beyond social networks

Small world phenomenon:

high clustering $C_{\text{network}} >> C_{\text{random graph}}$

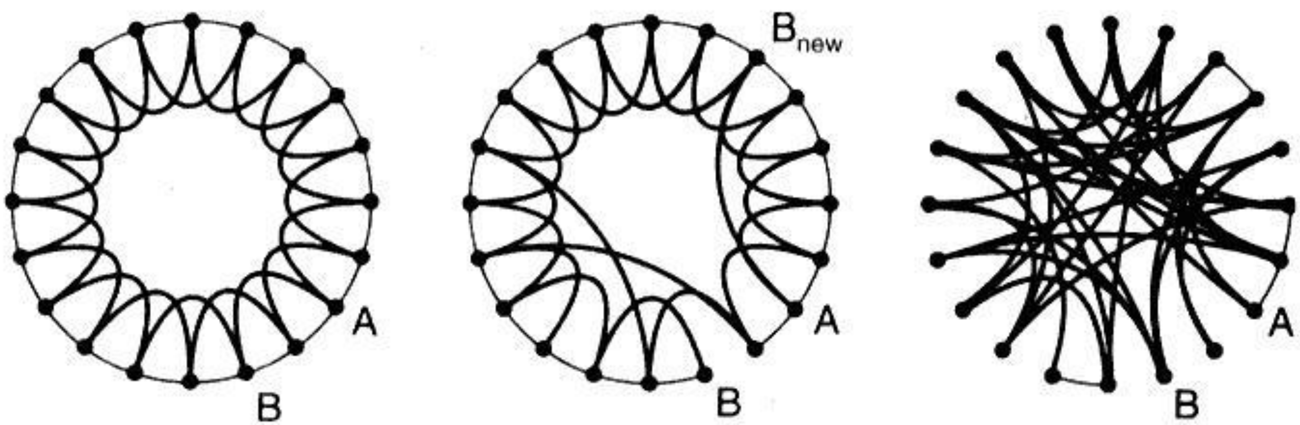
low average shortest path $l_{\text{network}} \approx \ln(N)$

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

Small world phenomenon: Watts/Strogatz model

Reconciling two observations:

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



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

Watts-Strogatz model:
Generating small world graphs

Select a fraction p of edges
Reposition on of their endpoints

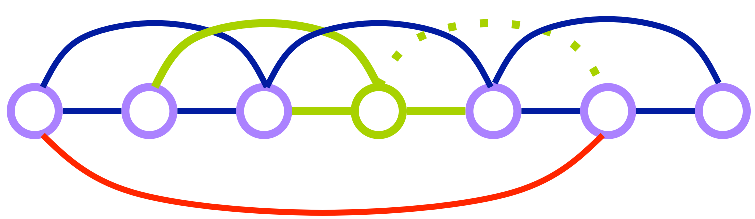
Add a fraction p of additional edges leaving underlying lattice intact

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

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

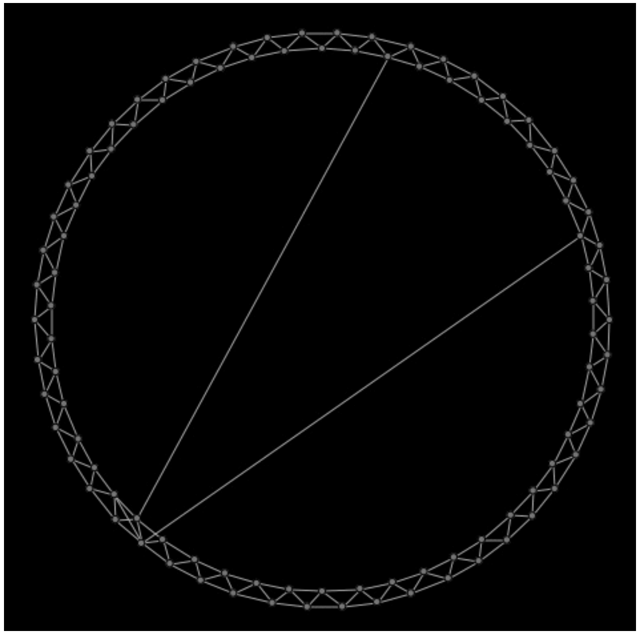
Watts-Strogatz model:
Generating small world graphs

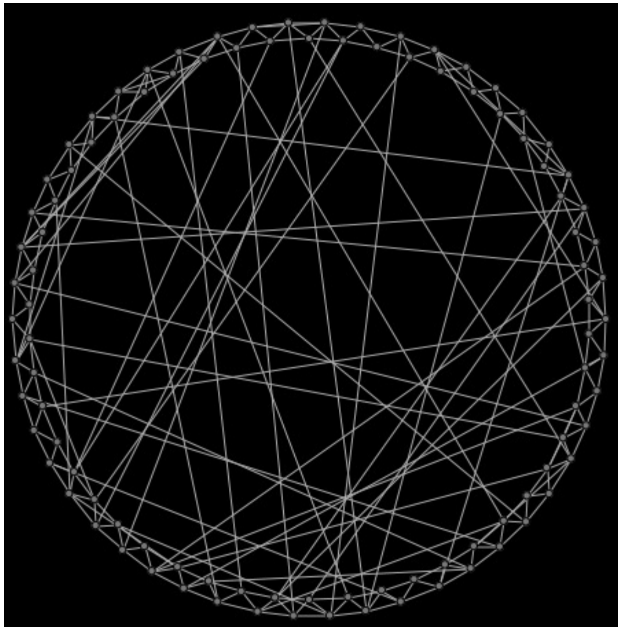
- Each node has K>=4 nearest neighbors (local)
- tunable: vary the probability *p* of rewiring any given edge
- small *p*: regular lattice
- large *p*: classical random graph



Quiz question:

- Which of the following is a result of a higher rewiring probability?

