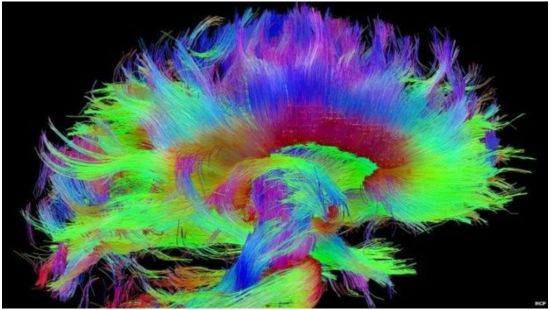
CS 4884: Erdös-Renyi and Small World Networks

T. M. Murali

February 3 and 8, 2022



Collective dynamics of 'small-world' networks

Duncan J. Watts & Steven H. Strogatz

Nature **393**, 440–442 (04 June 1998)
doi:10.1038/30918
Download Citation
Received: 27 November 1997
Accepted: 06 April 1998
Published online: 04 June 1998

Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs.

Specifically, we require $n \gg k \gg \ln(n) \gg 1$, *where* $k \gg \ln(n)$ *guarantees that a random graph will be connected.*

Collective dynamics of 'small-world' networks

Duncan J. Watts & Steven H. Strogatz

Nature **393**, 440–442 (04 June 1998)
doi:10.1038/30918
Download Citation
Received: 27 November 1997
Accepted: 06 April 1998
Published online: 04 June 1998

Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs.

Specifically, we require $n \gg k \gg \ln(n) \gg 1$, *where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*

Random Graps

- What is a random graph?

Random Graps

- What is a random graph?
- How do we create a random unweighted, undirected graph on n nodes?

Random Graps

- What is a random graph?
- How do we create a random unweighted, undirected graph on n nodes?
- Question is under-specified. There are many approaches:

Random Graps

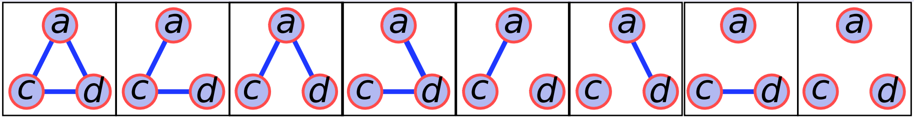
- What is a random graph?
- How do we create a random unweighted, undirected graph on n nodes?
- Question is under-specified. There are many approaches:
  1. Idea 1: From the set of all graphs of n nodes, pick one uniformly at random.
  2. Idea 2: Specify the number of edges m. From the set of all graphs of n nodes and m edges, pick one uniformly at random.
  3. Idea 3: Specify a probability $0 \leq p \leq 1$. For every pair of nodes, add an edge between the nodes with probability p.

Idea 1 for Creating Random Graphs

From the set of all graphs of n nodes, pick one uniformly at random.

Idea 1 for Creating Random Graphs

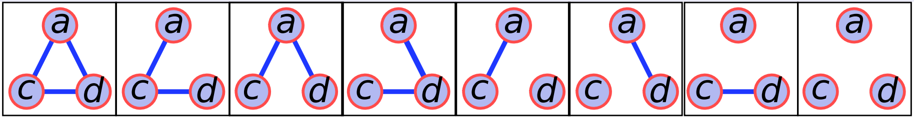
From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?

Idea 1 for Creating Random Graphs

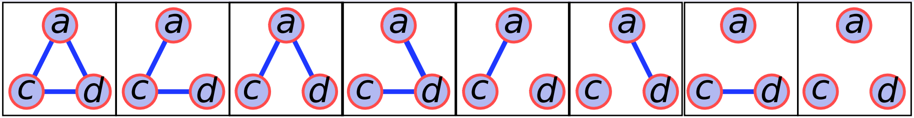
From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?
  - To make a graph, we have two options for each edge: include it or exclude it.
  - Therefore, there are $2^{\binom{n}{2}}$ graphs possible on n nodes.

Idea 1 for Creating Random Graphs

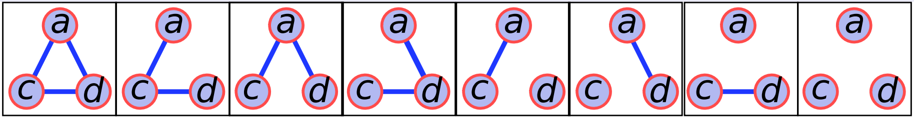
From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?
  - To make a graph, we have two options for each edge: include it or exclude it.
  - Therefore, there are $2^{\binom{n}{2}}$ graphs possible on n nodes.
- How do we implement Idea 1? How do we select one of these graphs with probability $\frac{1}{2^{\binom{n}{2}}}$?

Idea 1 for Creating Random Graphs

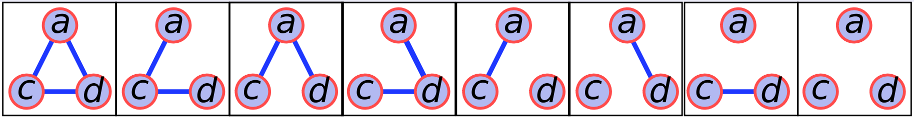
From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?
  - To make a graph, we have two options for each edge: include it or exclude it.
  - Therefore, there are $2^{\binom{n}{2}}$ graphs possible on n nodes.
- How do we implement Idea 1? How do we select one of these graphs with probability $\frac{1}{2^{\binom{n}{2}}}$?
- Explicitly construct all $2^{\binom{n}{2}}$ and then select one uniformly at random.

Idea 1 for Creating Random Graphs

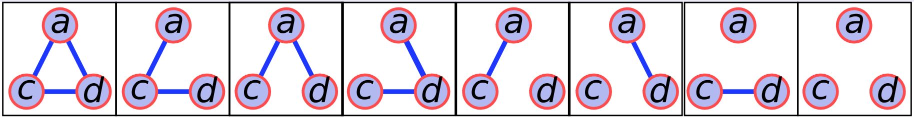
From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?
  - To make a graph, we have two options for each edge: include it or exclude it.
  - Therefore, there are $2^{\binom{n}{2}}$ graphs possible on n nodes.
- How do we implement Idea 1? How do we select one of these graphs with probability $\frac{1}{2^{\binom{n}{2}}}$?
- Explicitly construct all $2^{\binom{n}{2}}$ and then select one uniformly at random. Running time is $O(n^2 2^{\binom{n}{2}})$. Too slow!

