



query hardness estimation

Javed Aslam

Virgil Pavlu



why query hardness

- Users can reformulate their query.
- Systems could employ alternate search strategies
- Distributed retrieval systems could more accurately combine their input results



query hardness

- specific to collection searched
 - clarity score = KL distance between query and collection
 - ranking robustness = stability with respect to collection perturbations
- specific to an IR system
 - stability with respect to query perturbations



background



	What makes a query difficult ? [Carmel...]	Predicting query performance [Townsend, Zhou, Croft]	Ranking robustness [Zhou, Croft]	Learning to estimate query difficulty [Yom-Tov...]	Overview of TREC 2004 Robust track [Voorhees]	Query difficulty, robustness [Amati et al]
define Q hardness	Q median AP JURU AP	multinom language model AP	Q-likelihood model	JURU AP JURU PC@10	self (sys AP)	Query AP
predict Q hardness	distance -d(Q,C) -d(Q,Rel) -d(Rel,C)	clarity score	rank robust.	overlap between top N results on sub-queries	self (sys) prediction	InfoDFR
evaluate prediction	ρ Spearman	Spearman	ρ Spearman R^2 regress. coeff	τ	τ	correlation with AP
collection tested	Terabyte04-05	TREC4-8 AP 88-89	TREC 1-5 Robust 04 Terabyte04-05	TREC 8 WT10G	Robust 04	TREC 4,5



query:

"CIKM 2003"

- Is it difficult ?

query:
"CIKM 2003"

● Is it difficult ?



The screenshot shows a Google search interface with the query "CIKM 2003". The results list several links related to the CIKM 2003 conference, including the homepage, call for papers, registration, and mailing lists.

Google ^{Adv} ^{cik}

Web Images Groups Dire
Searched the web for **CIKM 2003**

CIKM 2003 Homepage
... 03, Hotel Inter-Continental, Novem
Conference on Information and Know
Description: ACM Conference on Info
Category: [Reference > Knowledge M](#)
bit.csc.lsu.edu/~cikm2003/ - 9k - [Ca](#)

ACM CIKM 2003 Call Fo
... Management (CIKM03). Ac
highest caliber papers submit
bit.csc.lsu.edu/~cikm2003/CI
[[More results from bit.csc.lsu](#)

CIKM-2003 Registration
... All Amounts in USD. Attendee Sta
one author of an accepted paper mus
www.cikmreg.org/ - 13k - [Cached](#) - [S](#)

ACM CIKM 2003 Call For Pap
ACM CIKM 2003 -Call for Industry T
Sponsored by ACM SIGIR and ACM
www.cs.wisc.edu/dbworld/messages

CIKM 2002 Homepage
... 03), Las Vegas, Nevada, June 23-
Conference on Information and Know
Description: SAIC Headquarters, Mc
Category: [Business > Management](#)
www.cikm.org/2002/ - 16k - [Cached](#) -

Conference on Informatio
... The CIKM 2003 web page i
devoted to emerging areas of
www.cikm.org/ - 6k - [Cached](#)

[Asis-] **CIKM 2003**
[Asis-] CIKM 2003. Padmini Srinivas
Mon, 29 Sep 2003 12:59:36 -0500: P
mail.asis.org/pipermail/asis-l/2003-S

(DBWORLD) **CIKM 2003** (Pad
(DBWORLD) CIKM 2003 (Padmini S
CIKM 2003; From: "Padmini Srinivas
wwwiti.cs.uni-magdeburg.de/MailArc

bridge-cikm-2003
bit.csc.lsu.edu/~cikm2003.
www.logistics-2000-versailles.net/site

Collaborative Filtering Mailing
[collab@sims] CFP: **CIKM 2003** ...

query:

"CIKM 2003"

● Is it difficult ?

The image displays three overlapping search engine results for the query "cikm 2003".

- Google (top):** Shows search results for "cikm 2003". The first result is "CIKM 2003 Homepage" with a description: "... 03, Hotel Inter-Continental, Novem... Conference on Information and Know... Description: ACM Conference on Info... Category: Reference > Knowledge M... bit.csc.lsu.edu/~cikm2003/ - 9k - Ca...". Other results include "ACM CIKM 2003 Call For...", "CIKM-2003 Registration", "ACM CIKM 2003 Call For Paper", "CIKM 2002 Homepage", "Conference on Information", "[Asis-] CIKM 2003", "(DBWORLD) CIKM 2003 (Padmini S", "bridge-cikm-2003", and "Collaborative Filtering Mailing".
- Altavista (middle):** Shows search results for "cikm 2003". The first result is "CIKM 2003 Homepage" with a description: "... 12th international conference on info... bit.csc.lsu.edu/~cikm2003 • Refresh... More pages from bit.csc.lsu.edu". Other results include "ACM CIKM 2003 Call For...", "CIKM 2002 Homepage", "Conference on Information", "CIKM", and "SIGIR Information Server".
- alltheweb (bottom):** Shows search results for "cikm 2003". The first result is "CIKM 2003 Homepage" with a description: "... Information and Knowledge Management CIKM... an international ... Description: The ACM Conference on Information... and knowledge management, as well as recent ac... more hits from: http://bit.csc.lsu.edu/~cikm2003/...". Other results include "IVML Call For Papers Archive: ACM CI...", "CIKM 2002 Homepage", "ACM WIDM'2003", "MMDB'03", "D-Lib Workshops and Conferences: 20...", and "Invited paper to be presented in CIKM".



● Is it difficult ?

- **NO**- if the same documents were retrieved in the same order by all systems
- **YES** - if the documents were retrieved in different order

[illegible]

query:

"CIKM 2003"

● Is it difficult ?

- **NO**- if the documents were retrieved in the same order by all systems
- **YES** - if the documents were retrieved in different order

The collage displays search results for the query "cikm 2003" from three different search engines: Google, Altavista, and AllTheWeb. Each engine's results are shown in a separate window, with the top result "CIKM 2003 Homepage" highlighted in green. The results show that while the top result is the same, the order of subsequent results varies between the different search engines, illustrating the difficulty of achieving consistent retrieval order across systems.

Google Results:

- CIKM 2003 Homepage
- ACM CIKM 2003 Call For Papers
- CIKM 2003 Registration
- ACM CIKM 2003 Call For Papers
- CIKM 2002 Homepage
- Conference on Information and Knowledge Management (CIKM)
- [Asis-] CIKM 2003
- (DBWORLD) CIKM 2003 (Padmini Srinivas)
- bridge-cikm-2003
- Collaborative Filtering Mailing List

Altavista Results:

- CIKM 2003 Homepage
- ACM CIKM 2003 Call For Papers
- CIKM 2002 Homepage
- Conference on Information and Knowledge Management (CIKM)
- CIKM
- 7. CIKM 1998: Bethesda, Maryland
- SIGIR Information Server
- Web Caching Publications/Venue

AllTheWeb Results:

- CIKM 2003 Homepage
- IVML Call For Papers Archive: ACM CIKM
- CIKM 2002 Homepage
- ACM WIDM'2003
- MMDB'03
- D-Lib Workshops and Conferences: 2003
- Invited paper to be presented in CIKM
- Internet Conferences 2003

query:

"CIKM 2003"

● Is it difficult ?

