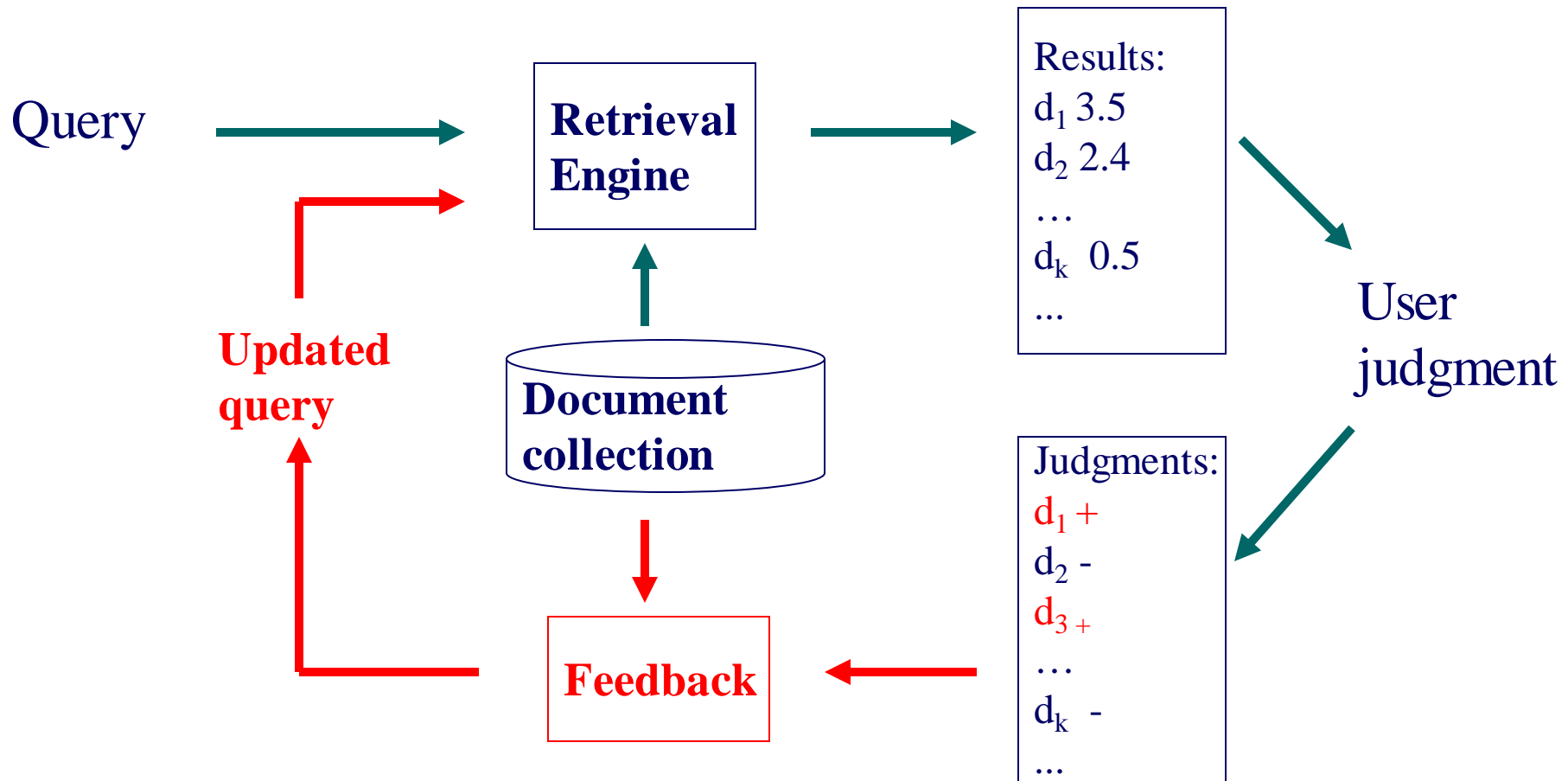


# Implicit User Feedback

Hongning Wang

CS@UVa

# Explicit relevance feedback




# Relevance feedback in real systems

- Google used to provide such functions

[Personalization](#) - Wikipedia, the free encyclopedia  

**Personalization** involves using technology to accommodate the differences between individuals. Once confined mainly to the Web, it is increasingly becoming a ...

[en.wikipedia.org/wiki/Personalized](#) - 42k - [Cached](#) - [Similar pages](#) - 

Relevant

[Personalized Gifts from Personalization Mall](#)  

It shows you went out of your way to find the perfect gift and to **personalize** it to make it theirs alone! At PersonalizationMall.com, we design most of our ...

[www.personalizationmall.com/Default.aspx?&did=111028](#) - 47k -

[Cached](#) - [Similar pages](#) - 

Nonrelevant

[What is personalization?](#) - a definition from Whatis.com  

Mar 6, 2007 ... On a Web site, **personalization** is the process of tailoring pages to individual users' characteristics or preferences.

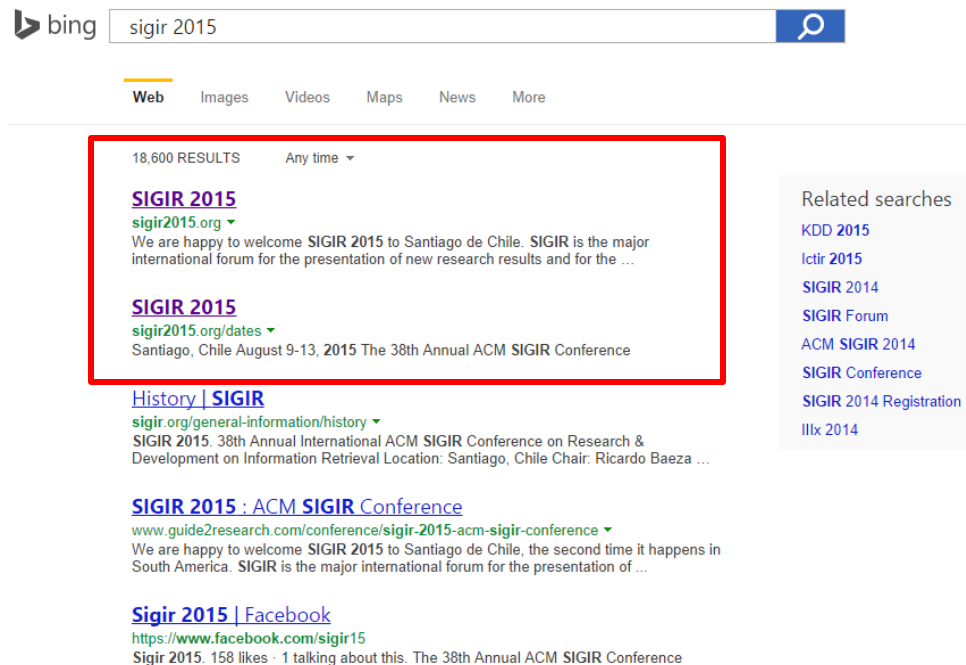
[searchcrm.techtarget.com/sDefinition/0,,sid11\\_gci532341,00.html](#) - 72k -

[Cached](#) - [Similar pages](#) - 

– Vulnerable to spammers

# How about using clicks

- Clicked document as relevant, non-clicked as non-relevant
  - Cheap, largely available



# Is click reliable?

- Why do we click on the returned document?
  - Title/snippet looks attractive
    - We haven't read the full text content of the document
  - It was ranked higher
    - Belief bias towards ranking
  - We know it is the answer!

# Is click reliable?

- Why do not we click on the returned document?
  - Title/snippet has already provided the answer
    - Instant answers, knowledge graph
  - Extra effort of scrolling down the result page
    - The expected loss is larger than skipping the document
  - We did not see it....

