Real-Time Bidding
A New Frontier of Computational Advertising Research

Jun Wang and Shuai Yuan
University College London
With Invited Speaker

Kaihua Cai
AppNexus
About us

- Department of Computer Science, University College London
  - Media future research group
  - Information Retrieval and Computational advertising

- Dr. Shuai Yuan
  - Lead Data Scientist, MediaGamma
  - PhD in Computational Advertising
  - Winner of 3rd season of iPinyou Global Bidding Algorithm Competition in 2013 (Weinan Zhang), and the Best Paper Award of ADKDD 2014

- Dr. Jun Wang
  - Senior Lecturer (Associate Professor)
  - Information retrieval, collaborative filtering, computational advertising
AppNexus is one of the largest online advertising exchanges

- Offers one of the most powerful, open and customizable advertising technology platforms for both the buy and sell sides;
- Serves Google AdX, Microsoft Advertising Exchange, Interactive Media (Deutsche Telekom), Collective Exchange, and a lot more

Speaker: Dr. Kaihua Cai

- PhD in Mathematics from Caltech
- Research fellow at MSRI in Berkeley, California and Institute for Advanced Study in Princeton, New Jersey;
- Worked in finance at Chatham Financial, Goral Trading, and IV Capital;
- Joined AppNexus as a Data Scientist in 2012
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

      -----Break (20min)-----

3. Fighting publisher fraud

4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
The iPad is Apple’s new tablet computer.

Steven P. Jobs positioned the iPad as a device that sits between the laptop and the smart phone - and which does certain things better than both of them, like browsing the Web, reading e-books and playing video. There was enormous anticipation for them to charge for no nothing.

The iPad’s features and specifications, once the stuff of Internet myth, are now sharply in focus: The half-inch thick, 1.5-pound device will feature a

### Ads:

Tablet PCs, mobile phone etc.

### Ads by Google

**BlackBerry® Curve™ 8900**
The Thinnest & Lightest Full-QWERTY BlackBerry Smartphone Available.
www.blackberry.com/curve

**Tablet PC’s**
Mobile Tablet PC Solutions, Digitizer, Touch and Rugged Devices
www.camtechsystems.co.uk

**Win an Amazing New iPad**
Real-time Advertising: Selling ad slot per impression targeted to the user
Real-time Advertising: 
Selling ad slot per impression targeted to the user

Online advertising is now one of the fastest advancing areas in IT industry. In display and mobile advertising, the most significant development in recent years is the growth of Real-Time Bidding (RTB), which allows selling and buying online display advertising in real-time one ad impression at a time. Since then, RTB has fundamentally changed the landscape of the digital media market by scaling the buying process across a large number of available inventories. It also encourages behaviour (re-)targeting, and makes a significant shift toward buying focused on user data, rather than contextual data. A report from IDC shows that in 2011, global RTB based display ad spend increased by 237% compared to 2010, with the U.S.’s $2.2 billion RTB display spend leading the way. The market share of RTB-based spending of all display ad spending will grow from 10% in 2011 to 27% in 2016, and its share of all indirect spending will grow from 28% to 78%.

Scientifically, the further demand for automation, integration and optimization in RTB brings new research opportunities in the CIKM field. For instance, the much enhanced flexibility of
RTB workflow: less than 100ms

1. Bid Request (user, context)
2. Bid Response (ad, bid)
3. Ad Auction
4. Win Notice (paying price)
5. Ad (with tracking)
6. User Feedback (click, conversion, etc.)

0. Ad Request

User Information
- User Demography: Male, 25, Student, etc.
- User Segmentations: Ad science, London, etc.

Data Management Platform

Demand-Side Platform
- Advertiser: Booking.com

Webpage

User
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

   -----Break (20min)------

3. Fighting publisher fraud

4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 1.1 The background of Computational advertising

- Advertising has a long history
  - Egyptians used papyrus to make sales messages and wall posters (4000 BCE)
  - In the 18th century, ads started to appear in weekly newspapers in England
  - Thomas J. Barratt has been called "the father of modern advertising"
Glossaries

• **Real-Time Bidding** is an important aspect of **Programmatic buying**, which is getting more and more popular in **Display (related) advertising**. Another major part of **Online advertising** is **Sponsored search**

• An **Impression** is an ad display opportunity which generates when a **User** visits a webpage containing ad **Placements**

• The Publisher sends a bid request of this impression to an **Ad network**, or an **Ad exchange** via his **Supply side platform (SSP)**, then to **Demand side platforms (DSP)** to reach **Advertisers**

• Usually, DSPs contact **Data management platform (DMP)** to check the **Segments** of the current user, i.e., his intents or interests. Then a bid will be computed for the **Campaign**

• The payment among these entities is usually in **Cost per mille (CPM)**, but sometimes could be **Cost per click (CPC)** or **Cost per acquisition (CPA)**

• If the advertiser wins the impression, his **Creative** will be displayed to the user
The fundamental challenges

• To find the best **match** between a given user in a given context and a suitable advertisement?
• To achieve the best campaign **performance** (e.g., ROI) within the budget constraint?
• To generate the most **revenue** given the traffic and demand?
• To maintain a **healthy environment** so that users get less annoyed (both quality and quantity)?

Computational advertising, AZ Border, 2008
Dynamics of bid optimization in online advertisement auctions, C Borges et al. 2007
Dynamic revenue management for online display advertising, G Roels and K Fridgeirsdottir, 2009
Advertising in a pervasive computing environment, A Ranganathan and RH Campbell, 2002
Direct sales (since 1994)

- Advertisers and publishers talk to (4A) agencies
- Still popular in today’s marketplace
- Getting back to the market as the Programmatic Guaranteed

27th Oct 1994, AT & T on HotWired.com (78% CTR)

courtesy of Ad Age
Trading in ad networks (since 1996)

- After direct sales, some impressions will remain unsold (remnants)
- Small publishers cannot find buyers directly
Introducing the ad exchange (since 2009)

single ad network is easy

a few ad networks are manageable

hundreds of ad networks are nightmare

courtesy of www.liesdamnedlies.com
Introducing the ad exchange contd.

