JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Generating scale-free networks with high clustering

Marcel Llopis

Johns Hopkins
Whiting School of Engineering

Outline

The objective of this project is to find a way to find a way to create a graph that is both (1) scale free, and (2) with high clustering.

The proposed overall process is as follows:

1. Create a scale-free graph
2. Increase clustering of scale-free graph

Creating a scale-free graph

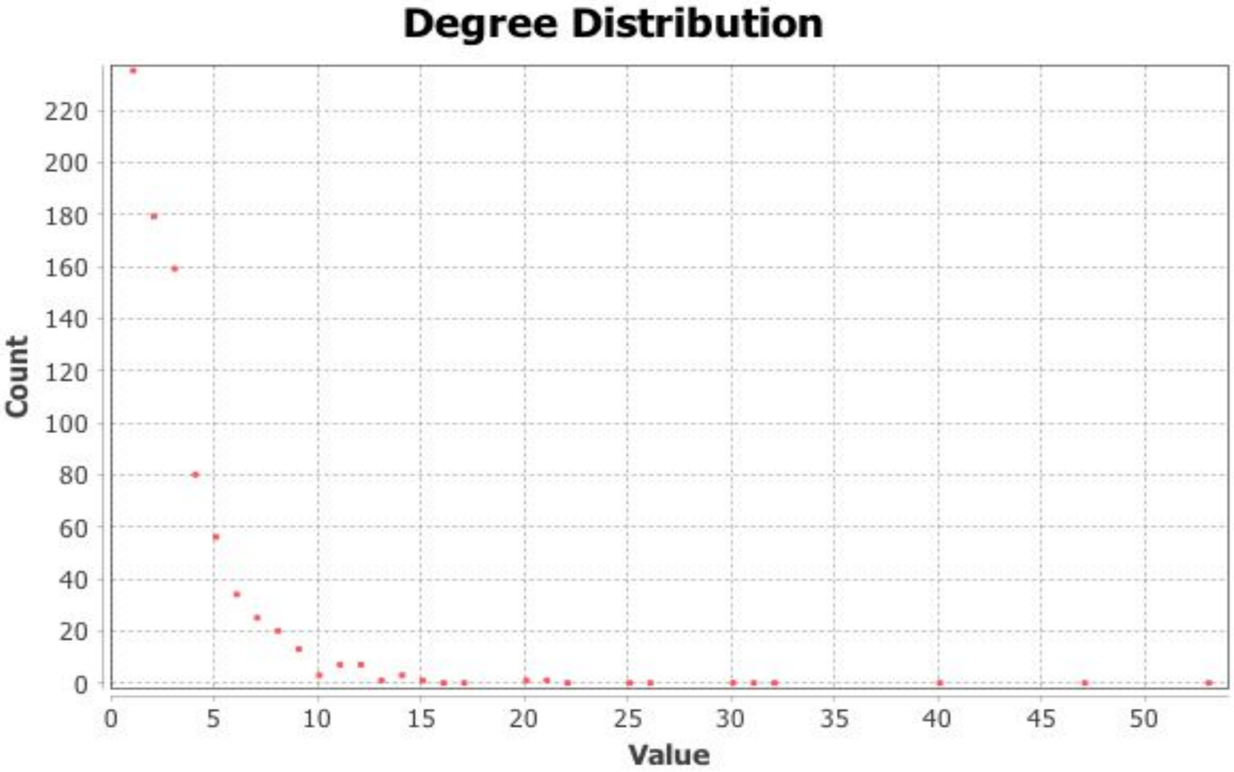
We know that the Barabasi-Albert method lets us create scale-free graphs. Scale-freeness -> Rich get richer!

According to to original paper below, attachment of a new vertex with another *v* occurs with a probability p defined as $p(v)=degree(v)/|E|$.

A more robust p can be defined as $p(v)=(degree(v)+1)/(|E|+|V|)$. This ensures the probability of attachment for any existing isolated vertex would be >0.

Creating a scale-free graph

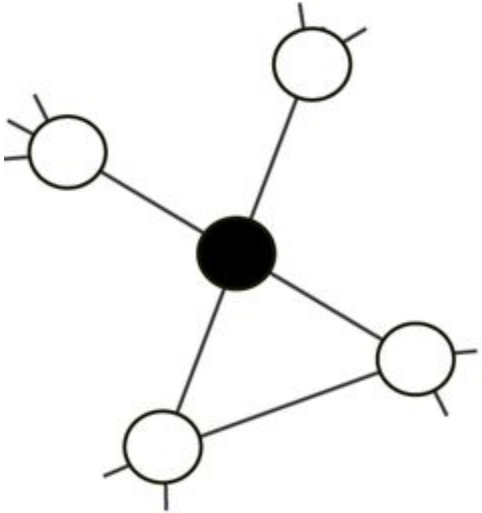
With this approach, as an example, we create a ~1000 node graph with the following degree distribution.



