Small-Worlds & Co

Céline Comte, Fabien Mathieu

Master ACN
Paris, France

October 13, 2016

Roadmap

Introduction

Taxonomy

Properties and models

The Small-World Effect

Two strangers meet somewhere...

- As they talk they discover a mutual friend!
- What are the odds?
  - Small but memorable (cognitive bias)?
  - Greater than you think? Why?

The Small-World Effect: origins

Chain-Links, Frigyes Karinthy, 1929

- Technology turns the world into a global village

Planet Earth has never been as tiny as it is now. It shrunk – relatively speaking of course – due to the quickening pulse of both physical and verbal communication. This topic has come up before, but we had never framed it quite this way. We never talked about the fact that anyone on Earth, at my or anyone's will, can now learn in just a few minutes what I think or do, and what I want or what I would like to do. If I wanted to convince myself of the above fact: in couple of days I could be – Hocus pocus! – where I want to be.

The Small-World Effect: origins

Chain-Links, Frigyes Karinthy, 1929

- Technology turns the world into a global village
- First version of the «six degrees of separation»

One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth – anyone, anywhere at all. He bet us that, using no more than five individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances.

The Small-World Effect: origins

Chain-Links, Frigyes Karinthy, 1929

- Technology turns the world into a global village
- First version of the «six degrees of separation»
- How should it work?

Finding a Nobel Prize

1. I play tennis with Béla Kehrling who plays tennis with
2. King Gustav of Sweden who gave the Nobel Prize to
3. Selma Lagerlöf

The Small-World Effect: origins

Chain-Links, Frigyes Karinthy, 1929

- Technology turns the world into a global village
- First version of the «six degrees of separation»

Finding an anonymous riveter at the Ford Company

1. I have a friend called Árpád Páztor who is friend with
2. the director of Hearst publishing who plays tennis with
3. Henry Ford who is the boss of
4. the riveter's foreman who is the boss of
5. the riveter.

The Small-World Effect: scientists like that!

Contacts and Influence, Pool & Kochen, 1950's

- Probability that two strangers have a mutual friend?
- If none, how long would the "chain-link" be?
- Can we understand underlying causes?

The Small-World Effect: experiments

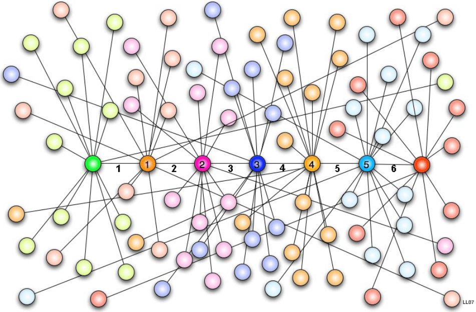
Milgram experiment: an interlude

- Stanley Milgram (1933–1984): social psychologist
- Known for experiments on obedience (1960–1963)
  - 60% of "success" (average
  - Up to 92.5% (37/40) for some settings
  - I as in Icarus (French Movie with Yves Montand) (`http://www.youtube.com/watch?v=EAqJXF32470`)

The Small-World Effect: experiments

Stanley Milgram's 1967 experiment

- Goal: demonstrate «six degrees»
- Protocol:
  - Transmit a letter (name, profession, town in the US)
  - Only transmit to people known on a first-name basis



The Small-World Effect: experiments

Mixed results

First experiment: 60 people (50)

- One chain with 3 hops
- Average length of 8 hops

The Small-World Effect: experiments

Mixed results

First experiment: 60 people (50)

- One chain with 3 hops
- Average length of 8 hops
- 3 successes only (5%)

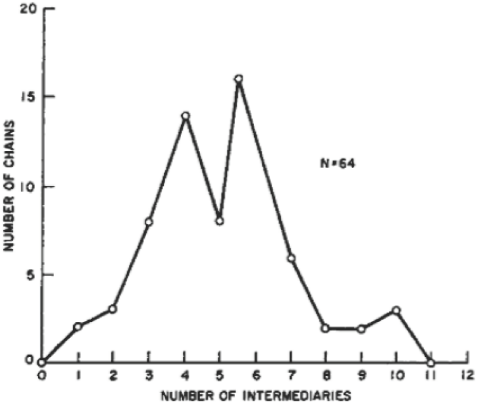
The Small-World Effect: experiments

Mixed results

- Experiments #2 and #3: failures (unpublished)
- Experiment #4: 44/160
- Goes better and better (*An Experimental Study of the Small World Problem, 1969*).