*Can we trust click as relevance feedback?*



# Accurately Interpreting Clickthrough Data as Implicit Feedback [Joachims SIGIR'05]

- Eye tracking, click and manual relevance judgment to answer
  - Do users scan the results from top to bottom?
  - How many abstracts do they read before clicking?
  - How does their behavior change, if search results are artificially manipulated?

# Which links do users view and click?

- Positional bias

*Fixations: a spatially stable gaze lasting for approximately 200-300 ms, indicating visual attention*

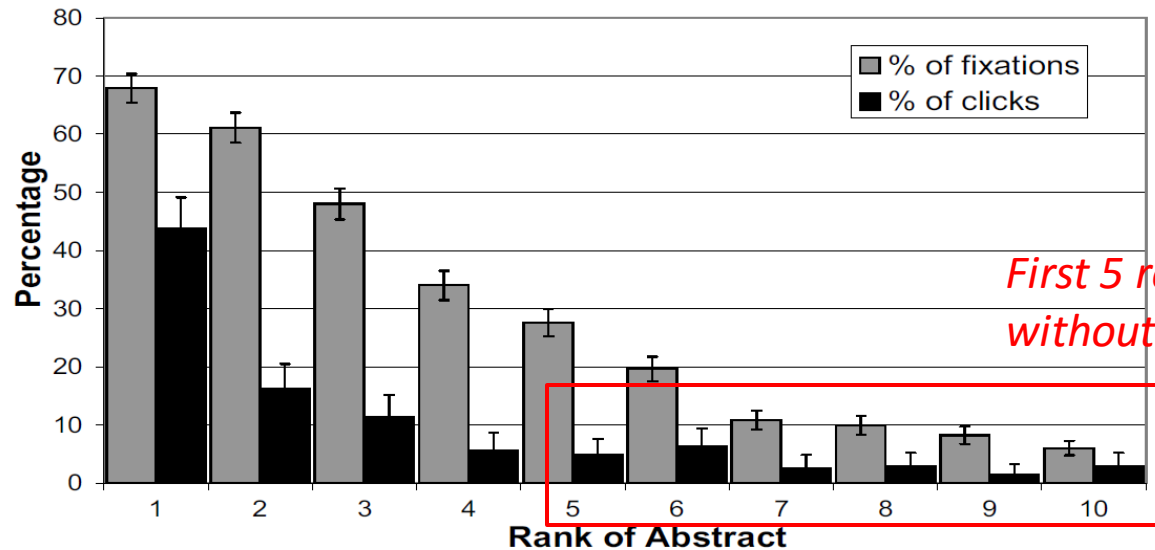
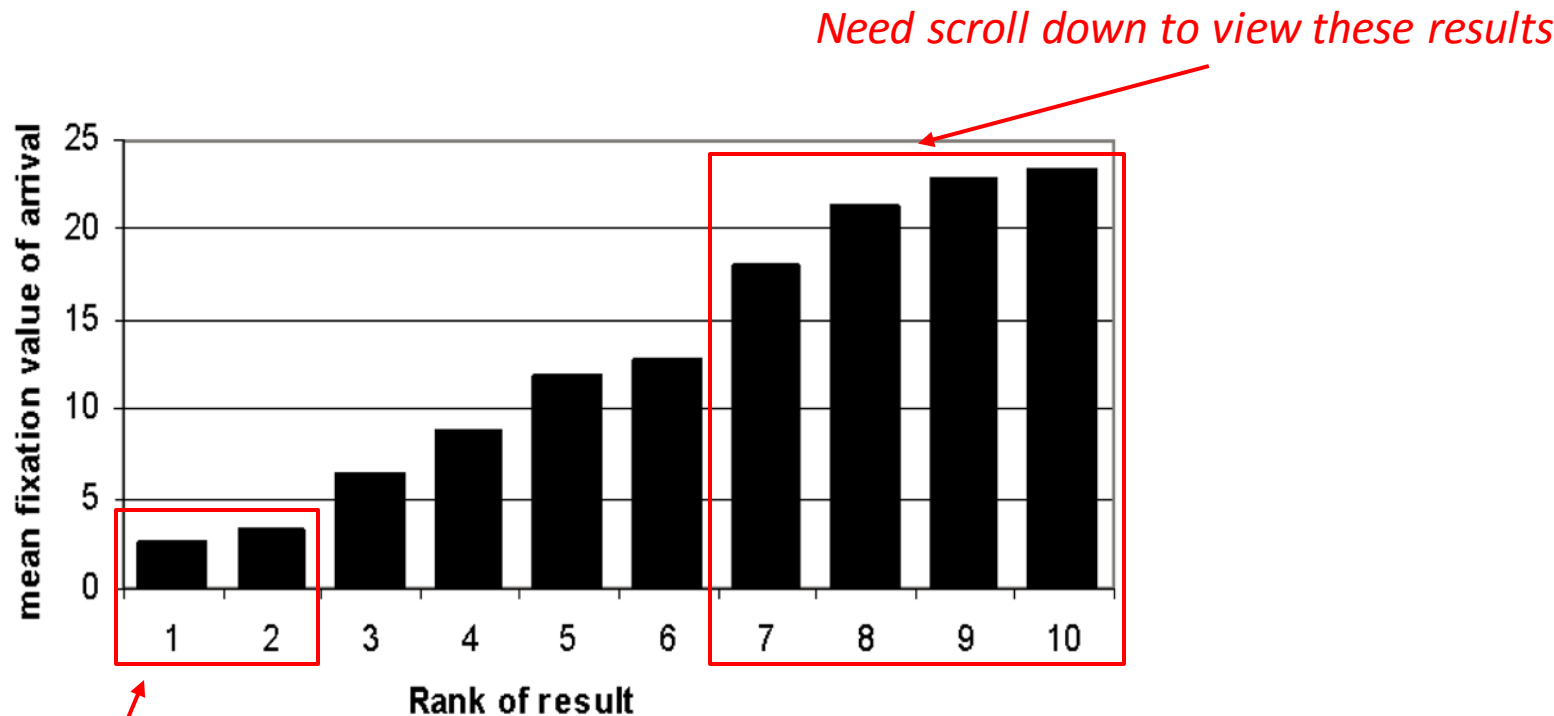


Figure 1: Percentage of times an abstract was viewed/clicked depending on the rank of the result.



# Do users scan links from top to bottom?



**Figure 2:** Mean time of arrival (in number of previous fixations) depending on the rank of the result.

*View the top two results within the second or third fixation*

# Which links do users evaluate before clicking?

- The lower the click in the ranking, the more abstracts are viewed before the click

**Table 2:** Percentage of times the user viewed an abstract at a particular rank before he clicked on a link at a particular rank.

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%

# Does relevance influence user decisions?

- Controlled relevance quality
  - Reverse the ranking from search engine
- Users' reactions
  - Scan significantly more abstracts than before
  - Less likely to click on the first result
  - Average clicked rank position drops from 2.66 to 4.03
  - Average clicks per query drops from 0.8 to 0.64

# Are clicks absolute relevance judgments?

- Position bias

- Focus on position one and two, equally likely to be viewed

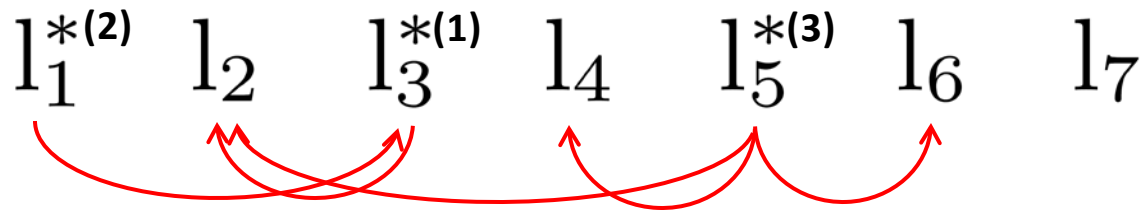
“normal”	$l_1^-, l_2^-$	$l_1^+, l_2^-$	$l_1^-, l_2^+$	$l_1^+, l_2^+$	total
$\text{rel}(l_1) > \text{rel}(l_2)$	15	19	1	1	36
$\text{rel}(l_1) < \text{rel}(l_2)$	11	5	2	2	20
$\text{rel}(l_1) = \text{rel}(l_2)$	19	9	1	0	29
total	45	33	4	3	85

