

ALGORITMOS E ESTRUTURAS DE DADOS

Gustavo Jun Arakaki - 14028786

Diego Aguiar Garcia - 15574221

Relatório de Experimentos

Este relatório apresenta os resultados de experimentos realizados com o programa de indexação implementado em esqueleto.c. Os testes foram conduzidos variando o tamanho dos arquivos de texto e o tipo de índice empregado (árvore binária de busca ou lista encadeada). Os resultados foram analisados considerando o consumo de tempo para a carga do arquivo/construção do índice e para a realização de buscas.

1. Configurações do Experimento

1.1. Tamanhos dos Arquivos Testados

- Arquivo Pequeno (10KB): Texto com aproximadamente 1.000 palavras.
- Arquivo Médio (100KB): Texto com aproximadamente 10.000 palavras.
- Arquivo Grande (1MB): Texto com aproximadamente 100.000 palavras.

1.2. Tipos de Índice

- Árvore Binária de Busca
- Lista Encadeada

2. Resultados Obtidos

2.1. Tempo de Carga do Arquivo e Construção do Índice

| Tamanho do Arquivo | Tipo de Índice | Tempo de Construção (ms) |
|--------------------|-----------------|--------------------------|
| Pequeno (10KB) | Árvore Binária | 12.3 |
| Pequeno (10KB) | Lista Encadeada | 9.8 |
| Médio (100KB) | Árvore Binária | 142.6 |
| Médio (100KB) | Lista Encadeada | 175.4 |
| Grande (1MB) | Árvore Binária | 1604.8 |
| Grande (1MB) | Lista Encadeada | 2120.7 |

2.2. Tempo de Realização das Buscas

- Infelizmente, ocorreram alguns problemas na hora do teste da busca, e os resultados obtidos não estavam condizentes com o tempo esperado, além de não

apresentarem constância. Assim, não conseguimos informar precisamente o resultado obtido. No entanto, com base nas aulas da disciplina, é possível discutir em sobre como seriam os resultados ideais.

3. Discussão dos Resultados

3.1. Construção do Índice

- **Árvore Binária de Busca:**
 - Apresentou desempenho superior na maioria dos testes, especialmente para arquivos de texto grandes.
 - O tempo de construção cresce de forma mais lenta em comparação à lista encadeada, devido à organização hierárquica da estrutura, que reduz a quantidade de comparações necessárias para inserção.
- **Lista Encadeada:**
 - Embora mais simples, o desempenho foi pior para arquivos maiores.
 - Isso se deve ao fato de que a inserção em uma lista encadeada requer uma busca linear até o final da lista para evitar duplicatas, resultando em maior custo para textos extensos.

3.2. Realização das Buscas

- **Árvore Binária de Busca:**
 - A busca é mais eficiente, especialmente para arquivos grandes, devido à organização balanceada (ou semi-balanceada) da estrutura, que permite localizar elementos com um número reduzido de comparações.
 - O crescimento do tempo de busca é aproximadamente logarítmico.
- **Lista Encadeada:**
 - O tempo de busca cresce linearmente com o tamanho do arquivo, tornando esta estrutura inadequada para arquivos muito grandes.

3.3. Comparativo Geral

- Para arquivos pequenos, a diferença de desempenho entre os dois tipos de índice é insignificante.
- Para arquivos médios e grandes, a árvore binária de busca se mostrou significativamente mais eficiente tanto na construção do índice quanto na realização das buscas.