Class 4: Random Networks & Small World Networks
(Chapters 3-4 in Textbook)

Boleslaw Szymanski

based on slides by Albert-László Barabási & Roberta Sinatra
In the random graph literature, it is often assumed that the connection probability $p(N)$ scales as $N^z$, where $z$ is a tunable parameter between $-\infty$ and 0. In this language Erdős and Rényi discovered that as we vary $z$, key properties of random graphs appear quite suddenly.

A graph has a given property $Q$ if the probability of having $Q$ approaches 1 as $N \to \infty$. That is, for a given $z$ either almost every graph has the property $Q$ or almost no graph has it. For example, for $z$ less than $-3/2$ almost all graphs contain only isolated nodes and pairs of nodes connected by a link. Once $z$ exceeds $-3/2$, most networks will contain paths connecting three or more nodes, see image below.

\[ p = \frac{\langle k \rangle}{N-1} \approx \frac{\langle k \rangle}{N} = N^z \text{ so } \langle k \rangle = N^{z+1} \text{ and } z = \ln(\langle k \rangle) - 1 \]
Real networks are supercritical
Network Science: Random Networks

**Section 7**

### Bar Chart

- **Subcritical**
  - Internet
  - Power Grid
  - Science Collaboration
  - Actor Network
  - Yeast Protein Interactions

- **Supercritical**
  - Fully Connected

### Table: Network Statistics

<table>
<thead>
<tr>
<th>Network</th>
<th>$N$</th>
<th>$L$</th>
<th>$&lt;k&gt;$</th>
<th>$\ln N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet</td>
<td>192,244</td>
<td>609,066</td>
<td>6.34</td>
<td>12.17</td>
</tr>
<tr>
<td>Power Grid</td>
<td>4,941</td>
<td>6,594</td>
<td>2.67</td>
<td>8.51</td>
</tr>
<tr>
<td>Science Collaboration</td>
<td>23,133</td>
<td>186,936</td>
<td>8.08</td>
<td>10.04</td>
</tr>
<tr>
<td>Actor Network</td>
<td>212,250</td>
<td>3,054,278</td>
<td>28.78</td>
<td>12.27</td>
</tr>
<tr>
<td>Yeast Protein Interactions</td>
<td>2,018</td>
<td>2,930</td>
<td>2.90</td>
<td>7.61</td>
</tr>
</tbody>
</table>

Network Science: Random Networks
The measurements indicate that real networks extravagantly exceed the $\langle k \rangle = 1$ threshold. Sociologists estimate that an average person has around 1,000 acquaintances; a typical neuron is connected to dozens of other neurons, some to thousands; in our cells, each molecule takes part in several chemical reactions, some, like water, in hundreds.

The average degree of real networks is well beyond the $\langle k \rangle = 1$ threshold, implying they all have a giant component.

Do we have single component (if $\langle k \rangle > \ln N$), or multiple components (if $\langle k \rangle < \ln N$)? For social networks this requires $\langle k \rangle \geq \ln(7 \times 10^9) \approx 22.7$; so nearly two dozens acquaintances per person; with $\langle k \rangle \approx 1,000$ this is clearly satisfied. Most real networks do not satisfy this criteria, e.g., the Internet implying some routers are disconnected, so of little utility!

Most real networks are in the supercritical regime. This means that these networks have a giant component, but it coexists with many disconnected components and nodes, but only if real networks are accurately described by the Erdős-Rényi model, i.e. are random.

Today, we will further discuss the structure of real networks, we will understand why real networks can stay connected despite failing the $k > \ln N$ criteria.
Small worlds
SIX DEGREES

SMALL WORLDS

Frigyes Karinthy, 1929
Stanley Milgram, 1967

Network Science: Small World Networks
"Look, Selma Lagerlöf just won the Nobel Prize for Literature, thus she is bound to know King Gustav of Sweden, after all he is the one who handed her the Prize, as required by tradition. King Gustav, to be sure, is a passionate tennis player, who always participates in international tournaments. He is known to have played Mr. Kehrling, whom he must therefore know for sure, and as it happens I myself know Mr. Kehrling quite well."

