Structure and dynamics of complex networks

March 19, 2020

Network models

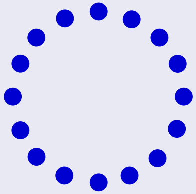
THE ERDŐS-RÉNYI MODEL

The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently with probability p.
- This is also called as $G(N, p)$ model.
- The $G(N, M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

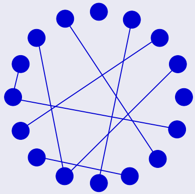


The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently with probability p.
- This is also called as $G(N,p)$ model.
- The $G(N,M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

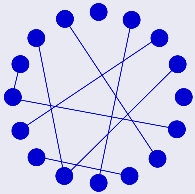


The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently with probability p.
- This is also called as $G(N, p)$ model.
- The $G(N, M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

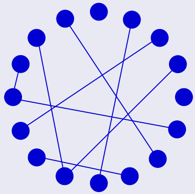


The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently with probability p.
- This is also called as $G(N, p)$ model.
- The $G(N, M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.



Properties of the Erdős-Rényi model

Properties of the E-R model:

- average degree?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- average degree: $\langle k \rangle = (N-1)p \simeq Np$,

Properties of the Erdős-Rényi model

Properties of the E-R model:

- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links: $L = pN(N-1)/2$.

Properties of the Erdős-Rényi model

Properties of the E-R model:

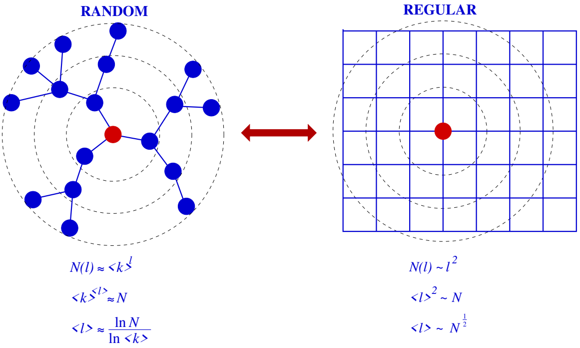
- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links: $L = pN(N-1)/2$.
- small world property?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links: $L = pN(N-1)/2$.
- small world property?

→ as discussed on Lecture no.3, due to its random nature, it is also small world:



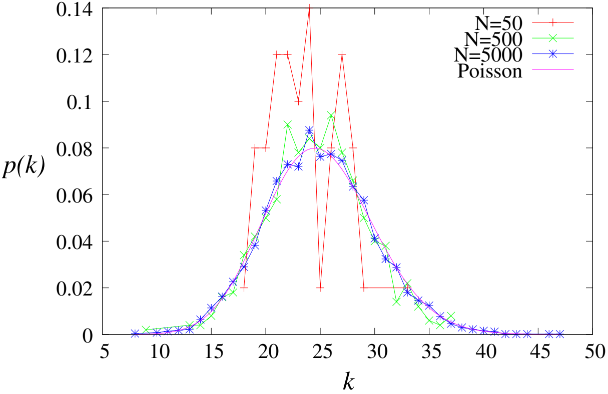
Degree distribution of the E-R graph

Degree distribution of the E-R graph

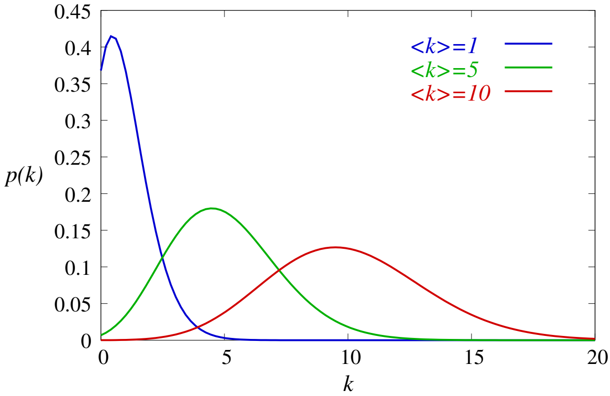
We have already derived the $p(k)$ of the E-R graph earlier,

$$\begin{aligned} p(k) &= \binom{N-1}{k} p^k (1-p)^{N-1-k} \simeq \binom{N}{k} p^k (1-p)^{N-k} && \text{(binomial)} \\ &\simeq \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} && \text{(Poisson)} \end{aligned}$$

Degree distribution of the E-R graph



Degree distribution of the E-R graph



Degree distribution of the E-R graph

Variance

- What is the **variance** of $p(k)$?

