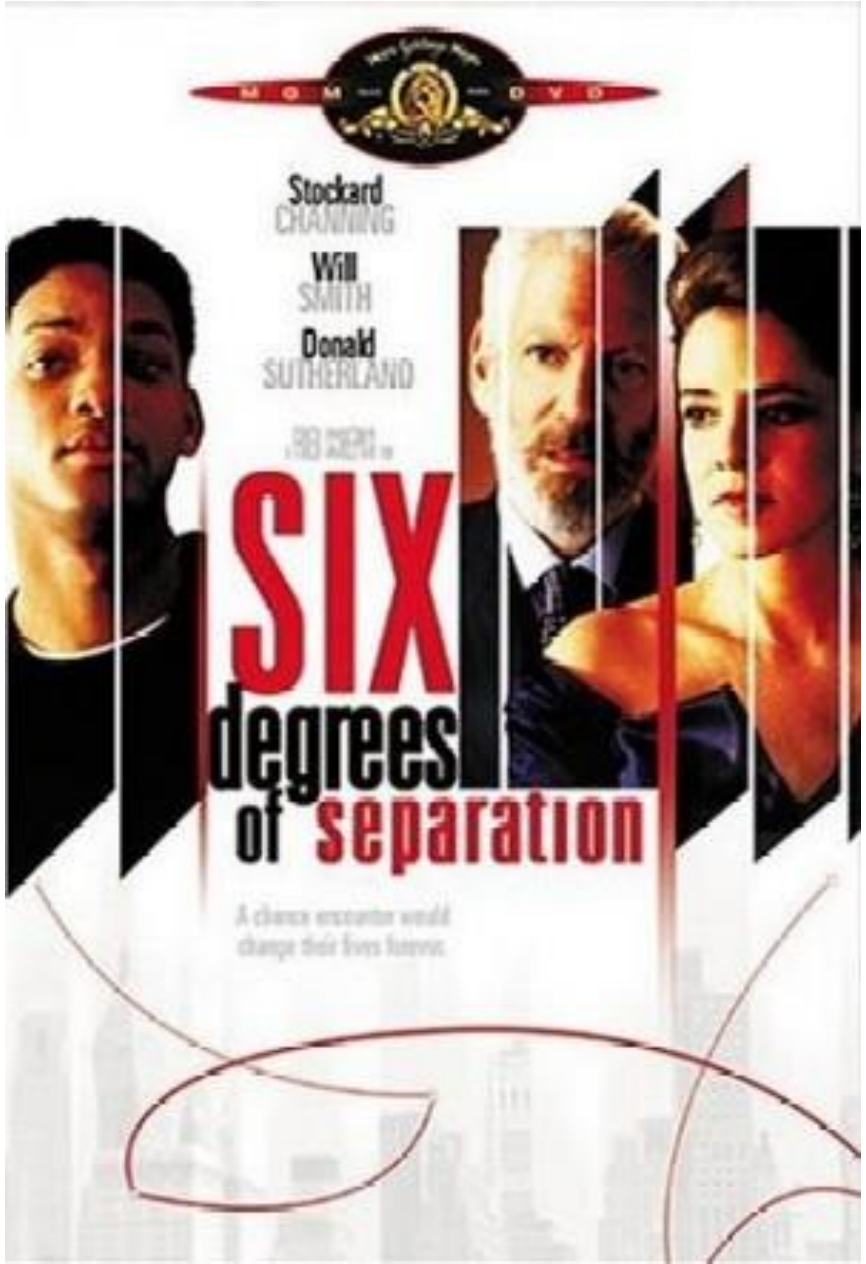
Lecture 7: Network Models II

Small-World Networks

Six Degrees of Separation

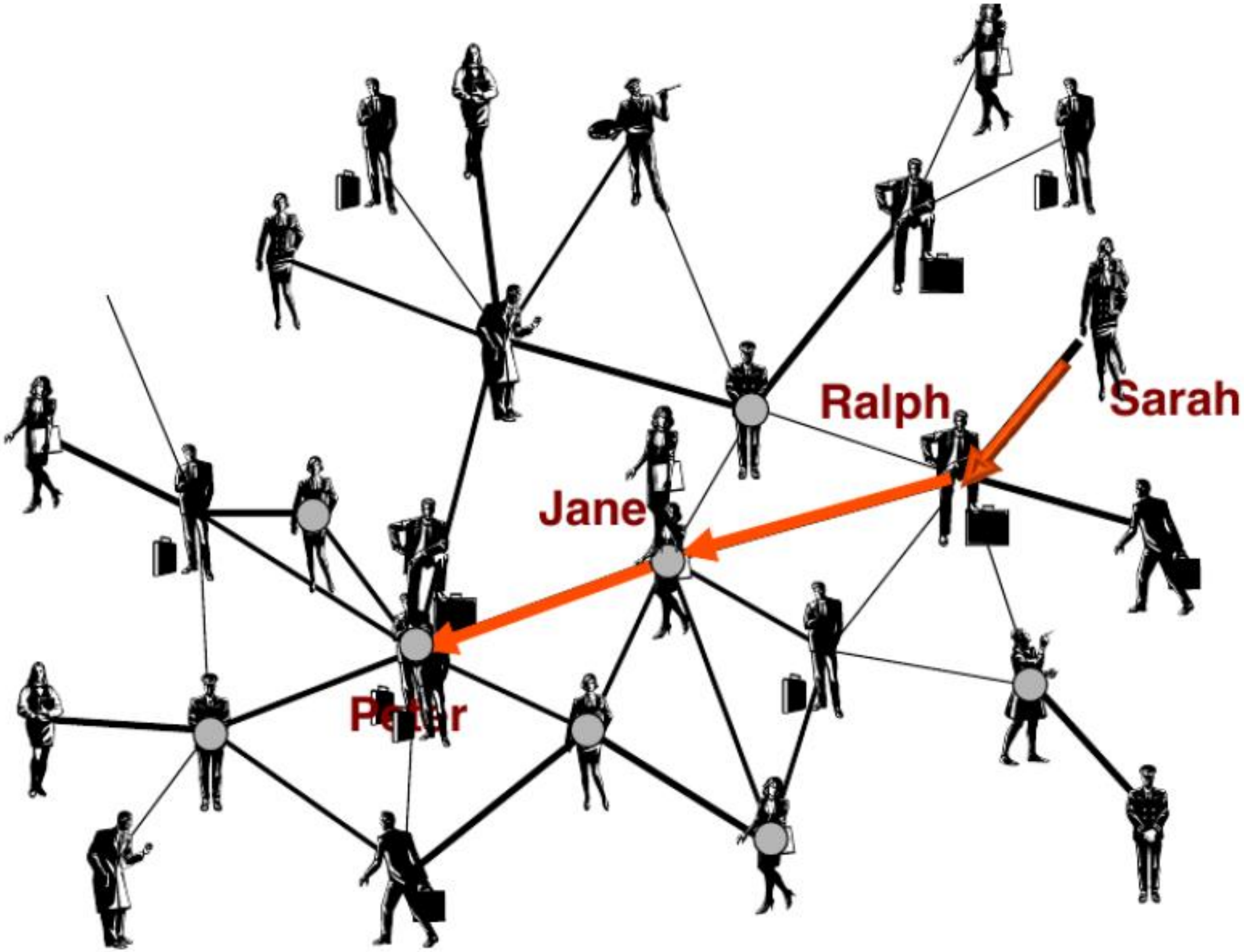


Six Degrees of Separation

Six degrees of separation refers to the idea that everyone is on average approximately six steps away from any other person on Earth, so that a chain of, “a friend of a friend” statements can be made, on average, to connect any two people in six steps or fewer. It was originally set out by *Frigyes Karinthy* (1929) and popularized by a play written by *John Guare*.

It was tested experimentally by *Stanley Milgram* (1967). In 1998, *Duncan J. Watts* and *Steven Strogatz* published the first small-world network model by analogy with the small-world phenomenon (six degrees of separation).

Six Degrees of Separation



Six Degrees of Separation

sport football opinion culture business lifestyle fashion environment tech

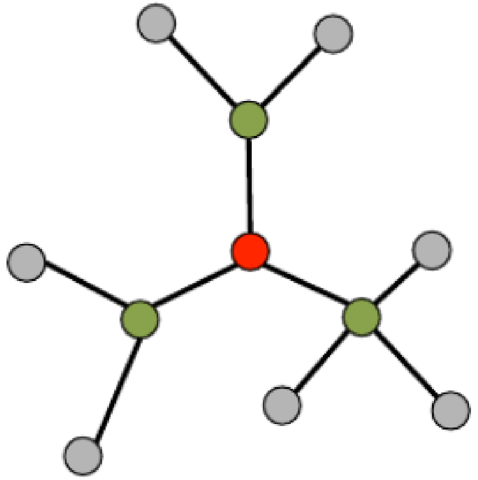
Facebook brings the world three-and-a-bit degrees of separation closer

The social media platform used its friend graph to calculate the degrees separating its 1.6 billion members and found it is as few as 3.57 people

Bringing the world together: Facebook says every person in the world is connected to every other person by an average of three and a half other people. Photograph: Alamy

Six Degrees of Separation

Random graphs tend to have a tree-like topology with almost constant node degrees.



- Number of first neighbors : $N_1 \cong < k >$
- Number of second neighbors : $N_2 \cong < k >^2$
- Number of neighbors at distance d : $N_d \cong < k >^d$
- Estimated maximum distance d_{max} :

$$N = 1 + \langle k \rangle + \langle k \rangle^2 + \ldots + \langle k \rangle^{d_{\max}} = \frac{\langle k \rangle^{d_{\max}+1} - 1}{\langle k \rangle - 1} \approx \langle k \rangle^{d_{\max}} \Rightarrow d_{\max} = \frac{\log N}{\log \langle k \rangle}$$

Diameter

Six Degrees of Separation

Network Name	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\frac{\log N}{\log\langle k \rangle}$
Internet	192,244	609,066	6.34	6.98	26	6.59
WWW	325,729	1,497,134	4.60	11.27	93	8.32
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	186,936	8.08	5.35	15	4.81
Actor Network	212,250	3,054,278	28.78	-	-	-
Citation Network	449,673	4,707,958	10.47	11.21	42	5.55
E Coli Metabolism	1,039	5,802	5.84	2.98	8	4.04
Yeast Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Small-World phenomenon: Applicable to real world networks

