

ABSTRAK

Kompresi sangat berkaitan dengan masalah efisiensi penyimpanan dan penurunan *latency* pada saat transmisi. Penelitian ini berfokus pada kompresi data teks menggunakan algoritma *ternary Huffman codes* terhadap dua bahasa daerah populer di Indonesia, yakni Bahasa Minangkabau dan Bahasa Melayu. Tujuan utama dari studi ini adalah untuk menganalisis kinerja kompresi *ternary Huffman codes*. Metode *ternary Huffman* dipilih karena memiliki keunggulan dalam menghasilkan kompresi yang optimal dan dapat didekodekan kembali secara akurat. Hasil analisis menunjukkan bahwa Bahasa Minangkabau memiliki nilai *entropy* sebesar 2,634314585 bit/symbol, *expected code length* 2,703805092 bit/symbol, efisiensi 97%, *compression ratio* 34%, dan *compression rate* sebesar 66%. Sedangkan Bahasa Melayu menunjukkan nilai *entropy* sebesar 2,661970699 bit/symbol, *expected code length* 2,728489484 bit/symbol, efisiensi 97,56%, *compression ratio* 34,11%, dan *compression rate* sebesar 65,89%. Penelitian ini menunjukkan bahwa algoritma *ternary Huffman* dapat mengompres data teks dengan efisiensi yang cukup tinggi. Hasil ini membuka peluang untuk penerapan lebih lanjut dalam sistem komunikasi modern, termasuk pada teknologi seperti 5G dan generasi berikutnya. Selain itu, pengujian *outage probability* pada skenario komunikasi antar pengguna yang menggunakan bahasa berbeda menunjukkan adanya fenomena *error floor* akibat interferensi dari salah satu pengguna. Besarnya *error floor* dipengaruhi oleh nilai *entropy* dan *compression ratio*. Pengguna yang bergerak dengan nilai rendah cenderung mengalami peningkatan *error floor* jika berinterferensi dengan pengguna diam yang memiliki nilai tinggi, dan sebaliknya.

Kata Kunci: Kompresi, Bahasa Minangkabau, Bahasa Melayu, *Ternary Huffman Codes*, *outage probability*.



ABSTRACT

Compression is closely related to the issue of storage efficiency and latency reduction during transmission. This study focuses on text data compression using the ternary Huffman codes algorithm for two popular regional languages in Indonesia, namely Minangkabau and Malay. The main objective of this study is to analyze the performance of ternary Huffman codes compression. The ternary Huffman method was chosen because it has the advantage of producing optimal compression and can be decoded accurately. The analysis results show that Minangkabau has an entropy value of 2.634314585 bits/symbol, an expected code length of 2.703805092 bits/symbol, an efficiency of 97%, a compression ratio of 34%, and a compression rate of 66%. Meanwhile, Malay shows an entropy value of 2.661970699 bits/symbol, an expected code length of 2.728489484 bits/symbol, an efficiency of 97.56%, a compression ratio of 34.11%, and a compression rate of 65.89%. This research demonstrates that the ternary Huffman algorithm can compress text data with relatively high efficiency. These results open up opportunities for further application in modern communication systems, including technologies such as 5G and future generations. Furthermore, outage probability testing in communication scenarios between users using different languages indicates the presence of an error floor phenomenon due to interference from one of the users. The magnitude of the error floor is influenced by the entropy value and compression ratio. Mobile users with low values tend to experience an increase in error floor if they interfere with stationary users with high values, and vice versa.

Keywords: Compression, Minangkabau Language, Melayu Language, Ternary Huffman Codes, Outage Probability

