Network models

Empirical network features:
- Power-law (heavy-tailed) degree distribution
- Small average distance (graph diameter)
- Large clustering coefficient (transitivity)
- Giant connected component, hierarchical structure, etc

Generative models:
- Random graph model (Erdos & Renyi, 1959)
- Preferential attachment model (Barabasi & Albert, 1999)
- Small world model (Watts & Strogatz, 1998)
Most of the networks we study are evolving over time, they expand by adding new nodes:

- Citation networks
- Collaboration networks
- Web
- Social networks
Preferential attachment model

Barabasi and Albert, 1999
Dynamic growth model
Start at \( t = 0 \) with \( n_0 \) nodes and some edges \( m_0 \geq n_0 \)

1. **Growth**
   At each time step add a new node with \( m \) edges \( (m \leq n_0) \), connecting to \( m \) nodes already in network \( k_i(i) = m \)

2. **Preferential attachment**
   The probability of linking to existing node \( i \) is proportional to the node degree \( k_i \)
   \[
   \Pi(k_i) = \frac{k_i}{\sum_i k_i}
   \]

after \( t \) timesteps: \( t + n_0 \) nodes, \( mt + m_0 \) edges
Preferential attachment model

Scale-Free Model

Barabasi, 1999
Preferential attachment

Continues approximation: continues time, real variable node degree
\( \langle k_i(t) \rangle \) - expected value over multiple realizations

Time-dependent degree of a single node:

\[
k_i(t + \delta t) = k_i(t) + m \Pi(k_i) \delta t
\]

Initial conditions:

\[ k_i(t = i) = m \]

Solution:

\[ k_i(t) = m \left( \frac{t}{i} \right)^{1/2} \]
Node degree $k$ as function of time $t$

$$k_i(t) = m \left( \frac{t}{i} \right)^{1/2}$$
Preferential attachment
Preferential attachment

Time evolution of a node degree

\[ k_i(t) = m \left( \frac{t}{i} \right)^{1/2} \]

Nodes with \( k_i(t) \leq k \):

\[ m \left( \frac{t}{i} \right)^{1/2} \leq k \]
\[ i \geq \frac{m^2}{k^2} t \]

Probability of randomly selected node to have \( k' \leq k \) (fraction of nodes with \( k' \leq k \))

\[ F(k) = P(k' \leq k) = \frac{n_0 + t - m^2 t/k^2}{n_0 + t} \approx 1 - \frac{m^2}{k^2} \]

Distribution function:

\[ P(k) = \frac{d}{dk} F(k) = \frac{2m^2}{k^3} \]
Preferential attachment vs random graph

BA: \( P(k) = \frac{2m^2}{k^3} \), \quad ER: \ P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad \langle k \rangle = pn
Preferential attachment vs random graph

**Node degree distribution**

BA: \( P(k) = \frac{2m^2}{k^3} \),

ER: \( P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!} \), \( \langle k \rangle = pn \)
Preferential attachment vs random graph
Preferential attachment model

\[ m = 1 \quad m = 2 \quad m = 3 \]
Growing random graph

1. **Growth**
   
   At each time step add a new node with $m$ edges ($m \leq n_0$), connecting to $m$ nodes already in network $k_i(i) = m$

2. **Preferential attachment Uniformly at random**

   The probability of linking to existing node $i$ is
   
   $$\Pi(k_i) = \frac{1}{n_0 + t - 1}$$

   Node degree growth:
   
   $$k_i(t) = m \left(1 + \log \left(\frac{t}{i}\right)\right)$$

   Node degree distribution function:
   
   $$P(k) = \frac{e}{m} \exp \left(-\frac{k}{m}\right)$$
Preferential attachment

- Power law distribution function:
  \[ P(k) = \frac{2m^2}{k^3} \]

- Average path length (analytical result):
  \[ \langle L \rangle \sim \frac{\log(N)}{\log(\log(N))} \]

- Clustering coefficient (numerical result):
  \[ C \sim N^{-0.75} \]
Many more models

Some other models that produce scale-free networks:

- Non-linear preferential attachment
- Link selection model
- Copying model
- Cost-optimization model
- ...

Leonid E. Zhukov (HSE)
Historical note

- Polya urn model, George Polya, 1923
- Yule process, Udny Yule, 1925
- Distribution of wealth, Herbert Simon, 1955
- Evolution of citation networks, cumulative advantage, Derek de Solla Price, 1976
- Preferential attachment network model, Barabasi and Albert, 1999
Motivation: keep high clustering, get small diameter

Clustering coefficient $C = 1/2$
Graph diameter $d = 8$
Watts and Strogatz, 1998

Single parameter model, interpolation between regular lattice and random graph

- start with regular lattice with \( n \) nodes, \( k \) edges per vertex (node degree), \( k << n \)
- randomly connect with other nodes with probability \( p \), forms \( p nk/2 \) "long distance" connections from total of \( nk/2 \) edges
- \( p = 0 \) regular lattice, \( p = 1 \) random graph
Small world

Watts, 1998
Small world model

- Node degree distribution: Poisson like
- Ave. path length $\langle L(p) \rangle$:
  - $p \to 0$, ring lattice, $\langle L(0) \rangle = 2n/k$
  - $p \to 1$, random graph, $\langle L(1) \rangle = \log(n)/\log(k)$
- Clustering coefficient $C(p)$:
  - $p \to 0$, ring lattice, $C(0) = 3/4 = \text{const}$
  - $p \to 1$, random graph, $C(1) = k/n$

Watts, 1998
Small world model

20% rewiring:
ave. path length = 3.58  →  ave. path length = 2.32
clust. coeff = 0.49      →  clust. coeff = 0.19
## Model comparison

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>BA model</th>
<th>WS model</th>
<th>Empirical networks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P(k)</strong></td>
<td>$\frac{\lambda^k e^{-\lambda}}{k!}$</td>
<td>$k^{-3}$</td>
<td>poisson like</td>
<td>power law</td>
</tr>
<tr>
<td></td>
<td>$\langle k \rangle / N$</td>
<td>$N^{-0.75}$</td>
<td>const</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>$\log(N)$</td>
<td>$\log(N)$</td>
<td>$\log(N)$</td>
<td>small</td>
</tr>
<tr>
<td></td>
<td>$\log(\langle k \rangle)$</td>
<td>$\log \log(N)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>\langle L \rangle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Leonid E. Zhukov (HSE)  
Lecture 4  
06.02.2016  
23 / 23