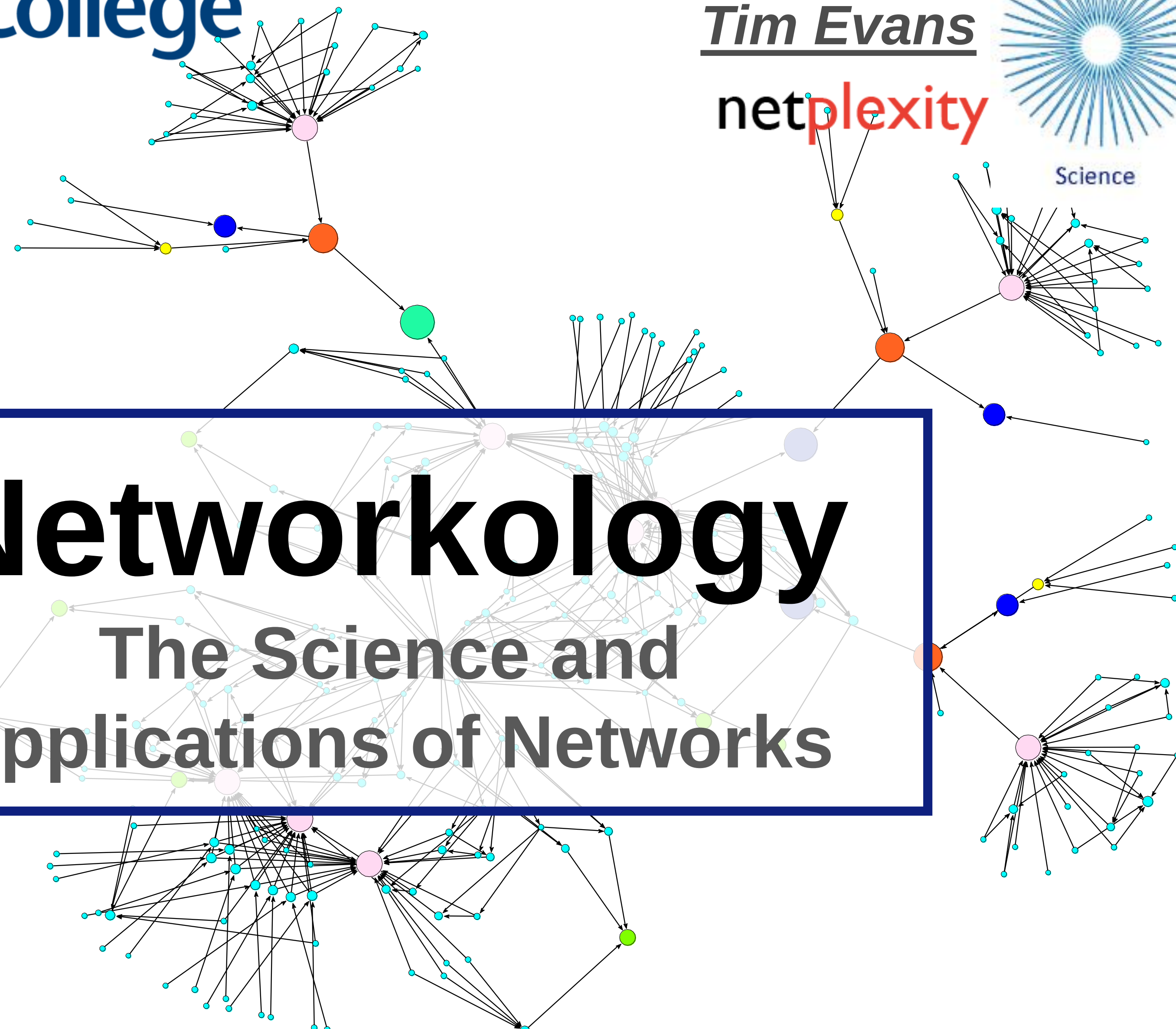
Imperial College London

Talk given at Science for Fiction,
https://davecl.wordpress.com/2016/04/28/science-for-fiction2016/
4th July2012

Tim Evans

netplexity

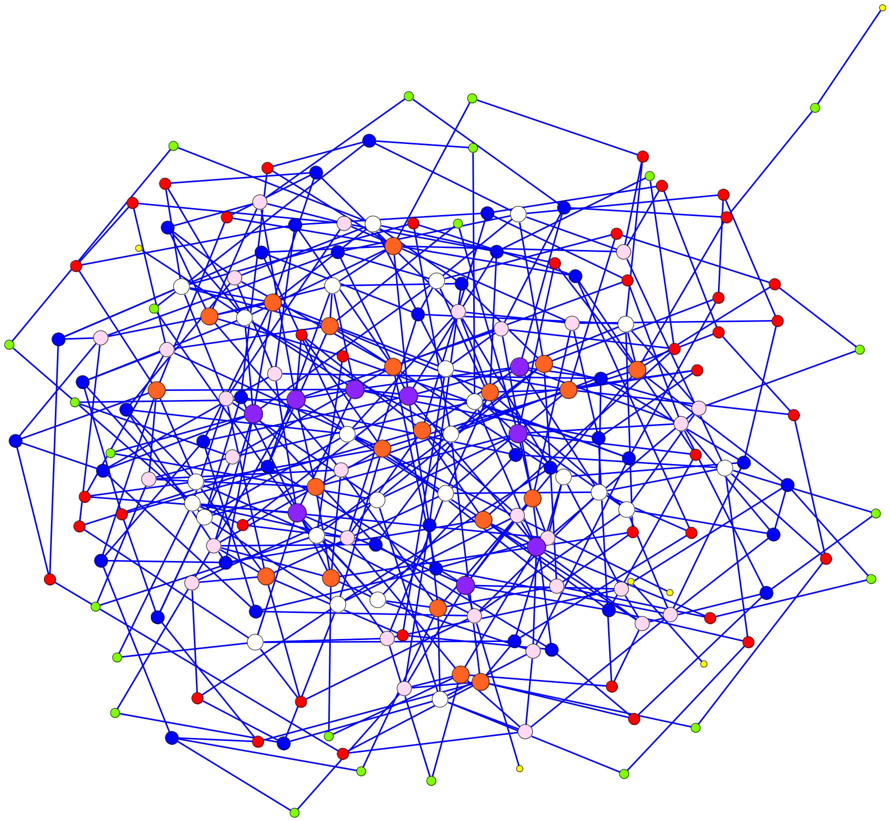
Complexity Science



Networkology

The Science and Applications of Networks

Slides available from
http://dx.doi.org/10.6084/m9.figshare.3467066

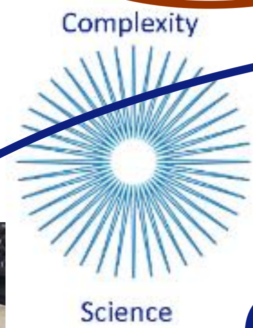


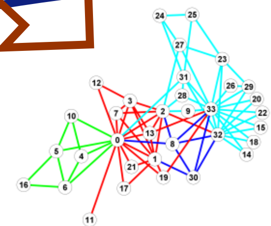
INTRODUCTION

Condensed Matter Physics

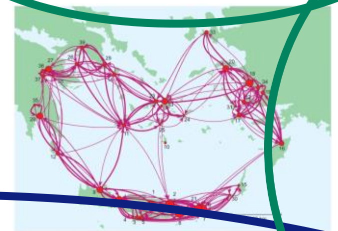
Particle Physics

Statistical Physics

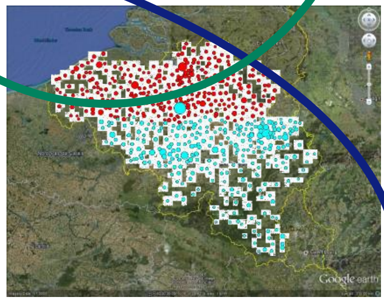


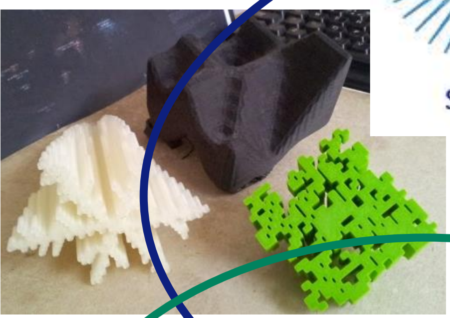


Archaeology



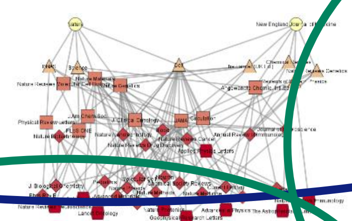
Spatial Networks





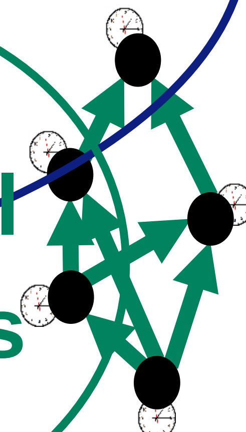
Complexity & Networks

Evolution of Complex Systems



Innovation, Bibliometrics

Temporal Networks

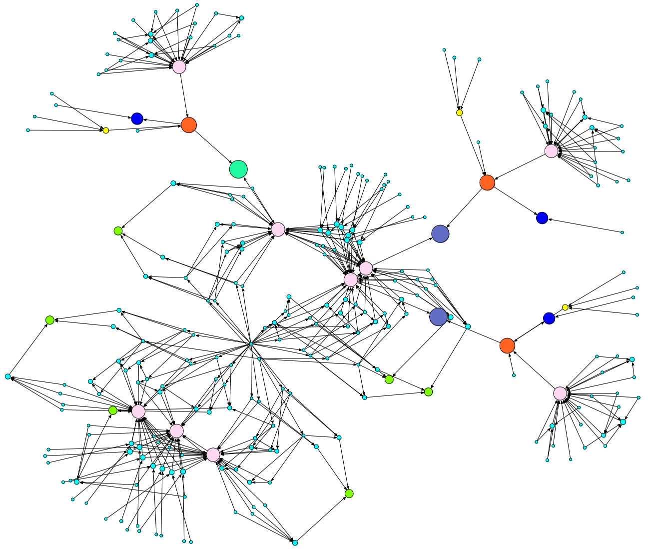


Complexity is

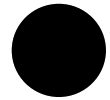
- Interactions occur defined at small, local scales
- Emergence of large scale phenomena

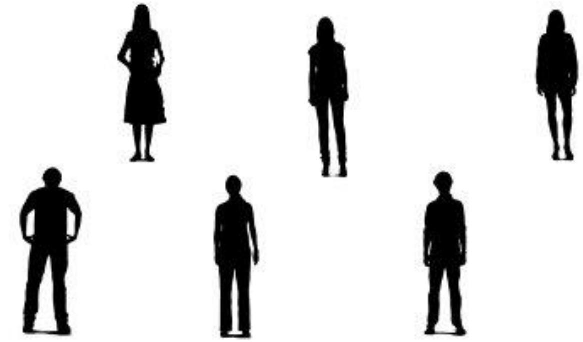
  - ➢ Just statistical mechanics applied to new problems?

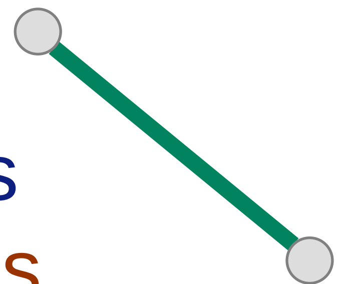
Networks are part of this wider complexity programme



Definition of a Network

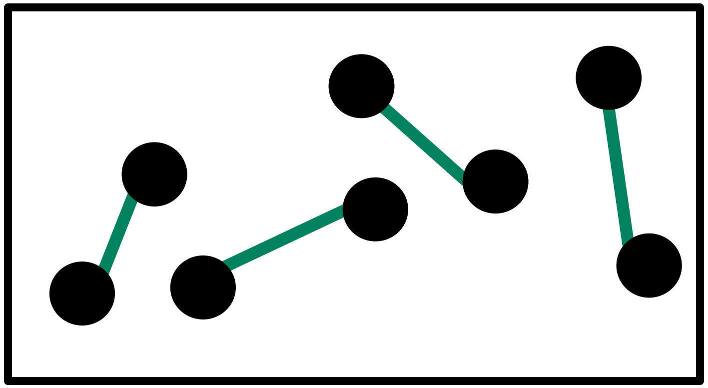
- A set of **nodes** 
  e.g. people



- A set of **edges** between nodes 
  e.g. friendships
  between people

No Need For Network Analysis?

Edges describe bilateral relationships between nodes



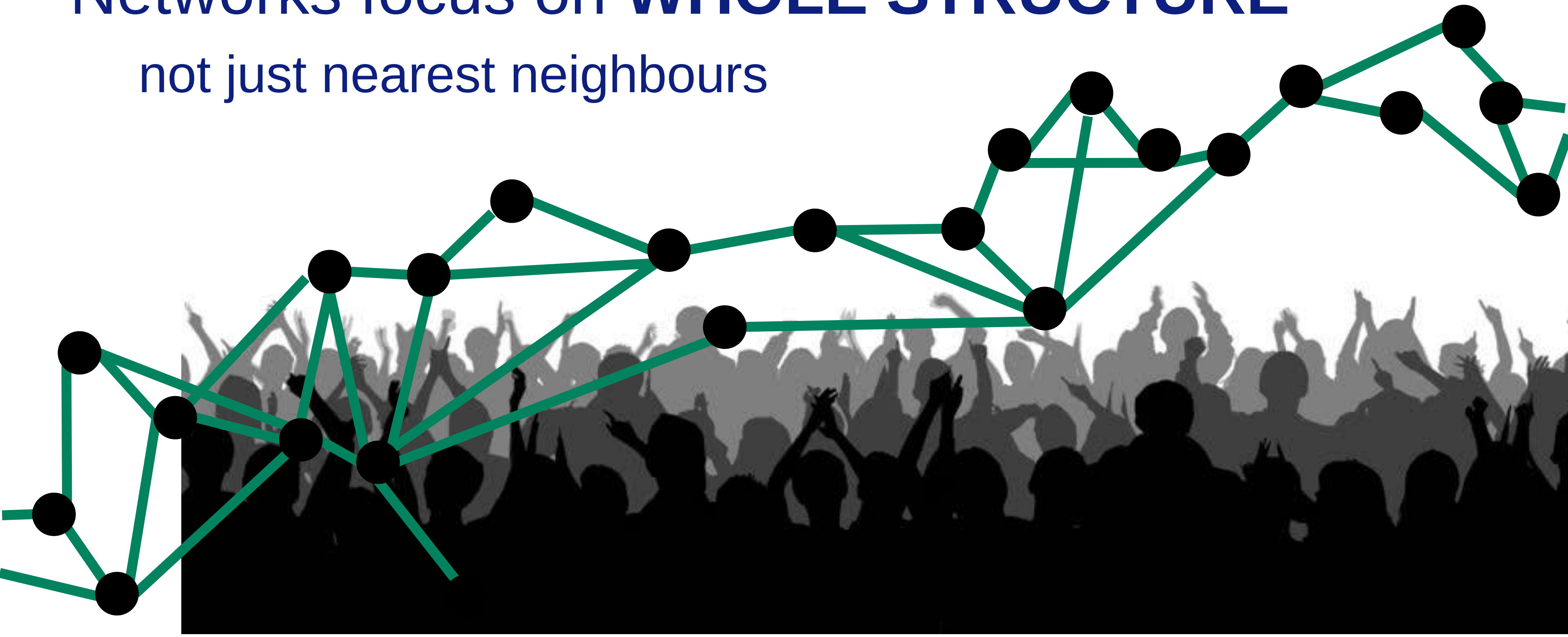
Just analyse the **statistics of these pairs** using usual statistical methods

Network Analysis

Adds new insights when large scales are relevant

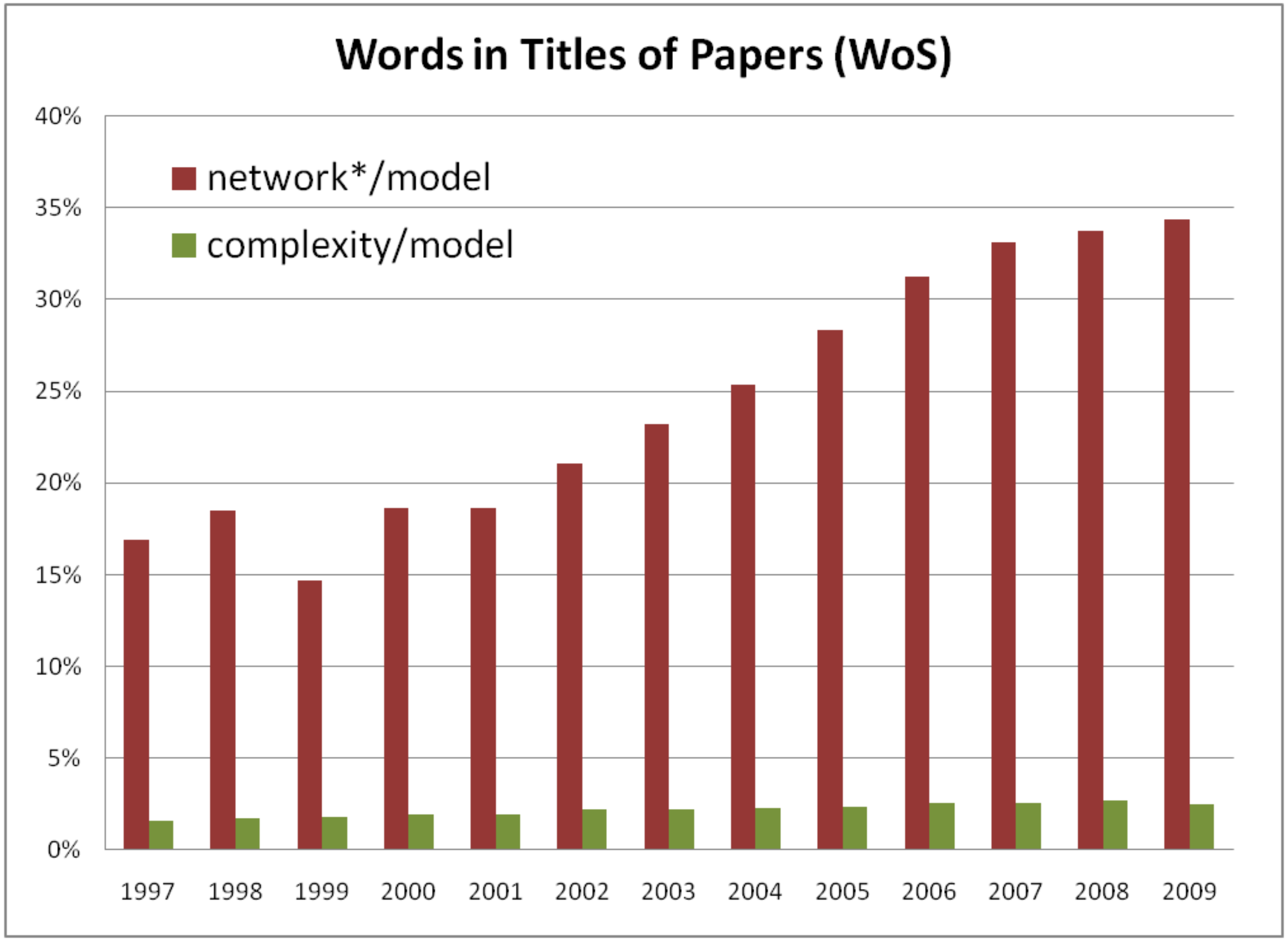
Networks focus on **WHOLE STRUCTURE**

not just nearest neighbours



[Watts & Strogatz, Nature 1998]

Explosion of interest since 1998

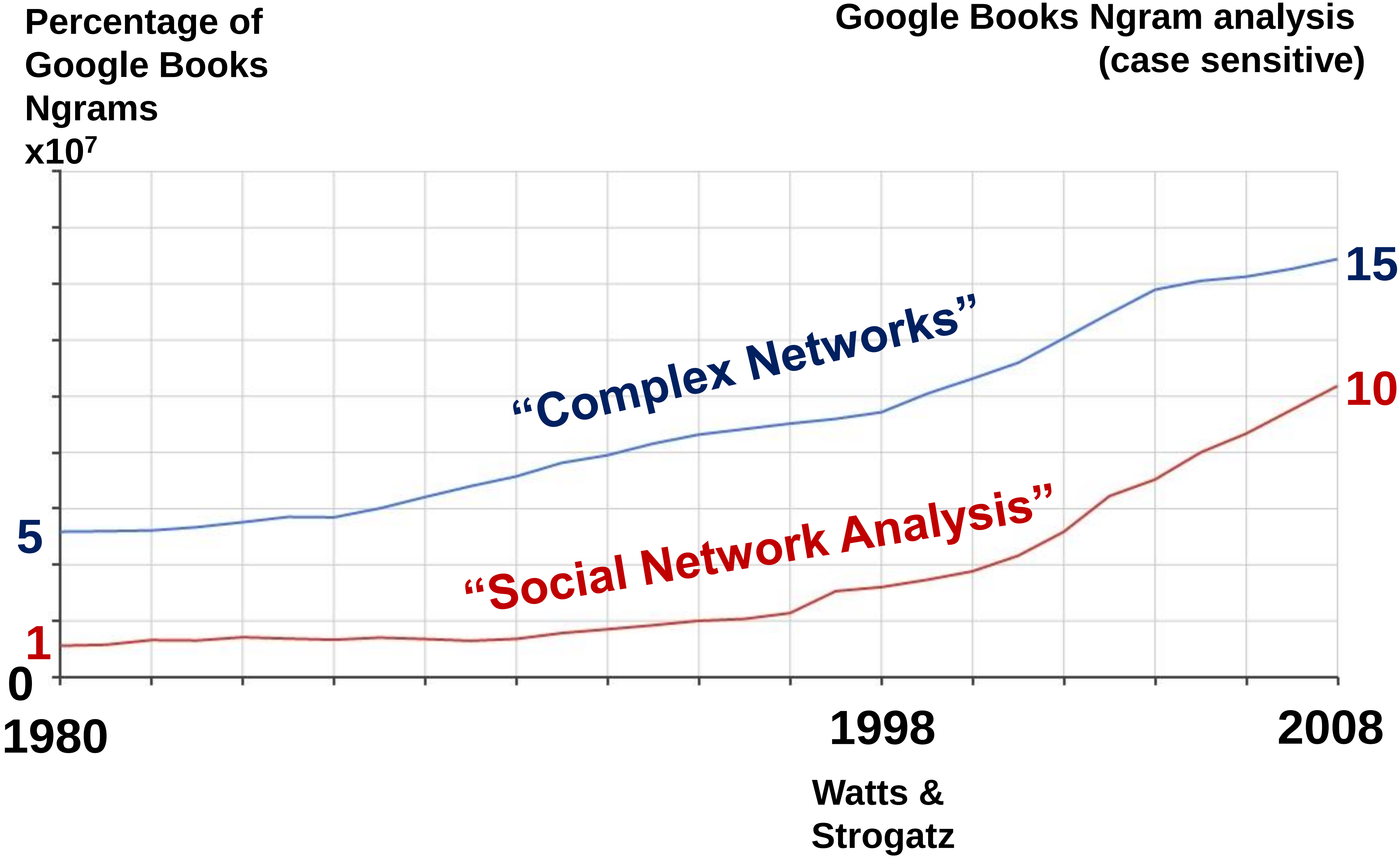


Fraction of papers with word starting "NETWORK" in title compared to number of papers with word "MODEL"

➡ WHY?

