University of Stuttgart
Germany

Complex Network Systems

Small worlds

Ilche Georgievski
ilche.georgievski@iaas.uni-stuttgart.de
Room: U38 0.353

2019/2020
Winter

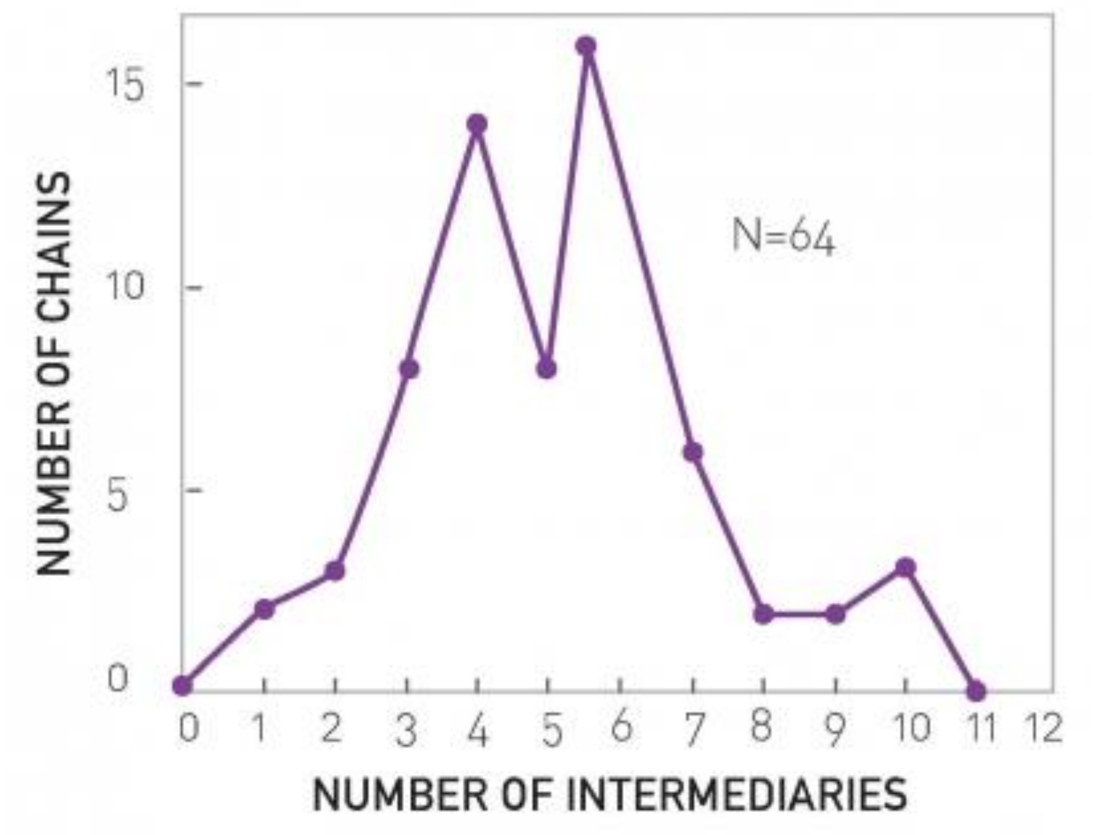
SMALL-WORLD PHENOMENON

Milgram’s experiment

Stanley, Milgram, The Small-World Problem, *Psychology Today,* May 1967, 1(1): 61-67.

Instructions:
Given a target individual (stockbroker in Boston), pass the message to a person you correspond with who is “closest” to the target.

