Project Report – Gossip & PushSum Simulator

Ammar Amjad 5992-1730, Mohammad Anas 5981-5998

October 10, 2022

Task: To determine the convergence time of the Gossip and Push-Sum algorithm for multiple topologies through a simulator.
Bonus Task: Determine convergence of Gossip Algorithm for multiple topologies using dropout ie. temporarily making actors unavailable. Bonus task at end of the report.

Introduction:

An erlang based simulator is created to complete the above task. A comparative study was done on the different topologies. The convergence time graphs are shown in figure 1 and figure 2. It was observed that the convergence time is directly proportional to the number of nodes and inversely to the number of neighbours, i.e, the more neighbours a node in a network has, the less time it takes to converge.

Our Simulator has N actors, equal to the number of nodes. As soon as the initial message is delivered to an actor, the actors proceed to send messages further to their neighbours randomly until they converge. As a result, the entire network can receive the message.

The termination criteria for the Gossip algorithm is that when a node in the gossip network has heard the rumour 3 times, that node has met the convergence condition. For the Push-Sum, the termination condition is when the ratio changes by less than 10 raised to power -10 over the course of three rounds. When all nodes have converged, the execution ends.

For the **Gossip** algorithm, in figure 1, we can see that the Full network takes the least amount of time. After the Full network, Imperfect 3D followed by 2D takes the least amount of time while the line takes the most. This is because the line has only two neighbours and it restricts its ability to communicate to a larger audience. On the other hand, in a full network, every node is connected to another node and this allows a single node to send a rumour to a large number of nodes. This makes the diffusion of messages very fast. This was also confirmed by our figure.

