Abstract

- Online advertising is a form of promotion that uses the Internet and World Wide Web for the expressed purpose of delivering marketing messages to attract customers. Examples of online advertising include text ads that appear on search engine results pages, banner ads, in-text ads, or Rich Media ads that appear on regular web pages, portals or applications. Since its inception over 15 years ago, online advertising has grown rapidly and currently accounts for 13% of the overall advertising spend (which is approximately $700 billion worldwide). A large part of the more recent success in this field has come from the following key factors:
  - Personalization: offline advertising (via broadcast TV, radio, newspaper etc.) is largely a broadcast form of communication where as digital advertising is much more targeted and thus enables a personalized, and possibly informative, message to consumers.
  - Interactivity: internet advertising is becoming increasingly interactive with the advent of new forms of advertising such as social advertising; this is enables advertisers and consumers to operate in a more conversant manner.
  - Engagement: consumers are spending more time online than with any other form of media thereby enabling a broader reach and deeper connection with consumers.
  - Explainability: advertisers are beginning to understand their consumers better.
- This shift in focus in digital advertising from location (i.e., publisher web pages) to personalization has brought with it numerous challenges some of which have received a lot of research attention in the DIK communities (Data, Information and Knowledge communities) over the past 10-20 years. In this talk I will review, along the dimensions outlined above, some of these key technical problems and challenges that arise when advertising becomes personal. This will be done within the context of the elaborate (and ever-evolving) ecosystems of modern day digital advertising where one has to capture, store, and process petabytes of data within the constraints of a, sometimes, sequential workflow. The ultimate goal to is provide millisecond-based decision-making at each step of this workflow that enables customizable and engaging consumer experiences.
Brief Bio James G. Shanahan

• 20 years in the field AI and information management
  – Principal and Founder, Boutique Data Consultancy
    • Clients include: Digg, AT&Ti, Ancestry, SearchMe, SkyGrid, MyOfferPal
  – Lecturer, University of California Santa Cruz (UCSC)
  – Chief Scientist, Turn Inc. (A CPX ad network → DSP)
  – Principal Scientist, Clairvoyance Corp (CMU spinoff; sister lab to JRC)
  – Research Scientist, Xerox Research
  – Research Engineer, Mitsubishi Group
  – PhD in machine learning (1998), University of Bristol, UK;
    B.Sc. Comp. Science (1989), Uni. of Limerick, Ireland
• Now: Machine Learning Consultant (San Francisco)
  – IF (you have large data problems and need a consultant)
    THEN {email me at James.Shanahan_AT_gmail.com}
  – Where problems ∈ {web search, online advertising, machine learning, ranking, user modeling, optimization, social networks}

Disclaimer

• The Author retains all rights, including copyrights and distribution rights.
• No publication or further distribution in full or in part permitted without explicit written permission from the author
• Living vicariously!
Outline

- **Background**: digital advertising, the 0.2% opportunity (of GDP)
- **3rd Generation DA (Personalization)**
  - Explainability: trading desks for the advertiser
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
- **Summary**

What marketers want?

- Deliver marketing messages and attract customers and sell products/services (long term vs. short term)

<table>
<thead>
<tr>
<th>Goal</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduce:Reach</td>
<td>Media Planning</td>
</tr>
<tr>
<td>Influence:Brand</td>
<td>Ad Effectiveness (CTR, site visits)</td>
</tr>
<tr>
<td>Close</td>
<td>Marketing Effectiveness (Transactions, ACR, Credit Assignment)</td>
</tr>
<tr>
<td>Network Effect</td>
<td>Referrals/Advocacy</td>
</tr>
</tbody>
</table>
It annoys but it yet it works…..

• Online advertising is a form of advertising utilizing the Internet and World Wide Web in order to deliver marketing messages and attract customers [wikipedia.com]

• Advertising annoys people! Advertising works!
  – "Half the money I spend on advertising is wasted; the trouble is, I don't know which half." - John Wanamaker, father of modern advertising. [Credit assignment]
  – "I do not regard advertising as entertainment or an art form, but as a medium of information...“, “Ogilvy on Advertising” by David Ogilvy

• Goals of Online advertising
  A – Deliver/push an advertiser’s message
  A+P – Generate ROI for the advertiser and revenue for the publisher
  P+C – Enable ads as a medium of information (true in the case of search)

Advertising makes up ~2% of US GDP

Despite its problems (lack of credit assignment etc.)

• US GDP = $14.1 Trillion (Global $56 Trillion, 56x10^{12})
• US Advertising Spend
  – ~$275 Billion across all media (2% of GDP since the early 1900s)
  – ~$23 Billion in Digital Advertising (8.4% of overall spend)
• In 2008, Worldwide online advertising was $65B
• I.e., about 10% of all ad spending across all media [IDC, 2008]