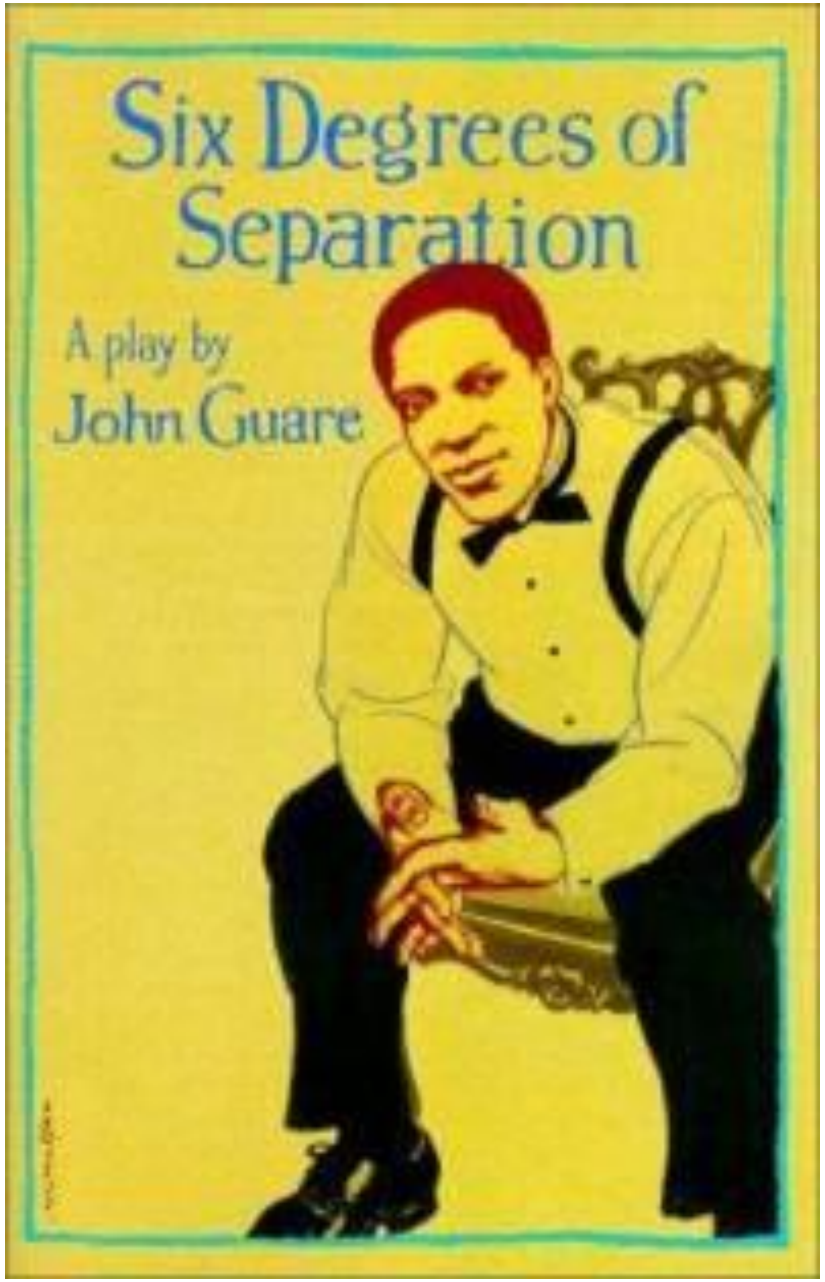
Milgram's experiment



Outcome:
20% of initiated chains reached target

average chain length = 6.5
mean chain length = 5.2

Stanley, Milgram, The Small-World Problem, *Psychology Today,* May 1967, 1(1): 61-67.



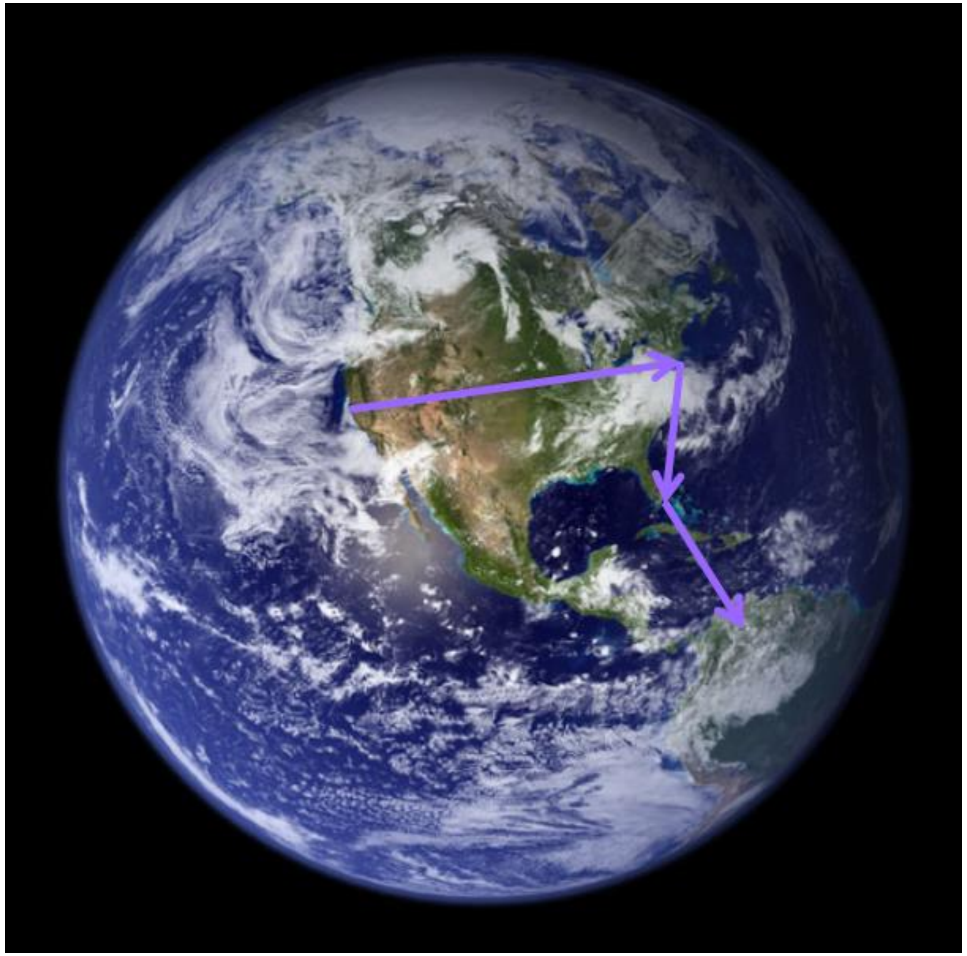
"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice…. It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. How every person is a new door, opening up into other worlds."

Interpreting Milgram's experiment

Is 6 is a ***surprising*** number?

In the 1960s? Today? Why?

