

Simple performance evaluation of Baidu, sogou, Bing

with three Chinese queries

IR-HW-05-朱海峰-190110716

methods

Queries used in this evaluation:

1. 哈工大、韩杰才、校长
2. 哈尔滨、地铺、热
3. 出省、核酸、深圳

For each query, the user need:

1. Any information that 韩杰才 becomes new 校长 of 哈工大.
2. The news that students of some colleges in 哈尔滨 make 地铺 because of the weather which is too 热.
3. The latest official policy whether 核酸 test is needed when 出省 from 深圳.

Procedure

The first query

For the first query, 哈工大、韩杰才、校长, on each search engine with top 20 documents returned, I manually judge the relevance.

Documents are just too many to list, so I just take a screenshot as an example, shown as figure

1.

Relevant documents are marked with “1” while the others I leave blank or mark with “0”, just as in figure 2.

哈工大 韩杰才 校长

Q 网页 资讯 视频 图片 知道 文库 贴吧 地图 采购

百度为您找到相关结果约375,000个

哈工大常务副校长韩杰才院士走进焊接基层党支部讲党课

2018年12月24日 2018年12月20日下午,常务副校长韩杰才教授走进材料学院焊接系,为支部成员讲授了题为《做新时代的先锋》的专题党课。本次会议由焊接系主任曹健教授主持,材料学院党委书记刘伟老...
weld.hit.edu.cn/2018/1224/c164... 百度快照

韩杰才任哈工大常务副校长 曹喜滨刘宏任哈工大常委...

【东方】2019年2月26日 哈工大报讯(童笑梅/文)日前,工业和信息化部党组决定,韩杰才任哈尔滨工业大学常务副校长(正局级),曹喜滨任哈尔滨工业大学党委书记,刘宏任哈尔滨工业大学副校长。...
news.hit.edu.cn/2019/0226/c151... 百度快照

哈工大新校长韩杰才院士? - 知乎

9小时前 首页 会员 发现 等你来答 登录加入知乎 哈工大新校长韩杰才院士? 关注问题写回答 哈尔滨工业大学(HIT) 校长 校园热点 哈工大新校长 韩杰才院士? 鱼香 关注者23 被浏览7,062 关注问题...
知乎 百度快照

其他人还在搜

哈工大韩杰才到处跑 韩杰才新校长 哈尔滨工业大学韩杰才院士 哈工大一共多少院士
哈工大换校长 哈工大博导名单 韩杰才数学学院 哈工大校长助理范峰

有人说哈工大副校长韩杰才回来我电当校长?【985吧】 - 百...

回复数: 14 发帖时间: 2018年01月19日

1楼: 这是真的吗? 大家咋看呀?

figure 1 documents returned

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Baidu			1																	
sogou				1																
bing																				

Total documents relevant:1.

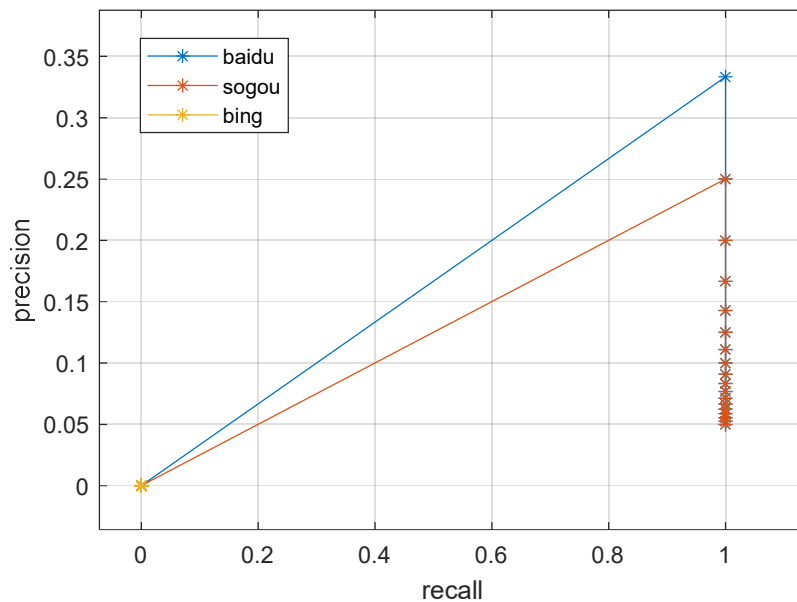


figure 2 P/R graph for the first query

$$\text{baidu:MAP} = \frac{1}{1} \sum_1^1 1 = 1$$

$$\text{sogou:MAP} = \frac{1}{1} \sum_1^1 1 = 1$$

$$\text{bing:MAP} = 0$$

The second query

For the second query, 哈尔滨、地铺、热, methods remains the same:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Baidu	0	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	0	0	0	0
sogou	1	0	0	0	1	1	0	1	0	1	1	1	1	0	0	0	1	1	1	0
bing	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

Total documents relevant:22.

