Abstract

Plagiarism detection systems is one tools we needed, especially in the academic system. This is due to many acts of plagiarism committed by students. Moreover, teacher, or assignments corrector must handle a lot of documents. Because the document is too many, almost impossible to detect any acts of plagiarism in these tasks.

Plagiarism detection system that already exists mostly doing a lot of comparison to many for all documents, and then provide a list of similarity values are used to determine which documents are contain plagiarism act with other documents. However , this system is not really a solution when the document to be checked too much, because it will be an explosion of combinations between documents.

The solution offered in this thesis is to use indexing for document and map the elimination of documents based on these maps. In addition, the matching process for plagiarism detection is using document index graph technique, so that the matching can be performed in parallel. With a combination like this, it is expected processing time can be minimized with a system that produces faster processing time.

It is shown that the trend of the time with the implementation of an index makes moving the query processing time is linear, not exponential moving like before.

Keywords: plagiarism detection system, indexing, scapegoat tree, document indexing graph