Abstract

The goal of Information retrieval systems is to provide useful information for user's information need. For a collection of documents and a given query, an IR system returns a ranked list of documents. Different IR systems based on IR models such as Vector Space, Smart Vector Space, Extended Boolean, Latent Semantic Indexing etc. produce different text documents for the same query. They rarely return the same documents in response to the same queries. This has led to the field of "data fusion", which seeks to improve the quality of results being presented to user, by combining the outputs of multiple IR algorithms or systems into a single result set. CombMNZ is a score-based fusion algorithm which adds all the reported scores for a document and multiplies the sum value to the number of retrieval models that have returned that document. This paper focuses on Norm_CombMNZ algorithm which normalizes the result obtained from CombMNZ, so that scores lie in 0 to 1 common range and better ranking judgment can be made. The performance of individual IR system is compared with the performance of data fusion system using performance measures such as recall and precision. The graphical result shows that Norm_CombMNZ provides fused resulting text documents to the user, in the form of effective text retrieval.

Index Terms

Computer Science

Information Sciences
Keywords
Data fusion  information retrieval  performance measures  IR models.