CMS/CS/EE 144

Networks: Structure & Economics

Administrivia

1) QUIZ TODAY
2) HW3 is due Thursday
3) HW 2 solutions are up front
4) Lunch Bunch today – Vision & Deep Learning
5) Office hours Tue/Wed 7-10
6) Fill out the “time spent” polls…
7) Don’t forget about blog posts…

So far:

Four “universal” properties of networks

1) A “giant” connected component
2) Small diameter
3) Heavy-tailed degree distribution
4) High clustering coefficient

We’re trying to understand:

Why are these properties “universal”?

…the last two lectures:

Why is there a giant component?

Where do heavy-tails come from?

…this time:

Why is it a “small world”?

RECAP

We've seen 4 nearly "universal" properties of networks

Indep. random choices – G(n,p)

"rich get righer"

1) A "giant" connected component
2) Small diameter
3) Heavy-tailed degree distribution
4) High clustering coefficient

RECAP

We've seen 4 nearly "universal" properties of networks

Indep. random choices – G(n,p)

"rich get righer"

1) A "giant" connected component
2) Small diameter
3) Heavy-tailed degree distribution
4) High clustering coefficient

Alone these two properties aren't very surprising.

High clustering coefficient

Friends of my friends are likely to be my friends.

Small diameter

I have ~100 friends,
who each have ~100 friends,
and so on…

So, I can reach everyone in
s steps where $100^s = n$

$\longrightarrow$ $s = \log(n)$

Alone these two properties aren't very surprising.
Together, they are.

High clustering coefficient

Friends of my friends are likely to be my friends.

Small diameter

I have ~100 friends,
who each have ~100 friends,
and so on…

~~So, I can reach everyone in~~
~~s steps where $100^s = n$~~
~~→ $s = \log(n)$~~

Graphs that are highly clustered and still have small diameters are called:

SMALL WORLD GRAPHS

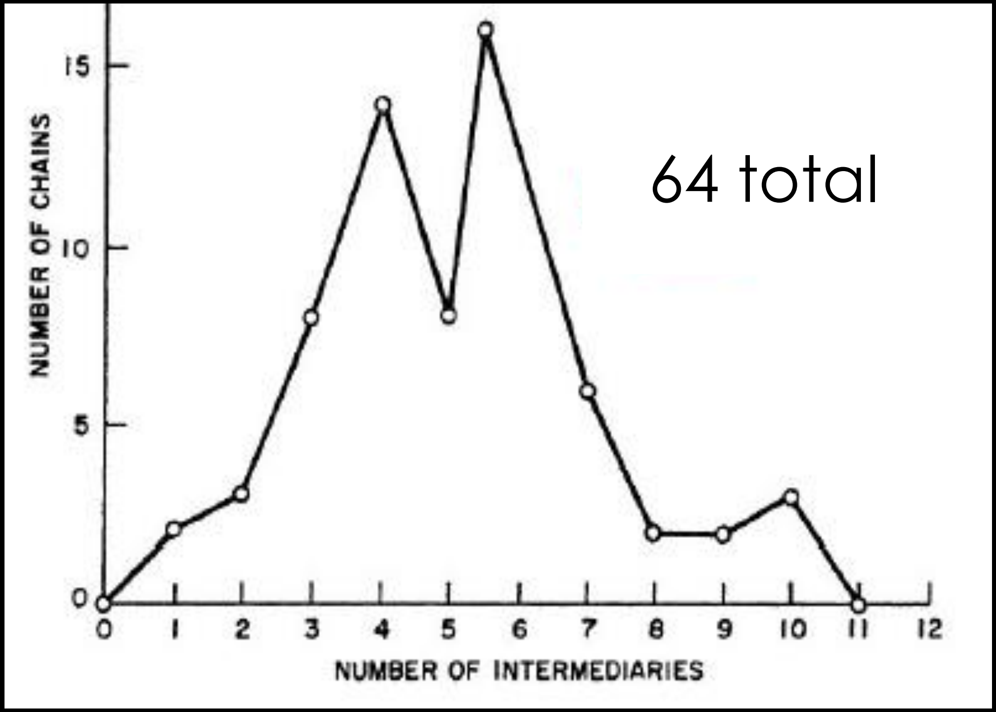
Stanley Milgram

1967

The first small-world experiment

Pick 300 people in midwest

Target → stockbroker in boston



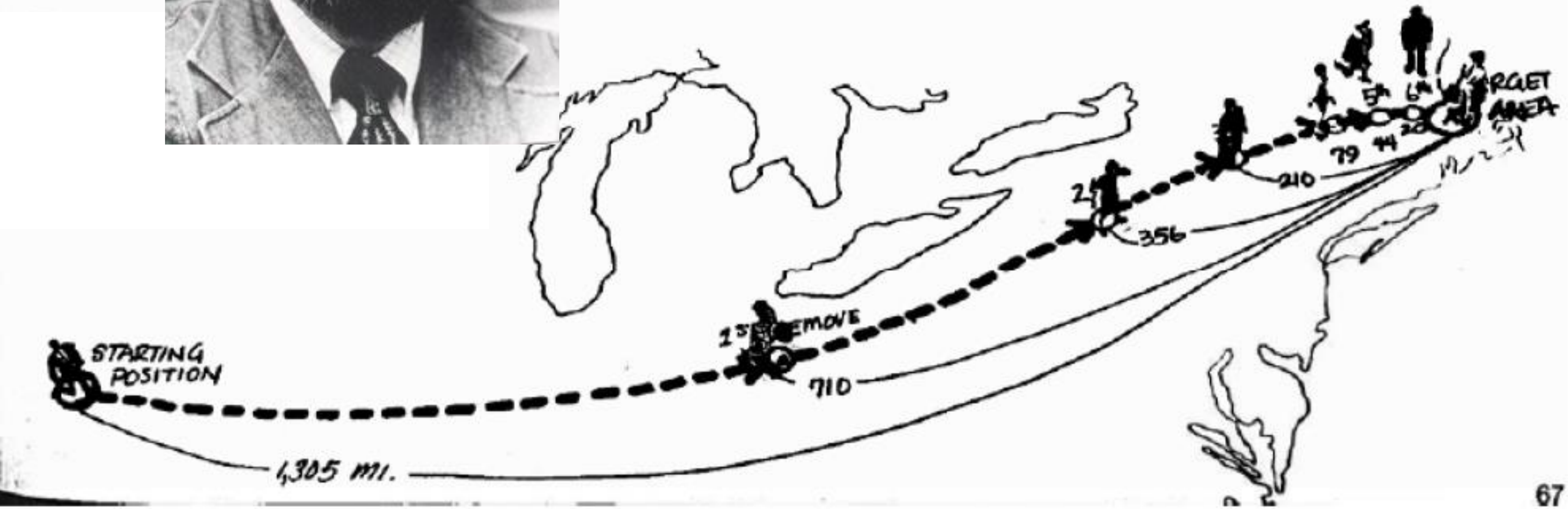
“six degrees of separation”

Stanley Milgram

The first small-world experiment

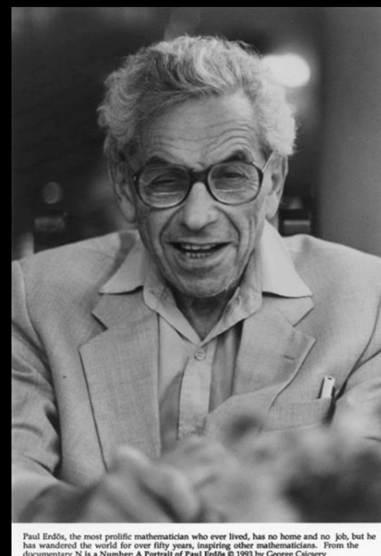
Pick 300 people in midwest

Target → stockbroker in boston



Small worlds in pop culture

Six degrees of
Kevin Bacon



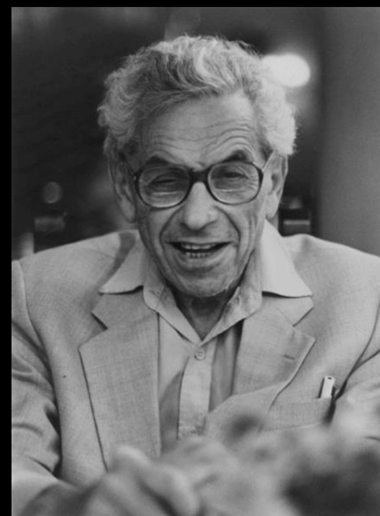
Erdos number

Will Smith movie

Erdos-Bacon numbers
Erdos # + Bacon #

Small worlds in pop culture

Six degrees of Kevin Bacon



Paul Erdös, the most prolific mathematician who ever lived, has no home and no job, but he has wandered the world for over fifty years, inspiring other mathematicians. From the documentary N is a Number: A Portrait of Paul Erdös © 1993 by George Csicsery

Erdos number

Erdos-Bacon numbers

Erdos # + Bacon #

- Erdos – 3 or 4
- Daniel Kleitman – 3
- Natalie Portman – 6
- Richard Feynman – 6
- Colin Firth – 7

Also see:

Erdos-Bacon-Sabbath numbers!

- Stephen Hawking – 8
- Daniel Levitin – 8
- Natalie Portman – 10
- Richard Feynman – 10
- Colin Firth – 11

Our small world experiment

→Wikipedia
→Citation network
→Product co-purchasing

Q: What were your path lengths?

	first,	shortest,	min
3D printing → Gwynne Shotwell:	5.3,	4,	3
Steven Low → Star Wars:	3.9,	3.5,	3
Adam → Feynman:	6.06,	6.04,	4
Adam's lecture → All Star:	19,	18,	3
Guardians of the Galaxy → Toothbrush:	15.6,	5.5,	6

Q: How did you find the paths?

A previous year's small world experiment

→Facebook

3 targets in 3 countries

~100 message chains
Only 5 reached their targets
Average path length 3.1

Q: Why was the success rate so small?

