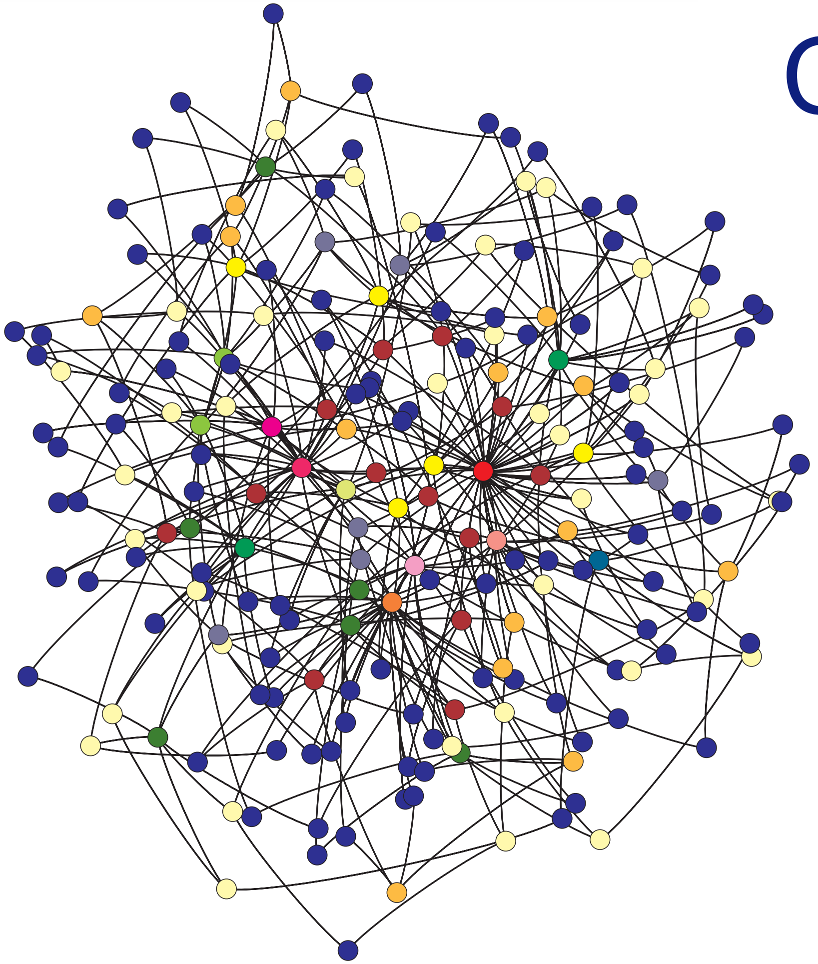
Imperial College London

Dept Colloquium 9th March 05
Tim Evans
Theoretical Physics



Complex Networks

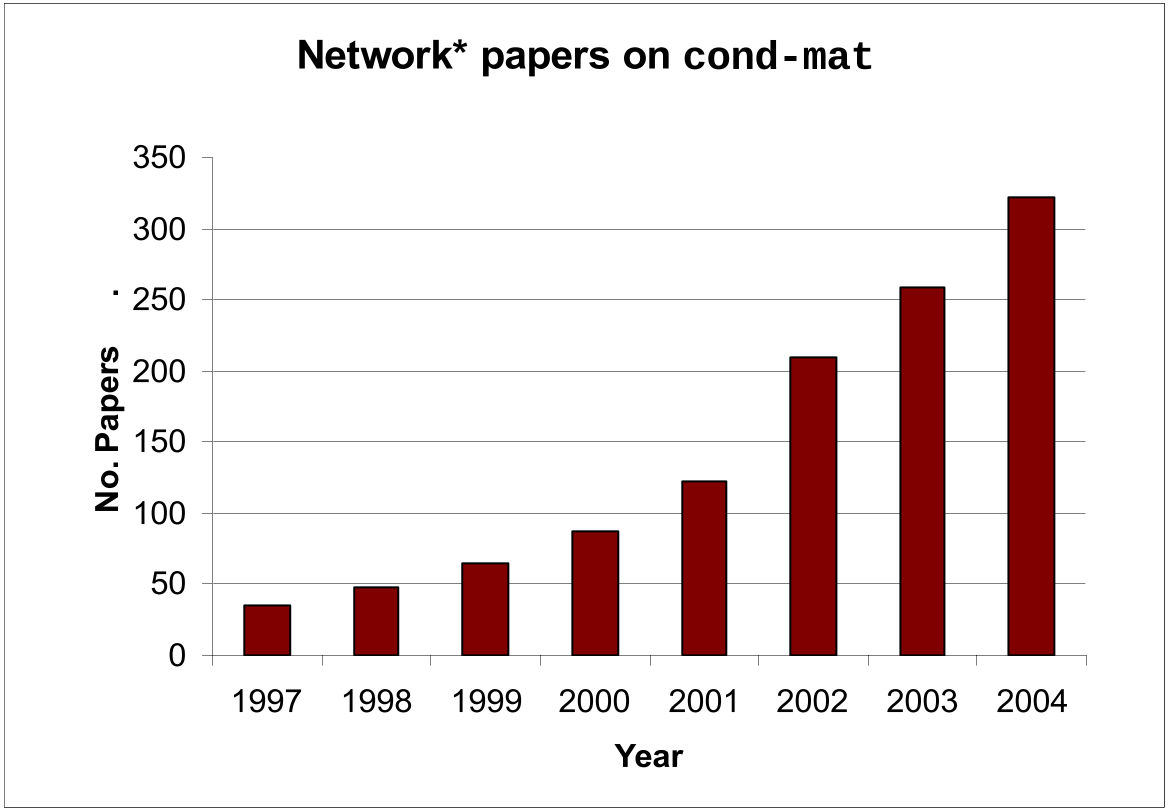
Six Degrees of Separation and all That

T.S.Evans
"Complex Networks"
Contemporary Physics
45 (2004) 455 - 474

Explosion of interest – WHY?

Since 1997 there has been an explosion of interest in networks by physicists.

For instance the condensed matter electronic preprint archives have gone from 35 papers in 1997 with a word starting with Network in their title to 322 last year, an increase of 800%



➡ **WHY?**

Multidisciplinary Nature

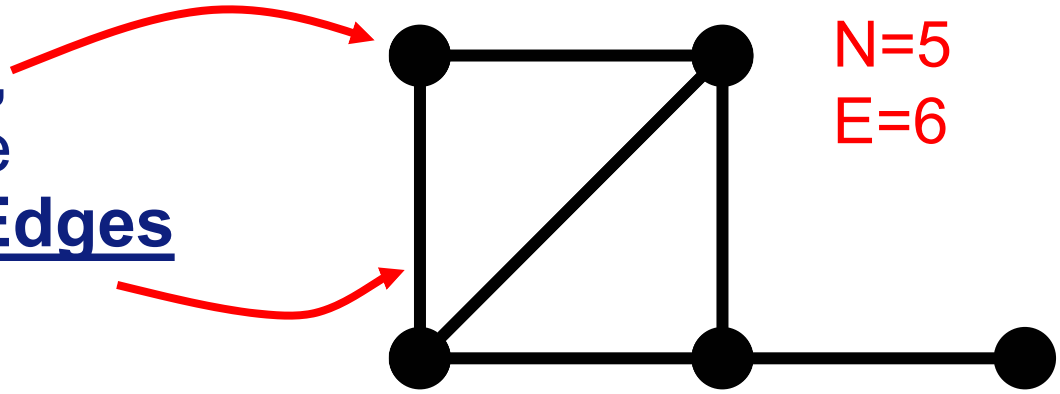
- Mathematics (Graph Theory, Dynamical Systems)
- Physics (Statistical Physics)
- Biology (Genes, Proteins, Disease Spread, Ecology)
- Computing (Web search and ranking algorithms)
- Economics (Knowledge Exchange in Markets)
- Geography (Transport Networks, City Sizes)
- Anthropology (Social Networks)
- Archaeology (Trade Routes)

http://www.iscom.unimo.it/

Basic Definitions

A **Network** or **Graph** is a collection of

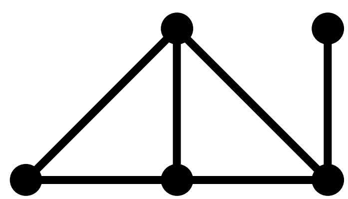
N **Vertices** (nodes), pairs of which are connected by **E** **Edges**



This is a SIMPLE graph, it has no other information.

In particular the *same* network can be shown in several identical ways.

=

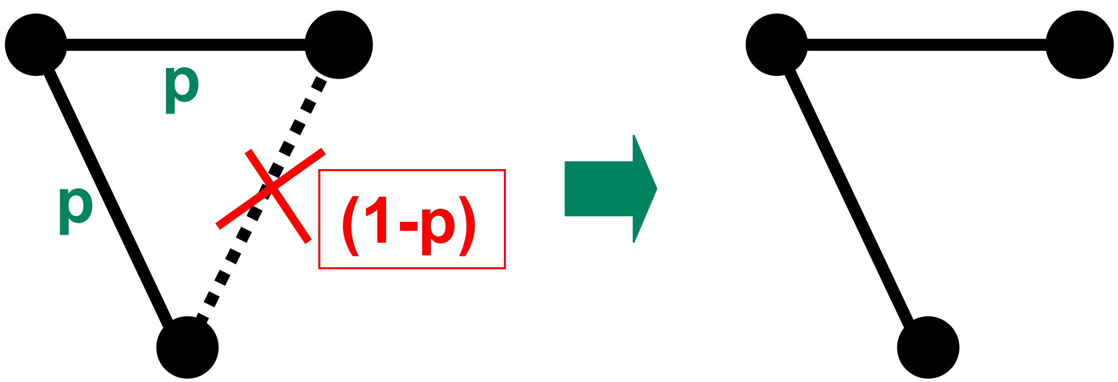


In general networks may have arrows on the edges (directed graphs), different values on edges (weighted graphs) or values to the vertices (coloured graphs).

Random Networks

- Take **N** vertices then consider every pair of vertices and connect each with probability **p**
  Erdős-Reyní (1959).
  This is the opposite of the perfectly ordered lattice.

