Abhishikth Daniel Merugu (NUID:001548340)
Ram Charan Teja Moodapally (NUID:002950328)

INFO 6205: Program Structures & Algorithms

Fall 2021

FINAL PROJECT (MSD STRING SORT)

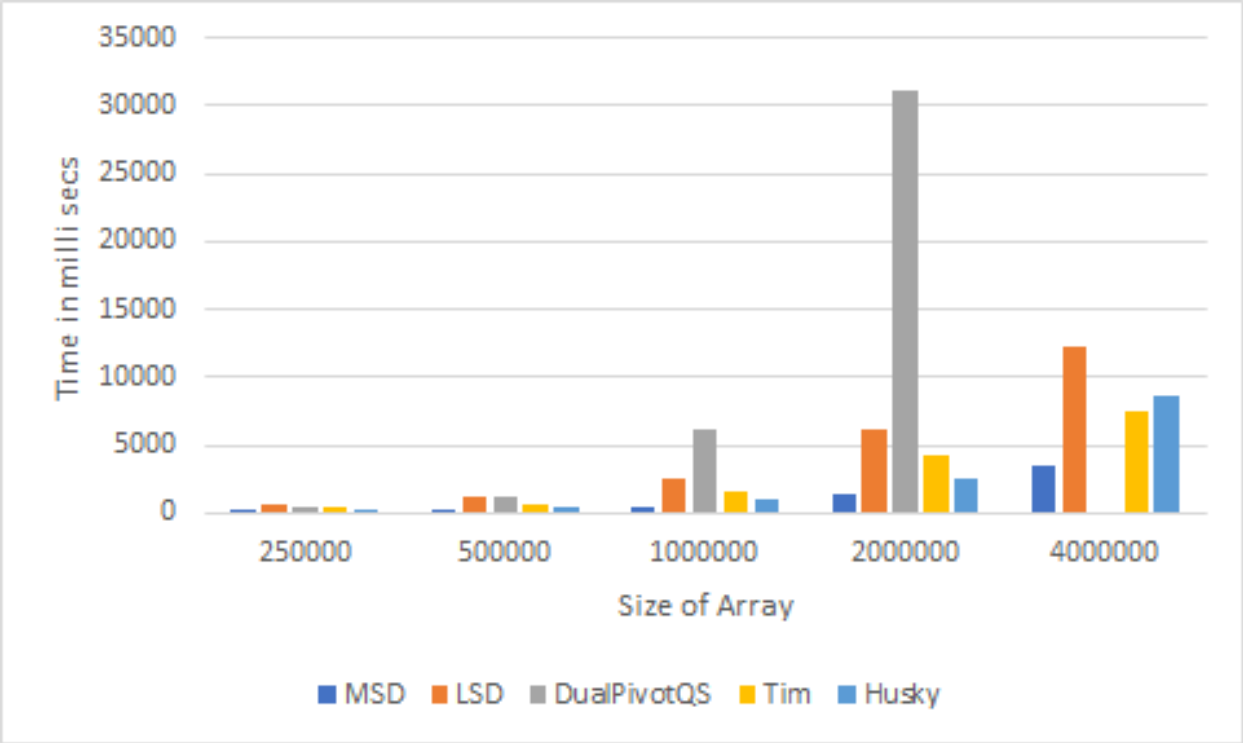
TASK:

We present and evaluate implementation technique for string sorting. We will be using the sorting technique for sorting "Telugu Language" String which uses Unicode characters. Our experiment results indicate the comparisons of the sorting efficiency with different sorting algorithm like Tim Sort, Dual Pivot Quick Sort, Husky Sort, and LSD(Least Significant Digit) Radix sorting using the Benchmark Technique and suggest the optimization techniques and method used for sorting based on the input data.

