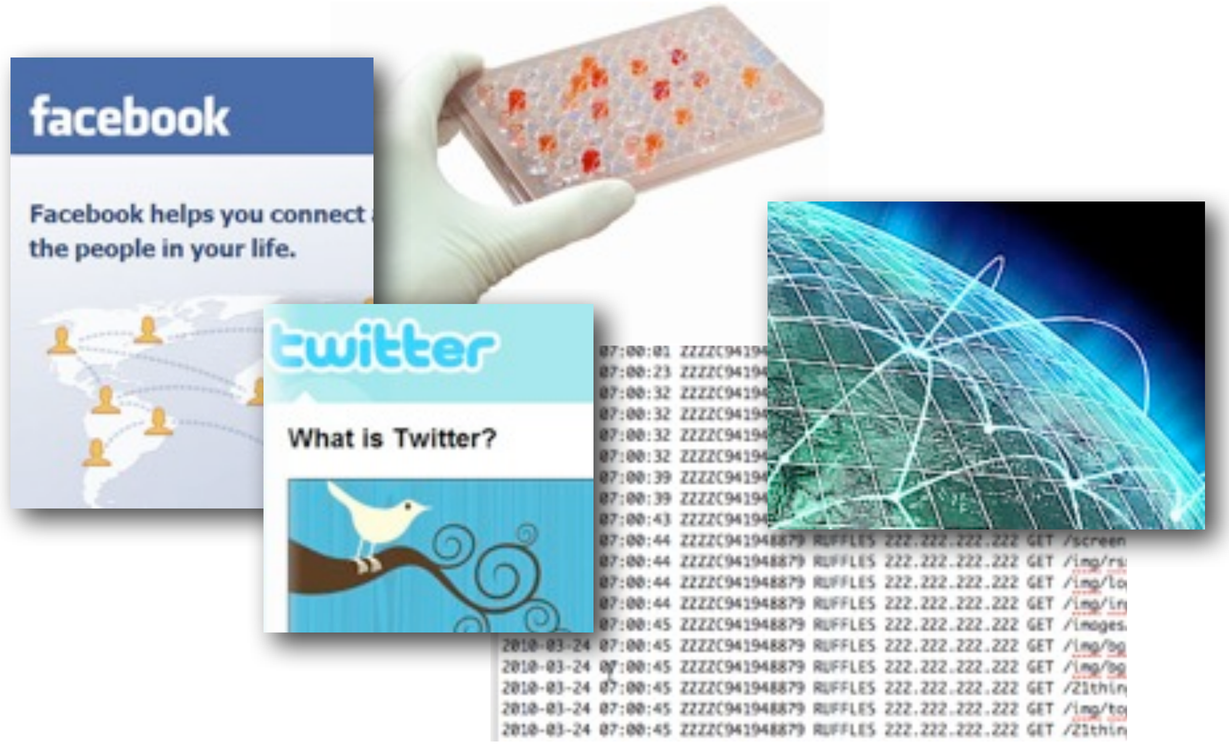
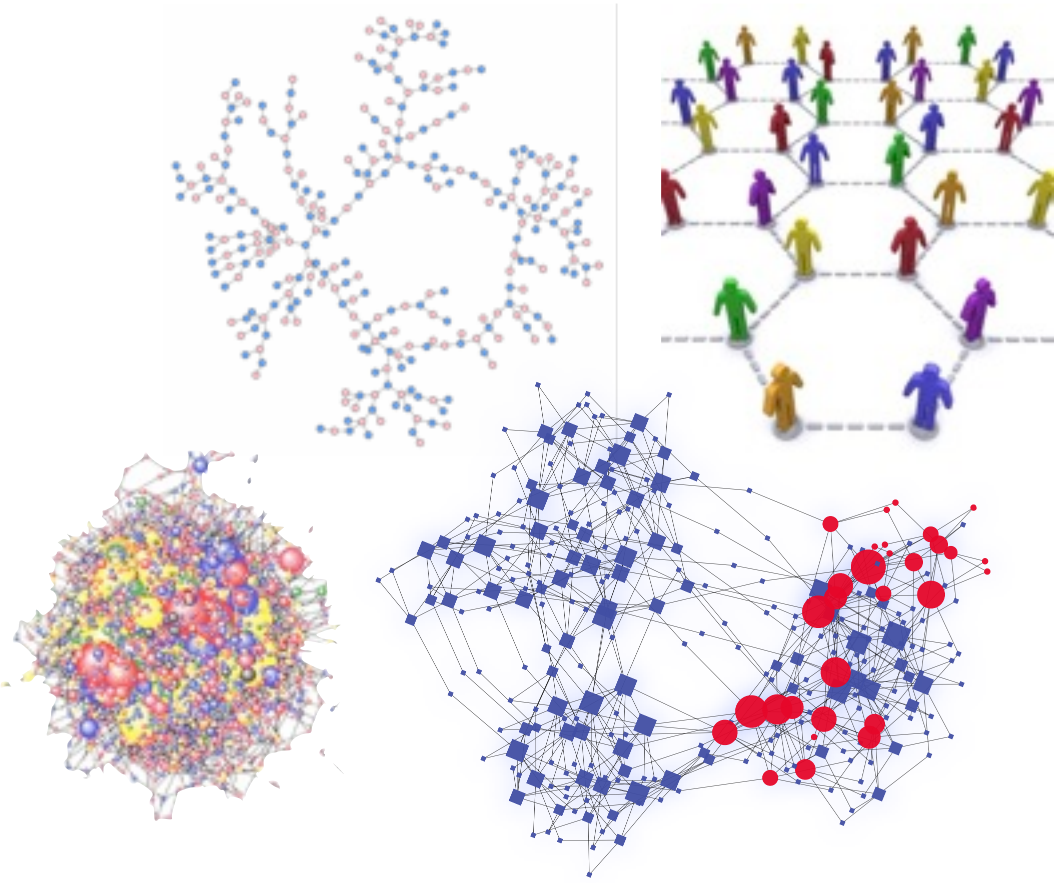


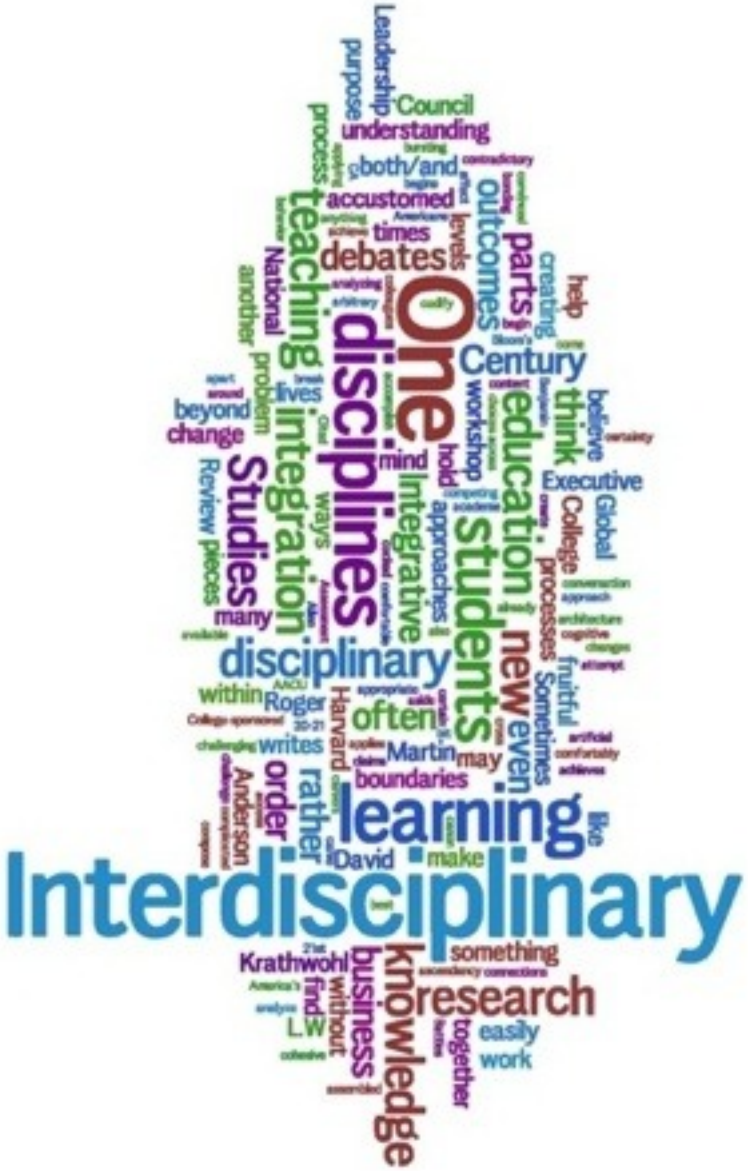


Introduction to Networks

Jim Bagrow

NetSci 2012 School
June 18, 2012

Networks



Sociology

Physics

Computer science

Biology

Economics

Math

Outline

Part I

- History
- Network examples/data
- Why study networks?
- Network quantifiers (**jargon**!)
- Types of networks
- Random network models

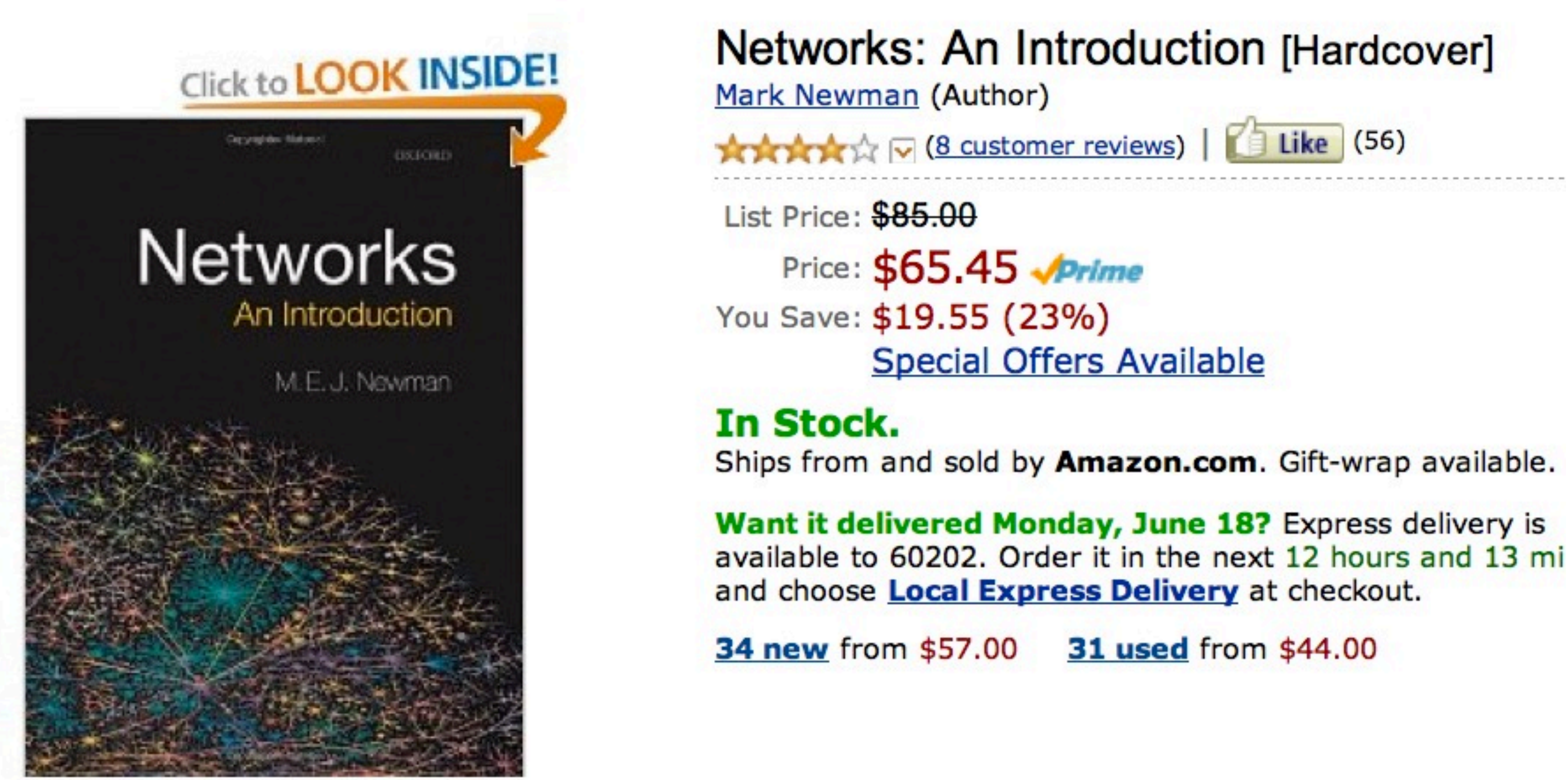
Part II

- Getting started on a computer
- Network search
- Network robustness
- Dynamics on networks

slides will be on

bagrow.com

Good reference



Networks: An Introduction [Hardcover]

Mark Newman (Author)

★★★★☆ (8 customer reviews) | Like (56)

List Price: ~~$85.00~~

Price: $65.45 ✓Prime

You Save: $19.55 (23%)

Special Offers Available

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered Monday, June 18? Express delivery is available to 60202. Order it in the next 12 hours and 13 mi and choose **Local Express Delivery** at checkout.

34 new from $57.00 **31 used** from $44.00

Part I

History

Leonhard Euler

Swiss mathematician

- Exceptionally prolific and influential
- He introduced **functions**!
  - First to write *f*(*x*)

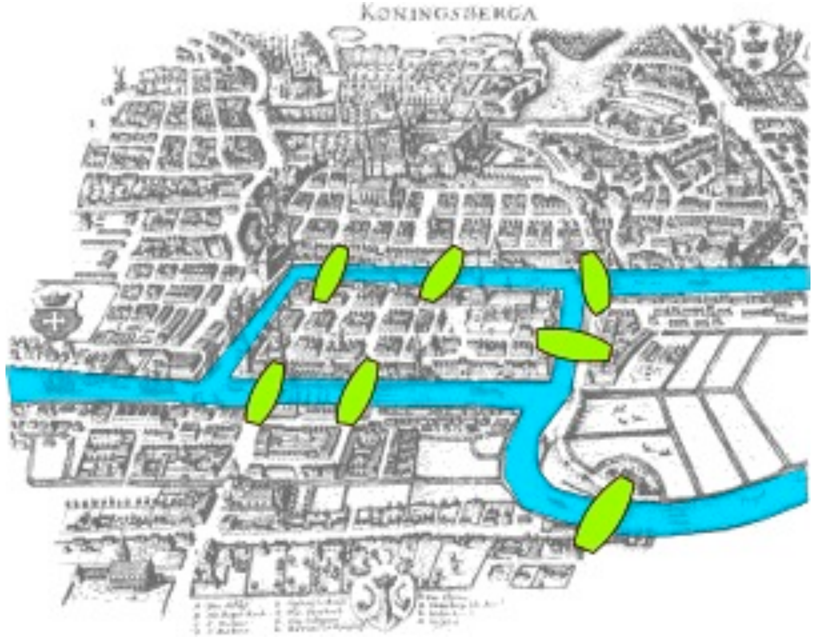
866 works, first publication age 19

The Euler Archive

Seven bridges

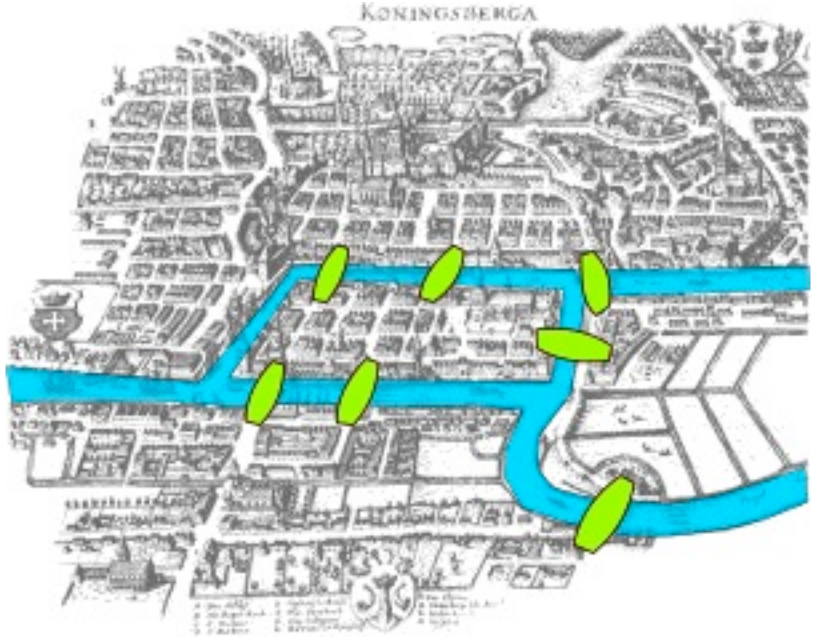
Seven bridges

Königsberg



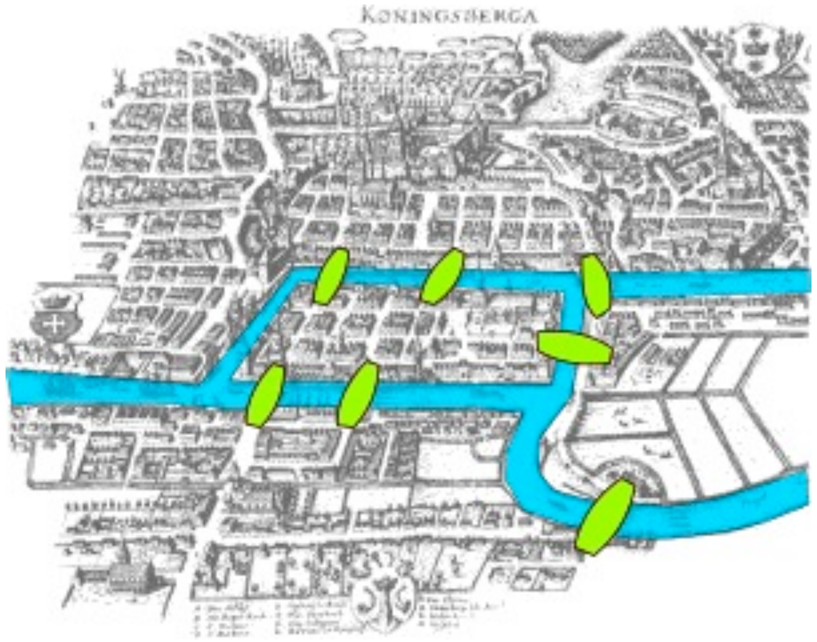
Seven bridges

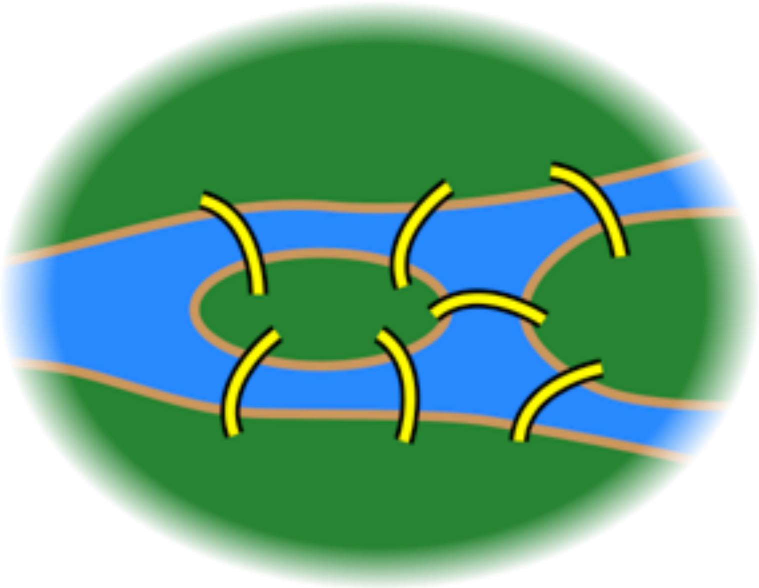
Königsberg



"Can I walk through the city and cross each of the seven bridges **exactly once**?"

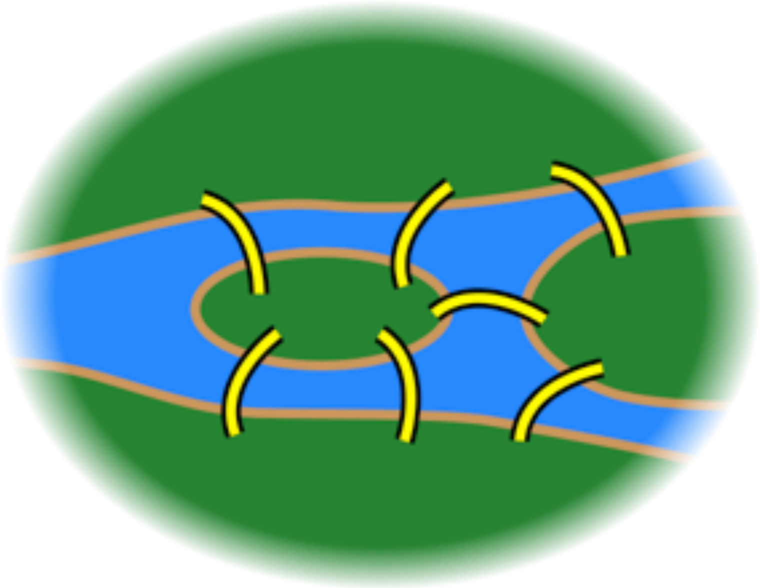
Seven bridges

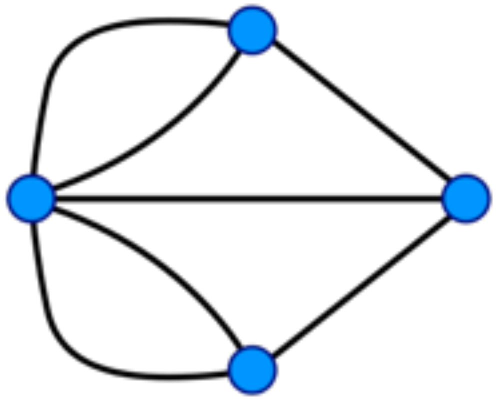




Abstract away details

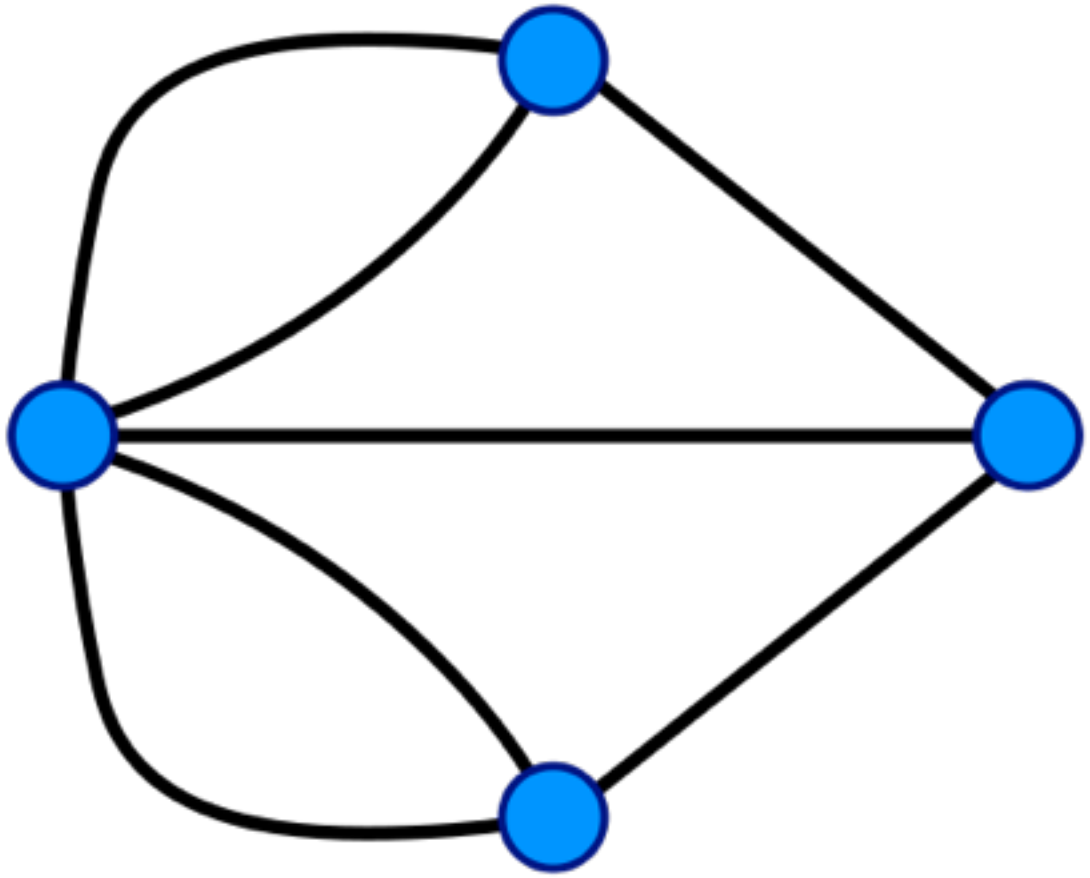
Seven bridges





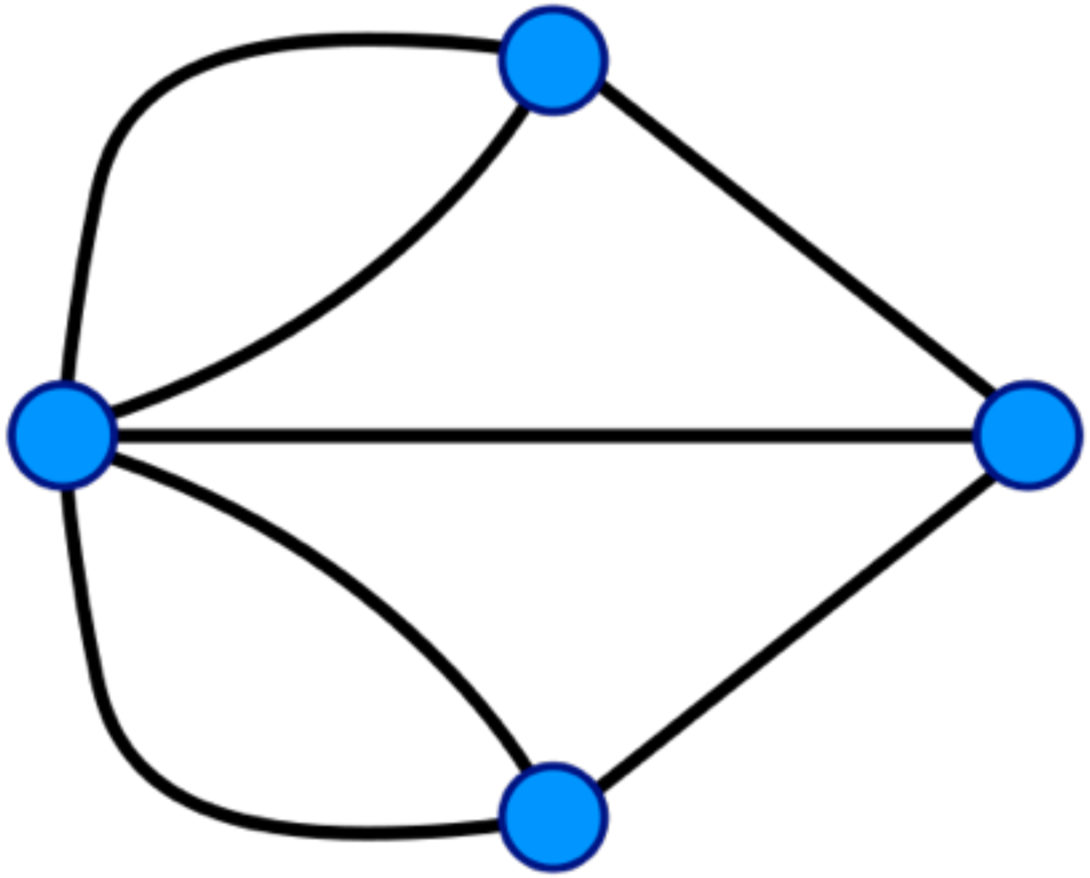
Seven bridges

Graph theory!



Seven bridges

Graph theory!

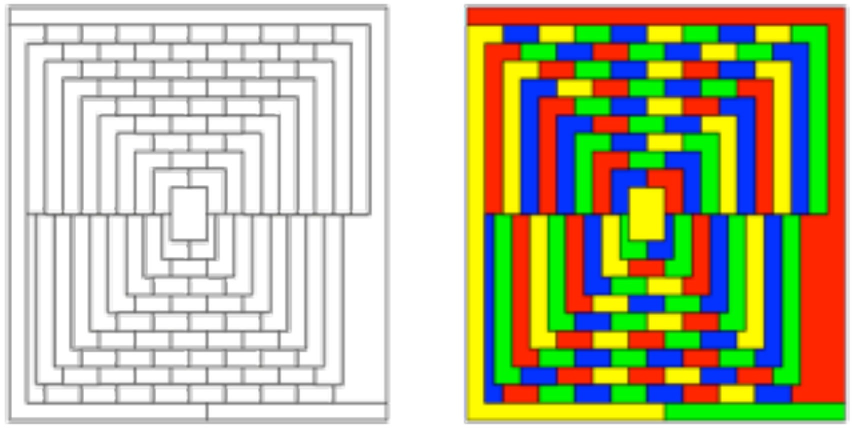


A **graph** is an object consisting of:

- ● **nodes** (vertices)
- ╲ **links** (edges) between those nodes

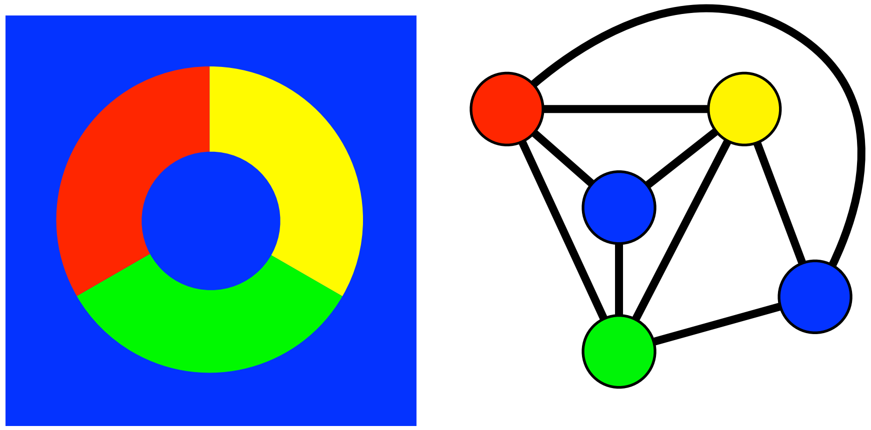
Four-color theorem

"To color any **map** of countries without adjacent countries sharing the same color requires only **four colors**"



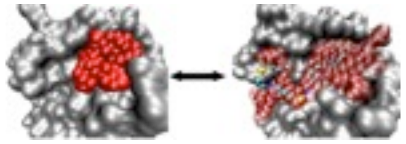
Four-color theorem

“To color any **map** of countries without adjacent countries sharing the same color requires only **four colors**”



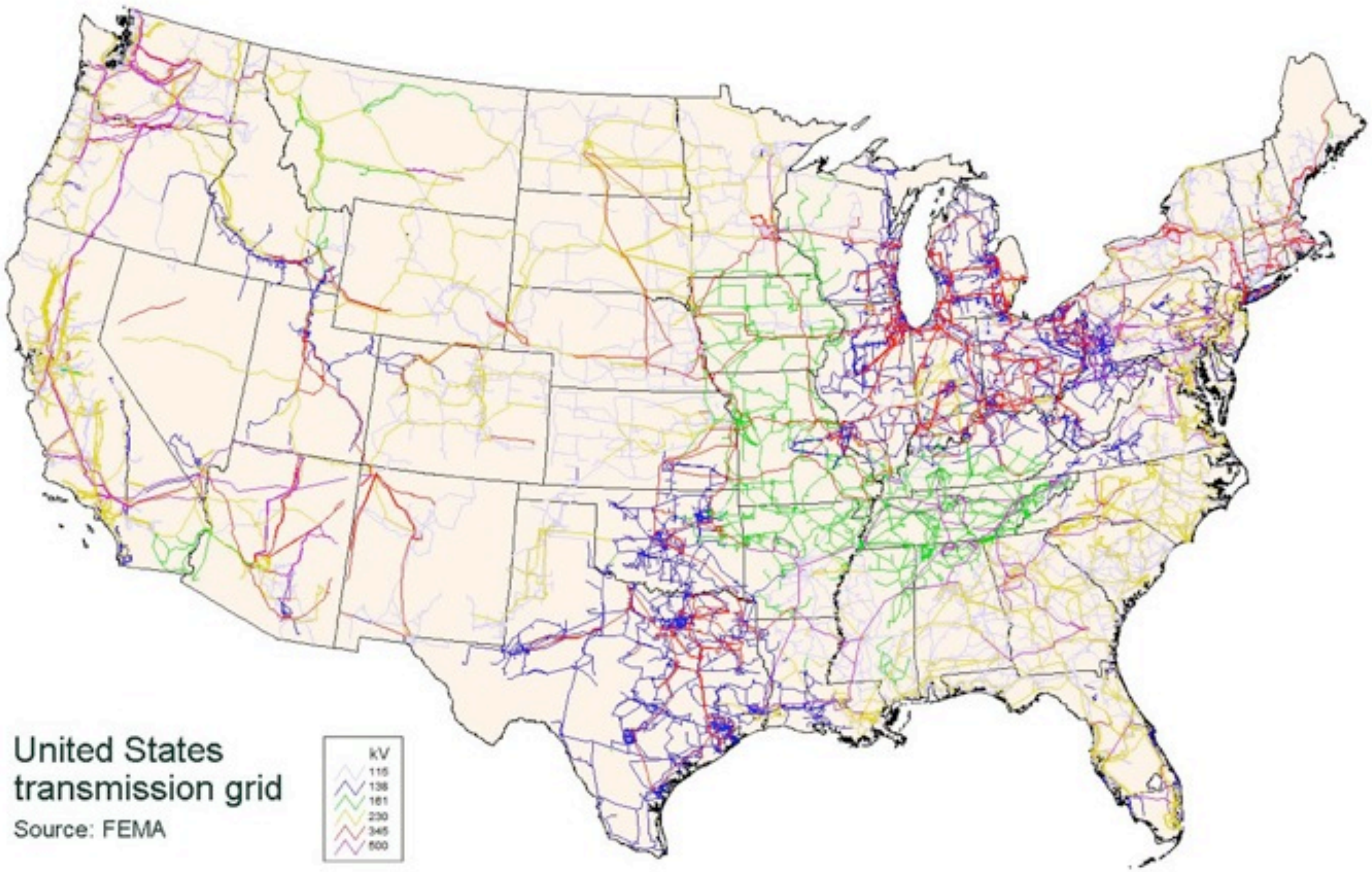


Examples of networks and network data


Technology & Infrastructure

Power grid



United States transmission grid

Source: FEMA

Power grid

Nodes: power generators/ consumers

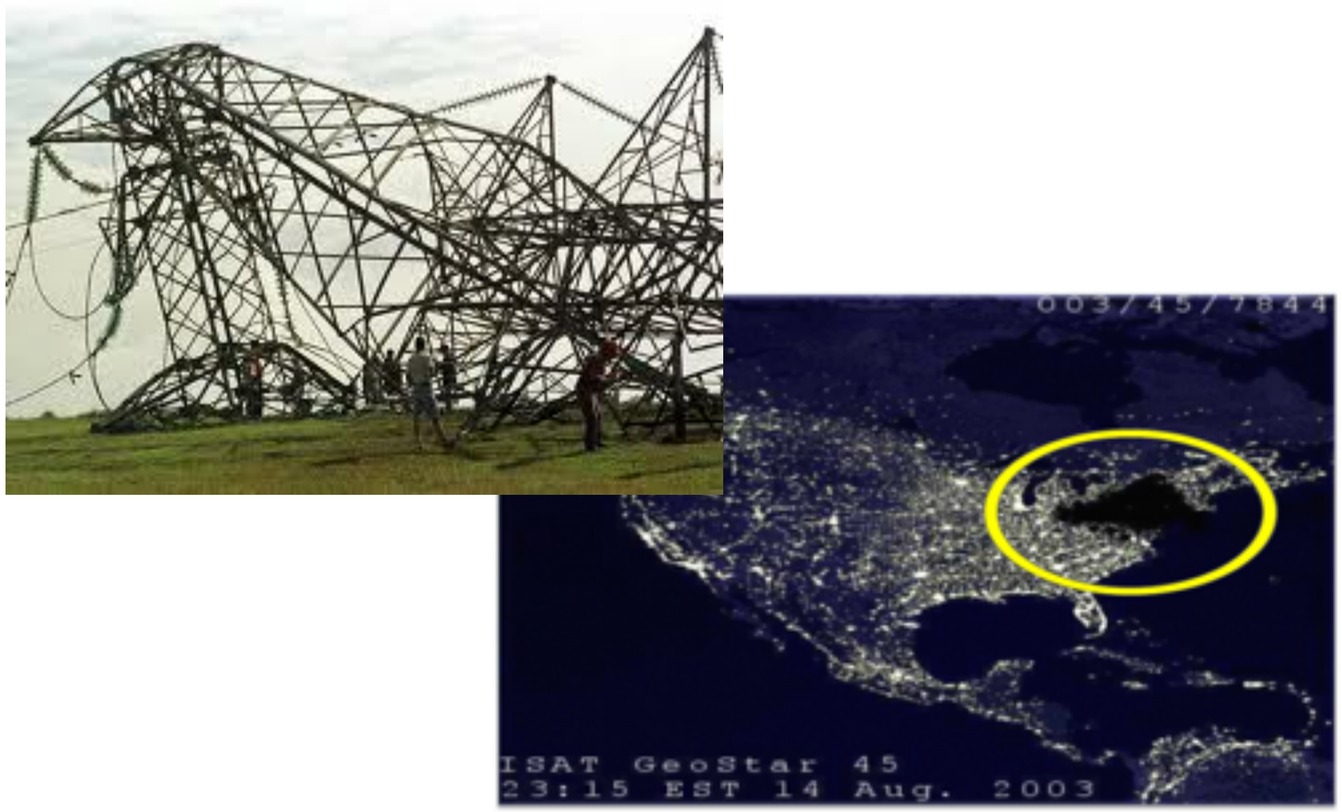
Power grid

Nodes: power generators/ consumers

Links: transmission lines

Power grid

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



Road networks

International roads 1990

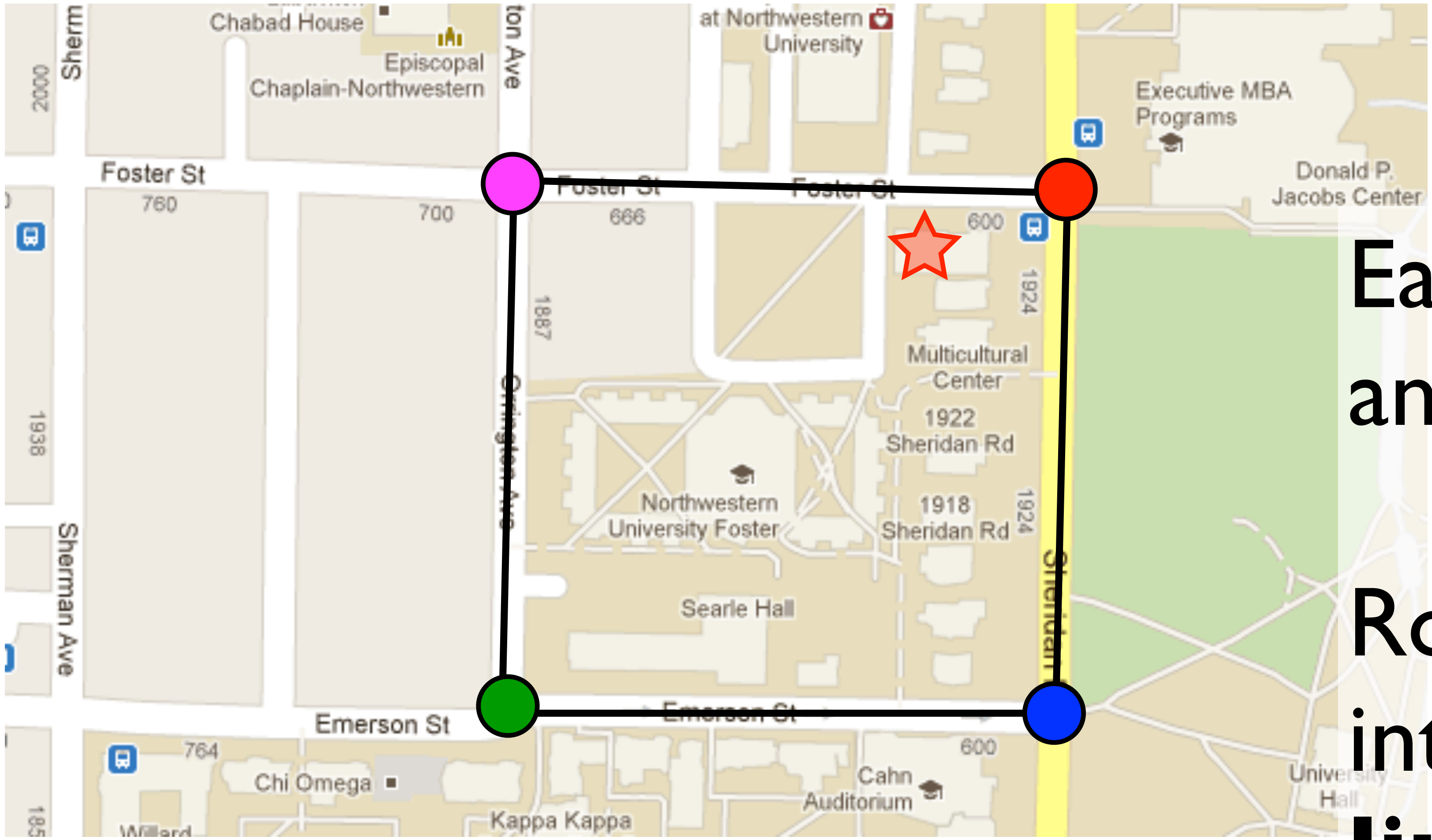
energy.gov

Road networks

Build network?

google maps

Road networks



Build network?

Each **node** represents an intersection

Roads connecting intersections form **links**

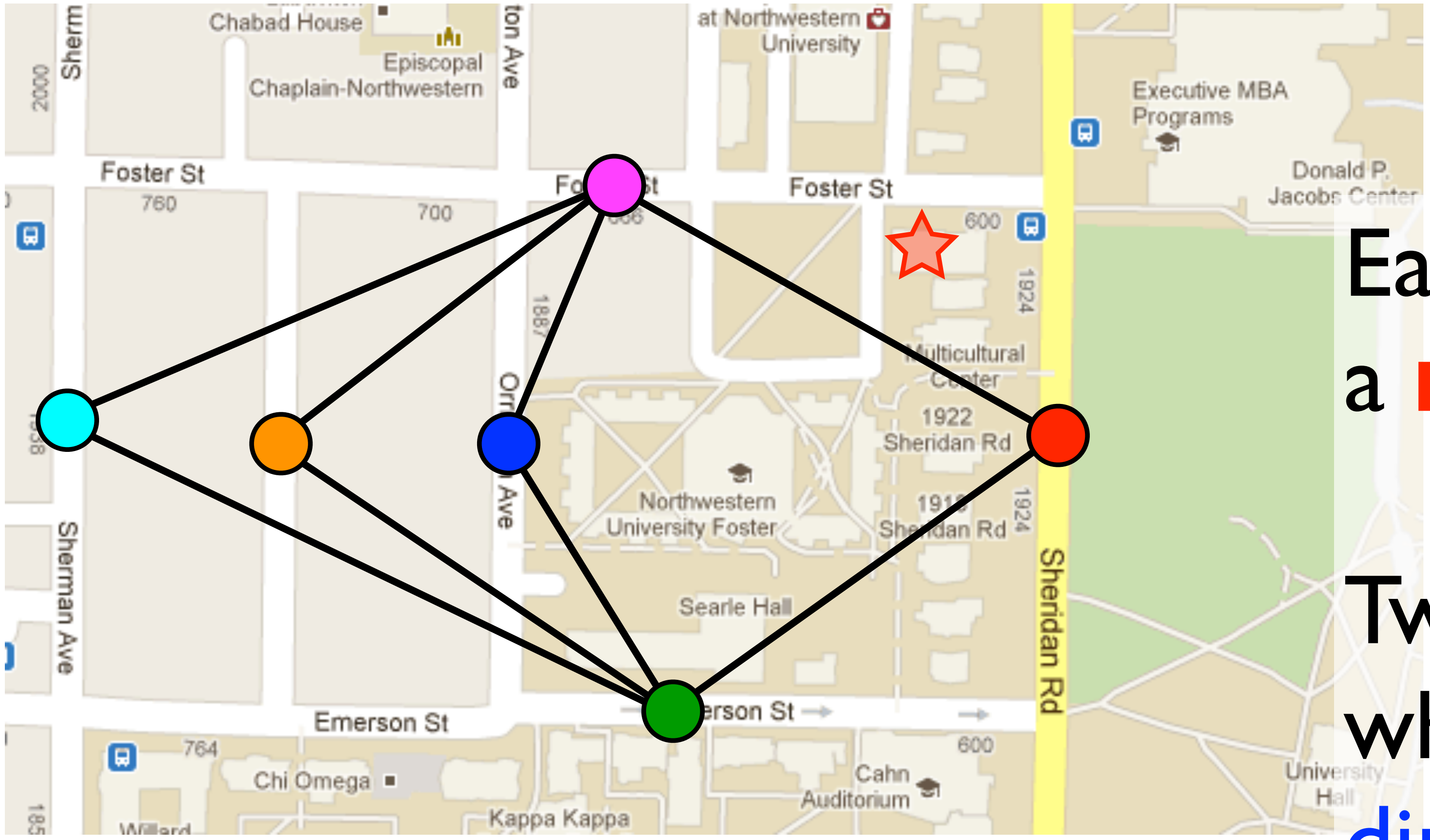
google maps

Road networks

Alternative

google maps

Road networks



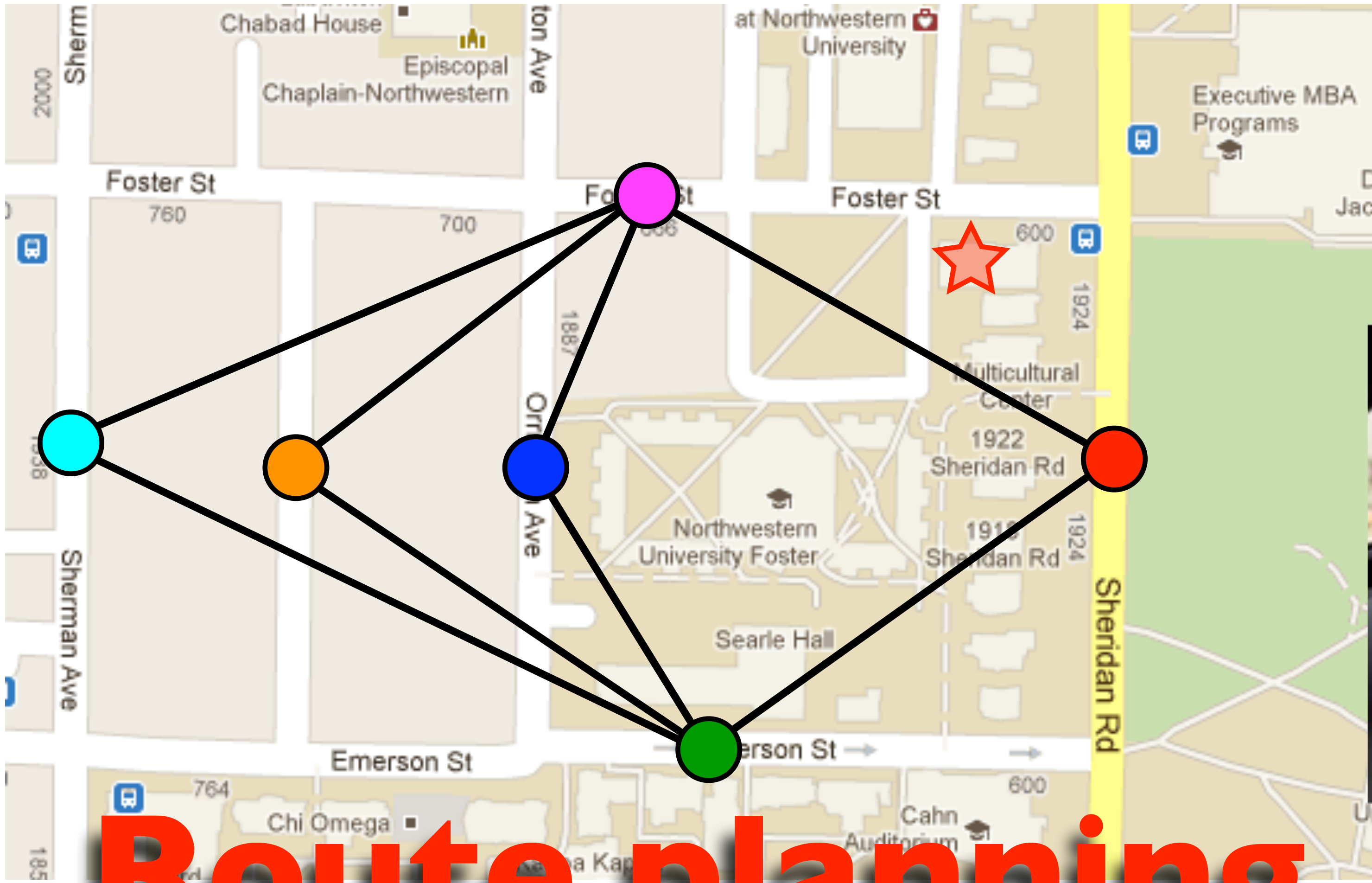
Alternative

Each node represents a **road**

Two roads are linked when you can drive directly from one to the other

google maps

Road networks



Route planning

Alternative

Each node represents

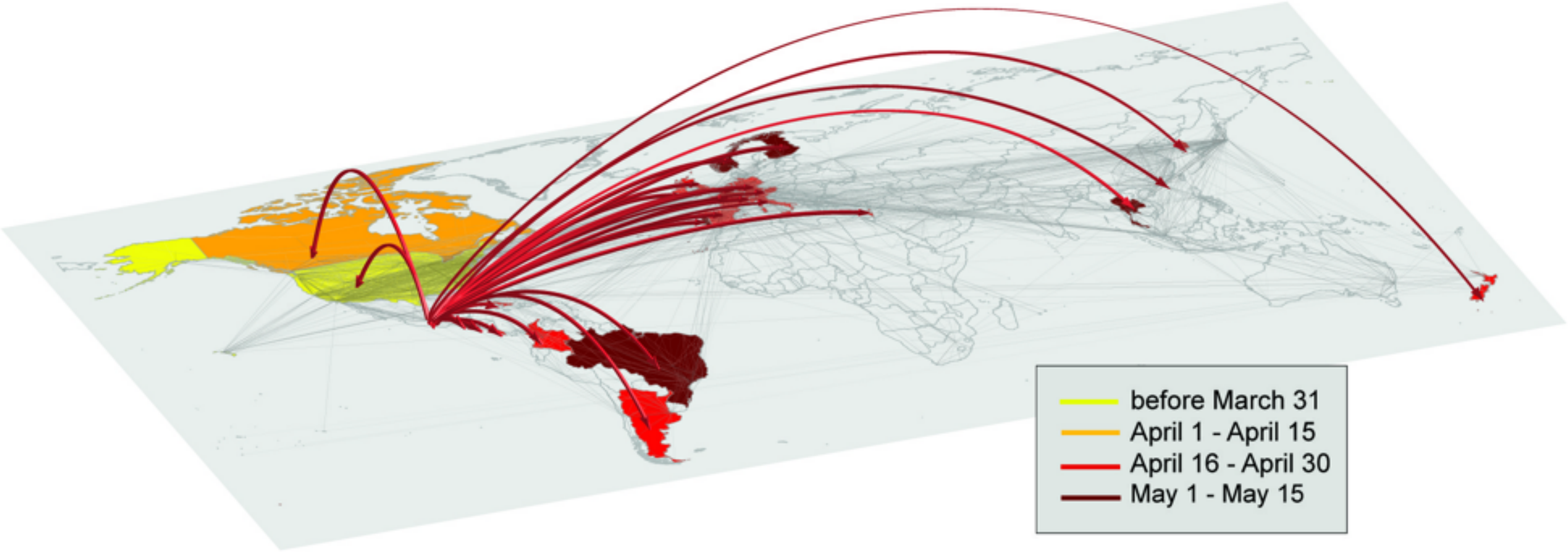
inked

rive

directly from one to the other

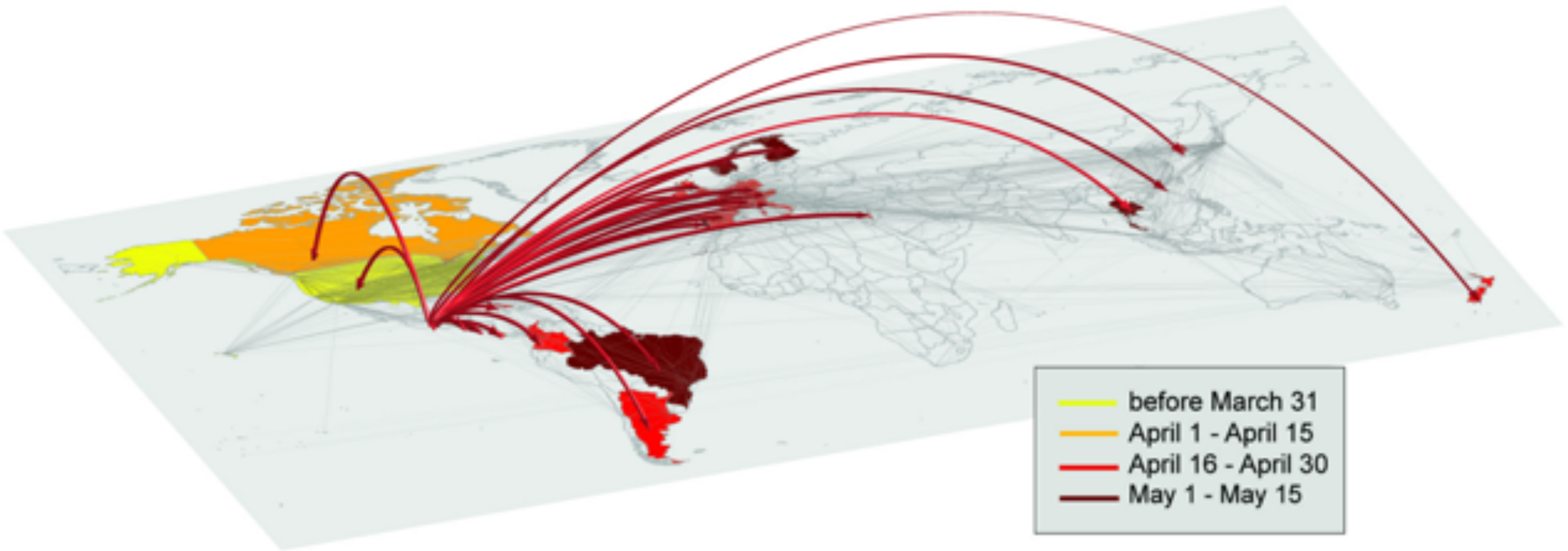
google maps

Air travel

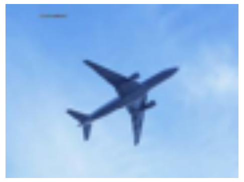


gleamviz.org

Air travel



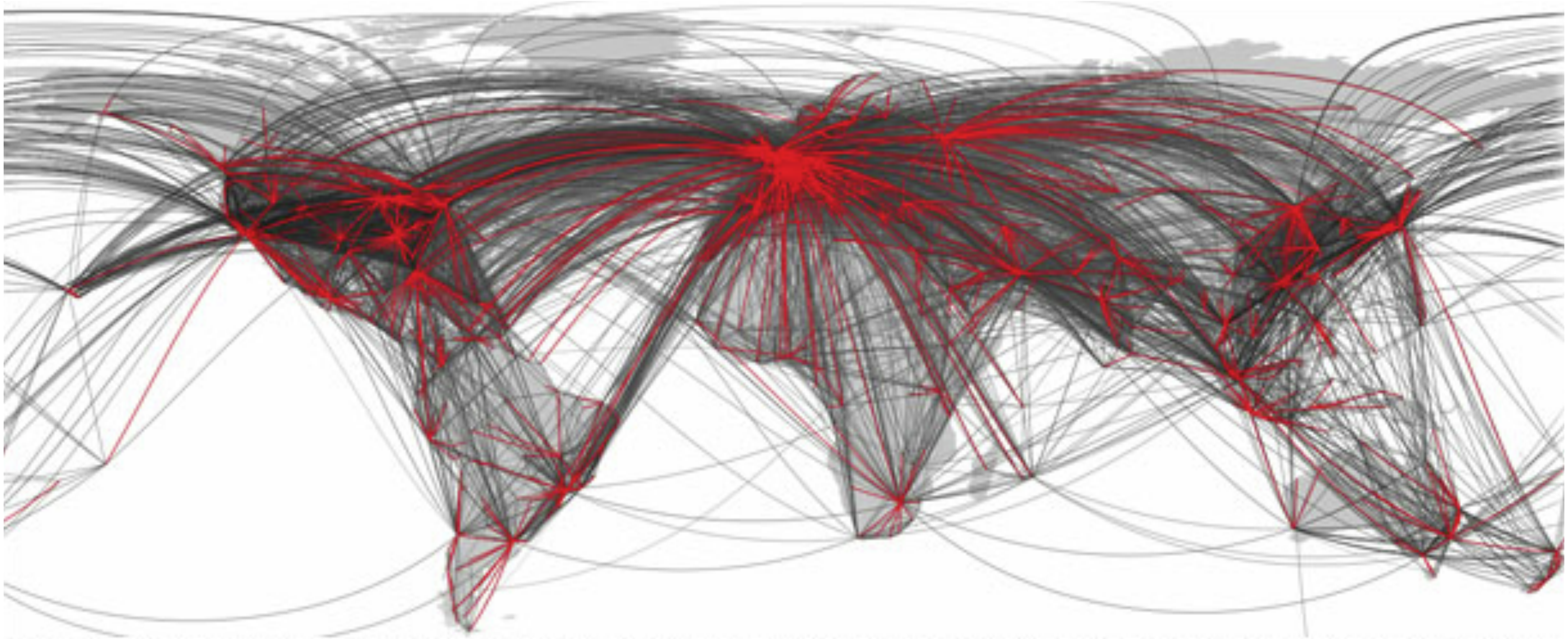
nodes = airports



links = direct flights
between airports

gleamviz.org

Air travel

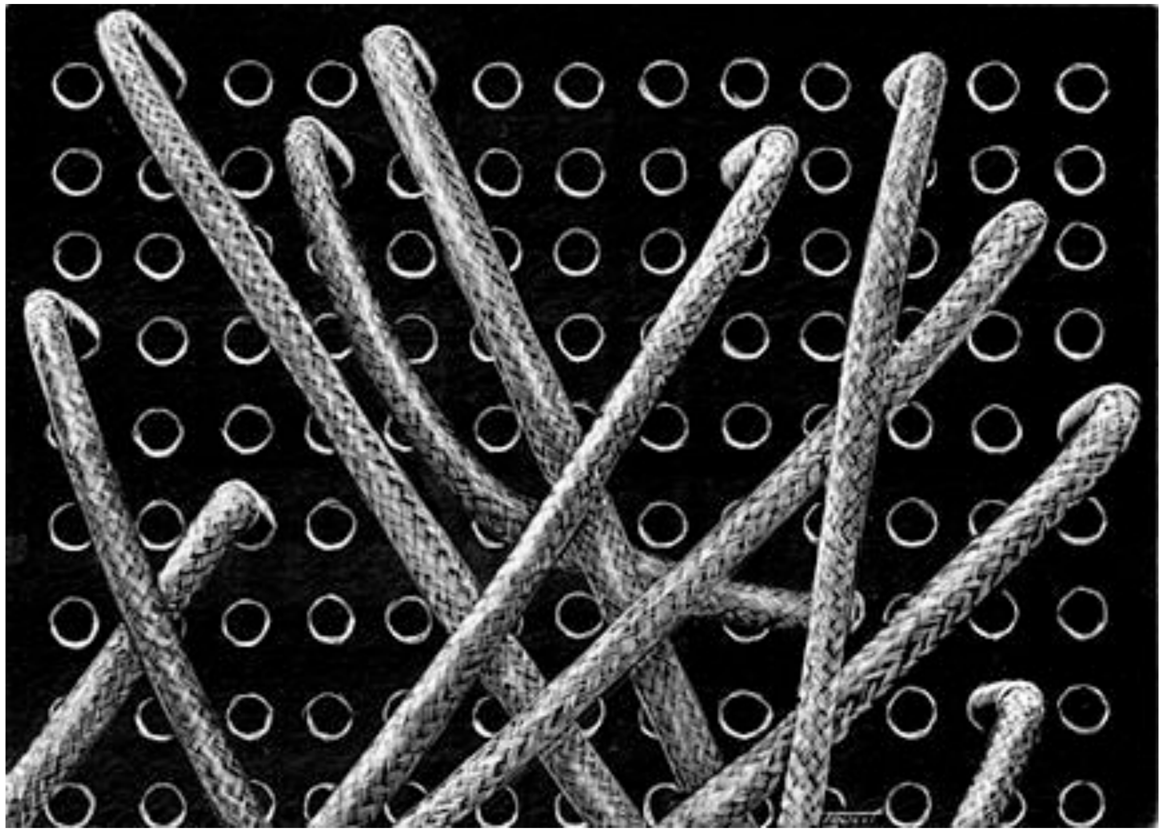


Grady, *et al*. Nat Commun, 2012

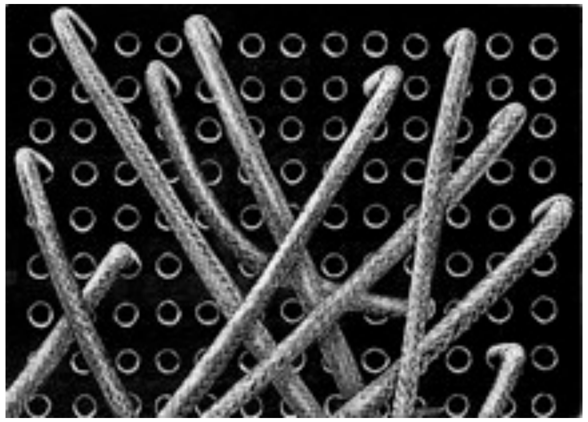
Air travel

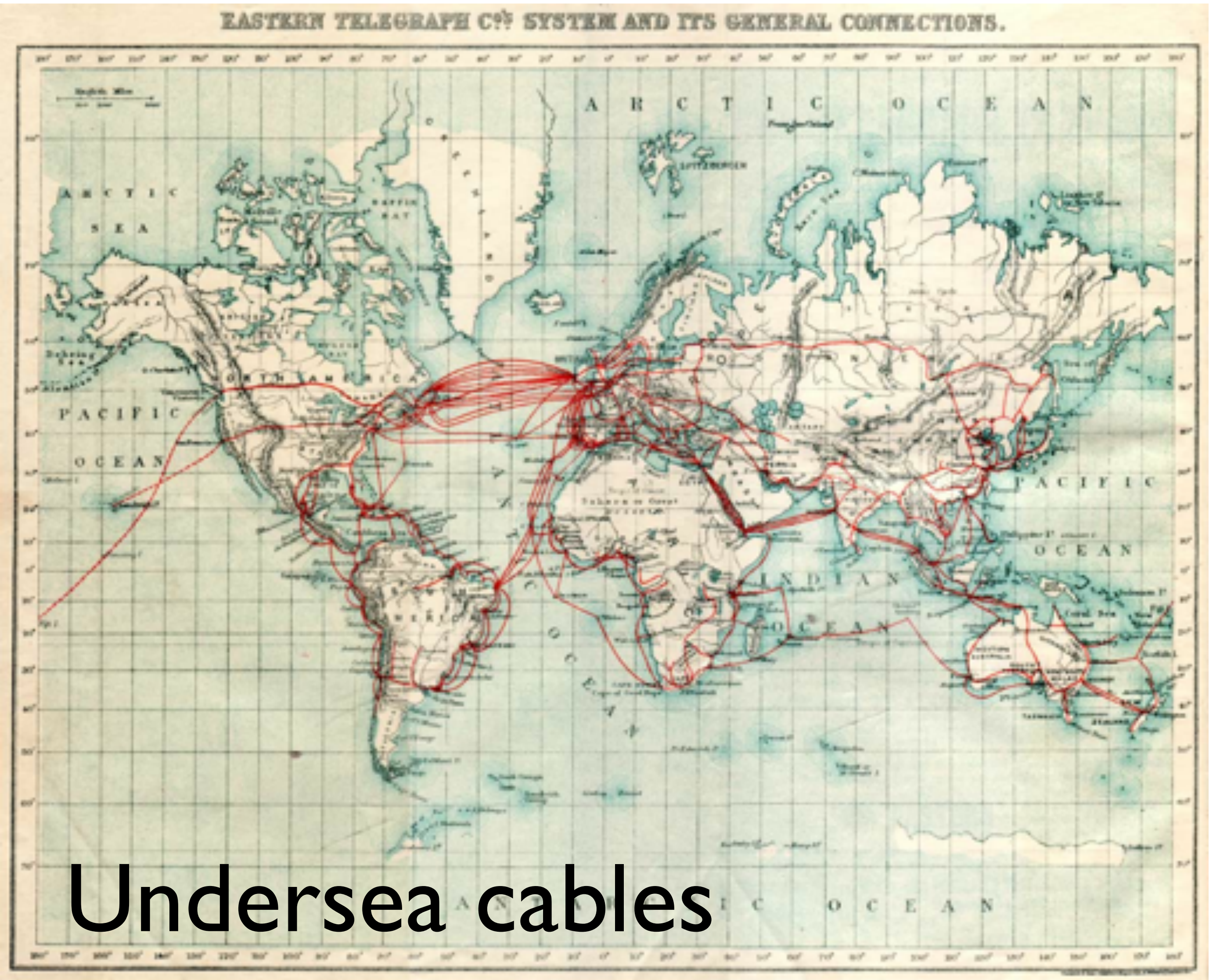
Disease spreading

Telecommunications



Telecommunications

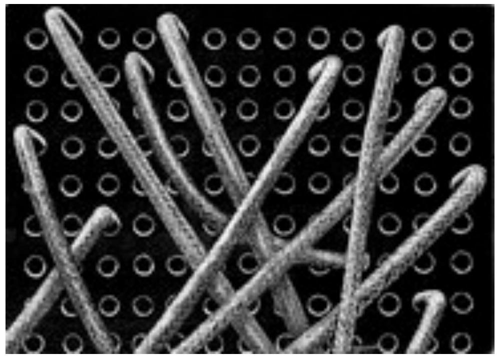




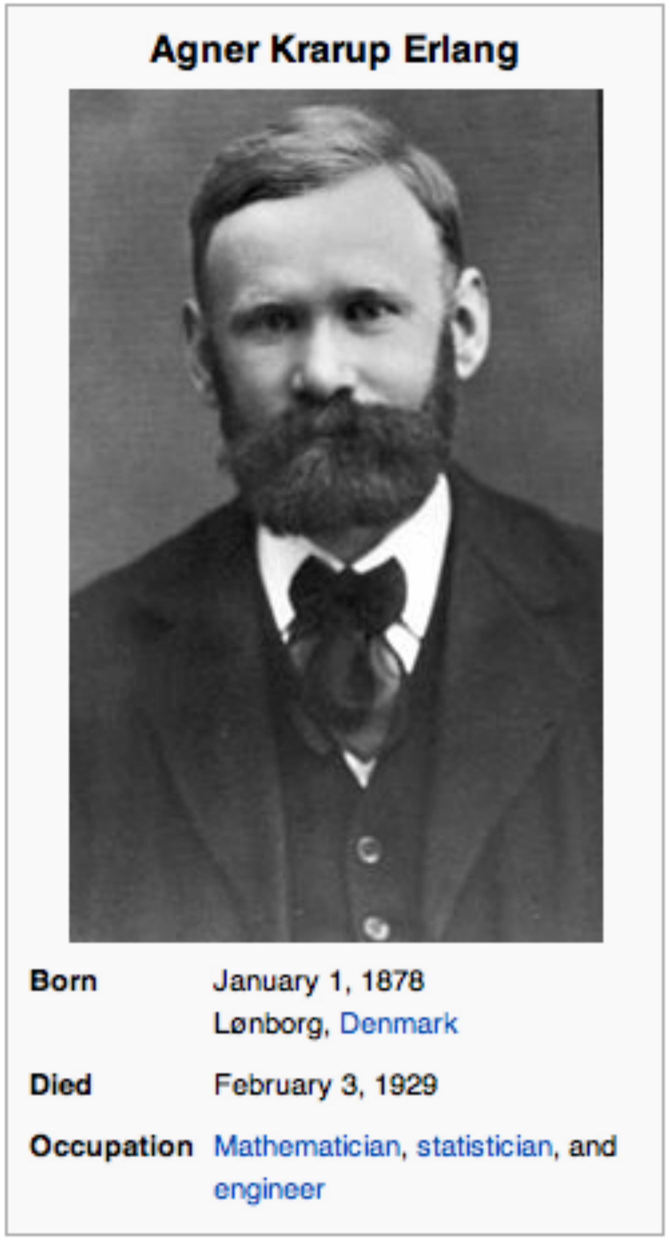
Undersea cables

1901!

Telecommunications



How many lines do we need for our phone calls?



Agner Krarup Erlang

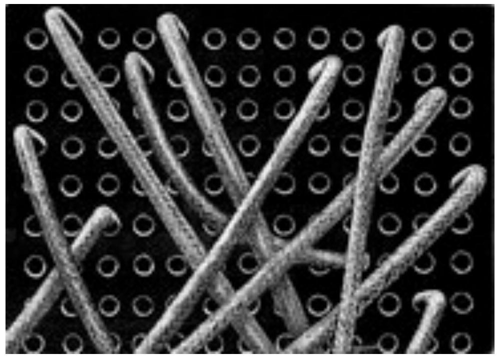
Born January 1, 1878
Lønborg, Denmark

Died February 3, 1929

Occupation Mathematician, statistician, and engineer

wikipedia

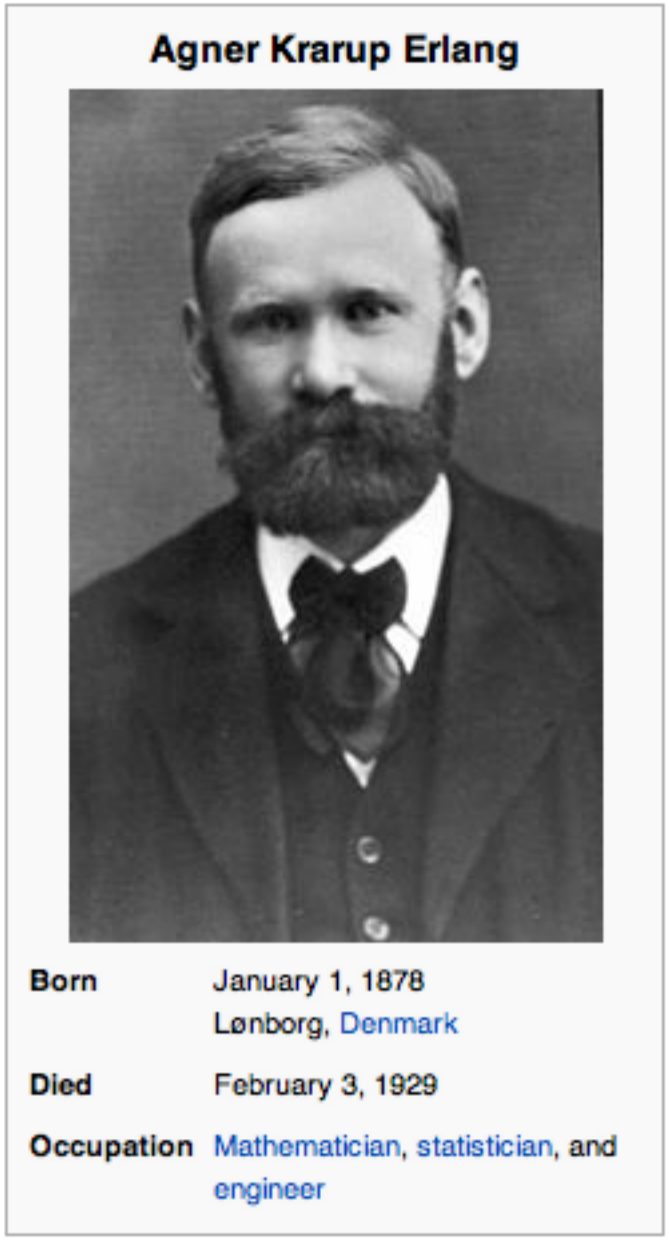
Telecommunications



How many lines do we need for our phone calls?

↓

Queueing Theory



Agner Krarup Erlang

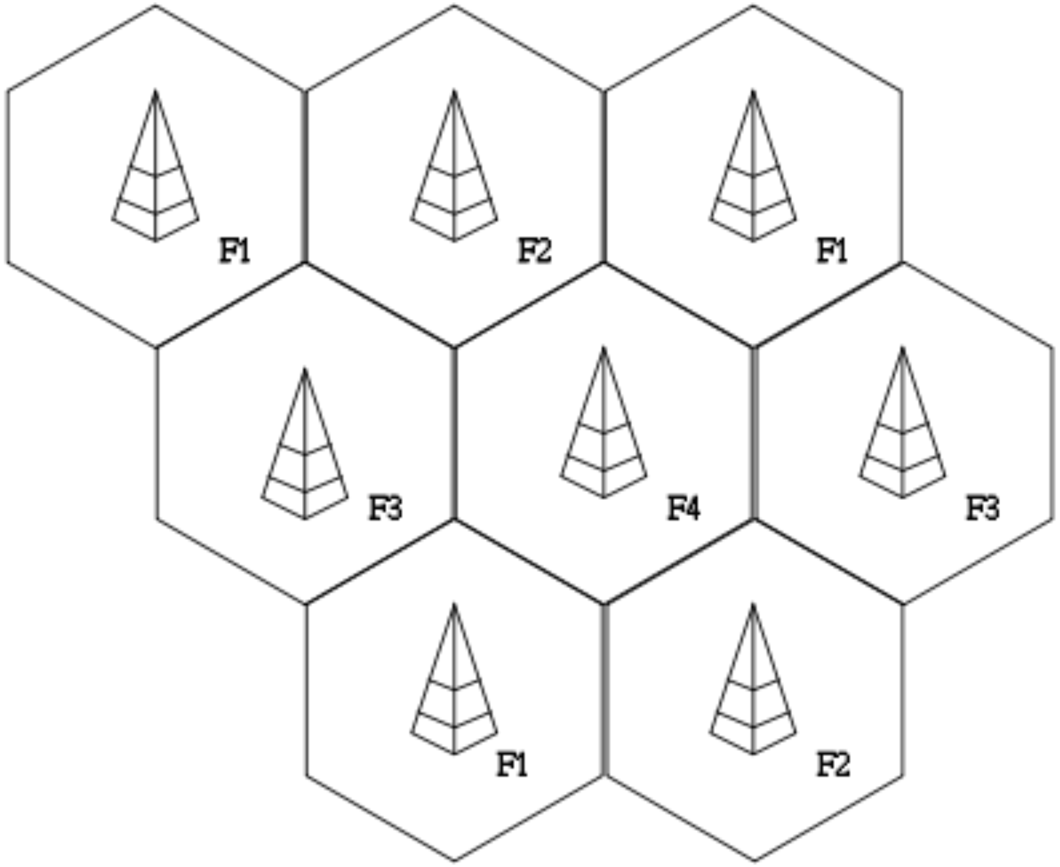
Born	January 1, 1878 Lønborg, Denmark
Died	February 3, 1929
Occupation	Mathematician, statistician, and engineer

wikipedia

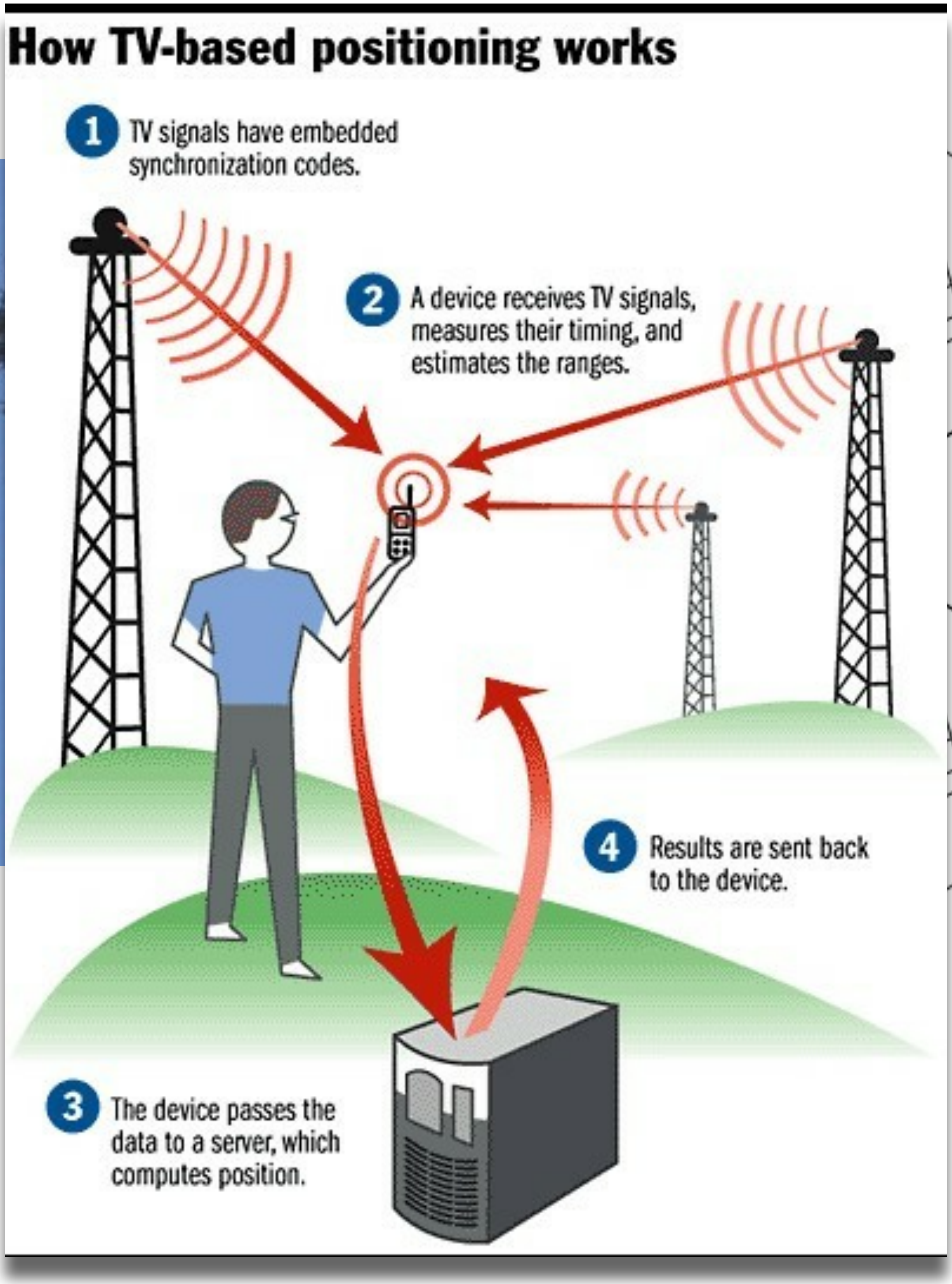

Agner Krarup Erlang


FÜR DIE WOHLFAHRTSPFLEGE
Fernsprechvermittlungsstelle um 1890
80
+35
DEUTSCHE BUNDESPOST
1990

Mobile phones



Mobile phones



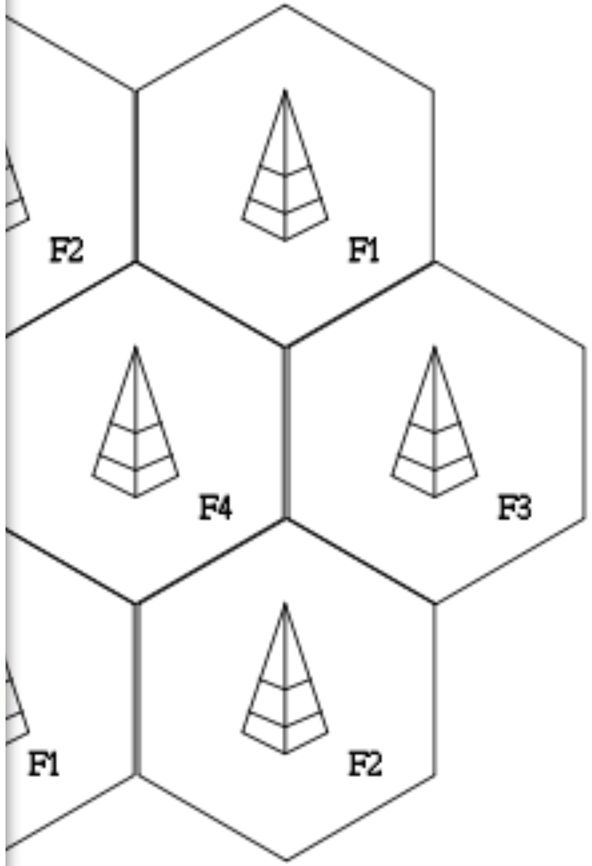
How TV-based positioning works

1 TV signals have embedded synchronization codes.

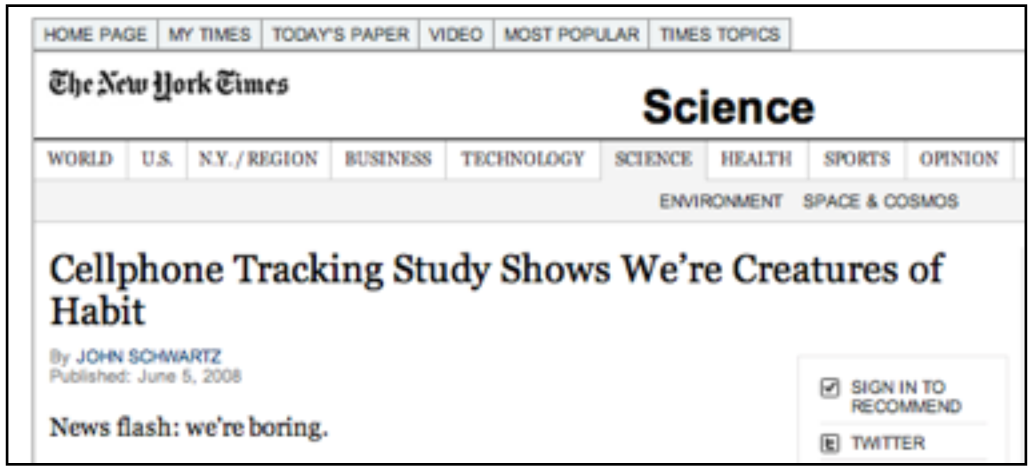
2 A device receives TV signals, measures their timing, and estimates the ranges.

3 The device passes the data to a server, which computes position.

4 Results are sent back to the device.



Mobile phone data



HOME PAGE | MY TIMES | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times

Science

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

ENVIRONMENT SPACE & COSMOS

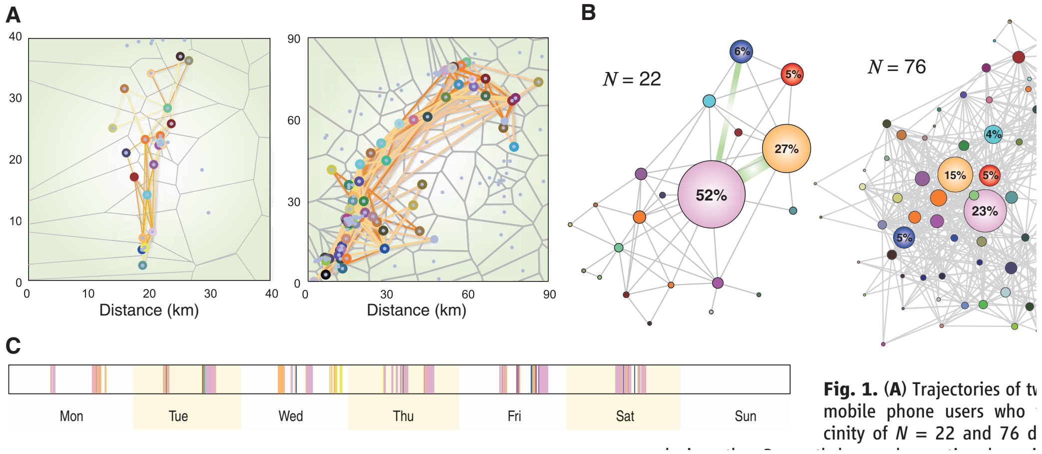
Cellphone Tracking Study Shows We're Creatures of Habit

By JOHN SCHWARTZ
Published: June 5, 2008

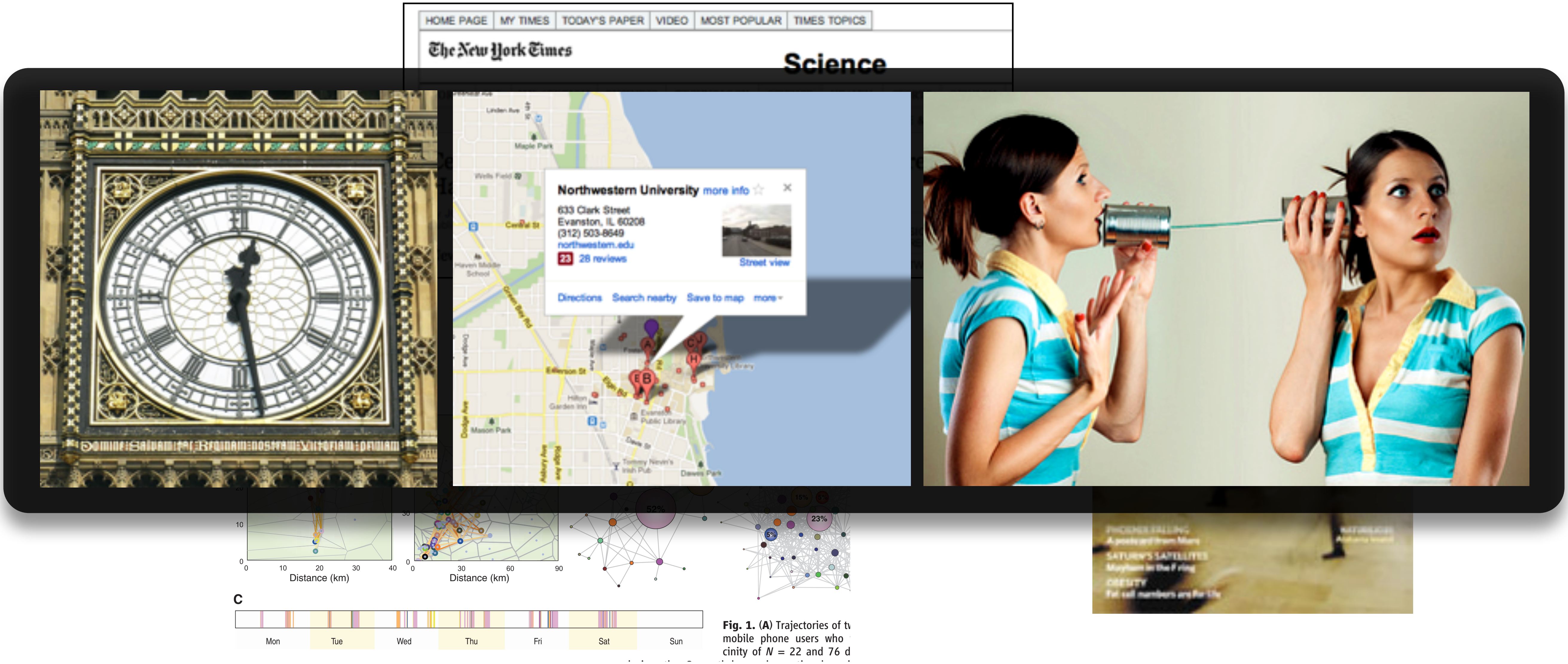
News flash: we're boring.

