

PAS: a Personal Alert System for Information Retrieval in CRISs

Germán Hurtado Martín^{1,2}, Chris Cornelis²

¹ *Department of Industrial Sciences, Hogeschool Gent
Schoonmeersstraat 52, 9000 Gent, Belgium
E-mail: German.HurtadoMartin@HoGent.be*

² *Department of Applied Mathematics and Computer Science, Ghent University
Krijgslaan 281 (S9), 9000 Gent, Belgium
E-mail: Chris.Cornelis@UGent.be*

Current Research Information Systems (CRISs) usually contain large amounts of heterogeneous and distributed data, which makes finding specific information difficult for a user. Leaving aside problems like the fact that users are not always able to express their information need precisely, there is also the complication that most CRISs are constantly being updated. As a consequence, users may not even be aware of information relevant to them, let alone be able to compose an effective query for it. In order to prevent the user from missing potentially useful information, the concept of a personal search agent, proactively informing the user about newly available information, becomes more and more popular.

At Hogeschool Gent, a Personal Alert System (PAS) is being developed along these lines. PAS combines profiles of the staff and their research activities, with documents about funding possibilities to support research. Both information sources are mapped to a common ontology, which is currently a version of the three-level IWETO taxonomy [1], enriched with a fourth level containing free keywords obtained from staff profiles. The main goal of the system is then to alert users whenever a funding opportunity can be matched to their research interests by the search agent. Flexibility in the matching process is obtained by using fuzzy [2] and rough [3] set theories: the fuzzy component allows for expressing partial relationships (between terms in the taxonomy, as well as between users/funding documents and taxonomy terms), while the rough component provides mechanisms for query expansion: a user profile and a funding document may still be matched when they refer to different, but related, keywords. Currently, a basic prototype has already been developed, based on the algorithms in [4], and in which the user can also influence term relations through an easy feedback process.

References

- [1] URL: <http://www.iweto.be>
- [2] L. A. Zadeh: *Fuzzy Sets*, Information and Control, 8 (1965) 338–353
- [3] Z. Pawlak: *Rough Sets*, International Journal of Computer and Information Sciences, 11(5) (1982) 341-356
- [4] P. Srinivasan, M. E. Ruiz, D. H. Kraft, J. Chen: *Vocabulary mining for information retrieval: rough sets and fuzzy sets*, Information Processing and Management, 37(1) (2001) 15–38.