“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

# Are clicks relative relevance judgments?

- Clicks as pairwise preference statements
  - Given a ranked list and user clicks



- Click > Skip Above
- Last Click > Skip Above
- Click > Earlier Click
- Last Click > Skip Previous
- Click > Skip Next

# Clicks as pairwise preference statements

- Accuracy against manual relevance judgment

Explicit Feedback Data Strategy	Abstracts				
	Phase I “normal”	“normal”	Phase II “swapped”	“reversed”	all
Inter-Judge Agreement	89.5	N/A	N/A	N/A	82.5
Click > Skip Above	80.8 ± 3.6	88.0 ± 9.5	79.6 ± 8.9	83.0 ± 6.7	83.1 ± 4.4
Last Click > Skip Above	83.1 ± 3.8	89.7 ± 9.8	77.9 ± 9.9	84.6 ± 6.9	83.8 ± 4.6
<del>Click &gt; Earlier Click</del>	<del>67.2 ± 12.3</del>	<del>75.0 ± 25.8</del>	<del>36.8 ± 22.9</del>	<del>28.6 ± 27.5</del>	<del>46.9 ± 13.9</del>
Click > Skip Previous	82.3 ± 7.3	88.9 ± 24.1	80.0 ± 18.0	79.5 ± 15.4	81.6 ± 9.5
<del>Click &gt; No Click Next</del>	<del>84.1 ± 4.9</del>	<del>75.6 ± 14.5</del>	<del>66.7 ± 13.1</del>	<del>70.9 ± 15.7</del>	<del>70.4 ± 8.0</del>

# How accurately do clicks correspond to explicit judgment of a document?

- Accuracy against manual relevance judgment

Explicit Feedback Data Strategy	Pages Phase II all
Inter-Judge Agreement	86.4
Click > Skip Above	78.2 $\pm$ 5.6
Last Click > Skip Above	80.9 $\pm$ 5.1
<del>Click &gt; Earlier Click</del>	<del>64.3 <math>\pm</math> 15.4</del>
Click > Skip Previous	80.7 $\pm$ 9.6
<del>Click &gt; No Click Next</del>	<del>67.4 <math>\pm</math> 8.2</del>

# What do we get from this user study?

- Clicks are influenced by the relevance of results
  - Biased by the trust over rank positions
- Clicks as relative preference statement is more accurate
  - Several heuristics to generate the preference pairs