☑ SIGN IN TO RECOMMEND

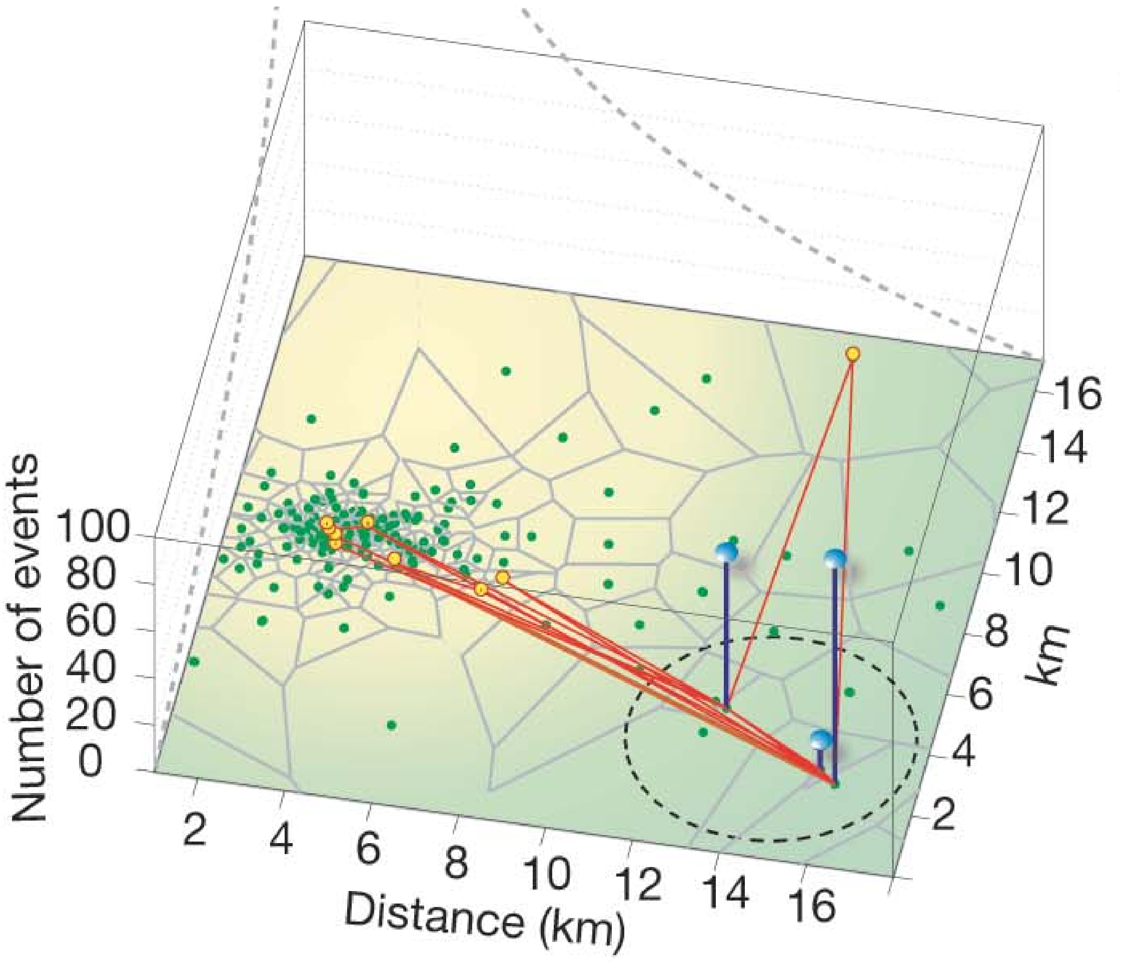
TWITTER

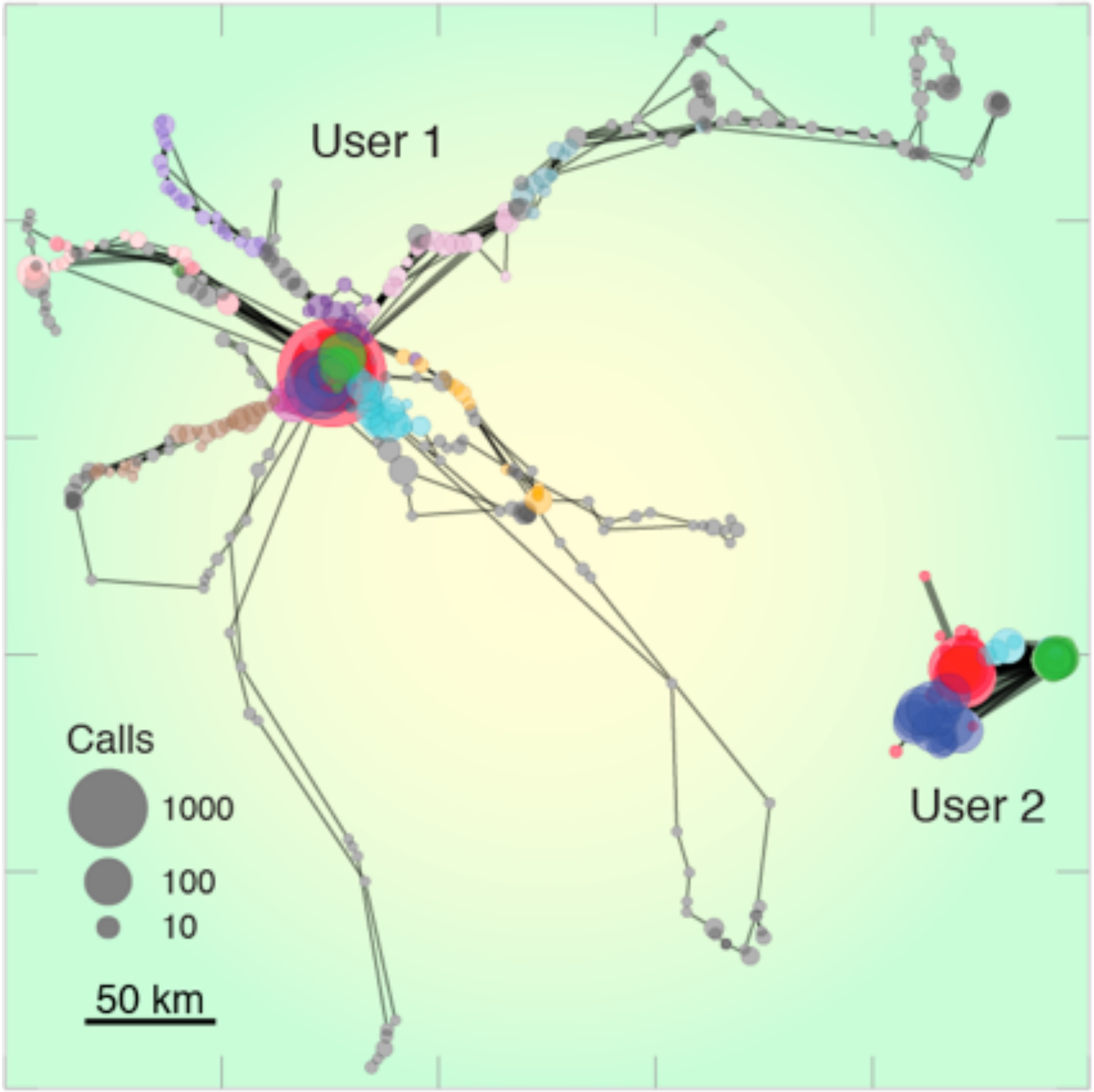


Mobile phone data



Human mobility

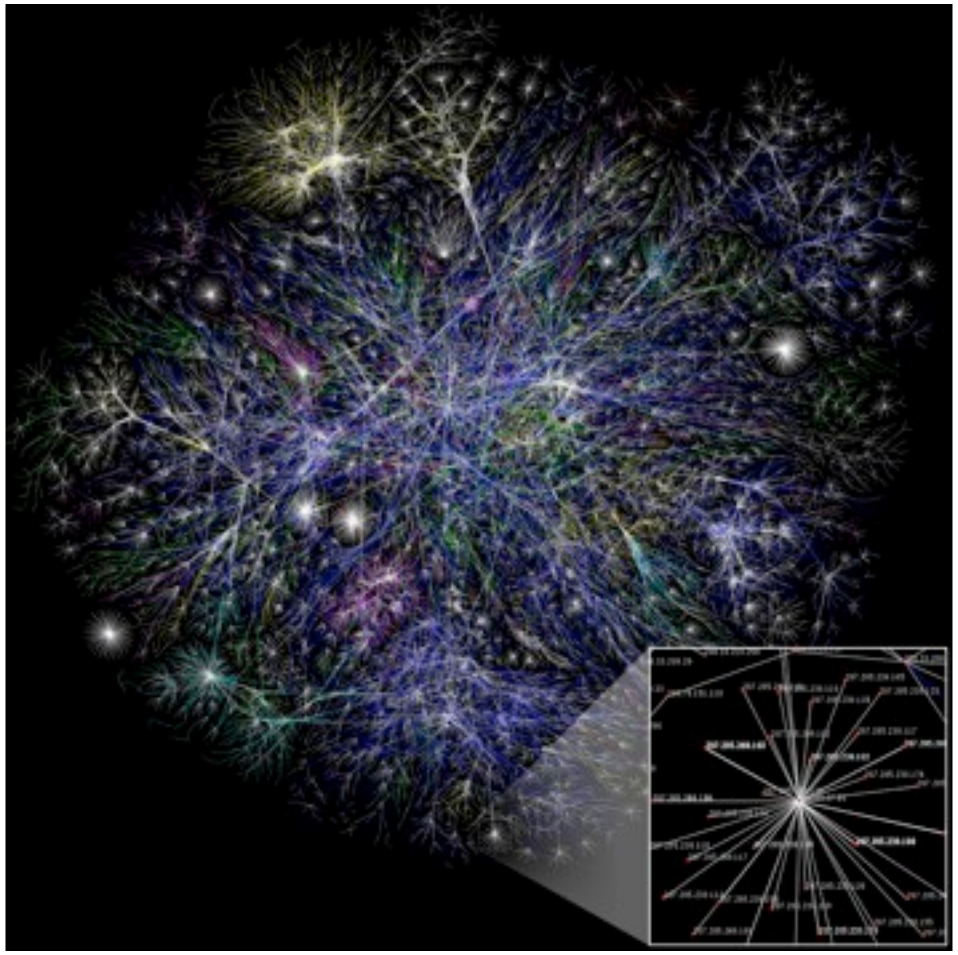


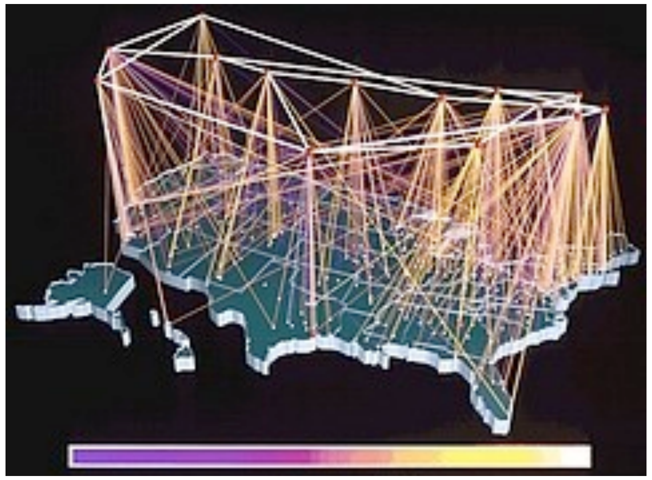


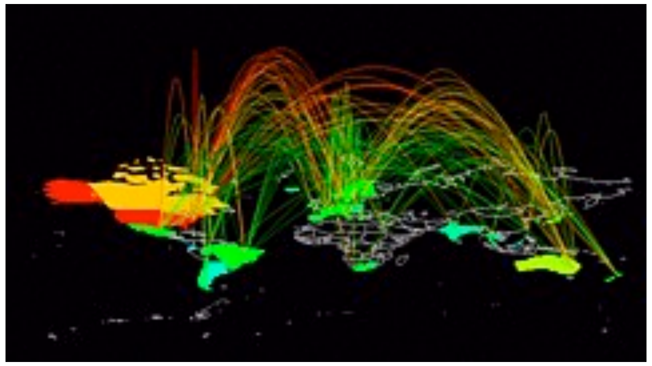
M. C. González, et al., 2008

Bagrow and Lin, 2012

Internet

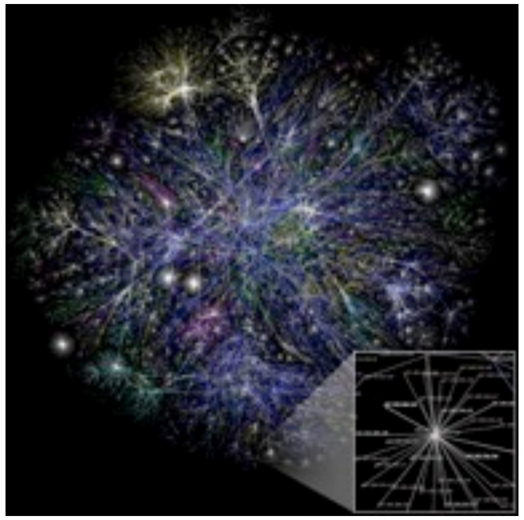


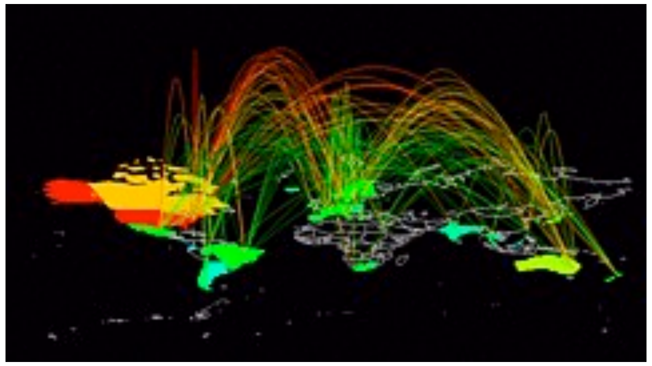


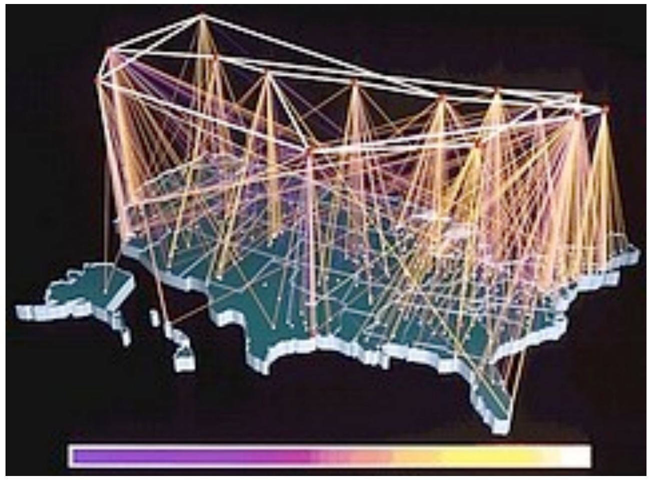


infovis.info > Atlas of cyberspace

Internet







Nodes: Computers, routers, subnetworks, etc.

Links: Connect nodes that **share data**

Web

WWW sits **on top of** the Internet

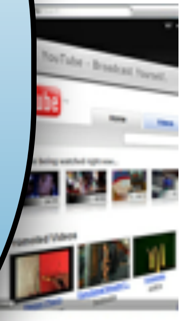
First web server

Nodes	Web pages
Links	Hyperlinks

Web

the

Novelty of the **web** led to the **explosion** of networks research in the late 90s



First web se

Links Hyperlinks

Email

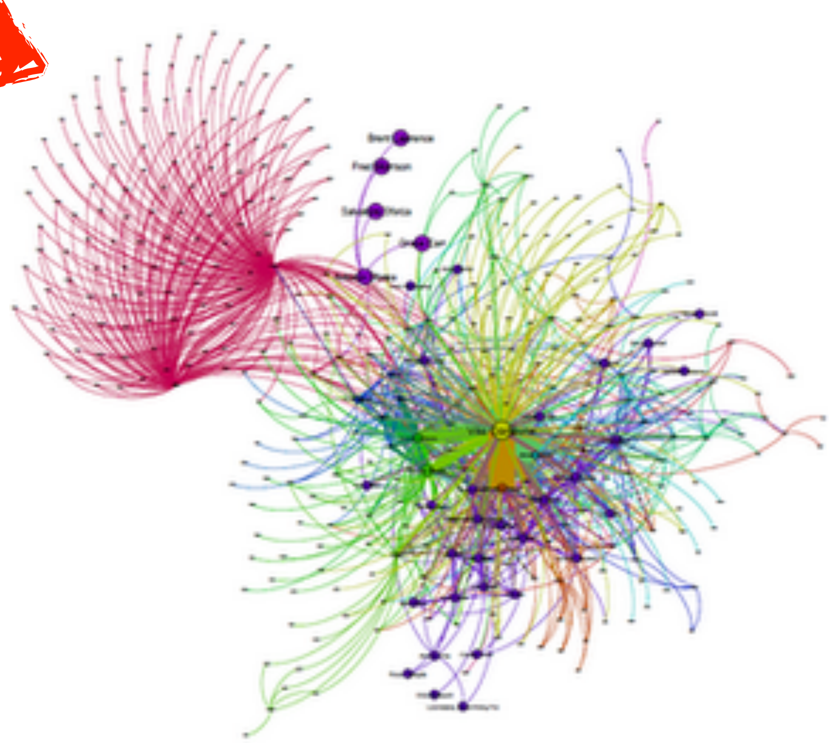
```
MIME-Version: 1.0
Date: Fri, 30 Mar 2012 12:34:07 -0500
Subject: Re: Eqs
From: Jim Bagrow <bagrowjp@gmail.com>
To: Dirk Brockmann <brockmann@northwestern.edu>
...
```

Email

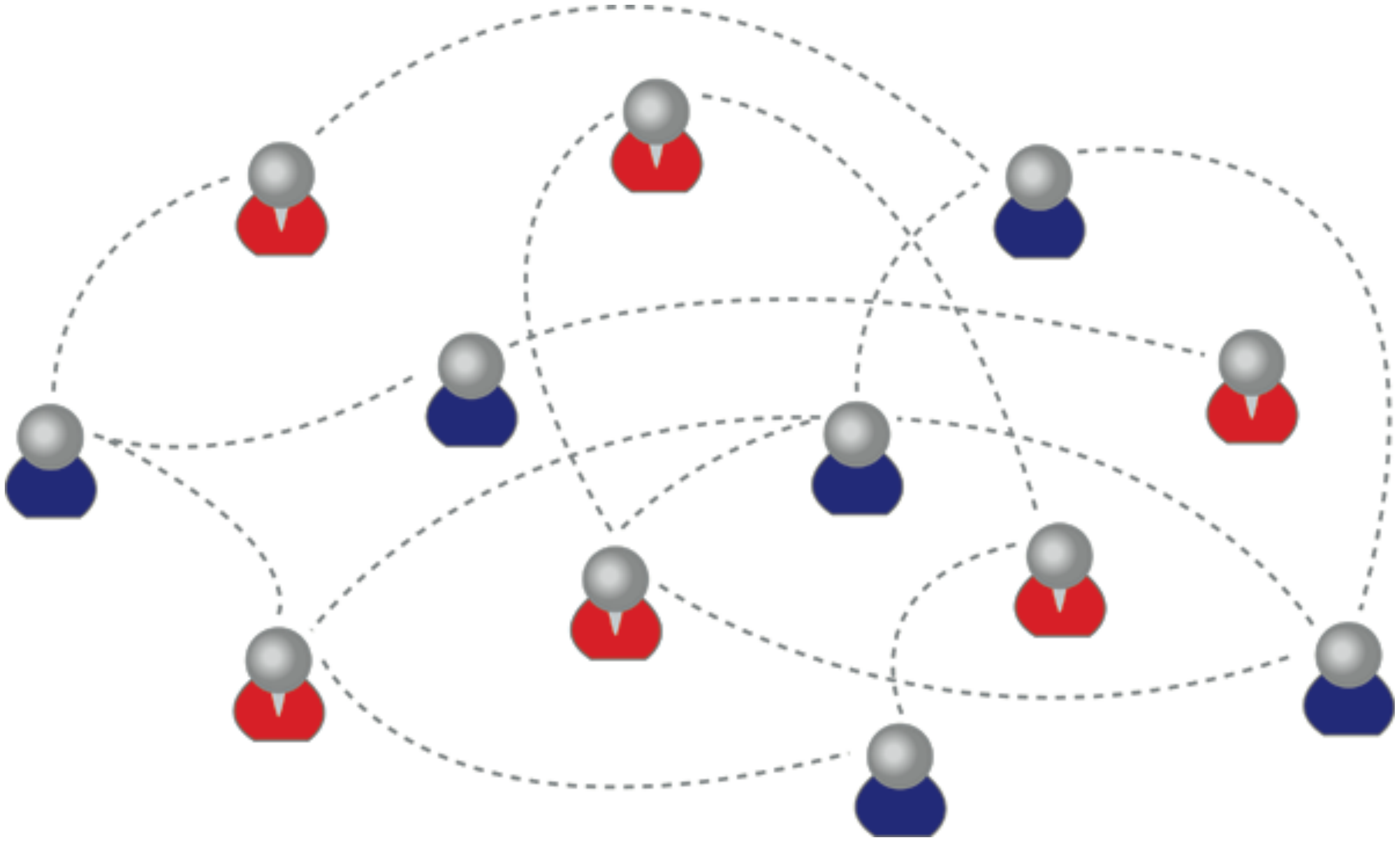
```
MIME-Version: 1.0
Date: Fri, 30 Mar 2012 12:34:07 -0500
Subject: Re: Eqs
From: Jim Bagrow <bagrowjp@gmail.com>
To: Dirk Brockmann <brockmann@northwestern.edu>
...
```

Email

```
MIME-Version: 1.0
Date: Fri, 30 Mar 2012 12:34:07 -0500
Subject: Re: Eqs
From: Jim Bagrow <bagrowjp@gmail.com>
To: Dirk Brockmann <brockmann@northwestern.edu>
...
```



Social networks



Nodes

People

Links

Relationships

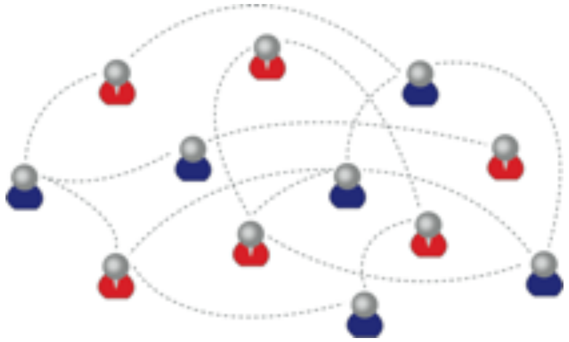
between people

Huge area

Information spreading

Disease spreading

Sociology

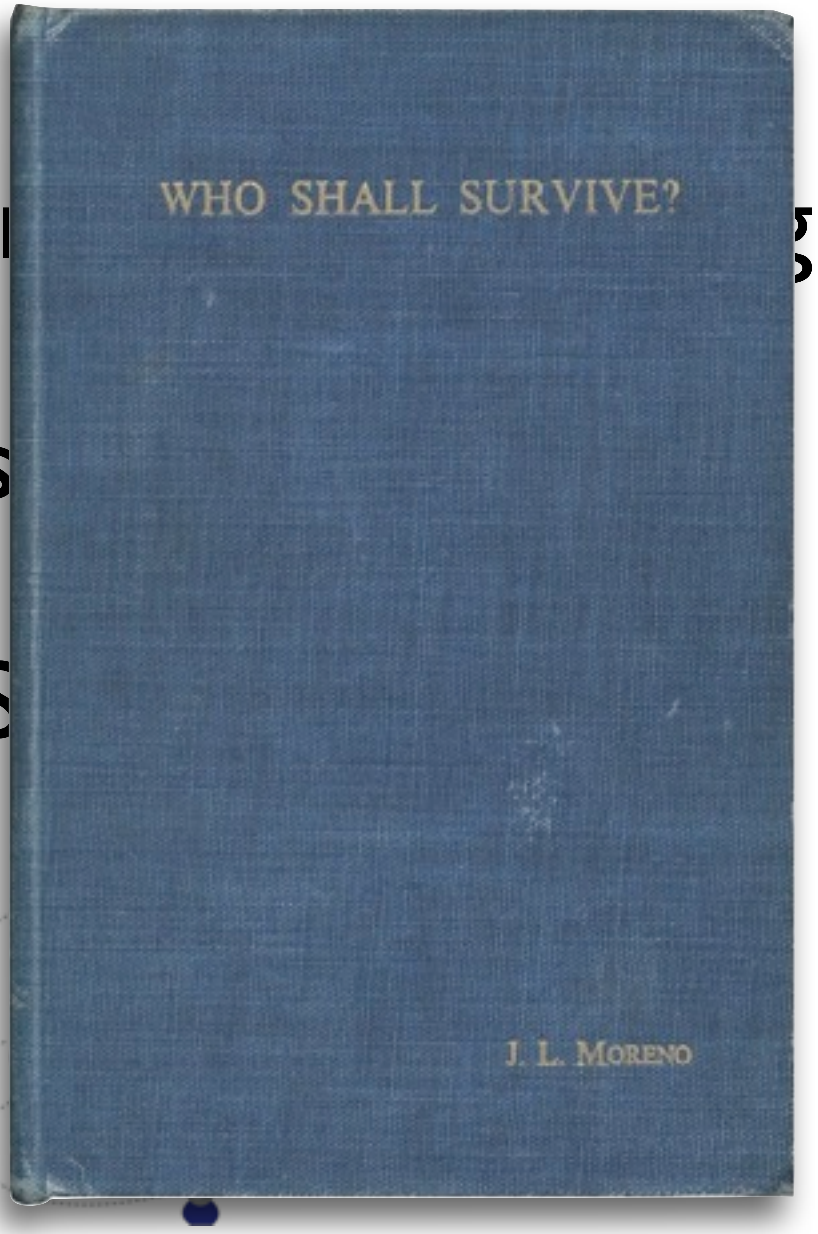


Huge area

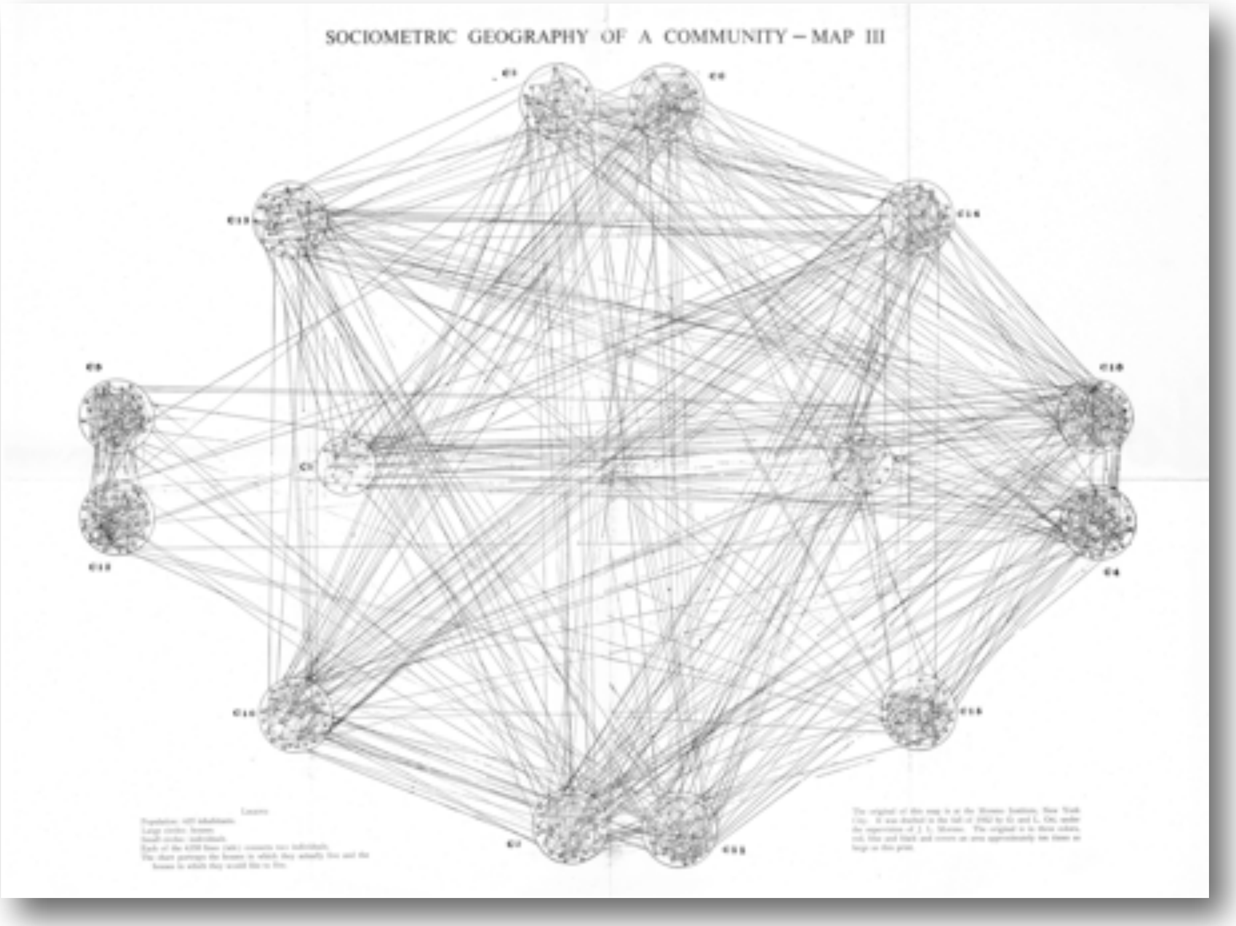
Info g

Dis

S



1934, 1953 (2nd ed)



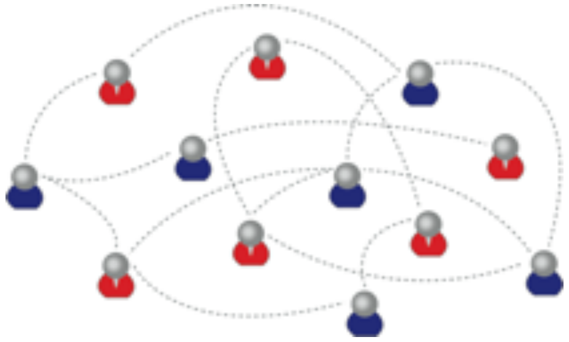
1932

Huge area

Information spreading

Disease spreading

Sociology



Applications

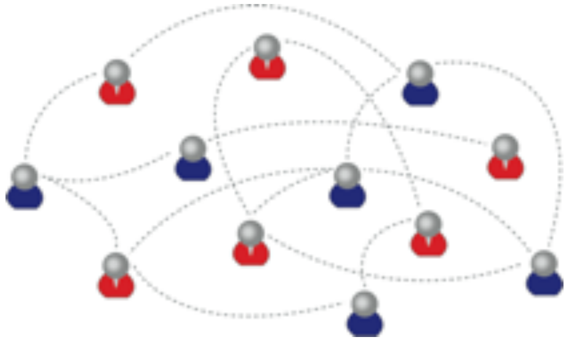
Marketing
Vaccine distribution
Social media
Emergency response
....

Huge area

Information spreading

Disease spreading

Sociology



Applications

Marketing
Vaccine distribution
Social media
Emergency response
.

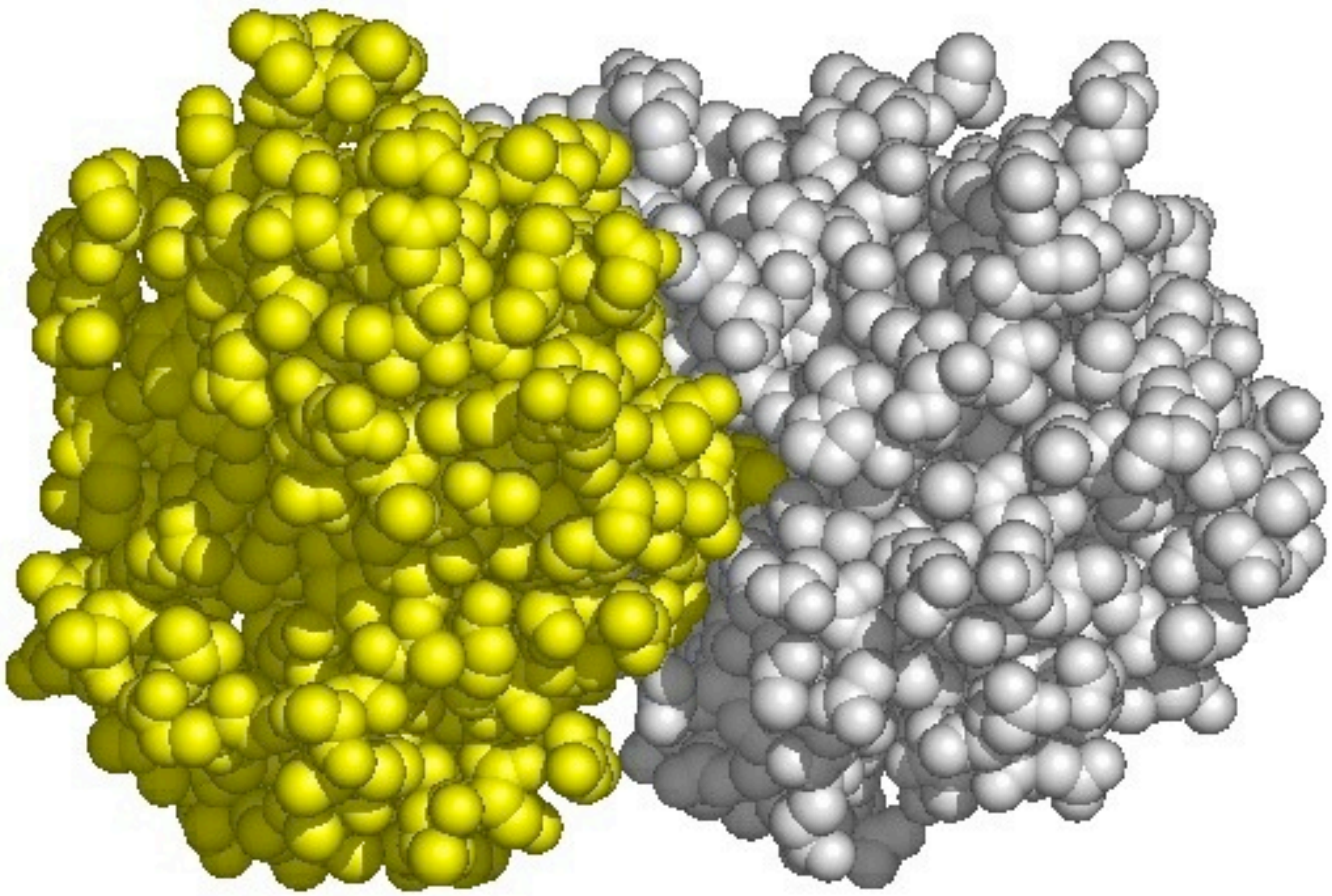
See next NetSci school session

Biological networks

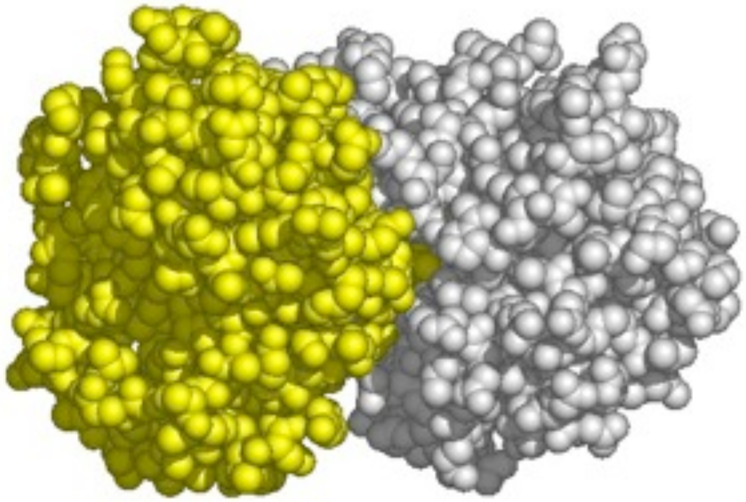
Another **HUGE** area

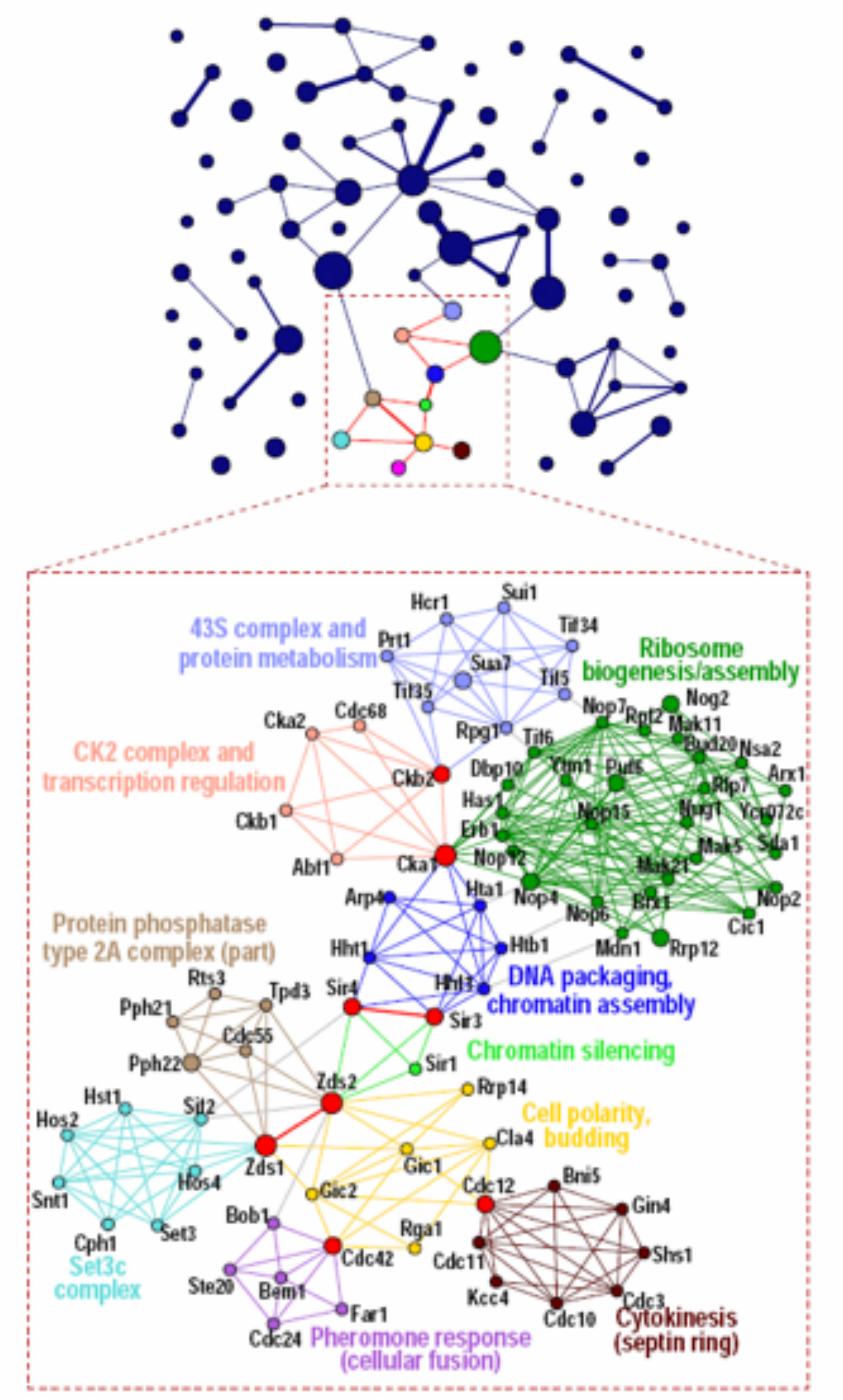
Systems biology

Protein-Protein
Interaction
networks

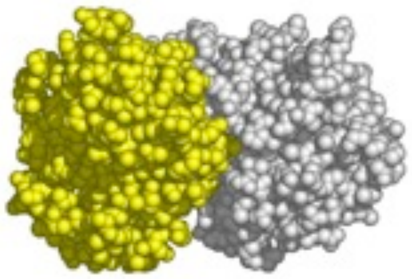


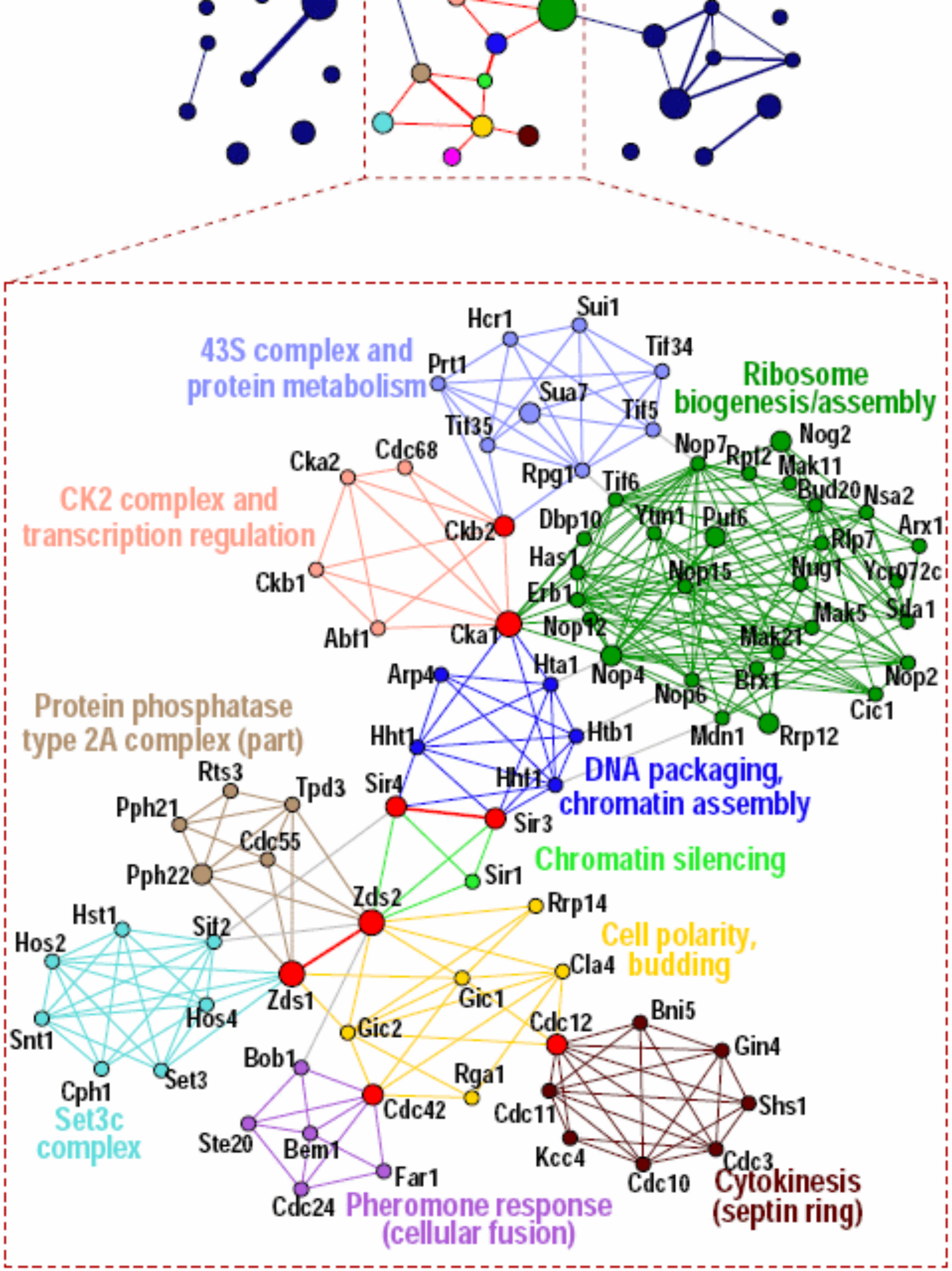
PPI networks





Palla, et al. Nature, 2005





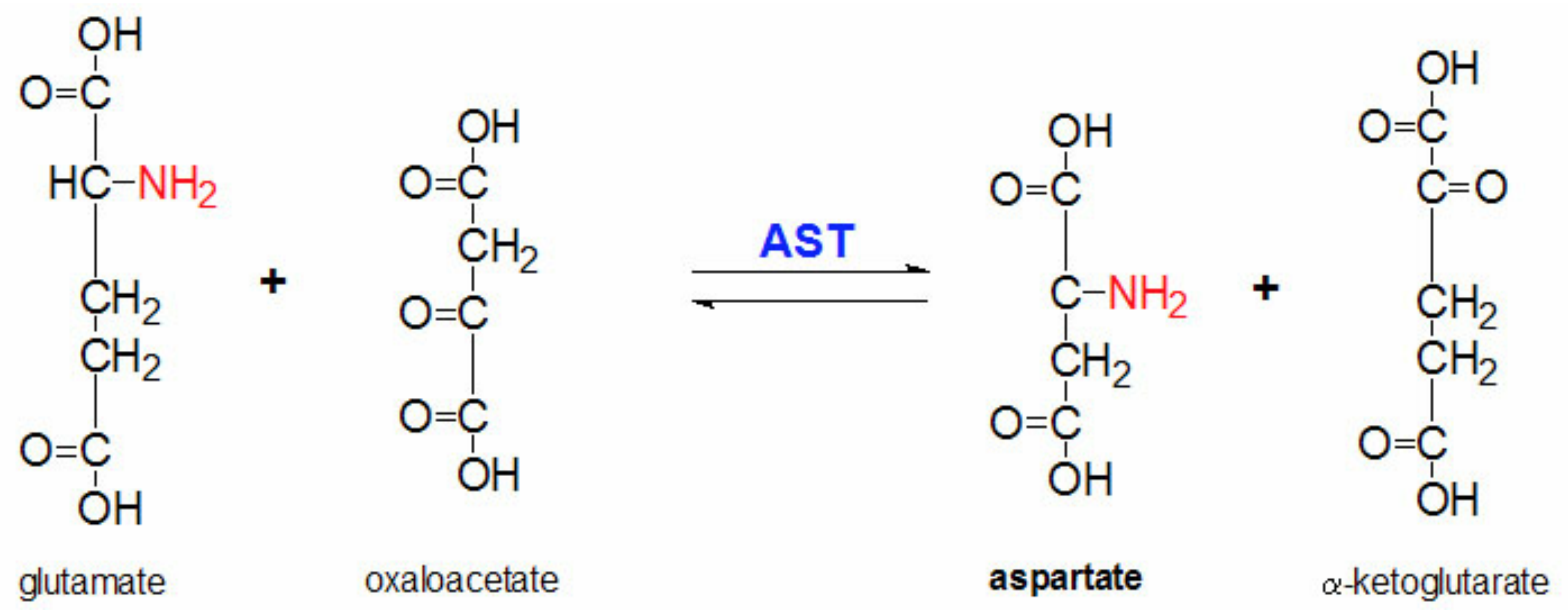
Palla, et al. Nature, 2005

Metabolic networks

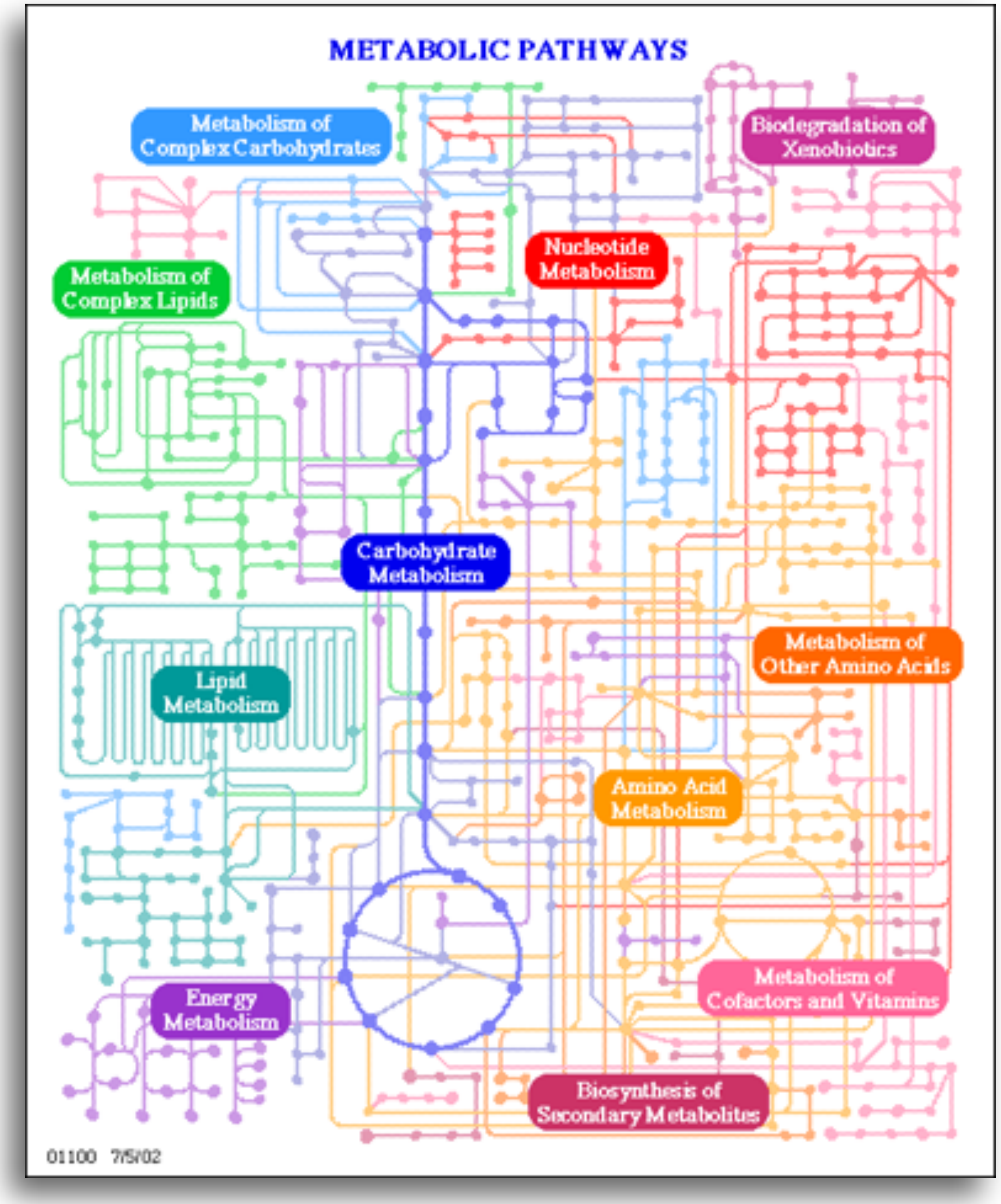
Metabolic networks

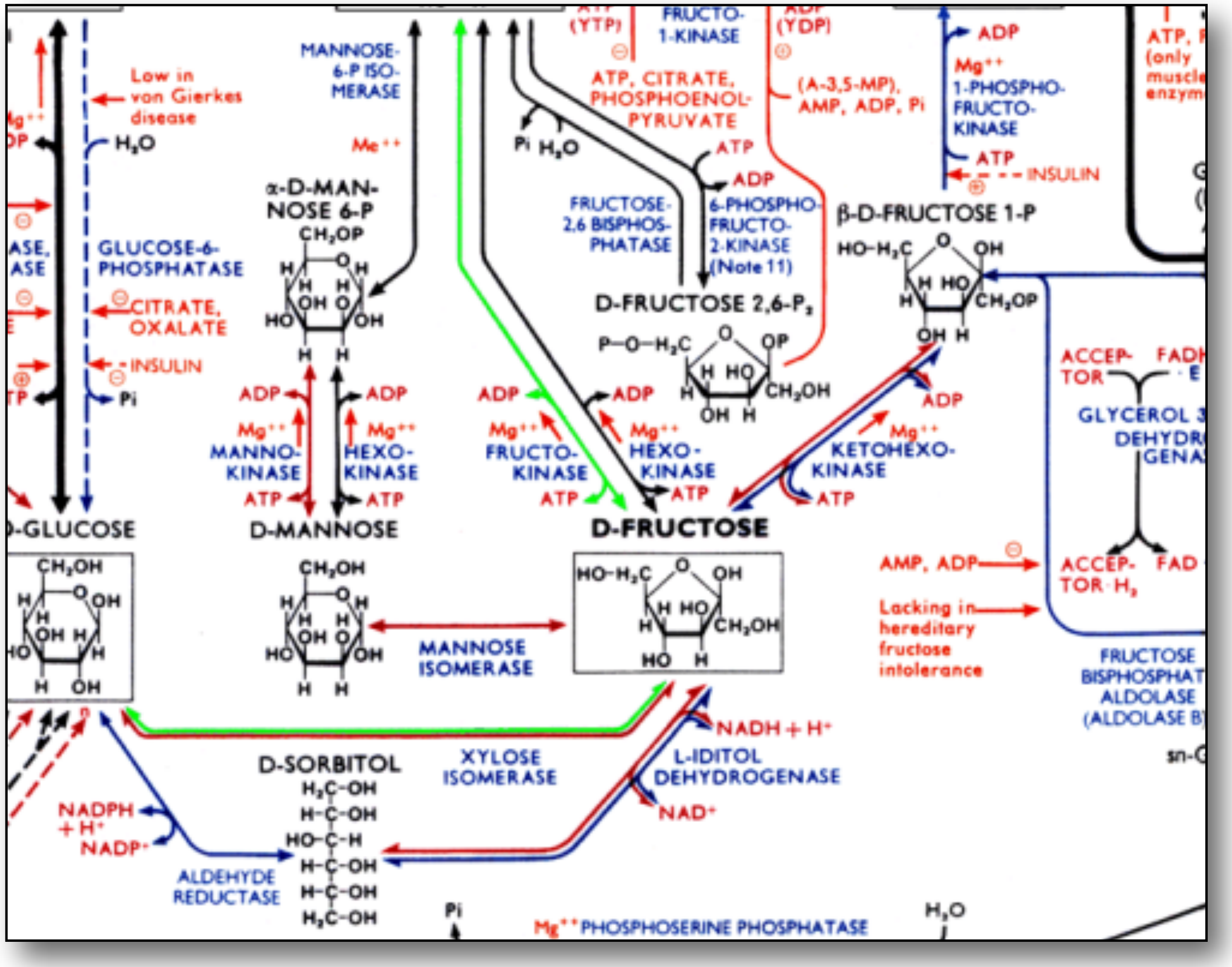
nodes: Metabolites (chemicals)

links: Reactions involving metabolites



Metabolic networks





Kyoto Encyclopedia of Genes and Genomes

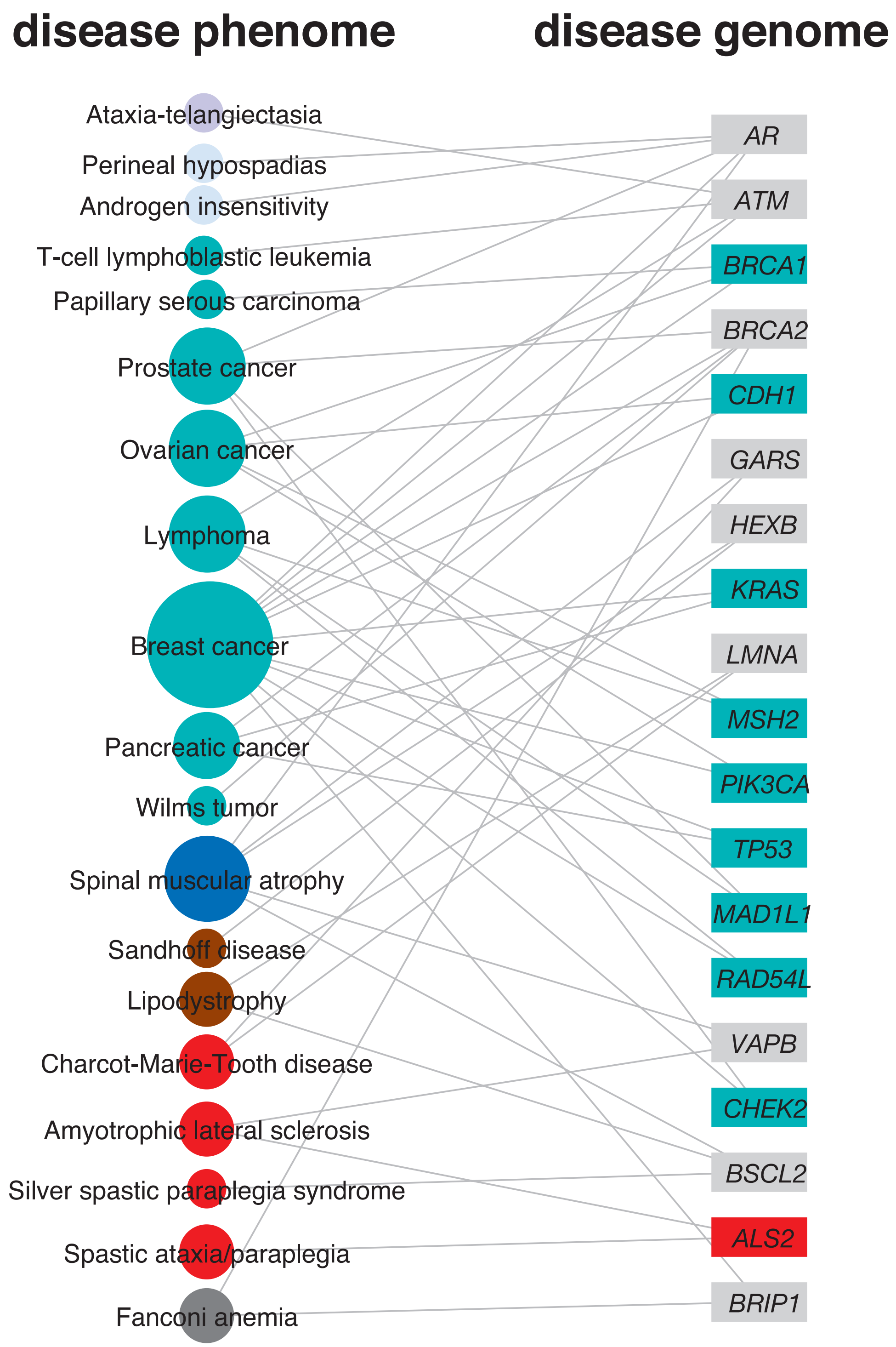
http://www.expasy.org/tools/pathways/

“Diseaseome”

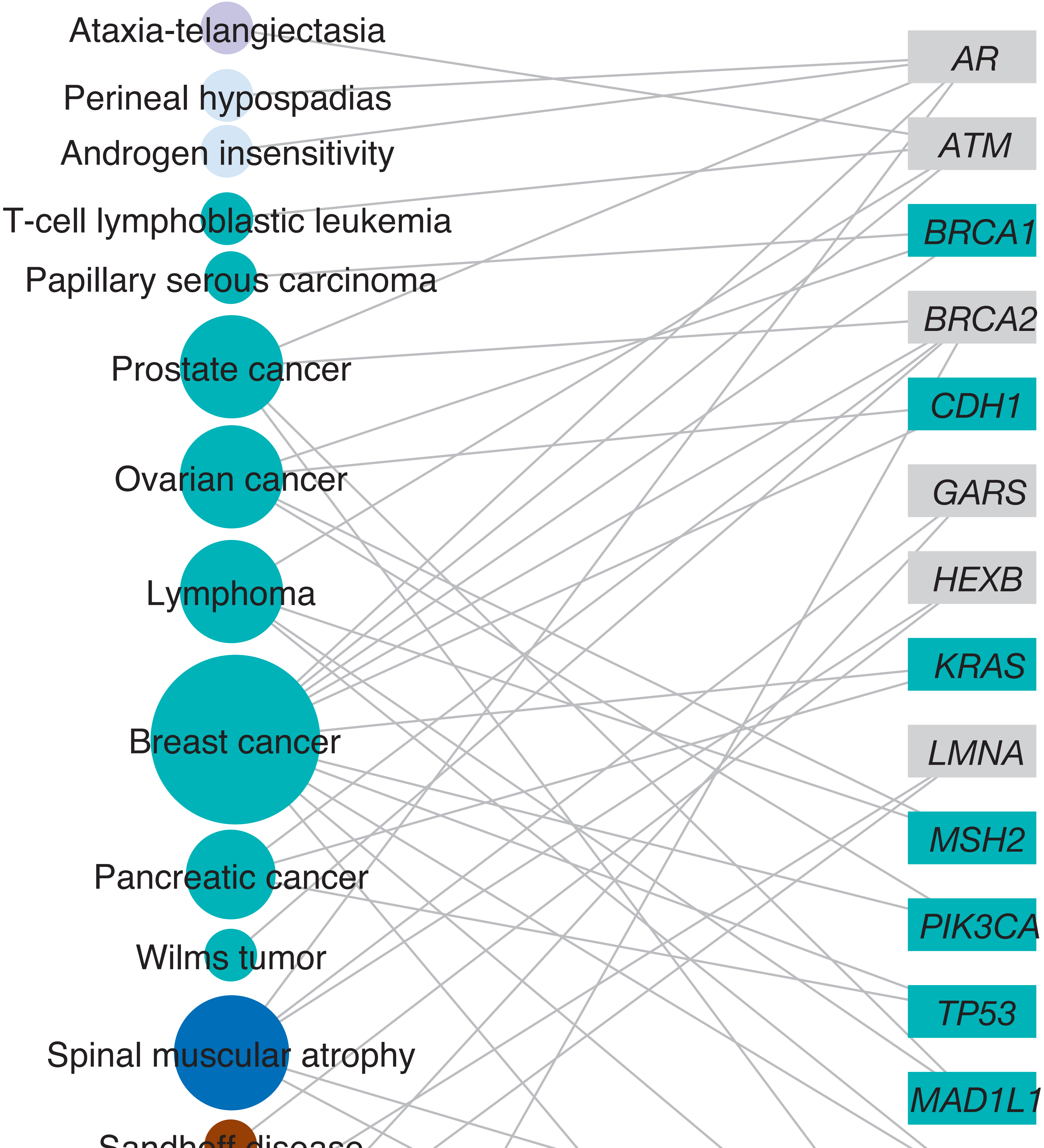
Goh, et al., *PNAS* 2007

“Diseaseome”

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



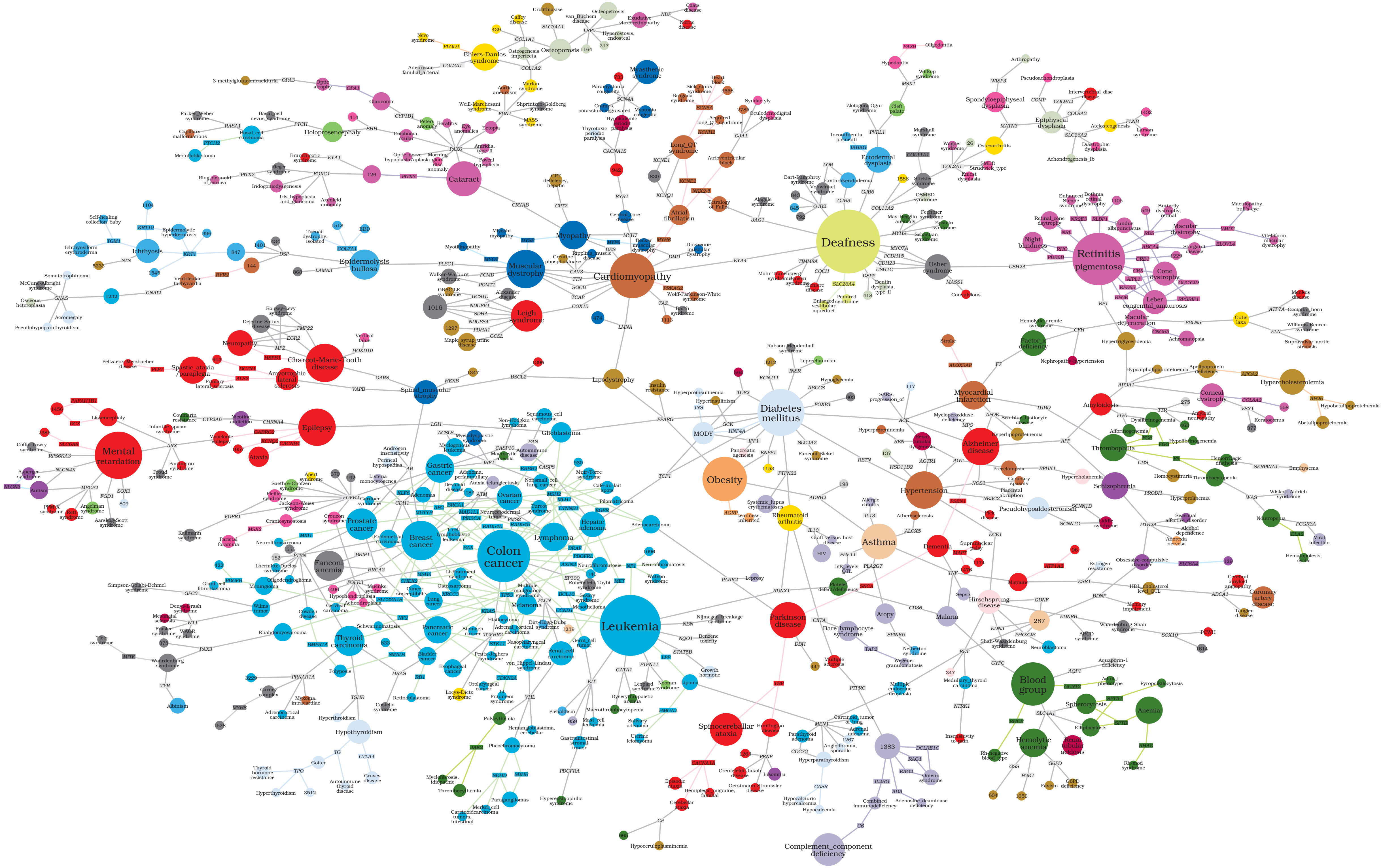
Goh, et al., *PNAS* 2007


Ataxia-telangiectasia
Perineal hypospadias
Androgen insensitivity
T-cell lymphoblastic leukemia
Papillary serous carcinoma
Prostate cancer
Ovarian cancer
Lymphoma
Breast cancer
Pancreatic cancer
Wilms tumor
Spinal muscular atrophy
Sandhoff disease
AR
ATM
BRCA1
BRCA2
CDH1
GARS
HEXB
KRAS
LMNA
MSH2
PIK3CA
TP53
MAD1L1

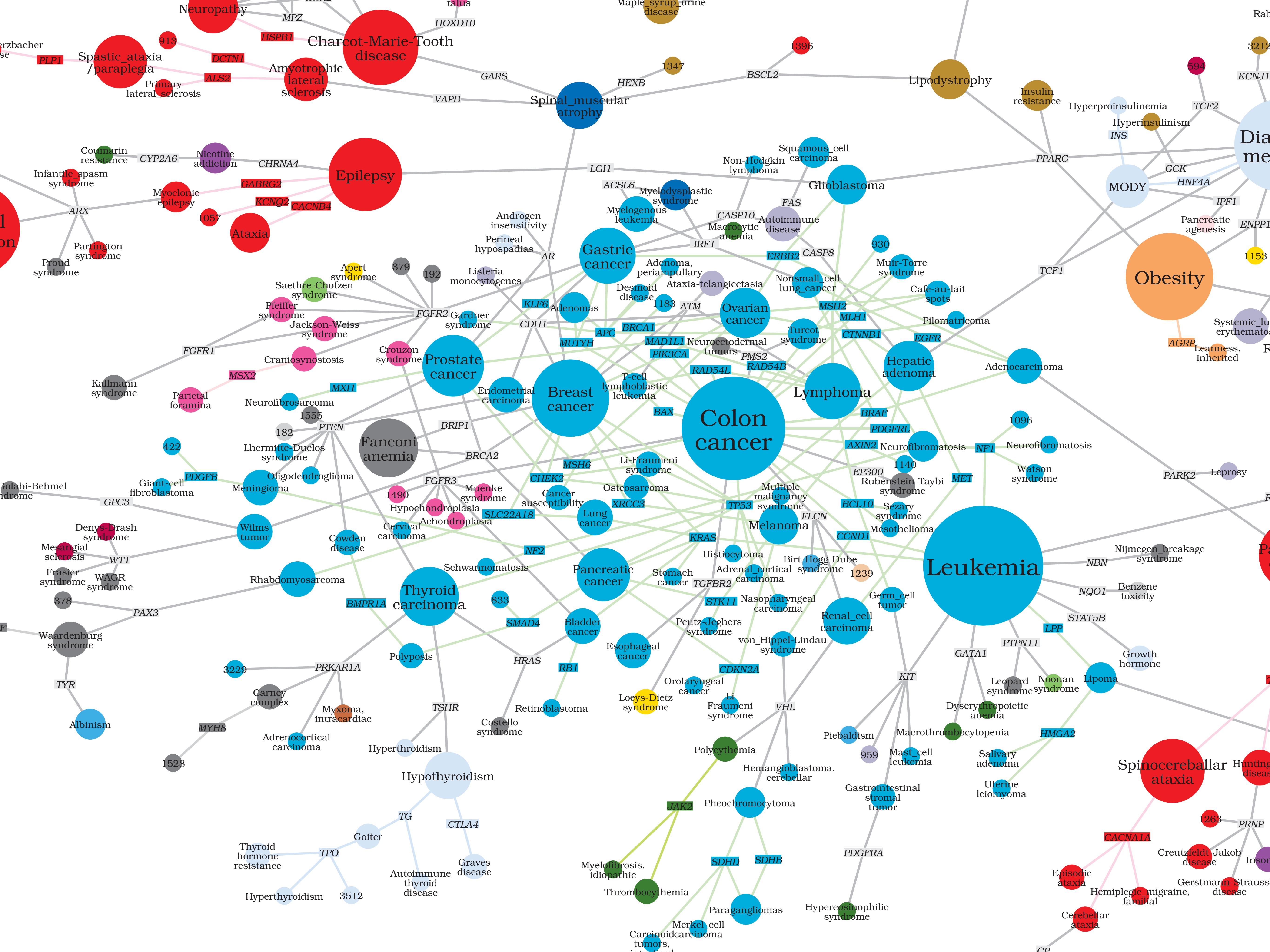
Goh, et al.,

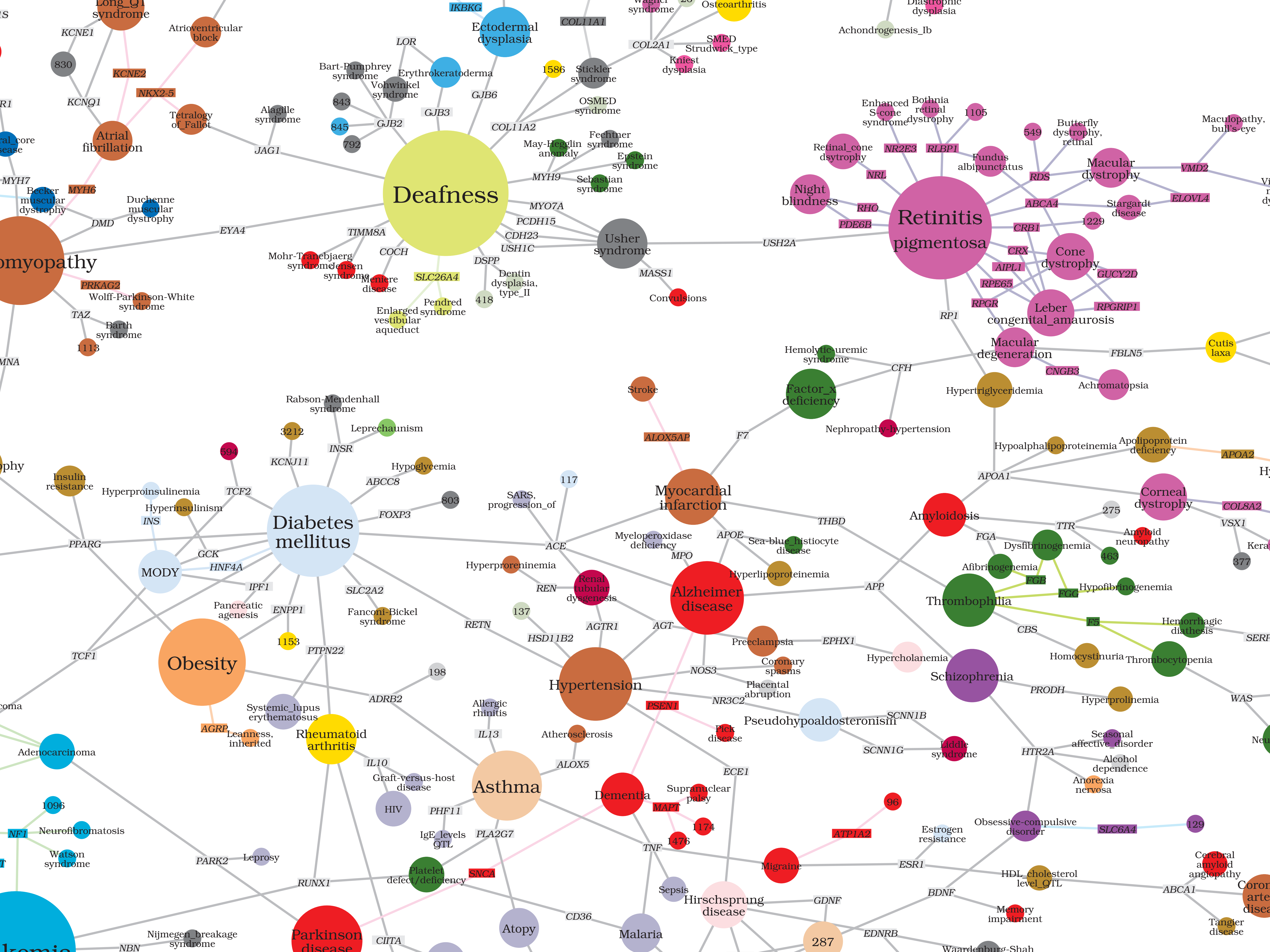
The human disease network

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



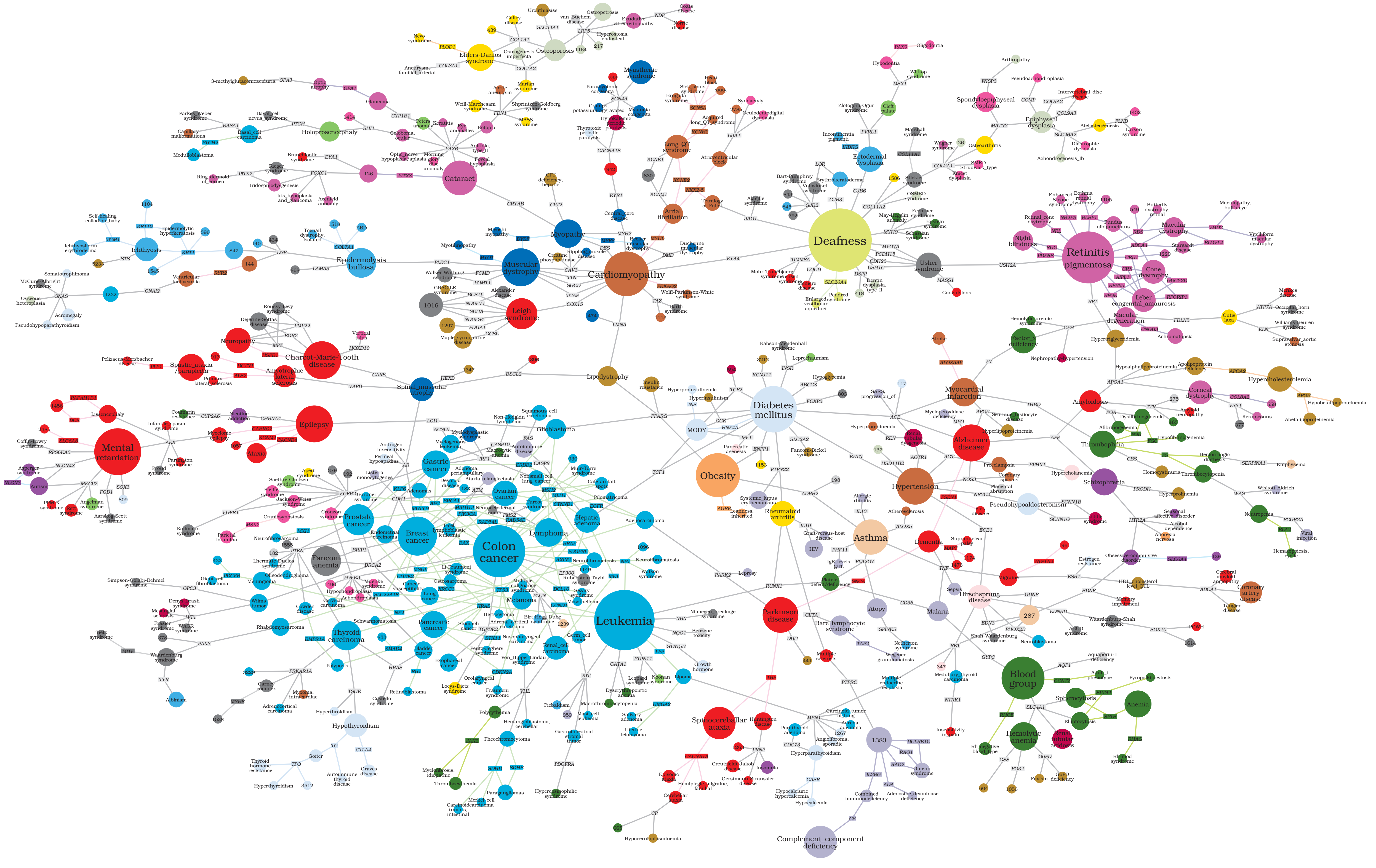
Goh, et al., *PNAS* 2007


Neuropathy
MPZ
talus
HOXD10
Maple_syrup_urine disease
Rab
HSPB1
Charcot-Marie-Tooth disease
913
PLP1
Spastic_ataxia /paraplegia
DCTN1
Amyotrophic lateral sclerosis
ALS2
Primary lateral_sclerosis
GARS
VAPB
Spinal_muscular atrophy
HEXB
1347
1396
BSCL2
Lipodystrophy
3212
594
KCNJ1
Insulin resistance
Hyperproinsulinemia
TCF2
Hyperinsulinism
INS
Coumarin resistance
CYP2A6
Nicotine addiction
CHRNA4
Epilepsy
LGI1
Squamous_cell carcinoma
Non-Hodgkin lymphoma
PPARG
GCK
MODY
HNF4A
Infantile_spasm syndrome
GABRG2
Myoclonic epilepsy
KCNQ2
CACNB4
1057
Ataxia
ARX
Partington syndrome
Proud syndrome
ACSL6
Myelodysplastic syndrome
Myelogenous leukemia
Glioblastoma
FAS
IPF1
Pancreatic agenesis
ENPP1
Androgen insensitivity
CASP10
Autoimmune disease
Macrocytic anemia
Perineal hypospadias
AR
IRF1
Gastric cancer
ERBB2
CASP8
930
1153
Obesity
Apert syndrome
379
192
Listeria monocytogenes
Adenoma, periampullary
Muir-Torre syndrome
TCF1
Saethre-Chotzen syndrome
Desmoid disease
Ataxia-telangiectasia
Nonsmall_cell lung_cancer
Cafe-au-lait spots
KLF6
1183
ATM
Ovarian cancer
MSH2
Pfeiffer syndrome
FGFR2
Gardner syndrome
Adenomas
MLH1
Pilomatricoma
Systemic_lu erythemato
Jackson-Weiss syndrome
CDH1
APC
BRCA1
Turcot syndrome
CTNNB1
AGRP
Leanness, inherited
MUTYH
MAD1L1
Neuroectodermal tumors
EGFR
FGFR1
Crouzon syndrome
Prostate cancer
PIK3CA
PMS2
Hepatic adenoma
Adenocarcinoma
Craniosynostosis
RAD54B
MSX2
RAD54L
Kallmann syndrome
Parietal foramina
MXI1
Endometrial carcinoma
Breast cancer
T-cell lymphoblastic leukemia
Lymphoma
Neurofibrosarcoma
1555
BAX
Colon cancer
BRAF
182
PTEN
Fanconi anemia
BRIP1
PDGFRL
1096
422
Lhermitte-Duclos syndrome
BRCA2
AXIN2
Neurofibromatosis
NF1
Neurofibromatosis
MSH6
Li-Fraumeni syndrome
1140
Golabi-Behmel syndrome
PDGFB
Giant-cell fibroblastoma
Oligodendroglioma
FGFR3
CHEK2
EP300
Rubenstein-Taybi syndrome
MET
Watson syndrome
PARK2
Leprosy
Meningioma
Muenke syndrome
Osteosarcoma
Multiple malignancy syndrome
GPC3
1490
Hypochondroplasia
Cancer susceptibility
XRCC3
TP53
BCL10
Sezary syndrome
Wilms tumor
Cervical carcinoma
Achondroplasia
SLC22A18
Lung cancer
FLCN
Melanoma
Denys-Drash syndrome
Cowden disease
Mesothelioma
CCND1
Mesangial sclerosis
WT1
NF2
KRAS
Histiocytoma
Birt-Hogg-Dube syndrome
Nijmegen_breakage syndrome
Frasier syndrome
WAGR syndrome
Schwannomatosis
Pancreatic cancer
Stomach cancer
Adrenal_cortical carcinoma
1239
Leukemia
NBN
Rhabdomyosarcoma
TGFBR2
Thyroid carcinoma
378
BMPR1A
833
STK11
Nasopharyngeal carcinoma
Germ_cell tumor
NQO1
Benzene toxicity
PAX3
SMAD4
Bladder cancer
Peutz-Jeghers syndrome
Renal_cell carcinoma
STAT5B
Waardenburg syndrome
LPP
Polyposis
Esophageal cancer
von_Hippel-Lindau syndrome
PTPN11
GATA1
Growth hormone
3229
PRKAR1A
HRAS
RB1
CDKN2A
KIT
TYR
Carney complex
Oropharyngeal cancer
Leopard syndrome
Noonan syndrome
Lipoma
Loeys-Dietz syndrome
Li Fraumeni syndrome
Dyserythropoietic anemia
Albinism
MYH8
Myxoma, intracardiac
TSHR
Retinoblastoma
VHL
Costello syndrome
Macrothrombocytopenia
Piebaldism
HMGA2
Adrenocortical carcinoma
Polycythemia
959
Mast_cell leukemia
Salivary adenoma
Spinocereballar ataxia
Hunting disease
1528
Hyperthroidism
Hemangioblastoma, cerebellar
Hypothyroidism
Uterine leiomyoma
Gastrointestinal stromal tumor
JAK2
Pheochromocytoma
1263
TG
CTLA4
PRNP
CACNA1A
Goiter
Thyroid hormone resistance
TPO
PDGFRA
Creutzfeldt-Jakob disease
Graves disease
Autoimmune thyroid disease
Myelofibrosis, idiopathic
SDHD
SDHB
Episodic ataxia
Inson
Gerstmann-Strauss disease
Hyperthyroidism
3512
Thrombocythemia
Hemiplegic_migraine, familial
Paragangliomas
Hypereosinophilic syndrome
Merkel_cell carcinoma
Carcinoid tumors, intestinal
Cerebellar ataxia
CP

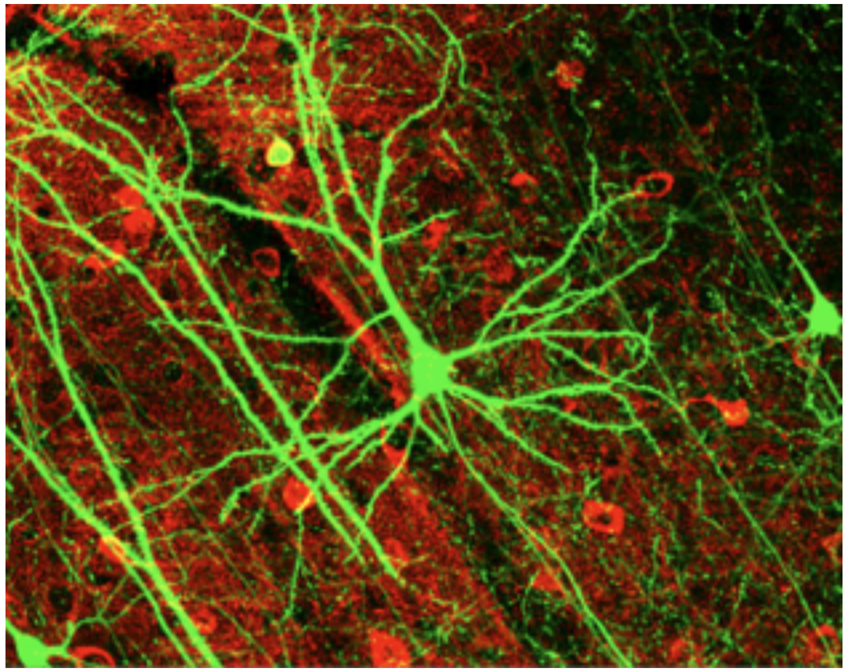

Long_QT syndrome
KCNE1
Atrioventricular block
830
KCNE2
KCNQ1
NKX2-5
Tetralogy of_Fallot
Atrial fibrillation
MYH7
MYH6
Becker muscular dystrophy
Duchenne muscular dystrophy
DMD
EYA4
omyopathy
PRKAG2
Wolff-Parkinson-White syndrome
TAZ
Barth syndrome
1113
MNA
IKBKG
Ectodermal dysplasia
COL11A1
LOR
Bart-Pumphrey syndrome
Erythrokeratoderma
Vohwinkel syndrome
843
GJB3
GJB6
845
GJB2
792
Alagille syndrome
JAG1
Deafness
1586
Stickler syndrome
COL11A2
OSMED syndrome
May-Hegglin anomaly
Fechtner syndrome
Epstein syndrome
MYH9
Sebastian syndrome
MYO7A
PCDH15
CDH23
USH1C
Usher syndrome
TIMM8A
COCH
Mohr-Tranebjaerg syndrome
Jensen syndrome
Meniere disease
SLC26A4
DSPP
Dentin dysplasia, type_II
418
Pendred syndrome
Enlarged vestibular aqueduct
MASS1
Convulsions
USH2A
Wagner syndrome
Osteoarthritis
COL2A1
SMED Strudwick_type
Kniest dysplasia
Diastrophic dysplasia
Achondrogenesis_Ib
Enhanced S-cone syndrome
Bothnia retinal dystrophy
1105
549
Butterfly dystrophy, retinal
Maculopathy, bull's-eye
Retinal_cone dystrophy
NR2E3
RLBP1
Fundus albipunctatus
RDS
Macular dystrophy
VMD2
NRL
Night blindness
RHO
PDE6B
ABCA4
Stargardt disease
ELOVL4
Retinitis pigmentosa
CRB1
1229
CRX
Cone dystrophy
AIPL1
GUCY2D
RPE65
RPGR
Leber congenital_amaurosis
RPGRIP1
RP1
Macular degeneration
FBLN5
Cutis laxa
CNGB3
Achromatopsia
Hemolytic-uremic syndrome
CFH
Factor_x deficiency
Hypertriglyceridemia
Nephropathy-hypertension
Stroke
ALOX5AP
F7
Rabson-Mendenhall syndrome
3212
Leprechaunism
594
KCNJ11
INSR
Hypoglycemia
ABCC8
803
FOXP3
117
SARS, progression_of
phy
Insulin resistance
Hyperproinsulinemia
TCF2
Hyperinsulinism
INS
PPARG
Diabetes mellitus
GCK
MODY
HNF4A
IPF1
Pancreatic agenesis
ENPP1
SLC2A2
Fanconi-Bickel syndrome
1153
PTPN22
TCF1
Obesity
ACE
Myocardial infarction
THBD
Myeloperoxidase deficiency
APOE
Sea-blue_histiocyte disease
MPO
Hyperproreninemia
REN
Renal tubular dysgenesis
Alzheimer disease
Hyperlipoproteinemia
APP
137
RETN
HSD11B2
AGTR1
AGT
Preeclampsia
EPHX1
Hypercholanemia
Coronary spasms
NOS3
Placental abruption
Hypertension
NR3C2
PSEN1
Pseudohypoaldosteronism
SCNN1B
SCNN1G
Liddle syndrome
Pick disease
Hypoalphalipoproteinemia
Apolipoprotein deficiency
APOA2
APOA1
Amyloidosis
Corneal dystrophy
275
TTR
COL8A2
VSX1
Amyloid neuropathy
FGA
Dysfibrinogenemia
463
377
Afibrinogenemia
FGB
FGG
Hypofibrinogenemia
Thrombophilia
F5
Hemorrhagic diathesis
CBS
Homocystinuria
Thrombocytopenia
Schizophrenia
PRODH
Hyperprolinemia
WAS
coma
198
ADRB2
Systemic_lupus erythematosus
AGRP
Leanness, inherited
Rheumatoid arthritis
Allergic rhinitis
IL13
Atherosclerosis
ALOX5
ECE1
Adenocarcinoma
IL10
Graft-versus-host disease
Asthma
Dementia
Supranuclear palsy
MAPT
HIV
PHF11
IgE_levels QTL
PLA2G7
1174
1476
96
ATP1A2
TNF
Migraine
ESR1
Estrogen resistance
Seasonal affective_disorder
HTR2A
Alcohol dependence
Anorexia nervosa
Obsessive-compulsive disorder
SLC6A4
129
1096
Neurofibromatosis
NF1
Watson syndrome
PARK2
Leprosy
RUNX1
Platelet defect/deficiency
SNCA
Sepsis
Hirschsprung disease
GDNF
BDNF
Memory impairment
HDL_cholesterol level_QTL
Cerebral amyloid angiopathy
ABCA1
Tangier disease
CD36
Atopy
Malaria
287
EDNRB
Waardenburg-Shah
Parkinson disease
CIITA
Nijmegen_breakage syndrome
NBN
kemia

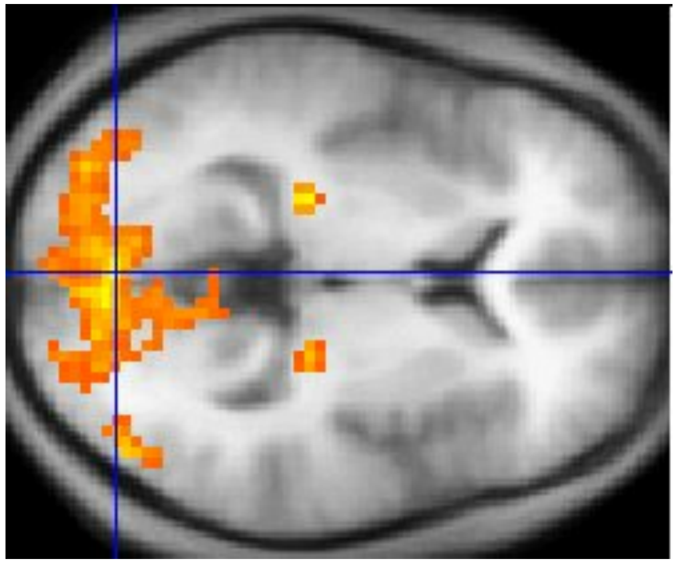
The human disease network

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

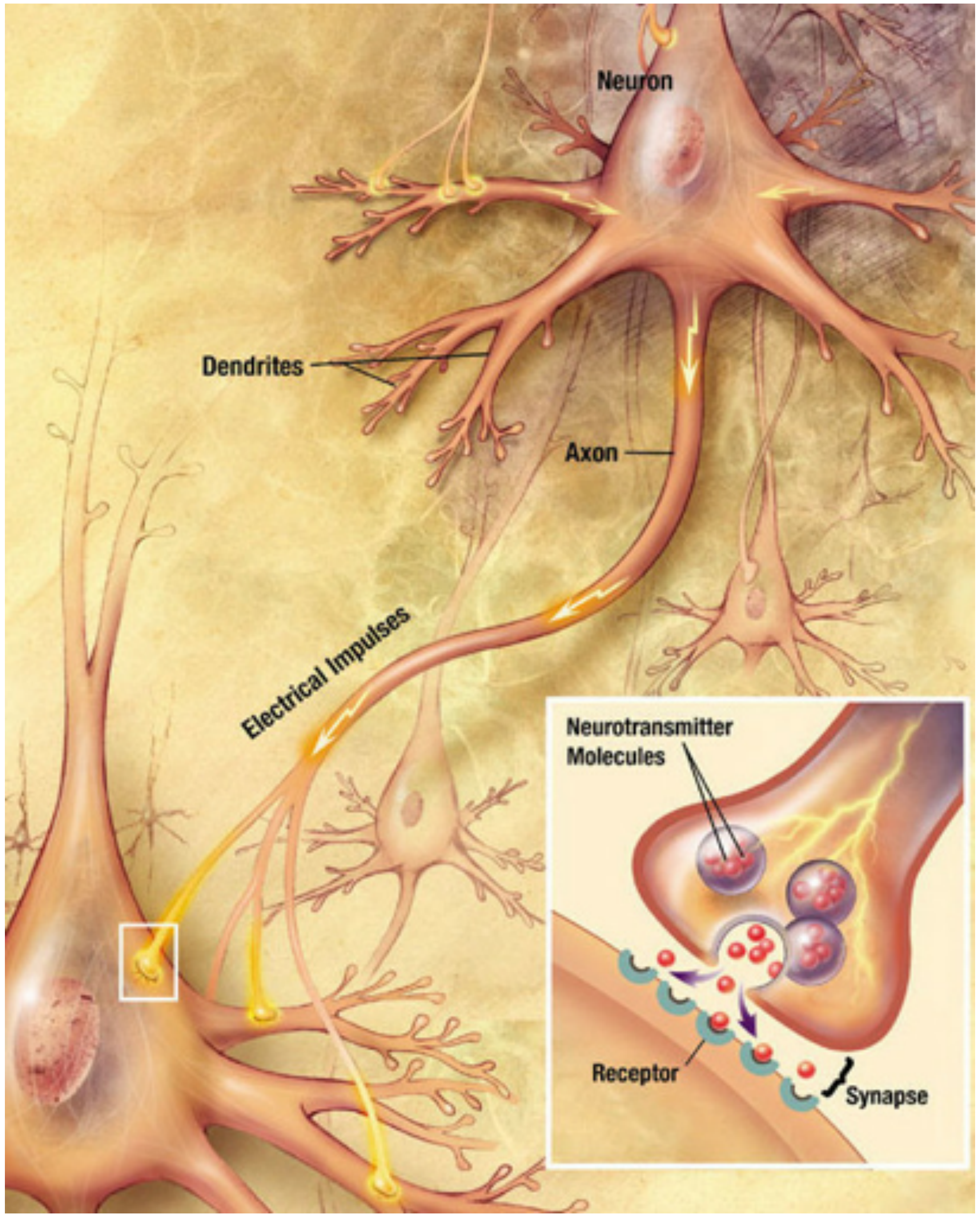


Neuroscience

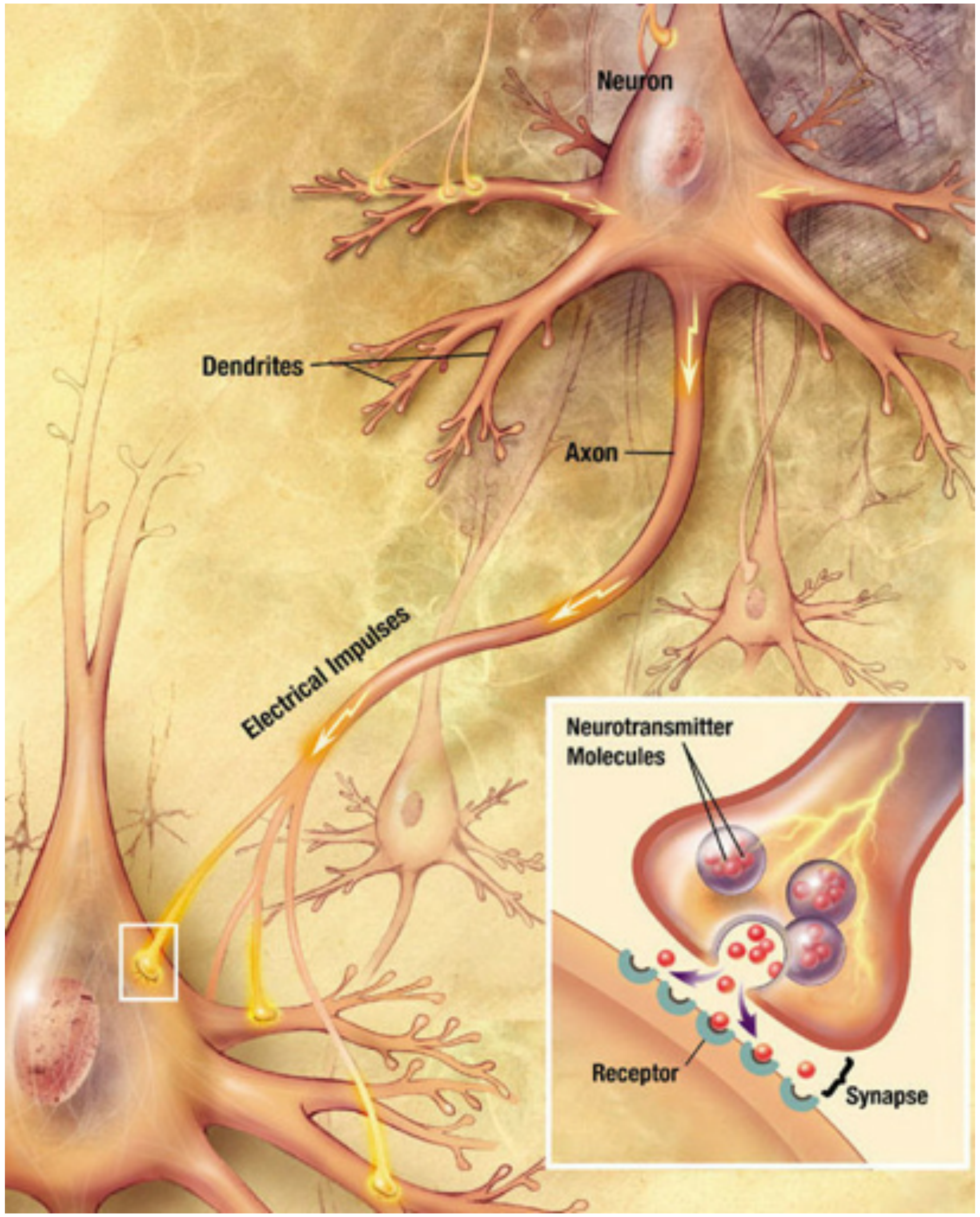




Neuroscience



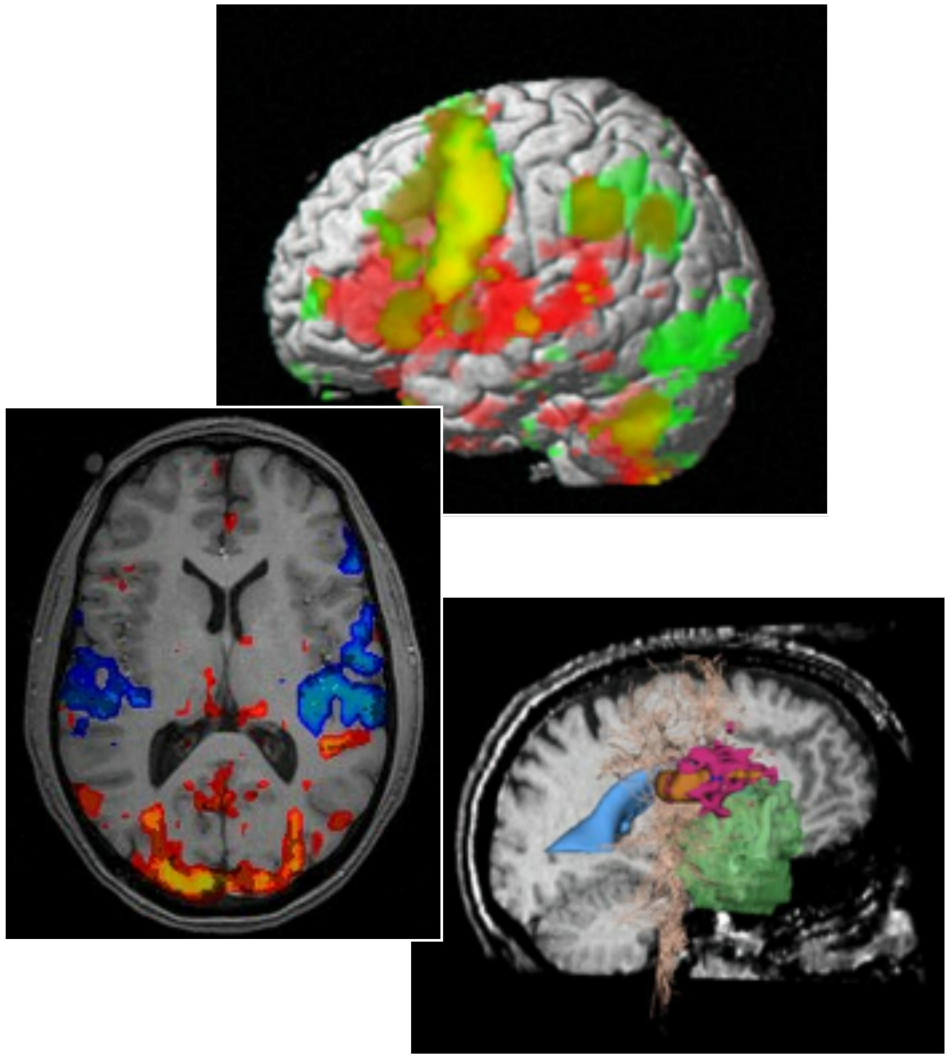
Neuroscience



Network between **neurons** transmitting electrochemical **signals**

Neuroscience

Networks from **Neuroimaging**

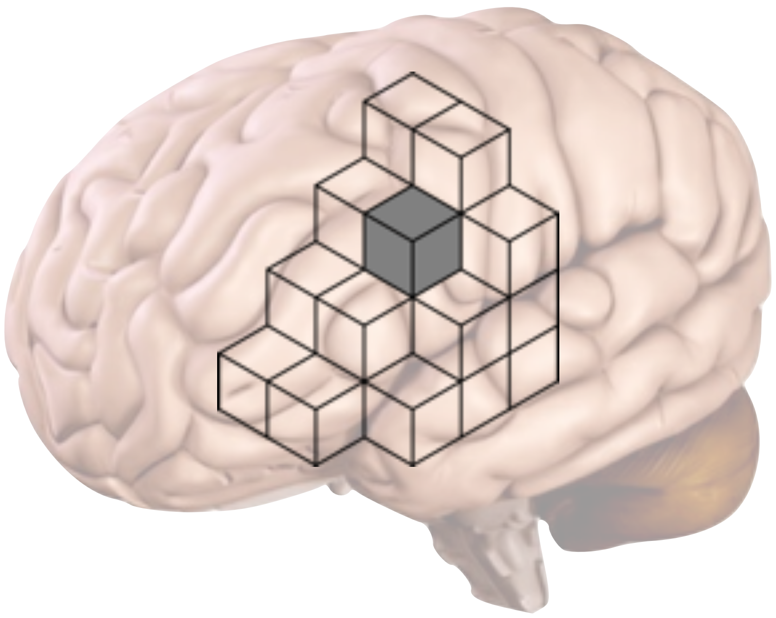


Neuroscience

Networks from **Neuroimaging**

fMRI

Divide the brain
into **voxels**

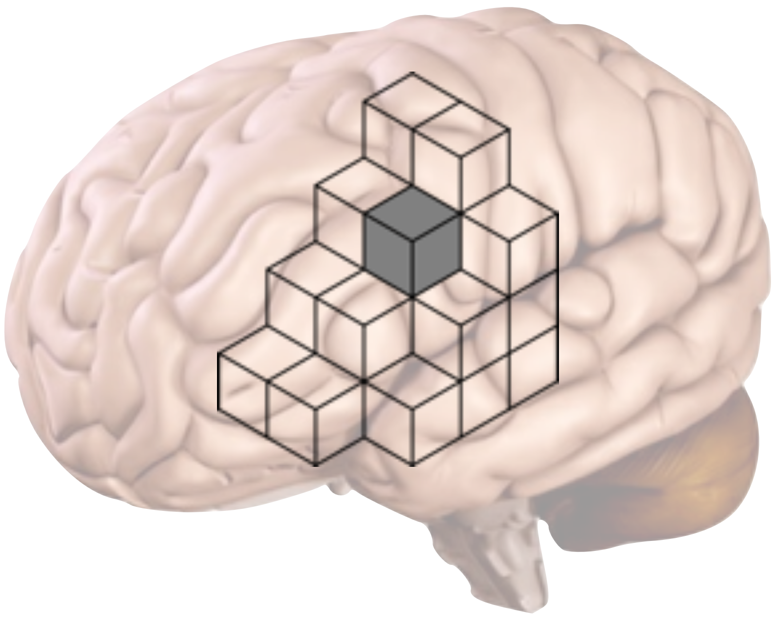


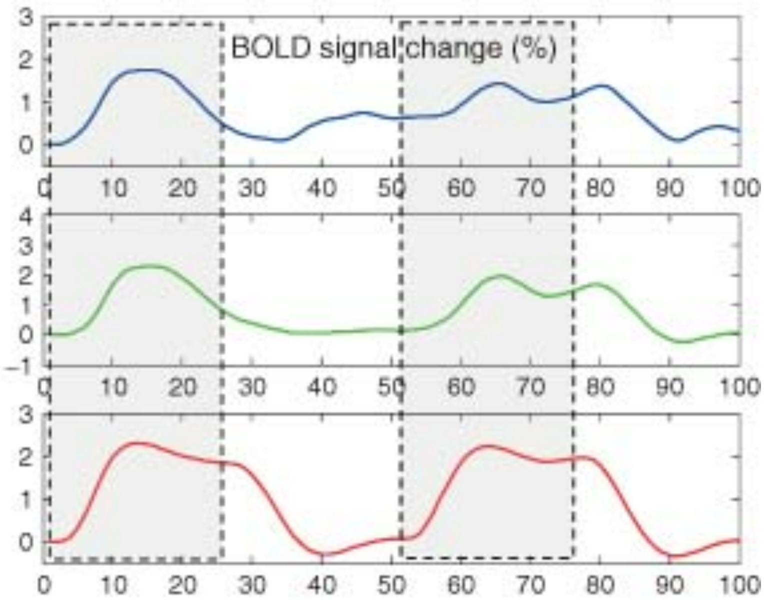
Neuroscience

Networks from **Neuroimaging**

fMRI

Divide the brain into **voxels** → Measure time series of blood flow inside each voxel (BOLD)

