Optimal Bidding for Mobile-Ad Campaigns
Research in Progress

Manmohan Aseri
University of Texas at Dallas
manmohan.aseri@utdallas.edu

Milind Dawande
University of Texas at Dallas
milind@utdallas.edu

Ganesh Janakiraman
University of Texas at Dallas
ganesh@utdallas.edu

Vijay Mookerjee
University of Texas at Dallas
vijaym@utdallas.edu

Abstract

Self-service advertising platforms such as Cidewalk enable advertisers to directly launch their individual mobile advertising campaigns. These platforms contract with advertisers to provide a certain number of impressions on mobile apps in a specific geographic location (usually a town or a zip code) within a fixed time period (usually a day); this is referred to as a campaign. To meet the commitment for a campaign, the platform bids on an ad-exchange to win the required number of impressions from the desired area within the time period of the campaign. We address the platform's problem of deciding its bidding policy to minimize the expected cost in fulfilling the campaign.

Key Words: *Internet Advertising, Self-Service Platforms, Mobile-Ad Campaigns, Optimal Bidding.*

1. Introduction

The increase in the use of online media – personal computers, mobile phones, tablets, etc. – for advertising has been tremendous over the past decade (Central Market Research, 2012; Lieberman, 2013). The rate of growth of online advertising has been particularly impressive in the past five years; e.g., the total revenue from online advertising in the United States in 2013 was about 17% more than that in 2012 (Interactive Advertising Bureau, 2014). In the United Kingdom, Internet advertising revenue increased 12.5% over that in 2011 (Interactive Advertising Bureau UK 2013a,b). It is estimated that revenues from Internet advertising will reach $76 billion by 2016 (Hof, 2011; eMarketer, 2012).

The focus of the current study is on advertising on a mobile device (e.g., a smart phone, or a tablet). Specifically, we consider ads that are displayed on a mobile application (hereafter abbreviated as an app), such as an app for weather, stocks, or a game. This form of advertising is on the increase. In addition to the end user of the mobile app, there are at least two other parties that are involved in the process of mobile in-app advertising: (i) The advertiser, who provides ads created to promote products or services, and (ii) the publisher, or the owner of the mobile app. The advertiser and the publisher often interact with one another through their respective agents. Advertisers (who ultimately generate demand for advertising space) are usually represented by *demand-side* agents (or ad aggregators). Demand-side agents provide advertisers the access to a variety of publishers for the appropriate exposure of their ads. On the other hand, *supply-side* agents represent publishers that supply the space for ad-display. Supply-side agents help monetize the space owned by publishers and earn revenue for them.

Till recently, most of the supply of ad space from apps was sold on a *contractual* basis. That is, supply-side agents entered into relatively long-term contractual arrangements with app owners to sell their space for the

display of ads. These contracts were often drawn on a revenue-sharing basis, i.e., app owners got a proportion of the revenue generated from their ad-space. In recent times, with the advent of *mobile ad exchanges* (e.g., Nexage, AdMarvel, etc.) the supply of ad-space on apps is fast becoming a commodity. There are no long-term contracts; rather, each opportunity to display an ad (called an *impression*) is auctioned off on a mobile ad exchange. Impressions are typically sold on a *cpm* (cost per thousand impressions) basis.

The growth of mobile ad exchanges has also led to changes in the demand-side of the industry. Demand aggregators are now able to directly buy supply from exchanges rather than acessing the supply via a supply aggregator. While the advertising needs of large advertisers (such as GE or SONY) are often addressed in an "ad agency" mode providing end-to-end service, demand aggregators are moving to a "self-service" mode for small advertisers. Here, advertisers directly launch their individual campaigns using a self-service platform. The self-service mode is suitable for thousands of small advertisers without deep pockets and offers an attractive advertising solution for a hitherto underserved segment of the market. A prominent example is Facebook that offers a self-service advertising solution, albeit the service is restricted to the supply within Facebook.

Self-service platforms scale well and provide access to supply via the use of ad exchanges. The key expertise underlying the platform is in the ability to bid intelligently for supply on a mobile ad exchange. Also, there needs to be a high level of integration between the ad exchange and the platform to complete the bidding process in real-time and render the winner's ad on a mobile app. Such technical expertise is usually not possessed by small advertisers. Hence there is a niche for firms that possess sophisticated integration skills and fast, real-time analytic abilities to buy supply at affordable prices and deliver ad campaigns at a net profit.

