Web Mining
Lecture 15: Introduction to Computational Advertising (Part 1)

Manish Gupta

20th Sep 2013

Slides borrowed (and modified) from
http://pierre.senellart.com/enseignement/2009-2010/inf396/4-Recommandation/CompAdv.pdf
http://www.stanford.edu/class/msande239/

Recap of Lecture 14: Analysis of Microblogs (Part 2): Location Prediction

- Location Prediction using Tweet Content
- Location Prediction using Social Ties
- Applications of Location Prediction

Announcements

- Schedule change
  - 28th Sep class moved to 30th Sep 6-7:30pm

Today's Agenda

- Introduction to Computational Advertising
- Display Ads
- Textual Ads
- Auctions

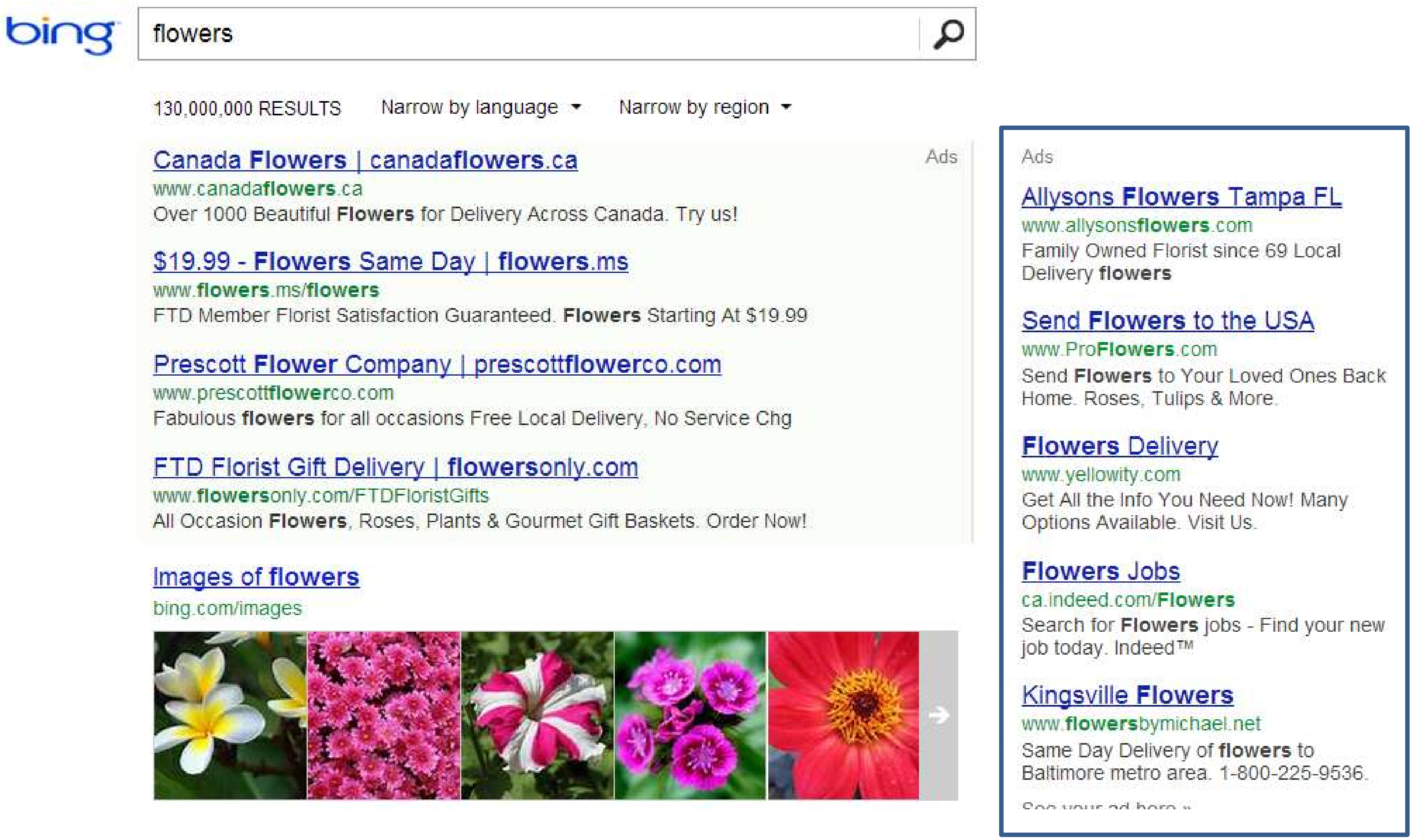
Today's Agenda

- **Introduction to Computational Advertising**
- Display Ads
- Textual Ads
- Auctions

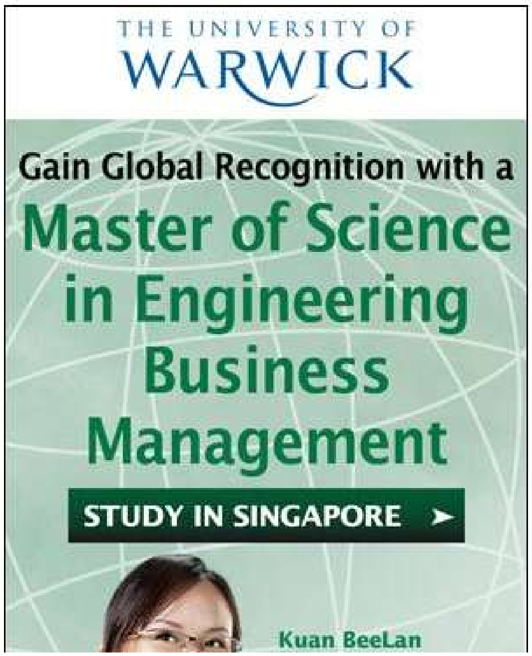
Introduction

- Why is Computational Advertising Important?
  - Multi-billion dollar business
  - Supplies most of the revenue for search engines
- Why does it work?
  - massive scale, automated
  - “learn from the data”
- Ad types
  - Sponsored search which serves ads in response to search queries
  - Content match which places ads on third-party pages
  - Display advertising or banner ads or graphical ads

