Models of Complex Networks

Last updated: 2023/08/22, 11:48:21 EDT

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number,
2023–2024 | @pocsvox

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

Outline

Models

Some important models:

1. Generalized random networks
2. Scale-free networks
3. Small-world networks
4. Statistical generative models (p^*)
5. Generalized affiliation networks

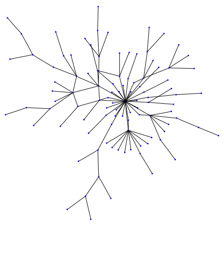
Models

1. Generalized random networks:

- Arbitrary degree distribution P_k.
- Wire nodes together randomly.
- Create ensemble to test deviations from randomness.
- Interesting, applicable, rich mathematically.
- Much fun to be had with these guys...

Models

2. 'Scale-free networks':



γ = 2.5
$\langle k \rangle$ = 1.8
$N = 150$

- Due to Barabasi and Albert [2]
- Generative model
- Preferential attachment model with growth
- P[attachment to node i] $\propto k_i^{\alpha}$.
- Produces $P_k \sim k^{-\gamma}$ when $\alpha = 1$.
- Trickiness: other models generate skewed degree distributions...

Models

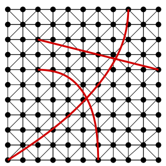
3. Small-world networks

- Due to Watts and Strogatz [18]

Two scales:
- local regularity (high clustering—an individual's friends know each other)
- global randomness (shortcuts).

Strong effects:

- Shortcuts make world 'small'
- Shortcuts allow disease to jump
- Facilitates synchronization [8]



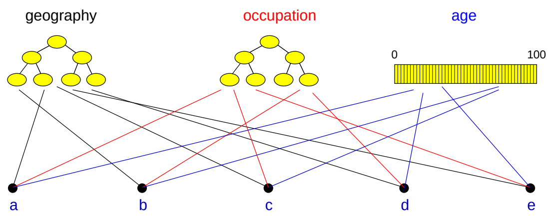
Models

4. Generative statistical models

- Idea is to realize networks based on certain tendencies:
  - Clustering (triadic closure)..
  - Types of nodes that like each other..
  - Anything really...
- Use statistical methods to estimate 'best' values of parameters.
- Drawback: parameters are not real, measurable quantities.
- Non-mechanistic and blackboxish.
- c.f., temperature in statistical mechanics.

Models

5. Generalized affiliation networks

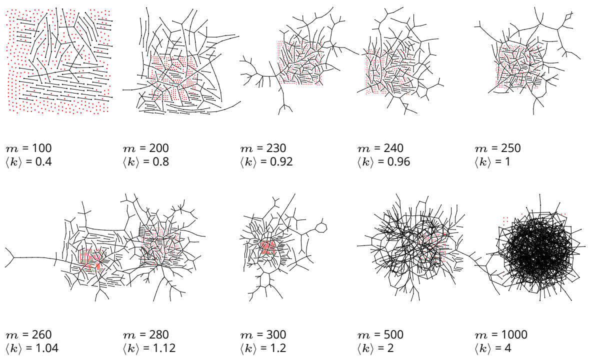


- Blau & Schwartz [3], Simmel [15], Breiger [4], Watts *et al.* [17]

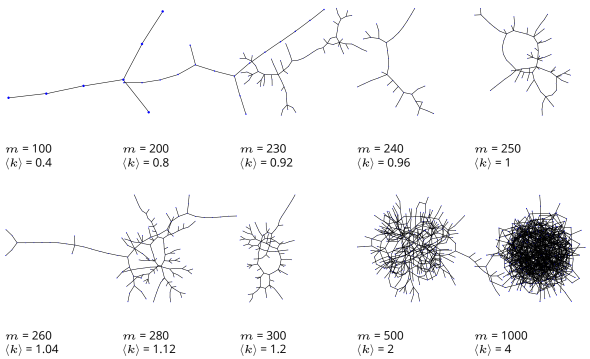
Pure, abstract random networks:

- Consider set of all networks with N labelled nodes and m edges.
- Horribly, there are $\binom{\binom{N}{2}}{m}$ of them.
- Standard random network = randomly chosen network from this set.
- To be clear: each network is equally probable.
- Known as Erdős-Rényi random networks
- Key structural feature of random networks is that they locally look like branching networks
- (No small cycles and zero clustering).

