Overview of the TREC-2004 Web Track

Nick Craswell
MSR Cambridge, UK
nickcr@microsoft.com

David Hawking
CSIRO, Australia
david.hawking@csiro.au

1 Introduction

This year's main experiment involved processing a mixed query stream, with an even mix of each query type studied in TREC-2003: 75 homepage finding queries, 75 named page finding queries and 75 topic distillation queries. The goal was to find ranking approaches which work well over the 225 queries, without access to query type labels.

We also ran two small experiments. First, participants were invited to submit classification runs, attempting to correctly label the 225 queries by type. Second, we invited participants to download the new W3C test collection, and think about appropriate experiments for the proposed TREC-2005 Enterprise Track. This is the last year for the Web Track in its current form, it will not run in TREC-2005.

2 Mixed query task

The mixed query task was conducted using the 18 gigabyte, 1.25 million document crawl of the .GOV domain. Last year's tasks involved queries of three types:

Topic distillation The query describes a general topic, e.g. 'electoral college', the system should return homepages of relevant sites.

Homepage finding The query is the name of a site that the user wishes to reach, e.g. 'Togo embassy', and the system should return the URL of that site's homepage at (or near) rank one.

Named page finding The query is the name of a non-homepage that the user wishes to reach, e.g. 'Ireland consular information sheet', and the system should return the URL of that page at (or near) rank one.

There are several possible approaches to dealing with the mixed query stream. One is to find a robust ranking method which works well for all three types. Another is to find specialised methods e.g. one for TD, one for NP and one for HP. Specialised methods could be combined, for example by interleaving ranks or combining scores. Combination can either be done uniformly for all queries or based on query classification, preferring the specialist method which seems most appropriate for the current query.

2.1 Judging and Measures

Since each NP and HP topic is developed with a URL in mind, the only judging task is to identify URLs of equivalent (near-duplicate) pages. For example identifying that `http://xyz.gov/` and `http://xyz.gov/index.html` are equivalent answers. TD judging is more time consuming. Finding URLs which are homepages of relevant sites involves a relevance judgment combined with understanding of site structure, which can be gained by navigating between pages and looking at URL(s).

Judges found 1763 relevant[1] pages: 80 for NP (5 extra), 83 for HP (8 extra) and 1600 for TD. For distillation, the mean number of results per query was $1600/75 = 21.3$, with a median of 13. Topic distillation 2003 had mean 10.3 and median 8. Because there were no major changes in query development and judging methods, we believe the 2003 and 2004 sets are matching and reusable test sets for topic distillation.

We have four measures which we can apply to all query types:

MAP and MRR Mean average precision (MAP) and

[1] Varying the definition of relevant according to the query type.

mean reciprocal rank of the first correct answer (MRR) are standard TREC measures. They are related measures, in that they are exactly equivalent for queries with one correct answer. The problem with applying MAP globally is that some NPHP queries have multiple answers and we only care about the first correct answer. Therefore we apply MAP to TD queries and MRR to NPHP queries. Both measures are calculated on the whole run (1000 ranks), but both put a natural emphasis on the top-ranked documents.

Success@1 The proportion of queries for which a good answer was at rank 1 (the first result the user sees).

Success@5 The proportion of queries for which one or more good answers were in the top 5. The top 5 is what might typically appear on the results page of a web search system, without the user needing to scroll ("above the fold"). If a correct answer appears in the top 5 for 90 of 225 queries, then S@5=0.4.

Success@10 This measure indicates how often a system found something in the top 10, which typically is the first page of web search results. This can also be thought of as a failure measure, because $1 - \text{S@10}$ is the proportion of queries with nothing in the top 10.

We also apply Precision@10 and Recall@1000 to the topic distillation queries.

2.2 Results per query type

Table 3 presents the results for the 75 distillation queries. Considering the MAP and P@10 measures, the top two groups tied, only differing by 0.0011 in MAP and 0.0014 in P@10. Groups 3 and 4 are also very close to each other.

Table 1 has the results for the 75 named page queries. This year's NP MRR scores are higher than last year's, but a striking difference is that the gap between NP and HP has closed. This is illustrated in Figure 1 which, compared to a similar plot last year, has a much smaller gap between HP and NP for the top-scoring runs. This could reflect a better balance between 'relevance' and homepage bias (too much homepage bias hurts NP performance).