Sponsored Search Ad



Context Match Ad



FIIs infuse Rs 6,000cr in Indian capital market in two weeks

Tweet

PTI | Sep 15, 2013, 08.56PM IST

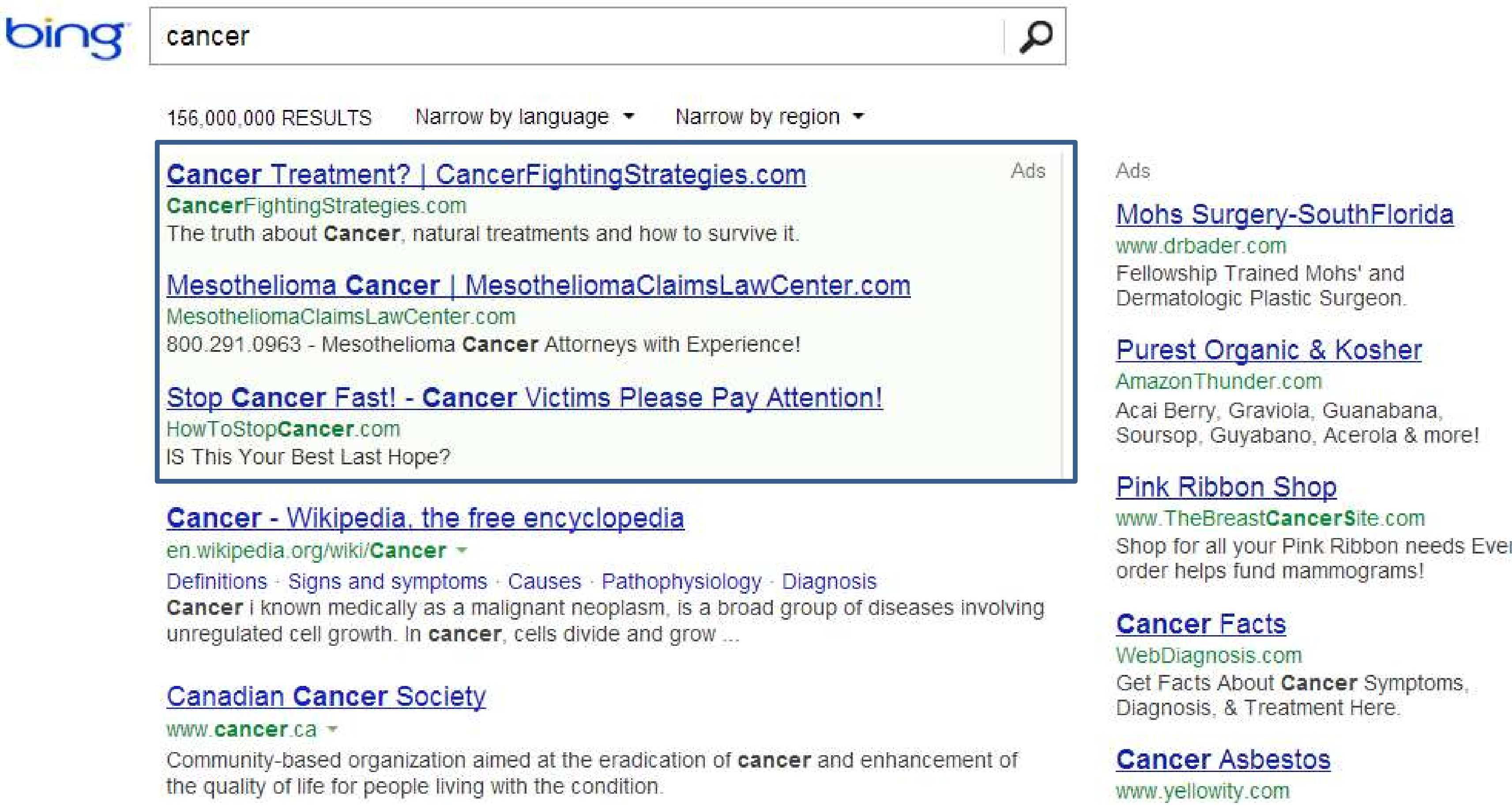
NEW DELHI: Overseas investors have pumped in nearly Rs 6,000 crore in the Indian capital markets in a fortnight ended September 13, mainly on the back of new RBI governor Raghuram Rajan announcing various measures to boost the depreciating local currency and revive economic growth.

Inflows in equities were about Rs 6,372 crore ($966 million) during September 2-13, while there was a pull-out of Rs 382 crore ($64 million) from the debt market, translating into net inflows of Rs 5,990 crore ($922 million), as per latest data available with market regulator Sebi.

In August there was a net withdrawal of nearly Rs 16,000 crore (about $2.5 billion) from the domestic capital markets.

