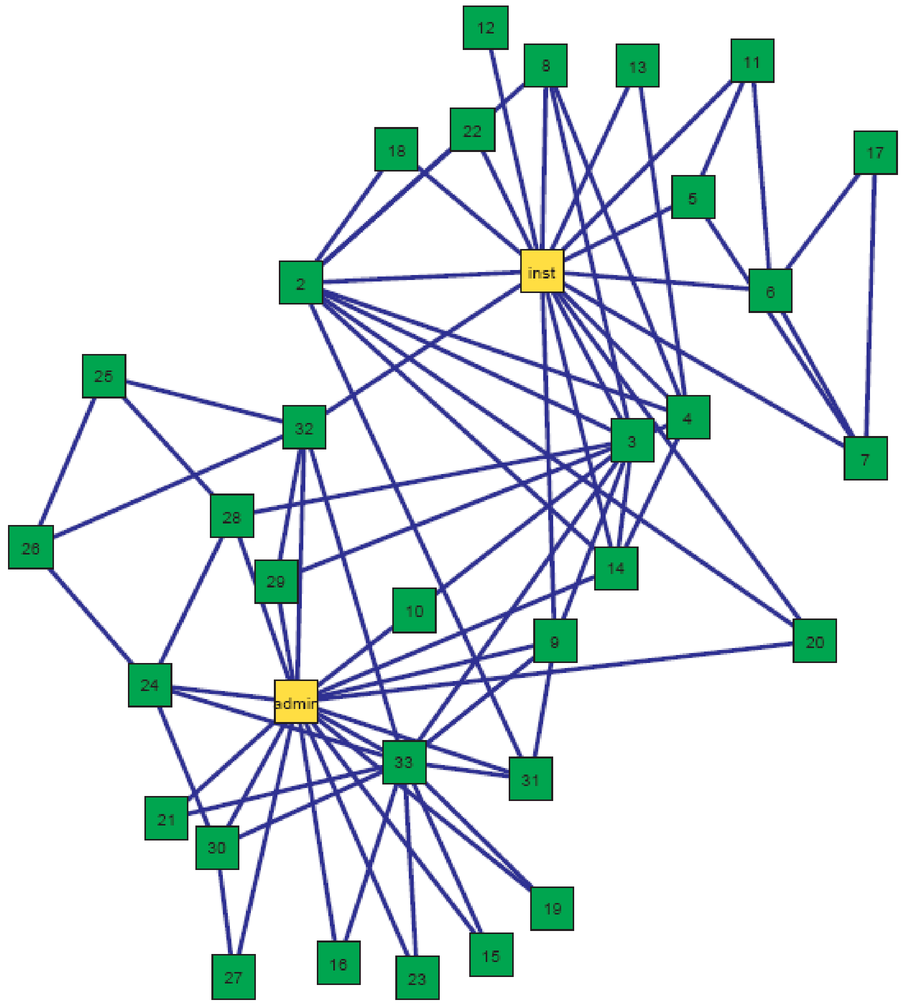
Distributed Computing over
Communication Networks:

Social Networks

Social Relationships in Karate Club (1970)



Observations?
Dispute caused split:
guess where!

Small World Networks: Phenomena, Vision, and Science

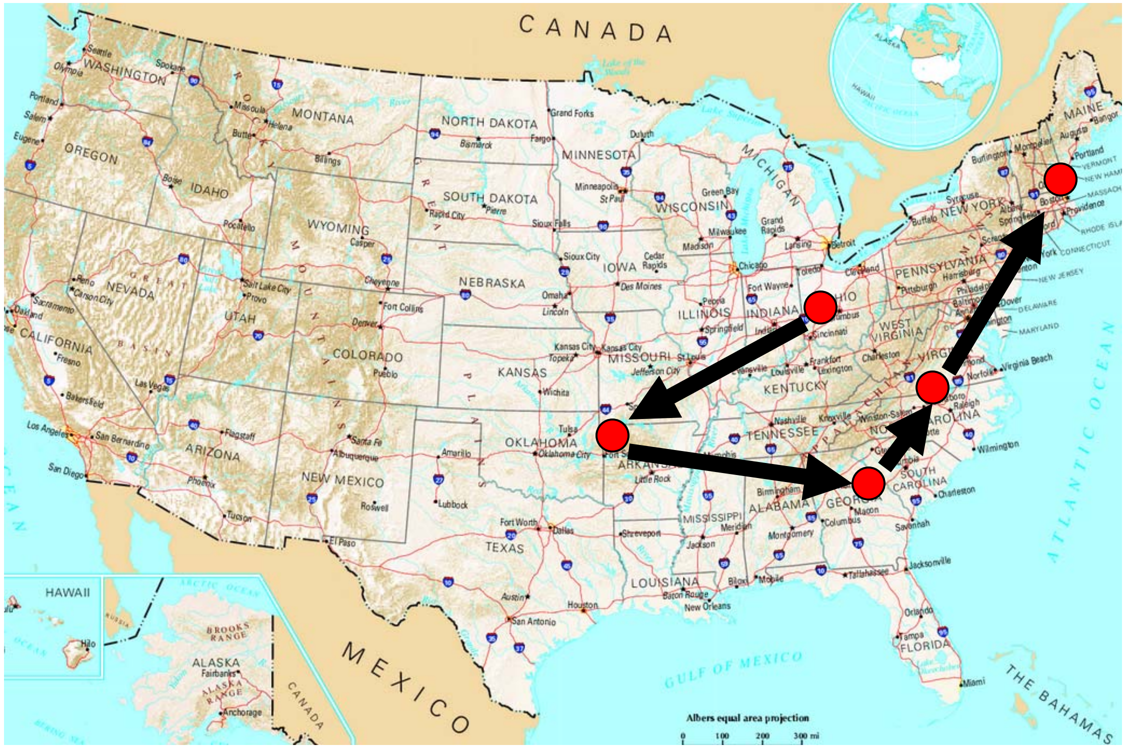
Karinthy (1929): „World is shrinking as humans are more connected!“
(inspired by Marconi?)

