AI meets Real-Time Bidding: A Case Study

Jian Xu

Big Data and Artificial Intelligence Lab
TouchPal Inc.

July 12, 2017

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Online Advertising Landscape

The compound annual revenue growth rate over the past ten years for online advertising is 16%. Mobile accounting for 50.52% of all revenue in 2016, surpassing non-mobile revenue for the first time. [1]

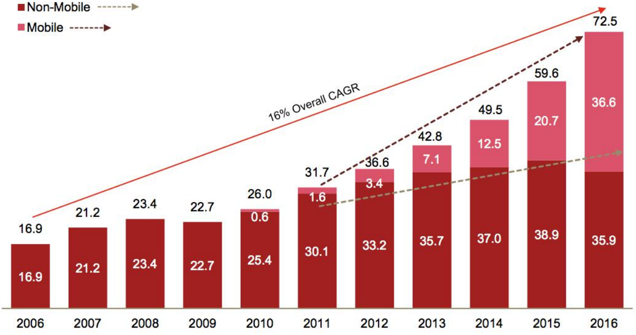

Figure: US online advertising annual revenue.

[1] Source: IAB/PwC Internet Ad Revenue Report, FY 2016

Online Advertising Landscape

Search ads (46%) and display ads (45%) contribute the majority of the online ad revenue. [2]

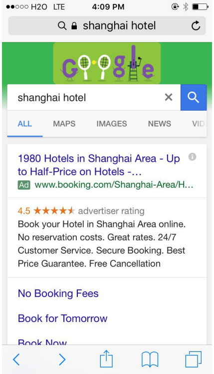

Figure: An search ad

Figure: A "native" display ad

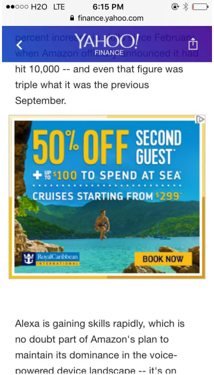

Figure: A traditional display ad

[2] Source: IAB/PwC Internet Ad Revenue Report, FY 2016

Online Advertising Landscape

- 80% US online display ad revenue comes from programmatic buying.
- 50% of the programmatic ad revenue comes from Real-Time Bidding (RTB) [3].

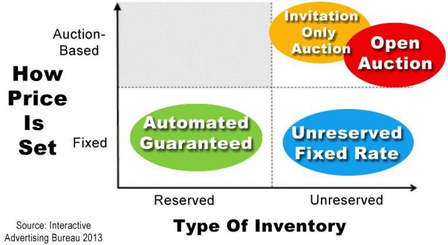

Figure: How display ad opportunities (inventory) are sold via programmatic buying.

RTB is deeply powered by AI.

[3] Source: https://www.emarketer.com/Article/eMarketer-Releases-New-Programmatic-Advertising-Estimates/1015682

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Real-Time Bidding Mechanism

Advertising buyers bid on an impression and, if the bid is won, the buyer's ad is instantly displayed on the publisher's site.

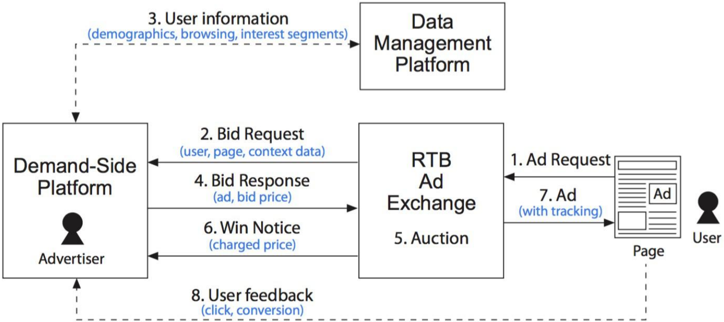

Figure: An illustration of how an ad impression is sold in real-time bidding. [4]

[4] Source: Weinan Zhang, Optimal Real-Time Bidding for Display Advertising, PhD Thesis, 2016.

