New Horizons in Search Theory, 3rd Workshop; Newport, RI

Social and Organizational Search

Collective Dynamics Group, Columbia University

D. J. Watts, C. F. Sabel, M. E. J. Newman, R. Muhamad, G. Kossinets, P. S. Dodds

> NSF
> Legg Mason
> Hewlett Foundation
> McDonnell Foundation
> Office of Naval Research

Outline:

I. Social search: The Small World Phenomenon

a. Theory

b. Experiment

II. Organizations: Information exchange, Searchability, and Robustness.

I. Social search:

The Small World Phenomenon

a. Theory

Social Search:

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

Social Search:

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

A. Yes (apparently):
The small world phenomenon
or
"Six Degrees of Separation."

The problem:

Stanley Milgram et al.
Late 1960's.

- Target person worked in Boston as a stockbroker.
- 296 senders from Boston and Omaha.
- 20% of senders reached target.
- average chain length $\simeq 6.5$.

The problem:

Two significant features characterize a small-world network:

1. Short paths exist.

The problem:

Two significant features characterize a small-world network:

1. Short paths exist.

2. People are good at finding them.

Previous work—short paths:

Connected random networks
have short average path lengths:

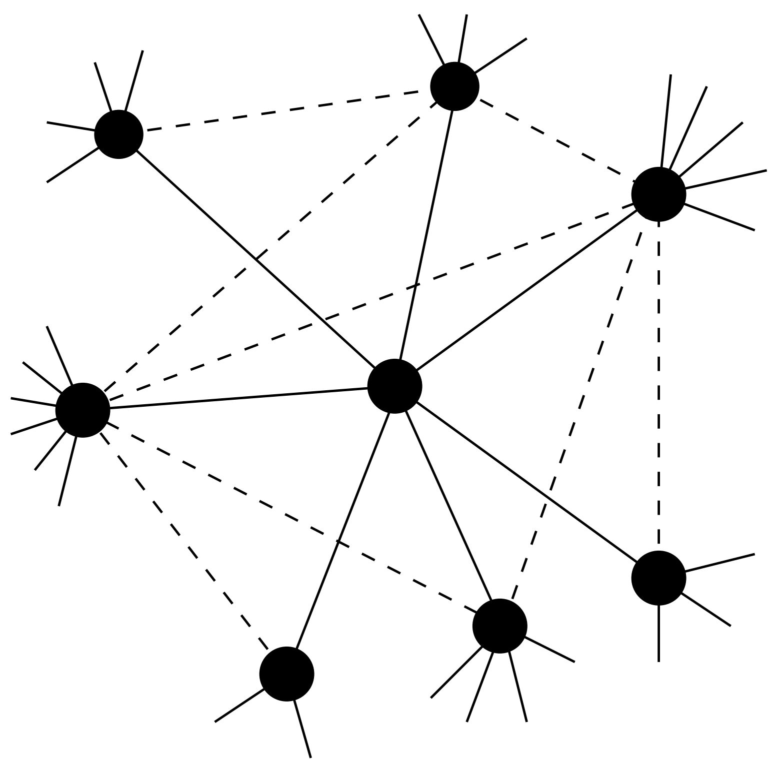
$\langle d_{AB} \rangle \sim \log(N)$

N = population size,
d_{AB} = distance between nodes A and B.

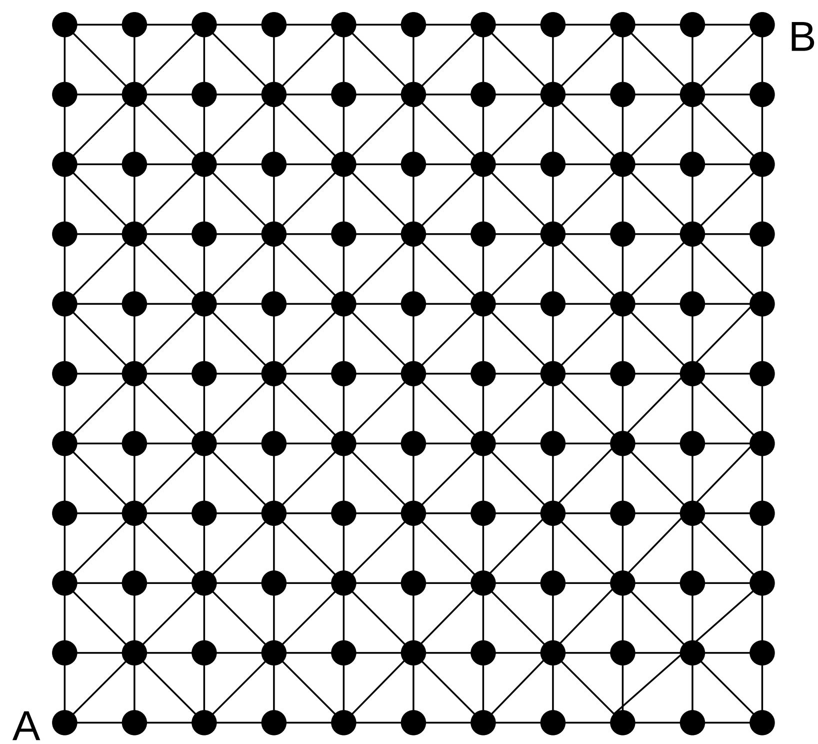
But: social networks aren't random.

Previous work—short paths:

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

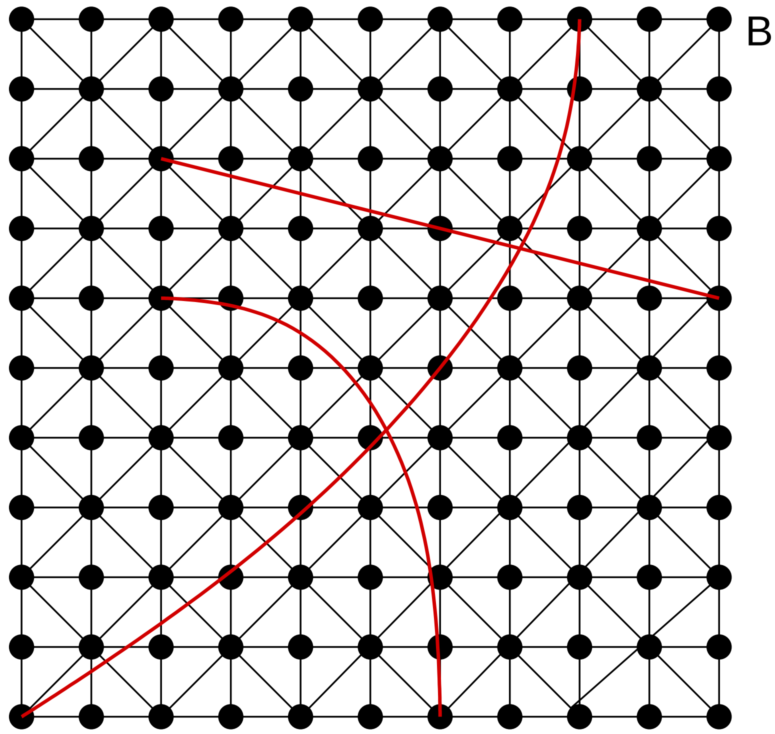


Non-randomness gives clustering:



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

Randomness + regularity:



Now have $d_{AB} = 3$

$\langle d \rangle$ decreases overall

Previous work—short paths:

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

Small-world networks found everywhere:

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

Previous work—short paths:

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

Small-world networks found everywhere:
- neural network of C. elegans,
- semantic networks of languages,
- actor collaboration graph,
- food webs.

Very weak requirements:
local regularity
+
random short cuts.

Previous work—finding short paths:

But are these short cuts findable?

Previous work—finding short paths:

But are these short cuts findable?

No.

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

Previous work—finding short paths:

Jon Kleinberg (Nature, 2000),
"Navigation in a small world."