Idea 1 for Creating Random Graphs

From the set of all graphs of n nodes, pick one uniformly at random.



- How many graphs can there be on n nodes?
  - To make a graph, we have two options for each edge: include it or exclude it.
  - Therefore, there are $2^{\binom{n}{2}}$ graphs possible on n nodes.
- How do we implement Idea 1? How do we select one of these graphs with probability $\frac{1}{2^{\binom{n}{2}}}$?
- Explicitly construct all $2^{\binom{n}{2}}$ and then select one uniformly at random. Running time is $O(n^2 2^{\binom{n}{2}})$. Too slow!
- For every pair of nodes, add an edge with probability $1/2$. Running time is $O(n^2)$.

Properties of Random Graphs Created by Idea 1

From the set of all graphs of n nodes, pick one uniformly at random.

- What is the expected degree of a node?

Properties of Random Graphs Created by Idea 1

From the set of all graphs of n nodes, pick one uniformly at random.

- What is the expected degree of a node? $(n-1)/2$.

Properties of Random Graphs Created by Idea 1

From the set of all graphs of n nodes, pick one uniformly at random.

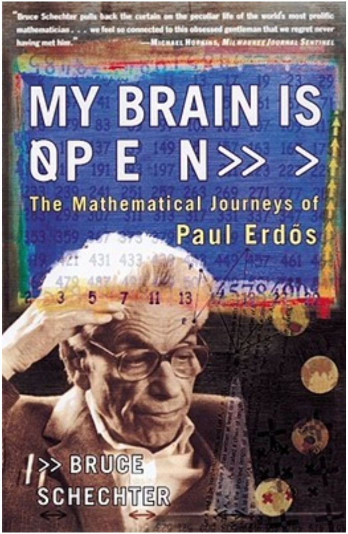
- What is the expected degree of a node? $(n-1)/2$.
- What is the expected number of edges in the graph?

Properties of Random Graphs Created by Idea 1

From the set of all graphs of n nodes, pick one uniformly at random.

- What is the expected degree of a node? $(n-1)/2$.
- What is the expected number of edges in the graph? $n(n-1)/4$.
- On average, these graphs are very dense.

Erdős-Rényi Graphs



A mathematician is a device for turning coffee into theorems.

Erdős-Rényi Graphs

Idea 3: Specify a probability $0 \leq p \leq 1$.
For every pair of nodes, add an edge between the nodes with probability p.

- Series of papers in the 1960s setting the foundation of random graph theory.
- Framework for generating a random graph.
- $G(n, p)$: an undirected, unweighted graph (family) with n nodes.
- To generate a graph in $G(n, p)$:
  - For each pair (u, v) of $\binom{n}{2}$ node pairs, connect u and v by an edge with probability p.

On the evolution of random graphs, P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci. 5, 17–61, 1960.

Erdős-Rényi Graphs

Idea 3: Specify a probability $0 \leq p \leq 1$.
For every pair of nodes, add an edge between the nodes with probability p.

- Series of papers in the 1960s setting the foundation of random graph theory.
- Framework for generating a random graph.
- $G(n, p)$: an undirected, unweighted graph (family) with n nodes.
- To generate a graph in $G(n, p)$:
  - For each pair (u, v) of $\binom{n}{2}$ node pairs, connect u and v by an edge with probability p.
  - How do you "do something" with probability p?

On the evolution of random graphs, P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci. 5, 17–61, 1960.

Erdős-Rényi Graphs

Idea 3: Specify a probability $0 \leq p \leq 1$.
For every pair of nodes, add an edge between the nodes with probability p.

- Series of papers in the 1960s setting the foundation of random graph theory.
- Framework for generating a random graph.
- $G(n, p)$: an undirected, unweighted graph (family) with n nodes.
- To generate a graph in $G(n, p)$:
  - For each pair (u, v) of $\binom{n}{2}$ node pairs, connect u and v by an edge with probability p.
  - How do you "do something" with probability p?
  - Generate a random number x between 0 and 1 under the uniform distribution. If $x \leq p$, then "do something", else "do the other thing".

On the evolution of random graphs, P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci. 5, 17–61, 1960.

Degrees and Connectivity in Erdős-Rényi Graphs

- To generate a graph in $G(n, p)$: For each pair (u, v) of $\binom{n}{2}$ nodes, connect u and v by an edge with probability p.
- How many edges does this graph have on average?

Degrees and Connectivity in Erdős-Rényi Graphs

- To generate a graph in $G(n, p)$: For each pair (u, v) of $\binom{n}{2}$ nodes, connect u and v by an edge with probability p.
- How many edges does this graph have on average? $n(n-1)p/2$.
- What is the expected degree of a node?

Degrees and Connectivity in Erdős-Rényi Graphs

- To generate a graph in $G(n, p)$: For each pair (u, v) of $\binom{n}{2}$ nodes, connect u and v by an edge with probability p.
- How many edges does this graph have on average? $n(n-1)p/2$.
- What is the expected degree of a node? $(n-1)p$.

Degree Distribution

- What is the degree distribution of $G(n, p)$?

Degree Distribution

- What is the degree distribution of $G(n, p)$?
- What is the probability that a node v has degree k?