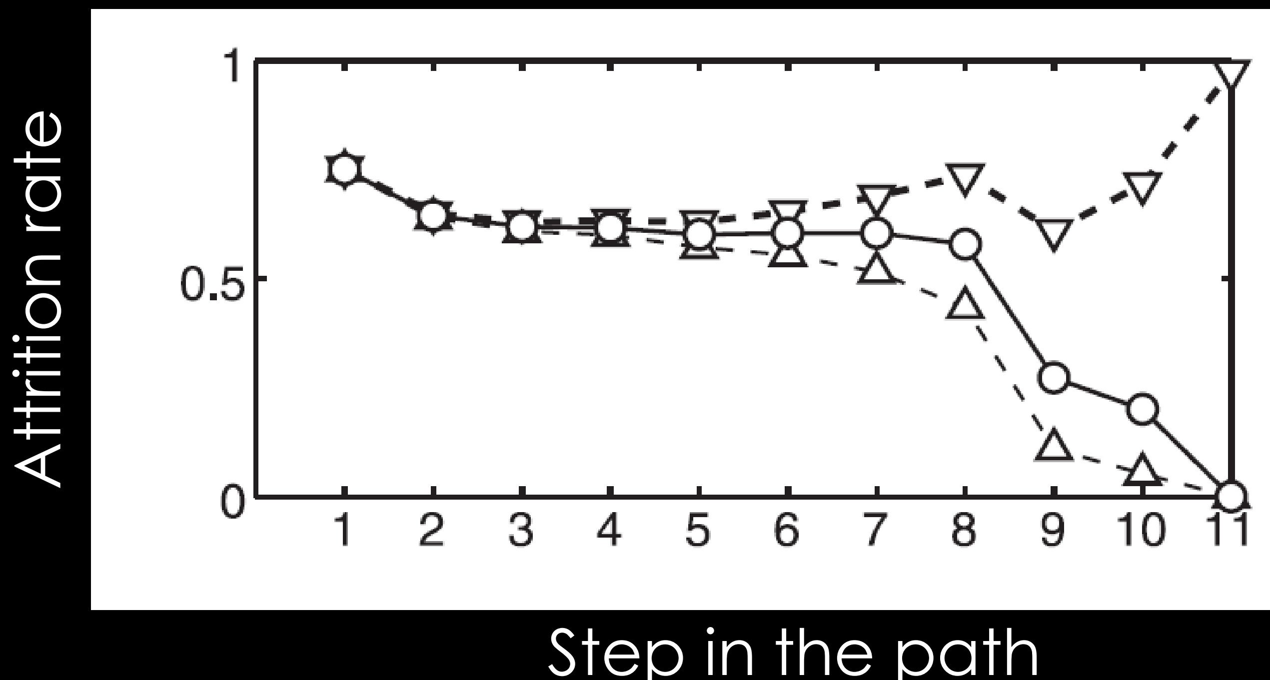
⟶ Attrition rate was very high

Large-scale small world experiments

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

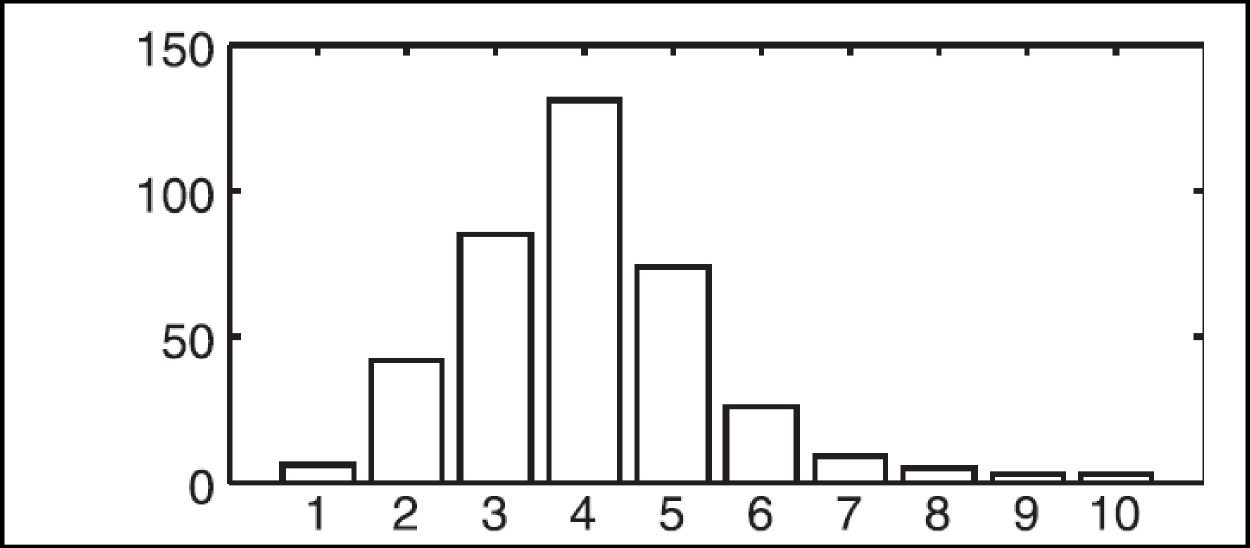
Over email – 18 targets in 13 countries

24,163 message chains
384 reached their targets (1.5%)
Average path length 4.0



Attrition rate
1
0.5
0
1
2
3
4
5
6
7
8
9
10
11
Step in the path

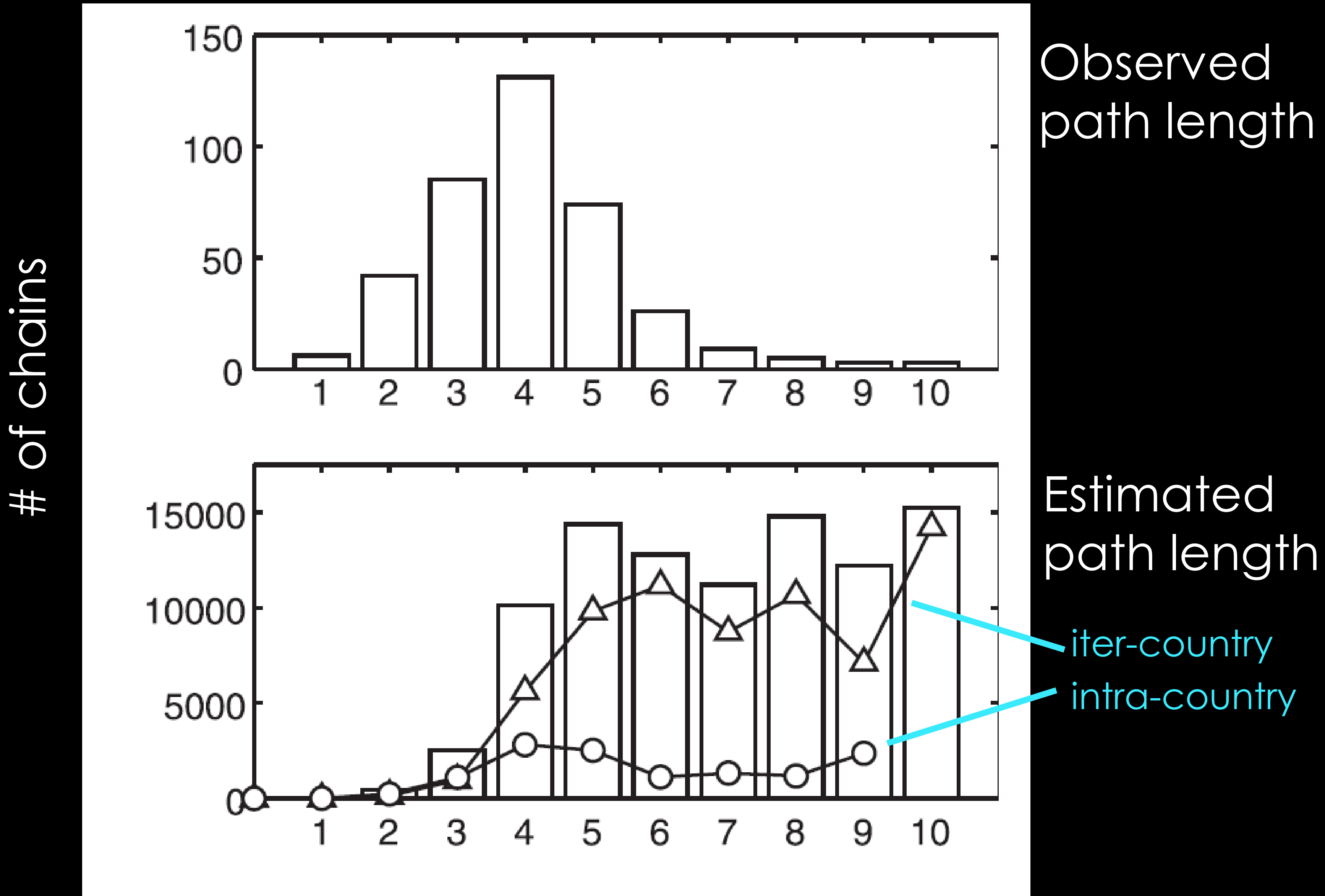