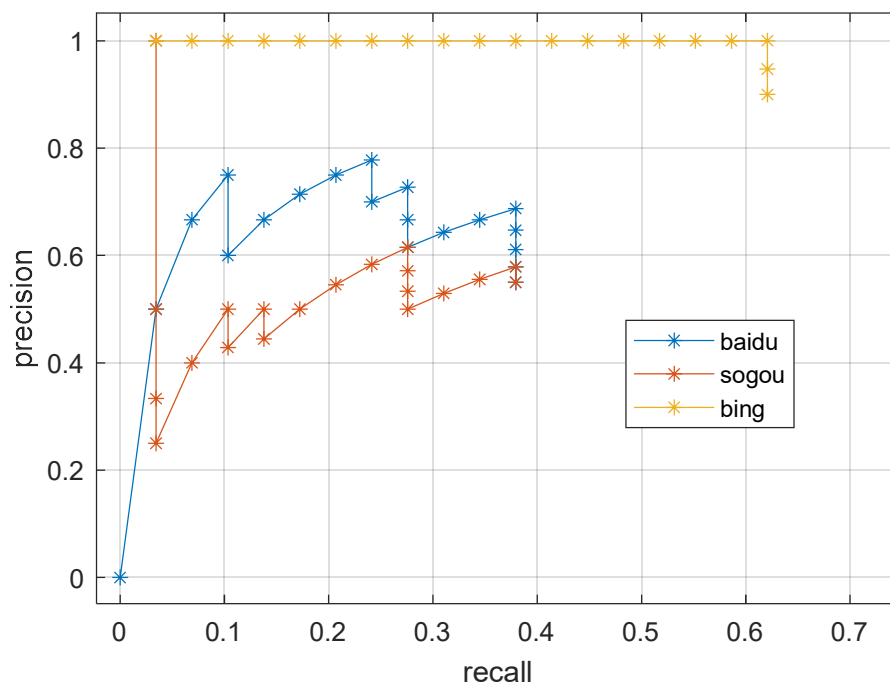


figure 3 P/R graph for the second query

$$\text{baidu:MAP} = \frac{1}{11} \sum_{i=1}^{11} \left(\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{6} + \frac{5}{7} + \frac{6}{8} + \frac{7}{9} + \frac{8}{11} + \frac{9}{14} + \frac{10}{15} + \frac{11}{16} \right) = 0.68633576$$

$$\text{sogou: MAP} = 0.573462471$$

$$\text{bing: MAP} = 1$$

The third query

For the third query, 出省、核酸、深圳, methods remains the same:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Baidu	1	1	1	0	1	1	1	0	1	1	1	0	1	0	1	0	0	0	0	0
sogou	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0
bing	1	1	0	1	0	1	0	0	0	1	0	1	0	0	1	1	0	1	1	0

Total documents relevant:25.

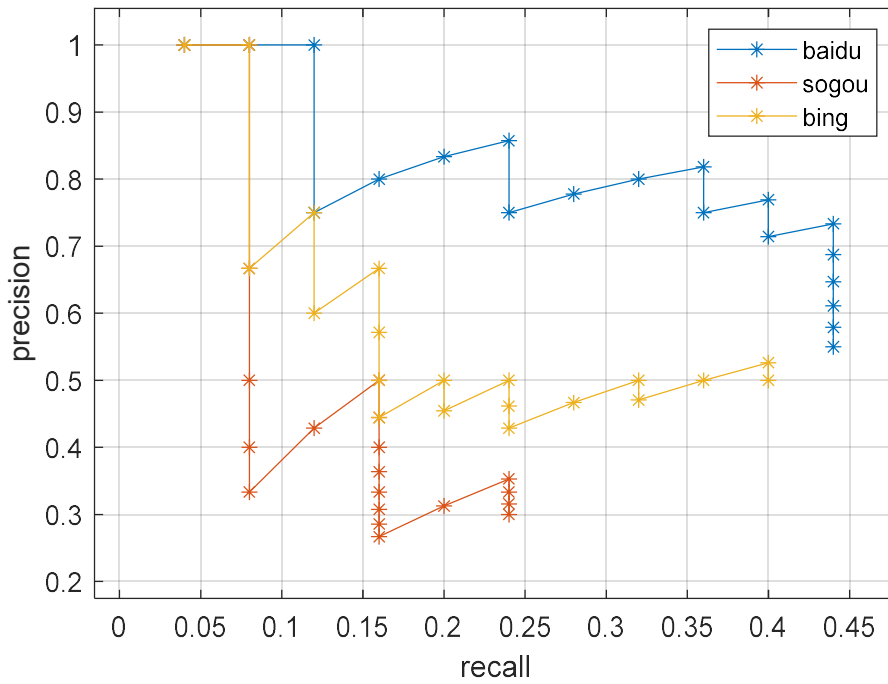


figure 4 P/R graph for the third query

baidu:MAP = 0.853545444

sogou: MAP = 0.599002101

bing: MAP = 0.640964912

conclusion

we can get the general MAP:

$$\text{MAP}_{\text{baidu}} = \frac{1}{3} \sum_1^3 (1 + 0.6863 + 0.8535) = 0.8466$$

$$\text{MAP}_{\text{sogou}} = \frac{1}{3} \sum_1^3 (1 + 0.5735 + 0.5995) = 0.7243$$

$$\text{MAP}_{\text{bing}} = \frac{1}{3} \sum_1^3 (0 + 1 + 0.6410) = 0.547$$

With different types of need, different search engines perform differently. As I used some specific queries, the result is not as reliable as estimated.

Trying to explain for the results above, baidu, undoubtedly performs the best as well as stably, while sogou may be affected by some commercial advertisements. Bing can return many relevant documents, though, it missed information from Chinese BBS, like, zhihu, bilibili, etc.