Bid Optimizing and Inventory Scoring in Targeted Online Advertising

Claudia Perlich,Brian Dalessandro,
Rod Hook,Ori Stitelman,Troy Raeder
m6d research
37 East 18th Street
New York, NY, 10003
claudia,briand,rod,ori,troy@m6d.com

Foster Provost
NYU / m6d research
New York, NY, 10012
fprovost@stern.nyu.edu

ABSTRACT

Billions of online display advertising spots are purchased on a daily basis through real time bidding exchanges (RTBs). Advertising companies bid for these spots on behalf of a company or brand in order to purchase these spots to display banner advertisements. These bidding decisions must be made in fractions of a second after the potential purchaser is informed of what location (Internet site) has a spot available and who would see the advertisement. The entire transaction must be completed in near real-time to avoid delays loading the page and maintain a good users experience. This paper presents a bid-optimization approach that is implemented in production at Media6Degrees for bidding on these advertising opportunities at an appropriate price. The approach combines several supervised learning algorithms, as well as second price auction theory, to determine the correct price to ensure that the right message is delivered to the right person, at the right time.

Categories and Subject Descriptors

I.5.4 [**Computing Methodologies**]: Pattern Recognition-Applications

Keywords

Probability Estimation, Online Advertising, Causality, Bid Optimization, Logistic Regression

1. INTRODUCTION

In recent years online advertising has seen a major shift toward real time bidding (RTB) ad-exchanges [7]. As the term suggests, RTBs allow advertisers to bid on the opportunity to show an ad to a specific browser on a specific site with a display ad slot. Advertisers are integrated into the exchanges through APIs and collect different data to decide whether or not they want to bid and at what price. RTBs

provide liquidity on both supply and demand and have created a complex economic ecosystem of publishers, advertisers, marketers, and data providers [12]. Media6Degrees (m6d) is one of the players in the online display targeting sector with a main focus on prospecting — finding new customers for a product or brand. We find and target browsers for over 100 marketers delivering millions of ads daily. Before we get into some of the inner workings of this ecosystem, lets have a short and intuitive clarification of the terminology of the players in the display advertising ecosystem:

- The **marketer, brand, or product** is the one who is interested in advertising. It could for instance be Nike that wants to promote a new running shoe.

- The **advertiser** is the entity charged with executing an advertising campaign. This includes finding the right people and showing them the ad and would for instance be m6d.

- The **publisher or inventory** is the place where the ad will be shown. In particular, the online version of the New York Times is a publisher, but so are blogs and many other places online that have room for ads. We use the term inventory whenever we are looking at a particular publisher site appearing in an online display auction.

- The **browser or cookie** is conceptually the person who is going to see the ad. While there typically is a real person on the other side somewhere (unless it is a bot), from the advertiser's perspective it is really the cookie that is currently associated with the auction. There is not necessarily a one-to-one mapping between a person and a cookie.

- **Data providers** are in the business of annotating the inventory, the browser or both and selling this data to advertisers to help them make better targeting decisions. There are for instance sets of people labeled "credit card shoppers" that can be purchased. This information is typically the result of some modeling and is a prediction rather than ground truth.

- **Ad exchanges** bring together advertisers, browsers, publishers and data providers by running auctions. The result of an auction is an **impression**: the showing of an ad by the advertiser with the highest bid on the publisher inventory to the browser. Advertisers are

typically connected to multiple exchanges. In particular, m6d currently works with more than 15 different RTBs.

Strategies for targeting and bidding vary greatly between different advertisers and are primarily a function of the data an advertiser is willing and capable of using. The ultimate goal from the marketer's perspective is the holy grail of advertising:

> We want to address the **right browser** with the **right message** at the **right moment** and preferably at the **right price**.

Consider for instance search-based advertising: When a browser is entering "car insurance" as a search term, it is pretty straightforward to decide that this is a great opportunity for Geico or Allstate to show an ad for their insurance products. However, matching a search term to a product is typically not that simple. The 'right browser' is any browser who is currently searching for a keyword that is matched either manually or using machine learning [19] to the product or brand. The browser also defines the right time: the time of the search since he is obviously thinking about the product at that moment.

Contextual advertising has a similar flavor: if the browser is currently reading a review about used cars, right now might be a good opportunity to show him a car or car insurance ad. The challenge is again to associate the content of the page with the product or message. This works well only for a small set of well-understood content and very clearly-defined products.

Both of the above strategies suffer from high competition and limited scale. The number of total searches is finite and determined by the browsers. In cases where the match between the keyword and the product is obvious, the competition for showing the browser an ad is high and as a result the winning bid price in the auction is high as well. The same is true for very well defined content with clear product association. In addition, the latter scenario is typically managed through a direct agreement between a larger publisher and the relevant marketers. As a result, few such opportunities appear in the auctions. For the majority of products or things one might want to advertise, the association of content and product is highly non-trivial [19].

The m6d targeting strategy goes beyond this information. Through our data partners we collect historical information on browsers and use machine learning to build high-dimensional supervised ranking models to identify browsers with a high interest in the brand or high likelihood of buying the product. The actual targeting model is not the focus of this paper and is presented in more detail in [14, 15]. An important distinction between the above strategies and the m6d approach is that the latter targets browsers, not just singluar impression opportunities. In other words, even before the browser gets to the inventory, we already know some measure of fit between him and all our marketers. Using supervised machine learning on data that includes positively labeled browsers who already took some relevant brand action typically provides more relevant information (for conversion prediction) than the match between the content of the current inventory and an advertiser. At the point of the auction, we already have a good measure on the fit between browser and marketer, what we now want to do is adjust the bid price to reflect the last piece of the puzzle: is this the right place and time?