Table 2 shows results for HP queries. Although the results are high, they are not as high as last year's best HP

Run	MRR	S@1	S@5	S@10
MSRC04B2S	0.731	0.653	0.827	0.880
MSRAx4	0.685	0.587	0.787	0.853
UAmsT04MSind	0.640	0.507	0.800	0.867
uogWebSelAnL	0.619	0.493	0.787	0.840
THUIRmix045	0.619	0.493	0.787	0.867
MeijiHILw1	0.611	0.480	0.800	0.867
ICT04CIIS1AT	0.606	0.480	0.760	0.880
humW04pl	0.569	0.480	0.667	0.760
wdf3oks0a	0.545	0.413	0.693	0.760
SJTUINCMIX2	0.543	0.387	0.733	0.787
VTOK5	0.511	0.400	0.640	0.733
csiroatnist	0.456	0.320	0.613	0.680
mpi04web08	0.423	0.347	0.507	0.547
MU04web5	0.411	0.333	0.493	0.560
LamMcm1	0.323	0.213	0.440	0.547
fdwiedf0	0.276	0.147	0.453	0.533
irtbow	0.159	0.120	0.173	0.293
XLDBTumba01	0.068	0.067	0.067	0.080

Table 1: Named page results.

Run	MRR	S@1	S@5	S@10
MSRC04C12	0.749	0.653	0.840	0.880
MSRAx2	0.729	0.653	0.867	0.907
UAmsT04MSinu	0.659	0.560	0.760	0.827
THUIRmix045	0.626	0.533	0.733	0.787
uogWebSelAnL	0.625	0.493	0.813	0.840
csiroatnist	0.568	0.467	0.680	0.747
ICT04MNZ3	0.563	0.467	0.653	0.747
MU04web1	0.553	0.467	0.667	0.693
SJTUINCMIX3	0.489	0.400	0.613	0.667
humW04rdpl	0.479	0.373	0.587	0.693
MeijiHILw1	0.473	0.360	0.640	0.680
wdf3oks0brr1	0.421	0.320	0.493	0.640
mpi04web08	0.379	0.307	0.467	0.493
fdwiedf0	0.379	0.333	0.413	0.493
LamMcm1	0.326	0.267	0.413	0.453
VTOK5	0.270	0.173	0.373	0.427
irttil	0.090	0.053	0.120	0.173
XLDBTumba01	0.004	0.000	0.013	0.013

Table 2: Homepage results.

Run	MAP	P@10	R@1000	S@1	S@5	S@10
uogWebCAU150	0.179	0.249	0.777	0.507	0.773	0.893
MSRAmixed1	0.178	0.251	0.815	0.387	0.720	0.880
MSRC04C12	0.165	0.231	0.744	0.387	0.747	0.800
humW04rdpl	0.163	0.231	0.808	0.373	0.787	0.907
THUIRmix042	0.147	0.205	0.761	0.213	0.587	0.747
UAmsT04MWScb	0.146	0.209	0.786	0.360	0.667	0.760
ICT04CIIS1AT	0.141	0.208	0.785	0.333	0.640	0.787
SJTUINCMIX5	0.129	0.189	0.748	0.293	0.573	0.720
MU04web1	0.115	0.199	0.647	0.333	0.640	0.760
MeijiHILw3	0.115	0.153	0.547	0.307	0.547	0.640
csiroatnist	0.111	0.205	0.261	0.320	0.693	0.853
mpi04web01	0.106	0.177	0.453	0.240	0.640	0.787
VTOK5	0.101	0.135	0.721	0.187	0.493	0.533
fdwiedf0	0.090	0.117	0.536	0.293	0.493	0.587
wdf3oks0brr1	0.085	0.124	0.720	0.120	0.413	0.573
LamMcm1	0.049	0.087	0.270	0.173	0.400	0.467
irttil	0.018	0.029	0.147	0.067	0.147	0.173
XLDBTumba01	0.003	0.011	0.008	0.040	0.093	0.107

Table 3: Distillation results.

Run	Average	TD MAP	NP MRR	HP MRR	S@1	S@5	S@10
MSRC04B2S	0.546	0.162	0.731	0.745	0.564	0.809	0.862
MSRAx4	0.527	0.175	0.685	0.721	0.516	0.796	0.871
UAmsT04MSind	0.477	0.133	0.640	0.657	0.453	0.733	0.818
uogWebSelAn	0.466	0.166	0.615	0.617	0.444	0.760	0.818
THUIRmix045	0.457	0.126	0.619	0.626	0.409	0.702	0.778
ICT04MNZ3	0.435	0.137	0.603	0.563	0.440	0.689	0.769
MeijiHILw1	0.398	0.110	0.611	0.473	0.364	0.671	0.738
SJTUINCMIX2	0.385	0.125	0.543	0.487	0.347	0.618	0.689
csiroatnist	0.378	0.111	0.456	0.568	0.369	0.662	0.760
humW04rdpl	0.375	0.163	0.484	0.479	0.369	0.671	0.782
wdf3oks0arr1	0.344	0.085	0.542	0.404	0.276	0.542	0.653
MU04web1	0.343	0.115	0.362	0.553	0.356	0.587	0.662
mpi04web08	0.295	0.082	0.423	0.379	0.298	0.520	0.564
VTOK5	0.294	0.101	0.511	0.270	0.253	0.502	0.564
fdwiedf0	0.248	0.090	0.276	0.379	0.258	0.453	0.538
LamMcm1	0.232	0.049	0.323	0.326	0.218	0.418	0.489
irtbow	0.086	0.012	0.159	0.086	0.071	0.133	0.231
XLDBTumba01	0.025	0.003	0.068	0.004	0.036	0.058	0.067

