can be reformulated such that the bid engine becomes a three player non-cooperative game. The three players that combined produce bids on behalf of an advertiser are represented by impression valuation, campaign control, and bid shading optimization. Impression valuation computes the expected value of an impression conditioned on it being awarded to the campaign [1]–[3]. Campaign control makes bid adjustments to satisfy campaign delivery constraints (more on that later) [4]–[8]. And bid shading optimization computes the final bid by taking into account how much other campaigns are expected to bid for the same impression [9]–[12].

There is a rich literature on strategies for the three players, and improvements to the strategies remain an active research area in both academia and industry. This paper in particular deals with strategies for the campaign controller player.

Campaign control is a critical feature of the bid computation, but is challenging due to plant uncertainties, nonlinearities, time-variance, and noise. It is difficult to devise a solution that is simultaneously robust and that performance-wise is near-optimal. The nature of the control problem is described in [13], [14], and some techniques on how to overcome the challenges are discussed in [8], [15]. These techniques involve bid randomization, which ensures local linearity of the plant. However, the linear approximation is difficult to identify online, making stabilization of the closed loop system non-trivial.

Moreover, advertisers are becoming increasingly sophisticated and expect their campaigns to satisfy a growing number of optimization constraints [16], [17]. Each added constraint translates to a control problem, and as the dimension of the control problem grows, stability is harder to guarantee for a plant that is nonlinear and uncertain.

Our contribution is the introduction of a companion throttling control signal that creates a useful plant linearity. We redefine the control problem in such a way that the linearity is exploited for improved feedback control. A dual lever control algorithm is designed and evaluated in simulations, with promising results.

The article is organized as follows. Section II defines the optimization and control problems under investigation. Thereafter, Section III describes a sensible plant model used for design and simulation. The control design itself is developed in Section IV, whereas Section V summarizes the control algorithm. Basic simulation results are provided in Section VI, and conclusions and ideas of future work are discussed in Section VII.

II. Problem Formulation

We first we briefly describe the optimization problem that justifies the control problem considered in the remainder of the paper. The optimization objective is to maximize the cost-discounted profit $\mathcal{J} := \mathrm{E}V - \alpha \mathrm{E}C$, where $\mathrm{E}V$ is the total expected daily advertiser value, $\mathrm{E}C$ is the total expected daily cost of awarded impressions, and $\alpha \in (0, 1]$ is a given cost-discount parameter. The problem is subject to a spend constraint $\mathrm{E}C \leq \xi$, where $\xi \in [0, \infty)$ is a daily budget.

Let Ω denote the set of all eligible impression opportunities for the day. Impressions $i \in \Omega$ are awarded via sequential real time bidding in an open impression exchange, and sold in the manner of a first or second price cost model [18]. The cost model for each impression opportunity is decided by the seller of the impression and is known before a bid is computed. For a first price impression ($i \in \Omega_1$) the winner pays an amount equal to its own bid, whereas for a second price impression ($i \in \Omega_2$) the winner pays an amount equal

[1]N. Karlsson (nkarlsson8 at gmail.com) is Senior Principal Research Scientist with Amazon Ads, Palo Alto, California, USA.

to the second highest bid. Assume Let $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Omega_1 \cup \Omega_2 = \Omega$. The decision variables of the problem are given by the bid prices $b_i \in [0, b_{max}]$, for all $i \in \Omega$, where $b_{max} > 0$ is an advertiser specified *max bid*. Let the expected value of impression i (if awarded) be denoted $v_i \in [0, \infty)$. It represents the expected performance and/or branding value, and may encode the probability of a click or conversion (a sale), and the branding value of reaching prospective future customers with the advertisement message.

The optimization problem is mathematically defined by

$$\underset{\{b_i | 0 \leq b_i \leq b_{max}, \forall i \in \Omega\}}{\text{maximize}} \quad \mathrm{E}V - \alpha \mathrm{E}C \tag{1}$$

subject to $\mathrm{E}C \leq \xi$. An optimal bidding strategy for this problem is well-known (see e.g. [16], [17]) and satisfies

$$b_i^{opt} = \begin{cases} \underset{0 \leq b \leq b_{max}}{\text{argmax}} \ (b_i^u - b)F_i(b), & i \in \Omega_1, \\ b_i^u, & i \in \Omega_2, \end{cases}$$

where the *private value*, b_i^u is defined by

$$b_i^u = \frac{v_i}{\lambda}; \tag{2}$$

and where $\mathrm{E}C \leq \xi$, $(\lambda - \alpha)(\mathrm{E}C - \xi) = 0$, and $\lambda \geq \alpha$. Function $F_i(b)$ is the cumulative density of the highest competing bid price (the win rate function), and λ is a unitless shadow price. Assume $F_i(b)$ and v_i are known (produced by a different system). It remains to find the value of λ, which is independent of individual impressions.

