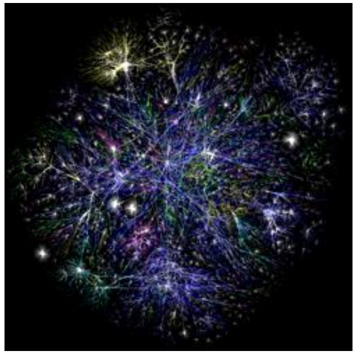
Information Networks

Small World Networks

Lecture 5



Announcement

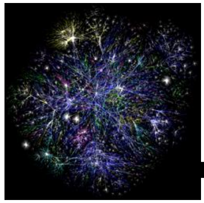
- The first assignment is out
- There will be a tutorial this Monday, April 4 where Evimaria will present some helpful material and you can also ask questions about the assignment

Small world Phenomena

- So far we focused on obtaining graphs with power-law distributions on the degrees. What about other properties?
  - Clustering coefficient: real-life networks tend to have high clustering coefficient
  - Short paths: real-life networks are "small worlds"
  - Can we combine these two properties?

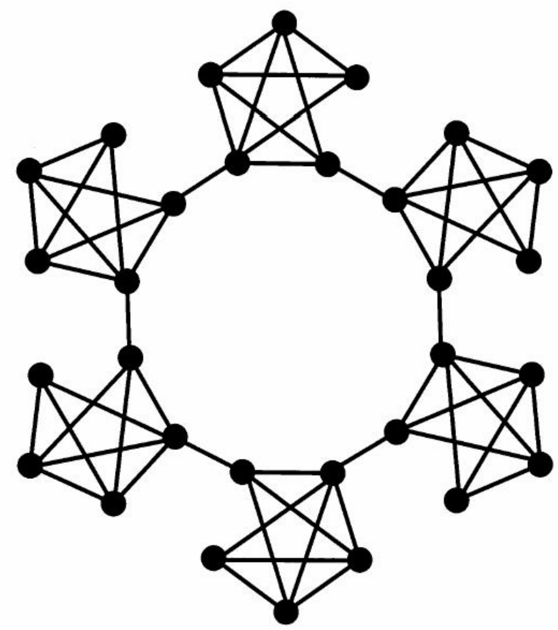
Small-world Graphs

- According to Watts [W99]
  - Large networks (n >> 1)
  - Sparse connectivity (avg degree k << n)
  - No central node (k_{max} << n)
  - Large clustering coefficient (larger than in random graphs of same size)
  - Short average paths (~log n, close to those of random graphs of the same size)



The Caveman Model [W99]

- The random graph
  - edges are generated completely at random
  - low avg. path length $L \leq \log n / \log k$
  - high clustering coefficient $C \sim k/n$
- The Caveman model
  - edges follow a structure
  - high avg. path length $L \sim n/k$
  - high clustering coefficient $C \sim 1\text{-}O(1/k)$



- Can we interpolate between the two?

Mixing order with randomness

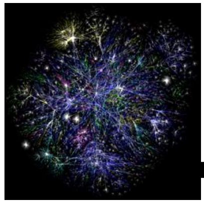
- Inspired by the work of Solmonoff and Rapoport
  - nodes that share neighbors should have higher probability to be connected
- Generate an edge between i and j with probability proportional to R_{ij}

$$R_{ij} = \begin{cases} 1 & \text{if } m_{ij} \ge 0 \\ \left(\frac{m_{ij}}{k}\right)^{\alpha}(1-p)+p & \text{if } 0 < m_{ij} < k \\ p & \text{if } m_{ij} = 0 \end{cases}$$

