Network Analysis
Statistical Approaches and Modeling

Basic Models

Vladimir Batagelj

NRU HSE Moscow, IMFM Ljubljana and IAM UP Koper

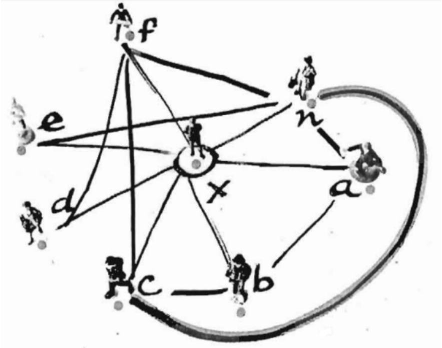
Master's programme

Applied Statistics with Social Network Analysis

International Laboratory for Applied Network Research

NRU HSE, Moscow 2019

Outline



group inbreeding in small world

Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Current version of slides (May 16, 2019 at 12 : 40): slides PDF

Models

Approaches to network analysis can be classified in different ways. One of them is:

- **Analysis of a given network:** general properties (size, type, components, distributions, . . .), important elements (nodes or links) and subnetworks, position of selected elements in a network, etc.
- **Analysis of families of networks:** derivation/explanation of general properties of networks from a family, position of a given network (unusual or anomalous property value) in the family, role of an element/subnetwork with respect to the family, etc. Creation/evolution of networks from the family.

Emergent properties in complex systems: Netlogo: Earth Science/Fire – phase transition.

Random network models

We want to have formal processes which can give rise to networks with specific properties (degree distribution, transitivity, diameter etc.). These models and their features can help us understand how the properties of a network (network structure) arise.

Intuitively we can think about a model in which pairs of nodes are connected with some probability. That is, if we start with a collection of n nodes and for each of the $n(n–1)/2$ possible links, we connect a pair of nodes u, v with certain probability $p_{u,v}$. Then, if we consider a set of network parameters to be fixed and allow the links to be created by a random process, we can create models that permit us to understand the influence of these parameters on the structure of networks.

Random network models

Let G denote a graph and let $p(G)$ be a probability distribution over all such graphs. The typical or expected value of some network measure x is then given by

$$\mathrm{E}(x) = \sum_G x(G) \cdot p(G)$$

where $x(G)$ is the value of the measure x on a particular graph G. This equation has the usual form of an average, but is calculated by summing over the combinatoric space of graphs. If some observed value is very different from the value expected from the model, then we may conclude that the true generating process for the data is different (more interesting) than the simple random process we assumed. This approach to classifying properties as interesting or not treats the random graph as a null model, which is a classic approach in the statistics.

Random network models

For example, the diameter of class $\mathbf{G}$, would be the diameter of a graph $G \in \mathbf{G}$, averaged over the class

$$\overline{\text{diam}} = \sum_{G \in \mathbf{G}} \text{diam}(G) p(G)$$

This approach is in general convenient:

- Often allows analytical calculation
- We can see the typical properties of the network model we consider
- The distribution of many network metrics, at the limit of large n, is sharply peaked around the mean value. Hence in the limit of large n we expect to see behaviors very close to the mean of the class.

Erdös–Rényi model

In this model, $\mathbf{G}_{ER}(n, p)$, we start with n isolated nodes. We then pick a pair of nodes and with a fixed probability p we add a link between them. For each pair of nodes we generate a random number, r, uniformly from $[0, 1]$ and if $p > r$ we add a link between them. Consequently, if we select $p = 0$ the network will remain fully disconnected forever and if $p = 1$ we end up with a complete graph.

This model was proposed by Gilbert in 1959 [8]. Erdös and Rényi (1959) [6] proposed a slightly different model, $\mathbf{G}_{ER}(n, m)$, in which a uniformly distributed random graph with n nodes and exactly m links is obtained. It turns out that graphs obtained on either way have almost the same properties [4].