One important consideration in the discussion of the *value* of an advertising opportunity is the evaluation. We may have some high-level intuition about this, but in reality we have to focus on measurable outcomes to evaluate opportunities as well as advertising strategies. Primarily due to convenience, the industry has historically focused on clicks. This is rather unfortunate, as many studies have shown a rather loose if not inverse relationship between clicks and purchases for display advertising [1, 4]. Instead m6d measures *post-view* conversions that allow for a certain time period between the ad impression and some relevant action (e.g. visiting the marketer's website, downloading code, subscribing to some service, or purchasing a product). Keep in mind that both the advertiser and the marketer can trace the events of a browser through time using cookies (of course only as long as they do not get deleted).

Yet another layer to the optimization puzzle is attribution. My incentive as an advertiser is not just to increase the conversion rate or brand recognition of my customer, but to be fairly compensated for it. Attribution in the context of multi-touch and view through conversion is an entire research topic by itself [16, 5] and we will leave it for future work to tie our bid strategy to the attributed value. The biggest problem from our perspective is the complete lack of visibility. We do not know who else might have shown an ad to the browser and ultimately receives the credit for the conversion. Unfortunately, short-sighted attribution can have a strong impact on optimization strategies that are overall suboptimal for the marketer (and the industry in general). Last touch attribution leads for instance to a 'carpet bombing' strategy where some advertisers might try to buy huge volumes of cheap inventory that is not even seen by the browser. Even though attribution is not the focus of this paper, we are assuming a 'fair' attribution method where the advertiser has the incentive to evaluate opportunities with respect to the impact on a browser's conversion probability.

So to recast, we want to bid on the **right browser** with the **right message** at the **right inventory** at a price that reflects the true value of the opportunity to the marketer.

The rest of the paper is organized as follows: Section 2 will give a brief overview of the m6d targeting and bidding to clarify the constraints under which the bid optimization solution was developed. Section 3 initially presents consideration on *why* and *how* inventory affects the value of an opportunity and proposes a formal measure for opportunity value. Section 4 covers the estimation task and some modeling results. Finally, section 5 focuses on translating model predictions into bid strategies and measuring the actual performance impact on 15 advertising campaigns.

2. THE FLOW OF M6D TARGETING

In order to perform bid optimization one must first be able to determine the value of a particular inventory that is placed up for auction. However, determining the quality of a particular inventory is a challenging task because past decisions to show ads on a given inventory depend in part on the browser's propensity to convert. In other words, inventory quality is confounded by the many different decisions

that m6d has made as a company that make it possible to operate our large scale targeting machine. As a result, the quality of browsers that we observe at a piece of inventory is not equal across all inventory. Understanding the different pieces of the m6d infrastructure and how the components of that infrastructure contribute the quality of browsers at a particular inventory is essential for valuing each piece of inventory. For the purpose of this paper we will take a browser centric view and visit the sequence of events that ultimately lead to an impression of a particular ad to a particular user at a particular inventory. By illuminating this process one can better understand what issues contribute to the differential quality of browsers across the different inventories on which m6d receives bid requests. We are deliberately skipping the core m6d targeting component that builds a conversion propensity model for every campaign. Details can be found in [15]. The following steps highlight the process of browsers through the m6d system:

1. **Initiate** - A browser is first observed by m6d either through data partners or the marketer. They typically place a pixel on their page allowing m6d to write a cookie on the browsers computer. This cookie will contain in our case a random identifier that is assigned to the browser and allows us to recognize him again.

2. **Monitor** - m6d anonymously tracks the browser's Internet surfing behavior.

3. **Score and Segment** - The browsing history is used to score the browser for hundreds of active advertising campaigns. For all those campaigns that the browser scores high enough, he is placed into a discrete 'segment'. Each campaign has multiple segments for different thresholds and different models. Typically only about 1% of browsers qualify for any segment of a given campaign. The notion of discrete segments rather than continuous scores is a remnant of traditional demographic and behavioral targeting. It is still useful to discretely limit the set of eligible browsers, for which the RTB forwards bid requests.

4. **Sync with Exchange** - Once a browser is placed into a relevant segment, we have to communicate to the exchange that we would like to receive bid requests for him from now on. This process is called syncing.

5. **Activate Segment** - Account managers determine what segments to activate for a particular campaign, what base price to assign to each segment, and in which exchanges to run them. The base bid price is determined by the account manager and reflects the specifics of the contract with the customer, as well as margin and performance considerations. These decisions are assisted by reports that notify the account manager of the predicted conversion rate for the browsers in the segment and the anticipated number of browsers in each segment.

6. **Receive Bid Request** - Subsequently, m6d begins receiving bid requests for the browser. At the time of the request we know the identity of the browser, the current inventory, as well as additional generic information about the user agent, etc.