http://en.wikipedia.org/wiki/Advertising
Why online?: ROI versus CPM

<table>
<thead>
<tr>
<th></th>
<th>CPM</th>
<th>ROI per $1 Spent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outdoor</td>
<td>$1-5</td>
<td>$6.81</td>
</tr>
<tr>
<td>Cable TV</td>
<td>$5-8</td>
<td>$8.60</td>
</tr>
<tr>
<td>Radio</td>
<td>$8</td>
<td>$8.60</td>
</tr>
<tr>
<td>Online</td>
<td></td>
<td>$19.83</td>
</tr>
<tr>
<td>–Display</td>
<td>$5-$30</td>
<td>$19.78</td>
</tr>
<tr>
<td>–Contextual</td>
<td>$1-$5</td>
<td></td>
</tr>
<tr>
<td>–Search</td>
<td>$1 - $200</td>
<td>$21.84</td>
</tr>
<tr>
<td>Social</td>
<td>$0.2-$5</td>
<td>$12.57</td>
</tr>
<tr>
<td>Mobile</td>
<td>$0.2-$6</td>
<td>$7.50</td>
</tr>
<tr>
<td>Email</td>
<td></td>
<td>$44.93</td>
</tr>
<tr>
<td>Network/Local TV</td>
<td>$20</td>
<td>$6.81</td>
</tr>
<tr>
<td>Magazine</td>
<td>$10-30</td>
<td>$10.11</td>
</tr>
<tr>
<td>Newspaper</td>
<td>$30-35</td>
<td>$12.77</td>
</tr>
<tr>
<td>Direct Mail</td>
<td>$250</td>
<td>$7.34-$15.28</td>
</tr>
</tbody>
</table>

TV = 7:1 ROI
Online = 20:1 ROI

[ROI numbers based on DMA Power of direct marketing 2010 Edition]

Big lag in ad dollars spend online

- Typically ad dollars have followed media consumption
- Despite people spending more time on the internet, there is a $40 billion deficit (in US)

$65B $23B

TV = 7:1 ROI
Online = 20:1 ROI

[JP Morgan, 2010]
$14 Billion

$8 Billion

- Volume 16 times Search
- 4Trillion Impression
- Display CPMs: $3-$5 (US)
- Search RPS: $70 (US)
- Search RPM: $10-15
Big questions

- $40 Billion online gap

- $80 billion display gap

Why the big lag in ad dollars spend online?

- **Advertisers and agencies are old fashioned**
  - Don’t use mathematically sophisticated systems; human experts
  - Advertisers are broadcast oriented (TV, Newspaper, Radio etc)
  - Huge culture change
    - Real-time versus weekly/monthly
    - They speak the language of audiences and panels (not data mining, machine learning); social scientists

- **Accountability: Internet is still relatively new and largely unknown**
  - Apart from sponsored search (2% of online activity) other media types are largely not accountable and underexploited
  - Poor targeting
  - Guaranteed markets at spot market prices

- **Huge opportunity if history is anything to go by and ....**
Outline

• Background: digital advertising, the 0.2% opportunity (of GDP)

• 3rd Generation DA (Personalization)
  – Explainability: trading desks for the advertiser
  – Behavioral targeting
  – Interactivity: thru social ads
  – Engagement: transformative ad formats

• Summary

From Mad Men To Wall Street and beyond!

• Set in New York City, *Mad Men* begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.

2007

Data

Personalization

Increasingly

Human Intensive
Lots of guess work
Forward Market

Technology

Based on

Forward Market

Spot Markets

Double digit growth

1st Generation

Advertisers still in broadcast mode

3rd Generation

2nd Generation
Audience Participation
Augmented Reality
Share with my friends: socialize products

Audience Participation
NYC’s Sexiest Billboard Doesn’t Care For Dumbphones


Audience Participation
NYC’s Sexiest Billboard Doesn’t Care For Dumbphones


From Mad Men To Wall Street and beyond!

- Set in New York City, *Mad Men* begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.

2007

Data

Personalization

Increasingly

Double digit growth

Human Intensive
Lots of guess work
Forward Market

Technology
Integrated
Forward Market
Spot Markets

Advertisers still in broadcast mode

1st Generation
2nd Generation
3rd Generation
Advertising: a supply-demand marketplace

DEMAND
Advertiser wishes to reach consumers

SUPPLY
Ad Slots for sale

Formal Relationship

Advertising Agency: creates & traffics ads

DEMAND
Advertiser wishes to reach consumers

SUPPLY
Ad Slots for sale

An advertising agency or ad agency is a service business dedicated to creating, planning and handling advertising for its clients. E.g., SEMs, Omnicom Group...
Advertising Agency: creates & traffics ads

DEMAND
Advertiser wishes to reach consumers

SUPPLY
- Sponsored Search
  - (208 Billion, US, 2009)
- Non Search
  - 4 Trillion, US, 09
- Contextual
- Display
- Classified
- Email
- Social
- Mobile

Consumers

2nd Generation

- CPC, CPA
- Quant driven and quant support

- Supply can be fragmented ➔ Ad Networks
  - Outside of search supply can be fragmented
  - Publishers maybe small and not have a sales team
Advertising Network: Aggregates Publishers

Online Advertising is a Frenemic Network Play
• New more efficient market places
  – Ad Exchanges
  – Data exchanges
• Audience-based targeting
• Very complex pipeline
  – Yield mgt and Demand side platforms
Ad Exchanges: a new SD Marketplace

- The ad exchange is a real time marketplace
  - with an auction-based system where the participants - advertisers and publishers – transact on a common platform to purchase and sell online graphical advertising.

- Currently, publishers sell remnant inventory
  - on the exchange for advertisers to purchase through bidding on a user-friendly interface.