m_{ij} = number of common neighbors of i and j

p = very small probability

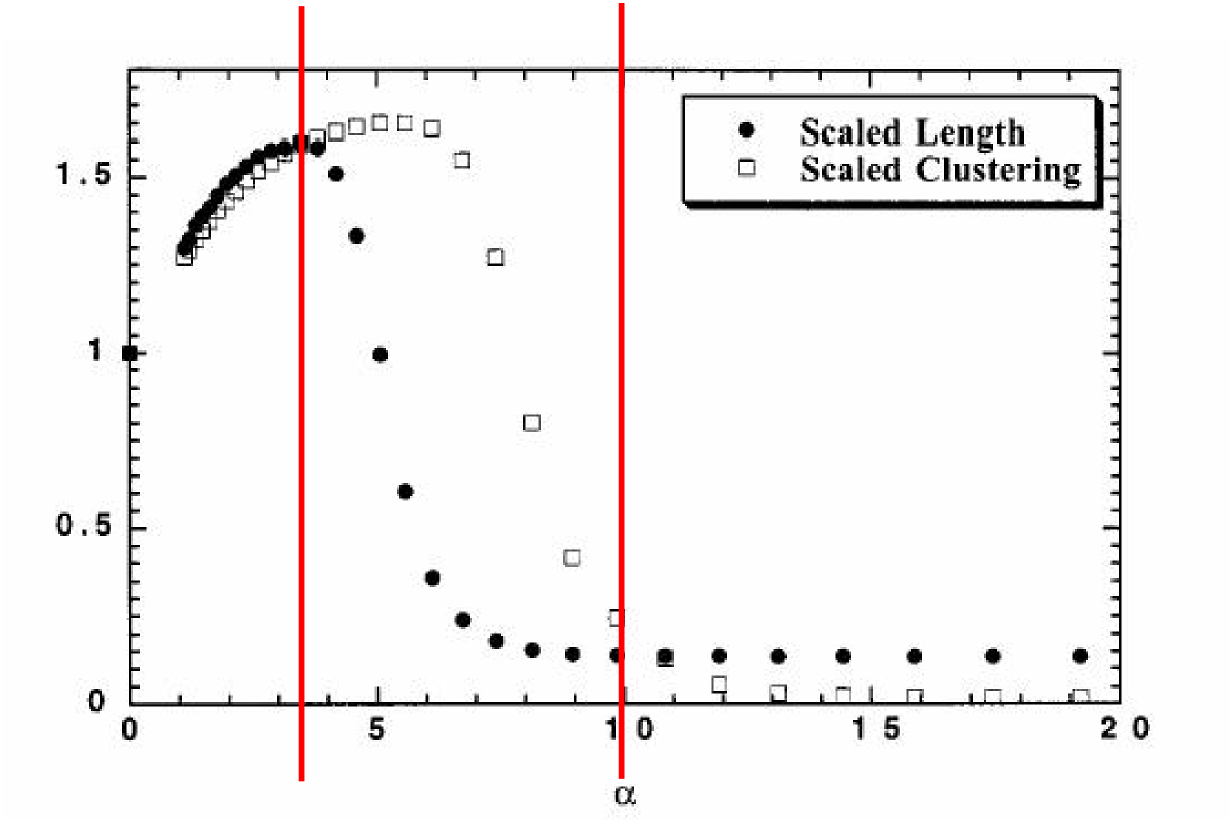
- When $\alpha = 0$, edges are determined by common neighbors
- When $\alpha = \infty$ edges are independent of common neighbors
- For intermediate values we obtain a combination of order and randomness



Algorithm

- Start with a ring
- For $i = 1 \ldots n$
  - Select a vertex j with probability proportional to R_{ij} and generate an edge (i,j)
- Repeat until k edges are added to each vertex

