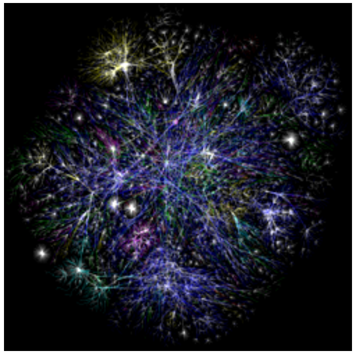
Navigation and Propagation in Networks

Michael Goodrich

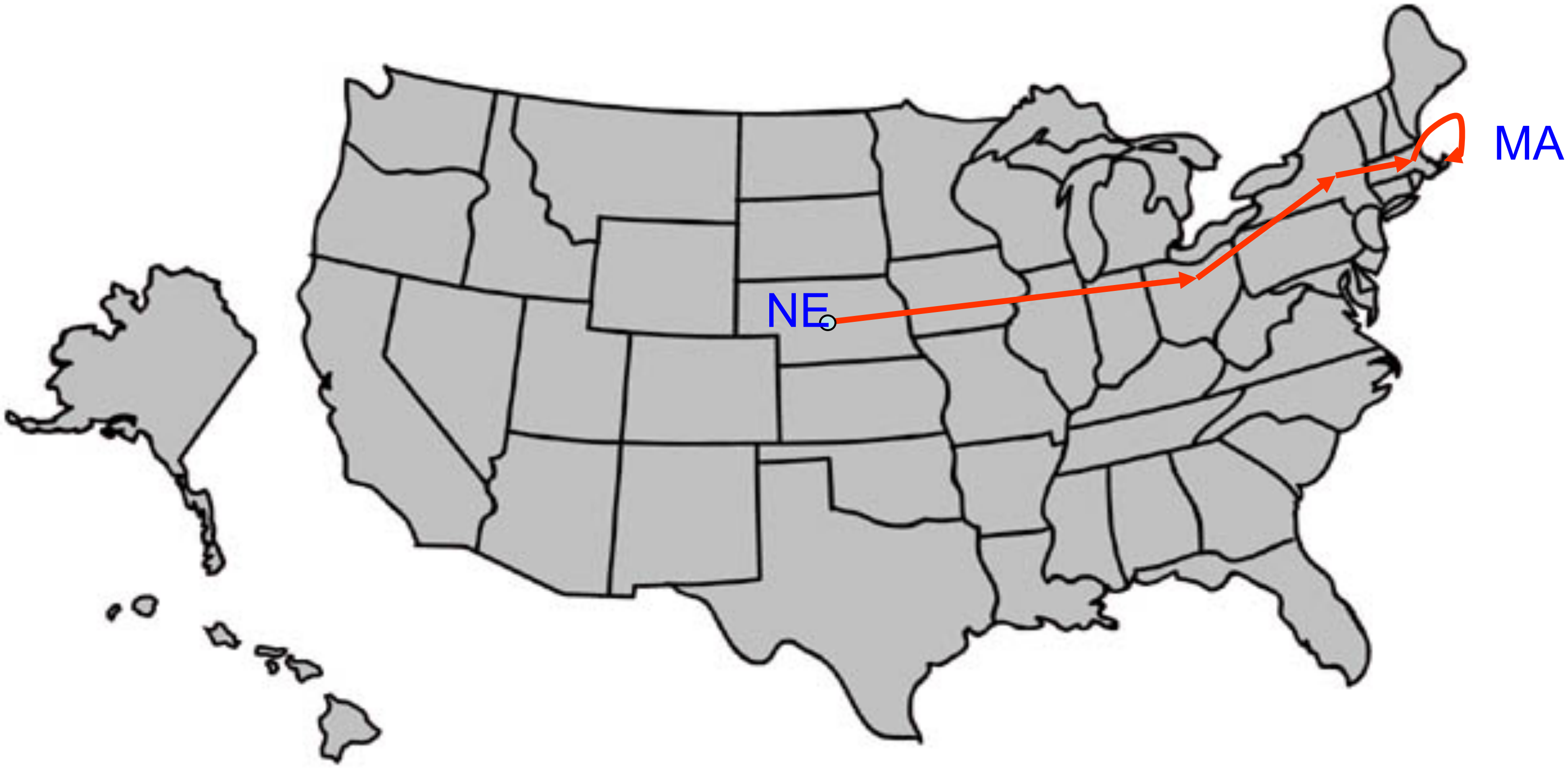


Some slides adapted from slides by Jean Vaucher, Panayiotis Tsaparas, Jure Leskovec, and Christos Faloutsos