N=3
p~2/3

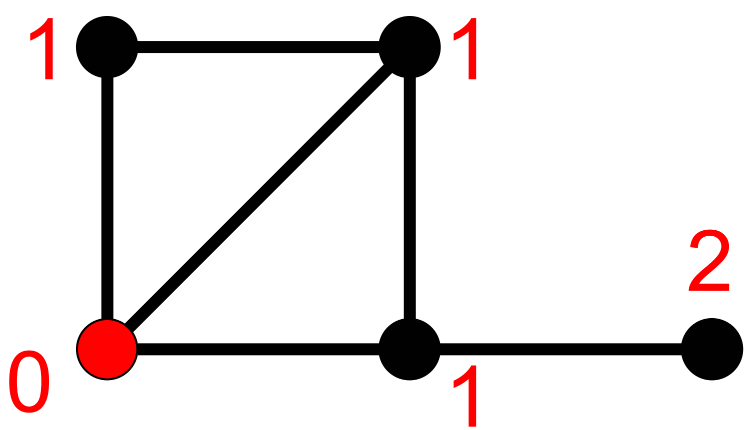


- How do we characterise the differences between these types of network?

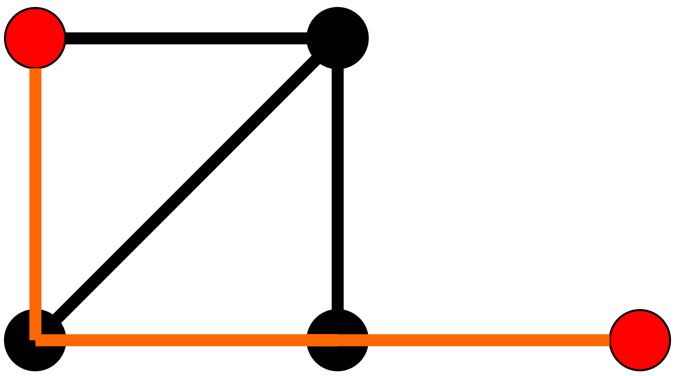
➡We need some more concepts first…

Network Distance

- Counting one for each edge traversed, we can find the shortest path between any two vertices, giving a distance between the two.
- The longest of these shortest paths is the diameter.



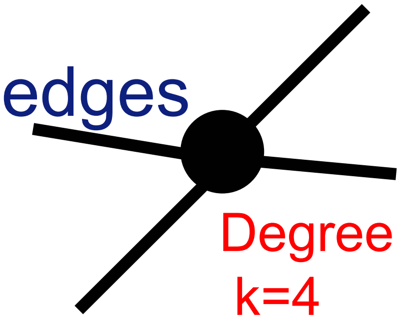
Distances from red vertex



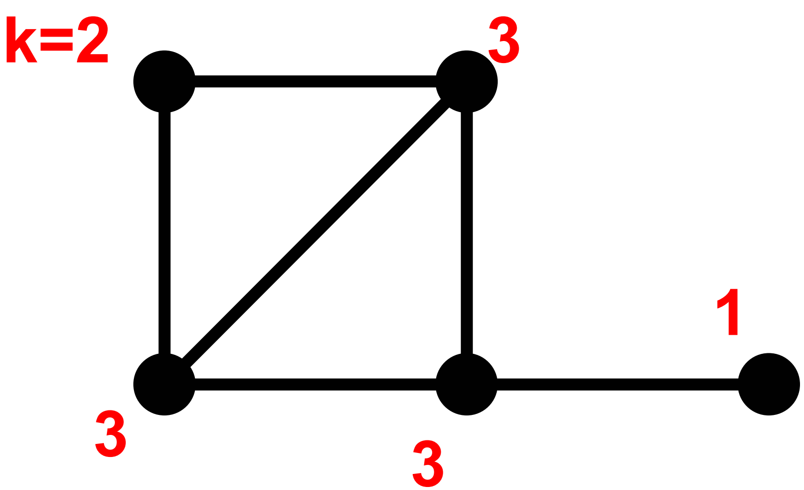
Diameter is **3**, between red vertices

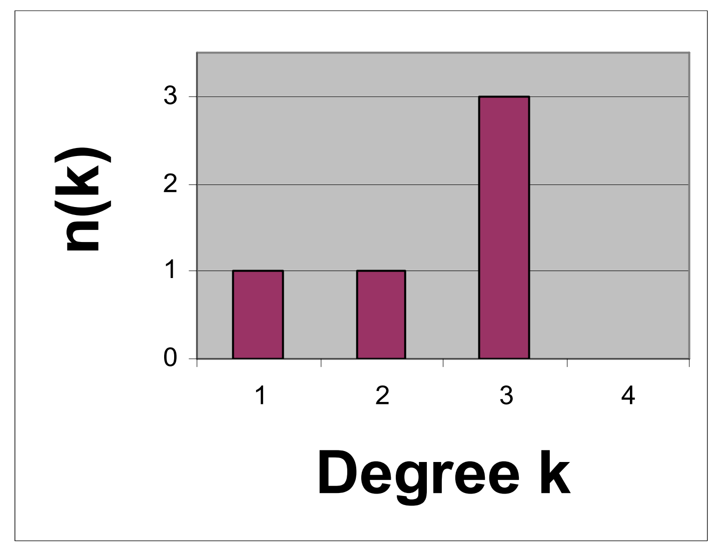
Degree (connectivity)

- The **<u>Degree</u> k** of a vertex is the number of edges attached to it.



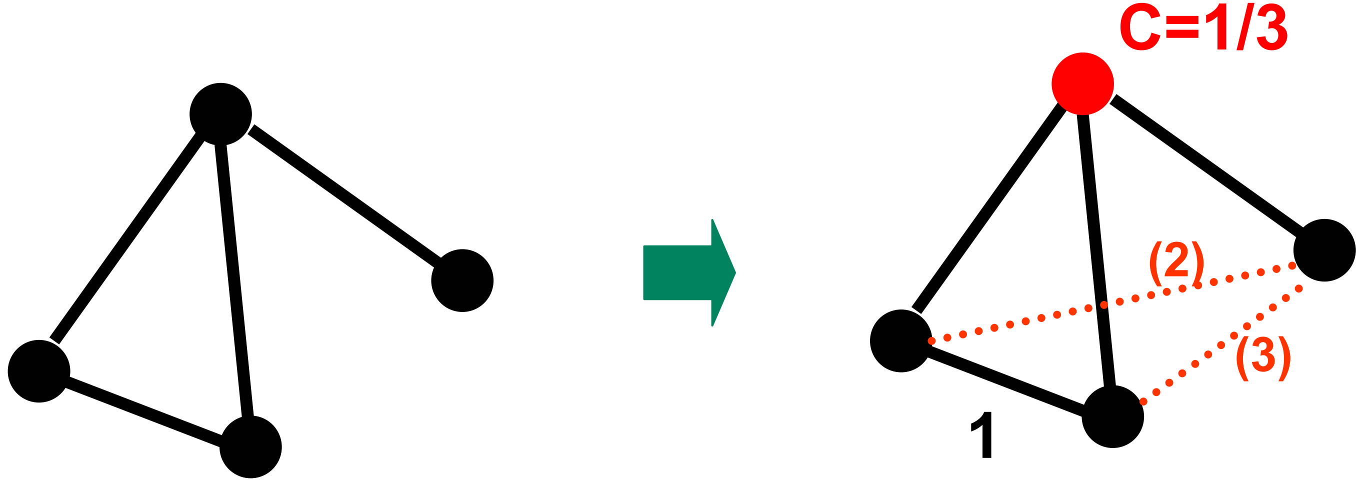
- The **<u>Degree Distribution</u> n(k)** is the number of vertices with degree k





Cluster Coefficient

- <u>**Clustering coefficient**</u> **c**:
  Fraction of the neighbours which are themselves connected
  Simple measure of how much local structure there is in a network



Lattice vs. Random Networks

- Lattices are Large $d \sim N^{1/dim}$
  Random Networks are Small $d \sim \log(N)$

- Lattices have fixed degrees $k = k_0$
  Random Networks have a small range $k < \sim\log(N)$

- Lattices have large cluster coef. $c \sim O(1)$
  Random Networks have very small clustering coefficients $c \sim \langle k \rangle / N$

But is the real world ordered or random?

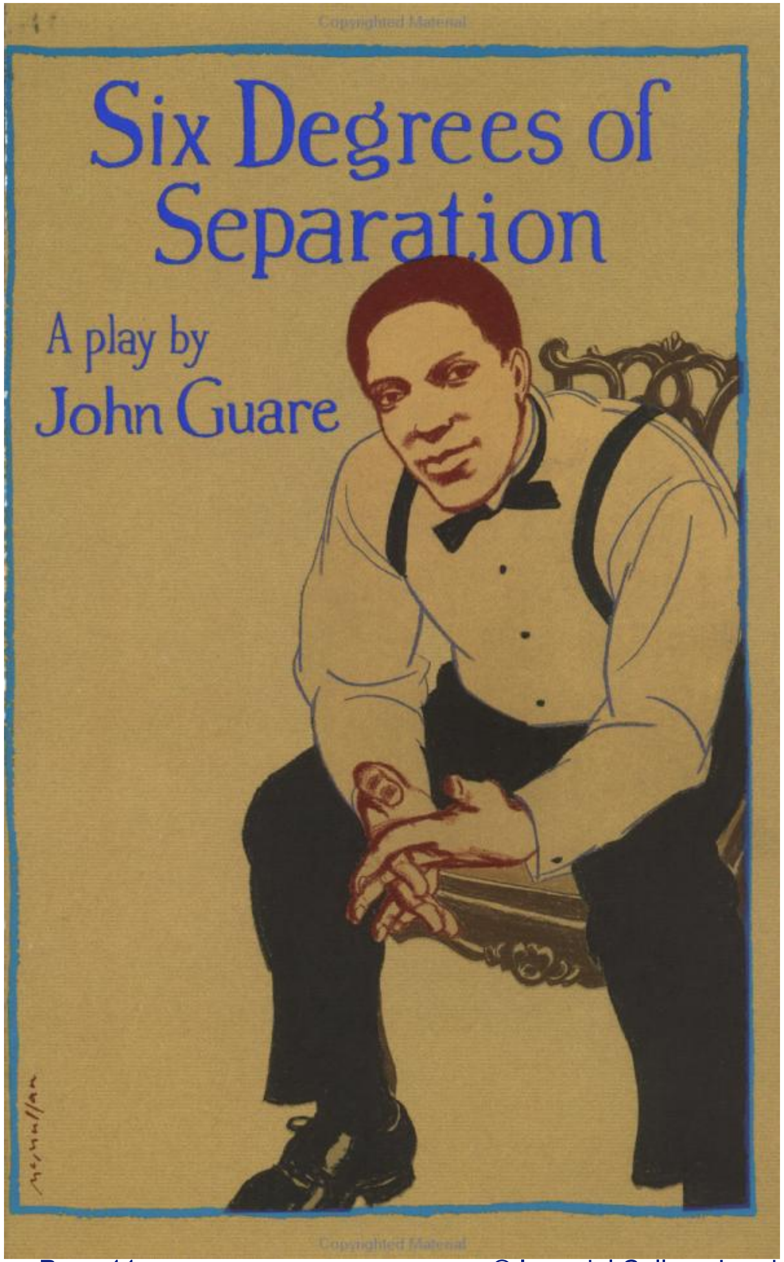
Social Networks and Small Worlds

- Let each vertex represent a person, and let the edges represent friendships
- Friendships are not limited by physical distance

