WARWICK
THE UNIVERSITY OF WARWICK

Random Graphs

Susana Gomes

Last updated November 18, 2023

Plan for the rest of the module (there might not be time for everything)

1. Introduction to Random graphs
2. Erdős–Rényi random graph (including percolation)
3. Spectra of random matrices
4. Configuration model
5. Preferential attachment
6. Small-world Networks and the Watts–Strogatz model
7. What else can we do?
   a) Short discussion on community detection
   b) Dynamics on networks: some examples.
8. Final remarks :)

Introduction to Random Graphs

Why do we need random graphs?

When we are analysing a graph or network, it is very useful to be able to compare its properties to those of an appropriate “benchmark" (a network with some kind of *“expected behaviour"*).

Random graphs provide us with such benchmarks. There are several examples and we will see a few of them for the rest of this module. Most of the time, they are split between two types:

- Graphs where existing edges follow certain (random) constraints.
  - ⋆ Erdős–Rényi graph (today!)
  - ⋆ Configuration model (next week)
- Mechanistic models, defined by simple rules on how nodes and links are created or destroyed in the course of time.
  - ⋆ Growing network with preferential attachment (next week)
  - ⋆ Watts–Strogatz model (next week)

Some comments

It is important to note that (see Lambiotte's lecture notes for more details)

- these random models don't provide explanations for the values taken by the parameters or the reason for certain constraints to emerge in a “real-life" network.
- they have nice mathematical and statistical properties, making them a useful baseline for deciding whether patterns observed in empirical data are significant.
- In practice, if a value of a measurement observed in empirical data is significantly different from the expected value for the random graph model, the model does not represent the process behind the empirical data.
- Comparing these models and real-life networks allows us to identify potential mechanisms driving the evolution of the empirical networks.

Erdős–Rényi Random Graphs

The simplest example of a random graph is the Erdős–Rényi random graph (sometimes also called the Poisson or binomial random graph), and it was introduced by the Hungarian mathematicians Erdős and Rényi in 1959.

The ER graph can be defined in two ways (the second one being the most common):

Definition (E-R graph)

An **Erdős–Rényi (short E-R) random graph** $G \sim \mathcal{G}_{N,K}$ is a graph with uniform distribution on the set of all undirected graphs with N vertices and $K = |E|/2$ edges, i.e.

$$\mathbb{P}_{N,k}[G = (V, E)] = 1 \Big/ \binom{\frac{N(N-1)}{2}}{K}.$$

More commonly, an **(E-R) random graph** $G \sim \mathcal{G}_{N,p}$ is a graph with N vertices and where each (undirected) edge is present independently with probability $p \in [0, 1]$, i.e.

$$\mathbb{P}_{N,p}[G = (V, E)] = p^K (1 - p)^{\frac{N(N-1)}{2} - K} .$$

Some notes about the ER graph

- $\mathcal{G}_{N,p}$ does not represent a single network, but a random ensemble of them in the probabilistic sense.
- Any network without multiple edges or self-loops can be generated by the random graph $G(N; p)$ as long as $0 < p < 1$.
- However, note that the probability that the model generates a target network, or a similar network, may be tiny.
- Most of the time with these networks, our goal is to predict the average behaviour of certain network metrics and, if possible, their variance.
- As mentioned above, it is easier to work with $\mathcal{G}_{N,p}$ than $\mathcal{G}_{N,K}$. For N, K large, $\mathcal{G}_{N,K}$ is largely equivalent to $\mathcal{G}_{N,p}$ with $p = \frac{2K}{(N(N-1))}$.
- Since edges are present independently of each other, graphs $G \in \mathcal{G}_{N,p}$ should typically be uncorrelated, and indeed one can show that its correlation coefficient, $\chi(G)$, satisfies

$$\mathbb{E}[\chi] \to 0 \quad \text{as} \quad N \to \infty.$$

Properties of E-R graphs (edges and degrees)

As stated above, we wish to characterise the behaviour of these networks as best we can. From the definition, we can already say a lot:

- The **number of undirected edges** for $G \sim \mathcal{G}_{N,p}$ is random, with distribution $K \sim \mathrm{Bi}\left(\frac{N(N-1)}{2}, p\right)$, i.e.

$$\mathbb{P}\left(\frac{|E|}{2} = K\right) = \binom{\frac{N(N-1)}{2}}{K} p^K (1-p)^{\frac{N(N-1)}{2} - K}.$$

- From this, we can already predict the **expected number of edges**, which is $\mathbb{E}(K) = p\frac{N(N-1)}{2}$.
- Similarly, we can also predict the **average degree**. It is also random, and is given by

$$\langle k \rangle = \frac{2K}{N} \quad \text{with} \quad \mathbb{E}[\langle k \rangle] = (N-1)p.$$

- By homogeneity, this means that, for all i, we have $k_i \sim \mathrm{Bi}(N-1, p)$ and $\mathbb{E}[\langle k \rangle] = \mathbb{E}[k_i] = (N-1)p$.