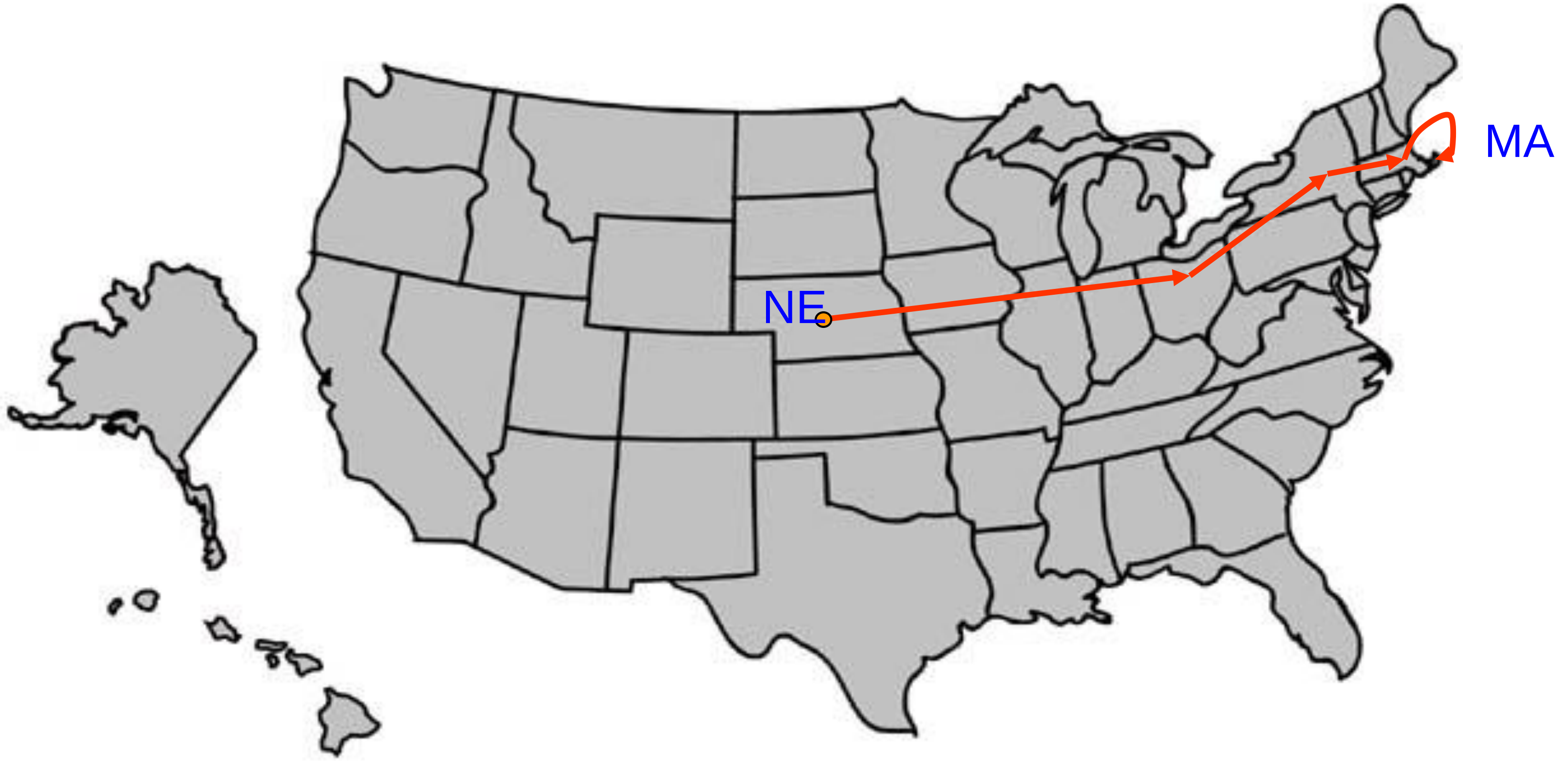
Same pattern:

high clustering $C_{network} \gg C_{random\ graph}$

low average shortest path $l_{network} \approx \ln(N)$

- neural network of C. elegans
- semantic networks of languages
- actor collaboration graph
- food webs

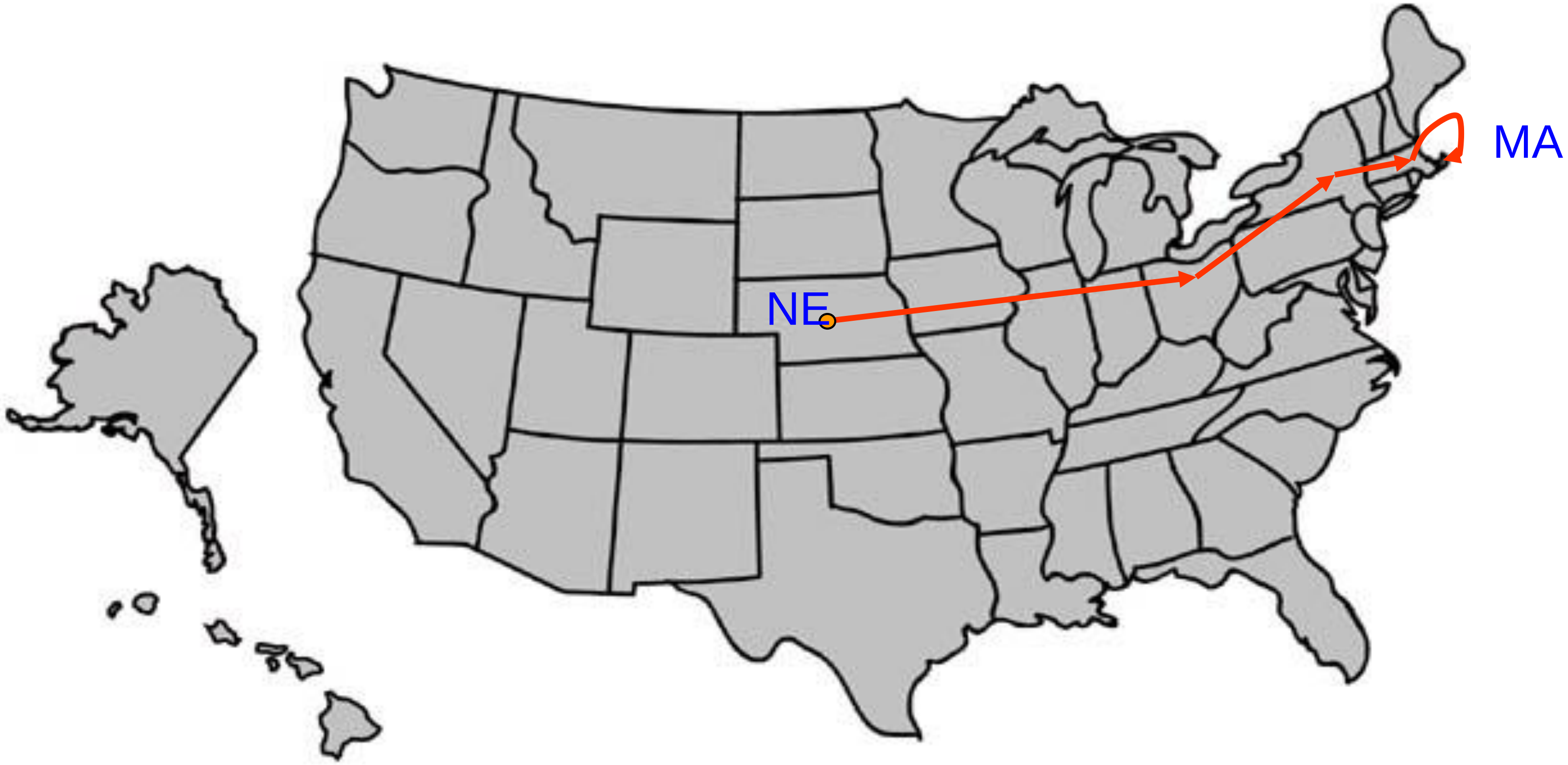
Small-World phenomenon: Milgram's experiment



A group of people from Omaha (Nebraska) and Wichita (Kansas) was asked to send a letter to an unknown person in Boston (Massachussetts).

Rule of the experiment: People should forward the letter to a person that they consider closer to the target person.

Small-World phenomenon: Milgram's experiment



Outcome:

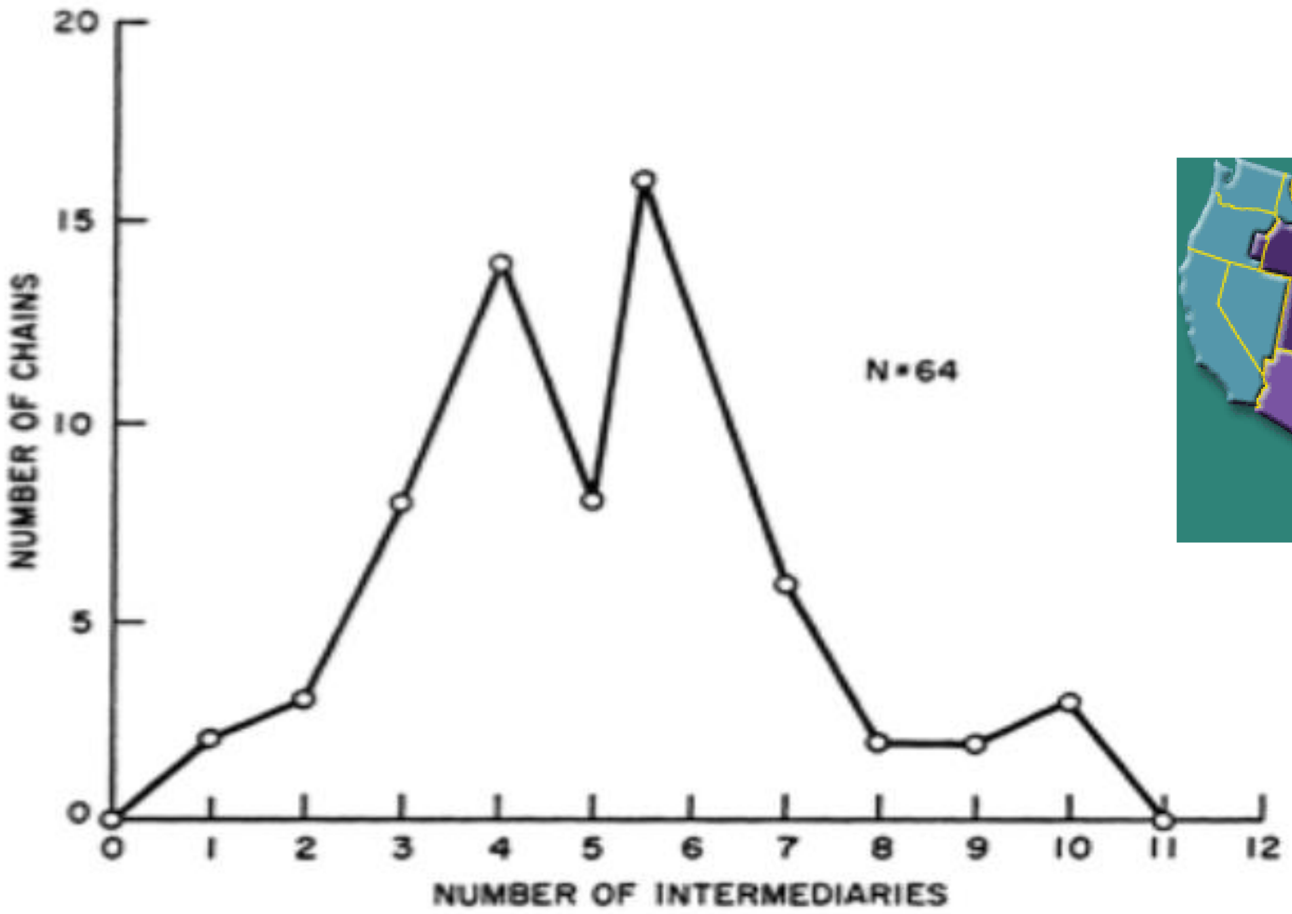
20% of initiated chains reached target (64 out of 296 letters (with path lengths from 1 to 10)

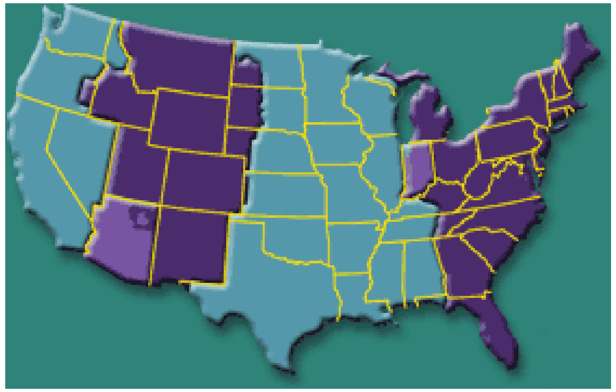
average path length ≈ 6

"Six degrees of separation"

Small-World phenomenon: Milgram's experiment

Distribution of Intermediaries in Milgram's experiment





"Six degrees of separation"

Small-World phenomenon: Interpreting Milgram's experiment

- Is 6 a ***surprising*** number?
  - In the 1960s? Today? Why?