Table 4: Overall results. Average is the mean of the TD MAP, NP MRR and HP MRR.

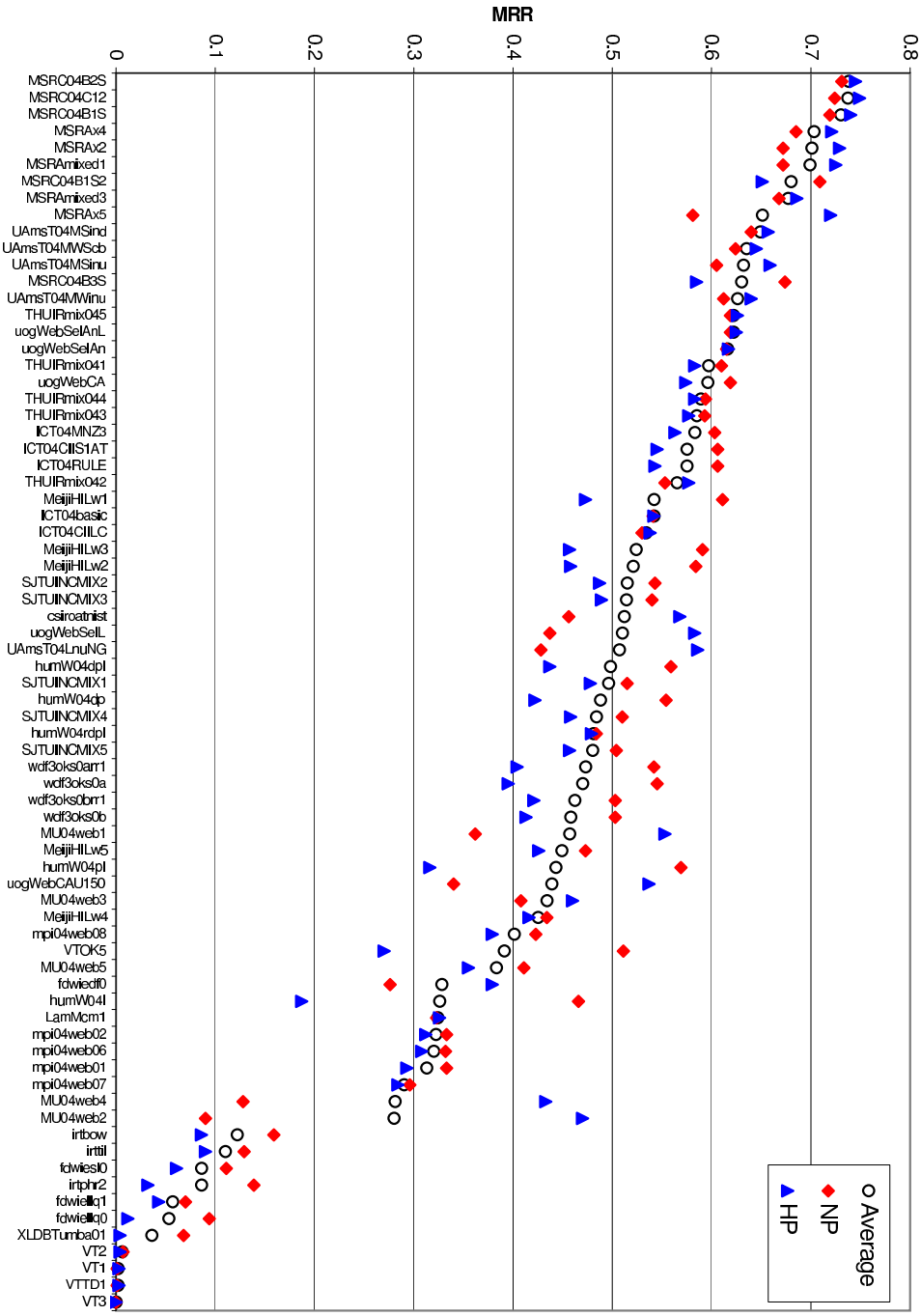

Figure 1: This year the top runs had less of a gap between HP and NP performance (compared to a plot in last year's overview).