Properties of E-R graphs (scaling limits)

The ER model is mostly used for sparse networks, where the total number of edges scales linearly with the number of nodes, N.

This means that the average degree of the nodes $\langle k \rangle = p(N-1)$ should not depend on N and so we usually consider a small value of p, such as $q \propto \frac{1}{N}$.

In the limit of large networks, we consider $p = \frac{\langle k \rangle}{N-1}$, and in this case, the binomial degree distribution is well approximated by the Poisson distribution:

$$\mathbb{E}[p(k)] = \mathbb{P}[k_i = k] = \binom{N-1}{k} p_N^k (1-p_N)^{N-1-k} \to \frac{z^k}{k!}\, e^{-z}.$$

This is why E-R $\mathcal{G}_{N,p}$ graphs are sometimes called **Poisson random graphs**.

Properties of E-R graphs (clustering, other measures)

We can derive a lot of other properties ofthe ER random graph thanks to the independence of edges:

- The **average distance** of the ER random graph (hard to derive) is

  $$L \approx \frac{\log N}{\log \langle k \rangle},$$

  for $p > O\left(\frac{\log N}{N}\right)$ as $N \to \infty$ (ensures that the network is connected).
- We can also count the **number of triangles**, $\binom{N}{3} p^3$, and the **number of triples**, $\binom{N}{3} 3p^2$.
- Using this, we can find the **clustering coefficient**: $C = \frac{\binom{N}{3} p^3}{\binom{N}{3} p^2} \to p$
- When $\langle k \rangle$ does not depend on N, we obtain $C \to 0$ as $N \to \infty$.
- This can be generalised to counts of loops or cliques of larger size: the density of such structures decays to zero as $N \to \infty$.
- This means that as $N \to \infty$ the network has a locally tree-like structure, i.e., connected components are tree subgraphs with probability 1.

Percolation and E-R graphs

Percolation

Percolation studies how robust the connectivity properties of graphs are under deletion of edges or vertices (e.g. random attacks or immunization).

To motivate the following definitions, think of an E-R random graph $\mathcal{G}_{N,p}$ as a *complete graph* which had edges removed with probability $1-p$. We can do this for any random graph.

Take any connected, undirected graph $G=(V,E)$ and choose $p \in (0,1)$. We can construct a subgraph of G by removing edges (independently) with probability $1-p$.

- The edges we removed are called **closed**.
- The remaining edges are called **open** and we define

$$E^o = \{\text{set of open edges}\}.$$

Note that we choose to call the edges open and closed by making the analogy of being “open to transport" (so an edge exists) or closed otherwise.

We denote the (random) subgraph of G containing only open edges by

$$G^o = (V, E^o) \quad \text{with} \quad E^o = \{(i,j) \in E : e_{ij} = 1\} \subset E.$$

Percolation

The probability distribution associated with the initial graph $G(V,E)$ is called **bond percolation**. It is a static probabilistic model with state space

$$S = \Omega = \{0,1\}^E = \left\{e_{ij} \in \{0,1\} : (i,j) \in E\right\}.$$

and distribution $p = \mathbb{P}[e_{ij} = 1] = 1 - \mathbb{P}[e_{ij} = 0]$, i.e. $e_{ij} \sim \mathrm{Be}(p)$ iid with $p \in [0,1]$.

Definition:

We say that a sequence of connected graphs G_N with increasing size $|V_N| = N$ exhibits **percolation with parameter p** if

$$\frac{|\bar{C}_N^o|}{N} \geq c > 0 \quad \text{as } N \to \infty \quad \text{with probability } 1,$$

where we define $|\bar{C}_N^o| = \max\limits_{i \in V_N} |C_i^o|$ to be the size of the largest connected component $\bar{C}_N^o$ of G_N^o.

An alternative/similar concept is that of **site percolation**. Here, we remove vertices at random, and delete their adjacent edges.

Percolation and E-R graphs

As mentioned before, we can think of E-R random graphs $\mathcal{G}_{N,p}$ as having the same distribution as open subgraphs $G^o = (V, E^o)$ with $E^o \subset K_N$ under **percolation on the complete graph $G_N = (V, K_N)$** with parameter p.

Using this one can derive the following result.

