

Small-world networks

Last updated: 2024/11/14, 19:57:24 EST

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number, 2024–2025

Prof. Peter Sheridan Dodds


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

 Licensed under the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

The PoCSverse
Small-world networks
1 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

Social Search

A small slice of the pie:



 **Q.** Can people pass messages between distant individuals using only their existing social connections?

 **A.** Apparently yes ...

The PoCSverse
Small-world networks
6 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

Six Degrees of Kevin Bacon:




 It's a game : “Kevin Bacon is the Center of the Universe”

 The Oracle of Bacon 



Six Degrees of Paul Erdős:





 Academic papers.



 Erdős Number 

 Erdős Number Project 

 So naturally we must have the Erdős-Bacon Number .

 One Story Lab alum has $EB\# < \infty$.

 Natalie Herschlag's (Portman's) $EB\# = 5 + 2 = 7$.

 The EBS# is also a thing: erdosbaconsabbath.com .

The PoCSverse
Small-world networks
9 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

Outline

Small-world networks

Experiments

Theory

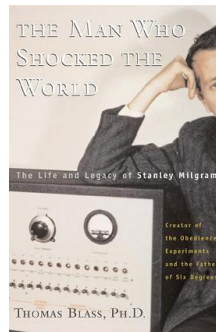
Generalized affiliation networks

Nutshell


References


The PoCSverse
Small-world networks
2 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References


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 ≈ 6.5 .

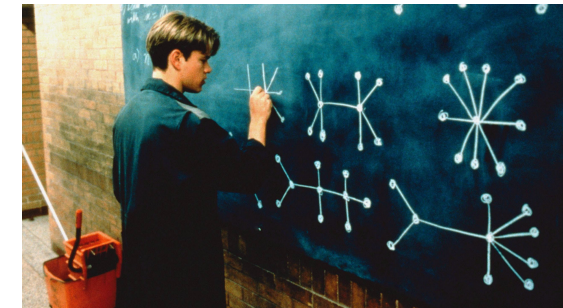
Popular terms:

 The Small World Phenomenon;



 “Six Degrees of Separation.”

The PoCSverse
Small-world networks
7 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

Good Will Hunting:




 Boardwork by Dan Kleitman , $EB\# = 1 + 2 = 3$.


 See Kleitman's sidebar in Mark Saul's Movie Review  (Notices of the AMS, Vol. 45, 1998.)

The PoCSverse
Small-world networks
10 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References


People thinking about people:


How are social networks structured?


 How do we define and measure connections?


 Methods/issues of self-report and remote sensing.

What about the dynamics of social networks?


 How do social networks/movements begin & evolve?


 How does collective problem solving work?

 How does information move through social networks?

 Which rules give the best ‘game of society?’

Sociotechnical phenomena and algorithms:

 What can people and computers do together? (google)

 Use Play + Crunch to solve problems. Which problems?

The PoCSverse
Small-world networks
5 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

From Frigyes Karinthy's “Chain-links”  in “Everything is Different”, 1929:

‘A fascinating game grew out of this discussion. One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth—anyone, anywhere at all. He bet us that, using no more than five individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances. For example, “Look, you know Mr. X.Y., please ask him to contact his friend Mr. Q.Z., whom he knows, and so forth.”’

The PoCSverse
Small-world networks
8 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

You may already be a winner in NSA's “three-degrees” surveillance sweepstakes!



NSA's probes could cover hundreds of millions of Americans. Thanks, Kevin Bacon.

by Sean Gallagher - July 18 2013, 4:00pm EDT

BIG DATA **IT'S**



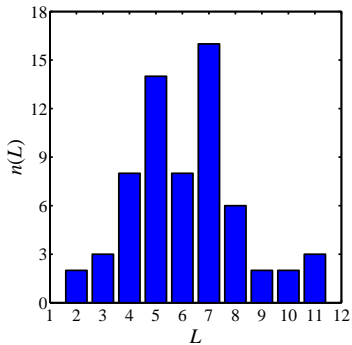
 Aurich Lawson

 Many people  are within three degrees from a random person ...

The PoCSverse
Small-world networks
11 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nutshell
References

The problem

Lengths of successful chains:



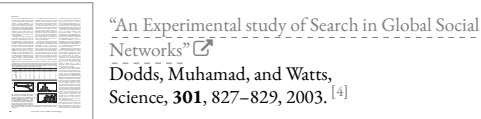
From Travers and Milgram (1969) in Sociometry: [9]
“An Experimental Study of the Small World Problem.”