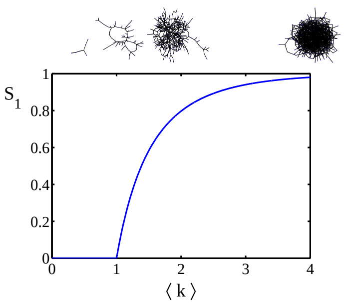
Random networks: examples for N=500



Random networks: largest components



Giant component:



- S_1 = fraction of nodes in largest component.
- Old school phase transition.
- Key idea in modeling contagion.

Properties

But:

- Erdős-Rényi random networks are a *mathematical construct*.
- Real networks are a microscopic subset of all networks...
- ex: 'Scale-free' networks are growing networks that form according to a plausible mechanism.

But but:

- Randomness is out there, just not to the degree of a completely random network.

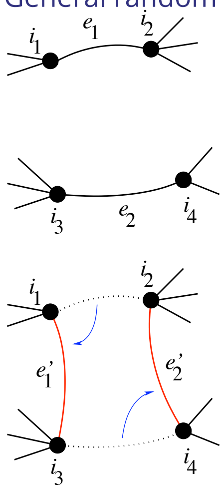
General random networks

- So... standard random networks have a Poisson degree distribution
- Can happily generalize to arbitrary degree distribution P_k.
- Also known as the configuration model. [12]
- Can generalize construction method from ER random networks.
- Assign each node a weight w from some distribution and form links with probability

$$P(\text{link between } i \text{ and } j) \propto w_i w_j.$$

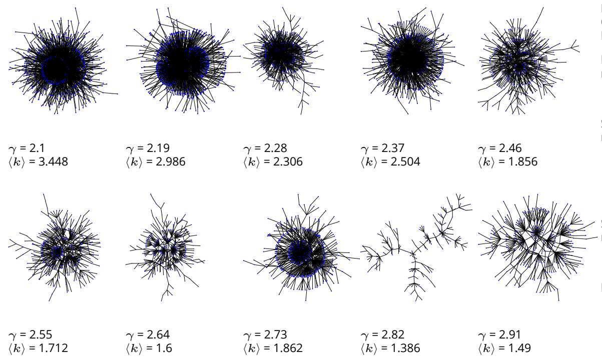
- A more useful way:
  1. Randomly wire up (and rewire) already existing nodes with fixed degrees.
  2. Examine mechanisms that lead to networks with certain degree distributions.

General random rewiring algorithm



- Randomly choose two edges.
  (Or choose problem edge and a random edge)
- Check to make sure edges are disjoint.
- Rewire one end of each edge.
- Node degrees do not change.
- Works if e_1 is a self-loop or repeated edge.
- Same as finding on/off/on/off 4-cycles. and rotating them.

Random networks: largest components



The edge-degree distribution:

- The degree distribution P_k is fundamental for our description of many complex networks
- A related key distribution:
  R_k = probability that a friend of a random node has k other friends.
- $$R_k = \frac{(k+1)P_{k+1}}{\sum_{k'=0}(k'+1)P_{k'+1}} = \frac{(k+1)P_{k+1}}{\langle k \rangle}$$
- Natural question: what's the expected number of other friends that one friend has?
- Find

$$\langle k \rangle_R = \frac{1}{\langle k \rangle}\left(\langle k^2 \rangle - \langle k \rangle\right)$$

- True for all random networks, independent of degree distribution.

Giant component condition

- If:

$$\langle k \rangle_R = \frac{1}{\langle k \rangle}\left(\langle k^2 \rangle - \langle k \rangle\right) > 1$$

then our random network has a giant component.

- Exponential explosion in number of nodes as we move out from a random node.
- Number of nodes expected at n steps:

$$\langle k \rangle \cdot \langle k \rangle_R^{n-1} = \frac{1}{\langle k \rangle^{n-2}}\left(\langle k^2 \rangle - \langle k \rangle\right)^{n-1}$$

- We'll see this again for contagion models...

Mild weirdness...

