

# Retroactive Answering of Search Queries

Beverly Yang

Glen Jeh

*Google*

# Personalization

- Provide more relevant services to specific user
  - Based on Search History
- Usually operates at a high level
  - e.g., Re-order search results based on a user's general preferences
  - Classic example:
    - User likes cars
    - Query: "jaguar"
- Why not focus on known, specific needs?
  - User likes cars
  - User is interested in the 2006 Honda Civic

# The QSR system

- QSR = **Q**uery-**S**pecific (Web) **R**ecommendations
- Alerts user when *interesting new results* to selected *previous queries* have appeared
- Example
  - Query: “britney spears concert san francisco”
  - No good results at time of query (*Britney not on tour*)
  - One month later, new results (*Britney is coming to town!*)
  - User is automatically notified

## Recommended Links

- **[Britney Spears at Warfield 2/16/06](#)**

**Britney Spears** is to appear at the Warfield in a concert to promote her latest album...

(for query "[britney spears concert sf](#)" on 10/12/05)

- **[Google Reader](#)**

**Reader** makes it easier to keep up with your ever expanding reading-list...

(for query "[rss reader](#)" on 7/06/07)

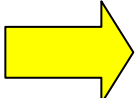
- **[McAfee's Virex does not work with Tiger](#)**

McAfee's anti-virus software, **Virex**, does not work properly on OS X (**Tiger**)...

(for query "[virex tiger](#)" on 9/13/05)

- Query treated as ***standing query***
- New results are ***web page recommendations***

# Challenges

- How do we identify queries representing standing interests?
    - Explicit – Web Alerts. But no one does this
    - Want to automatically identify
  - How do we identify interesting new results?
    - Web alerts: change in top 10. But that's not good enough
-  Focus: addressing these two challenges