- Ad exchanges are marketplaces
- Advertisers and publishers have to rely on tools to connect
- Real-Time Bidding promotes user-oriented bidding
The simplified history of online (display) advertising

1. Direct sales (private contracts)
   - Real-time bidding flow
   - Not established

2. Centralised marketplace and real-time bidding since 2009

3. Aggregated demand and supply since 1996

Real-time Bidding for Online Advertising: Measurement and Analysis, S Yuan et al., 2013
The complex display ad eco-system

courtesy of LUMAscape 2014
A great visualisation

Behind the banner

http://o-c-r.org/adcells/
A visualization of the adtech ecosystem, Adobe, 2013
Part 1.2 RTB and an empirical study

- To understand the bidding behaviours in RTB auctions
- To present some research challenges
- To help to get familiar with RTB in the real-world

- The data is from production DSP & SSP based in UK

Real-time Bidding for Online Advertising: Measurement and Analysis, S Yuan et al., 2013
Impressions and clicks

The numbers of imp (left) and click (right) both show strong daily and weak weekly patterns, corresponding to the normal human activity.
Conversions

Daily periodic patterns for conv (left) and cvr (right) show that people are less likely to convert during late night.

The post-view conversions are NOT negligible.
Conversions (Frequency distribution)

The frequency against CVR plot from two different campaigns
Campaign 1 sets a frequency cap of 2-5 -> poor performance
Campaign 2 sets a frequency cap of 6-10 -> waste of budget
**Conversions** (Recency distribution)

The recency factor affects the CVR (left)
Campaign 1 sets a long recency cap -> waste of budget
Campaign 2 sets a short recency cap -> poor performance

The wide conversion window (right) challenges attribution models
Auctions (Bidding competition)

The winning bids peak at 8-10am due to intensive competition
Auctions (Change of winner)

The more bidders, the higher chance of winner change.
Auctions (Bids’ distribution)

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Accepted (p&gt;0.05)</th>
<th>Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD test per auction</td>
<td>0.343</td>
<td>0.657</td>
</tr>
<tr>
<td>AD test per placement</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>CQ test per auction</td>
<td>0.068</td>
<td>0.932</td>
</tr>
</tbody>
</table>

The commonly adopted assumption of Uniform distribution or Log-normal distribution were mostly rejected

- Anderson-Darling test for Normality
- Chi-squared test for Uniformity

Finding the best fit of bids’ distribution is important:
- Optimal reserve price
- Bid landscape forecasting
- etc.

And what’s the granularity? (placement, geographical location, time & weekday, etc.)
Budgeting and daily pacing

(a) Premature Stop
(b) Fluctuating Budget
(c) Uniform Pacing
(d) Traffic Based Pacing
(e) Performance Based Pacing

Real Time Bid Optimization with Smooth Budget Delivery in Online Advertising, KC Lee et al., 2013
Budgeting and daily pacing

No pacing (premature stop)

Even-daily pacing
A mixture of first and second price auctions

- A high soft floor price can make it first price auction (In RTB, floor prices are not always disclosed before auctions)
- In our dataset, 45% first price auctions consumed 55% budgets
- The complicated setting puts advertisers in an unfavourable position and could damage the ad eco-system

Assuming the soft floor can be zero.
References

• The History of Advertising: How Consumers Won the War for Their Attention, HubSpot, 2013
  blog.hubspot.com/the-history-of-advertising-war-for-consumer-attention-slideshare
• How Cluttered Is the Advertising Landscape, Really? HubSpot, 2013
  blog.hubspot.com/how-cluttered-is-advertising-landscape-timeline
• Navigating Planet Ad Tech, MIT Technology Review, 2013
  www.technologyreview.com/view/518551/the-evolution-of-ad-tech/
• Internet Advertising: An Interplay among Advertisers, Online Publishers, Ad Exchanges and Web Users, S Yuan et al., 2013
  arxiv.org/abs/1206.1754
• Ad exchanges: research issues, S Muthukrishnan, 2009
  sites.google.com/site/algoresearch/start2.pdf
• Behind the banner (A visualization of the adtech ecosystem), Adobe, 2013
  cmsummit.com/behindthebanner/
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

-----Break (20min)-----

   3. Fighting publisher fraud
   4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 2.1 Bidding strategy

• Bid optimisation methods
  – Introductions
  – CTR predictions
  – Bidders
Demand Side Platform (DSP)
Bidder in DSP

1. Call for Auction
2. Bid Request
3. (Ad, Bid Price)
4. Auction
5. Winning Ad

User

Web Page

Ad Exchange

DSP Bid Engine
Bidder in DSP

Relevant Data:
- Bid Request Data
- Campaign Ad Data

Real-Time Process:
- Feature Mapping
- Impression Evaluation
- Bid Calculation

Pre-Process:
- Campaign
- Target Rules
- And Data

CTR/CVR Estimator Training
Data Statistics
Bid Request Forecast
Winning Function Modelling

Bidding Function Optimisation

(Ad, Bid Price)
Input-process-output

• Input
  - logs for bid requests, impressions and events (browsing, clicks, conversions)
  - targeting rules
  - budgets and pacing preference
  - internal/external user data

• The Decision Engine
  - regression for effectiveness
  - cost efficiency

• Output
  - bid price
The bidding problem

• Objective
  – Bidding strategy $\pi^*$
  – Maximises the $KPI$ (usually – $CPA$)
  – Subject to the constraint $cost \leq budget$

