Measurement and Analysis of OSN Ad Auctions

Yabing Liu†  Chloe Kliman-Silver‡
Robert Bell§  Balachander Krishnamurthy§
Alan Mislove†

†Northeastern University
‡AT&T Labs–Research
§Brown University
Motivation

Online advertising networks are everywhere
Google earned over $50 billion in advertising in 2013

How is online advertising implemented?
Through Auctions!
Advertisers pick keywords, search terms and bid on ads
Advertising networks select the winning bidders and present ads to users
Two ways to pay
  - CPM: Cost Per Mille, the cost of 1,000 ad impressions
  - CPC: Cost Per Click
Motivation

The new OSN-based ad services became popular
Facebook had over $7.8 billion in advertising in 2013
Ask users to fill in their information

Significant data about the users
Personal information (demographics, interests, educational history, relationship status, etc)
Identities of friends
User activity

Target users directly (not keywords, or search terms)
Atlas to serve ads on non-OSN sites
across multiple devices
What has been studied?

Web-search-based advertising networks

"Estimated prices" from Google's Traffic Estimator Tool [Manage.Sci.'11]
Analytical models to predict the clicks, prices, CTR [WWW'14]
New models for conducting online auctions [EC'12]

User Value

Influential users in OSNs [EC'12]
The contribution of users to advertising revenue is skewed [IMC'13]
65% of ad categories received by users are targeting interests [HotNets'13]

Unfortunately

Little academic study of the OSN-based ad networks
OSNs have released little data about their advertising markets
This paper

Goal
- Develop techniques to measure and understand OSN ad markets
- Bring visibility to OSN ad markets, focusing on Facebook
- Research problem meaningful for advertisers, users, and other researchers

Assumption
- No current tool to measure Facebook ad market
- No visibility to Facebook internal system (as external researchers)
Outline

Motivation

Exploring suggested bid mechanism

How are suggested bids calculated?

Exploring user value
Facebook advertising
# Facebook's targeting parameters

<table>
<thead>
<tr>
<th>Basic Fields</th>
<th>Parameters/Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
<td>Country, State, City, Postal code</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>Male, Female, All</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>Range (from 13–65)</td>
</tr>
<tr>
<td><strong>Precise Interest</strong></td>
<td>Travel, Science, Music, ...</td>
</tr>
<tr>
<td><strong>Broad Category</strong></td>
<td>Cooking, Gardening, iPhone 5, ...</td>
</tr>
<tr>
<td><strong>Interested In</strong></td>
<td>Male, Female, All</td>
</tr>
<tr>
<td><strong>Relationship Status</strong></td>
<td>All, Single, In a relationship, Married, Engaged, Not specified</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>English, Spanish, French, ...</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td>Anyone, In high school, In College, College Grad</td>
</tr>
<tr>
<td><strong>Workplaces</strong></td>
<td>Google, Facebook, AT&amp;T, ...</td>
</tr>
</tbody>
</table>

**Notes**
- Target any combination of these parameters
- Required to specify at least one country
Facebook advertising

- Locations: United States, All United States
- Age: 13 - No max
- Gender: All, Men, Women
- Languages: Enter a language...

Potential Reach: 82,000,000 people

Your ad targets people:
- Who live in United States
- Who are male

Suggested Bid: $0.04–$1.23 USD
Methodology

What are suggested bids?
Facebook undocumented feature

"The suggested bid range you see when creating your ads is based on the bids that are currently winning the ad auction for the users you've chosen to target."

How to collect suggested bids in scale?
Programmatically send HTTP GET requests to the Facebook Ad Creation URL

Query:

https://graph.facebook.com/reachestimate?targeting_spec=
{
"countries": ["US"], "age_min": 21, "age_max": 30, genders=[1]
}&currency=USD&account_id=XXX&access_token=XXXX

Response:

{"data": {"users":62984500,"bid_estimations":
[{"location":3,"cpc_min":54,"cpc_median":82,"cpc_max":144,
"cpm_min":3,"cpm_median":14,"cpm_max":83}]}"}
Suggested bid data

Example Dataset

- 1,000 suggested bids
- Each of the 204 countries that Facebook supports
- Queries were roughly spaced 35 milliseconds apart
- U.S.: 159M; New Zealand: 2.2M; Antigua and Barbuda: 29K users

CPM ad prices

- reasoning about CPC requires knowing an advertiser's CTR
- CTR (click-through rate): the fraction of impressions that result in a click
Suggested bid observations

1 Skewed distribution

United States (159,115,060 users)
Suggested bid observations

1 Skewed distribution

2 Significant variance
Suggested bid observations

1 Skewed distribution

2 Significant variance

3 Variance independent of audience size
Suggested bid observations

1 **Skewed distribution**

2 **Significant variance**

3 Variance independent of **audience size**

4 Variance across **accounts**

![Graph showing CPM maximum over time for two accounts, Account1 and Account2, in the United States with 159,115,060 users.](image)
Suggested bid observations

1 Skewed distribution

2 Significant variance

3 Variance independent of audience size

4 Variance across accounts

5 Non-persistence of min or max
Outline

Motivation

Exploring suggested bid mechanism

How are suggested bids calculated?

