ABSTRACT

MOHAMAD REZA PANCAWAN. Query Expansion on Information Retrieval System Using Thesaurus and Intersection Weight. Supervised by AHMAD RIDHA.

Query expansion is a technique to change the initial query from user in order to improve and optimize result of a retrieval system by getting more relevant documents. Thesaurus is a collection of synonyms and antonyms that can be utilized to get additional terms in query expansion. We propose a method to choose additional terms for query expansion based on the intersection between results from original query and results from expanded query.

We experiment with synonyms and antonyms using 2095 documents and 20 queries that we made for this research. To make these queries, we use words that are generally different from the words that are contained in the document collection. The result of this research shows that the use of query expansion can improve the performance of the retrieval system in getting more relevant documents. It also shows that the use of two synonym terms result in the best query expansion, retrieving 98.1% relevant documents with an average precision of 0.252, while the use of an antonym is only able to retrieve 86.9% relevant documents with an average precision of 0.173.

Keywords: antonym, intersection weight, synonym, thesaurus, query expansion.