• Components
  – $x$, the bid request, user and ad features
  – $\theta(x)$, the prediction function (e.g. CTR and CVR)
  – $b(\theta(x), x)$, the bidding function
  – $w(b, x)$, the win function (chances to win the impression)
Typical features example

- Optimising from the advertisers’ perspective
  - Ads (creative) known
  - Historical performance known
  - First party audience data available
  - Competition unknown

<table>
<thead>
<tr>
<th>Col #</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bid ID</td>
<td>0153000008...3f5a4f5121</td>
</tr>
<tr>
<td>2</td>
<td>Timestamp</td>
<td>20130218001203638</td>
</tr>
<tr>
<td>3</td>
<td>Log type</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>iPinyou ID</td>
<td>35605620124122340227135</td>
</tr>
<tr>
<td>5</td>
<td>User-Agent</td>
<td>Mozilla/5.0 (compatible; \ MSIE 9.0; Windows NT \ 6.1; WOW64; Trident/5.0)</td>
</tr>
<tr>
<td>6</td>
<td>IP</td>
<td>118.81.189.*</td>
</tr>
<tr>
<td>7</td>
<td>Region</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>City</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>Ad exchange</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>Domain</td>
<td>e80f4ec7...c01cd1a049</td>
</tr>
<tr>
<td>11</td>
<td>URL</td>
<td>hz55500000...3d6f275121</td>
</tr>
<tr>
<td>12</td>
<td>Anonymous URL ID</td>
<td>Null</td>
</tr>
<tr>
<td>13</td>
<td>Ad slot ID</td>
<td>2147689_8764813</td>
</tr>
<tr>
<td>14</td>
<td>Ad slot width</td>
<td>300</td>
</tr>
<tr>
<td>15</td>
<td>Ad slot height</td>
<td>250</td>
</tr>
<tr>
<td>16</td>
<td>Ad slot visibility</td>
<td>SecondView</td>
</tr>
<tr>
<td>17</td>
<td>Ad slot format</td>
<td>Fixed</td>
</tr>
<tr>
<td>18</td>
<td>Ad slot floor price</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>Creative ID</td>
<td>e39e178ffd...1ee56bcd</td>
</tr>
<tr>
<td>20</td>
<td>Bidding price</td>
<td>753</td>
</tr>
<tr>
<td>21</td>
<td>Paying price</td>
<td>15</td>
</tr>
<tr>
<td>22</td>
<td>Key page URL</td>
<td>a8be178ffd...1ee56bcd</td>
</tr>
<tr>
<td>23</td>
<td>Advertiser ID</td>
<td>2345</td>
</tr>
<tr>
<td>24</td>
<td>User Tags</td>
<td>123,5678,3456</td>
</tr>
</tbody>
</table>
Predicting CTR or CVR

• Regression
  – Generalised linear regression models (Logistic, Bayesian probit, FTRL-Proximal, etc.)
  – Rule based
    • [Dembczynski2008]
  – Tree based models (Random forest, Gradient boosting regression tree, etc.)
    • [Mohan2011]
  – Neural networks and deep learning
    • [Corrado2012]

\[ \hat{y}_i = \frac{1}{1 + \exp\{-w^T x_i\}} \]

\[ L(w) = \sum_i (-y_i \log(\hat{y}_i) - (1 - y_i)\log(1 - \hat{y}_i)) + \frac{\lambda}{2} \|w\|_2 \]
Bidding strategy

- **Baseline** (constant or random, usually for exploration)
- **Linear bidder** (proportional to the effect of the inventory)
- **Heuristic bidder** (linear with capping)
- **Multiplicative bidder** (a modifier vector applied to basic predictions)
- **Uniform bidder** (fixed bids based on the bid landscape)
- **Optimal bidder** (combining prediction with bid landscape)
Linear bidder

• To modify the bid proportional to the effect of the inventory

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

\[ B^* = B \times \Phi^* \]

• Strategy 2 (aggressive bidding)
  - \( \phi < 0.8 \rightarrow \phi = 0 \) (not bidding)
  - \( 0.8 \leq \phi \leq 1.2 \rightarrow \phi = 1 \)
  - \( \phi > 1.2 \rightarrow \phi = 2 \)

C Perlich et al., Bid Optimizing and Inventory Scoring in Targeted Online Advertising, 2012
A heuristic bidder

- To calculate the bid price based on CTR prediction with capping

\[
b_i = \min \left\{ \left( \frac{\hat{y}_i}{y_0} \right)^a y_0 b + c, d \right\}
\]

- \(a, b, c, d\) are empirical parameters learned from the training dataset
A heuristic bidder

\[ y_0 = 10^{-4} \]

\[ a = 2.8, b = 0.3, c = 10, d = 150 \]
Multiplicative bidder

The market prices can be well represented by the multiplication of two vectors

Features: country & hour of the day

Two bid adjustment dimensions with $m$ and $n$ possible settings

For each entry $(i,j) \in [m] \times [n]$, an advertiser is given a price $p_{ij}$ and value $v_{ij}$

He is required to specify a bid multiplier $r_i$ for each row $i$ and $c_j$ for each column $j$

The bid is then $r_i \cdot c_j$

Bateni et al., Multiplicative bidding in online advertising, 2014
An example

Brain algorithm from MediaMath

Advertisers can specify a modifier for a targeting combination
Multiplicative bidder

- A cell \((i,j)\) is captured if \(r_i \cdot c_j \geq p_{ij}\)
- The budget constraint \(\sum p_{ij} \leq B\)
- Objective to maximise \(\sum v_{ij}\)

**Algorithm 2:** Overview of \(O(\log m)\) approximation

Step 1:
Round down all of the \(p_i\)’s to the nearest powers of 2;
Cluster together rows with the same \(p_i\);
Reorder the clusters in increasing \(p_i\);

Step 2:
Reorder the rows within each cluster in increasing values;
Compute \(\text{OPT}(B/4)\), the individual bidding optimum with budget \(B/4\);