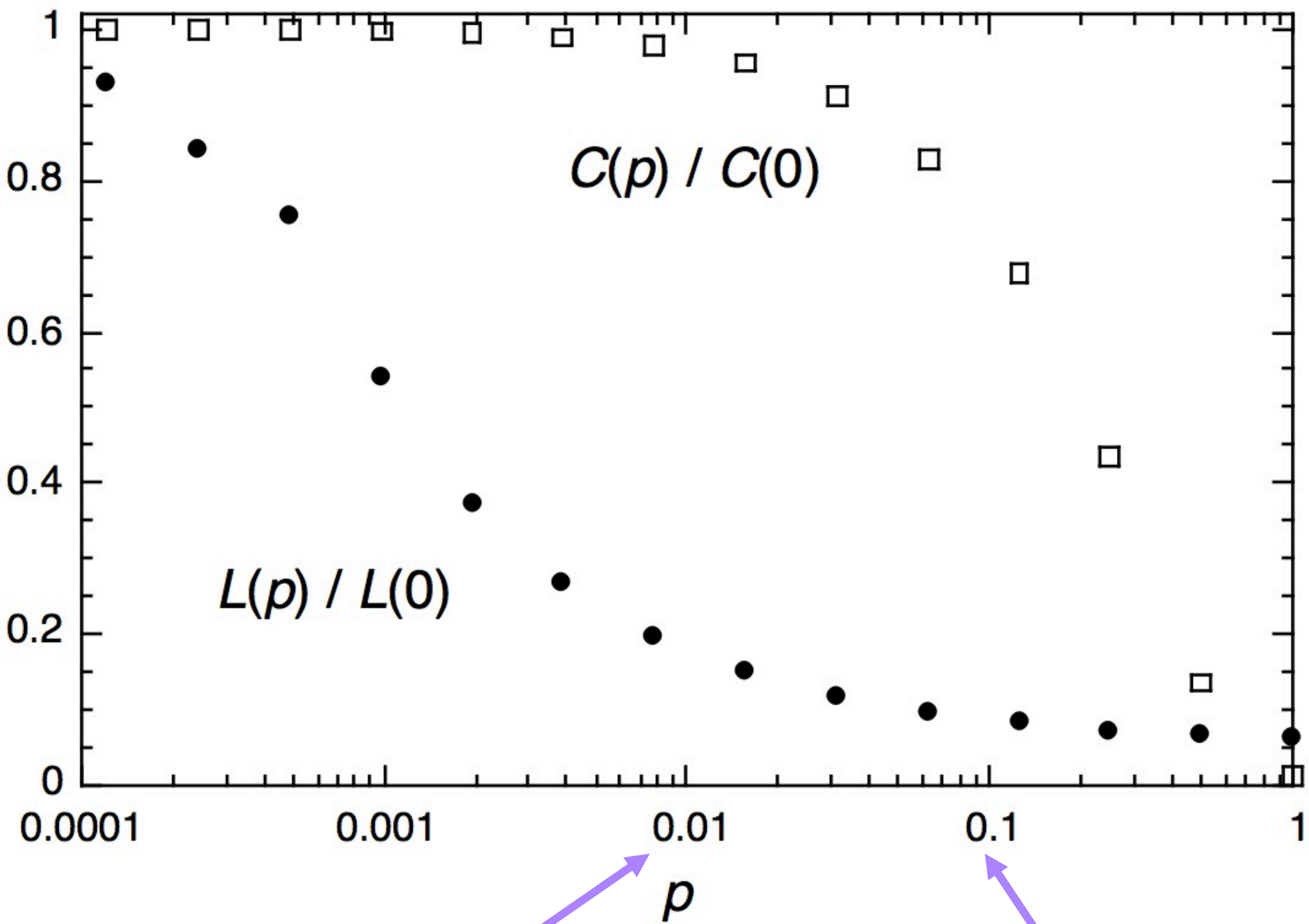




What happens in between?