Average Degree: 3.667

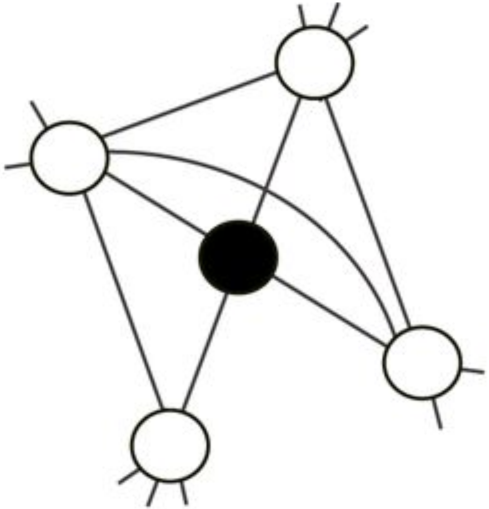
Increasing clustering

We know clustering increases when a node’s neighbors are more connected



$$k_i = 4$$

$$C_i = \frac{1}{6}$$

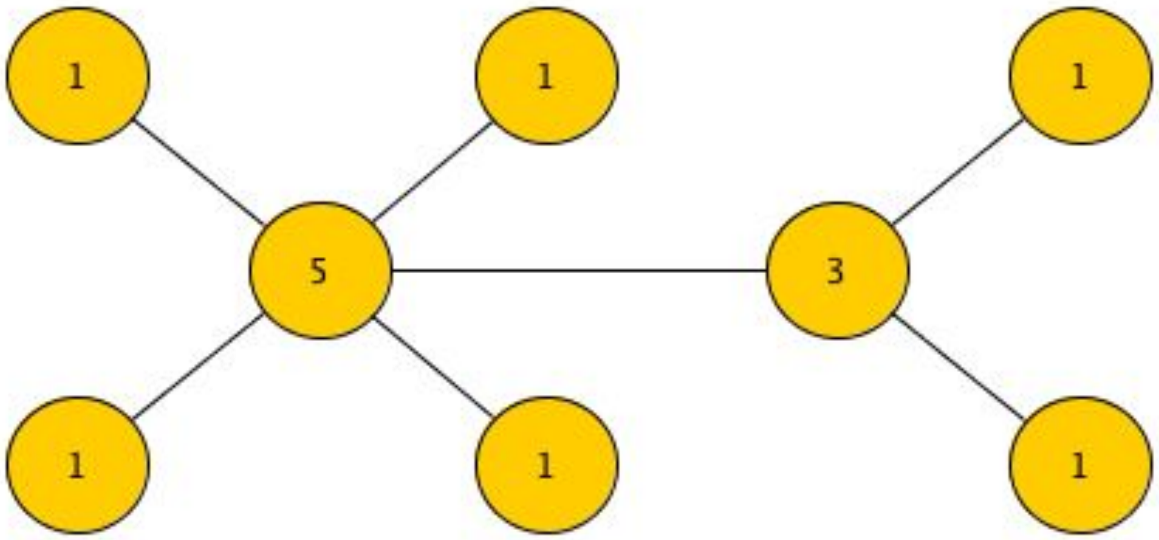


$$k_i = 4$$

$$C_i = \frac{2}{3}$$

Increasing clustering

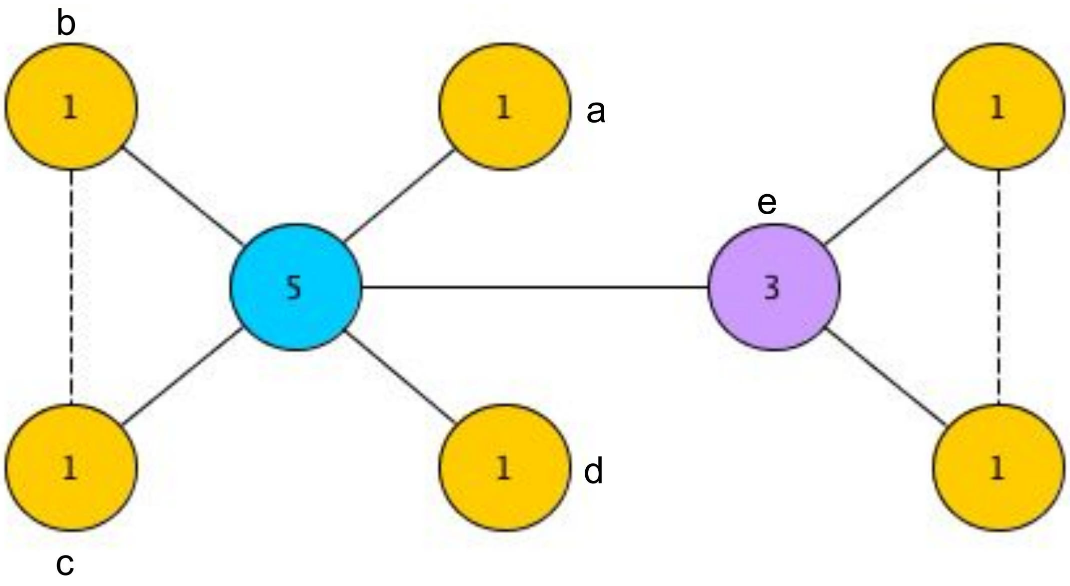
Following the "the rich get richer" approach. We could increase clustering by increasing the connections among the neighbors of "richer" (higher degree) nodes.