Step 3:
\textbf{for} \(h=1,2,\ldots,m\) \textbf{do}
\hspace{1cm} \text{ALG}_h \leftarrow \emptyset;
\hspace{1cm} \text{Insert into ALG}_h \text{ all height-}h \text{ towers of OPT}(B/4);
\hspace{1cm} \text{Insert into ALG}_h \text{ all height-}h \text{ towers in the strips above the ones in the last line;}
\textbf{end}

Output the best \(\text{ALG}_h\) as \(\text{ALG}\)
Uniform bidder (in sponsored search)

- A bipartite graph G on the two vertex sets K and Q
- Matches of \( q \in Q \) are neighbours of \( q \) in K
- Click function: \( \text{clicks}_q(\cdot) \)
- Cost function: \( \text{cost}_q(\cdot) \)
- Objective to maximise: \( \Sigma_q \text{clicks}_q(b_q) \)
- Subject to: \( \Sigma_q \text{cost}_q(b_q) \leq U \)
- Nearly all formulation of this optimisation problem is NP-hard
Uniform bidder (in sponsored search)

- An approximation: two-bid uniform strategy
  - To bid $b_1$ or $b_2$ on all keywords randomly

A bid landscape example, where clicks are plotted as a function of cost; The effective bids are discrete.

Similarly, a single-bid uniform strategy could be found with worse approximation but better computational complexity.

Feldman et al., Budget Optimization in Search-Based Advertising Auctions, 2008
Optimal Bidder: Problem Definition

Input: bid request include
Cookie information
(anonymous profile), website
category & page, user
terminal, location etc
Output: bid price
Considerations: Historic data,
CRM (first party data), DMP
(3rd party data from Data
Management Platform)

What is the optimal bidder given
a budget constraint?
e.g., Maximise

\[ R = \sum (Clk + Conv \times weight) \]

Subject to the budget constraint
Optimal bidder: the formulation

- **Functional Optimisation Problem**
  
  - **Dependency assumption:** $x \rightarrow \theta \rightarrow b \rightarrow w$

  \[
  b(\_)_{\text{ORTB}} = \arg \max_{b(\_)} NT \int_\theta \theta w(b(\theta))p_\theta(\theta)d\theta \\
  \text{subject to} \quad NT \int_\theta b(\theta)w(b(\theta))p_\theta(\theta)d\theta \leq B: \leq B
  \]

- **Solution: Calculus of variations**

  \[
  \mathcal{L}(b(\theta), \lambda) = \int_\theta \theta w(b(\theta))p_\theta(\theta)d\theta - \lambda \int_\theta b(\theta)w(b(\theta))p_\theta(\theta)d\theta + \frac{\lambda B}{NT} \\
  \lambda w(b(\theta)) = \left[ \theta - \lambda b(\theta) \right] \frac{\partial w(b(\theta))}{\partial b(\theta)}
  \]

W Zhang et al., Optimal Real-Time Bidding for Display Advertising, KDD 2014
Win function
Optimal bidder: the solution

(a) Winning function 1.

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

(b) Bidding function 1.

\[ b_{ORTB1}(\theta) = \sqrt{\frac{c}{\lambda} \theta + c^2 - c} \]
Optimal bidder: the solution

(a) Winning function 2.

\[ w(b(\theta)) = \frac{b^2(\theta)}{c^2 + b^2(\theta)} \]

(b) Bidding function 2.

\[ b_{ORTB2}(\theta) = c \cdot \left[ \left( \frac{\theta + \sqrt{c^2 \lambda^2 + \theta^2}}{c \lambda} \right)^{\frac{1}{3}} - \left( \frac{c \lambda}{\theta + \sqrt{c^2 \lambda^2 + \theta^2}} \right)^{\frac{1}{3}} \right] \]
Optimal bidder: the solution
Experiments

Winner of the first global Real-time Bidding algorithm contest 2013-2014

W Zhang et al., Optimal Real-Time Bidding for Display Advertising, KDD 2014
Beyond CTR: Alternative metrics

• **Top funnel metrics** (to gain brand awareness)
  - brand recall (awareness uplift)
  - branded search
  - direct website traffic

• **Mid funnel metrics** (to educate and engage the prospects)
  - cost per new website visitor
  - page view & form uplift

• **Bottom funnel metrics** (to generate value both online and offline)
  - total conversion
  - cost per conversion
  - opportunity contribution (interested but not converted yet)
  - revenue
Beyond CTR: Transfer learning

- The Problem
  - CTR is no good metrics but CVR is too low

- Task
  - To train on site visits

- Challenge
  - Which site visits, and weights?
  - Data availability

- Solution
  - Similarity (contextual as a priori, Bayesian)

Dalessandro et al., Evaluating and Optimizing Online Advertising: Forget the Click, But There are Good Proxies, 2012
References

- Optimal bidding on keyword auctions, B Kitts and B Leblanc, 2004
- Stochastic gradient boosted distributed decision trees, J Ye et al., 2009
- Web-scale bayesian click-through rate prediction for sponsored search advertising in Microsoft's Bing search engine, T Graepel et al., 2010
- A novel click model and its applications to online advertising, ZA Zhu et al., 2010.
- Web-search ranking with initialized gradient boosted regression trees, A Mohan et al., 2011
- A Gentle introduction to random forests, ensembles, and performance metrics in a commercial system, D Benyamin, 2012, goo.gl/XpbkDY
- Deep networks for predicting ad click through rates, G Corrado, 2012
- Click modeling for display advertising, O Chapelle, 2012, goo.gl/fNpZHv
- Causal reasoning and learning systems, L Bottou and E Portugaly, 2012
- Ad click prediction: a view from the trenches, HB McMahan et al., 2013
- Deep learning, yesterday, today, and tomorrow, K Yu et al., 2013
- Deep learning of representations: looking forward, Y Bengio, 2013
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