Allowed to vary:

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

Previous work—finding short paths:

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 between nodes i and j with probability

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

Previous work—finding short paths:

Theoretical optimal search:

1. "Greedy" algorithm.

2. $\alpha = d$.

Search time grows like $\log^2(N)$.

For $\alpha \neq d$, polynomial factor N^{β} appears.

But: social networks aren't lattices plus links.

Previous work—finding short paths:

If networks have hubs can also search well (Adamic et al.)

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

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

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

But: hubs in social networks are limited.

The problem:

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

Which friend is closest to the target?

What does closest mean?

How to measure 'social distance' accurately?

The model:

One solution: incorporate identity.

Identity is formed from attributes such as:

1. Geographic location,

2. Type of employment,

3. Religious beliefs,

4. Recreational activities.

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

The model:

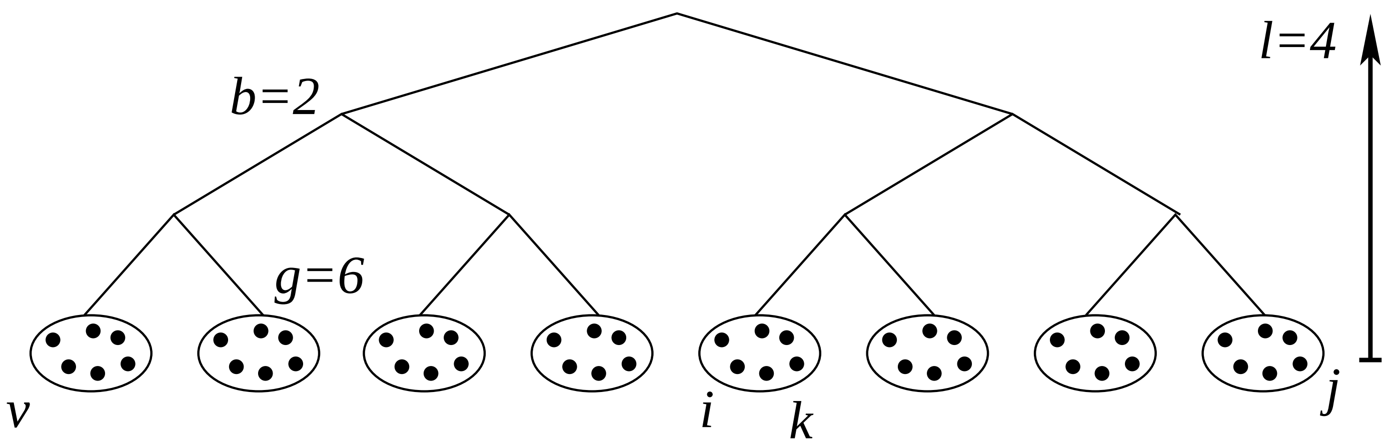
Six propositions about social networks:

P1: Individuals have identities and belong to various groups that reflect these identities.

P2: Individuals break down the world into a hierarchy of categories.

The model:

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



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

The model:

P3: Individuals are more likely to know each other the closer they are within a hierarchy.

Construct z connections for each node using

$$p_{ij} = c\,\mathrm{exp}\{-\alpha x_{ij}\}.$$

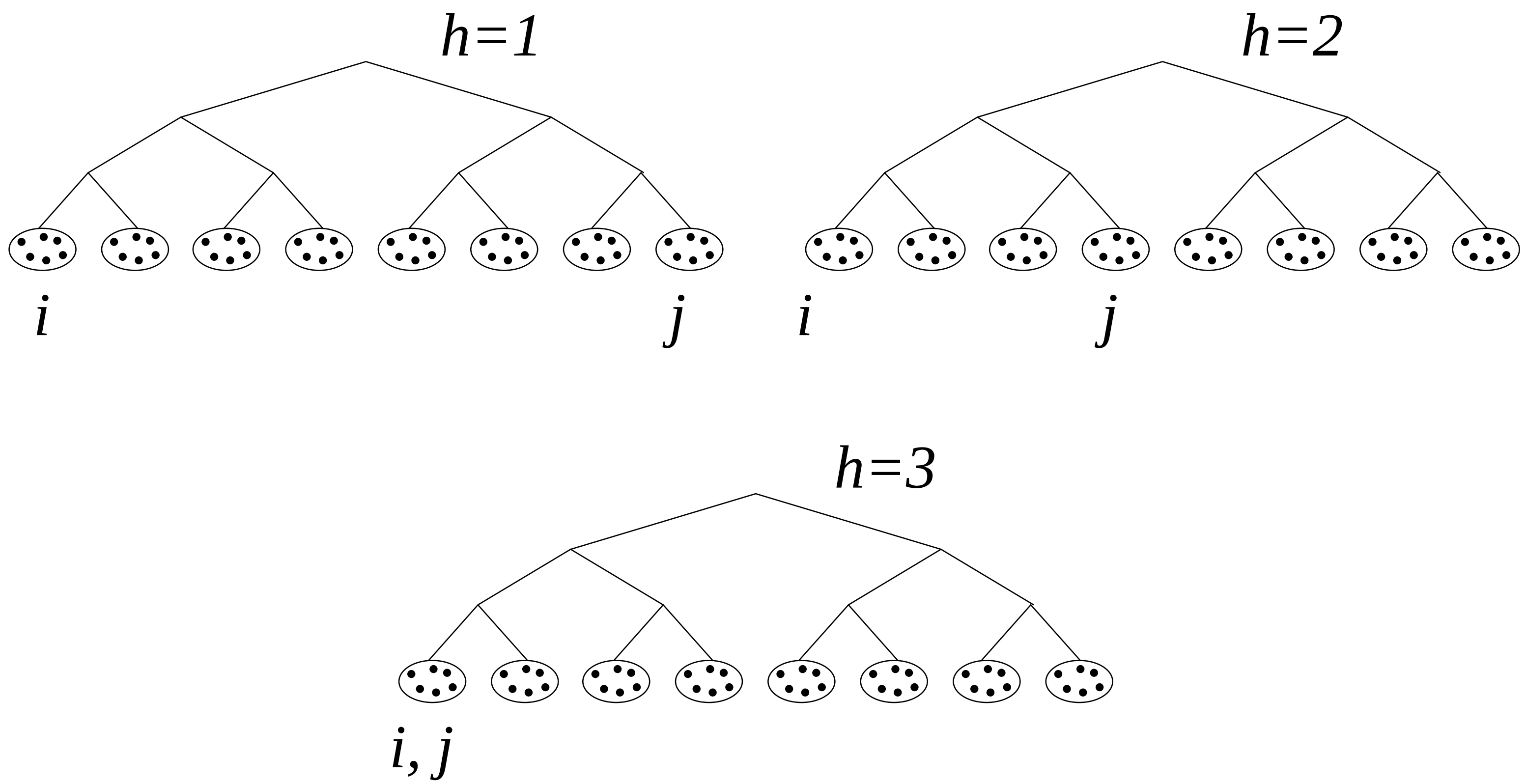
$\alpha = 0$: random connections.

α large: local connections.

The model:

P4: Each attribute of identity $\equiv$ hierarchy.

The model:



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

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

The model:

P5: "Social distance" is the minimum distance between two nodes in all hierarchies.

