Research Frontier of Real-Time Bidding based Display Advertising

Weinan Zhang
University College London
w.zhang@cs.ucl.ac.uk
http://www0.cs.ucl.ac.uk/staff/w.zhang

August 2015
Basic RTB Process

0. Ad Request

User Information
User Demography:
- Male, 26, Student
User Segmentations:
- Ad science, London traveling

1. Bid Request
   (user, page, context)

Demand-Side Platform
Advertiser

2. Bid Response
   (ad, bid price)

Data Management Platform

3. Ad Auction

RTB Ad Exchange

4. Win Notice
   (charged price)

Page

5. Ad
   (with tracking)

User

6. User Feedback
   (click, conversion)
Model Bidding Strategy

- A function mapping from bid request feature space to a bid price
- Design this function to optimise the advertising key performance indicators (KPIs)
Bidding Strategy in Practice

Bidding Strategy

- Feature Eng.
- Frequency Capping
- Retargeting
- Budget Pacing
- Bid Landscape
- Whitelist / Blacklist
- CTR / CVR Estimation
- Campaign Pricing Scheme
- Bid Calculation

Bid Request (user, ad, page, context)

Bid Price
Bidding Strategy in Practice: New Perspective

Bid Request
(user, ad, page, context)

CTR, CVR, revenue

Bidding Function

Preprocessing

Utility Estimation

Cost Estimation

Bid landscape

Bid Price
Discussed Topics of This Talk

Fundamentals
- CTR/CVR Estimation
- Bid Landscape Forecasting
- Bidding Strategies

Advances
- Arbitrage
- Unbiased Training and Optimisation
- Conversion Attribution
CTR/CVR Estimation

• A seriously unbalanced-label binary regression problem

$$\min_w \sum_{(y, x) \in D} \mathcal{L}(y, \hat{y}) + \lambda \Phi(w)$$

– Negative down sampling, calibration

• Logistic Regression

[Lee et al. Estimating Conversion Rate in Display Advertising from Past Performance Data. KDD 12]

$$\min_w \sum_{(y, x) \in D} \log(1 + e^{-yw^T x}) + \frac{\lambda}{2} \|w\|_2^2$$
CTR/CVR Estimation

• Follow-The-Regularised-Leader (FTRL) regression
  [McMahan et al. Ad Click Prediction : a View from the Trenches. KDD 13]

\[
\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left( \mathbf{g}_{1:t} \cdot \mathbf{w} + \frac{1}{2} \sum_{s=1}^{t} \sigma_s \| \mathbf{w} - \mathbf{w}_s \|_2^2 + \lambda_1 \| \mathbf{w} \|_1 \right)
\]

\[
\mathbf{g}_{1:t} = \sum_{s=1}^{t} \mathbf{g}_s \quad \sigma_s = \sqrt{s} - \sqrt{s - 1}
\]

Closed-form solution

\[
\omega_{t+1,i} = \begin{cases} 
0 & \text{if } |z_{t,i}| \leq \lambda_1 \\
-\eta_t (z_{t,i} - \text{sgn}(z_{t,i}) \lambda_1) & \text{otherwise.}
\end{cases}
\]

\[
z_{t-1} = \mathbf{g}_{1:t-1} - \sum_{s=1}^{t-1} \sigma_s \mathbf{w}_s
\]
CTR/CVR Estimation

- Factorisation Machines
  
  [Oentaryo et al. Predicting response in mobile advertising with hierarchical importance-aware factorization machine. WSDM 14]

  \[ \hat{y}(\mathbf{x}) = \sigma \left( w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i x_j \mathbf{v}_i^T \mathbf{v}_j \right) \]

  - Explicitly model feature interactions
  - Empirically better than logistic regression
  - A new way for user profiling

- GBDT+FM
  
Deep Learning Models [our working project]

CTR

- Fully Connected
- Hidden Layer (L2)
- Fully Connected
- Hidden Layer (L1)
- Fully Connected

Dense real Layer (Z)

Embedding via Factorization Machine

Fully connected within each feature

Sparse binary features (X)

Global

Feature 1

Feature 2
Bid Landscape Forecasting

Win probability:

\[ w(b) = \int_{z=0}^{b} p(z) \, dz \]