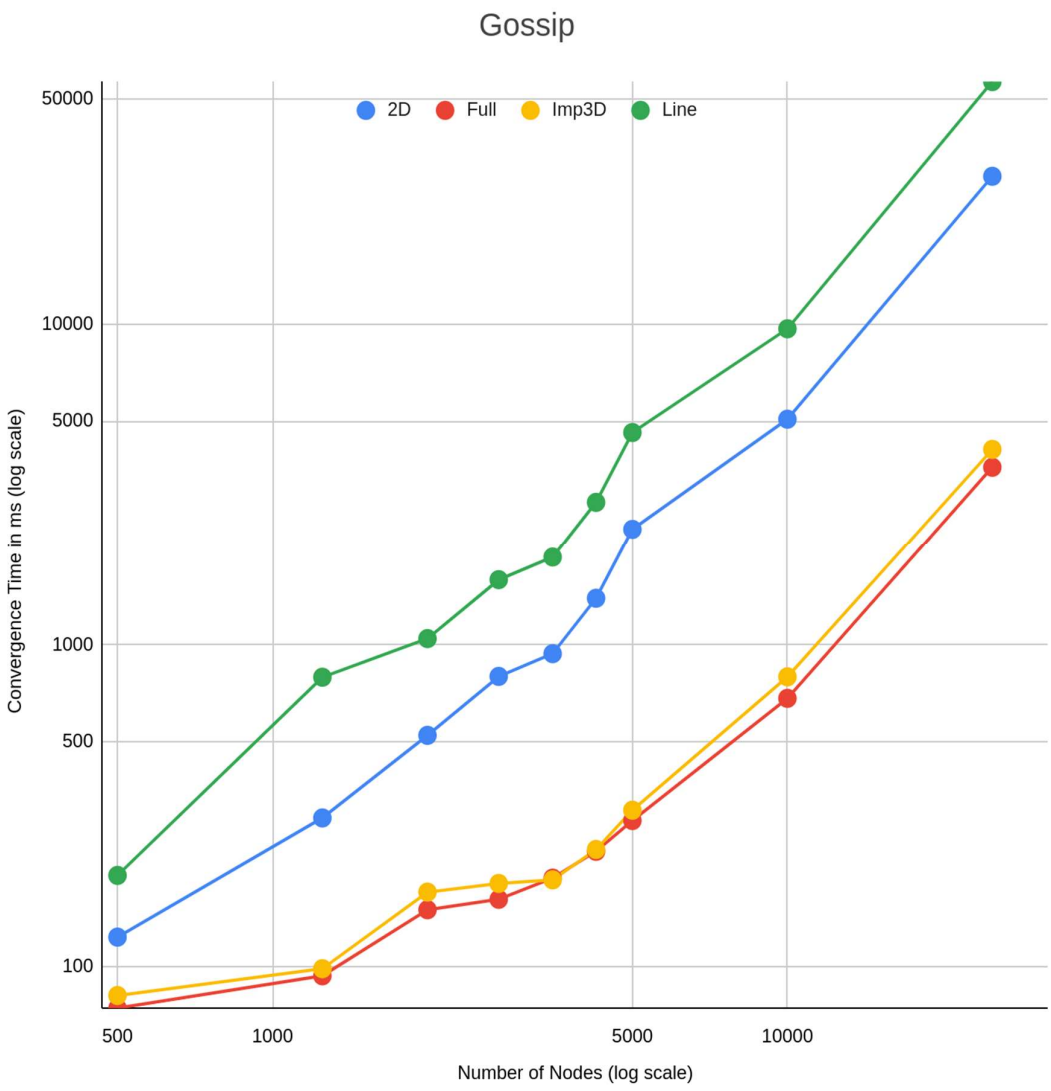

Figure 1

For the **Push-Sum** algorithm, in figure 2, we observe a similar trend. The full network takes the least time while the Line takes the most. The Imperfect 3D takes more time than a Full network and 2D takes more time than an Imperfect 3D network. However, the gap between the plots is larger in the Push-sum as compared to Gossip as the difference between individual convergence times of the algorithms increases.