- **NO**- if the same documents were retrieved in the same order by all systems
- **YES** - if the documents were retrieved in different order

Google Advanced Search Preferences Language cikm 2003

Web Images Groups Directories Searched the web for cikm 2003

CIKM 2003 Homepage
... 03, Hotel Inter-Continental, Nov...
Conference on Information and Know...
Description: ACM Conference on In...
Category: Reference > Knowledge

ACM **CIKM 2003** Call For...
... Management (CIKM03). A...
highest caliber papers submit...
bit.csc.lsu.edu/~cikm2003/CI...
[More results from bit.csc.lsu.edu]

CIKM-2003 Registration
... All Amounts in USD. Attendee Sta...
one author of an accepted paper mus...
www.cikmreg.org/ - 13k - Cached - S...

ACM **CIKM 2003** Call For Pa...
ACM **CIKM 2003** -Call for Industry T...
Sponsored by ACM SIGIR and ACM...
www.cs.wisc.edu/dbworld/messages...

CIKM 2002 Homepage
... 03), Las Vegas, Nevada, June 23...
Conference on Information and Know...
Description: SAIC Headquarters, Mc...
Category: Business > Management

Conference on Information...
... The **CIKM 2003** web page i...
devoted to emerging areas of...
www.cikm.org/ - 6k - Cached

[Asis-] **CIKM 2003**
[Asis-] **CIKM 2003**. Padmini Srinivas...
Mon, 29 Sep 2003 12:59:36 -0500: P...
mail.asis.org/pipermail/asis-l/2003-S...

(DBWORLD) **CIKM 2003** (Pad...
(DBWORLD) **CIKM 2003** (Padmini S...
CIKM 2003; From: "Padmini Srinivas...
www.witi.cs.uni-magdeburg.de/MailArc...

bridge-cikm-2003
bit.csc.lsu.edu/~cikm2003.
www.logistics-2000-versailles.net/site...

Collaborative Filtering Mailing...
[collab@sim] CFP: **CIKM 2003** ...

altavista Web Images Searched the web for cikm 2003

CIKM 2003 Homepage
12th international conference on in...
bit.csc.lsu.edu/~cikm2003 • Refres...
More pages from bit.csc.lsu.edu

ACM **CIKM 2003** Call For...
... Management (CIKM03) A...
caliber papers submitted to C...
bit.csc.lsu.edu/~cikm2003/C...
More pages from bit.csc.lsu.edu

CIKM 2002 Homepage
... 2002 Advance Technical Progra...
Treasurer, Dr. Nicholas, the **CIKM**...
www.cikm.org/2002 • Related Page...

Conference on Information...
CIKM CIKM 2003 Topi...
Knowledge Management The...
www.cikm.org • Related Page...
More pages from www.cikm.org

CIKM
... Information and Knowledge Mana...
CIKM 2003 Home Page 11. **CIKM 2**...
www.informatik.uni-trier.de/~ley/db/c...
More pages from www.informatik.uni...

7. **CIKM 1998: Bethesda**, Ma...
7. **CIKM 1998: Bethesda**, Ma...
Copyright © by Michael Ley...
www.informatik.uni-trier.de/~l...
More pages from www.inform...

SIGIR Information Server
... award. Upcoming SIGIR Spons...
2003. JCDL 2004 (Tucson, AZ) - Jur...
www.acm.org/sigir • Refreshed in pa...
More pages from www.acm.org

Web Caching Publications/Venu...

alltheweb advanced se cikm 2003 Results in: C

Web News Pictures Video Audio

CIKM 2003 Homepage
... Information and Knowledge Management **CIKM**...
an international ...
Description: The ACM Conference on Information...
and knowledge management, as well as recent ac...
more hits from: http://bit.csc.lsu.edu/~cikm2003/

IVML Call For Papers Archive: ACM **CI**...
Fo ACM **CIKM 2003** Preliminary Call For Papers...
message ... Digital Games Conference 2003" Pre...
Description: Previous message: Kostas Karpouzis...
Conferences - SPIE Web...
http://www.image.ntua.gr/cfp/archive/2003/0234.htm

CIKM 2002 Homepage
... Information and Knowledge Management (**CIKM**...
been ... Information and Knowledge Management...
Description: SAIC Headquarters, McLean, Virginia...
more hits from: http://www.cikm.org/2002/ - 15 k

ACM WIDM'2003
... **CIKM 2003** ... Information and Knowledge Man...
of the workshop ...
Description: WIDM03 is the fifth in a series of wor...
Information and Knowledge Management (**CIKM**),...
more hits from: http://www.cais.ntu.edu.sg/widm2...

MMDB'03
... Programs ACM - MMDB 2003 The First ACM In...
Knowledge Management (ACM **CIKM 2003**). One...
Description: The First ACM International Worksho...
university researchers, scientists, industry profess...
more hits from: http://www.cs.fiu.edu/mmdb03/ -

D-Lib Workshops and Conferences: 20...
... proceedings, and other documents. 2004 2003...
November 2003, Sicily, Italy ACM **CIKM 2003**, 12...
Description: Below is a listing of digital library mee...
Click on a year to find additional events with links...
more hits from: http://www.dlib.org/groups.html -

Invited paper to be presented in **CIKM** (...
Invited paper to be presented in **CIKM** (Conference...
Privacy Policy - Copyright © 2000-2003 NEC and...
Description: Abstract: The main objective of a virtu...
resources contributed by the organizations toward...
more hits from: http://gunther.smeal.psu.edu/199...

Internet Conferences 2003
... 2000 Books 2001 Books 2002 Books 2003 net...

from ranked list to distribution

$$\begin{aligned} \text{weight}(r) &= \\ &= \frac{1}{2c} \left(1 + \frac{1}{r} + \frac{1}{r+1} + \dots + \frac{1}{c} \right) \\ &\approx \frac{1}{2c} \log \frac{c}{r} \end{aligned}$$

Google

Web Images Groups Dir
Searched the web for **cikm 2003**.

CIKM 2003 Homepage

... 03, Hotel Inter-Continental, Novem
Conference on Information and Know
Description: ACM Conference on Info
Category: Reference > Knowledge M
bit.csc.lsu.edu/~cikm2003/ - 9k - Ca

ACM CIKM 2003 Call For

... Management (CIKM03). AC
highest caliber papers submit
bit.csc.lsu.edu/~cikm2003/CI
[More results from bit.csc.lsu

CIKM-2003 Registration

... All Amounts in USD. Attendee Sta
one author of an accepted paper mus
www.cikmreg.org/ - 13k - Cached - S

ACM CIKM 2003 Call For Paper

ACM CIKM 2003 -Call for Industry T
Sponsored by ACM SIGIR and ACM
www.cs.wisc.edu/dbworld/messages

CIKM 2002 Homepage

... 03), Las Vegas, Nevada, June 23-
Conference on Information and Know
Description: SAIC Headquarters, Mc
Category: Business > Management
www.cikm.org/2002/ - 16k - Cached

Conference on Information

... The CIKM 2003 web page i
devoted to emerging areas of
www.cikm.org/ - 6k - Cached

[Asis-l] CIKM 2003

[Asis-l] CIKM 2003. Padmini Srinivas
Mon, 29 Sep 2003 12:59:36 -0500: P
mail.asis.org/pipermail/asis-l/ 2003-S

(DBWORLD) CIKM 2003 (Pad

(DBWORLD) CIKM 2003 (Padmini S
CIKM 2003; From: "Padmini Srinivas
wwwiti.cs.uni-magdeburg.de/MailArc

bridge-cikm-2003

bit.csc.lsu.edu/~cikm2003.
www.logistics-2000-versailles.net/site


Collaborative Filtering Mailing

[collab@sims] CFP: CIKM 2003 ...



$$\text{weight}(r) =$$

$$\frac{1}{2c} \left(1 + \frac{1}{r} + \frac{1}{r+1} + \dots + \frac{1}{c} \right) \approx \frac{1}{2c} \log \frac{c}{r}$$



[Web](#)
[Images](#)

SEARCH:
☒ Worldwide
 ☐ U.S.