"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Arpad Pasztor, someone I not only know, but to the best of my knowledge a good friend of mine. So I could easily ask him to send a telegram via the general director telling Ford that he should talk to the manager and have the worker in the shop quickly hammer together a car for me, as I happen to need one."
HOW TO TAKE PART IN THIS STUDY

1. ADD YOUR NAME TO THE ROSTER AT THE BOTTOM OF THIS SHEET, so that the next person who receives this letter will know who it came from.

2. DETACH ONE POSTCARD. FILL IT AND RETURN IT TO HARVARD UNIVERSITY. No stamp is needed. The postcard is very important. It allows us to keep track of the progress of the folder as it moves toward the target person.

3. IF YOU KNOW THE TARGET PERSON ON A PERSONAL BASIS, MAIL THIS FOLDER DIRECTLY TO HIM (HER). Do this only if you have previously met the target person and know each other on a first name basis.

4. IF YOU DO NOT KNOW THE TARGET PERSON ON A PERSONAL BASIS, DO NOT TRY TO CONTACT HIM DIRECTLY. INSTEAD, MAIL THIS FOLDER (POST CARDS AND ALL) TO A PERSONAL ACQUAINTANCE WHO IS MORE LIKELY THAN YOU TO KNOW THE TARGET PERSON. You may send the folder to a friend, relative or acquaintance, but it must be someone you know on a first name basis.
Network Science: Small World Networks

SIX DEGREES

1967: Stanley Milgram

![Diagram showing the number of chains vs. the number of intermediaries. The peak occurs at 6 intermediaries with N=64.]
"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice.... It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. How every person is a new door, opening up into other worlds."
WWW: 19 DEGREES OF SEPARATION

Image by Matthew Hurst
Blogosphere

Network Science: Small World Networks
Small World graphs tend to have a tree-like topology with almost constant node degrees.

\[ N = 1 + \langle k \rangle + \langle k \rangle^2 + ... + \langle k \rangle^{d_{\text{max}}} \approx \frac{\langle k \rangle^{d_{\text{max}}+1} - 1}{\langle k \rangle - 1} \approx \langle k \rangle^{d_{\text{max}}} \]

- \( \langle k \rangle \) nodes at distance one (\( d = 1 \)).
- \( \langle k \rangle^2 \) nodes at distance two (\( d = 2 \)).
- \( \langle k \rangle^3 \) nodes at distance three (\( d = 3 \)).
- ... \( \langle k \rangle^d \) nodes at distance \( d \).

\[ d_{\text{max}} = \frac{\log N}{\log \langle k \rangle} \]
We will call the small world phenomena the property that the average path length or the diameter depends logarithmically on the system size. Hence, “small” means that $\langle d \rangle$ is proportional to $\log N$, rather than $N$.

The $1/\log\langle k \rangle$ term implies that denser the network, the smaller will be the distance between the nodes.

$$d_{\text{max}} = \frac{\log N}{\log\langle k \rangle}$$

$$\langle d \rangle = \frac{\log N}{\log\langle k \rangle}$$

In most networks this offers a better approximation to the average distance between two randomly chosen nodes, $\langle d \rangle$, than to $d_{\text{max}}$. 