Review:
Milgram's experiment

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

Small world phenomenon:
Milgram's experiment



Outcome:
average chain length was between 5 and 6

"Six degrees of separation"

Small world phenomenon:
Milgram's experiment repeated

email experiment
Dodds, Muhamad, Watts, Science 301, (2003)

- 18 targets
- 13 different countries

- 60,000+ participants
- 24,163 message chains
- 384 reached their targets
- average path length 4.0

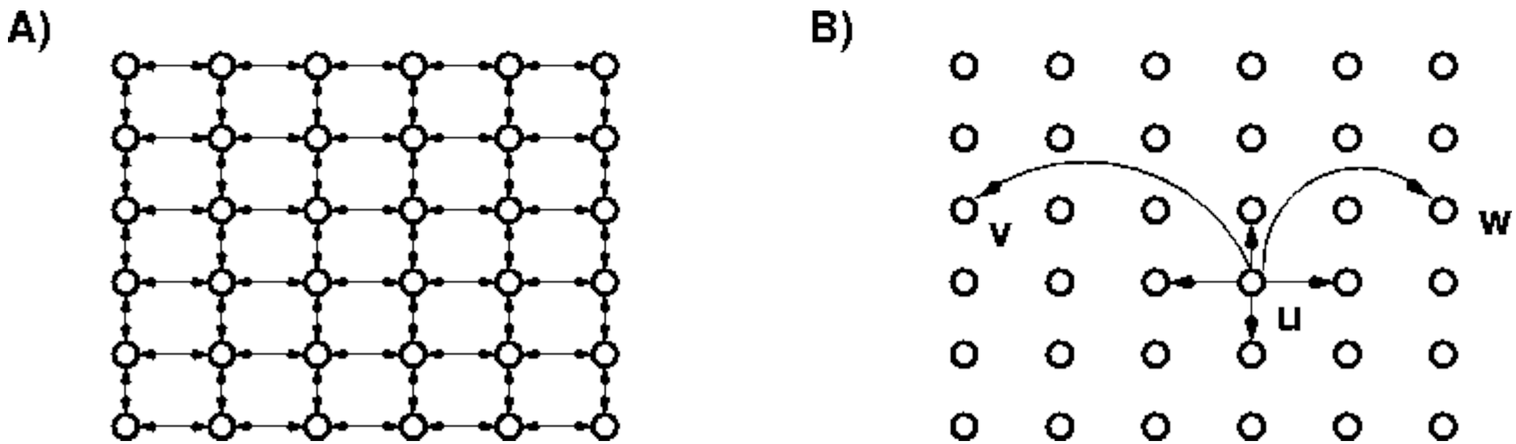
Source: NASA, U.S. Government; http://visibleearth.nasa.gov/view_rec.php?id=2429

Milgram's experiment revisited

- What did Milgram's experiment show?
  - (a) There are short paths in large networks that connect individuals
  - (b) People are able to find these short paths using a simple, greedy, decentralized algorithm
- Small world models take care of (a)
- Kleinberg: what about (b)?