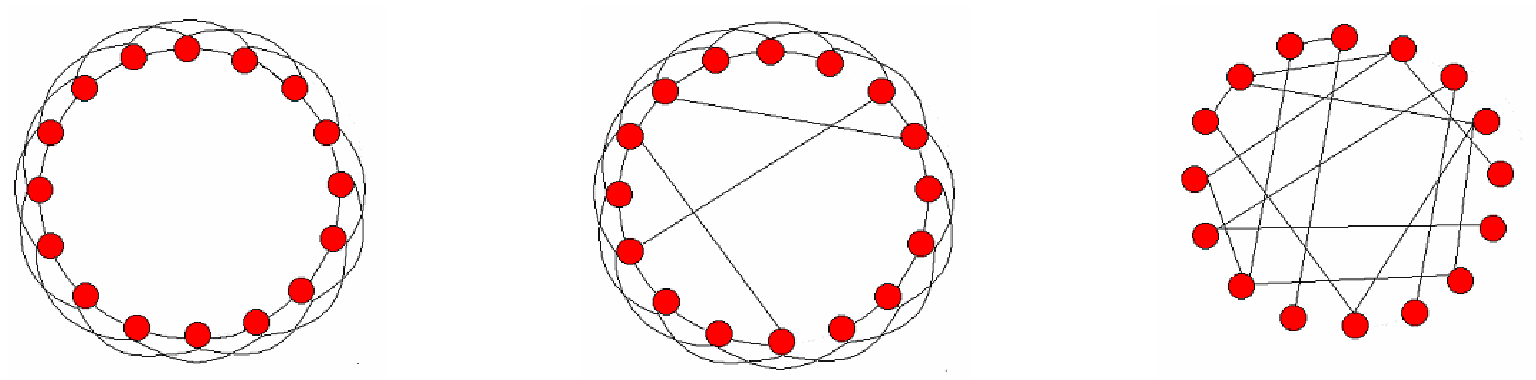
Clustering coefficient – Avg path length



small world graphs

Watts and Strogatz model [WS98]

- Start with a ring, where every node is connected to the next k nodes
- With probability p, rewire every edge (or, add a shortcut) to a uniformly chosen destination.
  - Granovetter, “The strength of weak ties”



order

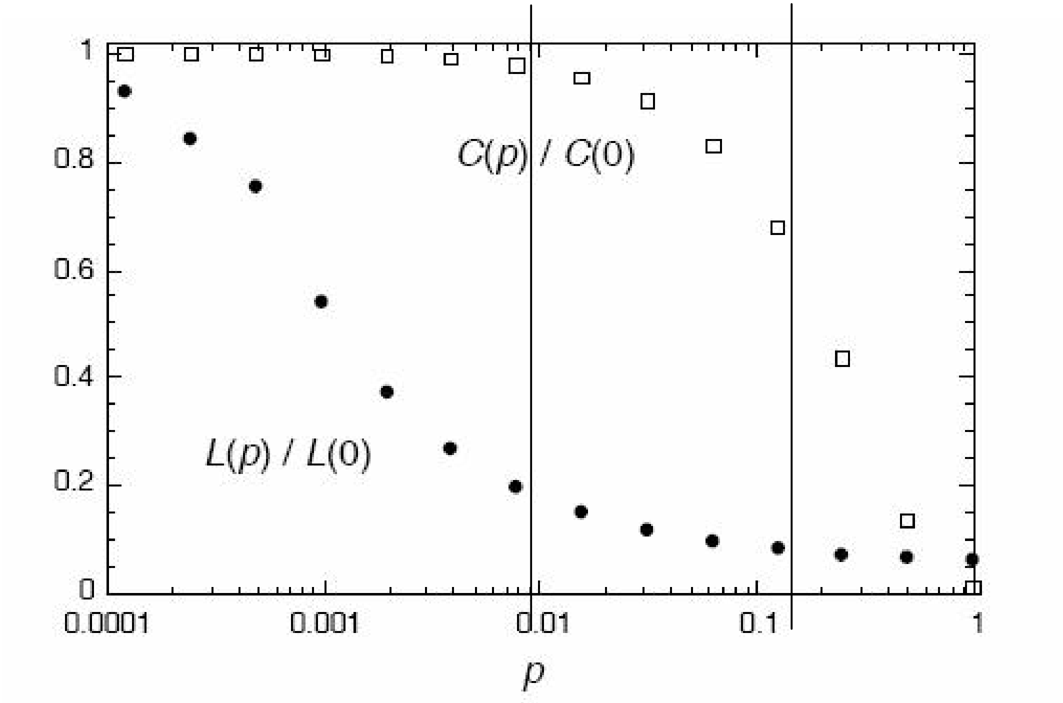
randomness

p = 0

0 < p < 1

p = 1

Clustering Coefficient – Characteristic Path Length



log-scale in p

When p = 0, C = 3(k-2)/4(k-1) ~ ¾
L = n/k

For small p, C ~ ¾
L ~ logn

Graph Theory Results

- Graph theorist failed to be impressed. Most of these results were known.
- Bolobas and Chung 88
  - superimposing a random matching to a ring yields diameter O(logn)