It is easily shown that $\mathrm{E}V$ and $\mathrm{E}C$ are monotonic decreasing functions of λ, and that the optimal λ equals the smallest value, but no smaller than α, for which $\mathrm{E}C \leq \xi$ [8].

Impressions Ω do not occur all at once but over time, which suggests a feedback control solution might be applicable. The goal is to compute a *price control* signal $\lambda(t)$ that converges sufficiently fast towards a constant that satisfies $(\lambda - \alpha)(\mathrm{E}C - \xi) = 0$ and $\mathrm{E}C \leq \xi$ at the end of the day.

Consider a time-sampled implementation of the control system with equidistant time points indexed $t = 1, 2, \ldots$, where the sampling time Δ is, for example, 1/30 hours. The number of samples per day equals $T = 24/\Delta$. Distribute the daily budget uniformly throughout the day as $\bar{u}_c(t) = \xi/T$, and let this represent the *command signal*. Furthermore, let the observed spend in time interval t be denoted $y(t)$. It follows that $\xi = \sum_{t=1}^{T} \bar{u}_c(t)$ and $c = \sum_{t=1}^{T} y(t)$, where c is an observation of the daily spend C (note, C is random).

Although one might consider designing a feedback controller that adjusts only the scalar price control signal $\lambda(t)$ towards its optimal value, this turns out to be a remarkably difficult problem [8]. The challenge stems from a highly nonlinear and uncertain relationship between λ and y, which makes the robustness versus performance trade-off difficult. To ensure stability, this approach often requires a conservative control system with low performance.

Therefore, consider a complementary *rate control* signal $u(t) \in [0, 1]$, which is a lever that indiscriminately and randomly throttles impression opportunities. For example, if $u(t) = 0.8$, then for each individual impression opportunity, a bid is submitted with a 0.8 probability. In other words, $u(t)$ is the probability to participate in the bidding.

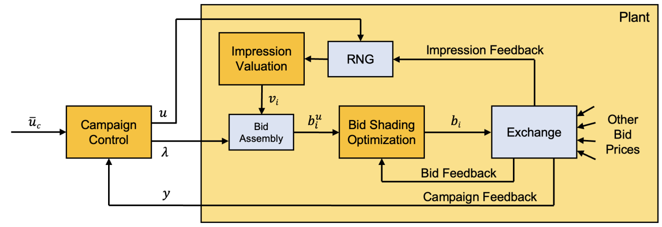

Fig. 1. The closed loop bid optimization system from the vantage point of the feedback controller, where impression valuation and bid shading optimization in charge of computing v_i and b_i are parts of the plant.

Figure 1 depicts the closed loop bid optimization system from the perspective of the feedback controller. Impression valuation and bid shading optimization, which compute v_i and b_i (outside the scope of the paper), are components of the plant and the details of these sub systems are irrelevant for as long as a useful model of the relationship from control signals $u(t)$ and $\lambda(t)$ to plant output $y(t)$ is available.

The control problem now is to adjust $u(t)$ and $\lambda(t)$ simultaneously toward their optimal values. Note, the optimal value of $u(t) = 1$, but allowing it to operate slightly away from one leads to a less difficult control problem.

III. Plant Model

The plant is non-linear, time-varying, and stochastic; but is typically subject to only a short delay. Assume a one-step delay. It follows that $y(t+1) = f(t, u(t), \lambda(t), w(t))$, where $w(t)$ is stochastic noise. The (unknown) function f is monotonic increasing in $u(t)$, monotonic decreasing in $\lambda(t)$, and produces a non-negative output.

The explicit time dependency of $f(\cdot)$ is due to a time-varying Internet traffic and competitive landscape. Assume it is T-periodic, scale-invariant, and independent of all other inputs. Since bid throttling is applied indiscriminately, we may assume $y(t+1)$ is proportional to $u(t)$. On the other hand, the relationship between $\lambda(t)$ and $y(t+1)$ is nonlinear, poorly known, and potentially discontinuous. Finally, assume the stochastic noise is scale-invariant and white. Combined, the plant model is given by

$$y(t+1) = (1 + h(t))\tilde{f}(\lambda(t))u(t)(1 + w(t)), \tag{3}$$

where *seasonality function* $h(t+T) = h(t)$, $\sum_{t=1}^{T} h(t) = 0$, and $h(t) > -1$, $\forall t$, $\tilde{f}(\cdot)$ is non-negative and monotonic decreasing, and $w(t)$ is mean zero white noise. The time-periodicity $h(t)$ can be estimated with reasonable accuracy in advance, and enhanced online via standard system identification techniques [19]. Here we assume $h(t)$ to be known.