Did you mean: [ci km 2003](#)

AltaVista found 5,186 results [About](#)

[CIKM 2003 Homepage](#)
 12th international conference on information and knowledge management
bit.csc.lsu.edu/~cikm2003 • [Refresh](#)
[More pages from bit.csc.lsu.edu](#)

[ACM CIKM 2003 Call For Papers](#)
 ... Management (CIKM03) Accepted Papers
 caliber papers submitted to CIKM 2003
bit.csc.lsu.edu/~cikm2003/CallForPapers
[More pages from bit.csc.lsu.edu](#)

[CIKM 2002 Homepage](#)
 ... 2002 Advance Technical Program
 Treasurer, Dr. Nicholas, the CIKM 2002
www.cikm.org/2002 • [Related Pages](#)

[Conference on Information and Knowledge Management](#)
CIKM CIKM 2003 CIKM Topics
 Knowledge Management Theoretical Foundations
www.cikm.org • [Related Pages](#)
[More pages from www.cikm.org](#)

[CIKM](#)
 ... Information and Knowledge Management
CIKM 2003 Home Page 11. CIKM 2003
www.informatik.uni-trier.de/~ley/db/Conf/CIKM
[More pages from www.informatik.uni-trier.de](#)

[7. CIKM 1998: Bethesda, Maryland](#)
 7. CIKM 1998: Bethesda, Maryland
 Copyright © by Michael Ley
www.informatik.uni-trier.de/~ley/db/Conf/CIKM
[More pages from www.informatik.uni-trier.de](#)

[SIGIR Information Server](#)
 ... award. Upcoming SIGIR Sponsor
2003: JCDL 2004 (Tucson, AZ) - June 2004
www.acm.org/sigir • [Refreshed in page](#)
[More pages from www.acm.org](#)

[Web Caching Publications/Venues](#)
 ... Publications/Venues



$$\text{weight}(r) =$$

$$\frac{1}{2c} \left(1 + \frac{1}{r} + \frac{1}{r+1} + \dots + \frac{1}{c} \right)$$
$$\approx \frac{1}{2c} \log \frac{c}{r}$$

what motivates
this formula ?
-> next slide

<p>alltheweb • • • find it all • • •</p> <p>Web News Pictures Vid</p> <p>1 - 10 of 6,293 Results for cikm 2003</p>	
<p>CIKM 2003 Homepage ... Information and Knowledge Management ... Description: The ACM Conference on Information and Knowledge Management, as well as the ACM SIGMOD Conference on Management of Data more hits from: http://bit.csc.lsu.edu/cikm2003/</p>	
<p>IVML Call For Papers Archived For ACM CIKM 2003 Preliminary Call for Papers ... Digital Games Conference Description: Previous message: Kostas S. Papadimitriou - SPIE Web http://www.image.ntua.gr/cfp/archive/</p>	
<p>CIKM 2002 Homepage ... Information and Knowledge Management ... Description: SAIC Headquarters, Madrid more hits from: http://www.cikm.org/</p>	
<p>ACM WIDM'2003 ... CIKM 2003 ... Information and Knowledge Management ... Description: WIDM03 is the fifth in a series of workshops on Information and Knowledge Management more hits from: http://www.cais.ntu.edu.sg/widm03/</p>	
<p>MMDB'03 ... Programs ACM - MMDB 2003 The First ACM International Conference on Knowledge Management (ACM CIKM) ... Description: The First ACM International Conference on Knowledge Management (ACM CIKM) ... more hits from: http://www.cs.fiu.edu/~mmdb/</p>	
<p>D-Lib Workshops and Conferences ... proceedings, and other documents ... November 2003, Sicily, Italy ACM Conference on Information and Knowledge Management Description: Below is a listing of digital libraries ... Click on a year to find additional events more hits from: http://www.dlib.org/</p>	
<p>Invited paper to be presented at CIKM 2003 Invited paper to be presented in CIKM 2003 Privacy Policy - Copyright © 2000-2003 Description: Abstract: The main objective of the conference is to provide a forum for the exchange of ideas and resources contributed by the organization more hits from: http://gunther.smea.com/</p>	
<p>Internet Conferences 2003 ... 2000 Books 2001 Books 2002 Books 2003 Books 2004 Books 2005 Books 2006 Books 2007 Books 2008 Books 2009 Books 2010 Books 2011 Books 2012 Books 2013 Books 2014 Books 2015 Books 2016 Books 2017 Books 2018 Books 2019 Books 2020 Books 2021 Books 2022 Books 2023 Books 2024 Books 2025 Books 2026 Books 2027 Books 2028 Books 2029 Books 2030 Books 2031 Books 2032 Books 2033 Books 2034 Books 2035 Books 2036 Books 2037 Books 2038 Books 2039 Books 2040 Books 2041 Books 2042 Books 2043 Books 2044 Books 2045 Books 2046 Books 2047 Books 2048 Books 2049 Books 2050 Books 2051 Books 2052 Books 2053 Books 2054 Books 2055 Books 2056 Books 2057 Books 2058 Books 2059 Books 2060 Books 2061 Books 2062 Books 2063 Books 2064 Books 2065 Books 2066 Books 2067 Books 2068 Books 2069 Books 2070 Books 2071 Books 2072 Books 2073 Books 2074 Books 2075 Books 2076 Books 2077 Books 2078 Books 2079 Books 2080 Books 2081 Books 2082 Books 2083 Books 2084 Books 2085 Books 2086 Books 2087 Books 2088 Books 2089 Books 2090 Books 2091 Books 2092 Books 2093 Books 2094 Books 2095 Books 2096 Books 2097 Books 2098 Books 2099 Books 2100 Books 2101 Books 2102 Books 2103 Books 2104 Books 2105 Books 2106 Books 2107 Books 2108 Books 2109 Books 2110 Books 2111 Books 2112 Books 2113 Books 2114 Books 2115 Books 2116 Books 2117 Books 2118 Books 2119 Books 2120 Books 2121 Books 2122 Books 2123 Books 2124 Books 2125 Books 2126 Books 2127 Books 2128 Books 2129 Books 2130 Books 2131 Books 2132 Books 2133 Books 2134 Books 2135 Books 2136 Books 2137 Books 2138 Books 2139 Books 2140 Books 2141 Books 2142 Books 2143 Books 2144 Books 2145 Books 2146 Books 2147 Books 2148 Books 2149 Books 2150 Books 2151 Books 2152 Books 2153 Books 2154 Books 2155 Books 2156 Books 2157 Books 2158 Books 2159 Books 2160 Books 2161 Books 2162 Books 2163 Books 2164 Books 2165 Books 2166 Books 2167 Books 2168 Books 2169 Books 2170 Books 2171 Books 2172 Books 2173 Books 2174 Books 2175 Books 2176 Books 2177 Books 2178 Books 2179 Books 2180 Books 2181 Books 2182 Books 2183 Books 2184 Books 2185 Books 2186 Books 2187 Books 2188 Books 2189 Books 2190 Books 2191 Books 2192 Books 2193 Books 2194 Books 2195 Books 2196 Books 2197 Books 2198 Books 2199 Books 2200 Books 2201 Books 2202 Books 2203 Books 2204 Books 2205 Books 2206 Books 2207 Books 2208 Books 2209 Books 2210 Books 2211 Books 2212 Books 2213 Books 2214 Books 2215 Books 2216 Books 2217 Books 2218 Books 2219 Books 2220 Books 2221 Books 2222 Books 2223 Books 2224 Books 2225 Books 2226 Books 2227 Books 2228 Books 2229 Books 2230 Books 2231 Books 2232 Books 2233 Books 2234 Books 2235 Books 2236 Books 2237 Books 2238 Books 2239 Books 2240 Books 2241 Books 2242 Books 2243 Books 2244 Books 2245 Books 2246 Books 2247 Books 2248 Books 2249 Books 2250 Books 2251 Books 2252 Books 2253 Books 2254 Books 2255 Books 2256 Books 2257 Books 2258 Books 2259 Books 2260 Books 2261 Books 2262 Books 2263 Books 2264 Books 2265 Books 2266 Books 2267 Books 2268 Books 2269 Books 2270 Books 2271 Books 2272 Books 2273 Books 2274 Books 2275 Books 2276 Books 2277 Books 2278 Books 2279 Books 2280 Books 2281 Books 2282 Books 2283 Books 2284 Books 2285 Books 2286 Books 2287 Books 2288 Books 2289 Books 2290 Books 2291 Books 2292 Books 2293 Books 2294 Books 2295 Books 2296 Books 2297 Books 2298 Books 2299 Books 2300 Books 2301 Books 2302 Books 2303 Books 2304 Books 2305 Books 2306 Books 2307 Books 2308 Books 2309 Books 2310 Books 2311 Books 2312 Books 2313 Books 2314 Books 2315 Books 2316 Books 2317 Books 2318 Books 2319 Books 2320 Books 2321 Books 2322 Books 2323 Books 2324 Books 2325 Books 2326 Books 2327 Books 2328 Books 2329 Books 2330 Books 2331 Books 2332 Books 2333 Books 2334 Books 2335 Books 2336 Books 2337 Books 2338 Books 2339 Books 2340 Books 2341 Books 2342 Books 2343 Books 2344 Books 2345 Books 2346 Books 2347 Books 2348 Books 2349 Books 2350 Books 2351 Books 2352 Books 2353 Books 2354 Books 2355 Books 2356 Books 2357 Books 2358 Books 2359 Books 2360 Books 2361 Books 2362 Books 2363 Books 2364 Books 2365 Books 2366 Books 2367 Books 2368 Books 2369 Books 2370 Books 2371 Books 2372 Books 2373 Books 2374 Books 2375 Books 2376 Books 2377 Books 2378 Books 2379 Books 2380 Books 2381 Books 2382 Books 2383 Books 2384 Books 2385 Books 2386 Books 2387 Books 2388 Books 2389 Books 2390 Books 2391 Books 2392 Books 2393 Books 2394 Books 2395 Books 2396 Books 2397 Books 2398 Books 2399 Books 2400 Books 2401 Books 2402 Books 2403 Books 2404 Books 2405 Books 2406 Books 2407 Books 2408 Books 2409 Books 2410 Books 2411 Books 2412 Books 2413 Books 2414 Books 2415 Books 2416 Books 2417 Books 2418 Books 2419 Books 2420 Books 2421 Books 2422 Books 2423 Books 2424 Books 2425 Books 2426 Books 2427 Books 2</p>	



