

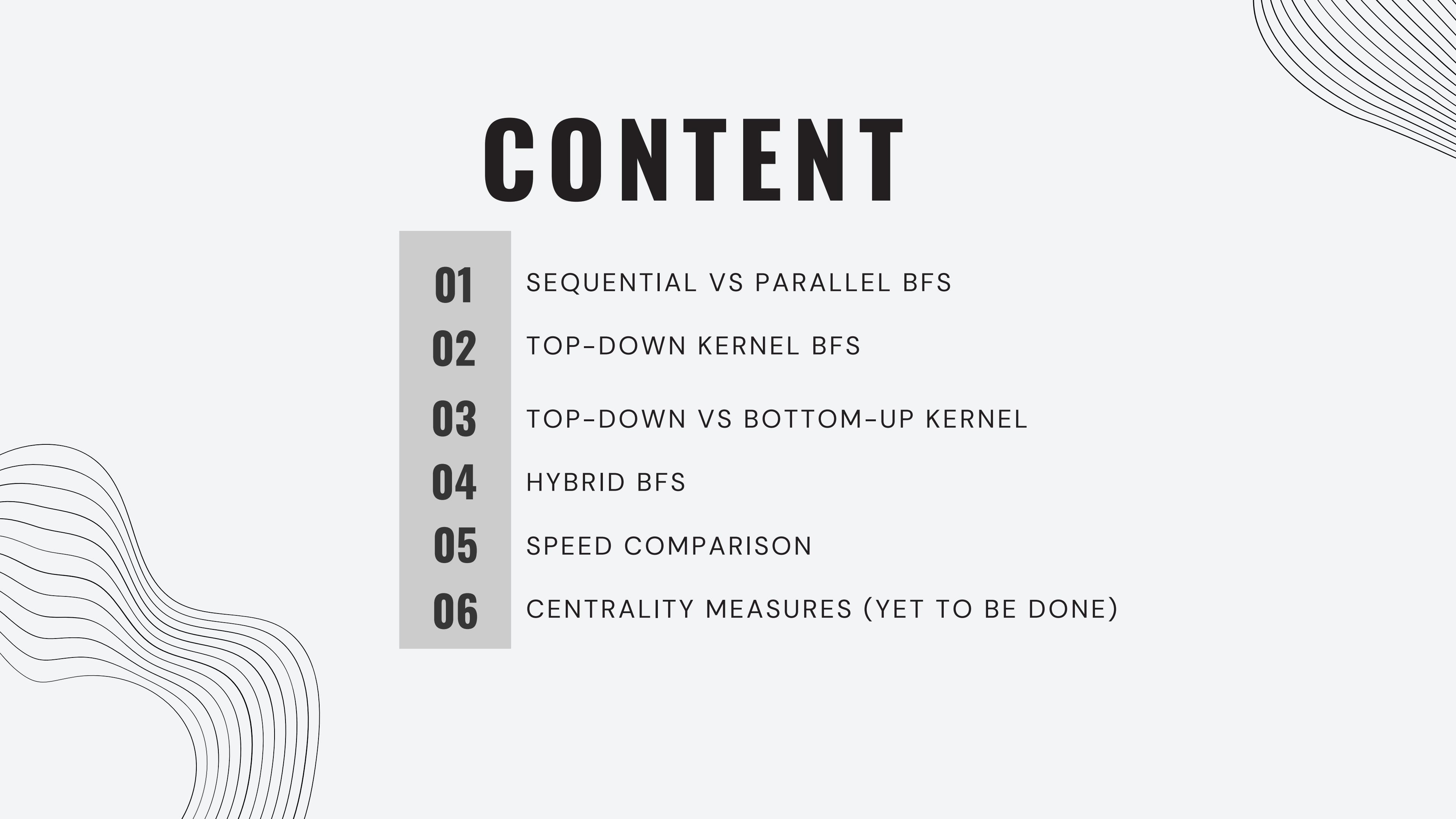
DC PROJECT

PARALLEL BFS

DISHIT SHARMA (B22CS082)

[GITHUB LINK](#)

CONTENT

- 
- 01** SEQUENTIAL VS PARALLEL BFS
 - 02** TOP-DOWN KERNEL BFS
 - 03** TOP-DOWN VS BOTTOM-UP KERNEL
 - 04** HYBRID BFS
 - 05** SPEED COMPARISON
 - 06** CENTRALITY MEASURES (YET TO BE DONE)

SEQUENTIAL VS PARALLEL BFS



Here, my goal was to compare and see the speed boost obtained by the parallelization of code using CUDA.

