Group Zipf: First DC Homework

Dilara Isikli 1891292
Alexandru Melnic 1692625
Vedat Katirci 1908990
Lorenzo Loretucci 1903794

April 13, 2020

PART 1

1.1 Complexities

In the following graph for each of the 3 methods: *BFS*, *Laplacian Eigenvalues* and *Adjancency matrix irreducibility* we measured the running time of the algorithms varying the number of nodes K.

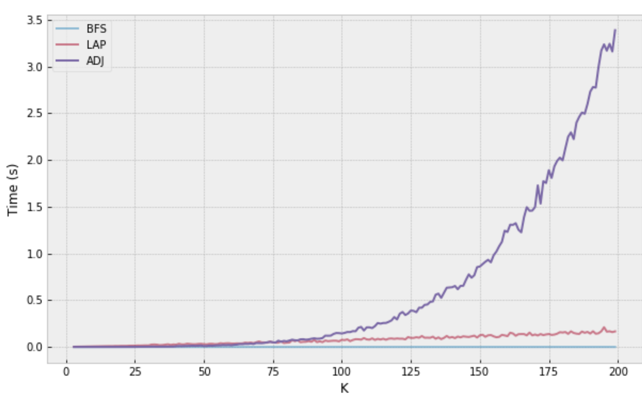

Figure 1: Complexities for the three methods

1.2 p-ER random graph

For this Monte Carlo simulation for each p ranging from 0 to 0.1 a p-ER random graph with number of nodes K=100 was generated 1000 times, and at each step the connectivity condition was checked through the `is_connected()` function from the `networkx` library. At the end, for each p, the fraction, out of the 1000 iterations, of the connected graphs was saved and plotted. Plot in Figure 2.

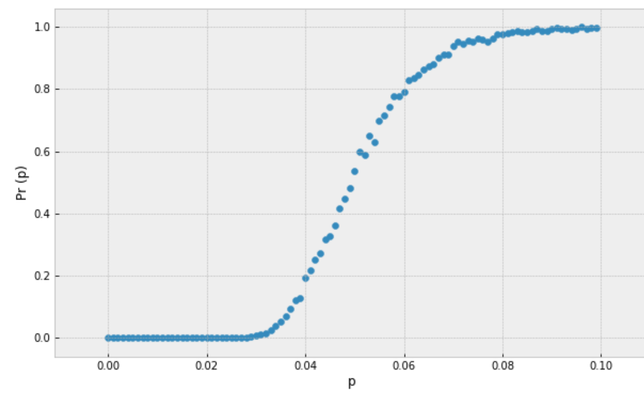

Figure 2: p-ER random graph

1.3 r-regular random graph

The MC simulation was conducted as before for a 2-regular random graph and for a 8-regular random graph, with K, the number of nodes, ranging from 10 to 100. Plots in Figure 3.

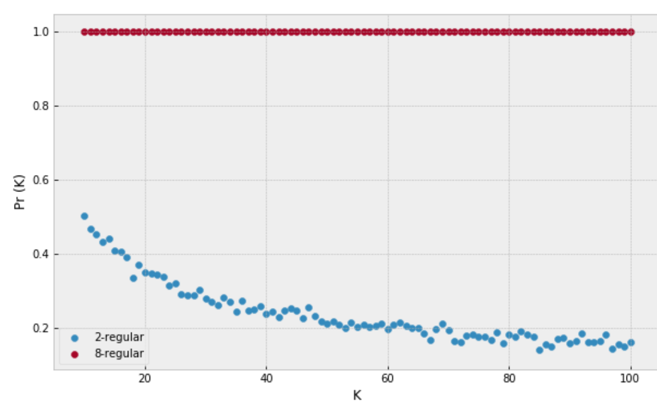

Figure 3: 2-regular and 8-regular random graphs

PART 2

2.1 Confronting Jellyfish and Fat tree topologies

For an r-regular random graph, the number of servers N that can be placed is given by the expression :

$$N = S(n - r)$$

where S is the number of the n ports switches. For a 3-layer Fat Tree topology the number of switches and servers is strictly dependent on the number of ports of the switches: $N = n^3/4$, $S = 5n^2/4$. Fixing the number of switches and the number of servers of the Fat Tree for the Jellyfish topology and expliciting r we get:

$$r = n - \frac{N}{S} = n - \frac{n^3/4}{5n^2/4} = \frac{4n}{5}$$

2.2 Throughput expression

The general bound for the Throughput is:

$$TH \leq \frac{L}{\overline{h}\nu_f}$$

where L is the number of links, $\overline{h}$ the average shortest path and ν_f is the number of traffic flows, that for an all-to-all traffic is: $\nu_f = N(N-1)/2$, with N the number of servers. We will only treat the all-to-all traffic case.
We want to compare the performances of Fat Tree and Jellyfish topologies with same equipment, for this reason the TH bound, as a function of only n and $\overline{h}$, is the same for both topologies. Only $\overline{h}$ will be different, as we will see in the next section.
In a Fat Tree the number of links L is only dependent on the number of ports in the switches n: $L = 3N = 3n^3/4$, with N the number of servers. Substituting this value to the general TH expression we get:

$$TH \leq \frac{L}{\overline{h}\nu_f} = \frac{6}{\overline{h}(N-1)} = \frac{24}{\overline{h}(n^3-4)}$$

2.3 Table with computed values

Average shortest path for Fat Tree

Starting the path from 1 server, to reach another server attached in the same switch the shortest path length is 2; to reach another server in the same pod and not attached to the same switch the path length is 4; to reach another server located in a different pod the path length is 6. Keeping in mind that the number of servers attached to an edge layer switch is equal to $n/2$: for the first case the numbers of reachable servers is $N^* = \frac{n}{2} - 1$; for the second case the number of reachable servers is $N^* = \frac{n^2}{4} - \frac{n}{2}$; for the last case the number of reachable servers is $N^* = \frac{n^3}{4} - \frac{n^2}{4}$.
Assuming that starting from one server the reachable nodes are all equally likely, we can evaluate the average shortest path $\overline{h}$ with a weighted average, where the weights are the fractions, out of the total servers, for the 3 different path lengths:

$$\overline{h} = \frac{2(n/2-1) + 4(n^2/4 - n/2) + 6(n^3/4 - n^2/4)}{n^3/4 - 1} = \\ = \frac{6n^3 - 2n^2 - 4n - 8}{n^3 - 4}$$

Table

The $\overline{h}$ for an r-regular random graph is computed through the Cerf et alii (1974) lower bound. We decided to round to the floor the N value and to the ceil the S and L values.
In the following table n is number of ports in switch, N is number of servers, S is number of switches, L is the number of links and the TH values are the Throughputs for the Fat Tree and the Jellyfish topologies.

n	N	S	L	$TH_{FatTree}$	$TH_{r-regular}$
5	31	32	94	0.035714	0.084805
10	250	125	750	0.004172	0.009959
15	843	282	2531	0.001216	0.002913
20	2000	500	6000	0.000509	0.001222
25	3906	782	11719	0.000259	0.000624
30	6750	1125	20250	0.000149	0.000360
35	10718	1532	32157	0.000094	0.000227
40	16000	2000	48000	0.000063	0.000151
45	22781	2532	68344	0.000044	0.000106
50	31250	3125	93750	0.000032	0.000078