Hot topic in 60s:
- McLuhan coined „Global Village“
- Milgram‘s experiment: average path length between „random people“

Milgram‘s experiment:
- Choose random people from US Midwest
- Tell them to send letter to some guy in Boston
- They are only allowed to forward letter to someone they know on first name basis!
- How many „hops“ until letters arrive? (What do you think?)

Milgram‘s Experiment

How many hops until letters arrive?



Results and Discussion

Results:

- Many letters got lost
- But for the ones that arrived, the average hop distance was 5.5!
- „six-degrees-of-separation“ / „small world“

How to explain? Not only small diameter but also „navigable“!

Still an important research question!

- E.g., concept of „power-law“ graphs: node degrees are distributed according to power law, i.e., number of nodes with degree δ is proportional to δ^{α} for some α>0
- Power law graphs have been observed also in the Internet, in biology, in physics, etc.
- Kleinberg‘s explanation: regular grid with a small number of random links!

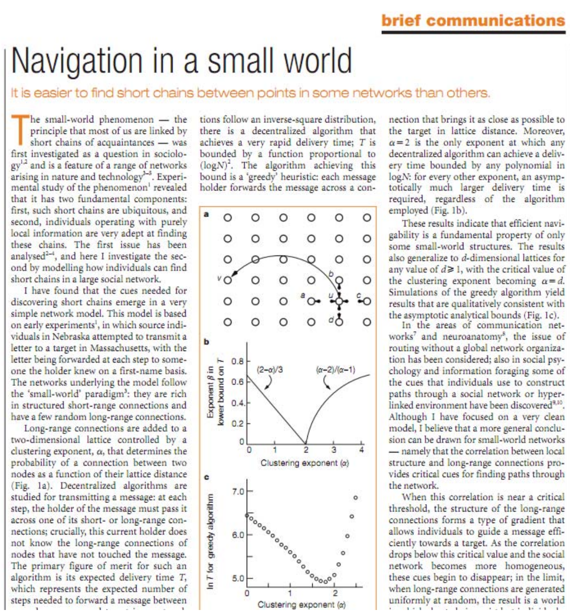
Discussion

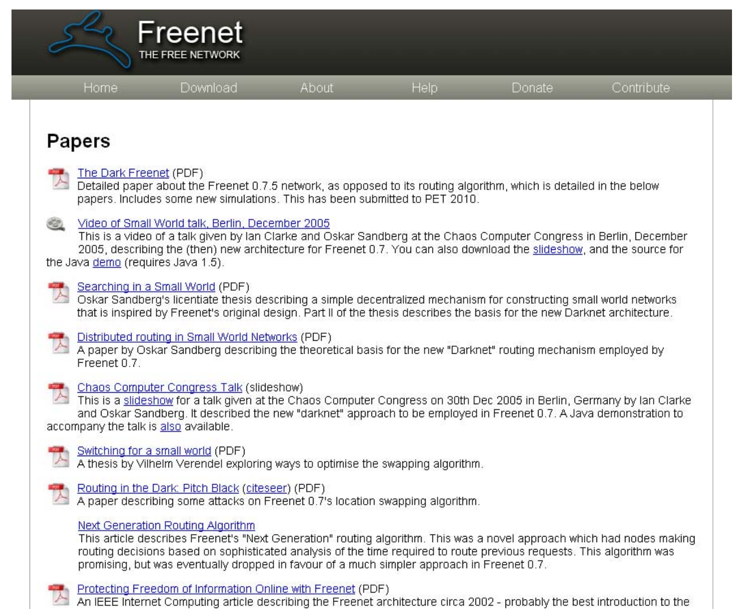
- Implications of people being closely connected?!
- Many people criticize experiment, or generalization of insight: e.g., letter from white person to black person needs more hops, etc.

