Random graph theory


- fixed node number $N$
- connecting pairs of nodes with probability $p$

$p = 0$  $p = 0.1$  $p = 0.15$

Expected number of edges:  
\[ n = p \frac{N(N - 1)}{2} \]

Random graph theory studies the properties of graphs with $N \to \infty$
The properties of random graphs depend on $p$

Properties studied:
- does the graph contain cliques (complete subgraphs)?
- does the graph contain a large cluster?
- is the graph connected?
- what is the diameter of the graph?

Some of these properties appear suddenly, at a threshold $p_c(N)$

$$
\lim_{N \to \infty} P_{N,p}(Q) = \begin{cases} 
0 & \text{if } \frac{p(N)}{p_c(N)} \to 0 \\
1 & \text{if } \frac{p(N)}{p_c(N)} \to \infty
\end{cases}
$$
Evolution of a random graph

Assume that the connection probability is a power-law of N, \( p = cN^z \)
Assume that \( z \) increases from \( -\infty \) to 0
Look for trees, cycles (circuits) and cliques in the graph.

Appearance thresholds: \( p \sim N^z \)

<table>
<thead>
<tr>
<th>( z )</th>
<th>(-\infty)</th>
<th>(-2)</th>
<th>(-\frac{3}{2})</th>
<th>(-\frac{4}{3})</th>
<th>(-\frac{5}{4})</th>
<th>(-1)</th>
<th>(-\frac{2}{3})</th>
<th>(-\frac{1}{2})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
<td>( \cdot )</td>
</tr>
</tbody>
</table>

The graph contains cycles of any length if \( z \geq -1 \)
Clusters in a random graph

- For $p < N^{-1}$ the graph contains only isolated trees.
- If $p = cN^{-1}$ with $c < 1$ the graph has isolated trees and cycles.
- At $p = cN^{-1}$ with $c = 1$ a giant cluster appears.
- The size of the giant cluster approaches $N$ rapidly as $c$ increases.

\[ S = (f(1) - f(c))N \]

- The graph becomes connected if $p \geq \ln N / N$
Node degrees in random graphs

- average degree: \( \langle k \rangle = \frac{2n}{N} \approx pN \)

- degree distribution:

\[
P(k) \approx C_{N-1}^k p^k (1-p)^{N-1-k}
\]

Most of the nodes have approximately the same degree. The probability of very highly connected nodes is exponentially small.
Path lengths in random graphs

Random graphs tend to have a tree-like topology with almost constant node degrees.

- nr. of first neighbors: \( N_1 \cong \langle k \rangle \)
- nr. of second neighbors: \( N_2 \cong \langle k \rangle^2 \)
- estimate average path length:

\[
\sum_{i=0}^{l} \langle k \rangle^i = N \quad \Rightarrow \quad l = \frac{\ln N}{\ln \langle k \rangle}
\]

This scaling was proven by Chung and Lu, Adv. Appl. Math 26, 257 (2001).
There is no local order in random graphs

Measure of local order: \[ C_i \equiv \frac{n_i}{k_i(k_i - 1)/2} \]

Since edges are independent and have the same probability \( p \),

\[ n_i \approx p \frac{k_i(k_i - 1)}{2} \implies C \approx p \]

The clustering coefficient of random graphs is small.
Are real networks like random graphs?

As quantitative data about real networks becomes available, we can compare their topology with that of random graphs.
Starting measures: N, <k> for the real network.
Determine l, C and P(k) for a random graph with the same N and <k>.

\[ l_{\text{rand}} = \frac{\ln N}{\ln \langle k \rangle} \quad C_{\text{rand}} = p = \frac{\langle k \rangle}{N} \]

\[ P_{\text{rand}}(k) \approx C_{N-1}^k p^k (1 - p)^{N-1-k} \]

Measure l, C and P(k) for the real network. Compare.
Path length and order in real networks

\[ l_{\text{rand}} = \frac{\ln N}{\ln \langle k \rangle} \]

\[ C_{\text{rand}} = \frac{\langle k \rangle}{N} \]

Real networks have short distances like random graphs yet show signs of local order. New model?
Small-world networks

Real networks resemble both regular lattices and random graphs – perhaps they are in between.


- lattice with $K$ neighbors
- rewire edges with probability $p$

\[ l = \frac{N}{2K}, \quad C = \frac{3(K-2)}{4(K-1)} \]

Is there a regime with small $l$ and large $C$?
Transition from a lattice to a small world

There is a broad interval of $p$ over which $C(p) \approx C(0)$ but $l(p) \approx l(1)$
The onset of the small-world behavior depends on the system size

\[ l(N,p) \approx \frac{N^{1/d}}{K} f(pKN) \]

\[ f(u) = \begin{cases} 
\text{const} & \text{if } u << 1 \\
\ln u / u & \text{if } u >> 1 
\end{cases} \]

\[ C(p) = C(0)(1 - p)^3 \]

These results cannot be directly compared to most real networks because the rewiring probability \( p \) is not known.
Degree distribution of a small-world network

 rewiring does not change the average degree, but modifies the degree distribution.

\[ \langle k \rangle = K \]

\( P(k) \) depends on the rewiring parameter \( p \), but is always centered around \( <k> \).

Degree distribution similar to that of a random graph, with exponentially small probability for very highly connected nodes.
Ex. 1
A random graph has average degree $\langle k \rangle = 10$. How much is the average distance between nodes for $N=10^n$, where $n=2, 3, \ldots$?

Ex. 2
The same question for a ring lattice where every node has degree 10.

Ex. 3
A variant of the Watts-Strogatz model adds random edges to a regular lattice with probability $p$. Start with a ring lattice with 12 nodes. For each node, throw with a dice, and if the number is even, throw with two dice, and connect the original node with the node whose number you obtained. What is the degree distribution before and after? What is the largest distance from node 1 before and after?
The degree distribution of real networks is not peaked at all.

Nodes with small degrees are most frequent.
The fraction of highly connected nodes decreases fast.
The degree distribution of the WWW is a power-law

\[ P_{\text{out}}(k) = k^{-2.45} \]
\[ P_{\text{in}}(k) = k^{-2.1} \]


Power-law degree distributions were found in diverse networks.

\[ P(k) \approx k^{-2.4} \]

\[ P(k) \approx k^{-2.3} \]

Networks of science collaborations also have power-law degree distributions

\[ P(k) \approx k^{-1.2} \]

\[ P(k) \approx k^{-2.1} \]


A.-L. Barabási et al., cond-mat/0104162 (2001)
Even metabolic networks have a power-law degree distribution

\[ P(k) \approx k^{-2.2} \]


Archaeoglobus f.

E. coli

C. elegans
The scale-free degree distribution indicates a heterogeneous topology. New models are needed to reproduce this feature.