Expected cost:

\[ c(b) = \frac{\int_{z=0}^{b} zp(z) \, dz}{\int_{z=0}^{b} p(z) \, dz} \]
Bid Landscape Forecasting

- Log-Normal Distribution

[Cui et al. Bid Landscape Forecasting in Online Ad Exchange Marketplace. KDD 11]

\[
f_s(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0
\]
Bid Landscape Forecasting

- Price Prediction via Linear Regression

[Wu et al. Predicting Winning Price in Real Time Bidding with Censored Data. KDD 15]

\[ z = \beta^T x + \epsilon \quad \max_{\beta} \sum_{i \in W} \log \phi \left( \frac{z_i - \beta^T x_i}{\sigma} \right) \]

- Modelling censored data in lost bid requests

\[ P(b_i < z_i) = \Phi \left( \frac{\beta^T x_i - b_i}{\sigma} \right) \]

\[ \max_{\beta} \sum_{i \in W} \log \phi \left( \frac{z_i - \beta^T x_i}{\sigma} \right) + \sum_{i \in L} \log \Phi \left( \frac{\beta^T x_i - b_i}{\sigma} \right) \]
Bidding Strategies

• How much to bid for each bid request?

• Bid to optimise the KPI with budget constraint

\[
\begin{align*}
\text{max} & \quad \text{bidding strategy} \\
\text{KPI} & \quad \text{subject to} \quad \text{cost} \leq \text{budget}
\end{align*}
\]
Bidding Strategies

• Truthful bidding in second-price auction
  [Chen et al. Real-time bidding algorithms for performance-based display ad allocation. KDD 11]
  – Bid the true value of the impression

\[
\text{bid} = r_{\text{conv}} \times \text{CVR} \quad \text{or} \quad \text{bid} = r_{\text{click}} \times \text{CTR}
\]

• Non-truthful linear bidding
  [Perlich et al. Bid Optimizing and Inventory Scoring in Targeted Online Advertising. KDD 12]
  – With budget and volume consideration

\[
\text{bid} = \text{base\_bid} \times \frac{\text{predicted\_CTR}}{\text{base\_CTR}}
\]
Bidding Strategies

- Direct functional optimisation

  \[ b_{\text{ORTB}}(\theta) = \arg \max_{b(\theta)} N_T \int_0^\theta \theta w(b(\theta)) p_{\theta}(\theta) d\theta \]

  subject to \[ N_T \int_0^\theta b(\theta) w(b(\theta)) p_{\theta}(\theta) d\theta \leq B \]

- Solution: Calculus of variations

  \[ \mathcal{L}(b(\theta), \lambda) = \int_0^\theta \theta w(b(\theta)) p_{\theta}(\theta) d\theta - \lambda \int_0^\theta b(\theta) w(b(\theta)) p_{\theta}(\theta) d\theta + \frac{\lambda B}{N_T} \]

  \[ \frac{\partial \mathcal{L}(b(\theta), \lambda)}{\partial b(\theta)} = 0 \quad \Rightarrow \quad \lambda w(b(\theta)) = \left[ \theta - \lambda b(\theta) \right] \frac{\partial w(b(\theta))}{\partial b(\theta)} \]
Optimal Bidding Strategy Solution

(a) Winning function 1.

\[ w(b(\theta)) = \frac{b(\theta)}{c + b(\theta)} \]

(b) Bidding function 1.

\[ b_{\text{ORTB1}}(\theta) = \sqrt{\frac{c}{\lambda \theta} + c^2 - c} \]

[Zhang et al. Optimal real-time bidding for display advertising. KDD 14]
Overall Performance – Optimising Clicks or Conversions

iPinYou dataset

[Zhang et al. Optimal real-time bidding for display advertising. KDD 14]
Discussed Topics of This Talk

**Fundamentals**
- CTR/CVR Estimation
- Bid Landscape Forecasting
- Bidding Strategies

**Advances**
- Arbitrage
- Unbiased Training and Optimisation
- Conversion Attribution
Discussed Topics of This Talk

Fundamentals

• CTR/CVR Estimation
• Bid Landscape Forecasting
• Bidding Strategies

Advances

• Arbitrage
• Unbiased Training and Optimisation
• Conversion Attribution
Display Advertising Intermediaries