7. **Bid** - m6d first determines internally the set of active segments of the browser and checks against campaign volume goals and individual browser frequency cap. The segment with the highest price is then chosen and a bid is submitted on that campaign's behalf (this is all done in milliseconds). Every segment has a base price that is set by the account manager. In the case of bid optimization the price is adjusted to reflect the quality of the inventory as described in the following sections.

8. **Show Impression** - If the m6d bid was the highest across all bids for this auction, we get to show the creative for the campaign and pay the price equal to the second highest bid.

9. **Track Conversion** - We observe the browser for some campaign-specific conversion period, recording if the browser converts.

10. **The Cycle** - The entire process repeats where more data is collected, the browser is re-scored and more bid requests are received and ads shown.

11. **Cookie Deletion** - The browser at some point deletes its cookie and is no longer seen by the system. If we observe him again, he will receive a new cookie with a new random identifier and we start over again.

3. INVENTORY AND AD EFFECTIVENESS

We typically observe browsers/users multiple times during the day browsing different sites. The goal of this work is to quantify the value of a specific opportunity O where an opportunity is given by a triplet $O = (U, A, I)$ of an ad A for a marketer to a user U at a particular inventory I. Given the value of this opportunity we want to adjust the bid price B accordingly. So what is the relationship between the inventory and the browser's probability of converting? Inventory relates to the probability of conversion in two distinct ways. The first is independent of the ad whereas the second reflects an interaction between the inventory and the ad experience.

Organic Propensity

In some cases, the mere fact that the person visits a particular inventory can change our best estimate of his conversion probability. Reasons for this are:

1. **Contextual Relevance:** There is a correlation between the content of the inventory and the product or the brand. A purely-contextual targeting approach would take only the current content of the inventory URL into account when making a targeting decision. In the m6d context it is just one more data point and can be seen as an incremental update of the previous m6d propensity score. The main difference between the last content and the earlier ones is the relative timeliness that makes it potentially more relevant.

2. **Current Intentions:** If the browser is already engaged in online shopping, the probability that he might purchase from our marketer (regardless of the impression) is notably higher than if he was currently involved in some activity completely unrelated to shopping.

3. **Life Expectation:** Presumably, there can be a correlation between the kind of inventory a user sees and his probability of deleting the cookie. While this relationship is unlikely to be significant for most inventories, it might be notable for some. As a result of an increased probability of cookie deletion, the probability of *observed* conversion is lower.

Causal Impact
The objective of advertising is of course to increase the browser's propensity to convert. The effects below capture interactions between the inventory and the causal effect of showing an ad.

1. **Perceptiveness:** Even beyond the current intentions, the inventory can be related to the probability that a browser perceives an ad. While looking for the correct syntax of a SQL query on a forum page, a user is much less likely to take note of the impression on the side than if he is browsing around in news stories. Psychological experiments (the 'gorilla in the room' [3]) show that perception of external stimuli decreases significantly when the mind is focused on a specific task.

2. **Impression Quality:** There has been recently a lot of concern about the actual time that an impression is 'in view'. According to a recent study [9] as few as 37% of all impressions are actually seen (defined as being in the visible part of the page and actually displayed for at least 2 seconds) by the browser. Potential reasons for lack of viewability are that it may take too long for the impression to load or that the impression is 'displayed' below the fold or in some other invisible part of the page.

While it might be possible to quantify and model all 5 different factors independently, the true reason why a piece of inventory is particularly good is not nearly as relevant as a reliable measurement of the *relative* value of one piece of inventory over another.

Optimization Objective
So far we have focused on the conversion probability. However, the value of an opportunity depends on the objectives of the marketer and the advertiser. In an auction-theoretic framework [18], the optimal bid price of a second-price auction is determined by the expected value of the outcome

$$P(c|u, a, i) * V(c) \qquad (1)$$

where $V(c)$ is the value of a conversion. In reality, for online targeting, $V(c)$ ends up playing only a minor role in bid optimization. For all intents and purposes, $V(c)$ is typically unknown or simply assumed to be constant across all users. In that case, the value of the opportunity is proportional to the probability of conversion. This implies that the bid price should be proportional to the conversion probability as well.

In conclusion, we want to modify our bid proportionally to the effect of the inventory. If showing an ad on a particular piece of inventory doubles the probability of conversion over some random inventory, the bid price should be twice the average bid price as well.

$$\Phi^* = \frac{p(c|u, i, a)}{E_j[p(c|u, j, a)]}. \qquad (2)$$

The formulation in Equation 2 has another advantage: it is independent of a campaign's base conversion rate and can be handled by our system the same way no matter what the base rate of the campaign is. In particular, we can easily integrate it with the base bid price B determined by the account management team as a multiplier:

$$B^* = B * \Phi^*. \qquad (3)$$

One last observation on Equation 2 is its close relationship to the notion of relative impact in the context of causal analysis. Our definition of value compares the outcome of serving an ad at inventory i to the expectations over all counterfactual events of showing the ad at different inventories j. However, in order for a counterfactual analysis to produce unbiased estimates of a true causal effect (both in observational and experimental settings), several strict assumptions must be made about the data that we are not fully considering here [17].

