Searching and Finding in a Marketplace: A Long Tail TellTale

Neel Sundaresan

http://labs.ebay.com
When asked if he understood that the laser pointer was broken, the buyer said “Of course, I’m a collector of broken laser pointers”
Divine Reward!

Neel Sundaresan
What wags the Tail?

eBay users trade about $1,400 worth of goods on the site every second.

On an average day on eBay...

A vehicle sells every 2 minutes

A part or accessory sells every 3 seconds

Diamond jewelry sells every 83 seconds

A Timberland shoe sells every 10 minutes

A trading card sells every 6 seconds

Neel Sundaresan
The Long Tail Nature

- Buyers outnumber sellers 5:1
- Seller-items sold has a power law distribution
- Seller-revenue has a power law distribution
- Buyer-items bought has a power law distribution
  - For a sample period over 1 month mean is 3 and median is 1
- Buyer-spent money has a power law distribution
  - For a sample period of 1 month mean is 98$ and median was 45$
- Categories browsed
  - mean 10.5, median 5
- ~10M new items a day, most items eventually sell, items last from a day to 30 days, most items not cataloged, some auction-some fixed price
Search: Beyond Text Relevance…

- Trust
- Value
- Diversity
Items Won by Searcher vs. Search Result Display Rank

\[ y = 2121.7x^{0.753} \]

\[ R^2 = 0.9752 \]

Boyd et al. 2009
An iPhone is not a Vintage Barbie!

Singh, et al. 2010

Neel Sundaresan
Tusting… Familiar Strangers…

- Stamps

- Antiques

Shen, Sundaresan 2007

Neel Sundaresan
Text Relevance

• Near similar titles
  “Apple IPOD Nano 4GB Black NEW! Great Deal!”
  “Apple IPOD Nano 4GB Black NEW! Skin Great Deal!”

What is someone querying for “IPOD Nano” looking for?
Desirability = Demand - Supply
Desirability

- Desirability can be captured using Desirability events like Click-throughs, Bid-throughs, Purchases with positive and negative weights with dampening

\[
\text{Desirability} = \text{Demand} - \text{Supply}
\]

\[
\text{Desirability}(t,w) = \text{Demand}(t,w) - \text{Supply}(t,w)
\]

\[
\text{Demand}(t+k,w) = \text{Supply}(t,w) + \text{Desirability}(t,w)
\]

Where \( w \) is the demand entity (viz. Query)

Gupta, et al 2007
Bridging the Gap – Helping the Seller Enhanced Cranfield Model!

• Item title is a limited set of words
  “Antique Barbie doll”

• A seller should know their potential buyers
  What the queries are likely to be?
  “vintage barbie”

• What terms in the title truthfully represent the item for sale and are also preferred by the buyers queries
  “vintage Barbie Box Mint Mattel NRFB”

http://labs.ebay.com/bayestimate/
What’s That in English?

- Avon SSS 6.7 FL OZ Shave Gel NIB FS GWP

- SKIN SO SOFT
- NEW IN BOX
- Free Shipping
- Gift With Purchase
Metamorphosis of an Acronym: Skin So Soft = SSS

A Translation Model:
Frequency of acronym terms in query
Frequency of expansion in title
co-occurrence of (acronym, expansion) in query sessions
co-occurrence of (acronym, expansion) in title

Neel Sundaresan

Mauge’ et al. 2009
What’s in a word?

- Spelling corrections for Swarowski?
  64 of them!

swarovski, swaroski, swavorski, swarovsky, swarsi, swarowski, swarofski,
swarvski, sworovski, swarovksi, swarosky, svarovski, swarowsky, swarkovski, swarovki,
svarowski, swaraski, swaroviski, swarovski, swaravski, swarorski, swartski, swarovsk,
swaovski, swarvosky, sworoski, swalovski, svorvski, swavroski, swrovski,
swaorvski, swavoski, sawrovski, swarovsi, swarovasaki, swaroski, swarozski, swarouski,
swarokski, swarvaski, sworvoski, swarvowski, swarosvki, svarovsky, swaravoski, swarovky,
sawarski, sarovski, swarozvki, swarocski, swarovskl, swarovski, swarovsky, swarovski,
swarvske, swarovvski, swarovvski, saworvski, swarosvski, swarovrski, swarivski, swarovo ski

K. Mauge’ 2008

Neel Sundaresan
Is Warren Buffet Predictable?

Neel Sundaresan

Parikh, Sundaresan 2010
Trending Haiti – Influence on Relevance

Parikh, Sundaresan 2010
Beyond Relevance - Diversity

- Borrowed from Ecology on Species Diversity Index - Simpson Index

\[ D = 1 - \frac{\sum_{i=1}^{S} n_i(n_i - 1)}{N(N - 1)} \]

Seller Diversity
\[ D_s = 1 - \frac{\sum_{|s|} n(n - 1)}{N(N - 1)} \]

Format Diversity
\[ D_f = 1 - \frac{\sum_{|f|} n(n - 1)}{N(N - 1)} \]

Item Diversity is computed from the Desirability Index

Overall Diversity is a composite of the feature-specific diversity measures

Neel Sundaresan
# Measuring Diversity When Relevant

<table>
<thead>
<tr>
<th>Query</th>
<th>Diversity Score</th>
<th>Some observations from Result Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miss depression glass</td>
<td>0.45</td>
<td>Found salad plates but did not find iced tea glasses.</td>
</tr>
<tr>
<td>Hootie the blowfish</td>
<td>0.46</td>
<td>Found concert tickets but did not find jerseys.</td>
</tr>
<tr>
<td>Peru</td>
<td>0.46</td>
<td>Found collection of stamps but did not find showpieces.</td>
</tr>
<tr>
<td>Sten</td>
<td>0.48</td>
<td>Found british WWII sten manual but did not find mkii and mkiii stens.</td>
</tr>
<tr>
<td>Aloha</td>
<td>0.48</td>
<td>Found Hawaiian vintage, but did not find classic shirts and cribs.</td>
</tr>
<tr>
<td>Buffets</td>
<td>0.59</td>
<td>Found oak sideboards and clarinets but did not find mirrors and saxophones.</td>
</tr>
<tr>
<td>Roger federer</td>
<td>0.65</td>
<td>Found shirts and cards but no rackets and photographs.</td>
</tr>
</tbody>
</table>

Neel Sundaresan
Finding Without Searching – Towards Serendipity

Neel Sundaresan
From Searching To Finding

- Relevance as a derivative of query
- Diversity as a derivative of relevance
- Serendipity as a derivative of diversity
Findables that are not Searchables

Popular Queries → Relevant Items w/ Diversity

User Behavior

Personalized Serendipitous Findables → Recommended Items

User Behavior
Along diversity dimensions

Neel Sundaresan
Thanks - Questions?

Contact: nsundareshan@ebay.com

Neel Sundaresan