The problem

- Two features characterize a social ‘Small World’:
- 1. Short paths exist, (= Geometric piece) and
 - 2. People are good at finding them. (= Algorithmic piece)


Social Search

Milgram’s small world experiment with email:



Social search—the Columbia experiment

- 60,000+ participants in 166 countries
- 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, and
 - a veterinarian in the Norwegian army.
- 24,000+ chains

We were lucky and contagious (more later):
“Using E-Mail to Count Connections” , Sarah Milstein, New York Times, Circuits Section (December, 2001)

Social search—the Columbia experiment

- Motivation/Incentives/Perception matter.
- If target *seems* reachable
⇒ participation more likely.
- Small changes in attrition rates
⇒ large changes in completion rates
- e.g., ↘ 15% in attrition rate
⇒ ↗ 800% in completion rate

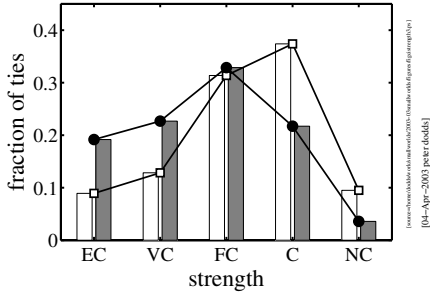
All targets:

Table S1										
Target	City	Country	Occupation	Gender	N	N (%)	# (rel)	<L>		
1	Novosibirsk	Russia	PhD student	F	8234	200 (24)	64 (76)	4.05		
2	New York	USA	Writer	F	6044	31 (0.51)	65 (73)	3.61		
3	Bandung	Indonesia	Unemployed	M	8151	0	66 (76)	n/a		
4	New York	USA	Journalist	F	5090	44 (0.77)	60 (72)	3.9		
5	Ithaca	USA	Professor	M	5855	168 (2.87)	54 (71)	3.84		
6	Melbourne	Australia	Travel Consultant	F	5597	20 (0.36)	60 (71)	5.2		
7	Bardufoss	Norway	Army veterinarian	M	4343	16 (0.37)	63 (76)	4.25		
8	Perth	Australia	Police Officer	M	4485	4 (0.09)	64 (75)	4.5		
9	Omaha	USA	Life Insurance Agent	F	4562	2 (0.04)	66 (79)	4.5		
10	Welwyn Garden City	UK	Retired	M	6593	1 (0.02)	68 (74)	4		
11	Paris	France	Librarian	F	4198	3 (0.07)	65 (75)	5		
12	Tallinn	Estonia	Archival Inspector	M	4530	8 (0.18)	63 (79)	4		
13	Munich	Germany	Journalist	M	4350	32 (0.74)	62 (74)	4.66		
14	Split	Croatia	Student	M	6629	0	63 (77)	n/a		
15	Gurgaon	India	Technology Consultant	M	4510	12 (0.27)	67 (78)	3.67		
16	Managua	Nicaragua	Computer analyst	M	6547	2 (0.03)	68 (78)	5.5		
17	Katikati	New Zealand	Potter	M	4091	12 (0.3)	62 (74)	4.33		
18	Edmonton	USA	Lutheran Pastor	M	4438	9 (0.21)	68 (76)	4.33		
Totals					98,847	384 (0.4)	63 (75)	4.05		

Social search—the Columbia experiment

Comparing successful to unsuccessful chains:

- Successful chains used relatively weaker ties:



Social search—the Columbia experiment

- Milgram’s participation rate was roughly 75%
- Email version: Approximately 37% participation rate.
- Probability of a chain of length 10 getting through:

$.37^{10} \approx 5 \times 10^{-5}$

- ⇒ 384 completed chains (1.6% of all chains).

Social search—the Columbia experiment

Successful chains disproportionately used:

- Weak ties, Granovetter [5]
- Professional ties (34% vs. 13%)
- Ties originating at work/college
- Target’s work (65% vs. 40%)

...and disproportionately avoided

- hubs (8% vs. 1%) (+ no evidence of funnels)
- family/friendship ties (60% vs. 83%)

Geography → Work

Social search—the Columbia experiment

Senders of successful messages showed little absolute dependency on

- age, gender
- country of residence
- income
- religion
- relationship to recipient

Range of completion rates for subpopulations: 30% to 40%

The PoCSverse
Small-world networks
21 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Social search—the Columbia experiment

Mildly bad for continuing chain: choosing recipients because “they have lots of friends” or because they will “likely continue the chain.”