We want to pay per action
Advertisers with ad budget

CPA contracts
Intermediary making the match and arbitrages

CPM inventory
Publishers with ad inventory

We want to earn per impression

This work: Intermediary arbitrage algorithms in RTB display advertising.
[Zhang et al. Statistical Arbitrage Mining for Display Advertising. KDD 15]
Intermediary’s Statistical Arbitrage via RTB

- Statistical arbitrage opportunity occurs, e.g., when \((\text{CPM})\) cost per conversion < \((\text{CPA})\) payoff per conversion
- 1000 impressions * 5 cent < 8000 cent for 1 conversion
Statistical Arbitrage Mining

• Expected utility (net profit) and cost on multiple campaigns

\[ \mathbb{E}[R(\mathbf{v}, b(\theta, r))] = T \sum_{i=1}^{M} v_i \int_{\theta} \left( \theta r_i - b(\theta, r_i) \right) w(b(\theta, r_i)) p^{i}_{\theta}(\theta) d\theta \]

\[ \mathbb{E}[C(\mathbf{v}, b(\theta, r))] = T \sum_{i=1}^{M} v_i \int_{\theta} b(\theta, r_i) w(b(\theta, r_i)) p^{i}_{\theta}(\theta) d\theta \]

Bid request vol.  Est. payoff  winning function  CVR estimation

bidding function

Prob. of selecting Campaign i  Cost upper bound
Statistical Arbitrage Mining

- Optimising net profit by tuning bidding function and campaign volume allocation

\[ b_{\text{SAM}}(\cdot), v^* = \arg\max_{b(\cdot),v} \mathbb{E}[R] \]  

Total arbitrage net profit

\[ \mathbb{E}[C] \leq B \]  
Total cost constraint

\[ \text{Var}[R] \leq h \]  
Risk control

\[ 0 \leq v \leq 1 \]  

\[ v^T 1 = 1 \]  

- Solve it in an EM fashion

M-Step

E-Step
M-Step: Bidding function optimisation

- Fix \( \mathbf{v} \) and tune \( b() \)

\[
\max_{b()} \quad T \sum_{i=1}^{M} v_i \int_{\theta} \left( \theta r_i - b(\theta, r_i) \right) w(b(\theta, r_i)) p_{\theta}(\theta) d\theta
\]

subject to \[
T \sum_{i=1}^{M} v_i \int_{\theta} b(\theta, r_i) w(b(\theta, r_i)) p_{\theta}(\theta) d\theta \leq B.
\]

\[
\mathcal{L}(b(), \mathbf{v}) \quad = \quad 0 \quad \Rightarrow \quad \left( \frac{\theta r_i}{1 + \lambda} - b(\theta, r_i) \right) \frac{\partial w(b(\theta, r_i))}{\partial b(\theta, r_i)} = w(b(\theta, r_i))
\]

\[
\downarrow
\]

\[
w(b(\theta, r)) = \frac{b(\theta, r)}{l} \quad \Rightarrow \quad b_{\text{sam1}}(\theta, r) = \frac{r \theta}{2(1 + \lambda)}
\]

\[
w(b(\theta, r)) = \frac{b(\theta, r)}{b(\theta, r) + l} \quad \Rightarrow \quad b_{\text{sam2}}(\theta, r) = \sqrt{\frac{rl \theta}{1 + \lambda} + l^2 - l}
\]
E-Step: Campaign volume allocation

- Multi-campaign portfolio optimisation

\[
\max_{\mathbf{v}} \mathbf{v}^T \mu(b) - \alpha \mathbf{v}^T \Sigma(b) \mathbf{v},
\]

where

\[
\text{s.t. } \mathbf{v}^T \mathbf{1} = 1, \quad 0 \leq \mathbf{v} \leq \mathbf{1}
\]

