Data-Driven Reserve Prices for Social Advertising Auctions at LinkedIn

Tingting Cui
Lijun Peng
David Pardoe
Kun Liu
Deepak Kumar
Deepak Agarwal

Relevance @ LinkedIn
KDD 2017
Introduction
LinkedIn news feeds consist of both organic updates and sponsored content (SC)

- The number of SC LinkedIn can show to members is limited

- Different positions have different desirability

- Auctions: allocating positions
How Sponsored Content Auction Works

What location do you want to target? (required)

Select specific targeting criteria to zero in on your ideal audience:

- Company name
- Company industry
- Company size
- Job title
- Job function
- Job seniority
- Member schools
- Fields of study
- Degrees
- Member skills
- Member groups
- Member gender
- Member age
- Years of experience
- Company followers
- Company connections
How Sponsored Content Auction Works

**Targeting**

- Advertisers
- Geo: US
- Title: SWE
- Skill: Java

**GSP Auction**

- Advertiser — Ad
- QWB(1)
- QWB(2)
- QWB(3)

**Serving**

- User

What & Why Reserve Price

• What is reserve price
  • The minimum bid to enter auction
  • The minimum price to pay

• Why reserve price
  • Protect valuable inventory and optimize revenue
    • Too high - advertisers are discouraged from participating in auctions, resulting in low sell-through rate and revenue
    • Too low - poor price support in lack of competition
Why Reserve Price for Sponsored Content

- Goal: scalable data-driven reserve price system
  - **Data-driven** – Rate card based reserve prices were used when LinkedIn first launched SC, which does not reflect market dynamics now
  - **Pricing support** – Protect valuable LinkedIn inventory, especially in regional markets with low liquidity
  - **Scalable** - The scale of LinkedIn’s social advertising imposes significant challenges in designing an effective system to compute & serve reserve prices

<table>
<thead>
<tr>
<th>Members</th>
<th>Monthly Active Members</th>
<th>Daily Ad Requests</th>
</tr>
</thead>
<tbody>
<tr>
<td>500+M</td>
<td>200+M</td>
<td>100+M</td>
</tr>
</tbody>
</table>
This Talk

• A scalable **regression model** which predicts the distribution of bidders’ valuations to derive revenue maximizing **reserve price at the user level**

• A **novel mechanism** that produces the **segment-level reserve price** considering the trade-off between our revenue and advertisers’ satisfaction

• **Field experiments** from emerging and developed markets show that reserve prices improve revenue metrics and auction health
Reserve Price Optimization
Two Stage Reserve Prices

User-level Prices

Campaign Target

Quantile

Campaign-level Prices
Reserve Price Optimization

- Assumptions
  1. Advertiser valuation distribution is known to LinkedIn and advertisers
  2. Advertiser valuation distribution is log normal
  3. Advertisers bid their true valuation
    - Click probability declines more dramatically by position
    - Advertisers have an incentive to bid their true valuation

- The revenue optimizing reserve price \( r^* \) (Myerson 1981)
  \[ r^* = \frac{1 - F(r^*)}{f(r^*)} \]
  where \( F \) and \( f \) are CDF/PDF of valuation distribution
Fitting Valuation Distribution

- Fit valuation distribution for a **user** via linear regression
  - Fit log of valuations ($V$) against users’ profile attributes ($X$) via linear regression
    \[
    \log V = X^T \beta + \varepsilon, \; \varepsilon \sim N(0, \Sigma).
    \]
  - $X$: a user-by-attribute binary matrix indicating the absence/presence of profile attributes for a user.
  - Following the assumption that bids ($B$) are asymptotically equal to valuations
    \[
    \hat{\beta} = (X^T X + \lambda I)^{-1} (X^T \log B).
    \]
- Fit separate regression models for different geographic markets to reflect different market dynamics
User-Level Reserve Prices

- Run optimization at user level
  - Indivisible and mutually exclusive unit

- Linear regression model to predict valuation distribution for each

- Numerically solve \( r^* = (1 - F(r^*)) / f(r^*) \) for each user with fitted \( F \) and \( f \) to find the optimal reserve price
Campaign-Level Reserve Prices

- Serve at campaign level
  - Easy to communicate with advertisers
  - Regulate bidding behavior
  - Discourage cherry-picking