   ------Break (20min)------

3. Fighting publisher fraud

4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 2.2
Inventory management and reserve price optimisation

• Typical revenue models
• Ad density optimal control
• Reserve price optimisation
Typical revenue models for the supply side

- Subscription access to content (FT.com)
- Pay Per View access to document (Downloading a paper outside the campus)
- CPM display advertising on site
- CPC advertising on site (Google AdSense)
- Sponsorship of site sections or content types (typically fixed fee for a period)
- Affiliate revenue (Compare shopping, CPA/CPC)
  
- Subscriber data access for marketing (VISA & MasterCard)
- User contributed data for marketing (Surveys)
Examples

Yield management tools in AppNexus

Ad density

The task:
• To find the optimal advertising density (number of ad placements) for a given website

The challenges:
• Users’ preference model
• Expected CPM
• Competition

The assumption:
• Using real-time bidding only
Some websites do not rely on ads to compensate the maintenance cost

- Government
- Education
- Most of .org
All ads

- Created by Alex Tew in 2005
- Selling 100k 100-pixels at $100 each
- Sold out in 4 months
- Almost 0% CTR

courtesy of www.milliondollarhomepage.com
Ad density: example

What is another name for freezing point?

Another name for freezing point is melting point since the temperature at which a substance freezes is also the temperature at which it melts, going in the other direction.

Can you answer these science of matter and energy questions?

1. How many atoms does a carbohydrate have?
2. What accounts for variable re-absorption of water and re-absorption or secretion of sodium potassium hydrogen and bicarbonate ions?
3. How many atoms are in one molecule of carbohydrate?
4. Can substances be compounds and elements?

Research your answer:
Ad density: example

Question: is it reasonable to put on so many more ads?
Models and solutions

• Assume the following
  – Unit payoff (CPM) $c$
  – Ad density $\rho$
  – Impressions $x$
  – Content cost $k$

• The cumulative revenue

$$R(T) = \sum_{t}^{T} \left( cx\rho - \frac{k(1 - \rho)^2}{2} \right)$$

• The state transition function

$$x(t) = \frac{1}{L-1} \sum_{l}^{L} \rho_{l}(t)x_{l}(t) - m\rho_{i}(t)x_{i}(t) + h$$
A monopoly case

Dewan et al., Management and valuation of advertisement-supported web sites, 2003
Reference

- Management and valuation of advertisement-supported web sites, RM Dewan, 2003
- Optimal pricing and advertising policies for web services, S Kumar et al., 2004
- Is revamping your web site worthwhile? EY Huang, 2005
- An economic analysis of ad-supported software, BJ Jiang, 2007
- Dynamic pricing and advertising for web content providers, S Kumar and SP Sethi, 2009
- Pricing display ads and contextual ads: Competition, acquisition, and investment, YM Li and JH Jhang-Li, 2009
- Dynamic ad layout revenue optimization for display advertising, H Cheng et al., 2012
- Automatic ad format selection via contextual bandits, L Tang et al., 2013
Reserve price optimisation

The task:
• To find the optimal reserve prices

The challenge:
• Practical constraints v.s common assumptions (bids’ distribution, bidding private values, etc.)

S Yuan et al., An Empirical Study of Reserve Price Optimisation in Display Advertising, 2014
Why

• Suppose it is second price auction
  - Normal case: \( b_2 \geq \alpha \)
  - Preferable case: \( b_1 \geq \alpha > b_2 \) (it increases the revenue)
  - Undesirable case: \( \alpha > b_1 \) (but there is risk)
An example

• Suppose: two bidders, private values drawn from Uniform[0, 1]

• Without a reserve price (or $a = 0$), the payoff $r$ is:

$$r = E[\min(b_1, b_2)] = 0.33$$

• With $a = 0.2$:

$$r = E[\min(b_1, b_2) | b_1 > 0.2, b_2 > 0.2] + 0.32 \times 0.2 = 0.36$$

• With $a = 0.5$:

$$r = E[\min(b_1, b_2) | b_1 > 0.5, b_2 > 0.5] + 0.5 \times 0.5 = 0.42$$

• With $a = 0.6$:

$$r = E[\min(b_1, b_2) | b_1 > 0.6, b_2 > 0.6] + (0.6 \times 0.4) \times 2 \times 0.6 = 0.405$$

Paying the second highest price  Paying the reserve price
The optimal auction theory

- In the second price auctions, advertisers bid their private values 
  \([b_1, \ldots, b_K]\)
- Private values -> Bids’ distributions \(F(b) = F_1(b_1) \times \cdots \times F_K(b_K)\)
  - Uniform
  - Log-normal
- The publisher also has a private value \(V_p\)
- The optimal reserve price is given by:
  \[\alpha - \frac{1 - F(b)}{F'(b)} - V_p = 0\]

Questions:
- How are advertisers bidding?
- Does Uniform/Log-normal fit well?
Bidding could be irrational

- They usually use a private regression model (No access to publishers)
- Perhaps they don’t even know it! (Just try to maximise the ROI)

Many advertisers bid at fixed values
(Think about a decision tree)

And they come and go
(with different lifetime span)

S Yuan et al., An Empirical Study of Reserve Price Optimisation in Display Advertising, 2014
Uniform/Log-normal distributions do NOT fit well

Test at the placement level
(because we usually set reserve prices on placements)

Test at the auction level

- Chi-squared test for Uniformity
- Anderson-Darling test for Normality

S Yuan et al., An Empirical Study of Reserve Price Optimisation in Display Advertising, 2014
Results from a field experiment

