Modeling Consumer Search for Making Online Advertising Decisions

Alan Montgomery
Associate Professor
Carnegie Mellon University
Tepper School of Business

e-mail: alanmontgomery@cmu.edu
web: http://www.andrew.cmu.edu/user/alm3

HKU SPACE, 14 July 2008,
United Centre, Room #610, Admiralty, HK

© 2008 by Alan Montgomery, All rights reserved.

Outline

• Online Advertising
  – Background
  – Search Advertising
• Search Advertising
• Consumer Search
• Conclusions

Online Advertising

Industry Background

The first banner ad in 1994

AT&T paid HotWired to display the above banner ad on October 25, 1994
A Brief History of Online Advertising

Internet Advertising is Growing!

<table>
<thead>
<tr>
<th>Medium</th>
<th>2003E Advertising Spending ($B)</th>
<th>Households (MM)</th>
<th>Ad Spending / Household ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promotions</td>
<td>$97</td>
<td>99</td>
<td>$976</td>
</tr>
<tr>
<td>Newspapers</td>
<td>45</td>
<td>50</td>
<td>886</td>
</tr>
<tr>
<td>Classifieds</td>
<td>15</td>
<td>55</td>
<td>289</td>
</tr>
<tr>
<td>Direct Telephone</td>
<td>65</td>
<td>99</td>
<td>663</td>
</tr>
<tr>
<td>Direct Mail</td>
<td>48</td>
<td>99</td>
<td>483</td>
</tr>
<tr>
<td>Broadcast TV</td>
<td>43</td>
<td>99</td>
<td>429</td>
</tr>
<tr>
<td>Radio</td>
<td>20</td>
<td>60</td>
<td>328</td>
</tr>
<tr>
<td>Cable TV</td>
<td>16</td>
<td>70</td>
<td>226</td>
</tr>
<tr>
<td>Yellow Pages</td>
<td>14</td>
<td>60</td>
<td>164</td>
</tr>
<tr>
<td>Internet / Online</td>
<td>7</td>
<td>60</td>
<td>120</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$374</strong></td>
<td><strong>735</strong></td>
<td><strong>$4,469</strong></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>42</strong></td>
<td><strong>82</strong></td>
<td><strong>487</strong></td>
</tr>
</tbody>
</table>

Source: The Decade in Advertising, DoubleClick (2005), Figure 2

Online Advertising in 2008

<table>
<thead>
<tr>
<th></th>
<th>Online (% of total media)</th>
<th>Online (% of total media)</th>
<th>Online (% of total media)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>$16.9</td>
<td>$281.6</td>
<td>6.0%</td>
</tr>
<tr>
<td>2007</td>
<td>$21.1</td>
<td>$283.9</td>
<td>7.4%</td>
</tr>
<tr>
<td>2008</td>
<td>$25.9</td>
<td>$293.3</td>
<td>8.8%</td>
</tr>
<tr>
<td>2009</td>
<td>$30.0</td>
<td>$299.0</td>
<td>10.0%</td>
</tr>
<tr>
<td>2010</td>
<td>$35.0</td>
<td>$307.0</td>
<td>11.4%</td>
</tr>
<tr>
<td>2011</td>
<td>$41.0</td>
<td>$318.0</td>
<td>13.0%</td>
</tr>
<tr>
<td>2012</td>
<td>$51.0</td>
<td>$332.0</td>
<td>15.4%</td>
</tr>
</tbody>
</table>

Source: eMarketer, March 2008

Dominant Players in Online Advertising

- **Google**
  - The de facto leader: biggest advertising network.
  - Search ads product is Google Adwords.
  - Content ads product is Google Adsense.

- **Yahoo! Search**
  - The number 2 player.
  - Search ads product is Yahoo! Search Marketing (YSM).
  - Content ads product is Yahoo! Publisher Network (YPN).

- **MSN/Live**
  - Up-and-coming player.
  - Search ads product is MSN AdCenter Search.
  - Content ads product is MSN ContentAds.
Search Advertising

How does Search Advertising Work?