- If social networks were random… ?
  - Pool and Kochen (1978) - ~500-1500 acquaintances/person
  - ~ 1,000 choices 1st link
  - ~ 1000^2 = 1,000,000 potential 2nd links
  - ~ 1000^3 = 1,000,000,000 potential 3rd links

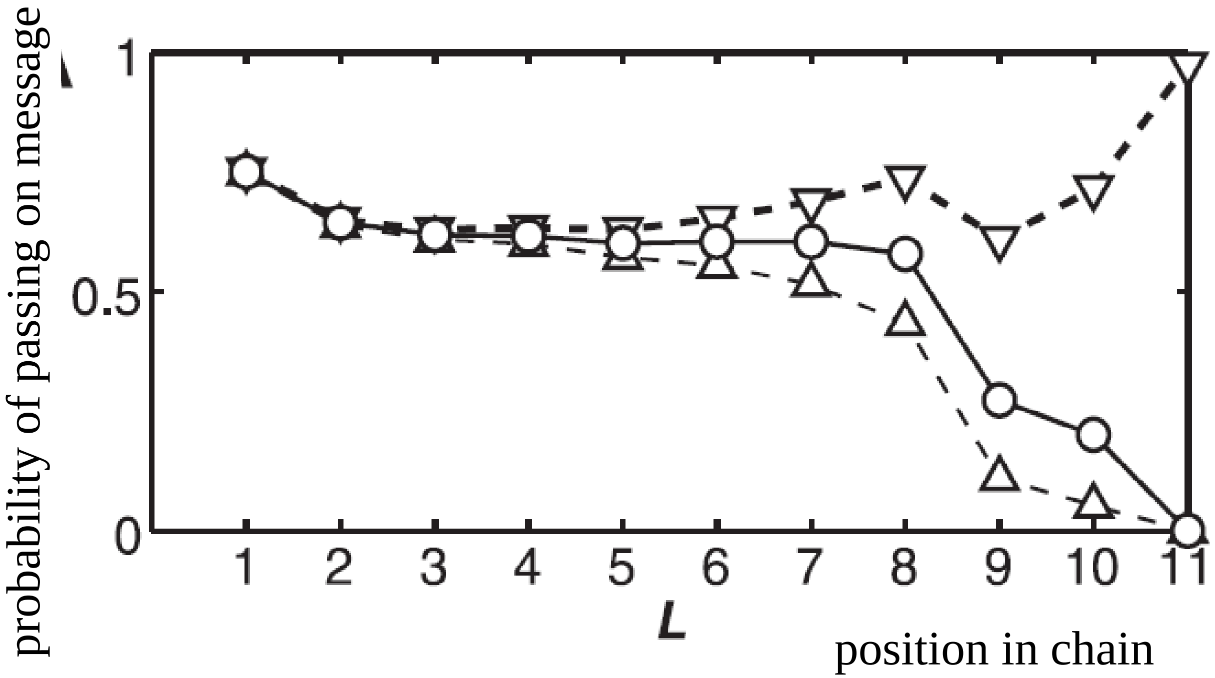
- If networks are completely cliquish?
  - all my friends' friends are my friends
  - what would happen?

Small-World experiment: Accuracy of distances

- Is 6 an ***accurate*** number?

- What bias is introduced by uncompleted chains?
  - are longer or shorter chains more likely to be completed?
  - if each person in the chain has 0.5 probability of passing the letter on, what is the likelihood of a chain being completed
    - of length 2?
    - of length 5?

Small-World experiment accuracy: Attrition rate approximately constant

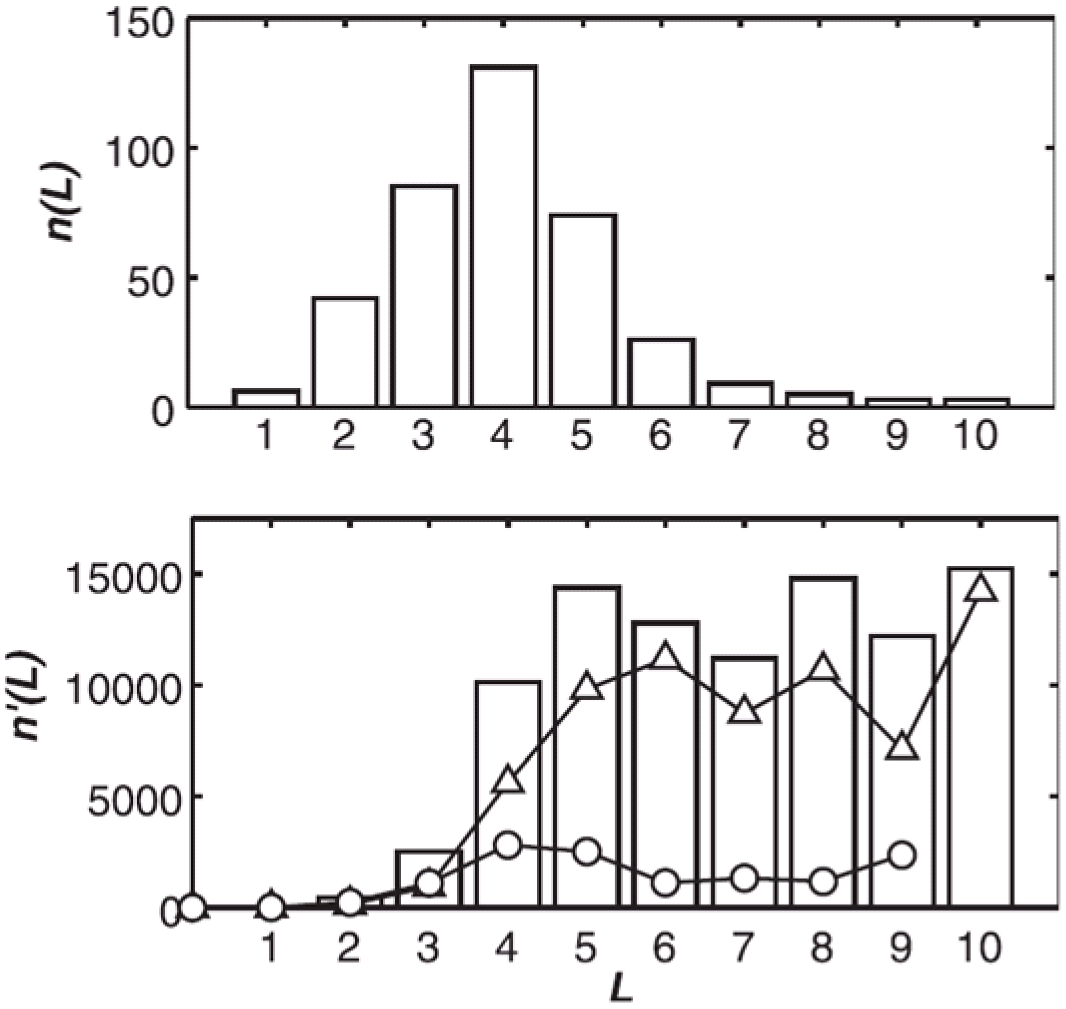


○ average

△ 95 % confidence interval

Source: An Experimental Study of Search in Global Social Networks: Peter Sheridan Dodds, Roby Muhamad, and Duncan J. Watts (8 August 2003); Science 301 (5634), 827.

Small-World experiment accuracy: Estimating true distance distribution



Above:

observed chain lengths

Below:

'recovered' histogram of path lengths

○ intra-country

△ inter-country

Source: An Experimental Study of Search in Global Social Networks: Peter Sheridan Dodds, Roby Muhamad, and Duncan J. Watts (8 August 2003); Science 301 (5634), 827.

Small-World experiment: Accuracy of distances

- Is 6 an ***accurate*** number?