The degree distribution is binomial,

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k},$$

→ the average and variance for a binomial distribution in general is

$$\begin{aligned} \langle k \rangle &= Np, \\ \langle k^2 \rangle &= Np(1-p) + p^2N^2, \\ \mathrm{Var}(k) &= \langle k^2 \rangle - \langle k \rangle^2 = Np(1-p), \\ \sigma(k) &= \sqrt{\mathrm{Var}(k)} = \sqrt{Np(1-p)}. \end{aligned}$$

(For more details about the binomial distribution see e.g. https://en.wikipedia.org/wiki/Binomial_distribution)

Degree distribution of the E-R graph
Variance

- What is the **variance** of $p(k)$?

  The degree distribution is binomial,

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k},$$

→ the average and variance for a binomial distribution in general is

$$\begin{aligned}
\langle k \rangle &= Np, \\
\langle k^2 \rangle &= Np(1-p) + p^2N^2, \\
\mathrm{Var}(k) &= \langle k^2 \rangle - \langle k \rangle^2 = Np(1-p), \\
\sigma(k) &= \sqrt{\mathrm{Var}(k)} = \sqrt{Np(1-p)}.
\end{aligned}$$

(For more details about the binomial distribution see e.g. `https://en.wikipedia.org/wiki/Binomial_distribution`)

Degree distribution of the E-R graph

Variance

- What happens in the $N \to \infty$ limit for a “realistic” E-R graph? (I.e., an E-R graph modeling a real system).

- What happens in the $N \to \infty$ limit for a "realistic" E-R graph? (I.e., an E-R graph modeling a real system).

It must remain sparse! $\rightarrow$ $\langle k \rangle$ =const.,

$$\left.\begin{array}{l} N \to \infty \\ \langle k \rangle = Np \to \text{const.} \end{array}\right\} \Rightarrow p \to 0$$

$$\mathrm{Var}(k) = Np(1-p) = \langle k \rangle (1-p)$$

$$\rightarrow \quad \mathrm{Var}(k) \to \langle k \rangle = \text{const.}$$

Degree distribution of the E-R graph

Variance

- What happens in the $N \to \infty$ limit for a “realistic” E-R graph? (I.e., an E-R graph modeling a real system).

It must remain sparse! $\rightarrow$ $\langle k \rangle$ =const.,

$$\left.\begin{array}{l} N \to \infty \\ \langle k \rangle = Np \to \text{const.} \end{array}\right\} \Rightarrow p \to 0$$

$$\mathrm{Var}(k) = Np(1-p) = \langle k \rangle (1-p)$$

$$\rightarrow \quad \mathrm{Var}(k) \to \langle k \rangle = \text{const.}$$

The variance is **constant**, thus, it becomes **negligible compared to the system size**!

Degree distribution of the E-R graph
Variance

- We have seen that HUBS are important in scale-free networks.

Degree distribution of the E-R graph
Variance

- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?

Degree distribution of the E-R graph

Variance

- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.

Degree distribution of the E-R graph

Variance

- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.
- OK, but how can we calculate the probability for finding 'extreme large' degrees, i.e., a degree larger than some k_0?

Degree distribution of the E-R graph
Variance

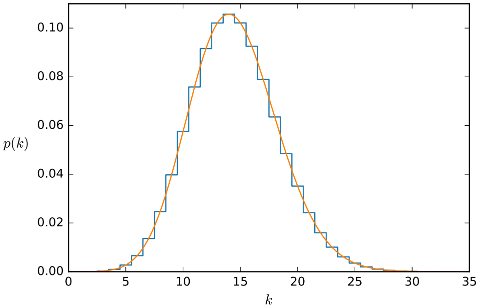
- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.
- OK, but how can we calculate the probability for finding 'extreme large' degrees, i.e., a degree larger than some k_0?

→ It is given by the sum (or integral) of $p(k)$ from k_0 to infinity.

Degree distribution of the E-R graph
Variance

- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.
- OK, but how can we calculate the probability for finding 'extreme large' degrees, i.e., a degree larger than some k_0?

