A Survey of Sponsored Search Advertising in Large Commercial Search Engines

George Trimponias, CSE
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Introduction

• Commercial search engines act as information gateways for millions of Internet users.
• In response to a user’s query, search engines generate a ranked list of results.
• These results have a critical power in shaping the Web users’ actions.
• This unique role as an intermediary between Internet users and the vast web content has created a new market for sponsored search advertising.
Market for Sponsored Search

- Advertisers pay a fee to be displayed alongside the web search results for keyword queries that they have expressed interest in.
- When a user enters a query, a small number of paid (sponsored) links appears on top or to the right side of the unpaid (organic) results.
- The search engine relies on the market to determine the advertisers that win ad slots and the prices by using auctions among the advertisers.
- Payments are based on the pay per click model.
Top organic and sponsored results for *melanoma treatment* using Google

- **Melanoma Treatment (PDQ®) - National Cancer Institute**
  - www.cancer.gov/cancertopics/pdq/treatment/melanoma/.../page1 - Cached
  - 1 Doc 2010 – Treatment Options for Recurrent Melanoma ... (See the PDQ summary on Intracocular (Eye) Melanoma Treatment for more information.) ... Stages of Melanoma - Skin Cancer Treatment (PDQ®) - Treatment Option Overview

- **Treatment of melanoma by stage**
  - www.cancer.org/.../melanoma-skin-cancer-treating-by-... - Cached
  - 17 Aug 2011 – The type of treatment(s) your doctor recommends will depend on the stage and location of the melanoma and on your overall health.

- **Melanoma (Skin Cancer) Signs, Symptoms, Treatment, Prognosis**
  - www.medicinenet.com/.../melanoma/index - Cached
  - 1 day ago – Get information on melanoma signs, diagnosis, staging, prognosis and symptoms. Learn about early detection, treatment side effects, ...

- **Health | Melanoma treatment expands with two new drugs | Seattle ...**
  - seattlehealthsnews.com/html/health/20160515_melanoma.html - Cached
  - 2 Aug 2011 – In March, the FDA approved another late-stage melanoma treatment, Bristol-Myers Squibb's Yervoy, which acts to trigger the body’s anticancer ...

- **Melanoma — Treatment Options at Mayo Clinic**
  - www.mayoclinic.org/melanoma/ - Cached
  - Melanoma Learn about melanoma treatment options at Mayo Clinic.
A thriving multibillion dollar industry

• The three most prominent players are Google’s AdWords, Yahoo! Search Marketing, and MSN AdCenter.
• Google’s total revenue in fiscal 2010 was over $29 billion; over 96% of it was related to advertising.
• Google currently boasts a market capitalization of over $170 billion; General Motors’ is less than $32 billion.
• In 2009 only, Google generated a total of $54 billion of economic activity for American businesses.
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Advertisers

• They declare to the search engine a set of keywords that relevant users may search for.
• For each keyword, they determine a maximum bid, or, maximum cost per click (maximum CPC).
• The maximum CPC corresponds to their maximum willingness to pay for a click.
  – The actual CPC may be lower.
• Daily budgets.
• Brand awareness, conversions.
The Search Engine

• Enables interaction between advertisers and users.
• Ongoing conflict between revenue maximization and high-quality sponsored results.
• The most prominent measure of ad quality is the clickthrough rate (CTR).
• Large search engines utilize a quality score to measure how relevant the keyword is to the ad text and to the user’s query.
• Reserve prices.
Users

• Users constitute through their actions the commodity that advertisers bid on.
• By posing textual queries, they express to the search engine an intention.
• The advertisers declare keywords to the search engine based on that intention.
  – Often, without success.
• Even today, there is a very limited understanding of user behavior and interaction with the search engine.
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Ranking and Pricing Schemes

• Ads are sorted in decreasing order of their rank.
  – Their rank depends on both the maximum CPC, and the ad quality score.
  – Google: Ad Rank = Maximum CPC × Quality Score.
• The ad with the highest rank gains the first position, and so on down the page.
• Pricing is based on a generalized second-price auction (GSP).
• Under GSP, the advertiser is charged the minimum amount required to maintain their ad’s position (plus a tiny increment).
  – Google: The price for advertiser $i$ is determined by advertiser $i+1$ below them, and is equal to $(\text{Maximum CPC})_{i+1} \times (\text{Quality Score})_{i+1}/(\text{Quality Score})_{i}$. 
Generalized-Second-Price Auction