- Campaign-level reserve price: quantile of member-level reserve prices
  - The reserve price for a campaign targeting a user segment $S$:
    $$\hat{r}_s = \sup\{r > 0|\Pr\{R_S \leq r\} \leq p\},$$
    $0 < p < 1$ is the quantile of choice
Implementation
Engineering Implementation

- **Challenge** – Scale of LinkedIn’s user base and ads business
- **Component** - Offline Hadoop pipeline + online web service

**Offline Hadoop Pipeline**

- Read the latest member profile and ad auction logs
- Fit the bidder valuation distributions & compute user-level reserve prices
- Store the optimal reserve price for each user in **Pinot**, a realtime distributed OLAP datastore, which is used at LinkedIn to deliver scalable real time analytics with low latency

**Online Web Service**

- Ad server calls Pinot store to retrieve campaign level floor price at serving time
- Campaigns with bid below the reserve price for the visiting member are removed from the auction
- The remaining campaigns are charged by $\text{Max(second price cost, campaign level floor price)}$
Architecture of Reserve Price System

- Auction Log
- User Profile
- Offline Data Pipeline
- Linear Regression
- User Level Reserve Price
- User Dimension Aggregator
- Real-time distributed OLAP data store
- Advertiser Campaign Create/Update Requests
- Campaign Reserve Price
- Online Campaign Service
- Campaign Level Reserve Price
Results
Experiment in Emerging Markets

- Emerging markets where sell-through-rates are relatively low
- Compared against the legacy rate-card based approach
- Results
  - **Lower reserve prices:** 20-60% drop depending on geographic market
  - **Significant increase in demand:** the percent of auctions with at least one participant increased by 30-60%
  - **Positive revenue impact:** the increased demand quickly made up for the lower price
Results from Emerging Markets

Figure 1: Percentage of auctions with at least one participant in emerging markets, normalized so the starting value is 1.0.
Experiment in Developed Markets

- Developed markets where sell-through-rates are relatively high
  - Report results from CPC campaigns targeting the US market only
  - Stratified sampling to balance advertiser’s type and remove outlier campaigns

- Revenue-related metrics - direct revenue impact
  - +1.7% lift in median bid, +2.2% lift in median CPC for campaigns bidding above reserve prices

- Advertiser-centric metrics – advertiser experience
  - +17% reduction in churn rate, mainly attributed to campaigns bidding at the reserve price, as they now tend to submit more realistic bids => more likely to win in auctions and stay active
Results from Developed Markets

<table>
<thead>
<tr>
<th>Campaign group</th>
<th>Increase in median bid</th>
<th>Increase in median revenue per click</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bid at reserve price</td>
<td>36.0%</td>
<td>36.0%</td>
</tr>
<tr>
<td>Bid above reserve price</td>
<td>1.7%</td>
<td>2.2%</td>
</tr>
</tbody>
</table>

Table 1: Changes in median bid and revenue per click, treatment v.s. control.

<table>
<thead>
<tr>
<th>Campaign group</th>
<th>Abandonment rate</th>
<th>Churn rate</th>
<th>New campaigns per advertiser</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1.03</td>
<td>0.83</td>
<td>1.07</td>
</tr>
<tr>
<td>Control</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 2: Advertiser-Centric Metrics, normalized so that the control group always have values of 1.0.
Future Work

• Address Overestimation
  • Valuation is overestimated, as valuation below existing floors are not observed
  • Overestimation is more severe if auction is thinner
  • Current heuristic approach - Apply a discount factor depending on sell-through rates of different regional markets given the trade-off between revenue and efficiency
  • Future - Improve the estimation of bidders’ valuations
Thank you
Online Social Advertising

• Distinct features of social advertising
  • **Rich user profile** – work experience, industry, skill, interests, education…
  • **More effective targeting** – users are usually required to log in, “I know what you did last night”
Generalized Second Price Auction

• Auction Mechanism
  • Generalized first price (GFP)
  • Vickrey–Clarke–Groves (VCG)
  • Generalized second price (GSP)

• GSP: widely used in industry & less susceptible to gaming
  • Ads are ranked by their quality-weighted bids
  • The price that an advertiser pays for a click is determined by the next highest bid (the minimum necessary to retain its position)
  • If there are fewer advertisers than slots, the last advertiser pays a reserve price $r$
Implementation

Auction log → Model → Member database

Offline: Advertisers → Pinot data store
Online: Member floor
Results from Developed Markets