Ads by Google

Display Ad (Banner Ad)

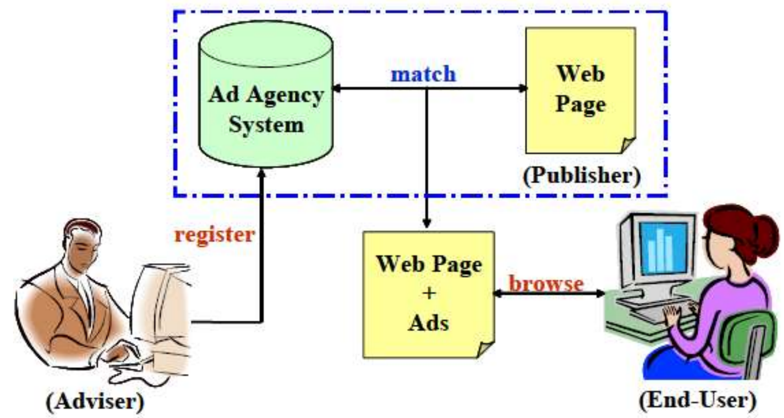


Comparison of Computational Ads with Traditional Ads

- Classical
  - Relatively few venues –magazines, billboards, newspapers, handbills, TV, etc
  - High cost per venue ($3Mil for a Super Bowl TV ad)
  - No personalization possible
  - Targeting by the wisdom of ad-people
  - Hard to measure ROI
- Computational –almost the exact opposite
  - Billions of opportunities
  - Billions of creatives
  - Totally personalizable
  - Tiny cost per opportunity
  - Much more quantifiable

Contextual Ad Basics

- Four entities
  - The publisher is the owner of Web pages on which advertising is displayed.
  - The advertiser provides the supply of ads.
  - The ad network (exchange) is a mediator between the advertiser and the publisher, who selects the ads that are put on the pages.
  - End-users visit the Web pages of the publisher and interact with the ads.



- Revenue models
  - CPM: Cost Per iMpression
  - CPC: Cost Per Click
  - CPV: Cost Per Visitor
  - CPA/CPT: Cost Per Action/Transaction

Problems

- Manual or automated review process to ensure that advertiser content is in fact relevant to the target keyword
- Matching advertiser content to user queries as they are received
- Displaying advertiser content in some rank order
- Gathering data, measuring clicks, charging advertisers based on consumer clicks, etc.

Revenue Models

- Under CPM: Revenue = N * CPM
- Under CPC: Revenue = N * CTR * CPC
  - CPC depends on auction mechanism
- Under CPA: Revenue = N * CTR * Conv. Rate * CPA
- Revenue dependence
  - CPM: website traffic
  - CPC: + ad relevance
  - CPA: + landing page quality
- From 1st to 3rd: more relevant for advertisers, bigger prices and bids!

Today's Agenda

- Introduction to Computational Advertising
- **Display Ads**
- Textual Ads
- Auctions

Display Ads/Banner Ads/Graphical Ads

- Just pick ads
  - graphically displayed
  - mostly for brand awareness
  - revenue model is typically CPM
- Ads are targeted at particular demographics
  - GM ads on Yahoo autos shown to "males above 55"
  - Mortgage ad shown to "everybody Yahoo Front page"
- Book a slot well in advance
  - "2M impressions in Jan next year"
  - impressions guaranteed by the ad network!

Display Ads Problems

- Two types of online graphical advertising
  - Guaranteed delivery (GD)
  - Performance graphical advertising (non-guaranteed delivery, NGD)
- Guaranteed delivery
  - Contract booked based on targeting attributes of an impression: age, income, location,…
  - Each contract has a duration and a desired number of impressions
  - Issues in GD
    - Contract pricing
    - Traffic forecasting
    - Optimal impression allocation to the active contracts
    - Demographics overlap
    - How much will advertisers want each demographic

Performance of Display Ads

- Graphical ads can also be placed based on performance –CPM/CPC/CPA
- Optimization Problem Definition = Max CTR
- Matching approaches
  - Reactive: explore the placement of a particular ad on different pages; for each page observe achieved CTR; once the CTRs are learned, given page, pick the ad with highest observed CTR
  - Predictive: generate features for the ad using related ads (same advertiser), landing page, or advertiser metadata – predict performance based on page and ad features
  - Hybrid: (1) and (2) are complementary and can be combined

Today's Agenda

- Introduction to Computational Advertising
- Display Ads
- **Textual Ads**
- Auctions

Textual Ads: Content Match and Sponsored Search Ads

- Content match
  - Pick ads by matching them to content
  - The user intent is unclear
  - The user intent is unclear
- Sponsored search
  - Given a search query
  - Pick ads by matching them to the query
  - User declares her intention
  - Query is short and less noisy than Content Match