# Outline

- Introduction

- ➔ Basic QSR Architecture

- Identifying Standing Interests

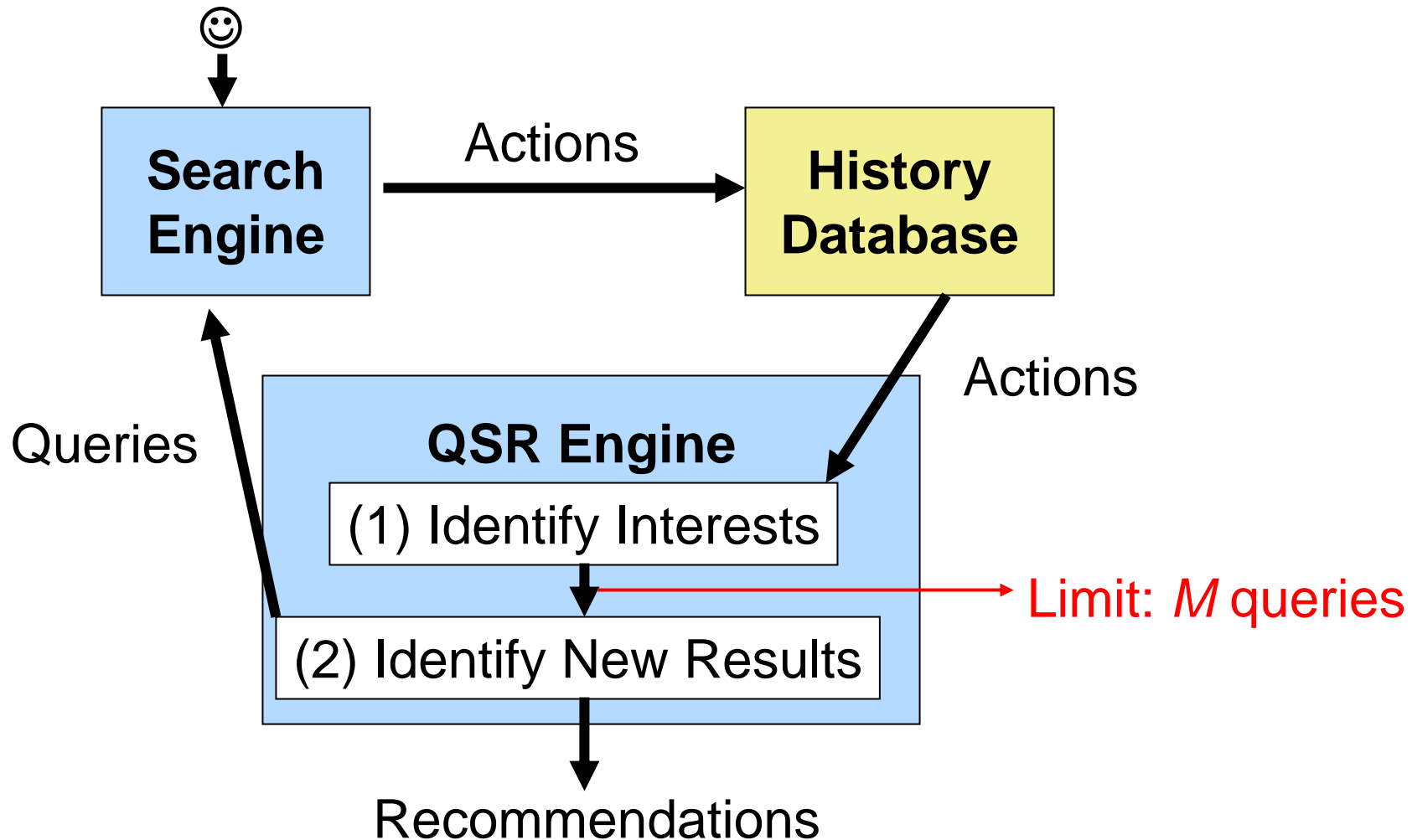
- Determining Interesting Results

**Heuristic**

- User Study Setup

- Results

# Architecture



# Related Work

## ■ Identifying User Goal

- [Rose & Levinson 2004], [Lee, Liu & Cho 2005]
- At a higher, more general level

## ■ Identifying Satisfaction

- [Fox, et. al. 2005]
- One component of identifying standing interest
- Specific model, holistic rather than considering strength and characteristics of each signal

## ■ Recommendation Systems

- Too many to list!

# Outline

- Introduction
- Basic QSR Architecture
- ➔ Identifying Standing Interests
  - Determining Interesting Results
  - User Study Setup
  - Results

# Definition

- A user has a *standing interest* in a query if she *would be interested in seeing new interesting results*
- Factors to consider:
  - Prior fulfillment/Satisfaction
  - Query interest level
  - Duration of need or interest

# Example

→ QUERY (8s) -- html encode java

→ ■ RESULTCLICK (91s) - 2. <http://www.java2html.de/ja...>

→ ■ RESULTCLICK (247s) - 1. <http://www.javapractices/...>

→ ■ RESULTCLICK (12s) - 8. <http://www.trialfiles.com/...>

→ ■ NEXTPAGE (5s) - start = 10

→ □ RESULTCLICK (1019s) - 12. <http://forum.java.su...>

→ □ REFINEMENT (21s) - html encode java utility

→ ■ RESULTCLICK (32s) - 7. <http://www.javapracti...>

→ □ NEXTPAGE (8s) - start = 10

→ ■ NEXTPAGE (30s) - start = 20

# Example

QUERY (8s) -- html encode java

- RESULTCLICK (91s) - 2. <http://www.java2html.de/ja...>
- RESULTCLICK (247s) - 1. <http://www.javapractices/...>
- RESULTCLICK (12s) - 8. <http://www.trialfiles.com/...>
- NEXTPAGE (5s) - start = 10
  - RESULTCLICK (1019s) - 12. <http://forum.java.su...>
  - REFINEMENT (21s) - html encode java utility
    - RESULTCLICK (32s) - 7. <http://www.javapracti...>
      - NEXTPAGE (8s) - start = 10
        - NEXTPAGE (30s) - start = 20

# Signals

- Good ones:

- # terms
- # clicks, # refinements
- History match
- Repeated non-navigational

- Other:

- Session duration, number of long clicks, etc.