→ It is given by the sum (or integral) of $p(k)$ from k_0 to infinity. Illustration:

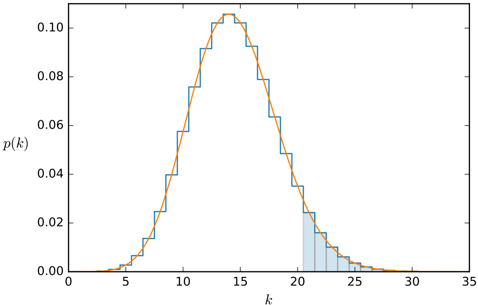


Degree distribution of the E-R graph

Variance

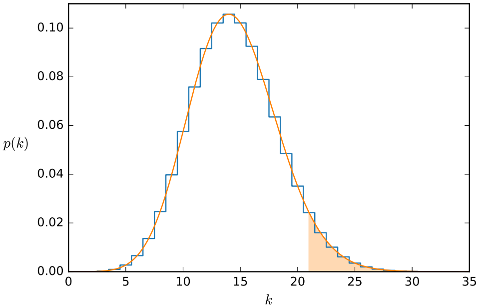
- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.
- OK, but how can we calculate the probability for finding 'extreme large' degrees, i.e., a degree larger than some k_0?

→ It is given by the sum (or integral) of $p(k)$ from k_0 to infinity. Illustration:



- We have seen that HUBS are important in scale-free networks.
- What about HUBS in an E-R graph?
- In short, there are no HUBS in an E-R graph; the degree of any node is close to the average degree.
- OK, but how can we calculate the probability for finding 'extreme large' degrees, i.e., a degree larger than some k_0?

→ It is given by the sum (or integral) of $p(k)$ from k_0 to infinity. Illustration:



Degree distribution of the E-R graph
Variance

- Thus, in mathematical terms, the probability for finding nodes with degrees above k_0 can be written as

$$\mathcal{P}(k > k_0) = \sum_{k=k_0+1}^{\infty} p(k) \simeq \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k\rangle^k}{k!} e^{-\langle k\rangle} dk.$$

- E.g, for $\langle k\rangle = 10$ by evaluating the integral we obtain:
  - the prob. to find a node with $k \geq 20$ is 0.00158826,
  - the prob. to find a node with $k \leq 1$ is 0.00049,
  - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

→ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

→ A random society would be extremely homogeneous, with no outliers!

Degree distribution of the E-R graph
Variance

- Thus, in mathematical terms, the probability for finding nodes with degrees above k_0 can be written as

$$\mathcal{P}(k > k_0) = \sum_{k=k_0+1}^{\infty} p(k) \simeq \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} dk.$$

- E.g, for $\langle k \rangle = 10$ by evaluating the integral we obtain:
  - the prob. to find a node with $k \geq 20$ is 0.00158826,
  - the prob. to find a node with $k \leq 1$ is 0.00049,
  - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

→ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

→ A random society would be extremely homogeneous, with no outliers!

Degree distribution of the E-R graph
Variance

- Thus, in mathematical terms, the probability for finding nodes with degrees above k_0 can be written as

$$\mathcal{P}(k > k_0) = \sum_{k=k_0+1}^{\infty} p(k) \simeq \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} dk.$$

- E.g, for $\langle k \rangle = 10$ by evaluating the integral we obtain:
  - the prob. to find a node with $k \geq 20$ is 0.00158826,
  - the prob. to find a node with $k \leq 1$ is 0.00049,
  - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

→ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

→ A random society would be extremely homogeneous, with no outliers!

Clustering coefficient in the E-R graph

- What is the clustering coefficient in the E-R graph?

Clustering coefficient in the E-R graph

- What is the clustering coefficient in the E-R graph?
  - The E-R graph is very democratic, and we expect all nodes to have more or less the same C, thus, $C_i \simeq \langle C \rangle$.
  - C_i can be also interpreted as the probability of the neighbors of i being connected. Since in the E-R model we link every pair independently with uniform probability p, the neighbors of any node shall be linked also with probability p.
  - Thus, in the E-R graph $\langle C \rangle = p$.

E-R graph vs real networks

How do the properties of the E-R model compare to that of real networks?

E-R graph vs real networks

How do the properties of the E-R model compare to that of real networks?

- We have seen that real networks are
  - **sparse**,
  - **small world**,
  - **highly clustered**,
  - **scale-free**.

E-R graph vs real networks

How do the properties of the E-R model compare to that of real networks?

- We have seen that real networks are
  - **sparse**,
  - **small world**,
  - **highly clustered**,
  - **scale-free**.

→ How good is the E-R model in reproducing these features?

E-R graph vs real networks

Sparseness?