Exploring user value
Reverse-engineering suggested bids

Goal:
The suggested bid algorithm is a **black box**
Look for the most reasonable explanation

Hypothesis 1: Winning bids change rapidly
Derived from the most-recent-k winning bids for the target users

Not true: significant variance observed on very short timescales
Reverse-engineering suggested bids

Hypothesis 2: Adding random noise

In order to obfuscate the true value

Statistical tests: if the data matched a number of common statistical distributions (Uniform random, Gaussian, Cauchy, Log-Normal or Logistic)

Example probability distribution function of CPM max values for 20,000 suggested bids.

Fails statistical tests.

Not true: poor fit for all distributions, with a p-value of less than $10^{-16}$
Reverse-engineering suggested bids

Hypothesis 3: Sampling winning bids
Sampling from the recent-k winning bids
Reporting the min, median, and max of the sample

Logical mechanism for calculating suggested bids
Consistent with all the five properties of suggested bids

Suggested bid is most likely sampled from the recent-k winning bids
How real auctions affect suggested bids?

Changes to the market

Actively participate in the advertising market
See how quickly we can affect the ad market
Chose a small country (Seychelles, 26K users) with low suggested CPM
Bid a higher CPM ($1.00) than the suggested CPM max ($0.16) from 3 accounts
Ran campaigns concurrently for 8 hours

Changes to the ad market are reflected in the suggested bids.
How suggested bids correlate with revenue?

Comparison with Facebook's revenue

The ground truth: Facebook's SEC filings

Average Revenue Per User (ARPU) at the granularity of regions

Aggregate our CPM median data into the same regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Facebook ARPU</th>
<th>Suggested Bid</th>
</tr>
</thead>
<tbody>
<tr>
<td>US, Canada</td>
<td>$3.50</td>
<td>1.00</td>
</tr>
<tr>
<td>Europe</td>
<td>$1.60</td>
<td>0.45</td>
</tr>
<tr>
<td>Asia</td>
<td>$0.64</td>
<td>0.18</td>
</tr>
<tr>
<td>Rest of World</td>
<td>$0.50</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Notes

Rank the regions in the same order

Europe and Rest of the World regions at approximately the same ratios

The suggested bid data at least correlates with the distribution of Facebook's revenue.
How researchers use suggested bids

Suggested bid data is most likely calculated by sampling from the recent winning bids for the target users.

Multiple samples
- Extract the overall min, median, and max from the collated samples

Convergence
- How many samples to collate together?

We use 25 collated suggested bids
Outline

Motivation

Exploring suggested bid mechanism

How are suggested bids calculated?

Exploring user value
Location

How the location influences the ad auction winning bids?

GDP (Gross Domestic Product) per capita for 204 countries
The output of a country's economy per person

Note
Observed a correlation of 0.37 between the GDP per capita and CPM values

Dramatic differences in ad auction prices across different countries
How is CPM median price correlated with user age?

Select the same three countries (US, NZ, AG)

The smallest age is 13, while targeting age 65 encompasses all users 65 and over

Notes

For U.S. and NZ, as age increases, the CPM median price increase as well
Less clear trend for AG
How is CPM median price correlated with user age?

Select the same three countries (US, NZ, AG)

The smallest age is 13, while targeting age 65 encompasses all users 65 and over.

Notes

For U.S. and NZ, as age increases, the CPM median price increase as well.
Less clear trend for AG

Less differences in ad auction prices across different ages.
Price stability

How **stable** are the prices for different target demographics **over time**?

Select four different sets of targeting parameters

Retrieve 25 suggested bids each hour for a period of 3 weeks (Apr. 3~23, 2013)

**Notes**

G1 shows a periodic increase per week

G2 and G3 shows a multi-day increase starting on 04/16

G4 does not vary much over the study period

**Significant long-term dynamics present in Facebook's ad auctions.**
Summary

Identify the **suggested bid** mechanism to measure Facebook ad market

Validate and show how researchers can **use** the suggested bid data

Analyze how **different users** contribute to Facebook's revenue

- **Dramatic differences** in ad auction prices across different locations, interest
- Significant **long-term** dynamics present in Facebook's ad network
Questions?

Our suggested bid collection code and collected data available at http://osn-ads.ccs.neu.edu