Milgram (1967) asked people in Omaha (Nebraska) and Wichita (Kansas) to send packets to people in Cambridge MA specified by name, profession and rough location only. Packets were only swapped between people who knew each other by first name. If the packets arrived at the correct person, they had been through about five intermediaries.

➡ It's a **Small World**

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 everybody else on this planet. The president of the United States. A gondolier in Venice. … It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. … How every person is a new door, opening to other worlds. Six degrees of separation between me and everyone else on this planet."

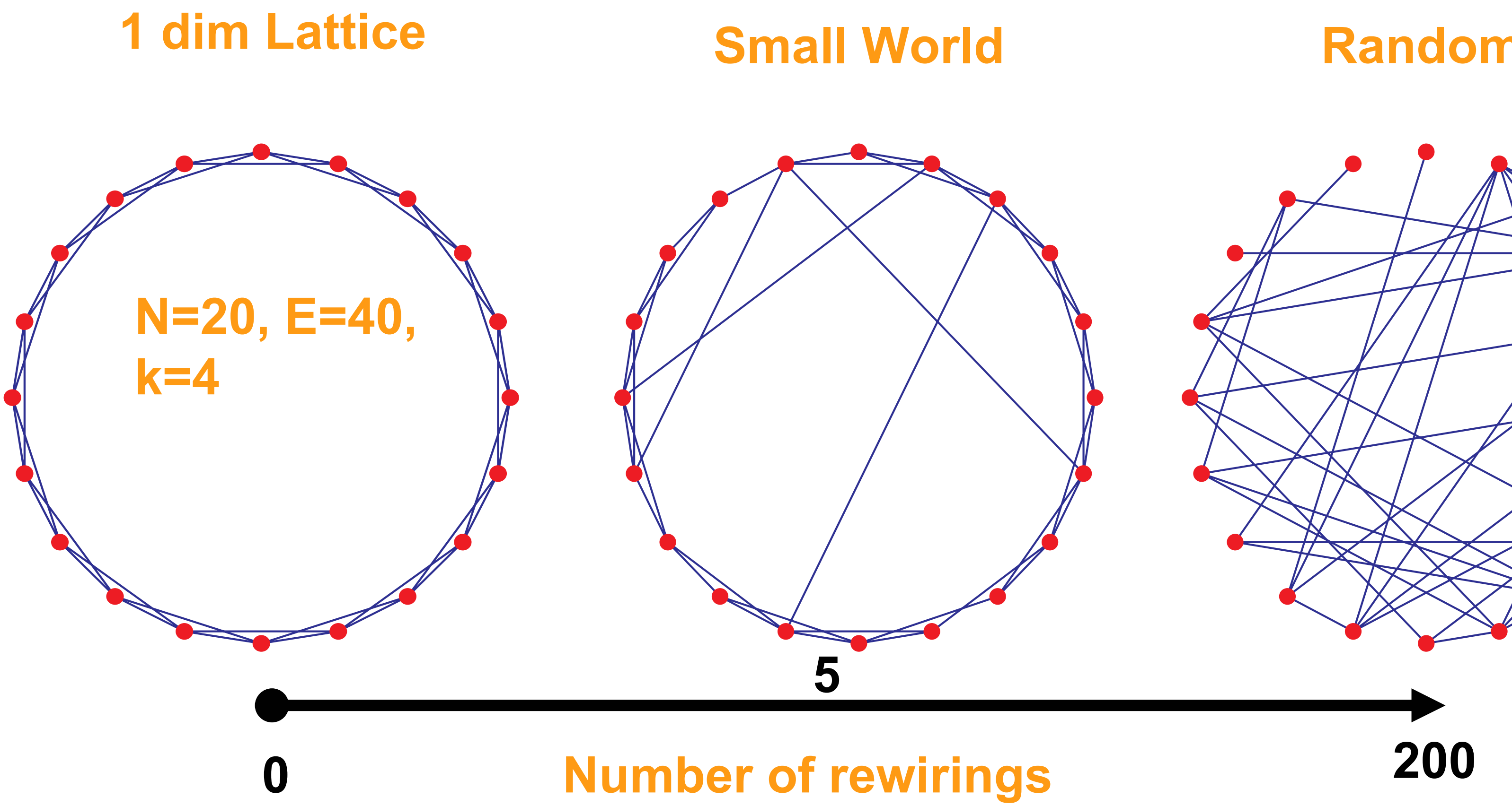
What sort of network is a social network?

- Social Networks have short distances like a random network but unlike a lattice (Milgram)
- Social Networks also have local structure as my friends often become (or were already) friends of each other – high cluster coefficients unlike random networks but like a lattice
- Friendships are neither purely random nor precisely ordered

➡Need something new…

Watts and Strogatz's Small World Network (1998)

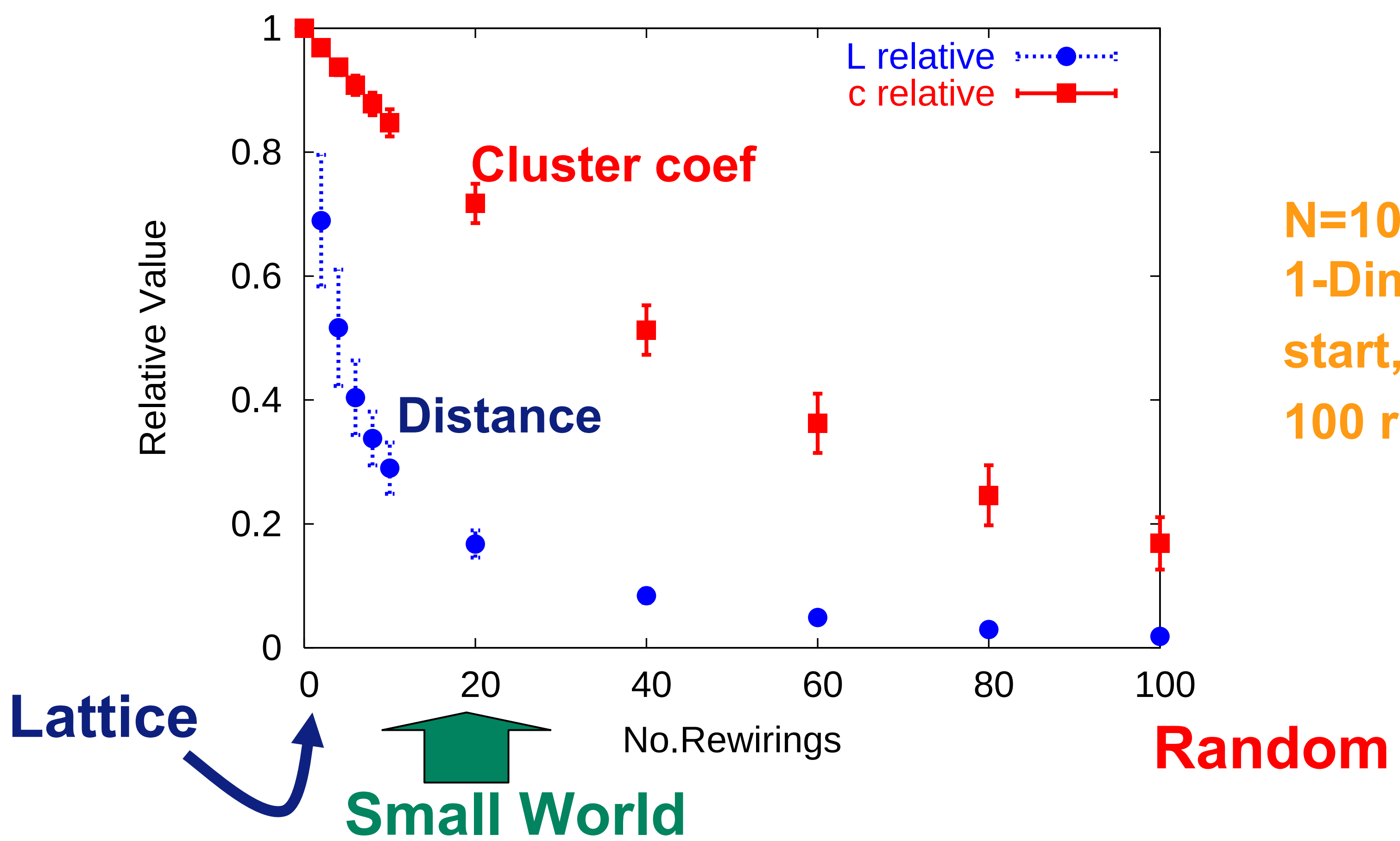
- Start with lattice, pick random edge and rewire – move it to two link two new vertices chosen at random.