- Average # friends of friends per node is

$$\langle k_2 \rangle = \langle k^2 \rangle - \langle k \rangle.$$

- Average depends on the 1st and 2nd moments of P_k and not just the 1st moment.
- Three peculiarities:
  1. We might guess $\langle k_2 \rangle = \langle k \rangle(\langle k \rangle - 1)$ but it's actually $\langle k(k-1) \rangle$.
  2. If P_k has a large second moment, then $\langle k_2 \rangle$ will be big.
  3. Your friends have more friends than you...

Size distributions

The sizes of many systems' elements appear to obey an inverse power-law size distribution:

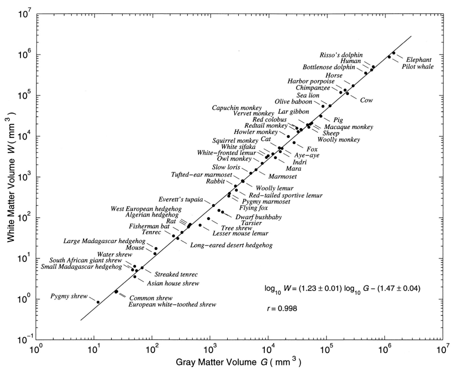
$$P(\text{size} = x) \sim c\, x^{-\gamma}$$

where $x_{\text{min}} < x < x_{\text{max}}$ and $\gamma > 1$.

- x can be continuous or discrete.
- Typically, $2 < \gamma < 3$.
- No dominant internal scale between x_{min} and x_{max}.
- If $\gamma < 3$, variance and higher moments are 'infinite'
- If $\gamma < 2$, mean and higher moments are 'infinite'
- Negative linear relationship in log-log space:

$$\log P(x) = \log c - \gamma \log x$$

A beautiful, heart-warming example:



from Zhang & Sejnowski, PNAS (2000) [19]

$\alpha \simeq 1.23$

gray matter: 'computing elements'

white matter: 'wiring'

Size distributions

Power law size distributions are sometimes called Pareto distributions after Italian scholar Vilfredo Pareto.

- Pareto noted wealth in Italy was distributed unevenly (80–20 rule).
- Term used especially by economists

Size distributions

Examples:

- Earthquake magnitude (Gutenberg Richter law): $P(M) \propto M^{-3}$
- Number of war deaths: $P(d) \propto d^{-1.8}$ [14]
- Sizes of forest fires
- Sizes of cities: $P(n) \propto n^{-2.1}$
- Number of links to and from websites

Size distributions

Examples:

- Number of citations to papers: $P(k) \propto k^{-3}$.
- Individual wealth (maybe): $P(W) \propto W^{-2}$.
- Distributions of tree trunk diameters: $P(d) \propto d^{-2}$.
- The gravitational force at a random point in the universe: $P(F) \propto F^{-5/2}$.
- Diameter of moon craters: $P(d) \propto d^{-3}$.
- Word frequency: e.g., $P(k) \propto k^{-2.2}$ (variable)

Note: Exponents range in error; see M.E.J. Newman arxiv.org/cond-mat/0412004v3

History

- Random Additive/Copying Processes involving Competition.
- Widespread: Words, Cities, the Web, Wealth, Productivity (Lotka), Popularity (Books, People, ...)
- Competing mechanisms (more trickiness)

Work of Yore

- 1924: G. Udny Yule [?]: # Species per Genus
- 1926: Lotka [10]: # Scientific papers per author (Lotka's law)
- 1953: Mandelbrot [11]: Optimality argument for Zipf's law; focus on language.
- 1955: Herbert Simon [16, 20]: Zipf's law for word frequency, city size, income, publications, and species per genus.
- 1965/1976: Derek de Solla Price [5, 13]: Network of Scientific Citations.
- 1999: Barabasi and Albert [2]: The World Wide Web, networks-at-large.

Not everyone is happy...


Mandelbrot vs. Simon:

- Mandelbrot (1953): "An Informational Theory of the Statistical Structure of Languages" [11]
- Simon (1955): "On a class of skew distribution functions" [16]
- Mandelbrot (1959): "A note on a class of skew distribution function: analysis and critique of a paper by H. A. Simon"
- Simon (1960): "Some further notes on a class of skew distribution functions"

Not everyone is happy... (cont.)