- Why:
- Specificity important
 - Successful links used relevant information. (e.g. connecting to someone who shares same profession as target.)

The PoCSverse
Small-world networks
22 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Social search—the Columbia experiment

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

The PoCSverse
Small-world networks
23 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Usefulness:

- Harnessing social search:
- Can distributed social search be used for something big/good?
 - What about something evil? (Good idea to check.)
 - What about socio-inspired algorithms for information search? (More later.)
 - For real social search, we have an incentives problem.
 - Which kind of influence mechanisms/algorithms would help propagate search?
 - Fun, money, prestige, ...?
 - Must be ‘non-gameable.’

The PoCSverse
Small-world networks
24 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Finding red balloons:

- The winning team and strategy:
- MIT’s Media Lab won in less than 9 hours.
 - Pickard et al. “Time-Critical Social Mobilization,” Science Magazine, 2011.
 - People were virally recruited online to help out.
 - Idea: Want people to both (1) find the balloons, and (2) involve more people.
 - Recursive incentive structure with exponentially decaying payout:
 - \$2000 for correctly reporting the coordinates of a balloon.
 - \$1000 for recruiting a person who finds a balloon.
 - \$500 for recruiting a person who recruits the balloon finder, ...
 - (Not a Ponzi scheme.)
 - True victory: Colbert interviews Riley Crane

The PoCSverse
Small-world networks
27 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Red balloons:

- A Grand Challenge:
- 1969: The Internet is born (the ARPANET—four nodes!).
 - Originally funded by DARPA who created a grand Network Challenge for the 40th anniversary.
 - Saturday December 5, 2009: DARPA puts 10 red weather balloons up during the day.
 - Each 8 foot diameter balloon is anchored to the ground somewhere in the United States.
 - Challenge: Find the latitude and longitude of each balloon.
 - Prize: \$40,000.

*DARPA = Defense Advanced Research Projects Agency.

The PoCSverse
Small-world networks
25 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Finding balloons:

- Clever scheme:
- Max payout = \$4000 per balloon.
 - Individuals have clear incentives to both
 - involve/source more people (spread), and
 - find balloons (goal action).
 - Gameable?
 - Limit to how much money a set of bad actors can extract.

- Extra notes:
- MIT’s brand helped greatly.
 - MIT group first heard about the competition a few days before. Ouch.
 - A number of other teams did well.
 - Worthwhile looking at these competing strategies.

The PoCSverse
Small-world networks
28 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Where the balloons were:



The PoCSverse
Small-world networks
26 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Collective Detective:

- Finding an errant panda
- Nature News: “Crowdsourcing in manhunts can work: Despite mistakes over the Boston bombers, social media can help to find people quickly” by Philip Ball (April 26, 2013)
- Motherboard, Vice: One Degree of Separation in the Forever War by Brian Castner (November 11, 2015)

Once again, social media proved to be a powerful dragnet. Around 1:35 p.m., a Washingtonian posted a picture on Twitter of Rusty in a patch of weeds in the Adams Morgan district, not far from the 163-acre zoo, which was created in 1889 by an act of Congress. “Red panda in our neighborhood,” wrote Ashley Fonglitz, who identified herself as a singer, actress and traveler. “Please come save him!”

Another neighbor posted a photograph of two zoo workers, one in safari shorts standing on a rooftop, one holding a giant butterfly net. Soon the zoo announced: “Rusty the red panda has been recovered, crated & is headed safely back to the National Zoo!”

The PoCSverse
Small-world networks
29 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The social world appears to be small ...why?

Theory: how do we understand the small world property?

☞ Connected **random networks** have short average path lengths:

$$\langle d_{AB} \rangle \sim \log(N)$$

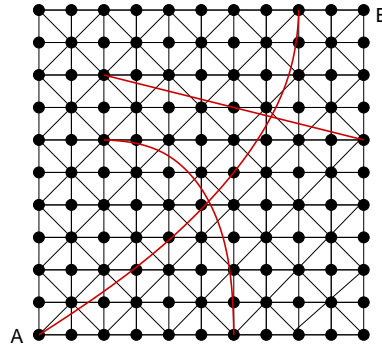
N = population size,

d_{AB} = distance between nodes A and B .