Theorem (Giant component for E-R graphs)

Consider $G_{N,p} \sim \mathcal{G}_{N,p}$ with $p = \frac{z}{N} = \frac{\langle k \rangle}{N}$ and its **maximal connected component** $\bar{C}_{N,p}$. Then

$$|\bar{C}_{N,p}| = \begin{cases} O(\log N) & \text{, for } z < 1 \\ O(N^{2/3}) & \text{, for } z = 1 \\ O(N) & \text{, for } z > 1 \end{cases} \quad \text{and} \quad \theta(z) = \begin{cases} 0 & \text{, for } z \leq 1 \\ > 0 & \text{, for } z > 1 \\ \to 1 & \text{, for } z \to \infty \end{cases}$$

where $\theta(z) := \lim_{N\to\infty} \dfrac{|\bar{C}_{N,p}|}{N}$ is a continuous, monotone increasing function of z. For $z > 1$, $\bar{C}_{N,p}$ is the **only giant component**, and the second largest has size $O(\log N)$.

Some intuition on the previous theorem

This theorem actually shows that the E-R random graph exhibits a **phase transition**. Phase transitions appear in a lot of real-life systems, and represent qualitative changes in the behaviour of a system due to nonlinearities. They imply that **small changes of a parameter** may have **drastic consequences** on the organisation and dynamics of a system.

To see what this means for us, assume that we are looking at the size of the largest connected component in the network (the **order parameter**) as a function of the mean degree $\langle k \rangle$ (the **control parameter**).

- If $\langle k \rangle = 0$, then the E-R graph is simply N disconnected nodes.
- On the other hand, if $\langle k \rangle = N - 1$, then the resulting graph is complete (or quite close).
- What the theorem says is that between these two cases, the network **does not change smoothly** in terms of the largest component size.
- Instead, a **giant component**, i.e., a component whose size is the largest and proportional to N, suddenly appears as $\langle k \rangle$ increases, marking a phase transition.

An illustration of this phase transition

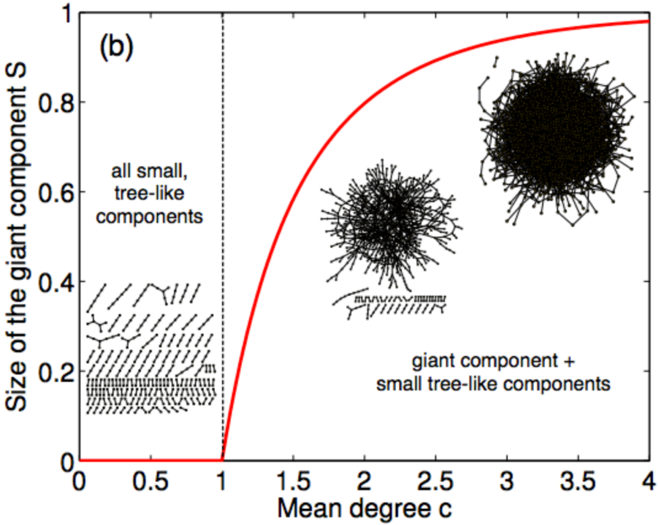


Illustration of the different regimes of the E-R model (from Lambiotte's lecture notes)

A note on numerical simulation of ER graphs

[We will see more about this in the next support class]

You can simulate the E-R random graph in many different ways

1. One can perform a Bernoulli test of probability p for all possible pairs of nodes and define the edges using this. This requires $O(N^2)$ operations, which can be very computationally expensive for large networks.
2. You can draw the degrees of the N nodes from the Poisson distribution and build the graph using the configuration model (which we will see next), which respects the degree constraint. This is substantially faster for large networks

Spectra of random graphs

Spectra of random graphs

Similarly to what we have seen before, we can have a look at the spectrum of a random graph. For this, it is convenient to look at Wigner matrices.

Definition (Wigner matrix)

A **Wigner matrix** $A = (a_{ij} \ : i, j = 1, \dots, N)$ is a real, symmetric matrix with iid entries a_{ij} which has finite moments, and for which $\mathbb{E}[a_{ij}] = 0$ and $\mathrm{Var}[a_{ij}] = \sigma^2$.

The Wigner semi-circle law tells us things about the spectral density of these matrices.

Theorem (Wigner semi-circle law)

The spectral density ρ_N of the matrix $A/\sqrt{N}$ (where A is a Wigner matrix) converges in distribution to

$$\rho_N(\lambda) \to \rho_{sc}(\lambda) := \begin{cases} (2\pi\sigma^2)^{-1}\sqrt{4\sigma^2 - \lambda^2} & \text{, if } |\lambda| < 2\sigma \\ 0 & \text{, otherwise} \end{cases} .$$