I manually created graphs with number of vertices 15, 50, 600, 6000 and 1000 with edge probabilities of 0.01, 0.20, 0.40 and 0.60.

This enabled me to do the comparison across various different graph sizes and edge densities.

I observed almost 10x speed boost for the bigger graphs.

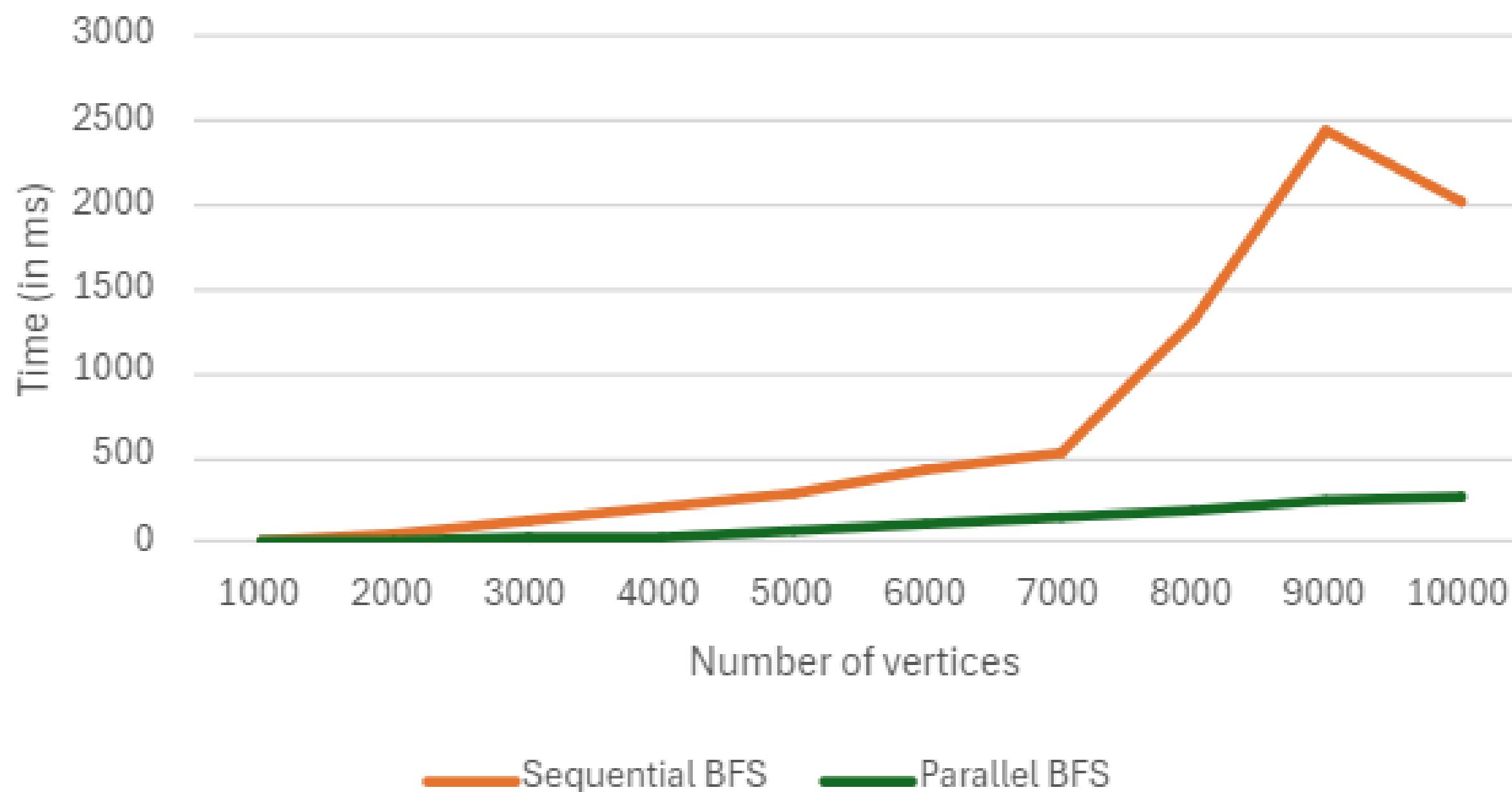
OUTPUT

PROBABILITY: 0.01		
Vertices: 15	Sequential: 0.20 ms	Parallel: 0.30 ms
Vertices: 50	Sequential: 0.00 ms	Parallel: 0.30 ms
Vertices: 600	Sequential: 0.20 ms	Parallel: 0.40 ms
Vertices: 6000	Sequential: 11.10 ms	Parallel: 2.30 ms
Vertices: 10000	Sequential: 20.80 ms	Parallel: 5.60 ms
No of nodes visited by parallel BFS: 1, 1, 597, 6000, 10000		
PROBABILITY: 0.20		
Vertices: 15	Sequential: 0.00 ms	Parallel: 0.60 ms
Vertices: 50	Sequential: 0.10 ms	Parallel: 0.50 ms
Vertices: 600	Sequential: 2.70 ms	Parallel: 2.00 ms
Vertices: 6000	Sequential: 157.60 ms	Parallel: 30.10 ms
Vertices: 10000	Sequential: 446.00 ms	Parallel: 91.60 ms
No of nodes visited by parallel BFS: 12, 50, 600, 6000, 10000		
PROBABILITY: 0.40		
Vertices: 15	Sequential: 0.00 ms	Parallel: 0.60 ms
Vertices: 50	Sequential: 0.10 ms	Parallel: 0.50 ms
Vertices: 600	Sequential: 6.50 ms	Parallel: 2.80 ms
Vertices: 6000	Sequential: 310.30 ms	Parallel: 80.50 ms
Vertices: 10000	Sequential: 1719.80 ms	Parallel: 205.60 ms