- Ad Exchanges do not compete with ad networks
  - targeting technologies, or publishers, but rather serve as a more efficient way for the exchange of inventory within these groups

- Googles acquired DoubleClick, Yahoo acq RightMedia, etc.. $11 in M&A in 2007
Ad Exchange: auctioneer-centric marketplace

Demand-Side Platform: A trading desk for Adv.
Demand side platforms

Demand Side, Supply side platforms

A DSP is a demand side trading desk; is a SaaS
media suppliers
Campaign management tools to manually
target or automatically optimize campaigns.
Automated bid management capabilities
(RTB).
Advanced analysis and “decisioning” about
the value and desirability of ad impression
opportunities.
Key Features of DSP

- Advanced and accurate audience targeting capabilities
- Easy-to-use inventory control
- Bidding dashboards
- Ability to set frequency caps on the ads being served
  - reaching the "right consumer" too many times can lead to a significant decline in interest

Open research Areas

- Forecasting
- Targeting
- Mechanism design
- Realtime bidding
  - [Selective call out and real time bidding, Chakraborty et al., 2010]
Look-a-like modeling: explainable Knowledge

• Rule-based targeting
  – Advertiser/Agency specifies knowledge of target audience (lives in metropolitan area, 25-44, income, reads business sections)

• Data driven from consumers who bought already
  – Observe individuals who just transacted; find more who look-like these
  – Positive and negative examples
  – Active learning: Instead of assuming that all of the training examples are given at the start, active learning algorithms interactively collect new examples, typically by making queries to a human user. Often, the queries are based on unlabeled data, which is a scenario that combines semi-supervised learning with active learning.


• Forecasting (see CIKM 2010 papers and posters)
  – How many of uniques? What-if I increase by bid price?

Look-a-like modeling: Challenges

• Modeling user based on interaction patterns

• Data e.g., from QuantCast
  – Millions of partner sites
  – 10 Billion weblog records (ad tag firing events from publisher); 250 Billion per month
  – 1 Billion users globally
  – 15 terabytes per day of new data

• Advertisers need to define, refine their models of a typical consumer daily, hourly, secondly

• Forecasting (CIKM 2010)

• Map-Reduce for offline, real-time modeling
Outline

- Background: digital advertising, the 0.2% opportunity (of GDP)
- 3rd Generation DA (Personalization)
  - Explainability: trading desks for the advertiser
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
- Summary

Advent of WWW: media became personal

- Advent of the World Wide Web media consumption starts to become more personalized
  - people could pull the media they wanted to read/watch/listen
  - Media pushed also
- And the Internet banner ad was a marketing revolution.
  - We could not only measure the size of the audience that saw our ads but we had a real-time (or near real-time) feedback mechanism in the click-through rate
- WSJ.com, for example, charges advertisers as much as $64.60 to show a banner ad to 1,000 viewers. $5 CPM

[A Pricing Revolution Looms in Online Advertising, Business Week, 4/2009]
The dreaded banner is a big money winner

WSJ.com, for example, charges advertisers as much as $64.60 to show a banner ad to 1,000 viewers. ➔ $5 CPM
[A Pricing Revolution Looms in Online Advertising, Business Week, 4/2009]

But banner ads are mass marketing

• Measure the size of the audience and CTRs
• However, various studies showed banner blindness
• E.g., comScore/Fox study
  – 84 percent of Web users never click on banners.
  – 12 percent not very active
  – 4% percent of Web users that click on banner ads, produce 67 percent of all clicks.

http://www.comscore.com/Press_Events/Press_Releases/2008-02/Display_Ad_Click-Through_Behavior
Behavioral Targeting: Modeling The User

• Target ads based on user’s browsing behavior
• Commonly used by:
  – Website owners (e.g., e-commerce websites)
  – Ad networks
• Key players include:
  – E-commerce websites such as Amazon
  – Blue Lithium (acq by Yahoo!, $300M), Tacoda (acq by AOL, $275M), Burst, Phorm and Revenue Science, Turn.com, and others...

[For more background see: http://en.wikipedia.org/wiki/Behavioral_targeting]

Personalization via BT

• Intuition:
  – the users who share similar Web browsing behaviors will have similar preference over ads
• Selling Audiences (and not sites)
  – Traditionally did this based on panels (user surveys or using Comscore/NetRatings); very broad and not very accurate
  – Through a combination of cookies and log analysis BT enables very specific segmentation
• Domains of Application
  – Sponsored search
  – Non-Sponsored search (e.g., contextual, display)
Behavioral Targeting: Main Types

- **Website/web page visited**
  - E.g., Users who frequent Orbitz.com and Expedia.com or the travel section of USAToday.com would become part of the “Travel Shoppers” segment. Then, these users are re-targeted when they are found on other more general content type sites.

- **Keyword profile**
  - using recent searches or content that was read

- **Retarget past visitors to your website (surrogate modeling)**
  - The goal is to use BT to let you send the appropriate message to each user based on where they are in your product sales cycle.

