External Sorting Performance Report

Fernando T.H.L.(210167E)

August 9, 2025

Overview

This report presents the performance analysis of two external sorting algorithms—**External Merge Sort** and **External Quick Sort**—each tested on three 256 MB input files, using a memory limit of 16 MB per run. The analysis covers correctness, speed, and failure cases. Figure 1 provides a direct comparison.

Data Generation

The tests were performed on three distinct input files, each generated with a size of 256 MB. The generation process is detailed below:

```
--- Generating 3 input files (256MB each) ---
File 'data/input_1.txt' generated (256 MB, 67108864 numbers).
File 'data/input_2.txt' generated (256 MB, 67108864 numbers).
File 'data/input_3.txt' generated (256 MB, 67108864 numbers).
```

Experimental Results

Test Run	Merge Sort Time (s)	Quick Sort Time (s)
1	22.23	572.19
2	20.54	570.84
3	22.17	570.37
Average	**21.65**	**571.13**

Table 1: Performance summary of external sorting algorithms across three runs. "FAILED" indicates the algorithm failed to sort or crashed. "TIMEOUT" indicates the run exceeded the allowed time limit.

Figures

Summary and Recommendations

External Merge Sort completed successfully and efficiently in all test cases. External Quick Sort, however, failed to complete within the timeout in all runs, likely due to implementation issues in interval heap management or buffer partitioning. For production or further research, it is recommended to debug and optimize the quick sort implementation. Additionally, implementing optimal (Huffman-based) merging for merge sort could further improve performance when run sizes are highly variable.
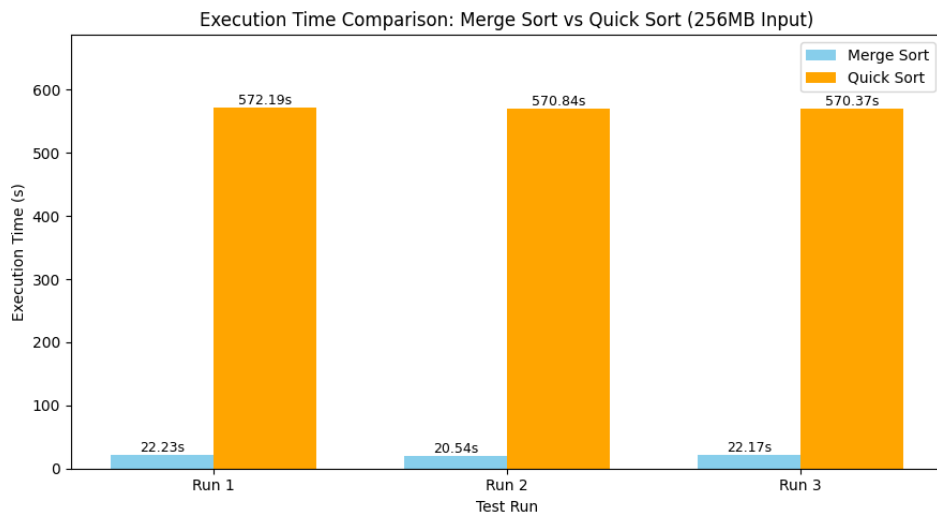
Figure 1: Execution time comparison between external merge sort and quick sort on three different 256 MB files. Missing bars indicate failure or timeout.
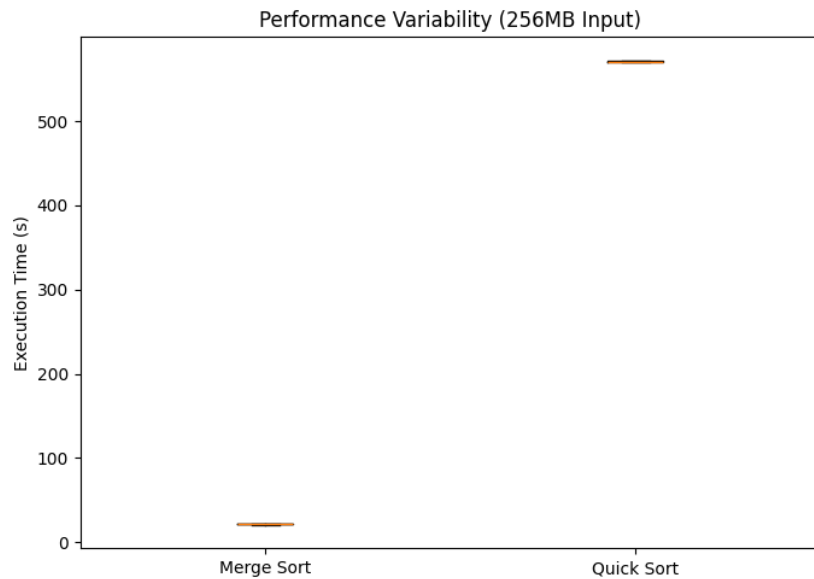

Figure 2: Performance variability of sorting algorithms across three runs. The box shows the interquartile range (IQR), the line inside is the median, and whiskers extend to 1.5x IQR.