International J. of Engg. Research & Indu. Appls. (IJERIA).
ISSN 0974-1518, Vol.1, No. VII (2008), pp 383-393

A NEW COMPRESSION METHOD USEFUL FOR DIRECT SEARCHING

UMESH S. BHADADE AND A.I. TRIVEDI

Abstract

The direct search of compressed text files is a very useful technique. It not only reduces the amount of storage space required for a text file, but it also speeds up the search time. Generally text is coded in ASCII format. Huffman coding or any other run length encoding techniques compresses the plain text. In this paper, a new compression and decompression technique for plain ASCII texts is presented. The merits of our method are: (1) The word-based matching process can be performed directly on the compressed text directly; (2) the decompression process can be started at the position of the search results. The coding scheme successfully codes a group of 1-2-3-4 text characters into a single code. In this paper, the implementation of the plain text compression using a new coding scheme is illustrated, which is based upon static dictionary in a matrix form.

Keywords: Data compression, Text searching.