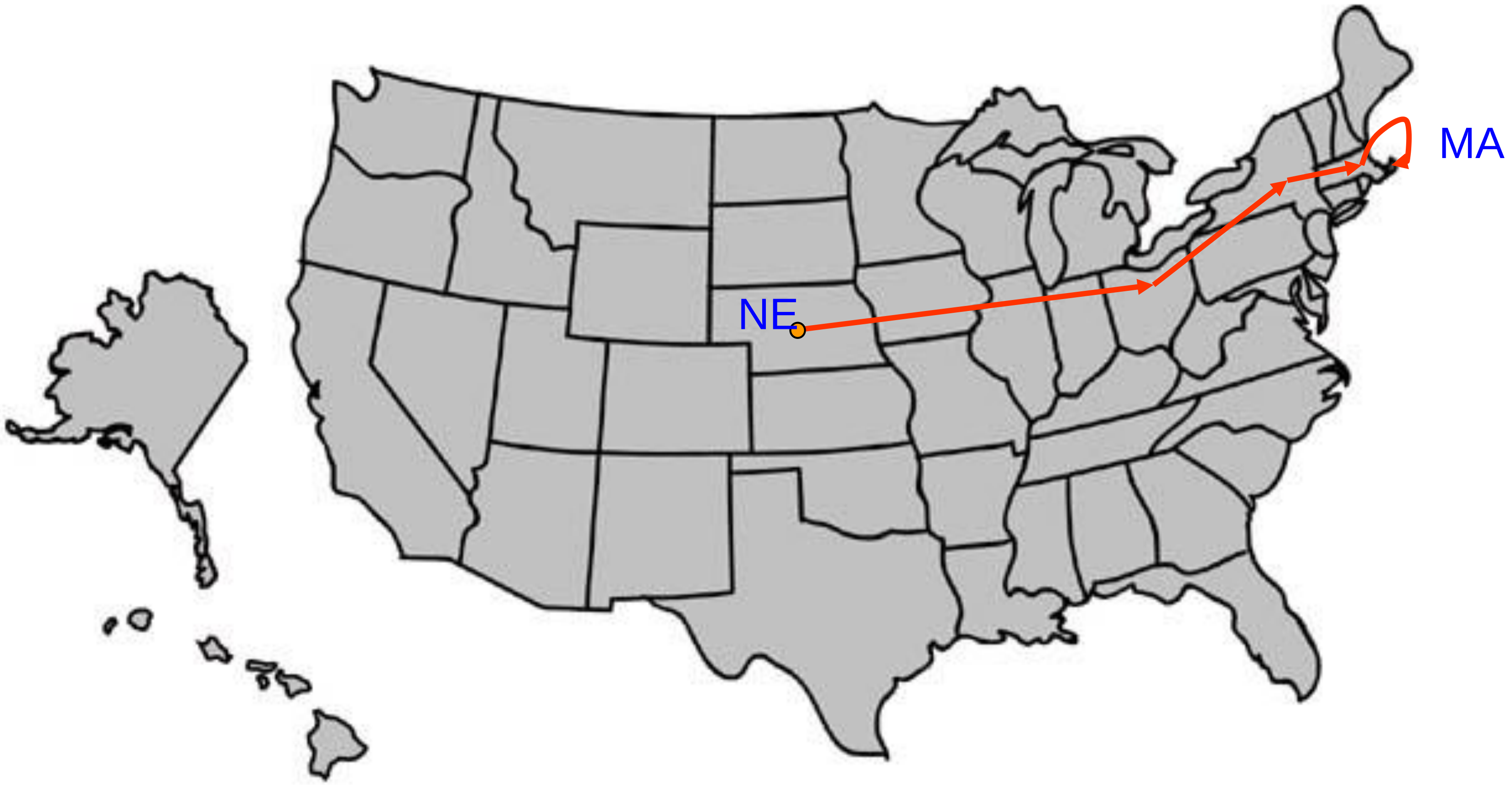
- Do people find the ***shortest*** paths?

  - The accuracy of small-world chains in social networks by Killworth et.al.
  - less than optimal choice for next link in chain is made ½ of the time

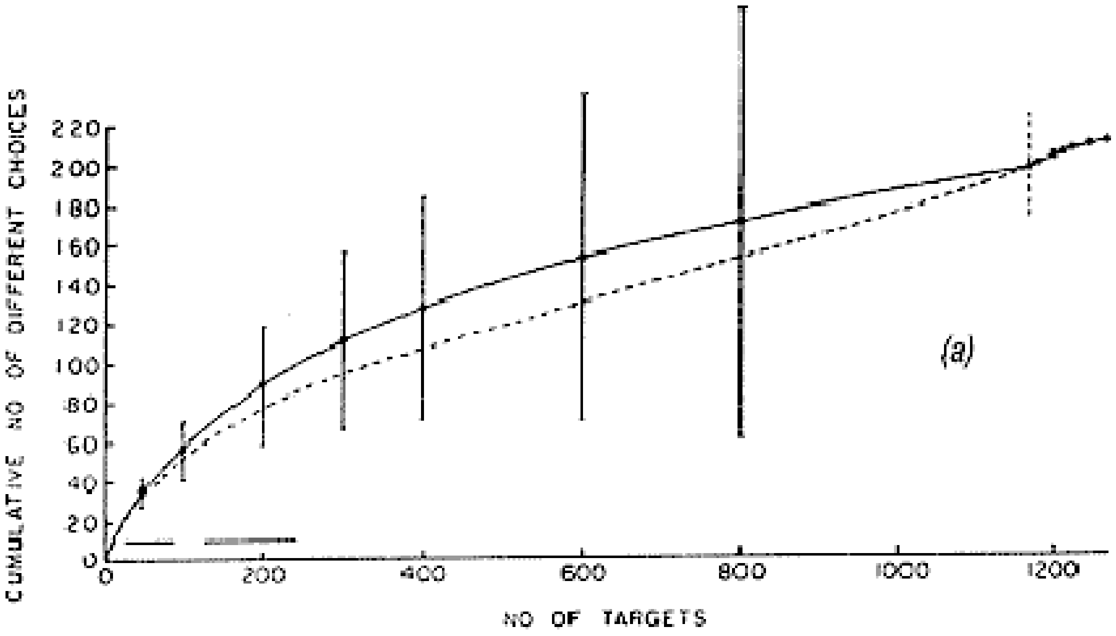
Geographical small world models: What if long range links depend on distance?

"The geographic movement of the [message] from Nebraska to Massachusetts is striking. There is a progressive closing in on the target area as each new person is added to the chain"

S.Milgram 'The small world problem', Psychology Today 1,61,1967



Navigability and Search Strategy: Reverse small world experiment



- Killworth & Bernard (1978)
- Given hypothetical targets (name, occupation, location, hobbies, religion…) participants choose an acquaintance for each target
- Acquaintance chosen based on
- (most often) Name, occupation and geographic location
- *Simple greedy algorithm*: most similar acquaintance

Source: 1978 Peter D. Killworth and H. Russell Bernard. "*The Reverse Small World Experiment Social Networks*" 1:159–92.

Small-World phenomenon: Repeating Milgram's experiment

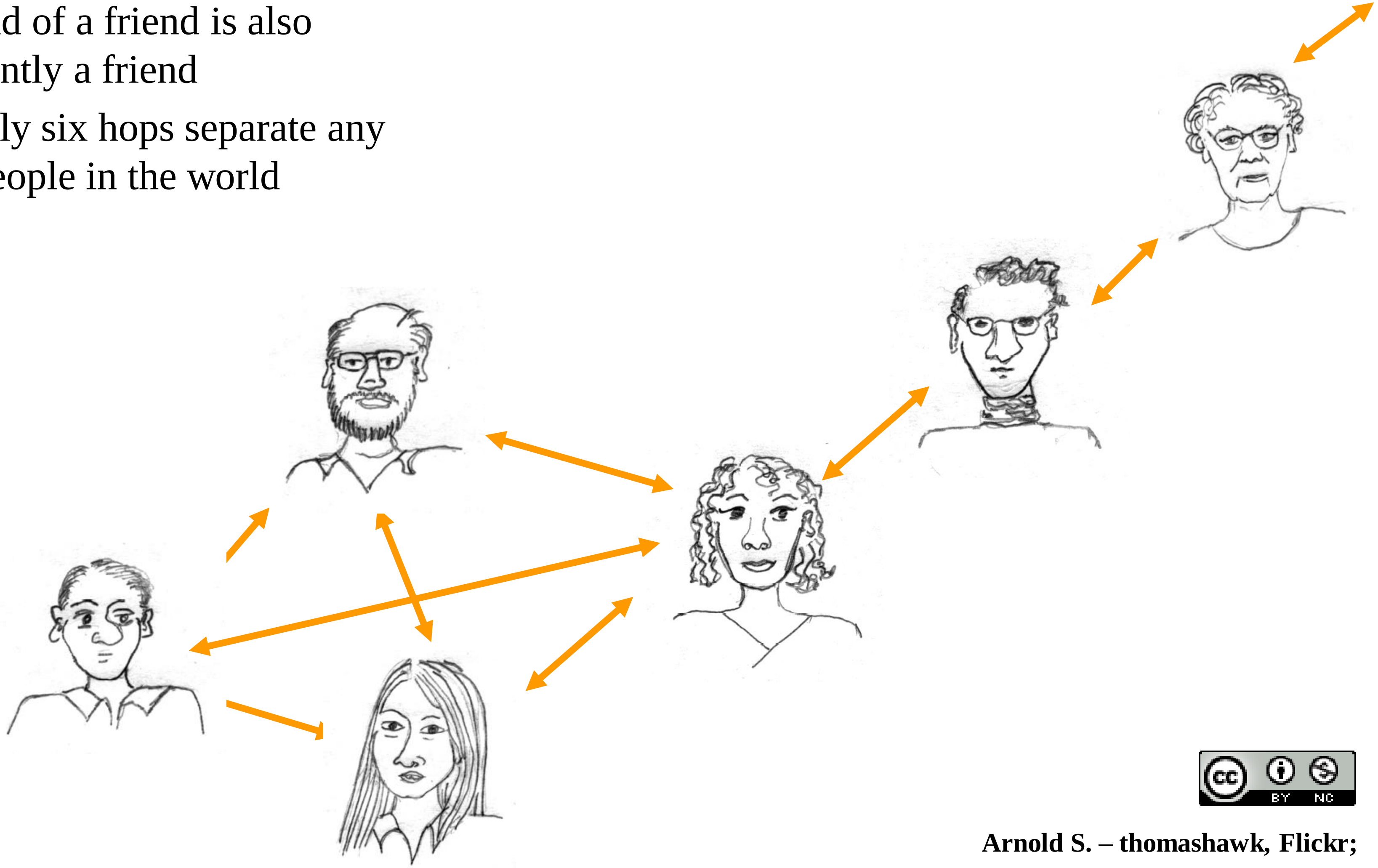
email experiment
Dodds, Muhamad, Watts,
Science 301, (2003)

•18 targets
•13 different countries

•60,000+ participants
•24,163 message chains
•384 reached their targets
•average path length 4.0

Modeling networks: Small Worlds

- Small worlds
  - a friend of a friend is also frequently a friend
  - but only six hops separate any two people in the world



Small-World Networks

A network is called a small-world network by analogy with the small-world phenomenon (popularly known as six degrees of separation). Historically, the small world hypothesis, which was first described by the Hungarian writer Frigyes Karinthy in 1929, and tested experimentally by Stanley Milgram (1967), is the idea that two arbitrary people are connected by only six degrees of separation. In 1998, Duncan J. Watts and Steven Strogatz published the first small-world network model, which through a single parameter smoothly interpolates between a random graph and a lattice.

