Lift-Based Bidding in Ad Selection

Jian Xu*, Xuhui Shao, Jianjie Ma, Kuang-chih Lee, and Quan Lu

*TouchPal Inc.
Yahoo Inc.

02/14/2016
Motivation

Macro assumption: an action (conversion) can happen even if the user has not been exposed to an ad.

A tiny example

Two users: $a$ and $b$
$AR_a$: 0.04 if exposed to the ad, 0.03 if not;
$AR_b$: 0.02 if exposed to the ad, 0.001 if not.

🌟 If only one of them can be exposed to the ad, who will you select?
A not-so-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: 0.19 conversions.

Expected attribution to Bidder_1: 0.04 conversions;

Expected attribution to Bidder_2: 0.02 conversions.

💡 Prevalent bidding strategy does not optimize campaign performance;

💡 Bidders are not rewarded fairly.
"Obama’s campaign focused on swing state voters the campaign had scored as "persuadable," and voters who were supporters but needed to be encouraged to turn out at the polls"

– Carol Davidsen, ran the campaign’s television ad "Optimizer” project
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, \(\Delta p\) 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 \(\Delta p\) 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?
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\(_2\): \(\sum_k p_k\)

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 \(\frac{\sum_j p_j + \sum_k (p_k - \Delta p_k)}{\sum_j p_j} < 2 - \frac{\alpha}{\beta} < \frac{\sum_j (p_j - \Delta p_j) + \sum_k p_k}{\sum_k p_k} \).
Theorem 2: Bidder\textsubscript{2} costs more than Bidder\textsubscript{1} when they are attributed same amount of credit.

\(i\): the index of all the users  
\(j\): the index of those users that Bidder\textsubscript{1} wins (i.e., \(\alpha \times p_j > \beta \times \Delta p_j\))  
\(k\): the index of those users that Bidder\textsubscript{2} wins (i.e., \(\alpha \times p_k < \beta \times \Delta p_k\))

Expected attribution to Bidder\textsubscript{1}: \(\sum_j p_j\)  
Expected attribution to Bidder\textsubscript{2}: \(\sum_k p_k\)  
Expected cost of Bidder\textsubscript{1}: \(\sum_j \beta \times \Delta p_j\)  
Expected cost of Bidder\textsubscript{2}: \(\sum_k \alpha \times p_k\)

When the same amount of actions is attributed to Bidder\textsubscript{1} and Bidder\textsubscript{2} (i.e., \(\sum_j p_j = \sum_k p_k\)), we have \(\frac{\sum_j \beta \times \Delta p_j}{\sum_j p_j} < \frac{\sum_j \alpha \times p_j}{\sum_j p_j} = \frac{\sum_k \alpha \times p_k}{\sum_k p_k}\).
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

\[
\Delta p = p(action|s_+(ad)) - p(action|s) \tag{1}
\]

**How to learn from data and predict \( \Delta p \)?**

💡 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

\[
\hat{\Delta} p = \hat{P}(action|F(s_+(ad))) - \hat{P}(action|F(s)) \tag{2}
\]
Traditional AR prediction models are trained based on impression/click events. Some concerns:

- Not generalized enough to learn $p(\text{action}|s)$,
- 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.
Our approach:

- **Training sample generation**
  - Generate training samples on user + time-stamp level,
  - Emphasize users with high ad request volume,
  - Fully leverage usage of action data (conversion pixel firings)

- **Features**
  - \{Frequency, Recency\} $\times$ \{impression, click, retargeting, conversion\}
  - \{Frequency, Recency\} $\times$ \{search, page view\} $\times$ \{Topic$_1$, ..., Topic$_M$\}
  - Demo, geo, 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

<table>
<thead>
<tr>
<th>Adv</th>
<th>No bid</th>
<th>Value-based bidding</th>
<th>Incremental action</th>
<th>Action lift</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># imps</td>
<td># actions</td>
<td># imps</td>
<td># actions</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>642</td>
<td>53,396</td>
<td>714</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>823</td>
<td>298,333</td>
<td>896</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1,438</td>
<td>11,048,583</td>
<td>1,477</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1,892</td>
<td>3,915,792</td>
<td>2,016</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>5,610</td>
<td>6,015,322</td>
<td>6,708</td>
</tr>
</tbody>
</table>

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

Xu et al. AAAI’16 (TouchPal Inc.)
Online A/B test (cont.)

No show vs. lift-based bidding

<table>
<thead>
<tr>
<th>Adv</th>
<th>No bid</th>
<th>Lift-based bidding</th>
<th>Incremental action</th>
<th>Action lift</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># imps</td>
<td># actions</td>
<td># imps</td>
<td># actions</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>642</td>
<td>59,703</td>
<td>826</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>823</td>
<td>431,637</td>
<td>980</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1,438</td>
<td>11,483,360</td>
<td>1509</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1892</td>
<td>4,368,441</td>
<td>2,471</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>5,610</td>
<td>8,770,935</td>
<td>8,291</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Adv</th>
<th>Value-based bidding</th>
<th>Lift-based bidding</th>
<th>Action lift</th>
<th>Lift-over-lift</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># imps</td>
<td># actions</td>
<td>Action lift (vs “No bid&quot;)</td>
<td># imps</td>
</tr>
<tr>
<td>1</td>
<td>53,396</td>
<td>714</td>
<td>11.2%</td>
<td>59,703</td>
</tr>
<tr>
<td>2</td>
<td>298,333</td>
<td>896</td>
<td>8.9%</td>
<td>431,637</td>
</tr>
<tr>
<td>3</td>
<td>11,048,583</td>
<td>1,477</td>
<td>2.7%</td>
<td>11,483,360</td>
</tr>
<tr>
<td>4</td>
<td>3,915,792</td>
<td>2,016</td>
<td>6.6%</td>
<td>4,368,441</td>
</tr>
<tr>
<td>5</td>
<td>6,015,322</td>
<td>6,708</td>
<td>19.6%</td>
<td>8,770,935</td>
</tr>
</tbody>
</table>

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”.

Xu et al. AAAI’16 (TouchPal Inc.)
Value-based bidding vs. lift-based bidding - DSP’s perspective.

<table>
<thead>
<tr>
<th>Adv</th>
<th>Value-based bidding</th>
<th>Lift-based bidding</th>
<th>Inventory-cost diff</th>
<th>Cost-per-imp diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># imps</td>
<td># attrs</td>
<td>Inventory cost</td>
<td># imps</td>
</tr>
<tr>
<td>1</td>
<td>53,396</td>
<td>50</td>
<td>$278.73</td>
<td>59,703</td>
</tr>
<tr>
<td>2</td>
<td>298,333</td>
<td>80</td>
<td>$1,065.05</td>
<td>431,637</td>
</tr>
<tr>
<td>3</td>
<td>11,048,583</td>
<td>240</td>
<td>$25,522.22</td>
<td>11,483,360</td>
</tr>
<tr>
<td>4</td>
<td>3,915,792</td>
<td>200</td>
<td>$10,846.74</td>
<td>4,368,441</td>
</tr>
<tr>
<td>5</td>
<td>6,015,322</td>
<td>500</td>
<td>$19,296.51</td>
<td>8,770,935</td>
</tr>
</tbody>
</table>

**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.
Lift-based bidding benefits advertisers but may hurt DSPs with industry standard attribution model.

Lift can be estimated/predicted using an "omnipotent" AR prediction model.

AR prediction model could be established in a different way from conventional approaches.

The key to move DSPs to lift-based bidding is the attribution model.
Thank you!