BT Technology

- **Segmenting users**
  - Clustering, data mining, classification
  - Rule-based system, hybrid systems
  - Segmenting publisher real estate into categories

- **Collaborative filtering**
  - People who bought this also bought X…
  - Yehuda Koren, Factorization meets the neighborhood: a multifaceted collaborative filtering model. KDD 2008 426-434
  - Remember NetFlix problem??
Organic versus Sponsored Clicks

Organic Clicks

Sponsored Clicks

User Query

[How much the Behavioral Targeting can Help Online Advertising? — Jun Yan, et al., WWW 2009]

Represent user behavior: terms or click’d URLs

- Represent the users by their behaviors using BT strategies
  - Queries: leading to a User by Query Term (query unigram) matrix
    - Use all user queries (in unigram fashion)
    - Stem; remove stopwords and words with a frequency Less than or Eq 1

\[
\begin{array}{c|c|c|c|c}
\text{User} & \text{Q Term}_1 & \text{Term}_2 & \ldots & \text{Term}_l \\
\hline
\text{User}_1 & \text{TF-IDF}_1 & \ldots & & \\
\text{User}_2 & & \ldots & & \\
\text{User}_n & & & \text{TD-IDF}_n & \\
\end{array}
\]

- Clicked page: represent user in terms of URLs of pages she clicked

\[
\begin{array}{c|c|c|c|c}
\text{User} & \text{URL}_1 & \text{URL}_2 & \ldots & \text{URL}_l \\
\hline
\text{User}_1 & \text{U}_1 & \ldots & & \\
\text{User}_2 & & \ldots & & \\
\text{User}_n & & & \text{U}_n & \\
\end{array}
\]

where

\[
U_j = (\log(\text{User}_j \text{ click } \text{URL}_j) + 1) \times \text{IDF(\text{URL})}
\]
Cluster users based on query unigrams

<table>
<thead>
<tr>
<th>User × Term</th>
<th>Q Term₁</th>
<th>Term₂</th>
<th>......</th>
<th>Termₙ</th>
</tr>
</thead>
<tbody>
<tr>
<td>User₁</td>
<td>TF_IDF₁₁</td>
<td>....</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Userₙ</td>
<td></td>
<td></td>
<td></td>
<td>TD_IDFₙ₁</td>
</tr>
</tbody>
</table>

Cluster1

Cluster20

Add ad clicks to matrix measure CTRs

<table>
<thead>
<tr>
<th>User × Term</th>
<th>Q Term₁</th>
<th>Term₂</th>
<th>......</th>
<th>Ad₁</th>
<th>......</th>
<th>Adₘ</th>
</tr>
</thead>
<tbody>
<tr>
<td>User₁</td>
<td>TF_IDF₁₁</td>
<td>....</td>
<td></td>
<td>ClickCnt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User₂</td>
<td></td>
<td></td>
<td></td>
<td>ClickCnt</td>
<td></td>
<td>No Click</td>
</tr>
<tr>
<td>Userₙ</td>
<td></td>
<td></td>
<td></td>
<td>No click</td>
<td></td>
<td>No click</td>
</tr>
</tbody>
</table>

Ads

C1

C20
Add ad clicks to matrix: compare CTRs

User x Term | Q Term_1 | Term_2 | .... | Ad1 | .... | Ad_m
---|---|---|---|---|---|---
User_1 | TF_IDF 11 | .... | ClickCnt | CTR_{AD1,c1} | No Click
User_2 | .... | CTR_{AD1,c1} >> CTR_{AD1,c20} | No click | No click
User_n | .... | CTR_{AD1,c20} | No click | No click

Cluster organic behaviors using K-means

- **Represent the users by their organic behaviors**
  - Queries terms
  - Clicked URLs
- **Cluster users based on organic behaviour**
  - using k-means or CLUTO
- **Evaluate clusters using clicks on sponsored ads**
  - Is there a tendency to group all clicks into one (or small number of organic clusters)?
  - Evaluate how much BT can help online advertising by delivering ads to good user segments.
- **Metrics**
  - Inter and intra cluster, CTR, F-measure, Ads Click entropy
**Short/long term AND Query/Page**

- **7 Days**
  - **Page model**
    - LP: using Long term user behavior all through the seven days and representing the user behavior by Page-views;
  - **Query model**
    - LQ: using Long term user behavior all through the seven days and representing the user behavior by Query terms;

- **1 Day**
  - **Page model**
    - SP: using Short term user behavior (1 day) and representing user behavior by Page-views;
  - **Query model**
    - SQ: using Short term user behavior (1 day) and representing user behavior by Query terms.

**BT Dataset**

- **Focused on English queries**

- **6,426,633 unique users 17,901 ads**
  - Originally 335,170 unique ads within the seven days.
  - Remove anyone who has more than 100 clicks a day
  - But filter out all the ads that have less than 30 clicks within these seven days,
Log File Format

Table 1. Format of click-through log used in our study.