☞ But: social networks aren't random ...

The PoCSverse
Small-world networks
32 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

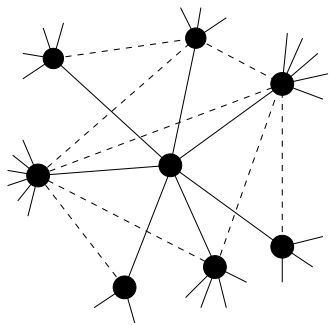
Randomness + regularity



Now have $d_{AB} = 3$

$\langle d \rangle$ decreases overall

Simple socialness in a network:



Need “**clustering**” (your friends are likely to know each other):

The PoCSverse
Small-world networks
33 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Small-world networks

Introduced by Watts and Strogatz (Nature, 1998)^[11]
“Collective dynamics of ‘small-world’ networks.”

Small-world networks were found everywhere:

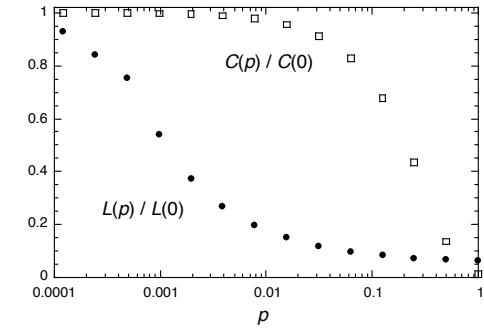
- ☞ neural network of *C. elegans*,
- ☞ semantic networks of languages,
- ☞ actor collaboration graph,
- ☞ food webs,
- ☞ social networks of comic book characters, ...

Very weak requirements:

- ☞ **local regularity** + **random short cuts**

The PoCSverse
Small-world networks
35 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The structural small-world property:

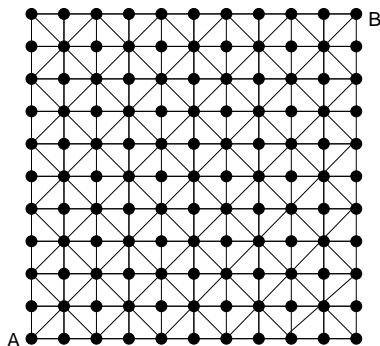


☞ $L(p)$ = average shortest path length as a function of p

☞ $C(p)$ = average clustering as a function of p

The PoCSverse
Small-world networks
38 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

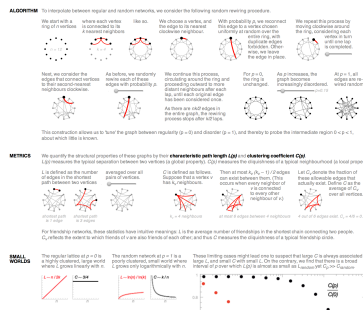
Non-randomness gives clustering:



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

The PoCSverse
Small-world networks
34 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Papers should be apps:



- ☞ Bret Victor's Scientific Communication As Sequential Art
- ☞ Interactive figures and tables = windows into large data sets (empirical or simulated).

The PoCSverse
Small-world networks
36 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Previous work—finding short paths

But are these short cuts findable?

Nope.^[6]

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

Need a more sophisticated model ...

The PoCSverse
Small-world networks
39 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Previous work—finding short paths

- ☞ What can a local search method reasonably use?
- ☞ How to find things without a map?
- ☞ Need some measure of distance between friends and the target.

Some possible knowledge:

- ☞ Target's identity
- ☞ Friends' popularity
- ☞ Friends' identities
- ☞ Where message has been

The PoCSverse
Small-world networks
40 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Previous work—finding short paths

Jon Kleinberg (Nature, 2000) ^[6]
“Navigation in a small world.”

Allowed to vary:

- 1. local search algorithm
and
- 2. network structure.

Previous work—finding short paths

Kleinberg’s Network:

- 1. Start with regular d-dimensional cubic lattice.
- 2. Add local links so nodes know all nodes within a distance q .
- 3. Add m short cuts per node.
- 4. Connect i to j with probability

$$p_{ij} \propto x_{ij}^{-\alpha}.$$

- ⚙️ $\alpha = 0$: random connections.
- ⚙️ α large: reinforce local connections.
- ⚙️ $\alpha = d$: connections grow logarithmically in space.