The Small-World Effect: experiments

Mixed results



The Small-World Effect: celebrity

Six Degrees of Separation, John Guare, 1990

I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation between us and everyone else on this planet. The President of the United States, a gondolier in Venice, just fill in the names. I find it

- *extremely comforting that we're so close*

The Small-World Effect: celebrity

Six Degrees of Separation, John Guare, 1990

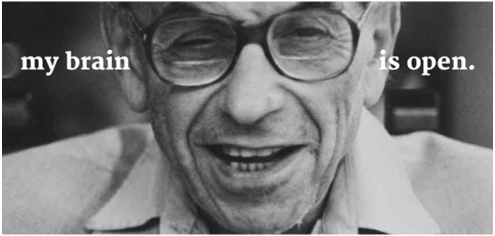
I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation between us and everyone else on this planet. The President of the United States, a gondolier in Venice, just fill in the names. I find it

- *extremely comforting that we're so close*
- *like Chinese water torture that we're so close because **you have to find** the right six people to make the right connection*

The Small-World Effect: games

The Erdös number

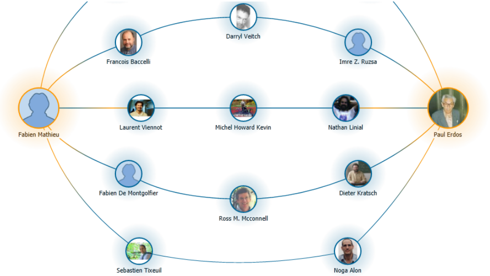
- Paul Erdös (1913–1996), influential, multidisciplinary, prolific mathematician: $>$1500 papers, $>$500 papers
- "And what is your Erdös number?", Casper Goffman, 1969
- Rule: co-authorship



The Small-World Effect: games

The Erdös number

- Paul Erdös (1913–1996), influential, multidisciplinary, prolific mathematician: >1500 papers, >500 papers
- "And what is your Erdös number?", Casper Goffman, 1969
- Rule: co-authorship
- Mine is 3



The Small-World Effect: games

Six Degrees of Kevin Bacon

- Kevin Bacon (1958?-), plays in lots of movies (79 credits)
- “Kevin Bacon is the Center of the Universe”, 1994
- Rule: play in the same movie

The Small-World Effect: games

Six Degrees of Kevin Bacon

- Kevin Bacon (1958?-), plays in lots of movies (79 credits)
- "Kevin Bacon is the Center of the Universe", 1994
- Rule: play in the same movie

Jean Lefebvre

Kevin Bacon

The Small-World Effect: games

Six Degrees of Kevin Bacon

- Kevin Bacon (1958?-), plays in lots of movies (79 credits)
- "Kevin Bacon is the Center of the Universe", 1994
- Rule: play in the same movie

Jean Lefebvre

Le bourgeois gentil mec (1969)

Paul Demange

Kevin Bacon

The Small-World Effect: games

Six Degrees of Kevin Bacon

- Kevin Bacon (1958?-), plays in lots of movies (79 credits)
- "Kevin Bacon is the Center of the Universe", 1994
- Rule: play in the same movie

Jean Lefebvre

Le bourgeois gentil mec (1969)

Paul Demange

Il giudizio universale (1961)

Vittorio Gassman

Kevin Bacon

The Small-World Effect: games

Six Degrees of Kevin Bacon

- Kevin Bacon (1958?-), plays in lots of movies (79 credits)
- “Kevin Bacon is the Center of the Universe”, 1994
- Rule: play in the same movie

Jean Lefebvre

Le bourgeois gentil mec (1969)

Paul Demange

Il giudizio universale (1961)

Vittorio Gassman

Sleepers (1996)

Kevin Bacon

The Small-World Effect: games

The Erdös-Bacon-Sabbath game

- Rule: find People with low (finite) sum of
  - Erdös number
  - Bacon number
  - Black Sabbath number (musical collaborations)

The Small-World Effect: games

The Erdös-Bacon-Sabbath game

- Rule: find People with low (finite) sum of
  - Erdös number
  - Bacon number
  - Black Sabbath number (musical collaborations)

- Surprising results
  - Albert Einstein (2+4+5)
  - Condoleeza Rice (6+3+4)
  - Natalie Portman (5+2+3)
  - Stephen Hawking (4+2+2)
  - Terry Pratchett (4+2+3)

The Small-World Effect: scientists love that!