Email experiment



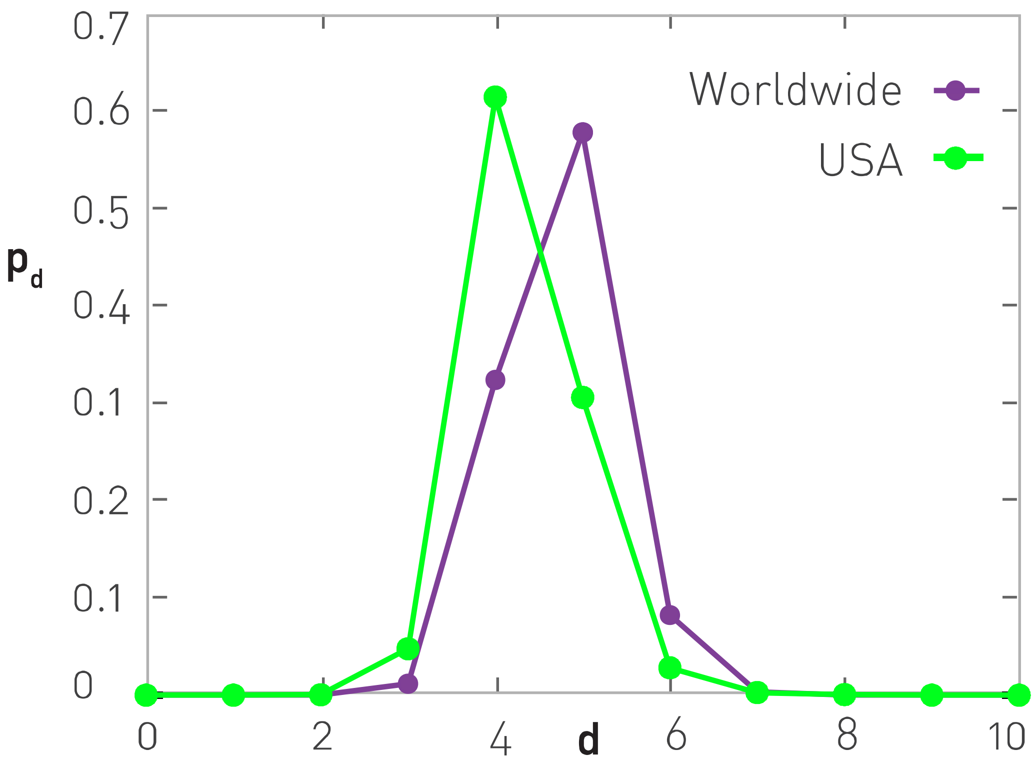
18 targets
13 different countries

60 000+ participants
24 163 message chains
384 reached their targets

average path length = 4.0

Dodds, Peter Sheridan, Muhamad, Roby, Watts, Duncan J., An Experimental Study of Search in Global Social Networks, *Science,* 08 Aug 2003 : 827-829.

Facebook experiment



721 million active users
68 billion symmetric friendship links

average path length = 4.74

L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. In ACM Web Science 2012: Conference Proceedings, pages 45−54. ACM Press, 2012.

Interpreting Milgram's experiment

- If social networks were random…?
  - ~1000 acquaintances/person
  - ~1000 choices 1st link
  - $\sim 1000^2$ = 1,000,000 potential 2nd degree neighbors
  - $\sim 1000^3$ = 1,000,000,000 potential 3rd degree neighbors

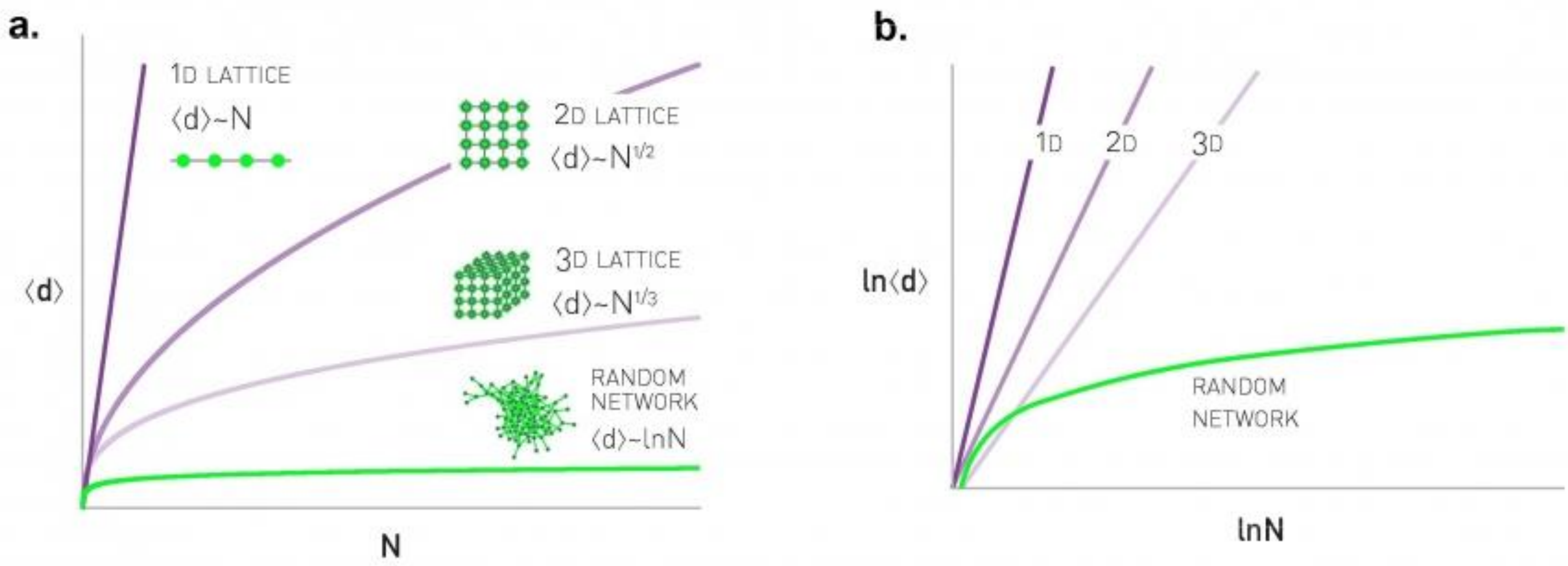
de Sola Pool, Ithiel and Kochen, Malfred, Contacts and Influence, *Social Networks,* 1(1978/79): 5-51.

Question

- If the network were completely cliquish, that is all your friends of friends were also directly your friends, what would be true:
  a) None of your friendship edges would be part of a triangle (closed triad)
  V b) It would be impossible to reach any node outside the clique by following directed edges
  c) Your shortest path to your friends' friends would be 2

Why are small worlds surprising?

