Towards Nootropia: a Non-Linear Approach to Adaptive Document Filtering

Nikolaos Nanas
Knowledge Media Institute
8 December 2003
Supervisors: Dr Victoria S. Uren, Professor Anne De Roeck and Dr John Domingue

In recent years, it has become increasingly difficult for users to find relevant information within the accessible glut. Research in Information Filtering (IF) tackles this problem through a tailored representation of the user interests, a user profile. Traditionally, IF inherits techniques from the related and more well established domains of Information Retrieval and Text Categorisation. These include, linear profile representations that exclude term dependencies and may only represent a single topic of interest, and linear learning algorithms that achieve a steady profile adaptation pace. We argue that these practices are not attuned to the dynamic nature of user interests. A user may be interested in more than one topic in parallel, and both frequent variations and occasional radical changes of interests are inevitable over time. With our experimental system "Nootropia", we achieve adaptive document filtering with a single, multi-topic user profile. A hierarchical term network that takes into account topical and lexical correlations between terms and identifies topic-subtopic relations between them, is used to represent a user's multiple topics of interest and distinguish between them. A series of non-linear document evaluation functions is then established on the hierarchical network. Experiments using a variation of TREC's routing subtask to test the ability of a single profile to represent two and three topics of interest, reveal the approach's superiority over a linear profile representation. Adaptation of this single, multi-topic profile to a variety of changes in the user interests, is then achieved through a process of self-organisation that constantly readjusts the profile stucturally, in response to user feedback. We used virtual users and another variation of TREC's routing subtask to test the profile on two learning and two forgetting tasks. The results clearly indicate the profile's ability to adapt to both frequent variations and radical changes in user interests.