- Portfolio margin mean
- Portfolio margin variance
- Net profit margin on each campaign

\[
\mu(b) = (\mu_1(b), \mu_2(b), \ldots, \mu_M(b))^T
\]

\[
\Sigma(b) = \{\sigma_{i,j}(b)\}_{i=1 \ldots M, j=1 \ldots M}
\]

\[
\mu_i(b) = \mathbb{E}[\gamma_i] = \mathbb{E} \left[ \frac{R_i(\mathbf{v}_{i=1}, b)}{C_i(\mathbf{v}_{i=1}, b)} \right], \quad \sigma_i^2(b) = \mathbb{E} \left[ \frac{R_i(\mathbf{v}_{i=1}, b)^2}{C_i(\mathbf{v}_{i=1}, b)^2} \right] - \mathbb{E} \left[ \frac{R_i(\mathbf{v}_{i=1}, b)}{C_i(\mathbf{v}_{i=1}, b)} \right]^2
\]
### Campaign Portfolio Optimisation Results

<table>
<thead>
<tr>
<th>strategies</th>
<th>easy payoff</th>
<th></th>
<th>hard payoff</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>profit (CNY)</td>
<td>margin</td>
<td>profit (CNY)</td>
<td>margin</td>
</tr>
<tr>
<td>bid. cam. algo. select.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lin greedy</td>
<td>501.12</td>
<td>6.63</td>
<td>68.59</td>
<td>0.91</td>
</tr>
<tr>
<td>lin portfolio</td>
<td>925.45</td>
<td>13.11</td>
<td>181.54</td>
<td>2.50</td>
</tr>
<tr>
<td>lin uniform</td>
<td>747.00</td>
<td>9.53</td>
<td>127.14</td>
<td>1.62</td>
</tr>
<tr>
<td>ortb greedy</td>
<td>517.02</td>
<td>6.65</td>
<td>70.96</td>
<td>0.91</td>
</tr>
<tr>
<td>ortb portfolio</td>
<td>802.15</td>
<td>10.32</td>
<td>146.13</td>
<td>1.88</td>
</tr>
<tr>
<td>ortb uniform</td>
<td>765.12</td>
<td>9.89</td>
<td>133.16</td>
<td>1.72</td>
</tr>
<tr>
<td>sam1 greedy</td>
<td>966.02</td>
<td>20.81</td>
<td>230.38</td>
<td>11.13</td>
</tr>
<tr>
<td><strong>sam1 portfolio</strong></td>
<td><strong>1,037.98</strong></td>
<td><strong>15.84</strong></td>
<td><strong>240.63</strong></td>
<td><strong>7.96</strong></td>
</tr>
<tr>
<td>sam1 uniform</td>
<td>768.38</td>
<td>9.78</td>
<td>172.43</td>
<td>7.57</td>
</tr>
<tr>
<td>sam2 greedy</td>
<td>961.68</td>
<td>28.73</td>
<td>235.31</td>
<td>24.00</td>
</tr>
<tr>
<td><strong>sam2 portfolio</strong></td>
<td><strong>983.01</strong></td>
<td><strong>17.21</strong></td>
<td><strong>248.65</strong></td>
<td><strong>13.61</strong></td>
</tr>
<tr>
<td>sam2 uniform</td>
<td>774.09</td>
<td>10.32</td>
<td>168.15</td>
<td>5.16</td>
</tr>
<tr>
<td>truth greedy</td>
<td>787.10</td>
<td>14.69</td>
<td>227.86</td>
<td>29.05</td>
</tr>
<tr>
<td>truth portfolio</td>
<td>787.10</td>
<td>14.69</td>
<td>242.07</td>
<td>18.34</td>
</tr>
<tr>
<td>truth uniform</td>
<td>326.57</td>
<td>4.14</td>
<td>101.12</td>
<td>5.36</td>
</tr>
</tbody>
</table>
Dynamic Portfolio Optimisation

- Campaign Volume Allocation Against Time
- Estimated Campaign Margin Against Time
- Empirical Net Profit Against Time
Online A/B Test on BigTree™ DSP

- 23 hours, 13-14 Feb. 2015, with $60 budget each
Discussed Topics of This Talk

Fundamentals
• CTR/CVR Estimation
• Bid Landscape Forecasting
• Bidding Strategies

Advances
• Arbitrage
• Unbiased Training and Optimisation
• Conversion Attribution
Problem of Training Data Bias