<table>
<thead>
<tr>
<th>UserID</th>
<th>UID030608473X</th>
<th>A user ID for each unique user.</th>
</tr>
</thead>
<tbody>
<tr>
<td>QueryText</td>
<td>xbox</td>
<td>The detailed query text used by the user.</td>
</tr>
<tr>
<td>QueryTime</td>
<td>08-06-03 21:15:47</td>
<td>The time when the query was issued.</td>
</tr>
<tr>
<td>ClickTime</td>
<td>08-06-03 21:16:02</td>
<td>The time when the click occurred after the query was issued.</td>
</tr>
<tr>
<td>ClickURL</td>
<td><a href="http://www.xbox365.com">http://www.xbox365.com</a></td>
<td>The URL which has been clicked by the user.</td>
</tr>
<tr>
<td>IsAd</td>
<td>0</td>
<td>A Boolean value to show the clicked URL is an ad or not.</td>
</tr>
<tr>
<td>NumberAd</td>
<td>3</td>
<td>The number of ads displayed in the search results.</td>
</tr>
<tr>
<td>DisplayAd</td>
<td><a href="http://videogames.half.ebay.com/">http://videogames.half.ebay.com/</a> <a href="http://accessories.us.dell.com/">http://accessories.us.dell.com/</a> <a href="http://www.gamefly.com">http://www.gamefly.com</a></td>
<td>The URL list of all the ads that displayed by the query. (To save space, we only reserve top domain of the ad URL in this example.)</td>
</tr>
</tbody>
</table>

Day-based query model works best

Table 2. Within- and between-ads user similarity.

<table>
<thead>
<tr>
<th></th>
<th>$S_w$</th>
<th>$S_h$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP</td>
<td>0.1417</td>
<td>0.0252</td>
<td>28.9217</td>
</tr>
<tr>
<td>LQ</td>
<td>0.2239</td>
<td>0.0196</td>
<td>44.2908</td>
</tr>
<tr>
<td>SP</td>
<td>0.1532</td>
<td>0.0281</td>
<td>24.5086</td>
</tr>
<tr>
<td>SQ</td>
<td>0.2594</td>
<td>0.0161</td>
<td>91.1890</td>
</tr>
</tbody>
</table>

(a) User clustering by k-means

NIPS MLOAD, 2010 Whistler. Digital advertising: broadcast to personalized advertising © 2010 James G. Shanahan
Paper Conclusions

• **Representing short term behaviors (1 day) is better than 7 days**
  – Query words, URLs models are similar; a user’s interest tends to diverge over long periods so the shorter the period the better

• **Results suffer from some biases**
  – Post-hoc analysis; therefore targeting was not part of study
  – Results measured using 7 days of historical logs (not an AB test)
  – Queries have a strong correlation to the clicked ads but URLs have less of a correlation
  – (Ads are targeted by keywords so users clustered using clicked ads and search terms would yield similar clusters)

• **Having said that, very strong signal: 670% CTR improvement**

How to accomplish BT?

• **BT Scenario**
  – Given positive examples (query terms, clicked results, ignored results)

• **Much research**
  – Filtering [TREC]
  – Adaptive Filter [TREC]
  – Interactive tracks

• **Can use adaptive approaches**
  – Rocchio: weight terms (using base + positive exs + negative exs)
  – Active Learning (base machine learning algorithm with explore+exploit capabilities)
    • Bandit problem; Bayesian updating
  – Offline categorization (and then constrain results set); e.g., SVMs
  – WWW, SIGIR, CIKM, TREC, EC etc.
  – 100 milliseconds
BT at Google Adsense + AdWords

- **Ad network**
  - Google segments users along 20 categories and nearly 600 subcategories.
  - Assign users based on users browser behavior
    - It does not plan to associate the cookie of users with search data or with information from other Google services, like Gmail.
- **Query history to target ads on Google SERPs**
  - When determining which ads to show on a Google search result page, the AdWords system evaluates some of the user's previous queries during their search session as well as the current search query.
  - If the system detects a relationship, it will show ads related to these other queries, too.

Google: editing my BT interests

[Image of Google Adsense interface showing personalization settings]
Personalized Sponsored Search via BT

- **Google uses query history to target ads on Google SERPs**
  - When determining which ads to show on a Google search result page, the AdWords system evaluates some of the user’s previous queries during their search session as well as the current search query.
  - If the system detects a relationship, it will show ads related to these other queries, too.

- **This feature is an enhancement of broad match.**
  - It works by generating similar terms for each search query based on the content of the current query and, if deemed relevant, the previous queries in a user’s search session.
  - Your ad will potentially show if one of your broad-matched keywords matches any of these similar terms.
  - **Win-win**: quality score not affected; Google accepts the cost

http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=74246
BT had a bad name

- **Gaining respect**
  - Google began showing ads on 3/11/2009 to people based on their previous online activities in a form of advertising known as behavioral targeting
  - Interest-based advertising

- **Privacy Protection**
  - Edit BT profiles
    - give users the ability to see and edit the information that it has compiled about their interests for the purposes of behavioral targeting
  - Opt-out easily
    - Google and many others is a participating member of the Network Advertising Initiative.
  - Remove cookies


The Network Advertising Initiative (NAI)

A consortium of approximately 30 companies that use BT technology. Opt-out easily!!
BT Works

- Behavioral targeting tries to model user needs and intents based on past Web-surfing behavior
  - Sponsored search, display, and contextual advertising
- Advertiser reduce CPMs
  - Reduce ad costs from $60 to $3 CPM, a 95% savings
- Publisher can generate more and higher RPMs
  - Find unsold segments of readers and sell them to advertisers that previously couldn’t access them.
  - Use data on past reader "click streams" from other Web sites to help advertisers reach the consumers who are most interested in their products.