Mandelbrot vs. Simon:

- Mandelbrot (1961): "Final note on a class of skew distribution functions: analysis and critique of a model due to H.A. Simon"
- Simon (1961): "Reply to 'final note' by Benoit Mandelbrot"
- Mandelbrot (1961): "Post scriptum to 'final note'"
- Simon (1961): "Reply to Dr. Mandelbrot's post scriptum"

Not everyone is happy... (cont.)

Mandelbrot:

"We shall restate in detail our 1959 objections to Simon's 1955 model for the Pareto-Yule-Zipf distribution. Our objections are valid quite irrespectively of the sign of p-1, so that most of Simon's (1960) reply was irrelevant."

Simon:

"Dr. Mandelbrot has proposed a new set of objections to my 1955 models of the Yule distribution. Like his earlier objections, these are invalid."

Essential Extract of a Growth Model

Random Competitive Replication (RCR):

1. Start with 1 element of a particular flavor at $t = 1$
2. At time $t = 2, 3, 4, \ldots$, add a new element in one of two ways:
   - With probability ρ, create a new element with a new flavor
     ➤ Mutation/Innovation
   - With probability $1 - \rho$, randomly choose from all existing elements, and make a copy.
     ➤ Replication/Imitation
   - Elements of the same flavor form a group

Random Competitive Replication

Example: Words in a text

- Consider words as they appear sequentially.
- With probability ρ, the next word has not previously appeared
  ➤ Mutation/Innovation
- With probability $1 - \rho$, randomly choose one word from all words that have come before, and reuse this word
  ➤ Replication/Imitation
- Please note: authors do not do this...

Random Competitive Replication

- Competition for replication between elements is random
- Competition for growth between groups is not random
- Selection on groups is biased by size
- Rich-gets-richer story
- Random selection is easy
- No great knowledge of system needed

Random Competitive Replication

- After some thrashing around, one finds:

$$P_k \propto k^{-\frac{(2-\rho)}{(1-\rho)}} = k^{-\gamma}$$

$$\boxed{\gamma = 1 + \frac{1}{(1-\rho)}}$$

- See γ is governed by rate of new flavor creation, ρ.

Evolution of catch phrases

- Yule's paper (1924) [?]:
  "A mathematical theory of evolution, based on the conclusions of Dr J. C. Willis, F.R.S."
- Simon's paper (1955) [16]:
  "On a class of skew distribution functions" (snore)
- Price's term: Cumulative Advantage

Evolution of catch phrases

- Robert K. Merton: the Matthew Effect
- Studied careers of scientists and found credit flowed disproportionately to the already famous

  From the Gospel of Matthew:
  "For to every one that hath shall be given...
  (Wait! There's more....)
  but from him that hath not, that also which he seemeth to have shall be taken away.
  And cast the worthless servant into the outer darkness; there men will weep and gnash their teeth."

Evolution of catch phrases

Merton was a catchphrase machine:

1. self-fulfilling prophecy
2. role model
3. unintended (or unanticipated) consequences
4. focused interview → focus group

And just to rub it in...

Merton's son, Robert C. Merton, won the Nobel Prize for Economics in 1997.

Evolution of catch phrases

- Barabási and Albert [2]—thinking about the Web
- Independent reinvention of a version of Simon and Price's theory for networks
- Another term: "Preferential Attachment"
- Basic idea: a new node arrives every discrete time step and connects to an existing node i with probability $\propto k_i$.
- Connection:
  Groups of a single flavor $\sim$ edges of a node
- Small hitch: selection mechanism is now non-random
- Solution: Connect to a random node (easy)
- + Randomly connect to the node's friends (also easy)
- Scale-free networks = food on the table for physicists

Scale-free networks

- Networks with power-law degree distributions have become known as scale-free networks.
- Scale-free refers specifically to the degree distribution having a power-law decay in its tail:

$$P_k \sim k^{-\gamma} \text{ for 'large' } k$$

- Please note: not every network is a scale-free network...