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

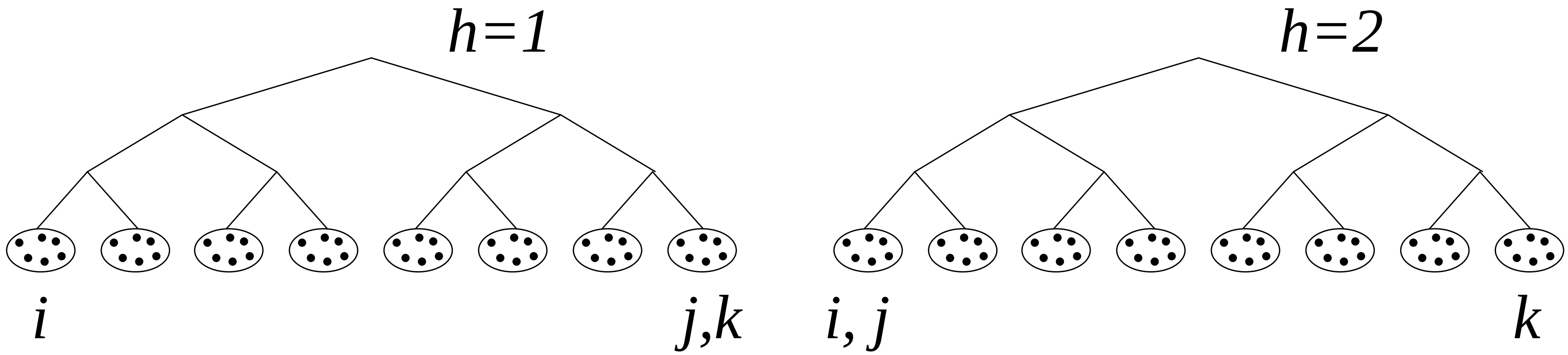
Previous slide:

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

$$\Rightarrow y_{ij} = 1.$$

The model:

Triangle inequality doesn't hold:



$$y_{ik} = 4 > y_{ij} + y_{jk} = 1 + 1 = 2.$$

The model:

P6: Individuals know the identity vectors of

1. themselves,

2. their friends,

and

3. the target.

Individuals can estimate the social distance between their friends and the target.

Use a greedy algorithm.

The model:

Define q as probability of an arbitrary message chain reaching a target.

Definition of a searchable network:

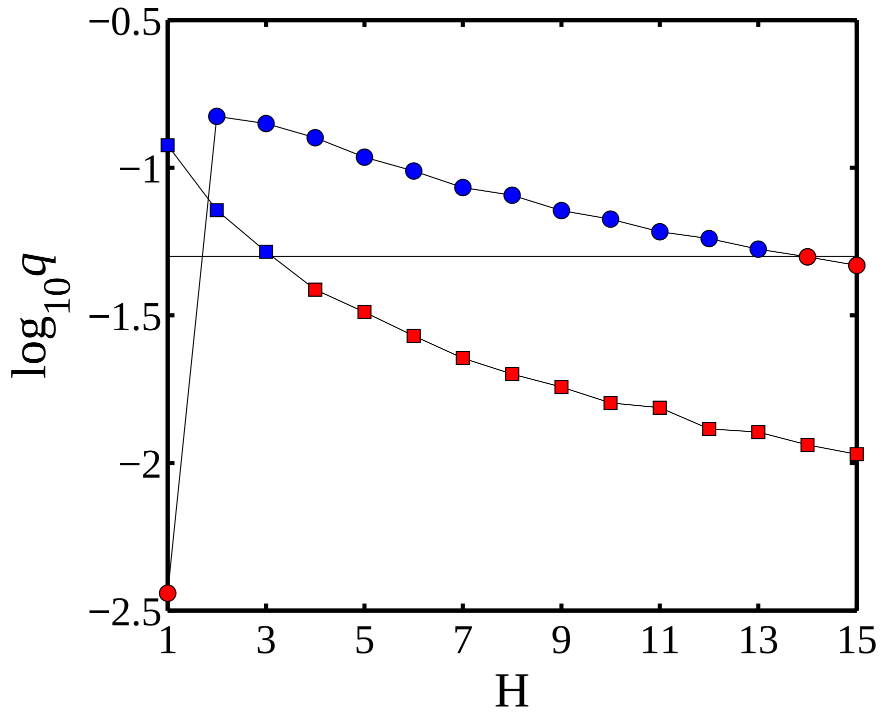
Any network for which

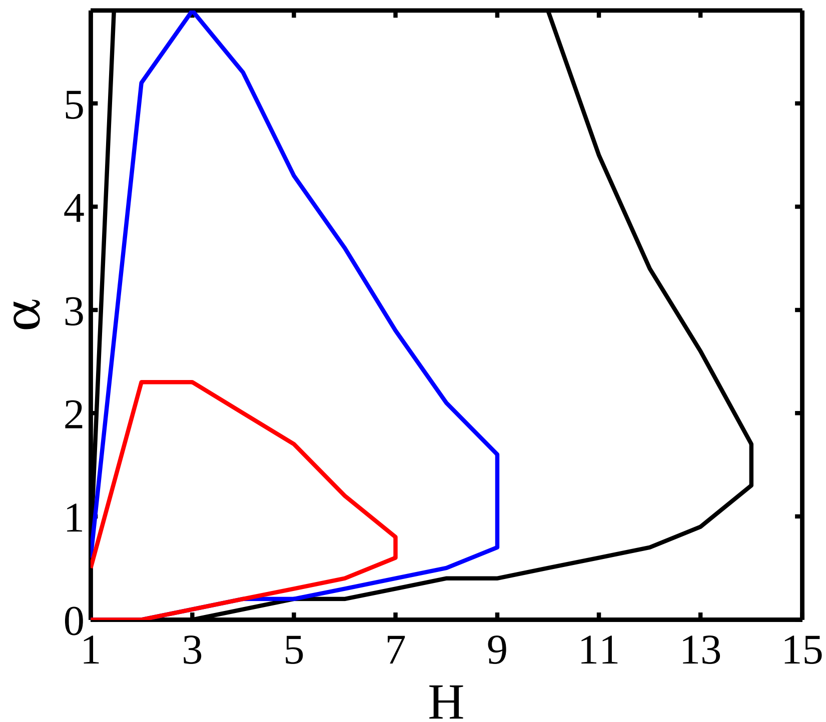
$$\boxed{q \geq r}$$

for a desired r.

The model-results:

$\alpha = 0$ versus $\alpha = 2$ for N=102400:





N=102400

N=204800

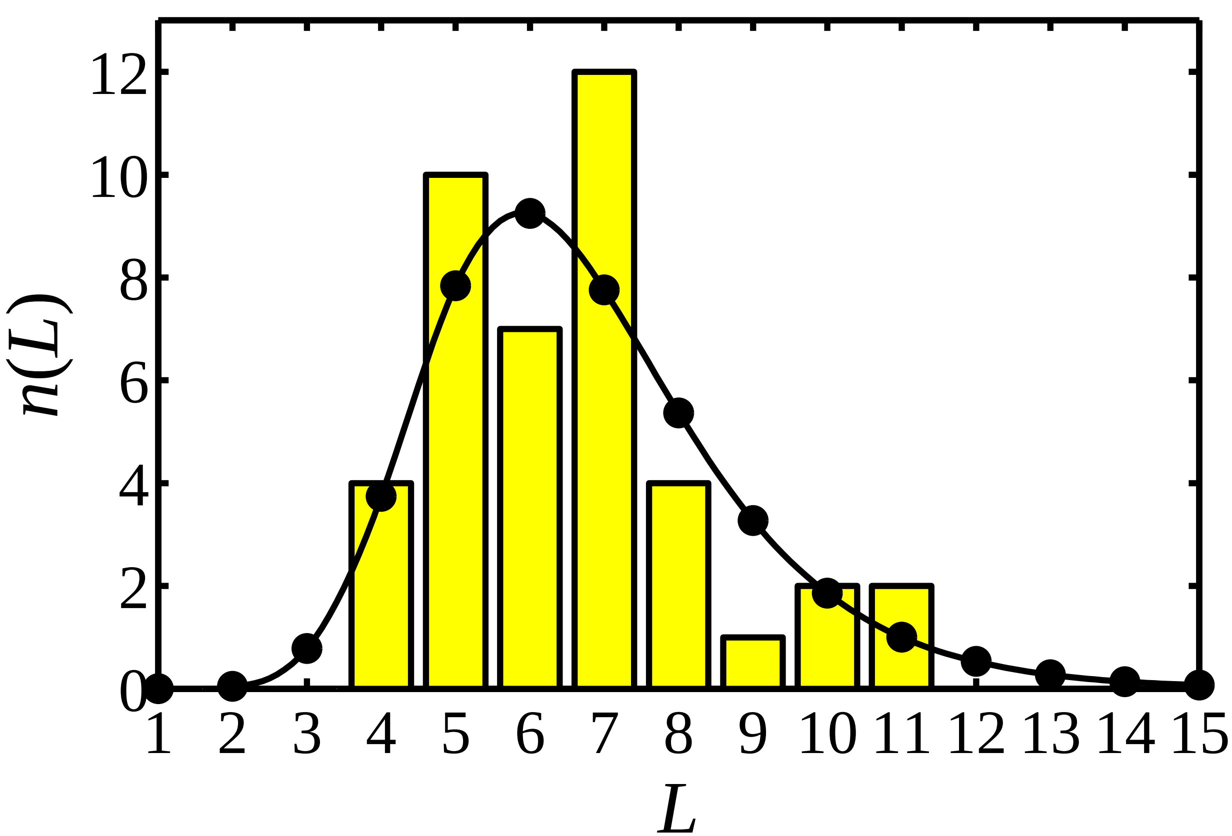
N=409600

p=0.25, r=0.05

b=2, g=100, z=99

The model-results:

Milgram's Nebraska-Boston data:



Conclusions:

- Bare networks are not enough.

- Paths are findable if nodes understand how network is formed.

- Importance of identity.

Applications:

- Improved social network models.

- Construction of peer-to-peer networks.

- Construction of searchable information databases.

I. Social search:

The Small World Phenomenon

b. Experiment

Social search—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

Social search—Experiment:

Approximately 37% participation rate.

Probability of a chain of length 10 getting through:

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

$\Rightarrow$ 384 completed chains (1.6% of all chains).

Social search—Experiment:

Motivation/Incentives/Perception matter.

If target *seems* reachable
$\Rightarrow$ participation more likely.

Small changes in attrition rates
$\Rightarrow$ large changes in completion rates

e.g., $\searrow$ 15% in attrition rate
$\Rightarrow$ $\nearrow$ 800% in completion rate

Social search—Experiment:

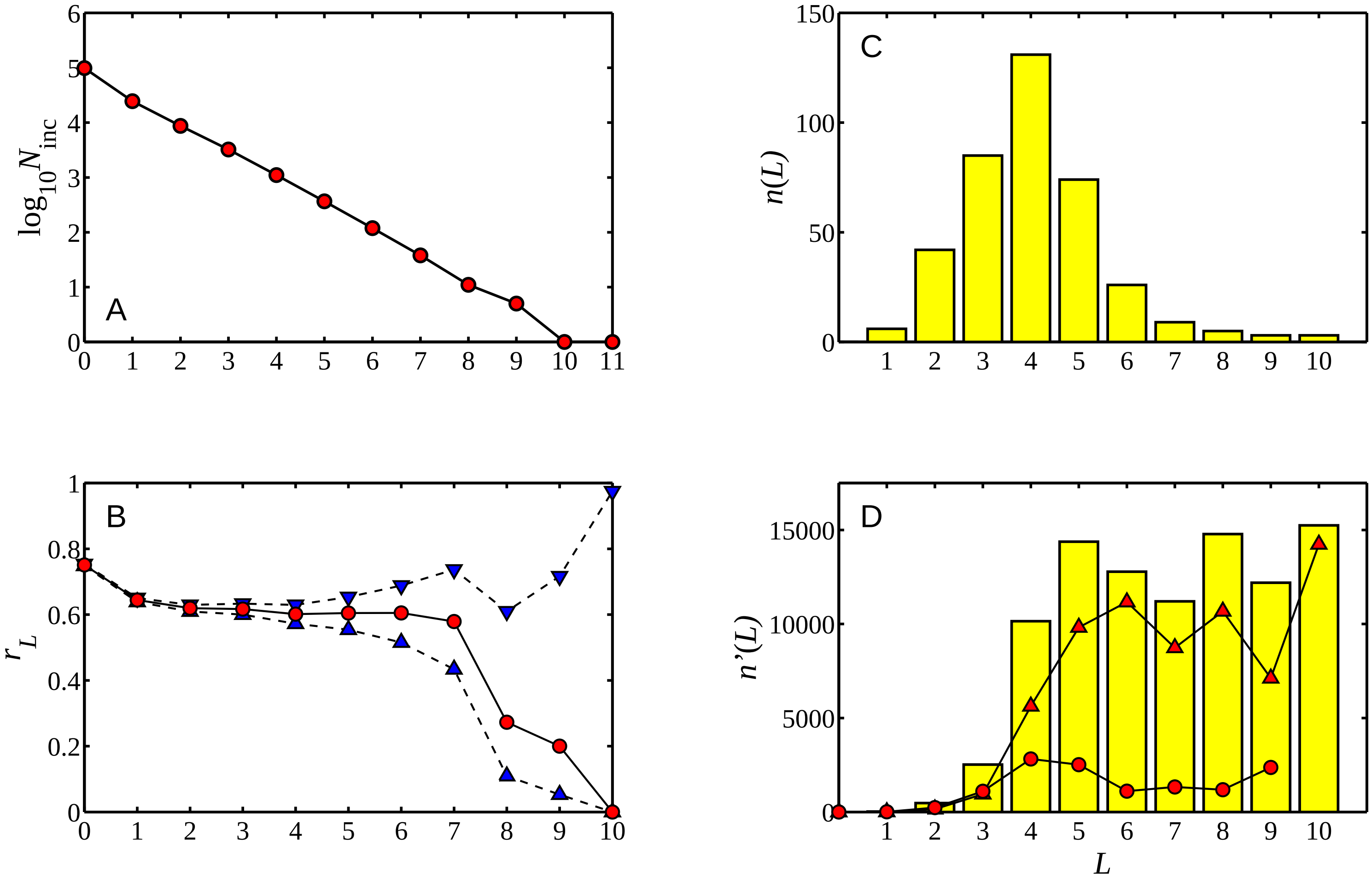
Successful chains disproportionately used

- weak ties
- professional relationships
- target's work

. . . and disproportionately avoided

- hubs
- friends
- target's location

Social search—Experiment:



Social search—Experiment:

$\langle L \rangle = 4.05$ for all completed chains

Social search—Experiment:

$\langle L \rangle = 4.05$ for all completed chains

$L_* =$ Estimated 'true' median chain length

Intra-country chains: $L_* = 5$

Inter-country chains: $L_* = 7$

All chains: $L_* = 7$

Social search—Experiment:

$\langle L \rangle = 4.05$ for all completed chains

$L_* =$ Estimated 'true' median chain length

Intra-country chains: $L_* = 5$

Inter-country chains: $L_* = 7$

All chains: $L_* = 7$

Milgram: $L_* \simeq 8$–9

Social search—Experiment:

Other experiments:

1. Small World Experiment II (now running)

2. The People Finder project

3. Expert search

II. Organizations:

Information exchange, Searchability, and Robustness.

Organizations—February 1997:

Aisin, maker of brake valve parts for Toyota, burns to ground.

Organizations—February 1997:

Aisin, maker of brake valve parts for Toyota, burns to ground.

- 4 hours supply ("just in time").

Organizations—February 1997:

Aisin, maker of brake valve parts for Toyota, burns to ground.

- 4 hours supply ("just in time").
- 14,000 cars per day $\rightarrow$ 0 cars per day.

Organizations—February 1997:

Aisin, maker of brake valve parts for Toyota, burns to ground.

- 4 hours supply ("just in time").
- 14,000 cars per day $\rightarrow$ 0 cars per day.
- 6 months before new machines would be ready.

Organizations—February 1997:

Aisin, maker of brake valve parts for Toyota, burns to ground.

- 4 hours supply ("just in time").

- 14,000 cars per day $\rightarrow$ 0 cars per day.