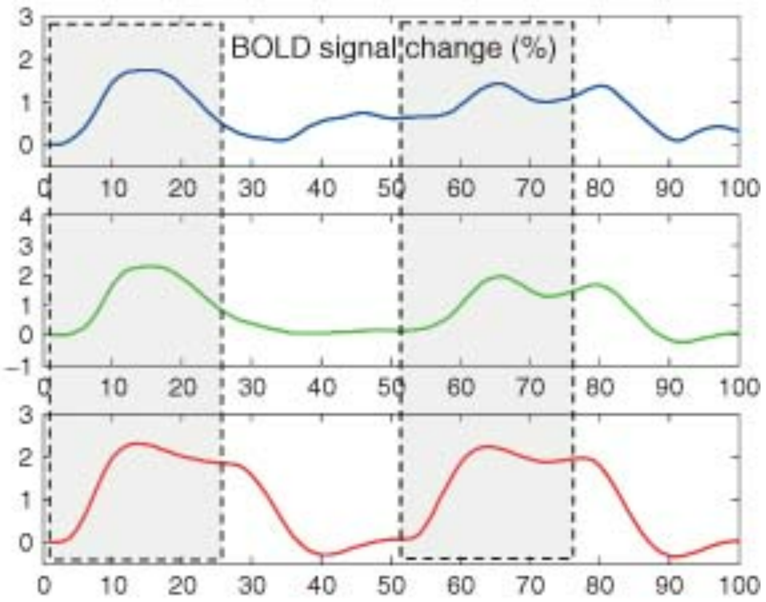


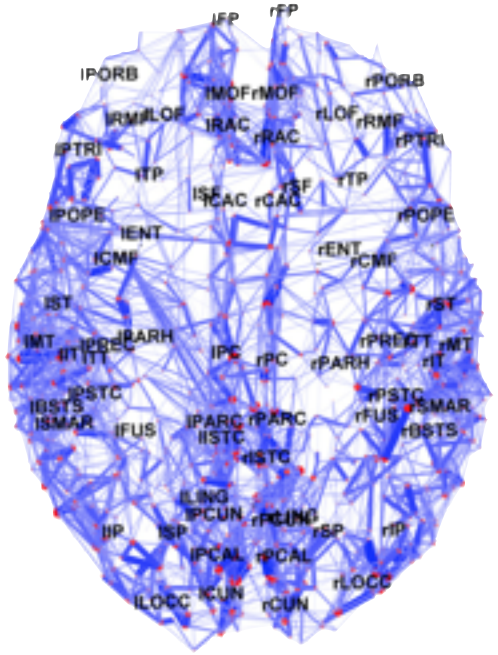


Neuroscience

Networks from **Neuroimaging** fMRI

Measure time series of blood flow inside each voxel (BOLD) ⟶ Links between voxels with temporally correlated time series



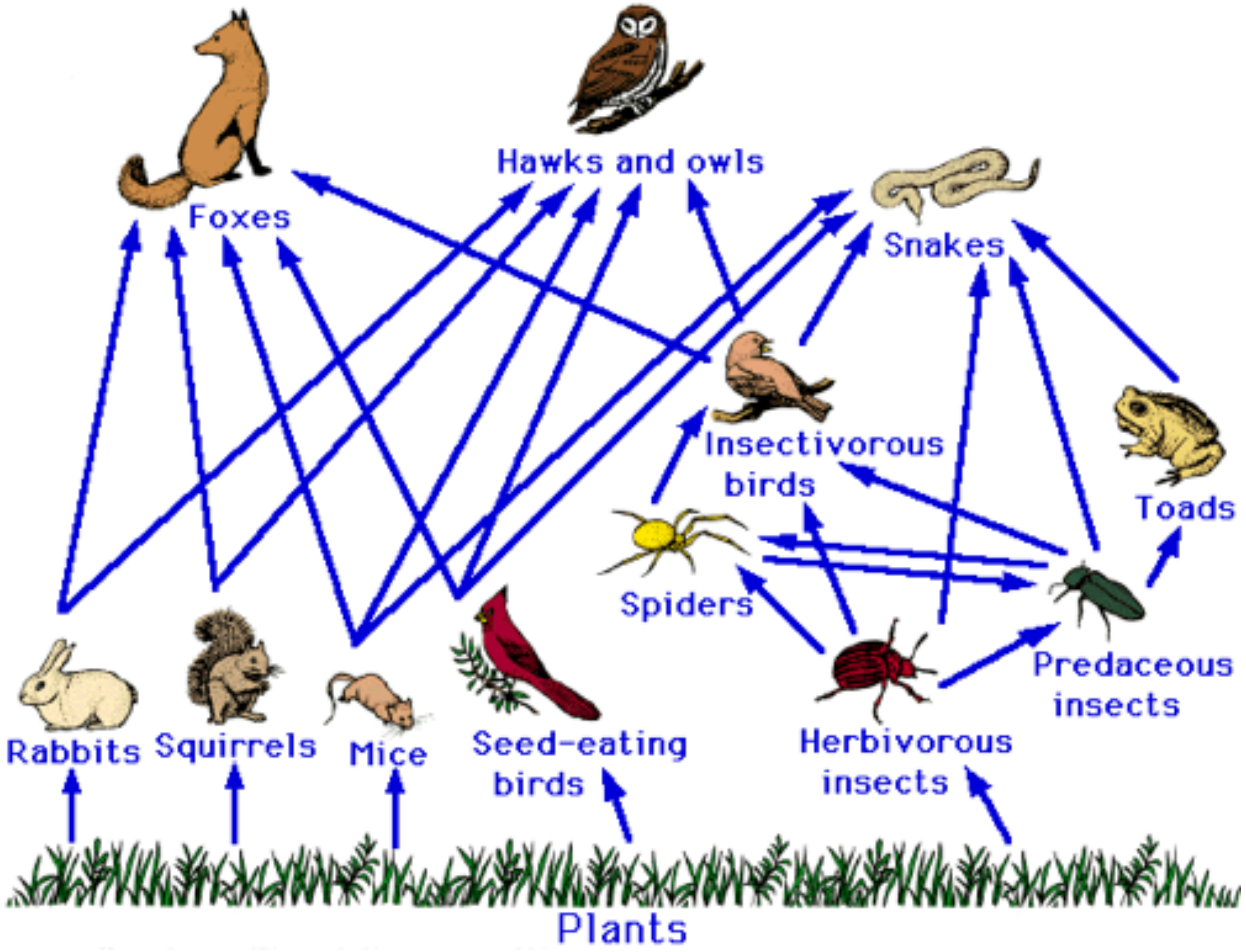


Ecology

Food webs

Ecology

Food webs

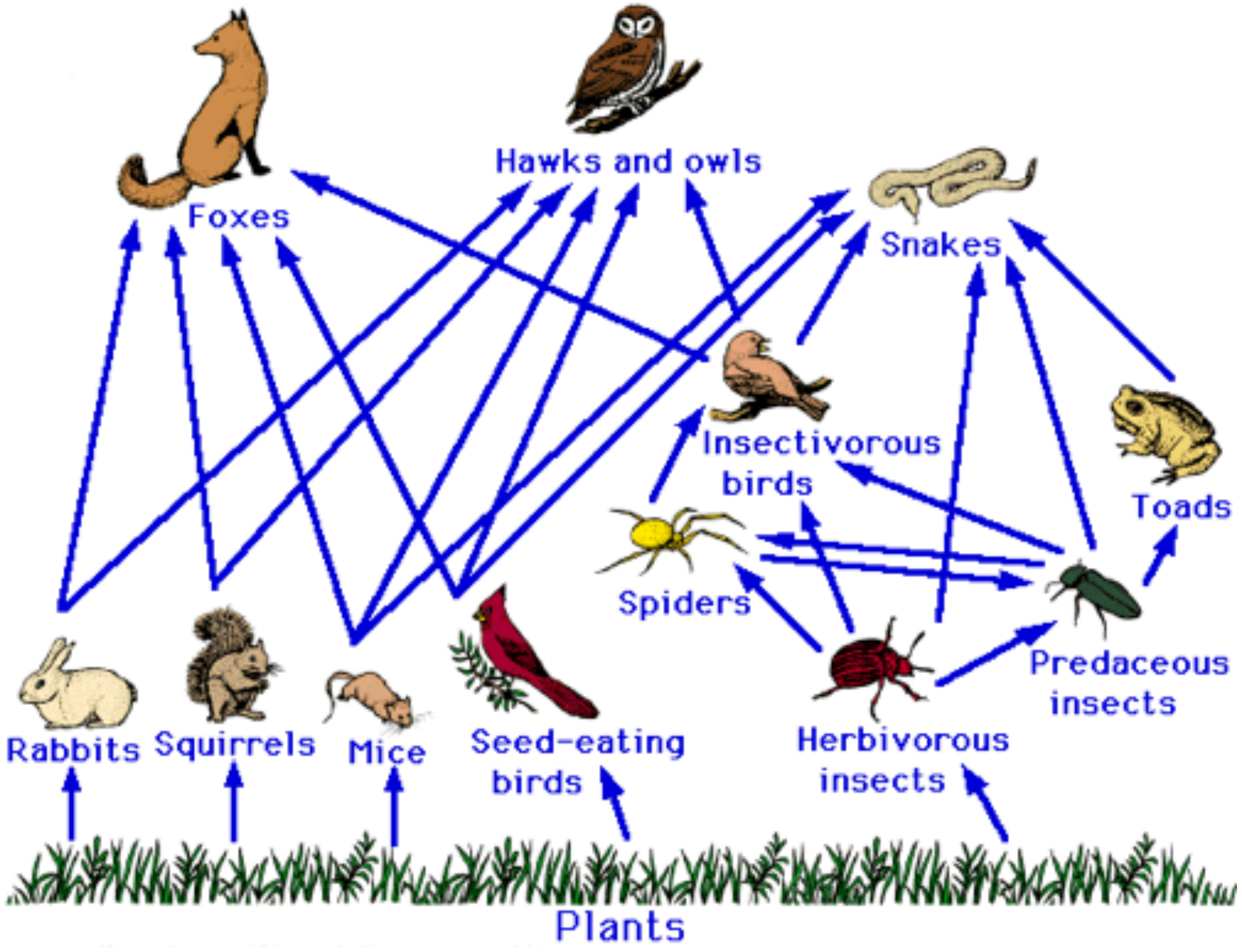


Ecology

Food webs

Nodes:

- Organisms (compartments)



Ecology

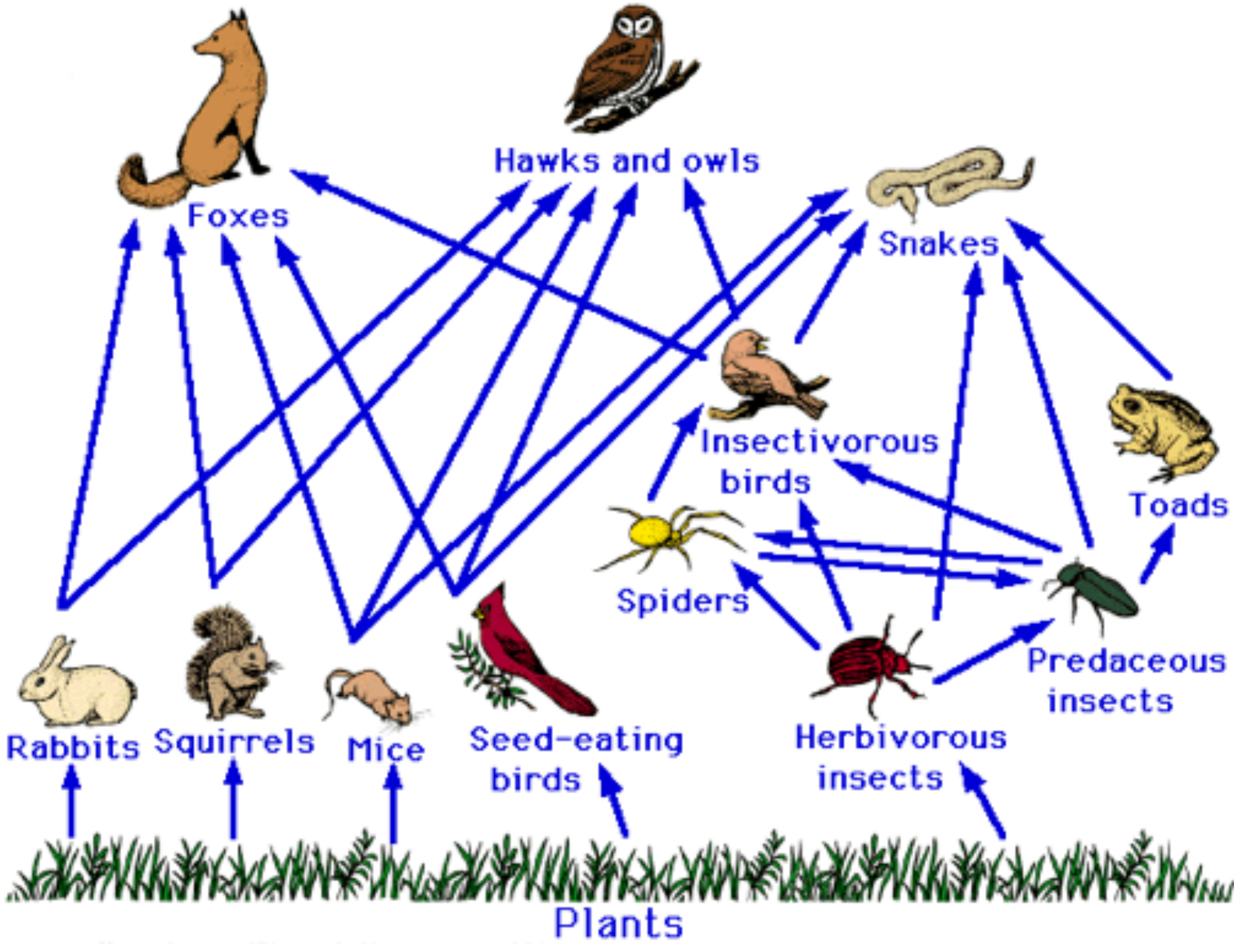
Food webs

Nodes:

- Organisms (compartments)

Links:

- Organisms **exchange carbon**



Ecology

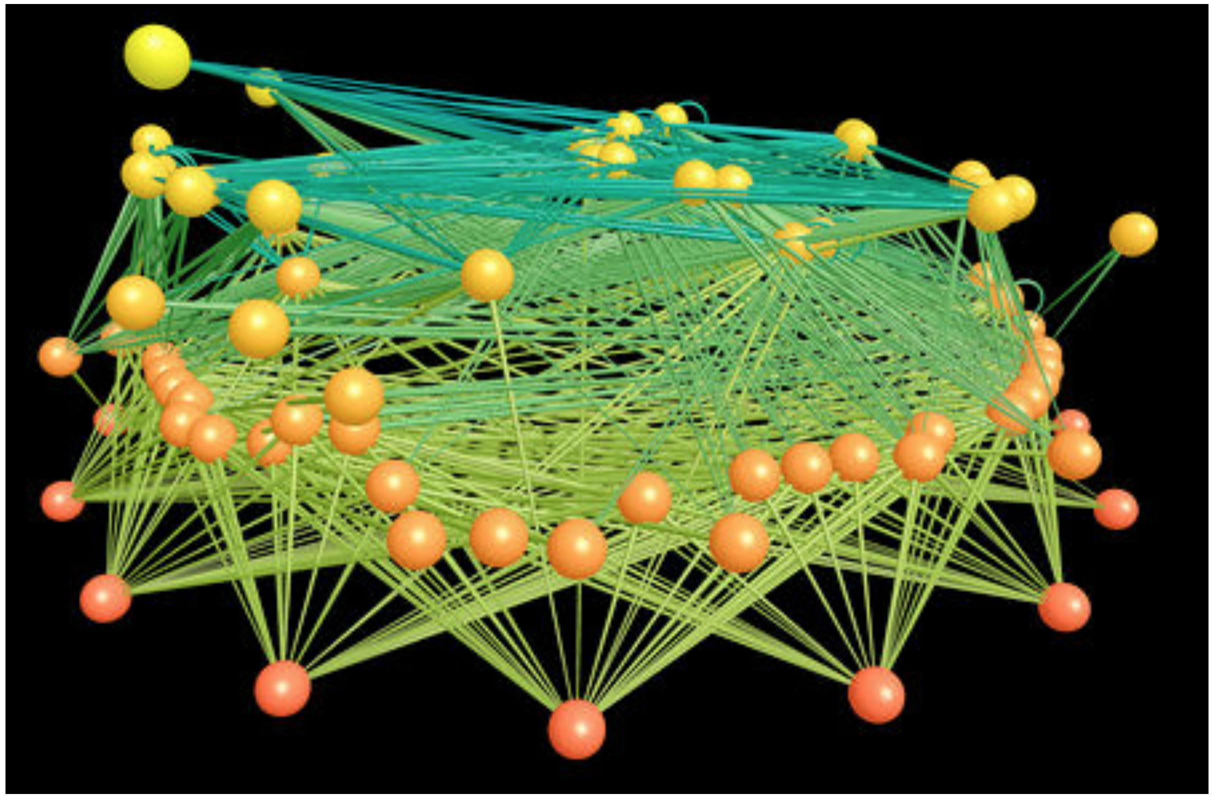
Food webs

Nodes:

Organisms (compartments)

Links:

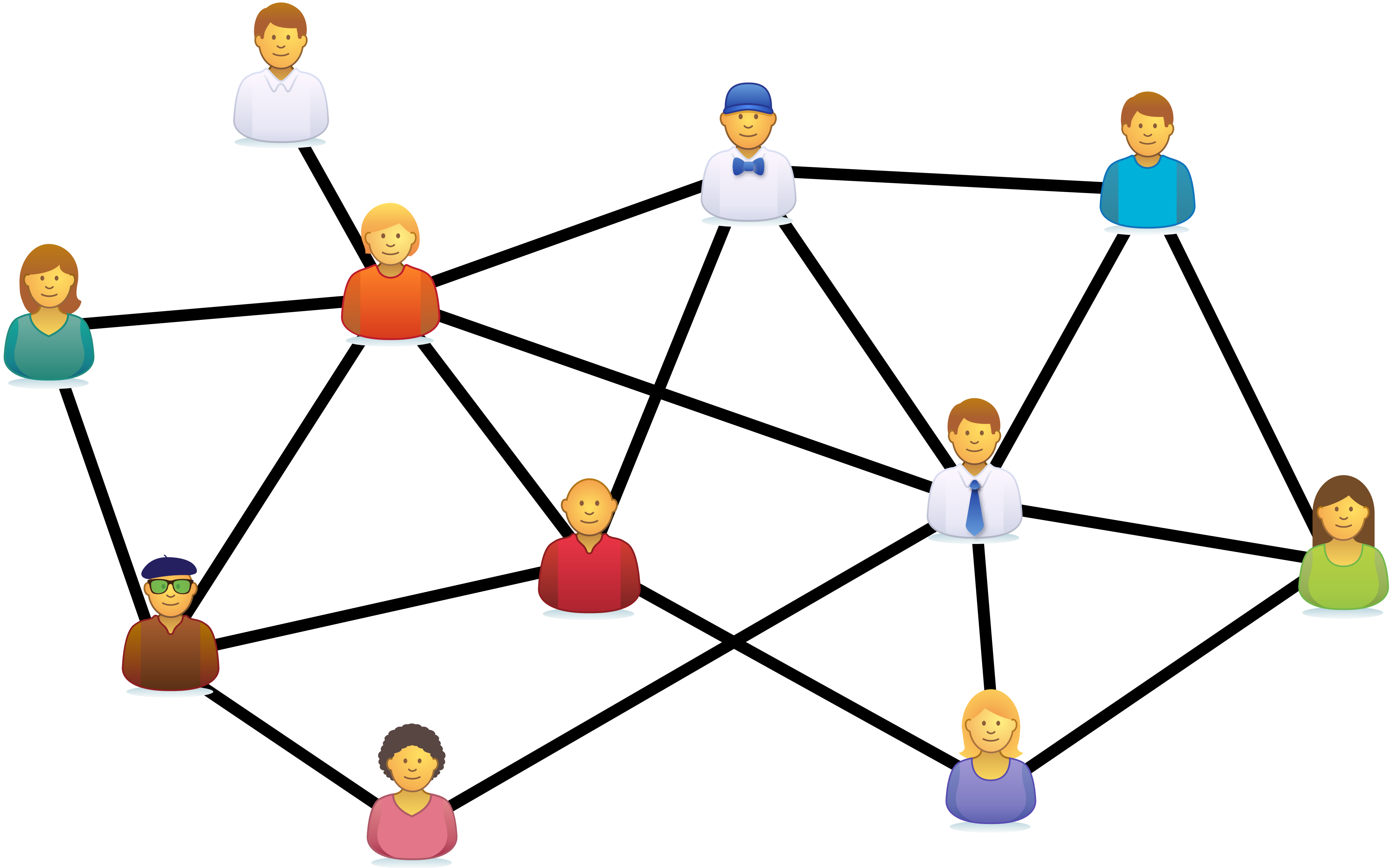
Organisms **exchange carbon**



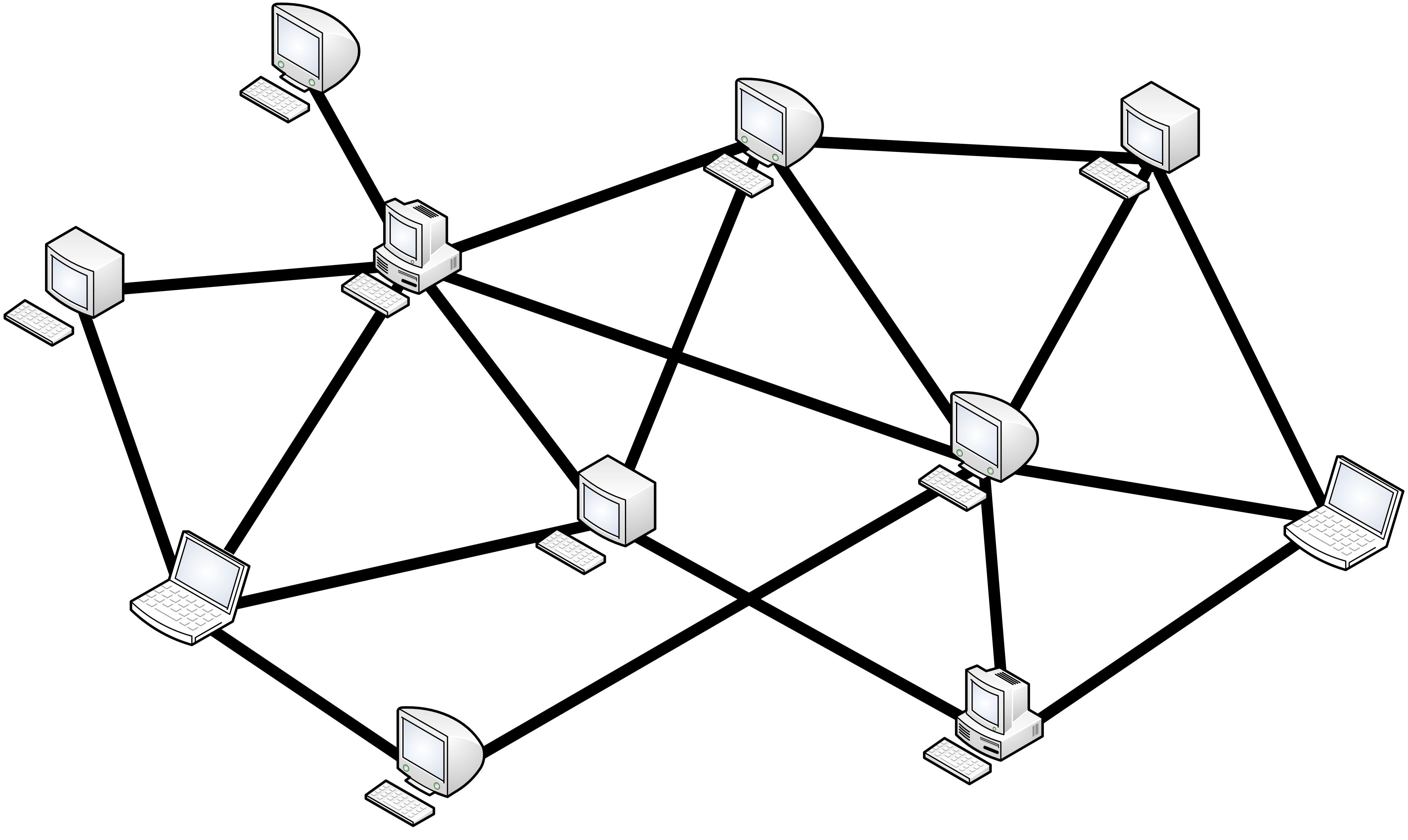
N. Martinez / Science Photo Library

Summary so far

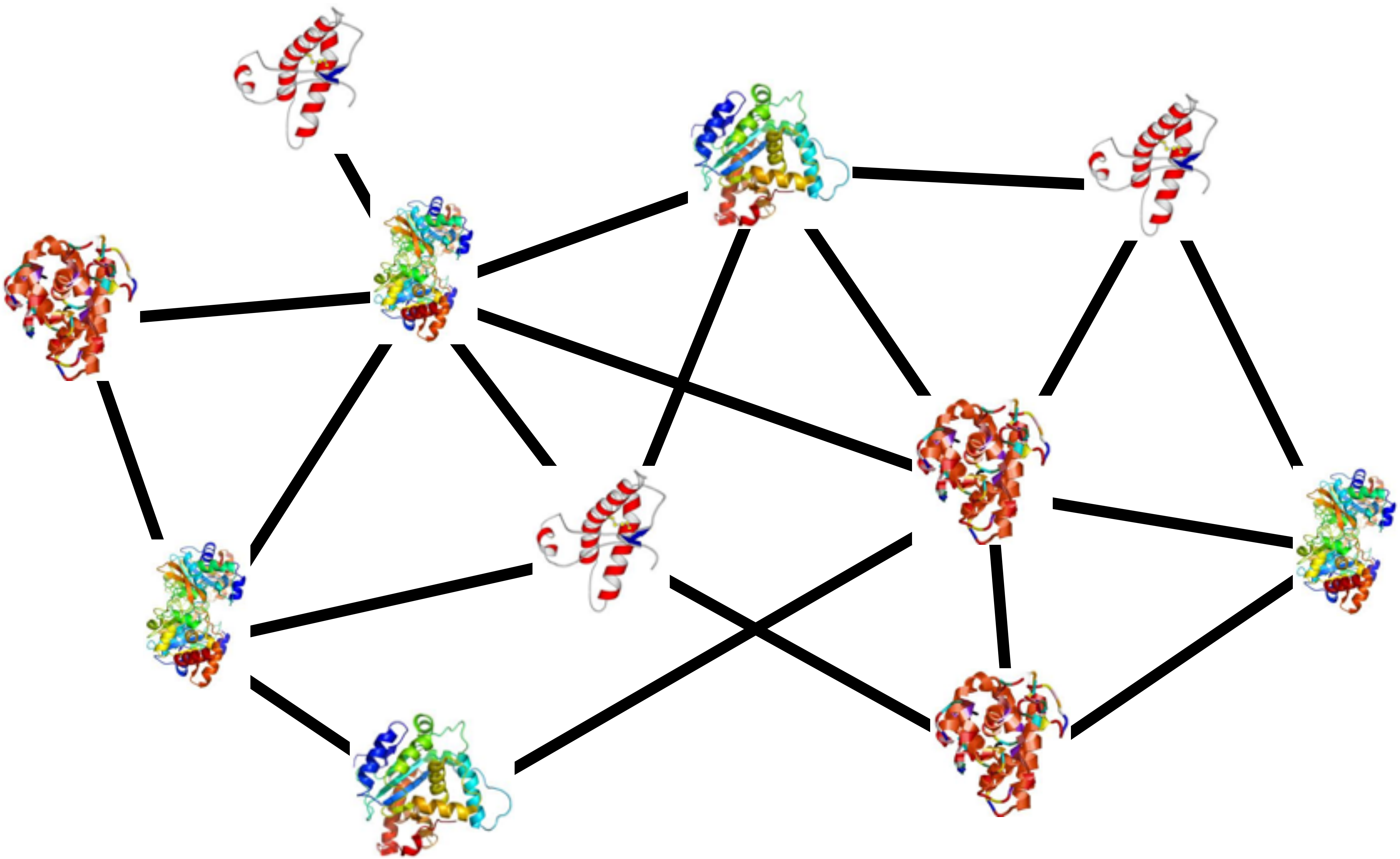
Networks



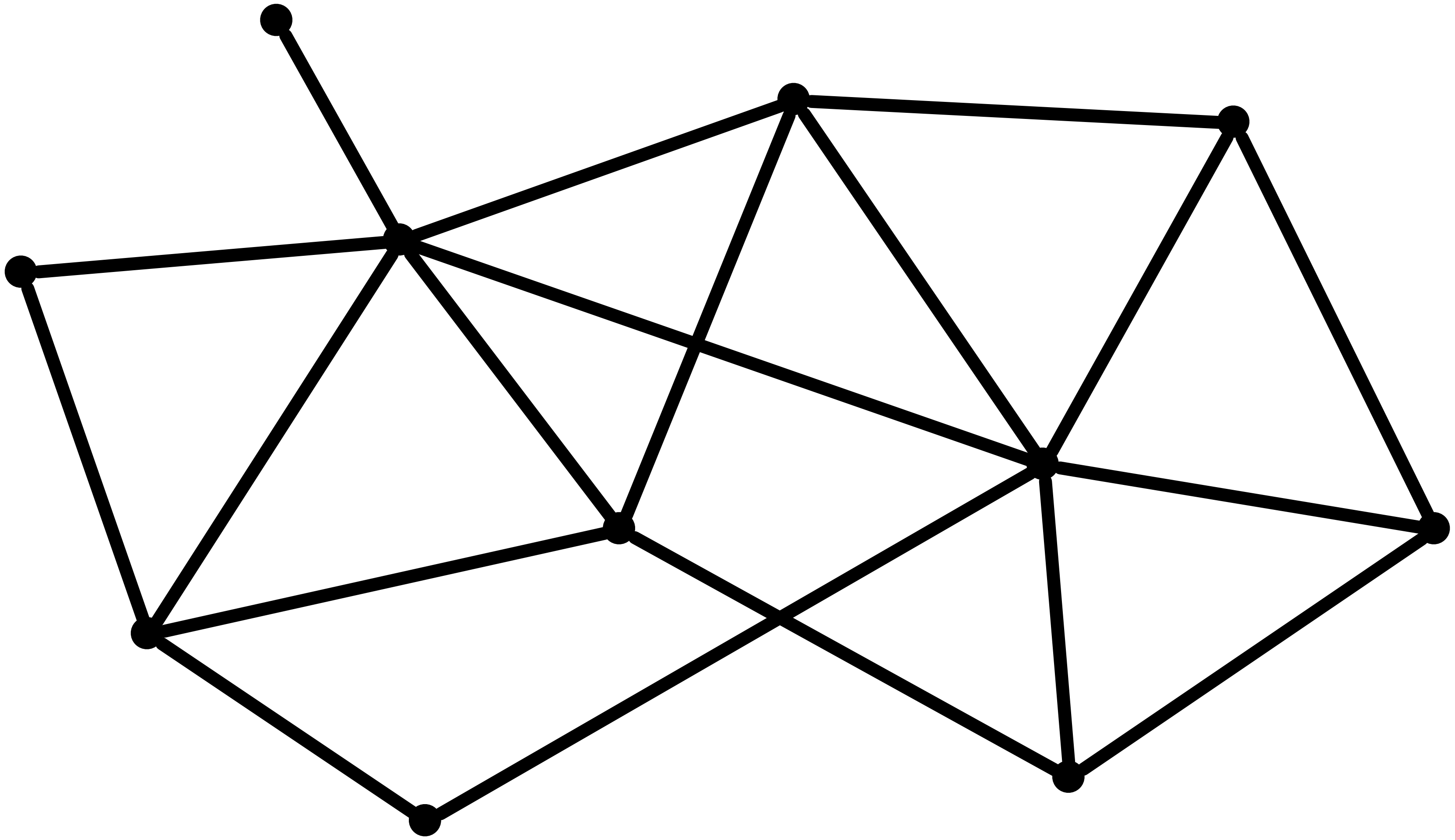
Networks



Networks



Networks

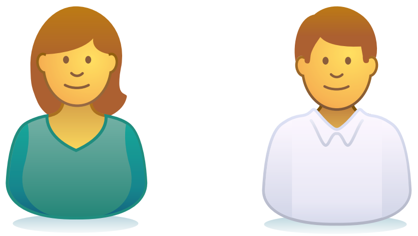


very
general

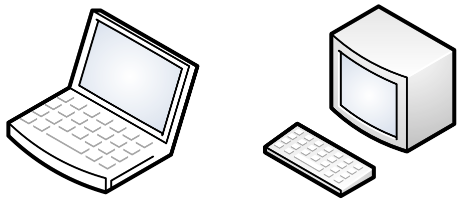
objects and the
relationships
between them

Networks

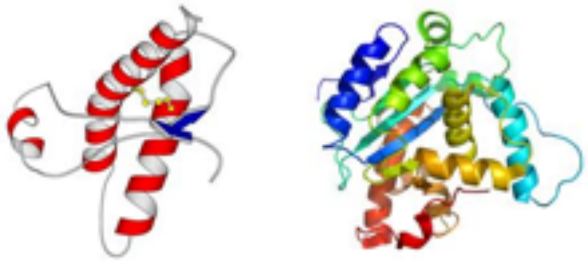
objects **relationships**



friendship, family, sexual



transmit data, shared power

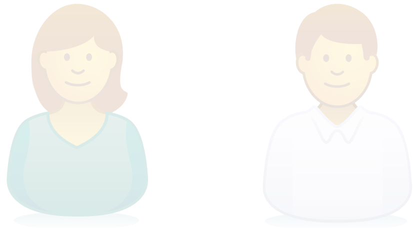


bind together, signal transduction

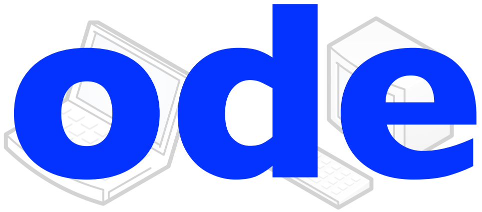
Networks

objects

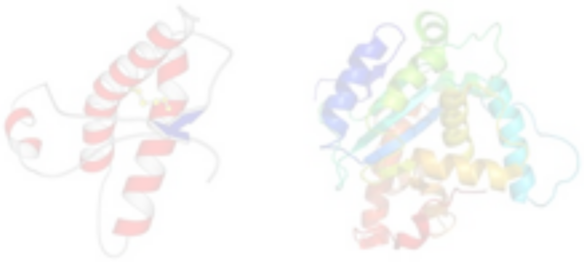
relationships



friendship, family, sexual



transmit data, shared power



bind together, signal transduction

nodes

links

Why study networks?

Why study networks?

Simple components

Interacting

Complex systems

Network Quantifiers

(Advanced terminology)

(and some major results)

Networks and graphs

$$G = \{V, E\}$$

Graph for
the network

Set of nodes
(vertices)

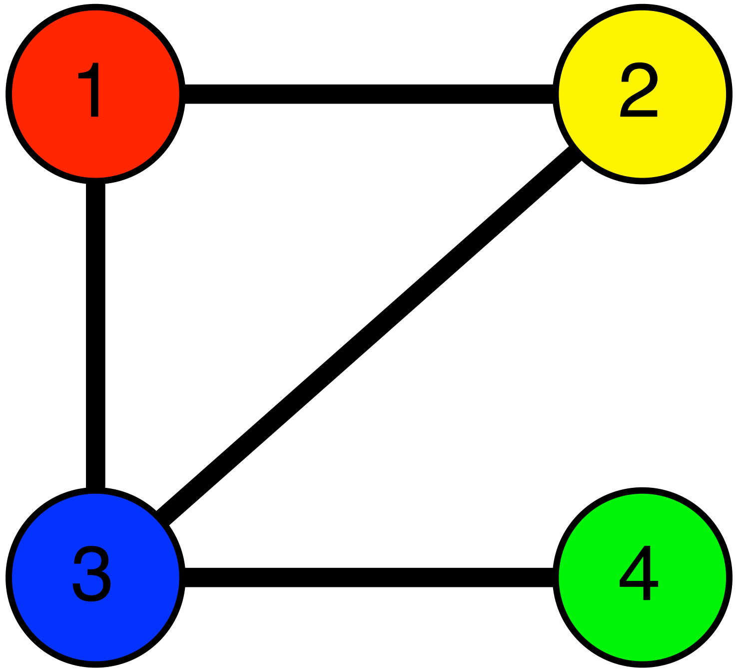
Set of links
(edges)

Adjacency matrix

Matrix of 1's and 0's storing which nodes are connected

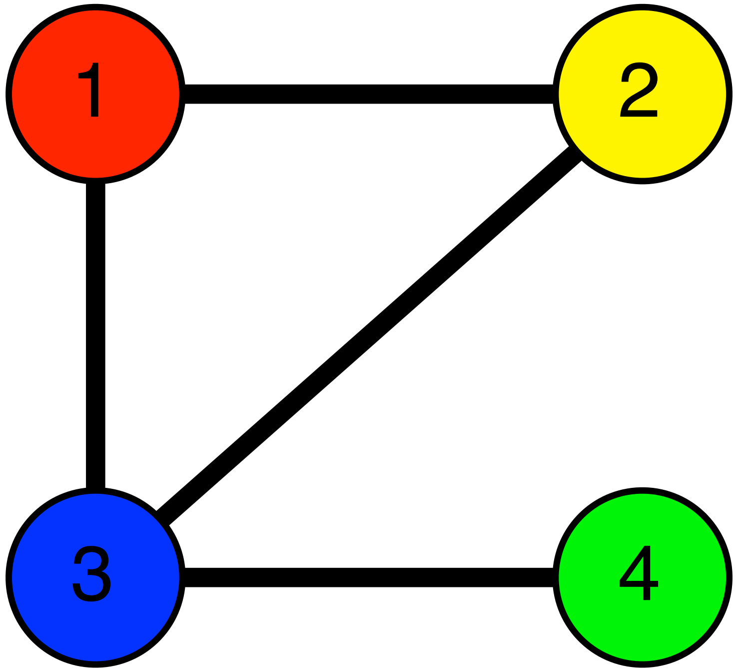
Adjacency matrix

Matrix of 1's and 0's storing which nodes are connected



Adjacency matrix

Matrix of 1's and 0's storing which nodes are connected



	1	2	3	4
1	0	1	1	0
2	1	0	1	0
3	1	1	0	1
4	0	0	1	0

Adjacency matrix

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 \end{array} = A$$

Adjacency matrix

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 \end{array} = A$$

For a network with N nodes, A is NxN

Adjacency matrix

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 \end{array} = A$$

For a network with
N nodes, A is NxN

$A_{ij} = 1$ if nodes i and j
are linked

Adjacency matrix

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 \end{array} = A$$

For a network with *N* nodes, *A* is *N*x*N*

A_{ij} = 1 if nodes *i* and *j* are linked

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_i k_i$$

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_i k_i = \sum_i \sum_j A_{ij}$$

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_i k_i = \sum_i \sum_j A_{ij} = \text{twice the number of links}$$

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_i k_i = \sum_i \sum_j A_{ij} = \text{twice the number of links}$$

M links $\longrightarrow$ $M = \frac{1}{2} \sum_{i,j} A_{ij}$

$$\sum_j A_{ij} = \text{number of neighbors of } i = k_i$$

Degree of node *i*

$$\sum_i k_i = \sum_i \sum_j A_{ij} = \text{twice the number of links}$$

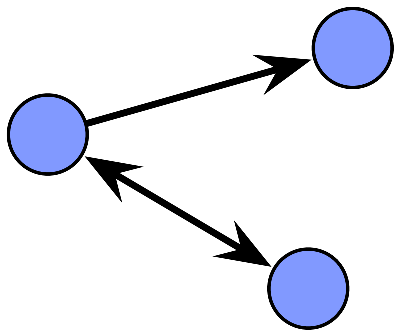
M links ⟶ $M = \frac{1}{2} \sum_{i,j} A_{ij}$

N nodes

Generalizations

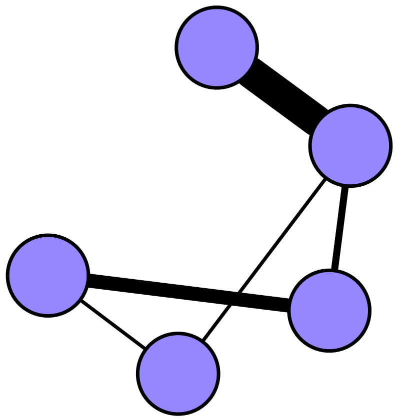
A does not have to be **symmetric**

$A_{ij} \neq A_{ji} \longrightarrow$ **directed** network



Elements of *A* don't have to be 1's and 0's

weighted network



(w_{ij})

Degree **distribution**

Degree–(perhaps) most fundamental property of a node

Degree **distribution**

Degree–(perhaps) most fundamental property of a node

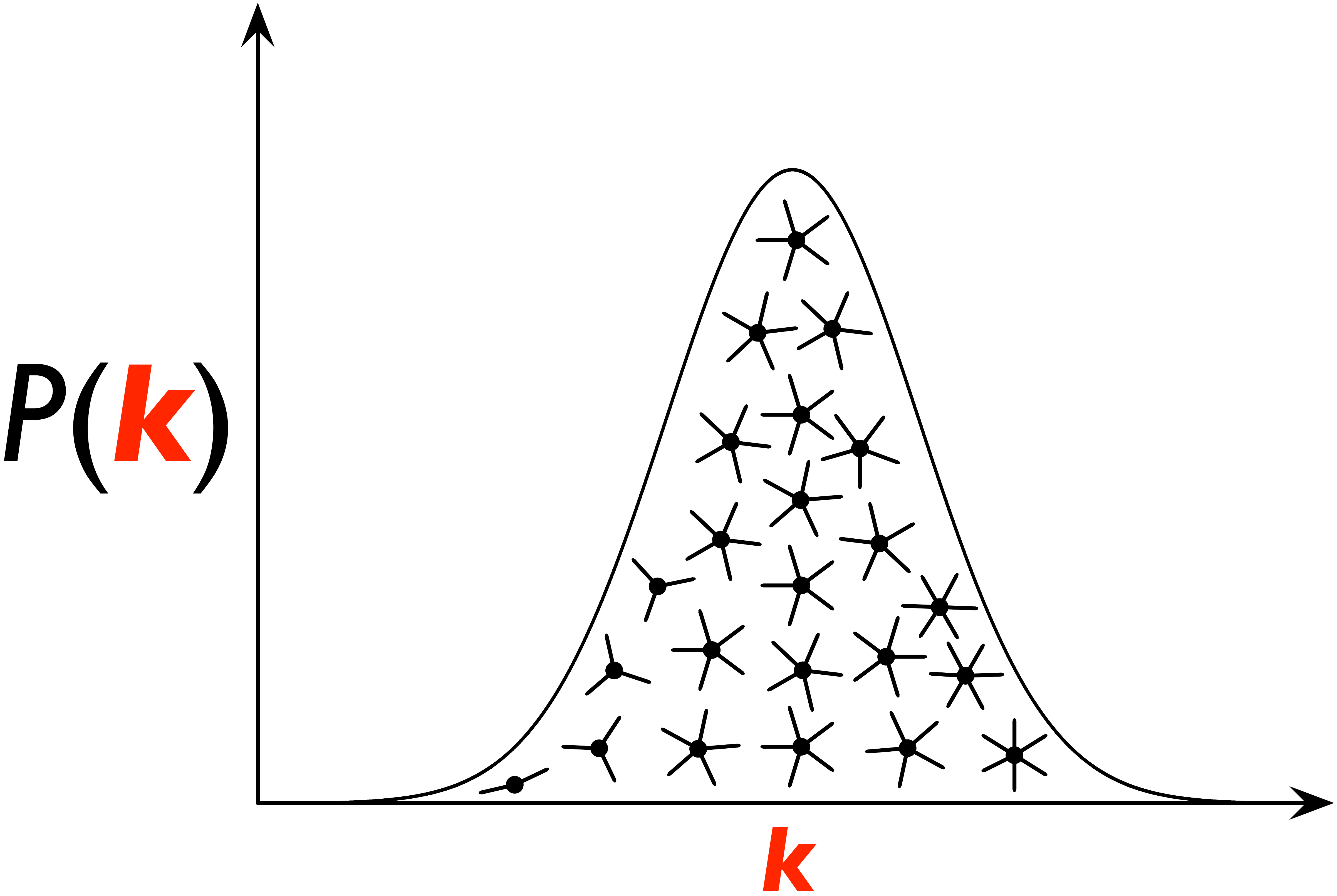
What happens if I ask a random node, "what's your degree?"

Degree **distribution**

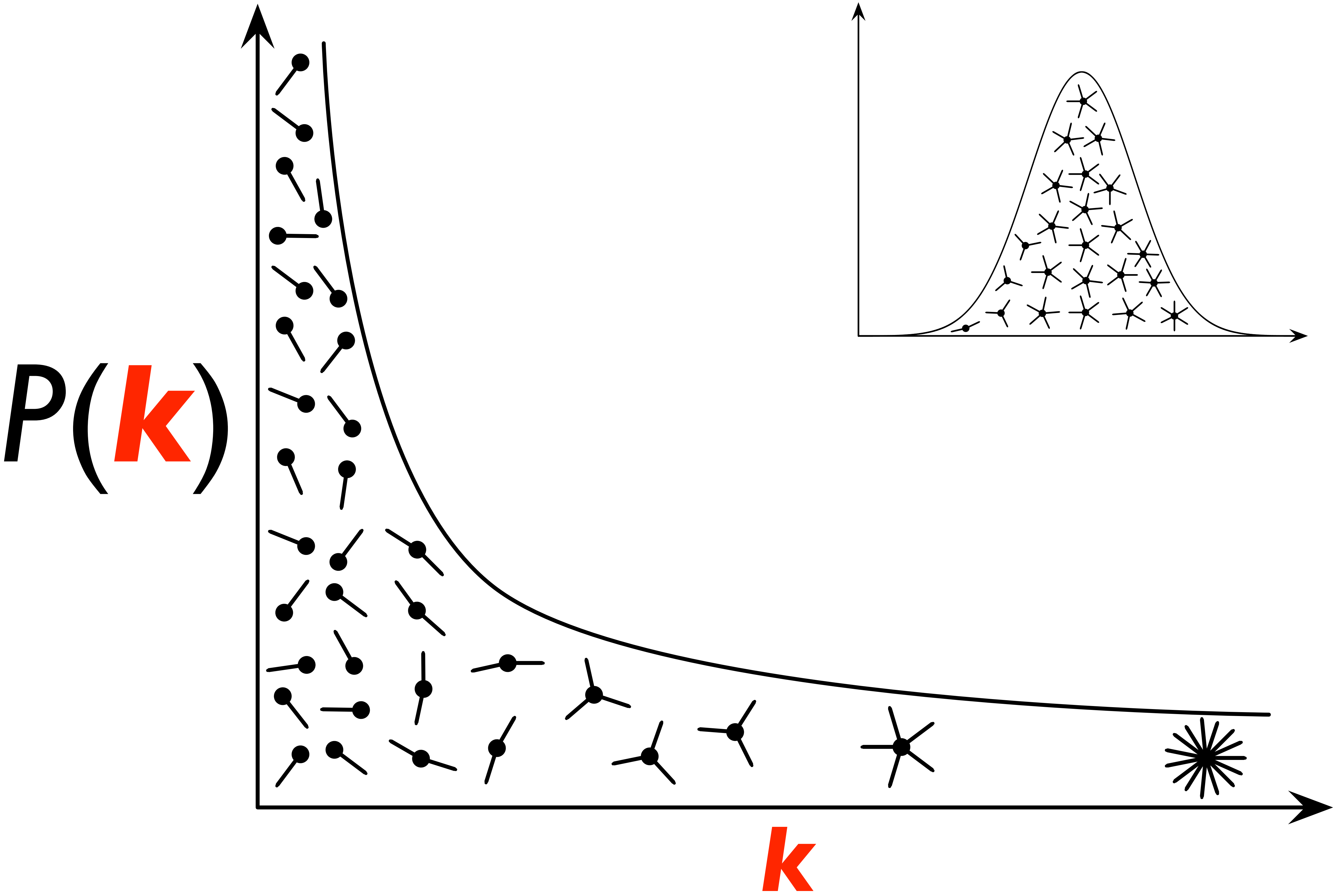
$P(k)$

Probability that a random node has degree k

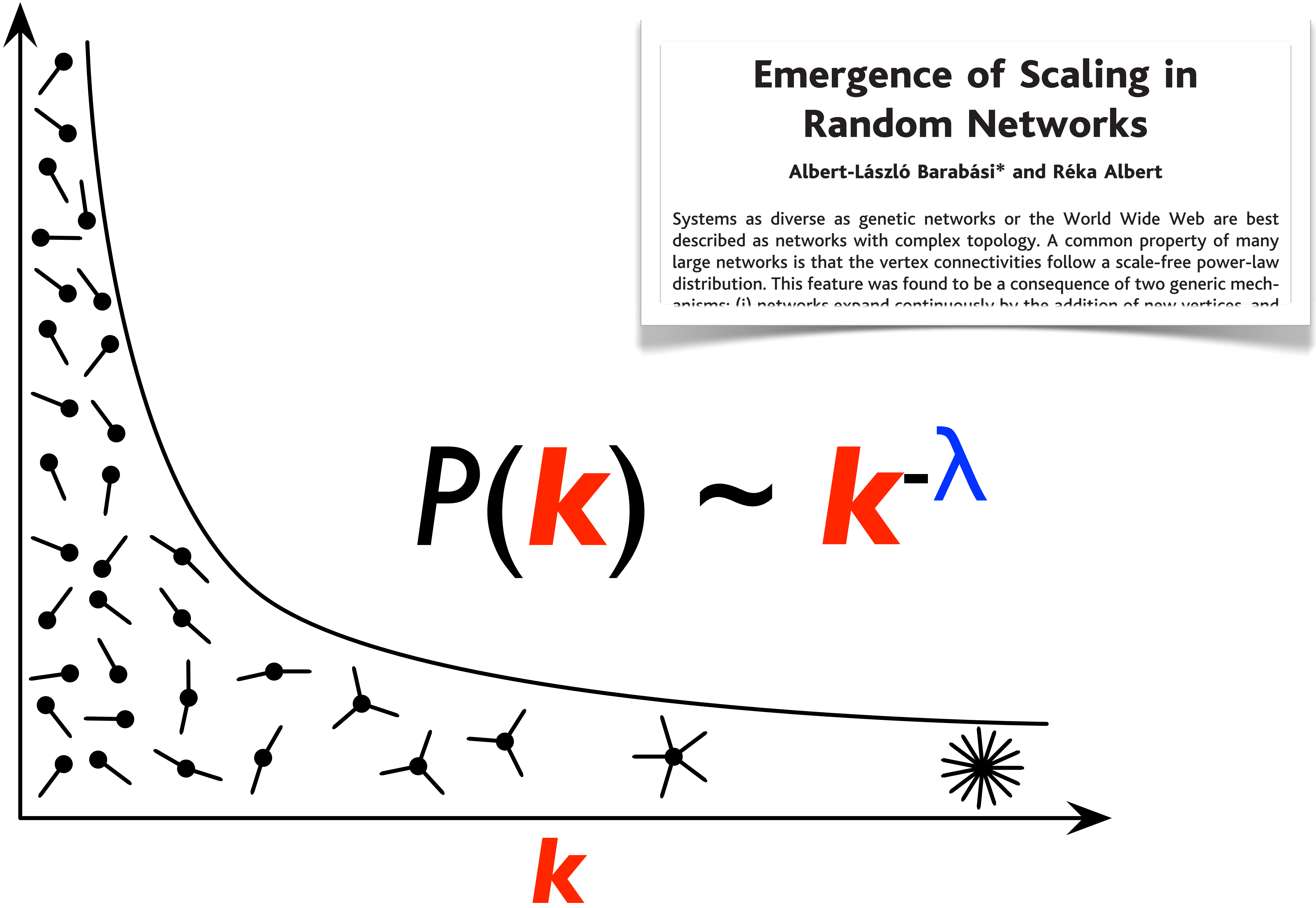
Degree **distribution**



scale-free network



scale-free network



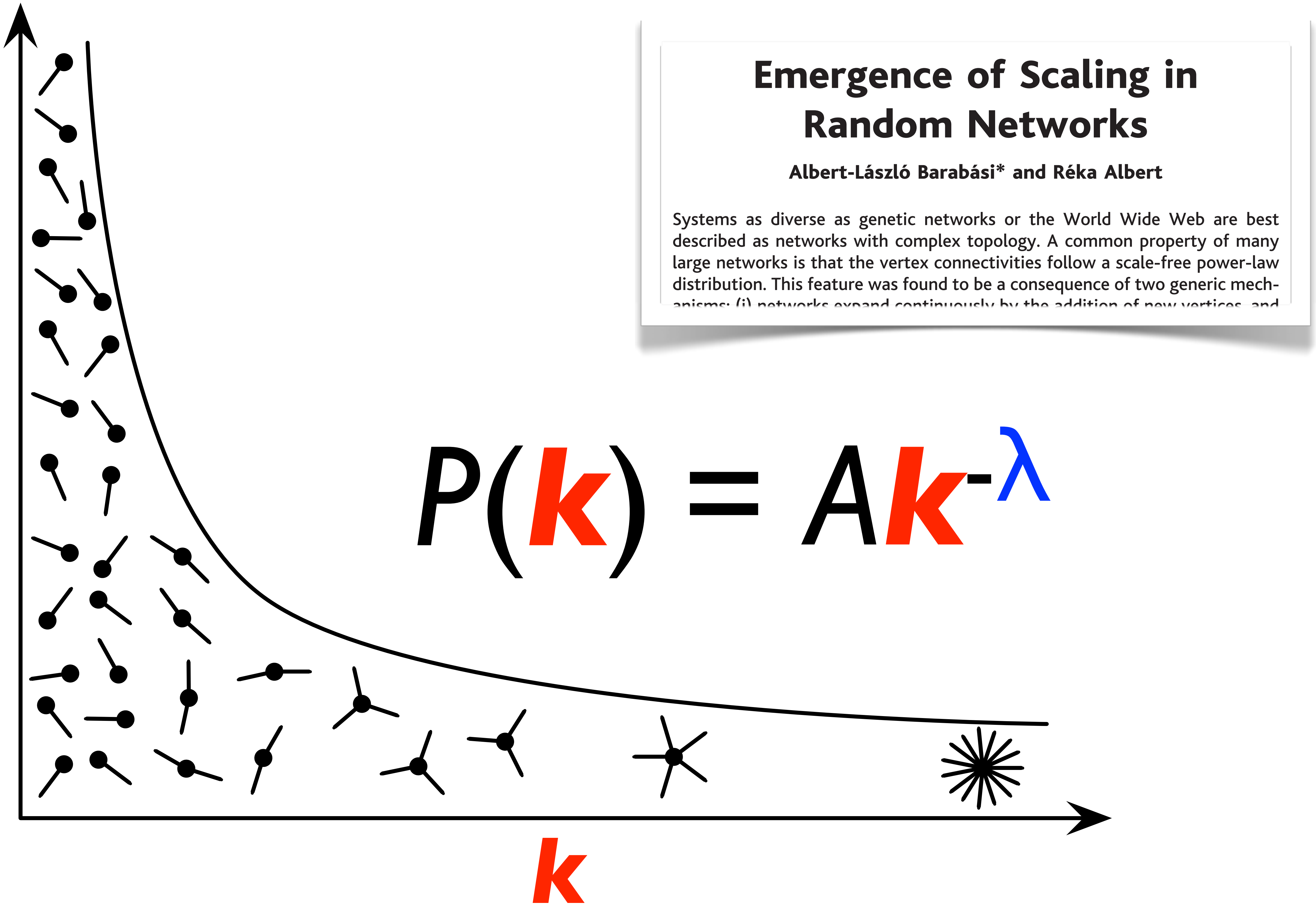
$$P(k) \sim k^{-\lambda}$$

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

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

$$P(k) = Ak^{-\lambda}$$

Aside: log-log axes

Some random data:

```
1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.73299464231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
      ⋮
```

Aside: log-log axes

Some random data:

```
1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.73299464231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
      ⋮
```

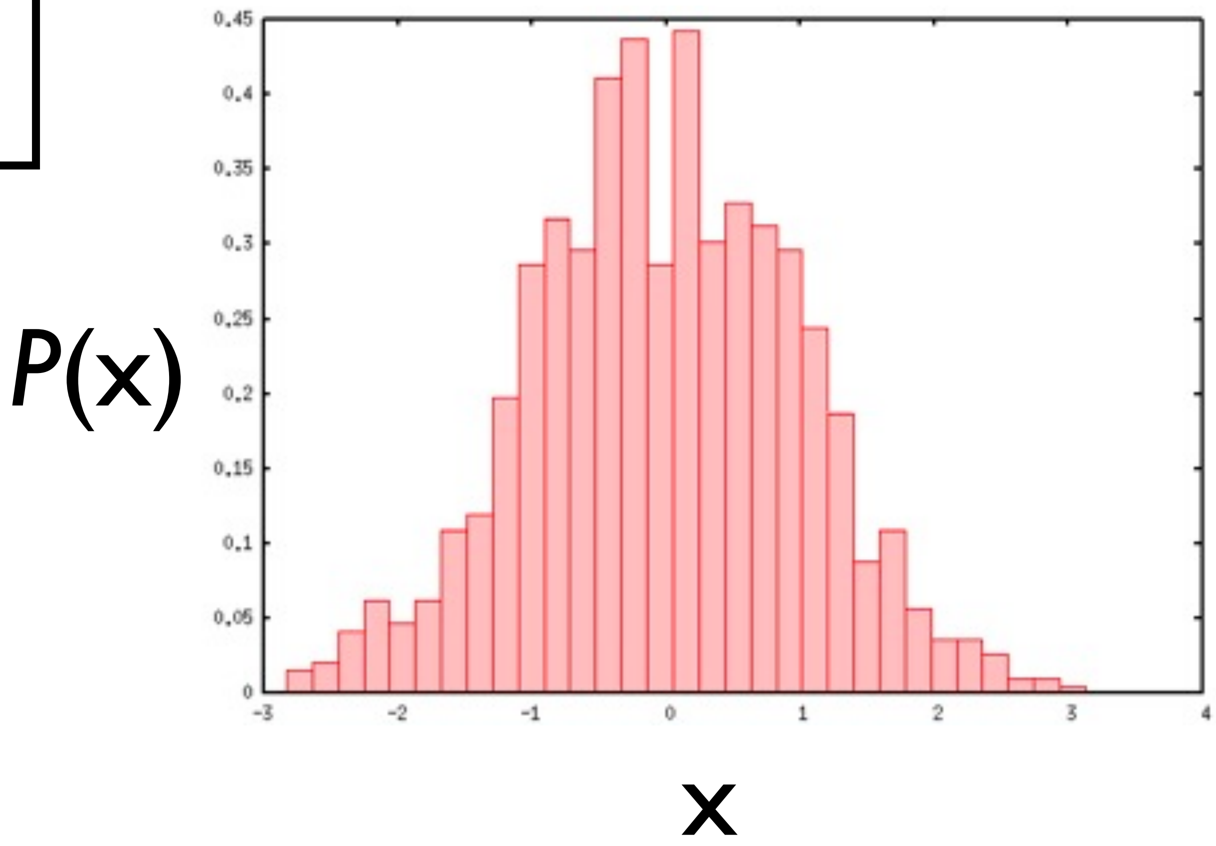
Histogram

Aside: log-log axes

Some random data:

1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.7329964231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
⋮

Histogram

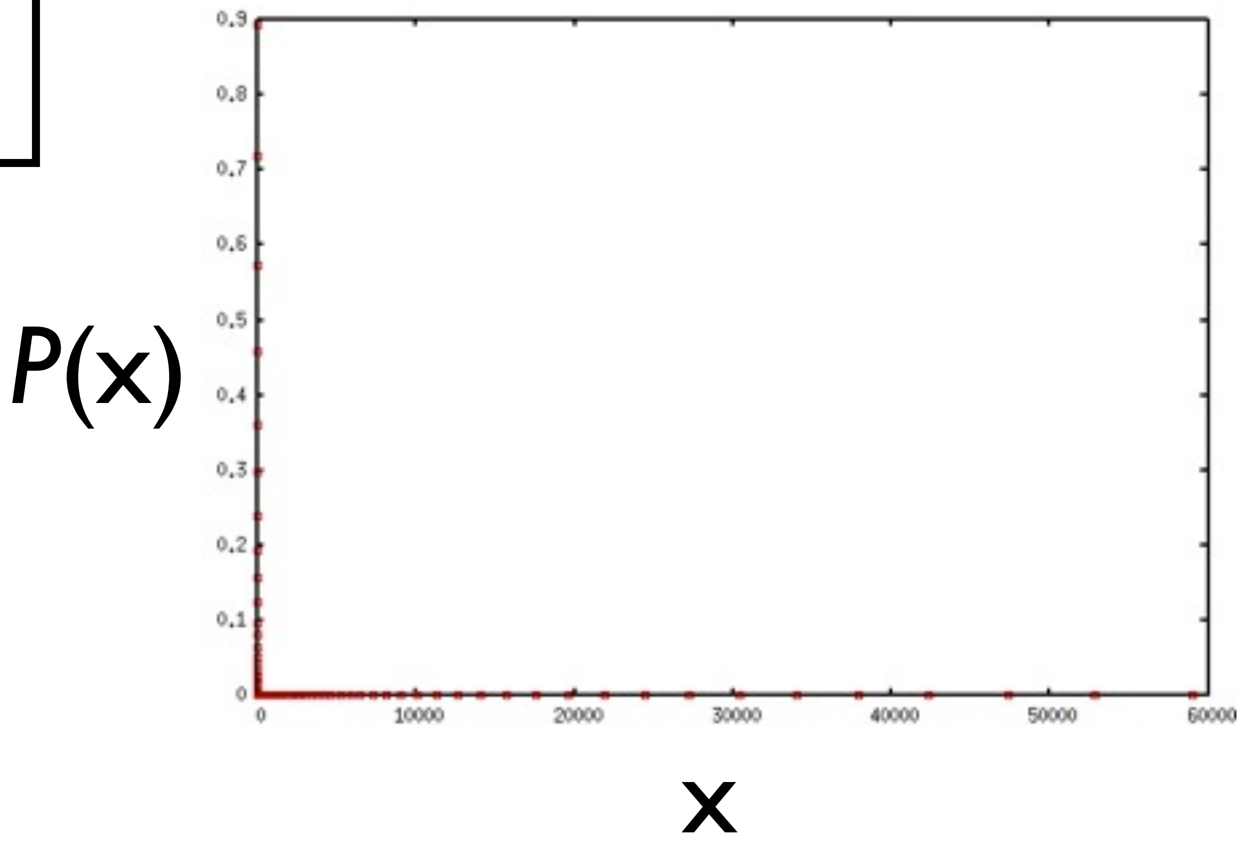


Aside: log-log axes

Some random data:

1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.7329964231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
⋮

Histogram

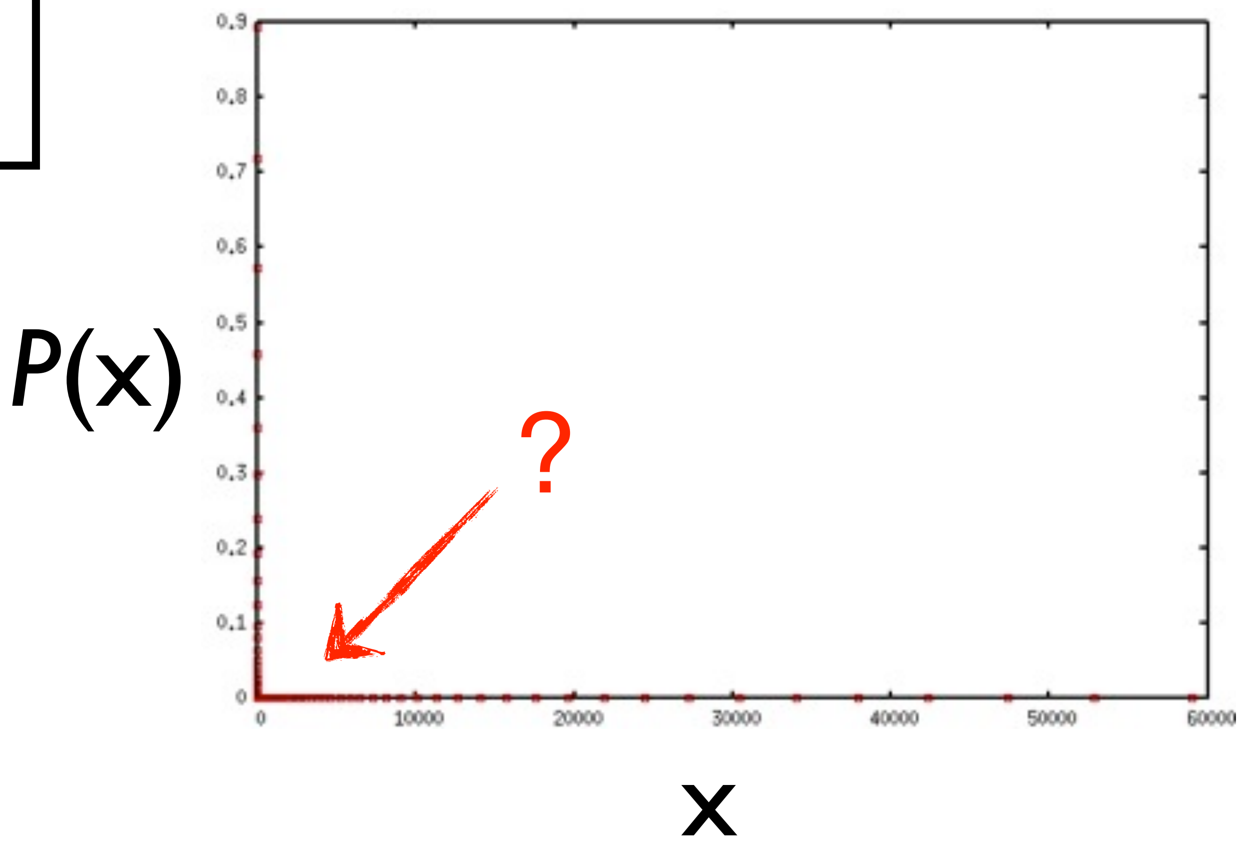


Aside: log-log axes

Some random data:

1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.73299464231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
⋮

Histogram

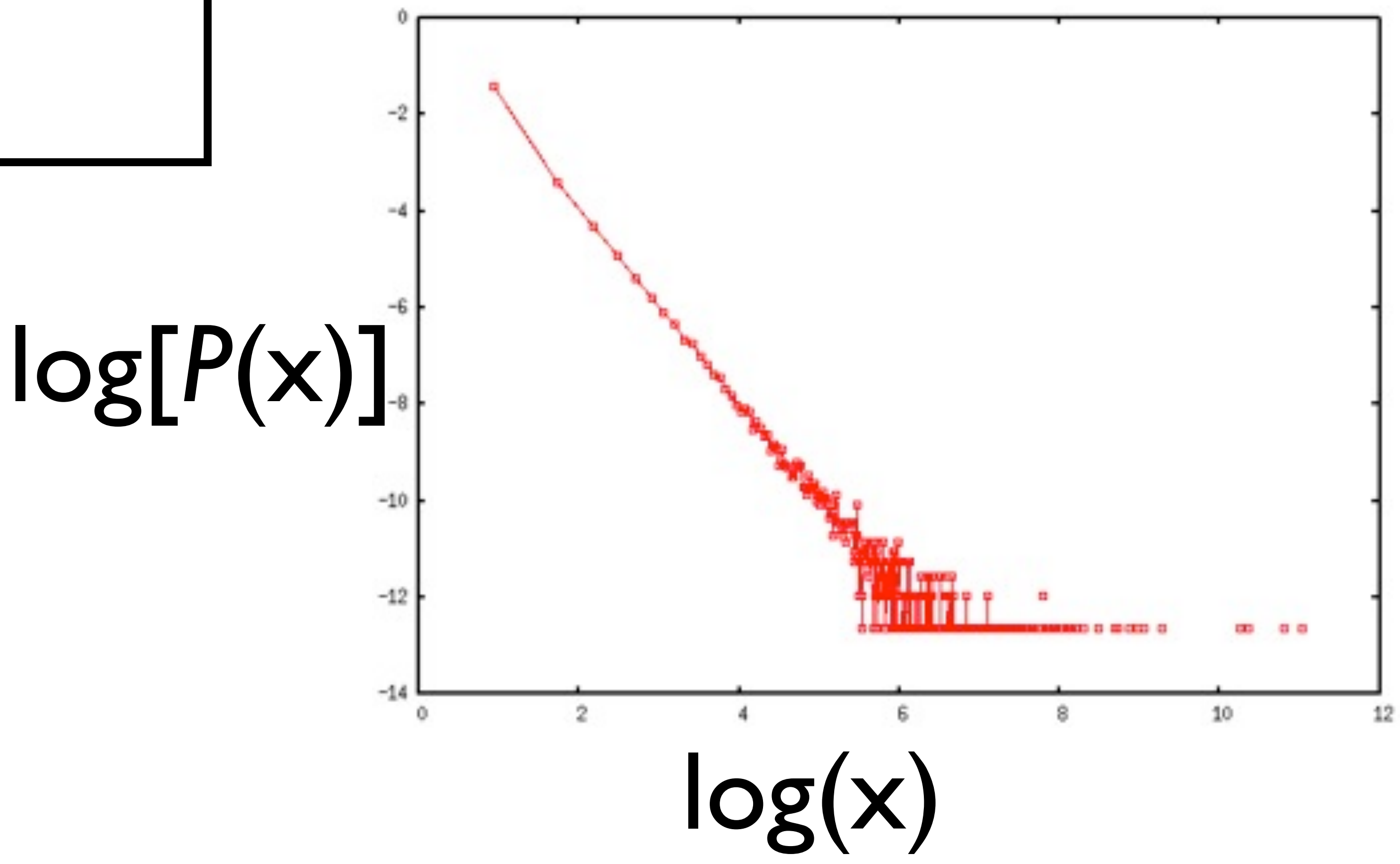


Aside: log-log axes

Some random data:

```
1.29766326467
1.27551208591
3.02324018927
1.57103746721
9.83447258547
1.7329964231
1.7323287015
5.38670711152
1.59810623031
1.49517306783
      ⋮
```

Histogram

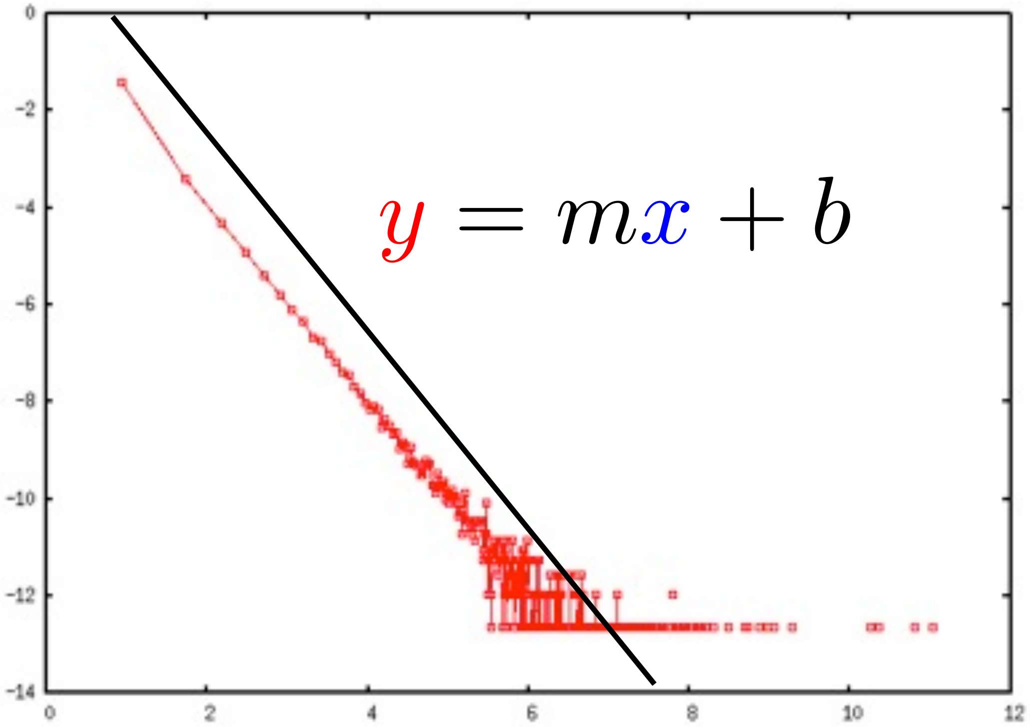


Aside: log-log axes

power law

Why?

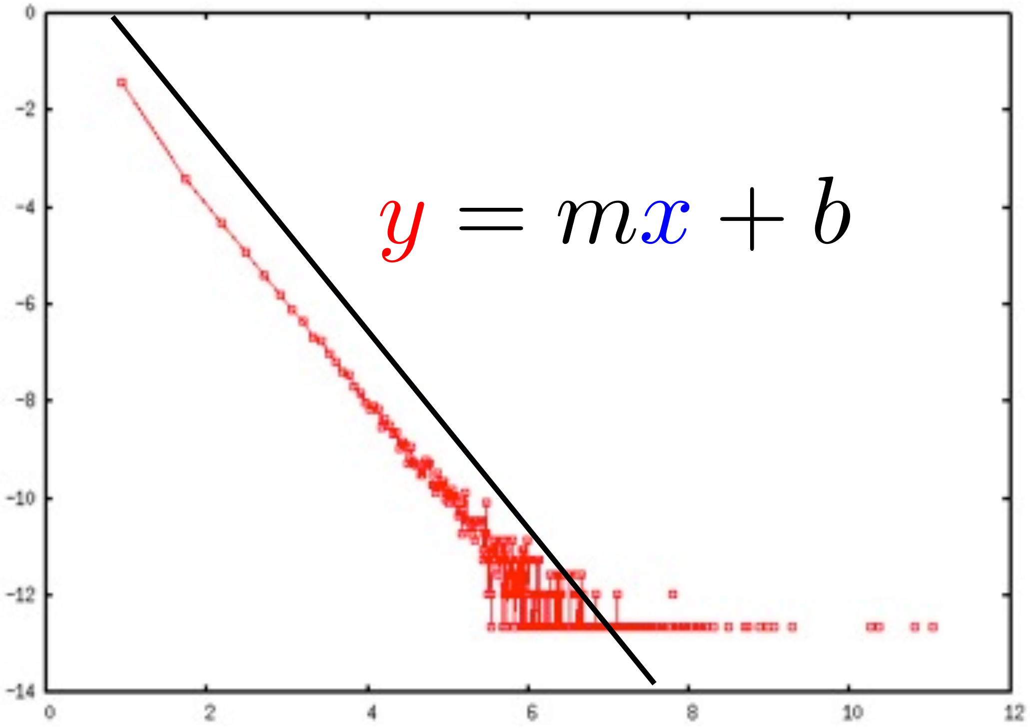
$$y = Ax^b$$



$$y = mx + b$$

Aside: log-log axes

Why?



$$y = mx + b$$

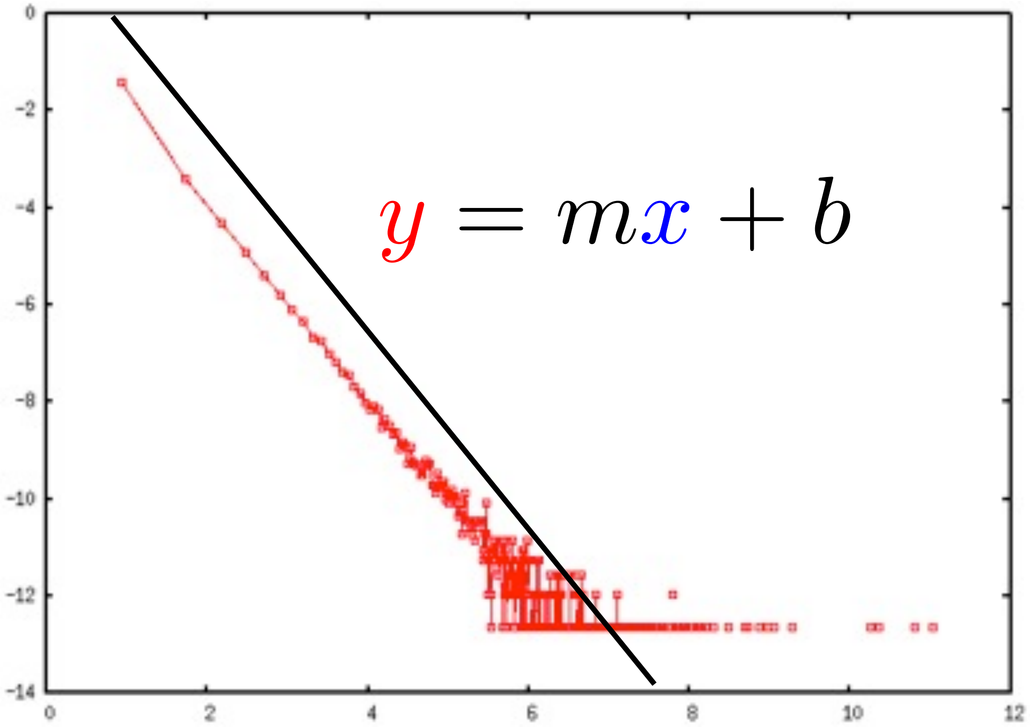
power law

$$y = Ax^b$$

$$\log(y) = \log(Ax^b)$$

Aside: log-log axes

Why?



$$y = mx + b$$

power law

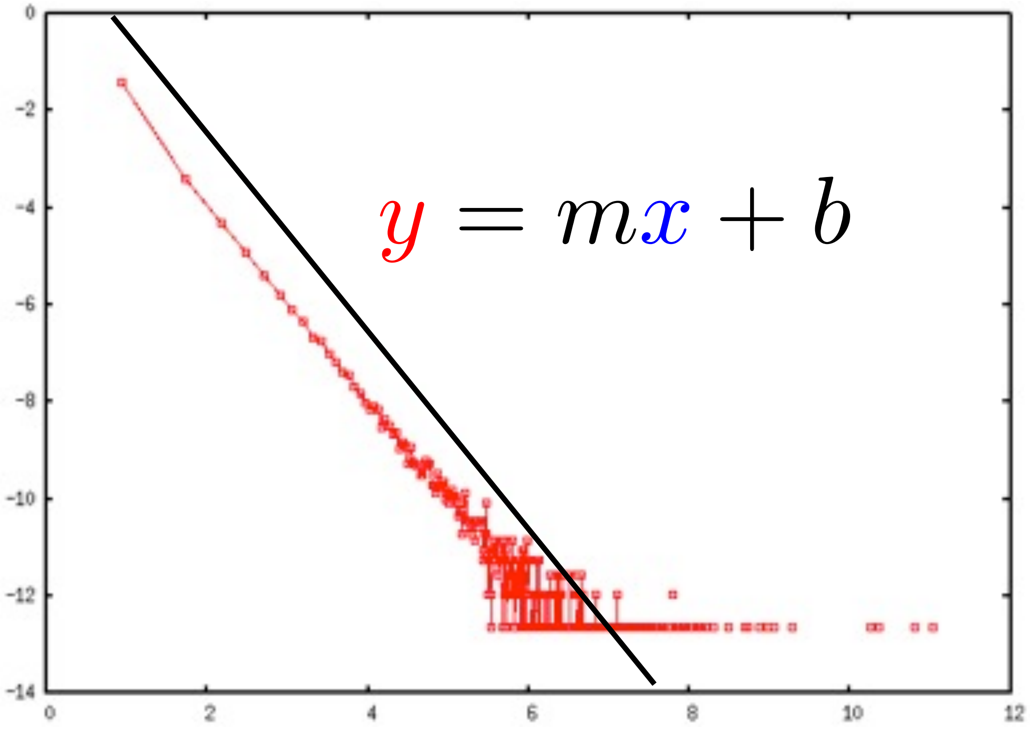
$$y = Ax^b$$

$$\log(y) = \log(Ax^b)$$

$$\log(y) = \log(A) + b\log(x)$$

Aside: log-log axes

Why?



$$y = mx + b$$

power law

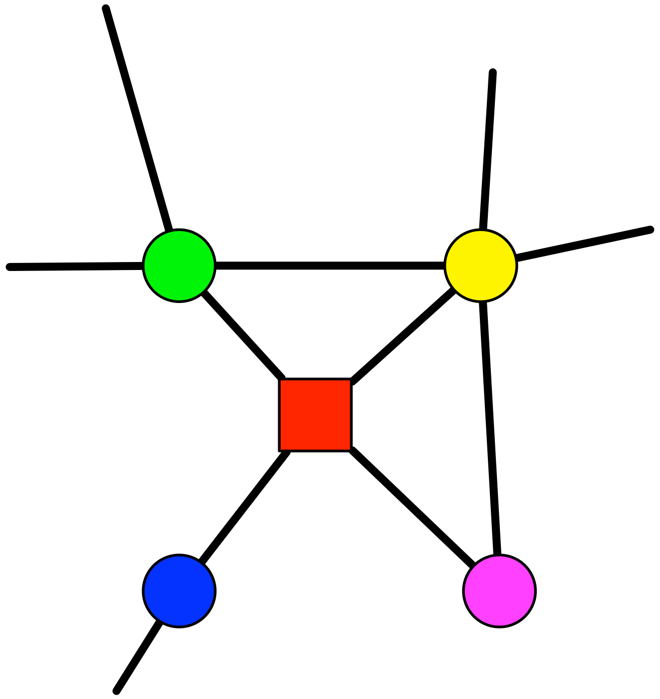
$$y = Ax^b$$

$$\log(y) = \log(Ax^b)$$

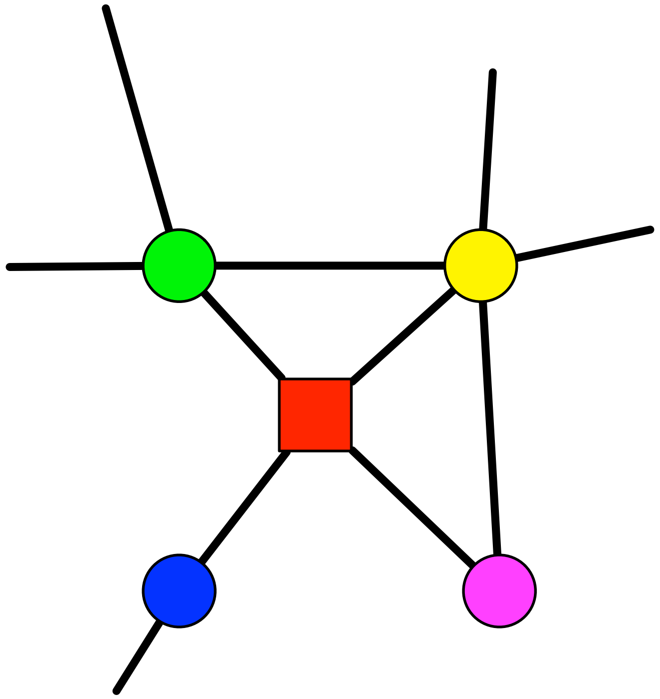
$$\begin{tikzcd}
\underline{\log(y)} \arrow[dr] & = & \underline{\log(A)} \arrow[d] & +\, b & \underline{\log(x)} \arrow[dl] \\
& y' = & A' & +\, bx' &
\end{tikzcd}$$

What **causes** these **power laws**?