ρ_{sc} is called the semi-circle distribution

Consequences of the Wigner semi-circle law

- The bulk of eigenvalues of unscaled Wigner matrices typically lies in the interval $[-2\sqrt{N}\sigma, 2\sqrt{N}\sigma]$.
- Adjacency matrices A of $\mathcal{G}_{N,p}$ **random graphs** are symmetric with iid $\mathrm{Be}(p)$ entries with $\mathbb{E}[a_{ij}] = p$ and $\mathrm{Var}[a_{ij}] = p(1-p)$, so are **not Wigner matrices**.
- For $\mathcal{G}_{N,p}$, A has a **maximal Perron-Frobenius eigenvalue** of order pN, but all other eigenvalues have modulus of order $\sqrt{N}$.
- For fixed $p > 0$ or scaled $p = p_N \gg p_c = 1/N$ the Wigner semi-cirlce law holds for $N \to \infty$ with support width $4\sqrt{N}\sigma_N$.
- However, for $p = p_N \ll p_c = 1/N$ the asymptotic spectral density deviates from ρ_{sc}.
- There is a related **circular law** for **non-symmetric Wigner matrices**.

Configuration model

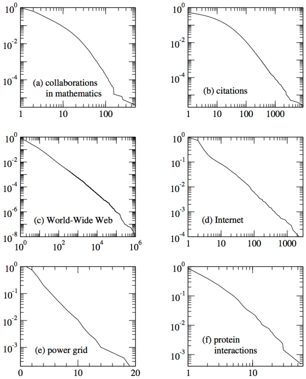
Motivation

The Erdős-Rényi random graph plays a fundamental role in network science, as it helps us understand real-world phenomena, such as the small average distance L between any 2 nodes which we discuss before.

However, it also produces unrealistic patterns, such as locally tree-like structure and the Poisson degree distribution.

In fact, the majority of real-life networks has many triangles (i.e., it is not tree-like) and has a long-tailed degree distribution.

This motivates the definition of other random graphs.



Configuration model

The configuration model can be seen as a generalisation of the E-R random graph, but where we prescribe the **degree distribution** instead of the probability of having an edge.

It is used to study the effect of “heterogeneous" degree distributions (e.g. if all nodes have the same degree) on networks, and doesn’t have a lot of other clustering properties, like the E-R graph does.

Definition (Configuration model):

The **configuration model** $\mathcal{G}^{\text{conf}}_{N,D}$ is defined as the uniform distribution among all undirected graphs with N vertices with the constraint that the degree of node i is a given k_i, i.e., with a given degree sequence $D = (k_1, \ldots, k_N)$, such that $\sum\limits_{i \in V} k_i = 2K$.

Some notes on the configuration model

- The degree sequence $D = (k_1, \ldots, k_N)$ is usually obtained from a chosen degree distribution $p(k)$, under the constraint that the sum of degrees must be an even number (this is a requirement for all graphs, and is known as the *handshaking lemma*).
- For directed versions with D^{in} and D^{out} we need $\sum_{i \in V} k_i^{\mathrm{in}} = \sum_{i \in V} k_i^{\mathrm{out}}$.
- Not all sequences D that sum to an even number are **graphical** (i.e., they don't always correspond to a graph).
- E-R graphs are given by $\mathcal{G}_{N,p} = \mathcal{G}_{N,D}^{\mathrm{conf}}$ with random degree sequence D consisting of iid degrees $k_i \sim \mathrm{Bi}(N-1, p)$.
- General **randomized graphs** with given degree distribution $p(k)$ can be sampled in the same way.

Simulating the configuration model

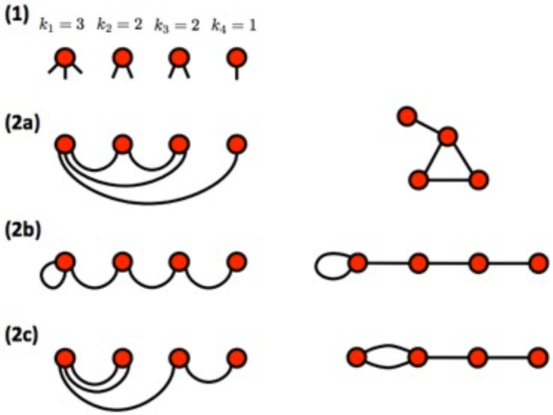
To generate an instance of the configuration model for a given degree sequence, we proceed as follows:

1. Create **stubs** (half-edges) at each node *i* such that its number is equal to k_i.

2. Randomly select pairs of stubs one by one to **connect them as an edge**.

3. We must avoid the case in which the link creation stops in the middle:

   For example, if there remain three nodes which have 1, 2, and 3 unused stubs, we have to create three more edges because there are six stubs remaining. However, we cannot do that without a self-loop or multiple edge.



Different realisations of the configuration model given the sequence of node degrees (1). Taken from Lambiotte's lecture notes.