4. ESTIMATION DETAILS AND MODELING RESULTS

Equation 2 states a well defined supervised modeling problem. We need to estimate $p(c|u, i, a)$ for all pieces of inventory i. We will estimate this quantity for each campaign a independently leaving us with $p_a(c|u, i)$. The inventory i is originally provided in form of an URL. M6d parses the URL to a meaningful level — typically the hostname. A subset of relatively common canonical URLs are assigned unique inventory IDs and the remaining long tail distribution of URLs is collected in an exchange-specific DEFAULT inventory bucket. Note that the inventory is exchange-specific. This implies that the same hostname (e.g., `www.facebook.com`) can map into multiple inventories. In total we maintain about 5,000 unique inventories.

4.1 Reducing the User Information

What do we know about the user at the time of the auction? We know the identity of the browser and could in principle retrieve all relevant information from the cookie to estimate $p_a(c|u, i)$. But is a tall order for a very high-dimensional targeting model to score in real time having to look up all the browsing history and the corresponding model parameters. On the other hand, we already have a highly optimized and pre-calculated estimate of the user's propensity to convert $p_a(c|u)$ from the m6d targeting that has assigned the browser into a segment of only the good prospects for a given campaign. So under the time constraint of a real time auction — where one of the worst case scenarios is timing out and not being able to bid — the only really relevant browser information is his segment s. Since trafficking decisions and base price are affected by segment s, we would have to control for it anyway. Consider if all the best prospects are targeted in exchange E_1 and the lower prospects in E_2. This would induce a major bias that makes the inventory in exchange E_1 look great unless we control during model estimation for all the variables that drive the trafficking decision.

4.2 Data Preparation

To estimate inventory scores for a given campaign, we need a sample of impressions delivered across multiple inventories, along with the segment information of each targeted
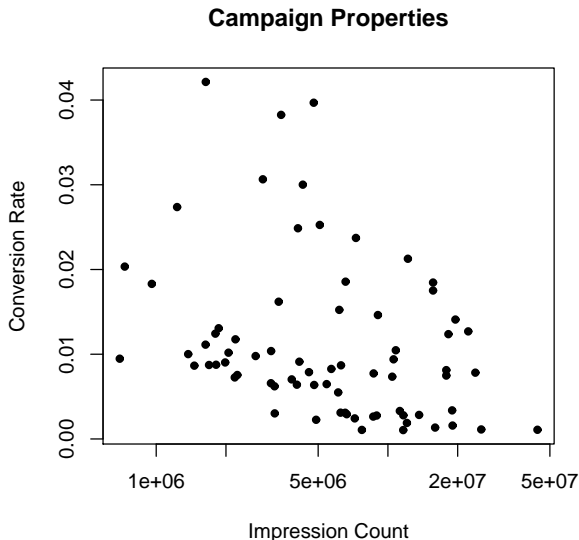
Campaign Properties

Figure 1: Number of impressions in a 3 week period and conversion rates for the 100 largest campaigns.

	AUC	Lift 10	Lift 20	Lift 50
Segment	0.729	1.802	1.744	1.461
Seg&Inventory	0.7636	2.121	1.952	1.524
Delta	0.0346	0.318	0.207	0.062
% Improved	97%	96%	95%	95%
% Delta	4.7%	17.6%	11.8%	4.3%
p-value	0	0	0	0

Table 1: Out of sample performance of the segment & inventory models is significantly better across different metrics (AUC and Lift at different percentage cutoffs) than the segment-only model. The p-value is for a paired t-test is always below 10^{-15}. Accounting for inventory almost never hurts targeting performance.

user. In addition, we track for each impression whether or not the user converted according to the rules of the campaign — typically a 7 day view through conversion. For a campaign where the marketer insists on measuring the click through rate, we can alternatively use only clicks as positives. Note that this leaves us with an interesting sparse representation where every example has exactly 2 binary non-zero features: one for the segment and one for the inventory.

Before we discuss the data for estimation, let us consider the dimensionality of the problem. We currently have 5,000 unique inventories across 15 exchanges, about 1,000 of which have notable volume. In addition, the average campaign uses between 10 and 50 different segments. The estimation task in its simplest form has an effective dimensionality of $\geq 1,000$. If one was to consider interaction terms, the dimensionality grows easily by an order of magnitude.

Conversion rates on the other hand are typically low and range between 0.04 in the best case and below 0.001 for others. This implies that the desirable amount of data for estimation is rather large. As a result, we typically use three weeks of impression data. Figure 1 shows typical impression counts and conversion rates for several of our campaigns. On average we show 15 million impressions, observe 250,000 conversions (mostly site visits) at an average conversion rate of 1%.

Our models are by default built against post-view conversion for a 7 day conversion window. The m6d system automatically assigns conversions to the last impression prior to the conversion. At this point it can be debated whether the implicit 'last touch' attribution is adequate. To make the estimation process efficient, we downsample the negative set from to a base conversion rate of 20% in the training set of each campaign. Note that this will affect the raw predictions and in turn the performance ratio, so we have to post-correct the model predictions [10].

Model Estimation and Feature Selection
We estimate $p(c|s, i)$ using L1-constrained logistic regression (with the default penalty weight) as implemented in the `bbr` software package [8]. Having an L1 penalty implicitly selects only the relevant features and provides a sparse solution. For many inventories, the parameter is set to zero and as a result the estimated impact factor is 1 and will not change the bid price (see the formal explanation below in Equation 4). Additionally, we applied an initial filter that included an inventory feature only if its volume was sufficient to yield at least 5 conversions in expectation given the base conversion rate of the campaign. Otherwise it was assigned to a default 'other' bucket. We need to include all observations to ensure the calibration of the model when we implicitly integrate over all inventories to calculate the expected value.

