Lesson 4

Random graphs

Sergio Barbarossa
Graph models

1. Uncorrelated random graph (Erdős, Rényi)

N nodes are connected through n edges which are chosen randomly from the \( \frac{N(N - 1)}{2} \) possible configurations.

2. Binomial model (Gilbert model)

Every pair of nodes is connected with probability p.

The total number of edges is a random variable with expected value

\[
p \, \frac{N(N - 1)}{2}
\]
Graph models

One of the most interesting features of random graphs is that there exists a critical probability scaling law $p_c(N)$ such that:

If $p(N)$ grows faster than $p_c(N)$, then almost every graph has property Q (like, e.g., connectivity)

If $p(N)$ grows slower than $p_c(N)$, then almost every graph fails to have property Q

This feature establishes a link with percolation theory
Graph features

**Degree distribution**

In a random graph with connection probability $p$, the degree $k_i$ of a node $i$ follows a binomial distribution

$$P(k_i = k) = C^k_{N-1} p^k (1 - p)^{N-1-k}$$

With a good approximation, this is also the degree distribution of a random graph

For large $N$ and infinitesimal $p$, such that $\lim_{N \to \infty} p(N)N = \lambda$

The degree distribution can be approximated by the Poisson law

$$P(k) \approx e^{-pN} \frac{(pN)^k}{k!}$$
Graph features

Diameter

Def.: The diameter of a graph is the maximal distance between any pair of nodes

Denoting with $k_{ave}$ the average degree

If $k_{ave} < 1$ the graph is composed of isolated trees

If $k_{ave} > 1$ a giant cluster appears

If $k_{ave} > \log N$ the graph is totally connected and the diameter is concentrated around

$$D \propto \frac{\log(N)}{\log(k_{ave})}$$
Clustering coefficient

The clustering coefficient $C_i$ for a vertex $v_i$ is given by the proportion of links between the vertices within its neighborhood divided by the max number of links that could possibly exist between them.

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i - 1)} : v_j, v_k \in N_i, e_{jk} \in E$$

The clustering coefficient for the whole system is the average of the clustering coefficients:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^{n} C_i$$
Centrality of a node

**Degree centrality**

\[
\frac{d_i}{n - 1}
\]

**Closeness centrality**

\[
\frac{n - 1}{\sum_{j \neq i} l(i, j)}
\]

where \( l(i, j) \) denotes the number of links in the shortest path between \( i \) and \( j \), or

\[
\sum_{j \neq i} \delta^{l(i,j)}
\]

with \( 0 < \delta < 1 \) and \( l(i, j) = \infty \) if \( i \) and \( j \) are not path-connected.
Centrality of a node

*Betweenness centrality*

\[
\sum_{k \neq j: i \notin \{k, j\}} \frac{P_j(kj)/P(kj)}{(n - 1)(n - 2)/2}
\]

where \( P_j(kj) \) denotes the number of geodesics (shortest paths) between \( k \) and \( j \), that \( i \) lies on, whereas \( P(kj) \) is the number of geodesics between \( k \) and \( j \).
Centrality of a node

Betweenness centrality – Example: fifteenth century Florence

\[ BC(\text{Medici}) = 0.522 \]
\[ BC(\text{Strozzi}) = 0.103 \]
\[ BC(\text{Guadagni}) = 0.255 \]
Spectrum

The spectrum of an undirected graph is the set of eigenvalues of its adjacency matrix

If \( p(N) = cN^{-z} \), with \( z < 1 \), the spectral density converges to the semicircle law (Wigner 1955)

Graph features

- \( N = 100, p = 0.1 \)
- \( N = 100, p = 0.02 \)
Motivation

A small-world network is a type of mathematical graph in which most nodes are not neighbors of one another, but they can be reached from every other by a small number of hops.

Purely random graphs exhibit a small average shortest path length (varying typically as the logarithm of the number of nodes) along with a small clustering coefficient.

However, many real-world networks have a small average shortest path length, but also a clustering coefficient significantly higher than expected by random chance.
Small-world networks

Watts and Strogatz model:

(i) a small average shortest path length,
(ii) a large clustering coefficient

- starting from a regular graph
- rewiring edges with equal and independent probability $p_r$

regular  small-world  (uncorrelated) random

$p_r = 0$  increasing randomness  $p_r = 1$
Small-world networks

Watts and Strogatz model

(i) a small average shortest path length,
(ii) a large clustering coefficient

for intermediate values of $p_r$:

**small-world behavior:**

- average clustering (C) high
- average distance (L) low
Scale-free networks

The distinguishing characteristic of scale-free networks is that their degree distribution follows a power law relationship defined by

\[ P(k) \sim k^{-\gamma} \]

In words, some nodes act as "highly connected hubs" (high degree), but most nodes have a low degree.

The scale-free model has a systematically shorter average path length than a random graph (thanks to the hub nodes).
Scale-free networks

Network building rules (dynamic)

1. The network begins with an initial network of $m_0 (>1)$ nodes

2. **Growth**: New nodes are added to the network one at a time

3. **Preferential attachment**: Each new node is connected to $m$ of the existing with a probability proportional to the number of links that the existing node already has. Formally, the probability $p_i$ that the new node is connected to node $i$ is

   $$p_i = \frac{k_i}{\sum_j k_j}$$

   where $k_i$ is the degree of node $i$ (rich gets richer)
Random geometric graphs

A random geometric graph is a random undirected graph drawn on a bounded region

It is generated by:

1. Placing vertices at random uniformly and independently on the region

2. Connecting two vertices, $u$, $v$ if and only if the distance between them is smaller than a threshold $r$

$$d(u, v) \leq r$$
Def.: A graph is said to be \( k\)-connected \((k=1,2,3,...)\) if for each node pair there exist at least \( k \) mutually independent paths connecting them.

Equivalently, a graph is \( k\)-connected if and only if no set of \((k-1)\) nodes exists whose removal would disconnect the graph.

The maximum value of \( k \) for which a connected graph is \( k\)-connected is the connectivity \( \kappa \) of \( G \). It is the smallest number of nodes whose failure would disconnect \( G \).

As \( r_0 \) increases, the resulting graph becomes \( k\)-connected at the moment it achieves a minimum degree \( d_{\text{min}} \) equal to \( k \).
Random geometric graphs

Thm (Gupta & Kumar): Graph $G(n, r_0)$, with

$$\pi r_n^2 = \frac{\log n + c_n}{n},$$

is connected with probability one as $n$ goes to infinity if and only if

$$c = \lim_{n \to \infty} c_n.$$
References