Network Science: Small World Networks
## DISTANCES IN Small Word Nets

### Comparison with real networks

<table>
<thead>
<tr>
<th>Network</th>
<th>N</th>
<th>L</th>
<th>$\langle k \rangle$</th>
<th>$\langle d \rangle$</th>
<th>$d_{\text{max}}$</th>
<th>$\ln N / \ln \langle k \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet</td>
<td>192,244</td>
<td>609,066</td>
<td>6.34</td>
<td>6.98</td>
<td>26</td>
<td>6.58</td>
</tr>
<tr>
<td>WWW</td>
<td>325,729</td>
<td>1,497,134</td>
<td>4.60</td>
<td>11.27</td>
<td>93</td>
<td>8.31</td>
</tr>
<tr>
<td>Power Grid</td>
<td>4,941</td>
<td>6,594</td>
<td>2.67</td>
<td>18.99</td>
<td>46</td>
<td>8.66</td>
</tr>
<tr>
<td>Mobile-Phone Calls</td>
<td>36,595</td>
<td>91,826</td>
<td>2.51</td>
<td>11.72</td>
<td>39</td>
<td>11.42</td>
</tr>
<tr>
<td>Email</td>
<td>57,194</td>
<td>103,731</td>
<td>1.81</td>
<td>5.88</td>
<td>18</td>
<td>18.4</td>
</tr>
<tr>
<td>Science Collaboration</td>
<td>23,133</td>
<td>93,437</td>
<td>8.08</td>
<td>5.35</td>
<td>15</td>
<td>4.81</td>
</tr>
<tr>
<td>Actor Network</td>
<td>702,388</td>
<td>29,397,908</td>
<td>83.71</td>
<td>3.91</td>
<td>14</td>
<td>3.04</td>
</tr>
<tr>
<td>Citation Network</td>
<td>449,673</td>
<td>4,707,958</td>
<td>10.43</td>
<td>11.21</td>
<td>42</td>
<td>5.55</td>
</tr>
<tr>
<td>E. Coli Metabolism</td>
<td>1,039</td>
<td>5,802</td>
<td>5.58</td>
<td>2.98</td>
<td>8</td>
<td>4.04</td>
</tr>
<tr>
<td>Protein Interactions</td>
<td>2,018</td>
<td>2,930</td>
<td>2.90</td>
<td>5.61</td>
<td>14</td>
<td>7.14</td>
</tr>
</tbody>
</table>
Introduction to Scale Free Networks
Why are small worlds surprising? Suprising compared to what?

1d lattice: $\langle d \rangle \sim N$
2d lattice: $\langle d \rangle \sim N^{1/2}$
3d lattice: $\langle d \rangle \sim N^{1/3}$
Random Network: $\langle d \rangle \sim \log N$
Three, Four or Six Degrees?

For the globe’s social networks:

\[ \langle k \rangle \approx 10^3 \]

\[ N \approx 7 \times 10^9 \] for the world’s population.

\[ \langle d \rangle = \frac{\ln(N)}{\ln\langle k \rangle} = 3.28 \]
“The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but to the best of my knowledge a good friend of mine.”

Karinthy, 1929

“Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice. It’s not just the big names. It’s anyone. A native in a rain forest. A Tierra del Fuego. An Eskimo. I am bound to everyone on this planet by a trail of six people. It’s a profound thought. How every person is a new door, opening up into other worlds.”

Guare, 1991

Frigyes Karinthy [1887-1938]
Hungarian writer, journalist and playwright, the first to describe the small world property. In his short story entitled 'Láncczemek' [Chains] he links a worker in Ford’s factory to himself [23, 24].

Manfred Kochen [1928-1989], Ihthiel de Sola Pool [1917-1984]
Scientific interest in small worlds started with a paper by political scientist Ihthiel de Sola Pool and mathematician Manfred Kochen. Written in 1958 and published in 1978, their work addressed in mathematical detail the small world effect, predicting that most individuals can be connected via two to three acquaintances. Their paper inspired the experiments of Stanley Milgram.

Stanley Milgram [1933-1984]
American social psychologist who carried out the first experiment testing the small-world phenomena. (BOX 3.6)

John Guare [1938]
The phrase ‘six degrees of separation’ was introduced by the playwright John Guare, who used it as the title of his Broadway play.

The Facebook Data Team measures the average distance between its users, finding “4 degrees” (BOX 3.6).

Duncan J. Watts [1971].
Steven Strogatz [1959].
A new wave of interest in small worlds followed the study of Watts and Strogatz, finding that the small world property applies to natural and technological networks as well.
Clustering coefficient
Since edges are independent and have the same probability $p$, 

$$<L_i> = p \frac{k_i(k_i - 1)}{2} \implies C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$ 

- The clustering coefficient of random graphs is small.
- For fixed degree $C$ decreases with the system size $N$.
- $C$ is independent of a node’s degree $k$. 

Network Science: Small World Networks
C decreases with the system size $N$.