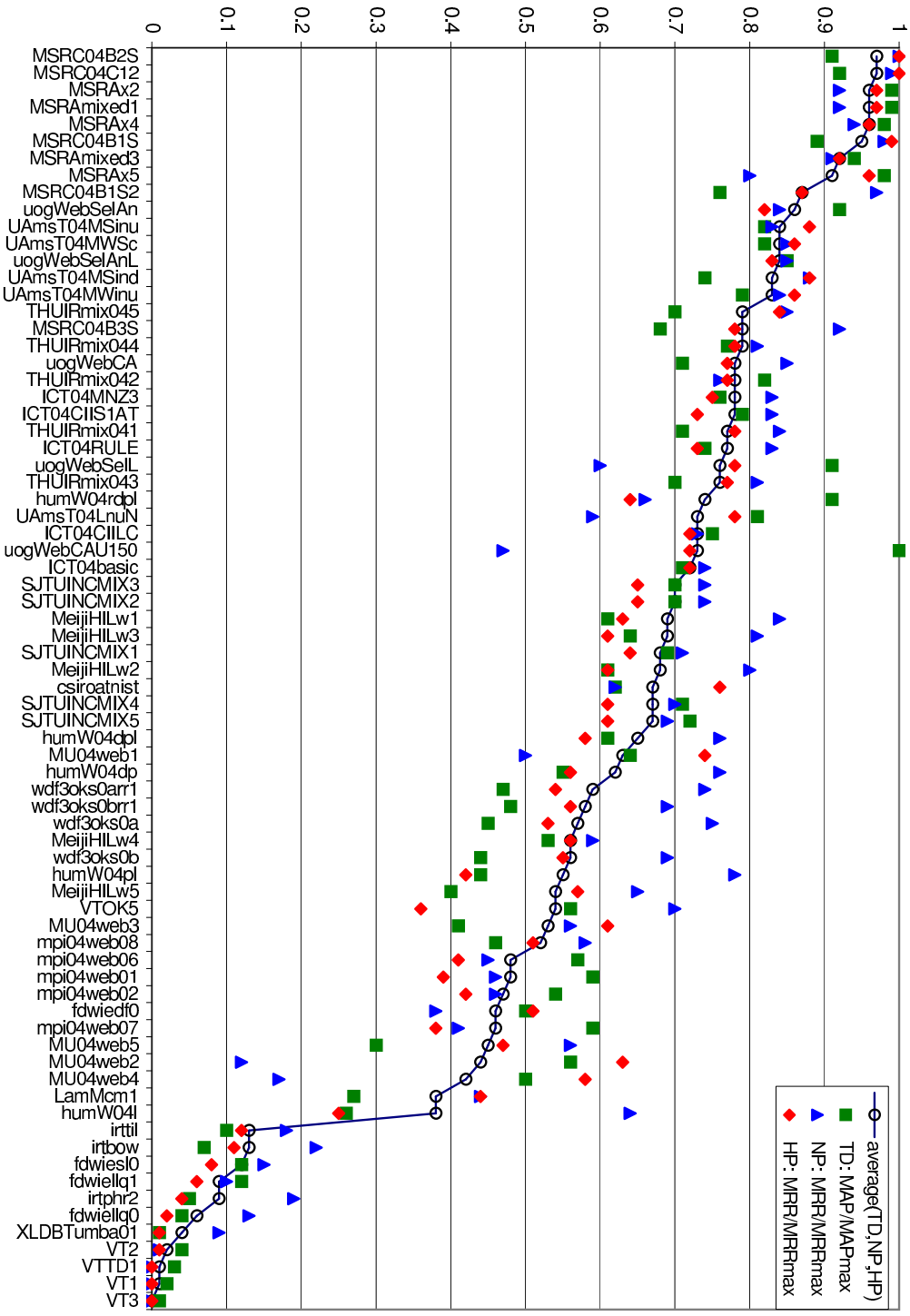

Figure 2: Performance of all runs, based on ratios with the best run of each type.