D. J. Watts and S. H. Strogatz., *"Collective dynamics of 'small-world' networks"*, Nature 393 (1998) 440-442.

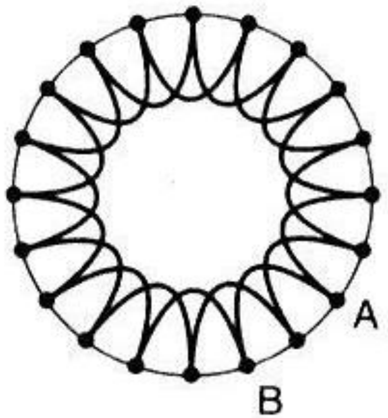
Small-World Networks

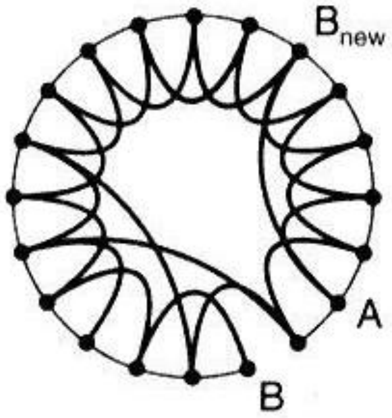
- Duncan Watts and Steven Strogatz

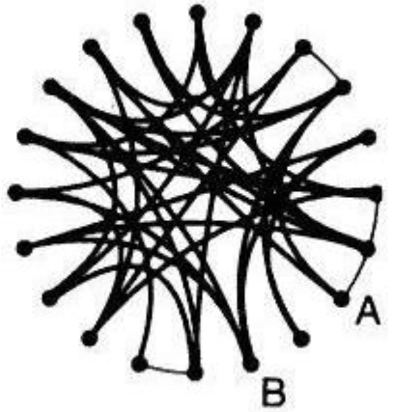
Reconciling two observations:

- **High clustering:** my friends' friends tend to be my friends
- **Short average paths**

- a few random links in an otherwise structured graph make the network a small world: average shortest path is short







regular lattice*:* my friend's friend is always my friend

small world*:* mostly structured with a few random connections

random graph*:* all connections random

D. J. Watts and S. H. Strogatz., *"Collective dynamics of 'small-world' networks"*, Nature 393 (1998) 440-442.

Watts-Strogatz model: Generating small-world graphs

Select a fraction p of links

Rewiring the links to other endpoints

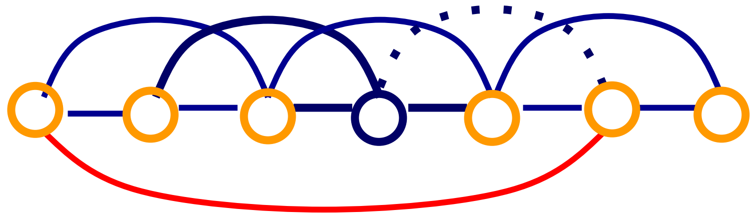
Add a fraction p of additional links leaving underlying lattice intact

- As in many network generating algorithms
  - Disallow self-edges
  - Disallow multiple edges

D. J. Watts and S. H. Strogatz., *“Collective dynamics of 'small-world' networks”*, Nature 393 (1998) 440-442.

Watts-Strogatz model: Generating small-world graphs

- Each node has $k \geq 4$ nearest neighbors (local)
- tunable: vary the probability p of rewiring any given edge
- small p: regular lattice
- large p: classical random graph

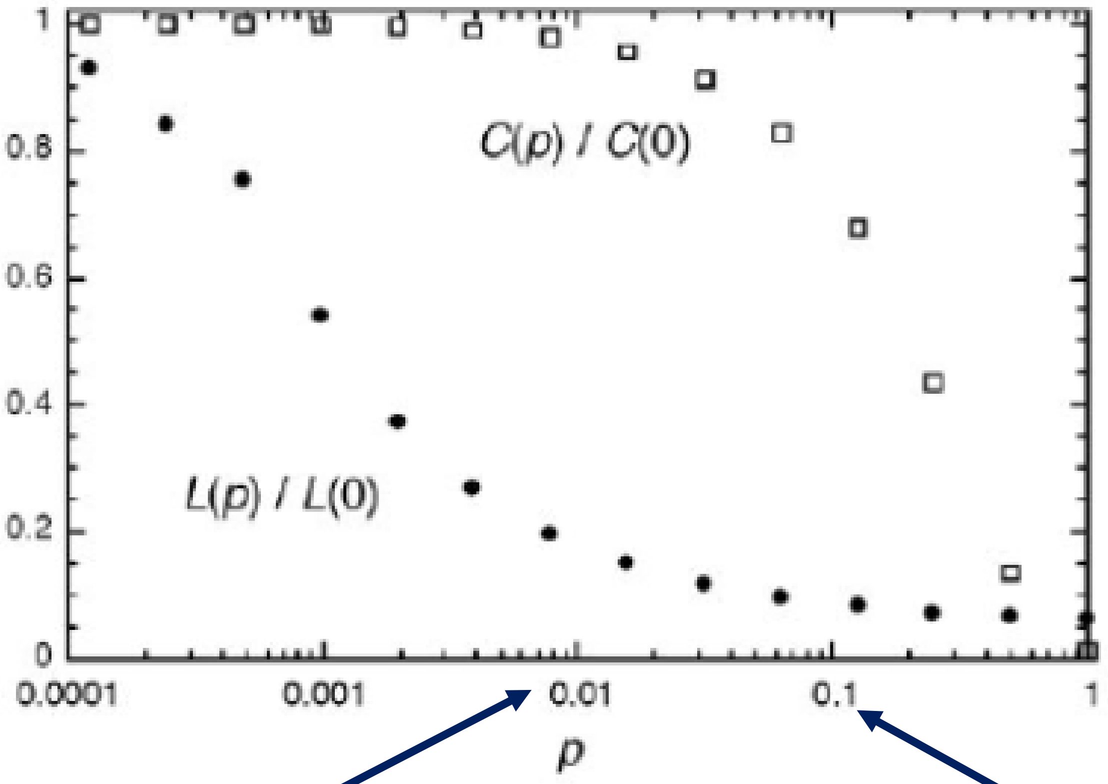


Watts-Strogatz model: What happens in between?

- Small shortest path means small clustering?
- Large shortest path means large clustering?
- Through numerical simulation
  - As we increase p from 0 to 1
    - Fast decrease of mean distance
    - Slow decrease in clustering

Watts-Strogatz model:

Change in clustering coefficient and average path length as a function of the proportion of rewired edges



1% of links rewired

10% of links rewired

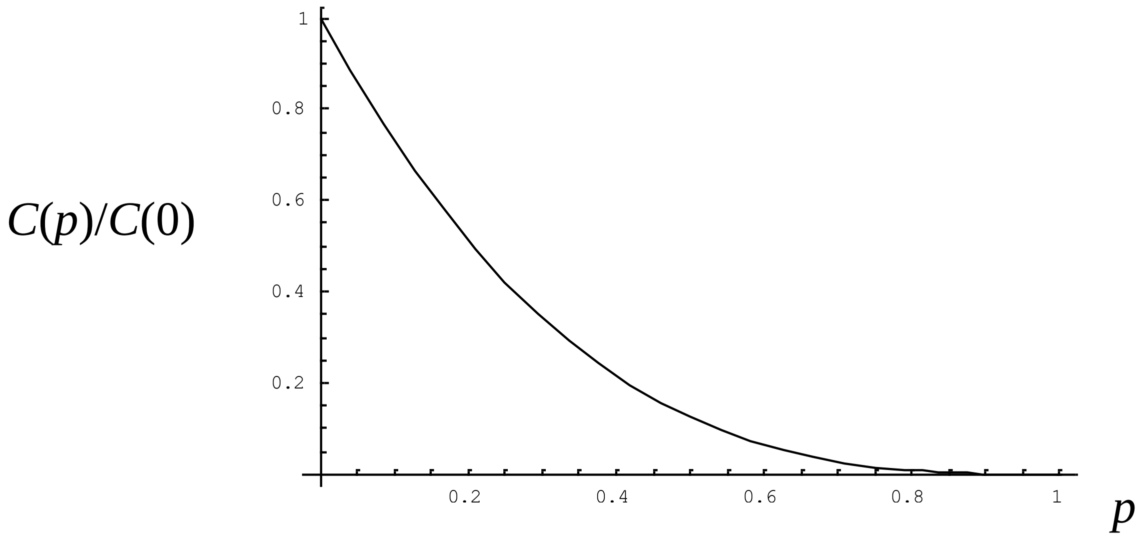
D. J. Watts and S. H. Strogatz., *"Collective dynamics of 'small-world' networks"*, Nature 393 (1998) 440-442.

Watts-Strogatz model:

Clustering coefficient can be computed for WS model with rewiring

- The probability that a connected triple stays connected after rewiring
  - probability that none of the 3 edges were rewired $(1-p)^3$
  - probability that edges were rewired back to each other is very small, can ignore

- Clustering coefficient = $C(p) = C(0)(1-p)^3$

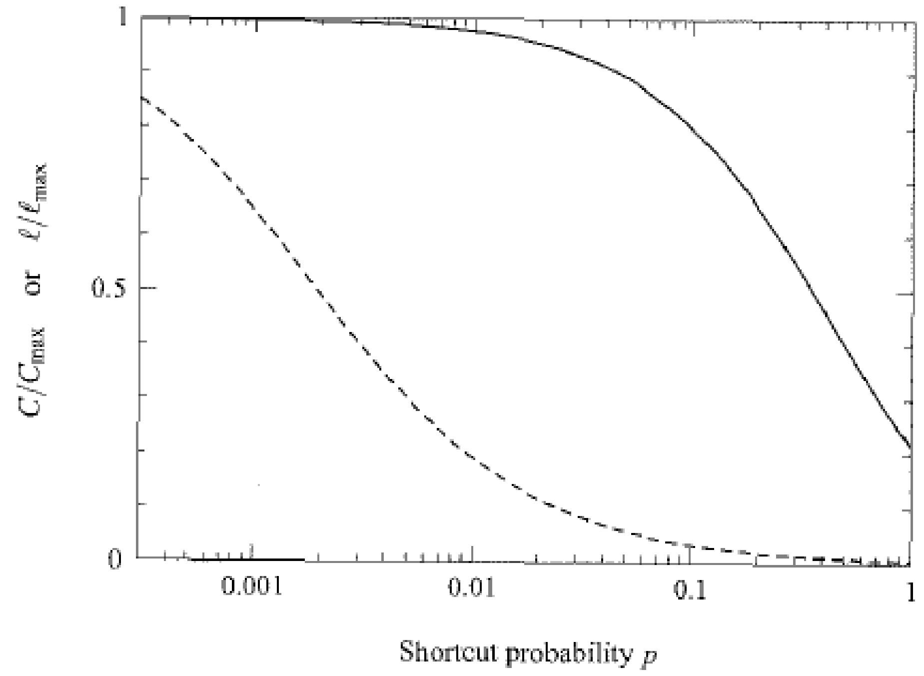


D. J. Watts and S. H. Strogatz., *"Collective dynamics of 'small-world' networks"*, Nature 393 (1998) 440-442.

Watts-Strogatz model:

Clustering coefficient: addition of random edges

- How does *C'* depend on *p*?
- $C'(p) = \frac{3 \times (number\ of\ triangles)}{(number\ of\ connected\ triples)}$
- *C'*(*p*) computed analytically for the small world model without rewiring



$$C'(p) = \frac{3(c-2)}{4(c-1) + 4cp(p+2)}$$

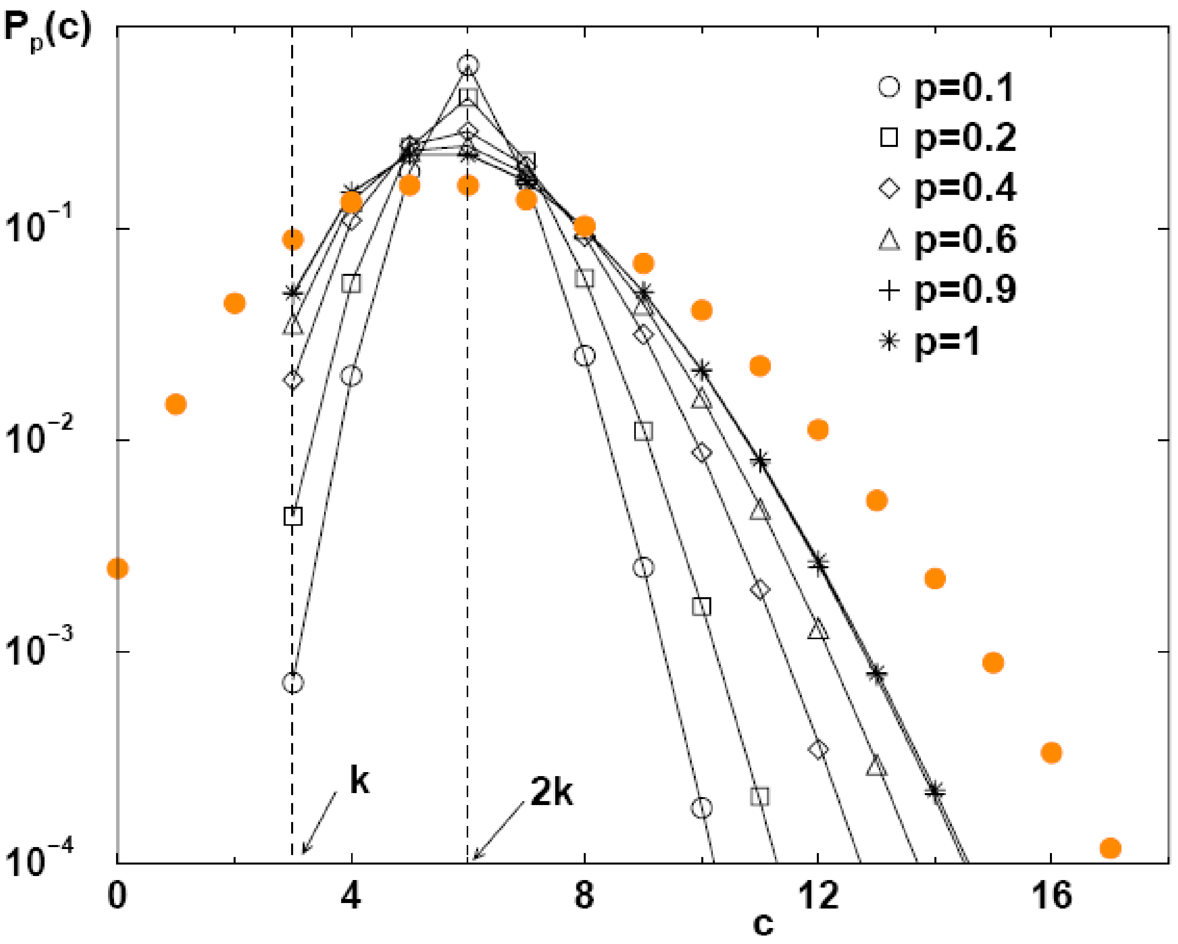
Newman, M. E. J., Watts, D. J., Phys. Lett. A 263(4–6), 341–346 (1999).

Watts-Strogatz model: Degree distribution

- $p = 0 \rightarrow \delta$-function
- $p > 0$ broadens the distribution
- Edges left in place with probability $(1-p)$
- Edges rewired towards i with probability $1/n$

Watts-Strogatz model:

Model: small-world with probability p of rewiring



1000 vertices

random network with average connectivity k

visit nodes sequentially and rewire links

D. J. Watts and S. H. Strogatz., *"Collective dynamics of 'small-world' networks"*, Nature 393 (1998) 440-442.

Watts-Strogatz model:
Average path lengths: addition of random edges

No exact expression for mean distance has yet been found, but some approximate expressions are known and have been found in simulations of the model to be reasonably accurate. One thing that is known about path lengths in the model is how they *scale* with the model parameters:

$$\frac{\boldsymbol{cl}}{\boldsymbol{n}} = \boldsymbol{f}(\boldsymbol{ncp})$$

where

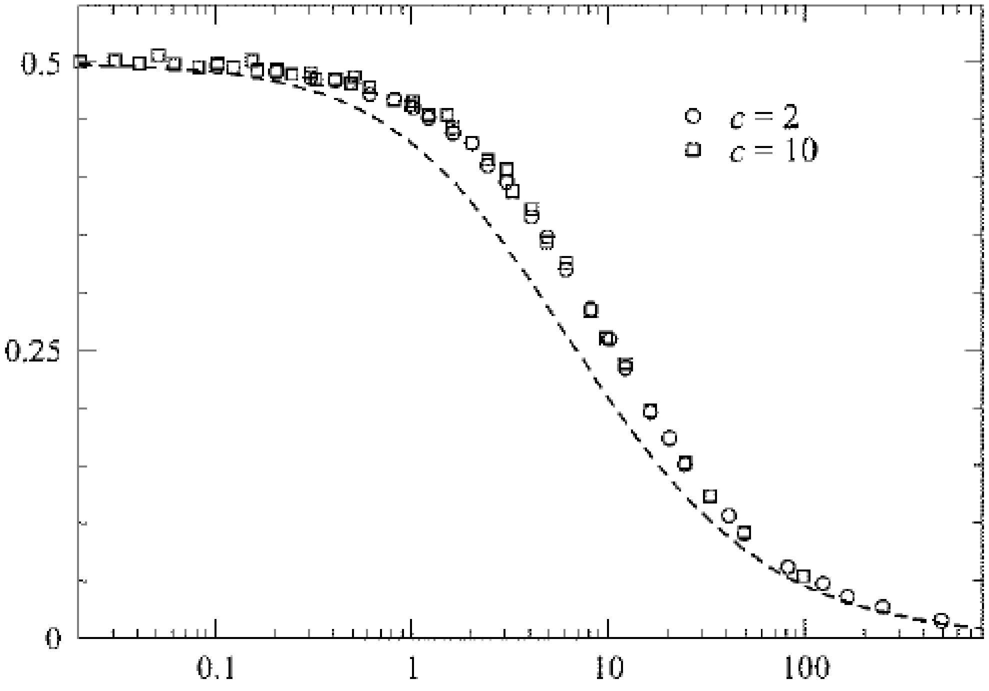
l is the average path lengths of the network
n is the total number of vertices in the network
c is the degree of the vertices (before adding random edges)
p is the probability of the addition of shortcuts (random links)
f is a universal scaling function depending on the dimensionless parameter *x* (=*ncp*)

Barthelemy, M. and Amaral, L.A.N., *Phys. Rev, Lett*, 82,3180-3183 (1999).
Newman, M.E.J., Moore, C., and Watts, D.J., *Phys. Rev. Lett.* 84, 3201-3204 (2000).

Watts-Strogatz model:

Average path lengths: addition of random edges

$\frac{cl}{n}$



ncp

$$f(x) = \frac{2}{\sqrt{x^2+4x}} tanh^{-1}\sqrt{\frac{x}{x+4}}$$

Scaling function for the small-world model. The points show numerical results for *cl/n* as a function of *ncp* for the small-world model with n = 128 to 32768 and p = 0.000001 to 0.03, and two different values of c as marked. Each point is averaged over 1000 networks with the same parameter values. The points collapse, to a reasonable approximation, onto a single scaling function *f(ncp)*. The dashed curve is the mean-field approximation to the scaling function given by the equation on the right.

Small-World Networks

Comparison with "random graph" to determine whether real-world networks are "small-world"

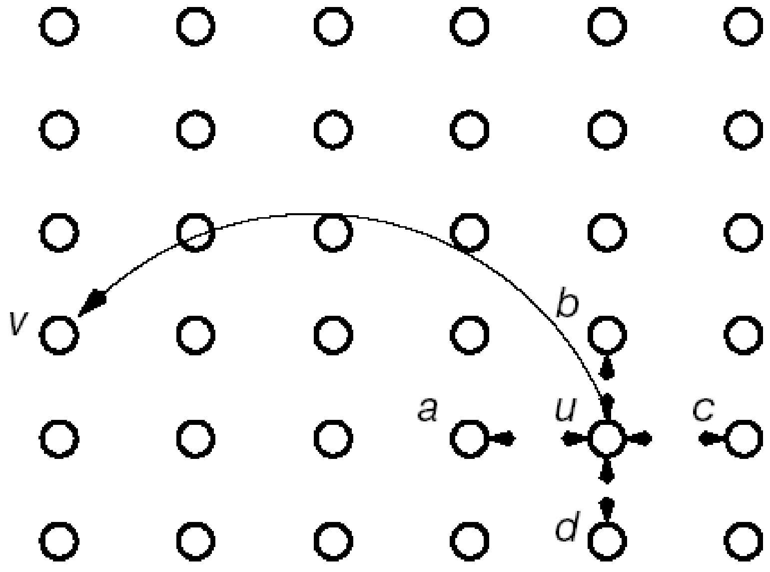
Network	**size**	**av. shortest path**	**Shortest path in fitted random graph**	**Clustering (averaged over vertices)**	**Clustering in random graph**
Film actors	225,226	3.65	2.99	0.79	0.00027
MEDLINE co-authorship	1,520,251	4.6	4.91	0.56	1.8×10^{-4}
E.Coli substrate graph	282	2.9	3.04	0.32	0.026
C.Elegans	282	2.65	2.25	0.28	0.05

Small-World Networks

What features of real social networks are missing from the WS small-world model?