**Step 1:**
Order all advertisers by $QS_i \times b_i$

- 1
- 2
- $n - 1$
- $n$

**Step 2:**
Advertiser $i (i \leq k)$ gets slot $i$ for a price $p_i$

- 1
- 2
- $k - 1$
- $k$
- $k + 1$
Pricing per click

Pay the *minimum* amount of money needed to attain your position!

\[ p_1 = \frac{QS_2 \times b_2}{QS_1} \]

\[ QS_i \times p_i = QS_{i+1} \times b_{i+1} \]

\[ p_k = \frac{QS_{k+1} \times b_{k+1}}{QS_k} \]
Account Structure

- An advertiser may have one or more accounts.
- An account contains one or more campaigns.
- A campaign is created to achieve a clear marketing goal, and is made up of one or more ad groups.
- An ad group contains a set of ads and a keyword list that triggers these ads to show.
- Budgets are applied at the campaign level, whereas bids are applied at the ad group level.
# Account Structure

<table>
<thead>
<tr>
<th>Account Name:</th>
<th>Electronics-Planet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campaign Name:</td>
<td>Monthly Sales Promotion</td>
</tr>
<tr>
<td>Ad Group Name:</td>
<td>Laptop Computers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Ads</th>
<th>Destination URLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>pc laptops</td>
<td>Electronics-Planet Has Laptop Computers</td>
<td><a href="http://www.electronics-planet.com">http://www.electronics-planet.com</a></td>
</tr>
<tr>
<td>notebook computers</td>
<td>Huge selection of popular laptop computers. Save up to 10%.</td>
<td></td>
</tr>
<tr>
<td>laptop</td>
<td>Low Prices on Laptop Computers</td>
<td><a href="http://www.electronics-planet.com/ship_4free.html">http://www.electronics-planet.com/ship_4free.html</a></td>
</tr>
<tr>
<td>computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>laptop accessories</td>
<td>Free shipping on popular laptop computers at electronics-planet.com</td>
<td></td>
</tr>
</tbody>
</table>
Ad Structure

• Headline.
• Lines of text.
• Display URL.
• Destination URL (Landing Page).

Electronics-Planet Has Notebook Computers
Huge selection of popular notebook computers. Save up to 10%.
http://www.electronics-planet.com
Click Probabilities

• Search engines compute the CTR of an ad without taking into account the ad position.
• Empirically, ads at the top receive more clicks.
• Moreover, ad externalities play a significant role.
• Various models that incorporate positional information have been proposed.
  – Separability Assumption: \( \text{CTR} = \text{Ad Specific CTR} \times \text{Position Specific Visibility Factor} \).
  – Cascading: Ad Specific CTR, Ad Continuation Probability.
Payment Model

• Advertisers and the search engine have conflicting views on when a payment should occur.
  • Search engine: pay per impression.
  • Advertisers: pay per conversion.
  • The pay per click model reconciles the two opposing viewpoints.
Click Fraud

• Type of Internet crime.
• A sponsored link is intentionally clicked with no intention of generating value.
• Can be done automatically by computer scripts or directly by humans.
• Possible reasons:
  – A competitor may desire to minimize the impact of an ad campaign.
  – The search engine may want to increase its income.
• Very difficult to identify.
Click Fraud

• Common tools to deal with this problem:
  – Aggressive monitoring
  – Improved automated filters that use sophisticated data mining technology.
  – Pay per conversion.
Keyword Match

• Advertisers have to come up with the right sets of keywords.

• Users searching for the keyword may use:
  – Singular or plural, synonyms and other variations, misspellings, extensions, word reorderings.

• Search engines alleviate this problem by providing 4 matchtypes.

• *Exact matchtype*: the user searches for the specific keyword in this order, and without any other terms in the search query.
Keyword Match

• **Phrase matchtype**: the search query contains the keyword, with the terms in that order, but can additionally contain other terms.

• **Broad matchtype**: a user’s query contains all terms in the keyword in any order, possibly along with other terms; or, the afore listed variations.