Behavioral Targeting is growing

![Graph showing the growth of US Behaviorally Targeted Online Advertising Spending, 2006-2012 (millions)]
Behavioural Targeting Summary

- BT works for both sponsored search and non-sponsored search
  - CTRs >> E.g., 11% to 840% higher than average
- Holds huge promise to monetize longtail of mostly remnant inventory (such as Web2.0)
  - $8 Billion US Display ad market
  - Network effect will increase the value of this inventory (reach more people)
- Privacy concerns need to be addressed before consumers, advertisers and publishers embrace this direction in a big way

Personalized Ads Example

- “The RHS ad was customized and chosen from thousands of different creative elements, automatically and in real-time, by machine-learning algorithms developed by Teracent, a San Mateo, California startup”
Outline

- Background: digital advertising, the 0.2% opportunity (of GDP)
- 3rd Generation DA (Personalization)
  - Explainability: trading desks for the advertiser
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
- Summary

Social Advertising

- Introduced in the context of community-oriented websites
  - such as Facebook and Digg. And more recently Twitter
- Leverages historically "offline" dynamics
  - such as peer-pressure, recommendations, and other forms of social influence
- In the case of Digg, Digg or Bury the ad
  - users determine what ads appear on the website news streams by voting up or "burying" ads,
  - much the way they can digg or "bury" organic news items
  - recently they have rolled this out for graphic ads also

This enables advertisers and consumers to operate in a more conversant manner.
Social Online Advertising
Advertiser-Publisher-User-Community

User+Community+Diggs

Select Ads

Advertiser
Interactive Social Graphic Ads

Social Ads Facebook
Ad Reach within top Social Networks

Twitter ads are limited to search streams

Social Advertising Summary

• Social advertising has the potential to be more conversant with the consumer
  thereby increase the advertisers trust in display-advertising in particular

• Still very new but with huge reach

• Social ads successfully deployed
  much more popular than pure display ads

• 500 million users on Facebook, $1 Billion in revenue
  $2 per user revenue
Outline

• Background: digital advertising
• Personalization thru:
  – Explainability: trading desks for the advertiser
  – Behavioral targeting
  – Interactivity: thru social ads
  – Engagement: the untethered consumer
    • Smarter and more portable devices (3G, 4G, broadband)
    • Social networking is becoming more pervasive
    • Location based services (LBS), mobile video and apps, augmenting websites with social capabilities
    • Transformative advertising
    • Augmented reality and quick response (QR) codes

• Summary

Smartphones 50% share in mid2011

USA only
Pandora

App: UrbanSpoon

http://searchengineland.com/google-adsense-ads-making-their-way-to-iphone-apps-16969
Mobile Web

Tablet PCs

• **Large Format Benefit**
  - Enhanced mobile apps
  - Increased content display opportunities
  - Video optimization
  - Interactive graphics

• **Growth**
  - Total media tablets device market
    • 15MM units in 2010 growing
    • 28MM in 2011 (ABI, 2010; Barclays Capital, 2010)
    • 46MM units in 2014 (IDC, May 10, 2010)
  - Apple expected to sell ~20MM iPads in 2011 (Barclays Capital, July 2010)
Mobile Use & Video

Smartphones driving video explosion
- 51% mobile video growth since 2009 (Nielsen, 2010)
- # of mobile video users will rise by 34% annually to 95MM in 2015 (Coda Research Consultancy, 2010)
- By 2015, US mobile consumers are expected to use 327,000 terabytes of mobile data monthly, rising at a CAGR of more than 117% (Coda Research Consultancy, 2010)
- Mobile video revenue set to exceed $2B in 2013, thanks to 4G mobile broadband (ABI Research, 2010)

Tipping Points

- The true turning point for the industry was the introduction of the smartphone, heightened by Apple’s iPhone launch in mid-2007.
  - Tearing down the carrier wall
- Better connectivity
  - 3G, 4G mobile phones led to better connection speeds
  - Wi-Fi connectivity and the rise of mobile Internet browsing.
- Richer Content
  - Apps, For mobile web pages, Video, TV, Games
- In developing markets, mobile is the primary gateway to the Internet not PC!
Mobile Monetization

- All the internet monetization models are applicable to mobile.
- New monetization models that leverage unique mobile features e.g. pay per call.

<table>
<thead>
<tr>
<th></th>
<th>Search</th>
<th>Contextual</th>
<th>Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPM</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>CPC</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>CPA</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>PPC</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Select Best Ads to Display

- **Requires:**
  - Ads to select from (i.e., ads with budget remaining)
  - Ability to determine best ad to show (see below);
  - Serve ad or not (social responsibility)

- **Selecting Best Ad(s)**
  - Estimate the CTR of each ad using available features.
  - A number of techniques for estimating CTR:
    - Regularized Regression models, Multi-armed bandit models
    - Non parametric: collaborative filtering, nearest neighbor methods.
  - Features for estimating CTR:
    - Location based features,
    - Social network features
    - Site characteristics
    - Ad Characteristics
    - User Characteristics, user behavior
Mobile Ad Spend will triple by 2012

Outline

- Background: digital advertising
- Personalization thru:
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
    - Smarter and more portable devices (3G, 4G, broadband)
    - Social networking is becoming more pervasive
      - Location based services (LBS), mobile video and apps, augmenting websites with social capabilities
      - Transformative advertising
      - Augmented reality and quick response (QR) codes
    - Explainability: trading desks for the advertiser
- Summary
Mobile, smart devices enable social

