

Abstract

Search engines are designed to help people search files and documents from many other documents and then stored in a computer, such as a server. String matching algorithm is the basic component for data searching in search engines. The main problem in the search string is finding a string that consists of several characters (called pattern) in a large number of texts [8]. As a result it takes a lot of time to get the search done. It requires an algorithm that can work quickly and can sort the documents according to the level of compatibility. One of the algorithms that match is the Mixed Heuristic. These algorithms perform a search pattern or query not just against a word, but can be a sentence of more than one word. The matching string system makes the pattern shift as much as one word, if the match is not found. In addition, these algorithms also perform ranking of relevant documents. In this thesis, it shows the analysis of the level of accuracy using precision and recall of the results given by search engines by using the Mixed Heuristic algorithms for string matching, and the analysis of documents from the results given by search engines.

Keyword: *search engine, string matching, mixed heuristic*