Anatomy of a Sponsored Search Ad

Bid phrase: sigir 2010
Bid: $1

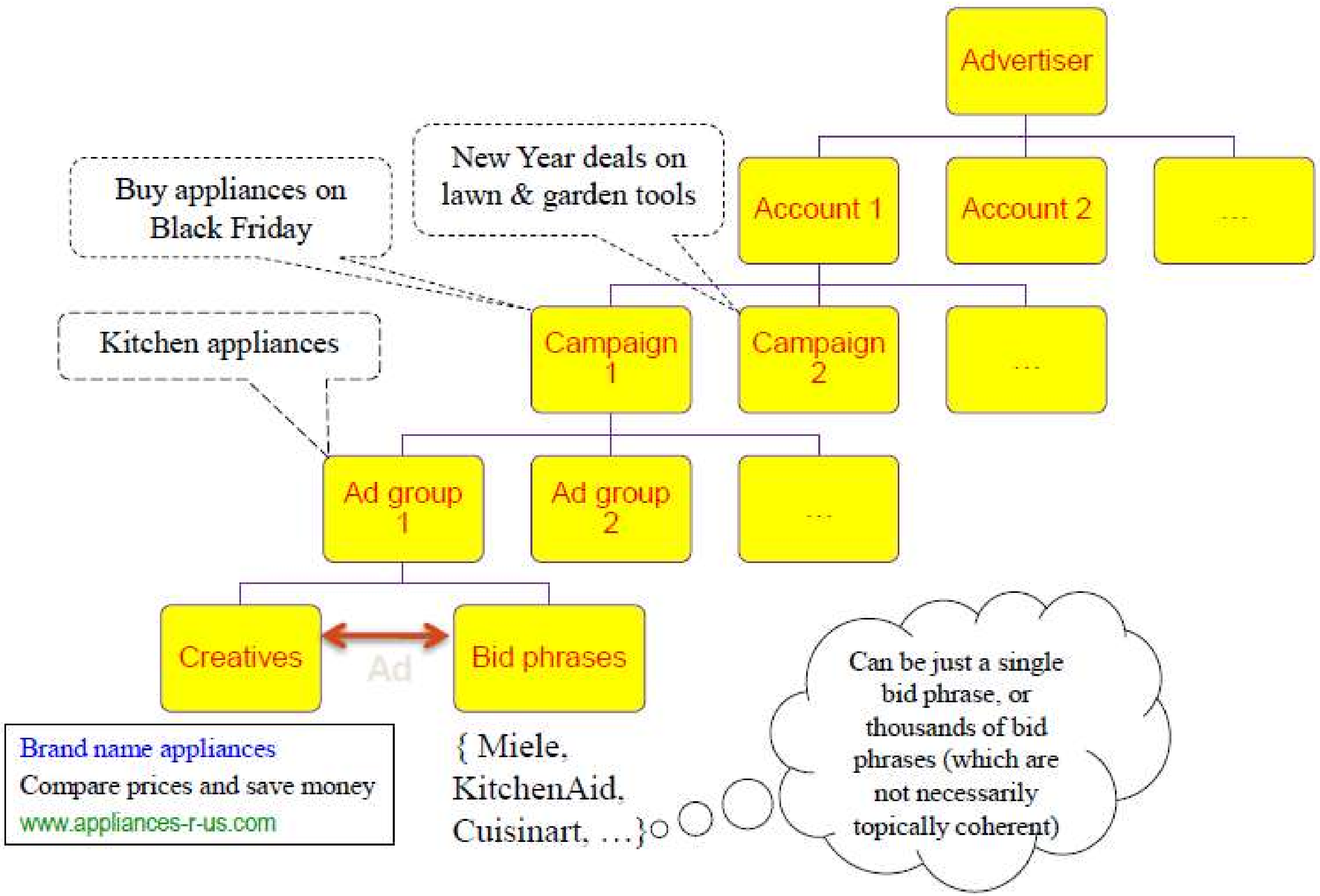
Title — Tutorial at **SIGIR 2010**

Creative — Information Retrieval Challenges in Computational Advertising

Display URL — research.yahoo.com/tutorials/**sigir**

Landing URL: http://research.yahoo.com/tutorials/sigir10_compadv/

Textual Ad Schema



Main Issues

- Given a query
  - Select the top-k ads to be shown on the k slots in order to maximize total expected revenue
- What affects the total revenue
  - Relevance of the ad to the query
  - Bids on the ads
  - User experience on the landing page (ad quality)

Selecting an Ad

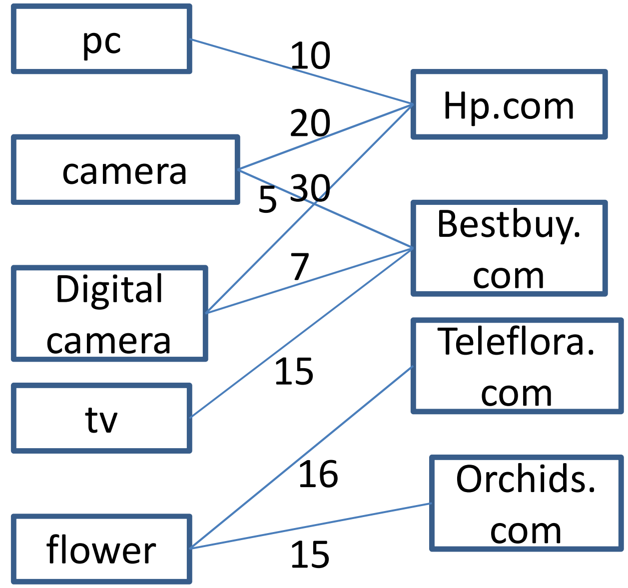
- Each participant has its own utility
  - Advertisers wants ROI and volume
  - User wants relevance
  - Publisher wants revenue per impressions/search
  - Ad network wants revenue and growth
- Ad selection: optimize for a goal that balances the utilities of the four participants
- IR based Ad Relevance Computation
  - Use a search engine to match ads to context
    - Ads are the “documents”
    - Context (user query or webpage content) are the query
  - Problem: word matches might not always work
  - Need to extract topical information
- Bid value for ad (usually second price auction is used)
- Machine learning from clicks
  - Estimate CTR=Pr(click|ad, query, user)
  - Ad-Ad similarity & collaborative filtering