Scale-free networks

- Term 'scale-free' is somewhat confusing...
- Scale-free networks are not fractal in any sense.
- Usually talking about networks whose links are abstract, relational, informational, ...(non-physical)
- Main reason is link cost.
- Primary example: hyperlink network of the Web
- Much arguing about whether or networks are 'scale-free' or not...

Scale-free networks

The big deal:

- We move beyond describing networks to finding mechanisms for why certain networks arise.

A big deal for scale-free networks:

- How does the exponent γ depend on the mechanism?
- Do the mechanism's details matter?
- We know they do for Simon's model...

Real data (eek!)

From Barabási and Albert's original paper [2]:

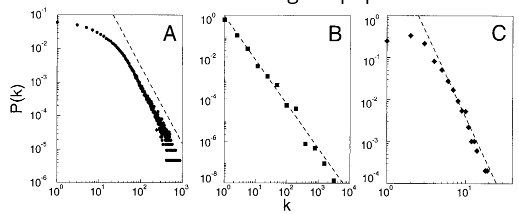


Fig. 1. The distribution function of connectivities for various large networks. **(A)** Actor collaboration graph with $N = 212{,}250$ vertices and average connectivity $\langle k \rangle = 28.78$. **(B)** WWW, $N = 325{,}729$, $\langle k \rangle = 5.46$ (6). **(C)** Power grid data, $N = 4941$, $\langle k \rangle = 2.67$. The dashed lines have slopes (A) $\gamma_{\text{actor}} = 2.3$, (B) $\gamma_{\text{www}} = 2.1$ and (C) $\gamma_{\text{power}} = 4$.

- But typically for real networks: $2 < \gamma < 3$.
- (Plot C is on the bogus side of things...)

Generalized model

Fooling with the mechanism:

- 2001: Redner & Krapivsky (RK) [9] explored the general attachment kernel:

$$\textbf{Pr}(\text{attach to node } i) \propto A_k = k_i^{\nu}$$

  where A_k is the attachment kernel and $\nu > 0$.
- RK also looked at changing very subtle details of the attachment kernel.
- e.g., keep $A_k \sim k$ for large k but tweak A_k for low k.
- RK's approach is to use rate equations.

Universality?

- Consider $A_1 = \alpha$ and $A_k = k$ for $k \geq 2$.
- Some unsettling calculations leads to $P_k \sim k^{-\gamma}$ where

$$\gamma = 1 + \frac{1 + \sqrt{1 + 8\alpha}}{2}.$$

- We then have

$$0 \leq \alpha < \infty \Rightarrow 2 \leq \gamma < \infty$$

- Craziness...

Sublinear attachment kernels

- Rich-get-somewhat-richer:

$$A_k \sim k^{\nu} \text{ with } 0 < \nu < 1.$$

- General finding by Krapivsky and Redner: [9]

$$P_k \sim k^{-\nu} e^{-c_1 k^{1-\nu} + \text{correction terms}}.$$

- Weibull distribution*ish* (truncated power laws).
- Universality: now details of kernel do not matter.

Superlinear attachment kernels

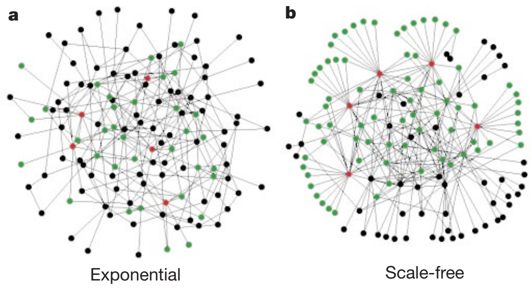
- Rich-get-much-richer:

$$A_k \sim k^{\nu} \text{ with } \nu > 1.$$

- Now a winner-take-all mechanism.
- One single node ends up being connected to almost all other nodes.
- For $\nu > 2$, all but a finite # of nodes connect to one node.