- 6 months before new machines would be ready.

Recovered in 5 days.

Organizations—February 1997:

- 36 suppliers, 150 subcontractors
- 50 supply lines
- sewing machine maker produced 40 valves a day

Recovery depended on horizontal links:

- robustness
- searchability

Nishiguchi and Beaudet (1997)

Real organizations:

Extremes:

Hierarchy

- maximum efficiency
- suited to static environment
- brittle

Real organizations:

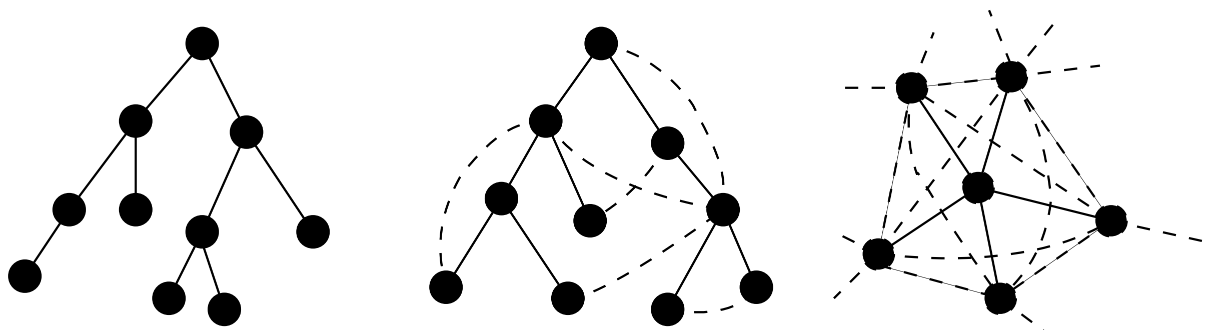
Extremes:

Hierarchy
- maximum efficiency
- suited to static environment
- brittle

'Market'
- resilient
- suited to rapidly changing environment
- requires low cost interactions

Real organizations:

Organizations are in the middle...



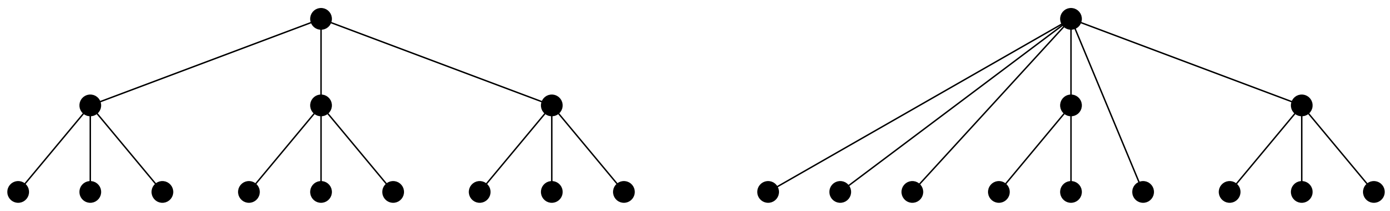
“Heterarchies” (D. Stark, 1999)

Organizations as efficient hierarchies:

Economics: Organizations ≡ Hierarchies.

e.g., Radner (1993), Van Zandt (1998)

Hierarchies performing associative operations:



Desirable organizational qualities:

1. Ultra-robustness:

I. Congestion robustness
(Resilience to failure due to information exchange)

Desirable organizational qualities:

1. Ultra-robustness:

I. Congestion robustness
(Resilience to failure due to information exchange)

$+$

II. Connectivity robustness
(Recoverability in the event of failure)

Desirable organizational qualities:

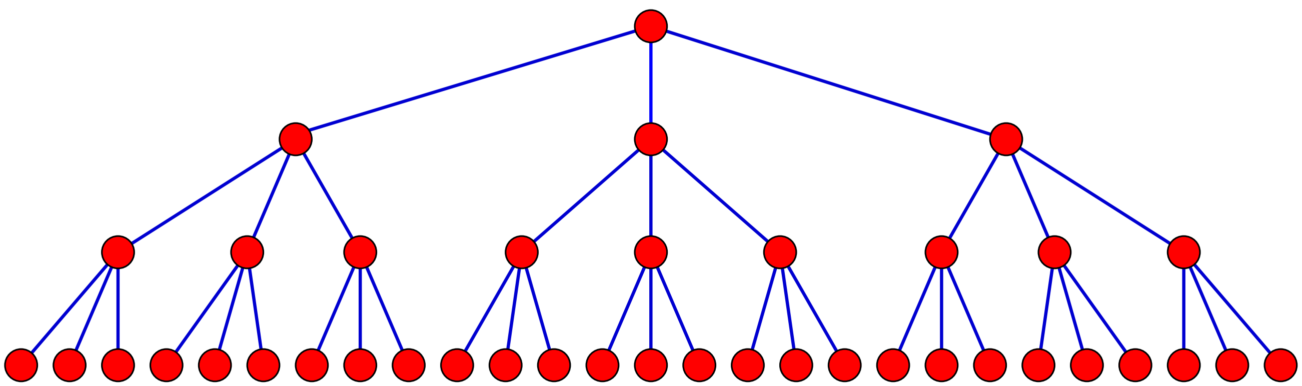
2. Low cost (requiring few links)

3. Scalability

4. Ease of construction—existence is plausible

5. Searchability—creativity, problem-solving

Model—underlying hierarchy:



$$b = 3, \quad L = 4, \quad N = 40$$

Model:

Formal organizational structure:

- Underlying hierarchy

  branching ratio b
  depth L
  $N = (b^L - 1)/(b - 1)$ nodes
  $N - 1$ links

Model:

Formal organizational structure:

- Underlying hierarchy

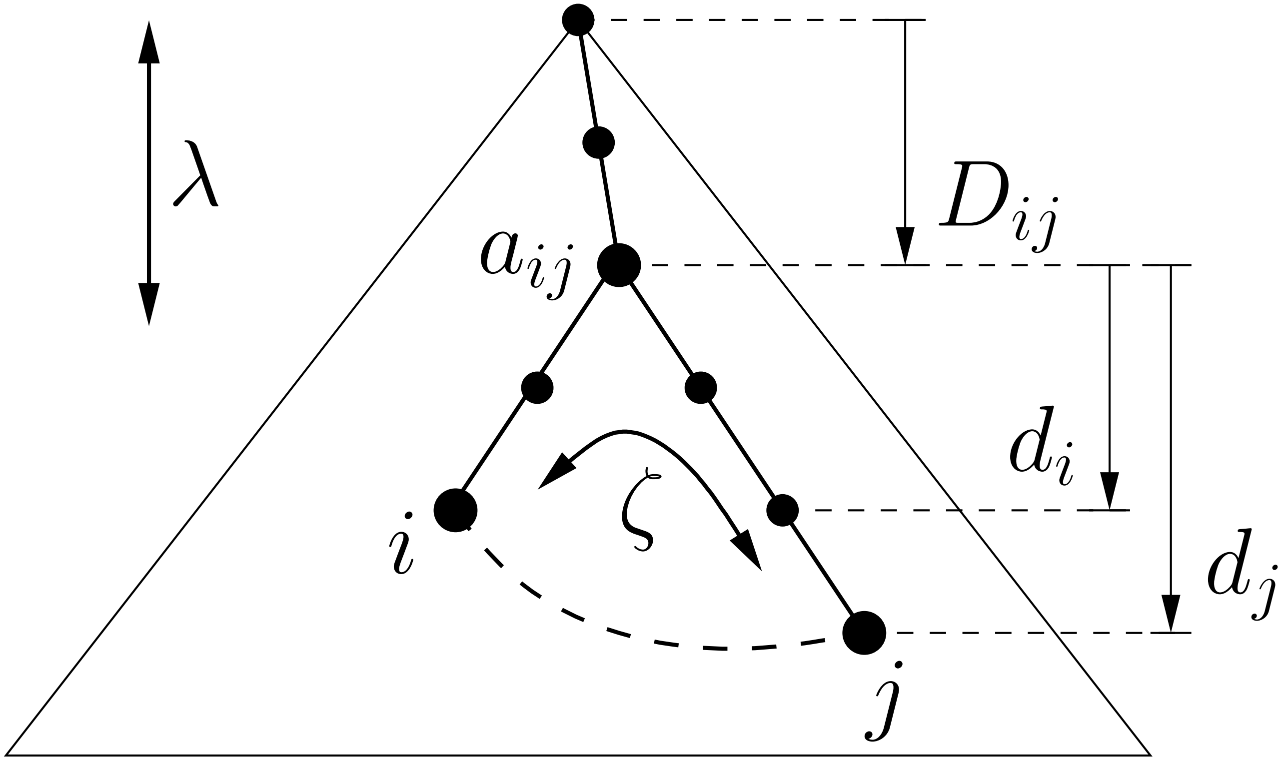
  branching ratio b
  depth L
  $N = (b^L - 1)/(b - 1)$ nodes
  $N - 1$ links