No of nodes visited by parallel BFS: 15, 50, 600, 6000, 10000		
PROBABILITY: 0.60		
Vertices: 15	Sequential: 0.00 ms	Parallel: 0.40 ms
Vertices: 50	Sequential: 0.10 ms	Parallel: 0.70 ms
Vertices: 600	Sequential: 11.40 ms	Parallel: 3.20 ms
Vertices: 6000	Sequential: 1094.00 ms	Parallel: 116.70 ms
Vertices: 10000	Sequential: 1904.70 ms	Parallel: 273.10 ms
No of nodes visited by parallel BFS: 15, 50, 600, 6000, 10000		

VISUALIZATION USING GRAPHS

PROBABILITY: 0.6

Vertices: 1000	Sequential: 17.00 ms	Parallel: 2.80 ms
Vertices: 2000	Sequential: 53.10 ms	Parallel: 11.20 ms
Vertices: 3000	Sequential: 131.70 ms	Parallel: 19.40 ms
Vertices: 4000	Sequential: 212.60 ms	Parallel: 38.00 ms
Vertices: 5000	Sequential: 298.90 ms	Parallel: 59.50 ms
Vertices: 6000	Sequential: 423.30 ms	Parallel: 116.50 ms
Vertices: 7000	Sequential: 535.00 ms	Parallel: 155.70 ms
Vertices: 8000	Sequential: 1320.70 ms	Parallel: 182.40 ms
Vertices: 9000	Sequential: 2452.80 ms	Parallel: 252.80 ms
Vertices: 10000	Sequential: 2029.20 ms	Parallel: 266.80 ms



TOP-DOWN BFS



Here, my goal was to use frontiers (queues) to explore the graph of very big sizes utilizing the top-down approach

I utilized the CSR (compressed sparse row) format to store the .egr graphs which I took from SNAP

This enabled me to do the comparison across various different graph sizes.

