

Big Data Computing - HW 3 - G24

June 10, 2025

Jianan E

2144167

Gianfranco Mauro

2160477

Davide Volpi

2140728

1 Test 1

The goal is to assess the accuracy of the CM and CS estimations for the top-K heavy hitters for varying K and fixed D=9, W=100, T=100000, and port 8886. As per the assignment, we ran the code 3 times for each configuration.

k	Average Relative Error for top-K heavy hitters with CM			Average Relative Error for top-K heavy hitters with CS		
	RUN 1	RUN 2	RUN 3	RUN 1	RUN 2	RUN 3
10	0.0236	0.0236	0.0231	0.0252	0.0251	0.0250
20	95.6347	282.2221	194.8625	103.9688	316.2384	222.1598
50	430.3376	282.3198	206.9503	472.9426	316.3337	234.4987
100	396.3007	231.0537	261.9498	437.2601	260.5731	295.3719

2 Test 2

The goal is to assess the accuracy of the CM and CS estimations for the top-K heavy hitters for varying columns W and fixed D=9, K=10, T=100000, and port 8886. As per the assignment, we also ran the code 3 times for these configurations.

W	Average Relative Error for top-K heavy hitters with CM			Average Relative Error for top-K heavy hitters with CS		
	RUN 1	RUN 2	RUN 3	RUN 1	RUN 2	RUN 3
5	0.8061	0.8046	0.8128	0.2023	0.2035	0.2036
10	0.2485	0.2481	0.2483	0.9529	0.9538	0.9540
20	0.1211	0.1225	0.1211	0.4252	0.4253	0.4251
50	0.0477	0.0476	0.0482	0.0504	0.0506	0.0501