Neighborhoods



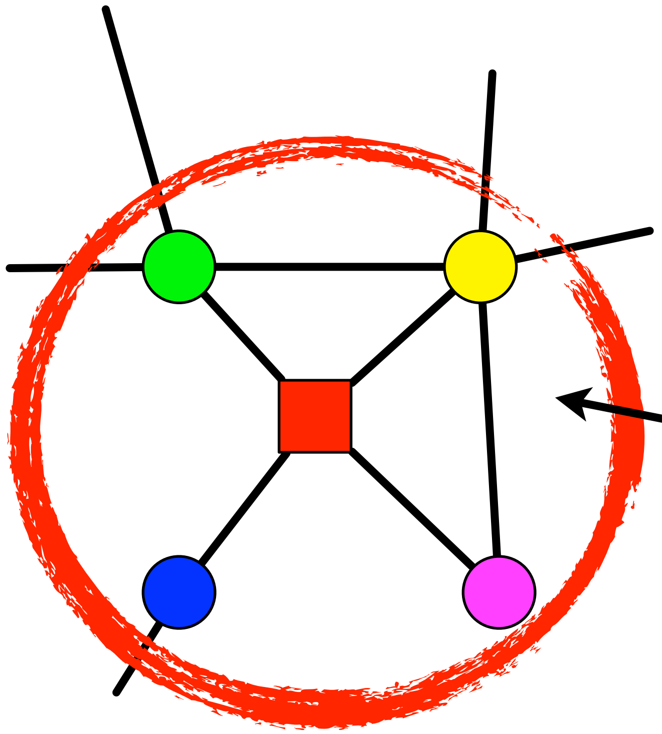
Neighborhoods



Neighbors(🟥) ={ 🔵 🟢 🟣 🟡 }

Neighborhoods

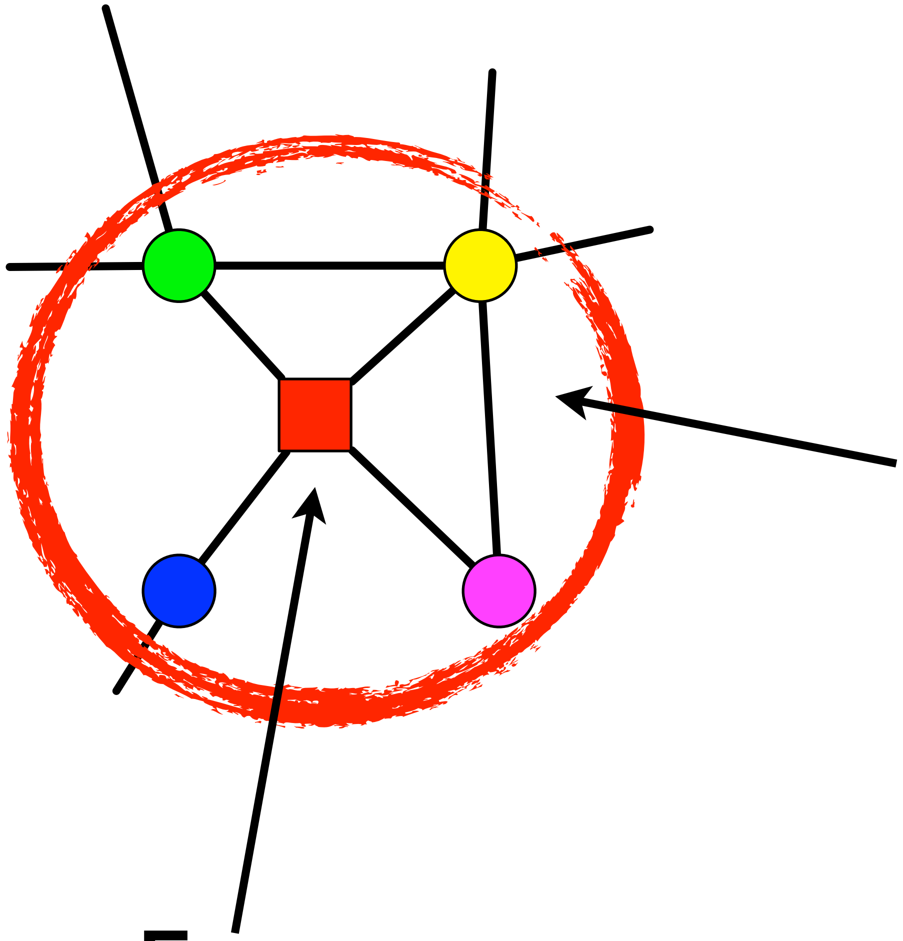
Neighbors(🟥) ={ 🔵 🟢 🟣 🟡 }



Neighborhood

Neighborhoods

Neighbors(🟥) ={ 🔵 🟢 🟣 🟡 }

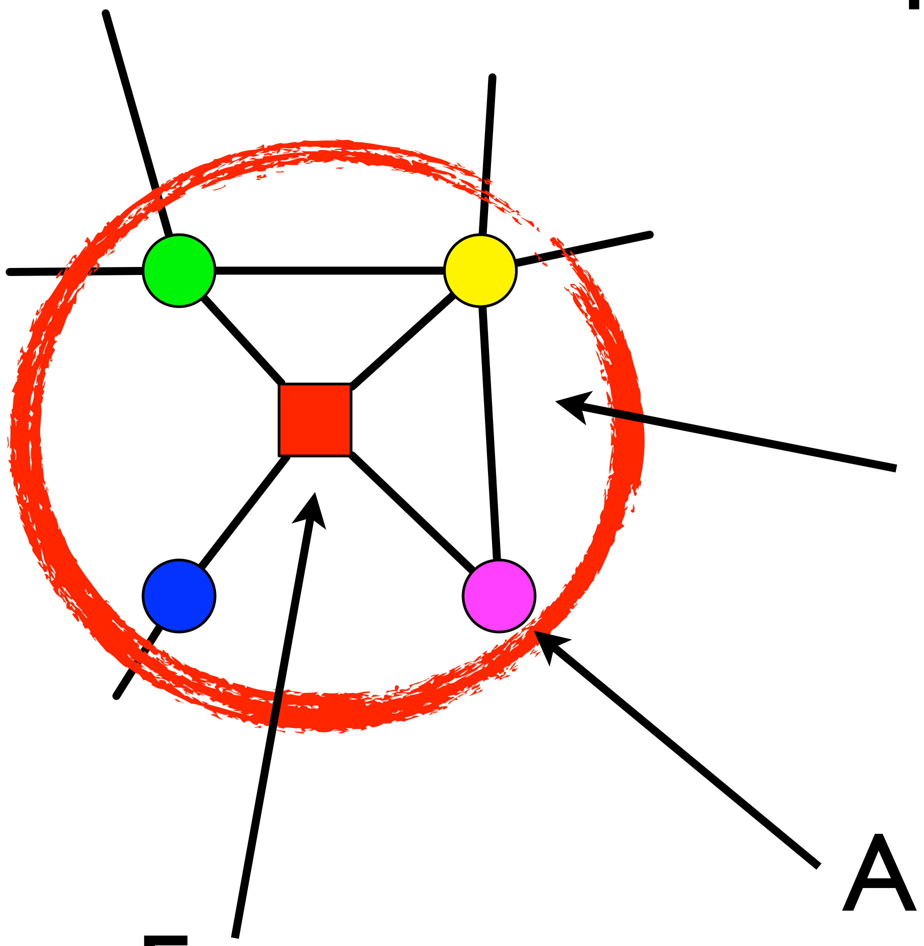


Neighborhood

Ego

Neighborhoods

Neighbors(🟥) ={ 🔵 🟢 🟣 🟡 }

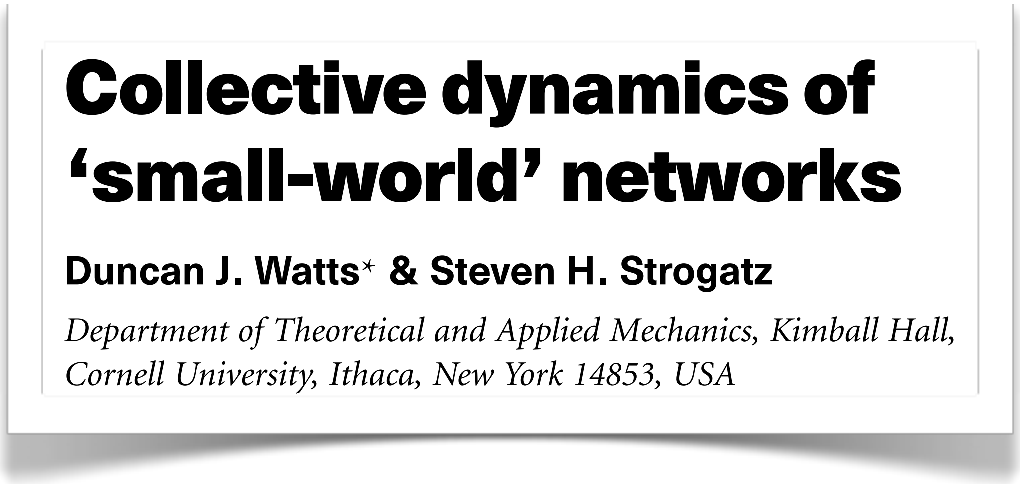


Neighborhood

Alter

Ego

Clustering coefficient



Collective dynamics of 'small-world' networks

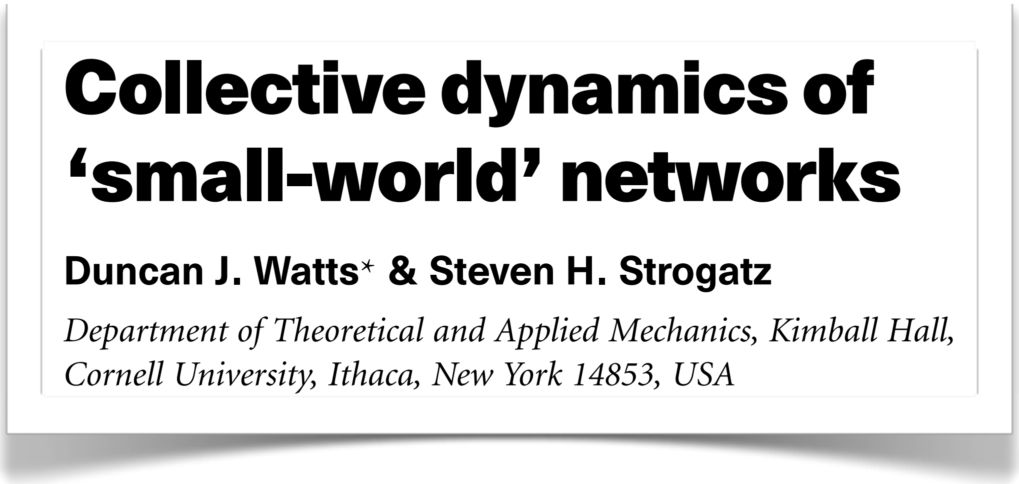
Duncan J. Watts* & Steven H. Strogatz

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

Feature of network neighborhoods

Nature, 1998

Clustering coefficient



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?

Nature, 1998

Clustering coefficient

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

How many triangles **are**
possible?

Nature, 1998

Clustering coefficient

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

How many triangles **are possible**?

$$C_i = \frac{\text{number of triangles}}{\text{maximum number of triangles}} = \frac{T_i}{\frac{k_i(k_i-1)}{2}}$$

Nature, 1998

Clustering coefficient

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

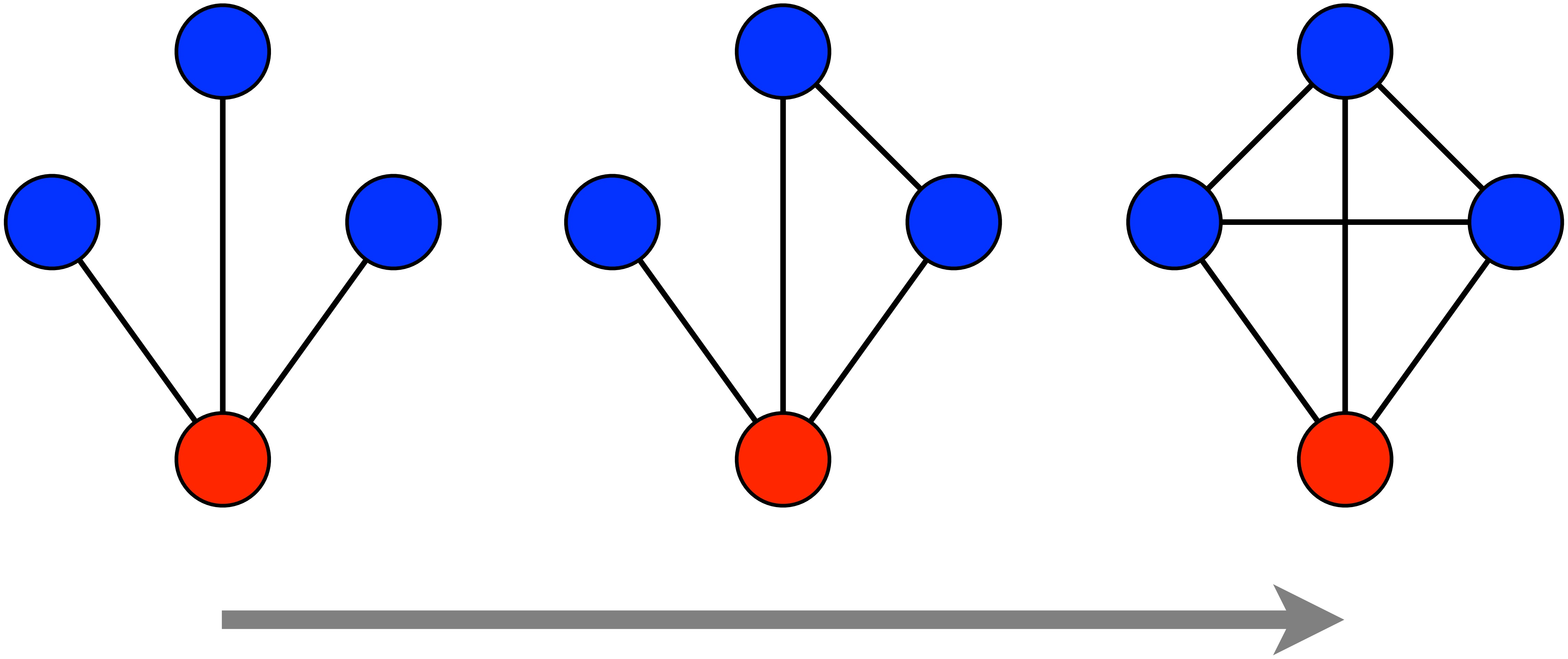
How many triangles **are possible**?

$$C_i = \frac{\text{number of triangles}}{\text{maximum number of triangles}} = \frac{T_i}{\frac{k_i(k_i-1)}{2}}$$

$$C = \frac{1}{N}\sum_i C_i$$

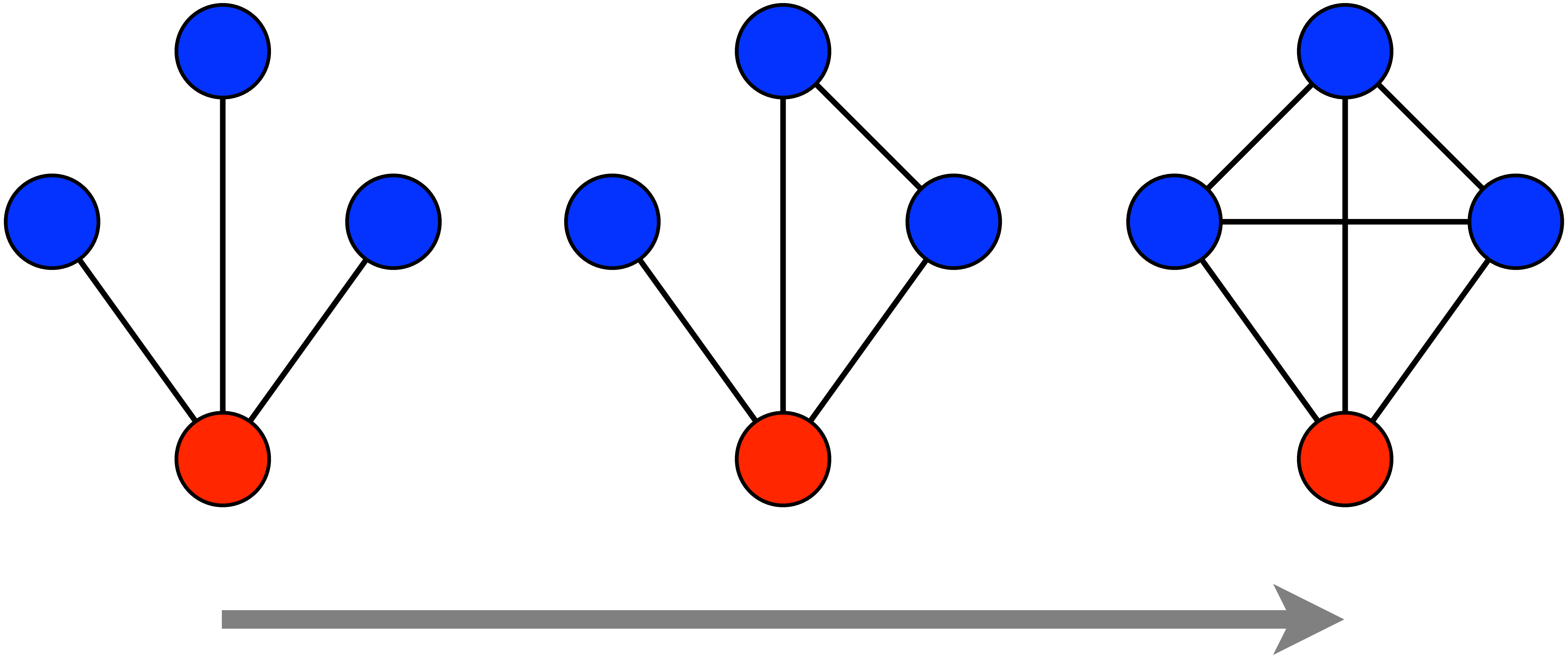
Nature, 1998

Clustering coefficient



$C_i = 0$ $C_i = 1/3$ $C_i = 1$

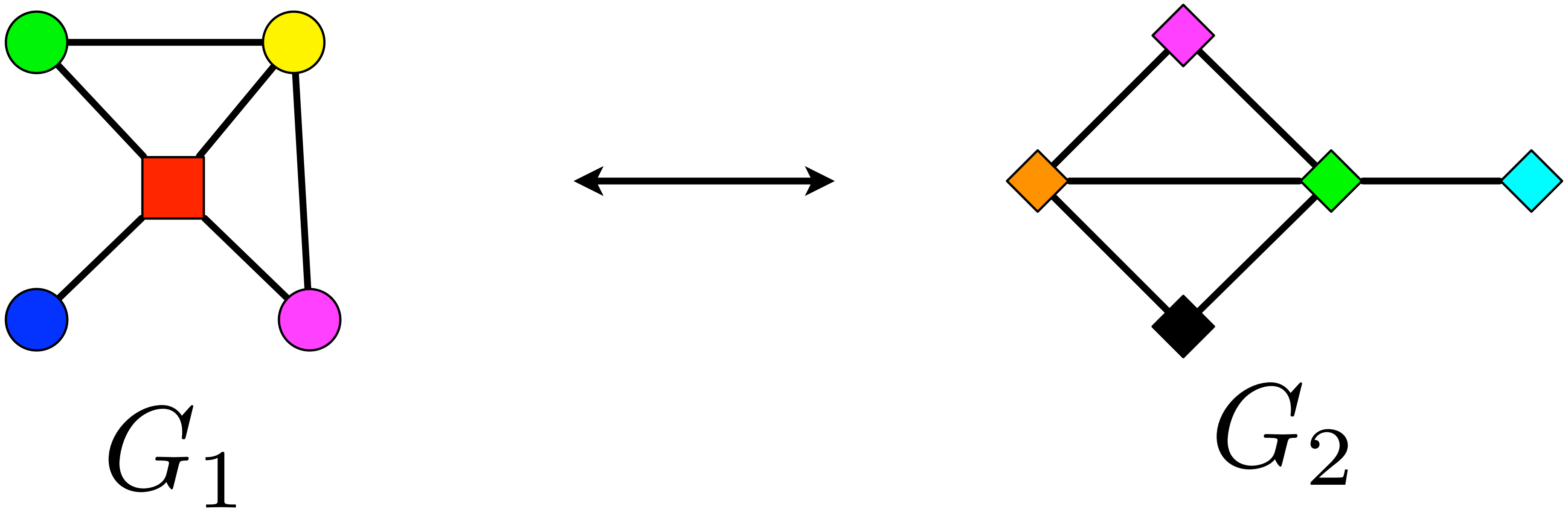
Clustering coefficient



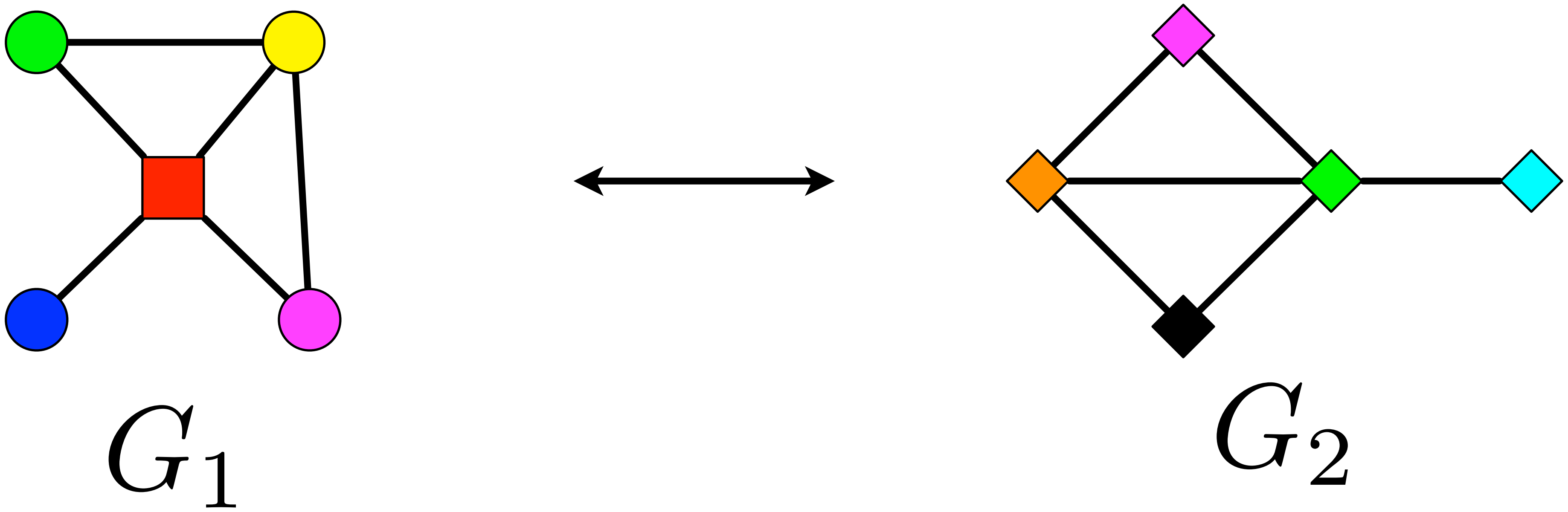
Real networks ➟ **more** **triangles** than expected!

Motifs

Motifs

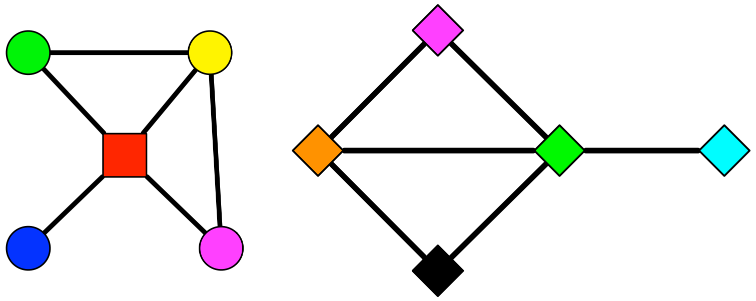


Motifs



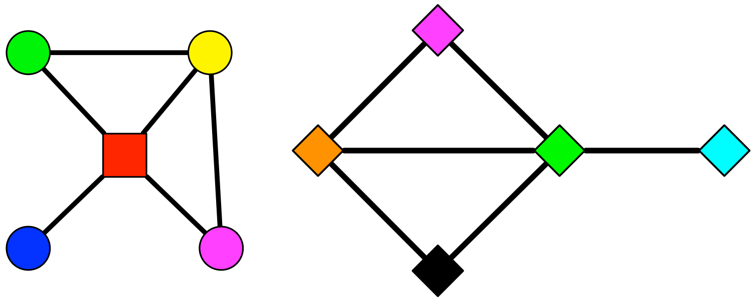
G_1 and G_2 are **isomorphic**!

Motifs

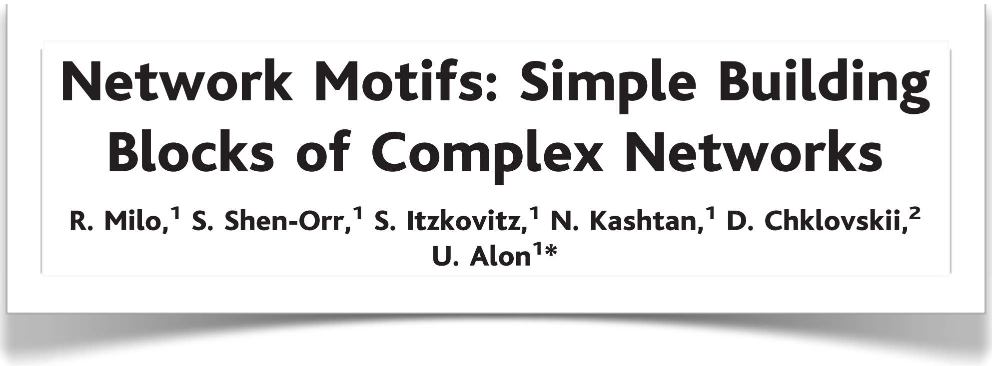


Motifs: Frequently occurring isomorphic subgraphs

Motifs



Motifs: Frequently occurring isomorphic subgraphs



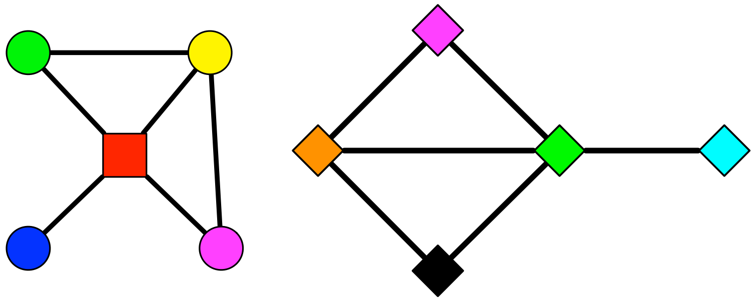
Network Motifs: Simple Building Blocks of Complex Networks

R. Milo,[1] S. Shen-Orr,[1] S. Itzkovitz,[1] N. Kashtan,[1] D. Chklovskii,[2] U. Alon[1*]

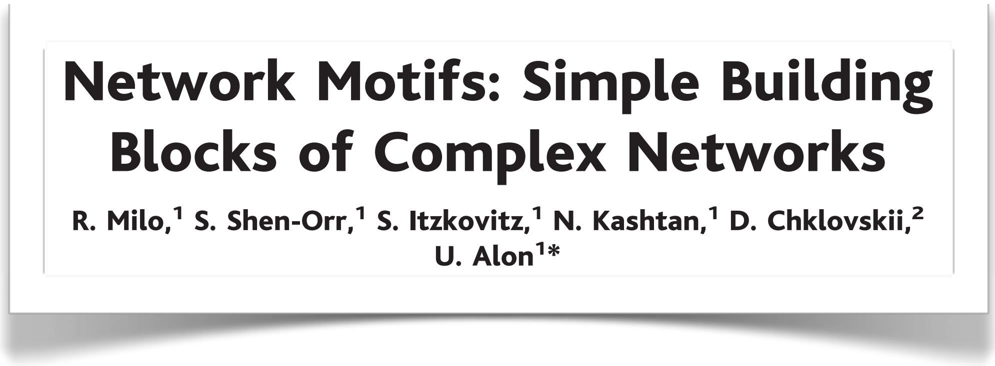
Indicate the network possesses non-trivial **structure**

Milo, et al. *Science*, 2002

Motifs



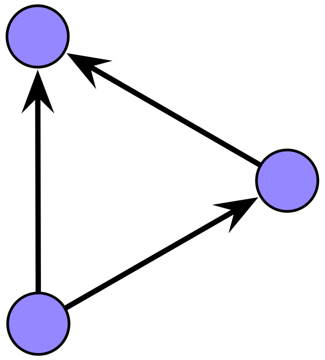
Motifs: Frequently occurring isomorphic subgraphs



Network Motifs: Simple Building Blocks of Complex Networks

R. Milo,[1] S. Shen-Orr,[1] S. Itzkovitz,[1] N. Kashtan,[1] D. Chklovskii,[2] U. Alon[1*]

Indicate the network possesses non-trivial **structure**



feedforward loop

Milo, et al. *Science*, 2002

Assortativity

How do links form in the network?

Newman, Assortative mixing in networks, *Phys Rev Lett*, 2002

Assortativity

How do links form in the network?

Do **similar** nodes connect to one another (homophily)?

Newman, Assortative mixing in networks, *Phys Rev Lett*, 2002

Assortativity

How do links form in the network?

Do **similar** nodes connect to one another (homophily)?

Do links form between **different** nodes?

Newman, Assortative mixing in networks, *Phys Rev Lett*, 2002

Assortativity

How do links form in the network?

Example: **degree**

r = assortativity coefficient

Measures correlation in degrees of linked nodes

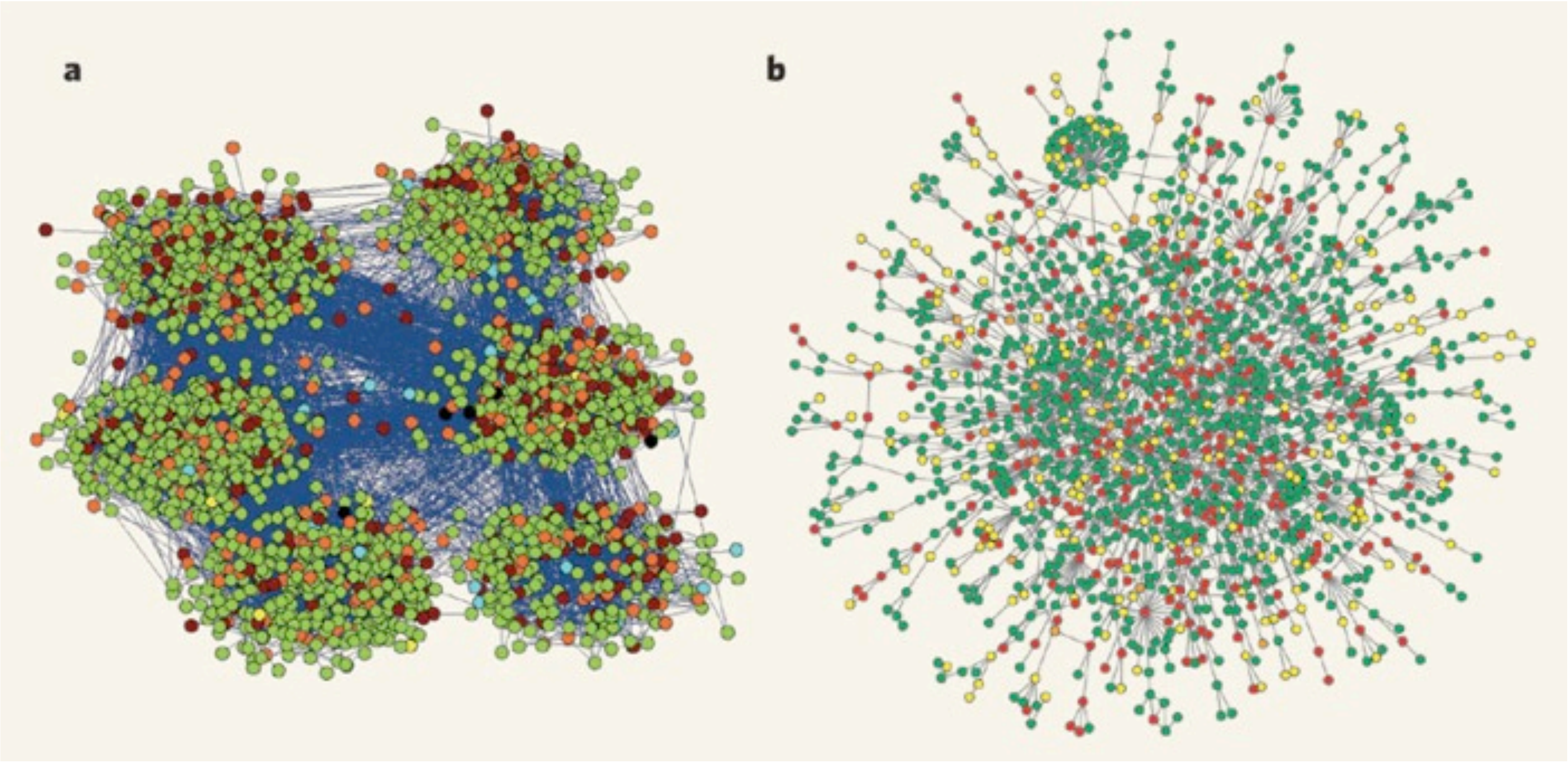
$$-1 \leq r \leq 1$$

Dissortative

Assortative

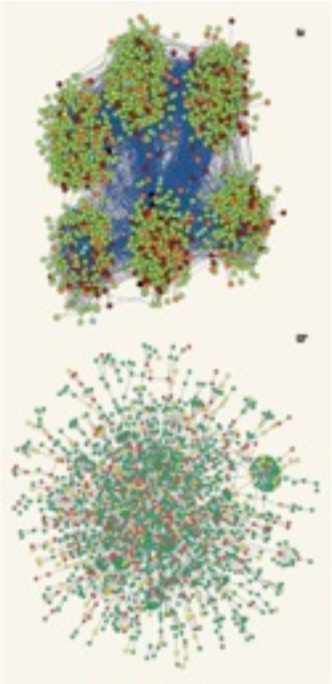
Assortativity

Assortative Dissortative



Redner, *Nature*, 2008

Assortativity



	network	n	r
real-world networks	physics coauthorship[a]	52 909	0.363
	biology coauthorship[a]	1 520 251	0.127
	mathematics coauthorship[b]	253 339	0.120
	film actor collaborations[c]	449 913	0.208
	company directors[d]	7 673	0.276
	Internet[e]	10 697	−0.189
	World-Wide Web[f]	269 504	−0.065
	protein interactions[g]	2 115	−0.156
	neural network[h]	307	−0.163
	food web[i]	92	−0.276

Newman, Assortative mixing in networks, *Phys Rev Lett*, 2002

Distances and Networks

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

Distances and Networks

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

Space lets us tell how far apart things are

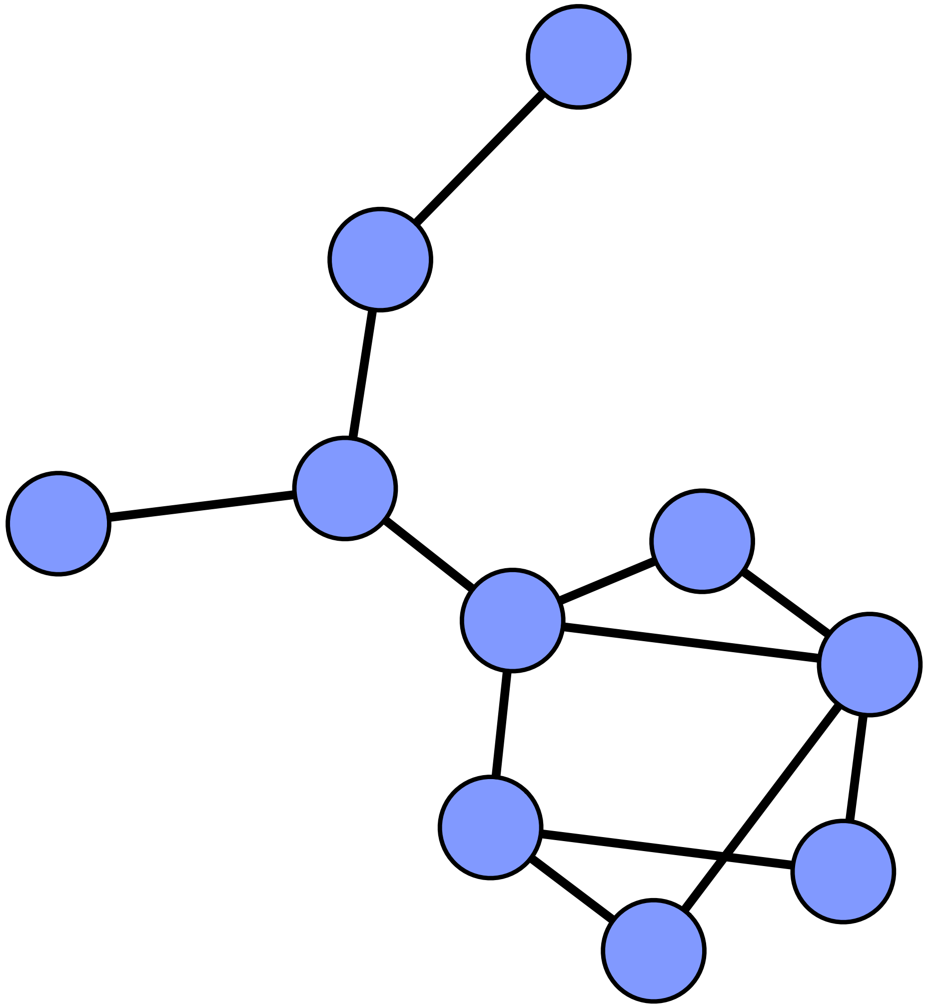
Distances and Networks

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

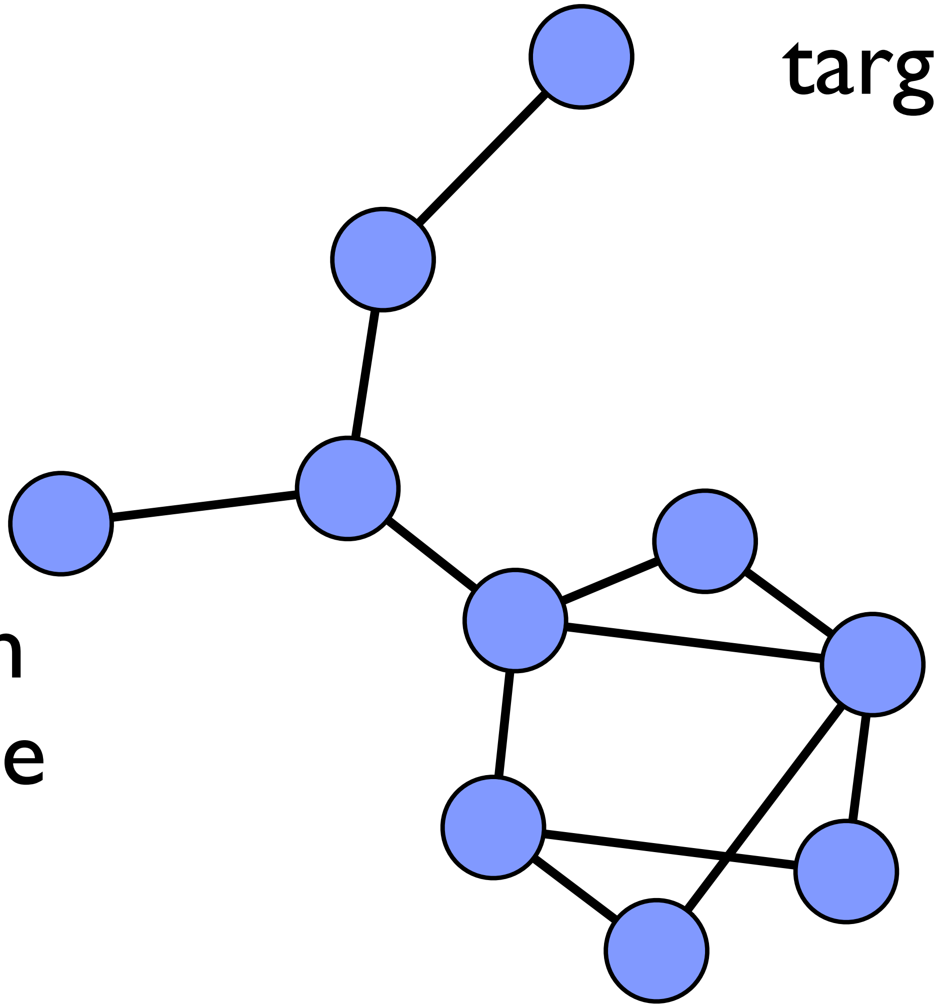
Space lets us tell how far apart things are

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

Paths



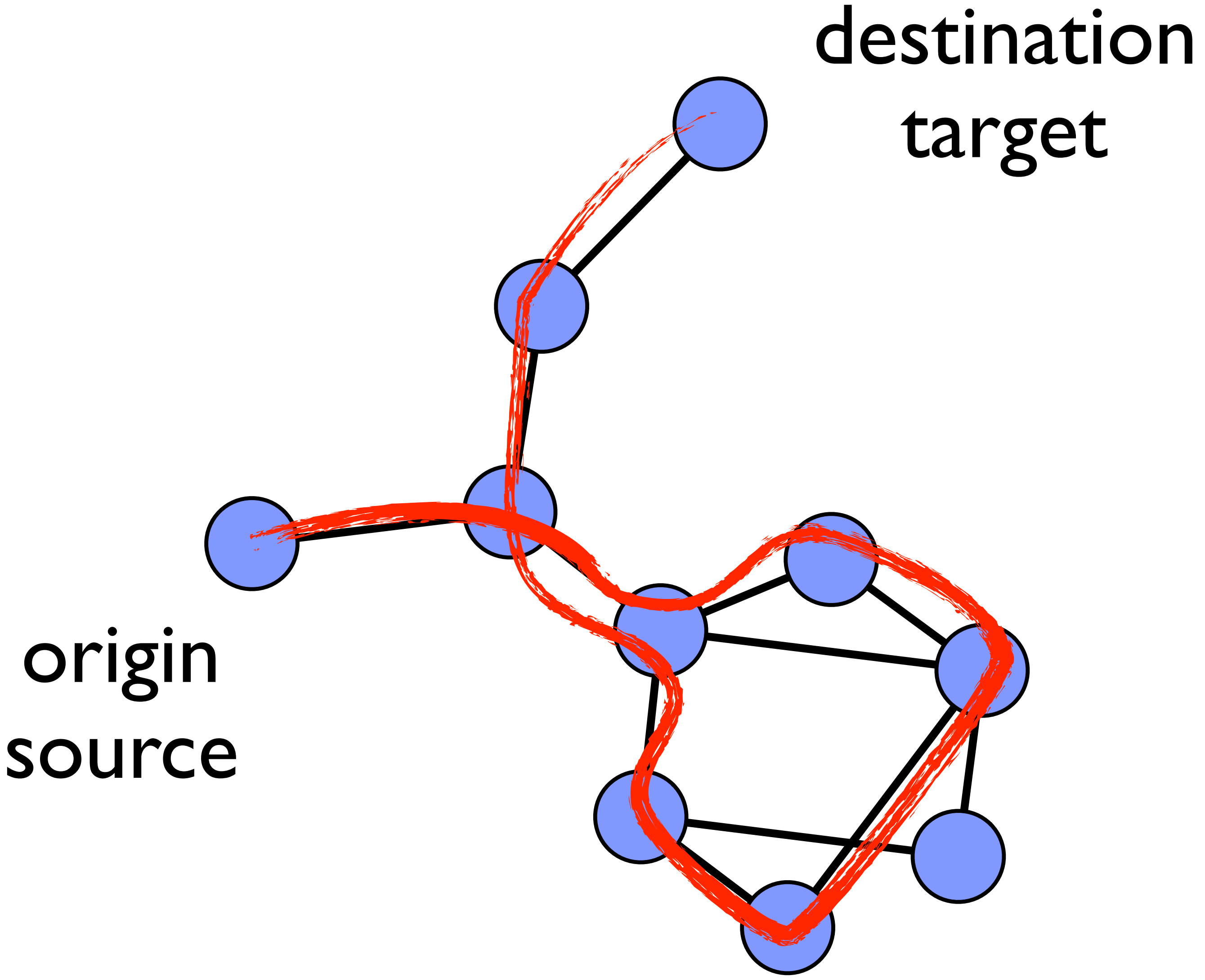
Paths



destination
target

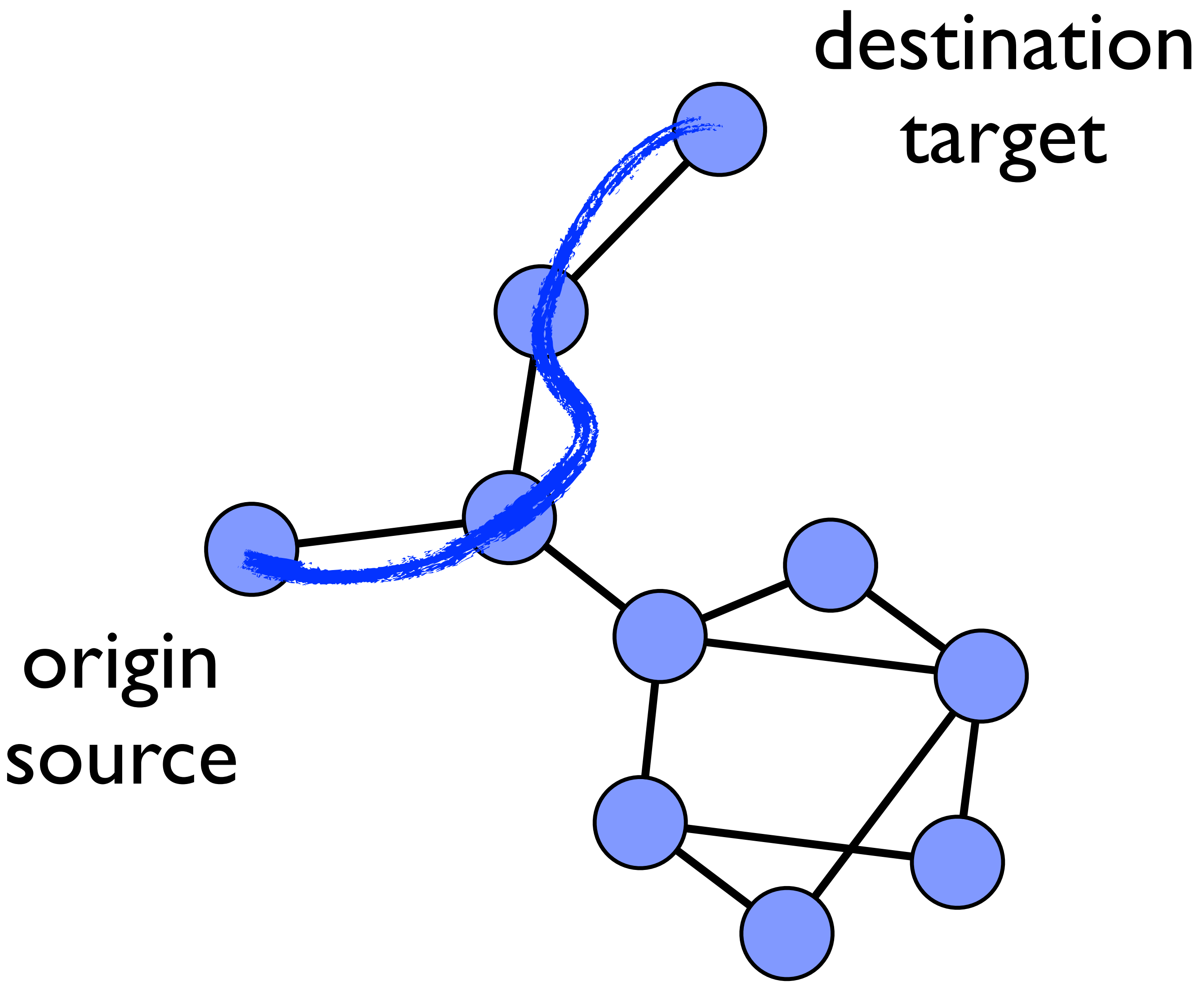
origin
source

Paths



many paths exist

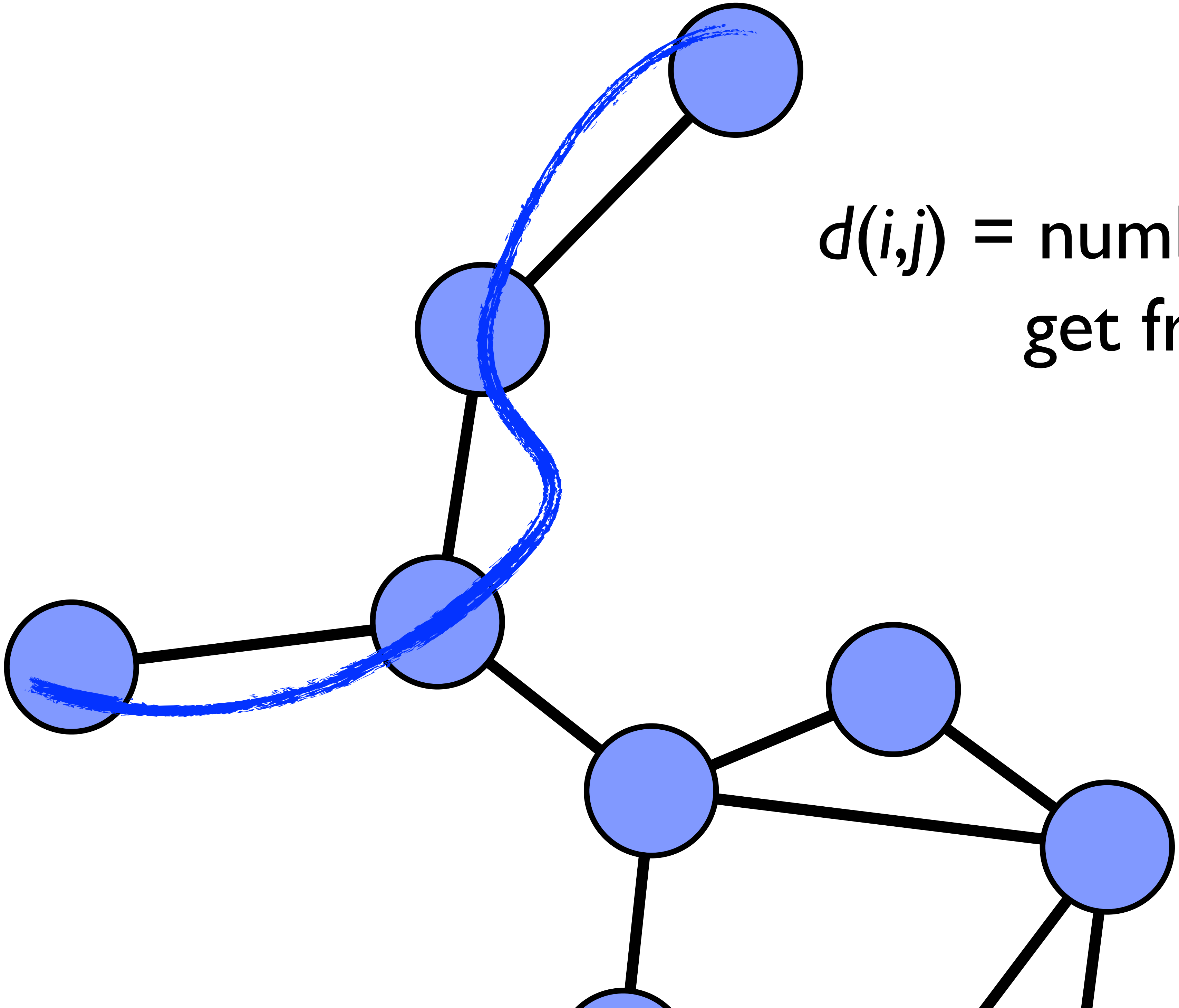
Paths



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

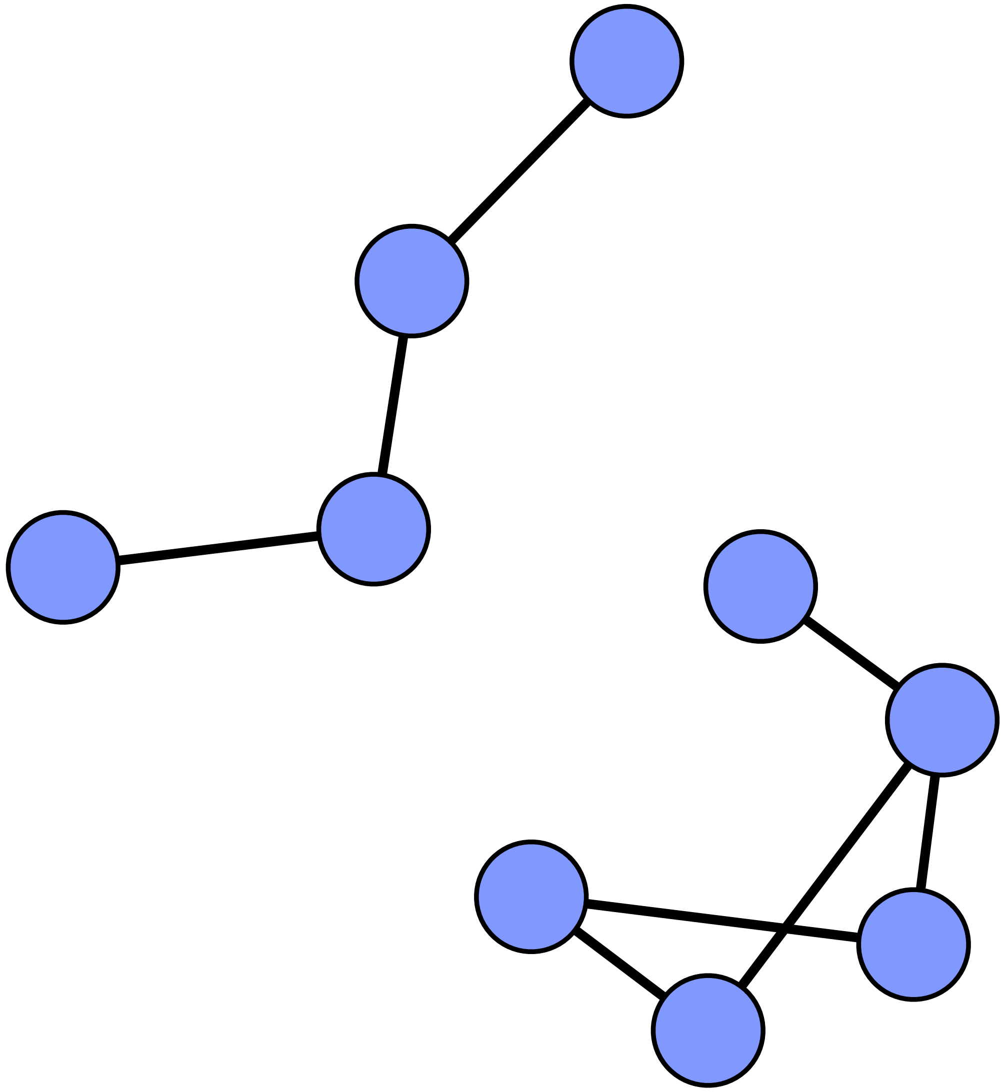
Eccentricity of a node

Longest shortest path starting from that node

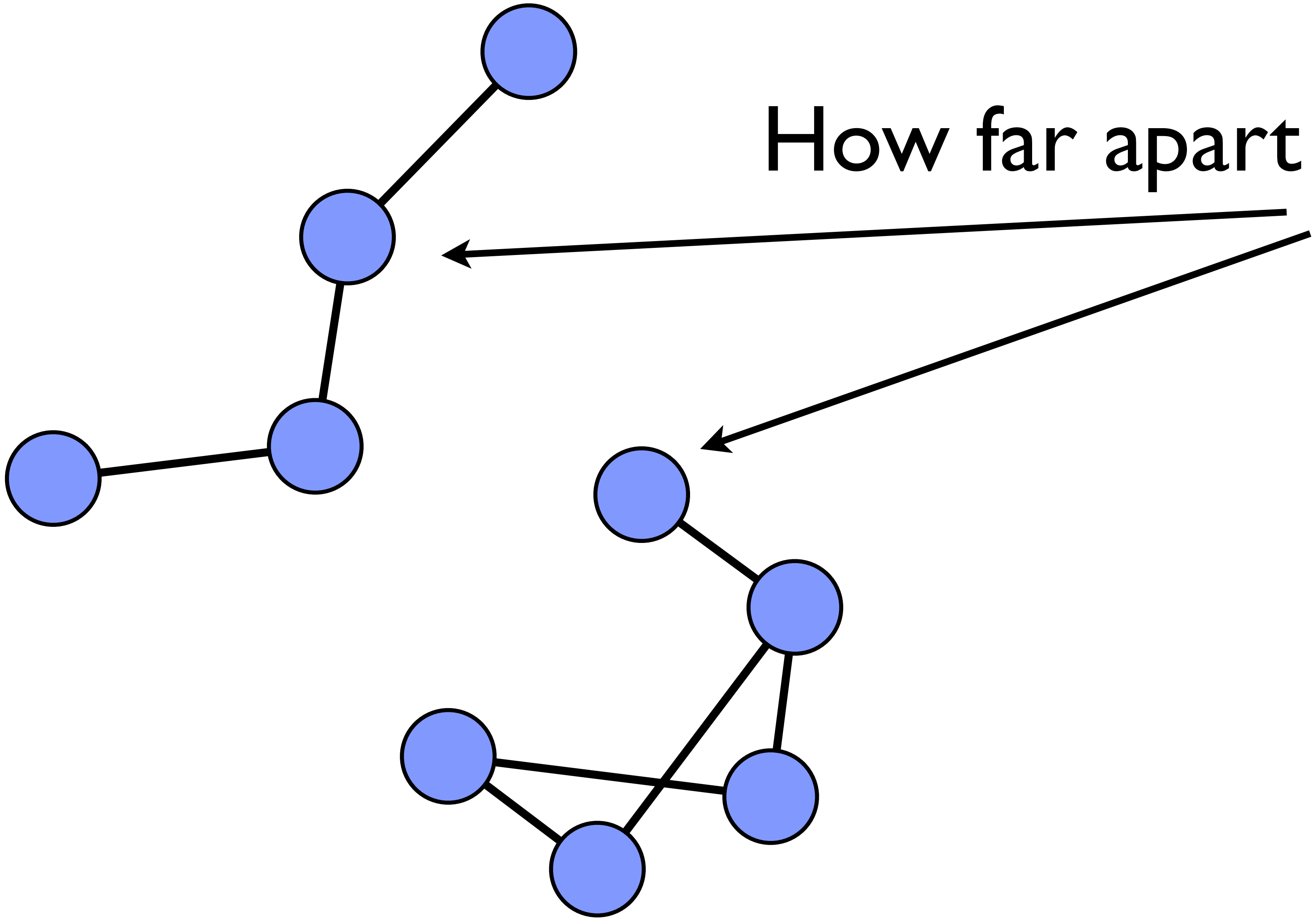
Diameter of a network

Longest of **all** shortest paths

How big can distances be?

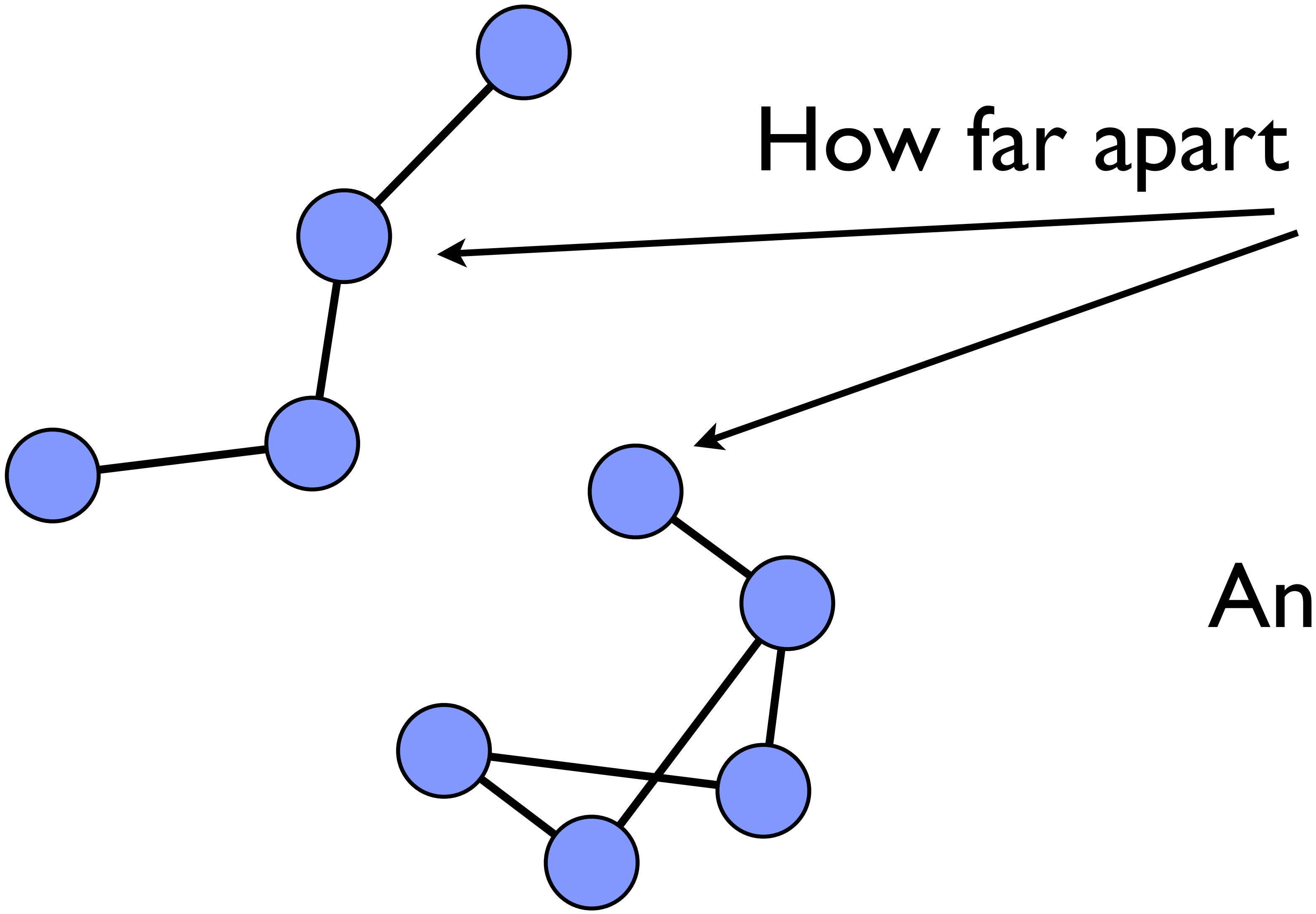


How big can distances be?



How far apart are these nodes?

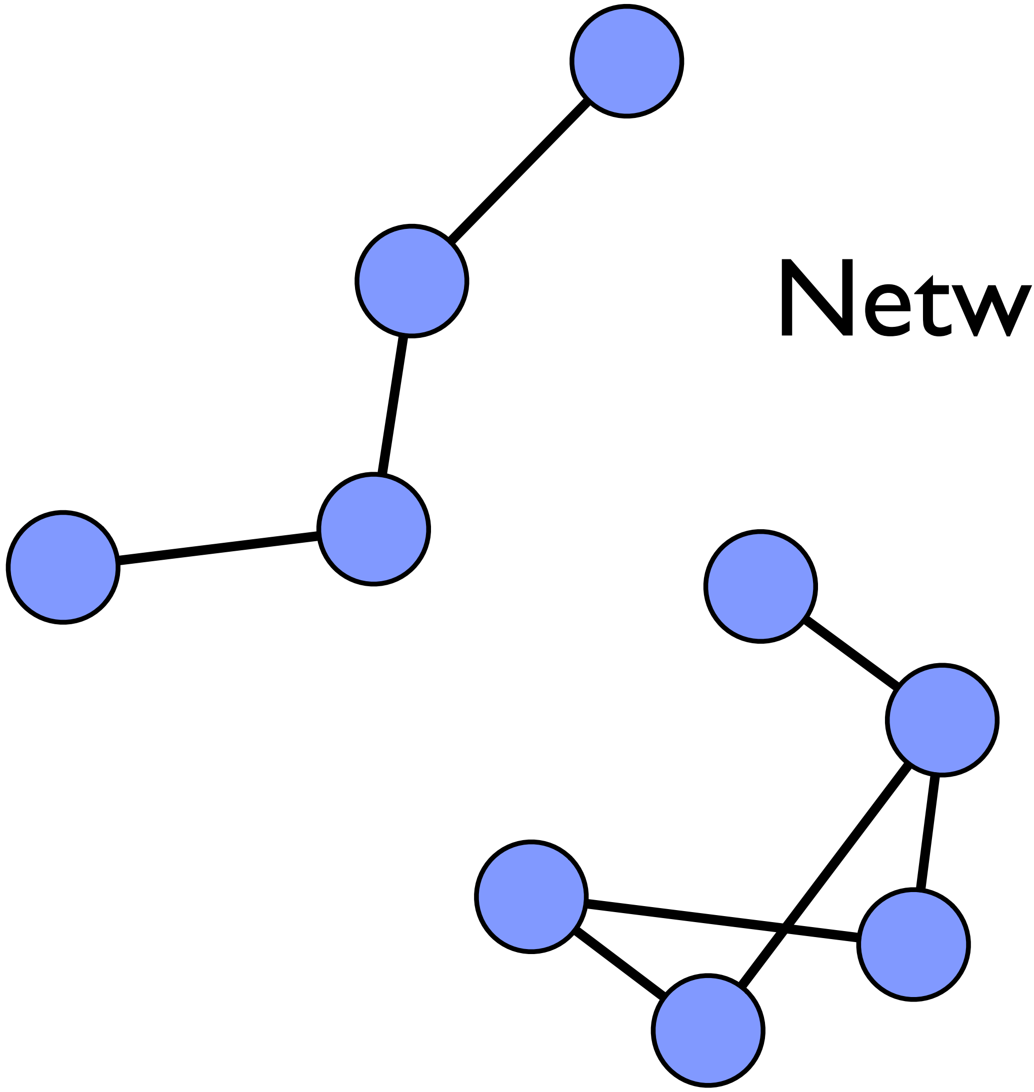
How big can distances be?



How far apart are these nodes?

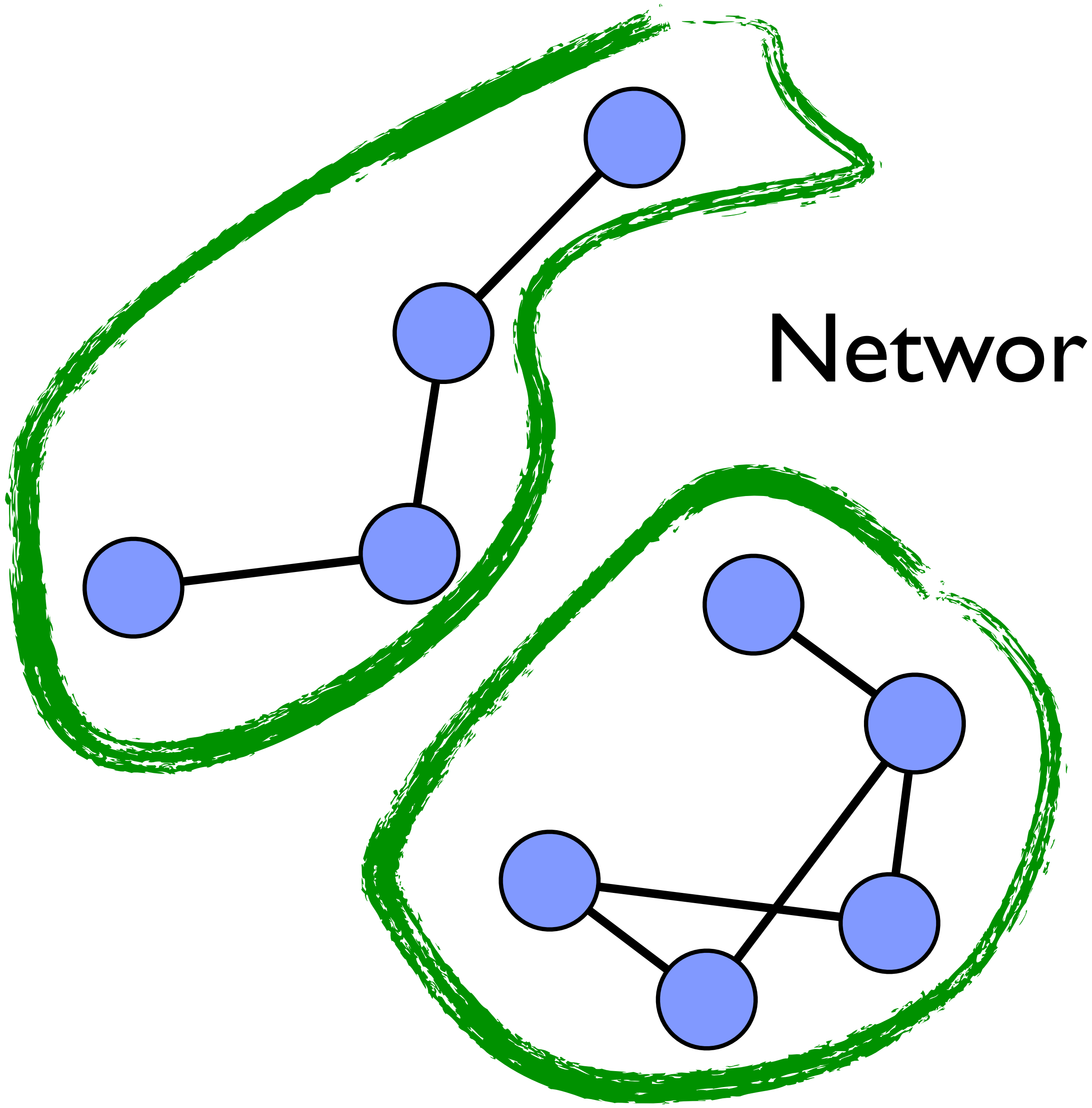
Ans: $d = \infty$

How big can distances be?



Networks can be **disconnected**
or **disjoint**

How big can distances be?



Networks can be **disconnected**

or **disjoint**

Components

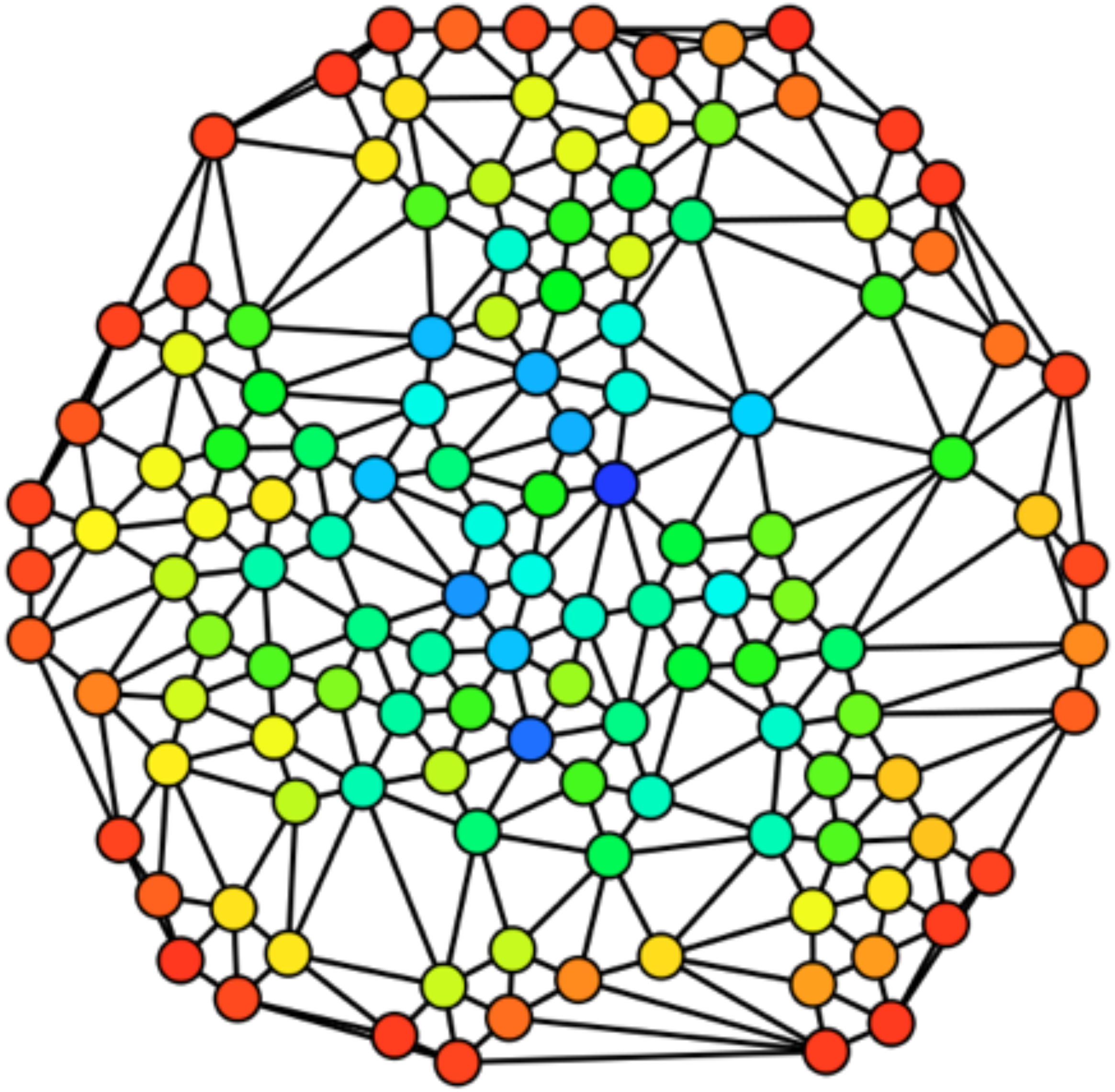
(Connected components)

Centrality

What nodes are “**important**”?

Centrality

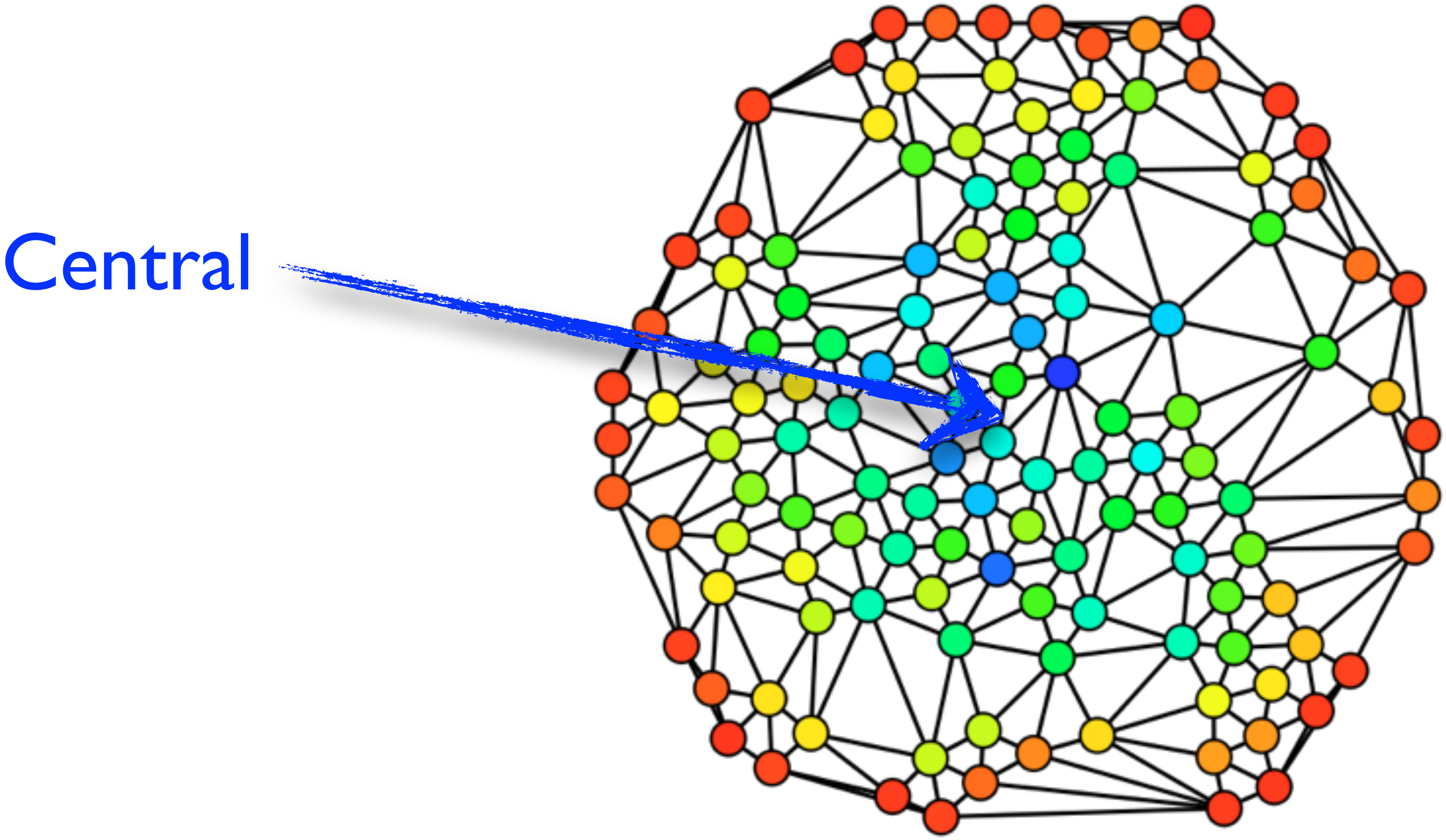
What nodes are “**important**”?



wikipedia.org

Centrality

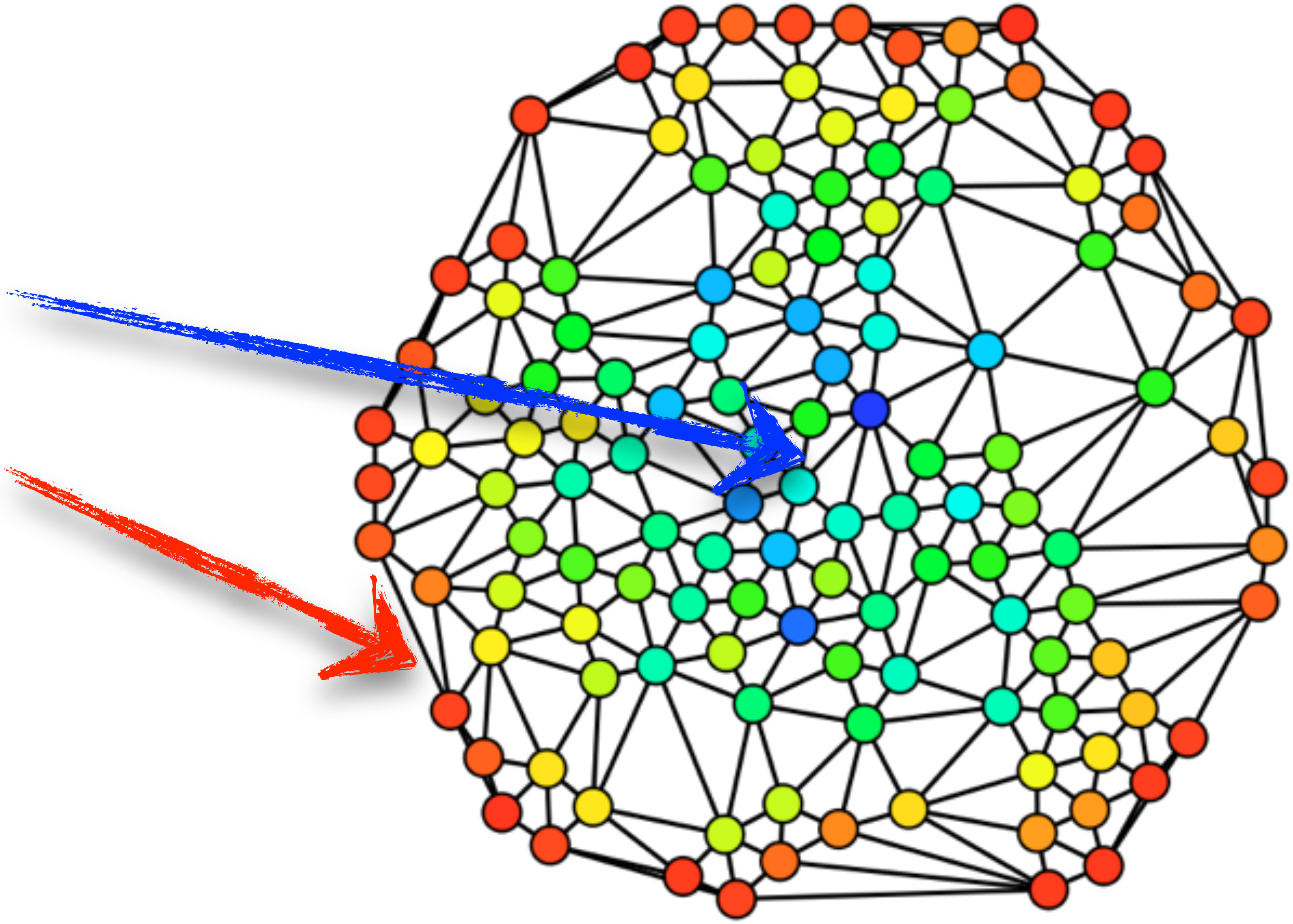
What nodes are "**important**"?



wikipedia.org

Centrality

What nodes are "**important**"?



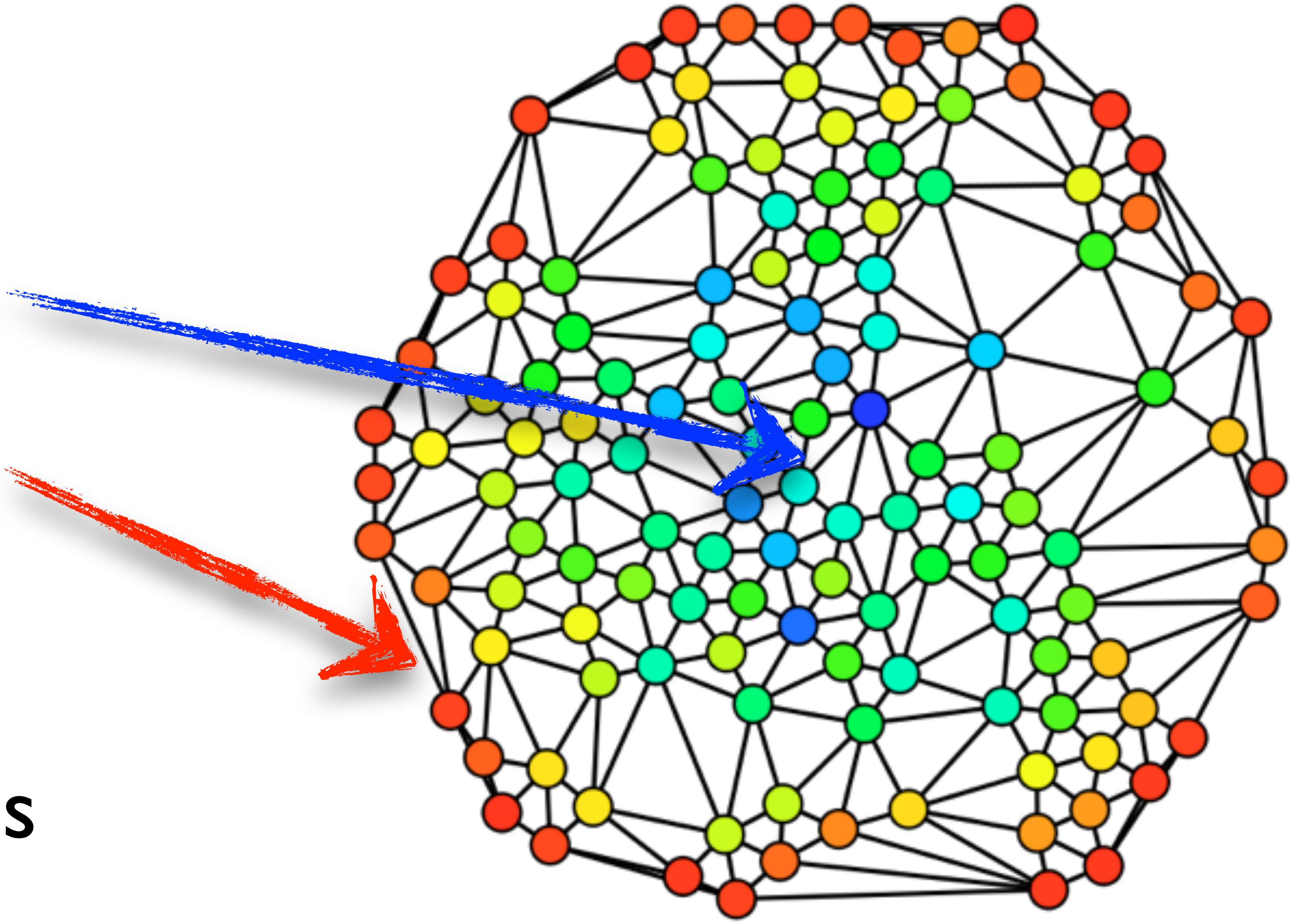
Central

Peripheral

wikipedia.org

Centrality

What nodes are “**important**”?