- The latest wave of post-iPhone smart phones have put mobile social networking, search, location services and even gaming and video into the fast lanes of user adoption.
Outline

• **Background: digital advertising**
• **Personalization thru:**
  – Behavioral targeting
  – Interactivity: thru social ads
  – Engagement: transformative ad formats
    • Smarter and more portable devices (3G, 4G, broadband)
    • Social networking is becoming more pervasive
    • Location based services (LBS), mobile video and apps, augmenting websites with social capabilities
    • Transformative advertising
    • Augmented reality and quick response (QR) codes
  – Explainability: trading desks for the advertiser
• **Summary**
Top Tier Location Based Social Networks

Layering on Social: Google Connect

- “Google Friend Connect instantly awakens and strengthens the community that visits your site by enriching it with social features”.
Add social gadgets to your website

- Choose from a range of social gadgets that let users post comments, post links, rate, review, and more.
- The gadget gallery includes gadgets created by Google and OpenSocial developers.

Simply add HTML code to webpage
Hot Areas: Model network behavior

• Need new tools to model information (product) diffusion in social networks (email, Facebook, MySpace, Digg etc.)

• Influence models
  – Investigate the direct peer influence on information diffusion among social networks.
  – Explore different features, both associated with the users or the information, correlate with information diffusion.
  – [User grouping behavior in online forums. Shi, Zhu, Cai, Zhang, KDD’09.]
  – [Information diffusion in computer science citation networks, Shi, Tseng, Adamic, ICWSM, 2009]

• Other Diffusion models

Vertex-Importance Graph Synopsis (VIGS)

[The very small world of the well-connected. Shi, Bonner, Adamic, Gilbert, ACM Hypertext, 2008]

• Create sub-graphs of important vertices to study both key vertices and the entire graph

• Which vertices are important?
  – People with most friends on facebook? The most quoted blogs?
  – Standard, well-established measures: Degree, Betweenness, Closeness, PageRank
Important vertices in real networks are densely connected, no matter which importance measure we use.

Properties of VIGS in real networks

Compare the properties of important vertices in subgraphs and in the original graphs.

<table>
<thead>
<tr>
<th></th>
<th>Erdos-Renyi</th>
<th>BuddyZoo</th>
<th>TREC</th>
<th>Web</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Density</strong></td>
<td>Less dense</td>
<td>Denser</td>
<td>Denser</td>
<td>Denser</td>
</tr>
<tr>
<td><strong>Connectivity</strong></td>
<td>Disconnected when the number of vertices is small</td>
<td>Almost connected</td>
<td>Always connected</td>
<td>Always connected</td>
</tr>
<tr>
<td><strong>Average Shortest Path</strong></td>
<td>Much longer</td>
<td>Almost the same</td>
<td>Almost the same</td>
<td>Almost the same</td>
</tr>
<tr>
<td><strong>Rank of Importance</strong></td>
<td>Very different</td>
<td>Similar</td>
<td>Similar</td>
<td>Similar</td>
</tr>
</tbody>
</table>

Get good idea of relationships between important vertices using only the subgraphs.
Diffusion in Blogs

- Data – Blogs:
  - We crawled 45,000 blogs for 1 year
  - 10 million posts and 350,000 cascades

Blogs: Cascade shapes

The probability of observing a cascade on n nodes follows a Zipf distribution:
\[ p(n) \sim n^{-2} \]
**Next: Creating big cascades**

- Blogs – information epidemics
  - Which are the influential/infectious blogs?
- Viral marketing
  - Who are the trendsetters?
  - Influential people?
- Disease spreading
  - Where to place monitoring stations to detect epidemics?

---

**Diffusion in partially observable networks**

- In many applications, the underlying network over which the diffusions and propagations spread is actually *unobserved*.
- Develop a method for tracing paths of diffusion and influence through networks and inferring the networks over which contagions propagate.
- Given the times when nodes adopt pieces of information or become infected, identify the optimal network that best explains the observed infection times.
- This optimization problem is NP-hard to solve exactly
  - develop an efficient approximation algorithm that scales to large datasets and in practice gives provably near-optimal performance.


---

NIPS MLOAD, 2010 Whistler. Digital advertising: broadcast to personalized advertising © 2010 James G. Shanahan
Graph Algs are Computational Expensive

<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>d1</td>
<td>the quick brown fox</td>
</tr>
<tr>
<td>d2</td>
<td>the fox ate the mouse</td>
</tr>
<tr>
<td>d3</td>
<td>how now brown cow</td>
</tr>
</tbody>
</table>

Assume block size is 20 Characters

> hadoop dfs -put f1.txt exampleDir

Outline

- **Background**: digital advertising
- **Personalization thru**:
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
    - Smarter and more portable devices (3G, 4G, broadband)
    - Social networking is becoming more pervasive
    - Location based services (LBS), mobile video and apps, augmenting websites with social capabilities
  - Transformative advertising
    - Augmented reality and quick response (QR) codes
  - Explainability: trading desks for the advertiser
- **Summary**
Quick Response Codes (QR)

What is a QR code?
- A QR Code is a matrix barcode (or two-dimensional code), readable by QR scanners, mobile phones with camera, and smartphones

