Trabalho Aeds 3: Estudo de Artigo

Grupo

1. Breno Lopes
2. João Pedro Barroso
3. Lucas Carvalho
4. Lucas Vinícius
5. Vinícius Giovanini

Detalhes do Artigo:

Título: Ferramenta para Análise de Desempenho e Comparação dos Algoritmos de Compactação sem Perda Huffman e LZW

Publicação: periodicos.unesc.net/Google Scholar

Autores:

- **Bruno Casemiro Kurzawe**
- **Christine Vieira Scarpato**

Contextualização do Artigo

A ideia da compressão de dados vem de uma área da computação chamada Teoria da Informação, e desde 1950 são estudados métodos para realizar essa compactação, essa área está preocupada com a quantificação para fins de armazenamento ou comunicação, e todos os dias surgem novas técnicas de processamento de dados a fim de encontrar melhorias para armazena-los, e nas épocas de 80 e 90 essa área teve um grande crescimento, com o desenvolvimento de diversos algoritmos que são utilizados até os dias atuais como Huffman, LEMPEL-ZIV & WELCH LZW, Burrows-Wheeler Transform BWT entre outros. Dessa maneira o artigo aborda como compactadores atuais como o Winrar e o Winzip trabalham utilizando esse cálculo, e o desempenho desses algoritmos com determinados tipos de arquivo.

Objetivo do Artigo

A compactação de dados é importante principalmente nos dias atuais para reduzir a redundância de informações transmitidas, visto que atualmente o trafego de informações pela rede de internet é muito frequente, e também presente no cotidiano com o ZIP, RAR entre outros. Dessa forma o artigo tem como objetivo analisar através de técnicas como a taxa, velocidade de compactação e uso da memória, o melhor algoritmo parta cada tipo de arquivo, detalhando principalmente os algoritmos de Huffman e LZW.

O que foi desenvolvido no artigo

O artigo analisou o algoritmo de Huffman, que é um algoritmo sem perdas, que consiste que ele entrega compactado 100% do que foi destinado para a compressão, com um custo maior, diferente dos algoritmos com perdas que entregam uma compactação final sem conter todos os dados inicias, porém com um custo mais baixo. Huffman trabalha com a técnica de probabilidade de ocorrência, desenvolvido por David A.

Huffman em 1952, tornando-se um dos mais famosos no meio da compactação, sendo utilizado em diversas faculdades pelo mundo.

O algoritmo LZW também foi detalhado, explicando que ele surgiu através do conceito de dicionário de dados, revolucionando novamente a compactação por todo o mundo, e hoje em dia a maioria dos apps utilizam algum algoritmo da familia LZW.

Foi desenvolvido uma ferramenta chamada **FADA-CHL** que implementa o algoritmo de Huffman e o algoritmo LZW, analisando os resultados com arquivos de benchmark da **Canterbury** (Universidade de CanterBury da Nova Zelândia). A aplicação desenvolvida foi divida em duas partes, onde a primeira tem a interface e na segunda o núcleo funcional que será a parte interna do sistema. O usuário interage com a interface selecionando o arquivo e o algoritmo desejado.

Ligação do Artigo com a matéria em Sala de Aula

O artigo relata a origem do algoritmo de Huffman e seu detalhamento, e o mesmo com LZW, trazendo comparações entre eles através de uma ferramenta desenvolvida que utiliza a tabela Canterbury Corpus como parâmetro.

Conclusão

A aplicação desenvolvida para comparar o algoritmo de Huffman com LZW, foi testada com a tabela da universidade de Canterbury, e pode concluir que o Huffman foi mais eficiente, reduzindo cerca de 43, 71% dos arquivos em média, e o LZW foi responsável por uma redução em média de 38, 58% dos arquivos. Percebendo assim que ao escolher o método de compactação deve ser analisados alguns parâmetros, que na maioria dos casos o algoritmo mais eficiente é escolhido pelo tipo de arquivo que está compactando. Também foi analisado que tratando de arquivos de texto LZW e Huffman tem quase a mesma taxa de compactação, mas o tempo de execução do LZW é maior que o Huffman. Quando comparado o uso de memória através da aplicação FADA-CHL foi observado que os dois usam quase a mesma quantidade de memória, e quando existe alguma variação do consumo é relacionado ao uso do dicionário e da árvore binária em maior escala.