<d> is diameter



In regular lattices path lengths are significantly longer than in a random network

Average path length or diameter depends logarithmically on the system size

Accuracy of distances

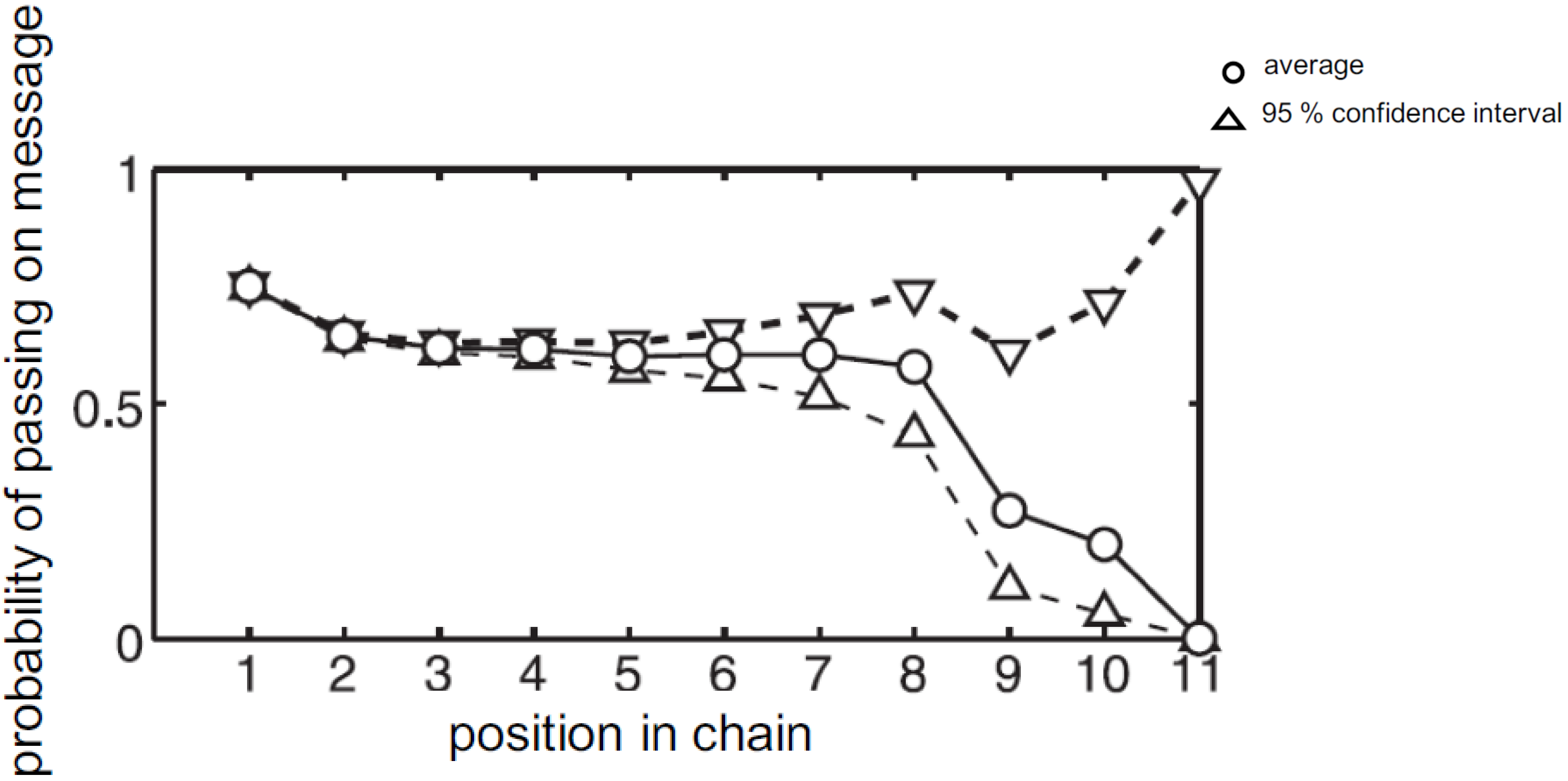
- Is 6 an ***accurate*** number?

    no, there are survival bias

- What bias is introduced by uncompleted chains?
    – are longer or shorter chains more likely to be completed?

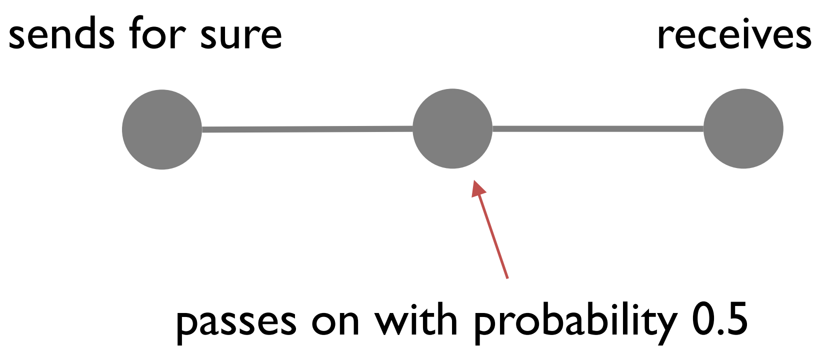
Dodds, Peter Sheridan, Muhamad, Roby, Watts, Duncan J., An Experimental Study of Search in Global Social Networks, *Science,* 08 Aug 2003 : 827-829.

Attrition



Question

– If each person in the chain has 0.5 probability of passing the message on, what is the likelihood of a chain of length 2 being completed?