- On Yahoo! Sponsored search
- Using the Optimal Auction Theory

<table>
<thead>
<tr>
<th>Table 7: Restricted sample (optimal reserve price &lt; 20¢)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Number of keywords (T - treatment group)</td>
</tr>
<tr>
<td>Number of keywords (C - control group)</td>
</tr>
<tr>
<td>(Mean change in depth in T) - (mean change in depth in C)</td>
</tr>
<tr>
<td>(Mean change in revenue in T) - (mean change in revenue in C)</td>
</tr>
<tr>
<td>Estimated impact of reserve prices on revenues</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 8: Restricted sample (optimal reserve price ≥ 20¢)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Number of keywords (T - treatment group)</td>
</tr>
<tr>
<td>Number of keywords (C - control group)</td>
</tr>
<tr>
<td>(Mean change in depth in T) - (mean change in depth in C)</td>
</tr>
<tr>
<td>(Mean change in revenue in T) - (mean change in revenue in C)</td>
</tr>
<tr>
<td>Estimated impact of reserve prices on revenues</td>
</tr>
</tbody>
</table>

Levin and Smith, Optimal Reservation Prices in Auctions, 1996
Our solution

• A dynamic and one-shot game between the winner (w) and the publisher (p)

• Extension form representation
  
  – Information nodes:
    • $I_1$: Auction succeeded: the winning bid $b_1$ is higher
    • $I_2$: Auction failed: the reserve price $\alpha$ is higher
  
  – Actions:
    • $a_{w1}$: to increase $b_1$ so that $b_1 \geq \alpha$
    • $a_{w2}$: to increase $b_1$ so that $b_1 < \alpha$
    • $a_{w3}$: to decrease $b_1$ so that $b_1 \geq \alpha$
    • $a_{w4}$: to decrease $b_1$ so that $b_1 < \alpha$
    • $a_{p1}$: to increase $\alpha$ so that $\alpha \geq b_1$
    • $a_{p2}$: to increase $\alpha$ so that $\alpha < b_1$
    • $a_{p3}$: to decrease $\alpha$ so that $\alpha \geq b_1$
    • $a_{p4}$: to decrease $\alpha$ so that $\alpha < b_1$
1) Expected payoff of advertiser, publisher

2) Payoff for the advertiser could be negative if one has been bidding the max price ($a_{w1}$: to increase $b_1$ so that $b_1 \geq \alpha$)

3) One won’t do that, so discounted publisher’s payoff
Findings

12.3% better than the second best
28.5% better than the optimal auction theory
The continuous bidding activity

An outlier
(Triggered by some random action)

The unchanged budget allocation

The unchanged bidding pattern

S Yuan et al., An Empirical Study of Reserve Price Optimisation in Display Advertising, 2014
References

• On optimal reservation prices in auctions, Engelbrecht-Wiggans, 1987
• Optimal reservation prices in auctions, Levin and Smith, 1996
• Auction theory: a guide to the literature, Klemperer, 1999
• Reserve prices in internet advertising auctions: a field experiment, Ostrovsky and Schwarz, 2009
• Auction theory 2nd edition, Krishna, 2009
• Optimal reserve price for the generalized second-price auction in sponsored search advertising, Xiao et al., 2009
• Optimal auction design and equilibrium selection in sponsored search auctions, Edelman and Schwarz, 2010
• Optimal auction design in two-sided markets, R Gomes, 2011
Let’s have a break

20min
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

------Break (20min)------

3. Fighting publisher fraud

4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 2.3 Fighting publisher fraud

- Overview
- Detecting fraud
- A classification algorithm
- A graph clustering algorithm
Publisher fraud in Online advertising

- **Definition**
  - Publisher generates non-human traffic to sell for money

- **How to fraud**
  - Publisher sets up bots to visit its own websites
  - Publisher installs malware to other’s machine and the malware visits the websites without the machine’s owner knowing it.

- **Category**
  - Impression fraud
  - Click fraud

- **Impact**
  - 10% - 30% of RTB impressions estimated to be bot traffic
  - Cost advertisers billions of dollar in 2014
Characteristics of publisher fraud

- Bot is Non-human
  - Bot usually uses old version OS/browser
  - Bot generates traffic 24 hours a day
- Greedy
  - It has to be large scale to be profitable
- Require high liquidity
  - Usually sold cheap: as cheap as $1 for 100K imps
  - Has long daisy chain, deeply embedded
- Little transparency
  - Domain obfuscation
Fighting publisher fraud: the old fashion way

- Put the police on the street
  - Manually eyeball the webpage
  - Verify the address on the Google map
- Follow how the money flows
- This approach just can’t scale and is not sustainable
- Better to hand over this problem to data scientist
Data Scientists’ Toolbox

- Anomaly detection
  - Online algorithm
  - Offline algorithm
- Classification algorithm
  - Human traffic vs. bot traffic
  - Human clicks vs. bot clicks
- Clustering algorithm
  - Bots could display dramatically different behavior.
- Language process technique
  - Fraudulent websites often scrape content from each other or legit websites
Problem set

• What can be fraudulent
  - Cookie ID
  - IP address: both audience IP and web site host IP
  - URL
  - Ad placement
  - Publisher
Features

- Audience-related
- Content-related
- Business-related
- Audience-content interaction
- Audience-business interaction
Applying tools to problems

- Cookie ID
  - Anomaly detection both online and offline
- Audience IP address
  - Anomaly detection both online and offline
- Web site host IP with Cookie ID
  - Bipartite graph, clustering algorithm on graph
  - Both supervised and unsupervised
- URL
  - Outlier detecting against Alexa ranking
- Ad placement, Publisher
  - Classification with SVM or logistic regression
Example #1: Classification on publisher

- Off-the-shelf algorithm
  - For example Python scikit-learn
  - Logistic regression with L1 regularization
  - Support vector machine
- Training data
  - Known good and bad publishers from other algorithms
  - Collect multiple daily data point for each publisher
- Features
  - Cover both impression and clicks
- Performance
  - 96% prediction accuracy
Example #1: Result and Caveat

- SVM performs slightly better than Logistic Regression (LR), but LR is preferred for its transparency.
- Produces quality score for each publisher.
- The power of the algorithm is as good as the training data.
  - Good at combining the strength of individual detection algorithms.
  - Unlikely to find brand-new fraud pattern.
- More scrutiny and manual checking is needed before blacklisting fraudulent publishers.
Example #2: Graph clustering algorithm

- Key observation:
  • Even the major sites only share at most 20% cookie_id within a few hours, let alone those long tail sites.

