Bing: User Intent and Decision Engine

Harry Shum, PhD
Corporate Vice President
Microsoft Corporation
Why Decision Engine
Bing Demos
Search Interaction model
Data-driven Research Problems
Q & A
Opportunities for Search Innovation

1. Imprecise Results
2. Research Sessions
3. Inform Decisions
Users Relying More on Search

- 66% of people are using search more frequently to make decisions.

Q. In the past six months have you used a search engine to help inform your decisions for the following tasks?

<table>
<thead>
<tr>
<th>Task</th>
<th>Queries/Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Purchase</td>
<td>75%</td>
</tr>
<tr>
<td>Local Activity</td>
<td>62%</td>
</tr>
<tr>
<td>Flight or Hotel</td>
<td>45%</td>
</tr>
<tr>
<td>Healthcare Research</td>
<td>43%</td>
</tr>
</tbody>
</table>

Decision Sessions are Lengthy

- Length of Sessions by Type:
  - Navigational: 1.90 mins
  - Transactional: 4.76 mins
  - Informational: 6.03 mins

- 8.87 mins

- 9.60 mins

Users need help with tasks and making decisions

Complex task and decision sessions could be easier

Sources: Microsoft Internal Research conducted by iPos 2009; Live Search Session Analysis
Demos

Frederick Savoye
Senior Director
Bing Product Marketing
Search User Interaction Model
Search User Interaction Model

Search engine
- Objective: getting the user relevant information (a website)
- Model: getting out of search results page with a simple click
- Challenge: query – URL matching

Decision engine
- Objective: completing the task by fulfilling user intent
- Model: exploring search results by clicking and browsing
- Challenge: whole page relevance
Bing interaction model

Explore pane (or left rail)
- TOC: verify user intent
- Related search: expand user intent
- Search history: remind user intent

Task completion
- D-cards: showing structural information for the site
- Hover: showing more information for the site
- Simple tasks: e.g. Fedex tracking number
- ...
Data-driven Research Problems
A lot of data…
- Billions of query logs, documents, pictures, clicks, etc.
- Processing them is costly and takes time

Statistical learning + distributed computing
- Can we train 1 Billion samples (query – URL pairs)
- Within a few hours? No over-fitting?

Two examples of Bing-MSR
- “Bing-it-on” N-gram
- Statistical model for log mining
LM Applications in Search
- Query processing: alterations, expansions, suggestions
- Document processing: classification, clustering
- Matching and ranking
- Powerset.com

Leading technology: N-gram
- P(next word | N-1 preceding words)
- Better “understanding” = model predicts better
Challenges to build N-gram for Search

- High quality model needs lots of data at web-scale:
  - Billions of documents, trillions of words, PetaBytes of storage
- Smoothing:
  - How to deal with a very long tail
- Freshness:
  - Web contents are added and revised rapidly

Bing-MSR innovation: Constantly Adapting LM (CALM)
- Highly parallel algorithm designed for cloud computing
- Refine LM as soon as documents are crawled
"Bing-It-On" Ngram Services

We are sharing our resources
- For details, go to [http://www.facebook.com/microsoftresearch](http://www.facebook.com/microsoftresearch) and follow “Bing-It-On Ngram”
- Compare Bing-It-On with Google’s Release

<table>
<thead>
<tr>
<th>Content Types</th>
<th>Google</th>
<th>Bing-It-On Ngram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Types</td>
<td>Raw Count only</td>
<td>Count and smoothed models</td>
</tr>
<tr>
<td>Highest order N</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Training Size (Body)</td>
<td>~ 1.0 trillion words</td>
<td>&gt; 1.3 trillion</td>
</tr>
<tr>
<td># of 1-gram (Body)</td>
<td>13 million</td>
<td>1 billion</td>
</tr>
<tr>
<td># of 5-gram (Body)</td>
<td>1 billion</td>
<td>237 billion</td>
</tr>
<tr>
<td>Availability</td>
<td>DVDs from LDC</td>
<td>On demand web services hosted by MS</td>
</tr>
<tr>
<td>Update</td>
<td>September 2006</td>
<td>Monthly</td>
</tr>
</tbody>
</table>
Data driven decision is first class citizen in search
  Toolbar and search logs: ~10 TB/Day each

Bing uses log mining to
  Understand what users want
  Assess how we are doing
  Quickly improve Bing
    Query Processing
    Ranking
    User experience

Examples:
  Relevance inference using Bayesian Browsing Model (BBM)
  User behavior understanding with Hidden Markov Model (HMM)
Mining search log with HMM

- Search log records only clicked results
  - Skipped results are not explicitly recorded
- Hidden data mining
  - Model viewed results as Markov chain
  - Skipped results = hidden data
- How well does it work?
  - Very close to eye tracking data
What’s In the Name?

- approachable
- global
- memorable
- fresh
- can be used as a verb
- easy to spell
- the Sound of Found
Thank you!