Pajek: `Network/Create random network/Total no of arcs` or `Network/Create random network/Bernoulli`
R: sna: `rgraph`, `rgnm`
`igraph::erdos.renyi.game`

Erdös–Rényi model

1. The expected number of edges is $\overline{m} = \frac{1}{2}n(n\text{–}1)p$
2. The expected node degree is $\overline{d} = (n\text{–}1)p$
3. The degrees follow a binomial distribution

$$p(d) = \binom{n-1}{d} p^d (1-p)^{n-1-d}$$

approximated with Poisson distribution $p(d) = \frac{1}{d!}e^{-\overline{d}}\overline{d}^d$ or normal distribution $N(mean(d), sd(d))$. R

4. The probability of drawing at random a graph with m edges from the $\mathbf{G}(n,p)$ is:

$$p(m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}$$

A formal derivation of 1: $\overline{m} = \sum_{i=1}^{\binom{n}{2}} m \cdot p(m)$.

Erdös–Rényi model

5. In a random graph the probability that **any** two vertices are linked is equal to p.
6. The average clustering coefficient is $\overline{cl}(G) = p$. For constat $\overline{d}$ therefore $\overline{cl}(G) = \frac{\overline{d}}{n-1}$ tends to 0 with growing n. Not true for most real-life networks.
7. The average path length for large n is

$$\bar{l}(G) = \frac{\ln n - \gamma}{\ln(pn)} + \frac{1}{2} \approx \frac{\ln n}{\ln \overline{d}}$$

where $\gamma \approx 0.577$ is the Euler–Mascheroni constant.

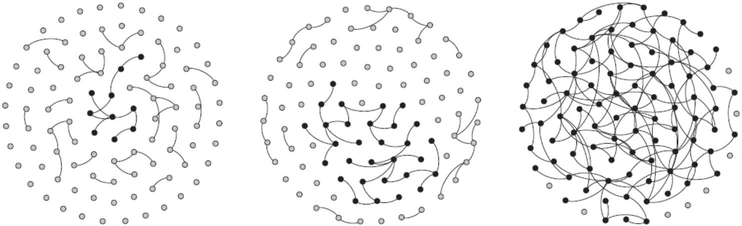
8. As p increases, most nodes tend to be clustered in one giant component, while the rest of nodes are isolated in very small components. Netlogo ER Phase transition!

Erdös–Rényi model

9. The structure of $G_{ER}(n, p)$ changes as a function of $\overline{d} = p \cdot (n–1)$, giving rise to the following three stages.
   a. Subcritical $\overline{d} < 1$, where all components are simple and very small. The size of the largest component is $S = O(\ln n)$.
   b. Critical $\overline{d} = 1$, where the size of the largest component is $S = O(n^{2/3})$.
   c. Supercritical $\overline{d} > 1$, where the probability that $(f–\varepsilon)n < S < (f + \varepsilon)n$ is 1 when $n \to \infty$, $\varepsilon > 0$, and where $f = f(\overline{d})$ is the positive solution of the equation $f = 1–e^{-\overline{d}f}$. The rest of the components are very small, with the second largest having size about $\ln n$. Small components are almost acyclic – mostly trees.

Erdös–Rényi model

emmergence of the giant component



(a) $p = 0.0075, S = 8$ (b) $p = 0.01, S = 21$ (c) $p = 0.025, S = 91$

S is the size of the largest component.

```
> n <- 100; p <- c(0.0075, 0.01, 0.025)
> (da <- p*(n-1))
[1] 0.7425 0.9900 2.4750
```