OUTPUT

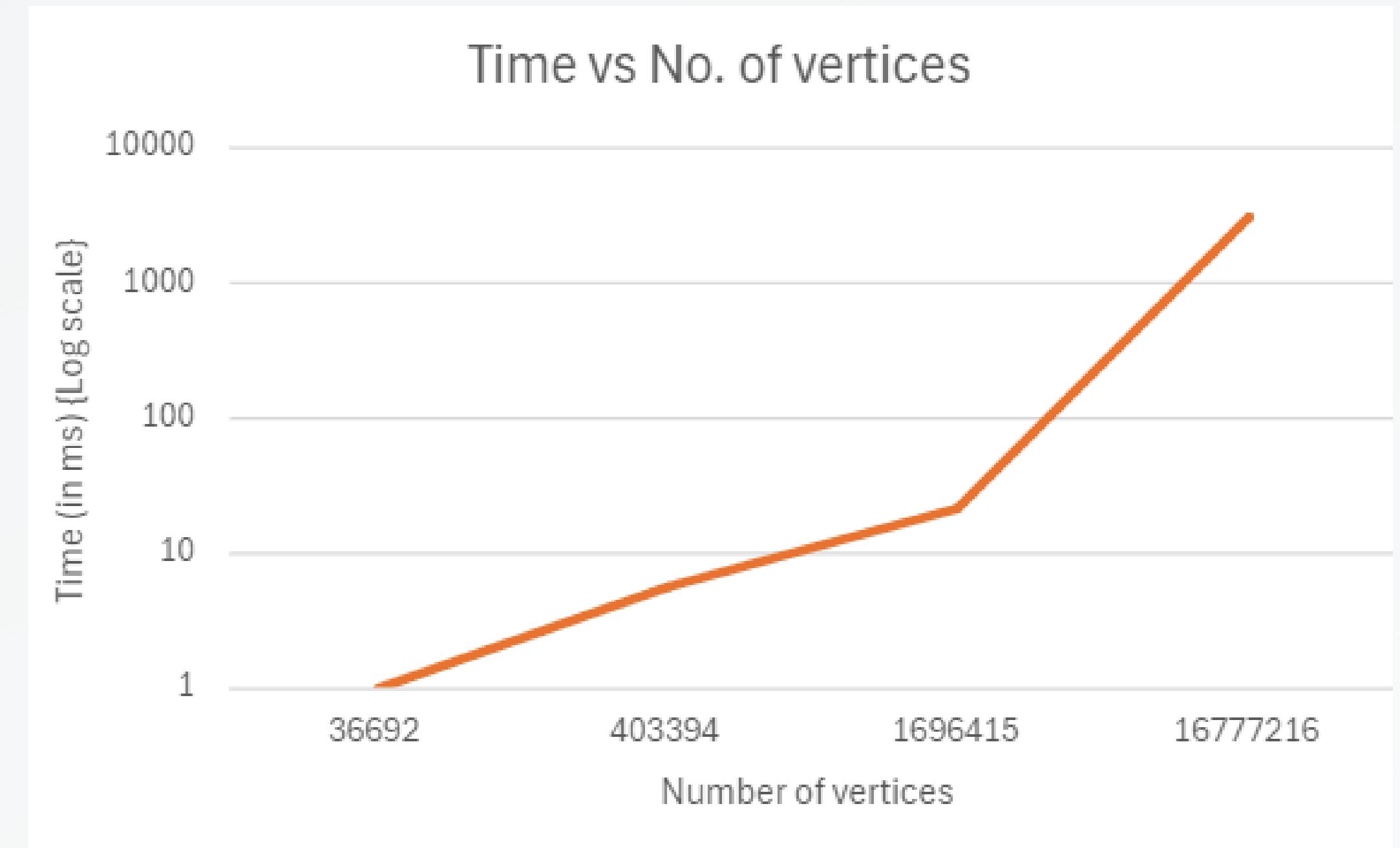
```
Number of nodes in graph: 36692
Number of edges in graph: 367662
Time taken in preprocessing: 553.239 ms
Time taken by BFS: 1.0082 ms
Number of nodes visited by BFS: 33696
```

```
Number of nodes in graph: 403394
Number of edges in graph: 4886816
Time taken in preprocessing: 3.3325 ms
Time taken by BFS: 5.5887 ms
Number of nodes visited by BFS: 403364
```

```
Number of nodes in graph: 1696415
Number of edges in graph: 22190596
Time taken in preprocessing: 15.564 ms
Time taken by BFS: 21.5769 ms
Number of nodes visited by BFS: 1694616
```

```
Number of nodes in graph: 16777216
Number of edges in graph: 100663202
Time taken in preprocessing: 69.2494 ms
Time taken by BFS: 3163.52 ms
Number of nodes visited by BFS: 16777216
```

VISUALIZATION USING GRAPHS



TOP-DOWN VS BOTTOM-UP KERNEL



Here, my goal was to compare and see the speed differences between the top-down and the bottom-up kernels

I utilized the CSR (compressed sparse row) format to store the .egr graphs which I took from SNAP

This enabled me to do the comparison across various different graph sizes.