Increasing clustering

Probability of adding an edge between neighbors of vertex *v*

p(v)=degree(v)/sumDegreesOverOne()



p(blue) = 5/(5+3) = 0.625

p(purple) = 3/(5+3) = 0.375

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

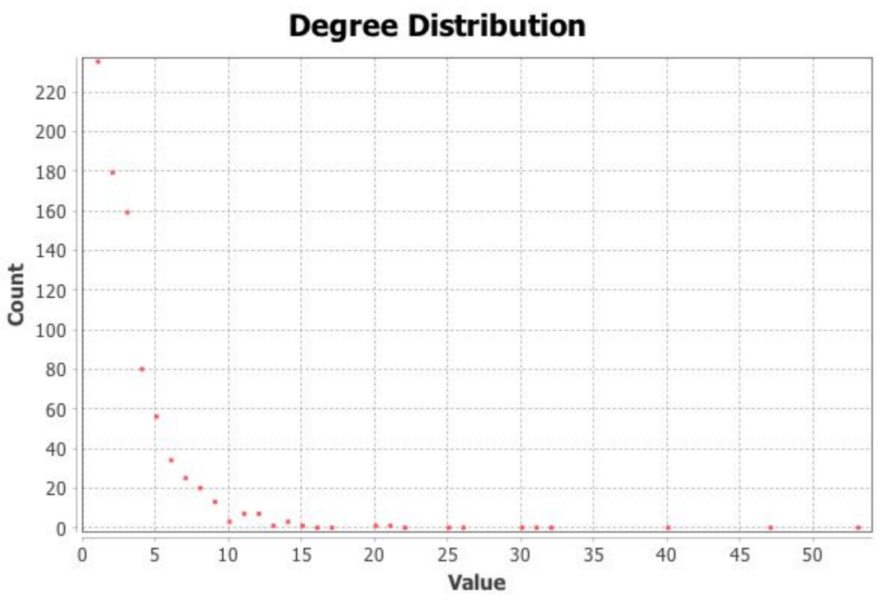
Increasing clustering

Analysis:

- The initial ~1000 node scale-free graph generated in the first step had a clustering coefficient of 0.072. Running the clustering-enhancing algorithm raises that to 0.142.
- The algorithm can be run iteratively until one reaches the clustering coefficient desired. Run #10 provides a clustering coefficient of 0.502.
- Scale-freeness is maintained!