# Outline

- Introduction
- Basic QSR Architecture
- Identifying Standing Interests
- ➔ Determining Interesting Results
  - User Study Setup
  - Results

# Web Alerts

- Heuristic: new result in top 10
- Query: “beverly yang”
  - Alert 10/16/2005:  
<http://someblog.com/journal/images/04/0505/>
  - Seen before through a web search
  - Poor quality page
  - Alert repeated due to ranking fluctuations

# QSR Example

Query: "rss reader"

(not real)

Rank	URL	PR score	Seen
1	www.rssreader.com	3.93	Yes
2	blogspot.com/rss/readers	3.19	Yes
3	www.feedreader.com	3.23	Yes
4	<b>www.google.com/reader</b>	<b>2.74</b>	<b>No</b>
5	www.bradsoft.com	2.80	Yes
6	www.bloglines.com	2.84	Yes
7	www.pluck.com	2.63	Yes
8	sage.mozdev.org	2.56	Yes
9	www.sharpreader.net	2.61	Yes

# Signals

- Good ones:
  - History presence
  - Rank (inverse!)
  - Popularity and relevance (PR) scores
  - Above dropoff
    - PR scores of a few results are much higher than PR scores of the rest
  - Content match
- Other:
  - Days elapsed since query, sole changed

# Outline

- Introduction
- Basic QSR Architecture
- Identifying Standing Interests
- Determining Interesting Results
- ➔ User Study Setup
- Results

# Overview

- Human subjects: Google Search History users
- Purpose:
  - Demonstrate *promise* of system effectiveness
  - Verify intuitions behind heuristics
- Many disclaimers:
  - Study conducted internally!!!
  - 18 subjects!!!
  - Only a fraction of queries in each subject's history!!!
  - Need additional studies over broader populations to generalize results

# Questionnaire

1) Did you find a satisfactory answer for your query?

Yes   Somewhat   No   Can't  
Remember

2) How interested would you be in seeing a new high-quality result?

Very   Somewhat   Vaguely   Not

3) How long would this interest last for?

Ongoing   Month   Week   Now

4) How good would you rate the quality of this result?

Excellent   Good   Fair   Poor

QUERY (8s) -- html encode java

■RESULTCLICK (91s) - 2. http://

■RESULTCLICK (247s) - 1. http://

■RESULTCLICK (12s) - 8. http://

■NEXTPAGE (5s) - start = 10

□RESULTCLICK (1019s) - 12.

□REFINEMENT (21s) - html er

■RESULTCLICK (32s) - 7.

□NEXTPAGE (8s) - sta

■NEXTPAGE (30s) -

# Outline

- Introduction
- Basic QSR Architecture
- Identifying Standing Interests
- Determining Interesting Results
- User Study Setup

 Results



# Questions

- Is there a need for automatic detection of standing interests?
- Which signals are useful for indicating standing interest in a query session?
- Which signals are useful for indicating quality of recommendations?

# Is there a need?

How many Web alerts have you ever registered?

**0: 73%**    **1: 20%**    **2: 7%**    **>2: 0%**

Of the queries marked “very” or “somewhat” interesting (154 total), how many have you registered?