End of 90's, Watts & Strogatz revived small-world studies

- Social networks easier to obtain
  - IMDB (thanks to Kevin)
  - Co-autorship (thanks to Paul)
  - ...
- Curiousity and Crickets
- Computer Science / Maths approach
- *Collective dynamics of small-world networks*, Nature, 1998

We hope that our work will stimulate further studies of small-world networks (...). Although small-world architecture has not received much attention, we suggest that it will probably turn out to be widespread in biological, social and made-made systems, often with important dynamical consequences.

The Small-World Effect: today

- Acces to real, large, social networks
  - Facebook
  - Twitter
  - ...
- Big economic stakes
- Very active research field for more than 15 ans

Three worlds that are very close

Social Networks

Graphs representing social interactions (local choices)

Three worlds that are very close

Social Networks

Graphs representing social interactions (local choices)

Real-World graphs

Graphs from real networks (biology, infrastructures, CS)

Three worlds that are very close

Social Networks

Graphs representing social interactions (local choices)

Real-World graphs

Graphs from real networks (biology, infrastructures, CS)

Small-Worlds

Graphs with properties that have been observed in most social networks (and in some real-world graphs)

Three worlds that are very close

Social Networks

Graphs representing social interactions (local choices)

Real-World graphs

Graphs from real networks (biology, infrastructures, CS)

Small-Worlds

Graphs with properties that have been observed in most social networks (and in some real-world graphs)

All graphs are not small-worlds

Three worlds that are very close

Social Networks

Graphs representing social interactions (local choices)

Real-World graphs

Graphs from real networks (biology, infrastructures, CS)

Small-Worlds

Graphs with properties that have been observed in most social networks (and in some real-world graphs)

All graphs are not small-worlds

- Family tree, organization chart
- New-York streets

Examples of Small-Worlds

Social Networks

- Facebook
- Twitter
- Specific communities
  - actors
  - researchers
  - musicians
- Web (to some extend)

Examples of Small-Worlds

Real-World graphs

- C. Elegans “brain” (302 neurons)
- Electrical power grid
- Airport network
- Distributed Hash Tables

Examples of Small-Worlds

Artificial models

- Albert & Barabasi's preferential attachment
- Watts & Strogatz' ring
- Kleinberg's grid
- etc etc

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- Short distances
- Scale-free
- High Clustering
- Navigable

Properties of Small-Worlds

Six properties characterize small-worlds:

- **Large number of nodes**
- Low density
- Short distances
- Scale-free
- High Clustering
- Navigable

Small-Worlds are **not** small

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- **Low density**
- Short distances
- Scale-free
- High Clustering
- Navigable

Average degree is low, $O(1)$ or $O(\log(n))$

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- **Short distances**
- Scale-free
- High Clustering
- Navigable

Distance between nodes is low: $O(\log n)$

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- Short distances
- **Scale-free**
- High Clustering
- Navigable

- Degree distribution is Heavy-Tail: a few nodes are far more connected than others
- Example: Power Law ($\approx K\frac{1}{d^{\alpha}}$ nodes of degree d)

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- Short distances
- Scale-free
- **High Clustering**
- Navigable

- Clustering coefficient: probability of triangles in the graph
- High: compared to equivalent graph with totally random edges
- Indicator of the "mutual friend" paradox

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- Short distances
- Scale-free
- High Clustering
- **Navigable**

- **you have to find** the right six people
- Navigability: nodes successfully find short paths
- Explains the Milgram experiment

Properties of Small-Worlds

Six properties characterize small-worlds:

- Large number of nodes
- Low density
- Short distances
- Scale-free
- High Clustering
- Navigable

Not all are required to qualify as a small-world:

- Minimum : size, density, distances, and one more
- Most social networks have the first 5

Intuition: properties come from the way interactions are made.

Properties of Small-Worlds

Remark: a large random graph (Erdös-Rényi) with a logarithmic degree is not a Small-World

- will have
  - Large size
  - Low degree
  - Short distances
- will fail
  - Clustering coefficient
  - Scale-free
  - Navigability

Low density: intuition

- n is big

Low density: intuition

- n is big
- Connexions are local

Low density: intuition

- n is big
- Connexions are local
- Each node has a limited capacity

Low density: intuition

- n is big
- Connexions are local
- Each node has a limited capacity
- Dunbar's number (148)

Short distances: intuition

Short distances: intuition

- Chaos (Erdös-Rényi)

Short distances: intuition

- Chaos (Erdös-Rényi)
- Order (Trees)

Short distances: intuition

- Chaos (Erdös-Rényi)
- Order (Trees)
- Concentration (Hubs born from scale-free distribution)

Heavy tail distribution: definition

Degree distribution

Gives for each degree d the proportion $P(d)$ of nodes (probability) that have degree d.