- Long range links not as likely as short range ones
- Hierarchical structure / groups
- Hubs

Kleinberg's geographical small-world model



nodes are placed on a lattice and
connect to nearest neighbors

exponent that will determine navigability

additional links placed with

$$p(\text{link between } u \text{ and } v) = (\text{distance}(u,v))^{-\alpha}$$

Source: Kleinberg, 'Navigation in a small world'

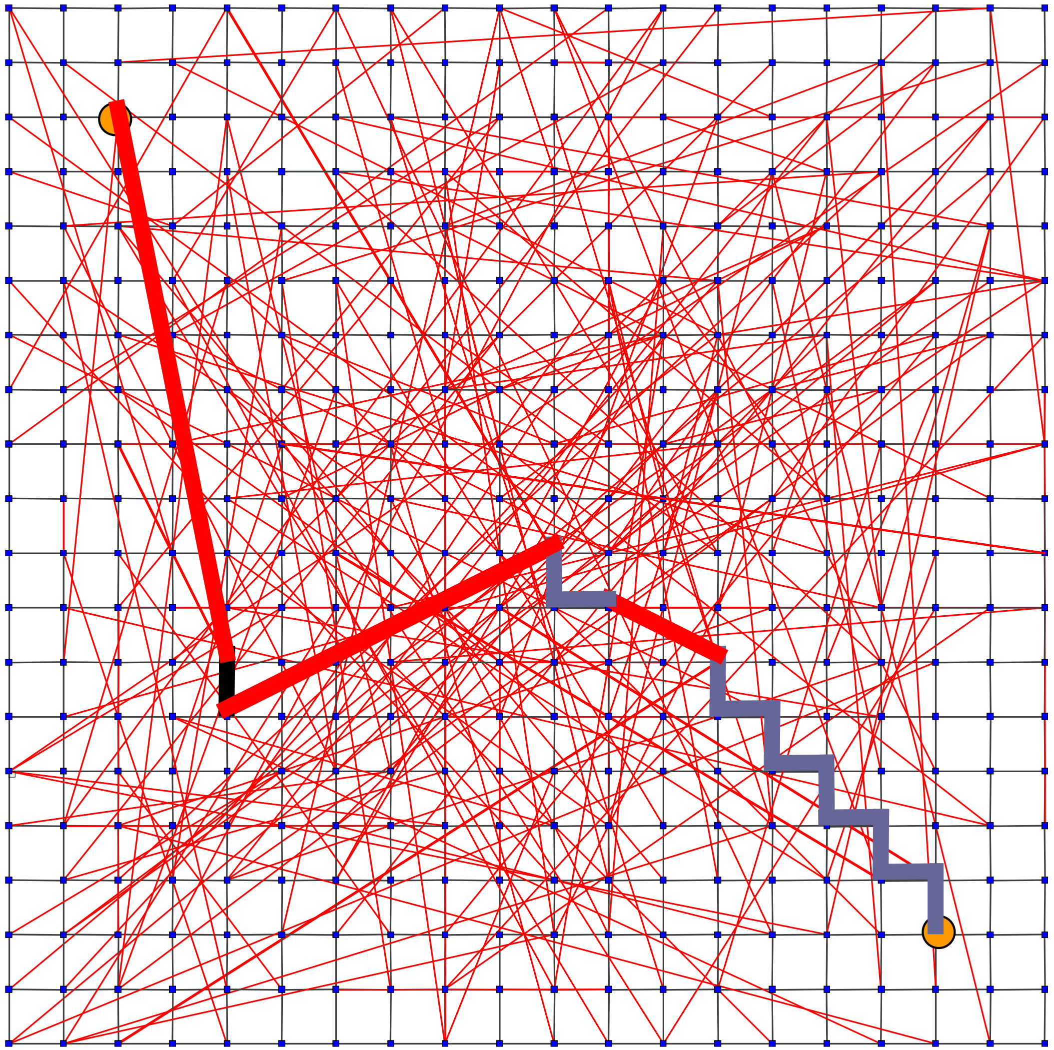
Kleinberg's geographical small-world model:

Geographical search when network lacks locality

When **α=0**, links are randomly distributed, *average shortest path* ~ **log(*n*)**, *n* is size of the grid

When **α=0**, expected search time ~$a_0 n^{2/3}$

$p \sim p_0$

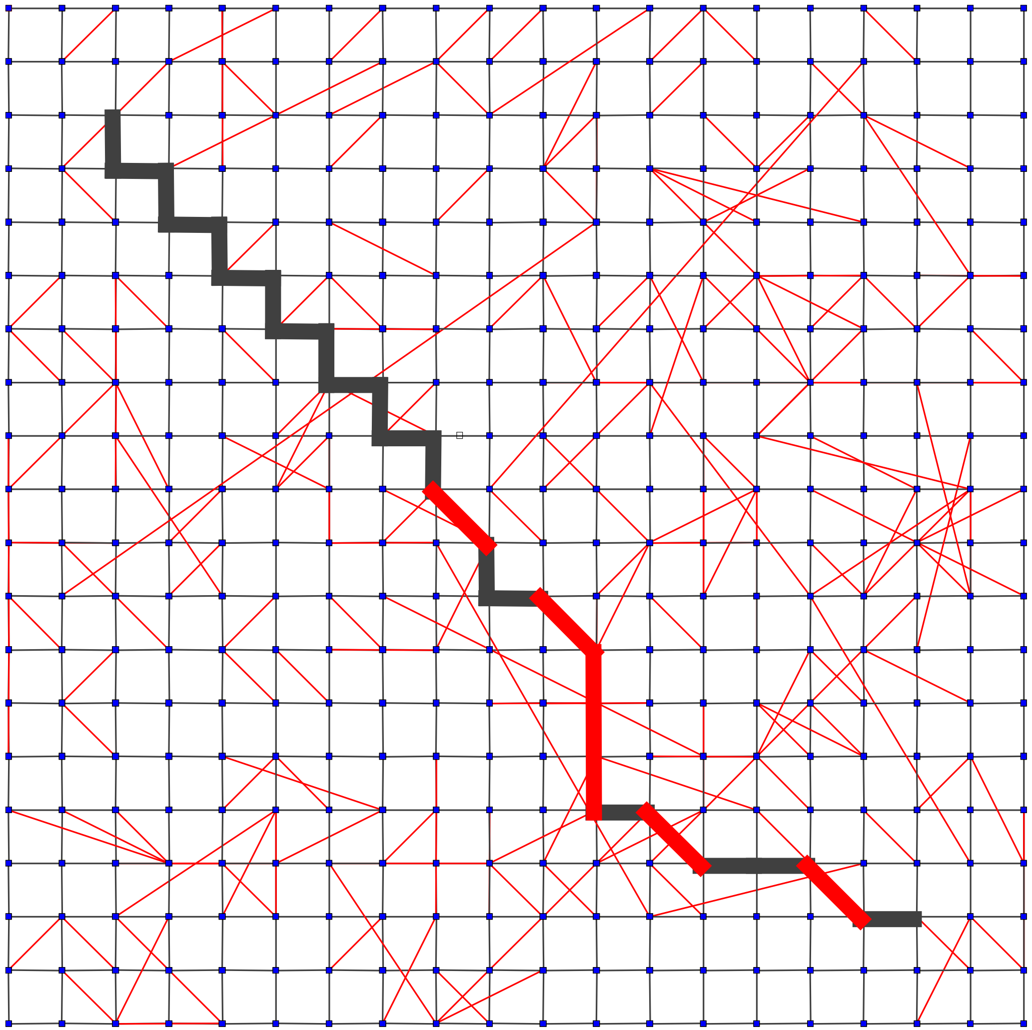


When **α<2**,

expected time at least $a_\alpha n^{(2-\alpha)/3}$

Kleinberg's geographical small-world model:
Overly localized links on a lattice

When **α>2** expected search time ~ $\boldsymbol{n^{(\alpha-2)/(\alpha-1)}}$



In this example,

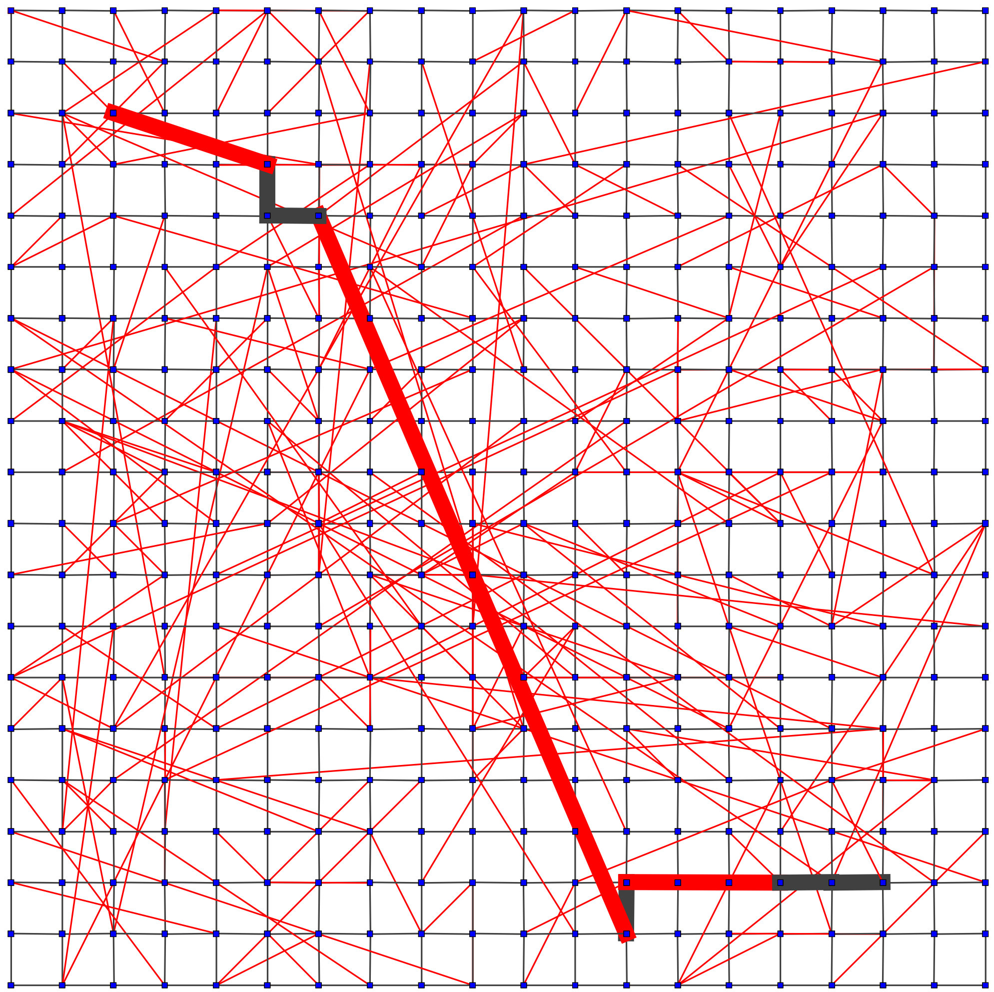
$\alpha = 4, \quad p \sim \frac{1}{d^4}$

Kleinberg's geographical small-world model:

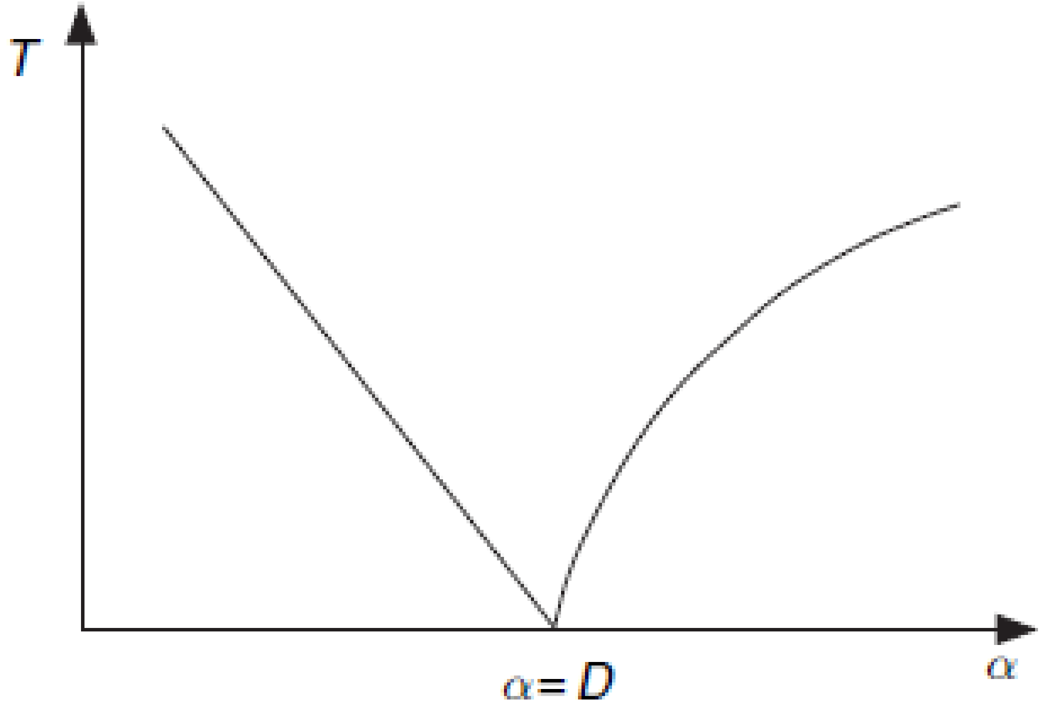
Links balanced between long and short range

When **α=2,** expected search time ~ ***C*** **(log *n*)^2**

$$p \sim \frac{1}{d^2}$$



Kleinberg's geographical small-world model: Results



D is the spatial dimension

Schematic evolution of the delivery time as a function of α in Kleinberg's model of a small-world network. Figure adapted from Kleinberg (2000).