In this study, we consider the problem faced by one such ad firm (Cidewalk; http://www.cidewalk.com) of optimizing the bidding policy for mobile ad delivery to support self-service ad campaigns. Cidewalk enters into contracts with advertisers under which they have to provide a certain number of impressions (which we refer to as a "campaign") on mobile apps in a specific geographic location (usually a town or a zipcode) within a fixed time period (usually a day). Cidewalk bids on ad-exchanges to win these impressions: The more the bid, the higher is the probability of winning an impression. The objective is to win a contracted number of bids (to place impressions) over a given time period at minimum cost.

2. Model

The model we study is the following. The firm (Cidewalk, in the description above) has made a commitment to a customer to deliver C impressions, as part of a campaign within a certain period of time, to users of mobile apps in a certain geographical area. For example, a newly opened restaurant may want to run an ad-campaign targeting 10,000 potential customers within a certain zip-code within one week. Specifically, the customer is interested in displaying the ad on mobile devices when users open an app – we refer to the event of a mobile device user in the desired area opening an app as an *impression*. These impressions are auctioned in real-time in an ad-exchange. Advertisers or advertising companies who act on behalf of advertisers bid for these impressions. The highest bidder wins the impression and pays the price that she bids to the exchange. Clearly, the higher the bid, the greater the probability of winning an impression. We model this using a *win-curve*, that is a function $p(b) : [0, b_{max}] \to [0, 1]$ which specifies the probability of winning an impression by bidding an amount b. It is reasonable and convenient to assume that b_{max} is a large enough value that an impression will definitely be won with a bid of that value. In order to avoid the possibility of incurring a very high cost for procuring C impressions to meet a strict guarantee of delivering that many impressions to the customer, the firm specifies a probability α (very close to 1) and promises the customer that it will deliver C impressions with a probability of α or more (a typical value of α could be 0.99, i.e., 99%).[1] We use T to denote the number of *time slots* in the desired time period; for example, a time slot could be a millisecond. By *time slot*, we mean a sufficiently small interval of time in which the probability of more than one impression arriving from the desired geographical area is zero. Let q denote the probability that an impression from the desired area arrives in a time slot; thus, $1 - q$ is the probability that no such

---

[1]Since our firm and its customers expect to engage in many such campaigns over a year, customers will have the ability to assess whether the firm is meeting their "probabilistic guarantee".

impression arrives in a time slot. We focus on the firm's problem for a single campaign. The decisions are the bids to place on the impressions that (possibly) arrive from the desired area in time slots $\{1, 2, \ldots, T\}$. Notice that these bids can be *dynamic*; for example, if many impressions have been won early on within the campaign period, the firm may start bidding low on subsequent impressions. The objective is to minimize the expected cost while the constraint is to obtain C impressions with a probability of α or more. We denote this problem by $\mathcal{P}^{Prob}(C, \alpha)$ (the superscript "Prob" denotes the probabilistic constraint).

Below, we present a mathematical formulation of $\mathcal{P}^{Prob}(C, \alpha)$. To do this, it is useful to define $b(x) : [0, 1] \to [0, b_{max}]$ as the inverse of the function p; that is, $b(x) = p^{-1}(x)$. We also define $f(x) = x \times b(x)$; this is the expected cost associated with choosing a win-probability of x for an impression. It should be clear that our problem can also be formulated as one in which the decisions are the win-probabilities to use on the impressions that arrive. Now, a policy can be formally defined as a matrix $\mathbf{x}$ of win-probabilities, whose entries are $\{x(t, j_t) : 1 \le t \le T, 0 \le j_t \le t - 1\}$; here, $x(t, j_t)$ specifies the win-probability for the impression that could possibly arrive in time slot t if the firm has won j_t impressions in the first $t - 1$ time slots. If $x(t, j_t)$ is constant (i.e., independent of t and j_t), we refer to $\mathbf{x}$ as a static policy. In general, we refer to $\mathbf{x}$ as a dynamic policy. The stochastic process $\{j_t : 1 \le t \le T\}$, under a given policy $\mathbf{x}$, evolves as follows:

$$j_{t+1} = j_t + 1 \text{ if an impression arrives in slot } t \text{ and the bid of } x(t, j_t) \text{ wins impression } i,$$
$$= j_t \text{ otherwise.}$$

Then, $\mathcal{P}^{Prob}(C, \alpha)$ can be written as follows:

$$\min_{\mathbf{x}} \quad E\left[\sum_{t=1}^{T} q f(x(t, j_t))\right] \tag{1}$$

$$\text{s.t.} \quad P[j_{T+1} \ge C] \ge \alpha, \text{ and} \tag{2}$$

$$x(t, j_t) \in [0, 1] \text{ for all } 1 \le t \le T \text{ and all } j_t \in \{0, 1, \ldots, t - 1\}, \tag{3}$$

where j_{T+1} is the number of impressions won by the policy $\mathbf{x}$ at the end of the desired time period (i.e., end of slot T).

