

Evaluating the Accuracy of Implicit Feedback from Clicks and Query Reformulations in Web Search

T. Joachims, L. Granka, Bing Pan, H. Hembrooke,
F. Radlinski, G. Gay

**Dept. of Computer Science & Dept. of Information Science
Cornell University**

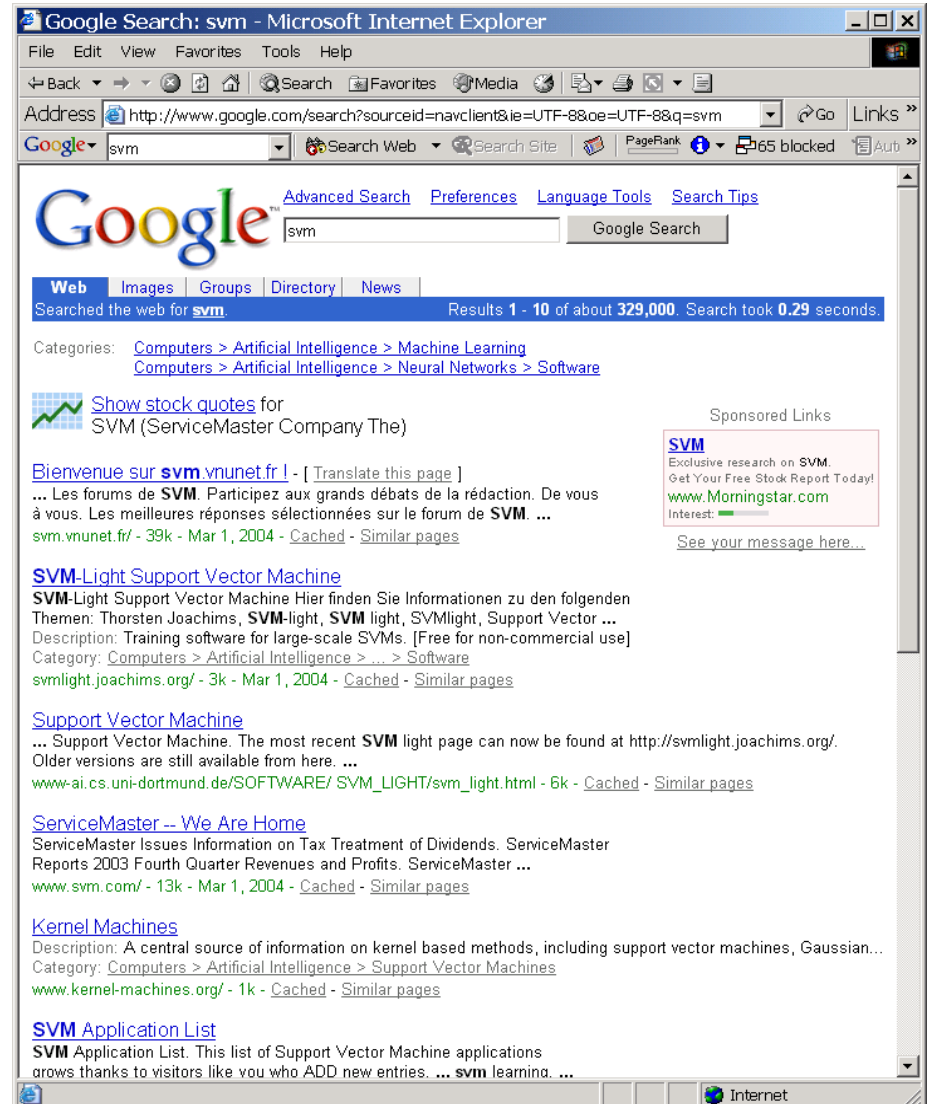
Overview of Talk

- ➔ • **Understanding how users act**
 - User study of Web search behavior using eye-tracking
- **How clicks relate to relevance**
 - Interpreting clicks as relative vs. absolute feedback
 - Dealing with presentation bias
 - Accuracy of feedback strategies
- **Learning from user behavior**
 - Learning ranking functions: Ranking SVM

Sources of Feedback

- ~~Explicit Feedback~~
 - Overhead for user
 - Only few users give feedback

⇒ not representative
- Implicit Feedback
 - Queries, clicks, time, mousing, scrolling, etc.
 - Personalized, democratic, timely, cheap, abundant
 - More difficult to interpret



Is Implicit Feedback Reliable?