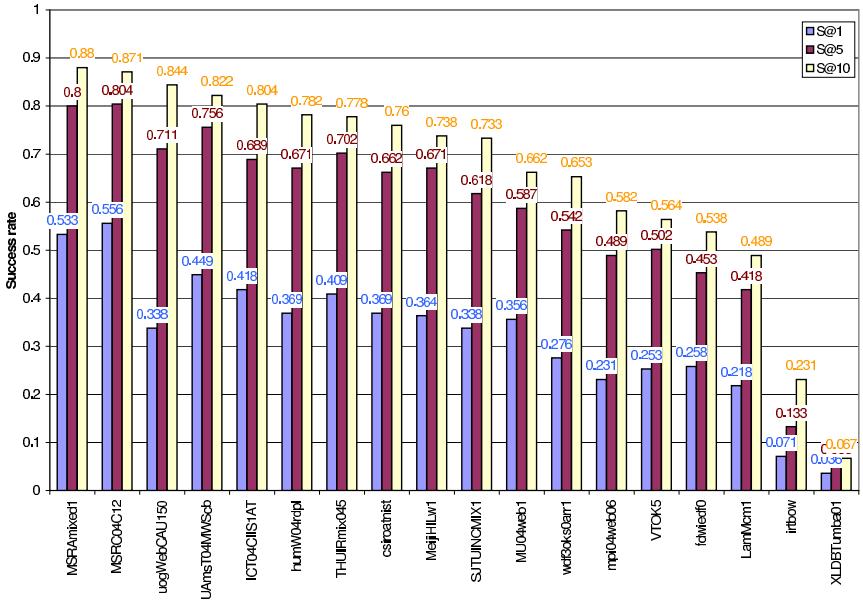

Figure 3: Success rate results. Best run from each group, by S@10.

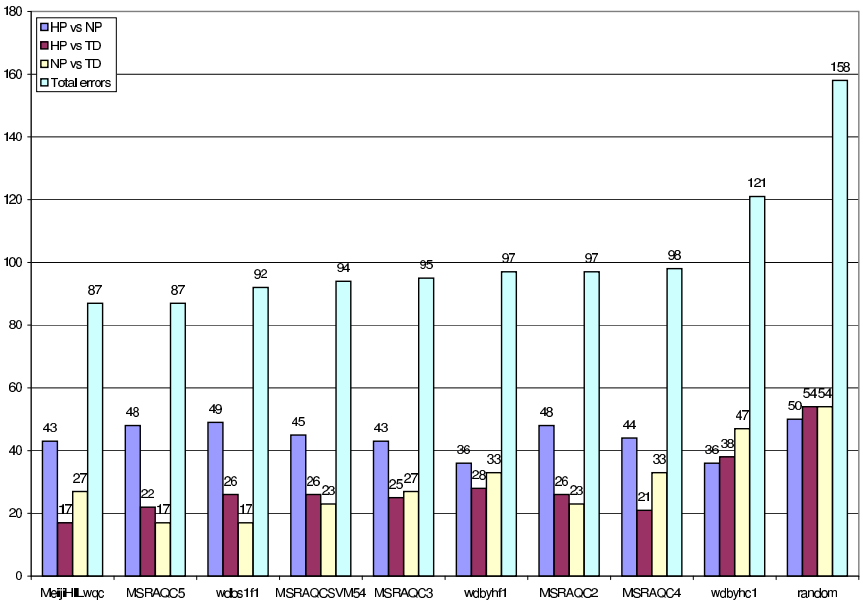

Figure 4: Results of query classification runs. Three types of error and total error.

Run	Avg	TD MAP	NP MRR	HP MRR	Anc	Lnk	Strc	ULen	UOth	QCls
MSRC04C12	0.97	0.92 (0.165)	0.99 (0.724)	1.00 (0.749)	yes	yes	yes	yes	no	no
MSRAx2	0.96	0.99 (0.177)	0.92 (0.672)	0.97 (0.729)	yes	yes	yes	yes	yes	no
uogWebSelAn	0.86	0.92 (0.166)	0.84 (0.615)	0.82 (0.617)	yes	no	yes	yes	no	yes
UAmsT04MWScb	0.84	0.82 (0.146)	0.85 (0.624)	0.86 (0.645)	yes	yes	yes	yes	no	no
THUIRmix045	0.79	0.70 (0.126)	0.85 (0.619)	0.84 (0.626)	yes	no	yes	no	no	no
ICT04CIIS1AT	0.78	0.79 (0.141)	0.83 (0.606)	0.73 (0.545)	yes	no	yes	no	no	no
humW04rdpl	0.74	0.91 (0.163)	0.66 (0.484)	0.64 (0.479)	no	no	yes	yes	yes	no
SJTUINCMIX3	0.70	0.70 (0.125)	0.74 (0.540)	0.65 (0.489)	yes	no	yes	no	no	yes
MeijiHILw1	0.69	0.61 (0.110)	0.84 (0.611)	0.63 (0.473)	yes	yes	yes	yes	no	no
csiroatnist	0.67	0.62 (0.111)	0.62 (0.456)	0.76 (0.568)	yes	yes	yes	yes	yes	no
MU04web1	0.63	0.64 (0.115)	0.50 (0.362)	0.74 (0.553)	yes	yes	yes	yes	yes	no
wdf3oks0arr1	0.59	0.47 (0.085)	0.74 (0.542)	0.54 (0.404)	yes	no	yes	yes	yes	no
VTOK5	0.54	0.56 (0.101)	0.70 (0.511)	0.36 (0.270)	yes	no	yes	no	yes	no
mpi04web08	0.52	0.46 (0.082)	0.58 (0.423)	0.51 (0.379)	yes	yes	yes	yes	yes	no
fdwiedf0	0.46	0.50 (0.090)	0.38 (0.276)	0.51 (0.379)	no	no	no	yes	yes	no
LamMcm1	0.38	0.27 (0.049)	0.44 (0.323)	0.44 (0.326)	yes	yes	yes	yes	yes	no
irtbow	0.13	0.07 (0.012)	0.22 (0.159)	0.11 (0.086)	no	no	no	no	no	no
XLDBTumba01	0.04	0.01 (0.003)	0.09 (0.068)	0.01 (0.004)						