Characteristic path length of the graph

- If N is large, when $\langle k^2 \rangle$ is finite, the average distance of the configuration model is given by

$$L = 1 + \frac{\frac{\log N}{\langle k \rangle}}{\log \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}}.$$

- We discussed power-law degree distributions ($p(k) \propto k^{-\gamma}$) before. In that case, we have

$$L \propto \begin{cases} \log \log N, & \text{if } 2 < \gamma < 3, \\ \frac{\log N}{\log \log N}, & \text{if } \gamma = 3, \\ \log N, & \text{if } \gamma > 3. \end{cases}$$

We note that in this case, we only have $\langle k^2 \rangle < \infty$ for $\gamma > 3$ which is consistent with the previous comment.
If $\gamma \leq 3$ L is very small, and in this case the network is called **ultra-small world** (we will see more about this next week)

Clustering coefficient and other measures

The clustering coefficient is given by

$$C = \sum_{k'=1}^{\infty} \sum_{k''=1}^{\infty} \frac{k' p(k')}{\langle k \rangle} \frac{k'' p(k'')}{\langle k \rangle} \frac{(k'-1)(k''-1)}{\langle k \rangle N} = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^3 N}$$

This is small unless the degree distribution is highly heterogeneous. In this case, $\langle k^2 \rangle \ll \langle k \rangle^2$

A stub emanating from i is connected to j with probability $\frac{k_j}{2K}$. Since i owns k_i stubs, the expected number of edges between i and j is $A^*_{ij} = \frac{k_i k_j}{2K}$ where A^*_{ij} represents the statistical average of the adjacency matrix.

For this annealed adjacency matrix, the largest eigenvalue can be evaluated and is

$$\lambda_1 = \frac{\langle k^2 \rangle}{\langle k \rangle}.$$

This diverges for networks with power-law degree distributions $p(k) \propto k^{-\gamma}$ and $\gamma \leq 3$. In this case, the largest eigenvalue diverges as $N \to \infty$.

Giant components

Finally, if $k_{\mathsf{max}} = \mathsf{max}_i\, k_i$ is bounded, one can show that graphs obtained from the configuration model exhibit a giant (connected) component of size $O(N)$ if

$$Q := \sum_{k \geq 0} k(k-2)p(k) > 0.$$

If $Q < 0$ the largest component is of size $O(k_{\mathrm{max}}^2 \log N)$.

Network growth and Preferential attachment

Network growth

The prevalence of power-law degree distributions in real-life complex networks can be attributed to growth mechanisms subject to **preferential attachment**, which we will look at next.

The majority of networks we study grows in time both in terms of the number of nodes N and the number of edges - an example which we talked about is the network of collaborations or citations in science.

An easy example to think of is a phenomenon called **triadic closure**, in which if there is an edge between vertices v_1 and v_2, and another between v_2 and v_3, then a new edge (v_1, v_3) is likely to form, yielding a high clustering coefficient.

The main example of a network with preferential attachment is the **Barabási–Albert random graph** which we will see next.

Preferential attachment

Definition

The **Barabási-Albert** random graph, denoted by $\mathcal{G}^{\mathrm{BA}}_{N,K}$ is an undirected graph with N nodes and $K = m_0(m_0 - 1)/2 + m(N - m_0)$ edges which is defined as follows:

1. Start with a complete graph (V_0, E_0) with $|V_0| = m_0$ nodes.
2. At each time step $t = 1, \dots, N - m_0$, add a new node $j = t + m_0$.
3. Add $m \leq m_0$ undirected edges with existing nodes $i \in V_{t-1}$ with a probability proportional to their degree:

   $$\pi_{j \leftrightarrow i} = \frac{k_i}{\sum_{l \in V_t} k_l}$$

   This probability is called preferential attachment.

This model is known for a long time (even though it is named after Barabási and Albert, who proposed it more recently). It belongs to the same family of *multiplciative stochastic models* as, e.g. the Pólya urn model we saw in assignment 1.

Example

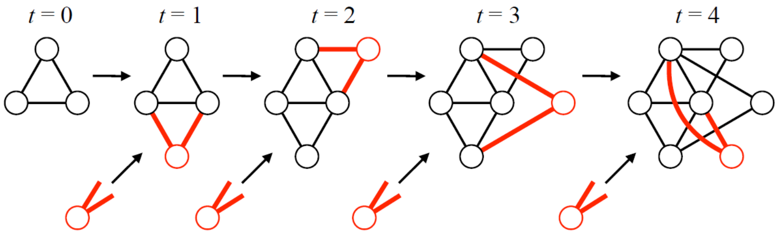


FIG. 17. First several stages of the BA model. The bold lines represent new links. We set $m_0 = 3$ and $m = 2$.