How do users choose where to click?

- How many abstracts do users evaluate before clicking?
- Do users scan abstracts from top to bottom?
- Do users view all abstracts above a click?
- Do users look below a clicked abstract?

How do clicks relate to relevance?

- Absolute Feedback:
Are clicked links relevant? Are not clicked links not relevant?
- Relative Feedback:
Are clicked links more relevant than not clicked links?

1. Kernel Machines
<http://www.kernel-machines.org/>
2. Support Vector Machine
<http://jbolivar.freesevers.com/>
3. SVM-Light Support Vector Machine
http://ais.gmd.de/~thorsten/svm_light/
4. An Introduction to SVMs
<http://www.support-vector.net/>
5. Support Vector Machine and ...
<http://svm.bell-labs.com/SVMrefs.html>
6. Archives of SUPPORT-VECTOR...
<http://www.jisc.ac.uk/lists/SUPPORT...>
7. Lucent Technologies: SVM demo applet
<http://svm.bell-labs.com/SVMsvt.html>
8. Royal Holloway SVM
<http://svm.dcs.rhnc.ac.uk>
9. SVM World
<http://www.svmworld.com>
10. Fraunhofer FIRST SVM page
<http://svm.first.gmd.de>

User Study: Eye-Tracking and Relevance

- **Scenario**

- WWW search
- Google search engine
- Subjects were not restricted
- Answer 10 questions

- **Eye-Tracking**

- Record the sequence of eye movements
- Analyze how users scan the results page of Google

- **Relevance Judgements**

- Ask relevance judges to explicitly judge the relevance of all pages encountered
- Compare implicit feedback from clicks to explicit judgments



Web Images Groups Directory News
Searched the web for michael jordan statistician

[The Statistician - Things That Make You Go Hmmm](#)
... Things That Make You Go Hmmm... Our intent in developing this section of The Statistician was to provide the webmasters with diverse information. ...
www.thestatistician.com/archives/060802/hmmm.html - 41k - Cached - Similar pages

[The Statistician - Things That Make You Go Hmmm](#)
... Nothing, 1,302,540, 0.501177394. The \$20 Million A Year Man "I can accept failure, but I can't accept not trying." - Michael Jordan Michael Jeffrey Jordan ...
www.thestatistician.com/archives/060801/page32.html - 20k - Cached - Similar pages

[WebSeer ...](http://infolab.cs.uchicago.edu/webseer/)
... Searching for "Michael Jordan, photograph" retrieves images of Jordan kissing the ... Carnegie-Mellon's. University of Chicago statistician Yali Amit ...
www-news.uchicago.edu/releases/96/961120.webseer.shtml - 9k - Cached - Similar pages

[2003 schedule](#)
... Head Coach/Offensive Coordinator: Michael Esposito, Defensive Coordinator: Brad Winder, Defensive Line ... Quarterbacks: Jordan Haylor, Head Statistician: Steve St ...
www.footballme.com/coaches%20list.htm - 17k - Cached - Similar pages

Eye Tracking Measurements

- **Lookzone for each result**
- **Data capture**
 - Eyetracker:
 - Fixations per lookzone
 - Clicks
 - Typing
 - HTTP-Proxy
 - Remove ads
 - All pages viewed
 - All pages in results list

The screenshot shows a Google search for "Cornell HCI". The search bar contains "Cornell HCI" and the "Google Search" button is visible. Below the search bar, there are navigation tabs for "Web", "Images", "Groups", "Directory", and "News". The search results are displayed in a list format, with several results highlighted by red boxes:

- Cornell Human-Computer Interaction Group**
The HCI Group at Cornell University is an interdisciplinary research team investigating social, psychological, and design issues surrounding the use of ...
Description: HCI Group working out of Cornell University in Ithaca NY, United States.
Category: Computers > Human-Computer Interaction > Departments
www.hci.cornell.edu/ - 3k - Cached - Similar pages
- Cornell HCI Group: Projects**
Context Aware CampusAware An electronic hand-held campus tour system MUSE Utilizing hand-held computers in museums Wireless NOMAD Studies wireless computing in ...
www.hci.cornell.edu/projects.php - 3k - Cached - Similar pages
[More results from www.hci.cornell.edu]
- Cornell HCI Group: People: Michael Stefanone**
... In 2001, he completed the Masters degree program at Cornell University, focusing on communication technology and human-computer interaction (HCI). ...
www.cs.cornell.edu/boom/2002sp/extproj/www.hci.cornell.edu/people.php?id=5.htm - 4k - Cached - Similar pages
- Cornell HCI Group: People: Geri Gay**
... Geri Gay is director of the Human Computer Interaction Group (HCI Group) and a professor at Cornell University in the department of Communication and ...
www.cs.cornell.edu/boom/2002sp/extproj/www.hci.cornell.edu/people.php?id=3.htm - 4k - Cached - Similar pages
[More results from www.cs.cornell.edu]
- Cornell HCI Group: Publications**
Journal Publications Hembrooke, H. & Gay, G. (in press) The Lecture and the Laptop: The Effects of Multitasking in the Classroom. ...
www.img.cornell.edu/publications.php - 9k - Cached - Similar pages
- Search Result for Geri Gay**
... Cornell HCI Group: People: Geri Gay <http://www.hci.cornell.edu/people.php?id=3>
171 Points (Preselect:72) Source: FAST(25) FAST: Geri Gay is director of the ...
hpsearch.uni-trier.de/hp/a-tree/g/Gay:Geri.html - 7k - Cached - Similar pages

Experiment Setup

- **Task**
 - Answer 10 questions
 - Start with Google search, no restrictions
 - Users unaware of study goal
- **10 Questions**
 - Balanced informational and navigational
- **Study (Phase I)**
 - 36 subjects
 - Undergraduate students
 - Familiar with Google

Who discovered the first modern antibiotic?

Find the homepage of Emeril - the chef who has a TV cooking program.

What actor starred as the main character in the original 'Time Machine' movie?

Find the page displaying the routemap for Greyhound buses.

You are excited to cast your vote in the democratic presidential primary - when can you do so in NY?

Find the homepage of Michael Jordan, the statistician.

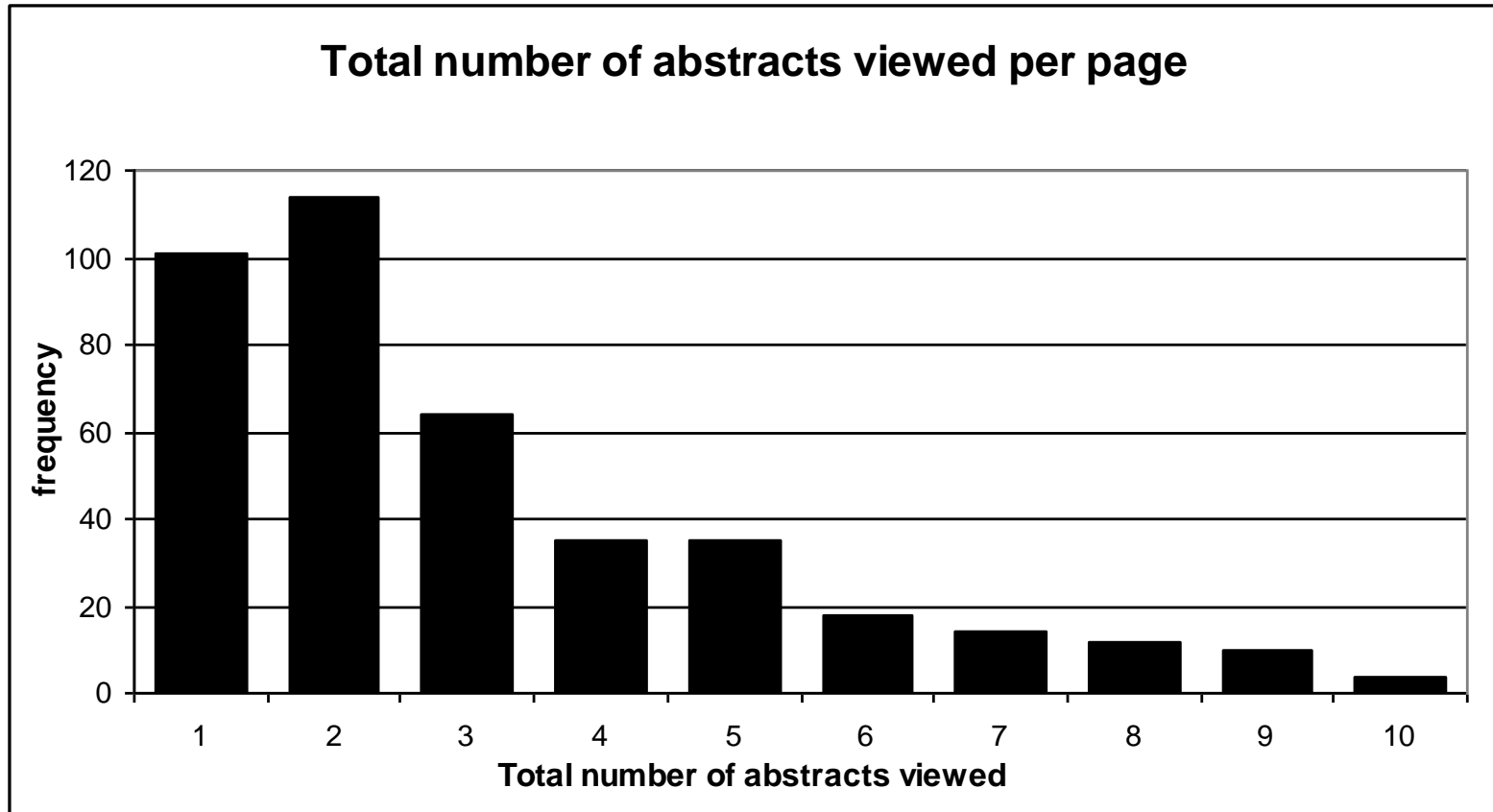
Where is the tallest mountain in NY located?

