Bid Shading by Win-Rate Estimation and Surplus Maximization*

Shengjun Pan
Verizon Media
AdKDD ’20

* Work by: S Pan, B Kitts, T Zhou, H He, B Shetty, A Flores, D Gligorijevic, J Pan, T Mao, S Gultekin and J Zhang
What is bid shading?

Key difference: Bidding strategy for First-Price Auction relies on competing bidders.
Bid Shading: $3 → $2.01

Second-Price Auction:
Bidder 1: $1, Pay = $2
Bidder 2: $2, Pay = $2
Bidder 3: $3, Pay = $2

First-Price Auction:
Bidder 1: $1, Pay = $3
Bidder 2: $2, Pay = $3
Bidder 3: $3, Pay = $3
Surplus

Tradeoff between shading and winning:
more shading, more saving, less likely to win

\[ \text{surplus} = (V - b)I(b > mb2w) = \begin{cases} V - b, & \text{if } b > mb2w, \\ 0, & \text{otherwise.} \end{cases} \]

\[ V = \text{value of current ad opportunity} \]
\[ mb2w = \text{minimum bid to win (highest competing bid price)} \]
Optimal bid price

Competing bid prices unknown → Estimate distribution instead

\[ V = \text{value of current ad opportunity} \]
\[ \text{mb2w} = \text{minimum bid to win (highest competing bid price)} \]
## Previous approaches

<table>
<thead>
<tr>
<th>Segment</th>
<th>Machine-Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Distribution of mb2w* → Maximize surplus</td>
</tr>
<tr>
<td>Models</td>
<td>One model per segment</td>
</tr>
</tbody>
</table>
| Cons    | ● Cross-segment info unused  
         | ● Too many small segments;  
         | ● System burden; | ● Fixed bid for a segment;  
         | ● Mb2w must be given. |

* mb2w = minimum bid to win

Publisher + user dimensions
A unified approach: joint distribution estimation

Publisher features → Model → Would win or lose?
User features → Model
bid price → Model

Bid price:
- known at training;
- to be searched for at serving

Essentially estimate win-rate CDFs for all segments simultaneously.
A unified approach: formulation, bisection search

\[
Pr(\text{win}) = \left(1 + e^{-\left(w_0 + \sum_{i=1}^{k} w_i x_i + \beta \log b\right)}\right)^{-1}
\]

\[
b^* = \arg \max_{b > 0} (V - b) \text{ logistic} \left(w_0 + \sum_{i=1}^{k} w_i x_i + \beta \log b\right)
\]

\[
= \arg \max_{b > 0} (V - b) \left(1 + e^{-w_0 - \sum_{i=1}^{k} w_i x_i - \beta \log b}\right)^{-1}
\]

\[
= \arg \max_{b > 0} \frac{V - b}{1 + e^{-\alpha b - \beta}}
\]

where \(\alpha = w_0 + \sum_{i=1}^{k} w_i x_i\).
A/B testing

Budget controller:
Shaded → Surplus (saved cost) → Reinvested → Better eCPX

<table>
<thead>
<tr>
<th>A/B Testing</th>
<th>Spend</th>
<th>Surplus</th>
<th>eCPM</th>
<th>eCPC</th>
<th>eCPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>WR v.s. LR</td>
<td>+1.3%</td>
<td>+1.4%</td>
<td>-7.4%</td>
<td>-4.5%</td>
<td>-2.7%</td>
</tr>
<tr>
<td>WR v.s. SEG</td>
<td>+1.2%</td>
<td>+2.5%</td>
<td>-5.4%</td>
<td>-5.5%</td>
<td>-3.9%</td>
</tr>
</tbody>
</table>

Table 3: Improvements on Business Metrics

LR: logistic regression, predicting optimal bidding factor
SEG: per-segment distribution estimation
WR: our method, win-rate based unified approach
Challenges

- It is not obvious “what’s the right thing to do”
  - Different choices for formulation, e.g., what distribution to use
  - Different choices for metrics

- Convoluted with budget controllers
  - Offline experiments are not sufficient
  - Dependency on controllers

- Speed requirement at serving time
  - Inference or maximization must be fast at serving
Thank you!

Questions?