Ad Selection Approaches

- Exact match
  - The ad's bid phrase matches the query
  - Need query normalization
  - Cannot bid on all feasible queries
- Broad match: translate the query into bid phrases
  - The ad platform finds good ads for a given query (the advertiser did not bid on that specific keyword, but the query is deemed of interest to the advertiser)
  - Pricing can be misleading
  - Significant portion of the traffic has no bids L

Query Rewriting

- Rewrite the user query q into $Q' = (q_1, q_2, L)$
- Use exact match to select ads for Q'
- Offline vs online
  - Offline can be done only for queries that repeat often
  - Online
    - For rare queries offline not practical or simply does not work
    - Lot less time to do analysis (a few ms)
- Using Search logs (frequent rewrites from query logs)
  - insertions: game codes -> video game codes
  - substitutions: john wayne bust -> john wayne statue
  - deletions: skateboarding pics -> skateboarding
  - spell correction: real eastate -> real estate
  - specialization: jobs -> marine employment
- Using Clicks
  - SimRank on bipartite graph of queries and ads
  - Edge weights could be #clicks for (ad, query) pair or CTR
  - Iterative computation
    - “Two queries are similar if they are connected to similar ads”
    - “Two ads are similar if they are connected to similar queries”



Similar Queries
Camera – Digital Camera
pc – camera
pc – digital camera
tv – camera
tv – digital camera
pc-tv

Ad Relevance by Online Learning

- Offline (batched) learning
  - Learned from historical data
  - Slow response to emerging patterns
  - Initial biases never corrected
    - if the system never showed “golf classes” for “iPod” it can never learn if this matching is good.
- Online Learning
  - combine exploitation with exploration
  - pick ads that are good according to current model
  - pick ads that increase your knowledge about the entire space of ads

Online Content Matching

- Web advertising for two types of Web pages
  - Static page (Offline): Matching of ads can be based on prior analysis of their entire content
    - Works well for static content pages that are displayed repeatedly
  - Dynamic page (Online): Ads need to be matched to the page while it is being served to the end-user. Thus, limiting the amount of time allotted for its content analysis.
- When a user views a page, the ad selection engine has only a couple hundred milliseconds to provide the ads.

Collaborative Filtering Connection

- Traditional IR based on fixed query-result relevance
- Ads: Rank by CTR probability
  - Continuous CTR feedback for each (query, ad) pair
  - Learn the "best match between a user in a given context and a suitable advertisement"
- Data is sparse, in order to get the best match, we need to find similar ads, pages, and users
- Make use of dyadic interaction systems (recommendation systems)
  - Note dyad is a pair: (user, movie), (user, ad), etc.
  - Predict response to unknown dyads using collaborative fitering

Sponsored Search (Big Picture)

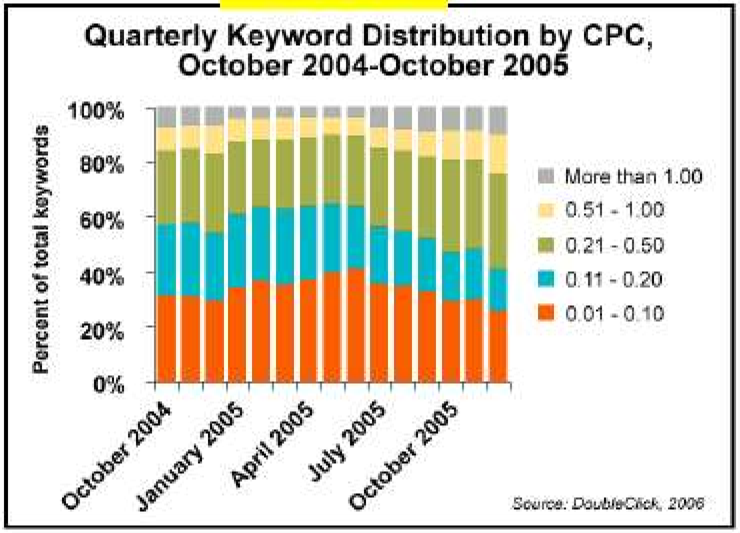
- Ads corpus = Bid phrases + Title + URL + landing page
- Ad query = Search keywords + context (location, user profile, search history)
  - Sponsored Search: Context = Web search results
  - Content match, banners: Context=Publisher page
- Ad search is similar to web search but with these differences
  - Ad database is smaller
  - Ad database entries are small
  - Ranking depends also on bids and CTRs
  - The query (current page) can be much larger than the target document
- Finding the best textual ad is an information retrieval problem with multiple, possible contradictory utility functions

Interactions in Sponsored Search

- Advertisers
  - Submit ads associated to certain bid phrases
  - Bid for position
  - Pay CPC
- Users
  - Make queries to search engine, expressing some intent
- Search engine
  - Executes query against web corpus + other data sources
  - Executes query against the ad corpus
  - Displays a Search Results Page (SERP) = integration of web results, other data, and ads

CPC Costs

2005



2009

U.S. Average Search CPC by Category, September and October

Category	CPC Sep ($)	CPC Oct ($)	Change (%)
Automotive	0.47	0.50	6.3
Finance	1.80	1.63	-9.4
Retail	0.40	0.43	7.5
Travel	0.55	0.54	-1.8

Source: efficient frontier 2009 (Via ClikZ.com)