Find the homepage for graduate housing at Carnegie Mellon University.

A friend told you that Mr. Cornell used to live close to campus - between University and Stewart Aves - does anyone live in his house now; if so, who?

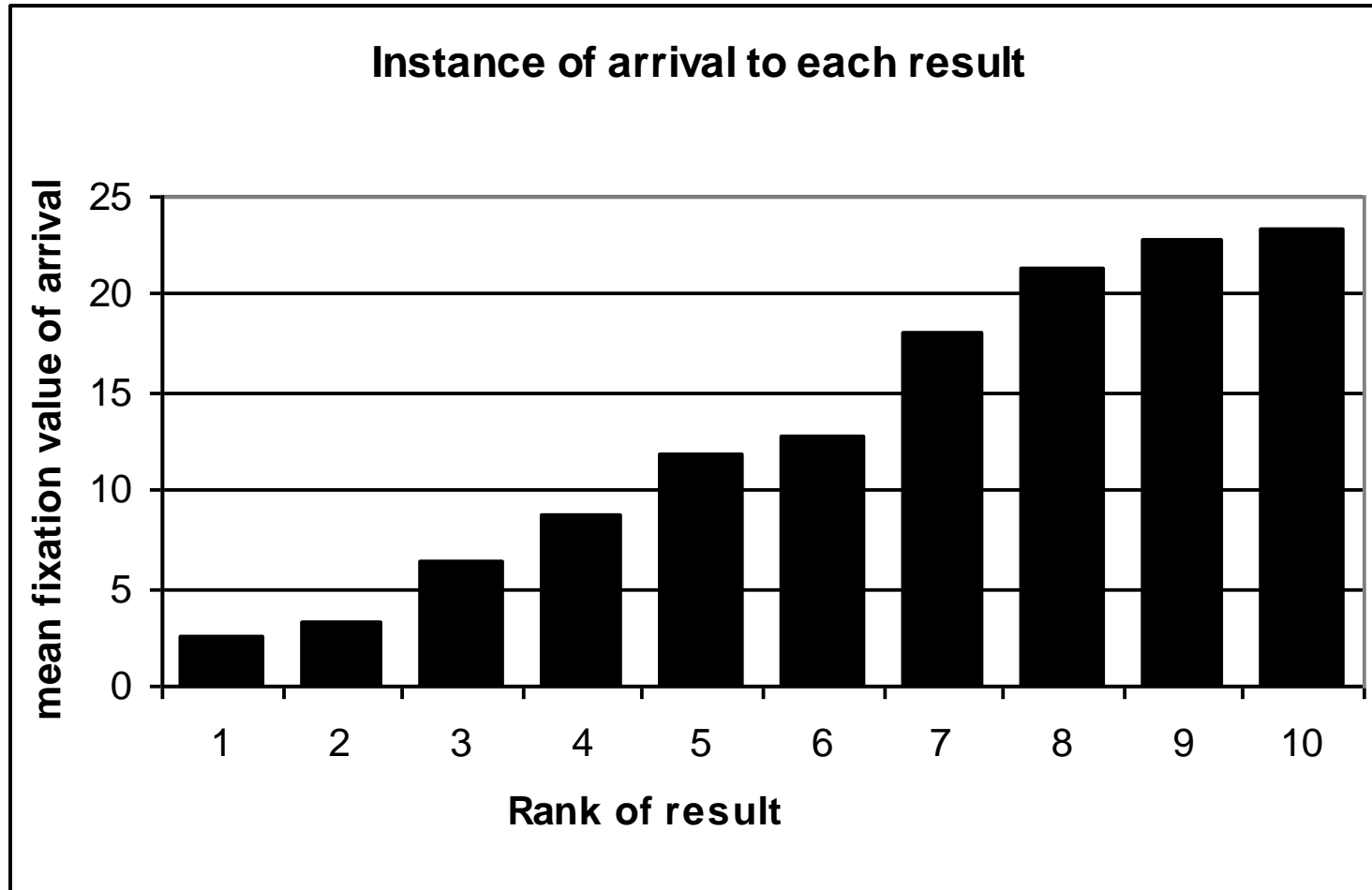
Find the homepage of the 1,000 Acres Dude Ranch.