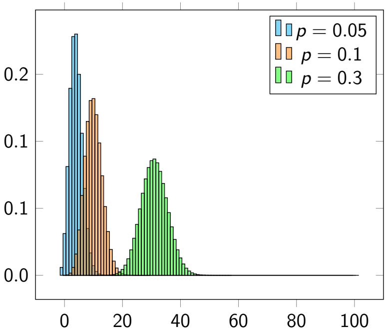
Degree Distribution

- What is the degree distribution of $G(n, p)$?
- What is the probability that a node v has degree k?
  - Connect (probability p) v to k neighbours out of $n-1$ nodes and not connect (probability $1-p$) to the rest.
  - Probability that v has degree k follows the binomial distribution
    $$\Pr(d(v) = k) = \binom{n-1}{k} p^k (1-p)^{n-k-1}$$

Degree Distribution

- What is the degree distribution of $G(n, p)$?
- What is the probability that a node v has degree k?
  - Connect (probability p) v to k neighbours out of $n-1$ nodes and not connect (probability $1-p$) to the rest.
  - Probability that v has degree k follows the binomial distribution

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



Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} =$$

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

$$\big(p + (1-p)\big)^{n-1} = 1$$

- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

$$\big(p + (1-p)\big)^{n-1} = 1$$

- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

$$\sum_{k=0}^{n} k \binom{n-1}{k} p^k (1-p)^{n-k-1} =$$

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

$$\big(p + (1-p)\big)^{n-1} = 1$$

- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

$$\sum_{k=0}^{n} k \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} k \Pr(d(v) = k)$$

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

$$\big(p + (1-p)\big)^{n-1} = 1$$

- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

$$\sum_{k=0}^{n} k \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} k \Pr(d(v) = k)$$

$$E[d(v)]$$

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$

$$\big(p + (1-p)\big)^{n-1} = 1$$

- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

$$\sum_{k=0}^{n} k \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} k \Pr(d(v) = k)$$

$$(n-1)p = E[d(v)]$$

- The expected degree of a node is $(n-1)p$.

Binomial Identities

$$\sum_{k=0}^{n} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} \Pr(d(v) = k)$$
$$\big(p + (1-p)\big)^{n-1} = 1$$

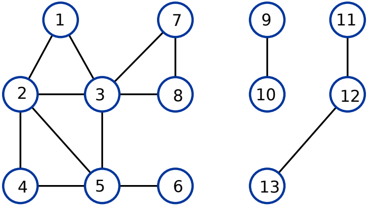
- Simply stating the degree of a node must take exactly one value between 0 and $n-1$.

$$\sum_{k=0}^{n} k \binom{n-1}{k} p^k (1-p)^{n-k-1} = \sum_{k=0}^{n-1} k \Pr(d(v) = k)$$
$$(n-1)p = E[d(v)]$$

- The expected degree of a node is $(n-1)p$.
- The expected number of edges in $G(n,p)$ is $n(n-1)p/2$.

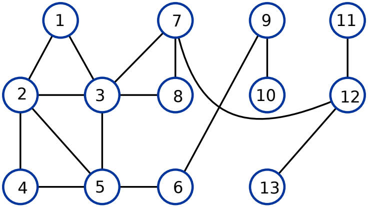
Guarantee that a Random Graph is Connected

Specifically, we require $n \gg k \gg \ln(n) \gg 1$, *where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*



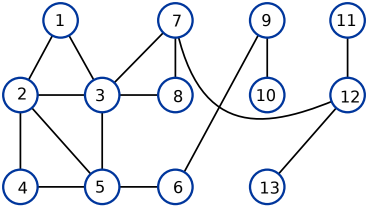
Guarantee that a Random Graph is Connected

Specifically, we require $n \gg k \gg \ln(n) \gg 1$*, where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*



Guarantee that a Random Graph is Connected

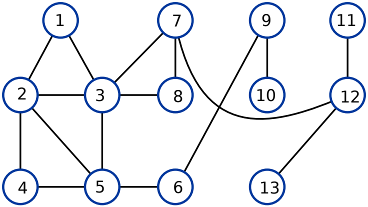
Specifically, we require $n \gg k \gg \ln(n) \gg 1$, *where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*



- Consider the evolution of $G(n, p)$ as p increases.

Guarantee that a Random Graph is Connected

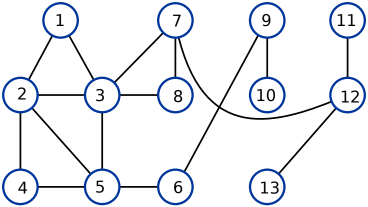
Specifically, we require $n \gg k \gg \ln(n) \gg 1$*, where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*



- Consider the evolution of $G(n, p)$ as p increases.
- When p is close to 0, graph has many small connected components.
- When p is close to 1, graph is very dense (has almost all the edges).
- When do all nodes in the graph become connected into one component?

Guarantee that a Random Graph is Connected

Specifically, we require $n \gg k \gg \ln(n) \gg 1$, *where* $k \gg ln(n)$ *guarantees that a random graph will be connected.*



- Consider the evolution of $G(n, p)$ as p increases.
- When p is close to 0, graph has many small connected components.
- When p is close to 1, graph is very dense (has almost all the edges).
- When do all nodes in the graph become connected into one component?

The evolution of the $G(n, p)$ random graph (Video, 4 min 51 sec)

Phase Transitions

Value of p	Property of $G(n, p)$
$p = 0$	
$p < \frac{(1-\varepsilon)}{n}$ $p > \frac{(1+\varepsilon)}{n}$	
$p < \frac{(1-\varepsilon)\ln n}{n}$ $p > \frac{(1+\varepsilon)\ln n}{n}$	.
$p = 1$	

Phase Transitions

Value of p	Property of $G(n, p)$
$p = 0$	Has no edges
$p < \frac{(1-\varepsilon)}{n}$ $p > \frac{(1+\varepsilon)}{n}$	
$p < \frac{(1-\varepsilon)\ln n}{n}$ $p > \frac{(1+\varepsilon)\ln n}{n}$	.
$p = 1$	Is a complete graph.

Phase Transitions

Value of p	**Property of** $G(n, p)$
$p = 0$	Has no edges
$p < \frac{(1-\varepsilon)}{n}$	All connected components are of size $\log n$.
$p > \frac{(1+\varepsilon)}{n}$	Has a unique connected component containing a positive fraction of the nodes (*giant component*)!
$p < \frac{(1-\varepsilon)\ln n}{n}$	
$p > \frac{(1+\varepsilon)\ln n}{n}$	.
$p = 1$	Is a complete graph.