• Data observation process

\[ p_x(\mathbf{x})w(b_x) = q_x(\mathbf{x}) \]

• We want to train the model

\[
\min_\theta \mathbb{E}_{\mathbf{x} \sim p_x(\mathbf{x})} [\mathcal{L}(y, f_\theta(\mathbf{x}))] + \lambda \Phi(\theta)
\]

• But we train on the biased data

\[
\min_\theta \mathbb{E}_{\mathbf{x} \sim q_x(\mathbf{x})} [\mathcal{L}(y, f_\theta(\mathbf{x}))] + \lambda \Phi(\theta)
\]

[Zhang et al. Learning and Optimisation with Censored Auction Data in Display Advertising. AAAI 2016 Submission]
Unbiased Training

- Training target

\[
\min_{\theta} \mathbb{E}_{x \sim p_x(x)}[\mathcal{L}(y, f_\theta(x))] + \lambda \Phi(\theta)
\]

- Eliminate the data bias via importance sampling

\[
\mathbb{E}_{x \sim p_x(x)}[\mathcal{L}(y, f_\theta(x))] = \int_x p_x(x) \mathcal{L}(y, f_\theta(x)) \, dx
\]

\[
= \int_x q_x(x) \frac{\mathcal{L}(y, f_\theta(x))}{w(x, b_x)} \, dx = \mathbb{E}_{x \sim q_x(x)} \left[ \frac{\mathcal{L}(y, f_\theta(x))}{w(x, b_x)} \right]
\]

- Modelling winning probability via bid landscape

\[
w(x, b_x) = \int_0^{b_x} p^x_z(z) \, dz
\]
Unbiased Training

• Modelling winning probability via bid landscape

\[ w(\boldsymbol{x}, b_{\boldsymbol{x}}) = \int_{0}^{b_{\boldsymbol{x}}} p_{z}^{\boldsymbol{x}}(z) \, dz \]

• Only use observed impression data [UOMP]

\[ w_{o}(b_{\boldsymbol{x}}) = \frac{\sum_{(y, \boldsymbol{x}) \in D} \delta(z_{\boldsymbol{x}} < b_{\boldsymbol{x}})}{|D|} \]

• Also use lost bid request data (censored data) [KMMP]

\[ w(b_{\boldsymbol{x}}) = 1 - \prod_{b_{j} < b_{\boldsymbol{x}}} \frac{n_{j} - d_{j}}{n_{j}} \]

nj: # \{winning prices > bj\} \quad dj: # \{winning prices = bj\}
Experimental Results

- Winning probability estimation

![Graphs showing estimated win probability for different campaign numbers](image)
**Experimental Results**

- **CTR estimation: immediate performance improvement**

<table>
<thead>
<tr>
<th>Camp.</th>
<th>BIAS</th>
<th>UOMP</th>
<th>KMMP</th>
<th>FULL</th>
<th>BIAS</th>
<th>UOMP</th>
<th>KMMP</th>
<th>FULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1458</td>
<td>98.26</td>
<td>98.56</td>
<td>99.13</td>
<td>98.57</td>
<td>2.42</td>
<td>2.39</td>
<td>2.39</td>
<td>2.32</td>
</tr>
<tr>
<td>2259</td>
<td>60.27</td>
<td>60.94</td>
<td>62.00</td>
<td>67.37</td>
<td>4.04</td>
<td>4.03</td>
<td>4.02</td>
<td>4.00</td>
</tr>
<tr>
<td>2261</td>
<td>57.49</td>
<td>58.86</td>
<td>59.05</td>
<td>60.91</td>
<td>3.75</td>
<td>3.74</td>
<td>3.74</td>
<td>3.72</td>
</tr>
<tr>
<td>2821</td>
<td>59.25</td>
<td>59.69</td>
<td>60.28</td>
<td>62.36</td>
<td>7.07</td>
<td>7.06</td>
<td>7.04</td>
<td>6.92</td>
</tr>
<tr>
<td>2997</td>
<td>59.35</td>
<td>60.50</td>
<td>60.79</td>
<td>59.28</td>
<td>32.89</td>
<td>32.84</td>
<td>32.81</td>
<td>32.38</td>
</tr>
<tr>
<td>3358</td>
<td>96.59</td>
<td>96.78</td>
<td>97.01</td>
<td>97.32</td>
<td>4.48</td>
<td>4.47</td>
<td>4.38</td>
<td>4.36</td>
</tr>
<tr>
<td>3386</td>
<td>73.74</td>
<td>74.01</td>
<td>74.16</td>
<td>78.23</td>
<td>8.84</td>
<td>8.83</td>
<td>8.83</td>
<td>8.64</td>
</tr>
<tr>
<td>3427</td>
<td>96.04</td>
<td>96.42</td>
<td>96.78</td>
<td>97.02</td>
<td>3.37</td>
<td>3.37</td>
<td>3.33</td>
<td>3.31</td>
</tr>
<tr>
<td>3476</td>
<td>93.66</td>
<td>93.55</td>
<td>92.19</td>
<td>95.93</td>
<td>4.35</td>
<td>4.34</td>
<td>4.34</td>
<td>4.08</td>
</tr>
<tr>
<td>all</td>
<td>71.76</td>
<td>73.84</td>
<td>74.80</td>
<td>78.38</td>
<td>7.71</td>
<td>7.61</td>
<td>7.55</td>
<td>7.31</td>
</tr>
</tbody>
</table>
Discussed Topics of This Talk