Sparseness?

- For any fixed system size N we can tune the link density with the help of the parameter p. For example, the average degree is $\langle k \rangle = (N-1)p$.

Sparseness?

- For any fixed system size N we can tune the link density with the help of the parameter p. For example, the average degree is $\langle k \rangle = (N-1)p$.
- Thus, **the E-R can be made sparse or dense** at will with appropriate choice of the parameter p.

E-R graph vs real networks

Small world property?

E-R graph vs real networks

Small world property?

Comparison of the average distance in real and E-R networks:

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153,127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	225,226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998
LANL coauthorship	52,909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b
MEDLINE coauthorship	1,520,251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b
SPIRES coauthorship	56,627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c
NCSTRL coauthorship	11,994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b
Math coauthorship	70,975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001
Neurosci. coauthorship	209,293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001
Words, synonyms	22,311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001
Power grid	4,941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998

E-R graph vs real networks

Prediction for $\langle \ell \rangle$ in the E-R model:

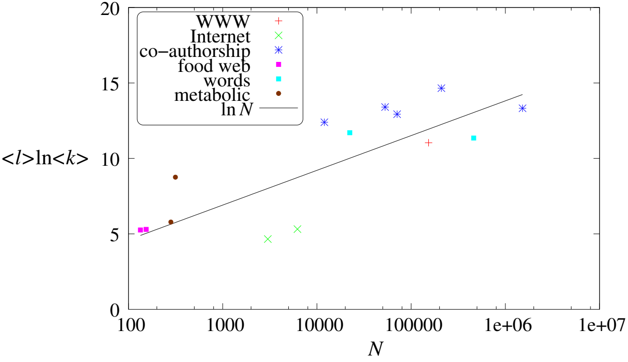
$$\langle \ell \rangle \simeq \frac{\ln N}{\ln \langle k \rangle} \quad \rightarrow \quad \langle \ell \rangle \ln \langle k \rangle \simeq \ln N$$

E-R graph vs real networks

Prediction for $\langle \ell \rangle$ in the E-R model:

$$\langle \ell \rangle \simeq \frac{\ln N}{\ln \langle k \rangle} \quad \rightarrow \quad \langle \ell \rangle \ln \langle k \rangle \simeq \ln N$$

Data from the table in the previous slide:

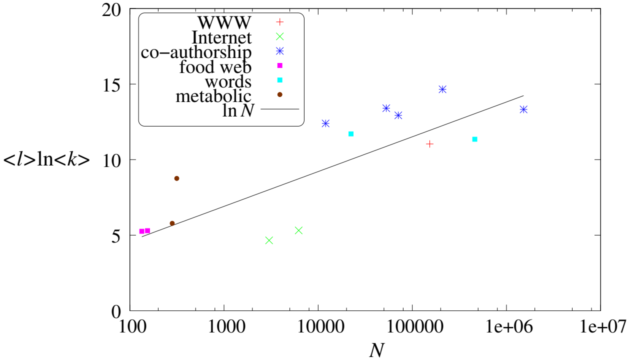


E-R graph vs real networks

Prediction for $\langle \ell \rangle$ in the E-R model:

$$\langle \ell \rangle \simeq \frac{\ln N}{\ln \langle k \rangle} \quad \rightarrow \quad \langle \ell \rangle \ln \langle k \rangle \simeq \ln N$$

Data from the table in the previous slide:



→ **The E-R is OK in reproducing the small world property.**

E-R graph vs real networks

High clustering?

E-R graph vs real networks

High clustering?

Comparison for C between real and E-R networks:

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153,127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	225,226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998
LANL coauthorship	52,909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b
MEDLINE coauthorship	1,520,251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b
SPIRES coauthorship	56,627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c
NCSTRL coauthorship	11,994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b
Math coauthorship	70,975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001
Neurosci. coauthorship	209,293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001
Words, synonyms	22,311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001
Power grid	4,941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998

E-R graph vs real networks

Prediction for $\langle C \rangle$ in the E-R model:

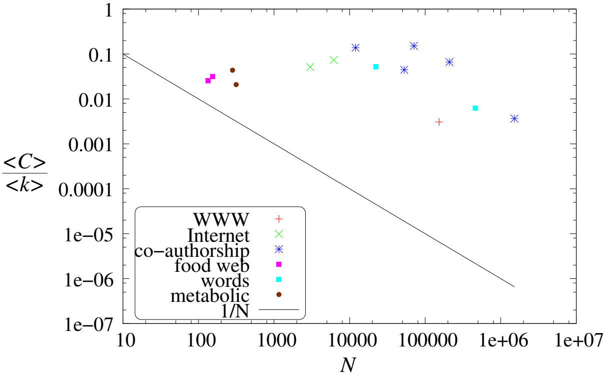
$$\langle C \rangle \simeq \frac{\langle k \rangle}{N} \quad \rightarrow \quad \frac{\langle C \rangle}{\langle k \rangle} \simeq \frac{1}{N}$$

E-R graph vs real networks

Prediction for $\langle C \rangle$ in the E-R model:

$$\langle C \rangle \simeq \frac{\langle k \rangle}{N} \quad \rightarrow \quad \frac{\langle C \rangle}{\langle k \rangle} \simeq \frac{1}{N}$$

Data from the table in the previous slide:

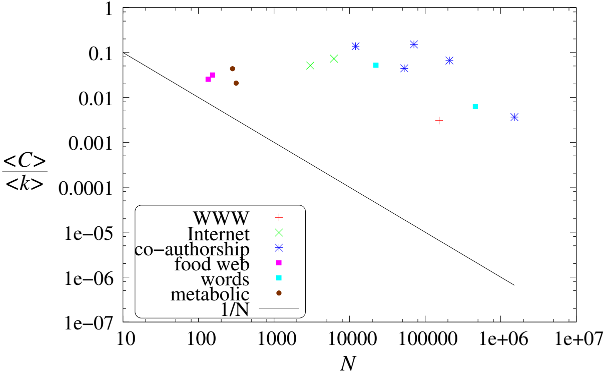


E-R graph vs real networks

Prediction for $\langle C \rangle$ in the E-R model:

$$\langle C \rangle \simeq \frac{\langle k \rangle}{N} \quad \rightarrow \quad \frac{\langle C \rangle}{\langle k \rangle} \simeq \frac{1}{N}$$

Data from the table in the previous slide:



→ **The E-R cannot reproduce the high clustering seen in real networks!**

E-R graph vs real networks

- What happens in the $N \to \infty$ limit if the E-R graph is sparse?

E-R graph vs real networks

- What happens in the $N \to \infty$ limit if the E-R graph is sparse?
- Since $\langle k \rangle = Np$, we have $p \to 0$.
- Since $\langle C \rangle = p$, also $\langle C \rangle \to 0$.

E-R graph vs real networks

Scale-free $p(k)$?

E-R graph vs real networks

Scale-free $p(k)$?

The degree distribution in the E-R model:

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k} \simeq \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}$$

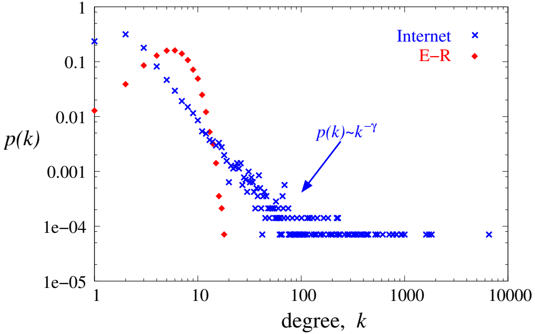
E-R graph vs real networks

Scale-free $p(k)$?

The degree distribution in the E-R model:

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k} \simeq \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}$$

In contrast, real data shows usually that $p(k) \sim k^{-\gamma}$.



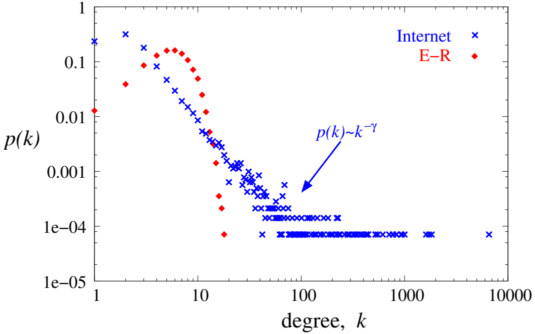
E-R graph vs real networks

Scale-free $p(k)$?

The degree distribution in the E-R model:

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k} \simeq \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}$$

In contrast, real data shows usually that $p(k) \sim k^{-\gamma}$.