The careful reader might have noticed some built-in dependence at this point. In the current setting, the sign of the inventory effect is bound to be the same across all segments. In a linear model without interaction terms, the relative effect of the inventory will always point in the same direction (either greater or smaller than one) because the estimated parameter on the inventory indicator is either positive or negative. As a result, for a given campaign, the inventory score from Equation 2 can only differ in relative magnitude across segments, but this is purely a function of the different conversion rates of the segments. This effect is intuitively appealing if we consider the drivers of inventory impact from Section 3. It is hard to imagine that there is a very strong non-monotonic interaction effect between the conversion rate $p(c|s)$ and the inventory i. Either the inventory is of high quality with the impression in view or not. Something similar can be said for the current intention of the browser and his state of mind. Nevertheless we also build the models allowing for interactions between the segment and the inventory. Note that this increases the parameter space significantly (about a factor of 10) and typically hurts model performance due to overfitting.

Estimating Φ_i^*
We do not need to explicitly integrate over all inventories to estimate $E_j[p(c|u, j, a)]$. Instead we use the empirical distribution of j in the data. All that is needed is to build the same models as before on the exact same dataset but using only the segment variables as features.

$$E_j[p(c|u, j, a)] \sim \hat{p_a}(c|s) \qquad (4)$$

Inventory	Hotel $\hat{\Phi}_i^*$
Social Media Site	0.3
Travel Site	3.9
Dating Site	0.6
Movie Site	1.2
Video Site	1.4

Table 2: Example inventory scores $\hat{\Phi}_i^*$ for a hotel chain campaign.

This is also a relevant baseline to evaluate whether there is indeed any predictive information in the inventory i. Instead of looking at the absolute out-of-sample performance of the inventory + segment model, we compare its performance to this model. We know that inventory can be predictive — this is the basis of contextual targeting. However, bid-optimization is only effective if the inventory can provide *additional* explanatory power above and beyond the segment.

At this point we have two models for each marketer a, one that can predict $\hat{p_a}(c|s,i)$ and one that predicts $\hat{p_a}(c|i)$. Both are miscalibrated due to the initial downsampling of negatives. We correct both probability estimates according to the sample rate [10] and arrive at:

$$\hat{\Phi}_a^*(i,s) = \frac{\hat{p_a}(c|s,i)}{\hat{p_a}(c|s)} \qquad (5)$$

Model Performance
Before using this score in production we would like to assess how much information the inventory can provide beyond ordinary m6d targeting. Table 1 compares the out-of-sample performance (using cross validation) of the estimates of $\hat{p}(c|s,i)$ and $\hat{p}(c|s)$. The results for both AUC and lift at different percentage cutoffs are very encouraging with significantly higher performance of the inventory model. The AUC increases on average by 0.03, which corresponds to a 5% increase due to the additional (real-time) information from inventory. Similarly, lift (more relevant to targeting) increases as well, obviously more so for higher cutoffs. 'Lift 20' calculates the lift for the top 20% (ranked by the respective model) of browsers who saw an ad compared to all browsers who were shown an ad. Note that this reported lift is only within the 1-3% browsers that actually saw an ad for the specific campaign out of all browsers known to our system. Recall that the set of eligible browsers are identified by our main targeting model based on their browsing history and are selected because they have the highest conversion probabilities. So the reported lift is **not** the lift of our full targeting and bid optimization solution, but only the lift of inventory selection on top of the regular selection. As a results, this is a vast understatement of the true lift of our campaign where the baseline set of browsers contains the other 97-99% that we deliberately chose not to show ads to.

Example Predictions
Figure 2 shows the distribution of the estimates of $\hat{\Phi}_i^*$ across all inventories i for a particular hotel chain campaign. The scores are well-calibrated around 1 and show some promising signal for inventory that represents travel sites. Table 2 shows some example scores for the hotel chain campaign. Not surprisingly, travel sites have an extremely high score.
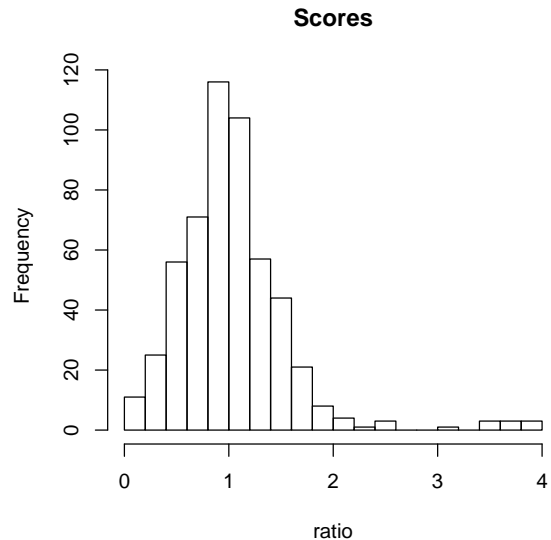

Scores

Figure 2: Distribution of the estimates of $\hat{\Phi}_i^*$ for a hotel chain campaign. The mean score was 1.03. The high-scoring inventories are all travel sites.

