An Implicit System for Predicting Interests

Ryen W. White

Department of Computing Science
University of Glasgow
Glasgow, Scotland, G12 8RZ.
ryen@dcs.gla.ac.uk

Joemon M. Jose

Department of Computing Science
University of Glasgow
Glasgow, Scotland, G12 8RZ.
jj@dcs.gla.ac.uk

ABSTRACT

We demonstrate an adaptive search system that works proactively to help searchers find relevant information. The system observes searcher interaction, uses what it sees to model information needs and chooses additional query terms. The system watches for changes in the topic of the search and selects retrieval strategies that reflect the extent to which the topic is seen to change.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**];

General Terms

Design, Human Factors

1. INTRODUCTION

Searchers often cannot devise queries that express their needs, are reluctant to provide explicit indications on what meets these needs and may find documents too large to give feedback accurately. Traditional relevance feedback systems suffer because they only address the first of these problems. Our system uses implicit feedback methods to help searchers locate useful information. It makes inferences from searcher behaviour, builds a model of their need and uses this model to predict their interests and changes in these interests.

2. THE SYSTEM

The system monitors interaction with a content-rich search interface designed to generate evidence for the techniques it employs. The interface (shown in Figure 1) uses a variety of document representations, all related to the initial query. These include a list of *top-ranking sentences*, extracted from the top documents and ranked independently of source: titles, summaries, individual sentences from the summaries and these sentences in their original document context. These representations combine to form an interactive search path at the interface. These paths and representations are similar to the interface components presented in White et al. [1].

The system is divided into two parts; *need prediction* and *need tracking*.

Need prediction: The system monitors searcher interaction with the results interface, and chooses terms that describe the information viewed. We use a term selection model based on *Jeffrey's rule of conditioning* (presented in earlier work [2]) to choose terms that represent this information and could be useful to improve the original query statements.
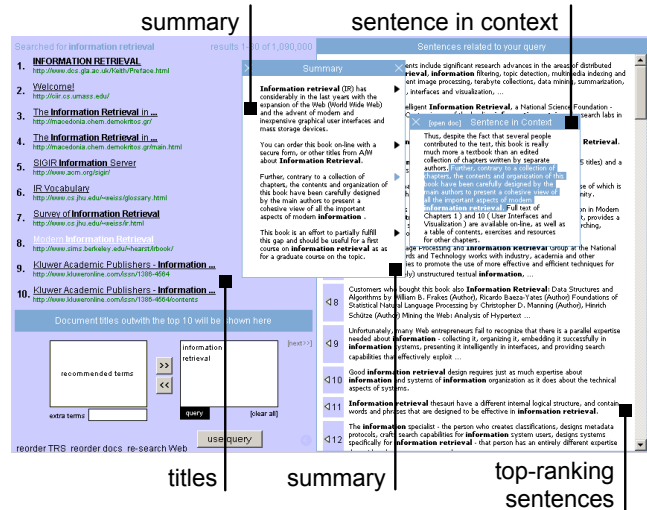
Figure 1. Search interface

Need tracking: During a search session the system passively observes the searcher, watching for changes in the information viewed. The system takes 'snapshots' of the information need during the course of the search. Each shot comprises a ranked list of terms that reflects the need at that point. This list is ranked based on the scores generated by the need prediction component.

It compares each shot with the need near the beginning of the search. For this comparison it uses the Pearson product moment correlation coefficient, and the statistical significance of the coefficient as a decision metric. The system uses this metric to decide between four possible retrieval strategies that reflect the perceived degree of search topic change: *re-search Web*, *reorder top-ranked documents*, *reorder top-ranking sentences* and *no action*.

In this demonstration we present a semi-automatic version of our system that recommends the most useful additional terms and retrieval strategies to the searcher. In this demonstration we work with the live Internet and give participants a chance to use their choice of query. We foresee that this will facilitate a clear exposition of our system.

3. REFERENCES

[1] White, R.W., Jose, J.M. and Ruthven, I. (2003) 'An Approach for Implicitly Detecting Information Needs'. *Proceedings of the 12th ACM CIKM*, 504-507.

[2] White, R.W., Jose, J.M., Van Rijsbergen, C.J. and Ruthven, I. (2004) 'A Simulated Study of Implicit Feedback Models'. *Proceedings of the 26th European Conference on Information Retrieval*, 310-325.