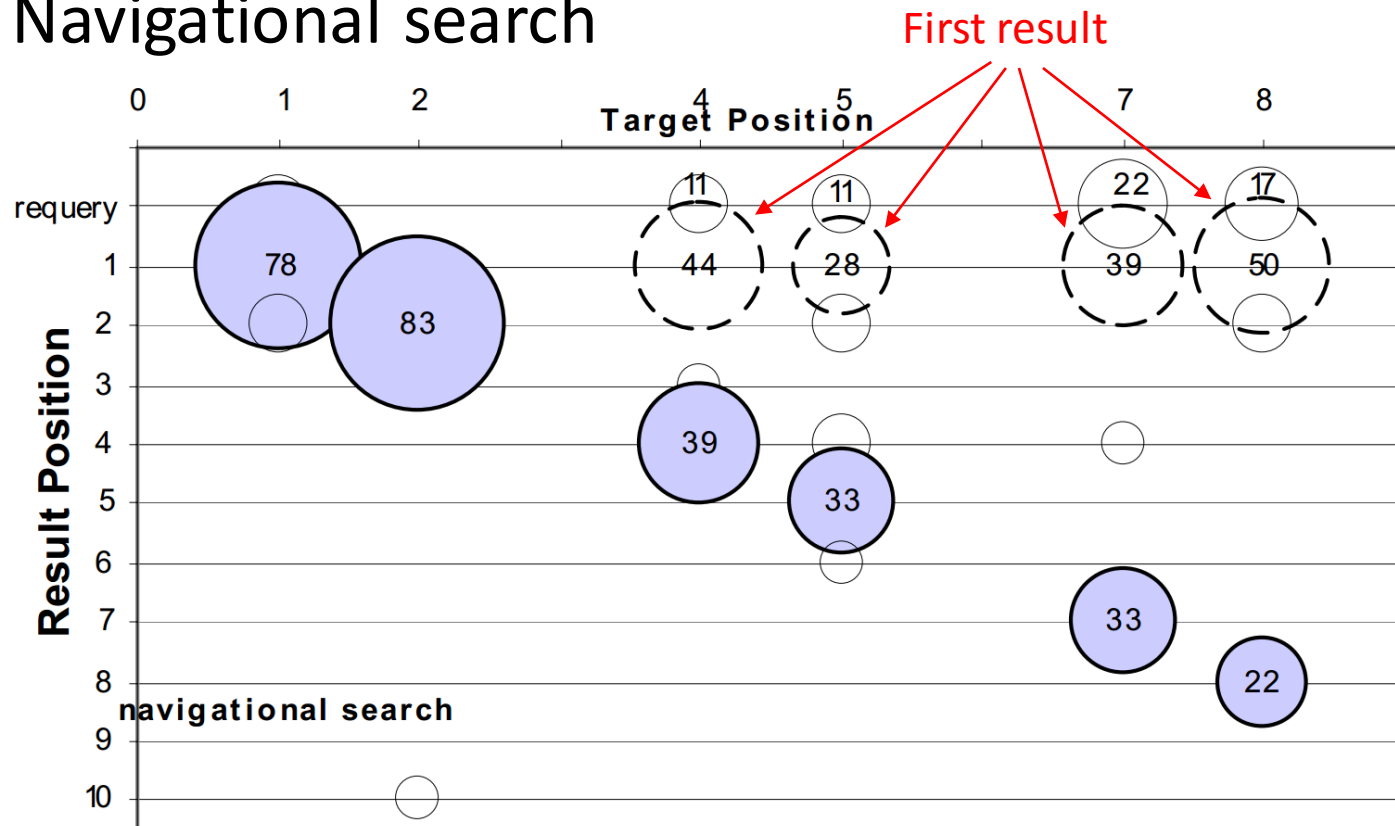


# How to utilize such preference pairs?

- Pairwise learning to rank algorithms
  - Will be covered later

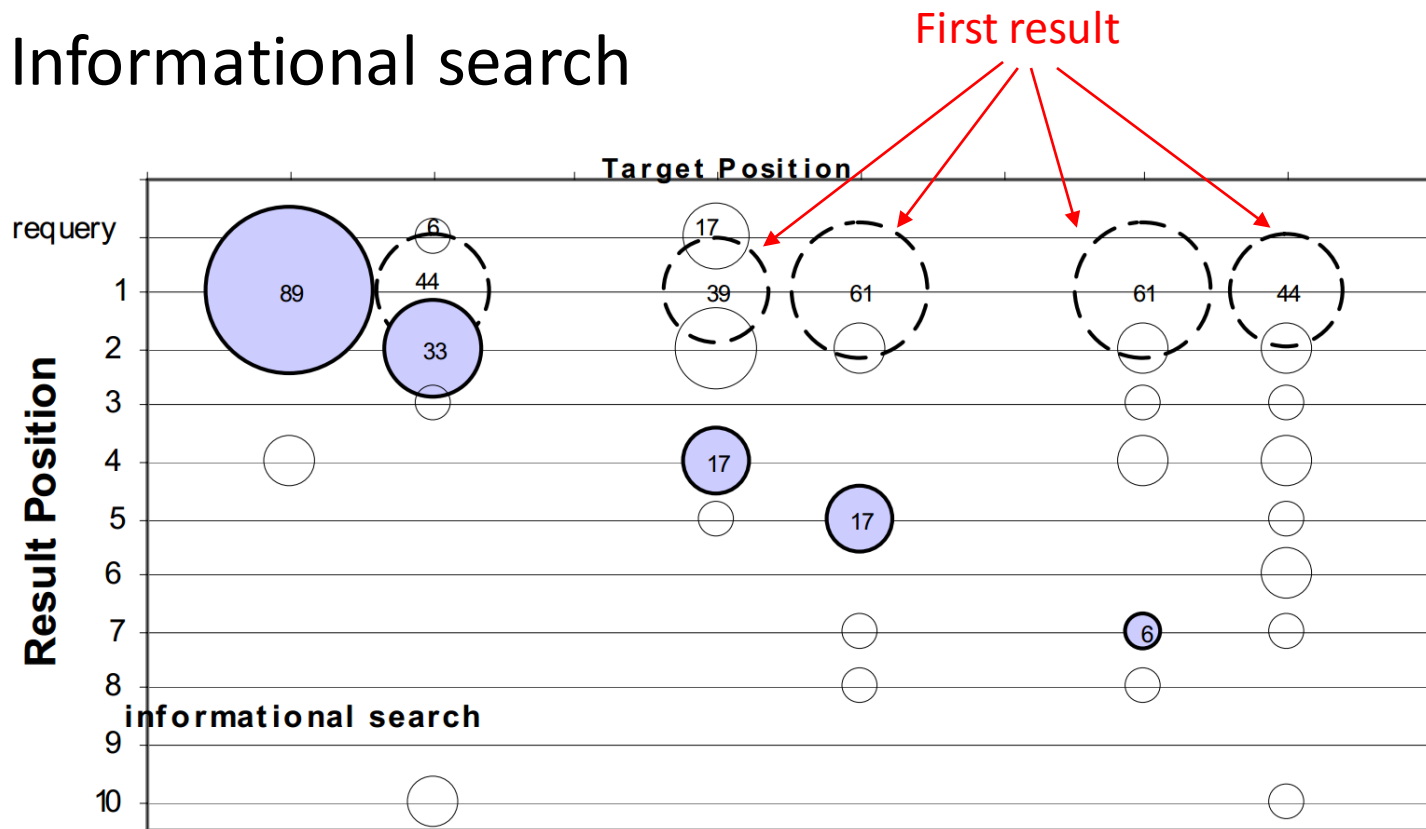
# An eye tracking study of the effect of target rank on web search [Guan CHI'07]

- Break down of users' click accuracy
  - Navigational search



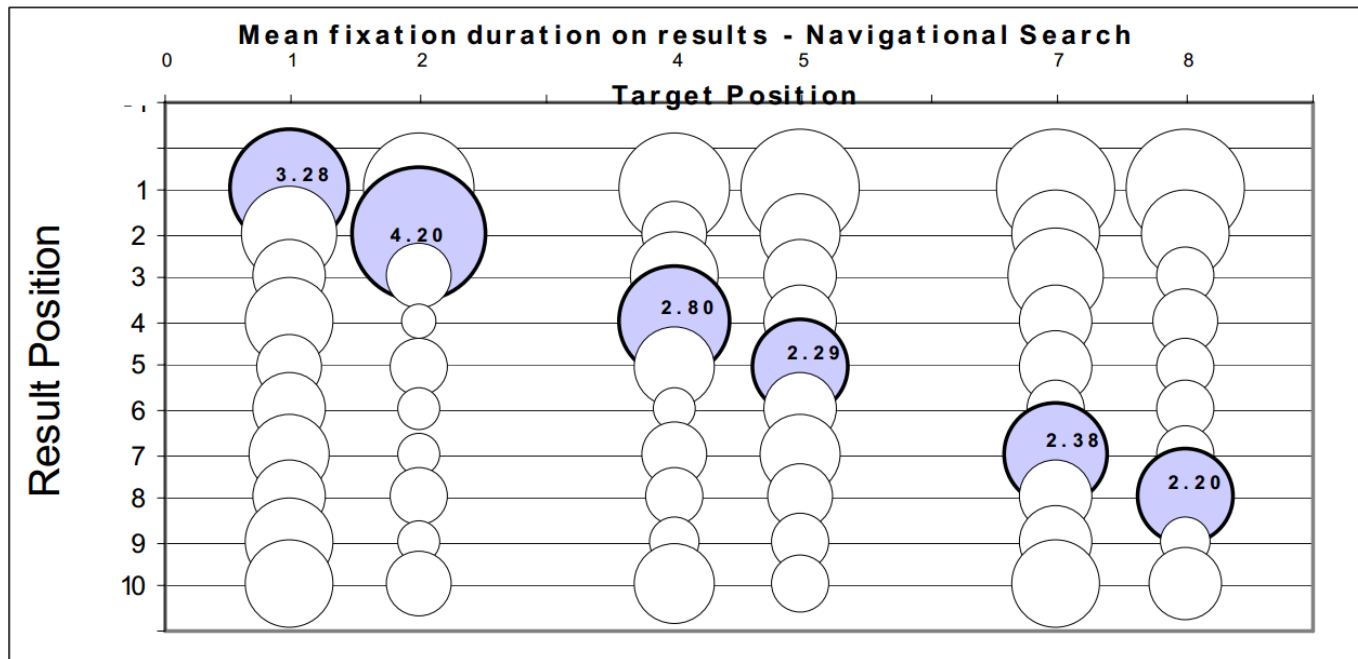
# An eye tracking study of the effect of target rank on web search [Guan CHI'07]

- Break down of users' click accuracy
  - Informational search



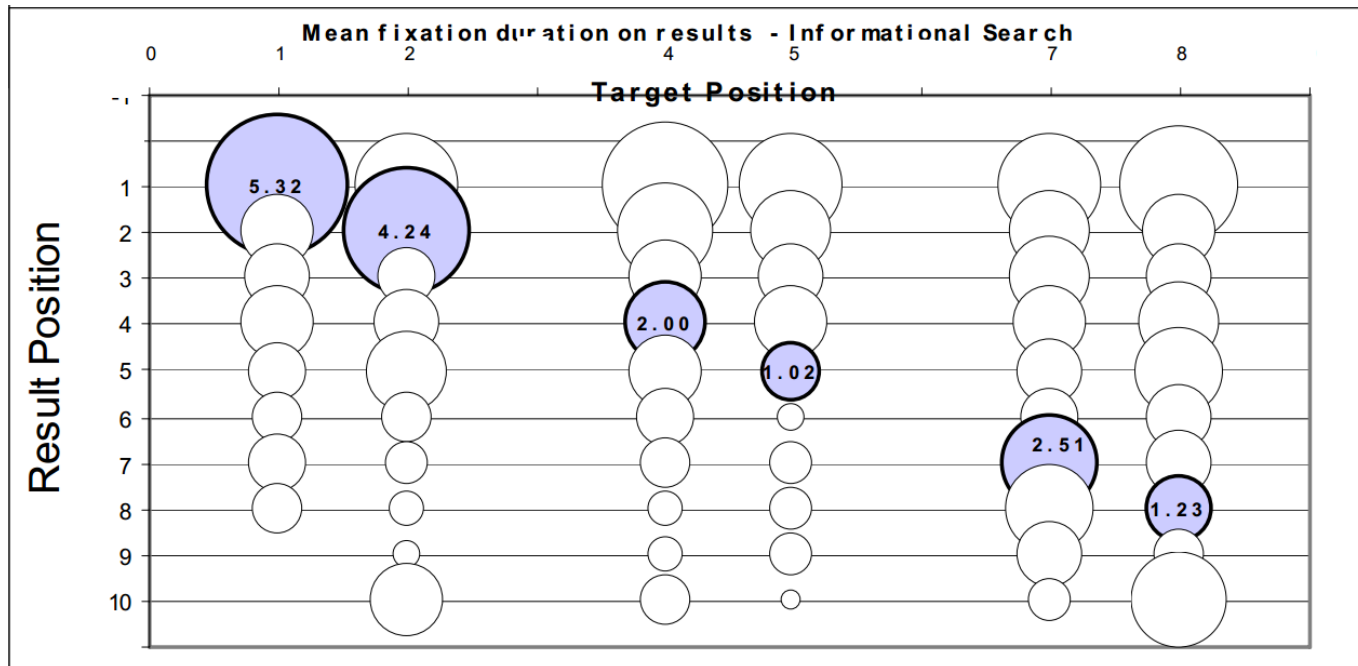
# Users failed to recognize the target because they did not read it!