How do people use QR codes?
- Mobile users scan or take a picture of a QR code to trigger an action:
  - Display web or mobile pages
  - Send an SMS/MMS
  - Multi-action menu
  - Display coupons/offers
  - Save a contact
  - Send an e-mail
  - Calendar entry
  - Generate video
  - Make a voice call
  - Promotion entry/data collection
  - Send tweets
  - Register for/enter loyalty programs

NIPS MLOAD, 2010 Whistler. Digital advertising: broadcast to personalized advertising © 2010 James G. Shanahan
Augmented reality (AR)

- AR is a layering technology; used to overlay tagged images, video files, applications and data to display information over video, image, and camera displays on PC’s, mobile phones and TV/monitors
  - E.g., NFL game broadcasts use the yellow electronic line to indicate where the first down is located
- AR can use GPS and location-based targeting, as well as image and facial recognition
- Different formats commonly used?
  - Print-Based
    - Leverages print to drive consumer and brand engagement
    - 3-D Image & Video Use
- Smartphone Layering
  - Utilizes smartphone camera to
  - Geo-location content
AR +QR: Virtual Dressing Room

Engagement

- Smarter and more portable devices (3G, 4G, broadband)
- Social networking is becoming more pervasive
  - location based services (LBS), mobile video and apps, augmenting websites with social capabilities
- Transformative advertising
  - augmented reality and quick response (QR) codes

- How many more rewrites
  - Basic webpages
  - Web2.0
  - Mobile
  - IPTV
Outline

• Background: digital advertising
• Personalization thru:
  – Behavioral targeting
  – Interactivity: thru social ads
  – Engagement: transformative ad formats
  – Explainability: trading desks for the advertiser
• Other issues
  – Credit assignment
• Summary

Credit Assignment is a new area...

• Giving credit where credit is due (offline and online)
• Marketers know that customers rarely travel along a purchase path in a neat, tidy, linear fashion.
  – Typically, they can do up to five searches before getting close to a page to convert, and in the middle of those searches, can look at an e-mail blast promotion, go to a store and look at product in person, and perhaps visit influential social media sites along the way.
  – Which of these activities had a true effect on a purchase and the relative influence of each.
  – In the past, analytics relied on a “last click” model, assigning 100% of the purchasing influence to the last click a consumer made before purchasing.

Credit Assignment: The Customer Buying Cycle

Outline

- Background: digital advertising, the 0.2% opportunity (of GDP)
- 3rd Generation DA (Personalization)
  - Explainability: trading desks for the advertiser
  - Behavioral targeting
  - Interactivity: thru social ads
  - Engagement: transformative ad formats
- Summary
Key Enablers

- Offline advertising is largely broadcast
  - (via broadcast TV, radio, newspaper etc.) is largely a broadcast form of communication whereas digital advertising is much more targeted and thus enables personalized, and possibly informative, message to consumers.

- "Always connected" has graduated from desktop to handheld
  - Bridging offline with online
  - Smartphones (>50% will have one next year), Tablet
  - Life-logging capabilities
  - IP-based TV

- Storage and Data requirements
  - Store, Process data
  - Algorithms make all this happen

Always connected → Lifelogging

http://www.imrc.kist.re.kr/wiki/LifeLog
Research Opportunities

- Data sparsity
- Feature engineering
- Optimization theory and objectives
- Metrics
- Data storage and throughput
- User modeling
- Credit assignment
- Economic modeling, game theory, mechanism design

From Mad Men To Wall Street and beyond!

- Set in New York City, Mad Men begins in 1960 at the fictional Sterling Cooper advertising agency on New York City's Madison Avenue.

2007

Data

Personalization

Human Intensive
Lots of guess work
Forward Market

Technology
Data Driven
Forward Market
Spot Markets

Advertisers still in broadcast mode

1st Generation

2nd Generation

3rd Generation
Executive Summary: Digital Advertising

- **3rd Generation Online advertising: Location → personalization**
  - Advertisers deliver a message to consumers via publishers
  - $23.4 billion in USA (2008), $65B worldwide (10% of overall ad spend)
  - Online advertising was location, location and location ($11B M&A, 2007)
  - But it is more about personalization, personalization, personalization
    - Behavioral Targeting, Engaging Ads, AR, Trading Benches
    - and data, features, objectives, optimization, and emergent behaviors
  - Machine learning is a key driver in the success of online advertising

- **As a discipline**
  - Business informatics/engineering/analytics
  - Bottom up: Economics, IR, ML, OR, Game Theory, Stats, Social Sciences
  - Top-down: create (under)grad programs
    - UC Santa Cruz (ISM209, ISM250, ISM251), Stanford University

ICWSM (Social Media) 2011 in Barcelona

Co-located with IJCAI

**IMPORTANT DATES:**
- Sponsorship immediately!!
- Tutorial proposals due: Feb 18, 2011
- Workshop proposals due: Feb 18, 2011
- Papers due: Jan 31, Feb 6, 2010
- Conference July 17-20
Elizabeth: Thanks

EMAIL:
James_DOT_Shanahan_AT_gmail_DOT_com

Digital advertising is a frenemy culture

• Digital advertising is “frenemic” (players are simultaneously a partner and a competitor).