Explosion of interest

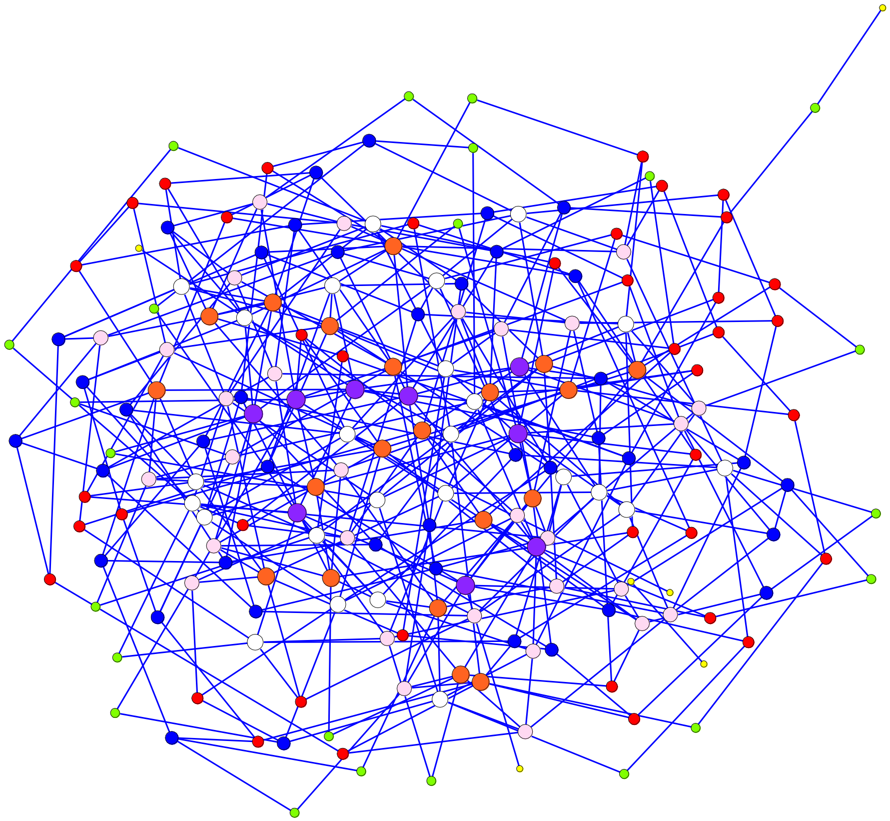
Google Books Ngram analysis (case sensitive)



Multidisciplinary Nature of Network Research

Field	*Since*
• Mathematics (Graph Theory, Dynamical Systems)	1930's
• Physics (Statistical Physics)	2000's
• Biology (Genes, Proteins, Disease Spread, Ecology)	2000's
• Computing (Web search and ranking algorithms)	1970's
• Economics (Knowledge Exchange in Markets)	1990's
• Geography (Transport Networks, City Sizes)	1960's
• Social Science (Social Networks)	1960's
• Archaeology (Trade Routes)	1970's

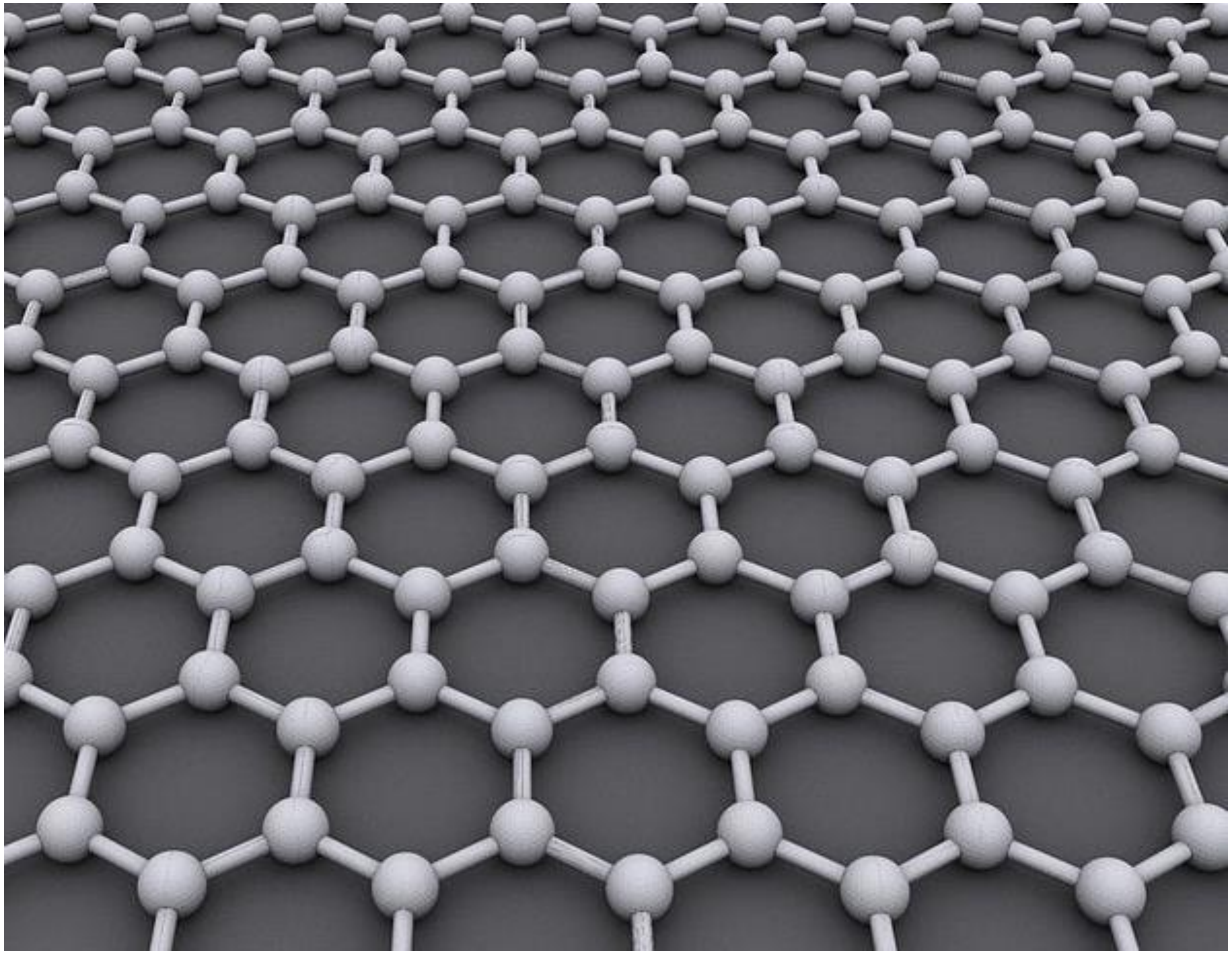
http://www.iscom.unimo.it/



EXAMPLES OF NETWORKS

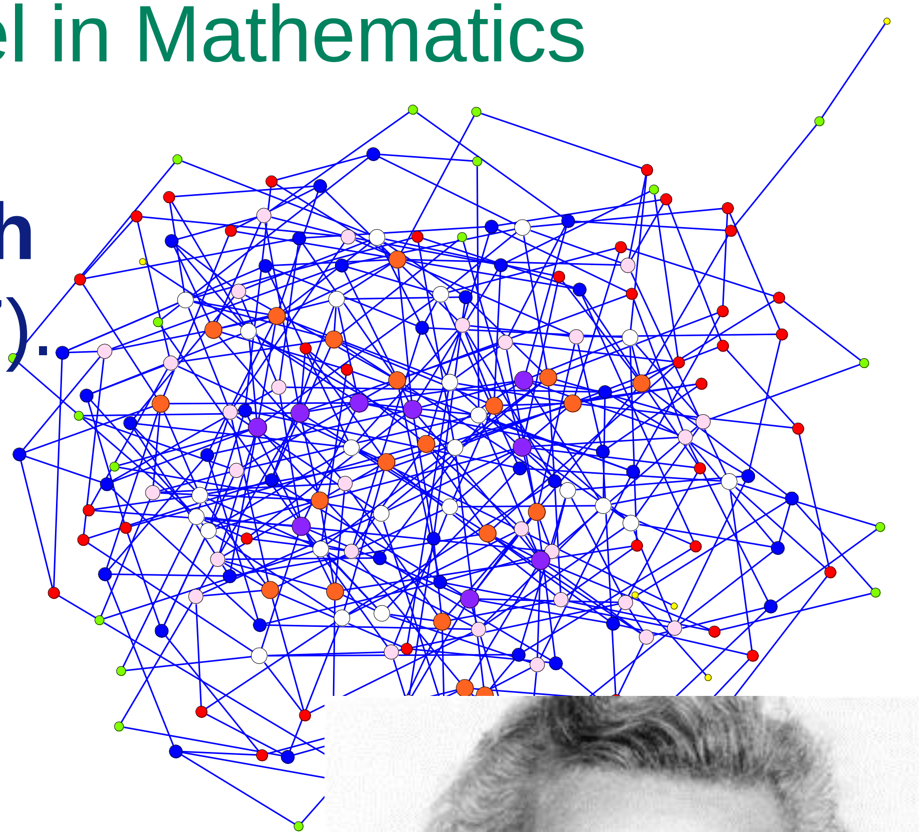
Traditional Physics networks

Graphene - regular lattice of atoms in two dimensions



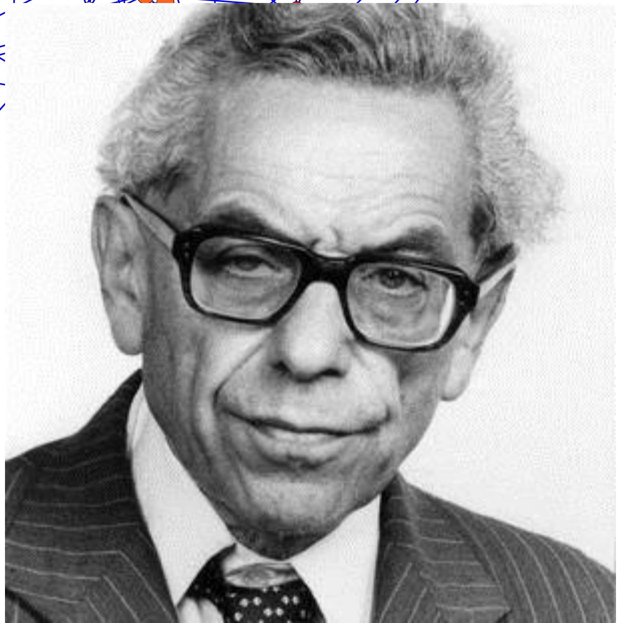
Traditional Network Model in Mathematics

Classical Random graph
of Erdős & Reyní (1957).



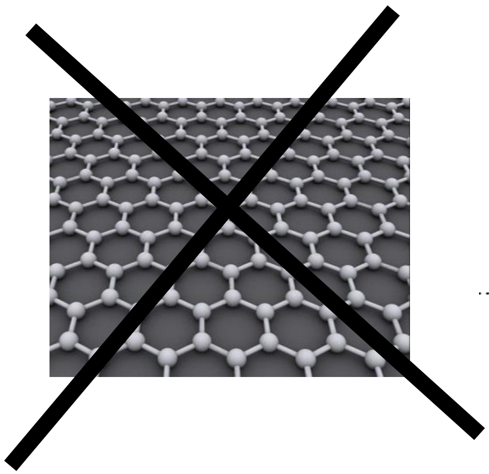
1. Take ***N*** nodes,
2. Throw down
   E edges at random

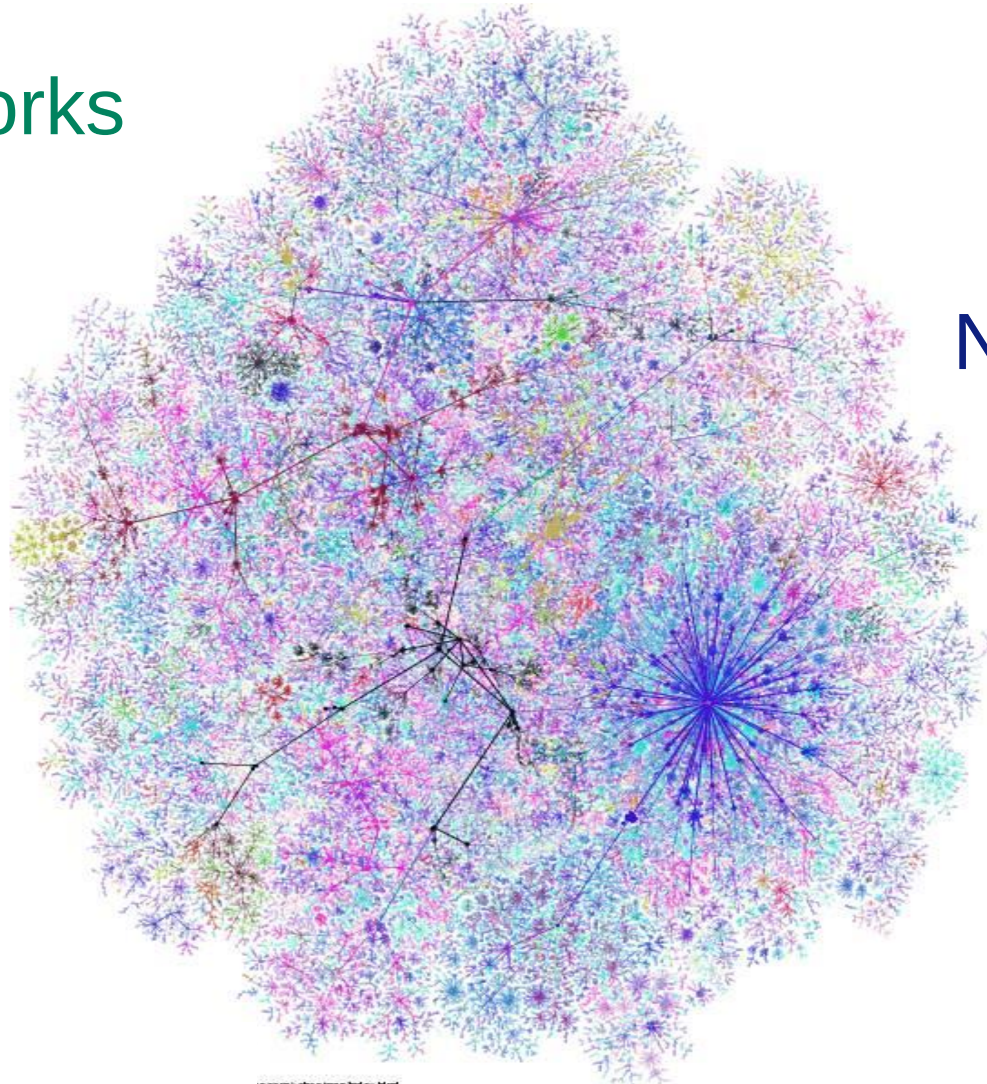
Paul Erdős
(1913-1996)
Itinerant, prolific, eccentric



Real Networks

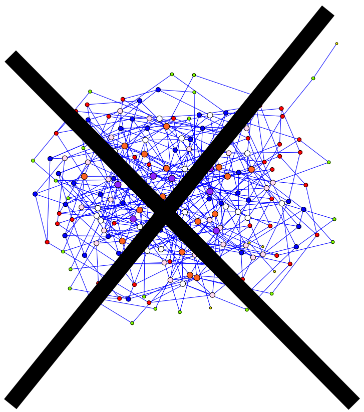
Neither perfectly regular like atoms





Router Level Map of Internet
[Burch & Cheswick, Internet Mapping Project]

Nor perfectly random like Classical Random graphs

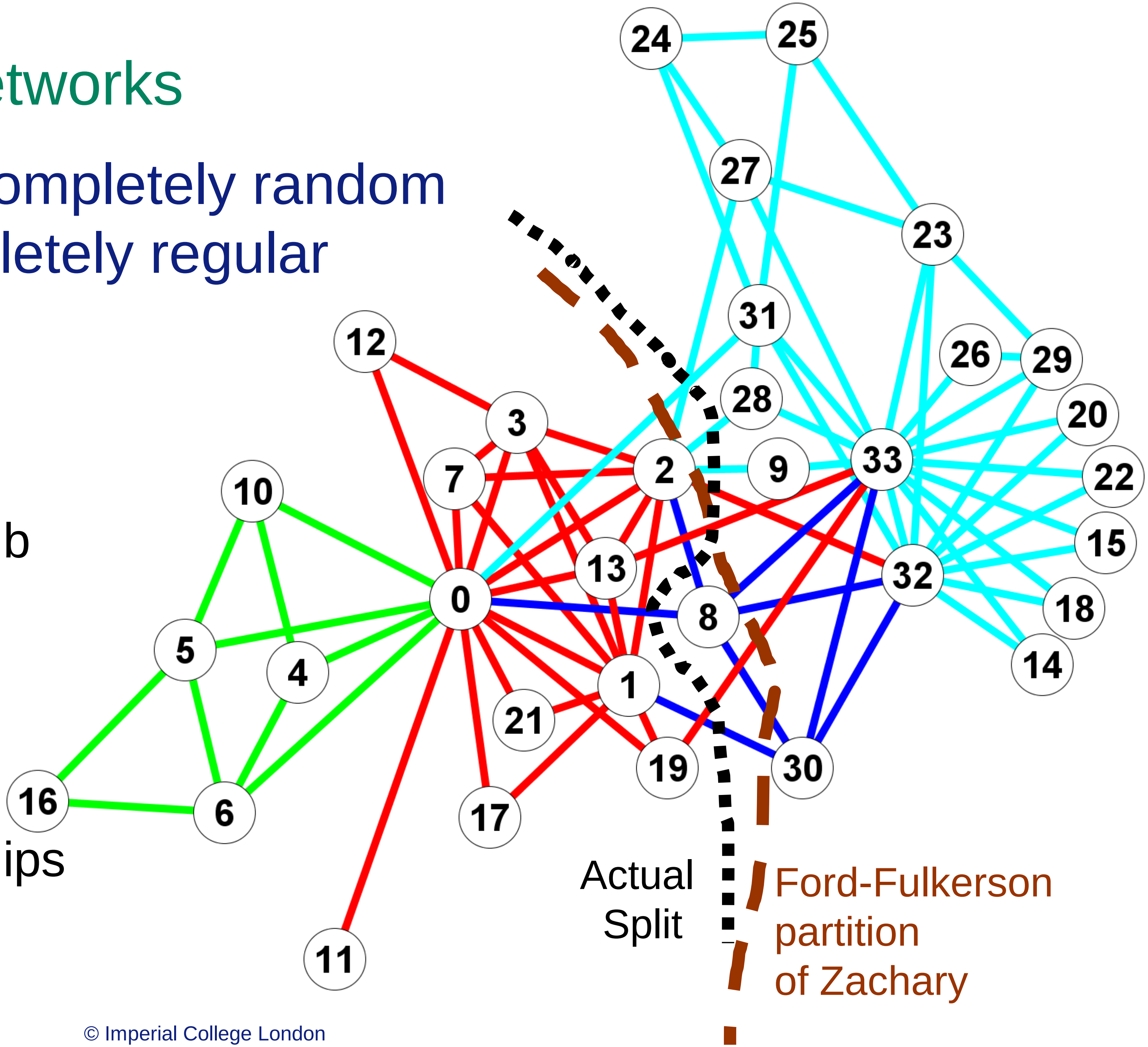


Social Networks

Neither completely random nor completely regular

Nodes = Karate Club Members

Edges = Observed Relationships ON or OFF



[Zachary 1977,TSE & Lambiotte, 2009]

