

DISTRIBUTED OPERATING SYSTEM PRINCIPLES

Project 2 : Gossip Algorithm

Sumanth Devarasetty - UFID: 57829275 ; Praveen Kumar Dande UFID: 62316196

Gossip Algorithm for information Propagation

The Gossip algorithm involves the following:

•Starting:

A participant(actor) is told/sent a rumor(fact) by the main process

•Step:

Each actor selects a random neighbor and tells it the rumor

•Termination:

Each actor keeps track of rumors and how many times it has heard the rumor. It stops transmitting once it has heard the rumor 10 times (10 is arbitrary, you can select other values).

Instructions to run the code.

- Download project2.erl from the zip file.
- On the device with the file, move to the relevant directory and run the following commands in the terminal.

```
erl
c(project2).
project2:start(numNodes, topology, algorithm).
```

- numNodes is the Number of Nodes, topology is one of ["line", "two_d", "full", "imp_3d"]
- Algorithm is either "gossip" or "pushsum"

1. What is working

Topologies:

- Line: All the actors are arranged in a line, Each actor has only 2 neighbors (one left and one right, unless it is the first or last element of the line).
- Full Network: Every actor is a neighbor of all other actors i.e. every actor can talk directly to any other actor.
- 2D Grid: Actors form a 2D grid. The actors can only talk to the grid neighbors.
- Imperfect 3D Grid: Grid arrangement but one random other neighbor is selected from the list of all actors.

We were able to run all the line, full, 2D, and Imp3D topologies in any combination with gossip or push-sum protocol. The convergence in Gossip protocol is achieved when all the nodes have converged. A node is set to be converged when it listens to the message for the 10th time.

After convergence, the node stops transmitting the message to its neighbor. Once the network is converged i.e. all nodes are converged, the total time for convergence is printed out.

Gossip

- Line topology is the slowest to converge in both the algorithms. This is due to the fact that it has only access to 2 neighbors (left and right node).
- Full topology is the fastest to converge till 2000 nodes. This is because it has the highest number of neighbours when compared to all the topologies, so the convergence is faster to achieve in this case. After 2000+ nodes we found that imperfect 3D is faster than full topology. We think the reason for this might be the increase in distance between the nodes in full topology as the number of nodes increases.
- Generally both 2D and imperfect 3D achieved the convergence time in between the line and full with imperfect 3D performing slightly better than 2D performance.

Push Sum

The PushSum network works by sending message s and w as parameter to an actor. The initial value of s is equal to the index of the actor and $w = 1$. The propagation stops when an actor's s/w ratio does not change for 3 times in a row (i.e stays within limit of 10^{-10})

The push sum algorithm is able to converge in all topologies but the convergence time is higher than gossip. We believe this is due to the fact we try to reduce the value s/w ratio till it does stop changing the ratio for consecutive three times. This allows in more messages being transmitted compared to the gossip algorithm allowing the network to converge which in turn causes increased convergence time.

2. What is the largest network you managed to deal with for each type of topology and algorithm?

The largest network that we have managed to solve is 30000 nodes for each topology with Gossip algorithm. For Pushsum algorithm We decided to limit it to 10000 nodes as the time for convergence gets higher with the increase in number of nodes when compared to gossip. The code works fine for even more number of nodes.

All the sample outputs and graphs are in the Report.pdf