Previous work—finding short paths

Theoretical optimal search:

- ⚙️ “Greedy” algorithm.
- ⚙️ Number of connections grow logarithmically (slowly) in space: $\alpha = d$.
- ⚙️ Social golf.

Search time grows slowly with system size (like $\log^2 N$).

But: social networks aren’t lattices plus links.

Previous work—finding short paths

⚙️ If networks have hubs can also search well: Adamic et al. (2001) ^[1]

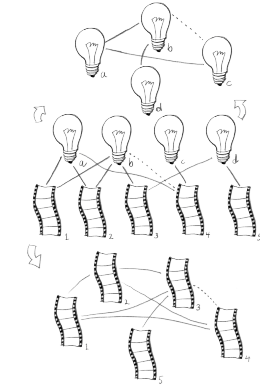
$$P(k_i) \propto k_i^{-\gamma}$$

where k = degree of node i (number of friends).

⚙️ Basic idea: get to hubs first (airline networks).

⚙️ But: hubs in social networks are limited.

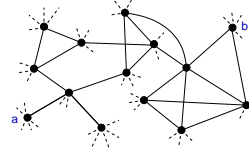
Bipartite affiliation structures:



- ⚙️ Many real-world networks have an underlying multi-partite structure.
 - ⚙️ Stories-tropes.
 - ⚙️ Boards and directors.
 - ⚙️ Films-actors-directors.
 - ⚙️ Classes-teachers-students.
 - ⚙️ Upstairs-downstairs.
- ⚙️ Unipartite networks may be induced or co-exist.

The problem

If there are no hubs and no underlying lattice, how can search be efficient?



Which friend of a is closest to the target b ?
What does ‘closest’ mean?
What is ‘social distance’?

Models

One approach: incorporate identity.

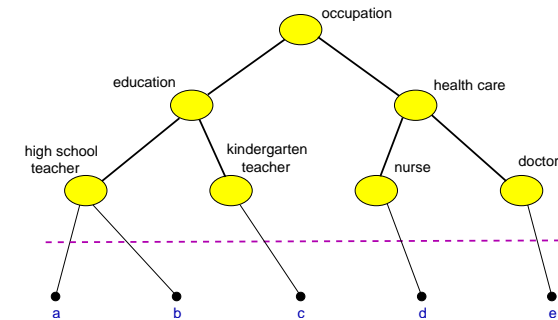
Identity is formed from attributes such as:

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

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

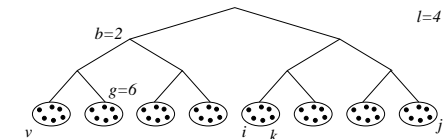
Attributes \Leftrightarrow Contexts \Leftrightarrow Interactions \Leftrightarrow Networks.

Social distance—Context distance



Models

Distance between two individuals x_{ij} is the height of lowest common ancestor.



$$x_{ij} = 3, x_{ik} = 1, x_{iv} = 4.$$

Models

- Individuals are more likely to know each other the closer they are within a hierarchy.

- Construct z connections for each node using

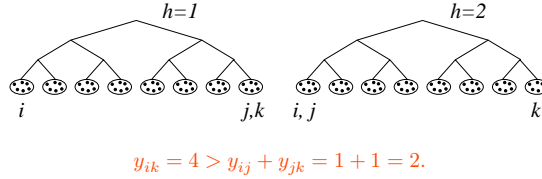
$$p_{ij} = \text{cexp}\{-\alpha x_{ij}\}.$$

- $\alpha = 0$: random connections.
- α large: local connections.