Origins of small-worlds: Other generative models

- Hierarchical small-world models (Kleinberg; Watts, Dodds & Newman; …)

- Assign properties to nodes
  - e.g. spatial location, group membership

- Add or rewire links according to some rules
  - optimize for a particular property
    - simulated annealing

  - add links with probability depending on property of existing nodes, edges
    - preferential attachment, link copying

  - simulate nodes as agents ‘deciding’ whether to rewire or add links

Summary

- The world is small!
- Simple models to explain why
- Other models incorporate geography and hierarchical social structure
- Small worlds may evolve from different constraints
  - navigation, constraint optimization, group affiliation

Exponential Random Graphs

Exponential Random Graph Models

Exponential random graph models (***ERGMs***) are statistical models for network structure, permitting inferences about how network ties are patterned. They are largely studied in social network analysis (Holland and Leinhardt, 1981; Frank and Strauss, 1986). This approach considers the adjacency matrix – also called the ***sociomatrix*** in the social network literature – $G = \{g_{ij}\}$ characterizing the graph of size *N* as a random matrix whose realization occurs with a probability *P*(*G*) defined in the sample space of all possible graphs. Here, one wants to create an ensemble of networks with a given set of properties, such as a given number of edges or a given value of the clustering coefficient. The approach is to fix the *average* value of the property or properties of interest.

Suppose one has a set of network measures of which the numerical values we want to fix. For example, number of edges or mean degree of a vertex, degrees of individual vertices, number of triangles or clustering coefficient, etc. Denote these measures by $x_1, x_2, \ldots$. Now consider the set $\mathcal{G}$ of all simple graphs with *n* vertices and let us define an ensemble by giving each graph *G* in the set a probability *P*(*G*), normalized so that

$$\sum_{G \in \mathcal{G}} P(G) = 1$$

Exponential Random Graph Models (cont'd)

The expectation value $\langle x_i \rangle$ of a network measure x_i within this ensemble is given by

$$\langle x_i \rangle = \sum_{G \in \mathcal{G}} P(G) x_i(G)$$

where $x_i(G)$ is the value of x_i measured on the graph G. The probability distribution $P(G)$ is unknown and needs to be fixed. One way is to employ the ***Principle of Maximum Entropy*** to obtain $P(G)$, i.e., to maximize the Gibbs entropy

$$S = -\sum_{G \in \mathcal{G}} P(G) \ln P(G)$$

subject to the known constraints. The optimum is the set of values of the $P(G)$ that maximizes the quantity

$$-\sum_{G \in \mathcal{G}} P(G) \ln P(G) - \alpha \left[1 - \sum_{G \in \mathcal{G}} P(G)\right] - \sum_i \beta_i \left[\langle x_i \rangle - \sum_{G \in \mathcal{G}} P(G) x_i(G)\right]$$

where α and β_i are Lagrange multipliers whose values will be determined shortly.

Exponential Random Graph Models (cont'd)

Differentiate with respect to $P(G)$,

$$-\ln P(G) - 1 - \alpha + \sum_i \beta_i x_i(G) = 0$$

$$\longrightarrow \quad P(G) = \exp\left[\alpha - 1 + \sum_i \beta_i x_i(G)\right] \rightarrow \; P(G) = \frac{e^{H(G)}}{Z}$$

Where $Z = e^{1-\alpha}$ is called the *partition function*, and $H(G) = \sum_i \beta_i x_i(G)$, is the *graph Hamiltonian.*

Using

$$\sum_{G\in\mathcal{G}} P(G) = \frac{1}{Z}\sum_{G\in\mathcal{G}} e^{H(G)} = 1 \;\rightarrow Z = \sum_{G\in\mathcal{G}} e^{H(G)}$$

There is no equivalent general formula for β*.* They are determined from the expectation value $\langle x_i \rangle$ of the network measures x_i.

Exponential Random Graph Models (cont'd)

Once we determine $P(G)$, we can calculate the expectation value of a quantity y in the ensemble

$$\langle y \rangle = \sum_{G \in \mathcal{G}} P(G) y\,(G) = \frac{1}{Z} \sum_{G \in \mathcal{G}} e^{H(G)}\, y(G)$$

If y is one of the network measures x_i above, we would have

$$\langle x_i \rangle = \frac{1}{Z} \sum_{G \in \mathcal{G}} e^{H(G)}\, x_i(G) = \frac{1}{Z} \frac{\partial}{\partial \beta_i} \sum_{G \in \mathcal{G}} e^{\sum_i \beta_i x_i(G)} = \frac{1}{Z} \frac{\partial Z}{\partial \beta_i} = \frac{\partial \ln Z}{\partial \beta_i} = \frac{\partial F}{\partial \beta_i}$$

where the quantity $F = \ln Z$ is called the *free energy* of the ensemble.

Exponential Random Graph Models (cont'd)

Example: Simple graph, undirected and unweighted

H = βm, where m is the number of edges. Then,

$$P(G) = \frac{e^{\beta m}}{Z} \quad , \qquad \text{where} \quad Z = \sum_G e^{\beta m}$$

Higher values of β correspond to denser networks. To calculate Z, we write $m = \sum_{i<j} A_{ij}$ and the partition function is

$$Z = \sum_{\{A_{ij}\}} \exp\left(\beta \sum_{i<j} A_{ij}\right) = \sum_{\{A_{ij}\}} \prod_{i<j} e^{\beta A_{ij}} = \prod_{i<j} \sum_{A_{ij}=0,1} e^{\beta A_{ij}} = \prod_{i<j} \left(1 + e^{\beta}\right) = \left(1 + e^{\beta}\right)^{\binom{n}{2}}$$

where $\{A_{ij}\}$ denotes summation over all allowed values of the adjacency matrix.

$$F = \ln Z = \binom{n}{2} \ln\left(1 + e^{\beta}\right)$$

$$\langle m \rangle = \frac{\partial F}{\partial \beta} = \binom{n}{2} \frac{1}{1 + e^{-\beta}} \quad ; \qquad \beta = \ln \frac{\langle m \rangle}{\binom{n}{2} - \langle m \rangle}$$

Exponential Random Graph Models (cont'd)

Example: Simple graph, undirected and unweighted (cont'd)

One can also calculate, e.g. the probability of having an edge between a particular pair of vertices *u* and *v*,

$$p_{uv} = \langle A_{uv} \rangle = \frac{1}{Z} \sum_{\{A_{ij}\}} A_{uv} \exp\left(\beta \sum_{i<j} A_{ij}\right) = \frac{\sum_{A_{uv}=0,1} A_{uv} e^{\beta A_{uv}}}{\sum_{A_{uv}=0,1} e^{\beta A_{uv}}} = \frac{1}{1+e^{-\beta}} = \frac{\langle m \rangle}{\binom{n}{2}}$$

Here, the probability of an edge between a given pair of vertices is the same for every pair. In other words, this model is just the ordinary Poisson random graph of with $p = \langle m \rangle / \binom{n}{2}$. The *random graph*, therefore, is a special case of the more general *exponential random graph model*.

Exponential Random Graph Models (cont'd)

Example: Undirected and unweighted graph, each vertex i has its degree k_i

Here, the Hamiltonian, $H = \sum_i \beta_i k_i$, and $k_i = \sum_j A_{ij}$ and hence the Hamiltonian becomes

$$H = \sum_i \beta_i A_{ij} = \sum_{i<j} \beta_i A_{ij} + \sum_{i>j} \beta_i A_{ij} = \sum_{i<j} \beta_i A_{ij} + \sum_{i<j} \beta_j A_{ji} = \sum_{i<j} (\beta_i + \beta_j) A_{ij}$$

The partition function takes the form

$$Z = \sum_{\{A_{ij}\}} \exp\left(\sum_{i<j} (\beta_i + \beta_j) A_{ij}\right) = \prod_{i<j} \sum_{A_{ij}=0,1} e^{(\beta_i + \beta_j) A_{ij}} = \prod_{i<j} \left(1 + e^{(\beta_i + \beta_j)}\right)$$

and the probability of an edge between vertices *u* and *v* is

$$p_{uv} = \langle A_{uv} \rangle = \frac{1}{Z} \sum_{\{A_{ij}\}} A_{uv} \exp\left(\sum_{i<j} (\beta_i + \beta_j) A_{ij}\right) = \frac{\sum_{A_{uv}=0,1} A_{uv} e^{(\beta_u + \beta_v) A_{uv}}}{\sum_{A_{uv}=0,1} e^{(\beta_u + \beta_v) A_{uv}}} = \frac{1}{1 + e^{-(\beta_u + \beta_v)}}$$

Exponential Random Graph Models (cont'd)

Example: Undirected and unweighted graph, each vertex i has its degree k_i (cont'd)

If it is a sparse network, $p_{uv} \ll 1$, or $e^{-(\beta_u+\beta_v)} \gg 1$, which means $p_{uv} \approx e^{\beta_u}e^{\beta_v}$.
Now,

$$\langle k_u \rangle = \sum_v p_{uv} = e^{\beta_u} \sum_v e^{\beta_v}$$

so that

$$e^{\beta_u} = C\langle k_u \rangle\,;\ \ C = 1/\sum_v e^{\beta_v}$$

Thus, $p_{uv} = C^2\langle k_u \rangle\langle k_v \rangle$ and since we begin with $2\langle m \rangle = \sum_u k_u = \sum_{uv} p_{uv}$, we therefore have

$$p_{uv} = \frac{\langle k_u \rangle\langle k_v \rangle}{2\langle m \rangle}$$

which is the random model that we have seen before.