Kleinberg's model

- Consider a directed 2-dimensional lattice
- For each vertex u add q shortcuts
  - choose vertex v as the destination of the shortcut with probability proportional to $[d(u,v)]^{-r}$
  - when $r = 0$, we have uniform probabilities

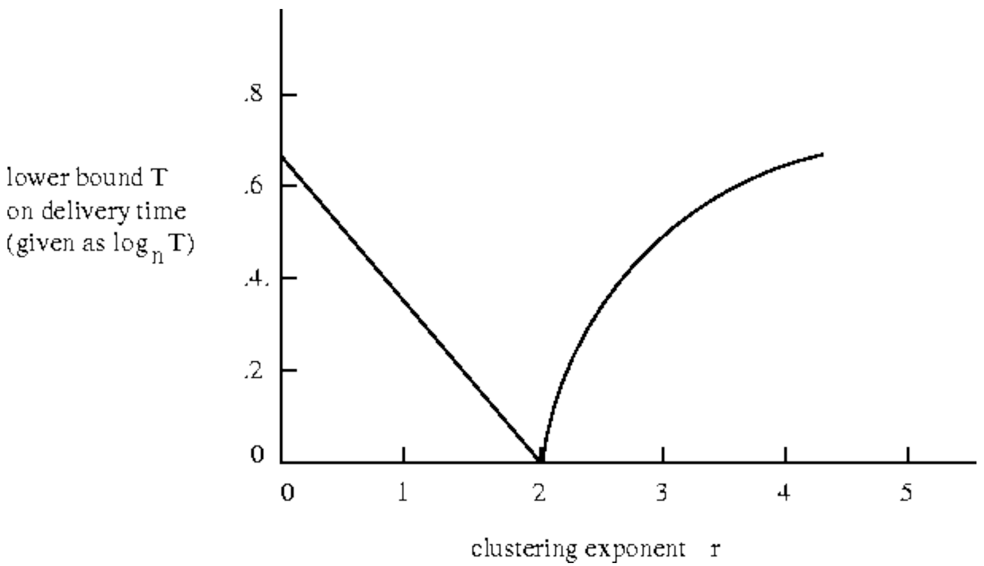


Searching in a small world

- Given a source s and a destination t, define a greedy local search algorithm that
  1. knows the positions of the nodes on the grid
  2. knows the neighbors and shortcuts of the current node
  3. knows the neighbors and shortcuts of all nodes seen so far
  4. operates greedily, each time moving as close to t as possible
- Kleinberg proved the following
  - When $r=2$, an algorithm that uses only local information at each node (not 2) can reach the destination in expected time $O(\log^2 n)$.
  - When $r<2$ a local greedy algorithm (1-4) needs expected time $\Omega(n^{(2-r)/3})$.
  - When $r>2$ a local greedy algorithm (1-4) needs expected time $\Omega(n^{(r-2)/(r-1)})$.

Searching in a small world

- For $r < 2$, the graph has paths of logarithmic length (small world), but a greedy algorithm cannot find them
- For $r > 2$, the graph does not have short paths
- For $r = 2$ is the only case where there are short paths, and the greedy algorithm is able to find them