• **Negative matchtype**: the ad is not shown upon occurrence of certain query terms.
Bidding Expressivity

• How to best translate advertiser needs into an appropriate bidding language.
• Commercial search engines allow the advertisers to fine-tune their ads by
  – Specific locations.
  – Days of the week.
  – Time of the day.
  – Demographic (gender and age) groups.
  – Languages.
• Inherent tradeoff between bidding expressivity and computational cost of underlying auction.
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Early Internet Advertising

• Began in 1994.
• Pay per impression basis.
• Advertisers paid flat fees for a thousand showings.
• Contracts on a case-by-case basis.
• High minimum contracts, slow entry.
Generalized First-Price Auction

- Introduced in 1997 by Overture.
- Every advertiser submits a maximum cost per click bid for a keyword.
- Advertisers are ranked in decreasing order of their bids.
- Whenever their ad is clicked, the advertiser is charged an amount equal to their latest bid.
- Generalization of first-price auction for single items.
Generalized First-Price Auction

• Advertisers better target their ads.
• Ease of use, very low entry costs, mechanism transparency.
• On the other hand, this auction induces a highly unstable dynamics
  – Bidders do not state their true valuations (non-truthful).
  – Bidders keep changing their bids in response to other bidders’ behavior.
  – Unstable cycling (“sawtooth”) patterns emerge.
  – Efficiency is compromised.
  – Volatile prices are translated into revenue losses.
Generalized First-Price Auction
Generalized Second-Price Auction

• Introduced by Google in February, 2002.
• Advertisers now pay the minimum possible price to attain their position.
• Price is independent of advertiser’s bid; depends on the next-highest bidder only.
• New auction structure is much less susceptible to gaming.
• Yahoo!/Overture also switched to GSP.
Generalized Second-Price Auction

• Generalization of the single-item second-price auction, also called Vickrey auction.
• Vickrey auction is truthful; GSP is not.
• The non-truthfulness property cannot guarantee stable system dynamics.
• How to generalize the second-price auction in way that achieves truthfulness and efficiency?
The VCG Auction

• Named after Vickrey, Clarke, and Groves.
• Ranking: bidders are ranked in decreasing order of their bids.
• Payment: every bidder pays the *externality* that he imposes on the others, i.e., the decrease in the valuations of other bidders because of his presence.
• Truthful and socially optimal.
Market Development and Assessment

• Consolidation and Convergence.
• GSP rather than VCG is employed in practice, possibly because:
  – VCG is hard to communicate to advertisers.
  – VCG may entail substantial transition costs.
  – The revenue consequences of switching to VCG are largely unpredictable.
  – Ad quality scores diminish the potential for gaming the system, since advertisers have no control over the ad quality scores.
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Matching Markets

- A market refers to a variety of systems, institutions, procedures, social relations, and infrastructures, whereby parties engage in an exchange.
- Two-sided markets with a set of agents, and a set of buyers.
- Every buyer has valuations for each of the items.
- Assume each valuation is a nonnegative whole number.
- How to determine prices?
Matching Markets

• Assume every seller puts their item up for sale for some price.

• Given a set of prices, a buyer will buy from the seller that maximizes his payoff, i.e., valuation – price.
  – These seller(s) are called the preferred sellers of the buyer.

• Market-clearing prices: each item can be assigned to a different buyer.
  – Perfect matching in the bipartite graph of sellers and buyers.
# Matching Markets

<table>
<thead>
<tr>
<th>Prices</th>
<th>Sellers</th>
<th>Buyers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>a</td>
<td>x</td>
<td>10, 3, 2</td>
</tr>
<tr>
<td>1</td>
<td>b</td>
<td>y</td>
<td>6, 6, 6</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>5, 4, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Valuations</th>
</tr>
</thead>
<tbody>
<tr>
<td>a: 12, 4, 2</td>
</tr>
<tr>
<td>b: 8, 7, 6</td>
</tr>
<tr>
<td>c: 7, 5, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prices</th>
<th>Sellers</th>
<th>Buyers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>a</td>
<td>x</td>
<td>7, 2, 2</td>
</tr>
<tr>
<td>2</td>
<td>b</td>
<td>y</td>
<td>3, 5, 6</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>2, 3, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a: 9, 3, 2</td>
</tr>
<tr>
<td>b: 5, 6, 6</td>
</tr>
<tr>
<td>c: 4, 4, 2</td>
</tr>
</tbody>
</table>

Diagram:

- At price 2, a matches x.
- At price 1, b matches y.
- At price 0, c matches z.
Matching Markets

• For any set of buyer valuations, there always exists a set of market-clearing prices.

• For any set of market-clearing prices, a perfect matching in the resulting preferred-seller graph has the maximum total valuation of any assignment of sellers to buyers.

• Proved by economists Damange, Gale, and Sotomayor in 1986.