Heavy tail distribution: definition

Degree distribution

Gives for each degree d the proportion $P(d)$ of nodes (probability) that have degree d.

Remark

$D = \sum_{d \geq 1} dP(d)$

Heavy tail distribution: definition

Degree distribution

Gives for each degree d the proportion $P(d)$ of nodes (probability) that have degree d.

Remark

$D = \sum_{d \geq 1} dP(d)$

Heavy tail distribution

- There exist nodes with a very high degree compared to D.
- Not verified by usual random distributions: binomial, Poisson, geometric, ...
- Verified by power law distribution : $P(d) = \frac{K}{d^{\alpha}}$
- Warning, it is very easy to tell stupid things when dealing with heavy tail distributions.

Heavy tail distribution: intuition

Algorithm (preferential attachment)

- Start with two connected nodes (each has degree 1)

Heavy tail distribution: intuition

Algorithm (preferential attachment)

- Start with two connected nodes (each has degree 1)
- Add one (disconnected) new node

Heavy tail distribution: intuition

Algorithm (preferential attachment)

- Start with two connected nodes (each has degree 1)
- Add one (disconnected) new node
- Choose an existing node at random <u>according to the degree distribution</u>

Heavy tail distribution: intuition

Algorithm (preferential attachment)

- Start with two connected nodes (each has degree 1)
- Add one (disconnected) new node
- Choose an existing node at random <u>according to the degree distribution</u>
- Connect the two nodes

Heavy tail distribution: intuition

Algorithm (preferential attachment)

- Start with two connected nodes (each has degree 1)
- Add one (disconnected) new node
- Choose an existing node at random according to the degree distribution
- Connect the two nodes
- Iterate...

Heavy tail distribution: intuition

Preferential attachment

- Was introduced by Albert and Barabasi to model the growth of a network.
- Motivation: degree $\approx$ popularity

Heavy tail distribution: intuition

Preferential attachment

- Was introduced by Albert and Barabasi to model the growth of a network.
- Motivation: degree $\approx$ popularity

Properties of the resulting graph:

- Average popularity stays bounded: $D \to 2$.
- The popularity distribution converges to a power law.

Heavy tail distribution: intuition

Preferential attachment

- Was introduced by Albert and Barabasi to model the growth of a network.
- Motivation: degree $\approx$ popularity

Properties of the resulting graph:

- Average popularity stays bounded: $D \to 2$.
- The popularity distribution converges to a power law.

Interpretation

Models a snowball effect: *the rich get richer*.

Heavy-tail distribution: intuition

Preferential attachment

- Models the growth of a network
- Was introduced by Albert and Barabasi to explain Web degree distribution
- Recipe: nodes are more attracted to nodes with high degree
- Produces Heavy-tailed distributions

Similar SnowBall effects

- Wealth (rich people get richer)
- City sizes
- Forest fires areas

Clustering: intuition

Several probable causes for clustering have been proposed

Clustering: intuition

Several probable causes for clustering have been proposed

Cloning
: Rookies try to make one friend and join his/her community

Clustering: intuition

Several probable causes for clustering have been proposed

- **Cloning** Rookies try to make one friend and join his/her community
- **LinkedIn** Starting from existing network, reach friends of friends

Clustering: intuition

Several probable causes for clustering have been proposed

- **Cloning** Rookies try to make one friend and join his/her community
- **LinkedIn** Starting from existing network, reach friends of friends
- **Space** Close (spatially) people tend to be close (socially)

Clustering: intuition

Several probable causes for clustering have been proposed

- **Cloning** Rookies try to make one friend and join his/her community
- **LinkedIn** Starting from existing network, reach friends of friends
- **Space** Close (spatially) people tend to be close (socially)
- **Communities** Edges correspond to small communities with high density

Clustering: intuition

Several probable causes for clustering have been proposed

Cloning Rookies try to make one friend and join his/her community

LinkedIn Starting from existing network, reach friends of friends

Space Close (spatially) people tend to be close (socially)