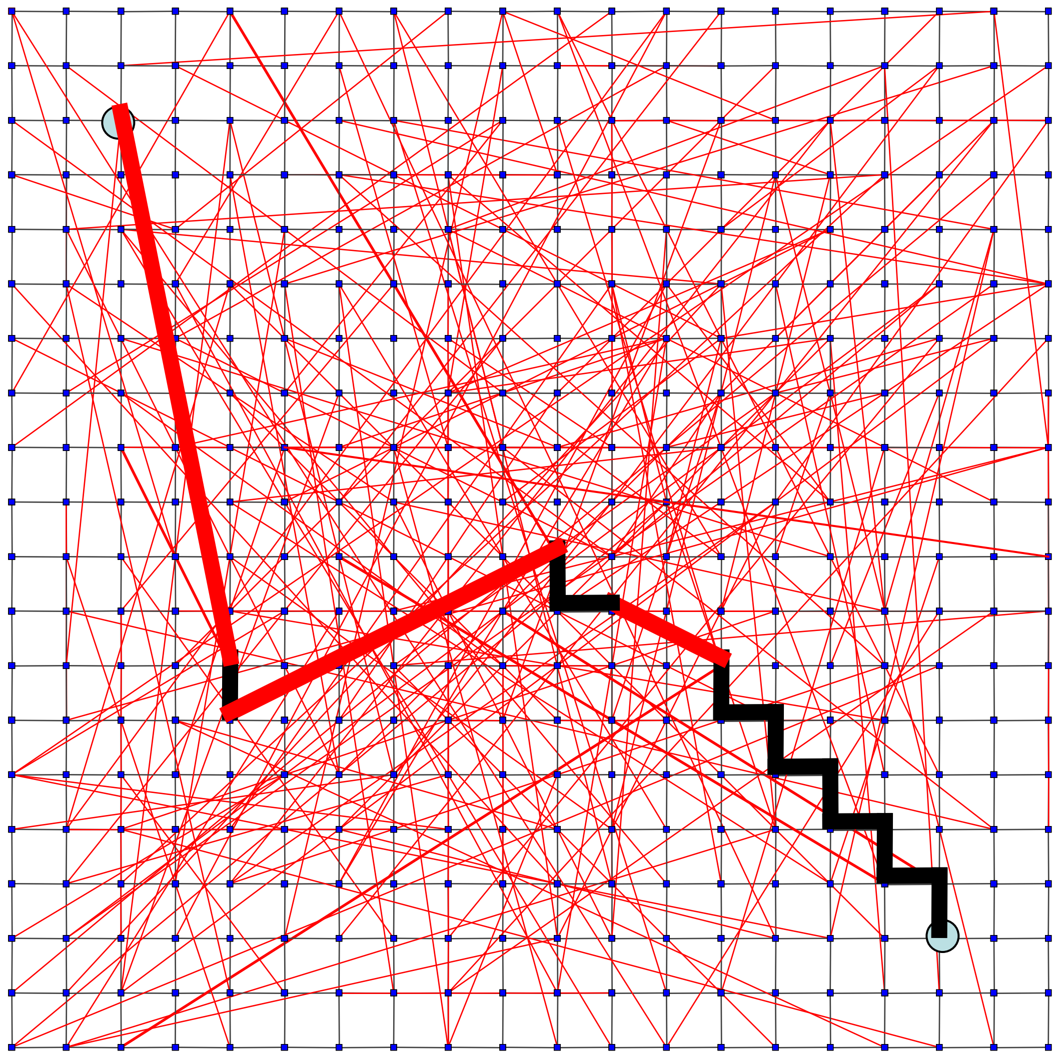


geographical search when network lacks locality

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

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

$p \sim p_0$

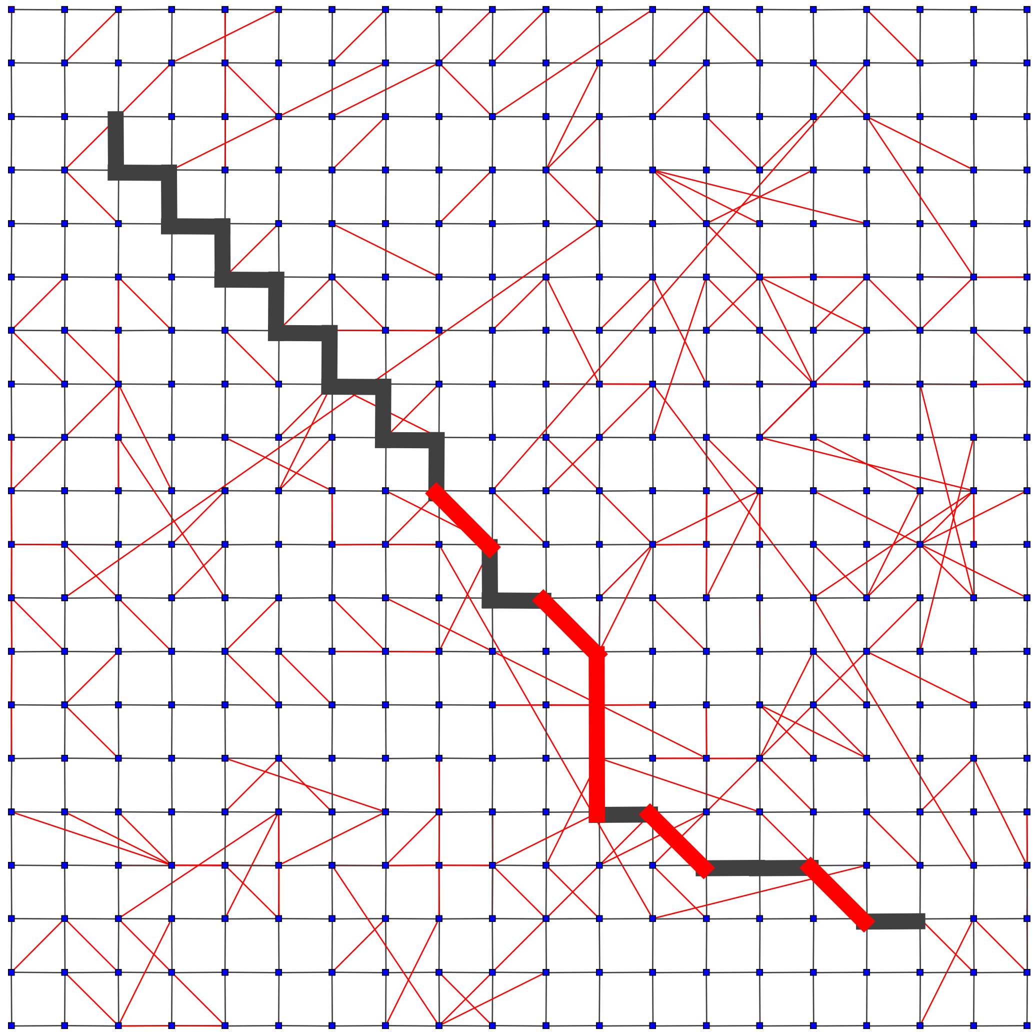


When **r<2**, expected time at least $\alpha_r n^{(2-r)/3}$

Overly localized links on a lattice

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

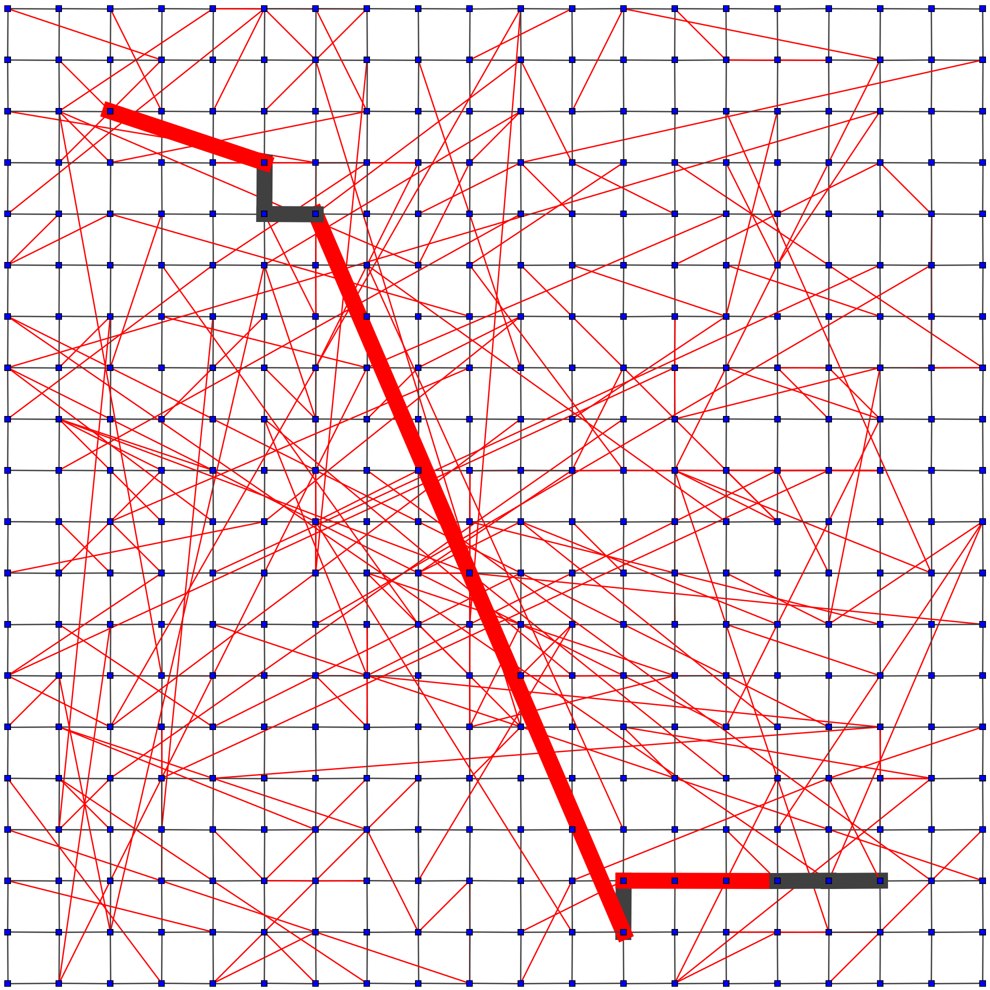
$$p \sim \frac{1}{d^4}$$



geographical small world model
Links balanced between long and short range

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

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



Extensions

- If there are log n shortcuts, then the search time is $O(\log n)$
  - we save the time required for finding the shortcut