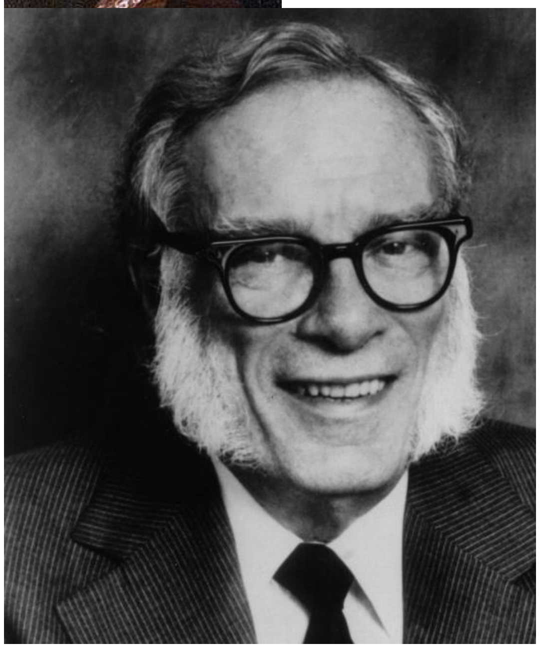
SocioPhysics?

- Thomas Hobbes
  ‘*Leviathan*’ (1651)

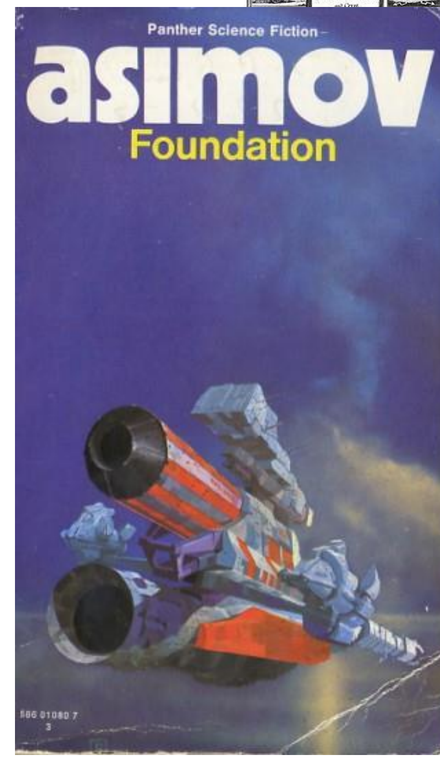
- Asimov, Foundation Trilogy (1940’s+)

- Philip Ball
  ‘Critical Mass’ (2004)

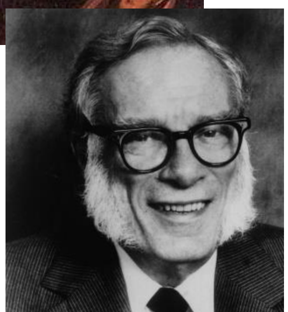
SocioPhysics?



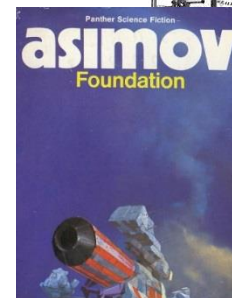
- Thomas Hobbes ‘Leviathan’ (1651)
- Asimov, Foundation Trilogy (1940’s+)
- Philip Ball ‘Critical Mass’ (2004)



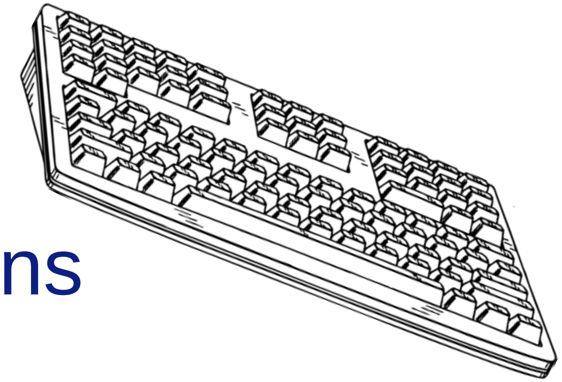
SocioPhysics?



- Thomas Hobbes ‘Leviathan’ (1651)
- Asimov, Foundation Trilogy (1940’s+)
- Philip Ball ‘Critical Mass’ (2004)

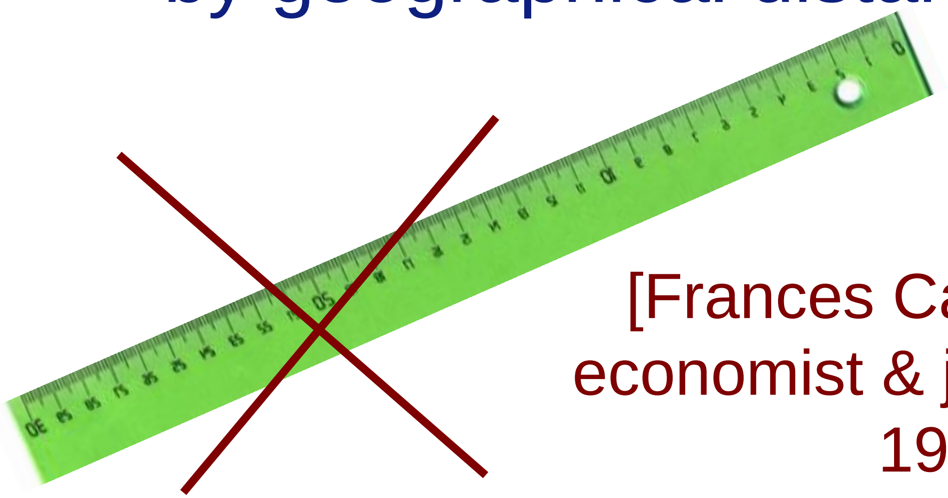


The “Death of Distance”



Modern electronic communications

Links no longer hindered by geographical distance



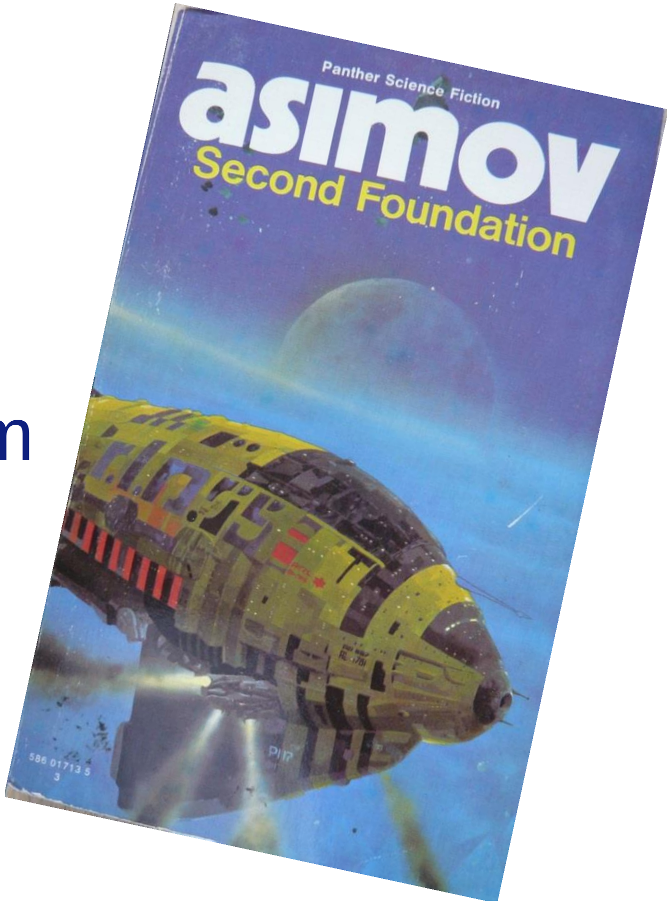
[Frances Cairncross, economist & journalist, 1995, 1997]

Distance & Space – Geographical or Social?

The two foundations lie at
"opposite ends of the galaxy"

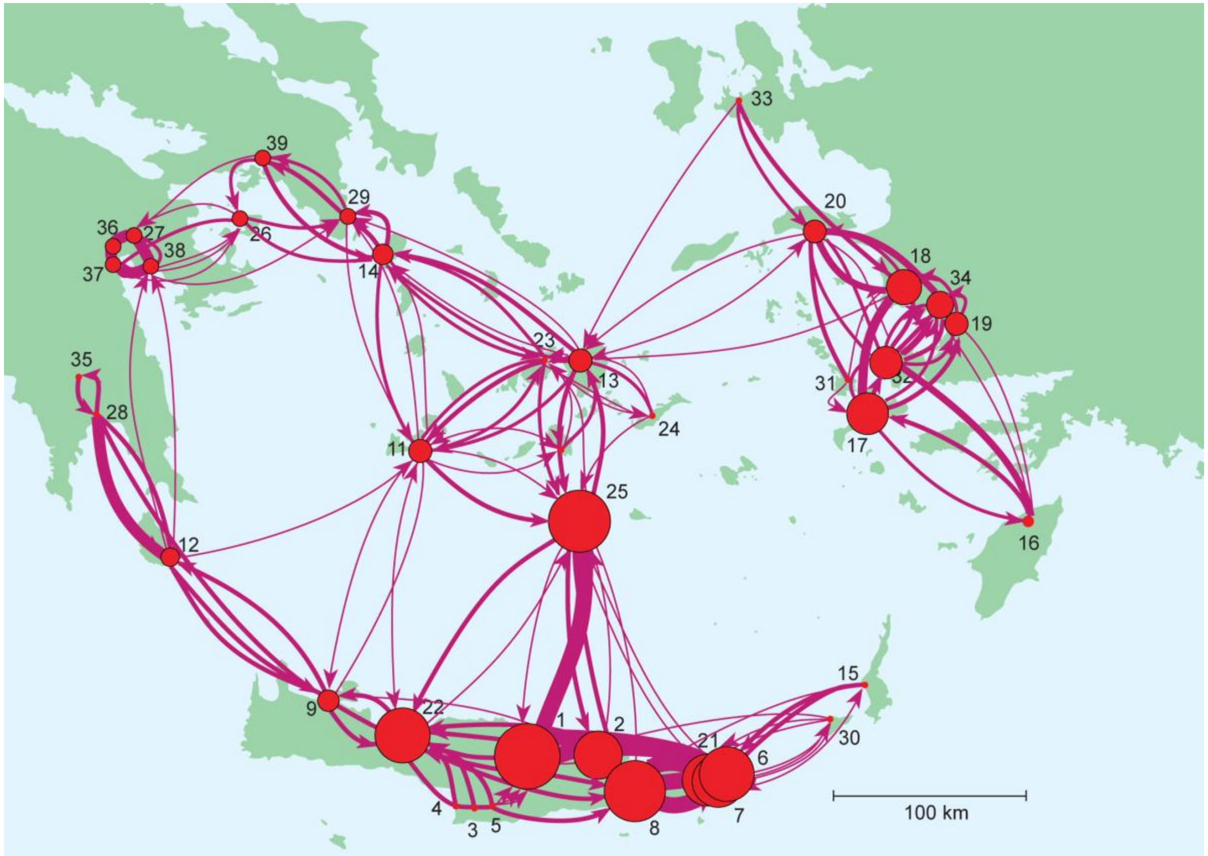
- Two ends of a spiral
  ⇒ **Physical separation**

- Two ends of the political spectrum
  ⇒ **Social separation**



Physical Networks: Minoan Aegean

Historical era geography dominates

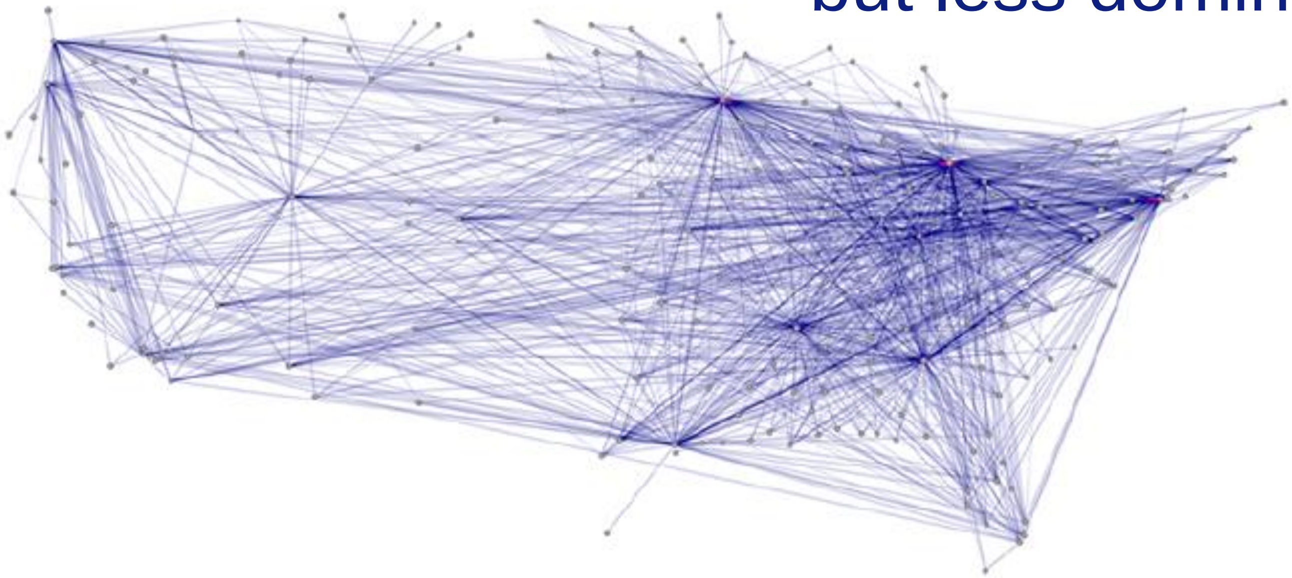


Stochastic network model representing links in Minoan Aegean (c1500BC)

[TSE, Knappett and Rivers, 2008,2009]

Transport – Airline Map

Modern era geography still plays a role
but less dominant



nodes = airports, geographical location

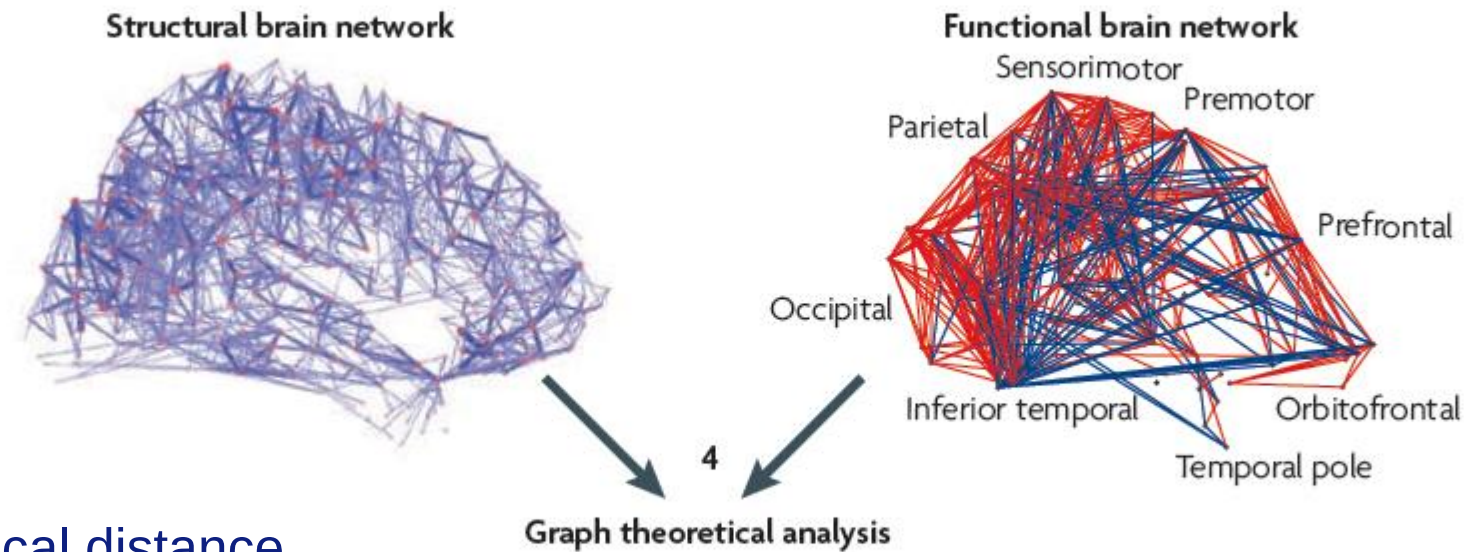
Edges= flights from/to, thickness~passengers

[Holten & van Wijk 2009]

Biological Networks:- Neuroscience

Edges from anatomy

Edges from activity



Physical distance still carries a cost

Network Analysis

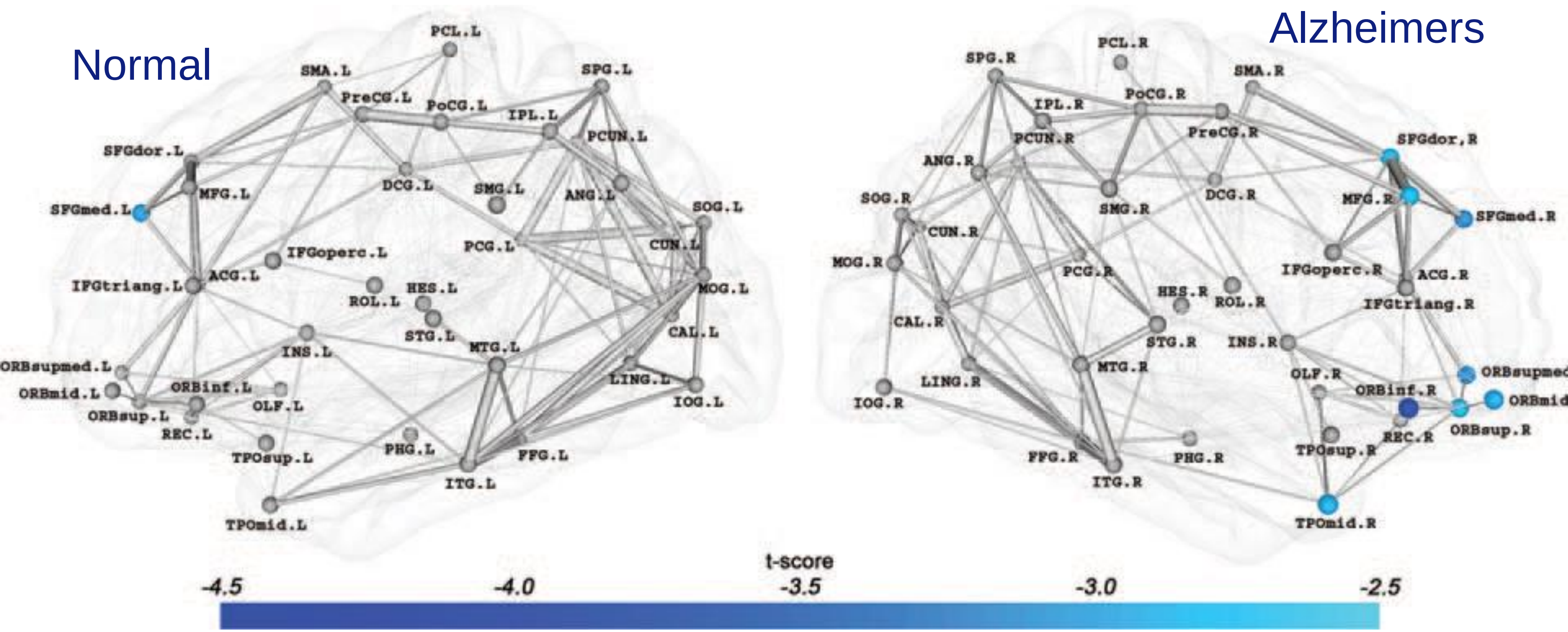
[Bullmore & Sporns, 2009]

