Complex Networks

Master of Science in Electrical Engineering

Erivelton Geraldo Nepomuceno

Department of Electrical Engineering
Federal University of São João del-Rei

May 23, 2019

5. Generalised Random Graphs

- Many real networks, such as the network of hyperlinks in the World Wide Web, have a degree distribution p_k that follows a power law for large values of the degree k.
- This means that, although most nodes have a small number of links, the probability of finding nodes with a very large degree is not negligible.

- Both ER random graphs models small-world models studied in the previous chapter produce instead networks with a degree distribution peaked at the average degree $\langle k \rangle$ and rapidly decreasing for large values of k.
- Configuration model: a method to generate random graphs with any form of p_k.
- Properties of random graphs with power-law degree distributions $p_k \approx k^{-\gamma}$, with $2 < \gamma \leq 3$.

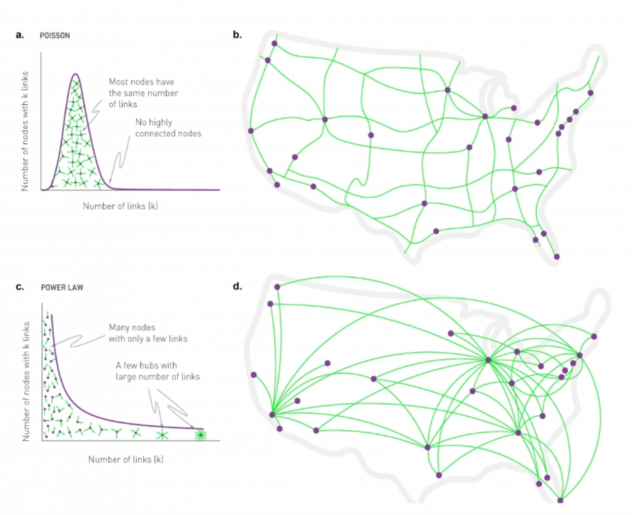

Figure 1: In a random network most nodes have comparable degrees and hence hubs are forbidden. Hubs are not only tolerated, but are expected in scale-free networks. Source: http://networksciencebook.com

5.1 The World Wide Web

- We have shown that various real-world networks have short characteristic path lengths and high clustering coefficients.
- The WWW is a directed network: given two webpages *i* and *j*, there can be a hyperlink pointing from *i* to *j*, a hyperlink pointing from *j* to *i*, or both.

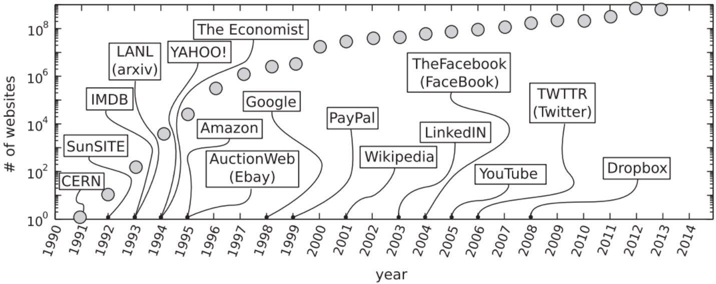

Figure 2: The growth of the World Wide Web over the years from its initial creation at CERN in 1991.

- In a directed graph we have to distinguish between the in-degree k_i^{in} and the out-degree k_i^{out} of a vertex i.
- Therefore, in any network sample of the WWW, there are two degree distributions: p_k^{in} and p_k^{out}.
- We can extract empirically the in- and the out degree distribution by counting respectively the number $N_{k^{in}=k}$ of nodes with k ingoing links, and the number $N_{k^{out}=k}$ of nodes with k outgoing links.

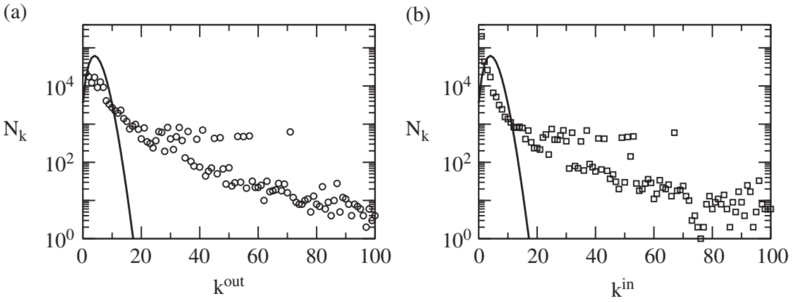

Figure 3: Plot, as a function of k, the histograms of $N_{k^{in}=k}$ and $N_{k^{out}=k}$ obtained for the Notre Dame network, altogether with the Poisson distribution expected in a random graph with the same total number of nodes and links. The degree distributions of the WWW are totally different from that of a random graph.

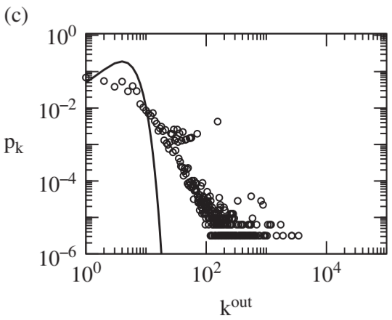

Figure 4: To observe entirely the long tail of the distribution in a single plot it is much better to use a log–log scale, plotting directly the degree distributions. The network of the WWW is the result of billions of agents acting independently, so there were no a priori reasons to find such large deviations from a random graph!?

- The fact that the points stay on a straight line in a log–log plot means that we can write

$$\log p_k = a + b \log k \tag{1}$$

- This is equivalent to say that the probability p_k of having a node with degree k is a power law:

$$p_k = ck^{-\gamma} \tag{2}$$

where $c = 10^a$ and $\gamma = -b > 0$.

Computational Exercise

1. Plot the histogram of degree distribution of the following networks:
   - Kindergarten
   - Movie actors
   - Astrophysics coauthorship
   - WWW NotreDame
2. Create node equivalent ER model.
3. Compare the histograms.
4. Using functions `hist` and `polyfit` estimate the power law of these 8 networks.
5. Discuss with a colleague the results.