

Chapter XVI

User Relevance Feedback in Semantic Information Retrieval

Antonio Picariello

Università di Napoli Federico II, Italy

Antonio M. Rinaldi

Università di Napoli Federico II, Italy

ABSTRACT

The user dimension is a crucial component in the information retrieval process and for this reason it must be taken into account in planning and technique implementation in information retrieval systems. In this paper we present a technique based on relevance feedback to improve the accuracy in an ontology based information retrieval system. Our proposed method combines the semantic information in a general knowledge base with statistical information using relevance feedback. Several experiments and results are presented using a test set constituted of Web pages.

INTRODUCTION

A user is a fundamental component in the information retrieval process and we can affirm that the goal of an information retrieval system is to satisfy a user's information needs. In several contexts, with the Web it can be very hard to satisfy

completely the request of a user, given the great amount of information and the high heterogeneity in the information structure. On the other hand, users find it difficult to define their information needs, either because of the inability to express information need or just insufficient knowledge about the domain of interest hence they use just

a few keywords. In this context, it is very useful to define the concept of *relevance information*. We can divide relevance into two main classes (Harter, 1992; Saracevic, 1975; Swanson, 1986) called *objective* (system-based) and *subjective* (human (user)-based) relevance respectively. The objective relevance can be viewed as a topicality measure, i.e. a direct match of the topic of the retrieved document and the one defined by the query. Several studies on the human relevance show that many other criteria are involved in the evaluation of the IR process output (Barry, 1998; Park, 1993; Vakkari & Hakala, 2000). In particular the subjective relevance refers to the intellectual interpretations carried out by users and it is related to the concepts of *aboutness* and *appropriateness* of retrieved information. According to Saracevic (1996) five types of relevance exist: an *algorithmic relevance* between the query and the set of retrieved information objects; a *topicality-like type*, associated with the concept of aboutness; *cognitive relevance*, related to the user information need; *situational relevance*, depending on the task interpretation; and *motivational and affective relevance*, which is goal-oriented. Furthermore, we can say that relevance has two main features defined at a general level: *multidimensional relevance*, which refers to how relevance can be perceived and assessed differently by different users; *dynamic relevance*, which instead refers to how this perception can change over time for the same user. These features have great impact on information retrieval systems which generally have not a user model and are not adaptive to individual users. It is generally acknowledged that some techniques can help the user in information retrieval tasks with more awareness, such as *Relevance Feedback* (RF). Relevance feedback is a means of providing additional information to an information retrieval system by using a set of results provided by a classical system by means of a query (Salton & Buckley, 1990). In the RF context, the user feeds some judgment back to the system to improve the initial search results.

The system can use this information to retrieve other documents similar to the relevant ones or rank the documents on the basis of user clues. In this paper we use the second approach. A user may provide the system with relevance information in several ways. He may perform an *explicit feedback* task, directly selecting documents from list results, or an *implicit feedback* task, where the system tries to estimate the user interests using the relevant documents in the collection. Another well known technique is the *pseudo-relevance feedback* where the system chooses the top-ranked documents as the relevant ones. This paper is organized as follows: in section 2 some related work about relevance feedback techniques and different methods and contexts are presented; section 3 briefly summarizes the fundamental theoretical background used in this work; several novel similarity metrics are then introduced in section 4; in section 5 we describe our Web information retrieval system based on ontologies and user feedback, while evaluations, experiments and conclusions are described in section 6 and 7 respectively.

RELATED WORKS

Relevance feedback techniques have been investigated for more than 30 years (Spink & Losee, 1996) and several papers show that they are effective for improving retrieval performance (Harman, 1992; Rocchio, 1971). From a general point of view RF techniques refer to the measure of relevance. In this context an end-user bases his judgment on the expected contribution of the analyzed document to his task. In Resnick et al. (1994) is presented GroupLens, a collaborative filter-based system which ranks the documents on the basis of numeric ratings explicitly assigned by the user. The basic idea is that people who agreed with the evaluation of past articles are likely to agree again in the future. Moreover the system tries to predict user's agreement using the ratings

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/user-relevance-feedback-semantic-information/10203

Related Content

U.S. Disabilities Legislation Affecting Electronic and Information Technology

Deborah Bursa, Lorraine Justice and Mimi Kessler (2005). *Encyclopedia of Information Science and Technology, First Edition* (pp. 2916-2920).

www.irma-international.org/chapter/disabilities-legislation-affecting-electronic-information/14718

A Literature Review on MOOCs Integrated With Learning Analytics

Zhonggen Yu (2021). *Journal of Information Technology Research* (pp. 67-84).

www.irma-international.org/article/a-literature-review-on-moocs-integrated-with-learning-analytics/274279

A University/Community Partnership to Bridge the Digital Divide

David David Ruppel and Cynthia Ruppel (2009). *Encyclopedia of Information Science and Technology, Second Edition* (pp. 3880-3889).

www.irma-international.org/chapter/university-community-partnership-bridge-digital/14156

Process-Aware Information Systems for Virtual Teamwork

Schahram Dustdar (2005). *Encyclopedia of Information Science and Technology, First Edition* (pp. 2314-2320).

www.irma-international.org/chapter/process-aware-information-systems-virtual/14605

A Scholar Case Study of IT Education Knowledge Through Practice for Development in a Post-Conflict Region Kosovo: Increased Investment in Private Education

Halime Cocaj (2023). *Journal of Cases on Information Technology* (pp. 1-66).

www.irma-international.org/article/a-scholar-case-study-of-it-education-knowledge-through-practice-for-development-in-a-post-conflict-region-kosovo/330127