Communities Edges correspond to small communities with high density

- Example: IMDB, each movie creates a small complete graph

Clustering: Watts & Strogatz ring

Goal

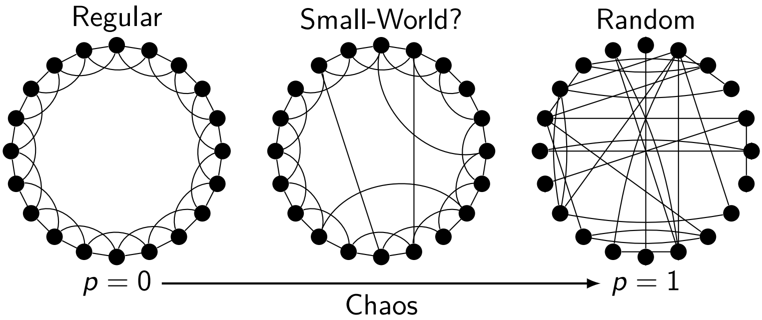
- Get a model for artificial random small-worlds
- Focus on clustering and short distances

Assumptions

- Clustering comes from order (space)
- Short distances come from chaos (short-cuts)
- Can we find a trade-off?

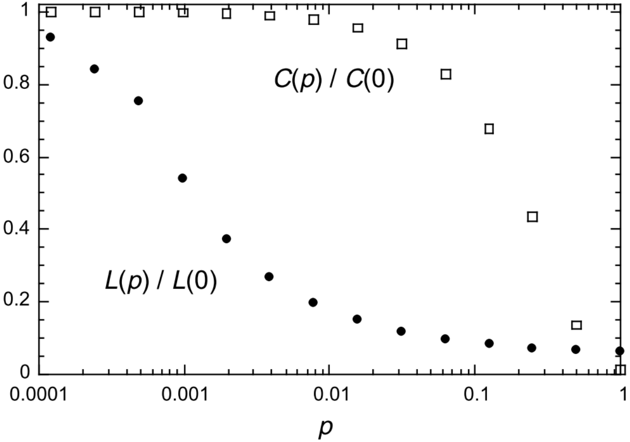
Clustering: Watts & Strogatz ring

- A large ring of nodes
- Each node is connected to its k closest neighbors (4–10)
- Each edge is randomly rewired with probability p
- $p = 0$: totally structured graph (high clustering, high distances)
- $p = 1$: totally random graph



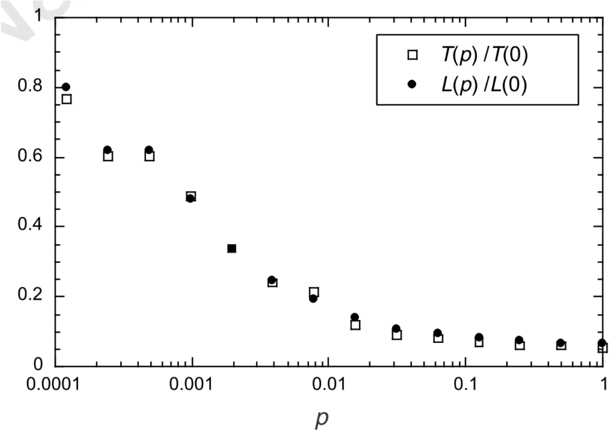
Clustering: Watts & Strogatz ring

- Diameter collapses fast
- Clustering lasts longer
- Small $p > 0$: Small-World!



Clustering: Watts & Strogatz ring

Epidemic susceptibility close to diameter (practical if time)



Clustering: Watts & Strogatz ring

- A small article (2.5 pages) in Nature
- 15 years of research in maths, CS, social, eco. . .

We hope that our work will stimulate further studies of small-world networks (. . .). Although small-world architecture has not received much attention, we suggest that it will probably turn out to be widespread in biological, social and made-made systems, often with important dynamical consequences.

Navigability: intuition

Navigability: intuition

Hubs All roads lead to Rome, so aim at Kevin Bacon

Navigability: intuition

Hubs	All roads lead to Rome, so aim at Kevin Bacon
Hierarchy	Follow the chain of command

Navigability: intuition

Hubs	All roads lead to Rome, so aim at Kevin Bacon
Hierarchy	Follow the chain of command
Chaos	Find useful random shortcuts

Navigability: intuition

Hubs	All roads lead to Rome, so aim at Kevin Bacon
Hierarchy	Follow the chain of command
Chaos	Find useful random shortcuts (is it possible?)

