



ANALISANDO O TEMPO DE EXECUÇÃO DO RADIXSORT LSD DECIMAL AO ORDENAR MÚLTIPLOS DÍGITOS A CADA ITERAÇÃO

VÍCTOR ROMULO CAPOBIANCO SÁ¹; JOSÉ RUI CASTRO DE SOUSA²

Algoritmos de ordenação existem para que dados possam ser ordenados o mais rápido possível em computadores. No entanto, a eficiência de cada método varia de acordo com o contexto onde será utilizado. Devido às demandas de big data, a importância de algoritmos adequados para entradas volumosas, de objetos complexos, cresce a cada dia. É desejável que esses métodos apresentem um baixo consumo de memória, para permitir a ordenação interna das informações. Um algoritmo adequado para esse cenário é o radixsort LSD para listas, desde que ordene-se múltiplos dígitos por iteração. Isso porque seu uso de memória é relativamente baixo e, por ser um algoritmo de hashing, apresenta grande flexibilidade nos tipos de dados que pode ordenar. Implementa-se o radixsort, o bottom-up mergesort e o quicksort em três partições em C++, na base decimal. São geradas mais de trezentas e cinquenta entradas, de cinco tipos distintos e quantidades de dígitos do maior número variando de um a vinte. É coletado o tempo de execução do radixsort, com quantidade de dígitos ordenados simultaneamente variando de um a oito. Também coleta-se dados dos outros algoritmos implementados e do std::list::sort, método de ordenação de listas padrão do C++. Utilizar múltiplos dígitos simultaneamente melhora o desempenho do radixsort de duas a vinte vezes, dependendo do caso e do tamanho da entrada. Isso se traduz em um algoritmo várias vezes mais rápido do que os procedimentos por comparação selecionados. E a diferença cresce de acordo com a quantidade de elementos da entrada com, por exemplo, o mergesort sendo 1,25 vezes mais lento na de tamanho 106 preenchida com inteiros de vinte dígitos? 1,57 vezes na que possui 107 registros e 2,15 vezes quando se tem 108 elementos. A única exceção ocorre ao se comparar o radixsort com o quicksort no seu melhor caso (todas as chaves são iguais). No entanto, em todos os outros casos o algoritmo se mostrou ser a melhor escolha para entradas de tamanhos iguais e, por indução, superiores aos testados. Sendo muito eficiente para bases de dados volumosas é, portanto, compatível com as demandas de big data.

PALAVRAS CHAVE: Ordenação interna, Radixsort, Multi-dígito.

Apoio(s): IF Sudeste MG

¹Ex-Aluno - IF Sudeste MG/Campus Rio Pomba - victorcapobianco01@yahoo.com.br

² Orientador - IF Sudeste MG - jose.castro@ifsudestemg.edu.br