The TREC-9 Filtering Track

David Hull
Xerox Research Centre Europe
Meylan, France

Stephen Robertson
Microsoft Research
Cambridge, UK

A text filtering system sifts through a stream of arriving information to find documents relevant to a set of user profiles. Unlike the traditional search query, user profiles are persistent, and tend to reflect a long term information need. With user feedback, the system can learn a better profile, and improve its performance over time. The TREC filtering track tries to simulate on-line time-critical text filtering applications, where the value of a document decays rapidly with time. This means that potentially relevant documents must be
presented immediately to the user. There is no time to accumulate and rank a set of documents according to their relevance. Evaluation is based only on the quality of the retrieved set, which is scored using a utility measure. The utility measure assigns a positive score for each relevant document retrieved and a negative score to each retrieved document that is not relevant.

The TREC filtering track is broken down into three subtasks: adaptive filtering, batch filtering, and routing. In adaptive filtering, the system starts with only a user profile and must immediate begin filtering documents. In this simulation, each retrieved document is immediately judged for relevance, and this information can be used by the system to adaptively update the filtering profile. Batch filtering is identical to adaptive filtering, except the system also starts with a large sample of evaluated training documents. This makes it much more like the text categorization task. Routing is very similar to batch filtering, but in this case, the system is expected to return a ranked list of documents which will be evaluated according to traditional IR methods. Research groups are free to choose the subtasks they wish to participate in.

The TREC-9 filtering track will follow much the same format as last year, with the same three subtasks. However, we plan to make the adaptive filtering task easier this year by providing systems with one
to four relevant training examples per topic. Setting a filtering threshold with almost no prior information is extremely difficult. In the past, systems achieved the best performance by relying on extremely conservative filtering strategies, retrieving few if any documents per topic. The goal of the track is to enable research on adaptive learning algorithms, which means that systems should be encouraged to retrieve enough documents to make adaptive learning possible. We are also thinking about moving to a text categorization data set such as OHSUMED for TREC-9. Most participants have expressed a strong desire to try their text filtering systems on a new domain, and these collections provide a very rich set of pseudo-topics to choose from. Finally, we plan to explore some new user models, such as: find n relevant documents as fast as possible, or return at least k documents per unit time. It is difficult to compare systems using the current utility measure and it is unclear how it relates to the real world filtering task. The new evaluation models will hopefully add more realism to the text filtering track.

TREC provides a unique opportunity to compare your text filtering or text categorization system in a fair competition to some of the best research systems in the world. If interested in participating, please see the call for participation on the TREC web site (http://trec.nist.gov).