Central

Peripheral

Rank nodes

wikipedia.org

Centrality

Rank nodes

Centrality

Rank nodes

- Degree centrality: rank nodes by their degree
    - Hubs are most central

Centrality

Rank nodes

Degree centrality: rank nodes by their degree

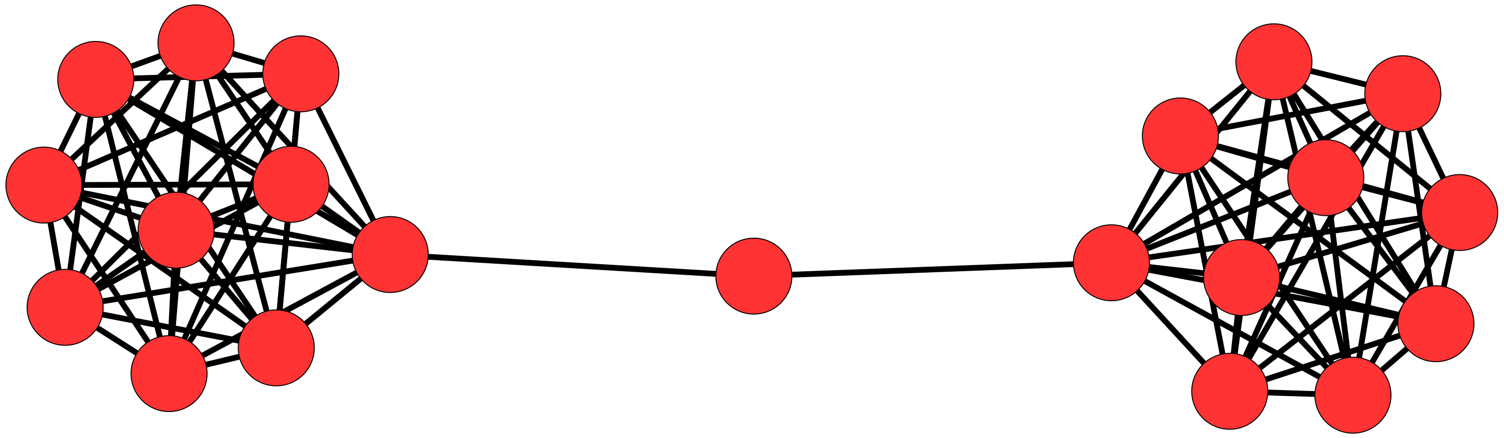
Hubs are most central

Betweenness centrality: rank nodes (or links) by number of shortest paths

Centrality

Rank nodes

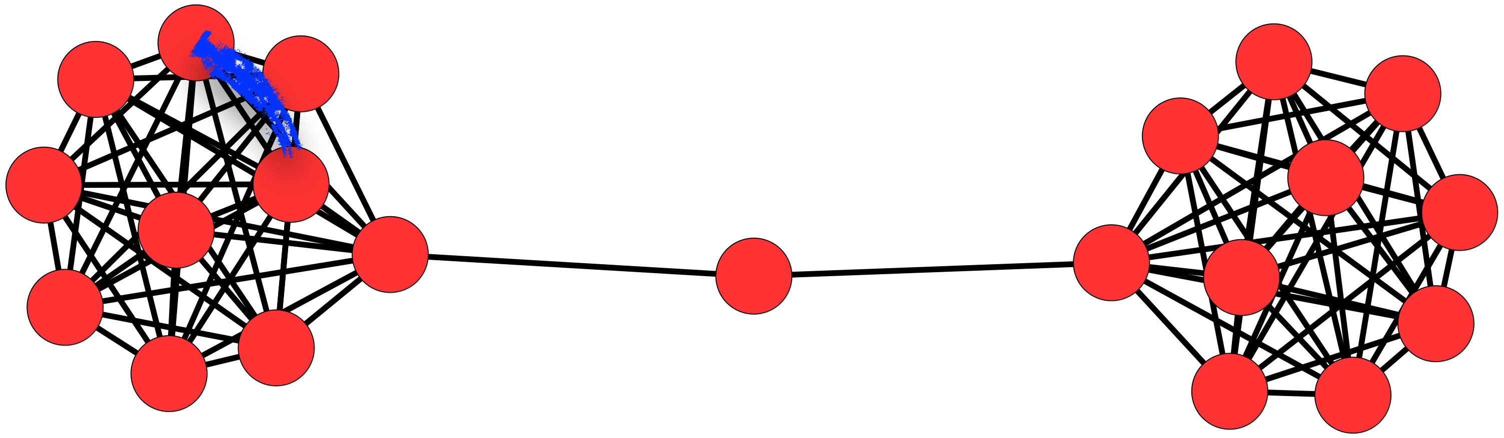
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

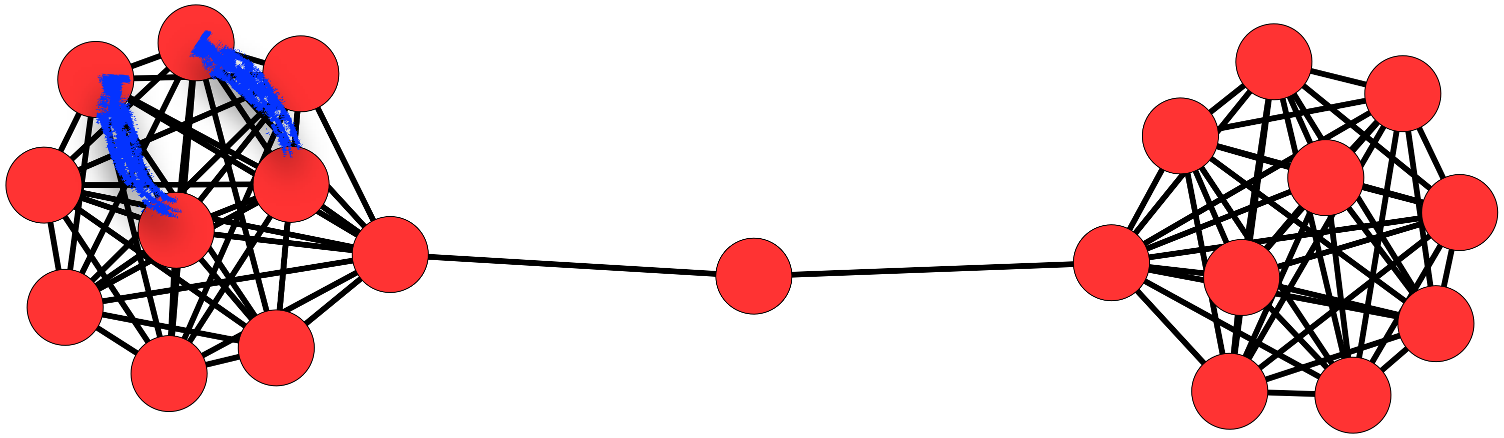
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

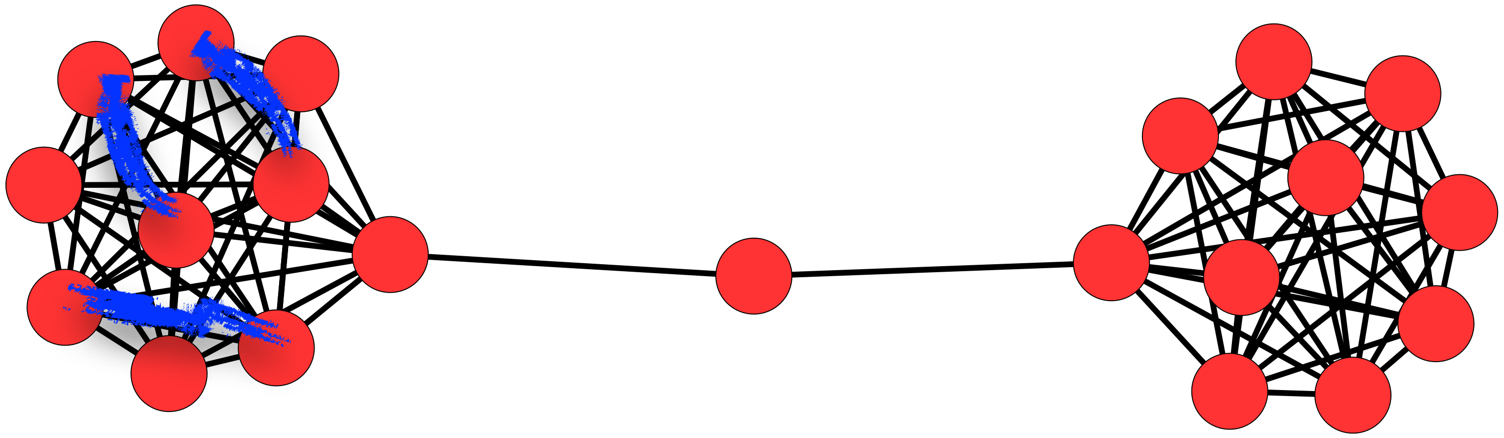
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

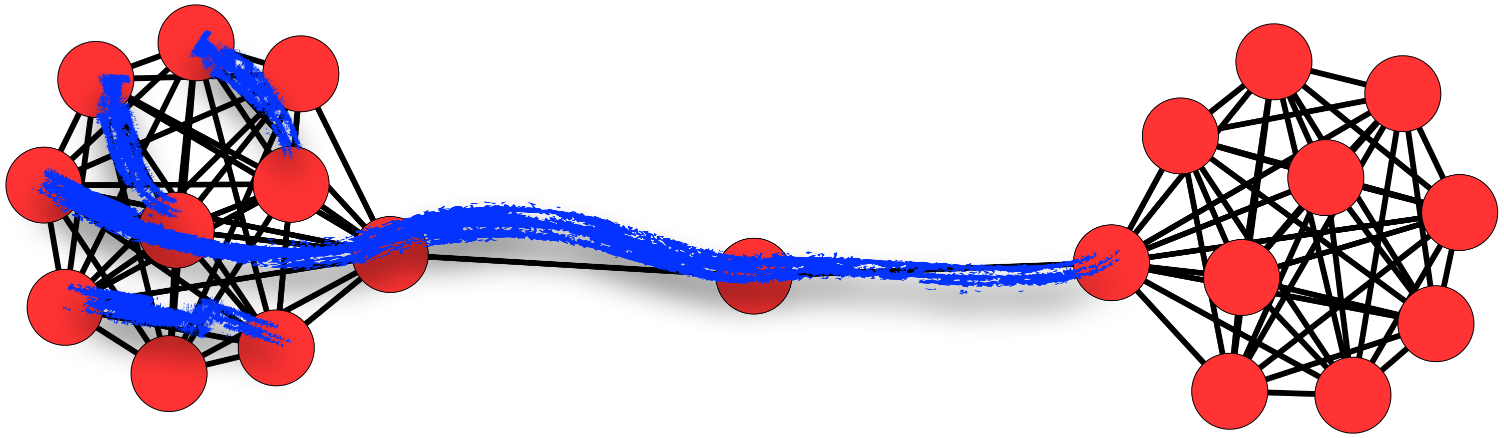
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

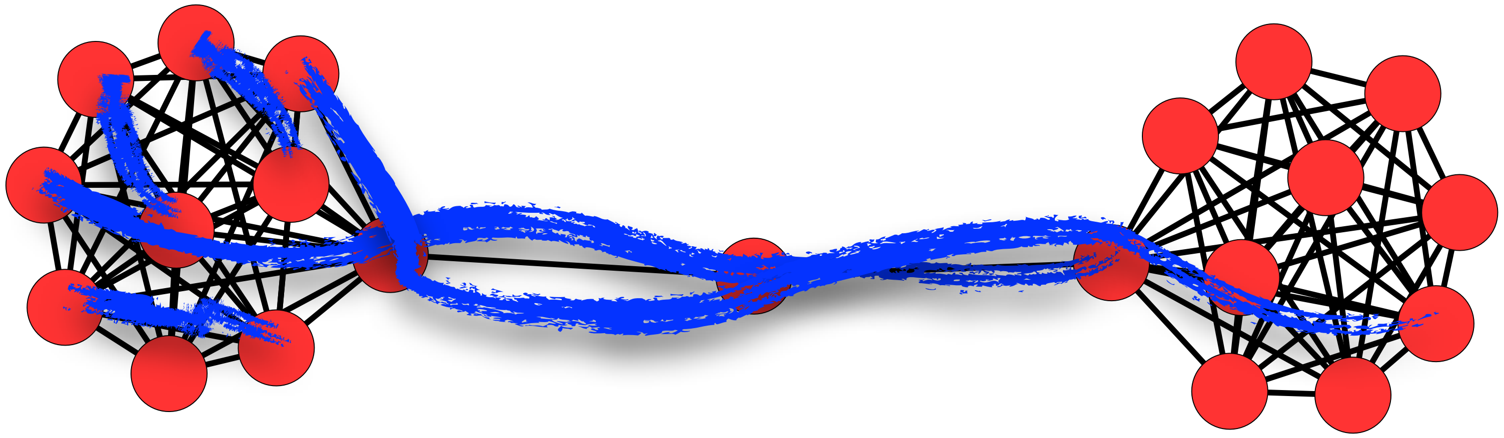
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

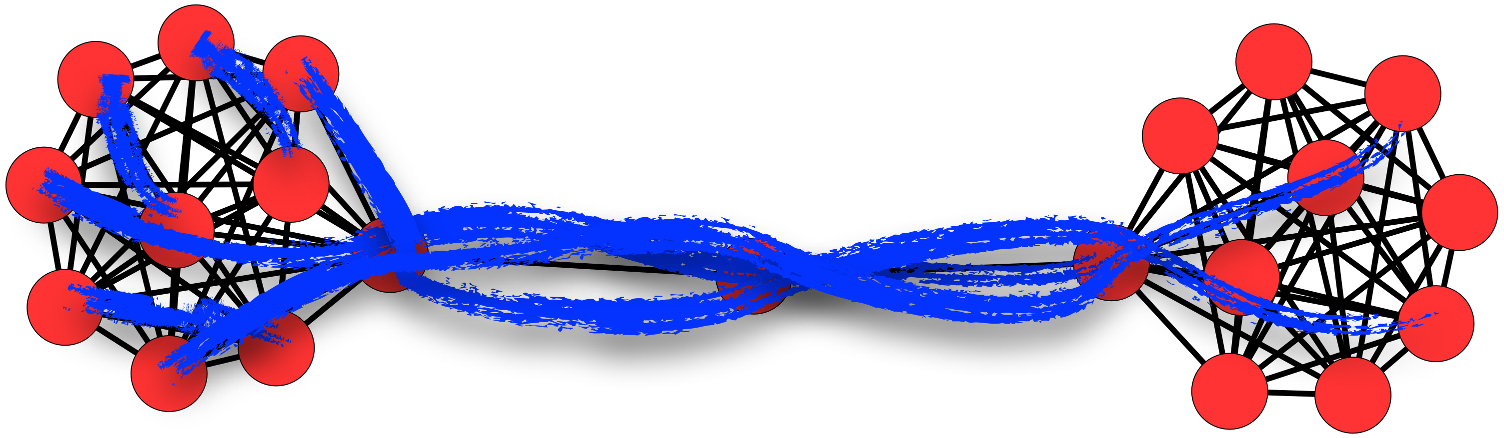
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

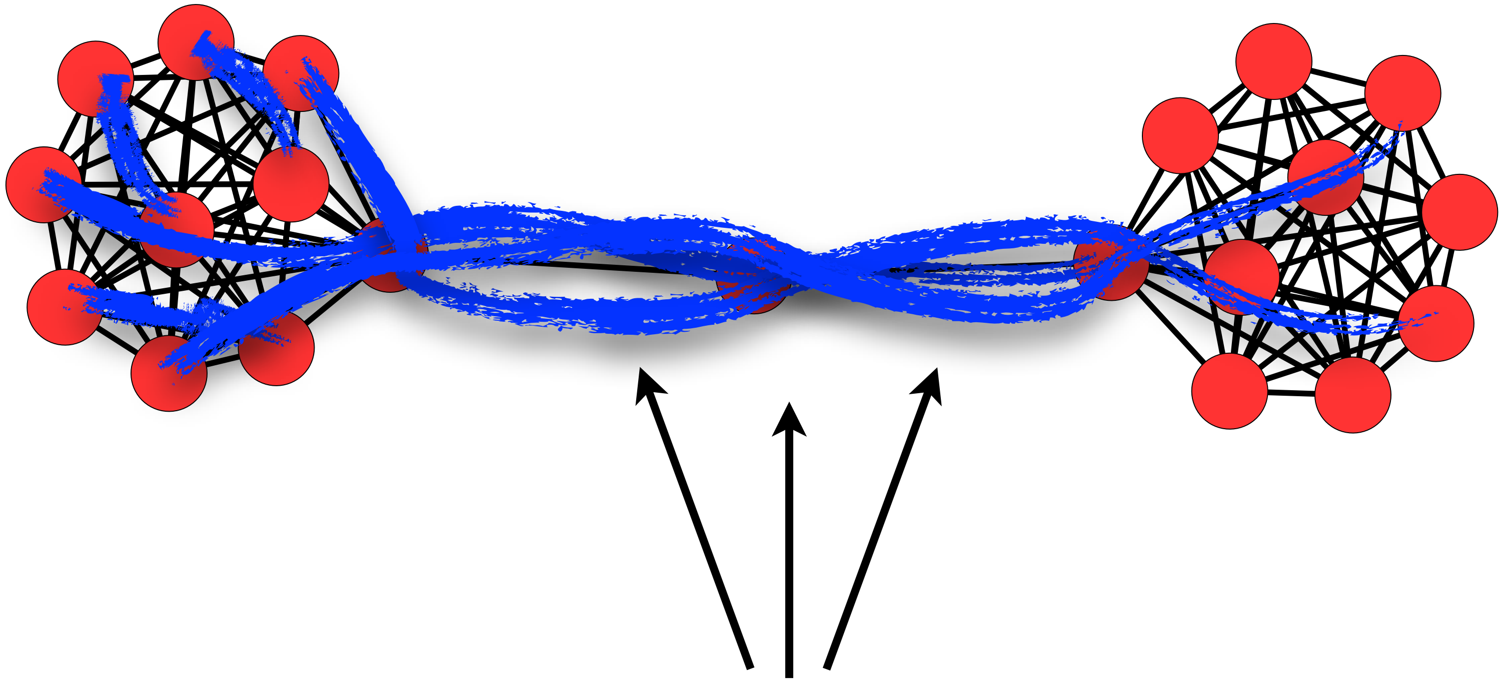
Betweenness centrality: rank nodes (or links) by number of shortest paths



Centrality

Rank nodes

Betweenness centrality: rank nodes (or links) by number of shortest paths

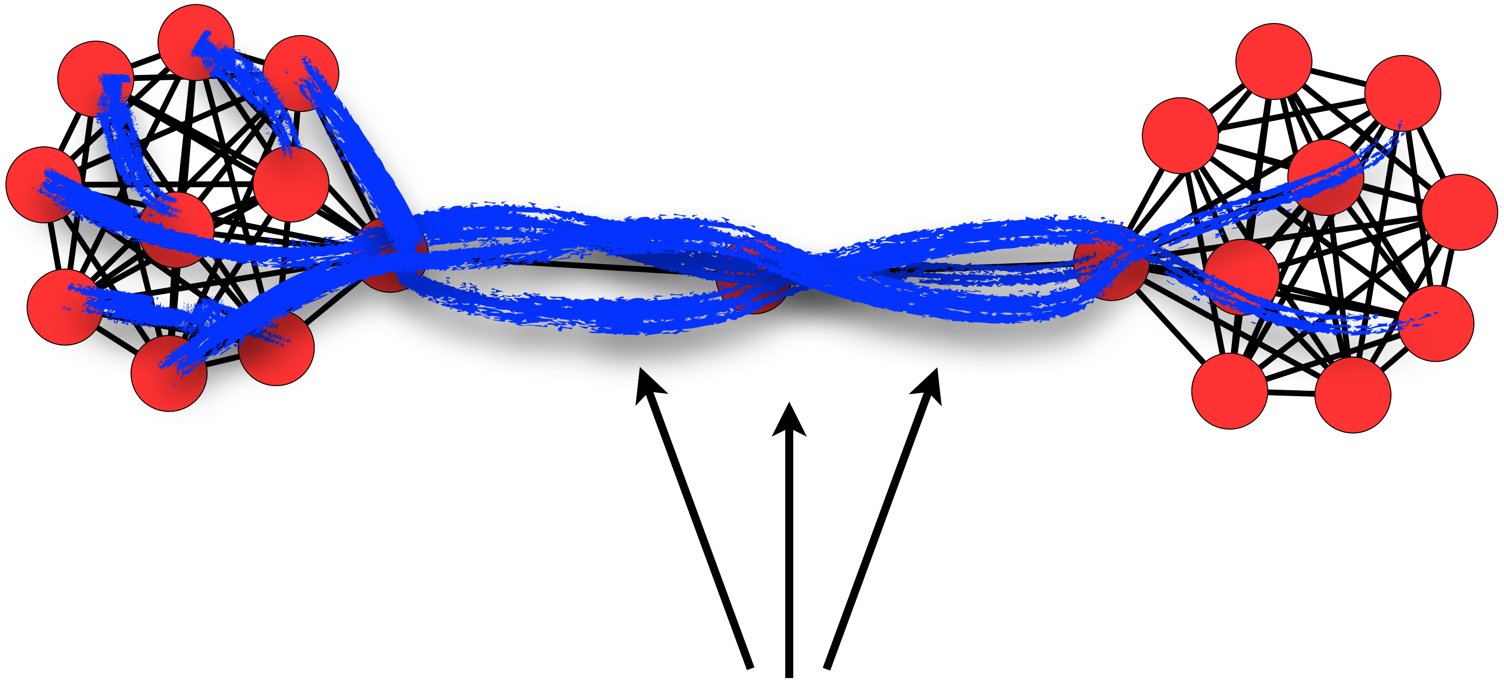


Very central bridge

Centrality

Rank nodes

Betweenness centrality: rank nodes (or links) by number of shortest paths



but **low** degree centrality!!

Centrality

Rank nodes

- PageRank

Centrality

Rank nodes

PageRank ⟶ Google

Centrality

Rank nodes

PageRank ⟶ Google

Random **web surfer**

Centrality

Rank nodes

PageRank → Google

Random **web surfer**

> “If I move around at
> random, where will I
> tend to **find myself**?”

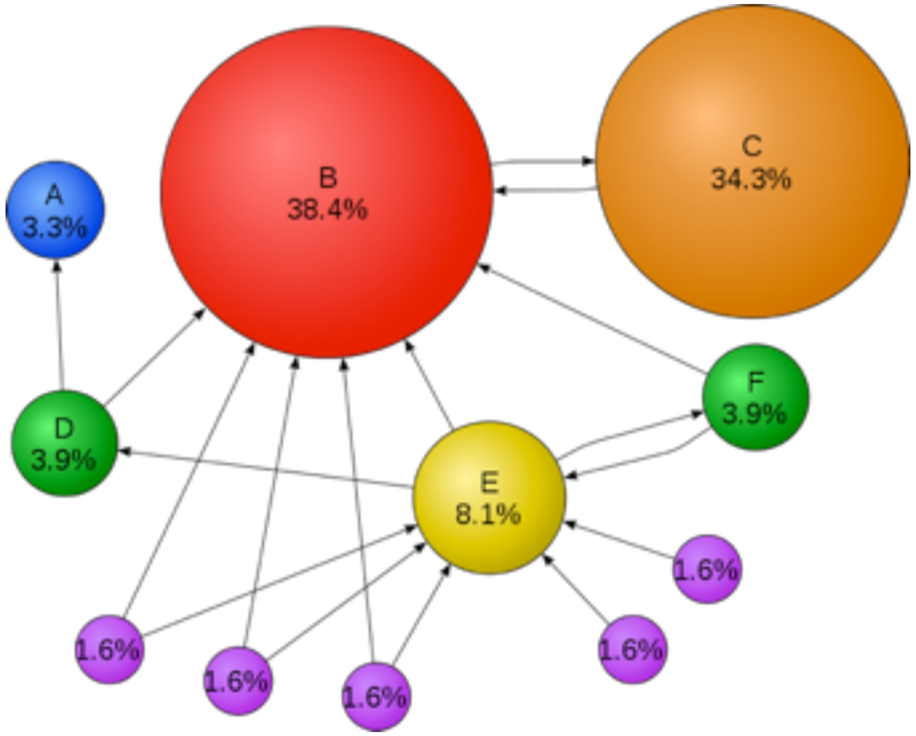
Centrality

Rank nodes

PageRank ⟶ Google

Random **web surfer**

"If I move around at random, where will I tend to **find myself**?"

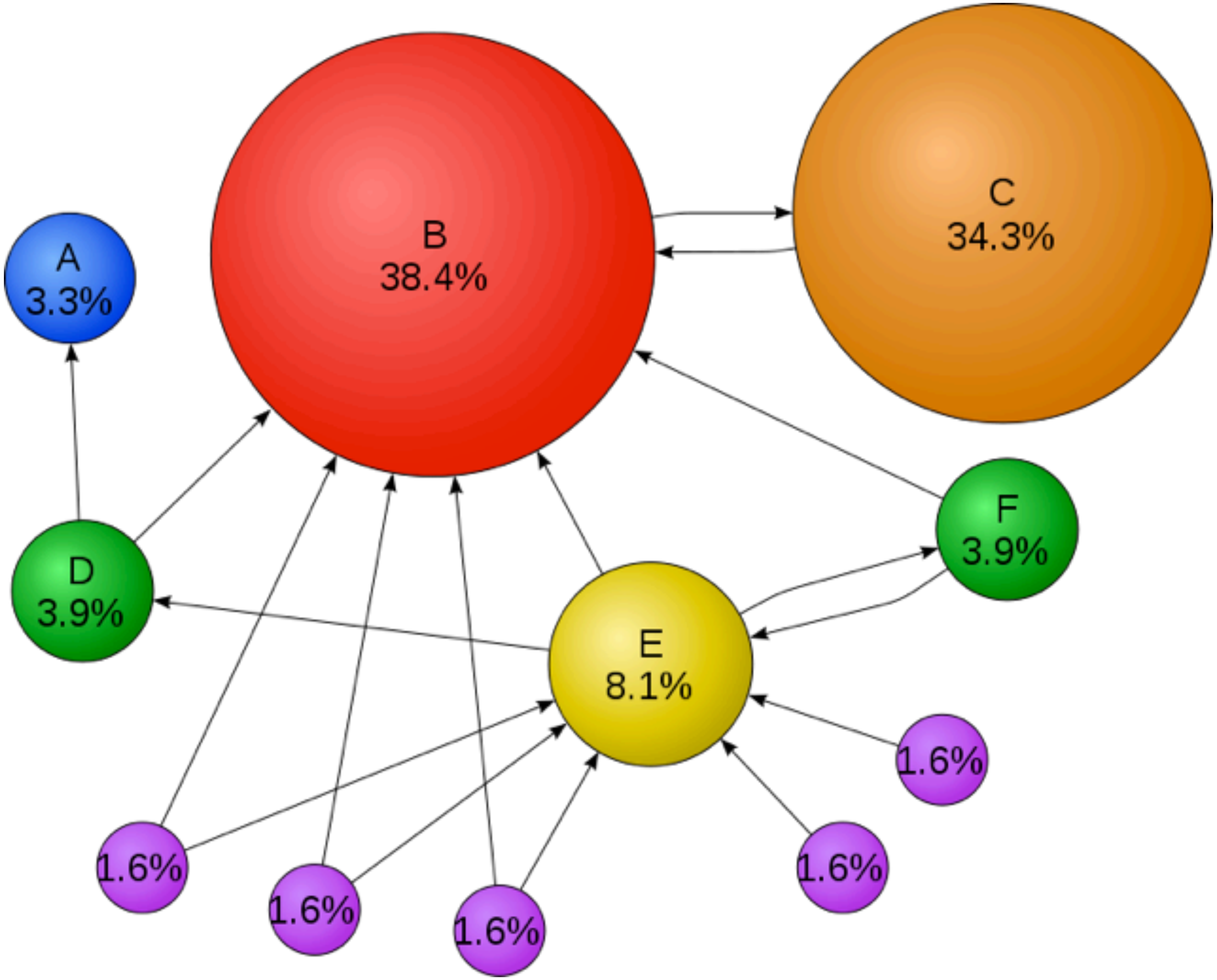


wikipedia.org

Centrality

Rank nodes

PageRank



wikipedia.org

Sparse and **dense** networks

Sparse and **dense** networks

Sparse network has a sparse adjacency matrix (**mostly zeros**)

Sparse and **dense** networks

Sparse network has a sparse adjacency matrix (**mostly zeros**)

typical degree << N

Sparse and **dense** networks

Sparse network has a sparse adjacency matrix (**mostly zeros**)

typical degree << N

max degree << N

Sparse and dense networks

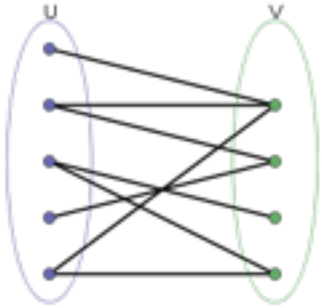
Sparse network has a sparse adjacency matrix (**mostly zeros**)

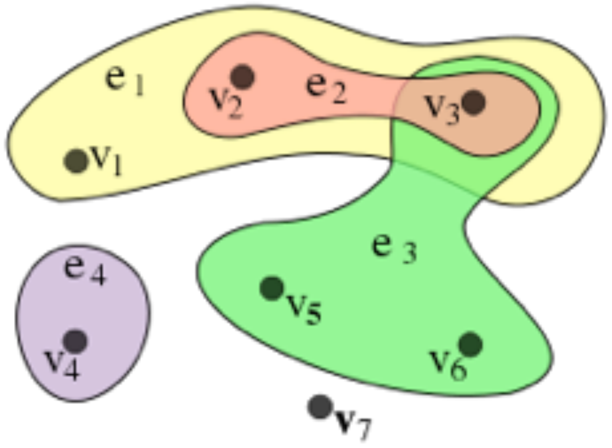
typical degree << N

max degree << N

“**dense**” is sometimes abused

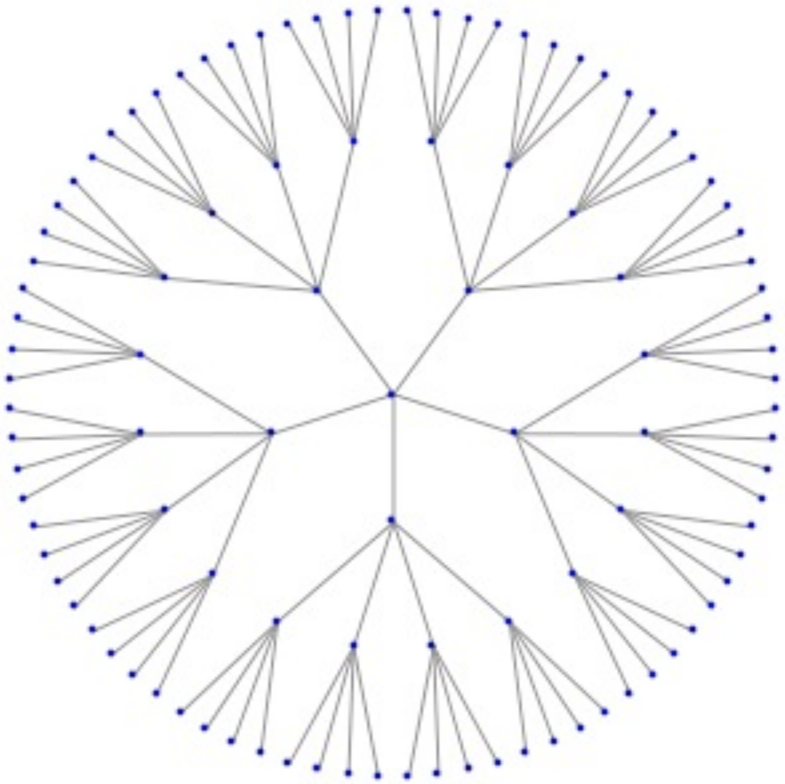
Types of networks and subnetworks





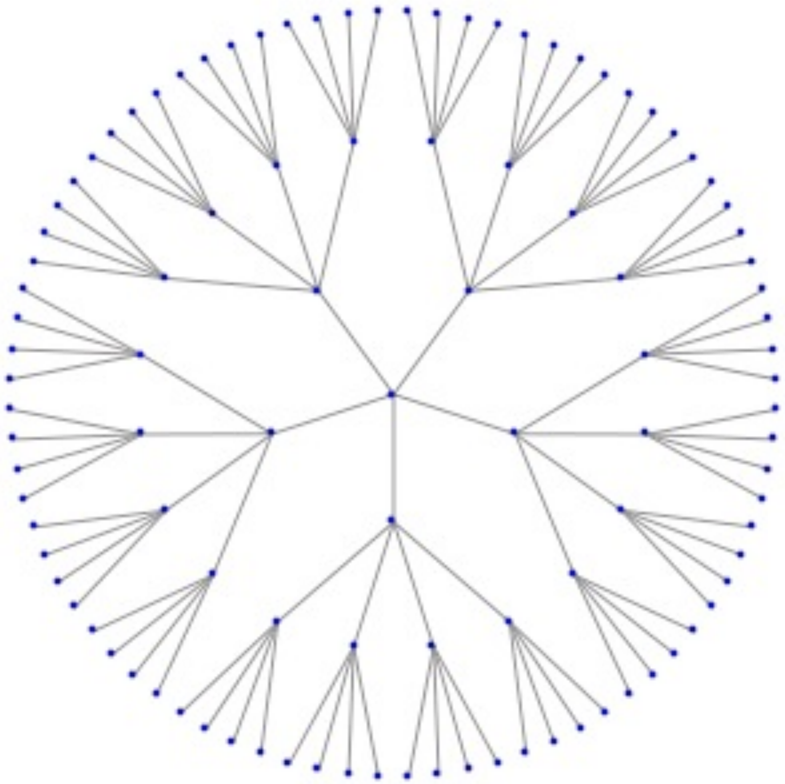
Network **zoology**

Trees Networks with **no loops**

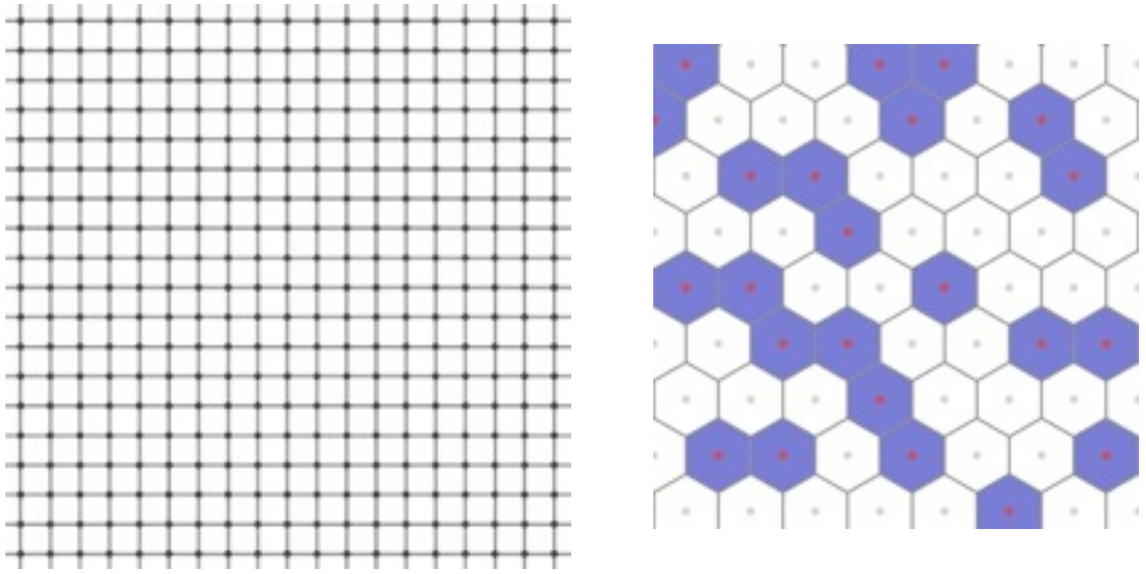


Network **zoology**

Trees Networks with **no loops**

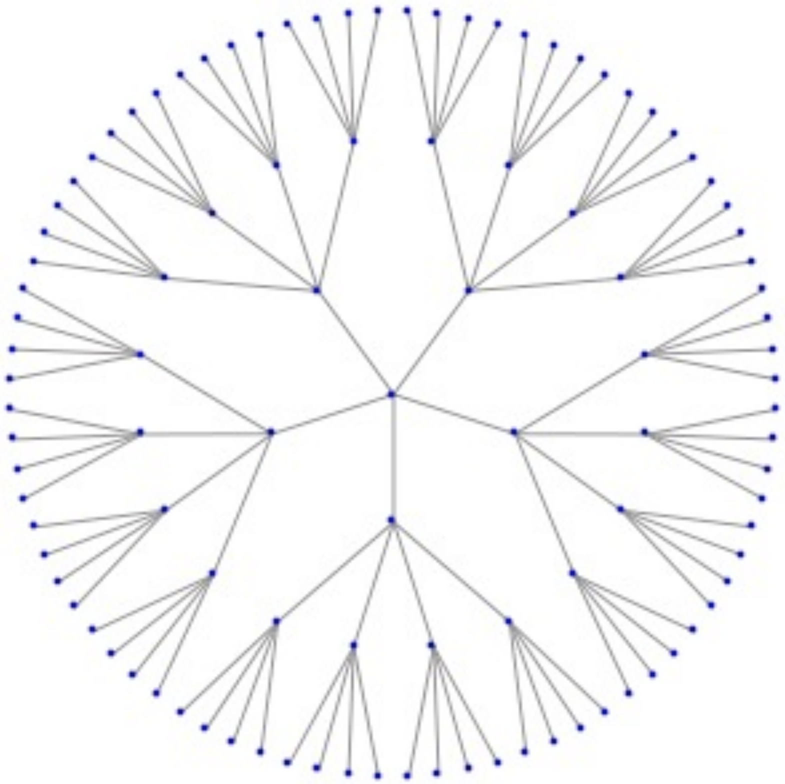


Lattices

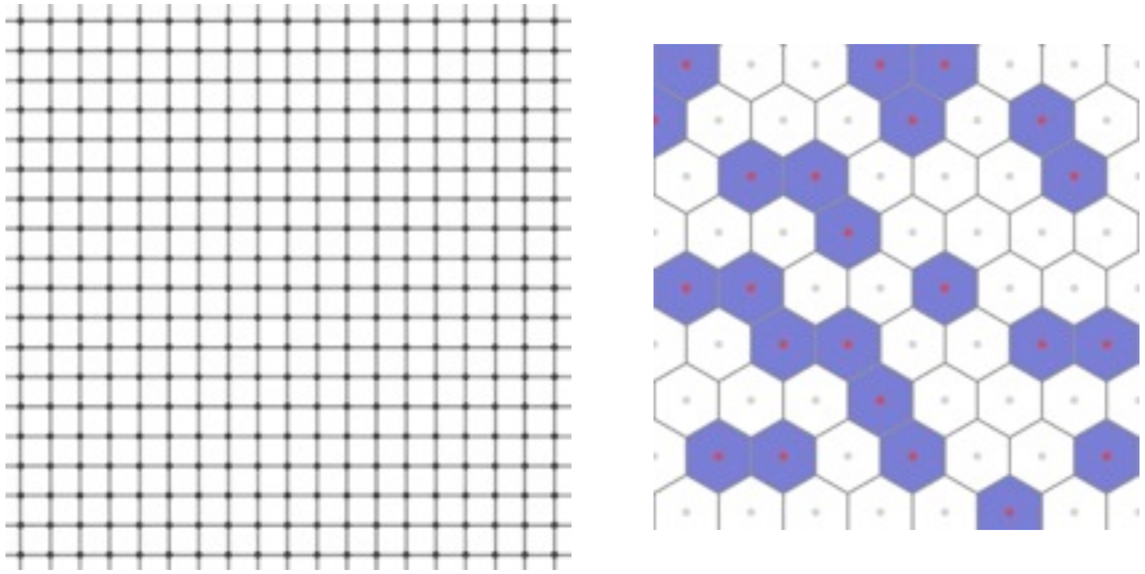


Network **zoology**

Trees Networks with **no loops**



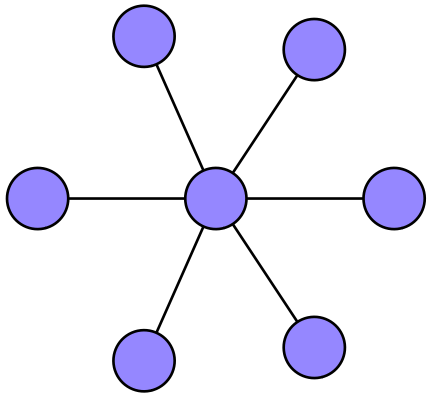
Lattices

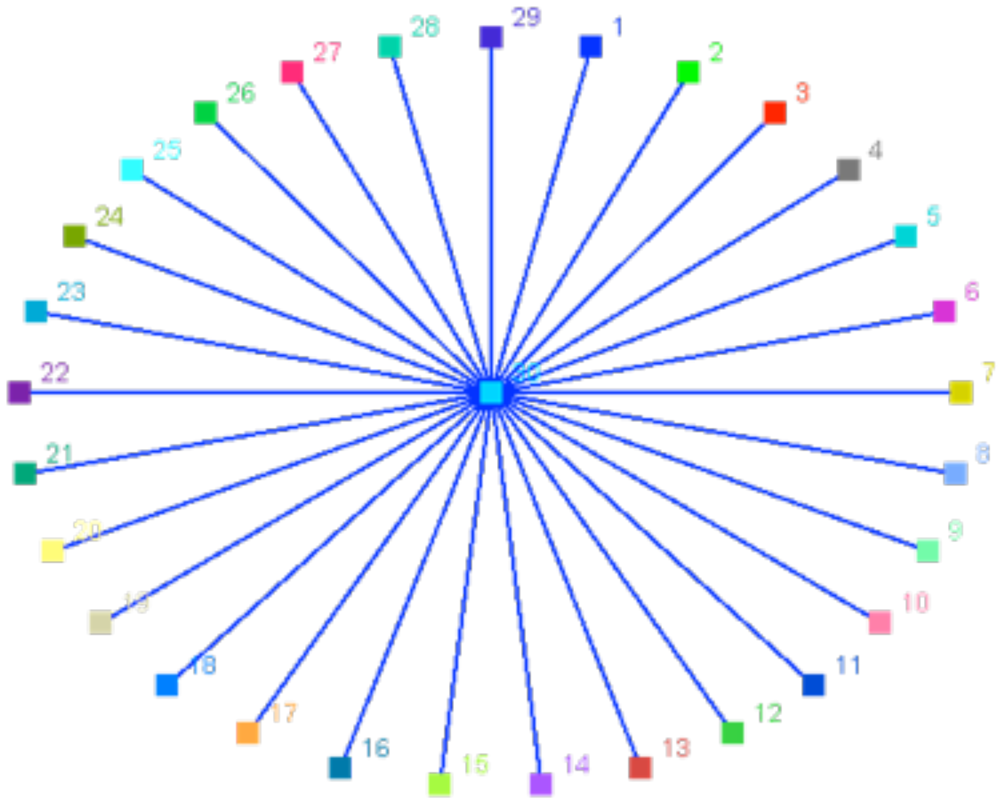


Regular Every node has the same degree

Network **zoology**

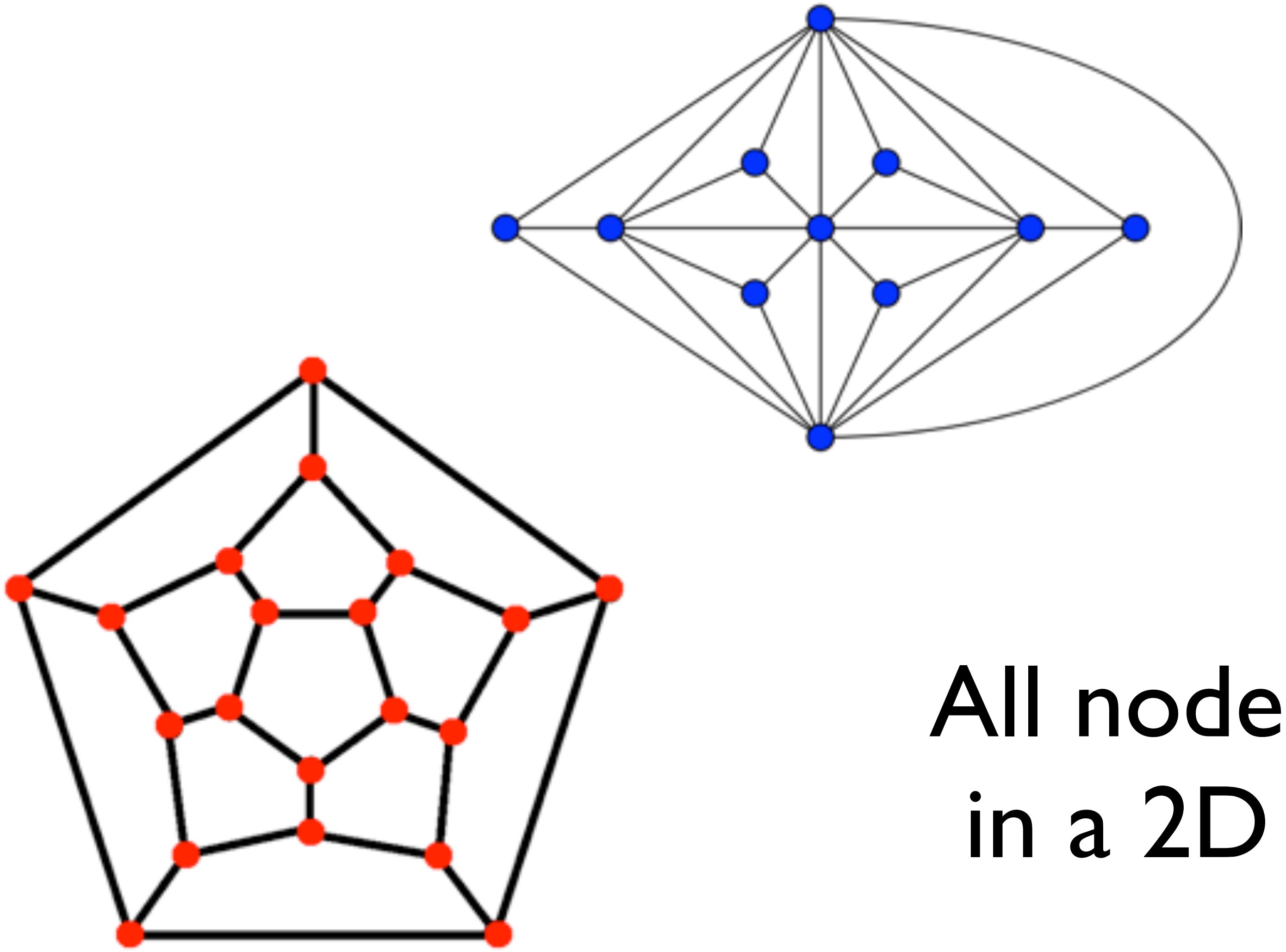
Star graph

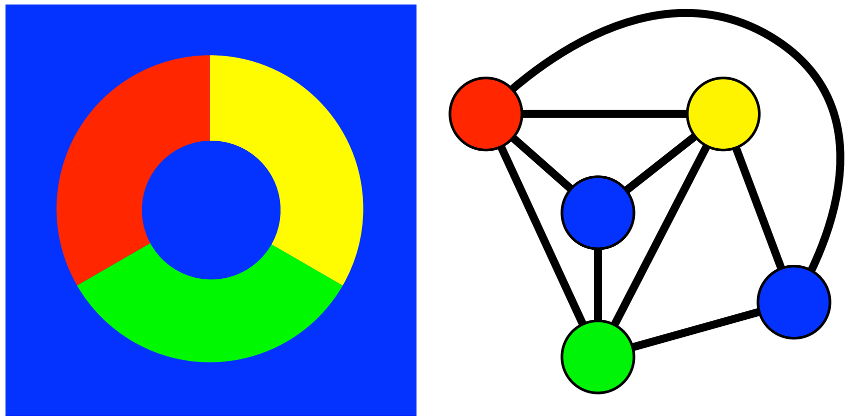




Network **zoology**

Planar graph





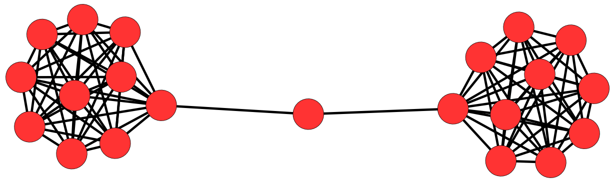
four color theorem

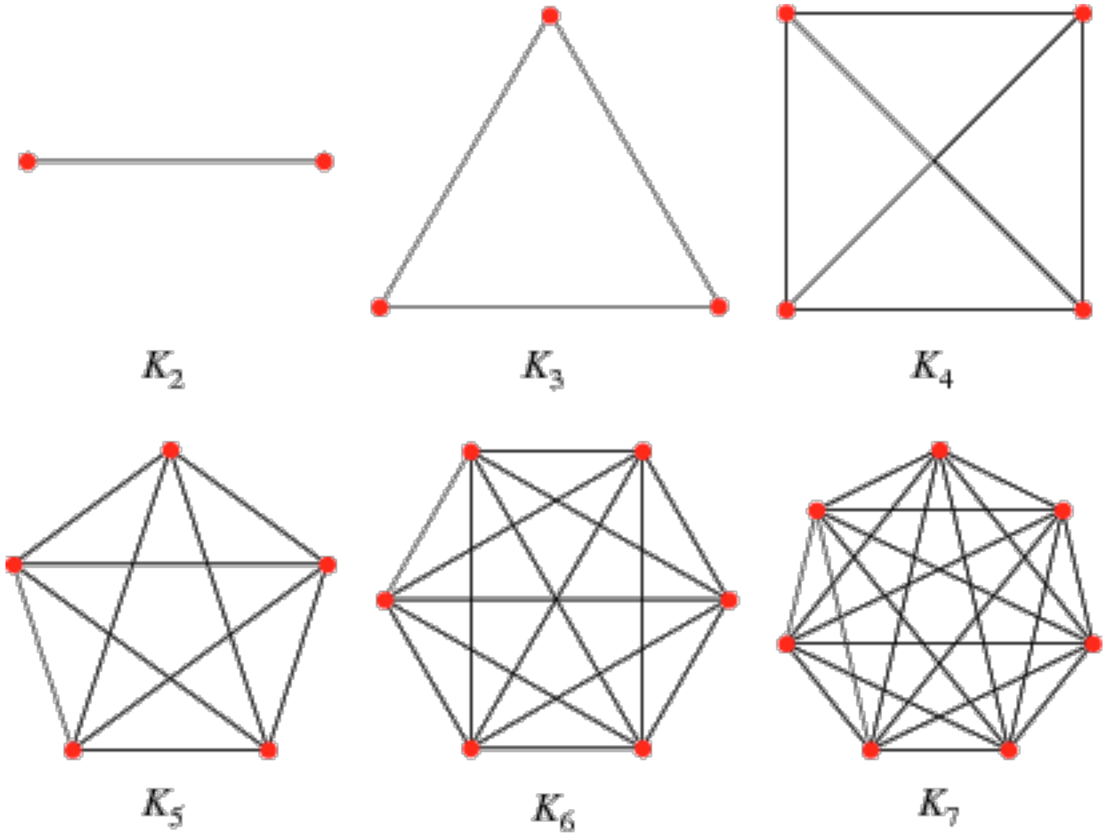
All nodes can be embedded in a 2D plane such that **no links cross**

Network **zoology**

Complete graph

Clique





Mathworld

Network **zoology**

Bipartite graph

Network **zoology**

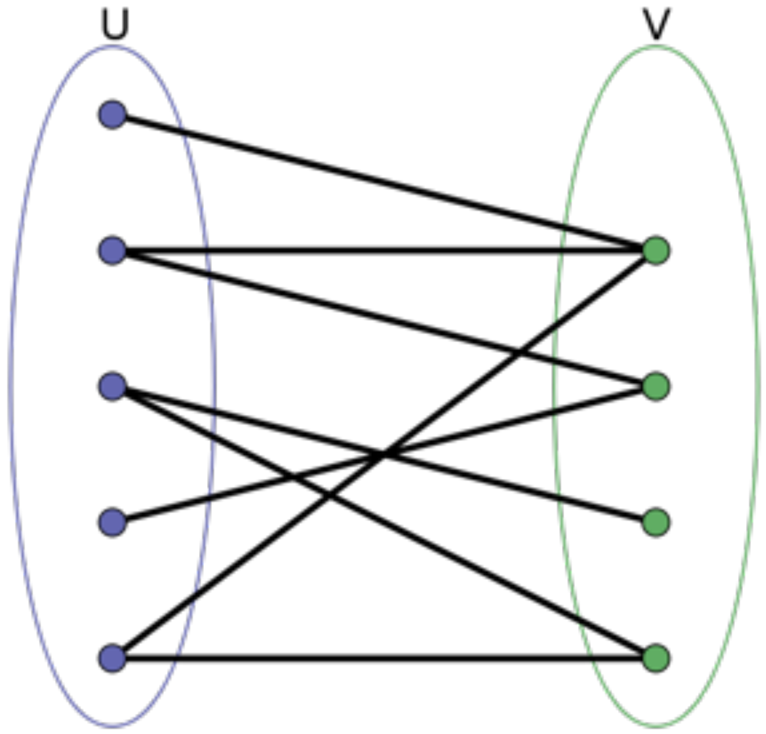
Bipartite graph Two **types** of nodes

Network **zoology**

Bipartite graph

Two **types** of nodes

Links only between nodes of **different types**

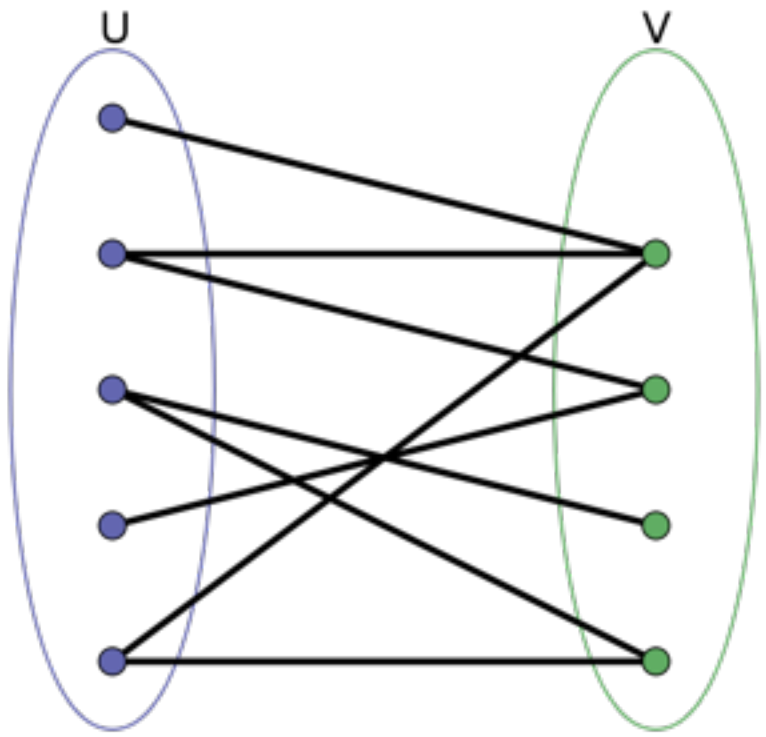


Network **zoology**

Bipartite graph

Two **types** of nodes

Links only between nodes of **different types**



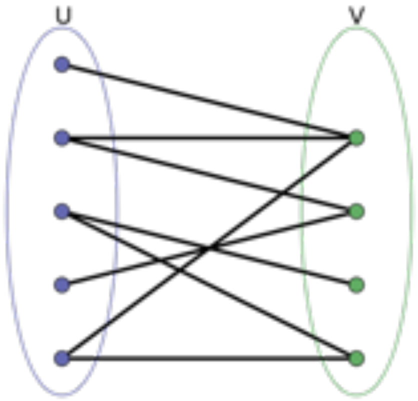
Movies ⟷ Actors

Enzymes ⟷ Metabolites

Scientists ⟷ Papers

Network **zoology**

Bipartite **projection**

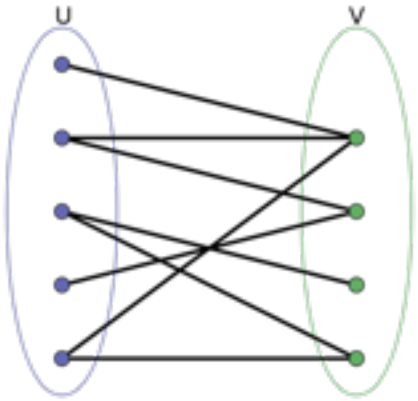


Movies ⟷ Actors

Network **zoology**

Bipartite **projection**

Link nodes in one group that have **common neighbors** in the other group

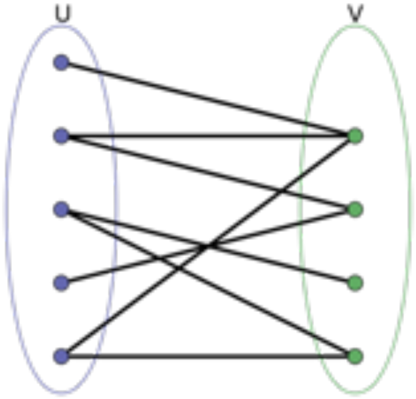


Movies ⟷ Actors

Network **zoology**

Bipartite **projection**

Link nodes in one group that have **common neighbors** in the other group



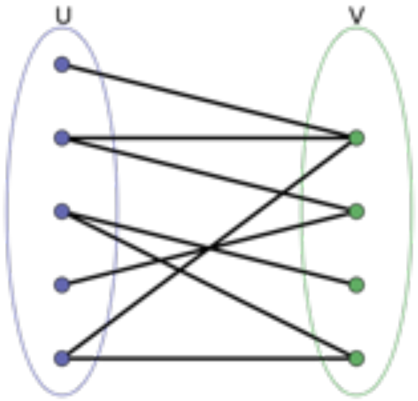
Movies ⟷ Actors

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

Network **zoology**

Bipartite **projection**

Link nodes in one group that have **common neighbors** in the other group



Movies ⟷ Actors

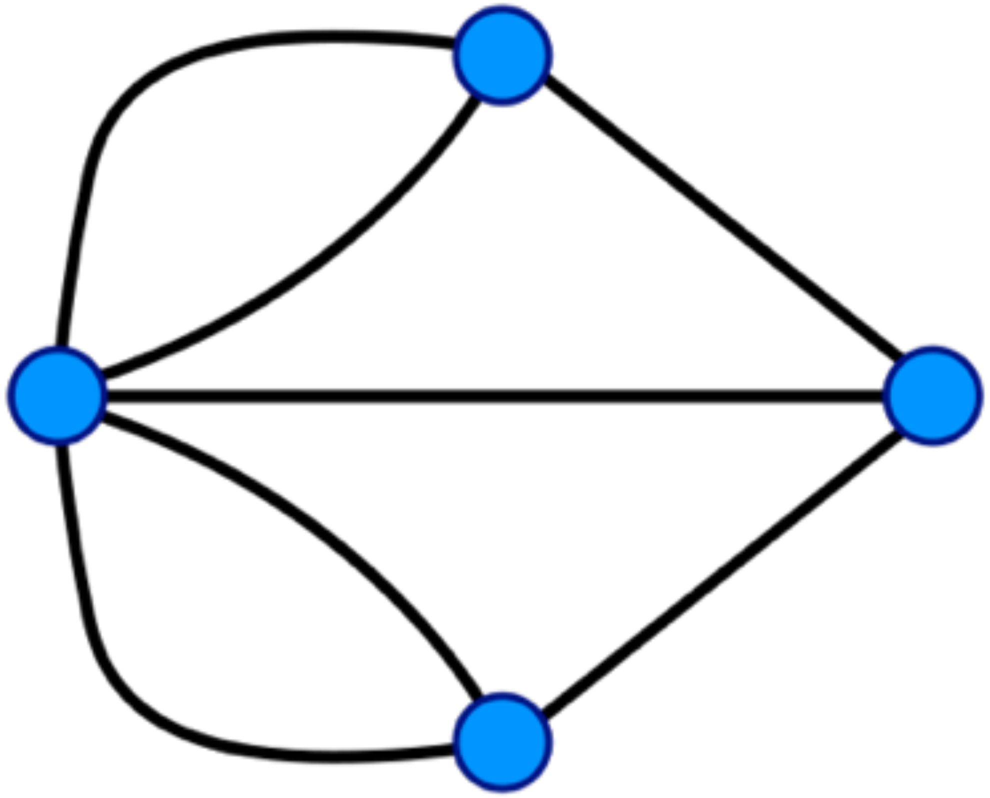
"**Movies** that star the same actor(s)"

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

Network **zoology**

Multigraphs

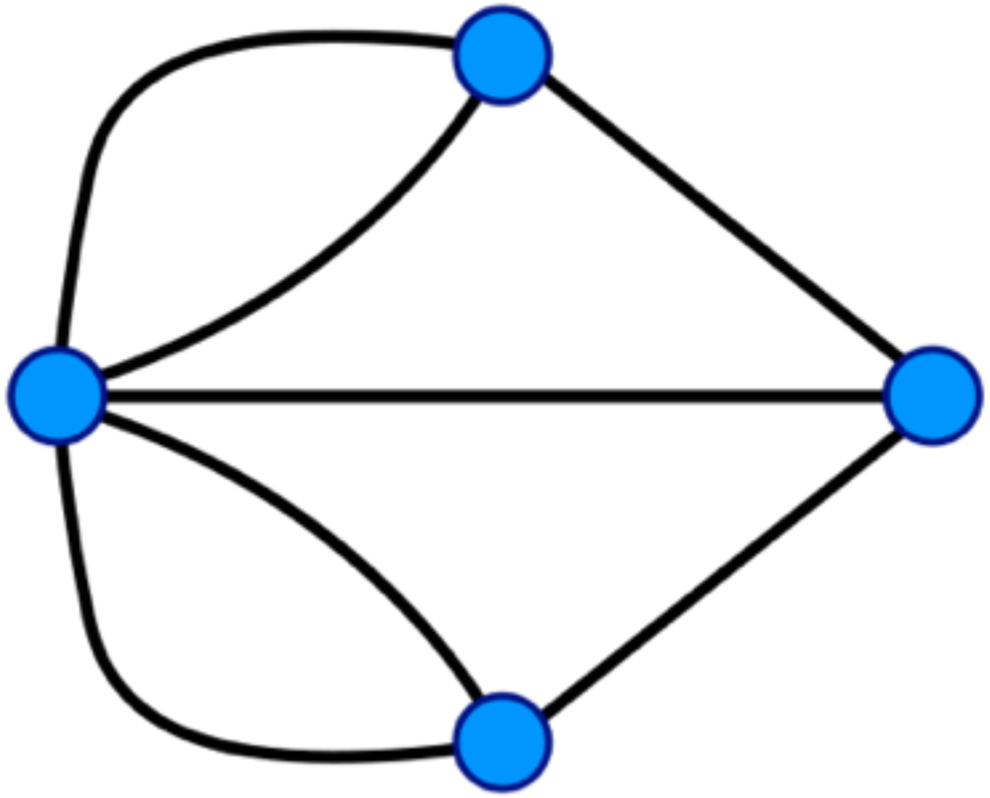
More than one link between node pairs



Network **zoology**

Multigraphs

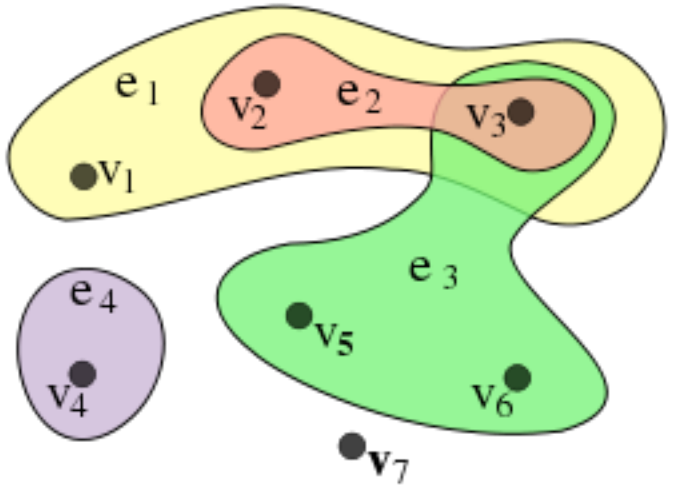
More than one link between node pairs



Network **zoology**

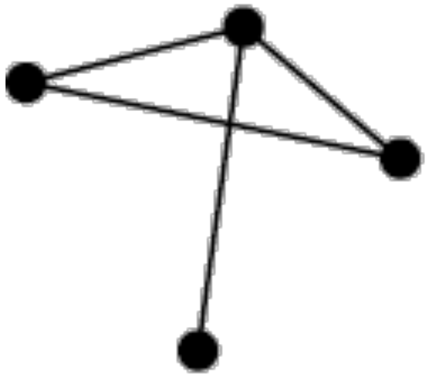
Hypergraphs

Links between **more than two nodes**

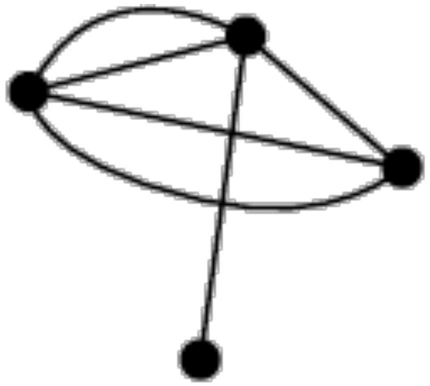


Network **zoology**

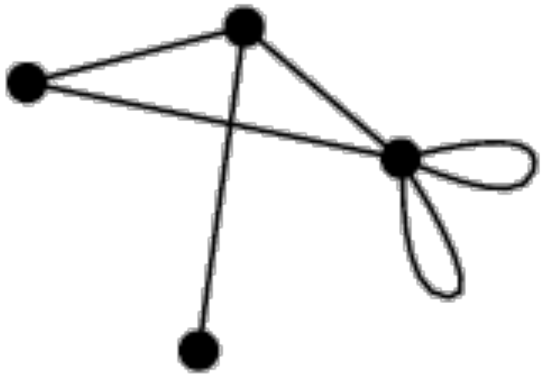
Etc.



simple graph



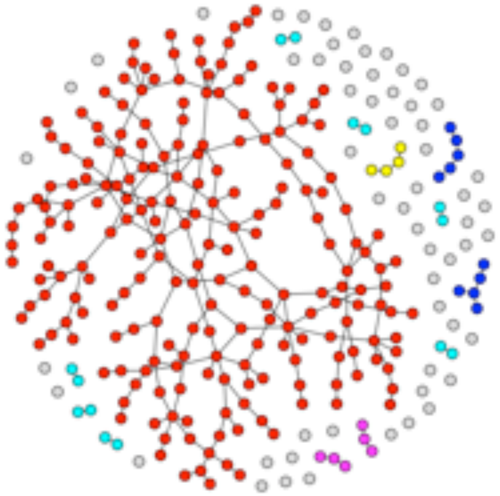
nonsimple graph
with multiple edges



nonsimple graph
with loops

Mathworld

Random network models



Modeling networks

Modeling networks

So much data
nowadays


facebook

Facebook helps you connect
the people in your life.

twitter

What is Twitter?


The Economist

Obama the warrior
Misgoverning Argentina
The economic shift from West to East
Genetically modified crops blossom
The right to eat cats and dogs

The data deluge

AND HOW TO HANDLE IT: A 14-PAGE SPECIAL REPORT

2010

Modeling networks

So much data
nowadays

Why turn to **models**?

facebook

Facebook helps you connect
the people in your life.

twitter

What is Twitter?

Google

The Economist

Obama the warrior
Misgoverning Argentina
The economic shift from West to East
Genetically modified crops blossom
The right to eat cats and dogs

The data deluge

AND HOW TO HANDLE IT: A 14-PAGE SPECIAL REPORT

2010

Modeling networks

So much data **nowadays**

Why turn to **models**?

Google

The Economist

Obama the warrior
Misgoverning Argentina
The economic shift from West to East
Genetically modified crops blossom
The right to eat cats and dogs

The data deluge

AND HOW TO HANDLE IT: A 14-PAGE SPECIAL REPORT

2010

A1. Wasn't **always** this much data

Modeling networks

So much data
nowadays

Google

2010

Why turn to **models**?

A1. Wasn't **always**
this much data

A2.

Modeling networks

Why turn to **models**?

A2.

Modeling networks

Why turn to **models**?

A2. To try to understand underlying **principles** or organizing **laws**

Modeling networks

Why turn to **models**?

A2. To try to understand
underlying **principles**
or organizing **laws**

Build **simplified** networks
preserving/destroying certain
features or properties

Modeling networks

Why turn to **models**?

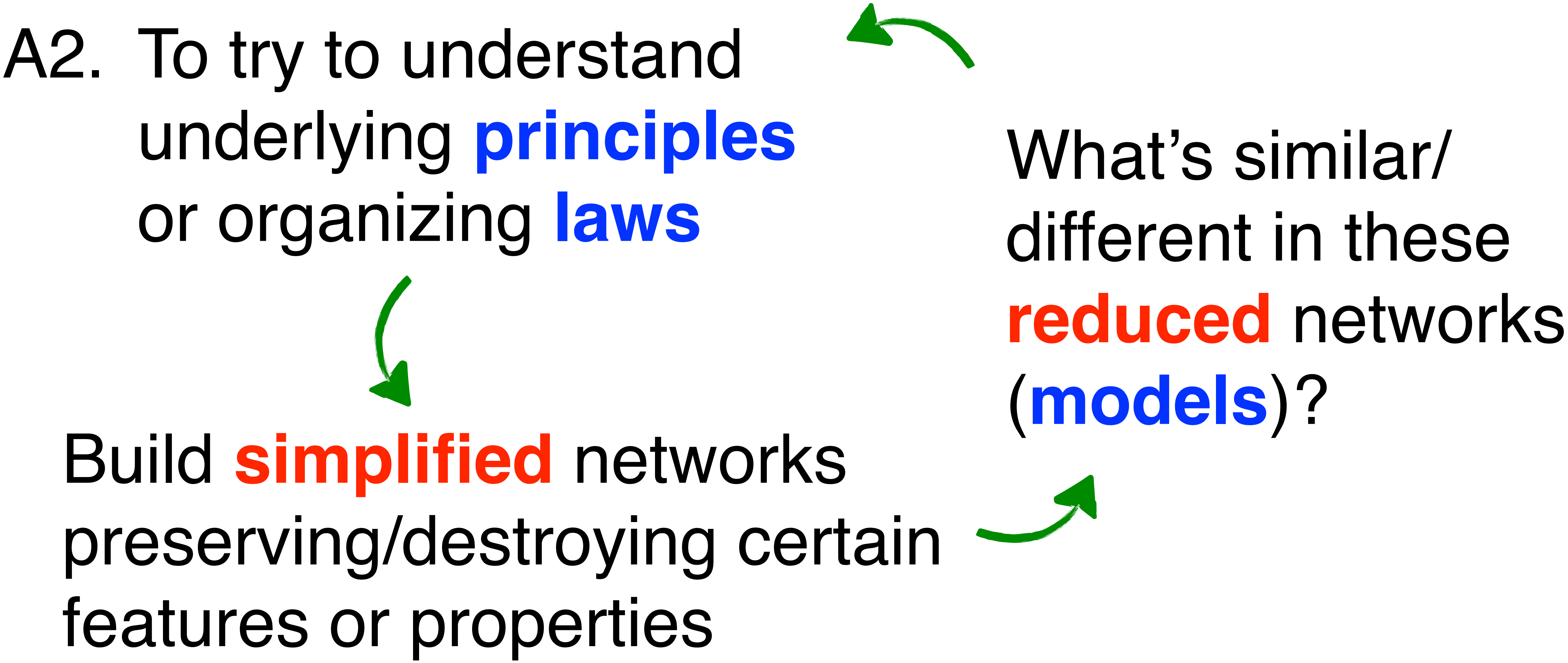
A2. To try to understand underlying **principles** or organizing **laws**

Build **simplified** networks preserving/destroying certain features or properties

What's similar/ different in these **reduced** networks (**models**)?

Modeling networks

Why turn to **models**?



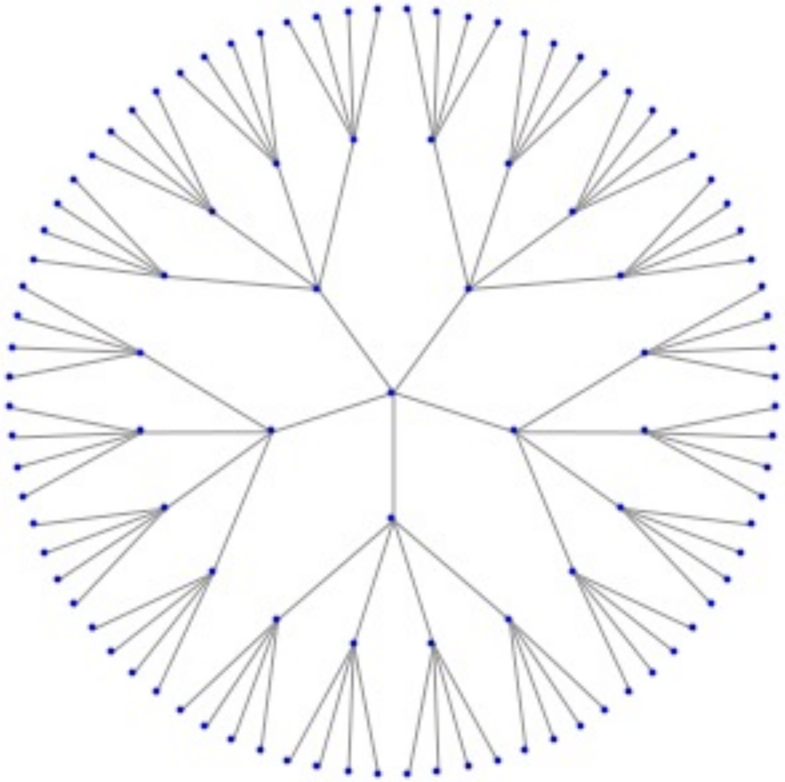
A2. To try to understand underlying **principles** or organizing **laws**

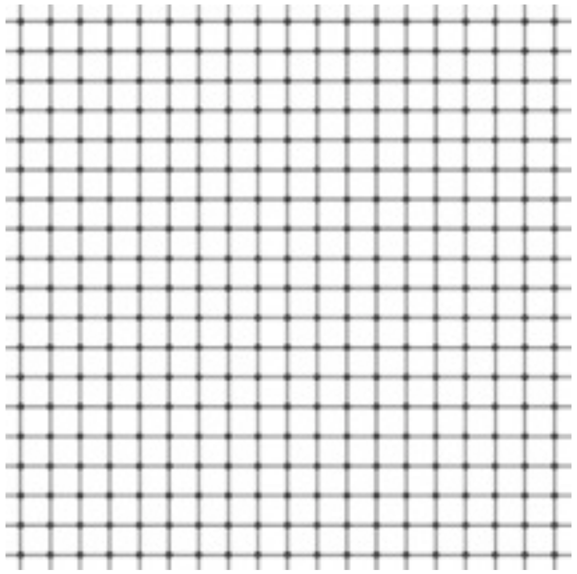
Build **simplified** networks preserving/destroying certain features or properties

What's similar/ different in these **reduced** networks (**models**)?

Trees and lattices

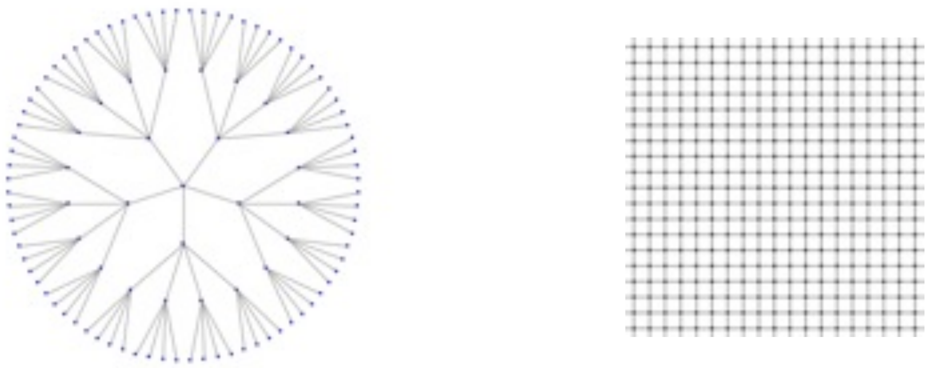
Simple models





Trees and lattices

Simple models



Completely regular or **ordered**

Randomness

replace **overwhelming details** with simple
probabilistic rules (coin flips)

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

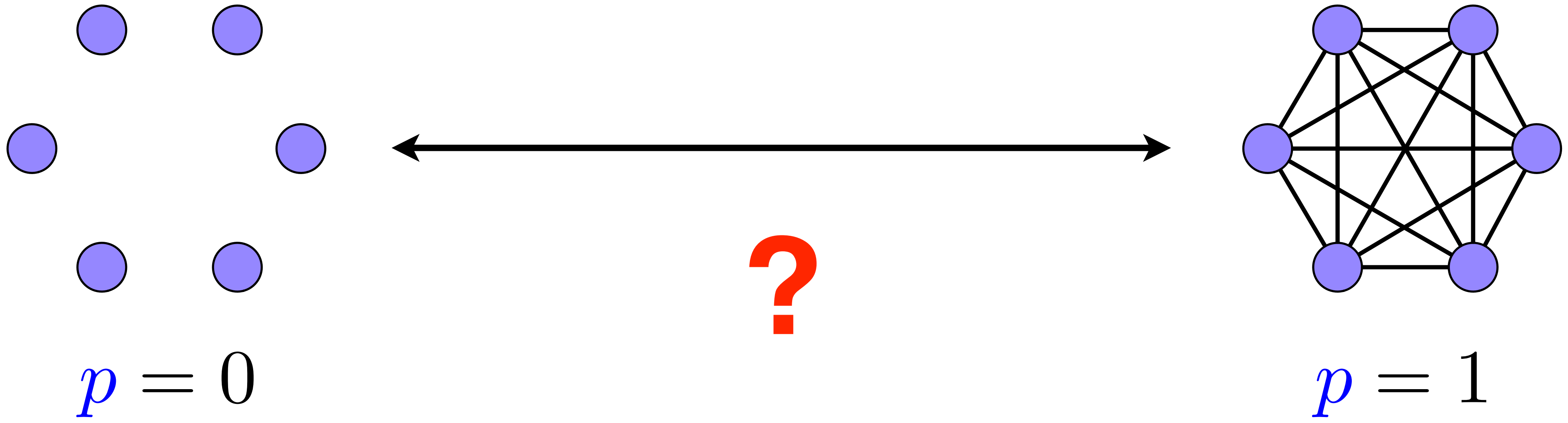
1. Start with an **empty graph** of N nodes

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

Erdős-Rényi Graph

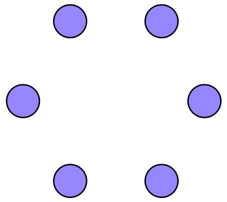
1. Start with an **empty graph** of *N* nodes

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

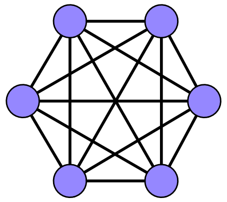


Erdős-Rényi Graph

$p = 0$



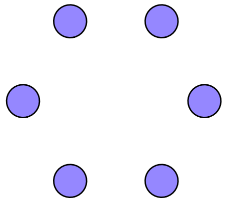
?



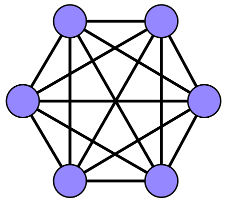
$p = 1$

Erdős-Rényi Graph

$p = 0$



?

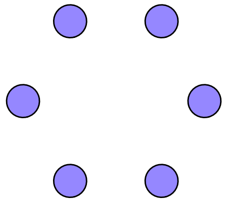


$p = 1$

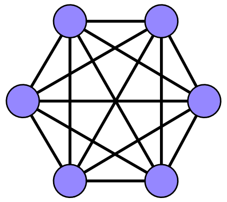
If $Np < 1$

Erdős-Rényi Graph

$p = 0$

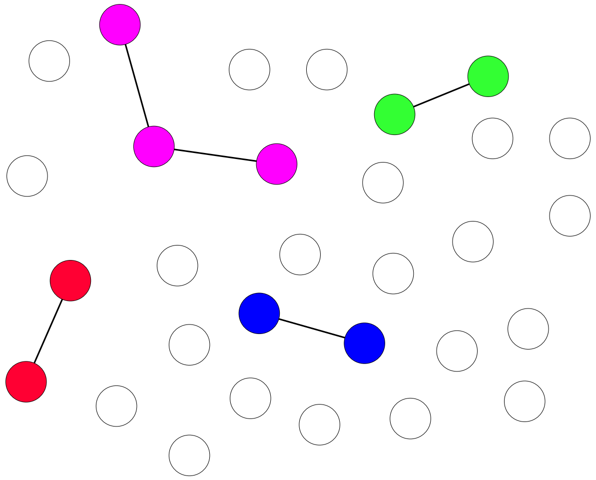


?

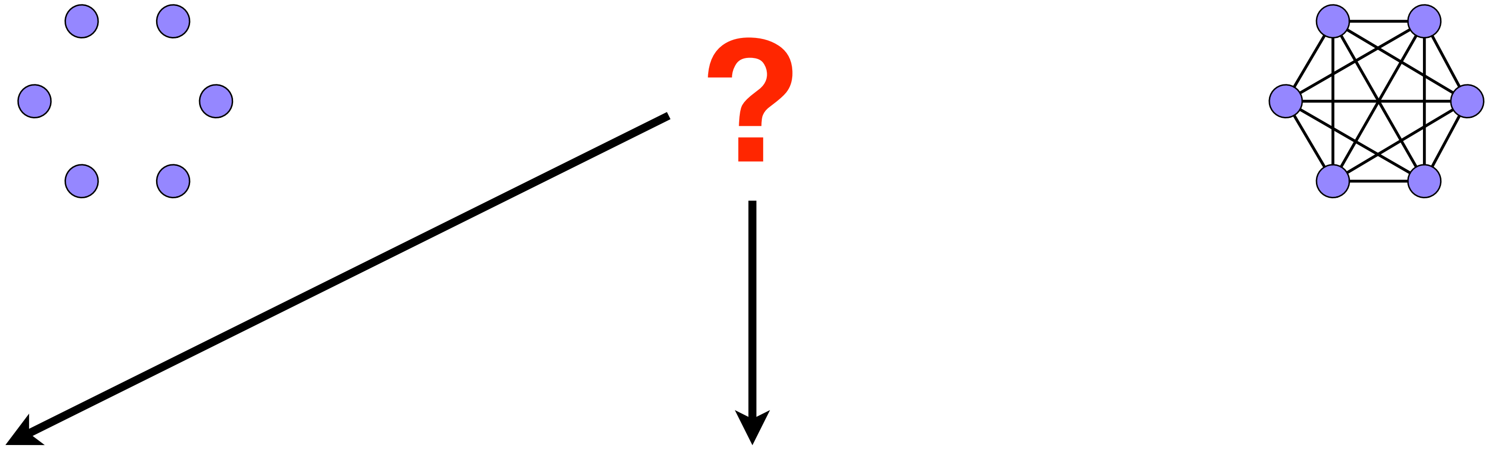


$p = 1$

If $Np < 1$



Erdős-Rényi Graph



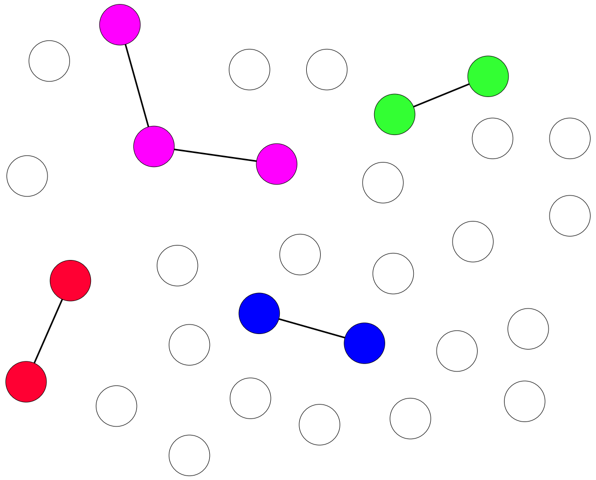
$p = 0$

?

$p = 1$

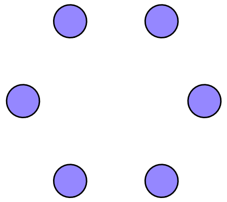
If $Np < 1$

If $Np = 1$

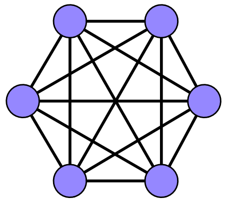


Erdős-Rényi Graph

$p = 0$



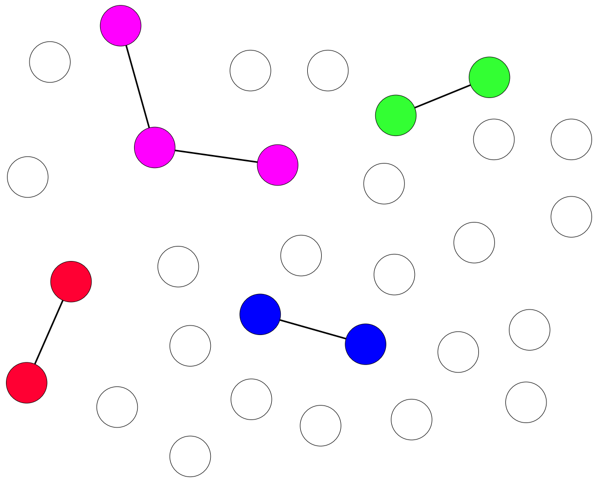
?

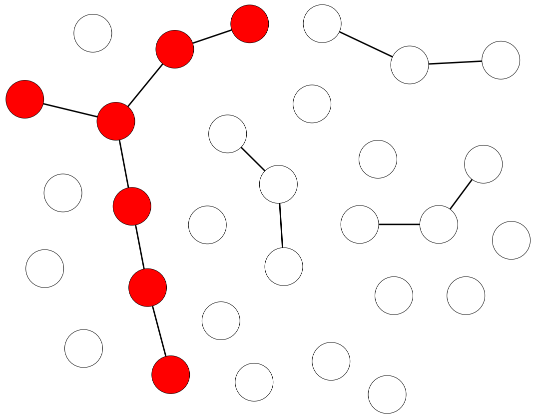


$p = 1$

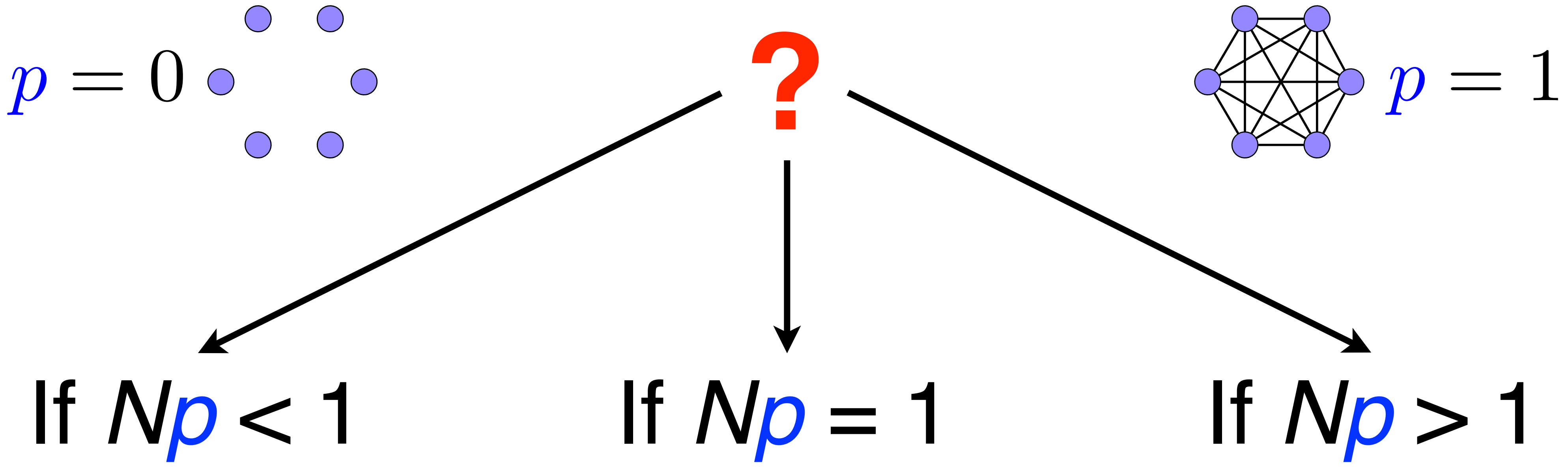
If $Np < 1$

If $Np = 1$





Erdős-Rényi Graph



$p = 0$

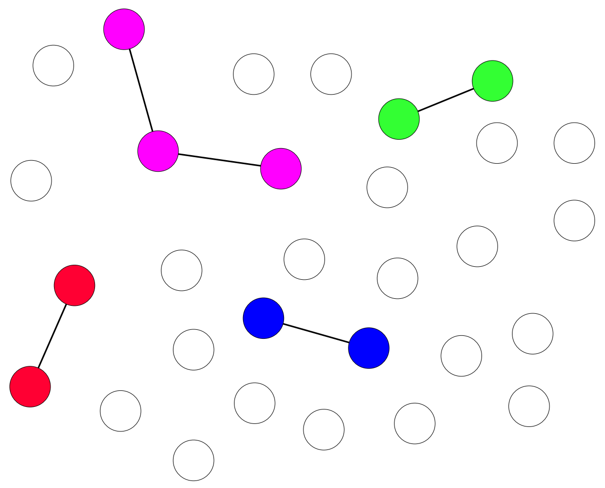
?

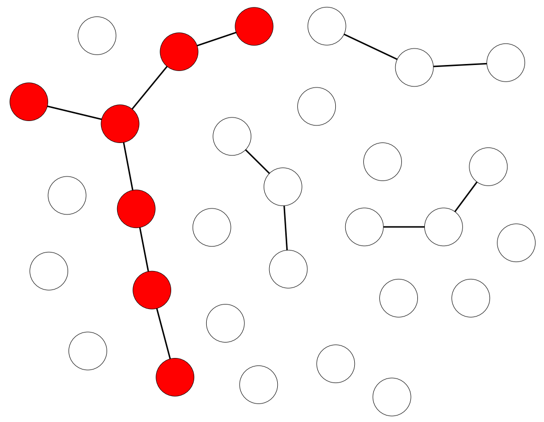
$p = 1$

If $Np < 1$

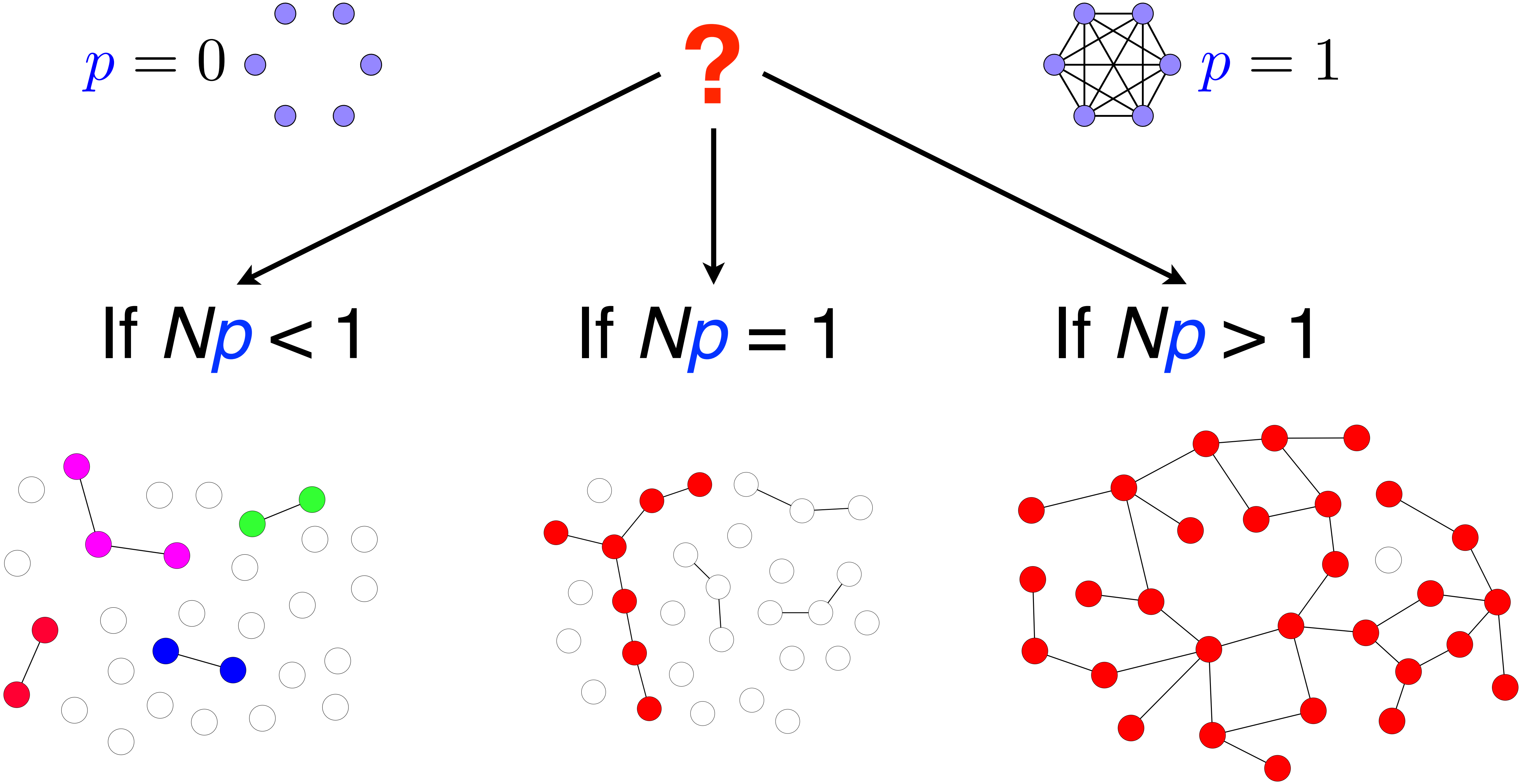
If $Np = 1$

If $Np > 1$





Erdős-Rényi Graph



$p = 0$

?

$p = 1$

If $Np < 1$

If $Np = 1$

If $Np > 1$

Erdős-Rényi Graph

Degree distribution

Erdős-Rényi Graph

Degree distribution

- Np looks important

Erdős-Rényi Graph

Degree distribution

Np looks important ⟶ What is it?

Erdős-Rényi Graph

Degree distribution

Np looks important → What is it?

↓

average degree

Erdős-Rényi Graph

Degree distribution

Np looks important → What is it?

↓

$\langle k \rangle = (N\text{-}1)p \xleftarrow{\text{well.....}}$ **average** degree

Erdős-Rényi Graph

Degree distribution

Np looks important → What is it?

↓

$\langle k \rangle = (N\text{-}1)p$ ← well..... **average** degree

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Erdős-Rényi Graph

Degree distribution

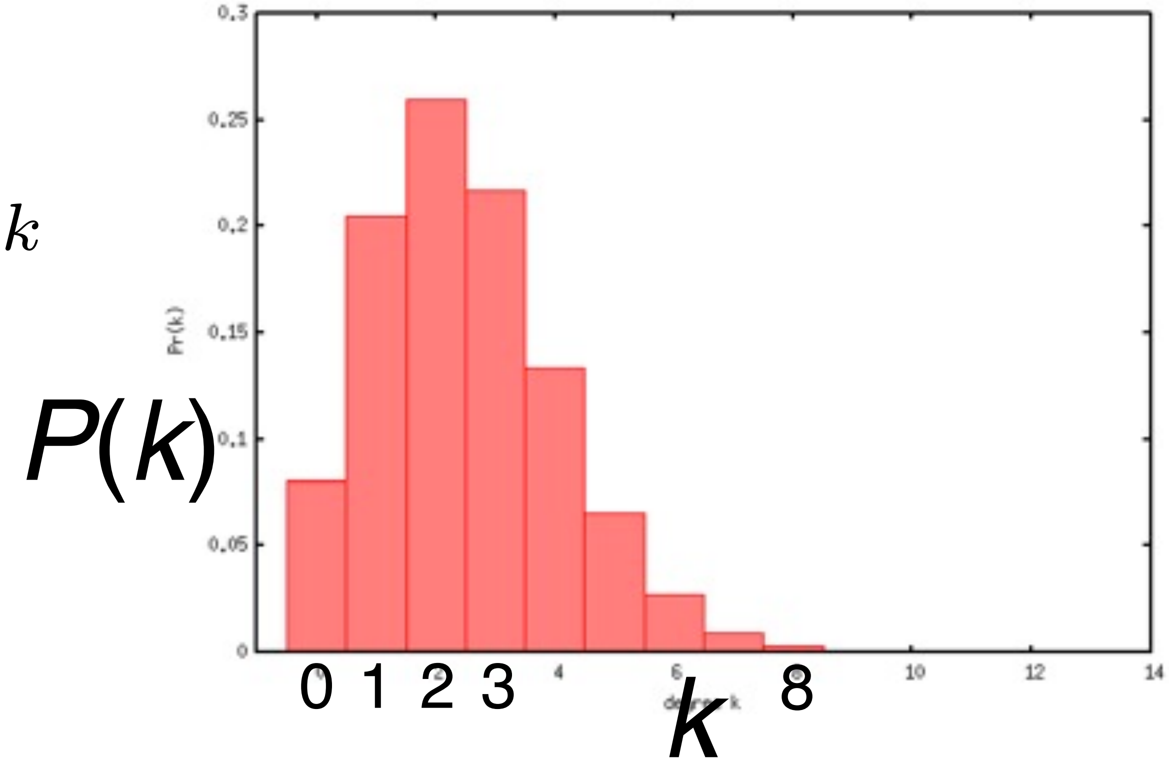
Np looks important → What is it?

↓

well..... ← **average** degree

$\langle k \rangle = (N\text{-}1)p$

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$



Erdős-Rényi Graph

Degree distribution

Np looks important → What is it?

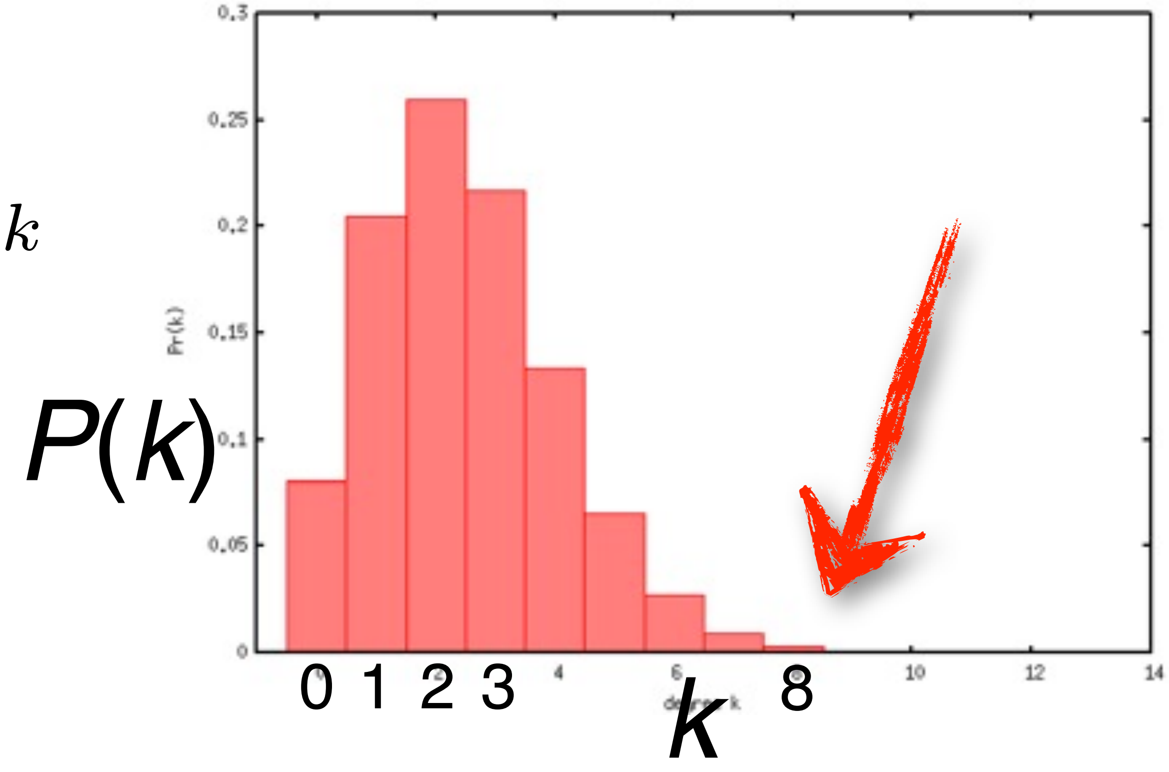
↓

average degree

well..... ←

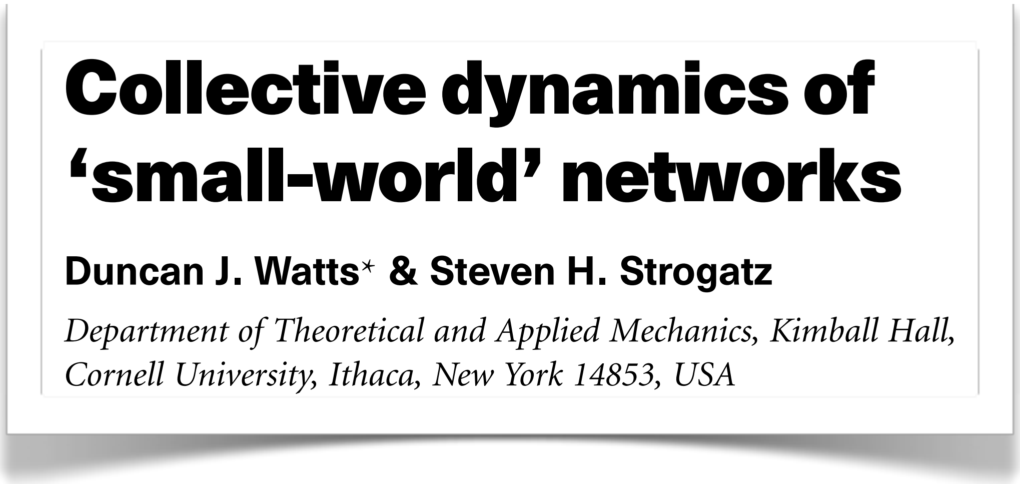
$\langle k \rangle = (N\text{-}1)p$

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$



Watts-Strogatz

Entering the **modern era**, 1998

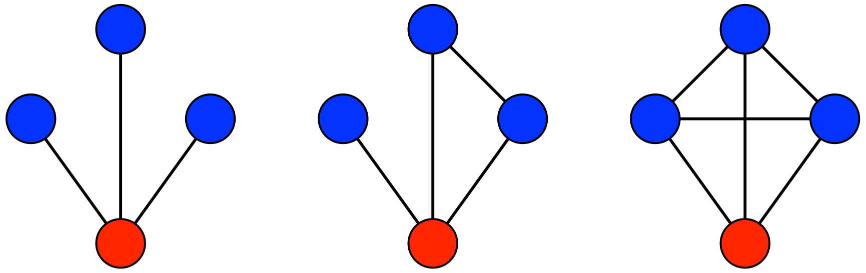


Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

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

Recall:



Introduced clustering coefficient

Nature, 1998

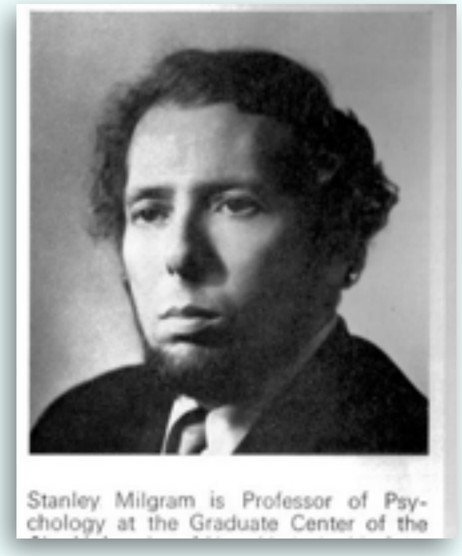
Watts-Strogatz Model

Besides triangles, they were interested in **distances**

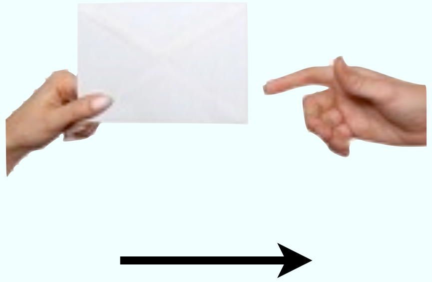
Watts-Strogatz Model

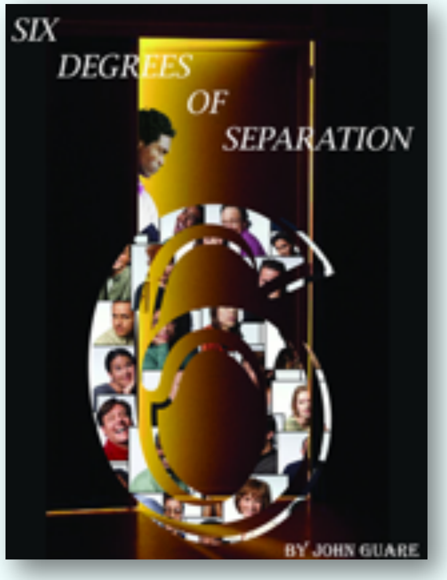
Besides triangles, they were interested in **distances**

1960s **Milgram** asked “How far apart are we?”