How do we account for attrition?



# of chains

Observed path length

How do we account for attrition?



The scientific questions:

1) Why do short paths exist?
2) How can people find short paths with so little information?

The engineering/business question:

How can we exploit short paths?

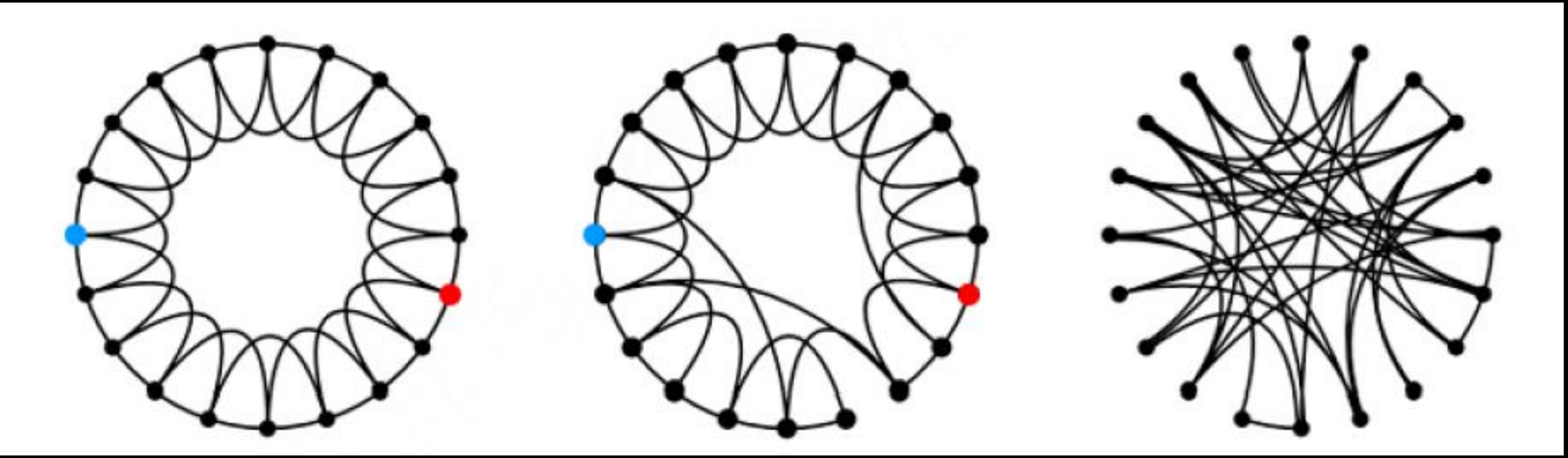
Q: What is a model that has high clustering and short paths?