IV. Control Design

We propose a control design that involves a T-periodic feedforward controller, a plant gain estimator, and two separate, but connected, feedback controllers.

First, the feedforward controller adjusts the uniform budget allocation encoded in $\bar{u}_c(t)$ throughout the day based on the plant periodicity and computes an adjusted command signal $u_c(t)$. Thereafter, the rate feedback controller adjusts $u(t)$ dynamically such that $y(t) \to u_c(t)$. In the process,

$u(t) \to u_e(t)$, where $u_e(t)$ is the steady-state bid rate control. Meanwhile, the plant gain estimator identifies the input-output sensitivity of the plant (relative the rate control signal). The plant gain estimate is used by the rate feedback controller to ensure desired robustness versus performance trade-off. A somewhat slower price feedback controller has a dead-zone; i.e., the control signal is updated only if the target signal is outside a prescribed interval, with hysteresis. This controller adjusts price control signal $\lambda(t) \in [\lambda_{min}, \lambda_{max}]$ such that $u_e(t)$ converges to a neighborhood of $u_t \in [u_l, u_h]$, which is a configured target bid rate control signal.

Typically, $\lambda_{min} \equiv \alpha$, which provides an interpretation of the lower bound of $\lambda(t)$ as the cost-discount parameter of objective function (1) in the original optimization problem. The upper bound λ_{max} is introduced on practical grounds to avoid an unnecessarily wide range of price control values. Scenarios where a larger value of $\lambda(t)$ is otherwise warranted, may be handled by a reduced value of $u(t)$.

Figure 2 shows a block diagram of the control system

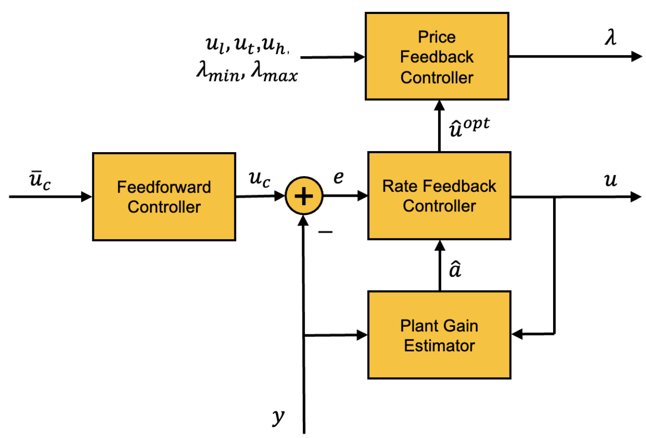

Fig. 2. Block diagram of the control system with the four main components and their interfaces.

with the four main components and their interfaces. Each component is described in detail in the following subsections.

A. Feedforward Controller

The feedforward controller computes an adjusted command signal $u_c(t)$ from $\bar{u}_c(t)$ according to

$$u_c(t) = \big(1 + h(t-1)\big)\bar{u}_c(t). \tag{4}$$

As shown next, this makes stabilization of the closed loop system possible. The idea is that the daily budget via $u_c(t)$ is distributed throughout the day to make convergence of the control signals possible.

B. Rate Feedback Controller

The plant in expected sense ($\mathrm{E}w(t) = 0$) is given by $y(t+1) = (1 + h(t))\tilde{f}(\lambda(t))u(t)$. Assume $\tilde{f}(\lambda(t))$ by virtue of how $\lambda(t)$ is updated is at most slowly varying. We may then adopt the certainty equivalence principle [20] during the rate control design and pretend $\tilde{f}(\lambda(t)) := a$ (the *plant gain*) is constant. It follows that

$$y(t+1) = \big(1 + h(t)\big)au(t). \tag{5}$$

Assume for a moment that a is known. To serve as a proof of concept, and to permit stability analysis (see [8]), consider a pure integral error feedback controller, which in recursive form satisfies

$$u(t) = u(t-1) + c_{\mathrm{I}}e(t), \tag{6}$$

where c_{I} is the *integral gain*, and the error signal $e(t)$ is

$$e(t) = u_c(t) - y(t). \tag{7}$$

Since $u(t)$ cannot take values outside the interval $[0, 1]$, in an implementation each computed value $u(t)$ is projected to $[0, 1]$ to prevent impossible bid rate control signal values from being produced. Combine (5)-(7) to obtain

$$\begin{aligned} u(t) &= u(t-1) + c_{\mathrm{I}}\big(u_c(t) - y(t)\big) \\ &= u(t-1) + c_{\mathrm{I}}\Big(u_c(t) - \big(1 + h(t-1)\big)au(t-1)\Big). \end{aligned} \tag{8}$$