AP induced rank importance

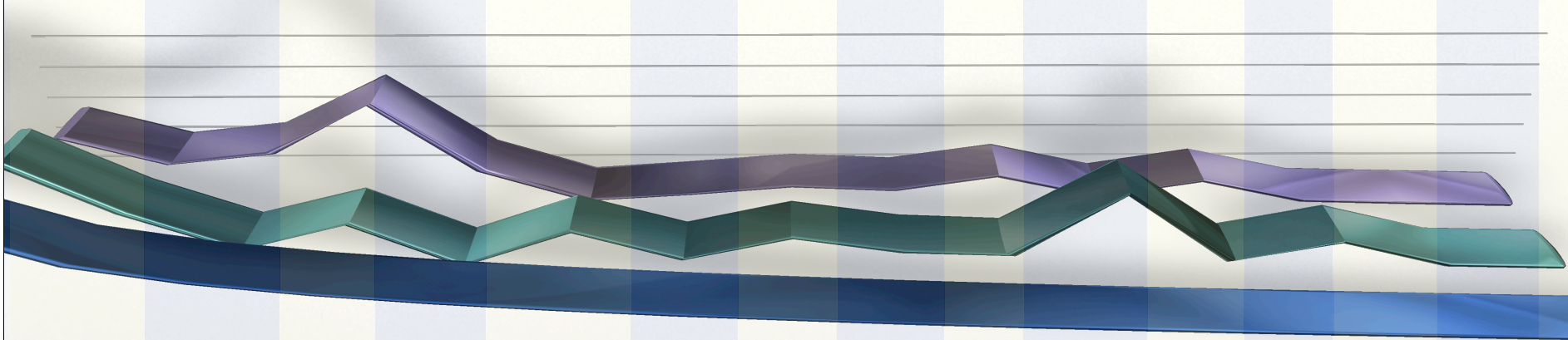
$$AP = \frac{1}{R} \sum_{1 \leq j \leq i \leq Z} \frac{1}{i} \cdot rel(i) \cdot rel(j)$$

	1	2	3	...	Z
1	1				
2	1/2	1/2			
3	1/3	1/3	1/3		
⋮					
Z	1/Z	1/Z	1/Z	...	1/Z

	1	2	3	...	Z
1	2	1/2	1/3	...	1/Z
2	1/2	1	1/3	...	1/Z
3	1/3	1/3	2/3	...	1/Z
⋮					
Z	1/Z	1/Z	1/Z	...	2/Z



compare distributions



CIKM 2003 Homepage
... 03, Hotel Inter-Continental, Novem
Conference on Information and Know
Description: ACM Conference on Info
Category: [Reference > Knowledge M](#)
[bit.csc.lsu.edu/~cikm2003/](#) - 9k - [Ca](#)

ACM CIKM 2003 Call For Papers
... Management (CIKM03). A
highest caliber papers submit
[bit.csc.lsu.edu/~cikm2003/CI](#)
[[More results from bit.csc.lsu](#)

CIKM-2003 Registration
... All Amounts in USD. Attendee Sta
one author of an accepted paper mus
[www.cikmreg.org/](#) - 13k - [Cached](#) - [S](#)

ACM CIKM 2003 Call For Papers
ACM CIKM 2003 –Call for Industry T
Sponsored by ACM SIGIR and ACM
[www.cs.wisc.edu/dbworld/messages](#)

CIKM 2002 Homepage
... 03), Las Vegas, Nevada, June 23-
Conference on Information and Know
Description: SAIC Headquarters, Mc
Category: [Business > Management](#)
[www.cikm.org/2002/](#) - 16k - [Cached](#) - [S](#)

Conference on Information and Knowledge Management
... The CIKM 2003 web page is
devoted to emerging areas of
[www.cikm.org/](#) - 6k - [Cached](#) - [S](#)

[Asis-] CIKM 2003
[Asis-] CIKM 2003. Padmini Srinivas
Mon, 29 Sep 2003 12:59:36 -0500: P
[mail.asis.org/pipermail/asis-l/2003-S](#)

(DBWORLD) CIKM 2003 (Padmini Srinivas)
(DBWORLD) CIKM 2003 (Padmini S
CIKM 2003; From: "Padmini Srinivas
[wwwiti.cs.uni-magdeburg.de/MailArc](#)

bridge-cikm-2003
[bit.csc.lsu.edu/~cikm2003](#)
[www.logistics-2000-versailles.net/site](#)