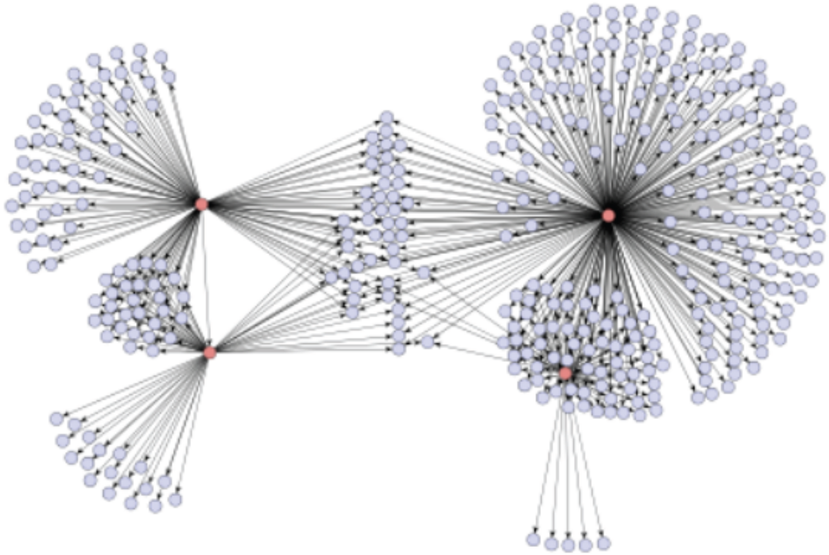
- If we know the shortcuts of log n neighbors the time becomes $O(\log^{1+1/d} n)$

Small Worlds

& Epidemic diseases

- Nodes are living entities
- Link is contact
- 3 States
  - Uninfected
  - Infected
  - Recovered (or dead)

Diffusion in Social Networks



- One of the networks is a spread of a disease, the other one is product recommendations
- Which is which?

Diffusion in Social Networks

- A fundamental process in social networks: Behaviors that cascade from node to node like an epidemic
  - News, opinions, rumors, fads, urban legends, …
  - Word-of-mouth effects in marketing: rise of new websites, free web based services
  - Virus, disease propagation
  - Change in social priorities: smoking, recycling
  - Saturation news coverage: topic diffusion among bloggers
  - Internet-energized political campaigns
  - Cascading failures in financial markets
  - Localized effects: riots, people walking out of a lecture

Failures in networks

- Fault propagation or viruses
- Scale-free networks are far more resistant to random failures than ordinary random networks
  - because of most nodes are leaves
- But failure of hubs can be catastrophic vulnerable or targets of deliberate attacks
  - which may make scale-free networks *more* vulnerable to deliberate attacks
- Cascades of failures

Effect of peers & pundits (hubs and authorities)

- People's decisions are affected by what others do and think
  - Pressure to conform?
- Efficient strategy when insufficient knowledge or expertise
  - Ex: picking a restaurant
- Google's PageRank is a score for influential nodes in a network (the WWW)