Real-Time Bidding Mechanism

From the buyers'(advertisers, DSPs) perspective, it is important to bid smartly. There are two widely adopted problem settings:

No budget limit, maximize ROI

Suppose the return of an action (conversion) is *CPA*, the estimated action rate is $\hat{AR}$. The optimal bidding strategy is surprisingly simple

$$b_{OPT} = CPA \cdot \hat{AR} \tag{1}$$

With budget and lifetime limit, maximize some KPI

To optimize total action number within budget B and bid request volume T, Suppose p and F are p.d.f. and C.D.F. of the action rate respectively and the estimated action rate is $\hat{AR}$. The optimal bidding strategy is [a]

$$b_{OPT} = \frac{B}{T \int_{\theta} \int_{0}^{\theta} \vartheta (n-1) F(\vartheta)^{n-2} p(\vartheta) d\vartheta \, p(\theta) d\theta} \cdot \hat{AR} \tag{2}$$

[a] Source: Weinan Zhang, Optimal Real-Time Bidding for Display Advertising, PhD Thesis, 2016.

Real-Time Bidding Mechanism

RTB has to be accomplished with the help of AI:

- $<100ms$ latency requirement,
- $>1M$ bid requests per second (QPS)
- AI implements the bidding strategies by estimating $\hat{AR}$ in real-time; and even more ...

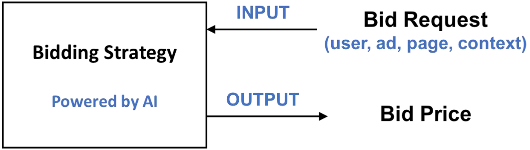

Figure: Real-Time Bidding as an AI task

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Case Study: AI Enables Better Bidding Strategy

The effective impact of an ad (AR lift) is more important than the absolute AR value.

A tiny example

Two users: *a* and *b*, campaign CPA: $100
AR_a: 0.04 if exposed to the ad, 0.03 if not (lift: 0.01); AR_b: 0.02 if exposed to the ad, 0.001 if not (lift: 0.019).
Bidder_1 bids prop. to AR assuming exposed: $4 for *a*, $2 for *b*;
Bidder_2 bids prop. to AR lift: $2 for *a*, $3.8 for *b*. Incremental value from Bidder $_1$: 0.01 conversions;
Incremental value from Bidder $_2$: 0.019 conversions.

- *If only one of them can be exposed to the ad, who will you select?*
- *Prevalent bidding strategy does not optimize campaign performance.*

Predicting AR lift

Let *ad* be an ad, *s* be the *state* of a user at ad request time, and $s_+(ad)$ be the *state* of the user if *ad* is shown. Let $p(action/s)$ be the AR of the user if *ad* is not shown and $p(action/s_+(ad))$ be the AR if *ad* is shown, the AR lift is

$$\Delta p = p(action/s_+(ad)) - p(action/s) \tag{3}$$

How to learn from data and predict Δp?

- Learn an *omnipotent* model to predict AR given any arbitrary state.
  - Use a function F to map a state to a set of features;
  - A generic AR prediction model $\hat{P}$ is built upon the derived feature set;
  - Finally the AR lift can be estimated as

$$\hat{\Delta p} = \hat{P}(action/F(s_+(ad))) - \hat{P}(action/F(s)) \tag{4}$$

Predicting AR lift (cont.)

Traditional AR prediction models are trained based on impression/click events. Some concerns about the traditional models [5]:

- Not generalized enough to learn $p(action|s)$, Suffering from survival bias,
- Not leveraging all the action information.

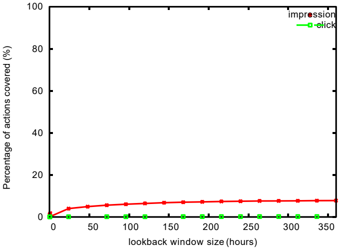

Figure: Only less than 10% of the converted users had been exposed to the ad of the advertiser.

[5] Jian Xu, et al. Lift-Based Bidding in Ad Selection. AAAI 2016.

Predicting AR lift (cont.)