- Navigational search



# Users did not click because they did not read the results!

- Informational search



# Predicting clicks: estimating the click-through rate for new ads [Richardson WWW'07]

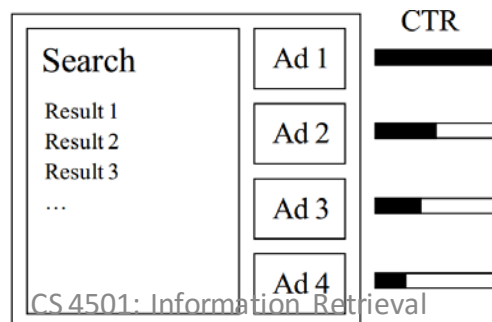
- To maximize ad revenue

$$- E_{ad}[Revenue] = \sum_{ad} p(click|ad)CPC(ad)$$

Estimated click-through rate

*Cost per click: basic business model in search engines*

- Position-bias is also true in online ads
  - Observed low CTR is not just because of ads' quality, but also their display positions!



# Combat position-bias by explicitly modeling it

- Being clicked is related to its quality and position

$$\begin{aligned} - p(click|ad, pos) &= p(click|ad, pos, seen)p(seen|pos) \\ &= \underline{p(click|ad, seen)} \underline{p(seen|pos)} \end{aligned}$$

Calibrated CTR for ads ranking

Discounting factor

$$- p(click = 1|ad, seen = 0) = 0$$

$$- p(click = 1|ad, seen = 1) = \frac{1}{1 + \exp(-w^t f_{ad})}$$

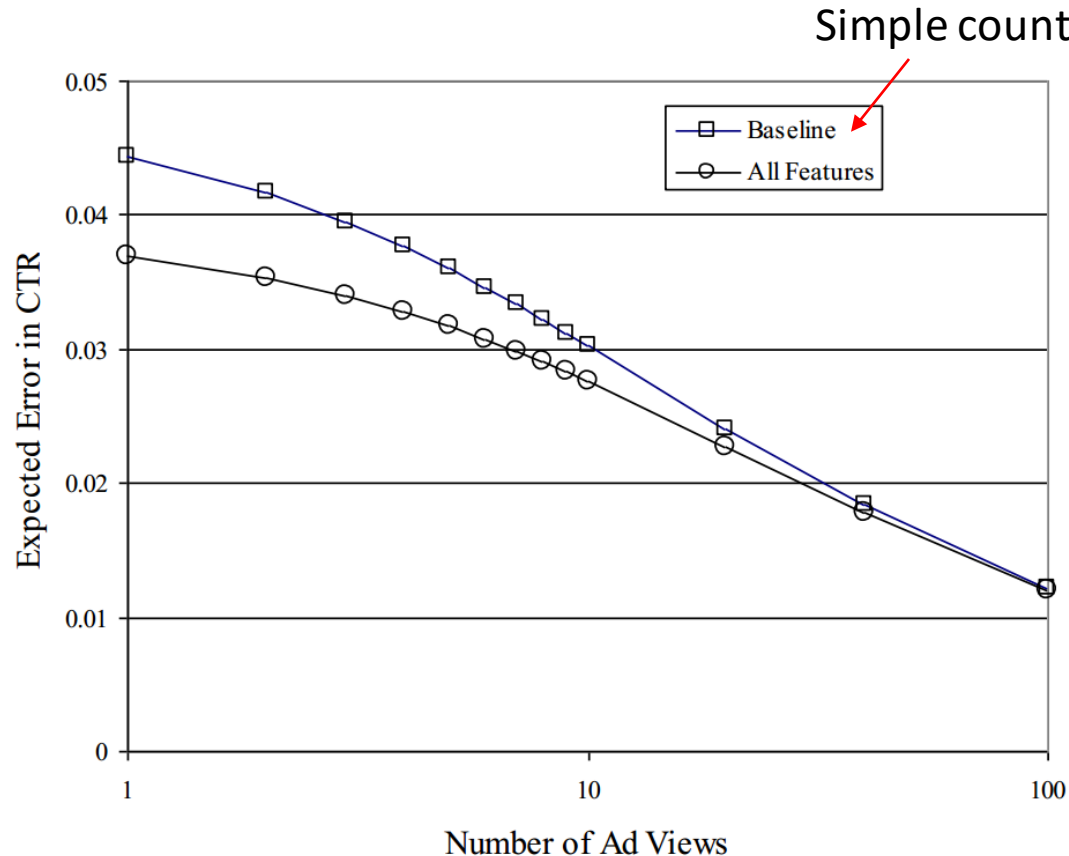
Logistic regression by features of the ad

# Parameter estimation

- Discounting factor
  - Approximation: positions being clicked must be seen already
    - $p(\textit{seen}|\textit{pos}) \propto \# \textit{clicks\_at\_pos}$
- Calibrated CTR
  - Maximum likelihood for  $w$  with historic clicks
    - $\hat{w} = \operatorname{argmax}_w \sum_{ad} \log p(\textit{click}|ad, pos)$



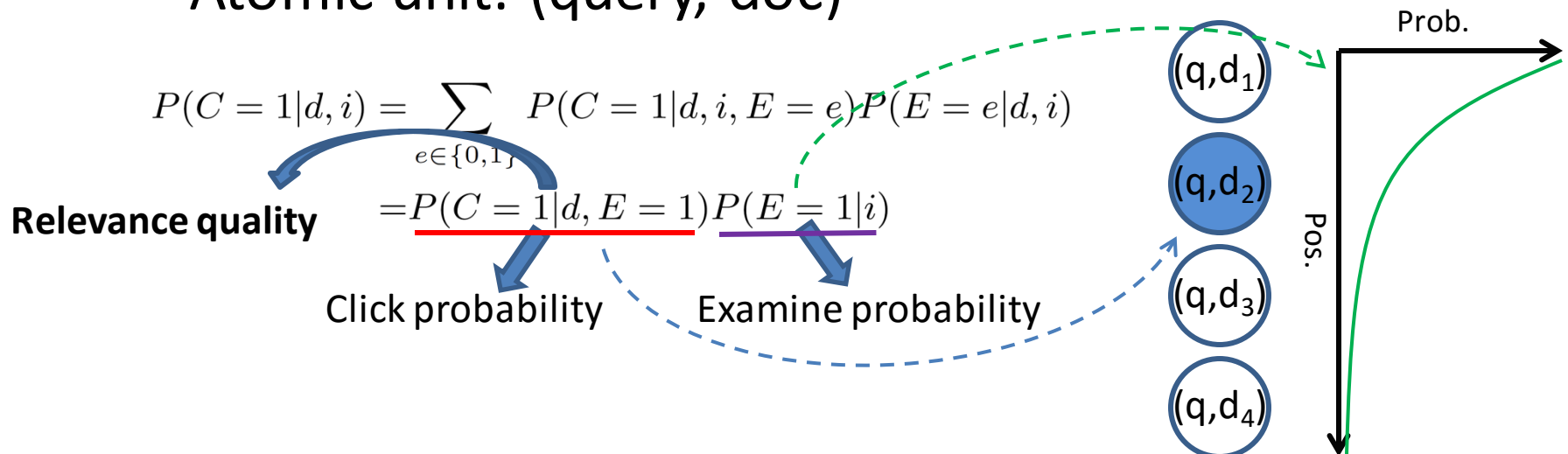
# Calibrated CTR is more accurate for new ads