They meet most of the influence criteria outlined in section 3 for the organic category. Dating and social media sites, on the other hand, do not seem very suitable for hotel advertisement. They probably score low across multiple dimensions of inventory impact (bad target, low on current intent and low on perceptiveness).

5. CAMPAIGN PERFORMANCE IMPACT

Satisfied by the above results, we integrated the inventory scores into the m6d production system and studied the effectiveness of several inventory-based bid strategies on campaign performance. We present results on 15 campaigns come from a variety of industries including telecommunications, services, retail, travel, and others. The main selection criterion for choosing campaigns was a sufficient post-view site-visit rate to allow for significance in the results on approximately 2 weeks of data. We anticipated a performance increase of 5% lift. We back out the required number of impressions given the empirical conversion rate that we see for that campaign and include only campaigns that deliver a sufficient number of impressions such that a 5% increase should be significant under the assumption of a binomial distribution. Some subset of campaigns meeting this criterion were excluded since the account management team felt that they were unfit for experimentation. The characteristics of the the trial campaigns are very similar to those shown in Figure 1.

5.1 Bid Strategies

Using the inventory score in real time requires a very short response time. The technical implementation uses a lookup table by campaign, inventory, and segment that only keeps non-one scores to limit the size and increase the speed of the lookup. We have not observed a notable increase in response

time or timeouts since we implemented the inventory-based bid strategies.

- **Strategy 0:** The old strategy will serve as baseline where we always bid the base price of the segment as determined by the account managers. Note that this corresponds to a constant score of 1 across all inventories and is consistent with an uninformative inventory model.

- **Strategy 1:** The first optimization strategy takes an auction-theoretic view. Given that it is a second price auction, all players have the incentive to bid what the opportunity is worth to them. So we multiply the base price by the score ratio of the particular inventory.

- **Strategy 2:** The second optimization strategy is more aggressive towards performance than the previous one. Instead of bidding what we think an opportunity is worth, we apply a step function to the model ratio to translate it into a factor applied to the price. Any ratio below 0.8 yields a bid price of 0 (so not bidding), ratios between 0.8 and 1.2 are set to 1 and ratios above 1.2 bid twice the base price.

M6d has a number of other bid strategies that are inventory independent and are not discussed further. However, the design of the bid optimization as a multiplicative modifier allows for any arbitrary combination of strategies provided they are all centered at 1 for neutrality.

During fall 2011, m6d implemented a flexible bidding platform that can test multiple bid strategies on the same campaign simultaneously. For this purpose, we assign each browser randomly to exactly one strategy even if they qualify for multiple segments of the same campaign.

5.2 Campaign Results

We evaluate the bid strategies with respect to:

1. **Conversion Rate (PVSVR):** Percentage of impressions leading to site visits within 7 days. Higher conversion rates are better.

2. **Cost per Acquisition (CPA):** This metric combines cost and conversion rate and looks at the total cost of impressions for a given strategy relative to the total number of conversion. Lower CPA is better.

Both measures are of practical relevance and related to the previous discussions. The majority of m6d's customer agreements specify a cost per mille (CPM) payment model where we are paid independently of the conversion rate. However, the conversion rate is considered by the customer as a performance metric to compare us against other vendors. CPA is directly relevant to our margin. For a given conversion performance we would like to minimize the CPA or alternatively for a given cost, we will try to deliver more conversions. Note that from the marketer's perspective on a CPM campaign there is a direct relationship between the PVSVR the CPA. In more general terms, CPA measures economic efficiency and should be closely related to strategy 1. Lets assume for the moment that the actual cost is equal or proportional to the bid price. Under strategy 1 we would expect that the tradeoff between CPA and PVSVR is better than the suboptimal strategy 0 – either CPA is lower for approximately the same conversion rate or higher conversion rate for similar

Campaign	Measure	Str 0	Str 1	Str 2
1	PVSVR	0.00274	0.00294	**0.00333**
2	PVSVR	0.00080	0.00087	**0.00097**
3	PVSVR	0.00031	0.00033	**0.00041**
4	PVSVR	0.00044	0.00047	**0.00053**
5	PVSVR	0.00183	0.00193	**0.00229**
6	PVSVR	0.00037	**0.00041**	0.00040
7	PVSVR	**0.00526**	0.00432	0.00441
8	PVSVR	0.00020	**0.00026**	0.00025
9	PVSVR	0.00185	NA	**0.00205**
10	PVSVR	0.00439	NA	**0.00600**
11	PVSVR	0.00104	NA	**0.00119**
12	PVSVR	0.00889	NA	**0.01088**
13	PVSVR	0.00387	NA	**0.00464**
14	PVSVR	0.00204	NA	**0.00299**
15	PVSVR	0.00292	NA	**0.00319**
1	CPA	**0.4434**	0.5048	0.5640
2	CPA	1.3406	**1.2799**	1.4807
3	CPA	**4.5124**	4.6714	5.2446
4	CPA	**2.3304**	2.5129	2.7400
5	CPA	0.4854	**0.4785**	0.5100
6	CPA	2.7487	**2.6566**	3.3143
7	CPA	**0.1201**	0.1382	0.192
8	CPA	6.1238	**5.6980**	7.9599
9	CPA	0.4764	NA	**0.4261**
10	CPA	**0.1966**	NA	0.2242
11	CPA	**1.1136**	NA	1.5764
12	CPA	**0.1804**	NA	0.2086
13	CPA	**0.3524**	NA	0.4354
14	CPA	0.6074	NA	**0.5497**
15	CPA	**0.4698**	NA	0.6207

Table 3: Bid optimization performance in terms of PVSVR and CPA for identifying new customers (Prospecting). Bold indicates better performance, while not necessarily significant.

