Lecture Title and Date

25m10c - Networks - Network Generation Models (03/05/2025)

Objectives of the Lecture

- Introduction to and overview of models for network generation.
- Historical references to previously applied models.
- Modern networks modeling biological complexity.

Key Concepts and Definitions

- Network Topology -
  A major building block for data communication. Consist of nodes and paths often called links. Communicates a physical/logical representation of relationships between nodes.
- Erdos-Renyi [ER] model -
  Model developed in 1960. Starts with N vertices and no edges. Connect each pair of vertices with probability P_{ER}.
- Small world network -
  A simple connected graph that exhibits two properties:
    1. A large clustering coefficient.
    2. Small characteristic path length.
- Watts-Strogatz [WS] model -
  Model developed in 1998. Start with a simple network of N vertices. Rewire each edge with probability P_{ws}. P_{ws} can have a value between 0 and 1.
- Scale-free network -
  A network generation method that uses two basic mechanisms to establish a link to a new node: growth and preferential attachment. This method leads to systems that are dominated by hubs of nodes that have a large number of links.
  Think "rich get richer".
- Barabasi-Albert [BA] model -
  Scale-free network model developed in 1999. Has a variable number of vertices and the probability of finding a highly connected node decreases exponentially with K in the following $P(K) \sim K^{-y}$.

Main Content/Topics

To begin to understand complex network relationships it is important to first understand what exactly network topology accomplishes. The topology of a network includes two simple characteristics: nodes and paths. As a baseline example we can think of two dots (nodes) and a line (path) drawn between them. This would mean in network analysis that the two dots share some sort of relationship. This establishes foundational knowledge that network topology communicates both a physical and logical representation of relationships between nodes in a network.

Now, rather than our simple two node example, let's take a look at a 10 node network. In this 10 node network, we want to generate a random linkage to better understand the effect of randomness on a system. If we assume that each node connection has a similar probability, then we can use the Erdos-Renyi [ER] model. The ER model starts with N vertices and zero edges. Then each pair of nodes is connected with probability P_{ER}. It is important to emphasize that many properties in these graphs appear suddenly, more specifically at a threshold of $P_{ER}(N)$.

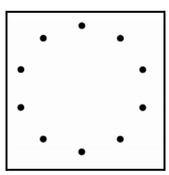
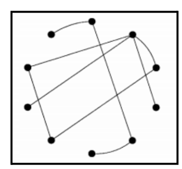
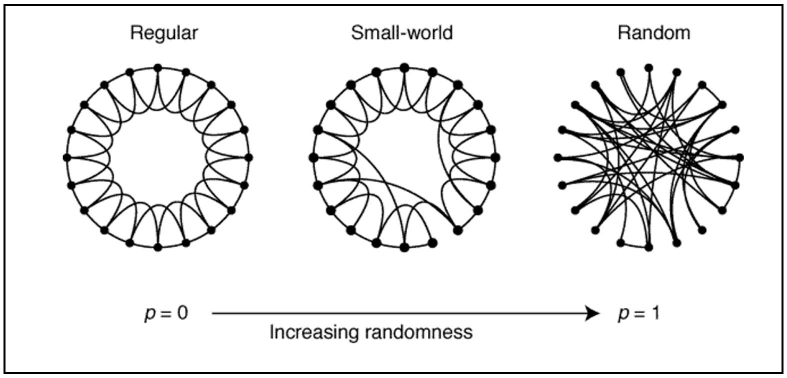
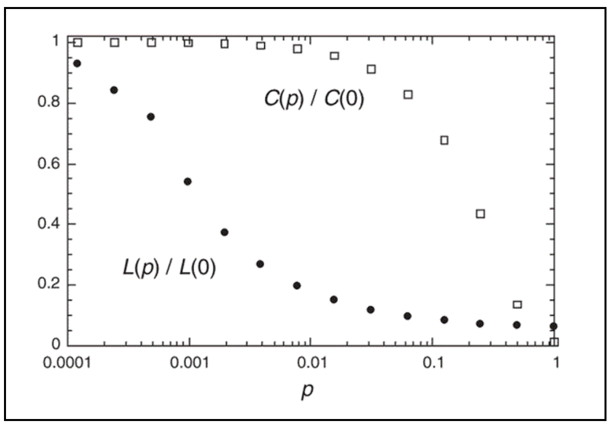
$N = 10, P_{ER} = 0$	$N = 10, P_{ER} = 0.2$
	

Figure 1. Represented network graphs from Erdos-Renyi model with N nodes and probability of P_{ER}.