Nabigability: Kleinberg's grid

Goal

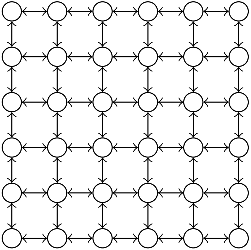
- Understand Milgram
- Prove navigability can emerge from randomness

Assumptions

- Nodes only know their neighbors
- They have a notion of "spatial" distance
- Decentralized greedy routing: send the letter to the «nearest» neighbor

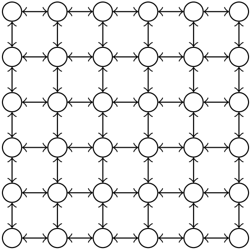
Nabigability: Kleinberg's grid

- A Manhattan nXn grid



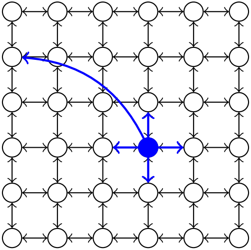
Nabigability: Kleinberg's grid

- A Manhattan nXn grid
- Each node is connected to neighbors at distance $\leq p$ ($p = 1$)



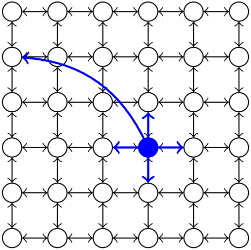
Nabigability: Kleinberg's grid

- A Manhattan nXn grid
- Each node is connected to neighbors at distance $\leq p$ ($p = 1$)
- Shortcuts: each node is connected to q "shortcuts" ($q = 1$) chosen with probability $\equiv \frac{1}{d^r}$



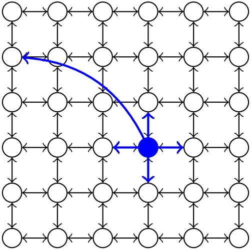
Nabigability: Kleinberg's grid

- A Manhattan *nXn* grid
- Each node is connected to neighbors at distance $\leq p$ ($p = 1$)
- Shortcuts: each node is connected to q "shortcuts" ($q = 1$) chosen with probability $\equiv \frac{1}{d^r}$
- Greedy routing



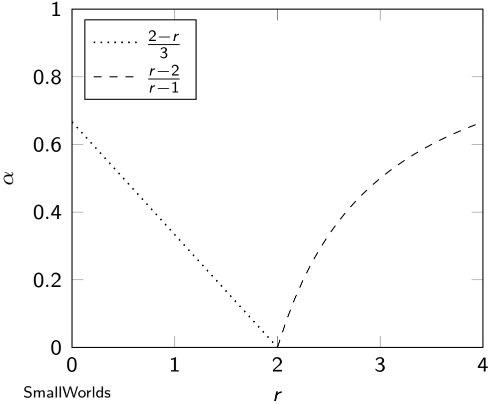
Nabigability: Kleinberg's grid

- A Manhattan *nXn* grid
- Each node is connected to neighbors at distance $\leq p$ ($p = 1$)
- Shortcuts: each node is connected to q "shortcuts" ($q = 1$) chosen with probability $\equiv \frac{1}{d^r}$
- Greedy routing
- Can we find short routes?



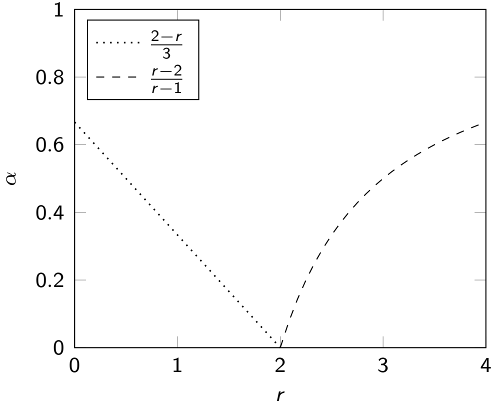
Nabigability: Kleinberg's grid

- $p = q = 1$ (minimum), $r = 2$: greedy routing is $O(\log^2(n))$
- $0 \leq r < 2$: any decentralized algorithm takes $\Omega(f(p,q)n^{(2-r)/3})$.
- $r > 2$: any decentralized algorithm takes $\Omega(f(p,q)n^{(r-2)/(r-1)})$.



Nabigability: Kleinberg's grid

- Proofs and simulations: final practical
- Greedy routing finds short paths, i.e. $O(\log^2(n))$
- Requires a proper tuning of the randomness?



Small-World Take-Away

- No unique model can explain how all connections are made

Small-World Take-Away

- No unique model can explain how all connections are made
- Most connections can be explained by (at least) one simple reason: locality, hierarchy, popularity. . . and randomness!

Small-World Take-Away

- No unique model can explain how all connections are made
- Most connections can be explained by (at least) one simple reason: locality, hierarchy, popularity. . . and randomness!

$\rightarrow$ Most social networks are small-worlds, but most small-world simple models only focus on specific properties.