Implication for Computer Science?

The good properties of social networks have inspired scientists and practitioners to build computer networks with similar characteristics!

Example: Jon Kleinberg at Cornell discovered navigable („greedy routable“) social networks, which inspired Clarke/Sandberg/... to build the peer-to-peer system Freenet.





Watts-Strogatz Model

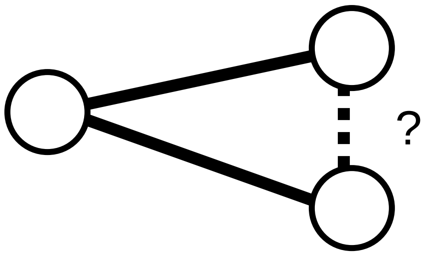
Idea: Model social network as combination of two networks!

(1) As a basis network with large cluster coefficient,
(2) and then add random links (e.g., constant number of random nodes all over the graph): long-range acquaintances

Important concept in social networks:

Cluster Coefficient

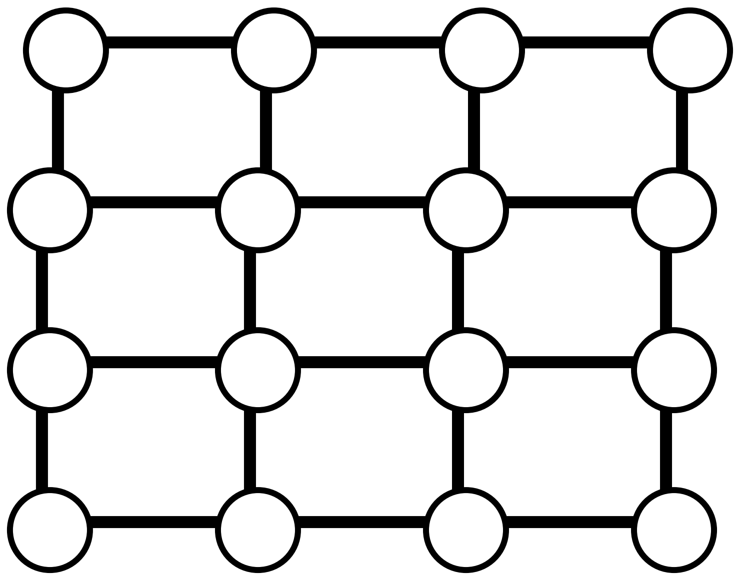
Probability that two friends of a node are also friends, summed up over all nodes.



Problem:
Is clustering a good measure?
And how to navigate??

Cluster Coefficient

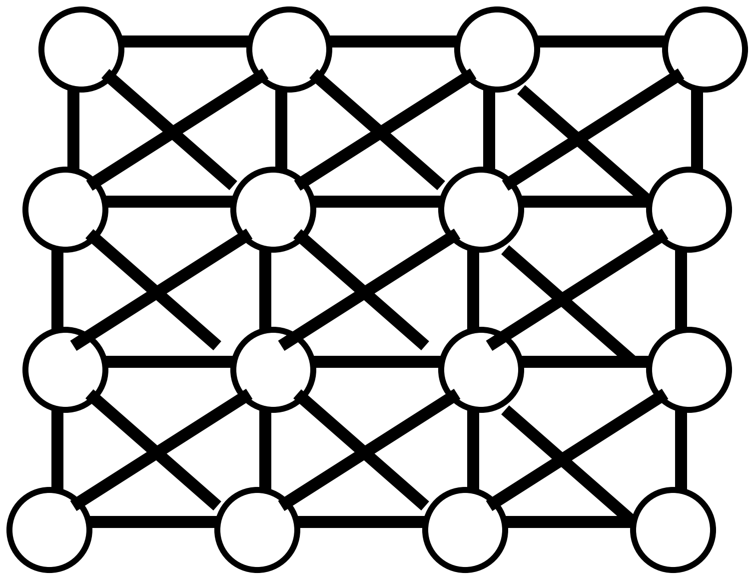
Careful, not so useful in the „worst-case“: cluster coefficient of grid?



Coefficient: 0

Cluster Coefficient

And here?



Coefficient: 3/7

(out of 28 pairs of neighbors, 12 are neighbors)

But how to route...?!