ER models are a good example of what random networks may look like, but with a few hangups that prevent it from being completely accurate. This leads us to the next modeling theory, the small world network and the Watts-Strogatz [WS] model. The WS model has edges installed at the regular state, meaning there is a high clustering coefficient and highly characteristic path length that is consistent between similar nodes. The WS model then propagates the regular network and rewires each edge with probability p. If p stays at 0, there will be no change in the network, thus it is a regular predictive network. As p approaches 1, the randomness of the network increases. This is reflected in having a low clustering coefficient and having an uncharacteristic path length. For values of p that are between 0 and 1, these networks are called small world networks.



In small world networks, there are a few important characteristics to consider. First is that there is a broad interval of p for which the path character is low while the clustering coefficient remains large. These networks are noted for their local connectivity and global reach and coincidently these networks are common.

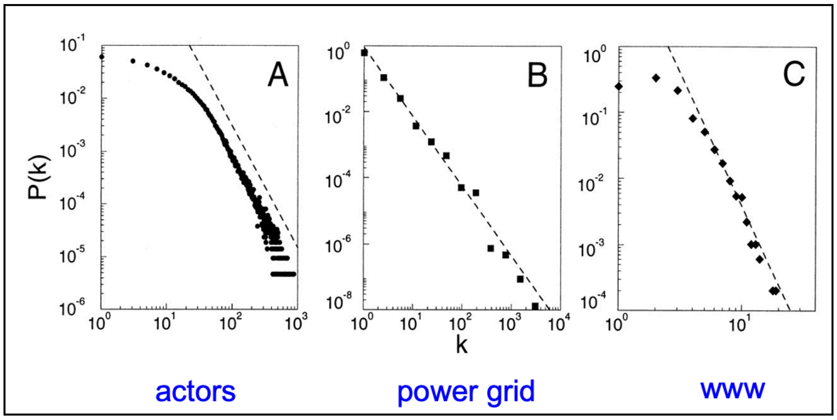


Although small world networks based on the WS model are common, there are two major problems that are not addressed in the WS nor ER model. The number of nodes is consistent and the probability that two nodes connect is uniform. This is addressed by the Barabasi-Albert [BA] model. In a BA model there is growth, where at every timestamp there is an addition of another node. How this node connects is based on the concept of preferential attachment, where the probability that a new node is connected to node i depends on the overall connectivity of that node. This is shown in the following equation:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

What isn't so surprising about this relation is that the probability of finding a highly connected node decreases exponentially as the number of nodes increases. This is an example of scale-free networks, where the number of nodes can increase infinitely, but the connections to each node subsequently decreases. This is shown below in the equation and figure.

$$P(K) \sim K^{-\gamma}$$



Since scale-free network evolves primarily through preferential attachment, it can be used to describe some biological phenomena such as gene duplication. From the lecture it is stated that in an interaction network, gene duplication followed by mutation of the duplicated gene leads to preferential attachment. Meaning that the partners of a hub are more likely to be duplicated than the partners of a non-hub. An easy way to remember this is to think of “the rich get richer”.

List all suggested reading here and please answer:

1. Albert-László Barabási,Réka Albert,Emergence of Scaling in Random Networks.*Science***286**,509-512(1999).DOI:10.1126/science.286.5439.509
2. Watts, D., Strogatz, S. Collective dynamics of ‘small-world’ networks. *Nature* 393, 440–442 (1998). https://doi.org/10.1038/30918
3. Barabási AL, Bonabeau E. Scale-free networks. Sci Am. 2003 May;288(5):60-9. doi: 10.1038/scientificamerican0503-60. PMID: 12701331.
4. Albert, Réka, and Albert-László Barabási. “Statistical Mechanics of Complex Networks.” *Rev. Mod. Phys.* 74, no. 1 (January 2002): 47–97. https://doi.org/10.1103/RevModPhys.74.47.

These readings are helpful since they are source material for a ton of seminal information. These papers lay the foundation for our current understanding of how networks generation models were developed overtime. Seeing the transition and development of models is essential for critically evaluating current models to further improve them for more accurate predictions.