How Many Links do Users View?



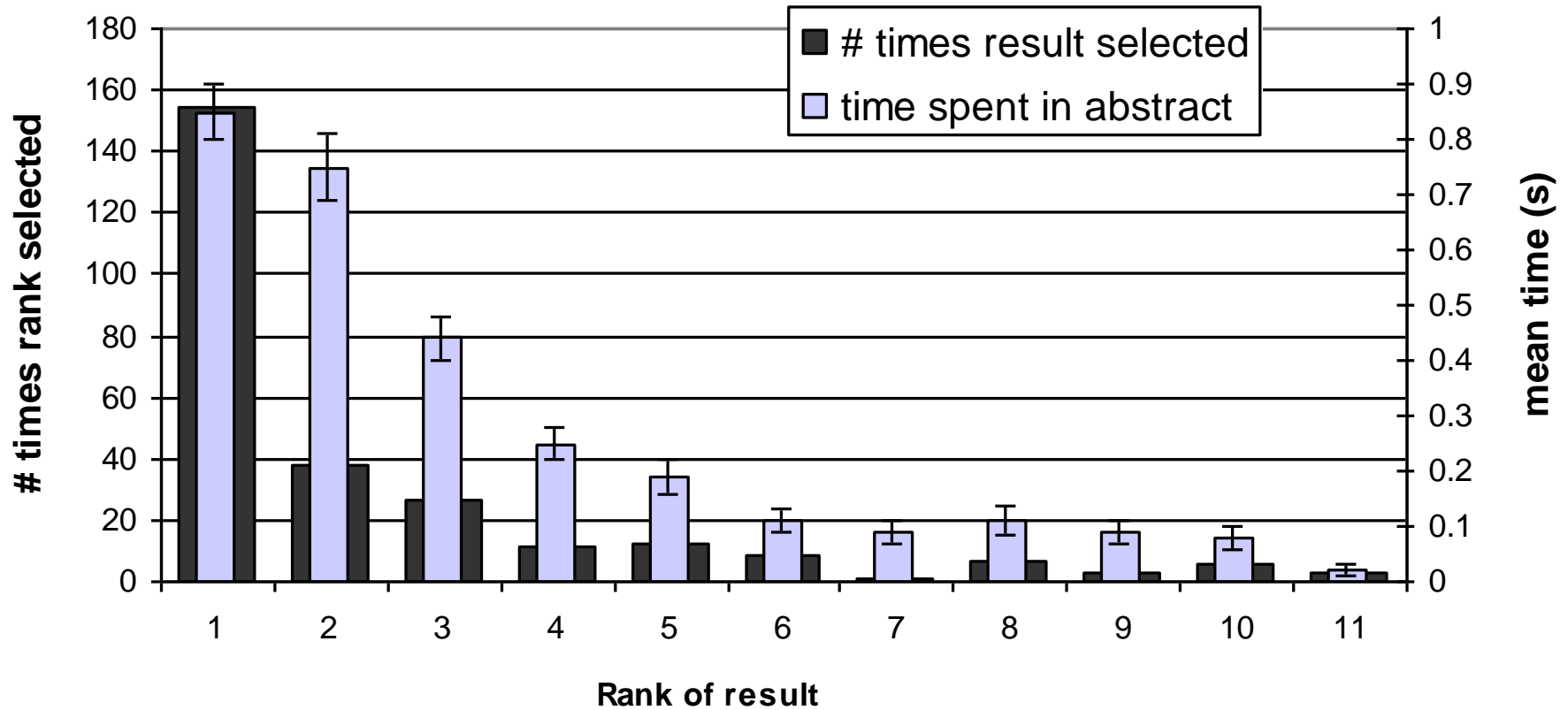
Mean: 3.07 Median/Mode: 2.00

In Which Order are the Results Viewed?



=> Users tend to read the results in order

Looking vs. Clicking



=> Users view links one and two more thoroughly / often

=> Users click most frequently on link one

Do Users Look Below the Clicked Link?

| Viewed Rank | Clicked Rank | | | | | |
|-------------|--------------|-------|-------|-------|--------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 90.6% | 76.2% | 73.9% | 60.0% | 54.5% | 45.5% |
| 2 | 56.8% | 90.5% | 82.6% | 53.3% | 63.6% | 54.5% |
| 3 | 30.2% | 47.6% | 95.7% | 80.0% | 81.8% | 45.5% |
| 4 | 17.3% | 19.0% | 47.8% | 93.3% | 63.6% | 45.5% |
| 5 | 8.6% | 14.3% | 21.7% | 53.3% | 100.0% | 72.7% |
| 6 | 4.3% | 4.8% | 8.7% | 33.3% | 18.2% | 81.8% |

=> Users typically do not look at links below before they click (except maybe the next link)

Conclusions: Decision Process

- Users most frequently view two abstracts
- Users typically view results in order from top to bottom
- Users view links one and two more thoroughly and often
- Users click most frequently on link one
- Users typically do not look at links below before they click (except maybe the next link)

=> **Design strategies for interpreting clickthrough data that respect these properties!**

Overview of Talk