Erdös–Rényi model

10 The largest eigenvalue of the adjacency matrix A in an ER network grows proportionally to n so that

$$\lim_{n\to\infty} \frac{\lambda_1(A)}{n} = p$$

11 The second largest eigenvalue grows more slowly than λ_1. In fact,

$$\lim_{n\to\infty} \frac{\lambda_2(A)}{n^\varepsilon} = 0$$

for every $\varepsilon > 0.5$

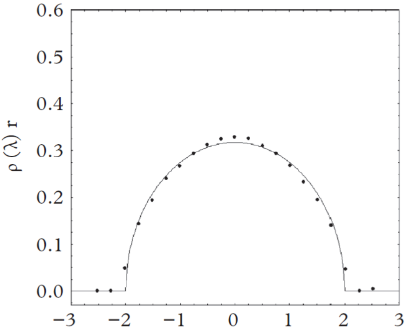
12 The most negative eigenvalue grows in a similar way to $\lambda_2(A)$. Namely,

$$\lim_{n\to\infty} \frac{\lambda_n(A)}{n^\varepsilon} = 0$$

for every $\varepsilon > 0.5$

Erdös–Rényi model

spectral density



The spectral density of an ER random network follows Wigner's semicircle law. That is, almost all of the eigenvalues of an ER random network lie in the range $[-2r, 2r]$ where $r = \sqrt{np(1-p)}$ and within this range the density function is given by

$$\rho(\lambda) = \frac{\sqrt{4-\lambda^2}}{2\pi}$$

Configuration model

Configuration model is a generalization of the Erdös–Rényi model $\mathbf{G}(n, m)$. It defines a class of random graphs, $\mathbf{G}(n, \mathbf{d})$, with fixed degree sequence $\mathbf{d} = (d_1, d_2, d_3, \ldots, d_n)$. Note that the number of edges is also fixed $m = \frac{1}{2} \sum d_i$.

Not all sequences of non-negative integers are degree sequences: Erdős–Gallai theorem
A sequence of non-negative integers ($d_1 \geq \cdots \geq d_n$) can be represented as the degree sequence of a finite simple (undirected) graph on n nodes if and only if $\sum d_i$ is even and

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(d_i, k)$$

holds for every k in $1 \leq k \leq n$.

Configuration model

For simple directed graphs the answer is given by Fulkerson–Chen–Anstee theorem.

A sequence $((a_1, b_1), \ldots, (a_n, b_n))$ of nonnegative integer pairs with $a_1 \geq \cdots \geq a_n$ is digraphic if and only if $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$ and the following inequality holds for k such that $1 \leq k \leq n$:

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} \min(b_i, k-1) + \sum_{i=k+1}^{n} \min(b_i, k)$$

When the degrees are determined from an example graph we know that at least one such graph exists.

Configuration model

The configuration model can serve as a null model for investigating the structure of a real network H. That is, it allows us to quantitatively answer the question of

How much of some observed pattern is driven by the degrees alone?

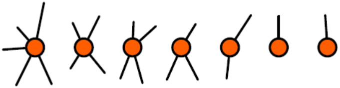
The configuration model defines a probability distribution over graphs $p(G|\mathbf{d})$ that has the same degrees as the original network H. Thus, if we can compute a function f on H, we can compute the same function on a graph drawn from this configuration model $f(G)$. And, because G is a random variable, we can compute the entire distribution $p(f(G)|\mathbf{d})$. For simple functions and simple specifications of the configuration model, we can often compute these distributions analytically.

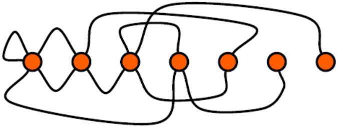
For more complicated functions or for a configuration model specified with an empirical degree sequence, we can compute $p(f(G)|\mathbf{d})$ numerically, by drawing many graphs $\{G_1, G_2, \ldots\}$ from the model, computing f on each, and tabulating the results. If the empirical value $f(H)$ is unusual relative to this distribution, we can conclude that it is a property of H that is not well explained by the degrees alone.

Configuration model

generating, Molloy-Reed method

Molloy-Reed method: Randomly (uniformly) link the semilinks in stubs:





1 1 1 1 1 2 2 2 2 3 3 3 3 4 4 4 5 5 6 7

1 4 1 2 2 3 2 5 1 2 3 7 3 4 3 5 1 1 4 6

Loops and parallel edges might appear. They distroy the uniformity of graphs distribution. But, since their density tends to 0 with increasing n, the variations in their probabilities are expected to be small.

