Heidelberg University Institute of Computer Science Database Systems Research Group

Lecture: Complex Network Analysis

Prof. Dr. Michael Gertz

Assignment 4 Scale-Free Networks

https://github.com/nilskre/CNA_assignments

Team Member: Patrick Günther, 3660886,

Applied Computer Science rh269@stud.uni-heidelberg.de

Team Member: Felix Hausberger, 3661293,

Applied Computer Science eb260@stud.uni-heidelberg.de

Team Member: Nils Krehl, 3664130,

Applied Computer Science pu268@stud.uni-heidelberg.de

1 Problem 4-1 Power laws

Consider the degree distribution functions p_k of the following two undirected networks:

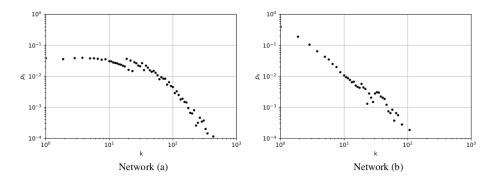


Figure 1: Degree distribution of two networks.

1. One of these networks is approximately scale-free, the other is not. Identify the scale-free network and explain how you came to your conclusion.

Network b is is scale free. Figure 1 shows the log-log plots of the degree distribution of the two networks a and b. For scale-free networks the degree distribution forms a straight line across the diagonal. This is given for network b.

2. A particular network is believed to have a degree distribution that follows a power law. A random sample of nodes is taken and their degrees are measured. The degrees of the first twenty nodes with degrees 10 or greater are:

16 17 10 26 13 14 28 45 10 12

12 10 136 16 25 36 12 14 22 10

For $K_{min} = 10$, estimate the exponent γ of the power law using the estimation method presented in the lecture.

Hint: you only have to estimate the exponent (see Slide 4-33)!

Calculate the error σ of your estimation using the equation $\sigma = \frac{\gamma - 1}{\sqrt{N}}$

The following solution is based on Slide 4-28 Estimating the degree exponent. From step 1. (Estimate value of the degree exponent corresponding to K_{min}) following formula is used:

$$\gamma = 1 + N \left[\sum_{i=0}^{N} \left(\ln\left(\frac{k_i}{K_{min} - \frac{1}{2}}\right) \right) \right]^{-1}$$
 (1)

Now the values N=20, k_i from the list above and $K_{min}=10$ are inserted:

$$\gamma = 1 + 20 \left[ln\left(\frac{16}{20 - \frac{1}{2}}\right) + ln\left(\frac{17}{20 - \frac{1}{2}}\right) + ln(...) + ... \right]^{-1}$$
$$\gamma \approx -14, 17$$

Based on N and γ the error σ is calculated:

$$\sigma = \frac{\gamma - 1}{\sqrt{N}} = \frac{-14, 17 - 1}{\sqrt{20}} = -3, 39$$

Now, one would normally iteratively increase $K_m in$ from $K_m in = k_m in$ to $K_m in = k_m ax$ and then choose $K_m in$ such that it minimizes the error σ . Then the resulting power law distribution would fit best to the real distribution. But as the task description states to execute this procedure only for $K_m in = 10$, we stop here.

2 Problem 4-2 Friendship Paradox

Let p_k be the probability of a randomly chosen node to have the degree k. If we instead randomly choose an edge, let $q_k = \frac{1}{C} \cdot k \cdot p_k$ be the probability that a node at one of its ends has degree k, where $\frac{1}{C}$ is a normalization factor.

- 1. Assuming that the network has a power-law degree distribution with $2 < \gamma < 3$, minimum degree k_{min} , and maximum degree k_{max} , show how to derive the normalization factor C.
- 2. q_k is also the probability that a randomly chosen node has a neighbor with degree k. Show how to compute the average degree of the neighbors of a randomly chosen node.
- 3. Given a power-law degree distribution network, with $N=10^4$, $\gamma=2.3$, $k_{min}=1$ and $k_{max}=1000$. Compute the average degree of the neighbors of a randomly chosen node and compare this to $\langle k \rangle$.
- 4. Try to explain the paradox of subtask 3, i.e., explain why on average, the neighbors of a node have more neighbors than the node itself?

3 Problem 4-3 Network Measures of Real Graphs

See problem_3.ipynb.