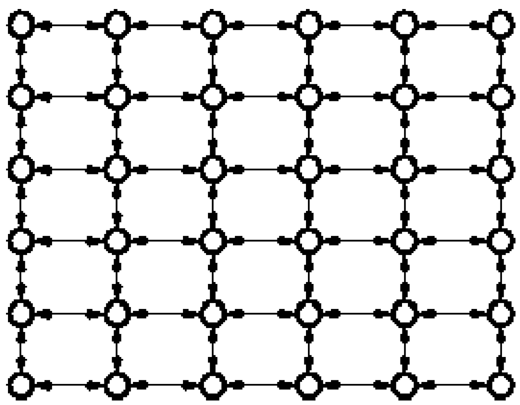
Milgram's experiment revisited

- 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)?

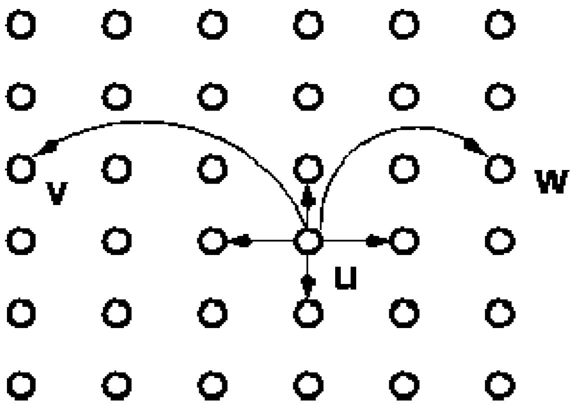
Kleinberg's model

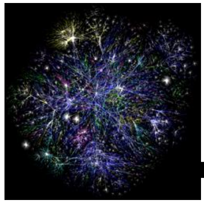
- Consider a directed 2-dimensional lattice
- For each vertex u add q shortcuts
  - choose vertex v as the destination of the shortcut with probability proportional to $[d(u,v)]^{-r}$
  - when $r = 0$, we have uniform probabilities

A)



B)