1. **(Expected) probability of an edge between nodes u and v :** There are $\deg(u)$ stubs at node u and$\deg(v)$ at v. The probability that one of the $\deg(u)$ stubs of node u connects with one of the stubs of node v is $\deg(v)/(2m-1)$. Since there are $\deg(u)$ possible stubs for node u the overall probability is:

$$p(u,v) = \frac{\deg(u)\deg(v)}{2m-1} \approx \frac{\deg(u)\deg(v)}{2m}$$

Configuration model

generating, Chung-Lu method

If we do not insist on exact degree sequence but on expected degree distribution (analogous to $\mathbf{G}(n,p)$) we can use the *Chung-Lu* method: Select each link (u,v) with a probability:
undirected network:

$$p(u,v) = \deg(u)\frac{\deg(v)}{\sum_t \deg(t)} = \frac{\deg(u)\cdot\deg(v)}{2m}$$

directed network:

$$p(u,v) = \text{outdeg}(u)\frac{\text{indeg}(v)}{\sum_t \text{indeg}(t)} = \frac{\text{outdeg}(u)\cdot\text{indeg}(v)}{m}$$

R

igraph: `degree.sequence.game`

Configuration model

2. **Expected number of parallel edges in the network:** $\frac{1}{2}(\frac{\overline{d^2} - \overline{d}}{\overline{d}})^2$
3. **Expected number of loops in the network:** $\frac{\overline{d^2} - \overline{d}}{2\overline{d}}$
4. **Expected number n_{uv} of common neighbors between nodes u and v:**

$$n_{uv} = p_{uv}\frac{\overline{d^2} - \overline{d}}{\overline{d}}$$

Configuration model

The friendship paradox

5. **Neighbor's degree distribution:** there is a $d/(2m-1) \approx d/(2m)$ probability the edge we follow to end to a specific node of degree d. The total number of nodes with degree d is $np(d)$. Hence the probability that a neighbor of a node has degree d is:

$$\frac{d}{2m} np(d) = \frac{dp(d)}{\overline{d}} \quad \text{since} \quad 2m = n\overline{d}$$

6. **Average degree of a neighbor:**

$$\sum_d d \frac{dp(d)}{\overline{d}} = \frac{\overline{d^2}}{\overline{d}}$$

Therefore

$$\frac{\overline{d^2}}{\overline{d}} - \overline{d} = \frac{1}{\overline{d}}(\overline{d^2} - \overline{d}^2) = \frac{\sigma(d)^2}{\overline{d}} \geq 0$$

Friendship paradox: Your friends have more friends than you!

Configuration model

7. Expected value of clustering coefficient $\mathrm{E}(cl) = \frac{1}{n}\frac{(\overline{d^2} - \overline{d})^2}{\overline{d}^3}$
8. Average number of k-hop neighbors c_k:

$$c_k = c_{k-1}\frac{c_2}{c_1} = (\frac{c_2}{c_1})^{k-1} c_1, \quad k = 1, 2, \ldots$$

and $c_1 = \overline{d}$ and $c_2 = \overline{d^2} - \overline{d}$.
Giant component iff $c_2 > c_1$; or equivalently $\overline{d^2} > 2\overline{d}$.

Configuration model

The configuration model is an improvement over the simple random graph model in that it allows us to specify its degree structure. As a null model, this property is often suficient for us to use the model to decide whether some other property of a network could be explained by its degree structure alone.

Small worlds

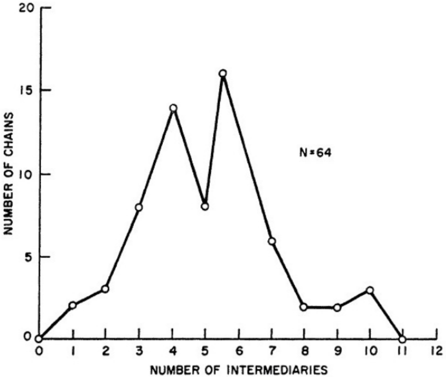
F. Karinthy, probably inspired by Marconi, in the 1920s observed a 'shrinking' modern world due to increased human connectedness' – strangers being linked by a short chain of acquaintances. He proposed a challenge to find another person to whom he could not be connected through at most five people.

