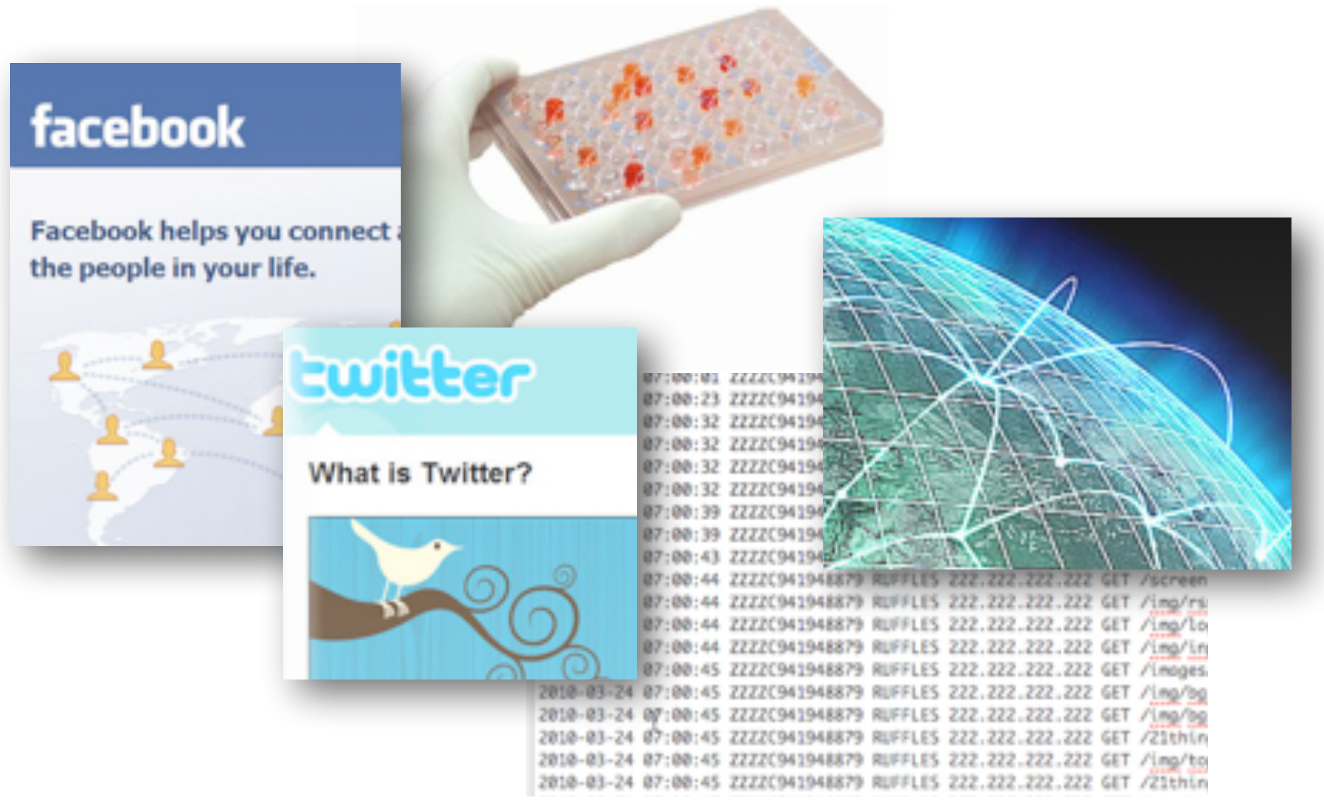
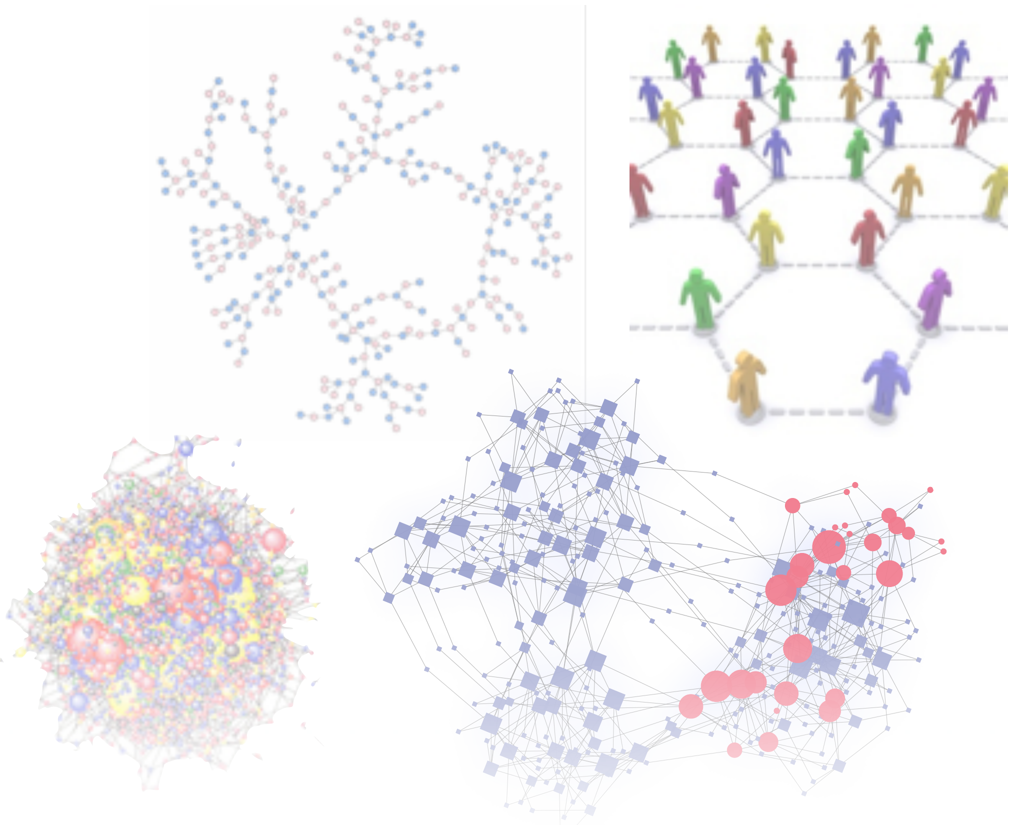
Introduction to Network Science

Jim Bagrow

NERCCS 2020 — April 1, 2020





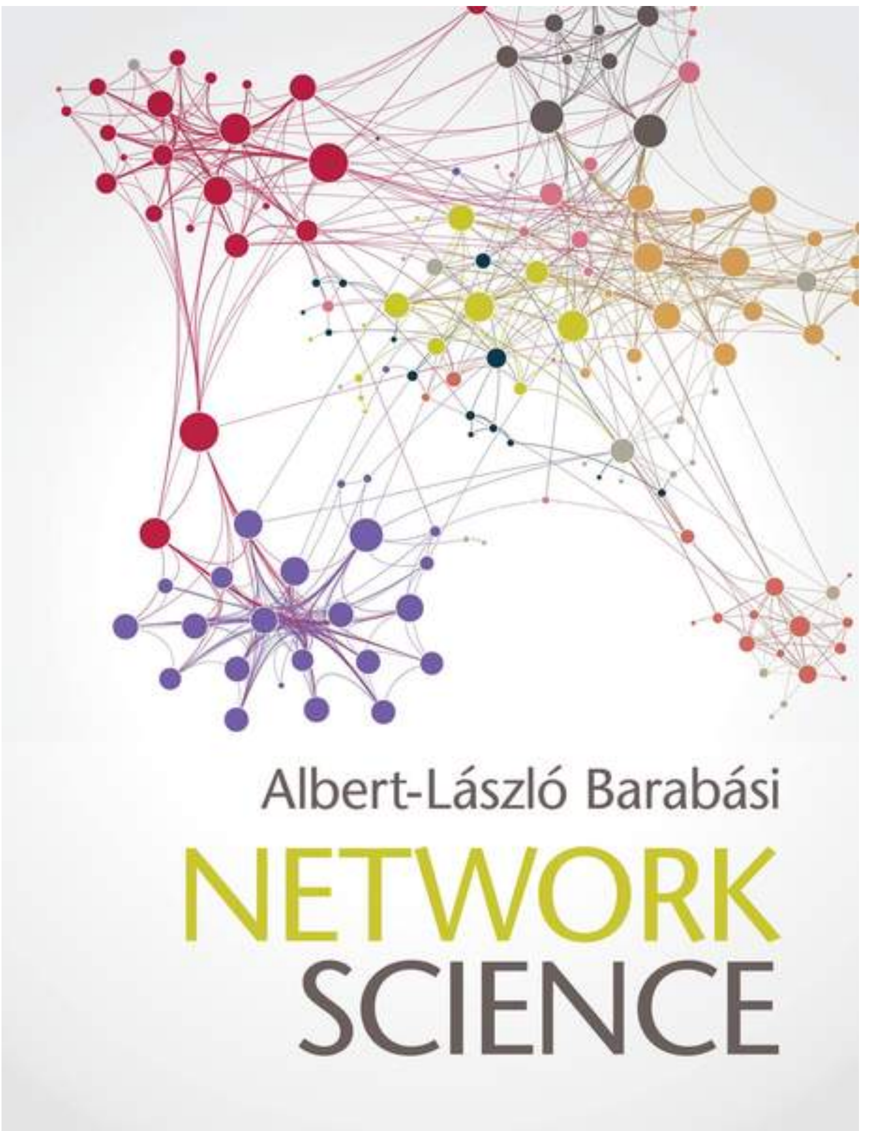
The University of Vermont

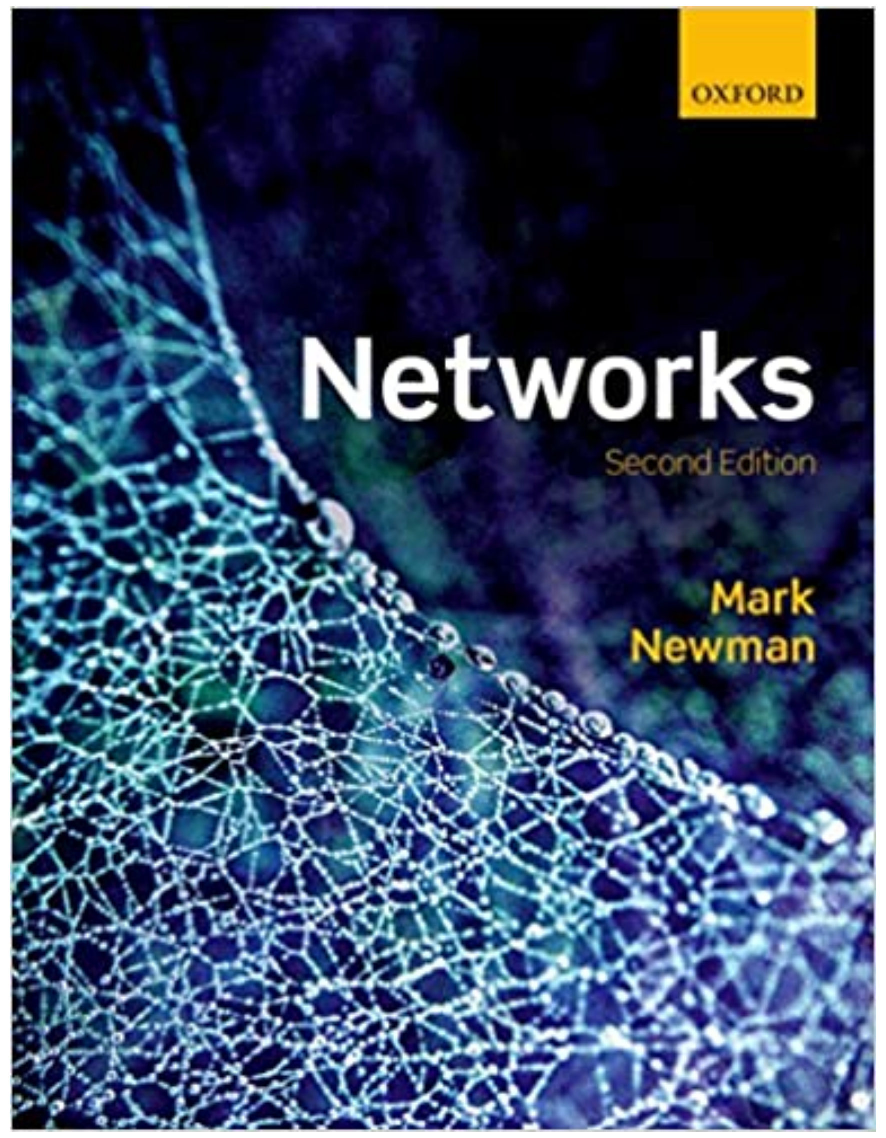
Vermont Complex Systems Center

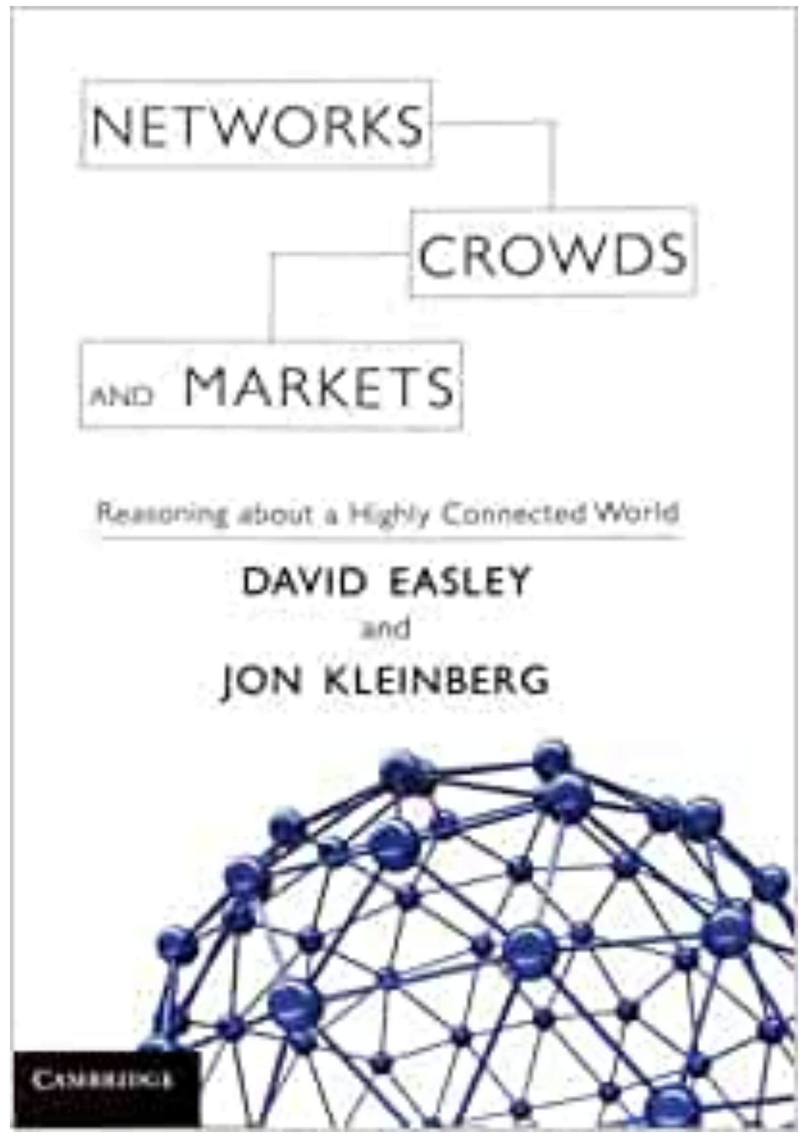
Outline

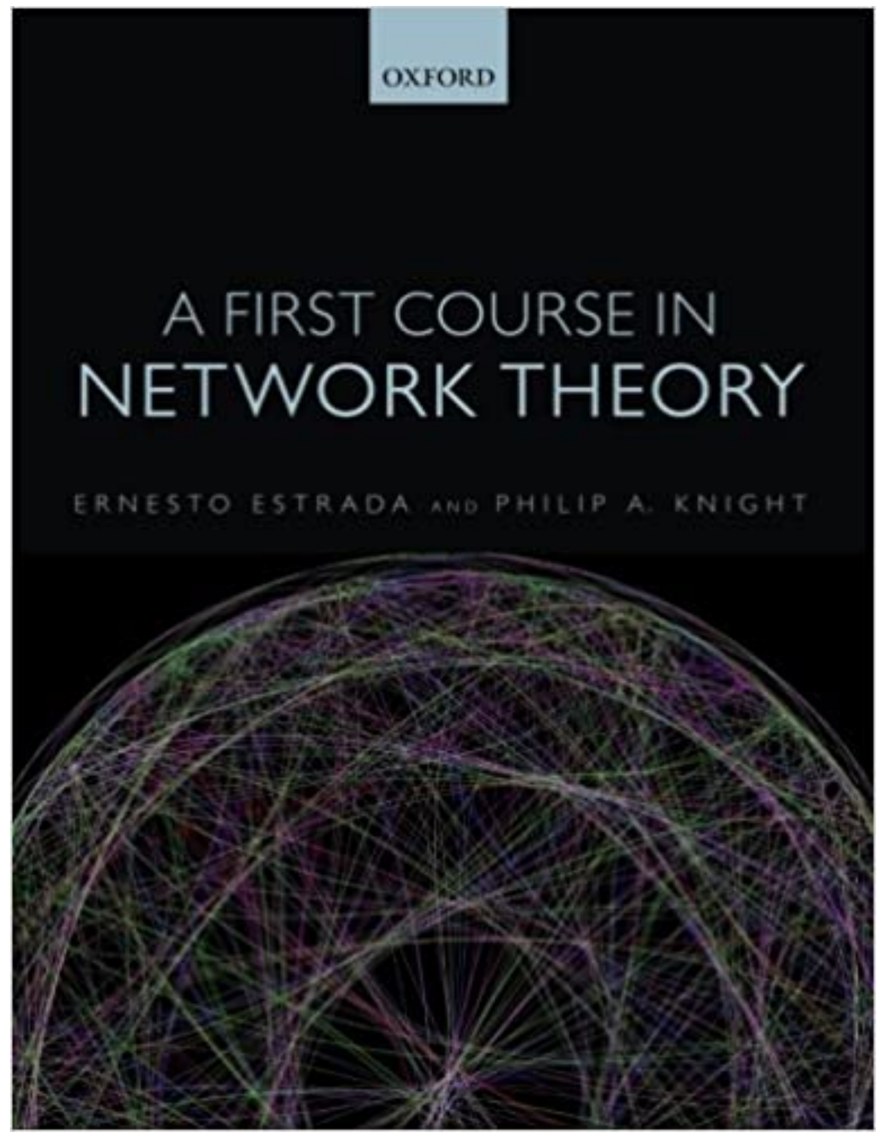
- Network examples/data
- Why study networks?
- Types of networks
- Network quantifiers (jargon!)
- Random network models
- Network robustness
- Dynamics on networks
- Future of network science

Some good references



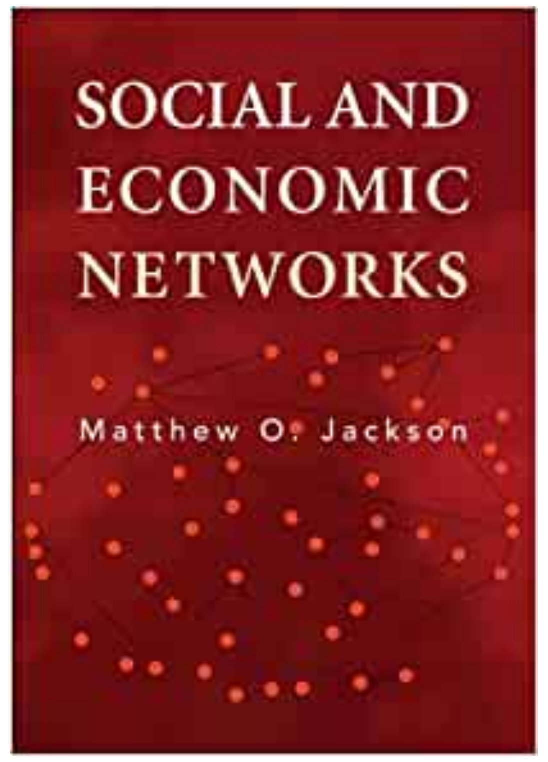


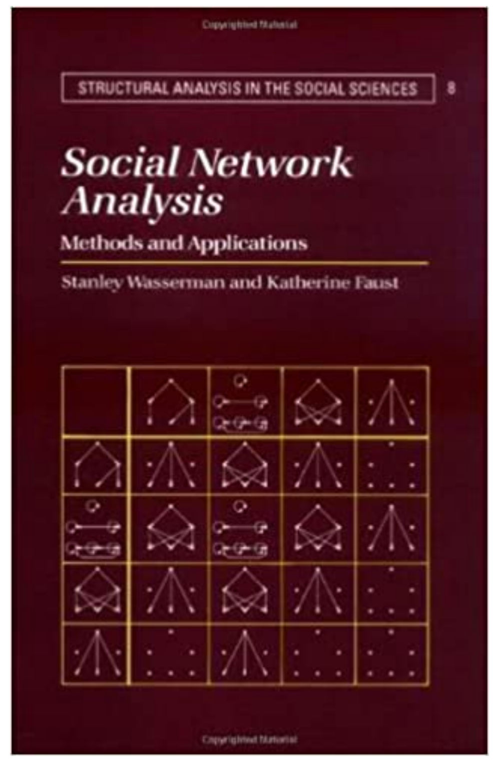


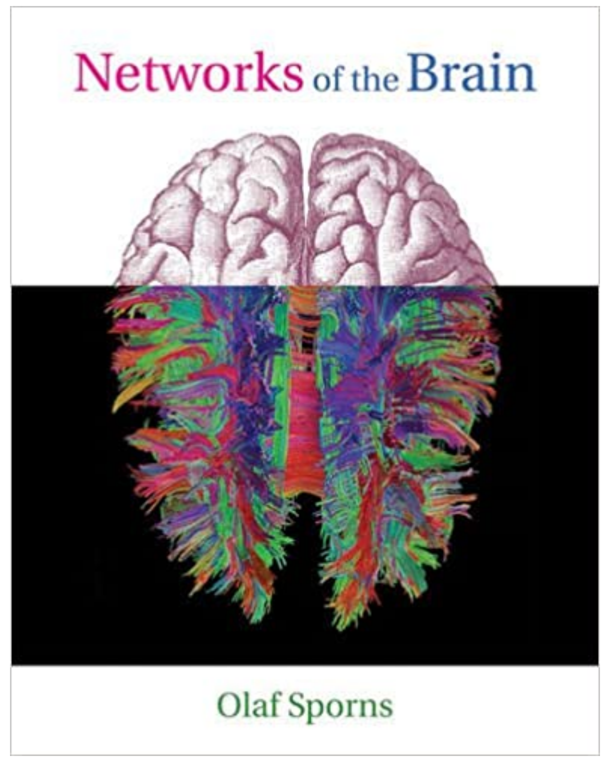


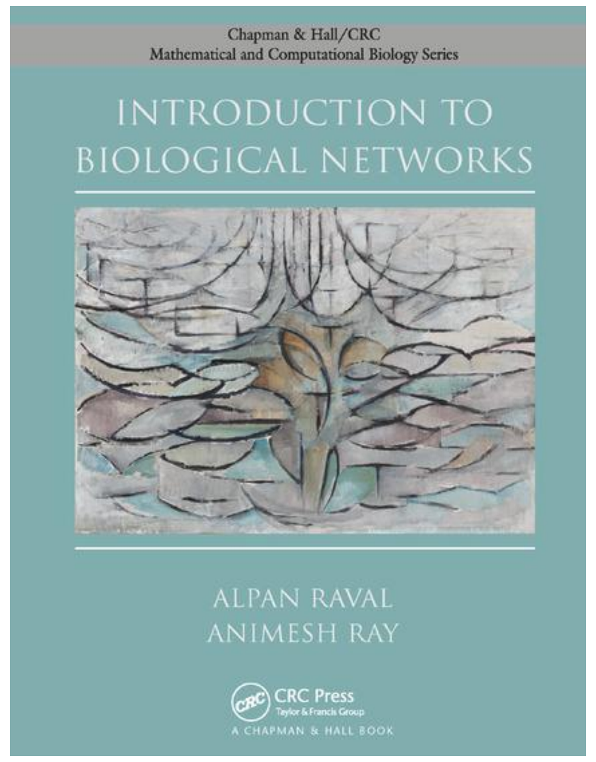
+ many more!

domain-specific texts:

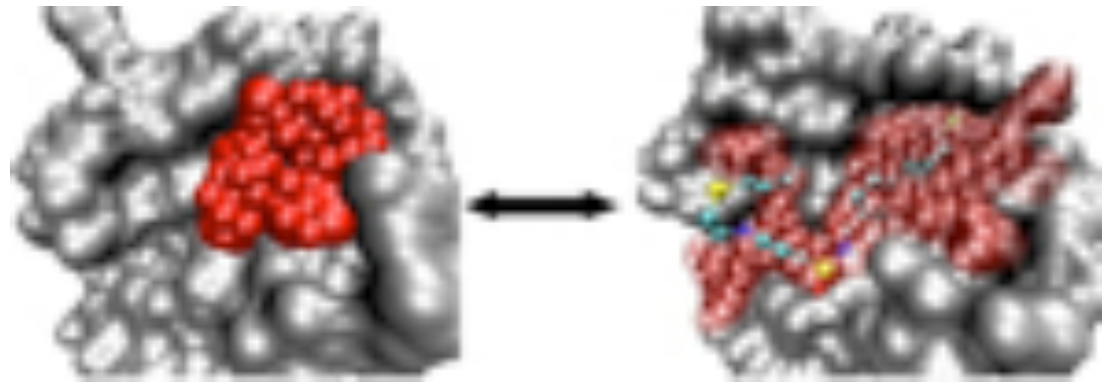




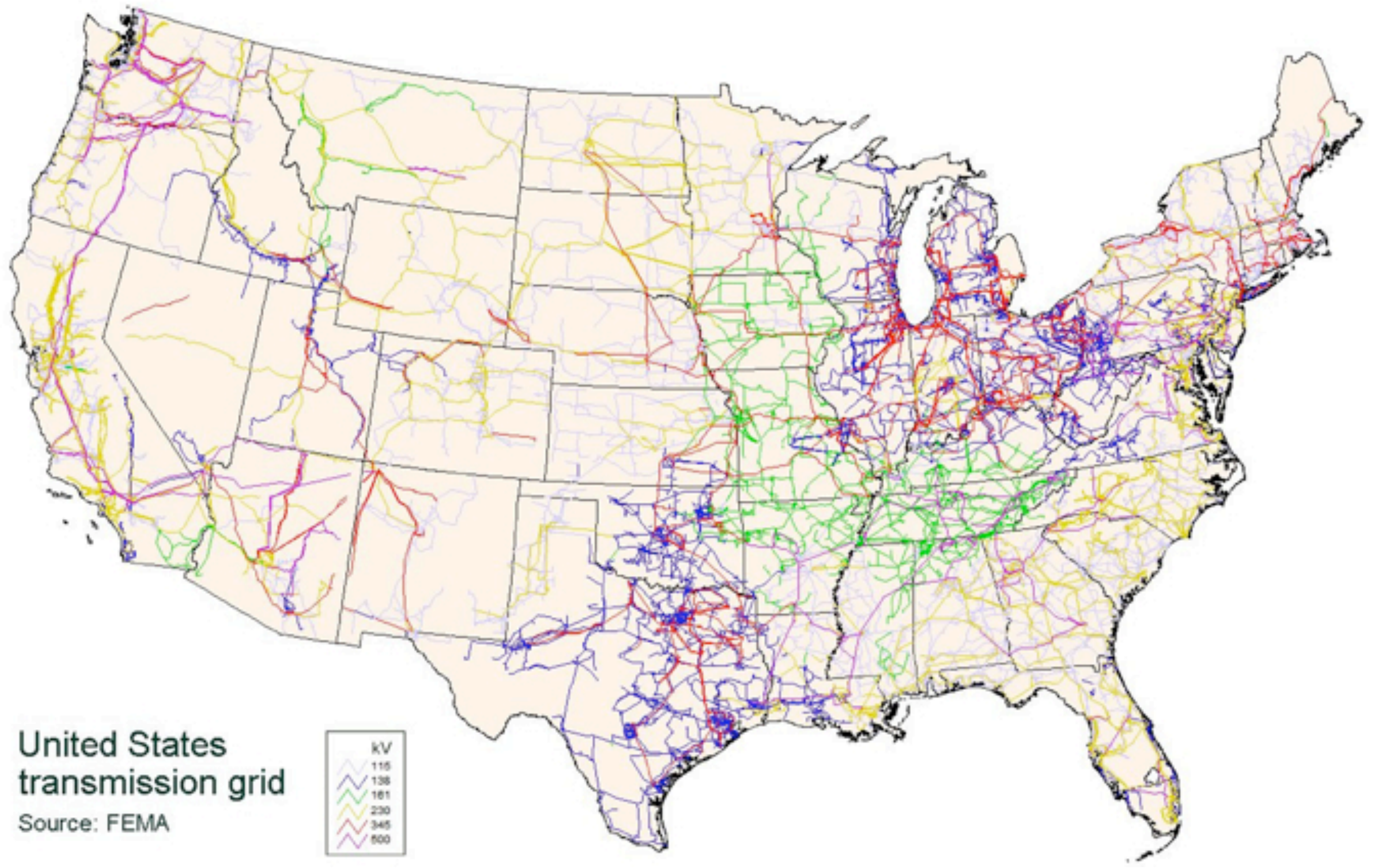




Examples of networks and network data



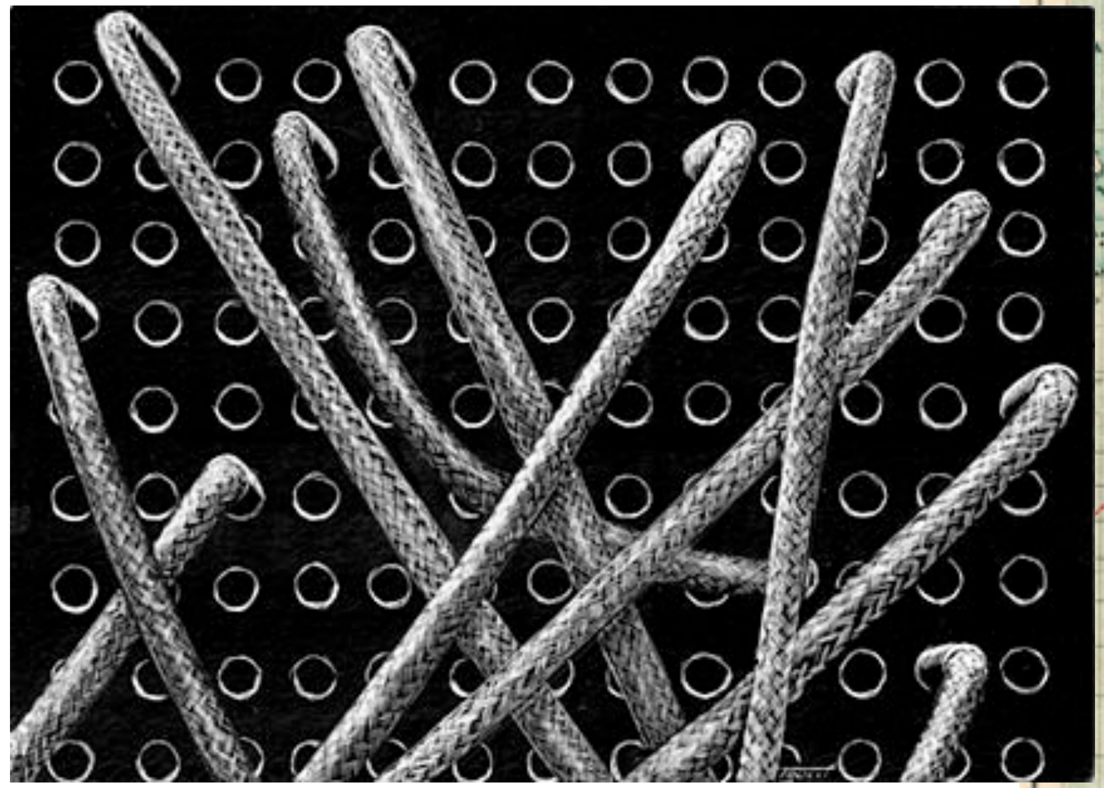
Technology & Infrastructure

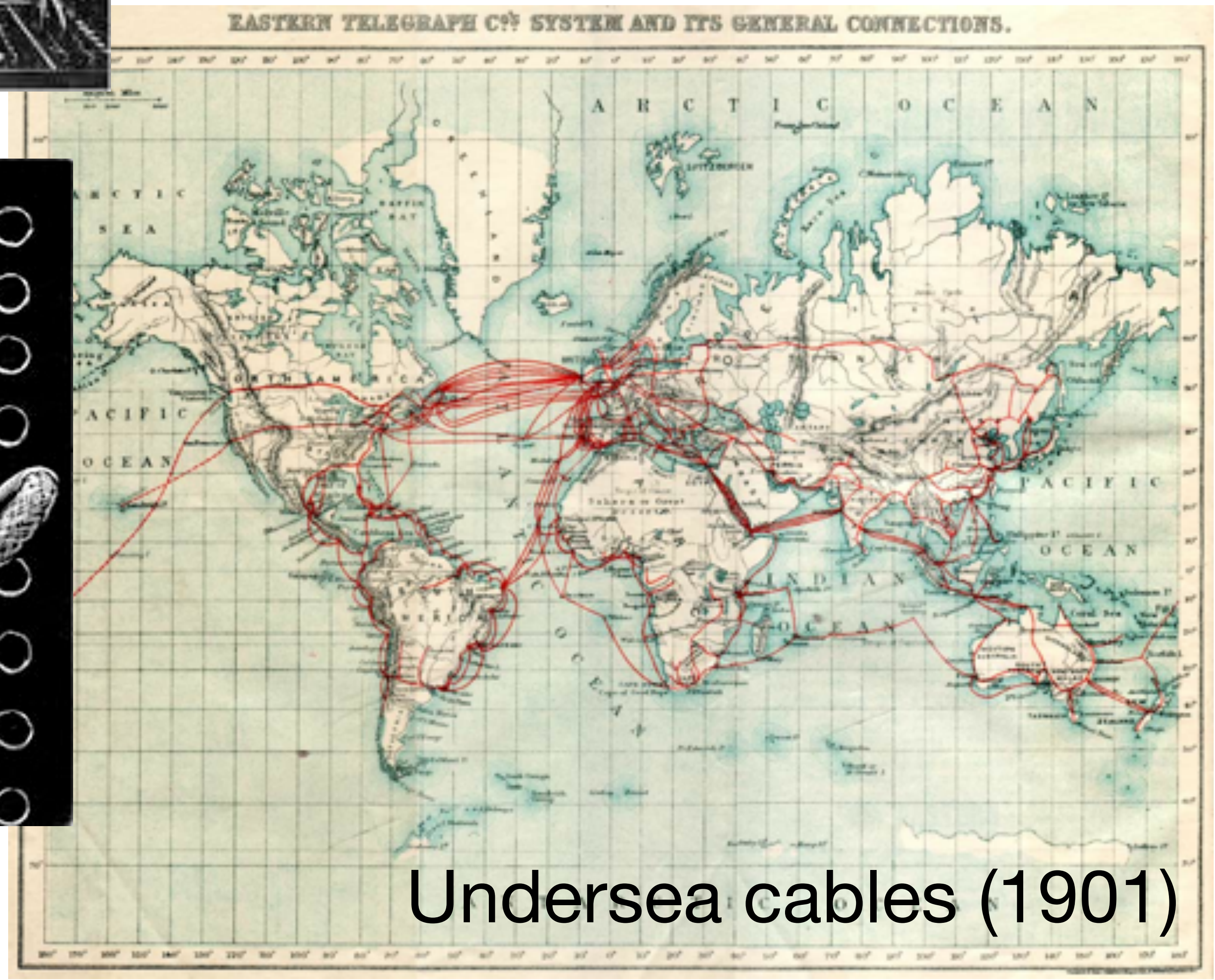


Power grid

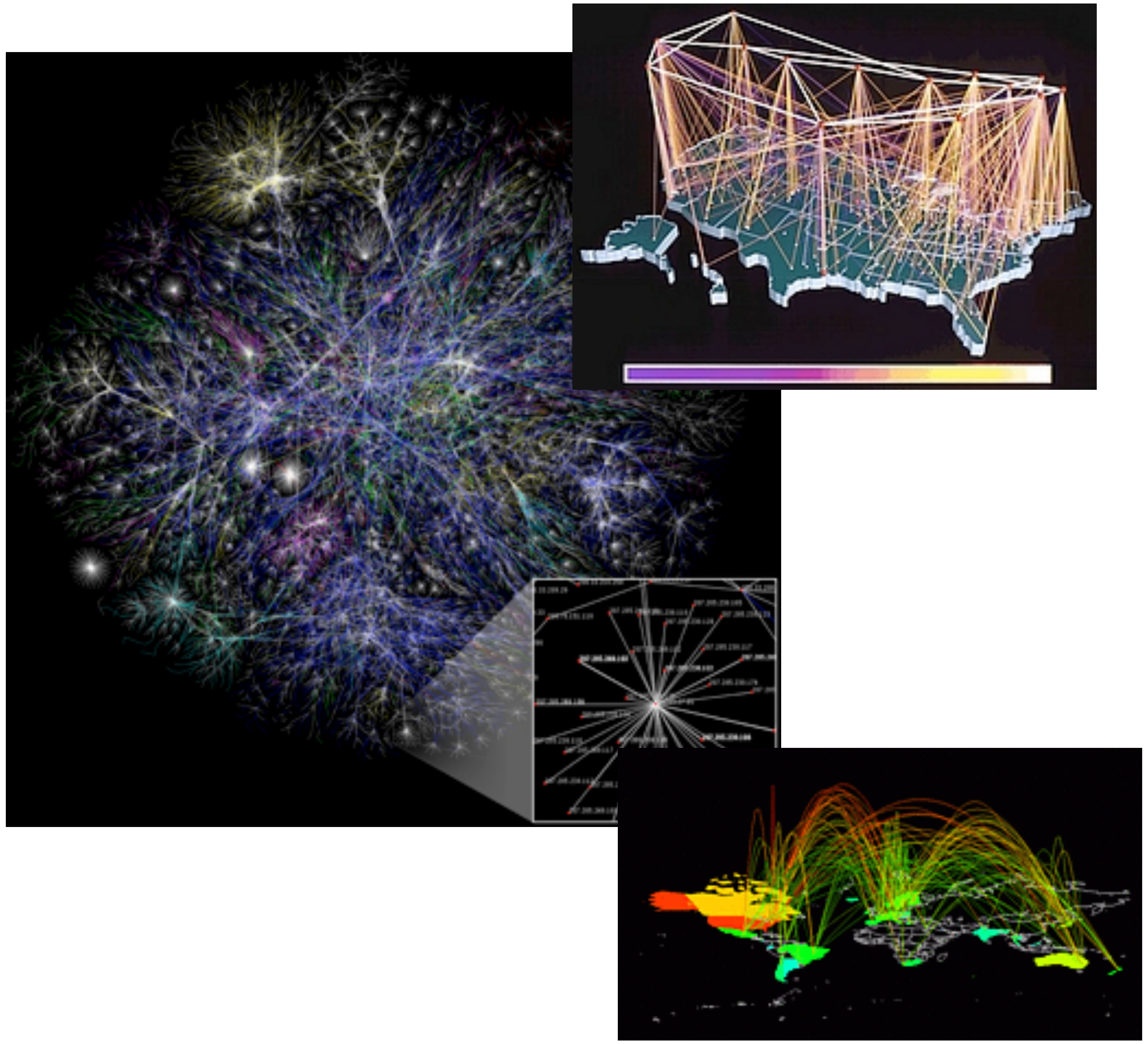
Understand **cascading failures** and **blackouts**