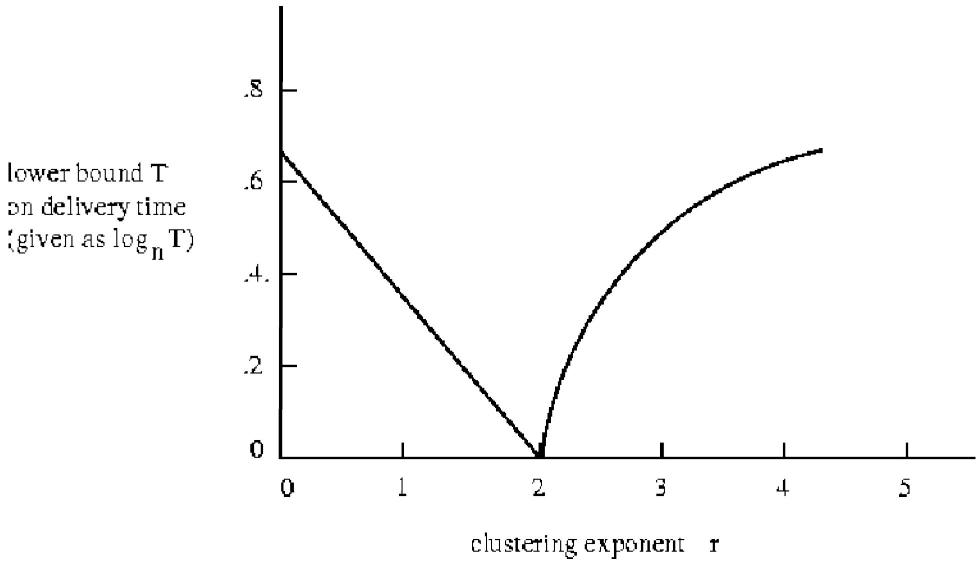


Searching in a small world

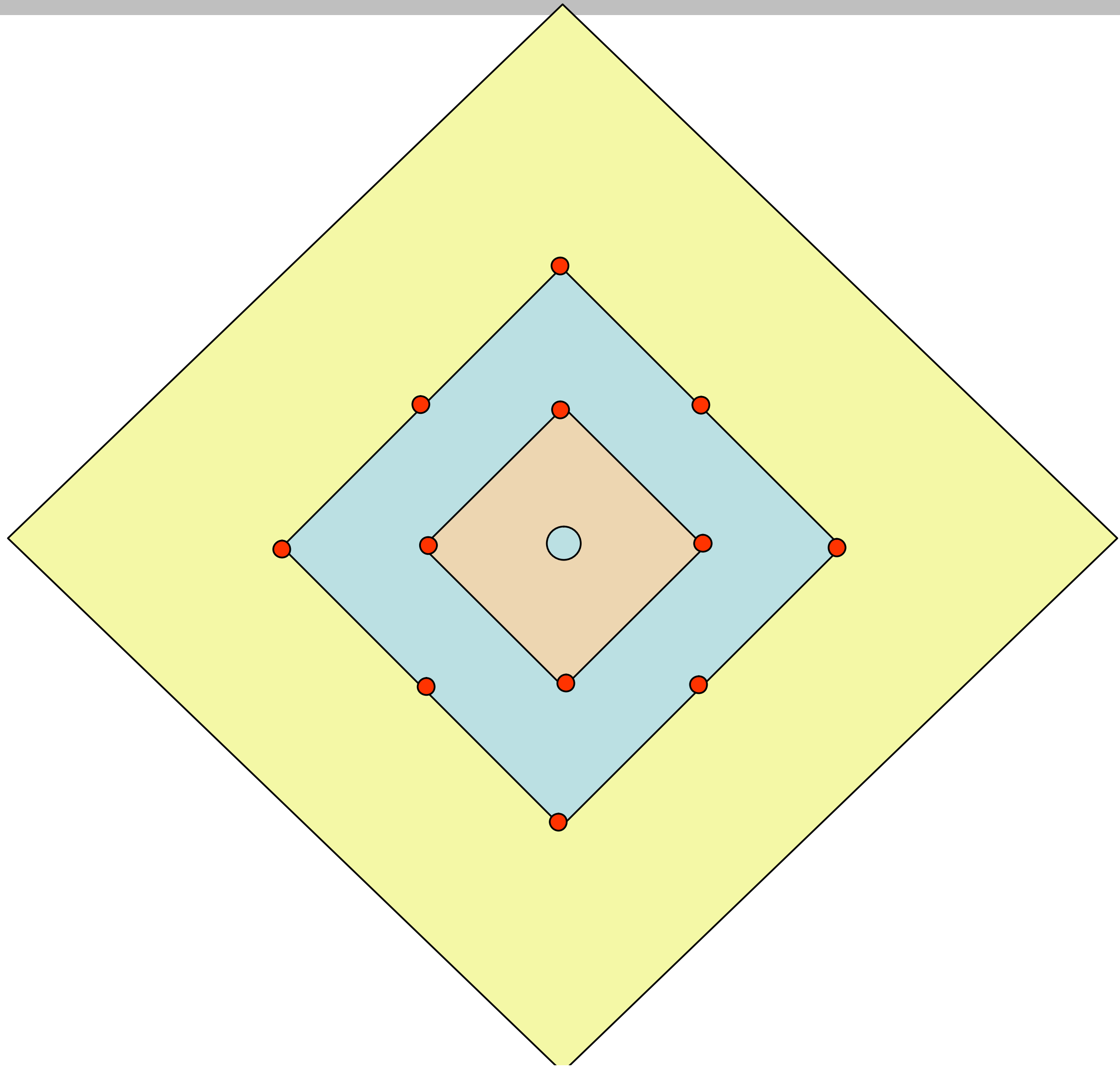
- Given a source s and a destination t, define a greedy local search algorithm that
  1. knows the positions of the nodes on the grid
  2. knows the neighbors and shortcuts of the current node
  3. knows the neighbors and shortcuts of all nodes seen so far
  4. operates greedily, each time moving as 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 $\Omega(n^{(r-2)/(r-1)})$.
  - Generalizes for a d-dimensional lattice, when $r=d$ (query time is independent of the lattice dimension)
    - $d = 1$, the Watts-Strogatz model

