Abstract

String matching is one of the most important concepts used in computer science in various real-life applications like Intrusion detection system, Data mining, Plagiarism detection system. There are many string matching algorithms which help to find patterns from the text. These algorithms are categorized in single string matching and multiple string matching. The Wu-Manber (WM) algorithm is a multiple patterns algorithm which is the finest string matching algorithm. The performance of WM depends on various tables built in the preprocessing phase; these are the prefix table, shift table, and hash table. We introduce a new algorithm namely the Efficient Wu Manber algorithm (EWM) algorithm, which is an advanced version of Wu Manber algorithm with respect to time. Efficient Wu-Manber Algorithm eliminates the prefix table which is unused most of the cases in Wu Manber, constructs two shift tables instead of a single shift table, and uses a nonlinear data structure, i.e., AVL tree instead of a linear data structure, i.e., linked list used in WM in the hash table, which reduces the traversed number of nodes to find an exact match. The experimental results and analysis show that EWM algorithm has better performance as compared to WM and its existing improved algorithm and also better from various string matching algorithms.
Efficient Wu Manber String Matching Algorithm for Large Number of Patterns

tools.

References


Index Terms

Computer Science

Algorithms
Keywords

Wu-Manber, String Matching, Single pattern matching, Multiple pattern matching, Boyer Moore, KMP, Advance Wu Manber.