

Abstract

Plagiarism detection system is needed in many areas . One of them are rife at this time, the system is used to indicate a document plagiarism in massive document. Document plagiarism indication process is using string matching method by matching each input document with the document corpus. In the process of course irrelevant documents at there is any possibility of being matched, this causes the system runs less effective. It can be seen that the matching documents in massive document on plagiarism detection system will certainly require large computational time which is proportional to the large number of documents. It is of course necessary to have a approached method to reduce irrelevant documents and the searching process needs to be fast in the plagiarism detection system. Indexing data structure using multi-level hash table with the FNV-1a hash functions can be implemented for fast searching process. In the structure of the data is also used the threshold value to reduce the irrelevant documents. The result of implementing the indexing process using multi-level hash tables provide stable performance in a document similarity search. Implementation of the indexing process can speed up the process of LCS plagiarism detection system significantly, amounting to 84.14 % using 0,25 of threshold value of symmetric similarity document search method.

Keywords: Indexing, FNV-1a, Hash Table Multi-Level, Plagiarism Detection System.