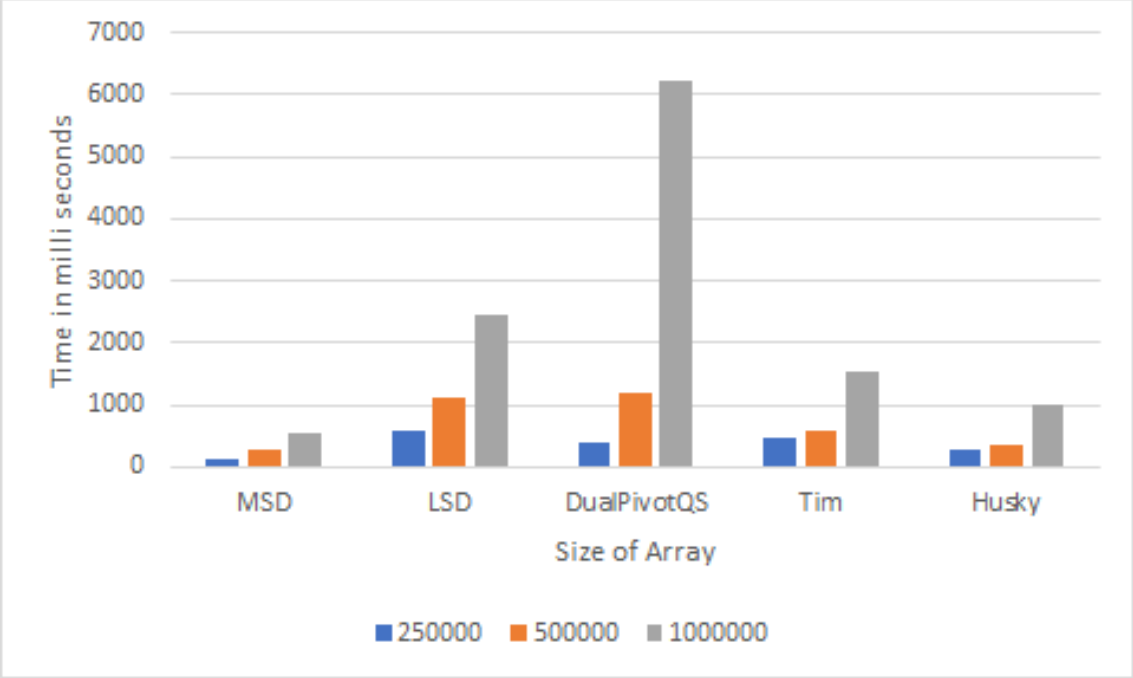
OUTPUT:

Input size (No. of words)	MSD Time in (milliseconds)	LSD (milliseconds)	DualPivotQS (milliseconds)	Tim (milliseconds)	Husky (milliseconds)
250000	138.2	577.2	386.4	481.2	274.2
500000	289.2	1128.2	1205.6	576.2	373.4
1000000	526.6	2456	6206.6	1524.4	1018.2
2000000	1339	6143.8	31116.2	4336.2	2524.6
4000000	3415.8	12220		7465	8692

GRAPHS:



Graph for 1M input data,



Results are generated on the machine of below configurations
OS: Windows
Process: intel i5, 11th Generation
RAM: 8GB

OUTPUT:

2021-12-05 22:44:58 INFO Benchmark_Timer - Begin run: MSDStringSort on 250000 words with 5 runs
138.2
2021-12-05 22:45:07 INFO Benchmark_Timer - Begin run: LSDStringSort on 250000 words with 5 runs
577.2
2021-12-05 22:45:19 INFO Benchmark_Timer - Begin run: DualPivotQuickSort on250000 words with 5 runs
386.4
2021-12-05 22:45:28 INFO Benchmark_Timer - Begin run: TimSort on250000 words with 5 runs
481.2
2021-12-05 22:45:36 INFO Benchmark_Timer - Begin run: HuskySort on250000 words with 5 runs
274.2
2021-12-05 22:45:47 INFO Benchmark_Timer - Begin run: MSDStringSort on 500000 words with 5 runs
289.2
2021-12-05 22:45:55 INFO Benchmark_Timer - Begin run: LSDStringSort on 500000 words with 5 runs
1128.2
2021-12-05 22:46:09 INFO Benchmark_Timer - Begin run: DualPivotQuickSort on500000 words with 5 runs
1205.6
2021-12-05 22:46:21 INFO Benchmark_Timer - Begin run: TimSort on500000 words with 5 runs
576.2
2021-12-05 22:46:31 INFO Benchmark_Timer - Begin run: HuskySort on500000 words with 5 runs
373.4
2021-12-05 22:46:37 INFO Benchmark_Timer - Begin run: MSDStringSort on 1000000 words with 5 runs
526.6
2021-12-05 22:46:45 INFO Benchmark_Timer - Begin run: LSDStringSort on 1000000 words with 5 runs
2456.0
2021-12-05 22:47:07 INFO Benchmark_Timer - Begin run: DualPivotQuickSort on1000000 words with 5 runs
6206.6
2021-12-05 22:47:52 INFO Benchmark_Timer - Begin run: TimSort on1000000 words with 5 runs

1524.4
2021-12-05 22:48:09 INFO Benchmark_Timer - Begin run: HuskySort on1000000 words with 5 runs
1018.2
2021-12-05 22:48:20 INFO Benchmark_Timer - Begin run: MSDStringSort on 2000000 words with 5 runs
1339.0
2021-12-05 22:48:34 INFO Benchmark_Timer - Begin run: LSDStringSort on 2000000 words with 5 runs
6143.8
2021-12-05 22:49:20 INFO Benchmark_Timer - Begin run: DualPivotQuickSort on2000000 words with 5 runs
31116.2
2021-12-05 22:53:07 INFO Benchmark_Timer - Begin run: TimSort on2000000 words with 5 runs
4336.2
2021-12-05 22:53:46 INFO Benchmark_Timer - Begin run: HuskySort on2000000 words with 5 runs

CONCLUSION:

From the above graph we can conclude that, for sorting Telugu 1M words MSD performs better followed by Husky then Tim, LSD and in the last DualPivotQuick Sort

MSD < Husky < Tim < LSD < DualPivot