Neuroscience Networks

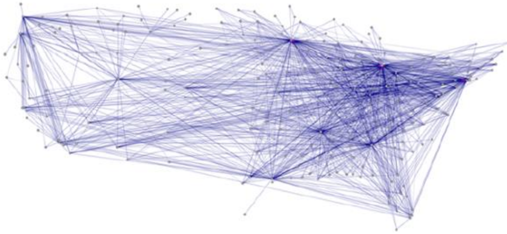
nodes = Regions,
Edges= Physical connections

Normal

Alzheimers



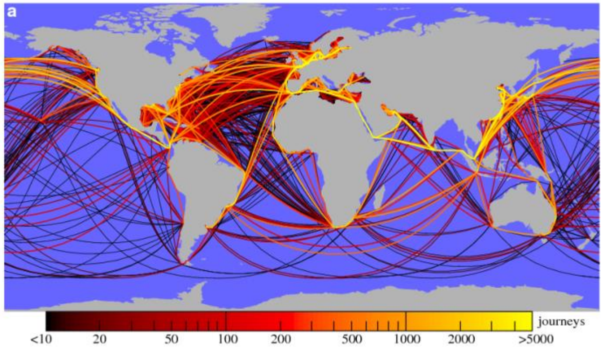
Lo et al, 2010

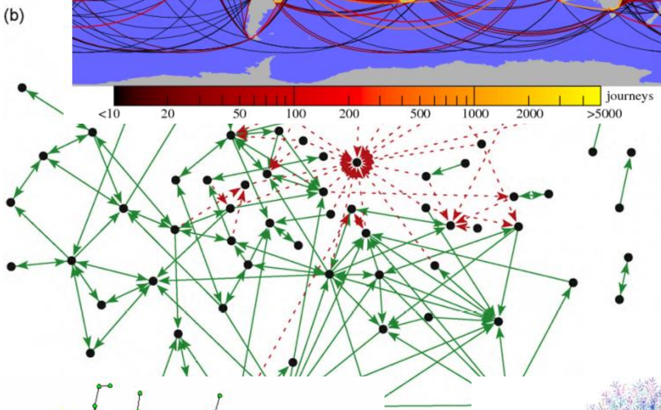


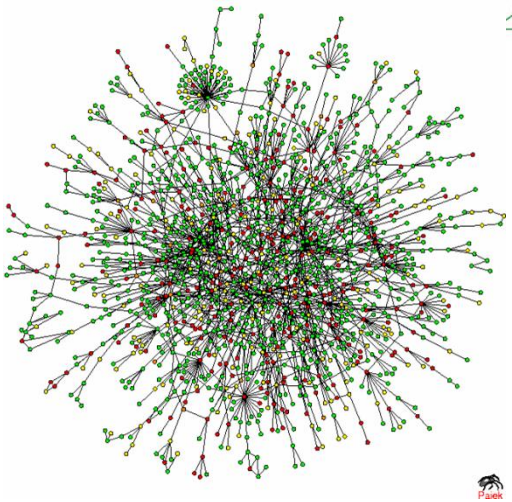
So many networks

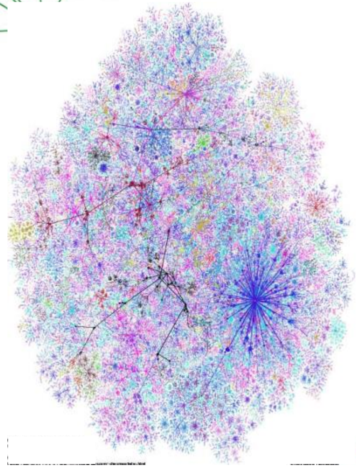
Networks are a useful way to describe many different data sets

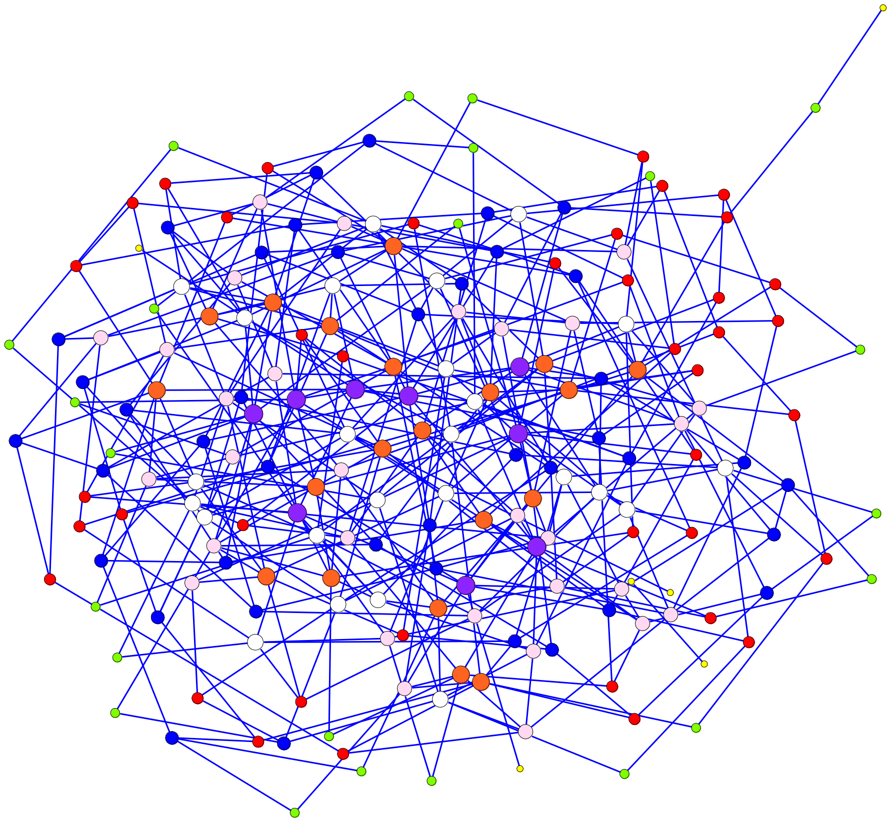
- Physical links/Hardware based
- Biological Networks
- Social Networks
- Information Networks











REPRESENTATIONS

Representations

- Data often has a `natural' network

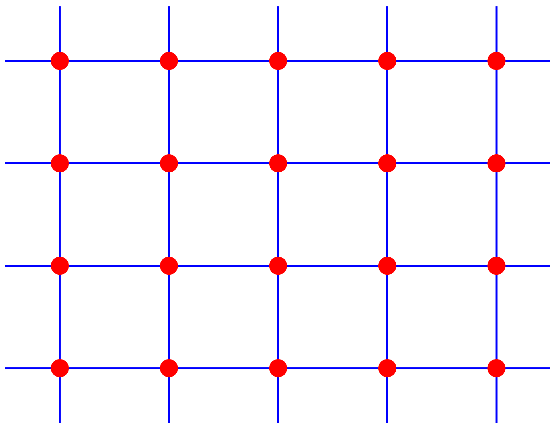
- There is no one way to view this natural network $\Rightarrow$ ***visualisation***

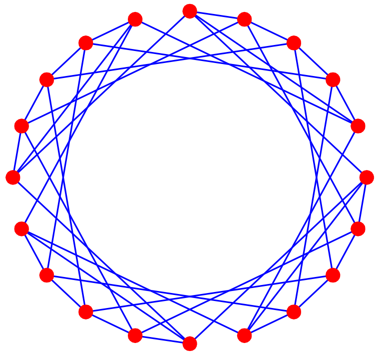
- There are always many different networks representing the same data

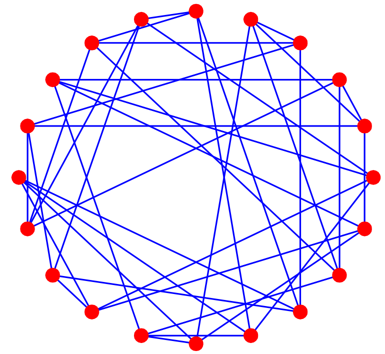
[Evans, 2004]

Visualisation – Topology vs Space

In a network the location of a node is defined only by its neighbours $\Rightarrow$ *topology*







Identical Networks

Periodic Lattice

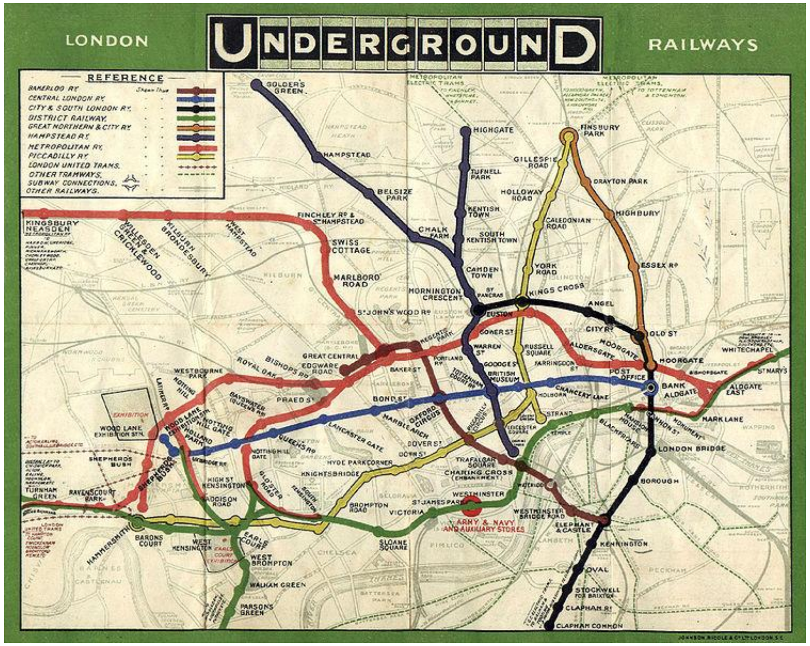
N=20, E=40

Same network with nodes arranged in regular order.

Same network with nodes arranged in random order

Choose visualisation to reveal spatial structure

Visualisation – Space before Topology



Tube lines laid out in geographical space (1908)

Visualisation – Topology over Space

Modern map preserves relationships between stations, not geography

[Derived from Beck 1931]

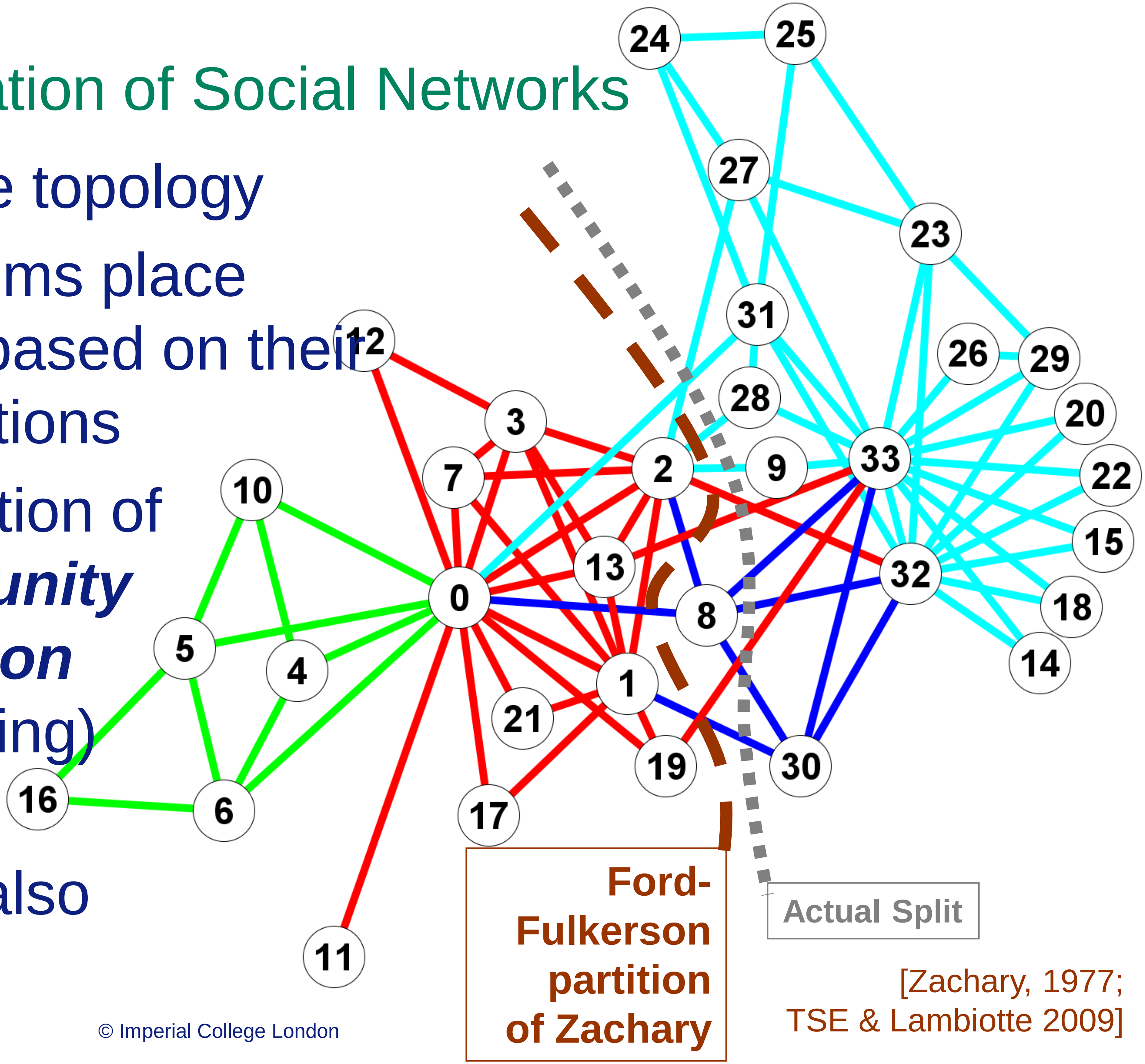
Visualisation – no space, only topology

Most networks are not embedded in any space e.g. social networks

Visualisation of Social Networks

Only have topology

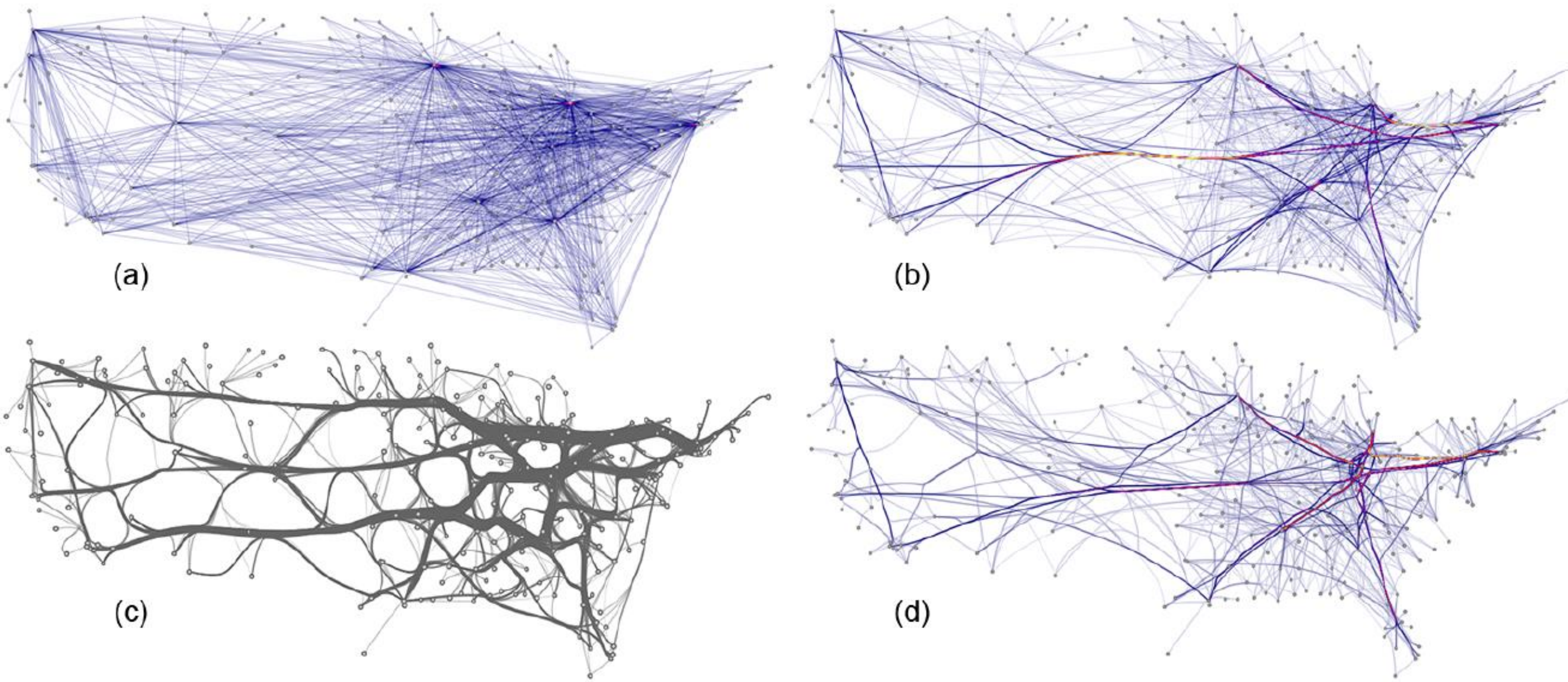
- Algorithms place nodes based on their connections
- Application of ***Community Detection*** (clustering)
- Edge colour also



[Zachary, 1977; TSE & Lambiotte 2009]

Visualisation

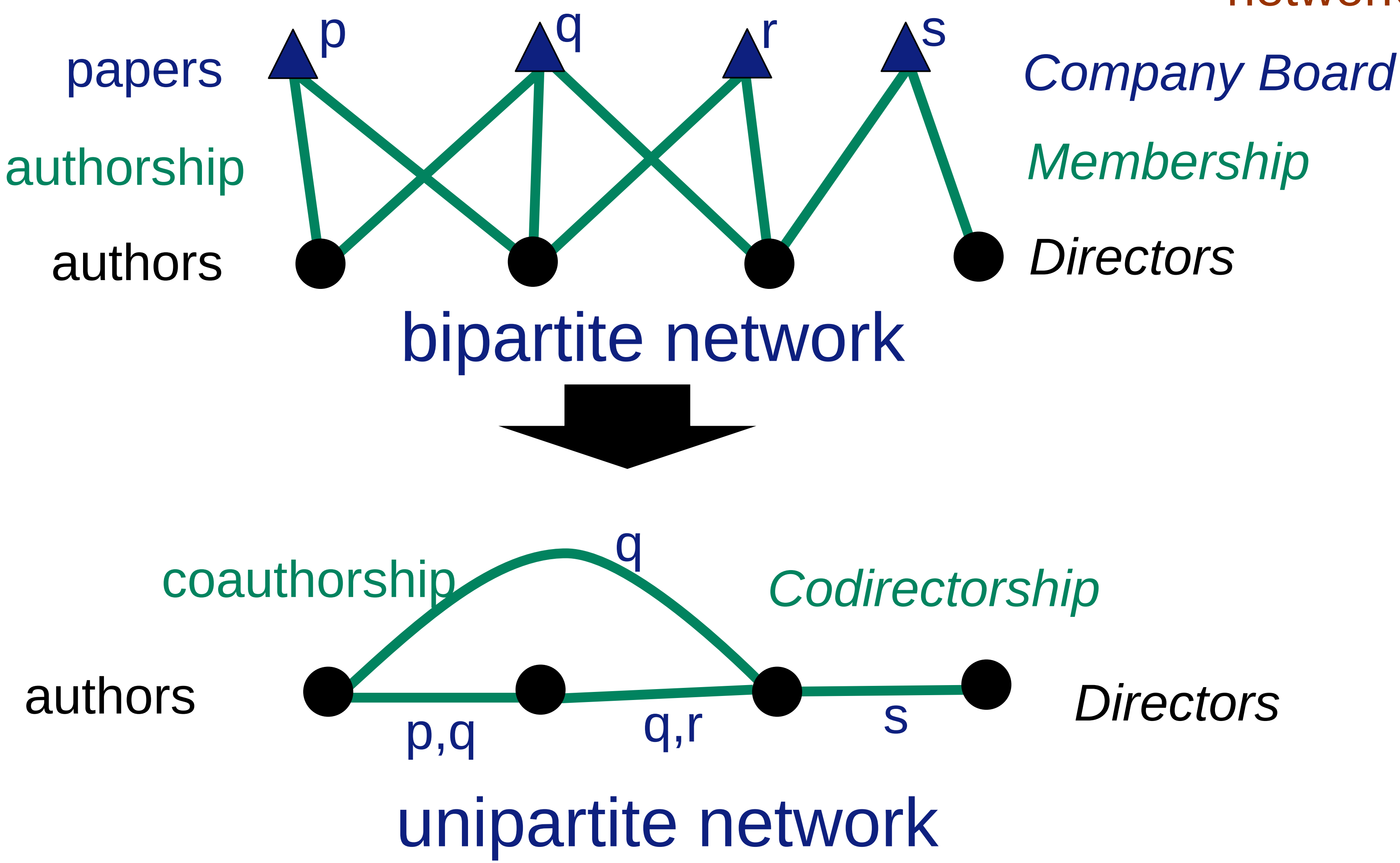
Choosing the right visualisation is a powerful practical tool, and its not just the nodes ...



e.g. Edge bundling for air network [Holten & van Wijk, 2009]

Same Data, Many Networks

e.g. Co-membership networks



Network of Streets

Some networks are embedded in geographical space.