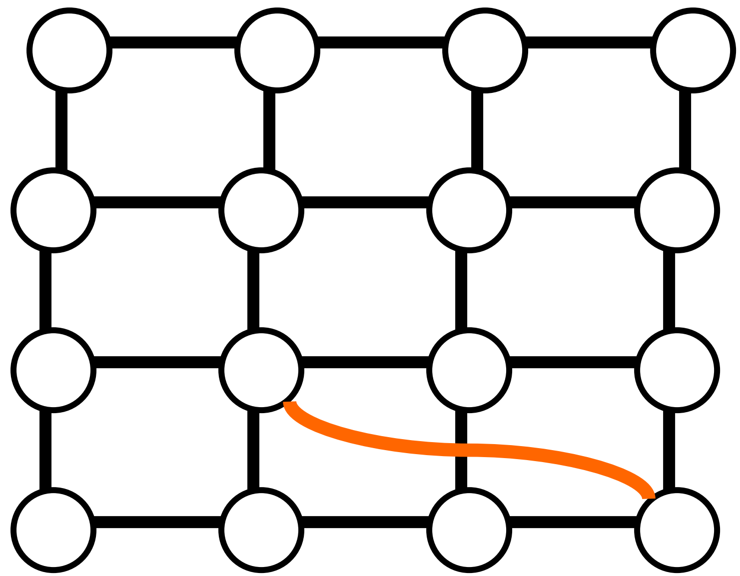
Having a graph model is good, but how did the mail arrive?!

Jon Kleinberg: most simple „navigable"social networks

(1) Nodes forward letter to neighbor, without really knowing whether neighbor is closer to destination!
(2) Only greedy routing can explain phenomenon...

Jon Kleinberg's graph model:
the augmented m x m grid

A simple grid with some additional random links per node with specific random distribution.



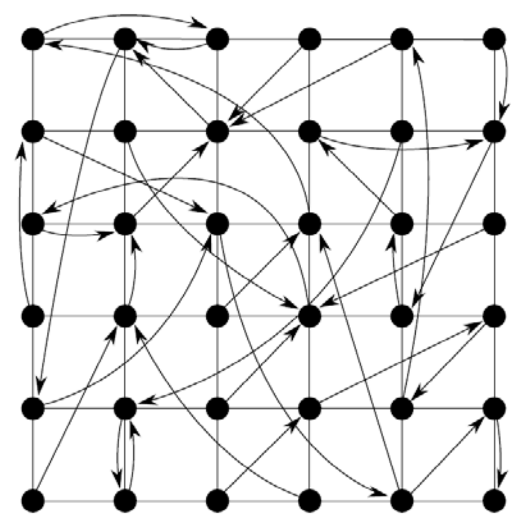
Kleinberg's Augmented Grid

Augmented Grid

Consider an $(m \times m)$ grid of $n=m^2$ nodes, where each node has a directed edge to each lattice neighbor (local contacts). In addition, each node has an additional random link (long-range contact).
For all u and v, the long-range contact of u to points to node v with probability $d(u,v)^{-\alpha}/\sum_{w \in V \setminus \{u\}} d(u,w)^{-\alpha}$, where d() is the distance in the grid and α is a parameter.

Geographic interpretation?
Interpretation of α?

Parameter α=0 means uniform at random (indep. of distance!); larger α make long-range links shorter.
One can show: if $\alpha \leq 2$, the diameter is polylogarithmic. (α=0 implies log diameter: proof as exercise?) But what about routing?



How to navigate??

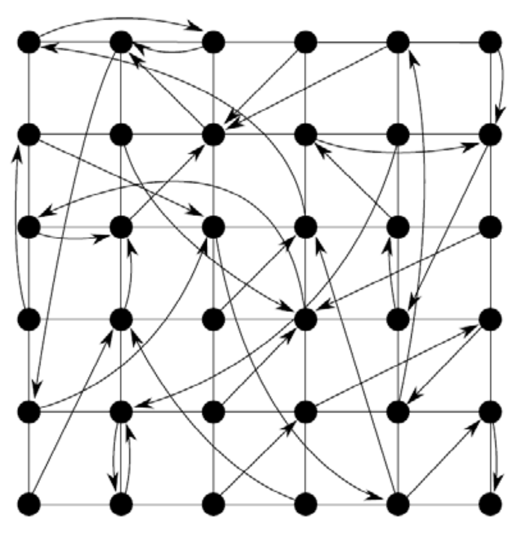
Navigability

Idea how to route on Kleinberg's graph?

Greedy Routing

while (not at destination):
    go to neighbor which is closest to destination
    (considering grid distance only)

Runtime?