A: One possible answer comes from Watts & Strogatz 1998

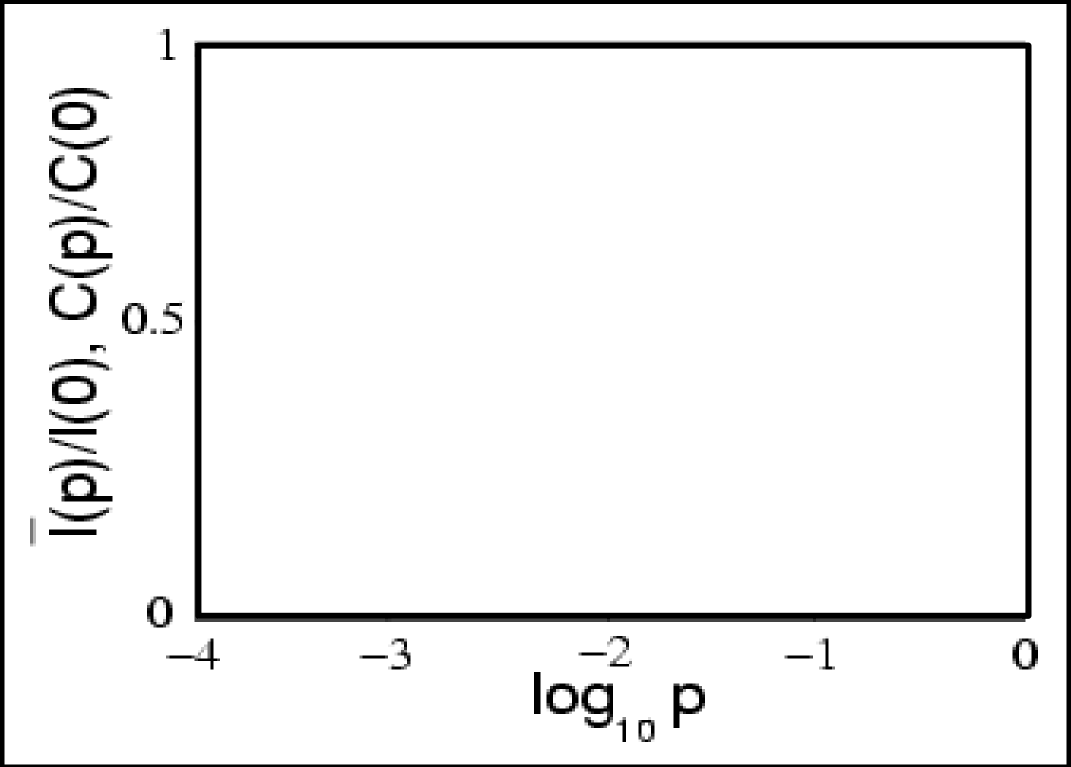
Arrange nodes in a d-dim lattice and

1) <u>Correlated local connections</u>
   add edges to all lattice points within distance 2
2) <u>Random long range connections</u>
   rewire local edges with probability p to connect to a random other node (uniform at random)