Table 5: Normalised overall results with indication of methods used. Anc: Anchor text used? Lnk: Other link structure used? Strc: Document structure used? ULen: URL length used? UOth: Other URL features used? QCls: Special processing for different query types?

performance, of nearly 0.80. Similarly to last year, S@10 performance seems to max out at around 90%.

2.3 Overall results

Table 4 presents the best run from each group, judged on the average of TD MAP, NP MRR and HP MRR. Although the magnitude for TD is much less than NP and HP, MAP and MRR are related measures so it makes sense to look at the average.

Another way to get an overall score out of TD MAP, NP MRR and HP MRR is to normalise each query type according to the maximum score. This gives each run three scores between 0 and 1, and the average of these three scores is an overall score. Such scores are presented in Figure 2 and Table 5.

A third way to look at the overall result is by success rate. Success at 10 is an interesting number, because it is different from MAP and MRR which give a lot of weight to rank one, and it indicates how often a user reads a whole page of results without finding a good answer. Figure 3 presents success rate figures for the best run from each group, according to S@10 across all queries. The best S@10=0.88 measure gives the user no useful documents for 12% of queries, although perhaps this is acceptable if we assume that in those cases the user reformulates their query.

2.4 What worked

Table 5 indicates which technologies were used by the best run from each group. It is clear that most groups use document structure and many use anchor text. It also seems useful to use link structure and URL length. Other URL features and query classification were not necessary for good performance, but if groups had their best run using such methods they may well be helpful.

We also present information on methods used by the best run from several groups. (Full information is in Appendix A.)

1. **MSRC04C12** Interleaving of stem and nostem runs, each using structure, URL length and PageRank.

3. **MSRAx2** We interpolated relevance scores on the fields of title, body, anchor, url and merged the former four together. The score functions include BM25, proximity and a new proposed URL score function. And the final score combines relevance score and a HostRank that is a PageRank-like value.

10. **uogWebSelAn** Content and anchor-text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme, URL-length reranking, Selecting between content and anchor-text retrieval, or content with anchor-text and URL-length reranking

11. **UAmsT04MWScb** CombMNZ (non-normalized, non-weighted) of stemmed and non-stemmed runs, each using a mixture language model on stemmed full-text, titles, and anchor texts, using both an indegree and URL prior.

16. **THUIRmix045** Word pair weighting based on another run, which used content retrieval in full text and in-link anchor, with a larger weight in fields of Title, head, Bold and first line of page content.

20. **ICT04CIIS1AT** Anchor text forward propagation, page title text back propagation, combination of anchor text ,key words ,h1 text etc. ,different pivoted weigth function for different part

27. **humW04rdpl** Plain content search including linguistic expansion from English inflectional stemming, extra weight on properties such as Title and Metadata, lower url depth and root urls

3 Query classification runs

Three groups submitted a total of 9 query classification runs. Results are presented in Figure 4. Random classification of 225 queries into three types would tend to lead to about 150 errors, so classification runs were able to do significantly better than random. The best run MeijiHILwqc was a manual run. The most common type of error was confusing HP and NP (either by classifying HP as NP or classifying NP as HP).

4 W3C Investigation

Workshop participants proposed a variety of new experiments, for example relevance ranking in email, or searching for people who are experts in a particular topic area. We plan to pursue such ideas using the W3C dataset in the TREC-2005 Enterprise Track.