Searching in a small world

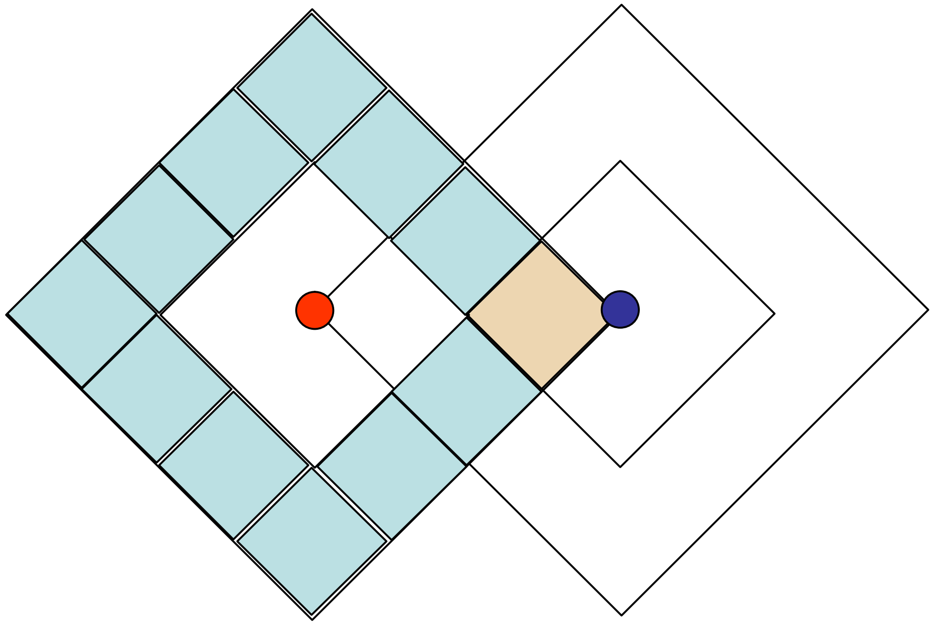
- For r < 2, the graph has paths of logarithmic length (small world), but a greedy algorithm cannot find them
- For r > 2, the graph does not have short paths
- For r = 2 is the only case where there are short paths, and the greedy algorithm is able to find them