Our approach to build a machine learning model to predict AR lift:

- Training sample generation
  - Generate training samples on user + time-stamp level, Leverage all the action data (conversion pixel firings).
- Features
  - $\{\text{Frequency, Recency}\} \times \{\text{impression, click, conversion}\}$
  - $\{\text{Frequency, Recency}\} \times \{\text{search, page view}\} \times \{\text{Topic}_1, \ldots, \text{Topic}_M\}$
  - Demographics, geolocation, device, etc.
- Model fitting
  - GBDT for rank order,
  - Isotonic regression for calibration.

Online A/B test

Three equal-sized random buckets: No show: does not bid at all,
Value-based bidding (50% budget assigned),
Lift-based bidding (50% budget assigned).

No show vs. value-based bidding

Adv	No bid		Value-base d bidding		Incremental action	**Action lift**
	# imps	# actions	# imps	# actions		
1	0	642	53,396	714	72	**11.2%**
2	0	823	298,333	896	73	**8.9%**
3	0	1,438	11,048,583	1,477	39	**2.7%**
4	0	1892	3,915,792	2,016	124	**6.6%**
5	0	5,610	6,015,322	6,708	1,098	**19.6%**

Table: Blind A/B test on five pilot advertisers - Value-based bidding v.s. “No bid”.

Online A/B test (cont.)

No show vs. lift-based bidding

Adv	No bid		Lift-based bidding		Incremental action	**Action lift**
	# imps	# actions	# imps	# actions		
1	0	642	59,703	826	184	**28.7%**
2	0	823	431,637	980	157	**19.1%**
3	0	1,438	11,483,360	1509	71	**4.9%**
4	0	1892	4,368,441	2,471	579	**30.6%**
5	0	5,610	8,770,935	8,291	2,681	**47.8%**

Table: Blind A/B test on five pilot advertisers - Lift-based bidding v.s. “No bid”.

Value-based bidding vs. lift-based bidding - Advertiser’s perspective

Adv	Value-based bidding			Lift-based bidding			Action lift	**Lift-over-lift**
	# imps	# actions	Action lift (vs “No bid”)	# imps	# actions	Action lift (vs “no bid”)		
1	53,396	714	11.2%	59,703	826	28.7%	13.6%	**156%**
2	298,333	896	8.9%	431,637	980	19.1%	9.4%	**115%**
3	11,048,583	1,477	2.7%	11,483,360	1509	4.9%	2.2%	**82%**
4	3,915,792	2,016	6.6%	4,368,441	2,471	30.6%	22.6%	**367%**
5	6,015,322	6,708	19.6%	8,770,935	8,291	47.8%	23.6%	**144%**

Table: “Action lift” is the absolute # actions difference between lift-based bidding and value-based bidding. “Lift-over-lift” is comparing the their action lifts over “no bid”.

Online A/B test (cont.)

DSPs get rewarded (paid) by advertisers based on the number of actions attributed to them. So there is a gap between DSPs and advertisers that must be abridged with attribution models.

Value-based bidding vs. lift-based bidding - DSP's perspective.

Adv	Val ue-based b dding			Li ft-based bi ding			Inventory-cost diff	Cost-per-imp diff
	# imps	# attrs	Inventory cost	# imps	# attrs	Inventory cost		
1	53,396	50	$278.73	59,703	50	$300.31	**7.7%**	**-3.6%**
2	298,333	80	$1,065.05	431,637	80	$1,467.57	**37.8%**	**-4.8%**
3	11,048,583	240	$25,522.22	11,483,360	240	$25,837.56	**1.2%**	**-2.6%**
4	3,915,792	200	$10,846.74	4,368,441	200	$11,183.21	**3.1%**	**-7.6%**
5	6,015,322	500	$19,296.51	8,770,935	500	$23,501.90	**21.8%**	**-16.5%**

Table: Both bidders spent out equal amount of assigned budget, so the # attributions are always the same. Cost-per-impression is the inventory cost averaged by # impressions.

Agenda

Online Advertising Landscape

Real-Time Bidding Mechanism

Case Study: AI Enables Better Bidding Strategies

Takeaways

Summary

- RTB contributes a large portion of the online advertising revenue. It is the most dynamic marketplace that AI has to be leveraged.
- AI technologies are the building blocks to implement advanced bidding strategies.
- The effective impact of an ad can be modeled and predicted, with which a lift-based bidding strategy can bring advertisers true values.
- Lift-based bidding may hurt DSPs with industry standard attribution model. The key to move DSPs to lift-based bidding is the attribution model.

