

Advantages of Query Biased Summaries in Information Retrieval

by Anastasios Tombros and Mark Sanderson

Elif Eser

Caner Mercan

Outline

- Introduction
- System Overview
- Experiments
- Conclusion

Introduction

- Typical IR systems have predefined static summary of the document as output.
 - Users cannot get a clear view from the title & first sentences.
- Query biased summaries! ← Contribution!
 - Users see the parts of the documents related to the query terms.

Typical Output of an IR System

727 results returned, ranked by relevance.

☐ Reviewed Sites Only ☐ Green Light sites only
☒ The Entire Web [Search Tips and Help](#)

[Landings: Aircraft Manufacturers](#) 91% [\[find similar\]](#)
Aviation's Directory Aerobatics-Flying . Aircraft-Sales . Aircraft-Service/Parts . Aircraft-Manufacturers . A/C-Values .
Airlines . Airports . Airshows/Events . Aviation-BBSs . Aviation-Images . Aviation-Lists . Aviation-News-Groups .
Avionics . Classifieds . Companies . Databases . Flight-Schools/FBOs . Flying-Clubs . Flight-Planning .
General-Aviation . Government/Research . GPS.
http://www.landings.com/_landings/pages/aircraft-manuf.html

[Landings: Avionics](#) 90% [\[find similar\]](#)
Digifly USA Icc aviation electronic instruments GPS moving map, flight and engine instruments. Skylight Avionics
Manufacturer of electronics interfaces and cockpit information displays for business and commuter aircraft.
http://www.landings.com/_landings/pages/avionics.html

[Landings: Aviation related companies on ...](#) 90% [\[find similar\]](#)
Aero Appraisal Service: Is an aircraft appraisal company based in Dallas, Texas. Maryland-based FBO with an FAA
approved 141 flight school, offering full-service maintenance, fuel, aircraft rentals and charters, computerized testing
and weather services.
http://www.landings.com/_landings/pages/companies.html

[Landings: Aviation Image Archives: airp...](#) 90% [\[find similar\]](#)
Russian Aircraft Resource: Information about Russian aircraft with pictures, etc. Russian Aviation Museum
Information about over 500 Russian aircraft.
http://www.landings.com/_landings/pages/images.html

Summarization System

- Title Method
 - Terms that occur in the title section of the documents are assigned a positive weight.
- Location Method
 - An ordinal weight is assigned to the first two sentences of each articles.
 - A heading score is assigned to each one of the sentences comprising a heading.

Summarization System

- Term Occurrence Information
 - The system locates clusters of significant words within sentences and assign scores to them.
- Biasing Summaries Towards Queries
 - The larger the number of query terms in a sentence, the more likely that sentence conveys a significant amount of the information.

Experimental Design

- Two groups consisting of 10 subjects
 - One of them for static summaries
 - The other one for query biased summaries
- Experimental sessions with the same environment
- 50 queries from TREC test collection

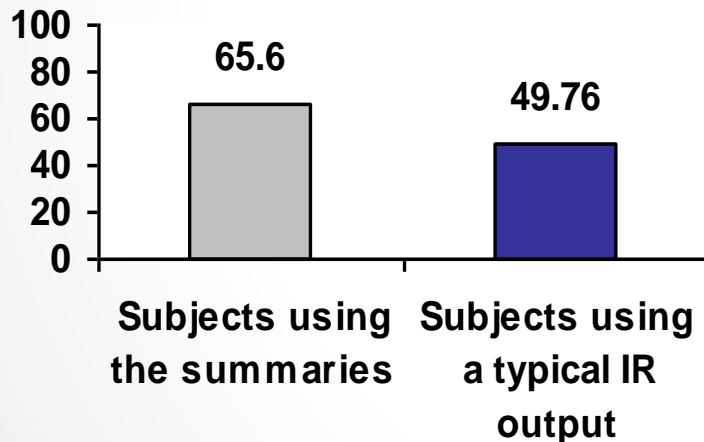
Aim of the Experiment

- To measure how query-biased summaries affect:
 - the accuracy of the judgements (Precision & Recall)
 - the speed of the users' judgements
 - the need to refer to the full text of the documents
 - the opinion of the subjects for the IR system they are using

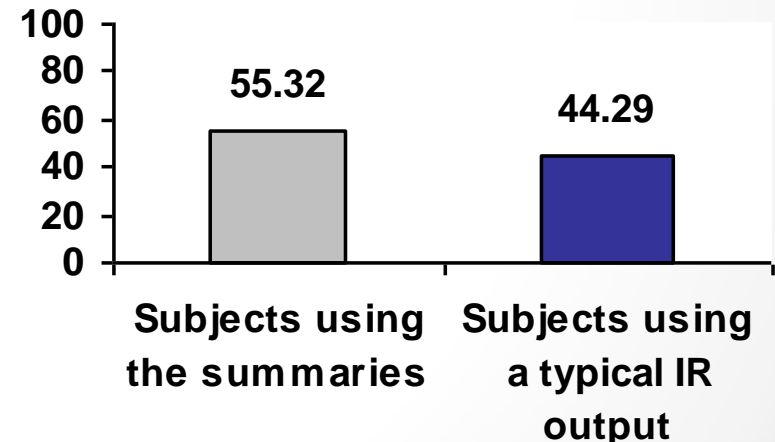
Recall and Precision

$recall = \frac{\text{number of relevant docs correctly identified for a query}}{\text{total number of relevant docs for that query}}$