chain of length 2

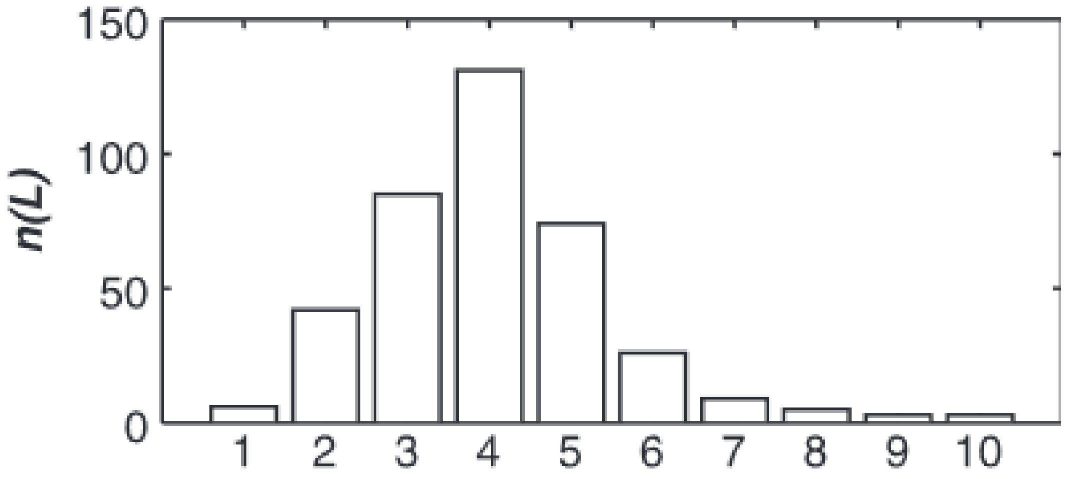
Question

– If each person in the chain has 0.5 probability of passing the message on, what is the likelihood of a chain of length 5 being completed?

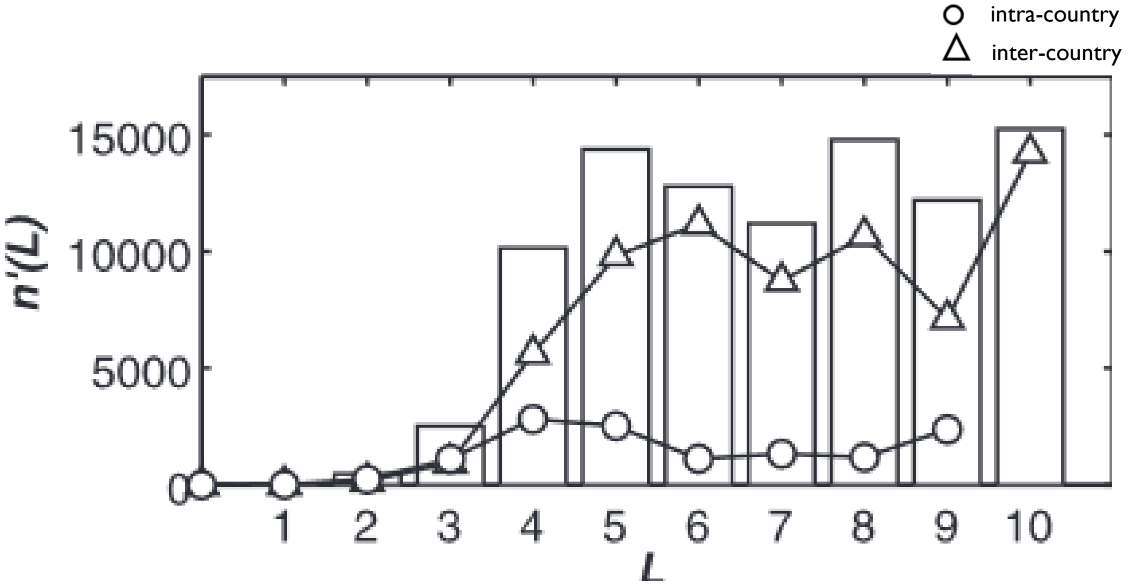
a) 1/2
b) 1/4
c) 1/8
d) 1/16

Dodds, Peter Sheridan, Muhamad, Roby, Watts, Duncan J., An Experimental Study of Search in Global Social Networks, *Science,* 08 Aug 2003 : 827-829.

Estimating the true distance



observed chain lengths



'recovered' histogram of path lengths

Accuracy of distances

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

- Do people find the **shortest** paths?
  - accuracy of small-world chains in social networks by Killworth et al.: less than optimal choice for the next link in chain is made ½ of the time

