

ALGORITHM PROJECT PROPOSAL

Anshika Bang - 19410004
Anshul Jain - 19411007
Prateek Somani - 19411022

Topic- Data Compression Analysis

Introduction-

As the name suggests the process of data compression refers to the process of reducing the size of data file. Compression is useful because it reduces resources required to store and transmit data. Computational resources are consumed in the compression process and, usually, in the reversal of the process (decompression). So an efficient way to compress and decompress is necessary for increased efficiency.

There are two types of compression algorithms available till date- Lossless and Lossy compression.

Motivation-

Data compression finds its applications everywhere - easing data transfer, file encryption, and many more. Data compression involves encoding information using fewer bits than original representation.

Objective-

The main designing technique used in Compression algorithms is greedy.

We will try to lay emphasis on Lossless compression algorithms. The two main algorithms in this category are LZW ,PDLZW and LZSS .Also a trivial greedy algorithm Huffman coding comes to help.

- The main aim will be to implement these algorithms and also try decrease the space and time complexity of the final implemented technique. LZW is based on dictionary keys and LZSS is based on the concept of sliding window. Whereas Huffman coding is based simply on mapping the text to bits in such a way that the most frequent character is assigned the least length bit. LZW and LZSS complexity can be reduced by an intelligent use of hash tables and tries. PDLZW is based on a parallel dictionary set.
- Also we will try to apply the designed implementation on a text file to see if it actually works.

Another aspect of compression algorithms is Lossy compression, wherein the original data cannot be fully regained by decompressing. This technique is used mostly in image compression where there is a trade off between the actual pixels of the original image and the pixels of compressed image.

References to be used

- data compression journals by SIAM and ACM.
- A Lossless Data Compression and Decompression Algorithm and Its Hardware Architecture (ieee xplore).
- Research papers of various universities.