Telecommunications & Internet





Undersea cables (1901)



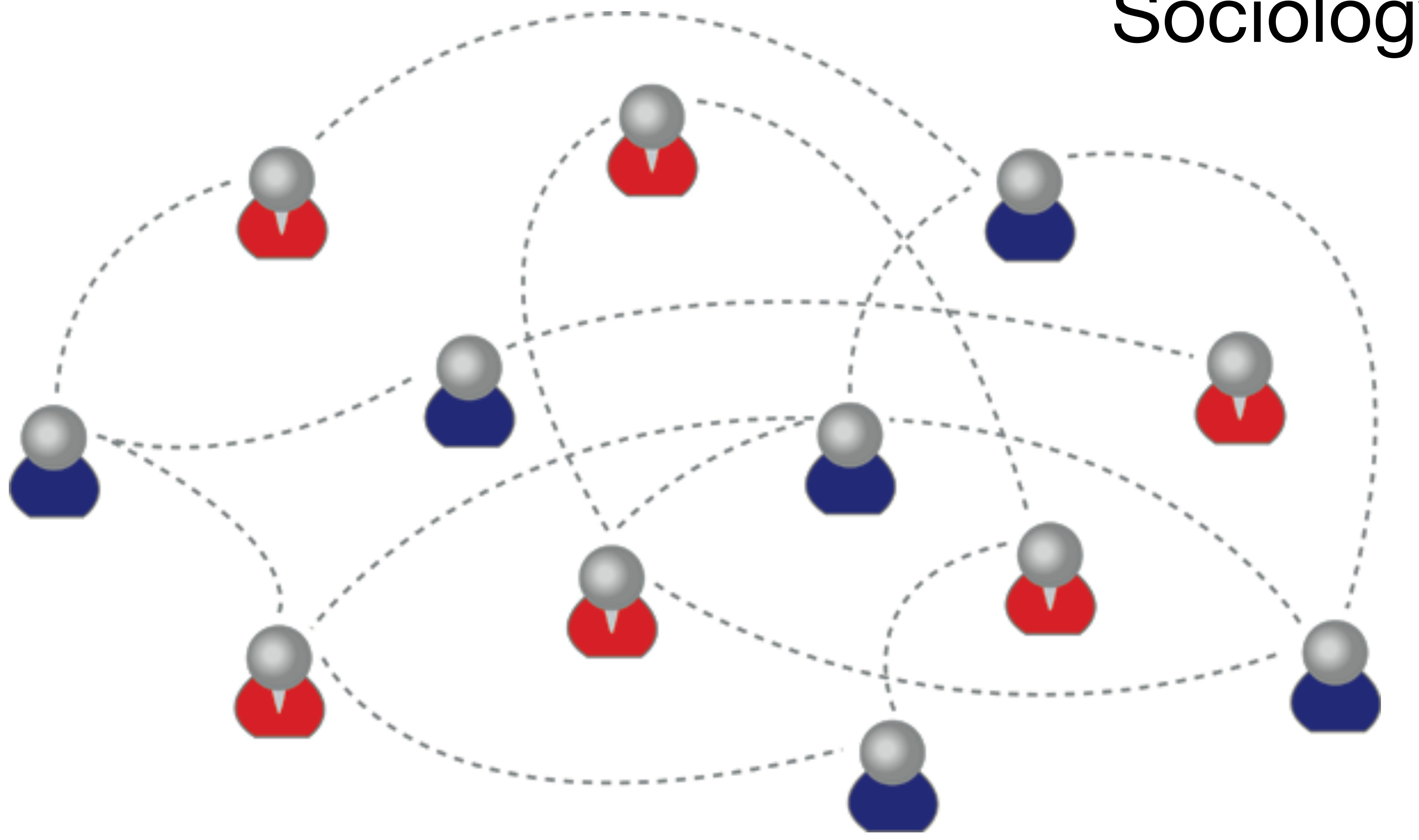
infovis.info > Atlas of cyberspace

Air travel network

Disease spreading

gleamviz.org

Social networks



Information spreading

Disease spreading

Sociology

Applications

Marketing

Vaccine distribution

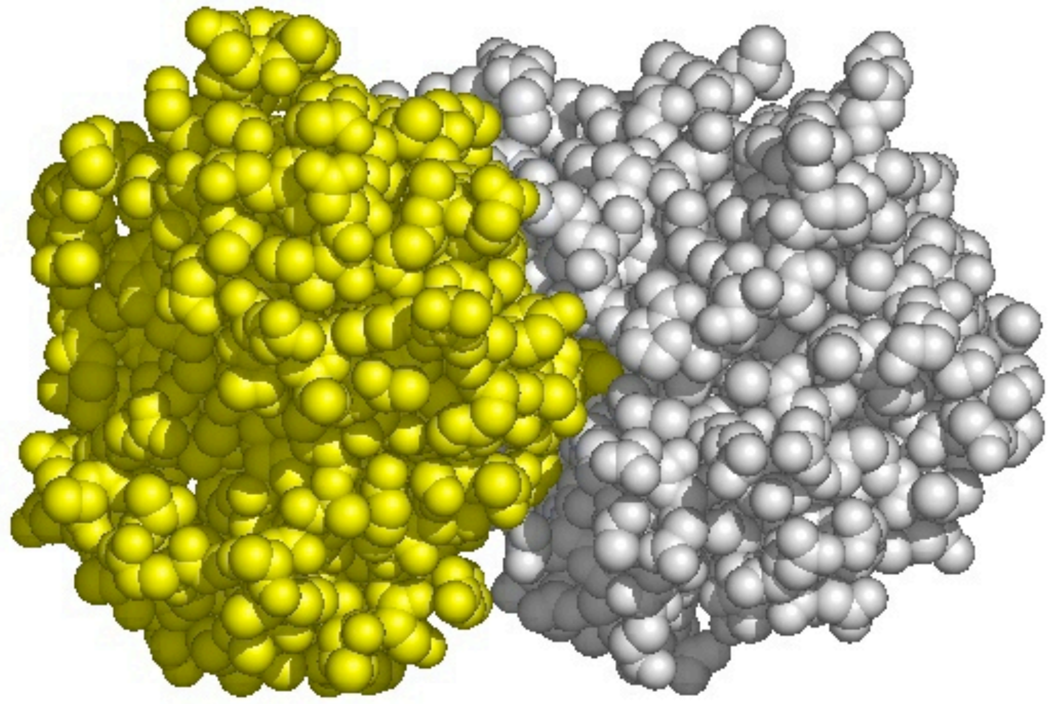
Social media

Emergency response

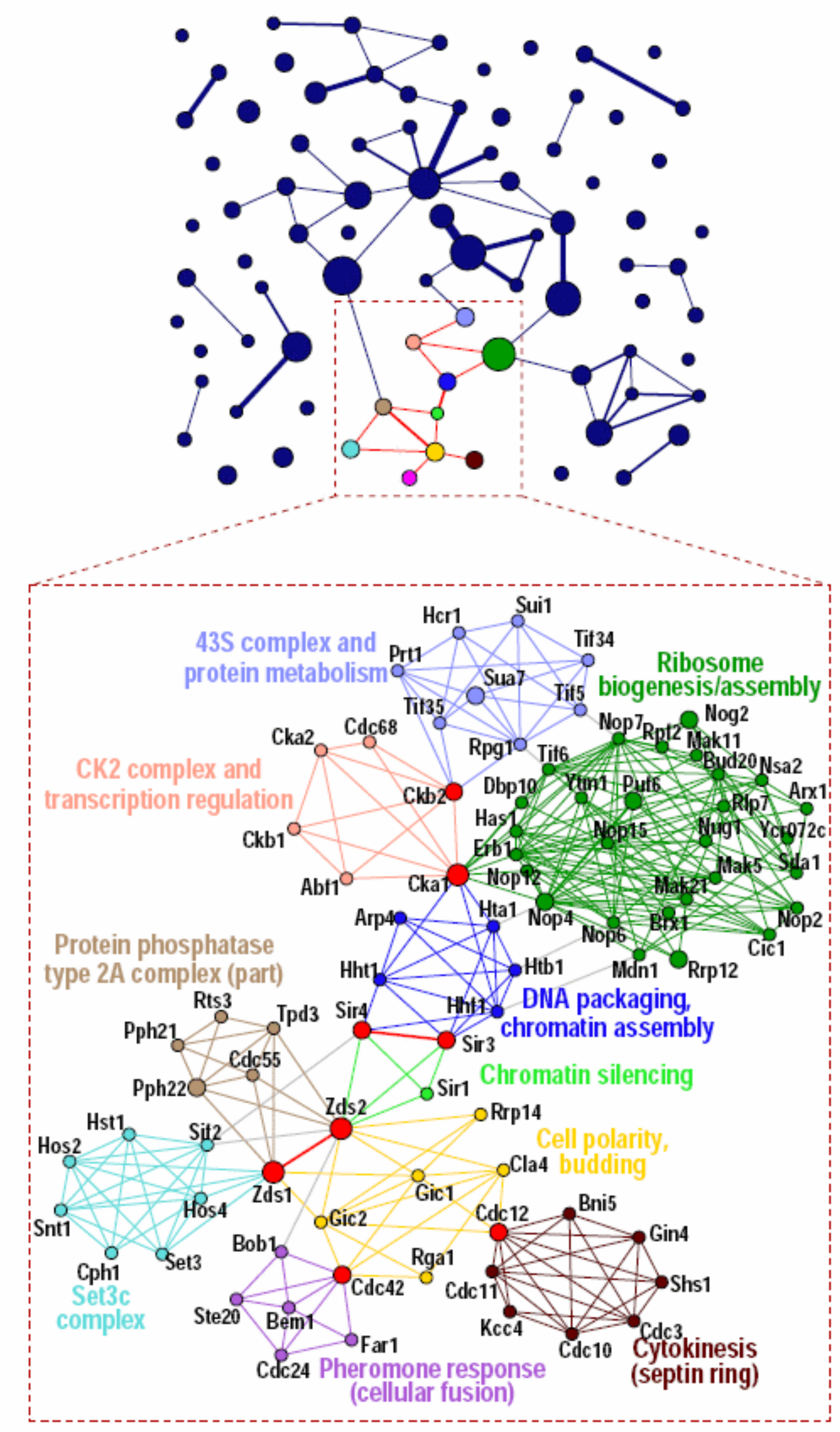
....

Biological networks

Another **HUGE** area

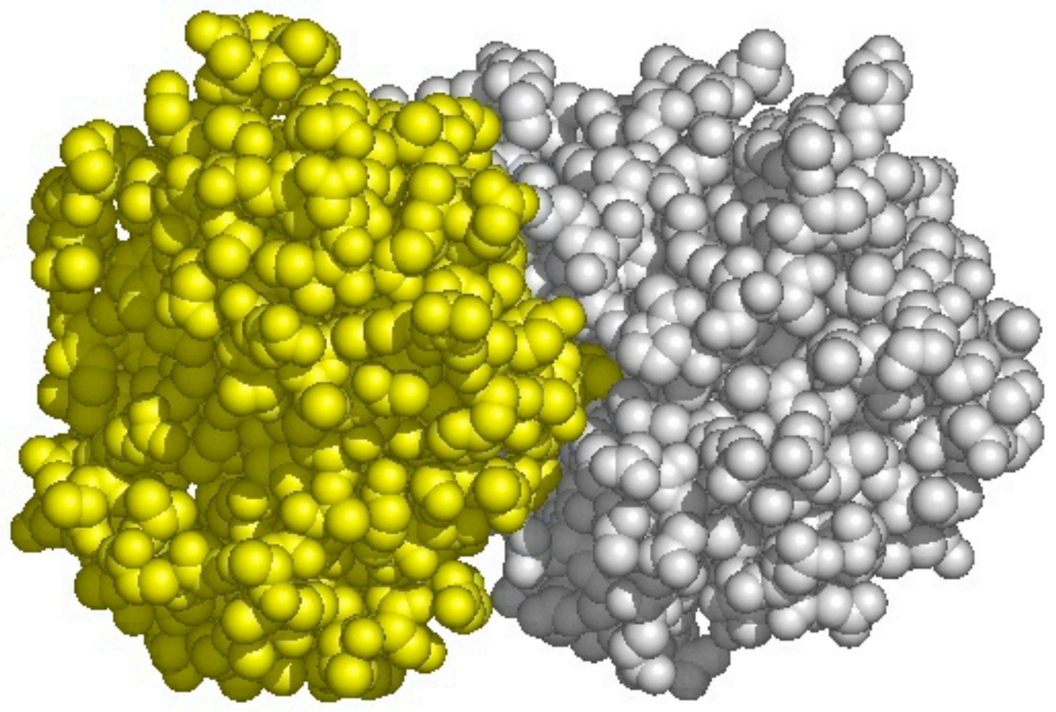


Protein-Protein
Interaction networks

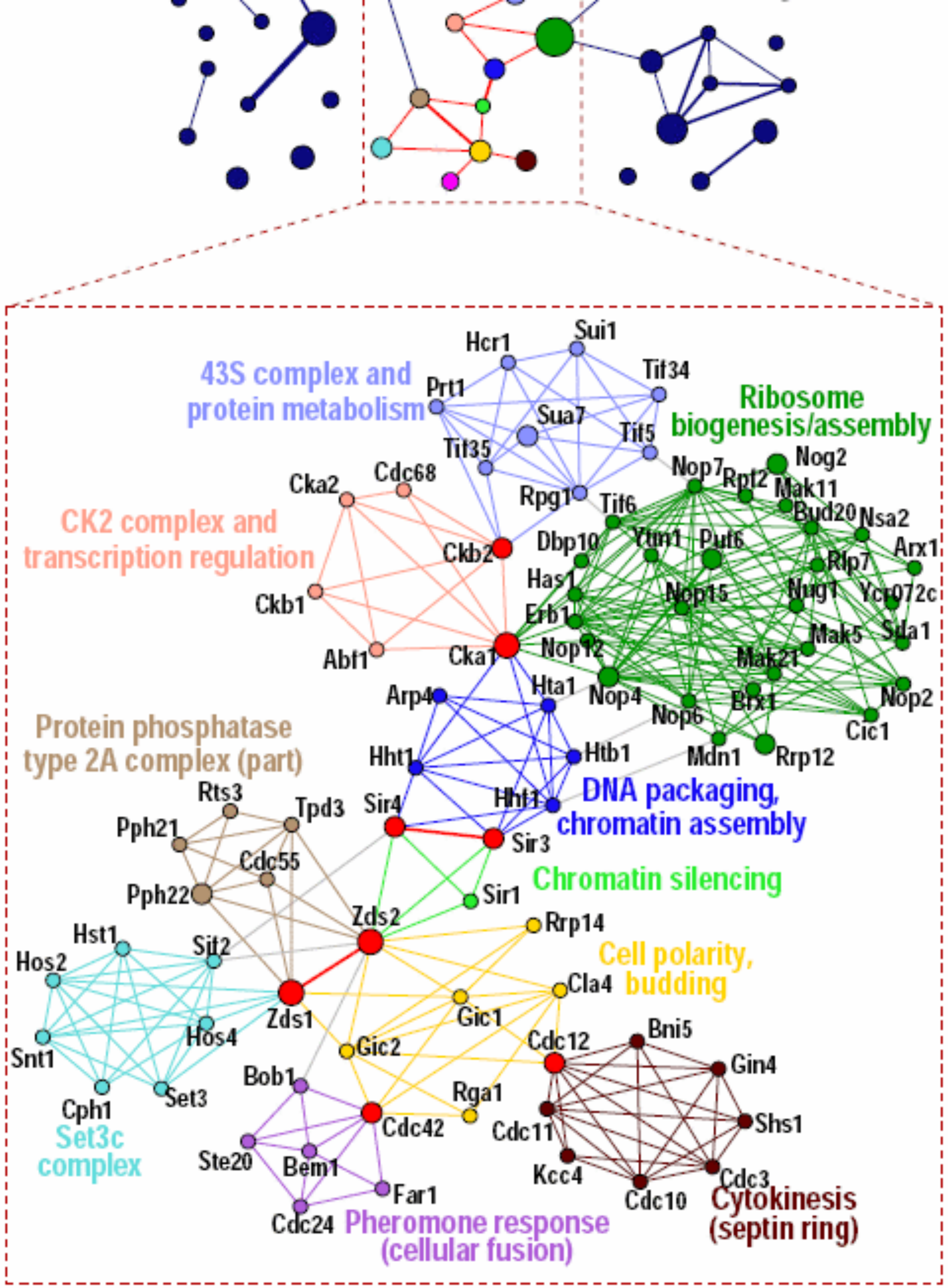


Palla, et al. Nature, 2005

Biological networks



Protein-Protein
Interaction networks

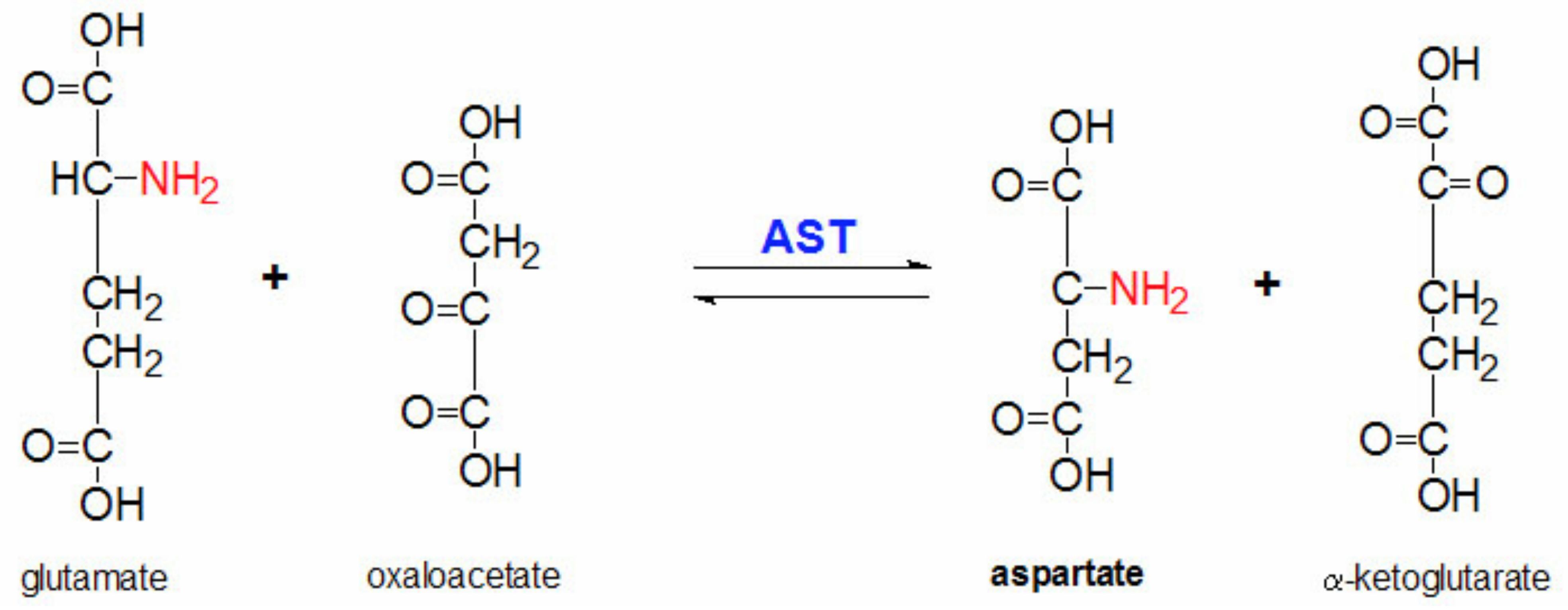


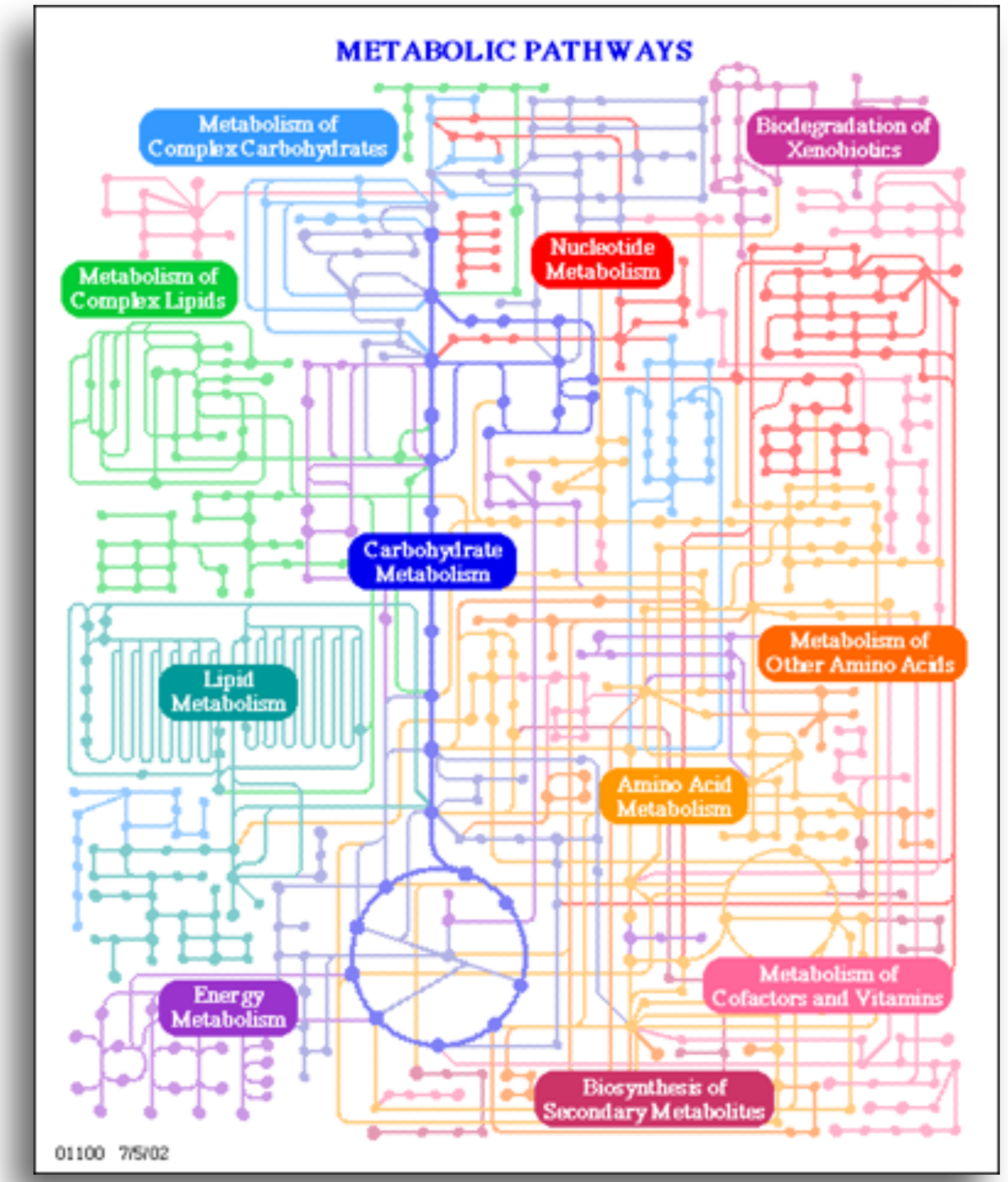
Palla, et al. Nature, 2005

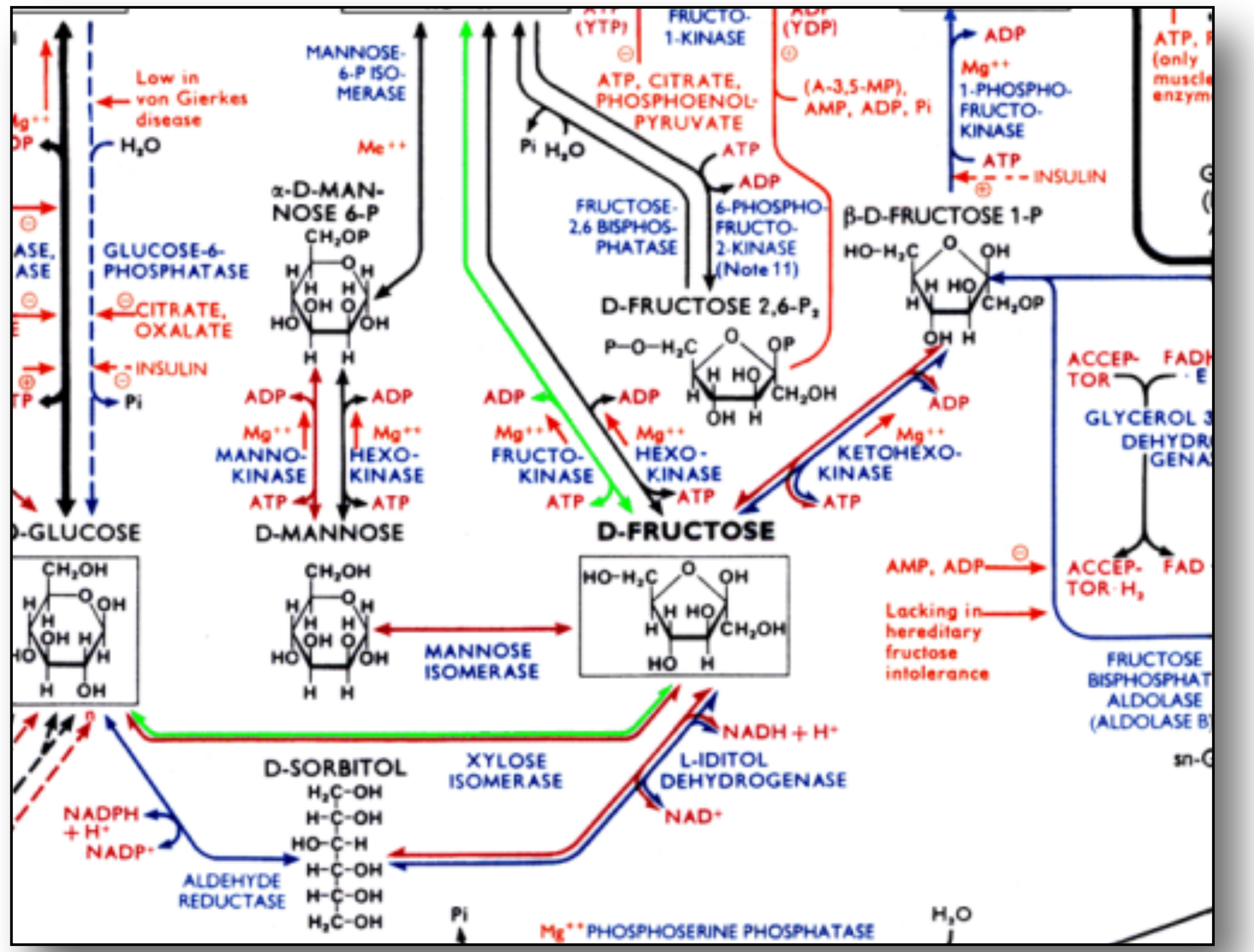
Metabolic networks

Metabolites (chemicals)

Reactions involving metabolites



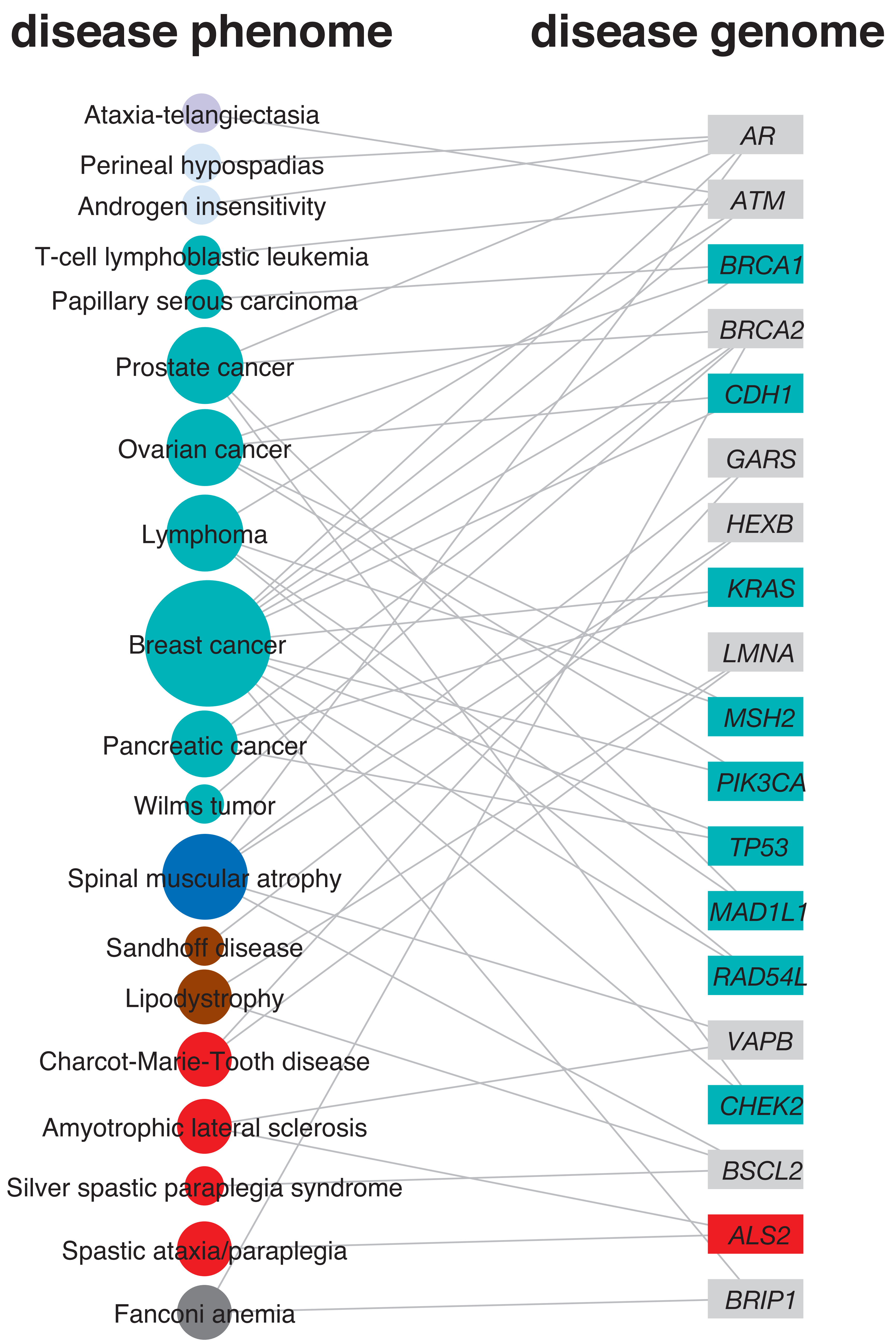




Kyoto Encyclopedia of Genes and Genomes *http://www.expasy.org/tools/pathways/*

"Diseaseome"

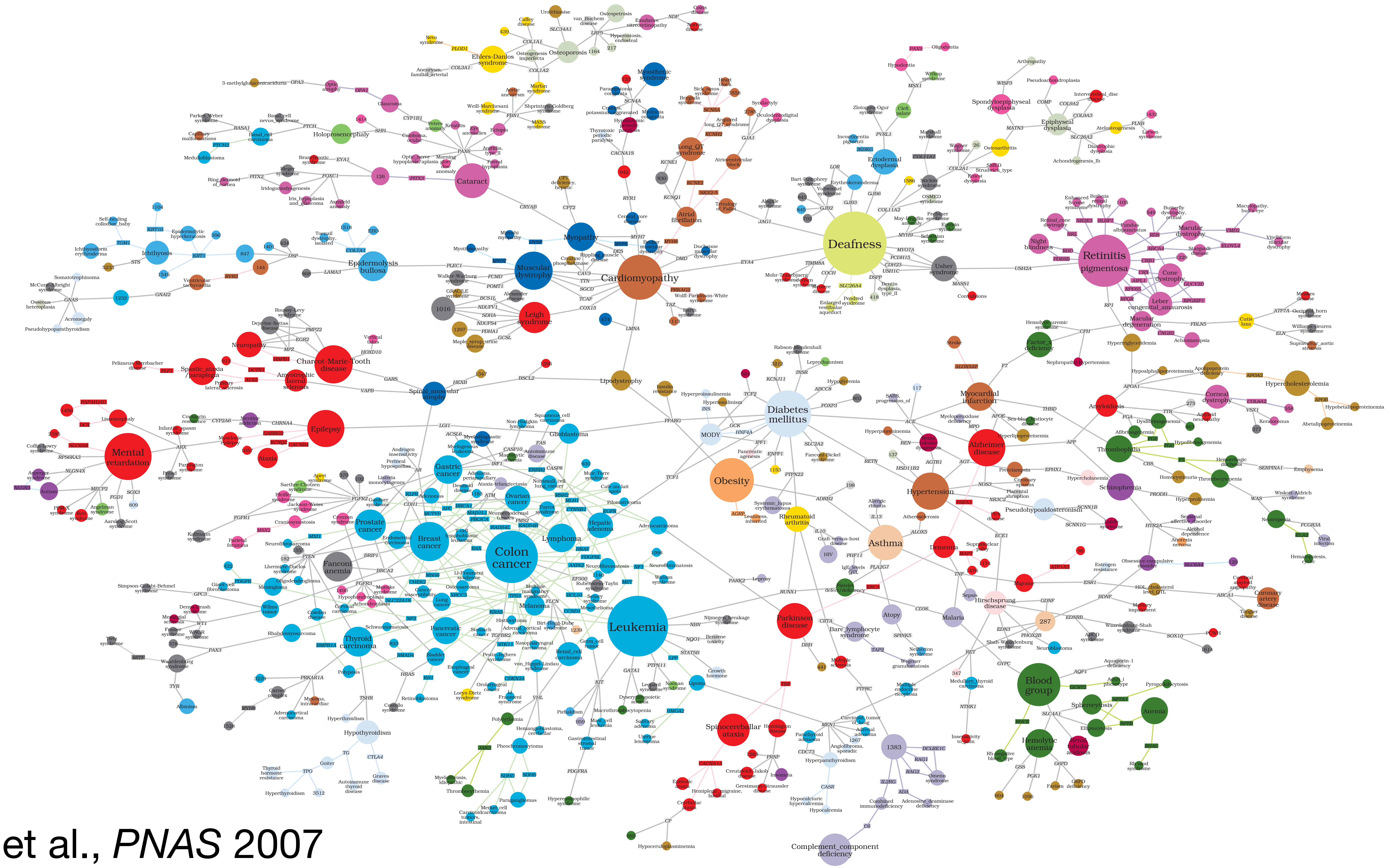
Network between **diseases** and **genes** associated with those diseases



Goh, et al., *PNAS* 2007

The human disease network

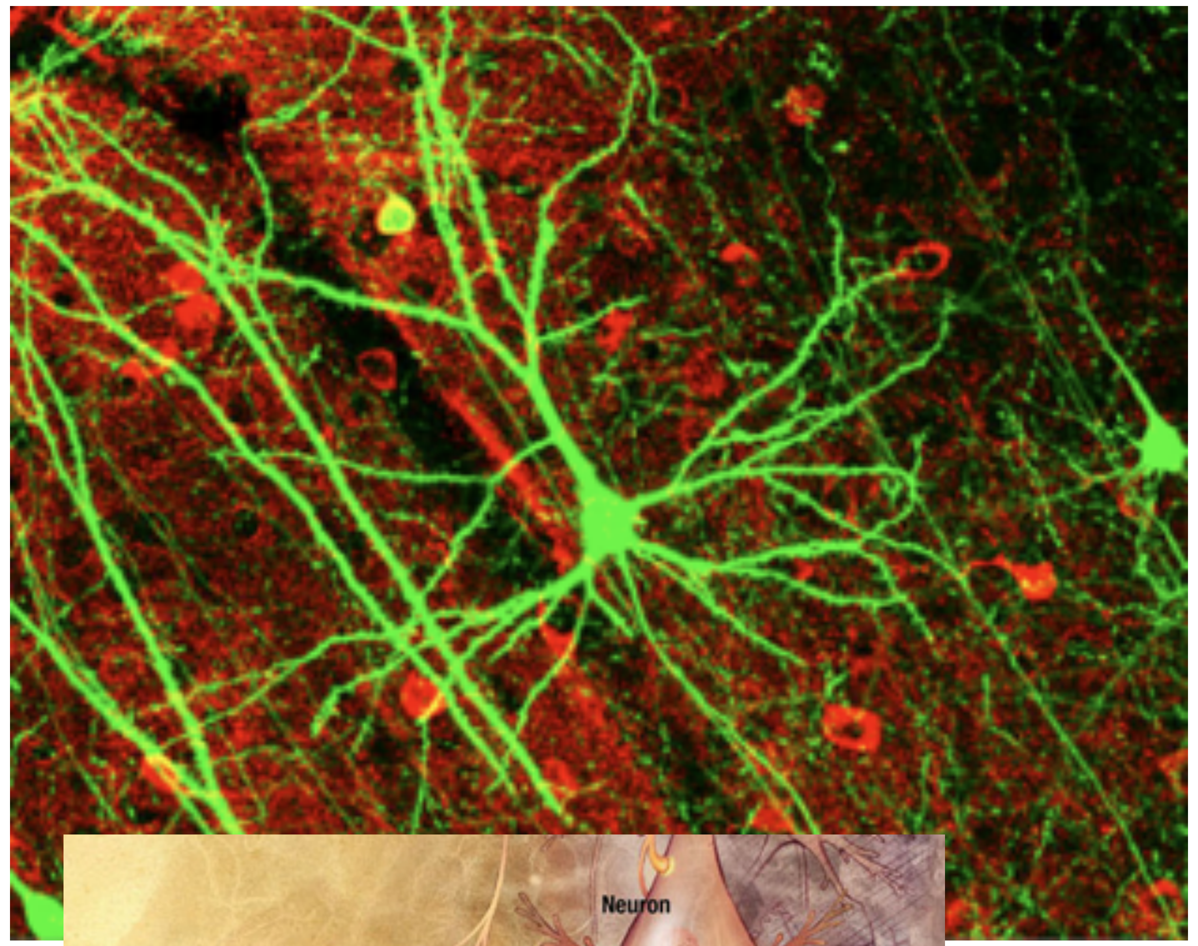
Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L (2007) *Proc Natl Acad Sci USA* 104:8685-8690