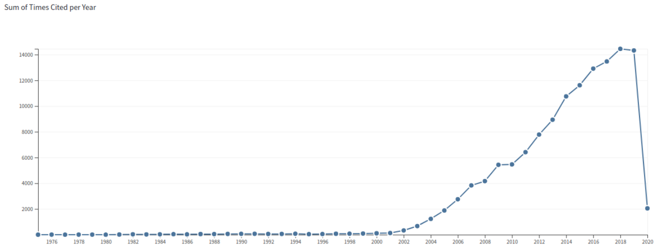
→ **The E-R model cannot reproduce the scale-free property of real networks!**

The E-R model compared to real networks:

- sparseness, $\langle k \rangle$?	**OK**
-the small-world effect, $\langle \ell \rangle$?	**OK**
-large local clustering coeff., $\langle C \rangle$?	**NO!**
-scale-free $p(k)$?	**NO!**

Why is the E-R model so important?

Yearly citations of the original Erdős-Rényi paper:



- Pure random graph: extremely democratic, extremely random.
- Analytic results.

→ Very important **reference** system.

Further reading on the Erdős-Rényi model (not compulsory):

- Network science book by A.-L. Barabási, chapter 3:
  `http://networksciencebook.com/chapter/3`
- Wikipedia:
  `https://en.wikipedia.org/wiki/Erdos-Renyi_model`

THE WATTS-STROGATZ MODEL

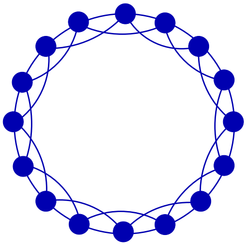
Duncan Watts and Steven H. Strogatz, (1998):

Can we have both the **small-world effect** and **local clustering** in a simple random graph model?

The Watts-Strogatz-model

The Watts-Strogatz-model(1998)

- Start from a regular ring of nodes in which the q first neighbors are linked.
- Rewire each link randomly with probability β.



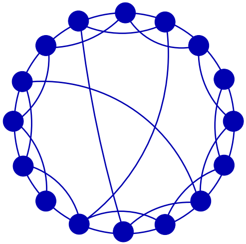
(The parameter q is equal to $q = 2$ in the illustration above).

D. Watts and S. H. Strogatz, *Nature* **393**,409 (1998)

The Watts-Strogatz-model

The Watts-Strogatz-model(1998)

- Start from a regular ring of nodes in which the q first neighbors are linked.
- Rewire each link randomly with probability β.



(The parameter q is equal to $q = 2$ in the illustration above).

D. Watts and S. H. Strogatz, *Nature* **393**,409 (1998)

We can calculate (or at least estimate) the **average distance** $\langle \ell \rangle$ and the **average clustering coefficient** $\langle C \rangle$ in the two extreme limits of the model, corresponding to $\beta = 0$ and $\beta = 1$, respectively.

The Watts-Strogatz model
Limiting cases

- **Let's start with $\beta = 0$, corresponding to no rewiring at all!**

- **Let's start with $\beta = 0$, corresponding to no rewiring at all!**
- In this case for any node, the most distant node is at $N/2$ steps if we walk along the perimeter of the ring. However, due to the extra links compared to a simple ring, we can actually pass by q nodes in one step, thus, the most distant node is only $N/(2q)$ steps away.

- **Let's start with $\beta = 0$, corresponding to no rewiring at all!**
- In this case for any node, the most distant node is at $N/2$ steps if we walk along the perimeter of the ring. However, due to the extra links compared to a simple ring, we can actually pass by q nodes in one step, thus, the most distant node is only $N/(2q)$ steps away.
- Due to the regular structure of the network, the average distance of all other nodes from any chosen node is going to equal to simply half of the largest distance, thus, the average distance is

$$\langle \ell \rangle = \frac{N}{4q}.$$

- **Let's start with $\beta = 0$, corresponding to no rewiring at all!**
- In this case for any node, the most distant node is at $N/2$ steps if we walk along the perimeter of the ring. However, due to the extra links compared to a simple ring, we can actually pass by q nodes in one step, thus, the most distant node is only $N/(2q)$ steps away.
- Due to the regular structure of the network, the average distance of all other nodes from any chosen node is going to equal to simply half of the largest distance, thus, the average distance is

$$\langle \ell \rangle = \frac{N}{4q}.$$

- The calculation of the clustering coefficient is more elaborate, and it is given in details in a separate document (also uploaded to Moodle), yielding

$$C = \frac{3q-3}{4q-2}.$$

The Watts-Strogatz model
Limiting cases

- **What about the $\beta = 1$ case, corresponding to rewiring all links at random?**

- **What about the $\beta = 1$ case, corresponding to rewiring all links at random?**
- The network after the rewiring is very much like an Erdős-Rényi random graph!