(figure taken from Lambiotte’s lecture notes)

Some known results

- As $N \to \infty$, the average degree of the BA graph is $\langle k \rangle = 2m$.
- Similarly, as $N \to \infty$ the degree distribution $p_N(k)$ converges to $p(k) \propto k^{-3}$, i.e., to a **power law** distribution, with an exponent which is close to exponents observed for real-world networks.
  - ⋆ This is independent of the parameters m_0 and m!
  - ⋆ For a proof, see Lambiotte's lecture notes
- The BA graph is **not homogeneous**: the expected degree of a node increases with age. It is also **uncorrelated**.
- We can compute the characteristic path length L and clustering coefficient C for the BA graph. They typically behave like

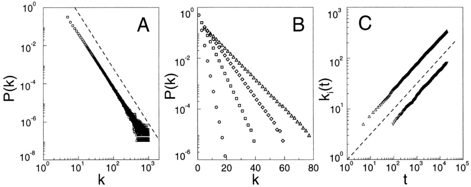
$$L \propto \begin{cases} \log N & \text{if } m = 1 \\ \frac{\log N}{\log \log N} & \text{if } m \geq 2 \end{cases}, \qquad C \approx \frac{m-1}{8} \frac{(\log N)^2}{N}.$$

- The above equation implies that the BA model lacks the clustering property because $C \to 0$ as $N \to \infty$. Various extensions of the BA model realise a non-vanishing C value as $N \to \infty$.

Preferential attachment

We can construct variations of the BA model.[1] Examples include:

- Connecting to vertices i with probability proportional to $k_i + k_0$ (for some k_0), leading to power law degree distributions with $\alpha = 3 + \frac{k_0}{m}$.
- **Nonlinear preferentail attachment**, with probability proportional to k_i^γ , where we get:
  - ⋆ for $\gamma \in [0, 1)$, $p(k)$ has a **stretched exponential tail** $\exp(-Ck^{1-\gamma})$ and the graph is **assortative**.
  - ⋆ for $\gamma > 1$, all vertices connect to m **super vertices** and the graph is **disassortative**.



(A) power law for $\gamma = 1$, $m_0 = m = 5$, $N = 200K$, (B) exponential tail for $\gamma = 0$, $m_0 = m = 1, 3, 5, 7$, (C) degree increasing with time for $t_1 = 5$, $t_2 = 95$, taken from [A.-L. Barabási, R. Albert, Science **286**(5439), 509-512 (1999)]

[1] See **this review by Newman** for more examples.

Small-World Networks

The small-world property

An important property to consider when studying real-world networks is the small-world property.

Definition:

A sequence of connected graphs G_N with increasing size $|V_N| = N$ exhibits the **small-world property**, if the characteristic path length $L(G_N) = O(\log N)$.

More generally, we say that a network has the small-world property if the average distance between nodes increases sufficiently slowly as a function of the number of nodes in the network.

Some examples which we have seen already include

- Trees with degrees $k_i \geq 3$,
- The giant or largest component in E-R random graphs.
- In most graph models, small-worldness is paired with low clustering coefficients, e.g. 0 for trees and p for $\mathcal{G}_{N,p}$ graphs.
- However, many real examples of small world networks also exhibit **large clustering coefficients**, such as networks of social contacts.

Watts-Strogatz random graph

Most of the time, when people refer to a small-world network, they actually refer to the Watts-Strogatz random graph. [2]

Definition:

Consider a $2m$-regular ring graph - this is a graph with N nodes and a total number $K = mN$ of undirected edges, and with adjacency matrix

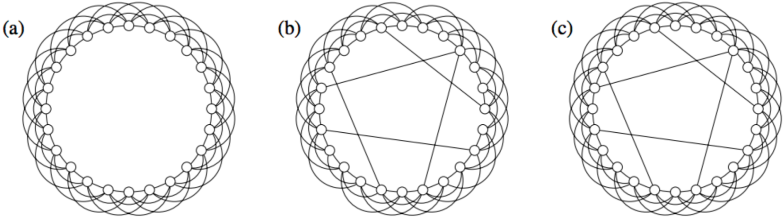
$$a_{ij} = \begin{cases} 1 & , |i - j| \leq m \\ 0 & , \text{otherwise.} \end{cases}$$

For all nodes i, each edge (i, j) with a clockwise neighbour $j > i$ is **rewired** with probability $p \in [0, 1]$, i.e. replaced by an edge (i, l) where l is chosen uniformly among vertices not adjacent to i.

The resulting graph is a **Watts-Strogatz random graph**, denoted by $\mathcal{G}^{\mathrm{WS}}_{N,K}$.

[2] See this **Scholarpedia article by Mason Porter** about small-world networks or this **Wikipedia page on the W-S model** - the Watts here is the author of the book I mentioned a couple of weeks ago, and Strogatz is the Steven Strogatz you might have heard of before.