OUTPUT

An unknown error came when using bottom-up kernel for one of the graph. For the rest, it worked just fine.

TOP-DOWN KERNEL:

```
Number of nodes in graph: 36692
Number of edges in graph: 367662
Time taken in preprocessing: 529.704 ms
Time taken by Hybrid BFS: 2.1122 ms
Number of nodes visited by BFS: 33696
```

```
Number of nodes in graph: 403394
Number of edges in graph: 4886816
Time taken in preprocessing: 4.0869 ms
Time taken by Hybrid BFS: 8.6472 ms
Number of nodes visited by BFS: 403364
```

```
Number of nodes in graph: 1696415
Number of edges in graph: 22190596
Time taken in preprocessing: 14.9596 ms
Time taken by Hybrid BFS: 30.2039 ms
Number of nodes visited by BFS: 1694616
```

```
Number of nodes in graph: 16777216
Number of edges in graph: 100663202
Time taken in preprocessing: 84.4632 ms
Time taken by Hybrid BFS: 3176.31 ms
Number of nodes visited by BFS: 16777216
```

BOTTOM-UP KERNEL:

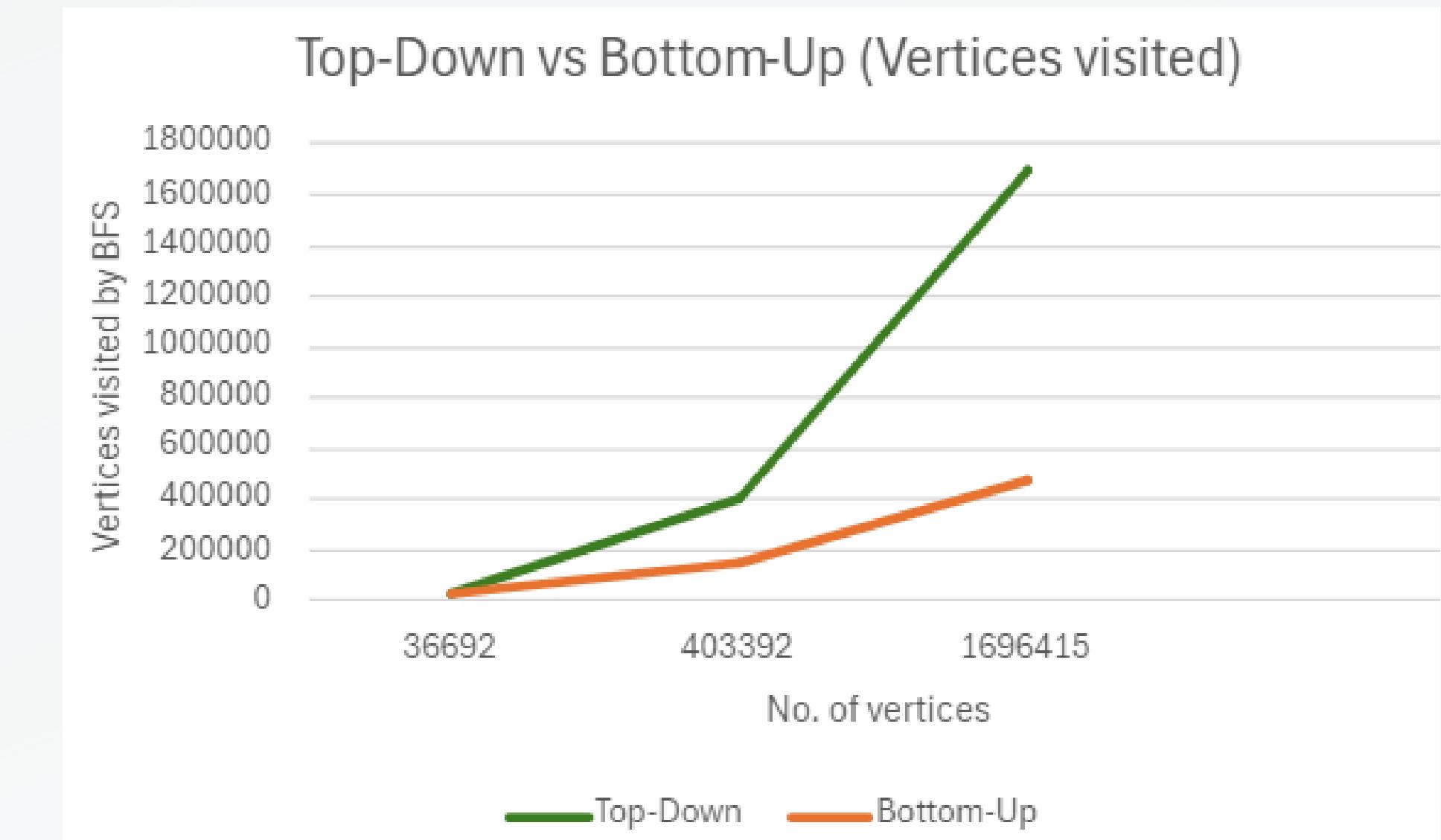
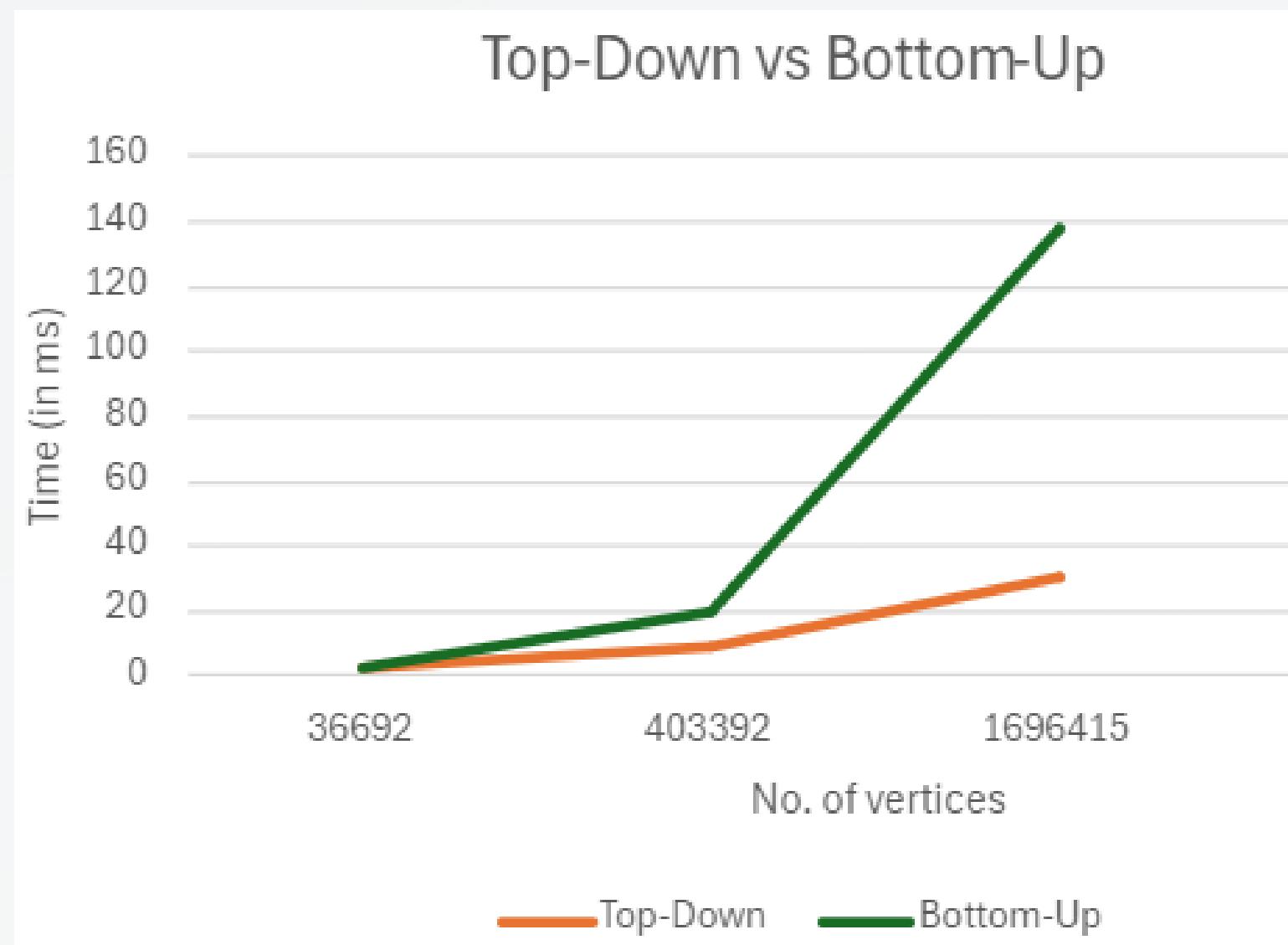
```
Number of nodes in graph: 36692
Number of edges in graph: 367662
Time taken in preprocessing: 0.6999 ms
Time taken by Hybrid BFS: 2.7418 ms
Number of nodes visited by BFS: 25579
```

```
Number of nodes in graph: 403394
Number of edges in graph: 4886816
Time taken in preprocessing: 4.294 ms
Time taken by Hybrid BFS: 20.0909 ms
Number of nodes visited by BFS: 152974
```

```
Number of nodes in graph: 1696415
Number of edges in graph: 22190596
Time taken in preprocessing: 14.1947 ms
Time taken by Hybrid BFS: 138.109 ms
Number of nodes visited by BFS: 477825
```

```
Number of nodes in graph: 16777216
Number of edges in graph: 100663202
Time taken in preprocessing: 78.8027 ms
Time taken by Hybrid BFS: 18.7718 ms
Number of nodes visited by BFS: 1
```