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

Clust coeff. and ASP as rewiring increases



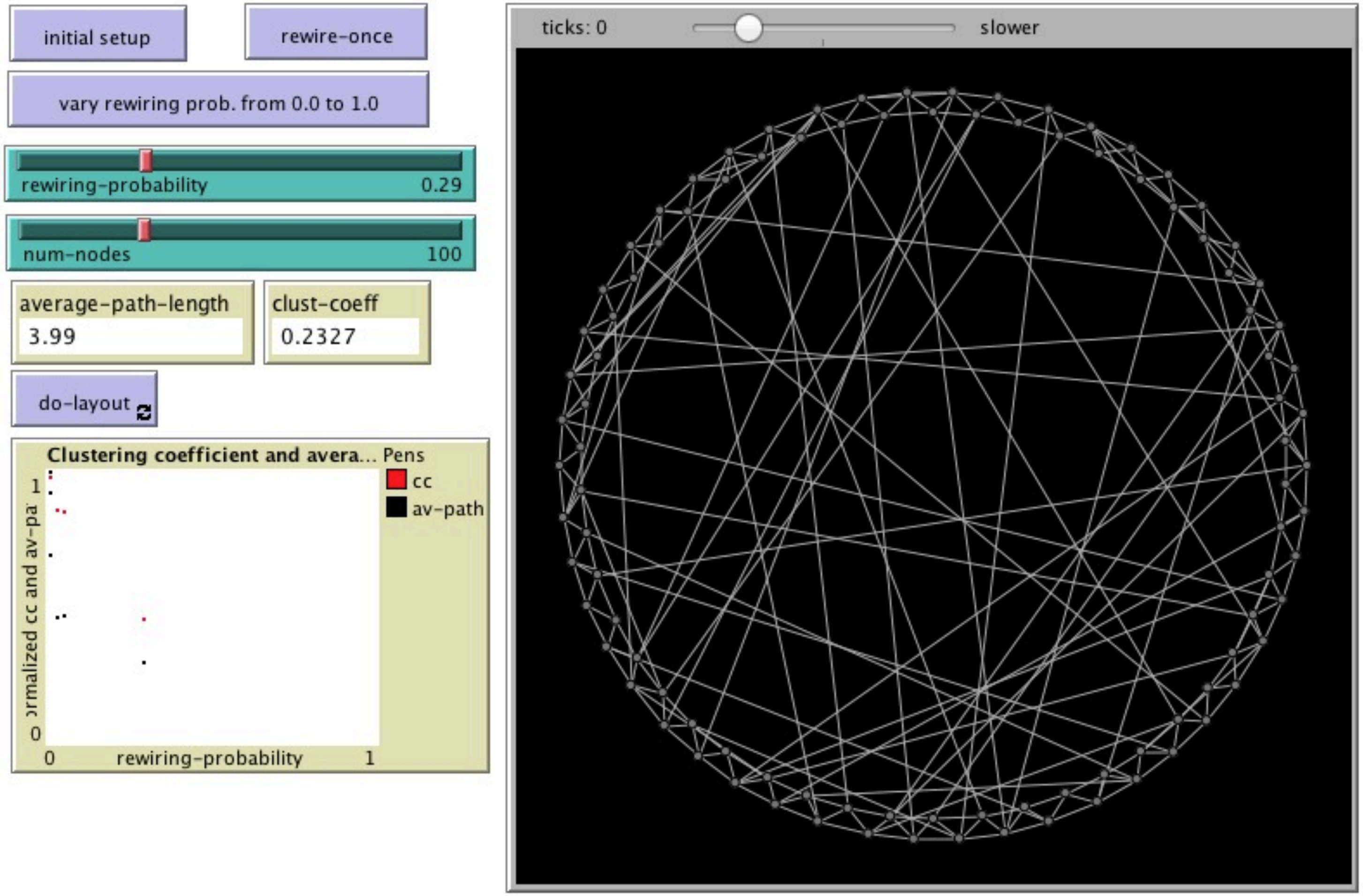
1% of links rewired

10% of links rewired

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

Trying this with NetLogo

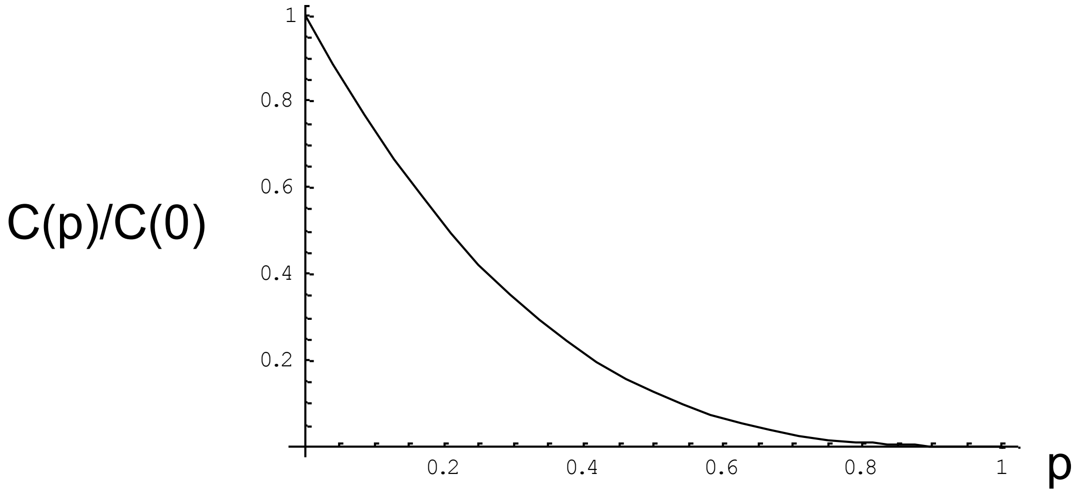
http://www.ladamic.com/netlearn/NetLogo4/SmallWorldWS.html



WS model clustering coefficient

- 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 very small, can ignore

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



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

Comparison with "random graph" used to determine whether real-world network is "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

Quiz Q

- Which of the following is a description matching a small-world network?

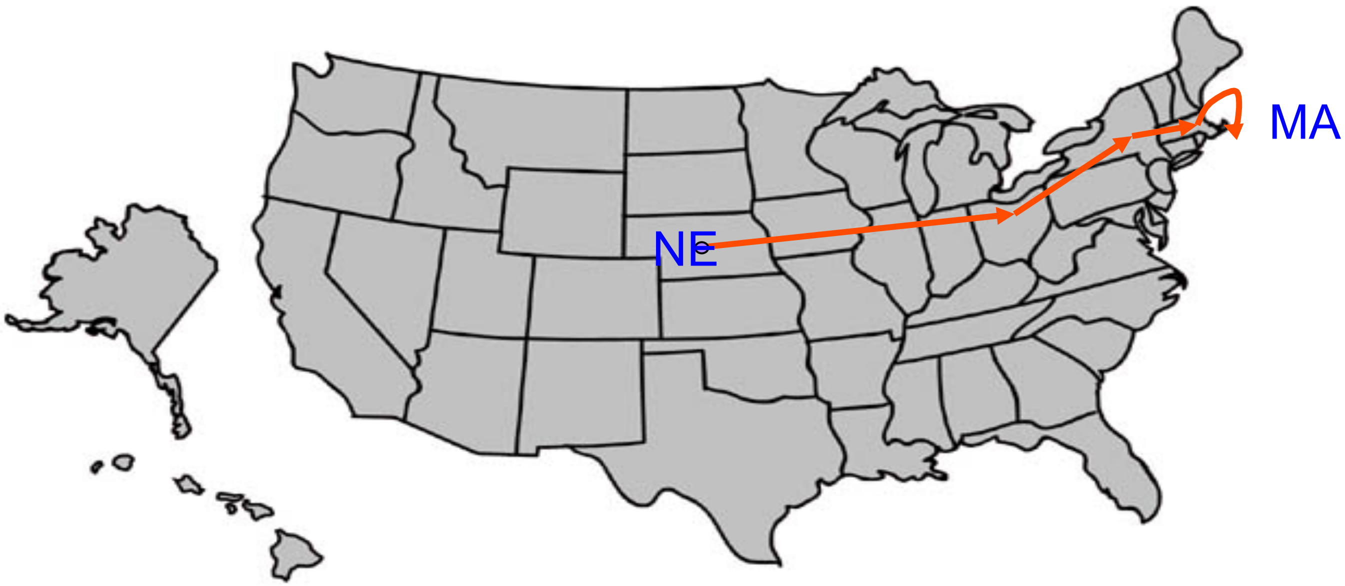
WS Model: What's missing?