The PoCServe
Small-world networks
51 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The model

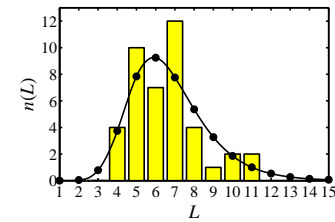
Triangle inequality doesn't hold:



The PoCServe
Small-world networks
54 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The model-results

Milgram's Nebraska-Boston data:



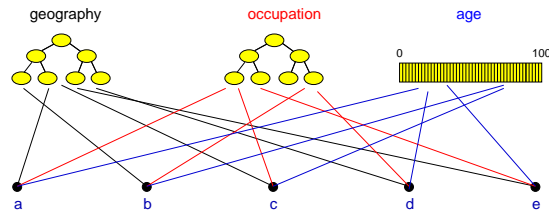
Model parameters:

- $N = 10^8$,
- $z = 300, g = 100$,
- $b = 10$,
- $\alpha = 1, H = 2$;
- $\langle L_{\text{model}} \rangle \approx 6.7$
- $L_{\text{data}} \approx 6.5$

The PoCServe
Small-world networks
57 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Models

Generalized affiliation networks



- Blau & Schwartz [2], Simmel [8], Breiger [3], Watts *et al.* [10]; see also Google+ Circles.

The PoCServe
Small-world networks
52 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The model

- Individuals know the identity vectors of
 - themselves,
 - their friends,
 - and
 - the target.
- Individuals can estimate the social distance between their friends and the target.
- Use a greedy algorithm + allow searches to fail randomly.

The PoCServe
Small-world networks
55 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

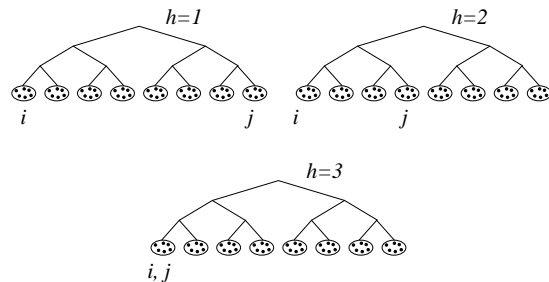
Social search—Data

Adamic and Adar (2003)

- For HP Labs, found probability of connection as function of organization distance well fit by exponential distribution.
- Probability of connection as function of real distance $\propto 1/r$.

The PoCServe
Small-world networks
58 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The model



$$\vec{v}_i = [1 \ 1 \ 1]^T, \vec{v}_j = [8 \ 4 \ 1]^T$$

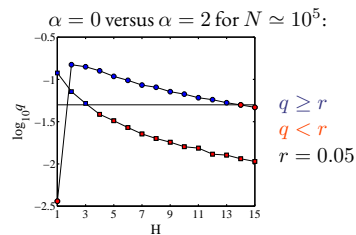
$$x_{ij}^1 = 4, x_{ij}^2 = 3, x_{ij}^3 = 1.$$

Social distance:

$$y_{ij} = \min_h x_{ij}^h.$$

The PoCServe
Small-world networks
53 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

The model-results—searchable networks



- q = probability an arbitrary message chain reaches a target.
- A few dimensions help.
- Searchability decreases as population increases.
- Precise form of hierarchy largely doesn't matter.

The PoCServe
Small-world networks
56 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Social Search—Real world uses

- Tags create identities for objects
- Website tagging: bitly.com
- (e.g., Wikipedia)
- Photo tagging: flickr.com
- Dynamic creation of metadata plus links between information objects.
- Folksonomy: collaborative creation of metadata

The PoCServe
Small-world networks
59 of 66
Small-world networks
Experiments
Theory
Generalized affiliation networks
Nashall
References

Recommender systems:

- Amazon uses people’s actions to build effective connections between books.
- Conflict between ‘expert judgments’ and tagging of the hoi polloi.

[1] L. Adamic, R. Lukose, A. Puniyani, and B. Huberman. Search in power-law networks. Phys. Rev. E, 64:046135, 2001. pdf ↗

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

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

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

[9] J. Travers and S. Milgram. An experimental study of the small world problem. Sociometry, 32:425–443, 1969. pdf ↗

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

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

Nutshell for Small-World Networks:

- Bare networks are typically unsearchable.
- Paths are findable if nodes understand how network is formed.
- Importance of identity (interaction contexts).
- Improved social network models.
- Construction of peer-to-peer networks.
- Construction of searchable information databases.

[5] M. Granovetter. The strength of weak ties. Am. J. Sociol., 78(6):1360–1380, 1973. pdf ↗

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

[7] G. Pickard, W. Pan, I. Rahwan, M. Cebrian, R. Crane, A. Madan, and A. Pentland. Time-critical social mobilization. Science, 334:509–512, 2011. pdf ↗

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