- Unfortunately, their evaluation criterion is still based on biased clicks in testing set

# Click models

- Decompose relevance-driven clicks from position-driven clicks
  - Examine: user reads the displayed result
  - Click: user clicks on the displayed result
  - Atomic unit: (query, doc)



# Cascade Model [Craswell et al. WSDM'08]

- Sequential browsing assumption
  - At each position decides whether to move on
    - $p(C_i = 1) = p(R_i = 1) \prod_{j=1}^{i-1} (1 - p(R_j = 1))$
    - Assuming  $R_i = 1 \rightarrow C_i = 1$
  - Only one click is allowed on each search result page

*Kind of “Click > Skip Above”?*

# User Browsing Model [Dupret et al. SIGIR'08]

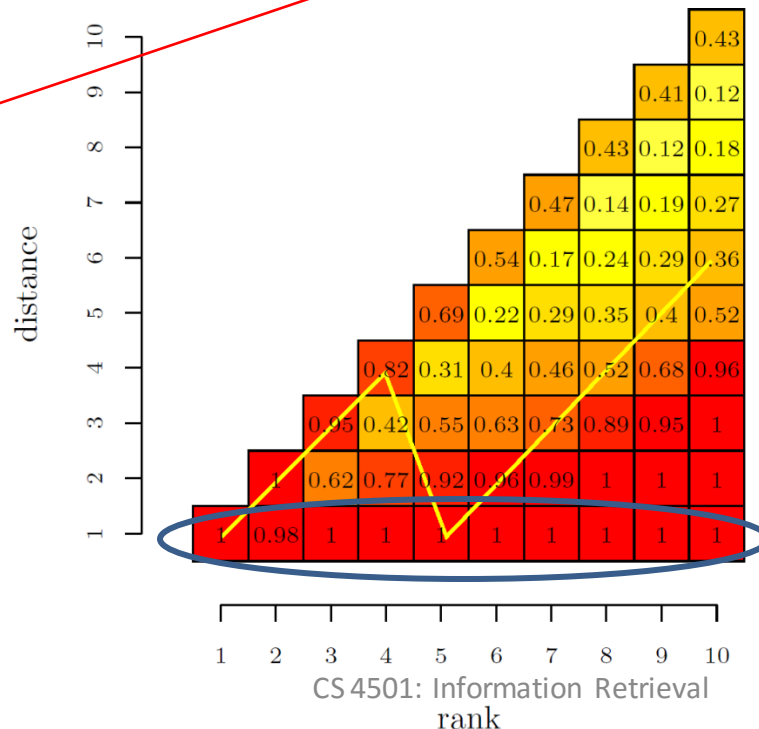
- Examination depends on distance to the last click