**0: 100%**

# Effectiveness of Signals

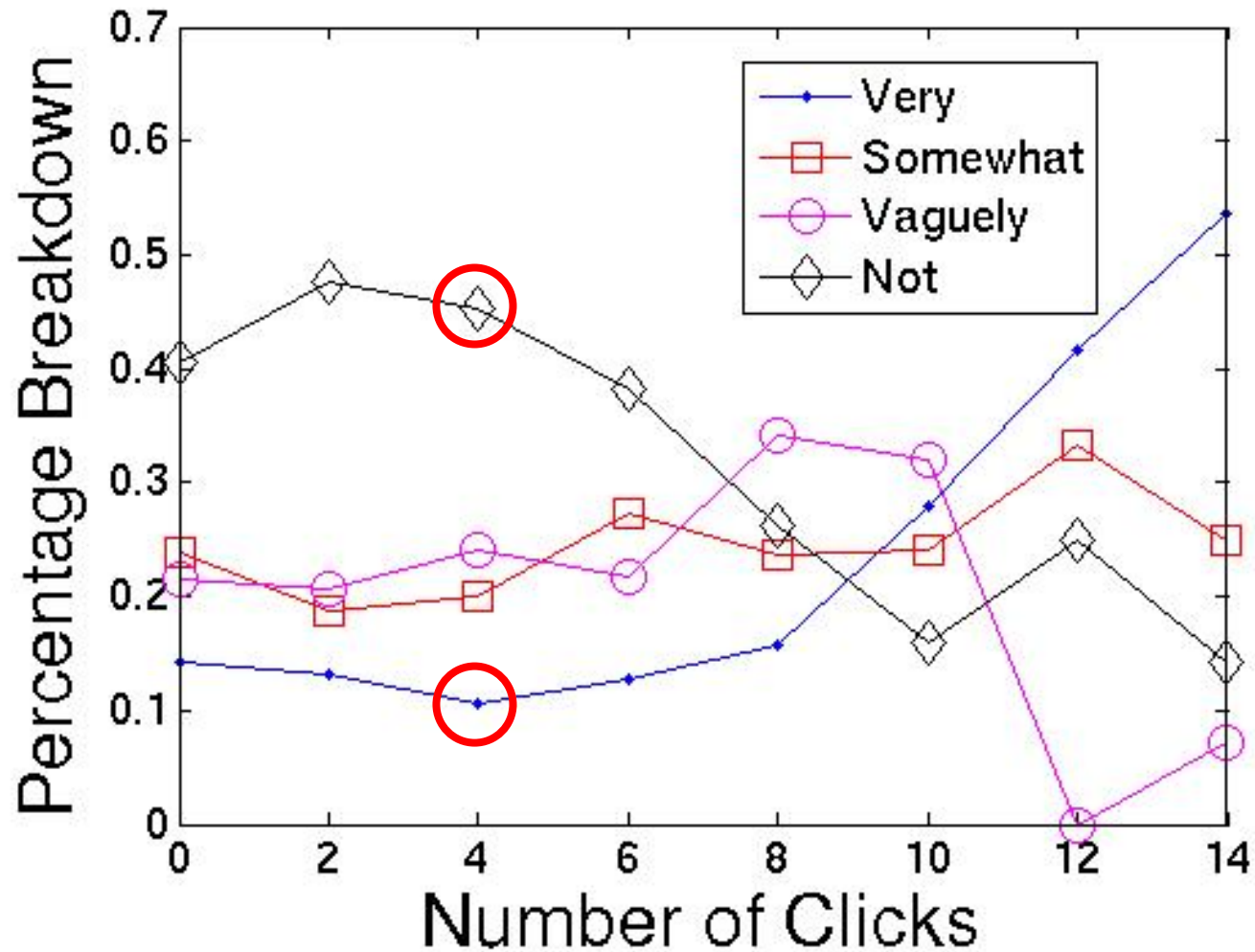
## ■ Standing interests

- # clicks (> 8)
- # refinements (> 3)
- History match
- Also: repeated non-navigational, # terms (> 2)

