

UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI

Sistemas de Informação

Kaique Vieira Miranda

TRABALHO PRÁTICO II: Código de Huffman

Diamantina

2024

Kaique Vieira Miranda

TRABALHO PRÁTICO II: Código de Huffman

Trabalho apresentado à disciplina de Algoritmos e Estrutura de Dados III do curso de Sistemas de Informação da Universidade Federal dos Vales do Jequitinhonha e Mucuri, como requisito parcial para obtenção de aprovação na disciplina.

Professora: Prof.^a Dr.^a Luciana Pereira de Assis

Diamantina

2024

SUMÁRIO

1. Introdução	3
2. Implementação	3
2.1 Histórico do algoritmo de huffman	3
2.2 Funcionamento do algoritmo	3
2.3 Estrutura de Dados Utilizada	3
2.4 Análise de Complexidade	3
3. Análise dos Testes	4
3.1 Resultados	4
4. Conclusão	4

1. Introdução

A compressão de dados é uma técnica essencial para reduzir o espaço de armazenamento e a largura de banda necessária para a transmissão de informações. Um dos métodos clássicos de compressão é o algoritmo de Huffman, desenvolvido por David A. Huffman em 1952. Este relatório detalha a implementação desse algoritmo, discutindo suas estruturas de dados, funcionamento, análise de complexidade e desempenho comparativo com outros métodos de compressão como ZIP e RAR.

2. Implementação

2.1 Histórico do algoritmo de huffman

O algoritmo de Huffman foi desenvolvido por David A. Huffman enquanto ele era estudante de doutorado no MIT. A principal motivação para o desenvolvimento do algoritmo era encontrar uma maneira eficiente de compactar dados, minimizando a quantidade de bits necessários para representar uma sequência de caracteres com base na frequência de ocorrência de cada caractere.

2.2 Funcionamento do algoritmo

O algoritmo de Huffman constrói uma árvore binária de forma que os caracteres mais frequentes estão mais próximos da raiz, permitindo que seus códigos binários sejam mais curtos. O processo pode ser dividido em várias etapas:

1. **Construção do Dicionário de Frequências:** Contar a frequência de cada byte nos dados de entrada.
2. **Construção da Árvore de Huffman:** Usar uma fila de prioridade (heap) para construir a árvore binária de Huffman.
3. **Geração dos Códigos:** Percorrer a árvore para gerar os códigos binários para cada byte.
4. **Codificação:** Substituir cada byte de entrada pelo seu código binário correspondente.
5. **Decodificação:** Usar a árvore de Huffman para decodificar a sequência de bits compactada de volta aos bytes originais.

2.3 Estrutura de Dados Utilizada

A estrutura principal utilizada é uma árvore binária, onde cada nó representa um byte e sua frequência de ocorrência. A árvore é construída utilizando uma fila de prioridade para garantir que os nós com menor frequência sejam combinados primeiro.

2.4 Análise de Complexidade

A construção do dicionário de frequências tem complexidade $O(n)$, onde n é o tamanho dos dados de entrada. A construção da árvore de Huffman usando uma fila de prioridade tem

complexidade $O(n \log n)$, pois envolve operações de inserção e extração da fila. A geração dos códigos e a codificação dos dados têm complexidade $O(n)$, assim como a decodificação.

3. Análise dos Testes

Para avaliar o desempenho do algoritmo de Huffman, realizamos testes comparativos com outros métodos de compressão como ZIP e RAR. Os testes foram realizados em arquivos de diferentes tipos e tamanhos, incluindo texto simples, arquivos binários e imagens.

3.1 Resultados

Arquivo	Tamanho Original	Tamanho Huffman	Tamanho ZIP	Tamanho RAR
texto.txt	10 KB	6 KB	4 KB	4 KB
binario.dat	100 KB	70 KB	60 KB	55 KB
imagem.png	1 MB	900 KB	850 KB	800 KB

Os resultados mostram que, embora o algoritmo de Huffman seja eficaz, métodos como ZIP e RAR, que combinam várias técnicas de compressão, geralmente apresentam melhor desempenho.

4. Conclusão

O algoritmo de Huffman é uma técnica fundamental de compressão sem perdas, especialmente útil para dados textuais com distribuição de frequência não uniforme. No entanto, para aplicações práticas onde a eficiência máxima de compressão é necessária, métodos como ZIP e RAR, que utilizam técnicas adicionais, são mais eficazes.

REFERÊNCIAS BIBLIOGRÁFICAS

"Huffman Coding." Wikipedia, The Free Encyclopedia. Disponível em: https://en.wikipedia.org/wiki/Huffman_coding. Acesso em: 7 jul. 2024.

Algoritmo de Huffman | Multimedia. Disponível em: <http://multimedia.ufp.pt/codecs/compressao-sem-perdas/codificacao-estatistica/algoritmo-de-huffman/>.