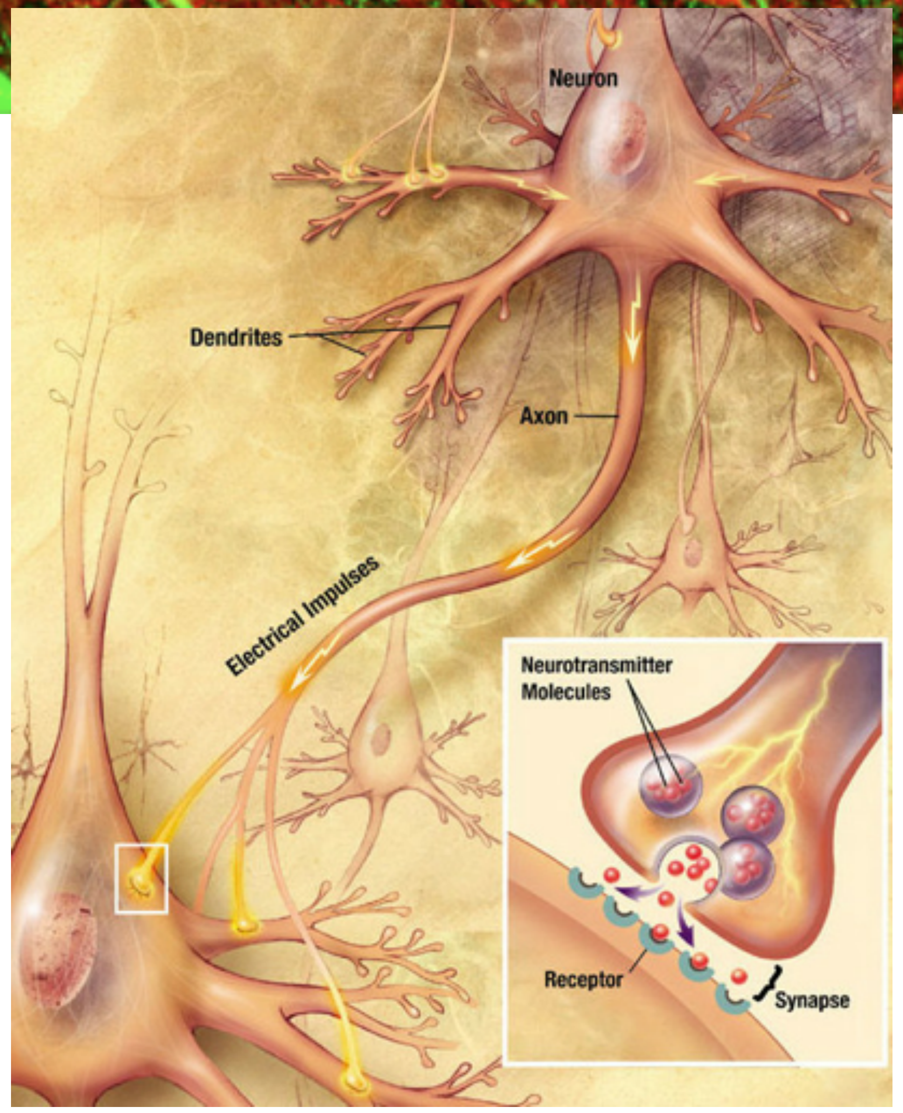


Goh, et al., *PNAS* 2007

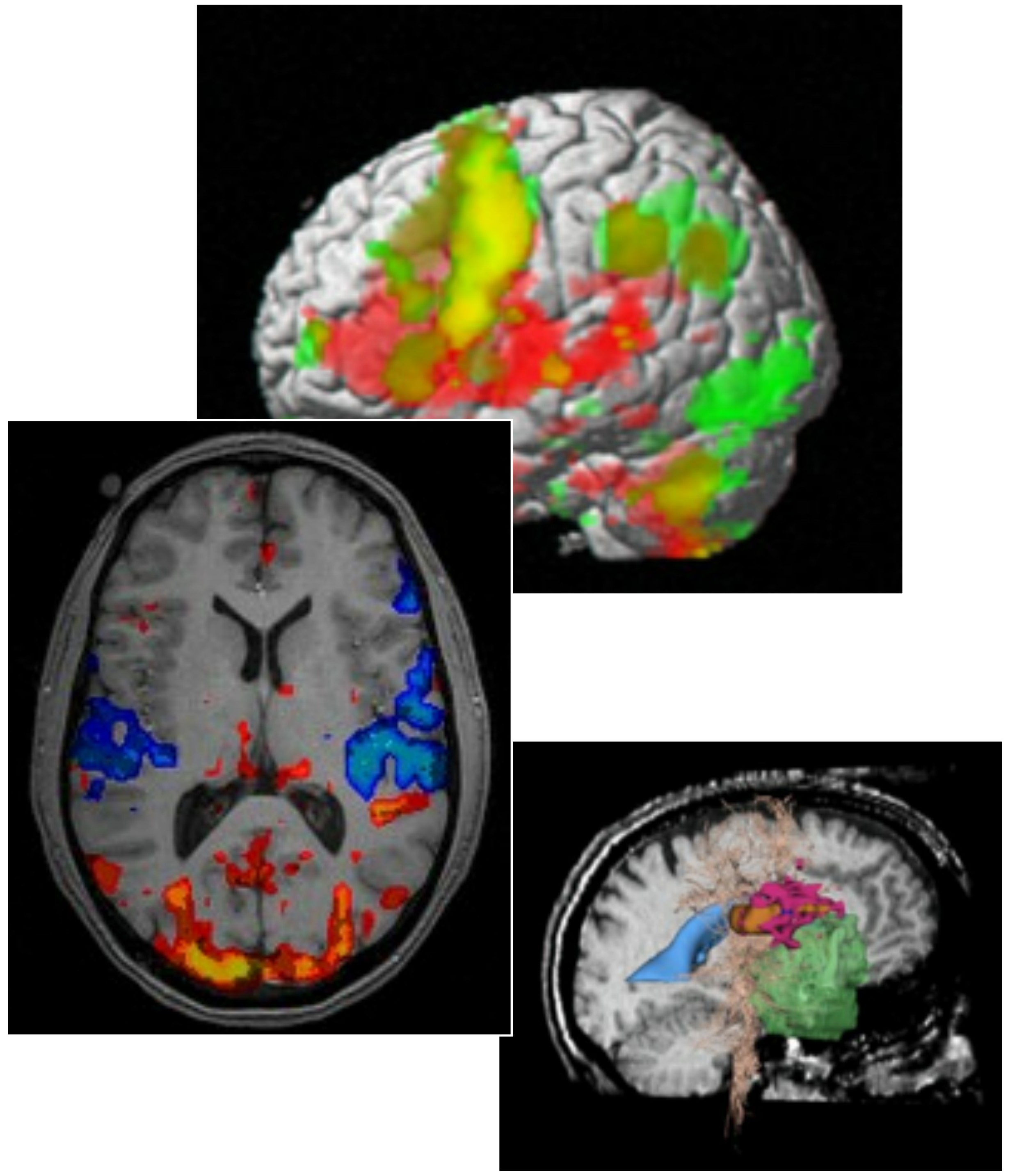
lateral_sclerosis
sclerosis
VAPB
Spinal_muscular atrophy
Insulin resistance
Hyperproinsulinemia
Hyperinsulinism
INS
TCF2
ABCC8
803
FOXP3
Diabetes mellitus
Coumarin resistance
CYP2A6
Nicotine addiction
CHRNA4
Epilepsy
Infantile_spasm syndrome
GABRG2
Myoclonic epilepsy
KCNQ2
CACNB4
1057
Ataxia
ARX
Partington syndrome
Proud syndrome
LGI1
PPARG
MODY
GCK
HNF4A
IPF1
Hyperpro
ACSL6
Myelodysplastic syndrome
Non-Hodgkin lymphoma
Squamous_cell carcinoma
Glioblastoma
Myelogenous leukemia
FAS
CASP10
Macrocytic anemia
Autoimmune disease
Androgen insensitivity
Perineal hypospadias
AR
IRF1
Gastric cancer
ERBB2
CASP8
930
Muir-Torre syndrome
Pancreatic agenesis
ENPP1
SLC2A2
Fanconi-Bickel syndrome
RETN
1153
PTPN22
TCF1
Obesity
Apert syndrome
379
192
Saethre-Chotzen syndrome
Pfeiffer syndrome
Listeria monocytogenes
Adenoma, periampullary
Ataxia-telangiectasia
Desmoid disease
1183
ATM
Nonsmall_cell lung_cancer
Cafe-au-lait spots
198
ADRB2
Allergic rhinitis
KLF6
Adenomas
Ovarian cancer
MSH2
MLH1
FGFR2
Gardner syndrome
CDH1
Pilomatricoma
Systemic_lupus erythematosus
Jackson-Weiss syndrome
APC
BRCA1
Turcot syndrome
CTNNB1
EGFR
MUTYH
MAD1L1
Neuroectodermal tumors
PIK3CA
PMS2
Hepatic adenoma
IL13
AGRP
Leanness, inherited
Rheumatoid arthritis
FGFR1
Crouzon syndrome
Prostate cancer
Craniosynostosis
RAD54L
RAD54B
Adenocarcinoma
MSX2
T-cell lymphoblastic leukemia
Lymphoma
IL10
Kallmann syndrome
Parietal foramina
MXI1
Endometrial carcinoma
Breast cancer
Colon cancer
Graft-versus-host disease
Neurofibrosarcoma
1555
BAX
BRAF
1096
HIV
PTEN
182
BRIP1
PDGFRL
PHF11
Fanconi anemia
AXIN2
Neurofibromatosis
NF1
Neurofibromatosis
422
Lhermitte-Duclos syndrome
BRCA2
MSH6
Li-Fraumeni syndrome
EP300
1140
IgE_levels QTL
PLA2G7
PDGFB
Oligodendroglioma
CHEK2
Rubenstein-Taybi syndrome
MET
Watson syndrome
Leprosy
PARK2
n-Golabi-Behmel syndrome
Giant-cell fibroblastoma
Meningioma
FGFR3
Muenke syndrome
Osteosarcoma
Multiple malignancy syndrome
Platelet defect/deficiency
SNCA
GPC3
1490
Hypochondroplasia
Cancer susceptibility
XRCC3
TP53
BCL10
Sezary syndrome
RUNX1
SLC22A18
Lung cancer
FLCN
Denys-Drash syndrome
Wilms tumor
Achondroplasia
Melanoma
Mesothelioma
Cervical carcinoma
KRAS
CCND1
Mesangial sclerosis
WT1
Cowden disease
NF2
Histiocytoma
Birt-Hogg-Dube syndrome
Leukemia
NBN
Nijmegen_breakage syndrome
Parkinson disease
CIITA
Bare_lymphocyte syndrome
Frasier syndrome
WAGR syndrome
Schwannomatosis
Pancreatic cancer
Adrenal_cortical carcinoma
1239
Thyroid carcinoma
Rhabdomyosarcoma
Stomach cancer
TGFBR2
NQO1
Benzene toxicity
378
833
Nasopharyngeal carcinoma
Germ_cell tumor
DBH
STK11
Renal_cell carcinoma
STAT5B
PAX3
BMPR1A
Bladder cancer
Peutz-Jeghers syndrome
TAP2
MITF
SMAD4
Esophageal cancer
LPP
Waardenburg syndrome
von_Hippel-Lindau syndrome
GATA1
PTPN11
Growth hormone
Multiple sclerosis
Polyposis
HRAS
CDKN2A
441
3229
PRKAR1A
RB1
KIT
Lipoma
PTPRC
TYR
Carney complex
Orolaryngeal cancer
Leopard syndrome
Noonan syndrome
TBP
Myxoma, intracardiac
TSHR
Loeys-Dietz syndrome
Li Fraumeni syndrome
VHL
Dyserythropoietic anemia
Albinism
MYH8
Costello syndrome
Retinoblastoma
Piebaldism
Macrothrombocytopenia
Adrenocortical carcinoma
Hyperthroidism
Polycythemia
959
Mast_cell leukemia
HMGA2
Carcinoid_tumor of_lung
1528
Hypothyroidism
Hemangioblastoma, cerebellar
Salivary adenoma
Spinocereballar ataxia
Huntington disease
MEN1
Adrenal adenoma
Parathyroid adenoma
1267
Gastrointestinal stromal tumor
Uterine leiomyoma
Angiofibroma, sporadic
TG
CTLA4
JAK2
Pheochromocytoma
CDC73
Goiter
1263
PRNP
Hyperparathyroidism
Thyroid hormone resistance
TPO
PDGFRA
CACNA1A
Graves disease
Autoimmune thyroid disease
Myelofibrosis, idiopathic
SDHD
SDHB
Creutzfeldt-Jakob disease
Insomnia
Episodic ataxia
CASR
Hyperthyroidism
3512
Thrombocythemia
Hemiplegic_migraine, familial
Gerstmann-Straussler disease
Paragangliomas
Hypereosinophilic syndrome
Hypocalciuric hypercalcemia
Cerebellar ataxia
Merkel cell

Neuroscience



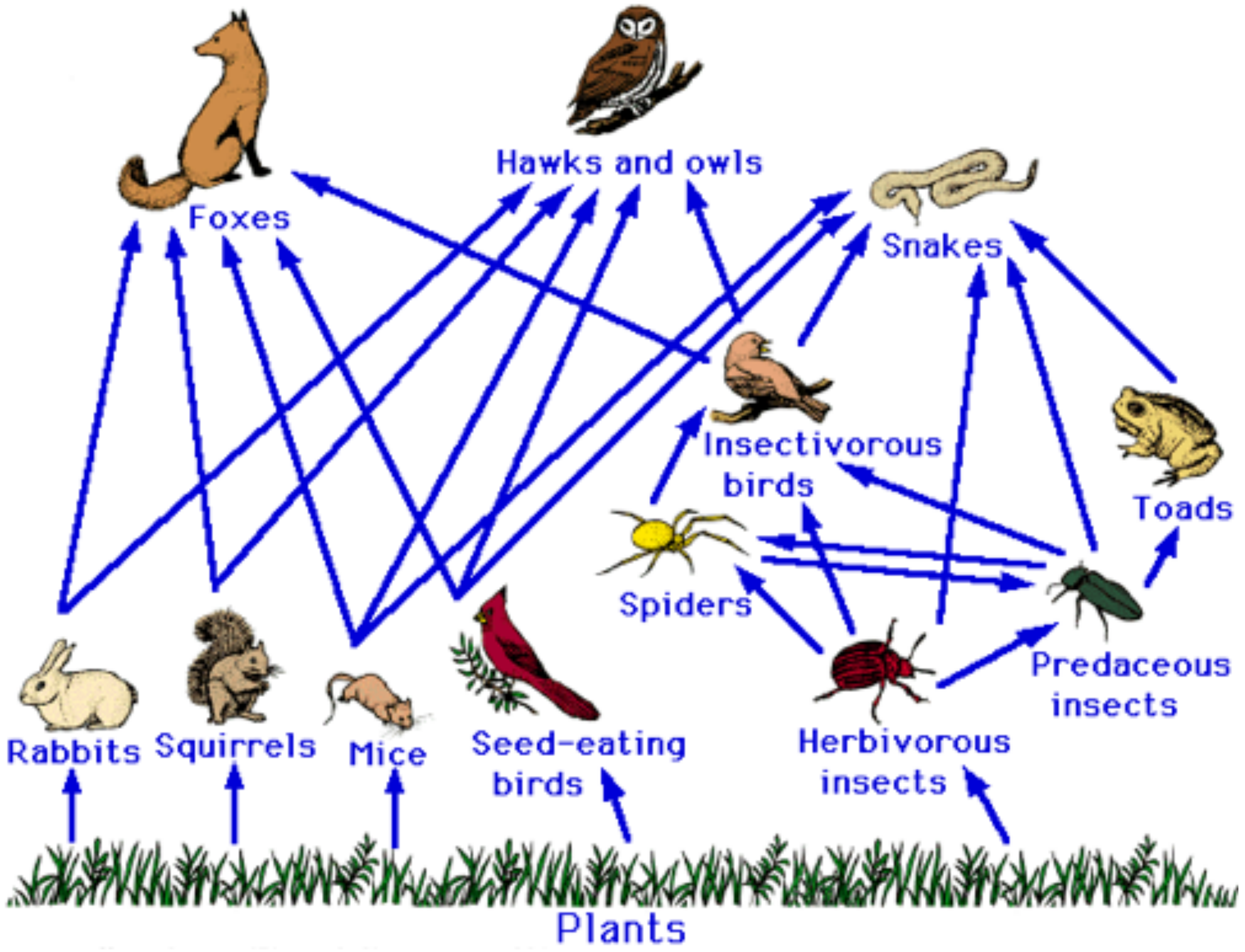


Networks from **Neuroimaging**

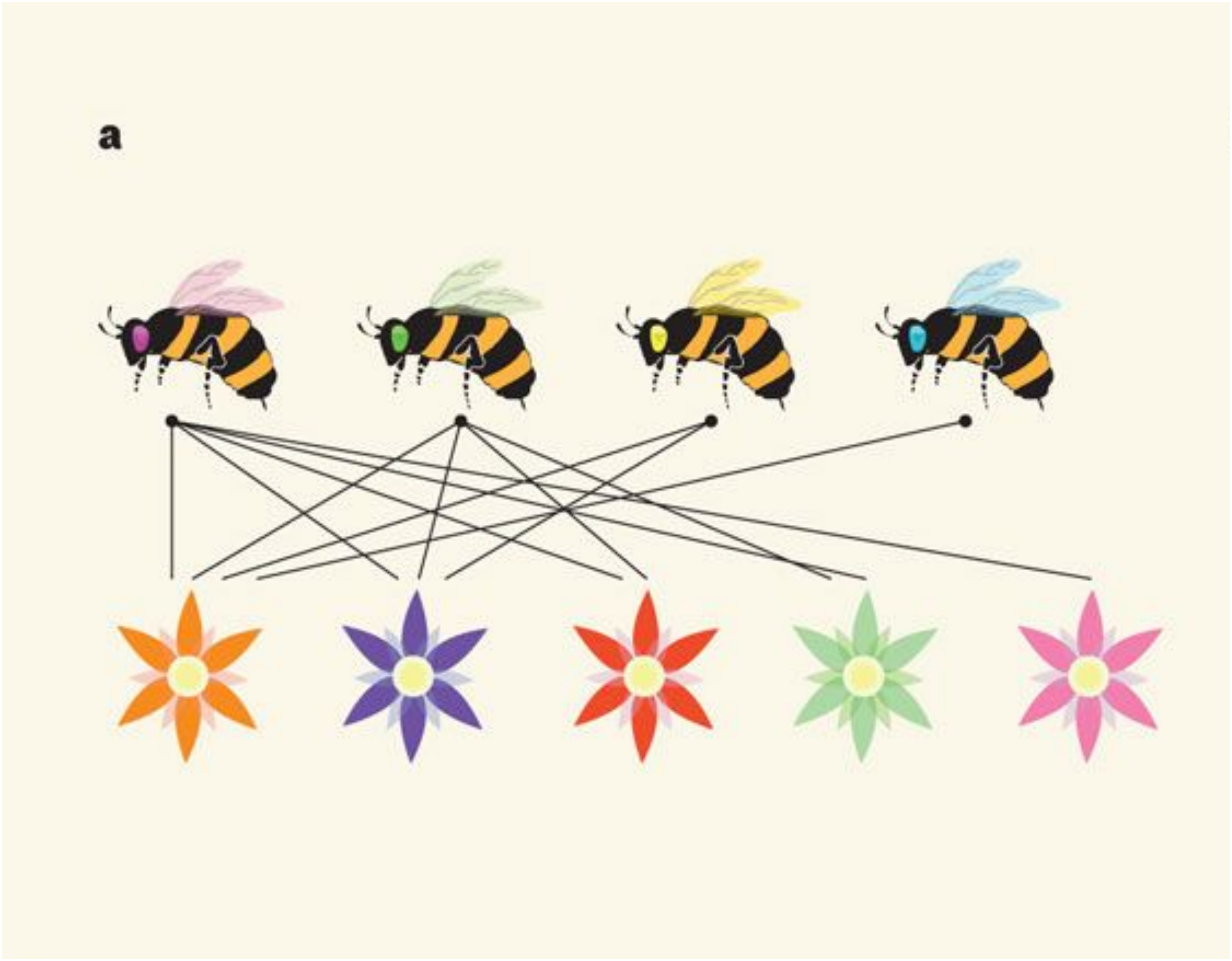


Ecology

Food webs



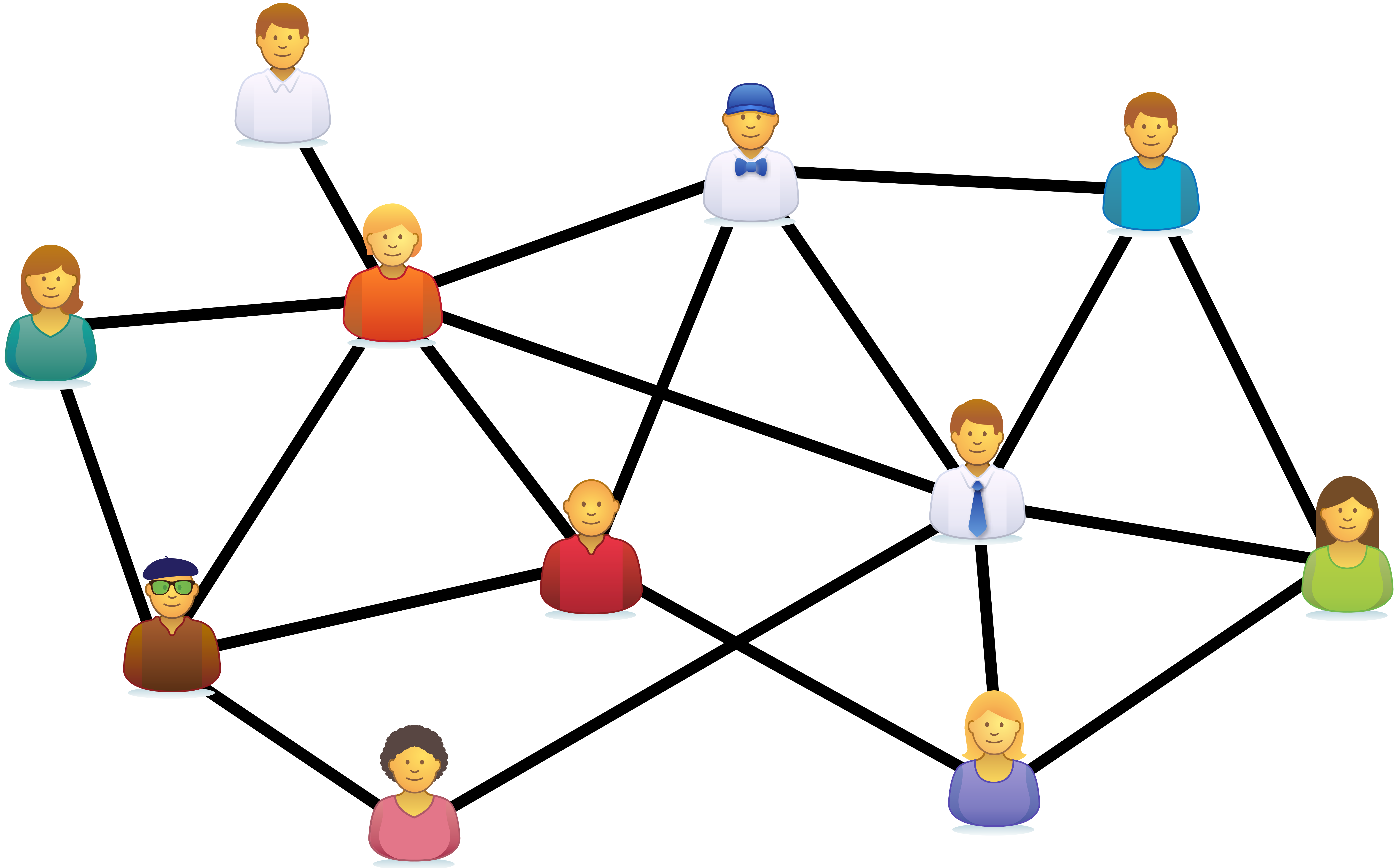
Mutualistic Networks



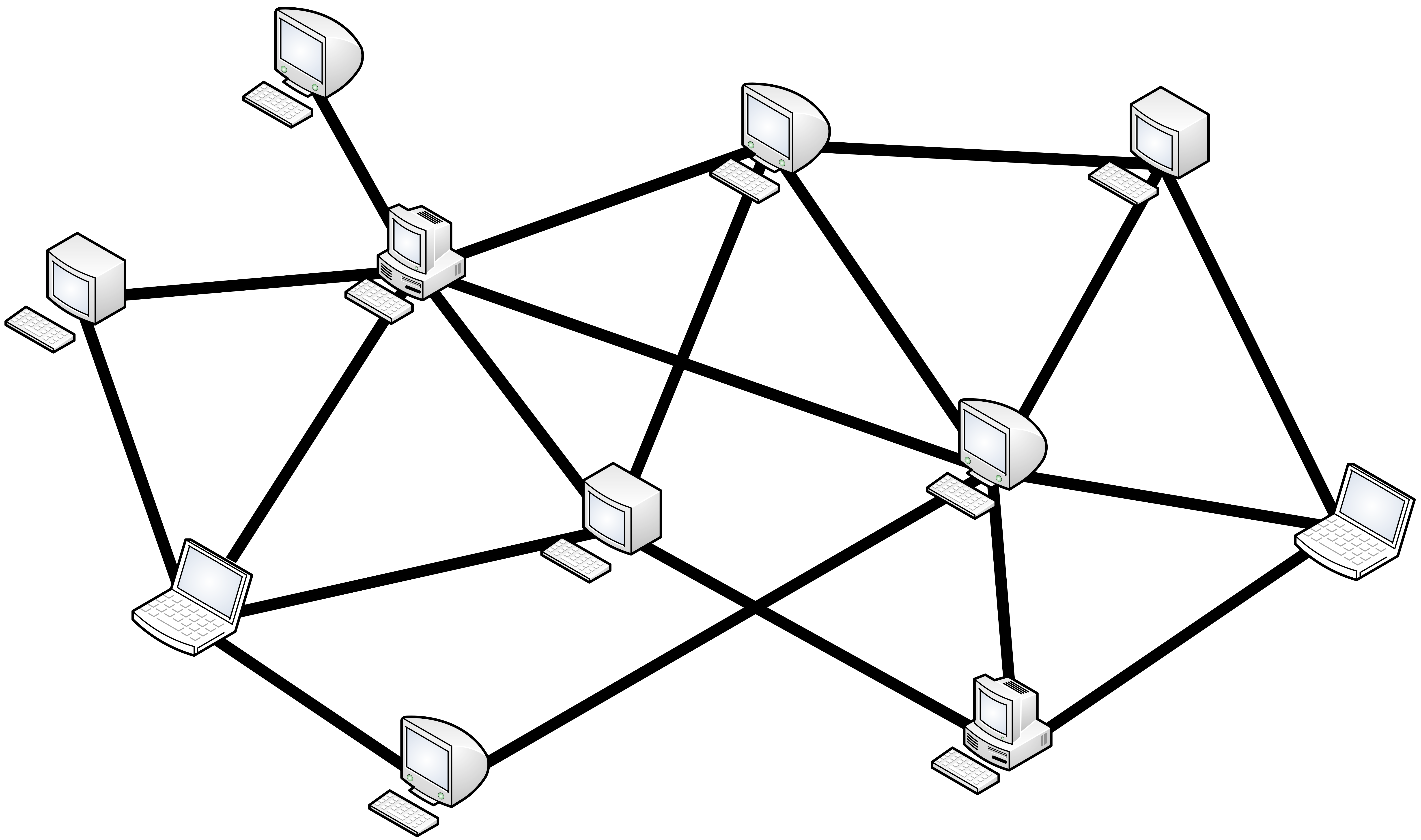
Allesino 2012

Summary so far

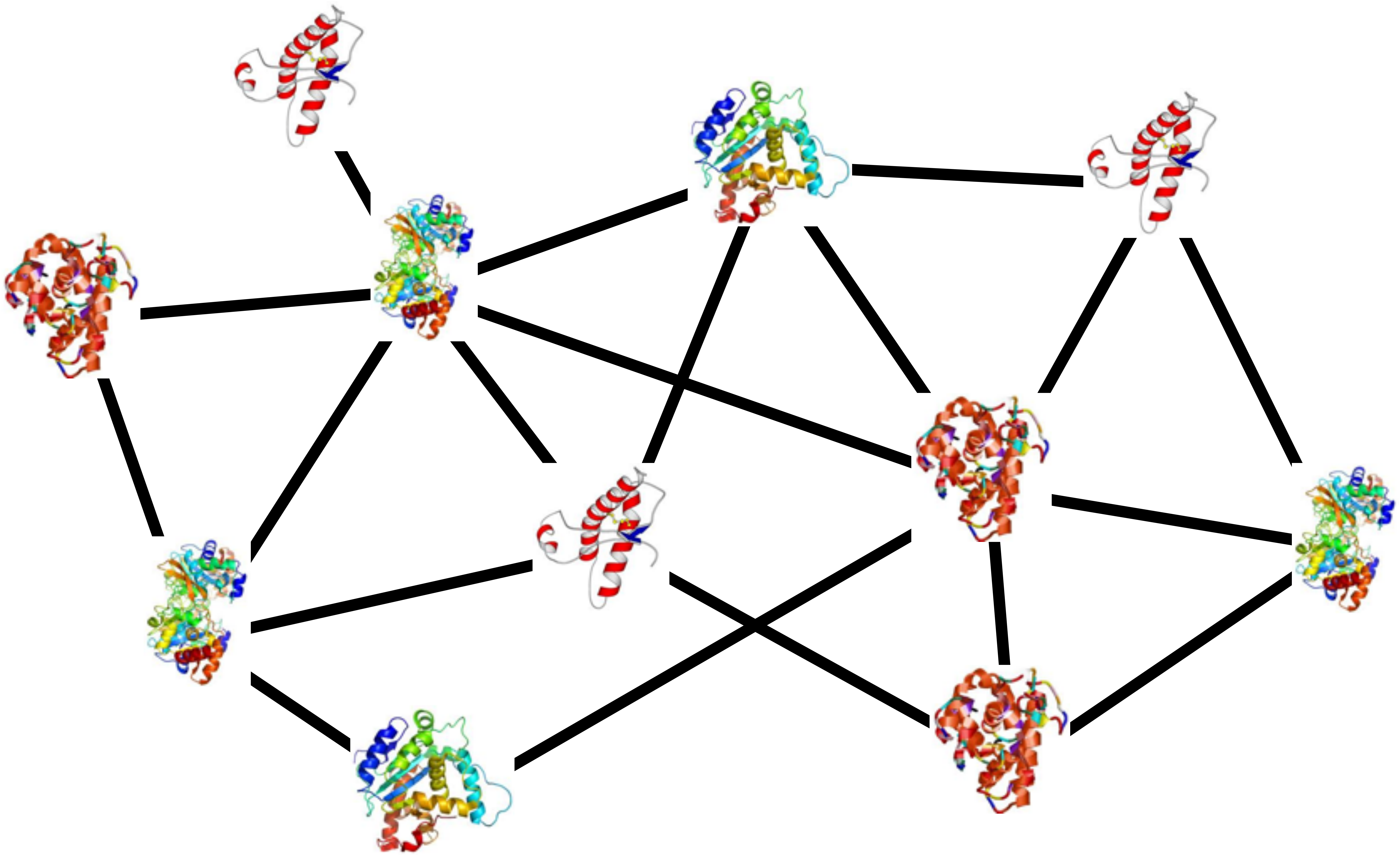
Networks



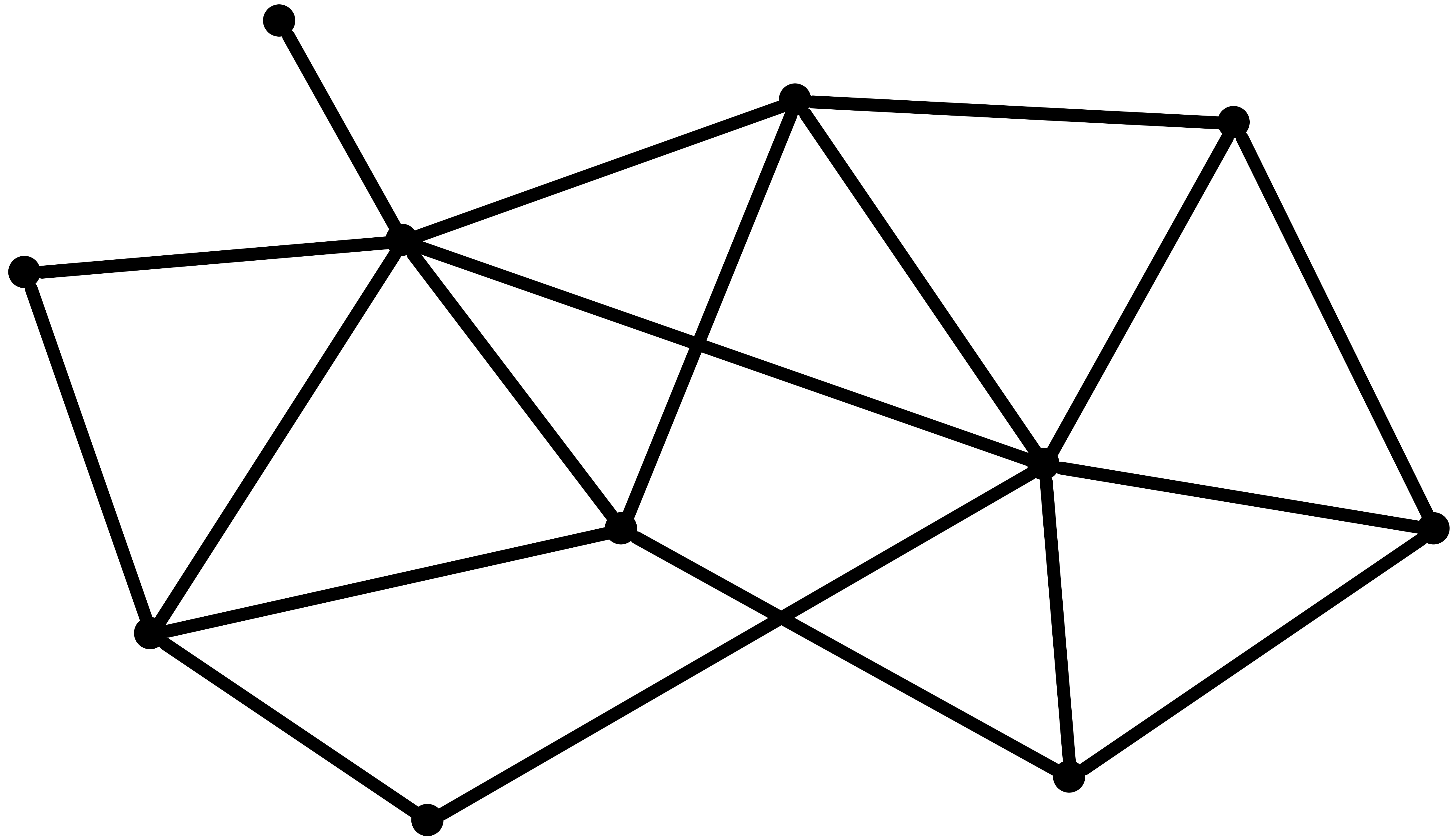
Networks



Networks



Networks

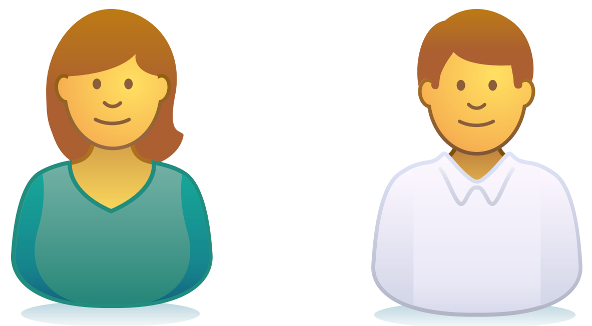
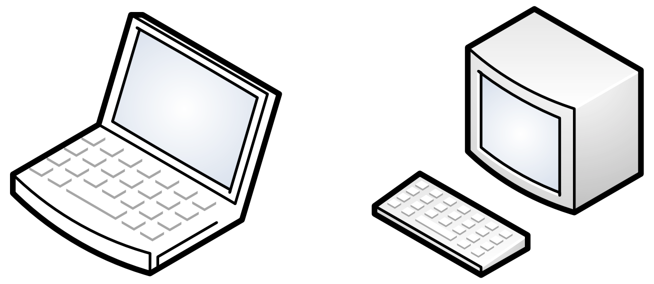
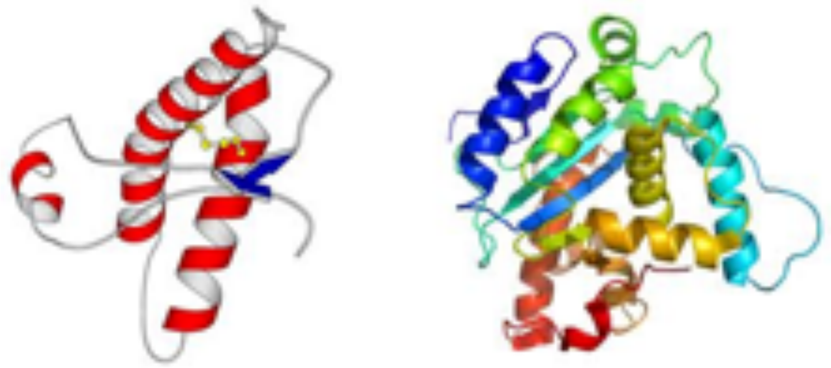


very

general

entities and the **relationships** between them

Networks

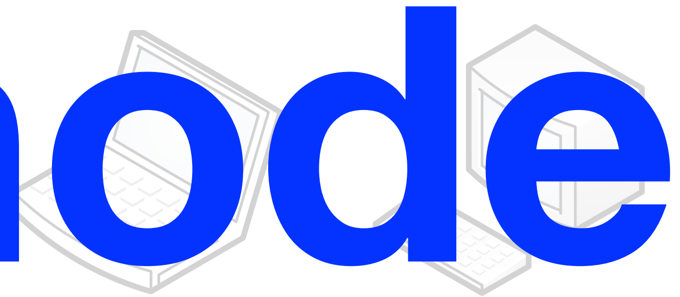
entities	relationships
	friendship, family, sexual
	transmit data, shared power
	bind together, signal transduction

Networks

entities

relationships

friendship, family, sexual



nodes

links

transmit data, shared power

bind together, signal transduction

Why study networks?