Additional informal ties:

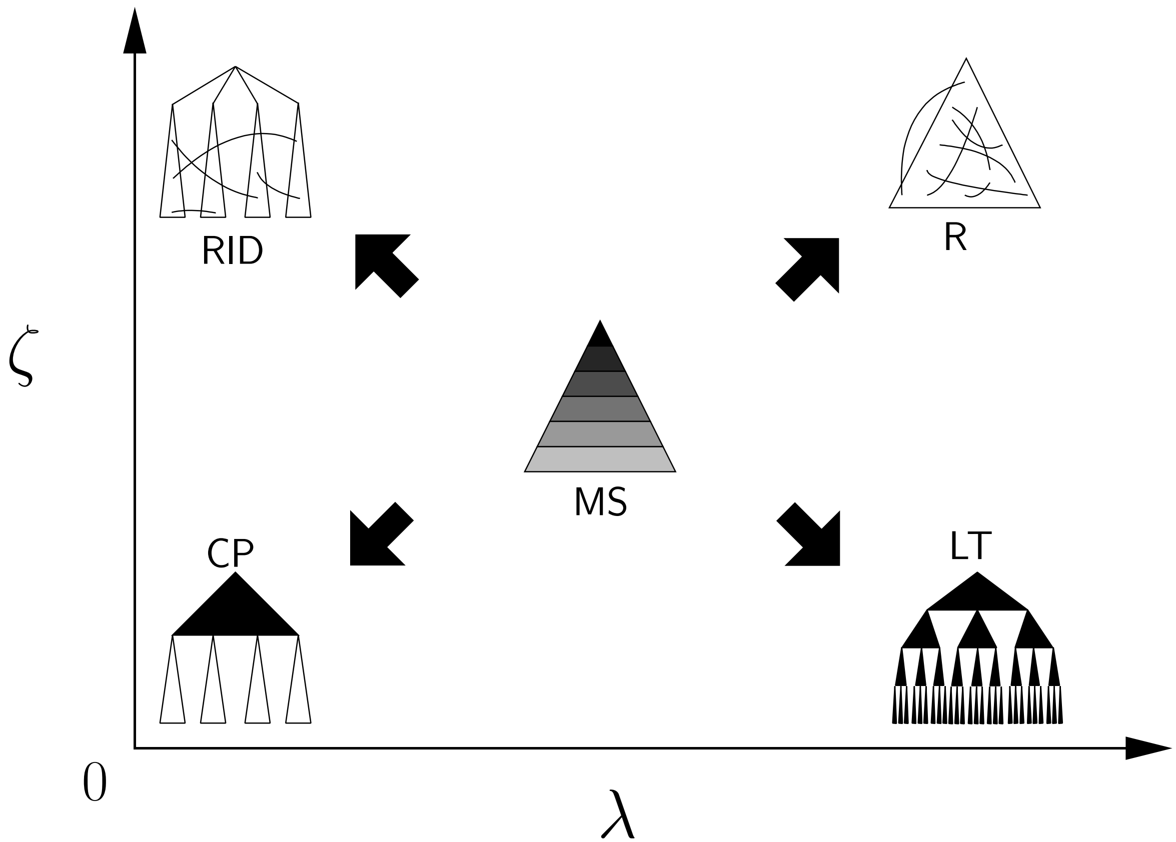
- Choose m links according to a two parameter probability distribution

  $0 \leq m \leq (N-1)(N-2)/2$

Model—construction:



Model—limiting cases:



Message passing pattern:

- Recipient of message chosen based on distance from sender.

$$P(\text{recipient at distance } d) \propto e^{-d/\xi}$$

$\xi = 0$: local message passing
$\xi = \infty$: random message passing

Message passing pattern:

Interpretations:

1. Sender knows specific recipient.

Message passing pattern:

Interpretations:

1. Sender knows specific recipient.

2. Sender requires certain kind of recipient.

Message passing pattern:

Interpretations:

1. Sender knows specific recipient.

2. Sender requires certain kind of recipient.

3. Sender seeks specific information but recipient unknown.

Message passing pattern:

Interpretations:

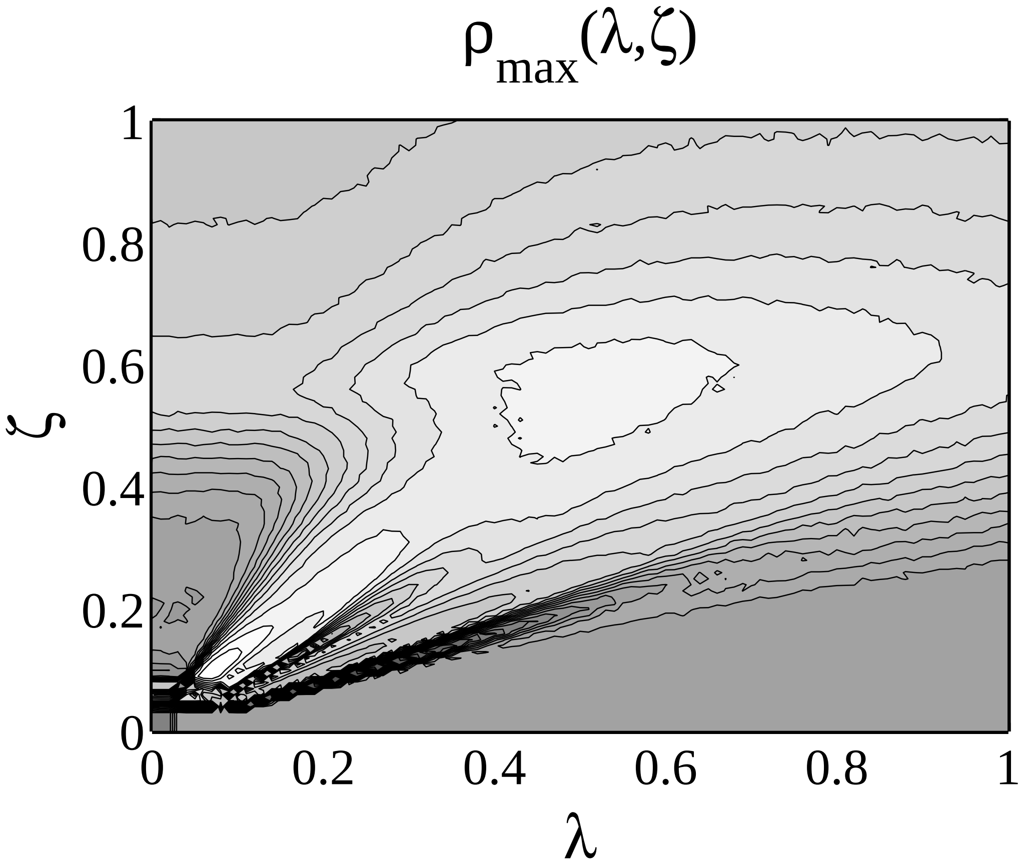
1. Sender knows specific recipient.