Proof of the upper bound



Proof of the upper bound



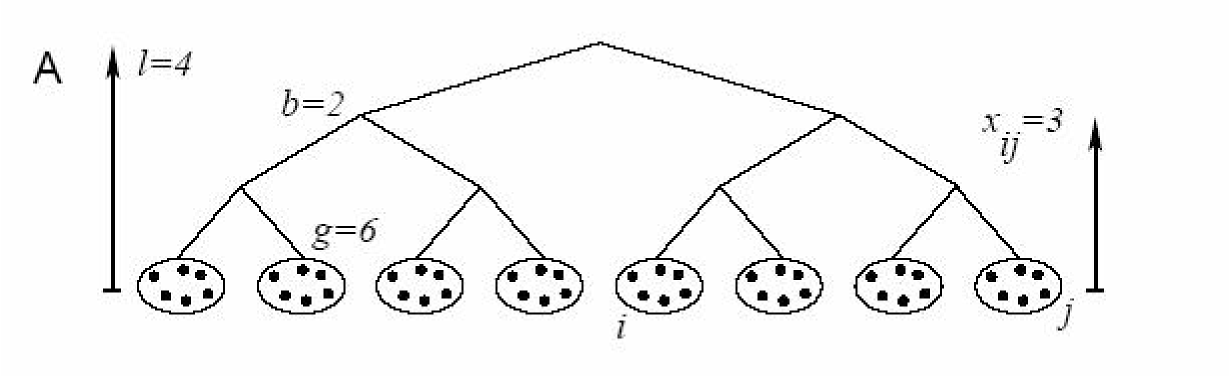
Extensions

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

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

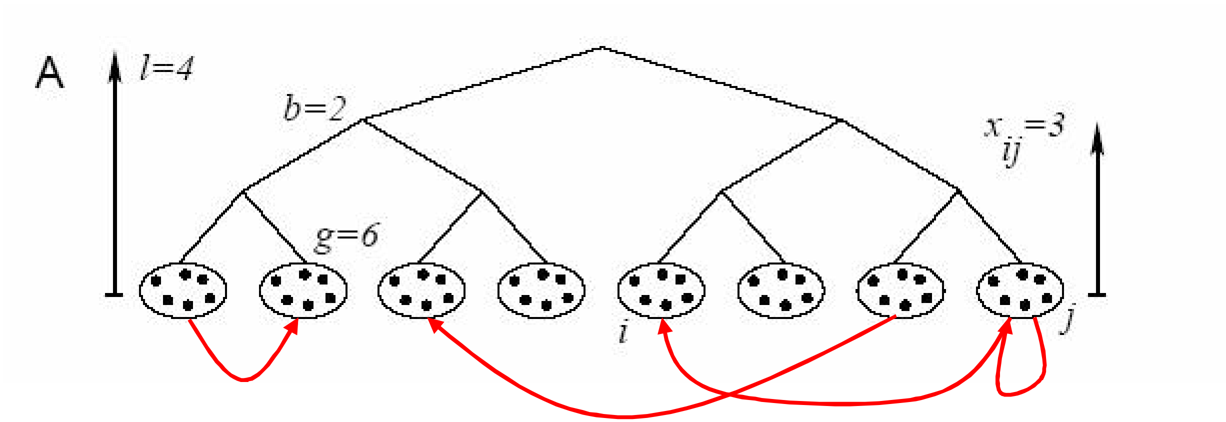
Other models

- Lattice captures geographic distance. How do we capture social distance (e.g. occupation)?
- Hierarchical organization of groups
  - distance h(i,j) = height of Least Common Ancestor



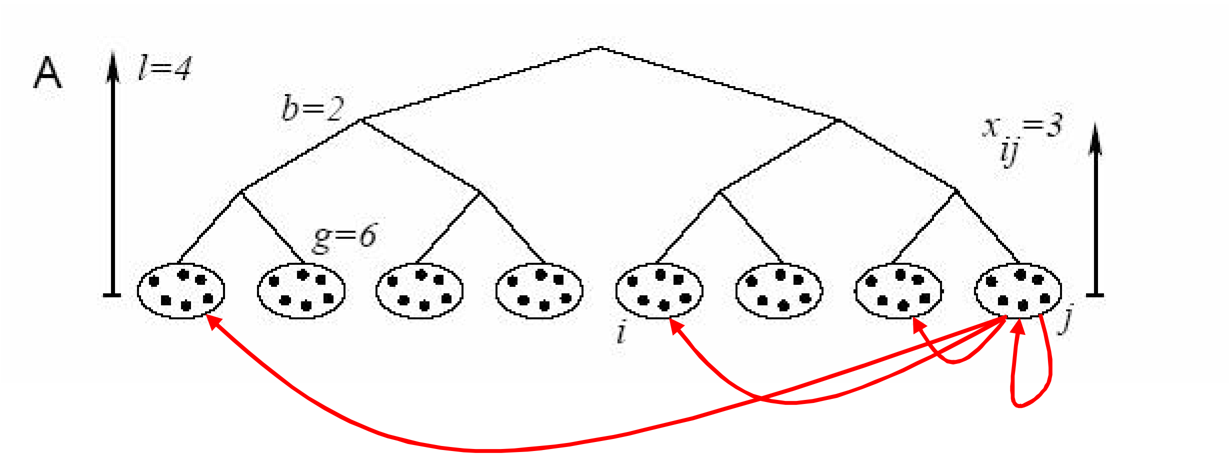
Other models

- Generate links between leaves with probability proportional to $b^{-\alpha h(i,j)}$
  - $b=2$ the branching factor