Example in 1-d

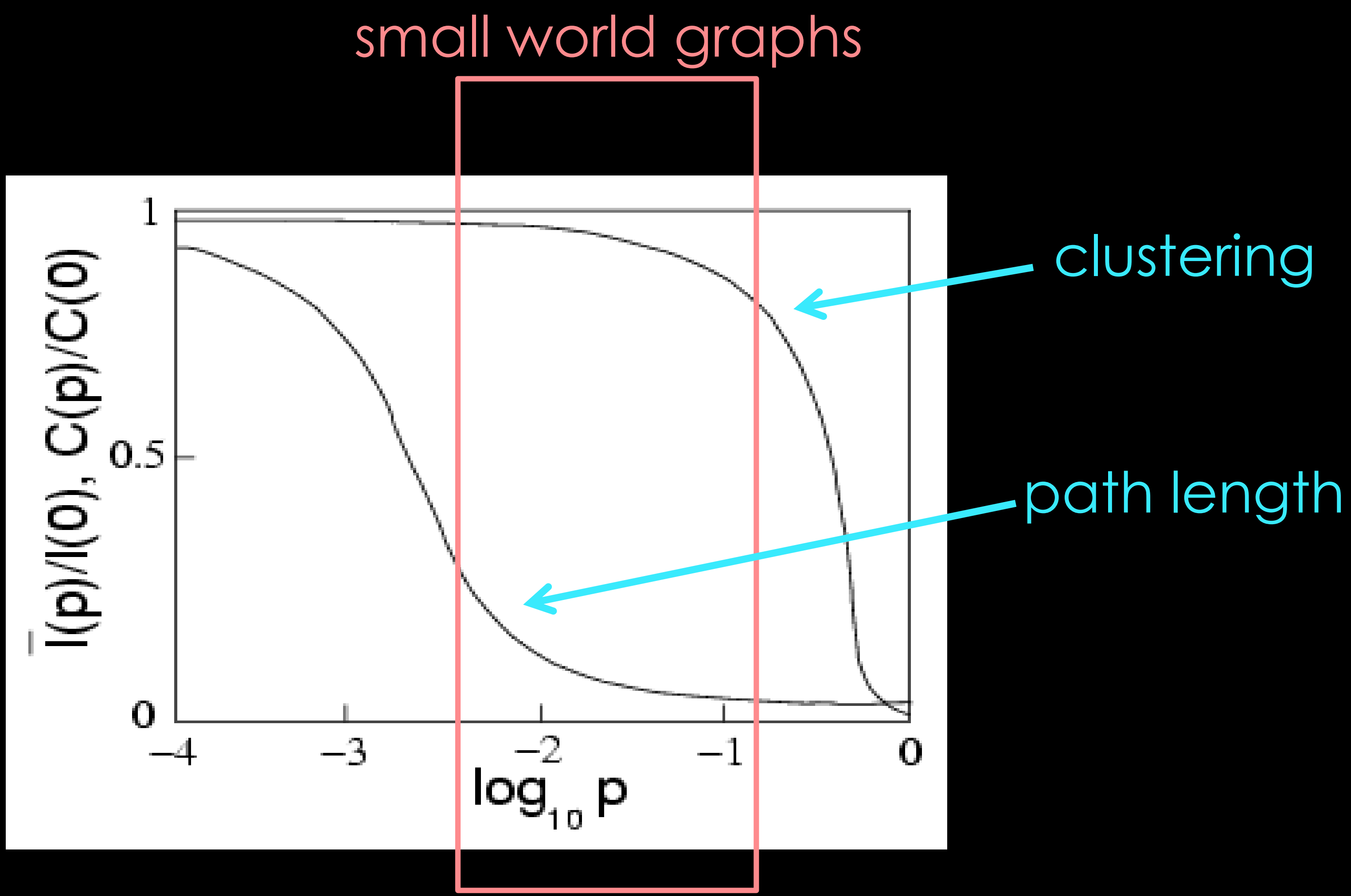


p=0 → p=1



C(p) = avg. clustering coefficient of with prob p

l(p) = avg. path length with prob p



C(p) = avg. clustering coefficient of with prob p

l(p) = avg. path length with prob p

So, one explanation for "small worlds" is to:

1) have correlated local connections
   to ensure high clustering, &
2) random long range connections
   to ensure small diameter

→ Can short paths be found in these networks?

Path lengths close to the diameter

Q: Can “myopic” “distributed” agents find “short” paths?

Agents send packets to neighbor that is closest to the target

Agents have no global information other than destination of packet

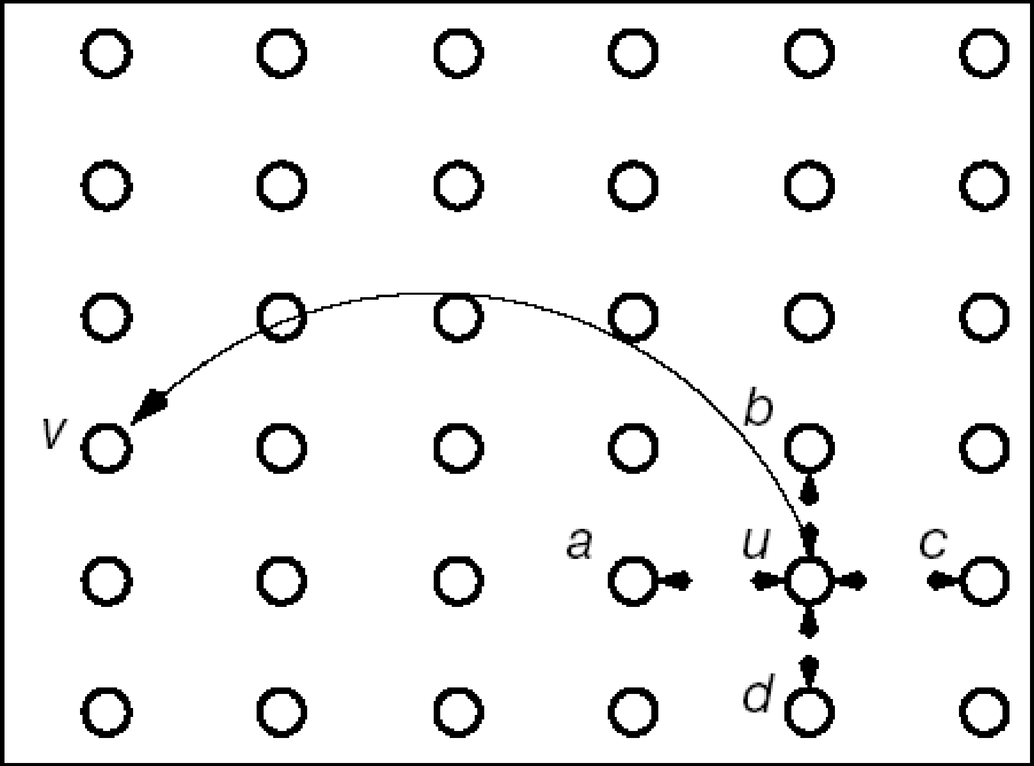
We’ll look at a slight variation of the model (to make analysis easier)