Fundamentals
- CTR/CVR Estimation
- Bid Landscape Forecasting
- Bidding Strategies

Advances
- Arbitrage
- Unbiased Training and Optimisation
- Conversion Attribution
Conversion Attribution

- Assign credit% to each channel according to contribution
- Current solution: last-touch attribution
  [Shao et al. Data-driven multi-touch attribution models. KDD 11]
Multi-Touch Attribution

- **How to estimate the contribution of each channel?**
  [Shao et al. Data-driven multi-touch attribution models. KDD 11]

\[
P(y|x_i) = \frac{N_{\text{positive}}(x_i)}{N_{\text{positive}}(x_i) + N_{\text{negative}}(x_i)}
\]

\[
P(y|x_i, x_j) = \frac{N_{\text{positive}}(x_i, x_j)}{N_{\text{positive}}(x_i, x_j) + N_{\text{negative}}(x_i, x_j)}
\]

\[
V(x_i) = \frac{1}{2} P(y|x_i) + \frac{1}{2N_{j\neq i}} \sum_{j\neq i} \left( P(y|x_i, x_j) - P(y|x_j) \right)
\]

- **A more general formula**
  [Dalessandro et al. Casually Motivated Attribution for Online Advertising. ADKDD 11]

\[
V(x_i) = \sum_{S \subseteq I \setminus i} w_{S,i} (P(y|S, x_i) - P(y|S))
\]
<table>
<thead>
<tr>
<th>Channel</th>
<th>MTA Total</th>
<th>LTA Total</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Click</td>
<td>17,494</td>
<td>17,017</td>
<td>97%</td>
</tr>
<tr>
<td>Email Click</td>
<td>6,938</td>
<td>7,340</td>
<td>106%</td>
</tr>
<tr>
<td>Display Network A</td>
<td>5,567</td>
<td>8,148</td>
<td>146%</td>
</tr>
<tr>
<td>Display Network G</td>
<td>2,037</td>
<td>470</td>
<td>23%</td>
</tr>
<tr>
<td>Display Network B</td>
<td>1,818</td>
<td>1,272</td>
<td>70%</td>
</tr>
<tr>
<td>Display Trading Desk</td>
<td>1,565</td>
<td>1,367</td>
<td>87%</td>
</tr>
<tr>
<td>Display Network C</td>
<td>1,494</td>
<td>1,373</td>
<td>92%</td>
</tr>
<tr>
<td>Display Network D</td>
<td>1,491</td>
<td>1,233</td>
<td>83%</td>
</tr>
<tr>
<td>Email View</td>
<td>1,420</td>
<td>458</td>
<td>32%</td>
</tr>
<tr>
<td>Display Network E</td>
<td>1,187</td>
<td>1,138</td>
<td>96%</td>
</tr>
<tr>
<td>Brand Campaign</td>
<td>907</td>
<td>1,581</td>
<td>174%</td>
</tr>
<tr>
<td>Social</td>
<td>768</td>
<td>1,123</td>
<td>146%</td>
</tr>
<tr>
<td>Display Network H</td>
<td>746</td>
<td>284</td>
<td>38%</td>
</tr>
<tr>
<td>Display Network F</td>
<td>673</td>
<td>787</td>
<td>117%</td>
</tr>
<tr>
<td>Display Network I</td>
<td>489</td>
<td>136</td>
<td>28%</td>
</tr>
<tr>
<td>Retail Email Click</td>
<td>483</td>
<td>491</td>
<td>102%</td>
</tr>
<tr>
<td>Display Network J</td>
<td>222</td>
<td>92</td>
<td>41%</td>
</tr>
<tr>
<td>Retail Email</td>
<td>168</td>
<td>110</td>
<td>66%</td>
</tr>
<tr>
<td>Social Click</td>
<td>133</td>
<td>153</td>
<td>115%</td>
</tr>
<tr>
<td>Video</td>
<td>58</td>
<td>31</td>
<td>54%</td>
</tr>
</tbody>
</table>