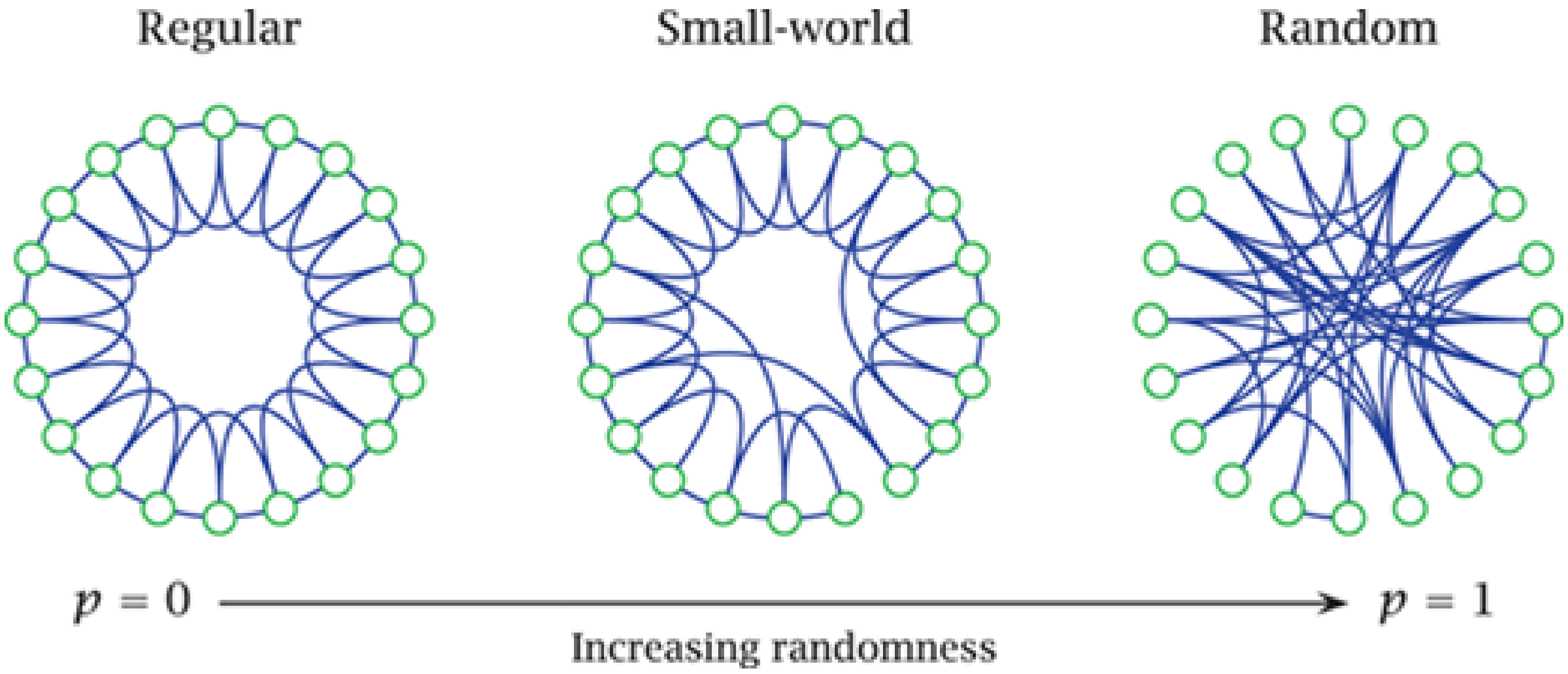
Watts-Strogatz model

SMALL-WORLD MODEL

Models

- Watts and Strogatz
  – Clustering and short paths
- Kleinberg
  – Geographical structure
- Watts, Dodds and Newman
  – Hierarchical structure

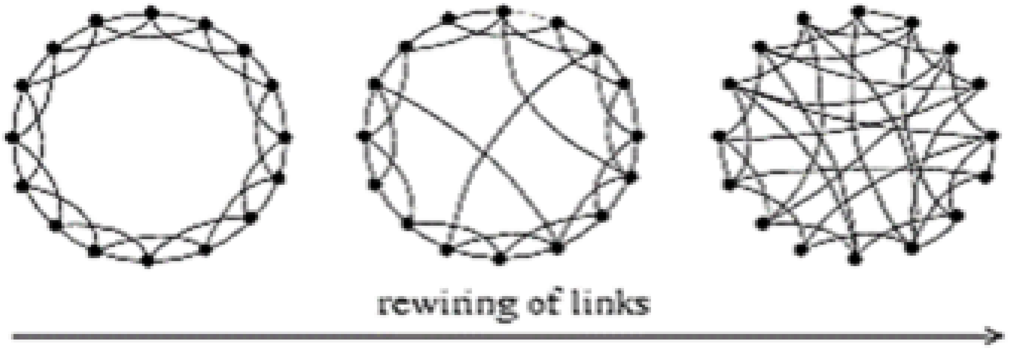
Watts-Strogatz model



Olofsen, Erik and Dahan, Albert, Big Brain, Small World?, *Anesthesiology,* 2015, 122(1):8-11.

Generating a small-world graph

- The lattice has a high, but fixed, clustering coefficient
- The lattice has a high average path length
- The model begins with a regular lattice and starts adding random edges (through rewiring)
  - **Rewiring**: take an edge, change one of its end-points randomly



Constructing small-world networks

Algorithm 4.1 Small-World Generation Algorithm

Require: Number of nodes $|V|$, mean degree k, parameter p

return A small-world graph $G(V, E)$

G = A regular ring lattice with $|V|$ nodes and degree k

for node v_i (starting from v_1), and all edges $e(v_i, v_j)$, $i < j$ **do**

    v_k = Select a node from V uniformly at random.

    if rewiring $e(v_i, v_j)$ to $e(v_i, v_k)$ does not create loops in the graph or multiple edges between v_i and v_k **then**

        rewire $e(v_i, v_j)$ with probability p: $E = E - \{e(v_i, v_j)\}$, $E = E \cup \{e(v_i, v_k)\}$;

    end if

end for

Return $G(V, E)$

As in many network generation algorithms:

- Disallow loops
- Disallow multiple edges

SMALL-WORLD MODEL PROPERTIES

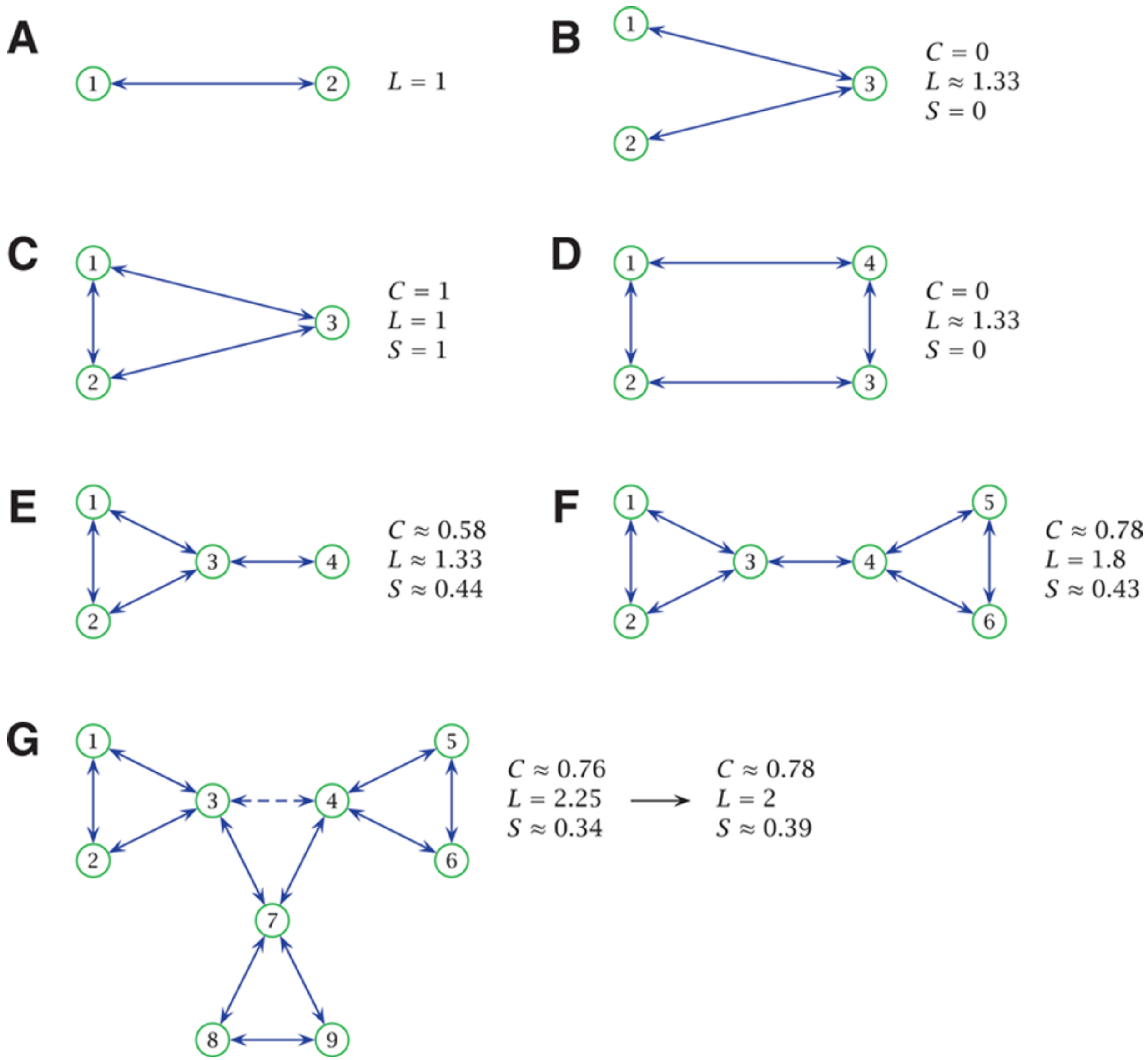
Degree distribution

- What can we say about distribution of degrees in small-world graphs?
- In the graph generated by the small-world model, most nodes have <u>similar degrees</u> due to the underlying lattice

Regular lattice vs. random graph

- Regular lattice
  - Clustering coefficient: high
  - Average path length: high

- Random graph
  - Clustering coefficient: low
  - Average path length: ok



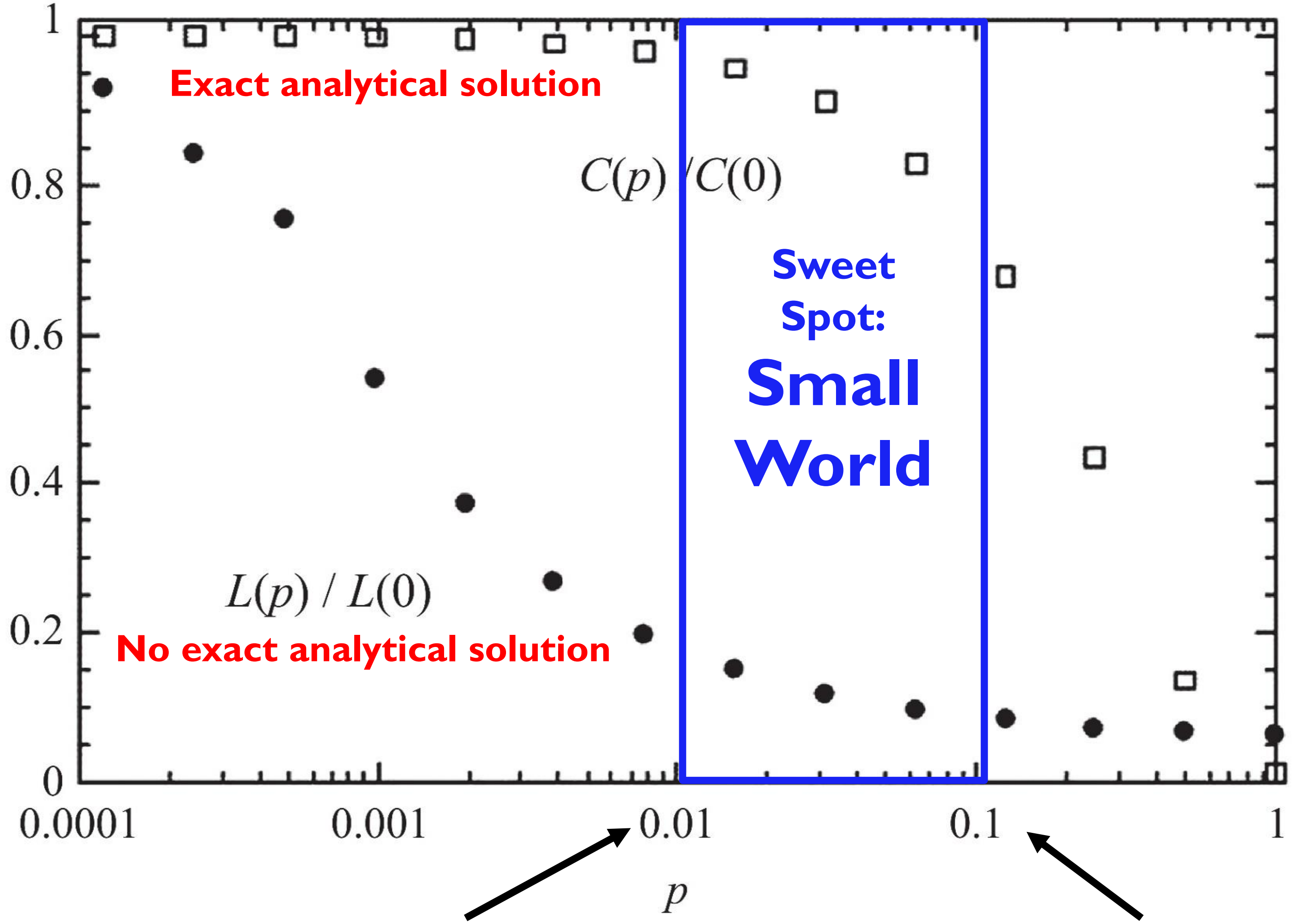
Olofsen, Erik and Dahan, Albert, Big Brain, Small World?, *Anesthesiology*, 2015, 122(1):8-11.