Simple components

↓

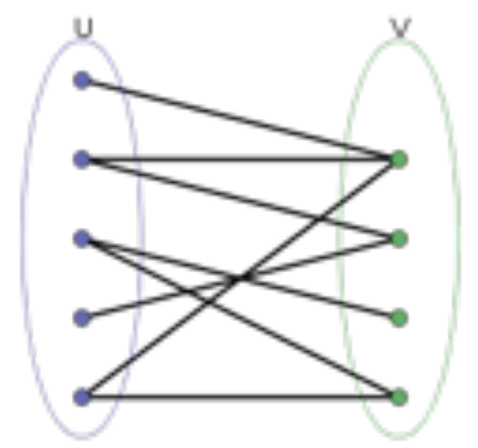
Interacting

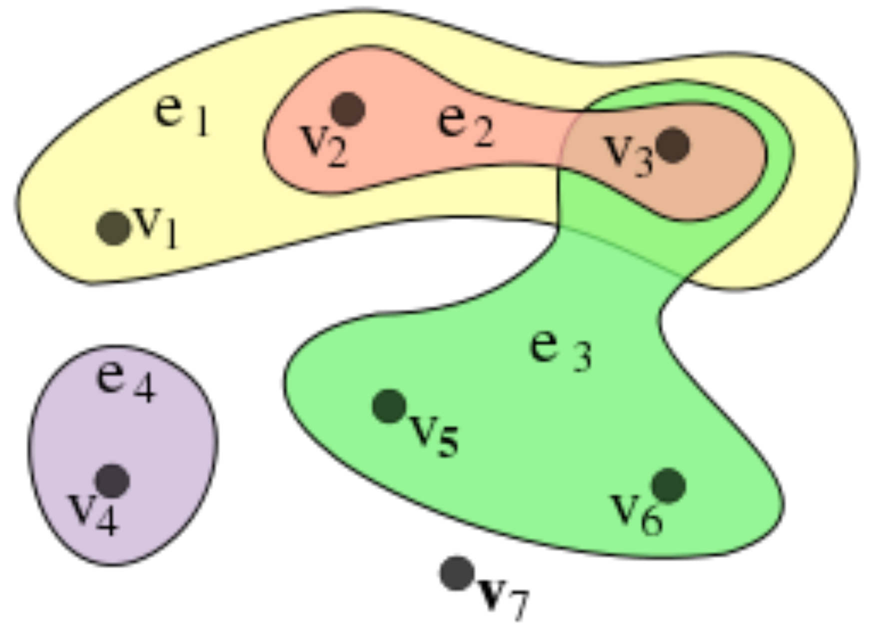
↓

Complex systems

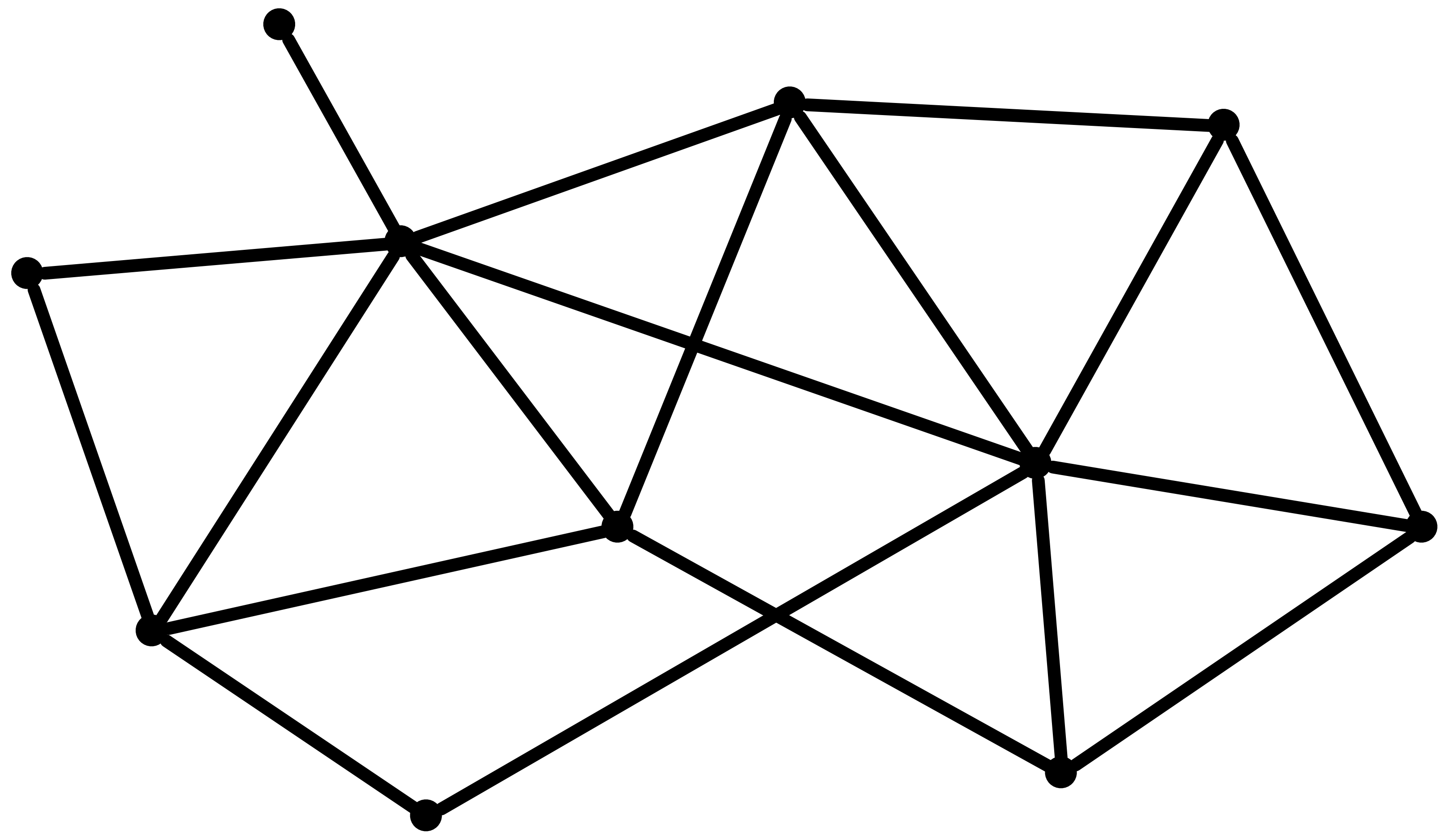
Many complex systems are amenable to network representations

Types of networks and subnetworks





Basic network



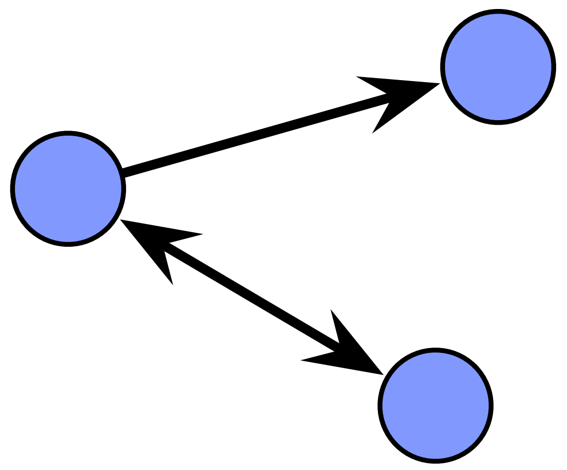
Static network

Links are bidirectional

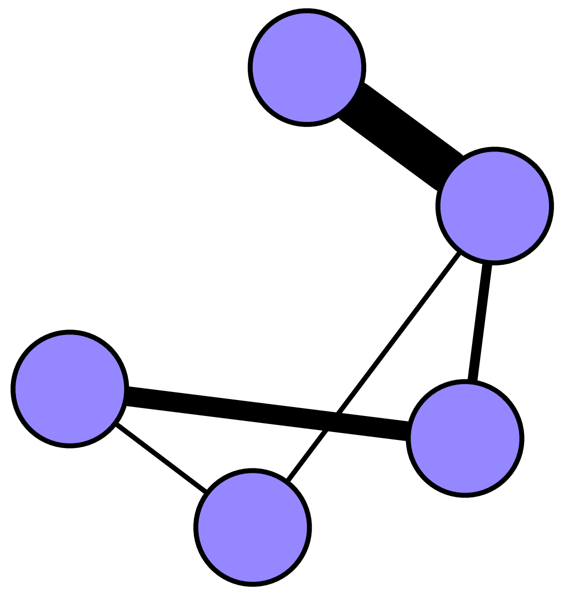
Links equal "strength"

Types and Generalizations

directed network



weighted network



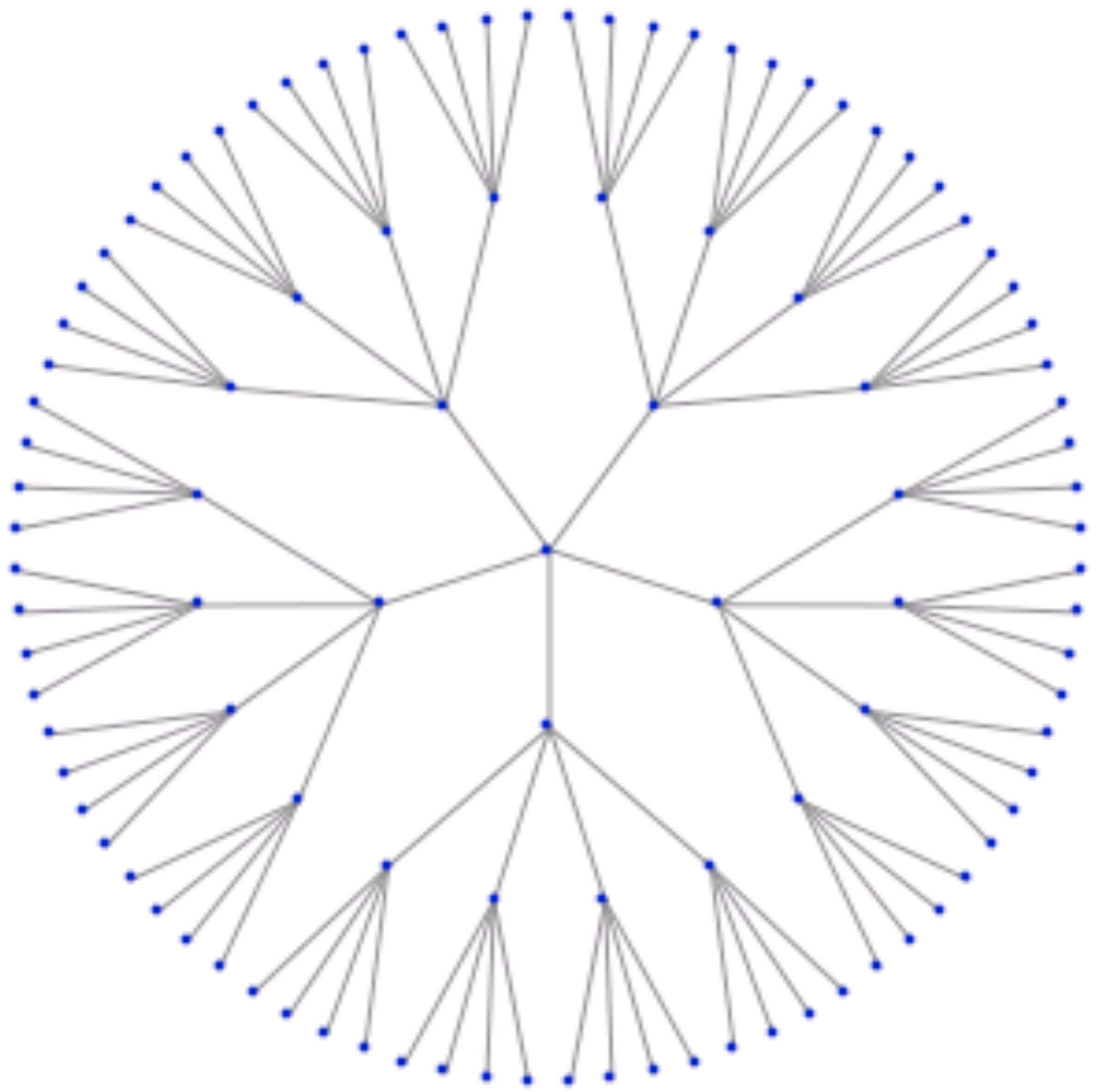
Temporal network

Multilayer/multiplex network

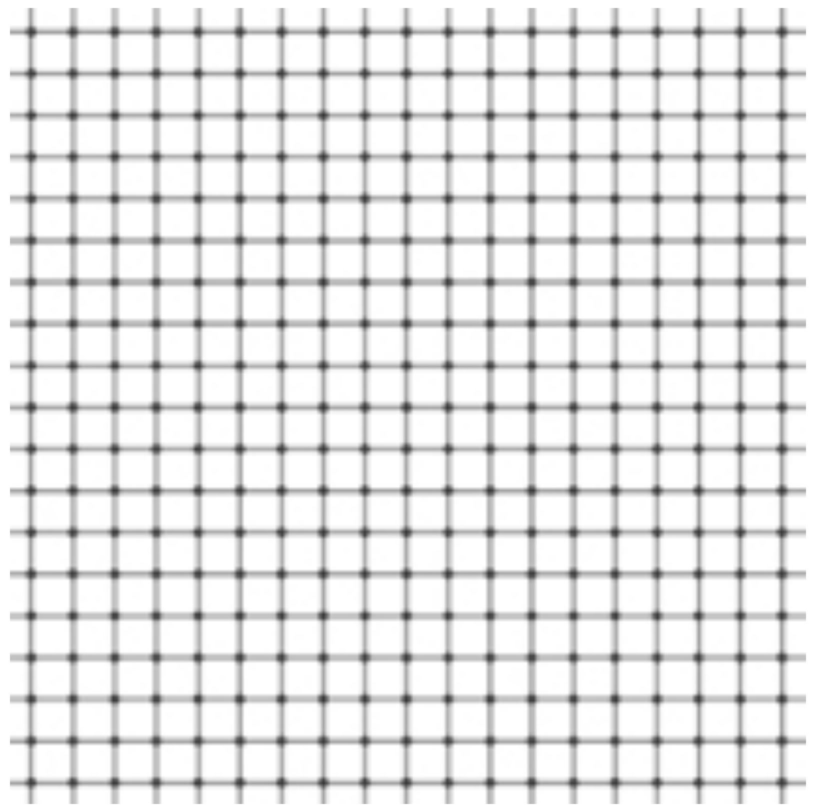
Network zoology

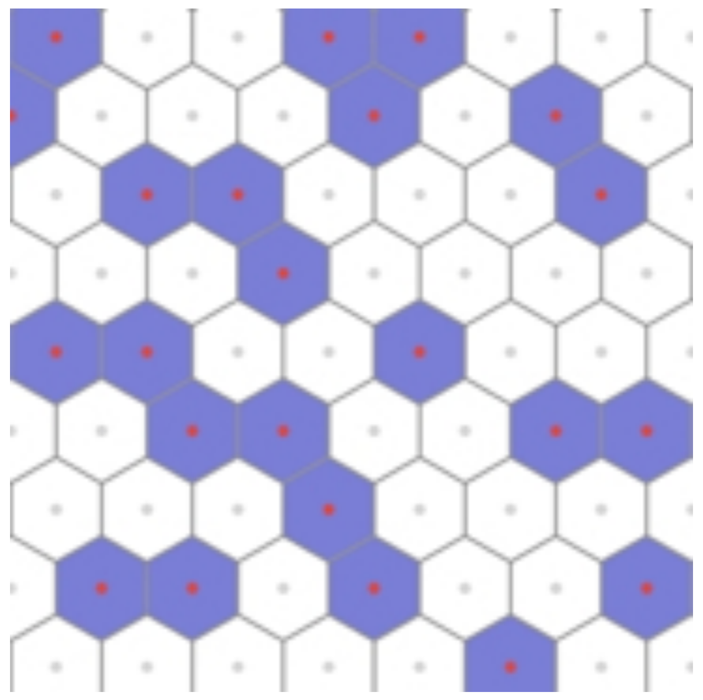
Trees

Networks with **no loops**



Lattices

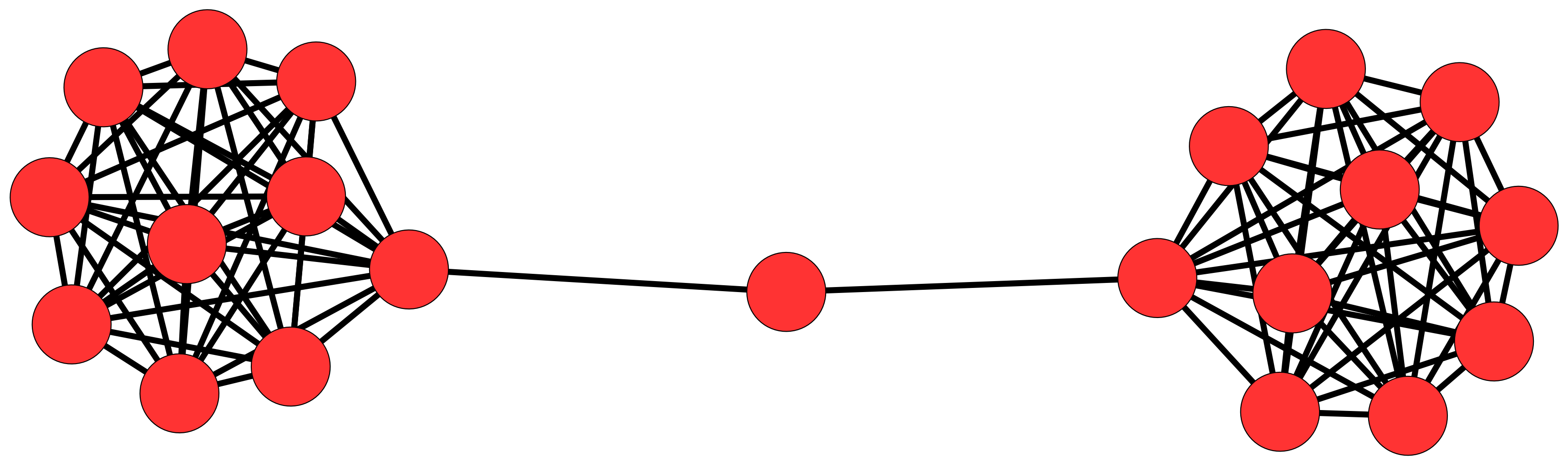


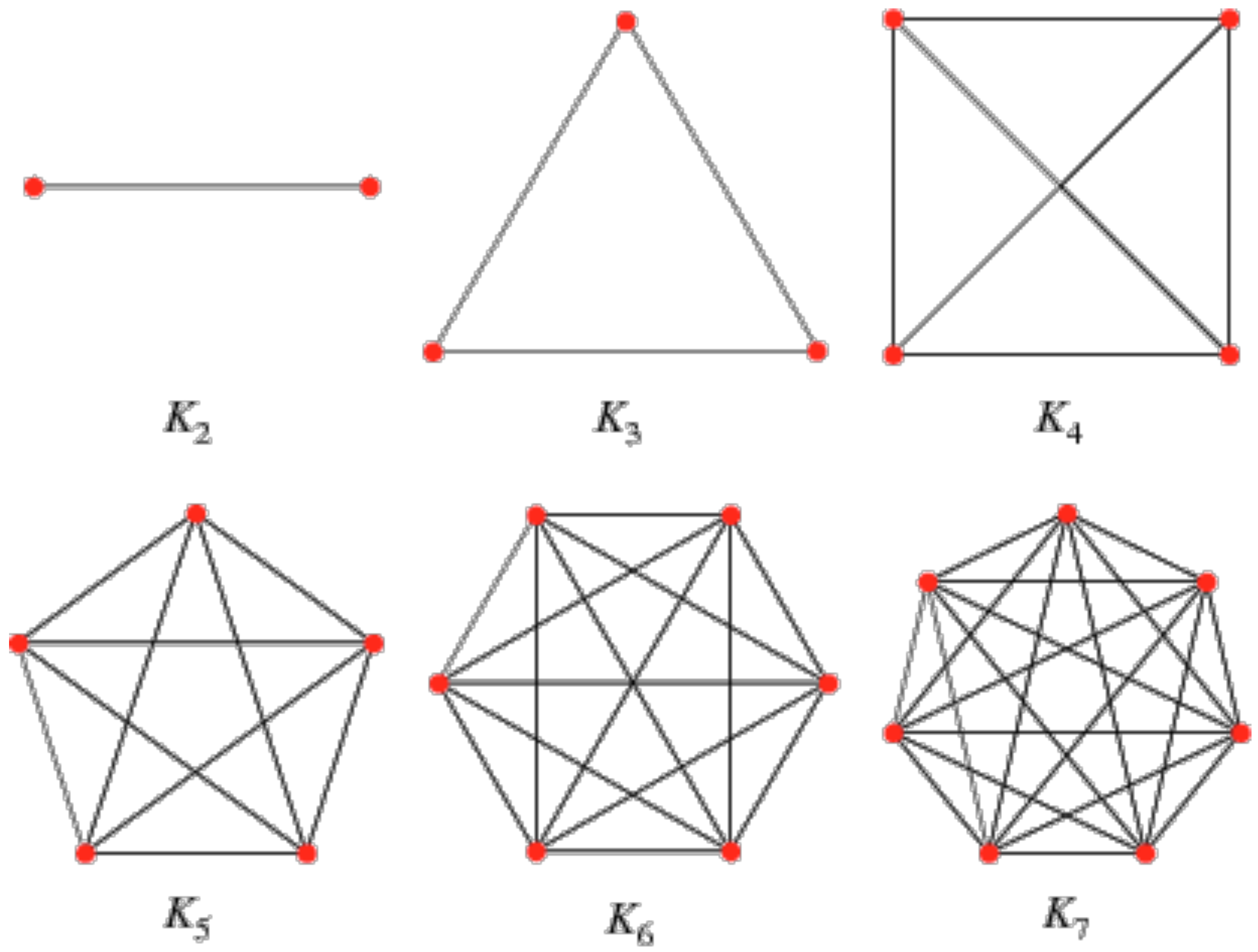


Network **zoology**

Complete graph

Clique





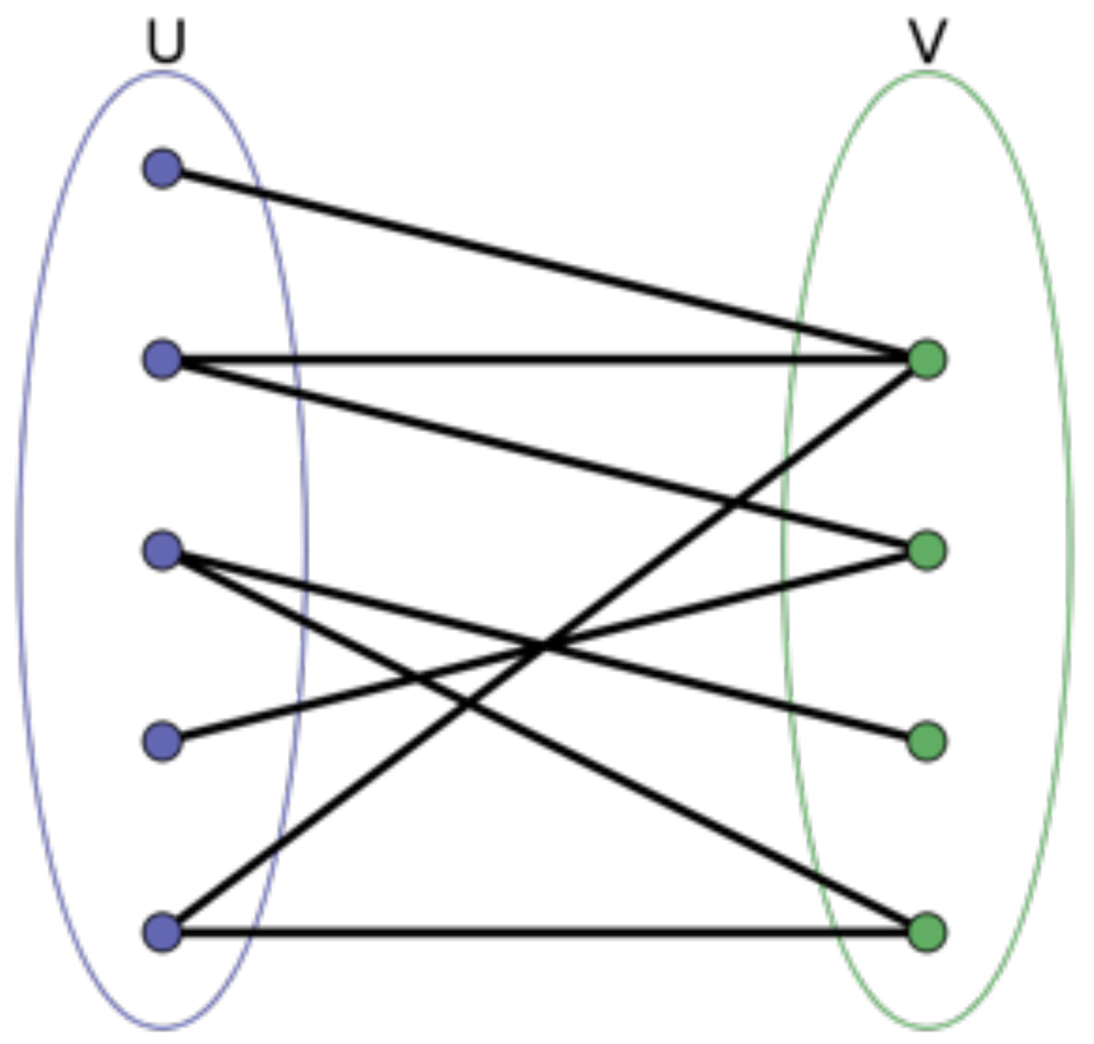
Mathworld

Network **zoology**

Bipartite graph

Two **types** of nodes

Links only between nodes of **different types**

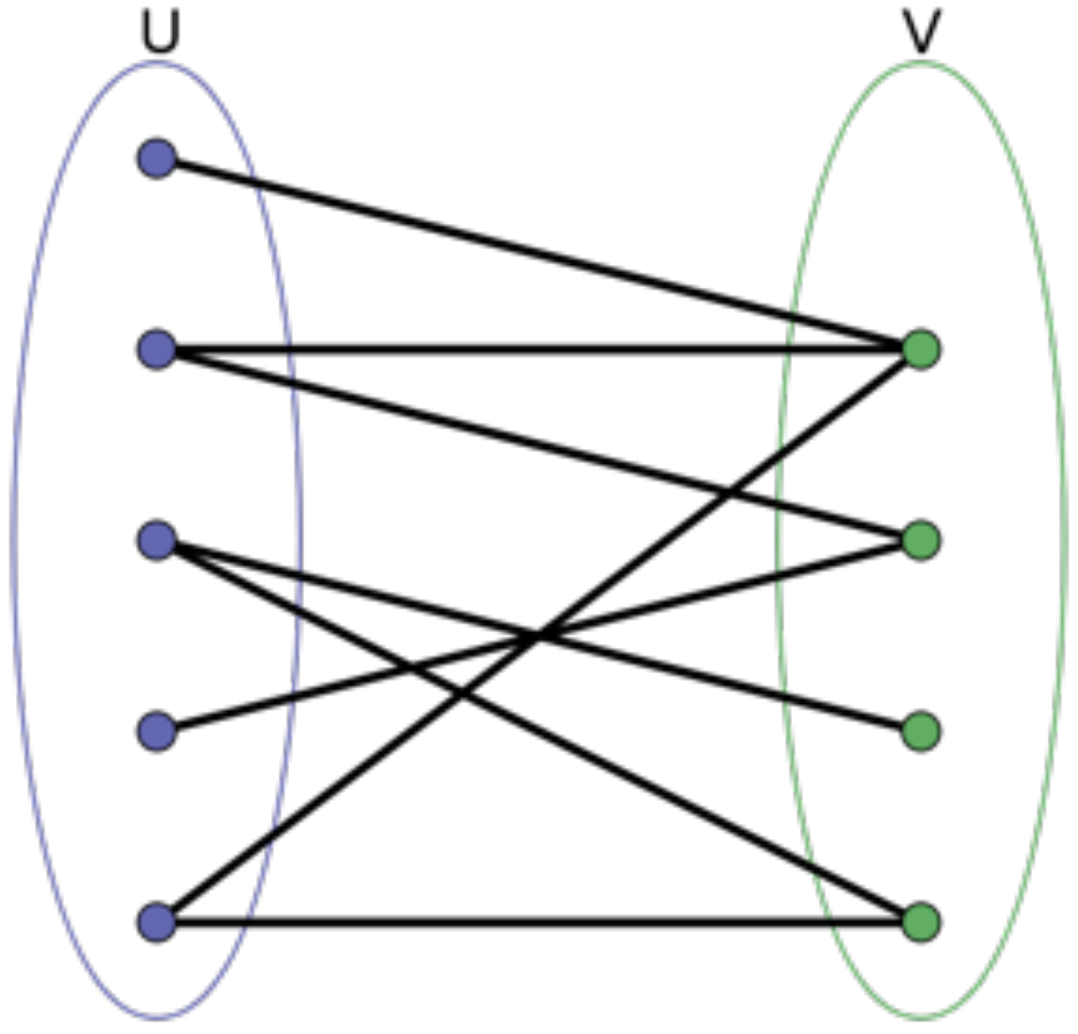


Network **zoology**

Bipartite graph

Two **types** of nodes

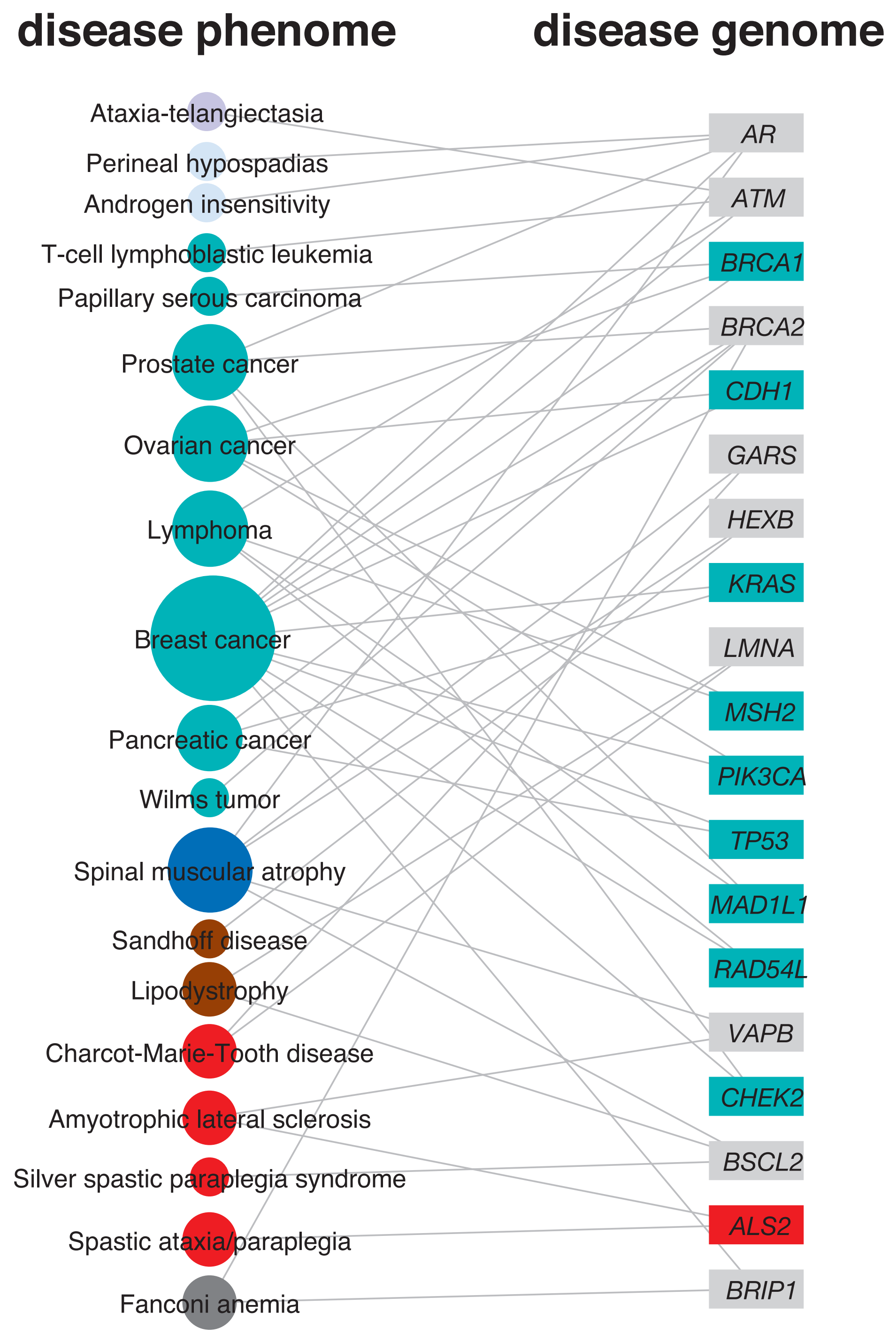
Links only between nodes of **different types**

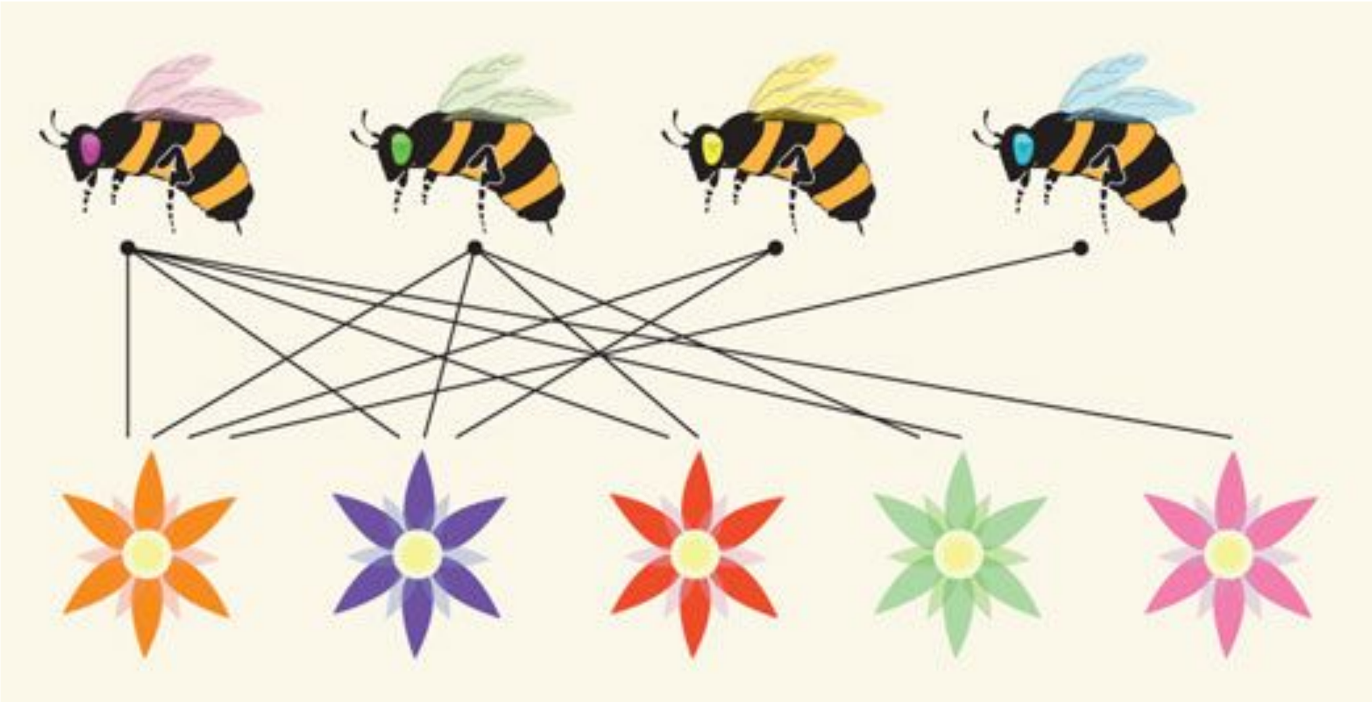


Movies ⟷ Actors

Enzymes ⟷ Metabolites

Scientists ⟷ Papers

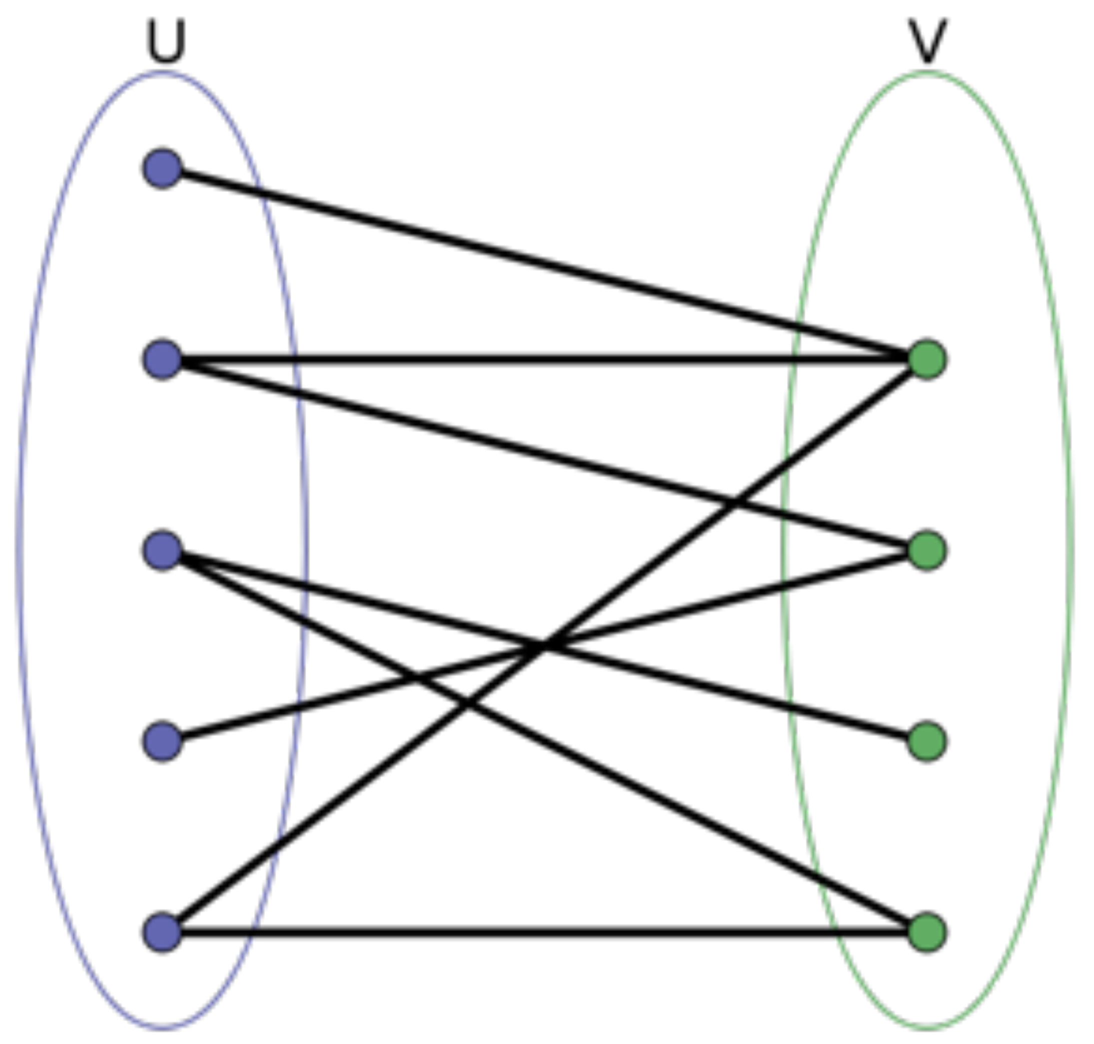




Network **zoology**

Bipartite **projection**

connect nodes in one group that have **common nodes** in the other group



Movies ⟷ Actors

"**Movies** that star the same actor(s)"

