

A Bit-level Text Compression Scheme Based on the ACW Algorithm

Hussein Al-Bahadili¹ Shakir M. Hussain²

¹Arab Academy for Banking and Financial Sciences, Amman 11942, Jordan

²Faculty of Information Technology, Petra University, Amman 11196, Jordan

Abstract: This paper presents a description and performance evaluation of a new bit-level, lossless, adaptive, and asymmetric data compression scheme that is based on the adaptive character wordlength (ACW(n)) algorithm. The proposed scheme enhances the compression ratio of the ACW(n) algorithm by dividing the binary sequence into a number of subsequences (s), each of them satisfying the condition that the number of decimal values (d) of the n -bit length characters is equal to or less than 256. Therefore, the new scheme is referred to as ACW($n; s$), where n is the adaptive character wordlength and s is the number of subsequences. The new scheme was used to compress a number of text files from standard corpora. The obtained results demonstrate that the ACW($n; s$) scheme achieves higher compression ratio than many widely used compression algorithms and it achieves a competitive performance compared to state-of-the-art compression tools.

Keywords: Data compression, bit-level text compression, ACW(n) algorithm, Huffman coding, adaptive coding.