- Advertisers (directly or through a SEM firm) buy a set of keywords (the keywords typed in by users of a search engine) by allocating a total budget & a max. bid for each keyword. e.g. Apple buying “Jay Leno” keyword from Google using AdWords.

- When a user types/searches for “Jay Leno” in Google, an automatic auction will take place: the highest bidder will have its ad shown in the result page.

AdWords Pricing

- Google assigns a quality score based on keywords selected by advertiser.
  - Quality score depends on keyword’s CTR, relevance of ad text, historical keyword performance, and other relevancy factors
  - Higher quality score = lower minimum bid and costs.

- The higher the rank number, the higher the position of the ad.
  - Calculated in part by multiplying keyword’s max. CPC times its Quality Score.
  - Every time your ad is

- Advertising price is determined by Auction Market.

- If your maximum bid is less than the minimum rate needed to trigger an ad, you either need to raise your bid or refine your key words.
  - You can set a maximum bid that Google will never exceed.
  - You can set a daily budget. Google will adjust ad frequency based on this budget.
More on Search Advertising

• In practice, Google also adds its own algorithm to ensure that the ad is most relevant to the keyword, and to determine the order/position of the ads (who goes first, second, etc). How?
  - CPM (Cost per Milieu) = cost per thousand impression/click
  - Clickthrough Rate (CTR)
  - Other factors

• Terms to know: Searches/Queries, Clicks, CTR, Cost-per-Click, Conversion Rate, PPC (Pay-per-Click)...

Google’s Adsense allows ads to be placed almost anywhere

• Google: Since we let advertisers put ads on our search results pages, why not create a solution for advertiser to put ads on other websites? Result = AdSense. Advertisers subscribed to AdWords have the option to sign up for AdSense.

• Adsense is not keyword-based. AdSense must “sense” the context of a specific website in order to determine what ads to show. How?
  - Text content, images meta-data, URL, etc etc.

• In AdSense, I get paid a commission everytime visitors to my site clicks on the Google ads on my website.
  - Arbitrage opportunity: clickfraud problem – I can click on the ads on my website and get paid for it.
  - Some solutions: CPA (Cost-per-Action) model
Purchase Conversion Prediction

Keywords to Target

Advertising Decisions

1. Keywords to Target
2. Landing Page
3. Ad Text Copy/Design
4. How much to bid per-click

Conversion Rate

Simple, Naïve Approach

- Enumerate keywords and experiment
  - Google can suggest keywords
  - Check organic keywords
  - Look at competitors search terms
  - Use a search engine optimizer
- Problems:
  - Many keywords have sparse responses, (ex., one click for “Embroidered Harley Jacket with Diamonds”)
  - May miss good keywords (“Lether jacket”) or reject marginal ones too soon (0 out of 10 clicks for one week)
  - Experimentation is Expensive

Generating Keywords

- Exploit the long-tails in keyword search
  - “massage” costs $5 per click
  - “lomilomi massage” costs $.20 per click
  - “traditional hawaiian massage” costs $.05 per click
- Some automated approaches:
  - Google’s Adword Tool looks for past queries that contain the search terms
  - Proximity based methods: query search engine for seed keyword and then append it with words near it. (Cannot generate keywords that do not contain original term.)
  - WordTracker: use meta-tag spidering
  - TermsNet: exploit semantic relationships
  - Wordy: Look for relevant keywords that might be cheaper
**Expert Approach**

- Classify keywords by type and use historical data to estimate conversion models
- Example:
  - “coats” → Generic search
  - “large leather jackets” → Size, Type
  - “Gap leather coat” → Brand, Type
  - “Black Izod Lambskin Leather” → Color, Brand, Type
  - “Wilson Leather.com” → Web address
- Important attributes: Brand, Size, Color, Price, … Looking for attributes that connote specificity