C is independent of a node’s degree $k$. 

$$C_i = \frac{2\langle L_i \rangle}{k_i (k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$
Real networks are not random
As quantitative data about real networks became available, we can compare their topology with the predictions of random graph theory.

Note that once we have $N$ and $\langle k \rangle$ for a random network, from it we can derive every measurable property. Indeed, we have:

**Average path length:**

$$\langle l_{\text{rand}} \rangle \approx \frac{\log N}{\log \langle k \rangle}$$

**Clustering Coefficient:**

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i-1)} = p = \frac{\langle k \rangle}{N}.$$  

**Degree Distribution:**

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$
Prediction:

\[<d> = \frac{\log N}{\log \langle k \rangle}\]

Real networks have short distances like random graphs.
Prediction:

\[ C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}. \]

\( C_{rand} \) underestimates with orders of magnitudes the clustering coefficient of real networks.
THE DEGREE DISTRIBUTION

Prediction:

\[ P(k) = e^{-\langle k \rangle} \frac{(\langle k \rangle)^k}{k!} \]

Data:

\[ P(k) \approx k^{-\gamma} \]
As quantitative data about real networks became available, we can compare their topology with the predictions of random graph theory.

Note that once we have $N$ and $<k>$ for a random network, from it we can derive every measurable property. Indeed, we have:

**Average path length:**

$$<l_{\text{rand}}> \approx \frac{\log N}{\log <k>}$$

**Clustering Coefficient:**

$$C_i = \frac{2\langle L_i \rangle}{k_i (k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$  

**Degree Distribution:**

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$
(B) Most important: we need to ask ourselves, are real networks random?

The answer is simply: NO

There is no network in nature that we know of that would be described by the random network model.
It is the reference model for the rest of the class.

It will help us calculate many quantities, that can then be compared to the real data, understanding to what degree is a particular property the result of some random process.

Patterns in real networks that are shared by a large number of real networks, yet which deviate from the predictions of the random network model.

In order to identify these, we need to understand how would a particular property look like if it is driven entirely by random processes.

While WRONG and IRRELEVANT, it will turn out to be extremly USEFUL!
Summary
Erdös-Rényi MODEL (1960)
1951, Rapoport and Solomonoff:
→ first systematic study of a random graph.
→ demonstrates the phase transition.
→ natural systems: neural networks; the social networks of physical contacts (epidemics); genetics.

1959: $G(N,p)$

Why do we call it the Erdos-Renyi random model?
Erdos:
1,400 papers
507 coauthors

Einstein: EN=2
Paul Samuelson EN=5

ALB: EN=3
Collaboration Network:

**Nodes:** Scientists

**Links:** Joint publications

**Physical Review:** 1893 – 2009.

**N=449,673**

**L=4,707,958**

See also Stanford Large Network database

Scale-free Hierarchical Network Science: Small World Networks

Science Collaboration

Scale-free

Hierarchical
WORLD WIDE WEB

Power laws and scale-free networks
Nodes: WWW documents
Links: URL links

Over 3 billion documents

ROBOT: collects all URL’s found in a document and follows them recursively

**Discrete Formalism**
As node degrees are always positive integers, the discrete formalism captures the probability that a node has exactly $k$ links:

$$p_k = Ck^{-\gamma}.$$  
$$\sum_{k=1}^{\infty} p_k = 1.$$  
$$C \sum_{k=1}^{\infty} k^{-\gamma} = 1$$  
$$C = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)},$$  
$$P_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

**Continuum Formalism**
In analytical calculations it is often convenient to assume that the degrees can take up any positive real value:

$$p(k) = Ck^{-\gamma}.$$  
$$\int_{k_{\text{min}}}^{\infty} p(k)dk = 1$$  
$$C = \frac{1}{\int_{k_{\text{min}}}^{\infty} k^{-\gamma}dk} = (\gamma - 1)k_{\text{min}}^{\gamma-1}$$  
$$p(k) = (\gamma - 1)k_{\text{min}}^{\gamma-1}k^{-\gamma}.$$  
$$\int_{k_{\text{min}}}^{k_2} p(k)dk$$