Clustering and Length Scale in WS network

- As you *rewire*, distance drops very quickly, clustering does not

➡ Find **Small World networks** with short distances of random network, large clustering and local structure like a lattice



N=100, k=4, 1-Dim lattice start, 100 runs

Network Comparison so far

Network Type	Distance d	Cluster Coef. c
Lattice	Large $d \sim N^{1/dim}$	~1
WS Small World	Small $d \sim \log(N)$	~ 1
Random	Small $d \sim \log(N)$	~ 1/N

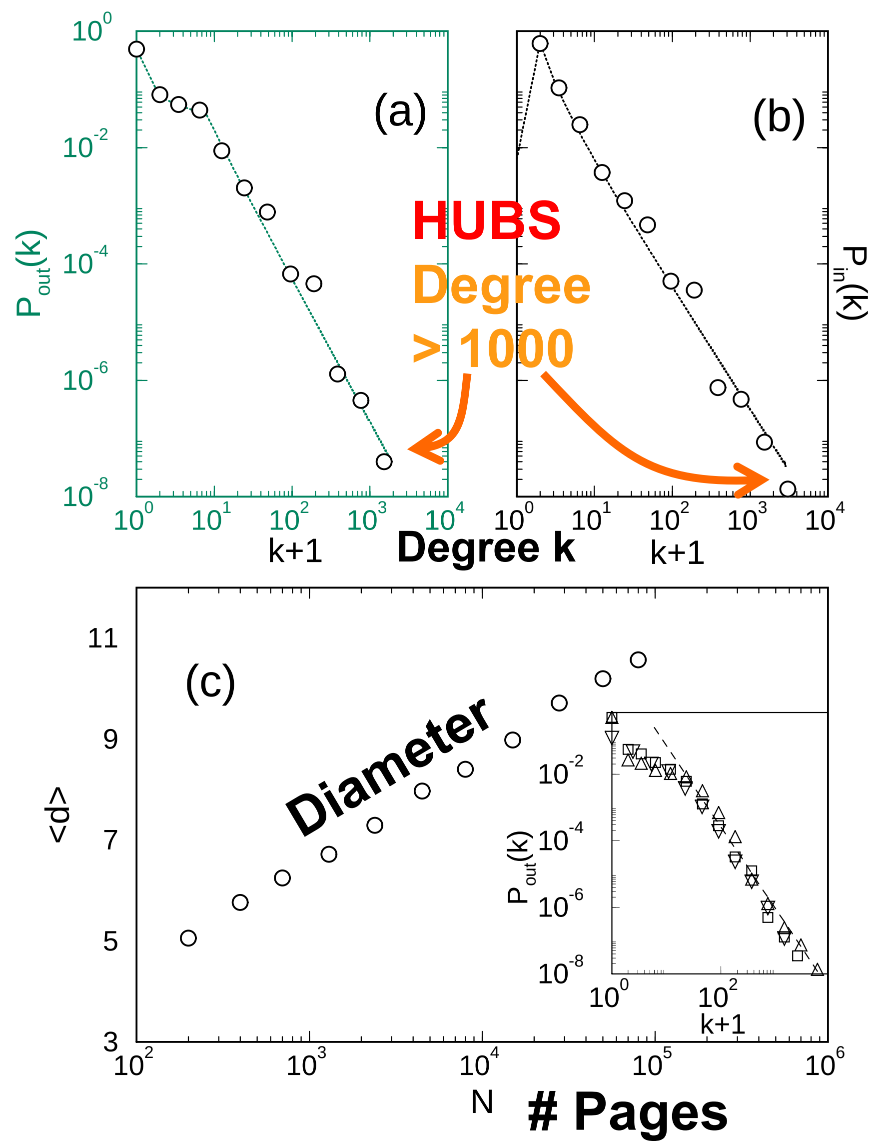
So WS Small World might be good for social networks

But what about other physical networks?

The World Wide Web

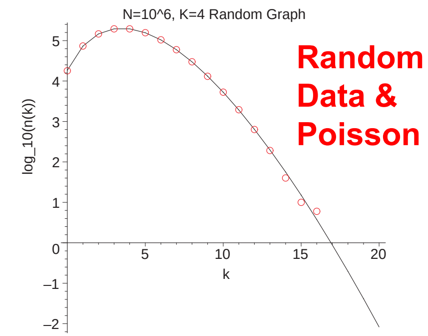
- Every web page is a vertex, every link is an edge
- A few pages have a tremendous amount of links to them e.g. college home page, eBay, Google
These are **Hubs** and they are a key aspect of how we navigate and use the web

Log-Log plot of degree distribution of nd.edu

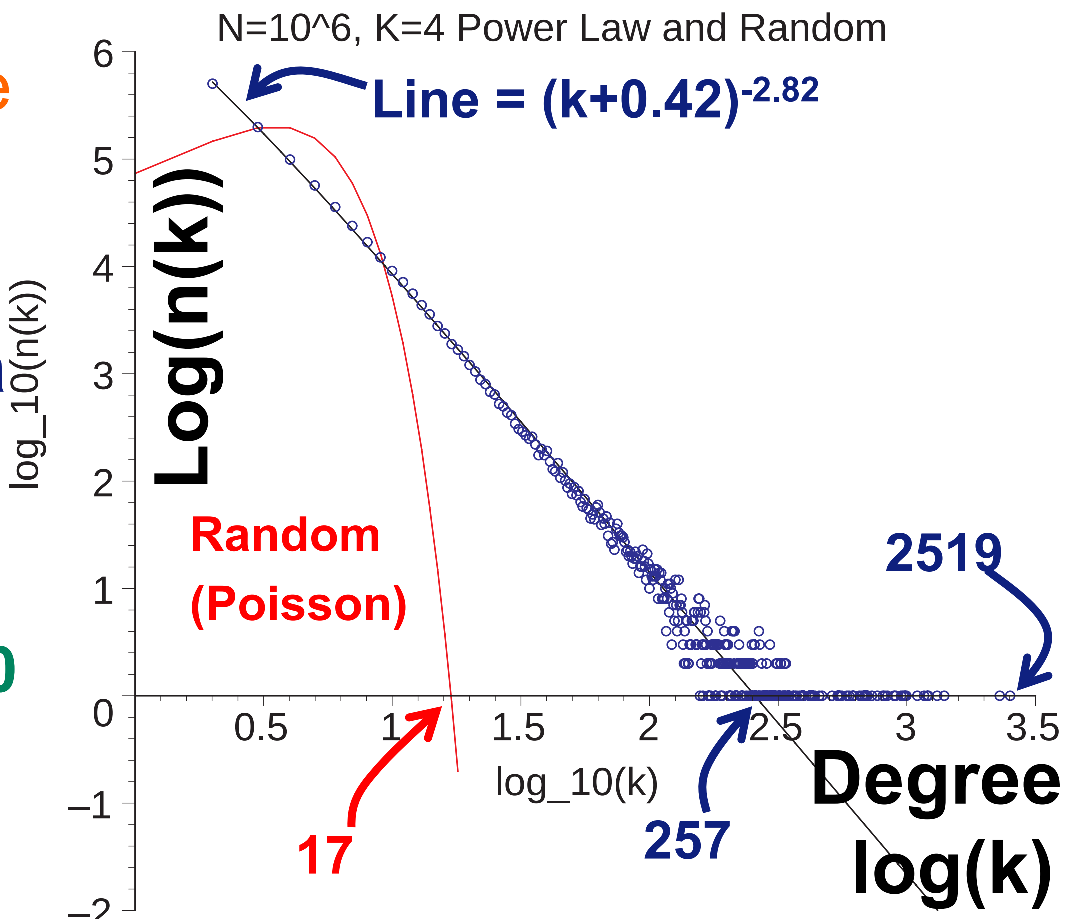


(Barabasi, Albert, Jeong 1999)

What sort of network has hubs?



- Lattices, WS (Watts-Strogatz) Small World and random networks have no hubs, e.g. the largest degree is **17** for a random network with $N=10^6$, <k>=4
- Want a network with a **long tailed** degree distribution e.g. power law ~ k^{-3} has max. degree ~**2520** for $N=10^6$ <k>=4



Scale Free Networks

- Any network with a **power law degree distribution** for large degrees $\lim_{k \to \infty} [n(k)] \propto k^{-\gamma}$
- Always have many large **Hubs** nodes with many edges attached – e.g. routers in the internet
- **Scale Free** means the number of vertices of degree 2k with those of degree k, always the same whatever k , that is there is no scale for degree

$$\frac{n(2k)}{n(k)} = \text{constant}$$

- In practice there are at least two scales for finite N: $O(1) \sim k_{min} \leq k \leq k_{max} \sim O(N^{1/(\gamma-1)})$

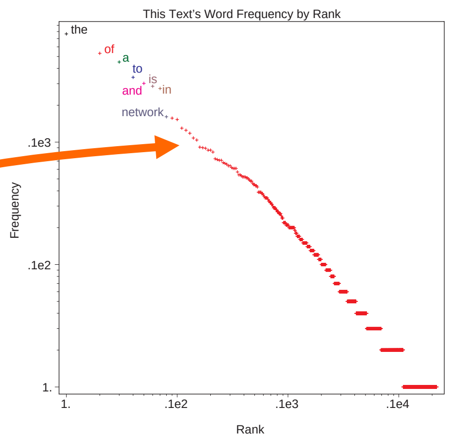
Network Comparison

Network Type	Distance d	Degree Distrib. n(k)	Maximum Degree k_{max}	Cluster Coef. c
Lattice	Large $d \sim N^{1/dim}$	No Tail $\delta(k-k_0)$	Fixed k_0	~1
WS Small World	Small $d \sim \log(N)$	No Tail $\sim \delta(k-k_0)$	V.Small $\sim k_0$	~ 1/N
Random	Small $d \sim \log(N)$	Short Tail Poisson $\langle k\rangle^k e^{-\langle k\rangle} / k!$	Small $\sim\log(N)$	~ 1/N
Scale-Free	Small $d \sim \log(N)$	Long Tail $\sim k^{-\gamma}$	Large = HUBS $\sim k^{1/(\gamma-1)}$	

How often do we see Scale-Free networks?