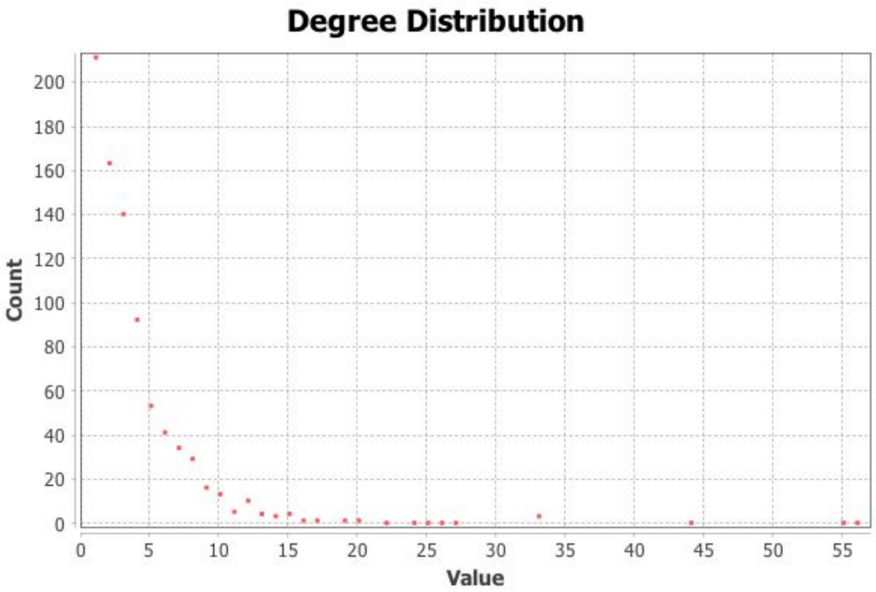
Increasing clustering

Scale-free graph



853 vertices, 1564 edges
Avg. degree: 3.667, Clustering Coefficient: 0.072

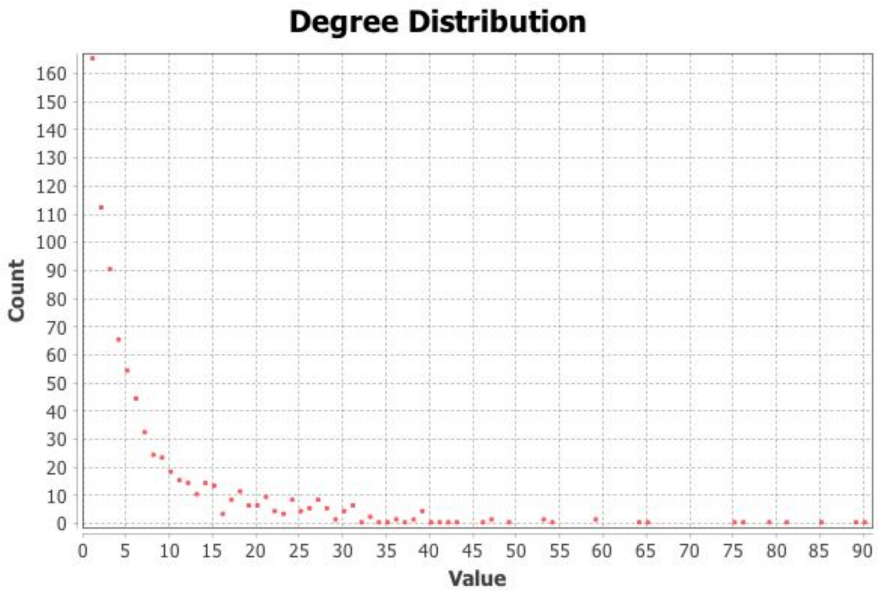
Clustering run #1



853 vertices, 1792 edges
Avg. degree: 4.202, Clustering Coefficient: 0.142

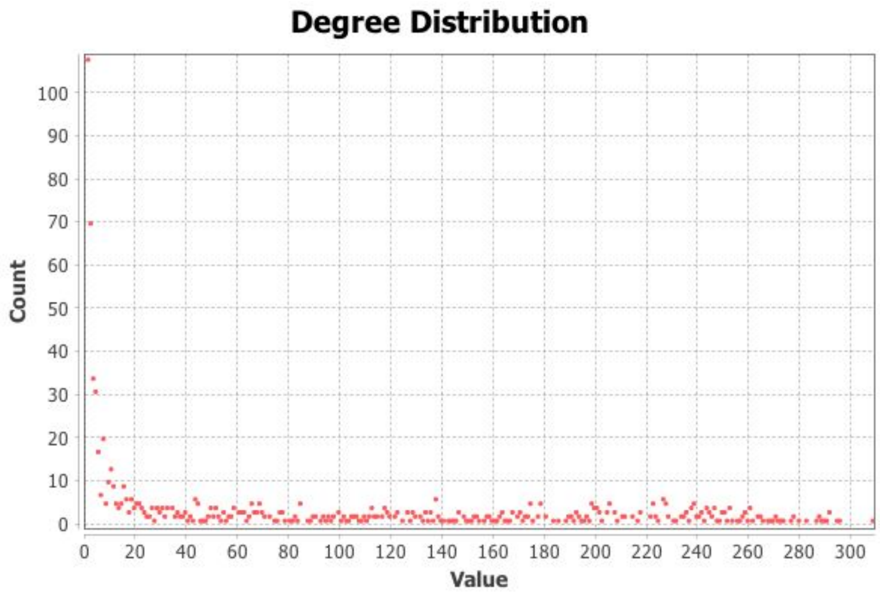
Increasing clustering

Clustering run #5



853 vertices, 3799 edges
Avg. degree: 8.907, Clustering Coefficient: 0.255

Clustering run #10



853 vertices, 35658 edges
Avg. degree: 83.606, Clustering Coefficient: 0.502

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Conclusion

- Scale-freeness can be maintained while increasing clustering in a 2-step algorithm.
- The clustering increasing algorithm:
  - Can be parallelized
  - Benefits more vertices with higher degrees - the rich get richer again.
  - Increases clustering fast initially but not particularly fast after a few runs.
  - Reaches a point of diminishing returns if run several times.

Thank you!!! I can provide the code and GraphML datasets created with the code if you drop me an e-mail.