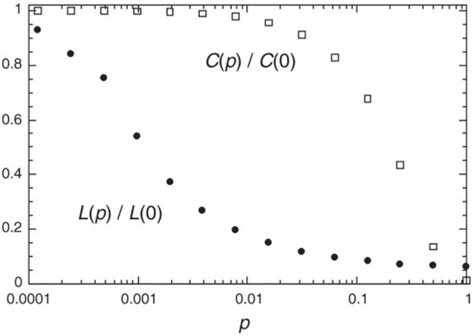
Illustration of this graph



(a) A one-dimensional lattice with connections between all vertex pairs separated by k or fewer lattice spacing, with k = 3 in this case. (b) The small-world model of Watts and Strogats is created by choosing at random a fraction p of the edges in the graph and moving one end of each to a new location, also chosen uniformly at random. (c) A slight variation on the model in which shortcuts are added randomly between vertices, but no edges are removed from the underlying one-dimensional lattice. Taken from Newman, Mark EJ. "The structure and function of complex networks." SIAM review 45.2 (2003): 167-256.

Properties of the W-S model

- The W-S model was introduced to try and “fix" one of the main ussues of E-R random graphs: it exhibits high clustering coefficients while having short average path lengths.
- W-S random graphs interpolate between a regular lattice for $p = 0$ and a $\mathcal{G}_{N,K}$ E-R random graph conditioned on the event that all vertices have degree $k_i \geq m$.
- The expected clustering coefficient $\mathbb{E}[C(p)]$ and average path length $\mathbb{E}[L(p)]$ are decreasing functions of p and show the following behaviour:



Taken from [D.J. Watts, S.H. Strogatz, Nature **393**, 440-442 (1998)]

A (very biased!) overview of what else is there

Community detection motivation

In many real-world scenarios, networks can exhibit a community structure.

This happens when the network is composed of groups of nodes such that the nodes are densely connected within the same group and relatively sparsely connected across different groups.

In some cases (which we have seen when discussing graph Laplacians), there are only two groups, but there is no reason for this to always be the case.

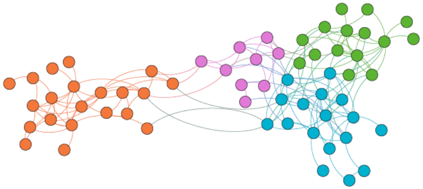


FIG. 18. A social network of bottlenose dolphins. Four communities detected by the Louvain algorithm implemented on gephi (http://www.gephi.org) are shown by different colours.

Graph partitioning and modularity

There are plenty of algorithms for community detection, and most of them are based on the concepts of graph partitioning and modularity.

- We already saw **spectral partitioning**, where a graph is split into 2 communities using the **Graph Laplacian** (and its **Fiedler vector**).
  - ⋆ See Lambiotte's notes, section VI.A. for an explanation of why this works
  - ⋆ More complicated versions of this method, partitioning a graph with N nodes into c groups, are common in computer science, but are highly computationally expensive (with a complexity of N^{c^2})
- This requires knowing the number of groups, c, and their size - a lot of the time this is not known. In this case, we can use **modularity** to quantify how good a partition is.
  - ⋆ Modularity compares the number of edges within a community with, e.g., the edges a *null model* would have.
  - ⋆ Modularity can be optimised, in the same spirit as spectral partitioning.

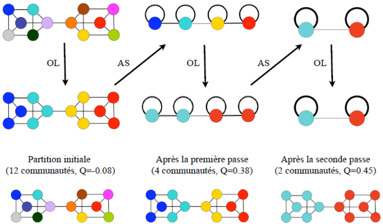
Louvain algorithm

The previous methods for community detection are computationally expensive since they require computing eigenvalues and eigenvectors of large matrices.

An alternative to this is to use *greedy algorithms*, and an example of this is the **Louvain method**, which is implemented in most libraries and packages.

The algorithm consists of two phases, which are repeated until an optimal partition (in terms of modularity) is obtained:[3]

1. local optimisation
2. “merging" of vertices



[3] Figure from Lambiotte’s lecture notes

Dynamics on networks

We have mentioned before that some of the interest on identifying structures of netowrks is to see how they evolve in time.

One simple example is the effect of adding a node to a network (e.g. citations or preferential attachment) but actually we can have more complicated things going on, like :

- the evolution of opinions and formation of consensus on a social network, leading to, for example, results on elections,
- evolution of epidemics on a network (recall contact process)
- many other examples

Opinion dynamics and consensus

One of the most famous problems you can study on a network is that of consensus dynamics.
This is because of the simplicity of the basic models, which lead to analytic tractability.
This is used, for example

- In socio-economics, consensus provides a model for opinion formation in a society of individuals.
- In engineering, it can be used as a building block for an efficient distributed computation of global functions in networks of sensors, robots, or other agents.
- In collective dynamics, for models of movement of groups of animals.