Power Laws in the Real World

- 2^{nd} Order Phase Transitions
  (e.g. superconductors, superfluids,…)
  Long range order = no scale = physical insight
  Critical Phenomena – Renormalisation Group
- Scaling in Particle Physics
- Biology
  – Kleiber's Law (1930's) metabolic rate $\mathbf{r} \propto \mathbf{m}^{3/4}$ body mass,
  explained (West, Brown, Enqvist 1997)
- Social Sciences
  – Zipf's Law (1949)
  City sizes,
  Word frequency, …
  ➔file compression

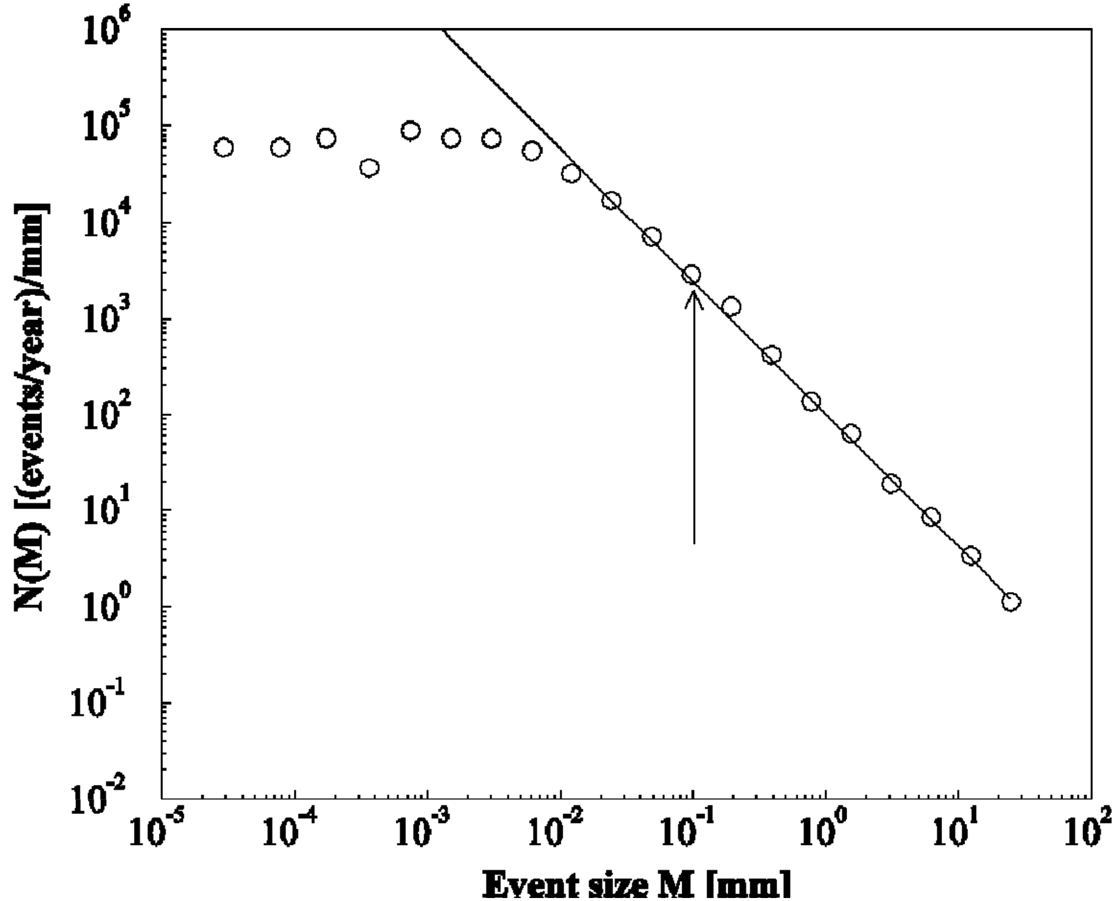


Scaling in Complex Systems

- Earthquakes
  (Gutenberg-Richter Law),
  forest fires,
  rice piles,
  rainfall distributions,
  etc etc

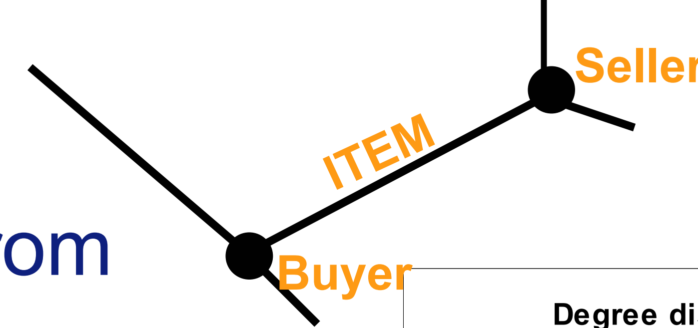
  Self-Organised Criticality

- **Still leading to further physical insights**

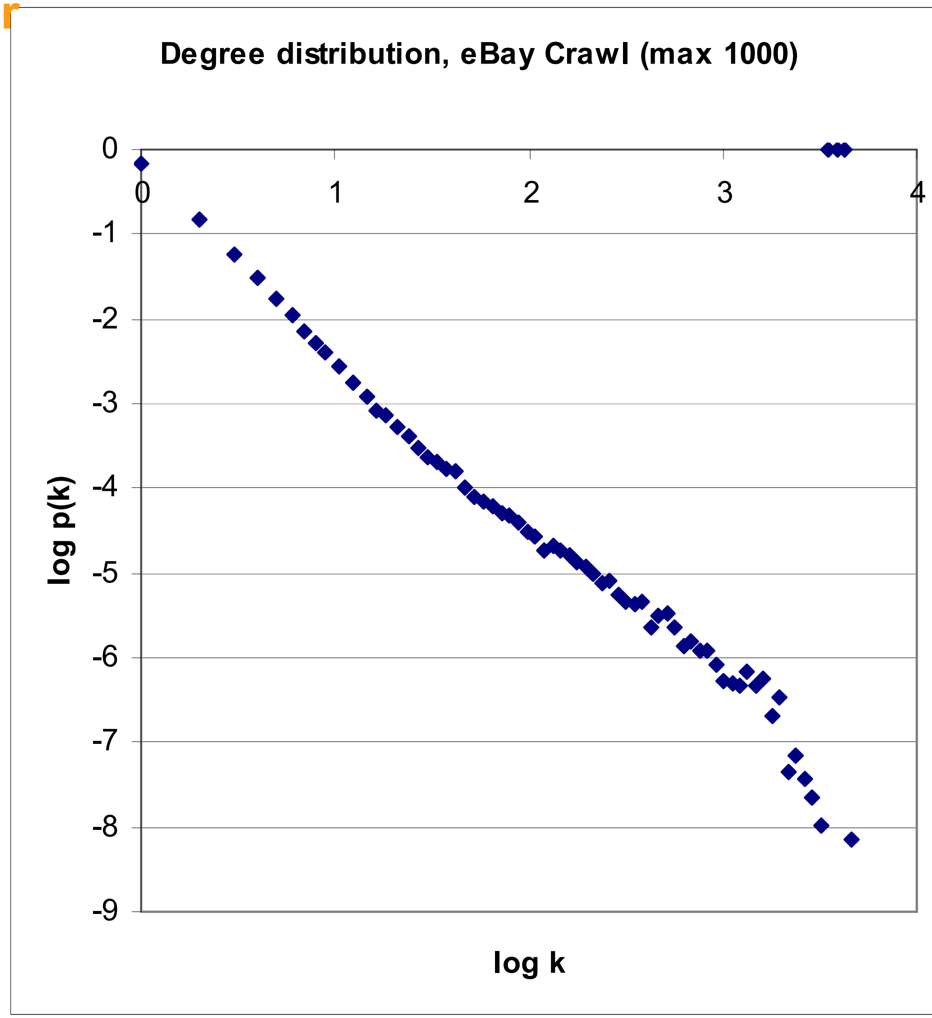


(Peters, Hertlein, Christensen 2002)

eBay



- Network from buyer/seller feedback links
- eBay is dominated by a few very large hubs.
  - The slight curvature due to crawling method.
  - Fetched 5,000 pages and built up a network of 318,000 nodes and 670,000 edges
- $\gamma \approx 2.3$



Sooman, Warren, TSE (2004)

Scaling with every network

- Friendship networks - Kevin Bacon game
- Scientific Collaboration Networks - Erdős number
- Scientific Citation Networks
- Word Wide Web
- Internet
- Food Webs
- Language Networks
- Protein Interaction Networks
- Power Distribution Networks
- Imperial Library Lending Data (Laloe, Lunkes, Sooman, Warren, Hook, TSE)
- eBay relationships (Sooman, Warren, TSE)
- Greek Gods
- Marvel Comic Heroes

Scaling – a health warning

Almost every network is scale free if you believe the literature but

- Not many decades of data
  e.g. 10^6 vertex scale free network has largest vertex about 1000 so at most two decades of large degree scaling
- Data often a single data set no repeats
- Errors unknown in much social science data
- Other long tailed distributions have hubs too

Models of Scale Free Networks

- Growing with Preferential Attachment

  (Barabasi,Albert 1999; Simon 1955)

  Add new vertex with <k>/2 new edges

  Attach to existing vertices chosen with probability Π proportional to their degree $\Pi \propto k$

  - *but* $\Pi \propto k^{\alpha}$ for any $\alpha \neq 1$ *fails*

- The Walk (TSE, Klauke 2002; Saramäki, Kaski 2004;TSE, Saramäki 2004)

  Add a new vertex with <k>/2 new edges

  Attached to existing vertices found by executing a random walk on the network

  - automatically generates scale-free

  A self-organising mechanism?

  e.g. rumours propagating on a friendship network

Applications

- Understand better how some complex systems work if we can understand the patterns in their networks
  e.g. proteins in biological systems
- Spread of viruses better studied on more realistic networks, better preventative methods?
- Search Algorithms – how did Milgram's letters reach their destination? Why did 80% fail to arrive?