- **Understanding how users act**
 - User study of Web search behavior using eye-tracking
- • **How clicks relate to relevance**
 - Interpreting clicks as relative vs. absolute feedback
 - Dealing with presentation bias
 - Accuracy of feedback strategies
- **Learning from user behavior**
 - Learning ranking functions: Ranking SVM

Feedback from Clickthrough Data

Relative Feedback:

Clicks reflect preference between observed links.

Absolute Feedback:

The clicked links are relevant to the query.

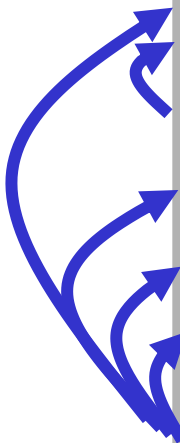
(3 < 2),

(7 < 2),

(7 < 4),

(7 < 5),

(7 < 6)



1. **Kernel Machines**
<http://svm.first.gmd.de/>
2. Support Vector Machine
<http://jbolivar.freeservers.com/>
3. **SVM-Light Support Vector Machine**
<http://ais.gmd.de/~thorsten/svm light/>
4. An Introduction to Support Vector Machines
<http://www.support-vector.net/>
5. *Support Vector Machine and Kernel ... References*
<http://svm.research.bell-labs.com/SVMrefs.html>
6. Archives of SUPPORT-VECTOR-MACHINES ...
<http://www.jiscmail.ac.uk/lists/SUPPORT...>
7. **Lucent Technologies: SVM demo applet**
<http://svm.research.bell-labs.com/SVT/SVMsvt.html>
8. Royal Holloway Support Vector Machine
<http://svm.dcs.rhbnc.ac.uk>

Rel(1),
NotRel(2),
Rel(3),
NotRel(4),
NotRel(5),
NotRel(6),
Rel(7)

User Study:

How do Clicks Relate to Relevance?