Rearrange (8) and substitute $u_c(t)$ with (4) to obtain

$$u(t) = \psi(t)u(t-1) + \phi(t)\bar{u}_c(t), \tag{9}$$

where

$$\begin{aligned} \psi(t) &= 1 - c_{\mathrm{I}}a\big(1 + h(t-1)\big) \quad \text{and} \\ \phi(t) &= c_{\mathrm{I}}\big(1 + h(t-1)\big). \end{aligned}$$

Equation (9) is a first order T-periodic linear difference equation, and it can be shown (Theorem 3, [8]) it is globally asymptotically stable if and only if $|\prod_{t=1}^{T} \psi(t)| < 1$. A sufficient condition for global asymptotic stability is

$$0 < c_{\mathrm{I}}a < \frac{2}{1 + h(t)}, \qquad \text{for all } t, \tag{10}$$

(Corollary 1, [8]). If the stability condition is met and $a \geq \bar{u}_c(t)$ (via adjustments to $\lambda(t)$), then

$$u(t) \to \frac{\phi(t)\bar{u}_c(t)}{1 - \psi(t)} = \frac{\bar{u}_c(t)}{a} \in [0, 1], \quad \big(\equiv u_e(t)\big). \tag{11}$$

It follows from (5) and (11) that $y(t) \to (1+h(t-1))\bar{u}_c(t)$, which implies that $\mathrm{E}C = \sum_{t=1}^{T} y(t) \to \sum_{t=1}^{T}(1 + h(t-1))\bar{u}_c = \sum_{t=1}^{T}(1 + h(t-1))\xi/T = \xi$.

In light of the stability condition, a sensible design parameter is given by the *closed loop gain* $c_{\mathrm{CL}} := c_{\mathrm{I}}a$ that is chosen to satisfy (10), whereas the integral gain is computed from $c_{\mathrm{I}} = c_{\mathrm{CL}}/a$. In practice a is unknown and must be estimated, which is the topic of next section.

Besides computing $u(t)$, which by design evolves gracefully to ensure stability under noise and uncertainty, a delay-free estimate of the steady state bid rate control is given by

$$\hat{u}^{opt}(t) = \frac{\bar{u}_c(t)}{a} \in [0, \infty).$$

It is distinguished from $u_e(t)$ in that the latter is the actual steady state value bounded between zero and one, while $\hat{u}^{opt}(t)$ is a fictional rate control signal that may exceed one.

C. Rate Plant Gain Estimator

The rate feedback controller makes updates to $u(t)$ based on $e(t)$ and c_{I}, but to ensure stability c_{I} must be chosen relative a, which in practice is unknown and may vary over time. Let $\hat{a}$ denote an estimate of a, and consider plant model (3), which can be expressed as

$$y(t) \quad = \quad (1+h(t-1))u(t-1)\big(a+aw(t-1)\big),$$

Technically, for as long as $u(t-1) > 0$, an unbiased estimate of a can be computed based on one pair of $u(t-1)$ and $y(t)$, by recognizing that

$$\tilde{y}(t) \quad = \quad a+\epsilon(t),$$

where $\epsilon(t)$ is mean zero white noise, and

$$\begin{aligned} \tilde{y}(t) &= \frac{y(t)}{(1+h(t-1))u(t-1)}, \\ \epsilon(t) &= aw(t-1). \end{aligned}$$

Indeed, a possible estimate is $\hat{a} = \tilde{y}(t)$; however, such an estimate would have a very high variance. To account for noise and other plant uncertainties, we propose a recursive estimator defined by

$$\hat{a}(t) = \begin{cases} \gamma\hat{a}(t-1) + \dfrac{(1-\gamma)y(t)}{(1+h(t-1))u(t-1)}, & \text{if } u(t-1) > 0, \\ \hat{a}(t-1), & \text{otherwise,} \end{cases}$$

where $\gamma \in [0,1)$ is a *plant gain estimation forgetting factor*.

D. Price Feedback Controller

The responsibility of the price feedback controller is to adjust $\lambda(t) \in [\lambda_{min}, \lambda_{max}]$ such that $u_e(t)$ converges to a neighborhood of the *target bid rate control signal* $u_t \in [u_l, u_h]$. In order to do so with an element of foresight, the controller is designed to drive $\hat{u}^{opt}(t)$ towards u_t. The plant as perceived by this controller is the mapping from $\lambda(t)$ to $\hat{u}^{opt}(t)$, which can be shown to be monotonic increasing, but otherwise heavily dependent on the nonlinear and difficult to estimate $\tilde{f}(\lambda(t))$.