Phase Transitions

Value of p	Property of $G(n, p)$
$p = 0$	Has no edges
$p < \frac{(1-\varepsilon)}{n}$	All connected components are of size $\log n$.
$p > \frac{(1+\varepsilon)}{n}$	Has a unique connected component containing a positive fraction of the nodes (*giant component*)!
$p < \frac{(1-\varepsilon)\ln n}{n}$	Has at least one isolated node.
$p > \frac{(1+\varepsilon)\ln n}{n}$	Is connected! k in Watts-Strogatz is np.
$p = 1$	Is a complete graph.

Phase Transitions

Value of p	**Property of $G(n, p)$**
$p = 0$	Has no edges
$p < \frac{(1-\varepsilon)}{n}$	All connected components are of size $\log n$.
$p > \frac{(1+\varepsilon)}{n}$	Has a unique connected component containing a positive fraction of the nodes (*giant component*)!
$p < \frac{(1-\varepsilon)\ln n}{n}$	Has at least one isolated node.
$p > \frac{(1+\varepsilon)\ln n}{n}$	Is connected! k in Watts-Strogatz is np. The average shortest path length is $\frac{\ln n}{\ln(1+\varepsilon)+\ln\ln n}$. Path lengths are logarithmic in the number of nodes!
$p = 1$	Is a complete graph.

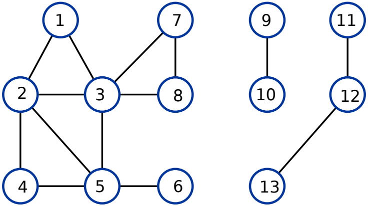
Phase Transitions

Value of p	**Property of $G(n, p)$**
$p = 0$	Has no edges
$p < \frac{(1-\varepsilon)}{n}$	All connected components are of size $\log n$.
$p > \frac{(1+\varepsilon)}{n}$	Has a unique connected component containing a positive fraction of the nodes (*giant component*)!
$p < \frac{(1-\varepsilon)\ln n}{n}$	Has at least one isolated node.
$p > \frac{(1+\varepsilon)\ln n}{n}$	Is connected! k in Watts-Strogatz is np. The average shortest path length is $\frac{\ln n}{\ln(1+\varepsilon)+\ln\ln n}$. Path lengths are logarithmic in the number of nodes!
$p = 1$	Is a complete graph.

Statements hold with high probability, e.g., if $p > \frac{(1+\varepsilon)\ln n}{n}$, then

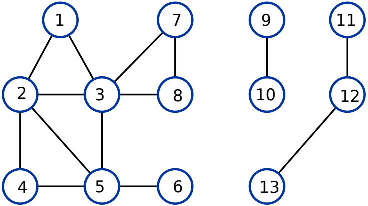
$$\Pr\{G(n, p) \text{ is not connected }\} \approx \frac{1}{e^{n^{\varepsilon}}}.$$

Clustering Coefficient



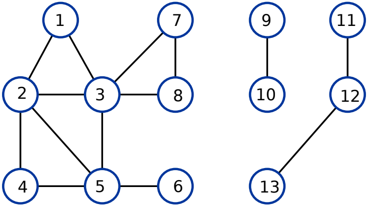
- Measures the extent of clusters/cliques around a node, on average.
- *Clustering coefficient $c(v)$* for a node v is the fraction of pairs of its neighbours that are themselves connected.
- *Clustering coefficient $c(G)$* of a graph G is the average of the clustering coefficients of its nodes.
  - Note that I am using lowercase c (since c is a number), whereas the paper uses uppercase C.
  - Technically, $c(v)$ should have the graph G as an argument, but we will be sloppy and ignore it.

Clustering Coefficient



- Measures the extent of clusters/cliques around a node, on average.
- *Clustering coefficient* $c(v)$ for a node v is the fraction of pairs of its neighbours that are themselves connected.
- *Clustering coefficient* $c(G)$ of a graph G is the average of the clustering coefficients of its nodes.
  - Note that I am using lowercase c (since c is a number), whereas the paper uses uppercase C.
  - Technically, $c(v)$ should have the graph G as an argument, but we will be sloppy and ignore it.
- What is the clustering coefficient of a lattice? A complete graph?

Clustering Coefficient



- Measures the extent of clusters/cliques around a node, on average.
- *Clustering coefficient* $c(v)$ for a node v is the fraction of pairs of its neighbours that are themselves connected.
- *Clustering coefficient* $c(G)$ of a graph G is the average of the clustering coefficients of its nodes.
  - Note that I am using lowercase c (since c is a number), whereas the paper uses uppercase C.
  - Technically, $c(v)$ should have the graph G as an argument, but we will be sloppy and ignore it.
- What is the clustering coefficient of a lattice? A complete graph? 0 and 1, respectively.

Clustering Coefficient of an Erdős-Rényi Graph

- Assume $p > \frac{(1+\varepsilon)\ln n}{n}$.
- We know that the average shortest path length in $G(n, p)$ is $\approx \frac{\ln n}{\ln(1+\varepsilon)}$.

Clustering Coefficient of an Erdős-Rényi Graph

- Assume $p > \frac{(1+\varepsilon)\ln n}{n}$.
- We know that the average shortest path length in $G(n, p)$ is $\approx \frac{\ln n}{\ln(1+\varepsilon)}$.
- What is the clustering coefficient of $G(n, p)$?

Clustering Coefficient of an Erdős-Rényi Graph

- Assume $p > \frac{(1+\varepsilon)\ln n}{n}$.
- We know that the average shortest path length in $G(n, p)$ is $\approx \frac{\ln n}{\ln(1+\varepsilon)}$.
- What is the clustering coefficient of $G(n, p)$?
  - A node u has $(n-1)p$ nodes on average.
  - What is the probability that two neighbours v and w are connected?

Clustering Coefficient of an Erdős-Rényi Graph

- Assume $p > \frac{(1+\varepsilon)\ln n}{n}$.
- We know that the average shortest path length in $G(n, p)$ is $\approx \frac{\ln n}{\ln(1+\varepsilon)}$.
- What is the clustering coefficient of $G(n, p)$?
  - A node u has $(n-1)p$ nodes on average.
  - What is the probability that two neighbours v and w are connected? p!
  - Hence, the clustering coefficient of $G(n, p)$ is $p < 1$.

Milgram's Experiment

It's a small world! (Video, 1 min 36 sec)

Milgram's Experiment

It's a small world! (Video, 1 min 36 sec)

Criticisms

- Overestimates path lengths.
- Underestimates path lengths.