- **Experiment (Phase II)**
 - Additional 16 subjects
 - Experiment setup same at Phase I
- **Manipulated Rankings**
 - **Normal:** Google's ordering
 - **Swapped:** Top Two Swapped
 - **Reversed:** Ranking reversed
 - Manipulations not detected by subjects
- **Manually Judged Relevance**
 - Abstract
 - Page

1. **Kernel Machines**
<http://www.kernel-machines.org/>
2. Support Vector Machine
<http://jbolivar.freesevers.com/>
3. **SVM-Light Support Vector Machine**
http://ais.gmd.de/~thorsten/svm_light/
4. An Introduction to SVMs
<http://www.support-vector.net/>
5. Support Vector Machine and ...
<http://svm.bell-labs.com/SVMrefs.html>
6. Archives of SUPPORT-VECTOR...
<http://www.jisc.ac.uk/lists/SUPPORT...>
7. **Lucent Technologies: SVM demo applet**
<http://svm.bell-labs.com/SVMsvt.html>
8. Royal Holloway SVM
<http://svm.dcs.rhbnc.ac.uk>
9. SVM World
<http://www.svmworld.com>
10. Fraunhofer FIRST SVM page
<http://svm.first.gmd.de>

Presentation Bias

Hypothesis: Order of presentation influences where users look, but not where they click!

| | l_1^-, l_2^- | l_1^+, l_2^- | l_1^-, l_2^+ | l_1^+, l_2^+ | total |
|-----------|----------------|----------------|----------------|----------------|-------|
| “normal” | 45 | 33 | 4 | 3 | 85 |
| “swapped” | 64 | 36 | 11 | 3 | 114 |

=> Users appear to have trust in Google’s ability to rank the most relevant link first.

Presentation Bias

~~Hypothesis: Order of presentation influences where users look, but not where they click!~~

| "normal" | l_1^-, l_2^- | l_1^+, l_2^- | l_1^-, l_2^+ | l_1^+, l_2^+ | total |
|-------------------------------------|----------------|----------------|----------------|----------------|-------|
| $\text{rel}(l_1) > \text{rel}(l_2)$ | 15 | 19 | 1 | 1 | 36 |
| $\text{rel}(l_1) < \text{rel}(l_2)$ | 11 | 5 | 2 | 2 | 20 |
| $\text{rel}(l_1) = \text{rel}(l_2)$ | 19 | 9 | 1 | 0 | 29 |
| total | 45 | 33 | 4 | 3 | 85 |
| "swapped" | l_1^-, l_2^- | l_1^+, l_2^- | l_1^-, l_2^+ | l_1^+, l_2^+ | total |
| $\text{rel}(l_1) > \text{rel}(l_2)$ | 11 | 15 | 1 | 1 | 28 |
| $\text{rel}(l_1) < \text{rel}(l_2)$ | 17 | 10 | 7 | 2 | 36 |
| $\text{rel}(l_1) = \text{rel}(l_2)$ | 36 | 11 | 3 | 0 | 50 |
| total | 64 | 36 | 11 | 3 | 114 |

Quality-of-Context Bias

~~Hypothesis: Clicking depends only on the link itself, but not on other links.~~

| | Rank of clicked link as sorted by relevance judges |
|-------------------------|---|
| Normal + Swapped | 2.67 |
| Reversed | 3.27 |

=> Users click on less relevant links, if they are embedded between irrelevant links.

Are Clicks Absolute Relevance Judgments?

- **Clicks depend not only on relevance of a link, but also**
 - On the position in which the link was presented
 - The quality of the other links
- => Interpreting Clicks as absolute feedback extremely difficult!**

Strategies for Generating Relative Feedback

Strategies

- “Click > Skip Above”
 - (3>2), (5>2), (5>4)
- “Last Click > Skip Above”
 - (5>2), (5>4)
- “Click > Earlier Click”
 - (3>1), (5>1), (5>3)
- “Click > Skip Previous”
 - (3>2), (5>4)
- “Click > Skip Next”
 - (1>2), (3>4), (5>6)