7. CIKM 1998: Bethesda, Maryland
7. CIKM 1998: Bethesda, Ma
Copyright © by Michael Ley (C
[www.informatik.uni-trier.de/~](#)
More pages from [www.inform](#)

SIGIR Information Server
... award. Upcoming SIGIR Sponsor
2003. JCDL 2004 (Tucson, AZ) - Jun
[www.acm.org/sigir](#) • Refreshed in pa
More pages from [www.acm.org](#)

Knowledge Management (ACM CIKM 2003)
Description: The First ACM International Worksho
university researchers, scientists, industry profess
[more hits from:](#) [http://www.cs.fu.edu/mmbd03/](#) - [S](#)

D-Lib Workshops and Conferences: 2003
proceedings and other documents: 2004 2003
November 2003, Sicily, Italy ACM CIKM 2003, 12
Description: Below is a listing of digital library me
Click on a year to find additional events with links
[more hits from:](#) [http://www.dlib.org/groups.html](#) - [S](#)

Invited paper to be presented in CIKM (Conference on Information and Knowledge Management)
Invited paper to be presented in CIKM (Conference
Privacy Policy - Copyright © 2000-2003 NEC and
Description: Abstract: The main objective of a virt
resources contributed by the organizations toward
[more hits from:](#) [http://gunter.sneak.net/psu.edu/199](#)

Internet Conferences 2003
... 2000 Books 2001 Books 2002 Books 2003 net



Jensen-Shannon divergence

$$JS(\mathbf{p}_1, \dots, \mathbf{p}_m) = \frac{1}{m} \sum_j KL(\mathbf{p}_j || \bar{\mathbf{p}})$$

- generalized, symmetrized version of KL
 - 0 iff all distributions are identical
- avg. distance to avg. distribution
- vastly used
- well studied



how to test

- compare JS divergence
 - computed for 2,5,10,20 systems
- with defined query hardness as AVERAGE (Or MEDIAN or MEDIAN-system) measured performance off all systems for that query



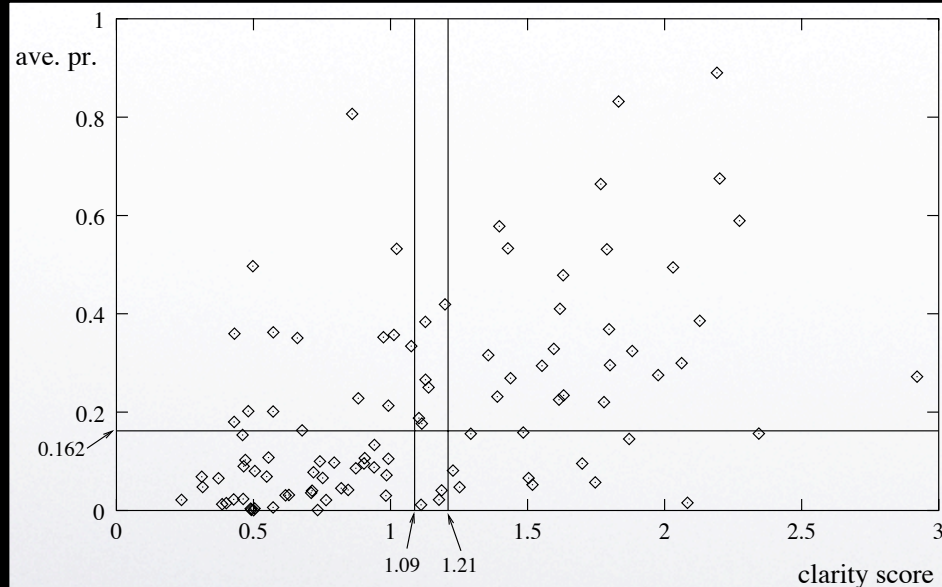
how to test



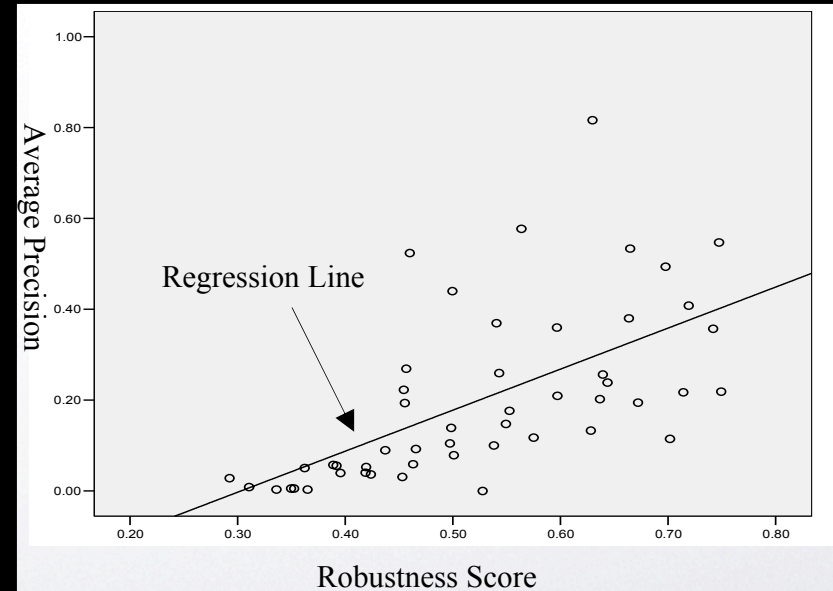
	What makes a query difficult ? [Carmel...]	Predicting query performance [Townsend, Zhou, Croft]	Ranking robustness [Zhou, Croft]	Learning to estimate query difficulty [Yom-Tov...]	Overview of TREC 2004 Robust track [Voorhees]	Query difficulty, robustness [Amati et al]	JS divergence among systems [Aslam, Pavlu]
define Q hardness	Q median AP JURU AP	multinom language model AP	Q-likelihood model	JURU AP JURU PC@10	self (sys AP)	Query AP	Q avg AP Q med AP Q medsys AP
predict Q hardness	distance -d(Q,C) -d(Q,Rel) -d(Rel,C)	clarity score	rank robust.	overlap between top N results on sub-queries	self (sys) prediction	InfoDFR	JS divergence
evaluate prediction	ρ Spearman	Spearman	ρ Spearman R^2 regress.	τ	τ	correlation with AP	ρ τ
collection tested	Terabyte04-05	TREC4-8 AP 88-89	TREC 1-5 Robust 04 Terabyte04-05	TREC 8 WT10G	Robust 04	TREC 4,5	TREC 5-8 Robust 04 Terabyte 04-05



calibration: how difficult is this problem ?