- **What about the $\beta = 1$ case, corresponding to rewiring all links at random?**
- The network after the rewiring is very much like an Erdős-Rényi random graph!

→ We can use our knowledge about the E-R graph to calculate $\langle \ell \rangle$ and $\langle C \rangle$.

The Watts-Strogatz model
Limiting cases

- **What about the $\beta = 1$ case, corresponding to rewiring all links at random?**
- The network after the rewiring is very much like an Erdős-Rényi random graph!

→ We can use our knowledge about the E-R graph to calculate $\langle \ell \rangle$ and $\langle C \rangle$.

- The average distance is simply

$$\langle \ell \rangle \simeq \frac{\ln N}{\langle k \rangle} = \frac{\ln N}{2q}.$$

- **What about the $\beta = 1$ case, corresponding to rewiring all links at random?**
- The network after the rewiring is very much like an Erdős-Rényi random graph!

→ We can use our knowledge about the E-R graph to calculate $\langle \ell \rangle$ and $\langle C \rangle$.

- The average distance is simply

$$\langle \ell \rangle \simeq \frac{\ln N}{\langle k \rangle} = \frac{\ln N}{2q}.$$

- The average clustering coefficient is equal to the p parameter of the E-R model, which we can obtain based on e.g., the average degree as $2q = \langle k \rangle = (N-1)p_{\rm ER}$. Thus, we find

$$\langle C \rangle = p_{\rm ER} = \frac{2q}{N-1}.$$

The Watts-Strogatz model
Limiting cases

In summary:

- $\beta = 0$:

$$\begin{aligned} \langle \ell \rangle &\simeq \frac{N}{4q} \\ \langle C \rangle = C &= \frac{q(q-1)\frac{3}{2}}{q(2q-1)} = \frac{3q-3}{4q-2} \end{aligned}$$

(A network with high clustering, but definitely not small world).

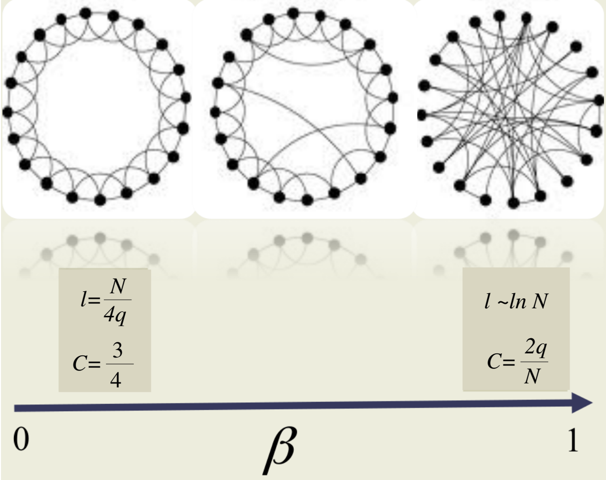
- $\beta = 1$: E-R model, G(N,M) version.

$$\begin{aligned} \langle \ell \rangle &\sim \log N \\ \langle C \rangle &= p_{\mathrm{ER}} = \frac{2q}{N-1} \end{aligned}$$

(A small world network that is definitely not highly clustered).

The Watts-Strogatz-model

Between limiting cases



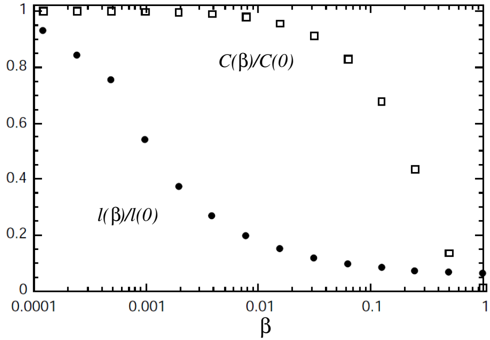
(from the slides of A.-L. Barabási)

Network models

E-R model
Definition
p(k) in the E-R model
C in the E-R model
E-R graph vs real networks

W-S model
The model
Limiting cases
Between limiting cases

Original figure from the paper by Watts and Strogatz:



- There is a wide range of β values where C is still relatively high, whereas ℓ is already relatively low!