In 1967 Milgram made a " small-world" experiment. A **friend** is someone known on a first-name basis. He sent 296 letters to people in Wichita, KS and Omaha, NE. Letters indicated a (unique) **contact** person in Boston, MA. He asked them to forward the letter to the contact, following rules:

Rule 1: If contact is a friend then send her/him the letter; else
Rule 2: Relay to friend most-likely to be a contact's friend.

S. Milgram, The small-world problem," Psychology Today, vol. 2, pp. 60-67, 1967.

Milgram's experiment



64 of 296 letter reached the destination, average path length $l = 6.2$ – *six degrees of separation*. There was a large group inbreeding, which resulted in acquaintances of one individual feeding a letter back into his/her own circle, thus usually eliminating new contacts.

The six degrees of separation were popularized by a play of Guare in 1990.

Small worlds

A *small-world network* is defined to be a network where the typical distance L between two randomly chosen nodes grows proportionally to the logarithm of the number of nodes n in the network:

$$L \propto \log n$$

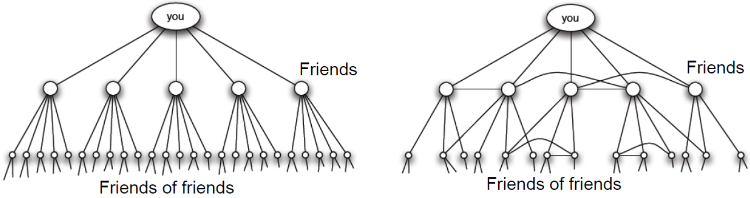
while the clustering coefficient is not small.

Examples of small world networks are: Internet, Wikipedia, gene networks.

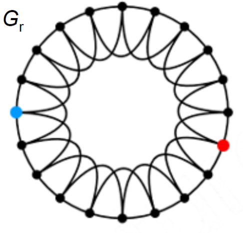
Small worlds

Is the small-world model reasonable?

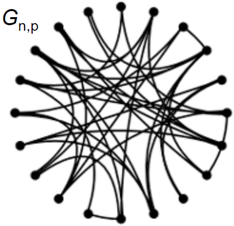
We have 100 friends, each of them has 100 other friends, ... After 5 degrees we get 10^{10} friends $>$ the Earth's population ($7.6\,10^9$ in 2018). There should be many cross-links forming *shortcuts*.



The Watts-Strogatz model



High clustering and diameter



Low clustering and diameter

A regular circulant graph G_r on n nodes: each node is linked to its $2r$ closest neighbors (r to each side).

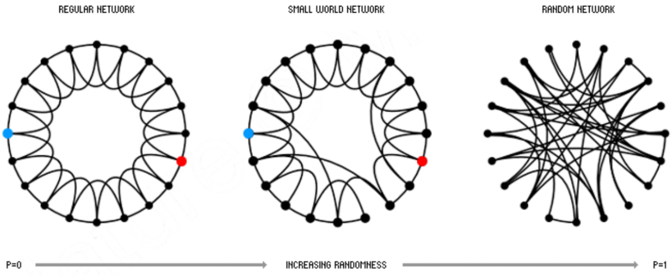
G_r (**structure**) yields high clustering and high diameter: $cl(G_r) = (3r - 3)/(4r - 2)$ and $diam(G_r) = n/(2r)$.

A random graph $ER(n, p)$ with $p = 2r/(n - 1) = O(1/n)$ (**randomness**) yields low clustering and low diameter: $cl(ER(n, p)) = O(1/n)$ and $diam(ER(n, p)) = O(\log n)$.