Applications: Network Resilience

- Resilience of Networks, if remove vertices (hardware failure, virus attack) when does a network split into small disconnected pieces?
- For random vertex removal:
  Random Network fails when $2E/N = \langle k \rangle = 1$
  A Scale Free $\gamma < 3$ **never** fails
- Remove biggest hubs first:
  Scale Free Networks fail very quickly

Build a scale free network but make sure the hubs have the best protection!

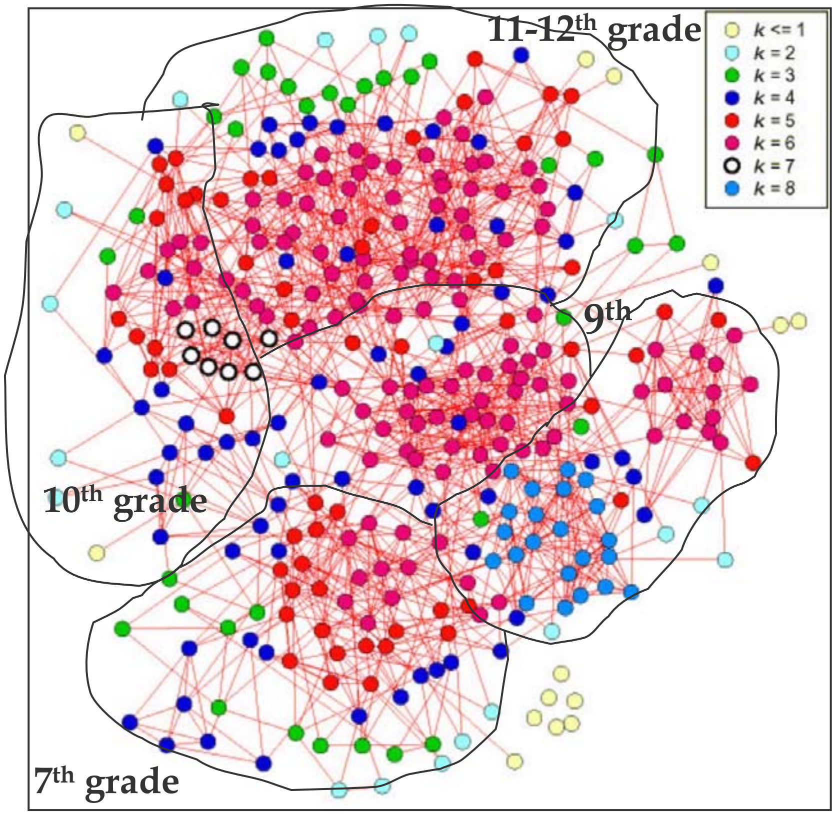
Applications: Communities

- Communities – close relationships vertices revealed only though analysis of network

  e.g. undiscovered collaboration opportunities from library lending patterns (Sooman, Warren, TSE)

  e.g. Quantitative tools to analyse social networks (can use Q-state Potts model)

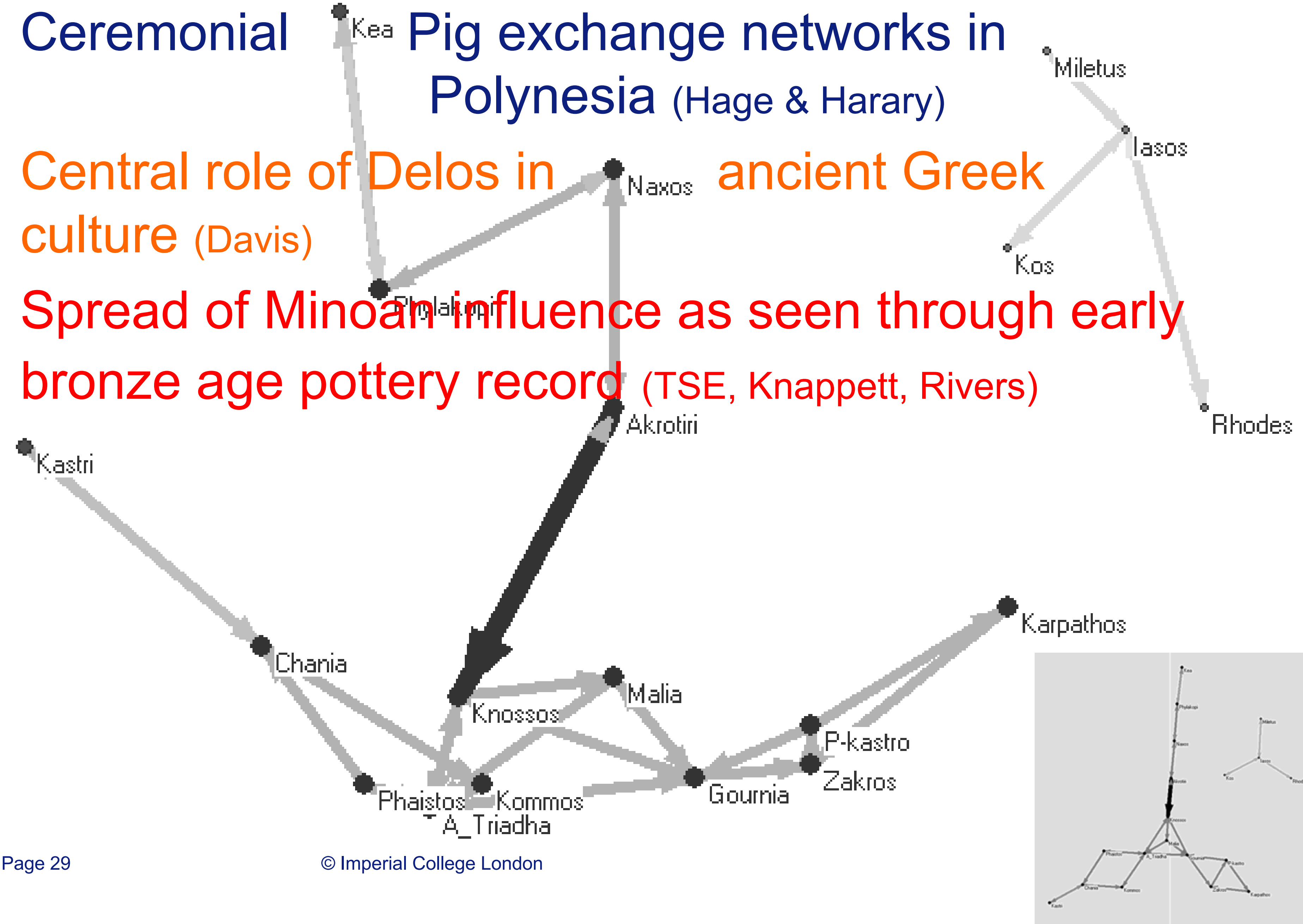
Friendship Cohesion in an American high school
Moody, White (2003)



Social embeddedness in nested cohesive subgroups applied to high school friendship network. Red individuals are cores of most cohesive subgroups, only trans-grade for older kids

Applications: Historical patterns

- Ceremonial Pig exchange networks in Polynesia (Hage & Harary)
- Central role of Delos in ancient Greek culture (Davis)
- Spread of Minoan influence as seen through early bronze age pottery record (TSE, Knappett, Rivers)



Summary

- New network models over last 7 years
  often more realistic than older ones
- New tools to analyse networks
  many from statistical physics
- Extensive experience in other fields
  not to be dismissed as methodology techniques often quite different
- New results, new applications, new views of old problems
- Deep insights still awaited
  where do the power laws in social sciences come from?

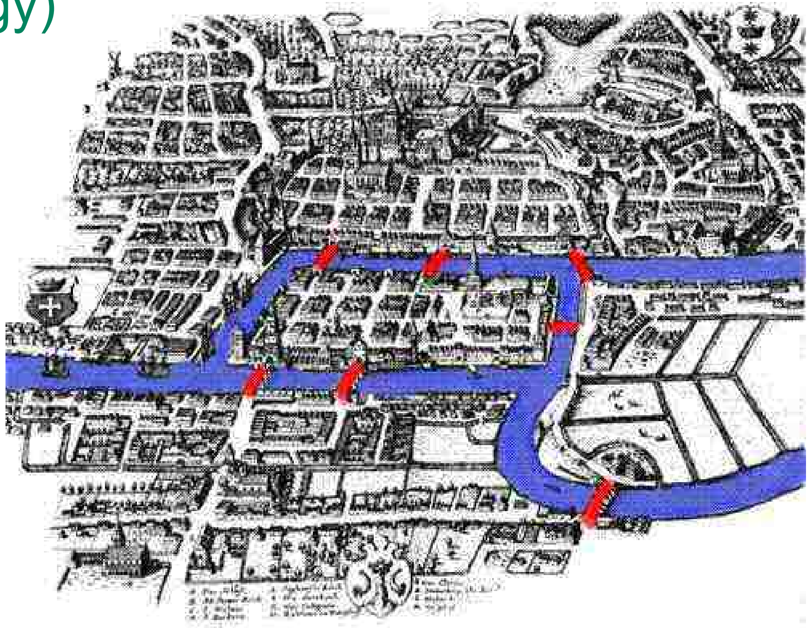
I couldn't have done this without …

- Project Students
  Seb Klauke, JB Laloë, Christian Lunkes, Karl Sooman, Alex Warren
- ISCOM organisers
  David Lane, Sander van der Leeuw, Geoff West and all the ISCOM participants
- Collaborators
  Daniel Hook, Carl Knappet, Ray Rivers, Jari Saramäki

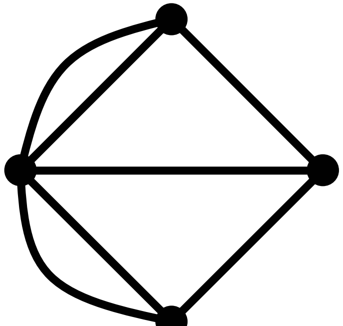
T.S.Evans, "Complex Networks", Contemporary Physics **45** (2004) 455 - 474

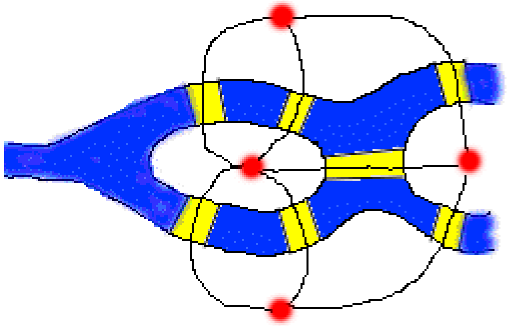
… but I've only got myself to blame

The start of graph theory (and topology)



In 1735 the great Swiss mathematician Euler showed that it was impossible to walk around the city of Königsberg (now Kaliningrad) crossing the 7 bridges over the river Pregel once and only once.