$\rightarrow$ **HIGH CLUSTERING and SMALL WORLD!**

Why does it work?

Why does it work?

- It turns out that already a few 'long distance' random links can drastically reduce $\langle \ell \rangle$. Thus, a relatively small number of rewired links can turn the system into a small world network.

Why does it work?

- It turns out that already a few 'long distance' random links can drastically reduce $\langle \ell \rangle$. Thus, a relatively small number of rewired links can turn the system into a small world network.
- However, in order to destroy the large amount of triangles responsible for the high $\langle C \rangle$ in the initial state, we would need significantly more rewiring steps!

The Watts-Strogatz-model

Why does it work?

Take home message of the W-S model:

It takes a lot of randomness to ruin the clustering, but a very small amount to overcome locality.

When does the small-world effect take place?

- Assuming that we would like to generate a network that is already small world but in the mean time, as highly clustered as possible, what would be the optimal value for β?

When does the small-world effect take place?

- Assuming that we would like to generate a network that is already small world but in the mean time, as highly clustered as possible, what would be the optimal value for β?
- If β is too small, the network is not small-world...

When does the small-world effect take place?

- Assuming that we would like to generate a network that is already small world but in the mean time, as highly clustered as possible, what would be the optimal value for β?
- If β is too small, the network is not small-world...
- If β is too large, the network is small-world, but the clustering coefficient is also reduced due to the rewirings...

When does the small-world effect take place?

- Assuming that we would like to generate a network that is already small world but in the mean time, as highly clustered as possible, what would be the optimal value for β?
- If β is too small, the network is not small-world...
- If β is too large, the network is small-world, but the clustering coefficient is also reduced due to the rewirings...
- In the following we give a simple estimate for the lowest possible β, where we can expect the network to be already small-world.

When does the small-world effect take place?

- Number of random links?

When does the small-world effect take place?

- Number of random links: $\beta q N$.

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- what if $\beta q N << 1$?

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.
- what if $\beta q N >> 1$?

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.
- if $\beta q N >> 1$ then many random links and $\langle \ell \rangle \sim \ln N$.

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.
- if $\beta q N >> 1$ then many random links and $\langle \ell \rangle \sim \ln N$.
- simple guess for the transition point?

When does the small-world effect take place?

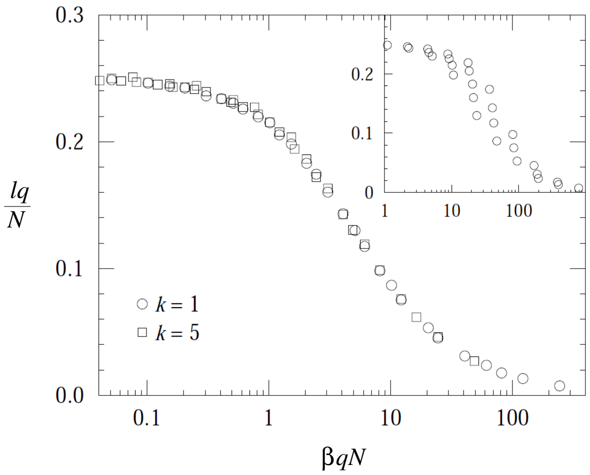
- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.
- if $\beta q N >> 1$ then many random links and $\langle \ell \rangle \sim \ln N$.
- the transition point is somewhere around $\beta_c q N = 1$.

When does the small-world effect take place?

- Number of random links: $\beta q N$.
- if $\beta q N << 1$ then no random links and $\langle \ell \rangle \sim N$.
- if $\beta q N >> 1$ then many random links and $\langle \ell \rangle \sim \ln N$.
- the transition point is somewhere around $\beta_c q N = 1$.

→ It can be derived analytically that
  - The transition is really at $\beta_c q N = 1$,
  - and the $\ell(\beta)$-curve for all W-S graphs can be collapsed into a single curve via
$$\ell = \frac{N}{q} f(\beta q N),$$
where $f(x)$ is a universal function.

The Watts-Strogatz model
Data collapse



M. E. J. Newman and D. J. Watts, *Physics Letters A* **263**, 341–346 (1999)

Further reading about the Watts-Strogatz model (non compulsory):

- Wikipedia: `https://en.wikipedia.org/wiki/Watts-Strogatz_model`
- Network science book by A.-L. Barabási, Chapter 3, Box 3.9: `http://networksciencebook.com/chapter/3#clustering-3-9`