In consideration of the high degree of plant uncertainty, and to reduce the interference with the rate feedback controller near the steady-state, we propose a design, where the price controller starts making updates to $\lambda(t)$ only if $\hat{u}^{opt}(t)$ departs from the interval $[u_l, u_h]$, and then continues making updates until $\hat{u}^{opt}(t)$ reaches u_t, which is an interior point of the interval. This control strategy makes the design a dead-zone controller with hysteresis. Figure 3 illustrates the concept of ideal operating regime. The control system strives to operate with control signals $u(t)$ and $\lambda(t)$ in the green-shaded area.

In this paper, each update to $\lambda(t)$ is a fixed multiplicative adjustment δ or $1/\delta$, where $\delta > 1$ is close to one. Not making an adjustment of $\lambda(t)$ relative to the difference $u_t - \hat{u}^{opt}(t)$ reflects the assumed nonlinear and poorly understood plant. However, assume we know how to select δ to avoid instability. This typically requires a conservative choice.

Let $\eta(t) \in \{-1, 0, 1\}$ denote an internal *bid heading* state. It is set to -1/+1 when the controller detects that the bid price needs to decrease/increase and $\lambda(t)$ needs to increase/decrease. This happens when $\hat{u}^{opt}(t)$ exits $[u_l, u_h]$. If $\eta(t)$ is set to -1 or +1, then it will remain at that value until $\hat{u}^{opt}(t)$ reaches u_t, at which point $\eta(t)$ is set to zero.

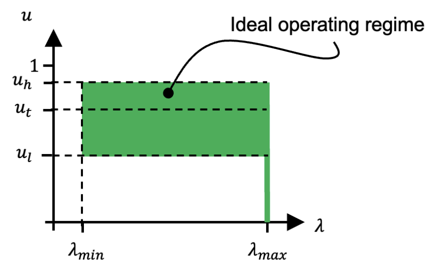

Fig. 3. Illustration of ideal operating regime. The control system strives to operate with control signals $u(t)$ and $\lambda(t)$ in the green-shaded area.

The control signal update is described mathematically as follows: First, if $\big(\eta(t-1) = -1$ and $\hat{u}^{opt}(t) \geq u_t\big)$ or $\big(\eta(t-1) = 1$ and $\hat{u}^{opt}(t) \leq u_t\big)$, then $\eta(t) = 0$. Thereafter,

$$\eta(t) = \begin{cases} -1, & \text{if } \hat{u}^{opt}(t) < u_l, \\ +1, & \text{if } \hat{u}^{opt}(t) > u_h. \end{cases}$$

Finally, increase $\lambda(t)$ by a factor δ, if $\eta(t) = -1$; and by a factor $1/\delta$, if $\eta(t) = 1$.

$$\lambda(t) \quad = \quad \text{Proj}_{[\lambda_{min}, \lambda_{max}]}\Big(\delta^{-\eta(t)}\lambda(t-1)\Big),$$

where the projection operator Proj caps its argument from above and below to remain in the interval $[\lambda_{min}, \lambda_{max}]$.

V. Implementation

The controller derived in Section IV is summarized in Algorithm 1. A sampling time-invariant implementation is obtained by defining parameters $\gamma, c_{\text{CL}}, \delta$ in terms of the sampling time Δ, time constants $\tau_{\text{CL}}, \tau_\gamma$, and max hourly bid price increase δ_{hr}. In particular, $\gamma = \exp(-\tau_\gamma/\Delta)$, $c_{\text{CL}} = \exp(-\tau_{CL}/\Delta)$, and $\delta = \delta_{hr}^{1/\Delta}$. Assume $h(t) = h_\theta(t)$ is fully determined by θ, a known *seasonality parameter*.

VI. Simulation Results

Based on first principle reasoning and real-world observations, [15], [19], consider a plant where $\Delta = 1/30$ hrs,

$$h(t) \quad = \quad 0.58\sin\left(\frac{2\pi t}{T} - 1.6\right) + 0.32\sin\left(\frac{4\pi t}{T} - 1.5\right),$$

and

$$\tilde{f}(\lambda) \quad = \quad \frac{1000(\pi/2 + \arctan(-(\log(\lambda)+20)/2))}{\pi}.$$

Time t is now in the unit of hours. It is no longer just an integer index.

Assume the daily budget changes over time according to

$$\xi = \begin{cases} 135, & \text{if } 0 \leq t \leq 72, \\ 300, & \text{if } 72 < t \leq 120, \\ 80, & \text{if } 120 < t \leq 216, \\ 200, & \text{if } 216 < t \leq 288, \\ 130, & \text{if } 288 < t \leq 384, \\ 400, & \text{if } 384 < t \leq 408. \end{cases}$$

Consider two simulated scenarios.