This street network is embedded in real two-dimensional space

Same Data, Many Networks - Line Graphs

nodes = Intersections

Edges = Streets

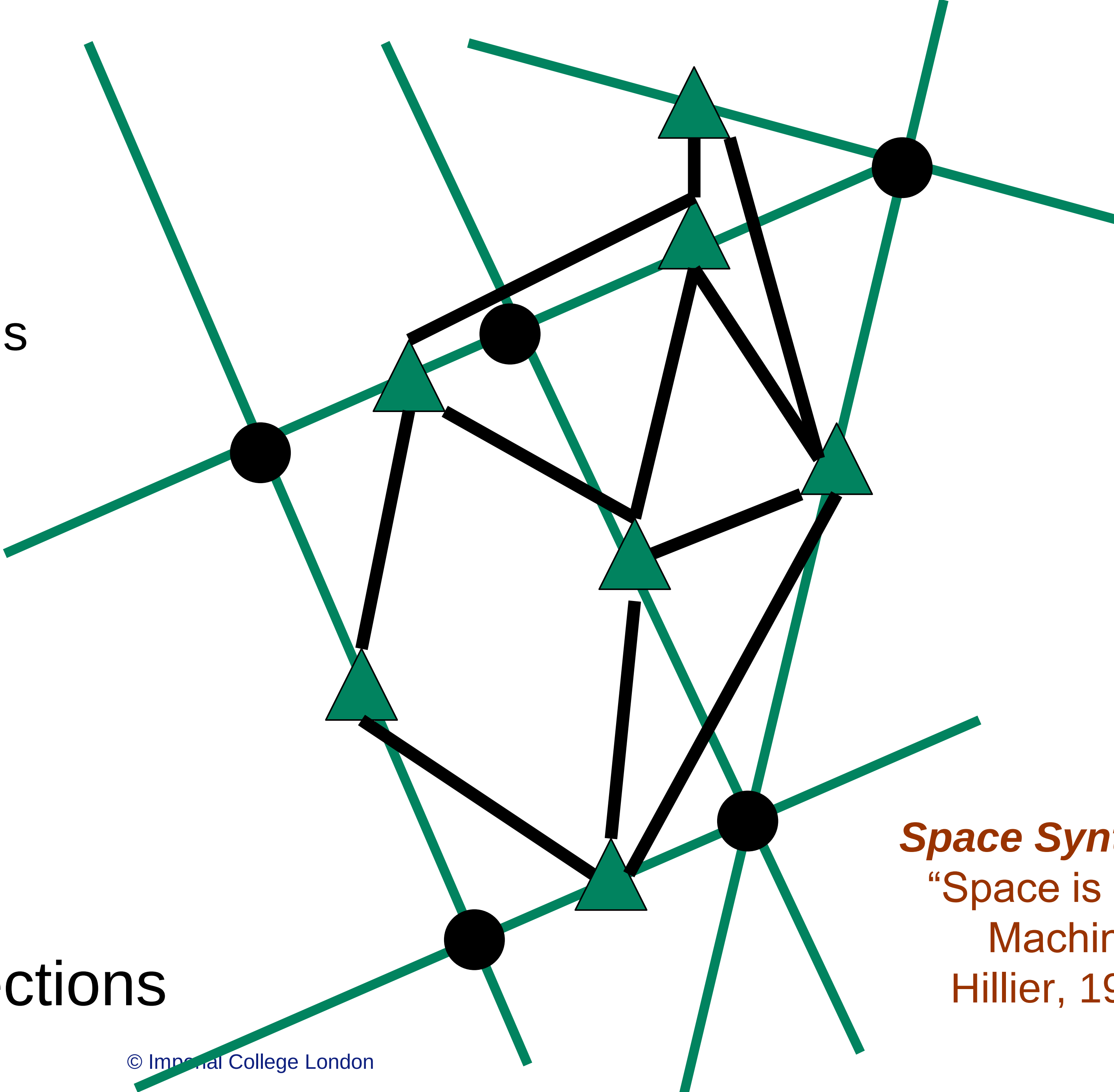
Line Graphs

nodes = Intersections

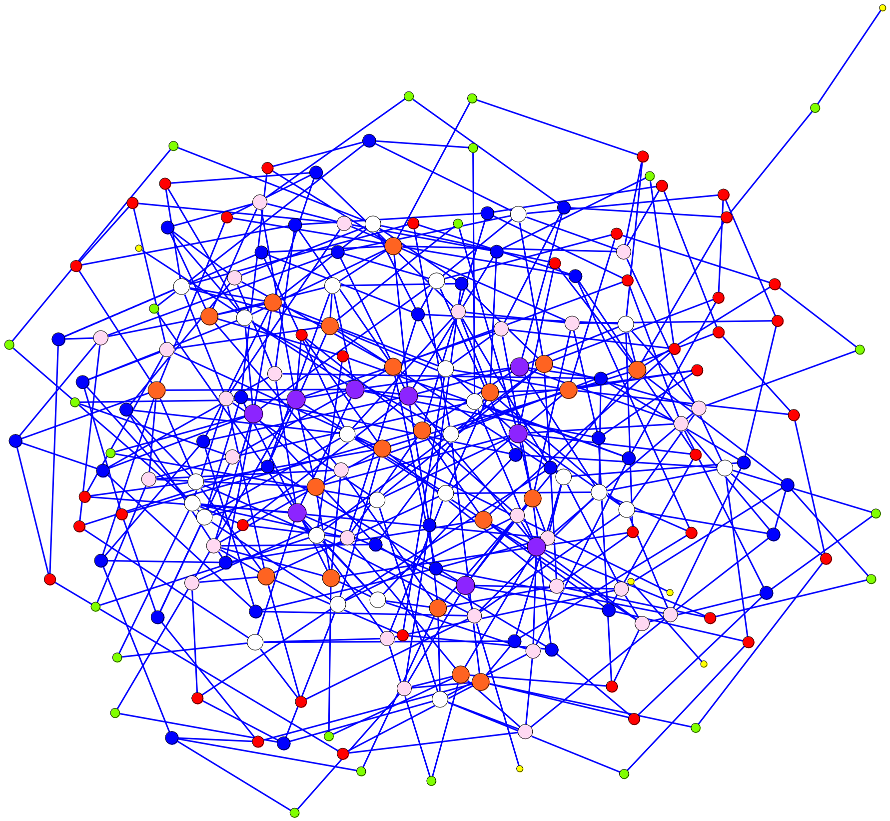
Edges = Streets

nodes = Streets

Edges = Intersections



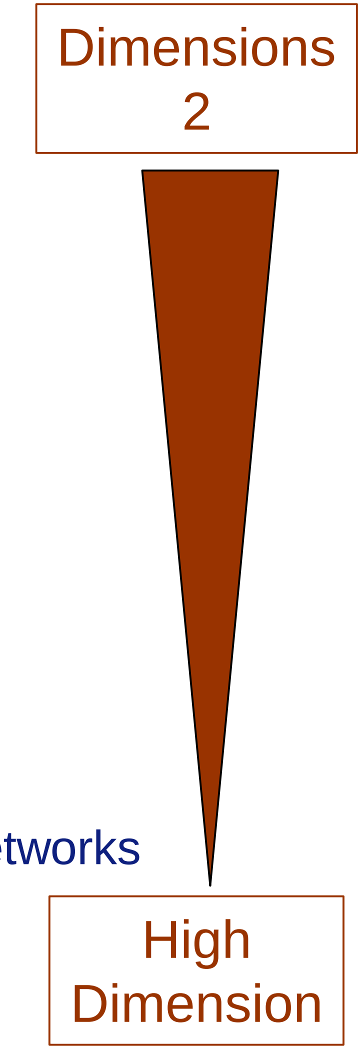
Space Syntax
"Space is the Machine", Hillier, 1996



THE DIMENSION OF NETWORKS

Types of Network – by application

- Physical links/Hardware based
  - telephone links, internet hardware, power lines, transport
- Biological Networks
  - neural, biochemical, protein, ecological
- Social Networks
  - Questionaires, observation, electronic social networks
- Information Networks
  - academic papers, patents, keywords, web pages, artefact networks



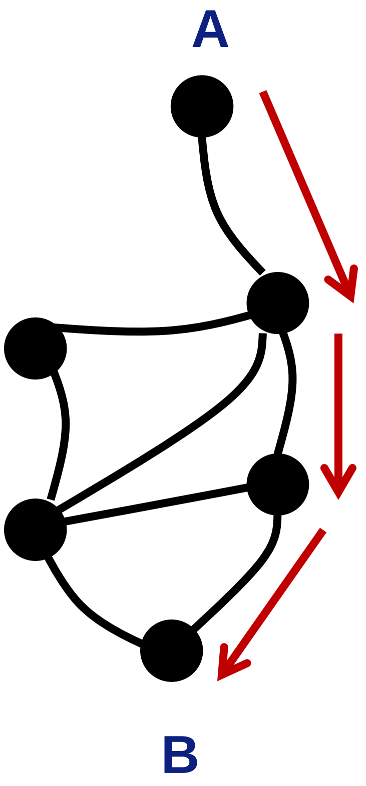
Death of Distance?

[Cairncross 1997]

Length of a Path

The length of a path between two nodes is the number of edges in that path

e.g. The shortest path from A to B has length 3.
There are two such paths of length 3 from A to B and many longer ones.



Dimension

For regular lattice in ***d*** dimensions we find that the shortest path between two points scales with the number of nodes ***N*** as

$$\mathrm{Volume} \sim N = (L)^d$$

Shortest Path Length

$$L_{\min} \propto (N)^{1/d}$$

The length of shortest path between two nodes is the number of edges traversed.

Dimension of Random Graphs

For random graphs we find that the shortest path between two points scales with the number of nodes ***N*** as

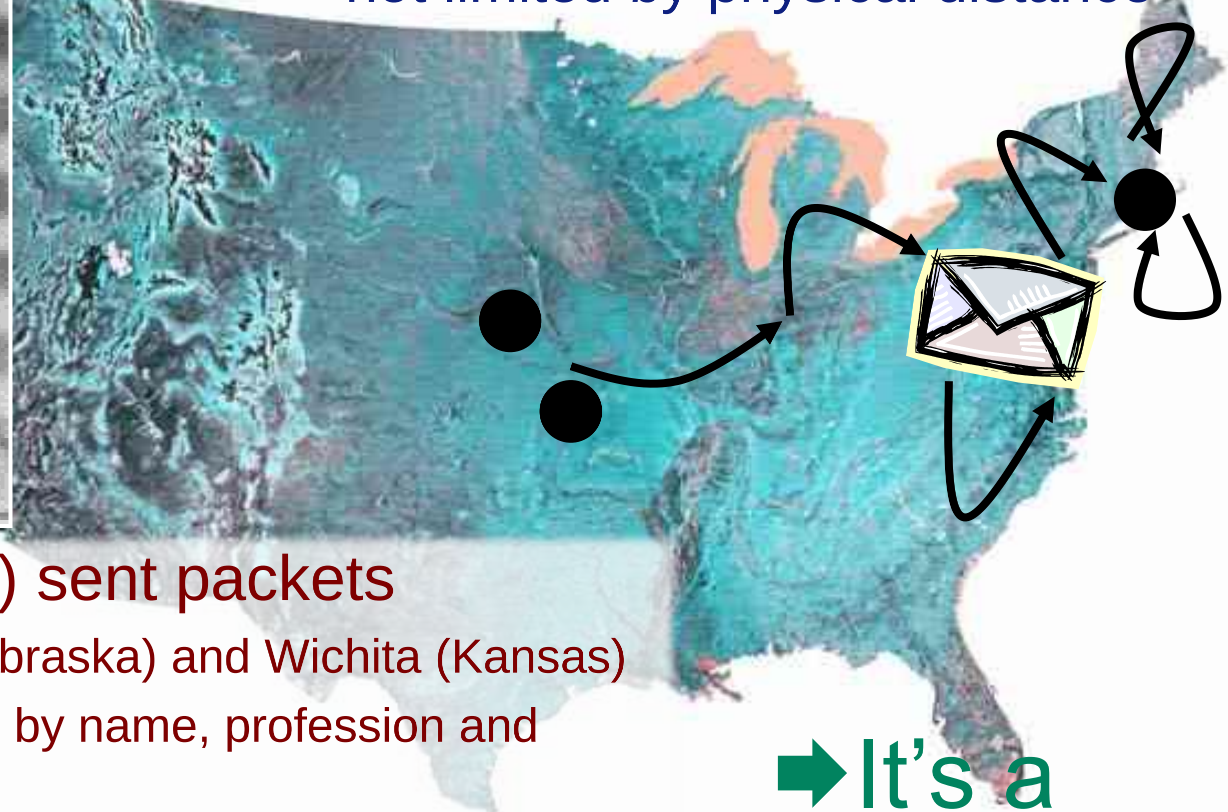
$$L_{\min} \propto \log(N)$$

$$\approx \lim_{d\to\infty}\left(N^{1/d}\right)$$

Random graphs look like infinite dimensional networks

Social Networks and Small Worlds

- node = Person
- Edges = Friendship
  not limited by physical distance



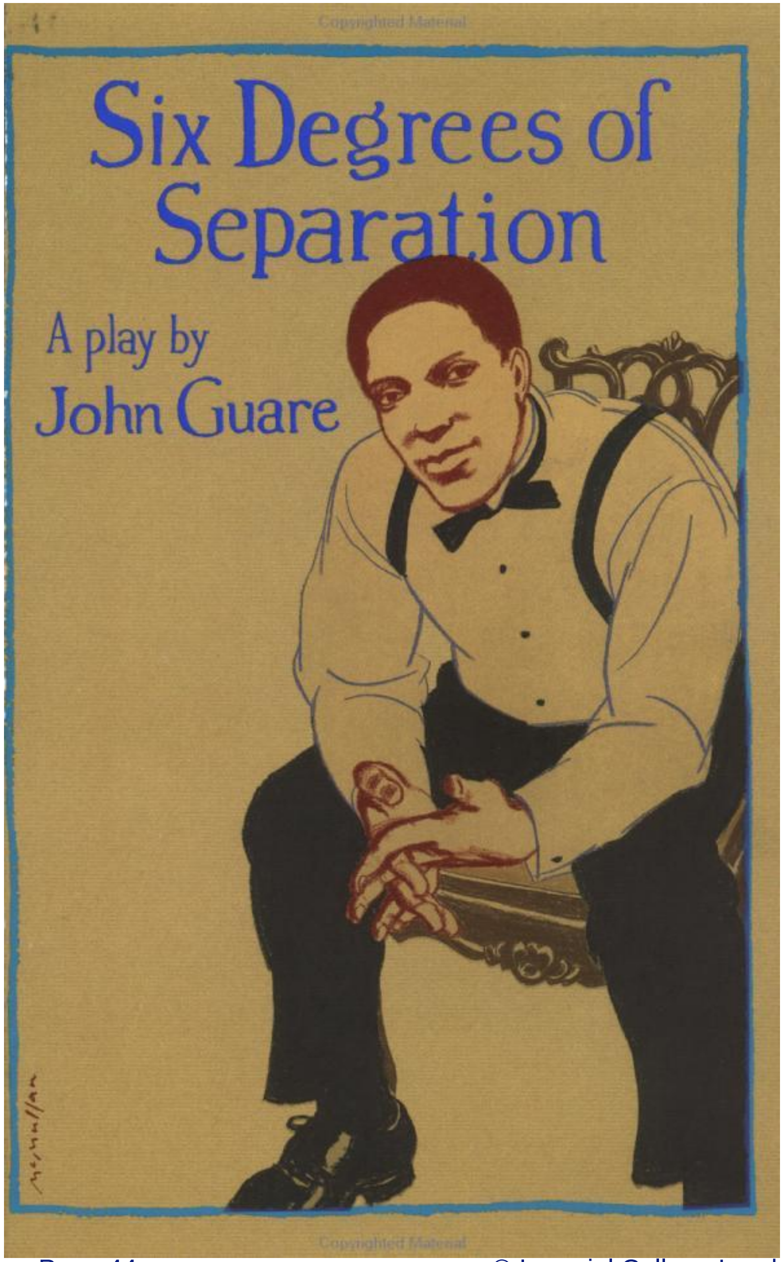
Stanley Milgram (1967) sent packets

- From people in Omaha (Nebraska) and Wichita (Kansas)
- To Cambridge MA specified by name, profession and rough location
- Only swapped between people on first name terms
- Packets which arrived averaged **FIVE intermediaries**

➡ It's a

Small World

Six Degrees of Separation



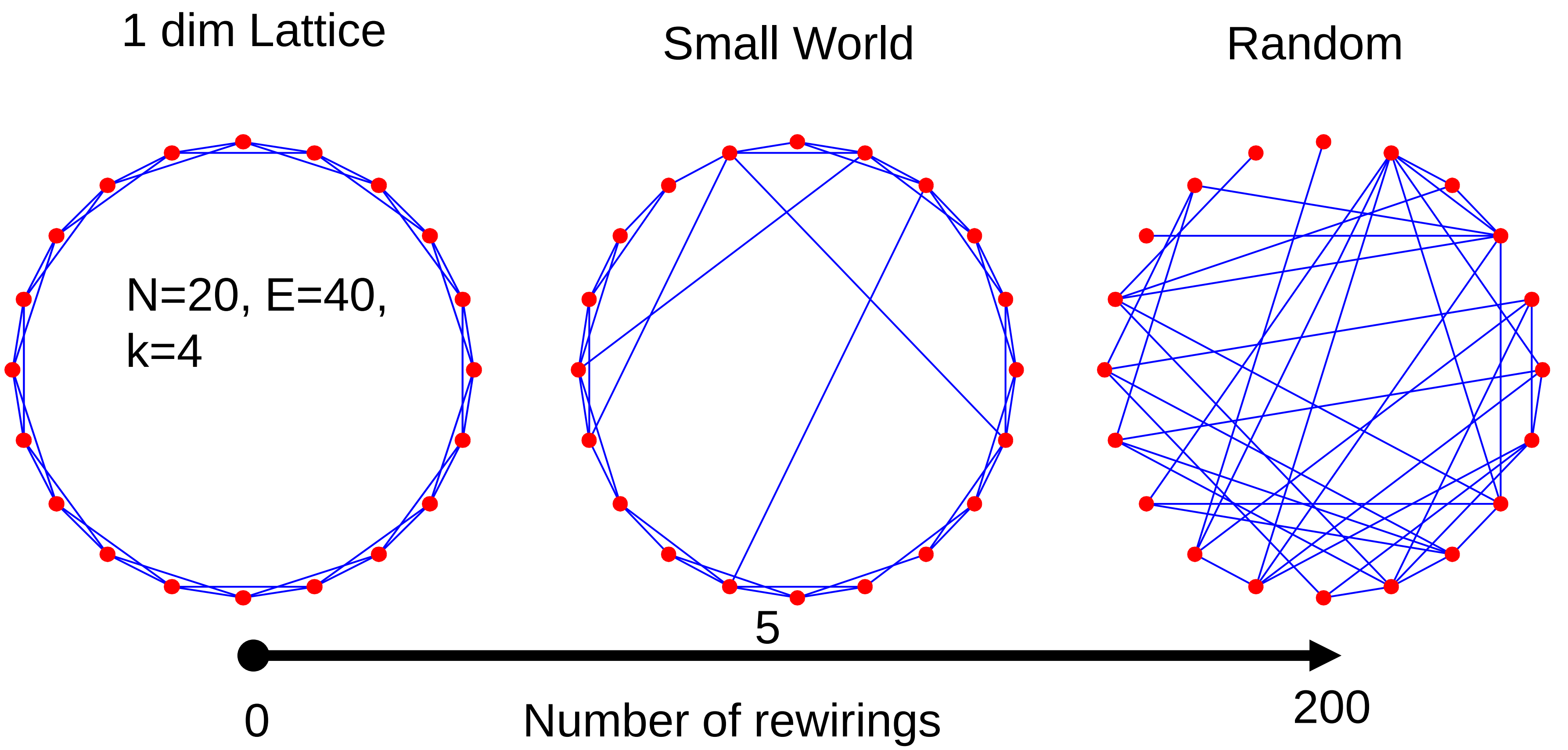
"I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation."