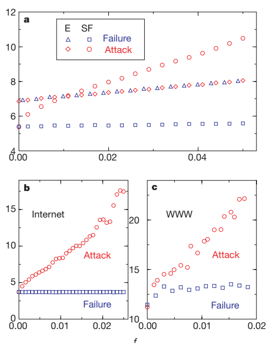
Robustness

- Standard random networks (Erdős-Rényi) versus Scale-free networks



from Albert et al., 2000 "Error and attack tolerance of complex networks"[1]

Robustness



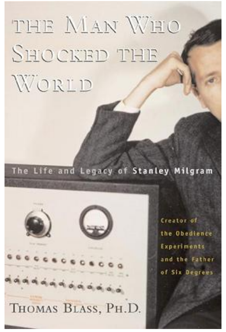
from Albert et al., 2000

- Plots of network diameter as a function of fraction of nodes removed
- Erdős-Rényi versus scale-free networks
- blue symbols = random removal
- red symbols = targeted removal (most connected first)

Robustness

- Scale-free networks are thus robust to random failures yet fragile to targeted ones.
- All very reasonable: Hubs are a big deal.
- But: next issue is whether hubs are vulnerable or not.
- Representing all webpages as the same size node is obviously a stretch (e.g., google vs. a random person's webpage)
- Most connected nodes are either:
  1. Physically larger nodes that may be harder to 'target'
  2. or subnetworks of smaller, normal-sized nodes.
- Need to explore cost of various targeting schemes.

Milgram's social search experiment (1960s)



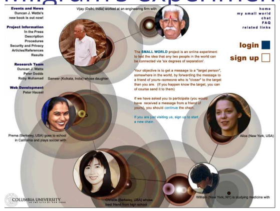
http://www.stanleymilgram.com

- Target person = Boston stockbroker.
- 296 senders from Boston and Omaha.
- 20% of senders reached target.
- chain length $\simeq 6.5$.

Popular terms:

- The Small World Phenomenon;
- "Six Degrees of Separation."

Milgram's experiment with e-mail [6]



Participants:

- 60,000+ people in 166 countries
- 24,000+ chains
- Big media boost...

18 targets in 13 countries including

- a professor at an Ivy League university,
- an archival inspector in Estonia,
- a technology consultant in India,
- a policeman in Australia,
- a potter in New Zealand,
- a veterinarian in the Norwegian army.

Social search—the Columbia experiment

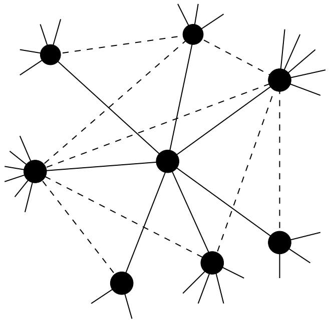
The world is smaller:

- $\langle L \rangle = 4.05$ for all completed chains
- L_* = Estimated 'true' median chain length (zero attrition)
- Intra-country chains: $L_* = 5$
- Inter-country chains: $L_* = 7$
- All chains: $L_* = 7$
- c.f. Milgram (zero attrition): $L_* \simeq 9$

Previous work—short paths

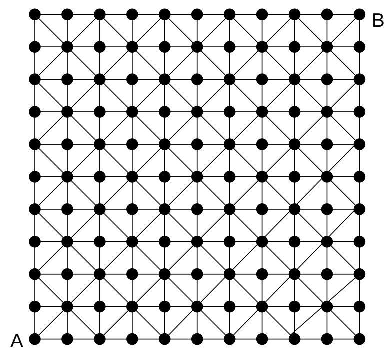
- Connected random networks have short average path lengths:
$$\langle d_{AB} \rangle \sim \log(N)$$
N = population size,
d_{AB} = distance between nodes A and B.
- But: social networks aren't random...

Previous work—short paths



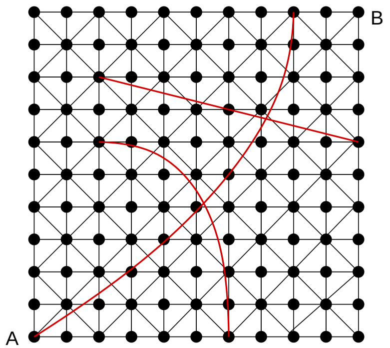
- Need "clustering" (your friends are likely to know each other):
- Randomly connecting people gives short path lengths ... weird.

Non-randomness gives clustering



$d_{AB} = 10 \rightarrow$ too many long paths.