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

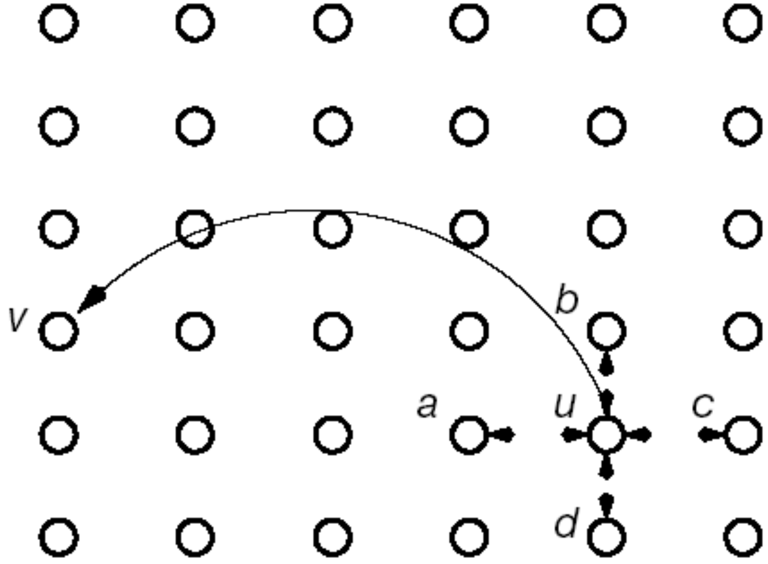
Ties and geography

"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



Kleinberg's geographical small world model



nodes are placed on a lattice and connect to nearest neighbors

exponent that will determine navigability

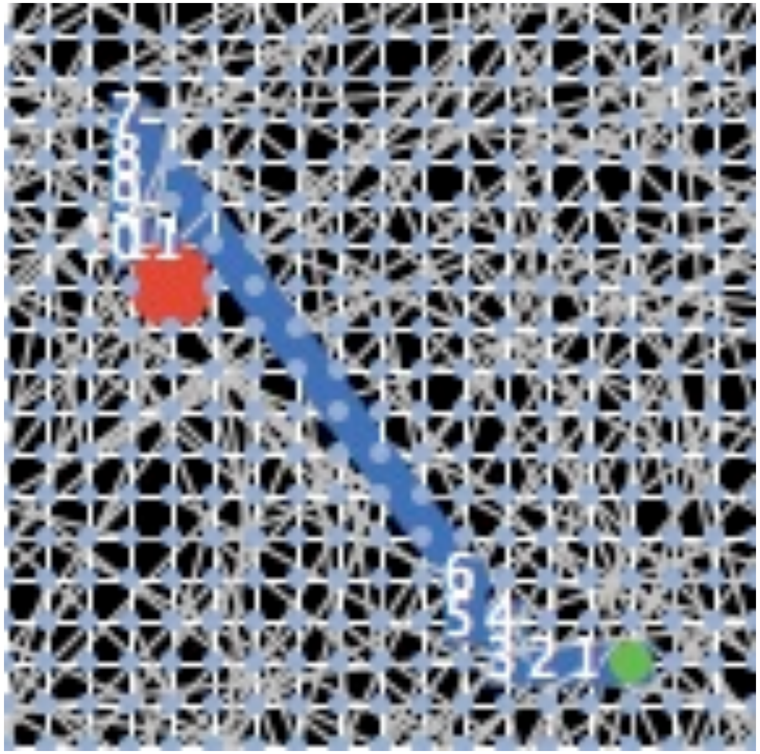
additional links placed with

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

Source: Kleinberg, 'The Small World Phenomenon, An Algorithmic Perspective' (Nature 2000).

NetLogo demo

- how does the probability of long-range links affect search?