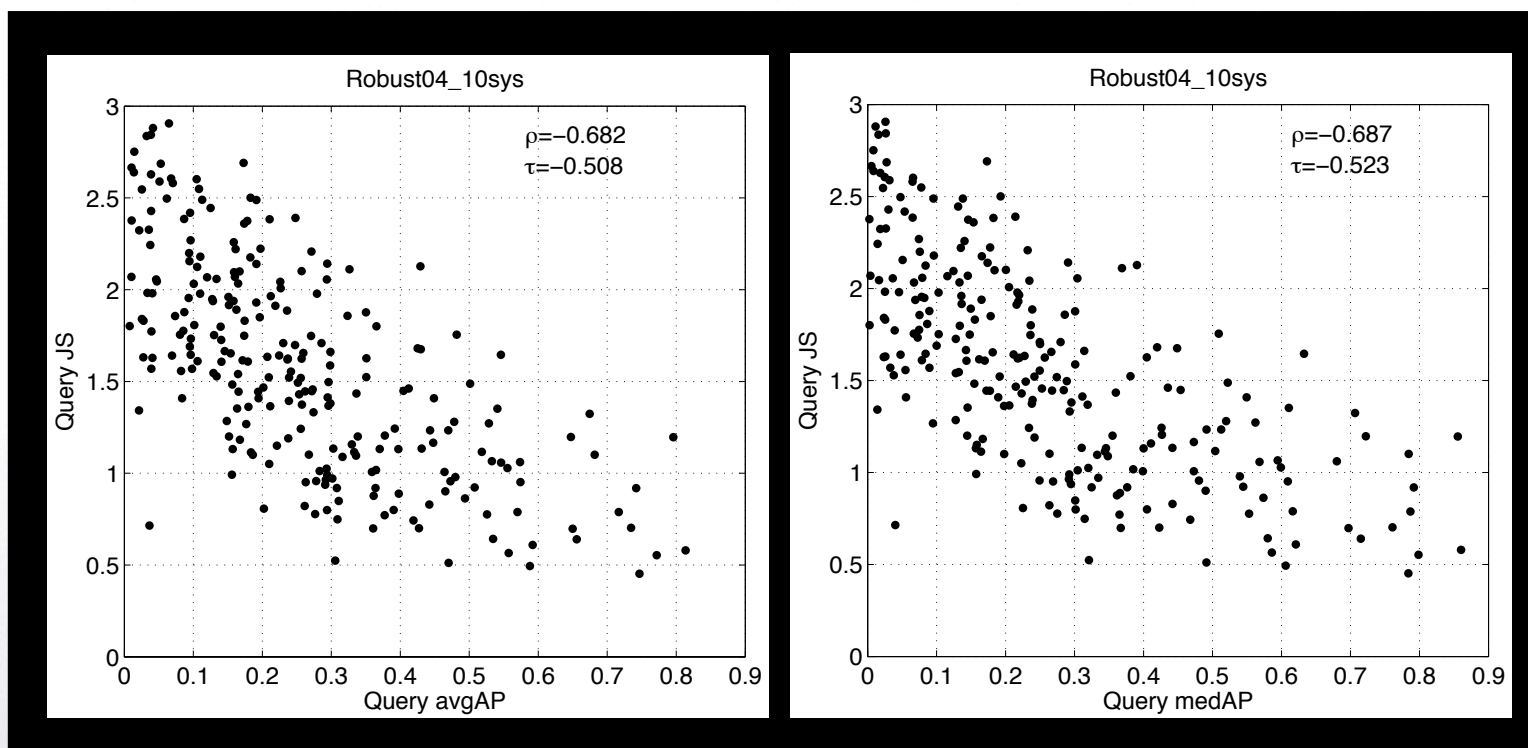
2002



2006



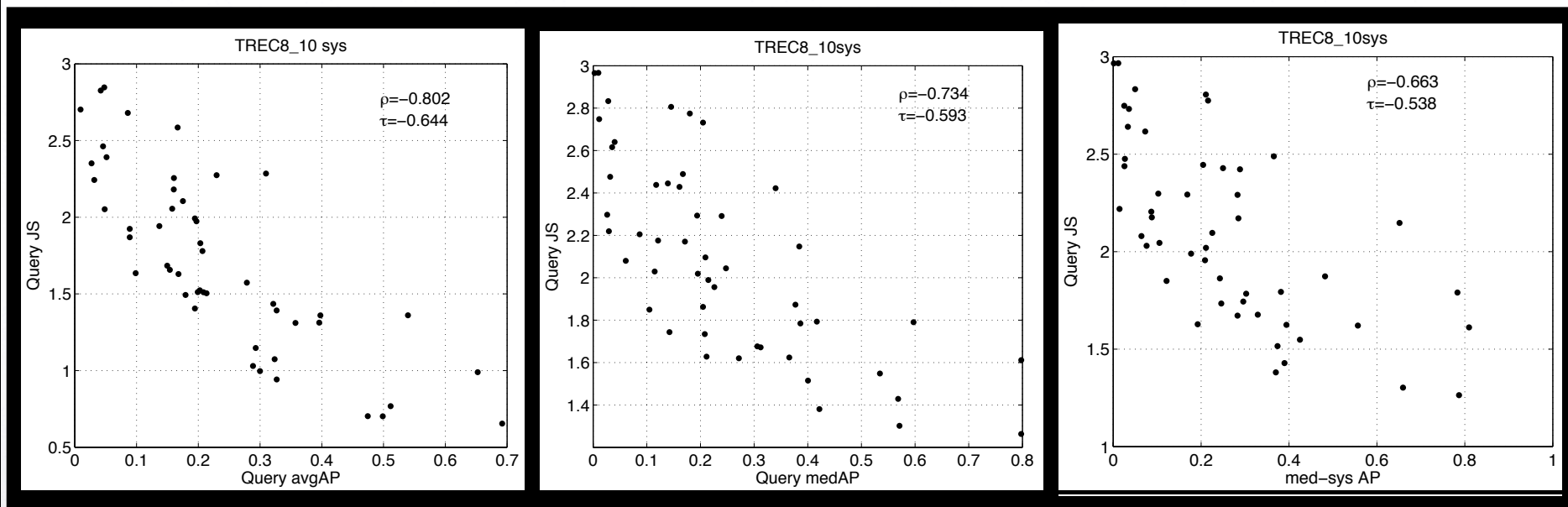
results



Robust Track, TREC2004, 249 queries



results



adhoc track, TREC8, 50 queries



results

Prediction method	TREC5	TREC6	TREC7	TREC8	Robust04	TB04	TB05
[7] clarity	0.366				0.507	0.305	0.206
[7] robust+clarity	0.469				0.613	0.374	0.362
JS (2 systems)	0.425	0.294	0.553	0.595	0.542	0.435	0.338
JS (5 systems)	0.537	0.459	0.609	0.676	0.645	0.490	0.467
JS (10 systems)	0.556	0.469	0.639	0.707	0.659	0.566	0.524
JS (20 systems)	0.562	0.479	0.679	0.724	0.665	0.585	0.545
JS (all systems)	0.567	0.497	0.657	0.702	0.677	0.603	0.541

correlation coefficient (ρ)