2011

Average Search Cost per Click (CPC) in the US, by Industry, March-April 2010 & March-April 2011

	March 2010	April 2010	March 2011	April 2011	% change (April vs. March 2011)	% change (April 2011 vs. April 2010)
Finance	$1.45	$1.61	$2.03	$2.26	11.3%	40.4%
Automotive	$0.55	$0.52	$0.53	$0.54	1.9%	3.8%
Retail	$0.42	$0.43	$0.43	$0.45	4.7%	4.7%

Source: Efficient Frontier as cited in company blog, May 20, 2010

128040 www.eMarketer.com

Budget and Other Factors

- Advertisers can specify budgets
- Budgets can be implemented as follows
  - Spend it quickly till out of money
  - Spend it slowly till end-of-day
  - Spend it as the search engine sees fit
  - Spend it on a certain demography of users only
- There are sometimes "reserve prices" = minimum cost to be shown on a given kw (depends on kw)
- There are sometimes "minimum bids" = minimum bid required to participate in action (could depend on advertiser and keyword)

Three Problems for a Search Engine

- Ad retrieval
  - Match to query/context
- Ordering the ads
- Pricing on a click-through

Today’s Agenda

- Introduction to Computational Advertising
- Display Ads
- Textual Ads
- **Auctions**

Short Introduction to Game Theory

- Set of players.
- A set of strategies available to those players (each has its own set)
- A specification of payoffs for each player for each combination of strategies.
- Each player's payoff depends on the strategy chosen by every other player!
- Dominant strategy
  - Strategy = a complete definition of how a player will play a game.
  - Strategy X (for a player) dominates another strategy Y if for all choices by other player(s), X yields at least as much payoff as Y.
  - Strategy X is dominant if it dominates all other strategies.

Nash Equilibrium

- Nash equilibrium= choice of strategies in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing his own strategy unilaterally.
- Pure strategy= deterministic definition of how a player will play a game
- Mixed strategy = an assignment of probabilities to each pure strategy --the players throw coins to pick the strategy they follow
- A game could have many Nash equilibria or none, if players must follow pure strategies.
- Nash theorem: In every n-player game in which every player has finitely many pure strategies there exists a set of mixed strategies that forms a Nash equilibrium.

Game Theory for Ads

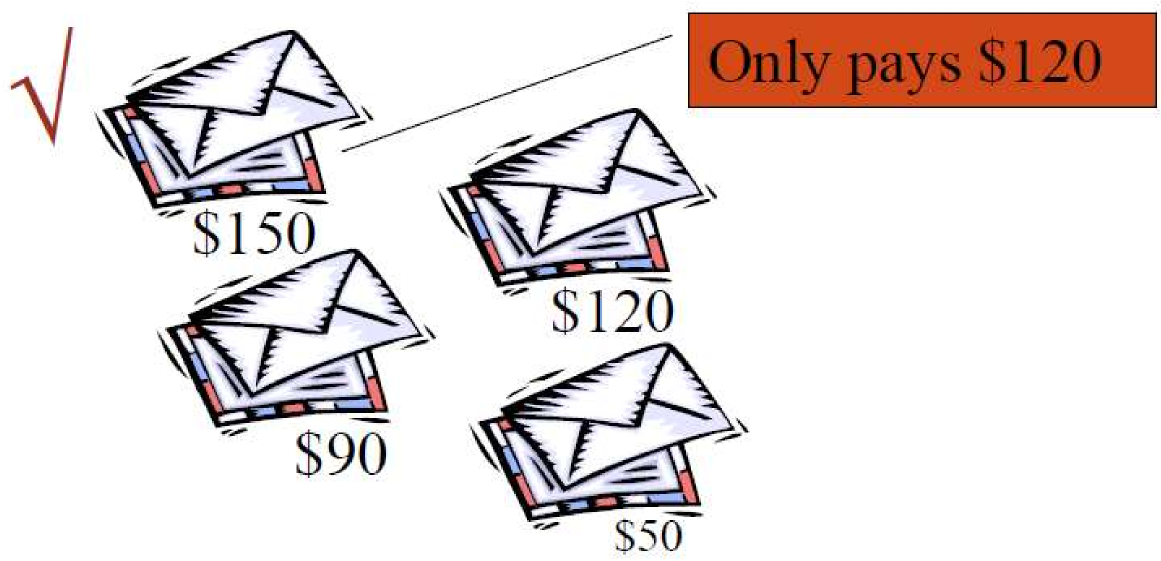
- Advertisers make bids (their moves)
- Advertiser seek attention and volume – prefer higher positions
- Engines order ads and price clicks according to some rules known to all bidders
- The bidders can all keep reacting to each other

Types of Auctions

- First-price sealed-bid
  - Bidders place their bid in a sealed envelope
  - Simultaneously give them to the auctioneer.
  - Highest bidder wins, pays his bid.
- Second-price sealed-bid auctions (Vickrey auctions)
  - Bidders place their bid in a sealed envelope
  - Simultaneously give them to the auctioneer.
  - Highest bidder wins, pays price equal to the second highest bid.
- Open Ascending-bid auctions (English auctions)
  - Price is steadily raised by the auctioneer
  - Bidders drop out once the price becomes too high.
  - Eventually there is only one bidder who wins the auction at the current price.
- Open Descending-bid auctions (Dutch auctions)
  - Price starts at infinity and is steadily lowered by the auctioneer
  - The first bidder to accept the current price, wins
  - Pays the current price.