The Watts-Strogatz model

Watts-Strogatz small-world model: blend of structure with little randomness:

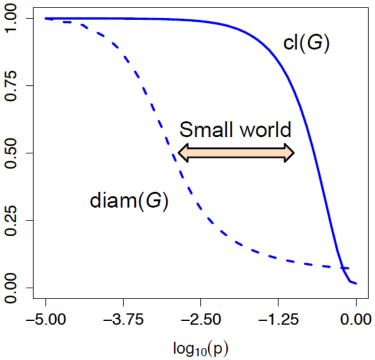
Start with regular lattice that has desired clustering.
Introduce randomness to generate shortcuts in the graph – each edge is randomly rewired (one of its end points moved to a new randomly chosen node) with (small) probability *p*.



Netlogo Small world

The Watts-Strogatz model

Simulation of Watts-Strogatz model with $n = 1000$ and $r = 6$:



Broad range of $p \in [10^{-3}, 10^{-1}]$ yields small $diam(G)$ and high $cl(G)$.

Small worlds

1. For large n it holds $cl(G) \approx cl(G_r)(1 - p^3)$.
2. degree distribution concentrated around $2r$
3. The average path length decays very fast from that of a circulant graph to approach that of a random network.

Is my network a small world network?

Small worlds

Small worlds are a combination of two basic social-network ideas:

- **Homophily:** the principle that we connect to others who are like ourselves, and hence creates many triangles.
- **Weak ties:** the links to acquaintances that connect us to parts of the network that would otherwise be far away, and hence the kind of widely branching structure that reaches many nodes in a few steps

Small-world graph models are particularly relevant to 'communication' in a broad sense:

- spread of news, gossip, rumors;
- spread of natural diseases and epidemics;
- search of content in peer-to-peer networks.

Generative models

The ER model generates networks with Poisson degree distributions. However, it has been empirically observed that many networks in the real-world have a fat-tailed degree distribution of some kind, which varies greatly from the distribution observed for ER random networks.

In the *static* network models we have seen until now some parameters are fixed (e.g., number of nodes, number of edges, degree distribution etc.) and we study the properties of the graph (e.g., path lengths, component sizes etc.) .

The *generative* network models model the mechanism that drive the network formation. If the structures resemble real world structures, then this mechanism **might** be at work in real networks.

Price model

Price [1] proposed an elegant model of formation of citation networks. His model was inspired by the work of Hebert Simon, an economist who proposed an explanation for the wealth distribution: people who have more money already, gain more at a rate proportional to how much they already have. This can lead to power law distribution for the wealth (Rich-get-richer, cumulative advantage, preferential attachment, Matthew effect).

- Every new paper (node) cites on average c (outdegree) other papers;
- This newly appearing paper cites previously published papers:
  - with probability a a uniformly selected random paper, or otherwise
  - at random with probability proportional to the number of citations those previous papers have.

Price model

1. Price model creates acyclic graphs.
2. expected number of nodes with in-degree q
   Let $p_q(n)$ be the fraction of vertices that have in-degree q in a network with n nodes (after n steps). It satisfies the *master equation*

$$(n+1)p_q(n+1) = np_q(n) + \frac{c(q-1+a)}{c+a}p_{q-1}(n) - \frac{c(q+a)}{c+a}p_q(n)$$

with a special case, for $q = 0$

$$(n+1)p_0(n+1) = np_0(n) + 1 - \frac{ca}{c+a}p_0(n)$$

Price model

3. For asymptotic behavior of the degree distribution we use the shorthand $p_q = p_q(\infty)$. From the master equation we get the solution