Our Goal: It is easy to see that the optimal policy for problem $\mathcal{P}^{Prob}(C, \alpha)$ will, in general, be a non-trivial function of the states (t, j_t); in other words, the optimal policy is *state-dependent*. In addition to being difficult to obtain, such a policy is also cumbersome to use since the number of impressions that arrive at the ad exchange over the desired time period (say, one day) is very high – and so is the number of states. Our goal in this paper is to obtain a policy that is *simple* to compute and is provably *near-optimal*. Indeed, the policy that will result from our analysis is *state-independent* and is described by a single number.

3. A Real-World Policy

As mentioned earlier, the advertising platform Cidewalk (Cidewalk, Inc. 2015) faces problem $\mathcal{P}^{Prob}(C, \alpha)$ when it bids for the impressions needed to satisfy a user's campaign. Cidewalk uses the following bidding policy:

Cidewalk's Policy: Bid $b\left(\min\left\{\frac{C - j_t}{q(T - t)}, 1\right\}\right)$ in state (t, j_t).

The numerator in the ratio above, i.e., $C - j_t$ is the remaining number of impressions to be won in state (t, j_t), while the denominator $q(T - t)$ is the expected number of impressions yet to arrive over the desired time period. This is clearly a dynamic policy; e.g., the bids become progressively aggressive as the expected number of impressions to arrive approaches the remaining number of impressions to be won. If the remaining impressions to be won exceeds the expected number of impressions to arrive in the remainder of the period, then the policy bids b_{max}.

In the next section, we will analyze a relaxation of the problem $\mathcal{P}^{Prob}(C, \alpha)$. We will obtain a state-independent optimal policy for this relaxation. Then, in Section (5), we will exploit this simple policy to obtain a feasible and near optimal policy for our original problem $\mathcal{P}^{Prob}(C, \alpha)$.

4. A Relaxation and Its Optimal Solution

We start by defining a new problem, similar to problem $\mathcal{P}^{Prob}(C, \alpha)$, in which the constraint is to deliver a certain number of impressions, say $\beta \geq 0$, *in expectation.* More formally, we denote by $\mathcal{P}^E(\beta)$ the following problem:

$$\min_{\mathbf{x}} \quad E\left[\sum_{t=1}^T qf(x(t, j_t))\right] \tag{4}$$

$$\text{s.t.} \quad E[j_{T+1}] \geq \beta \,, \quad \text{and} \tag{5}$$

$$x(t, j_t) \in [0, 1] \text{ for all } 1 \leq t \leq T \text{ and all } j_t \in \{0, 1, \dots, t-1\} \,. \tag{6}$$

Next, we know from Markov's inequality that

$$P\big[j_{T+1} \geq C\big] \leq \frac{E[j_{T+1}]}{C}.$$

Also, we need $P[j_{T+1} \geq C] \geq \alpha$; thus,

$$\alpha \leq P\big[j_{T+1} \geq C\big] \leq \frac{E[j_{T+1}]}{C}.$$

Therefore, the probabilistic guarantee in problem $\mathcal{P}^{Prob}(C, \alpha)$, that is, the constraint $P[j_{T+1} \geq C] \geq \alpha$ implies the inequality

$$E[j_{T+1}] \geq C\alpha. \tag{7}$$

Thus, using the choice $\beta = C\alpha$, Problem $\mathcal{P}^E(C\alpha)$ is a relaxation of Problem $\mathcal{P}^{Prob}(C, \alpha)$. We now proceed to solve this relaxed problem. We begin with an assumption on the win-curve, $p(\cdot)$.

Assumption 1 *The function $p(\cdot)$ is strictly increasing and concave.*

The following result is a consequence of Assumption 1. Its proof is a straightforward exercise in calculus and is, hence, omitted.

Claim 1 *Under Assumption 1, the functions $p^{-1}(\cdot)$ and $f(\cdot)$ are both increasing and convex.*

Next, we observe from the definition of j_t and $x(t, j_t)$ that $E[j_{T+1}] = \sum_{t=1}^T E[x(t, j_t)]$. Thus, Problem $\mathcal{P}^E(\beta)$ can be written as

$$\min_{\mathbf{x}} \quad E\left[\sum_{t=1}^T qf(x(t, j_t))\right] \tag{8}$$

$$\text{s.t.} \quad \sum_{t=1}^T E[x(t, j_t)] \geq \beta \,, \quad \text{and} \tag{9}$$

$$x(t, j_t) \in [0, 1] \text{ for all } 1 \leq t \leq T \text{ and all } j_t \in \{0, 1, \dots, t-1\} \,. \tag{10}$$

Using the convexity of f shown earlier and a standard Lagrangian analysis, we obtain the solution to Problem $\mathcal{P}^E(\beta)$ as follows:

$$x(t, j_t) = \frac{\beta}{Tq}, \text{ for all } t = 1 \dots T \,, \; j_t = 1 \dots (t-1). \tag{11}$$

We formally state this below.