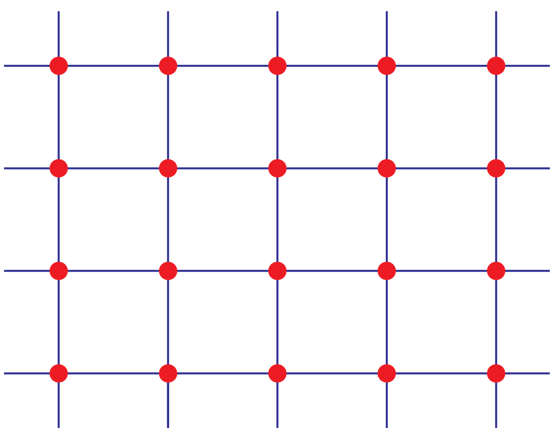
A Network or Graph

Visualisation

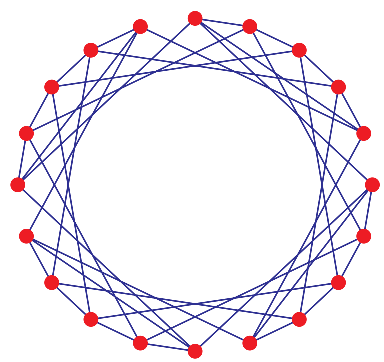
In many networks the location of vertices in figures has no meaning. A Graph records relationships between vertices as edges which may have no simple numerical value or at least no simple interpretation as a distance (no metric)

e.g. friendship networks.

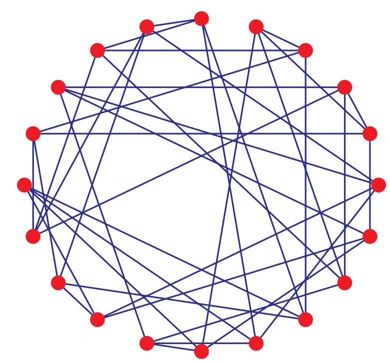
➡ Networks have no dimension, live in no obvious (Euclidean) space



Periodic Lattice



Same network with vertices arranged in regular order.

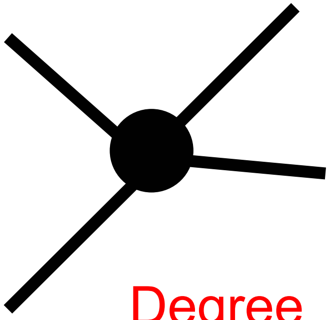


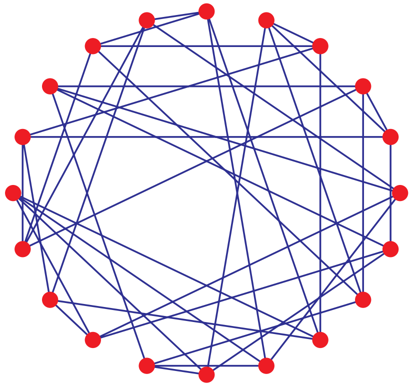
Same network with vertices arranged in random order

N=20, E=40

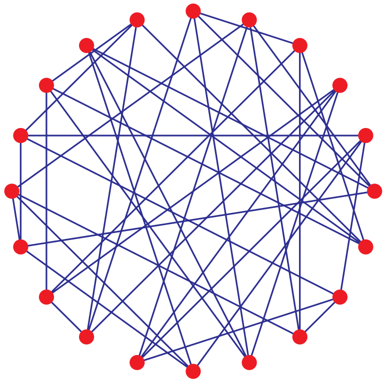
Degree

The Degree k of a vertex is the number of edges attached to it.

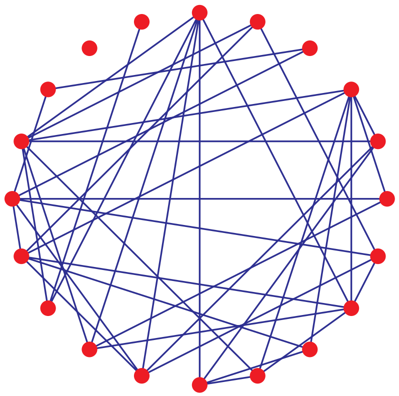




Periodic so always degree 4

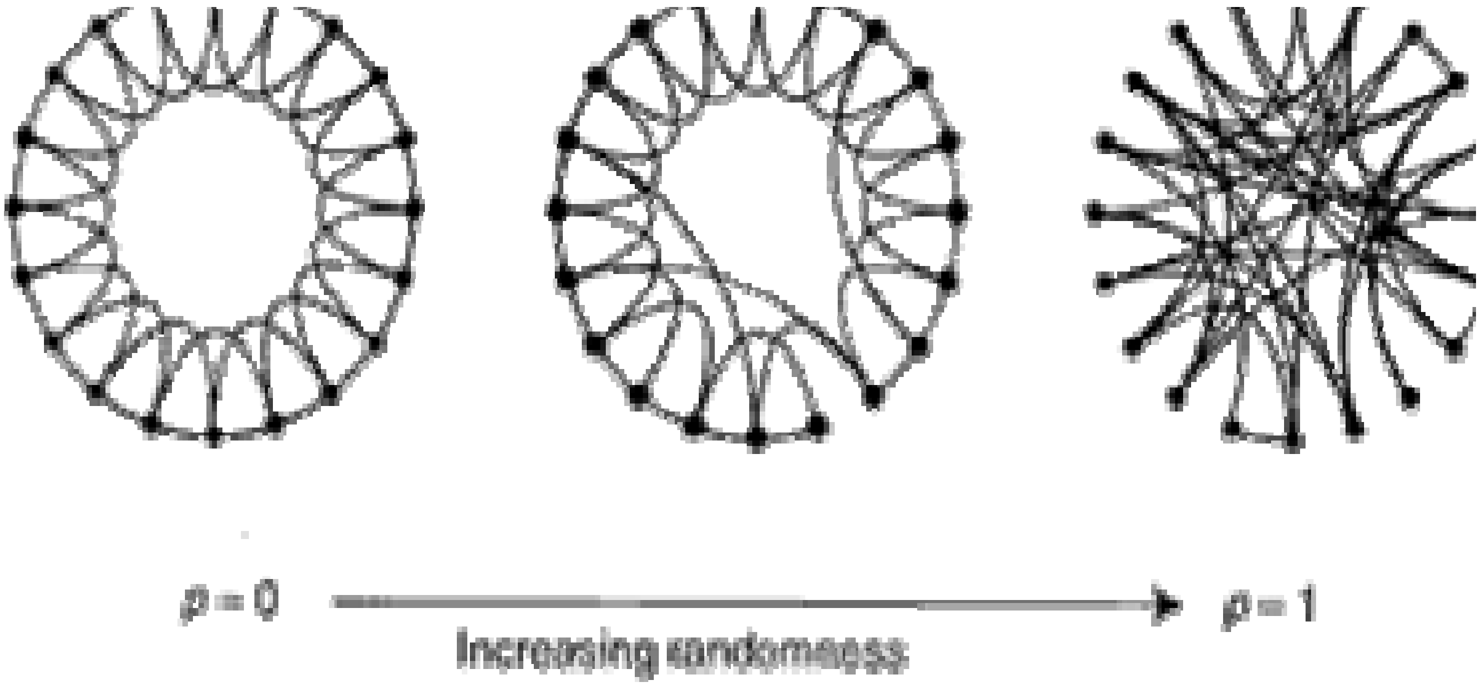


Random but always degree 4



Random but average degree 4

Watts and Strogatz’s original diagram



Scaling in Biology

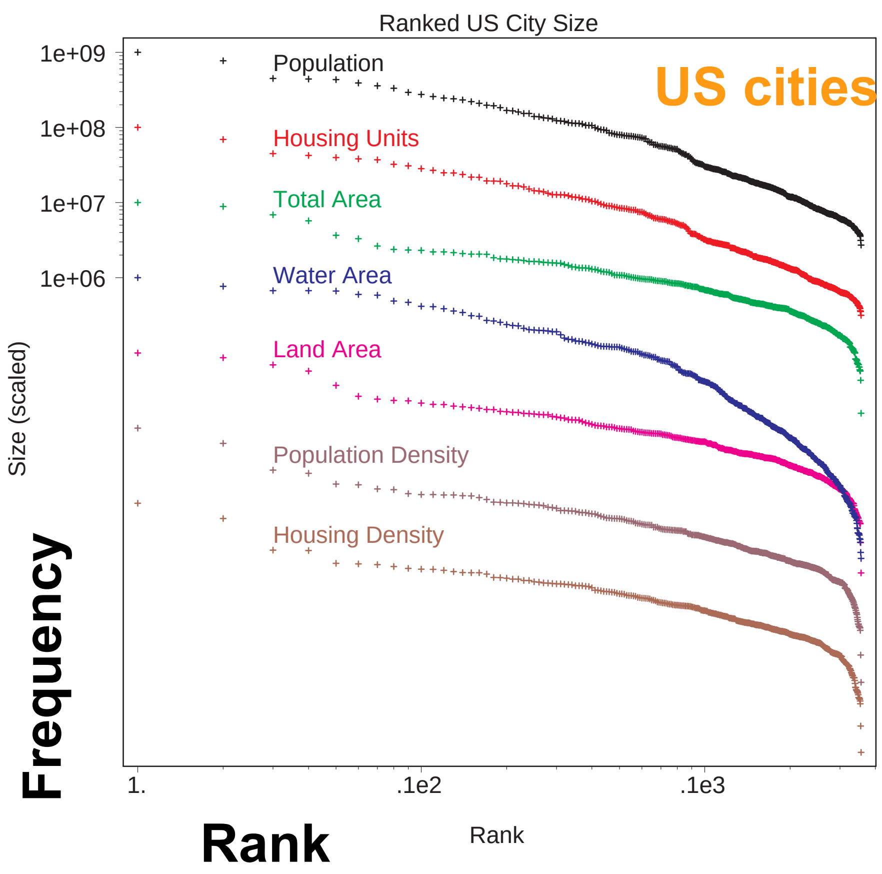
- e.g. Kleiber's Law (1930's)
  From cells to blue whales
  metabolic rate r, body mass m