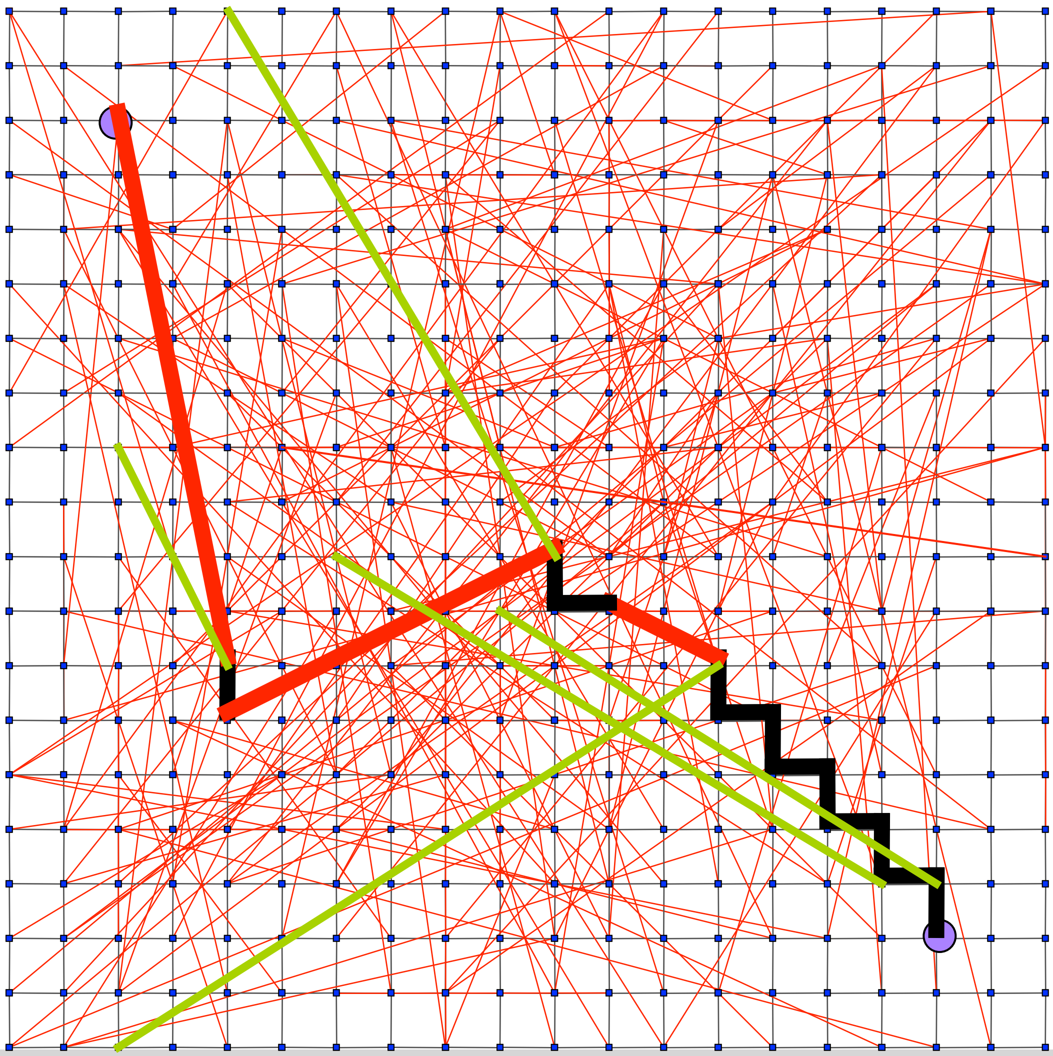
http://www.ladamic.com/netlearn/NetLogo4/SmallWorldSearch.html