[Shao et al. Data-driven multi-touch attribution models. KDD 11]
Bidding in Multi-Touch Attribution Mechanism

• Current bidding strategy
  – Driven by last-touch attribution
    \[
    \text{bid} = r_{\text{conv}} \times \text{CVR} \quad \text{or} \quad \text{bid} = r_{\text{click}} \times \text{CTR}
    \]

• A new bidding strategy
  – Driven by multi-touch attribution
    \[
    \text{bid} = r_{\text{conv}} \times \text{CVR} \times P(\text{attribution}|\text{conversion})
    \]
    [Xu et al. Lift-Based Bidding in Ad Selection. ArXiv 1507.04811. 2015]

\[
\Delta P = P(y|S, a) - P(y|S) \\
\text{bid} = \Delta P \times \text{base_bid}
\]
Value-based bidding v.s. Lift-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><strong># imps</strong></td>
<td><strong># actions</strong></td>
<td><strong># imps</strong></td>
<td><strong># actions</strong></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 2. Blind A/B test on five pilot advertisers - Value-based bidding v.s. “No bid”.

<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><strong># imps</strong></td>
<td><strong># actions</strong></td>
<td><strong># imps</strong></td>
<td><strong># actions</strong></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>1,892</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 3. Blind A/B test on five pilot advertisers - Lift-based bidding v.s. “No bid”.
Value-based bidding v.s. Lift-based bidding

| Adv | Value-based bidding | | Lift-based bidding | | | Inventory-cost diff | Cost-per-imp diff |
|-----|---------------------|---|-------------------|---|------------------|------------------|
| # imps | # attrs | Inventory cost | # imps | # attrs | Inventory cost | 7.7% | -3.6% |
| 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% |

• Comparison
  – Lift-based bidding help brings more conversions to advertisers
  – but its eCPA is higher than value-based bidding because of last-touch attribution

• Lift-based bidding with multi-touch attribution could bring a better eco-system
Taking-home Messages

• **Statistical Arbitrage Mining:** The internal auction selects the ad with highest arbitrage margin instead of the highest bid price.

• **Unbiased Training:** Add the weight to each instance to eliminate the auction-selection bias.

• **Attribution and Bidding:** Bidding proportional to the CVR lift instead of CVR value.
Computational Advertising Research in Academia

Disadvantages

- Lack of data and online test platform
- Lack of specific domain knowledge

Advantages

- Good at mathematic modelling
- Focus on knowledge collection and communication
- More research human resource
OpenBidder Project: www.openbidder.com

- Online open-source benchmarking project
  - Bid optimisation, CTR estimation, Bid landscape etc.
- Bridge academia and industry research on computational advertising
Collaborations

- Collaborations are more than welcome!

UK: adform, MediaGamma, US: Yahoo Labs

CN: iPinyou, YOYi, BigTree DSP, TalkingData, Tukmob
Thank You!
Questions?

http://www.computational-advertising.org
http://www0.cs.ucl.ac.uk/staff/w.zhang

Ad Science
WeChat Group