2. Sender requires certain kind of recipient.

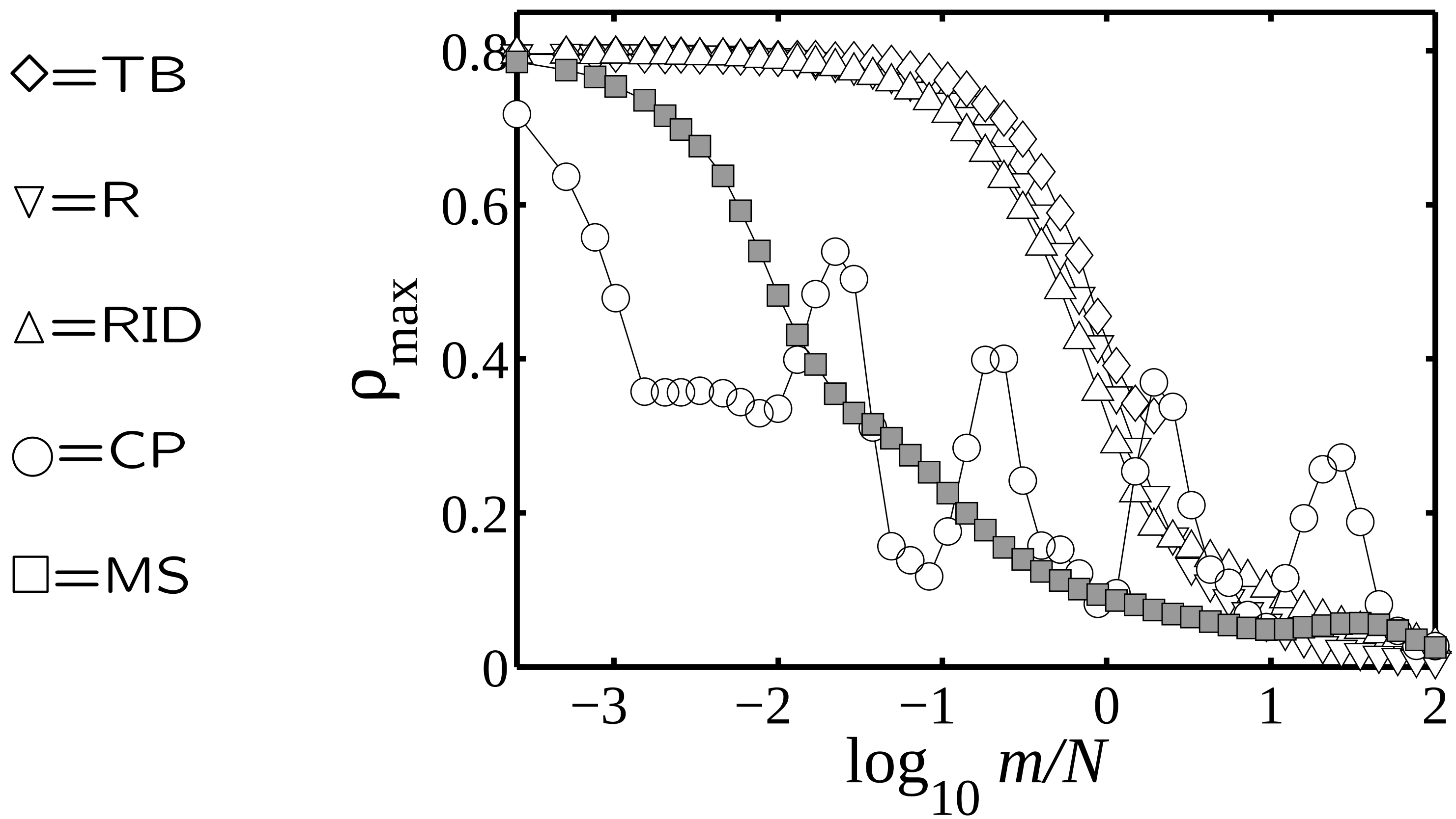
3. Sender seeks specific information but recipient unknown.

4. Sender has a problem but information/recipient unknown.

Results—congestion robustness:



Results—varying number of links added:



Results—varying message passing pattern:

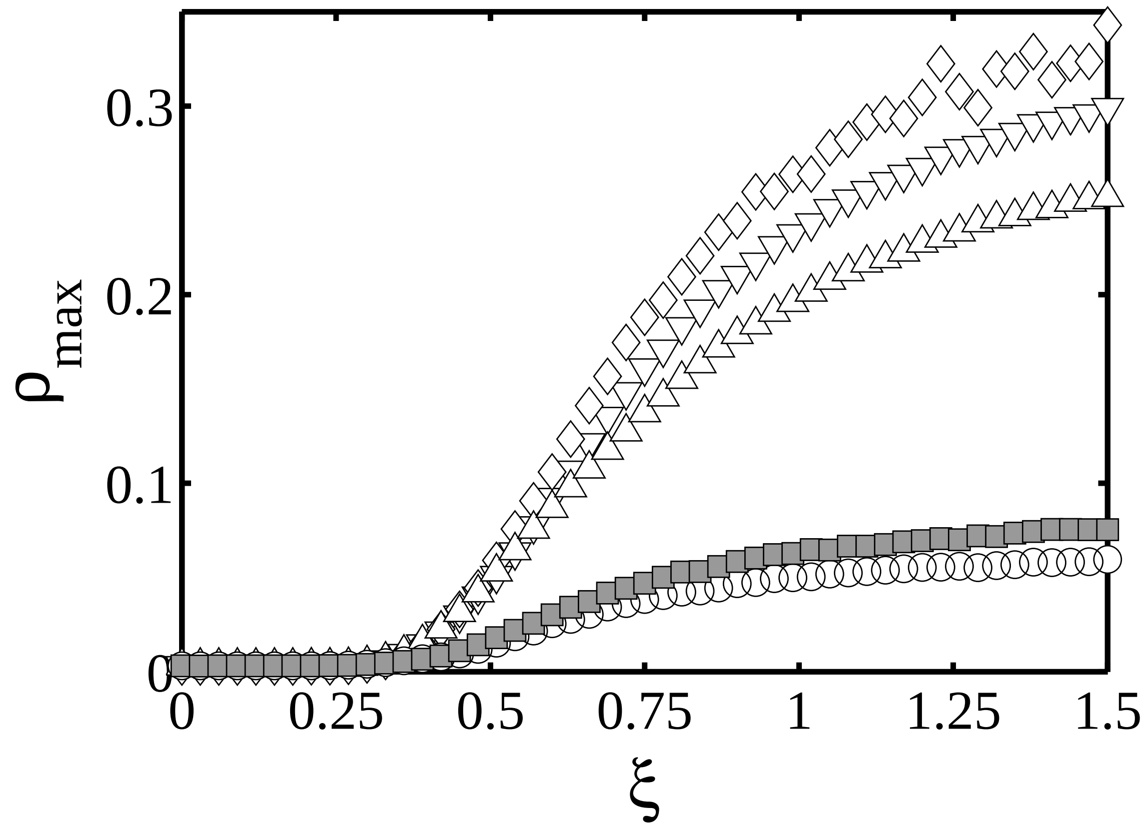
◇=TB

▽=R

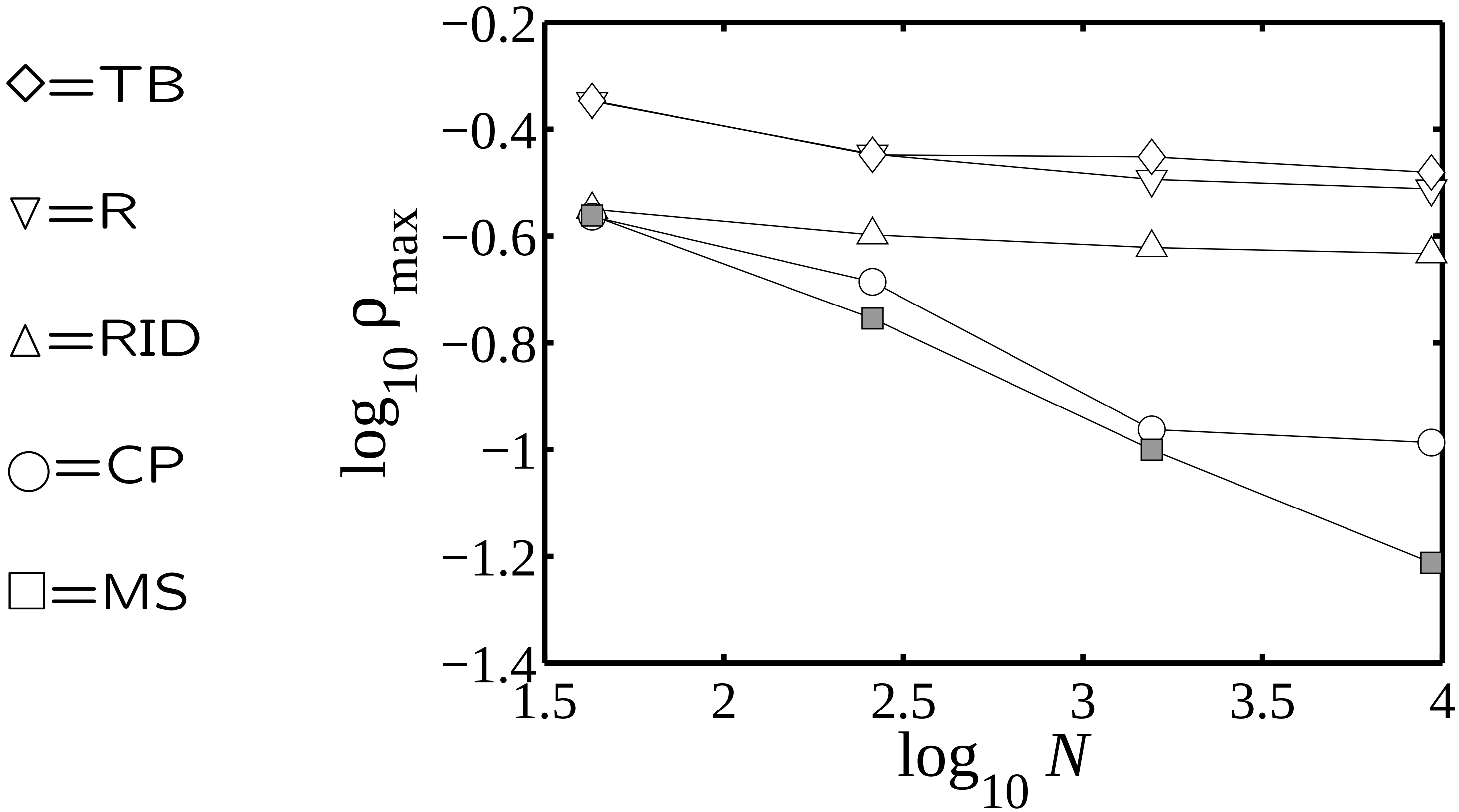
△=RID

○=CP

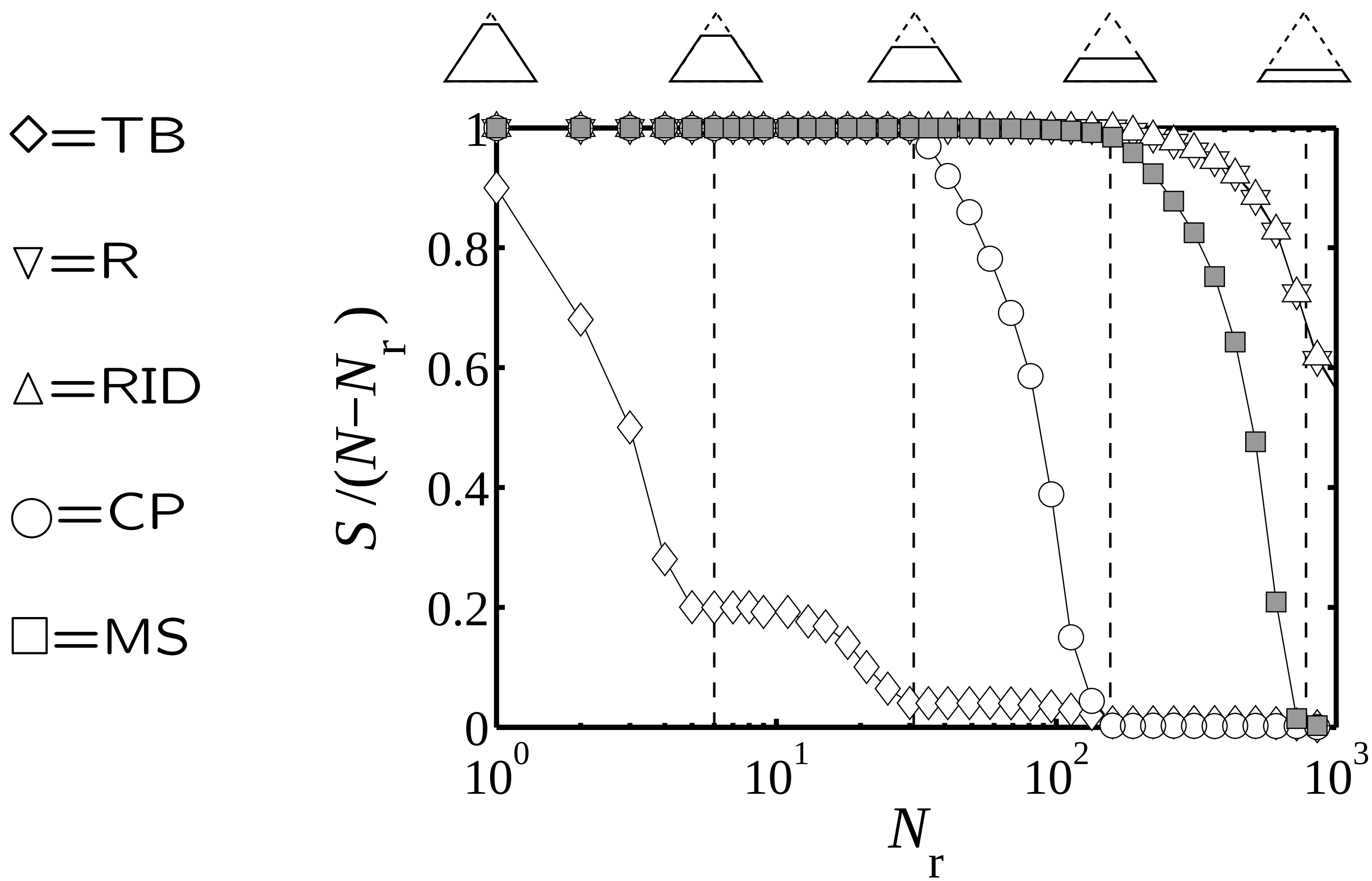
□=MS



Results—Scalability:



Results—connectivity robustness:



Some next steps:

- Explore searchability in setting of partial knowledge.
- Abstract model of query reformulation.
- Optimal arrangement of generalists and specialists.
- How many managers does an organization need?

References:

D. J. Watts, P. S. Dodds, & M. E. J. Newman.
"Identity and Search in Social Networks"
Science, **296**, 1302–1305, 2002.

P. S. Dodds, R. Muhamad, & D. J. Watts.
"An Experimental study of Search in Global Social Networks"
Science, **301**, 827–829, 2003.

P. S. Dodds, D. J. and Watts, & C. F. Sabel.
"Information Exchange and the Robustness of Organizational Networks."
Proc. Natl. Acad. Sci., (to appear September, 2003).