$precision = \frac{\text{number of relevant docs correctly identified for a query}}{\text{total number of marked relevant docs for that query}}$

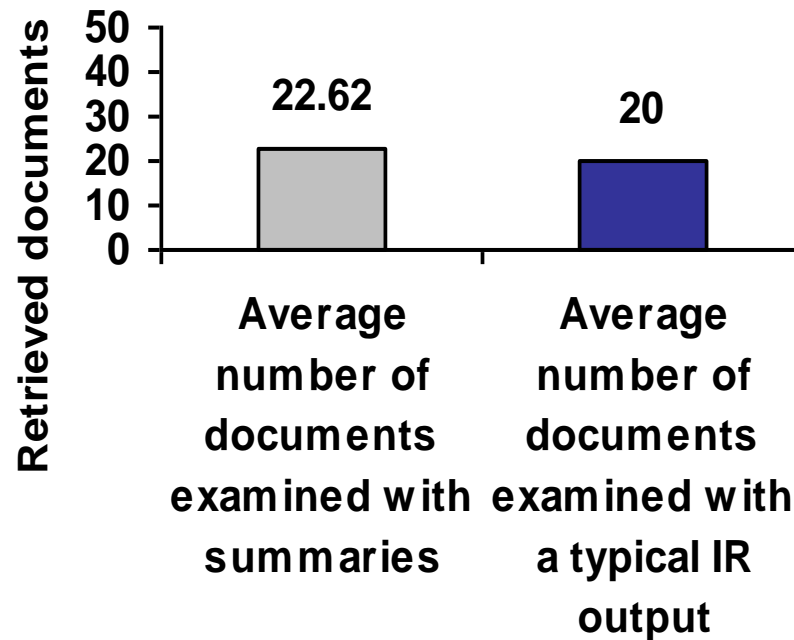


Recall values of 2 groups



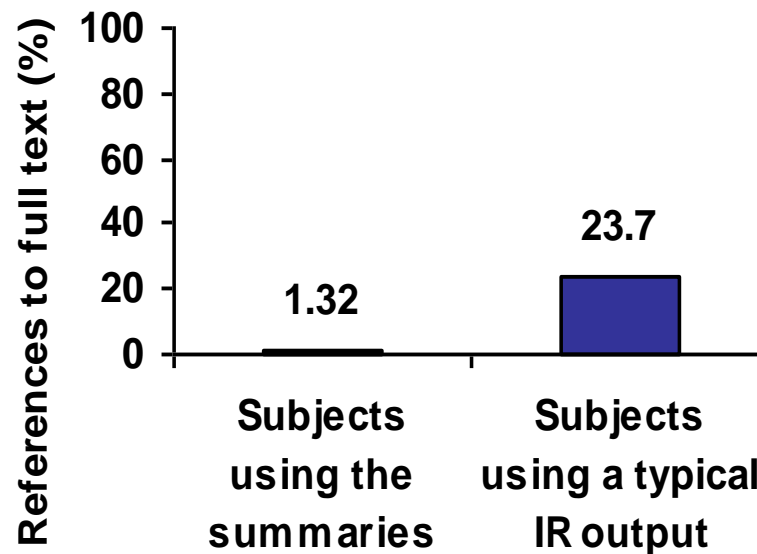
Precision values of 2 groups

Speed of Judgements



The number of examined documents for each group

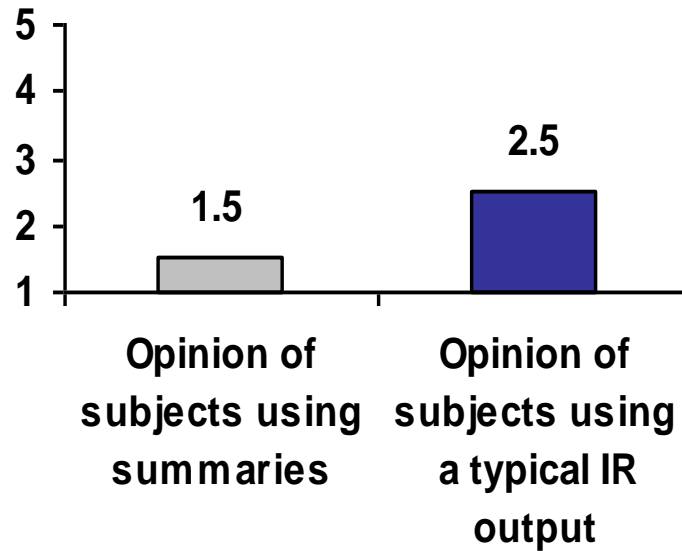
Reference to the Full Text of the Documents



Average number of references to the full text of documents per query

Opinion of the Users

1: positive opinion
5: negative opinion



Average number of subjective opinion of the users

Conclusion and Future Works

- Users need more relevance clues from IR systems. With query biased systems, they can get more relevant documents without checking the whole retrieved documents.
- As future work
 - These summarization method can be applied to web search engines
 - Experiments can be done with different approaches to summarisation