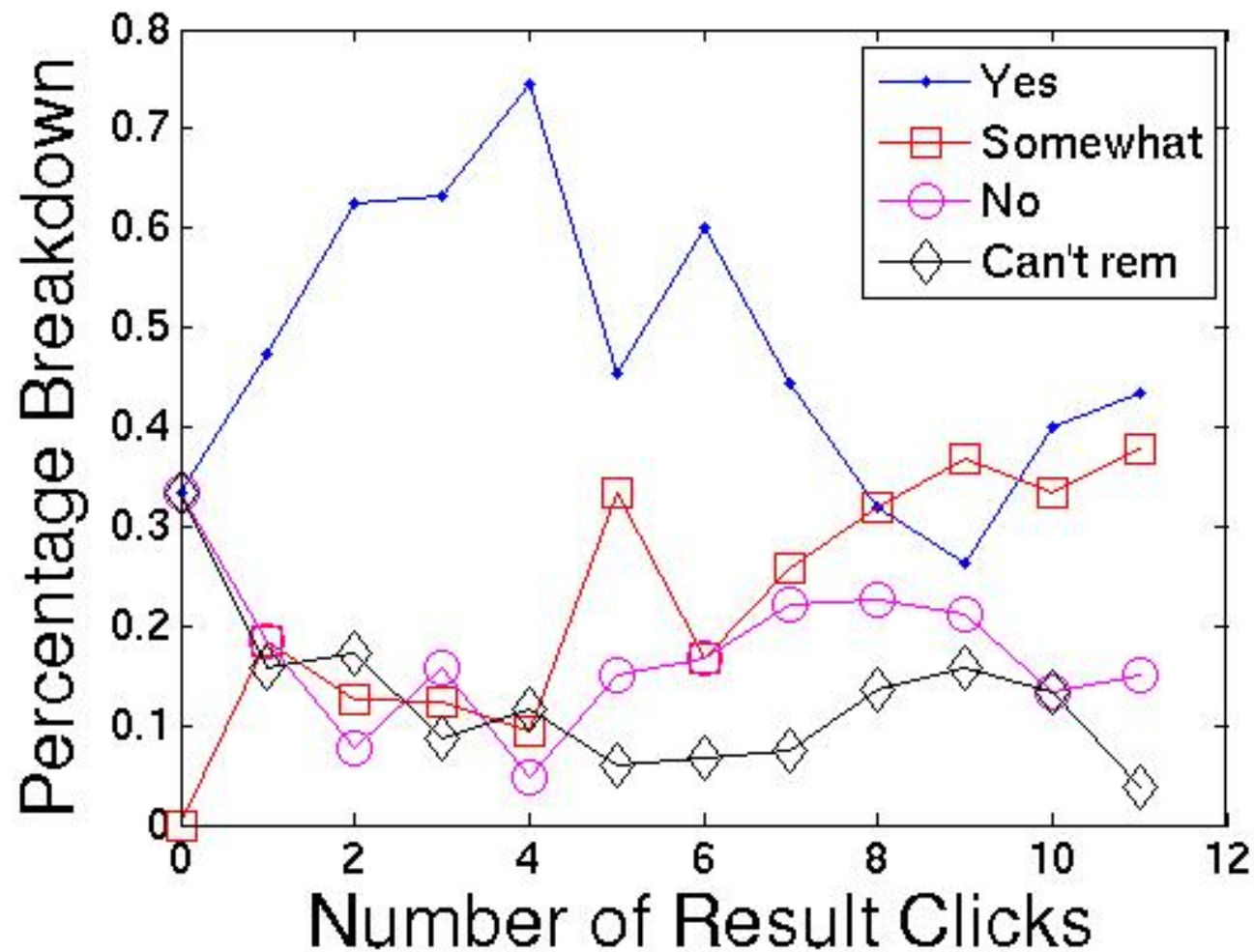
## ■ Quality Results

- PR score (high)
- Rank (low!!)
- Above Dropoff

# Standing Interest



# Prior Fulfillment



# Interest Score

- Goal: capture the relative standing interest a user has in a query session

```
iscore =  
    a * log(# clicks + # refinements) +  
    b * log(# repetitions) +  
    c * (history match score)
```

