Abstract

The rampant of plagiarism nowadays makes the need of a system that can detect plagiarism. To find *similarity* of the documents query with other documents must be done by compare with all documents in the database. The process of matching two documents that require a long time to become a problem. The more number of documents in the database makes the computation time becomes very long. To deal with this long process of matching, proposed an indication of plagiarism detection system using the method of *indexing* with a 2-3 tree data structure. This method aims to eliminate documents in the database that actually has nothing to do with the query document, so that the matching process is only done on the documents that has any relevance to the query document. The result of this research is to process a query and 10000 database documents need 59 seconds for system with *indexing* and 134 seconds for system without *indexing*. While the *f-measure*ment value, average of *precision* and *recall*, of the system with *indexing* is 0,7387 with *threshold* of elimination 0,15 and 0,000428 for system without *indexing*.

Keywords: plagiarism, *indexing*, 2-3 tree, document, fingerprint, query