geographical search when network lacks locality

When **r=0**, links are randomly distributed, ASP ~ **log(n)**, n size of grid

When **r=0**, any decentralized algorithm is at least $\mathbf{a_0 n^{2/3}}$

$p \sim p_0$

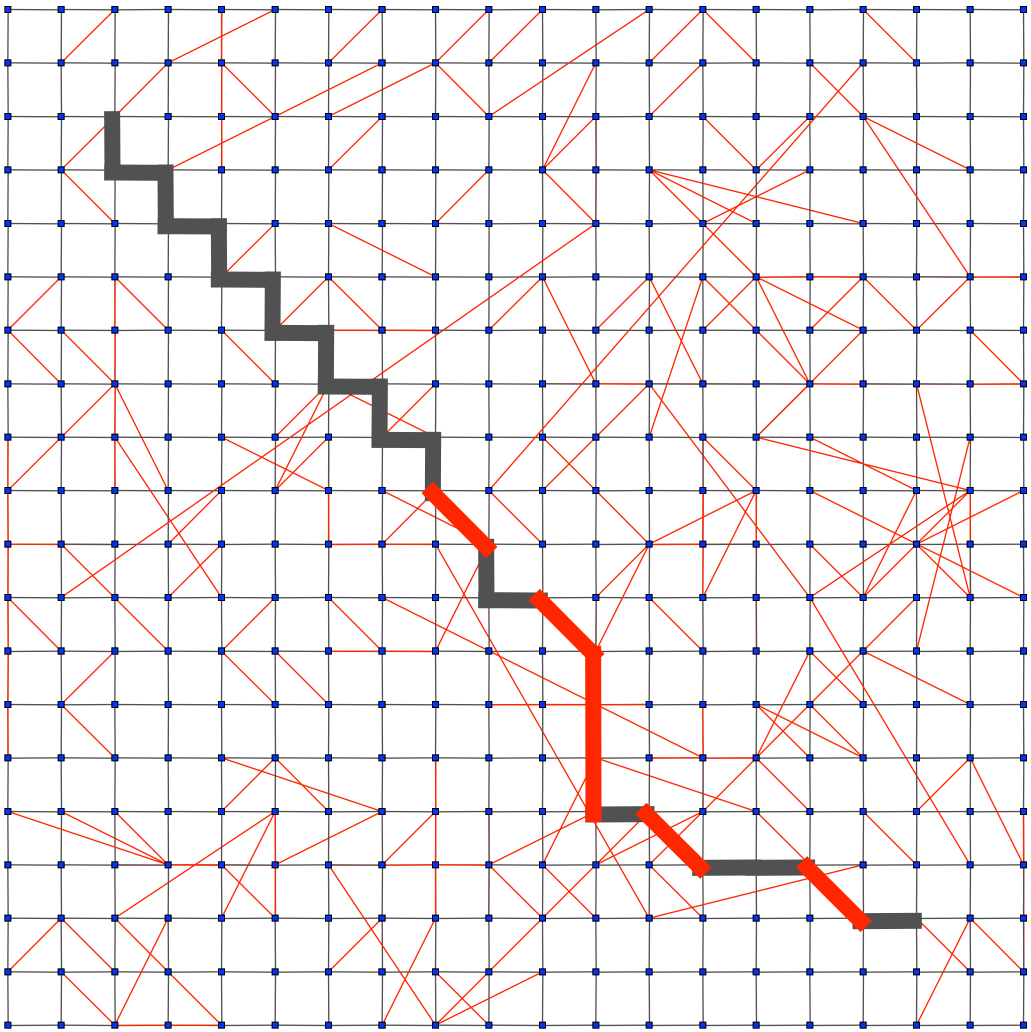


When **r<2**, expected time at least $\alpha_r n^{(2-r)/3}$

Overly localized links on a lattice

When **r>2** expected search time ~ $\mathbf{N^{(r-2)/(r-1)}}$

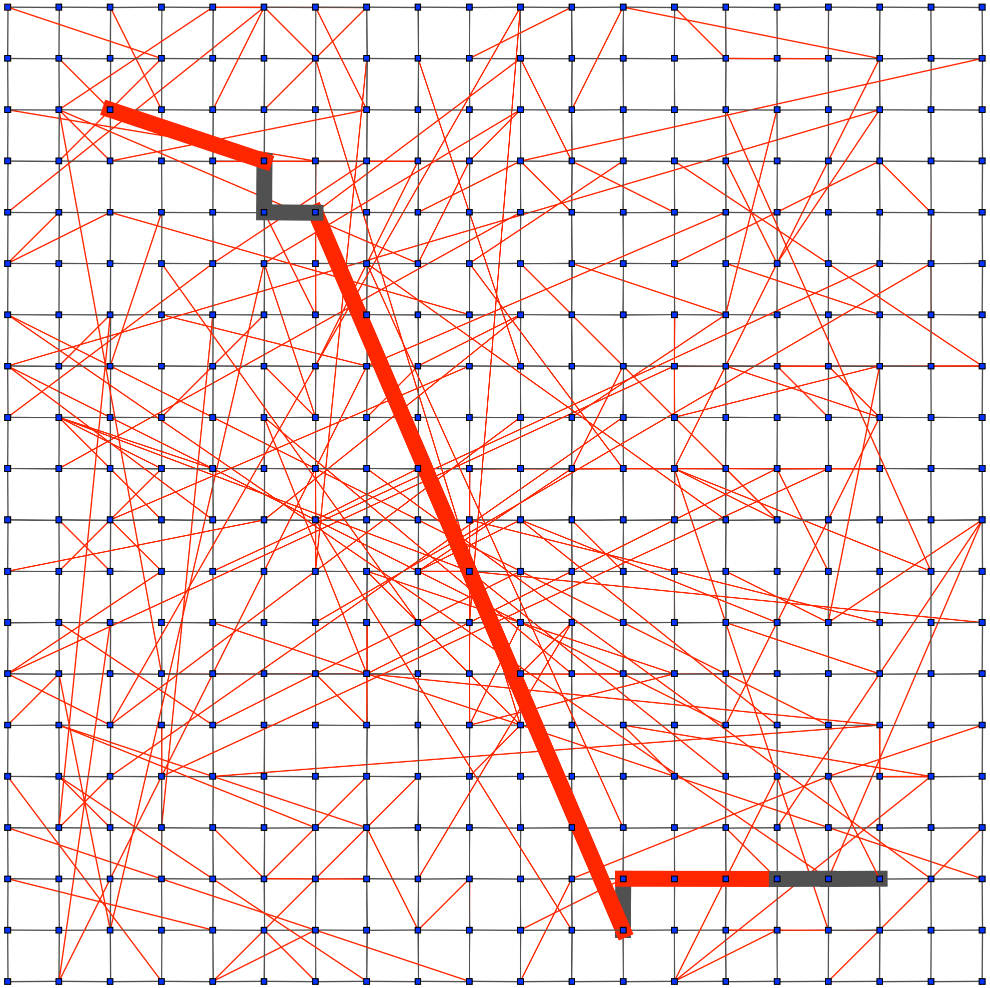
$$p \sim \frac{1}{d^4}$$



Just the right balance

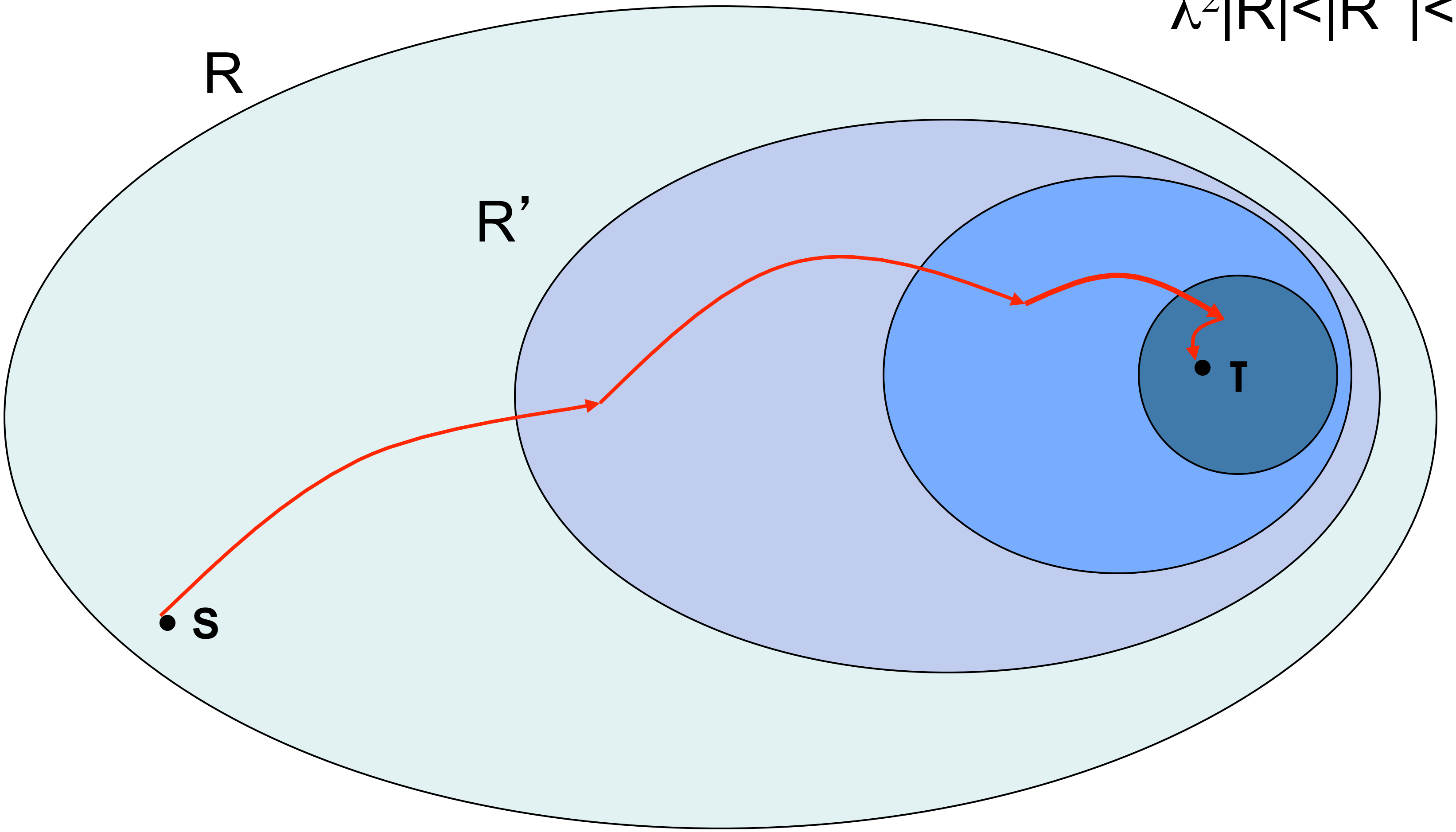
When **r=2,** expected time of a DA is at most **C (log N)^2**