Opinion dynamics and consensus

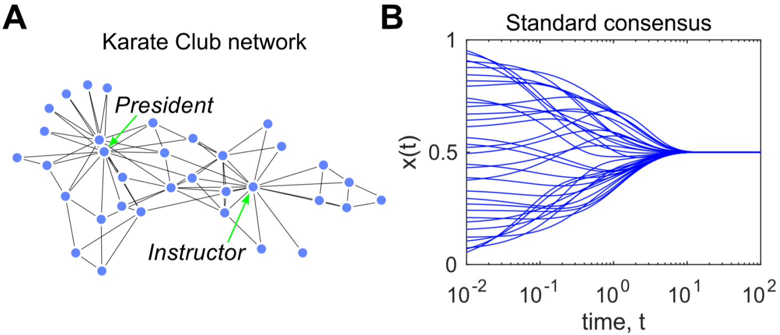
The dynamics is defined for a connected network of N nodes with adjacency matrix A, and where the node i has a value $x_i \in \mathbb{R}$ associated to it (its opinion, it's position, speed, direction ...). The (average) consensus dynamics on such a network is then defined as

$$\dot{\mathbf{x}} = L\mathbf{x} \quad \text{or} \quad \dot{x}_i = \sum_{j=1}^{n} A_{ij}(x_i - x_j)$$

where L is the (possibly weighted) graph Laplacian. This means that each node adjusts its state such that the difference to its neighbours is reduced.

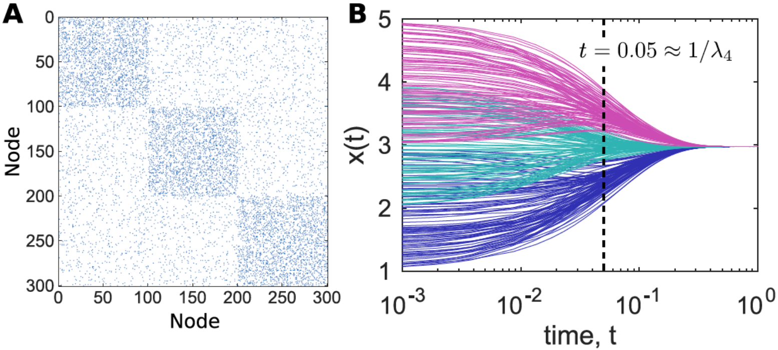
We can see that if the system starts at $x(0) = x_0$, then the differential equation drives the state to a global 'consensus state' in which the state variables of all nodes are equal, $x_i = x_\star$ for all i as $t \to \infty$, where $x_\star = \mathbf{1}^T x_0 / n$.

Consensus



(From Lambiotte's Lecture notes:) FIG. 22. Illustration of a consensus dynamics on the Karate Club network. A: Karate Club network orginally analysed by Zachary. B: Consensus dynamics on the Karate club network starting from a random initial condition.

Consensus - another example



(From Lambiotte's Lecture notes:) FIG. 23. Illustration of a consensus dynamics on a structured network with time-scale separation. A: Adjacency matrix (unweighted) of a structured network with 5 groups. B: A consensus dynamics on this network displays a time-scale separation: until around t = 0:05, approximate consensus is reached within each group (groups indicated by color); then a consensus is reached between the groups.

Dynamics on networks - random walks

Another important type of dynamics on a network is **diffusion**. This describes how an entity randomly explores the underlying structure of the network.

Consider a walker diffusing on an undirected network: at each step, the walker is located on a node i, selects one of the edges leaving i at random and jumps to an adjacent node, say j. This process is equivalent to a Markov chain and can be completely described by the $N \times N$ transition matrix P which we described before!

Using the Markov chain techniques from the beginning of the module, we can describe all sorts of things, like the probability that the walker visits the $i-$th node after n steps, $\pi_n(i)$, where

$$\pi_n = \pi_0 P^n.$$

It can be shown that the transition matrix of this process is very similar to the normalised graph Laplacian which we saw before!

This justifies the **theory behind PageRank**. (for more on this, see section VIII in Lambiotte's lecture notes)

Other examples and more resources

There are many more examples of things we could do which we do not have time to cover. You may want to look at Lambiotte's lecture notes (sections listed below) or Newmann's book.

- You can look more in detail into time-scale separation (Section VII: C-E)
- Or how to use all of this to reveal hidden structures in your networks (Section IX, and recall MathSys Forum in November)
- And many other things, like dynamics of epidemic processes in networks (Section X)

There are also lots of resources online which you can have a look. For example, there is this **Network database** with loads of data, or **this website** with a lot of analysis on networks.

There is much more we can discuss, but for now, that's it! I hope you enjoyed this module, and are curious to see what's next!