Runtime of Greedy Routing

Greedy Routing

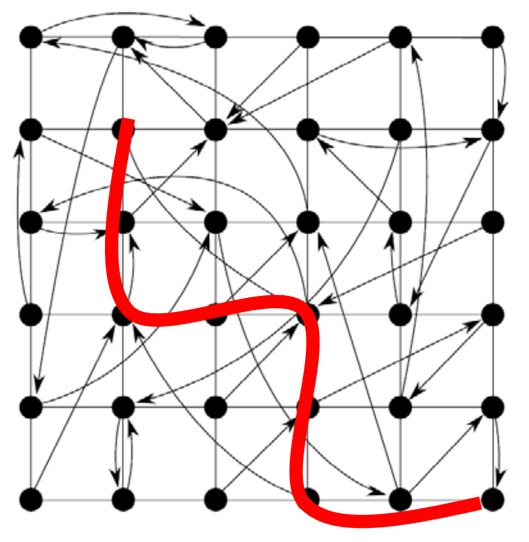
The greedy algorithm finds a routing path of length at most O(sqrt(n)).

Proof idea?

There are always neighbors closer to the destination. We can reduce the distance in at least one grid dimension...

QED

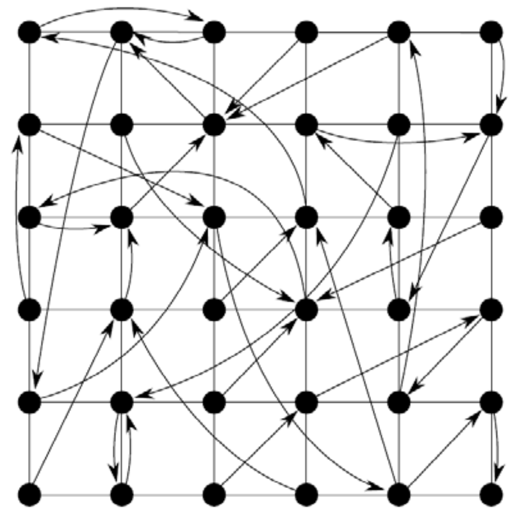
But Milgram promises more!
How do random links speed it up?



Random Links

A node u's random link leads to a node v with probability

$\Theta(1/(d(u,v)^{\alpha} m^{2-\alpha}))$	if $\alpha<2$
$\Theta(1/(d(u,v)^{2} \log n))$	if $\alpha=2$
$\Theta(1/d(u,v)^{\alpha})$	if $\alpha>2$



Proof?

since d(.,.) $\geq$ 1, and there are
at most O(r) new nodes at distance r

For $\alpha \neq 2$:
$$\sum_{w \in V \setminus \{u\}} \frac{1}{d(u,w)^{\alpha}} \in \sum_{r=1}^{m} \frac{\Theta(r)}{r^{\alpha}} = \Theta\left(\int_{r=1}^{m} \frac{1}{r^{\alpha-1}}\, dr\right) = \Theta\left(\left[\frac{r^{2-\alpha}}{2-\alpha}\right]_1^m\right)$$

For $\alpha<2$ this is $\Theta(m^{2-\alpha})$, for $\alpha>2$ this is $\Theta(1)$.

For $\alpha=2$ we have:
$$\sum_{w \in V \setminus \{u\}} \frac{1}{d(u,w)^{\alpha}} \in \sum_{r=1}^{m} \frac{\Theta(r)}{r^{2}} = \Theta(1) \cdot \sum_{r=1}^{m} \frac{1}{r} = \Theta(\log m) = \Theta(\log n)$$

QED

What does it mean?

Discussion

What does it mean?

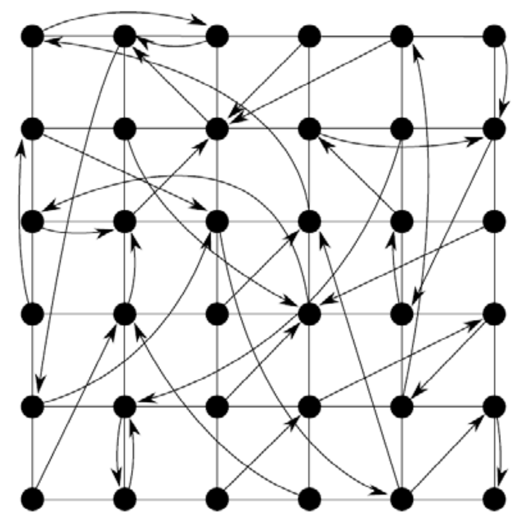
m=sqrt(n)...

$\Theta(1/(d(u,v)^{\alpha} m^{2-\alpha}))$ **if α<2** ☹: for small α, likely long paths...