- Define a graph:
  • Node: site
  • Weighted edge: user overlap ratio of two sites

- Cluster this weighted undirected graph

- Any big cluster with long tail sites are all fraud.

- Main reference for this approach:
  • Using Co-visitation Networks For Classifying Non-Intentional Traffic by six authors at m6d Research
Example #2: algorithm implementation

Cluster nodes on weighted undirected graph

Let $e_{ij}$ be the fraction of weights of edges connecting community $i$ and community $j$

Define $a_i = \sum_j e_{ij}$ and $Q = \sum_i (e_{ii} - a_i^2)$

Goal: maximize $Q$

Note: $a_ia_j = e_{ij}$ on average for random graph

"Fast algorithm for detecting community structure in networks" by M. E. J. Newman
Example #2: How to maximize $Q$

- A bottom-up greed algorithm starts with each node as its own community.
- At each step, merge two communities which increase $Q$ most
  - The change of $Q$ at each step could be negative
- Pick communities for the maximum of $Q$
- Result in dendrograph
Example #2: A beautiful picture
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation

   ------Break (20min)------

3. Fighting publisher fraud

4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 2.4 Programmatic Guaranteed and Ad Options

- RTB
  - Combined with the forward market
  - Combined with the futures & options exchange
(RTB) Ads prices are volatile

The price movement of a display opportunity from Yahoo! ads data
Under GSP (generalized second price auction)
Hedge the price risk

• Need Ad’s Futures Contract and Risk-reduction Capabilities
  – Technologies are constrained mainly to “spots” markets, i.e., any transaction where delivery takes place right away (in Real-time Advertising and Sponsored Search)
  – No principled technologies to support efficient forward pricing & risk management mechanisms
Solution 1:
Combine RTB with **Forward Market**, which pre-sell inventories in advance with a fixed price.

Solution 2:
If we got **Futures Exchange** or provide **Option Contracts**, advertisers could *lock in* the campaign cost and Publishers could *lock in* a profit in the future.
Solution 1:
RTB with Forward Programmatic Guaranteed Market

Figure 1: A systematic view of programmatic guarantee (PG) in display advertising: \([t_0, t_n]\) is the time period to sell the guaranteed impressions that will be created in future period \([t_n, t_{n+1}]\).
Optimization objective

\[
\max \left\{ \int_0^T (1 - \omega \kappa) p(\tau) \theta(\tau, p(\tau)) f(\tau) d\tau + \left( S - \int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau \right) \phi(\xi) \right\},
\]

\[G = \text{Expected total revenue from guaranted selling minus expected penalty of failing to delivery}\]

\[H = \text{Expected total revenue from RTB}\]

\[
s.t. \quad p(0) = \begin{cases} 
\phi(\xi) + \lambda \psi(\xi), & \text{if } \pi(\xi) \geq \phi(\xi) + \lambda \psi(\xi) \\
\pi(\xi), & \text{if } \pi(\xi) < \phi(\xi) + \lambda \psi(\xi),
\end{cases}
\]

where

\[
\xi = \frac{\text{Remaining demand in } [t_n, t_{n+1}]}{\text{Remaining supply in } [t_n, t_{n+1}]} = \frac{Q - \int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau}{S - \int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau}.
\]
Figure 9: An empirical example of AdSlot14: (a) the optimal dynamic guaranteed prices; (b) the estimated daily demand; (c) the daily demand calculated based on the actual bids in RTB on the delivery date; (d) the winning bids and payment price in RTB on the delivery date; (e) the comparison of revenues [see Table 8 for summary of notations B-I, B-II, B-III, R-I, R-II]. The parameters are: \( Q = 17691; S = 2847; \alpha = 2.0506; \beta = 0.2; \zeta = 442; \eta = 0.2; \omega = 0.05; \kappa = 1; \gamma = 0.4240; \lambda = 2. \)
An **Ad Option** is a contract in which the option publisher grants the advertiser **the right but not the obligation to** enter into a transaction either buy or sell an underlying ad slot at a specified price on or before a specified date.

The specified pre-agreed price is called **strike price** and the specified date is called **expiration date**. The option seller grants this right in exchange for a certain amount of money at the current time is called **option price**.

---

J Wang and B Chen, Selling futures online advertising slots via option contracts, WWW 2012.
Ad options: Benefits

<table>
<thead>
<tr>
<th>Advertisers</th>
<th>Publishers</th>
</tr>
</thead>
<tbody>
<tr>
<td>• secure impressions delivery</td>
<td>• sell the inventory in advance</td>
</tr>
<tr>
<td>• reduce uncertainty in auctions</td>
<td>• have a more stable and predictable revenue over a long-term period</td>
</tr>
<tr>
<td>• cap cost</td>
<td>• increase advertisers’ loyalty</td>
</tr>
</tbody>
</table>
Ad options contd.

Sells a list of ad keywords via a multi-keyword multi-click option

<table>
<thead>
<tr>
<th>multi-keyword multi-click option (3 month term)</th>
</tr>
</thead>
<tbody>
<tr>
<td>upfront fee (m = 100)</td>
</tr>
<tr>
<td>£5</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Submits a request of guaranteed ad delivery for the keywords ‘MSc Web Science’, ‘MSc Big Data Analytics’ and ‘Data Mining’ for the future 3 month term [0, T], where T = 0.25.