Milgram's Experiment

It's a small world! (Video, 1 min 36 sec)

Criticisms

- Overestimates path lengths.
- Underestimates path lengths.

Conclusions. Which is correct?

- Some paths in social networks are short.
- All paths between all pairs of nodes are short.
- The average shortest path length is small. Average taken over all pairs of nodes.

Milgram's Experiment

It's a small world! (Video, 1 min 36 sec)

Criticisms

- Overestimates path lengths.
- Underestimates path lengths.

Conclusions. Which is correct?

- Some paths in social networks are short.
- All paths between all pairs of nodes are short.
- The average shortest path length is small. Average taken over all pairs of nodes.

Burning question

How do networks with small average shortest path length arise?

Motivation

- Consider two measures for a graph G:
  - $l(G)$, the average shortest path length in G.
  - $c(G)$, the clustering coefficient of G.

Motivation

- Consider two measures for a graph G:
  - $l(G)$, the average shortest path length in G.
  - $c(G)$, the clustering coefficient of G.
- $G(n, p)$, $p > \frac{(1+\varepsilon)\ln n}{n}$:

$$l(G) = \frac{\ln n}{\ln np} \text{ (small)} \qquad c(G) = p \text{ (small)}$$

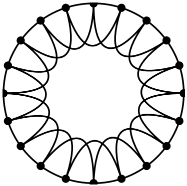
Motivation

- Consider two measures for a graph G:
  - $l(G)$, the average shortest path length in G.
  - $c(G)$, the clustering coefficient of G.
- $G(n, p)$, $p > \frac{(1+\varepsilon)\ln n}{n}$:

$$l(G) = \frac{\ln n}{\ln np} \text{ (small)} \qquad c(G) = p \text{ (small)}$$

- *Regular ring graph*: n nodes in a ring, each node connected to the next $k/2$ nodes appearing in clockwise order around the ring.

$$l(G) = \qquad\qquad c(G) =$$



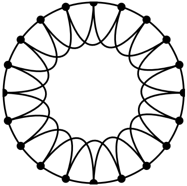
Motivation

- Consider two measures for a graph G:
  - $l(G)$, the average shortest path length in G.
  - $c(G)$, the clustering coefficient of G.
- $G(n, p)$, $p > \frac{(1+\varepsilon)\ln n}{n}$:

$$l(G) = \frac{\ln n}{\ln np} \text{ (small)} \qquad c(G) = p \text{ (small)}$$

- *Regular ring graph*: n nodes in a ring, each node connected to the next $k/2$ nodes appearing in clockwise order around the ring.

$$l(G) = n/2k \text{ (large)} \qquad c(G) =$$



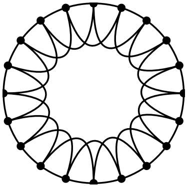
Motivation

- Consider two measures for a graph G:
  - $l(G)$, the average shortest path length in G.
  - $c(G)$, the clustering coefficient of G.
- $G(n, p)$, $p > \frac{(1+\varepsilon)\ln n}{n}$:

$$l(G) = \frac{\ln n}{\ln np} \text{ (small)} \qquad c(G) = p \text{ (small)}$$

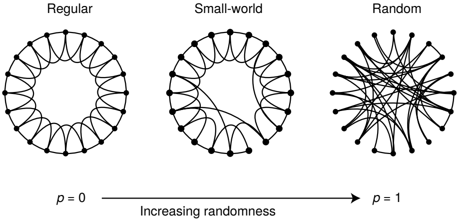
- *Regular ring graph*: n nodes in a ring, each node connected to the next $k/2$ nodes appearing in clockwise order around the ring.

$$l(G) = n/2k \text{ (large)} \qquad c(G) = \approx 3/4 \text{ (large)}$$



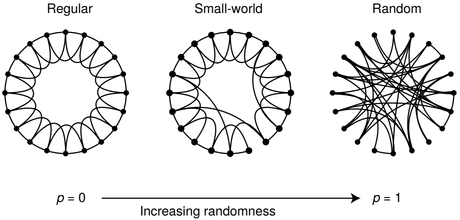
- Real world networks have small average shortest path lengths (like $G(n, p)$) but large clustering coefficients (like ring graph).

Watts-Strogatz Model



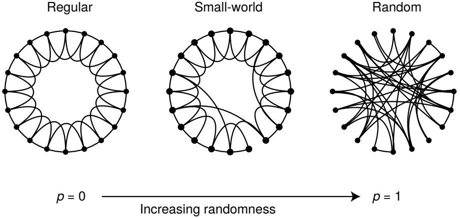
- Three parameters: n, number of nodes; k: degree of each node; p: rewiring probability. This p is different from the p in E-R graphs.
- Rewire regular ring graph in $k/2$ rounds. In round j,
  1. For each node i, consider edge $(i, i+j)$.
  2. Pick a candidate node l uniformly at random between 1 and n.
  3. With probability p, replace $(i, i+j)$ with (i, l) if

Watts-Strogatz Model



- Three parameters: n, number of nodes; k: degree of each node; p: rewiring probability. This p is different from the p in E-R graphs.
- Rewire regular ring graph in $k/2$ rounds. In round j,
  1. For each node i, consider edge $(i, i+j)$.
  2. Pick a candidate node l uniformly at random between 1 and n.
  3. With probability p, replace $(i, i+j)$ with (i, l) if $i \neq l$ and (i, l) not already in graph.