"**Actors** that appeared in the same movie(s)"

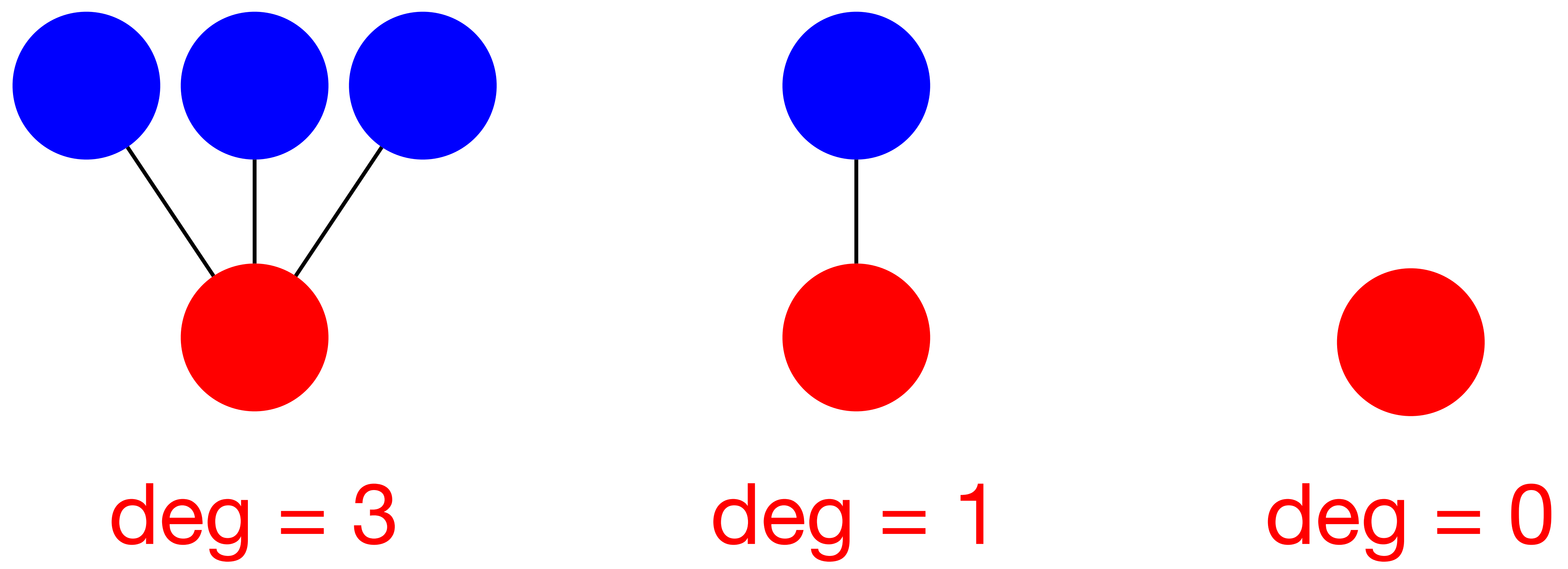
Network Quantifiers

(Advanced terminology)

Degree distribution

Degree — (perhaps) most fundamental property of a node

Number of neighbors connected to a node

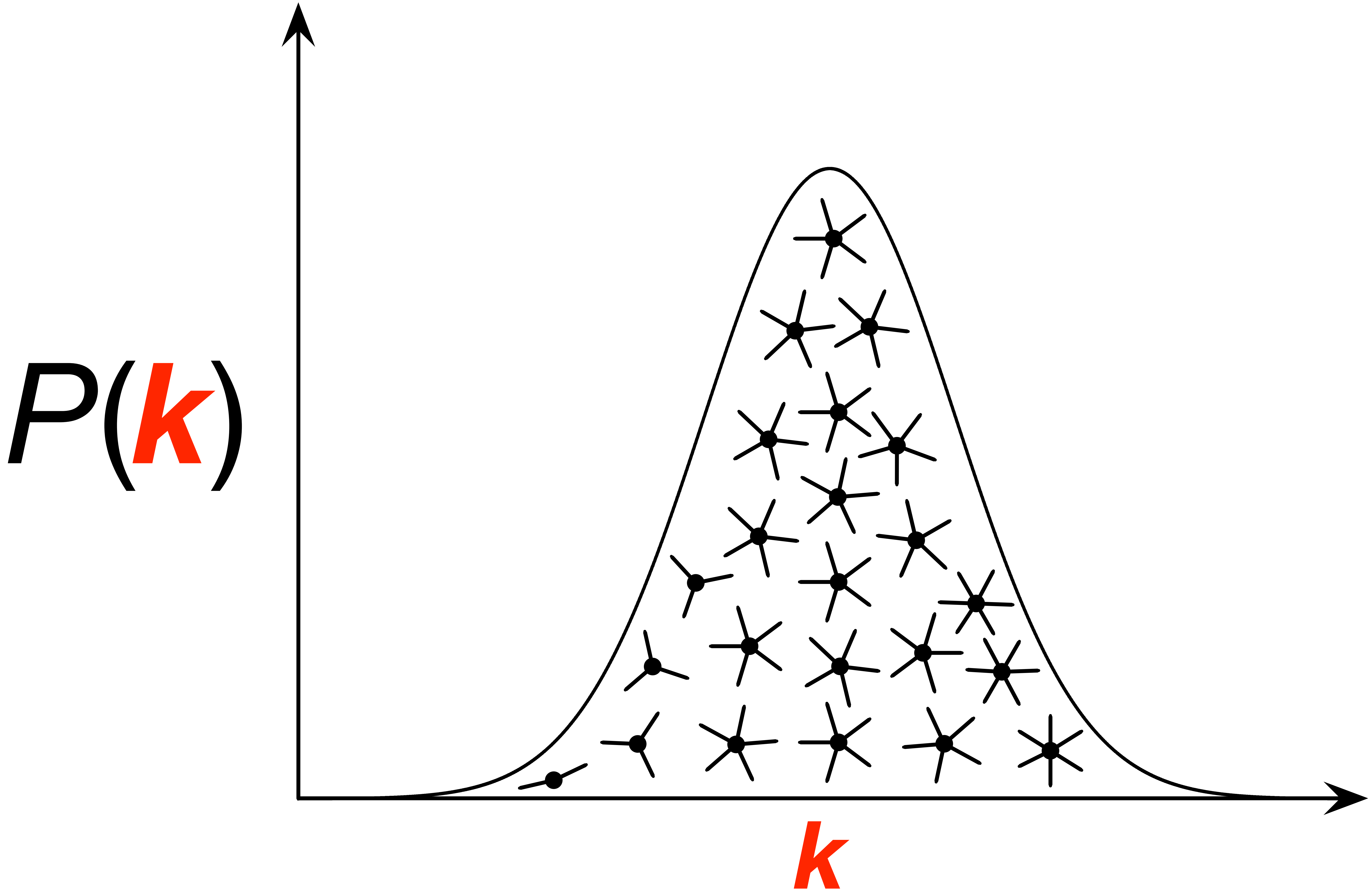


Degree **distribution**

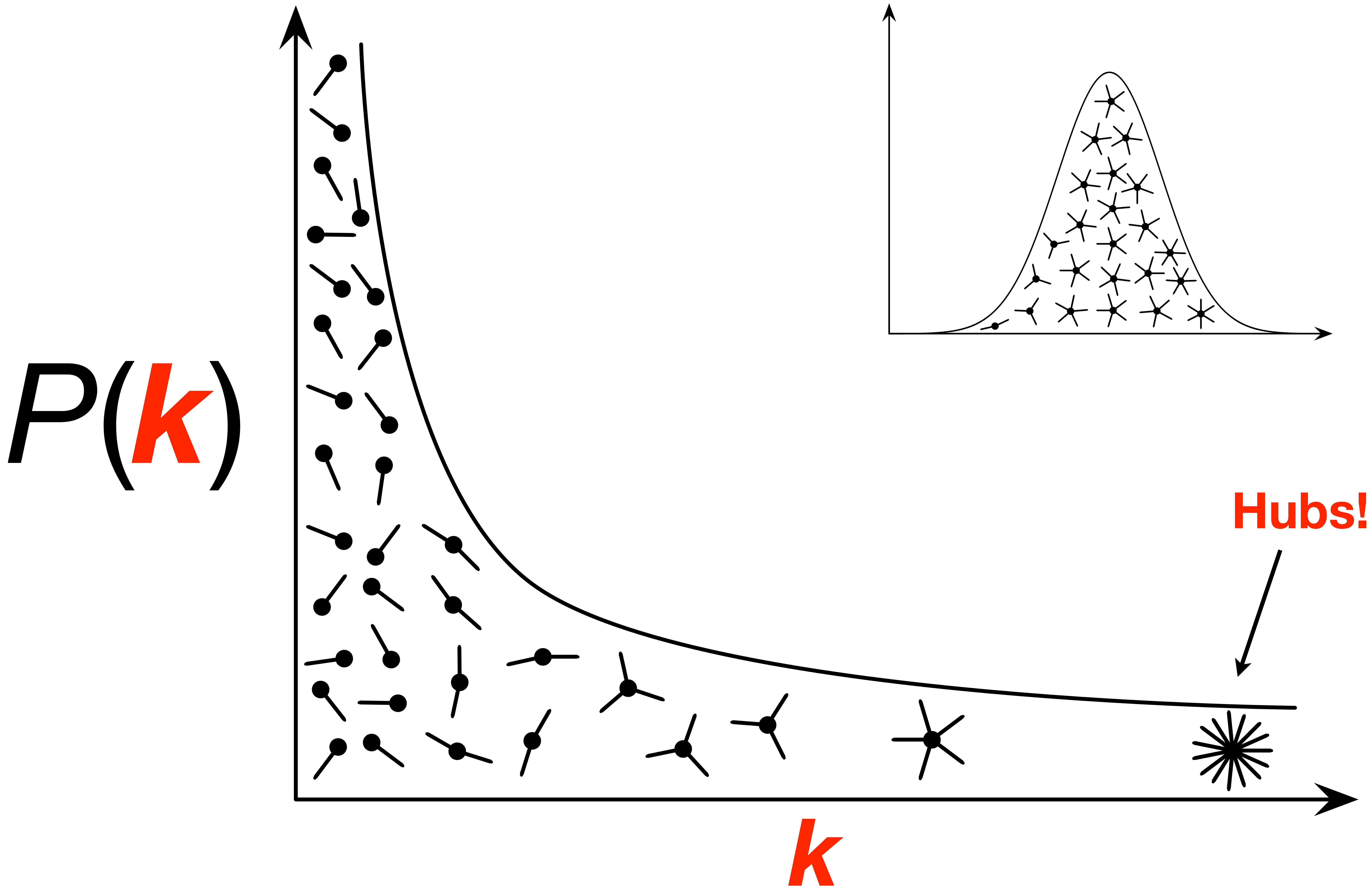
$P(k)$

Probability that a random node has degree k

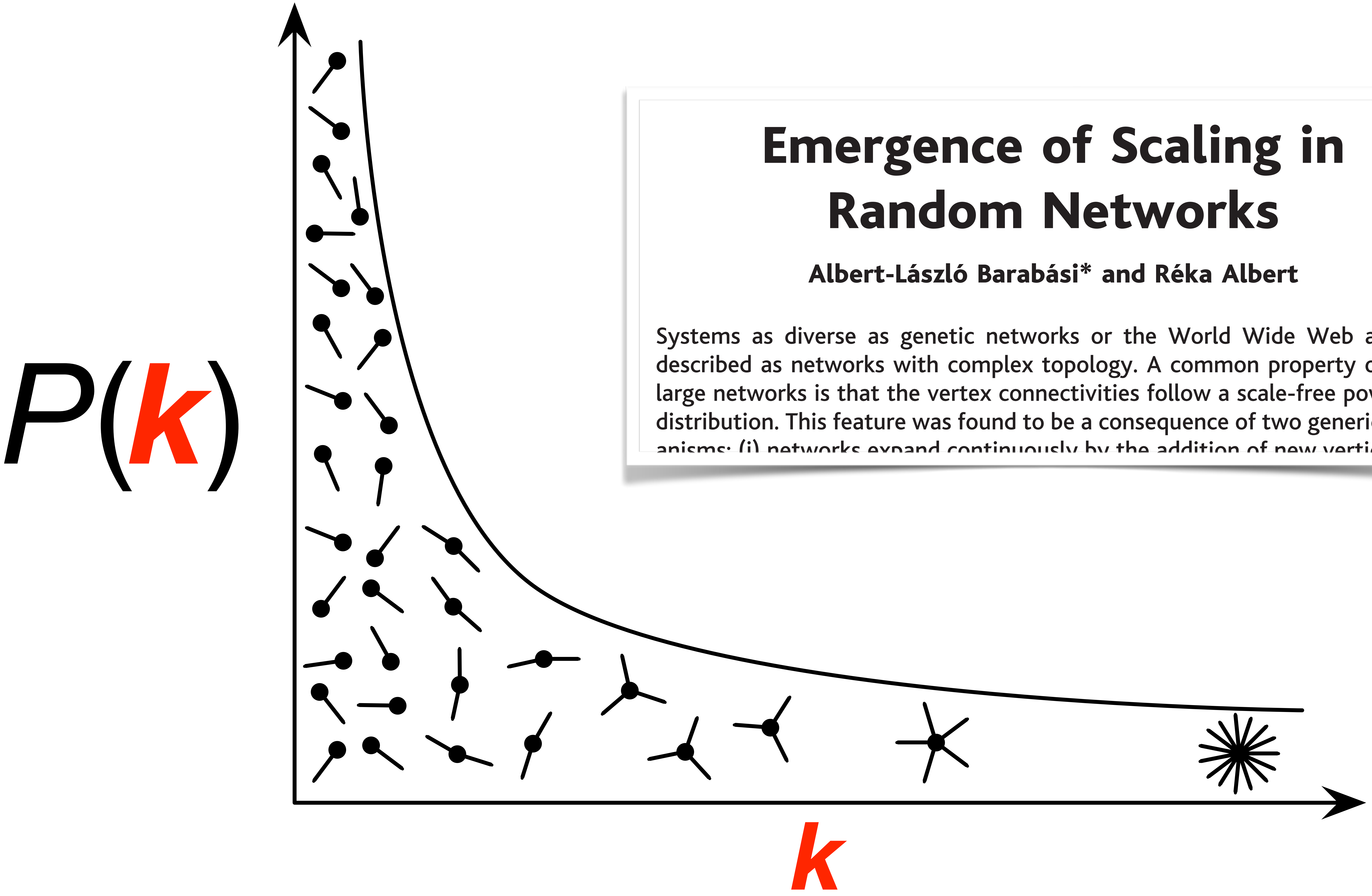
Degree distribution



scale-free network



scale-free network

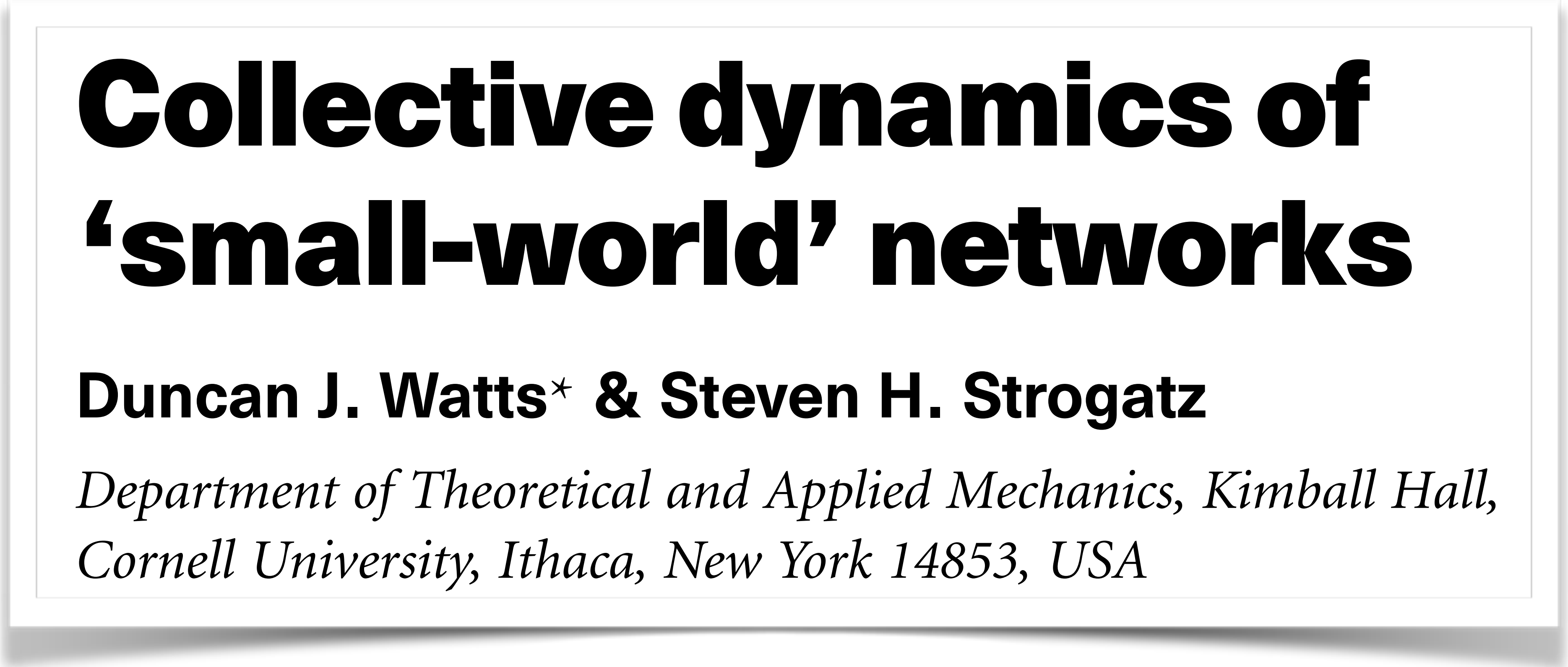


Emergence of Scaling in Random Networks

Albert-László Barabási* and Réka Albert

Systems as diverse as genetic networks or the World Wide Web are best described as networks with complex topology. A common property of many large networks is that the vertex connectivities follow a scale-free power-law distribution. This feature was found to be a consequence of two generic mechanisms: (i) networks expand continuously by the addition of new vertices, and

Clustering coefficient



Collective dynamics of ‘small-world’ networks

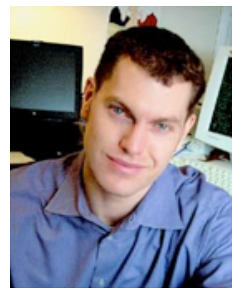
Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Feature of network neighborhoods

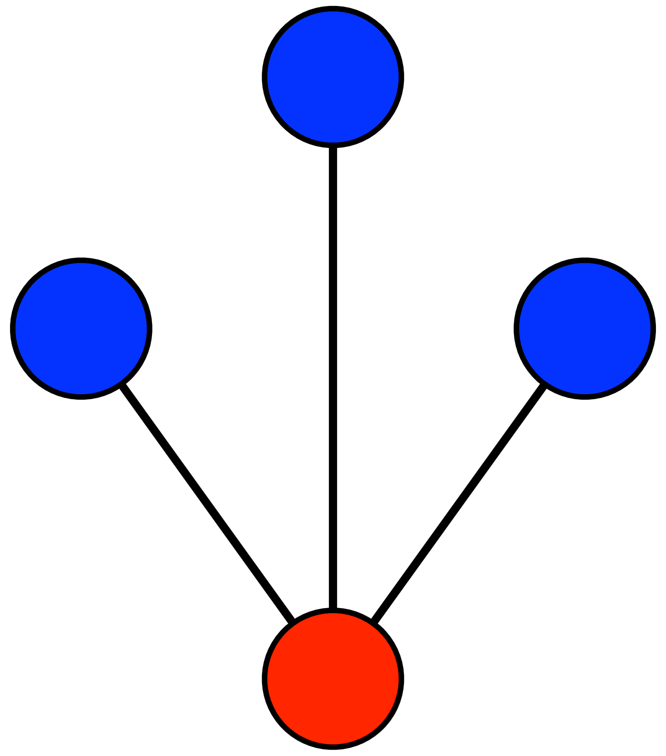
Nature, 1998

Clustering coefficient

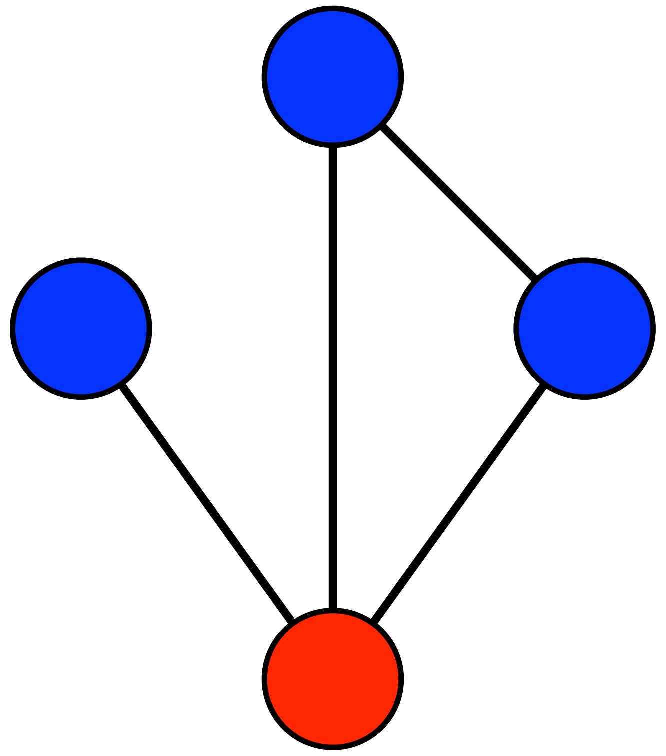
 

How many **triangles** are in the neighborhood?

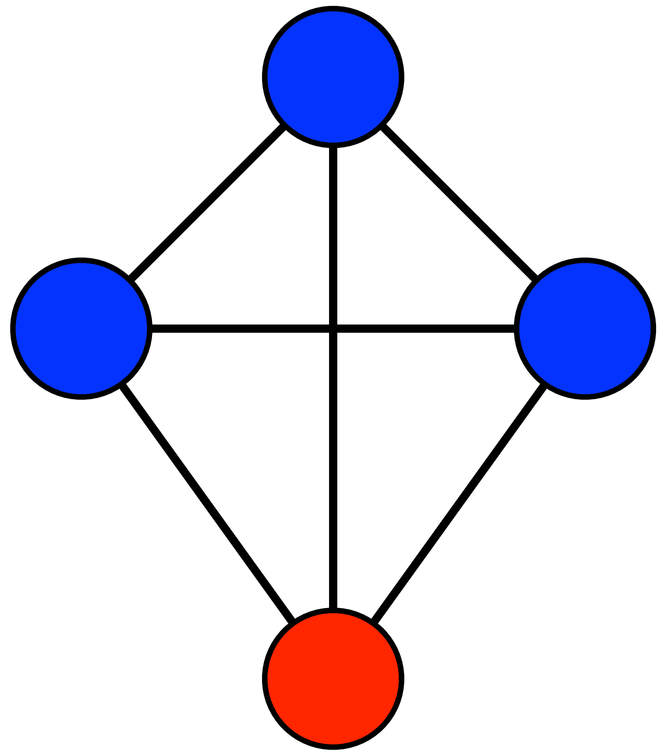
How many triangles **are possible**?



$C_i = 0$



$C_i = 1/3$



$C_i = 1$

Real networks ➟ **more triangles** than expected!

Nature, 1998

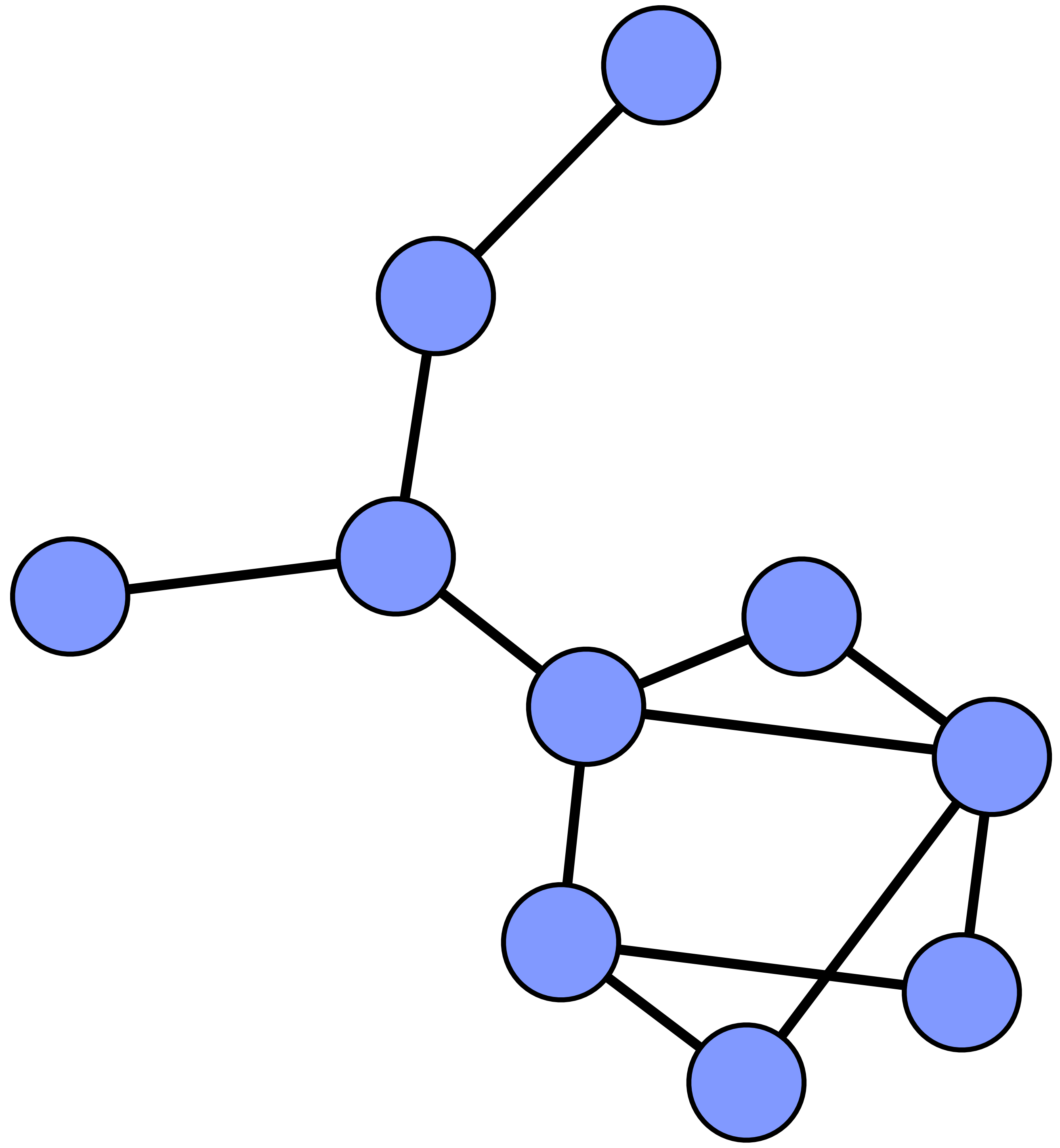
Distances and Networks

Networks aren’t though of as existing in **ordinary space**

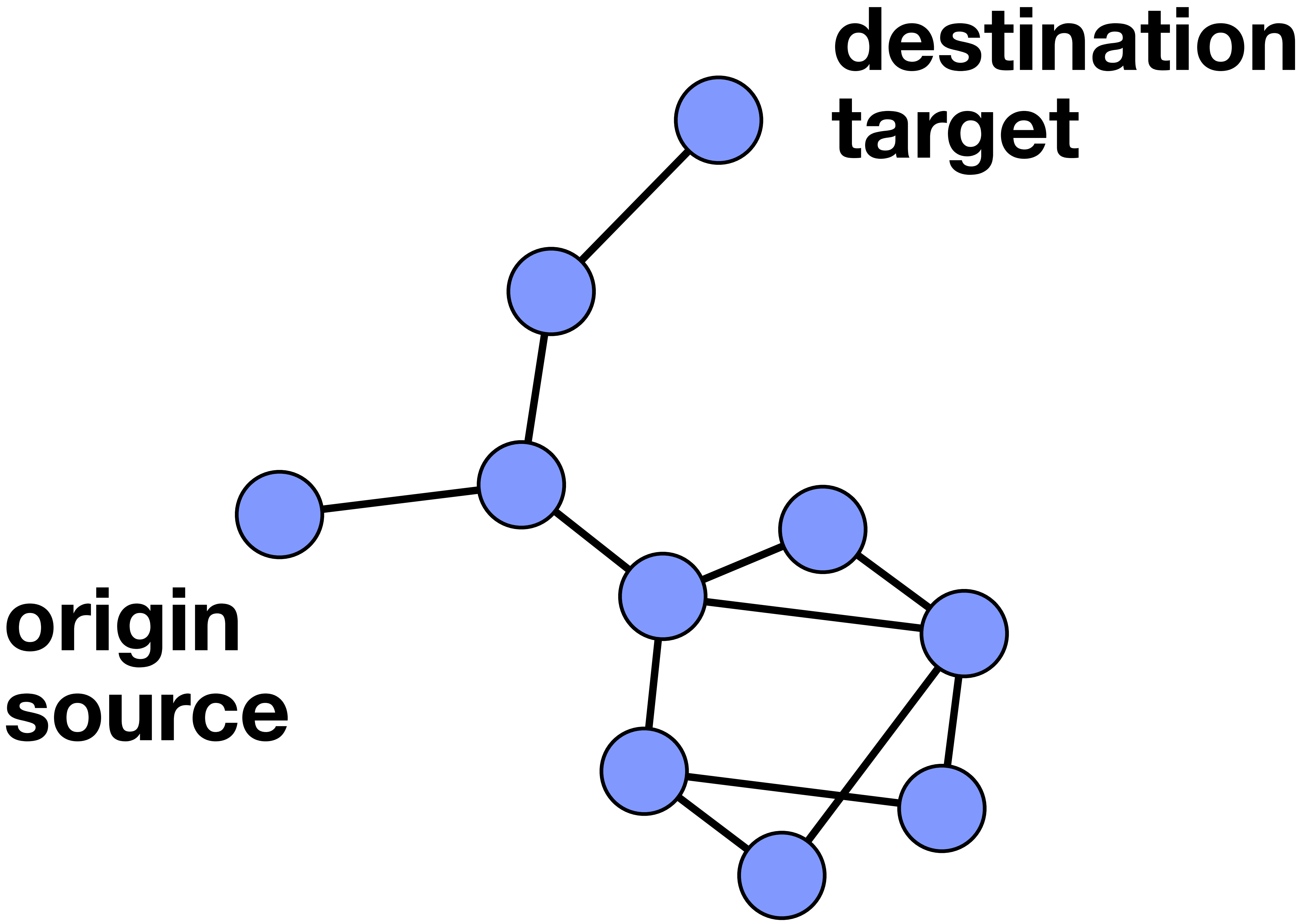
Space lets us tell how far apart things are

How to measure **distance** in a network?