$\Theta(1/(d(u,v)^{2} \log(n)))$ **if α=2** ☺?

$\Theta(1/d(u,v)^{\alpha})$ **if α>2** ☹: for large α, likely long paths...

exponential in α



α<2: Links have roughly same distance. We have to go far until finding a link that points close to destination („too random“)! Until then, we walk on grid...

α>2: Mostly short links only, and we have to go far until finding a link that reaches far („too focused“)! Until then, we don‘t make much progress wrt distance...

Navigability: α=2

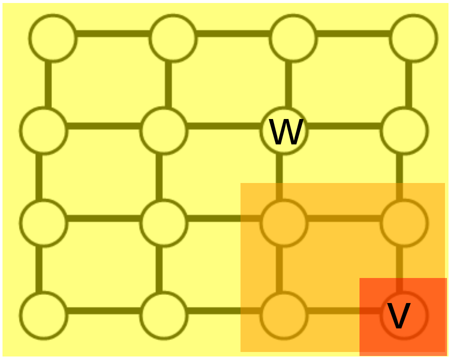
So let's study the case α=2!

Phase

Consider routing from node u to v, and assume we are at some intermediate node w. We say we are in phase j at node w if the lattice distance d(w,v) to the target v is between $2^j < d(w,v) \leq 2^{j+1}$.

So we count-down phases!
How many phases are there at most?

Logarithmic in max routing distance, i.e., O(log m)=O(log n).



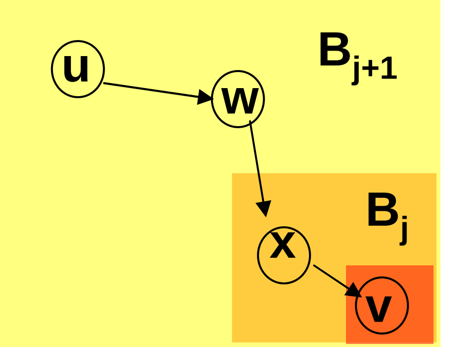
Navigability: α=2

Lemma

Assume we are at node w in phase j (on the way from u to v). The probability of getting to phase j-1 in one step is at least $\Omega(1/\log n)$.

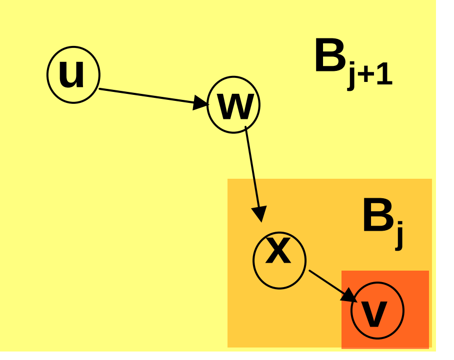
Proof idea?

Let $\mathbf{B_j}$ denote the nodes with $\mathbf{d(x,v) \leq 2^j}$. We get to phase j-1 from phase j if the long-range contact of w points to some node in B_j. Since we have not been at w before, its link points to a random node, independent of path to w.



Navigability: α=2

We know that for all nodes $x \in B_j$,
$d(w,x) \leq d(w,v)+d(v,x) \leq 2^{j+1}+2^j < 2^{j+2}$
(triangle inequality).
So for each $x \in B_j$, the probability that
w points to x is
$\Omega(1/(2^{2j+4}\log n))$.
Why?



See our lemma...

So probability to hit one of them?
Sum over all nodes that are in B_j...
And B_i (number of nodes at distance) grows quadratic, so at least $(2^j)/2$.
So:

$$\Omega\left(|B_j| \cdot \frac{1}{2^{2j+4}\log n}\right) = \Omega\left(\frac{2^{2j-1}}{2^{2j+4}\log n}\right) = \Omega\left(\frac{1}{\log n}\right)$$

QED

Navigability: α=2

Theorem

The expected path length is O(log^2 n).

Proof idea?

We have O(log n) phases. We proceed from one phase to the next with probability at least Ω(1/log n).
And linearity of expectation... ☺

QED

Yay, it‘s polylogarithmic!
So does it explain Milgram? ☺

Propagation Studies

Where to put ice cream stand?

In the middle...