VISUALIZATION USING GRAPHS



HYBRID BFS



Here, my goal was to come up with a switching condition b/w top-down and bottom-up kernels to optimize performance.

I utilized the CSR (compressed sparse row) format to store the .egr graphs which I took from SNAP

This enabled me to do the comparison across various different graph sizes.

After reading through research papers... I finalised alpha = 14 and beta = 24 as the value of the parameters.

OUTPUT

TOP-DOWN KERNEL:

Number of nodes in graph: 36692
Number of edges in graph: 367662
Time taken in preprocessing: 1682.8 ms
Time taken by Hybrid BFS: 2.1319 ms
Number of nodes visited by BFS: 33696

Number of nodes in graph: 403394
Number of edges in graph: 4886816
Time taken in preprocessing: 3.9483 ms
Time taken by Hybrid BFS: 8.2439 ms
Number of nodes visited by BFS: 403364

Number of nodes in graph: 1696415
Number of edges in graph: 22190596
Time taken in preprocessing: 17.5838 ms
Time taken by Hybrid BFS: 30.428 ms
Number of nodes visited by BFS: 1694616

Number of nodes in graph: 16777216
Number of edges in graph: 100663202
Time taken in preprocessing: 83.921 ms
Time taken by Hybrid BFS: 3287.74 ms
Number of nodes visited by BFS: 16777216

HYBRID BFS:

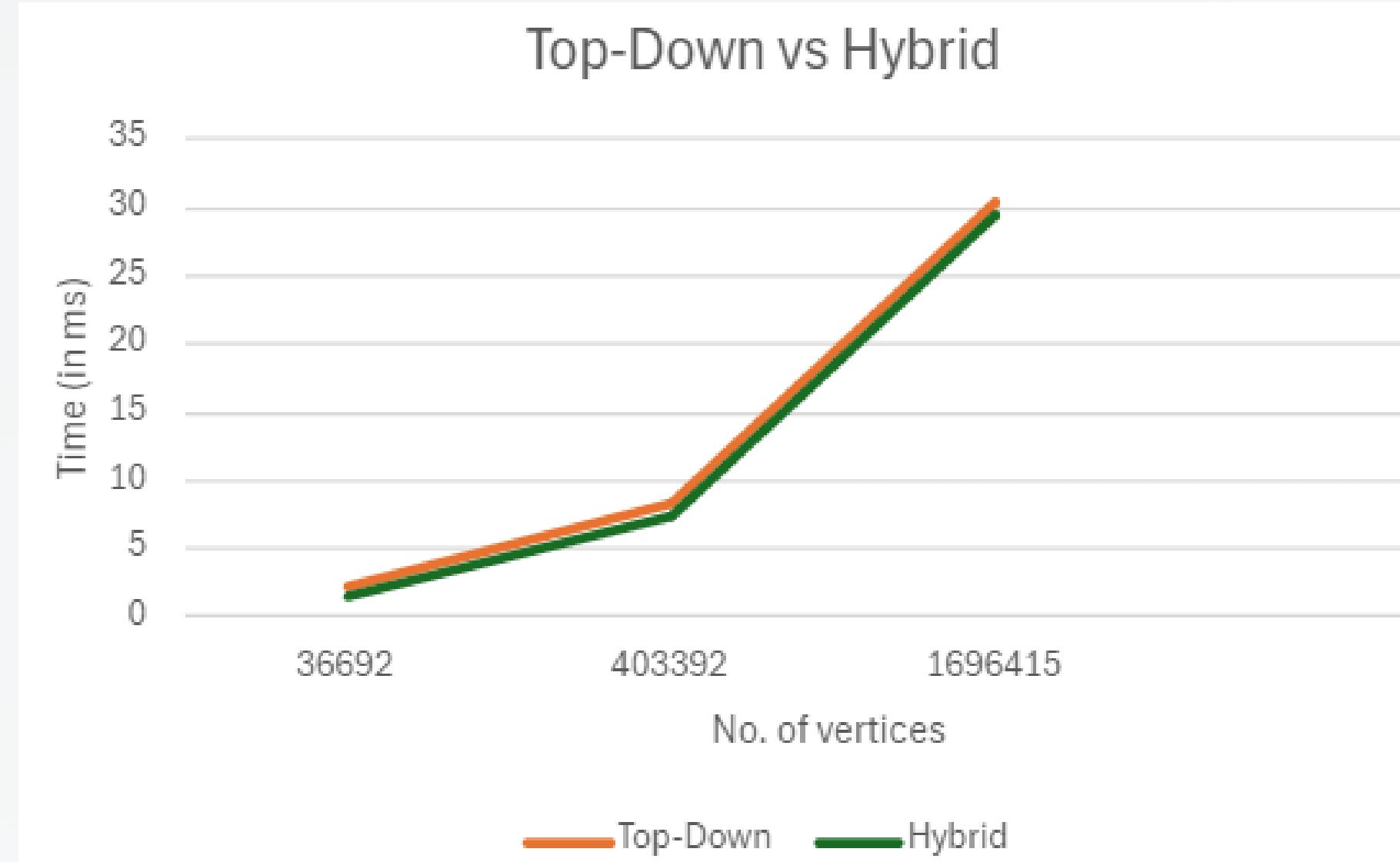
Going from top-down to bottom-up
Going from bottom-up to top-down
Number of nodes in graph: 36692
Number of edges in graph: 367662
Time taken in preprocessing: 0.7687 ms
Time taken by Hybrid BFS: 1.5299 ms
Number of nodes visited by BFS: 32198

Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Number of nodes in graph: 403394
Number of edges in graph: 4886816
Time taken in preprocessing: 4.0997 ms
Time taken by Hybrid BFS: 7.3618 ms
Number of nodes visited by BFS: 393050

Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Going from top-down to bottom-up
Going from bottom-up to top-down
Number of nodes in graph: 1696415
Number of edges in graph: 22190596
Time taken in preprocessing: 15.6013 ms
Time taken by Hybrid BFS: 29.6639 ms
Number of nodes visited by BFS: 1137979

Going from top-down to bottom-up
Going from bottom-up to top-down
Number of nodes in graph: 16777216
Number of edges in graph: 100663202
Time taken in preprocessing: 80.8725 ms
Time taken by Hybrid BFS: 3240.48 ms
Number of nodes visited by BFS: 16777216

VISUALIZATION USING GRAPHS



As of right now, hybrid is only slightly outperforming top-down kernel. This is because I am yet to find the optimal switching conditions for all the different sizes of graphs.

THANK YOU