Six Degrees of Separation, John Guare (1990)

https://www.youtube.com/watch?v=NJQ6VRixNfY

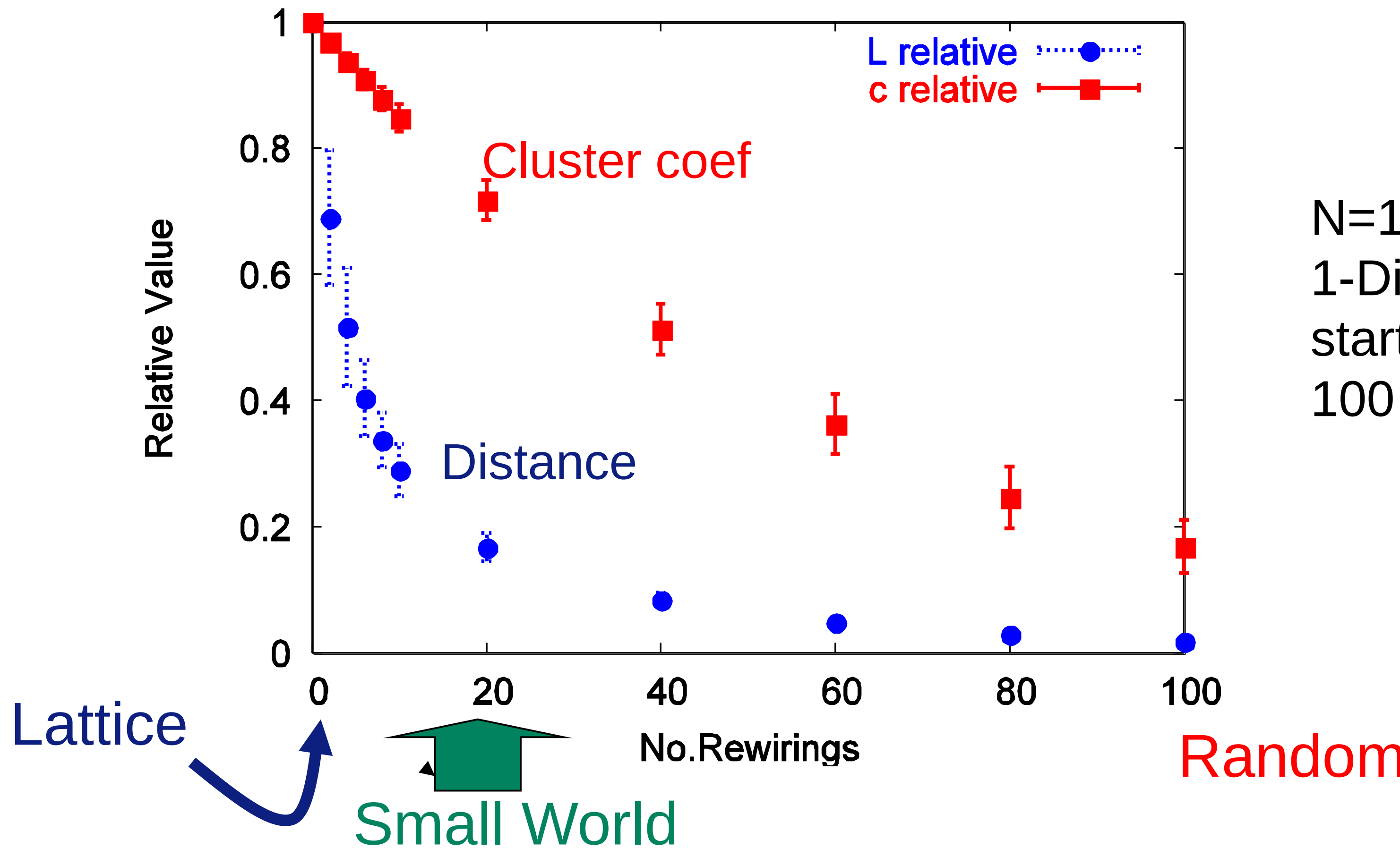
Watts and Strogatz's Small World Network (1998)

Start with regular lattice, pick random edge and move it to link two new nodes chosen at random.



Clustering and Length Scale in WS network

Find **Small World networks** with short distances of random network, local structure (clustering) like a lattice

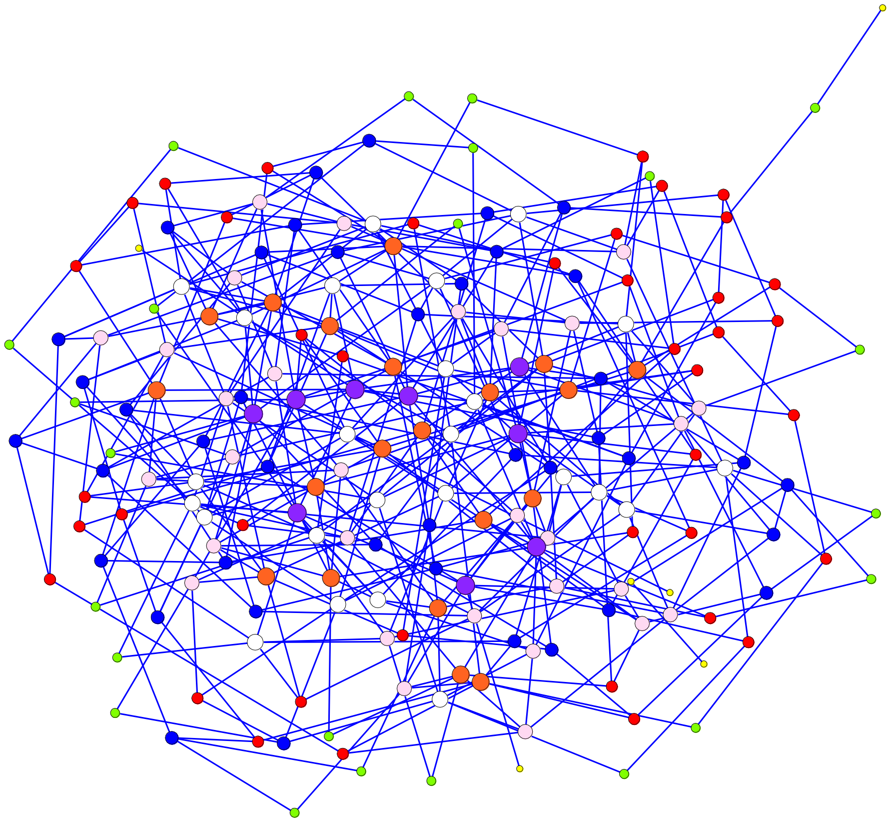


N=100, k=4,
1-Dim lattice
start,
100 runs

Real Network Summary

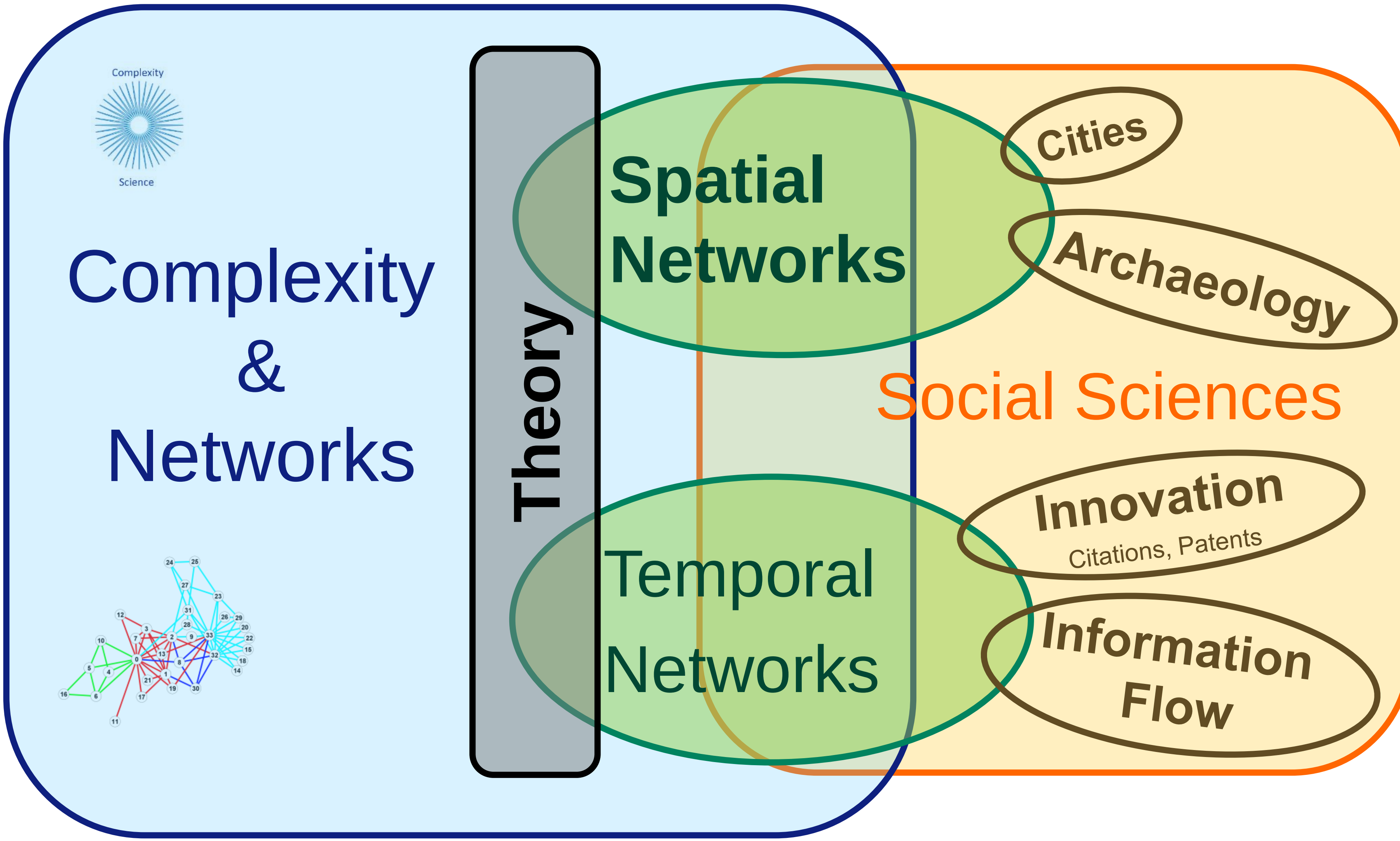
Real Networks have a lot of structure yet are not perfectly regular

- Short distances common like random graphs
- Local structure like regular lattice common
- Many have hubs – a few nodes have a large fraction of edges
  - fat tails for degree distributions, **Scale-Free networks**



TIM EVANS' CURRENT RESARCH

Tim Evans' Current Research (2015)



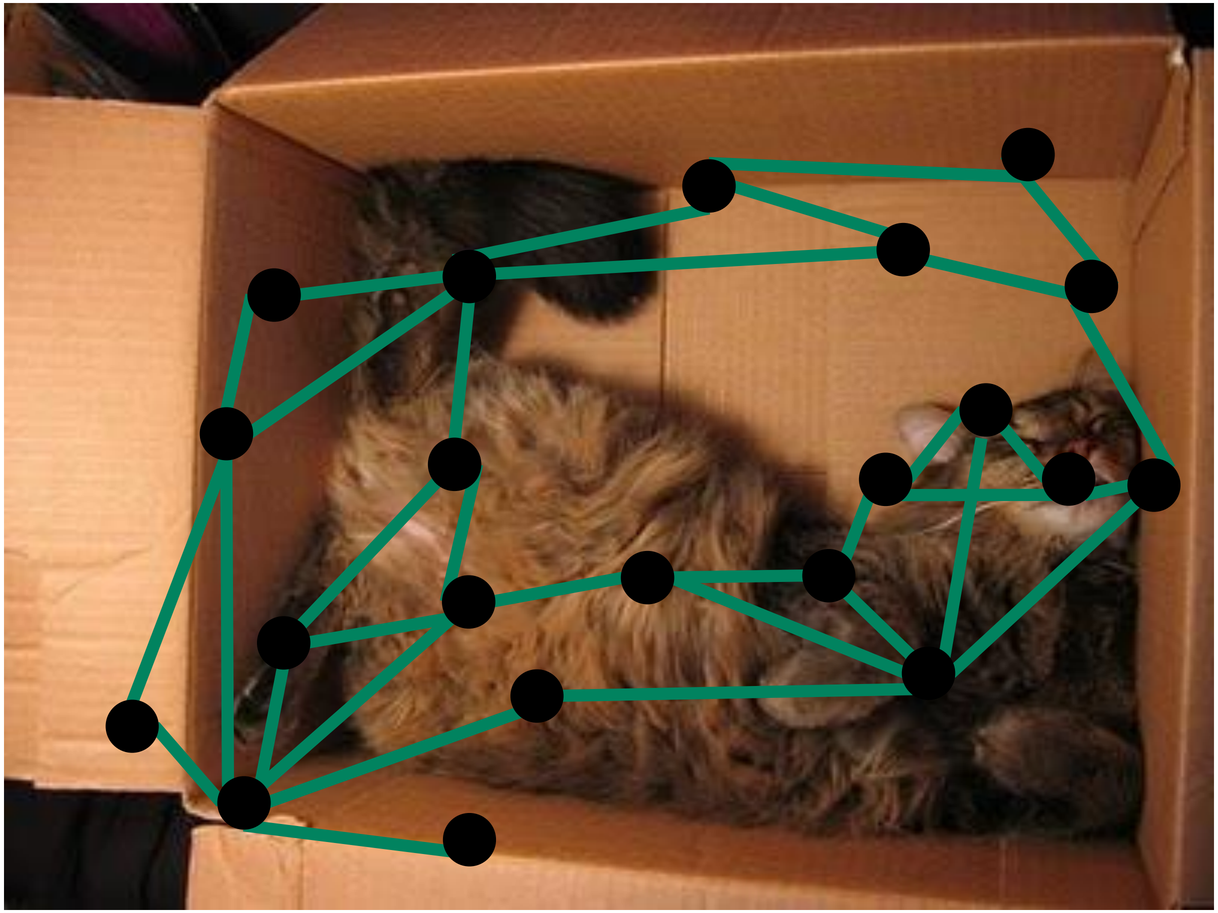
Basic Philosophy

Measurements on a network must be adapted to take account of all constraints

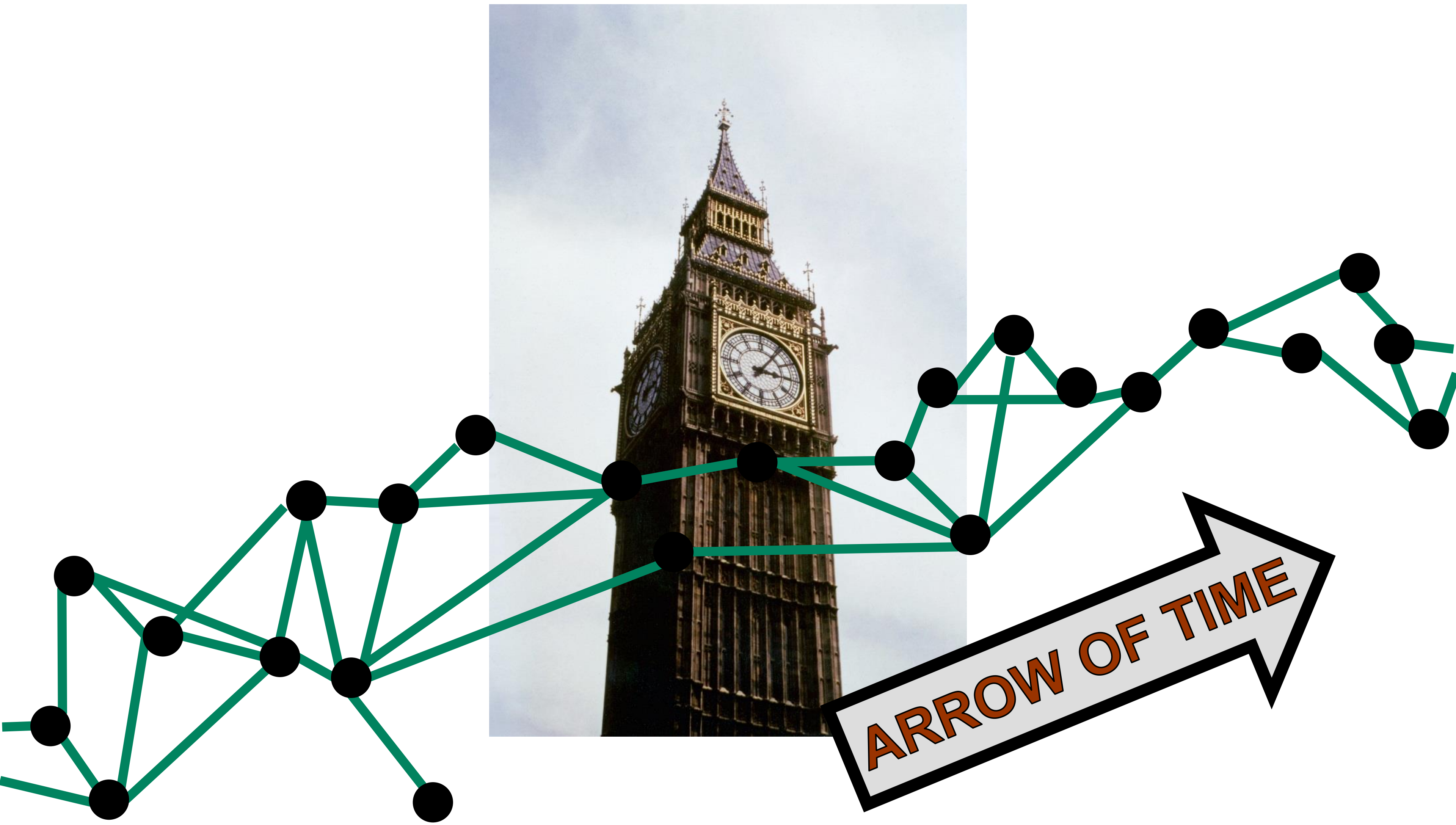
e.g.

