

Atividade Acadêmica - Estrutura de Dados I

Paulo Francisco da Silva Freitas

¹Departamento de Tecnologias e Linguagens -
Universidade Federal Rural do Rio de Janeiro (UFRRJ)
R. Governador Roberto Silveira S/N - Nova Iguaçu -
Rio de Janeiro - RJ - Brasil

paulofrsf@gmail.com

Resumo. Neste relatório será apresentado a proposta de trabalho e os resultados obtidos com a compressão de dados a partir do método de Codificação de Huffman.

1. Problema

Compactar um arquivo de texto usando a codificação de Huffman, com isso é gerada uma árvore que nos permite codificar os caracteres de acordo com a sua frequência. Quanto maior a frequência do caracter, maior é a prioridade dele.

2. Codificação de Huffman

O Método de Huffman consiste em organizar os caracteres em uma árvore de acordo com a sua frequência, e ao percorrê-la registra 1 se percorrer o nó esquerdo/direito (depende da implementação) e 0 se percorrer o outro caminho até chegar ao nó folha e repetir isso de maneira recursiva para todos os nós.

O resultado será uma sequência binária, que após gerada, será guardada em um array, dividida e convertida para representações de 1 byte, que será salva no arquivo como a versão compactada do texto.

3. Testes

Seguem Testes:

Caso 1.

Texto "Hello World"

Tamanho original: 11bytes

Tamanho compactado: 4bytes

Taxa: 63.64

Caso 2.

Texto "A programming language is a formal computer language or constructed language designed to communicate instructions to a machine, particularly a computer. Programming languages can be used to create programs to control the behavior of a machine or to express algorithms."

Tamanho original: 268bytes

Tamanho compactado: 139bytes

Taxa: 48.13

4. Conclusão

A codificação de Huffman é uma boa opção para a compactação de dados, quanto maior for a variedade de caracteres diferentes no texto que se deseja compactar, menor é a sua eficiência.