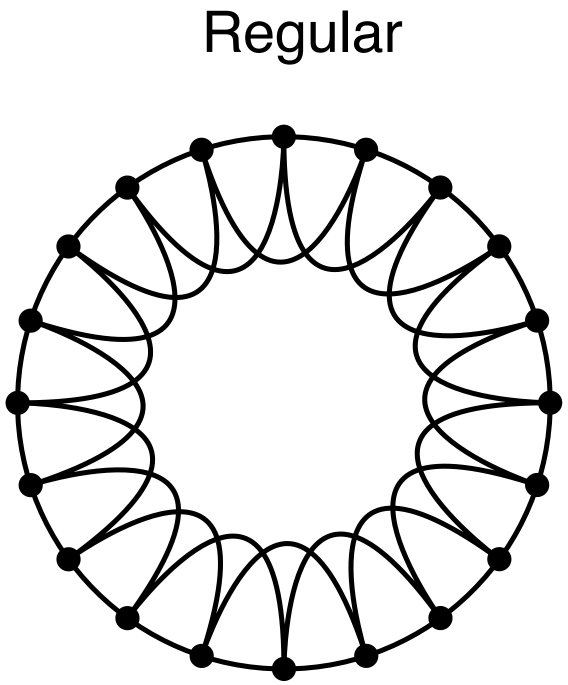
Stanley Milgram is Professor of Psychology at the Graduate Center of the





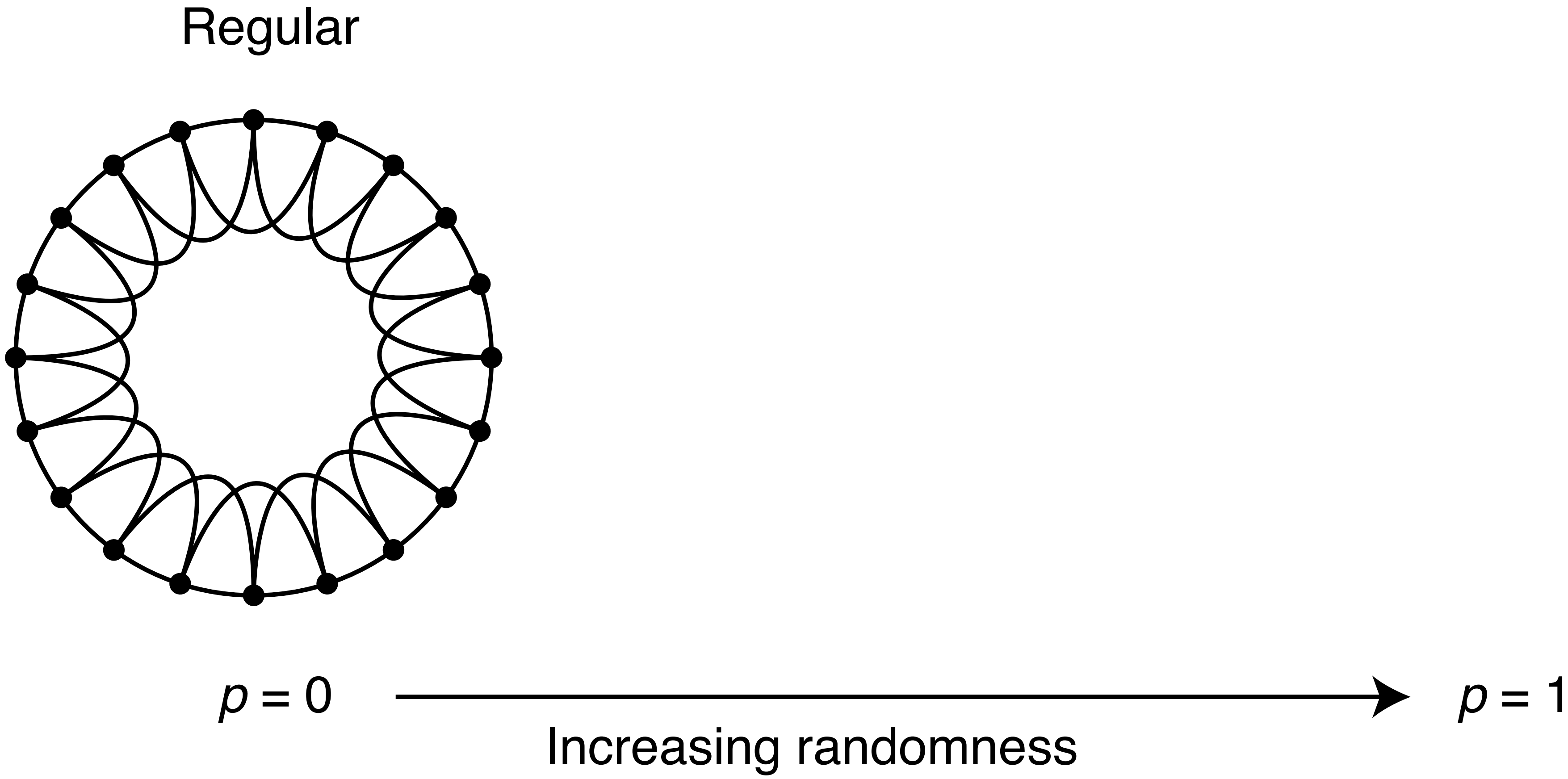
Watts-Strogatz Model

Besides triangles, they were interested in **distances**



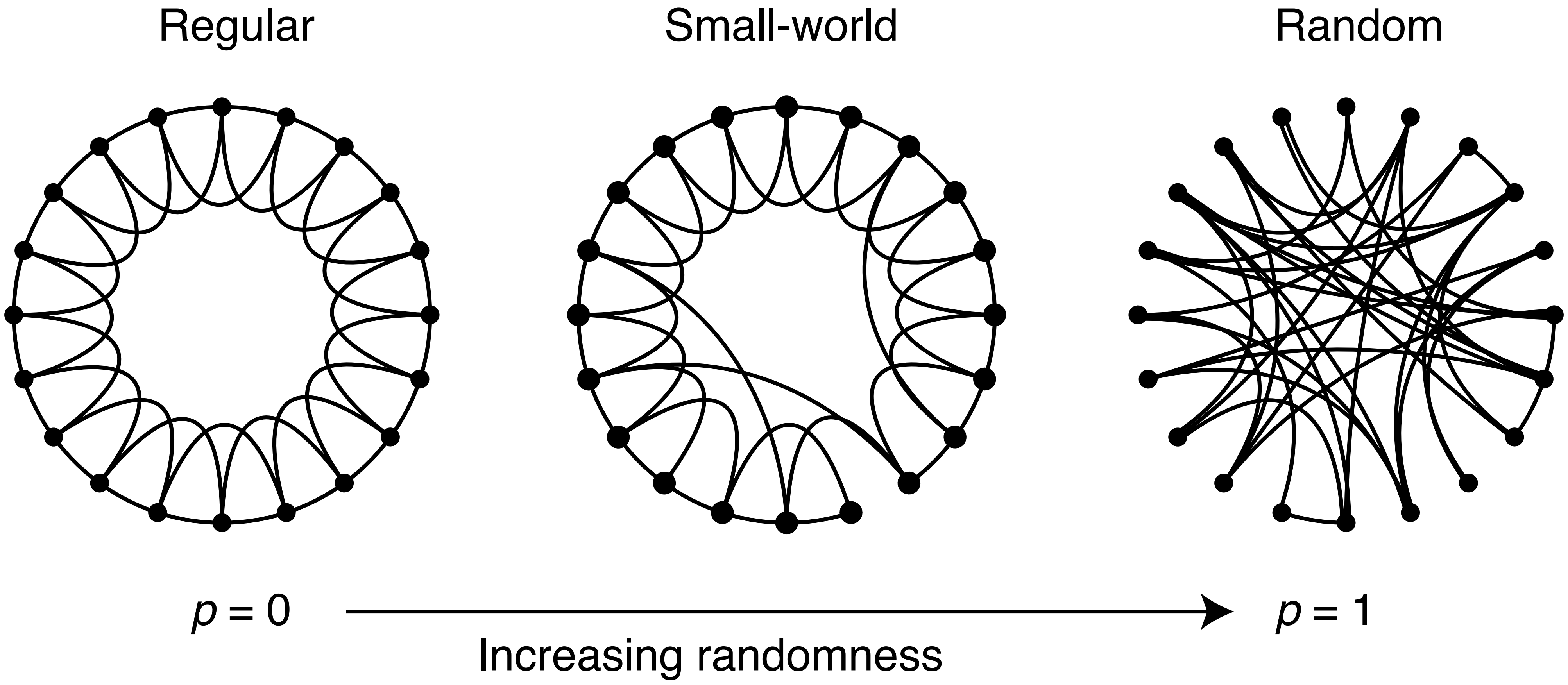
Watts-Strogatz Model

Besides triangles, they were interested in **distances**



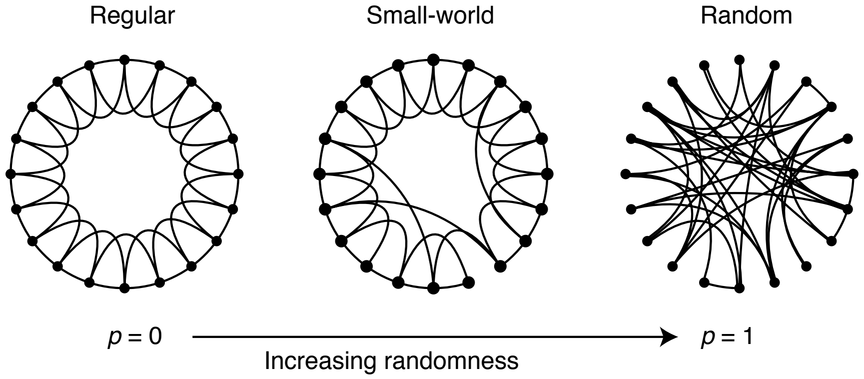
Watts-Strogatz Model

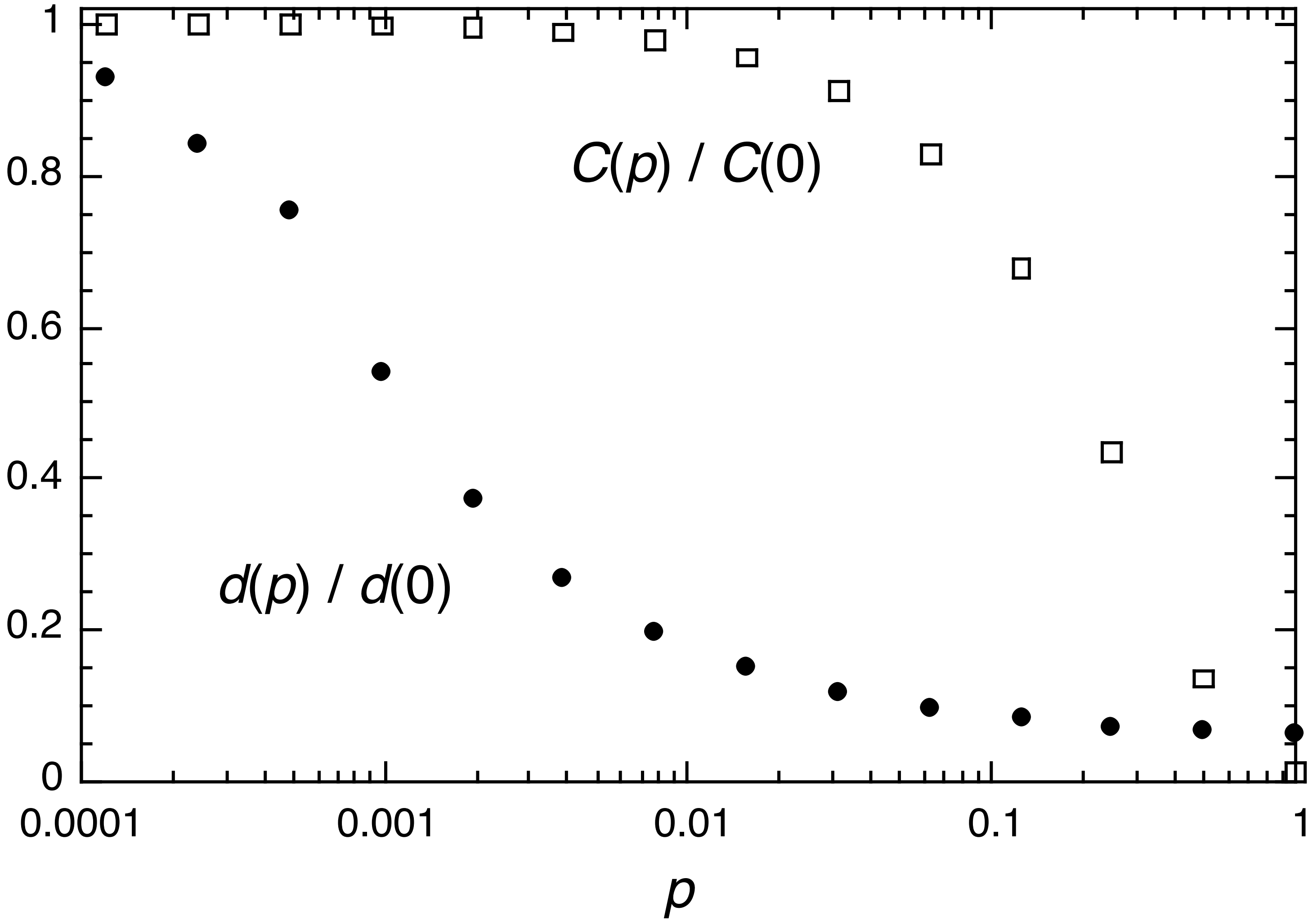
Besides triangles, they were interested in **distances**



Watts-Strogatz Model

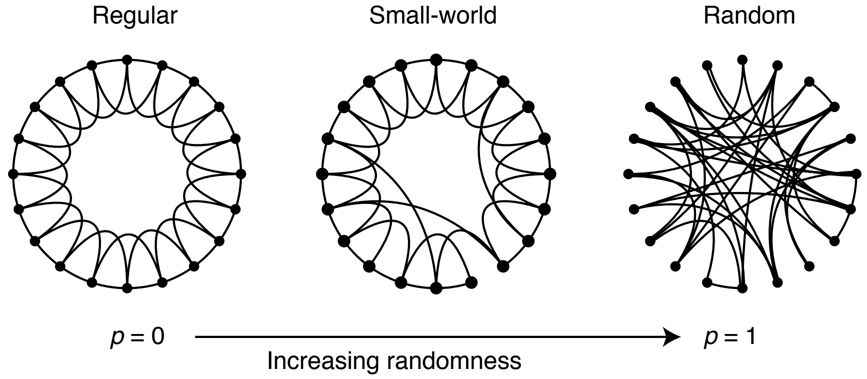
Besides triangles, they were interested in **distances**

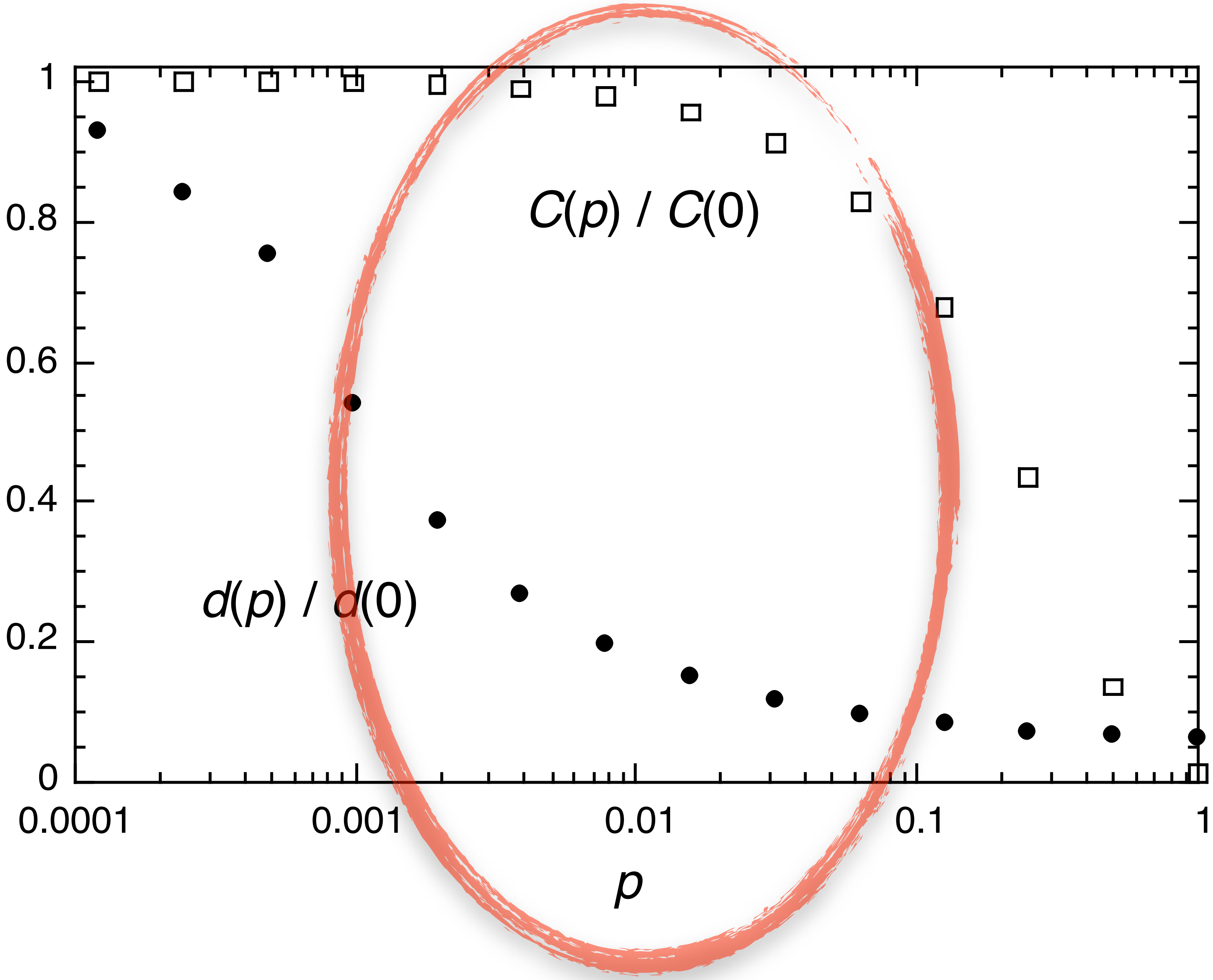




Watts-Strogatz Model

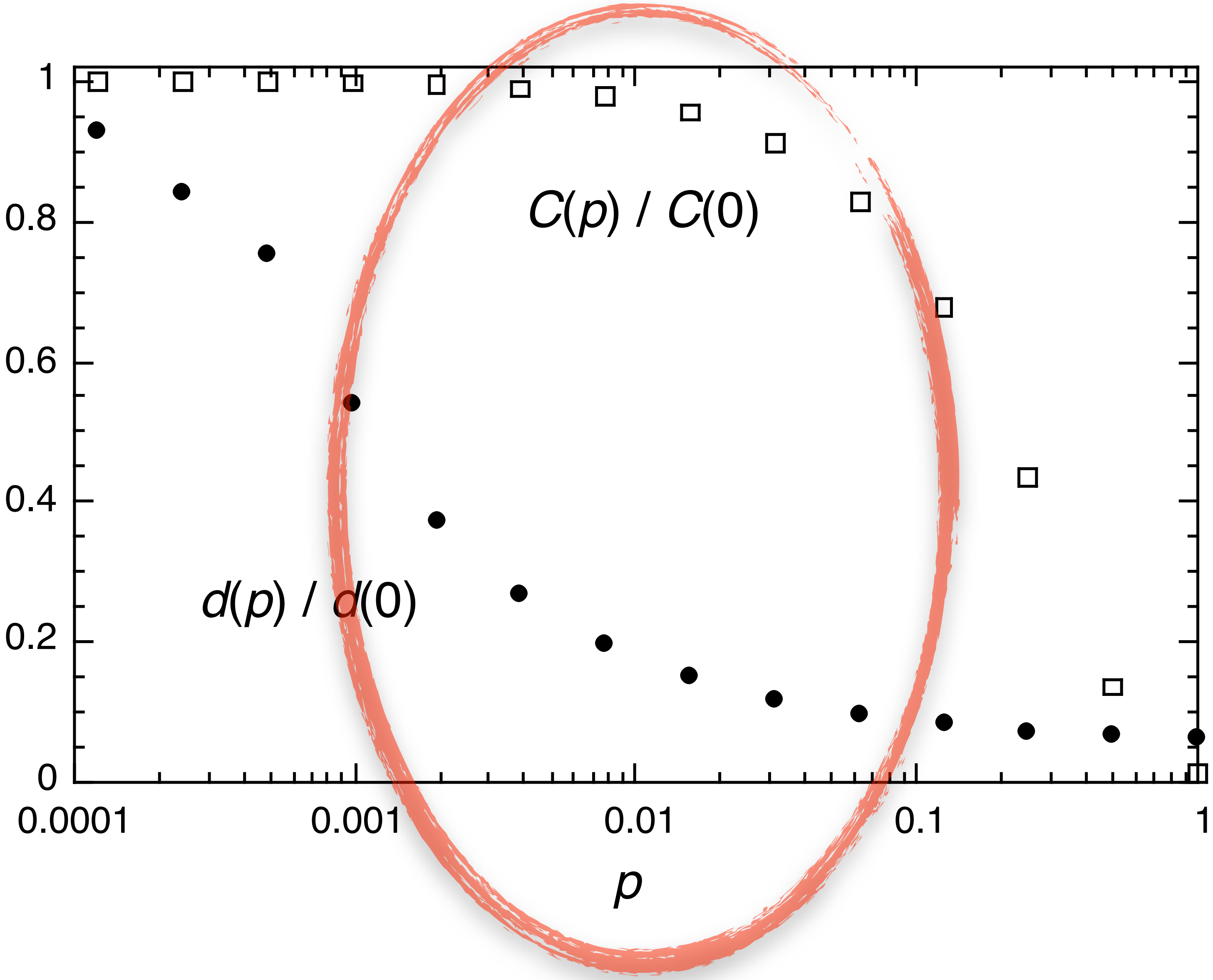
Besides triangles, they were interested in **distances**





Watts-Strogatz Model

Besides triangles, they were interested in **distances**

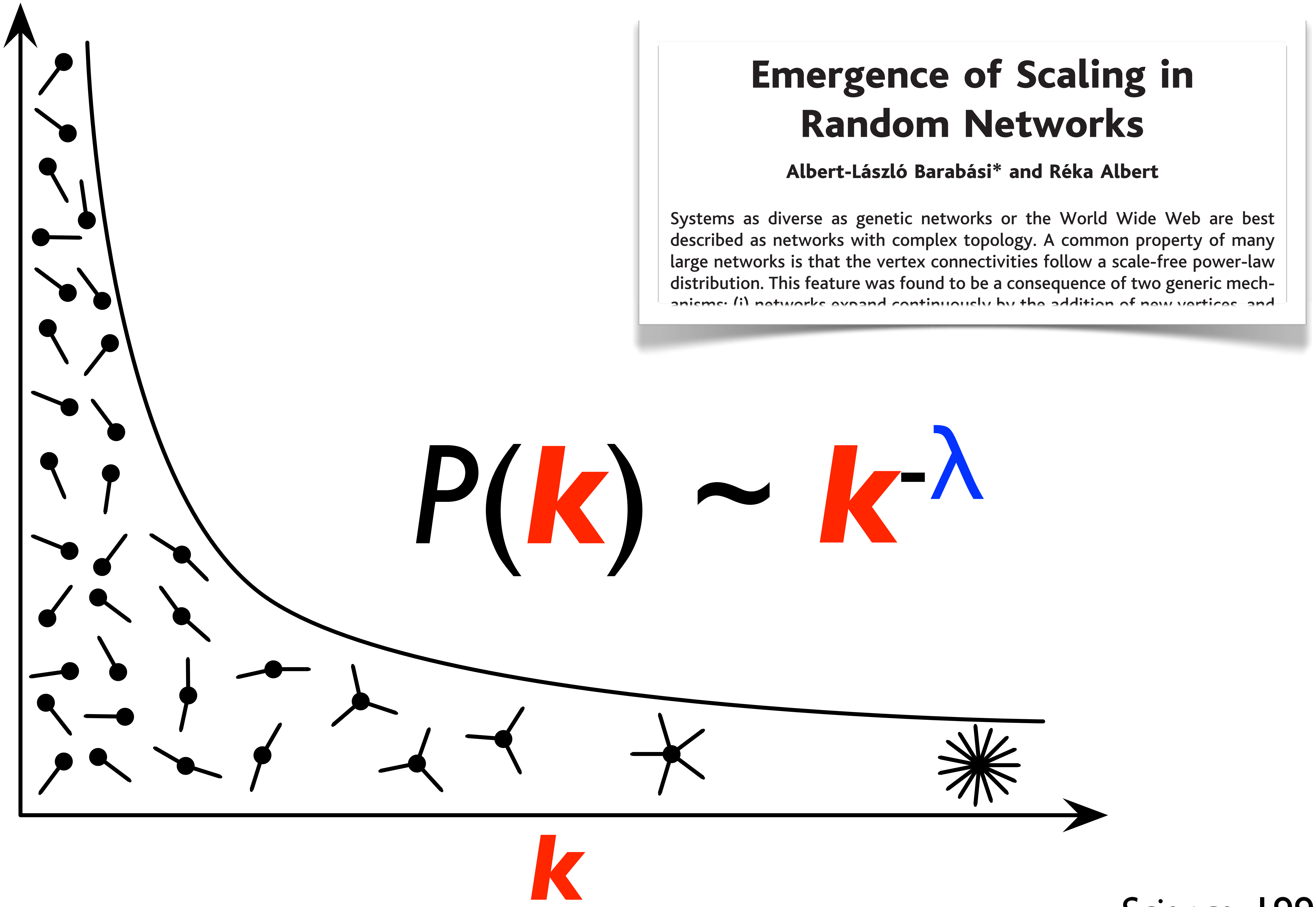


Small-world:

Diameter much smaller than number of nodes N

$$D \sim \log(N)$$

Scale-free networks



Emergence of Scaling in Random Networks

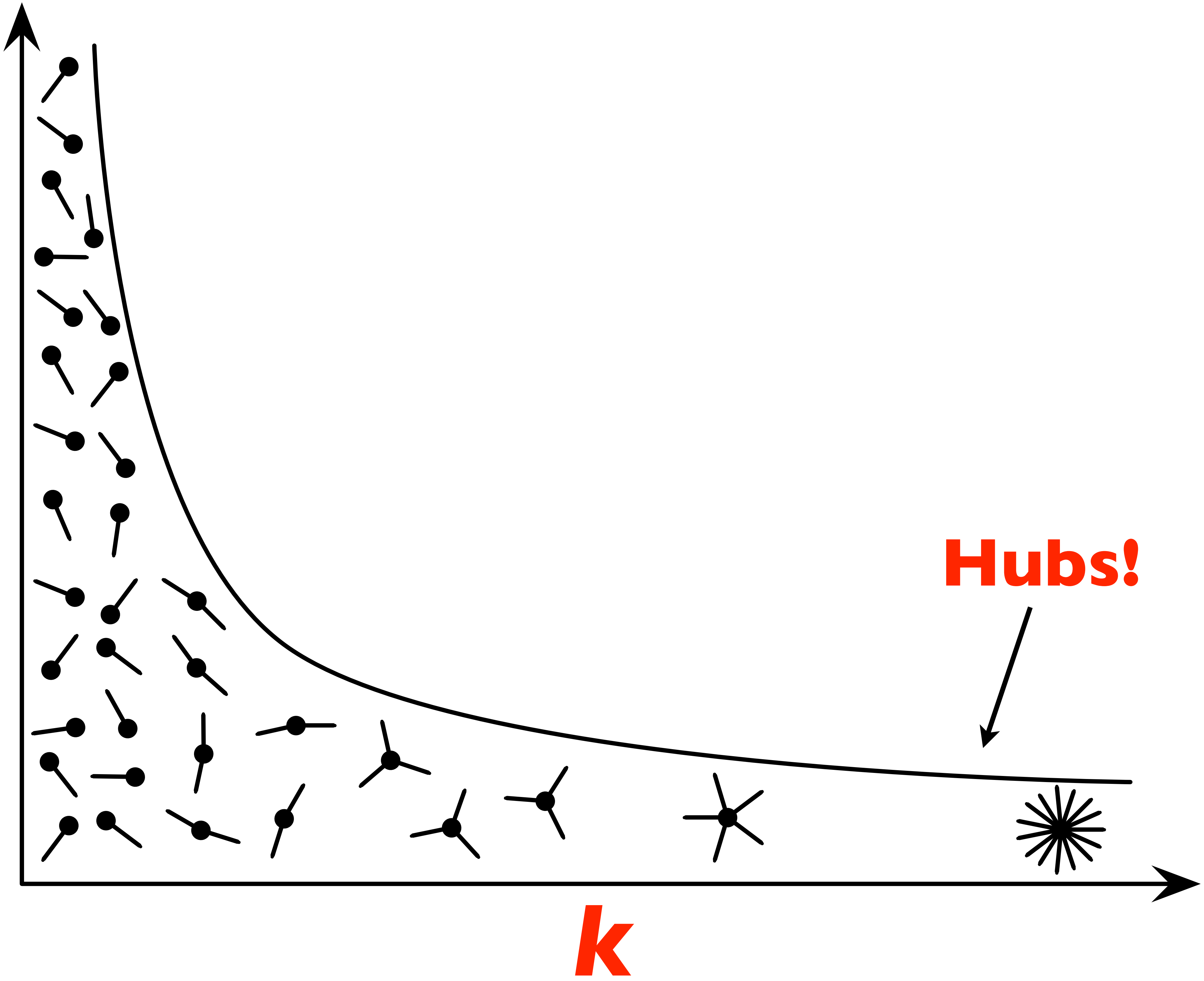
Albert-László Barabási* and Réka Albert

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

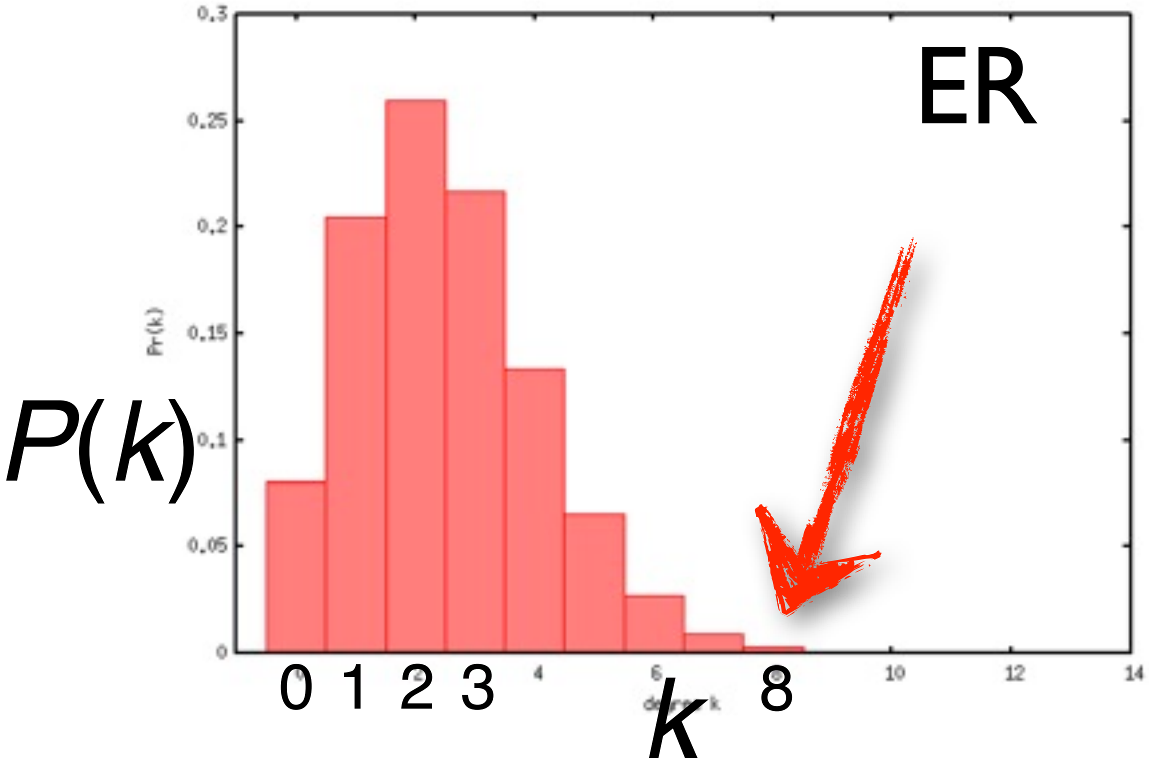


Science, 1999

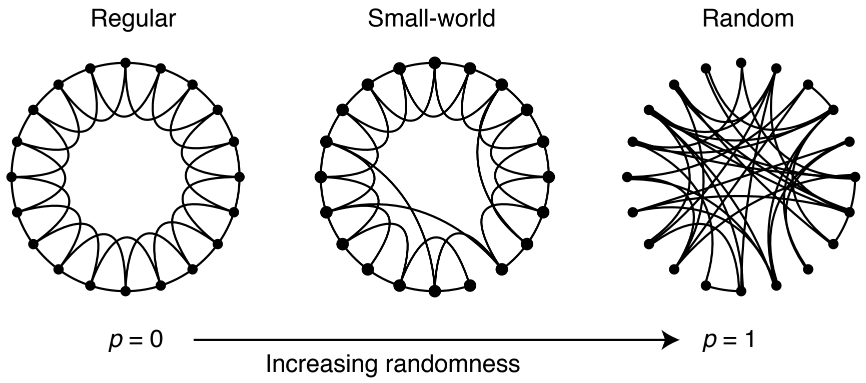
Scale-free networks

Hubs!

Earlier models → No hubs



WS



Barabási-Albert Model

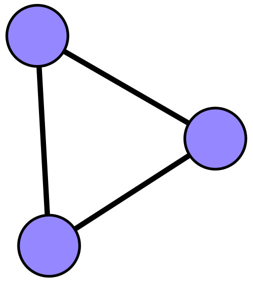
Barabási-Albert Model

⟶ **Growing** network model

Barabási-Albert Model

⟶ **Growing** network model

1. start with a **seed graph**



Barabási-Albert Model

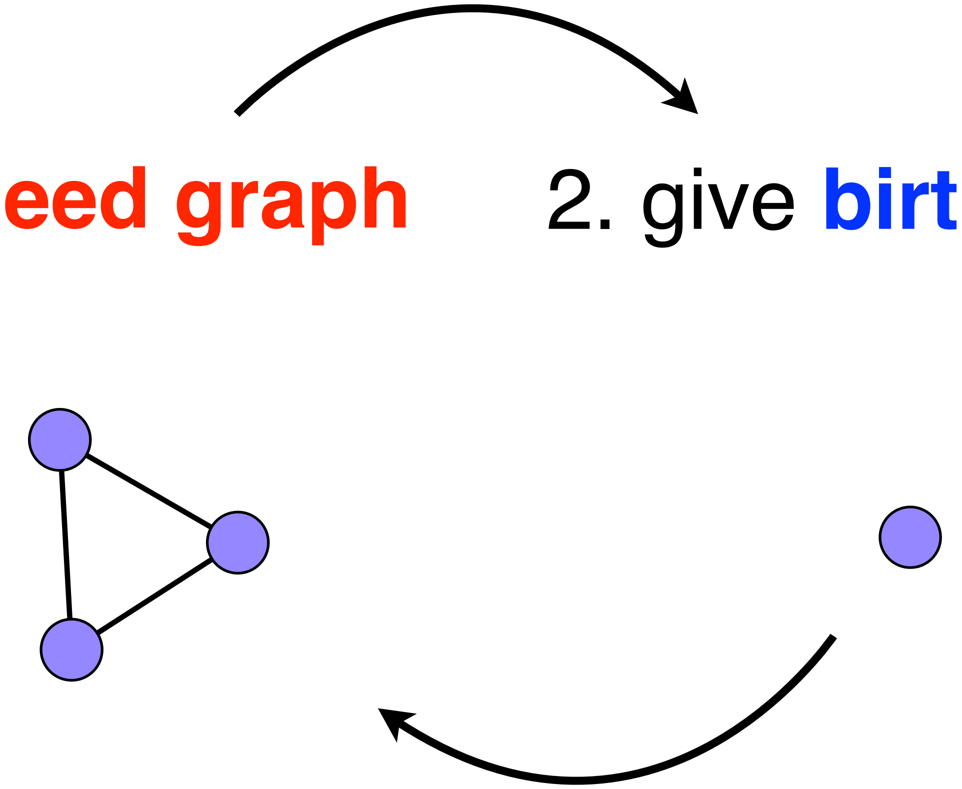
→ **Growing** network model

1. start with a **seed graph**

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

Barabási-Albert Model

→ **Growing** network model



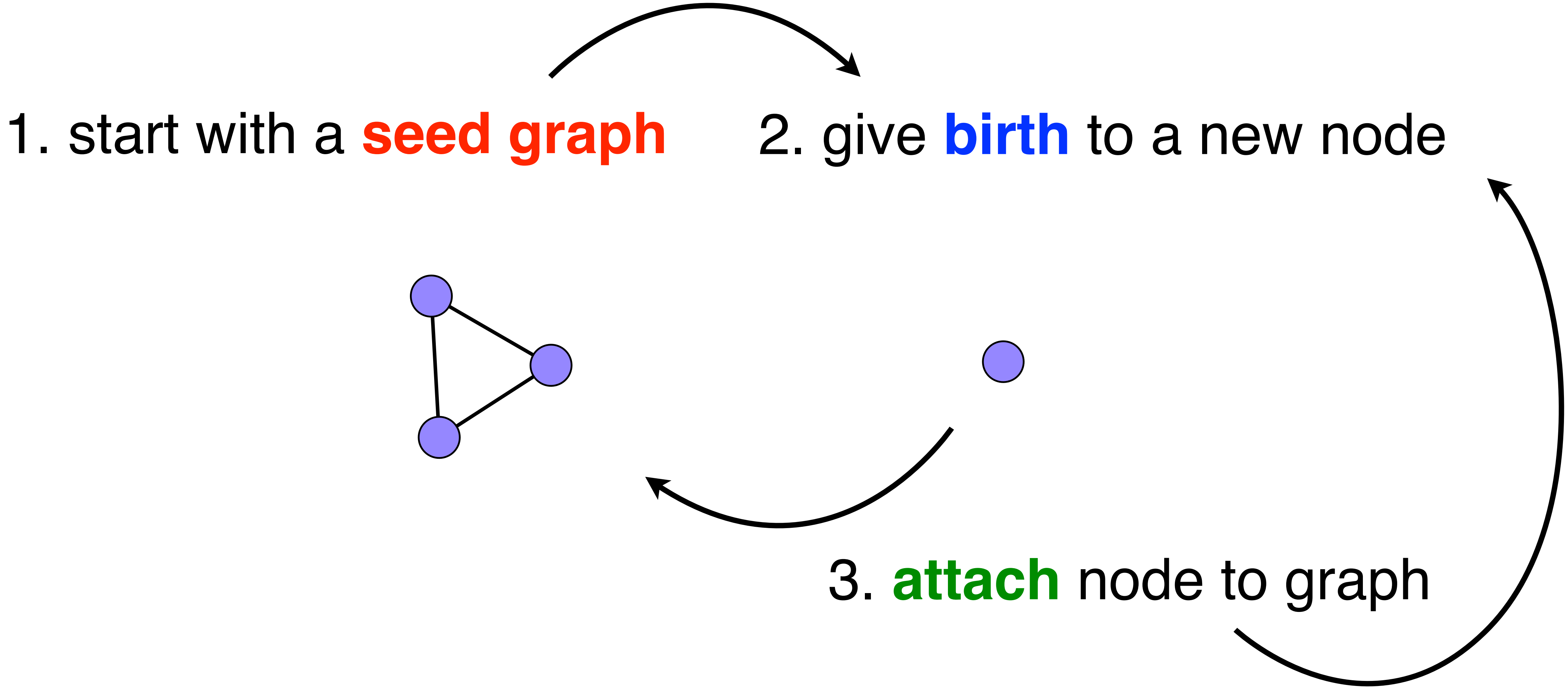
1. start with a **seed graph**

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

3. **attach** node to graph

Barabási-Albert Model

→ **Growing** network model

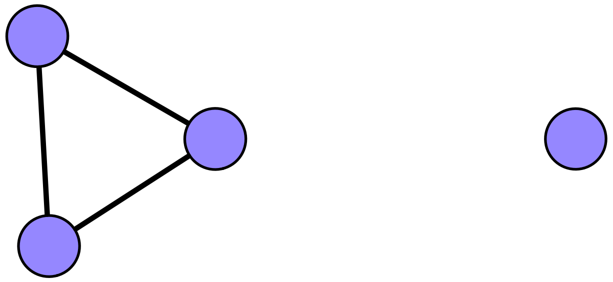


1. start with a **seed graph**

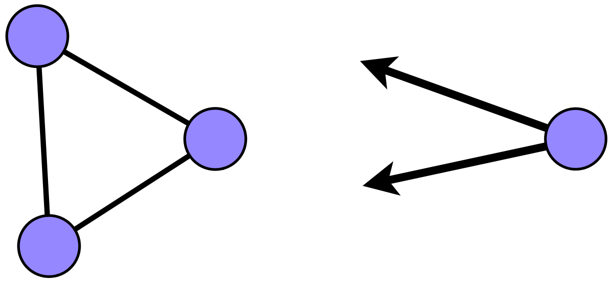
2. give **birth** to a new node

3. **attach** node to graph

Barabási-Albert Model

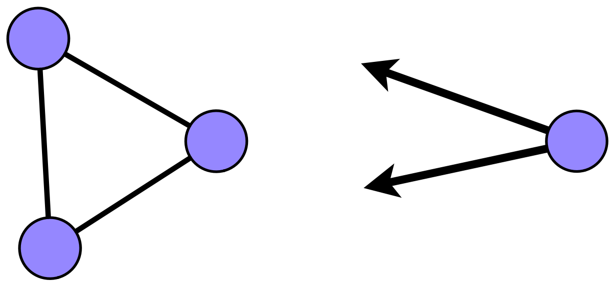


Barabási-Albert Model



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

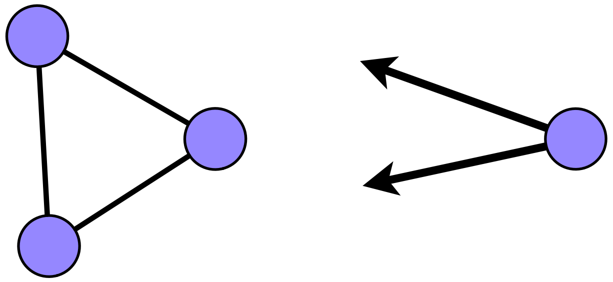
Barabási-Albert Model



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

How?

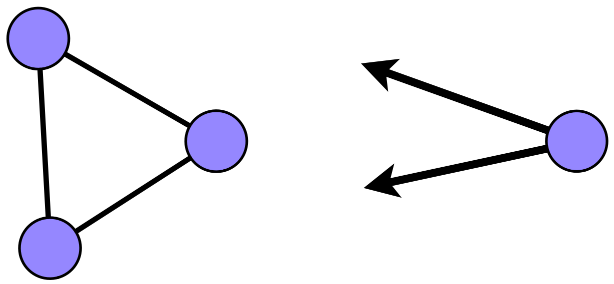
Barabási-Albert Model



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

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

Barabási-Albert Model



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

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

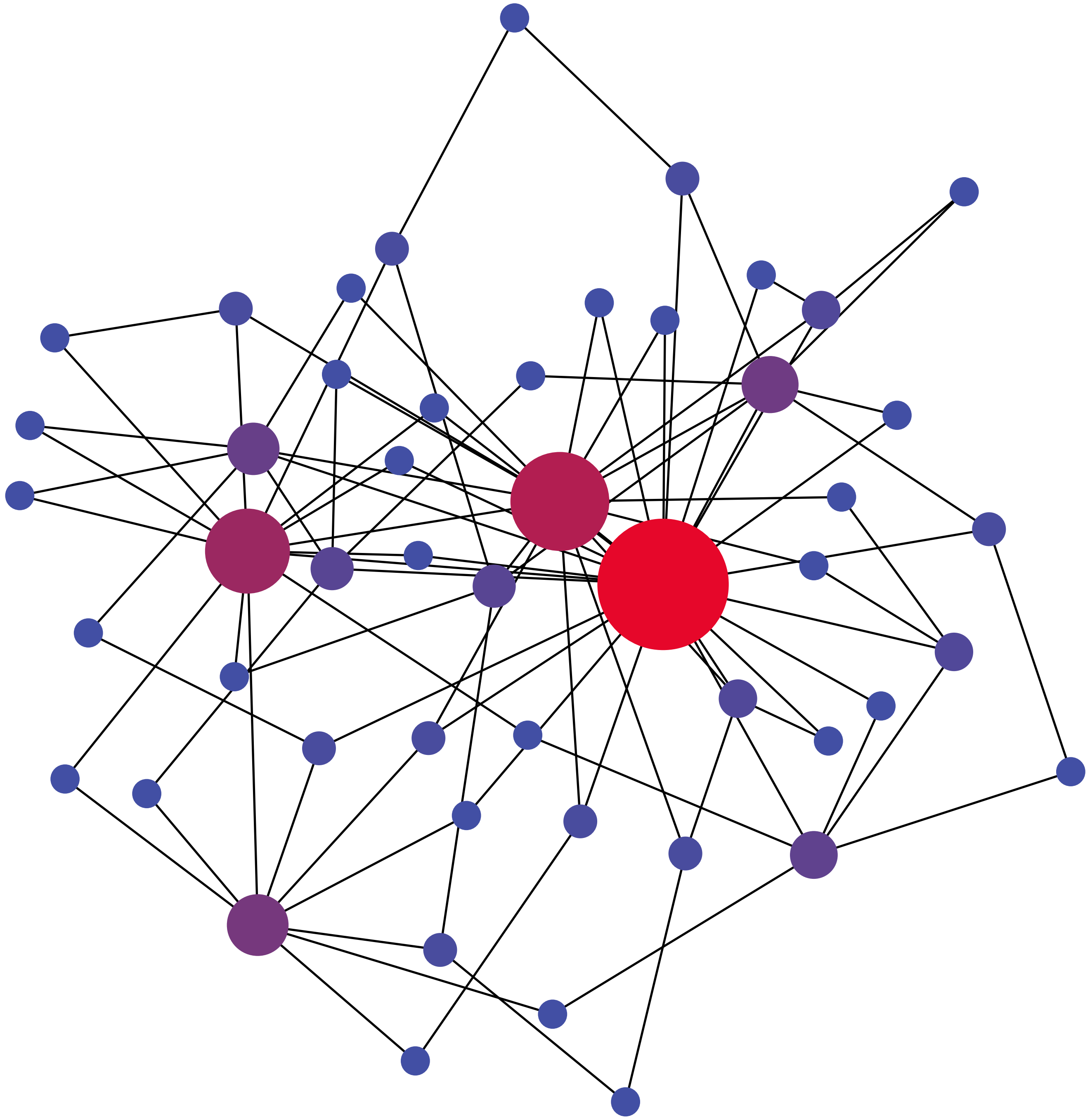
Link to existing node *i* with probability:

$$P_{\text{attach}}(i) = \frac{k_i}{\sum_j k_j}$$

Barabási-Albert Model

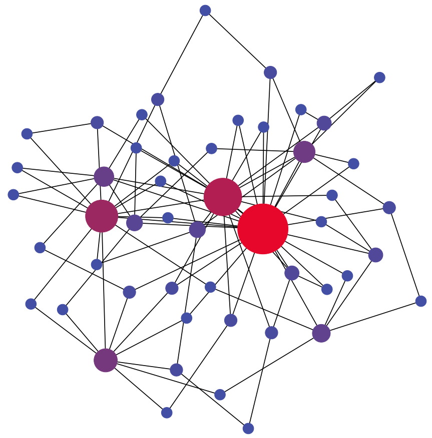
Preferential Attachment

$$P_{\text{attach}}(i) = \frac{k_i}{\sum_j k_j}$$



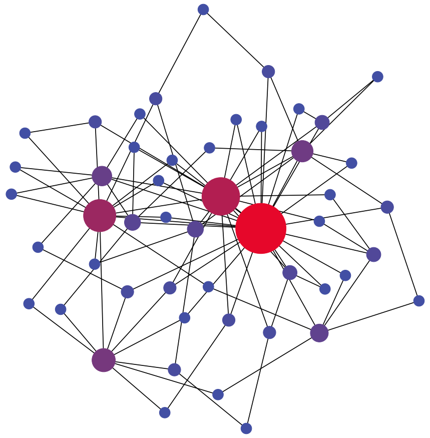
Barabási-Albert Model

Preferential Attachment + **Growth** =



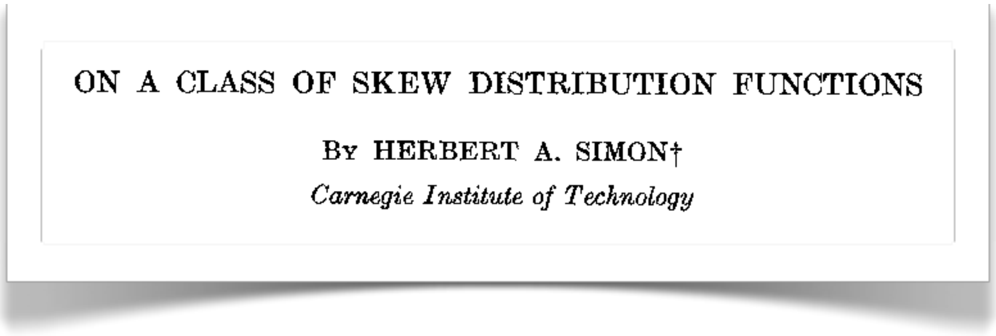
Barabási-Albert Model

Preferential Attachment + **Growth** =



Not the first:

1955!



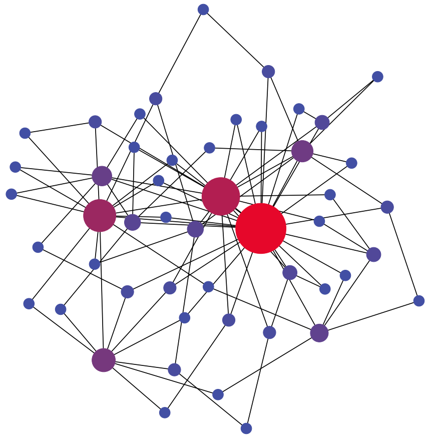
ON A CLASS OF SKEW DISTRIBUTION FUNCTIONS

By HERBERT A. SIMON†

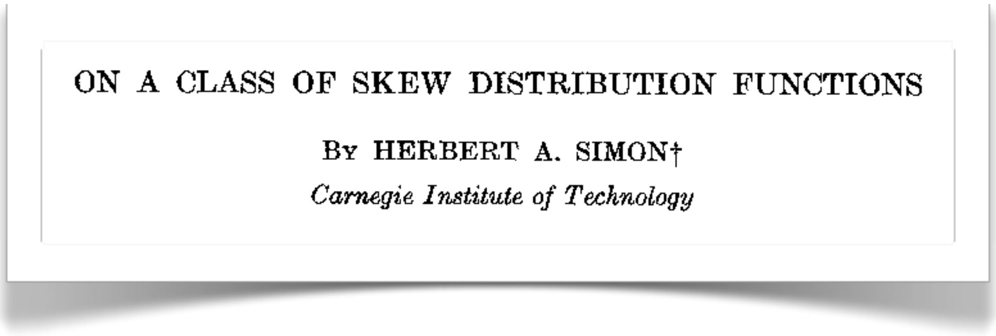
Carnegie Institute of Technology

Barabási-Albert Model

Preferential Attachment + **Growth** =



Not the first: 1955!

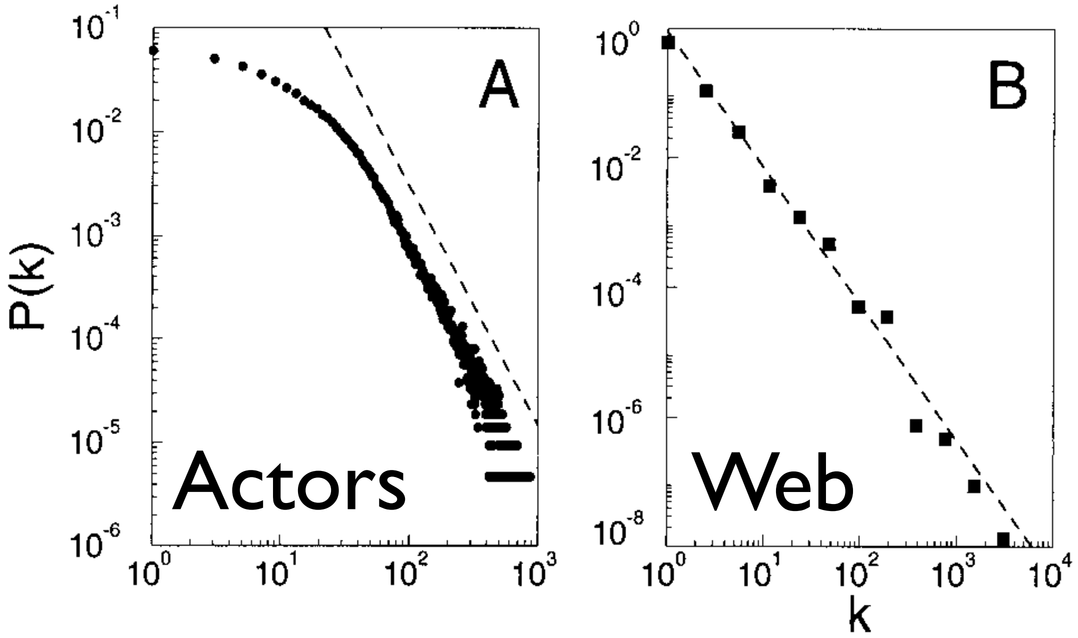


ON A CLASS OF SKEW DISTRIBUTION FUNCTIONS

By HERBERT A. SIMON†

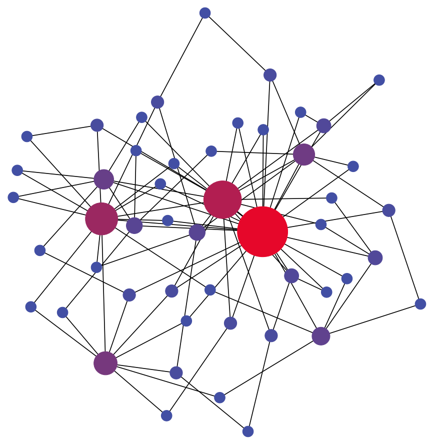
Carnegie Institute of Technology

But they found it in **new systems**



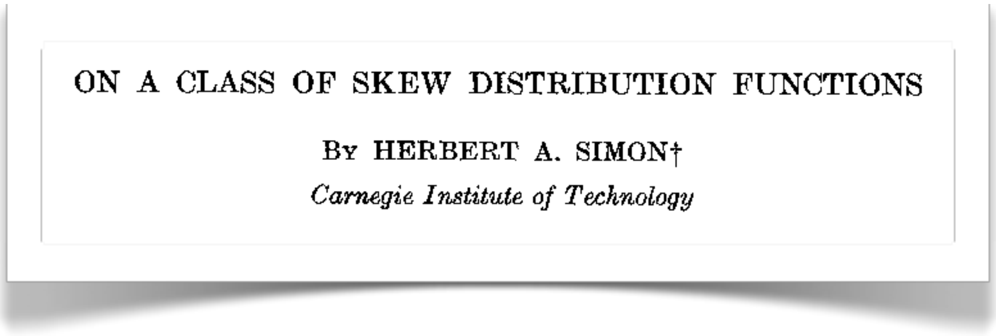
Barabási-Albert Model

Preferential Attachment + **Growth** =



Not the first:

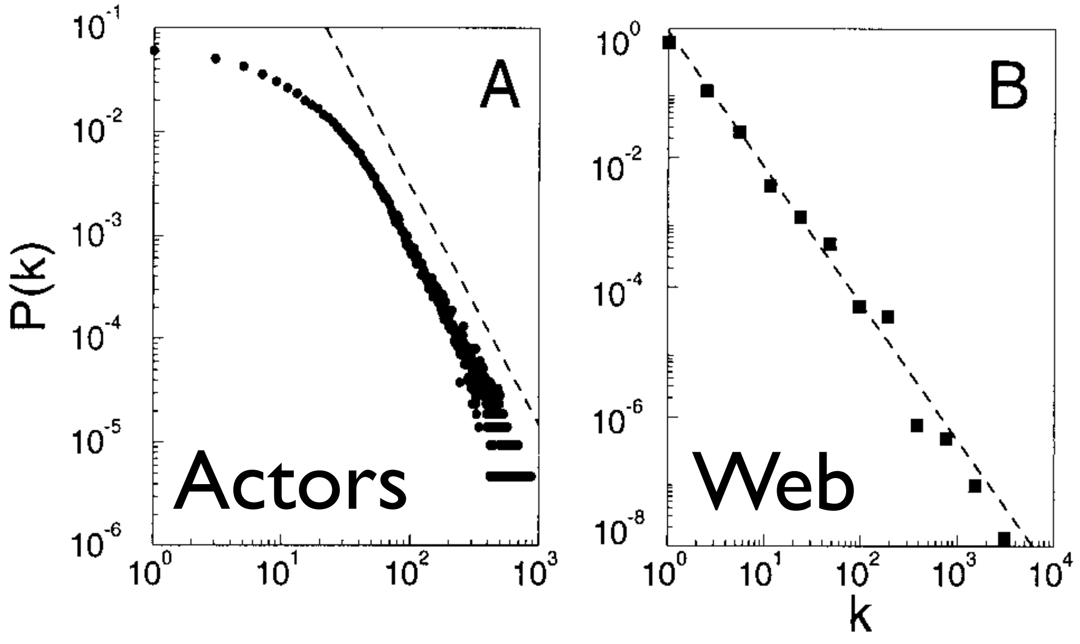
1955!

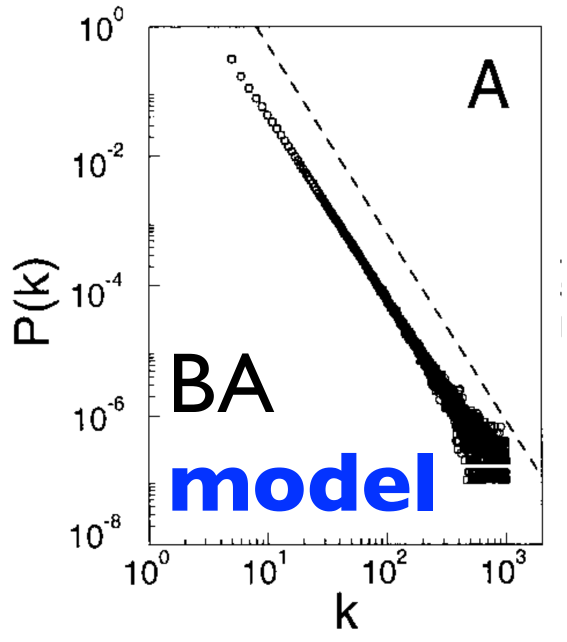


ON A CLASS OF SKEW DISTRIBUTION FUNCTIONS

By HERBERT A. SIMON†

Carnegie Institute of Technology





Heavy-tailed,
scale-free,
power-law,
degree distributions

Where does **PA** come from?

Preferential Attachment

$$P_{\text{attach}}(i) = \frac{k_i}{\sum_j k_j}$$

Where does **PA** come from?

Preferential Attachment

$$P_{\text{attach}}(i) = \frac{k_i}{\sum_j k_j}$$

Where does **PA** come from?

Preferential Attachment

$$P_{\text{attach}}(i) = \frac{k_i}{\sum_j k_j}$$

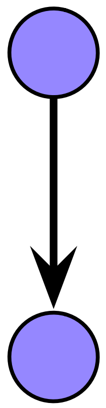
Requires global information

Krapivsky-Redner Model

Redirection model

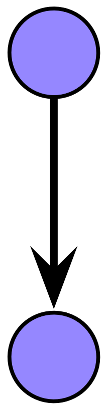
Krapivsky-Redner Model

Redirection model



Krapivsky-Redner Model

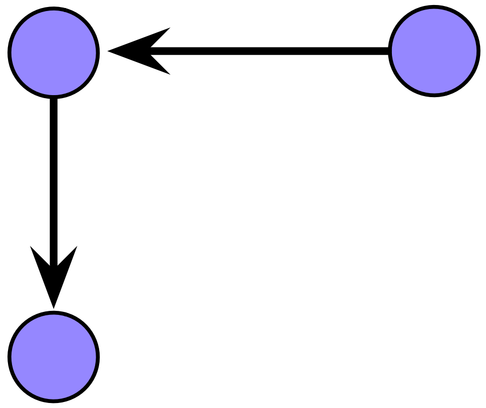
Redirection model



New node: pick a **random** node and attach

Krapivsky-Redner Model

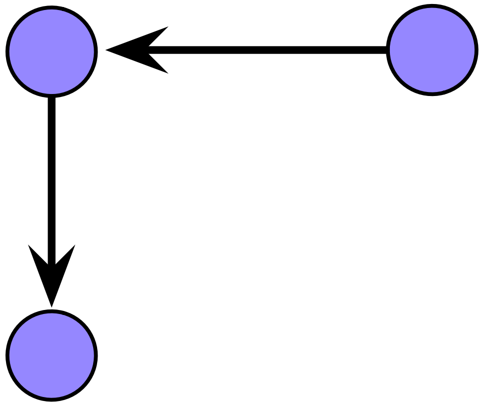
Redirection model



New node: pick a **random** node and attach

Krapivsky-Redner Model

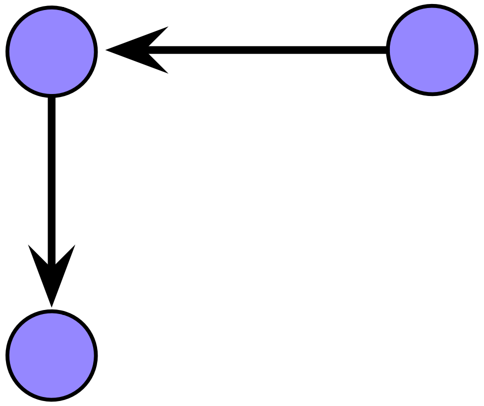
Redirection model



Then, flip a coin

Krapivsky-Redner Model

Redirection model

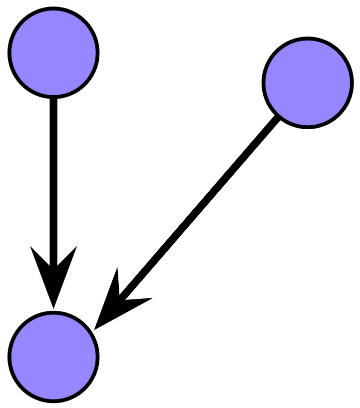


Then, flip a coin

With prob p you redirect the new link to an **ancestor**

Krapivsky-Redner Model

Redirection model

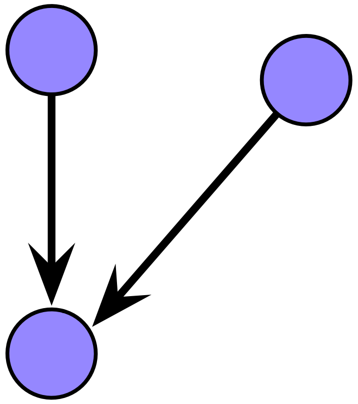


Then, flip a coin

With prob p you redirect the new link to an **ancestor**

Krapivsky-Redner Model

Redirection model



Then, flip a coin

With prob p you redirect the new link to an **ancestor**

Continue

Krapivsky-Redner Model

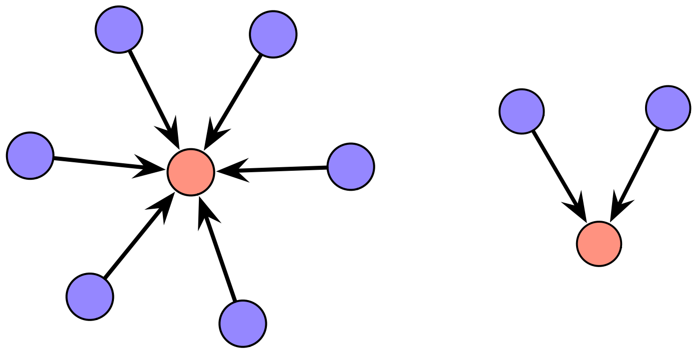
Redirection model

Nodes with more descendants are **more likely** to gain **new** descendants through redirection

Krapivsky-Redner Model

Redirection model

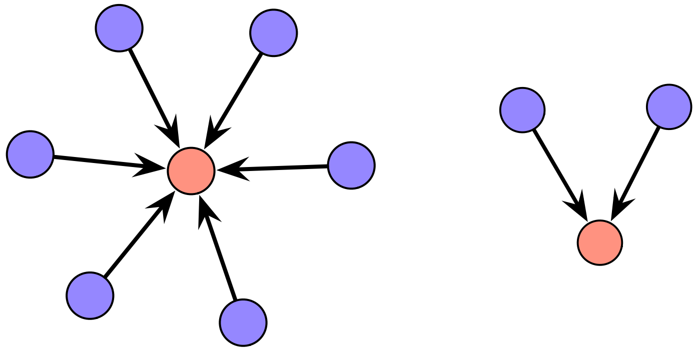
Nodes with more descendants are **more likely** to gain **new** descendants through redirection



Krapivsky-Redner Model

Redirection model

Nodes with more descendants are **more likely** to gain **new** descendants through redirection



Rich-get-richer is **automatic!**

One more

Configuration model

Configuration model

Generate degrees:

Configuration model

Generate degrees:

5

2

1

3

1

2

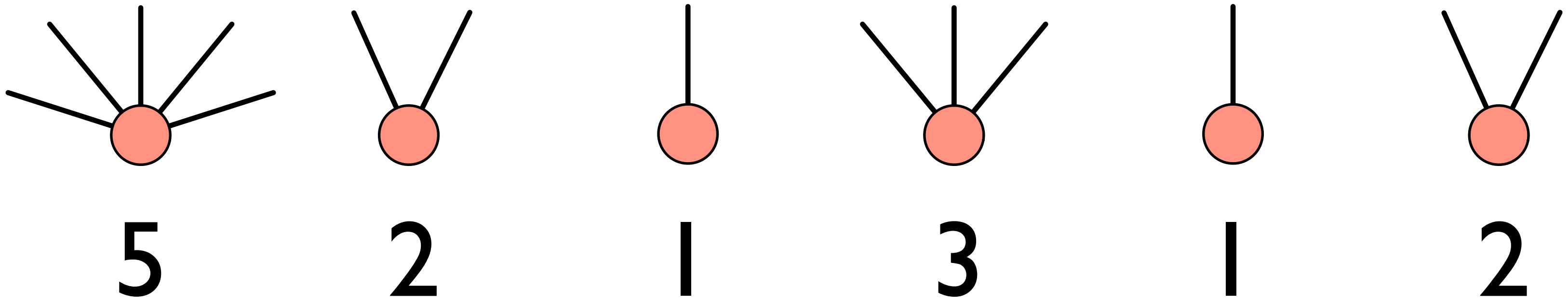
Configuration model

Make empty **stubs**

5 2 1 3 1 2

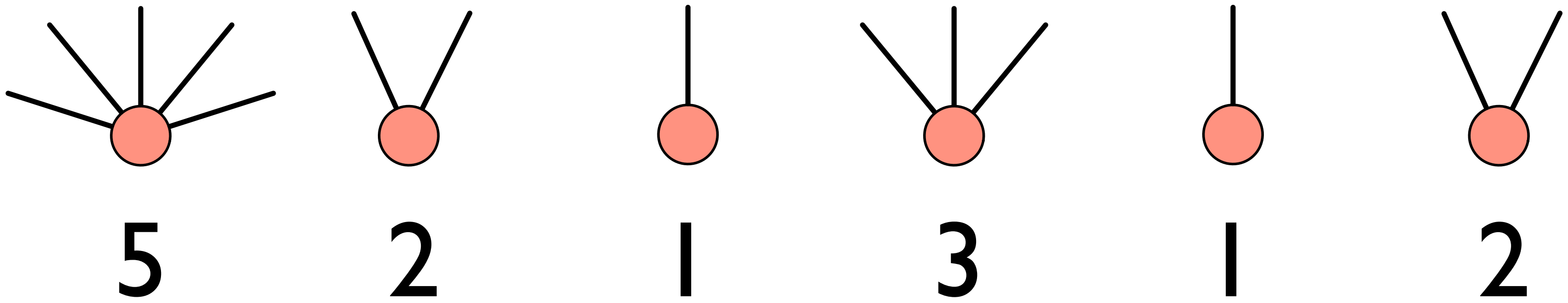
Configuration model

Make empty **stubs**



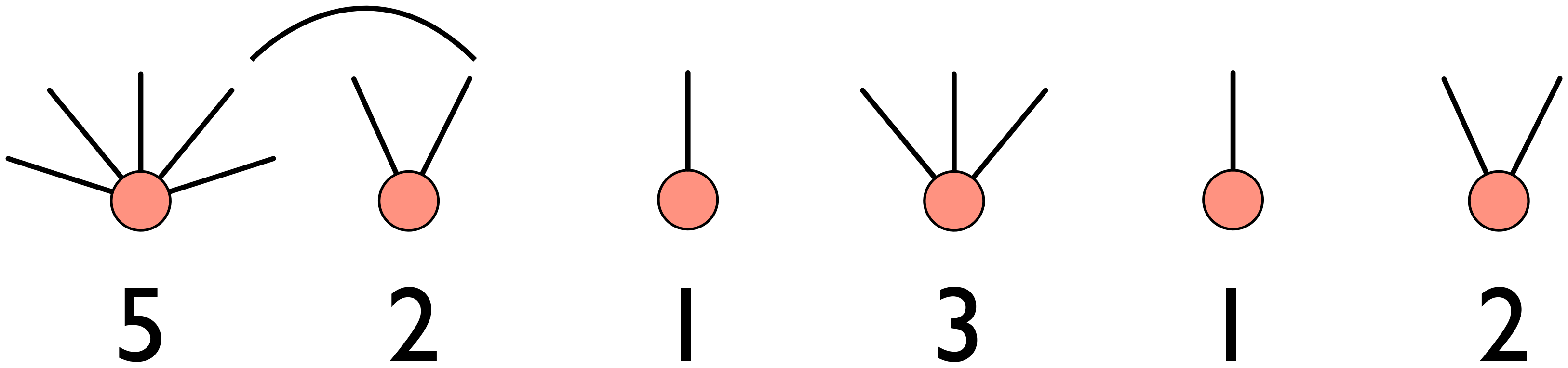
Configuration model

Randomly drop **links** between stubs



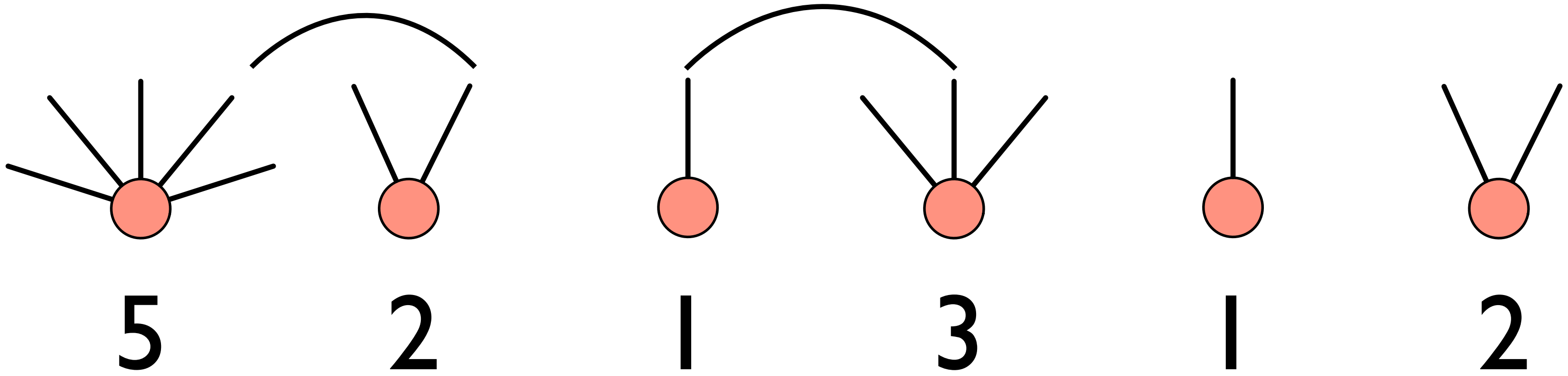
Configuration model

Randomly drop **links** between stubs



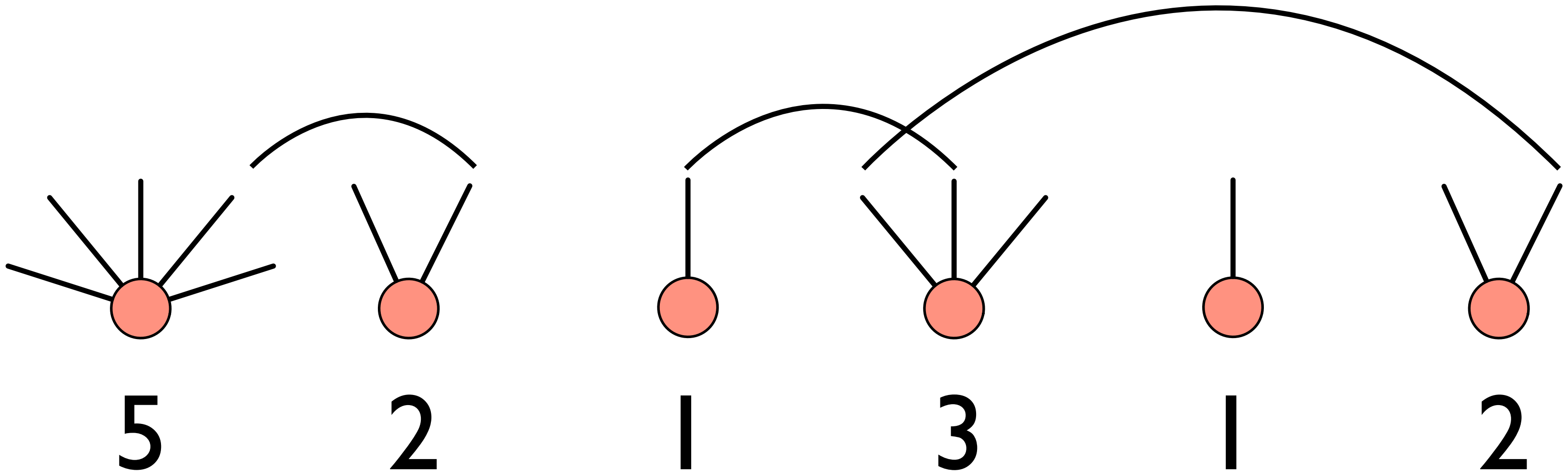
Configuration model

Randomly drop **links** between stubs

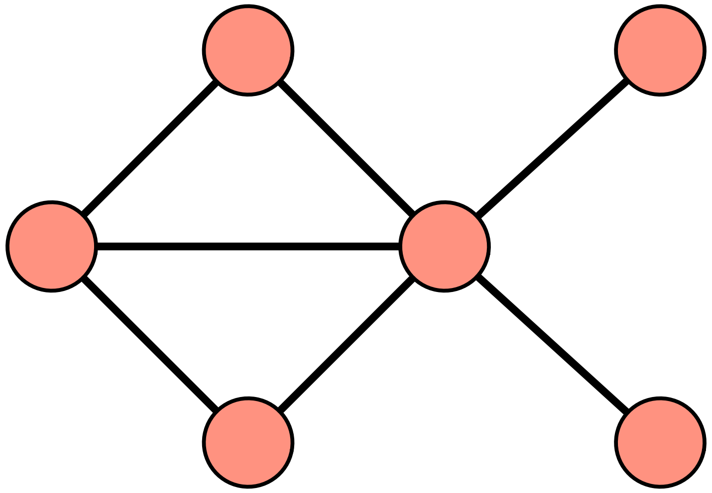


Configuration model

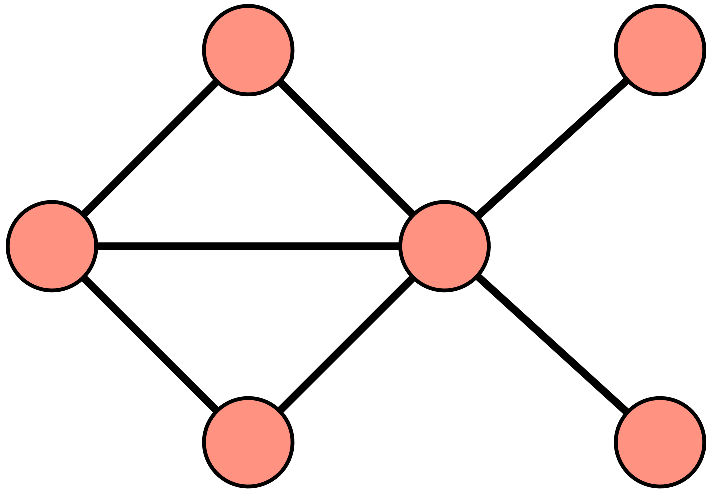
Randomly drop **links** between stubs



Configuration model

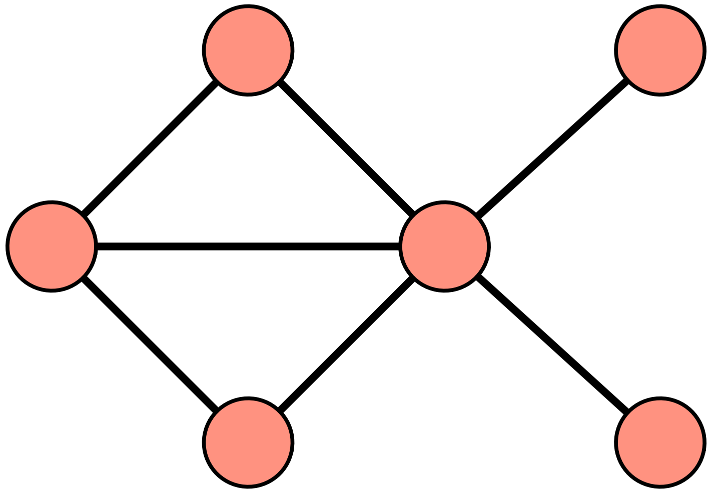


Configuration model



Doesn't just **preserve** degree distribution
but **degree sequence**

Configuration model



Doesn't just **preserve** degree distribution
but **degree sequence**

Useful **null model**: all structure
destroyed except degrees

Part II

Getting started on a computer

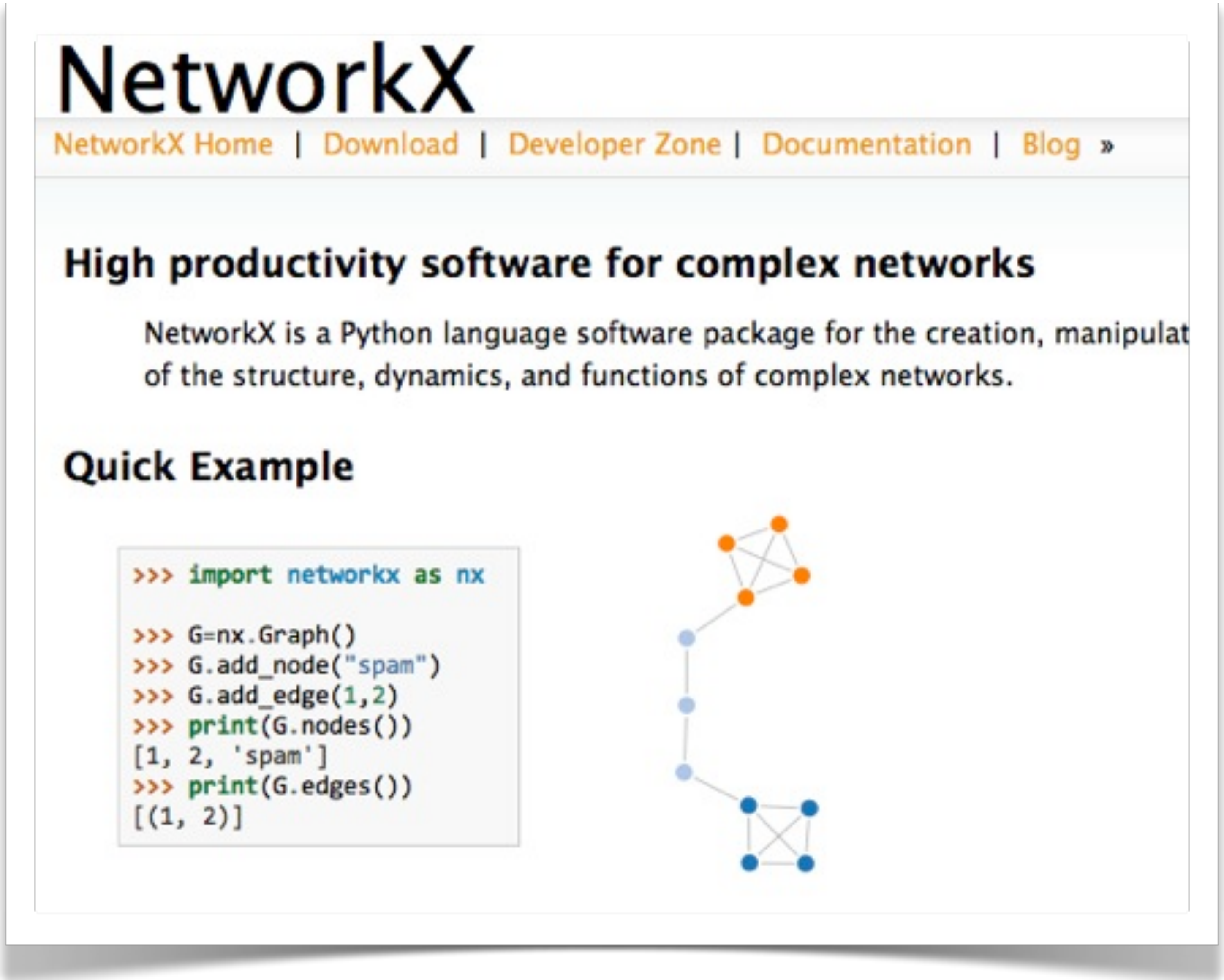
→

```
drwxr-xr-x  2 hkong hkong 4096 2007-10-21 00:54 video/
drwxr-xr-x  2 hkong hkong 4096 2007-11-24 18:03 vmware/
drwxr-xr-x  4 hkong hkong 4096 2007-10-21 00:54 wallpaper/
[~]14> ls
backup  Desktop   download  eggdrop  hi.py           photo        sh        video
bin     Dev       dump      games    levels.pixelman pix          snapshot  vmwar
bt      document  dwhelper  Hex      mount           programming  tex       wallp
[~]15> ls -F
backup/   Dev/        dwhelper/  hi.py            pix/          tex/
bin/      document/   eggdrop/   levels.pixelman  programming/  video/
bt/       download/   games/     mount/           sh/           vmware/
Desktop/  dump/       Hex/       photo/           snapshot/     wallpaper/
[~]16> cd programming/python/
[~/programming/python]17> ls
comicdl.py  except.py     os_doc.txt  testnum.py   web.py
dl-mod      funcreate.py  pygame      ucpy         wxPython-demo-2.8.7.1.tar.bz2
errlog.txt  glade         pytest      useclass.py
examples    ioread.py     temp.out    useclass.pyc
[~/programming/python]18> ls -F
comicdl.py*  except.py      os_doc.txt*  testnum.py*   web.py*
dl-mod/      funcreate.py*  pygame/      ucpy/         wxPython-demo-2.8.7.1.tar.
errlog.txt   glade/         pytest/      useclass.py*
examples/    ioread.py*     temp.out*    useclass.pyc*
[~/programming/python]19> █
```

→

Demo time

Demo time



NetworkX

NetworkX Home | Download | Developer Zone | Documentation | Blog »

High productivity software for complex networks

NetworkX is a Python language software package for the creation, manipulat
of the structure, dynamics, and functions of complex networks.

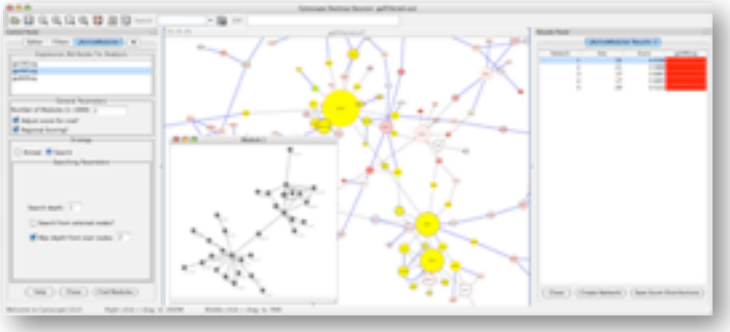
Quick Example

```
>>> import networkx as nx

>>> G=nx.Graph()
>>> G.add_node("spam")
>>> G.add_edge(1,2)
>>> print(G.nodes())
[1, 2, 'spam']
>>> print(G.edges())
[(1, 2)]
```

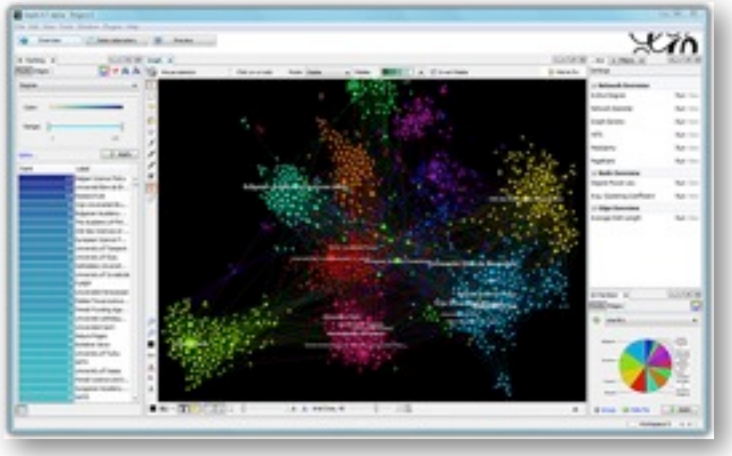
networkx.lanl.gov
python.org

Cytoscape



cytoscape.org

Gephi



gephi.org

Network search

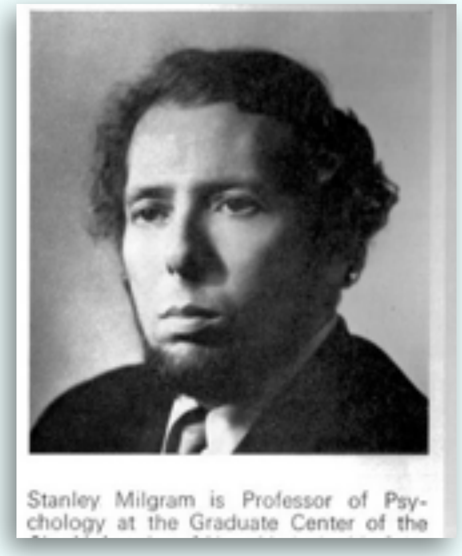
How to find stuff

Recall

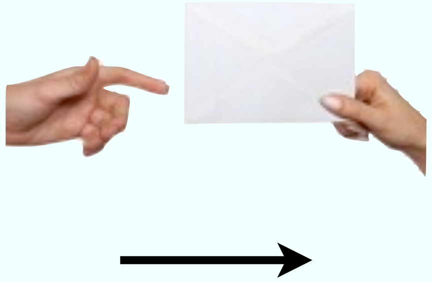
How to find stuff

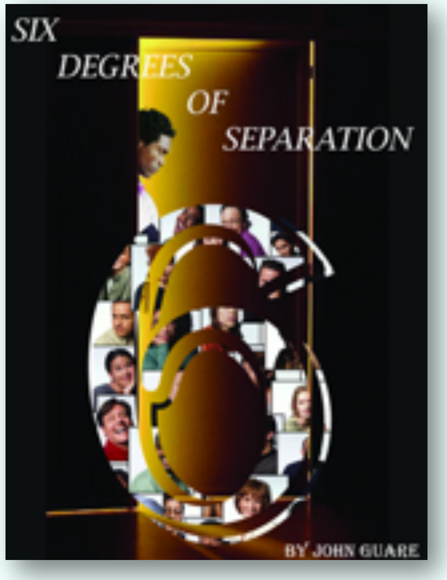
Recall

1960s **Milgram** asked “How far apart are we?”



Stanley Milgram is Professor of Psychology at the Graduate Center of the



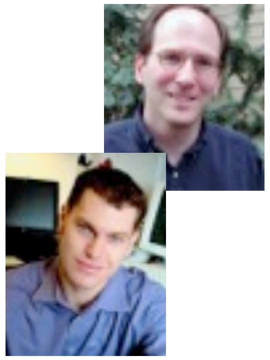


Network search

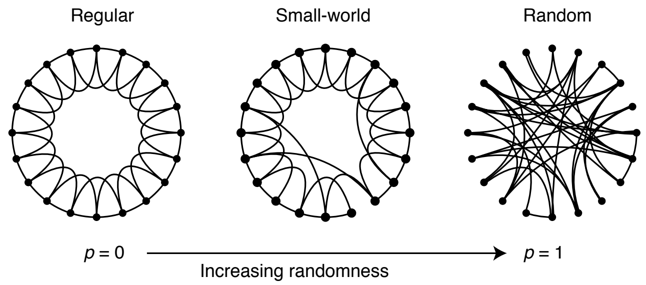
Can I **find things** on a network with only **local information**?

Network search

Can I **find things** on a network with only **local information**?

