



INDIAN INSTITUTE OF TECHNOLOGY DELHI

COL380

Report

Assignment – 3

Parallel HNSW

Abstract

We implemented Parallel HNSW an algorithm to quickly and efficiently search for closest vectors in a given search space. Our implementation is parallelized and is scalable, we present the results of our analysis to corroborate the same.

Harsh Agrawal	2019CS10431
---------------	-------------

Arnav Tuli	2019CS10424
------------	-------------

1 Analysis

Below is the analysis for our implementation on different experiments. We have reported the time it takes for the program to execute on N nodes, M cores, with program producing K recommendations. The memory reported is a cumulative sum of the memory used on each node. We used the `time` utility to measure the time taken for one run to complete. The verbose flag in the utility also lists the memory used by the program. Since the utility has no way of knowing the memory usage across nodes, we had the memory usage for a single node. The implementation divides the work almost symmetrically and it is expected that the memory usage will be same for each node. The total memory is thus calculated by scaling the memory usage on one node with the number of nodes.

We also measured the time taken to complete the data setup part. This program takes as input the text files in a folder and converts them to binary. The total time taken to convert all the text files to binary is around 10 minutes.

We then calculated the precision and recall of our recommendations with the ground truth news ids. The following are the results for these metrics.

K	Precision (%)	Recall (%)
5	17.42	12.31
10	16.35	18.79
15	15.00	22.83

Table 1: Precision and Recall metrics for predicted news items

Parallel HNSW

Num nodes	Num cores	K	Time(s)	Total Memory (GB)
2	5	10	0:18.76	13.56
2	5	15	0:22.87	13.42
2	5	5	0:17.77	13.55
2	10	10	0:10.44	13.56
2	10	15	0:11.93	13.56
2	10	5	0:10.88	13.57
2	24	10	0:08.41	13.58
2	24	15	0:09.16	13.58
2	24	5	0:09.10	13.58
5	5	10	0:25.96	28.34
5	5	15	0:18.31	28.34
5	5	5	0:18.51	28.33
5	10	10	0:08.65	28.33
5	10	15	0:09.46	28.36
5	10	5	0:15.51	28.35
5	24	10	0:08.81	28.35
5	24	15	0:08.97	28.36
5	24	5	0:14.83	28.36
10	5	10	0:11.33	52.95
10	5	15	0:12.30	52.94
10	5	5	0:18.62	52.95
10	10	10	0:15.64	52.88
10	10	15	0:16.37	52.84
10	10	5	0:41.60	52.84
10	24	10	0:09.92	52.97
10	24	15	0:13.83	52.99
10	24	5	0:18.09	52.96

Table 2: Time and Memory analysis for various runs of Parallel HNSW