Theorem 1 *The policy $x(t, j_t) = \frac{\beta}{Tq}$ is optimal for problem $\mathcal{P}^E(\beta)$.*

Note that the policy above is static and, moreover, is described by a single number. In the next section, we will use the above analysis to obtain a near-optimal solution to problem $\mathcal{P}^{Prob}(C, \alpha)$.

5. A Near-Optimal Solution to Problem $\mathcal{P}^{Prob}(C, \alpha)$

Motivated by the preceding analysis, our hope is to obtain a static policy that is feasible for problem $\mathcal{P}^{Prob}(C, \alpha)$ and has a near-optimal expected cost. Since the expected cost $f(x)$ is increasing in the associated win-probability x, we will aim to obtain a static policy with a win-probability x_α, i.e., $x(t, j_t) = x_\alpha; t = 1, \ldots, T; j_t = 1, \ldots, (t-1)$, such that Constraint (2) is tight. That is, $P[j_{T+1} \geq C] = \alpha$. With a (static) win-probability of x_α, the number of impressions won over T periods is a binomially-distributed random variable with a trial success probability of $\hat{p} = x_\alpha q$; let $\Phi_{Bin(T,\hat{p})}$ denote the c.d.f. of this distribution. Thus, to satisfy Constraint (2), we impose

$$\bar{\Phi}_{Bin(T,\hat{p})}(C) = \alpha. \tag{12}$$

Using a normal approximation to $Bin(T, \hat{p})$ (which is appropriate since T is large in practice), we obtain

$$\bar{\Phi}_N \left(\frac{C - Tx_\alpha q}{\sqrt{Tx_\alpha q(1 - x_\alpha q)}} \right) = \alpha,$$

where $\Phi_N(.)$ is the c.d.f. of the standard normal distribution. Thus, we obtain the following quadratic equation in x_α:

$$(C - Tx_\alpha q)^2 = z_\alpha^2 Tx_\alpha q(1 - x_\alpha q),$$

where $z_\alpha = \Phi_N^{-1}(\alpha)$. The relevant solution to this quadratic equation is

$$x_\alpha = \frac{(2C + z_\alpha^2) + \sqrt{(2C + z_\alpha^2)^2 - 4\left(1 + \frac{z_\alpha^2}{T}\right)C^2}}{2(Tq + z_\alpha^2 q)}.$$

In practice, C is of the order of thousands. Consequently, for practically relevant values of α (e.g., 0.99), we have $C \gg z_\alpha$. Therefore, the expression above for x_α can be approximated as

$$x_\alpha \approx \frac{C}{Tq} + \frac{z_\alpha \sqrt{C}}{Tq}.$$

We now proceed to show that the cost incurred by following the static policy with win-probability x_α is very close to the lower bound on the optimal cost of problem $\mathcal{P}^{Prob}(C, \alpha)$. Let $\text{Cost}(x_\alpha)$ represent the cost incurred by following the static policy with win-probability x_α and let $\text{Opt}(C, \alpha)$ be the optimal cost associated with problem $\mathcal{P}^{Prob}(C, \alpha)$. We calculate the ratio

$$\frac{\text{Cost}(x_\alpha)}{\text{Opt}(C, \alpha)}.$$

We know that $\text{Cost}(x_\alpha) = Tqf\left(\frac{C}{Tq} + \frac{z_\alpha \sqrt{C}}{Tq}\right)$ and $\text{Opt}(C, \alpha) \geq Tqf\left(\frac{C\alpha}{Tq}\right)$. Therefore, we have

$$\frac{\text{Cost}(x_\alpha)}{\text{Opt}(C, \alpha)} \leq \frac{f\left(\frac{C}{Tq} + \frac{z_\alpha \sqrt{C}}{Tq}\right)}{f\left(\frac{C\alpha}{Tq}\right)}. \tag{13}$$

Theorem 2 *The policy $x(t, j_t) = x_\alpha, \forall\, t, j_t$, achieves the following performance guarantee for problem $\mathcal{P}^{Prob}(C, \alpha)$:*

$$\frac{Cost(x_\alpha)}{Opt(C, \alpha)} \leq \frac{f\left(\frac{C}{Tq} + \frac{z_\alpha\sqrt{C}}{Tq}\right)}{f\left(\frac{C\alpha}{Tq}\right)}.$$