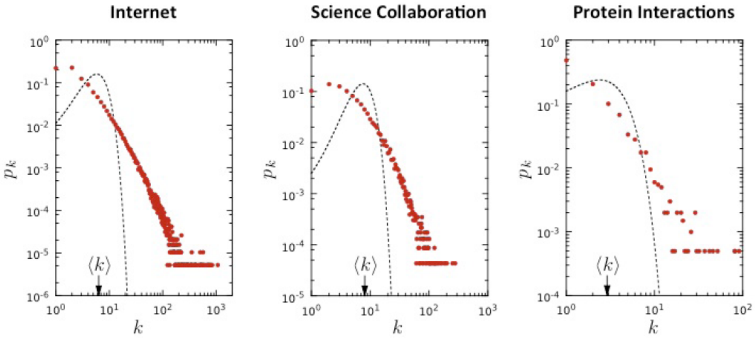
$$p_q = \frac{B(q+a, 2+a/c)}{B(a, 1+a/c)}$$

where $B(x, y)$ is the Euler's beta function.

Since for large x, $B(x, y) \approx x^{-y}\Gamma(y)$; for large values of in-degree q: $p_q \approx (q+a)^{-\alpha}$ or simply $p_q \approx q^{-\alpha}$, where $\alpha = 2 + a/c$.

4. the probability that an outgoing arc attaches to vertex u is: $\dfrac{q_u + a}{n(c+a)}$

Barabási–Albert model



Barabási–Albert model

A simple model to generate networks in which the probability of finding a node of degree d decays as a power law of the degree was put forward by Barabási and Albert in 1999. We initialize with a small network with m_0 nodes. At each step we add a new node u to the network and connect it to $m \leq m_0$ of the existing nodes $v \in V$. The probability of attaching node u to node v is proportional to the degree of v. That is, we are more likely to attach new nodes to existing nodes with high degree. This process is known as *preferential attachment*.

Netlogo: Barabási-Albert

Barabási–Albert model

1. The BA model is a special case of Price model with $a = c$.
2. The probability that a node has degree $d \geq d_0$ is given by

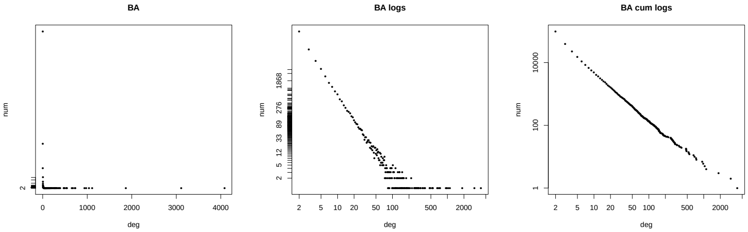
$$p(d) = \frac{2d_0(d_0-1)}{d(d+1)(d+2)} \approx d^{-3}$$

   That is, the distribution is close to a power law as illustrated in figure (left and middle).
3. The cumulative degree distribution is $P(d) \approx d^{-2}$, illustrated on a log–log scale in figure (right)
4. The expected value for the clustering coefficient, $\overline{cl}$, approximates

$$\frac{d-1}{8} \frac{\log^2 n}{n}$$

   as $n \to \infty$

Barabási–Albert model



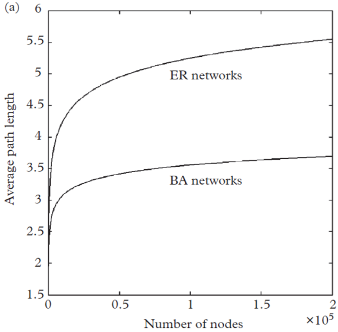
Barabási–Albert model

5. The average path length is given by

$$\bar{l} = \frac{\ln n \; - \; ln(d/2) \; - 1 - \gamma}{\ln \ln n + \ln(d/2)} + \frac{3}{2}$$

where again γ is the Euler-Mascheroni constant.

For the same number of nodes and average degree, BA networks have smaller average path length than their ER analogues. The figure shows the change in the average path length of random networks created with the BA and ER models as the number of nodes increases.

