

Query 19:

“AGREEMENT BY THE UNITED ARAB REPUBLIC AND SAUDI ARABIA TO WITHDRAW THEIR FORCES FROM YEMEN, WHICH INVOLVES OBSERVERS FROM THE UNITED NATIONS EXPEDITIONARY FORCE BEING SENT TO YEMEN .”

Iteration 1:

User Feedback system

Precision – 0.5

Recall – 0.5

MAP – 0.69

Positive feedback docs provided – 99, 100

Negative feedback docs provided – 36, 86

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.8

MAP – 0.69

Iteration 2:

User Feedback system

Precision – 0.5

Recall – 0.5

MAP – 0.7675

Positive feedback docs provided – 100, 195

Negative feedback docs provided – 121, 55

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.8

MAP – 0.4825

Iteration 3:

User Feedback system

Precision – 0.5

Recall – 0.5

MAP – 0.95

Positive feedback docs provided – 195, 267

Negative feedback docs provided – 115, 159

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.8

MAP – 0.4825

Iteration 4:

User Feedback system

Precision – 0.625

Recall – 0.625

MAP – 0.975

Positive feedback docs provided – 99, 100, 344

Negative feedback docs provided – 55, 121

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.8

MAP – 0.4825

Iteration 5:

User Feedback system

Precision – 0.625

Recall – 0.625

MAP – 1

Positive feedback docs provided – 195, 267

Negative feedback docs provided – 55

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.8

MAP – 0.4825

Comparing the systems – Query 19

When comparing the MAP value, precision, and recall between the user feedback system and the pseudo relevance feedback system for query 19 initially the MAP value and precision are the same, and the recall is significantly better in the pseudo system. As the iterations continue precision and recall remain stagnant to their respective feedback types, but the MAP value decreases significantly in the pseudo system, to the point that it is less than half of what the value is for the user feedback system.

Query 31:

“LEADERS WHICH FIGURE IN DISCUSSIONS OF THE FUTURE OF THE WEST GERMAN CHANCELLORSHIP .”

Iteration 1:

User Feedback system

Precision – 0.3

Recall – 0.43

MAP – 0.512

Positive feedback docs provided – 47, 56

Negative feedback docs provided – 39, 9

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.3

Recall – 0.43

MAP – 0.512

Iteration 2:

User Feedback system

Precision – 0.2

Recall – 0.29

MAP – 1

Positive feedback docs provided – 47, 56

Negative feedback docs provided – 48, 290

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.6

Recall – 0.86

MAP – 0.67

Iteration 3:

User Feedback system

Precision – 0.2

Recall – 0.29

MAP – 1

Positive feedback docs provided – 47, 56

Negative feedback docs provided – 48, 202, 121, 1, 53

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.6

Recall – 0.86

MAP – 0.67

Iteration 4:

User Feedback system

Precision – 0.2

Recall – 0.29

MAP – 1

Positive feedback docs provided – 47, 56

Negative feedback docs provided – 48, 202, 121, 1, 53, 157, 343, 290

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.6

Recall – 0.86

MAP – 0.67

Iteration 5:

User Feedback system

Precision – 0.2

Recall – 0.29

MAP – 1

Positive feedback docs provided – 47, 56

Negative feedback docs provided – 48, 202

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.6

Recall – 0.86

MAP – 0.67

Query Drift

This query experienced query drift. Because the “user” did not input the details of all the non-relevant documents those non-relevant documents were continuously being returned. After the first iteration the relevant document 81 never reared its head again even though it was the second ranking document, and highest-ranking relevant document at iteration one. This is because document 81 was not marked as relevant in the first iteration. After iteration one none of the relevant documents other than 47 and 56 showed up in the top k (10) documents, but 47 and 56 were the top two documents for iterations 2-5.

Comparing the systems – Query 31

In this situation recall, precision, and MAP are the same for both methods on the first iteration. On iteration 2 precision and recall are about 3 times greater using the pseudo relevance feedback method, but the MAP value is one and a half times greater using the user feedback method. This stays consistent throughout all 5 iterations.

Query 58:

“OTHER NATIONS POSSESSING U.S . POLARIS MISSILES FOR THEIR NUCLEAR SUBMARINE FLEETS .”

Iteration 1:

User Feedback system

Precision – 0.5

Recall – 0.75

MAP – 1

Positive feedback docs provided – 135, 89, 1

Negative feedback docs provided – 38, 148

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.5

Recall – 0.75

MAP – 1

Iteration 2:

User Feedback system

Precision – 0.58

Recall – 0.875

MAP – 1

Positive feedback docs provided – 135, 247, 47

Negative feedback docs provided – 402, 53

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.67

Recall – 1

MAP – 0.912

Iteration 3:

User Feedback system

Precision – 0.58

Recall – 0.875

MAP – 1

Positive feedback docs provided – 1, 247, 135

Negative feedback docs provided – 121

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.67

Recall – 1

MAP – 0.912

Iteration 4:

User Feedback system

Precision – 0.58

Recall – 0.875

MAP – 1

Positive feedback docs provided – 47, 89, 1, 247

Negative feedback docs provided – 121, 48, 228

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.67

Recall – 1

MAP – 0.912

Iteration 5:

User Feedback system

Precision – 0.58

Recall – 0.875

MAP – 1

Positive feedback docs provided – 47

Negative feedback docs provided – 121

New query and weights – (see results.txt)

Pseudo Relevance System

Precision – 0.67

Recall – 1

MAP – 0.912

Comparing the systems – Query 58

As with all of the queries I have tested today, iteration 1 obviously has the exact same values for precision, recall, and MAP. Iterations 2-5 have better precision, and total recall when using the pseudo relevance feedback method. Iterations 2-5 have a slightly lower MAP when using the pseudo relevance feedback method, but the fact that user feedback method has a MAP of 1 and the pseudo relevance method has a MAP of 0.912 means that their differences are essentially negligible.

Overall comparison of the systems

I would say that the pseudo relevance feedback system is the right system to use if you are looking for maximum recall and/or precision values. If you are looking for a system that presents the most relevant documents as the first several documents returned by a query, then the user feedback system would be better.

Earlier this semester we had a question about what type of search results an attorney who is paid by the hour would prefer. Since that attorney would prefer the best relevance and to a lesser extent precision, the pseudo relevance system would be best.

A user who is casually searching for information would probably prefer the results of the user feedback system, though. This system has a higher MAP value, which means that the more relevant documents are presented first. Most users click on the first few links, so even though they might be getting documents at later positions that are no relevant, they likely will not click those links anyways.