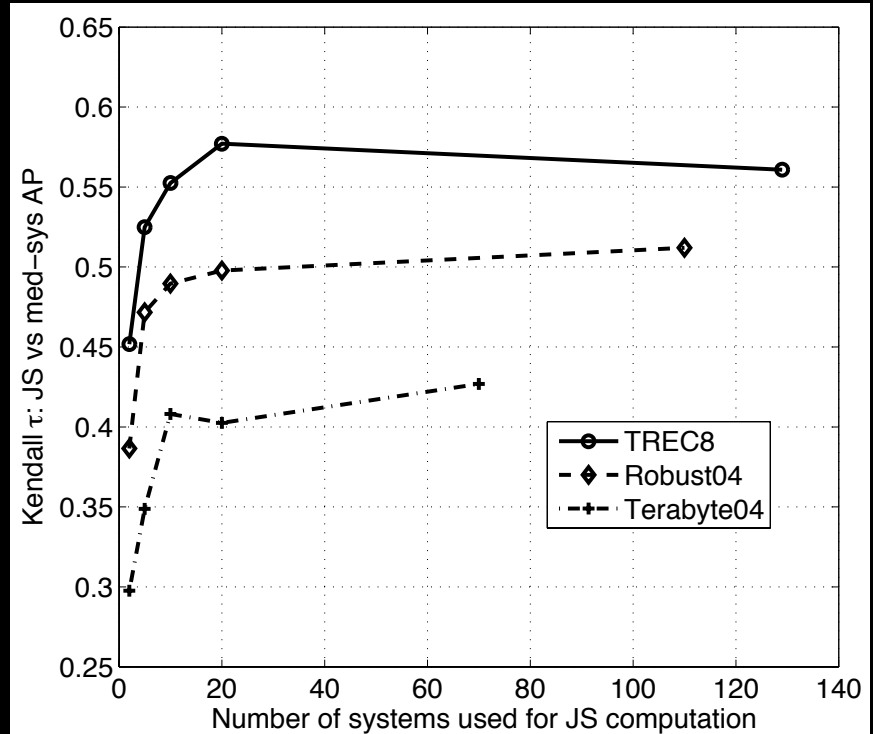
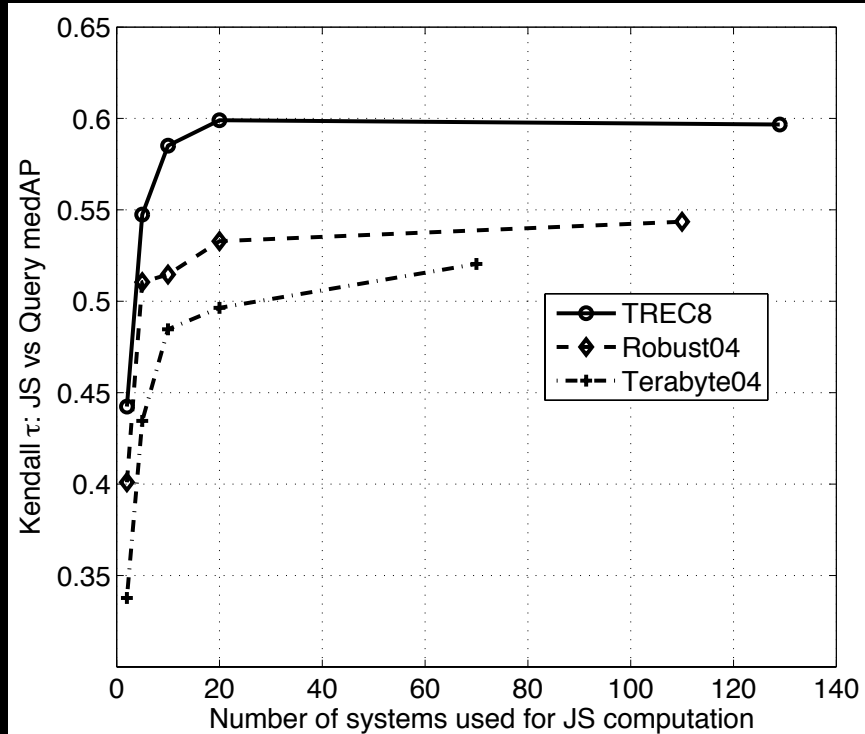
results

Prediction Method	TREC5	TREC6	TREC7	TREC8	Robust04	TB04	TB05
[7] clarity	0.311				0.412	0.134	0.171
[7] robust+clarity	0.345				0.460	0.226	0.252
[1] hist. boost. class.				.439			
JS (2 systems)	0.260	0.276	0.384	0.452	0.387	0.298	0.241
JS (5 systems)	0.350	0.370	0.427	0.525	0.472	0.349	0.318
JS (10 systems)	0.355	0.334	0.476	0.552	0.490	0.408	0.359
JS (20 systems)	0.339	0.365	0.516	0.577	0.498	0.403	0.380
JS (all systems)	0.363	0.355	0.509	0.561	0.512	0.427	0.381