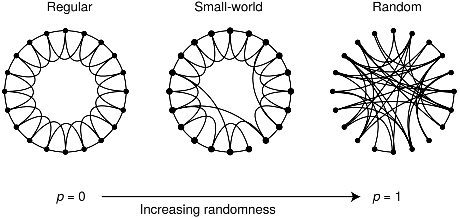
l and c for Watts-Strogatz Graphs



$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) =$$

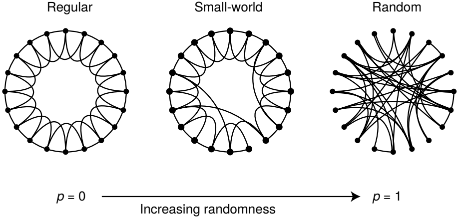
l and c for Watts-Strogatz Graphs



$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) =$$

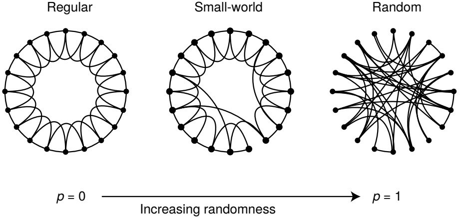
l and c for Watts-Strogatz Graphs



$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) = \approx 3/4$$

l and c for Watts-Strogatz Graphs



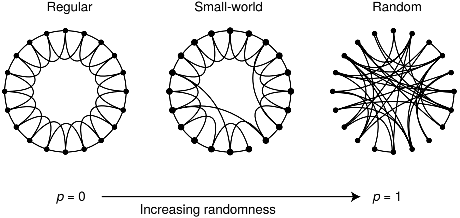
$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) = \approx 3/4$$

Ring lattice is large-world and highly clustered.

$$l(1) =$$

l and c for Watts-Strogatz Graphs



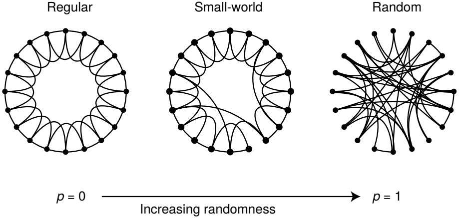
$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) = \approx 3/4$$

Ring lattice is large-world and highly clustered.

$$l(1) = \ln n / \ln k \qquad c(1) =$$

l and c for Watts-Strogatz Graphs



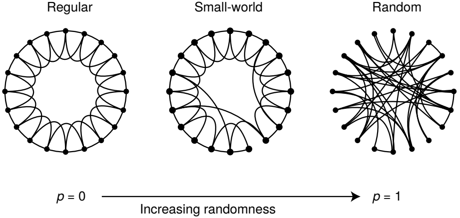
$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) = \approx 3/4$$

Ring lattice is large-world and highly clustered.

$$l(1) = \ln n / \ln k \qquad c(1) = k/n$$

l and c for Watts-Strogatz Graphs



$l(p)$: average shortest path length for ring graph rewired with prob. p.
$c(p)$: average clustering coefficient for ring graph rewired with prob. p.

$$l(0) = n/2k \qquad c(0) = \approx 3/4$$

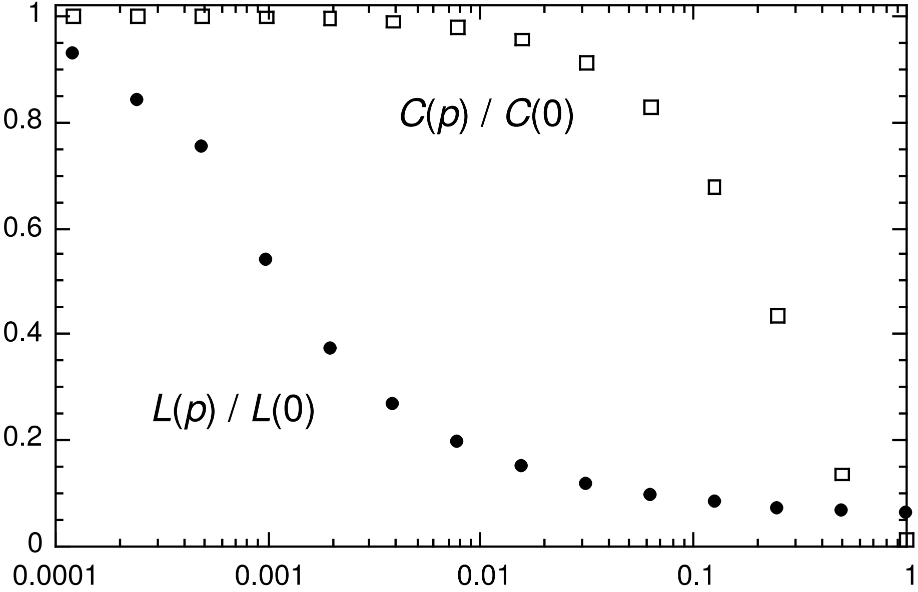
Ring lattice is large-world and highly clustered.

$$l(1) = \ln n / \ln k \qquad c(1) = k/n$$

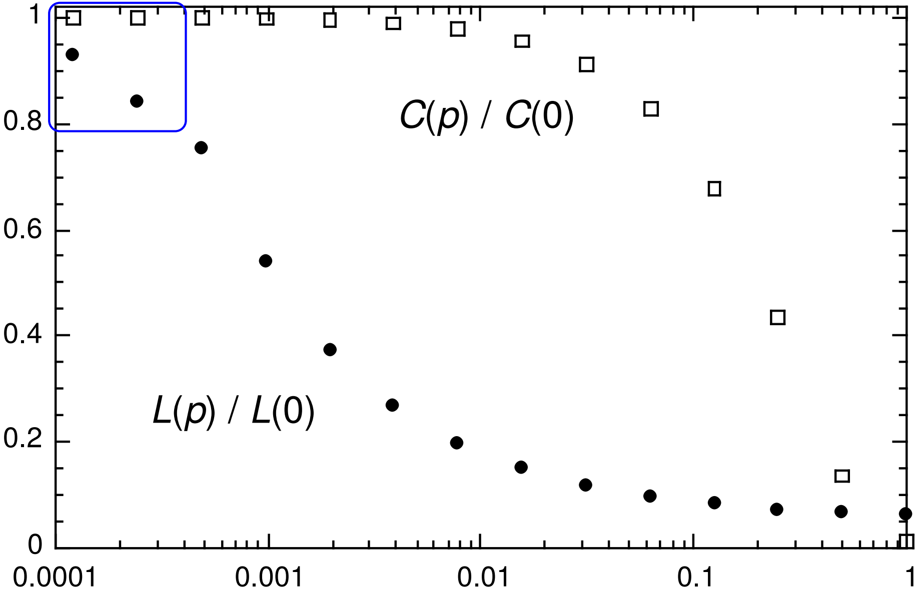
Random ring graph is small-world but poorly clustered.