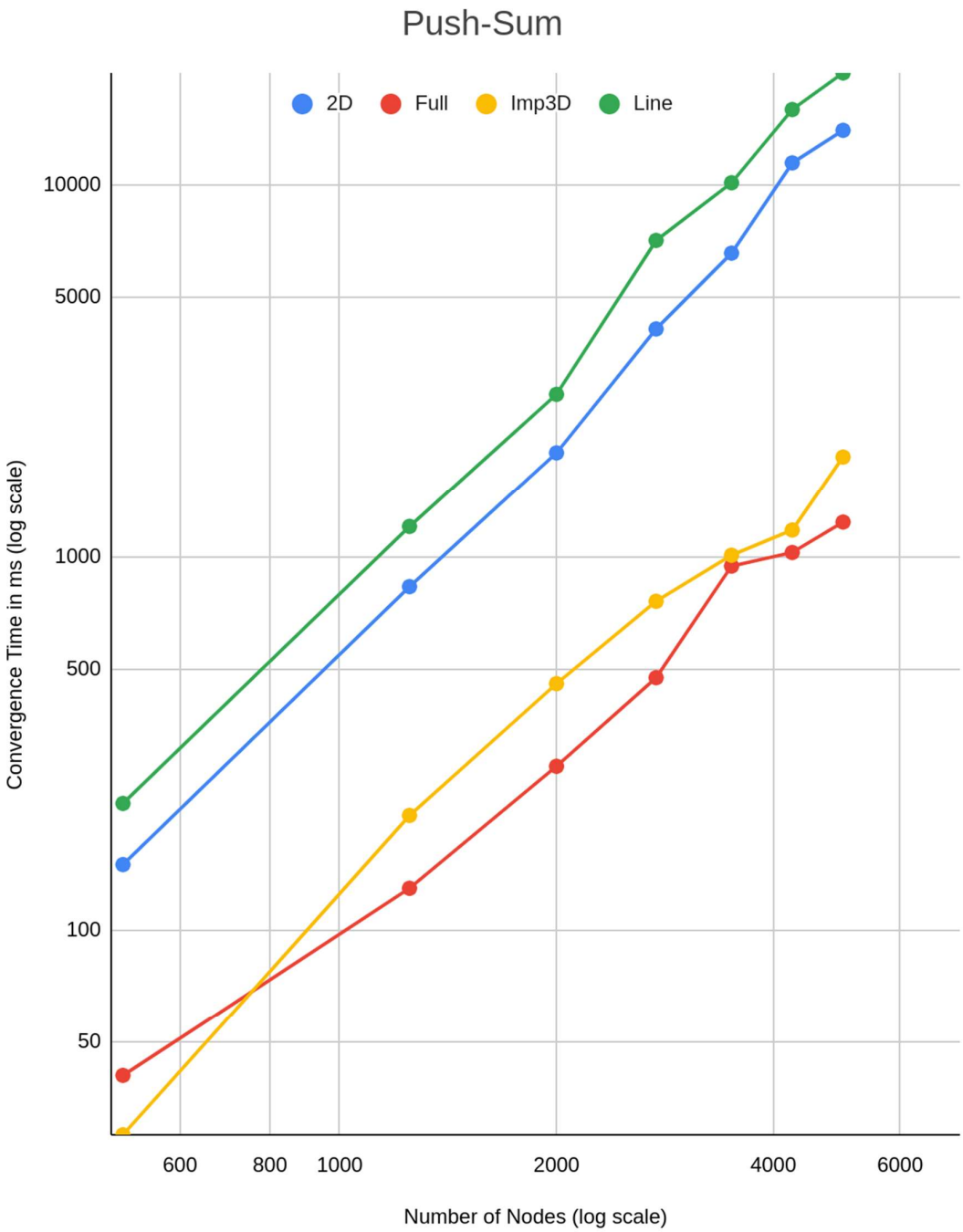

Figure 2

Bonus: Gossip DropOut Simulator

We implemented a novel Dropout based Simulator which also takes into consideration the failure of nodes. For this, we created a system in which some percentage of nodes become unavailable (fail). First, we tested our simulator with 20% of nodes unavailable. This means that 20% of nodes will not participate in the spread of rumours for a period of time. Similarly, we performed our experiment for 50% and 80% of nodes for different network topologies.

We observed from the below graphs that, as the % of failure in a topology increases the time taken by the algorithm to converge also increases. However, the effect of failure is different in different topologies, i.e, the effect of failure is inversely proportional to the number of neighbours in the network. The topologies with a high number of neighbours are more fault tolerant. This is justified because as the number of neighbours of a particulate node increases it has more options to mitigate the effects of a failed neighbour. For example, it was observed from plots that the Line topology is most affected due to failure. Since in Line topology, each node has only 2 neighbours so the failure of one node reduces the probability of propagation of the message by half. It was also observed that as the % of failure in node increases the time taken by the algorithm to converge also increases.

Please see graphs below:

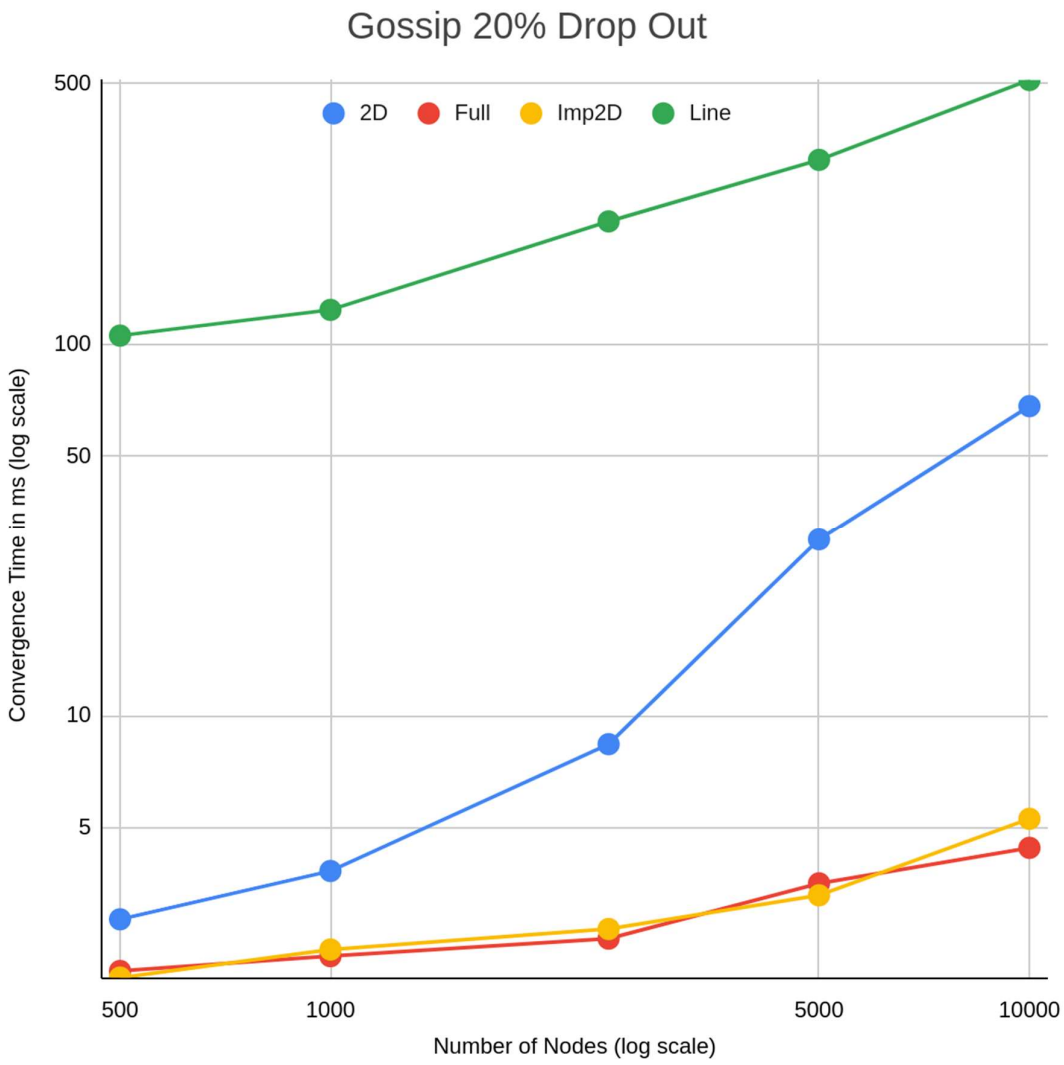
Gossip 20% Drop Out
2D
Full
Imp2D
Line
500
100
50
10
5
Convergence Time in ms (log scale)
500
1000
5000
10000
Number of Nodes (log scale)