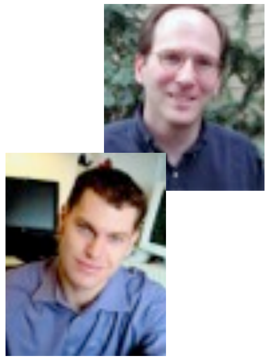


small world

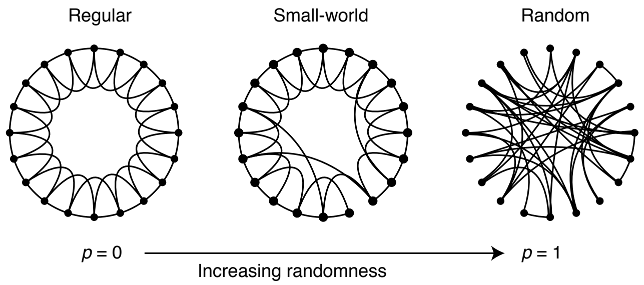


Network search

Can I **find things** on a network with only **local information**?



small world

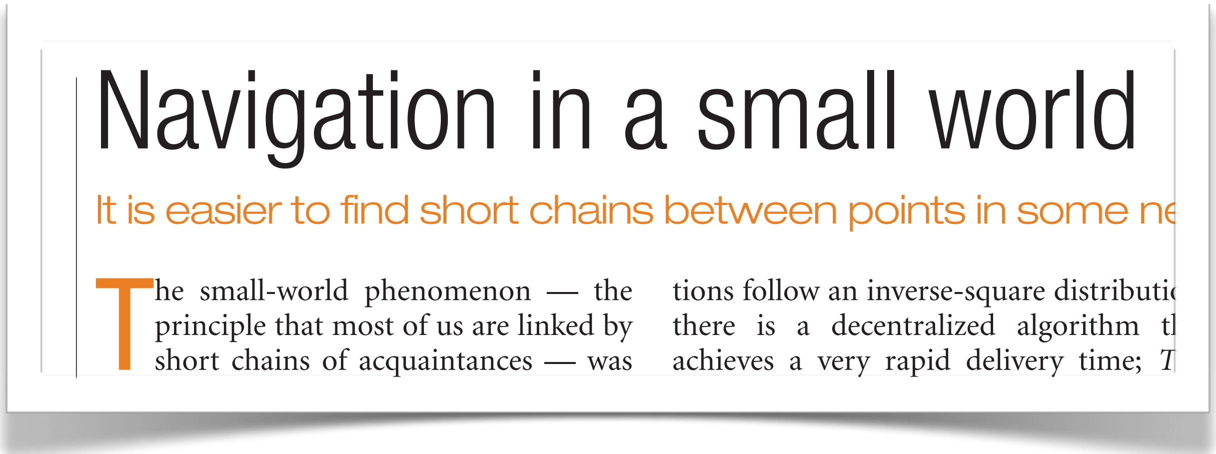


→

Kleinberg

Network search

Kleinberg



Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

tions follow an inverse-square distributi there is a decentralized algorithm t achieves a very rapid delivery time; T

Networks exist in space

Network search

Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

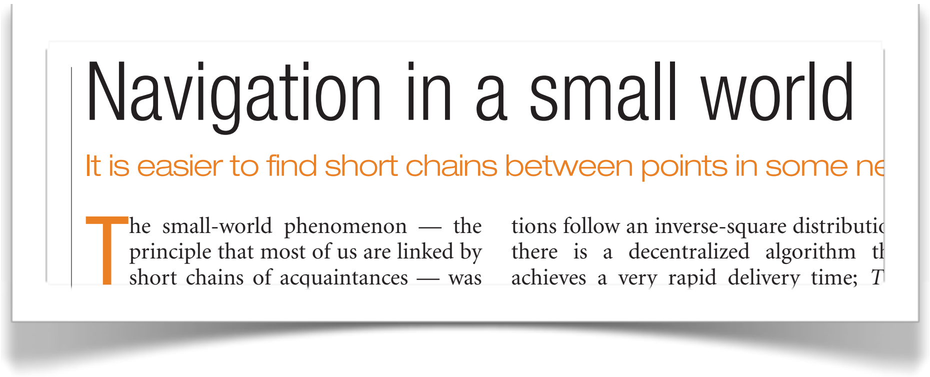
tions follow an inverse-square distributi
there is a decentralized algorithm t
achieves a very rapid delivery time;

→

Algorithmic analysis of the **Milgram** letter passing experiment

How **long** does it take to send a letter from a source to target with only local information?

Network search



Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

tions follow an inverse-square distributi
there is a decentralized algorithm th
achieves a very rapid delivery time; T

→ Algorithmic analysis of the **Milgram** letter passing experiment

Kleinberg model

Network search

Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

tions follow an inverse-square distributi there is a decentralized algorithm t achieves a very rapid delivery time; T

Algorithmic analysis of the **Milgram** letter passing experiment

Kleinberg model

1) 2D lattice

Network search

Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

tions follow an inverse-square distributic there is a decentralized algorithm th achieves a very rapid delivery time; T

→

Algorithmic analysis of the **Milgram** letter passing experiment

Kleinberg model

1) 2D lattice

2) Add one long-range link to each node

Network search

Navigation in a small world

It is easier to find short chains between points in some ne

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was

tions follow an inverse-square distributi there is a decentralized algorithm th achieves a very rapid delivery time; T

→

Algorithmic analysis of the **Milgram** letter passing experiment

Kleinberg model

1) 2D lattice

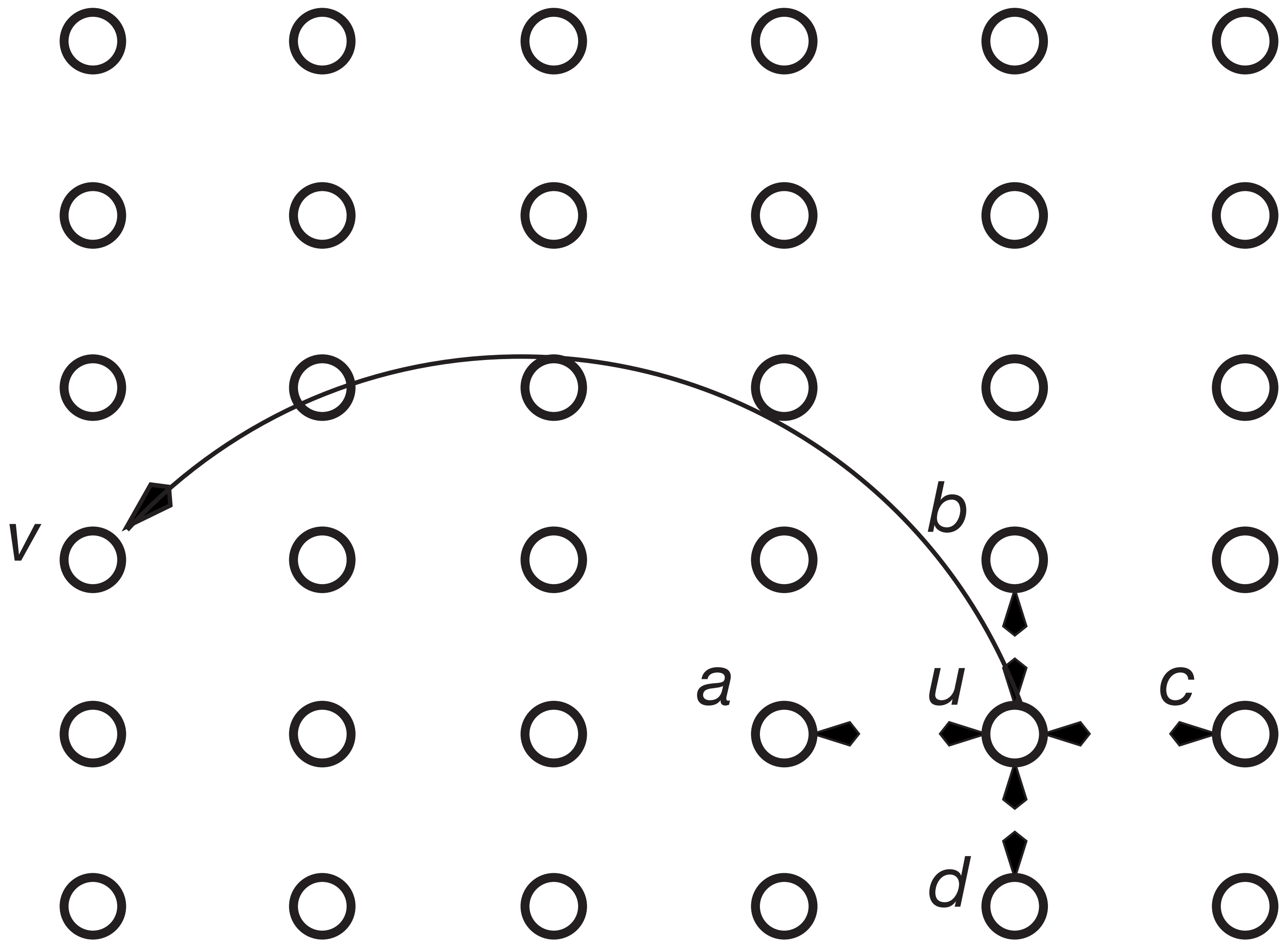
2) Add one long-range link to each node

3) Each node is a person and only knows the location of five neighbors

Network search

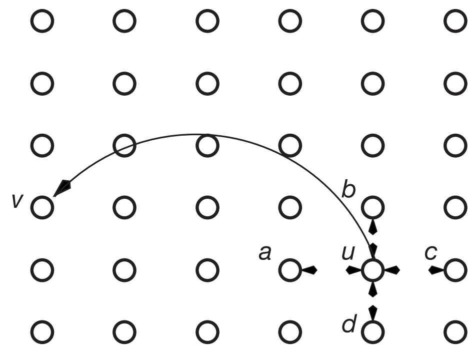
Kleinberg model

Models **local** neighbors and **long-distance** friendships



Network search

Kleinberg model

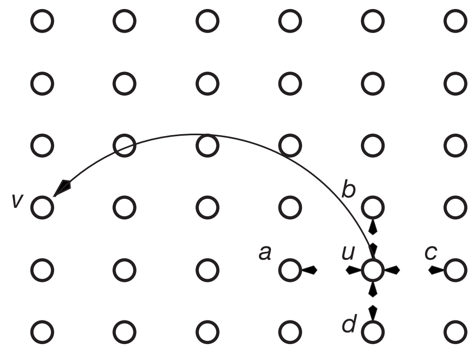


Models **local** neighbors and **long-distance** friendships

Long-range links

Network search

Kleinberg model



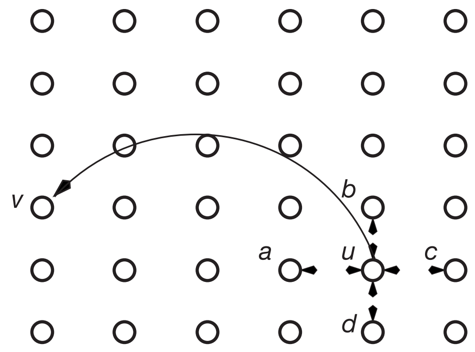
Models **local** neighbors and **long-distance** friendships

Long-range links

Neighbor at distance *r* linked with probability:

Network search

Kleinberg model



Models **local** neighbors and **long-distance** friendships

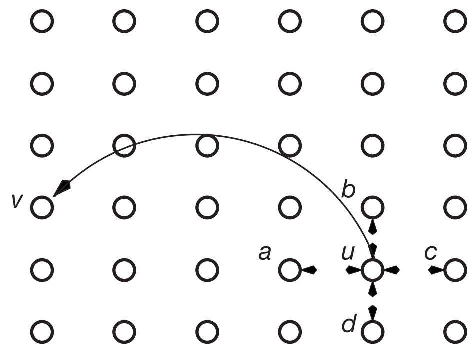
Long-range links

Neighbor at distance *r* linked with probability:

$$P_{\text{link}}(r) = \frac{r^{-\alpha}}{\sum_{\ell} \ell^{-\alpha}}$$

Network search

Kleinberg model



Long-range links

Neighbor at distance *r* linked with probability:

$$P_{\text{link}}(r) = \frac{r^{-\alpha}}{\sum_\ell \ell^{-\alpha}}$$

α “clustering exponent”

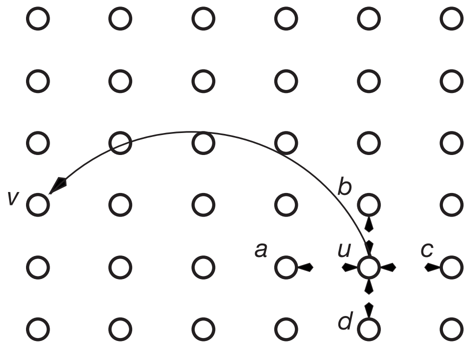
$\alpha = 0$

$\alpha \to \infty$

Network search

Kleinberg model

$$P_{\text{link}}(r) = \frac{r^{-\alpha}}{\sum_{\ell} \ell^{-\alpha}}$$



Navigation

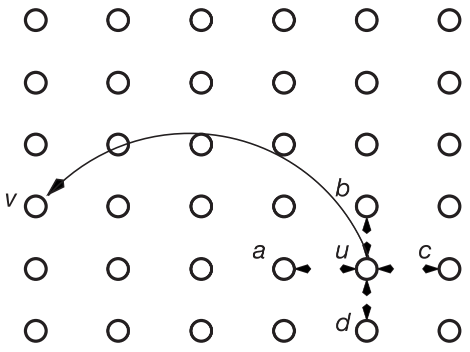
Pass message from source to target along neighbors

Only told location of target

Network search

Kleinberg model

$$P_{\text{link}}(r) = \frac{r^{-\alpha}}{\sum_{\ell} \ell^{-\alpha}}$$



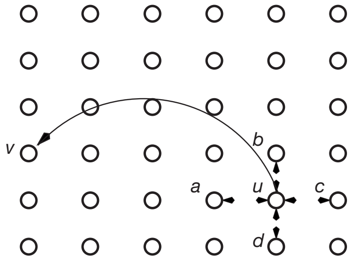
Only told location of target

Use **greedy** algorithm:

Each message holder sends message to neighbor **closest** to the target

Network search

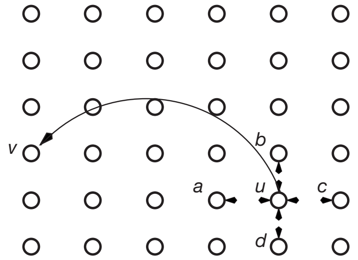
Kleinberg
model



How many times *T* must the message be passed?

Network search

Kleinberg model



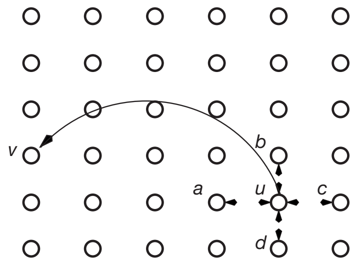
How many times *T* must the message be passed?

$\alpha \to \infty$

No long-range links: **slow**

Network search

Kleinberg
model



How many times T must the message be passed?

$\alpha \to \infty$

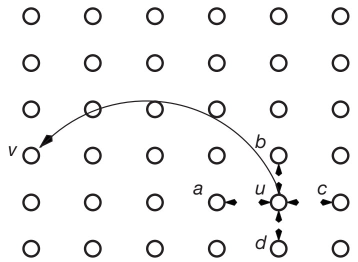
No long-range links: **slow**

$\alpha = 0$

Long-range links too long: **slow**

Network search

Kleinberg model



How many times *T* must the message be passed?

$\alpha \to \infty$

No long-range links: **slow**

?

$\alpha = 0$

Long-range links too long: **slow**

Network search

Kleinberg model

How many times T must the message be passed?

Greedy navigation is fastest when $\alpha = 2$

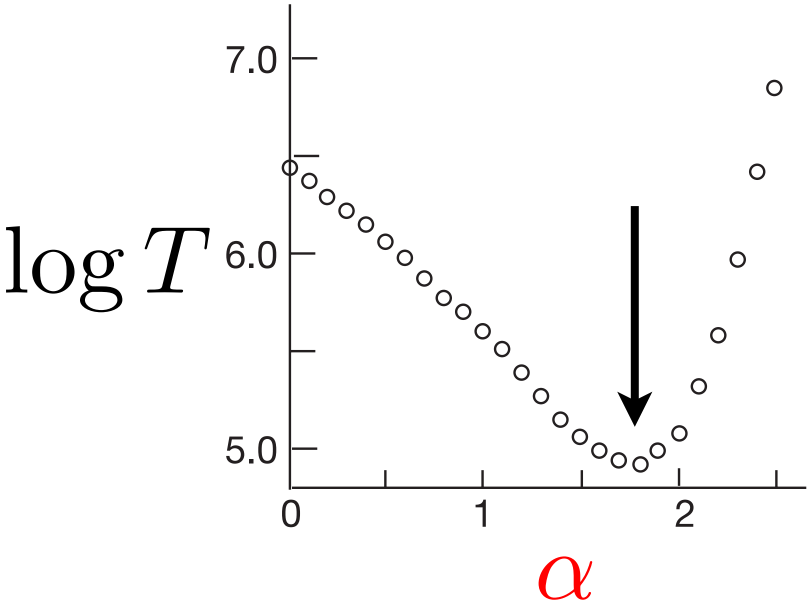
Network search

Kleinberg model

How many times *T* must the message be passed?

Greedy navigation is fastest when $\alpha = 2$

Simulations



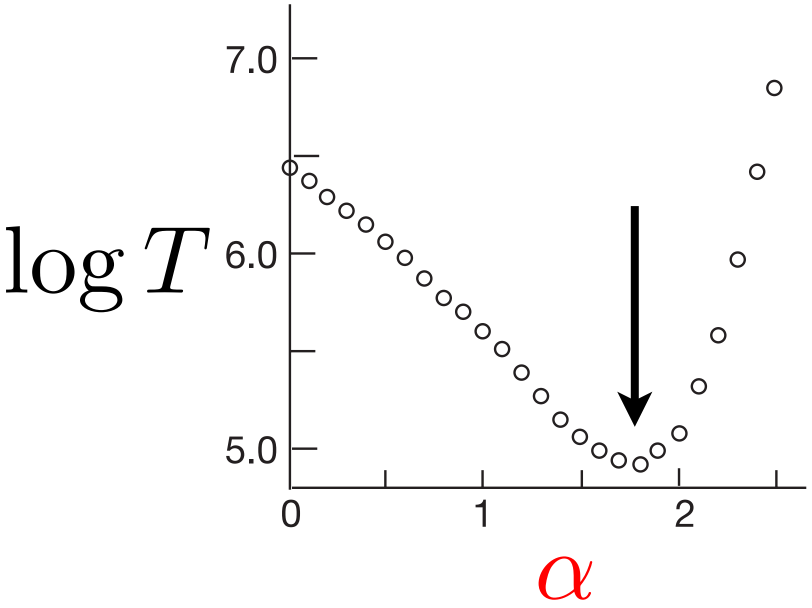
Network search

Kleinberg model

How many times *T* must the message be passed?

Greedy navigation is fastest when $\alpha = 2$

Simulations



Proof

$$T \geq CN^{\beta}, \qquad \text{for } \alpha \neq 2$$

$$T \geq C\left(\log N\right)^{\gamma}, \text{for } \alpha = 2$$

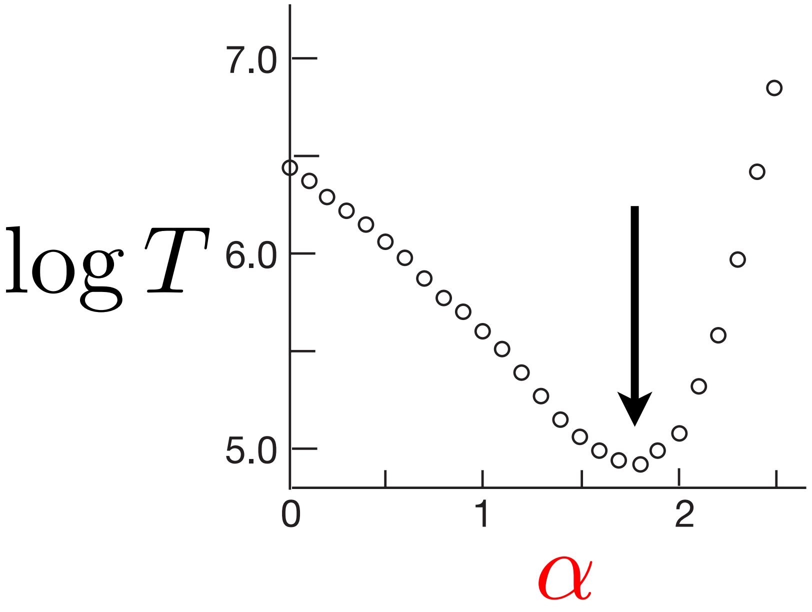
Network search

Kleinberg model

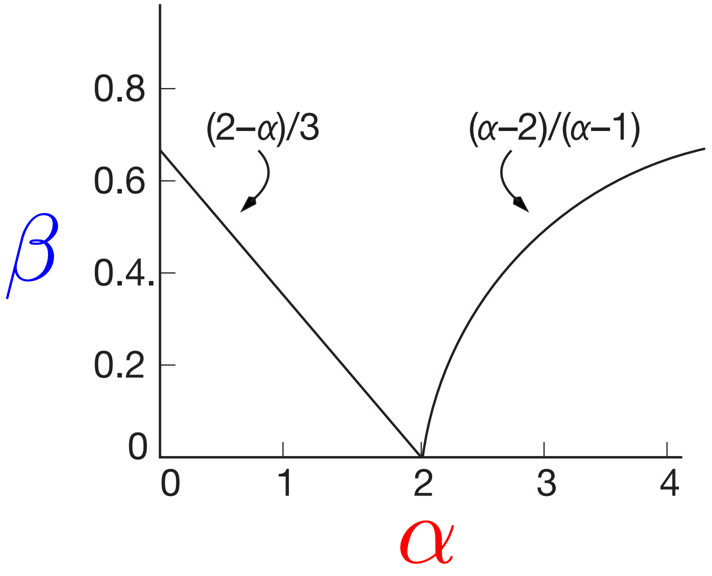
How many times *T* must the message be passed?

Greedy navigation is fastest when $\alpha = 2$

Simulations



Proof $T \geq CN^{\beta},$



The strength of weak ties

Algorithms are fine but there are also **sociological aspects** to navigation/ message passing

Q: **which links** are used to navigation?

The strength of weak ties

Granovetter, 1973

How do people **find jobs**?

The strength of weak ties

Granovetter, 1973

How do people **find jobs**?

Not from **best** friends—they have the same info you do!

The strength of weak ties

Granovetter, 1973

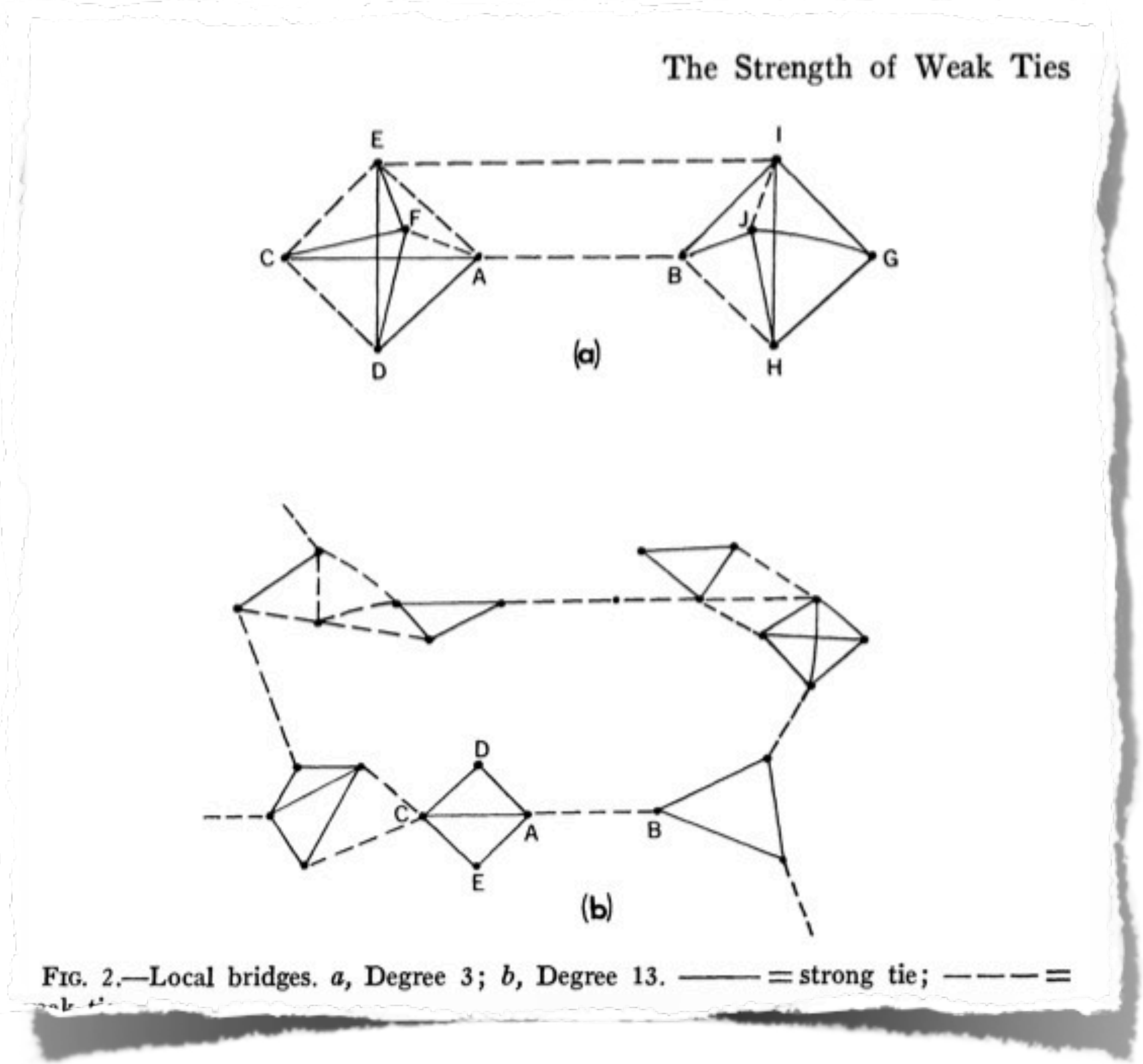
How do people **find jobs**?

Not from **best** friends—they have the same info you do!

Infrequent contacts—new pools of info

The strength of weak ties

Granovetter, 1973

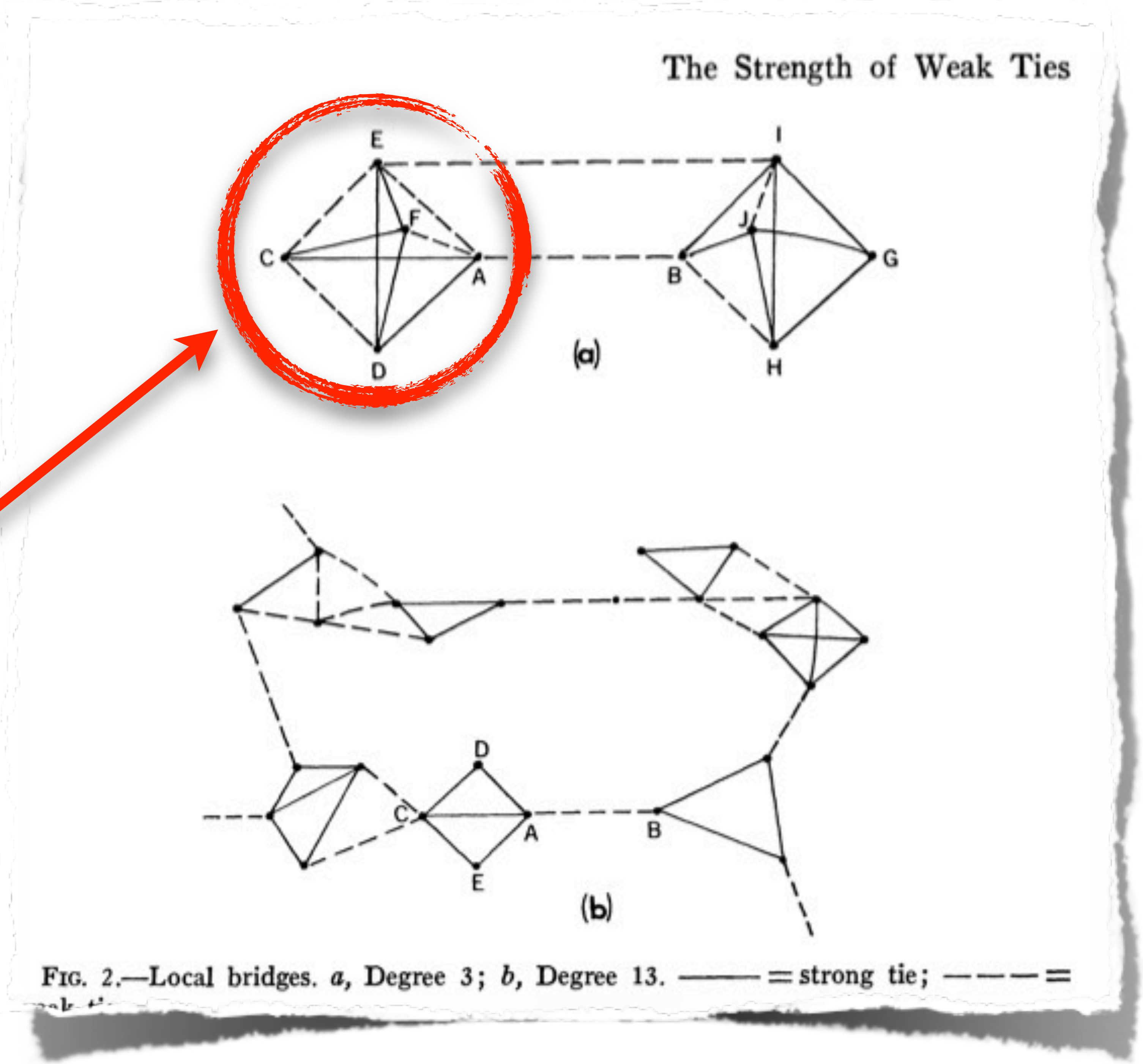


The Strength of Weak Ties

FIG. 2.—Local bridges. *a*, Degree 3; *b*, Degree 13. ——— = strong tie; — — — =

The strength of weak ties

Granovetter, 1973



The Strength of Weak Ties

FIG. 2.—Local bridges. *a*, Degree 3; *b*, Degree 13. ——— = strong tie; – – – =

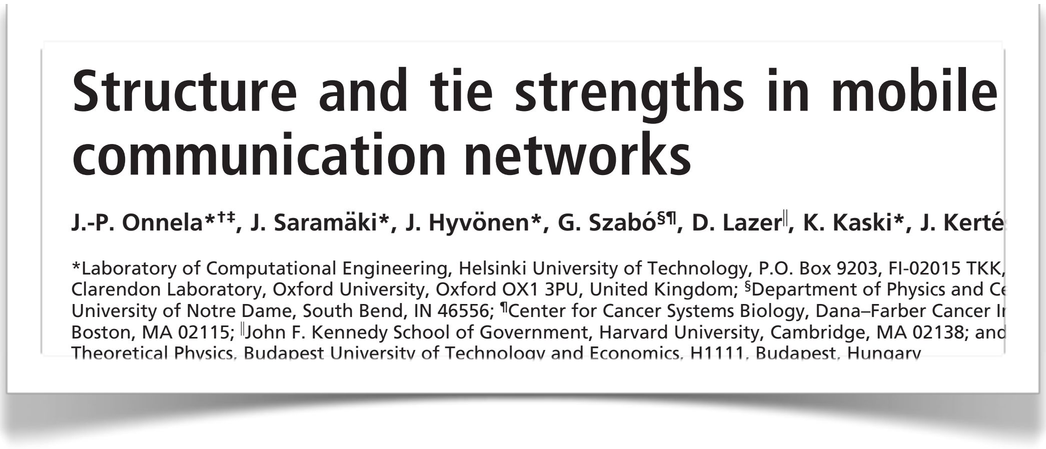
communities

The strength of weak ties

Modern validation ⟶ **Mobile phones**

The strength of weak ties

Modern validation → **Mobile phones**



Structure and tie strengths in mobile communication networks

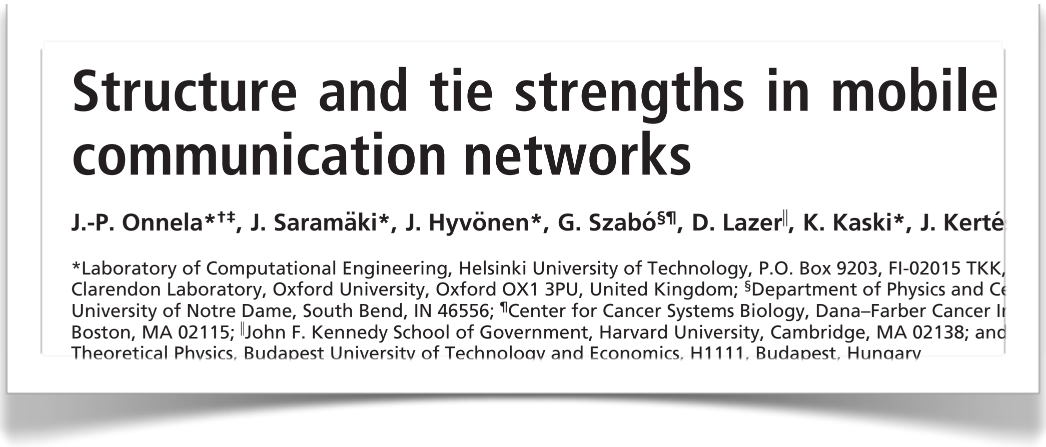
J.-P. Onnela*[†‡], J. Saramäki*, J. Hyvönen*, G. Szabó[§¶], D. Lazer[‖], K. Kaski*, J. Kerté

*Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, FI-02015 TKK,
Clarendon Laboratory, Oxford University, Oxford OX1 3PU, United Kingdom; [§]Department of Physics and C
University of Notre Dame, South Bend, IN 46556; [¶]Center for Cancer Systems Biology, Dana–Farber Cancer I
Boston, MA 02115; [‖]John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138; and
Theoretical Physics, Budapest University of Technology and Economics, H1111, Budapest, Hungary

PNAS, 2007

The strength of weak ties

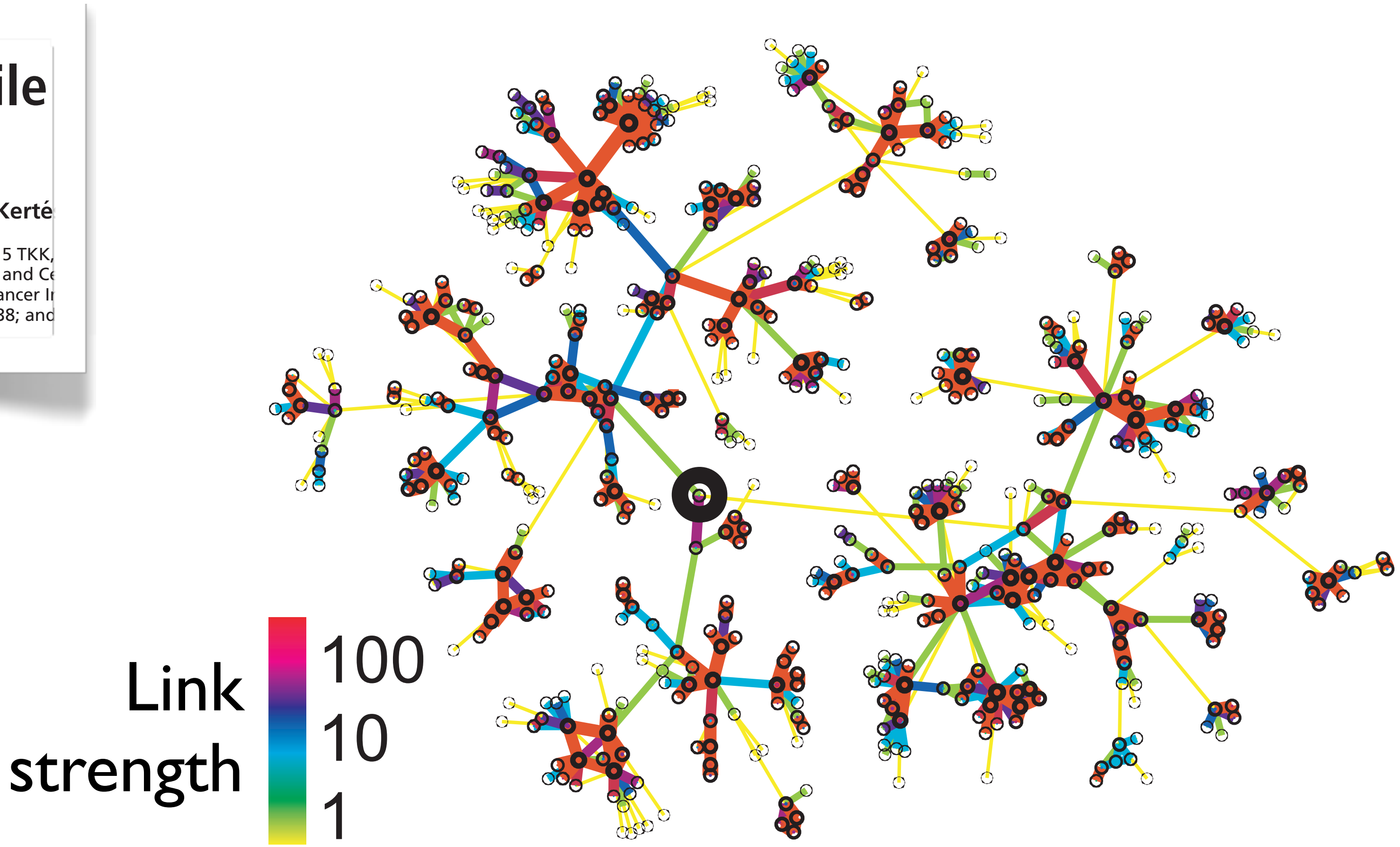
Modern validation ⟶ **Mobile phones**



Structure and tie strengths in mobile communication networks

J.-P. Onnela*[†‡], J. Saramäki*, J. Hyvönen*, G. Szabó[§¶], D. Lazer[‖], K. Kaski*, J. Kerté

*Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, FI-02015 TKK,
Clarendon Laboratory, Oxford University, Oxford OX1 3PU, United Kingdom; [§]Department of Physics and Ce
University of Notre Dame, South Bend, IN 46556; [¶]Center for Cancer Systems Biology, Dana–Farber Cancer I
Boston, MA 02115; [‖]John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138; and
Theoretical Physics, Budapest University of Technology and Economics, H1111, Budapest, Hungary



PNAS, 2007

Network robustness

Network robustness

Networks are composed of many simple interacting **elements**

Network **robustness**

Networks are composed of many simple interacting **elements**

Elements can **fail**

Network **robustness**

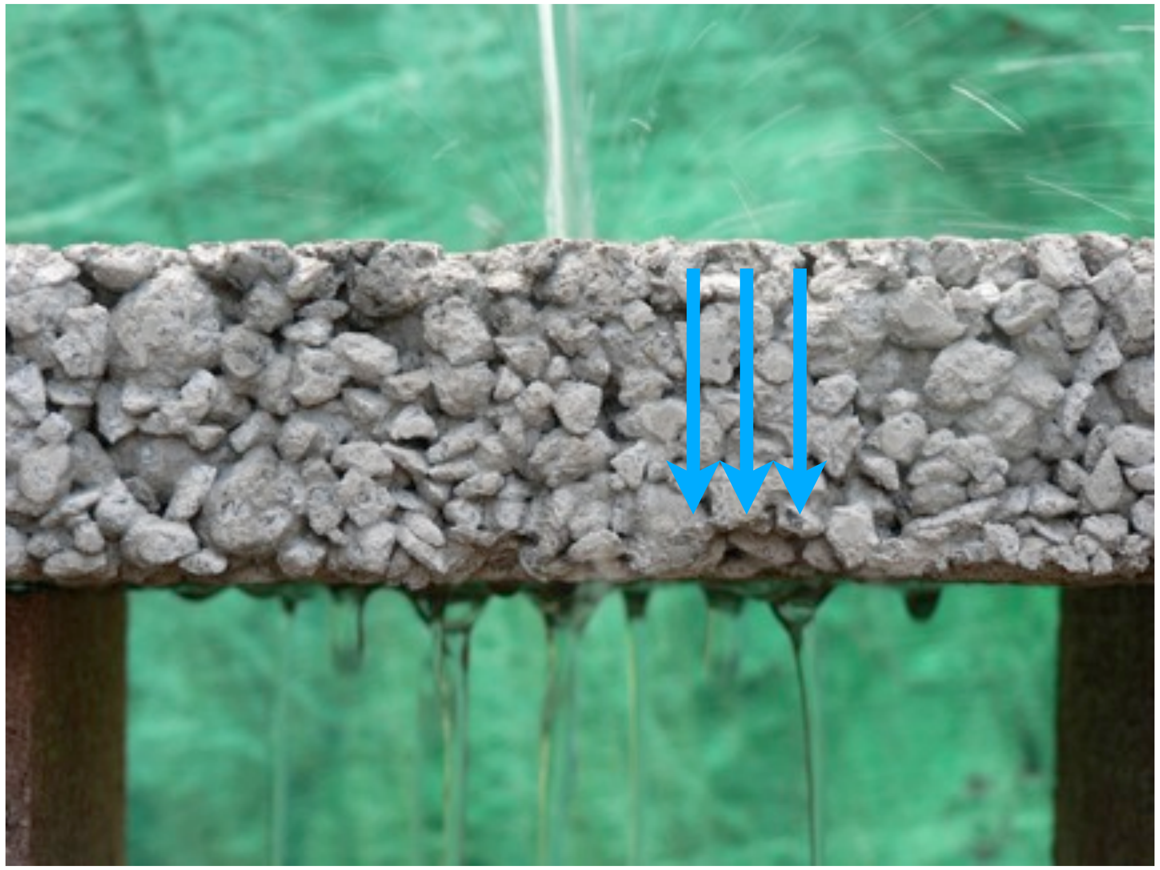
Networks are composed of many simple interacting **elements**

Elements can **fail**

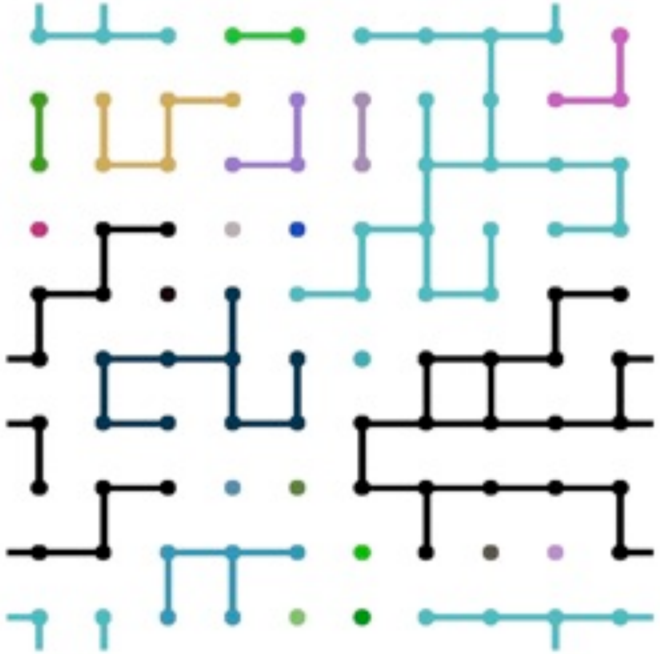
How do **networks** fail?

Percolation

Model fluid percolating through porous soil



Model soil with **lattice**

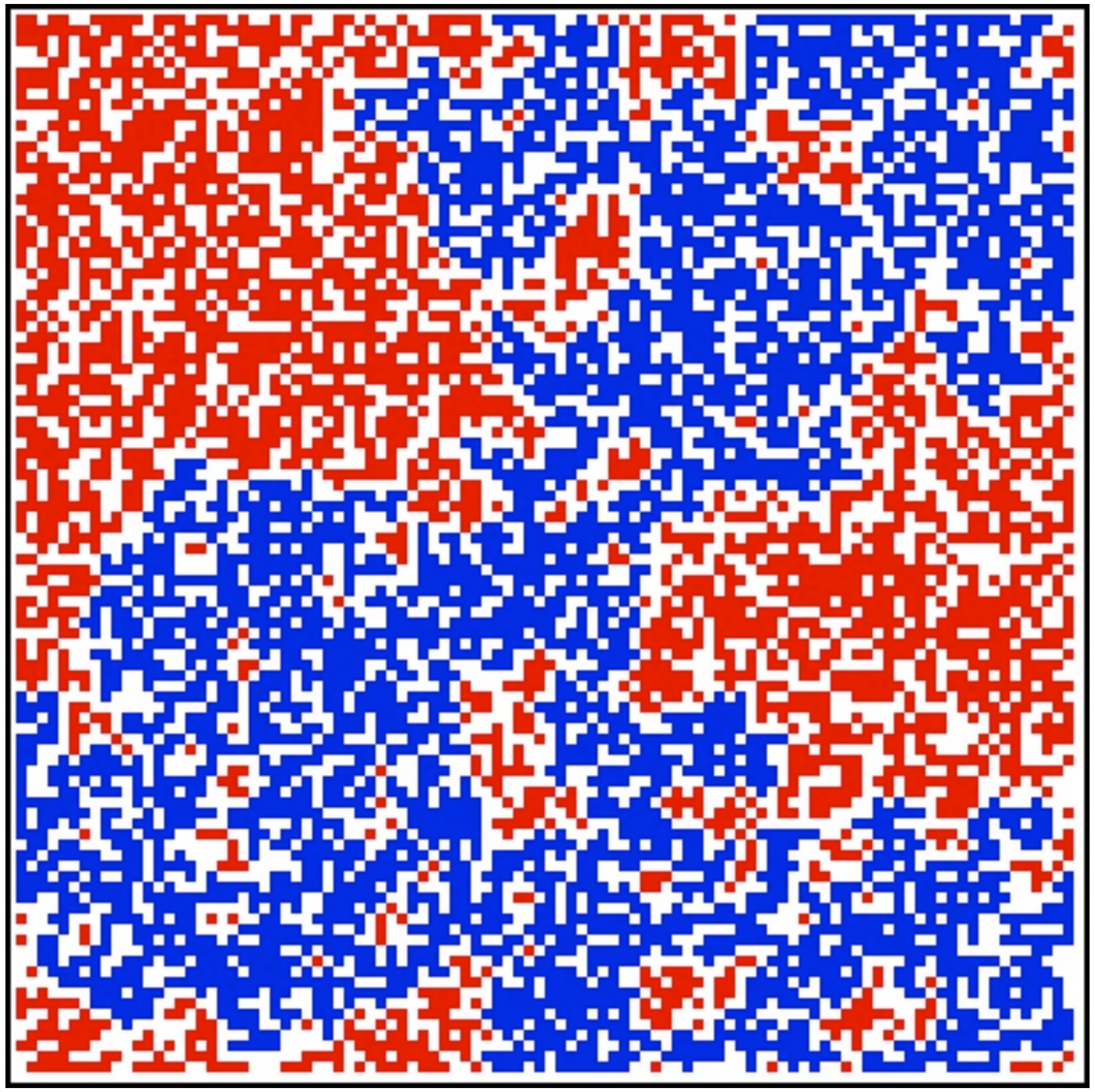


Blockages ⟶ Cut links (bonds)

Percolation

How many bonds should I cut until water won't flow?

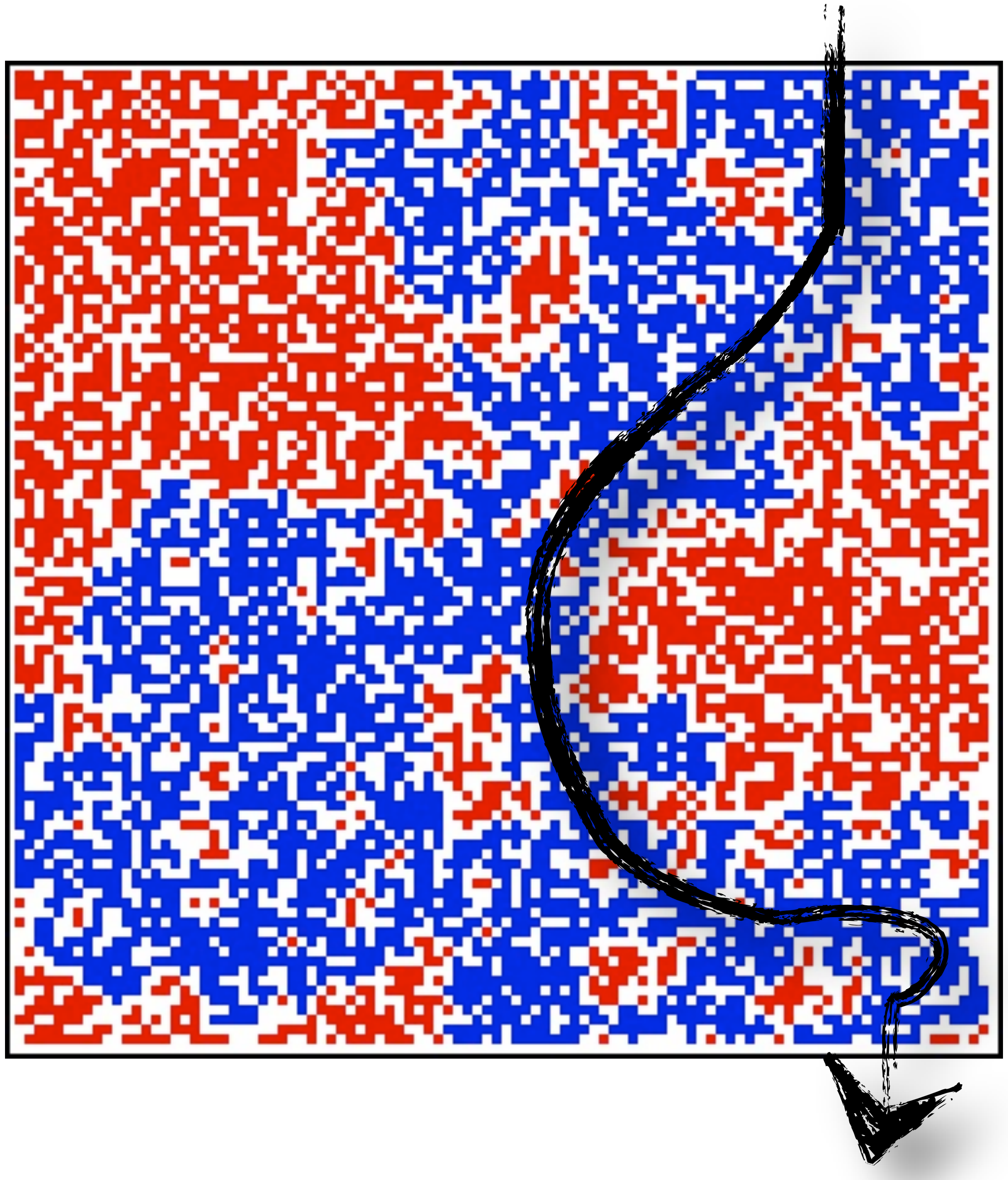
In other words, when is there a **spanning cluster**?



Percolation

How many bonds should I cut until water won’t flow?

In other words,
when is there a
spanning
cluster?



Percolation

In the context of **networks**

Have a random graph:

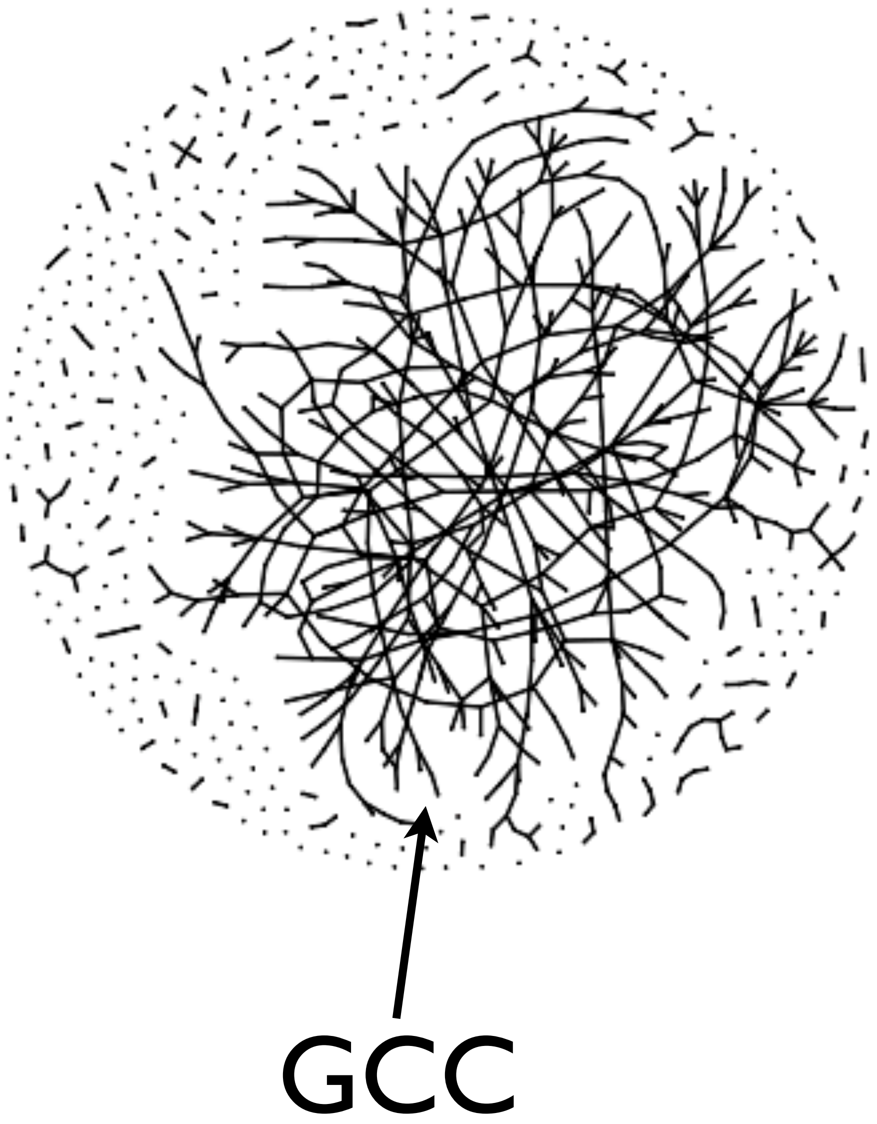
- Random **nodes fail** (site percolation)
- Random **links fail** (bond percolation)

p

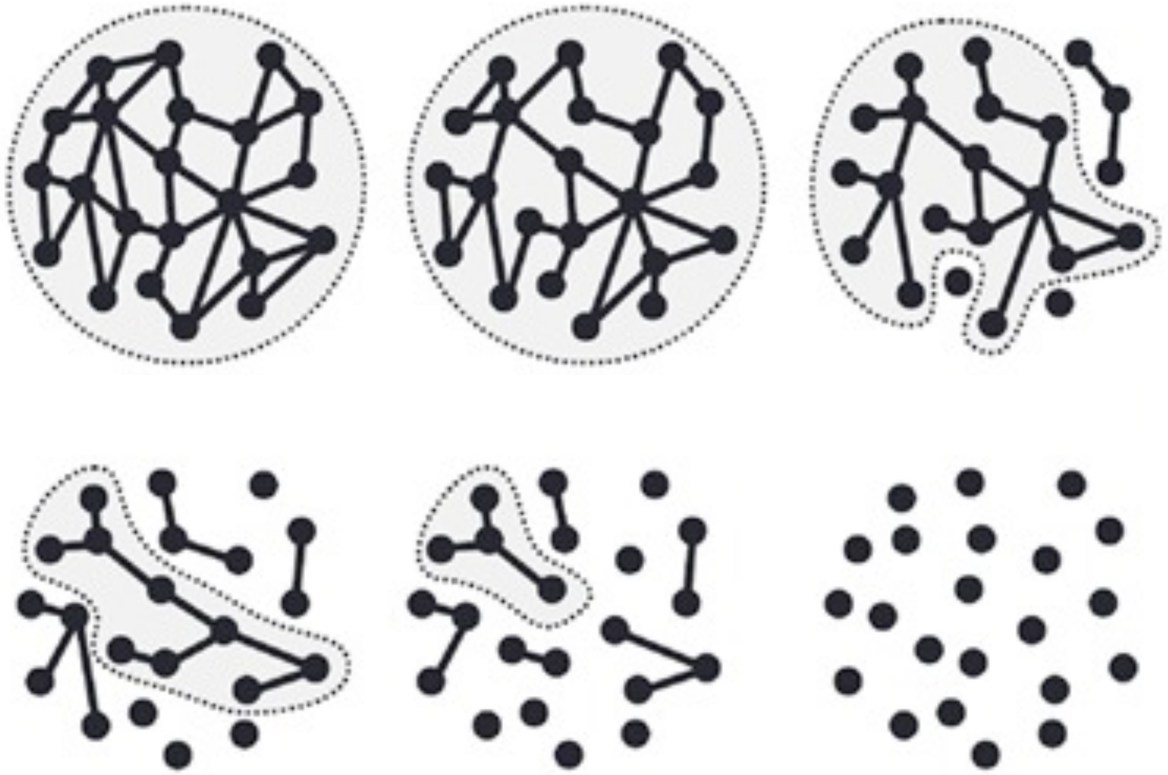
Does the network lose **global connectivity**?

No spanning cluster → No **Giant Connected Component** (GCC)

Percolation

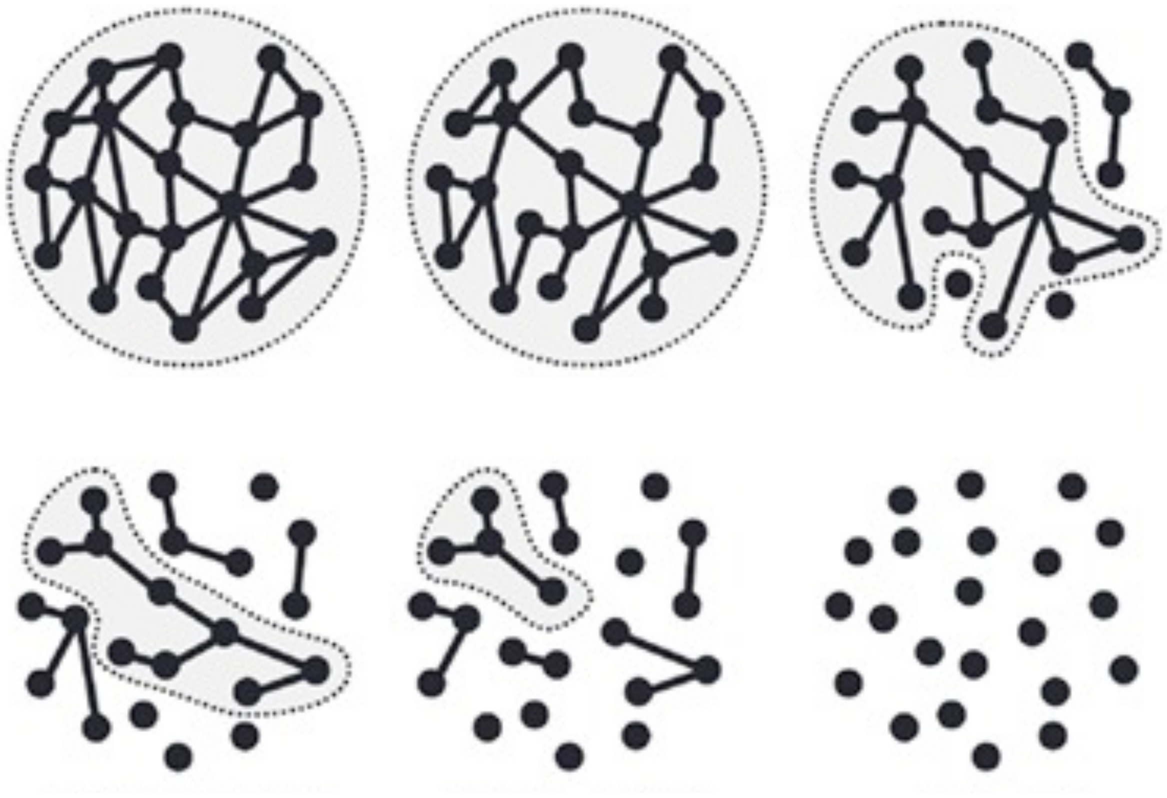


How do networks respond to random failures?



Percolation

How do networks respond to random failures?



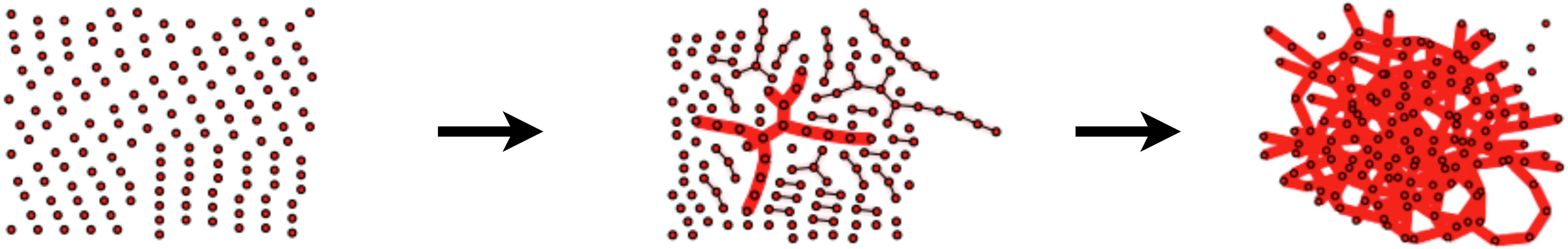
Is global damage **gradual** or **sudden**?

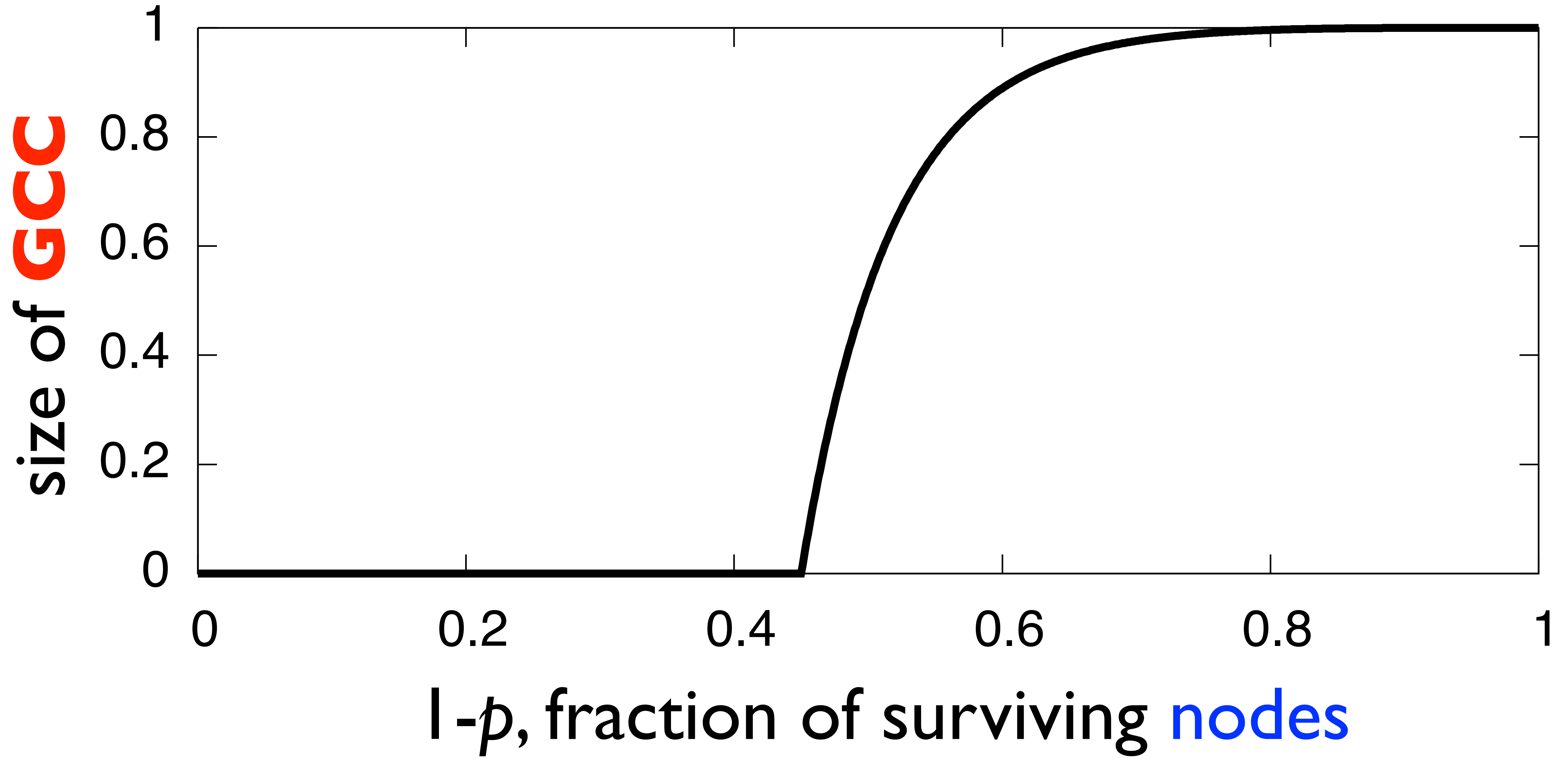
Percolation

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

Percolation transition

Percolation



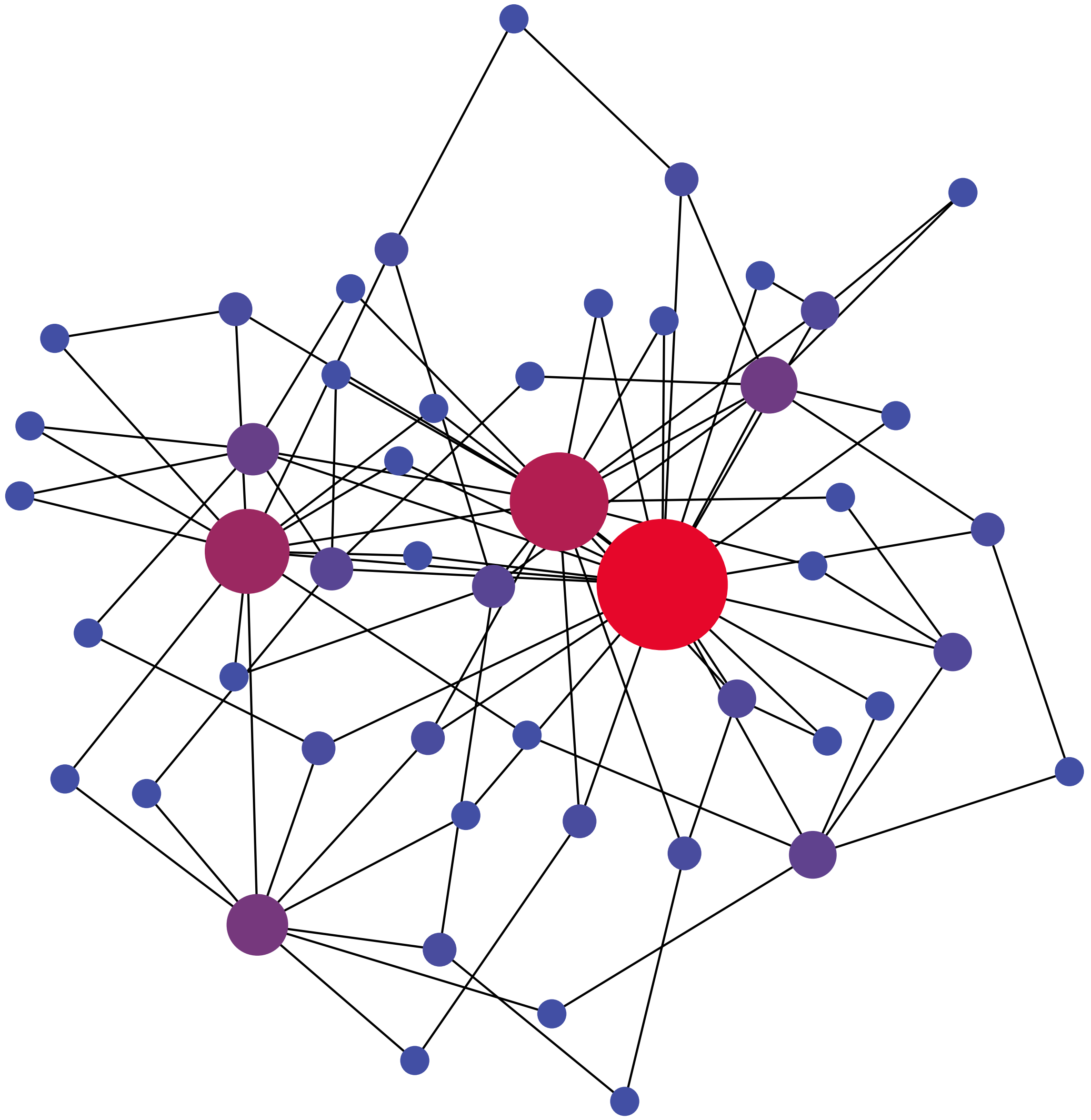


Some networks are special

Some networks are special

Scale-free graphs

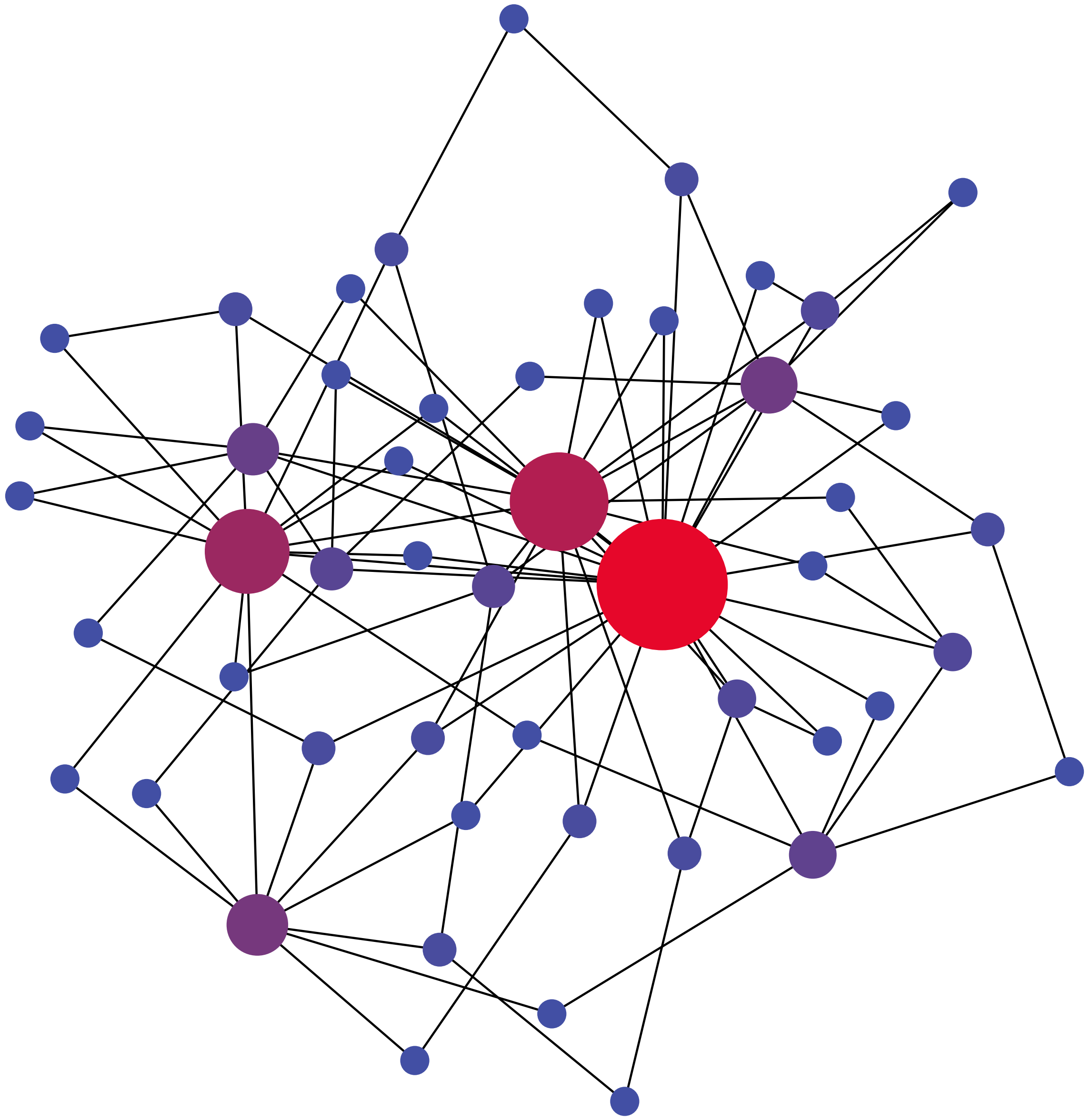
Hubs!



Some networks are special

Scale-free graphs

Hubs!

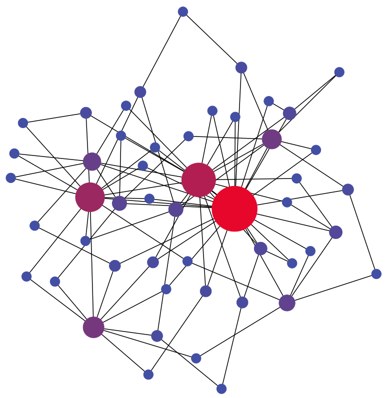


Imagine I randomly remove nodes from a very large scale-free network

Some networks are special

Scale-free graphs

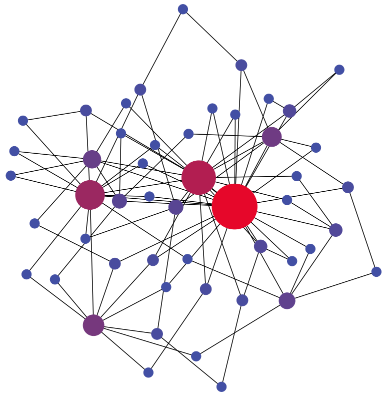
Hubs!



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

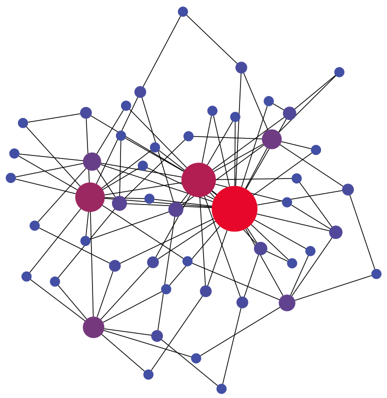
Unlikely I'll hit all the hubs

Hubs do a disproportionate job gluing the network together

Some networks are special

Scale-free graphs

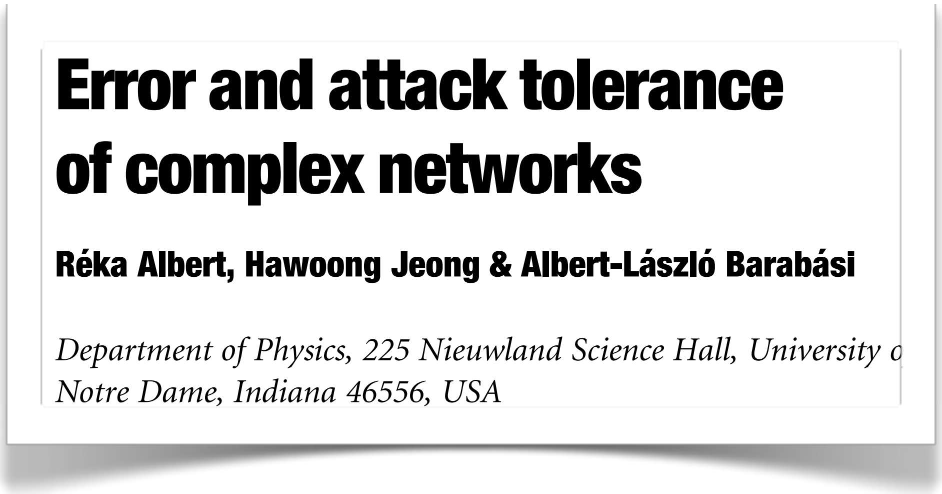
Hubs!



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 graphs are **robust** against random failures!

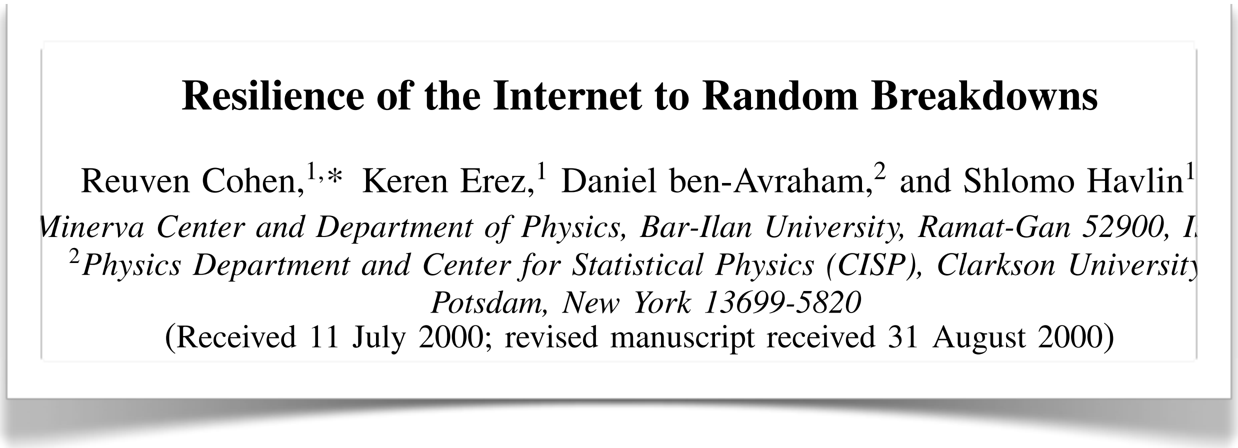


Error and attack tolerance of complex networks

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)

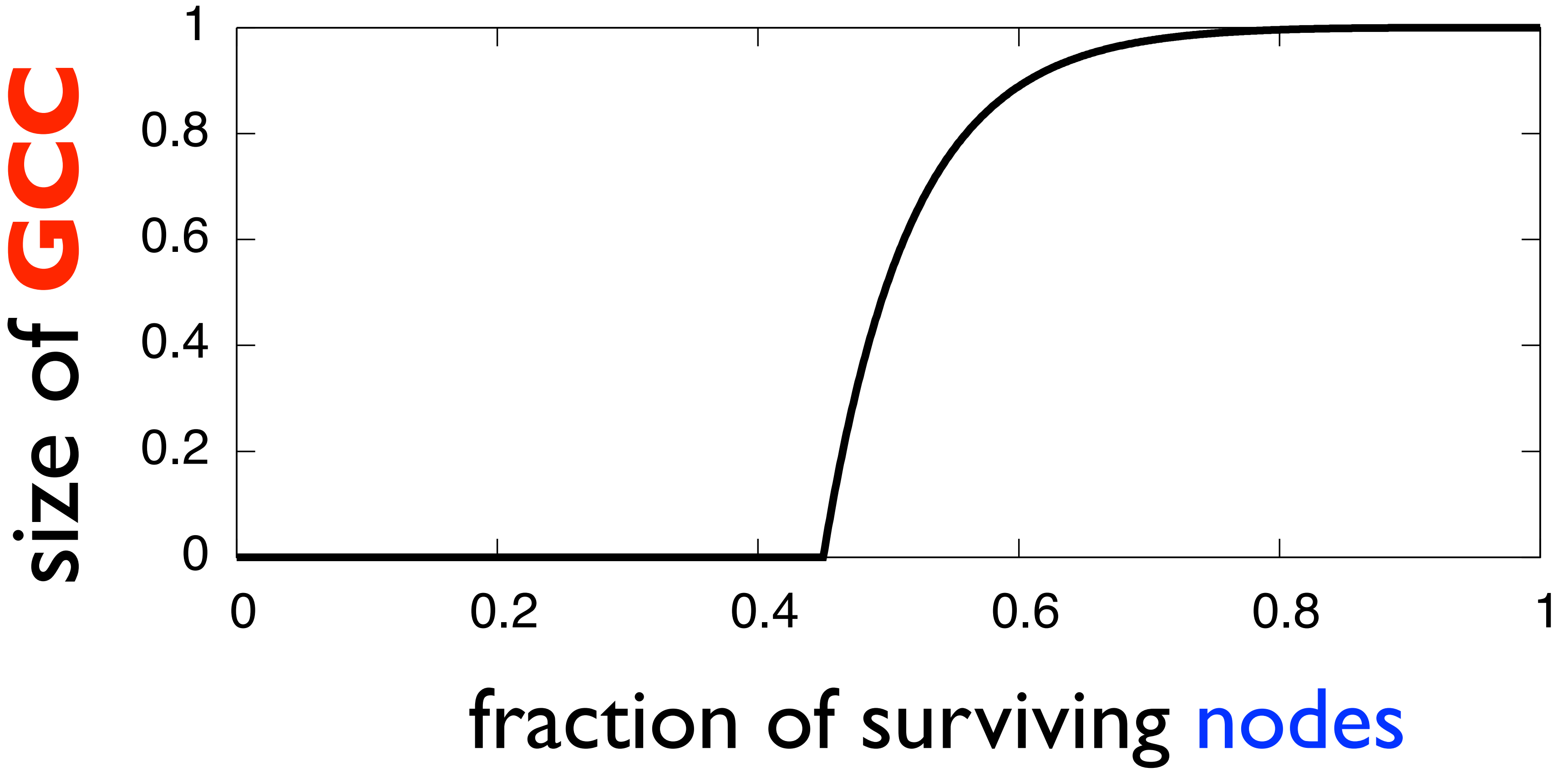
Phys Rev Lett, 2000

Some networks are special

Scale-free graphs are **robust** against random failures!

Some networks are special

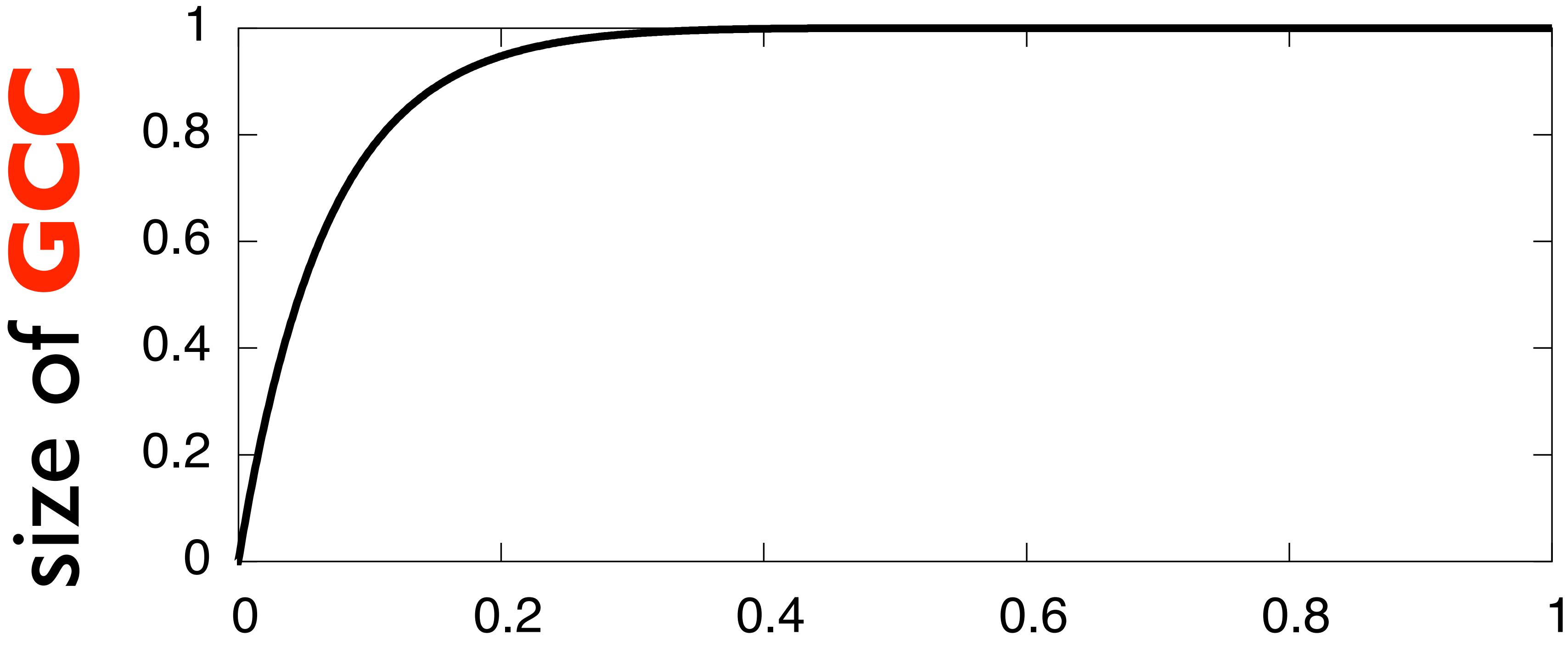
Scale-free graphs are **robust** against random failures!



Some networks are special

Scale-free graphs

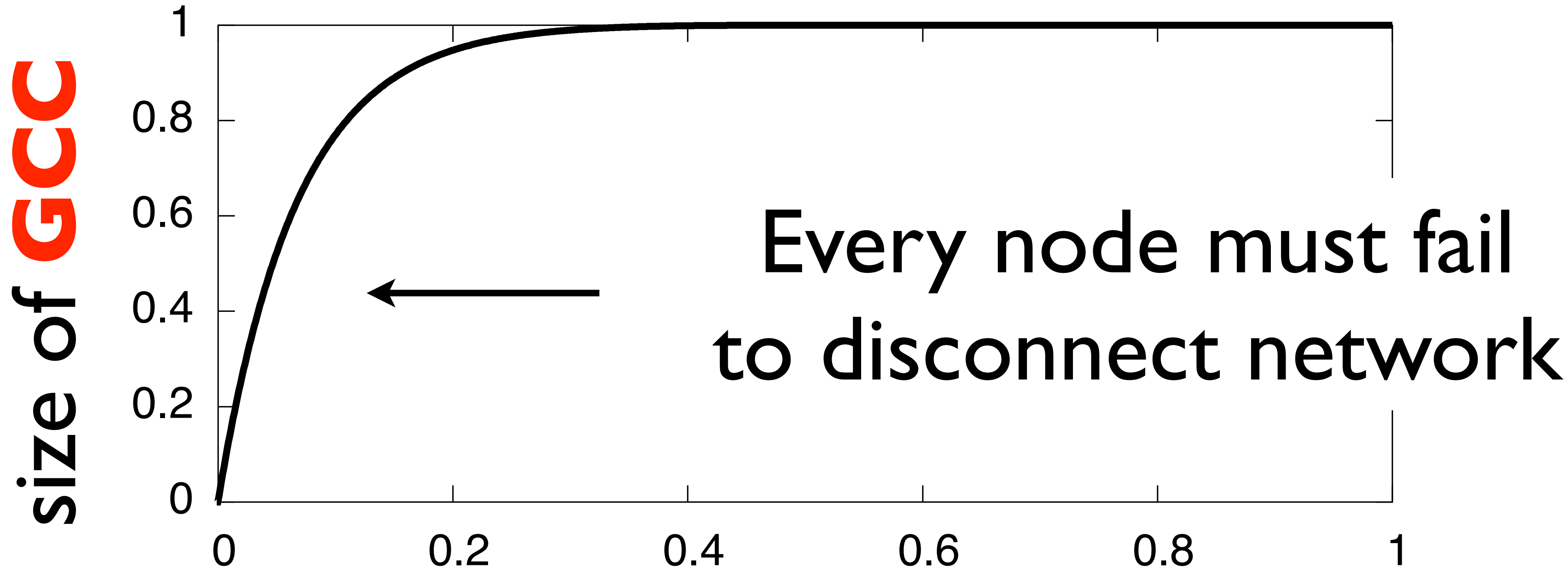
are **robust** against random failures!



fraction of surviving nodes

Some networks are special

Scale-free graphs are **robust** against random failures!



Every node must fail to disconnect network

fraction of surviving nodes

But there's a **price**

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

Do failures need to be **random**?

But there’s a **price**

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

Do failures need to be **random**?

Attack the network

But there's a **price**

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

Do failures need to be **random**?

Attack the network
Attack the hubs

But there’s a **price**

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

Do failures need to be **random**?

Attack the network
Attack the hubs
Hubs are **more likely** to fail

But there’s a **price**

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

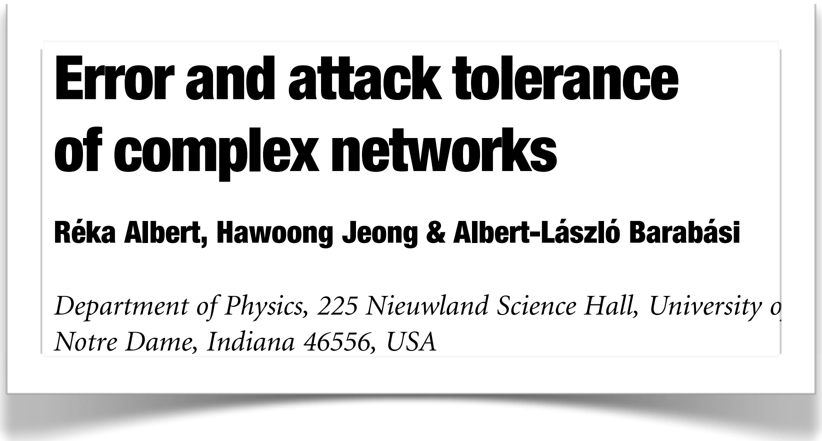
But there's a **price**

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

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

But there’s a **price**

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

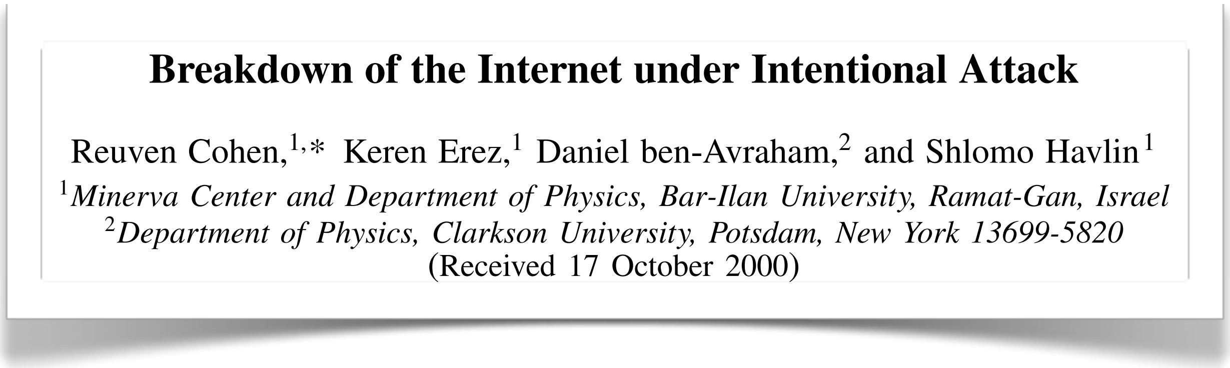


Error and attack tolerance of complex networks

Réka Albert, Hawoong Jeong & Albert-László Barabási

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

Nature, 2000



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

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

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

Percolation is a huge area

Tons of results and variants

Percolation is a huge area

Tons of results and variants

- Explosive percolation
- Cascading failures
- ....

Percolation is a huge area

Tons of results and variants

Explosive percolation
Cascading failures
....

Lots of applications

Percolation is a huge area

Tons of results and variants

- Explosive percolation
- Cascading failures
- ....

Lots of applications

Many complex problems map into simple percolation models

Percolation is a huge area

Tons of results and variants

Explosive percolation
Cascading failures
....