5 Conclusion

The main experiment showed that, on a mixed query set, effective retrieval is possible without query classification. Topic distillation is still by far the most difficult query type. Query classification runs showed that it is indeed possible to tell the difference. The most common classification mistake was to confuse NP and HP queries.

The other effect of the mixed query task is to consolidate the findings of previous Web Track years. There are web search information needs which are based on a page's position (a 'homepage') and importance, rather than just the page's text. To answer these information needs, it is not sufficient to search on content alone: use of 'Web evidence' based on structure, links and URLs is necessary. This evidence may be effectively used in an enterprise-scale crawl, of a million pages. The Web Track collections are now reusable resources for new experiments with TD, NP, HP and mixed query streams.

Of course there is also more work to be done in developing evaluation methodologies. Future web experiments could model other user needs, for example transactional search, and refine solutions to tricky issues such as distillation judging and scoring of near-duplicate results. Another direction would be to venture into the wider Web, where adversarial information retrieval is an issue, and many pages are there to manipulate the ranking rather than provide useful information. These can be eliminated or down-weighted via analysis at crawl time or query time. Finally, having so far considered enterprise-scale webs in the Web Track, it is interesting consider ranking with other forms of enterprise information such as mailing list archives and document shares/archives, and a search across a mixture of web and non-web enterprise data.

A All run descriptions

The a description of each run as submitted, sorted as in Figure 2. Each group's best run is marked with a *.