No triangles

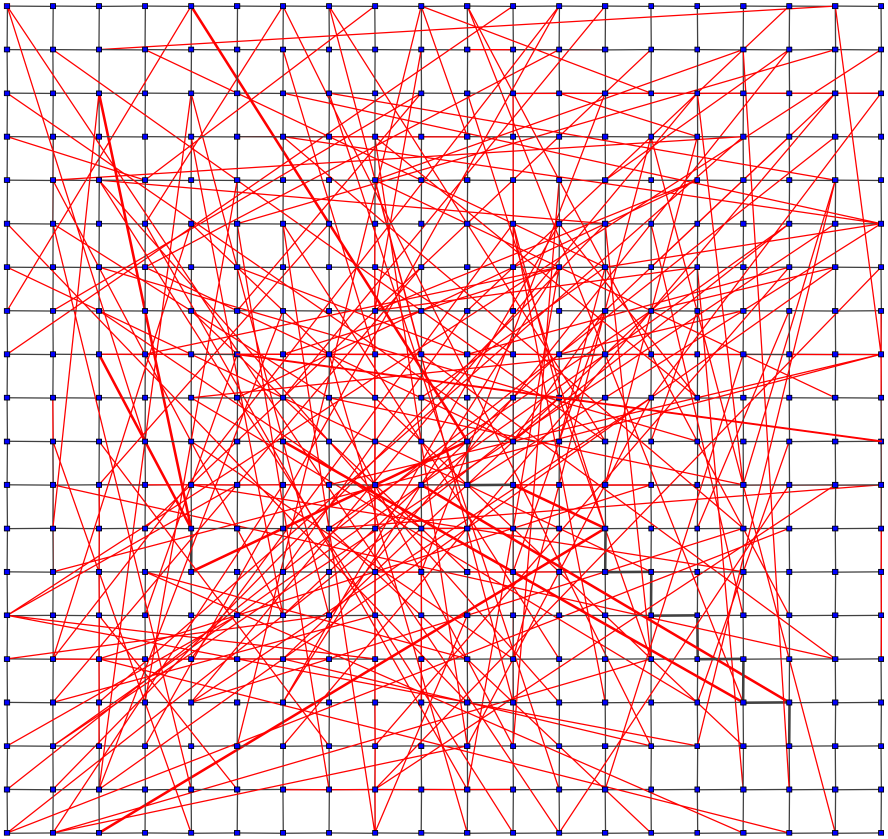
Kleinberg's variation (2001):

→ Local connections are only to distance 1 lattice points

→ Instead of rewiring local links to create long range links. Just add one long range Link for every node and choose it's endpoint randomly (uniformly for now)



An example:



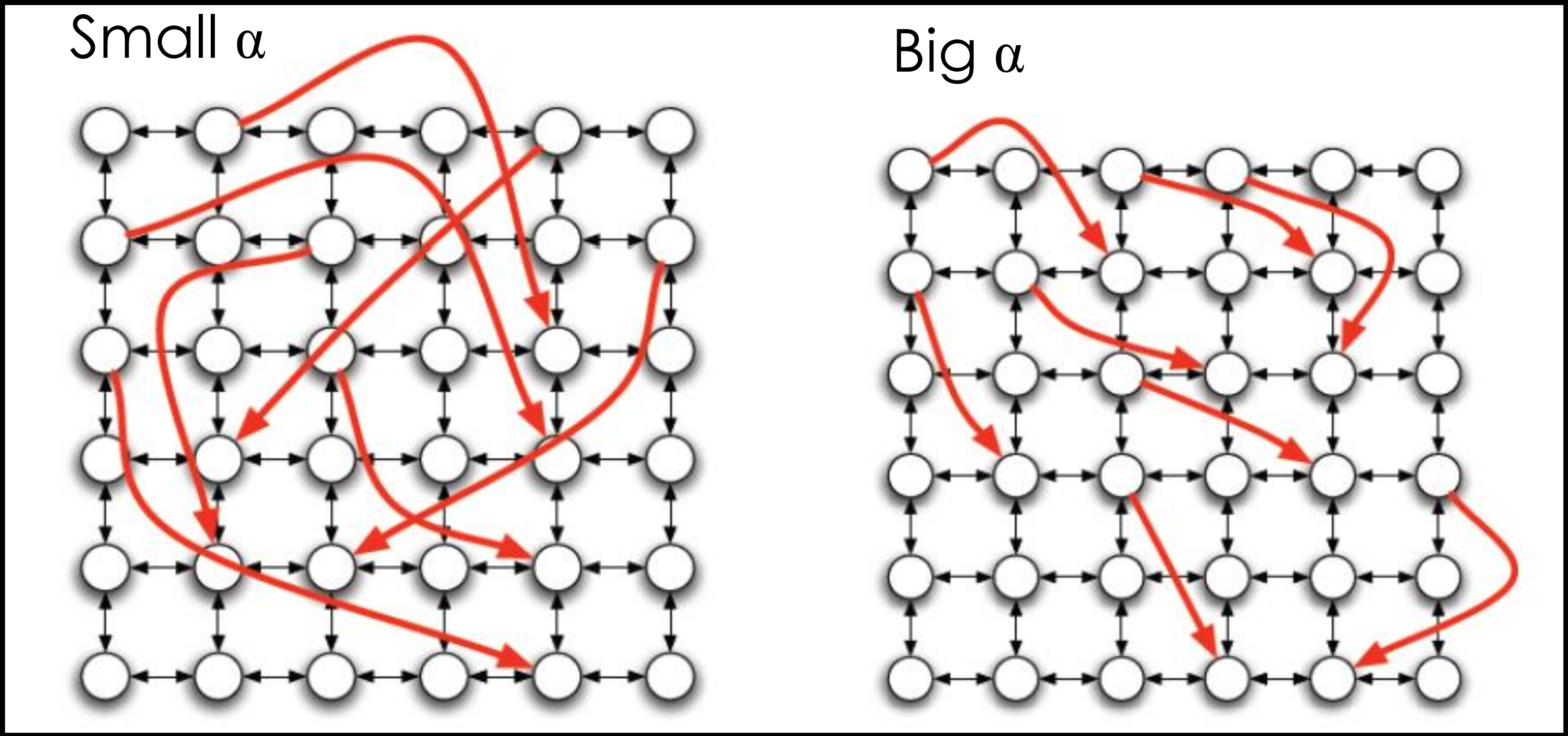
Path lengths close to the diameter (log(n))

Q: Can “myopic” “distributed” agents find “short” paths?