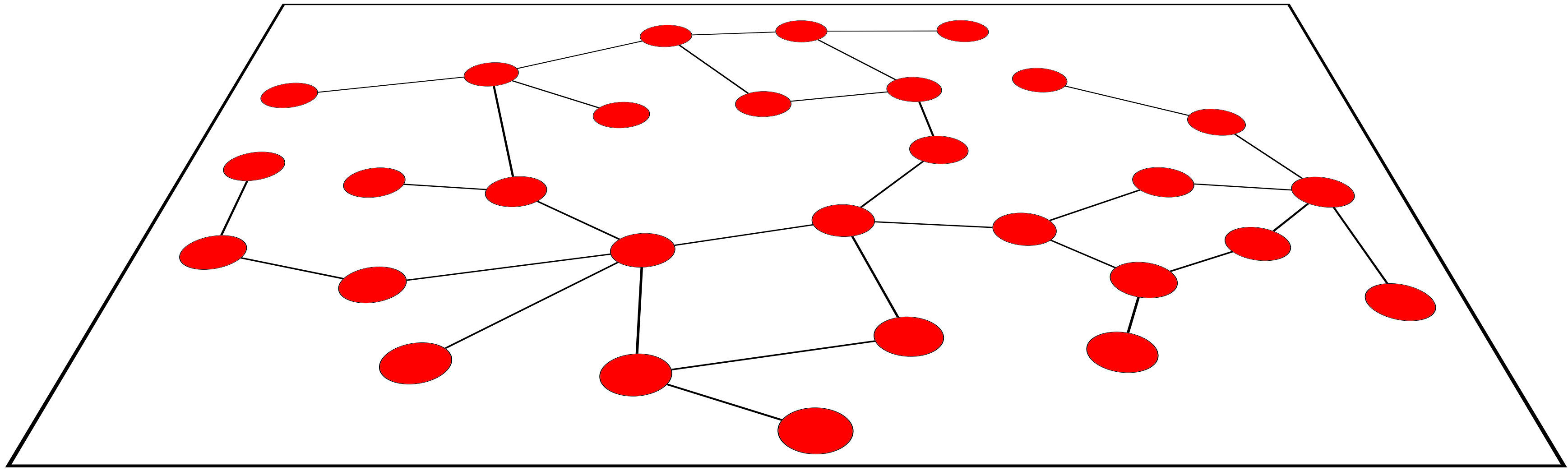
Lots of applications

Many complex problems map into simple percolation models

Epidemics / Vaccinations

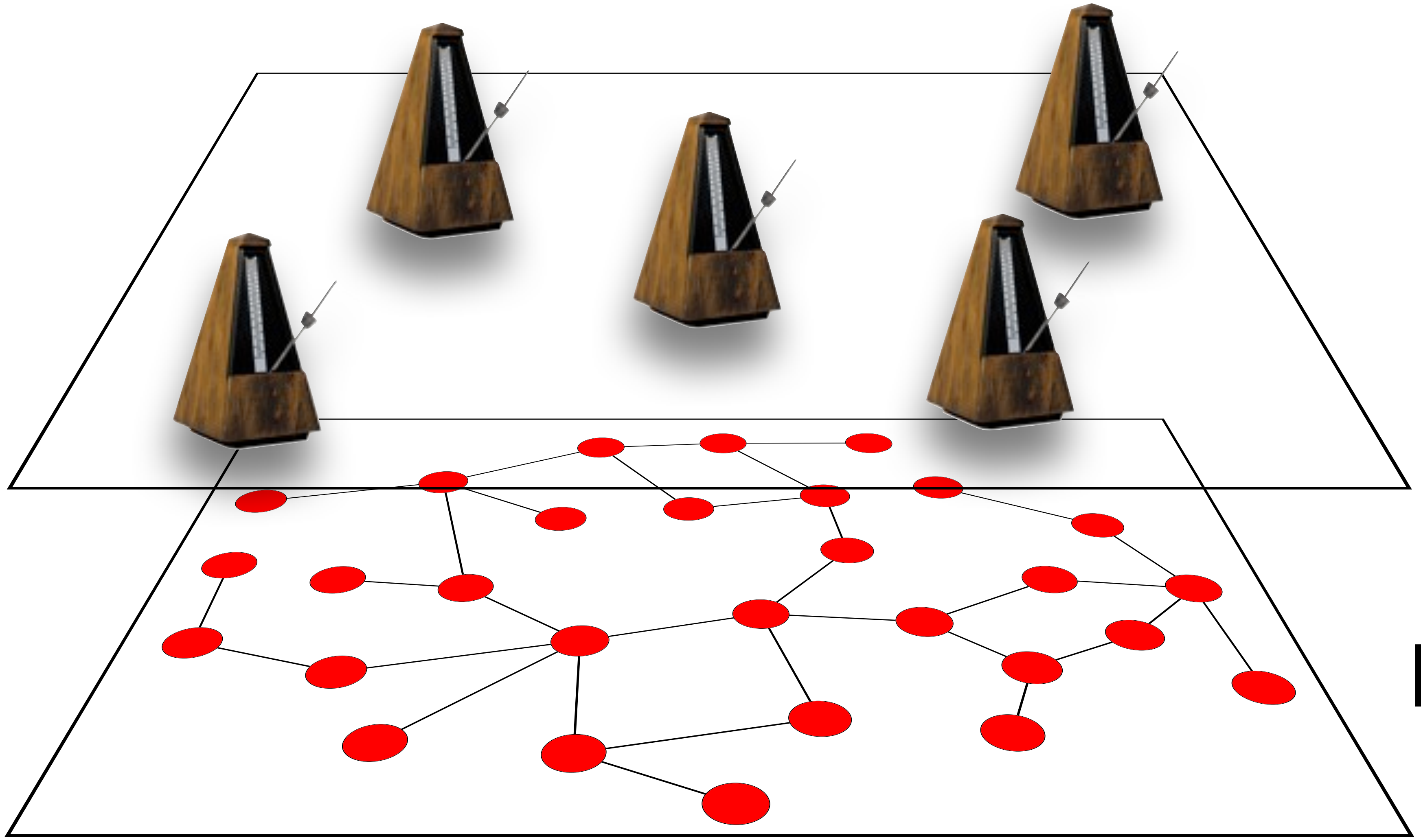
Dynamics on networks

Dynamics on networks



Network substrate

Dynamics on networks



Dynamical system on top

↑

Network substrate

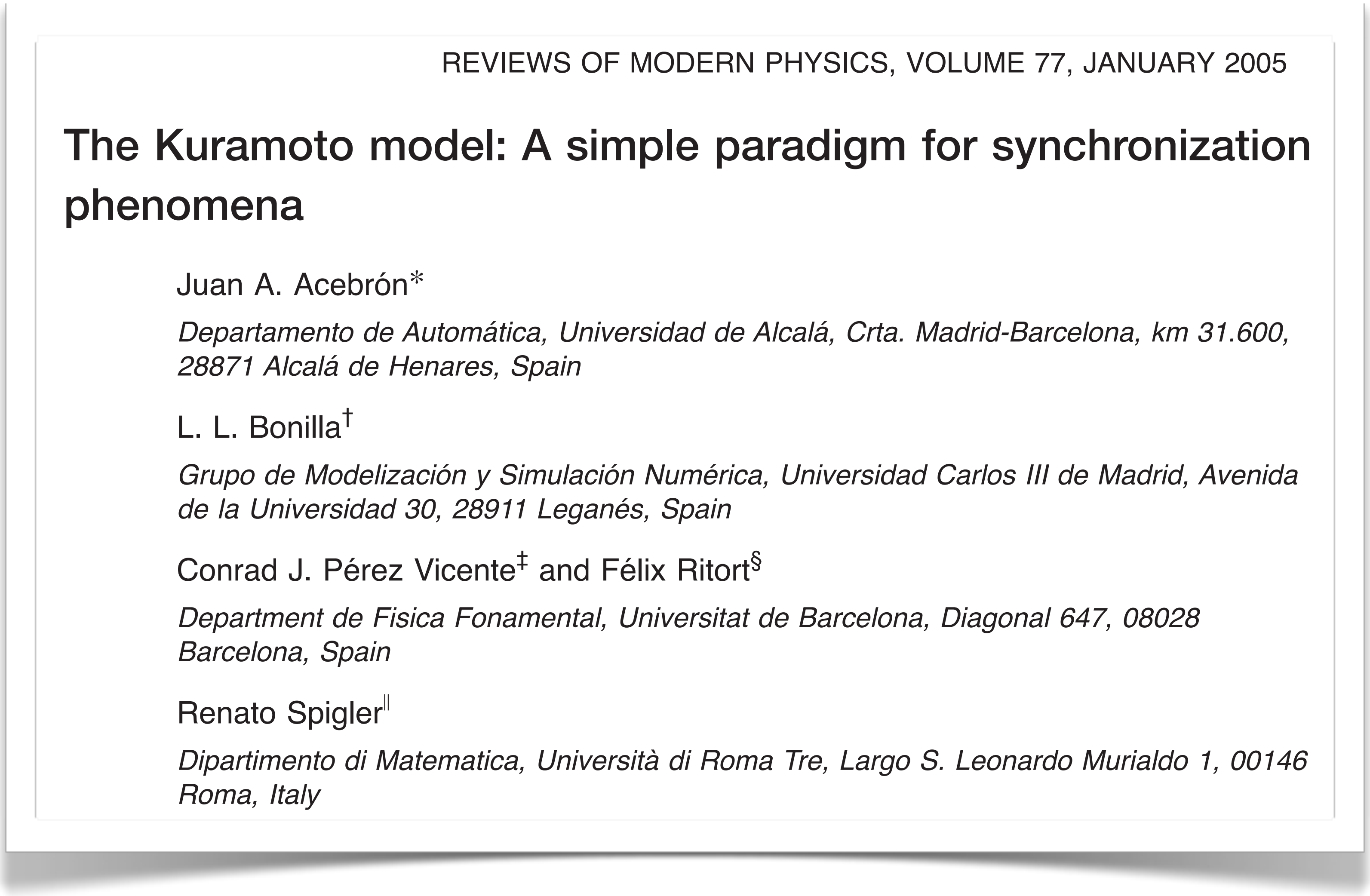
Dynamics on networks

Oscillators at each node, coupled with neighbors $\longrightarrow$ Can they **synchronize**?

Dynamics on networks

Oscillators at each node,
coupled with neighbors → Can they **synchronize**?

HUGE
area



REVIEWS OF MODERN PHYSICS, VOLUME 77, JANUARY 2005

The Kuramoto model: A simple paradigm for synchronization phenomena

Juan A. Acebrón*

Departamento de Automática, Universidad de Alcalá, Crta. Madrid-Barcelona, km 31.600,
28871 Alcalá de Henares, Spain

L. L. Bonilla†

Grupo de Modelización y Simulación Numérica, Universidad Carlos III de Madrid, Avenida
de la Universidad 30, 28911 Leganés, Spain

Conrad J. Pérez Vicente‡ and Félix Ritort§

Department de Fisica Fonamental, Universitat de Barcelona, Diagonal 647, 08028
Barcelona, Spain

Renato Spigler‖

Dipartimento di Matematica, Università di Roma Tre, Largo S. Leonardo Murialdo 1, 00146
Roma, Italy

Dynamics on networks

Dynamics on networks

Something simple ⟶ **Spreading**

Dynamics on networks

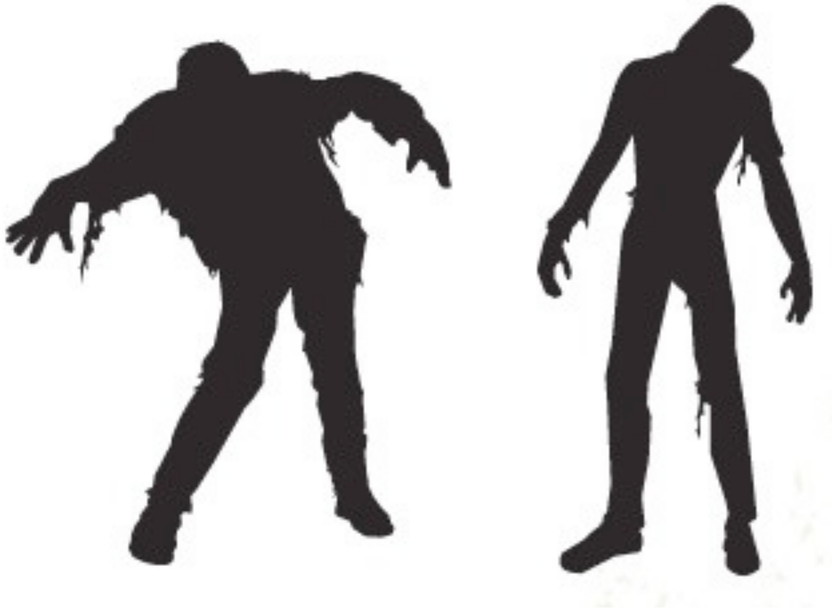
Something simple ⟶ **Spreading**

Disease outbreak
viral marketing
information cascade

Dynamics on networks

Something simple ⟶ **Spreading**

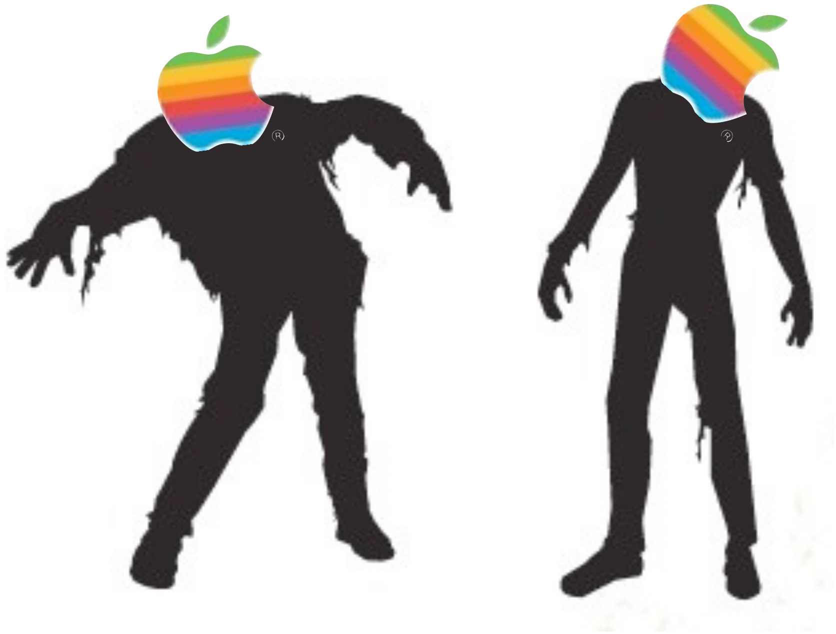
Disease outbreak
viral marketing
information cascade



Dynamics on networks

Something simple ⟶ **Spreading**

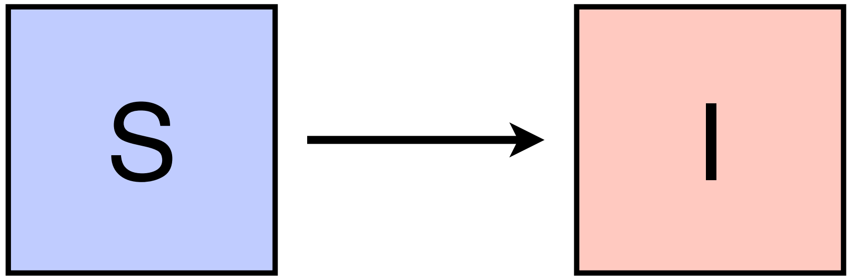
Disease outbreak
viral marketing
information cascade



Mathematical epidemiology

Compartmental models No network...

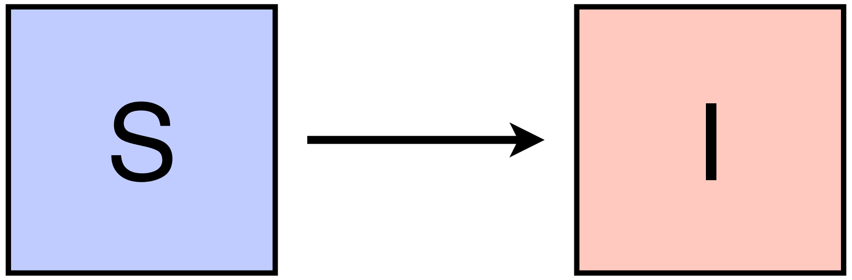
SI model People are **Susceptible** or **Infected**

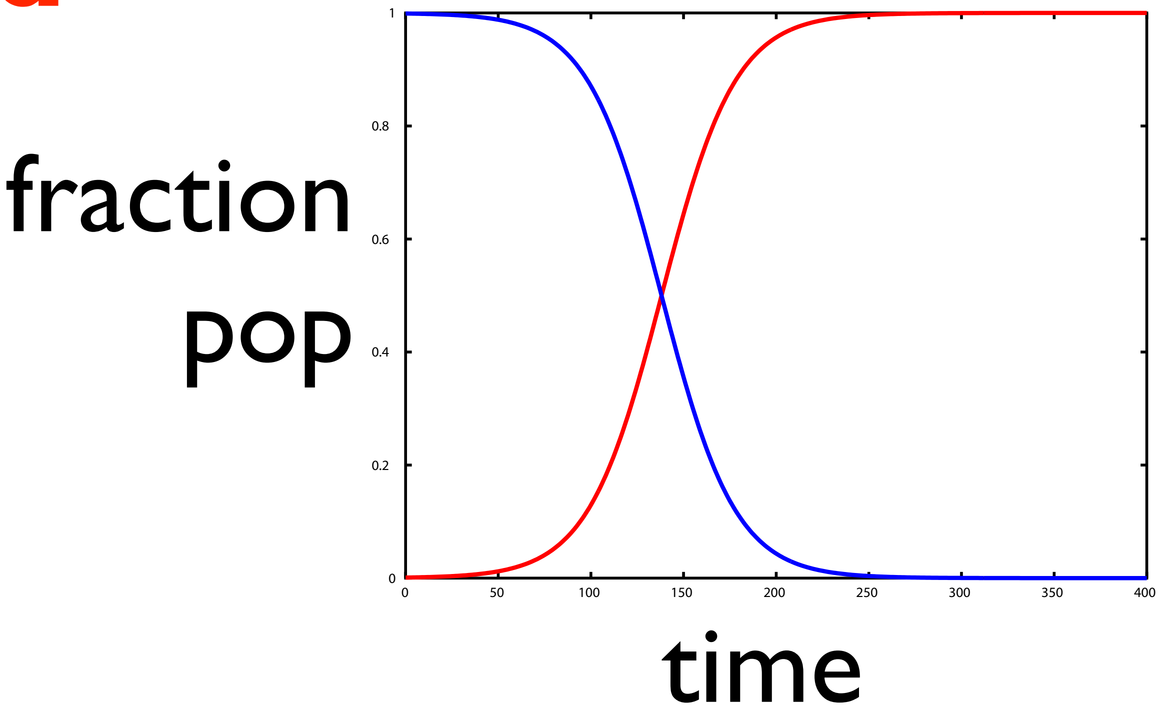


Mathematical epidemiology

Compartmental models   No network...

SI model   People are **Susceptible** or **Infected**

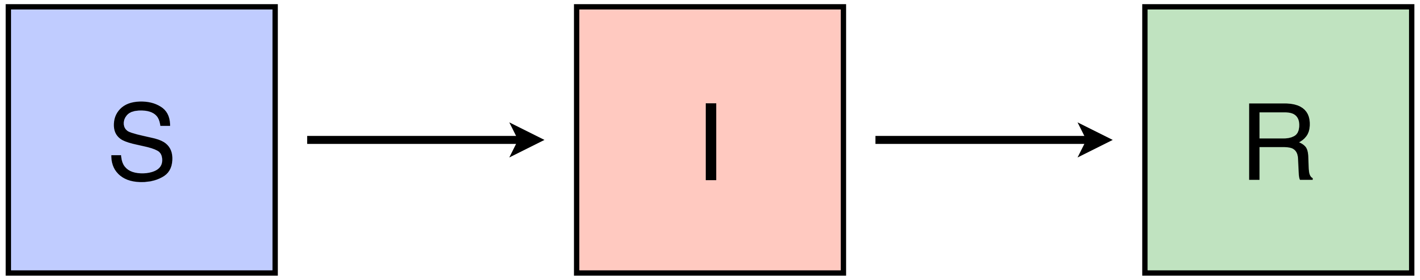




Mathematical epidemiology

Compartmental models No network...

SIR model People are **Susceptible**, **Infected**, or **Recovered**



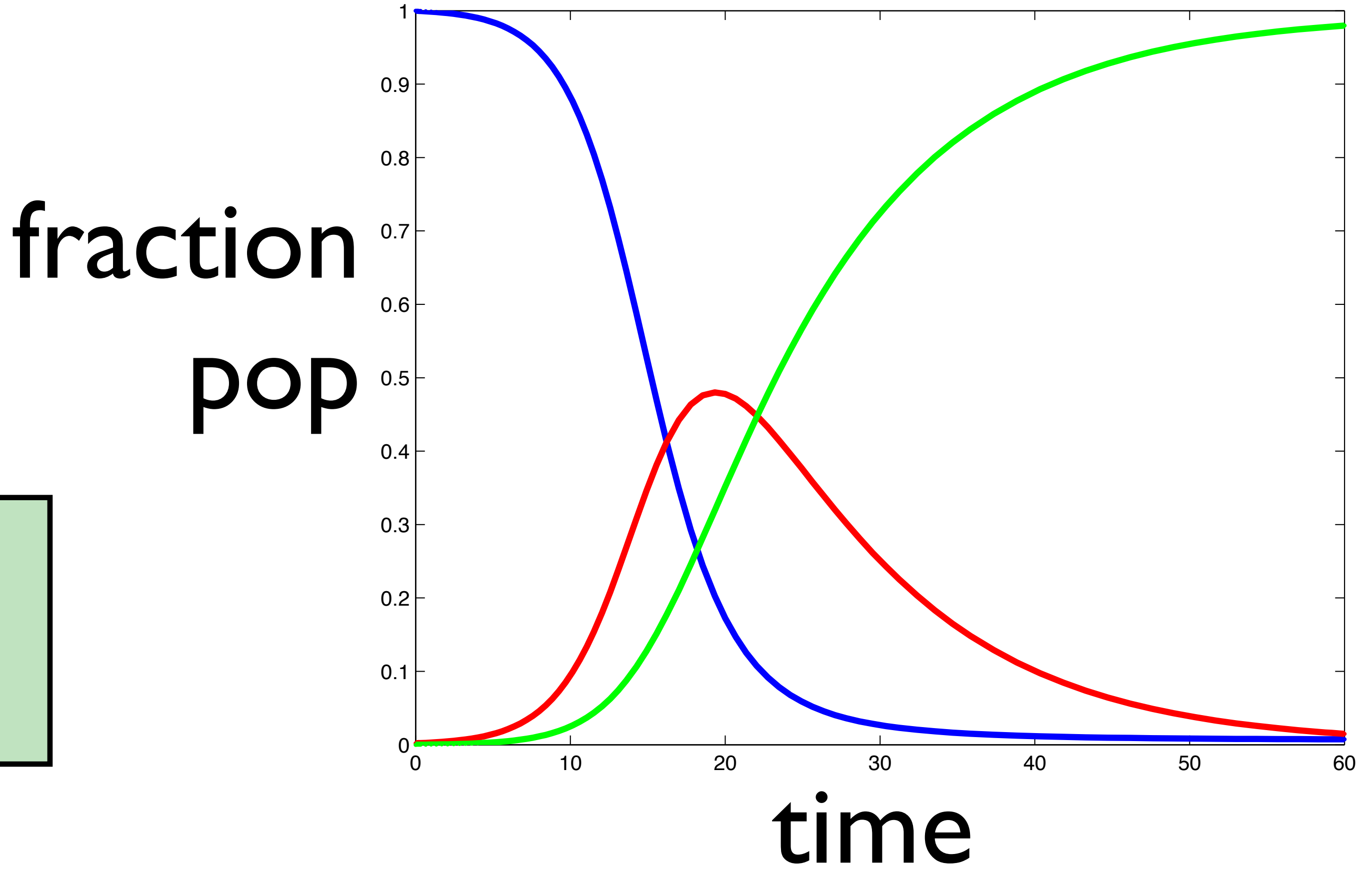
Mathematical epidemiology

Compartmental models

No network...

SIR model

S I R



Network epidemics

SI on a **network**

Unlike compartments, an outbreak might not infect **everyone**

If the network is **disconnected**

↓

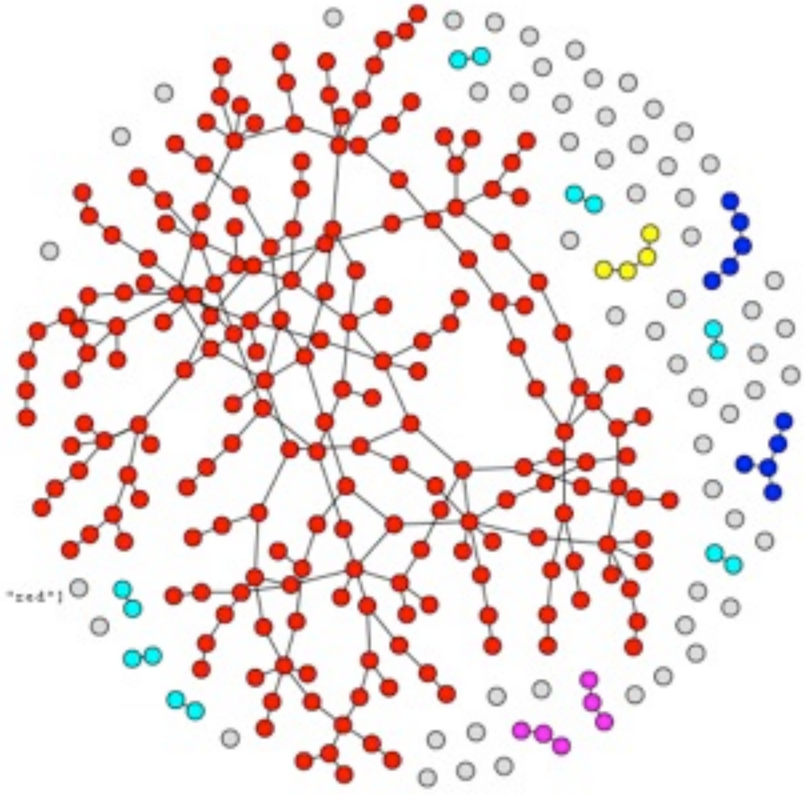
Percolation

Network epidemics

SI on a **network**

Distribution of **component sizes** tells us how big outbreak will be

(more than just GCC)



Network epidemics

SIR on a **network**

People recover after a certain **time**

An infected node might recover **before** coming into **contact** with a neighbor

ϕ Transmission probability

Network epidemics

SIR on a **network**

ϕ Transmission probability

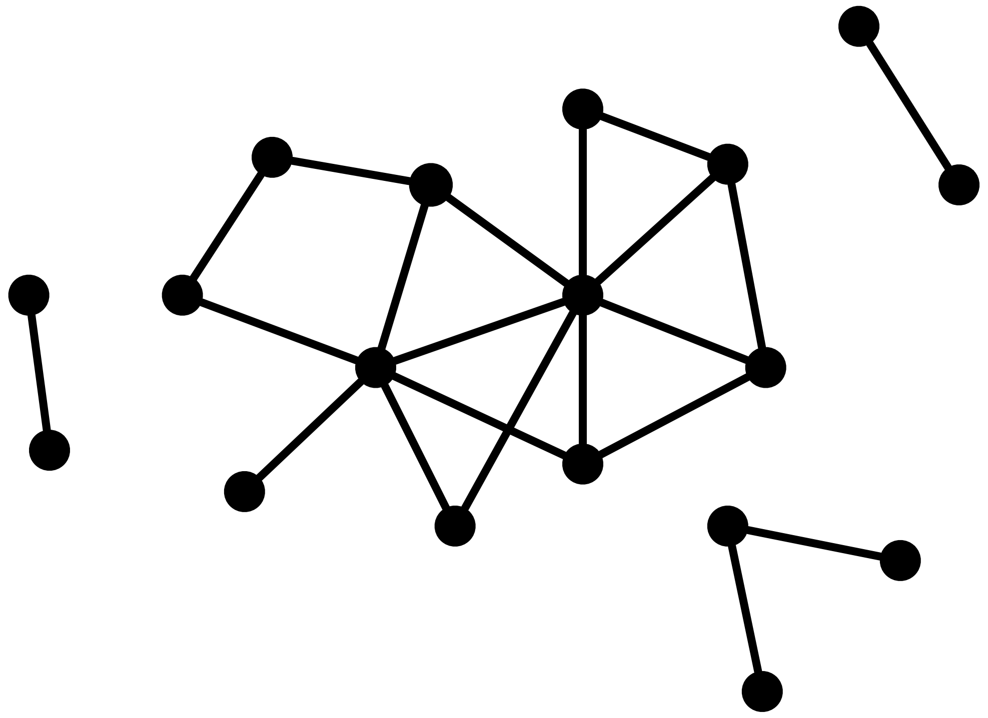
Disease not being transmitted = cut link

Random disease = bond percolation

Network epidemics

SIR on a **network** ϕ Transmission probability

Random disease = bond percolation



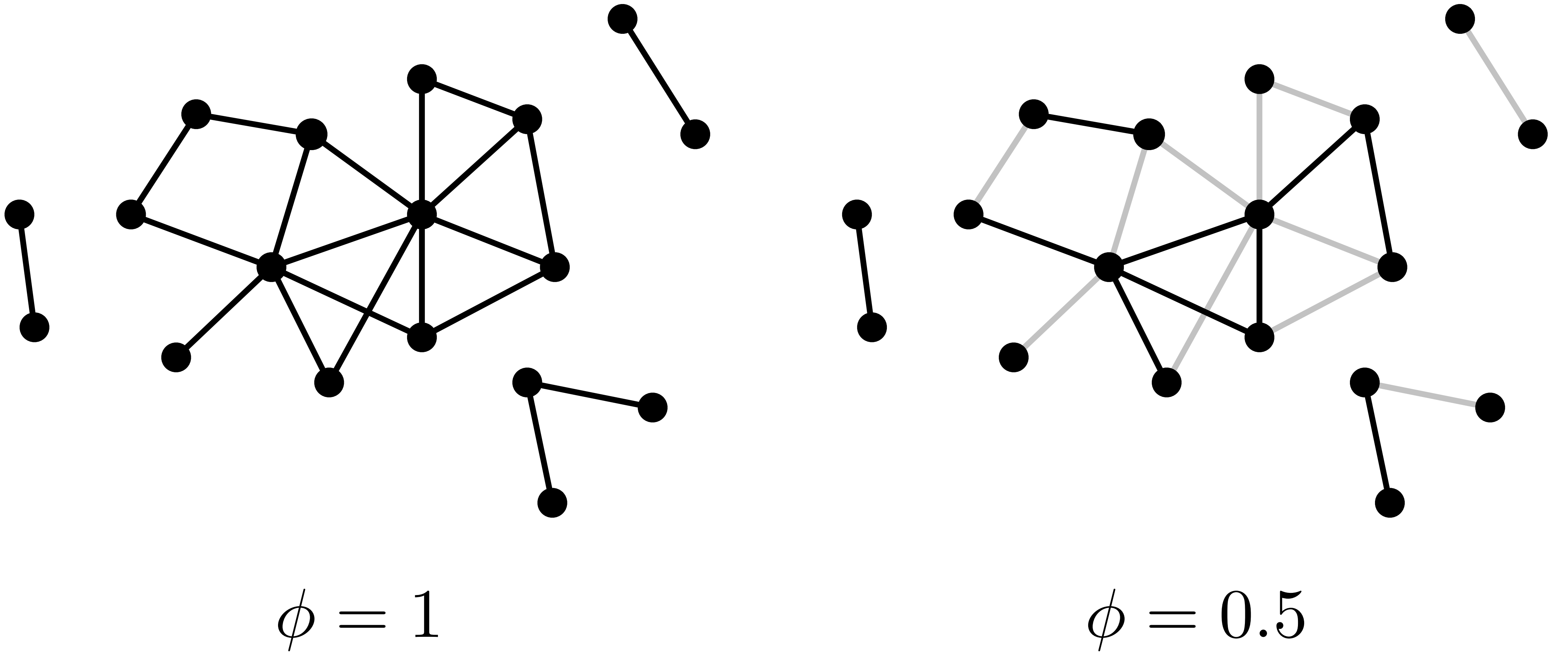
$\phi = 1$

Network epidemics

SIR on a **network**

ϕ Transmission probability

Random disease = bond percolation

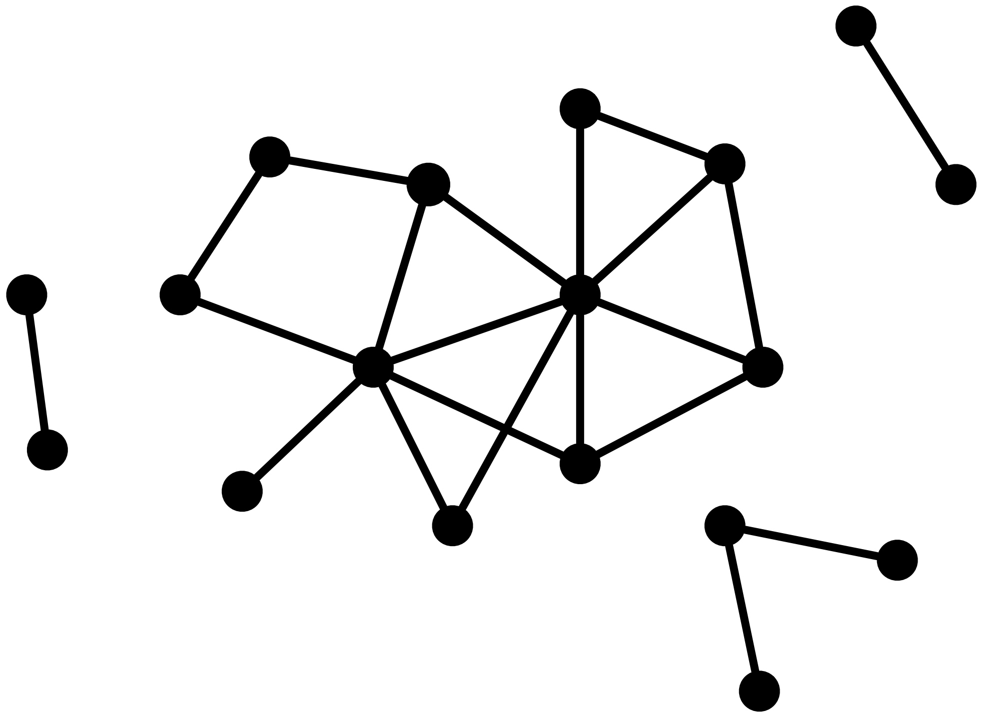


Network epidemics

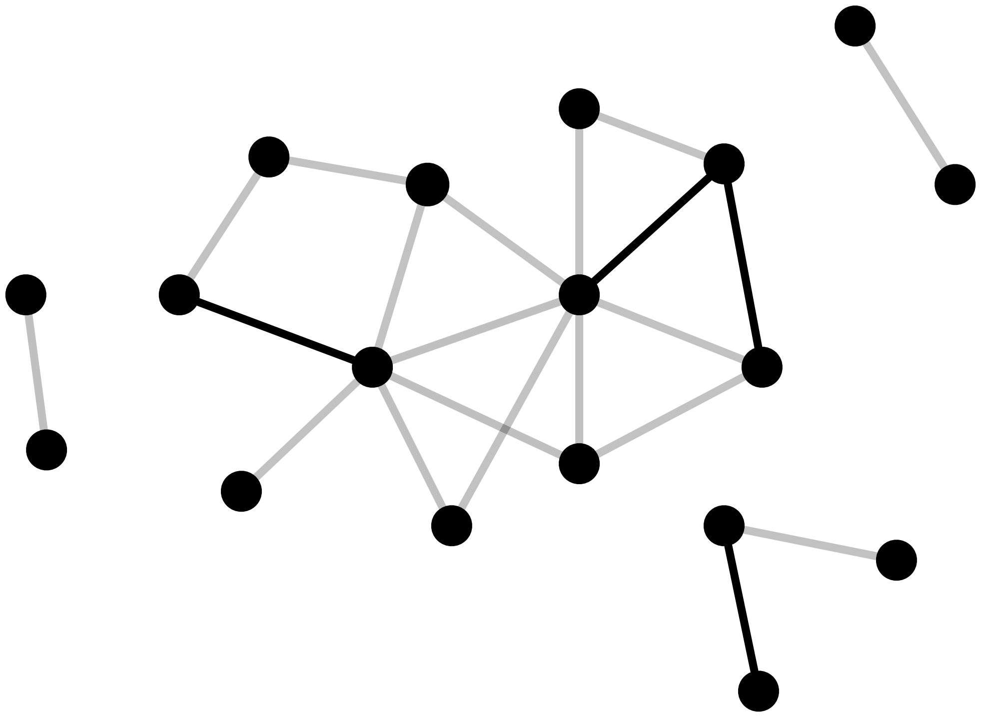
SIR on a **network**

ϕ Transmission probability

Random disease = bond percolation



$\phi = 1$



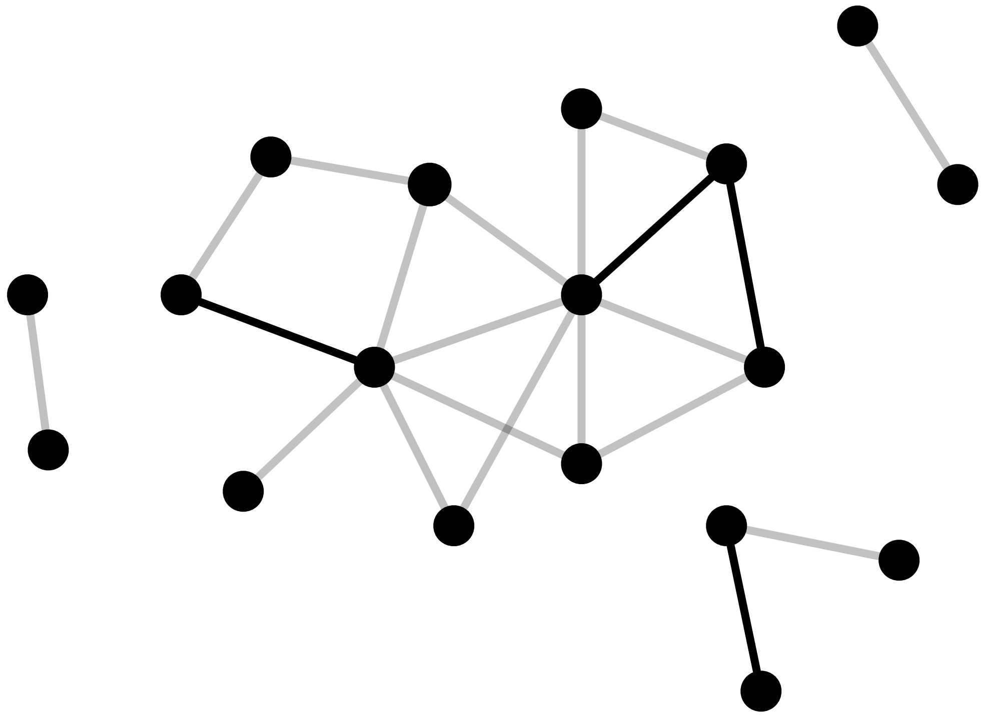
$\phi = 0.2$

Network epidemics

SIR on a **network**

ϕ Transmission probability

Random disease = bond percolation



$\phi = 0.2$

Expected statistics of bond percolation clusters

Teach us about **expected behavior** of SIR model on the network

Network epidemics

What if network is connected?

Network epidemics

What if network is connected?

Deleting nodes = vaccinating people

Network epidemics

What if network is connected?

Deleting nodes = vaccinating people

Limited number of vaccines: who to vaccinate?

Network epidemics

What if network is connected?

Deleting nodes = vaccinating people

Limited number of vaccines: who to vaccinate?

Attack the **hubs**!

Shatter the network = prevent **large** outbreak

Network epidemics

How to find the hubs when you don’t know the network

Network epidemics

How to find the hubs when you don't know the network

1. Pick someone at **random**
2. Ask who their **friends** are
3. Vaccinate **friends**

Network epidemics

How to find the hubs when you don't know the network

1. Pick someone at **random**
2. Ask who their **friends** are
3. Vaccinate **friends**

Mimics Krapivsky-Redner **redirection**

Network epidemics

How to find the hubs when you don't know the network

1. Pick someone at **random**
2. Ask who their **friends** are
3. Vaccinate **friends**

Mimics Krapivsky-Redner **redirection**

Vaccines distributed by **preferential attachment**

Communities

Mesoscopic Structure

Mesoscopic Structure

local

global

Mesoscopic Structure

local

motifs

navigation

global

degree

mixing

hierarchy

degree distribution

clustering

modules

robustness

Mesoscopic Structure

local

global

modules

Mesoscopic Structure

local

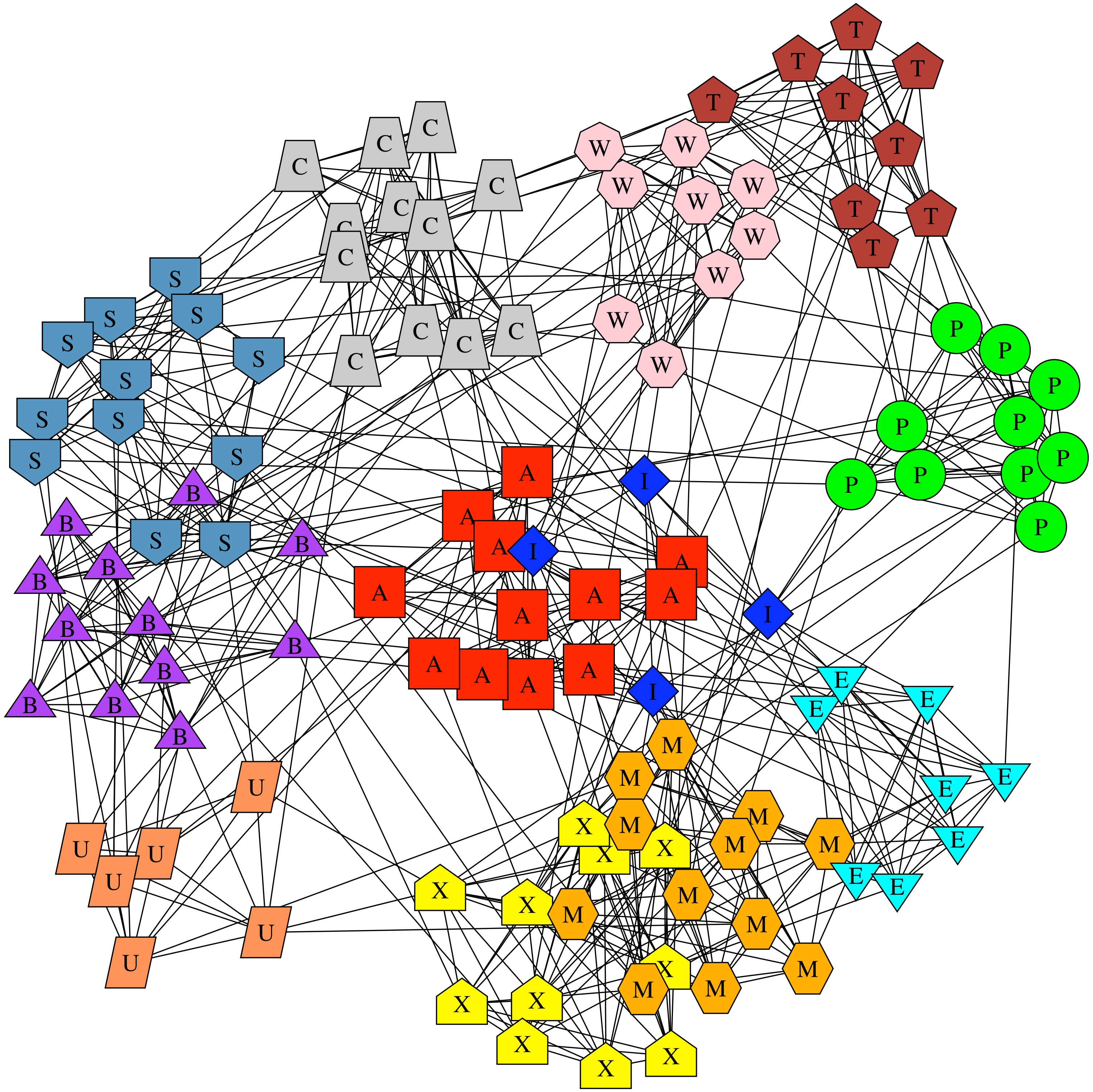
global

modules

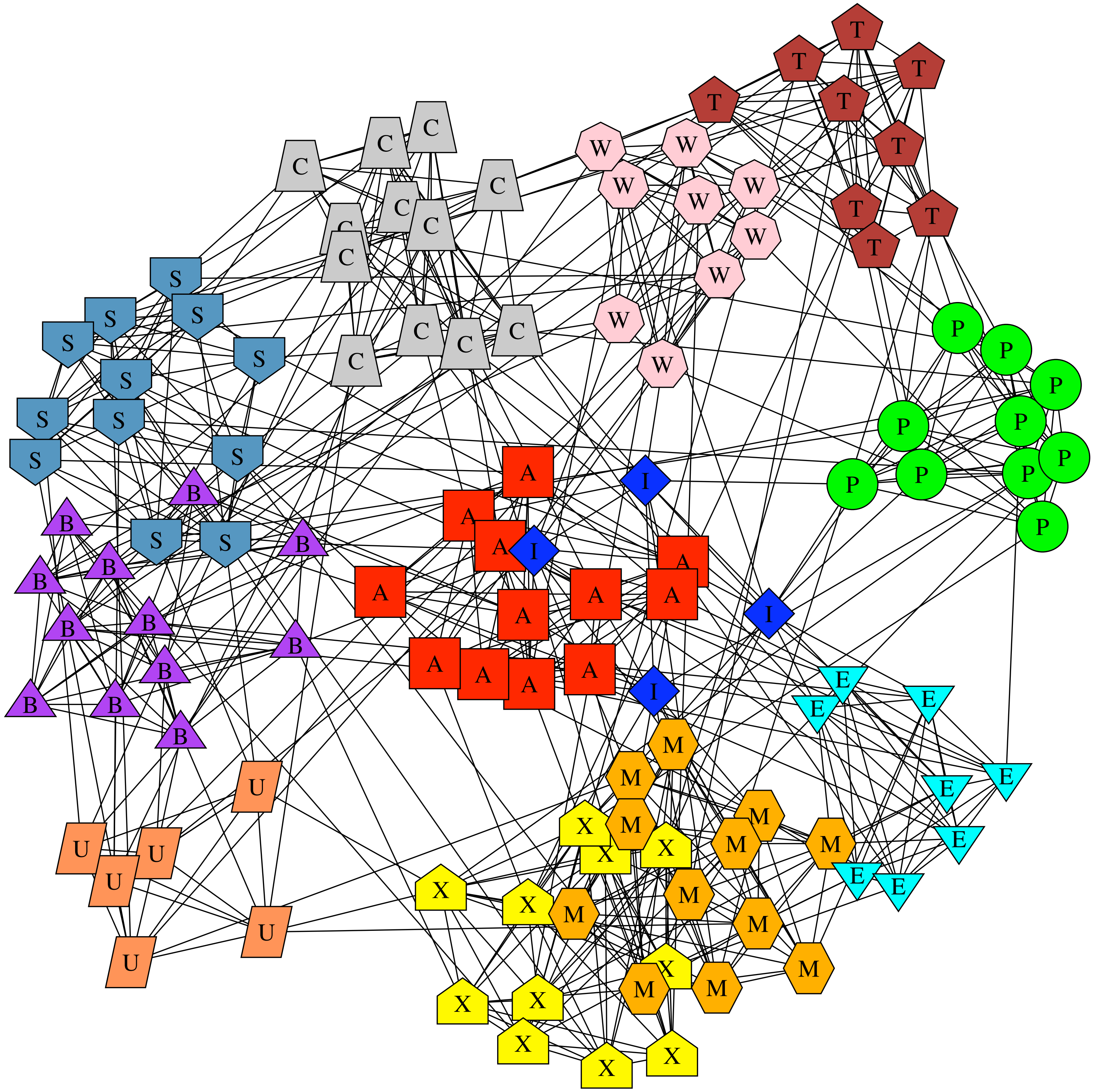
↓

communities

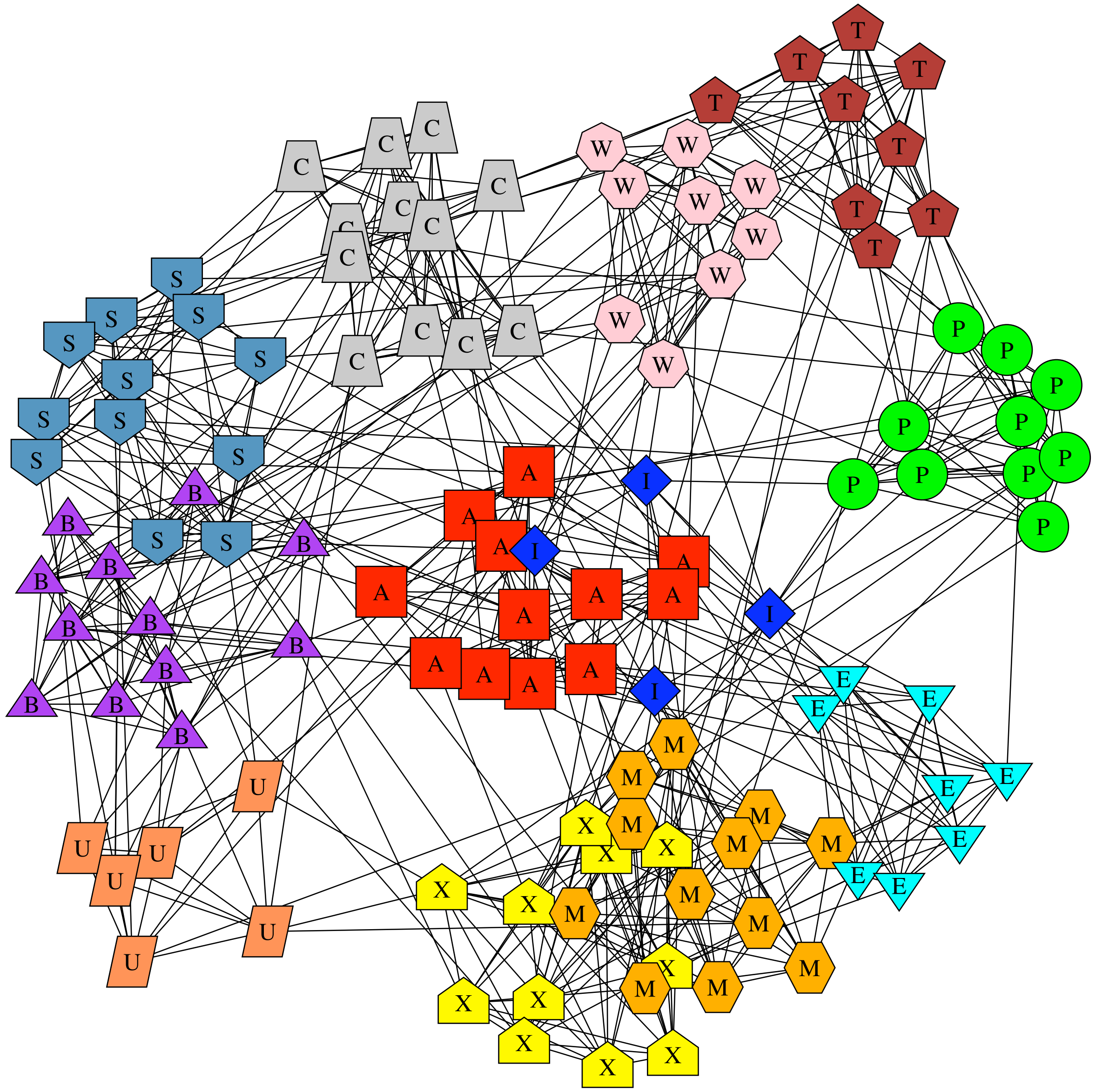
Mesoscopic Structure



Mesoscopic Structure



Communities



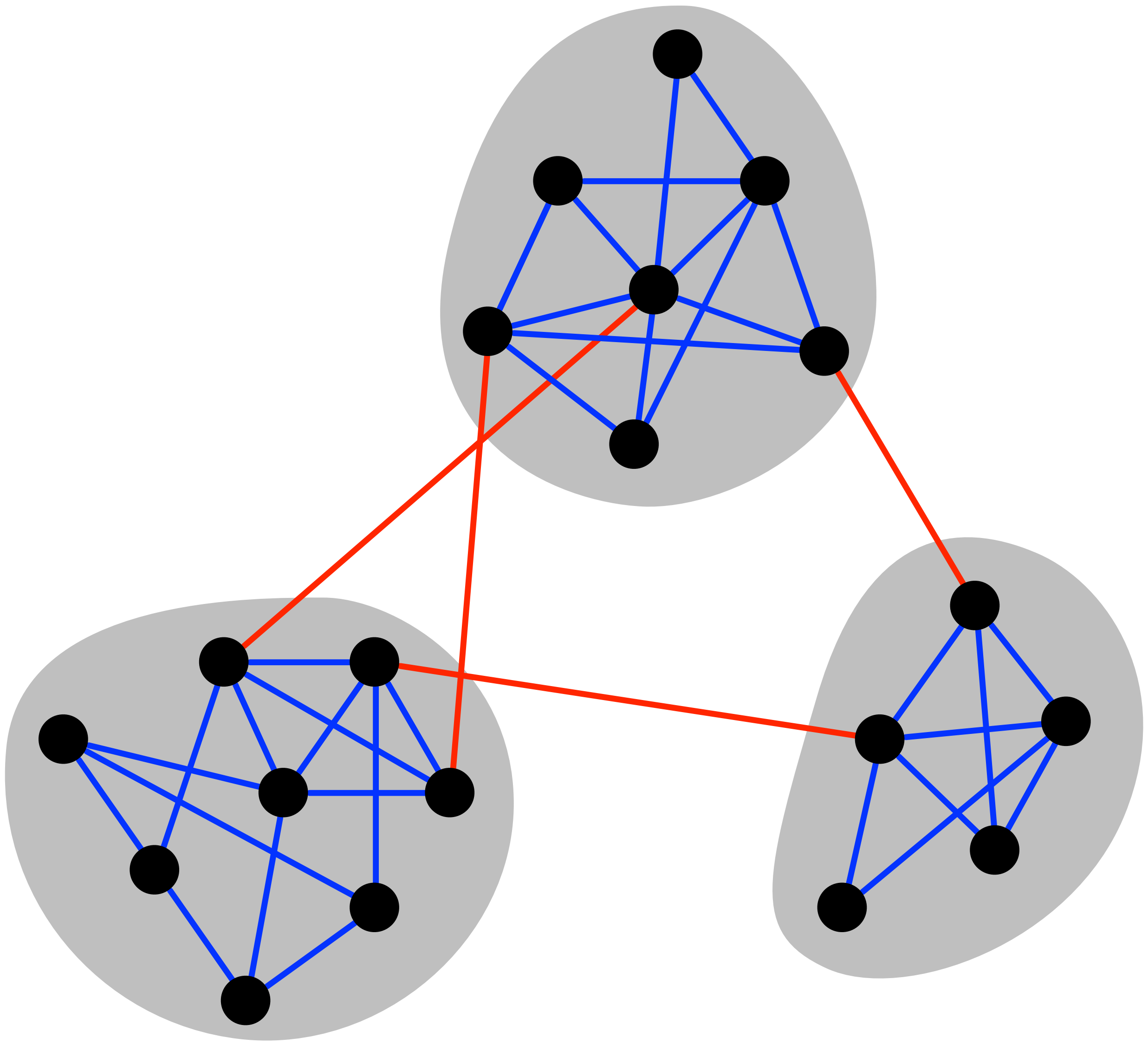
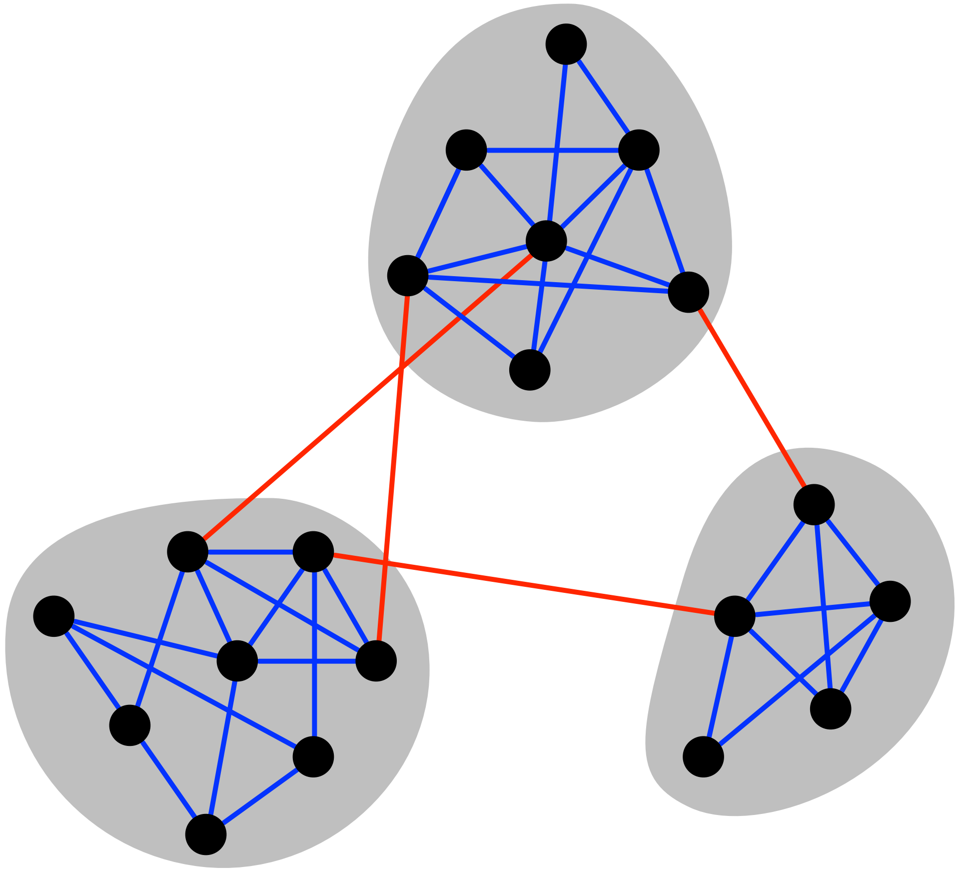


Illustration. Newman, PNAS **103,** 8577-8582 (2006).

Communities

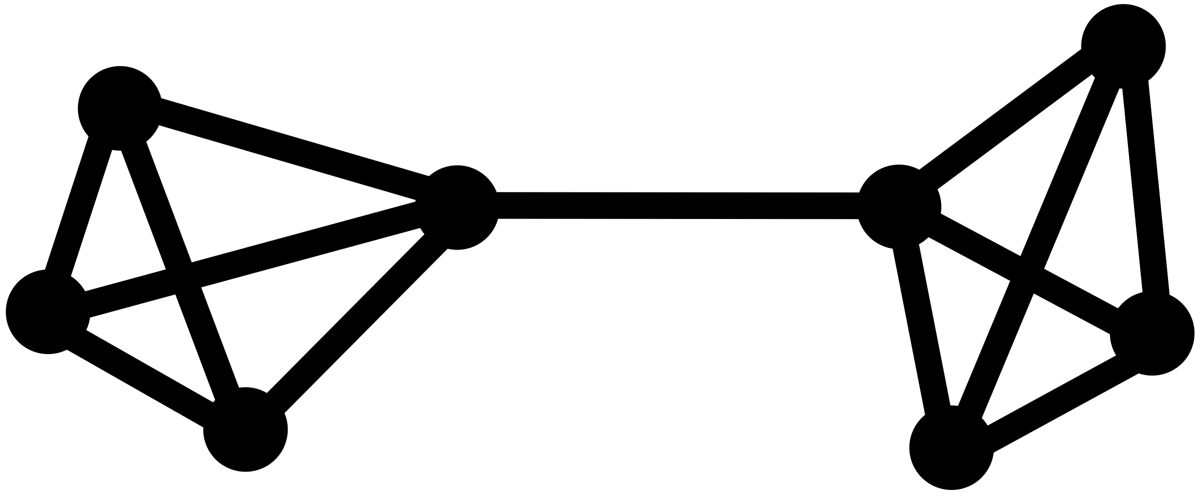


No **precise** definition

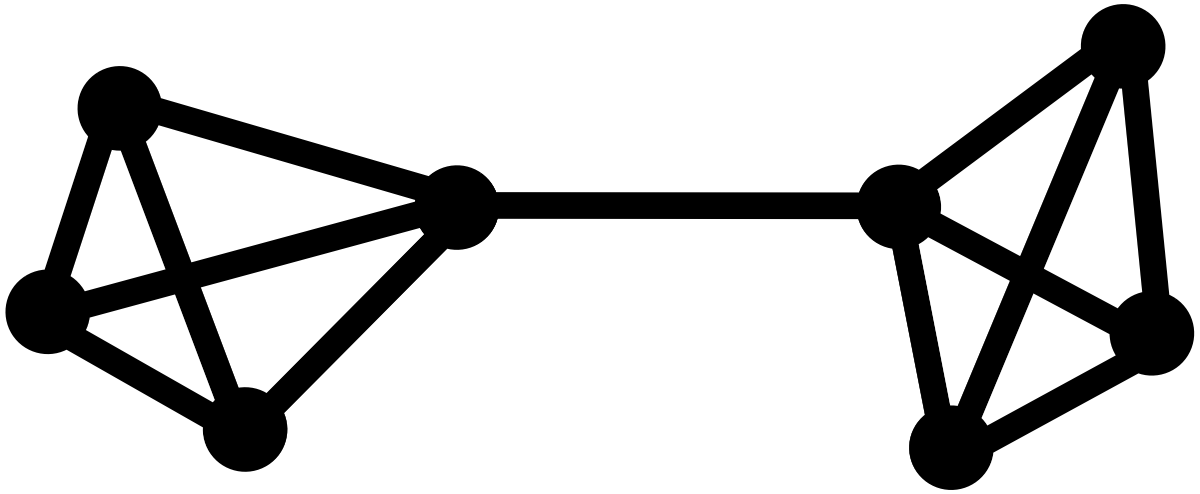
groups of nodes with

- many **internal links**
- few **external links**

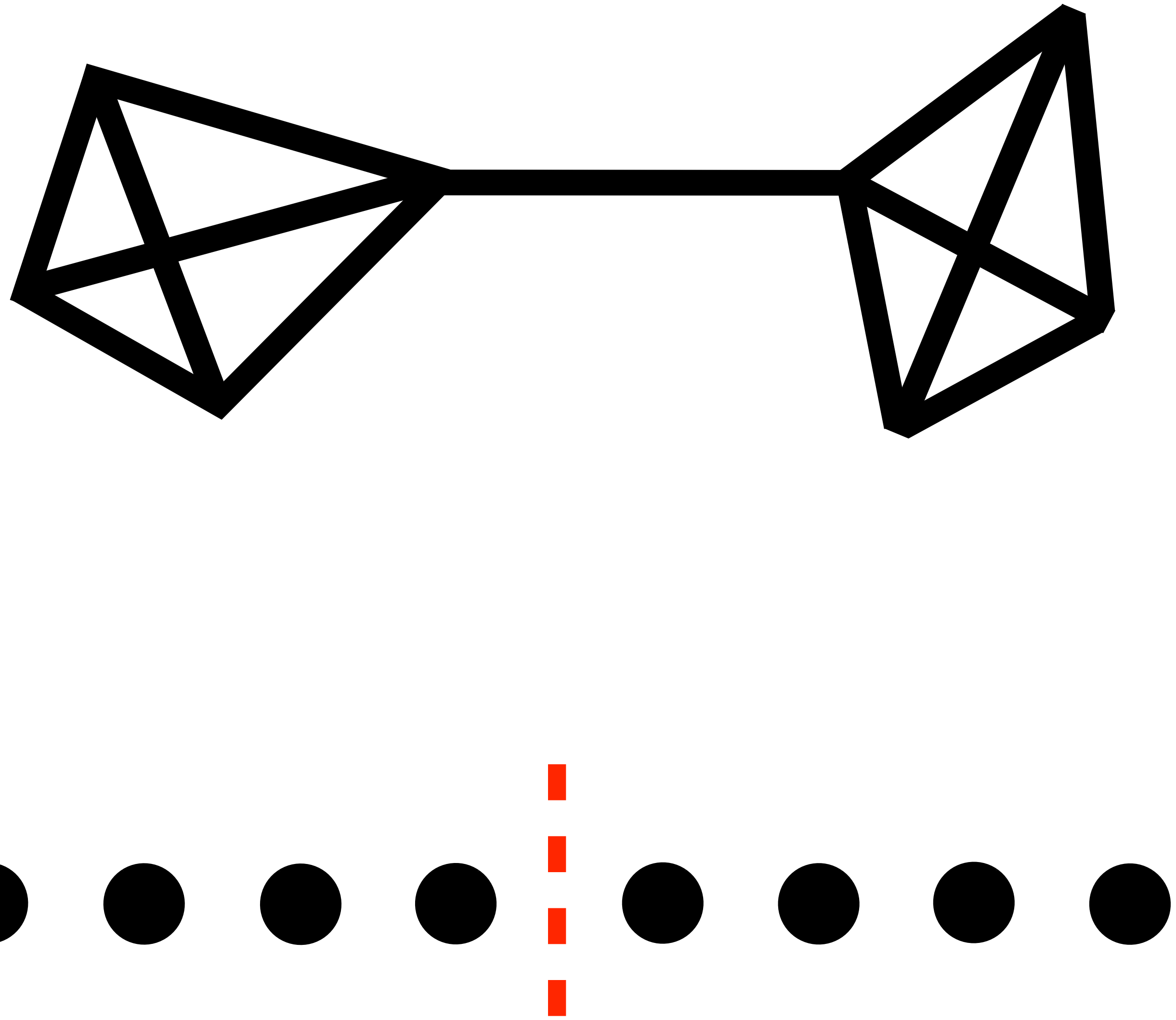
Community problem



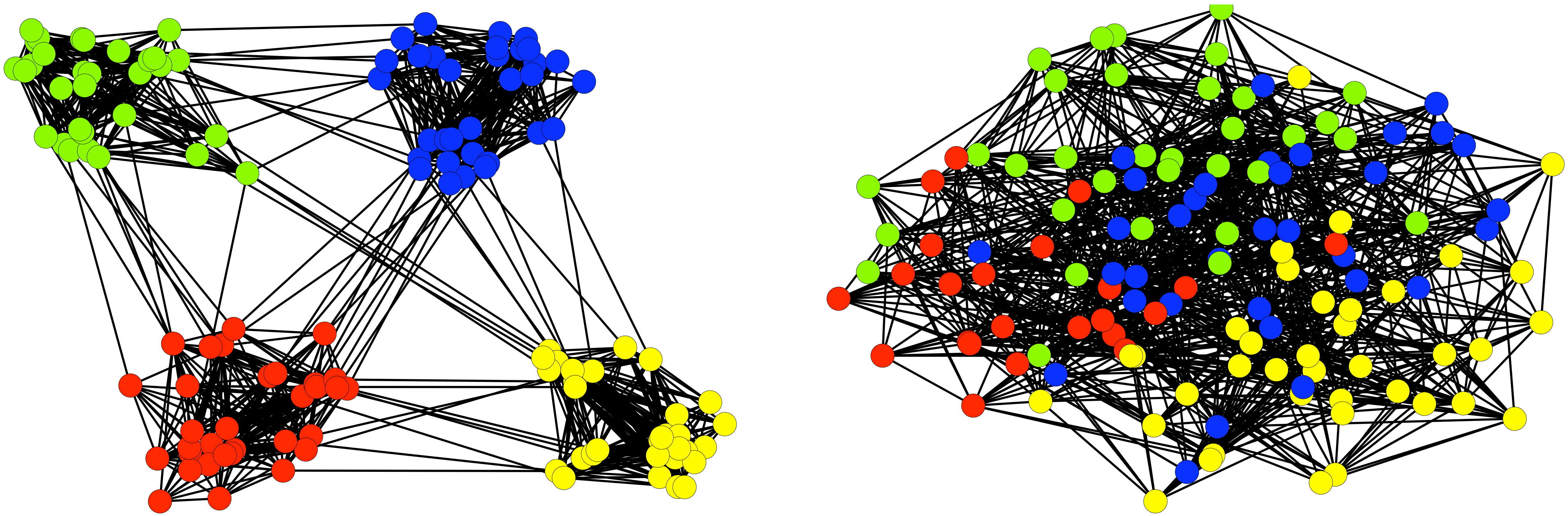
Community problem



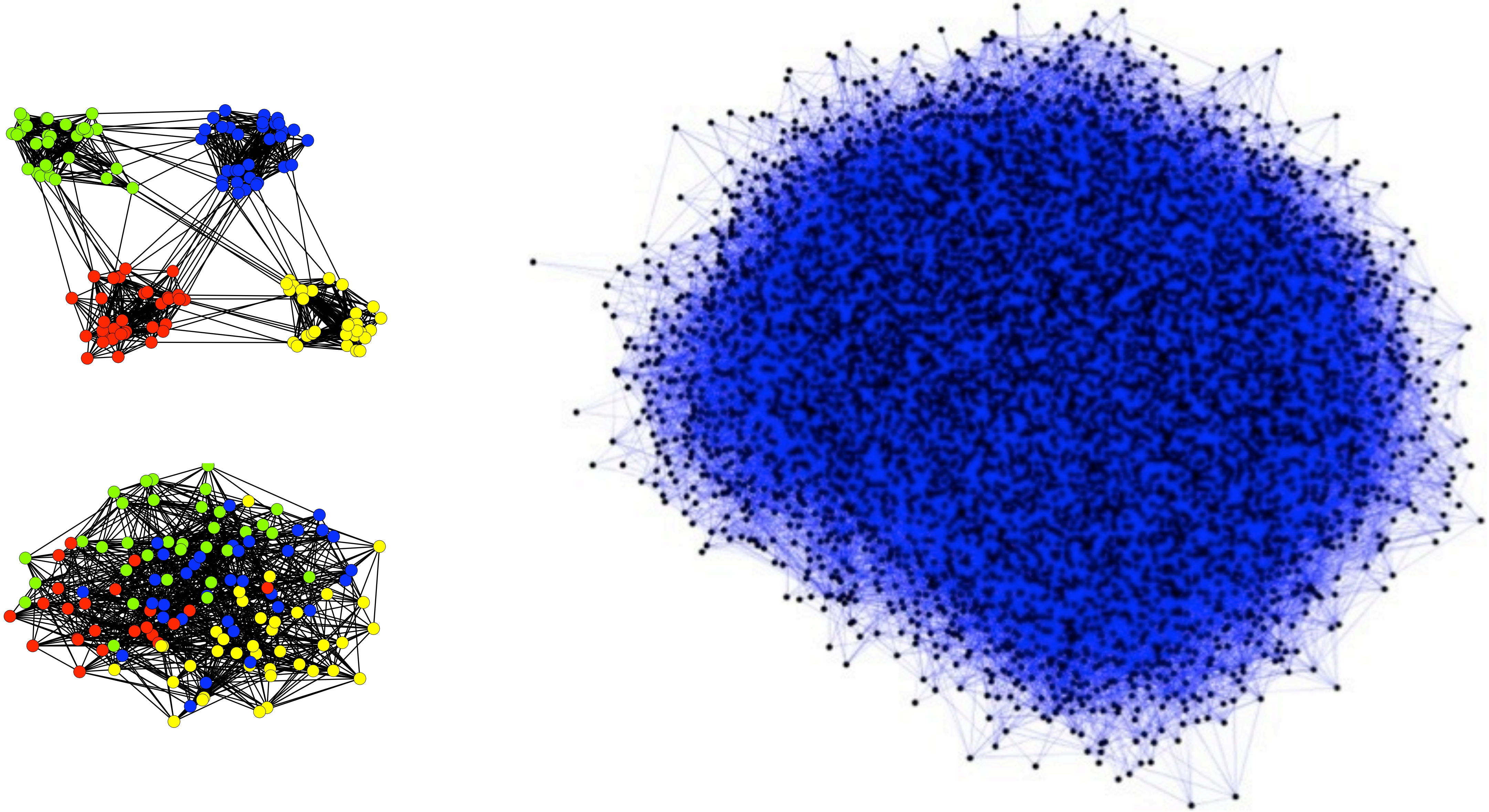
Community problem



Not so easy!



Not so easy!



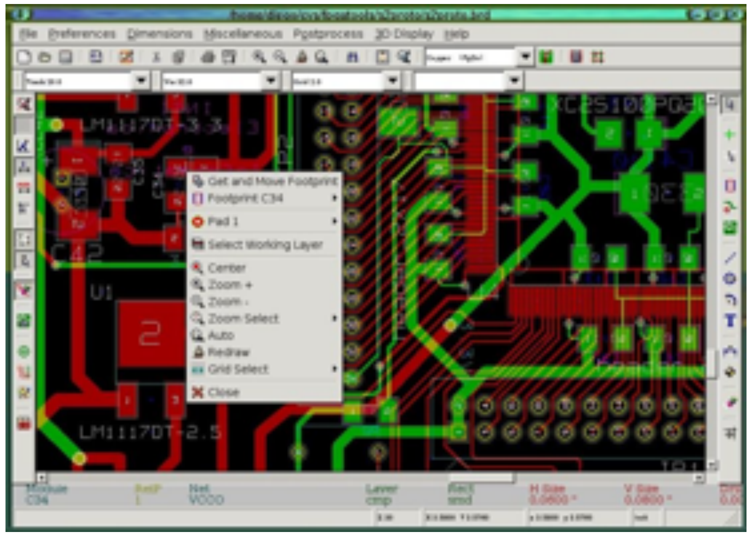
Why?

Why?

Why?

Optimization

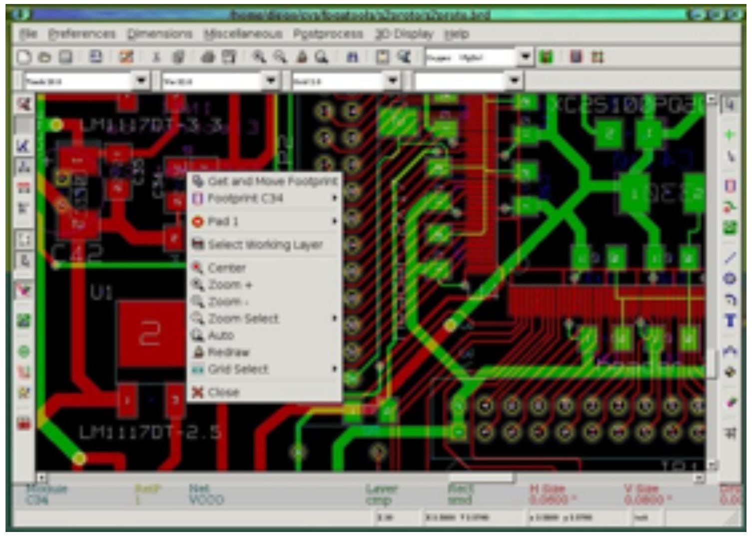
Circuit boards
Load balancing
Disease spreading



Why?

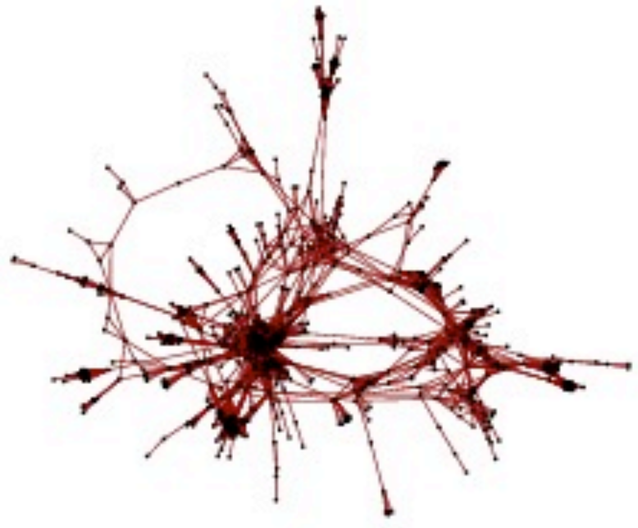
Optimization

- Circuit boards
- Load balancing
- Disease spreading



Biology

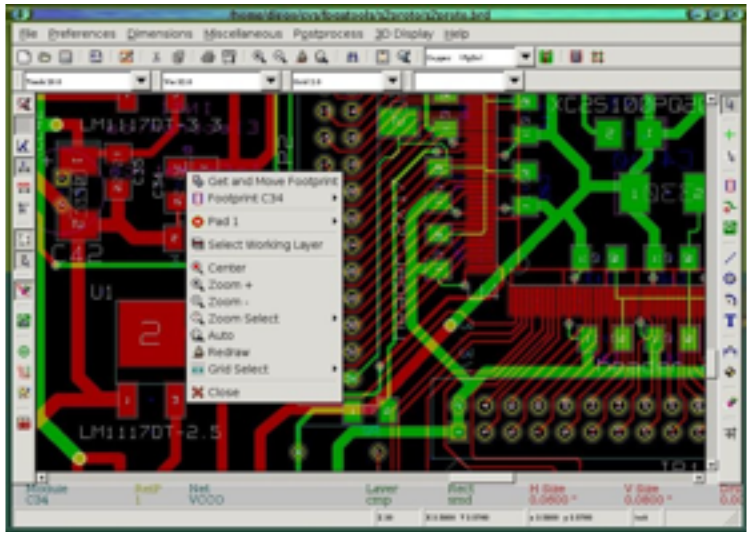
- Protein complexes
- Functional modules
- Neuronal clusters



Why?

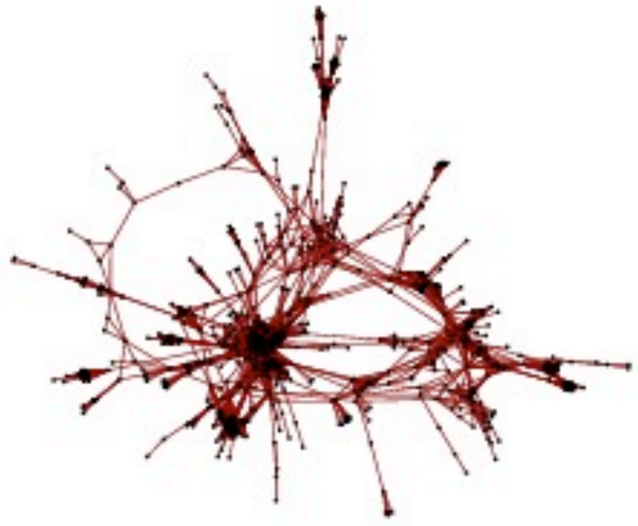
Optimization

Circuit boards
Load balancing
Disease spreading



Biology

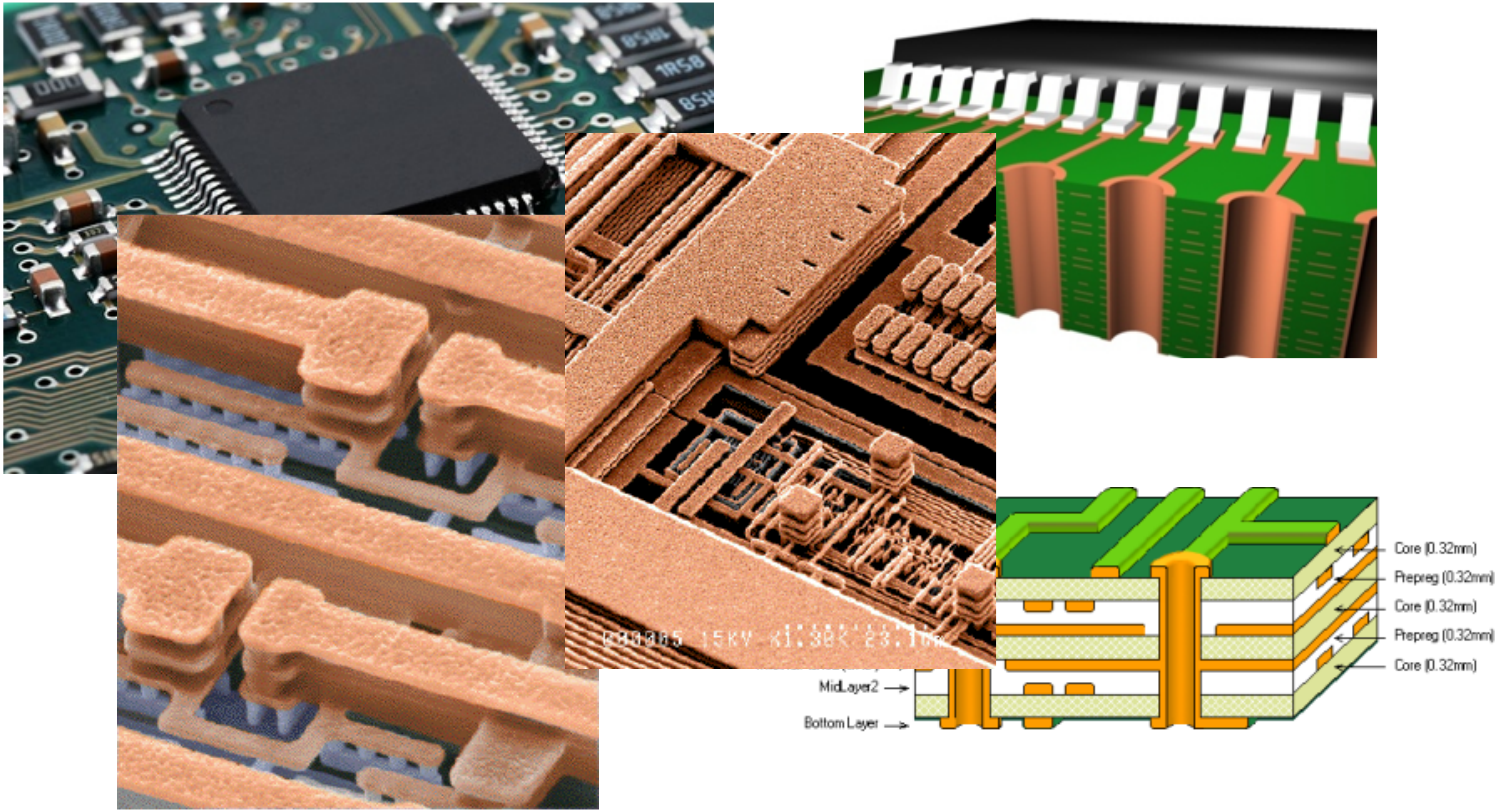
Protein complexes
Functional modules
Neuronal clusters



Inferring hidden relationships

Dynamics

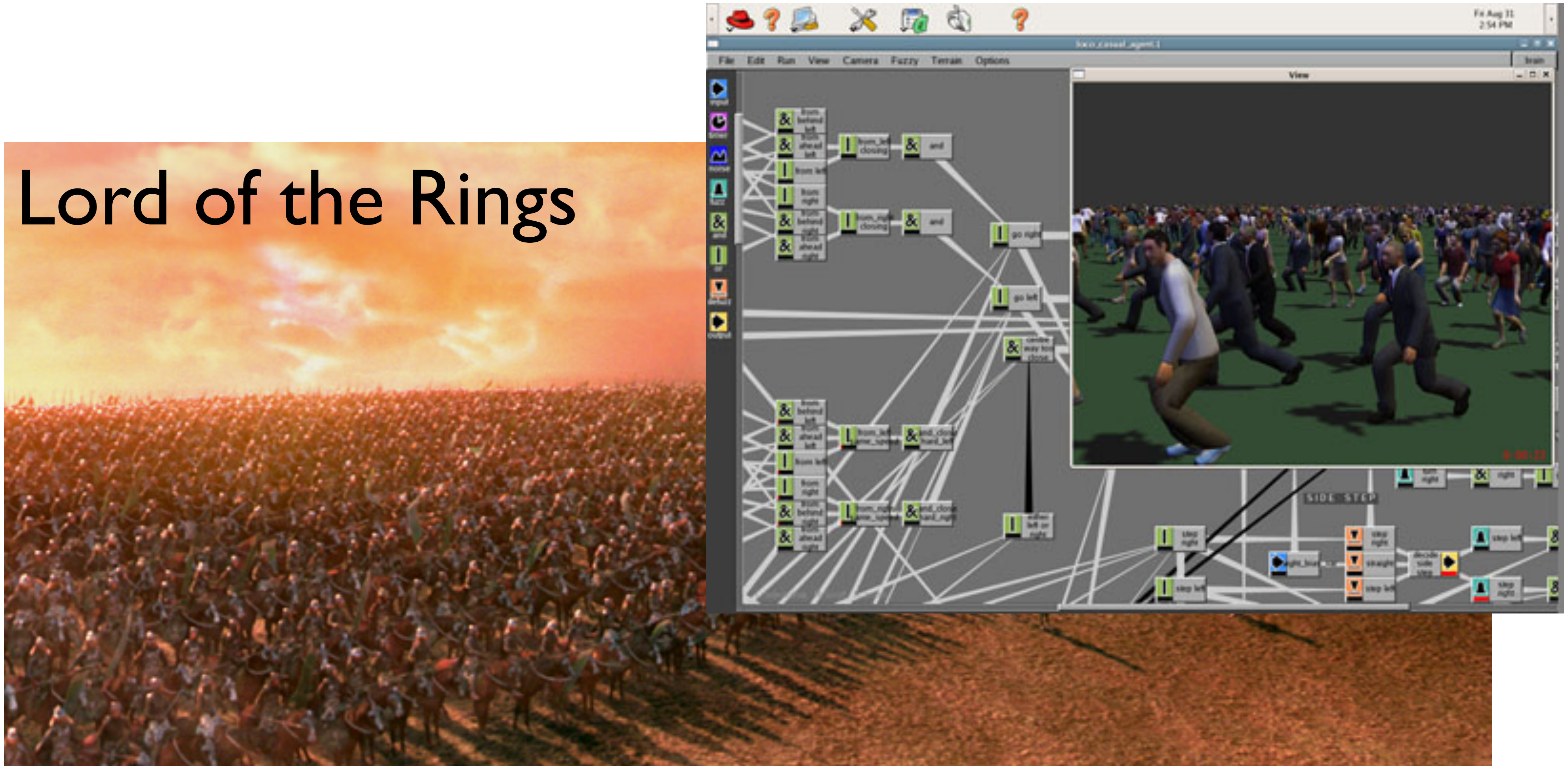
Circuit Boards



Distributed Computing

Lord of the Rings

Distributed Computing

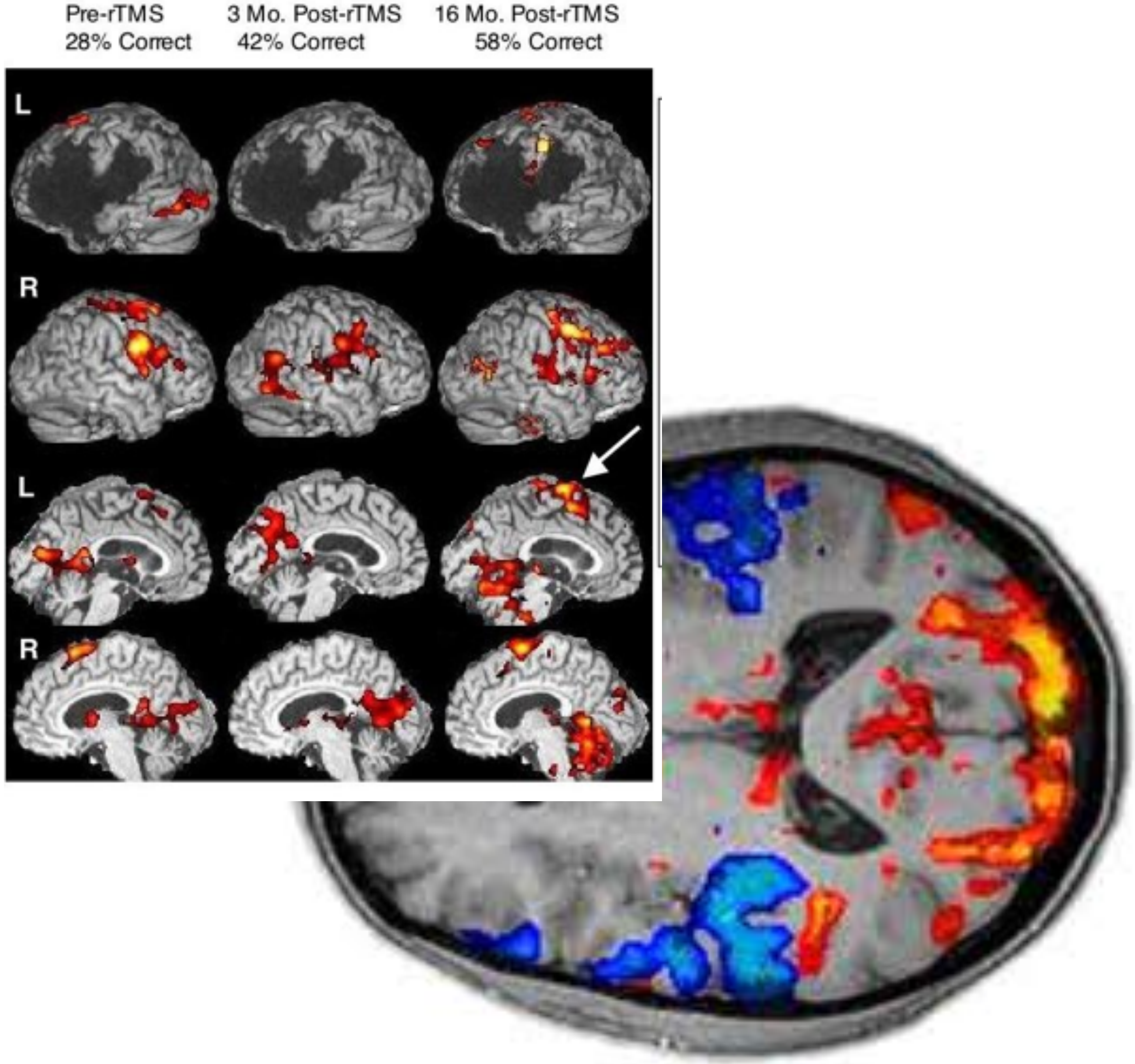


Distributed Computing

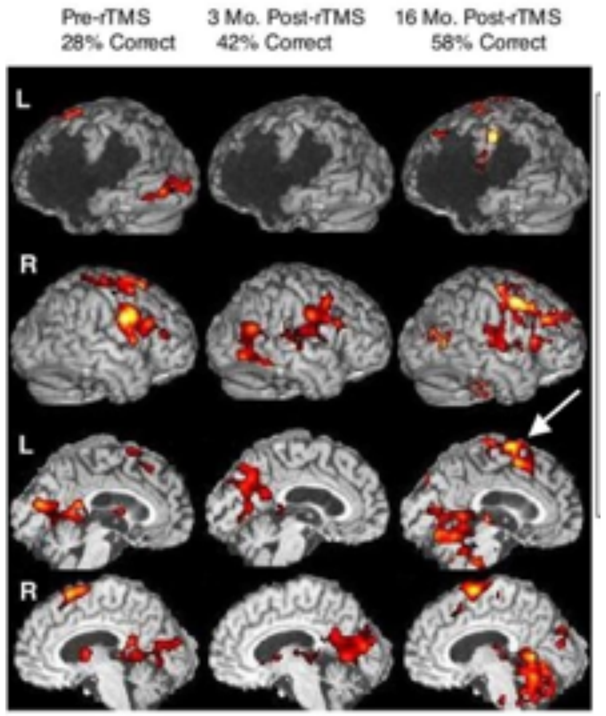
Lord of the Rings

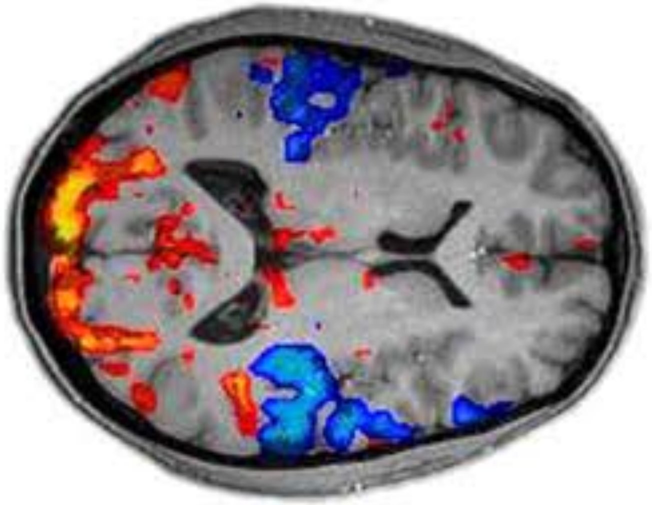
Life Sciences

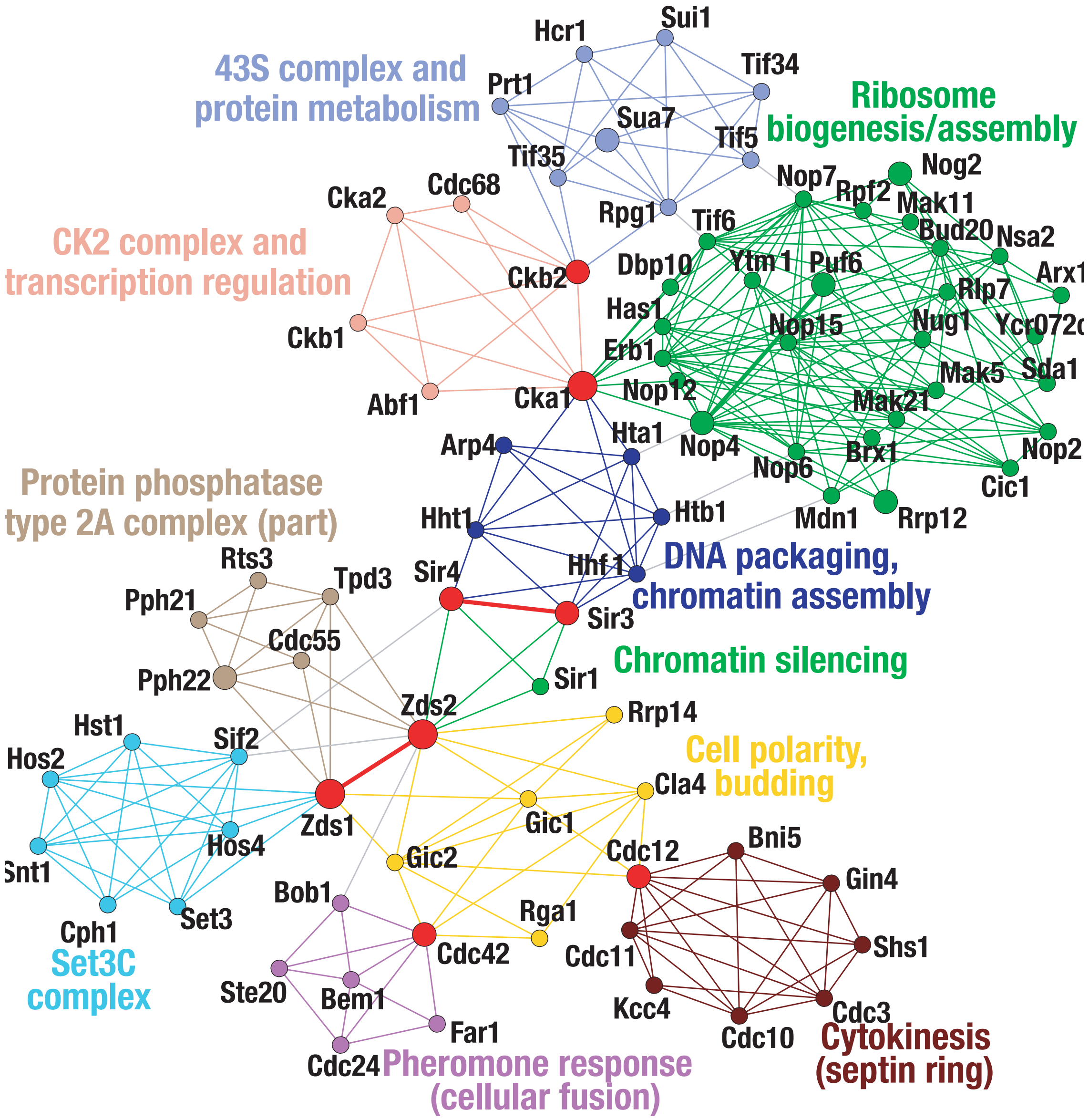
Life Sciences



Life Sciences







(Palla *et al.*, 2005)

Summary

Outline

Part I

- History
- Network examples/data
- Why study networks?
- Network quantifiers (jargon!)
- Types of networks
- Random network models

Part II

- Getting started on a computer
- Network search
- Network robustness
- Dynamics on networks

slides will be on
bagrow.com