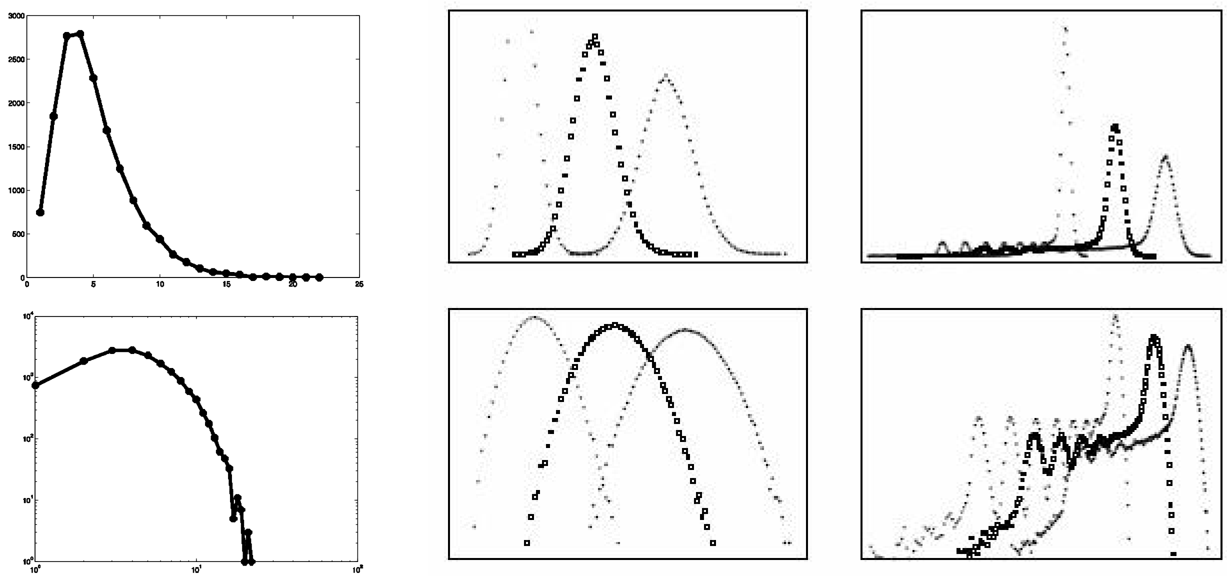
Other models

- Theorem: For α=1 there is a polylogarithimic search algorithm. For α≠1 there is no decentralized algorithm with poly-log time
  - note that α=1 and the exponential dependency results in uniform probability of linking to the subtrees



Degree distributions

- The small world models do not exhibit power law distributions



- Recently there are efforts towards creating scale free small world networks

Searching Power-law networks

- Kleinberg considered the case that you can fix your network as you wish. What if you cannot?
- [Adamic et al.] Instead of performing simple BFS flooding, pass the message to the neighbor with the highest degree
- Reduces the number of messages to $O(n^{(a-2)/(a-1)})$

References

- M. E. J. Newman, The structure and function of complex networks, SIAM Reviews, 45(2): 167-256, 2003
- B. Bollobas, Mathematical Results in Scale-Free random Graphs
- D.J. Watts. Networks, Dynamics and Small-World Phenomenon, American Journal of Sociology, Vol. 105, Number 2, 493-527, 1999
- Watts, D. J. and S. H. Strogatz. Collective dynamics of 'small-world' networks. Nature 393:440-42, 1998
- J. Kleinberg. The small-world phenomenon: An algorithmic perspective. Proc. 32nd ACM Symposium on Theory of Computing, 2000
- J. Kleinberg. Small-World Phenomena and the Dynamics of Information. Advances in Neural Information Processing Systems (NIPS) 14, 2001.
- Renormalization group analysis of the small-world network model, M. E. J. Newman and D. J. Watts, Phys. Lett. A 263, 341-346 (1999).
- Identity and search in social networks, D. J. Watts, P. S. Dodds, and M. E. J. Newman, Science 296, 1302-1305 (2002).
- Search in power-law networks, Lada A. Adamic, Rajan M. Lukose, Amit R. Puniyani, and Bernardo A. Huberman, Phys. Rev. E 64, 046135 (2001)