What happens in between?

- Small average path length means small clustering?
- Large average path length means large clustering?

- Numerical simulation
  - As we increase p from 0 to 1, we observe
    - Fast decrease of average distance
    - Slow decrease in clustering coefficient

Change in clustering coefficient/Avg. path length



1% of links rewired

10% of links rewired

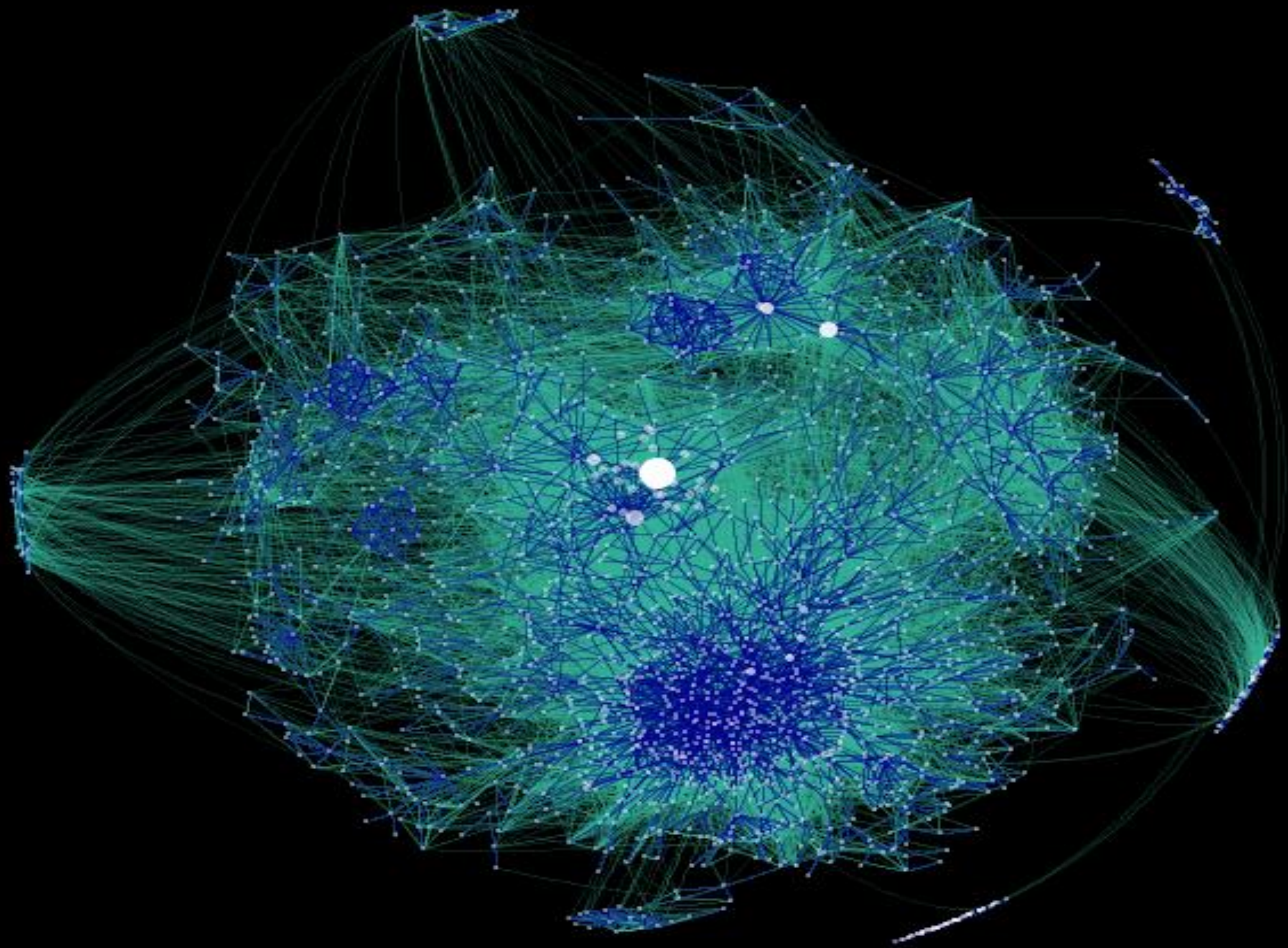
Modelling with the small-world model

- Given a real-world network in which average degree k and clustering coefficient C are given
  - we set $C(p) = C$
  - we calculate p using

$$C(p) \approx (1-p)^3 C(0)$$

- Given $p, k,$ and $n = |V|$, we can simulate the small-world model

Image by **Matthew Hurst**
Blogosphere



Network	N	L	‹k›	‹d›	d_{max}	$\ln N / \ln \langle k \rangle$
Internet	192,244	609,066	6.34	6.98	26	6.58
WWW	325,729	1,497,134	4.60	11.27	93	8.31
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	93,437	8.08	5.35	15	4.81
Actor Network	702,388	29,397,908	83.71	3.91	14	3.04
Citation Network	449,673	4,707,958	10.43	11.21	42	5.55
E. Coli Metabolism	1,039	5,802	5.58	2.98	8	4.04
Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Sources

- Adamic, L. Social and Information Network Analysis, CS224W, Stanford University (2015), http://snap.stanford.edu/class/cs224w-2015/
- Zafarani, R., Abbasi, M. A. and Liu, H. *Social Media Mining: An Introduction*, Cambridge University Press, 2014.
- Barabási, A. Network Science, http://networksciencebook.com