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



Navigability

$\lambda^2|R|<|R'|<\lambda|R|$



$k = c \log^2 n$

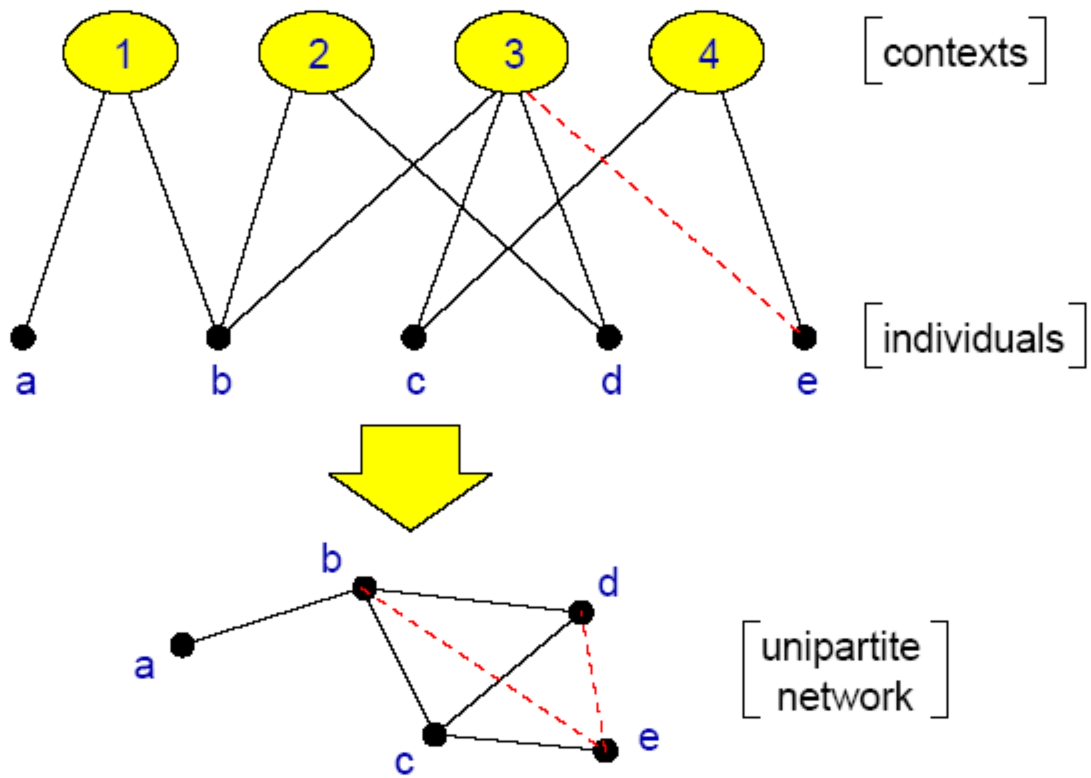
calculate probability that s fails to have a link in R'

Quiz Q:

- What is true about a network where the probability of a tie falls off as distance^{-2}

Origins of small worlds: group affiliations

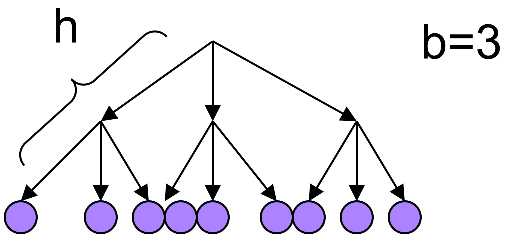
Social distance—Bipartite networks:



hierarchical small-world models: Kleinberg

Hierarchical network models:

Individuals classified into a hierarchy,
h_{ij} = height of the least common ancestor.