Algorithm 1 Simultaneous Rate and Price Control

Parameters: $\theta, c_{\text{CL}}, \gamma, \lambda_{min}, \lambda_{max}, u_l, u_t, u_h, \delta$
Input: $\bar{u}_c, y$
Output: λ, u
State: $u, \lambda, \hat{a}, \eta$
$u(0) = u_0$, $\lambda(0) = \lambda_0$, $\hat{a}(0) = \hat{a}_0$, $\eta(0) = \eta_0$
for $t = 1, 2, \dots$ **do**
  $u_c(t) = \big(1 + h_\theta(t-1)\big)\bar{u}_c(t)$
  if $u(t-1) > 0$ **then**
  $\hat{a}(t) = \gamma\hat{a}(t-1) + (1-\gamma)y(t)/[(1+h_\theta(t-1))u(t-1)]$
  else
    $\hat{a}(t) = \hat{a}(t-1)$
  $e(t) = u_c(t) - y(t)$
  $c_{\text{I}} = c_{\text{CL}}/\hat{a}(t)$
  $u(t) = \text{Proj}_{[0,1]}\big(u(t-1) + c_{\text{I}}e(t)\big)$
  $\hat{u}^{opt}(t) = \bar{u}_c(t)/\hat{a}(t)$
  if $(\eta(t-1) = -1$ and $\hat{u}^{opt}(t) \geq u_l)$ or
    $(\eta(t-1) = 1$ and $\hat{u}^{opt}(t) \leq u_l)$ **then**
    $\eta(t) = 0$
  if $\hat{u}^{opt}(t) < u_l$ **then**
    $\eta(t) = -1$
  else if $\hat{u}^{opt}(t) > u_h$ **then**
    $\eta(t) = 1$
  $\lambda(t) = \text{Proj}_{[\lambda_{min},\lambda_{max}]}\big(\delta^{-\eta(t)}\lambda(t-1)\big)$

A. Example 1

Let the plant be noise free ($w(t) \equiv 0$), and the controller be configured by $\gamma = 0.5$, $\tau_{\text{CL}} = 1.5$, $\tau_\gamma = 0.75$, $\delta_{hr} = 2.7$, $u_l = 0$, $u_t = 0.85$, $u_h = 1$, $\lambda_{min} = 6.3e-16$, $\lambda_{max} = 1$.

A representative closed loop result is shown in Figure 4. The top-left panel displays the daily budget (black dashed curve) and the intraday cumulative spend (red solid curve). It is noted that the campaign spends each daily budget at the end almost every day of the 17 days long flight in spite of five budget changes. The bottom-left panel shows the adjusted command signal $u_c(t)$ (black dashed curve), which is the budget allocated to each sampling interval; and the observed spend $y(t)$ (red solid curve). As shown, the observed spend accurately tracks the adjusted command signal. The top-right panels displays the predicted optimal bid rate $\hat{u}^{opt}(t)$ (black curve) and the bid rate control $u(t)$ (red curve). Note how, for each budget level, $u(t)$ and $\hat{u}^{opt}(t)$ have the same steady-state value, but that $\hat{u}^{opt}(t)$ moves faster. Finally, the bottom-right panel shows the price control signal $\lambda(t)$, which in this example stays at a constant value since the configured ideal operating regime of the rate control signal $[u_l, u_h] = [0, 1]$. Indeed, as long as campaign is able to deliver the budget using a bid rate $u(t) \in [0, 1]$, no adjustment of $\lambda(t)$ is triggered. The control system is highly responsive to budget changes. This is thanks to a fast bid rate feedback controller, which is made possible by an accurate estimation of the corresponding plant gain.

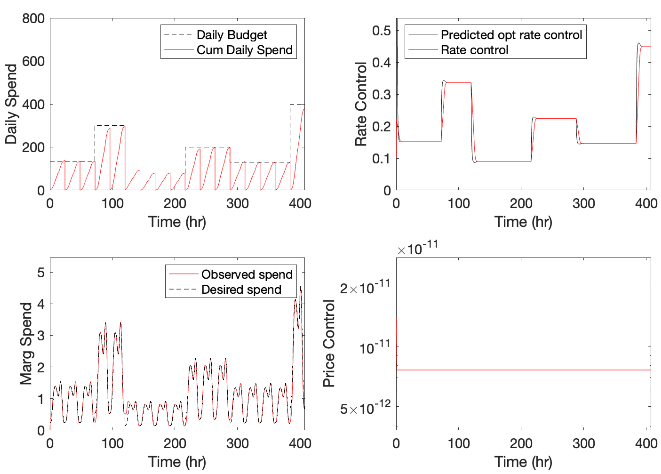

Fig. 4. Top-left: daily budget (black dashed curve) and intraday cumulative spend (red solid curve). Bottom-left: adjusted command signal $u_c(t)$ (black dashed curve) and observed spend $y(t)$ (red solid curve). Top-right: predicted optimal bid rate $\hat{u}^{opt}(t)$ (black curve) and bid rate control $u(t)$ (red curve). Bottom-right: price control signal $\lambda(t)$.