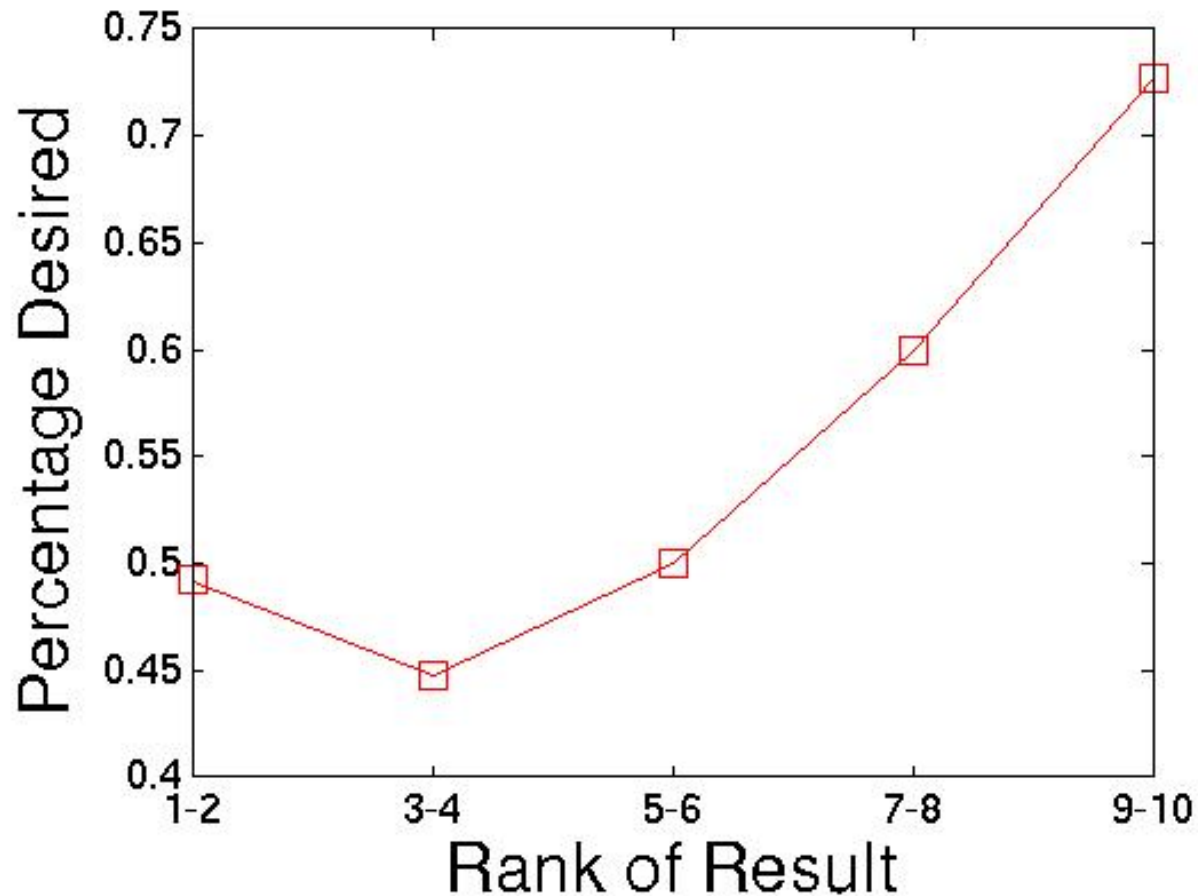
- Select query sessions with  $\text{iscore} > t$

# Effectiveness of iscore

- Standing Interest:
  - Sessions for which user is *somewhat* or *very* interested in seeing further results
- Select query sessions with  $\text{iscore} > t$ 
  - Vary  $t$  to get precision/recall tradeoff
    - 90% precision, 11% recall
    - 69% precision, 28% recall
  - Compare: 28% precision by random selection
  - Recall – percentage of standing interest sessions *that appeared in the survey*

# Quality of Results

“Desired”: marked in survey as “good” or “excellent”