1. Kernel Machines
<http://www.kernel-machines.org/>
2. Support Vector Machine
<http://jbolivar.freesevers.com/>
3. SVM-Light Support Vector Machine
http://ais.gmd.de/~thorsten/svm_light/
4. An Introduction to SVMs
<http://www.support-vector.net/>
5. Support Vector Machine and ...
<http://svm.bell-labs.com/SVMrefs.html>
6. Archives of SUPPORT-VECTOR...
<http://www.jisc.ac.uk/lists/SUPPORT...>
7. Lucent Technologies: SVM demo applet
<http://svm.bell-labs.com/SVMsvt.html>
8. Royal Holloway SVM
<http://svm.dcs.rhbnc.ac.uk>
9. SVM World
<http://www.svmworld.com>
10. Fraunhofer FIRST SVM page
<http://svm.first.gmd.de>

Comparison with Explicit Feedback

| Explicit Feedback Data Strategy | Abstracts Phase I "normal" |
|---------------------------------|----------------------------|
| Inter-Judge Agreement | 89.5 |
| Click > Skip Above | 80.8 ± 3.6 |
| Last Click > Skip Above | 83.1 ± 3.8 |
| Click > Earlier Click | 67.2 ± 12.3 |
| Click > Skip Previous | 82.3 ± 7.3 |
| Click > No Click Next | 84.1 ± 4.9 |

=> All but "Click > Earlier Click" appear accurate

Is Relative Feedback Affected by Bias?

| Explicit Feedback Data Strategy | Abstracts Phase II | | |
|---------------------------------------|-----------------------|-------------|-------------|
| | “normal” | “swapped” | “reversed” |
| Click > Skip Above | 88.0 ± 9.5 | 79.6 ± 8.9 | 83.0 ± 6.7 |
| Last Click > Skip Above | 89.7 ± 9.8 | 77.9 ± 9.9 | 84.6 ± 6.9 |
| Click > Earlier Click | 75.0 ± 25.8 | 36.8 ± 22.9 | 28.6 ± 27.5 |
| Click > Skip Previous | 88.9 ± 24.1 | 80.0 ± 18.0 | 79.5 ± 15.4 |
| Click > No Click Next | 75.6 ± 14.5 | 66.7 ± 13.1 | 70.0 ± 15.7 |

⇒ **Significantly better than random in all conditions, except “Click > Earlier Click”**

How Well Do Users Judge Relevance Based on Abstract?

| Explicit Feedback Data Strategy | Abstracts | Pages |
|---------------------------------|-------------|-------------|
| | Phase II | |
| | all | all |
| Inter-Judge Agreement | 82.5 | 86.4 |
| Click > Skip Above | 83.1 ± 4.4 | 78.2 ± 5.6 |
| Last Click > Skip Above | 83.8 ± 4.6 | 80.9 ± 5.1 |
| Click > Earlier Click | 46.9 ± 13.9 | 64.3 ± 15.4 |
| Click > Skip Previous | 81.6 ± 9.5 | 80.7 ± 9.6 |
| Click > No Click Next | 70.4 ± 8.0 | 67.4 ± 8.2 |

⇒ **clicks based on abstracts reflect relevance of the page well**

Feedback across Query Chains

The image displays two side-by-side screenshots of Microsoft Internet Explorer, illustrating a search query chain. The left window shows a search for "svm" on the MSN Search engine. The search results page lists several links, including "Buy SVM Stock for \$4", "ServiceMaster: In-depth Company Info", "ServiceMaster -- We Are Home", "SVM-World.de", "SVM srl", "School of Volunteer Management", and "SV Mattersburg Online". A black box with the word "reformulate" in white text has an arrow pointing from the search bar area of the left window to the search bar area of the right window. The right window shows a search for "support vector machine" on the MSN Search engine. The search results page lists several links, including "Programming Vector File Format Support", "Support Vector Machines", "Buy 'Support Vector Machines' at BN.com", "Support Vector Machines - The Book - Support Vector", "Support Vector Machine - The Software", "Support vector machine - Wikipedia, the free encyclopedia", and "GIST: Support Vector Machine 1.0 - Data submission". The browser's address bar in both windows shows the MSN Search URL with the respective query. The browser's search bar also contains the query. The browser's status bar at the bottom of both windows shows "Internet".

Conclusions: Implicit Feedback

- **Interpreting clicks as absolute feedback is difficult**
 - Presentation Bias
 - Quality-of-Context Bias
- **Relative preferences derived from clicks are accurate**
 - “Click > Skip Above”
 - “Last Click > Skip Above”
 - “Click > Skip Previous”

Model of User Behavior

Users select the most promising (biased) action among the alternatives they observed.