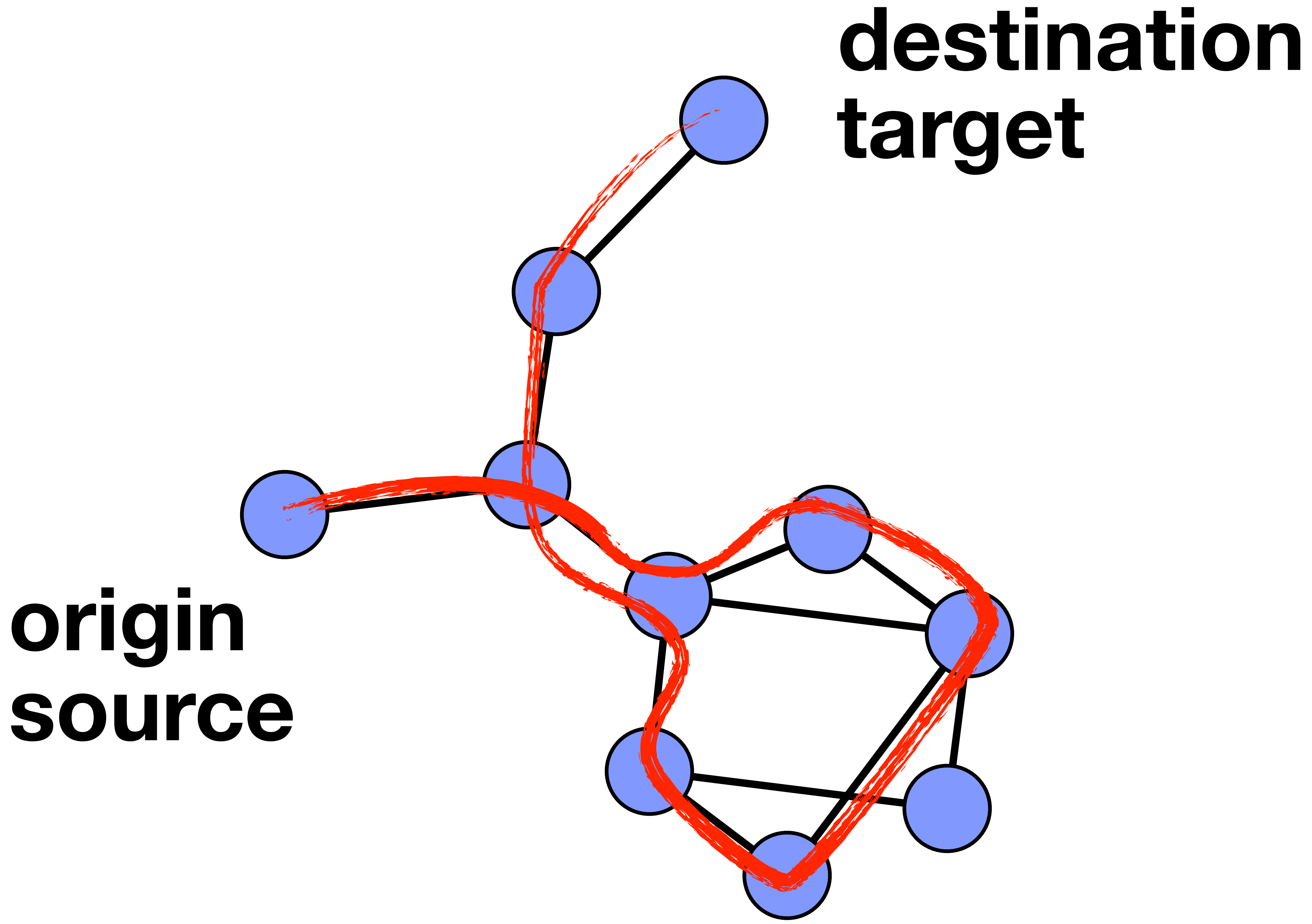
Paths



Paths

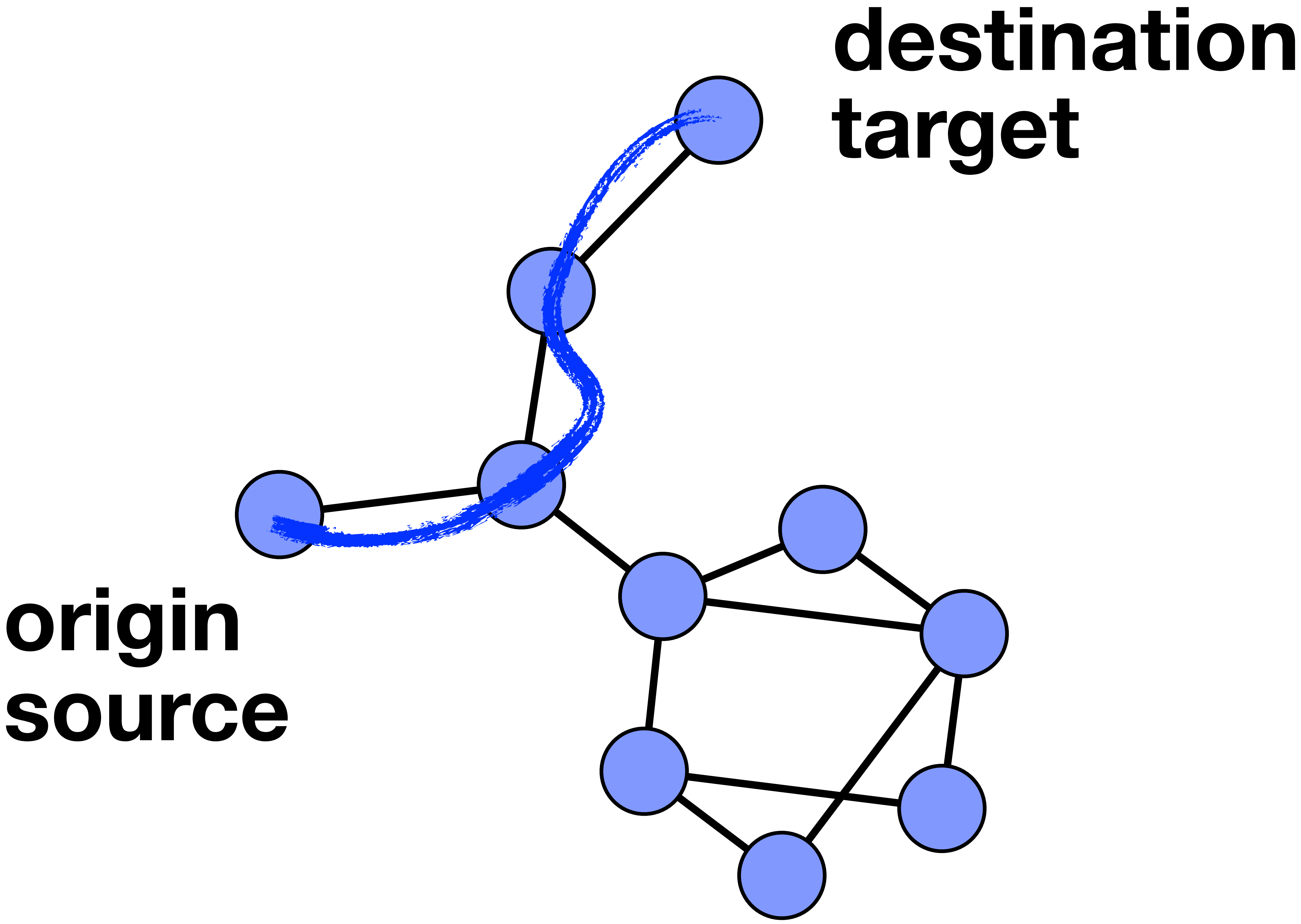


Paths



many paths exist

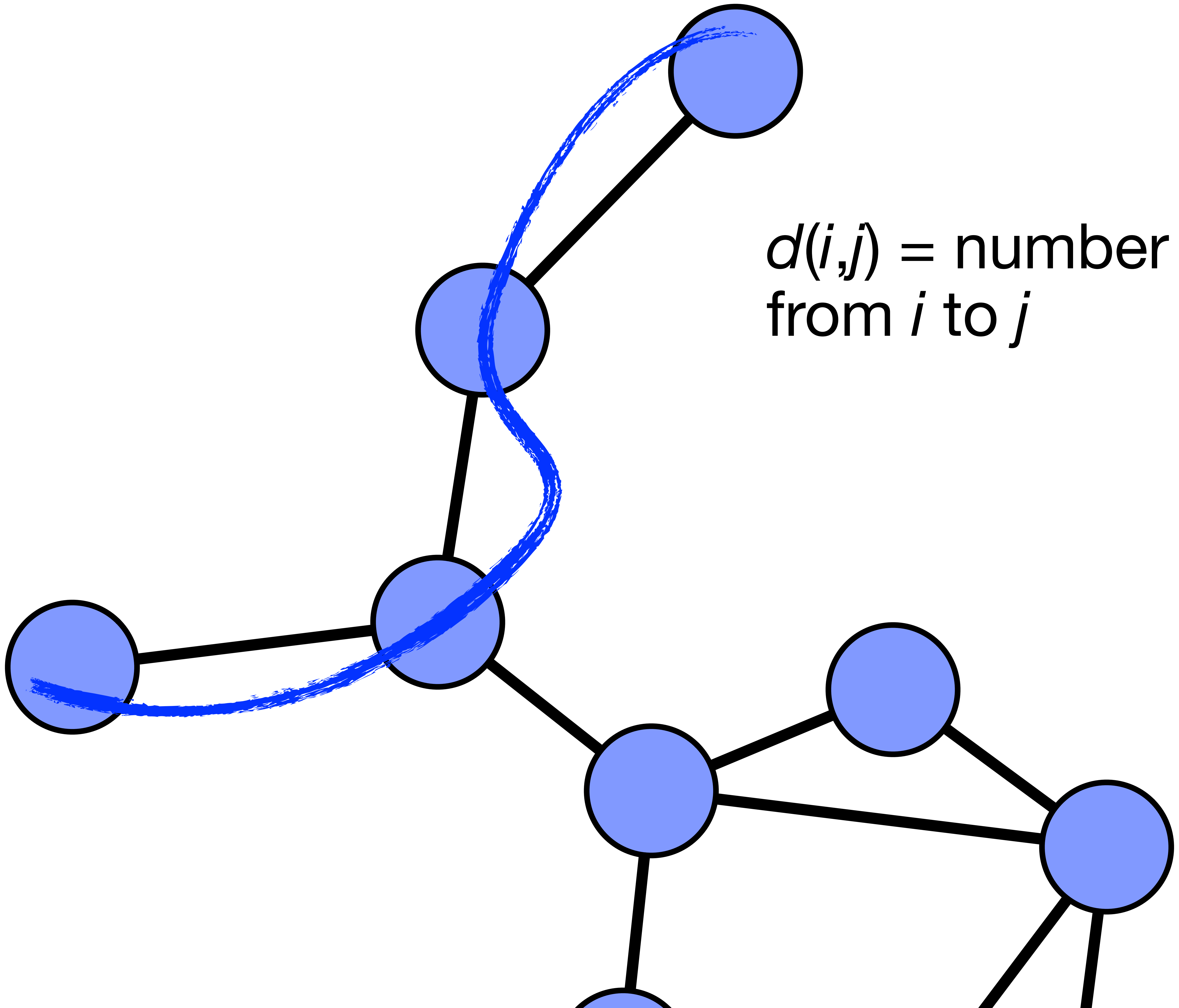
Paths



many paths exist

we want the
shortest path

Path Length



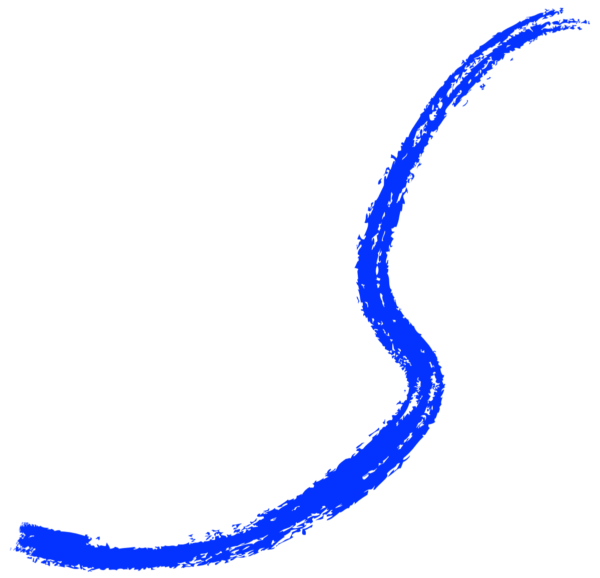
$d(i,j)$ = number of **hops** to get from i to j

$$d = 3$$

Set of **all** paths

Compute shortest path from a node to **every other** node

Set of **all** paths



Compute shortest path from a node to **every other** node

Eccentricity of a node

Longest shortest path starting from that node

Set of **all** paths

Compute shortest path from a node to **every other** node

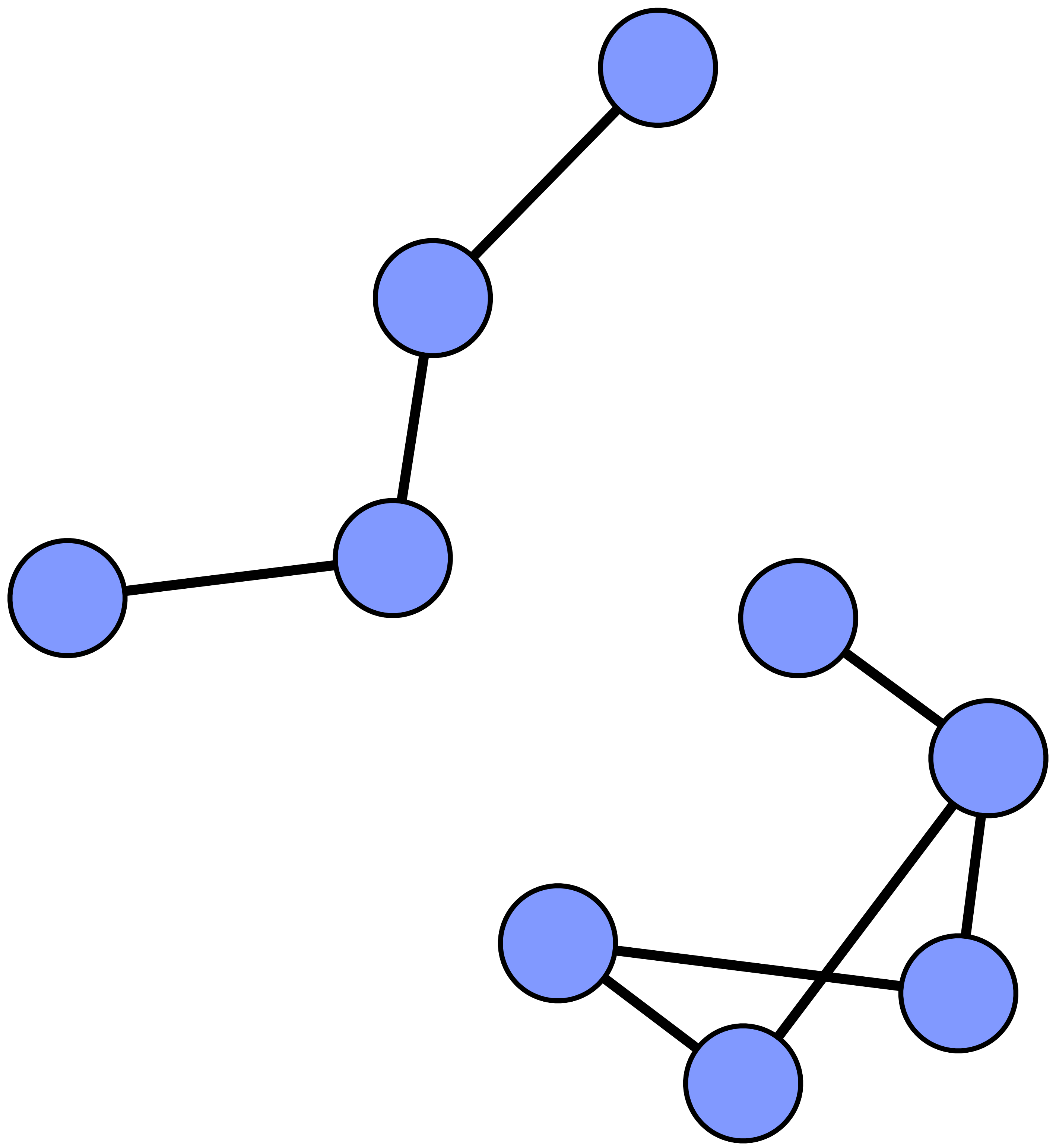
Eccentricity of a node

Longest shortest path starting from that node

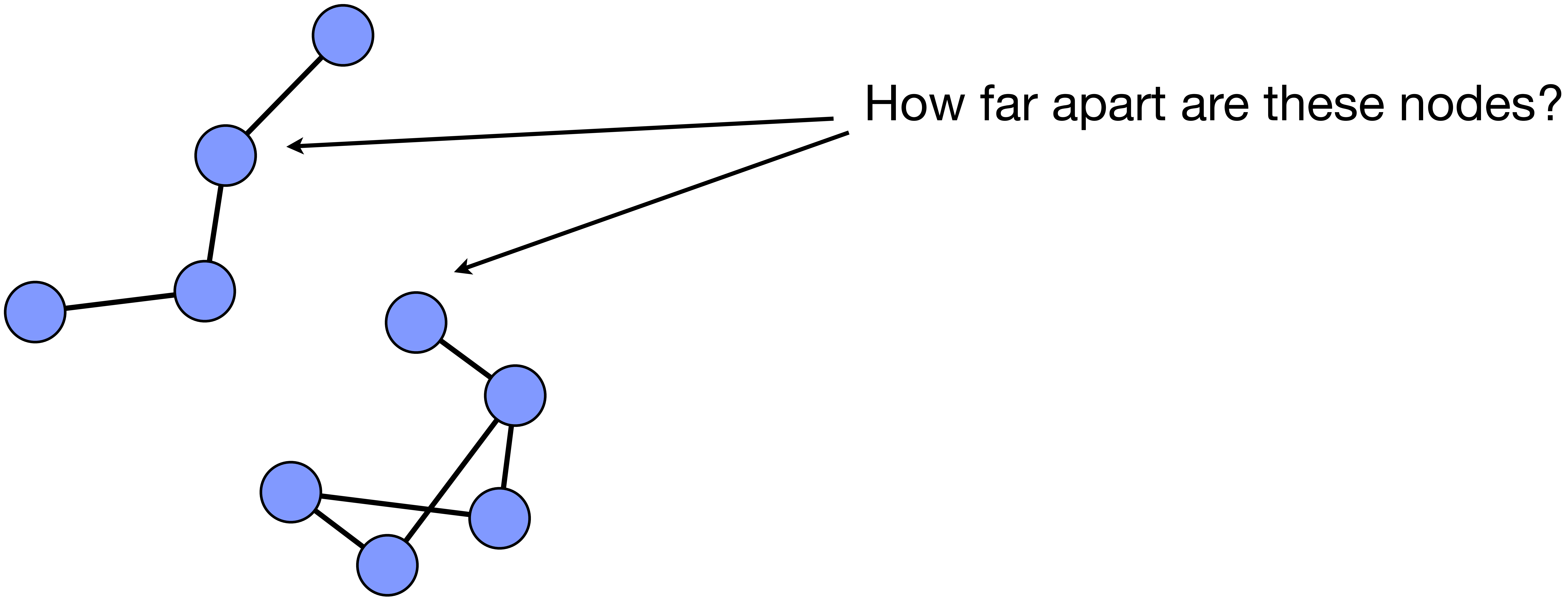
Diameter of a network

Longest of **all** shortest paths

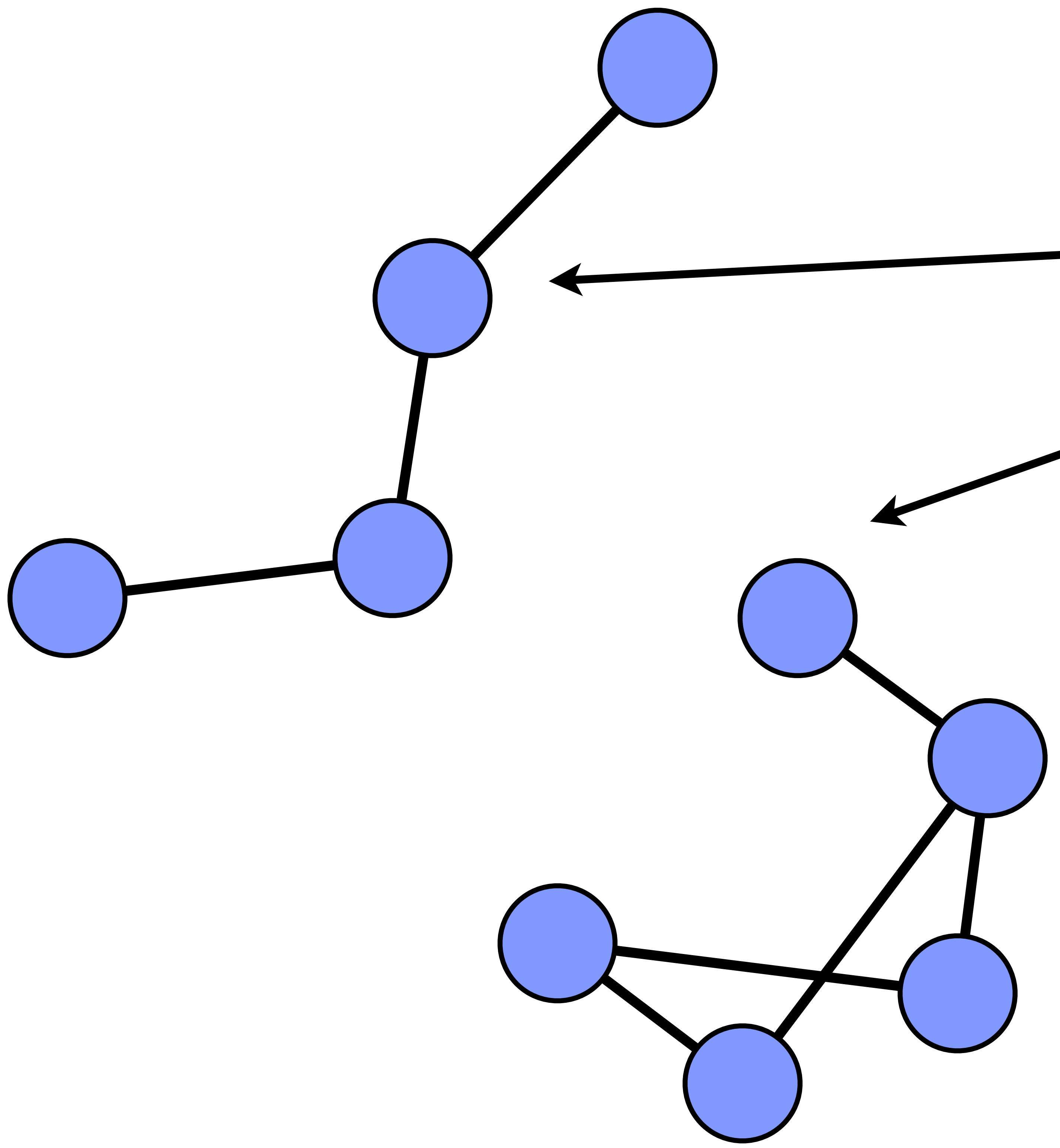
How big can distances be?



How big can distances be?



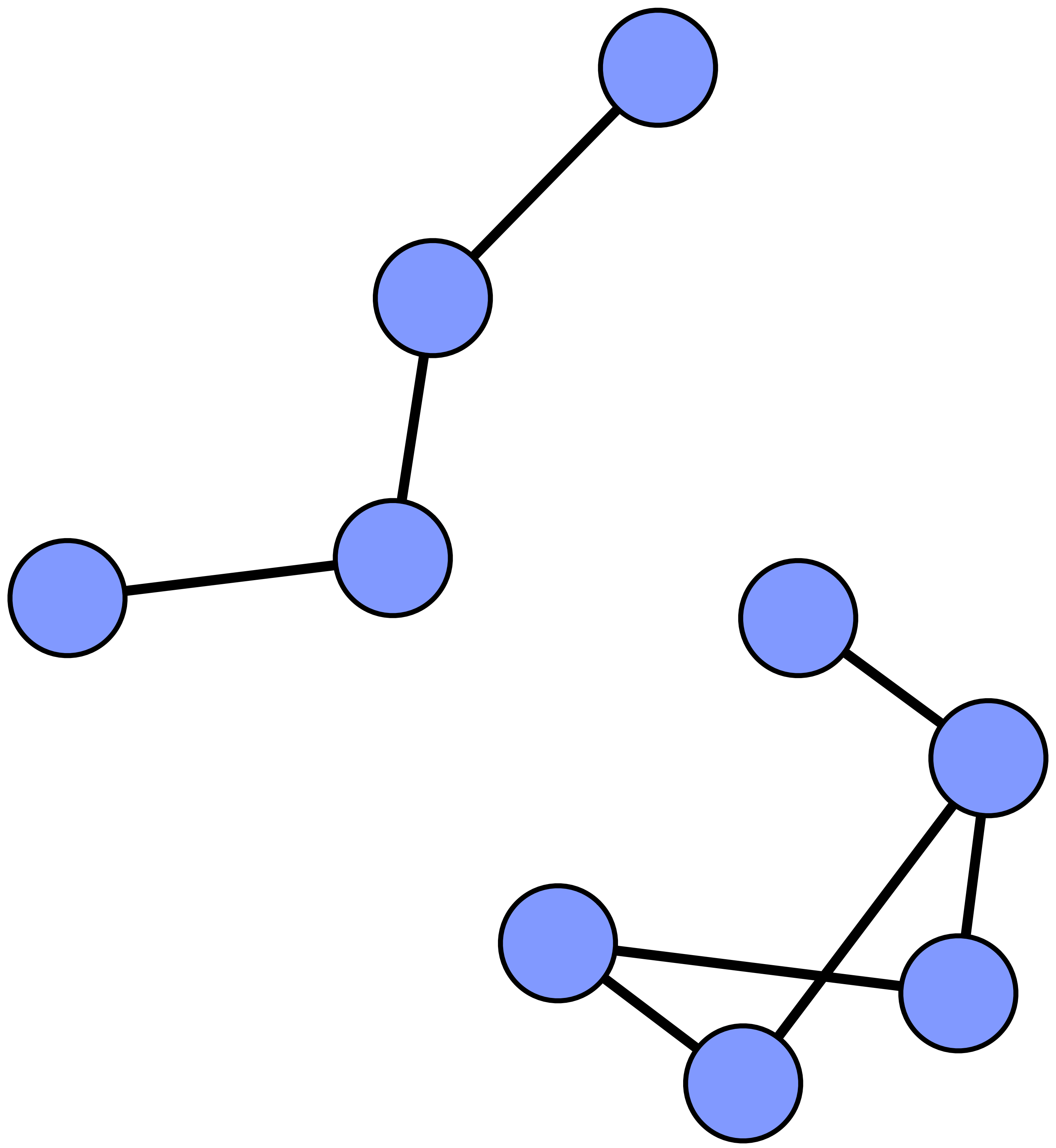
How big can distances be?



How far apart are these nodes?

Answer: $d = \infty$

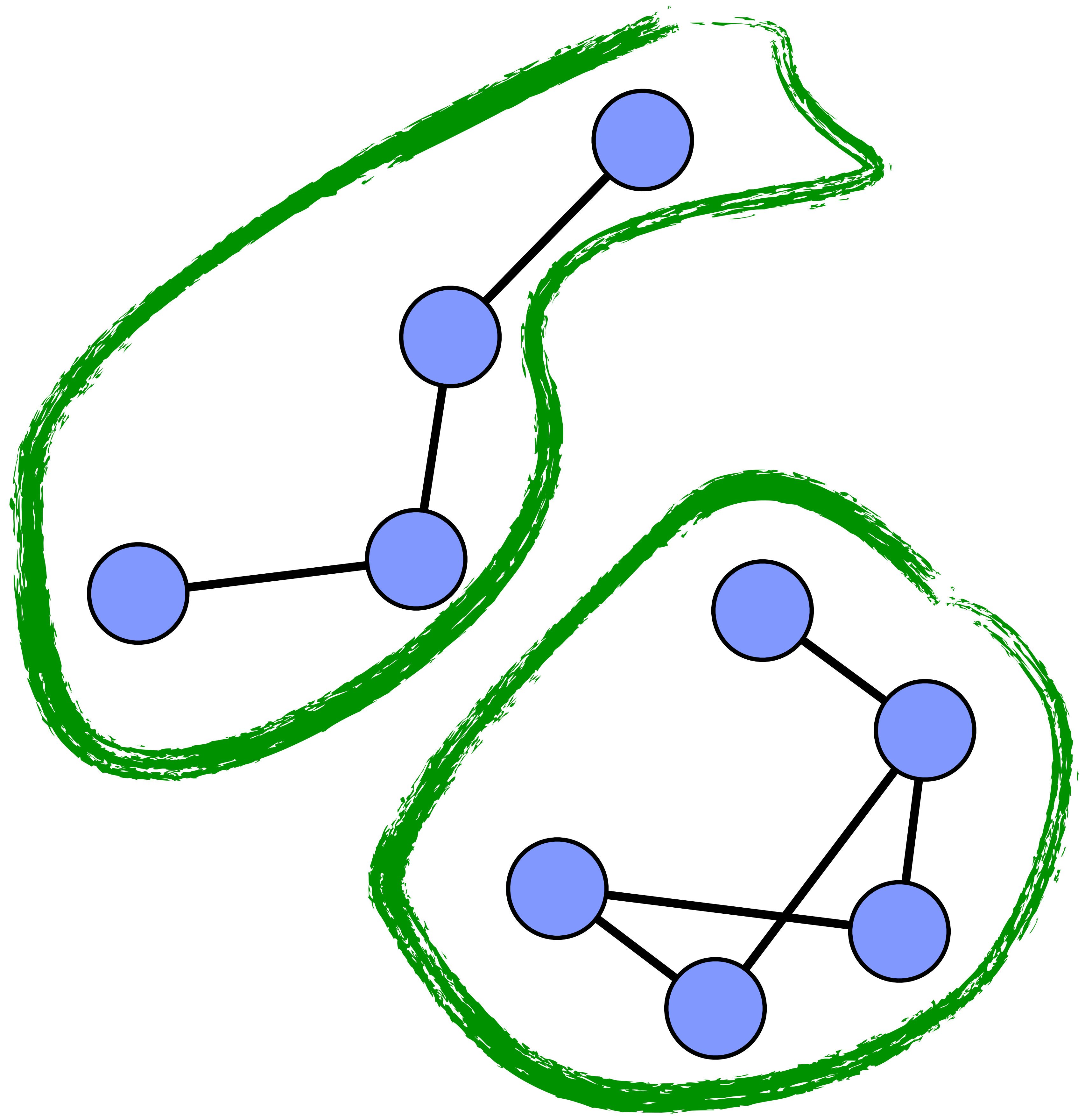
How big can distances be?



Networks can be **disconnected**

or **disjoint**

How big can distances be?



Networks can be **disconnected**

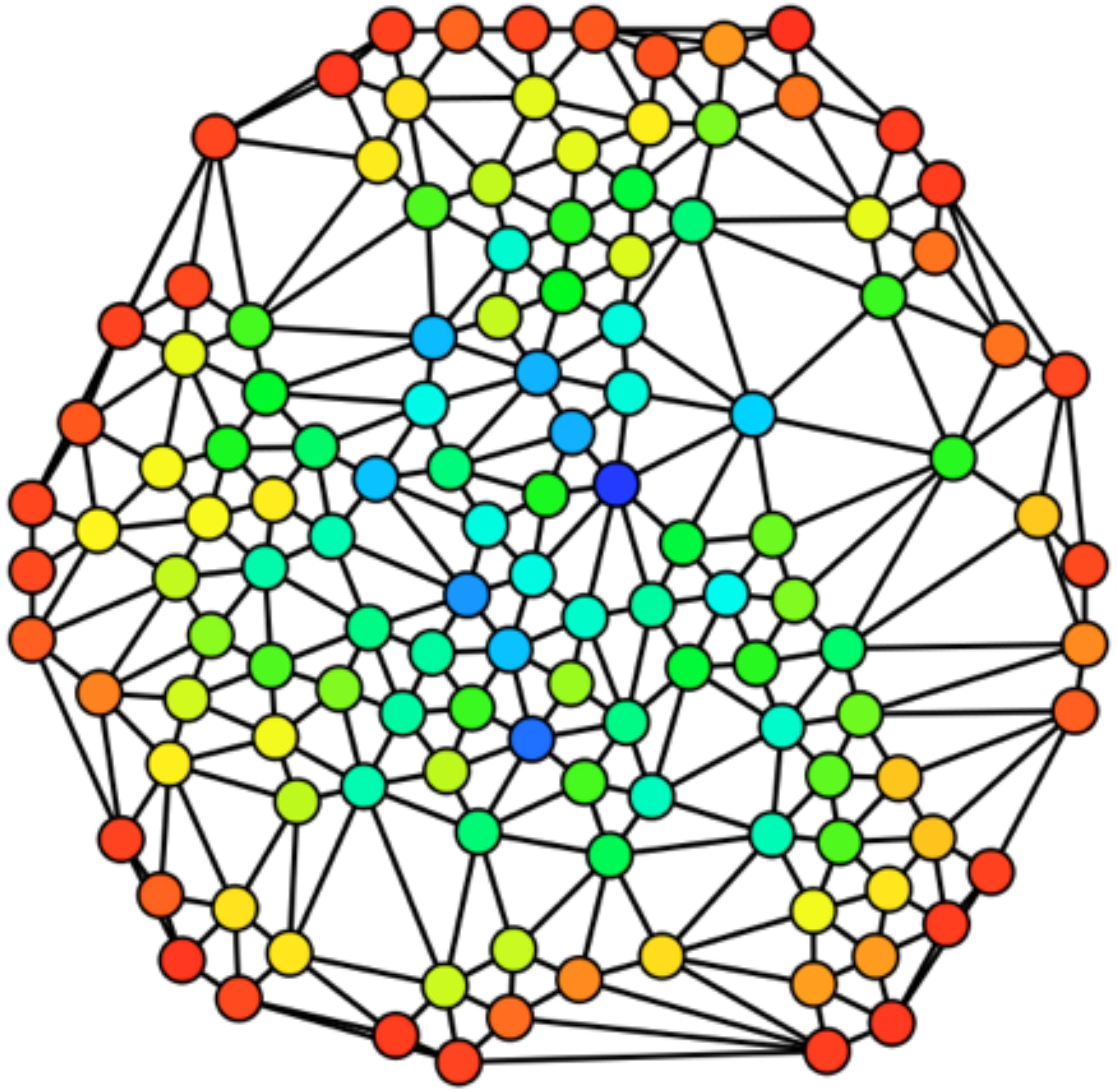
or **disjoint**

Components

(Connected components)

Centrality

What nodes are "**important**"?



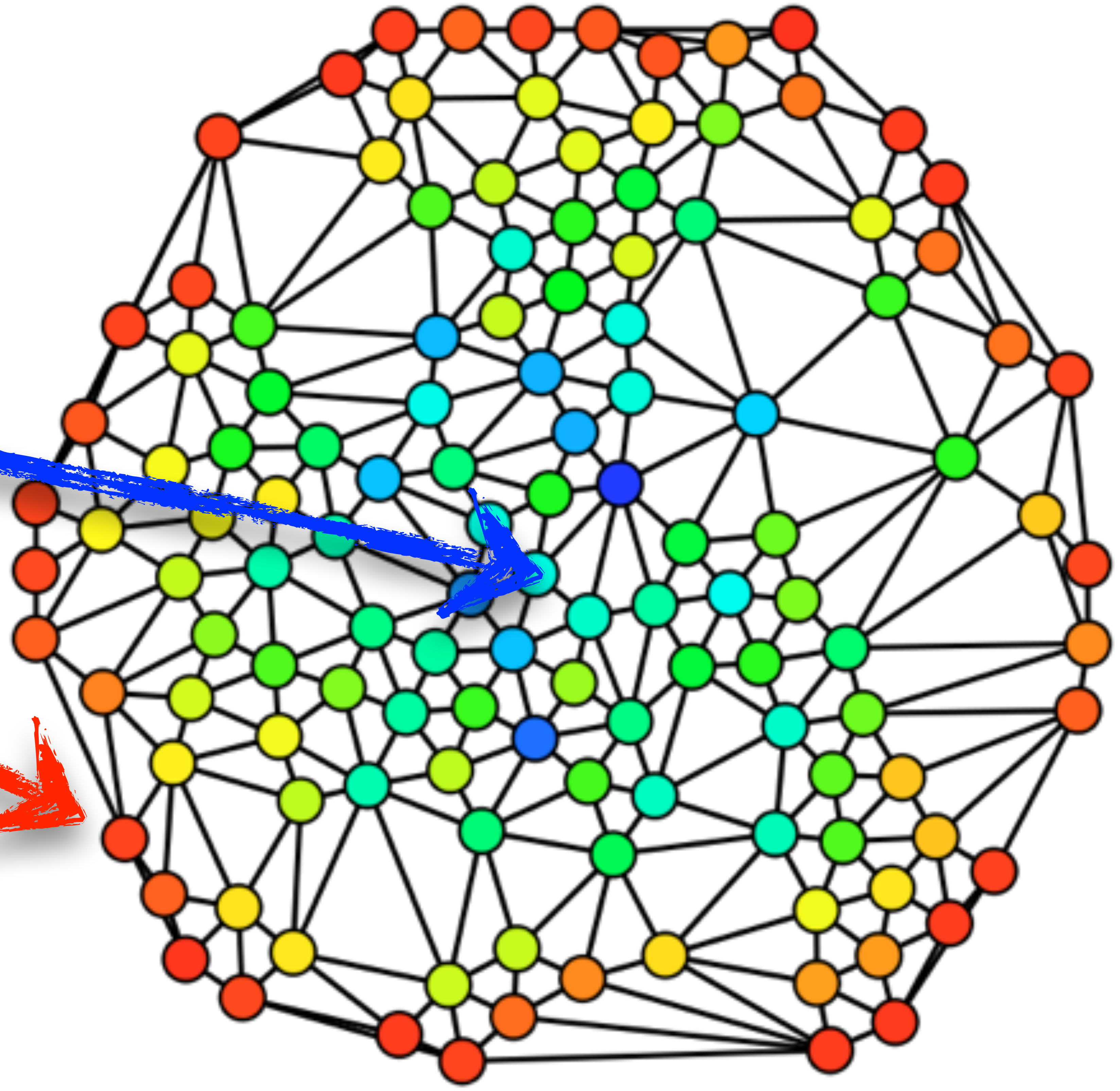
wikipedia.org

Centrality

What nodes are "**important**"?

