

# CSC 591 Algorithms for Data Guided Business Intelligence

## Project- Network properties

Tushar Himmat Dahibhate  
tdahibh

### Degree Distribution:

**1. Generate a few random graphs. You can do this using networkx's random graph generators. Do the random graphs you tested appear to be scale free? (Include degree distribution with your answer).**

a. gnp1

$\gamma = 4.939081128857231$  (Calculated using the power law package)

As  $2 < 3 < \gamma$ , we can say that the gnp1 is NOT scale-free.

b. gnp2

$\gamma = 54.58226131097677$  (Calculated using the power law package)

As  $2 < 3 < \gamma$ , we can say that the gnp2 is NOT scale-free.

c. gnm1

$\gamma = 2.887544867499438$  (Calculated using the power law package)

As  $2 < \gamma < 3$ , we can say that the gnm1 is scale-free.

d. gnm2

$\gamma = 9.620663765606892$  (Calculated using the power law package)

As  $2 < 3 < \gamma$ , we can say that the gnm2 is NOT scale-free.

**2. Do the Stanford graphs provided to you appear to be scale-free?**

a. amazon.graph.small

$\gamma = 2.3948604146303003$  (Calculated using the power law package)

As  $2 < \gamma < 3$ , we can say that the amazon.graph.small is scale-free.

b. dblp.graph.small

$\gamma = 1.6077866723222844$  (Calculated using the power law package)

As  $\gamma < 2 < 3$ , we can say that the dblp.graph.small is NOT scale-free.

c. youtube.graph.small

$\gamma = 1.367441761614622$  (Calculated using the power law package)

As  $\gamma < 2 < 3$ , we can say that the youtube.graph.small is NOT scale-free.

d. amazon.graph.large

$\gamma = 1.3255773302967864$  (Calculated using the power law package)

As  $\gamma < 2 < 3$ , we can say that the amazon.graph.large is NOT scale-free.

e. dblp.graph.large

$\gamma = 1.3143917250729842$  (Calculated using the power law package)

As  $\gamma < 2 < 3$ , we can say that the dblp.graph.large is NOT scale-free.

f. youtube.graph.large

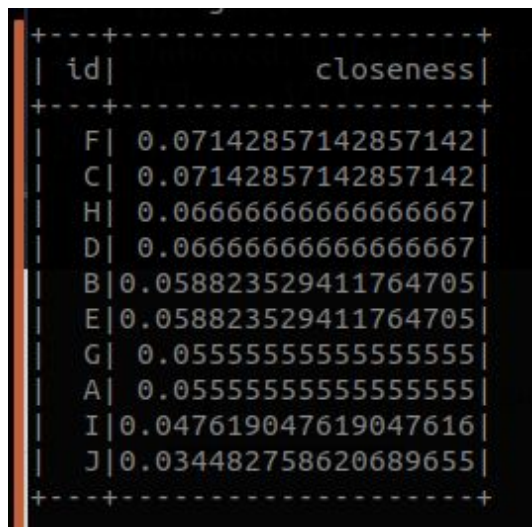
$\gamma = 1.560511013438854$  (Calculated using the power law package)

As  $\gamma < 2 < 3$ , we can say that the youtube.graph.large is NOT scale-free.

### Centrality:

1. Rank the nodes from highest to lowest closeness centrality.

**Answer:**



```
+---+-----+
| id|      closeness|
+---+-----+
| F| 0.07142857142857142|
| C| 0.07142857142857142|
| H| 0.06666666666666667|
| D| 0.06666666666666667|
| B| 0.058823529411764705|
| E| 0.058823529411764705|
| G| 0.05555555555555555|
| A| 0.05555555555555555|
| I| 0.047619047619047616|
| J| 0.034482758620689655|
+---+-----+
```

id	closeness
F	0.07142857142857142
C	0.07142857142857142
H	0.06666666666666667
D	0.06666666666666667
B	0.058823529411764705
E	0.058823529411764705
G	0.05555555555555555
A	0.05555555555555555
I	0.047619047619047616
J	0.034482758620689655

2. Suppose we had some centralized data that would sit on one machine but would be shared with all computers on the network. Which two machines would be the best

**candidates to hold this data based on other machines having few hops to access this data?**

**Answer:**

The nodes F(0.07412857142857142) and C(0.07412857142857142) will be the best candidates as they have the highest closeness.

**Articulation Points:**

- 1. In this example, which members should have been targeted to best disrupt communication in the organization?**

**Answer:**

```
Execution time: 285.234828949 seconds
Articulation points:
+-----+-----+
|id      |articulation|
+-----+-----+
|Mohamed Atta      |1|
|Usman Bandukra    |1|
|Mamoun Darkazanli |1|
|Essid Sami Ben Khemais|1|
|Djamal Beghal     |1|
|Nawaf Alhazmi     |1|
|Raed Hijazi       |1|
+-----+-----+
-----
```