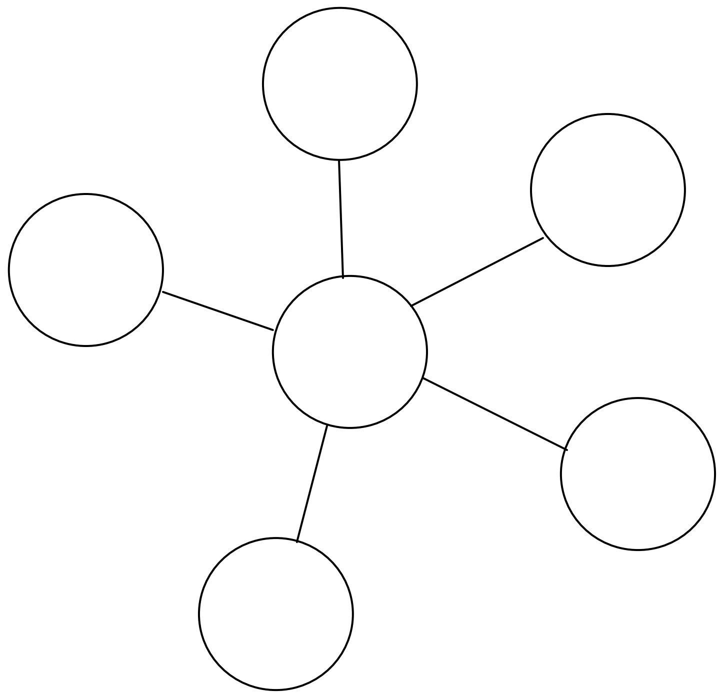
Where to put the second stand?

Right next to it?

Does the first player always have an advantage (cover more customers)?

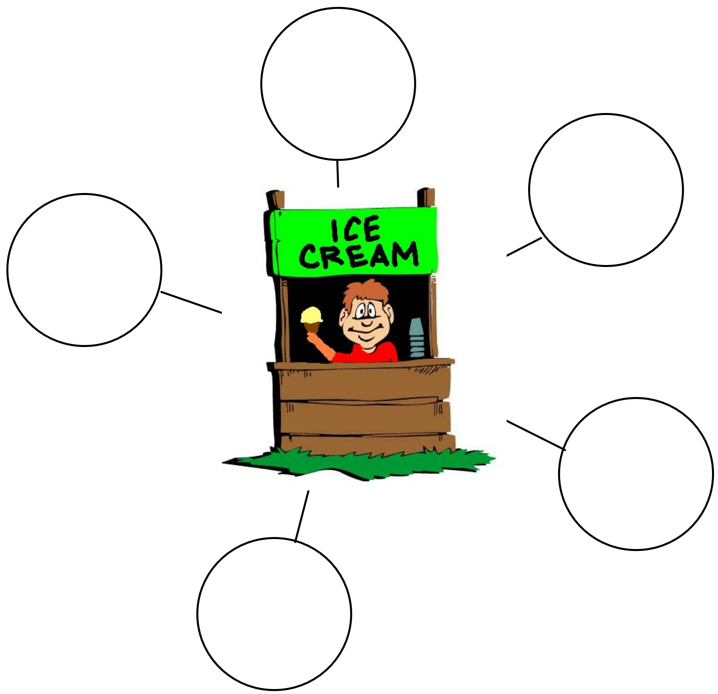
Propagation Studies

Here?



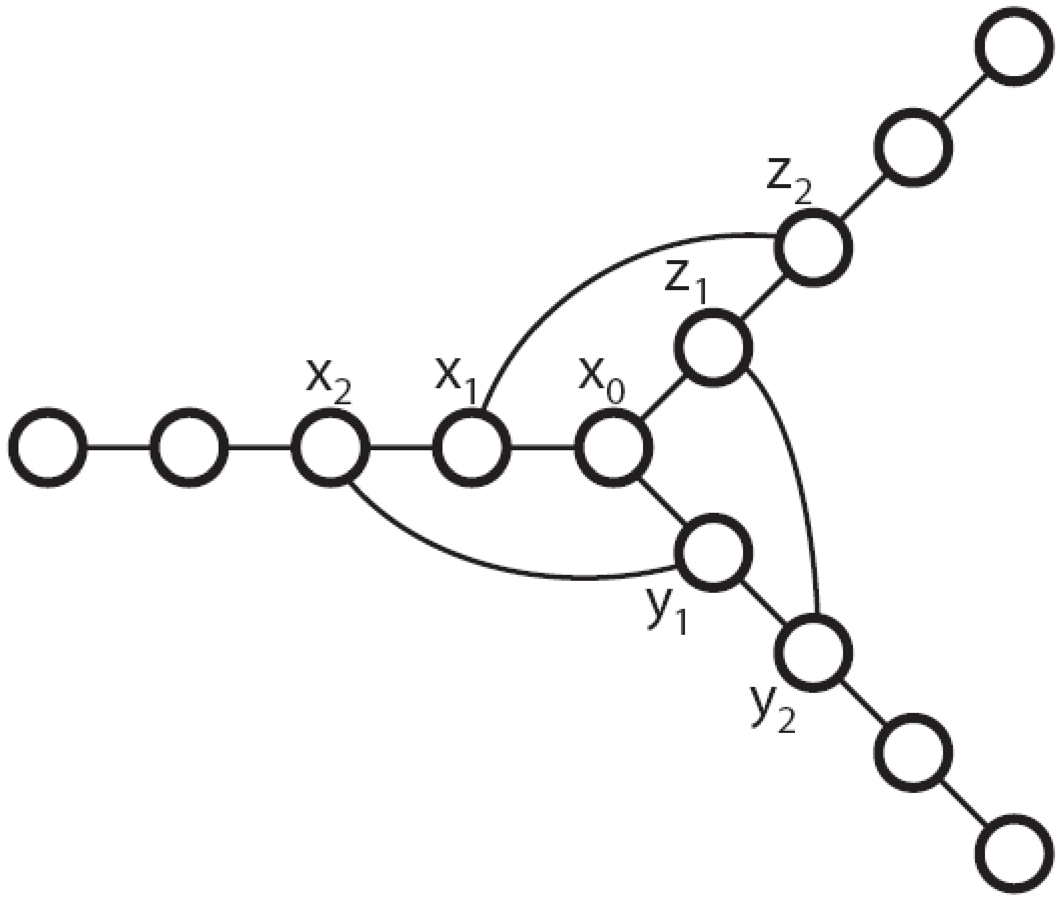
Always first mover advantage?