*Attractiveness, determined by query and URL*

*Examination, determined by position and distance to last click*

$$- P(c = 1 | u, q, r, d) = \alpha_{uq} \gamma_{rd}$$

*EM for parameter estimation*

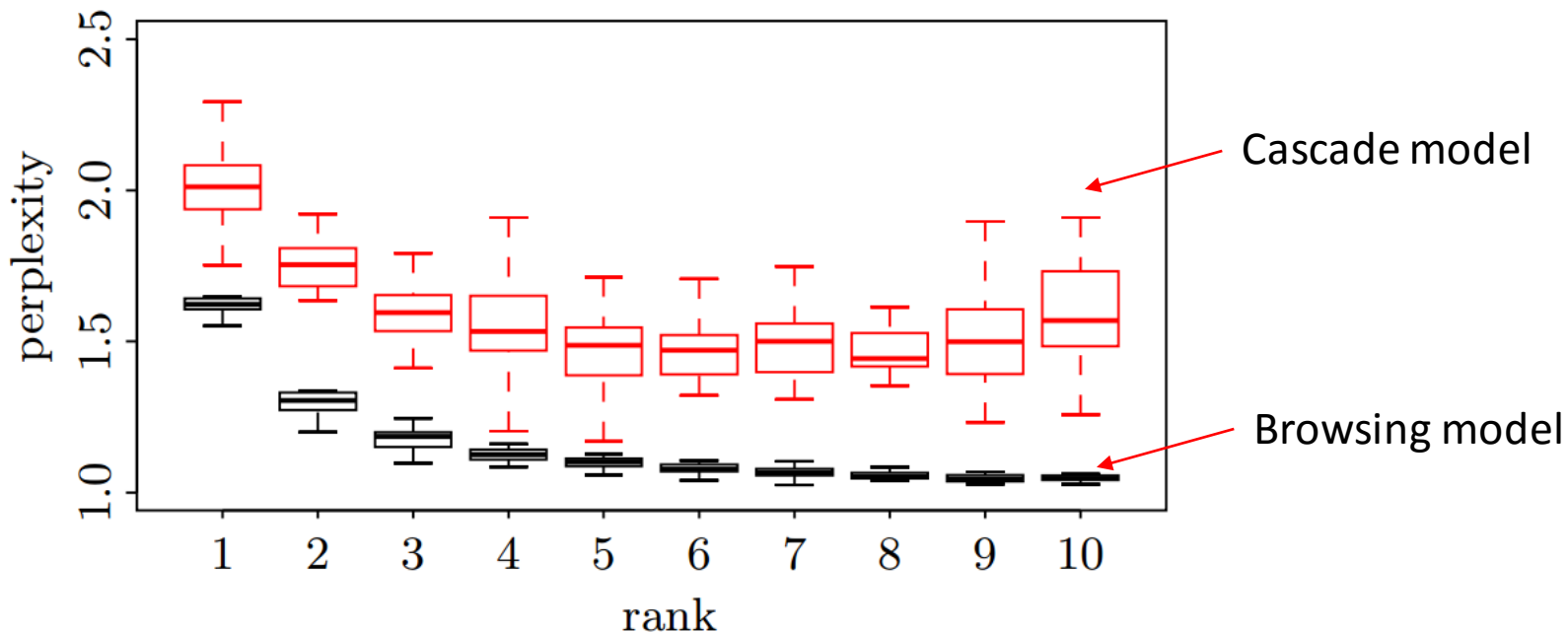


From absolute discount to relative discount

*Kind of "Click > Skip Next" + "Click > Skip Above"?*

# More accurate prediction of clicks

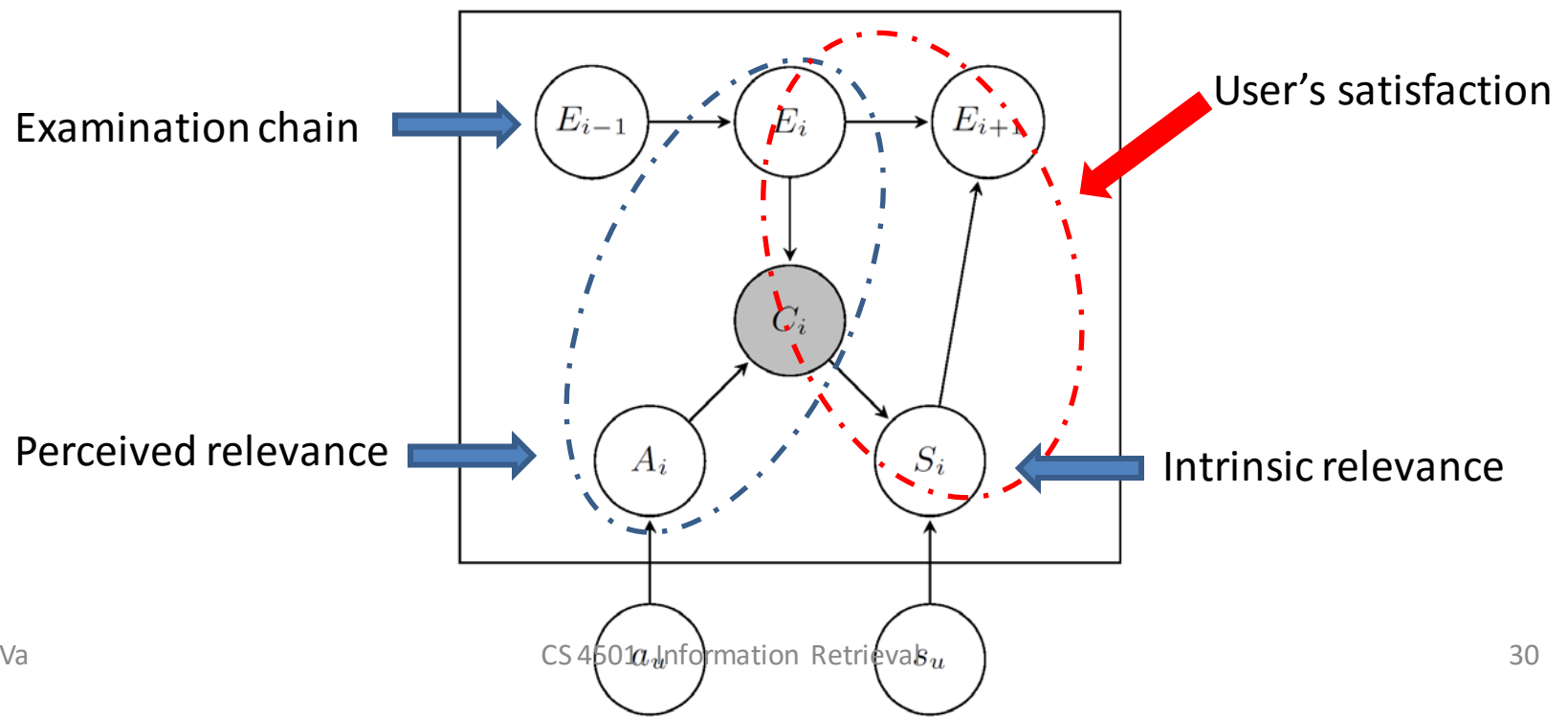
- Perplexity – randomness of prediction



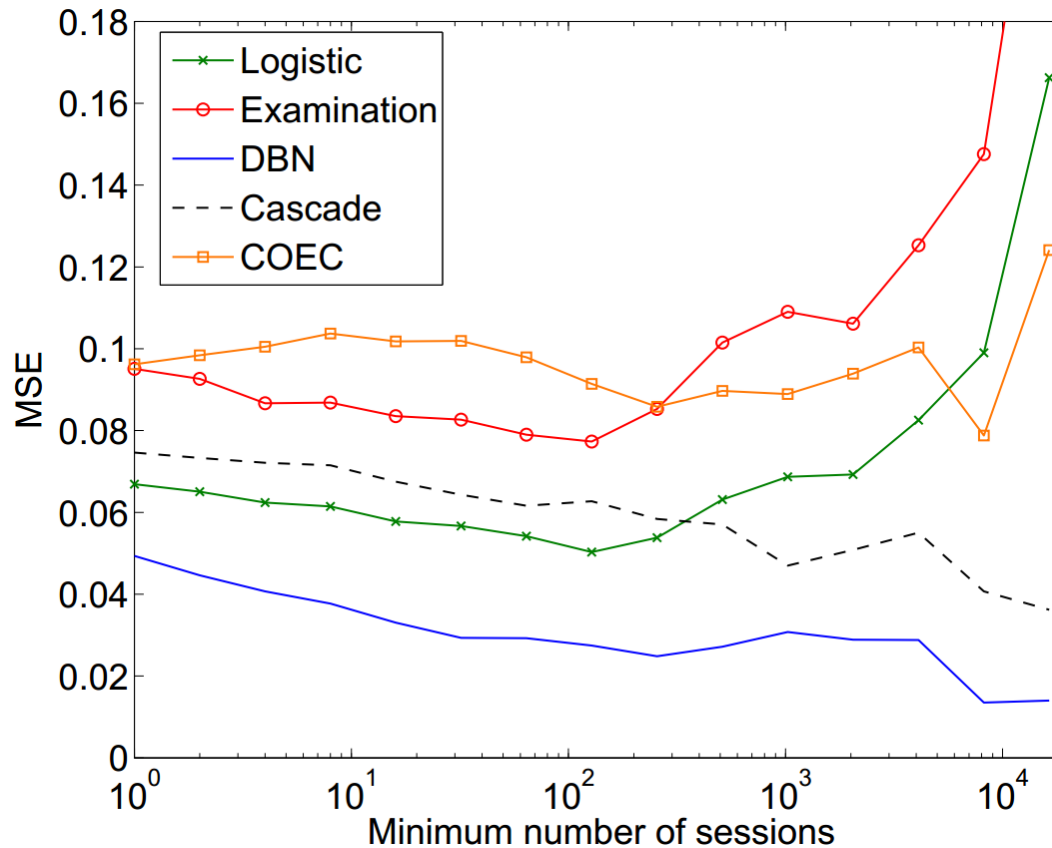
# Dynamic Bayesian Model [Chapelle et al. WWW'09]

- A cascade model
  - Relevance quality:

$$P(E_{i+1} = 1 | E_i = 1, S_i = 0) = \gamma$$
$$S_i = 1 \Rightarrow E_{i+1} = 0$$



# Accuracy in predicting CTR



# Revisit User Click Behaviors

**YAHOO! NEWS**

Search:

10,181 results

WEB IMAGES VIDEO **NEWS** MORE ▾

**Also try:** [jeremy lin knicks](#), [jeremy lin shoe](#), [More...](#)

Sort Results by: [Relevance](#) | [Time](#)

**Knicks-Pistons Linsanity Live Stream: Where To Watch Jeremy Lin For Free Online**  
Can **Jeremy Lin** and the New York Knicks get back to their winning ways under new coach Mike Woodson? We'll find out Saturday night as they host the lowly Detroit Pistons at Madison Square Garden.  
International Business Times - Mar 24 02:08pm

**Bats: Less of a Frenzy for Lin in Toronto This Time**  
**Jeremy Lin's** second trip to Toronto with the Knicks has been a less intense than his first.  
New York Times - Mar 23 10:58am

[Off the Dribble: Less of a Frenzy for Lin in Toronto This Time...](#) - New York Times  
[Less of a Frenzy for Lin in Toronto This Time](#) - New York Times  
[Grange on NBA: The life of Jeremy Lin](#) - Sportsnet.ca  
[all 142 news articles...](#)

**Jeremy Lin Update: Knicks Win, Lin Still Plagued by Turnovers | The Harvard Crimson**  
**Jeremy Lin** and the Knicks got back on track at Madison Square Garden on Saturday against the Detroit Pistons. The Pistons stayed within close range for most of the first three quarters, but New York pulled away in the fourth to cruise to a 101-79 victory.  
The Harvard Crimson - Mar 24 11:35pm

[New Coach Runs Similar Show, and Lin Stays in Picture](#) - New York Times  
[Jeremy Lin's Brand Thriving With Knicks New Coach](#) - Forbes  
[NBA: Rajon Rondo prepared for showdown with Jer...](#) - Worcester Telegram & Gazette  
[all 14 news articles...](#)

**FILTER BY SOURCE**  
**All sources**  
Yahoo! Sports (425)  
Sports Illustrated (203)  
Yahoo! News (144)  
New York Times (126)  
More...