Central

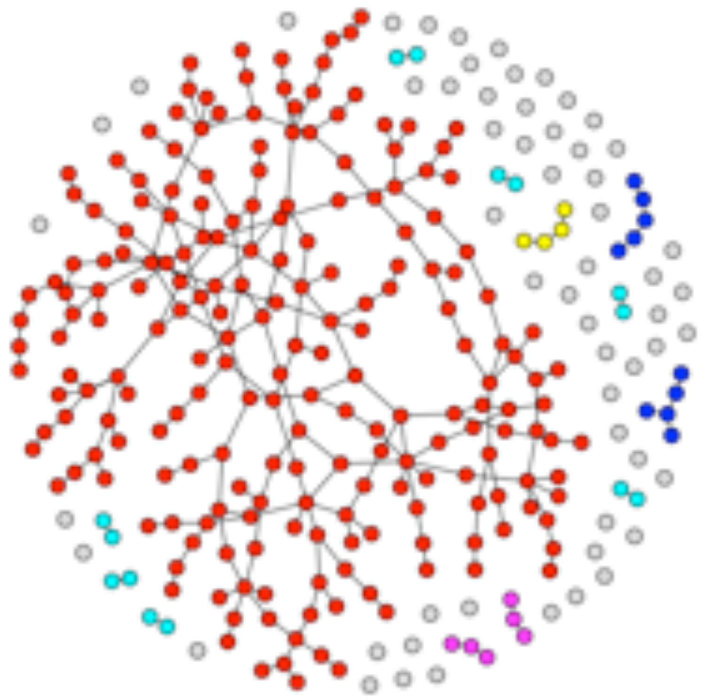
Peripheral



Rank nodes

wikipedia.org

Random network models



Random graphs

1736 Graph theory

Euler

Random graphs

1736 Graph theory

Euler

1959 Random graph theory

Erdős

Rényi

Gilbert

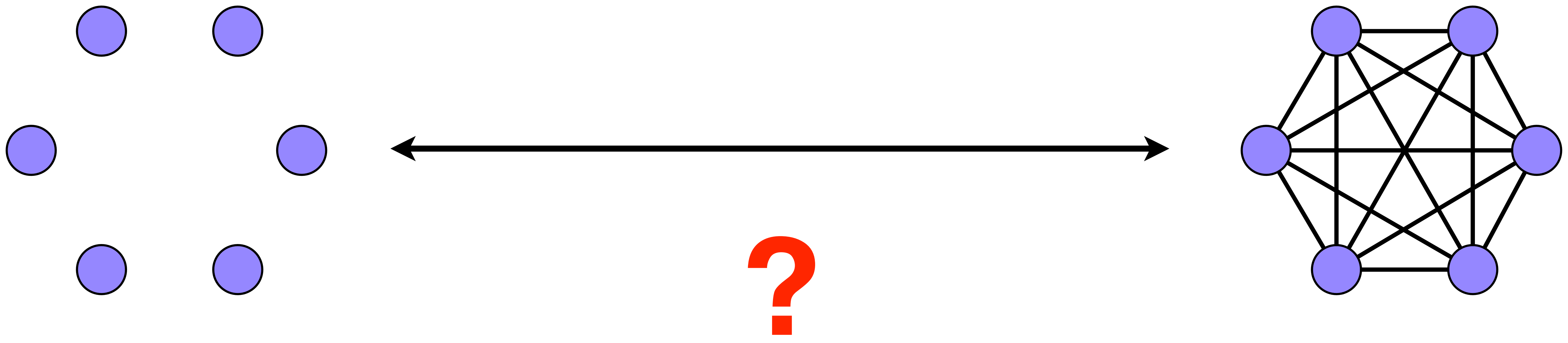
Erdős-Rényi Graph

1. Start with an **empty graph** of N nodes

Erdős-Rényi Graph

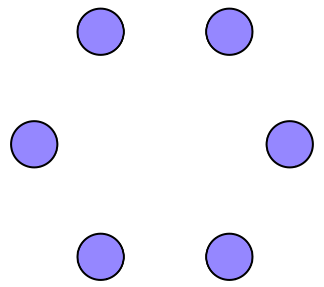
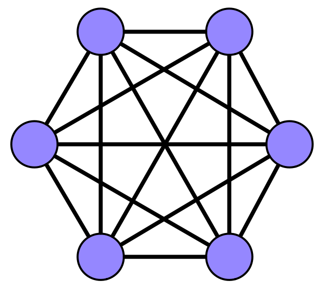
1. Start with an **empty graph** of *N* nodes

2. Look at **every pair** of nodes:
   With probability ***p*** connect
   that pair with a link

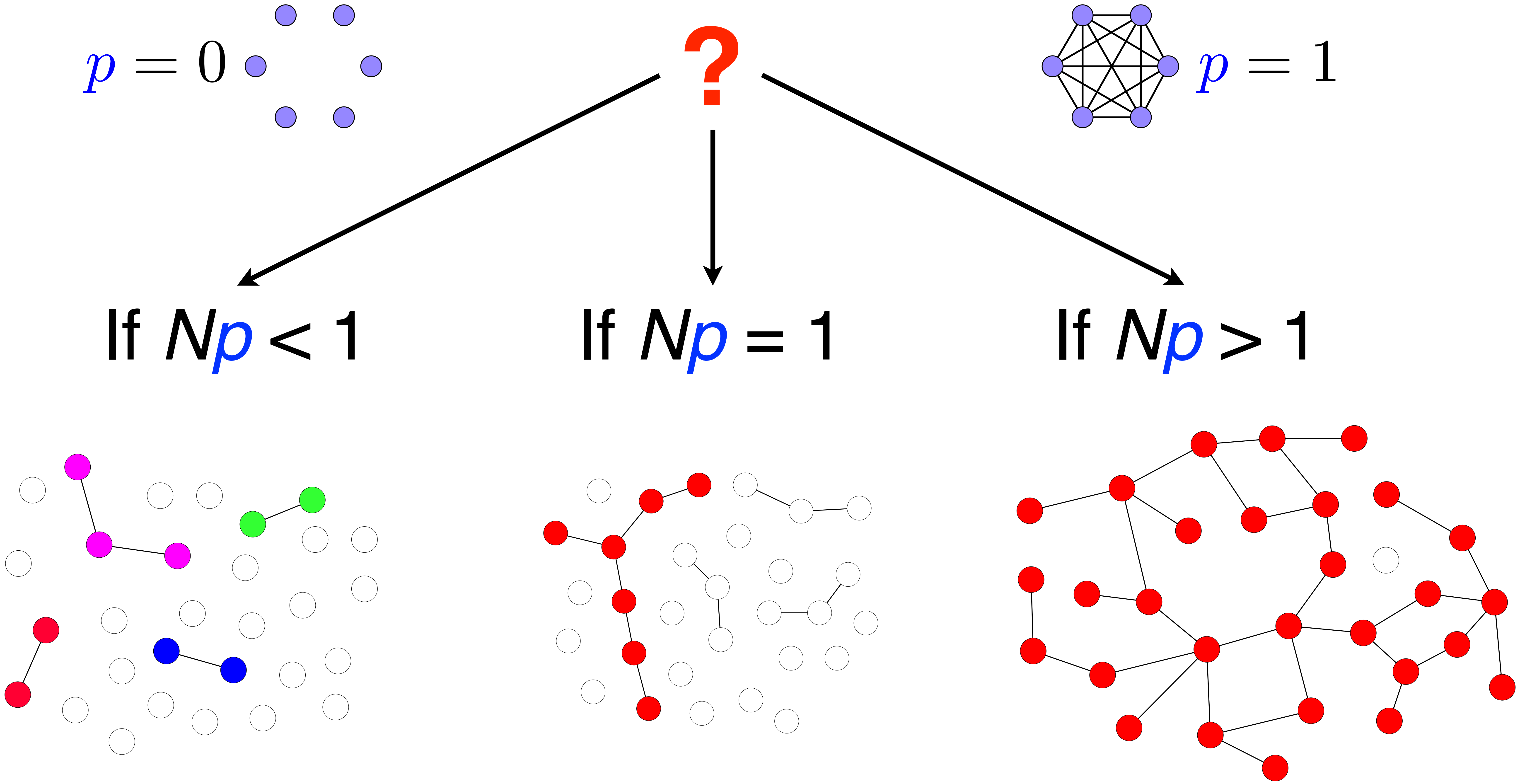


$p = 0$ $p = 1$

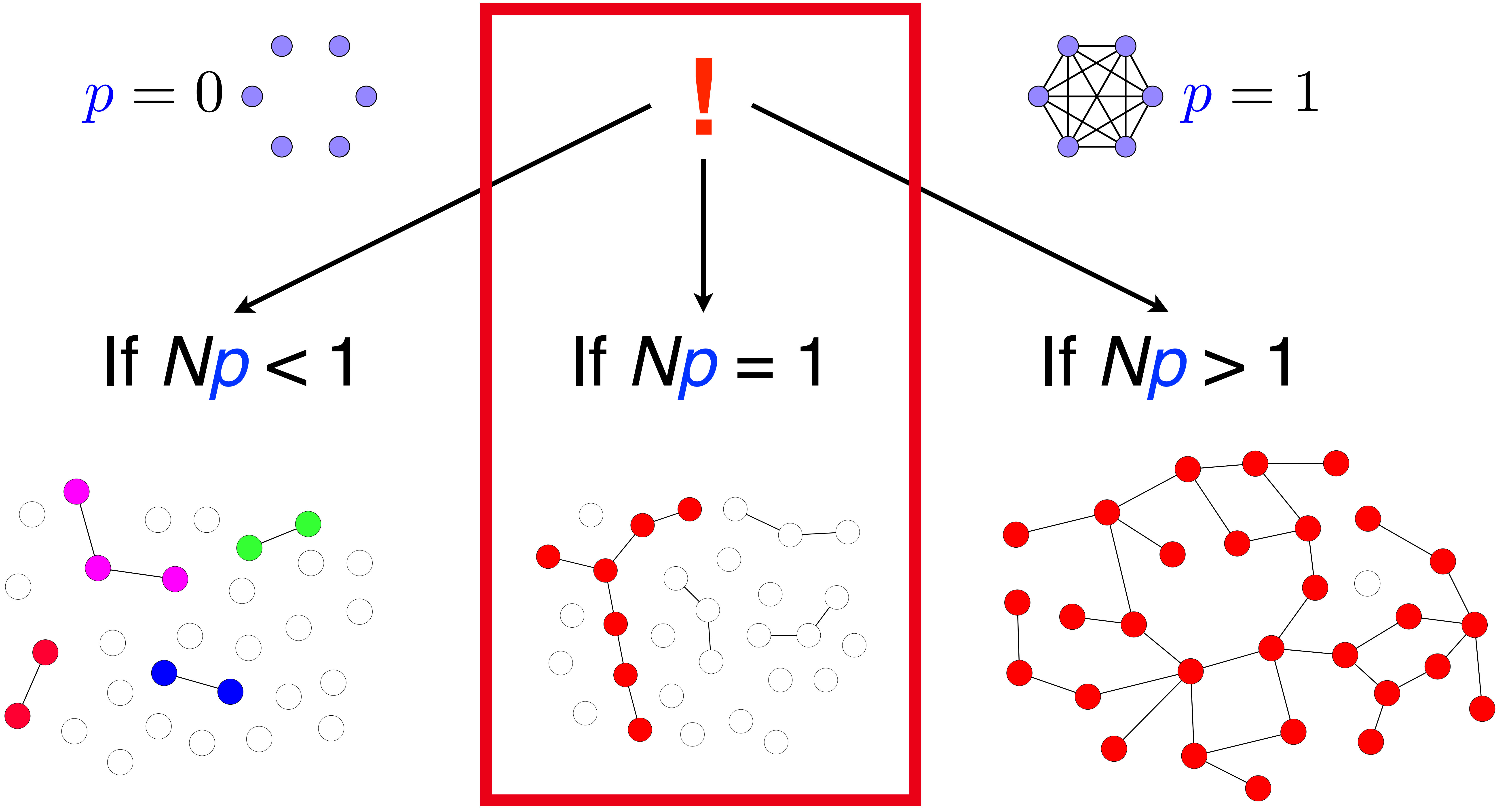
Erdős-Rényi Graph

$p = 0$  **?**  $p = 1$

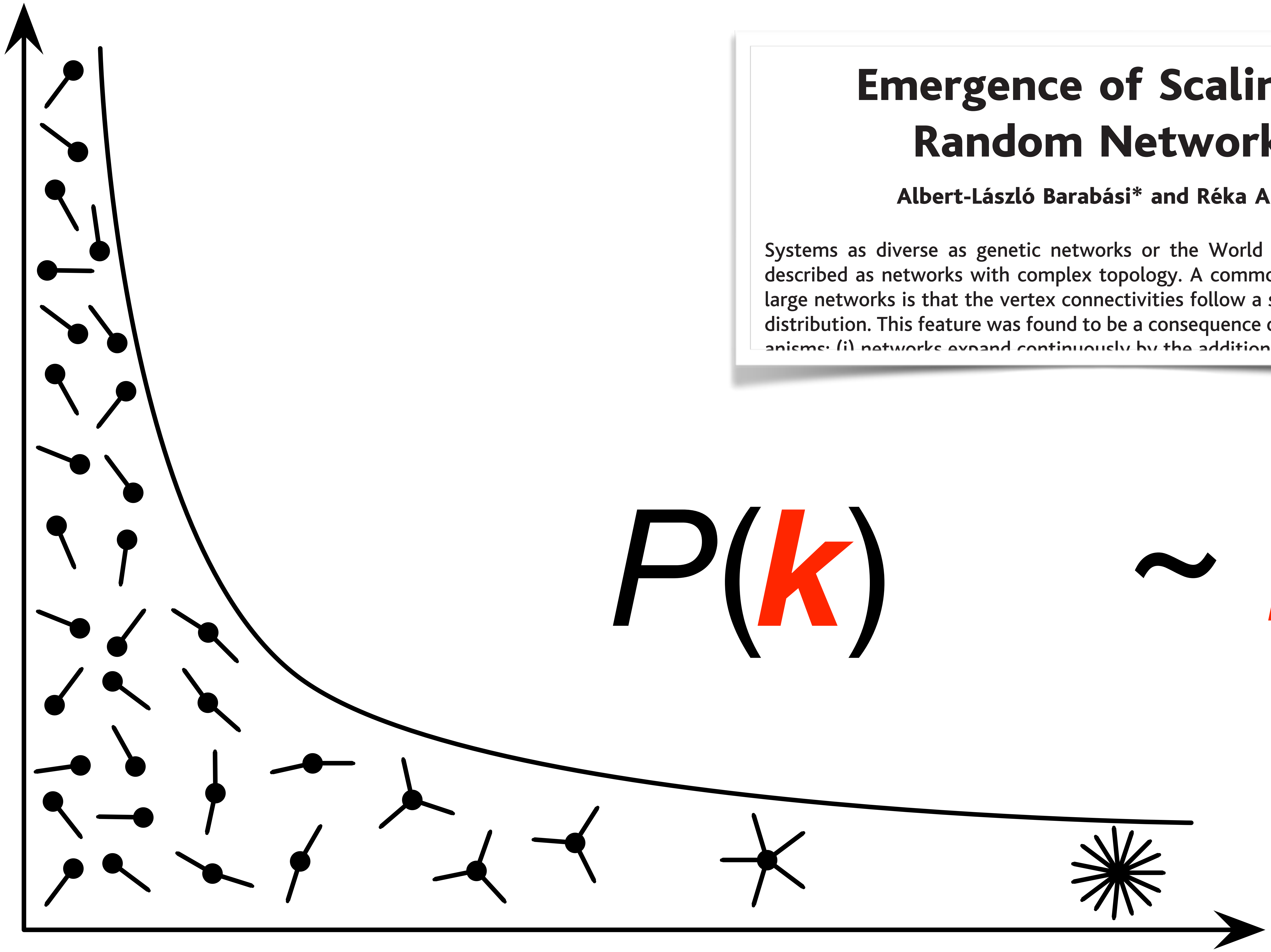
Erdős-Rényi Graph



Erdős-Rényi Graph



Scale-free networks



k

Emergence of Scaling in Random Networks

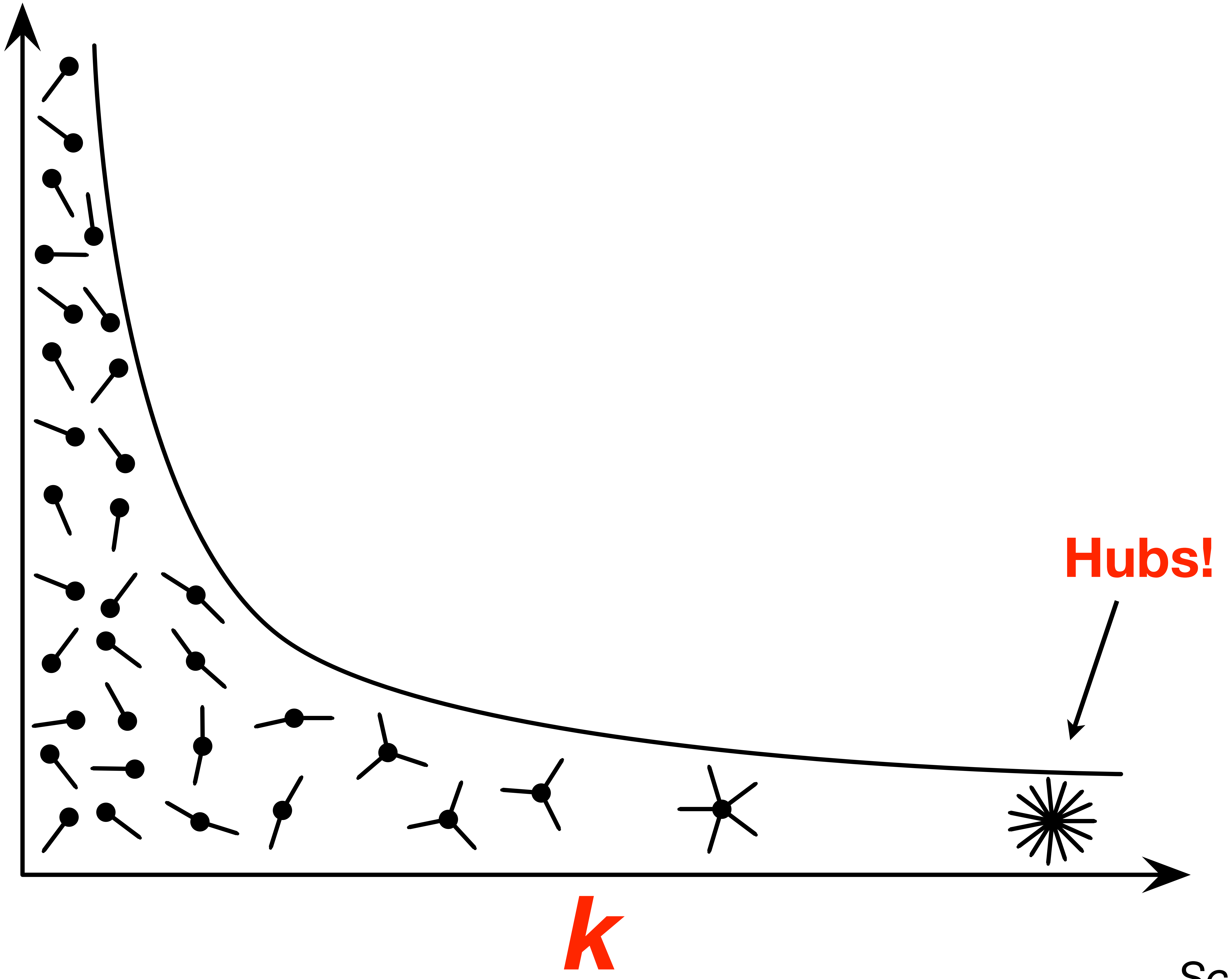
Albert-László Barabási* and Réka Albert

Systems as diverse as genetic networks or the World Wide Web are best described as networks with complex topology. A common property of many large networks is that the vertex connectivities follow a scale-free power-law distribution. This feature was found to be a consequence of two generic mechanisms: (i) networks expand continuously by the addition of new vertices, and

$$P(k) \sim k^{-\lambda}$$

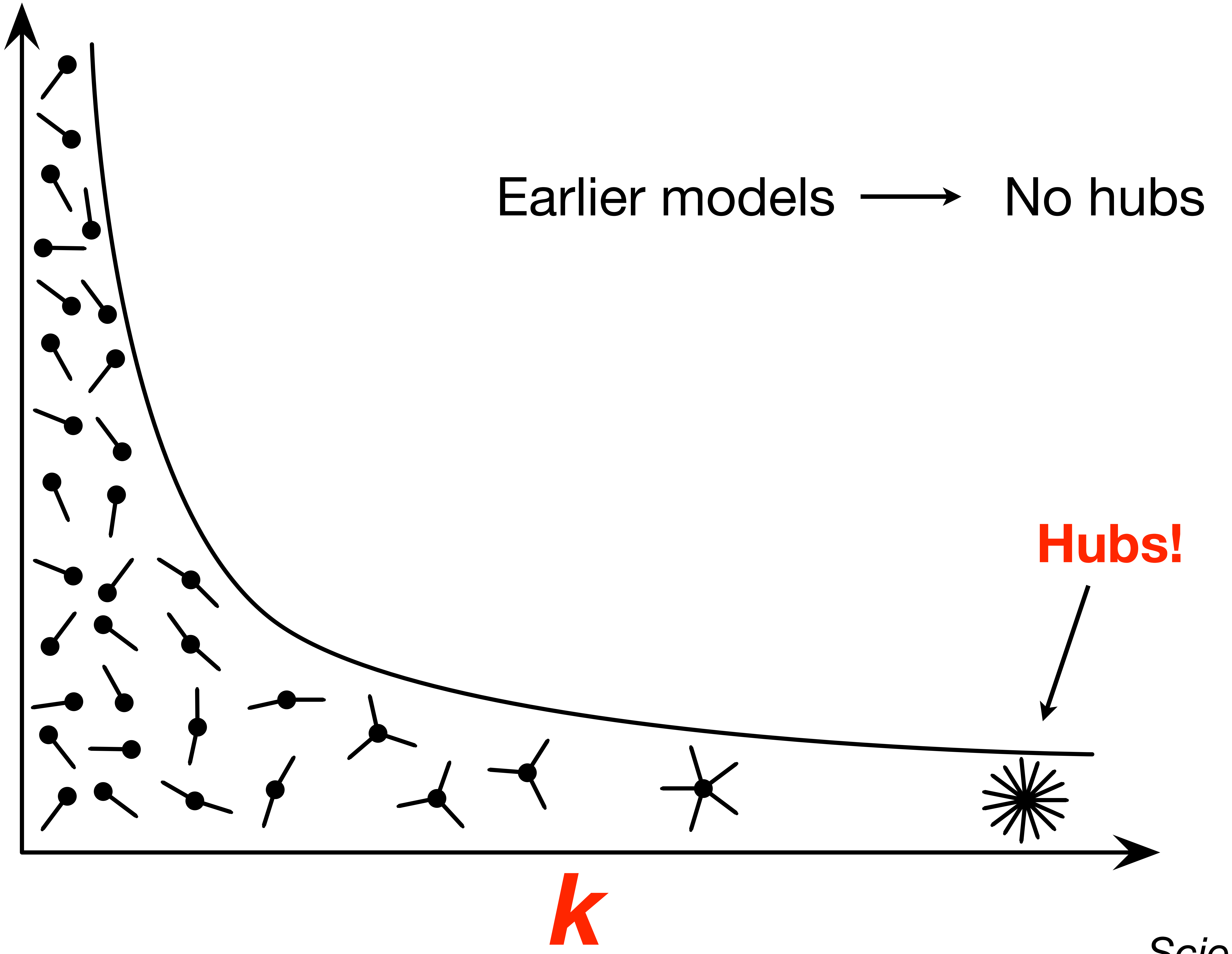
Science, 1999

Scale-free networks



Science, 1999

Scale-free networks



Science, 1999

Barabási-Albert Model

Barabási-Albert Model

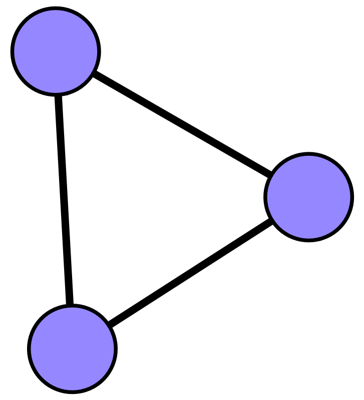
→ **Growing** network model

Barabási-Albert Model

⟶ **Growing** network model

1. start with initial graph



Barabási-Albert Model

Growing network model

1. start with initial graph

2. give **birth** to a new node

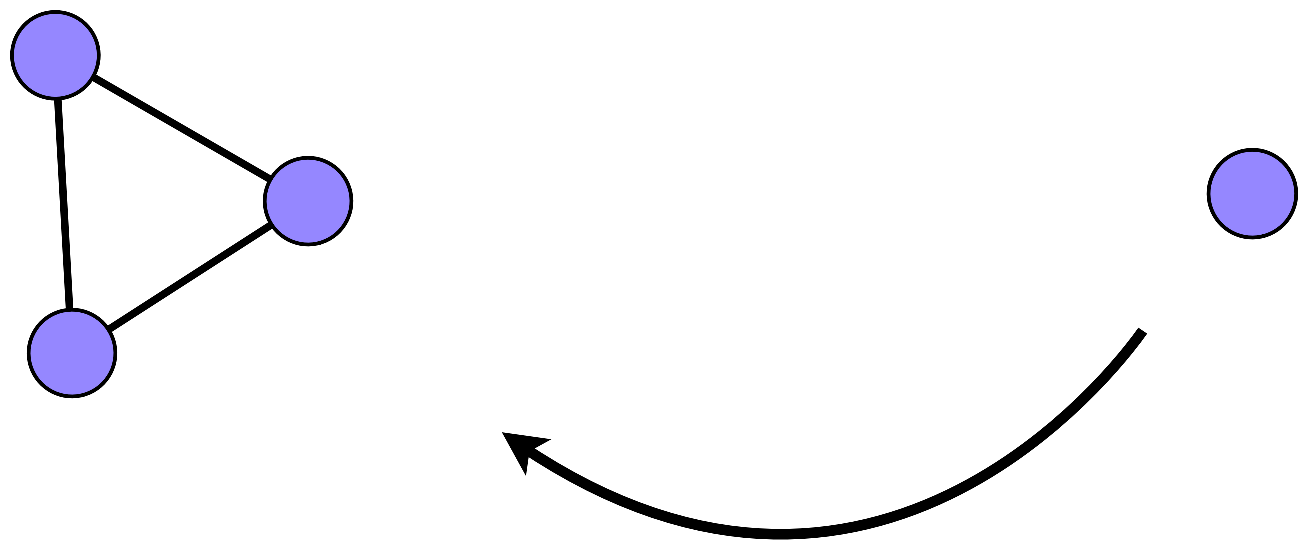
Barabási-Albert Model

→ **Growing** network model

1. start with initial graph

2. give **birth** to a new node



3. **attach** node to graph

Barabási-Albert Model

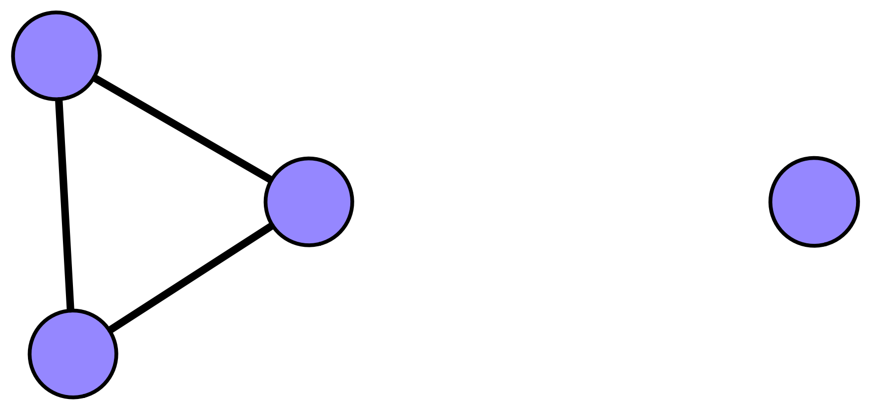
→ **Growing** network model

1. start with initial graph

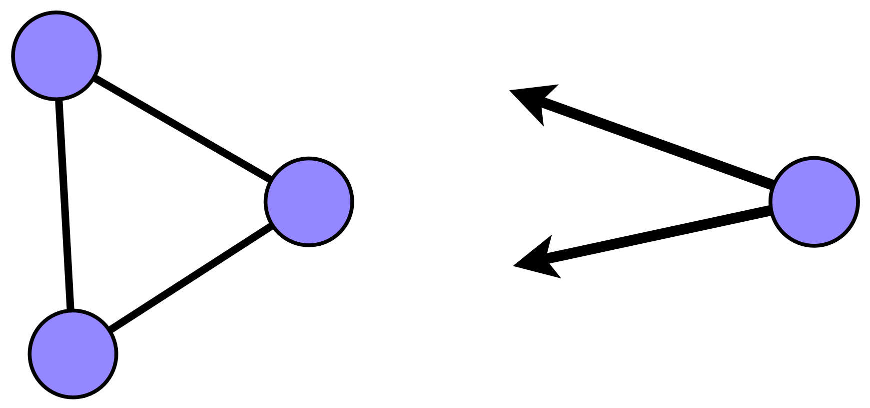
2. give **birth** to a new node

3. **attach** node to graph

Barabási-Albert Model

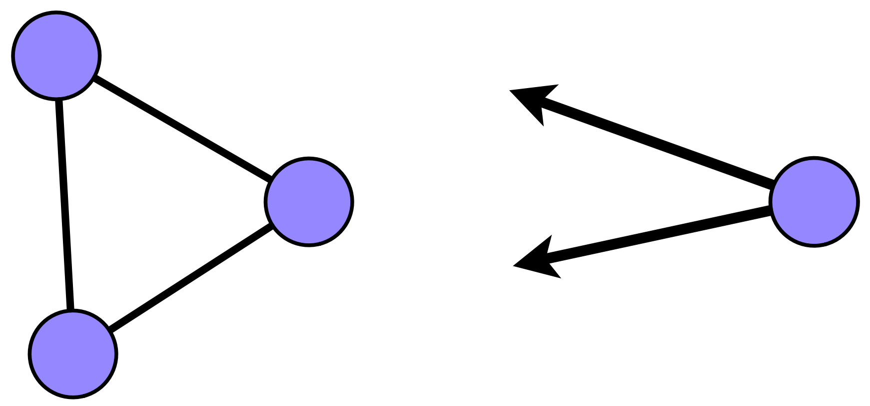


Barabási-Albert Model



Each timestep new node **attaches** to existing nodes

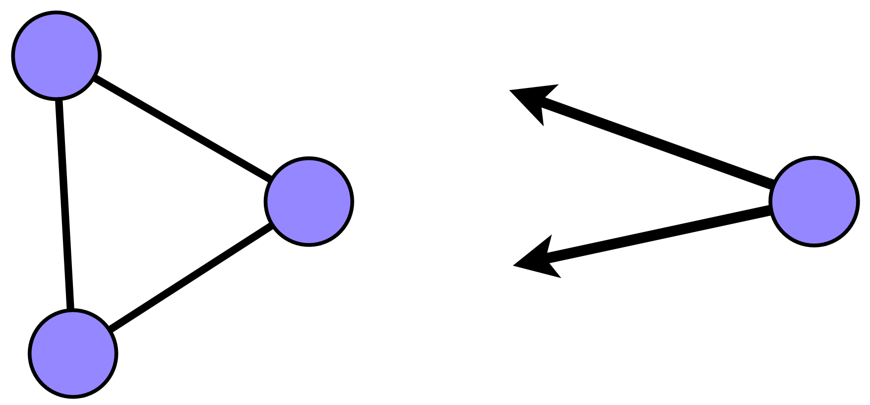
Barabási-Albert Model



Each timestep new node **attaches** to existing nodes

How?

Barabási-Albert Model

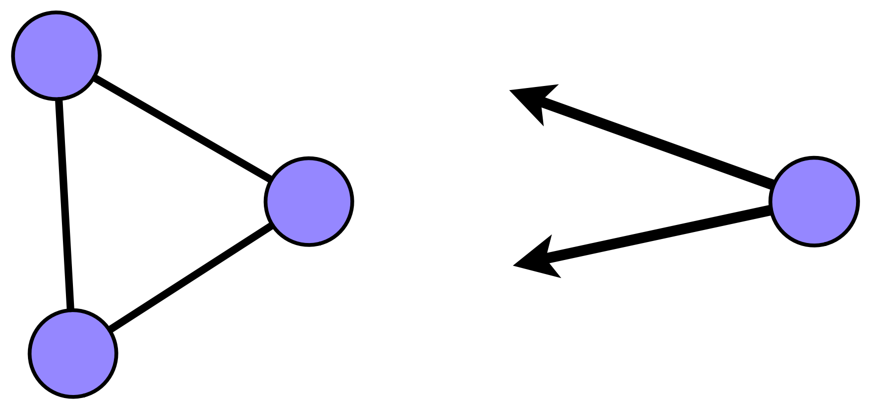


Each timestep new node **attaches** to existing nodes

How?

Rich-get-richer ⟶ **Preferential Attachment**

Barabási-Albert Model



Each timestep new node **attaches** to existing nodes

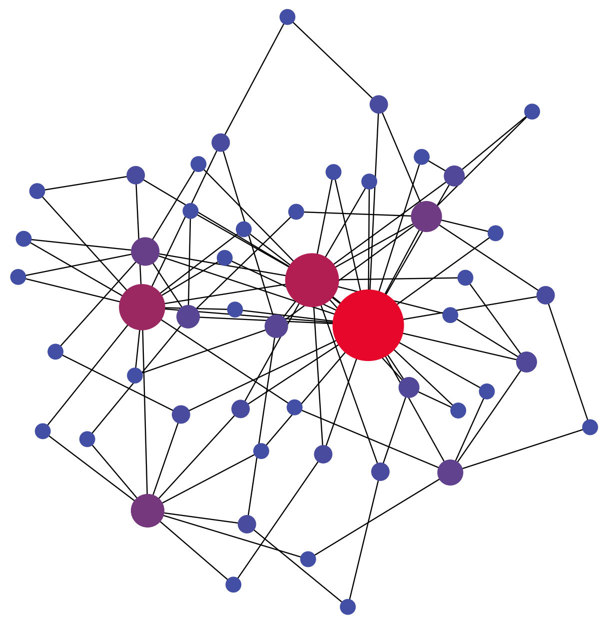
How?

Rich-get-richer ⟶ **Preferential Attachment**

Link to existing node *i* with probability proportional to degree of i

Barabási-Albert Model

Growth + **Preferential Attachment** =

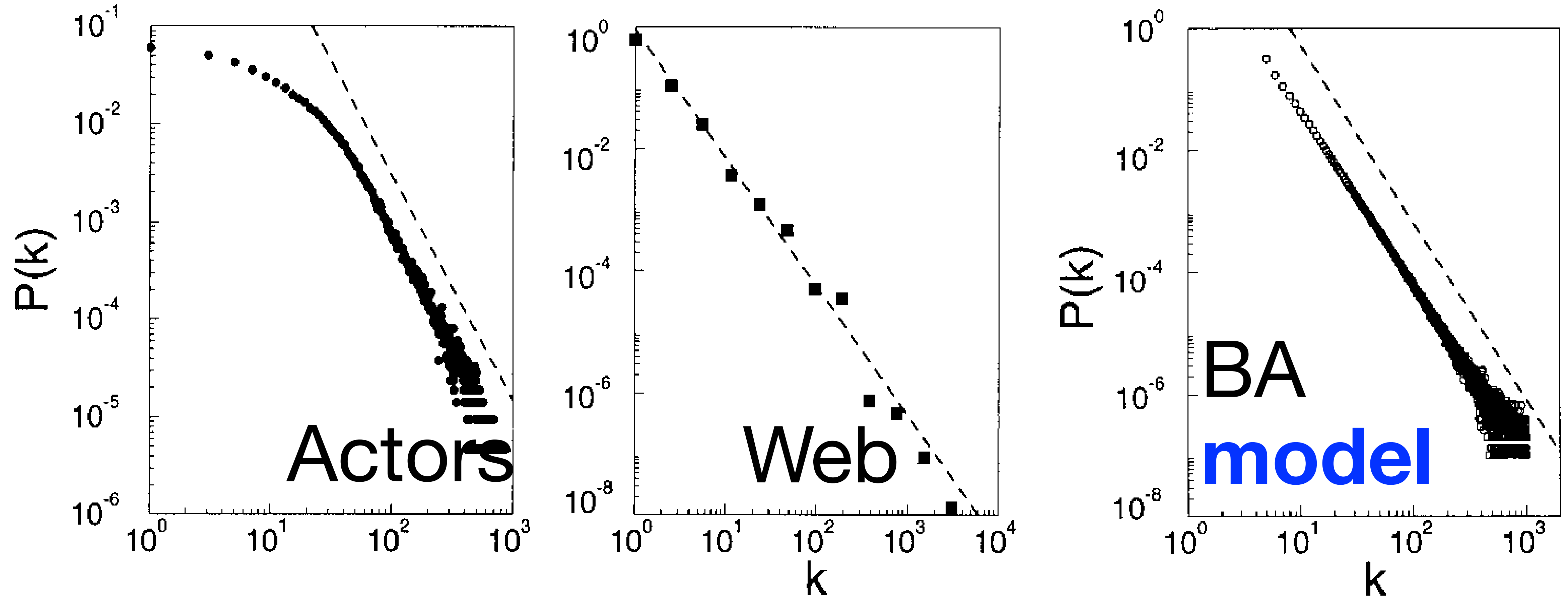


Not the first:

Yule 1924

Simon 1955

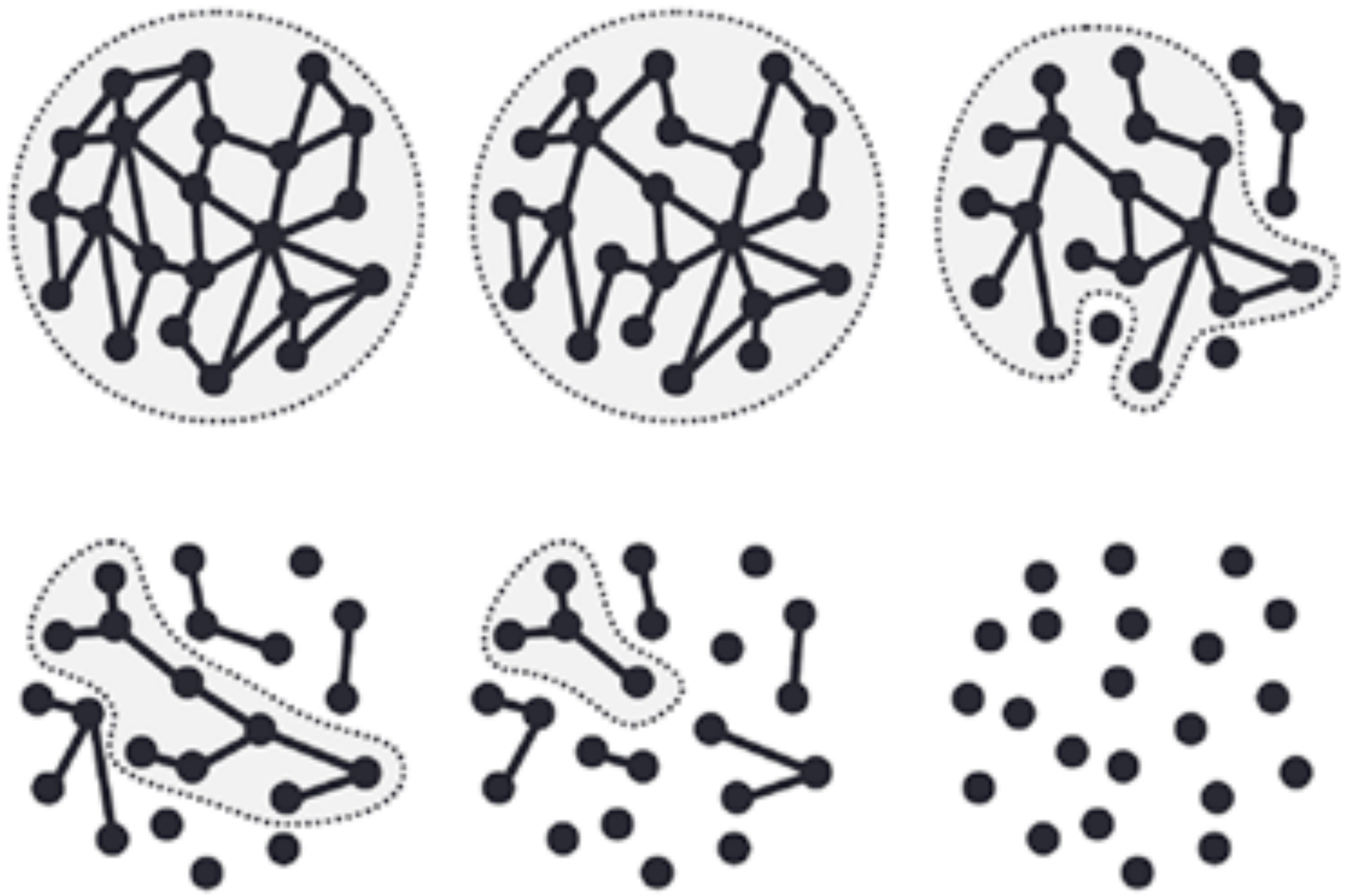
But they found it in **new systems**



Network robustness

Percolation

How do networks respond to random failures?



Is global damage **gradual** or **sudden**?

Percolation

How do networks respond to random failures?