Thank you! Questions?

Value-based bidding vs. lift-based bidding

Definition (*AR, background AR, and AR Lift*)

Given ad request, user, advertiser triplet (q, u, A), AR w.r.t. (q, u, A) is the probability that u will take the desired action defined by A after the ad of A is served to q, background AR w.r.t. (q, u, A) is the probability that u will take the desired action if the ad of A is not served to q, and AR lift as the difference between AR and background AR. We denote by p the AR, Δp the AR lift, and $p - \Delta p$ the background AR.

Definition (*Value-Based Bidding*)

Let p be the AR of a user if the advertiser's ad is shown, value-based bidding places a bid price of $\alpha \times p$ to acquire an impression from this user for the advertiser, where $\alpha > 0$.

Definition (*Lift-Based Bidding*)

Let Δp be the AR lift of a user if the advertiser's ad is shown, lift-based bidding places a bid price of $\beta \times \Delta p$ to acquire an impression from this user for the advertiser, where $\beta > 0$.

Bidder$_1$ is a value-based bidder, Bidder$_2$ is a lift-based bidder. What if they bid for the same advertiser simultaneously?

Value-based bidding vs. lift-based bidding (cont.)

Lemma

$Bidder_1$ wins the auction for u_i at the cost of $\beta \times \Delta p_i$ if $\alpha \times p_i > \beta \times \Delta p_i$; $Bidder_2$ wins the auction for u_i at the cost of $\alpha \times p_i$ if $\alpha \times p_i < \beta \times \Delta p_i$.

Theorem 1: $Bidder_2$ yields more actions than $Bidder_1$ when they are attributed same amount of credit.

i : the index of all the users

j : the index of those users that $Bidder_1$ wins (i.e., $\alpha \times p_j > \beta \times \Delta p_j$)

k : the index of those users that $Bidder_2$ wins (i.e., $\alpha \times p_k < \beta \times \Delta p_k$)

Expected attribution to $Bidder_1$: $\sum_j p_j$

Expected attribution to $Bidder_2$: $\sum_k$

Expected # of actions if only $Bidder_1$ is considered: $\sum_j p_j + \sum_k (p_k - \Delta p_k)$

Expected # of actions if only $Bidder_2$ is considered: $\sum_j (p_j - \Delta p_j) + \sum_k p_k$

When the same amount of actions is attributed to $Bidder_1$ and $Bidder_2$ (i.e., $\sum_j p_j = \sum_k p_k$), we have $\frac{\sum_j p_j + \sum_k (p_k - \Delta p_k)}{\sum_j p_j} < 2 - \frac{\alpha}{\beta} < \frac{\sum_j (p_j - \Delta p_j) + \sum_k p_k}{\sum_k p_k}$.

Value-based bidding vs. lift-based bidding (cont.)

Theorem 2: Bidder$_2$ costs more than Bidder$_1$ when they are attributed same amount of credit.

i : the index of all the users
j : the index of those users that Bidder$_1$ wins (i.e., $\alpha \times p_j > \beta \times \Delta p_j$)
k : the index of those users that Bidder$_2$ wins (i.e., $\alpha \times p_k < \beta \times \Delta p_k$)
Expected attribution to Bidder$_1$: $\sum_j p_j$
Expected attribution to Bidder$_2$: $\sum_k p_k$
Expected cost of Bidder$_1$: $\sum_j \beta \times \Delta p_j$ Expected cost of Bidder$_2$: $\sum_k \alpha \times p_k$

When the same amount of actions is attributed to Bidder$_1$ and Bidder$_2$ (i.e., $\sum_j p_j = \sum_k p_k$), we have $\frac{\sum_j \beta \times \Delta p_j}{\sum_j p_j} < \frac{\sum_j \alpha \times p_j}{\sum_j p_j} = \frac{\sum_k \alpha \times p_k}{\sum_k p_k}$.