**FILTER BY TIME**  
**Any time**  
Past hour  
Past day  
Past week

**RELATED CONCEPTS**  
golden state warrior...  
rex walters  
harvard basketball



Match my query?

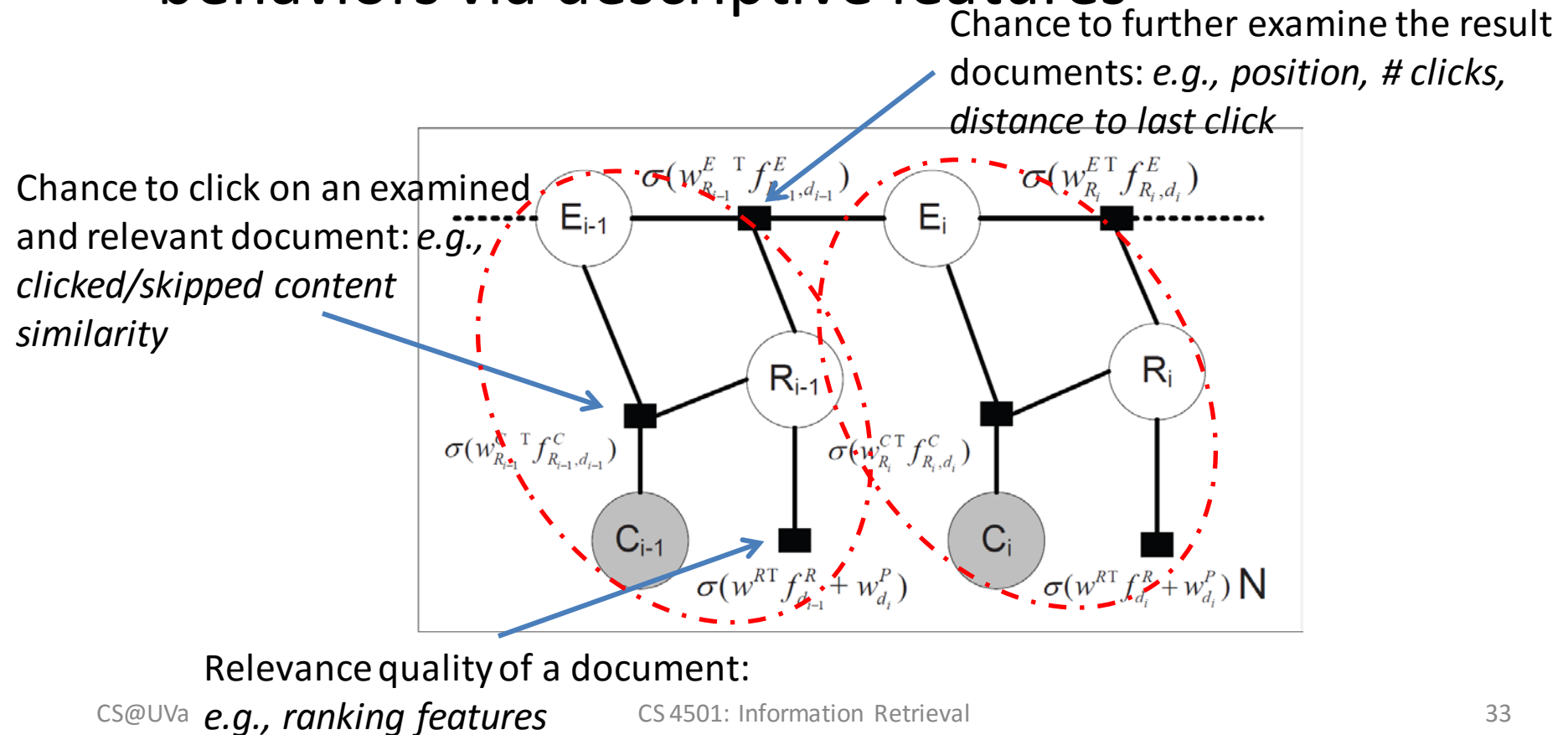
Redundant doc?

Shall I move on?



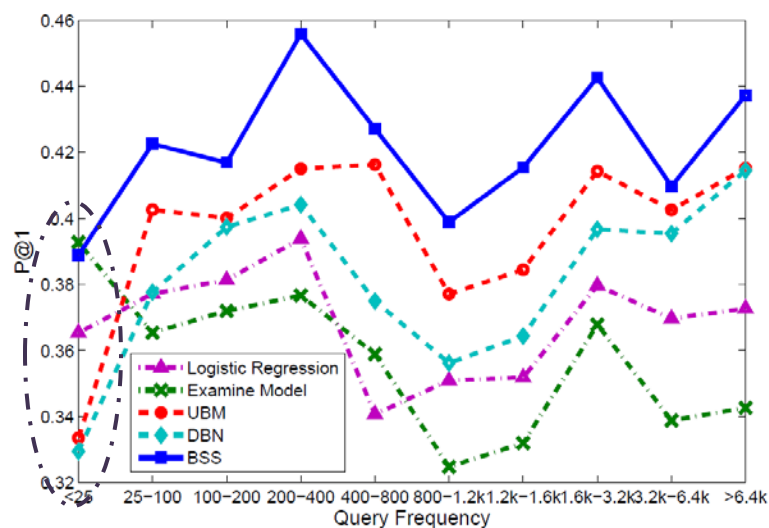
# Content-Aware Click Modeling [Wang et al. WWW'12]

- Encode dependency within user browsing behaviors via descriptive features

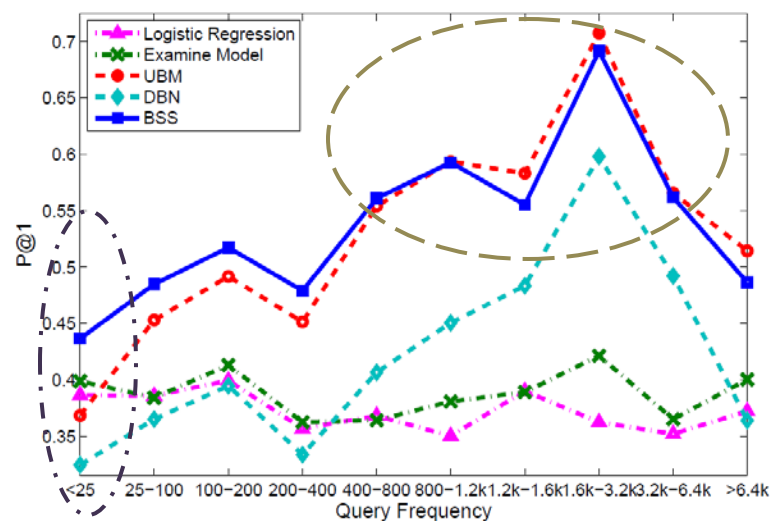


# Quality of relevance modeling

- Estimated relevance for ranking



(a) P@1 ranking performance under different query frequency categories on the random bucket click set



(b) P@1 ranking performance under different query frequency categories on the normal click set

# Understanding user behaviors

- Analyzing factors affecting user clicks

$f^R$	age	authority	title match	abs. match	body match
$w^R$	-0.839	0.007	0.098	0.167	0.020
$f^C$	pos	# click	dis. to last click	query length	bias
$w_{R=0}^C$	-1.133	-0.351	-0.445	-3.659	-4.654
$w_{R=1}^C$	0.149	0.335	0.415	3.707	4.405
$f^E$	pos	# click	dis. to last click	avg cont. sim.	bias
$w_{R=0}^E$	1.807	-0.418	0.684	2.947	5.325
$w_{R=1}^E$	-1.381	0.665	-3.395	-2.237	3.266

# What you should know

- Clicks as implicit relevance feedback
- Positional bias
- Heuristics for generating pairwise preferences
- Assumptions and modeling approaches for click models