CPA. Setting the bid price proportional to the conversion probability does not maximize PVSVR. As a matter of fact, the optimal strategy for maximizing the conversion rate is to neglect any cost calculation and simply bid the practical equivalent of infinity for every opportunity above a certain threshold that is determined by the total desired number of impressions.

However, this is clearly not in our best interest considering margins. Strategy 2 is closer to this extreme whereas strategy 1 is aimed at economic efficiency in terms of the cost of acquisition. We have a number of campaigns that are evaluated with respect to the cost per acquisition.

Tables 3 and 4 show detailed results for the three bid strategies. A 'NA' indicates that this marketer did not use that particular strategy. Given the results m6d decided that overall strategy 2 held more promise and ran only strategy 2 during the second stage of our experiment. We have separated the results into two browser groups: those with very high conversion rate (retargeting) and those with lower conversion rates (prospecting). Recall that the model is estimated jointly over both populations.

In both populations we observe that the more aggressive strategy 2 outperforms strategies 0 and 1 in terms of the post-view site visit rate (PVSVR) by a large margin. This

Campaign	Measure	Str 0	Str 1	Str 2
1	PVSVR	0.0062	0.0065	**0.0126**
2	PVSVR	0.0025	0.0026	**0.0030**
3	PVSVR	0.0013	0.0014	**0.0018**
4	PVSVR	0.0055	0.0056	**0.0064**
5	PVSVR	0.0192	0.0193	**0.0236**
6	PVSVR	0.0016	0.0016	**0.0026**
7	PVSVR	0.0168	0.0183	**0.0218**
8	PVSVR	0.0026	0.0026	**0.0028**
9	PVSVR	0.0094	NA	**0.0118**
10	PVSVR	0.0092	NA	**0.0154**
11	PVSVR	0.0283	NA	**0.0338**
12	PVSVR	0.0126	NA	**0.0163**
13	PVSVR	0.0270	NA	**0.0316**
14	PVSVR	**0.0075**	NA	0.0074
15	PVSVR	0.0309	NA	**0.0404**
1	CPA	0.3407	0.3074	**0.1854**
2	CPA	0.9108	**0.8995**	0.9430
3	CPA	**1.3419**	1.3667	1.3526
4	CPA	0.4459	0.4285	**0.3928**
5	CPA	0.0677	0.0693	**0.0666**
6	CPA	**1.7629**	1.7813	1.1997
7	CPA	0.1785	0.1665	**0.1551**
8	CPA	**0.9503**	0.9619	1.0223
9	CPA	**0.2058**	NA	0.2109
10	CPA	0.1155	NA	**0.1000**
11	CPA	**0.0803**	NA	0.0837
12	CPA	0.1580	NA	**0.1478**
13	CPA	**0.0631**	NA	0.0708
14	CPA	**0.2141**	NA	0.2823
15	CPA	**0.0444**	NA	0.0480

Table 4: Bid optimization performance in terms of PVSVR and CPA for predicting the return of an existing customer (Retargeting). Bold indicates better performance, while not necessarily significant.

is very consistent with our expectations. The performance increase over the baseline is on average 24% for retargeting and 21% for prospecting. For comparison, a lift of 20% is considered quite substantial in marketing and this does not even include the core m6d targeting component. So the performance gains realized through strategy 2 are very impressive and as a result this strategy got rolled out to a larger number of campaigns than strategy 1.

In terms of CPA, we see that strategy 2 is on average 5% lower than strategy 0 in retargeting, beating the baseline in 7 of 15 cases, but 18% higher for prospecting and exceeds the baseline in all but one campaign. The decline in the CPA performance was expected for strategy 2. It is a positive surprise that strategy 2 is actually better both in terms of CPA and PVSVR for the retargeting group.

The reason for the CPA increase for prospecting is that the model estimates are affected more by the retargeting population due to the notably higher conversion rate and as a result, the model is more accurate for retargeting. But even in the prospecting case, the cost increase for strategy 2 reflects only a 10-20% reduction of our margin, which is acceptable when we need to drive conversions. This suggests that we may not implement this strategy across all campaigns but rather use it as a tool to manage performance

more effectively when renewal or the size of a followup deal is under consideration.

Strategy 1 beats strategy 0 on 7 out of 8 campaigns in terms of PVSVR for both browser groups and realizes an average lift of 6% on prospecting and 3% on retargeting. However, it provides economic efficiency gains (CPA) over the the baseline in only 4 out of 8 cases - clearly not significant in either group. In summary we observe a higher conversion rate at equal CPA: an overall better economic position.

We would like to finish the discussion with an interesting observation. In most CPM campaigns where we are paid for a fixed number of impressions independently of the conversions, we are still subject to comparative evaluation against other advertisers in terms of the implied CPA. Note that our PVSVR translates directly into the marketer's measure of cost per acquisition. So while our direct incentive may not be to increase the conversion rate, in a repeated game where we have to win renewal or a larger budget in direct comparison to other advertisers, it is very much in our interest to trade short term margin for long term growth.