To assess the quality of this upper bound, consider the following example with realistic values of the parameters:

- Let $p(b) = \sqrt[3]{b}, 0 \leq b \leq 1$. Therefore, $b(x) = x^3$.
- Let each time slot be of duration 1 millisecond. Thus, corresponding to an eight-hour-campaign, the total number of time slots is $T = 259,200,000$.
- Assume one impression arrives every second. Thus, $q = \frac{1}{1000}$.
- Let the total number of impressions to be won over the campaign be 100,000. Thus, $C = 100,000$.
- Let $\alpha = 0.99$. Therefore, $z_\alpha = 2.326$.

For the above values of the parameters, the win-probability of our static policy is $x_\alpha = 0.38865$ and the value of the upper bound in (13) is about 1.072. That is, the total cost corresponding to the static policy with win-probability x_α is at most 7.2% higher than that of an optimal policy.

Remark: In the analysis above, to solve the feasibility equation (12) for our policy $x(t, j_t) = x_\alpha, \forall\, t, j_t$, it was convenient to use the normal approximation to the c.d.f. of the binomial distribution – this approximation resulted in a closed-form expression for the win-probability x_α. The approximation, however, may lead to the policy acquiring slightly fewer than the number of impressions required to guarantee feasibility. This infeasibility can be easily avoided if closed-form expressions are not required. One can directly use the c.d.f. of the binomial distribution to numerically solve $\bar{\Phi}_{Bin(T,\hat{p})}(C) = \alpha$ for the value of $\hat{p}$. If we use this direct approach in the numerical example above, then the value of the upper bound in (13) is 1.0721. That is, the total cost corresponding to the static policy $x(t, j_t) = x_\alpha$ is at most 7.21% (instead of 7.20% with the Normal approximation) higher than that of an optimal policy for $\mathcal{P}^{Prob}(C, \alpha)$. One can also ensure feasibility by using sharp bounds on the c.d.f. of the binomial distribution (see, e.g., Zubkov and Serov 2013) at the expense of losing our simple closed-form expressions.

The above analysis can be extended in a straightforward manner when the arrival probability depends on time; that is, the arrival probability is $q(t)$.

6. Ongoing Research

The problem formulated in this paper addresses the service platform's bidding policy for a *single* campaign. In practice, one would expect several campaigns that are simultaneously active in the same geographical area. Generalizing further, each of these campaigns could be interested in advertising in multiple geographical areas (e.g., towns or zip-codes). Thus, a multi-campaign, multi-location extension of the problem requires the platform to win a sufficient number of impressions from a *set* of locations for each campaign. We are actively working on this extension.

References

[Central Market Research 2012.] Central Market Research. Online Advertising One of the Fastest Growing Advertising Segments. *Available at:*
http://finance.yahoo.com/news/Online-Advertising-One-iw-991038561.html (Accessed May 4, 2015).

[Lieberman, D. 2013.] D. Lieberman. TV Advertising Is Surprisingly Weak Due To Internet And Economy: Analyst. *Available at:*
http://www.deadline.com/2013/03/tv-advertising-weakness-economy-internet-analyst-report/ (Accessed May 4, 2015).

[Hof, R. 2011.] R. Hof. Online Ad Spend to Overtake TV by 2016. *Available at: http://www.forbes.com/sites/roberthof/2011/08/26/online-ad-spend-to-overtake-tv/   (Accessed May 4, 2015).*

[eMarketer 2012.] eMarketer. US Online Advertising Spending to Surpass Print in 2012. *Available at: http://www.emarketer.com/Article/US-Online-Advertising-Spending-Surpass-Print-2012/1008783 (Accessed May 4, 2015).*

[Interactive Advertising Bureau UK 2013a.] Interactive Advertising Bureau UK. 2012 Full Year Digital Adspend Results. *Available at: http://www.iabuk.net/research/library/2012-full-year-digital-adspend-results (Accessed May 4, 2015).*

[Interactive Advertising Bureau UK 2013b.] Interactive Advertising Bureau UK. *UK Digital Adspend Up 12.5% to Almost £5.5bn. Available at: http://www.iabuk.net/about/press/archive/uk-digital-adspend-up-125-to-almost-55bn   (Accessed May 4, 2015).*

[Cidewalk, Inc. 2015.] Private Communications.

[Zubkov and Serov. 2013] A Complete Proof of Universal Inequalities for the Distribution Function of the Binomial Law. *Theory of Probability & Its Applications* 57 (3): 539–544.