Randomness + regularity



$d_{AB} = 10$ without random paths
$d_{AB} = 3$ with random paths

$\langle d \rangle$ decreases overall

Theory of Small-World networks

Introduced by
Watts and Strogatz (Nature, 1998) [18]
"Collective dynamics of 'small-world' networks."

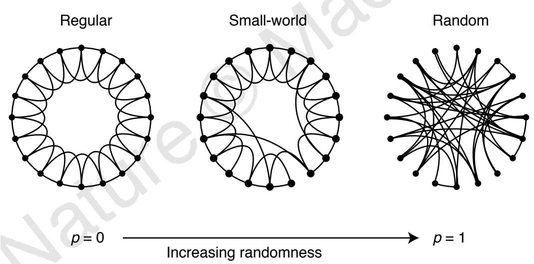
Small-world networks are found everywhere:

- neural network of C. elegans,
- semantic networks of languages,
- actor collaboration graph,
- food webs,
- social networks of comic book characters,...

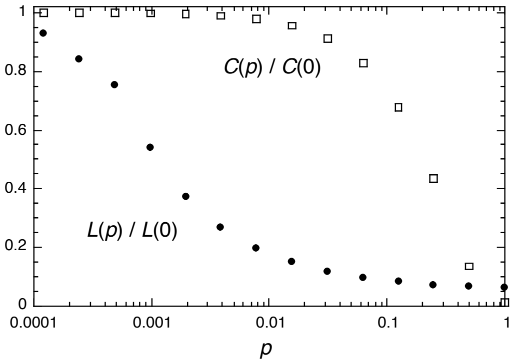
Very weak requirements:

- local regularity + random short cuts

Toy model



The structural small-world property



The structural small-world property

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

Characteristic path length L and clustering coefficient C for three real networks, compared to random graphs with the same number of vertices (n) and average number of edges per vertex (k). (Actors: $n = 225{,}226$, $k = 61$. Power grid: $n = 4{,}941$, $k = 2.67$. *C. elegans*: $n = 282$, $k = 14$.) The graphs are defined as follows. Two actors are joined by an edge if they have acted in a film together. We restrict attention to the giant connected component[16] of this graph, which includes ~90% of all actors listed in the Internet Movie Database (available at http://us.imdb.com), as of April 1997. For the power grid, vertices represent generators, transformers and substations, and edges represent high-voltage transmission lines between them. For *C. elegans*, an edge joins two neurons if they are connected by either a synapse or a gap junction. We treat all edges as undirected and unweighted, and all vertices as identical, recognizing that these are crude approximations. All three networks show the small-world phenomenon: $L \gtrsim L_{random}$ but $C \gg C_{random}$.

Previous work—finding short paths

But are these short cuts findable?

No!

Nodes cannot find each other quickly with any local search method.

- Jon Kleinberg (Nature, 2000) [7] "Navigation in a small world."
- Only certain networks are navigable
- So what's special about social networks?

The model

One approach: incorporate identity.
(See "Identity and Search in Social Networks." Science, 2002, Watts, Dodds, and Newman [17])

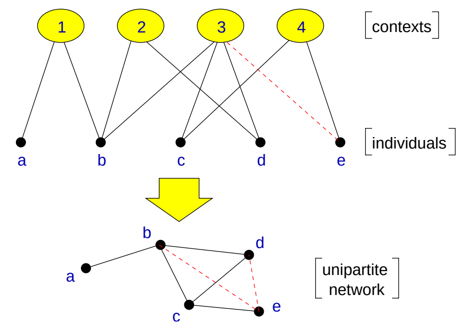
Identity is formed from attributes such as:

- Geographic location
- Type of employment
- Religious beliefs
- Recreational activities.

Groups are formed by people with at least one similar attribute.

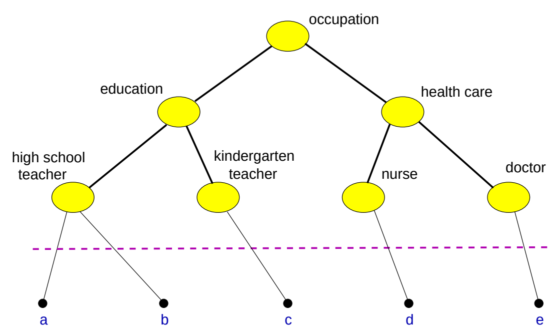
Attributes ⇔ Contexts ⇔ Interactions ⇔ Networks.