Figure 3

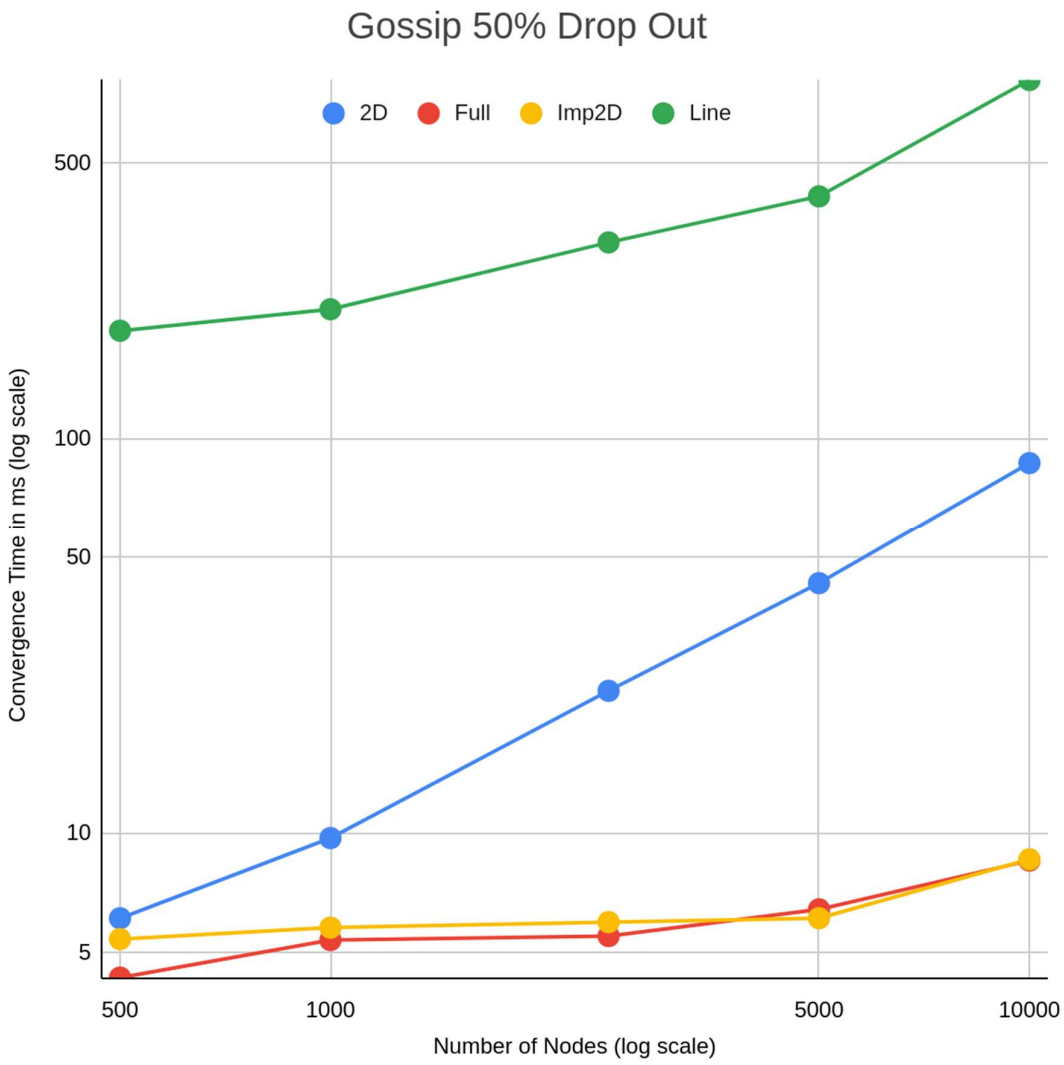
Gossip 50% Drop Out
2D
Full
Imp2D
Line
500
100
50
10
5
Convergence Time in ms (log scale)
500
1000
5000
10000
Number of Nodes (log scale)