B. Example 2

Now assume the plant is subject to noise given by $w \sim \text{Gaussian}(0, 0.1^2)$, and let the controller be configured by $\gamma = 0.5$, $\tau_{\text{CL}} = 1.5$, $\tau_\gamma = 0.75$, $\delta_{hr} = 2.7$, $u_l = 0.6$, $u_t = 0.85$, $u_h = 1$, $\lambda_{min} = 6.3e-16$, $\lambda_{max} = 1$

Figure 5. The top-left panel demonstrates how the campaign delivers each daily budget in full almost exactly at the end of each day. Furthermore, the bottom-left panel shows how the observed spend $y(t)$ (though being noisy) tracks the adjusted command signal well. The top-right panels displays the predicted optimal bid rate $\hat{u}^{opt}(t)$ and the bid rate control $u(t)$. It is noted that $\hat{u}^{opt}(t)$ occasionally exceeds one, and that $u(t)$ almost always operate within $[u_l, u_h]$ and near the target bid rate u_t. Finally, the bottom-right panel shows the price control signal $\lambda(t)$, which is subject to adjustments only intermittently. For long stretches of time, $\lambda(t)$ is constant. Keep in mind that the price control signal $u(t)$, due to (2), acts like a break pedal. The larger its value, the smaller is the final bid prices submitted by the campaign for impression opportunities.

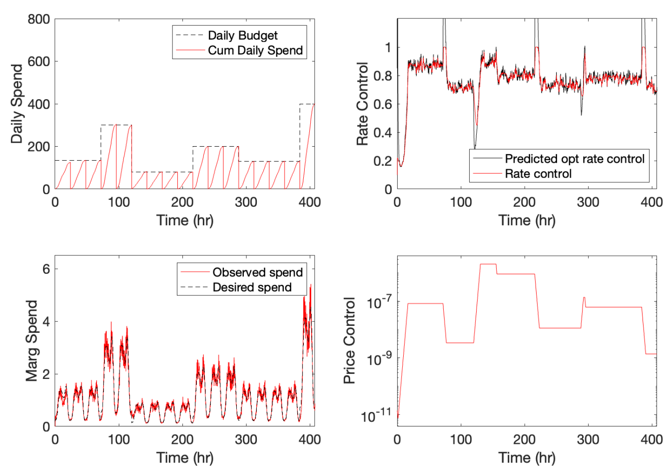

Fig. 5. Top-left: daily budget (black dashed curve) and intraday cumulative spend (red solid curve). Bottom-left: adjusted command signal $u_c(t)$ (black dashed curve) and observed spend $y(t)$ (red solid curve). Top-right: predicted optimal bid rate $\hat{u}^{opt}(t)$ (black curve) and bid rate control $u(t)$ (red curve). Bottom-right: price control signal $\lambda(t)$.

In summary, the control system is highly responsive to budget changes and noise, thanks to a fast bid rate feedback controller, made possible by an accurate estimation of the corresponding plant gain. Near optimal bidding is achieved by the price feedback controller that adjusts $\lambda(t)$ in such a way that the bid rate control signal operates near one.

VII. Conclusions and Future Work

We have proposed a dual lever feedback control algorithm for optimization and pacing of online advertising campaigns. It involves a periodic feedforward controller, a plant gain estimator, and two separate feedback controllers. The first and fast feedback controller adjusts a throttling lever $u(t)$ to make an observed spend output $y(t)$ rapidly track a desired reference signal $u_c(t)$. A separate and less aggressive feedback controller adjust a price control signal $\lambda(t)$ to drive the steady-state value of $u(t)$ into a desired operating interval $[u_l, u_h]$ (typically near one). The price controller is a dead-zone controller with hysteresis.

The plant on which the control system operates is periodic, nonlinear, and stochastic; but the introduction of a companion throttling control signal creates a linearity that is exploited for efficient and robust control. Stability results by the author in prior work has been leveraged to support the proposed algorithm.

Simulation results demonstrate stability and good performance under a wide range of operating conditions, including a nonlinear plant, significant noise, and budget changes.

Immediate future work includes the experimental validation of the proposed control system. Real advertising campaigns are difficult to model, hence, simulation results cannot capture all important and complex behaviors that are present in the real plant.

Future work also includes the design of a more sophisticated price control algorithm. The algorithm in the present paper is naive and makes almost no use of information that may be available about the relationship between $\lambda(t)$ and $y(t)$. Incorporating such information, for example via a price plant gain estimate (an estimate of $\partial y(t)/\partial\lambda(t)$), is likely to offer significant improvement in terms of campaign performance and versatility of the control system. Other future work includes identifying the seasonality function $h(t)$, which was assumed known in this paper, and accounting for the imperfection of any estimate $\hat{h}(t)$ by adequate calibration of the control system design parameters.