Second Price Auction (Vickrey Auction)

- All buyers submit their bids privately
- Buyer with the highest bid wins; pays the price of the second highest bid



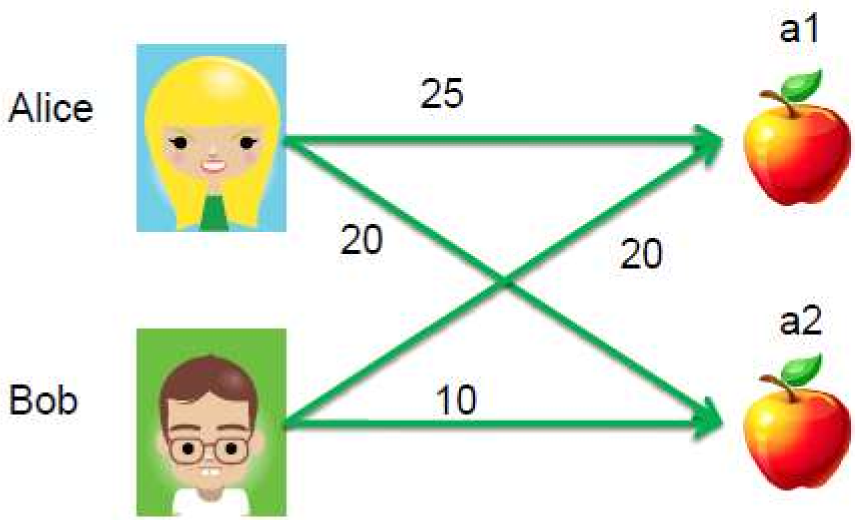
Truthfulness (Incentive Compatibility) of Vickrey Auction

- An auction mechanism is truthful, if the dominant strategy for every player is to truthfully bid their own value.
- Telling the truth is optimal in second-price (Vickrey) auction
- Suppose your value for the item is $100;if you win, your net gain (loss) is $100-price
- If you bid more than $100
  - You increase your chances of winning at price >$100
  - You do not improve your chance of winning for < $100
- If you bid less than $100
  - You reduce your chances of winning at price < $100
  - There is no effect on the price you pay if you do win
- Dominant optimal strategy: bid $100
- Key: The price you pay is out of your control
- Vickrey's Nobel Prize due in large part to this result!

Vickrey-Clark-Groves (VCG)

- Generalization of Vickrey
- Works for arbitrary number of goods, including allowing combination bids
- Auction procedure:
  - Collect bids
  - Allocate goods to maximize total social value (goods go to those who claim to value them most) = maximum weighted matching
  - Payments: Each bidder b pays his externality = (max TSV without b's participation) –(max TSV for everyone else when b participates)
  - NB: (max TSV for everyone else when b participates) = max weighted matching without b & without b's items.
- Incentive compatible (truthful) = all the bidders do best when they bid their true value i.e. reveal their private information

VCG Example



Max matching = 40 → A gets a2, B gets a1
Max matching without Bob = 25
Max matching without Bob, without a1 = 20
- **Bob pays 5**

Max matching without Alice = 20
Max matching without Alice, without a2 = 20
- **Alice pays 0**

- Max matching without Alice does not depend on her bids
- Max matching without Alice and her assigned apple does not depend on her bids
- Price paid by Alice for her apple does not depend on her bid
  - Should not bid more than her value –might pay too much!
  - Should not bid less –might not get it!
- Thus VCG leads to truthfulness

How does the sponsored search auction work

- Search engines
  - run keyword auctions to sell available inventory of ad positions
- Advertisers
  - submit bids which indicate their willingness-to-pay per click
    - for example, bid of $1.75 per click for the keyword “laptop”
- The search engine orders the ads in descending order
  - Bid is a key determinant of ad position
  - Other factors such as CTR are also factored in

Unique Features of the Market for Internet Ads

- Bidding takes place continuously
- The search engines effectively sell flows (clicks/hour)
- Not unlike electricity markets unused capacity is wasted
- On the other hand, user utility might be impaired by excessive advertisement

"Unit" of Advertisement

- Advertiser's perspective: transaction is a "unit"
  - Pricing model: pay per transaction (CPT/CPA)
- Search engine's perspective: exposure is a "unit" (CPM)
  - Pricing model: pay per exposure
- Middle ground: click is a "unit"
  - Pricing model: pay per click (CPC)
- All three pricing models are widely used
- Pay per click dominates sponsored search

Generalized First-Price Auctions

- 1997 auction revolution by Overture (then GoTo.com, created at Idealab)
- Pay per-click for a particular keyword
  - Initially crazy idea, meant to combat search spam
  - Search engine "destination" that ranks results based on who is willing to pay the most
  - With algorithmic search engines out there, who would use it
  - Commercial web sites would! (Much better than to depend on ranking!)
- Results
  - Links arranged in descending order of bids
  - Pay your bid (First price)
  - Overture became a platform for Yahoo! and MSN--Imperfect mechanism: unstable due to dynamic nature of the environment
- Problem: GFP is unstable because bids can be adjusted dynamically

Example on GFP

- Two slots and three bidders
  - ad in first slot: 200 clicks per hour
  - ad in second slot: 100.

- Bidders 1, 2, and 3 have values: \$10, \$4, and \$2
- If bidder 2 bids \$2.01, to make sure he gets a slot
- Bidder 1 will not want to bid more than \$2.02
- Bidder 1 gets the top spot, but then bidder 2 will want to revise his bid to \$2.03 to get the top spot
- Bidder 1 will in turn raise his bid to \$2.04, and so on.

