

Abstract

Plagiarism is an action that copying or confessing someone's document as his or her document without permission. The practice of document plagiarism is often applied by both academics in school and university level. Sometimes the act of plagiarism was modified by replacing the words that contain synonyms, with the intention that looks different from the original article. The process of selecting the test documents based on keywords entered by the user to calculate the weights and similarity query / keyword with any documents using TF-IDF and the similarity between the documents can use the cosine similarity. Selected test documents are documents that are most relevant to the user input keyword. Plagiarism detection system uses an winnowing algorithm. Winnowing algorithm is one of the algorithm in document fingerprinting method. Output from document fingerprinting process is a set of hash value, called a fingerprint. Fingerprint is what will be the basis of a comparison between the test documents that have been entered. The existence of the concept synonym recognition is intended to be able to recognize words that contain synonyms as an act of plagiarism. Synonym recognition process uses an Asymmetric Co-occurrence method. Output from Asymmetric Co-occurrence method is a similarity weight to each term partner. Detecting plagiarism using synonyms get a higher percentage than without using synonyms.

Keyword : *Plagiarism, TF-IDF, cosine similarity, winnowing, synonym recognition.*