Searching for Expertise using the Terrier Platform

Craig Macdonald, Iadh Ounis
Department of Computing Science
University of Glasgow, Glasgow, Scotland, UK
{craigm, ounis}@dcs.gla.ac.uk

Categories and Subject Descriptors:
H3.3 **[Information Storage and Retrieval]:** Information Search and Retrieval. H3.4 **[Information Storage and Retrieval]:** Systems and Software – *User profiles and alert services*.

General Terms: Algorithms, Design

Keywords: Expert Finding, Expert Search Information Retrieval, Terrier, Expertise Modelling

1. EXTENDED ABSTRACT

In large Enterprise organisations, users often seek other members of their organisation to collaborate with. The use of an *expert search* system can help users in finding people that have relevant expertise to their need.

In the TREC 2005 Enterprise track, we developed an expert search system based on the Terrier IR platform [1]. In this work, we demonstrate an improved version of our technology, applied in a different Enterprise setting, namely the University of Glasgow, Dept. of Computing Science.

From the intranet of the organisation, we crawl all documents, such as mailing lists, staff Web homepages, CVs, research publications and administrative and teaching material. Using the staff list, our expert search system builds a profile of expertise evidence for each candidate expert, by associating documents to each candidate.

The expert search system uses a retrieval methodology based on the Divergence from Randomness framework to rank candidate profiles in response to a user query. Each document is represented by a set of fields (e.g. Content, Title and Anchor text of incoming hyperlinks), and the used matching function takes these various fields of the documents into account when ranking candidate profiles. Based on our experience from the TREC 2005 Enterprise track, we assign different importance weights on the various sources of evidence. For example, the homepage or CV of a candidate are good sources of expertise evidence.

Figure 1 presents the results for a typical query to the expert search system. A ranking of candidate experts for the query is displayed, along with a concise description of their job interests. In contrast to [2], the top-ranked documents for each candidate are also displayed, as found by the retrieval methodology. These related documents are important, as they allow the user to quickly assess the relevance of a candidate by scanning the related documents. This is similar to a user scanning summaries to assess the relevance of documents in classical retrieval.

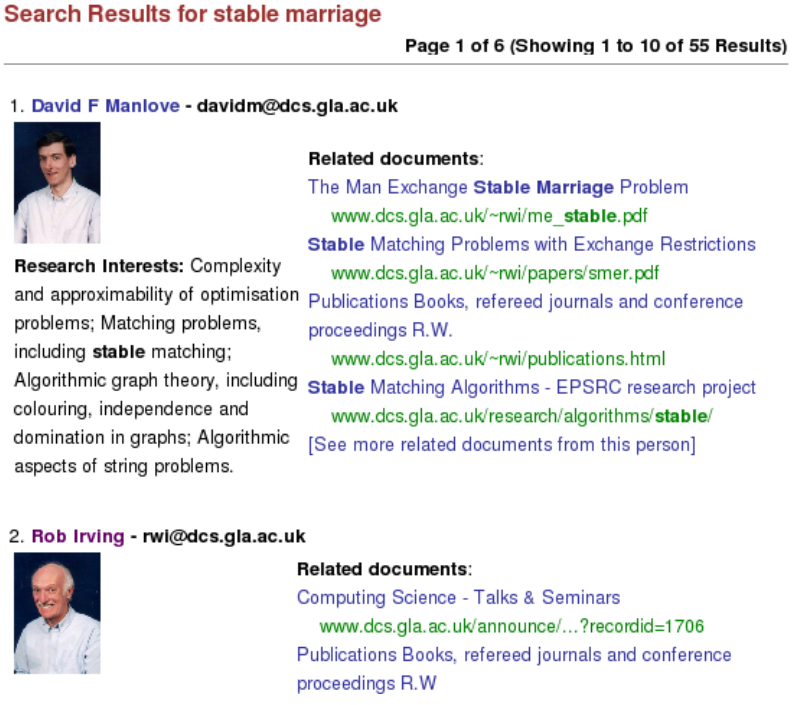

Figure 1: Results presentation for an expert search query, showing the ranked candidates and their related documents. The query in this case is 'stable marriage'.

2. REFERENCES

[1] C. Macdonald, V. Plachouras, B. He, and I. Ounis. University of Glasgow at TREC 2005: Experiments in Terabyte and Enterprise tracks with Terrier. In *TREC-2005 Proc*.

[2] M. Maybury, R. D'Amore, and D. House. Automating Expert Finding. In *Intern. J. Technology Research and Manage*. 43,6 (Nov-Dec 2000), 12-15

SIGIR'06, August 6-11, 2006, Seattle, Washington, USA.
ACM 1-59593-369-7/06/0008.