1. **MSRC04C12*** Interleaving submissions MSRC04B1S and MSRC04B2S
2. **MSRC04B2S** Weighted Field BM25 (fields title, body & anchor) optimised on the Named Page 2003 task, with linear adition of non-linear PageRank and URL features. Stemming.
3. **MSRAx2*** relevance propagation + HostRank (more details in Section 2.4 above)
4. **MSRAmixed1** fields weighting + proximity + a new importance named HostRank
5. **MSRAx4** URL match and level + BM25 + HostRank
6. **MSRC04B1S** Weighted Field BM25 (fields title, body & anchor) optimised on the Named Page 2003 task, with linear adition of non-linear PageRank and URL features. No stemming.
7. **MSRAmixed3** BM2500 + Proximity
8. **MSRAx5** relevance propagation + HostRank
9. **MSRC04B1S2** Weighted Field BM25 (fields title, body & anchor) optimised on the Topic Distillation 2003 task, with linear adition of non-linear Click-Distance and URL features. No stemming.
10. **uogWebSelAn*** content and anchor-text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme, URL-length reranking, Selecting between content and anchor-text retrieval, or content with anchor-text and URL-length reranking
11. **UAmsT04MWScb*** CombMNZ (non-normalized, non-weighted) of runs UAmsT04MWinu and UAmsT04MSinu.
12. **uogWebSelAnL** content and anchor-text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme, URL-length reranking, Selecting between content and anchor-text retrieval, or content with anchor-text and URL-length reranking
13. **UAmsT04MSinu** Mixture language model on stemmed full-text, titles, and anchor texts, using both an indegree and URL prior.
14. **UAmsT04MSind** Mixture language model on stemmed full-text, titles, and anchor texts, using an indegree prior.
15. **UAmsT04MWinu** Mixture language model on non-stemmed full-text, titles, and anchor texts, using both an indegree and URL prior.
16. **THUIRmix045*** Word pair weighting based on THUIRmix041.
17. **MSRC04B3S** Weighted Field BM25 (fields title, body & anchor) optimised on the Topic Distillation 2003 task, with linear adition of non-linear Click-Distance No stemming.
18. **THUIRmix044** Query classification with query length and named entity information. TD topics are assigned to THUIRmix042, while the others are retrieved on THUIRmix041.
19. **THUIRmix042** Content retrieval in full text and in-link anchor of Key resource pages. Key resource pages are selected with non-content features using clustering technologies.
20. **ICT04CIIS1AT*** anchor text forward propagation , page title text back propagation, combination of anchor text ,key words ,h1 text etc. ,different pivoted weigth function for different part
21. **ICT04MNZ3** CombMNZ for combination of anchor text retrieval result ,structure info retrieval result and content retrieval result. anchor text forward propagation , page title text back propagation.
22. **uogWebCA** content and anchor text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme
23. **THUIRmix041** Content retrieval in full text and in-link anchor, with a larger weight in fields of Title, head, Bold and first line of page content.
24. **ICT04RULE** rerank the result by some heuristic strategies make use of the url depth,url works,anchkor text, site compression like trick.
25. **uogWebSelL** content and anchor-text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme, URL-length reranking, Selecting between content and anchor-text retrieval, or content with anchor-text and URL-length reranking
26. **THUIRmix043** THUIRmix041 + primary space model weighting in in-link anchor text and contents of Title, head, Bold and first line of page content.
27. **humW04rdpl*** same as humW04dpl except extra weight for root urls
28. **ICT04CIILC** comparable run with ICT04basic, using a different weighted function for Content text, others just the same as ICT04basic
29. **uogWebCAU150** content and anchor text retrieval, Porter Stemming, Divergence From Randomness PL2 weighting scheme, URL-length reranking
30. **UAmsT04LnuNG** Lnu.ltc run with word n-gram boosting, using document structure and anchor texts.
31. **ICT04basic** vector space content model, baseline for all the runs, using combination of anchor text and some simplest page structure info. not stems,not feedback and classification of queries
32. **SJTUINCMIX3*** BM25
33. **SJTUINCMIX2** Task classification,BM25
34. **MeijiHILw1*** Vector space model. Using anchor text, url-depth and title text. Outdegree reranking.
35. **MeijiHILw3** Vector space model. Using anchor text, url-depth and title text. Outdegree reranking. Query Classified based on last year's queries. Document vector modification by Relevance-based Superimposition Model(RSModel).
36. **SJTUINCMIX1** task classification,BM25,minimal span weighting reRank
37. **MeijiHILw2** Vector space model. Using anchor text, url-depth and title text. Outdegree reranking. Query Classified based on last year's queries.
38. **SJTUINCMIX5** Task classification,BM25,Site Unit
39. **SJTUINCMIX4** Task classification,BM25,PageRank reRank
40. **csiroatnist*** This is a baseline run obtained by submitting the query titles to the Panoptic (CSIRO software) search service at ir.nist.gov. Note that an error with topic 179 resulted in no documents retrieved. To pass the submission checking script, the 30th result for topic 178 was arbitrarily inserted as the first for 179.
41. **humW04dpl** same as humW04pl except extra weight for lower url depth
42. **MU04web1*** Vector Space Model + Document-centric impact + pagerank + URL depth
43. **humW04dp** same as humW04dpl except linguistic expansion from stemming disabled
44. **wdf3oks0arr1*** result merging, okapi, simple stemmer, homepage rank boosting
45. **wdf3oks0brr1** result merging, okapi, combo stemmer, homepage rank boosting
46. **wdf3oks0a** result merging, okapi, simple stemmer
47. **MeijiHILw4** Vector space model. Using anchor text, url-depth and title text. Outdegree reranking. Query Classified based on last year's queries.Query expansion using Conceptual Fuzzy Sets(CFS).
48. **wdf3oks0b** result merging, okapi, combo stemmer
49. **humW04pl** same as humW04l except extra weight on properties such as Title and Metadata
50. **VTOK5*** BASELINE
51. **MeijiHILw5** Vector space model. Using anchor text, url-depth and title text. Outdegree reranking. Query Classified based on last year's queries.Query expansion using Conceptual Fuzzy Sets(CFS). Document vector modification by Relevance-based Superimposition Model(RSModel).
52. **MU04web3** Vector Space Model + Document-centric impacts + Pagerank
53. **mpi04web08*** Automatic phrase detection, Anchor text reranking, PageRank, Stemming
54. **mpi04web01** our baseline plain keyword queries from title PageRank Stemming
55. **mpi04web06** Autmatic query expansion + phrase detection PageRank Stemming
56. **mpi04web02** Autmatic query expansion + phrase detection PageRank Stemming
57. **fdwiedf0*** hammingbird algorithm
58. **mpi04web07** Automatic phrase detection, PageRank, Stemming
59. **MU04web5** Vector space model + document-centric impacts
60. **MU04web2** Vector Space Model + Document-centric impacts + URL depth
61. **MU04web4** Vector space model + document-centric impact + pagerank + URL depth
62. **LamMcm1*** Multicriteria analysis Lovins Stemming Kleinberg authority scores
63. **humW04l** plain content search including linguistic expansion from English inflectional stemming
64. **irtbow*** bag of words but with added weighting for query term order and proximity; Lnu.Ltc weighting.
65. **irttil** title only; Lnu.Ltc weighting
66. **fdwiesl0** improved okpai method
67. **irtphr2** phrase search (not useful for single-term queries); Lnu.Ltc weighting.
68. **fdwiellq1** anchro-text ranking
69. **fdwiellq0** okpai model
70. **XLDBTumba01***
71. **VT2** Ranking tuning using linear fusion
72. **VTTD1** TD tuning
73. **VT1** best trial
74. **VT3** Ranking tuning using linear fusion