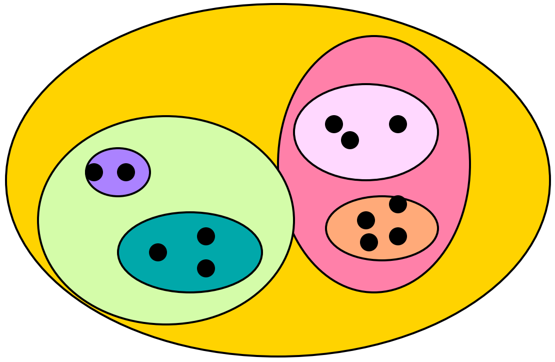


$$p_{ij} : b^{-\alpha h_{ij}}$$

e.g. state-county-city-neighborhood
industry-corporation-division-group

Group structure models:
Individuals belong to nested groups
q = size of smallest group that v,w belong to

$$f(q) \sim q^{-\alpha}$$

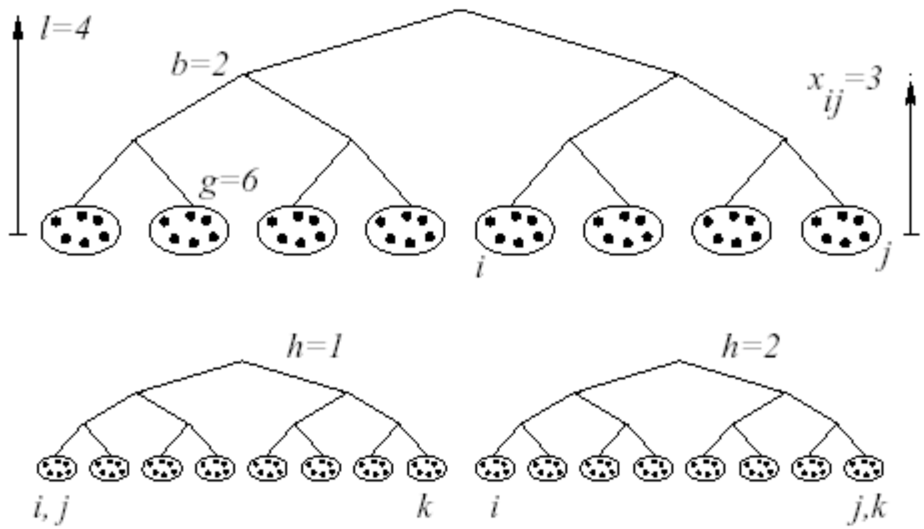


Source: Kleinberg, 'Small-World Phenomena and the Dynamics of Information' NIPS 14, 2001.

hierarchical small-world models: WDN

Watts, Dodds, Newman (Science, 2001)

individuals belong to hierarchically nested groups

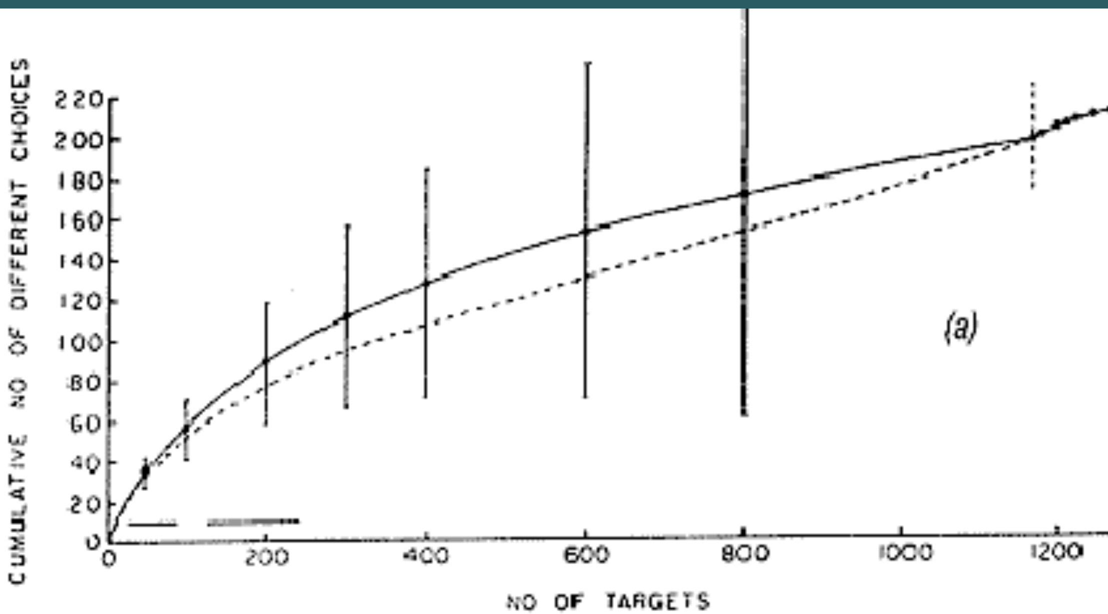


$p_{ij} \sim \exp(-\alpha x)$

multiple independent hierarchies h=1,2,..,H coexist corresponding to occupation, geography, hobbies, religion…

Source: Identity and Search in Social Networks: Duncan J. Watts, Peter Sheridan Dodds, and M. E. J. Newman; Science 17 May 2002 296: 1302-1305. < http://arxiv.org/abs/cond-mat/0205383v1 >

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
- based on (most often) occupation, geography
- only 7% because they "know a lot of people"
- Simple greedy algorithm: most similar acquaintance
- two-step strategy rare

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

Navigability and search strategy: Small world experiment @ Columbia

Successful chains disproportionately used

- weak ties (Granovetter)
- 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%)

Strategy: Geography -> Work