• The price construction equivalent to the Hungarian method by Egervary in 1916.
Sponsored Search as a Matching Market

• The ad slots are the items, and the advertisers are the potential buyers.

• The valuation of an advertiser for a certain slot depends on:
  – his revenue per click, and
  – the clickthrough rate of the slot.
VCG Prices and the Market-Clearing Property

• Recall the two important properties of the VCG mechanism:
  – VCG is socially optimal, i.e., it maximizes the total valuation of buyers.
  – Truthful reporting in VCG is a dominant strategy.

• What is the relationship of VCG prices and market-clearing prices?
VCG Prices and the Market-Clearing Property

• First, note that VCG are *personalized* prices, whereas market-clearing prices are *posted* prices.

• However, deep connections between the two models exist:
  – VCG prices are market-clearing.
  – VCG prices form the unique set of market-clearing prices of minimum total sum.

• Proved by Leonard, and Damange.
Properties of the GSP Protocol

- The GSP procedure suffers from several pathologies.

<table>
<thead>
<tr>
<th>Clickthrough rate</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Revenues per click</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Valuations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>a</td>
<td>x</td>
<td>7</td>
<td>a</td>
<td>x</td>
<td>70, 28, 0</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>y</td>
<td>6</td>
<td>b</td>
<td>y</td>
<td>60, 24, 0</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>1</td>
<td>c</td>
<td>z</td>
<td>10, 4, 0</td>
</tr>
</tbody>
</table>
Properties of the GSP Protocol

- Truth-telling may not constitute a Nash equilibrium.

<table>
<thead>
<tr>
<th>Prices</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>a</td>
<td>x</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>y</td>
<td>20</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>0</td>
</tr>
</tbody>
</table>

- Bid Vector: (7,6,1)

<table>
<thead>
<tr>
<th>Prices</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>a</td>
<td>x</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>y</td>
<td>10</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>0</td>
</tr>
</tbody>
</table>

- Bid Vector: (5,6,1)
Properties of the GSP Protocol

- GSP may accept multiple Nash equilibria, among which some may be socially suboptimal.

<table>
<thead>
<tr>
<th>Prices</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>a</td>
<td>x</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>b</td>
<td>y</td>
<td>16</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prices</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>a</td>
<td>x</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>y</td>
<td>30</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>0</td>
</tr>
</tbody>
</table>

Bid Vector: (5,4,2)  
Bid Vector: (3,5,1)
Properties of the GSP Protocol

- VCG prices in this example.

<table>
<thead>
<tr>
<th>Prices</th>
<th>Slots</th>
<th>Advertisers</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>a</td>
<td>x</td>
<td>30, 24, 0</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>y</td>
<td>20, 20, 0</td>
</tr>
<tr>
<td>0</td>
<td>c</td>
<td>z</td>
<td>-30, 0, 0</td>
</tr>
</tbody>
</table>
Properties of the GSP Protocol

• Search engine revenue depends on which equilibrium is actually played in the game.

• Indeed, in our example:
  – In the former Nash equilibrium, revenue is 48.
  – In the latter Nash equilibrium, revenue is 34.
  – VCG revenue is 44.
GSP and Matching Markets

• There always exists a set of bids that forms a socially optimal Nash equilibrium for the GSP procedure.
• We can construct such a set of bids from any set of market-clearing prices in the corresponding matching market.
• The derived equilibria are called symmetric Nash equilibria.
  – Stronger than Nash equilibria.
  – No only does a bidder not want to change his bid, but does not want to get another slot at its current price.
Outline

• Introduction
• Structure of Sponsored Search Advertising (SSA)
• Practical Issues in SSA
• Historical Evolution of SSA
• Marching Markets, VCG, and the GSP Procedure
• Conclusion
Challenges

• Location-based sponsored search
  – Large search engines treat mobile users in exactly the same way as other users.
  – But mobile users are more valuable advertising targets.
• Explosion of social networking sites and social media forums
  – Personalized ads based on user profiles (privacy concerns).
  – A key role in influencing the level of buzz about firms, brands, and products in the online world.
Open Problems

• Parameter Estimation
  – Through logs of search and ad traffic over significant periods of time.
  – Very hard to model (think, e.g., of human physiology, psychology).
  – Exploitation vs. Exploration Tradeoff.

• Incomplete Knowledge.

• Sponsored Search Auction Design
  – GSP suffers from numerous pathologies.
  – VCG may not be sufficiently expressive.
  – Grand simulation platforms and grand simulation models could help in this direction.
Questions?

Thank you!