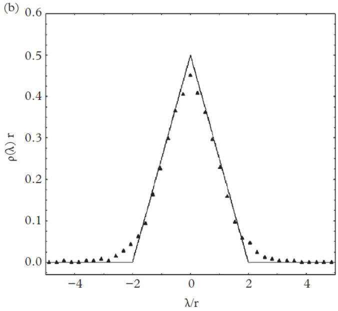


Barabási–Albert model

6. The density of eigenvalues follows a triangle distribution

$$\rho(\lambda) = \begin{cases} (\lambda + 2)/4 & -2 \leq \lambda/r \leq 0 \\ (2 - \lambda)/4 & 0 \leq \lambda/r \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

See figure.



7. these networks are resilient against random vertex or edge removals (random attacks), but quickly become disconnected when large degree nodes (*Achilles' heel*) are removed (targeted attacks).

Scale-free networks

The BA model can be generalized to fit general power-law distributions where the probability of finding a node with degree d decays as a negative power of the degree: $p(d) \approx d^{-\gamma}$.

Because for these networks their degree distribution has no natural scale they were named *scale free* networks.

$$\frac{p(ax_1)}{p(ax_2)} = \frac{p(x_1)}{p(x_2)}$$

For a discussion about the notion of scale-free network see Li et al.

. . . Scale-free networks

In real-life networks vertices often also form groups – clustering.

Several improvements and alternative models were proposed that also produce scale-free networks with some additional properties characteristic for real-life networks: copying (Kleinberg 1999), combining random and preferential attachment (Pennock et al. 2002), R-mat (Chakrabarti et al. 2004), forest fire (Leskovec et al. 2005), aging, fitness, nonlinear preferences, . . .

. . . Scale-free networks – exponent

Problems: large variability at the end, line only on an interval, nonuniform data density, . . .

Also the **complementary** cummulative distribution function is power law

$$\int_x^\infty Cx^{-\tau} = C\frac{x^{1-\tau}}{1-\tau}$$

Newman's estimate

$$\tau = 1 + n(\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}})^{-1}$$

M. E. J. Newman: Power laws, Pareto distributions and Zipf's law and Power-law distributions in empirical data. Packages in R: igraph. power, Pareto
Clauset: Toolkit for fitting, testing, and comparing power-law distributions in empirical data; Power-law distributions in empirical data; Scale-free networks are rare

See also

Matthew Effect: Wikipedia, When Do Matthew Effects Occur?
Epidemies: Barthélemy, Barrat, Pastor-Sattoras, Vespignani, Complex Networks Collaboratory.
Searching: Adamic et al.
General: Center for Complex Network Research, Newman, Borner, Sanyal, Vespignani.

Resources I

- Allison, PD, de Solla Price, D, Belver C, Griffith MJ, Moravcsik JA (May 1976): Lotka's Law: A Problem in Its Interpretation and Application. Social Studies of Science, Vol. 6, No. 2, pp. 269-276
- Barabási, A.-L. and Albert, R., Emergence of scaling in random networks, Science 286:509–512, 1999.
- Batagelj, V, Brandes, U: Efficient generation of large random networks. PHYS REV E 71 (3): - Part 2, 036113, 2005
- Bollobás, B., Random Graphs, Cambridge University Press, 2001.
- Clauset, Aaron (2017): CSCI 5352: Network Analysis and Modeling WWW
- Erdõs, P, Rényi A (1959): Publ. Math. Debrecen, 6, 290 .
- Estrada, Ernesto and Knight, Philip A.: A First Course in Network Theory. Oxford University Press, 2015.

Resources II

- Gilbert, E. N. (1959): Ann. Math. Stat. 30, 1141.
- Goldenberg, Anna, Zheng, Alice X., Fienberg, Stephen E., Airoldi, Edoardo M. (2009): A Survey of Statistical Network Models. ArXiV.
- Mateos, Gonzalo (2018): Course ECE442 – Network Science Analytics. University of Rochester. WWW
- Murphy, Phil: rPubs. WWW
- Pelechrinis, Konstantinos (2015): 2125: Network Science and Analysis. WWW
- Watts, D.J., Strogatz, S.H., Collective dynamics of 'small-world' networks, Nature 393:440–442, 1998.