Agents send packets to neighbor that is closest to the target

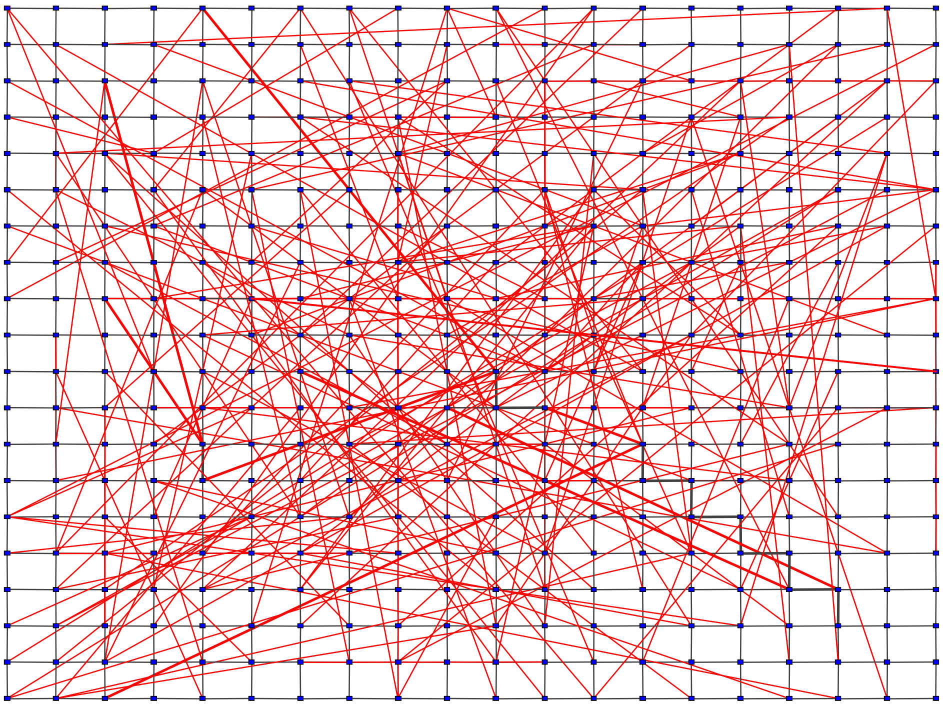
Agents have no global information other than destination of packet

Q: How can we "fix" the model?

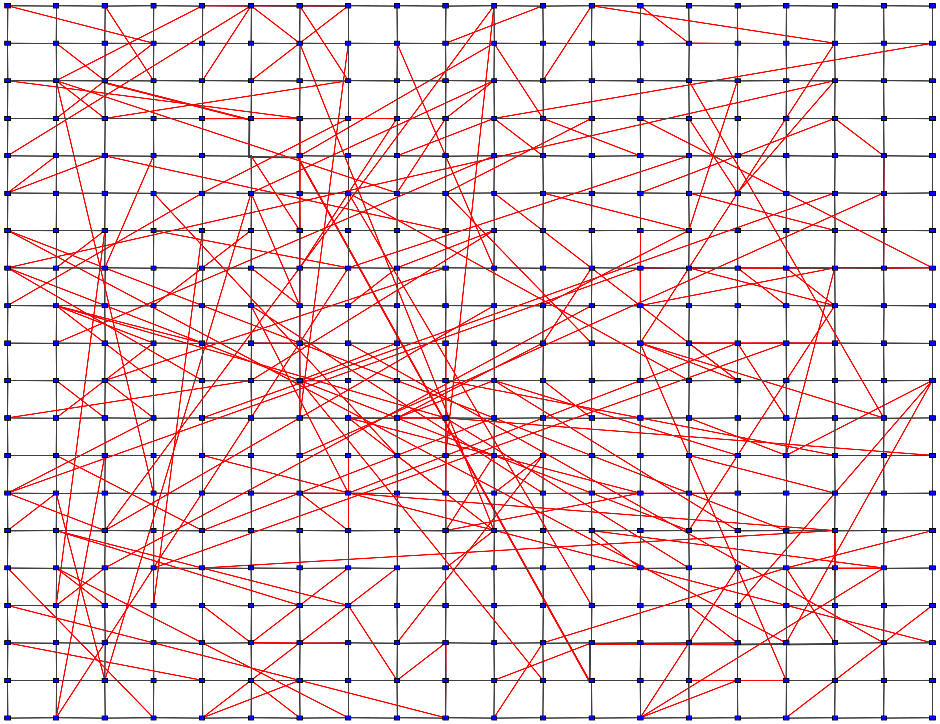


Small α
Big α

$p \sim p_0$



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



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