Generalized Second-price Auctions

- Generalized Second-Price (GSP) Auctions
- 2002 GSP implemented by Google
- Yahoo!/Overture and others switched to GSP
- Two way of generalizing:
  - Bid ranking: Order the ads by their bids. Rename ads so ad i ends in position i. Bidder in position I pays bid(i+1).
  - Revenue ranking: Order the ads by expected revenue in position i assuming maximum bids, that is by b(i)*ctr(i) .
    - Rename ads so ad i ends in position i.
    - Bidder in position i pays bid(i+1)*ctr(i+1)/ctr(i)
    - Note that bidder i pays less than bid(i) since bid(i)*ctr(i) > bid(i+1)*ctr(i+1)
    - If all CTRs are the same, revenue ranking is the same as bid ranking!
- CTR can be estimated for an advertiser based on click history

GSP Example

- Same example under GSP mechanism with bid ranking
- Two slots and three bidders.
  - First slot 200 clicks per hour regardless of ad
  - Second slot gets 100 regardless of ad
  - Bidders 1, 2, and 3 have values per click of $10, $4, and $2, respectively.
- If all advertisers bid truthfully, then bids are $10, $4, $2.
  - Payments for slot one and two are $4 and $2 per click.
  - Total payment of bidder 1 is $800 = $1200 pay-off
  - Total payment of bidder 2 is $200 = $200 pay-off
  - In this example truth-telling is an equilibrium because no bidder can benefit by changing his bid.

Is GSP a VCG?

- GSP is not VCG -- GSP has no dominant strategies
- Truth-telling is generally not an equilibrium
- With only one slot, VCG and GSP are identical
- With several slots, the mechanisms are different
  - GSP charges bidder i the bid of bidder i+1 (In practice + $0.01)
  - VCG charges bidder i for his externality

Truth-telling is not a dominant strategy under GSP

- Proof: Example with three bidders and two slots
- Per click values are \$10, \$4, and \$2
- CTR's are 200 and 199
- (Assume all ads are equally attractive)
- If all bid truthfully bidder 1 bids \$10 and pays \$4 so his payoff is
  - (\$10-\$4)*200=\$1200
- If bidder 1 bids \$3 (and pays \$2) his payoff is:
  - (\$10-\$2)*199=\$1592>\$1200

Same Example using VCG

- Let us compute VCG payments for the example considered before.
  - Two slots and three bidders.
    - First slot 200 clicks per hour
    - Second slot gets 100.
    - Bidders 1, 2, and 3 have values per click of \$10, \$4, and \$2, respectively.
- The second bidder's payment is \$200, as before (externality imposed on 3 who loses \$200 = value for him of the slot he does not get!)
- However, the payment of the first advertiser is now \$600
  - \$200 for the externality that he imposes on bidder 3 (by forcing him out of position 2) +
  - \$400 for the externality that he imposes on bidder 2 (by moving him from position 1 to position 2 and thus causing him to lose (200-100)=100 clicks per hour).
- Note that in this example, revenues under VCG are lower than under truth telling equilibrium of GSP!

Nash Equilibrium for GSP

- Locally envy-free equilibrium
  - Fixed point where bidders don't want to move up or down
- Bidders first choose the optimal position for them: position i
- Within range of bids that land them in position i, bidder chooses point of indifference between staying in current position and swapping up with bidder in position i-1

What do we have?

- GFP is not stable
- Choose between GSP and VCG
  - GSP is not truthful
  - VCG is truthful and stable but not really used (revenue?)
- Adaptation of VCG
  - The higher the bid, the better the position
  - The last bidder to get a slot pays same as GSP
  - Total payment of bidder in position i under VCG, p(i)
  - $p(i) = (\alpha_i - \alpha_{i+1})b_{i+1} + p(i+1)$
    - α_i is expected number of clicks at position i
    - b_i is bid of i^{th} highest bidder

Take-away Messages

- Computational Advertising is a new growing field with lots of interesting problems.
- Two main types of ads are display ads and textual ads.
- We studied interesting problems in displaying graphic ads and textual ads.
- Finally, we discussed various auction mechanisms like GFP, GSP, VCG and an adaptation of VCG

Further Reading

- Algorithmic Challenges in Online Advertising, Deepak Agrawal and Deepayan Chakrabarti. CIKM 2008 Tutorial
- Computational Advertising course @ Stanford: http://www.stanford.edu/class/msande239/
- Edelman, Ostrovsky, and Schwarz, Internet Advertising and the Generalized Second Price Auction, 2005
- Varian, Position Auctions, 2006
- Lahaie, Pennock, Saberi, Vohra, Sponsored Search, Chapter 28 in Algorithmic Game Theory, Cambridge University Press, 2007

Preview of Lecture 16: Computational Advertising (Part 2): Contextual Advertising

- Contextual advertising basics
- Ad selection in contextual advertising
- Phrase Extraction for Contextual Advertising
- IR methods for content match ad retrieval
- Holistic view at the page in Contextual Advertising
- When to advertise
- Search-based ad selection for sponsored search
- Predicting clicks

Disclaimers

- This course represents opinions of the instructor only. It does not reflect views of Microsoft or any other entity (except of authors from whom the slides have been borrowed).
- Algorithms, techniques, features, etc mentioned here might or might not be in use by Microsoft or any other company.
- Lot of material covered in this course is borrowed from slides across many universities and conference tutorials. These are gratefully acknowledged.

Thanks!