

INTISARI

Perkembangan data digital yang semakin pesat menuntut teknik kompresi yang efisien, terutama pada teks naratif yang memiliki struktur kalimat kompleks dan pola pengulangan beragam. Penelitian ini bertujuan menganalisis efektivitas kombinasi tiga algoritma kompresi lossless, yaitu Huffman, LZW, dan Burrows Wheeler Transform (BWT), dalam meningkatkan rasio kompresi tanpa mengorbankan keutuhan data. Variabel yang dikaji meliputi rasio kompresi dan waktu kompresi, dengan validasi integritas data menggunakan hash SHA-256 untuk memastikan perfect reconstruction. Metode yang digunakan adalah eksperimen kuantitatif terhadap teks naratif berformat plain text dengan variasi ukuran ± 100 KB, ± 500 KB, dan ± 1 MB, yang diuji menggunakan algoritma tunggal, kombinasi dua algoritma, serta kombinasi tiga algoritma dalam berbagai urutan eksekusi. Hasil penelitian menunjukkan bahwa urutan algoritma berpengaruh signifikan terhadap kinerja kompresi; pada algoritma tunggal LZW memberikan rasio kompresi lebih baik dibanding Huffman dan BWT, sedangkan pada kombinasi dua tahap urutan BWT $\rightarrow$ LZW menunjukkan hasil paling efisien. Pada kombinasi tiga tahap, urutan Huffman $\rightarrow$ BWT $\rightarrow$ LZW secara konsisten menghasilkan rasio kompresi tertinggi hingga 2,60 pada dataset 1 MB, sementara beberapa urutan lain justru menyebabkan pembesaran ukuran file. Temuan ini menegaskan bahwa pendekatan stacked compression akan optimal apabila disusun dengan urutan yang tepat, sehingga mampu meningkatkan efisiensi kompresi teks naratif tanpa mengurangi integritas data.

Kata kunci: Kompresi Lossless, Huffman, LZW, BWT, Teks Naratif.

ABSTRACT

The rapid growth of digital data demands efficient compression techniques, especially for narrative texts that feature complex sentence structures and diverse repetition patterns. This study aims to analyze the effectiveness of combining three lossless compression algorithms Huffman, LZW, and Burrows Wheeler Transform (BWT) in improving compression ratios without compromising data integrity. The variables examined include compression ratio and compression time, with data integrity validated using the SHA-256 hash to ensure perfect reconstruction. The method employed is a quantitative experiment on plain text narrative data, with varying sizes of approximately 100 KB, 500 KB, and 1 MB. Tests were conducted using single algorithms, combinations of two algorithms, and combinations of all three algorithms in various execution orders. The results show that the order of algorithms significantly affects compression performance; among single algorithms, LZW provides a better compression ratio compared to Huffman and BWT, while in two-stage combinations, the BWT → LZW order yields the most efficient results. For three-stage combinations, the Huffman → BWT → LZW sequence consistently achieves the highest compression ratio, reaching up to 2.60 on a 1 MB dataset, whereas some other sequences actually increase file size. These findings confirm that a stacked compression approach can be optimized if arranged in the correct order, thus enhancing the efficiency of narrative text compression without sacrificing data integrity.

Keywords: Lossless Compression, Huffman, LZW, BWT, Narrative Text.