

4037547SCCS/M : MAJOR : COMPUTER SCIENCE ; M.Sc. (COMPUTER SCIENCE)  
KEY WORDS : SIGNATURE FILE / WILDCARD / TEXT PATTERN SEARCH  
ATHIWAT ARPAPONGSAK : PARTIAL STRING MATCHING USING BIT-  
SLICED SIGNATURE FILES. THESIS ADVISOR : DAMRAS WONGSAWANG Ph.D.  
82 p. ISBN 974-662-307-9

The partial string searching using text pattern matching in unformatted data normally requires much processing time because it compares an indicated query with all data which are usually in large size. A signature file represents an actual file in searching. Since signature file size is much smaller than an actual file's size, the processing time is faster. However, signature file algorithm cannot be directly used in partial searching. This thesis proposes a new approach to the use of signature file algorithm in partial string searching. We develop a searching algorithm called Wildcard Searching with Signature File (WSSF). WSSF creates two signature files that will be used in partial string searching. From theoretical analysis and experimentation of WSSF with actual data, we found that WSSF is more efficient than any existing text pattern matching algorithms when applied to partial string searching.

This thesis presents how WSSF works and its procedures in detail. The researches on WSSF, its experimentation with real data, and its results are discussed. Moreover, suggestions and comments for improving WSSF in using this algorithm in the real world are also presented.