Furthermore, since advertisers are becoming increasingly sophisticated and expect their campaigns to satisfy many constraints beyond the budget, it is important to generalize the application of the proposed control system to multi-constraint optimization problems, and to produce stability results for this more general setup.

VIII. Acknowledgments

I thank Neal Richter and Sebastian Bock at Amazon Ads for inspiring and thought provoking discussions on the merits of throttling-based control in online advertising, and the Amazon senior leadership team for nurturing an environment of innovation and for recognizing the importance of science publishing as a means to raise the bar and to attract talent.

References

[1] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. Deep & cross network for ad click predictions. *CoRR*, abs/1708.05123, 2017.

[2] Junwei Pan, Jian Xu, Alfonso Lobos Ruiz, Wenliang Zhao, Shengjun Pan, Yu Sun, and Quan Lu. Field-weighted factorization machines for click-through rate prediction in display advertising. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, pages 1349–1357, 2018.

[3] Yuya Yoshikawa and Yusaku Imai. A nonparametric delayed feedback model for conversion rate prediction. *CoRR*, abs/1802.00255, 2018.

[4] Deepak Agarwal, Souvik Ghosh, Kai Wei, and Siyu You. Budget pacing for targeted online advertisements at linkedin. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1613–1619, New York, NY, USA, 2014.

[5] Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. *Proceedings of the 9th ACM International Conference on Web Search and Data Mining, New York, NY, USA*, pages 407–416, 2018.

[6] Niklas Karlsson. Control of periodic systems in online advertising. In *Proceedings of the 2018 IEEE 57th Conference on Decision and Control, Miami, FL, USA*, Dec 17-19 2018.

[7] Niklas Karlsson. Adaptive optimization and control in online advertising. In *Proceedings of the 2019 IEEE 58th Conference on Decision and Control, Nice, France*, Dec 11-13 2019.

[8] Niklas Karlsson. Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77, 2020.

[9] Djordje Gligorijevic, Tian Zhou, Bharatbhushan Shetty, Brendan Kitts, Shengjun Pan, Junwei Pan, and Aaron Flores. Bid shading in the brave new world of first-price auctions. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, Oct 2020.

[10] Shengjun Pan, Brendan Kitts, Tian Zhou, Hao He, Bharatbhushan Shetty, Aaron Flores, Djordje Gligorijevic, Junwei Pan, Tingyu Mao, San Gultekin, and Jianlong Zhang. Bid shading by win-rate estimation and surplus maximization. 2020.

[11] Niklas Karlsson and Qian Sang. Adaptive bid shading optimization of first-price ad inventory,. In *Proceedings of the 2021 American Control Conference, New Orleans, LA, USA*, ACC '21, pages 4971–4978, May 26–28 2021.

[12] Tian Zhou, Hao He, Shengjun Pan, Niklas Karlsson, Bharatbhushan Shetty, Brendan Kitts, Djordje Gligorijevic, San Gultekin, Tingyu Mao, Junwei Pan, Jianlong Zhang, and Aaron Flores. An efficient deep distribution network for bid shading in first-price auctions. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'21), August 14-18, 2021, Singapore*, 2021.

[13] Niklas Karlsson and Jianlong Zhang. Applications of feedback control in online advertising. In *Proceedings of the 2013 American Control Conference, Washington, DC, June 17 - 19*, pages 6008 – 6013, 2013.

[14] Niklas Karlsson. Control problems in online advertising and benefits of randomized bidding strategies. *European Journal of Control*, 30:31—49, July 2016.

[15] Niklas Karlsson. Adaptive control using Heisenberg bidding. In *Proceedings of the 2014 American Control Conference, Portland, USA, June 4-6, 2014*, pages 1304–1309, 2014.

[16] Niklas Karlsson. Scalable multi-objective optimization in programmatic advertising via feedback control. In *Proceedings of the 2021 60th IEEE Conference on Decision and Control (CDC), Austin, TX, USA*, pages 1363–1370, Dec 2021.

[17] Niklas Karlsson. Feedback control-based multiobjective optimization in programmatic advertising involving a cost per bid constraint. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 505–510, 2022.

[18] Vijay Krishna. *Auction Theory*. Academic Press, 2002.

[19] Hao He and Niklas Karlsson. Identification of seasonality in internet traffic to support control of online advertising. In *2019 American Control Conference (ACC)*, pages 3835–3840, 2019.

[20] Karl Johan Åström and Björn Wittenmark. *Adaptive Control (2 ed.)*. Addison-Wesley, 1995.