Social distance—Bipartite affiliation networks

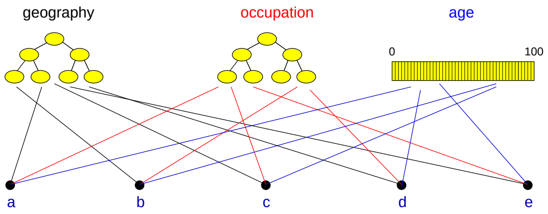


Bipartite affiliation networks: boards and directors, movies and actors.

Social distance as a function of identity



Homophily



(Blau & Schwartz, Simmel, Breiger)

- Networks built with 'birds of a feather...' are searchable.
- Attributes ⇔ Contexts ⇔ Interactions ⇔ Networks.

Social Search—Real world uses

- Tagging: e.g., Flickr induces a network between photos
- Search in organizations for solutions to problems
- Peer-to-peer networks
- Synchronization in networked systems
- Motivation for search matters...

References I

[1] R. Albert, H. Jeong, and A.-L. Barabási.
Error and attack tolerance of complex networks.
Nature, 406:378–382, 2000. pdf

[2] A.-L. Barabási and R. Albert.
Emergence of scaling in random networks.
Science, 286:509–511, 1999. pdf

[3] P. M. Blau and J. E. Schwartz.
Crosscutting Social Circles.
Academic Press, Orlando, FL, 1984.

[4] R. L. Breiger.
The duality of persons and groups.
Social Forces, 53(2):181–190, 1974. pdf

[5] D. J. de Solla Price.
Networks of scientific papers.
Science, 149:510–515, 1965. pdf

References II

[6] P. S. Dodds, R. Muhamad, and D. J. Watts.
An experimental study of search in global social networks.
Science, 301:827–829, 2003. pdf

[7] J. Kleinberg.
Navigation in a small world.
Nature, 406:845, 2000. pdf

[8] G. Korniss, M. A. Novotny, H. Guclu, Z. Toroczkai, and P. A. Rikvold.
Suppressing roughness of virtual times in parallel discrete-event simulations.
Science, 299:677–679, 2003. pdf

[9] P. L. Krapivsky and S. Redner.
Organization of growing random networks.
Phys. Rev. E, 63:066123, 2001. pdf

References III

[10] A. J. Lotka.
The frequency distribution of scientific productivity.
Journal of the Washington Academy of Science, 16:317–323, 1926.

[11] B. B. Mandelbrot.
An informational theory of the statistical structure of languages.
In W. Jackson, editor, Communication Theory, pages 486–502. Butterworth, Woburn, MA, 1953. pdf

[12] M. E. J. Newman.
The structure and function of complex networks.
SIAM Rev., 45(2):167–256, 2003. pdf

References IV

[13] D. D. S. Price.
A general theory of bibliometric and other cumulative advantage processes.
Journal of the American Society for Information Science, pages 292–306, 1976. pdf

[14] L. F. Richardson.
Variation of the frequency of fatal quarrels with magnitude.
J. Amer. Stat. Assoc., 43:523–546, 1949.

[15] G. Simmel.
The number of members as determining the sociological form of the group. I.
American Journal of Sociology, 8:1–46, 1902.

References V

[16] H. A. Simon.
On a class of skew distribution functions.
Biometrika, 42:425–440, 1955. pdf

[17] D. J. Watts, P. S. Dodds, and M. E. J. Newman.
Identity and search in social networks.
Science, 296:1302–1305, 2002. pdf

[18] D. J. Watts and S. J. Strogatz.
Collective dynamics of 'small-world' networks.
Nature, 393:440–442, 1998. pdf

[19] K. Zhang and T. J. Sejnowski.
A universal scaling law between gray matter and white matter of cerebral cortex.
Proceedings of the National Academy of Sciences, 97:5621–5626, 2000. pdf

References VI

[20] G. K. Zipf.
Human Behaviour and the Principle of Least-Effort.
Addison-Wesley, Cambridge, MA, 1949.