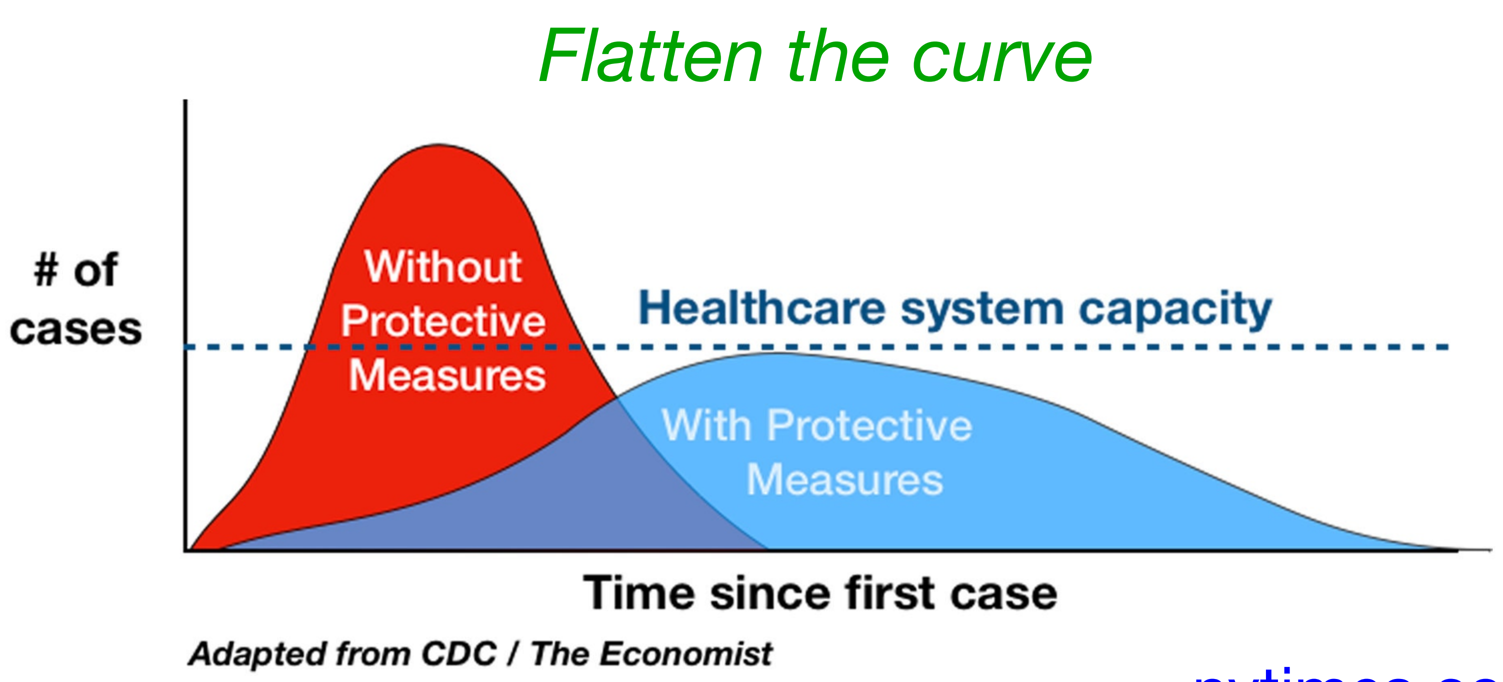
Affects dynamics

Contact network — pandemic

easy to spread

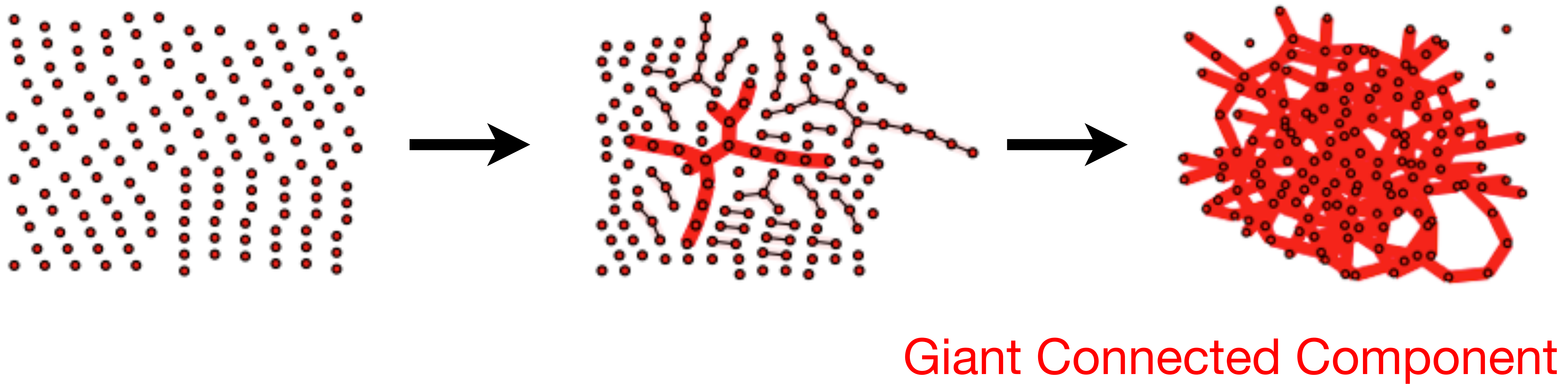
hard to spread

social distancing ≈ percolation 🤔



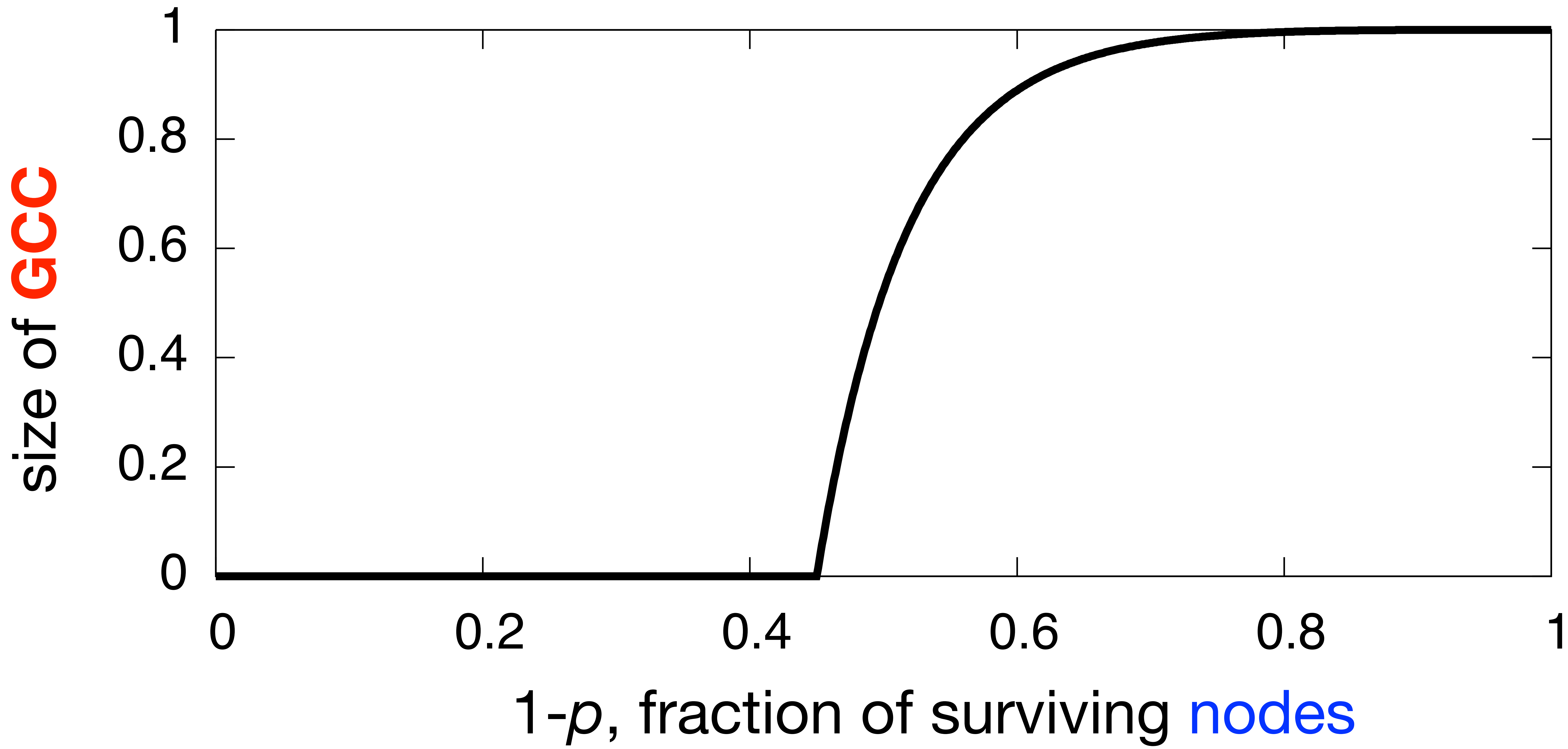
nytimes.com

Percolation

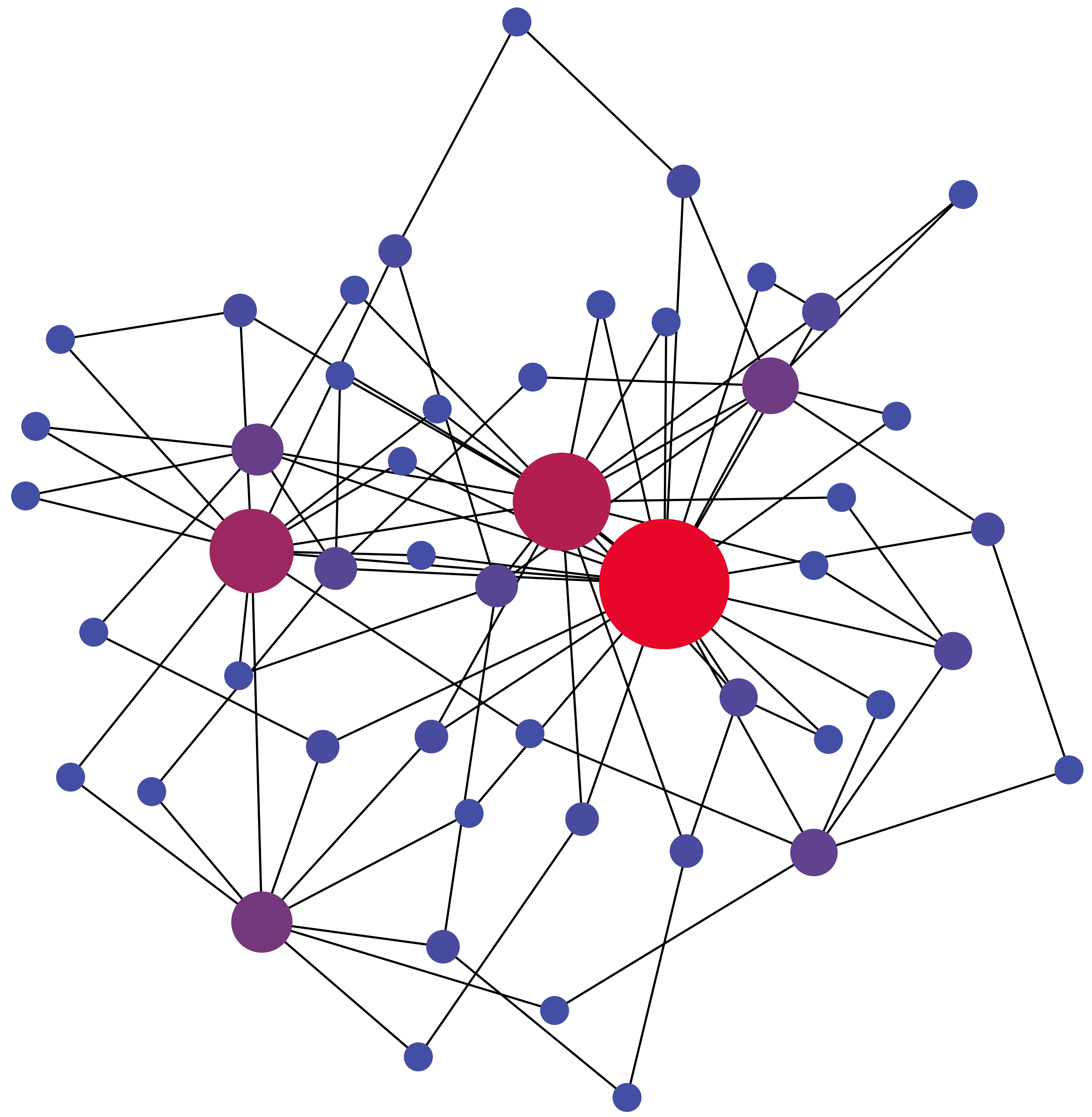


Many systems show a **sharp transition** in connectivity

Percolation transition



Some networks are special



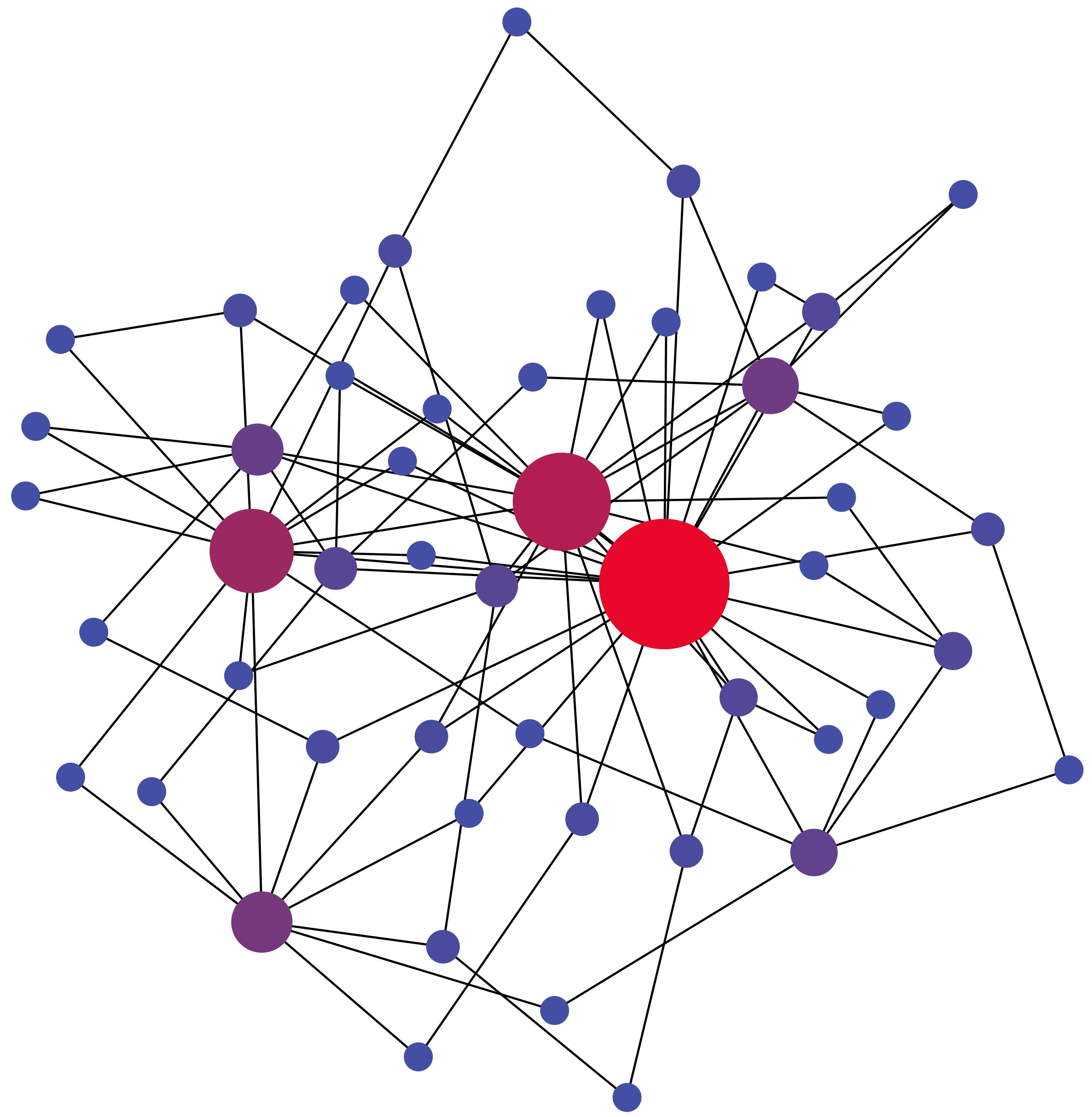
Scale-free graphs

Hubs!

Imagine I randomly remove nodes from a very large scale-free network

Unlikely I'll hit all the hubs

Some networks are special



Scale-free graphs	Hubs!
Imagine I randomly remove nodes from a very large scale-free network	Unlikely I'll hit all the hubs
Hubs do a disproportionate job gluing the network together	Very unlikely I can make the network fall apart

Some networks are special

Scale-free networks robust to **random failures**

Error and attack tolerance of complex networks

Réka Albert, Hawoong Jeong & Albert-László Barabási

Department of Physics, 225 Nieuwland Science Hall, University o
Notre Dame, Indiana 46556, USA

Nature, 2000

Resilience of the Internet to Random Breakdowns

Reuven Cohen,[1,*] Keren Erez,[1] Daniel ben-Avraham,[2] and Shlomo Havlin[1]
Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, I
[2]*Physics Department and Center for Statistical Physics (CISP), Clarkson Universit*
Potsdam, New York 13699-5820
(Received 11 July 2000; revised manuscript received 31 August 2000)

Phys Rev Lett, 2000

But there’s a price

Scale-free networks robust to **random failures**

But there's a price

Scale-free networks robust to **random failures**

Scale-free networks **vulnerable** to **targeted attacks**!

But there's a **price**

Scale-free networks robust to **random failures**

Error and attack tolerance of complex networks

Réka Albert, Hawoong Jeong & Albert-László Barabási

Department of Physics, 225 Nieuwland Science Hall, University o
Notre Dame, Indiana 46556, USA

Nature, 2000

Scale-free networks **vulnerable** to **targeted attacks**!

Breakdown of the Internet under Intentional Attack

Reuven Cohen,[1,*] Keren Erez,[1] Daniel ben-Avraham,[2] and Shlomo Havlin[1]

[1]*Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan, Israel*
[2]*Department of Physics, Clarkson University, Potsdam, New York 13699-5820*
(Received 17 October 2000)

Phys Rev Lett, 2001

Deleting a **small number of hubs** will **drastically disconnect** the network

Percolation

Tons of results and variants

Explosive percolation

Cascading failures

....

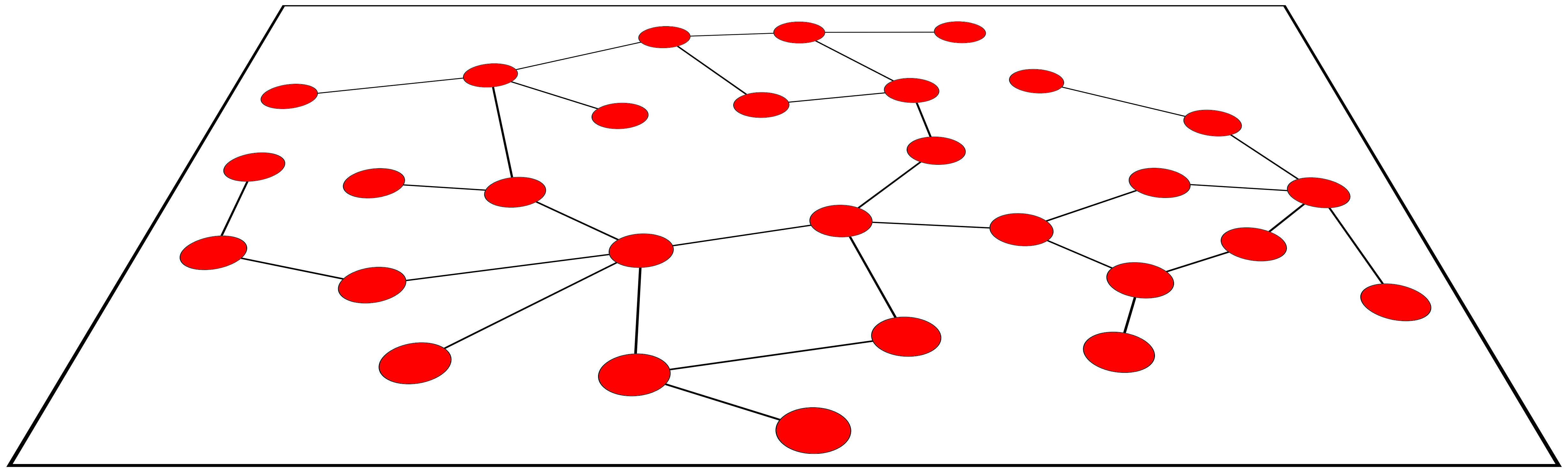
Lots of applications

Many complex problems map into simple percolation models

Epidemics: Vaccinations, Social distancing

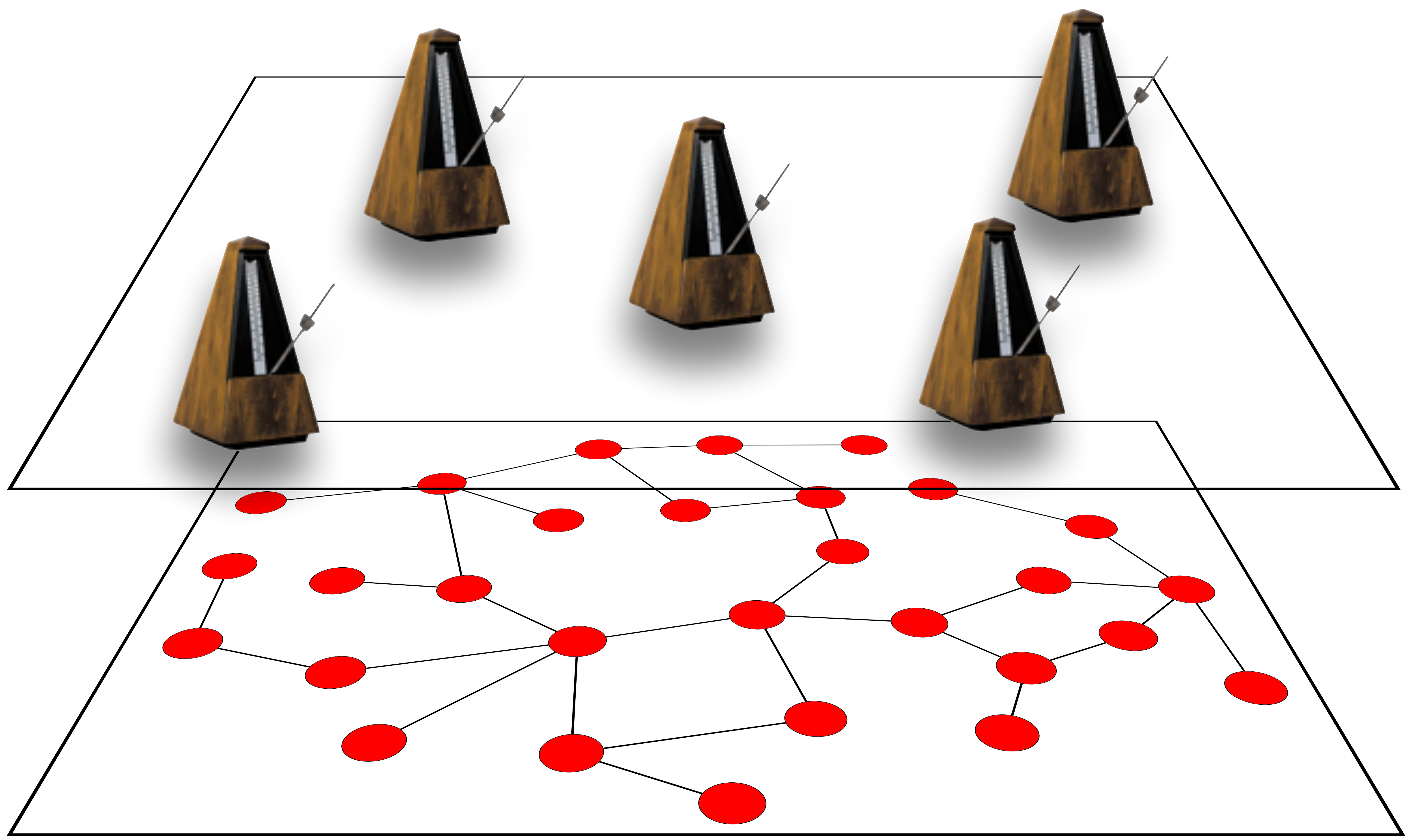
Statistics: Missing data

Dynamics on networks



Network substrate

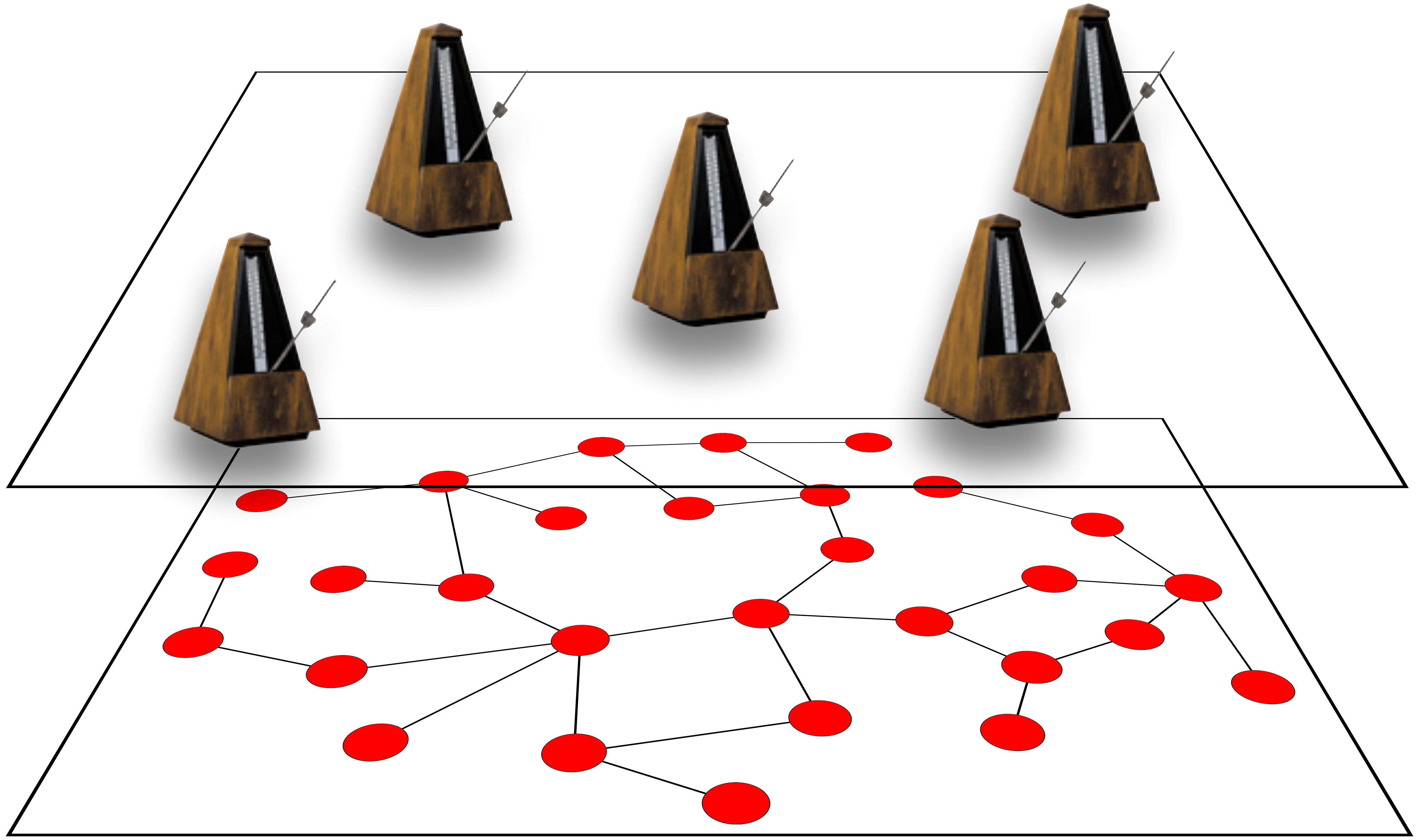
Dynamics on networks



Dynamical system
on top

Network substrate

Dynamics on networks



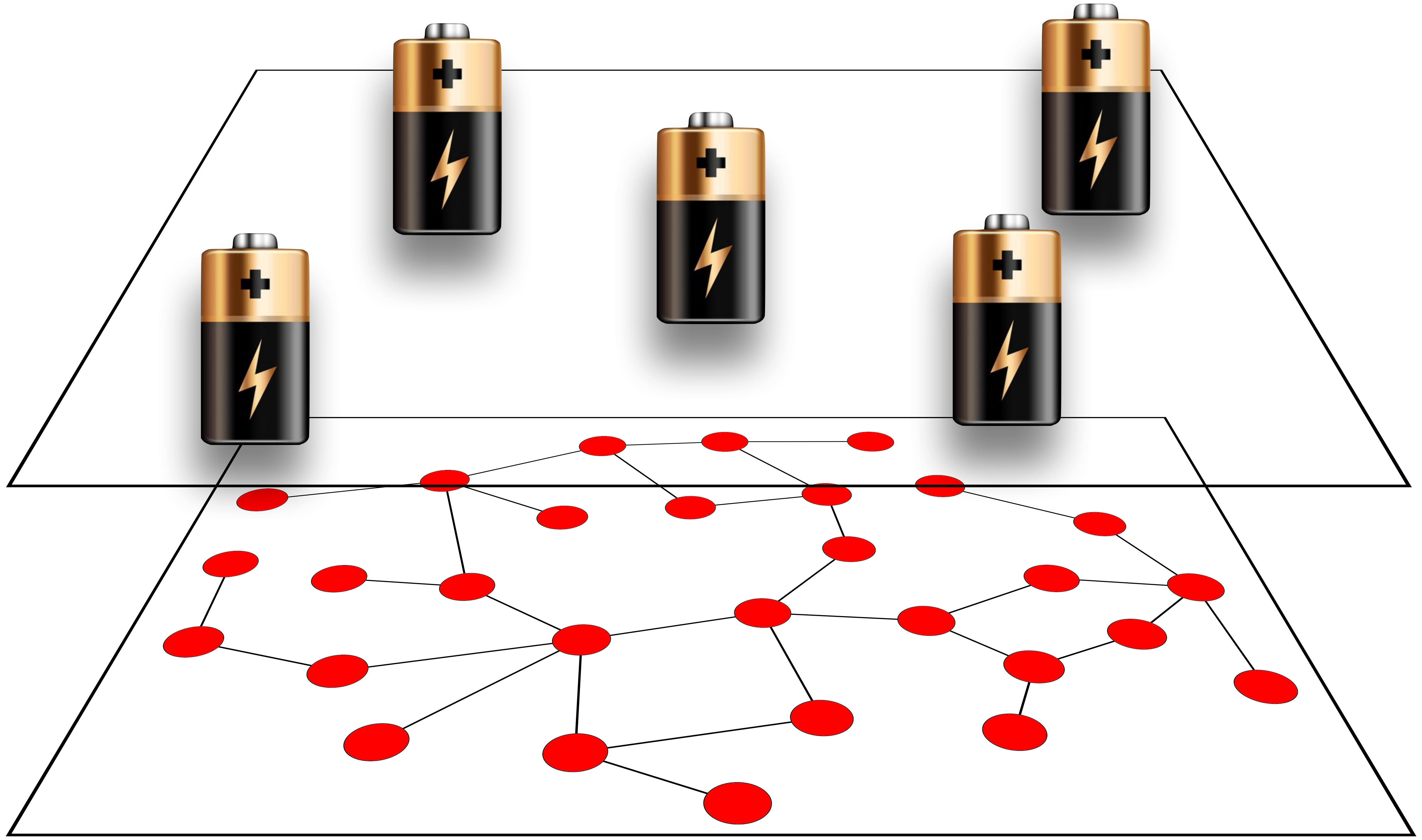
Synchronization of coupled oscillators

↑

Network substrate

Ex: Acebrón et al. (2005)

Dynamics on networks

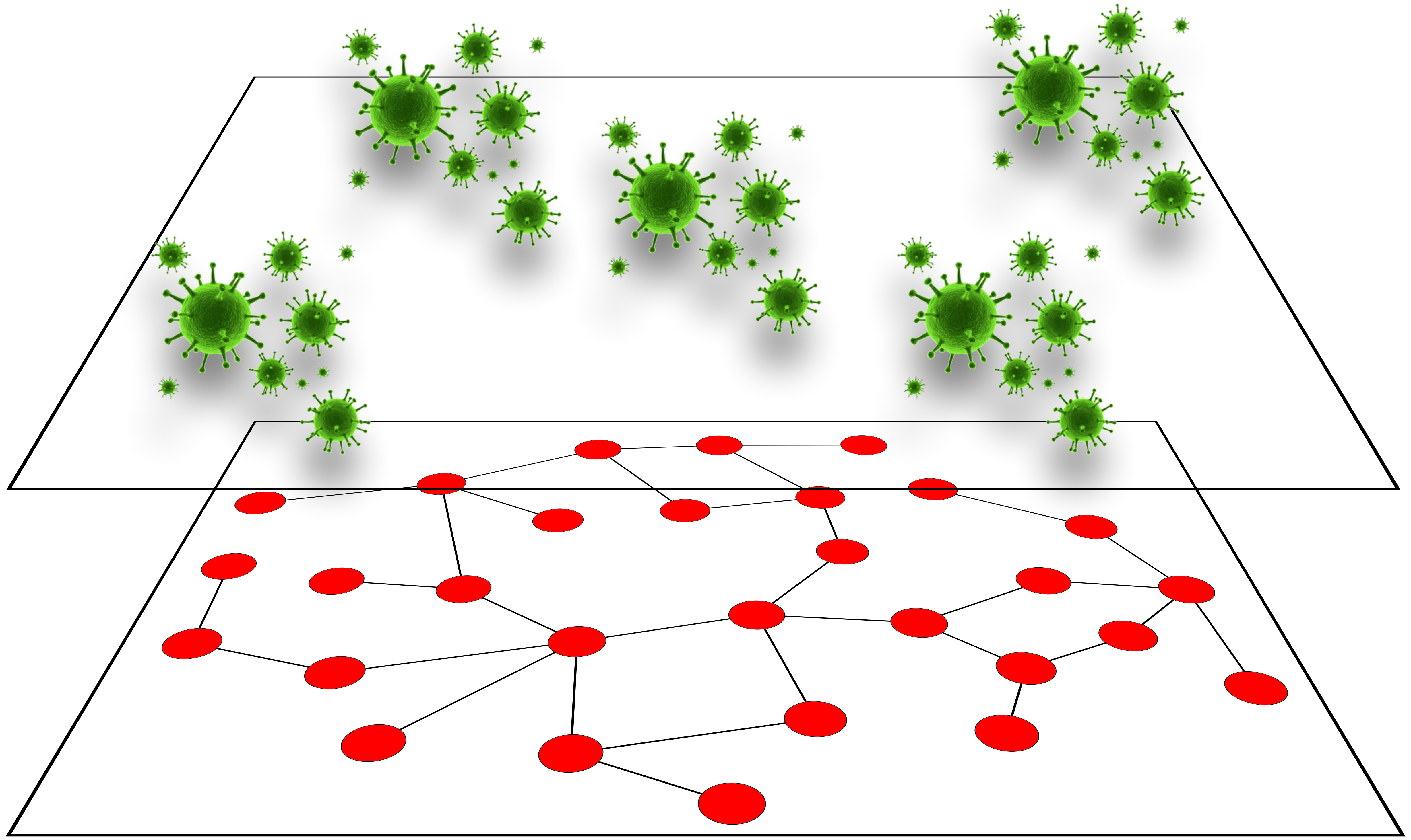


DC Load Flow

Network substrate

Ex: Dobson et al. (2007)

Dynamics on networks



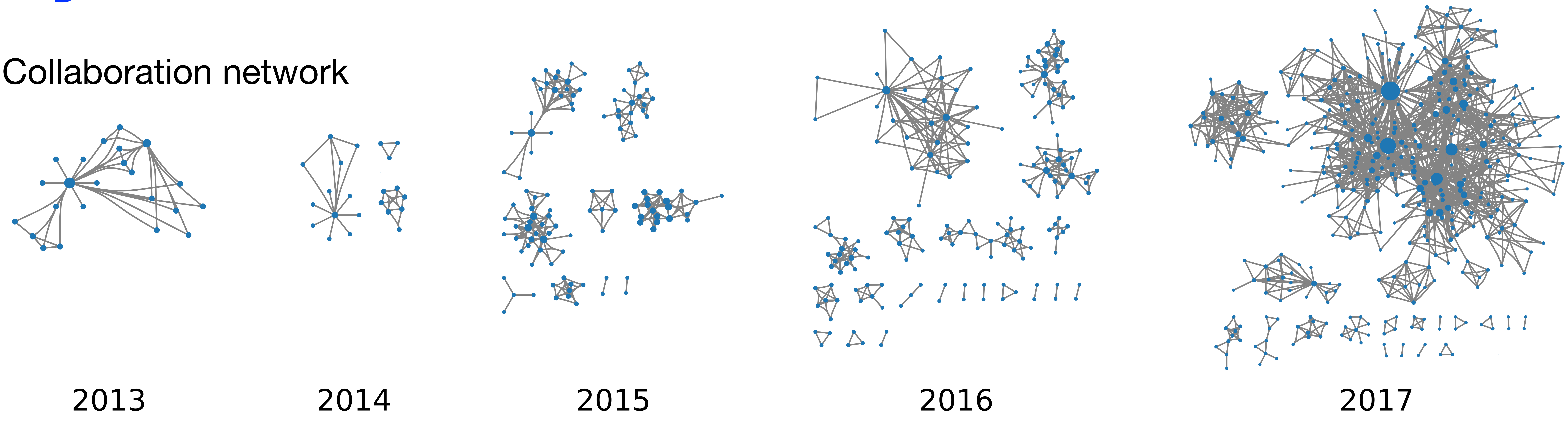
Contagion

Network substrate

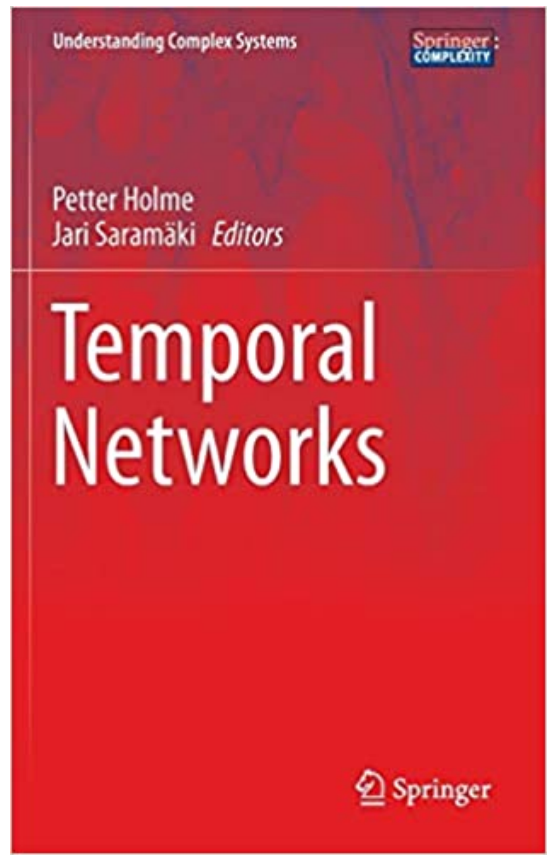
Ex: Pastor-Satorras and Vespignan

Dynamics *of* networks

Collaboration network



Bagrow and Bollt (2019)



Holme, Saramäki
eds (2013)

see also:

multiplex, multilayer networks

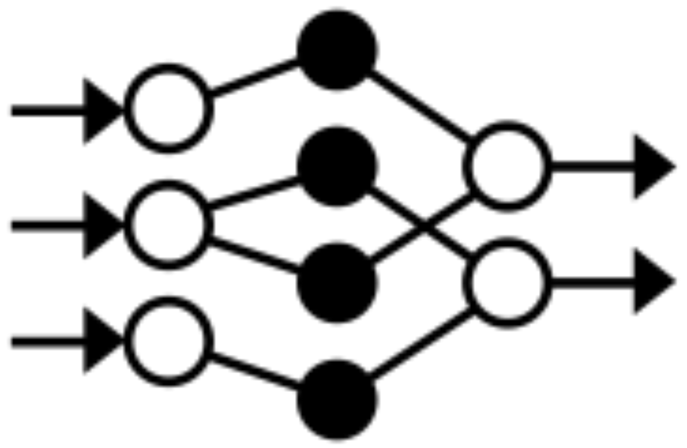
What's next?

Future of network science

Brain imaging, large-scale studies (ABCD study)

Single cell genomics and spatial gene expression

Social information flow, misinformation, disinformation



Machine Learning: Graph Neural Networks

→ DATA

Summary

Introduction to Network Science

- Network examples/data
- Why study networks?
- Types of networks
- Network quantifiers (jargon!)
- Random network models
- Network robustness
- Dynamics on networks
- Future of network science

Introduction to Network Science

- Network examples/data
- Why study networks?
- Types of networks
- Network quantifiers (jargon!)
- Random network models
- Network robustness
- Dynamics on networks
- Future of network science

THANKS!!!