Figure 4

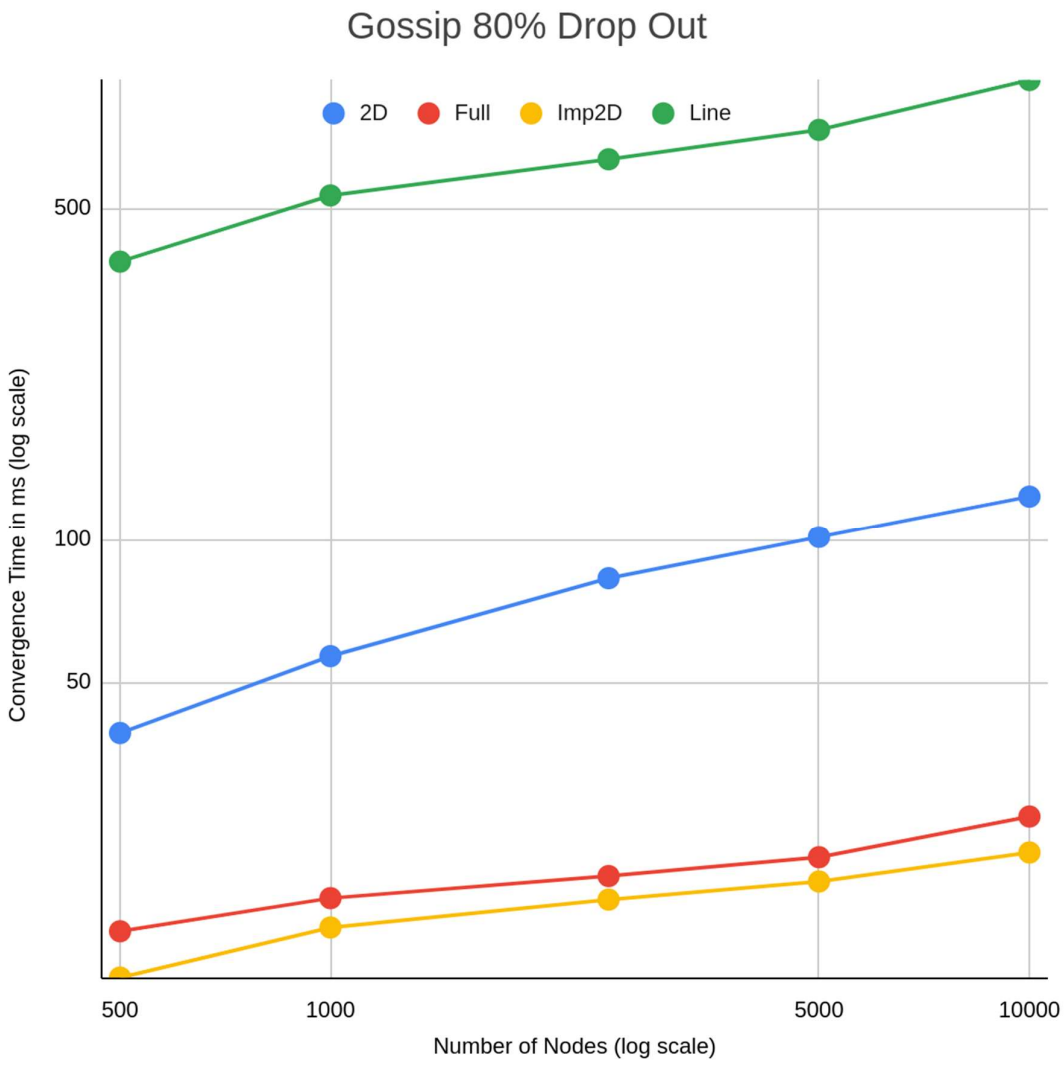
Gossip 80% Drop Out
2D
Full
Imp2D
Line
500
100
50
Convergence Time in ms (log scale)
500
1000
5000
10000
Number of Nodes (log scale)

Figure 5