Are there values of p for which $l(p)$ is small but $c(p)$ is large?

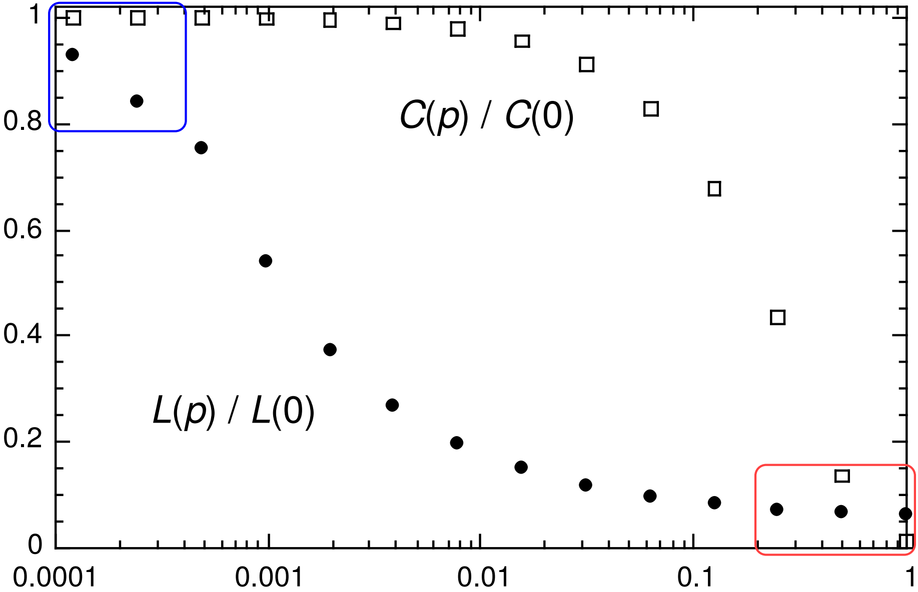
l and c for Watts-Strogatz Graphs



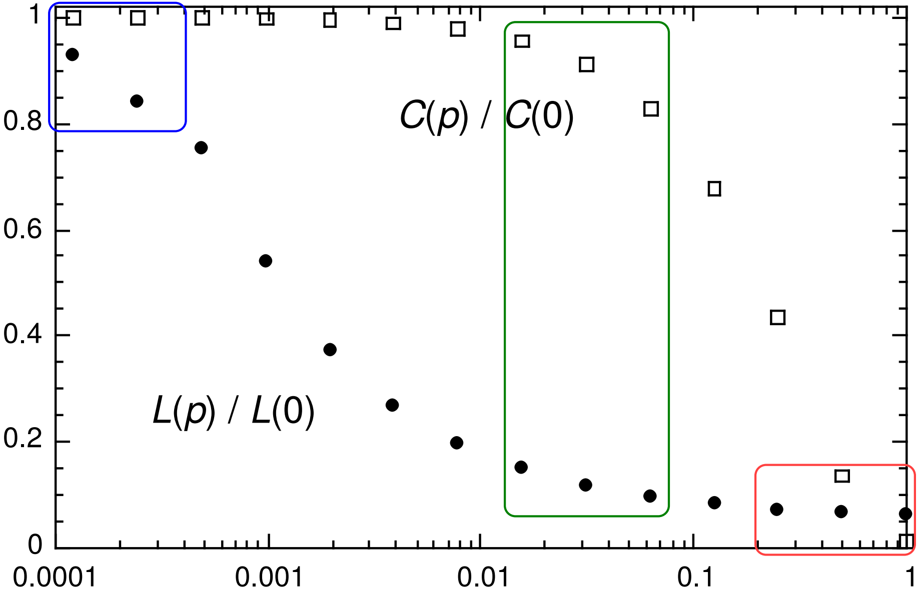
l and c for Watts-Strogatz Graphs



l and c for Watts-Strogatz Graphs



l and c for Watts-Strogatz Graphs



Observations

- $l(p)$ becomes small due to the addition of a small number of "long-range" edges.
- These short cuts connect nodes that would otherwise be very far apart.
- Non-linear effect on $l(p)$: Short cuts also contract the distance between neighbours of the connected nodes, their neighbours, and so on.

Observations

- $l(p)$ becomes small due to the addition of a small number of "long-range" edges.
- These short cuts connect nodes that would otherwise be very far apart.
- Non-linear effect on $l(p)$: Short cuts also contract the distance between neighbours of the connected nodes, their neighbours, and so on.
- Linear effect on $c(p)$: Removal of an edge from a node's neighbourhood has a linear effect on $c(p)$.
- At the local level, transition to a small world is almost undetectable.

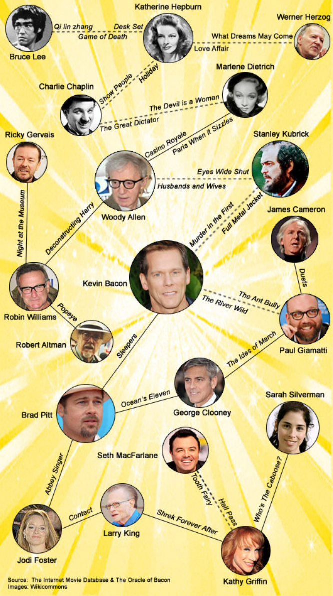
Observations

- $l(p)$ becomes small due to the addition of a small number of "long-range" edges.
- These short cuts connect nodes that would otherwise be very far apart.
- Non-linear effect on $l(p)$: Short cuts also contract the distance between neighbours of the connected nodes, their neighbours, and so on.
- Linear effect on $c(p)$: Removal of an edge from a node's neighbourhood has a linear effect on $c(p)$.
- At the local level, transition to a small world is almost undetectable.

Do real-world networks have small l and large c?