6. RELATED WORK

Ad exchanges and real time bidding are fairly new and have revolutionized the advertising industry only in the last 1-2 years. Most reports are short blogs or white papers with very limited technical and performance details. The majority of published related work is in the more established application area of search advertising [2, 11, 6, 13]. In search term bidding the auction is for a given search term and the bid price determines the relative position in the ranking of the returned offers. One major difference in this setting is the payout structure. In our case, we always pay if we have the winning bid whereas in most of search advertising the bidder only pays his exact bid price if the offer was clicked on by the browser. That puts a larger burden of analytical optimization on the search provider (Google, Yahoo, etc.) and gives them the incentive to develop more elaborate auction styles where the bid price may not be the sole decision criterion. In addition, the bid price is not really determined in real time and is not browser specific. The advertisers could typically submit and change their bid price for a given term at any time throughout the day. The predictive modeling component typically evolves around estimating the conditional click propensity and on the side of the advertiser the expected revenue given click.

7. CONCLUSION AND FUTURE WORK

Bid optimization has become one standard component of the production system of m6d targeting. We developed a predictive modeling approach that estimates reliably and efficiently the relative impact of inventory on the conversion probability. It is interesting to note the scale of the underlying models: We are estimating close to 1,000 parameters by downsampling from an average of 50 million impressions per campaign semi-automatically with a model refresh rate of about 2 weeks. The scores can be applied as factors to the base-prices previously set by the account management team. The empirical results suggest a clear improvement of the campaign performance in terms of conversion rate that can be used to tune the campaign to satisfy the performance expectations of the marketer. One factor that could

currently limit our performance is the internal selection of the campaign with the highest (post-adjustment) bid price. Since we are picking the max of a distribution, this can lead to an over-inflation of the bid price and amplify overfitting (the highest scores in most models are not necessarily the most reliable). Understanding this effect is one of out immediate future work efforts and should lead to reliable cross-campaign optimization. In addition we are embarking on a deep dive into the modeling process with a stronger focus on the prospecting group. The main advantage of pooling the two groups is a favorable bias - variance tradeoff. However we think there is potentially a better solution that can improve the prospecting CPA further. Finally we see great potential in expanding the scope of bid optimization to bring together in real time additional information from the cookie, such as age and activity level, and combine the performance-centric optimization approach with other tuning parameters. One such parameter is the frequency of ad exposure for a given marketer. There is presumably a diminishing causal impact of each subsequent impression on browser conversion. Future bid optimization will consider the impression history for that browser in addition to inventory considerations.

8. ACKNOWLEDGEMENT

9. REFERENCES

[1] L. Abraham. Is the click the right currency for display advertising effectiveness. ComScore, 2011. http://paidcontent.org/files/uploads/ DLD_2011_Linda_Abraham.pdf.

[2] G. Aggarwal, A. Goel, and R. Motwani. Truthful auctions for pricing search keywords. In *Procedings of ACM Electronic Commerce*, 2006.

[3] C. Chabris and D. Simons. *The Invisible Gorilla*. Broadway, 2011.

[4] B. Dalessandro, F. Provost, C. Perlich, and R. Hook. Surrogate training for online conversion modeling. In *EMPGENS2*, 2012.

[5] B. Dalessandro, O. Stitelman, C. Perlich, and F. Provost. Causally motivated attribution for online advertising. In *Proceedings of the Workshop on Data Mining and Audience Intelligence for Online Advertising at KDD*, 2012.

[6] B. Edelman and M. Ostrovsky. Strategic bidder behavior in sponsored search auctions. In *Proceedings of ACM Electronic Commerce*, 2005.

[7] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, 2007.

[8] A. Genkin, D. D. Lewis, and D. Madigan. Large-scale bayesian logistic regression for text categorization. *Technometrics*, 49(3):291–304, 2007.

[9] M. Hughes. Taking issues with viewable impressions, 2012. http://www.adexchanger.com/data-driven-thinking/taking-issue/.

[10] G. King and L. Zeng. Logistic regression in rare events data. *Political Analysis*, 9(2):137–163, 2001.

[11] B. Kitts and B. LeBlanc. Optimal bidding on keyword auctions. *Electronic Markets*, 14(3):186–201, 2004.

[12] M. Muthukrishnan. Data mining problems in internet ad systems. In *Tutorial at ACM SIGKDD*, 2011.

[13] B. Edelman M. Ostrovsky and M. Schwartz. Internet advertising and the generalized second prize auction: Selling billions of dollars worth of keywords. *American Economic Review*, 2006.

[14] F. Provost, B. Dalessandro, R. Hook, X. Zhang, and A. Murray. Audience selection for on-line brand advertising: privacy-friendly social network targeting. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge discovery and data mining*, 2009.

[15] T. Raeder, B. Dalessandro, O. Stitelman, C. Perlich, and F. Provost. Design principles of massive, robust prediction systems. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012.

[16] X. Shao and L. Li. Data-driven multi-touch attribution models. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 258–264, 2011.

[17] A.A. Tsiatis. *Semiparametric theory and missing data*. Springer Verlag, 2006.

[18] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, (16):8–27, 1961.

[19] C. Xiong, T. Wang, W. Ding, Y. Shen, and T.Y. Liu. Relational click prediction for sponsored search. In *Proceedings of WSDM*, 2012.