**Modeling our Expert Approach**

- Text Classification Model for Search Strings:
  \[ C: "search string" \text{ for site } s \text{ and visit } i \rightarrow \mathbf{x}_{si} \]
- Hierarchical Logistic Regression to Predict Purchase Conversion:
  \[
  \text{logit}\{\Pr(\text{convert}_{si})\} = \mathbf{\beta}' \mathbf{x}_{si} \\
  \mathbf{\beta}_{s} = \mathbf{Z}_{s} \mathbf{\theta} + \mathbf{u}_{s}
  \]

  Allows us to score arbitrary keywords and predict conversion, Hierarchical nature allows us to borrow information across sites

**Expert Approach Example**

- Suppose a consumer searches for:
  “large Wilsons leather jacket”
- We can identify the following components and they relative contributions to the odds ratio:
  “large” → Size +10% on log odds
  “Wilson” → Brand +50% on log odds
  “leather jacket” → Base Conversion of 1%
- Predicted overall conversion rate: 2%

**Expert Approach Summary**

- Advantages:
  - Robust to many sites within a category
  - Much better forecasts for infrequent words
  - Could be used to predict new sites that are similar
- Disadvantages:
  - Requires expert to create categories
  - Which can be slow and costly
  - Careful attention to how words are categorized
Machine Learning Approach

- Define a language model for search queries. Alternative generate a sample of search results for each string, and use this sample of results to define term-document approach, and use vector-space representation of search.

  "search string"  → \( \{ r_{s1}, r_{s2}, \ldots, r_{sM} \} \)

  \( V: \{ r_{s1}, r_{s2}, \ldots, r_{sM} \} \rightarrow \gamma_{si} \)

- Many choices for our dimension reduction function \( V \), principal component, clustering, LSI, ... We can use the search string representation in this new space in our logistic regression model.

  Problem: Statistical properties of the sampling process

Machine Learning Example

- Extract words frequently associated with "large Wilsons leather jacket" using a search engine (top 100 matches)
  - Words are weighted by frequency

- Compare this phrase with others for similarity. For example "large leather jacket"
  - Vocabulary for "large leather jacket": "leather", "jacket", "clothing", "style", "fashion", "coats", "Men", "women", "classy", "comfortable", "prices", "discounted", "parkas", "luggage", ...

- Find that "large Wilsons leather jacket" is more similar to "Wilsons leather jacket" than "big red leather jacket"

Machine Learning Approach Summary

- Advantages:
  - Automated and does not require human expert to categorize
  - Can be replicated over large number of sites
  - Yields similar accuracy to expert approach

- Disadvantages:
  - Training is computational intensive and more complex
  - Statistical properties of query expansions not well understood
  - Does not take advantage of natural language structure

Empirical Application
**Keyword Conversion**

- ~15,000 purchases
- ~3,000 keywords
- ~5 purchases per keyword, but highly skewed as well
- 2,175 (71%) keywords don't generate any sales
- The top 10 keywords attract 10,292 (72%) sales

**Predictions**

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Conversion</td>
<td>Conversion</td>
</tr>
<tr>
<td>Experts</td>
<td>27430</td>
<td>10693</td>
</tr>
<tr>
<td>ML Approach</td>
<td>26348</td>
<td>11775</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Conversion</td>
<td>Conversion</td>
</tr>
<tr>
<td>Experts</td>
<td>539</td>
<td>2298</td>
</tr>
<tr>
<td>ML Approach</td>
<td>505</td>
<td>2332</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Overall Accuracy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>ML Approach</td>
<td>66%</td>
<td>34%</td>
</tr>
</tbody>
</table>

**Overall Accuracy**

- 73%

**Online Advertising**

- Search advertising has arisen as a dominate form of online advertising
- Presented an approach for predicting purchase conversion before any keyword bid is made by using experience with past keyword bids
- Next steps:
  - Need to generalize to all facets of search engine marketing (landing pages, advertising text)
  - Relate to previous searches, understand general to specific searching