Here yes:



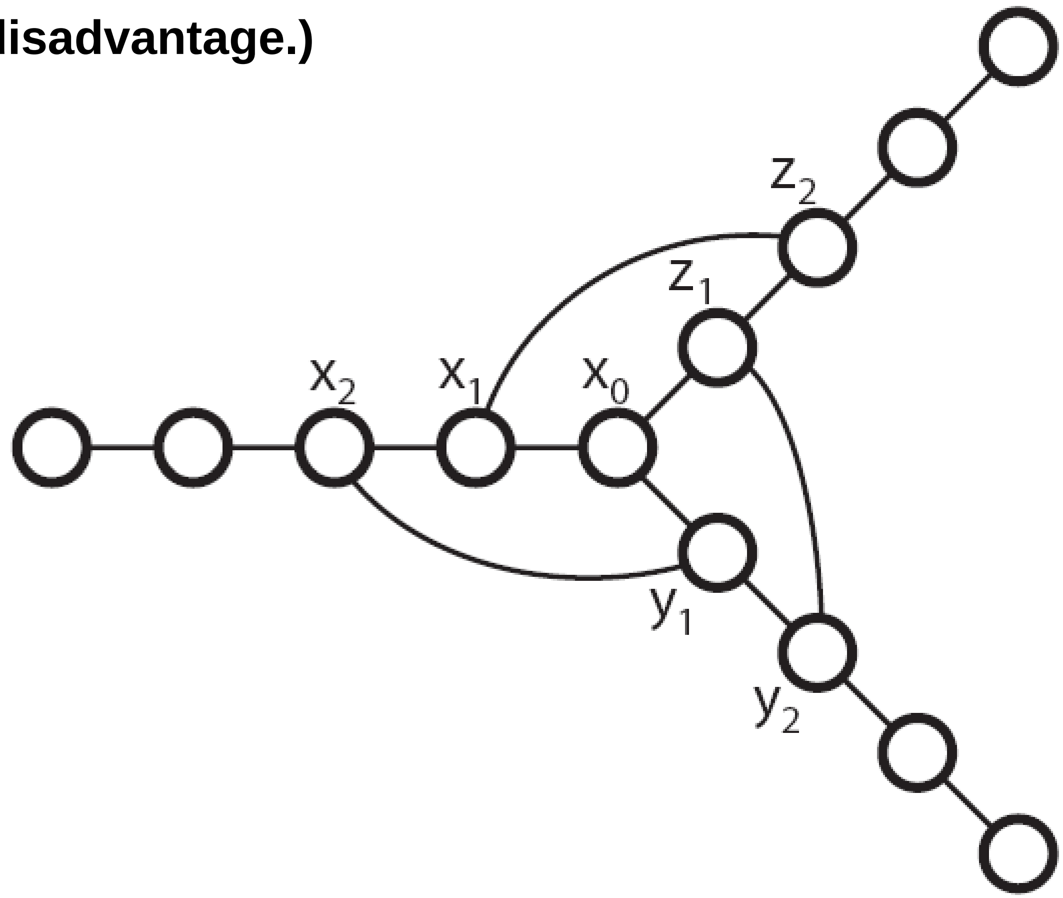
Always first mover advantage?

Here?



Always first mover advantage?

No advantage!
(But also no disadvantage.)



Literature for further reading:

- not Peleg‘s book ☺

End of lecture