Pays £5 upfront option price to obtain the option.

Timeline

search engine

online advertiser

B Chen et al., Multi-Keyword Multi-Click Advertisement Option Contract for Sponsored Search, 2013
Exercising the option

Pays £1.80 to the search engine for each click until the requested 100 clicks are fully clicked by Internet users.

Exercises 100 clicks of ‘MSc Web Science’ via option.

Reserves an ad slot of the keyword ‘MSc Web Science’ for the advertiser for 100 clicks until all the 100 clicks are fully clicked by Internet users.
Not exercising the option

If the advertiser thinks the fixed CPC £8.67 of the keyword ‘Data Mining’ is expensive, he/she can attend keyword auctions to bid for the keyword as other bidders, say £8.

Pays the GSP-based CPC for each click if winning the bid.

Selects the winning bidder for the keyword ‘Data Mining’ according to the GSP-based auction model.

Timeline

online advertiser

search engine
Risk Hedge when Ad Options and RTB spot are combined

B Chen and J Wang, A Lattice Framework for Pricing Display Ad Options with the Stochastic Volatility Underlying Model, Technical Report, 2014
An empirical example of search engine's revenue for the keyword *equity loans*

Additional references

- Internet advertising and the generalized second price auction: selling billions of dollars worth of keywords, B Edelman, 2005
- Price cycles in online advertising auctions, X Zhang and J Feng, 2005
- Budget optimization in search-based advertising auctions, J Feldman et al., 2007
- The economics of the online advertising industry, DS Evans, 2008
- Expressive banner ad auctions and model-based online optimization for clearing, C Boutilier et al., 2008
- Computational advertising, AZ Broder, 2008
- Algorithmic methods for sponsored search advertising, J Feldman and S Muthukrishnan, 2008
- Internet ad auctions: Insights and directions, S Muthukrishnan, 2008
- The online advertising industry: economics, evolution, and privacy, DS Evans, 2009
- Ad exchanges: Research issues, S Muthukrishnan, 2009
- Adaptive bidding for display advertising, A Ghosh et al., 2009
- The arrival of real-time bidding, Google, 2011
Additional references

• Algorithms and strategies for web advertising, P Papadimitriou, 2011
• OpenRTB API specification, IAB, 2012
• Targeted, not tracked: client-side profiles and privacy-friendly behavioral advertising, M Bilenko and M Richardson, 2012
• Computational advertising in social networks, A Bhasin, 2012
• Size, labels, and privacy in targeted display advertising, C Perlich, 2012
• Estimating conversion rate in display advertising from past performance data, K Lee et al., 2012
• Handling forecast errors while bidding for display advertising, KJ Lang et al., 2012
• Marketing campaign evaluation in targeted display advertising, J Barajas et al., 2012
• Ad exchange-proposal for a new trading agent competition game, M Schain and Y Mansour, 2012
• Auctions for online display advertising exchanges: approximations and design, S Balseiro et al., 2012
• Real-time bidding for online advertising: measurement and analysis, S Yuan et al., 2013
Additional references

• Impression fraud in on-line advertising via pay-per-view networks, K Springborn and P Barford, 2013
• An overview of computational challenges in online advertising, RE Chatwin, 2013
• Competition and yield optimization in ad exchanges, SR Balseiro, 2013
• Internet advertising revenue report, IAB and PwC
Outline

1. The background of Computational advertising

2. Research problems and techniques
   1. Bidding strategy Optimisation
   2. Inventory management and floor prices optimisation
      ------Break (20min)------
   3. Fighting publisher fraud
   4. Programmatic Guaranteed and Ad Options

3. Datasets, tools, and platforms
Part 3.1 Datasets

- iPinYou Real-Time Bidding Dataset
  - data.computational-advertising.org
  - 5.87GB
  - Bid requests, impressions, clicks, conversions

- Criteo conversion logs
  - labs.criteo.com/2014/08/criteo-release-public-datasets
  - O Chapelle, Modeling Delayed Feedback in Display Advertising, 2014
  - 534MB
  - Conversion logs
Datasets

- **Criteo - Kaggle Display Advertising Challenge Dataset**
  - 4.3G
  - CTR prediction
- **Avazu – Kaggle CTR Prediction Challenge**
  - www.kaggle.com/c/avazu-ctr-prediction
  - Deadline: 9th Feb 2015
- **Open advertising dataset**
  - code.google.com/p/open-advertising-dataset
  - Data from Google AdWords
Datasets

- Internet Advertisements Data Set
  - [archive.ics.uci.edu/ml/datasets/Internet+Advertisements](archive.ics.uci.edu/ml/datasets/Internet+Advertisements)
  - Display ads, webpage and ad context

- Farm Ads Data Set
  - [archive.ics.uci.edu/ml/datasets/Farm+Ads](archive.ics.uci.edu/ml/datasets/Farm+Ads)
  - Ad creative and landing pages
Datasets

• Webscope from Yahoo!
  – webscope.sandbox.yahoo.com
  – Sponsored search, CTR prediction

• KDD CUP 2012 Track 2
  – Tencent search engine, sponsored search, CTR prediction
Part 3.2 Tools

• OpenRTB API specification
  - openrtb.github.io/OpenRTB
  - A good description of protocols and data exchanges

• RTBKit
  - rtbkit.org
  - An open source Real-time bidding framework
  - No (intelligent) bidding algorithms included
  - Production level design and implementation
  - Takes effort to setup
Tools

- Revive Adserver
  - www.revive-adserver.com
  - An open source Ad server
  - Formerly known as OpenX Source

- Orbit Open Ad Server
  - orbitopenadserver.com
  - An open source Ad server

- mAdserver
  - www.madserve.org
  - An open source mobile Ad server
  - No longer being developed