The Science of Six Degrees of Separation (Video, 9 min 22 sec)

Actor Network



Node $\equiv$

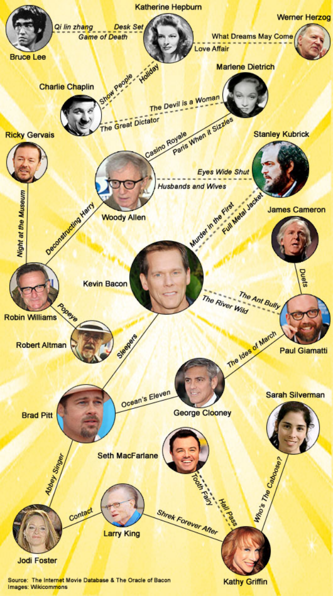
Edge $\equiv$

Edge weight $\equiv$

$n =$

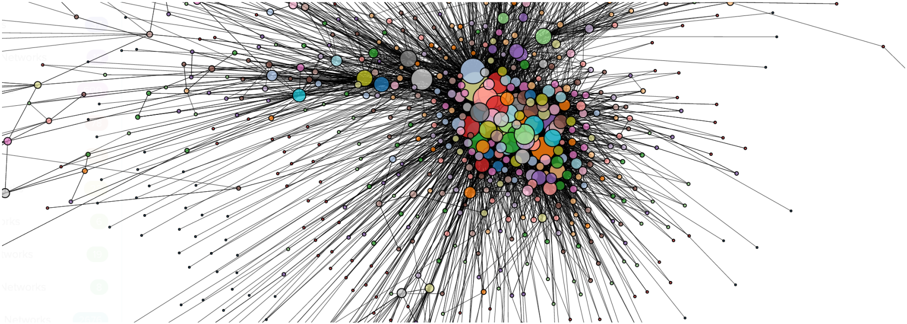
$m =$

Actor Network



Node $\equiv$ Actor
Edge $\equiv$ Collaboration
Edge weight $\equiv 1$
$n = 225,226$
$m = (225,226 \times 61)/2 = 6,869,393$

Power Network



Node $\equiv$

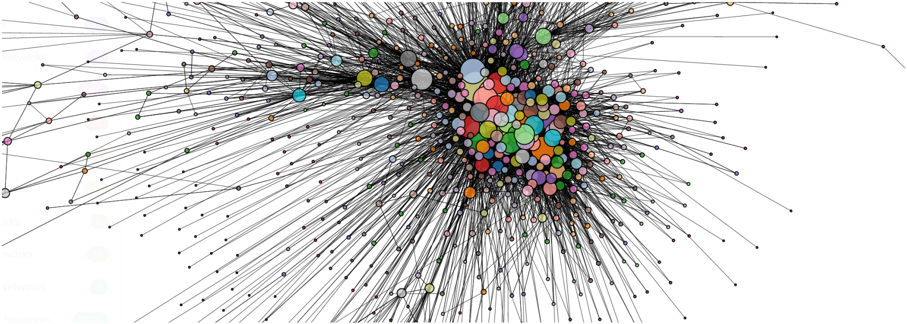
Edge $\equiv$

Edge weight $\equiv$

$n =$

$m =$

Power Network



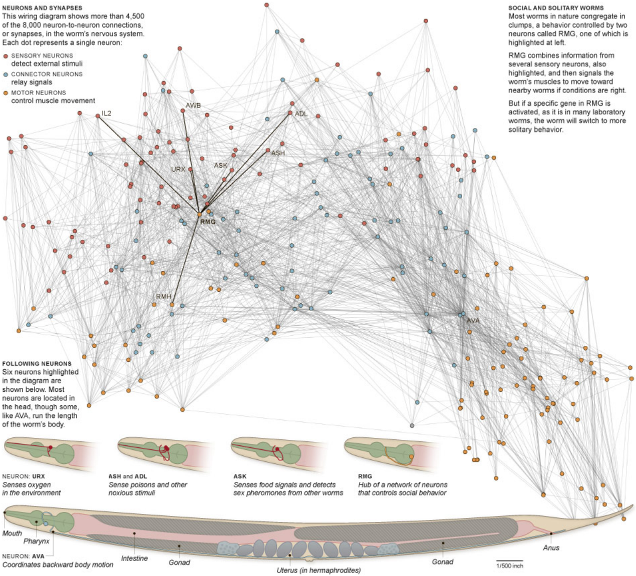
Node ≡ Generators, transformers, and substations

Edge ≡ High-voltage transmission line

Edge weight ≡ 1

$n = 4,941$
$m = (4,941 \times 2.67)/2 = 6,596$

C elegans connectome



Node $\equiv$

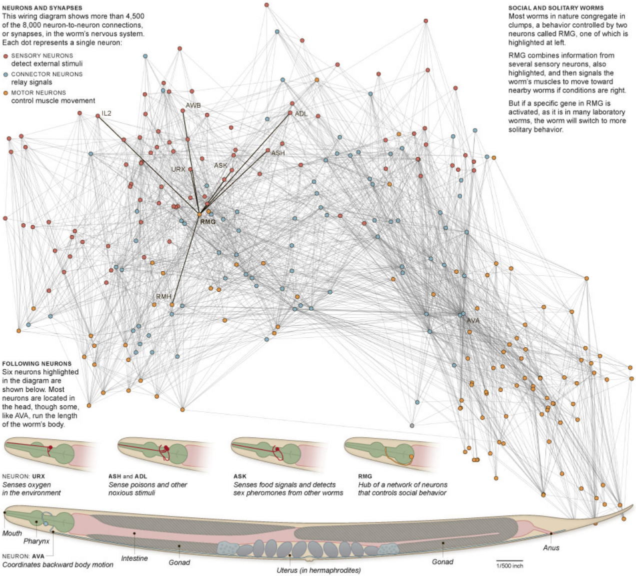
Edge $\equiv$

Edge weight $\equiv$

$n =$

$m =$

C elegans connectome



Node $\equiv$ Neuron

Edge $\equiv$ Synpase

Edge weight $\equiv 1$

$n = 282$

$m = (282 \times 14)/2 = 1974$

Real-world Networks are Small World

Table 1 Empirical examples of small-world networks

	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
C. elegans	2.65	2.25	0.28	0.05

Real-world Networks are Small World

Table 1 Empirical examples of small-world networks

	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
C. elegans	2.65	2.25	0.28	0.05

The pattern in Nature's networks (Video, 3 min 25 sec)