**INTERPRETATION:**
Vilfredo Federico Damaso Pareto (1848 – 1923), Italian economist, political scientist and philosopher, who had important contributions to our understanding of income distribution and to the analysis of individuals choices. A number of fundamental principles are named after him, like Pareto efficiency, Pareto distribution (another name for a power-law distribution), the Pareto principle (or 80/20 law).
The difference between a power law and an exponential distribution
Let us use the WWW to illustrate the properties of the high-\(k\) regime. The probability to have a node with \(k \approx 100\) is

• About \(p_{100} \approx 10^{-30}\) in a Poisson distribution

• About \(p_{100} \approx 10^{-4}\) if \(p_k\) follows a power law.

• Consequently, if the WWW were to be a random network, according to the Poisson prediction we would expect \(10^{-18}\) \(k > 100\) degree nodes, or none.

• For a power law degree distribution, we expect about \(N_{k > 100} = 10^9\) \(k > 100\) degree nodes
All real networks are finite → let us explore its consequences.
→ We have an expected maximum degree, $k_{\text{max}}$

**Estimating $k_{\text{max}}$**

$$\int_{k_{\text{max}}}^{\infty} P(k) \, dk \approx \frac{1}{N}$$

Why: the probability to have a node larger than $k_{\text{max}}$ should not exceed the prob. to have one node, i.e. $1/N$ fraction of all nodes

$$\int_{k_{\text{max}}}^{\infty} P(k) \, dk = (\gamma - 1) k_{\text{min}}^{\gamma - 1} \int_{k_{\text{max}}}^{\infty} k^{-\gamma} \, dk = \frac{(\gamma - 1)}{(-\gamma + 1)} k_{\text{min}}^{\gamma - 1} \left[ k^{-\gamma + 1} \right]_{k_{\text{max}}}^{\infty} = \frac{k_{\text{min}}^{\gamma - 1}}{k_{\text{max}}^{\gamma - 1}} \approx \frac{1}{N}$$

$$k_{\text{max}} = k_{\text{min}} N^{\gamma - 1}$$
The size of the biggest hub

\[ k_{\text{max}} = k_{\text{min}} N^{\gamma^{-1}} \]

To illustrate the difference in the maximum degree of an exponential and a scale-free network let us return to the WWW sample of Figure 4.1, consisting of \( N \approx 3 \times 10^5 \) nodes. As \( k_{\text{min}} = 1 \), if the degree distribution were to follow an exponential, (4.17) predicts that the maximum degree should be \( k_{\text{max}} \approx 13 \). In a scale-free network of similar size and \( \gamma = 2.1 \), (4.18) predicts \( k_{\text{max}} \approx 85,000 \), a remarkable difference. Note that the largest in-degree of the WWW map of Figure 4.1 is 10,721, which is comparable to \( k_{\text{max}} \) predicted by a scale-free network. This reinforces our conclusion that in a random network hubs are effectively forbidden, while in scale-free networks they are naturally present.
Finite scale-free networks

Expected maximum degree, $k_{\text{max}}$

$$k_{\text{max}} = k_{\text{min}} \frac{1}{N^{\gamma-1}}$$

- $k_{\text{max}}$, increases with the size of the network
  - The larger a system is, the larger its biggest hub
- For $\gamma > 2$, $k_{\text{max}}$ increases slower than $N$
  - The largest hub will contain a decreasing fraction of links as $N$ increases.
- For $\gamma = 2$, $k_{\text{max}} \sim N$.
  - The size of the biggest hub is $O(N)$
- For $\gamma < 2$, $k_{\text{max}}$ increases faster than $N$: condensation phenomena
  - The largest hub will grab an increasing fraction of links. Anomaly!
The size of the largest hub

\[ k_{\text{max}} = k_{\text{min}} \left( \frac{N}{k_{\text{max}}} \right)^{\gamma - 1} \]

\[ k_{\text{max}} \sim N^{\frac{1}{\gamma - 1}} \]

\[ (N - 1) \]

SCALE-FREE

RANDOM NETWORK

Network Science: Scale-Free Networks