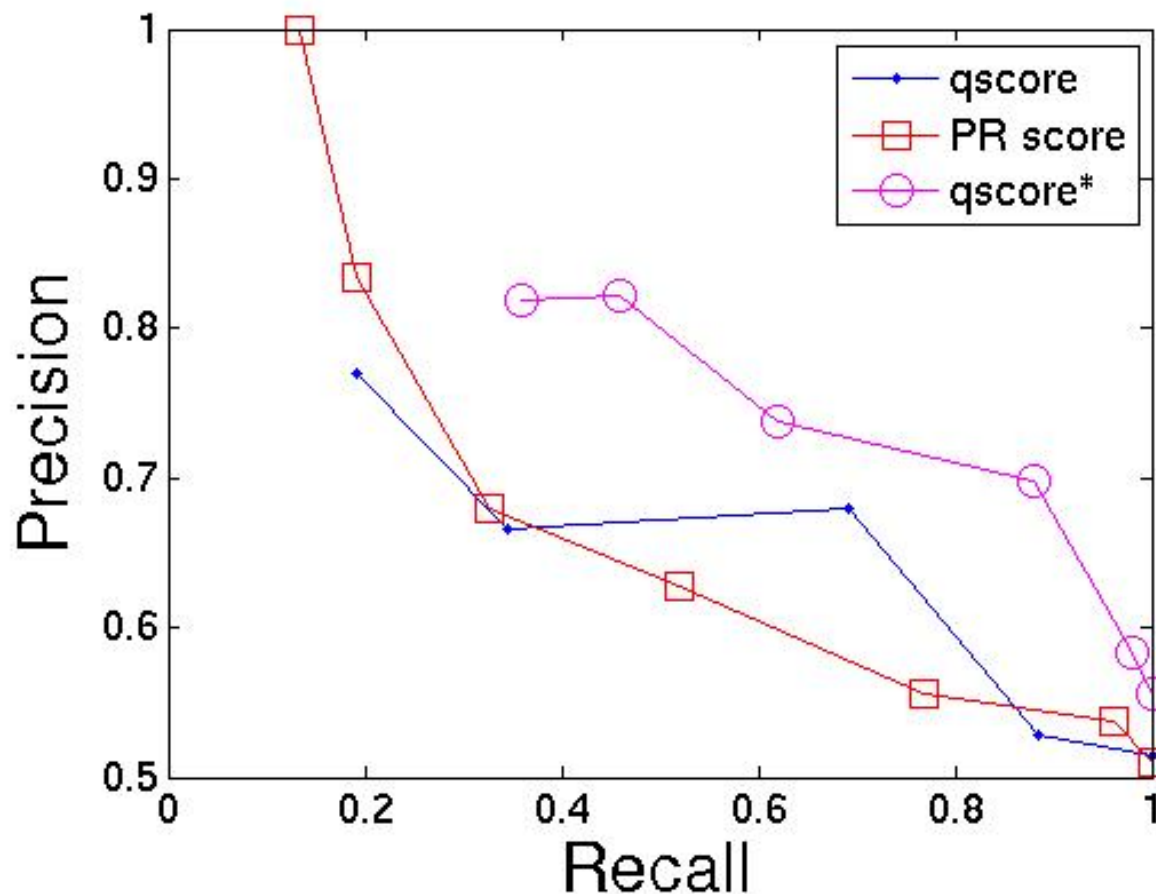
# Quality Score

- Goal: capture relative quality of recommendation
  - Apply score after result has passed a number of boolean filters

$$\text{qscore} = a * \text{PR score} + b' * \frac{1}{\text{rank}}$$
$$c * \text{topic match}$$

# Effectiveness of qscore

Select URLs with score  $> t$



Recall:  
Percentage of  
URLs in the  
survey marked  
as “good” or  
“excellent”

# Conclusion

- Huge gap:
  - Users' standing interests/needs
  - Existing technology to address them
- QSR: Retroactively answer search queries
  - Automatic identification of standing interests and unfulfilled needs
  - Identification of interesting new results
- Future work
  - Broader studies
  - Feedback loop



Thank you!

# Selecting Sessions

- Users may have thousands of queries
  - Must only show 30
  - Try to include a mix of positive and negative sessions
  - Prevents us from gathering some stats
- Process
  - Filter special-purpose queries (e.g., maps)
  - Filter sessions with 1-2 actions
  - Rank sessions by iscore
    - Take top 15 sessions by score
    - Take 15 randomly chosen sessions



# Selecting Recommendations

- Tried to only show good recommendations
  - Assumption: some will be bad
- Process
  - Only consider sessions with history presence
  - Only consider results in top 10 (Google)
  - Must pass at least 2 boolean signals
  - Select top 50% according to qscore



# 3<sup>rd</sup>-Person study

- Not enough recommendations in 1<sup>st</sup>-person study
- Asked subjects to evaluate recommendations made for other users' sessions