Algorithm programming assignment – 3

# Shyam Kantesariya(40042715)

PageRank (PR) is an algorithm used by Google Search to rank websites in their search engine results – Wikipedia.

As part of this assignment I have implemented Pagerank algorithm in Java with following equation to calculate pagerank value.



Here d = scaling factor

PR(B) = pagerank of node B

L(B) = #references B has, that is #links from B to other nodes

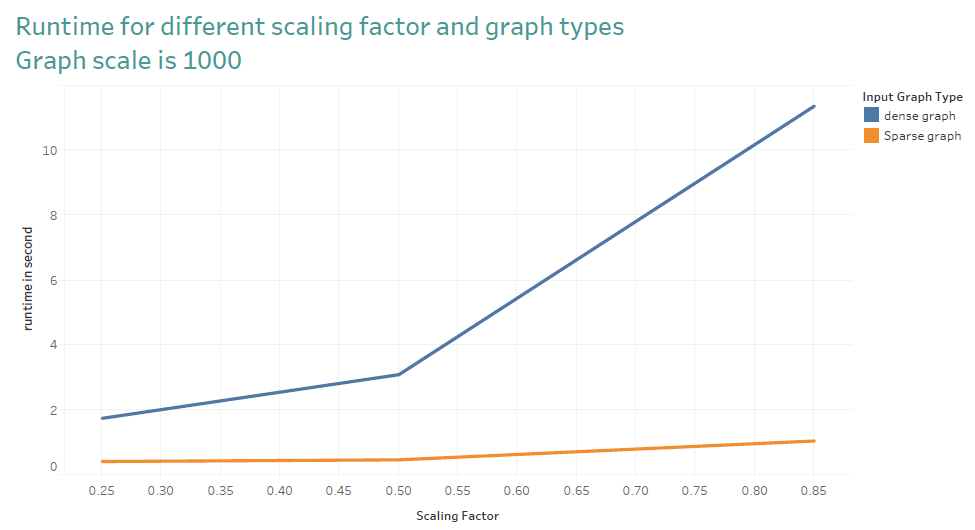
B, C and D are the nodes who have links referring to A.

**Complexity**: O(EV) for each iteration

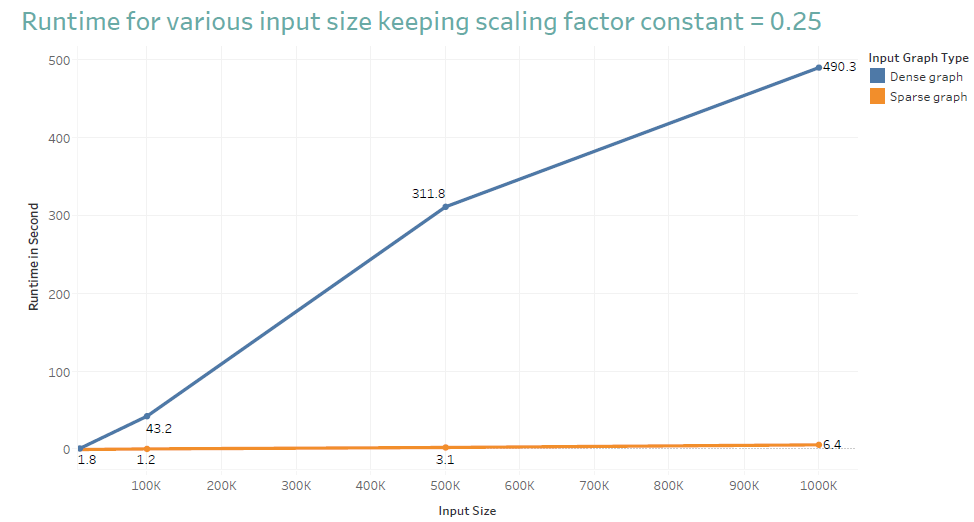
Following are the runtimes observed for different combinations of Graph types, scaling factor and input size.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TC#** | **Graph Type** | **Graph Scale** | **Scaling Factor** | **Runtime in ms** |
| 1 | Sparse graph | 10000 | 0.25 | 420 |
| 2 | Sparse graph | 10000 | 0.5 | 472 |
| 3 | Sparse graph | 10000 | 0.85 | 1052 |
| 4 | Sparse graph | 100000 | 0.25 | 1221 |
| 5 | Sparse graph | 100000 | 0.5 | 1815 |
| 6 | Sparse graph | 100000 | 0.85 | 4411 |
| 7 | Sparse graph | 500000 | 0.25 | 3054 |
| 8 | Sparse graph | 500000 | 0.5 | 4680 |
| 9 | Sparse graph | 500000 | 0.85 | 15088 |
| 10 | Sparse graph | 1000000 | 0.25 | 6433 |
| 11 | Sparse graph | 1000000 | 0.5 | 9826 |
| 12 | Sparse graph | 1000000 | 0.85 | 32327 |
| 13 | Dense graph | 10000 | 0.25 | 1750 |
| 14 | Dense graph | 10000 | 0.5 | 3094 |
| 15 | Dense graph | 10000 | 0.85 | 11360 |
| 16 | Dense graph | 100000 | 0.25 | 43162 |
| 17 | Dense graph | 100000 | 0.5 | 86032 |
| 18 | Dense graph | 100000 | 0.85 | 119596 |
| 19 | Dense graph | 500000 | 0.25 | 311792 |
| 20 | Dense graph | 500000 | 0.5 | 572779 |
| 21 | Dense graph | 500000 | 0.85 | 825673 |
| 22 | Dense graph | 1000000 | 0.25 | 490344 |
| 23 | Dense graph | 1000000 | 0.5 | 799457 |
| 24 | Dense graph | 1000000 | 0.85 | 1083293 |

# Effect of scaling factor



# Effect of input size

Pagerank distribution for different graph types

|  |  |
| --- | --- |
| Sparse graph (N links) rank distribution | Dense graph (N\*(N-1)/2 links) rank distribution |
| 9,0.8031255956592775  8,0.768383053716797  7,0.7275094749609377  6,0.6794229117187501  5,0.6228504843750001  4,0.5562946875000001  3,0.47799375000000005  2,0.385875  1,0.2775 | 1,0.9999946142894082  2,0.9999946142894082  3,0.9999946142894082  4,0.9999946142894082  5,0.9999946142894082  6,0.9999946142894082  7,0.9999946142894082  8,0.9999946142894082  9,0.9999946142894082 |

# Conclusion

|  |  |
| --- | --- |
| **Parameter** | **Observation** |
| Scaling factor | As we increase the scaling factor, convergence time increases. Increase in convergence time is exponential in Dense graph compared to Sparse graph |
| Input size | As we increase the input size, convergence time increases. Increase in convergence time is exponential in Dense graph compared to Sparse graph |
| Pagerank distribution for various Graph types | Rank distribution is even for dense graph though that’s not the case for sparse graph |