- **in-degree** and **out-degree** used when we have *directed* graphs
- Use **strength** not degree for *weighted* graphs

Spatial Constraints



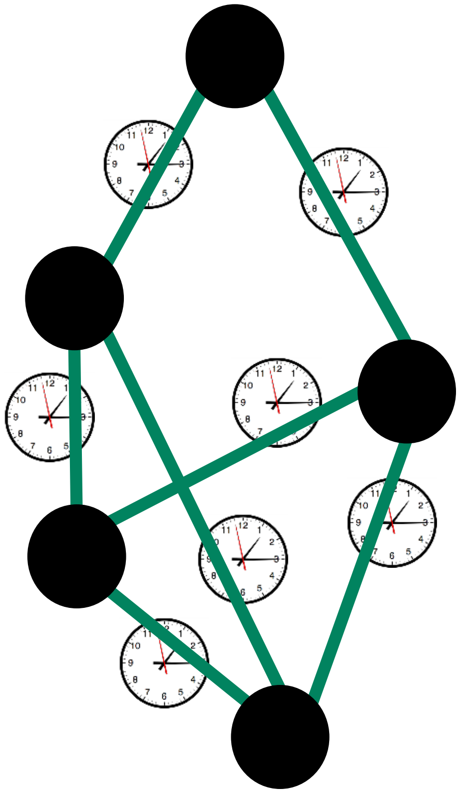
TimE and Networks



Timed Edges

Communication Networks

- Email
- Phone
- Letters



e.g. see [Holme, Saramäki, 2012]

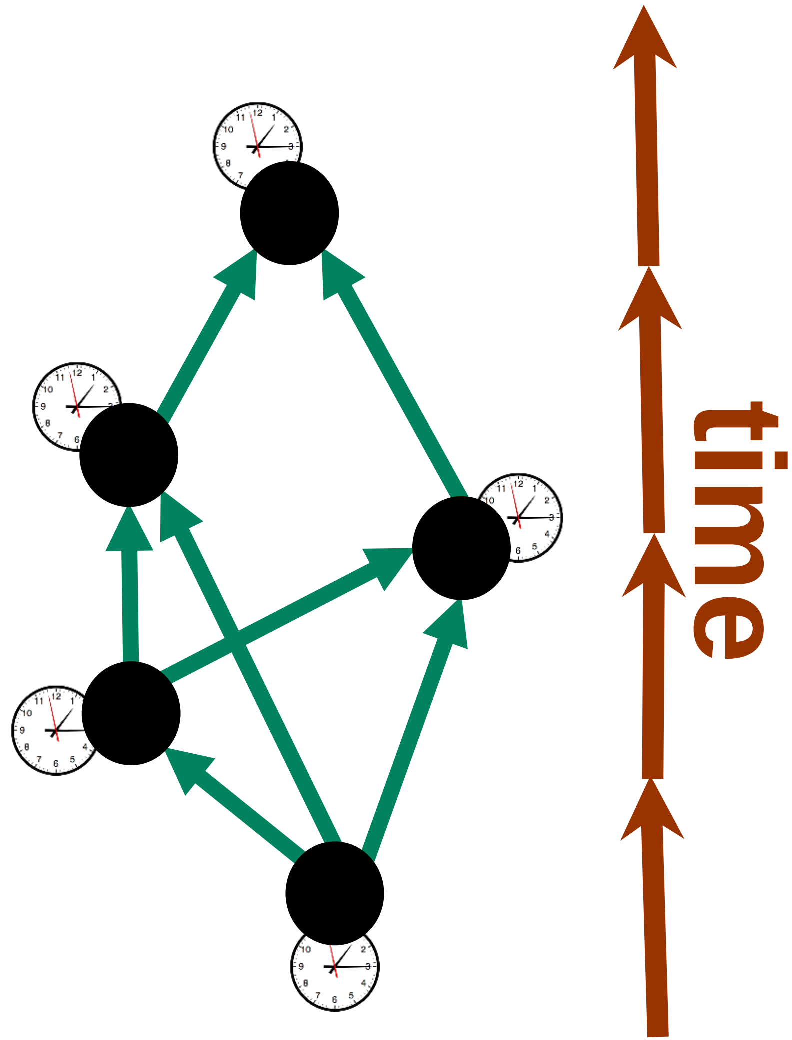
Timed nodes

Each node represents an event occurring at one time

⇒Time constrains flows

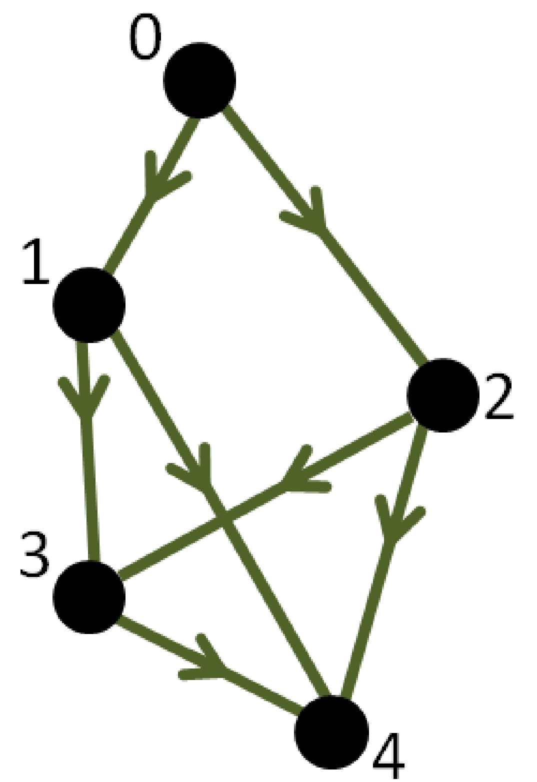
⇒ Edges point in one direction only

⇒ No Loops



DAG – Directed Acyclic Networks

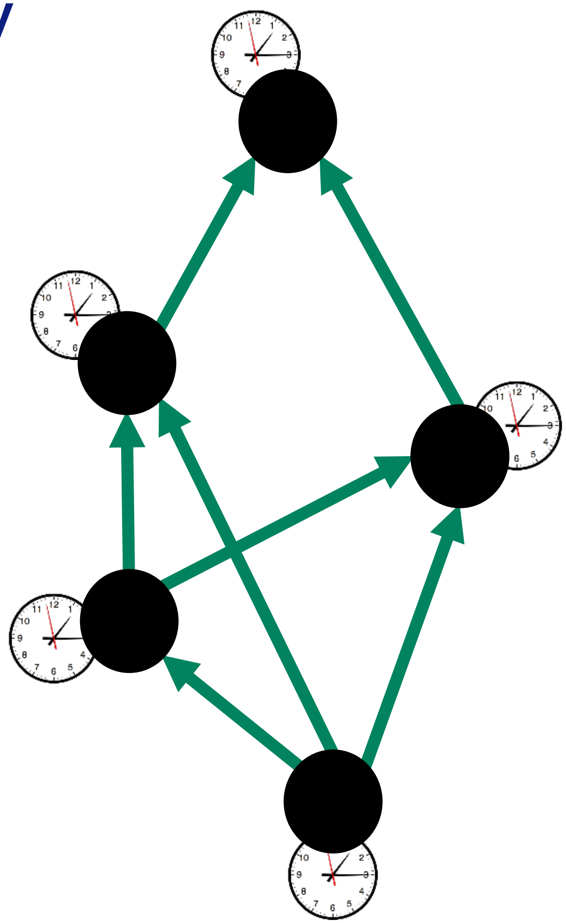
- Citation Graphs
- Patents
- Court Judgements
- Physical Flow
  - Raw materials to finished goods
- Logical Flow
  - Spreadsheets
  - All Maths
- Space-Time
  - Causal Set approach to quantum gravity [Dowker 2004]



Innovations

Each node represents a discovery

- Patent
- Academic Paper
- Law Judgment

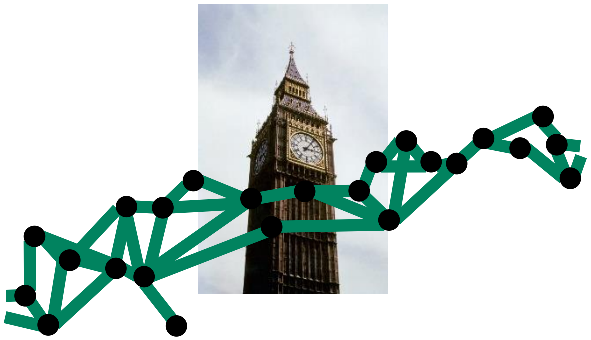


Information is now **copied** from one node to all connected events in the future

Different process from a random walk

Similar to epidemics but different network

TimE and Networks



- Transitive Reduction
- Dimension

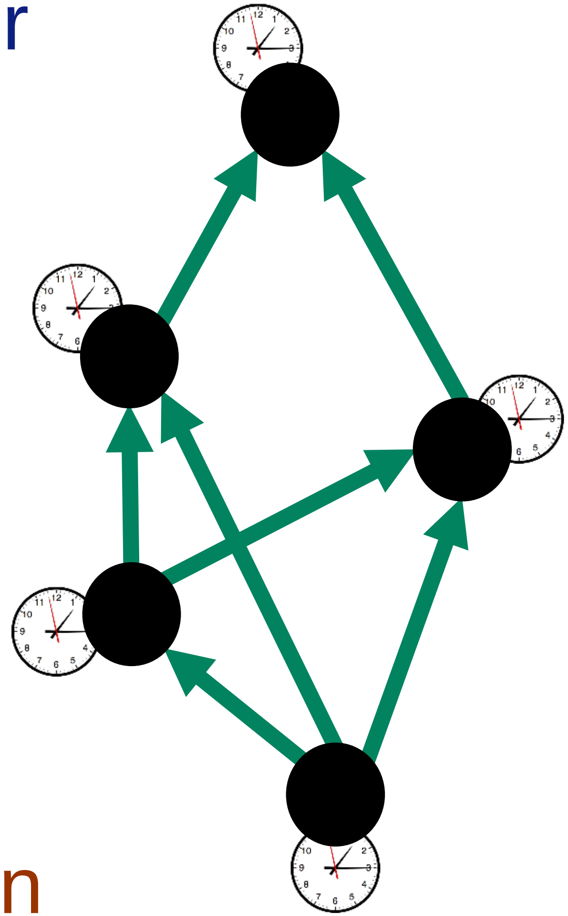
Transitive Reduction

Remove all edges not needed for causal links

- Uniquely defined because of causal structure

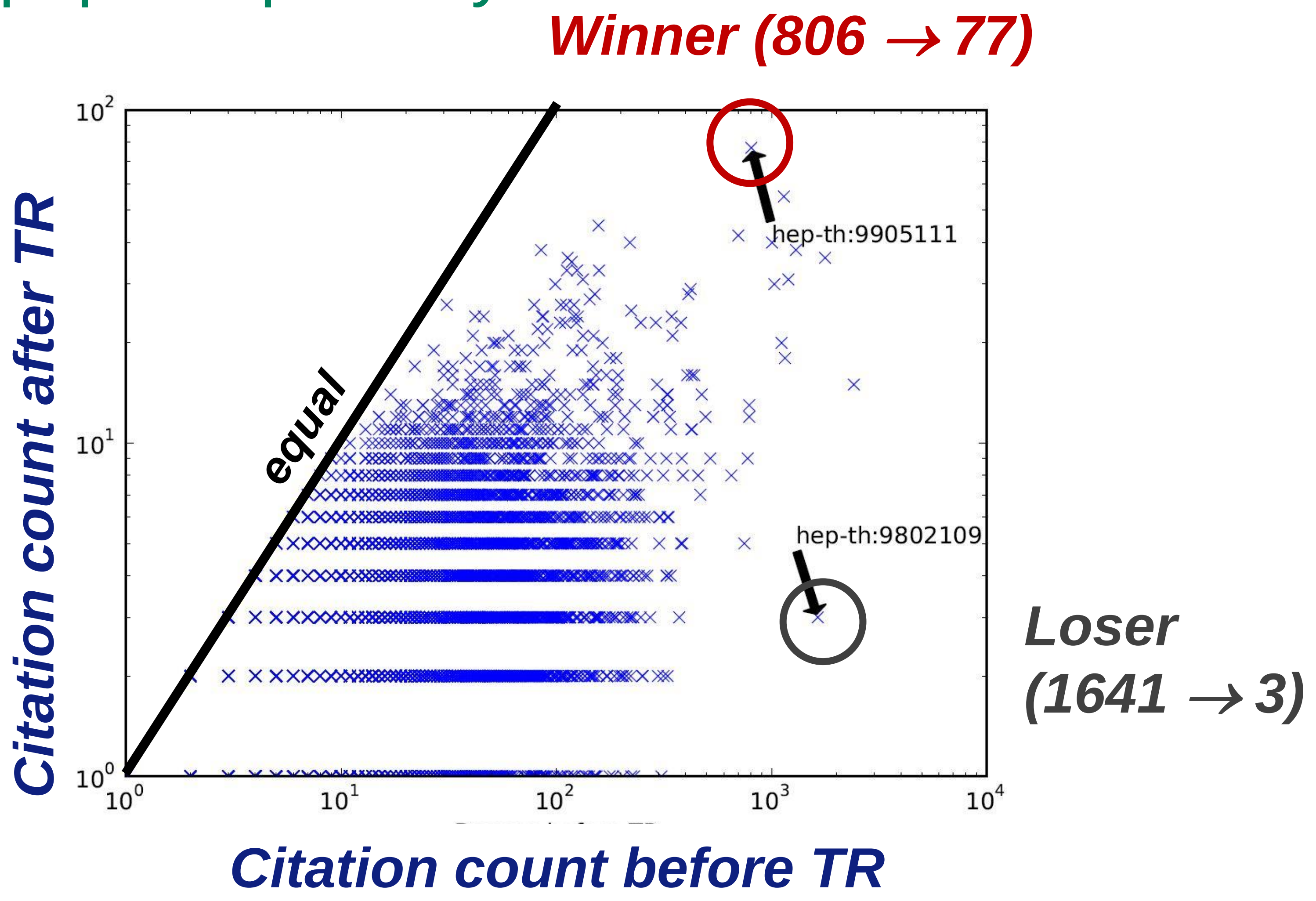
Conjecture:

This removes unnecessary or indirect influence on innovation



[Clough et al. 2014]

arXiv paper repository



Transitively Reducing a Citation Network

- The Transitively Reduced citation network has the same causal structure as the original data.
  - Information can still flow

- Transitive Reduction removes **~80%** edges
  - Simkin & Roychowdury 2002; Goldberg, Anthony, TSE 2015

- Conjecture: the edges in Transitive Reduced citation network are the most useful.

Winner and Losers

- ‘Winners’ = papers whose citation rank increases after Transitive Reduction
  $\Rightarrow$ Most useful to many different areas
  $\Rightarrow$ More impact

- ‘Losers’ = papers that drop relative to others after Transitive Reduction
  $\Rightarrow$ papers cited within a narrow group
  $\Rightarrow$ Less impact

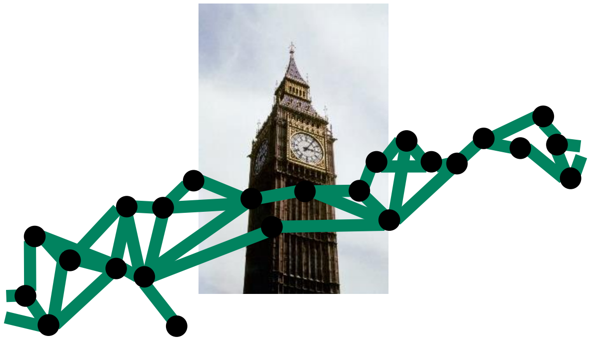
Citation Counts

Citations of academic papers not always made for good reasons

- Cite your own paper
- Cite a standard paper because everyone does
- Copy a citation from another paper (80%? [Simkin and Roychowdhury 2003])