Kendall's τ

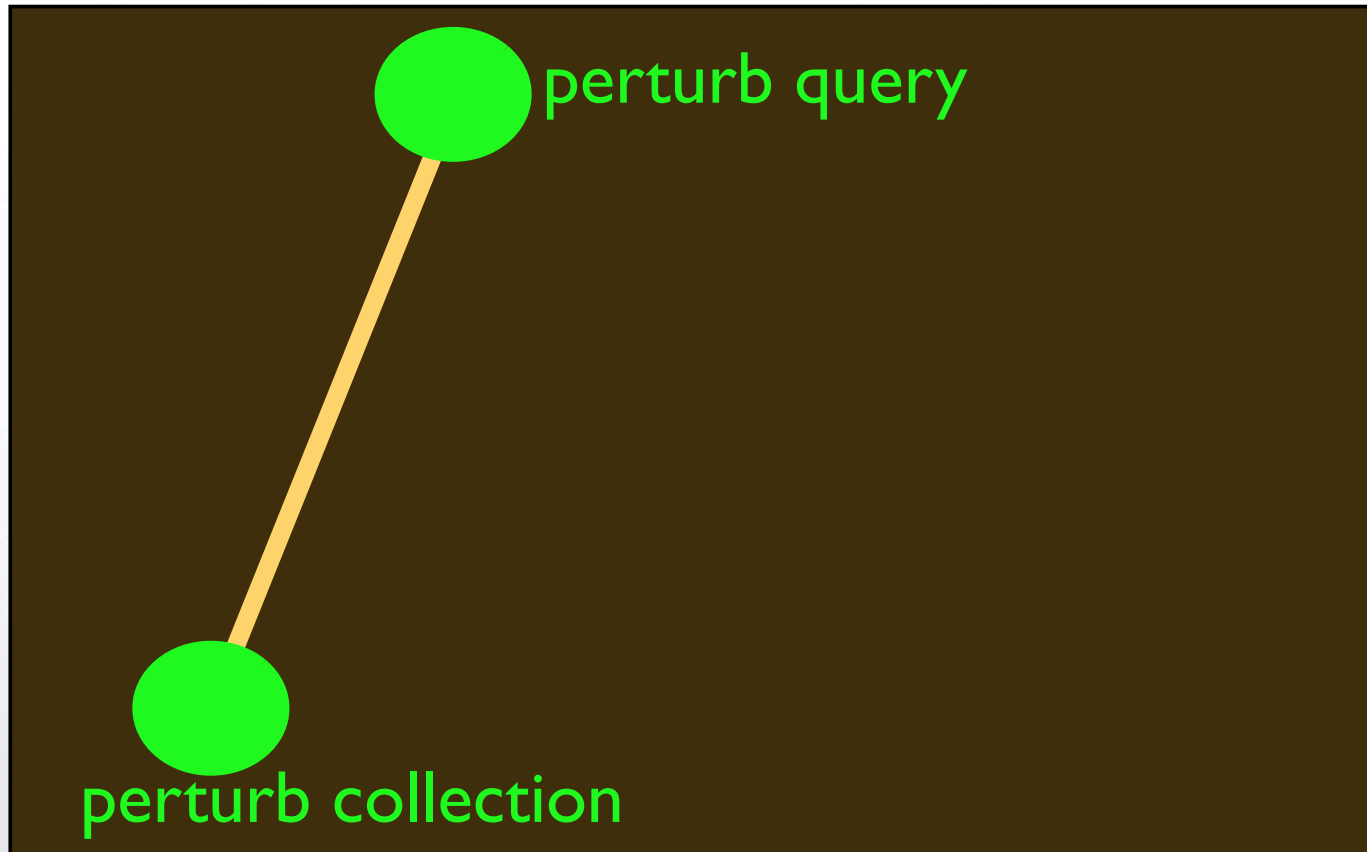


results



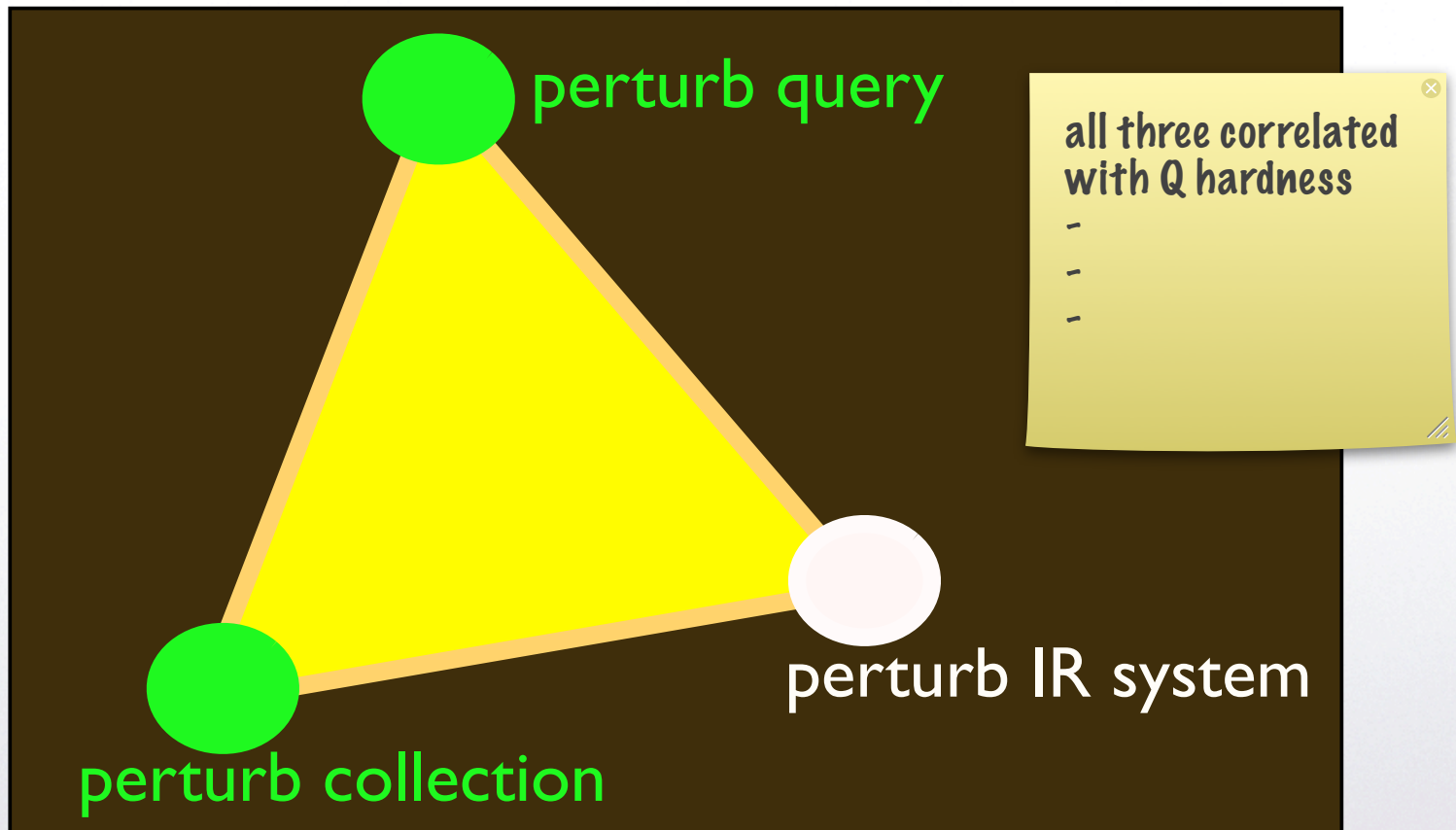


query hardness estimation



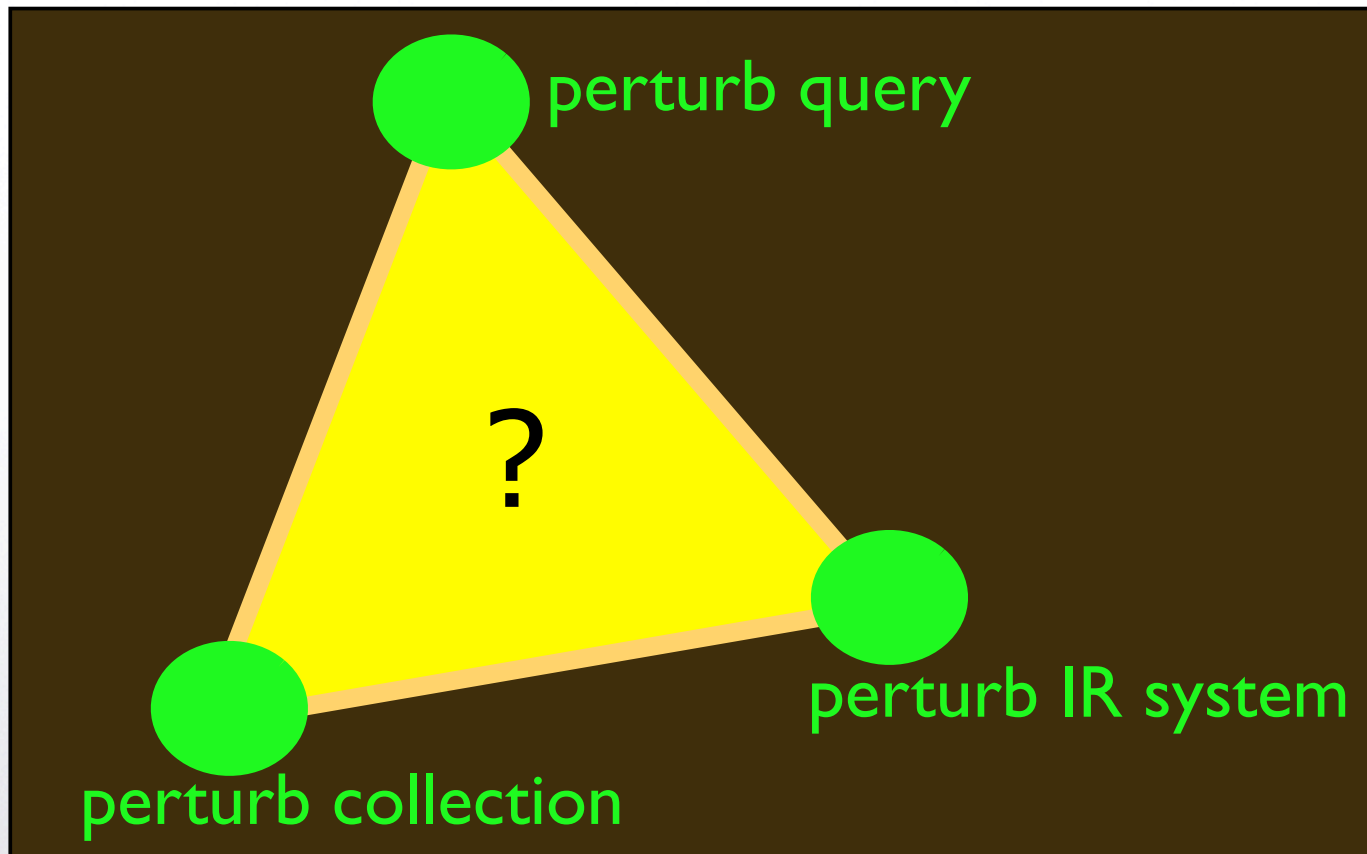


query hardness estimation





query hardness estimation





thank you

- questions ?