Abstract

In this article, an empirical investigation was conducted to determine whether merging search results generated by multiple query variants with the same information need can improve the retrieval performance in patient centered health information retrieval. In addition, this approach was compared with the selective collection enrichment approach, where only the results generated by a single query, which was predicted to perform better on the local collection is used. Three different results merging strategies predominantly used in distributed search environments with large overlapping databases were used in this study. The results of this investigation suggests that merging results using multiple query variants with the same information need can improve the retrieval performance. Also it was observed that the choice of an external collection used in generating these query variants can have an impact in the retrieval performance as it can sometimes lead to a degradation in the retrieval performance. When a comparison was made between results merging strategies and the selective collection enrichment approach, it was observed that the selective collection enrichment approach ranks fewer and highly relevant documents in the top 10 retrieved documents while the results
merging strategies ranks more and slightly relevant documents in the top 10 retrieved documents.

References


Index Terms

Computer Science

Information Sciences
Keywords

Query Variants, Distributed Information Retrieval, Results Merging