Transitive reduction removes poor citations

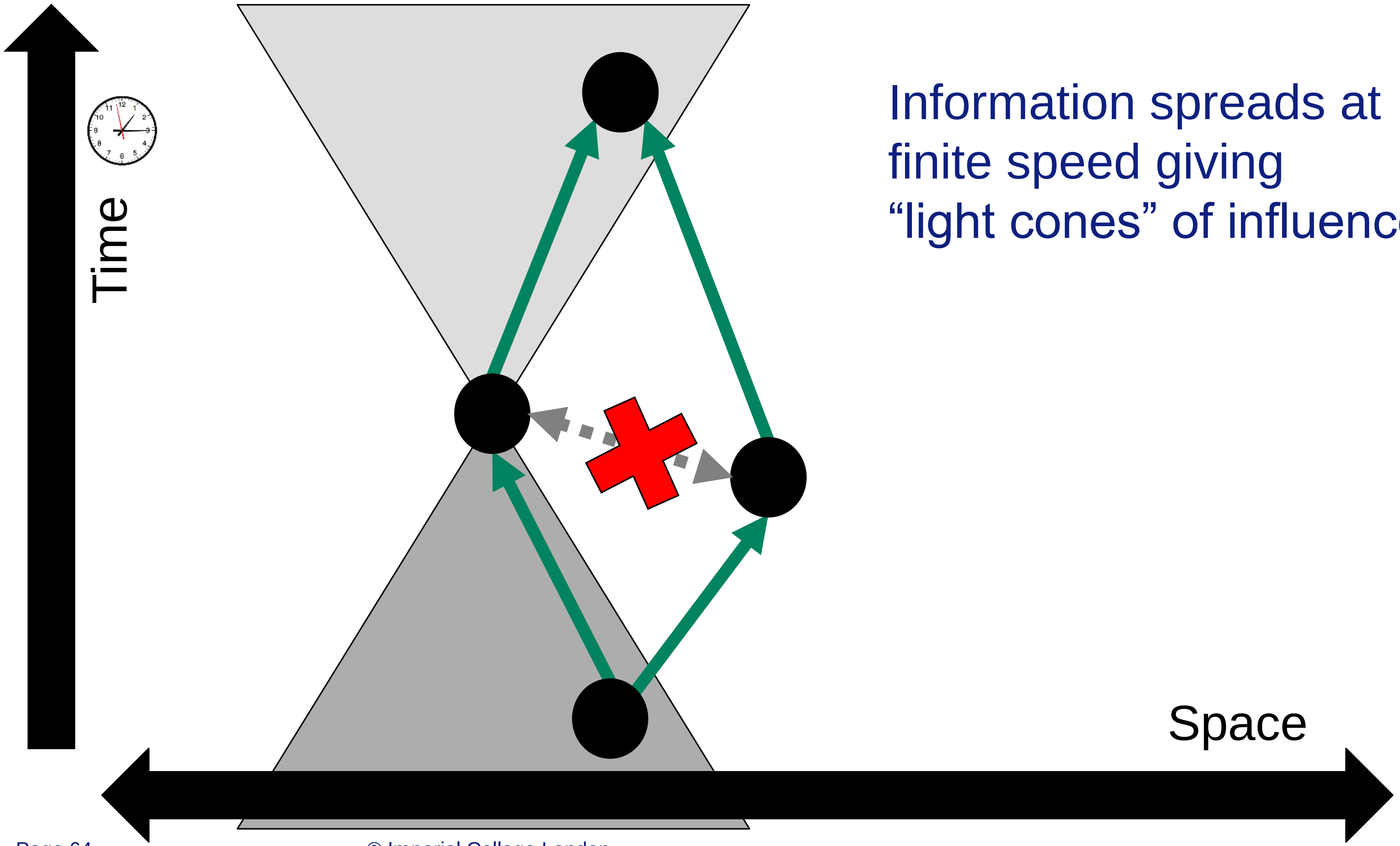
TimE and Networks



- Transitive Reduction
- Space-Time Dimension

Space-Time

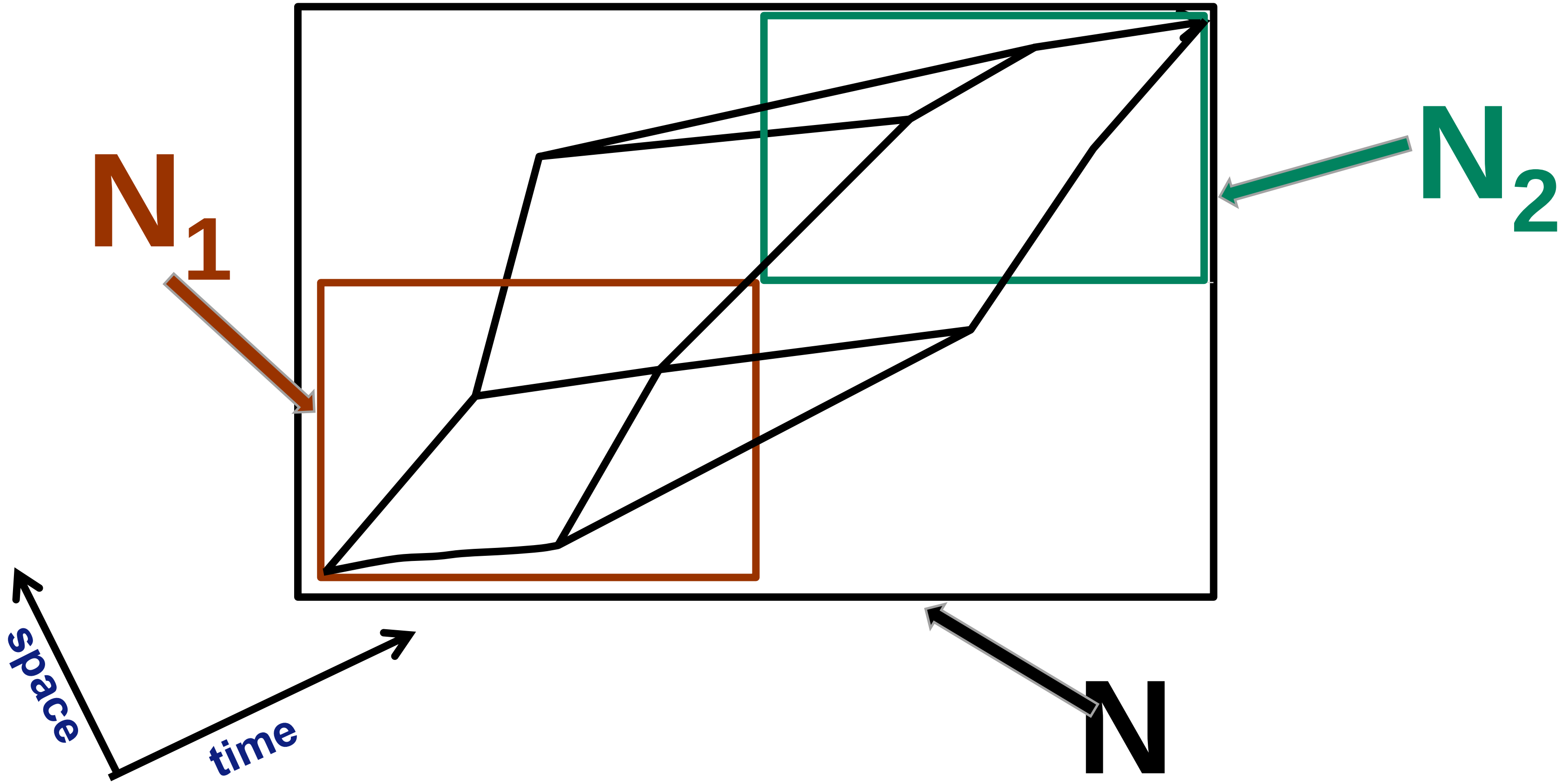
Minkowski space-time



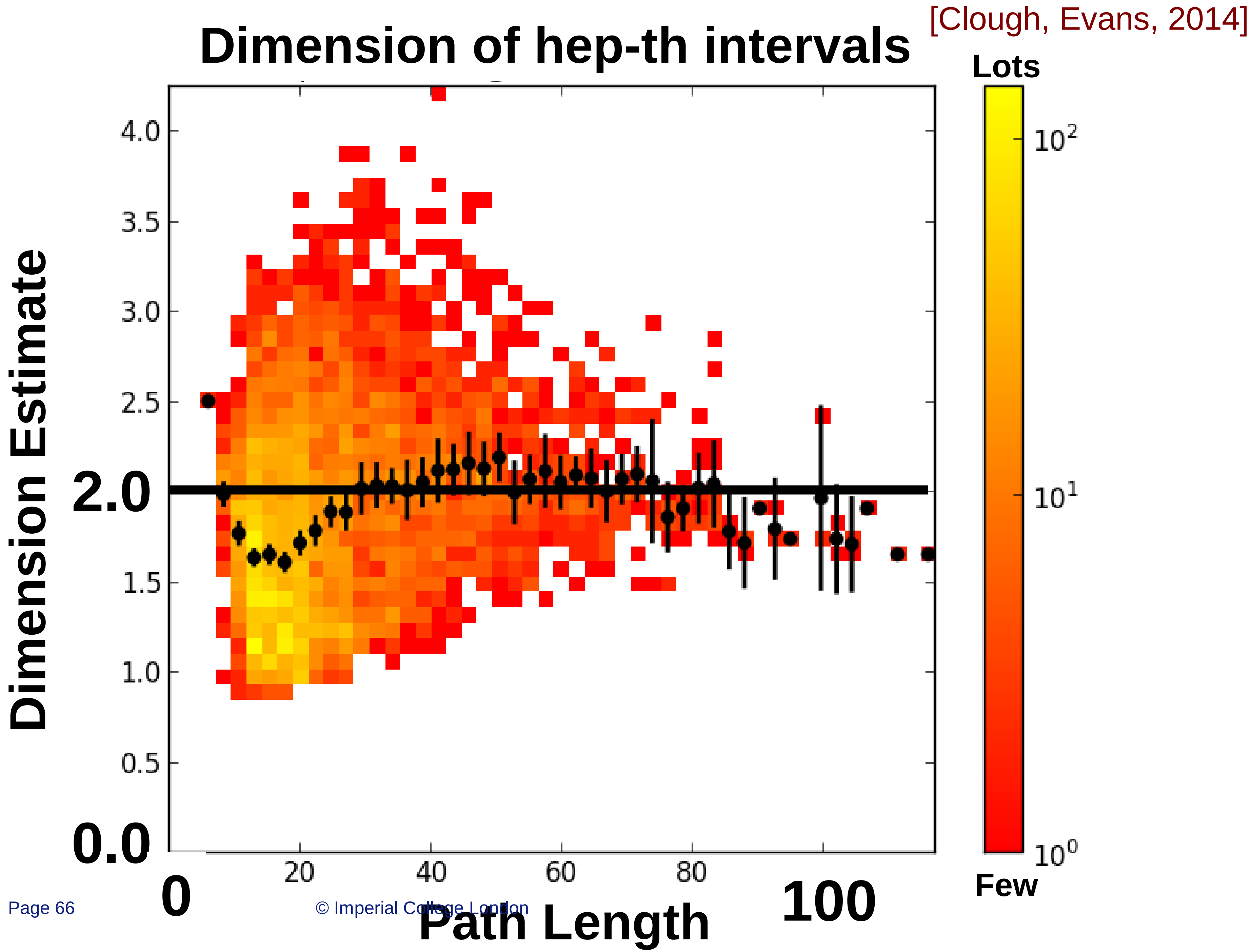
Information spreads at finite speed giving "light cones" of influence

Midpoint Scaling Dimension

- We should expect to see $N_1 \approx N_2 \approx N \times 2^{-d}$

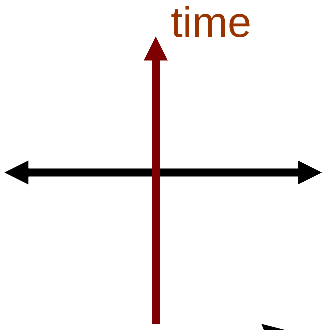


Dimension of hep-th intervals

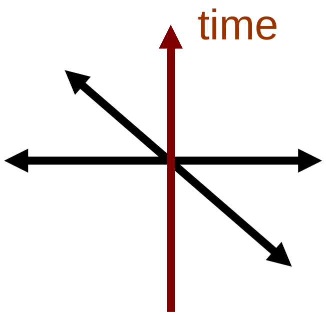


What does this tell us?

For two different dimension measures on two **`arXiv`** subsets:



- **`hep-th`** has dimension **2**
  - 1 time, 1 “topic direction”
- **`hep-ph`** has dimension **3**
  - 1 time, 2 “topic directions”



Dimension of a DAG gives a new insight into differences between research topics

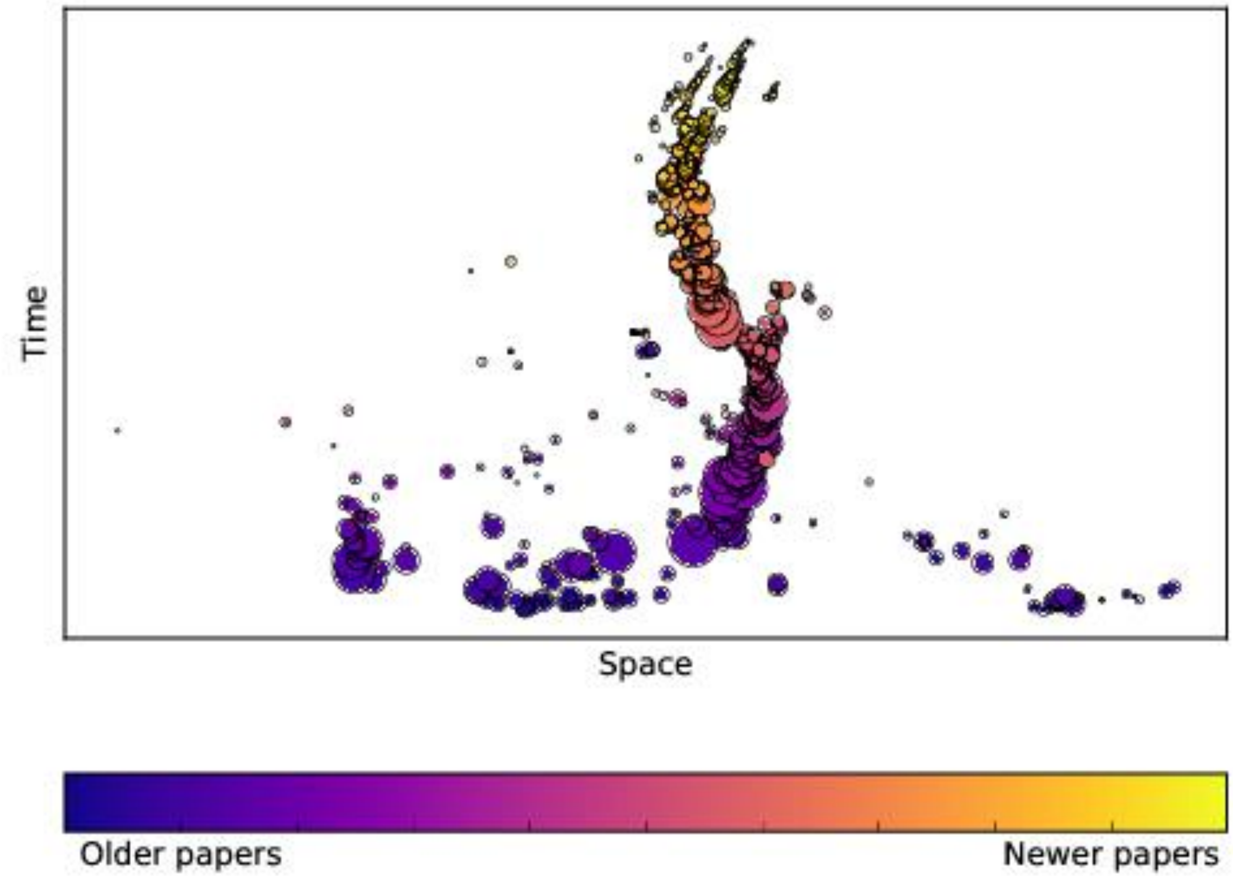
[Clough & Evans 2014]

Dimensions

Data	Dimension
hep-th (String Theory)	2
hep-ph (Particle Physics)	3
quant-ph (quantum physics)	3
astro-ph (astrophysics)	3.5
US Patents	>4
US Supreme Court Judgments	3 (short times), 2 (long times)

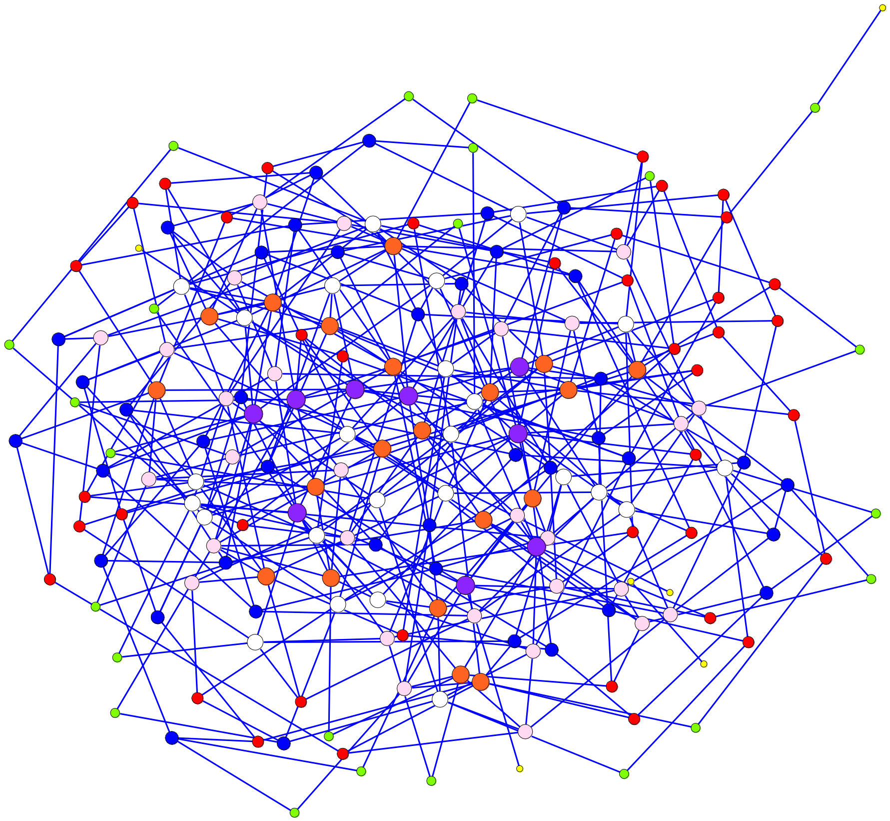
String theory appears to be a narrow field

Embedding hep-th papers



Future Directions: DAG Temporal Networks

- Refine existing measures
- Interpretation of measures
  - Social science input
- Develop practical citation network tools
- Applications other than citation networks
- Theoretical models and Mathematical results
  - Connections to directed percolation
  - Connections to work in partially ordered sets



NETWORK SCIENCE

What is “Network Science”?

- Based on analysis through networks
  - Graphs, hypergraphs
- Part of wider studies in complexity
  - Local interactions produce emergent phenomena
- Not new
  - Social Network Analysis since 50’s
  - Mathematical graph theory since Euler in 1735
- New aspect is Information Age
  - Large data sets and their analysis now possible
- Multidisciplinary
  - Communication difficult between fields

Does “Network Science” really exist?

Possible criticisms:-

- No coherent definition
- Too broad to be a single area
- New name for old work = ***Hype***
- Too early to say
- No need to define a new field

[“Network Science”, nap.edu, 2005]

Are networks providing new insights?

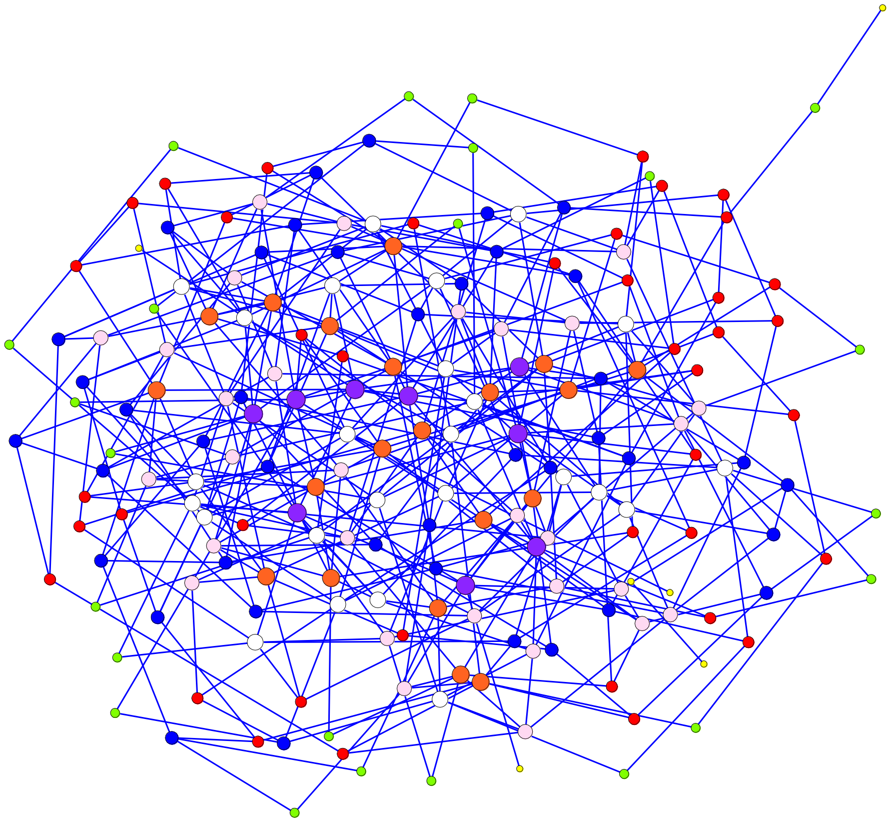
- Another approach to statistical analysis and data mining
- Sometimes this is a better way to analyse
  - Gives new questions e.g. Small world definition
  - Gives new answers e.g. Small world models
- Brings the tools of Complexity
  - Scaling

Imperial College London

Complexity Science

Social & Cultural Analytics Lab,
DSI, Imperial College London

netplexity



THANKS

See **netplexity.org**

or search for *Tim Evans Networks*

Slides available from
http://dx.doi.org/10.6084/m9.figshare.3467066

Bibliography: Free General Network Reviews

- Evans, T.S. “Complex Networks”, *Contemporary Physics,* **2004***, 45*, 455-474 [arXiv.org/cond-mat/0405123]
- Easley, D. & Kleinberg, J. Networks, “Crowds, and Markets: Reasoning About a Highly Connected World Cambridge University Press, 2010” [http://www.cs.cornell.edu/home/kleinber/networks-book/]
- van Steen, M. “Graph Theory and Complex Networks” Maarten van Steen, 2010 [http://www.distributed-systems.net/gtcn/]
- Hanneman, R. A. & Riddle, M. “Introduction to social network methods” 2005 Riverside, CA: University of California, Riverside [http://faculty.ucr.edu/~hanneman/]
- Brandes, U. & Erlebach, T. (ed.) “Network Analysis: Methodological Foundations” 2005 [http://www.springerlink.com/content/nv20c2jfpf28/]

Bibliography for this talk

- Evans, T.S. “Complex Networks”, *Contemporary Physics,* **2004***, 45*, 455-474 [arXiv.org/cond-mat/0405123]
- Newman, M. “Networks: An Introduction”, OUP, 2010
- Watts, D. J. & Strogatz, S. H., “Collective dynamics of 'small-world' networks”, *Nature,* **1998***, 393*, 440-442.
- F. Cairncross (1997) The Death of Distance, Harvard University Press, Cambridge MA.
- Zachary, W. “Information-Flow Model For Conflict And Fission In Small-Groups” *J. Anthropological Research,* **1977***, 33*, 452-473
- Hal Burch, Bill Cheswick's “Internet Mapping Project” http://personalpages.manchester.ac.uk/staff/m.dodge/cybergeography/atlas/lumeta_large.jpg
- P.Kaluza, A.Kölzsch, M.T.Gastner, B.Blasius, J. R. Soc. Interface 2010, **7**, 1093-1103

Bibliography for this talk

- P.Kaluza, A.Kölzsch, M.T.Gastner, B.Blasius, J. R. Soc. Interface 2010, **7**, 1093-1103
- Holten, D. & van Wijk, J. J.”Force-Directed Edge Bundling for Graph Visualization”, Eurographics/ IEEE-VGTC Symposium on Visualization 2009, H.-C. Hege, I. Hotz, and T. Munzner (Guest Editors) 28 (2009), Number 3
- Chun-Yi Zac Lo, Yong He, Ching-Po Lin, “Graph theoretical analysis of human brain structural networks”, Reviews in the Neurosciences 2011 (on line)
- Hillier, “Space is the Machine”, 1996

Bibliography for this talk

- T.S.Evans, J.Saramäki, “Scale Free Networks from Self-Organisation”, *Phys Rev E,* **2005***, 72*, 026138 [arXiv:cond-mat/0411390]
- D.S. Price, Networks of Scientific Papers, Science, 1965, **149**, 510-515.
- Barabási, A. & Albert, R. “Emergence of scaling in random networks”, Science, 1999, **286**, 173.
- T.S.Evans, R.Lambiotte, “Line Graphs, Link Partitions and Overlapping Communities”, *Phys.Rev.E,* **2009***, 80*, 016105 [arXiv:0903.2181]
- Evans, T.S. “Clique Graphs and Overlapping Communities”, *J.Stat.Mech,* **2010**, P12037 [arxiv.org:1009.063]
- Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R. & Lefebvre, E., “Fast unfolding of community hierarchies in large networks”, *J STAT MECH-THEORY E,* **2008**, P10008 [arXiv.org:0803.0476]
- Barabási, A. & Albert, R. “Emergence of scaling in random networks”, Science, 1999, 286, 173
- Map of Complexity, http://commons.wikimedia.org/wiki/File:Complexity-map_Castellani.jpg