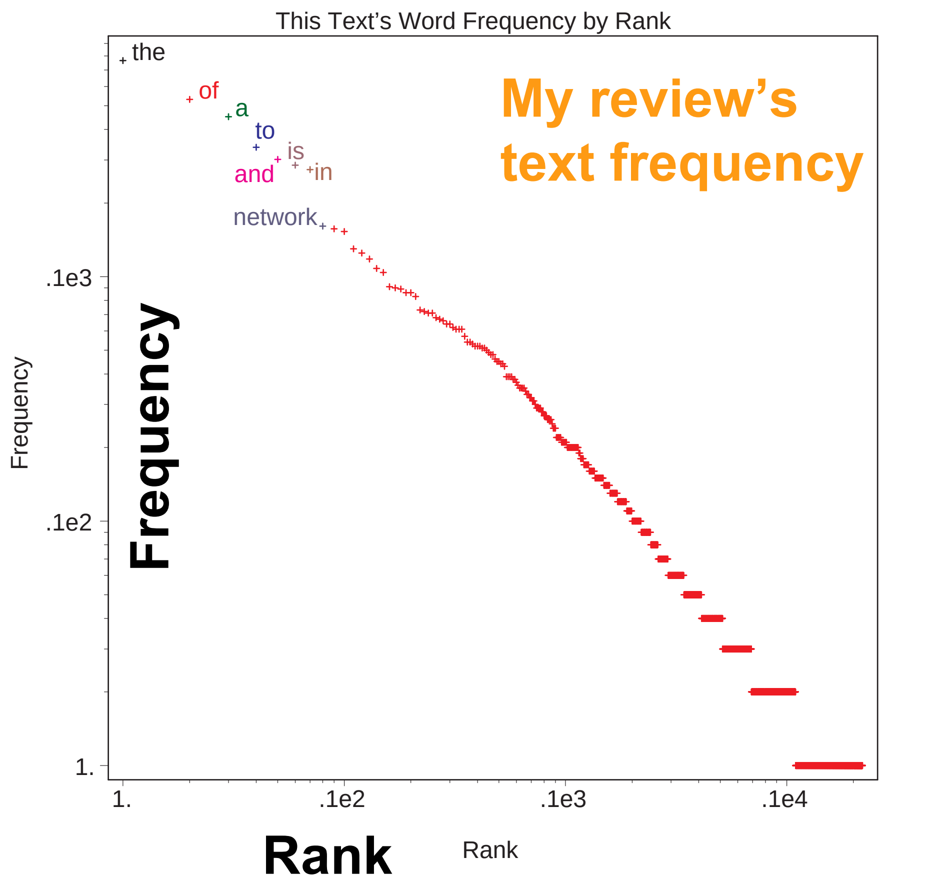
$$r \propto m^{3/4}$$

- Explanation in terms of physics of space-filling networks needed to deliver energy to all parts of an organism, from particle theorist (where scaling is a vital concept)
  Geoff West
  and biology colleagues
  (West, Brown, Enqvist 1997)

Scaling in Social Sciences

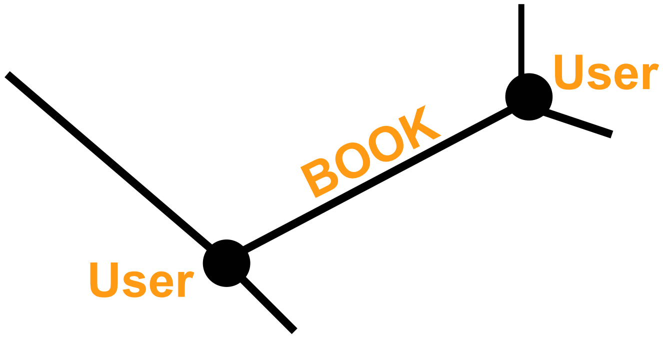
- Zipf law (1949) – City Sizes, Text Frequencies,…
- Pareto's 80:20 rule (1890's)

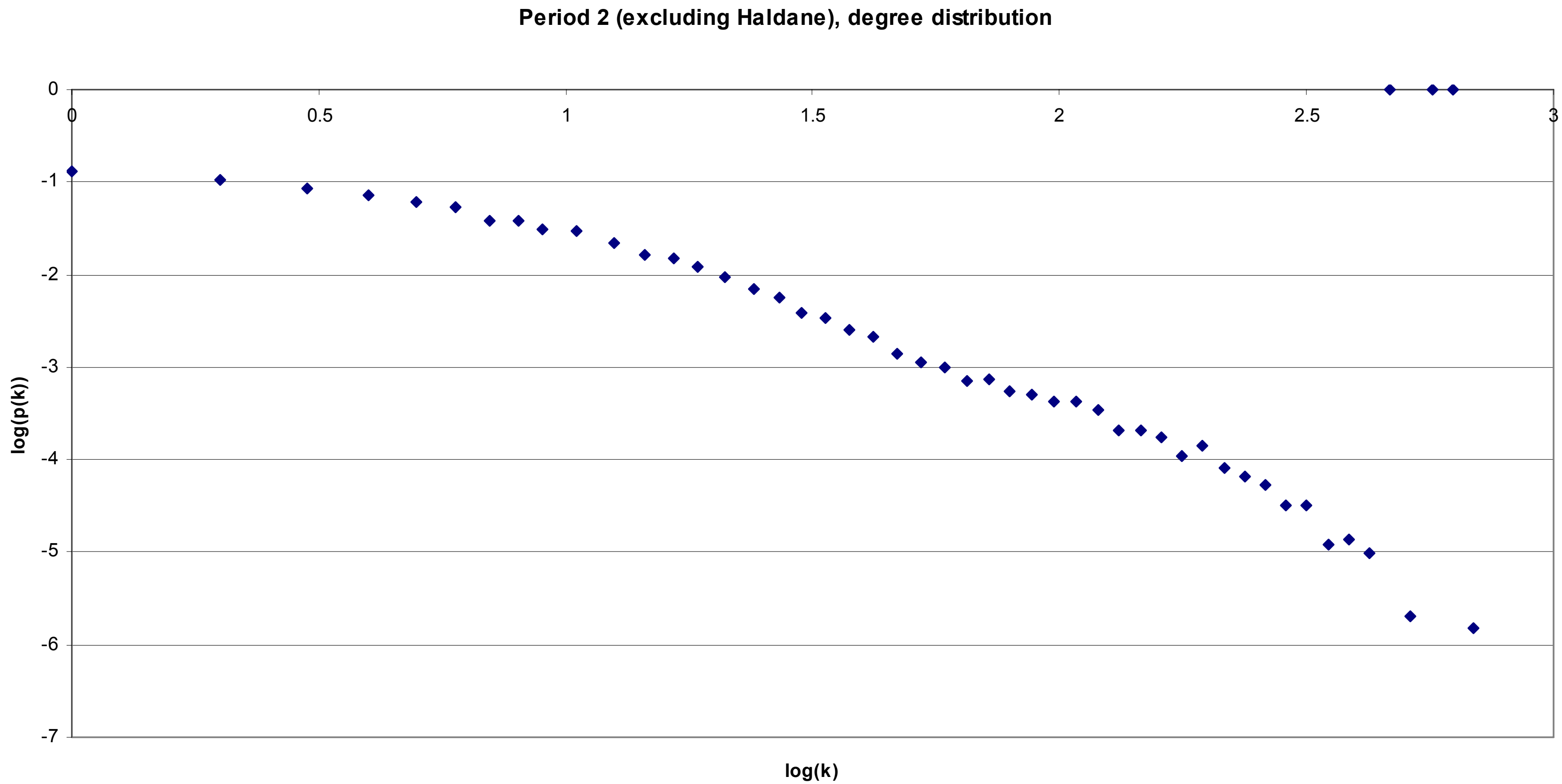




Imperial Library

- Used to detect groups from lending patterns





Sooman, Warren, Hook, TSE (2004)

Network Comparison

Network Type	Distance d	Degree Distrib. n(k)	Maximum Degree k_{max}	Cluster Coef. c
Lattice	Large $d \sim N^{1/dim}$	No Tail $\delta(k-k_0)$	Fixed k_0	~1
WS Small World	Small $d \sim \log(N)$	No Tail $\sim \delta(k-k_0)$	V.Small $\sim k_0$	~ 1/N
Random	Small $d \sim \log(N)$	Short Tail Poisson e^{-k}	Small $\sim\log(N)$	~ 1/N
Scale-Free	Small $d \sim \log(N)$	Long Tail $\sim k^{-\gamma}$	Large HUBS $\sim k^{1/(\gamma-1)}$	~

But is the real world ordered or random?

Network Comparison

Network Type	Distance d	Degree Distrib. n(k)	Maximum Degree k_{max}	Cluster Coef. c
Lattice	Large $d \sim N^{1/dim}$	No Tail $\delta(k-k_0)$	Fixed k_0	~1
WS Small World	Small d ~ log(N)	No Tail $\sim \delta(k-k_0)$	V.Small $\sim k_0$	~ 1/N
Random	Small d ~ log(N)	Short Tail Poisson e^{-k}	Small ~log(N)	~ 1/N
Scale-Free	Small d ~ log(N)	Long Tail $\sim k^{-\gamma}$	Large HUBS $\sim k^{1/(\gamma-1)}$	~

But is the real world ordered or random?

Power Laws in the Real World

These are a sure way to get a physicist over excited, and increasingly others too

- 2nd Order Phase Transitions
  (e.g. superconductors, superfluids,…)

- **Long range order** = no scale = physical insight
  Critical Phenomena – Renormalisation Group